ESTIMATION OF THE HURST PARAMETER FROM DISCRETE NOISY DATA

ARNAUD GLOTER, MARC HOFFMANN

Laboratoire d'Analyse et de Mathématique Appliquées, CNRS UMR 8050 et Université de Marne-la-Vallée

Abstract. We estimate the Hurst parameter H of a fractional Brownian motion from discrete noisy data, observed along a high frequency sampling scheme. The presence of the systematic experimental noise makes harder the recovering of H since relevant information is mostly contained in the high frequencies of the signal.

We quantify the difficulty of the statistical problem in the minimax sense, and prove that the rate $n^{-1/(4H+2)}$ is optimal for estimating H. Our estimators are based on adaptive estimation of quadratic functionals using wavelets.

1. Introduction

1.1. **Motivation.** Many phenomena arising in physics, molecular biology or traffic networks, possess, or are suspected to possess self-similar properties that are essential for their understanding or modelling. Recovering these so-called scaling exponents from experimental data is a challenging and ongoing issue (Abry and Veitch [2], Willinger *et al.* [29], West and Grigolini [27], Scafetta *et al.* [22] and the references therein). The purpose of this paper is to investigate a new statistical method for estimating self-similarity, based on adaptive estimation of quadratic functionals of the noisy data using wavelets. We keep to dimension 1 and focus on the paradigmatic example of fractional Brownian motion.

1.2. **Statistical model.** Let X be a one-dimensional process of interest, that has the form

$$X_t = \sigma W_t^H,$$

where W^H is a fractional Brownian motion defined on the real line, with self-similar index (or Hurst parameter) $H \in (0,1)$ and scaling parameter $\sigma \in (0,+\infty)$. Both H and σ are unknown.

In practice, it is unrealistic to assume that a sample path of X can be observed (in which case the parameters H and σ would be identified). Instead, X is rather observed at discrete times with frequency n over a time interval, say $[0,1]$. The problem of estimating H (and σ) in this context has

2000 *Mathematics Subject Classification.* 60G18; 62G99; 62F12; 62M09.

Key words and phrases. Scaling exponent; Noisy data; High frequency data; Fractional Brownian motion; Adaptive estimation of quadratic functionals, Wavelet methods.

Acknowledgement: we are grateful to Yuri Golubev, François Roueff and Philippe Soulier for helpful discussions and comments.

been paid considerable attention (see e.g. Dahlhaus [6], Istas and Lang [15], Ludeña [18] among many others and the references therein).

In this paper, we take the next logical step: we assume that each observation is contaminated by noise, so we observe

$$Y_i^n = X_{i/n} + a(X_{i/n})\,\xi_i^n, \tag{1}$$

for $i = 0, \dots, n$, where the ξ_i^n are centered noise terms and $x \rightsquigarrow a(x)$ is an unknown variance function. Throughout, we assume for simplicity that $n = 2^N$. Recovering the Hurst parameter H from the data (Y_i^n) is our objective.

1.3. **Results.** We show in Theorems 1 and 2 that the rate

$$v_n(H) = n^{-1/(4H+2)}$$

is optimal for estimating H. This means that (2) and (3) below agree, with $v_n = v_n(H)$. The accuracy $v_n(H)$ is slower by a polynomial order than the usual $n^{-1/2}$ obtained in the absence of noise. The difficulty lies in the fact the the information about H is contained in the high frequencies of the signal $t \rightsquigarrow X_t$. Although the high frequency sampling rate n usually allows to recover H with the classical rate $n^{-1/2}$ when $X_{\cdot/n}$ is directly observed (e.g. by means of quadratic variations, see [15]) the presence of the ξ_i^n in our context significantly alters the nature of the problem.

2. Main results

2.1. **Methodology.** We denote by $\mathcal{D} \subset (0,1) \times (0,+\infty)$ the parameter set in which lies (H,σ). The process X and the noise variables (ξ_i^n) are defined on some common probability space endowed with a probability measure $\mathbb{P}^n_{H,\sigma}$. Denote by $\mathcal{Y}$ the sigma field generated by the observations (1). The joint law of (Y_i^n) is thus the restriction $\mathbb{P}^n_{H,\sigma|\mathcal{Y}}$ so the sequence of statistical experiments

$$\mathcal{E} = \left(\mathbb{P}^n_{H,\sigma|\mathcal{Y}}, (H,\sigma) \in \mathcal{D}\right)_{n \geq 1}$$

specifies our mathematical model. Note that $\mathcal{E}$ also implicitly depends on the choice of $x \rightsquigarrow a(x)$ and the conditional joint law of the (ξ_i^n) given X.

A rate $v_n \to 0$ is said to be achievable over $\mathcal{D}$ if there exists a (sequence of) estimator(s) $\widehat{H}_n$ such that the (sequence of) normalized error(s)

$$v_n^{-1}(\widehat{H}_n - H) \tag{2}$$

is bounded in $\mathbb{P}^n_{H,\sigma}$–probability, uniformly over $\mathcal{D}$. The rate v_n is said to be a lower rate of convergence over $\mathcal{D}$ if there exists $c > 0$ such that

$$\liminf_{n\to\infty} \inf_F \sup_{(H,\sigma)\in\mathcal{D}} \mathbb{P}^n_{H,\sigma}\{v_n^{-1}|F - H| \geq c\} > 0, \tag{3}$$

where the infimum is taken over all estimators (i.e. $\mathcal{Y}$-measurable variables).

2.2. **The estimation strategy.** The fact that X is a fractional Brownian motion enables to predict that its energy levels

$$Q_j := \sum_k d_{j,k}^2 := \sum_k \left(\int_{\mathbb{R}} X_s \, \psi_{j,k}(s) \mathrm{d}s\right)^2 \tag{4}$$

scale[1] with a ratio related to H:

$$\frac{Q_{j+1}}{Q_j} \sim 2^{-2H}, \tag{5}$$

up to an error term that vanishes as the frequency level j increases. Here, $d_{j,k}$ is the random wavelet coefficient of X related to a certain wavelet basis $(\psi_{j,k}, j \geq 0, k \in \mathbb{Z})$. In Section 3.2 below, we construct a procedure

$$(Y_{i/n}) \rightsquigarrow (\widehat{d_{j,k,n}^2}, k = 0, \ldots, 2^j - 1, 0 \leq j \leq J_n) \tag{6}$$

that processes the data into estimates of the squared wavelet coefficients $d_{j,k}^2$ up to the maximal resolution level $J_n = [\frac{1}{2}\log_2(n)]$. We obtain a family of estimators for H by setting

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

with

$$\widehat{Q}_{j,n} = \sum_k \widehat{d_{j,k,n}^2}.$$

The ratio level j between two estimated energy levels that contains maximal information about H is chosen by means of a block thresholding rule see below. The rule is inspired by the methodology intoduced for the adaptive estimation of quadratic functionals (among others: Efromovich and Low [8], Gayraud and Tribouley [10], Tribouley [24] and the references therein).

2.3. **Statement of the results.** We consider parameter sets of the form:

$$\mathcal{D} := [H_-, H_+] \times [\sigma_-, \sigma_+] \subset (\tfrac{1}{2}, 1) \times (0, +\infty). \tag{7}$$

Assumption A. *(i) The function $x \rightsquigarrow a(x)$ is bounded, continuously differentiable with a bounded derivative.*
(ii) The continuous time process X is $\mathcal{F}^n$-adapted with respect to a filtration $\mathcal{F}^n = (\mathcal{F}_t^n, t \geq 0)$.
(iii) The noise term ξ_i^n at time i/n is $\mathcal{F}_{(i+1)/n}^n$-measurable. Moreover:

$$\mathbb{E}_{H,\sigma}^n\{\xi_i^n \,|\, \mathcal{F}_{i/n}^n\} = 0, \qquad \mathbb{E}_{H,\sigma}^n\big\{(\xi_i^n)^2 \,|\, \mathcal{F}_{i/n}^n\big\} = 1,$$

and $\sup_{(H,\sigma)\in\mathcal{D}} \sup_{i,n} \mathbb{E}_{H,\sigma}^n\big\{(\xi_i^n)^4\big\} < +\infty$.

Theorem 1. *Grant Assumption A. The rate $v_n(H) := n^{-1/(4H+2)}$ is achievable for estimating H, over any parameter space of the form* (7)*. Moreover, the estimator constructed in Section 3 and given by* (9)–(11) *below achieves the rate $v_n(H)$.*

---

[1] As for the approximation symbol $\sim$, we do not yet specify; see Proposition 1 below.

This rate is indeed optimal as soon as the noise process enjoys some regularity:

Assumption B. *(i) We have* $\inf_x a(x) > 0$.
(ii) Conditional on X*, the variables* ξ_i^n *are independent, absolutely continuous with* $\mathcal{C}^2$ *densities* $x \rightsquigarrow \exp\big(-v_{i,n}(x)\big)$ *vanishing at infinity (together with their derivatives) at a rate strictly faster than* $1/x^2$ *and:*

$$\sup_{i,n} \mathbb{E}\big\{\big(\tfrac{d}{dx}v_{i,n}(\xi_i^n)\big)^2(1+|\xi_i^n|^2)\big\} < +\infty. \tag{8}$$

Moreover, the functions $x \rightsquigarrow \frac{d^2}{dx^2}v_{i,n}(x)$ *are Lipschitz continuous, with Lipschitz constants independent of* i, n.

Theorem 2. *Grant Assumptions A and B. For estimating* H*, the rate* $v_n(H) := n^{-1/(4H+2)}$ *is a lower rate of convergence over any parameter set of the form* (7) *with non empty interior.*

We complete this section by an ancillary result about the estimation of the scaling parameter σ, although we are primarily interested in recovering H. The estimation of σ has been adressed by Gloter and Jacod [12] for the case $H = 1/2$ and by Gloter and Hoffmann [11] in a slightly different model when $H \geq 1/2$ is known. Altogether, the rate $v_n(H)$ is proved to be optimal for estimating σ when H is *known*. Our next result shows that we loose a logarithmic factor when H is *unknown*.

Theorem 3. *Grant Assumptions A and B. For estimating* σ*, the rate* $n^{-1/(4H+2)}\log(n)$ *is a lower rate of convergence over any parameter set of the form* (7).

2.4. **Discussion.**

2.4.1. *About the rate.* We see that the presence of noise dramatically alters the accuracy of estimation of the Hurst parameter: the optimal rate $v_n(H) = n^{-1/(4H+2)}$ inflates by a polynomial order as H increases. In particular, the classical (parametric rate) $n^{-1/2}$ is obtained by formally letting H tend to 0 (a case we do not have here).

2.4.2. *About Theorem 1.* The restriction $H_- > 1/2$ is important here, and it is linked to the discretization effect of the estimator. Assumption A can easily be fulfilled in the case of a noise process that is independent of the signal X. It is not minimal: more general noise processes could presumably be considered, and, more interestingly, more general scaling processes than fractional Brownian motion as well. To this end, it is required that the energy levels of X satisfy Proposition 1 and that the empirical energy levels satisfy Proposition 2 in Section 4 below. We do not pursue that here. See also Lang and Roueff [17].

2.4.3. *About Theorem 2.* It should be emphasised that our lower bound is local, in the sense that $\mathcal{D}$ can be taken arbitrarily small in the class specified by (7). Observe that since the rate $v_n(H)$ depends on the parameter value, the min-max lower bounds (3) are only really meaningful for parameter sets $\mathcal{D}$ that are concentrated around some given value of H.

Assumption B (ii) is not minimal; it simply ensures that if we replace $X_{\cdot/n}$ by a single unknown value θ, each translation-dilatation model $\theta \mapsto \theta + a(\theta)\xi_i^n$ admit a finite Fisher information. It is satisfied in particular when the ξ_i^n are i.i.d. centered Gaussian. More generally, any noise process would yield the same lower bound as soon as Proposition 4 is satisfied (see Section 6.1).

2.4.4. *The stationary case.* Golubev [13] remarked that in the particular case of i.i.d. Gaussian noises, independent of W^H, a direct spectral approach is simpler. Indeed, the observation generated by the $Y_{i/n} - Y_{(i-1)/n}$ becomes stationary Gaussian, and a classical Whittle estimator shall do (Whittle [28] or Dahlhaus [6]). In particular, although some extra care has to be taken about the approximation in n, such an approach would certainly prove simpler in that specific context for obtaining the lower bound.

2.4.5. *Quadratic variations alternatives.* Our estimator (precisely constructed in Section 3 below) can be linked to more traditional quadratic variations methods. Indeed, the fundamental energy levels Q_j defined in (4) can be obtained from the quadratic variations of $X_{\cdot/n}$ in the particular case of the Schauder basis (which has not sufficiently many vanishing moments for our purpose). However, the choice of an optimal j remains and we were not able to obtain the exact rate of convergence by this approach. We do not pursue that here.

2.5. **Organisation of the paper.** In Section 3, we give the complete construction of an optimal estimator $\widehat{H}_n$ that achieves the minimax rate $v_n(H)$. Section 4 explores the properties of the energy levels of X (Proposition 1) as well as their empirical version (Proposition 2). Theorem 1 is proved in Section 5. Finally, Sections 6 and 7 are devoted to the lower bounds. It is noteworthy that the complex stochastic structure of our model due to the two sources or randomness requires particular efforts for the lower bound. Our strategy to obtain lower bounds is outlined in the Section 6, it requires a delicate 'coupling' result proved in Section 7. The proof of some technical results are leaved to the Appendix in Section 8.

3. Construction of an estimator

3.1. Pick a wavelet basis $(\psi_{j,k}, j \geq 0, k \in \mathbb{Z})$ generated by a mother wavelet ψ with two vanishing moments and compact support in $[0, S]$ where S is some integer. The basis is fixed throughout the Sections 3–5. Assuming we have estimators $\widehat{d^2_{j,k,n}}$ of the squared wavelet coefficients, recalling the definition (4) of the energy levels, we obtain a family of estimators for H by setting

$$\widehat{H}_{j,n} := -\frac{1}{2} \log_2 \frac{\widehat{Q}_{j+1,n}}{\widehat{Q}_{j,n}}, \quad j = \underline{J}, \ldots, J_n - 1,$$

with

$$\widehat{Q}_{j,n} = \sum_{k=0}^{2^{j-1}-1} \widehat{d^2_{j,k,n}},$$

where $J_n := [\frac{1}{2}\log_2(n)]$ is a maximum level of detail needed in our statistical procedure and $\underline{J} := [\log_2(S-1)] + 2$ is some (irrelevant) minimum level introduced to avoid border effect while computing wavelet coefficient corresponding to location on $[0, 1/2]$ from an observation corresponding to $[0, 1]$. Following Gayraud and Tribouley [10] in the context of adaptive estimation of quadratic functionals, we let

$$J_n^\star := \max\left\{j = \underline{J}, \ldots, J_n \ : \ \widehat{Q}_{j,n} \geq \tfrac{2^j}{n}\right\} \tag{9}$$

(and in the case the set above is empty we let $J_n^\star = \underline{J}$ so that everything remains meaningful in the sequel). Eventually, our estimator of H is

$$\widehat{H}_{J_n^\star,n}. \tag{10}$$

The performance of $\widehat{H}_{J_n^\star,n}$ is related to scaling properties of X and the accuracy of the procedure (6).

3.2. Preliminary estimation of the $d^2_{j,k}$. Since ψ has compact support in $[0, S]$, the wavelet coefficient $d_{j,k}$ writes:

$$d_{j,k} = \sigma \sum_{l=0}^{S2^{N-j}-1} \int_{k/2^j+l/2^N}^{k/2^j+(l+1)/2^N} \psi_{j,k}(t) W_t^H \mathrm{d}t.$$

This suggests the approximation:

$$\widetilde{d}_{j,k,n} = \sum_{l=0}^{S2^{N-j}-1} \Big(\int_{k/2^j+l/2^N}^{k/2^j+(l+1)/2^N} \psi_{j,k}(t)\mathrm{d}t \Big) \, Y^n_{k2^{N-j}+l},$$

for, $\underline{J} \leq j \leq J_n, 0 \leq k \leq 2^{j-1} - 1$. The difference $\tilde{d}_{j,k,n} - d_{j,k}$ splits into $b_{j,k,n} + e_{j,k,n}$, respectively a bias term and a centered noise term:

$$\begin{aligned}
b_{j,k,n} &= -\sum_{l=0}^{S2^{N-j}-1} \int_{k/2^j+l/2^N}^{k/2^j+(l+1)/2^N} \psi_{j,k}(t)(X_t - X_{k/2^j+l/2^N})\mathrm{d}t, \\
e_{j,k,n} &= \sum_{l=0}^{S2^{N-j}-1} \Big(\int_{k/2^j+l/2^N}^{k/2^j+(l+1)/2^N} \psi_{j,k}(t)\mathrm{d}t \Big) \, a(X_{k/2^j+l/2^N}) \xi^n_{k2^{N-j}+l} \, .
\end{aligned}$$

We denote by $v_{j,k,n}$ the variance of $e_{j,k,n}$, conditional on $\mathcal{F}^n_{k2^{-j}}$, which is equal to

$$v_{j,k,n} = \sum_{l=0}^{S2^{N-j}-1} \Big(\int_{k/2^j+l/2^N}^{k/2^j+(l+1)/2^N} \psi_{j,k}(t)\mathrm{d}t \Big)^2 \mathbb{E}^n_{H,\sigma}[a(X_{k/2^j+l/2^N})^2 \mid \mathcal{F}^n_{k2^{-j}}].$$

The conditional expectations appearing in this expression are close to $a(X_{k/2^j})^2$ and thus may be estimated from the observation without the knowledge of H, σ. We define

$$\widehat{a^2}_{k/2^j,n} := 2^{-N/2} \sum_{l'=1}^{2^{N/2}} (Y^n_{k2^{N-j}+l'})^2 - \Big(2^{-N/2} \sum_{l'=1}^{2^{N/2}} Y^n_{k2^{N-j}+l'} \Big)^2,$$

and we set

$$\overline{v}_{j,k,n} = \sum_{l=0}^{S2^{N-j}-1} \Big(\int_{k/2^j+l/2^N}^{k/2^j+(l+1)/2^N} \psi_{j,k}(t)\mathrm{d}t\Big)^2 \widehat{a^2}_{k/2^j,n}.$$

Eventually, we set

$$\widehat{d^2_{j,k,n}} := (\widetilde{d}_{j,k,n})^2 - \overline{v}_{j,k,n}. \tag{11}$$

and $\widehat{H}_{J_n^\star,n}$ is well-defined. Remark that if the function a is assumed known, one can considerably simplify the construction of the approximation $\widehat{a^2}_{k/2^j,n}$.

