

ASYMPTOTIC PROPERTIES OF U -PROCESSES UNDER LONG-RANGE DEPENDENCE

BY C. LÉVY-LEDUC, H. BOISTARD, E. MOULINES, M. S. TAQQU¹
AND V. A. REISEN

*CNRS/LTCI/TelecomParisTech, Toulouse School of Economics,
CNRS/LTCI/TelecomParisTech, 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

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

where I is an interval included in \mathbb{R} , and G is a symmetric function. In this paper, we provide central and noncentral 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 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

Received January 2010; revised September 2010.

¹Supported in part by NSF Grants DMS-07-06786 and DMS-10-07616 at Boston University.
MSC2010 subject classifications. 60F17, 62M10, 62G30, 62G20.

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

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, e.g., 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), Dehling and Taqqu (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 of extending the results of Borovkova, Burton and Dehling (2001) in order to address the long-range dependence case, second, of extending the results obtained in Dehling and Taqqu (1989) to functions of two variables and third, of 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 not necessarily Gaussian long-range dependent processes; whereas, in our paper, we establish the asymptotic properties of U -processes involving Gaussian long-range dependent processes. 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

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

where I is an interval included in \mathbb{R} , and G is a symmetric function; that is, $G(x, y) = G(y, x)$ 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

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

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

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

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. (2011a) 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. (2011b).

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 noncentral 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. (2011c)] contains the proofs of some of the lemmas. It contains also some numerical experiments.

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

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

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

$$(4) \quad h_1(x, r) = \int_{\mathbb{R}} h(x, y, r)\varphi(y) dy.$$

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

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

as

$$(6) \quad U_n(r) - U(r) = W_n(r) + R_n(r),$$

where

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

and

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

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

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

where

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

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

$$(11) \quad \mathbb{E}[h^2(X, Y, r)] = \sum_{p, q \geq 0} \frac{\alpha_{p, q}^2(r)}{p!q!}.$$

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 $\alpha_{p, q}(r) \neq 0$. Thus, (9) can be rewritten as

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

The Hermite rank m of the class of functions $\{h(\cdot, \cdot, r) - U(r), r \in I\}$ is the smallest index $m = p + q \geq 1$ such that $\alpha_{p,q}(r) \neq 0$ 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

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

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

$$(14) \quad \tau \geq m.$$

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

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

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} ,*

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

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

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

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

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

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

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

and

$$(19) \quad |h_1(x, s) - h_1(x, t)| \leq C|t - s|.$$

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

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

PROOF. 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. \square

REMARK 1. The examples of Section 4 satisfy conditions (15) to (19), for example, 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:

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

where

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

Introduce also the Beta function

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

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

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

and

$$(25) \quad Z_{2,D}(t) = \int_{\mathbb{R}^2} \left[\int_0^t (u-x)_+^{-(D+1)/2} (u-y)_+^{-(D+1)/2} du \right] dB(x) dB(y),$$

$0 < D < 1/2,$

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 ,*

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

- (ii) U is a Lipschitz function.
- (iii) The function $\tilde{\Lambda}$ defined, for all s in I , by

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

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

Then,

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

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

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

and to

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

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

$$(28) \quad k(D) = B((1 - D)/2, D),$$

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

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

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

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

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

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

with $k(D)$ defined by (28). See Remark 1 in Lévy-Leduc et al. (2011c) 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 (6) in the supplemental article [Lévy-Leduc et al. (2011c)].

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

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

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

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

and to

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

3. Asymptotic behavior of empirical quantiles. We shall apply Theorems 1 and 2 in the preceding section to empirical quantiles. Recall that if $V : I \rightarrow [0, 1]$ is a nondecreasing cadlag function, where I is an interval of \mathbb{R} , then its generalized inverse V^{-1} is defined by $V^{-1}(p) = \inf\{r \in I, V(r) \geq p\}$. This applies to $U_n(r)$ and $U(r)$ since these are nondecreasing 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 nonnull. 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. 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 nonnull. 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

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

and to

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

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

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

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,

$$\begin{aligned} \alpha_{1,0}(r) &= \alpha_{0,1}(r) = \mathbb{E}[A 1_{\{A+B \leq 2r\}}] = - \int_{\mathbb{R}} \varphi(2r - y) \varphi(y) dy \\ (31) \quad &= -\varphi(r\sqrt{2})/\sqrt{2}, \end{aligned}$$

