The Annals of Statistics
1996, Vol. 24, No. 3, 992–1024

ON THE ASYMPTOTIC EXPANSION OF THE EMPIRICAL PROCESS OF LONG-MEMORY MOVING AVERAGES

BY HWAI-CHUNG HO[1] AND TAILEN HSING[2]

Academia Sinica and Texas A & M University

Let $X_n = \sum_{i=1}^{\infty} a_i \varepsilon_{n-i}$, where the ε_i are iid with mean 0 and finite fourth moment and the a_i are regularly varying with index $-\beta$ where $\beta \in (1/2, 1)$ so that $\{X_n\}$ has long-range dependence. This covers an important class of the fractional ARIMA process. For $r \geq 0$, let $Y_{N,r} = \sum_{n=1}^{N} \sum_{1 \leq j_1 < \cdots < j_r} \prod_{s=1}^{r} a_{j_s} \varepsilon_{n-j_s}$, $Y_{N,0} = N$, $\sigma_{N,r}^2 = \mathrm{Var}(Y_{N,r})$ and $F^{(r)} =$ the rth derivative of the distribution function of X_n. The $Y_{N,r}$ are uncorrelated and are stochastically decreasing in r. For any positive integer $p < (2\beta - 1)^{-1}$, it is shown under mild regularity conditions that, with probability 1,

$$\sum_{n=1}^{N} I(X_n \leq x) = \sum_{r=0}^{p} (-1)^r F^{(r)}(x) Y_{N,r} + o\big(N^{-\lambda} \sigma_{N,p}\big)$$

uniformly for all $x \in \Re$ $\forall\, 0 < \lambda < (\beta - 1/2) \wedge (1/2 - p(\beta - 1/2))$.

This generalizes a host of existing results and provides the vehicle for a number of statistical applications.

1. Introduction. Let $\{\varepsilon_i\}$ be iid random variables with marginal distribution G, which has zero mean and finite variance. For some $\beta \in (1/2, 1)$, let a_i, $i \geq 1$, be regularly varying at ∞ with index $-\beta$, denoted by $a_i \in \mathrm{RV}_{-\beta}$ [i.e., $a_i = i^{-\beta} L(i)$ for some slowly varying function L; cf. Feller (1971)]. Define the moving-average process

$$X_n = \sum_{i \geq 1} a_i \varepsilon_{n-i}, \qquad n \geq 1. \tag{1.1}$$

Since $\sum_{i=1}^{\infty} a_i^2 < \infty$, $\{X_n\}$ is a well-defined, strictly stationary process. In this paper, we give an asymptotic expansion of the empirical process of $\{X_n\}$ and some ensuing applications. This paper is organized as follows. The assumptions and main results together with some remarks are stated in Section 2 which also contains a heuristic argument of why the main results can be expected. Applications to density estimation are given in Section 3, and

Received August 1994; revised June 1995.
[1]Research supported in part by National Science Council of the Republic of China, Grants NSC84-2121-M-001-015 and NSC84-2121-M-001-025.
[2]Research supported by NAVY–ONR Grant N00014-92-J-1007 and NSC Grant NSC83-0208-M-001-030.
AMS 1991 *subject classifications*. Primary 60G10; secondary 60G30, 60F17.
Key words and phrases. Asymptotic expansion, empirical process, fractional ARIMA process, long-range dependence, noncentral limit theorem.

applications to partial sums, empirical characteristic functions and U and von Mises statistics are given in Section 4. A Bahadur-type representation which gives an asymptotic expansion of the quantile process is given in Section 5. Proofs and technical details are delayed until Section 6.

It is clear that $\rho_j := E(X_0 X_j) \in \mathrm{RV}_{1-2\beta}$ is not summable, which roughly corresponds to the spectral density of $\{X_n\}$ having a pole at the zeroth frequency. Expressions commonly used to describe second-order stationary processes possessing this property include long memory, long-range dependence, strong dependence and so on.

The long-memory phenomena were first recorded quite some time ago. A practical motivation came from hydrology. The expression "Hurst effect" describes the long-range dependence of the so-called R/S statistic in hydrology. See Hurst (1951) and Mandelbrot and Taqqu (1979). On the other hand, a concrete example of a long-memory Gaussian sequence was first delivered by Rosenblatt (1961). Applying a certain nonlinear function on a Gaussian sequence which is not strongly mixing, he obtained a new type of non-Gaussian limit law for the partial sums. Nowadays models with long-range dependence are recognized to have an ever-increasing importance in various areas of human and natural sciences. This is reflected in the rapidly growing literature on that subject. See the review papers of Taqqu (1985), Künsch (1986), Beran (1992), Robinson (1994) and the references therein.

