*The Annals of Statistics*
1995, Vol. 23, No. 3, 1000–1014

# NONPARAMETRIC REGRESSION UNDER LONG-RANGE DEPENDENT NORMAL ERRORS

BY SÁNDOR CSÖRGŐ[1] AND JAN MIELNICZUK[2]

*University of Michigan and Polish Academy of Sciences*

We consider the fixed-design regression model with long-range dependent normal errors and show that the finite-dimensional distributions of the properly normalized Gasser–Müller and Priestley–Chao estimators of the regression function converge to those of a white noise process. Furthermore, the distributions of the suitably renormalized maximal deviations over an increasingly finer grid converge to the Gumbel distribution. These results contrast with our previous findings for the corresponding problem of estimating the marginal density of long-range dependent stationary sequences.

**1. Introduction.** Consider the fixed-design regression model

$$Y_{i,n} = g(i/n) + Z_{i,n}, \qquad i = 1, 2, \dots, n, \tag{1.1}$$

where $g\colon [0,1] \mapsto \mathbb{R}$ is some function with smoothness properties described later. For each $n$, we observe the random variables $Y_{1,n}, Y_{2,n}, \dots, Y_{n,n}$, and our aim is to estimate the unknown function $g$ based on this information. Here $Z_{1,n}, Z_{2,n}, \dots, Z_{n,n}$ is a Gaussian array such that for each $n$, the finite sequence $\{Z_{i,n}\}_{i=1}^n$ is stationary, $E(Z_{1,n}) = 0$, $\mathrm{Var}(Z_{1,n}) = 1$ and $r(k) := \mathrm{Cov}(Z_{1,n}, Z_{k+1,n}) = E(Z_{1,n}Z_{k+1,n}) = k^{-\alpha}L(k)$, $k = 1, 2, \dots$, where $0 < \alpha < 1$ is a fixed constant and $L$ is a function defined on $[0,\infty)$, slowly varying at infinity and positive in some neighborhood of infinity. The array $\{Z_{i,n}\}_{i=1}^n$ is long-range dependent in the sense that $\sum_{k=1}^\infty |r(k)| = \infty$. In what follows we suppress the dependence of $Y_{i,n}$ and $Z_{i,n}$ on $n$ in our notation. Let $K$ be a fixed density function and let $b_n \to 0$ be positive smoothing constants. (Any asymptotic statement or relation is meant as $n \to \infty$ unless otherwise specified.) We consider the kernel regression-function estimator $g_n$ of $g$ introduced by Gasser and Müller (1979), and the Nadaraya (1964) and Watson (1964) type estimator $\hat{g}_n$ proposed by Priestley and Chao (1972), given as

$$\begin{aligned} g_n(x) &= \sum_{i=1}^n \left\{ \frac{1}{b_n} \int_{(i-1)/n}^{i/n} K\left(\frac{x-t}{b_n}\right) dt \right\} Y_i \qquad \text{and} \\ \hat{g}_n(x) &= \frac{1}{nb_n} \sum_{i=1}^n K\left(\frac{x - i/n}{b_n}\right) Y_i, \end{aligned} \tag{1.2}$$

Received May 1993; revised June 1994.
[1]Partially supported by NSF Grant DMS-92-08067.
[2]Work done while this author was Visiting Associate Professor, Department of Statistics, University of Michigan.
*AMS* 1991 *subject classifications*. Primary 62G07, 62M99; secondary 60F17.
*Key words and phrases*. Long-range dependence, kernel regression estimators, fixed design, maximal deviations, finite-dimensional distributions.

for each $x \in [0,1]$. Both $g_n$ and $\hat{g}_n$ have been investigated by numerous authors when the errors are independent or weakly dependent. On the other hand, a recent broad statistical review of long-range dependence has been given by Beran (1992) with discussion and many references. In particular, Hall and Hart (1990) consider the model (1.1) with general second-order stationary errors for which $r(k) \sim Ck^{-\alpha}$ as $k \to \infty$ for some constant $C > 0$ and $\alpha \in (0,1]$, assuming that $g$ is twice differentiable. They prove that the smallest mean squared error of the Nadaraya–Watson kernel estimator is $\mathscr{O}(n^{-4\alpha/(4+\alpha)})$, and this is achieved with an optimal bandwidth that is proportional to $n^{-\alpha/(4+\alpha)}$.

The problem of estimating the marginal distribution function $F$ of a more general long-range dependent stationary sequence $\{X_i = G(Z_i)\}_{i=1}^{\infty}$, where $G$ is any real-valued Borel function on $\mathbb{R}$, has been considered in the pioneering work of Dehling and Taqqu (1989), using the empirical distribution function $F_n$ of $X_1, \dots, X_n$. Under the suitable long-range dependence condition for $G$, they determined norming constants $a_n$ such that $t_n := a_n\{F_n - F\}$ converges weakly with respect to the supremum distance to a process of the form $TY$, where $T$ is a deterministic function and $Y$ is a fixed random variable. Considering a kernel estimator $f_n(x) = \sum_{i=1}^{n} K((x - X_i)/b_n)/(nb_n)$, $x \in \mathbb{R}$, of the marginal density $f = F'$, Csörgő and Mielniczuk (1995a) proved under appropriate smoothness and other conditions that the density process $a_n\{f_n - f\}$ converges weakly, again with respect to the supremum distance, to the derivative $T'Y$ of the Dehling–Taqqu limiting process. Such a phenomenon is wholly impossible when the observations are independent or weakly dependent. The problem of density estimation with long-range dependent observations is structurally different from that of with independent observations.

In particular, if $X_1, X_2 \dots$ are independent, then under some regularity conditions the finite-dimensional distributions of the centered process $D_n^* := \sqrt{nb_n}\{f_n - E(f_n)\}$ converge to those of a stationary mean-zero Gaussian white noise process. Hence $D_n^*$ cannot converge weakly with respect to supremum distances. Supposing that $f$ is concentrated on $[0,1]$, say, under suitable regularity conditions, Woodroofe (1967) has proved that

$$\lim_{n\to\infty} P\left\{\left(2nb_n \log \frac{1}{b_n}\right)^{1/2} \max_{x \in \Pi_n} \frac{|f_n(x) - f(x)|}{\left[f(x)\int K^2(u)\,du\right]^{1/2}} - c_n \le x\right\} \tag{1.3}$$
$$= \exp(-2e^{-x}), \qquad x \in \mathbb{R},$$

where $\Pi_n$ is as in (2.5) and $c_n$ is as in Theorem 1 in Section 3 below. The celebrated result of Bickel and Rosenblatt (1973) then states that (1.3) remains valid even when $\max_{x \in \Pi_n}$ is replaced by $\sup_{0 \le x \le 1}$ and, in the case when $K$ satisfies (2.6) below, $c_n$ is changed to $\tilde{c}_n = 2\log(1/b_n) - a(K)$, where the constant $a(K)$ depends only on $K$.

