HAL
open science

One-step estimation for the fractional Gaussian noise at high-frequency

Alexandre Brouste, Marius Soltane, Irene Votsi

To cite this version:

Alexandre Brouste, Marius Soltane, Irene Votsi. One-step estimation for the fractional Gaussian noise at high-frequency. ESAIM: Probability and Statistics, 2020, 24, pp.827-841. 10.1051/ps/2020022 . hal-03022878

HAL Id: hal-03022878

https://hal.science/hal-03022878

Submitted on 25 Nov 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ESAIM: PS 24 (2020) 827–841
https://doi.org/10.1051/ps/2020022

ESAIM: Probability and Statistics
www.esaim-ps.org

ONE-STEP ESTIMATION FOR THE FRACTIONAL GAUSSIAN NOISE AT HIGH-FREQUENCY

Alexandre Brouste*, Marius Soltane and Irene Votsi

Abstract. The present paper concerns the parametric estimation for the fractional Gaussian noise in a high-frequency observation scheme. The sequence of Le Cam's one-step maximum likelihood estimators (OSMLE) is studied. This sequence is defined by an initial sequence of quadratic generalized variations-based estimators (QGV) and a single Fisher scoring step. The sequence of OSMLE is proved to be asymptotically efficient as the sequence of maximum likelihood estimators but is much less computationally demanding. It is also advantageous with respect to the QGV which is not variance efficient. Performances of the estimators on finite size observation samples are illustrated by means of Monte-Carlo simulations.

Mathematics Subject Classification. 62F12, 62M09, 65U05.

Received April 12, 2020. Accepted September 17, 2020.

1. Introduction

Beyond the classical model of independent and identically distributed Gaussian random variables for the returns of an asset price, several extensions have been proposed to describe the observed stylized facts by means of heteroscedasticity, non-Gaussianity and/or dependence. Here we focus on the fractional Gaussian noise (fGn) which is characterized by a standard deviation parameter (volatility) and a dependence parameter (Hurst exponent).

In the large sample statistical experiment where the mesh size is constant and the observation time horizon tends to infinity (as the sample size increases), the joint estimation for the volatility and the Hurst exponent has been well studied. The Local Asymptotic Normal (LAN) property of the likelihoods has been obtained in [6] for the fGn. The LAN theory allows to define the asymptotic efficiency of the sequences of estimators (for more information, see [13, 14, 18]). It is worth noting that the sequence of maximum likelihood estimators (MLE) is asymptotically efficient (see [7, 8]) and consequently presents the smallest possible variance in a minimax sense on local alternatives. The sequence of Whittle estimators, which is an approximation of the MLE, has been studied and shown to be efficient as well [10].

Contrary to the large sample setting, the LAN property of the likelihoods cannot be deduced from the previous results in the high-frequency scheme (infill asymptotics), where the time horizon is fixed and the mesh size tends to zero. In this setting, the joint estimation of the volatility and the self-similarity index is singular and therefore no confidence interval of the pair could be obtained [2, 16]. Recently, this singularity has been

Keywords and phrases: Fractional Gaussian noise, infill asymptotics, efficient estimation, Fisher scoring.

Laboratoire Manceau de Mathématiques, Le Mans Université, Le Mans, France.

* Corresponding author: alexandre.brouste@univ-lemans.fr

© The authors. Published by EDP Sciences, SMAI 2020

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

untied in [3] using non-diagonal norming rates and the asymptotic efficiency of the sequence of MLE has been proved in the high-frequency scheme. This singularity has been also observed in the symmetric stable model [1] when the scale parameter and the stability index are estimated jointly and has been untied in a similar way [4].

The sequence of Whittle type estimators is also efficient in the high-frequency fGn setting [11]. Other sequences of fast computable estimators have been studied including estimators based on quadratic generalized variations [15] and estimators based on wavelet coefficients [9]. Unfortunately, these sequences of estimators are generally not asymptotically efficient.

Although the sequence of MLE is asymptotically efficient in the high-frequency scheme, it is not expressed in a closed form and its computation is time consuming. The goal of the paper is the construction of a fast-computable sequence of estimators which is asymptotically equivalent to the MLE, the so-called sequence of one-step MLE (see [17, 19]). The latter sequence is based on an initial (rate efficient but not variance efficient) sequence of quadratic generalized variations-based estimators (QGV) and a single Fisher scoring step. QGV based estimators have been defined in [15] and a recent use in the context of the linear fractional stable motion is done in [20].

The one-step MLE presents certain advantages over the MLE and the QGV in terms of computational cost and asymptotic variance. It is much less computationally expensive than the MLE while it has the same rate and the same asymptotic variance. On the other hand, it is optimal in terms of asymptotic variance which is not the case for the QGV. Here we provide theoretical results to establish the asymptotic equivalence between the MLE and the one-step MLE and Monte-Carlo simulations are employed to illustrate finite size sample performances.

The remainder of the paper is organized as follows: Section 2 describes the fractional Gaussian noise model observed in the high-frequency scheme and the asymptotic properties of the sequence of MLE. In Section 3 the fast-computable sequence of one-step MLE is described and shown to be asymptotically efficient. Performances of the estimators for samples of medium size are illustrated by means of Monte-Carlo simulations.

2. Fractional Gaussian noise model at high-frequency

The model considered in this paper is the scaled fractional Gaussian noise that is,

$$X_i = \sigma \left(B_H \left(t_i^{(n)} \right) - B_H \left(t_{i-1}^{(n)} \right) \right), \quad i = 1, \ldots, n, \tag{2.1}$$

where $(B_H(t), t \geq 0)$ is a standard fractional Brownian motion of Hurst parameter $H \in (0,1)$, $\sigma > 0$ and $0 = t_0^{(n)} < t_1^{(n)} < \ldots < t_n^{(n)} = T$ is a regular time grid, where $t_i^{(n)} = \frac{iT}{n} = i\Delta_n$ for $0 \leq i \leq n$ and $\Delta_n = \frac{T}{n}$. In the high-frequency setting (infill asymptotics), the time horizon T is fixed and the mesh size $\Delta_n = t_1^{(n)} - t_0^{(n)} \longrightarrow 0$ as the sample size $n \longrightarrow \infty$. The joint estimation of the Hurst parameter H and the volatility parameter σ is considered.

The optimal rates for the estimation of the parameter $\vartheta = (H, \sigma)^* \in \Theta \subset (0,1) \times \mathbb{R}_*^+$, where $*$ denotes the transposition, have been recently obtained in [3]. For instance, the lower bounds for the risk of estimators are given for any $C > 0$ by

$$\liminf_{n \to \infty} \sup_{|\varphi_n(\vartheta_0)^{-1}(\vartheta - \vartheta_0)| \leq C} E_\vartheta^n \left[c \left(\varphi_n(\vartheta_0)^{-1} (\overline{\vartheta}_n - \vartheta) \right) \right] \geq \int_{\mathbb{R}^2} c \left(\mathcal{I}_\varphi(\vartheta_0)^{-1/2} z \right) \phi(z) \mathrm{d}z,$$

for any sequence of estimators $(\overline{\vartheta}_n, n \geq 1)$, some cost function c, some proper rate matrix φ_n and an asymptotic Fisher information matrix $\mathcal{I}_\varphi(\vartheta_0)$ described in the sequel. Here ϕ is the density of the 2-dimensional standard normal distribution.

