Adaptive estimator of the memory parameter and goodness-of-fit test using a multidimensional increment ratio statistic

Jean-Marc Bardet, Béchir Dola

Adaptive estimator of the memory parameter and goodness-of-fit test using a multidimensional increment ratio statistic

Jean-Marc Bardet and Béchir Dola

`bardet@univ-paris1.fr` , `bechir.dola@malix.univ-paris1.fr`

SAMM, Université Panthéon-Sorbonne (Paris I), 90 rue de Tolbiac, 75013 Paris, FRANCE

October 1, 2010

Abstract

The Increment Ratio (IR) statistic was first defined and studied in Surgailis *et al.* (2008) for estimating the long-memory parameter either of a stationary or an increment stationary Gaussian process. Three extensions, for stationary processes only, are proposed here. Firstly, a multidimensional central limit theorem is established for a vector composed by several IR statistics. Secondly, a χ^2-type test is deduced from this theorem. Finally, adaptive versions of the estimator and the test are studied in a general semiparametric frame. The adaptive estimator of the long-memory parameter is proved to follow an oracle property. Simulations attest of the accuracies and robustness of the estimator and test, even in the non Gaussian case.

Keywords: Long-memory Gaussian processes; goodness-of-fit test; estimation of the memory parameter; minimax adaptive estimator.

1 Introduction

After almost thirty years of intensive and numerous studies, the long-memory processes are now important particular cases of time series (see for instance the book edited by Doukhan *et al*, 2003). The most famous long-memory stationary time series are the fractional Gaussian noises (fGn) with Hurst parameter H and FARIMA(p, d, q) processes. For both these time series, the spectral density f in 0 follows a power law: $f(\lambda) \sim C\,\lambda^{-2d}$ where $H = d + 1/2$ in the case of the fGn. In the case of long memory process $d \in (0, 1/2)$ but a natural expansion to $d \in (-1/2, 0]$ (short memory) implied that d can be considered as a more general memory parameter.

There are a lot of statistical results relative to the estimation of this memory parameter d. First and main results in this direction are obtained for parametric models with the essential papers of Fox and Taqqu (1986) and Dahlhaus (1989) for Gaussian time series, Giraitis and Surgailis (1990) for linear processes and Giraitis and Taqqu (1999) for non linear functions of Gaussian processes.

However and especially for numerical applications, parametric estimators are not really robust and can induce no consistent estimations. Thus, the research is now rather focused on semiparametric estimators of the memory parameter. Different approaches were considered: the famous and seminal R/S statistic (see Hurst, 1951), the log-periodogram estimator (see Geweke and Porter-Hudack, 1983, notably improved by

Moulines and Soulier, 2003), the local Whittle estimator (see Robinson, 1995) or the wavelet based estimator (see Veitch *et al*, 2003, Moulines *et al*, 2007 or Bardet *et al*, 2008). All these estimators require the choice of an auxiliary parameter (frequency bandwidth, scales,...) but adaptive versions of original estimators are generally built for avoiding this choice. In a general semiparametric frame, Giraitis *et al* (1997) obtained the asymptotic lower bound for the minimax risk of estimating d, expressed as a function of the second order parameter of the spectral density expansion around 0. Thus, several adaptive semiparametric are proved to follow an oracle property up to multiplicative logarithm term. But simulations (see for instance Bardet *et al*, 2003 or 2008) show that the most accurate estimators are local Whittle, global log-periodogram or wavelet based estimators.

In this paper, we consider the IR estimator of long-memory parameter (see its definition in the next Section) for Gaussian time series introduced in Surgailis *et al.* (2008) and propose three extensions to it in this paper. Firstly, a multivariate central limit theorem is established under more general condition on the spectral density than in Surgailis *et al.* (2008) for a vector of IR statistics with different "windows" (see Section 2). Secondly, this multivariate result allows us to define an adaptive estimator of the memory parameter d based on IR statistics: an "optimal" window is automatically computed (see Section 3). This notably improves the results of Surgailis *et al.* (2008) in which the choice of m is either theoretical (and cannot be applied to data) or guided by empirical rules without justifications. Thirdly, an adaptive goodness-of-fit test is deduced and its convergence to a chi-square distribution is proved (see Section 3). This also allows us to propose a test of long-memory in the case where $d > 0$ which is more significant than a test on the value of d.
In Section 4, several Monte Carlo simulations are realized for optimizing the adaptive estimator and exhibiting the theoretical results. Then some numerical comparisons are made with the 3 semiparametric estimators previously mentioned (local Whittle, global log-periodogram and wavelet based estimators) and the results are even better than the theory seems to indicate: both in terms of convergence rate than that of the robustness (notably in case of trend or seasonal component), the adaptive IR estimator provides efficient results. Finally, all the proofs are grouped in Section 5.

2 The multidimensional increment ratio statistic and its statistical applications

Let $X = (X_k)_{k\in\mathbb{N}}$ be a Gaussian time series satisfying the following Assumption $S(d,\beta)$:

Assumption $S(d,\beta)$**:** *There exist $\varepsilon > 0$, $c_0 > 0$, $c_0' > 0$ and $c_1 \in \mathbb{R}$ such that $X = (X_t)_{t\in\mathbb{Z}}$ is a stationary Gaussian time series having a spectral density f satisfying for all $\lambda \in (-\pi,0)\cup(0,\pi)$*

$$f(\lambda) = c_0|\lambda|^{-2d} + c_1|\lambda|^{-2d+\beta} + O\big(|\lambda|^{-2d+\beta+\varepsilon}\big) \quad \text{and} \quad |f'(\lambda)| \le c_0'\,\lambda^{-2d-1}. \tag{2.1}$$

Remark 1. *Note that here we only consider stationary processes. However, as it was already done in Surgailis* et al. *(2008), it could be possible,* mutatis mutandis*, to extend our results to the case of processes having stationary increments. A forthcoming paper will be devoted to this extension and to its application to a test of stationarity of the process.*

Let $(X_1,\cdots,X_N)$ be a path of X. For $m\in\mathbb{N}^*$, define the random variable $IR_N(m)$ such as

$$IR_N(m) := \frac{1}{N-3m}\sum_{k=0}^{N-3m-1}\frac{\big|\big(\sum_{t=k+1}^{k+m}X_{t+m}-\sum_{t=k+1}^{k+m}X_t\big)+\big(\sum_{t=k+m+1}^{k+2m}X_{t+m}-\sum_{t=k+m+1}^{k+2m}X_t\big)\big|}{\big|\big(\sum_{t=k+1}^{k+m}X_{t+m}-\sum_{t=k+1}^{k+m}X_t\big)\big|+\big|\big(\sum_{t=k+m+1}^{k+2m}X_{t+m}-\sum_{t=k+m+1}^{k+2m}X_t\big)\big|}.$$

From Surgailis *et al.* (2008), with m such that $N/m\to\infty$ and $m\to\infty$,

$$\sqrt{\frac{N}{m}}\big(IR_N(m)-\mathrm{E}IR_N(m)\big) \xrightarrow[N\to\infty]{\mathcal{L}} \mathcal{N}(0,\sigma^2(d)),$$

where

$$\sigma^2(d) := 2\int_0^\infty \mathrm{Cov}\Big(\frac{|Z_d(0)+Z_d(1)|}{|Z_d(0)|+|Z_d(1)|}, \frac{|Z_d(\tau)+Z_d(\tau+1)|}{|Z_d(\tau)|+|Z_d(\tau+1)|}\Big)d\tau \quad (2.2)$$

$$\text{and} \quad Z_d(\tau) := \frac{1}{\sqrt{|4^{d+0.5}-4|}}\big(B_{d+0.5}(\tau+2) - 2\,B_{d+0.5}(\tau+1) + B_{d+0.5}(\tau)\big) \quad (2.3)$$

with B_H a standardized fractional Brownian motion with Hurst parameter $H \in (0,1)$.

Now, let $m_j = j\,m$, $j = 1,\cdots,p$ with $p \in \mathbb{N}^*$, and define the random vector $(IR_N(j\,m))_{1\le j\le p}$. We can establish a multidimensional central limit theorem satisfied by $(IR_N(j\,m))_{1\le j\le p}$:

Property 2.1. *Assume that Assumption $S(d,\beta)$ holds with $-0.5<d<0.5$ and $\beta>0$. Then*

$$\sqrt{\frac{N}{m}}\Big(IR_N(j\,m) - \mathrm{E}\big[IR_N(j\,m)\big]\Big)_{1\le j\le p} \xrightarrow[{[N/m]\wedge m\to\infty}]{\mathcal{L}} \mathcal{N}(0,\Gamma_p(d)) \quad (2.4)$$

with $\Gamma_p(d) = (\sigma_{i,j}(d))_{1\le i,j\le p}$ and $\sigma_{i,j}(d)$ defined in (5.3).

The proof of this property as well as all the other proofs are given in Appendix.
As in Surgailis *et al.* (2008), for $r\in(-1,1)$, define the function $\Lambda(r)$ by

$$\Lambda(r) := \frac{2}{\pi}\,\arctan\sqrt{\frac{1+r}{1-r}} + \frac{1}{\pi}\sqrt{\frac{1+r}{1-r}}\log\Big(\frac{2}{1+r}\Big). \quad (2.5)$$

and for $d\in(-0.5,1.5)$ the function $\Lambda_0(d)$ defined by

$$\Lambda_0(d) := \Lambda(\rho(d)) \quad \text{where} \quad \rho(d) := \frac{4^{d+1.5}-9^{d+0.5}-7}{2(4-4^{d+0.5})}. \quad (2.6)$$

The function $d\in(-0.5,1.5)\to\Lambda_0(d)$ is a $\mathcal{C}^\infty$ increasing function. Thus, using an expansion of $\mathrm{E}\big[IR_N(m)\big]$ and the Delta-method, we obtain:

Theorem 1. *Let $\widehat{d}_N(j\,m) := \Lambda_0^{-1}\big(IR_N(j\,m)\big)$. Assume that Assumption $S(d,\beta)$ holds with $-0.5<d<0.5$ and $\beta>0$. Then if $m\sim C\,N^\alpha$ with $C>0$ and $(1+2\beta)^{-1}\vee(4d+3)^{-1}<\alpha<1$ then*

$$\sqrt{\frac{N}{m}}\Big(\widehat{d}_N(j\,m) - d\Big)_{1\le j\le p} \xrightarrow[N\to\infty]{\mathcal{L}} \mathcal{N}\Big(0,(\Lambda_0'(d))^{-2}\,\Gamma_p(d)\Big). \quad (2.7)$$

Remark 2. *If $\beta<2d+1$, the estimator $\widehat{d}_N(m)$ is a semiparametric estimator of d and its asymptotic mean square error can be minimized with an appropriate sequence (m_N) reaching the well-known minimax rate of convergence for memory parameter d in this semiparametric setting (see for instance Giraitis* et al.*, 1997). Indeed, under Assumption $S(d,\beta)$ with $d\in(-0.5,0.5)$ and $\beta>0$ and if $m_N=[N^{1/(1+2\beta)}]$, then the estimator $\widehat{d}_N(m_N)$ is rate optimal in the minimax sense,* i.e.

$$\limsup_{N\to\infty}\sup_{d\in(-0.5,0.5)}\sup_{f\in S(d,\beta)} N^{\frac{2\beta}{1+2\beta}}\cdot\mathrm{E}[(\widehat{d}_N(m_N)-d)^2]<\infty.$$

From the multidimensional CLT (2.7) a pseudo-generalized least square estimation (LSE) of d is possible by defining the following matrix:

$$\widehat{\Sigma}_N(m) := (\Lambda_0'(\widehat{d}_N(m))^{-2}\,\Gamma_p(\widehat{d}_N(m)). \quad (2.8)$$

Since the function $d\in(-0.5,1.5)\mapsto\sigma(d)/\Lambda'(d)$ is $\mathcal{C}^\infty$ it is obvious that under assumptions of Theorem 1 then

$$\widehat{\Sigma}_N(m) \xrightarrow[N\to\infty]{\mathcal{P}} (\Lambda_0'(d))^{-2}\,\Gamma_p(d).$$

Then with the vector $J_p := (1)_{1\le j\le p}$ and denoting J_p' its transpose, the pseudo-generalized LSE of d is:

$$\widetilde{d}_N(m) := \big(J_p'\big(\widehat{\Sigma}_N(m)\big)^{-1}J_p\big)^{-1}\,J_p'\,\big(\widehat{\Sigma}_N(m)\big)^{-1}\big(\widehat{d}_N(m_i)\big)_{1\le i\le p}$$

It is well known (Gauss-Markov Theorem) that the Mean Square Error (MSE) of $\widetilde{d}_N(m)$ is smaller or equal than all the MSEs of $\widehat{d}_N(jm)$, $j = 1, \ldots, p$. Hence, we obtain:

$$\sqrt{\frac{N}{m}}\big(\widetilde{d}_N(m) - d\big) \underset{N\to\infty}{\overset{\mathcal{L}}{\longrightarrow}} \mathcal{N}\Big(0\,,\,\Lambda_0'(d)^{-2}\left(J_p'\,\Gamma_p^{-1}(d)J_p\right)^{-1}\Big), \tag{2.9}$$

and $\Lambda_0'(d)^{-2}\big(J_p'\,\Gamma_p^{-1}(d)J_p\big)^{-1} \leq \Lambda_0'(d)^{-2}\sigma^2(d)$.
Now, a χ^2-type goodness-of-fit test deduced from the multidimensional CLT (2.7) can be defined by:

$$\widehat{T}_N(m) := \frac{N}{m}\big(\widetilde{d}_N(m) - \widehat{d}_N(j\,m)\big)'_{1\leq j\leq p}\big(\widehat{\Sigma}_N(m)\big)^{-1}\big(\widetilde{d}_N(m) - \widehat{d}_N(j\,m)\big)_{1\leq j\leq p}.$$

Then the following limit theorem can be deduced from Theorem 1:

Proposition 1. *Under the assumptions of Theorem 1 then:*

$$\widehat{T}_N(m) \underset{N\to\infty}{\overset{\mathcal{L}}{\longrightarrow}} \chi^2(p-1).$$

Note that this test is also a test of long memory when $d > 0$ and it is very simple to be implemented .

