ELSEVIER Stochastic Processes and their Applications 59 (1995) 143–155

stochastic processes and their applications

Distant long-range dependent sums and regression estimation

Sándor Csörgő[a,*,1], Jan Mielniczuk[b,2]

[a]Department of Statistics, University of Michigan, 1440 Mason Hall, Ann Arbor, MI 48109-1027, USA
[b]Institute of Computer Science, Polish Academy of Sciences, Ordona 21, 01-237 Warsaw, Poland

Received 15 December 1993; revised 17 October 1994

Abstract

Consider a stationary sequence $G(Z_0), G(Z_1), \ldots$, where $G(\cdot)$ is a Borel function and $Z_0, Z_1, \ldots$ is a sequence of standard normal variables with covariance function $E(Z_0 Z_j) = j^{-\alpha} L(j)$, $j = 1, 2, \ldots$, where $E(G(Z_0)) = 0$, $E(G^2(Z_0)) < \infty$, $0 < \alpha < 1$ and $L(\cdot)$ varies slowly at infinity. Let $S_n(t) = \sum_{j=0}^{\lfloor nt \rfloor - 1} G(Z_j)$, $t \geqslant 0$, be the associated partial-sum process. The main result is that for any fixed $k \in \mathbb{N}$ and $0 < b < \infty$, a suitable norming sequence $a_n > 0$ and sequences of gap-lengths $l_{1,n}, \ldots, l_{k,n}$ such that $l_{1,n} \to \infty$ and $l_{j,n} - l_{j-1,n} \to \infty$, $j = 2, \ldots, k$, arbitrary slowly, the vector process $(S_n(t_0), S_n(l_{1,n} + t_1) - S_n(l_{1,n}), \ldots, S_n(l_{k,n} + t_k) - S_n(l_{k,n}))/a_n$, $0 \leqslant t_0, t_1, \ldots, t_k \leqslant b$, converges in distribution in $\mathscr{D}[0, b]^{k+1}$ to the vector of $k + 1$ independent Hermite processes with a rank given by $G(\cdot)$. As an application, the asymptotic behavior of the finite-dimensional distributions of kernel estimators is determined in the fixed-design regression model with errors of the form $G(Z_j)$, $j = 0, 1, \ldots$.

Keywords: Long-range dependence; Delayed sums; Joint weak convergence; Asymptotic independence; Non-parametric regression

1. Introduction

Let $Z_0, Z_1, \ldots$ be a stationary Gaussian sequence with mean $E(Z_0) = 0$, variance $E(Z_0^2) = 1$ and covariance function $r(j) := E(Z_0 Z_j) = L(j)\, j^{-\alpha}$, $j \in \mathbb{N}$, where $0 < \alpha < 1$ and $L(\cdot)$ is a function on $[0, \infty)$ slowly varying at infinity and positive in some neighborhood of infinity. Further, let $G(\cdot)$ be a Borel-measurable real function on $\mathbb{R}$ such that $E(G(Z_0)) = 0$ and $E(G^2(Z_0)) < \infty$. Consider the subordinated stationary process $G(Z_0), G(Z_1), \ldots$ and its partial sums defined by

$$S_n(t) := \sum_{j=0}^{\lfloor nt \rfloor - 1} G(Z_j), \quad t \geqslant 0, \ n \in \mathbb{N}, \tag{1.1}$$

*Corresponding author. E-mail: scsorgo@stat.lsa.umich.edu. Fax: 313 763 4676.
[1]Partially supported by NSF Grant DMS-9208067.
[2]Work partially done while this author was a Visiting Associate Professor at the Department of Statistics, University of Michigan.

SSDI 0304-4149(95)00032-1

where $\lfloor s \rfloor = \max\{k \in \mathbb{Z}: k \leqslant s\}$ denotes the integer part of $s \in \mathbb{R}$, and an empty sum is understood as zero. For $j = 0, 1, \ldots$, let $H_j(s) = (-1)^j \mathrm{d}^j \mathrm{e}^{-s^2/2}/\mathrm{d}s^j$, $s \in \mathbb{R}$, be the jth Hermite polynomial, and let $m \in \mathbb{N}$ be the smallest index for which the coefficient c_m is non-zero in the $\mathscr{L}^2(\mathbb{R}, \varphi)$-expansion of $G(\cdot)$ in (2.3) below, where $\varphi(\cdot)$ is the N(0, 1) density. Then m is called the Hermite rank of the function $G(\cdot)$ and plays an important role in the asymptotic distribution theory of the sums in (1.1). For $0 < b < \infty$, let $\mathscr{D}[0, b]$ be the metric space of all real functions defined on $[0, b]$ which are right-continuous and have left-hand limits, endowed with the σ-algebra generated by the open sets with respect to a metric that induces the Skorohod topology. Dobrushin and Major (1979) and Taqqu (1979) independently proved that $(n^{(2-m\alpha)/2}L^{m/2}(n))^{-1}S_n(\cdot)$ converges in distribution in $\mathscr{D}[0, b]$, as $n \to \infty$, to a constant multiple of $Y_m(\cdot)$, the so-called Hermite process of rank m, provided that $m\alpha < 1$. The latter is commonly referred to as the condition of long-range dependence in the underlying model. For $t \geqslant 0$, the Hermite process $Y_m(t)$ of rank m is given by the multiple Wiener–Itô–Dobrushin integral

$$Y_m(t) = C_{m,\alpha} \int_{\mathbb{R}^m}^{*} \frac{\exp\left(it\sum_{j=1}^m t_j\right) - 1}{\mathrm{i}\left(\sum_{j=1}^m t_j\right)\left(\prod_{j=1}^m |t_j|^{(1-\alpha)/2}\right)} W(\mathrm{d}t_1)\ldots W(\mathrm{d}t_m), \tag{1.2}$$

where i is the imaginary unit, $C_{m,\alpha} = (2\Gamma(\alpha)\cos(\alpha\pi/2))^{-m/2}$ with the usual gamma function $\Gamma(\cdot)$ and where $W(\cdot)$ is a complex Gaussian white noise measure such that, with $\lambda(\cdot)$ standing for Lebesgue measure and overlines producing complex conjugates, $W(B) = \overline{W(-B)}$ and $E(W(A)\overline{W(B)}) = \lambda(A \cap B)$ for all Borel sets A and B in $\mathbb{R}$, and the symbol $\int^*$ indicates that the hyperdiagonals $t_j = \pm t_k$, $j \neq k$, are excluded from the domain of integration. A different representation of $Y_m(\cdot)$ as a multiple integral with respect to ordinary standard Brownian motion is given by Taqqu (1979). The processes $Y_m(\cdot)$ have stationary increments and are self-similar, i.e. the distributional equality $\{Y_m(at): t \geqslant 0\} \stackrel{\mathscr{D}}{=} \{a^H Y_m(t): t \geqslant 0\}$ holds for $a \geqslant 0$, where $H = (2 - m\alpha)/2$ is the Hurst exponent. These properties imply that all Hermite processes have the same covariance structure determined by H:

$$E(Y_m(s)Y_m(t)) = \tfrac{1}{2}\{s^{2H} + t^{2H} - |s-t|^{2H}\}, \quad s, t \geqslant 0.$$

The version given by (1.2) is sample-continuous for all $m \in \mathbb{N}$. Note that $1/2 < H < 1$, and plugging formally $H = 1/2$ in the last equation yields the covariance function of a standard Brownian motion. The fractional Brownian motion $Y_1(\cdot)$ is still a Gaussian process, but the processes $Y_m(\cdot)$ are not Gaussian for $m > 1$. Long-range dependence is qualitatively different from various weak dependent situations when the suitably normed partial-sum process converges in distribution to standard Brownian motion.