4. The behaviour of the energy levels

We denote by $\mathbb{P}_{H,\sigma}$ the law of $X = \sigma W^H$, defined on an appropriate probability space. We recall the expression of the energy at level j:

$$Q_j = \sum_{k=0}^{2^{j-1}-1} d_{j,k}^2.$$

Proposition 1. *(i) For all $\varepsilon > 0$, there exists $r_-(\varepsilon) \in (0,1)$ such that*

$$\inf_{(H,\sigma)\in\mathcal{D}} \mathbb{P}_{H,\sigma}\big\{ \inf_{j\geq 1} 2^{2jH} Q_j \geq r_-(\varepsilon)\big\} \geq 1-\varepsilon. \tag{12}$$

(ii) The sequence

$$Z_j := 2^{j/2} \sup_{l\geq j} \big| \frac{Q_{l+1}}{Q_l} - 2^{-2H} \big| \tag{13}$$

is bounded in $\mathbb{P}_{H,\sigma}$-probability, uniformly over $\mathcal{D}$.

Proposition 2. *Let $j_n(H) := [\frac{1}{2H+1}\log_2(n)]$. Then $J_n \geq j_n(H)$ for all $H \in [H_-, H_+]$ and for any $L > 0$*

$$\Big\{ n 2^{j_n(H)/2} \sup_{J_n \geq j \geq j_n(H)-L} 2^{-j} \Big| \widehat{Q}_{j,n} - Q_j \Big| \Big\}_{n \geq (4(S-1)2^L)^{(1+2H)}}$$

is bounded in $\mathbb{P}^n_{H,\sigma}$-probability, uniformly over $\mathcal{D}$.

We shall see below that Proposition 1 and 2 together imply Theorem 1.

4.1. Fractional Brownian motion. The fractional Brownian motion admits the harmonizable representation

$$W_t^H = \int_{\mathbb{R}} \frac{e^{it\xi}-1}{(i\xi)^{H+1/2}} \mathrm{B}(\mathrm{d}\xi),$$

where B is a complex Gaussian measure (Samorodnitsky and Taqqu [21]). Another representation using a standard Brownian motion B on the real line is given by

$$W_t^H = \frac{1}{\Gamma(H+1/2)} \int_{-\infty}^{\infty} [(t-s)_+^{H-1/2} - s_+^{H-1/2}] \mathrm{d}B_s.$$

(Γ is the Euler function.) The process W^H is H self–similar and the covariance structure of W^H is explicitly given by

$$\mathrm{Cov}(W_s^H, W_t^H) = \tfrac{\kappa(H)}{2}\big\{|t|^{2H} + |s|^{2H} - |t-s|^{2H}\big\}$$

where $\kappa(H) = \frac{\pi}{H\Gamma(2H)\sin(\pi H)}$. Recall that $d_{j,k} = \int_{\mathbb{R}} \psi_{j,k}(s) X_s \mathrm{d}s$ denote the random wavelet coefficients of X, given a wavelet ψ with 2 vanishing moments. It can be seen, using the stationarity of the increments of W^H that, for a fixed level j, the sequence $(d_{j,k})_{k\in\mathbb{Z}}$ is centred Gaussian, and stationary with respect to the location parameter k. Moreover the coefficients have the self similarity property

$$(d_{j,k})_{k\in\mathbb{Z}} \overset{\text{law}}{=} 2^{-j(H+1/2)}(d_{0,k})_{k\in\mathbb{Z}},$$

see Kawasaki and Morita [16], Delbeke and Abry [7], Abry and Veitch [2], [25], Abry *et al.*, [26], [1]). Moreover,

$$\mathrm{Var}\{d_{j,k}\} = \sigma^2 c(\psi)\kappa(H) 2^{-j(1+2H)},$$

where $c(\psi) = \frac{1}{2}\int \psi(s)\psi(t)\{|t|^{2H} + |s|^{2H} - |t-s|^{2H}\}\mathrm{d}s\mathrm{d}t$, and the covariance

$$\mathrm{Cov}(d_{j,k}, d_{j,k'}) = 2^{-j(2H+1)}\mathrm{Cov}(d_{0,k}, d_{0,k'})$$

decays polynomially as $k - k' \to \infty$ due to the two vanishing moments of ψ and

$$|\mathrm{Cov}(d_{0,k}, d_{0,k'})| \leq c(1 + |k-k'|)^{2(H-2)},$$

for some c which does not depend on σ nor H. See also Tewfik and Kim [23], Hirchoren and D'Attelis [14], Istas and Lang [15], Gloter and Hoffmann [11].

Proposition 3. *We have*

$$\sup_{(H,\sigma)\in\mathcal{D}} \mathbb{E}_{H,\sigma}\left\{[Q_j - 2^{-2jH}\frac{\sigma^2}{2}c(\psi)\kappa(H)]^2\right\} \leq c2^{-j(1+4H)}.$$

Proof. Remark that by stationarity:

$$Q_j - 2^{-2jH}\frac{\sigma^2}{2}c(\psi)\kappa(H) = \sum_{k=0}^{2^{j-1}-1}(d_{j,k}^2 - \mathbb{E}_{H,\sigma}\{d_{j,k}^2\}).$$

Then the variance of the sum above is evaluated using the decorrelation property of the wavelet coefficients (similar computations can be found in Istas and Lang [15] or Gloter and Hoffmann [11]). □

4.2. Proof of Proposition 1. By Proposition 3, we derive in the same way as in Lemma II.4. of Ciesielski *et al.* [5] that, for all $\varepsilon > 0$:

$$\sum_{j\geq 0}\sup_{(H,\sigma)\in\mathcal{D}} \mathbb{P}_{H,\sigma}\{2^{2jH}Q_j \notin [\frac{\sigma^2}{2}c(\psi)\kappa(H) - \varepsilon, \frac{\sigma^2}{2}c(\psi)\kappa(H) + \varepsilon]\} < \infty,$$

from which (i) easily follows. By (i), the probability that $|Z_j|$ is greater that a constant M is less than

$$\varepsilon + \mathbb{P}_{H,\sigma}\{\sup_{l\geq j}|Q_{l+1} - 2^{-2H}Q_l|2^{2lH} \geq M2^{-\frac{j}{2}}r_-(\varepsilon)\}. \tag{14}$$

By self-similarity $\mathbb{E}_{H,\sigma}\{Q_{l+1} - 2^{-2H}Q_l\} = 0$. By Markov's inequality, (14) is less than

$$\varepsilon + [M^2 r_-^2(\varepsilon)]^{-1}\sum_{l\geq j}\mathrm{Var}_{H,\sigma}\left\{Q_{l+1} - 2^{-2H}Q_l\right\}2^{4lH}2^j.$$

By Proposition 3, the sum above can be made arbitrarily small for large enough M, which proves (ii).

4.3. Proof of Proposition 2. We first claim that the following estimates holds:

$$\sup_{J_n \geq j \geq j_n(H)-L} \sup_{(H,\sigma)\in\mathcal{D}} 2^{-j/2}\mathbb{E}^n_{H,\sigma}\{\big|\widehat{Q}_{j,n} - Q_j\big|\} \leq cn^{-1}. \tag{15}$$

Proposition 2 readily follows. To prove (15), we first split $\widehat{Q}_{j,n} - Q_j$ into $\sum_{u=1}^6 r^{(u)}_{j,n}$, with

$$r^{(1)}_{j,n} = \sum_k b^2_{j,k,n}, \qquad r^{(2)}_{j,n} = \sum_k (e^2_{j,k,n} - v_{j,k,n}),$$

$$r^{(3)}_{j,n} = \sum_k (v_{j,k,n} - \overline{v}_{j,k,n}), \qquad r^{(4)}_{j,n} = 2\sum_k b_{j,k,n} d_{j,k},$$

$$r^{(5)}_{j,n} = 2\sum_k e_{j,k,n} d_{j,k}, \qquad r^{(6)}_{j,n} = 2\sum_k b_{j,k,n} e_{j,k,n}.$$

Using that $\mathbb{E}_{H,\sigma}\{(X_t - X_s)^2\} \leq c(H)\sigma^2|t-s|^{2H}$, it is readily seen that $\mathbb{E}^n_{H,\sigma}\{(b_{j,k,n})^2\}$ is less than a constant times $2^{-j}n^{-2H}$. Summing in k shows that the term $r^{(1)}_{j,n}$ is negligible since $H > 1/2$.

Using that $e^2_{j,k,n} - v_{j,k,n}$ are uncorrelated for $|k - k'| \geq S$, we deduce that $\mathbb{E}^n_{H,\sigma}[(r^{(2)}_{j,n})^2]$ is bounded by a constant times $\sum_{k=0}^{2^{j-1}-1}\{\mathbb{E}^n_{H,\sigma}[e^4_{j,k,n}] + \mathbb{E}^n_{H,\sigma}[v^2_{j,k,n}]\}$. Then using the martingale increments structure of the sequence $a(X_{k2^{-j}+l2^{-N}})\xi^n_{k2^{-j}+l2^{-N}}$ for $l = 0, \dots, S2^{N-j}$ enables to apply the Burkölder-Davis inequality. This gives, by Assumption A: $\mathbb{E}^n_{H,\sigma}[e^4_{j,k,n}] \leq cn^{-2}$. Then since $x \rightsquigarrow a(x)$ is bounded and thus $v_{j,k,n} \leq cn^{-1}$ we obtain that $\mathbb{E}^n_{H,\sigma}[(r^{(2)}_{j,n})^2]$ has the right order $2^j n^{-2}$.

Using conditional centering of $e_{j,k,n}$ with the fact that the variance of $d_{j,k}$ is less than $c2^{-j(2H+1)}$ and the condition $j \geq j_n(H) - L = [\frac{1}{2H+1}\log_2(n)] - L$, one easily checks that the terms $r^{(4)}_{j,n}$, $r^{(5)}_{j,n}$ and $r^{(6)}_{j,n}$ have negligible order.

We finally turn to the important term $r^{(3)}_{j,n}$, which encompasses the estimation of a. We claim that, for $0 \leq l \leq S2^{N-j} - 1$, the following estimates holds:

$$\mathbb{E}^n_{H,\sigma}\{\big|\widehat{a^2}_{k/2^j,n} - \mathbb{E}^n_{H,\sigma}[a(X_{k/2^j+l/2^N})^2 \mid \mathcal{F}^n_{k2^{-j}}]\big|\} \leq cn^{-1/4}. \tag{16}$$

Summing in l and k yields the result for $r^{(3)}_{j,n}$ as soon as (16) is proved: Indeed, since $v_{j,k,n} - \overline{v}_{j,k,n}$ is equal to

$$\sum_{l=0}^{S2^{N-j}-1} \Big(\int_{k/2^j+l/2^N}^{k/2^j+(l+1)/2^N} \psi_{j,k}(t)\mathrm{d}t\Big)^2 [\widehat{a^2}_{k/2^j,n} - \mathbb{E}^n_{H,\sigma}[a(X_{k/2^j+l/2^N})^2 \mid \mathcal{F}^n_{k2^{-j}}]],$$

we have that $\mathbb{E}^n_{H,\sigma}\{|r^{(3)}_{j,n}|\}$ is less than $c2^{j/2}n^{-1}2^{j/2}n^{-1/4}$. Therefore, under the restriction $j \leq J_n \leq [\frac{1}{2}\log_2(n)]$, (15) holds. It remains to prove (16).

We have $\widehat{a^2}_{k/2^j,n} - \mathbb{E}^n_{H,\sigma}[a(X_{k/2^j+l/2^N})^2 \mid \mathcal{F}^n_{k2^{-j}}] = t^{(1)}_{k,n} + t^{(2)}_{k,l,n} + t^{(3)}_{k,n}$, with

$$t^{(1)}_{k,n} = 2^{-\frac{N}{2}} \sum_{l'=1}^{2^{N/2}} X^2_{k/2^j+l'/2^N} - \Big(2^{-\frac{N}{2}} \sum_{l'=1}^{2^{N/2}} Y^n_{k/2^{N-j}+l'}\Big)^2$$

$$t^{(2)}_{k,l,n} = 2^{-\frac{N}{2}} \sum_{l'=1}^{2^{N/2}} a(X_{k/2^j+l'/2^N})^2(\xi^n_{k2^{j-N}+l'})^2 - \mathbb{E}^n_{H,\sigma}[a(X_{k/2^j+l/2^N})^2 \mid \mathcal{F}^n_{k2^{-j}}],$$

$$t^{(3)}_{k,n} = 2^{-\frac{N}{2}+1} \sum_{l'=1}^{2^{N/2}} X_{k/2^j+l'/2^N} a(X_{k/2^j+l'/2^N})\xi^n_{k2^{j-N}+l'}.$$

Since the $\xi^n_{k2^{j-N}+l'}$ are uncorrelated and centered, we readily have that the expectation of $|t^{(3)}_{k,n}|$ is of order $2^{-N/4} = n^{-1/4}$. For the term $t^{(2)}_{k,l,n}$, we use the preliminary decomposition

$$t^{(2)}_{k,l,n} = 2^{-N/2} \sum_{l'=1}^{2^{N/2}} a(X_{k/2^j+l'/2^N})^2[(\xi^n_{k2^{j-N}+l'})^2 - 1]$$
$$+ \ 2^{-N/2} \sum_{l'=1}^{2^{N/2}} [a(X_{k/2^j+l'/2^N})^2 - \mathbb{E}^n_{H,\sigma}[a(X_{k/2^j+l/2^N})^2 \mid \mathcal{F}^n_{k2^{-j}}]].$$

The $\mathbf{L}^1$-norm of the first term above is of order $n^{-1/4}$ since the summands $a(X_{k/2^j+l'/2^N})^2[(\xi^n_{k2^{j-N}+l'})^2-1]$ are martingale increments with second-order moments by Assumption A. Likewise, since $x \rightsquigarrow a(x)$ has a bounded derivative and

$$\mathbb{E}_{H,\sigma}\{(X_{k/2^j+l'/2^N} - X_{k/2^j})^2\} \leq c(H)\sigma^2(2^{-N/2})^{2H},$$
$$\mathbb{E}_{H,\sigma}\{(X_{k/2^j+l/2^N} - X_{k/2^j})^2\} \leq c(H)\sigma^2(2^{-j/2})^{2H},$$

the second term in the expression of $t^{(2)}_{k,l,n}$ has $\mathbf{L}^1$-norm less than a constant times $(2^{-j/2})^H \leq 2^{j_n(H)H/2} = n^{-H/(1+2H)}$, and thus has the right order since $H \geq 1/2$.

Finally, we further need to split $t^{(1)}_{k,n}$ into

$$2^{-N/2} \sum_{l'=1}^{2^{N/2}} X^2_{k/2^j+l'/2^N} - \big(2^{-N/2} \sum_{l'=1}^{2^{N/2}} X_{k/2^j+l'/2^N}\big)^2$$
$$- \ \big(2^{-N/2} \sum_{l'=1}^{2^{N/2}} a(X_{k/2^j+l'/2^N})\xi^n_{k2^{N-j}+l'}\big)^2$$
$$- \ 2\big(2^{-N/2} \sum_{l'=1}^{2^{N/2}} X_{k/2^j+l'/2^N}\big)\big(2^{-N/2} \sum_{l'=1}^{2^{N/2}} a(X_{k/2^j+l'/2^N})\xi^n_{k2^{N-j}+l'}\big).$$

The first term and second term are easily seen to be of the right order respectively by the smoothness property of X and the uncorrelation property of the variables ξ^n_i. The third term is seen to have the right order after remarking that one can replace the first sum $2^{-N/2}\sum_{l'=1}^{2^{N/2}} X_{k/2^j+l'/2^N}$ by

$X_{k/2^j}$ up to a negligible error and then use the conditional uncorrelation of ξ_i^n again. Thus (16) is proved, hence (15) follows. The proof of Proposition 2 is complete.

5. Proof of Theorem 1

First we need the following result that states the the level $J_n^\star$, based on the data, is with large probability greater that some level based on the knowledge of H.

