Submitted to the Annals of Statistics

ASYMPTOTIC PROPERTIES OF U-PROCESSES UNDER LONG-RANGE DEPENDENCE

By C. Lévy-Leduc[*,‡], H. Boistard[§], E. Moulines[‡], M. S. Taqqu[¶,†] and V. A. Reisen[‖]

CNRS/LTCI/TelecomParisTech [‡], *Toulouse School of Economics* [§], *Boston University* [¶] *and Universidade Federal do Espírito Santo* [‖]

Let $(X_i)_{i\geq 1}$ be a stationary mean-zero Gaussian process with covariances $\rho(k) = \mathbb{E}(X_1 X_{k+1})$ satisfying: $\rho(0) = 1$ and $\rho(k) = k^{-D}L(k)$ where D is in $(0,1)$ and L is slowly varying at infinity. Consider the U-process $\{U_n(r),\ r \in I\}$ defined as

$$U_n(r) = \frac{1}{n(n-1)} \sum_{1\leq i\neq j\leq n} \mathbb{1}_{\{G(X_i,X_j)\leq r\}} ,$$

where I is an interval included in $\mathbb{R}$ and G is a symmetric function. In this paper, we provide central and non-central limit theorems for U_n. They are used to derive, in the long-range dependence setting, new properties of many well-known estimators such as the Hodges-Lehmann estimator, which is a well-known robust location estimator, the Wilcoxon-signed rank statistic, the sample correlation integral and an associated robust scale estimator. These robust estimators are shown to have the same asymptotic distribution as the classical location and scale estimators. The limiting distributions are expressed through multiple Wiener-Itô integrals.

1. Introduction.

Since the seminal work by Hoeffding (1948), U-statistics have been widely studied to investigate the asymptotic properties of many statistics such as the sample variance, the Gini's mean difference and the Wilcoxon one-sample statistic, see Serfling (1980) for other examples. One of the most powerful tools used to derive the asymptotic behavior of U-statistics is the Hoeffding's decomposition [Hoeffding (1948)]. In the i.i.d and

[*]Corresponding author

[†]Murad S. Taqqu would like to thank Telecom ParisTech for their hospitality. This research was partially supported by the NSF Grants DMS-0706786 and DMS-1007616 at Boston University.

AMS 2000 subject classifications: Primary 60F17, 62M10, 62G30, 62G20.

Keywords and phrases: Long-range dependence, U-process, Hodges-Lehmann estimator, Wilcoxon-signed rank test, sample correlation integral.

weak dependent frameworks, it provides a decomposition of a U-statistic into several terms having different orders of magnitudes, and in general the one with the leading order determines the asymptotic behavior of the U-statistic, see Serfling (1980), Borovkova, Burton and Dehling (2001) and the references therein for further details. A recent review of the properties of U-statistics in various frameworks is presented in Hsing and Wu (2004). In the case of processes having a long-range dependent structure, decomposition ideas are also crucial. However, in the case of Gaussian long-memory processes, the classical Hoeffding's decomposition may not provide the complete asymptotic behavior of U-statistics because all terms of this decomposition may contribute to the limit, see for example Dehling and Taqqu (1991). In this case, the asymptotic study of U-statistics can be achieved by using an expansion in Hermite polynomials, see Dehling and Taqqu (1989, 1991). For a large class of processes including linear and nonlinear processes, a new decomposition is discussed in Hsing and Wu (2004). These authors use martingale-based techniques to establish the asymptotic properties of U-statistics.

A very natural extension of U-statistics (which are random variables) is the notion of U-processes which encompasses a wide class of estimators. For example, Borovkova, Burton and Dehling (2001) study the Grassberger-Proccacia estimator which can be used to estimate the correlation dimension. In Section 5 of their work, the authors investigate the asymptotic properties of U-processes when the underlying observations are functionals of an absolutely regular process, that is, short-memory processes. As far as we know, the asymptotic properties of U-processes in the case of long-range dependence setting have not been established yet, and this is the heart of the research discussed in this paper. More precisely, our contribution consists first in extending the results of Borovkova, Burton and Dehling (2001) in order to address the long-range dependence case, second in extending the results obtained in Dehling and Taqqu (1989) to functions of two variables and third in extending the results of Hsing and Wu (2004) to U-processes. The authors of the latter paper establish the asymptotic properties of U-statistics involving causal but non necessarily Gaussian long-range dependent processes whereas, in our paper, we establish the asymptotic properties of U-processes involving Gaussian long-range dependent processes. The authors in Hsing and Wu (2004) use a martingale decomposition and we use a Hoeffding decomposition or a decomposition in Hermite polynomials. In the proof section, we also present an extension of some results of Soulier (2001).

Consider the U-process defined by

$$U_n(r) = \frac{1}{n(n-1)} \sum_{1 \leq i \neq j \leq n} \mathbb{1}_{\{G(X_i, X_j) \leq r\}} \ , \quad r \in I \tag{1}$$

where I is an interval included in $\mathbb{R}$, G is a symmetric function *i.e.* $G(x,y) = G(y,x)$ for all x, y in $\mathbb{R}$, and the process $(X_i)_{i \geq 1}$ satisfies the following assumption:

(A1) $(X_i)_{i\geq 1}$ is a stationary mean-zero Gaussian process with covariances $\rho(k) = \mathbb{E}(X_1 X_{k+1})$ satisfying:

$$\rho(0) = 1 \text{ and } \rho(k) = k^{-D} L(k), \ 0 < D < 1 \ ,$$

where L is slowly varying at infinity and is positive for large k.

Note that, for a fixed r, $U_n(r)$ is a U-statistic based on the kernel $h(\cdot, \cdot, r)$ where

$$h(x, y, r) = \mathbb{1}_{\{G(x,y) \leq r\}} \ , \forall x, y \in \mathbb{R} \text{ and } r \in I \ . \tag{2}$$

We show in this paper that the asymptotic properties of the U-process $U_n(\cdot)$ depends on the value of D and on the Hermite rank m of the class of functions $\{h(\cdot, \cdot, r) - U(r), r \in I\}$, defined in Section 2. We obtain the rate of convergence of $U_n(\cdot)$ and also provide the limiting process when $D > 1/2$, $m = 2$ and when $D < 1/m$, $m = 1, 2$. The convergence rate in the former case is of order $\sqrt{n}$ whereas it is of order $n^{mD/2}/L(n)^{m/2}$ in the latter. These results are stated in Theorems 1 and 2, respectively. They are applied to derive the asymptotic properties of well-known robust location and scale estimators such as the Hodges-Lehmann estimator [Hodges and Lehmann (1963)] and the Shamos scale estimator proposed by Shamos (1976) and analyzed by Bickel and Lehmann (1979). These properties are illustrated in Lévy-Leduc *et al.* (2010a) using numerical experiments. Theorems 1 and 2 allow us to establish novel asymptotic properties on these estimators in the long-range dependence context. The most striking result is that these robust estimators have the same asymptotic distribution as the classical estimators, see Propositions 5 and 8 in Section 4.

Theorems 1 and 2 have also been used to derive the asymptotic distribution of a robust scale estimator proposed by Rousseeuw and Croux (1993) and a robust autocovariance estimator introduced in Ma and Genton (2000). The robustness and efficiency properties of these estimators have also been investigated through numerical experiments and real data analysis. For further details on these theoretical and numerical studies, we refer the reader to Lévy-Leduc *et al.* (2010b).

The paper is organized as follows. In Section 2, the main theorems 1 and 2 are stated. In Section 3, we derive the asymptotic properties of some quantile estimators. Section 4 presents new asymptotic results in the context of long-range dependence. In this section, central and non-central limit theorems are provided for several statistics as an illustration of the theory presented in Sections 2 and 3. These statistics are the Hodges-Lehmann estimator [Hodges and Lehmann (1963)], the Wilcoxon-signed rank statistic [Wilcoxon (1945)], the sample correlation integral [Grassberger and Procaccia (1983)] and an associated scale estimator proposed by Shamos (1976) and Bickel and Lehmann (1979). Section 5 develops the proofs of the results stated in Section 2. A supplemental article Lévy-Leduc *et al.* (2010c) contains the proofs of some of the lemmas. It contains also Section 6 which concerns numerical experiments.

2. Main results.

We start by introducing the terms involved in the Hoeffding's decomposition [Hoeffding (1948)]. Recall the definition of $U_n(\cdot)$ in (1) and let $U(\cdot)$ be defined as

$$U(r) = \int_{\mathbb{R}^2} h(x, y, r)\varphi(x)\varphi(y)\mathrm{d}x\mathrm{d}y \ , \text{ for all } r \text{ in } I \ , \tag{3}$$

where φ denotes the p.d.f of a standard Gaussian random variable and h is given by (2). For all x in $\mathbb{R}$, and r in I, let us define

$$h_1(x, r) = \int_{\mathbb{R}} h(x, y, r)\varphi(y)\mathrm{d}y \ . \tag{4}$$

The Hoeffding decomposition amounts to expressing, for all r in I, the difference

$$U_n(r) - U(r) = \frac{1}{n(n-1)} \sum_{1 \leq i \neq j \leq n} [h(X_i, X_j, r) - U(r)] \ , \tag{5}$$

as

$$U_n(r) - U(r) = W_n(r) + R_n(r) \ , \tag{6}$$

where

$$W_n(r) = \frac{2}{n} \sum_{i=1}^{n} \{h_1(X_i, r) - U(r)\} \ , \tag{7}$$

and

$$R_n(r) = \frac{1}{n(n-1)} \sum_{1 \le i \ne j \le n} \{h(X_i, X_j, r) - h_1(X_i, r) - h_1(X_j, r) + U(r)\} \ . \tag{8}$$

We now define the Hermite rank of the class of functions $\{h(\cdot,\cdot,r) - U(r), r \in I\}$ which plays a crucial role in understanding the asymptotic behavior of the U-process $U_n(\cdot)$. We shall expand the function $(x, y) \mapsto h(x, y, r)$ in a Hermite polynomials basis of $L^2_\varphi(\mathbb{R}^2)$, that is, the L^2 space on $\mathbb{R}^2$ equipped with product standard Gaussian measures. We use Hermite polynomials with leading coefficients equal to one which are: $H_0(x) = 1$, $H_1(x) = x$, $H_2(x) = x^2 - 1$, $H_3(x) = x^3 - 3x, \dots$. We get

$$h(x, y, r) = \sum_{p,q \ge 0} \frac{\alpha_{p,q}(r)}{p!q!} H_p(x) H_q(y) \ , \text{ in } L^2_\varphi(\mathbb{R}^2) \ , \tag{9}$$

where

$$\alpha_{p,q}(r) = \mathbb{E}[h(X, Y, r) H_p(X) H_q(Y)] \ , \tag{10}$$

and where (X, Y) is a *standard Gaussian vector* that is X and Y are independent standard Gaussian random variables. Thus,

$$\mathbb{E}[h^2(X, Y, r)] = \sum_{p,q \ge 0} \frac{\alpha^2_{p,q}(r)}{p!q!} \ . \tag{11}$$

Note that $\alpha_{0,0}(r)$ is equal to $U(r)$ for all r, where $U(r)$ is defined in (3). The Hermite rank of $h(\cdot,\cdot,r)$ is the smallest positive integer $m(r)$ such that there exist p and q satisfying $p + q = m(r)$ and $\alpha_{p,q}(r) \ne 0$. Thus, (9) can be rewritten as

$$h(x, y, r) - U(r) = \sum_{\substack{p,q \ge 0 \\ p+q \ge m(r)}} \frac{\alpha_{p,q}(r)}{p!q!} H_p(x) H_q(y) \ , \text{ in } L^2_\varphi(\mathbb{R}^2) \ . \tag{12}$$

The Hermite rank m of the class of functions $\{h(\cdot,\cdot,r) - U(r) \ , r \in I\}$ is the smallest index $m = p + q \ge 1$ such that $\alpha_{p,q}(r) \ne 0$ for at least one r in I, that is, $m = \inf_{r \in I} m(r)$. By integrating with respect to y in (9), we obtain the expansion in Hermite polynomials of h_1 as a function of x:

$$h_1(x,r) - U(r) = \sum_{p \geq 1} \frac{\alpha_{p,0}(r)}{p!} H_p(x) \ , \ \text{in } L^2_\varphi(\mathbb{R}) \ , \tag{13}$$

where $L^2_\varphi(\mathbb{R})$ denotes the L^2 space on $\mathbb{R}$ equipped with the standard Gaussian measure. Let $\tau(r)$ be the smallest integer greater than or equal to 1 such that $\alpha_{\tau,0}(r) \neq 0$, that is, the Hermite rank of the function $h_1(\cdot,r) - U(r)$. The Hermite rank of the class of functions $\{h_1(\cdot,r) - U(r)\ , r \in I\}$ is the smallest index $\tau \geq 1$ such that $\alpha_{\tau,0}(r) \neq 0$ for at least one r. Since $\tau(r) \geq m(r)$, for all r in I, one has

$$\tau \geq m \ . \tag{14}$$

In the sequel, we shall assume that m is equal to 1 or 2. As shown in Section 4, this covers most of the situations of practical interest. Theorem 1, given below, establishes the central-limit theorem for the U-process $\{\sqrt{n}(U_n(r) - U(r)), r \in I\}$ when

$$D > 1/m \text{ and } m = 2 \ .$$

THEOREM 1. *Let I be a compact interval of $\mathbb{R}$. Suppose that the Hermite rank of the class of functions $\{h(\cdot,\cdot,r) - U(r)\ , r \in I\}$ as defined in (12) is $m = 2$ and that Assumption (A1) is satisfied with $D > 1/2$. Assume that h and h_1, defined in (2) and (4), satisfy the three following conditions:*

(i) There exists a positive constant C such that for all s, t in I, u, v in $\mathbb{R}$,

$$\mathbb{E}[|h(X+u, Y+v, s) - h(X+u, Y+v, t)|] \leq \ C|t-s| \ , \tag{15}$$

where (X,Y) is a standard Gaussian vector.

(ii) There exists a positive constant C such that for all $k \geq 1$,

$$\mathbb{E}[|h(X_1+u, X_{1+k}+v, t) - h(X_1, X_{1+k}, t)|] \leq \ C(|u| + |v|) \ , \tag{16}$$

$$\mathbb{E}[|h(X_1, X_{1+k}, s) - h(X_1, X_{1+k}, t)|] \leq C|t-s| \ . \tag{17}$$

(iii) There exists a positive constant C such that for all t, s in I, and x, u, v in $\mathbb{R}$,

$$|h_1(x+u,t) - h_1(x+v,t)| \leq C(|u| + |v|) \ , \tag{18}$$

and

$$|h_1(x,s) - h_1(x,t)| \leq C|t-s| \ . \tag{19}$$

Then the U-process

$$\{\sqrt{n}(U_n(r) - U(r)), r \in I\}$$

defined in (1) and (3) converges weakly in the space of cadlag functions $\mathcal{D}(I)$ equipped with the topology of uniform convergence to the zero mean Gaussian process $\{W(r), r \in I\}$ with covariance structure given by

$$\begin{aligned} \mathbb{E}[W(s)W(t)] &= 4\,\mathrm{Cov}(h_1(X_1,s), h_1(X_1,t)) \\ &+ 4\sum_{\ell \geq 1}\{\mathrm{Cov}(h_1(X_1,s), h_1(X_{\ell+1},t)) + \mathrm{Cov}(h_1(X_1,t), h_1(X_{\ell+1},s))\} \ . \end{aligned} \tag{20}$$

Proof of Theorem 1. The proof of the theorem follows from the decomposition (6) and Lemmas 9 and 10, given in Section 5.1. Lemma 9 states that $\{\sqrt{n}W_n(r), r \in I\}$ converges weakly in the space of cadlag functions $\mathcal{D}(I)$ equipped with the topology of uniform convergence. Lemma 10 states that $\sup_{r\in I}\sqrt{n}|R_n(r)| = o_P(1)$. Its proof uses Lemmas 11, 12 and 13. □

Remark 1. The examples of Section 4 satisfy the conditions (15) to (19), *e.g.* through the choice $G(x,y) = (x+y)/2$. More generally, suppose either:
(i) G is linear.
(ii) The function G can be written as $G(x,y) = g(L(x,y))$ where $L(x,y) = \alpha x + \beta y$ is some linear function of (x,y), α and β in $\mathbb{R}$ are such that $|\alpha| = |\beta|$ and g is an even function satisfying for some $\lambda_g > 0$: $\forall x,\ t \leq g(x) \leq s \implies \lambda_g t \leq |x| \leq \lambda_g s$.
(iii) $G \geq 0$ and satisfies the triangle inequality: $G(x+x', y+y') \leq G(x,y) + G(x',y')$ and there exists some constant C such that for all (x,y), $G(x,y) \leq C(|x|+|y|)$.
Then Condition (i) implies Conditions (15) to (19), Condition (ii) implies Conditions (15), (17) and (19) and Condition (iii) implies Conditions (16) and (18). The proofs are based on techniques similar to the verification of (27) in Section 4.1 where $G(x,y) = (x+y)/2$.

Remark 2. The set I in the previous theorem may be equal to $[-\infty, +\infty]$ which involves the two-point compactification of the real line. Since $[-\infty, +\infty]$

is compact, all functions in $\mathcal{D}([-\infty,+\infty])$ are bounded. In fact, that space is isomorphic to $\mathcal{D}[0,1]$.

When $D<1/m$, W_n and R_n are not the leading term and the remainder term, respectively. Note that, on one hand, for a fixed r, Corollary 2 of Dehling and Taqqu (1989) gives $R_n(r)=O_P(n^{-D}L(n))$ for any D in $(0,1)$. On the other hand, if $D<1/\tau$, where τ is defined in (14), Theorem 6 of Arcones (1994) implies that $W_n(r)=O_P(n^{-\tau D/2}L(n)^{\tau/2})$ and if D is in $(1/\tau,1/m)$, $W_n(r)=O_P(n^{-1/2})$ by Theorem 4 of Arcones (1994). Thus, if for instance, $\tau=m=2$, $W_n(r)$ and $R_n(r)$ may be of the same order $O_P(n^{-D}L(n))$. Hence, to study the case $D<1/m$, we shall introduce a different decomposition of $U_n(\cdot)$ based on the expansion of h in the basis of Hermite polynomials given by (9). Thus, $U_n(r)$ defined in (1) can be rewritten as follows

$$n(n-1)\{U_n(r)-U(r)\}=\widetilde{W}_n(r)+\widetilde{R}_n(r)\ , \tag{21}$$

where

$$\widetilde{W}_n(r)=\sum_{1\le i\neq j\le n}\ \sum_{\substack{p,q\ge 0\\ p+q\le m}}\frac{\alpha_{p,q}(r)}{p!q!}H_p(X_i)H_q(X_j)\ . \tag{22}$$

Introduce also the Beta function

$$\mathrm{B}(\alpha,\beta)=\int_0^\infty y^{\alpha-1}(1+y)^{-\alpha-\beta}\mathrm{d}y=\frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}\ ,\quad \alpha>0,\ \beta>0\ . \tag{23}$$

The limit processes which appear in the next theorem are the standard fractional Brownian motion (fBm) $(Z_{1,D}(t))_{0\le t\le 1}$ and the Rosenblatt process $(Z_{2,D}(t))_{0\le t\le 1}$. They are defined through multiple Wiener-Itô integrals and given by

$$Z_{1,D}(t)=\int_{\mathbb{R}}\left[\int_0^t(u-x)_+^{-(D+1)/2}\mathrm{d}u\right]\mathrm{d}B(x),\quad 0<D<1\ , \tag{24}$$

and

$$Z_{2,D}(t)=\int_{\mathbb{R}^2}'\left[\int_0^t(u-x)_+^{-(D+1)/2}(u-y)_+^{-(D+1)/2}\mathrm{d}u\right]\mathrm{d}B(x)\mathrm{d}B(y),\quad 0<D<1/2\ , \tag{25}$$

where B is the standard Brownian motion, see Fox and Taqqu (1987). The symbol $\int'$ means that the domain of integration excludes the diagonal. Note that $Z_{1,D}$ and $Z_{2,D}$ are dependent but uncorrelated. The following theorem treats the case $D<1/m$ where $m=1$ or 2.