2.1. Preliminaries on the fGn

For a fixed $n \in \mathbb{N}_*$, the distribution of the fractional Gaussian noise $(X_i, 1 \le i \le n)$ defined in (2.1) is a centered Gaussian distribution with autocovariance

$$r(k) = \mathbf{E} X_i X_{i+k} = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ikx} f_{\Delta_n}(x)\, dx,$$

for any $i \in \{1, \dots, n\}$ and $k \in \{0, \dots, n-i\}$, with

$$f_{\Delta_n}(x) = \sigma^2 C_H \Delta_n^{2H} 2(1-\cos(x)) \sum_{k \in \mathbb{Z}} |x + 2k\pi|^{-2H-1}, \quad x \in (-\pi, \pi],$$

where $C_H = \frac{\Gamma(2H+1)\sin(\pi H)}{2\pi}$. The spectral density $f_{\Delta_n}(x)$ can be rewritten as $f_{\Delta_n}(x) = \sigma^2 \Delta_n^{2H} f_H(x)$ where

$$f_H(x) = C_H 2(1-\cos(x)) \sum_{k \in \mathbb{Z}} |x + 2k\pi|^{-2H-1}, \quad x \in (-\pi, \pi].$$

The autocovariance function is explicitly given (see [22, Sect. 7.2.3]) by

$$r(k) = \frac{\sigma^2}{2} \Delta_n^{2H} \left(|k+1|^{2H} - 2|k|^{2H} + |k-1|^{2H} \right).$$

Consequently, for the observation sample $X^{(n)} = (X_1, \dots, X_n)$, the likelihood function in $\vartheta = (H, \sigma)^*$ admits a closed form given by

$$\mathcal{L}(\vartheta, X^{(n)}) = (2\pi)^{-\frac{n}{2}} (\det \Sigma_n)^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \langle X^{(n)}, \Sigma_n^{-1} X^{(n)} \rangle \right)$$

where $\Sigma_n = (r(i-j))_{i,j}$. Let us denote

$$A_n = \sqrt{n} \left\{ \frac{1}{n\sigma^2 \Delta_n^{2H}} \langle X^{(n)}, T_n(H)^{-1} X^{(n)} \rangle - 1 \right\}$$

and

$$B_n = \frac{1}{\sqrt{n}} \left\{ \frac{1}{2} \partial_H \log(\det T_n(H)) + \frac{1}{2\, \sigma^2 \Delta_n^{2H}} \langle X^{(n)}, \partial_H \{T_n(H)^{-1}\} X^{(n)} \rangle \right\},$$

where $T_n(H) = \sigma^{-2} \Delta_n^{-2H} \Sigma_n$. Then let us notice that, as shown in [6], the pair (A_n, B_n) converges in law under $P^n_{(H,\sigma)}$ for a nondegenerate centered Gaussian random variable (A, B) whose covariance matrix is given by

$$J(H) = \begin{pmatrix} 2 & -\frac{1}{2\pi} \int_{-\pi}^{\pi} \partial_H \log f_H(\lambda) \mathrm{d}\lambda \\ -\frac{1}{2\pi} \int_{-\pi}^{\pi} \partial_H \log f_H(\lambda) \mathrm{d}\lambda & \frac{1}{4\pi} \int_{-\pi}^{\pi} |\partial_H \log f_H(\lambda)|^2 \mathrm{d}\lambda \end{pmatrix}.$$

Here $P^n_{(H,\sigma)}$ stands for the probability measure associated to the fractional Gaussian noise with parameter (H, σ).

The log-likelihood function is further denoted by $\ell_n(\vartheta) = \log \mathcal{L}(\vartheta, X^{(n)})$ and the score function is given by

$$\nabla \ell_n(\vartheta) = \begin{pmatrix} \partial_H \ell_n(\vartheta) \\ \partial_\sigma \ell_n(\vartheta) \end{pmatrix} = \begin{pmatrix} A_n \sqrt{n} \log \Delta_n - B_n \sqrt{n} \\ A_n \sqrt{n} \sigma^{-1} \end{pmatrix}. \tag{2.2}$$

From the last expression, we clearly see that the leading terms of $\partial_H \ell_n(\vartheta)$ and $\partial_\sigma \ell_n(\vartheta)$ are linearly dependent, which is the reason why we obtain a singular Fisher information matrix when considering diagonal rate matrices as in [16]. This singularity could be untied using non-diagonal rate matrices as in [3].

2.2. Properties of the sequence of MLE

The sequence of maximum likelihood estimators $(\widehat{\vartheta}_n, n \geq 1)$ is defined by

$$\mathcal{L}\left(\widehat{\vartheta}_n, X^{(n)}\right) = \sup_{\vartheta \in \Theta} \mathcal{L}\left(\vartheta, X^{(n)}\right), \quad n \geq 1.$$

This sequence has been shown to be asymptotically normal and asymptotically efficient for different rate matrices in [3]. To be more precise, as $n \longrightarrow \infty$,

$$\varphi_n(\vartheta)^{-1}\left(\widehat{\vartheta}_n - \vartheta\right) \longrightarrow \mathcal{N}\left(0, \mathcal{I}_\varphi^{-1}(\vartheta)\right) \tag{2.3}$$

in law under $P^n_{(H,\sigma)}$ where

$$\varphi_n(\vartheta) = \frac{1}{\sqrt{n}} \begin{pmatrix} \alpha_n & \bar{\alpha}_n \\ \beta_n & \bar{\beta}_n \end{pmatrix} \tag{2.4}$$

satisfies the following properties:

1. The quantity ω defined by $\omega = n \det(\varphi_n(\vartheta)) = \alpha_n \bar{\beta}_n - \bar{\alpha}_n \beta_n$ verifies $\omega \neq 0$ and does not depend on n;
2. $\alpha_n \to \alpha$ for some $\alpha \in \mathbb{R}$;
3. $\bar{\alpha}_n \to \bar{\alpha}$ for some $\bar{\alpha} \in \mathbb{R}$;
4. $\gamma_n := \alpha_n \log \Delta_n + \beta_n \sigma^{-1} \to \gamma$ for some $\gamma \in \mathbb{R}$;
5. $\bar{\gamma}_n := \bar{\alpha}_n \log \Delta_n + \bar{\beta}_n \sigma^{-1} \to \bar{\gamma}$ for some $\bar{\gamma} \in \mathbb{R}$;
6. $\alpha\bar{\gamma} - \bar{\alpha}\gamma \neq 0$,

and where the Fisher information matrix can be computed as

$$\mathcal{I}_\varphi(\vartheta) = \begin{pmatrix} \gamma & -\alpha \\ \bar{\gamma} & -\bar{\alpha} \end{pmatrix} J(H) \begin{pmatrix} \gamma & \bar{\gamma} \\ -\alpha & -\bar{\alpha} \end{pmatrix}. \tag{2.5}$$

Following [3], several examples of rate matrices can be elicited. Let us briefly describe three examples.

1. A positive definite symmetric matrix for rate φ_n is

$$\varphi_n(\vartheta) = \frac{1}{\sqrt{n}} \begin{pmatrix} -\frac{2}{\log \Delta_n} & \sigma \\ \sigma & -\sigma^2 \log \Delta_n \end{pmatrix},$$

   for which $\alpha = 0$, $\bar{\alpha} = \sigma$, $\gamma = -1$, $\bar{\gamma} = 0$ and $\omega = \sigma^2$.
2. Two other examples of rate matrices will be used in the following. The first example is

$$\varphi_n(\vartheta) = \frac{1}{\sqrt{n}} \begin{pmatrix} 1 & 0 \\ -\sigma \log \Delta_n & 1 \end{pmatrix} \quad \text{with} \quad \varphi_n^{-1}(\vartheta) = \sqrt{n} \begin{pmatrix} 1 & 0 \\ \sigma \log \Delta_n & 1 \end{pmatrix}, \tag{2.6}$$

which implies $\alpha = 1$, $\bar{\alpha} = 0$, $\gamma = 0$, $\bar{\gamma} = \sigma^{-1}$ and $\omega = 1$. The second example is

$$\varphi_n(\vartheta) = \frac{1}{\sqrt{n}} \begin{pmatrix} \frac{1}{\log \Delta_n} & 1 \\ 0 & -\sigma \log \Delta_n \end{pmatrix} \quad \text{with} \quad \varphi_n^{-1}(\vartheta) = -\frac{\sqrt{n}}{\sigma} \begin{pmatrix} -\sigma \log \Delta_n & -1 \\ 0 & \frac{1}{\log \Delta_n} \end{pmatrix}, \tag{2.7}$$

implying $\alpha = 0$, $\bar{\alpha} = 1$, $\gamma = 1$, $\bar{\gamma} = 0$ and $\omega = -\sigma$.