For a variety of reasons, two classes of models are of distinct importance in the family of long-memory processes. They are moving averages and functions of Gaussian processes. The moving-average process $\{X_n\}$ considered in this paper covers an important subset of the so-called fractional ARIMA process. See Remark 2.4 and the review paper by Robinson (1994) which emphasizes the time series aspect of long-range dependence. A number of recent papers study the asymptotic behavior of partial sums and empirical processes of random variables from these models. Without intending to be complete, we mention the following papers which are the most relevant ones to the present theme. Davydov (1970) considers the partial sum process of a long-memory moving-average process. Rosenblatt (1961), Taqqu (1975, 1979) and Dobrushin and Major (1979) study partial sums of nonlinear functions of strongly dependent Gaussian sequences. Surgailis (1983) and Avram and Taqqu (1987) investigate the partial sums of certain smooth nonlinear functions (certain entire functions and Appell polynomials) of moving averages with long memory. Finally, Dehling and Taqqu (1989) consider the empirical process of nonlinear functions of Gaussian sequences with long-range dependence. They establish the weak convergence in $D([-\infty, +\infty] \times [0, 1])$ equipped with the sup-norm and obtain an orthogonal expansion for the empirical process. On the other end of the spectrum, results concerning empirical processes of "short"-memory moving averages can be found in Billingsley (1968), Chanda and Ruymggart (1990) and Hesse (1990a).

As can be seen, the asymptotic behavior of the empirical process of long-memory moving averages has been mostly left untouched. One would naturally like to know whether or not an expansion in orthogonal terms like that derived by Dehling and Taqqu (1989) exists under the (non-Gaussian) moving-average setting.

While a formal expansion of the empirical process is possible using orthogonal polynomials, as is done in Dehling and Taqqu (1989) in the Gaussian case, it is by no means the most natural approach in the present setting. We show, based on a conditioning argument and recursive Taylor expansions, that it is possible to approximate the empirical process by an expansion of $p + 1$ uncorrelated terms, where p depends primarily on β. Each term is the product of a derivative of F, the distribution function of X_n, and a random variable which (after normalization) converges in distribution to a limit expressible by a multiple Wiener–Itô integral. The error of the approximation is a.s. $o(\sigma_{N,p})$ in the sup-norm, where $\sigma_{N,p}$ is the standard deviation of the stochastically smallest term in the expansion with N denoting sample size. This description may give the impression that the expansion is similar in spirit to Taylor expansions. Indeed, what we do is essentially Taylor expansions of indicators, which is possible in the long-range setting.

Since the empirical process contains all the information of the sample, one expects that the asymptotic expansion can be used to derive, for example, the asymptotic distributions of a class of statistics. Among applications that directly or indirectly follow, we mention the Kolmogorov–Smirnov statistic, density estimation, partial sums, empirical characteristic functions, U and von Mises statistics and a Bahadur-type representation for the quantile process. Most of these results are new and, in some cases, extensions of existing ones. Naturally, these are just a few of an array of possibilities in which our expansion plays a role. We anticipate other applications to surface in due course.

2. Main results. Let $\{X_n\}$ be the moving-average process defined in (1.1). Recall that G is the distribution function of ε_n, and let F be the distribution of X_n. Define

$$Y_{N,0} = N,$$
$$Y_{N,r} = \sum_{n=1}^{N} \sum_{1 \le j_1 < \cdots < j_r} \prod_{s=1}^{r} a_{j_s} \varepsilon_{n-j_s}.$$

For any positive integer r such that $r(2\beta - 1) < 1$, we have (cf. Lemma 6.1)

$$\sigma_{N,r}^2 = \operatorname{Var}(Y_{N,r}) \sim N^{2-r(2\beta-1)} L^{2r}(N),$$

where the notation $b_N \sim c_N$ means $b_N / c_N \to 1$ as $N \to \infty$. Let

$$S_{N,p}(x) = \sum_{n=1}^{N} I(X_n \le x) - \sum_{r=0}^{p} (-1)^r F^{(r)}(x) Y_{N,r}, \qquad p = 1, 2, \dots .$$

Our main results are Theorems 2.1 and 2.2.