On the other hand, it is well known that the problem of estimating a density from independent observations and the problem of estimating a regression function with independent errors are completely analogous. In particular, replacing $Z_i$ in (1.1) by $X_i$, independent and identically distributed mean-zero errors with variance 1, the finite-dimensional distributions of the process $R_n^* := \sqrt{nb_n}\{\hat{g}_n - E(\hat{g}_n)\}$ converge to those of another stationary mean-zero Gaussian white noise process; compare Benedetti (1977) or Stadtmüller (1986), for example. Again, $R_n^*$ cannot converge weakly in the usual function space $\mathscr{C}[0,1]$. Also, under appropriate regularity conditions, Stadtmüller (1986) has established the Bickel–Rosenblatt analogue of (1.3), proving that

$$\lim_{n\to\infty} P\left\{\left(2nb_n \log \frac{1}{b_n}\right)^{1/2} \max_{0\le x\le 1} \frac{|\hat{g}_n(x) - g(x)|}{\left[\int K^2(u)\,du\right]^{1/2}} - \tilde{c}_n \le x\right\} = \exp(-2e^{-x}), \qquad x \in \mathbb{R}. \tag{1.4}$$

In this paper we show that, with norming constants depending on $\alpha$ and $L$, the behavior of $g_n$ and $\hat{g}_n$ under long-range dependent normal errors is similar to or parallel with the behavior of $g_n$ and $\hat{g}_n$ themselves under independent errors. Some auxiliary observations and groundwork are in the next section, leading to the main results in Section 3. These results are then discussed in Section 4. Throughout the paper, convergence in distribution and in probability will be denoted by $\to_{\mathscr{D}}$ and $\to_P$, respectively.

**2. Auxiliary results.** In this section we concentrate only on the Gasser–Müller estimator $g_n$ in (1.2), using the basic assumptions between (1.1) and (1.2). The usual decomposition of $g_n - g$ into a deterministic and a stochastic term for every $x \in [0,1]$ is

$$\begin{aligned} g_n(x) - g(x) &= \left\{\sum_{i=1}^{n}\left[\frac{1}{b_n}\int_{(i-1)/n}^{i/n} K\left(\frac{x-t}{b_n}\right)dt\right] g\left(\frac{i}{n}\right) - g(x)\right\} \\ &\quad + \sum_{i=1}^{n}\left[\frac{1}{b_n}\int_{(i-1)/n}^{i/n} K\left(\frac{x-t}{b_n}\right)dt\right]\left[Y_i - g\left(\frac{i}{n}\right)\right]. \end{aligned} \tag{2.1}$$

In the present section we restrict attention to the stochastic second term which equals

$$\sum_{i=1}^{n}\left[\frac{1}{b_n}\int_{(i-1)/n}^{i/n} K\left(\frac{x-t}{b_n}\right)dt\right] Z_i = \frac{1}{b_n}\int_0^1 K\left(\frac{x-t}{b_n}\right) Z_{\lfloor nt\rfloor+1}\,dt,$$

where $\lfloor t\rfloor = \max\{k \in \mathbb{Z}: k \le t\}$, $t \in \mathbb{R}$; $\mathbb{Z} = \{0, \pm 1, \pm 2, \dots\}$. For $x \in [0,1]$ we consider

$$\begin{aligned} W_n(x) &:= \left(\frac{(nb_n)^{\alpha}}{L(nb_n)}\right)^{1/2} \sum_{i=1}^{n}\left[\frac{1}{b_n}\int_{(i-1)/n}^{i/n} K\left(\frac{x-t}{b_n}\right)dt\right]\left[Y_i - g\left(\frac{i}{n}\right)\right] \\ &= \left(\frac{(nb_n)^{\alpha}}{L(nb_n)}\right)^{1/2} \frac{1}{b_n}\int_0^1 K\left(\frac{x-t}{b_n}\right) Z_{\lfloor nt\rfloor+1}\,dt. \end{aligned} \tag{2.2}$$

Since $\{W_n\}$ is a sequence of Gaussian processes on $[0,1]$, it is essential to explore the covariance structure. Putting $M_n = (nb_n)^{\alpha}/L(nb_n)$ and adopting the convention $L(0)/0^{\alpha} = 1$, for any $0 < x \le y < 1$, the covariance $\operatorname{Cov}(W_n(x), W_n(y))$ is

$$
\begin{aligned}
&M_n \int_0^1 \int_0^1 \frac{1}{b_n^2} K\left(\frac{x-s}{b_n}\right) K\left(\frac{y-t}{b_n}\right) \operatorname{Cov}\left(Z_{\lfloor ns \rfloor + 1}, Z_{\lfloor nt \rfloor + 1}\right) ds\, dt \\
&\quad = M_n \int_{(x-1)/b_n}^{x/b_n} \int_{(y-1)/b_n}^{y/b_n} K(u) K(v) \\
&\qquad \times \frac{L(|\lfloor ny - nvb_n \rfloor - \lfloor nx - nub_n \rfloor|)}{|\lfloor ny - nvb_n \rfloor - \lfloor nx - nub_n \rfloor|^{\alpha}}\, du\, dv.
\end{aligned}
\tag{2.3}
$$

An analysis of this formula is given in three lemmas under the assumptions that

$$b_n = k_n/n \quad \text{for some } k_n \in \mathbb{N},\ k_n \le n,\ n \in \mathbb{N}, \text{ where } \lim_{n \to \infty} k_n = \infty \tag{2.4}$$

and where $\mathbb{N} = \{1, 2, \dots\}$. The points $x, y$ will be taken from the grid

$$\Pi_n = \{b_n, 2b_n, 3b_n, \dots, p_n b_n\} \quad \text{where } p_n = \lfloor b_n^{-1} \rfloor - 1. \tag{2.5}$$

Furthermore, the following conditions will be used on $K$ and $L$, not necessarily always:

$$\text{the support of } K \text{ is contained in the open interval } (-1, 1); \tag{2.6}$$

$$K \text{ is differentiable on } \mathbb{R} \text{ with derivative } K'; \tag{2.7}$$

$$C_L := \sup\{|L(x)|/|L(nb_n)| \colon 1 \le x \le n,\ n \in \mathbb{N}\} < \infty; \tag{2.8}$$

$$C_L^* := \sup\{|L(x)|/|L(nb_n)| \colon 1 \le x \le nb_n,\ n \in \mathbb{N}\} < \infty. \tag{2.9}$$

LEMMA 1. *Suppose* (2.4) *and* (2.6). *Then for all $n$ large enough and grid points $x, y \in \Pi_n$, the covariance* $\operatorname{Cov}(W_n(x), W_n(y))$ *depends only on $y - x$.*

PROOF. Since $nx$ and $ny$ are integers,

$$\lfloor ny - nvb_n \rfloor - \lfloor nx - nub_n \rfloor = n(y - x) + \lfloor -nvb_n \rfloor - \lfloor -nub_n \rfloor,$$

and the statement follows from (2.3). This is because $b_n \to 0$, so by (2.5) and (2.6) the integration is on the square $[-1, 1]^2$ for all $n$ sufficiently large. □