Here we consider the following vector process: for a fixed $k \in \mathbb{N}$ and $0 < b < \infty$,

$$V_n(t_0, \ldots, t_k) := \frac{(S_n(t_0), S_n(l_{1,n} + t_1) - S_n(l_{1,n}), \ldots, S_n(l_{k,n} + t_k) - S_n(l_{k,n}))}{n^H L^{m/2}(n)}, \tag{1.3}$$

where $(t_0, t_1, \ldots, t_k) \in [0, b]^{k+1}$ and $l_{j,n} \in \mathbb{N}$, $j = 1, \ldots, k$, $n \in \mathbb{N}$. Put $l_{0,n} \equiv 0$. The question of interest is whether long-range dependence affects the joint distribution of the sums in (1.3) when these sums are distant in the sense that the gaps between them are wide relative to the number of terms in the sums. This means that we are interested in the asymptotic distribution of V_n when $l_{j,n} - l_{j-1,n} \to \infty$ for all $j = 1, \ldots, k$ as $n \to \infty$, but arbitrary slowly. Perhaps somewhat surprisingly at first encounter, we show that distant sums become asymptotically independent despite of the underlying long-range dependence. To be precise, let $\mathscr{D}[0, b]^{k+1}$ denote the $(k+1)$-fold Cartesian product of the space $\mathscr{D}[0, b]$ with itself, endowed with the product σ-algebra, which, since $\mathscr{D}[0, b]$ is separable, is the same as the σ-algebra generated by the open sets with respect to the product metric. Provided the spectral density $f(\cdot)$ pertaining to the covariance function $r(\cdot)$ exists and satisfies condition (2.2) below and $L(\cdot)$ also satisfies condition (2.3), we prove in Theorem 1 that $V_n(\cdot, \ldots, \cdot)$ converges in distribution in $\mathscr{D}[0, b]^{k+1}$ to the vector $(Y_{m,0}(\cdot), \ldots, Y_{m,k}(\cdot))$, where $Y_{m,0}(\cdot), \ldots, Y_{m,k}(\cdot)$ are independent copies of $Y_m(\cdot)$.

The distant-sum problem above was directly motivated by a problem in nonparametric regression with long-range dependent errors. As an application, we find the asymptotic form of the finite-dimensional distributions of regression function estimators. Namely, we consider the fixed-design regression model

$$Y_{j,n} = g(j/n) + G(Z_{j,n}), \quad j = 1, \ldots, n, \tag{1.4}$$

where $g(\cdot)$ is an unknown real-valued function on $[0, 1]$, $G(\cdot)$ is as above, and the triangular array $\{Z_{1,n}, \ldots, Z_{n,n}\}_{n=1}^{\infty}$ of errors is based on the Gaussian sequences $\{Z_{j,n}\}_{j=1}^{n}$ such that for each $n \in \mathbb{N}$ the finite sequence $\{Z_{j,n}\}_{j=1}^{n}$ is stationary, $E(Z_{1,n}) = 0$, $\mathrm{Var}(Z_{1,n}) = 1$ and the covariances $r(j) := E(Z_{1,n} Z_{j+1,n}) = L(j) j^{-\alpha}$, $j = 1, \ldots, n-1$, do not depend on n and satisfy the condition above, i.e. $\alpha \in (0, 1/m)$ and $L(\cdot)$ is an eventually positive function on $[0, \infty)$, slowly varying at infinity. That the covariance between the errors for $g(1/10)$ and $g(2/10)$ becomes different as the sample size $n = 10$ grows to $n = 100$, say, is reasonable in situations when, for instance, a measuring device commits the same type of errors so that the correlation between any two of them depends only on the number of time units separating the two measurements. The reader is referred to the rich survey by Robinson (1994). (In many practical situations repeated sampling with an increasing sample size n is impossible or expensive, and the statistician is faced with a single sequence of observations. This may be the reason that (1.4) is often written in the literature as $Y_j = g(j/n) + G(Z_j)$, $j = 1, \ldots, n$, as in Hall and Hart (1990), tacitly assuming the triangular-array formulation described above.) Our aim is to estimate the function $g(\cdot)$ using the Priestley–Chao estimator

$$\hat{g}_n(x) = \frac{1}{n b_n} \sum_{j=1}^{n} K\left(\frac{x - j/n}{b_n}\right) Y_{j,n}, \quad 0 \leqslant x \leqslant 1, \tag{1.5}$$

where $K(\cdot)$ is a fixed density function and $b_n > 0$ is a sequence of bandwidths tending to 0. The properties of $\hat{g}_n(\cdot)$ have been investigated by numerous authors in the case when the errors in the model (1.4) are independent or weakly dependent. On the other hand, we refer to Hall and Hart (1990) and Csörgő and Mielniczuk (1995a) for the discussion of some situations with long-range dependent errors as in (1.4). Let $0 < x_0 < x_1 \dots < x_k < 1$ be fixed for some $k \in \mathbb{N}$. It is shown in Theorem 2, as a consequence of Theorem 1, that under certain conditions imposed on $g(\cdot)$, $K(\cdot)$, $L(\cdot)$ and b_n, the vector

$$\frac{(nb_n)^{m\alpha/2}}{L^{m/2}(nb_n)}(\hat{g}_n(x_0) - g(x_0), \dots, \hat{g}_n(x_k) - g(x_k))$$

converges in distribution, as $n \to \infty$, to a $(k+1)$-dimensional random vector of components that are independent copies of a functional of the Hermite process of rank m.

Note that the norming sequence is $o((nb_n)^{1/2})$. Both the form of the limiting finite-dimensional distributions and the rate of convergence are in contrast with the independent case when the finite-dimensional distributions of the process $(nb_n)^{1/2}(\hat{g}_n(\cdot) - g(\cdot))$ converge to those of a Gaussian white noise process. On the other hand, the asymptotic independence of the components both under independence and long-range dependence indicates that the estimation problem is structurally the same in the two, very different situations. In particular, in neither situation can the empirical regression process converge weakly in $\mathscr{C}[0, 1]$. The result in Theorem 2 extends a result in Csörgő and Mielniczuk (1995a) from the case when the long-range dependent errors are normal, i.e. from the case when $G(x) = x$, $x \in \mathbb{R}$. In this special case, in view of the point-wise asymptotic normality of $\hat{g}_n(\cdot)$, Theorem 2 follows from the fact that the values of $\hat{g}_n(\cdot)$ at different points are asymptotically uncorrelated. Since, as shown by Taqqu (1975), the moments uniquely characterize the limiting process only for Hermite ranks $m \leqslant 2$, investigation of their convergence cannot yield a general result. This is why the solution of the distant-sum problem in Theorem 1 is the main tool to prove Theorem 2. The circumstance that the gaps between consecutive partial sums in Theorem 1 may widen arbitrary slowly allows the choice of the bandwidth parameter b_n as dictated solely by the bias term as long as $\lim_{n\to\infty} nb_n = \infty$; see also the remark at the end of Section 3.