It is worth noting that in all the previous examples the rate matrices are non-diagonal and depend on the parameter σ. Other examples could be found in [3].

3. Asymptotic efficiency of the one-step MLE

The numerical procedure for the evaluation of the MLE requires the computation of the inverse matrix $T_n(H)^{-1}$ and the use of an iterative numerical optimization algorithm. As a result, the maximum likelihood estimation is a computationally demanding and time-consuming task; therefore it is very important to propose less expensive estimation counterparts enjoying the same appealing asymptotic properties as the MLE. For this purpose we present the sequence of one-step MLE (OSMLE) which is asymptotically efficient and has a smaller computational cost than the MLE [3] and the Whittle estimator [11]. The OSMLE requires a single Fisher scoring iteration and a well-chosen initial sequence of guess estimators.

Let us first consider the matrix $\varphi_n(\vartheta)$ defined in (2.4). The corresponding non-degenerate Fisher information matrix is given by $\mathcal{I}_\varphi(\vartheta)$ defined in (2.5). We suppose that there exists an efficient rate sequence of estimators ("initial guess estimators") $\widetilde{\vartheta}_n = (\widetilde{H}_n, \widetilde{\sigma}_n)^*$ satisfying, as $n \longrightarrow \infty$,

$$\varphi_n(\vartheta)^{-1} \left(\widetilde{\vartheta}_n - \vartheta \right) \longrightarrow \mathcal{N}(0, \Gamma_\varphi(\vartheta))$$

in law under $P^n_{(H,\sigma)}$ for some (non-efficient) positive definite matrix $\Gamma_\varphi(\vartheta)$. Then, we consider the sequence of estimators $\overline{\vartheta}_n = (\overline{H}_n, \overline{\sigma}_n)^*$ defined by the one-step scoring procedure

$$\overline{\vartheta}_n = \widetilde{\vartheta}_n + \left(\varphi_n^{-*}(\widetilde{\vartheta}_n) \mathcal{I}_\varphi(\widetilde{\vartheta}_n) \varphi_n(\widetilde{\vartheta}_n)^{-1} \right)^{-1} \nabla \ell_n(\widetilde{\vartheta}_n). \tag{3.1}$$

Here A^{-*} denotes the transpose of the inverse matrix of A, namely $A^{-*} = (A^{-1})^*$.

In the next section, we present an initial guess estimator $(\widetilde{\vartheta}_n, n \geq 1)$ that could be used in the one-step procedure.

3.1. Initial guess estimator *via* quadratic generalized variations

First, let us denote the quadratic generalized variations

$$V_{n,1} = \frac{1}{n-1} \sum_{i=1}^{n-1} (X_{i+1} - X_i)^2 \quad \text{and} \quad V_{n,2} = \frac{1}{n-3} \sum_{i=1}^{n-3} (X_{i+3} + X_{i+2} - X_{i+1} - X_i)^2 .$$

Straightforward calculations lead to

$$\begin{aligned} \mathbf{E} V_{n,1} &= \frac{1}{n-1} \sum_{i=1}^{n-1} \mathbf{E} \left((X_{i+1} - X_i)^2 \right) = \mathbf{E} \left((X_2 - X_1)^2 \right) \\ &= 2 \left(r(0) - r(1) \right) = \sigma^2 \Delta_n^{2H} (4 - 2^{2H}). \end{aligned}$$

and

$$\begin{aligned}\mathbf{E}V_{n,2} &= \frac{1}{n-3}\sum_{i=1}^{n-3}\mathbf{E}\left((X_{i+3}+X_{i+2}-X_{i+1}-X_i)^2\right)\\ &= \sigma^2\Delta_n^{2H}2^{2H}(4-2^{2H}).\end{aligned}$$

Then, we define the sequence of initial guess estimators $(\widetilde{\vartheta}_n, n\geq 1)$ through quadratic generalized variations, where $\widetilde{\vartheta}_n = (\widetilde{H}_n, \widetilde{\sigma}_n)^*$, it is defined by

$$\widetilde{H}_n = \frac{1}{2}\log_2\frac{V_{n,2}}{V_{n,1}}$$

and

$$\widetilde{\sigma}_n = \sqrt{\frac{V_{n,1}}{\Delta_n^{2\widetilde{H}_n}}\cdot\frac{1}{4-2^{2\widetilde{H}_n}}}.$$

The next result states that the sequence of initial guess estimators $(\widetilde{\vartheta}_n, n\geq 1)$ is rate-efficient.

Theorem 3.1. *The sequence of quadratic generalized variations-based estimators $(\widetilde{\vartheta}_n, n\geq 1)$ is asymptotically normal, i.e. as $n\longrightarrow +\infty$,*

$$\varphi_n(\vartheta)^{-1}\left(\widetilde{\vartheta}_n-\vartheta\right)\longrightarrow \mathcal{N}(0,\Gamma_\varphi(\vartheta)),$$

in law under $P^n_{(H,\sigma)}$ where $\Gamma_\varphi(\vartheta)$ is a symmetric positive definite matrix defined in (3.5), which depends on σ and H.

Proof. First, let us denote

$$f(u,v) = \begin{pmatrix} h(u,v)\\ s(u,v)\end{pmatrix},$$

where $h(u,v)=\frac{1}{2}\log_2\left(\frac{v}{u}\right)$ and $s(u,v)=\sqrt{\frac{u}{w(u,v)}}$ with $w(u,v)=\Delta_n^{2h(u,v)}(4-2^{2h(u,v)})$. Direct computations lead to

$$\frac{\partial}{\partial u}f(u,v) = \begin{pmatrix} \frac{\partial}{\partial u}h(u,v)\\ -s(u,v)\log\Delta_n\frac{\partial}{\partial u}h(u,v)+\frac{s(u,v)2^{2h(u,v)}\log 2}{(4-2^{2h(u,v)})}\frac{\partial}{\partial u}h(u,v)+\frac{1}{2s(u,v)\Delta_n^{2h(u,v)}(4-2^{2h(u,v)})}\end{pmatrix}$$

and

$$\frac{\partial}{\partial v}f(u,v) = \begin{pmatrix} \frac{\partial}{\partial v}h(u,v)\\ -s(u,v)\log\Delta_n\frac{\partial}{\partial v}h(u,v)+\frac{s(u,v)2^{2h(u,v)}\log 2}{(4-2^{2h(u,v)})}\frac{\partial}{\partial v}h(u,v)\end{pmatrix}.$$

Second, we fix $u^\star=\sigma^2\Delta_n^{2H}(4-2^{2H})$ and $v^\star=\sigma^2\Delta_n^{2H}2^{2H}(4-2^{2H})$ such that

$$f(u^\star,v^\star) = \begin{pmatrix} H\\ \sigma\end{pmatrix}. \tag{3.2}$$

Taylor's expansion of f around $(u^\star, v^\star)$ gives

$$f(u,v) - f(u^\star,v^\star) = \frac{\partial}{\partial u} f(u^\star,v^\star)(u-u^\star) + \frac{\partial}{\partial v} f(u^\star,v^\star)(v-v^\star) + o\left(\left\|\begin{pmatrix} u-u^\star \\ v-v^\star \end{pmatrix}\right\|\right).$$