LEMMA 2. *Let $x, y \in \Pi_n$ such that $y - x = ib_n$ for some $i = 1, \dots, p_n - 1$, let $K$ be a bounded kernel satisfying* (2.6) *and suppose* (2.4) *and* (2.8) *hold. Then*

$$|\operatorname{Cov}(W_n(x), W_n(y))| \le \frac{C(\delta_{i1} + \delta_{i2})}{(nb_n)^{1-\alpha} |L(nb_n)|} + \frac{C^*}{i^{\alpha}}$$

*for all $n$ large enough, where $\delta_{ij}$ is Kronecker's symbol and $C$ and $C^*$ are positive constants depending only on $L$, $K$ and $\alpha$, but not on $i$ and $n$.*

PROOF. It follows from (2.3) that $|\mathrm{Cov}(W_n(x), W_n(y))| \le I_n(x, y) + J_n(x, y)$, where

$$I_n(x,y) = \frac{(nb_n)^\alpha}{|L(nb_n)|} \int\int_{|t-s|\le 2/n} \frac{1}{b_n^2} K\left(\frac{x-s}{b_n}\right) K\left(\frac{y-t}{b_n}\right) \times \frac{|L(\|\lfloor nt \rfloor - \lfloor ns \rfloor\|)|}{\|\lfloor nt \rfloor - \lfloor ns \rfloor\|^\alpha}\, ds\, dt, \tag{2.10}$$

$$J_n(x,y) = \frac{(nb_n)^\alpha}{|L(nb_n)|} \int\int_{|t-s|> 2/n} \frac{1}{b_n^2} K\left(\frac{x-s}{b_n}\right) K\left(\frac{y-t}{b_n}\right) \times \frac{|L(\|\lfloor nt \rfloor - \lfloor ns \rfloor\|)|}{\|\lfloor nt \rfloor - \lfloor ns \rfloor\|^\alpha}\, ds\, dt. \tag{2.11}$$

Since (2.6) holds, it can be seen using (2.4) that for all $n$ sufficiently large, $I_n(x, y) \ne 0$ only if $y - x = b_n$ or $y - x = 2b_n$. Meanwhile, if $|t - s| \le 2/n$, the quantities $\|\lfloor nt \rfloor - \lfloor ns \rfloor\|$ can only take the values $0, 1, 2$, so that for $i = 1$ or $i = 2$ and all $n$ large enough,

$$\begin{aligned} I_n(x,y) &\le C_1 \frac{(nb_n)^\alpha}{|L(nb_n)|} \int\int_{|y-x-b_n(v-u)|\le 2/n} K(u)K(v)\, du\, dv \\ &\le C_1 C_2^2 \frac{(nb_n)^\alpha}{|L(nb_n)|} \int\int_{\{u,v:\, |i-(v-u)|\le 2/nb_n\}\cap[-1,1]^2} du\, dv \\ &\le C_1 C_2^2 \frac{(nb_n)^\alpha}{|L(nb_n)|} \frac{4}{(nb_n)^i} \le \frac{4C_1C_2^2}{(nb_n)^{1-\alpha}|L(nb_n)|}, \end{aligned} \tag{2.12}$$

where $C_1 = \max\{|L(k)|/k^\alpha : k = 0, 1, 2\}$ and $C_2 = \sup\{K(u): -1 \le u \le 1\} < \infty$, since $nb_n \to \infty$ from (2.4). The proof that $J_n(x, y) \le C^* i^{-\alpha}$, for all $n$ large enough, follows from the first half of the proof of (2.14) in Lemma 3 below and is postponed until then. □

The proof of the next lemma requires the notion of a fractional Brownian motion $\{B_H(t): t \in \mathbb{R}\}$ with index $H = H(\alpha) = 1 - \alpha/2$, so that $\frac{1}{2} < H < 1$. It is a mean-zero Gaussian process with stationary increments determined by the covariance function $E(B_H(s)B_H(t)) = \{|s|^{2H} + |t|^{2H} - |t-s|^{2H}\}/2$ for $s, t \in \mathbb{R}$. Assume (2.6) and (2.7), so that $\int K'(u)\, du = 0$. Introduce the stationary Gaussian process

$$V_\alpha(x) = \left[\frac{2}{2H(2H-1)}\right]^{1/2} \int_{-\infty}^{\infty} K'(x-t) B_H(t)\, dt, \qquad x \in \mathbb{R}.$$

It is straightforward to see that its mean is zero and that

$$E(V_\alpha(x+y)V_\alpha(x)) = R_\alpha(y) := \int\int \frac{K(u)K(v)}{|y-(v-u)|^\alpha}\, du\, dv, \qquad x \in \mathbb{R},\ y \ge 0. \tag{2.13}$$

While the previous lemma gives a useful bound for $|\mathrm{Cov}(W_n(x), W_n(y))|$ when $y - x = ib_n$ for a large $i$, the last one takes care of the case when $i$ is small.

LEMMA 3. *If conditions* (2.4) *and* (2.6)–(2.8) *are satisfied, then for every* $j \in \mathbb{N}$ *there exist an* $n_0(j) \in \mathbb{N}$ *and a* $\delta > 0$ *such that*

$$\max_{i=1,\dots,j}\ \max_{x,y\in\Pi_n:\, y-x=ib_n} |\mathrm{Cov}(W_n(x), W_n(y))| \le R_\alpha(0) - \delta$$

*whenever* $n \ge n_0(j)$, *where* $R_\alpha(0)$ *is defined in* (2.13).

PROOF. Note first that $L(nb_n)(nb_n)^{1-\alpha} \to \infty$ since $nb_n \to \infty$ by (2.4) and $L$ is slowly varying. Hence, using the proof of Lemma 2, we see that it suffices to show the existence of $n_0(j) \in \mathbb{N}$ and $\varepsilon > 0$ such that

$$\max_{i=1,\dots,j}\ \max_{x,y\in\Pi_n:\, y-x=ib_n} J_n(x,y) < R_\alpha(0) - \varepsilon \quad \text{whenever } n \ge n_0(j), \tag{2.14}$$

where $J_n(x, y)$ is as in (2.11). Choose any $j \in \mathbb{N}$ and let $x, y \in \Pi_n$ be a pair of values such that $y - x = ib_n$ for some $1 \le i \le j$. We have

$$\begin{aligned} J_n(x,y) &\le \iint_{|y-x-b_n(w-u)|>2/n} K(u)K(w) \\ &\qquad \times \frac{|L(|\lfloor n(y-wb_n)\rfloor - \lfloor n(x-ub_n)\rfloor|)|}{(|(y-x)/b_n - (w-u)| - 2/(nb_n))^\alpha |L(nb_n)|}\,du\,dw \\ &= \iint_{i-(v-u)>0} K(u)K\left(v - \frac{2}{nb_n}\right) \\ &\qquad \times \frac{|L(|\lfloor n(y-vb_n)+2\rfloor - \lfloor n(x-ub_n)\rfloor|)|}{|i-(v-u)|^\alpha |L(nb_n)|}\,du\,dv \\ &\quad + \iint_{i-(v-u)<0} K(u)K\left(v + \frac{2}{nb_n}\right) \\ &\qquad \times \frac{|L(|\lfloor n(y-vb_n)-2\rfloor - \lfloor n(x-ub_n)\rfloor|)|}{|i-(v-u)|^\alpha |L(nb_n)|}\,du\,dv. \end{aligned}$$