THEOREM 2. *Let I be a compact interval of $\mathbb{R}$. Suppose that Assumption (A1) holds with $D < 1/m$, where $m = 1$ or 2 is the Hermite rank of the class of functions $\{h(\cdot,\cdot,r) - U(r)\ , r \in I\}$ as defined in (12). Assume the following:*

(i) There exists a positive constant C such that, for all $k \geq 1$ and for all s, t in I,

$$\mathbb{E}[|h(X_1, X_{1+k}, s) - h(X_1, X_{1+k}, t)|] \leq C|t - s| \ . \tag{26}$$

(ii) U is a Lipschitz function.

(iii) The function $\widetilde{\Lambda}$ defined, for all s in I, by

$$\widetilde{\Lambda}(s) = \mathbb{E}[h(X, Y, s)(|X| + |XY| + |X^2 - 1|)] \ , \tag{27}$$

where X and Y are independent standard Gaussian random variables, is also a Lipschitz function.

Then,

$$\left\{n^{mD/2} L(n)^{-m/2} \left(U_n(r) - U(r)\right); r \in I\right\}$$

converges weakly in the space of cadlag functions $\mathcal{D}(I)$, equipped with the topology of uniform convergence, to

$$\{2\alpha_{1,0}(r) k(D)^{-1/2} Z_{1,D}(1); r \in I\} \ , \quad \textit{if } m = 1 \ ,$$

and to

$$\{k(D)^{-1}\left[\alpha_{1,1}(r) Z_{1,D}(1)^2 + \alpha_{2,0}(r) Z_{2,D}(1)\right]; r \in I\} \ , \quad \textit{if } m = 2 \ ,$$

where the fractional Brownian motion $Z_{1,D}(\cdot)$ and the Rosenblatt process $Z_{2,D}(\cdot)$ are defined in (24) and (25) respectively and where

$$k(D) = \mathrm{B}((1 - D)/2, D) \ , \tag{28}$$

where B is the Beta function defined in (23).

The proof of Theorem 2 is given in Section 5.3.

REMARK 3. The processes $Z_{1,D}$ and $Z_{2,D}$ are self-similar with mean 0. They are, however, not normalized. One has

$$\mathbb{E}[Z_{1,D}(t) Z_{1,D}(s)] = \mathbb{E}[Z_{1,D}^2(1)] \frac{1}{2} \left\{t^{2H_1} + s^{2H_1} - |t - s|^{2H_1}\right\} \ ,$$

$$\mathbb{E}[Z_{2,D}(t)Z_{2,D}(s)] = \mathbb{E}[Z_{2,D}^2(1)]\frac{1}{2}\left\{t^{2H_2} + s^{2H_2} - |t-s|^{2H_2}\right\} ,$$

where $H_1 = 1 - D/2 \in (0, 1/2)$, $H_2 = 1 - D \in (0, 1/2)$ and

$$\mathbb{E}[Z_{1,D}^2(1)] = \frac{2k(D)}{(-D+1)(-D+2)} , \tag{29}$$

$$\mathbb{E}[Z_{2,D}^2(1)] = \frac{4k(D)^2}{(-2D+1)(-2D+2)} , \tag{30}$$

with $k(D)$ defined by (28). See Remark 5 below for justification. The non-Gaussian random variables $Z_{1,D}^2(1)$ and $Z_{2,D}(1)$ are dependent. Their joint cumulants are given in (89) in the supplemental article Lévy-Leduc *et al.* (2010c).

REMARK 4. The results of Theorem 2 can be extended to the two-parameter process $\{U_{[nt]}(r) - U(r); r \in I, 0 \le t \le 1\}$. One can show that

$$\left\{\frac{n^{mD/2}}{L(n)^{m/2}}\left(U_{[nt]}(r) - U(r)\right);\ r \in I, 0 \le t \le 1\right\}$$

converges weakly in $\mathcal{D}(I \times [0, 1])$, equipped with the topology of uniform convergence, to

$$\{2\alpha_{1,0}(r)k(D)^{-1/2}Z_{1,D}(t);\ r \in I, 0 \le t \le 1\} , \quad \text{if } m = 1 ,$$

and to

$$\{k(D)^{-1}\left[\alpha_{1,1}(r)Z_{1,D}(t)^2 + \alpha_{2,0}(r)Z_{2,D}(t)\right];\ r \in I, 0 \le t \le 1\} , \quad \text{if } m = 2 .$$

3. Asymptotic behavior of empirical quantiles. We shall apply Theorems 1 and 2 in the preceding section to empirical quantiles. Recall that if $V : I \longrightarrow [0, 1]$ is a non-decreasing cadlag function, where I is an interval of $\mathbb{R}$, then its generalized inverse V^{-1} is defined by $V^{-1}(p) = \inf\{r \in I,\ V(r) \ge p\}$. This applies to $U_n(r)$ and $U(r)$ since these are non-decreasing functions of r. We derive in the following corollaries the asymptotic behavior of the empirical quantile $U_n^{-1}(\cdot)$ using Theorems 1 and 2.

COROLLARY 3. *Let p be a fixed real number in $(0, 1)$. Assume that the conditions of Theorem 1 are satisfied. Suppose also that there exists some r*

in I such that $U(r) = p$, that U is differentiable at r and that $U'(r)$ is non null. Then, as n tends to infinity,

$$\sqrt{n}(U_n^{-1}(p) - U^{-1}(p)) \xrightarrow{d} -W(U^{-1}(p))/U'(U^{-1}(p)) \ ,$$

where W is a Gaussian process having a covariance structure given by (20).

Proof of Corollary 3. By Lemma 21.3 in van der Vaart (1998), the functional $T : V \mapsto V^{-1}(p)$ is Hadamard differentiable at V tangentially to the set of functions h in $\mathcal{D}([0,1])$ with derivative $T'_V(h) = -h(V^{-1}(p))/V'(V^{-1}(p))$. Applying the functional Delta method (Theorem 20.8 in van der Vaart (1998)) thus yields

$$\begin{aligned}\sqrt{n}(U_n^{-1}(p) - U^{-1}(p)) &= T'_U\{\sqrt{n}(U_n - U)\} + o_P(1) \\ &= -\sqrt{n}\frac{(U_n - U)(U^{-1}(p))}{U'(U^{-1}(p))} + o_P(1) \ .\end{aligned}$$

The corollary then follows from Theorem 1. $\square$

Corollary 4. *Let p be a fixed real number in $(0,1)$. Assume that the conditions of Theorem 2 are satisfied. Suppose also that there exists some r in I such that $U(r) = p$, that U is differentiable at r and that $U'(r)$ is non null. Then, as n tends to infinity,*

$$\frac{n^{mD/2}}{L(n)^{m/2}}(U_n^{-1}(p) - U^{-1}(p))$$

converges in distribution to

$$-2k(D)^{-1/2}\frac{\alpha_{1,0}(U^{-1}(p))}{U'(U^{-1}(p))}Z_{1,D}(1) \ , \ \text{if } m = 1 \ ,$$

and to

$$-k(D)^{-1}\left\{\alpha_{1,1}(U^{-1}(p))Z_{1,D}(1)^2 + \alpha_{2,0}(U^{-1}(p))Z_{2,D}(1)\right\}/U'(U^{-1}(p)) \ , \ \text{if } m = 2 \ ,$$

where $Z_{1,D}(\cdot)$ and $Z_{2,D}(\cdot)$ are defined in (24) and (25) respectively, $k(D)$ in (28) and $\alpha_{p,q}(\cdot)$ is defined in (10).

The proof of Corollary 4 is based on similar arguments as the proof of Corollary 3 and is thus omitted.

4. Applications. We shall use the results established in Sections 2 and 3 to study the asymptotic properties of several estimators based on U-processes in the long-range dependence setting.

4.1. *Hodges-Lehmann estimator.* Consider the problem of estimating the location parameter of a long-range dependent Gaussian process. Assume that $(Y_i)_{i\geq 1}$ satisfy $Y_i = \theta + X_i$ where $(X_i)_{i\geq 1}$ satisfy Assumption (A1). To estimate the location parameter θ, Hodges and Lehmann (1963) suggest using the median of the average of all pairs of observations. The statistic they propose is

$$\hat{\theta}_{HL} = \text{median}\left\{\frac{Y_i + Y_j}{2}; 1 \leq i < j \leq n\right\}$$
$$= \theta + \text{median}\left\{\frac{X_i + X_j}{2}; 1 \leq i < j \leq n\right\} .$$

Define the U-process $U_n(r)$, $r \in \mathbb{R}$ by (1), where $G(x, y) = (x + y)/2$. The Hodges-Lehmann estimator may be then expressed as

$$\hat{\theta}_{HL} = \theta + U_n^{-1}(1/2) .$$

If A and B are independent standard Gaussian variables,

$$\alpha_{1,0}(r) = \alpha_{0,1}(r) = \mathbb{E}[A \mathbb{1}_{\{A+B\leq 2r\}}] = -\int_{\mathbb{R}} \varphi(2r-y)\varphi(y)\mathrm{d}y = -\varphi(r\sqrt{2})/\sqrt{2} , \tag{31}$$

using $x\varphi(x) = -\dot{\varphi}(x)$, where $\dot{\varphi}$ denotes the first derivative of φ. The quantities in (31) are different from 0 for all r in $\mathbb{R}$ since φ is the p.d.f of a standard Gaussian random variable. Thus, the Hermite rank m of the class of functions $\{\mathbb{1}_{G(\cdot,\cdot)\leq r} - \alpha_{0,0}(r); r \in \mathbb{R}\}$ is equal to 1. In order to derive the asymptotic properties of $\hat{\theta}_{HL}$, we now check the conditions of Theorem 2. Let us check Condition (26). Note that for all $k \geq 1$, $X_1 + X_{1+k} \sim \mathcal{N}(0, 2(1+\rho(k)))$, thus if $t \leq s$,

$$\mathbb{E}[h(X_1, X_{1+k}, s) - h(X_1, X_{1+k}, t)] = \Phi\left(\frac{\sqrt{2}s}{\sqrt{1+\rho(k)}}\right) - \Phi\left(\frac{\sqrt{2}t}{\sqrt{1+\rho(k)}}\right)$$
$$\leq \frac{1}{\sqrt{\pi}} \frac{|t-s|}{\sqrt{1+\rho_\star}} ,$$

where Φ is the c.d.f of a standard Gaussian random variable and $\rho_\star = \inf_k \rho(k) > -1$. Hence (26) holds. Similarly, $|U(s) - U(t)| \leq |\Phi(\sqrt{2}s) -$

$\Phi(\sqrt{2}t)| \leq \pi^{-1/2}|t-s|$ and hence U is a Lipschitz function. Let us now check Condition (27). Note that, if $s \leq t$

$$\int\int \mathbb{1}_{\{s<x+y\leq t\}}(|x|+|xy|+|x^2-1|)\varphi(x)\varphi(y)\mathrm{d}x\mathrm{d}y = \int\left(\int_{s-x}^{t-x}\varphi(y)\mathrm{d}y\right)|x|\varphi(x)\mathrm{d}x + \int\left(\int_{s-x}^{t-x}|y|\varphi(y)\mathrm{d}y\right)|x|\varphi(x)\mathrm{d}x + \int\left(\int_{s-x}^{t-x}\varphi(y)\mathrm{d}y\right)|x^2-1|\varphi(x)\mathrm{d}x\ .$$

Using that $\varphi(\cdot)$ and $|.|\varphi(\cdot)$ are bounded and that the moments of Gaussian random variables are all finite, we get (27). The assumptions of Theorem 2 are thus satisfied with $m=1$ and hence we get that

$$\left\{n^{D/2}L(n)^{-1/2}\left(U_n(r)-U(r)\right); -\infty \leq r \leq +\infty\right\}$$

converges weakly in $D([-\infty,+\infty])$, equipped with the sup-norm, to

$$\{-\sqrt{2}k(D)^{-1/2}\varphi(r\sqrt{2})Z_{1,D}(1); -\infty \leq r \leq +\infty\}\ .$$

Here, $U(r) = \int\Phi(2r-x)\varphi(x)\mathrm{d}x$, $U'(r) = 2\int\varphi(2r-x)\varphi(x)\mathrm{d}x$, $U(0) = 1/2\int(\Phi(x)+\Phi(-x))\varphi(x)\mathrm{d}x = 1/2$, $U^{-1}(1/2) = 0$ and $U'(U^{-1}(1/2)) = U'(0) = 1/\sqrt{\pi}$. Since, by (31), $\alpha_{1,0}(U^{-1}(1/2)) = \alpha_{1,0}(0) = -(2\sqrt{\pi})^{-1}$, Corollary 4 implies that

$$n^{D/2}L(n)^{-1/2}(\hat{\theta}_{HL}-\theta) \xrightarrow{d} k(D)^{-1/2}Z_{1,D}(1)\ , \tag{32}$$

where using (29), $k(D)^{-1/2}Z_{1,D}(1)$ is a zero-mean Gaussian random variable with variance $2(-D+1)^{-1}(-D+2)^{-1}$. Let's now compare the asymptotic behavior of the Hodges-Lehmann estimator with that of the sample mean. Lemma 5.1 in Taqqu (1975) shows that the sample mean $\bar{Y}_n = n^{-1}\sum_{i=1}^n Y_i$ satisfies the following central limit theorem

$$n^{D/2}L(n)^{-1/2}(\bar{Y}_n-\theta) \xrightarrow{d} k(D)^{-1/2}Z_{1,D}(1)\ .$$

We have thus proved

PROPOSITION 5. *In the long-memory framework with $0<D<1$, the asymptotic behavior of the Hodges-Lehmann estimator is Gaussian and given by (32). It converges to θ at the same rate as the sample mean with the same limiting distribution. There is no loss of efficiency.*

A similar result was proved in Beran (1991) for location M-estimators.

4.2. *Wilcoxon-signed rank statistic.* Assume that $(Y_i)_{i\geq 1}$ satisfy $Y_i = \theta + X_i$ where $(X_i)_{i\geq 1}$ satisfy Assumption (A1). The Wilcoxon-signed rank statistic first proposed by Wilcoxon (1945) can be used to test the null hypothesis (H_0): " $\theta = 0$ " against the one-sided alternative (H_1): " $\theta > 0$ ", based on the observations $Y_1, \ldots, Y_n$. It is defined as

$$T_n = \sum_{j=1}^{n} R_j \mathbb{1}_{\{X_j>0\}} ,$$

where the R_j's are the ranks of $X_1, \ldots, X_n$. Thus T_n is the sum of the ranks of the positive observations. Let us study this statistic under the null hypothesis. One will reject the null hypothesis if the value of T_n is large. Following Dewan and Prakasa Rao (2005), T_n can be written as

$$T_n = \sum_{i=1}^{n} \mathbb{1}_{\{X_i>0\}} + \sum_{1\leq i<j\leq n} \mathbb{1}_{\{X_i+X_j>0\}} =: nU_{n,1} + \frac{n(n-1)}{2} U_{n,2} . \tag{33}$$

The Hermite rank of $\mathbb{1}_{\{\cdot>0\}} - \mathbb{P}(X_1 > 0)$ equals 1, because $\mathbb{E}[X_1(\mathbb{1}_{\{X_1>0\}} - \mathbb{P}(X_1 > 0))] > 0$. We then deduce from Theorem 6 of Arcones (1994) that

$$n^{D/2} L(n)^{-1/2} (U_{n,1} - \mathbb{P}(X_1 > 0)) = O_p(1) . \tag{34}$$

The asymptotic properties of $U_{n,2}$ can be derived from those of $U_n(0)$ where $U_n(\cdot)$ is the U-process defined in (1) with $G(x,y) = x + y$. Using the results obtained in the study of the Hodges-Lehmann estimator, we obtain that $\alpha_{1,0}(r) = \alpha_{0,1}(r) = -\varphi(r/\sqrt{2})/\sqrt{2}$, which is different from 0 for all r in $\mathbb{R}$ since φ is the p.d.f of a standard Gaussian random variable. Thus, the Hermite rank of the class of functions $\{\mathbb{1}_{G(\cdot,\cdot)\leq r} - \alpha_{0,0}(r); r \in \mathbb{R}\}$ is equal to 1. Using the same arguments as those used in the previous example, the assumptions of Theorem 2 are fulfilled with $m = 1$. Since $2\alpha_{1,0}(0) = -2\varphi(0)/\sqrt{2} = -1/\sqrt{\pi}$, we get

$$n^{D/2} L(n)^{-1/2} \left(U_{n,2} - U_2(0)\right) \xrightarrow{d} \frac{k(D)^{-1/2}}{\sqrt{\pi}} Z_{1,D}(1) , \tag{35}$$

where $U_2(0) = \int\int \mathbb{1}_{\{x+y>0\}}\varphi(x)\varphi(y)\mathrm{d}x\mathrm{d}y = 1/2$ and $k(D)$ is the constant given in (28). From (33), (34) and (35), we get

$$\begin{aligned}\frac{2n^{D/2}}{n(n-1)L(n)^{1/2}}&(T_n - n\mathbb{P}(X_1 > 0) - n(n-1)U_2(0)/2) \\ &= \frac{n^{D/2}}{L(n)^{1/2}} \left(U_{n,2} - U_2(0)\right) + o_p(1) \xrightarrow{d} \frac{k(D)^{-1/2}}{\sqrt{\pi}} Z_{1,D}(1) ,\end{aligned}$$

which can be rewritten as follows

$$n^{D/2}L(n)^{-1/2}\left(\frac{2}{n(n-1)}T_n-\frac{1}{n-1}-1/2\right)\xrightarrow{d}\frac{k(D)^{-1/2}}{\sqrt{\pi}}Z_{1,D}(1)\ . \tag{36}$$

We have thus proved

PROPOSITION 6. *In the long-memory case with $0<D<1$, the asymptotic behavior of the Wilcoxon-signed rank statistic T_n is Gaussian and given by (36).*

The Wilcoxon one-sample statistic $U_{n,2}$ was also studied by Hsing and Wu (2004) pp. 1617-1618 by using a different approach. We obtain the additional constant $k(D)^{-1/2}$ in the limiting distribution compared to their result.