2. Distant long-range dependent sums

Let $F(\cdot)$ denote the spectral measure corresponding to the covariance function $r(n) = L(n)n^{-\alpha}$ of the underlying normal sequence as in the introduction, i.e. the unique probability measure on $(-\pi, \pi]$ such that

$$r(n) = \int_{-\pi}^{\pi} e^{inx} F(dx), \quad n \in \mathbb{N}. \tag{2.1}$$

From now on we assume that the density $f(\cdot)$ of the spectral measure exists and satisfies the asymptotic equality

$$f(x) \sim C_\alpha L(1/|x|)|x|^{\alpha-1} \quad \text{as } x \to 0, \tag{2.2}$$

where $C_\alpha = (2\Gamma(\alpha)\cos(\alpha\pi/2))^{-1}$. This is equivalent to the condition $r(j) \sim j^{-\alpha}L(j)$, as $j \to \infty$, provided the covariance function $r(\cdot)$ is of bounded variation and is quasi-monotone (cf. Fox and Taqqu, 1985). A sufficient condition for (2.2) to hold can be found in Zygmund (1958, p. 187). It is that for $L(\cdot)$ in the covariance, the function $L(u)u^{\delta}$ is increasing while $L(u)\,u^{-\delta}$ is decreasing for u large enough for any fixed $\delta > 0$. Consider the Fourier–Hermite expansion of $G(\cdot)$ in Hermite polynomials. In view of the fact that $E(G(Z_0)) = 0$ and $E(G^2(Z_0)) < \infty$, for some real coefficients $c_1, c_2, \ldots$ we have

$$G(x) = \sum_{j=1}^{\infty} c_j H_j(x), \quad x \in \mathbb{R}, \tag{2.3}$$

where the convergence is in the weighted $\mathscr{L}^2$ space $\mathscr{L}^2(\mathbb{R}, \varphi)$. Let $m = \min\{k \in \mathbb{N}: c_k \neq 0\}$ be the Hermite rank of the function $G(\cdot)$. We assume throughout that $m\alpha < 1$. Furthermore, we assume that there exist $c, C > 0$ and $0 < \varepsilon < \alpha/2$ such that for every $a_1, \ldots, a_m > c$ and $b_1 > a_1, \ldots, b_m > a_m$,

$$\max_{1 \leqslant k \leqslant m} \left| \frac{L^{1/2}(b_1) \cdots L^{1/2}(b_k)}{L^{1/2}(a_1) \cdots L^{1/2}(a_k)} - 1 \right| \leqslant C \left(\frac{b_1 \cdots b_m}{a_1 \ldots a_m} \right)^{\varepsilon}. \tag{2.4}$$

Note that (2.4) is satisfied for the slowly varying function $L(s) = \max(1, \log s)$, $s \geqslant 0$, for any $0 < \varepsilon < \alpha/2$. Let $\xrightarrow{\mathscr{D}}$ denote convergence in distribution as $n \to \infty$, and $\stackrel{\mathscr{D}}{=}$ equality in distribution, and recall the definition of $V_n(\cdot, \ldots, \cdot)$ in (1.3).

Theorem 1. *Let $k \in \mathbb{N}$ be fixed and assume that conditions* (2.2) *and* (2.4) *are satisfied, $m\alpha < 1$, and $l_{1,n}, \ldots, l_{k,n}$ are sequences of positive integers such that $l_{j,n} - l_{j-1,n} \to \infty$ as $n \to \infty$, $j = 1, \ldots, k$, where $l_{0,n} \equiv 0$. Then*

$$V_n(\cdot, \ldots, \cdot) \xrightarrow{\mathscr{D}} c_m(Y_{m,0}(\cdot), \ldots, Y_{m,k}(\cdot)) \quad \text{in } \mathscr{D}[0, b]^{k+1},$$

where $Y_{m,0}(\cdot), \ldots, Y_{m,k}(\cdot)$ are independent copies of the Hermite process $Y_m(\cdot)$ of rank m, defined in (1.2).

Proof. Formally we prove the theorem only for $k = 1$; the same reasoning yields the general result. Let $l_n := l_{1,n}$ and $a_n := n^H L^{m/2}(n)$, $n \in \mathbb{N}$, where $H = 1 - m\alpha/2$ as before. Taqqu (1975) proves that the sequence $\{S_n(\cdot)/a_n\}_{n=1}^{\infty}$ is tight in $\mathscr{D}[0, b]$. Using the stationarity of the sequence $\{G(Z_j)\}$, the same is true for $\{(S_n(l_n + \cdot) - S_n(l_n))/a_n\}_{n=1}^{\infty}$. Thus $\{(S_n(\cdot), S_n(l_n + \cdot) - S_n(l_n))/a_n\}_{n=1}^{\infty}$ is tight in the product space $\mathscr{D}[0, b]^2$. So, it suffices to prove the convergence of finite-dimensional distributions (cf. Billingsley, 1968, Theorem 15.1). We prove that if $l_n \to \infty$, then

$$V_n(s, t) = (S_n(s), S_n(l_n + t) - S_n(l_n))/a_n \xrightarrow{\mathscr{D}} c_m(Y_{m,0}(s), Y_{m,1}(t)) \tag{2.5}$$

for any fixed $(s, t) \in (0, b]^2$, where, and in what follows, all convergence relations take place as $n \to \infty$ unless otherwise stated. The proof of the required general claim $V_n(s_1, \ldots, s_j, t_1, \ldots, t_r) \xrightarrow{\mathscr{D}} c_m(Y_{m,0}(s_1), \ldots, Y_{m,0}(s_j), Y_{m,1}(t_1), \ldots, Y_{m,1}(t_r))$ for any fixed

$(s_1, \ldots, s_j, t_1, \ldots, t_r) \in (0, b]^{j+r}$, $j, r \in \mathbb{N}$, where $V_n(s_1, \ldots, s_j, t_1, \ldots, t_r) := (S_n(s_1), \ldots, S_n(s_j), S_n(l_n + t_1) - S_n(l_n), \ldots, S_n(l_n + t_r) - S_n(l_n))/a_n$, is completely analogous. From now on, dealing with $V_n(s, t)$, we follow the basic outline of the proof of Theorem 1′ in Dobrushin and Major (1979). (This or Theorem 5.6 in Taqqu (1979) imply by stationarity that *individually* the two components of $V_n(s, t)$ converge in distribution to $Y_m(s)$ and $Y_m(t)$, respectively; the issue is the asymptotic independence in (2.5).)

Consider first the special case of (2.5) when $G(\cdot)$ is equal to mth Hermite polynomial $H_m(\cdot)$, and hence $c_m = 1$. Define $N_1 = \lfloor ns \rfloor$, $N_2 = \lfloor nt \rfloor$ and $N = \lfloor nl_n \rfloor - \lfloor ns \rfloor$. Thus N_1 and N_2 are the number of terms in $S_n(s)$ and $S_n(l_n + t) - S_n(l_n)$, respectively, and N is the length of the gap between them. Then the following representation based on Itô's formula holds (cf. Dobrushin and Major, 1979):

$$V_n^*(s, t) := \left(\frac{S_n(s)}{a_{N_1}}, \frac{S_n(l_n + t) - S_n(l_n)}{a_{N_2}} \right)$$

$$= \left(\frac{1}{a_{N_1}} \int_{\mathbb{R}^m} \frac{e^{iN_1(x_1 + \cdots + x_m)} - 1}{e^{i(x_1 + \cdots + x_m)} - 1} Z_F(dx_1) \ldots Z_F(dx_m), \right.$$

$$\left. \frac{1}{a_{N_2}} \int_{\mathbb{R}^m} \frac{e^{i(N_1 + N)(x_1 + \cdots + x_m)} [e^{iN_2(x_1 + \cdots + x_m)} - 1]}{e^{i(x_1 + \cdots + x_m)} - 1} Z_F(dx_1) \ldots Z_F(dx_m) \right),$$