5.1. A fundamental lemma.

For $\varepsilon > 0$, define

$$J_n^-(\varepsilon) := \max\Big\{ j \geq 1 \; ; \; r^-(\varepsilon) 2^{-2jH} \geq \frac{2^j}{n} \Big\}. \tag{17}$$

Lemma 1. *For all $\varepsilon > 0$, there exists $L(\varepsilon) > 0$ such that*

$$\sup_{(H,\sigma)\in\mathcal{D}} \mathbb{P}^n_{H,\sigma} \big\{ J_n^\star < J_n^-(\varepsilon) - L(\varepsilon) \big\} \leq \varepsilon + \varphi_n\big(\varepsilon\big),$$

where φ_n satisfies $\lim_{n\to\infty}\varphi_n(\varepsilon) = 0$.

Proof. Let $L, \varepsilon > 0$. By definition of $J_n^-(\varepsilon)$,

$$\tfrac{1}{2} r_-(\varepsilon)^{1/(1+2H)} n^{1/(1+2H)} \leq 2^{J_n^-(\varepsilon)} \leq r_-(\varepsilon)^{1/(1+2H)} n^{1/(1+2H)}$$

hence for large enough n, we have $\underline{J} \leq J_n^-(\varepsilon) - L \leq J_n$. Thus, by (9), $\mathbb{P}^n_{H,\sigma}\{J_n^\star \geq J_n^-(\varepsilon) - L\}$ is greater than

$$\mathbb{P}^n_{H,\sigma}\Big\{ \widehat{Q}_{J_n^-(\varepsilon)-L,n} \geq 2^{J_n^-(\varepsilon)-L} n^{-1} \Big\}$$

that we rewrite as

$$\mathbb{P}^n_{H,\sigma}\Big\{ \widehat{Q}_{J_n^-(\varepsilon)-L,n} - Q_{J_n^-(\varepsilon)-L} \geq 2^{J_n^-(\varepsilon)-L} n^{-1} - Q_{J_n^-(\varepsilon)-L} \Big\}$$

and that we bound from below by

$$\mathbb{P}^n_{H,\sigma}\Big\{ \widehat{Q}_{J_n^-(\varepsilon)-L,n} - Q_{J_n^-(\varepsilon)-L} \geq 2^{J_n^-(\varepsilon)-L} n^{-1} - 2^{-2(J_n^-(\varepsilon)-L)H} r_-(\varepsilon) \Big\} -$$

$$-\mathbb{P}_{H,\sigma}\{\inf_{j\geq 1} 2^{2jH} Q_j < r_-(\varepsilon)\}.$$

Proposition 1 (i) and the definition of $J_n^-(\varepsilon)$ yield that this last term is greater than:

$$\mathbb{P}^n_{H,\sigma}\{ \widehat{Q}_{J_n^-(\varepsilon)-L,n} - Q_{J_n^-(\varepsilon)-L} \geq r_-(\varepsilon)^{1/(2H+1)} n^{-2H/(2H+1)} (2^{-L} - 2^{2LH})\} - \varepsilon.$$

Then, if L is such that $2^L - 2^{2LH} \leq -1$, an assumption we shall make from now on, Lemma 1 is proved, provided we show that

$$\mathbb{P}^n_{H,\sigma}\Big\{ |\widehat{Q}_{J_n^-(\varepsilon)-L,n} - Q_{J_n^-(\varepsilon)-L}| \geq r_-(\varepsilon)^{1/(2H+1)} n^{-2H/(2H+1)} \Big\} \tag{18}$$

can be made arbitrarily small, uniformly in (H, σ). Using again

$$2^{J_n^-(\varepsilon)} > \tfrac{1}{2} n^{1/(2H+1)} r_-(\varepsilon)^{1/(2H+1)},$$

we can pick $L' = L'(\varepsilon) > 0$ independent of n such that

$$J_n^-(\varepsilon) - L \geq j_n(H) - L'(\varepsilon),$$

therefore (18) is less than

$$\mathbb{P}^n_{H,\sigma}\Big\{\sup_{J_n\geq j\geq j_n(H)-L'(\varepsilon)}|\widehat{Q}_{j,n}-Q_j|\geq r_-(\varepsilon)^{1/(2H+1)}n^{-2H/(2H+1)}\Big\},$$

that we rewrite as

$$\mathbb{P}^n_{H,\sigma}\Big\{n2^{j_n(H)/2}\sup_{J_n\geq j\geq j_n(H)-L'(\varepsilon)}2^{-(j_n(H)-L'(\varepsilon))}|\widehat{Q}_{j,n}-Q_j|\geq v_H(\varepsilon,n)\Big\},$$

where

$$v_H(\varepsilon,n):=2^{L'(\varepsilon)}r_-(\varepsilon)^{1/(2H+1)}n^{1/(4H+2)}$$

and where we use that $2^{j_n(H)}$ is of order $n^{1/(2H+1)}$. We conclude by applying Proposition 2, using that for fixed $\varepsilon>0$, $2^{L'(\varepsilon)}r_-(\varepsilon)^{1/(2H+1)}n^{1/(4H+2)}\to\infty$ as $n\to\infty$. The uniformity in (H,σ) is straightforward. $\square$

5.2. Proof of Theorem 1, completion. Since $t\rightsquigarrow 2^{-2t}$ is invertible on $(0,1)$ with inverse uniformly Lipschitz on the compact sets of $(0,1)$, it suffices to prove Theorem 1 with 2^{-2H} in place of H and $\widehat{Q}_{J_n^\star+1,n}/\widehat{Q}_{J_n^\star,n}$ in place of $\widehat{H}_{J_n^\star,n}$. First, we bound

$$\Big|\frac{\widehat{Q}_{J_n^\star+1,n}}{\widehat{Q}_{J_n^\star,n}}-2^{-2H}\Big|$$

by a "bias" and a variance term, namely

$$\Big|\frac{Q_{J_n^\star+1}}{Q_{J_n^\star}}-2^{-2H}\Big|+\Big|\frac{\widehat{Q}_{J_n^\star+1,n}}{\widehat{Q}_{J_n^\star,n}}-\frac{Q_{J_n^\star+1}}{Q_{J_n^\star}}\Big|=:|B_n|+|V_n|\,,$$

say. Second, we prove Theorem 1 for B_n and V_n separately. Remark that the "bias" term (deterministic conditionally to the signal X), Q_{j+1}/Q_j-2^{-2H}, decreases as the level j increases, while the variance term $\widehat{Q}_{j+1,n}/\widehat{Q}_{j,n}-Q_{j+1}/Q_j$ increases. They both match at the level $j=J_n^-(\varepsilon)$. On contrary to many "bias-variance" situation, the behavior of the variance term depends on the unknown regularity of the signal through the decrease rate of the denominators $\widehat{Q}_{j,n}$ and Q_j. This explains the choice made in (9) to control the estimated level of energy $\widehat{Q}_{J_n^\star,n}$ by below.

5.2.1. *The bias term.* Let $M>0$ and $\varepsilon>0$. By Lemma 1, we have

$$\begin{aligned}
&\mathbb{P}^n_{H,\sigma}\Big\{n^{1/(4H+2)}\,|B_n|\geq M\Big\}\\
\leq\;&\mathbb{P}^n_{H,\sigma}\Big\{n^{1/(4H+2)}\,|B_n|\geq M,\,J_n^\star\geq J_n^-(\varepsilon)-L(\varepsilon)\Big\}+\varepsilon+\varphi_n(\varepsilon)\\
\leq\;&\mathbb{P}^n_{H,\sigma}\Big\{n^{1/(4H+2)}2^{-J_n^-(\varepsilon)/2}2^{L(\varepsilon)/2}|Z_{J_n^-(\varepsilon)-L(\varepsilon)}|\geq M\Big\}+\varepsilon+\varphi_n(\varepsilon)\\
\leq\;&\mathbb{P}^n_{H,\sigma}\Big\{\sqrt{2}r_-(\varepsilon)^{-1/(4H+2)}2^{L(\varepsilon)/2}|Z_{J_n^-(\varepsilon)-L(\varepsilon)}|\geq M\Big\}+\varepsilon+\varphi_n(\varepsilon),
\end{aligned}$$

where we used for the last line that, by (17)

$$2^{-J_n^-(\varepsilon)}\leq 2r_-(\varepsilon)^{-1/(2H+1)}n^{-1/(2H+1)}.$$

We conclude by Proposition 1 (ii) and taking successively ε sufficiently small, M sufficiently large and n sufficiently large.

5.2.2. *The variance term.* We split the variance term into $V_n = V_n^{(1)} + V_n^{(2)}$, where

$$V_n^{(1)} := \frac{\widehat{Q}_{J_n^\star+1,n} - Q_{J_n^\star+1}}{\widehat{Q}_{J_n^\star,n}} \quad \text{and} \quad V_n^{(2)} := \frac{Q_{J_n^\star+1}(Q_{J_n^\star} - \widehat{Q}_{J_n^\star,n})}{\widehat{Q}_{J_n^\star,n}\, Q_{J_n^\star}}.$$

Having lemma 1 in mind, we bound for any $M > 0$ and L integer, the probability $\mathbb{P}^n_{H,\sigma}\left\{n^{1/(4H+2)}|V_n^{(1)}| \geq M\right\}$ by:

$$\mathbb{P}^n_{H,\sigma}\left\{n^{1/(4H+2)}|V_n^{(1)}| \geq M,\ J_n^\star \geq J_n^-(\varepsilon) - L\right\} + \mathbb{P}^n_{H,\sigma}\left\{J_n^\star < J_n^-(\varepsilon) - L\right\}.$$

Fix $\varepsilon > 0$ and pick $L = L(\varepsilon)$ as in Lemma 1 so that the second probability $\mathbb{P}^n_{H,\sigma}\left\{J_n^\star < J_n^-(\varepsilon) - L(\varepsilon)\right\}$ is bounded by $\varepsilon + \varphi_n(\varepsilon)$. It remains now to deal with the first probability. As soon as n is large enough, $J_n^-(\varepsilon) - L(\varepsilon) > \underline{J}$ and thus by definition of $J_n^\star$, the denominator of $V_n^{(1)}$ is bounded below by $2^{J_n^\star}/n$, this yields to the new bound for the first probability:

$$\mathbb{P}^n_{H,\sigma}\left\{n^{1/(4H+2)+1}2^{-J_n^\star}|\widehat{Q}_{J_n^\star+1,n} - Q_{J_n^\star+1}| \geq M,\ J_n^\star \geq J_n^-(\varepsilon) - L(\varepsilon)\right\}.$$

Recall that we defined $j_n(H) = [\frac{1}{2H+1}\log_2(n)]$ in Proposition 2 and by definition of $J_n^-(\varepsilon)$, we have

$$2^{J_n^-(\varepsilon)} > \tfrac{1}{2}n^{1/(2H+1)}r_-(\varepsilon)^{1/(2H+1)}.$$

Therefore, we can pick a positive $L' = L'(\varepsilon)$ independent of n such that

$$J_n^-(\varepsilon) - L(\varepsilon) \geq j_n(H) - L'(\varepsilon)$$

and then one can bound the first probability by:

$$\mathbb{P}^n_{H,\sigma}\Big\{n^{1/(4H+2)+1} \sup_{J_n \geq j \geq j_n(H) - L'(\varepsilon)} 2^{-j}|\widehat{Q}_{j,n} - Q_j| \geq M\Big\}$$

Next, using that $n^{1/(4H+2)+1}$ is of order $n\,2^{j_n(H)/2}$ and Proposition 2, this term can be made arbitrarily small (uniformly in n) by taking M large enough.

We now turn to the term $V_n^{(2)}$. Fix $\varepsilon > 0$ and $M > 0$. Recalling the definition of Z_j in Proposition 1, we have

$$\begin{aligned} &\mathbb{P}^n_{H,\sigma}\left\{n^{1/(4H+2)}|V_n^{(2)}| \geq M\right\} \\ \leq\ &\mathbb{P}^n_{H,\sigma}\Big\{n^{1/(4H+2)}\Big|\frac{Q_{J_n^\star} - \widehat{Q}_{J_n^\star,n}}{\widehat{Q}_{J_n^\star,n}}\Big|\left(2^{-2H} + Z_0\right) \geq M\Big\} \end{aligned}$$

Now the tightness of the sequence Z_j implies that for some fixed constant M', this probability is less than

$$\mathbb{P}^n_{H,\sigma}\Big\{n^{1/(4H+2)}\Big|\frac{Q_{J_n^\star} - \widehat{Q}_{J_n^\star,n}}{\widehat{Q}_{J_n^\star,n}}\Big| \geq \tfrac{M}{2^{-2H}+M'}\Big\} + \varepsilon.$$

Then the conclusion follows exactly as for $V_n^{(1)}$. The proof of Theorem 1 is complete.

6. Proof of Theorems 2 and 3

Consistently with Section 4, we denote by $\mathbb{P}_{H,\sigma}$ the probability measure on the Wiener space $\mathcal{C}_0$ of continuous functions on $[0,1]$ under which the canonical process X has the law of σW^H. We write $\mathbb{P}_f^n$ for the law of the data, conditional on $X = f$ thus $\mathbb{P}^n_{H,\sigma|\mathcal{Y}} = \int_{\mathcal{C}_0} \mathbb{P}_{H,\sigma}(\mathrm{d}f)\,\mathbb{P}_f^n$.

6.1. Preliminaries. Define, for $\alpha \in (0,1)$:

$$\|f\|_{\mathcal{H}^\alpha} := \|f\|_\infty + \sup_{0\le s<t\le 1} \frac{|f(t)-f(s)|}{|t-s|^\alpha}. \tag{19}$$

The total variation of a signed measure μ is

$$\|\mu\|_{\mathrm{TV}} = \sup_{\|f\|_\infty \le 1} \left|\int f \mathrm{d}\mu\right|.$$

If μ and ν are two probability measures, the total variation of $\mu - \nu$ is maximal when μ and ν have disjoint support, in which case $\|\mu-\nu\|_{\mathrm{TV}} = 2$.

Proposition 4. *Grant Assumptions A and B, then there exists some constant c such that:*

$$\|\mathbb{P}_f^n - \mathbb{P}_g^n\|_{TV} \le cn^{\frac12}\|f-g\|_\infty^{\frac12},$$

$$1 - \frac{1}{2}\|\mathbb{P}_f^n - \mathbb{P}_g^n\|_{TV} \ge R\left(cn\|f-g\|_2^2 + c\|f\|^2_{\mathcal{H}^{1/2}} + c\|g\|^2_{\mathcal{H}^{1/2}}\right),$$

where R is some (universal) non increasing >0 function.

Proof. Let $D(\mu,\nu) := \int (\log\frac{\mathrm{d}\mu}{\mathrm{d}\nu})\mathrm{d}\mu \le +\infty$ denote the Kullback-Leibler divergence between two probability measures μ and ν. We recall the classical Pinsker's inequality $\|\mu-\nu\|_{\mathrm{TV}} \le \sqrt{2}D(\mu,\nu)^{1/2}$.

Using Assumption B (ii) and the representation (1) we deduce,

$$\mathbb{E}_f^n\left\{\log\frac{\mathrm{d}P_f^n}{\mathrm{d}P_g^n}(Y_0^n,\ldots,Y_n^n)\right\} = \sum_{i=0}^n \mathbb{E}_f^n\left\{v_{i,n}(\xi_i^n + \Delta_{i,n}) - v_{i,n}(\xi_i^n) - \log\left(\frac{a(f_{i/n})}{a(g_{i/n})}\right)\right\},$$

where $\Delta_{i,n} = \xi_i^n\left(\frac{a(f_{i/n})}{a(g_{i/n})} - 1\right) + \frac{f_{i/n}-g_{i/n}}{a(g_{i/n})}$. By a second order Taylor expansion, this yields the expression for the Kullback–Leibler divergence:

$$D(\mathrm{d}P_f^n, \mathrm{d}P_g^n) = \sum_{i=0}^n \left\{\mathbb{E}_f^n[(\tfrac{\mathrm{d}}{\mathrm{d}x}v_{i,n})(\xi_i^n)\Delta_{i,n}] - \log\left(\tfrac{a(f_{i/n})}{a(g_{i/n})}\right)\right\} + \frac{1}{2}\sum_{i=0}^n \mathbb{E}_f^n[(\tfrac{\mathrm{d}^2}{\mathrm{d}x^2}v_{i,n})(\xi_i^n + \theta_{i,n}\Delta_{i,n})\Delta_{i,n}^2], \tag{20}$$

for some (random) $\theta_{i,n} \in (0,1)$. Using that $x \rightsquigarrow \exp\big(-v_{i,n}(x)\big)$ vanishes at infinity we have $\mathbb{E}_f^n[(\frac{\mathrm{d}}{\mathrm{d}x}v_{i,n})(\xi_i^n)] = 0$ and $\mathbb{E}_f^n[(\frac{\mathrm{d}}{\mathrm{d}x}v_{i,n})(\xi_i^n)\xi_i^n] = 1$ integrating by part. It follows that the terms in the first sum of (20) are equal to $\frac{a(f_{i/n})}{a(g_{i/n})} - 1 - \log\left(\frac{a(f_{i/n})}{a(g_{i/n})}\right)$. The assumptions on $x \rightsquigarrow a(x)$ yields that this quantity is less than some constant times $(f_{i/n} - g_{i/n})^2$.