3 Adaptive versions of the estimator and goodness-of-fit test

Theorem 1 and Proposition 1 are interesting but they require the knowledge of β to be used (and therefore an appropriated choice of m). We suggest now a new procedure for a data-driven selection of an optimal sequence (m_N). For $d \in (-0.5, 1.5)$ define

$$Q_N(\alpha, d) := \big(\widehat{d}_N(j\,N^\alpha) - d\big)'_{1\leq j\leq p}\big(\widehat{\Sigma}_N(N^\alpha)\big)^{-1}\big(\widehat{d}_N(j\,N^\alpha) - d\big)_{1\leq j\leq p}. \tag{3.1}$$

$Q_N(\alpha, d)$ corresponds to the sum of the pseudo-generalized squared distance. From previous computations, it is obvious that for a fixed $\alpha \in (0, 1)$, Q is minimized by $\widetilde{d}_N(N^\alpha)$ and therefore for $0 < \alpha < 1$ define

$$\widehat{Q}_N(\alpha) := Q_N(\alpha, \widetilde{d}_N(N^\alpha)).$$

It remains to minimize $\widehat{Q}_N(\alpha)$ on $(0, 1)$. However, since $\widehat{\alpha}_N$ has to be obtained from numerical computations, the interval $(0, 1)$ can be discretized as follows,

$$\widehat{\alpha}_N \in \mathcal{A}_N = \Big\{\frac{2}{\log N}\,,\,\frac{3}{\log N}\,,\ldots,\,\frac{\log[N/p]}{\log N}\Big\}.$$

Hence, if $\alpha \in \mathcal{A}_N$, it exists $k \in \{2, 3, \ldots, \log[N/p]\}$ such that $k = \alpha\,\log N$. Consequently, define $\widehat{\alpha}_N$ by

$$\widehat{Q}_N(\widehat{\alpha}_N) := \min_{\alpha\in\mathcal{A}_N}\widehat{Q}_N(\alpha).$$

Remark 3. *The choice of this set of discretization $\mathcal{A}_N$ is implied to proof the consistency of $\widehat{\alpha}_N$. If the interval $(0, 1)$ is stepped in N^c points, with $c > 0$, the used proof cannot attest this consistency. However $\log N$ may be replaced in the previous expression of $\mathcal{A}_N$ by any negligible function of N compared to functions N^c with $c > 0$ (for instance, $(\log N)^a$ or $a\log N$ can be used).*

From the central limit theorem (2.7) one deduces the following :

Proposition 2. *Assume that Assumption $S(d, \beta)$ holds with $-0.5 < d < 0.5$ and $\beta > 0$. Moreover, if $\beta > 2d + 1$, suppose that c_0, c_1, c_2, d, β and ε is such that Condition (5.10) or (5.11) holds. Then,*

$$\widehat{\alpha}_N \underset{N\to\infty}{\overset{\mathcal{P}}{\longrightarrow}} \alpha^* = \frac{1}{(1 + 2\beta)\wedge(4d + 3)}.$$

From a straightforward application of the proof of Proposition 2, the asymptotic behavior of $\widehat{a}_N$ can be specified, that is,

$$\Pr\Big(\frac{N^\alpha}{(\log N)^\lambda} \leq N^{\widehat{\alpha}_N} \leq N^\alpha \cdot (\log N)^\mu\Big) \xrightarrow[N\to\infty]{\mathcal{P}} 1, \tag{3.2}$$

for all positive real numbers λ and μ such that $\lambda > \frac{2\alpha^*}{(p-2)(1-\alpha^*)}$ and $\mu > \frac{12}{p-2}$. Consequently, the selected window $\widehat{m}_N$ is asymptotically equal to N^α up to a logarithm factor.

Finally, Proposition 2 can be used to define an adaptive estimator of d. First, define the straightforward estimator

$$\widehat{\widetilde{d}_N} = \widetilde{d}_N(N^{\alpha_N}),$$

which should minimize the mean square error using $\widehat{\alpha}_N$. However, the estimator $\widehat{\widetilde{d}_N}$ does not satisfy a CLT since $\Pr(\widehat{\alpha}_N \leq \alpha) > 0$ and therefore it can not be asserted that $\mathrm{E}(\sqrt{N/\widehat{a}_N}(\widehat{\widetilde{d}_N} - d)) = 0$. To establish a CLT satisfied by an adaptive estimator $\widetilde{\widetilde{d}_N}$ of d, an adaptive scale sequence $(\widetilde{m}_N) = (N^{\widetilde{\alpha}_N})$ has to be defined to ensure $\Pr(\widetilde{\alpha}_N \leq \alpha) \underset{n\to\infty}{\longrightarrow} 0$. The following theorem provides the asymptotic behavior of such an estimator,

Theorem 2. *Define,*

$$\widetilde{\alpha}_N = \widehat{\alpha}_N + \frac{6\,\widehat{\alpha}_N}{(p-2)(1-\widehat{\alpha}_N)} \cdot \frac{\log\log N}{\log N} \quad \textit{and} \quad \widetilde{d}_N = \widehat{d}_N(N^{\widetilde{\alpha}_N}).$$

Then, under assumptions of Proposition 2,

$$\sqrt{\frac{N}{N^{\widetilde{\alpha}_N}}}\big(\widetilde{\widetilde{d}_N} - d\big) \xrightarrow[N\to\infty]{D} \mathcal{N}\Big(0\,;\,\Lambda_0'(d)^{-2}\left(J_p'\,\Gamma_p^{-1}(d)J_p\right)^{-1}\Big). \tag{3.3}$$

Moreover, if $\beta \leq 2d+1$, $\quad \forall \rho > \dfrac{2(1+3\beta)}{(p-2)\beta}, \quad \dfrac{N^{\frac{\beta}{1+2\beta}}}{(\log N)^\rho} \cdot \big|\widetilde{\widetilde{d}_N} - d\big| \xrightarrow[N\to\infty]{\mathcal{P}} 0.$

Remark 4. *When* $\beta \leq 2d+1$, *both the adaptive estimators* $\widehat{\widetilde{d}_N}$ *and* $\widetilde{\widetilde{d}_N}$ *converge to* d *with a rate of convergence rate equal to the minimax rate of convergence* $N^{\frac{\beta}{1+2\beta}}$ *up to a logarithm factor (this result being classical within this semiparametric framework). Thus there exist* $\ell < 0$ *and* $\ell' < 0$ *such that*

$$N^{\frac{2\beta}{1+2\beta}}(\log N)^\ell \mathrm{E}(\widehat{\widetilde{d}_N} - d)^2 < \infty \quad \textit{and} \quad N^{\frac{2\beta}{1+2\beta}}(\log N)^{\ell'} \mathrm{E}(\widetilde{\widetilde{d}_N} - d)^2 < \infty$$

Therefore $\widehat{\widetilde{d}_N}$ *and* $\widetilde{\widetilde{d}_N}$ *satisfy an oracle property for the considered semiparametric model.*
If $\beta > 2d+1$, *the estimator is not rate optimal. However, simulations (see the following Section) will show that the rates of convergence of the adaptive estimators* $\widehat{\widetilde{d}_N}$ *and* $\widetilde{\widetilde{d}_N}$ *can be better than the one of the best known rate optimal estimators (local Whittle or global log-periodogram estimators).*

Moreover an adaptive version of the previous test of long-memory can be derived. Thus define

$$\widetilde{T}_N := \widehat{T}_N(N^{\widetilde{\alpha}_N}). \tag{3.4}$$

Then,

Proposition 3. *Under assumptions of Proposition 2,*

$$\widetilde{T}_N \xrightarrow[N\to\infty]{\mathcal{L}} \chi^2(p-1).$$

4 Simulations and Monte-Carlo experiments

In the sequel, the numerical properties (consistency, robustness, choice of the parameter p) of $\widehat{\widetilde{d}_N}$ are investigated. Then the simulation results of the estimator $\widehat{\widetilde{d}_N}$ are compared to those obtained with the best known semiparametric long-memory estimators.

Remark 5. *Performance of* $\widetilde{\widetilde{d}_N}$ *(the second estimator defined previously) are very close (just a little worse) to the ones of* $\widehat{\widetilde{d}_N}$. *Thus, in the sequel we will only consider* $\widehat{\widetilde{d}_N}$.

Remark 6. *Note that all the softwares (in Matlab language) used in this Section are available with a free access on* `http://samm.univ-paris1.fr/-Jean-Marc-Bardet`.

To begin with, the simulation conditions have to be specified. The results are obtained from 100 generated independent samples of each process belonging to the following "benchmark". The concrete procedures of generation of these processes are obtained from the circulant matrix method, as detailed in Doukhan *et al.* (2003). The simulations are realized for different values of d, N and processes which satisfy Assumption $S(d,\beta)$:

1. the fractional Gaussian noise (fGn) of parameter $H = d+1/2$ (for $-0.5 < d < 0.5$) and $\sigma^2 = 1$. A fGn is such that Assumption $S(d,2)$ holds;

2. the FARIMA$[p,d,q]$ process with parameter d such that $d \in (-0.5, 0.5)$, the innovation variance σ^2 satisfying $\sigma^2 = 1$ and $p, q \in \mathbb{N}$. A FARIMA$[p,d,q]$ process is such that Assumption $S(d,2)$ holds;

3. the Gaussian stationary process $X^{(d,\beta)}$, such that its spectral density is
$$f_3(\lambda) = \frac{1}{\lambda^{2d}}(1+\lambda^\beta) \quad \text{for } \lambda \in [-\pi, 0(\cup]0,\pi], \tag{4.1}$$
with $d \in (-0.5, 0.5)$ and $\beta \in (0,\infty)$. Therefore the spectral density f_3 is such that Assumption $S(d,\beta)$ holds.

A "benchmark" which will be considered in the sequel consists of the following particular cases of these processes for $d = -0.4, -0.2, 0, 0.2, 0.4$:

- fGn processes with parameters $H = d + 1/2$;
- FARIMA$[0,d,0]$ processes standard Gaussian innovations;
- FARIMA$[1,d,1]$ processes with standard Gaussian innovations and AR coefficient $\phi = -0.3$ and MA coefficient $\phi = 0.7$;
- $X^{(d,\beta)}$ Gaussian processes with $\beta = 1$.

4.1 Application of the IR estimator and tests applied to generated data

Choice of the parameter p: This parameter is important to estimate the "beginning" of the linear part of the graph drawn by points $(i, IR(im))_i$. On the one hand, if p is a too small a number (for instance $p = 3$), another small linear part of this graph (even before the "true" beginning N^{α^*}) may be chosen. On the other hand, if p is a too large a number (for instance $p = 50$ for $N = 1000$), the estimator $\widehat{\alpha}_N$ will certainly satisfy $\widehat{\alpha}_N < \alpha^*$ since it will not be possible to consider p different windows larger than N^{α^*}. Moreover, it is possible that a "good" choice of p depends on the "flatness" of the spectral density f, *i.e.* on β. We have proceeded to simulations for several values of p (and N and d). Only $\sqrt{MSE}$ of estimators are presented. The results are specified in Table 1.

Conclusions from Table 1: it is clear that $\widehat{\widetilde{d}}_N$ converges to d for the four processes, the faster for fGn and FARIMA$(0, d, 0)$. The optimal choice of p seems to depend on N for the four processes: $\widehat{p} = 10$ for $N = 10^3$, $\widehat{p} = 15$ for $N = 10^4$ and $\widehat{p} \in [15, 20]$ for $N = 10^5$. We will now adopt the choice $\widehat{p} = [1.5 \log(N)]$ reflecting these results.

Concerning the adaptive choice of m, the main point to be remarked is that the smoother the spectral density the smaller m; thus $\widehat{m}$ is smaller for a trajectory of a fGn or a FARIMA$(0, d, 0)$ than for a trajectory of a FARIMA$(1, d, 1)$ or $X^{(d,1)}$. The choice of p does not appear to significantly affect the value of $\widehat{m}$. Moreover, more detailed results show that the larger d included in $(-0.5, 0.5)$ the smaller $\widehat{m}$: for instance, for the fGn, $N = 10^4$ and $m = 15$, the mean of $\widehat{m}$ is respectively equal to 23.9, 8.3, 4.5, 4.2 and 3.8 for d respectively equal to -0.4, -0.2, 0, 0.2 and 0.4. This phenomena can be deduced from the theoretical study.

Finally, concerning the goodness-of-fit test, we remark that it is too conservative for $p = 5$ or 10 but close to the expected results for $m = 15$ and 20, especially for FARIMA$(1, d, 1)$ or $X^{(d,1)}$.

Asymptotic distributions of the estimator and test: Figure 2 provides the density estimations of $\widehat{\widetilde{d}}_N$ and $\widetilde{T}_N$ for 100 independent samples of FARIMA$(1, d, 1)$ processes with $d = -0.2$ and $N = 10^5$ for $p = 20$. The goodness-of-fit to the theoretical asymptotic distributions (respectively Gaussian and chi-square) is satisfying. However, following d and the studied process, a small bias can appear and can degrade this goodness-of-fit.