4.3. *Sample correlation integral.* In the past few years, a lot of attention has been paid to the estimation of the correlation dimension of a strange attractor. In many examples, the correlation dimension α of an invariant probability measure μ can be expressed through the *correlation integral* $C_\mu(r)=(\mu\times\mu)\{(x,y):\ |x-y|\leq r\}$ through $C_\mu(r)\approx Cr^\alpha$, as r tends to 0, where C is a constant. For further details on the correlation dimension and its applications, see Borovkova, Burton and Dehling (2001). Grassberger and Procaccia (1983) proposed an estimator of the correlation dimension based on the sample correlation integral $U_n(r)$, defined in (1) with $G(x,y)=|x-y|$. In this case, $\alpha_{1,0}(r)=\alpha_{0,1}(r)=\int_{\mathbb{R}}x\mathbb{1}_{\{|x-y|\leq r\}}\varphi(x)\varphi(y)\mathrm{d}x\mathrm{d}y$ $=\int_{\mathbb{R}}x[\Phi(x+r)-\Phi(x-r)]\varphi(x)\mathrm{d}x$, where, as before, φ and Φ are the p.d.f. and the c.d.f. of a standard Gaussian random variable, respectively. Using the symmetry of a standard Gaussian random variable, one gets $\alpha_{1,0}(r)=\alpha_{0,1}(r)=0$. Lengthy but straightforward computations lead to

$$\alpha_{2,0}(r)=\alpha_{0,2}(r)=-\alpha_{1,1}(r)=\dot{\varphi}(r/\sqrt{2})\ , \tag{37}$$

where $\dot{\varphi}$ denotes the first derivative of φ. It is non-null if $r\neq 0$. Thus, for any compact interval I which does not contain 0, the Hermite rank of the class of functions $\{\mathbb{1}_{G(\cdot,\cdot)\leq r}-\alpha_{0,0}(r),r\in I\}$ is equal to 2. Let us assume that $(X_i)_{i\geq 1}$ satisfy Assumption (A1). In the case where $D>1/2$, let us check the assumptions of Theorem 1. Conditions (15) and (16) can be easily checked and Condition (17) is fulfilled by using similar arguments as those used in the example of the Hodges-Lehmann estimator. Conditions (18) and (19) are satisfied since

$$h_1(x, r) = \int_{\mathbb{R}} \mathbb{1}_{\{|x-y| \leq r\}} \varphi(y) \mathrm{d}y = \Phi(x+r) - \Phi(x-r) , \tag{38}$$

where Φ is the c.d.f of a standard Gaussian random variable. Thus, in the case where $D > 1/2$,

$$\{\sqrt{n}(U_n(r) - U(r)), r \in I\}$$

converges weakly in $\mathcal{D}(I)$, equipped with the topology of uniform convergence, to the zero mean Gaussian process $\{W(r), r \in I\}$ with covariance structure given by

$$\begin{aligned}\mathbb{E}[W(s)W(t)] &= 4\,\mathrm{Cov}(h_1(X_1, s), h_1(X_1, t)) \\ &+ 4 \sum_{\ell \geq 1} \{\mathrm{Cov}(h_1(X_1, s), h_1(X_{\ell+1}, t)) + \mathrm{Cov}(h_1(X_1, t), h_1(X_{\ell+1}, s))\} ,\end{aligned} \tag{39}$$

where h_1 is given in (38). If $D < 1/2$, with similar arguments as those used in the example on the Hodges-Lehmann estimator, the assumptions of Theorem 2 are satisfied with $m = 2$ and we get using (37), that

$$\left\{k(D) n^D L(n)^{-1} \left(U_n(r) - U(r)\right); r \in I\right\}$$

converges weakly in $\mathcal{D}(I)$, equipped with the topology of uniform convergence, to

$$\{\dot{\varphi}(r/\sqrt{2})(Z_{2,D}(1) - Z_{1,D}(1)^2); r \in I\} . \tag{40}$$

where I is any compact set of $\mathbb{R}$ which does not contain 0. Thus

PROPOSITION 7. *In the long-memory case with $1/2 < D < 1$, the asymptotic behavior of the sample correlation integral $U_n(r)$, $r \in I$ is Gaussian with covariance (39). If $0 < D < 1/2$ and if I is a compact set in $\mathbb{R}$ which does not contain 0, then the limit is non-Gaussian and given in (40).*

4.4. *Shamos scale estimator.* Assume that $(Y_i)_{i \geq 1}$ satisfy $Y_i = \sigma X_i$ where $(X_i)_{i \geq 1}$ satisfy Assumption (A1). The results of the previous subsection can be used to derive the properties of the estimator of the scale σ proposed by Shamos (1976) and Bickel and Lehmann (1979). From $Y_1, \dots, Y_n$, it is defined by

$$\hat{\sigma}_{BL} = c \text{ median}\{|Y_i - Y_j|; 1 \le i < j \le n\} = c\,\sigma \text{ median}\{|X_i - X_j|; 1 \le i < j \le n\} ,$$

where $c = 1/(\sqrt{2}\Phi^{-1}(1/4) \approx 1.0483$ and Φ is the c.d.f of a standard Gaussian random variable to achieve consistency for σ in the case of Gaussian distribution. $\hat{\sigma}_{BL}$ involves the median of the distance between observations. As is the case for the standard deviation, if the Y_i's are transformed into $aY_i + b$, then $\hat{\sigma}_{BL}$ is multiplied by $|a|$. Here $G(x, y) = |x - y|$, $U(r) = \int[\Phi(x+r) - \Phi(x-r)]\varphi(x)\mathrm{d}x$, $U'(r) = 2\int \varphi(x+r)\varphi(x)\mathrm{d}x$, $U^{-1}(1/2) = 1/c$ and $U'(U^{-1}(1/2)) = U'(1/c) = \sqrt{2}\varphi(1/(c\sqrt{2}))$. By Corollary 3, we obtain that for $D > 1/2$,

$$\sqrt{n}(\hat{\sigma}_{BL} - \sigma) \xrightarrow{d} -\frac{c\sigma W(1/c)}{\sqrt{2}\varphi(1/(c\sqrt{2}))} , \tag{41}$$

where W is a Gaussian process having the covariance structure (20) with h_1 given in (38). Consider now the case $D < 1/2$. By (37), $\alpha_{2,0}(U^{-1}(1/2)) = -\alpha_{1,1}(U^{-1}(1/2)) = -\alpha_{1,1}(1/c) = \dot{\varphi}(1/(c\sqrt{2}))$. Hence, we deduce from Corollary 4 that, if $D < 1/2$,

$$\begin{aligned} k(D)n^D L(n)^{-1}(\hat{\sigma}_{BL} - \sigma) &\xrightarrow{d} \frac{c\sigma\dot{\varphi}(1/(c\sqrt{2}))}{\sqrt{2}\varphi(1/(c\sqrt{2}))}(Z_{1,D}(1)^2 - Z_{2,D}(1)) \\ &= \frac{\sigma}{2}(Z_{2,D}(1) - Z_{1,D}(1)^2) . \end{aligned} \tag{42}$$

Let us now compare the asymptotic behavior of the Shamos scale estimator with that of the square root of the sample variance estimator, $\hat{\sigma}_{n,Y} = (\sum_{i=1}^n (Y_i - \bar{Y})^2/(n-1))^{1/2}$. We have

$$n(n-1)(\hat{\sigma}_{n,Y}^2 - \sigma^2) = \sigma^2[n\sum_{i=1}^n (X_i^2 - 1) - \sum_{1 \le i,j \le n} X_i X_j + n] ,$$

so that by Lemma 15,

$$k(D)n^D L(n)^{-1}(\hat{\sigma}_{n,Y}^2 - \sigma^2) \xrightarrow{d} \sigma^2(Z_{2,D}(1) - Z_{1,D}(1)^2) .$$

We apply the Delta method to go from σ^2 to σ, setting $f(x) = \sqrt{x}$, so that $f'(\sigma^2) = 1/(2\sqrt{\sigma^2}) = 1/(2\sigma)$. We obtain

$$k(D)n^D L(n)^{-1}(\hat{\sigma}_{n,Y} - \sigma) \xrightarrow{d} \frac{\sigma}{2}(Z_{2,D}(1) - Z_{1,D}(1)^2) . \tag{43}$$

Thus,

PROPOSITION 8. *In the long-memory case with $1/2 < D < 1$, the asymptotic behavior of the Shamos scale estimator $\hat{\sigma}_{BL}$ is Gaussian and given in (41). If $0 < D < 1/2$, it is non-Gaussian and given by (42); in this case, $\hat{\sigma}_{BL}$ converges to σ at the same rate as the square root of the sample variance estimator with no loss of efficiency.*

5. Proofs of Theorems 1 and 2.

5.1. *Lemmas used in the proof of Theorem 1.*

LEMMA 9. *Under the assumptions of Theorem 1, the process $\{\sqrt{n}W_n(r), r \in I\}$, where $W_n(\cdot)$ is defined in (7), converges weakly in the space of cadlag functions equipped with the topology of uniform convergence to the zero mean Gaussian process $\{W(r), r \in I\}$ with covariance structure given by*

$$\begin{aligned}\mathbb{E}[W(s)W(t)] &= 4\,\mathrm{Cov}(h_1(X_1,s),h_1(X_1,t)) \\ &\quad + 4\sum_{\ell\geq 1}\{\mathrm{Cov}(h_1(X_1,s),h_1(X_{\ell+1},t)) + \mathrm{Cov}(h_1(X_1,t),h_1(X_{\ell+1},s))\}\ .\end{aligned}$$

The proof of Lemma 9 is in the supplemental article Lévy-Leduc *et al.* (2010c).

LEMMA 10. *Under the assumptions of Theorem 1,*

$$\sup_{r\in I}\sqrt{n}|R_n(r)| = o_P(1)\ ,$$

where R_n is defined in (8).

The proof of Lemma 10 is in the supplemental article Lévy-Leduc *et al.* (2010c).

LEMMA 11. *Under the assumptions of Theorem 1, there exist positive constants C and α such that, for large enough n,*

$$\mathbb{E}[\{R_n(t) - R_n(s)\}^2] \leq C\,\frac{|t-s|}{n^{1+\alpha}}\ ,\ \text{for all } s,t\in I\ , \tag{44}$$

where R_n is defined in (8).

The proof of Lemma 11 can be found at the end of this subsection and is based on the following lemmas 12, 13 and 14.

LEMMA 12. *Let $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ be a bounded function such that its derivative $\partial^6 f/\partial x^3 \partial y^3$ exists. Let (X, Y) be a standard Gaussian random vector. Assume that $\mathbb{E}\left[\left(\partial^{i+j} f(X,Y)/\partial x^i \partial y^j\right)^2\right] < \infty$, for all $1 \leq i, j \leq 3$, then the Hermite coefficients of f defined by $c_{p,q}(f) := \mathbb{E}[f(X,Y)H_p(X)H_q(Y)]$ satisfy, for $p, q \geq 3$*

$$|c_{p,q}(f)| \leq \mathbb{E}[(\partial^6 f(X,Y)/\partial x^3 \partial y^3)^2]^{1/2} \sqrt{(p-3)!}\sqrt{(q-3)!}\ . \tag{45}$$

The proof of Lemma 12 is in the supplemental article Lévy-Leduc *et al.* (2010c). The following lemma is an extension of Corollary 2.1 in Soulier (2001) and is proved in the supplemental article Lévy-Leduc *et al.* (2010c).

LEMMA 13. *Let f_1 and f_2 be two functions defined on $\mathbb{R}^{a_1}$ and $\mathbb{R}^{a_2}$, respectively. Let Γ be the covariance matrix of the mean-zero Gaussian vector $Y = (Y_1, Y_2)$ where Y_1 and Y_2 are in $\mathbb{R}^{a_1}$ and $\mathbb{R}^{a_2}$, respectively. Assume that there exists a block diagonal matrix Γ_0 of size $(a_1+a_2)\times(a_1+a_2)$ built from Γ with diagonal blocks $\Gamma_{0,1}$ and $\Gamma_{0,2}$ of size $a_1 \times a_1$ and $a_2 \times a_2$, respectively, such that $r^\star := \|\Gamma_0^{-1/2}(\Gamma_0 - \Gamma)\Gamma_0^{-1/2}\|_2 \leq (1/3 - \varepsilon)$, for some positive ε. In the previous inequality $\|B\|_2$ denotes the spectral radius of the symmetric matrix B. If at least one function f_i has an Hermite rank larger than τ, then there exists a positive constant $C(a_1, a_2, \varepsilon)$ such that*

$$|\mathbb{E}[f_1(Y_1)f_2(Y_2)]| \leq C(a_1, a_2, \varepsilon)\|f_1\|_{2,\Gamma_{0,1}}\|f_2\|_{2,\Gamma_{0,2}}(r^\star)^{[(\tau+1)/2]}\ , \tag{46}$$

where $\|f_i\|^2_{2,\Gamma_{0,i}} = (2\pi)^{-a_i/2}|\Gamma_{0,i}|^{-1/2}\int_{\mathbb{R}^{a_i}} f_i^2(x)\exp(-x^T\Gamma_{0,i}^{-1}x/2)\mathrm{d}x$, $i = 1, 2$ and $[x]$ denotes the integer part of x.

We shall use the following notation: for a Gaussian vector (X_1, X_2, X_3, X_4) with covariance matrix Γ and for any real-valued function of this vector, the expected value $\mathbb{E}[f(X_1, X_2, X_3, X_4)]$ will be denoted by $\mathbb{E}_\Gamma[f(X_1, X_2, X_3, X_4)]$.

LEMMA 14. *Let (X_1, X_2, X_3, X_4) be a Gaussian vector with mean 0 and covariance matrix*

$$\Gamma = \begin{bmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{bmatrix} = \left[\begin{array}{cc|cc} 1 & \rho_{12} & \rho_{13} & \rho_{14} \\ \rho_{12} & 1 & \rho_{23} & \rho_{24} \\ \hline \rho_{13} & \rho_{23} & 1 & \rho_{34} \\ \rho_{14} & \rho_{24} & \rho_{34} & 1 \end{array}\right]$$

and let J_a and J_b be functions from $\mathbb{R}^2$ to $\mathbb{R}$ such that $\mathbb{E}_\Gamma[J_a(X_1,X_2)^2]<\infty$ and $\mathbb{E}_\Gamma[J_b(X_3,X_4)^2]<\infty$. Then, there is a Gaussian vector $(\bar{X}_1,\bar{X}_2,\bar{X}_3,\bar{X}_4)$ with mean 0 and covariance matrix

$$\bar{\Gamma}=\begin{bmatrix}\bar{\Gamma}_{11} & \bar{\Gamma}_{12}\\ \bar{\Gamma}_{21} & \bar{\Gamma}_{22}\end{bmatrix}=\left[\begin{array}{cc|cc} 1 & 0 & \bar{\rho}_{13} & \bar{\rho}_{14}\\ 0 & 1 & \bar{\rho}_{23} & \bar{\rho}_{24}\\ \hline \bar{\rho}_{13} & \bar{\rho}_{23} & 1 & 0\\ \bar{\rho}_{14} & \bar{\rho}_{24} & 0 & 1\end{array}\right]$$

with $\bar{\rho}_{13}=\rho_{13}$, $\bar{\rho}_{14}=(\rho_{14}-\rho_{13}\rho_{34})/\sqrt{1-\rho_{34}^2}$, $\bar{\rho}_{23}=(\rho_{23}-\rho_{12}\rho_{13})/\sqrt{1-\rho_{12}^2}$,

$$\bar{\rho}_{24}=\frac{\rho_{24}+\rho_{12}\rho_{13}\rho_{34}-\rho_{12}\rho_{14}-\rho_{23}\rho_{34}}{\sqrt{1-\rho_{34}^2}\sqrt{1-\rho_{12}^2}}\,.$$

If $|\rho_{ij}|\le\rho^\star$ for all i,j, then $\bar{\rho}_{ij}\le\bar{\rho}^\star$ for all i,j, where

$$\bar{\rho}^\star=\frac{4\rho^\star}{1-(\rho^\star)^2}\,. \tag{47}$$

There are, moreover, functions $\bar{J}_a$ and $\bar{J}_b$ such that

$$\mathbb{E}_\Gamma[J_a(X_1,X_2)J_b(X_3,X_4)]=\mathbb{E}_{\bar{\Gamma}}[\bar{J}_a(\bar{X}_1,\bar{X}_2)\bar{J}_b(\bar{X}_3,\bar{X}_4)]\,,$$

$$\mathbb{E}_\Gamma[J_a(X_1,X_2)]=\mathbb{E}_{\bar{\Gamma}}[\bar{J}_a(\bar{X}_1,\bar{X}_2)]\,,\ \mathbb{E}_\Gamma[J_b(X_3,X_4)]=\mathbb{E}_{\bar{\Gamma}}[\bar{J}_b(\bar{X}_3,\bar{X}_4)]\,,$$

and

$$\mathbb{E}_\Gamma[J_a(X_1,X_2)^2]=\mathbb{E}_{\bar{\Gamma}}[\bar{J}_a(\bar{X}_1,\bar{X}_2)^2]\,,\ \mathbb{E}_\Gamma[J_b(X_3,X_4)^2]=\mathbb{E}_{\bar{\Gamma}}[\bar{J}_b(\bar{X}_3,\bar{X}_4)^2]\,.$$

If J_a and J_b are bounded, then $\bar{J}_a$ and $\bar{J}_b$ are bounded as well.

The proof of Lemma 14 is in the supplemental article Lévy-Leduc *et al.* (2010c).

PROOF OF LEMMA 11. Note that $R_n(t)-R_n(s)$ can be written as

$$R_n(t)-R_n(s)=\frac{1}{n(n-1)}\sum_{1\le i\ne j\le n}J(X_i,X_j)\,, \tag{48}$$

where

$$\begin{aligned} J(x,y)=J_{s,t}(x,y)&=\{h(x,y,t)-h(x,y,s)\}-\{h_1(x,t)-h_1(x,s)\}\\ &\quad-\{h_1(y,t)-h_1(y,s)\}+\{U(t)-U(s)\}. \end{aligned} \tag{49}$$

In the sequel, we shall drop for convenience the subscripts s and t. In view of the definition of h, h_1 and U in (2), (4) and (3) respectively, one has

$$\|J\|_\infty \leq 4 \,, \tag{50}$$

that is, J is bounded. Then, by Conditions (17) and (19), for any Gaussian vector (X_i, X_j, X_k, X_ℓ), one has

$$\mathbb{E}[|J(X_i, X_j)J(X_k, X_\ell)|] \leq C\ \mathbb{E}[|J(X_i, X_j)|] \leq C\ |t-s| \,, \tag{51}$$

for some positive constant C which may change from line to line. By the degeneracy of Hoeffding projections, expanding J into the basis of Hermite polynomials leads to:

$$J(x,y) = \sum_{p,q\geq 0} \frac{c_{p,q}(s,t)}{p!\ q!} H_p(x)H_q(y) \,, \text{ with } \ c_{0,p} = c_{p,0} = 0 \,, \forall p \geq 0 \,, \tag{52}$$

where

$$c_{p,q}(s,t) = \mathbb{E}[J(X,Y)H_p(X)H_q(Y)] \,, \tag{53}$$

X and Y being independent standard Gaussian random variables. Therefore, using (51),

$$|c_{p,q}| \leq \mathbb{E}[J(X,Y)^2]^{1/2}(p!\ q!)^{1/2} \leq C(p!\ q!)^{1/2}|t-s|^{1/2} \,. \tag{54}$$

Remark that the sum in (52) is over p and q such that $p+q \geq m$, since the Hermite rank of J is greater than or equal to the Hermite rank of h. Using (48), we obtain that

$$\mathbb{E}[\{R_n(t) - R_n(s)\}^2] \leq \frac{1}{n^2(n-1)^2} \sum_{\substack{1\leq i_1\neq i_2\leq n \\ 1\leq i_3\neq i_4\leq n}} \mathbb{E}[J(X_{i_1}, X_{i_2})J(X_{i_3}, X_{i_4})] \,. \tag{55}$$

We shall consider 3 cases depending on the cardinality of the set $\{i_1, i_2, i_3, i_4\}$.