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 $\{1_{G(\cdot, \cdot) \leq r} - \alpha_{0,0}(r); r \in \mathbb{R}\}$ 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 $t \leq s$,

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

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

$$\begin{aligned} &\iint 1_{\{s < x+y \leq t\}} (|x| + |xy| + |x^2 - 1|) \varphi(x) \varphi(y) \, dx \, dy \\ &= \int \left(\int_{s-x}^{t-x} \varphi(y) \, dy \right) |x| \varphi(x) \, dx + \int \left(\int_{s-x}^{t-x} |y| \varphi(y) \, dy \right) |x| \varphi(x) \, dx \\ &\quad + \int \left(\int_{s-x}^{t-x} \varphi(y) \, dy \right) |x^2 - 1| \varphi(x) \, dx. \end{aligned}$$

Using that $\varphi(\cdot)$ and $|\cdot| \varphi(\cdot)$ 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

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

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, $U(r) = \int \Phi(2r - x) \varphi(x) \, dx$, $U'(r) = 2 \int \varphi(2r - x) \varphi(x) \, dx$, $U(0) = 1/2 \times \int (\Phi(x) + \Phi(-x)) \varphi(x) \, dx = 1/2$, $U^{-1}(1/2) = 0$ 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

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

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 us 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, \dots, Y_n . It is defined as

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

where the R_j ’s are the ranks of X_1, \dots, 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

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

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

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

The asymptotic properties of $U_{n,2}$ can be derived from those of $U_n(\cdot)$ 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 $\{1_{G(\cdot, \cdot) \leq r} - \alpha_{0,0}(r); r \in \mathbb{R}\}$ 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

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

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

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

which can be rewritten as follows:

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

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), pages 1617 and 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, $\alpha_{1,0}(r) = \alpha_{0,1}(r) = \int_{\mathbb{R}} x 1_{\{|x-y|\leq r\}}\varphi(x)\varphi(y) dx dy = \int_{\mathbb{R}} x[\Phi(x+r) - \Phi(x-r)]\varphi(x) dx$, 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

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

where $\dot{\varphi}$ denotes the first derivative of φ . It is nonnull if $r \neq 0$. Thus, for any compact interval I which does not contain 0, the Hermite rank of the class of functions $\{1_{G(\cdot, \cdot) \leq r} - \alpha_{0,0}(r), r \in I\}$ 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

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

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

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

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

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

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

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

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

$$\begin{aligned} \hat{\sigma}_{BL} &= c \operatorname{median}\{|Y_i - Y_j|; 1 \leq i < j \leq n\} \\ &= c\sigma \operatorname{median}\{|X_i - X_j|; 1 \leq i < j \leq n\}, \end{aligned}$$

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 $G(x, y) = |x - y|$, $U(r) = \int [\Phi(x + r) - \Phi(x - r)]\varphi(x) dx$, $U'(r) = 2 \int \varphi(x + r)\varphi(x) dx$, $U^{-1}(1/2) = 1/c$ and $U'(U^{-1}(1/2)) = U'(1/c) = \sqrt{2}\varphi(1/(c\sqrt{2}))$. By Corollary 3, we obtain that for $D > 1/2$,

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

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 $D < 1/2$,

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

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

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

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

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

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*

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

The proof of Lemma 9 is in the supplemental article [Lévy-Leduc et al. (2011c)].

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. (2011c)].

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

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

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 $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ 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 $\mathbb{E}[(\partial^{i+j} f(X, Y) / \partial x^i \partial y^j)^2] < \infty$, 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 $p, q \geq 3$*

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

The proof of Lemma 12 is in the supplemental article [Lévy-Leduc et al. (2011c)].

The following lemma is an extension of Corollary 2.1 in Soulier (2001) and is proved in the supplemental article [Lévy-Leduc et al. (2011c)].