Using (2.4) and Corollary 1.2.1(4) of de Haan (1970) for the slowly varying $L$, the integrands in the last two integrals converge to $K(u)K(v)|i-(v-u)|^{-\alpha}$, except for points $(u, v)$ the planar Lebesgue measure of which is zero. Furthermore, using (2.6)–(2.8), we see that both integrands are dominated by the integrable function $C_L C_2^2 |i-(v-u)|^{-\alpha} I_{[-1,1]\times[-2,2]}(u,v)$, $(u,v) \in \mathbb{R}^2$, where $C_2 = \sup_{-1\le u\le 1} K(u) < \infty$. [This yields the bound for $J_n(x, y)$ needed in Lemma 2.] Thus the upper bound goes to

$$\iint_{i-(v-u)>0} \frac{K(u)K(v)}{|i-(v-u)|^{\alpha}}\,du\,dv + \iint_{i-(v-u)<0} \frac{K(u)K(v)}{|i-(v-u)|^{\alpha}}\,du\,dv = R_{\alpha}(i),$$

where $R_{\alpha}(i)$ is as in (2.13). Hence

$$\limsup_{n\to\infty} \max_{i=1,\dots,j} \max_{x,y\in\Pi_n: y-x=ib_n} J_n(x,y) \le \max_{i=1,\dots,j} R_{\alpha}(i).$$

So, (2.14) will follows if we show that $\max_{i=1,\dots,j} R_{\alpha}(i) \le R_{\alpha}(0) - \varepsilon$ for some $\varepsilon > 0$.

Suppose there is no such $\varepsilon$. Then there exists a $1 \le i \le j$ such that $R_{\alpha}(i) \ge R_{\alpha}(0)$. By (2.13) and the Schwarz inequality we must have $R_{\alpha}(i) = R_{\alpha}(0)$. Since $E(V_{\alpha}(i)) = 0 = E(V_{\alpha}(0))$, this implies that $P\{V_{\alpha}(i) = V_{\alpha}(0)\} = 1$. By stationarity, the same argument gives $P\{V_{\alpha}(2i) = V_{\alpha}(i)\} = 1$. Thus $P\{V_{\alpha}(2i) = V_{\alpha}(0)\} = 1$, and continuing this we obtain that $P\{V_{\alpha}(mi) = V_{\alpha}(0)\} = 1$ for every $m \in \mathbb{N}$. Hence $R_{\alpha}(mi) = R_{\alpha}(0)$ by (2.13) again, for every $m \in \mathbb{N}$. However, this is impossible since $R_{\alpha}(mi) = \mathscr{O}(m^{-\alpha})$ as $m \to \infty$. $\square$

The result below describes the convergence of finite-dimensional distributions of

$$W_n(x) = \left(\frac{(nb_n)^{\alpha}}{L(nb_n)}\right)^{1/2} \{g_n(x) - E(g_n(x))\}, \qquad 0 \le x \le 1, \tag{2.15}$$

the process given in (2.2). Condition (2.4) is *not* required in this proposition.

PROPOSITION 1. *If $nb_n \to \infty$ and conditions* (2.6), (2.7) *and* (2.9) *are satisfied, then for every fixed $k \in \mathbb{N}$ and every fixed points $0 < x_1 < \cdots < x_k < 1$,*

$$(W_n(x_1),\dots,W_n(x_k)) \to_{\mathscr{D}} \sigma(\alpha,K)(N_1,\dots,N_k),$$

*where $N_1,\dots,N_k$ are independent standard normal random variables and*

$$\sigma(\alpha,K) = \sqrt{R_{\alpha}(0)} = \left[\iint \frac{K(u)K(v)}{|u-v|^{\alpha}}\,du\,dv\right]^{1/2}. \tag{2.16}$$

PROOF. First note that $l_n(u,v) := \|ny - nvb_n\rfloor - \lfloor nx - nub_n\rfloor|/(nb_n) \to \infty$ uniformly in $(u,v) \in (-1,1)^2$ if $x \ne y$. Thus for any function $L$, slowly varying at infinity,

$$\sup_{-1\le u,v\le 1} \left| \frac{L(nb_n l_n(u,v))(nb_n)^{\alpha}}{L(nb_n)(nb_n l_n(u,v))^{\alpha}} - \frac{1}{l_n^{\alpha}(u,v)} \right| \to 0,$$

which follows again from Corollary 1.2.1(4) in de Haan (1970). Hence the last line of (2.3) yields the asymptotic equality

$$\operatorname{Cov}(W_n(x), W_n(y)) \sim \iint \frac{K(u)K(v)}{l_n^\alpha(u,v)}\,du\,dv, \qquad 0 < x < y < 1.$$

This implies that $\operatorname{Cov}(W_n(x), W_n(y)) \to 0$ for all $0 < x < y < 1$. Since the random variables $W_n(x_1), \dots, W_n(x_k)$ are normally distributed, it suffices, therefore, to show that $\operatorname{Var}(W_n(x)) \to R_\alpha(0)$ for every fixed $0 < x < 1$.

By (2.3) we have $\operatorname{Var}(W_n(x)) \le I_n(x, x) + J_n(x, x)$, the bounds given in (2.10) and (2.11). An easy version of (2.12) shows that $I_n(x, x) \le [8C_1C_2^2]/[(nb_n)^{1-\alpha}|L(nb_n)|] \to 0$. Also, using only (2.9), a version of the first part of the proof of (2.14), with $i = 0$, gives that $\limsup_{n\to\infty} J_n(x, x) \le R_\alpha(0)$. Thus, $\limsup_{n\to\infty} \operatorname{Var}(W_n(x)) \le R_\alpha(0)$.