Then, for $\varphi_n(\vartheta)$ given by (2.6), we have

$$\begin{aligned}\varphi_n(\vartheta)^{-1}\left(f(u,v)-f(u^\star,v^\star)\right) &= \frac{\sqrt{n}}{\omega}\begin{pmatrix} \overline{\beta}_n & -\overline{\alpha}_n \\ -\beta_n & \alpha_n \end{pmatrix}\left(f(u,v)-f(u^\star,v^\star)\right) \\ &= \begin{pmatrix} a_{n,\vartheta}\frac{\sqrt{n}}{\Delta_n^{2H}}(u-u^\star) + b_{n,\vartheta}\frac{\sqrt{n}}{\Delta_n^{2H}}(v-v^\star) \\ c_{n,\vartheta}\frac{\sqrt{n}}{\Delta_n^{2H}}(u-u^\star) + d_{n,\vartheta}\frac{\sqrt{n}}{\Delta_n^{2H}}(v-v^\star) \end{pmatrix} \\ &\quad + o\left(\varphi_n(\vartheta)^{-1}\left\|\begin{pmatrix} u-u^\star \\ v-v^\star \end{pmatrix}\right\|\right) \end{aligned} \tag{3.3}$$

where

$$a_{n,\vartheta} = \frac{1}{\omega}\left[\left(\overline{\beta}_n + \overline{\alpha}_n\sigma\log\Delta_n - \overline{\alpha}_n\frac{\sigma 2^{2H}\log 2}{4-2^{2H}}\right)\cdot\left(-\frac{1}{2\log 2\sigma^2(4-2^{2H})}\right) - \overline{\alpha}_n\frac{1}{2\sigma(4-2^{2H})}\right],$$

$$b_{n,\vartheta} = \frac{1}{\omega}\left(\overline{\beta}_n + \overline{\alpha}_n\sigma\log\Delta_n - \overline{\alpha}_n\frac{\sigma 2^{2H}\log 2}{4-2^{2H}}\right)\cdot\left(\frac{1}{2\log 2\sigma^2 2^{2H}(4-2^{2H})}\right),$$

$$c_{n,\vartheta} = \frac{1}{\omega}\left[\left(-\beta_n - \alpha_n\sigma\log\Delta_n + \alpha_n\frac{\sigma 2^{2H}\log 2}{4-2^{2H}}\right)\cdot\left(-\frac{1}{2\log 2\sigma^2(4-2^{2H})}\right) + \alpha_n\frac{1}{2\sigma(4-2^{2H})}\right],$$

and

$$d_{n,\vartheta} = \frac{1}{\omega}\left(-\beta_n - \alpha_n\sigma\log\Delta_n + \alpha_n\frac{\sigma 2^{2H}\log 2}{4-2^{2H}}\right)\cdot\left(\frac{1}{2\log 2\sigma^2 2^{2H}(4-2^{2H})}\right).$$

Due to the properties of the rate matrix $\varphi_n(\vartheta)$, we note that $a_{n,\vartheta}$, $b_{n,\vartheta}$, $c_{n,\vartheta}$ and $d_{n,\vartheta}$ converge, as $n \longrightarrow \infty$, to

$$a_\vartheta = \frac{1}{\omega}\left[\left(\overline{\gamma}\sigma - \overline{\alpha}\frac{\sigma 2^{2H}\log 2}{4-2^{2H}}\right)\cdot\left(-\frac{1}{2\log 2\sigma^2(4-2^{2H})}\right) - \overline{\alpha}\frac{1}{2\sigma(4-2^{2H})}\right],$$

$$b_\vartheta = \frac{1}{\omega}\left(\overline{\gamma}\sigma - \overline{\alpha}\frac{\sigma 2^{2H}\log 2}{4-2^{2H}}\right)\cdot\left(\frac{1}{2\log 2\sigma^2 2^{2H}(4-2^{2H})}\right),$$

$$c_\vartheta = \frac{1}{\omega}\left[\left(-\gamma\sigma + \alpha\frac{\sigma 2^{2H}\log 2}{4-2^{2H}}\right)\cdot\left(-\frac{1}{2\log 2\sigma^2(4-2^{2H})}\right) + \alpha\frac{1}{2\sigma(4-2^{2H})}\right],$$

and

$$d_\vartheta = \frac{1}{\omega}\left(-\gamma\sigma + \alpha\frac{\sigma 2^{2H}\log 2}{4-2^{2H}}\right)\cdot\left(\frac{1}{2\log 2\sigma^2 2^{2H}(4-2^{2H})}\right),$$

respectively.

For any $H \in (0,1)$, the quadratic generalized variations were proved to be asymptotically normal (see [15, Thm. 2 (ii)]), that is, as $n \longrightarrow \infty$

$$\frac{\sqrt{n}}{\Delta_n^{2H}} \begin{pmatrix} V_{n,1} - \mathbf{E}_\vartheta V_{n,1} \\ V_{n,2} - \mathbf{E}_\vartheta V_{n,2} \end{pmatrix} \longrightarrow \mathcal{N}(0, E(\vartheta)\Gamma_1(H)E(\vartheta)) \tag{3.4}$$

in law under $P^n_{(H,\sigma)}$ with

$$E(\vartheta) = \sigma^2(4 - 2^{2H}) \begin{pmatrix} 1 & 0 \\ 0 & 2^{2H} \end{pmatrix}$$

and

$$\Gamma_1(H) = \begin{pmatrix} \gamma(\mu,\mu) & \gamma(\mu,\nu) \\ \gamma(\nu,\mu) & \gamma(\nu,\nu) \end{pmatrix},$$

where $\mu = (1,-2,1)$, $\nu = (1,0,-2,0,1)$ and

$$\gamma(\mu,\nu) = \frac{2\sum_{j\in\mathbb{Z}} \left(\sum_{k,\ell} \mu_k \nu_\ell \left|j+k-\ell\right|^{2H}\right)^2}{\left(\sum_{k,\ell} \mu_k \mu_\ell \left|k-\ell\right|^{2H}\right)\left(\sum_{k,\ell} \nu_k \nu_\ell \left|k-\ell\right|^{2H}\right)}.$$

Consequently, by replacing $u = V_{n,1}$ and $v = V_{n,2}$ in (3.3), we obtain firstly

$$f(V_{n,1}, V_{n,2}) = \begin{pmatrix} h(V_{n,1}, V_{n,2}) \\ s(V_{n,1}, V_{n,2}) \end{pmatrix} = \begin{pmatrix} \widetilde{H}_n \\ \widetilde{\sigma}_n \end{pmatrix} = \widetilde{\vartheta}_n.$$

Together with (3.2) and (3.4), we get

$$\varphi_n(\vartheta)^{-1}\left(\widetilde{\vartheta}_n - \vartheta\right) = \begin{pmatrix} a_\vartheta & b_\vartheta \\ c_\vartheta & d_\vartheta \end{pmatrix} \frac{\sqrt{n}}{\Delta_n^{2H}} \begin{pmatrix} V_{n,1} - \mathbf{E}_\vartheta V_{n,1} \\ V_{n,2} - \mathbf{E}_\vartheta V_{n,2} \end{pmatrix} + o_{P^n_\vartheta}(1).$$

Then the application of the delta method (see for instance [24, Thm. 3.1]) leads directly to the desired result with asymptotic variance defined by

$$\Gamma_\varphi(\vartheta) = \begin{pmatrix} a_\vartheta & b_\vartheta \\ c_\vartheta & d_\vartheta \end{pmatrix} E(\vartheta)\Gamma_1(H)E(\vartheta) \begin{pmatrix} a_\vartheta & b_\vartheta \\ c_\vartheta & d_\vartheta \end{pmatrix}^*. \tag{3.5}$$

□

Let us mention that in the previous proof the quadratic generalized variations satisfy the central limit theorem for any value $H \in (0,1)$. The more classical quadratic variations statistic $\tilde{V}_{n,1} = \frac{1}{n}\sum_{i=1}^n X_i^2$ satisfies the central limit theorem only for $H < \frac{3}{4}$ (see [15]) and the corresponding estimators would present a different behavior. It is worth emphasizing that this result could be extended to other rate optimal initial guess sequences of estimators (quadratic generalized variations of sufficiently high order [15], power variations estimators [12] or wavelets-based estimators [9]). For a comparison of such initial sequences, the reader can refer for instance to [5].