where $Z_F(\cdot)$ denotes the random spectral measure pertaining to the spectral measure $F(\cdot)$ (cf. Major, 1981, Theorem 3.1). The respective change of variables $y_j = N_1 x_j$ and $y_j = N_2 x_j$, $j = 1, \ldots, m$, in the first and the second components yields

$$V_n^*(s, t) = \left(\int_{\mathbb{R}^m} \frac{e^{i(y_1 + \cdots + y_m)} - 1}{N_1 [e^{i(y_1 + \cdots + y_m)/N_1} - 1]} Z_{F_{N_1}}(dy_1) \cdots Z_{F_{N_1}}(dy_m), \right.$$

$$\left. \int_{\mathbb{R}^m} e^{iT_n(y_1 + \cdots + y_m)} \frac{e^{i(y_1 + \cdots + y_m)} - 1}{N_2 [e^{i(y_1 + \cdots + y_m)/N_2} - 1]} Z_{F_{N_2}}(dy_1) \ldots Z_{F_{N_2}}(dy_m) \right),$$

where $T_n = (N_1 + N)/N_2$ and $Z_{F_{N_j}}(\cdot)$, $j = 1, 2$, are particular elements of the sequence of transformed random measures $Z_{F_n}(\cdot)$ defined, for all large n to make $L(n) > 0$, by

$$Z_{F_n}(A) = \frac{n^{\alpha/2}}{L^{1/2}(n)} Z_F\left(\frac{A}{n}\right), \quad \text{for Borel sets } A \subset \mathbb{R},$$

and where, to also make the notation consistent, $F_n(\cdot)$ denotes the spectral measure that corresponds to the random spectral measure $Z_{F_n}(\cdot)$. Then using the fact that the sequence of measures $\{F_n(\cdot)\}$ converges vaguely to some measure $F_0(\cdot)$ (Major, 1981, Lemma 8.1), one can find a suitable fixed compact set $K \subset \mathbb{R}^m$ such that both integrals in $V_n^*(s, t)$ above, when taken over the complement of K, converge to zero in probability (cf. Dobrushin and Major, 1979). Further, setting $M := \sup_{(y_1, \ldots, y_m) \in K} |y_1 + \cdots + y_m|$ and using the inequality

$$\sup_{(y_1, \ldots, y_m) \in K} \left| \frac{i(y_1 + \cdots + y_m)}{n[e^{i(y_1 + \cdots + y_m)/n} - 1]} - 1 \right| \leqslant \frac{M\pi}{4\sqrt{2}n}, \quad \text{whenever } \frac{M}{n} \leqslant \frac{\pi}{2},$$

for both $n = N_1$ and $n = N_2$, we obtain

$$V_n^*(s,t) = \left(\int_K \frac{e^{i(y_1 + \cdots + y_m)} - 1}{i(y_1 + \cdots + y_m)} Z_{F_{N_1}}(dy_1) \ldots Z_{F_{N_1}}(dy_m), \right.$$

$$\left. \int_K e^{iT_n(y_1 + \cdots + y_m)} \frac{e^{i(y_1 + \cdots + y_m)} - 1}{i(y_1 + \cdots + y_m)} Z_{F_{N_2}}(dy_1) \ldots Z_{F_{N_2}}(dy_m) \right) + o_P(1).$$

Now let $U(\cdot)$ be the uniform distribution on the interval $(-\pi, \pi]$ and let $Z_U(\cdot)$ be the corresponding random spectral measure. For $v > 0$, set $U_v(A) := vU(A/v)$, so that $Z_{U_v}(A) = \sqrt{v}\, Z_U(A/v)$, as A runs through the Borel sets of $\mathbb{R}$. Put

$$g(\mathbf{x}) := C_{m,\alpha} |x_1|^{(\alpha-1)/2} \cdots |x_m|^{(\alpha-1)/2} \frac{e^{i(x_1 + \cdots + x_m)} - 1}{i(x_1 + \cdots + x_m)}, \quad \mathbf{x} = (x_1, \ldots, x_m) \in \mathbb{R}^m, \quad (2.6)$$

where the constant $C_{m,\alpha}$ is as in (1.2), and, for all n large enough,

$$h_n(\mathbf{x}) := \frac{L^{1/2}(n/|x_1|) \ldots L^{1/2}(n/|x_m|)}{L^{1/2}(n) \ldots L^{1/2}(n)}, \quad \mathbf{x} = (x_1, \ldots, x_m) \in \mathbb{R}^m. \quad (2.7)$$

If for the random variables ξ_n, η_n, ζ_n we have $\xi_n = \eta_n + o_P(1)$ and $\eta_n \overset{\mathcal{D}}{=} \zeta_n$, we write $\xi_n \overset{\mathcal{D}}{\simeq} \zeta_n$. Since, for all large n to make $L(n) > 0$, with the spectral density f of F in (2.1) we have

$$\frac{dF_n}{dU_n}(x) = \frac{n^{\alpha-1}}{L(n)} f\left(\frac{x}{n}\right), \qquad -n\pi < x \leqslant n\pi,$$

changing measures (Major, 1981, Theorem 4.4) from $Z_{F_{N_j}}(\cdot)$ to $Z_{U_{N_j}}(\cdot)$, $j = 1, 2$, gives

$$V_n^*(s,t) \overset{\mathcal{D}}{\simeq} \left(\int_K g(\mathbf{x})\, h_{N_1}(\mathbf{x})\, d\mathbf{Z}_{U_{N_1}}(d\mathbf{x}), \int_K g(\mathbf{x}) e^{iT_n(x_1 + \cdots + x_m)} h_{N_2}(\mathbf{x})\, d\mathbf{Z}_{U_{N_2}}(d\mathbf{x}) \right)$$

$$=: \overline{V}_n(s,t),$$

where $d\mathbf{Z}_{U_n}(d\mathbf{x}) = Z_{U_n}(dx_1) \ldots Z_{U_n}(dx_m)$, using also condition (2.2). If we now put

$$\tilde{V}_n(s,t) = \left(\int_K g(\mathbf{x})\, d\mathbf{Z}_{U_{N_1}}(d\mathbf{x}), \int_K g(\mathbf{x}) e^{iT_n(x_1 + \cdots + x_m)}\, d\mathbf{Z}_{U_{N_2}}(d\mathbf{x}) \right),$$

then $E(|\overline{V}_n(s,t) - \tilde{V}_n(s,t)|^2) \leqslant R_{N_1}^2(K) + R_{N_2}^2(K)$, where, writing $d\mathbf{x} = dx_1 \ldots dx_m$, this upper bound is given by $R_n^2(K) := \int_K |g(\mathbf{x})|^2 |h_n(\mathbf{x}) - 1|^2\, d\mathbf{x}$. Since both $N_1 \to \infty$ and $N_2 \to \infty$, Lemma 2 below yields

$$V_n^*(s,t) \overset{\mathcal{D}}{\simeq} \left(\int_K g(\mathbf{x})\, d\mathbf{Z}_{U_{N_1}}(d\mathbf{x}), \int_K g(\mathbf{x})\, e^{iT_n(x_1 + \cdots + x_m)} d\mathbf{Z}_{U_{N_2}}(d\mathbf{x}) \right).$$

