The Annals of Statistics
1989, Vol. 17, No. 4, 1767–1783

THE EMPIRICAL PROCESS OF SOME LONG-RANGE DEPENDENT SEQUENCES WITH AN APPLICATION TO U-STATISTICS

BY HEROLD DEHLING[1] AND MURAD S. TAQQU[2]

Boston University

Let $(X_j)_{j=1}^{\infty}$ be a stationary, mean-zero Gaussian process with covariances $r(k) = EX_{k+1}X_1$ satisfying $r(0) = 1$ and $r(k) = k^{-D}L(k)$ where D is small and L is slowly varying at infinity. Consider the two-parameter empirical process for $G(X_j)$,

$$\left\{ F_N(x, t) = \frac{1}{N} \sum_{j=1}^{[Nt]} \left[1\{G(X_j) \le x\} - P(G(X_1) \le x) \right]; \right.$$

$$\left. -\infty < x < +\infty, 0 \le t \le 1 \right\},$$

where G is any measurable function. Noncentral limit theorems are obtained for $F_N(x, t)$ and they are used to derive the asymptotic behavior of some suitably normalized von Mises statistics and U-statistics based on the $G(X_j)$'s. The limiting processes are structurally different from those encountered in the i.i.d. case.

1. Introduction. It is well known that if $(Y_j)_{j=1}^{\infty}$ are i.i.d. random variables with cumulative distribution function $F(x) = P(Y_1 \le x)$, then the normalized two-parameter empirical process $N^{-1/2}\sum_{j=1}^{[Nt]}\{1\{Y_j \le x\} - F(x)\}$ converges weakly to the Kiefer process [see Müller (1970) or Shorack and Wellner (1986)]. Moreover, von Mises and U-statistics defined in terms of the Y_j's converge weakly to Wiener–Itô integrals of the Kiefer process [see Denker, Grillenberger and Keller (1985) and Mandelbaum and Taqqu (1984)]. For a general discussion on empirical processes and symmetric statistics, see Dehling and Taqqu (1987).

We want to find out what happens when $(Y_j)_{j=1}^{\infty}$ is a strongly dependent stationary sequence (i.e., when its spectral density diverges at the origin). We focus on the case where Y_j is a nonlinear function of Gaussian variables. In that case, normalized sums of Y_j may themselves have non-Gaussian limits which are expressed as multiple Wiener–Itô integrals [see Dobrushin and Major (1979) and Taqqu (1979)]. Thus the study of von Mises and U-statistics defined in terms of the strongly dependent $\{Y_j\}_{j=1}^{\infty}$ may involve a combination of two sets of multiple integrals, one set resulting from the strong dependence, the other from the degree of degeneracy in the von Mises or U-statistics. It turns out that

Received December 1987; revised November 1988.
[1]Research supported by the Deutsche Forschungsgemeinschaft.
[2]Research supported by NSF Grant ECS-86-96090 at Boston University.
AMS 1980 *subject classifications*. Primary 60F17, 62G30.
Key words and phrases. von Mises statistics, U-statistics, Hermite polynomials, empirical process, long-range dependence.

strong dependence causes a separation of variables which prevents the interference of these two sets of multiple integrals. This paper deals with weak convergence. Extensions of the results to a functional law of the iterated logarithm can be found in Dehling and Taqqu (1988a).

Let $(X_j)_{j=1}^{\infty}$ be a stationary, mean-zero Gaussian process with covariances $r(k) = EX_{k+1}X_1$ satisfying

$$(1.1)\qquad \begin{aligned} r(0) &= 1,\\ r(k) &= k^{-D}L(k), \qquad 0 < D < 1, \end{aligned}$$

where L is slowly varying at infinity and is positive for all large k. Such an $(X_j)_{j=1}^{\infty}$ exhibits long-range dependence. We want to study the empirical process of $Y_j = G(X_j)$ where G is *any* measurable function. We will also obtain results on von Mises and U-statistics for the Y_j's.

Let

$$(1.2)\qquad F_N(x) = \frac{1}{N}\sum_{j=1}^{N} 1\{G(X_j) \le x\}$$

be the empirical distribution function of the observations $G(X_1), \dots, G(X_N)$ and let $F(x)$ denote the c.d.f. of $G(X_j)$. Since $r(k) \to 0$ as $k \to \infty$, both $\{X_j\}$ and $\{G(X_j)\}$ are ergodic sequences. The ergodic theorem ensures that $F_N(x) \to F(x)$ a.s. for fixed x, and a technique similar to that used for the classical Glivenko–Cantelli theorem yields

$$\sup_{-\infty < x < +\infty} |F_N(x) - F(x)| \to 0 \quad \text{a.s.}$$

We first want to examine whether a properly renormalized process $d_N^{-1}N(F_N - F)$ converges to a nondegenerate limiting process. For a *fixed* x, one can use the theory of nonlinear functionals of Gaussian processes to find the right normalization. To do so, we must expand the function $\Delta_x(\cdot) = 1\{G(\cdot) \le x\} - F(x)$ in Hermite polynomials. These are $H_0(X) = 1$, $H_1(X) = X$, $H_2(X) = X^2 - 1, \dots$. We get

$$1\{G(X) \le x\} - F(x) = \sum_{q=m}^{\infty} \frac{J_q(x)}{q!} H_q(X),$$

where $X \sim N(0,1)$, the convergence is in L^2 and where

$$(1.3)\qquad J_q(x) = E1\{G(X) \le x\}H_q(X).$$

The coefficient $m = m(x)$ is called the Hermite rank of the function $\Delta_x(\cdot) = 1\{G(\cdot) \le x\} - F(x)$. It is the index of the first nonzero coefficient in the expansion.

From now on, suppose that the exponent D in (1.1) satisfies

$$(1.4)\qquad 0 < D < \frac{1}{m},$$

so that the sequence $1\{G(X_j) \le x\} - F(x)$ also exhibits long-range dependence.