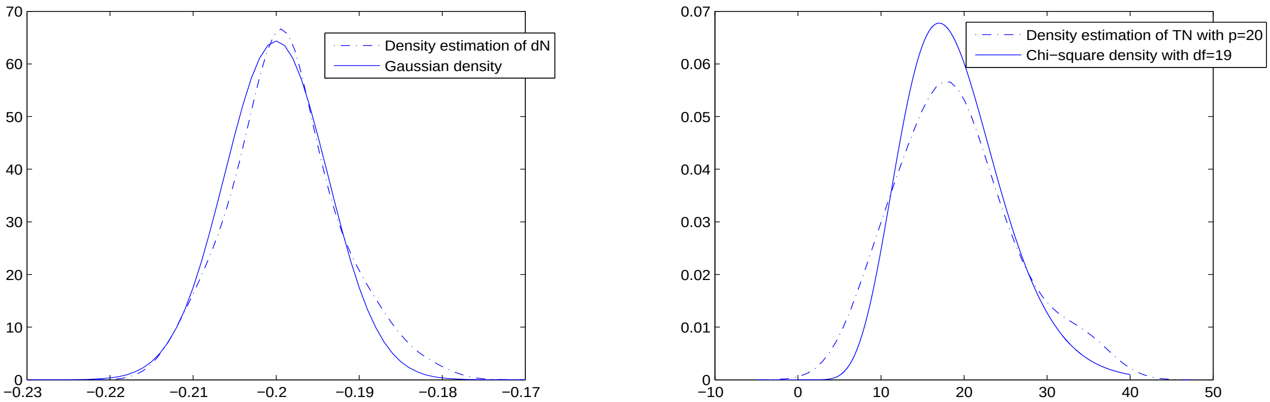

Figure 1: *Density estimations and corresponding theoretical densities of $\widehat{\widetilde{d}}_N$ and $\widetilde{T}_N$ for 100 samples of FARIMA$(1, d, 1)$ with $d = -0.2$ for $N = 10^5$ and $p = 20$.*

4.2 Comparison with other adaptive semiparametric estimator of the memory parameter

Consistency of semiparametric estimators: Here we consider the previous "benchmark" and apply the estimator $\widehat{\widetilde{d}}_N$ and 3 other semiparametric estimators of d known for their accuracies are considered:

- $\widehat{d}_{MS}$ is the adaptive global log-periodogram estimator introduced by Moulines and Soulier (1998, 2003), also called FEXP estimator, with bias-variance balance parameter $\kappa = 2$;

- $\widehat{d}_R$ is the local Whittle estimator introduced by Robinson (1995). The trimming parameter is $m = N/30$;
- $\widehat{d}_W$ is an adaptive wavelet based estimator introduced in Bardet *et al.* (2008) using a Lemarie-Meyer type wavelet (another similar choice could be the adaptive wavelet estimator introduced in Veitch *et al.*, 2003, using a Daubechie's wavelet, but its robustness property are quite less interesting).
- $\widetilde{\widehat{d}}_N$ defined previously with $p \sim [1.5 * \log(n)]$.

Simulation results are reported in Table 2.

Conclusions from Table 2: The adaptive IR estimator $\widetilde{\widehat{d}}_N$ numerically shows a convincing convergence rate with respect to the other estimators. Both the "spectral" estimator $\widehat{d}_R$ and $\widehat{d}_{MS}$ provide more stable results that do not depend very much on d and the process, while the wavelet based estimator $\widehat{d}_W$ and $\widetilde{\widehat{d}}_N$ are more sensible to the flatness of the spectral density. But, especially in the long memory case ($d > 0$) and smooth processes (fGn and FARIMA$(0, d, 0)$), $\widetilde{\widehat{d}}_N$ is a very accurate semiparametric estimator and is globally more efficient than the other estimators.

Robustness of the different semiparametric estimators: To conclude with the numerical properties of the estimators, five different processes not satisfying Assumption $S(d, \beta)$ are considered:

- a FARIMA$(0, d, 0)$ process with innovations satisfying a uniform law;
- a FARIMA$(0, d, 0)$ process with innovations satisfying a symmetric Burr distribution with cumulative distribution function $F(x) = 1 - \frac{1}{1+x^2}$ for $x \geq 0$ (and therefore $\mathrm{E}|X_i|^2 = \infty$ but $\mathrm{E}|X_i| < \infty$);
- a FARIMA$(0, d, 0)$ process with innovations satisfying a Cauchy distribution (thus $\mathrm{E}|X_i| = \infty$);
- a Gaussian stationary process (denoted $P4$) with a spectral density $f(\lambda) = ||\lambda| - \pi/2|^{-2d}$ for all $\lambda \in [-\pi, \pi] \setminus \{-\pi/2, \pi/2\}$. The local behavior of f in 0 is $f(|\lambda|) \sim (\pi/2)^{-2d} |\lambda|^{-2d}$ with $d = 0$, but the smoothness condition for f in Assumption $S(0, \beta)$ is not satisfied.
- a trended fGn with parameter $H = d + 0.5$ and an additive linear trend;
- a fGn ($H = d + 0.5$) with an additive linear trend and an additive sinusoidal seasonal component of period $T = 12$.

The results of these simulations are given in Table 3.

Conclusions from Table 3: The main advantages of $\widehat{d}_W$ and $\widetilde{\widehat{d}}_N$ with respect to $\widehat{d}_{MS}$ and $\widehat{d}_R$ is exhibited in this table: they are robust with respect to smooth trends. It has already been observed in Bruzaite and Vaiciulis (2008) for IR statistic (and even for certain discontinuous trends). Both those estimators are also robust with respect to seasonal component and this robustness would have been improved if we had chosen m (or scales) as a multiple of the period (which is generally known).
The second good surprise of these simulations is that the adaptive IR estimator $\widetilde{\widehat{d}}_N$ is also consistent for non Gaussian distributions even if the function Λ in (2.5) is typically obtained for a Gaussian distribution. This can be explained by the fact that the distribution of a FARIMA$(0, d, 0)$ with uniform or almost $\mathbb{L}^2$-innovations is close to be a Gaussian distribution (from central limit theorems). This property is no more valid for Cauchy innovations (without expectation) and the results of simulations prove this... Note that a recent study of IR statistic for heavy tailed processes have been done in Vaiciulis (2009).

5 Proofs

Proof of Property 2.1. We proceed in two steps.
Step 1: First, we compute the limit of $\frac{N}{m}\,\mathrm{Cov}\big(IR_N(jm), IR_N(j'm)\big)$ when N, m and $N/m \to \infty$. As in Surgailis *et al* (2008), define also for all $j = 1, \cdots, p$ and $k = 1, \cdots, N - 3m_j$ (with $m_j = jm$):

$$Y_j(k) \;:=\; \frac{1}{V_j} \sum_{t=k+1}^{k+m_j} (X_{t+m_j} - X_t) \;\; , \text{ with } \; V_j^2 := \mathrm{E}\Big[\sum_{t=k+1}^{k+m_j} (X_{t+m_j} - X_t)^2\Big] \tag{5.1}$$

$$\text{and} \quad \eta_j(k) \;:=\; \frac{|Y_j(k) + Y_j(k+m_j)|}{|Y_j(k)| + |Y_j(k+m_j)|}. \tag{5.2}$$

Note that $Y_j(k) \sim \mathcal{N}(0,1)$ for any k and j and

$$IR_N(m_j) = \frac{1}{N - 3m_j} \sum_{k=0}^{N-3m_j-1} \eta_j(k) \quad \text{for all } j = 1, \cdots p.$$

$$\begin{aligned}
\mathrm{Cov}(IR_N(m_j), IR_N(m_{j'})) &= \frac{1}{N-3m_j}\,\frac{1}{N-3m_{j'}} \sum_{k=0}^{N-3m_j-1} \sum_{k'=0}^{N-3m_{j'}-1} \mathrm{Cov}(\eta_j(k), \eta_{j'}(k'))) \\
&= \frac{1}{(\frac{N}{m_j}-3)(\frac{N}{m_{j'}}-3)} \int_{\tau=0}^{\frac{N-1}{m_j}-3} \int_{\tau'=0}^{\frac{N-1}{m_{j'}}-3} \mathrm{Cov}(\eta_j([m_j\tau]), \eta_{j'}([m_{j'}\tau']))) \, d\tau \, d\tau'.
\end{aligned}$$

Now according to (5.20) of the same article, with $\longrightarrow_{FDD}$ denoting the finite distribution convergence when $m \to \infty$,

$$Y_j([mj\tau]) \;\longrightarrow_{FDD}\; Z_d(j\,\tau) \quad \text{and} \quad Y_{j'}([mj'\tau']) \;\longrightarrow_{FDD}\; Z_d(j'\,\tau'),$$

where Z is defined in (2.3). As the function $\psi(x,y) = \frac{|x+y|}{|x|+|y|}$ is continue and bounded with $0 \le \psi(x,y) \le 1$ and since $\eta_j([mj\tau]) = \psi(Y_j[m_j\tau], Y_j[m_j(\tau+1)])$, then when $m \to \infty$:

$$\mathrm{Cov}\big(\eta_j([m_j\tau]), \eta_{j'}([m_{j'}\tau'])\big) \quad \underset{m\to\infty}{\longrightarrow} \quad \mathrm{Cov}\big(\psi(Z_d(j\,\tau), Z_d(j\,\tau+j)), \psi(Z_d(j'\,\tau'), Z_d(j'\,\tau'+j'))\big).$$

Thus, with $\gamma_d^{(j,j')}(t) = \mathrm{Cov}\big(\psi(Z_d(0), Z_d(j)), \psi(Z_d(t), Z_d(t+j'))\big)$ and the stationarity of the process Z_d, when N, m and $N/m \to \infty$,

$$\begin{aligned}
\frac{N}{m}\,\mathrm{Cov}(IR_N(jm), IR_N(j'm)) \quad &\sim \quad \frac{N}{m(\frac{N}{jm}-3)(\frac{N}{j'm}-3)} \\
&\qquad \times \int_0^{\frac{N-1}{jm}-3} \int_0^{\frac{N-1}{j'm}-3} \mathrm{Cov}\big(\psi(Z_d(j\,\tau), Z_d(j\,\tau+j)), \psi(Z_d(j'\,\tau'), Z_d(j'\,\tau'+j'))\big) d\tau d\tau' \\
&\sim \quad \frac{mN}{(N-3jm)(N-j'm)} \int_0^{\frac{N-1}{m}-3j} \int_0^{\frac{N-1}{m}-3j'} \gamma_d^{(j,j')}(s'-s)\, ds\, ds' \\
&\sim \quad \frac{m}{N} \int_{-\frac{N}{m}}^{\frac{N}{m}} \big(\frac{N}{m} - |u|\big)\, \gamma_d^{(j,j')}(u)\, du \\
&\longrightarrow \quad \int_{-\infty}^{\infty} \gamma_d^{(j,j')}(u)\, du = \sigma_{j,j'}(d)
\end{aligned} \tag{5.3}$$

(note that $\sigma_{j,j}(d) = j\,\sigma^2(d)$ for all $j \in \mathbb{N}$). This last limit is obtained, *mutatis mutandis*, from the relation (5.23) Surgailis *et al* (2008), and $\gamma_d^{(j,j')}(u) = C\,(u^{-2} \wedge 1)$. and therefore $\frac{m}{N} \int_{-\frac{N}{m}}^{\frac{N}{m}} |u|\, \gamma_d^{(j,j')}(u)\, du \underset{N,\, m,\, \frac{N}{m}\to\infty}{\longrightarrow} 0$. It achieves the first step of the proof.

Step 2: It remains to prove the multidimensional central limit theorem. Then consider a linear combination of $(IR_N(m_j))_{1\leq j\leq p}$, *i.e.* $\sum_{j=1}^p u_j\, IR_N(m_j)$ with $(u_1,\cdots,u_p)\in\mathbb{R}^p$. For ease of notation, we will restrict our purpose to $p=2$, with $m_i=r_im$ where $r_1\leq r_2$ are fixed positive integers. Then with the previous notations and following the notations and results of Theorem 2.5 of Surgailis *et al.* (2008):

$$u_1\, IR_N(r_1m)+u_2\, IR_N(r_2m)=u_1(\mathrm{E}IR_N(r_1m)+S_K(r_1m)+\widetilde{S}_K(r_1m)) \\ +u_2(\mathrm{E}IR_N(r_2m)+S_K(r_2m)+\widetilde{S}_K(r_2m)).$$

From (5.31) of Surgailis *et al.* (2008), we have $\widetilde{S}_K(m_1)=o(S_K(m_1))$ and $\widetilde{S}_K(m_2)=o(S_K(m_2))$ when $K\to\infty$ and from an Hermitian decomposition $(N/m)^{1/2}(u_1S_K(m_i)+u_2S_K(m_2))\to_D\mathcal{N}(0,\gamma_K^2)$ as N, m and $N/m\to\infty$ since the cumulants of $(N/m)^{1/2}(u_1S_K(m_i)+u_2S_K(m_2))$ of order greater or equal to 3 converge to 0 (since this result is proved for each $S_K(m_i)$). Moreover, from the previous computations, $\gamma_K^2\to(u_1^2\sigma_{r_1,r_1}(d)+2u_1u_2\sigma_{r_1,r_2}(d)+u_2^2\sigma_{r_2,r_2}(d))$ when $K\to\infty$. Therefore the multidimensional central limit theorem is established. □

Property 5.1. *Let X satisfy Assumption $S(d,\beta)$ with $-0.5<d<0.5$ and $\beta>0$. Then, there exists a constant $K(d,\beta)<0$ depending only on d and β such that*

$$\begin{aligned}\mathrm{E}\big[IR_N(m)\big] &= \Lambda(d)+K(d,\beta)\times m^{-\beta}+O\big(m^{-\beta-\varepsilon}+m^{-2d-1}\log(m)\big) && \text{if } -2d+\beta<1,\\ &= \Lambda(d)+K(d,\beta)\times m^{-\beta}\log(m)\quad+O\big(m^{-\beta}\big) && \text{if } -2d+\beta=1;\\ &= \Lambda(d)+O\big(m^{-2d-1}\big) && \text{if } -2d+\beta>1.\end{aligned}$$

Proof of Property 5.1. As in Surgailis *et al* (2008), we can write:

$$\mathrm{E}\big[IR_N(m)\big]=\mathrm{E}\Big(\frac{|Y^0+Y^1|}{|Y^0|+|Y^1|}\Big)=\Lambda\Big(\frac{R_m}{V_m^2}\Big)\quad\text{with}\quad\frac{R_m}{V_m^2}:=1-2\,\frac{\int_0^\pi f(x)\frac{\sin^6(\frac{mx}{2})}{\sin^2(\frac{x}{2})}dx}{\int_0^\pi f(x)\frac{\sin^4(\frac{mx}{2})}{\sin^2(\frac{x}{2})}dx}.$$

Therefore an expansion of R_m/V_m^2 will provide an expansion of $\mathrm{E}\big[IR_N(m)\big]$ when $m\to\infty$ and the multidimensional CLT (2.7) will be deduced from Delta-method.