For the second order terms, using the uniform Lipschitz assumption on $x \rightsquigarrow \frac{\mathrm{d}^2}{\mathrm{d}x^2} v_{i,n}(x)$ together with the uniform bound for $\mathbb{E}_f^n[|\xi_i^n|^3]$ gives,

$$|\mathbb{E}_f^n[(\tfrac{\mathrm{d}^2}{\mathrm{d}x^2} v_{i,n})(\xi_i^n + \theta_{i,n}\Delta_{i,n})\Delta_{i,n}^2]| \leq c|f_{i/n} - g_{i/n}|^3 + |\mathbb{E}_f^n[(\tfrac{\mathrm{d}^2}{\mathrm{d}x^2} v_{i,n})(\xi_i^n)\Delta_{i,n}^2]|.$$

Again we can bound $|\mathbb{E}_f^n[\frac{\mathrm{d}^2}{\mathrm{d}x^2} v_{i,n}(\xi_i^n)\Delta_{i,n}^2]|$ by a constant times $(f_{i/n} - g_{i/n})^2$ using that $\mathbb{E}_f^n[\frac{\mathrm{d}^2}{\mathrm{d}x^2} v_{i,n}(\xi_i^n)]$, $\mathbb{E}_f^n\big[\big(\frac{\mathrm{d}^2}{\mathrm{d}x^2} v_{i,n}(\xi_i^n)\big)\xi_i^n\big]$ and $\mathbb{E}_f^n\big[\big(\frac{\mathrm{d}^2}{\mathrm{d}x^2} v_{i,n}(\xi_i^n)\big)(\xi_i^n)^2\big]$ are controlled by $\sup_{i,n} \mathbb{E}_f^n\big\{\big(\frac{\mathrm{d}}{\mathrm{d}x} v_{i,n}(\xi_i^n)\big)^2(1 + |\xi_i^n|^2)\big\}$. Thus the divergence between the conditional laws is bounded by

$$D(\mathrm{d}P_f^n, \mathrm{d}P_g^n) \leq c \sum_{i=0}^n |f_{i/n} - g_{i/n}|^2,$$

and the first part of the proposition follows from Pinsker's inequality. For the second part of the proposition, we use

$$\sum_{i=0}^n |f_{i/n} - g_{i/n}|^2 \leq 4n \int_0^1 \big(f(x) - g(x)\big)^2 \mathrm{d}x + 8n^{1-2\alpha}(\|f\|_{\mathcal{H}^\alpha}^2 + \|g\|_{\mathcal{H}^\alpha}^2)$$

valid for any $\alpha \in (0,1)$ together with the fact that for two measures μ, ν the total variation $\|\mu - \nu\|_{\mathrm{TV}}$ remains bounded away from 2 when the divergences $D(\mu, \nu)$ and $D(\nu, \mu)$ are bounded away from $+\infty$. $\square$

The next result is the key to the lower bound. Its proof is delayed until Section 7. Let (σ_0, H_0) be a point in the interior of $\mathcal{D}$. Set, for $I > 0$, $\varepsilon_n := I^{-1} n^{-1/(4H_0+2)}$ and

$$H_1 := H_0 + \varepsilon_n, \quad \sigma_1 := \sigma_0\, 2^{j_0 \varepsilon_n},$$

where

$$j_0 = [\log_2(n^{1/(2H_0+1)})].$$

Proposition 5. *For I large enough, there exists a sequence of probability spaces $(\mathcal{X}^n, \mathfrak{X}^n, \mathbf{P}^n)$ on which can be defined two sequences of stochastic processes, $(\xi_t^{i,n})_{t \in [0,1]}$, $i = 0, 1$ such that:*

(i) For $1/2 \leq \alpha < H_0$, the sequences $\|\xi^{0,n}\|_{\mathcal{H}^\alpha}$ and $\|\xi^{1,n}\|_{\mathcal{H}^\alpha}$ are tight under $\mathbf{P}^n$.

(ii) Define $P^{i,n} = \int_{\mathcal{X}^n} \mathbf{P}^n(d\omega) \mathbb{P}^n_{\xi^{i,n}(\omega)}$. Then:

$$\lim_{n \to \infty} \|P^{i,n} - \mathbb{P}^n_{H_i, \sigma_i | \mathcal{Y}}\|_{TV} = 0, \quad i = 0, 1.$$

(iii) There exists a measurable transformation $T^n : \mathcal{X}^n \mapsto \mathcal{X}^n$ such that the sequence $n\|\xi^{1,n}(\omega) - \xi^{0,n}(T^n(\omega))\|_2^2$ is tight under $\mathbf{P}^n$.

(iv) If n is large enough, the probability measure $\mathbf{P}^n$ and its image measure $T^n\mathbf{P}^n$ are equivalent on $(\mathcal{X}^n, \mathfrak{X}^n)$. Moreover, for some $c^\star \in (0,2)$, we have

$$\|\mathbf{P}^n - T^n\mathbf{P}^n\|_{TV} \leq 2 - c^\star < 2$$

provided n is taken large enough.

Remark. The processes $\xi^{0,n}$ and $\xi^{1,n}$ play the role of an approximation for $\sigma_0 W^{H_0}$ and $\sigma_1 W^{H_1}$ respectively. The point (i) means that $\xi^{i,n}$ shares the same smoothness property as W^{H_i}; while (ii) implies that observing a noisy discrete sampling of $\sigma_i W^{H_i}$ $(i = 0, 1)$ or of its approximation is statistically equivalent as $n \to \infty$. Of course these points trivially hold in the case $\xi^{0,n} = \sigma_0 W^{H_0}$ and $\xi^{1,n} = \sigma_1 W^{H_1}$. However, a significant modification of this simple choice is needed in order to have the fundamental properties (iii) and (iv). These properties means that one can transform pathwise, using the application T^n, the process $\xi^{0,n}$ into approximate realizations of $\xi^{1,n}$; meanwhile this application essentially does not transform the measure $\mathbf{P}^n$ into a measure singular with it.

We next prove that Proposition 4 and 5 together imply Theorems 2 and 3.

6.2. Proof of Theorems 2 and 3. We prove Theorem 2 only. The proof of Theorem 3 is analogous after having remarked that the choice of H_i and σ_i entails that $\sigma_1 - \sigma_0$ is of order

$$\frac{\sigma_0 \log(n)}{I(1+2H_0)} n^{-1/(2+4H_0)}.$$

Pick n large enough so that that $(\sigma_1, H_1) \in \mathcal{D}$. Pick an arbitrary estimator $\widehat{H}_n$. Let $M > 0$, with $M < 1/(2I)$ for further purposes, we have:

$$\begin{aligned}
&\sup_{(H,\sigma)\in\mathcal{D}} \mathbb{P}^n_{H,\sigma}\Big\{n^{1/(4H+2)}|\widehat{H}_n - H| \geq M\Big\} \\
\geq\ & \tfrac{1}{2}\mathbb{P}^n_{H_0,\sigma_0}\Big\{n^{\frac{1}{4H_0+2}}|\widehat{H}_n - H_0| \geq M\Big\} + \tfrac{1}{2}\mathbb{P}^n_{H_1,\sigma_1}\Big\{n^{\frac{1}{4H_1+2}}|\widehat{H}_n - H_1| \geq M\Big\} \\
\geq\ & \tfrac{1}{2}P^{0,n}\Big\{n^{\frac{1}{4H_0+2}}|\widehat{H}_n - H_0| \geq M\Big\} + \tfrac{1}{2}P^{1,n}\Big\{n^{\frac{1}{4H_1+2}}|\widehat{H}_n - H_1| \geq M\Big\} + u_n,
\end{aligned}$$

where $u_n \to 0$ as $n \to \infty$ by (ii) of Proposition 5. By definition of $P^{i,n}$ and by taking n large enough, it suffices to bound from below

$$\tfrac{1}{2}\int_{\mathcal{X}^n}\big[\mathbb{P}^n_{\xi^0_n(\omega)}(A^0) + \mathbb{P}^n_{\xi^1_n(\omega)}(A^1)\big]\mathbf{P}^n(\mathrm{d}\omega), \tag{21}$$

where $A^i = \big\{n^{\frac{1}{2+4H_i}}|\hat{H}_n - H_i| \geq M\big\}$. By (iv) of Proposition 5, for n large enough:

$$\begin{aligned}
\int_{\mathcal{X}^n}\mathbb{P}^n_{\xi^{0,n}(\omega)}(A^0)\mathbf{P}^n(\mathrm{d}\omega) &= \int_{\mathcal{X}^n}\mathbb{P}^n_{\xi^{0,n}(\omega)}(A^0)\frac{\mathrm{d}\mathbf{P}^n}{\mathrm{d}T^n\mathbf{P}^n}(\omega)T^n\mathbf{P}^n(\mathrm{d}\omega) \\
= \ \int_{\mathcal{X}^n}\mathbb{P}^n_{\xi^{0,n}(T^n(\omega))}(A^0)\frac{\mathrm{d}\mathbf{P}^n}{\mathrm{d}T^n\mathbf{P}^n}(T^n\omega)\mathbf{P}^n(\mathrm{d}\omega).&
\end{aligned}$$

Thus (21) is equal to half the quantity

$$\int_{\mathcal{X}^n} \left[\mathbb{P}^n_{\xi^{0,n}(T^n\omega)}(A^0)\frac{\mathrm{d}\mathbf{P}^n}{\mathrm{d}T^n\mathbf{P}^n}(T^n\omega) + \mathbb{P}^n_{\xi^{1,n}(\omega)}(A^1)\right]\mathbf{P}^n(\mathrm{d}\omega)$$

$$\geq \quad e^{-\lambda}\int_{\mathcal{X}^n} \left[\mathbb{P}^n_{\xi^{0,n}(T^n\omega)}(A^0) + \mathbb{P}^n_{\xi^{1,n}(\omega)}(A^1)\right]1_{\frac{\mathrm{d}\mathbf{P}^n}{\mathrm{d}T^n\mathbf{P}^n}(T^n\omega)\geq e^{-\lambda}}\mathbf{P}^n(\mathrm{d}\omega)$$

$$\geq \quad e^{-\lambda}\int_{\mathcal{X}^n_r} \left[\mathbb{P}^n_{\xi^{0,n}(T^n\omega)}(A^0) + \mathbb{P}^n_{\xi^{1,n}(\omega)}(A^1)\right]1_{\frac{\mathrm{d}\mathbf{P}^n}{\mathrm{d}T^n\mathbf{P}^n}(T^n\omega)\geq e^{-\lambda}}\mathbf{P}^n(\mathrm{d}\omega).$$

for any $\lambda > 0$, and where $\mathcal{X}^n_r$ denotes the set of $\omega \in \mathcal{X}^n$ such that

$$n\|\xi^{0,n}(T^n\omega) - \xi^{1,n}(\omega)\|_2^2, \quad \big\|\xi^{0,n}(T^n\omega)\big\|_{\mathcal{H}^\alpha} \text{ and } \big\|\xi^{1,n}(\omega)\big\|_{\mathcal{H}^\alpha}$$

are bounded by $r > 0$.

Lemma 2. *For any $r > 0$ there exists $c(r) > 0$ such that, on $\mathcal{X}^n_r$:*

$$\mathbb{P}^n_{\xi^{0,n}(T^n\omega)}(A^0) + \mathbb{P}^n_{\xi^{1,n}(\omega)}(A^1) \geq c(r) > 0.$$

Lemma 3. *For large enough n, we have:*

$$\mathbf{P}^n\Big\{\mathcal{X}^n_r \cap \frac{d\mathbf{P}^n}{dT^n\mathbf{P}^n}(T^n\cdot) \geq e^{-\lambda}\Big\} \geq \mathbf{P}^n(\mathcal{X}^n_r) - e^{-\lambda} - 1 + c^\star/2.$$

Applying successively Lemma 2 and 3, we derive the following lower bound:

$$e^{-\lambda}c(r)[\mathbf{P}^n(\mathcal{X}^n_r) - e^{-\lambda} - 1 + c^\star/2].$$

Thus Theorem 2 is proved as soon as

$$\lim_{r\to\infty}\liminf_{n\to\infty}\mathbf{P}^n(\mathcal{X}^n_r) = 1. \tag{22}$$

It suffices then to take λ and r large enough. By (i) and (iii) of Proposition 5, (22) only amounts to show the tightness of $\|\xi^{0,n}(T^n\omega)\|_{\mathcal{H}^\alpha}$ under $\mathbf{P}^n$. For $L, L' > 0$, we have

$$\mathbf{P}^n\Big\{\|\xi^{0,n}\big(T^n(\omega)\big)\|_{\mathcal{H}^\alpha} \geq L\Big\} = \int_{\mathcal{X}^n} 1_{\{\|\xi^{0,n}(\omega)\|_{\mathcal{H}^\alpha}\geq L\}}\frac{\mathrm{d}T^n\mathbf{P}^n}{\mathrm{d}\mathbf{P}^n}(\omega)\mathbf{P}^n(\mathrm{d}\omega)$$

$$\leq L'\mathbf{P}^n\big\{\|\xi^{0,n}(\omega)\|_{\mathcal{H}^\alpha} \geq L\big\} + \mathbf{P}^n\Big\{\frac{\mathrm{d}T^n\mathbf{P}^n}{\mathrm{d}\mathbf{P}^n} \geq L'\Big\}$$

$$\leq L'\mathbf{P}^n\big\{\|\xi^{0,n}(\omega)\|_{\mathcal{H}^\alpha} \geq L\big\} + (L')^{-1}$$

by Chebyshev's inequality. The tightness of $\|\xi^{0,n}(T^n(\omega))\|_{\mathcal{H}^\alpha}$ then follows from the tightness of $\|\xi^{0,n}\|_{\mathcal{H}^\alpha}$. The proof of Theorem 2 is complete.

6.3. Proof of Lemmas 2 and 3.

6.3.1. *Proof of Lemma 2.* Since $H_0 < H_1$, it suffices to bound from below

$$\mathbb{P}^n_{\xi^{0,n}(T^n\omega)}\{n^{\frac{1}{4H_0+2}}|\hat{H}_n - H_0| \geq M\} + \mathbb{P}^n_{\xi^{1,n}(\omega)}\{n^{\frac{1}{4H_0+2}}|\hat{H}_n - H_1| \geq M\}.$$

Let

$$\mathrm{d}_{\mathrm{test}}(\mu,\nu) := \sup_{0\leq f\leq 1} |\int f\mathrm{d}\mu - \int f\mathrm{d}\nu|$$

denote the test distance between two probability measures μ and ν. The last term above is thus greater than

$$\mathbb{E}^n_{\xi^{1,n}(\omega)}\Big\{1_{n^{1/(4H_0+2)}|\hat{H}_n-H_0|\geq M} + 1_{n^{1/(4H_0+2)}|\hat{H}_n-H_1|\geq M}\Big\}$$

$$-\mathrm{d}_{\mathrm{test}}(\mathbb{P}^n_{\xi^{0,n}(T^n\omega)}, \mathbb{P}^n_{\xi^{1,n}(\omega)}).$$

Now since $M \leq 1/2I$ and by our choice for H_0 and H_1, one of the two events in the expectation above must occur with probability one. Using that $\mathrm{d}_{\mathrm{test}}(\mu,\nu) = \frac{1}{2}\|\mu-\nu\|_{\mathrm{TV}}$, the last term above is further bounded below by

$$1 - \tfrac{1}{2}\|\mathbb{P}^n_{\xi^{0,n}(T^n\omega)} - \mathbb{P}^n_{\xi^{1,n}(\omega)}\|_{\mathrm{TV}}.$$

We conclude by Proposition 4 together with the fact that $\omega \in \mathcal{X}^n_r$.

6.3.2. *Proof of Lemma 3.* It suffices to bound from below

$$\mathbf{P}^n\Big\{\mathcal{X}^n_r\Big\} - \int_{\mathcal{X}^n} 1_{\frac{\mathrm{d}\mathbf{P}^n}{\mathrm{d}T^n\mathbf{P}^n}(T^n\omega)\leq e^{-\lambda}}\mathbf{P}^n(\mathrm{d}\omega)$$

$$= \quad \mathbf{P}^n(\mathcal{X}^n_r) - \int_{\mathcal{X}^n} 1_{\frac{\mathrm{d}T^n\mathbf{P}^n}{\mathrm{d}\mathbf{P}^n}(\omega)\geq e^{\lambda}}T^n\mathbf{P}^n(\mathrm{d}\omega)$$

since $T^n\mathbf{P}^n$ and $\mathbf{P}^n$ are equivalent. We now replace the measure $T^n\mathbf{P}^n$ in the integral above by $\mathbf{P}^n$ with an error controlled by the test distance; the lower bounds becomes

$$\mathbf{P}^n\{\mathcal{X}^n_r\} - \mathbf{P}^n\Big\{\frac{\mathrm{d}T^n\mathbf{P}^n}{\mathrm{d}\mathbf{P}^n} \geq e^{\lambda}\Big\} - \mathrm{d}_{\mathrm{test}}(\mathbf{P}^n, T^n\mathbf{P}^n)$$

$$= \quad \mathbf{P}^n\{\mathcal{X}^n_r\} - \mathbf{P}^n\Big\{\frac{\mathrm{d}T^n\mathbf{P}^n}{\mathrm{d}\mathbf{P}^n} \geq e^{\lambda}\Big\} - \tfrac{1}{2}\|\mathbf{P}^n - T^n\mathbf{P}^n\|_{\mathrm{TV}}.$$

We conclude by using Chebyshev inequality and Proposition 5 (iv).