1) We first address the case where $i_1 = i_3$ and $i_2 = i_4$. Using (51), we get

$$\frac{1}{n^2(n-1)^2} \sum_{1\leq i_1\neq i_2\leq n} \mathbb{E}[J(X_{i_1}, X_{i_2})^2] \leq \frac{C}{n^2}|t-s| \,,$$

which is consistent with (44).

2) Let us now consider the case where the cardinality of the set $\{i_1, i_2, i_3, i_4\}$ is 3 and suppose without loss of generality that $i_1 = i_3$. Suppose also that ρ defined in Assumption (A1) has the following property: there exists some positive $\rho^\star$ such that

$$|\rho(k)| \leq \rho^\star < 1/13, \text{ for all } k \geq 1 . \tag{56}$$

If we apply the same arguments as in the previous case, we get a rate of order $1/n$ instead of the desired rate $1/n^{1+\alpha}$. To obtain the latter rate, we propose to approximate J by a smooth function J_ε using a convolution approach. More precisely, we define, for all x, y in $\mathbb{R}$,

$$J_\varepsilon(x, y) = \int J(x - \varepsilon z, y - \varepsilon z')\varphi(z)\varphi(z')\mathrm{d}z\mathrm{d}z' . \tag{57}$$

Thus,

$$\begin{aligned} \mathbb{E}[J(X_{i_1}, X_{i_2})J(X_{i_1}, X_{i_4})] = \mathbb{E}[J_\varepsilon(X_{i_1}, X_{i_2})J(X_{i_1}, X_{i_4})] \\ + \mathbb{E}[(J - J_\varepsilon)(X_{i_1}, X_{i_2})J(X_{i_1}, X_{i_4})] . \end{aligned} \tag{58}$$

Applying Lemma 12 to $f = J_\varepsilon$ and noting that, by Condition (15), $\|\partial^6 J_\varepsilon/\partial x^3 \partial y^3\| \leq C\varepsilon^{-6}|t - s|^{1/2}$, for some positive constant C, we obtain

$$\begin{aligned} &\mathbb{E}[J_\varepsilon(X_{i_1}, X_{i_2})J(X_{i_1}, X_{i_4})] \\ &\qquad \leq \mathbb{E}[\sum_{p,q\geq 1} \frac{|c_{p,q}(J_\varepsilon)|}{p!\, q!} |H_p(X_{i_1})J(X_{i_1}, X_{i_4})H_q(X_{i_2})|] \\ &\qquad \leq C\varepsilon^{-6}|t - s|^{1/2} \sum_{p,q\geq 3} (p!q!)^{-1} \sqrt{(p-3)!}\sqrt{(q-3)!} \\ &\qquad\qquad\qquad |\mathbb{E}[H_p(X_{i_1})J(X_{i_1}, X_{i_4})H_q(X_{i_2})]| , \end{aligned}$$

where $c_{p,q}(J_\varepsilon)$ is the (p, q)th Hermite coefficient of J_ε. We shall apply Lemma 13 with $Y_1 = (X_{i_1}, X_{i_4})$, $Y_2 = X_{i_2}$, $a_1 = 2$, $a_2 = 1$, $\Gamma_{0,1} = \mathrm{Id}$, $\Gamma_{0,2} = 1$, $f_1 = H_p J$ and $f_2 = H_q$. Observe that $\mathrm{Id} - \Gamma$ is a 3×3 matrix with ρ entries and hence $\|\mathrm{Id} - \Gamma\|_2 \leq (a_1 + a_2)\|\mathrm{Id} - \Gamma\|_\infty = 3\|\mathrm{Id} - \Gamma\|_\infty = 3\rho^\star$, where $\|A\|_\infty$ is defined for a matrix $A = (a_{i,j})_{i,j}$ by $\|A\|_\infty = \max_{i,j} |a_{i,j}|$. Hence, by (56), the condition on $r^\star$ of Lemma 13 is satisfied. Since J is bounded and f_2 is of Hermite rank larger than 2, Lemma 13 with $[(2 + 1)/2] = 1$ implies that there exists a positive constant C such that

$$|\mathbb{E}[H_p(X_{i_1})J(X_{i_1}, X_{i_4})H_q(X_{i_2})]| \leq C\sqrt{p!\, q!}\, (|\rho(i_4 - i_2)| \vee |\rho(i_2 - i_1)| \vee |\rho(i_4 - i_1)|) . \tag{59}$$

Hence,

(60)
$$\mathbb{E}[J_\varepsilon(X_{i_1},X_{i_2})J(X_{i_1},X_{i_4})] \le C\varepsilon^{-6}|t-s|^{1/2}\left(|\rho(i_4-i_2)|+|\rho(i_2-i_1)|+|\rho(i_4-i_1)|\right) .$$

Since, for example, $\sum_{1\le i_2\neq i_4\le n}|\rho(i_4-i_2)| \le n\sum_{|k|<n}|\rho(k)|$, and since there exist positive constants C and δ such that $|\rho(k)| \le C(1\wedge |k|^{-D+\delta})$, for all $k\ge 1$, we obtain that $\sum_{1\le i_2\neq i_4\le n}|\rho(i_4-i_2)| \le Cn^{2-D+\delta}$. Hence,

$$\frac{1}{n^2(n-1)^2}\sum_{\substack{1\le i_1\neq i_2\le n\\ 1\le i_1\neq i_4\le n}}\mathbb{E}[J_\varepsilon(X_{i_1},X_{i_2})J(X_{i_1},X_{i_4})] \le \frac{C\varepsilon^{-6}|t-s|^{1/2}}{n^{1+D-\delta}} . \tag{61}$$

We now focus on the last term in (58). By the Cauchy-Schwarz inequality and (51),

$$\begin{aligned}\frac{1}{n^2(n-1)^2}&\sum_{\substack{1\le i_1\neq i_2\le n\\ 1\le i_1\neq i_4\le n}}\mathbb{E}[(J-J_\varepsilon)(X_{i_1},X_{i_2})J(X_{i_1},X_{i_4})]\\ &\le C\frac{|t-s|^{1/2}}{n^2(n-1)}\sum_{1\le i_1\neq i_2\le n}\mathbb{E}[(J-J_\varepsilon)^2(X_{i_1},X_{i_2})]^{1/2} .\end{aligned} \tag{62}$$

Using (57), the Jensen's inequality and (50),

$$\begin{aligned}&\mathbb{E}[(J-J_\varepsilon)^2(X_{i_1},X_{i_2})]\\ &= \int_{\mathbb{R}^2}\left\{\int_{\mathbb{R}^2}[J(x,y)-J(x-\varepsilon z,y-\varepsilon z')]\varphi(z)\varphi(z')\mathrm{d}z\mathrm{d}z'\right\}^2 f_{i_1,i_2}(x,y)\mathrm{d}x\mathrm{d}y\\ &\le \int_{\mathbb{R}^2}\left\{\int_{\mathbb{R}^2}[J(x,y)-J(x-\varepsilon z,y-\varepsilon z')]^2 f_{i_1,i_2}(x,y)\mathrm{d}x\mathrm{d}y\right\}\varphi(z)\varphi(z')\mathrm{d}z\mathrm{d}z'\\ &\le C\int_{\mathbb{R}^2}\{\int_{\mathbb{R}^2}|J(x,y)-J(x-\varepsilon z,y-\varepsilon z')|f_{i_1,i_2}(x,y)\mathrm{d}x\mathrm{d}y\}\varphi(z)\varphi(z')\mathrm{d}z\mathrm{d}z' ,\end{aligned} \tag{63}$$

where f_{i_1,i_2} is the p.d.f of (X_{i_1},X_{i_2}). By (49), Conditions (16) and (18),

$$\begin{aligned}\int_{\mathbb{R}^2}\{\int_{\mathbb{R}^2}&|J(x,y)-J(x-\varepsilon z,y-\varepsilon z')|f_{i_1,i_2}(x,y)\mathrm{d}x\mathrm{d}y\}\varphi(z)\varphi(z')\mathrm{d}z\mathrm{d}z'\\ &\le C\varepsilon\int_{\mathbb{R}^2}(|z|+|z'|)\varphi(z)\varphi(z')\mathrm{d}z\mathrm{d}z' \le C\varepsilon .\end{aligned} \tag{64}$$

Using (62), (63) and (64), we get

(65)
$$\frac{1}{n^2(n-1)^2}\sum_{\substack{1\le i_1\neq i_2\le n\\ 1\le i_1\neq i_4\le n}}\mathbb{E}[(J-J_\varepsilon)(X_{i_1},X_{i_2})J(X_{i_1},X_{i_4})] \le C\frac{\varepsilon^{1/2}|t-s|^{1/2}}{n} .$$

Note that (61) involves the factor ε^{-6} and (65) involves the factor $\varepsilon^{1/2}$. By choosing $\varepsilon = \varepsilon_n = n^{-\nu}$ with $0 < \nu < (D - \delta)/6$ in (61) and (65), we obtain a result consistent with (44).

If Condition (56) is not satisfied then let τ be such that $\rho(k) \leq \rho^\star < 1/13$, for all $k > \tau$. In the case where, for instance, $|i_2 - i_4| \leq \tau$ then, using that J is bounded, Conditions (17) and (19), we get that

$$\frac{1}{n^2(n-1)^2} \sum_{\substack{1 \leq i_1 \neq i_2 \leq n \\ 1 \leq i_1 \neq i_4 \leq n, |i_2 - i_4| \leq \tau}} \mathbb{E}[J(X_{i_1}, X_{i_2}) J(X_{i_1}, X_{i_4})] \leq C \frac{\tau |t-s|}{n^2} ,$$

instead of (61), but the result is still consistent with (44). The same result holds when $|i_1 - i_4| \leq \tau$ or $|i_1 - i_2| \leq \tau$. Note also that the remaining sum over the indices such that $|i_1 - i_2| > \tau$, $|i_1 - i_4| > \tau$ and $|i_2 - i_4| > \tau$ can be addressed in the same way as when Condition (56) is satisfied.

3) Now, we assume that the cardinal number of the set $\{i_1, i_2, i_3, i_4\}$ equals 4 and that Condition (56) holds. By Lemma 14,

$$\mathbb{E}[J(X_{i_1}, X_{i_2}) J(X_{i_3}, X_{i_4})] = \mathbb{E}_{\bar{\Gamma}}[J_{i_1,i_2}(\bar{X}_{i_1}, \bar{X}_{i_2}) J_{i_3,i_4}(\bar{X}_{i_3}, \bar{X}_{i_4})] . \tag{66}$$

Here $(\bar{X}_{i_1}, \bar{X}_{i_2}, \bar{X}_{i_3}, \bar{X}_{i_4})$ is a Gaussian vector with mean 0 and covariance matrix $\bar{\Gamma}$ defined in Lemma 14 where $\rho_{ij} = \rho(i - j)$, $J_a = J_b = J$, $\bar{J}_a = J_{i_1,i_2}$ and $\bar{J}_b = J_{i_3,i_4}$. Since the covariance of $(\bar{X}_{i_1}, \bar{X}_{i_2})$ and $(\bar{X}_{i_3}, \bar{X}_{i_4})$ is the identity matrix, we can expand $J_{i_1,i_2}(\bar{X}_{i_1}, \bar{X}_{i_2})$ and $J_{i_3,i_4}(\bar{X}_{i_3}, \bar{X}_{i_4})$. $J_{i_1,i_2}(\bar{X}_{i_1}, \bar{X}_{i_2})$ is the limit in L^2, as $K \to \infty$, of

$$J^K_{i_1,i_2}(\bar{X}_{i_1}, \bar{X}_{i_2}) = \sum_{p=1}^{K} \frac{c^{i_1,i_2}_{p_1,p_2}}{\sqrt{p_1! p_2!}} H_{p_1}(\bar{X}_{i_1}) H_{p_2}(\bar{X}_{i_2}) , \tag{67}$$

with a similar expansion for $J^K_{i_3,i_4}(\bar{X}_{i_3}, \bar{X}_{i_4})$. Therefore,

$$\lim_{K \to \infty} \mathbb{E}_{\bar{\Gamma}}[J_{i_1,i_2}(\bar{X}_{i_1}, \bar{X}_{i_2}) J_{i_3,i_4}(\bar{X}_{i_3}, \bar{X}_{i_4}) - J^K_{i_1,i_2}(\bar{X}_{i_1}, \bar{X}_{i_2}) J^K_{i_3,i_4}(\bar{X}_{i_3}, \bar{X}_{i_4})] = 0 . \tag{68}$$

Thus it is enough to majorize

$$\begin{aligned} &\mathbb{E}_{\bar{\Gamma}}[J^K_{i_1,i_2}(\bar{X}_{i_1}, \bar{X}_{i_2}) J^K_{i_3,i_4}(\bar{X}_{i_3}, \bar{X}_{i_4})] \\ &\leq \sum_{1 \leq p_1, p_2 \leq K} \sum_{1 \leq p_3, p_4 \leq K} \frac{|c^{i_1,i_2}_{p_1,p_2}|}{p_1! p_2!} \frac{|c^{i_3,i_4}_{p_3,p_4}|}{p_3! p_4!} |\mathbb{E}_{\bar{\Gamma}}[H_{p_1}(\bar{X}_1) H_{p_2}(\bar{X}_2) H_{p_3}(\bar{X}_3) H_{p_4}(\bar{X}_4)]| . \end{aligned} \tag{69}$$

By Lemma 3.2 in Taqqu (1977), $\mathbb{E}_{\bar{\Gamma}}[H_{p_1}(\bar{X}_{i_1})H_{p_2}(\bar{X}_{i_2})H_{p_3}(\bar{X}_{i_3})H_{p_4}(\bar{X}_{i_4})]$ is zero if $p_1 + \cdots + p_4$ is odd. Otherwise it is bounded by a constant times a sum of products of $(p_1 + \cdots + p_4)/2$ covariances. These will be denoted $\bar{\rho}_{i,j} = \mathbb{E}(\bar{X}_i \bar{X}_j)$ and are given in Lemma 14. Since $\rho(k) \leq \rho^\star < 1/13$, we have that $\bar{\rho}_{i,j} \leq \bar{\rho}^\star < 1/3$, where $\bar{\rho}^\star = 4\rho^\star/(1-(\rho^\star)^2)$ by (47). Bounding, in each product of covariances, all the covariances but two, by $\bar{\rho}^\star < 1/3$, we get that
$\mathbb{E}_{\bar{\Gamma}}[H_{p_1}(\bar{X}_{i_1})H_{p_2}(\bar{X}_{i_2})H_{p_3}(\bar{X}_{i_3})H_{p_4}(\bar{X}_{i_4})]$ is bounded by

$$C\,(3\bar{\rho}^\star)^{\frac{p_1+p_2+p_3+p_4}{2}-2}\bar{A}(i_1,i_2,i_3,i_4)|\mathbb{E}[H_{p_1}(X)H_{p_2}(X)H_{p_3}(X)H_{p_4}(X)]|\ , \tag{70}$$

where, since $\bar{\rho}_{i_1,i_2} = \bar{\rho}_{i_3,i_4} = 0$,

$$\begin{aligned}\bar{A}(i_1,i_2,i_3,i_4) = \bar{\rho}_{i_1,i_3}\bar{\rho}_{i_2,i_4} + \bar{\rho}_{i_2,i_3}\bar{\rho}_{i_1,i_4} + \bar{\rho}_{i_1,i_3}\bar{\rho}_{i_2,i_3} + \bar{\rho}_{i_1,i_4}\bar{\rho}_{i_2,i_4} \\ + \bar{\rho}_{i_1,i_3}\bar{\rho}_{i_1,i_4} + \bar{\rho}_{i_2,i_3}\bar{\rho}_{i_2,i_4}\ ,\end{aligned} \tag{71}$$

and where X is a standard Gaussian random variable. Note also that the hypercontractivity Lemma 3.1 in Taqqu (1977) yields

$$|\mathbb{E}[H_{p_1}(X)H_{p_2}(X)H_{p_3}(X)H_{p_4}(X)]| \leq 3^{\frac{p_1+p_2+p_3+p_4}{2}}\sqrt{p_1!\,p_2!\,p_3!\,p_4!}\ . \tag{72}$$

Thus (69) is bounded by

$$C\bar{A}\left(\sum_{1\leq p_1,p_2\leq K}\frac{|c^{i_1,i_2}_{p_1,p_2}|}{\sqrt{p_1!\,p_2!}}(3\bar{\rho}^\star)^{\frac{p_1+p_2}{2}-1}\right)\left(\sum_{1\leq p_3,p_4\leq K}\frac{|c^{i_3,i_4}_{p_3,p_4}|}{\sqrt{p_3!\,p_4!}}(3\bar{\rho}^\star)^{\frac{p_3+p_4}{2}-1}\right)\ .$$