Step 1 Let f satisfy Assumption $S(d,\beta)$. Then we are going to establish that there exist positive real numbers C_1 and C_2 specified in (5.4) and (5.5) and such that:

1. if $-1<-2d<1$ and $-2d+\beta<1$, $\displaystyle\frac{R_m}{V_m^2}=\rho(d)\;+\;C_1(-2d,\beta)\;\;m^{-\beta}+O\big(m^{-\beta-\varepsilon}+m^{-2d-1}\log m\big)$
2. if $-1<-2d<1$ and $-2d+\beta=1$, $\displaystyle\frac{R_m}{V_m^2}=\rho(d)+C_2(1-\beta,\beta)\;m^{-\beta}\log m+O\big(m^{-\beta}\big)$
3. if $-1<-2d<1$ and $-2d+\beta>1$, $\displaystyle\frac{R_m}{V_m^2}=\rho(d)+O\big(m^{-2d-1}\big)$.

Indeed under Assumption $S(d,\beta)$ and with $J_j(a,m)$ defined in (5.20) in Lemma 5.1, it is clear that,

$$\frac{R_m}{V_m^2}=1-2\,\frac{J_6(-2d,m)+\frac{c_1}{c_0}J_6(-2d+\beta,m)+O(J_6(-2d+\beta+\varepsilon))}{J_4(-2d,m)+\frac{c_1}{c_0}J_4(-2d+\beta,m)+O(J_4(-2d+\beta+\varepsilon))},$$

since $\displaystyle\int_0^\pi O(x^{-2d+\beta+\varepsilon})\frac{\sin^j(\frac{mx}{2})}{\sin^2(\frac{x}{2})}dx=O(J_j(-2d+\beta+\varepsilon))$. Now we follow the results of Lemma 5.1,

1. Let $-1<-2d+\beta<1$. Then for any $\varepsilon>0$,

$$\begin{aligned}\frac{R_m}{V_m^2}&=1-2\,\frac{C_{61}(-2d)m^{1+2d}+C_{62}(-2d)+\frac{c_1}{c_0}\big(C_{61}(-2d+\beta)m^{1+2d-\beta}+C_{62}(-2d+\beta)\big)+O\big(m^{1+2d-\beta-\varepsilon}+\log m\big)}{C_{41}(-2d)m^{1+2d}+C_{42}(-2d)+\frac{c_1}{c_0}\big(C_{41}(-2d+\beta)m^{1+2d-\beta}+C_{42}(-2d+\beta)\big)+O\big(m^{1+2d-\beta-\varepsilon}+\log m\big)}\\ &=1-\frac{2}{C_{41}(-2d)}\Big[C_{61}(-2d)+\frac{c_1}{c_0}C_{61}(-2d+\beta)m^{-\beta}\Big]\Big[1-\frac{c_1}{c_0}\frac{C_{41}(-2d+\beta)}{C_{41}(-2d)}m^{-\beta}\Big]+O\big(m^{-\beta-\varepsilon}+m^{-2d-1}\log m\big)\\ &=1-\frac{2C_{61}(-2d)}{C_{41}(-2d)}+2\frac{c_1}{c_0}\Big[\frac{C_{61}(-2d)C_{41}(-2d+\beta)}{C_{41}(-2d)C_{41}(-2d)}-\frac{C_{61}(-2d+\beta)}{C_{41}(-2d)}\Big]m^{-\beta}+O\big(m^{-\beta-\varepsilon}+m^{-2d-1}\log m\big).\end{aligned}$$

As a consequence, with $\rho(d)$ defined in (2.6) and C_{j1} defined in Lemma 5.1,

$$\frac{R_m}{V_m^2} = \rho(d) \ + \ C_1(-2d,\beta) \ \ m^{-\beta} + \ O\Big(m^{-\beta-\varepsilon} + m^{-2d-1}\log m\Big) \quad (m\to\infty), \quad \text{with}$$

$$C_1(-2d,\beta) := 2\,\frac{c_1}{c_0}\,\frac{1}{C_{41}^2(-2d)}\big[C_{61}(-2d)C_{41}(-2d+\beta) - C_{61}(-2d+\beta)C_{41}(-2d)\big], \quad (5.4)$$

and numerical experiments proves that $C_1(-2d,\beta)/c_1$ is negative for any $d\in(-0.5,0.5)$ and $\beta>0$.
2. Let $-2d+\beta=1$.
Again with Lemma 5.1,

$$\begin{aligned}
\frac{R_m}{V_m^2} &= 1-2\frac{[C_{61}(-2d)m^{\beta}+C'_{61}\frac{c_1}{c_0}log(m\pi)+C_{62}(-2d)+\frac{c_1}{c_0}C'_{62}+O(1)]}{[C_{41}(-2d)m^{\beta}+C'_{41}\frac{c_1}{c_0}log(m\pi)+C_{42}(-2d)+\frac{c_1}{c_0}C'_{42}+O(1)]}\\
&= 1-\frac{2}{C_{41}(a)}\big[C_{61}(-2d)+\big(C'_{61}\frac{c_1}{c_0}\log(m)\big)m^{-\beta}\big]\big[1-\big(\frac{C'_{41}}{C_{41}(a)}\frac{c_1}{c_0}\log(m)\big)m^{-\beta}\big]+O\big(m^{-\beta}\big)\\
&= 1-\frac{2}{C_{41}(-2d)}\Big[C_{61}(-2d)-\frac{c_1}{c_0}\big(\frac{C_{61}(-2d)C'_{41}}{C_{41}(-2d)}-C'_{61}\big)\log(m)\ m^{-\beta}\Big]+O\big(m^{-\beta}\big).
\end{aligned}$$

As a consequence,

$$\frac{R_m}{V_m^2} = \rho(d) \ + \ C_2(-2d,\beta)m^{-\beta} \ \log m + O\big(m^{-\beta}\big) \quad (m\to\infty), \quad \text{with}$$

$$C_2(-2d,\beta) := 2\,\frac{c_1}{c_0}\,\frac{1}{C_{41}^2(-2d)}\Big(C'_{41}C_{61}(-2d)-C'_{61}C_{41}(-2d)\Big), \quad (5.5)$$

and numerical experiments proves that $C_2(-2d,\beta)/c_1$ is negative for any $d\in(-0.5,0.5)$ and $\beta=1-2d$.
3. Let $-2d+\beta>1$.
Once again with Lemma 5.1:

$$\begin{aligned}
\frac{R_m}{V_m^2} &= 1-2\frac{\big[C_{61}(-2d)m^{1+2d}+C_{62}(-2d)+\frac{c_1}{c_0}C''_{61}(-2d+\beta)+\frac{c_1}{c_0}C''_{62}(-2d+\beta)m^{1+2d-\beta}+O(1)\big]}{C_{41}(-2d)m^{1+2d}\big[1+\frac{C_{42}(-2d)}{C_{41}(-2d)}m^{-2d-1}+\frac{c_1}{c_0}\frac{C''_{41}(-2d+\beta)}{C_{41}(-2d)}m^{-2d-1}+\frac{c_1}{c_0}\frac{C''_{42}(-2d+\beta)}{C_{41}(-2d)}m^{-\beta}+O(m^{-2d-1})\big]}\\
&= 1-\frac{2}{C_{41}(-2d)}\big[C_{61}(-2d)+O\big(m^{-2d-1}\big)\big]\big[1-O\big(m^{-2d-1}\big)\big]\\
&= 1-\frac{2C_{61}(-2d)}{C_{41}(-2d)}+O\big(m^{-2d-1}\big).
\end{aligned}$$

Note that it is not possible possible to specify the second order term of this expansion as in both the previous cases. As a consequence,

$$\frac{R_m}{V_m^2} = \rho(d) \ + \ O\big(m^{-2d-1}\big) \quad (m\to\infty). \quad (5.6)$$

Step 2: A Taylor expansion of $\Lambda(\cdot)$ around $\rho(d)$ provides:

$$\Lambda\Big(\frac{R_m}{V_m^2}\Big) \simeq \Lambda\big(\rho(d)\big) + \Big[\frac{\partial\Lambda}{\partial\rho}\Big](\rho(d))\Big(\frac{R_m}{V_m^2}-\rho(d)\Big) + \frac{1}{2}\Big[\frac{\partial^2\Lambda}{\partial\rho^2}\Big](\rho(d))\Big(\frac{R_m}{V_m^2}-\rho(d)\Big)^2. \quad (5.7)$$

Note that numerical experiments show that $\Big[\dfrac{\partial\Lambda}{\partial\rho}\Big](\rho)>0.2$ for any $\rho\in(-1,1)$. As a consequence, using the previous expansions of R_m/V_m^2 obtained in Step 1 and since $\mathrm{E}\big[IR_N(m)\big]=\Lambda\big(R_m/V_m^2\big)$, then

$$\mathrm{E}\big[IR_N(m)\big] = \Lambda_0(d) + \begin{cases} c_1\,C'_1(d,\beta)\,m^{-\beta}+O\big(m^{-\beta-\varepsilon}+m^{-2d-1}\log m+m^{-2\beta}\big) & \text{if } \beta<1+2d\\ c_1\,C'_2(\beta)m^{-\beta}\log m+O(m^{-\beta}) & \text{if } \beta=1+2d\\ O\big(m^{-2d-1}\big) & \text{if } \beta>1+2d\end{cases},$$

with $C'_1(d,\beta)<0$ for all $d\in(-0.5,0.5)$ and $\beta\in(0,1+2d)$ and $C'_2(\beta)<0$ for all $0<\beta<2$.

□

Proof of Theorem 1. Using Property 5.1, if $m \simeq C\,N^\alpha$ with $C > 0$ and $(1+2\beta)^{-1} \wedge (4d+3)^{-1} < \alpha < 1$ then $\sqrt{N/m}\,\big(\mathrm{E}\big[IR_N(m)\big] - \Lambda_0(d)\big) \underset{N\to\infty}{\longrightarrow} 0$ and it implies that the multidimensional CLT (2.4) can be replaced by

$$\sqrt{\frac{N}{m}}\Big(IR_N(m_j) - \Lambda_0(d)\Big)_{1\le j\le p} \overset{\mathcal{L}}{\underset{N\to\infty}{\longrightarrow}} \mathcal{N}(0, \Gamma_p(d)). \tag{5.8}$$

It remains to apply the Delta-method with the function Λ_0^{-1} to CLT (5.8). This is possible since the function $d \to \Lambda_0(d)$ is an increasing function such that $\Lambda_0'(d) > 0$ and $\big(\Lambda_0^{-1}\big)'(\Lambda_0(d)) = 1/\Lambda_0'(d) > 0$ for all $d \in (-0.5, 0.5)$. It achieves the proof of Theorem 1. $\square$

Proof of Proposition 1. This result is obvious using the Cochran Theorem. $\square$

In Property 5.1, a second order expansion of $\mathrm{E}IR_N(m)$ can not be specified in the case $\beta > 2d+1$. In the following Property 5.2, we show some inequalities satisfied by $\mathrm{E}IR_N(m)$ which will be useful for obtaining the consistency of the adaptive estimator in this case.

Property 5.2. *Let X satisfy Assumption $S(d,\beta)$ with $-0.5 < d < 0.5$, $\beta > 1+2d$. Moreover, suppose that the spectral density of X satisfies Condition (5.10) or (5.11). Then there exists a constant $L(c_0, c_1, c_2, d, \beta, \varepsilon) > 0$ depending only on $c_0, c_1, c_2, d, \beta, \varepsilon$ such that*

$$\mathrm{E}\big[IR_N(m)\big] - \Lambda(d) \le -L\,m^{-2d-1} \quad \textit{or} \quad \mathrm{E}\big[IR_N(m)\big] - \Lambda(d) \ge L\,m^{-2d-1}. \tag{5.9}$$

Proof of Property 5.2. Using the expansion of $J_j(a,m)$ for $a > 1$ (see Lemma 5.1) and the same computations than in Property 5.1, we obtain:

$$\begin{aligned}
&-\frac{2}{C_{41}^2(-2d)}\Big[\big(C_{62}(-2d)C_{41}(-2d)-C_{42}(-2d)C_{61}(-2d)\big)+\frac{c_1}{c_0}\big(C''_{61}(-2d+\beta)C_{41}(-2d)-C''_{41}(-2d+\beta)C_{61}(-2d)\big)\\
&\qquad+\frac{|c_2|}{c_0}\big(C''_{61}(-2d+\beta+\varepsilon)C_{41}(-2d)+C''_{41}(-2d+\beta+\varepsilon)C_{61}(-2d)\big)\Big]\,m^{-2d-1}(1+o(1))\\
&\qquad\qquad\le \frac{R_m}{V_m^2} - \rho(d) \le\\
&-\frac{2}{C_{41}^2(-2d)}\Big[\big(C_{62}(-2d)C_{41}(-2d)-C_{42}(-2d)C_{61}(-2d)\big)+\frac{c_1}{c_0}\big(C''_{61}(-2d+\beta)C_{41}(-2d)-C''_{41}(-2d+\beta)C_{61}(-2d)\big)\\
&\qquad-\frac{|c_2|}{c_0}\big(C''_{61}(-2d+\beta+\varepsilon)C_{41}(-2d)+C''_{41}(-2d+\beta+\varepsilon)C_{61}(-2d)\big)\Big]\,m^{-2d-1}(1+o(1)).
\end{aligned}$$

Now, denote

$$\begin{aligned}
D_0(d) &:= C_{62}(-2d)C_{41}(-2d) - C_{42}(-2d)C_{61}(-2d) = \frac{C_{42}(-2d)C_{41}(-2d)}{48(1-2^{-1+2d})}\big(2^{4+2d}-5-3^{2+2d}\big),\\
D_1(d,\beta) &:= C_{62}(-2d+\beta)C_{41}(-2d) - C_{42}(-2d+\beta)C_{61}(-2d) = \frac{C_{42}(-2d+\beta)C_{41}(-2d)}{128(1-2^{-1+2d})}\big(2^{4+2d}-5-3^{2+2d}\big),\\
D_2(d,\beta,\varepsilon) &:= C''_{61}(-2d+\beta+\varepsilon)C_{41}(-2d) + C''_{41}(-2d+\beta+\varepsilon)C_{61}(-2d).
\end{aligned}$$

We have $D_0(d) > 0$, $D_1(d,\beta) > 0$ and $D_2(d,\beta,\varepsilon) > 0$ for all $d \in (-0.5, 0.5)$, $\beta > 0$ and $\varepsilon > 0$. Therefore, if $c_0, c_1, c_2, d, \beta, \varepsilon$ are such that

$$K_1(c_0,c_1,c_2,d,\beta,\varepsilon) := D_0(d) + \frac{c_1}{c_0}D_1(d,\beta) - \frac{|c_2|}{c_0}D_2(d,\beta,\varepsilon) > 0 \tag{5.10}$$

$$\text{or}\quad K_2(c_0,c_1,c_2,d,\beta,\varepsilon) := D_0(d) + \frac{c_1}{c_0}D_1(d,\beta) + \frac{|c_2|}{c_0}D_2(d,\beta,\varepsilon) < 0 \tag{5.11}$$

then $\dfrac{R_m}{V_m^2} - \rho(d) \le -\dfrac{K_1(c_0,c_1,c_2,d,\beta,\varepsilon)}{C_{41}^2(-2d)}\, m^{-2d-1}$ or $\dfrac{R_m}{V_m^2} - \rho(d) \ge -\dfrac{K_2(c_0,c_1,c_2,d,\beta,\varepsilon)}{C_{41}^2(-2d)}\, m^{-2d-1}$ for m large enough. Since the function $r \to \Lambda(r)$ is an increasing function and $\mathrm{E}IR_N(m) = \Lambda\big(\dfrac{R_m}{V_m^2}\big)$ then

$$\begin{array}{rcl} \mathrm{E}IR_N(m) & \le & \Lambda\Big(\rho(d) - \dfrac{K_1(c_0,c_1,c_2,d,\beta,\varepsilon)}{C_{41}^2(-2d)}\, m^{-2d-1}\Big) \\ \text{or} \quad \mathrm{E}IR_N(m) & \ge & \Lambda\Big(\rho(d) - \dfrac{K_2(c_0,c_1,c_2,d,\beta,\varepsilon)}{C_{41}^2(-2d)}\, m^{-2d-1}\Big). \end{array}$$

Now since Λ is a smooth function, using a Taylor expansion, inequalities (5.9) hold. $\square$

Proof of Proposition 2. Let $\varepsilon > 0$ be a fixed positive real number, such that $\alpha^* + \varepsilon < 1$.