We have [see Theorem 3.1 of Taqqu (1975)]

$$d_N^2 = \operatorname{Var}\left(\sum_{j=1}^{N} H_m(X_j) \right) \approx N^{2-mD} L^m(N), \tag{1.5}$$

where the symbol $\approx$ means asymptotically proportional to. [The constant of proportionality is actually $2m!(1 - mD)^{-1}(2 - mD)^{-1}$.] Using the weak reduction principle of Taqqu (1975), we get

$$d_N^{-1} N(F_N(x) - F(x)) - \frac{J_m(x)}{m!} d_N^{-1} \sum_{j \le N} H_m(X_j) \to 0$$

in L^2. Moreover the results of Taqqu (1975), Dobrushin and Major (1979) and Taqqu (1979) show that

$$\begin{aligned} &\left\{ d_N^{-1}[Nt]\left(F_{[Nt]}(x) - F(x)\right); 0 \le t \le 1 \right\} \\ &\qquad \to \left\{ \frac{J_m(x)}{m!} Z_m(t); 0 \le t \le 1 \right\} \end{aligned} \tag{1.6}$$

in the sense of weak convergence in $D[0,1]$ with the Skorohod topology. The limiting process $Z_m(t)$ is given through a representation involving a multiple Wiener–Itô–Dobrushin integral as

$$\begin{aligned} Z_m(t) = K(m, D) \int_{\mathbf{R}^m}'' &\frac{e^{i(\lambda_1 + \cdots + \lambda_m)t} - 1}{i(\lambda_1 + \cdots + \lambda_m)} \frac{1}{|\lambda_1|^{(1-D)/2}} \\ &\times \cdots \times \frac{1}{|\lambda_m|^{(1-D)/2}} \tilde{B}(d\lambda_1) \cdots \tilde{B}(d\lambda_m), \end{aligned} \tag{1.7}$$

where

$$K(m, D) = \left\{ \frac{(1 - (mD)/2)(1 - mD)}{m!\{2\Gamma(D)\sin([(1 - D)/2]\pi)\}^m} \right\}^{1/2}$$

and where $\tilde{B}$ is a Gaussian complex white noise measure satisfying $\tilde{B}(\Delta) = \bar{\tilde{B}}(-\Delta)$ and $E\tilde{B}(\Delta_1)\bar{\tilde{B}}(\Delta_2) = |\Delta_1 \cap \Delta_2|$ for all Borel sets Δ_1 and Δ_2 of $\mathbf{R}^1$ [see Taqqu (1979), Theorem 6.3]. The symbol $\int''$ means that the domain of integration excludes the hyperdiagonals $\{\lambda_i = \pm\lambda_j, i \ne j\}$. The process $Z_m(t)$ defined in (1.7) is called an mth order Hermite process. The normalization factor $K(m, D)$ ensures that $E(Z_m(1))^2 = 1$.

Recall that up to now we have assumed x to be fixed. We shall now study the joint convergence in both t and x.

Definition. The Hermite rank of the class of functions

$$\{\Delta_x(\cdot), -\infty < x < \infty\} = \{1\{G(\cdot) \le x\} - F(x), -\infty < x < \infty\},$$

is the smallest index $m \ge 1$ so that $J_m(x) \ne 0$ for at least one x, i.e., $m = \inf_x m(x)$.

Our first theorem states that (1.6) also holds in the sense of weak convergence of two-parameter processes.

THEOREM 1.1. *Let $(X_j)_{j=1}^{\infty}$ be a stationary, mean-zero Gaussian process with covariance* (1.1), *let the class of functions* $1\{G(X_j) \le x\} - F(x)$, $-\infty < x < \infty$, *have Hermite rank m and let* $0 < D < 1/m$. *Then*

$$\left\{d_N^{-1}[Nt]\left(F_{[Nt]}(x) - F(x)\right);\ -\infty \le x \le +\infty, 0 \le t \le 1\right\}$$

converges weakly in $D([-\infty, +\infty] \times [0,1])$, *equipped with the sup-norm, to*

$$\left\{\frac{J_m(x)}{m!}Z_m(t);\ -\infty \le x \le +\infty, 0 \le t \le 1\right\}, \tag{1.8}$$

where $Z_m(t)$ *is defined in* (1.7).

REMARK. $D([-\infty, +\infty] \times [0,1])$ is the natural generalization of $D[0,1]$, the space of all functions $f(t)$, $0 \le t \le 1$, that are right-continuous and have left limits. $D([-\infty, +\infty] \times [0,1])$ is the space of functions $f(x, t)$, $-\infty \le x \le +\infty$, $0 \le t \le 1$, that are upper right-continuous and have a limit in any of the other three quadrants. Here, $[-\infty, +\infty]$ involves the two-point compactification of the real line. Since $[-\infty, +\infty] \times [0,1]$ is compact, all functions in $D([-\infty, +\infty] \times [0,1])$ are bounded and hence have finite sup-norm. In fact, that space is isomorphic to $D([0,1]^2)$.

Since the space $D([-\infty, +\infty] \times [0,1])$, equipped with the sup-norm, is not separable, some difficulties arise with the definition of weak convergence of measures on that space. One defines the σ-field $\mathscr{B}$ on $D([-\infty, +\infty] \times [0,1])$ to be the σ-field generated by the open balls and not the open sets and one defines weak convergence of the measures μ_n to μ by requiring $\int f\, d\mu_n$ to converge to $\int f\, d\mu$ for all bounded, continuous $\mathscr{B}$-measurable functions f.

EXAMPLE 1. The function $G(x) = x$ gives rise to a class of functions $1\{X \le x\} - \Phi(x)$, $-\infty < x < \infty$, where Φ is the $N(0,1)$ c.d.f. This class has Hermite rank $m = 1$ because if $X \sim N(0,1)$ and if ϕ denotes the $N(0,1)$ density function,

$$J_1(x) = EH_1(X)\left(1\{X \le x\} - \Phi(x)\right) = EX1\{X \le x\} = -\phi(x)$$

is nonzero. Hence, the limit in the r.h.s. of (1.6) is $-\phi(x)Z_1(t)$, where $Z_1(t)$ is a Gaussian process with mean 0, stationary increments and covariance

$$EZ_1(s)Z_1(t) = \tfrac{1}{2}\left\{|s|^{2H} + |t|^{2H} - |s - t|^{2H}\right\}$$

with $H = 1 - D/2$. The process $Z_1(t)$ is called *fractional Brownian motion*.

EXAMPLE 2. The function $G(x) = x^2$ gives rise to a class of functions $1\{X^2 \le x\} - \{\Phi(\sqrt{x}) - \Phi(-\sqrt{x})\}$, $0 \le x < +\infty$, where $X \sim N(0,1)$. This class has Hermite rank 2, because for $x \ge 0$,

$$J_1(x) = EX\left(1\{-\sqrt{x} \le X \le \sqrt{x}\} + \left\{\Phi(\sqrt{x}) - \Phi(-\sqrt{x})\right\}\right) = 0$$

by symmetry and

$$\begin{aligned} J_2(x) &= E(X^2 - 1)\big(1\{-\sqrt{x} \le X \le \sqrt{x}\} - \{\Phi(\sqrt{x}) - \Phi(-\sqrt{x})\}\big) \\ &= \int_{-\sqrt{x}}^{\sqrt{x}} (s^2 - 1)\phi(s)\, ds \\ &= -s\phi(s)\big|_{-\sqrt{x}}^{+\sqrt{x}} \\ &= -\sqrt{2\pi^{-1}x}\, e^{-x/2}. \end{aligned}$$

Note that for $x \le 0$ both $J_1(x)$ and $J_2(x)$ are equal to 0.