7. Proof of Proposition 5.

The proof of Proposition 5 relies on the construction of the fractional Brownian motion by Meyer, Sellan and Taqqu [20]. In section 7.1, we recall the main steps of the construction and how to apply it to our framework. In section 7.2, we construct the sequence of spaces $(\mathcal{X}^n, \mathfrak{X}^n, \mathbf{P}^n)$. The proof of (i)–(iv) is delayed until sections 7.3–7.6

7.1. A synthesis of fractional Brownian motion. Consider, a scaling function ϕ whose Fourier transform has compact support as in Meyer's book [19], with the corresponding wavelet function $\psi \in \mathcal{S}(\mathbb{R})$. In [20] the authors introduced, for $d \in \mathbb{R}$, the following differentials of order d (via their Fourier transform):

$$\widehat{D^d\psi}(s) := (is)^d\hat{\psi}(s), \quad \widehat{\phi^{d,\Delta}}(s) := \left(\frac{is}{1-e^{is}}\right)^d \hat{\phi}(s),$$

where a determination of the argument on $\mathbb{C} \setminus \mathbb{R}_-$ with values in $(-\pi, \pi)$ is chosen. It is shown that the above formula is well defined and that $D^d\psi, \phi^{d,\Delta} \in \mathcal{S}(\mathbb{R})$. Define further, for $d = 1/2 - H \in (-1/2, 1/2)$:

$$\psi^H(t) := \int_{-\infty}^t D^d\psi(u)\mathrm{d}u = D^{d-1}\psi(t), \quad \psi^H_{j,k}(t) := 2^{\frac{j}{2}}\psi^H(2^j t - k),$$

$$\Theta^H_k(t) := \int_0^t \phi^{d,\Delta}(u-k)\mathrm{d}u, \quad \Theta^H_{j,k}(t) = 2^{\frac{j}{2}}\Theta^H_k(2^j t).$$

In their Theorem 2 in [20], Meyer *et al.* prove the following almost sure representation of fractional Brownian motion (on an appropriate probability space and uniformly over compact sets of $\mathbb{R}$):

$$W^H_t = \sum_{k=-\infty}^{\infty} \Theta^H_k(t)\epsilon^H_k + \sum_{j=0}^{\infty}\sum_{k=-\infty}^{\infty} 2^{-j(H+1/2)}\{\psi^H_{j,k}(t) - \psi^H_{j,k}(0)\}\epsilon_{j,k},$$

where $\epsilon^H_k = \sum_{l=0}^{\infty}\gamma_l\epsilon'_{k-l}$, and $(1-r)^d = \sum_{k=0}^{\infty}\gamma_k r^k$ near $r = 0$. The $\epsilon'_k, k \in \mathbb{Z}$, $\epsilon_{j,k}, j \geq 0, k \in \mathbb{Z}$ are i.i.d. $\mathcal{N}(0,1)$ random variables. Note that $\gamma_k = O(k^{-1+d})$ so the series above converges in quadratic mean and the time series obtained, $(\epsilon^H_k)_k$, has a spectral density equal to $|2\sin(\frac{v}{2})|^{1-2H_0}$. The scaling

$$W^H_t \overset{\text{law}}{=} 2^{-j_0 H} W^H_{2^{j_0}t}$$

gives yet another representation for W^H_t:

$$\sum_{k=-\infty}^{\infty} 2^{-j_0(H+1/2)}\Theta^H_{j_0,k}(t)\epsilon^H_k + \sum_{j=j_0}^{\infty}\sum_{k=-\infty}^{\infty} 2^{-j(H+1/2)}\{\psi^H_{j,k}(t) - \psi^H_{j,k}(0)\}\epsilon_{j,k}. \tag{23}$$

Comparing with other decompositions of fractional Brownian motion (for instances Ciesielski *et al.*, [5], Benassi *et al.* [3]) a particular feature is that the random variables appearing in the high frequency terms

$$\sum_{j=j_0}^{\infty}\sum_{k=-\infty}^{\infty} 2^{-j(H+1/2)}\{\psi^H_{j,k}(t) - \psi^H_{j,k}(0)\}\epsilon_{j,k}$$

are independent and independent of the low frequency terms.

A drawback is that the basis used depends on H and the functions appearing in the decomposition are not compactly supported. However one can explore the properties of this basis. In their paper [20], Meyer *et al.* shows that the differential of the initial wavelet functions generate a multiresolution of $\mathbf{L}^2(\mathbb{R})$ and state the following results.

Lemma 4 (Lemma 8 in [20])**.** 1) *There exists* $\mathcal{C}^\infty(\mathbb{R})$*,* 2π*-periodic functions* U_d *and* V_d *such that the following formulas hold:*

$$\widehat{\phi^{d,\Delta}}(s) = U_d(s/2)\widehat{\phi^{d,\Delta}}(s/2), \quad \widehat{D^d\psi}(s) = V_d(s/2)\widehat{\phi^{d,\Delta}}(s/2).$$

These 'filters' and U_d *and* V_d *vanishes respectively in a neighborhood of* π *and* 0*.*

2) *Let* $(c_k)_{k\in\mathbb{Z}} \in l^2(\mathbb{Z})$*, then the function* $\sum_k c_k 2\phi^{d,\Delta}(2t-k)$ *can be expressed with the basis* $\phi^{d,\Delta}(t-k)$ *and one level of detail:*

$$\sum_k c_k 2\phi^{d,\Delta}(2t-k) = \sum_k a_k \phi^{d,\Delta}(t-k) + \sum_k b_k D^d\psi(t-k), \tag{24}$$

where $(a_k)_{k\in\mathbb{Z}}$ *and* $(b_k)_{k\in\mathbb{Z}} \in l^2(\mathbb{Z})$*. Moreover* a *and* b *are given as follows: denote* A*,* B *and* C *the* 2π*-periodic extension of the discrete Fourier transforms of* a*,* b *and* c *we have*

$$A(s) = -4^{-d}[V_d(s/2+\pi)C(s/2) - V_d(s/2)C(s/2+\pi)]e^{is/2}, \tag{25}$$

$$B(s) = -4^{-d}[-U_d(s/2+\pi)C(s/2) + U_d(s/2)C(s/2+\pi)]e^{is/2}. \tag{26}$$

From these properties we can show the following lemma that will prove useful to control in $\mathcal{H}^\alpha$ norm the error made when we truncate the expansion. It also explores some properties of the basis when H varies.

Lemma 5. *Let* $H \in (0,1)$*. (i) If* u_k *and* $u_{j,k}$ *are two sequences such that* $|u_k| \le c(1+|k|)^c$ *and* $|u_{j,k}| \le c(1+j)^c(1+|k|)^c$*, then, for any* $\alpha \in [0,1)$ *and* $M \ge 0$*, there exists* $c(\alpha, M)$ *such that, for all* j_0*:*

$$\sum_{j=j_0}^{\infty} \sum_{|k|\ge 2^{j+1}} \left\| u_{j,k}\psi^H_{j,k} \right\|_{\mathcal{H}_\alpha} \le c(\alpha,M)2^{-Mj_0},$$

$$\sum_{|k|\ge 2^{j_0+1}} \left\| u_k \Theta^H_{j_0,k} \right\|_{\mathcal{H}_\alpha} \le c(\alpha,M)2^{-Mj_0}.$$

(ii) For all $M \ge 0$ *there exists* $c(M)$ *such that for all* $\varepsilon > 0$ *with* $H+\varepsilon < 1$ *and* $t \in \mathbb{R}$*:*

$$|\psi^{H+\varepsilon}(t) - \psi^H(t)| \le c(M)\frac{\varepsilon}{(1+|t|)^M}. \tag{27}$$

(iii) For all $\varepsilon > 0$ *with* $H+\varepsilon < 1$ *we have, for all* $k \in \mathbb{Z}$

$$\Theta^{H+\varepsilon}_k - \Theta^H_k = \sum_{l\in\mathbb{Z}} a_l(\varepsilon)\Theta^H_{k+l} + \sum_{l\in\mathbb{Z}} b_l(\varepsilon)\left\{\psi^H_{0,k+l}(t) - \psi^H_{0,k+l}(0)\right\}, \tag{28}$$

where the coefficients $a_l(\varepsilon)$ *and* $b_l(\varepsilon)$ *are such that for all* M*, there exists* $c(M)$ *such that for all* ε

$$\max\{|a_l(\varepsilon)|, |b_l(\varepsilon)|\} \le \varepsilon c(M)(1+|l|)^{-M}. \tag{29}$$

Moreover the 2π*-periodic function* B_ε *with Fourier coefficient* $b_l(\varepsilon)$ *vanishes in some neighborhood of zero independent of* ε*.*

Proof. See Section 8.1 in the Appendix. □

7.2. The space $(\mathcal{X}^n, \mathfrak{X}^n, \mathbf{P}^n)$**.** Let us recall that $H_1 = H_0 + \varepsilon_n$ where $\varepsilon_n = I^{-1}n^{-1/(2+4H_0)}$; $j_0 = [\log_2 n^{1/(1+2H_0)}]$ and $\sigma_1 = \sigma_0 2^{j_0\varepsilon_n}$.

7.2.1. We take for $\mathcal{X}^n$ an infinite product of real lines, endowed with the product sigma field $\mathfrak{X}^n$:

$$\mathcal{X}^n := \Big(\bigotimes_{k=-2^{j_0+1}}^{2^{j_0+1}} \mathbb{R} \Big) \otimes \Big(\bigotimes_{j=j_0}^{\infty} \underset{|k|\le 2^{j+1}}{\otimes} \mathbb{R} \Big) =: \mathcal{X}^n_{\mathbf{e}} \otimes \mathcal{X}^n_{\mathbf{d}}.$$

An element of $\mathcal{X}^n$ is denoted by $\omega = (\omega^{\mathbf{e}}, \omega^{\mathbf{d}})$ with $\omega^{\mathbf{e}} = (\omega^{\mathbf{e}}_k)_{|k|\le 2^{j_0+1}}$ and $\omega^{\mathbf{d}} = (\omega^{\mathbf{d}}_\lambda)_{\lambda=(j,k);j\ge j_0,|k|\le 2^{j+1}}$. The projection on the coordinates are denoted by $\epsilon_k(\omega) = \omega^{\mathbf{e}}_k$ for $|k| \le 2^{j_0+1}$ and $\epsilon_{j,k}(\omega) = \omega^{\mathbf{d}}_{j,k}$ for $j \ge j_0$, $|k| \le 2^{j+1}$.

On $\mathcal{X}^n$ we define the probability measure $\mathbf{P}^n := \mathbf{P}^n_{\mathbf{e}} \otimes \mathbf{P}^n_{\mathbf{d}}$, where $\mathbf{P}^n_{\mathbf{e}}$ is the unique probability on $\mathcal{X}^n_{\mathbf{e}}$ which makes the sequence (ϵ_k) a centred Gaussian stationary time series with spectral density $|2\sin(\frac{s}{2})|^{1-2H_0}$. The probability measure $\mathbf{P}^n_{\mathbf{d}}$ is the unique probability on $\mathcal{X}^n_{\mathbf{d}}$ that makes the sequence $(\epsilon_{j,k})$ i.i.d. $\mathcal{N}(0,1)$.

7.2.2. As suggested by section 7.1, we define an approximation of $\sigma_0 W^H_0$ by keeping a finite number of coefficients at each scale:

$$\begin{aligned}\xi^{0,n}(t) := &\sum_{|k|\le 2^{j_0+1}} \sigma_0 2^{-j_0(H_0+1/2)} \Theta^{H_0}_{j_0,k}(t)\epsilon_k \\ &+ \sum_{j\ge j_0} \sum_{|k|\le 2^{j+1}} \sigma_0 2^{-j(H_0+1/2)} \left\{ \psi^{H_0}_{j,k}(t) - \psi^{H_0}_{j,k}(0) \right\} \epsilon_{j,k}. \quad (30)\end{aligned}$$

Denote by $\mathcal{T}^{n,1}$ a linear mapping from $\mathcal{X}^n_{\mathbf{e}}$ to itself such that under the measure $\mathcal{T}^{n,1}\mathbf{P}^n_{\mathbf{e}}$, the coordinates (ϵ_k) form a centred Gaussian time series with spectral density $|2\sin(\frac{s}{2})|^{1-2H_1}$. Let

$$\epsilon'_k(\omega) := \epsilon_k(\mathcal{T}^{n,1}\omega). \quad (31)$$

We then define on the same space an approximation for $\sigma_1 W^H_1$. A natural choice would be to take again (30) with (σ_1, H_1) and ϵ'_k instead of (σ_0, H_0) and ϵ_k. We proceed a little bit differently: we replace all the $\Theta^{H_1}_{j_0,k}$ by their truncated expansion on $\Theta^{H_0}_{j_0,k+l}$ and $\psi^{H_0}_{j_0,k+l}$ using relation (28). We then reorder the sums and finally drop the terms with index k corresponding to the localization $k/2^j$ outside $[-2,2]$. The reason is that we want to use the same basis as in $\xi^{0,n}$, for the low frequency terms.

This leads us to the following approximation for $\sigma_1 W^{H_1}$:

$$\begin{aligned}\xi^{1,n}(t) := &\sum_{|k|\le 2^{j_0+1}} \sigma_1 2^{-j_0(H_1+1/2)} \Theta^{H_0}_{j_0,k}(t)\epsilon'_k \\ &+ \sum_{|l|\le 2^{j_0+1}} \sigma_1 2^{-j_0(H_1+1/2)} \Theta^{H_0}_{j_0,l}(t) \sum_{|k|\le 2^{j_0+1}} a_{l-k}\epsilon'_k \\ &+ \sum_{|l|\le 2^{j_0+1}} \sigma_1 2^{-j_0(H_1+1/2)} \{\psi^{H_0}_{j_0,l}(t) - \psi^{H_0}_{j_0,l}(0)\} \sum_{|k|\le 2^{j_0+1}} b_{l-k}\epsilon'_k \\ &+ \sum_{j\ge j_0} \sum_{|k|\le 2^{j+1}} \sigma_1 2^{-j(H_1+1/2)} \left\{ \psi^{H_1}_{j,k}(t) - \psi^{H_1}_{j,k}(0) \right\} \epsilon_{j,k}, \quad (32)\end{aligned}$$

where the coefficients $a = a(\varepsilon)$ and $b = b(\varepsilon)$ are defined by (28) with $H = H_0$, $H + \varepsilon = H_1$.

7.2.3. The last step is the construction of the mapping T^n from $(\mathcal{X}^n, \mathfrak{X}^n)$ to itself. Recalling (iii) of Proposition 5, we see that T^n should transform outcomes of $\xi^{0,n}$ into approximate outcomes of $\xi^{1,n}$. Thus we define the action of T^n on the the random space $(\mathcal{X}^n, \mathfrak{X}^n)$ by making the low frequency terms of $\xi^{0,n}(T^n\omega)$ exactly match the low frequency terms of $\xi^{1,n}(\omega)$.

We define $\mathcal{T}^{2,n}$ on $\mathcal{X}^n$ as the linear map such that:

$$\epsilon_l(\mathcal{T}^{2,n}\omega) = \sum_{|k|\leq 2^{j_0+1}} a_{l-k}\epsilon_k(\omega) + \epsilon_l(\omega), \tag{33}$$

$$\epsilon_{j_0,l}(\mathcal{T}^{2,n}\omega) = \sum_{|k|\leq 2^{j_0+1}} b_{l-k}\epsilon_k(\omega) + \epsilon_{j_0,l}(\omega), \tag{34}$$

$$\epsilon_{j,l}(\mathcal{T}^{2,n}\omega) = \epsilon_{j,l}(\omega) \text{ if } j > j_0. \tag{35}$$

Remark that the matrix of this linear map in the canonical basis of $\mathcal{X}^n$ is of course infinite, but $\mathcal{T}^{2,n}$ leaves invariant the finite dimensional subspace $\mathcal{X}^n_{\mathbf{e}} \otimes (\otimes_{|k|\leq 2^{j_0+1}} \mathbb{R}) \otimes (0, 0, \dots) \subset \mathcal{X}^n$ and is the identity on a supplementary space. On the finite dimensional subspace its matrix is $\mathrm{Id} + K^n$ where K^n is the square matrix of size $2[2^{j_0+2} + 1]$:

$$K^n = \begin{pmatrix} (a_{l-k})_{|l|,|k|\leq 2^{j_0+1}} & 0 \\ (b_{l-k})_{|l|,|k|\leq 2^{j_0+1}} & 0 \end{pmatrix}. \tag{36}$$

Finally, we set

$$T^n = \mathcal{T}^{n,2} \circ \mathcal{T}^{n,1}, \tag{37}$$

where we denote again by $\mathcal{T}^{n,1}$ the extension of $\mathcal{T}^{n,1}$ (previously defined only on $\mathcal{X}^n_{\mathbf{e}}$) to $\mathcal{X}^n$ such that it is the identity on $0_{\mathcal{X}^n_{\mathbf{e}}} \otimes \mathcal{X}^n_{\mathbf{d}}$.