By the Cauchy-Schwarz inequality, the first term in brackets is bounded by

$$\begin{aligned}\left(\sum_{1\leq p_1,p_2\leq K}\frac{(c^{i_1,i_2}_{p_1,p_2})^2}{p_1!\,p_2!}\right)^{1/2}\left(\sum_{1\leq p_1,p_2\leq K}(3\bar{\rho}^\star)^{p_1+p_2-2}\right)^{1/2} \\ \leq \mathbb{E}_{\mathrm{I}}\left[J_{i_1,i_2}(\bar{X}_{i_1},\bar{X}_{i_2})^2\right]^{1/2}\left(\sum_{p\geq 1}(3\bar{\rho}^\star)^{p-1}\right)\ ,\end{aligned} \tag{73}$$

where I is the identity matrix and similarly for the second term. Since $\bar{\rho}^\star < 1/3$, it follows from Lemma 14 that (69) is bounded by

$$\begin{aligned}&C\bar{A}\,\mathbb{E}_{\mathrm{I}}\left[J_{i_1,i_2}(\bar{X}_{i_1},\bar{X}_{i_2})\right]^{1/2}\mathbb{E}_{\mathrm{I}}\left[J_{i_3,i_4}(\bar{X}_{i_3},\bar{X}_{i_4})\right]^{1/2} \\ &\quad = C\bar{A}\,\mathbb{E}_{\Gamma_{11}}\left[J(X_{i_1},X_{i_2})^2\right]^{1/2}\mathbb{E}_{\Gamma_{22}}\left[J(X_{i_3},X_{i_4})^2\right]^{1/2} \leq C\bar{A}|t-s|\ ,\end{aligned} \tag{74}$$

where we used (51) and the fact that J is bounded. Thus, in wiew of (66), (68), (69) and (74), we have

$$\begin{aligned}(75)\quad &\frac{1}{n^2(n-1)^2}\sum_{\substack{1\le i_1,i_2,i_3,i_4\le n\\ |\{i_1,i_2,i_3,i_4\}|=4}} \mathbb{E}[J(X_{i_1},X_{i_2})J(X_{i_3},X_{i_4})]\\ &\qquad\le C\frac{|t-s|}{n^2(n-1)^2}\sum_{\substack{1\le i_1,i_2,i_3,i_4\le n\\ |\{i_1,i_2,i_3,i_4\}|=4}} \bar{A}(i_1,i_2,i_3,i_4)\,.\end{aligned}$$

We need to evaluate that sum. Recall that $\bar{A}=\bar{A}(i_1,i_2,i_3,i_4)$ is defined in (71) with the $\bar{\rho}_{i,j}$ defined in Lemma 14. We shall treat one summand in $\bar{A}$ (the others are treated in the same way). We have

$$\bar{\rho}_{i_1,i_3}\bar{\rho}_{i_2,i_4}\le C\rho(i_1-i_3)[\rho(i_3-i_4)+\rho(i_1-i_4)+\rho(i_3-i_4)+\rho(i_2-i_4)]\,.$$

Using that there exist positive constants C (changing from line to line) and ε such that $|\rho(k)|\le C(1\wedge|k|^{-D+\varepsilon})$, for all $k\ge 1$, we get

$$\begin{aligned}\sum_{\substack{1\le i_1,i_2,i_3,i_4\le n\\ |\{i_1,i_2,i_3,i_4\}|=4}}\rho(i_1-i_2)\rho(i_3-i_4)=(\sum_{1\le i_1\ne i_2\le n}\rho(i_1-i_2))^2\le n^2(\sum_{|k|<n}\rho(k))^2\\ \le Cn^{4-2D+2\varepsilon}\,,\end{aligned}$$

and

$$\begin{aligned}\sum_{\substack{1\le i_1,i_2,i_3,i_4\le n\\ |\{i_1,i_2,i_3,i_4\}|=4}}\rho(i_1-i_2)\rho(i_2-i_4)\le n\sum_{i_2=1}^{n}(\sum_{1\le i_1\ne i_2\le n}\rho(i_1-i_2))^2\\ =n\sum_{i_2=1}^{n}(\sum_{i_1=1}^{i_2-1}\rho(i_1-i_2)+\sum_{i_1=i_2+1}^{n}\rho(i_1-i_2))^2\le Cn\sum_{i_2=1}^{n}i_2^{2-2D+2\varepsilon}\le Cn^{4-2D+2\varepsilon}\,.\end{aligned}$$

Therefore, Relation (75) is bounded by $Cn^{-2D+2\varepsilon}|t-s|$, which is a result consistent with (44) with $2\varepsilon=D-1/2>0$.

If Condition (56) is not satisfied, then let τ be such that

$$\sup_{k>\tau}\rho(k)\le\frac{1}{13}(1-\sup_{1\le k\le\tau}\rho(k))\,. \tag{76}$$

In the case where, for instance, $|i_1 - i_3| \leq \tau$ and $|i_2 - i_4| \leq \tau$, there is no need to use (76) because we get using (51),

$$\frac{1}{n^2(n-1)^2} \sum_{\substack{1\leq i_1,i_2,i_3,i_4\leq n \\ |\{i_1,i_2,i_3,i_4\}|=4, |i_1-i_3|\leq\tau, |i_2-i_4|\leq\tau}} \mathbb{E}[J(X_{i_1}, X_{i_2})J(X_{i_3}, X_{i_4})] \leq C\frac{\tau^2|t-s|}{n^2}, \tag{77}$$

which is consistent with (44). In the case where, for instance, $|i_1-i_3| \leq \tau$ and the other distances are larger than τ, we apply the same method as in 2). What changes is the block diagonal matrix Γ_0 involved in Lemma 13. In fact, to evaluate $\mathbb{E}[J_\varepsilon(X_{i_1}, X_{i_2})J(X_{i_3}, X_{i_4})]$ we expand J_ε in Hermite polynomials, so that we need to control $\mathbb{E}[H_p(X_{i_1})J(X_{i_3}, X_{i_4})H_q(X_{i_2})]$. We want to apply Lemma 13 with $Y_1 = (X_{i_1}, X_{i_3}, X_{i_4})$, $Y_2 = X_{i_2}$, $f_1 = H_pJ$, $f_2 = H_q$. We let $\Gamma_{0,1}$ be a 3×3 block diagonal matrix with a first block corresponding to the covariance matrix of the vector (X_{i_1}, X_{i_3}) and a second block equal to 1, and we let $\Gamma_{0,2} = 1$, so that Γ_0 is a 4×4 matrix. Observe that $\|\Gamma_0^{-1}(\Gamma - \Gamma_0)\|_2 \leq 4\|\Gamma_0^{-1}(\Gamma-\Gamma_0)\|_\infty$, where $\|\Gamma_0^{-1}(\Gamma-\Gamma_0)\|_\infty \leq (\sup_{1\leq k\leq\tau}\rho(k))(1 - \sup_{k>\tau}\rho(k))^{-1} \leq 1/13$, by (76). Thus, $\|\Gamma_0^{-1}(\Gamma-\Gamma_0)\|_2 \leq 4/13 < 1/3-\eta$, for some positive η. Hence, the condition on $r^\star$ of Lemma 13 is satisfied. The remaining sum over indices where the distances between any two indices are larger than τ can be addressed in the same way as when Condition (56) is satisfied. □

5.2. *Lemmas used in the proof of Theorem 2.*

The following lemma proves joint convergence and provides the joint cumulants of the limits $(Z_{2,D}(1), (Z_{1,D}(1))^2)$.

LEMMA 15. *Let $(X_j)_{j\geq 1}$ be a stationary process satisfying Assumption (A1) with $D < 1/2$ and let a and b be two real constants. Then, as n tends to infinity,*

$$k(D)\frac{n^{D-2}}{L(n)}\left\{an\sum_{i=1}^n (X_i^2-1) + b\sum_{1\leq i,j\leq n} X_iX_j\right\} \xrightarrow{d} \left[aZ_{2,D}(1) + b(Z_{1,D}(1))^2\right],$$

where $\xrightarrow{d}$ denotes the convergence in distribution, $k(D) = \mathrm{B}((1-D)/2, D)$ where B denotes the Beta function, $Z_{1,D}(\cdot)$ and $Z_{2,D}(\cdot)$ are defined in (24) and (25) respectively. The cumulants of the limit process are given in (89).

The proof of Lemma 15 is in the supplemental article Lévy-Leduc *et al.* (2010c).

REMARK 5. It follows from Lemma 15 that $\mathbb{E}[Z_{D,1}(1)^2] = \sigma^2$ where σ^2 is given in (92). Moreover, setting $a = 1$, $b = 0$ in (89), we get the expression (30) for $\mathbb{E}[Z_{D,2}(1)^2]$.

LEMMA 16. *Under Assumption (A1) there exists a positive constant C such that, for n large enough,*

$$\frac{n^{D-2}}{L(n)}\mathbb{E}[\{\sum_{i=1}^{n} X_i\}^2] \leq C \ , \ \textit{when } D < 1 \ , \tag{78}$$

$$\frac{n^{2D-2}}{L(n)^2}\mathbb{E}[\{\sum_{i=1}^{n}(X_i^2 - 1)\}^2] \leq C \ , \ \textit{when } D < 1/2 \ , \tag{79}$$

and

$$\frac{n^{2D-4}}{L(n)^2}\mathbb{E}[\{\sum_{1\leq i\neq j\leq n} X_i X_j\}^2] \leq C \ , \ \textit{when } D < 1/2 \ . \tag{80}$$

The proof of Lemma 16 is in the supplemental article Lévy-Leduc *et al.* (2010c).

LEMMA 17. *Suppose that the assumptions of Theorem 2 hold, in particular $D < 1/m$ where $m = 1$ or 2. Then, $\widetilde{R}_n$ defined in (21) satisfies the following property. There exist positive constants α and C such that, for n large enough,*

$$a_n^2 \ \mathbb{E}[(\widetilde{R}_n(t) - \widetilde{R}_n(s))^2] \leq C\frac{|t-s|}{n^{\alpha}} \ , \ \textit{for all } s, t \in I \ , \tag{81}$$

where I is any compact interval of $\mathbb{R}$ and $a_n = n^{mD/2-2}L(n)^{-m/2}$.

The proof of Lemma 17 is in the supplemental article Lévy-Leduc *et al.* (2010c).

LEMMA 18. *Under the assumptions of Theorem 2, $\widetilde{R}_n$ defined in (21) satisfies, as n tends to infinity,*

$$\sup_{r\in I} a_n|\widetilde{R}_n(r)| = o_P(1) \ ,$$

where $a_n = n^{-2+mD/2} L(n)^{-m/2}$*, and* m *is the Hermite rank of the class of functions* $\{h(\cdot,\cdot,r) - U(r)\ , r \in I\}$ *which is equal to 1 or 2.*

The proof of Lemma 18 is in the supplemental article Lévy-Leduc *et al.* (2010c).

5.3. *Proof of Theorem 2.* Consider the decomposition (21). Since $G(x,y) = G(y,x)$, one has $\alpha_{1,0}(r) = \alpha_{0,1}(r)$, $\alpha_{2,0}(r) = \alpha_{0,2}(r)$ and $\widetilde{W}_n$ defined in (22) satisfies

$$\widetilde{W}_n(r) = 2(n-1)\alpha_{1,0}(r)\sum_{i=1}^{n} X_i\ ,\ \text{if } m = 1\ , \tag{82}$$

$$\widetilde{W}_n(r) = \alpha_{1,1}(r) \sum_{1\leq i\neq j\leq n} X_i X_j + (n-1)\alpha_{2,0}(r)\sum_{i=1}^{n}(X_i^2 - 1)\ ,\ \text{if } m = 2\ . \tag{83}$$

If $m = 1$, using Lemma 5.1 in Taqqu (1975), if r is fixed, $n^{D/2-2}L(n)^{-1/2}\widetilde{W}_n(r)$ defined in (82) converges in distribution to $2k(D)^{-1/2}\alpha_{1,0}(r)Z_{1,D}(1)$. Then, by the Cramer-Wold device, if $r_1, \dots, r_k$ are fixed real numbers, $k(D)^{1/2} n^{D/2-2}L(n)^{-1/2}(\widetilde{W}_n(r_1), \dots, \widetilde{W}_n(r_k))$ converges in distribution to $(2\alpha_{1,0}(r_1)Z_{1,D}(1), \dots, 2\alpha_{1,0}(r_k)Z_{1,D}(1))$. In the same way, if $m = 2$, using Lemma 15 in Section 5.2 and the Cramer-Wold device, $k(D)n^{D-2}L(n)^{-1}(\widetilde{W}_n(r_1), \dots, \widetilde{W}_n(r_k))$ converges in distribution to $(\alpha_{1,1}(r_1)(Z_{1,D}(1))^2 + \alpha_{2,0}(r_1)Z_{2,D}(1), \dots, \alpha_{1,1}(r_k)(Z_{1,D}(1))^2 + \alpha_{2,0}(r_k)Z_{2,D}(1))$. We now show that $\{n^{mD/2-2}L(n)^{-m/2}\widetilde{W}_n(r); r \in I\}$ is tight in $\mathcal{D}(I)$. We shall do it in the case $m = 1$. By (105), Lemma 16 in Section 5.2 and the fact that $\widetilde{\Lambda}$ is a bounded Lipschitz function, we get that there exists a positive constant C such that for all $r_1 < r_2$ in I,

$$(n^{D/2-2}L(n)^{-1/2})^2\mathbb{E}[\{\widetilde{W}_n(r_2) - \widetilde{W}_n(r_1)\}^2] \leq C(\widetilde{\Lambda}(r_2) - \widetilde{\Lambda}(r_1))^2 \leq C|r_2 - r_1|^2\ .$$

Using the Cauchy-Schwarz inequality, we obtain that for all r_1, r_2, r_3 in I, such that $r_1 < r_2 < r_3$,

$$(n^{D/2-2}L(n)^{-1/2})^2\mathbb{E}\left[\left|\widetilde{W}_n(r_2) - \widetilde{W}_n(r_1)\right|\ \left|\widetilde{W}_n(r_3) - \widetilde{W}_n(r_2)\right|\right] \\ \leq C|r_2 - r_1||r_3 - r_2| \leq C|r_3 - r_1|^2\ .$$

The tightness then follows from Theorem 15.6 of Billingsley (1968). A similar argument holds for $m = 2$. Thus, $\{n^{mD/2-2}L(n)^{-m/2}\widetilde{W}_n(r); r \in I\}$

converges weakly to $\{2\alpha_{1,0}(r)k(D)^{-1/2}Z_{1,D}(1); r \in I\}$, if $m = 1$ and to $\{k(D)^{-1}\left[\alpha_{1,1}(r)Z_{1,D}(1)^2 + \alpha_{2,0}(r)Z_{2,D}(1)\right]; r \in I\}$, if $m = 2$. To complete the proof of Theorem 2 use (21) and Lemma 18 in Section 5.2, which ensures that $\sup_{r\in I} n^{mD/2-2}L(n)^{-m/2}|\widetilde{R}_n(r)| = o_P(1)$, as n tends to infinity.

SUPPLEMENTARY MATERIAL

Proofs of Lemmas 9, 10, 12, 13, 14, 15, 16, 17 and 18 and some numerical experiments.
(http://lib.stat.cmu.edu/aoas/???/???). This supplement contains proofs of Lemmas 9, 10, 12, 13, 14, 15, 16, 17 and 18 and a section containing numerical experiments illustrating some results of Section 4.

References.

Arcones, M. (1994). Limit theorems for nonlinear functionals of a stationary Gaussian sequence of vectors. *Annals of Probability* **22** 2242–2274.

Beran, J. (1991). *M* Estimators of Location for Gaussian and Related Processes With Slowly Decaying Serial Correlations. *Journal of the American Statistical Association* **86** 704–708.

Bickel, P. J. and Lehmann, E. L. (1979). Descriptive statistics for nonparametric models IV: Spread. In *Contributions to Statistics, Hájek Memorial Volume* J. Jurečková ed. 33-40. Prague: Academia.

Billingsley, P. (1968). *Convergence of probability measures*. John Wiley & sons.

Borovkova, S., Burton, R. and Dehling, H. (2001). Limit theorems for functionals of mixing processes with applications to U-statistics and dimension estimation. *Transactions of the American Mathematical Society* **353** 4261–4318.

Dehling, H. and Taqqu, M. S. (1989). The empirical process of some long-range dependent sequences with an application to U-statistics. *Annals of Statistics* **17** 1767–1783.

Dehling, H. and Taqqu, M. S. (1991). Bivariate symmetric statistics of long-range dependent observations. *Journal of Statistical Planning and Inference* **28** 153–165.

Dewan, I. and Prakasa Rao, B. (2005). Wilcoxon-signed rank test for associated sequences. *Statistics & Probability Letters* **71** 131–142.

Fox, R. and Taqqu, M. S. (1987). Multiple stochastic integrals with dependent integrators. *Journal of multivariate analysis* **21** 105–127.

Grassberger, P. and Procaccia, I. (1983). Characterization of strange attractors. *Physical Review Letters* **50** 346–349.

Hodges, J. L. J. and Lehmann, E. L. (1963). Estimates of location based on rank tests. *Annals of Mathematical Statistics* **34** 598-611.

Hoeffding, W. (1948). A class of statistics with asymptotically normal distribution. *Annals of Mathematical Statistics* **19** 293–325.

Hsing, T. and Wu, W. (2004). On weighted U-statistics for stationary processes. *Annals of Probability* **32** 1600-1631.

Lévy-Leduc, C., Boistard, H., Moulines, E., Taqqu, M. S. and Reisen, V. A. (2010a). Large sample behavior of some well-known robust estimators under long-range dependence. *To appear in Statistics*.

Lévy-Leduc, C., Boistard, H., Moulines, E., Taqqu, M. S. and Reisen, V. A. (2010b). Robust estimation of the scale and of the autocovariance function of Gaussian short and long-range dependent processes. *To appear in Journal of Time Series Analysis*.

Lévy-Leduc, C., Boistard, H., Moulines, E., Taqqu, M. S. and Reisen, V. A. (2010c). Supplement to "Asymptotic properties of U-processes under long-range dependence".

Ma, Y. and Genton, M. (2000). Highly robust estimation of the autocovariance function. *Journal of Time Series Analysis* **21** 663–684.

Rousseeuw, P. and Croux, C. (1993). Alternatives to the median absolute deviation. *Journal of the American Statistical Association* **88** 1273–1283.

Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics*. John Wiley & Sons.

Shamos, M. I. (1976). Geometry and Statistics: Problems at the interface. In *New directions and Recent Results in Algorithms and Complexity* J. F. Traub ed. 251–280. New York: Academic Press.

Soulier, P. (2001). Moment bounds and central limit theorem for functions of Gaussian vectors. *Statistics & Probability Letters* **54** 193–203.

Stuart, A. and Ord, J. K. (1987). *Kendall's Advanced theory of statistics*. Charles Griffin & company limited, London.

Taqqu, M. S. (1975). Weak convergence to fractional Brownian motion and to the Rosenblatt process. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* **31** 287–302.

Taqqu, M. S. (1977). Law of the iterated logarithm for sums of non-linear functions of Gaussian variables that exhibit a long range dependence. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* **40** 203–238.

van der Vaart, A. W. (1998). *Asymptotic statistics*. *Cambridge Series in Statistical and Probabilistic Mathematics* **3**. Cambridge University Press, Cambridge. MR1652247 (2000c:62003)

Wilcoxon, F. (1945). Individual Comparisons by Ranking Methods. *Biometrics Bulletin* **1** 80–83.

Supplement to paper "Asymptotic properties of U-processes under long-range dependence"

This supplement contains proofs of Lemmas 9, 10, 12, 13, 14, 15, 16, 17 and 18 and a section containing numerical experiments illustrating some results of Section 4.