Conversely, using (2.3) and choosing $A > 2$ so that $L(x) \ge 0$ for all $x \ge A - 2$, for all $n$ large enough to make $L(nb_n) > 0$ we obtain $\operatorname{Var}(W_n(x)) \ge J_{n,A}(x) - I_{n,A}(x)$, where

$$\begin{aligned} J_{n,A}(x) &= \frac{(nb_n)^\alpha}{L(nb_n)} \iint_{|t-s|>A/n} \frac{1}{b_n^2} K\left(\frac{x-s}{b_n}\right) K\left(\frac{x-t}{b_n}\right) \frac{L(|\lfloor nt\rfloor - \lfloor ns\rfloor|)}{|\lfloor nt\rfloor - \lfloor ns\rfloor|^\alpha}\,ds\,dt \\ &\ge \iint_{|t-s|>A/n} \frac{1}{b_n^2} K\left(\frac{x-s}{b_n}\right) K\left(\frac{x-t}{b_n}\right) \frac{L(|\lfloor nt\rfloor - \lfloor ns\rfloor|)\,b_n^\alpha}{(|t-s| + 2/n)^\alpha L(nb_n)}\,ds\,dt \\ &\ge \iint_{|w-u|>A/nb_n} K(u)K(w) \\ &\qquad \times \frac{L(|\lfloor n(x-wb_n)\rfloor - \lfloor n(x-ub_n)\rfloor|)}{(|w-u| + A/(nb_n))^\alpha L(nb_n)}\,du\,dw \end{aligned}$$

and

$$\begin{aligned} I_{n,A}(x) &= \frac{(nb_n)^\alpha}{|L(nb_n)|} \iint_{|t-s|\le A/n} \frac{1}{b_n^2} K\left(\frac{x-s}{b_n}\right) K\left(\frac{x-t}{b_n}\right) \\ &\qquad \times \frac{|L(|\lfloor nt\rfloor - \lfloor ns\rfloor|)|}{|\lfloor nt\rfloor - \lfloor ns\rfloor|^\alpha}\,ds\,dt. \end{aligned}$$

Now $I_{n,A}(x) \to 0$ just as for $I_n(x, x)$. Also, the version of the first part of the proof of (2.14) just mentioned, with $A$ replacing 2, shows that the lower bound for $J_{n,A}(x)$ goes to $R_\alpha(0)$. Thus, $\liminf_{n\to\infty} \operatorname{Var}(W_n(x)) \ge R_\alpha(0)$. Therefore, $\operatorname{Var}(W_n(x)) \to R_\alpha(0)$. $\square$

The next proposition describes the asymptotic distribution of one- and two-sided maximal deviations of $g_n$ from $E(g_n)$ over the grid $\Pi_n$ in (2.5).

PROPOSITION 2. *If conditions* (2.4) *and* (2.6)–(2.8) *are satisfied, then*

$$P\left\{\sqrt{2\log p_n}\left(\frac{\max(W_n(b_n), W_n(2b_n), \dots, W_n(p_n b_n))}{\sigma(\alpha, K)} - c_n^*\right) \le x\right\}$$
$$\to \exp(-e^{-x}), \qquad x \in \mathbb{R},$$

*and*

$$P\left\{\sqrt{2\log p_n}\left(\frac{\max(|W_n(b_n)|, |W_n(2b_n)|, \dots, |W_n(p_n b_n)|)}{\sigma(\alpha, K)} - c_n^*\right) \le x\right\}$$
$$\to \exp(-2e^{-x}), \qquad x \in \mathbb{R},$$

*where* $\sigma(\alpha, K)$ *is as in* (2.16) *and*

$$c_n^* = \sqrt{2\log p_n} - [\log\log p_n + \log(4\pi)] / \left[2\sqrt{2\log p_n}\right].$$

PROOF. Setting $r_n(i) := \operatorname{Cov}(W_n(b_n)/\sigma_n, W_n((i+1)b_n)/\sigma_n)$, $i = 1, \dots, p_n - 1$, where $\sigma_n = \sqrt{r_n(0)}\,[\operatorname{Var}(W_n(b_n))]^{1/2}$, we will verify the condition

$$B_n^* := \sum_{i=1}^{p_n - 1} \frac{|r_n(i)|(p_n - i)}{\sqrt{1 - r_n^2(i)}} \frac{(\log p_n)^{1/(1+|r_n(i)|)}}{p_n^{2/(1+|r_n(i)|)}} \to 0.$$

Since $T_{1,n} := W_n(b_n)/\sigma_n$, $T_{2,n} := W_n(2b_n)/\sigma_n, \dots, T_{p_n,n} := W_n(p_n b_n)/\sigma_n$ is a stationary Gaussian sequence by Lemma 1, with mean zero and unit variance, in view of Theorem 9.2.1 of Berman (1992), this will imply the first statement with $\sigma(\alpha, K)$ replaced by $\sigma_n$. However, by (2.3), the proof of Lemma 1 and the proof of Proposition 1 for the convergence,

$$\sigma_n^2 = r_n(0) = \frac{(nb_n)^\alpha}{L(nb_n)} \int\!\!\int K(u)K(v) \frac{L(|\lfloor -nvb_n \rfloor - \lfloor -nub_n \rfloor|)}{|\lfloor -nvb_n \rfloor - \lfloor -nub_n \rfloor)|^\alpha}\, du\, dv \to \sigma^2$$
$$= \sigma^2(\alpha, K).$$

Thus the first assertion will follow by Slutsky's theorem if we show that $B_n^* \to 0$.

Fix any $0 < \varepsilon < \alpha$. Since $p_n \to \infty$, and $r_n(i) \to 0$ as $n \to \infty$ and $i \to \infty$ by Lemma 2 and the fact that $\sigma_n^2 \to \sigma^2$, we can select a $j = j(\varepsilon) \ge 2$ and an $n_* = n_*(\varepsilon) \ge n_0(j)$ such that $p_n \ge e$ and $2/[1 + |r_n(i)|] > 2 - \varepsilon$, and hence also $\varepsilon/(2 - \varepsilon) > |r_n(i)|$, for $n \ge n_*$ and $i \ge j$, where $n_0(j)$ is of Lemma 3. Increasing $p_n - i$ to $p_n$ and the exponent of $\log p_n$ to 1, decreasing the exponent of $p_n$ to $2 - \varepsilon$ and using Lemma 2, for $n \geq n_*$ we obtain

$$\begin{aligned}\sum_{i=j+1}^{p_n - 1} \frac{|r_n(i)|(p_n - i)}{\sqrt{1 - r_n^2(i)}} \frac{(\log p_n)^{1/(1+|r_n(i)|)}}{p_n^{2/(1+|r_n(i)|)}} &< \frac{C_\varepsilon}{\sigma_n^2} \sum_{i=j+1}^{p_n - 1} \left( \frac{C(\delta_{i1} + \delta_{i2})}{(nb_n)^{1-\alpha}|L(nb_n)|} + \frac{C^*}{i^\alpha} \right) \frac{\log p_n}{p_n^{1-\varepsilon}} \\ &< \frac{C_\varepsilon}{\sigma_n^2} \frac{C^* p_n^{1-\alpha}}{1 - \alpha} \frac{\log p_n}{p_n^{1-\varepsilon}} = \frac{C_\varepsilon}{\sigma_n^2} \frac{C^*}{1 - \alpha} \frac{p_n^\varepsilon \log p_n}{p_n^\alpha},\end{aligned}$$