3.2. Properties of the one-step MLE sequence

We present now one of the main results of the paper concerning the sequence of one-step MLE, $\overline{\vartheta}_n = (\overline{H}_n, \overline{\sigma}_n)^*$, defined in (3.1).

Theorem 3.2. *The sequence of one-step MLE $(\overline{\vartheta}_n, n \geq 1)$ is asymptotically normal, i.e., as $n \longrightarrow \infty$,*

$$\varphi_n(\vartheta)^{-1}\left(\overline{\vartheta}_n - \vartheta\right) \longrightarrow \mathcal{N}\left(0, \mathcal{I}_\varphi^{-1}(\vartheta)\right).$$

in law under $P^n_{(H,\sigma)}$ and asymptotically efficient (in Fisher's sense, i.e. with the same rate and the same asymptotic variance as the MLE sequence).

Proof. First, from (3.1), we have directly

$$\varphi_n(\vartheta)^{-1}\left(\bar{\vartheta}_n - \vartheta\right) = \varphi_n(\vartheta)^{-1}\left(\widetilde{\vartheta}_n - \vartheta\right) + \varphi_n(\vartheta)^{-1}\left(\varphi_n(\widetilde{\vartheta}_n)^{-*}\mathcal{I}_\varphi(\widetilde{\vartheta}_n)\varphi_n(\widetilde{\vartheta}_n)^{-1}\right)^{-1}\nabla\ell_n(\widetilde{\vartheta}_n).$$

Let us denote I_2 the 2×2 identity matrix. We consider a $\vartheta^* \in B(\vartheta, |\widetilde{\vartheta}_n - \vartheta|)$, where $B(\vartheta, r)$ is a ball of radius r centered in ϑ. Then Taylor's expansion of $\nabla\ell_n(\widetilde{\vartheta}_n)$ around ϑ gives

$$\begin{aligned}
\varphi_n(\vartheta)^{-1}\left(\bar{\vartheta}_n - \vartheta\right) &= \varphi_n(\vartheta)^{-1}\left(\widetilde{\vartheta}_n - \vartheta\right) \\
&\quad + \varphi_n(\vartheta)^{-1}\left(\varphi_n(\widetilde{\vartheta}_n)^{-*}\mathcal{I}_\varphi(\widetilde{\vartheta}_n)\varphi_n(\widetilde{\vartheta}_n)^{-1}\right)^{-1}\left(\nabla\ell_n(\vartheta) + \nabla^2\ell_n(\vartheta^*)\left(\widetilde{\vartheta}_n - \vartheta\right)\right) \\
&= \varphi_n(\vartheta)^{-1}\left(I_2 + \left(\varphi_n(\widetilde{\vartheta}_n)^{-*}\mathcal{I}_\varphi(\widetilde{\vartheta}_n)\varphi_n(\widetilde{\vartheta}_n)^{-1}\right)^{-1}\nabla^2\ell_n(\vartheta^*)\right)\left(\widetilde{\vartheta}_n - \vartheta\right) \\
&\quad + \varphi_n(\vartheta)^{-1}\left(\varphi_n(\widetilde{\vartheta}_n)^{-*}\mathcal{I}_\varphi(\widetilde{\vartheta}_n)\varphi_n(\widetilde{\vartheta}_n)^{-1}\right)^{-1}\nabla\ell_n(\vartheta) \\
&= \varphi_n(\vartheta)^{-1}\left(I_2 + \left(\varphi_n(\widetilde{\vartheta}_n)^{-*}\mathcal{I}_\varphi(\widetilde{\vartheta}_n)\varphi_n(\widetilde{\vartheta}_n)^{-1}\right)^{-1}\varphi_n(\vartheta)^{-*}\varphi_n(\vartheta)^*\nabla^2\ell_n(\vartheta^*)\varphi_n(\vartheta)\right. \\
&\qquad \left.\times\,\varphi_n(\vartheta)^{-1}\right)\left(\widetilde{\vartheta}_n - \vartheta\right) + \varphi_n(\vartheta)^{-1}\left(\varphi_n(\widetilde{\vartheta}_n)^{-*}\mathcal{I}_\varphi(\widetilde{\vartheta}_n)\varphi_n(\widetilde{\vartheta}_n)^{-1}\right)^{-1}\nabla\ell_n(\vartheta) \\
&= \varphi_n(\vartheta)^{-1}\left(I_2 - \left(\varphi_n(\widetilde{\vartheta}_n)^{-*}\mathcal{I}_\varphi(\widetilde{\vartheta}_n)\varphi_n(\widetilde{\vartheta}_n)^{-1}\right)^{-1}\varphi_n(\vartheta)^{-*}\mathcal{I}_n(\vartheta^*)\varphi_n(\vartheta)^{-1}\right) \\
&\quad \times\left(\widetilde{\vartheta}_n - \vartheta\right) + \varphi_n(\vartheta)^{-1}\left(\varphi_n(\widetilde{\vartheta}_n)^{-*}\mathcal{I}_\varphi(\widetilde{\vartheta}_n)\varphi_n(\widetilde{\vartheta}_n)^{-1}\right)^{-1}\nabla\ell_n(\vartheta)
\end{aligned} \tag{3.6}$$

where $\mathcal{I}_n(\vartheta) = -\varphi_n(\vartheta)\nabla^2\ell_n(\vartheta)\varphi_n(\vartheta)$.

If the following Sweeting's conditions (see for instance [23] or [4]) hold true, then the first term on the right-hand side of equation (3.6) tends in probability to 0 and the second term tends in law to a centered normal random variable with variance $\mathcal{I}_\varphi^{-1}(\vartheta)$. Sweeting's conditions are:

C1. The uniform convergence

$$\mathcal{I}_n(\vartheta) = -\varphi_n(\vartheta)\nabla^2\ell_n(\vartheta)\varphi_n(\vartheta) \longrightarrow_u \mathcal{I}_\varphi(\vartheta).$$

Here, we denote by $\rightarrow_u$ the ordinary uniform convergence with respect to ϑ over any compact set contained in $(0,1)\times(0,\infty)$.

C2. Control on the norming matrix φ_n in the following sense: for any $c > 0$

$$\sup_{\vartheta' \in \mathfrak{N}_n(c;\theta)} |\varphi_n(\vartheta')^{-1}\varphi_n(\vartheta) - I_2| \rightarrow_u 0,$$

where

$$\mathfrak{N}_n(c;\vartheta) := \left\{\vartheta' \in (0,1)\times(0,\infty);\ |\varphi_n(\vartheta)^{-1}(\vartheta'-\vartheta)| \le c\right\}$$

is a shrinking neighborhood of ϑ, and I_k denotes the $k\times k$-identity matrix.

C3. Another kind of uniform convergence of $\mathcal{I}_n(\theta)$, that is,

$$\sup_{\theta^1,\theta^2\in\mathfrak{N}_n(c;\theta)} \left|\varphi_n(\theta)^\top\{\nabla^2\ell_n(\theta^1,\theta^2)-\nabla^2\ell_n(\theta)\}\varphi_n(\theta)\right| \to_u 0$$

for each $c>0$, where

$$\nabla^2\ell_n(\theta^1,\theta^2) := \begin{pmatrix} \partial_H^2\ell_n(\theta^1) & \partial_H\partial_\sigma\ell_n(\theta^1) \\ \partial_H\partial_\sigma\ell_n(\theta^2) & \partial_\sigma^2\ell_n(\theta^2)\end{pmatrix}.$$

We go one step further and prove the aforementioned Sweeting's conditions.

C1. For the proof, see ([3], Thm. 3.1).