Proof of Lemma 9. Let us check that the assumptions of Theorem 9 in Arcones (1994) hold for the class $\mathcal{F}$ of functions $\{h_1(\cdot, r) : r \in I\}$ which is of rank $\tau \geq m = 2 > 1/D$. By Assumption (A1) and since $\tau > 1/D$, the condition (i") of Theorem 9 in Arcones (1994) is satisfied. We conclude the proof of the Lemma by observing that the condition (ii) of this Theorem is also fulfilled. To check this condition, we have to prove that

$$\int_0^\infty (N_{[\,]}^{(2)}(\varepsilon, \mathcal{F}))^{1/2} \mathrm{d}\varepsilon < \infty \,,$$

where $N_{[\,]}^{(2)}(\varepsilon, \mathcal{F})$ is the bracketing number of the class $\mathcal{F}$ as defined on page 2269 in Arcones (1994):

$$N_{[\,]}^{(2)}(\varepsilon, \mathcal{F}) = \min\{N : \exists \text{ measurable functions } f_1, \dots, f_N \text{ and } \Delta_1, \dots, \Delta_N \\ \text{such that for each } f \in \mathcal{F}, \exists i \leq N \text{ such that } |f_i - f| \leq \Delta_i \text{ and where} \\ \mathbb{E}(\Delta_i^2(X)) \leq \varepsilon^2 \text{ for each } i \leq N\} \,.$$

Let $\{r_i, i = 0, \dots, N\}$ be such that: for all i, $|r_i - r_{i-1}| \leq \varepsilon/C$ and for all $r \in I$, there exists i such that $|r - r_i| \leq \varepsilon$. The smallest N satisfying this property is at most equal to $[|I|C/\varepsilon]+1$, where $|I|$ denotes the length of I. Let us define for all $i \geq 1$, $f_i = h_1(\cdot, r_i)$ and $\Delta_i = h_1(\cdot, r_i) - h_1(\cdot, r_{i-1})$. Using (19) we first get $\mathbb{E}(\Delta_i^2(X)) = \mathbb{E}(\{h_1(X, r_i) - h_1(X, r_{i-1})\}^2) \leq C^2|r_i - r_{i-1}|^2 \leq \varepsilon^2$. Now, let $r \in I$ and $1 \leq i \leq N$ be such that, $r_{i-1} \leq r \leq r_i$. Then, using the fact that h_1 is increasing with respect to its second argument leads to $|f_i - h_1(\cdot, r)| \leq \Delta_i$. Thus, $N_{[\,]}^{(2)}(\varepsilon, \mathcal{F}) \leq [|I|C/\varepsilon] + 1$ which yields condition (ii) of Theorem 9 in Arcones (1994). $\square$

Proof of Lemma 10. We want to apply Lemma 5.2, P. 4307 of Borovkova, Burton and Dehling (2001) to $\{\sqrt{n}R_n(r), r \in I\}$. To do this, we prove that for all $s, t \in I, \delta > 0$ such that $s \leq t \leq s + \delta$ and $s + \delta \in I$:

$$\begin{aligned} &\sqrt{n}|R_n(t) - R_n(s)| \\ &\leq \sqrt{n}|R_n(s+\delta) - R_n(s)| + 2\sqrt{n}|W_n(s+\delta) - W_n(s)| + 4\sqrt{n}|U(s+\delta) - U(s)| \,. \end{aligned} \tag{84}$$

Using the definition of R_n given by (8) and the fact that h, h_1 and U are nondecreasing functions with respect to r, we get:

$$R_n(t) - R_n(s) \leq \frac{1}{n(n-1)} \sum_{1 \leq i \neq j \leq n} \{h(X_i, X_j, t) - h(X_i, X_j, s)\} + \{U(t) - U(s)\}$$
$$\leq \frac{1}{n(n-1)} \sum_{1 \leq i \neq j \leq n} \{h(X_i, X_j, s+\delta) - h(X_i, X_j, s)\} + \{U(s+\delta) - U(s)\} \,.$$

By adding and subtracting functions h_1 evaluated at $s+\delta$ and s, we obtain:

$$R_n(t) - R_n(s) \leq \{R_n(s+\delta) - R_n(s)\} + \frac{2}{n}\sum_{i=1}^{n}\{h_1(X_i, s+\delta) - h_1(X_i, s)\} .$$

Adding and subtracting $2(U(s) - U(s+\delta))$ leads to:

$$R_n(t)-R_n(s) \leq \{R_n(s+\delta)-R_n(s)\}+\{W_n(s+\delta)-W_n(s)\}+2\{U(s+\delta)-U(s)\} ,$$

where W_n is defined in (7). Thus,

(85)

$$R_n(t)-R_n(s) \leq |R_n(s+\delta)-R_n(s)|+|W_n(s+\delta)-W_n(s)|+2|U(s+\delta)-U(s)| .$$

Let us now find an upper bound for $R_n(s)-R_n(t)$. Starting with the expression (8) for $R_n(r)$ and setting $h(X_i, X_j, s) \leq h(X_i, X_j, s+\delta)$, $U(s) \leq U(s+\delta)$ and $h(X_i, X_j, t) \geq h(X_i, X_j, s)$, $U(t) \geq U(s)$ since h and U are non decreasing functions with respect to r, we obtain

$$\begin{aligned}
R_n(s) - R_n(t) \leq \frac{1}{n(n-1)} &\sum_{1\leq i\neq j\leq n} [\{h(X_i, X_j, s+\delta) - h(X_i, X_j, s)\} \\
&- 2\{h_1(X_i, s) - h_1(X_i, t)\}] + \{U(s+\delta) - U(s)\} \\
\leq \frac{1}{n(n-1)} &\sum_{1\leq i\neq j\leq n} [\{h(X_i, X_j, s+\delta) - h(X_i, X_j, s)\} \\
&- 2\{h_1(X_i, s+\delta) - h_1(X_i, s)\}] \\
+ \frac{2}{n(n-1)} &\sum_{1\leq i\neq j\leq n} [\{h_1(X_i, t) - h_1(X_i, s)\} + \{h_1(X_i, s+\delta) - h_1(X_i, s)\}] \\
&+ \{U(s+\delta) - U(s)\} \\
\leq \{R_n(s+\delta) - R_n(s)\} &+ \frac{4}{n}\sum_{1\leq i\leq n}\{h_1(X_i, s+\delta) - h_1(X_i, s)\} .
\end{aligned}$$

Adding and subtracting $4(U(s) - U(s+\delta))$ leads to:

(86)

$$R_n(s)-R_n(t) \leq |R_n(s+\delta)-R_n(s)|+2|W_n(s+\delta)-W_n(s)|+4|U(s+\delta)-U(s)| .$$

Combining (85) and (86), we get for all $s, t \in I, \delta > 0$ such that $s \leq t \leq s+\delta$ and $s+\delta \in I$:

$$\begin{aligned}
\sqrt{n}|R_n(t) - R_n(s)| \leq \sqrt{n}|R_n(s+\delta) - R_n(s)| &+ 2\sqrt{n}|W_n(s+\delta) - W_n(s)| \\
&+ 4\sqrt{n}|U(s+\delta) - U(s)| ,
\end{aligned}$$

which is (84). Remark that U is Lipschitz by Condition (19). In Lemma 5.2, P. 4307 of Borovkova, Burton and Dehling (2001), the monotone Lipschitz-continuous function Λ is here U, $\alpha = 1/2$ and the process $\{Y_n(t)\}$ is here $\{\sqrt{n}W_n(t)\}$. We shall now verify that conditions (i) and (ii) of that lemma are satisfied. Condition (i) holds because of Lemma 11. Condition (ii) involves $\{\sqrt{n}W_n(t)\}$. Applying inequality (2.43) of Theorem 4 in Arcones (1994) to $f(\cdot) = (h_1(\cdot, t) - h_1(\cdot, s)) - (U(t) - U(s))$, which is, by (14), of Hermite rank $\tau \geq 2 > 1/D$, we get using (19) that there exist some positive constants C and C' such that:

$$\begin{aligned}
&\mathbb{E}\left[|\sqrt{n}\{W_n(t) - W_n(s)\}|^2\right] \\
&\qquad = \mathbb{E}\left[\left\{\frac{2}{\sqrt{n}}\sum_{i=1}^{n}(h_1(X_i,t) - h_1(X_i,s)) - (U(t) - U(s))\right\}^2\right] \\
&\qquad \leq C\,\mathbb{E}\left[\{(h_1(X_1,t) - h_1(X_1,s)) - (U(t) - U(s))\}^2\right] \leq C'\,|t-s|^2 .
\end{aligned}$$

Thus condition (ii) of Lemma 5.2 in Borovkova, Burton and Dehling (2001) is satisfied with $r = 2$ and monotone function $g(t) = t$. An application of this lemma concludes the proof. $\square$

Proof of Lemma 12. Using that, for $n \geq 1$, $(H_{n-1}\varphi)' = -H_n\varphi$, where $'$ denotes the first derivative, and 6 integrations by parts (3 with respect to x and 3 with respect to y), we get that for $p, q \geq 3$,

$$c_{p,q}(f) = \int_{\mathbb{R}^2} \partial^6 f(x,y)/\partial x^3 \partial y^3 H_{q-3}(y)\varphi(y)H_{p-3}(x)\varphi(x)\mathrm{d}x\mathrm{d}y ,$$

where φ is the p.d.f of a standard Gaussian random variable. (45) then follows from the Cauchy-Schwarz inequality. $\square$

Proof of Lemma 13. Let f be a function defined on $\mathbb{R}^{a_1+a_2}$ such that $f(Y) = f(Y_1, Y_2) = f_1(Y_1)f_2(Y_2)$. Note that $\mathbb{E}[f(Y)] = \mathbb{E}[f(\Gamma_0^{1/2}Z)]$, where the covariance matrix of Z is equal to $\Gamma_0^{-1/2}\Gamma\Gamma_0^{-1/2}$. By the assumption on $r^\star$, the latter matrix is invertible and satisfies: $(\Gamma_0^{-1/2}\Gamma\Gamma_0^{-1/2})^{-1} = [I_{a_1+a_2} - \Gamma_0^{-1/2}(\Gamma_0 - \Gamma)\Gamma_0^{-1/2}]^{-1} = I_{a_1+a_2} + \sum_{k\geq 1}(\Gamma_0^{-1/2}(\Gamma_0 - \Gamma)\Gamma_0^{-1/2})^k$. Let $\Delta := (\Gamma_0^{-1/2}\Gamma\Gamma_0^{-1/2})^{-1} - I_{a_1+a_2}$. By definition of the density of the multivariate Gaussian distribution and the definition of the matrix

Δ, we obtain that

$$|\Gamma_0^{-1/2}\Gamma\Gamma_0^{-1/2}|^{1/2}\ \mathbb{E}[f(\Gamma_0^{1/2}Z)] \\ = \int_{\mathbb{R}^{a_1+a_2}} f(\Gamma_0^{1/2}z)\exp(-z^T\Delta z/2)\exp(-z^Tz/2)\frac{\mathrm{d}z}{(2\pi)^{(a_1+a_2)/2}} .$$

Expanding $\exp(-z^T\Delta z/2)$ in series leads to

$$|\Gamma_0^{-1/2}\Gamma\Gamma_0^{-1/2}|^{1/2}\ \mathbb{E}[f(\Gamma_0^{1/2}Z)] \\ = \sum_{k\geq 0}\frac{(-1/2)^k}{k!}\int_{\mathbb{R}^{a_1+a_2}} f(\Gamma_0^{1/2}z)(z^T\Delta z)^k\exp(-z^Tz/2)\frac{\mathrm{d}z}{(2\pi)^{(a_1+a_2)/2}} .$$

Set $\nu = [(\tau+1)/2]$, where $[x]$ denotes the integer part of x. Using that f is of Hermite rank at least τ and the previous equation, we get

$$|\Gamma_0^{-1/2}\Gamma\Gamma_0^{-1/2}|^{1/2}\ \mathbb{E}[f(\Gamma_0^{1/2}Z)] \\ = \sum_{k\geq \nu}\frac{(-1/2)^k}{k!}\int_{\mathbb{R}^{a_1+a_2}} f(\Gamma_0^{1/2}z)(z^T\Delta z)^k\exp(-z^Tz/2)\frac{\mathrm{d}z}{(2\pi)^{(a_1+a_2)/2}} .$$

Since $|\sum_{k\geq\nu}(-1/2)^k(z^T\Delta z)^k/k!| \leq |z^T\Delta z|^\nu\exp(|z^T\Delta z|/2)/(2^\nu\nu!)$, we obtain

$$|\Gamma_0^{-1/2}\Gamma\Gamma_0^{-1/2}|^{1/2}\ |\mathbb{E}[f(\Gamma_0^{1/2}Z)]| \\ \leq \frac{1}{2^\nu\nu!}\int_{\mathbb{R}^{a_1+a_2}} |f(\Gamma_0^{1/2}z)|\ |z^T\Delta z|^\nu\exp(|z^T\Delta z|/2)\exp(-z^Tz/2)\frac{\mathrm{d}z}{(2\pi)^{(a_1+a_2)/2}} .$$

Denoting by δ the spectral radius of Δ gives

$$|\Gamma_0^{-1/2}\Gamma\Gamma_0^{-1/2}|^{1/2}\ |\mathbb{E}[f(\Gamma_0^{1/2}Z)]| \\ \leq \frac{\delta^\nu}{2^\nu\nu!}\int_{\mathbb{R}^{a_1+a_2}} |f(\Gamma_0^{1/2}z)|(z^Tz)^\nu\exp\{(\delta/2-1/2)z^Tz\}\frac{\mathrm{d}z}{(2\pi)^{(a_1+a_2)/2}} .$$

By the Cauchy-Schwarz inequality, we get

$$|\Gamma_0^{-1/2}\Gamma\Gamma_0^{-1/2}|^{1/2}\ |\mathbb{E}[f(\Gamma_0^{1/2}Z)]| \\ \leq \frac{\delta^\nu}{2^\nu\nu!}\left(\int_{\mathbb{R}^{a_1+a_2}} f^2(\Gamma_0^{1/2}z)\exp(-z^Tz/2)\frac{\mathrm{d}z}{(2\pi)^{(a_1+a_2)/2}}\right)^{1/2} \\ \left(\int_{\mathbb{R}^{a_1+a_2}}(z^Tz)^{2\nu}\exp\{(\delta-1/2)z^Tz\}\frac{\mathrm{d}z}{(2\pi)^{(a_1+a_2)/2}}\right)^{1/2} . \tag{87}$$

By definition of Δ, the spectral radius δ of Δ satisfies $\delta \leq \sum_{k\geq 1}(r^\star)^k \leq r^\star/(1-r^\star)$, where $r^\star$ is the spectral radius of $\Gamma_0^{-1/2}(\Gamma_0-\Gamma)\Gamma_0^{-1/2}$. By assumption on $r^\star$, $\delta \leq 1/2 - 3\varepsilon/2$ which implies that the second integral in (87) is convergent. The first integral in (87) satisfies

$$\left(\int_{\mathbb{R}^{a_1+a_2}} f^2(\Gamma_0^{1/2}z)\exp(-z^T z/2)\frac{\mathrm{d}z}{(2\pi)^{(a_1+a_2)/2}}\right)^{1/2} = \|f_1\|_{2,\Gamma_{0,1}}\|f_2\|_{2,\Gamma_{0,2}}\ . \tag{88}$$

Finally, under the assumption on $r^\star$, the spectral radius δ_0 of $(\Gamma_0^{-1/2}\Gamma\Gamma_0^{-1/2})^{-1} = \sum_{k\geq 0}\{\Gamma_0^{-1/2}(\Gamma_0-\Gamma)\Gamma_0^{-1/2}\}^k$ satisfies $\delta_0 \leq \sum_{k\geq 0}(r^\star)^k = 1/(1-r^\star)$, so that $|\Gamma_0^{-1/2}\Gamma\Gamma_0^{-1/2}|^{-1/2} \leq (1-r^\star)^{-(a_1+a_2)/2} \leq (2/3+\varepsilon)^{-(a_1+a_2)/2} \leq (3/2)^{(a_1+a_2)/2}$. This establishes (46). □

PROOF OF LEMMA 14. Let I denote the 2×2 identity matrix. One can express Γ_{11} as $\Gamma_{11} = L_a L_a^T$, where T denotes the transpose so that the vector $(\bar{X}_1,\bar{X}_2)^T = L_a^{-1}(X_1,X_2)^T$ has covariance matrix $\bar{\Gamma}_{11} = L_a^{-1}\Gamma_{11}(L_a^{-1})^T = \mathrm{I}$. Similarly, $\Gamma_{22} = L_b L_b^T$ so that $(\bar{X}_3,\bar{X}_4)^T = L_b^{-1}(X_3,X_4)^T$ has covariance matrix $\bar{\Gamma}_{22} = \mathrm{I}$. Then

$$\mathbb{E}_\Gamma[J_a(X_1,X_2)J_b(X_3,X_4)] = \mathbb{E}_{\bar{\Gamma}}[\bar{J}_a(\bar{X}_1,\bar{X}_2)\bar{J}_b(\bar{X}_3,\bar{X}_4)]\ ,$$

where $\bar{J}_a = J_a\circ L_a$, $\bar{J}_b = J_b \circ L_b$ and

$$\bar{\Gamma} = \begin{bmatrix} L_a^{-1} & 0 \\ 0 & L_b^{-1}\end{bmatrix}\begin{bmatrix}\Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22}\end{bmatrix}\begin{bmatrix}(L_a^{-1})^T & 0 \\ 0 & (L_b^{-1})^T\end{bmatrix} = \begin{bmatrix}\mathrm{I} & L_a^{-1}\Gamma_{12}(L_b^{-1})^T \\ L_b^{-1}\Gamma_{21}(L_a^{-1})^T & \mathrm{I}\end{bmatrix}\ .$$

Observe that

$$L_a = \begin{bmatrix} 1 & 0 \\ \rho_{12} & \sqrt{1-\rho_{12}^2}\end{bmatrix}\ ,\ L_a^{-1} = \begin{bmatrix} 1 & 0 \\ -\rho_{12}/\sqrt{1-\rho_{12}^2} & 1/\sqrt{1-\rho_{12}^2}\end{bmatrix}$$

since $L_a L_a^T = \Gamma_{11}$. A similar expression holds for L_b with ρ_{12} replaced by