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 $r^* := \|\Gamma_0^{-1/2}(\Gamma_0 - \Gamma)\Gamma_0^{-1/2}\|_2 \leq (1/3 - \varepsilon)$, 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

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

where $\|f_i\|_{2,\Gamma_{0,i}}^2 = (2\pi)^{-a_i/2}|\Gamma_{0,i}|^{-1/2} \int_{\mathbb{R}^{a_i}} f_i^2(x) \exp(-x^T \Gamma_{0,i}^{-1}x/2) dx$, $i = 1, 2$, 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

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

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

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

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 $|\rho_{ij}| \leq \rho^*$ for all i, j , then $\bar{\rho}_{ij} \leq \bar{\rho}^*$ for all i, j , where

$$(47) \quad \bar{\rho}^* = \frac{4\rho^*}{1 - (\rho^*)^2}.$$

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

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

and

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

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. (2011c)].

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

$$(48) \quad R_n(t) - R_n(s) = \frac{1}{n(n-1)} \sum_{1 \leq i \neq j \leq n} J(X_i, X_j),$$

where

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

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

$$(50) \quad \|J\|_\infty \leq 4;$$

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

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

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

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

where

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

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

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

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

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

We shall consider three 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 ρ^* such that

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

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} ,

$$(57) \quad J_\varepsilon(x, y) = \int J(x - \varepsilon z, y - \varepsilon z')\varphi(z)\varphi(z') dz dz'.$$

Thus,

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

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

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

where $c_{p,q}(J_\varepsilon)$ is the (p, q) th Hermite coefficient of J_ε . We shall apply Lemma 13 with $Y_1 = (X_{i_1}, X_{i_4})$, $Y_2 = X_{i_2}$, $a_1 = 2$, $a_2 = 1$, $\Gamma_{0,1} = \text{Id}$, $\Gamma_{0,2} = 1$, $f_1 = H_p J$

and $f_2 = H_q$. Observe that $\text{Id} - \Gamma$ is a 3×3 matrix with ρ entries and hence $\|\text{Id} - \Gamma\|_2 \leq (a_1 + a_2)\|\text{Id} - \Gamma\|_\infty = 3\|\text{Id} - \Gamma\|_\infty = 3\rho^*$, 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^* 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

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

Hence,

$$(60) \quad \begin{aligned} & \mathbb{E}[J_\varepsilon(X_{i_1}, X_{i_2})J(X_{i_1}, X_{i_4})] \\ & \leq C\varepsilon^{-6}|t - s|^{1/2}(|\rho(i_4 - i_2)| + |\rho(i_2 - i_1)| + |\rho(i_4 - i_1)|). \end{aligned}$$

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

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

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

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

Using (57), Jensen’s inequality and (50),

$$(63) \quad \begin{aligned} & \mathbb{E}[(J - J_\varepsilon)^2(X_{i_1}, X_{i_2})] \\ & = \int_{\mathbb{R}^2} \left\{ \int_{\mathbb{R}^2} [J(x, y) - J(x - \varepsilon z, y - \varepsilon z')] \varphi(z) \varphi(z') \, dz \, dz' \right\}^2 \\ & \quad \times f_{i_1, i_2}(x, y) \, dx \, dy \\ & \leq \int_{\mathbb{R}^2} \left\{ \int_{\mathbb{R}^2} [J(x, y) - J(x - \varepsilon z, y - \varepsilon z')]^2 f_{i_1, i_2}(x, y) \, dx \, dy \right\} \\ & \quad \times \varphi(z) \varphi(z') \, dz \, dz' \end{aligned}$$

$$\leq C \int_{\mathbb{R}^2} \left\{ \int_{\mathbb{R}^2} |J(x, y) - J(x - \varepsilon z, y - \varepsilon z')| f_{i_1, i_2}(x, y) dx dy \right\} \times \varphi(z)\varphi(z') dz dz',$$

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

$$(64) \quad \int_{\mathbb{R}^2} \left\{ \int_{\mathbb{R}^2} |J(x, y) - J(x - \varepsilon z, y - \varepsilon z')| f_{i_1, i_2}(x, y) dx dy \right\} \varphi(z)\varphi(z') dz dz' \leq C\varepsilon \int_{\mathbb{R}^2} (|z| + |z'|)\varphi(z)\varphi(z') dz dz' \leq C\varepsilon.$$