where $C_\varepsilon = (2 - \varepsilon)/\sqrt{4 - 4\varepsilon}$, for $\sum_{i=j+1}^{p_n - 1} i^{-\alpha} \leq \int_1^{p_n} x^{-\alpha}\, dx < p_n^{1-\alpha}/(1 - \alpha)$. The bound tends to zero by the choice of $\varepsilon$. On the other hand, if $n \geq n_* \geq n_0(j)$, then $|r_n(i)| < \rho\sigma^2/\sigma_n^2$ for some $0 < \rho < 1$ for every $1 \leq i \leq j$ by Lemma 3. Since $\sigma_n^2 \to \sigma^2$, we have $|r_n(i)| < \tau$ for some $\rho < \tau < 1$ and all $n \geq n^*$ for some $n^* \geq n_*$. Thus for all $n \geq n^*$,

$$\begin{aligned}\sum_{i=1}^{j} \frac{|r_n(i)|(p_n - i)}{\sqrt{1 - r_n^2(i)}} \frac{(\log p_n)^{1/(1+|r_n(i)|)}}{p_n^{2/(1+|r_n(i)|)}} &\leq \frac{\tau}{\sqrt{1 - \tau^2}} \sum_{i=1}^{j} \frac{p_n \log p_n}{p_n^{2/(1+\tau)}} \\ &\leq \frac{j\tau}{\sqrt{1 - \tau^2}} \frac{\log p_n}{p_n^{(1-\tau)/(1+\tau)}},\end{aligned}$$

and this upper bound goes to zero. Hence $B_n^* \to 0$, proving the first statement.

To prove the second statement, let $w_n(x) = c_n^* + x(2 \log p_n)^{-1/2}$, where $x \in \mathbb{R}$ is arbitrarily fixed. Then, using the above notation, the stationary form of Lemma 11.1.2 of Leadbetter, Lindgren and Rootzén (1983) gives

$$\begin{aligned}&\left| P\left\{ \bigcap_{i=1}^{p_n} \{|T_{i,n}| \leq w_n(x)\} \right\} - P\left\{ \bigcap_{i=1}^{p_n} \{|N_i| \leq w_n(x)\} \right\} \right| \\ &\qquad \leq \frac{4}{\pi} \sum_{i=1}^{p_n - 1} \frac{|r_n(i)|(p_n - i)}{\sqrt{1 - r_n^2(i)}} \exp\left( -\frac{w_n^2(x)}{1 + |r_n(i)|} \right),\end{aligned}$$

where $N_1, \dots, N_{p_n}$ are independent standard normal variables. Since $B_n^* \to 0$, the upper bound here also goes to zero. So, if $P\{\max(|N_1|, \dots, |N_{p_n}|) \leq w_n(x)\} \to \exp(-2e^{-x})$, the second statement will also follow. However, this is a well-known result, following from the classical Fisher–Tippett theorem for $\max(N_1, \dots, N_n)$ and the asymptotic independence statement in Theorem 1.8.3 in Leadbetter, Lindgren and Rootzén (1983). □

**3. Main results.** Assume the regression model (1.1) with long-range dependent errors satisfying the conditions described between (1.1) and (1.2) and recall the definition of $\Pi_n$ in (2.5). Our findings for both the Gasser–

Müller estimator $g_n$ and the Priestley–Chao estimator $\hat{g}_n$, given in (1.2), are summarized in the following theorem.

THEOREM 1. (i) *Suppose that $g$ is twice differentiable on* $[0,1]$ *with both derivatives $g'$ and $g''$ bounded there. Assume that the function $L$ and the bandwidth sequence $\{b_n\}$ are such that $nb_n \to \infty$ and $nb_n^{1+4/\alpha}/L^{1/\alpha}(nb_n) \to 0$, and they satisfy condition* (2.9). *Finally, suppose that the kernel $K$ is a density such that $\int K(u)u\,du = 0$ and* (2.6) *and* (2.7) *hold. Then for every fixed $k \in \mathbb{N}$ and every fixed points $0 < x_1 < \cdots < x_k < 1$,*

$$\left(\frac{(nb_n)^\alpha}{L(nb_n)}\right)^{1/2}(g_n(x_1) - g(x_1),\dots,g_n(x_k) - g(x_k)) \\ \to_{\mathscr{D}} \sigma(\alpha,K)(N_1,\dots,N_k), \tag{3.1}$$

*where $N_1,\dots,N_k$ are independent standard normal variables and $\sigma(\alpha,K)$ is as in* (2.16).

(ii) *Let $g$ and $K$ be as in part* (i). *If the function $L$ and the bandwidth sequence $\{b_n\}$ are such that $nb_n^{1+4/\alpha}\log^{1/\alpha}(1/b_n)/L^{1/\alpha}(nb_n) \to 0$ and* (2.4) *and* (2.8) *hold, then*

$$P\left\{\left(2\frac{(nb_n)^\alpha}{L(nb_n)}\log\frac{1}{b_n}\right)^{1/2}\max_{x\in\Pi_n}\frac{\pm[g_n(x) - g(x)]}{\sigma(\alpha,K)} - c_n \le x\right\} \\ \to \exp(-e^{-x}), \qquad x \in \mathbb{R}, \tag{3.2}$$

*and*

$$P\left\{\left(2\frac{(nb_n)^\alpha}{L(nb_n)}\log\frac{1}{b_n}\right)^{1/2}\max_{x\in\Pi_n}\frac{|g_n(x) - g(x)|}{\sigma(\alpha,K)} - c_n \le x\right\} \\ \to \exp(-2e^{-x}), \qquad x \in \mathbb{R}, \tag{3.3}$$

*where $\sigma(\alpha,K)$ is as in* (2.16) *and*

$$c_n = 2\log(1/b_n) - \tfrac{1}{2}\log\log(1/b_n) + \tfrac{1}{2}\log(4\pi).$$

(iii) *If all the conditions in part* (i) *are satisfied, with the condition on how slow $b_n \to 0$ strengthened to $nb_n^{1+\eta} \to \infty$ for some $\eta > 0$, and the derivative $K'$ is bounded, then we have* (3.1) *for $\hat{g}_n$ as well.*

(iv) *If all the conditions in part* (ii) *are satisfied, with the condition on how slow $b_n \to 0$ strengthened to $nb_n^{1+\eta} \to \infty$ for some $\eta > 0$, and the derivative $K'$ is bounded, then we have* (3.2) *and* (3.3) *for $\hat{g}_n$ as well.*

Before proving the theorem we note that if $b_n = n^{-\delta}$ or, to also satisfy (2.4), $b_n = \lceil n^{1-\delta}\rceil/n$, $n \in \mathbb{N}$, where $\lceil y\rceil = \min\{k \in \mathbb{Z}\colon y \le k\}$, $y \in \mathbb{R}$, then both conditions on $\{b_n\}$ of (i) and (ii) are satisfied whenever $\alpha/(4+\alpha) < \delta < 1$. In this case, if the slowly varying function $L$ is nonnegative and nondecreasing on $[1,\infty)$, then $C_L^* \le C_L \le \sup\{L(n)/L(n^{1-\delta})\colon n \in \mathbb{N}\} =: C_L(\delta)$. Hence

$C_L(\delta) < \infty$ is a sufficient condition for (2.8) and (2.9). Note also that $C_L(\delta) = 1/(1-\delta)$ for the slowly varying function $L(x) = \log x,\ x > 0$. However, $L(x) = 1/\log x,\ x > 0$, does not satisfy even the weaker condition (2.9).