C2. *Proof.* For any positive functions $a_n(\vartheta)$ and $b_n(\vartheta)$, we write $a_n(\vartheta)\lesssim_u b_n(\vartheta)$ if there exists a universal constant C such that $\sup_{\vartheta\in K}|a_n(\vartheta)/b_n(\vartheta)|\le C$ for any compact set K contained in $(0,1)\times(0,\infty)$ and every n large enough. Moreover, for notational simplicity, we write $a_n(\vartheta,\vartheta')=o_{u,c}(1)$ if $\sup_{\vartheta'\in\mathfrak{N}_n(c;\vartheta)}|a_n(\vartheta,\vartheta')|\to_u 0$.

First, we set $\omega(\vartheta)=\alpha_{n,\vartheta}\bar\beta_{n,\vartheta}-\beta_{n,\vartheta}\bar\alpha_{n,\vartheta}$, $\delta_{n,\vartheta}=\alpha_{n,\vartheta}\bar\gamma_{n,\vartheta}-\bar\alpha_{n,\vartheta}\bar\gamma_{n,\vartheta}$ and $\delta(\vartheta)=\alpha_\vartheta\bar\gamma_\vartheta-\bar\alpha_\vartheta\bar\gamma_\vartheta$. Second, since

$$\varphi_n(\vartheta)^{-1}=\frac{\sqrt{n}}{\omega(\vartheta)}\begin{pmatrix}\bar\beta_{n,\vartheta} & -\bar\alpha_{n,\vartheta}\\ -\bar\beta_{n,\vartheta} & \bar\alpha_{n,\vartheta}\end{pmatrix},$$

we have directly

$$\varphi_n(\vartheta')^{-1}\varphi_n(\vartheta)=\frac{1}{\omega(\vartheta')}\begin{pmatrix}\bar\beta_{n,\vartheta'}\alpha_{n,\vartheta}-\bar\alpha_{n,\vartheta'}\beta_{n,\vartheta} & \bar\beta_{n,\vartheta'}\bar\alpha_{n,\vartheta}-\bar\alpha_{n,\vartheta'}\beta_{n,\vartheta}\\ \alpha_{n,\vartheta'}\beta_{n,\vartheta}-\beta_{n,\vartheta'}\alpha_{n,\vartheta} & \alpha_{n,\vartheta'}\bar\beta_{n,\vartheta}-\beta_{n,\vartheta'}\bar\alpha_{n,\vartheta}\end{pmatrix}.$$

Then since $\beta_{n,\vartheta}=\sigma\gamma_{n,\vartheta}-\alpha_{n,\vartheta}\sigma\log(\Delta_n)$ and $\bar\beta_{n,\vartheta}=\sigma\bar\gamma_{n,\vartheta}-\bar\alpha_{n,\vartheta}\sigma\log(\Delta_n)$, straightforward calculations lead to the following expression of the $(1,2)$ element of the aforementioned matrix:

$$\frac{\bar\alpha_{n,\vartheta}\bar\alpha_{n,\vartheta'}\log(\Delta_n)(\sigma-\sigma')}{\sigma'\delta(\vartheta')}+o_{u,c}(1).$$

We further notice that

$$\varphi_n(\vartheta')^{-1}(\vartheta'-\vartheta)=A_n(\vartheta)b_n(\vartheta,\vartheta'),$$

where

$$A_n(\vartheta)=\frac{1}{\sigma\delta_{n,\vartheta}}\begin{pmatrix}\bar\gamma_{n,\vartheta} & -\bar\alpha_{n,\vartheta}\\ -\gamma_{n,\vartheta} & \alpha_{n,\vartheta}\end{pmatrix}$$

and

$$b_n(\vartheta,\vartheta')=\begin{pmatrix}\sqrt{n}\sigma(H-H')\\ \sqrt{n}\sigma\log(\Delta_n)(H'-H)+\sqrt{n}(\sigma'-\sigma)\end{pmatrix}.$$

By the continuity in ϑ, we have

$$A_n(\vartheta)^* A_n(\vartheta) = \frac{1}{\sigma^2\delta^2(\vartheta)} \begin{pmatrix} \bar{\gamma}_\vartheta^2 + \gamma_\vartheta^2 & -(\alpha_\vartheta\gamma_\vartheta + \bar{\alpha}_\vartheta\bar{\gamma}_\vartheta) \\ \text{sym.} & \alpha_\vartheta^2 + \bar{\alpha}_\vartheta^2 \end{pmatrix} + o_{u,c}(1).$$

The smallest eigenvalue of the first term in the right-hand side matrix is

$$\lambda_\vartheta^{min} = \frac{1}{2\sigma^2\delta^2(\vartheta)}\big(f(\vartheta) - \sqrt{f(\vartheta)^2 - 4g(\vartheta)}\big),$$

where $f(\vartheta) = \bar{\gamma}_\vartheta^2 + \gamma_\vartheta^2 + \bar{\alpha}_\vartheta^2 + \alpha_\vartheta^2$ and $g(\vartheta) = (\alpha_\vartheta\gamma_\vartheta + \bar{\alpha}_\vartheta\bar{\gamma}_\vartheta)^2$. We further assume that $\inf_{\vartheta\in K}|\delta(\vartheta)| > 0$, for any compact set $K \subset (0,1)\times(0,\infty)$. Then since $|\lambda_\vartheta^{min}|^{-1} \lesssim_u 1$, we have that

$$\sup_{\vartheta'\in\mathfrak{N}_n(c;\vartheta)} |b_n(\vartheta,\vartheta')| \lesssim_u 1,$$

which implies

$$\sup_{\vartheta'\in\mathfrak{N}_n(c;\vartheta)} \left|\frac{\sqrt{n}}{\log(\Delta_n)}(\sigma' - \sigma)\right| \lesssim_u 1,$$

and

$$\sup_{\vartheta'\in\mathfrak{N}_n(c;\vartheta)} |\log(\Delta_n)(\sigma' - \sigma)| \to_u 0.$$

Treating the remaining elements of the matrix $\varphi_n(\vartheta')^{-1}\varphi_n(\vartheta)$ in a similar way, we have that

$$\sup_{\vartheta'\in\mathfrak{N}_n(c;\vartheta)} |\varphi_n(\vartheta')^{-1}\varphi_n(\vartheta) - I_2| \to_u 0.$$

□

C3. Proof. We first notice that

$$\begin{aligned} &\sup_{\vartheta^1,\vartheta^2\in\mathfrak{N}_n(c;\theta)} \left|\varphi_n(\vartheta)^*\{\nabla^2\ell_n(\vartheta^1,\vartheta^2) - \nabla^2\ell_n(\vartheta)\}\varphi_n(\vartheta)\right| \\ \lesssim_u &\sup_{\vartheta^1,\vartheta^2,\vartheta'\in\mathfrak{N}_n(c;\theta)} \left|\varphi_n(\vartheta)^*\{\partial_\vartheta^3\ell_n(\vartheta^1,\vartheta^2)[\vartheta' - \vartheta]\}\varphi_n(\vartheta)\right|. \end{aligned}$$

We further use the decomposition of the tensor $-\varphi_n^*\partial^3\ell_n\varphi_n$ described in [3, Thm. 3.1] and obtain the desired result.

□

Then Sweeting's conditions in (3.6) lead directly to

$$\varphi_n(\vartheta)^{-1}\left(\bar{\vartheta}_n - \vartheta\right) \longrightarrow \mathcal{N}(0, \mathcal{I}_\varphi^{-1}(\vartheta)).$$

as $n \longrightarrow \infty$ in law under $P^n_{(H,\sigma)}$. □

3.3. Simulation study

The objective of this section is to compare the performance of the three estimators of interest (QGV, MLE, OSMLE) for samples of medium size by means of Monte-Carlo simulations. The one-step MLE presents certain advantages over the MLE and the QGV in terms of computational burden and asymptotic variance, respectively. First, it is much less computationally expensive than the MLE while it has the same rate and the same asymptotic variance. Indeed, the computation of MLE necessitates iterative optimization procedures (as the Fisher scoring described in (3.1)) and a convergence criterion. On the contrary, with a proper initial condition, the OSMLE has the same appealing asymptotic properties as the MLE and, since it requires a single Fisher scoring step, it reduces significantly the computational cost. Second, the OSMLE is asymptotically efficient and outperforms the QGV in terms of asymptotic variance.