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

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

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 is such that $\rho(k) \leq \rho^* < 1/13$, 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,

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

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 $K \rightarrow \infty$, of

$$(67) \quad J_{i_1, i_2}^K(\bar{X}_{i_1}, \bar{X}_{i_2}) = \sum_{p=1}^K \frac{c_{p_1, p_2}^{i_1, i_2}}{\sqrt{p_1! p_2!}} H_{p_1}(\bar{X}_{i_1}) H_{p_2}(\bar{X}_{i_2})$$

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

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

Thus it is enough to majorize

$$(69) \quad \mathbb{E}_{\bar{\Gamma}}[J_{i_1,i_2}^K(\bar{X}_{i_1}, \bar{X}_{i_2})J_{i_3,i_4}^K(\bar{X}_{i_3}, \bar{X}_{i_4})] \leq \sum_{1 \leq p_1, p_2 \leq K} \sum_{1 \leq p_3, p_4 \leq K} \frac{|c_{p_1, p_2}^{i_1, i_2}|}{p_1! p_2!} \frac{|c_{p_3, p_4}^{i_3, i_4}|}{p_3! p_4!} \times |\mathbb{E}_{\bar{\Gamma}}[H_{p_1}(\bar{X}_{i_1})H_{p_2}(\bar{X}_{i_2})H_{p_3}(\bar{X}_{i_3})H_{p_4}(\bar{X}_{i_4})]|.$$

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 $p_1 + \dots + p_4$ is odd. Otherwise it is bounded by a constant times a sum of products of $(p_1 + \dots + p_4)/2$ covariances. These will be denoted $\bar{\rho}_{i,j} = \mathbb{E}(\bar{X}_i \bar{X}_j)$ and are given in Lemma 14. Since $\rho(k) \leq \rho^* < 1/13$, we have that $\bar{\rho}_{i,j} \leq \bar{\rho}^* < 1/3$, where $\bar{\rho}^* = 4\rho^*/(1 - (\rho^*)^2)$ by (47). Bounding, in each product of covariances, all the covariances but two, by $\bar{\rho}^* < 1/3$, 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

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

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

$$(71) \quad \bar{A}(i_1, i_2, i_3, i_4) = \bar{\rho}_{i_1,i_3}\bar{\rho}_{i_2,i_4} + \bar{\rho}_{i_2,i_3}\bar{\rho}_{i_1,i_4} + \bar{\rho}_{i_1,i_3}\bar{\rho}_{i_2,i_3} + \bar{\rho}_{i_1,i_4}\bar{\rho}_{i_2,i_4} + \bar{\rho}_{i_1,i_3}\bar{\rho}_{i_1,i_4} + \bar{\rho}_{i_2,i_3}\bar{\rho}_{i_2,i_4},$$

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

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

Thus (69) is bounded by

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

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

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

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

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

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

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

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

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

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

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

and

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

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

$$(76) \quad \sup_{k > \tau} \rho(k) \leq \frac{1}{13} \left(1 - \sup_{1 \leq k \leq \tau} \rho(k) \right).$$

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),

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

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 $Y_1 = (X_{i_1}, X_{i_3}, X_{i_4})$, $Y_2 = X_{i_2}$, $f_1 = H_p J$, $f_2 = H_q$. 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^* 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. \square

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,*

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

where \xrightarrow{d} denotes the convergence in distribution, $k(D) = B((1 - D)/2, D)$ 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 (6) in the supplemental article [Lévy-Leduc et al. (2011c)].

The proof of Lemma 15 is in the supplemental article [Lévy-Leduc et al. (2011c)].

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

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

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

and

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

The proof of Lemma 16 is in the supplemental article [Lévy-Leduc et al. (2011c)].

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

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

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. (2011c)].

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

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

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. (2011c)].