I. First, a bound of $\Pr(\widehat{\alpha}_N \le \alpha^* + \varepsilon)$ is provided. Indeed,

$$\begin{array}{rcl} \Pr\big(\widehat{\alpha}_N \le \alpha^*+\varepsilon\big) & \ge & \Pr\Big(\widehat{Q}_N(\alpha^*+\varepsilon/2) \le \displaystyle\min_{\alpha \ge \alpha^*+\varepsilon \text{ and } \alpha \in \mathcal{A}_N} \widehat{Q}_N(\alpha)\Big) \\ & \ge & 1 - \Pr\Big(\displaystyle\bigcup_{\alpha \ge \alpha^*+\varepsilon \text{ and } \alpha\in\mathcal{A}_N} \widehat{Q}_N(\alpha^*+\varepsilon/2) > \widehat{Q}_N(\alpha)\Big) \\ & \ge & 1 - \displaystyle\sum_{k=[(\alpha^*+\varepsilon)\log N]}^{\log[N/p]} \Pr\Big(\widehat{Q}_N(\alpha^*+\varepsilon/2) > \widehat{Q}_N\big(\frac{k}{\log N}\big)\Big). \end{array} \tag{5.12}$$

But, for $\alpha \ge \alpha^* + \varepsilon$,

$$\begin{array}{l} \Pr\Big(\widehat{Q}_N(\alpha^*+\varepsilon/2) > \widehat{Q}_N(\alpha)\Big) \\ \quad = \Pr\Big(\big\|\big(\widehat{d}_N(i\, N^{\alpha^*+\varepsilon/2})\big)_{1\le i\le p} - \widetilde{d}_N(N^{\alpha^*+\varepsilon/2})\big\|^2_{\widehat{\Sigma}_N(N^{\alpha^*+\varepsilon/2})} > \big\|\big(\widehat{d}_N(i\,N^\alpha) - \widetilde{d}_N(N^\alpha)\big)_{1\le i\le p}\big\|^2_{\widehat{\Sigma}_N(N^\alpha)}\Big) \end{array}$$

with $\|X\|^2_\Omega = X'\,\Omega^{-1}\,X$. Set $Z_N(\alpha) = \frac{N}{N^\alpha}\, \big\|\big(\widehat{d}_N(i\,N^\alpha)\big)_{1\le i\le p} - \widetilde{d}_N(N^\alpha)\big\|^2_{\widehat{\Sigma}_N(N^\alpha)}$. Then,

$$\begin{array}{rcl} \Pr\Big(\widehat{Q}_N(\alpha^*+\varepsilon/2) > \widehat{Q}_N(\alpha)\Big) & = & \Pr\Big(\dfrac{N^{\alpha^*+\varepsilon/2}}{N}\, Z_N(\alpha^*+\varepsilon/2) > \dfrac{N^\alpha}{N}\, Z_N(\alpha)\Big) \\ & = & \Pr\Big(Z_N(\alpha^*+\varepsilon/2) > N^{\alpha-(\alpha^*+\varepsilon/2)}\, Z_N(\alpha)\Big) \\ & \le & \Pr\Big(Z_N(\alpha^*+\varepsilon/2) > N^{(\alpha-(\alpha^*+\varepsilon/2))/2}\Big) + \Pr\Big(Z_N(\alpha) < N^{-(\alpha-(\alpha^*+\varepsilon/2))/2}\Big). \end{array}$$

From Proposition 1, for all $\alpha > \alpha^*$, $Z_N(\alpha) \xrightarrow[N\to\infty]{\mathcal{L}} \chi^2(p-1)$. As a consequence, for N large enough,

$$\begin{array}{rcl} \Pr\Big(Z_N(\alpha) \le N^{-(\alpha-(\alpha^*+\varepsilon/2))/2}\Big) & \le & 2\,\Pr\Big(\chi^2(p-1) \le N^{-(\alpha-(\alpha^*+\varepsilon/2))/2}\Big) \\ & \le & \dfrac{2}{2^{(p-1)/2}\Gamma((p-1)/2)} \cdot N^{-(\frac{p-1}{2})\frac{(\alpha-(\alpha^*+\varepsilon/2))}{2}}. \end{array}$$

Moreover, from Markov inequality and with N large enough,

$$\begin{array}{rcl} \Pr\Big(Z_N(\alpha^*+\varepsilon/2) > N^{(\alpha-(\alpha^*+\varepsilon/2))/2}\Big) & \le & 2\,\Pr\Big(\exp(\sqrt{\chi^2(p-1)}) > \exp\big(N^{(\alpha-(\alpha^*+\varepsilon/2))/4}\big)\Big) \\ & \le & 2\,\mathrm{E}(\exp(\sqrt{\chi^2(p-1)}))\,\exp\big(-N^{(\alpha-(\alpha^*+\varepsilon/2))/4}\big). \end{array}$$

Since $\mathrm{E}(\exp(\sqrt{\chi^2(p-1)})) < \infty$ does not depend on N and $\exp\big(-N^{(\alpha-(\alpha^*+\varepsilon/2))/4}\big) = o\big(N^{-(\frac{p-1}{2})\frac{(\alpha-(\alpha^*+\varepsilon/2))}{2}}\big)$ for $N \to \infty$, we deduce that there exists $M_1 > 0$ not depending on N, such that for large enough N,

$$\Pr\Big(\widehat{Q}_N(\alpha^*+\varepsilon/2) > \widehat{Q}_N(\alpha)\Big) \le M_1\, N^{-\left(\frac{p-1}{2}\right)\frac{(\alpha-(\alpha^*+\varepsilon/2))}{2}}.$$

Thus, the inequality (5.12) becomes, with $M_2 > 0$ and for N large enough,

$$\begin{aligned}
\Pr\big(\widehat{\alpha}_N \leq \alpha^* + \varepsilon\big) &\geq 1 - M_1 \sum_{k=[(\alpha^*+\varepsilon)\log N]}^{\log[N/p]} N^{-\frac{(p-1)}{4}\left(\frac{k}{\log N} - (\alpha^*+\varepsilon/2)\right)} \\
&\geq 1 - M_1\, N^{-\frac{(p-1)}{8}\varepsilon} \sum_{k=0}^{\infty} e^{-\frac{p-1}{4}k} \qquad (5.13)\\
&\geq 1 - M_2\, N^{-\frac{(p-1)}{8}\varepsilon}. \qquad (5.14)
\end{aligned}$$

II. Secondly, a bound of $\Pr(\widehat{\alpha}_N \geq \alpha^* - \varepsilon)$ can also be computed. Following the previous arguments and notations,

$$\begin{aligned}
\Pr\big(\widehat{\alpha}_N \geq \alpha^* - \varepsilon\big) &\geq \Pr\Big(\widehat{Q}_N(\alpha^* + \frac{1-\alpha^*}{2\alpha^*}\varepsilon) \leq \min_{\alpha \leq \alpha^* - \varepsilon \text{ and } \alpha \in \mathcal{A}_N} \widehat{Q}_N(\alpha)\Big) \\
&\geq 1 - \sum_{k=2}^{[(\alpha^*-\varepsilon)\log N]+1} \Pr\Big(\widehat{Q}_N(\alpha^* + \frac{1-\alpha^*}{2\alpha^*}\varepsilon) > \widehat{Q}_N\big(\frac{k}{\log N}\big)\Big), \qquad (5.15)
\end{aligned}$$

and as above,

$$\Pr\Big(\widehat{Q}_N(\alpha^* + \frac{1-\alpha^*}{2\alpha^*}\varepsilon) > \widehat{Q}_N(\alpha)\Big) = \Pr\Big(Z_N(\alpha^* + \frac{1-\alpha^*}{2\alpha^*}\varepsilon) > N^{\alpha - (\alpha^* + \frac{1-\alpha^*}{2\alpha^*}\varepsilon)} Z_N(\alpha)\Big). \qquad (5.16)$$

Now, in the case $\alpha < \alpha^*$, from the proof of Theorem 1 if $\beta \leq 2d+1$ then

$$\begin{aligned}
&\sqrt{\frac{N}{N^\alpha}}\big(\mathrm{E}\big[IR(i\,N^\alpha)\big] - \Lambda(d)\big) \simeq C\, i^{-(1-\alpha^*)/2\alpha^*}\, N^{(\alpha^*-\alpha)/2\alpha^*} (\log N)^{\mathbf{1}_{\beta=2d+1}} \\
&\Longrightarrow \sqrt{\frac{N}{N^\alpha}}\big(\widehat{d}(i\,N^\alpha) - d\big)_{1\leq i\leq p} \simeq C'\, N^{\frac{\alpha^*-\alpha}{2\alpha^*}} (\log N)^{\mathbf{1}_{\beta=2d+1}} \Big(i^{-\frac{1-\alpha^*}{2\alpha^*}}\Big)_{1\leq i\leq p} + \big(\widehat{\varepsilon}_N(i\,N^\alpha)\big)_{1\leq i\leq p},
\end{aligned}$$

and $\sqrt{\frac{N}{N^\alpha}}\big(\widetilde{d}_N(N^\alpha) - d\big) \simeq C'\, N^{\frac{\alpha^*-\alpha}{2\alpha^*}} (\log N)^{\mathbf{1}_{\beta=2d+1}} \big(J_p'\widehat{\Sigma}_N^{-1}(N^\alpha)J_p\big)^{-1} J_p'\widehat{\Sigma}_N^{-1}(N^\alpha)\big(i^{-\frac{1-\alpha^*}{2\alpha^*}}\big)_{1\leq i\leq p} + \big(\widetilde{\varepsilon}_N(i\,N^\alpha)\big)_{1\leq i\leq p}$

with $C \neq 0$, $C' = C\,(\Lambda_0'(d))^{-1} \neq 0$, and using Proposition 1, $\big(\widehat{\varepsilon}_N(i\,N^\alpha)\big)_{1\leq i\leq p} \xrightarrow[N\to\infty]{\mathcal{L}} \mathcal{N}\Big(0, (\Lambda_0'(d))^{-2}\,\Gamma_p(d)\Big)$ and $\big(\widetilde{\varepsilon}_N(i\,N^\alpha)\big)_{1\leq i\leq p} \xrightarrow[N\to\infty]{\mathcal{L}} \mathcal{N}\Big(0, (\Lambda_0'(d))^{-2}\big(J_p'\Gamma_p^{-1}(d)J_p\big)^{-1}\Big)$.

As a consequence, for $\alpha < \alpha^* - \varepsilon$,

$$\begin{aligned}
Z_N(\alpha) &\geq (C')^2\, N^{\frac{\alpha^*-\alpha}{\alpha^*}} (\log^2 N)^{\mathbf{1}_{\beta=2d+1}} \Big\|\big(J_p\big(J_p'\widehat{\Sigma}_N^{-1}(N^\alpha)J_p\big)^{-1}J_p'\widehat{\Sigma}_N^{-1}(N^\alpha) - I_p\big)\big(i^{-\frac{1-\alpha^*}{2\alpha^*}}\big)_{1\leq i\leq p}\Big\|_{\widehat{\Sigma}_N(N^\alpha)} \\
&\quad -2\frac{N}{N^\alpha}\,\|\widehat{\varepsilon}_N(i\,N^\alpha))\|^2_{\widehat{\Sigma}_N(N^\alpha)} - 2\frac{N}{N^\alpha}\,\|\widetilde{\varepsilon}_N(i\,N^\alpha))\|^2_{\widehat{\Sigma}_N(N^\alpha)}.
\end{aligned}$$

Now, since the vector $\big(i^{-\frac{1-\alpha^*}{2\alpha^*}}\big)_{1\leq i\leq p}$ is not in the subspace generated by J_p, we deduce that there exists $D > 0$ such that for N large enough and $\alpha < \alpha^* - \varepsilon$,

$$Z_N(\alpha) \geq D\, N^{\frac{\alpha^*-\alpha}{\alpha^*}} (\log^2 N)^{\mathbf{1}_{\beta=2d+1}} - 2\,\frac{N}{N^\alpha}\,\|\widehat{\varepsilon}_N(i\,N^\alpha))\|^2_{\widehat{\Sigma}_N(N^\alpha)} - 2\,\frac{N}{N^\alpha}\,\|\widetilde{\varepsilon}_N(i\,N^\alpha))\|^2_{\widehat{\Sigma}_N(N^\alpha)}.$$

with $\frac{N}{N^\alpha}\,\|\widehat{\varepsilon}_N(i\,N^\alpha))\|^2_{\widehat{\Sigma}_N(N^\alpha)} \xrightarrow[N\to\infty]{\mathcal{L}} \chi^2(p)$ and $\frac{N}{N^\alpha}\,\|\widetilde{\varepsilon}_N(i\,N^\alpha))\|^2_{\widehat{\Sigma}_N(N^\alpha)} \xrightarrow[N\to\infty]{\mathcal{L}} (\Lambda_0'(d))^{-2}\big(J_p'\Gamma_p(d)J_p\big)^{-1}\chi^2(1)$.