PROOF OF THEOREM 1. Observe first that Proposition 2 also holds true for the random variables $-W_n(b_n), \dots, -W_n(p_n b_n)$ replacing $W_n(b_n), \dots, W_n(p_n b_n)$. Set

$$\Delta_n := \left( \frac{(nb_n)^\alpha}{L(nb_n)} \right)^{1/2} \sup_{b_n \le x \le 1-b_n} |E(g_n(x)) - g(x)|.$$

Then, by (2.15) and Propositions 1 and 2, the latter also with $-W_n$, for (i) and (ii), it suffices to show that $\Delta_n \to 0$ and $\Delta_n^2 \log(1/b_n) \to 0$, respectively. This will give the statements in (ii) with $p_n$ replacing $1/b_n$ in the logarithms, but since $p_n b_n \to 1$, the desired forms follow. Since $K$ is a density function and $b_n \le x \le 1 - b_n$, by (2.1) we have

$$E(g_n(x)) - g(x) = \frac{1}{b_n}\int_0^1 K\left(\frac{x-t}{b_n}\right)\left[g\left(\frac{\lfloor nt \rfloor + 1}{n}\right) - g(t)\right] dt$$
$$+ \frac{1}{b_n}\int_0^1 K\left(\frac{x-t}{b_n}\right)[g(t) - g(x)]\, dt.$$

Now $\sup_{0 \le t \le 1}|g((\lfloor nt \rfloor + 1)/n) - g(t)| = \mathcal{O}(1/n)$ by the boundedness of $g'$, while the other conditions on $g$ imply that the second term is $\mathcal{O}(b_n^2)$ uniformly in $x \in [b_n, 1-b_n]$; compare Bretagnolle and Huber (1979). Hence $\sup_{b_n \le t \le 1-b_n}|E(g_n(x)) - g(x)| = \mathcal{O}(b_n^2 + n^{-1})$. Thus $\Delta_n \to 0$ and $\Delta_n^2 \log(1/b_n) \to 0$ under the respective conditions for (i) and (ii).

We prove (iii) and (iv) for the Priestly–Chao estimator $\hat{g}_n$ by showing that if $C_g := \sup\{|g(x)|: 0 \le x \le 1\} < \infty$ and if $K$ satisfies (2.6) and (2.7), then

$$\sup_{b_n \le x \le 1-b_n} |g_n(x) - \hat{g}_n(x)| = \mathcal{O}_P\left(\frac{\sqrt{\log n}}{nb_n}\right). \tag{3.4}$$

Since $nb_n^{1+\eta} \to \infty$ for some $\eta > 0$, this is of course more than enough to imply that

$$\left(\frac{(nb_n)^\alpha}{L(nb_n)} \log\frac{1}{b_n}\right)^{1/2} \sup_{b_n \le x \le 1-b_n} |g_n(x) - \hat{g}_n(x)| \to_P 0,$$

and hence the results for the Priestley–Chao estimator $\hat{g}_n$ in (iii) and (iv) follow from the respectively results in (i) and (ii) for the Gasser–Müller estimator $g_n$.

To prove (3.4), we write $g_n - \hat{g}_n = d_n + e_n$, where $d_n = [g_n - E(g_n)] - [\hat{g}_n - E(\hat{g}_n)]$ and $e_n = E(g_n) - E(\hat{g}_n)$. Using (1.2) and (2.1), an easy calculation gives

$$
\begin{aligned}
|d_n(x)| &\le \frac{1}{b_n} \sum_{i=1}^{n} \left| \int_{(i-1)/n}^{i/n} \left[ K\left(\frac{x-s}{b_n}\right) - K\left(\frac{x-i/n}{b_n}\right) \right] ds \right| |Z_i| \\
&\le \frac{1}{b_n} \sum_{i=\lceil n(x-b_n)\rceil+1}^{\lfloor n(x+b_n)\rfloor} \left\{ \int_{(i-1)/n}^{i/n} \left| K\left(\frac{x-s}{b_n}\right) - K\left(\frac{x-i/n}{b_n}\right) \right| ds \right\} |Z_i|.
\end{aligned} \tag{3.5}
$$

Since the integrands do not exceed $C_3/(nb_n)$, where $C_3 = \sup_{-1 \le x \le 1} |K'(x)| < \infty$, we have

$$\sup_{b_n \le x \le 1-b_n} |d_n(x)| \le C_3 \frac{2nb_n}{(nb_n)^2} \max(|Z_1|, \dots, |Z_n|) = \mathcal{O}_P\left( \frac{\sqrt{\log n}}{nb_n} \right).$$

This is because

$$\max(|Z_1|, \dots, |Z_n|) = \mathcal{O}_P\left(\sqrt{\log n}\right) \quad \text{for} \quad \max(Z_1, \dots, Z_n) = \mathcal{O}_P\left(\sqrt{\log n}\right).$$

The latter in turn comes from Theorem 9.2.2 in Berman (1992), for example, since the basic condition $r(k) = k^{-\alpha}L(k)$, $k \in \mathbb{N}$, implies that $r(n)\log n \to 0$. Since the second line of (3.5), with $|g(i/n)|$ replacing $|Z_i|$, is a bound for $|e_n(x)|$, the same argument gives $\sup_{b_n \le x \le 1-b_n} |e_n(x)| \le 2C_3 C_g/(nb_n)$. Putting the two bounds together yields (3.4). □

**4. Discussion.** Let $h(x) = x^{\alpha/2}/L^{1/2}(x)$, $q(x) = h(x)/x$, $x > 0$. The theorem shows that if $g_n^{\diamond}$ is any one of $g_n$ and $\hat{g}_n$, then, apart from the change of the norming factor $\sqrt{nb_n}$ to $h(nb_n)$, the behavior of the asymptotic distribution of $g_n^{\diamond} - g$ under long-range dependent Gaussian errors is completely analogous to its behavior under independent and identically distributed errors. The Hall and Hart (1990) bandwidth $b_n^* = C_0 n^{-\alpha/(4+\alpha)}$, where $C_0 = C_0(\alpha, g, K) > 0$ is some constant, is the unattainable boundary of all bandwidths of the form $b_n = n^{-\delta}$, $\alpha/(4+\alpha) < \delta < 1$, that satisfy the conditions of the theorem. The situation is the same with independent errors, when for the corresponding quantities we have $b_n^* = Cn^{-1/5}$ and $1/5 < \delta < 1$. This behavior is in contrast with that of the kernel density process $f_n - f$ based on long-range dependent observations described in the Introduction: While $f_n - f$ behaves differently under independence and long-range dependence, this is not so for $g_n^{\diamond} - g$. What is the heuristic reason for all this?