To get back to $V_n(s,t)$ in (2.5), we must renormalize the two components here by multiplying the first by $a_{N_1}/a_n = s^H + o(1)$ and the second by $a_{N_2}/a_n = t^H + o(1)$. Next, changing variables $y_j = x_j/s$ and $y_j = x_j/t$, $j = 1, \ldots, m$, in the two integrals, respectively, noticing that the corresponding integrals over the complement of K are negligible, and that $Z_{U_v}(\cdot)$ is a version of $W(\cdot \cap (-v\pi, v\pi])$ for all $v > 0$, where $W(\cdot)$ is the complex Gaussian white noise as in the introduction, we finally arrive at the representation $V_n(s,t) \overset{\mathscr{D}}{\simeq} C_{m,\alpha}(Y_m^*(s), Y_m^\diamond(t, tT_n))$, where

$$Y_m^*(s) = \int_{\mathbb{R}^m}^* \frac{e^{is(y_1+\cdots+y_m)} - 1}{i(y_1+\cdots+y_m)\prod_{j=1}^m |y_j|^{(1-\alpha)/2}} W(dy_1)\ldots W(dy_m) = \frac{Y_m(s)}{C_{m,\alpha}},$$

a meaningful formula for any $s \in \mathbb{R}$, and, for any $u, v \in \mathbb{R}$,

$$Y_m^\diamond(u,v) = \int_{\mathbb{R}^m}^* e^{iv(y_1+\cdots+y_m)} \frac{e^{iu(y_1+\cdots+y_m)} - 1}{i(y_1+\cdots+y_m)\prod_{j=1}^m |y_j|^{(1-\alpha)/2}} W(dy_1)\ldots W(dy_m).$$

Since $tT_n = t(N_1+N)/N_2 = t\lfloor nl_n \rfloor / \lfloor nt \rfloor \sim l_n$ and $l_n \to \infty$, (2.5) and hence the theorem follows now from Lemma 1 below in the special case when $G(\cdot) = H_m(\cdot)$. (The full force of Lemma 1 is actually not needed in view of the bracketed remark following (2.5), but the lemma is of interest in its own right.) The proof for a general function $G(\cdot)$ relies on showing that the asymptotic distribution of $S_n(t)/a_n$ is the same as that of $\tilde{S}_n(t)/a_n$, where $\tilde{S}_n(t) = \sum_{j=0}^{\lfloor nt \rfloor - 1} c_m H_m(Z_j)$ and m is the Hermite rank of $G(\cdot)$. But this is a center point of the standard theory. □

Lemma 1. *We have* $(Y_m^*(s), Y_m^\diamond(t, t_n)) \overset{\mathscr{D}}{\to} (Y_m^*(s), Y_m^\diamond(t))$ *for any fixed* $s, t \in \mathbb{R}$ *and for any sequence* $t_n \to \infty$, *where* $Y_m^\diamond(\cdot)$ *is an independent copy of* $Y_m^*(\cdot)$.

Proof. Taqqu (1979) shows that integrals of the type in $Y_m^*(s)$ and $Y_m^\diamond(t, t_n)$ have an alternative representation in terms of a standard Brownian motion $B(\cdot)$ on $\mathbb{R}$. His Lemma 6.2 implies that

$$(Y_m^*(s), Y_m^\diamond(t,t_n)) \overset{\mathscr{D}}{=} \left(\int_{\mathbb{R}^m} \psi_{0,s}(\boldsymbol{v})\,dB(v_1)\ldots dB(v_m), \int_{\mathbb{R}^m} \psi_{t_n,t}(\boldsymbol{v})\,dB(v_1)\ldots dB(v_m) \right),$$

where $\boldsymbol{v} = (v_1, \ldots, v_m)$ and, with I_A as the indicator of $A \subset \mathbb{R}$, for any $\tau, u \in \mathbb{R}$,

$$\psi_{\tau,u}(\boldsymbol{v}) = \psi_{\tau,u}(v_1,\ldots,v_m) = \int_\tau^{\tau+u} \prod_{j=1}^m \frac{I_{(v_j,\infty)}(v)}{(v-v_j)^{(1+\alpha)/2}}\,dv$$

$$= \int_0^u \prod_{j=1}^m \frac{I_{(v_j-\tau,\infty)}(w)}{(w-[v_j-\tau])^{(1-\alpha)/2}}\,dw.$$

Notice that the first component $Y_m^*(s)$ is measurable with respect to the generated σ-algebra $\sigma\{B(w) - B(v)\colon -\infty < v < w \leqslant s^+\}$, where $s^+ = \max(0, s)$. Setting $D(s) = \{(v_1,\ldots,v_m) \in \mathbb{R}^m\colon \min(v_1,\ldots,v_m) > s^+\}$ and $D^c(x) = \mathbb{R}^m \backslash D(s)$, we write $Y_m^\diamond(t,t_n) = X_{m,n}^\diamond(s,t) + \varepsilon_{m,n}(s,t)$, where $X_{m,n}^\diamond(s,t) = \int_{D(s)} \psi_{t_n,t}(\boldsymbol{v})\,dB(v_1)\ldots dB(v_m)$ and $\varepsilon_{m,n}(s,t) = \int_{D^c(s)} \psi_{t_n,t}(\boldsymbol{v})\,dB(v_1)\ldots dB(v_m)$. Here $X_{m,n}^\diamond(s,t)$ is measurable with respect to $\sigma\{B(w) - B(v)\colon \infty > w > v > s^+\}$ and so, since the increments of $B(\cdot)$ are independent, $X_{m,n}^\diamond(s,t)$ is independent of $Y_m^*(s)$. Thus it suffices to show that $\varepsilon_{m,n}(s,t)$ goes to zero in probability and $X_{m,n}^\diamond(s,t) \overset{\mathscr{D}}{=} Y_m^\diamond(t) \overset{\mathscr{D}}{\to} Y_m^*(t)$.

Since the function $\psi_{t_n,t}(\cdot, \dots, \cdot)$ is symmetric in its m arguments,

$$E(\varepsilon_{m,n}^2(s,t)) = m\,E\left(\left[\int_{\{\min(v_1,\dots,v_m)=v_1<s^+\}} \psi_{t_n,t}(\boldsymbol{v})\,\mathrm{d}B(v_1)\dots\mathrm{d}B(v_m)\right]^2\right)$$

$$= m \int_{\{\min(v_1,\dots,v_m)=v_1<s^+\}} \psi_{t_n,t}^2(\boldsymbol{v})\,\mathrm{d}\boldsymbol{v}$$

$$\leqslant m \int_{A(s^+)} \psi_{t_n,t}^2(\boldsymbol{v})\,\mathrm{d}\boldsymbol{v} =: m\Psi_n(s,t),$$

where $A(x) := \{(v_1, \dots, v_m): v_1 \leqslant x, v_2 \in \mathbb{R}, \dots, v_m \in \mathbb{R}\} \subset \mathbb{R}^m$, $x \in \mathbb{R}$. Using the formula for $\psi_{t_n,t}(\cdot)$ given above and substituting $\boldsymbol{w} = \boldsymbol{v} - t_n$, we see that $\Psi_n(s,t) = \int_{A(s^+-t_n)} \psi_{0,t}^2(\boldsymbol{w})\,\mathrm{d}\boldsymbol{w}$. Since $\int_{\mathbb{R}^m} \psi_{0,t}^2(\boldsymbol{w})\,\mathrm{d}\boldsymbol{w} < \infty$, as is necessary for the existence of the stochastic integrals in the representation of $(Y_m^*(s), Y_m^\diamond(t,t_n))$ above, and since $t_n \to \infty$, implying that $A(s^+ - t_n) \to \emptyset$, the dominated convergence theorem yields $\Psi_n(s,t) \to 0$.