As announced the choice of T^n, with (30)–(35) and the fact that $\sigma_1 2^{-j_0 H_1} = \sigma_0 2^{j_0\varepsilon_n} 2^{-j_0 H_1} = \sigma_0 2^{-j_0 H_0}$ yields

$$\begin{aligned}\xi^{1,n}(\omega) - \xi^{0,n}(T^n(\omega)) &= \sum_{j\geq j_0} \sum_{|k|\leq 2^{j+1}} \sigma_1 2^{-j(H_1+1/2)} \left\{\psi^{H_1}_{j,k}(t) - \psi^{H_1}_{j,k}(0)\right\} \epsilon_{j,k}(\omega) \\ &- \sum_{j\geq j_0} \sum_{|k|\leq 2^{j+1}} \sigma_0 2^{-j(H_0+1/2)} \left\{\psi^{H_0}_{j,k}(t) - \psi^{H_0}_{j,k}(0)\right\} \epsilon_{j,k}(\omega).\end{aligned} \tag{38}$$

We now have completed the setup of $(\mathcal{X}^n, \mathfrak{X}^n, \mathbf{P}^n)$ and it now remains to prove that Proposition 5 hold. Let us stress that the choice of j_0 is for that matter crucial. Clearly Proposition 5 (iii) requires that j_0 is large enough. Meanwhile, Proposition 5 (iv) requires that the number of components of $\mathcal{X}^n$ on which T^n is different from the identity is as small as possible, which requires that j_0 is not too large.

7.3. Proof of Proposition 5, (i). Define temporarily a probability space $(\tilde{\Omega}^0, \tilde{\mathcal{A}}^0, \tilde{\mathbb{P}}^0)$ with random variables $(\tilde{\epsilon}^0_{j,k})_{j\geq j_0, k\in\mathbb{Z}}$ and $(\tilde{\epsilon}^0_k)_{k\in\mathbb{Z}}$ as in the representation of section 7.1. Let $\tilde{W}^{H_0}$ be the corresponding fractional Brownian motion defined on this space. It is well known that $\|\tilde{W}^{H_0}\|_{\mathcal{H}^\alpha}$ is almost surely finite on this probability space. We write the decomposition

$$\sigma_0 \tilde{W}^{H_0} = \tilde{\xi}^{0,n} + \tilde{r}^{0,n},$$

where $\tilde{\xi}^{0,n}$ has an expression analogous to (30) and

$$\tilde{r}^{0,n} = \sum_{|k|>2^{j_0+1}} \sigma_0 2^{-j_0(H_0+1/2)} \Theta^{H_0}_{j_0,k}(t)\tilde{\epsilon}^0_k + \sum_{j\geq j_0} \sum_{|k|>2^{j+1}} \sigma_0 2^{-j(H_0+1/2)} \left\{ \psi^{H_0}_{j,k}(t) - \psi^{H_0}_{j,k}(0) \right\} \tilde{\epsilon}^0_{j,k}.$$

The tightness of $\|\xi^{0,n}\|_{\mathcal{H}^\alpha}$ under $\mathbf{P}^n$ is equivalent to that of $\|\tilde{\xi}^{0,n}\|_{\mathcal{H}^\alpha}$ under $\tilde{\mathbb{P}}^0$ and thus will follow from the study of $\|\tilde{r}^{0,n}\|_{\mathcal{H}^\alpha}$. It is sufficient to show that its expectation is bounded independently of j_0. Using the Gaussianity of the random variables (see lemma 11 below) there exists a positive $C(\omega)$ such that $\mathbb{E}\{C^p\}$ depends only on p and:

$$\forall j \geq j_0, \forall k \in \mathbb{Z},\ |\tilde{\epsilon}^0_{j,k}| \leq C(\omega)\log(2+j)\log(2+|k|) \text{ and } |\tilde{\epsilon}^0_k| \leq C(\omega)\log(2+|k|).$$

By Lemma 5 (i) we obtain that for arbitrarily large M

$$\left\|\tilde{r}^{0,n}\right\|_{\mathcal{H}_\alpha} \leq C(\omega)c(M)2^{-j_0 M}, \tag{39}$$

hence $\mathbb{E}\{\|\tilde{r}^{0,n}\|_{\mathcal{H}^\alpha}\} \leq c$ and the result is proved. We now turn to $\|\xi^{1,n}\|_{\mathcal{H}^\alpha}$. Define likewise on a probability space $(\tilde{\Omega}^1, \tilde{\mathcal{A}}^1, \tilde{\mathbb{P}}^1)$, the process $\tilde{W}^{H_1}_t$ as

$$\sum_{k=-\infty}^{\infty} 2^{-j_0(H_1+1/2)} \Theta^{H_1}_{j_0,k}(t)\tilde{\epsilon}^1_k + \sum_{j=j_0}^{\infty} \sum_{k=-\infty}^{\infty} 2^{-j(H_1+1/2)} \left\{ \psi^{H_1}_{j,k}(t) - \psi^{H_1}_{j,k}(0) \right\} \tilde{\epsilon}^1_{j,k}.$$

We use Lemma 5 (iii) for $|k| \leq 2^{j_0+1}$:

$$\begin{aligned}
\tilde{W}^{H_1}_t = &\sum_{|k|\leq 2^{j_0+1}} \tilde{\epsilon}^1_k 2^{-j_0(H_1+1/2)} \left(\sum_{l\in\mathbb{Z}} a_{l-k}\Theta^{H_0}_{j_0,l}(t) + \Theta^{H_0}_{j_0,k}(t) \right) \\
&+ \sum_{|k|\leq 2^{j_0+1}} \tilde{\epsilon}^1_k 2^{-j_0(H_1+1/2)} \sum_{l\in\mathbb{Z}} b_{l-k} \left\{ \psi^{H_0}_{j_0,l}(t) - \psi^{H_0}_{j_0,l}(0) \right\} \\
&+ \sum_{|k|>2^{j_0+1}} 2^{-j_0(H_1+1/2)} \Theta^{H_1}_{j_0,k}(t)\tilde{\epsilon}^1_k \\
&+ \sum_{j=j_0}^{\infty} \sum_{k=-\infty}^{\infty} 2^{-j(H_1+1/2)} \left\{ \psi^{H_1}_{j,k}(t) - \psi^{H_1}_{j,k}(0) \right\} \tilde{\epsilon}^1_{j,k}.
\end{aligned}$$

We derive the following decomposition

$$\sigma_1 \tilde{W}^{H_1} = \tilde{\xi}^{1,n} + \tilde{r}^{1,n}$$

where $\tilde{\xi}^{1,n}$ has under $\tilde{\mathbb{P}}^1$, the same law as $\xi^{1,n}$ under $\mathbf{P}^n$ and

$$\begin{aligned}\tilde{r}^{1,n}(t) := \sum_{|l|>2^{j_0+1}} & \sigma_1 2^{-j_0(H_1+1/2)}\Theta^{H_0}_{j_0,l}(t)\sum_{|k|\le 2^{j_0+1}} a_{l-k}\tilde{\epsilon}^1_k \\ &+ \sum_{|l|>2^{j_0+1}} \sigma_1 2^{-j_0(H_1+1/2)}\{\psi^{H_0}_{j_0,l}(t)-\psi^{H_0}_{j_0,l}(0)\}\sum_{|k|\le 2^{j_0+1}} b_{l-k}\tilde{\epsilon}^1_k \\ &\qquad+\sum_{|k|>2^{j_0+1}} \sigma_1 2^{-j_0(H_1+1/2)}\Theta^{H_1}_{j_0,k}(t)\tilde{\epsilon}^1_k \\ &\qquad+\sum_{j\ge j_0}\sum_{|k|>2^{j+1}} \sigma_1 2^{-j(H_1+1/2)}\left\{\psi^{H_1}_{j,k}(t)-\psi^{H_1}_{j,k}(0)\right\}\tilde{\epsilon}^1_{j,k}.\end{aligned}$$

Then using again Lemma 5 (i) and Lemma 11 in the Appendix we obtain that $\|\tilde{r}^{1,n}\|_{\mathcal{H}_\alpha} \le C(\omega)c(M)2^{-j_0M}$ for M arbitrarily large where C is a random variable with finite moment independent of j_0. The result follows.

7.4. Proof of Proposition 5 (ii). Let $A \in \mathcal{Y}$ be an event in the observation space. We have

$$P^n_{H_0,\sigma_0}(A) = \int_{\tilde{\Omega}^0} \mathbb{P}^n_{\sigma_0\tilde{W}^{H_0}(\omega)}(A)\,\tilde{\mathbb{P}}^0(\mathrm{d}\omega),\ P^{n,0}(A) = \int_{\tilde{\Omega}^0} \mathbb{P}^n_{\tilde{\xi}^{0,n}(\omega)}(A)\,\tilde{\mathbb{P}}^0(\mathrm{d}\omega).$$

It follows that

$$\|P^n_{H_0,\sigma_0|\mathcal{Y}} - P^{n,0}\|_{\mathrm{TV}} \le \int_{\tilde{\Omega}^0} \|\mathbb{P}^n_{\sigma_0\tilde{W}^{H_0}(\omega)} - \mathbb{P}^n_{\tilde{\xi}^{0,n}(\omega)}\|_{\mathrm{TV}}\tilde{\mathbb{P}}^0(\mathrm{d}\omega).$$

By Proposition 4 this is less than the expectation with respect to $\tilde{\mathbb{P}}^0$ of $c(n\|\sigma_0\tilde{W}^{H_0} - \tilde{\xi}^{0,n}\|^2_\infty)^{1/2} = c(n\|\tilde{r}^{0,n}\|^2_\infty)^{1/2}$. But, by (39), $n\|r^{0,n}\|^2_\infty \le n\|r^{0,n}\|^2_{\mathcal{H}_0}$ is bounded by some random variable with finite moment of any order under $\tilde{\mathbb{P}}^0$ times $n2^{-j_0M}$ for arbitrarily large M. Since 2^{-j_0} tends to zero as a negative power of n, if M is sufficiently large we obtain:

$$\lim_{n\to\infty} \|P^n_{H_0,\sigma_0|\mathcal{Y}} - P^{n,0}\|_{\mathrm{TV}} = 0.$$

One obtain analogously that $\lim_{n\to\infty}\|P^n_{H_1,\sigma_1|\mathcal{Y}} - P^{n,1}\|_{\mathrm{TV}} = 0$.

7.5. Proof of Proposition 5 (iii). From the choice of ε_n and j_0, all we need to prove is the tighness of $\varepsilon_n^{-2}2^{2j_0H_0}\|\xi^{1,n}(\omega) - \xi^{0,n}(T^n(\omega))\|^2_2$. We plan to use the following decomposition

$$\xi^{1,n}(\omega) - \xi^{0,n}\big(T^n(\omega)\big) = q_1(t) - q_1(0) + q_2(t) - q_2(0),$$

where

$$\begin{aligned} q_1(t) &:= \sum_{j\ge j_0}\sum_{|k|\le 2^{j+1}} \sigma_1 2^{-j(H_1+1/2)}\{\psi^{H_1}_{j,k}(t) - \psi^{H_0}_{j,k}(t)\}\epsilon_{j,k}(\omega), \\ q_2(t) &:= \sum_{j\ge j_0}\sum_{|k|\le 2^{j+1}} (\sigma_1 2^{-j(H_1+1/2)} - \sigma_0 2^{-j(H_0+1/2)})\psi^{H_0}_{j,k}(t)\epsilon_{j,k}(\omega).\end{aligned}$$

Using the independence of the $\epsilon_{j,k}$

$$\mathbb{E}_{\mathbf{P}^n}\{q_1(t)^2\} = \sum_{j\ge j_0}\sum_{|k|\le 2^{j+1}} \sigma_1^2 2^{-j(2H_1+1)}(\psi^{H_1}_{j,k}(t) - \psi^{H_0}_{j,k}(t))^2.$$

Then, by Lemma 5 (ii) for any $M \geq 0$, this is less than

$$c(M) \sum_{j \geq j_0} \sum_{|k| \leq 2^{j+1}} \sigma_1^2 2^{-j(2H_1+1)} \frac{2^j \varepsilon_n^2}{(1+|2^j t - k|)^{2M}}.$$

Since $H_1 \geq H_0$, by taking $M > 1/2$ we readily obtain $\mathbb{E}_{\mathbf{P}^n}\{q_1(t)^2\} \leq c\varepsilon_n^2 2^{-2j_0 H_0}$, hence

$$\mathbb{E}_{\mathbf{P}^n}\left(\|q_1(\cdot) - q_1(0)\|_2^2\right) \leq c\varepsilon_n^2 2^{-2j_0 H_0}. \tag{40}$$

Since $\sigma_1 2^{-j_0 \varepsilon_n} = \sigma_0$ we have

$$q_2(t) = \sum_{j \geq j_0} \sum_{|k| \leq 2^{j+1}} (2^{(j_0 - j)\varepsilon_n} - 1)\sigma_0 2^{-j(H_0+1/2)} \psi_{j,k}^{H_0}(t) \epsilon_{j,k}(\omega).$$

By the independence of the $\epsilon_{j,k}$ and the fact that $\psi^{H_0} \in \mathcal{S}(\mathbb{R})$, we derive, for any $M \geq 0$:

$$\mathbb{E}_{\mathbf{P}_n}\{q_2(t)^2\} \leq c(M) \sum_{j \geq j_0} \sum_{|k| \leq 2^{j+1}} |2^{(j_0-j)\varepsilon_n} - 1|^2 \sigma_0^2 2^{-j(2H_0+1)} \frac{2^j}{(1+|2^j t - k|)^{2M}}.$$

For $M > 1/2$, this quantity is smaller than $c\varepsilon_n^2 \sum_{j \geq j_0} (j - j_0)^2 2^{-2jH_0}$, hence

$$\mathbb{E}_{\mathbf{P}_n}\{q_2(t)^2\} \leq \varepsilon_n^2 2^{-2j_0 H_0},$$

from which we deduce a bound analogous to (40) for q_2.

7.6. Proof of Proposition 5 (iv).

7.6.1. Let us first briefly explain why

$$\mathcal{T}^n \mathbf{P}^n = \mathcal{T}^{n,2} \circ \mathcal{T}^{n,1} \mathbf{P}^n \sim \mathbf{P}^n.$$

Recall that the measure $\mathbf{P}^n = \mathbf{P}_{\mathbf{e}}^n \otimes \mathbf{P}_{\mathbf{d}}^n$ on $\mathcal{X}_{\mathbf{e}}^n \otimes \mathcal{X}_{\mathbf{d}}^n$ is such that $\mathbf{P}_{\mathbf{d}}^n$ is an infinite dimensional white noise and $\mathbf{P}_{\mathbf{e}}^n$ makes the components of $\mathcal{X}_{\mathbf{e}}^n$ a Gaussian time series with spectral density $|2\sin(\frac{s}{2})|^{1-2H_0}$. But the almost sure positivity of this spectral density implies that $\mathbf{P}_{\mathbf{e}}^n$ is equivalent to the Lebesgue measure $\mathrm{d}x$ on $\mathcal{X}_{\mathbf{e}}^n$ (see e.g. Brockwell and Davis p.137 [4]). Thus $\mathbf{P}^n \sim \mathrm{d}x \otimes \mathbf{P}_{\mathbf{d}}^n$. But the measure $\mathcal{T}^{n,1}\mathbf{P}^n$ has the same structure as $\mathbf{P}^n$ except that the components of $\mathcal{X}_{\mathbf{e}}^n$ now have the law of a time series with spectral density $|2\sin(\frac{s}{2})|^{1-2H_1} > 0$. Hence $\mathcal{T}^{n,1}\mathbf{P}^n \sim \mathrm{d}x \otimes \mathbf{P}_{\mathbf{d}}^n$ too. It follows that $\mathbf{P}^n \sim \mathcal{T}^{n,1}\mathbf{P}^n$.

Next, recall that $\mathcal{T}^{n,2}$ is a linear map on $\mathcal{X}^n = \mathcal{X}_{\mathbf{e}}^n \otimes (\bigotimes_{|k| \leq 2^{j_0+1}} \mathbb{R}) \otimes (\bigotimes_{j > j_0}^{\infty} \bigotimes_{|k| \leq 2^{j+1}} \mathbb{R})$ which has matrix $\mathrm{Id} + K^n$ (recall (36)) on the restriction $\mathcal{X}_{\mathbf{e}}^n \otimes (\bigotimes_{|k| \leq 2^{j_0+1}} \mathbb{R})$ and the identity on $\bigotimes_{j > j_0}^{\infty} \bigotimes_{|k| \leq 2^{j+1}} \mathbb{R}$. We deduce that a sufficient condition for the equivalence

$$\mathcal{T}^{n,2} \circ \mathcal{T}^{n,1}\mathbf{P}^n \sim \mathcal{T}^{n,1}\mathbf{P}^n$$

is that the matrix $\mathrm{Id} + K^n$ is non degenerate. But the summation of the coefficients along the lines and columns of K^n are bounded by $c\varepsilon_n$ and by Schur Lemma this is sufficient to imply that all the eigenvalues of K^n are at most of magnitude ε_n. Hence $\mathrm{Id} + K^n$ is invertible for large enough n and the equivalence is proved.

7.6.2. Using the triangle inequality Proposition 5 (iv) is proved as soon as

$$\|\mathbf{P}^n - \mathcal{T}^{n,1}\mathbf{P}^n\|_{\text{TV}} \tag{41}$$

and

$$\|\mathcal{T}^{n,2} \circ \mathcal{T}^{n,1}\mathbf{P}^n - \mathcal{T}^{n,1}\mathbf{P}^n\|_{\text{TV}} \tag{42}$$

can be made arbitrarily small for an appropriate choice of I and for large enough n. It will be convenient to use again the inequality $\|\mu - \nu\|_{\text{TV}} \leq \sqrt{2}D(\mu,\nu)^{1/2}$, where $D(\mu,\nu)$ denotes the Kullback-Leibler divergence between μ and ν.