The limiting process is $Z_2(t)$. It is called the Rosenblatt process [see Taqqu (1975)]. $Z_2(t)$ is non-Gaussian, has stationary increments and the same covariance as $Z_1(t)$ but with $H = 1 - D$, $0 < D < \frac{1}{2}$.

EXAMPLE 3. Are there functions G such that the class of functions $1\{G(X) \le x\} - F(x)$, $-\infty < x < \infty$, has Hermite rank $m > 2$? The answer is affirmative. In fact, for every $k < \infty$, there are G's such that $E(1\{G(X) \le x\} - F(x))H_j(x) = 0$ for $j = 0, 1, \dots, k$.

To see this, note first that for every $k \ge 1$ one can find a set $A = A(k)$ such that $\lambda(A) > 0$, $\lambda(\mathbf{R} \setminus A) > 0$ and $\int_A H_j(x)\phi(x)\,dx = 0$ for $j = 1, \dots, k$. (λ denotes the Lebesgue measure.) This result can be obtained as a corollary of Liapounov's theorem on the convexity of the range of a vector measure [see Rudin (1973), Theorem 5.5] in the following way: Define signed measures $\mu_0, \mu_1, \dots, \mu_k$ on $\mathbf{R}$ by $d\mu_j/dx = H_j(x)\phi(x)$ and let $\boldsymbol{\mu} = (\mu_0, \mu_1, \dots, \mu_k)^t$ be the corresponding vector measure. We have $\boldsymbol{\mu}(\varnothing) = (0, 0, \dots, 0)^t$ and $\boldsymbol{\mu}(\mathbf{R}) = (1, 0, \dots, 0)^t$ and, thus, by Liapounov's theorem, for every a in $[0, 1]$, there is a set A with vector measure $\boldsymbol{\mu}(A) = (a, 0, \dots, 0)^t$. By choosing an a different from 0 or 1, we obtain a nontrivial set A satisfying $\int_A H_j(x)\phi(x)\,dx = 0$ for $j = 1, \dots, k$. This set A has the claimed properties.

Now let $G = 1_A$. The level sets $1\{s\colon G(s) \le x\}$, $-\infty < x < \infty$, are $\varnothing$, A and $\mathbf{R}$. Since $E(1\{G(X) \le x\} - F(x))H_j(X) = 0$ for $j = 0, 1, \dots, k$, the Hermite rank of the class of function $1\{G(X) \le x\} - F(x)$, $\infty < x < \infty$, is $m > k$. Since $\lambda(A) > 0$, the coefficients in the Hermite expansion cannot all be zero, so that $m < \infty$.

Note that when $G = 1_A$, $F_N(x)$ takes only three values, 0, $(1/N)\sum_{i=1}^{[Nt]} 1_A(X_i)$ and 1. The limit result for $d_N^{-1}[Nt](F_{[Nt]} - F)$ does not require the full force of Theorem 1.1, but can be immediately derived from the results of Dobrushin and Major (1979).

Moreover, the preceding argument based on Liapounov's theorem does not provide a construction of G. A more explicit construction is provided in Dehling and Taqqu (1988b), where a larger class of examples can be found, including continuous functions G.

Structure of the paper. The rest of the paper is arranged as follows. Section 2 contains applications to von Mises and U-statistics. In Section 3 we show that, as $N \to \infty$,

$$\sup_{-\infty \le x \le \infty} \sup_{0 \le t \le 1} d_N^{-1} \left| [Nt]\left(F_{[Nt]}(x) - F(x)\right) - \frac{J_m(x)}{m!} \sum_{j \le [Nt]} X_j \right| \to 0$$

in probability (weak uniform reduction principle) and we use this result in Section 4 to establish Theorem 1.1.

2. Applications to *U*-statistics and von Mises statistics. Theorem 1.1 can be applied to obtain the limit distribution of degenerate U-statistics and von Mises statistics. Let h: $\mathbf{R}^k \to \mathbf{R}$ be a measurable function, invariant under permutation of its arguments, and let $Y_j = G(X_j)$, $j = 1, 2, \dots,$ be as in Section 1. The nonnormalized U-statistic $U_N(h)$ and von Mises statistic $V_N(h)$ with kernel h are defined by

$$U_N(h) = \sum_{\substack{1 \le j_1, \dots, j_k \le N \\ j_\mu \ne j_\nu \text{ for } \mu \ne \nu}} h(Y_{j_1}, \dots, Y_{j_k}), \tag{2.1a}$$

$$V_N(h) = \sum_{1 \le j_1, \dots, j_k \le N} h(Y_{j_1}, \dots, Y_{j_k}). \tag{2.1b}$$

Recall that $F(x) = P(Y_1 \le x) = P(G(X_1) \le x)$. Then h is called *degenerate* if

$$\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} |h(x_1, \dots, x_k)| \, dF(x_1) \cdots dF(x_k) < \infty \tag{2.2a}$$

and

$$\int_{-\infty}^{+\infty} h(x_1, x_2, \dots, x_k) F(dx_1) = 0, \qquad \forall\, x_2, \dots, x_k. \tag{2.2b}$$

$V_N(h)$ can then be written as a stochastic integral of $h(x_1, \dots, x_k)$ with integrator $F_N(x)$,

$$\begin{aligned} V_N(h) &= N^k \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} h(x_1, \dots, x_k) F_N(dx_1) \cdots F_N(dx_k) \\ &= N^k \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} h(x_1, \dots, x_k)(F_N(dx_1) - F(dx_1)) \\ &\qquad \cdots (F_N(dx_k) - F(dx_k)). \end{aligned} \tag{2.3}$$

The last step follows from the degeneracy of h. We would of course like to apply the weak convergence of $d_N^{-1}[Nt](F_{[Nt]}(x) - F(x))$, established in Theorem 1.1. Unfortunately, in general, the map $Q \to \int \cdots \int h(x_1, \dots, x_k) Q(dx_1) \cdots Q(dx_k)$ that maps elements Q of $D[-\infty, +\infty]$ into $\mathbf{R}$ is not continuous with respect to the sup-norm on $D[-\infty, +\infty]$. Actually this map is not even defined on all of $D[-\infty, +\infty]$. However, as long as h has bounded total variation, continuity holds, as can be seen through integration by parts.

COROLLARY 1. *Let h have bounded total variation and satisfy relations* (2.2). *Assume that $h(x_1, \dots, x_k)$ and $F(x_1) \cdots F(x_k)$ have no joint discontinuities.*

Then the $D[0,1]$-valued random elements

$$\left(d_N^{-k} V_{[Nt]}(h), 0 \le t \le 1\right)$$

converge weakly to

$$\left(\left(\frac{1}{m!}\right)^k \left(\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} h(x_1, \dots, x_k)\, dJ_m(x_1) \cdots dJ_m(x_k)\right)(Z_m(t))^k, \quad 0 \le t \le 1\right). \tag{2.4}$$

Here $Z_m(t)$ is the mth order Hermite process, as defined in (1.7).