On the other hand, substituting $(u_1, \dots, u_m) = \boldsymbol{u} = \boldsymbol{v} - t_n$, we have $X_{m,n}^\diamond(s,t) = \int_{D(s-t_n)} \psi_{0,t}(\boldsymbol{u})\,\mathrm{d}B(u_1 + t_n)\dots\mathrm{d}B(u_m + t_n)$, so that by the stationarity of the increments of $B(\cdot)$ we have $X_{m,n}^\diamond(s,t) \overset{\mathscr{D}}{=} \int_{D(s-t_n)} \psi_{0,t}(\boldsymbol{u})\,\mathrm{d}B(u_1)\dots\mathrm{d}B(u_m)$. Since $D(s - t_n) \to \mathbb{R}^m$, the last sequence of random variables converges to $Y_m^*(t)$ in the second mean, with reference to the dominated convergence theorem again, and hence also in probability. Thus $X_{m,n}^\diamond(s,t) \overset{\mathscr{D}}{\to} Y_m^*(t)$, regardless of the value of $s \in \mathbb{R}$. □

A version of the proof, in the first distributional equality of which $\{B(v): v \in \mathbb{R}\}$ is replaced by the standard Brownian motion $\{\bar{B}(v) = B(-v): v \in \mathbb{R}\}$, shows that Lemma 1 is also true for any sequence $t_n \to -\infty$.

Lemma 2. *Let $g(\boldsymbol{x})$ and $h_n(\boldsymbol{x}), \boldsymbol{x} \in \mathbb{R}^m$, be as in (2.6) and (2.7) and let $K \subset \mathbb{R}^m$ be a compact set. If condition (2.4) holds, then $R_n^2(K) = \int_K |g(\boldsymbol{x})|^2 |h_n(\boldsymbol{x}) - 1|^2\,\mathrm{d}\boldsymbol{x} \to 0$.*

Proof. Since L is slowly varying, by an application of Corollary 1.2.1 in de Haan (1970), for any $0 < \eta < 1$, the sequence $h_n(\boldsymbol{x})$ converges uniformly to 1 on the set $K \cap A_\eta$, where $A_\eta = \{(x_1, \dots, x_m): |x_j| > \eta, j = 1, \dots, m\}$. Thus $R_n^2(K \cap A_\eta) \to 0$. It remains to show that $\lim_{\eta\downarrow 0}\limsup_{n\to\infty} R_n^2(K_\eta) = 0$, where $K_\eta = K \cap (\mathbb{R}^m \setminus A_\eta)$. This will follow by the fact that for $0 < 2\varepsilon < \alpha$,

$$I_\beta(K) := \int_K |x_1|^{\beta-1} \cdots |x_m|^{\beta-1} \left|\frac{\mathrm{e}^{\mathrm{i}(x_1+\cdots+x_m)} - 1}{\mathrm{i}(x_1 + \cdots + x_m)}\right|^2 \mathrm{d}x_1 \dots \mathrm{d}x_m < \infty, \tag{2.8}$$

where $\beta = \alpha$ or $\beta = \alpha - 2\varepsilon$. The ε below is the one from condition (2.4).

Consider the decomposition $K_\eta = (\bigcup_{k=1}^{m-1} \bigcup_{1\leqslant j_1<\cdots<j_k\leqslant m} K_{j_1,\dots,j_k}^\eta) \cup K^\eta$, where $K_{j_1,\dots,j_k}^\eta = K \cap \{(x_1, \dots, x_m): |x_{j_l}| > \eta, |x_j| \leqslant \eta, j \neq j_l, l = 1, \dots, k, j = 1, \dots, m\}$ and $K^\eta = K \cap \{(x_1, \dots, x_m): |x_j| \leqslant \eta, j = 1, \dots, m\}$. By condition (2.4), $R_n^2(K^\eta) \leqslant C^2 I_{\alpha-2\varepsilon}(K^\eta)$ for all n large enough, so $\lim_{\eta\downarrow 0}\limsup_{n\to\infty} R_n^2(K^\eta) = 0$ by (2.8) as applied with $\beta = \alpha - 2\varepsilon$. Hence it suffices to show that $\lim_{\eta\downarrow 0}\limsup_{n\to\infty} R_n^2(K_{j_1,\dots,j_k}^\eta) = 0$ for any fixed $1 \leqslant j_1 < \cdots < j_k \leqslant m$, $k = 1, \dots, m-1$. This we show for the special case $j_1 = 1, \dots, j_k = k$, $k = 1, \dots, m-1$. The general case differs only in notation. Setting

$\ell_n(x) := L^{1/2}(n/|x|)/L^{1/2}(n)$, $x \neq 0$, and using (2.4), we see that $R_n^2(K_{1,\dots,k}^\eta)$ is not greater than

$$2\int_{K_{1,\dots,k}^\eta} |g(\mathbf{x})|^2 \left\{ \left| \prod_{j=1}^{k} \ell_n(x_j) - 1 \right|^2 \left| \prod_{j=k+1}^{m} \ell_n(x_j) \right|^2 + \left| \prod_{j=k+1}^{m} \ell_n(x_j) - 1 \right|^2 \right\} \mathrm{d}\mathbf{x}$$

$$\leqslant 2\int_{K_{1,\dots,k}^\eta} |g(\mathbf{x})|^2 \left\{ 2\left| \prod_{j=1}^{k} \ell_n(x_j) - 1 \right|^2 \left[1 + \left| \prod_{j=k+1}^{m} \ell_n(x_j) - 1 \right|^2 \right] + \left| \prod_{j=k+1}^{m} \ell_n(x_j) - 1 \right|^2 \right\} \mathrm{d}\mathbf{x}$$

$$\leqslant 2\int_{K_{1,\dots,k}^\eta} |g(\mathbf{x})|^2 \left\{ 2\left| \prod_{j=1}^{k} \ell_n(x_j) - 1 \right|^2 \left[1 + \frac{C^2}{\prod_{j=1}^m |x_j|^{2\varepsilon}} \right] + \frac{C^2}{\prod_{j=1}^m |x_j|^{2\varepsilon}} \right\} \mathrm{d}\mathbf{x}.$$

Thus, using again that L is slowly varying and (2.8) for both $\beta = \alpha$ and $\beta = \alpha - 2\varepsilon$, we obtain $\limsup_{n\to\infty} R_n^2(K_{1,\dots,k}^\eta) \leqslant 2C^2 I_{\alpha - 2\varepsilon}(K_{1,\dots,k}^\eta)$. A last application of (2.8) now yields $\lim_{\eta\downarrow 0}\limsup_{n\to\infty} R_n^2(K_{1,\dots,k}^\eta) = 0$. □

3. Regression estimation under long-range dependent errors

First note that one can modify the definition of $S_n(\cdot)$ given in (1.1) to having $S_n(t) := \sum_{j=1}^{\lfloor nt \rfloor} G(Z_j)$, $t \geqslant 0$, without affecting the result of Theorem 1 in any way. From now on we shall use this modified definition since it allows a simpler representation of regression estimates. Recall the regression model in (1.4) and the definition of $\hat{g}_n(\cdot)$ in (1.5). Besides the convention for $S_n(\cdot)$ just made, in this section we understand all $k+1$ component processes of the vector process $V_n(\cdot, \dots, \cdot)$ in (1.3) to be defined in terms of the variables $G(Z_{1,\lfloor n(l_{k,n}+b)\rfloor}), \dots, G(Z_{\lfloor n(l_{k,n}+b)\rfloor, \lfloor n(l_{k,n}+b)\rfloor})$, i.e. in terms of the $\lfloor n(l_{k,n}+b)\rfloor$th row of the array. Then, with the $r(\cdot)$ as at (1.4), Theorem 1 remains valid since for each $n \in \mathbb{N}$ the distribution of $V_n(\cdot, \dots, \cdot)$ is determined by $r(1), \dots, r(\lfloor n(l_{k,n}+b)\rfloor)$ and G.