In this context, the score function (2.2) cannot be defined explicitly since it relies on the quantities $\partial_H \log\det(T_n(H))$ and $\partial_H\{T_n(H)^{-1}\}$. In the numerical procedure, the following equalities are used

$$\partial_H\{T_n(H)^{-1}\} = -T_n(H)^{-1}\partial_H\{T_n(H)\}T_n(H)^{-1},$$

and

$$\partial_H \log\det(T_n(H)) = \text{ trace}\,(\partial_H\{T_n(H)^{-1}\}T_n(H)).$$

The non-singular Fisher information matrix is evaluated numerically through the method described in [11].

3.3.1. Statistical error for H

In order to compute the renormalized statistical errors $\sqrt{n}(\widetilde{H}_n - H)$ (QGV), $\sqrt{n}(\widehat{H}_n - H)$ (MLE) and $\sqrt{n}(\overline{H}_n - H)$ (OSMLE), we consider the matrix $\varphi_n(\vartheta)$ defined in (2.6). The corresponding non-degenerate Fisher information matrix is given by (2.3), that is,

$$\mathcal{I}_\varphi(\vartheta) = \begin{pmatrix} \frac{1}{4\pi}\int_{-\pi}^{\pi}|\partial_H \log f_H(\lambda)|^2 \mathrm{d}\lambda & \frac{1}{2\pi\sigma}\int_{-\pi}^{\pi}\partial_H \log f_H(\lambda)\mathrm{d}\lambda \\ \frac{1}{2\pi\sigma}\int_{-\pi}^{\pi}\partial_H \log f_H(\lambda)\mathrm{d}\lambda & \frac{2}{\sigma^2} \end{pmatrix}. \tag{3.7}$$

The asymptotic variance of the QGV given by (3.5) is computed by means of the terms

$$a_\vartheta = -\frac{1}{2\log 2\sigma^2(4-2^{2H})}, \quad b_\vartheta = \frac{1}{2\log 2\sigma^2 2^{2H}(4-2^{2H})},$$

$$c_\vartheta = \frac{4-2\cdot 2^{2H}}{2\sigma(4-2^{2H})^2} \quad \text{and} \quad d_\vartheta = \frac{1}{2\sigma(4-2^{2H})^2}.$$

A Monte-Carlo procedure (with 10000 iterations) was followed to explore the behavior of the different estimators (QGV, MLE and OSMLE) of the parameter H at the medium sample size $n = 2^8$. The results are depicted in Figure 1 for $\sigma = 0.5$ and for two different values of H : $H = 0.2$ (short-memory) and $H = 0.8 > \frac{3}{4}$ (long-memory). In particular, the computations were implemented in *R* [21].

The QGV estimator is clearly not optimal with respect to the asymptotic variance for both values of H. The theoretical asymptotic variance given in (3.5) and the asymptotic efficient variance given in (3.7) are superimposed. On the other side, the (asymptotically efficient) MLE of H has an appealing asymptotic behavior for the medium sample size $n = 2^8$ and small values of H. However, when greater values of H generate the data, the MLE presents a rather small bias which naturally disappears with the increase of the sample size.

The bias that characterizes the OSMLE for medium sample sizes, comes from the bias of the initial QGV estimate of σ (see Sect. 3.3.2) and disappears naturally as the sample size augments. To be more precise, the bias is transferred to both components of the OSMLE due to the mixing property of the non-diagonal rate matrix.

3.3.2. Statistical error for σ

Let us now consider the matrix $\varphi_n(\vartheta)$ defined in (2.7) with corresponding non-degenerate Fisher information matrix

$$\mathcal{I}_\varphi(\vartheta) = \begin{pmatrix} 2 & \frac{1}{2\pi}\int_{-\pi}^{\pi} \partial_H \log f_H(\lambda)\mathrm{d}\lambda \\ \frac{1}{2\pi}\int_{-\pi}^{\pi} \partial_H \log f_H(\lambda)\mathrm{d}\lambda & \frac{1}{4\pi}\int_{-\pi}^{\pi} |\partial_H \log f_H(\lambda)|^2\mathrm{d}\lambda \end{pmatrix}.$$

The asymptotic variance of the QGV could be computed by means of the terms

$$a_\vartheta = \frac{4 - 2\cdot 2^{2H}}{2\sigma^2(4-2^{2H})^2}, \quad b_\vartheta = \frac{1}{2\sigma^2(4-2^{2H})^2},$$

$$c_\vartheta = -\frac{1}{2\log 2\sigma^2(4-2^{2H})} \quad \text{and} \quad d_\vartheta = \frac{1}{2\log 2\sigma^2 2^{2H}(4-2^{2H})}.$$

Then by means of Monte-Carlo simulations, we generate 10 000 trajectories of the fGn model for $n = 2^8$, $\sigma = 0.5$ and two values of the Hurst parameter H: $H = 0.2$ and $H = 0.8$. Figure 1 displays the distribution of the three estimators of σ, for the simulated trajectories generated with $H = 0.2$ and $H = 0.8$.

We observe that the QGV estimator of σ is not optimal with respect to the asymptotic variance for both values of H and presents a visible bias at the medium sample size $n = 2^8$ (disappearing naturally at a large sample size n). In the short-memory setting, the MLE has appealing properties for a medium sample size $n = 2^8$. Nevertheless, for greater values of H, the MLE presents a visible bias, as it was also observed in [11]. The bias of the OSMLE at medium sample sizes n (disappearing naturally at large sample sizes n) derives from the bias of the initial QGV estimate of σ (see Sect. 3.3.2). Furthermore, it could be noticed that the behavior of the OSMLE is similar to the MLE (or the Whittle estimator).

3.3.3. Computation time

The calculation times needed for the Monte-Carlo procedure to generate random samples of the estimators were counted and displayed in Table 1. For $H = 0.2$ (resp. $H = 0.8$), the Monte-Carlo procedure for MLE (Col. 2) requires approximately 5.6 (resp. 3.4) times the calculation time of the OSMLE (Col. 1). This naturally shows that OSMLE outperforms the MLE in terms of computation time.

Table 1. Calculation time (in seconds) for generating random samples of OSMLE and MLE of size $n = 2^8$ for 10 000 Monte-Carlo simulations.