REMARK 1. Relation (2.2a) is always satisfied because h is bounded. Moreover, the integral in (2.4) is well-defined because $\|J_m\|_{\mathrm{TV}} \le \int_{-\infty}^{\infty} |H_m(s)|\phi(s)\, ds < \infty$, where $\|\ \|_{\mathrm{TV}}$ denotes the total variation norm and ϕ is the $N(0,1)$ density function.

REMARK 2. The integral in (2.4),

$$\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} h(x_1, \dots, x_k)\, dJ_m(x_1) \cdots dJ_m(x_k)$$

can be rewritten as

$$\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} h(G(u_1), \dots, G(u_k)) H_m(u_1) \cdots H_m(u_k) \times \phi(u_1) \cdots \phi(u_k) du_1 \cdots du_k.$$

To verify this, it suffices to show that for all bounded, measurable functions $f(x)$, the following holds:

$$\int_{-\infty}^{+\infty} f(x)\, dJ_m(x) = \int_{-\infty}^{+\infty} f(G(u)) H_m(u) \phi(u)\, du. \tag{2.5}$$

Assume first that $f(x) = 1\{x \le a\}$ for some $a \in \mathbf{R}$. Then the l.h.s. of (2.5) equals $J_m(a)$ and the r.h.s. equals $\int 1\{G(u) \le a\} H_m(u)\phi(u)\, du = J_m(a)$ by definition. For step functions f, the equation (2.5) follows by linearity and for general bounded measurable f, it follows via approximation by step functions.

For the proof of Corollary 1 we need an integration by parts formula for higher-dimensional integrals. It can be found in Young (1916). For simplicity of notation, we let $k = 2$.

Let H, K: $\mathbf{R}^2 \to \mathbf{R}$ be right-continuous functions with no common discontinuities. Then

$$\begin{aligned}\int_{a_1}^{b_1} \int_{a_2}^{b_2} H(x_1, x_2)\, dK(x_1, x_2) &= \int_{a_1}^{b_1} \int_{a_2}^{b_2} K(x_1, x_2)\, dH(x_1, x_2) - \int_{a_1}^{b_1} [K(x_1, x_2)\, dH(x_1, x_2)]_{x_2 = a_2}^{x_2 = b_2} \\ &\quad - \int_{a_2}^{b_2} [K(x_1, x_2)\, dH(x_1, x_2)]_{x_1 = a_1}^{x_1 = b_1} + [HK]_{x_1 = a_1,\, x_2 = a_2}^{x_1 = b_1,\, x_2 = b_2},\end{aligned}$$

where the [ ] symbols are defined by

$$[f(x_1, x_2)]_{x_1=a}^{x_1=b} = f(b, x_2) - f(a, x_2),$$

$$[F(x_1, x_2)]_{x_1=a_1,\, x_2=a_2}^{x_1=b_1,\, x_2=b_2} = \Big[[f(x_1, x_2)]_{x_1=a_1}^{x_1=b_1}\Big]_{x_2=a_2}^{x_2=b_2}$$
$$= f(b_1, b_2) - f(a_1, b_2) - f(b_1, a_2) + f(a_1, a_2).$$

PROOF OF COROLLARY 1. Again, for simplicity of notation, we let $k = 2$. We will apply the integration by parts formula to $H(x, y) = h(x, y)$ and $K(x, y) = (F_N - F)(x)(F_N - F)(y)$. The set of discontinuities of K lies on a countable union of lines parallel to one of the axes. Since the probability of a pair of observations (Y_j, Y_k) falling on a discontinuity point of H is 0, the functions H and K have almost surely no joint discontinuities.

Since h has bounded TV, we may let $a_1, a_2 \to -\infty$ and $b_1, b_2 \to +\infty$ in the integration by parts formula. Both $h(x, y)(F_n(x) - F(x))(F_n(y) - F(y))$ and $h(x, y)J_m(x)J_m(y)$ vanish at infinity so that

$$V_{[Nt]}(h) = N^2 \int\int_{\mathbf{R}^2} h(x, y) d(F_{[Nt]} - F)(x) d(F_{[Nt]} - F)(y)$$
$$= N^2 \int\int_{\mathbf{R}^2} (F_{[Nt]} - F)(x)(F_{[Nt]} - F)(y)\, dh(x, y).$$

Now the map Λ: $D([-\infty, +\infty] \times [0,1]) \to D[0,1]$ defined by

$$\Lambda(Q) = \int\int_{\mathbf{R}^2} Q(x, t) Q(y, t)\, dh(x, y)$$

is continuous with respect to the uniform topologies on both spaces. Hence we can apply Theorem 1 and the continuous mapping theorem to get

$$d_N^{-2} V_{[Nt]}(h) = \int\int_{\mathbf{R}^2} d_N^{-1}[Nt](F_{[Nt]} - F)(x) d_N^{-1}[Nt](F_{[Nt]} - F)(y)\, dh(x, y)$$
$$\to_w \left(\frac{1}{m!}\right)^2 \int\int_{\mathbf{R}^2} J_m(x) J_m(y)\, dh(x, y) \cdot (Z_m(t))^2$$
$$= \frac{1}{(m!)^2} \int\int_{\mathbf{R}^2} h(x, y)\, dJ_m(x)\, dJ_m(y) \cdot (Z_m(t))^2. \qquad \square$$

Let us now evaluate the limit behavior of the U-statistics $U_N(h)$, defined in (2.1a). The difference between $U_N(h)$ and $V_N(h)$ is given by

$$V_N(h) - U_N(h) = \sum_{\substack{1 \le j_1, \dots, j_k \le N \\ j_\mu = j_\nu \text{ for some } \mu \ne \nu}} h(Y_{j_1}, \dots, Y_{j_k}). \qquad (2.6)$$

One may think of this as being the sum over all the hyperdiagonal terms. We will show that $V_N(h) - U_N(h) = o_p(d_N^k)$ so that $d_N^{-k} V_{[Nt]}$ and $d_N^{-k} U_{[Nt]}$ have the same limit. This is due to the fact that $N = o(d_N^2)$ and hence the situation here is quite different from that encountered in the study of U-statistics of i.i.d. observations.

COROLLARY 2. *Under the same assumptions as in Corollary* 1 *the* $D[0,1]$-*valued random elements*

$$\left(d_N^{-k}U_{[Nt]}(h);\, 0 \le t \le 1\right)$$

converge in distribution to

$$\left(\left(\frac{1}{m!}\right)^k \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} h(x_1,\dots,x_k)\, dJ_m(x_1) \cdots dJ_m(x_k)\big(Z_m(t)\big)^k;\, 0 \le t \le 1\right).$$

Here $Z_m(t)$ *is the mth order Hermite process defined in* (1.7).