Theorem 2. *Assume that the conditions of Theorem 1 are satisfied. Furthermore suppose that the regression function $g(\cdot)$ is twice differentiable on $[0,1]$ with both derivatives bounded, the kernel $K(\cdot)$ is a density with a support contained in $(-1,1)$ and having a bounded first derivative, and that $nb_n^{(4+m\alpha)/m\alpha}/L^{1/\alpha}(nb_n) \to 0$ and $nb_n \to \infty$ as $n \to \infty$. Then for any $k \in \mathbb{N}$ and $0 < x_1 < \cdots < x_k < 1$,*

$$\frac{(nb_n)^{m\alpha/2}}{L^{m/2}(nb_n)} (\hat{g}_n(x_1) - g(x_1), \dots, \hat{g}_n(x_k) - g(x_k)) \xrightarrow{\mathcal{D}} (X_{m,1}, \dots, X_{m,k}),$$

where, with c_m as in Theorem 1,

$$X_{m,j} = c_m \int_0^2 K'(1-t)\, Y_{m,j}(t)\,\mathrm{d}t, \quad j = 1, \dots, k,$$

and $Y_{m,1}(\cdot), \dots, Y_{m,k}(\cdot)$ are independent copies of the Hermite process $Y_m(\cdot)$ of rank m.

We note before the proof that if $g_n(\cdot)$ denotes the Gasser–Müller kernel estimator of

$g(\cdot)$, defined and dealt with in Csörgő and Mielniczuk (1995a, b), then, as shown in the latter paper, $|g_n(x) - \hat{g}_n(x)| = \mathcal{O}_P((nb_n)^{-1})$ for any fixed $0 < x < 1$ whenever $K(\cdot)$ is as in Theorem 2 and $g(\cdot)$ is bounded on $[0, 1]$. Hence the result of Theorem 2 also holds true under the same conditions with g_n substituting $\hat{g}_n$.

Proof of Theorem 2. Under the conditions, two elements in the proof of the theorem in Csörgő and Mielniczuk (1995a) imply that $\sup_{b_n \leqslant x \leqslant 1-b_n}|E(\hat{g}_n(x)) - g(x)| = \mathcal{O}(b_n^2 + (nb_n)^{-1})$. Thus, by the conditions on $\{b_n\}$ and the fact that $(nb_n)^\varepsilon L^\eta(nb_n) \to \infty$ for all $\varepsilon, \eta > 0$, in particular, for $\varepsilon = H$ and $\eta = m/2$, it suffices to prove the statement when $g(x_j)$ is replaced by $E(\hat{g}_n(x_j))$, $j = 1, \ldots, k$.

First we show that the single components all converge as claimed, i.e. for $0 < x < 1$,

$$\frac{(nb_n)^{m\alpha/2}}{L^{m/2}(nb_n)}[\hat{g}_n(x) - E(\hat{g}_n(x))] \xrightarrow{\mathcal{D}} X_m = c_m \int_0^2 K'(1-t) Y_m(t)\,\mathrm{d}t. \tag{3.1}$$

Indeed, with the modified $S_n(\cdot)$, using the conditions on $K(\cdot)$, integrating by parts and subtracting $C(x) := (nb_n)^{-1}\int K'(u) S_n(x - b_n - 1/n)\,\mathrm{d}u = 0$, for all large n we obtain

$$\begin{aligned}
\hat{g}_n(x) - E(\hat{g}_n(x)) &= \frac{1}{nb_n}\sum_{j=1}^n K\left(\frac{x - j/n}{b_n}\right) G(Z_{j,n}) = \frac{1}{nb_n}\int_0^1 K\left(\frac{x-y}{b_n}\right)\mathrm{d}S_n(y) \\
&= \frac{1}{nb_n^2}\int_0^1 K'\left(\frac{x-y}{b_n}\right) S_n(y)\,\mathrm{d}y = \frac{1}{nb_n}\int_{-1}^1 K'(u) S_n(x - b_n u)\,\mathrm{d}u \\
&= \frac{1}{nb_n}\int_{-1}^1 K'(u)\left\{\sum_{j=\lfloor n(x-b_n)\rfloor}^{\lfloor n(x-b_n u)\rfloor} G(Z_{j,n})\right\}\mathrm{d}u \stackrel{\mathcal{D}}{=} \\
&\quad \frac{1}{nb_n}\int_{-1}^1 K'(u) S_{r_x(u)}(1)\,\mathrm{d}u,
\end{aligned}$$

by stationarity, where $r_x(u) = \lfloor n(x - b_n u)\rfloor - \lfloor n(x - b_n)\rfloor + 1$. Therefore,

$$\begin{aligned}
\frac{(nb_n)^{m\alpha/2}}{L^{m/2}(nb_n)}[\hat{g}_n(x) - E(\hat{g}_n(x))] &\stackrel{\mathcal{D}}{=} \int_0^2 K'(1-t)\frac{S_{r_x(1-t)}(1)}{\lfloor nb_n\rfloor^H L^{m/2}(\lfloor nb_n\rfloor)}\mathrm{d}t \\
&= \int_0^2 K'(1-t)\frac{S_{\lfloor nb_n\rfloor}(t)}{\lfloor nb_n\rfloor^H L^{m/2}(\lfloor nb_n\rfloor)}\mathrm{d}t + o_P(1),
\end{aligned}$$

where the last equation is obtained using the fact that $(nb_n)^H L^{m/2}(nb_n) \to \infty$, mentioned above. Since $nb_n \to \infty$, we have $S_{\lfloor nb_n\rfloor}(\cdot)/[\lfloor nb_n\rfloor^H L^{m/2}(\lfloor nb_n\rfloor)] \xrightarrow{\mathcal{D}} c_m Y_m(\cdot)$ in $\mathcal{D}[0, 2]$ by Theorem 5.6 in Taqqu (1979), or what is the same, by the one-component corollary of Theorem 1 here with $b = 2$. Since by the boundedness of $K'(\cdot)$ the functional $T(d) := \int_0^2 K'(1-t) d(t)\,dt$, $d \in \mathcal{D}[0, 2]$, is continuous, the mapping theorem (Billingsley, 1968, Theorem 5.1) implies (3.1).