ρ_{34}. Observe that

$$\bar{\Gamma}_{12} = L_a^{-1}\Gamma_{12}(L_b^{-1})^T = \begin{bmatrix} 1 & 0 \\ -\rho_{12}/\sqrt{1-\rho_{12}^2} & 1/\sqrt{1-\rho_{12}^2} \end{bmatrix} \begin{bmatrix} \rho_{13} & \rho_{14} \\ \rho_{23} & \rho_{24} \end{bmatrix} \begin{bmatrix} 1 & -\rho_{34}/\sqrt{1-\rho_{34}^2} \\ 0 & 1/\sqrt{1-\rho_{34}^2} \end{bmatrix} = \begin{bmatrix} \bar\rho_{13} & \bar\rho_{14} \\ \bar\rho_{23} & \bar\rho_{24} \end{bmatrix} ,$$

where the $\bar\rho_{ij}$ are given in the statement of the lemma. This characterizes the matrix $\bar\Gamma$ since $\bar\Gamma_{11} = \bar\Gamma_{22} = \mathrm{I}$ and $\bar\Gamma_{21} = \bar\Gamma_{12}^T$. The relations involving $\rho^\star$ and $\bar\rho^\star$ follow from those relating the $\bar\rho_{ij}$'s to the ρ_{kl}'s. Finally, one has

$$\mathbb{E}_{\mathrm{I}}[\bar J_a(\bar X_1, \bar X_2)^2] = \mathbb{E}_{\Gamma_{11}}[J_a(X_1,X_2)^2] ,$$

and also the other similar type relations. $\square$

Proof of Lemma 15. We first prove that, for $p \geq 2$, the pth cumulant κ_p of $aZ_{2,D}(1) + b(Z_{1,D}(1))^2$ is equal to

$$\kappa_p = 2^{p-1}(p-1)!\, k(D)^p \int_{[0,1]^p} \mathrm{d}u_1 \dots \mathrm{d}u_p \int_{[0,1]^p} \mathrm{d}v_1 \dots \mathrm{d}v_p \prod_{j=1}^p [a\delta(u_j - v_j) + b]\, |u_j - v_{j-1}|^{-D}, \text{ with } v_0 = v_p , \tag{89}$$

where $k(D) = \mathrm{B}((1-D)/2, D)$, $\delta(x) = 1$ if $x = 0$, and $\delta(x) = 0$ else. Using (24) and (25),

$$aZ_{2,D}(1) + b(Z_{1,D}(1))^2 = \int_{\mathbb{R}^2}' K(x,y)\mathrm{d}B(x)\mathrm{d}B(y) + b\sigma^2 ,$$

where

$$K(x,y) = \int_0^1 \int_0^1 [a\delta(u-v) + b](u-x)_+^{-(D+1)/2}(v-y)_+^{-(D+1)/2}\mathrm{d}u\mathrm{d}v ,$$

and

$$\sigma^2 = E[Z_{1,D}(1)^2] = \int_{\mathbb{R}} [\int_0^1 (u-x)_+^{-\frac{D+1}{2}}\mathrm{d}u]^2 \mathrm{d}x = \int_0^1 \int_0^1 [\int_{\mathbb{R}} (u-x)_+^{-\frac{D+1}{2}}(v-x)_+^{-\frac{D+1}{2}}\mathrm{d}x]\mathrm{d}u\mathrm{d}v .$$

Using that for $0 < \alpha < 1/2$,

$$\int_{\mathbb{R}} (u-x)_+^{\alpha-1}(v-x)_+^{\alpha-1}\mathrm{d}x = |u-v|^{2\alpha-1}\int_0^\infty y^{\alpha-1}(1+y)^{\alpha-1}\mathrm{d}y = |u-v|^{2\alpha-1}\mathrm{B}(\alpha,-2\alpha+1)\ , \tag{90}$$

where $\mathrm{B}(\cdot,\cdot)$ denotes the Beta function, we get

$$\int_{\mathbb{R}} (u-x)_+^{-(D+1)/2}(v-x)_+^{-(D+1)/2}\mathrm{d}x = \mathrm{B}((1-D)/2, D)|u-v|^{-D} = k(D)|u-v|^{-D}\ . \tag{91}$$

Thus,

$$\sigma^2 = k(D)\int_0^1\int_0^1 |u-v|^{-D}\mathrm{d}u\mathrm{d}v = \frac{2k(D)}{(-D+1)(-D+2)}\ . \tag{92}$$

Hence, using Proposition 4.2 in Fox and Taqqu (1987), we have that for $p \geq 2$,

$$\kappa_p = 2^{p-1}(p-1)!\int_{\mathbb{R}^p} K(x_1,x_2)K(x_2,x_3)\dots K(x_{p-1},x_p)K(x_p,x_1)\mathrm{d}x_1\dots\mathrm{d}x_p\ .$$

By definition of K, and with the convention $x_{p+1} = x_1$,

$$\begin{aligned}
&\int_{\mathbb{R}^p} K(x_1,x_2)K(x_2,x_3)\dots K(x_{p-1},x_p)K(x_p,x_1)\mathrm{d}x_1\dots\mathrm{d}x_p \\
&= \int_{[0,1]^p}\mathrm{d}u_1\dots\mathrm{d}u_p\int_{[0,1]^p}\mathrm{d}v_1\dots\mathrm{d}v_p\int_{\mathbb{R}^p}\mathrm{d}x_1\dots\mathrm{d}x_p \\
&\prod_{j=1}^p [a\delta(u_j-v_j)+b](u_j-x_j)_+^{-(D+1)/2}(v_j-x_{j+1})^{-(D+1)/2} \\
&= \int_{[0,1]^p}\mathrm{d}u_1\dots\mathrm{d}u_p\int_{[0,1]^p}\mathrm{d}v_1\dots\mathrm{d}v_p \\
&\prod_{j=1}^p [a\delta(u_j-v_j)+b]\int_{\mathbb{R}}(u_j-x_j)_+^{-(D+1)/2}(v_{j-1}-x_j)^{-(D+1)/2}
\end{aligned}$$

where $v_0 = v_p$ since x_j is associated with u_j and v_{j-1}. Using (91), we obtain the expression (89) for the cumulants κ_p, $p \geq 2$. Let us now compute the limit as n tends to infinity of the cumulants of

$$A_n = \frac{n^{D-2}}{L(n)}\left[X'(an\mathrm{I}+b\mathbf{1}\mathbf{1}')X\right] = \frac{n^{D-2}}{L(n)}\left[\widetilde{X}'\Sigma^{1/2}\left(an\mathrm{I}+b\mathbf{1}\mathbf{1}'\right)\Sigma^{1/2}\widetilde{X}\right]\ ,$$

where $X = (X_1, \dots, X_n)'$, $\mathbf{1} = (1, \dots, 1)'$, I is the $n \times n$ identity matrix, Σ is the covariance matrix of X and $\widetilde{X}$ is a standard Gaussian random vector. Using Stuart and Ord (1987), p. 488, the pth cumulant of A_n is equal to

$$cum_p = 2^{p-1}(p-1)!\ \mathrm{Tr}(B_n^p)\ ,$$

where $B_n = n^{D-2}L(n)^{-1}\left[\Sigma^{1/2}\left(an\mathrm{I} + b\mathbf{1}\mathbf{1}'\right)\Sigma^{1/2}\right]$. But

$$\begin{aligned}\mathrm{Tr}(B_n^p) &= \left(\frac{n^{D-2}}{L(n)}\right)^p \mathrm{Tr}\left[\{(an\mathrm{I} + b\mathbf{1}\mathbf{1}')\Sigma\}^p\right] \\ &= \left(\frac{n^{D-2}}{L(n)}\right)^p \sum_{\substack{1\le i_1, i_2, \dots, i_p \le n \\ 1 \le j_1, j_2, \dots, j_p \le n}} D_{i_1,j_1}\rho(j_1 - i_2)D_{i_2,j_2}\rho(j_2 - i_3)\dots \\ &\qquad\qquad D_{i_{p-1},j_{p-1}}\rho(j_{p-1} - i_p)D_{i_p,j_p}\rho(j_p - i_1)\ ,\end{aligned}$$

where ρ is defined in Assumption (A1) and $D_{i,j} = an\delta(i-j) + b$. With the convention $i_{p+1} = i_1$,

$$\begin{aligned}\mathrm{Tr}(B_n^p) &= \frac{1}{n^{2p}} \sum_{\substack{1\le i_1, i_2, \dots, i_p \le n \\ 1 \le j_1, j_2, \dots, j_p \le n}} \prod_{\ell=1}^p \left\{\frac{n^D}{L(n)}[an\delta(i_\ell - j_\ell) + b]\rho(j_\ell - i_{\ell+1})\right\} \\ &= \frac{1}{n^{2p}} \sum_{\substack{1\le i_1, i_2, \dots, i_p \le n \\ 1 \le j_1, j_2, \dots, j_p \le n}} \prod_{\ell=1}^p \left\{\frac{n^D}{L(n)}[an\delta(i_\ell - j_\ell) + b]\rho(j_{\ell-1} - i_\ell)\right\}\ ,\end{aligned}$$

where $j_0 = j_p$. Thus, as n tends to infinity,

$$\begin{aligned}\mathrm{Tr}(B_n^p) \to \int_{[0,1]^p} \mathrm{d}u_1 \dots \mathrm{d}u_p \int_{[0,1]^p} \mathrm{d}v_1 \dots \mathrm{d}v_p \\ \prod_{j=1}^p [a\delta(u_j - v_j) + b] \int_{\mathbb{R}} |u_j - v_{j-1}|^{-D}\ ,\end{aligned}$$

with the convention $v_0 = v_p$, which gives the expected result. $\square$

Proof of Lemma 16. By Assumption (A1), $\rho(k) = k^{-D}L(k)$. Using the adaptation of Karamata's theorem given in Taqqu (1975), one gets

$$\text{if } D < 1/m,\ \sum_{|k|<n} |\rho(k)|^m \sim \frac{2}{(1-mD)(2-mD)} n^{1-mD}(L(n))^m\ . \tag{93}$$

The bound (78) follows from

$$\mathbb{E}[\{\sum_{i=1}^{n} X_i\}^2] \leq (n + \sum_{1\leq i\neq j\leq n} |\rho(i-j)|) \leq n(1 + \sum_{|k|<n, k\neq 0} |\rho(k)|) ,$$

and by setting $m = 1$ in (93). Let us prove (79). Given that $\mathbb{E}[H_p(X_i)H_q(X_j)] = p!\,\delta(p-q)\rho(i-j)^p$, for all integers $p, q, i, j \geq 1$, we obtain

$$\begin{aligned}\mathbb{E}[\{\sum_{i=1}^{n}(X_i^2 - 1)\}^2] &= \mathbb{E}[\sum_{1\leq i,j\leq n} H_2(X_i)H_2(X_j)] \\ &= 2n + 2\sum_{1\leq i\neq j\leq n} \rho(i-j)^2 \leq 2n(1 + \sum_{|k|<n, k\neq 0} \rho(k)^2) .\end{aligned}$$

The bound (79) follows by using that $D < 1/2$ and (93) with $m = 2$. Let us now prove (80). Note that

$$\begin{aligned}&\mathbb{E}[(\sum_{1\leq i\neq j\leq n} X_iX_j)^2] = \sum_{\substack{1\leq i\neq j\leq n\\ 1\leq k\neq \ell\leq n}} \mathbb{E}(X_iX_jX_kX_\ell) \\ &= \sum_{1\leq i\neq j\leq n} \mathbb{E}(X_i^2X_j^2) + \sum_{\substack{1\leq i,j,k,\ell\leq n\\ |\{i,j,k,\ell\}|=4}} \mathbb{E}(X_iX_jX_kX_\ell) + 6\sum_{\substack{1\leq i,j,\ell\leq n\\ |\{i,j,\ell\}|=3}} \mathbb{E}(X_i^2X_jX_\ell) .\end{aligned} \tag{94}$$

Writing $X_i^2 = H_2(X_i) + 1$ and using that $\mathbb{E}[H_2(X_i)H_2(X_j)] = 2\rho(i-j)^2$, for all $i, j \geq 1$, the first term in the r.h.s of (94) satisfies

$$\sum_{1\leq i\neq j\leq n} \mathbb{E}(X_i^2X_j^2) \leq n^2 + 2n\sum_{|k|<n, k\neq 0} \rho(k)^2 .$$

Using Lemma 3.2 P. 210 in Taqqu (1977), the second term in the r.h.s of (94) satisfies, for some positive constant C,

$$\sum_{\substack{1\leq i,j,k,\ell\leq n\\ |\{i,j,k,\ell\}|=4}} \mathbb{E}(X_iX_jX_kX_\ell) \leq Cn^2(\sum_{|k|<n, k\neq 0} |\rho(k)|)^2 .$$

Writing $X_i^2 = H_2(X_i) + 1$ and using Lemma 3.2 P. 210 in Taqqu (1977), the third term in the r.h.s of (94) satisfies, for some positive constant C,

$$\sum_{\substack{1\leq i,j,\ell\leq n\\ |\{i,j,\ell\}|=3}} \mathbb{E}(X_i^2X_jX_\ell) \leq Cn(\sum_{|k|<n} |\rho(k)|)^2 + n^2\sum_{|k|<n} |\rho(k)| .$$

The last three inequalities lead to the expected result by using (93). $\square$

PROOF OF LEMMA 17. Set $\alpha_{p,q}(s,t) = \alpha_{p,q}(t) - \alpha_{p,q}(s)$ for all s, t in $\mathbb{R}$, where $\alpha_{p,q}(\cdot)$ is defined in (10). Then

$$\mathbb{E}[(\widetilde{R}_n(t) - \widetilde{R}_n(s))^2] = \sum_{\substack{1 \le i_1 \ne i_2 \le n \\ 1 \le i_3 \ne i_4 \le n}} \mathbb{E}[\widetilde{J}_{s,t}(X_{i_1}, X_{i_2})\widetilde{J}_{s,t}(X_{i_3}, X_{i_4})] \;, \tag{95}$$

where for all x, y in $\mathbb{R}$ and s, t in I,

$$\begin{aligned}\widetilde{J}_{s,t}(x,y) = (h(x,y,t) - h(x,y,s)) - (\alpha_{1,0}(t) - \alpha_{1,0}(s))(x+y) \\ - (U(t) - U(s)) \;, \text{ if } m = 1 \;,\end{aligned} \tag{96}$$

$$\begin{aligned}\widetilde{J}_{s,t}(x,y) = (h(x,y,t) - h(x,y,s)) - (\alpha_{1,1}(t) - \alpha_{1,1}(s))xy \\ - \frac{1}{2}(\alpha_{2,0}(t) - \alpha_{2,0}(s))(x^2 + y^2 - 2) - (U(t) - U(s)) \;, \text{ if } m = 2 \;.\end{aligned} \tag{97}$$

To obtain these relations express $\widetilde{R}_n(t) - \widetilde{R}_n(s)$ using (21), (5) and (22). We now consider 3 cases, depending on the cardinality of the set $\{i_1, i_2, i_3, i_4\}$.

1) We start with the case of cardinality 2. Let us address the case where the sum is over the set of indices $\{i_1, i_2, i_3, i_4\}$ such that $i_1 = i_3$ and $i_2 = i_4$. We shall only focus on the case where $m = 1$ because the case $m = 2$ could be addressed in the same way. We thus need to show that

$$\mathbb{E}[(\widetilde{R}_n(t) - \widetilde{R}_n(s))^2] \le C n^{4-D-\alpha} |t - s| \;. \tag{98}$$

Using that h, U and $\alpha_{1,0}$ are bounded functions, there exists a positive constant C such that

$$\begin{aligned}\sum_{1 \le i_1 \ne i_2 \le n} \mathbb{E}[\widetilde{J}^2_{s,t}(X_{i_1}, X_{i_2})] \le C \sum_{1 \le i_1 \ne i_2 \le n} \mathbb{E}\left[|h(X_{i_1}, X_{i_2}, t) - h(X_{i_1}, X_{i_2}, s)|\right. \\ \left. + |\alpha_{1,0}(t) - \alpha_{1,0}(s)| \{(X_{i_1} + X_{i_2})^2 + |X_{i_1} + X_{i_2}|\} + |U(t) - U(s)|\right] \;.\end{aligned}$$

Since Condition (26) holds and U, $\widetilde{\Lambda}$, defined in (27) are Lipschitz functions, there exist positive constants C_1 and C_2 such that

$$\begin{aligned}&\sum_{1 \le i_1 \ne i_2 \le n} \mathbb{E}[\widetilde{J}^2_{s,t}(X_{i_1}, X_{i_2})] \\ &\le C_1 n(n-1)|t - s| + C_2 |t - s| \sum_{1 \le i_1 \ne i_2 \le n} \mathbb{E}[(X_{i_1} + X_{i_2})^2 + |X_{i_2}| + |X_{i_2}|] \\ &\qquad \le C n(n-1)|t - s| \;,\end{aligned}$$

which gives (98).

2) Let us now address the case where the sum is over the set of indices $\{i_1, i_2, i_3, i_4\}$ having a cardinal number equal to 3 *i.e.* for instance when $i_3 = i_1$. As previously, we focus on the case where $m = 1$. Using that h, U and $\alpha_{1,0}$ are bounded functions, Condition (26), the Lipschitz property of U and $\widetilde{\Lambda}$, there exists a positive constant C such that

$$\sum_{\substack{1 \le i_1 \ne i_2 \le n \\ 1 \le i_1 \ne i_4 \le n}} \mathbb{E}[\widetilde{J}_{s,t}(X_{i_1}, X_{i_2})\widetilde{J}_{s,t}(X_{i_1}, X_{i_4})] \le Cn^3 |t - s| \ ,$$

which gives (98).