Let us recall the following fundamental fact on Gaussian measures: If μ_A and μ_B are two (centred) Gaussian measures on $\mathbb{R}^N$ with non degenerate covariance matrices A and B, then

$$D(\mu_A,\mu_B) = \text{Tr}(AB^{-1}) - \log\det(AB^{-1}) - \text{Tr}(\text{Id}).$$

On the vector space of matrices of size N, we will make use of the trace norm $\|A\|_{\text{tr}}^2 := \text{Tr}(A^*A)$ and the operator norm $\|A\|_{\text{op}}^2 = \sup_{\|x\|=1}\|Ax\|$, where $\|\cdot\|$ is the usual Euclidean norm on $\mathbb{R}^N$. Recall that

$$\|A\|_{\text{op}} \leq \|A\|_{\text{tr}} \leq \sqrt{N}\|A\|_{\text{op}}$$

and that

$$|\text{Tr}(AB)| \leq \|A\|_{\text{tr}}\|B\|_{\text{tr}}, \;\; \|AB\|_{\text{tr}} \leq \min\{\|A\|_{\text{op}}\|B\|_{\text{tr}}, \|A\|_{\text{tr}}\|B\|_{\text{op}}\}.$$

We further take $N := 2^{j_0+2}+1$, the dimensionality of $\mathcal{X}_{\mathbf{e}}^n$ and define

$$f_0(s) := |2\sin(\frac{s}{2})|^{1-2H_0} \quad \text{and} \quad f_1(s) := |2\sin(\frac{s}{2})|^{1-2H_1} \tag{43}$$

the two spectral densities involved in the definition of $\mathbf{P}^n$ and $\mathcal{T}^{n,1}\mathbf{P}^n$ respectively. Their restrictions to $\mathcal{X}_{\mathbf{e}}^n$ are Gaussian measures with covariance matrices given by the Toeplitz matrices $T_N(f_0)$ and $T_N(f_1)$ (the notation $T_N(f)$ means the matrix with entries $T_N(f)_{k,l} := \frac{1}{2\pi}\int_{-\pi}^{\pi} f(s)e^{i(k-l)s}\text{d}s$). Thus, we derive the following upper bound for the variation distance $\|\mathbf{P}^n - \mathcal{T}^{n,1}\mathbf{P}^n\|_{\text{TV}}$:

$$\sqrt{2}\left[\text{Tr}(T_N(f_1)T_N(f_0)^{-1}) - \log\det(T_N(f_1)T_N(f_0)^{-1}) - \text{Tr}(\text{Id})\right]. \tag{44}$$

But the inequality

$$f_1(s) = f_0(s)|2\sin(\frac{s}{2})|^{-2\varepsilon_n} \geq f_0(s)$$

implies, by Lemma 6 in the Appendix, that $T_N(f_1) \geq T_N(f_0)$. In particular, the eigenvalues $\lambda_1,\ldots,\lambda_N$ of the matrix $T_N(f_1)T_N(f_0)^{-1}$ are all greater than one, hence (44) becomes

$$\sqrt{2}\sum_{i=1}^N(\lambda_i - \log\lambda_i - 1) \leq \sum_{i=1}^N(\lambda_i-1)^2 = \text{Tr}\left([T_N(f_1)T_N(f_0)^{-1} - \text{Id}]^2]\right).$$

If we introduce the difference

$$g_{\varepsilon_n}(s) = f_1(s) - f_0(s) = \left[|2\sin(\frac{s}{2})|^{-2\varepsilon_n} - 1\right] f_0(s), \tag{45}$$

we obtain

$$\|\mathbf{P}^n - \mathcal{T}^{n,1}\mathbf{P}^n\|_{\text{TV}} \leq \text{Tr}(\left[T_N(g_{\epsilon_n})T_N(f_0)^{-1}\right]^2).$$

If we apply lemma 8 in Appendix below, we find that this quantity is less than $cN\varepsilon_n^2 \sim cI^{-2}$. This yields an arbitrary control on (41) if I is taken large enough.

In order to obtain a control on (42), we need to compare two Gaussian measures μ_C and μ_D on $\mathbb{R}^{N'}$ with now $N' = 2(2^{j_0+2}+1)$, $C = (\mathrm{Id} + K^n)D(\mathrm{Id}+K^n)^*$ and

$$D = \left(\begin{array}{c|c} T_N(f_1) & 0 \\ \hline 0 & \mathrm{Id} \end{array}\right),$$

as follows from (36). In the same way as for (41), we have

$$\|\mathcal{T}^{n,2}\circ\mathcal{T}^{n,1}\mathbf{p}^n - \mathcal{T}^{n,1}\mathbf{p}^n\|_{\mathrm{TV}} \le \sqrt{2}[\mathrm{Tr}(CD^{-1}) - \log\det(CD^{-1}) - \mathrm{Tr}(\mathrm{Id})].$$

Write

$$CD^{-1} = \mathrm{Id} + K^n + DK^{n*}D^{-1} + K^nDK^{n*}D^{-1}$$

hence $\|\mathcal{T}^{n,2}\circ\mathcal{T}^{n,1}\mathbf{P}^n - \mathcal{T}^{n,1}\mathbf{P}^n\|_{\mathrm{TV}}$ is bounded by

$$\sqrt{2}[2\,\mathrm{Tr}(K^n) - 2\log\det(\mathrm{Id}+K^n) + \mathrm{Tr}(K^nDK^{n*}D^{-1})].$$

Since all the eigenvalues of K^n are of magnitude ε_n the upper bound becomes

$$c[N\varepsilon_n^2 + \mathrm{Tr}(K^nDK^{n*}D^{-1})].$$

Since K^n is given by (36), it can be rewritten with the help of Toeplitz matrices as

$$K^n = \left(\begin{array}{c|c} T_N(A_\varepsilon) & 0 \\ \hline T_N(B_\varepsilon) & 0 \end{array}\right),$$

where A_ε and B_ε are the 2π–periodic function associated with the Fourier sequences $a_l(\varepsilon)$ and $b_l(\varepsilon)$ already introduced in the proof of Lemma 5. The product of these blockwise matrices shows that

$$\mathrm{Tr}(K^nDK^{n*}D^{-1}) = \mathrm{Tr}(T_N(A_\varepsilon)T_N(f_1)T_N(A_\varepsilon)^*T_N(f_1)^{-1}) + \mathrm{Tr}(T_N(B_\varepsilon)T_N(f_1)T_N(B_\varepsilon)^*).$$

These two traces of Toeplitz matrices are shown in Lemma 9–10 in Appendix below to be of order $cN\varepsilon_n^2 \sim cI^{-1}$. We finally get an arbitrary control on (42) for a large enough I.

8. Appendix

8.1. Proof of lemma 5. (i) Let $L \ge 2$ and k, $|k| \in [L2^j, (L+1)2^j]$. Since $\psi^H \in \mathcal{S}(\mathbb{R})$,

$$\sup_{t\in[0,1]} |\psi_{j,k}^H(t)| + |\tfrac{\partial}{\partial t}\psi_{j,k}^H(t)| \le c(M)2^{3/2j}(1+(L-1)2^j)^{-M}, \tag{46}$$

for an arbitrarily large M. The left hand side of the first inequality to be established is then less than

$$c(M)\sum_{j=j_0}^{\infty}\sum_{L=2}^{\infty} 2^{3/2j} \sum_{|k|\in[L2^j,(L+1)2^j]} |u_{j,k}(1+(L-1)2^j)^{-M}|.$$

Using the assumption on $u_{j,k}$, we bound this quantity by $c(M)\sum_{j=j_0}^{\infty}(1+j^c)\sum_{L=2}^{\infty} 2^{5/2j}(1+(L-1)2^j)^{c-M}$ for another constant $c(M)$. If $M > c+1$, the latter expression is smaller than $c(M)\sum_{j=j_0}^{\infty}(1+j^c)2^{5/2j}1/(M-c-1)(1+$

$2^j)^{c-M+1}$, which has the desired form. The second inequality is proved similarly, noticing that

$$\Theta^H_{j,k}(t) = 2^{3/2j} \int_0^t \phi^{d,\Delta}(2^j u - k)\mathrm{d}u,$$

where $\phi^{d,\Delta}$ is an element of $\mathcal{S}(\mathbb{R})$ and the forgoing arguments apply with $\Theta^H_{j,k}$ in place of $\psi^H_{j,k}$.

(ii) By definition of ψ^H, we have

$$\widehat{\psi^{H+\varepsilon}}(s) - \widehat{\psi^H}(s) = \left[(is)^{-\varepsilon} - 1\right] \widehat{\psi^H}(s) := g_\varepsilon(s).$$

For the Meyer wavelet, $\hat{\psi}$ vanishes in a neighbourhood of 0 and has compact support, so does $\widehat{\psi^H}(s)$ and the function $g_\varepsilon(s)$ is smooth with compact support. The classical relation between the norm of the derivatives of the Fourier transform of a function and the size of the function gives (27).

(iii) We have

$$\widehat{\phi^{d-\varepsilon,\Delta}}(s) - \widehat{\phi^{d,\Delta}}(s) = [(\frac{is}{1-e^{is}})^{-\varepsilon} - 1]\widehat{\phi^{d,\Delta}}(s).$$

In the construction of Meyer *et al.* $\widehat{\phi^{d,\Delta}}(s)$ vanishes if $|s| \geq 4\pi/3$, so we can find a smooth function $h_\varepsilon(s)$, with period 4π and equal to $\left(\frac{is}{1-e^{is}}\right)^{-\varepsilon} - 1$ if $|s| \leq 4\pi/3$ that vanishes in a neighborhood of 2π and such that

$$\widehat{\phi^{d-\varepsilon,\Delta}}(s) - \widehat{\phi^{d,\Delta}}(s) = h_\varepsilon(s)\widehat{\phi^{d,\Delta}}(s).$$

Using Lemma 4, $\widehat{\phi^{d-\varepsilon,\Delta}}(s) - \widehat{\phi^{d,\Delta}}(s) = C_\varepsilon(s/2)\widehat{\phi^{d,\Delta}}(s/2)$, where $C_\varepsilon(s) = h_\varepsilon(2s)U_d(s)$ is 2π-periodic. In the real variable domain:

$$\phi^{d-\varepsilon,\Delta}(t) - \phi^{d,\Delta}(t) = \sum_k c_k(\varepsilon) 2\phi^{d,\Delta}(2t-k),$$

where $c_k(\varepsilon)$ are the discrete Fourier coefficients of the function C_ε. By yet Lemma 4 again:

$$\phi^{d-\varepsilon,\Delta}(t) - \phi^{d,\Delta}(t) = \sum_l a_l(\varepsilon)\phi^{d,\Delta}(t-l) + \sum_l b_l(\varepsilon) D^d\psi(t-l), \qquad (47)$$

where the $a_l(\varepsilon)$, $b_l(\varepsilon)$ and $c_l(\varepsilon)$ are related by (25)-(26). Making the change of variable t by $t-k$ in (47) and then integrating between 0 and t gives (28). The next step is (29). It follows from the explicit expression for A_ε and B_ε and C_ε together with the fact that h_ε is easily bounded by a constant time ε. Finally the fact that B_ε vanishes on a neighborhood of 0 follows from (26) and the fact that U_d vanishes near π.

8.2. Technical lemmas. We start by some results on Toeplitz matrices of size N, which, given a function f, have entries defined by $T_N(f)_{k,l} = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(s)e^{i(k-l)s}\mathrm{d}s$.

Lemma 6. *Let f and g be two spectral densities with $f(s) \geq g(s) \geq 0$, then $T_N(f) \geq T_N(g)$. In particular, if $f(s) \geq c > 0$, then $T_N(f) \geq cId$.*

Proof. If $x \in \mathbb{R}^N$

$$\begin{aligned} x^* T_N(g)x &= \tfrac{1}{2\pi}\textstyle\int_{-\pi}^{\pi} g(s)|\sum_{k=1}^N e^{iks}x_k|^2 \mathrm{d}s \\ &\leq \tfrac{1}{2\pi}\textstyle\int_{-\pi}^{\pi} f(s)|\sum_{k=1}^N e^{iks}x_k|^2 \mathrm{d}s = x^* T_N(f)x. \end{aligned}$$

□

Lemma 7. *(i) Let f and h be 2π-periodic and positive. Assume that h is even and decreasing on $[0,\pi]$ then*

$$\|T_N(fh)^{1/2}T_N(f)^{-1/2}\|_{op}^2 \leq \int_{-\frac{1}{2N}}^{\frac{1}{2N}} f(s)h(s)ds\Big(\int_{-\frac{1}{2N}}^{\frac{1}{2N}} f(s)ds\Big)^{-1}. \qquad (48)$$

(ii) With the notation of section 7.6, we have

$$\|T_N(g_{\varepsilon_n})^{1/2}T_N(f_0)^{-1/2}\|_{op}^2 \leq c\varepsilon_n \log(N).$$

Proof. For the part (i) we follow closely the idea of Lemma 5.3 in Dahlhaus [6]:

$$\begin{aligned} \|T_N(fh)^{1/2}T_N(f)^{-1/2}\|_{\mathrm{op}}^2 &= \sup_{\|x\|=1} \frac{x^*T_N(fh)x}{x^*T_N(f)x} \\ &= \sup_{\|x\|=1} \frac{\int_{-\pi}^{\pi} f(s)h(s)|\sum_{k=1}^N e^{iks}x_k|^2 \mathrm{d}s}{\int_{-\pi}^{\pi} f(s)|\sum_{k=1}^N e^{iks}x_k|^2 \mathrm{d}s} \\ &\leq \sup_{\xi} \frac{\int_{-\pi}^{\pi} f(s)h(s)\xi(s)\mathrm{d}s}{\int_{-\pi}^{\pi} f(s)\xi(s)\mathrm{d}s} \end{aligned}$$

where the supremum in the last line above is taken for all functions $\xi \geq 0$, bounded by N, with $\int \xi = 1$. Using that h is decreasing on $[0,\pi]$ and symmetric, we see that the supremum is reached for $\xi = N 1_{[-N/2,N/2]}$ which exactly leads to (48). For (ii), we set $f(s) = f_0(s)$ and $f(s)h(s) = f_0(s)\big[|2\sin(\frac{s}{2})|^{-2\varepsilon_n} - 1\big]$, apply (i) and then evaluate the integrals. □

We now study the behaviour of $\mathrm{Tr}\left([T_N(g_\varepsilon)T_N(f_0)^{-1}]^2\right)$. Such kind of estimates are considered (for statistical purposes) in Dahlhaus [6] or Fox and Taqqu [9], but cannot be directly applied here, since the spectral density g_{ε_n} does not remains unchanged as the size of the matrix $N \to \infty$. We however heavily rely on their techniques.

Lemma 8. *There exists some constant c, such that for all $n \geq 0$,*

$$\mathrm{Tr}\left(\left[T_N(g_{\varepsilon_n})T_N(f_0)^{-1}\right]^2\right) \leq cN\varepsilon_n^2.$$

Proof. Set

$$C := T_N(g_\varepsilon)^{1/2}T_N(f_0)^{-1}T_N(g_\varepsilon)^{1/2}, \;\; D := T_N(g_\varepsilon)^{1/2}T_N(f_0^{-1})T_N(g_\varepsilon)^{1/2}.$$

The aim of this Lemma is to show $\mathrm{Tr}(C^2) \leq cN\varepsilon_n^2$. Note that

$$\mathrm{Tr}(C^2) \leq \|C\|_{\mathrm{tr}}^2 \leq \|T_N(g_\varepsilon)^{\frac{1}{2}}T_N(f_0)^{-1/2}\|_{\mathrm{op}}^2 \|T_N(f_0)^{-1/2}T_N(g_\varepsilon)^{\frac{1}{2}}\|_{\mathrm{tr}}^2.$$

Using that $\|\cdot\|_{\mathrm{tr}} \leq \sqrt{N}\|\cdot\|_{\mathrm{op}}$ and that the operator norm of a matrix and of its transpose are equal, we get: $\mathrm{Tr}(C^2) \leq N\|T_N(g_\varepsilon)^{1/2}T_N(f_0)^{-1/2}\|_{\mathrm{op}}^4$.

Applying Lemma 7 we derive

$$\mathrm{Tr}(C^2) \le cN\varepsilon^2 \log(N)^2.$$

Thus the result on $\mathrm{Tr}(C^2)$ follows from the considerations on the operator norm given in Lemma 7 up to a log factor. We now present how to get rid of this logarithmic factor by considering directly $\mathrm{Tr}(C^2)$. The proof is divided into two steps: first we show, as in Dahlhaus [6], that the difference between $\mathrm{Tr}(D^2)$ and $\mathrm{Tr}(C^2)$ is negligible. Then we study $\mathrm{Tr}(D^2)$.

First step. After some algebra, we have that

$$|\mathrm{Tr}(C^2) - \mathrm{Tr}(D^2)| \le (\|C\|_{\mathrm{op}} + \|D\|_{\mathrm{op}})\|C - D\|_{\mathrm{tr}}.$$

Then writing

$$C - D = T_N(g_\varepsilon)^{1/2} T_N(f_0)^{-1/2}\big(\mathrm{Id} - T_N(f_0)^{1/2} T_N(f_0^{-1}) T_N(f_0)^{1/2}\big) \\ T_N(f_0)^{-1/2} T_N(g_\varepsilon)^{1/2},$$

we deduce by Lemma 7 that

$$\|C - D\|_{\mathrm{tr}} \le c\varepsilon \log(N) \|\mathrm{Id} - T_N(f_0)^{1/2} T_N(f_0^{-1}) T_N(f_0)^{1/2}\|_{\mathrm{tr}}.$$

But it is proved in Lemma 5.2 of [6] that the trace norm of the matrix $\mathrm{Id} - T_N(f_0)^{1/2} T_N(f_0^{-1}) T_N(f_0)^{1/2}$ grows at most with rate N^δ for an arbitrary small $\delta > 0$. By Lemma 7 again, we have $\|C\|_{\mathrm{op}} \le c\varepsilon \log(N)$ hence $\|D\|_{\mathrm{op}} \le c\varepsilon \log(N) + \|C - D\|_{\mathrm{tr}} \le c\varepsilon \log(N) N^\delta$. This eventually gives:

$$|Tr(C^2) - \mathrm{Tr}(D^2)| \le c\varepsilon^2 \log(N)^2 N^{2\delta},$$

which is clearly negligible versus $\varepsilon^2 N$ if δ is small enough.