PROOF. It suffices to prove that $d_N^{-k} \max_{n=1,\dots,N}(V_n(h) - U_n(h)) \to 0$ in probability. The r.h.s. of (2.6) can be decomposed into a finite-sum of (nondegenerate) von Mises statistics of order lower than k, namely, according to which indices coincide. Then one can apply the Hoeffding decomposition to each of these statistics to finally represent $V_N - U_N$ as a sum of degenerate von Mises statistics, all of order lower than k. To these, Corollary 1 may be applied. Rather than going into the combinatorial details of the general case, let us consider a special, but representative, situation. Take $k = 4$ and let $j_1 = j_2$, i.e.,

$$\sum_{j_1, j_3, j_4} h\left(X_{j_1}, X_{j_1}, X_{j_3}, X_{j_4}\right).$$

Define $g(x_2,x_3) = \int h(x_1,x_1,x_2,x_3)\, dF(x_1)$ and $\bar{h}(x_1,x_2,x_3) = h(x_1,x_1,x_2,x_3) - g(x_2,x_3)$. Observe that both g and $\bar{h}$ are degenerate kernels and both have bounded total variation. We have then

$$\begin{aligned} d_N^{-4} &\max_{n=1,\dots,N} \sum_{j_1, j_3, j_4 \le n} h\left(X_{j_1}, X_{j_1}, X_{j_3}, X_{j_4}\right) \\ &= d_N^{-4} \max_{n=1,\dots,N}\left[\sum_{j_1, j_3, j_4 \le n} \bar{h}\left(X_{j_1}, X_{j_3}, X_{j_4}\right) + N \sum_{j_3, j_4 \le n} g\left(X_{j_3}, X_{j_4}\right)\right] \\ &\le d_N^{-4} \max_{n=1,\dots,N} \sum_{j_1, j_3, j_4 \le n} \bar{h}\left(X_{j_1}, X_{j_3}, X_{j_4}\right) + d_N^{-4} \max_{n=1,\dots,N} \sum_{j_3, j_4 \le n} g\left(X_{j_3}, X_{j_4}\right) \end{aligned}$$

and by Corollary 1 both terms on the r.h.s. converge to 0.

For general k, the worst situation involves k even and

$$d_N^{-k} \sum_{j_1,\dots,j_{k/2}} h\left(X_{j_1}, X_{j_1}, X_{j_2}, X_{j_2}, \dots, X_{j_{k/2}}, X_{j_{k/2}}\right).$$

The dominating term is the lowest order one in the Hoeffding decomposition and in this case it is Eh. However, as $N \to \infty$,

$$d_N^{-k} \max_{n=1,\dots,N} \sum_{1 \le i_1,\dots,i_{k/2} \le n} Eh \le d_N^{-k} N^{k/2} Eh \to 0. \qquad \square$$

Corollaries 1 and 2 assume that h has bounded total variation. We can relax this assumption in the case $k = 2$ by requiring that h has *locally* bounded variation and satisfies some additional condition. The details will be given in a subsequent paper.

We derived Corollaries 1 and 2 from Theorem 1 via a continuous mapping theorem. A referee remarked that this method may not be optimal since the integral in (2.4) can be 0, indicating that the normalizing factor d_N^k is not necessarily always the correct one.

3. The uniform weak reduction principle. In order to study the weak limit behavior of $F_N(x)$, properly normalized and centered, we have to look at the Hermite decomposition of $1\{Y_j \le x\} - F(x)$,

$$1\{Y_j \le x\} - F(x) = \sum_{q=m}^{\infty} \frac{J_q(x)}{q!} H_q(X_j). \tag{3.1}$$

Here m is the Hermite rank of the class of functions $1\{Y_j \le x\} - F(x)$, $-\infty \le x \le +\infty$. Then by Taqqu's (1975) weak reduction principle we know that for any fixed x the limiting distribution of

$$d_N^{-1} \sum_{j \le [Nt]} \big(1\{Y_j \le x\} - F(x)\big)$$

is the same as that of

$$d_N^{-1} \sum_{j \le [Nt]} \frac{J_m(x)}{m!} H_m(X_j).$$

It is our goal to extend this to a reduction principle uniformly in both $-\infty \le x \le +\infty$ and $0 \le t \le 1$.

THEOREM 3.1 (Weak uniform reduction principle). *There are absolute constants,* $C, \kappa > 0$ *such that for any* $0 < \varepsilon \le 1$,

$$P\left(\max_{n \le N} \sup_{-\infty \le x \le +\infty} d_N^{-1} \left| \sum_{j=1}^{n} \left[1\{Y_j \le x\} - F(x) - \frac{J_m(x)}{m!} H_m(X_j) \right] \right| > \varepsilon \right) \tag{3.2}$$
$$\le C N^{-\kappa}(1 + \varepsilon^{-3}).$$

For the proof, we have to introduce some notation:

$$F(x, y) = F(y) - F(x),$$

$$S_N(n; x) = d_N^{-1} \sum_{j \le n} \left[1\{Y_j \le x\} - F(x) - \frac{J_m(x)}{m!} H_m(X_j) \right], \tag{3.3}$$

$$S_N(n; x, y) = S_N(n; y) - S_N(n; x), \qquad x \le y.$$

Also recall the definition of $J_q(x)$:

$$J_q(x) = E1\{G(X_j) \le x\} H_q(X_j) = \int_{\{G(s) \le x\}} H_q(s)\phi(s)\, ds, \tag{3.4}$$

where $\phi(s) = (1/\sqrt{2\pi})e^{-s^2/2}$ is the standard normal density. Denote for $x \le y$,

$$J_q(x, y) = J_q(y) - J_q(x)$$
$$= \int_{\{x < G(s) \le y\}} H_q(s)\phi(s)\, ds.$$

Lemma 3.1. *There exist constants $\gamma > 0$ and C such that for $n \le N$,*

$$E|S_N(n; x, y)|^2 \le C\left(\frac{n}{N}\right) \cdot N^{-\gamma}(F(y) - F(x)).$$

Proof. Since $J_q(x) = 0$ for $q = 1, \dots, m-1$ and for all x, we have

$$1\{x < Y_j \le y\} - F(x, y) = \sum_{q=m}^{\infty} \frac{J_q(x, y)}{q!} H_q(X_j).$$

Observe that by orthogonality of the $H_q(X_j)$,

$$\begin{aligned}\sum_{q=m}^{\infty} \frac{J_q^2(x, y)}{q!} &= E\left(1\{x < Y_j \le y\} - F(x, y)\right)^2 \\ &= F(x, y)(1 - F(x, y)) \\ &\le F(x, y).\end{aligned} \tag{3.5}$$