Therefore, since $N^{\frac{\alpha^*-\alpha}{\alpha^*}} \xrightarrow[N\to\infty]{} \infty$ when $\alpha < \alpha^* - \varepsilon$,

$$\Pr\Big(Z_N(\alpha) \geq \frac{1}{2}\,D\,N^{\frac{\alpha^*-\alpha}{\alpha^*}}\Big) \xrightarrow[N\to\infty]{} 1.$$

Then, the relation (5.16) becomes for $\alpha < \alpha^* - \varepsilon$ and N large enough,

$$\begin{aligned}
\Pr\Big(\widehat{Q}_N(\alpha^* + \frac{1-\alpha^*}{2\alpha^*}\varepsilon) > \widehat{Q}_N(\alpha)\Big) &\leq \Pr\Big(\chi^2(p-1) \geq \big(\frac{1}{2}\,D\,N^{\frac{\alpha^*-\alpha}{\alpha^*}}\big)\,N^{\alpha - (\alpha^* + \frac{1-\alpha^*}{2\alpha^*}\varepsilon)}\Big) \\
&\leq \Pr\Big(\chi^2(p-1) \geq \frac{D}{2}\,N^{\frac{1-\alpha^*}{2\alpha^*}(2(\alpha^*-\alpha)-\varepsilon)}\Big) \\
&\leq M_2 \cdot N^{-(\frac{p-1}{2})\frac{1-\alpha^*}{2\alpha^*}\,\varepsilon},
\end{aligned}$$

with $M_2 > 0$, because $\frac{1-\alpha^*}{2\alpha^*}(2(\alpha^*-\alpha)-\varepsilon) \geq \frac{1-\alpha^*}{2\alpha^*}\varepsilon$ for all $\alpha \leq \alpha^* - \varepsilon$. Hence, from the inequality (5.15), for large enough N ,

$$\Pr\big(\widehat{\alpha}_N \geq \alpha^* - \varepsilon\big) \geq 1 - M_2 \cdot \log N \cdot N^{-(p-1)\frac{1-\alpha^*}{4\alpha^*}\varepsilon}. \tag{5.17}$$

If $\beta > 2d+1$, with $\alpha^* = (4d+3)^{-1}$ and from Property 5.2, we obtain an inequality (here we only consider the case $\leq$, the second case $\geq$ identically follows) instead of the equality (5.17):

$$\sqrt{\frac{N}{N^\alpha}}\big(\mathrm{E}\big[IR(i\,N^\alpha)\big] - \Lambda(d)\big) \leq -L(c_0,c_1,c_2,d,\beta,\varepsilon)\, i^{-(1-\alpha^*)/2\alpha^*}\, N^{(\alpha^*-\alpha)/2\alpha^*}. \tag{5.18}$$

Now, as previously and with the same notation, using a Taylor expansion,

$$\sqrt{\frac{N}{N^\alpha}}\big(\widetilde{d}_N(N^\alpha) - \widetilde{d}_N(N^\alpha)J_p\big) \simeq (\Lambda_0'(d))^{-1}\Big(J_p\big(J_p'\widehat{\Sigma}_N^{-1}(N^\alpha)J_p\big)^{-1}J_p'\widehat{\Sigma}_N^{-1}(N^\alpha) - I_p\Big)\big(\mathrm{E}\big[IR_N(iN^\alpha)\big] - \Lambda(d)\big)_{1\leq i\leq p} + \big(\widehat{\varepsilon}_N(i\,N^\alpha)\big)_{1\leq i\leq p} - \big(\widetilde{\varepsilon}_N(i\,N^\alpha)\big)_{1\leq i\leq p}. \tag{5.19}$$

Now the steps of the proof in the case $\beta \leq 2d+1$ can be followed and the same kind of bound (5.17) can be obtained.

Finally, the inequalities (5.14) and (5.17) imply that $\Pr\big(|\widehat{\alpha}_N - \alpha| \geq \varepsilon\big) \underset{N\to\infty}{\longrightarrow} 0.$ $\square$

Proof of Theorem 2. The results of Theorem 2 can be easily deduced from Theorem 1 and Proposition 2 (and its proof) by using conditional probabilities. $\square$

Proof of Proposition 3. Proposition 3 can be easily deduced from Theorem 2 using the Cochran Theorem. $\square$

Lemma 5.1. *For $j = 4, 6$, denote*

$$J_j(a,m) := \int_0^\pi x^a \frac{\sin^j(\frac{mx}{2})}{\sin^2(\frac{x}{2})}dx. \tag{5.20}$$

Then, we have the following expansion when $m \to \infty$:

1. *if* $-1 < a < 1$, $J_j(a,m) = C_{j1}(a)\,m^{1-a} + C_{j2}(a) + O\big(m^{-1-(a\wedge 0)}\big)$;
2. *if* $a = 1$, $J_j(a,m) = C_{j1}'\,\log(m) + C_{j2}' + O\big(m^{-1}\big)$;
3. *if* $a > 1$, $J_j(a,m) = C_{j1}''(a) + O\big(m^{1-a} + m^{-2}\big)$,

where constants $C_{j1}(a)$, $C_{j2}(a)$, $C_{j1}'(a)$, $C_{j2}'(a)$ and $C_{j1}''(a)$ are specified in the following proof and do not vanish for all a.

Proof of Lemma 5.1. 1. let $-1 < a < 1$.
We begin with the expansion of $J_4(a,m)$. First, decompose $J_4(a,m)$ as follows

$$J_4(a,m) = J_0(a,m) + \int_0^\pi \frac{x^a}{(\frac{x}{2})^2}\sin^4(\frac{mx}{2})dx$$
$$\text{with}\quad J_0(a,m) := 2^{a+1}\int_0^{\frac{\pi}{2}} y^a \sin^4(my)\Big[\frac{1}{\sin^2(y)} - \frac{1}{y^2}\Big]dy. \tag{5.21}$$

It is clear that using integrations by parts for $\sin^4(\frac{x}{2}) = \sin^2(\frac{x}{2}) - \frac{1}{4}\sin^2(x) = \frac{1}{8}\big(3 - 4\cos(y) + \cos(2y)\big)$ and when $m \to \infty$:

$$\begin{aligned}
\int_0^{\pi} \frac{x^a}{(\frac{x}{2})^2}\sin^4(\frac{mx}{2})dx &= 4\,m^{1-a}\int_0^{m\pi} y^{a-2}\sin^4(\frac{y}{2})dy \\
&= 4\,m^{1-a}\Big(\int_0^{\infty} y^{a-2}\sin^4(\frac{y}{2})dy - \int_{m\pi}^{\infty} y^{a-2}\sin^4(\frac{y}{2})dy\Big) \\
&= 4\,m^{1-a}\Big((1-\frac{1}{2^{1+a}})\int_0^{\infty}\frac{\sin^2(\frac{y}{2})}{y^{2(\frac{1-a}{2})+1}}dy - \frac{1}{8}\int_{m\pi}^{\infty} y^{a-2}\big(3-4\cos(y)+\cos(2y)\big)dy\Big) \\
&= \frac{\pi(1-\frac{1}{2^{1+a}})}{(1-a)\Gamma(1-a)\sin(\frac{(1-a)\pi}{2})}\,m^{1-a} - 3\,\frac{1}{2(1-a)}\pi^{a-1} + O(m^{-1})
\end{aligned}$$

where the left right side term of the last relation is obtained by integration by parts and the left side term is deduced from the following relation (see Taqqu *et al.*, 2003 , p. 31)

$$\int_0^{\infty} y^{-\alpha}\ \sin(y)\ dy \ = \ \frac{1}{2}\,\frac{\pi}{\Gamma(\alpha)\sin(\pi(\frac{\alpha}{2}))} \quad \text{for } 0<\alpha<2. \tag{5.22}$$

Moreover, with the linearization of $\sin^4 u$, let write

$$\frac{8}{2^{a+1}}J_0(a,m) \ = \ \underbrace{3\int_0^{\frac{\pi}{2}} y^a[\frac{1}{\sin^2(y)}-\frac{1}{y^2}]dy}_{J_{01}(a,m)} \ + \ \underbrace{\int_0^{\frac{\pi}{2}}(\cos(4my)-4\cos(2my))\ y^a\ [\frac{1}{\sin^2(y)}-\frac{1}{y^2}]dy.}_{J_{02}(a,m)}$$

From usual Taylor expansions,

$$\frac{1}{\sin^2(y)}-\frac{1}{y^2} \underset{y\to 0}{\sim} \frac{1}{3} \quad \text{and} \quad \frac{1}{y^3}-\frac{\cos(y)}{\sin^3(y)} \underset{y\to 0}{\sim} \frac{y}{15}.$$

From the first expansion we deduce that $J_{01}(a,m)$ exists since $-1<a<1$. Moreover, from an integration by parts,

$$J_{02}(a,m) \ = \ -\frac{1}{m}\int_0^{\frac{\pi}{2}}\Big(\frac{\sin(4my)}{4}-2\sin(2my)\Big)\Big(ay^{a-1}[\frac{1}{\sin^2(y)}-\frac{1}{y^2}]+2\,y^a\big[\frac{1}{y^3}-\frac{\cos(y)}{\sin^3(y)}\big]\Big)dy.$$

Hence, using $|\sin u|\leq 1\wedge |u|$, and $\big|\frac{1}{\sin^2(y)}-\frac{1}{y^2}\big|\leq 1$, $\big|\frac{1}{y^3}-\frac{\cos(y)}{\sin^3(y)}\big|\leq y$ for all $y\in[0,\pi/2]$, and integration by parts,

$$\begin{aligned}
\big|m\,J_{02}(a,m)\big| &\leq \frac{9}{4}\int_0^{\frac{\pi}{2}} 2\,y^a\big|\frac{1}{y^3}-\frac{\cos(y)}{\sin^3(y)}\big|dy + \Big|\int_0^{\frac{\pi}{2}}\Big(\frac{\sin(4my)}{4}-2\sin(2my)\Big)a\,y^{a-1}\Big(\frac{1}{\sin^2(y)}-\frac{1}{y^2}\Big)dy\Big| \\
&\leq \frac{9}{2(a+2)}(\frac{\pi}{2})^{a+2} + \Big|\int_0^{\frac{1}{m}} 5\,m\,a\,y^a dy\Big| + \frac{9}{4}\Big|\int_{\frac{1}{m}}^{\frac{\pi}{2}} a\,y^{a-1}dy\Big|
\end{aligned}$$

As a consequence, when $m\to\infty$, $J_{02}(a,m)=O\big(m^{-1-(a\wedge 0)}\big)$ and therefore

$$J_0(a,m) = \ 3\ \frac{2^{a+1}}{8}\int_0^{\frac{\pi}{2}} y^a[\frac{1}{\sin^2(y)}-\frac{1}{y^2}]dy + O\big(m^{-1-(a\wedge 0)}\big). \tag{5.23}$$

Finally, by replacing this expansion in (5.21), one deduces

$$J_4(a,m) = \int_0^{\pi} x^a\frac{\sin^4(\frac{mx}{2})}{\sin^2(\frac{x}{2})}dx = C_{41}(a)\,m^{1-a} + C_{42}(a) + O\big(m^{-1-(a\wedge 0)}\big) \quad (m\to\infty), \text{with}$$

$$C_{41}(a) := \frac{\pi(1-\frac{1}{2^{1+a}})}{(1-a)\Gamma(1-a)\sin(\frac{(1-a)\pi}{2})} \ \text{ and } C_{42}(a) := \frac{3}{2^{2-a}}\int_0^{\frac{\pi}{2}} y^a[\frac{1}{\sin^2(y)}-\frac{1}{y^2}]dy - \frac{3}{2(1-a)}\pi^{a-1}. \tag{5.24}$$

Note that $C_{41}(a) > 0$ and $C_{42}(a) < 0$ for all $0 < a < 1$, $C_{42}(a) > 0$ for all $-1 < a < 0$, $C_{42}(0) = 0$.

A similar expansion procedure of $J_6(a,m)$ with $\sin^6(\frac{mx}{2})$ instead of $\sin^4(\frac{mx}{2})$ can be provided. Let

$$J_6(a,m) := J_0'(a,m) + \int_0^\pi \frac{x^a}{(\frac{x}{2})^2} \sin^6(\frac{mx}{2})dx \quad \text{with} \quad J_0'(a,m) := 2^{a+1} \int_0^{\frac{\pi}{2}} y^a \sin^6(my)\Big[\frac{1}{\sin^2(y)} - \frac{1}{y^2}\Big]dy.$$

As previously with $\sin^6(\frac{y}{2}) = \frac{1}{32}\big(10 - 15\cos(y) + 6\cos(2y) - \cos(3y)\big)$, then, when $m \to \infty$,

$$J_6(a,m) = C_{61}(a)\, m^{1-a} + C_{62}(a) + O\big(m^{-1-(a\wedge 0)}\big),$$
$$\text{with } C_{61}(a) := \frac{\pi(15 + 3^{1-a} - 2^{1-a}6)}{16(1-a)\Gamma(1-a)\sin(\frac{\pi}{2}(1-a))} \text{ and } C_{62}(a) := \frac{5}{6}\, C_{42}(a).$$

Moreover it is clear that $C_{61}(a) > 0$.