Now let $k \geqslant 2$, and consider the points $0 < x_1 < \cdots < x_k < 1$. One can reason similarly as above. First from the lth component one subtracts $C(x_l) = 0$, $l = 1, \ldots, k$, and, in a distributional equality based on stationarity, shifts all k sums to the left upon decreasing all the indices by $\lfloor n(x_1 - b_n)\rfloor - 1$. Then one replaces the resulting sum $\sum_{j=\lfloor n(x_l-b_n)\rfloor-\lfloor n(x_1-b_n)\rfloor+1}^{\lfloor n(x_l-b_nu)\rfloor-\lfloor n(x_1-b_n)\rfloor+1} G(Z_{j,n})$ by $S_{\lfloor nb_n\rfloor}((x_l - x_1)/b_n + (1-u)) - S_{\lfloor nb_n\rfloor}((x_l - x_1)/b_n)$ in the lth component under the integral with variable $u \in (-1, 1)$ for each $l = 1, \ldots, k$. With the proper norming the error committed in the last step is $o_P(1)$, so that, meaning integration component-wise, a final change of variables $t = 1 - u$ yields

$$\frac{(nb_n)^{m\alpha/2}}{L^{m/2}(nb_n)}(\hat{g}_n(x_1) - E(g(x_1)), \ldots, \hat{g}_n(x_k) - E(g(x_k)))$$

$$\simeq \int_0^2 K'(1-t)\, V_{\lfloor nb_n\rfloor}(t, \ldots, t)\,\mathrm{d}t,$$

where, with $b = 2$ and n replaced by $\lfloor nb_n\rfloor$, the vector $V_{\lfloor nb_n\rfloor}(\cdot, \ldots, \cdot)$ here is the k-dimensional version of the one given in (1.3) with the conventions at the beginning of this section, in which $l_{j,\lfloor nb_n\rfloor} = \lfloor (x_{j+1} - x_1)/b_n\rfloor$, $j = 0, 1, \ldots, k-1$, so that $l_{0,\lfloor nb_n\rfloor} \equiv 0$ and $l_{j,\lfloor nb_n\rfloor} - l_{j-1,\lfloor nb_n\rfloor} + 1 \geq (x_{j+1} - x_j)/b_n$ for $j = 1, \ldots, k-1$. Since $x_1 < \cdots < x_k$ and $b_n \to 0$, we have $l_{j,\lfloor nb_n\rfloor} - l_{j-1,\lfloor nb_n\rfloor} \to \infty$ for all $j = 1, \ldots, k-1$. Hence Theorem 1 yields the desired result by considering the continuous mapping T: $\mathscr{D}[0,2]^k \mapsto \mathbb{R}^k$ given by $T(d_1, \ldots, d_k) = \int_0^2 K'(1-t)(d_1(t), \ldots, d_k(t))\,\mathrm{d}t$, $(d_1, \ldots, d_k) \in \mathscr{D}[0,2]^k$. □

Remark. Given $0 < x_1 < \cdots < x_k < 1$ for some $k \in \mathbb{N}$, since the support of $K(\cdot)$ is contained in $(-1, 1)$, the random variable $\hat{g}_n(x_j)$ is measurable with respect to $\mathscr{F}_{n,j} := \sigma\{\boldsymbol{Z}_{n,j}\}$ for all $j = 1, \ldots, k$, where $\boldsymbol{Z}_{n,j} := (Z_{\lfloor nx_j\rfloor - \lfloor nb_n\rfloor - 1}, \ldots, Z_{\lfloor nx_j\rfloor + \lfloor nb_n\rfloor + 1})$. Let $\boldsymbol{W}_{n,1}, \ldots, \boldsymbol{W}_{n,k}$ be *independent* normal vectors such that $\boldsymbol{W}_{n,j} \stackrel{\mathscr{D}}{=} \boldsymbol{Z}_{n,j}$, $j = 1, \ldots, k$. Then, using the normal comparison lemma (Leadbetter et al., 1983, p. 81) one can see that

$$|P\{\boldsymbol{Z}_{n,1} \leq z_{n,1}, \ldots, \boldsymbol{Z}_{n,k} \leq z_{n,k}\} - P\{\boldsymbol{W}_{n,1} \leq z_{n,1}, \ldots, \boldsymbol{W}_{n,k} \leq z_{n,k}\}|$$

$$= \mathcal{O}\left(\frac{n^2 b_n^2}{n^\alpha} L(n)\right)$$

for any vectors $z_{n,j} = (z_{\lfloor nx_j\rfloor - \lfloor nb_n\rfloor - 1}, \ldots, z_{\lfloor nx_j\rfloor + \lfloor nb_n\rfloor + 1}) \in \mathbb{R}^{2\lfloor nb_n\rfloor + 3}$, $j = 1, \ldots, k$. Hence if $b_n \to 0$ so fast that $n^{(2-\alpha)/2} b_n L^{1/2}(n) \to 0$, the σ-algebras $\mathscr{F}_{n,1}, \ldots, \mathscr{F}_{n,k}$ become asymptotically independent. Thus the statement of Theorem 2 follows, under the same conditions on $g(\cdot)$ and $K(\cdot)$, whenever $m\alpha \in (0, 1)$, $nb_n \to \infty$ and $nb_n^{2/(2-\alpha)} L^{1/(2-\alpha)}(n) \to 0$, *without* using Theorem 2, and hence without requiring conditions (2.2) and (2.4), only by proving (3.1) above by means of the original, one-dimensional Theorem 5.6 in Taqqu (1979).

However, since the long-range dependence condition is $m\alpha < 1$, and so $1 + 4/m\alpha - 2/(2-\alpha) > 5 - 2m/(2m-1) \geq 3$ for any Hermite rank $m \in \mathbb{N}$, the condition $nb_n^{2/(2-\alpha)} L^{1/(2-\alpha)}(n) \to 0$ is overly restrictive on the bandwidth sequence $\{b_n\}$ in comparison to that of Theorem 2, which requires only that $nb_n^{1+4/(m\alpha)}/L^{1/\alpha}(nb_n) \to 0$. (For instance, if $L(\cdot)$ is constant, $m = 2$ and $\alpha = 1/4$, then the former is satisfied when

$nb_n^{8/7} \to 0$, while the theorem requires only that $nb_n^9 \to 0$; choosing a very small b_n usually results in serious undersmoothing and hence a large variance.) Notice also that, ignoring $L(\cdot)$, the well-known condition $nb_n^5 \to 0$ under which the counterpart of Theorem 2 holds for independent errors may be looked upon as a limiting case of the condition $nb_n^{1+4/(m\alpha)}/L^{1/\alpha}(nb_n) \to 0$ of Theorem 2.

References

P. Billingsley, Convergence of Probability Measures (Wiley, New York, 1968).

S. Csörgő and J. Mielniczuk, Nonparametric regression under long-range dependent normal errors, Ann. Statist. 23 (1995a) to appear.

S. Csörgő and J. Mielniczuk, Close short-range dependent sums and regression estimation, Acta Sci. Math. (Szeged) 60 (1995b) to appear.

R.L. Dobrushin and P. Major, Non-central limit theorems for non-linear functionals of Gaussian fields, Z. Wahrsch. verw. Gebiete 50 (1979) 27–52.

S. Fox and M.S. Taqqu, Noncentral limit theorems for quadratic forms in random variables having long-range dependence, Ann. Probab. 13 (1985) 428–446.

L. de Haan, On Regular Variation and its Application to Weak Convergence of Sample Extremes (Mathematical Centre Tract 32, Amsterdam, 1970).

P. Hall and J.D. Hart, Nonparametric regression with long-range dependence, Stochastic Process. Appl. 36 (1990) 339–351.

M.R. Leadbetter, G. Lindgren and H. Rootzén, Extremes and Related Properties of Random Sequences and Processes (Springer, New York, 1983).

P. Major, Multiple Wiener-Itô Integrals, Lecture Notes in Mathematics, Vol. 848, (Springer, Berlin, 1981).

P.M. Robinson, Time series with strong dependence, in: C. A. Sims, ed., Advances in Econometrics: Sixth World Congress, Vol. I (Cambridge Univ. Press, Cambridge, 1994) pp. 47–95.

M.S. Taqqu, Weak convergence to the fractional Brownian motion and to the Rosenblatt process, Z. Wahrsch. verw. Gebiete 31 (1975) 287–302.

M.S. Taqqu, Convergence of integrated processes of arbitrary Hermite rank, Z. Wahrsch. verw. Gebiete 50 (1979) 53–83.

A. Zygmund, Trigonometric Series, Vol. I (Cambridge Univ. Press, Cambridge, 1958).