Second step. Following the method of Fox and Taqqu [9], we study the asymptotic behavior of

$$\mathrm{Tr}(D^2) = \mathrm{Tr}\left(\left[T_N(g_\varepsilon) T_N(f_0^{-1})\right]^2\right).$$

Let us define $h_\varepsilon(s) = [|2\sin(s/2)|^{-\varepsilon} - 1]$, so that $g_\varepsilon = h_\varepsilon f_0$ (recall (45)). We can expand the trace of D^2 using the spectral densities as:

$$\mathrm{Tr}(D^2) = (2\pi)^{-4} \int h_\varepsilon(y_1) f_0(y_1) f_0^{-1}(y_2) h_\varepsilon(y_3) f_0(y_3) f_0(y_4)^{-1} P_N(y) \mathrm{d}y$$

where the integration is over $y = (y_1, \dots, y_4) \in [-\pi, \pi]^4$ and

$$P_N(y) = \sum_{1 \le j_1, \dots j_4 \le N} e^{i(j_1 - j_2)y_1} \dots e^{i(j_4 - j_1)y_4}.$$

This kernel P_N can be rewritten in the convenient way:

$$P_n(y) = \Delta_N(y_1 - y_2)\Delta_N(y_2 - y_3)\Delta_N(y_3 - y_4)\Delta_N(y_4 - y_1),$$

where we have set $\Delta_N(s) = \sum_{n=1}^N e^{-ins}$. An important fact on this kernel P_N is that it concentrate on the diagonal $y_1 = \dots = y_4$ and we shall make use of the crucial bound $|\Delta_N(s)| \le cL_N(s)$ where L_N is the 2π–periodic extension of the function such that

$$L_N(s) = \begin{cases} 1/|s| & \text{if } \pi \ge |s| \ge 1/N \\ N & \text{if } |s| \le 1/N \end{cases}$$

Now let J_N be the integral $J_N = (2\pi)^{-4}\int h_\varepsilon(y_1)h_\varepsilon(y_3)P_N(y)\mathrm{d}y$, the proof of the lemma consists in showing the two following points:

$$|\mathrm{Tr}(D^2) - J_N| = o(\varepsilon^2 N) \tag{49}$$

$$|J_N| \leq c\varepsilon^2 N \tag{50}$$

For the proof of (49) we consider first the integral

$$J'_N = \int h_\varepsilon(y_1)\frac{f_0(y_1)}{f_0(y_2)}h_\varepsilon(y_3)\{\frac{f_0(y_3)}{f_0(y_4)} - 1\}P_N(y)\mathrm{d}y.$$

As stated in Dalhaus in can be seen that for all $\delta > 0$ small enough we have the bound, $|\frac{f_0(y_3)}{f_0(y_4)} - 1| \leq c\frac{|y_3-y_4|^{1-3\delta}}{|y_3|^{1-\delta}}$. Hence, using periodicity of f_0, for $y_3, y_4 \in [-\pi,\pi]^2$ we deduce:

$$|\frac{f_0(y_3)}{f_0(y_4)} - 1| \leq c\frac{L_N(y_3-y_4)^{-1+3\delta}}{|y_3|^{1-\delta}}.$$

Together with $|h_\varepsilon(s)| \leq c\varepsilon|s|^{-\delta/2}$ (for any $\delta > 0$), $|f_0(y_1)| \leq c|y_1|^{-1+2H_0}$ and the boundness of $1/f_0(y_2)$ we deduce that,

$$|J'_N| \leq c\varepsilon^2 \int |y_1|^{2H_0-1-\delta/2}|y_3|^{-1+\delta/2} \\ L_N(y_1-y_2)L_N(y_2-y_3)L_N^{3\delta}(y_3-y_4)L_N(y_4-y_1)\mathrm{d}y.$$

Now we use that $L_N(y_3 - y_4) \leq N$ and then that we can integrate with respect to y_4, $\int L_N(y_4-y_1)\mathrm{d}y_4 \leq c\log(N)$ to deduce

$$|J'_N| \leq c\varepsilon^2 N^{3\delta}\log(N)\int |y_1|^{2H_0-1-\delta/2}|y_3|^{-1+\delta/2} \\ L_N(y_1-y_2)L_N(y_2-y_3)\mathrm{d}y_1\mathrm{d}y_2\mathrm{d}y_3.$$

Now, for $y_1, y_2, y_3 \in [-\pi,\pi]$ we can write,

$$L_N(y_1-y_2) \leq cN^\delta\left[\frac{1}{|y_1-y_2+2\pi|} + \frac{1}{|y_1-y_2|} + \frac{1}{|y_1-y_2-2\pi|}\right]^{1-\delta}$$

$$L_N(y_2-y_3) \leq cN^{1-5\delta}\left[\frac{1}{|y_2-y_3+2\pi|} + \frac{1}{|y_2-y_3|} + \frac{1}{|y_2-y_3-2\pi|}\right]^{5\delta}$$

This enable us obtain the bound,

$$|J'_N| \leq c\varepsilon^2 N^{1-\delta}\log(N)\int|y_1|^{2H_0-1-\delta/2}|y_3|^{-1+\delta/2} \\ \times\left[\frac{1}{|y_1-y_2+2\pi|} + \frac{1}{|y_1-y_2|} + \frac{1}{|y_1-y_2-2\pi|}\right]^{1-\delta} \\ \times\left[\frac{1}{|y_2-y_3+2\pi|} + \frac{1}{|y_2-y_3|} + \frac{1}{|y_2-y_3-2\pi|}\right]^{5\delta}\mathrm{d}y_1\mathrm{d}y_2\mathrm{d}y_3$$

As soon as δ is small enough, the latter integral converges by power counting criteria (see Theorem 3.1 in [9]) and we deduce that $J'_N = o(\varepsilon^2 N)$.

We can obtain similarly an analogous bound for the integral

$$J''_N = \int h_\varepsilon(y_1)\left\{\frac{f_0(y_1)}{f_0(y_2)} - 1\right\}h_\varepsilon(y_3)P_N(y)\mathrm{d}y_1\mathrm{d}y_2\mathrm{d}y_3\mathrm{d}y_4,$$

and (49) follows.

We now turn to the proof of (50). Using the property $(2\pi)^{-1}\int_{-\pi}^{\pi}\Delta_N(u-v)\Delta_N(v-w)\mathrm{d}v = \Delta_N(u-w)$, we can rewrite J_N as

$$J_N = \frac{1}{(2\pi)^2}\int_{[-\pi,\pi]^2} h_\varepsilon(y_1)h_\varepsilon(y_3)\Delta_N(y_1-y_3)\Delta_N(y_3-y_1)\mathrm{d}y_1\mathrm{d}y_3,$$

which is equal to $\sum_{1\le j_1,j_2\le N}\alpha_\varepsilon(j_1-j_2)\alpha_\varepsilon(j_2-j_1)$ where α_ε are the Fourier coefficient of the periodic function h_ε. Hence, we deduce that:

$$|J_N| \le N\sum_{k\in\mathbb{Z}}\alpha_\varepsilon(k)^2 \le N\|h_\varepsilon\|^2_{\mathbf{L}^2}.$$

But we can easily find the bound,

$$\|h_\varepsilon\|^2_{\mathbf{L}^2} = (2\pi)^{-1}\int_{-\pi}^{\pi}\left[|2\sin(s/2)|^{2\varepsilon}-1\right]^2\mathrm{d}s \le c\varepsilon^2.$$

This gives (49). □

Lemma 9. *We have*

$$|\mathrm{Tr}(T_N(A_\varepsilon)T_N(f_1)T_N(A_\varepsilon)^*T_N(f_1)^{-1})| \le cN\varepsilon_n^2. \tag{51}$$

Proof. First, let us recall that A_ε is bounded by $c\varepsilon$. Moreover we may assume that $A_\varepsilon(0)=0$ by subtracting the true value of $A_\varepsilon(0)$ and notice that it does not affect (51). Then the proof relies on the quasi commutative property of Toeplitz matrices. For this, let us write the left hand side of (51) as

$$\mathrm{Tr}(T_N(A_\varepsilon)T_N(A_\varepsilon)^*)+\mathrm{Tr}\left[(T_N(A_\varepsilon)T_N(f_1)-T_N(f_1)T_N(A_\varepsilon))T_N(A_\varepsilon)^*T_N(f_1)\right].$$

But $\mathrm{Tr}(T_N(A_\varepsilon)T_N(A_\varepsilon)^*) \le N\sum_k a_k(\varepsilon)^2 \le N\varepsilon^2$ and the other trace is less than:

$$\|T_N(A_\varepsilon)T_N(f_1)-T_N(f_1)T_N(A_\varepsilon)\|_{\mathrm{tr}}\|T_N(A_\varepsilon)^*T_N(f_1)^{-1}\|_{\mathrm{tr}}.$$

The remainder of the proof is broken in two steps.

First step: For some $\delta>0$, $\|T_N(A_\varepsilon)^*T_N(f_1)^{-1}\|_{\mathrm{tr}} \le cN^{1-\delta}\varepsilon$.

Second step: For any $\delta>0$, $\|T_N(A_\varepsilon)T_N(f_1)-T_N(f_1)T_N(A_\varepsilon)\|_{\mathrm{tr}} \le cN^\delta\varepsilon$.

For the first step, using that $T_N(f_1)^{-1}$ is bounded in operator norm, yields $\|T_N(A_\varepsilon)^*T_N(f_1)^{-1}\|_{\mathrm{tr}} \le cN^{1/2}\varepsilon$.

For the second step, we evaluate $\|T_N(A_\varepsilon)T_N(f_1)-T_N(f_1)T_N(A_\varepsilon)\|_{\mathrm{tr}}$ as in lemma 8 by the integral:

$$\int[f_1(y_1)A_\varepsilon(-y_2)-A_\varepsilon(-y_1)f_1(y_2)][A_\varepsilon(y_3)f_1(y_4)-f_1(y_3)A_\varepsilon(y_4)]P_N(y)\mathrm{d}y.$$

Using that the function in the integral above vanishes on the diagonals $y_1=y_2$ and $y_3=y_4$, it can be shown that this integral is of order $\varepsilon_n^2N^\delta$. □

Lemma 10. *We have*

$$|\mathrm{Tr}[T_N(B_{\varepsilon_n})T_N(f_1)T_N(B_{\varepsilon_n})^*]| \le cN\varepsilon_n^2.$$

Proof. We evaluate again the trace by the integral

$$\int B_\varepsilon(y_1)f_1(y_2)B_\varepsilon(-y_3)P_N(y)\mathrm{d}y.$$

Using that $|B_\varepsilon|$ is less than ε and vanishes in a neighborhood of zero, it can be shown that the integral above is of order $N\varepsilon^2$. □

Lemma 11. *Let $(\epsilon_n(\omega))_{n\in\mathbb{Z}}$ be identically distributed random variables with law $\mathcal{N}(0,1)$. Then there exists $C(\omega)$, with finite moment of any order, such that almost surely*

$$|\epsilon_n(\omega)| \leq C(\omega)(\log(2+|n|))^{\frac{1}{2}}.$$

If the family of random variable is enumerated as $(\epsilon_{j,k}(\omega))_{j\geq 0,k\in\mathbb{Z}}$ then we have instead:

$$|\epsilon_{j,k}(\omega)| \leq C(\omega)\big(\log(2+j)\big)^{\frac{1}{2}}\big(\log(2+|k|)\big)^{\frac{1}{2}}.$$

Proof. See Lemma 3 in Meyer *et al.* [20]. □

References

[1] Abry, P; Gonçalvès, P.; Flandrin P. (1995) *Wavelet, spectrum estimation and 1/f processes*, in: A. Antoniadis, G. Oppenheim (Eds.), Wavelet and Statistics, Lectures note in Statistics, Vol 103, Springer, New York, 1995, p.15–30.

[2] Abry, P.; Veitch, D. (1998) *Wavelet analysis of long range dependence traffic.* IEEE Trans. Inform. Theory, 44(1), p.2–15.

[3] Benassi, A.; Jaffard, S.; Roux, D. (1997) *Elliptic Gaussian random processes.* Rev. Mat. Iberoamericana, 13(1), p. 19–90.

[4] Brockwell, P.J.; Davis R.A. (1991) Time series: theory and methods. Second edition. Springer series in statistics, Springer–Verlag, New York.

[5] Ciesielski, Z.; Kerkyacharian, G.; Roynette, B. (1993) *Quelques espaces fonctionnels associés à des processus gaussiens.* Studia Math., 107(2), p.171–204.

[6] Dahlhaus, R. (1989) *Efficient parameter estimation for self-similar processes.* Ann. Statist., 17(4), p. 1749–1766.

[7] Delbeke, L.; Abry, P. (2000) *Stochastic integral representation and properties of the wavelet coefficients of linear fractional stable process.* Stochastic Process. Appl., 86(2), p.177–182.

[8] Efromovich, S.; Low, M. (1996) *On optimal adaptive estimation of a quadratic functional.* Ann. Statist. 24(3), 1106–1125.

[9] Fox, R.; Taqqu, M. (1987) *Central limit theorems for quadratic forms in random variables having long-range dependence.* Probab. Theory Related Fields, 74(2), p.213–240.

[10] Gayraud, G.; Tribouley, K. (1999) *Wavelet methods to estimate an integrated functional: Adaptivity and asymptotic law.* Statist. Probab. Lett., 44(2), p.109–122.

[11] Gloter, A.; Hoffmann, M. (2004) *Stochastic volatility and fractional Brownian motion.* Stochastic Proc. Appl., 113(1), 143-172.

[12] Gloter, A.; Jacod, J. (2001) *Diffusions with measurement errors.* ESAIM Probab. Stat., 5, p. 225-242

[13] Golubev, Y. (2003) *Personal communication.*

[14] Hirchoren, G.A.; D'Attellis, C.E. (1997) *On the optimal number of scales in estimation of fractal signals using wavelets and filter banks.* Signal Processing, 63(1), p.55–63.

[15] Istas, J; Lang, G. (1997) *Quadratic variations and estimation of the local Hölder index of a Gaussian process.* Ann. Inst. H Poincaré Probab. Statist., 33(4), p.407–436.

[16] Kawasaki, S.; Morita, H. *Evaluation of wavelet–base estimators associated with fractional Brownian motion.* Preprint.

[17] Lang, G.; Roueff, F. (2001) *Semi-parametric estimation of the Hölder exponent of a stationary Gaussian process with minimax rates.* Stat. Inference Stoch. Process., (4)3, p.283–306.

[18] Ludeña, C. (2004) *Minimum contrast estimation for fractional diffusions.* Scand. J. Statist. 31(4), p.613-628.

[19] Meyer, Y. (1990) *Ondelettes et opérateurs. I. (French) [Wavelets and operators. I] Ondelettes. [Wavelets]* Hermann, Paris.

[20] Meyer, Y.; Sellan, F. and Taqqu M. (1999) *Wavelets, generalized white noise and fractional integration: the synthesis of fractional Brownian motion.* J. Fourier Anal. Appl.,5(5), p.465–494.

[21] Samorodnitsky, G; Taqqu, M. (1994) *Stable non Gaussian random processes: stochastic models with infinite variance.* Chapman & Hall, New York.

[22] Scafetta, N.; Latora, V.; Grigolini, P. (2002) *Lévy scaling: the diffusion entropy analysis applied to DNA sequences.* Phys. Rev. E (3) 66(3), 031906, 15 pp.

[23] Tewfik, A.H.; Kim, M. (1992) *Correlation structure of the discrete wavelet coefficients of fractional Brownian motion.* IEEE Trans. Inform. Theory 38(2), p.904–909

[24] Tribouley, K. (2000) *Adaptive estimation of integrated functionals.* Math. Methods Statist. 9, 19–38.

[25] Veitch, D.; Abry, P. (1998) *A Wavelet based joint estimator of the parameter of long range dependence.* IEEE Trans. Inform. Theory (Special issue on multiscale statistical signal analysis and its applications), 45(3), p.878–897.

[26] Veitch, D.; Taqqu, M.S.; Abry, P. (2000) *Meaningful MRA initialization for discrete time series.* Signal Processing, 80(9), p.1971–1983.

[27] West, B.J; Grigolini, P. (2000) *Fractional differences, derivatives and fractal time series.* Applications of fractional calculus in physics, p.171–201, World Sci. Publishing, River Edge, NJ.

[28] Whittle, P. (1953) *Estimation and information in time series.* Ark. Mat. 2, p.423–434

[29] Willinger, W., Taqqu, M.S. and Leland W.E. (1995) *Self-similarity in high speed packet traffic: analysis and modeling of ethernet traffic measurements.* Statist. Sci., 10(1), p.67–85.