2. let $a = 1$.
When $m \to \infty$ we obtain the following expansion:

$$\begin{aligned}
\int_0^\pi \frac{x\sin^4(\frac{mx}{2})}{\sin^2(\frac{x}{2})}dx &= 4\int_0^{m\pi} \frac{1}{x}\sin^4(\frac{x}{2})dx + \int_0^\pi x\sin^4(\frac{mx}{2})\Big(\frac{1}{\sin^2(\frac{x}{2})} - \frac{1}{(\frac{x}{2})^2}\Big)dx \\
&= \frac{1}{2}\Big(\int_0^{m\pi} \frac{\cos(2x)-1}{x}dx - 4\int_0^{m\pi} \frac{\cos(x)-1}{x}dx\Big) + 4\int_0^{\frac{\pi}{2}} y\sin^4(my)\Big(\frac{1}{\sin^2(y)} - \frac{1}{y^2}\Big)dy \\
&= \frac{1}{2}\Big(\int_0^{m\pi} \frac{\sin(2x)-2x}{2x^2}dx - 4\int_0^{m\pi} \frac{\sin(x)-x}{x^2}dx\Big) + 4\int_0^{\frac{\pi}{2}} y\sin^4(my)\Big(\frac{1}{\sin^2(y)} - \frac{1}{y^2}\Big)dy
\end{aligned}$$

But,

$$\begin{aligned}
\int_0^{m\pi} \frac{\sin(x)-x}{x^2}dx &= \int_0^1 \frac{\sin(x)-x}{x^2}dx + \int_1^{m\pi} \frac{\sin(x)}{x^2}dx - \int_1^{m\pi} \frac{1}{x}dx \\
&= \int_0^1 \frac{\sin(x)-x}{x^2}dx + \int_1^\infty \frac{\sin(x)}{x^2}dx + O(m^{-1}) - \log(m\pi).
\end{aligned}$$

From the same decomposition we obtain

$$\int_0^{m\pi} \frac{\sin(2x)-2x}{2x^2}dx - 4\int_0^{m\pi} \frac{\sin(x)-x}{x^2}dx = \frac{3}{2}\Big(\log(m\pi) + \int_1^\infty \frac{\sin y}{y^2}dy + \int_0^1 \frac{\sin y - y}{y^2}dy\Big) + O(m^{-1}).$$

Moreover from previous computations (see the case $a < 1$),

$$\int_0^{\frac{\pi}{2}} y\sin^4(my)\Big(\frac{1}{\sin^2(y)} - \frac{1}{y^2}\Big)dy = \frac{3}{8}\int_0^{\frac{\pi}{2}} y\Big(\frac{1}{\sin^2(y)} - \frac{1}{y^2}\Big)dy + O(m^{-1}).$$

As a consequence, when $m \to \infty$,

$$\int_0^\pi \frac{x\sin^4(\frac{mx}{2})}{\sin^2(\frac{x}{2})}dx = C'_{41}\,\log(m) + C'_{42} + O\big(m^{-1}\big), \quad \text{with} \quad C'_{41} := \frac{3}{2} \quad \text{and}$$
$$C'_{42} := \frac{3}{2}\Big(\log(\pi) + \int_0^{\frac{\pi}{2}} y\Big(\frac{1}{\sin^2(y)} - \frac{1}{y^2}\Big)dy + \int_1^\infty \frac{\sin y}{y^2}dy + \int_0^1 \frac{\sin y - y}{y^2}dy\Big).$$

Note that $C'_{41} > 0$ and $C'_{42} \simeq 2.34 > 0$.

In the same way , we obtain the following expansions when $m \to \infty$,

$$\int_0^\pi \frac{x\sin^6(\frac{mx}{2})}{\sin^2(\frac{x}{2})}dx = C'_{61}\,\log(m) + C'_{62} + O\big(m^{-1}\big) \quad \text{with} \quad C'_{61} := \frac{5}{4} \quad \text{and}$$
$$C'_{62} := \frac{5}{4}\log(\pi) + \frac{5}{4}\int_0^{\frac{\pi}{2}} y\Big(\frac{1}{\sin^2(y)} - \frac{1}{y^2}\Big)dy + \frac{1}{8}\int_1^\infty \frac{1}{y}\Big(-\cos(3y) + 6\cos(2y) - 15\cos(y)\Big)dy + 4\int_0^1 \frac{1}{y}\sin^6(\frac{y}{2})dy.$$

Note again that $C'_{61} > 0$ and numerical experiments show that $C'_{62} > 0$.

3. Let $a > 1$. Then, with the linearization of $\sin^4(u)$,

$$\begin{aligned}
\int_0^\pi \frac{x^a \sin^4(\frac{mx}{2})}{\sin^2(\frac{x}{2})}dx &= \frac{3}{8}\int_0^\pi \frac{x^a}{\sin^2(\frac{x}{2})}dx - \frac{1}{2}\int_0^\pi \frac{x^a}{\sin^2(\frac{x}{2})}\cos(mx)dx + \frac{1}{8}\int_0^\pi \frac{x^a}{\sin^2(\frac{x}{2})}\cos(2mx)dx \\
&= C''_{41}(a) + \frac{1}{m}\int_0^\pi \Big(\frac{\sin(mx)}{2} - \frac{\sin(2mx)}{16}\Big)\Big(g(x) + h(x)\Big)dx, \qquad (5.25)
\end{aligned}$$

with: $g(x) = \Big(\frac{ax^{a-1}}{\sin^2(\frac{x}{2})} - 4ax^{a-3}\Big) - \Big(\frac{x^a\cos(\frac{x}{2})}{\sin^3(\frac{x}{2})} - 8x^{a-3}\Big)$ and $h(x) = (4a-8)x^{a-3}$.

First, if $1 < a \leq 3$, with an integration by parts,

$$\begin{aligned}
I_m(h,a) &= \frac{1}{m}\int_0^\pi \Big(\frac{\sin(mx)}{2} - \frac{\sin(2mx)}{16}\Big)h(x)dx \\
&= \Big((4a-8)\int_0^\infty x^{a-3}\Big(\frac{\sin(x)}{2} - \frac{\sin(2x)}{16}\Big)dx\Big)m^{1-a} - \frac{(4a-8)}{m}\int_\pi^\infty x^{a-3}\Big(\frac{\sin(mx)}{2} - \frac{\sin(2mx)}{16}\Big)dx \\
&= O\big(m^{1-a} + m^{-2}\big). \qquad (5.26)
\end{aligned}$$

Now, if $a > 3$, the straightforward integration by parts is still possible and

$$\begin{aligned}
I_m(h,a) &= \frac{1}{m}\int_0^\pi \Big(\frac{\sin(mx)}{2} - \frac{\sin(2mx)}{16}\Big)h(x)dx \\
&= \frac{1}{m^2}\Big[\Big(-\frac{\cos(mx)}{2} + \frac{\cos(2mx)}{32}\Big)h(x)\Big]_0^\pi + \frac{4(a-2)(a-3)}{m^2}\int_0^\pi \Big(\frac{\cos(mx)}{2} - \frac{\cos(2mx)}{32}\Big)x^{a-4}dx \\
&= O\big(m^{-2}\big) \qquad (5.27)
\end{aligned}$$

Moreover,

$$\begin{aligned}
I_m(g,a) &= \frac{1}{m}\int_0^\pi \Big(\frac{\sin(mx)}{2} - \frac{\sin(2mx)}{16}\Big)g(x)dx \\
&= \frac{1}{m^2}\Big[\Big(-\frac{\cos(mx)}{2} + \frac{\cos(2mx)}{32}\Big)\ g(x)\Big]_0^\pi - \frac{1}{m^2}\int_0^\pi \Big(-\frac{\cos(mx)}{2} + \frac{\cos(2mx)}{32}\Big)\ g'(x)dx \\
&= \Big(\frac{1}{32} - \frac{(-1)^m}{2}\Big)\big(a\pi^2 - 4a + 8\big)\pi^{a-3}\frac{1}{m^2} - \frac{1}{m^2}\int_0^\pi \Big(-\frac{\cos(mx)}{2} + \frac{\cos(2mx)}{32}\Big)\ g'(x)dx
\end{aligned}$$

since $g(x) \underset{x=0^+}{\sim} \frac{a}{3}\, x^{a-1}$ and $g'(x) \underset{x=0^+}{\sim} \frac{a(a-1)}{3}x^{a-2}$. Therefore, if $1 < a < 3$,

$$I_m(g,a) = O(m^{-2}). \qquad (5.28)$$

If $a \geq 3$, another integration by parts is possible and

$$\begin{aligned}
I_m(g,a) &= \Big(\frac{1}{32} - \frac{(-1)^m}{2}\Big)\big(a\pi^2 - 4a + 8\big)\pi^{a-3}\frac{1}{m^2} - \frac{1}{m^3}\Big[\Big(-\frac{\sin(mx)}{2} + \frac{\sin(2mx)}{64}\Big)\ g'(x)\Big]_0^\pi \\
&\qquad + \frac{1}{m^3}\int_0^\pi \Big(-\frac{\sin(mx)}{2} + \frac{\sin(2mx)}{64}\Big)\ g''(x)dx \\
&= O\big(m^{-2}\big), \qquad (5.29)
\end{aligned}$$

since $g''(x) \underset{x=0^+}{\sim} \frac{1}{3}\, a(a-1)(a-2)x^{a-3}$.

In conclusion, for $1 < a$ we deduce,

$$\int_0^\pi \frac{x^a \sin^4(\frac{mx}{2})}{\sin^2(\frac{x}{2})}dx = C''_{41}(a) + O\big(m^{1-a} + m^{-2}\big) \quad \text{with} \quad C''_{41}(a) := \frac{3}{8}\int_0^\pi \frac{x^a}{\sin^2(\frac{x}{2})}dx. \qquad (5.30)$$

Similarly, for $1 < a < 3$ we deduce,

$$\int_0^\pi \frac{x^a \sin^6(\frac{mx}{2})}{\sin^2(\frac{x}{2})} dx = C''_{61}(a) + O\big(m^{1-a} + m^{-2}\big) \quad \text{with} \quad C''_{61}(a) := \frac{5}{16} \int_0^\pi \frac{x^a}{\sin^2(\frac{x}{2})} dx = \frac{5}{6}\, C''_{41}(a). \tag{5.31}$$

□

References

[1] Bardet, J.M. (2000). Testing for the presence of self-similarity of Gaussian time series having stationary increments, *J. Time Ser. Anal.* , 21, 497-516.

[2] Bardet, J.M., Bibi H. and Jouini, A. (2008). Adaptive wavelet-based estimator of the memory parameter for stationary Gaussian processes, *Bernoulli*, 14, 691-724.

[3] Bardet, J.M., Lang, G., Oppenheim, G., Philippe, A., Stoev, S. and Taqqu, M.S. (2003). Semiparametric estimation of the long-range dependence parameter: a survey. In *Theory and applications of long-range dependence*, Birkhäuser Boston, 557-577.

[4] Bardet J.M. and Surgailis, D. (2010). Measuring the roughness of random paths by increment ratios, To appear in *Bernoulli.*

[5] Bruzaite, K. and Vaiciulis, M. (2008). The increment ratio statistic under deterministic trends. *Lith. Math. J.* 48, 256-269.

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

[7] Doukhan, P., Oppenheim, G. and Taqqu M.S. (Editors) (2003). *Theory and applications of long-range dependence*, Birkhäuser.

[8] Fox, R. and Taqqu, M.S. (1986). Large-sample properties of parameter estimates for strongly dependent Gaussian time series. *Ann. Statist.* 14, 517-532.

[9] Geweke, J. and Porter-Hudak, S. (1983), The estimation and application of long-memory time-series models, *J. Time Ser. Anal.*, 4, 221-238.

[10] Giraitis, L., Robinson P.M., and Samarov, A. (1997). Rate optimal semiparametric estimation of the memory parameter of the Gaussian time series with long range dependence, *J. Time Ser. Anal.*, 18, 49-61.

[11] Giraitis, L. and Surgailis, D. (1990). A central limit theorem for quadratic forms in strongly dependent linear variables and its applications to the asymptotic normality of Whittle estimate. *Prob. Th. and Rel. Field.* 86, 87-104.

[12] Giraitis, L. and Taqqu, M.S. (1999). Whittle estimator for finite-variance non-Gaussian time series with long memory. *Ann. Statist.* 27, 178–203.

[13] Hurst, H. E. (1951) Long-term storage capacity of reservoirs, *Trans, Am. Soc. Civil Eng*, 116, 770-779.

[14] Moulines, E., Roueff, F. and Taqqu, M.S. (2007). On the spectral density of the wavelet coefficients of long memory time series with application to the log-regression estimation of the memory parameter, *J. Time Ser. Anal.*, 28, 155-187.

[15] Moulines, E. and Soulier, P. (2003). Semiparametric spectral estimation for fractionnal processes, In P. Doukhan, G. Openheim and M.S. Taqqu editors, *Theory and applications of long-range dependence*, 251-301, Birkhäuser, Boston.

[16] Robinson, P.M. (1995). Gaussian semiparametric estimation of long range dependence, *Ann. Statist.*, 23, 1630-1661.

[17] Surgailis, D., Teyssière, G. and Vaičiulis, M. (2007) The increment ratio statistic. *J. Multiv. Anal.* **99**, 510-541.

[18] Vaiciulis, M. (2009). An estimator of the tail index based on increment ratio statistics. *Lith. Math. J.* 49, 222-233.

[19] Veitch, D., Abry, P. and Taqqu, M.S. (2003). On the Automatic Selection of the Onset of Scaling, *Fractals*, 11, 377-390.