	OSMLE	MLE
$H = 0.2$	6280.627	35466.128
$H = 0.8$	5311.633	18231.750

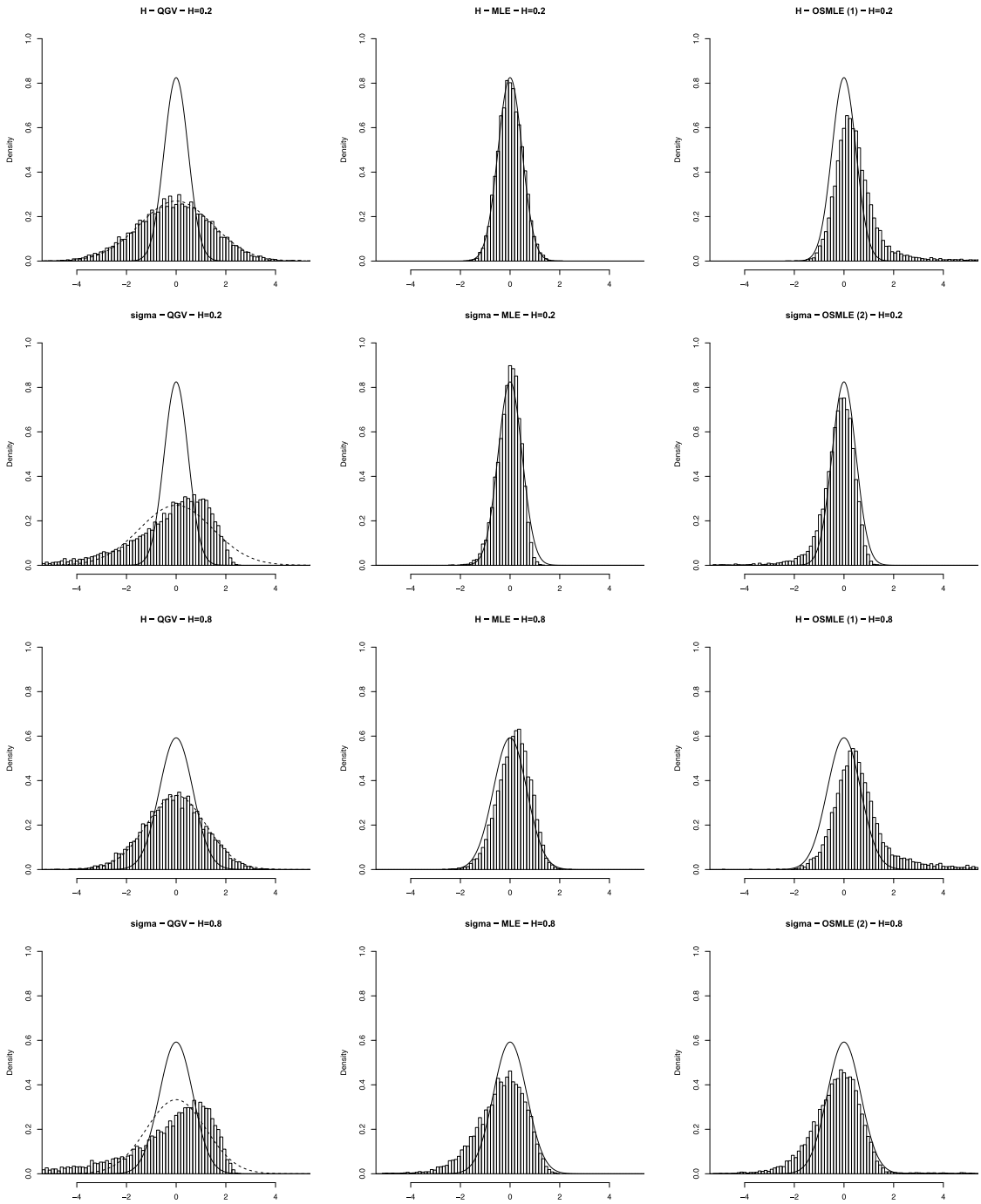

FIGURE 1. The renormalized statistical errors $\sqrt{n}(\widetilde{H}_n - H)$ and $\frac{\sqrt{n}}{\sigma \log \Delta_n}(\tilde{\sigma}_n - \sigma)$ on the left (QGV estimator), $\sqrt{n}(\widehat{H}_n - H)$ and $\frac{\sqrt{n}}{\sigma \log \Delta_n}(\widehat{\sigma}_n - \sigma)$ in the middle (MLE) and $\sqrt{n}(\overline{H}_n - H)$ and $\frac{\sqrt{n}}{\sigma \log \Delta_n}(\overline{\sigma}_n - \sigma)$ on the right (OSMLE) for $n = 2^8$. The two first lines present 10000 Monte-Carlo simulations for short-memory setting $H = 0.2$ and the two last lines for long-memory setting $H = 0.8$. Efficient asymptotic variance is superimposed (*plain line*). The theoretical QGV variance has been also added (*dashed line*).

Acknowledgements. The authors would like to thank M. Fukasawa and T. Takabatake for providing them their script of Paxson's method to evaluate the Fisher information. They also thank F. Karamé for fruitful discussions throughout the IRA-PANORisk seminar and the two anonymous referees for their valuable comments which improved the paper.

References

[1] A. Aït-Sahalia and J. Jacod, Fisher's information for discretely sampled Lévy processes. *Econometrica* **76** (2008) 727–761.
[2] C. Berzin and J. Léon, Estimation in models driven by fractional Brownian motion. *Ann. Inst. Henri Poincaré - Probab. Statist.* **44** (2008) 191–213.
[3] A. Brouste and M. Fukasawa, Local asymptotic normality property for fractional Gaussian noise under high-frequency observations. *Ann. Statist.* **46** (2018) 2045–2061.
[4] A. Brouste and H. Masuda, Efficient estimation of stable Lévy process with symmetric jumps. *Statist. Inference Stoch. Processes* **21** (2018) 289–307.
[5] J.F. Cœurjolly, Simulation and identification of the fractional Brownian motion: a bibliographical and comparative study. *J. Statist. Softw.* **5** (2000) 1–53.
[6] S. Cohen, F. Gamboa, C. Lacaux and J.-M. Loubes, LAN property for some fractional type Brownian motion. *ALEA: Latin Am. J. Probab. Math. Statist.* **10** (2013) 91–106.
[7] R. Dahlhaus, Efficient parameter estimation for self-similar processes. *Ann. Statist.* **17** (1989) 1749–1766.
[8] R. Dahlhaus, Correction efficient parameter estimation for self-similar processes. *Ann. Statist.* **34** (2006) 1045–1047.
[9] P. Flandrin, Wavelet analysis and synthesis of fractional Brownian motion. *IEEE Trans. Inform. Theory* **38** (1992) 910–917.
[10] R. Fox and M. Taqqu, Large-sample properties of parameter estimates for strongly dependent stationary Gaussian time series. *Ann. Statist.* **14** (1986) 517–532.
[11] M. Fukasawa and T. Takabatake, Asymptotically efficient estimators for self-similar stationary Gaussian noises under high frequency observations. *Bernoulli* **25** (2019) 1870–1900.
[12] X. Guyon and J. Léon, Convergence en loi des H-variations d'un processus gaussien stationnaire sur R. *Ann. Inst. Henri Poincaré - Probab. Statist. B* **25** (1989) 265–282.
[13] J. Hájek, Local asymptotic minimax and admissibility in estimation, in *Proceedings of 6th Berkeley Symposium on Math. Statist. Prob.* (1972) 175–194.
[14] I. Ibragimov and R. Has'minski, Statistical Estimation: Asymptotic Theory. Springer-Verlag (1981).
[15] J. Istas and G. Lang, Quadratic variations and estimation of the local Hölder index of a Gaussian process. *Ann. Instit. Henri Poincaré - Probab. Statist. B* **33** (1997) 407–436.
[16] R. Kawai, Fisher information for fractional Brownian motion under high-frequency discrete sampling. *Commun. Statist. - Theory Methods* **42** (2013) 1628–1636.
[17] Y. Kutoyants and A. Motrunich, On multi-step MLE-process for Markov sequences. *Metrika* **79** (2016) 705–724.
[18] L. Le Cam, Limits of experiments, in *Proceedings of the 6th Berkeley Symposium* (1972) 245–261.
[19] L. Le Cam, On the asymptotic theory of estimation and testing hypothesis, in *Proceedings of the 3rd Berkeley Symposium* (1956) 355–368.
[20] S. Mazur, D. Otryakhin and M. Podolskij, Estimation of the linear fractional stable motion. *Bernoulli* **26** (2020) 226–252.
[21] R Core Team, R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing Vienna, Austria. Available from: https://www.R-project.org/ (2016)
[22] G. Samorodnitsky and M. Taqqu, Stable Non-Gaussian Random Processes. Chapman & Hall (2000).
[23] T. Sweeting, Uniform asymptotic normality of the maximum likelihood estimator. *Ann. Statist.* **8** (1980) 1375–1381.
[24] A. van der Vaart, Asymptotic Statistics. Cambridge University Press (1998).