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 \tilde{W}_n defined in (22) satisfies

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

$$(83) \quad \tilde{W}_n(r) = \alpha_{1,1}(r) \sum_{1 \leq i \neq j \leq n} X_i X_j$$

$$+ (n-1)\alpha_{2,0}(r) \sum_{i=1}^n (X_i^2 - 1) \quad \text{if } m = 2.$$

If $m = 1$, using Lemma 5.1 in Taqqu (1975), if r is fixed, $n^{D/2-2}L(n)^{-1/2}\tilde{W}_n(r)$ defined in (82) converges in distribution to $2k(D)^{-1/2}\alpha_{1,0}(r)Z_{1,D}(1)$. Then, by the Cramér–Wold device, if r_1, \dots, r_k are fixed real numbers, $k(D)^{1/2}n^{D/2-2} \times L(n)^{-1/2}(\tilde{W}_n(r_1), \dots, \tilde{W}_n(r_k))$ 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 Cramér–Wold device, $k(D)n^{D-2}L(n)^{-1}(\tilde{W}_n(r_1), \dots, \tilde{W}_n(r_k))$ 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 $\{n^{mD/2-2}L(n)^{-m/2}\tilde{W}_n(r); r \in I\}$ is tight in $\mathcal{D}(I)$. We shall do it in the case $m = 1$. By (22) in the supplemental article [Lévy-Leduc et al. (2011c)], Lemma 16 in Section 5.2 and the fact that $\tilde{\Lambda}$ is a bounded Lipschitz function, we get that there exists a positive constant C such that for all $r_1 < r_2$ in I ,

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

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

$$(n^{D/2-2}L(n)^{-1/2})^2 \mathbb{E}[|\tilde{W}_n(r_2) - \tilde{W}_n(r_1)| |\tilde{W}_n(r_3) - \tilde{W}_n(r_2)|]$$

$$\leq C|r_2 - r_1||r_3 - r_2| \leq C|r_3 - r_1|^2.$$

The tightness then follows from Theorem 15.6 of Billingsley (1968). A similar argument holds for $m = 2$. Thus, $\{n^{mD/2-2}L(n)^{-m/2}\tilde{W}_n(r); r \in I\}$ converges weakly to $\{2\alpha_{1,0}(r)k(D)^{-1/2}Z_{1,D}(1); r \in I\}$, if $m = 1$ and to $\{k(D)^{-1}[\alpha_{1,1}(r) \times Z_{1,D}(1)^2 + \alpha_{2,0}(r)Z_{2,D}(1)]; r \in I\}$, if $m = 2$. To complete the proof of Theorem 2 use (21) and Lemma 18 in Section 5.2, which ensures that $\sup_{r \in I} n^{mD/2-2} \times L(n)^{-m/2} |\tilde{R}_n(r)| = o_P(1)$, as n tends to infinity.

Acknowledgment. Murad S. Taqqu would like to thank Telecom ParisTech for their hospitality.

SUPPLEMENTARY MATERIAL

Proofs of Lemmas 9, 10, 12, 13, 14, 15, 16, 17 and 18 and some numerical experiments (DOI: [10.1214/10-AOS867SUPP](https://doi.org/10.1214/10-AOS867SUPP); .pdf). 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. A. (1994). Limit theorems for nonlinear functionals of a stationary Gaussian sequence of vectors. *Ann. Probab.* **22** 2242–2274. [MR1331224](#)
- BERAN, J. (1991). M estimators of location for Gaussian and related processes with slowly decaying serial correlations. *J. Amer. Statist. Assoc.* **86** 704–708. [MR1147095](#)
- 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. Academia, Prague.
- BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. Wiley, New York. [MR0233396](#)
- BOROVKOVA, S., BURTON, R. and DEHLING, H. (2001). Limit theorems for functionals of mixing processes with applications to U -statistics and dimension estimation. *Trans. Amer. Math. Soc.* **353** 4261–4318. [MR1851171](#)
- DEHLING, H. and TAQQU, M. S. (1989). The empirical process of some long-range dependent sequences with an application to U -statistics. *Ann. Statist.* **17** 1767–1783. [MR1026312](#)
- DEHLING, H. and TAQQU, M. S. (1991). Bivariate symmetric statistics of long-range dependent observations. *J. Statist. Plann. Inference* **28** 153–165. [MR1115815](#)
- DEWAN, I. and PRAKASA RAO, B. L. S. (2005). Wilcoxon-signed rank test for associated sequences. *Statist. Probab. Lett.* **71** 131–142. [MR2126769](#)
- FOX, R. and TAQQU, M. S. (1987). Multiple stochastic integrals with dependent integrators. *J. Multivariate Anal.* **21** 105–127. [MR0877845](#)
- GRASSBERGER, P. and PROCACCIA, I. (1983). Characterization of strange attractors. *Phys. Rev. Lett.* **50** 346–349. [MR0689681](#)
- HODGES, J. L. J. and LEHMANN, E. L. (1963). Estimates of location based on rank tests. *Ann. Math. Statist.* **34** 598–611. [MR0152070](#)
- HOEFFDING, W. (1948). A class of statistics with asymptotically normal distribution. *Ann. Math. Statist.* **19** 293–325. [MR0026294](#)
- HSING, T. and WU, W. B. (2004). On weighted U -statistics for stationary processes. *Ann. Probab.* **32** 1600–1631. [MR2060311](#)
- LÉVY-LEDUC, C., BOISTARD, H., MOULINES, E., TAQQU, M. S. and REISEN, V. A. (2011a). Large sample behavior of some well-known robust estimators under long-range dependence. *Statistics* **45** 59–71.