	Model	Estimates	$p=5$	$p=10$	$p=15$	$p=20$
$N=10^3$	fGn $(H=d+1/2)$	$\sqrt{MSE}\ \widehat{\widetilde{d}}_N$	0.092	**0.074***	0.088	0.098
		mean$(\widehat{m})$	12.6	11.2	14.6	18.0
		$\widehat{proba}$	1.00	0.99	0.98	0.97
	FARIMA$(0,d,0)$	$\sqrt{MSE}\ \widehat{\widetilde{d}}_N$	0.103	0.095	**0.093***	0.100
		mean$(\widehat{m})$	12.2	11.6	14.3	17.0
		$\widehat{proba}$	1.00	0.99	0.98	0.98
	FARIMA$(1,d,1)$	$\sqrt{MSE}\ \widehat{\widetilde{d}}_N$	0.149	**0.145***	0.147	0.153
		mean$(\widehat{m})$	14.1	13.6	16.9	19.9
		$\widehat{proba}$	0.99	0.99	0.96	0.95
	$X^{(d,\beta)}$, $\beta=1$	$\sqrt{MSE}\ \widehat{\widetilde{d}}_N$	0.124	**0.117***	0.124	0.131
		mean$(\widehat{m})$	13.4	12.7	15.4	19.1
		$\widehat{proba}$	1.00	0.99	0.99	0.98

	Model	Estimates	$p=5$	$p=10$	$p=15$	$p=20$
$N=10^4$	fGn $(H=d+1/2)$	$\sqrt{MSE}\ \widehat{\widetilde{d}}_N$	0.028	0.021	0.018	**0.017***
		mean$(\widehat{m})$	11.1	9.0	9.0	8.1
		$\widehat{proba}$	1.00	0.99	0.98	0.97
	FARIMA$(0,d,0)$	$\sqrt{MSE}\ \widehat{\widetilde{d}}_N$	0.042	0.038	**0.036***	0.037
		mean$(\widehat{m})$	11.7	7.5	6.7	6.3
		$\widehat{proba}$	1.00	0.99	0.98	0.98
	FARIMA$(1,d,1)$	$\sqrt{MSE}\ \widehat{\widetilde{d}}_N$	0.071	0.067	**0.065***	0.068
		mean$(\widehat{m})$	17.3	15.2	13.4	12.3
		$\widehat{proba}$	0.99	0.98	0.96	0.94
	$X^{(d,\beta)}$, $\beta=1$	$\sqrt{MSE}\ \widehat{\widetilde{d}}_N$	0.074	**0.073***	**0.073***	0.075
		mean$(\widehat{m})$	14.8	12.9	10.9	10.3
		$\widehat{proba}$	1.00	0.98	0.96	0.95

	Model	Estimates	$p=5$	$p=10$	$p=15$	$p=20$
$N=10^5$	fGn $(H=d+1/2)$	$\sqrt{MSE}\ \widehat{\widetilde{d}}_N$	0.010	0.007	0.006	**0.005***
		mean$(\widehat{m})$	14.0	9.3	7.4	6.2
		$\widehat{proba}$	1.00	0.98	0.97	0.95
	FARIMA$(0,d,0)$	$\sqrt{MSE}\ \widehat{\widetilde{d}}_N$	0.024	0.022	**0.021***	**0.021***
		mean$(\widehat{m})$	13.8	10.8	8.5	7.0
		$\widehat{proba}$	0.99	0.97	0.97	0.93
	FARIMA$(1,d,1)$	$\sqrt{MSE}\ \widehat{\widetilde{d}}_N$	0.039	**0.038***	**0.038***	0.039
		mean$(\widehat{m})$	23.4	20.5	19.3	17.0
		$\widehat{proba}$	1.00	0.98	0.94	0.94
	$X^{(d,\beta)}$, $\beta=1$	$\sqrt{MSE}\ \widehat{\widetilde{d}}_N$	0.042	0.040	**0.039***	0.041
		mean$(\widehat{m})$	22.5	21.3	19.3	16.3
		$\widehat{proba}$	0.99	0.98	0.96	0.94

Table 1: $\sqrt{MSE}$ *of the estimator* $\widehat{\widetilde{d}}_N$*, sample mean of the estimator* $\widehat{m}_N$ *and sample frequency that* $\widehat{T_N} \leq q_{\chi^2(p-1)}(0.95)$ *following* p *from simulations of the different long-memory processes of the benchmark. For each value of* N *(*10^3, 10^4 *and* 10^5*), of* d *(*-0.4, -0.2, 0, 0.2 *and* 0.4*) and* p *(*5, 10, 15, 20*),* 100 *independent samples of each process are generated. The values* $\sqrt{MSE}\ \widehat{\widetilde{d}}_N$*, mean(*$\widehat{m}$*) and* $\widehat{proba}$ *are obtained from a mean on the different values of* d*.*

	Model	$\sqrt{MSE}$	$d=-0.4$	$d=-0.2$	$d=0$	$d=0.2$	$d=0.4$
$N=10^3 \longrightarrow$	fGn $(H=d+1/2)$	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.102	0.088	0.094	0.095	0.098
		$\sqrt{MSE}\ \widehat{d}_{R}$	0.091	0.108	0.106	0.117	0.090
		$\sqrt{MSE}\ \widehat{d}_{W}$	0.215	0.103	0.078	**0.073***	**0.061***
		$\sqrt{MSE}\ \widetilde{d}_{N}$	**0.071***	**0.074***	**0.075***	0.090	0.096
	FARIMA$(0,d,0)$	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.096	0.096	0.098	0.096	0.093
		$\sqrt{MSE}\ \widehat{d}_{R}$	0.094	0.113	0.107	0.112	0.084
		$\sqrt{MSE}\ \widehat{d}_{W}$	**0.069***	**0.073***	**0.074***	0.082	0.085
		$\sqrt{MSE}\ \widetilde{d}_{N}$	0.116	0.082	0.088	**0.081***	**0.081***
	FARIMA$(1,d,1)$	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.098	**0.092***	**0.089***	**0.088***	0.094
		$\sqrt{MSE}\ \widehat{d}_{R}$	**0.093***	0.110	0.115	0.110	**0.089***
		$\sqrt{MSE}\ \widehat{d}_{W}$	0.108	0.120	0.113	0.117	0.095
		$\sqrt{MSE}\ \widetilde{d}_{N}$	0.175	0.140	0.130	0.127	0.154
	$X^{(D,D')}$, $D'=1$	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.092	0.089	**0.113***	**0.107***	**0.100***
		$\sqrt{MSE}\ \widehat{d}_{R}$	0.093	0.111	0.129	0.124	0.111
		$\sqrt{MSE}\ \widehat{d}_{W}$	0.217	0.209	0.211	0.201	0.189
		$\sqrt{MSE}\ \widetilde{d}_{N}$	**0.070***	**0.088***	0.120	0.142	0.156

	Model	$\sqrt{MSE}$	$d=-0.4$	$d=-0.2$	$d=0$	$d=0.2$	$d=0.4$
$N=10^4 \longrightarrow$	fGn $(H=d+1/2)$	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.040	0.031	0.032	0.035	0.035
		$\sqrt{MSE}\ \widehat{d}_{R}$	0.040	0.027	0.029	0.031	0.030
		$\sqrt{MSE}\ \widehat{d}_{W}$	0.129	0.045	0.026	0.022	0.020
		$\sqrt{MSE}\ \widetilde{d}_{N}$	**0.023***	**0.022***	**0.017***	**0.018***	**0.017***
	FARIMA$(0,d,0)$	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.036	0.030	0.031	0.035	0.032
		$\sqrt{MSE}\ \widehat{d}_{R}$	0.031	0.028	0.027	0.029	0.029
		$\sqrt{MSE}\ \widehat{d}_{W}$	**0.020***	**0.018***	0.023	**0.025***	**0.028***
		$\sqrt{MSE}\ \widetilde{d}_{N}$	0.076	0.034	**0.018***	**0.025***	0.033
	FARIMA$(1,d,1)$	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.035	0.033	0.032	0.036	0.031
		$\sqrt{MSE}\ \widehat{d}_{R}$	**0.031***	**0.029***	**0.030***	**0.032***	**0.027***
		$\sqrt{MSE}\ \widehat{d}_{W}$	0.054	0.054	0.050	0.052	0.048
		$\sqrt{MSE}\ \widetilde{d}_{N}$	0.112	0.078	0.058	0.053	0.038
	$X^{(D,D')}$, $D'=1$	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.029	**0.037***	**0.035***	**0.041***	**0.038***
		$\sqrt{MSE}\ \widehat{d}_{R}$	0.032	0.041	0.037	**0.041***	0.039
		$\sqrt{MSE}\ \widehat{d}_{W}$	0.110	0.115	0.115	0.112	0.114
		$\sqrt{MSE}\ \widetilde{d}_{N}$	**0.019***	0.065	0.100	0.095	0.085

Table 2: *Comparison of the different log-memory parameter estimators for processes of the benchmark. For each process and value of d and N, $\sqrt{MSE}$ are computed from 100 independent generated samples.*

$N = 10^3 \longrightarrow$

Model+Innovation	$\sqrt{MSE}$	$d=-0.4$	$d=-0.2$	$d=0$	$d=0.2$	$d=0.4$
FARIMA$(0,d,0)$ Uniform	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.189	0.090	0.091	**0.082***	0.092
	$\sqrt{MSE}\ \widehat{d}_R$	0.171	0.104	0.109	0.102	0.086
	$\sqrt{MSE}\ \widehat{d}_W$	**0.111***	**0.066***	**0.072***	0.118	0.129
	$\sqrt{MSE}\ \widehat{\widetilde{d}}_N$	0.190	0.083	0.077	0.105	**0.083***
FARIMA$(0,d,0)$ Burr $(\alpha=2)$	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.181	0.087	0.092	0.084	**0.091***
	$\sqrt{MSE}\ \widehat{d}_R$	0.183	0.104	0.097	0.107	0.079
	$\sqrt{MSE}\ \widehat{d}_W$	**0.149***	**0.086***	0.130	0.101	0.129
	$\sqrt{MSE}\ \widehat{\widetilde{d}}_N$	0.230	0.106	**0.084***	**0.081***	0.110
FARIMA$(0,d,0)$ Cauchy	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.202	**0.080***	**0.069***	0.108	0.123
	$\sqrt{MSE}\ \widehat{d}_R$	**0.197***	0.093	0.088	**0.090***	**0.078***
	$\sqrt{MSE}\ \widehat{d}_W$	0.287	0.228	0.298	0.221	0.180
	$\sqrt{MSE}\ \widehat{\widetilde{d}}_N$	0.437	0.256	0.104	0.336	0.465
GARMA$(0,d,0)$	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.092	0.089	**0.113***	**0.107***	**0.100***
	$\sqrt{MSE}\ \widehat{d}_R$	0.093	0.111	0.129	0.124	0.111
	$\sqrt{MSE}\ \widehat{d}_W$	0.217	0.209	0.211	0.201	0.189
	$\sqrt{MSE}\ \widehat{\widetilde{d}}_N$	**0.070***	**0.088***	0.120	0.142	0.156
Trend	$\sqrt{MSE}\ \widehat{d}_{MS}$	1.307	0.891	0.538	0.290	0.150
	$\sqrt{MSE}\ \widehat{d}_R$	0.900	0.700	0.498	0.275	0.087
	$\sqrt{MSE}\ \widehat{d}_W$	**0.222***	**0.103***	0.083	**0.071***	0.059
	$\sqrt{MSE}\ \widehat{\widetilde{d}}_N$	1.79	0.588	**0.049***	0.089	**0.046***
Trend + Seasonality	$\sqrt{MSE}\ \widehat{d}_{MS}$	1.178	0.803	0.477	0.238	0.123
	$\sqrt{MSE}\ \widehat{d}_R$	0.900	0.700	0.498	0.284	**0.091***
	$\sqrt{MSE}\ \widehat{d}_W$	**0.628***	**0.407***	0.318	0.274	0.283
	$\sqrt{MSE}\ \widehat{\widetilde{d}}_N$	1.56	1.089	**0.301***	**0.151***	0.166

$N = 10^4 \longrightarrow$

Model+Innovation	$\sqrt{MSE}$	$d=-0.4$	$d=-0.2$	$d=0$	$d=0.2$	$d=0.4$
FARIMA$(0,d,0)$ Uniform	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.177	0.039	0.033	0.034	0.034
	$\sqrt{MSE}\ \widehat{d}_R$	0.171	0.032	0.030	**0.028***	**0.032***
	$\sqrt{MSE}\ \widehat{d}_W$	**0.125***	**0.027***	0.025	**0.028***	0.035
	$\sqrt{MSE}\ \widehat{\widetilde{d}}_N$	0.170	0.048	**0.019***	0.029	0.038
FARIMA$(0,d,0)$ Burr $(\alpha=2)$	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.18	0.036	0.041	0.033	0.032
	$\sqrt{MSE}\ \widehat{d}_R$	0.169	0.031	0.030	0.031	0.029
	$\sqrt{MSE}\ \widehat{d}_W$	**0.138***	0.068	0.065	0.076	0.066
	$\sqrt{MSE}\ \widehat{\widetilde{d}}_N$	0.218	0.075	**0.020***	0.037	0.071
FARIMA$(0,d,0)$ Cauchy	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.175	**0.028***	**0.013***	0.053	0.092
	$\sqrt{MSE}\ \widehat{d}_R$	**0.169***	0.025	**0.013***	**0.033***	**0.043***
	$\sqrt{MSE}\ \widehat{d}_W$	0.185	0.117	0.195	0.200	0.207
	$\sqrt{MSE}\ \widehat{\widetilde{d}}_N$	0.411	0.249	0.070	0.298	0.418
GARMA$(0,d,0)$	$\sqrt{MSE}\ \widehat{d}_{MS}$	0.092	0.089	**0.113***	**0.107***	**0.100***
	$\sqrt{MSE}\ \widehat{d}_R$	0.093	0.111	0.129	0.124	0.111
	$\sqrt{MSE}\ \widehat{d}_W$	0.217	0.209	0.211	0.201	0.189
	$\sqrt{MSE}\ \widehat{\widetilde{d}}_N$	**0.070***	**0.088***	0.120	0.142	0.156
Trend	$\sqrt{MSE}\ \widehat{d}_{MS}$	1.16	0.785	0.450	0.171	0.072
	$\sqrt{MSE}\ \widehat{d}_R$	0.900	0.700	0.431	0.192	0.067
	$\sqrt{MSE}\ \widehat{d}_W$	0.135	0.046	0.021	0.019	0.021
	$\sqrt{MSE}\ \widehat{\widetilde{d}}_N$	**0.018***	**0.017***	**0.020***	**0.017***	**0.017***
Trend + Seasonality	$\sqrt{MSE}\ \widehat{d}_{MS}$	1.219	0.841	0.474	0.194	0.099
	$\sqrt{MSE}\ \widehat{d}_R$	0.900	0.700	0.431	0.189	0.063
	$\sqrt{MSE}\ \widehat{d}_W$	**0.097***	**0.073***	0.063	0.065	**0.051***
	$\sqrt{MSE}\ \widehat{\widetilde{d}}_N$	0.811	0.152	**0.019***	**0.052***	0.057

Table 3: *Comparison of the different log-memory parameter estimators for processes of the benchmark. For each process and value of d and N, $\sqrt{MSE}$ are computed from 100 independent generated samples.*