Thus

$$\begin{aligned}&E\left(\sum_{j \le n} 1\{x < Y_j \le y\} - F(x, y) - \frac{J_m(x, y)}{m!} H_m(X_j)\right)^2 \\ &\quad = \sum_{q=m+1}^{\infty} \frac{J_q^2(x, y)}{q!} \cdot \frac{1}{q!} \sum_{j, k \le n} EH_q(X_j)H_q(X_k) \\ &\quad \le F(x, y) \cdot \sum_{j, k \le n} |r(j - k)|^{(m+1)},\end{aligned}$$

since $EH_q(X_j)H_q(X_k) = q!r^q(j - k)$. We now have to distinguish three cases, namely,

$$(m+1)D < 1: \qquad \sum_{j, k \le n} |r(j - k)|^{(m+1)} \approx n^{2-(m+1)D}(L(n))^{m+1},$$

$$(m+1)D > 1: \qquad \sum_{j, k \le n} |r(j - k)|^{(m+1)} \approx n,$$

$$(m+1)D = 1: \qquad \sum_{j, k \le n} |r(j - k)|^{(m+1)} \approx nL_0(n),$$

where L_0 is slowly varying. In general, we get

$$\sum_{j, k \le n} (r(j - k))^{m+1} \approx n^{1 \vee (2-(m+1)D)} L'(n),$$

where L' is slowly varying. Keeping in mind that $d_N^{-2} \approx N^{mD-2}(L(N))^{-m}$,

$$\begin{aligned}E|S_N(n; x, y)|^2 &\le CF(x, y) \cdot n^{1 \vee (2-(m+1)D)} N^{mD-2} L'(n)(L(N))^{-m} \\ &= CF(x, y)\left(\frac{n}{N}\right)^{1 \vee (2-(m+1)D)} N^{mD-1 \vee (-D)} L'(n)(L(N))^{-m}.\end{aligned}$$

The conclusion follows because for each $\lambda > 0$, $\max_{n \le N} L'(n)(L(N))^{-m} \le C(\lambda)N^{\lambda}$ for some constant $C(\lambda)$, since $\lim_{n\to\infty} L'(n)n^{-\lambda/2} = 0$ and $\lim_{N\to\infty}(L(N))^{-m}N^{-\lambda/2} = 0$ [see Bingham, Goldie and Teugels (1987), Proposition 1.3.6(v)]. □

LEMMA 3.2. *There exist constants $\rho > 0$ and C such that or $n \le N$ and* $0 < \varepsilon \le 1$,

$$P\Big\{\sup_x |S_N(n; x)| > \varepsilon\Big\} \le CN^{-\rho}\left\{\left(\frac{n}{N}\right)\varepsilon^{-3} + \left(\frac{n}{N}\right)^{2-mD}\right\}. \tag{3.6}$$

PROOF. Let

$$\Lambda(x) = F(x) + \int_{\{G(s)\le x\}} \frac{|H_m(s)|}{m!}\phi(s)\,ds$$

so that $F(x, y)$ and $(1/m!)J_m(x, y)$ are both bounded by $\Lambda(y) - \Lambda(x)$. The function Λ is monotone, $\Lambda(-\infty) = 0$, $\Lambda(+\infty) < \infty$. In what follows it is important to keep in mind that Λ could have discontinuities. We define refining partitions of $\mathbf{R}$,

$$-\infty = x_0(k) \le x_1(k) \le \cdots \le x_{2^k}(k) = +\infty, \qquad k = 0, 1, \dots, K,$$

by $x_i(k) = \inf\{x\colon \Lambda(x) \ge \Lambda(+\infty)i2^{-i}\}$, $i = 0, \dots, 2^k - 1$. The integer K will be chosen below (as a suitable function of N and ε).

If Λ has discontinuities larger than $\Lambda(+\infty)2^{-k}$, then a number of subdivision points $x_i(k)$ will be the same. But in all cases

$$\Lambda(x_i(k)-) - \Lambda(x_{i-1}(k)) \le \Lambda(+\infty)\cdot 2^{-k}.$$

For $k = 0$ we obtain the trivial partition

$$-\infty = x_0(0) < x_1(0) = +\infty.$$

For each x and for each $k = 0, 1, \dots, K$ define $i_k(x)$ by

$$x_{i_k(x)}(k) \le x < x_{i_k(x)+1}(k).$$

One can then define a chain, linking each point x to $-\infty$,

$$\begin{aligned}-\infty = x_{i_0(x)}(0) &\le x_{i_1(x)}(1) \le \cdots \le x_{i_{K-1}(x)}(K-1)\\ &\le x_{i_K(x)}(K) \le x < x_{i_K(x)+1}(K).\end{aligned}$$

Then

$$\begin{aligned}S_N(n; x) = {} & S_N\big(n; x_{i_0(x)}(0), x_{i_1(x)}(1)\big)\\ & + S_N\big(n; x_{i_1(x)}(1), x_{i_2(x)}(2)\big) + \cdots\\ & + S_N\big(n; x_{i_{K-1}(x)}(K-1), x_{i_K(x)}(K)\big) + S_N\big(n; x_{i_K(x)}(K), x\big).\end{aligned} \tag{3.7}$$

Let us first bound the last term in (3.7). Setting $\Lambda(x, y) = \Lambda(y) - \Lambda(x)$, we get

$$\begin{aligned}
\left|S_N\left(n; x_{i_K(x)}(K), x\right)\right| &= \left|d_N^{-1} \sum_{j \le n} \left[1\left\{x_{i_K(x)}(K) < Y_j \le x\right\} - F\left(x_{i_K(x)}(K), x\right)\right.\right. \\
&\qquad \left.\left. - \frac{1}{m!} J_m\left(x_{i_K(x)}(K), x\right) H_m(X_j)\right]\right| \\
&\le d_N^{-1} \sum_{j \le n} \left[1\left\{x_{i_K(x)}(K) < Y_j \le x\right\} + F\left(x_{i_K(x)}(K), x\right)\right] \\
&\quad + \left|\frac{1}{m!} J_m\left(x_{i_K(x)}(K), x\right)\right| d_N^{-1} \left|\sum_{j \le n} H_m(X_j)\right| \\
&\le d_N^{-1} \sum_{j \le n} \left[1\left\{x_{i_K(x)}(K) < Y_j < x_{i_K(x)+1}(K)\right\}\right. \\
&\qquad \left. + F\left(x_{i_K(x)}(K), x_{i_K(x)+1}(K) -\right)\right] \\
&\quad + \Lambda\left(x_{i_K(x)}(K), x_{i_K(x)+1}(K) -\right) d_N^{-1} \left|\sum_{j \le n} H_m(X_j)\right| \\
&\le \left|S_N\left(n; x_{i_K(x)}(K), x_{i_K(x)+1}(K) -\right)\right| \\
&\quad + 2n d_N^{-1} F\left(x_{i_K(x)}(K), x_{i_K(x)+1}(K) -\right) \\
&\quad + 2\Lambda\left(x_{i_K(x)}(K), x_{i_K(x)+1}(K) -\right) d_N^{-1} \left|\sum_{j \le n} H_m(X_j)\right| \\
&\le \left|S_N\left(n; x_{i_K(x)}(K), x_{i_K(x)+1}(K) -\right)\right| \\
&\quad + 2\Lambda(+\infty) n d_N^{-1} 2^{-K} + 2\Lambda(+\infty) 2^{-K} d_N^{-1} \left|\sum_{j \le n} H_m(X_j)\right|,
\end{aligned}$$