- LÉVY-LEDUC, C., BOISTARD, H., MOULINES, E., TAQQU, M. S. and REISEN, V. A. (2011b). Robust estimation of the scale and of the autocovariance function of Gaussian short and long-range dependent processes. *J. Time Ser. Anal.* **32** 135–156.
- LÉVY-LEDUC, C., BOISTARD, H., MOULINES, E., TAQQU, M. S. and REISEN, V. A. (2011c). Supplement to “Asymptotic properties of U -processes under long-range dependence.” DOI:10.1214/10-AOS867SUPP.
- MA, Y. and GENTON, M. (2000). Highly robust estimation of the autocovariance function. *J. Time Ser. Anal.* **21** 663–684. MR1801704
- ROUSSEEUW, P. J. and CROUX, C. (1993). Alternatives to the median absolute deviation. *J. Amer. Statist. Assoc.* **88** 1273–1283. MR1245360
- SERFLING, R. J. (1980). *Approximation Theorems of Mathematical Statistics*. Wiley, New York. MR0595165
- 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. Academic Press, New York. MR0431785
- SOULIER, P. (2001). Moment bounds and central limit theorem for functions of Gaussian vectors. *Statist. Probab. Lett.* **54** 193–203. MR1858634
- TAQQU, M. S. (1975). Weak convergence to fractional Brownian motion and to the Rosenblatt process. *Z. Wahrsch. Verw. Gebiete* **31** 287–302. MR0400329
- TAQQU, M. S. (1977). Law of the iterated logarithm for sums of non-linear functions of Gaussian variables that exhibit a long range dependence. *Z. Wahrsch. Verw. Gebiete* **40** 203–238. MR0471045
- VAN DER VAART, A. W. (1998). *Asymptotic Statistics. Cambridge Series in Statistical and Probabilistic Mathematics* **3**. Cambridge Univ. Press, Cambridge. MR1652247
- WILCOXON, F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin* **1** 80–83.

C. LÉVY-LEDUC
 E. MOULINES
 CNRS/LTCI/TELECOMPARISTECH
 46, RUE BARRAULT
 75634 PARIS CÉDEX 13
 FRANCE
 E-MAIL: celine.levy-leduc@telecom-paristech.fr
moulines@telecom-paristech.fr

H. BOISTARD
 TOULOUSE SCHOOL OF ECONOMICS
 GREMAQ, UNIVERSITÉ TOULOUSE 1
 MANUFACTURE DES TABACS
 BÂT. F, AILE J. J. LAFFONT
 21 ALLÉE DE BRIENNE
 31000 TOULOUSE
 FRANCE
 E-MAIL: helene@boistard.fr

M. S. TAQQU
 DEPARTMENT OF MATHEMATICS
 BOSTON UNIVERSITY
 111 CUMINGTON STREET
 BOSTON, MASSACHUSETTS 02215
 USA
 E-MAIL: murad@math.bu.edu

V. A. REISEN
 DEPARTAMENTO DE ESTATÍSTICA
 UNIVERSIDADE FEDERAL
 DO ESPÍRITO SANTO
 VITÓRIA/ES
 BRAZIL
 E-MAIL: valderio@cce.ufes.br