3) Let us now consider the case where the sum is over the indices i_1, i_2, i_3, i_4 such that the cardinal number of the set $\{i_1, i_2, i_3, i_4\}$ is equal to 4. This case is similar to the case 3) in the proof of Lemma 11. We need to show that

$$a_n^2 \sum_{\substack{1 \le i_1, i_2, i_3, i_4 \le n \\ |\{i_1, i_2, i_3, i_4\}| = 4}} \mathbb{E}[\widetilde{J}(X_{i_1}, X_{i_2})\widetilde{J}(X_{i_3}, X_{i_4})] \le C \frac{|t - s|}{n^\alpha} \ , \tag{99}$$

where $a_n = n^{mD/2-2} L(n)^{-m/2}$, $\widetilde{J} = \widetilde{J}_{s,t}$ is defined in (96) if $m = 1$ and in (97) if $m = 2$, $C > 0$ and $\alpha > 0$. We will only present the case $m = 1$. The idea is, once again, to replace $(X_{i_1}, X_{i_2}, X_{i_3}, X_{i_4})$ by $(\bar{X}_{i_1}, \bar{X}_{i_2}, \bar{X}_{i_3}, \bar{X}_{i_4})$ using Lemma 14, so that (66) holds with J, J_{i_1,i_2}, J_{i_3,i_4} replaced by $\widetilde{J}$, $\widetilde{J}_{i_1,i_2}$, $\widetilde{J}_{i_3,i_4}$ and to expand $\widetilde{J}_{i_1,i_2}(\bar{X}_{i_1}, \bar{X}_{i_2})$ and $\widetilde{J}_{i_3,i_4}(\bar{X}_{i_3}, \bar{X}_{i_4})$ in Hermite polynomials as in (67), with an expansion up to K, thus defining $\widetilde{J}^K_{i_1,i_2}$ and $\widetilde{J}^K_{i_3,i_4}$. Then, (68) holds with J replaced by $\widetilde{J}$. Denoting again the coefficients of the expansion by $c^{i_1,i_2}_{p_1,p_2}$ and $c^{i_3,i_4}_{p_3,p_4}$ respectively, one needs to majorize the right-hand side of (69). Assuming that $|i_1 - i_2|$ is the smallest distance between two different indices, namely $|i_1 - i_2| = \min\{|i_1 - i_2|, |i_1 - i_3|, |i_1 - i_4|, |i_2 - i_3|, |i_2 - i_4|, |i_3 - i_4|\}$, we obtain that the $\bar{\rho}_{ij}$'s are bounded by $4\rho(i_1 - i_2)/(1 - \rho(i_1 - i_2)^2)$. Using that there exist positive constants C and ε such that $|\rho(k)| \le C(1 \wedge |k|^{-D+\varepsilon}) =: \gamma(k)$, for all $k \ge 1$, we get, by (47), that the $\bar{\rho}_{ij}$'s are bounded by

$$4\gamma(i_1 - i_2)/(1 - \gamma(i_1 - i_2)^2) =: \bar{\gamma}(i_1 - i_2) \ . \tag{100}$$

By Lemma 3.2 in Taqqu (1977), we obtain

$$\begin{aligned} &\mathbb{E}_{\bar{\Gamma}}[H_{p_1}(\bar{X}_{i_1})H_{p_2}(\bar{X}_{i_2})H_{p_3}(\bar{X}_{i_3})H_{p_4}(\bar{X}_{i_4})] \\ &\qquad \le C\bar{\gamma}(i_1 - i_2)^{\frac{p_1+p_2+p_3+p_4}{2}} |\mathbb{E}[H_{p_1}(X)H_{p_2}(X)H_{p_3}(X)H_{p_4}(X)]| \ . \end{aligned}$$

Using (72), $\mathbb{E}_{\bar{\Gamma}}[\widetilde{J}^K_{i_1,i_2}(\bar{X}_{i_1},\bar{X}_{i_2})\widetilde{J}^K_{i_3,i_4}(\bar{X}_{i_3},\bar{X}_{i_4})]$ is bounded by

$$\sum_{1\leq p_1,p_2\leq K}\frac{|c^{i_1,i_2}_{p_1,p_2}|}{\sqrt{p_1!\,p_2!}}(3\bar{\gamma}(i_1-i_2))^{\frac{p_1+p_2}{2}}\sum_{1\leq p_3,p_4\leq K}\frac{|c^{i_3,i_4}_{p_3,p_4}|}{\sqrt{p_3!\,p_4!}}(3\bar{\gamma}(i_1-i_2))^{\frac{p_3+p_4}{2}}\ .$$

Using the Cauchy-Schwarz inequality, the first sum is bounded by

$$\mathbb{E}_{\mathrm{I}}[\widetilde{J}_{i_1,i_2}(\bar{X}_{i_1},\bar{X}_{i_2})^2]^{1/2}\left(\sum_{p\geq 1}(3\bar{\gamma}(i_1-i_2))^p\right)\ ,$$

and similarly for the second sum. It follows from Lemma 14 that $\mathbb{E}_{\bar{\Gamma}}[\widetilde{J}^K_{i_1,i_2}(\bar{X}_{i_1},\bar{X}_{i_2})\widetilde{J}^K_{i_3,i_4}(\bar{X}_{i_3},\bar{X}_{i_4})]$ is bounded by

$$\mathbb{E}_{\Gamma_{11}}[\widetilde{J}(X_{i_1},X_{i_2})^2]^{1/2}\mathbb{E}_{\Gamma_{22}}[\widetilde{J}(X_{i_3},X_{i_4})^2]^{1/2}\left(\sum_{p\geq 1}(3\bar{\gamma}(i_1-i_2))^p\right)^2\ .$$

Since Condition (26) holds and U, $\widetilde{\Lambda}$, defined in (27) are Lipschitz functions,

$$\mathbb{E}_{\Gamma_{11}}[\widetilde{J}(X_{i_1},X_{i_2})^2]^{1/2}\mathbb{E}_{\Gamma_{22}}[\widetilde{J}(X_{i_3},X_{i_4})^2]^{1/2}\leq C|t-s|\ .$$

We deduce from the previous inequalities that

$$a_n^2\sum_{\substack{1\leq i_1,i_2,i_3,i_4\leq n\\|\{i_1,i_2,i_3,i_4\}|=4}}\mathbb{E}[\widetilde{J}(X_{i_1},X_{i_2})\widetilde{J}(X_{i_3},X_{i_4})]$$
$$\leq Cn^3a_n^2|t-s|\sum_{\substack{|k|<n\\k\neq 0}}\left(\sum_{p\geq 1}(3\bar{\gamma}(k))^p\right)^2\ ,$$

where $\bar{\gamma}$ is defined in (100). Let η be a positive constant such that $\eta>3$. Then there is $K\geq 1$ such that $\eta\bar{\gamma}(k)\leq 1$, for all $k\geq K$. We may suppose without loss of generality that $K=1$, that is $\eta\bar{\gamma}(k)\leq 1$, for all $k\geq 1$. We then obtain that for large enough n,

(101)
$$a_n^2\mathbb{E}[(\widetilde{R}_n(t)-\widetilde{R}_n(s))^2]\leq C\eta^2|t-s|\ a_n^2\ n^3\left(\sum_{\substack{|k|<n\\k\neq 0}}\bar{\gamma}(k)^2\right)\left(\sum_{p\geq 1}\left(\frac{3}{\eta}\right)^p\right)\ .$$

Observe that $a_n^2n^3=n^{D-1}L(n)^{-1}$. Recall also that $\bar{\gamma}(k)$ is defined in (100) and note that $\sum_k\bar{\gamma}(k)^2$ may be finite or infinite. If $\sum_k\bar{\gamma}(k)^2<\infty$ then the result (81) follows with $\alpha = 1 - D$ which is positive since $D < 1$. If $\sum_k \bar{\gamma}(k)^2 = \infty$ then $\sum_{|k|<n} \bar{\gamma}(k)^2 \sim 16 n^{1-2D+2\varepsilon}$ and the result follows with $\alpha = D - 2\varepsilon$ if ε is chosen small enough to ensure that this quantity is positive. $\square$

PROOF OF LEMMA 18. We want to apply Lemma 5.2, P. 4307 of Borovkova, Burton and Dehling (2001) to $\{a_n \widetilde{R}_n(r), r \in I\}$. To do so, we first prove that for all $s, t \in I, \delta > 0$ such that $s \le t \le s + \delta$ and $s + \delta \in I$:

$$
\begin{aligned}
a_n|\widetilde{R}_n(t)-\widetilde{R}_n(s)| \le a_n|\widetilde{R}_n(s+\delta)-\widetilde{R}_n(s)| + 2a_n|\widetilde{\Lambda}(s+\delta)-\widetilde{\Lambda}(s)|\,|\sum_{1\le i\ne j\le n} X_i| \\
+ 2a_n n(n-1)|U(s+\delta)-U(s)|\ , \text{ if } m=1 \text{ and } D<1\ ,
\end{aligned} \tag{102}
$$

where $\widetilde{\Lambda}$ is defined in (27) and

$$
\begin{aligned}
a_n|\widetilde{R}_n(t)-\widetilde{R}_n(s)| \le a_n|\widetilde{R}_n(s+\delta)-\widetilde{R}_n(s)| + 2a_n|\widetilde{\Lambda}(s+\delta)-\widetilde{\Lambda}(s)| \\
[\,|\sum_{1\le i\ne j\le n} X_iX_j| + |\sum_{1\le i\ne j\le n}(X_i^2-1)|\,] + 2a_n n(n-1)|U(s+\delta)-U(s)|\ , \\
\text{if } m=2 \text{ and } D<1/2\ .
\end{aligned} \tag{103}
$$

Let us focus on the proof of (102), where $m = 1$ and $D < 1$, since the proof of (103) can be obtained by using similar arguments. In view of the definition (21) of $\widetilde{R}_n$ and the fact that U_n and U are non decreasing functions, we obtain

$$
\widetilde{R}_n(t)-\widetilde{R}_n(s) \le \widetilde{R}_n(s+\delta)-\widetilde{R}_n(s)+\widetilde{W}_n(s+\delta)-\widetilde{W}_n(t)+n(n-1)(U(s+\delta)-U(s))\ .
$$

Remark that the monotonicity of h in (27) implies that $\widetilde{\Lambda}$ is a non decreasing function and that for $p + q \le 2$,

$$
|\alpha_{p,q}(s) - \alpha_{p,q}(r)| \le |\widetilde{\Lambda}(s) - \widetilde{\Lambda}(r)|\ , \text{ for all } r, s\ . \tag{104}
$$

Since $p+q \le 2$ and $m = 1$, we need to consider only $p = 1$, $q = 0$ and $p = 0$, $q = 1$ in (22), we thus get

$$
\widetilde{W}_n(s+\delta) - \widetilde{W}_n(t) \le 2(\widetilde{\Lambda}(s+\delta) - \widetilde{\Lambda}(s))\,|\sum_{1\le i\ne j\le n} X_i|\ . \tag{105}
$$

In the same way, after switching s and t, we obtain

$$\widetilde{R}_n(s) - \widetilde{R}_n(t) \leq \widetilde{R}_n(s+\delta) - \widetilde{R}_n(s) + 2(\widetilde{\Lambda}(s+\delta) - \widetilde{\Lambda}(s))|\sum_{1\leq i\neq j\leq n} X_i| \\ + 2n(n-1)(U(s+\delta) - U(s)) ,$$

which gives (102). In Lemma 5.2, P. 4307 of Borovkova, Burton and Dehling (2001), the monotone Lipschitz-continuous function Λ is here $2U$, the process $\{Y_n(t)\}$ is $\{a_n\widetilde{\Lambda}(t)(\sum_{1\leq i\neq j\leq n} X_i)\}$ if $m = 1$, and if $m = 2$, the process $\{Y_n(t)\}$ is $\{a_n\widetilde{\Lambda}(t)(\sum_{1\leq i\neq j\leq n} X_iX_j + \sum_{1\leq i\neq j\leq n}(X_i^2 - 1))\}$. Using Lemma 16 and the fact that the function $\widetilde{\Lambda}$ defined in (27) is a Lipschitz function, the processes $\{Y_n(t), t \in I\}$ defined above satisfy the condition (ii) of Lemma 5.2 in Borovkova, Burton and Dehling (2001) with $r = 2$. Using Lemma 17, the condition (i) of Lemma 5.2 in Borovkova, Burton and Dehling (2001) is also satisfied. This concludes the proof. ☐

6. Numerical experiments.

In this section, we investigate the robustness properties of the Hodges-Lehmann and Shamos scale estimators defined in Section 4 using Monte Carlo experiments. We shall regard the observations X_t, $t = 1, \dots, n$, as a stationary series Y_t, $t = 1, \dots, n$, corrupted by additive outliers of magnitude ω. Thus we set

$$X_t = Y_t + \omega W_t, \tag{106}$$

where W_t are i.i.d. random variables. In Section 6.1, W_t are Bernoulli$(p/2)$ random variables. In Section 6.2, W_t are such that $\mathbb{P}(W_t = -1) = \mathbb{P}(W_t = 1) = p/2$ and $\mathbb{P}(W_t = 0) = 1 - p$, hence $\mathbb{E}[W_t] = 0$ and $\mathbb{E}[W_t^2] = \mathrm{Var}(W_t) = p$. Observe that, in this case, W is the product of Bernoulli(p) and *Rademacher* independent random variables; the latter equals 1 or -1, both with probability $1/2$. $(Y_t)_t$ is a stationary time series and it is assumed that Y_t and W_t are independent random variables. The empirical study is based on 5000 independent replications with $n = 600$, $p = 10\%$ and $\omega = 10$. We consider the cases where (Y_t) are Gaussian ARFIMA$(1, d, 0)$ processes, that is,

$$Y_t = (I - \phi B)^{-1}(I - B)^{-d} Z_t\ , \tag{107}$$

where B denotes the backward operator, $\phi = 0.2$ and $d = 0.1$, 0.35, corresponding respectively to $D = 0.8$, 0.3, where D is defined in (A1) and (Z_t) are i.i.d $\mathcal{N}(0, 1)$.

6.1. *Hodges-Lehmann estimator.* In this section, we illustrate the results of Proposition 5. In Figure 1, the empirical density functions of $\hat{\theta}_{HL}$ and $\bar{X}_n$ are displayed when X_t has no outliers with $d = 0.1$ (left) and $d = 0.35$ (right). In these cases both shapes are similar to the limit indicated in Proposition 5, that is, a Gaussian density with mean zero.

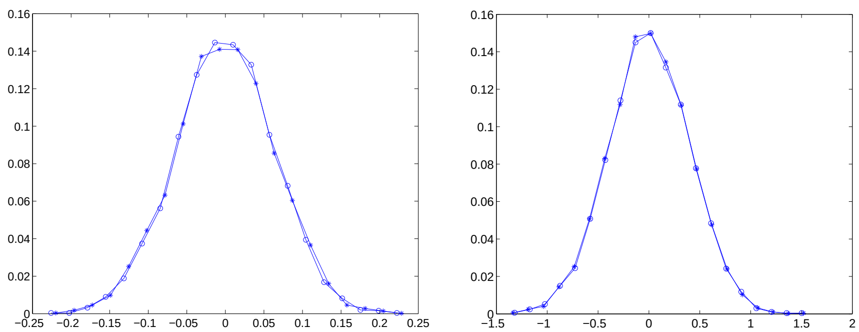

FIG 1. *Empirical densities of the quantities $\hat{\theta}_{HL}$ ('*') and $\bar{X}_n$ ('o') for the ARFIMA$(1, d, 0)$ model with $d = 0.1$ (left), $d = 0.35$ (right), $n = 600$ without outliers.*

Figure 2 displays the same quantities as in Figure 1 when X_t has outliers with $d = 0.1$ (left) and $d = 0.35$ (right). As expected, the sample mean is much more sensitive to the presence of outliers than the Hodges-Lehmann estimator. Observe that when the long-range dependence is strong (large d), the effect of outliers is less pronounced.

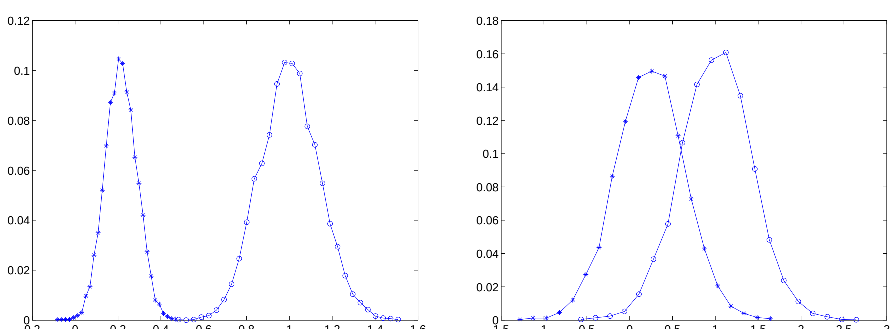

FIG 2. *Empirical densities of the quantities $\hat{\theta}_{HL}$ ('*') and $\bar{X}_n$ ('o') for the ARFIMA$(1, d, 0)$ model with $d = 0.1$ (left), $d = 0.35$ (right), $n = 600$ with outliers ($p = 10\%$ and $\omega = 10$).*

6.2. *Shamos scale estimator.* In this section, we illustrate the results of Proposition 8. In Figure 3, the empirical densities of $\hat{\sigma}_{BL} - \sigma$ and $\hat{\sigma}_{n,X} - \sigma$ are displayed when $d = 0.1$ without outliers (left) and with outliers (right). In the left part of this figure, we illustrate the results of the first part of Proposition 8 since both shapes are similar to that of Gaussian density with mean zero. On the right part of Figure 3, we can see that the classical scale estimator is much more sensitive to the presence of outliers than the Shamos-Bickel estimator.

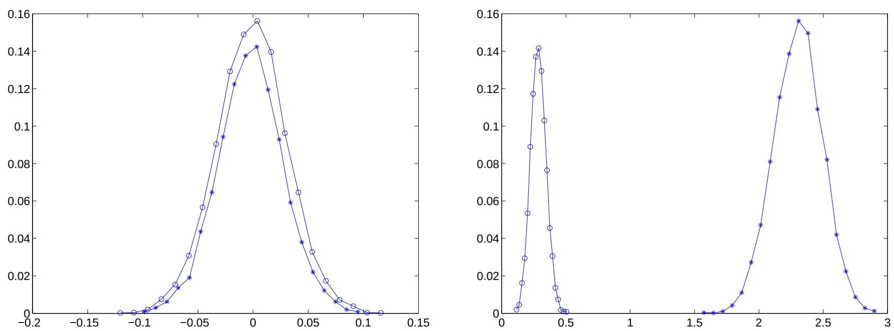

FIG 3. *Empirical densities of the quantities* $(\hat{\sigma}_{BL} - \sigma)$ *('*') and* $(\hat{\sigma}_{n,X} - \sigma)$ *('o') for the* $ARFIMA(1, d, 0)$ *model with* $d = 0.1$, $n = 600$ *without outliers (left) and with outliers* $p = 10\%$ *and* $\omega = 10$ *(right).*

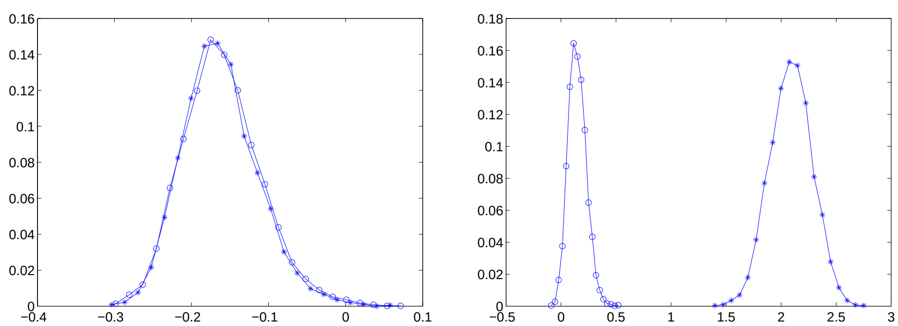

FIG 4. *Empirical densities of the quantities* $(\hat{\sigma}_{SB} - \sigma)$ *('*') and* $(\hat{\sigma}_{n,X} - \sigma)$ *('o') for the* $ARFIMA(1, d, 0)$ *model with* $d = 0.35$, $n = 600$ *without outliers (left) and with outliers* $p = 10\%$ *and* $\omega = 10$ *(right).*

Figure 4 (left) illustrates the second part of Proposition 8. $d = 0.35$ corresponds to $D = 0.3 < 1/2$. The right part of Figure 4 shows the robustness of the Shamos estimator with respect to the classical scale estimator in the presence of outliers.

Céline Lévy-Leduc (corresponding author)
CNRS/LTCI/TelecomParisTech
46, rue Barrault
75634 Paris Cédex 13, France.
E-mail: celine.levy-leduc@telecom-paristech.fr

Hélène Boistard
Toulouse School of Economics
GREMAQ, Université Toulouse 1
Manufacture des Tabacs
bât. F, aile J.J. Laffont
21 allée de Brienne
31000 Toulouse
E-mail: helene@boistard.fr

Eric Moulines
CNRS/LTCI/TelecomParisTech
46, rue Barrault
75634 Paris Cédex 13, France.
E-mail: moulines@telecom-paristech.fr

Murad S. Taqqu
Department of Mathematics
Boston University
111 Cumington Street
Boston, MA 02215, USA
E-mail: murad@math.bu.edu

Valderio Anselmo Reisen
Departamento de Estatística
Universidade Federal
do Espírito Santo
Vitória/ES, Brazil.
E-mail: valderio@cce.ufes.br