since $(1/m!)J_m(x, y) \le \Lambda(y) - \Lambda(x)$. Moreover, since $\sum_{k=0}^{\infty} \varepsilon/(k+3)^2 \le \varepsilon/2$, we get

$$\begin{aligned}
P\Big\{\sup_x |S_N(n; x)| > \varepsilon\Big\}& \\
\le P\Big\{\max_x \left|S_N\left(n; x_{i_0(x)}(0), x_{i_1(x)}(1)\right)\right| > \varepsilon/9\Big\}& \\
+ P\Big\{\max_x \left|S_N\left(n; x_{i_1(x)}(1), x_{i_2(x)}(2)\right)\right| > \varepsilon/16\Big\} + \cdots& \\
+ P\Big\{\max_x \left|S_N\left(n; x_{i_{K-1}(x)}(K-1), x_{i_K(x)}(K)\right)\right| > \varepsilon/(K+2)^2\Big\}& \qquad (3.8)\\
+ P\Big\{\max_x \left|S_N\left(n; x_{i_K(x)}(K), x_{i_K(x)+1}(K) -\right)\right| > \varepsilon/(K+3)^2\Big\}& \\
+ P\Big\{2\Lambda(+\infty)2^{-K} d_N^{-1} \Big|\sum_{j \le n} H_m(X_j)\Big| > (\varepsilon/2) - 2\Lambda(+\infty) n d_N^{-1} 2^{-K}\Big\}&.
\end{aligned}$$

Using Chebyshev's inequality and Lemma 3.1 we obtain

$$P\left\{\max_x \left|S_N\left(n; x_{i_k(x)}(k), x_{i_{k+1}(x)}(k+1)\right)\right| > \frac{\varepsilon}{(k+3)^2}\right\}$$

$$\geq \sum_{i=0}^{2^{k+1}-1} P\left\{S_N\left(n; x_i(k+1), x_{i+1}(k+1)\right) > \frac{\varepsilon}{(k+3)^2}\right\}$$

$$\leq C \sum_{i=0}^{2^{k+1}-1} \left(\frac{n}{N}\right) N^{-\gamma} \frac{(k+3)^4}{\varepsilon^2} \left(F(x_{i+1}(k+1)) - F(x_i(k+1))\right)$$

$$\leq C\left(\frac{n}{N}\right) N^{-\gamma}(k+3)^4 \varepsilon^{-2}.$$

In the same way we get

$$P\left\{\max_x \left|S_N\left(n; x_{i_K(x)}(K), x_{i_K(x)+1}(K)\,-\right)\right| > \frac{\varepsilon}{(K+3)^2}\right\}$$

$$\leq C\left(\frac{n}{N}\right) N^{-\gamma}(K+3)^4 \varepsilon^{-2}.$$

Choose now

$$K = \left[\log_2\left(\frac{8\Lambda(+\infty)}{\varepsilon} N d_N^{-1}\right)\right] + 1 \tag{3.9}$$

so that $2\Lambda(+\infty) \cdot N \cdot d_N^{-1} 2^{-K} \leq \varepsilon/4$. The last probability in (3.8) can then be bounded by

$$P\left\{d_N^{-1}\left|\sum_{j\leq n} H_m(X_j)\right| > \frac{\varepsilon}{4} \frac{2^{K-1}}{\Lambda(+\infty)}\right\} \leq C\left(\frac{d_n}{d_N}\right)^2 2^{-2K+2}\left(\frac{\varepsilon}{4}\right)^{-2} (\Lambda(+\infty))^2$$

$$\leq C\left(\frac{n}{N}\right)^{2-mD} N^{-2} d_N^2 \left(\frac{L(n)}{L(N)}\right)^m$$

$$\leq C\left(\frac{n}{N}\right)^{2-mD} N^{-mD+\lambda} \quad \text{for any } \lambda > 0.$$

Note that C is a universal constant, possibly changing from line to line. Hence we find

$$P\left\{\sup_x |S_N(n; x)| > \varepsilon\right\} \leq C\left(\frac{n}{N}\right) N^{-\gamma} \varepsilon^{-2} \sum_{k=0}^{K} (k+3)^4 + C\left(\frac{n}{N}\right)^{2-mD} N^{-mD+\lambda}$$

$$\leq C\left(\frac{n}{N}\right) N^{-\gamma} \varepsilon^{-2} (K+3)^5 + C\left(\frac{n}{N}\right)^{2-mD} N^{-mD+\lambda}.$$

By (1.5), $Nd_N^{-1} \approx N^{-mD/2}L^{-m/2}(N)$, so that the definition (3.9) of K yields

$$\begin{aligned}(K+3)^5 &\le C\left\{\left|\log\frac{1}{\varepsilon}\right|^5 + |\log N|^5\right\}\\ &\le C\varepsilon^{-1}\cdot N^{\delta} \quad \text{for any } \delta > 0, \text{ where } C = C(\delta).\end{aligned}$$

Now choose, for example, $\delta = \gamma/2$ and set $\rho = \min(\gamma - \delta, mD - \lambda) = \min(\gamma/2, mD - \lambda)$. Then

$$P\Big\{\sup_x |S_N(n;x)| > \epsilon\Big\} \le CN^{-\rho}\left\{\left(\frac{n}{N}\right)\epsilon^{-3} + \left(\frac{n}{N}\right)^{2-mD}\right\}. \qquad \square$$

Proof of Theorem 3.1. It obviously suffices to prove (3.2) with $N = 2^r$ for some integer r, since d_N/d_{2^r} is uniformly bounded away from 0 and ∞. For abbreviation, define

$$M_N(n) = \sup_x |S_N(n;x)|,$$
$$M_N(n_1, n_2) = M_N(n_2) - M_N(n_1).$$

Observe that $S_N(n;x) = \Phi(X_1,\dots,X_N)$ with $\Phi\colon \mathbf{R}^n \to D[-\infty,+\infty]$ measurable and that $S_N(n_2;x) - S_N(n_1;x) = \Phi(X_{n_1+1},\dots,X_{n_2})$. Stationarity of the underlying sequence (X_j) ensures that $S_N(n;x)$ has stationary increments in n, i.e.,

$$\{S_N(n_2;x) - S_N(n_1;x);\ x \in \mathbf{R}\} =_d \{S_N(n_2 - n_1;x);\ x \in \mathbf{R}\}.$$