Using $\hat{g}_n$ and the notation from the Introduction to answer this question, define $t_n^* = \sqrt{n}\{F_n - F\}$, $S_n^*(t) = n^{-1/2}\sum_{i=1}^{\lfloor nt \rfloor} X_i$, $S_n(t) = q(n)\sum_{i=1}^{\lfloor nt \rfloor} Z_i$, $0 \le t \le 1$, $D_n^0 = a_n\{f_n - E(f_n)\}$ and $R_n^0 = h(nb_n)\{\hat{g}_n - E(\hat{g}_n)\}$. Writing $K_n(y, x) = K([x-y]/b_n)$, we have $D_n^*(x) = b_n^{-1/2} \int K_n(y, x)\, dt_n^*(y)$ and $D_n^0(x) = b_n^{-1} \int K_n(y, x)\, dt_n(y)$, $x \in \mathbb{R}$. Also, $R_n^*(x) = b_n^{-1/2} \int K_n(t, x)\, dS_n^*(t)$ and, at least for $nb_n \in \mathbb{N}$, we have $R_n^0(x) = \int K_n(t, x)\, dS_{nb_n}(t/b_n)$, $x \in [0, 1]$.

Thus, if $X_1, X_2, \dots$ are independent, $D_n^*(\cdot) \approx b_n^{-1/2} \int K_n(F^{-1}(t), \cdot)\, dB_0(t)$, where $F^{-1}$ is the quantile function pertaining to $F$ and $B_0$ is a Brownian bridge on $[0,1]$, and $R_n^*(\cdot) \approx b_n^{-1/2} \int K_n(t, \cdot)\, dB(t)$ with a standard Brownian motion $B$. Since $B_0$ and $B$ behave locally the same way, and none of them is differentiable, the asymptotic properties of $D_n^*$ and $R_n^*$ become completely analogous. On the other hand, if $X_1 = G(Z_1)$, $X_2 = G(Z_2), \dots$ are long-range dependent, then $D_n^0(\cdot) \approx Y b_n^{-1} \int K_n(y, \cdot)\, dT(y) \approx T'(\cdot) Y$. However, if (1.1) holds for the regression problem, then, by Theorem 5.1 of Taqqu (1975), $S_n(\cdot)$ converges weakly to $C_H B_N(\cdot)$, where $C_H$ is a constant and $B_H$ is the fractional Brownian motion defined above (2.13). Hence if $nb_n$ is an integer, the self-similarity property of $B_H$ suggests that

$$R_n^0(\cdot) \approx C_H \int K_n(t, \cdot)\, dB_H(t/b_n) =_{\mathscr{D}} C_H b_n^{-H} \int K_n(t, \cdot)\, dB_H(t) =: R_n^{\diamond}(\cdot).$$

Therefore, the dissimilarity of $D_n^0$ and $R_n^0$ under long-range dependence results from the difference between the asymptotic behavior of the empirical process $t_n$ and the partial-sum process $S_n$. [If the former is based on the normal $Z_1, Z_2, \dots$ in (1.1), that is, $G(x) \equiv x$, then $a_n$ and $q(n)$ are proportional.] At the same time, the heuristic approximations for $R_n^*$ and $R_n^0$ are similar. (If we formally take $\alpha = 1$, so that $H = 1/2$, they become the same.) Thus the heuristics also explain that, modulo the norming constants, the regression problem under long-range dependence is still very much like under independence.

Parts (ii) and (iv) of the theorem present the maximal deviation theory of nonparametric regression with long-range dependent normal errors at the same level of sophistication where the corresponding theory of density estimation under independence was with Woodroofe's (1967) result in (1.3). It is then natural to ask whether (3.2) and (3.3) remain true with $\max_{x \in \Pi_n}$ replaced by $\sup_{0 \le x \le 1}$ as in the independent case, that is, whether the counterpart of (1.4) is feasible. In Csörgő and Mielniczuk (1995b), we prove the corresponding $\sup_{0 \le x \le 1}$ results for $R_n^{\diamond}$ in the above heuristic approximation, a kind of an extreme-value theory for smoothed fractional Brownian motions.

**Acknowledgments.** We thank Hans-Georg Müller for calling our attention to Stadtmüller (1986), and Peter Hall and Ulrich Stadtmüller for sending us their papers. The helpful remarks of two referees and an Associate Editor are greatly appreciated.

## REFERENCES

BENEDETTI, J. K. (1977). On the nonparametric estimation of regression functions. *J. Roy. Statist. Soc. Ser. B* **39** 248–253.

BERAN, J. (1992). Statistical methods for data with long-range dependence. *Statist. Sci.* **7** 404–427.

BERMAN, S. M. (1992). *Sojourns and Extremes of Stochastic Processes*. Wadsworth and Brooks/Cole, Pacific Grove, CA.

BICKEL, P. J. and ROSENBLATT, M. (1973). On some global measures of the deviations of density function estimates. *Ann. Statist.* **1** 1071–1095.

BRETAGNOLLE, J. and HUBER, C. (1979). Estimation de densités: risque minimax. *Z. Wahrsch. Verw. Gebiete* **47** 119–137.

CSÖRGŐ, S. and MIELNICZUK, J. (1995a). Density estimation under long-range dependence. *Ann. Statist.* **23** 990–999.

CSÖRGŐ, S. and MIELNICZUK, J. (1995b). Extreme values of derivatives of smoothed fractional Brownian motions. *Probab. Math. Statist.* **16**. To appear.

DE HAAN, L. (1970). *On Regular Variation and Its Application to the Weak Convergence of Sample Extremes. Math. Centre Tract* **32**. Math. Centrum, Amsterdam.

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.

GASSER, T. and MÜLLER, H.-G. (1979). Kernel estimation of regression functions. *Smoothing Techniques for Curve Estimation. Lecture Notes in Math.* **757** 23–68. Springer, New York.

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

LEADBETTER, M. R., LINDGREN, G. and ROOTZÉN, H. (1983). *Extremes and Related Properties of Random Sequences and Processes*. Springer, New York.

NADARAYA, E. A. (1964). On estimating regression. *Theory Probab. Appl.* **9** 141–142.

PRIESTLEY, M. B. and CHAO, M. T. (1972). Nonparametric function fitting. *J. Roy. Statist. Soc. Ser. B* **34** 385–392.

STADTMÜLLER, U. (1986). Asymptotic properties of nonparametric curve estimates. *Period. Math. Hungar.* **17** 83–108.

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

WATSON, G. S. (1964). Smooth regression analysis. *Sankhyā Ser. A* **26** 359–372.

WOODROOFE, M. (1967). On the maximum deviation of the sample density. *Ann. Math. Statist.* **38** 475–481.

DEPARTMENT OF STATISTICS
UNIVERSITY OF MICHIGAN
1444 MASON HALL
ANN ARBOR, MICHIGAN 48109-1027

INSTITUTE OF COMPUTER SCIENCE
POLISH ACADEMY OF SCIENCES
ORDONA 21
01-237 WARSAW
POLAND