Thus we get

$$\begin{aligned}M_N(n_1,n_2) &\le \sup_x |S_N(n_2,x) - S_N(n_1,x)|\\ &=_d \sup_x |S_N(n_2 - n_1;x)|\\ &= M_N(n_2 - n_1).\end{aligned}$$

By Lemma 3.2, we get for $k \le r$,

$$P\{M_N(2^k) > \varepsilon\} \le CN^{-\rho}\{2^{k-r}\varepsilon^{-3} + 2^{(k-r)(2-mD)}\}$$

and thus

$$\begin{aligned}P\Big\{\max_{j=1,\dots,2^{r-k}} |M_N((j-1)2^k, j2^k)| > \varepsilon\Big\} &\le \sum_{j=1}^{2^{r-k}} P\{|M_N(2^k)| > \varepsilon\}\\ &\le CN^{-\rho}\{\varepsilon^{-3} + 2^{(k-r)(1-mD)}\}.\end{aligned}$$

Let now $1 \le n \le N = 2^r$ be given. We can express $M_N(n)$ as a sum of increments over intervals of decreasing length, according to the dyadic expansion of n. More precisely, if n has dyadic expansion $n = \sum_{k=0}^r \sigma_k 2^{r-k}$, $\sigma_k \in \{0,1\}$, then

$$M_N(n) = \sum_{k=0}^{r} \sigma_k M_N((j_k - 1)2^{r-k}, j_k 2^{r-k}),$$

where $j_0 = 1$ and for $k = 1, \dots, r$,

$$(j_k - 1)2^{r-k} = \sigma_0 2^r + \sigma_1 2^{r-1} + \cdots + \sigma_{k-1} 2^{r-(k-1)}.$$

Hence $j_k \in \{1, \dots, 2^k\}$ and therefore

$$\max_{n \le N} |M_N(n)| \le \sum_{k=0}^{r} \max_{j=1,\dots,2^{r-k}} \left|\left(M_N(j-1)2^k, j2^k\right)\right|.$$

Thus,

$$\begin{aligned} P\left\{\max_{n \le N} |M_N(n)| > \varepsilon\right\} &\le \sum_{k=0}^{r} P\left\{\max_{j=1,\dots,2^{r-k}} \left|M_N\left((j-1)2^k, j2^k\right)\right| > \frac{\varepsilon}{(k+2)^2}\right\} \\ &\le CN^{-\rho}\left\{\sum_{k=0}^{\log_2 N} (k+2)^6 \varepsilon^{-3} + \sum_{k=0}^{\log_2 N} 2^{(k-r)(1-mD)}\right\} \\ &\le CN^{-\kappa}\{\varepsilon^{-3} + 1\} \end{aligned}$$

for some $0 < \kappa < \rho$. □

4. Proof of Theorem 1.1. From the results of Taqqu (1975, 1979) we know that

$$d_N^{-1} \sum_{j \le [Nt]} H_m(X_j) \to Z_m(t) \quad \text{(in distribution)}$$

in $D[0,1]$. Note that we equip $D[0,1]$ with the sup-norm topology and with the σ-field generated by the open balls. Since $Z_m(\cdot) \in C[0,1]$ a.s., we may apply the a.s. representation theorem of Skorohod and Dudley [Pollard (1984), page 71] to find a version $\tilde{Z}_{m,N}(t)$ of $d_N^{-1}\Sigma_{j \le [Nt]} H_m(X_j)$ and a version $\tilde{Z}_m(\cdot)$ of $Z_m(\cdot)$ such that

$$\left\|\tilde{Z}_{m,N}(\cdot) - \tilde{Z}_m(\cdot)\right\|_{D[0,1]} \to 0 \quad \text{a.s.}$$

This now implies

$$\left\| J_m(\cdot)\tilde{Z}_{m,N}(\cdot) - J_m(\cdot)\tilde{Z}_m(\cdot)\right\|_{D([-\infty,+\infty]\times[0,1])} \to 0 \quad \text{a.s.,}$$

which proves weak convergence in $D([-\infty, +\infty] \times [0,1])$ of

$$J_m(x) d_N^{-1} \sum_{j \le [Nt]} H_m(X_i)$$

to

$$J_m(x) Z_m(t).$$

Since we have shown in Theorem 3.1 that

$$\left\| d_N^{-1} \sum_{j \le [Nt]} \left(1\{Y_j \le x\} - F(x)\right) - J_m(x) d_N^{-1} \sum_{j \le [Nt]} H_m(X_j)\right\| \to 0$$

in probability, we get finally

$$d_N^{-1} \sum_{j \le [Nt]} \left(1\{Y_j \le x\} - F(x)\right) \to J_m(x) Z_m(t) \quad \text{in } D([-\infty, +\infty] \times [0,1]). \quad \square$$

Acknowledgments. This research was done while the first author was visiting Boston University. He thanks the Department of Mathematics for its hospitality. The authors would like to thank M. Keane and R. Vitale for making them aware of Liapounov's theorem and also an Associate Editor for his careful reading and his sense of humor.

REFERENCES

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

BINGHAM, N. H., GOLDIE, C. M. and TEUGELS, J. L. (1987). *Regular Variation*. Cambridge Univ. Press, Cambridge.

DEHLING, H. and TAQQU, M. S. (1987). The limit behavior of empirical processes and symmetric statistics. *Bull. Internat. Statist. Inst.* **52** (4) 217–234.

DEHLING, H. and TAQQU, M. S. (1988a). The functional law of the iterated logarithm for the empirical process of some long-range dependent sequences. *Statist. Probab. Lett.* **7** 81–85.

DEHLING, H. and TAQQU, M. S. (1988b). Continuous functions whose level sets are orthogonal to all polynomials of a given degree. *Acta Math. Hungar.* To appear.

DENKER, M., GRILLENBERGER, C. and KELLER, G. (1985). A note on invariance principles for von Mises statistics. *Metrika* **32** 197–214.

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

MANDELBAUM, A. and TAQQU, M. S. (1984). Invariance principle for symmetric statistics. *Ann. Statist.* **12** 483–496.

MÜLLER, D. W. (1970). On Glivenko–Cantelli convergence. *Z. Wahrsch. verw. Gebiete* **16** 195–210.

POLLARD, D. (1984). *Convergence of Stochastic Processes*. Springer, New York.

RUDIN, W. (1973). *Functional Analysis*. McGraw-Hill, New York.

SHORACK, G. R. and WELLNER, J. A. (1986). *Empirical Processes with Applications to Statistics*. Wiley, New York.

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

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

YOUNG, W. H. (1916). On multiple integration by parts and the second theorem of the mean. *Proc. London Math. Soc.* (2) **16** 273–293.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF GRONINGEN
P.O. BOX 800
9700AV GRONINGEN
THE NETHERLANDS

DEPARTMENT OF MATHEMATICS
111 CUMMINGTON STREET
BOSTON UNIVERSITY
BOSTON, MASSACHUSETTS 02215