THEOREM 2.1. *Assume that* $\int u\, dG(u) = 0$ *and* $\int u^4\, dG(u) < \infty$, *and* $a_i \in \mathrm{RV}_{-\beta}$ *for some* $\beta \in (1/2, 1)$. *Also assume that* G *is* $p + 3$ *times differentiable* *with bounded, continuous and integrable derivatives, where p is any positive integer less than* $(2\beta-1)^{-1}$. *Then, for any* $\zeta > 0$, *there exists a constant* $C < \infty$ *such that, for all* $b > 0$,

$$P\left\{\sigma_{N,p}^{-1} \sup_{x\in\Re} |S_{N,p}(x)| > b\right\} \le Cb^{-2}(1 \vee b^{-\zeta})N^{-\gamma(\beta,p)+\zeta},$$

where $\gamma(\beta,p) = (2\beta-1) \wedge (1-p(2\beta-1))$.

THEOREM 2.2. *Under the conditions of Theorem* 2.1, *for any* $\lambda < \gamma(\beta,p)/2$, *we have*

$$\frac{N^{\lambda}}{\sigma_{N,p}} \sup_{x\in\Re} |S_{N,p}(x)| \to 0 \quad a.s.$$

REMARK 2.1. The value of β restricts the possible number of terms p in the asymptotic expansion. The condition $p < (2\beta-1)^{-1}$ is needed to ensure that $N/\sigma_{N,p}^2 \to 0$, which is crucial for the expansion. It is clear, however, that a one-term expansion is always possible for any $\beta \in (1/2, 1)$.

REMARK 2.2. It is obvious that the $Y_{N,r}$ are uncorrelated. If $a_j \in \mathrm{RV}_{-\beta}$, $\beta \in (1/2,1)$ and $r < (2\beta-1)^{-1}$, then

$$\sigma_{N,r}^{-1} Y_{N,r} \to_d Z_r \quad \text{as } N \to \infty,$$

where the random variable Z_r can be represented by the multiple Wiener–Itô integral

$$Z_r = \kappa(\beta,r) \int_{-\infty<u_1<\cdots<u_r<1} \int \left\{\int_0^1 \prod_{j=1}^{r} \left[(v-u_j)^+\right]^{-\beta} dv\right\} dB(u_1)\cdots dB(u_r),$$

with B denoting standard Brownian motion and

$$\kappa(\beta,r) = \left\{r!\left(1-r\left(\beta-\frac{1}{2}\right)\right)(1-r(2\beta-1)) \times\left[\int_0^{\infty}(x+x^2)^{-\beta}\,dx\right]^{-r}\right\}^{1/2}, \tag{2.1}$$

ensuring $EZ_r^2 = 1$ (cf. Lemma 6.1). It can be shown that Z_1 is standard normal and Z_r is nonnormal for $r \ge 2$. See Taqqu (1979), Major (1981), Surgailis (1983) and Avram and Taqqu (1987).

REMARK 2.3. The conditions on the differentiability of G are needed for technical reasons. The important thing is that $F_j(\cdot) := P\{\sum_{i=1}^{j} a_i \varepsilon_{n-i} \le \cdot\}$ should converge to $F(\cdot)$ "fast" enough, and F should satisfy the differentiability conditions that are now assumed for G. Those do not seem too stringent in view of the discussion on page 273 of Hall and Hart (1990). As a result, we conjecture that our assumptions on G can be removed, or at least relaxed, with appropriate smoothing arguments. However, effort will not be made to justify that claim in this paper.

REMARK 2.4. A parsimonious device to model long-range dependence is the now well-known fractional ARIMA process introduced by Granger and Joyeux (1980) and Hosking (1981). A process $\{Y_n\}$ is said to be fractional ARIMA(p, d, q) with $d \in (-1/2, 1/2)$ if Y_n is a stationary solution of the difference equation [cf. Brockwell and Davis (1987), page 469]

$$\phi(B)\nabla^d Y_n = \theta(B) Z_n.$$

Here B is the backward shift operator, that is, $BY_n = Y_{n-1}$, $\{Z_n\}$ is white noise with $EZ_n = 0$ and $EZ_n^2 < \infty$, $\phi(z)$ and $\theta(z)$ are polynomials of degrees p and q, respectively, and in the form of

$$\phi(z) = 1 - \phi_1 z - \phi_2 z^2 - \cdots - \phi_p z^p$$

and

$$\theta(z) = 1 + \theta_1 z + \cdots + \theta_q z^q.$$

The fractional differencing operator $\nabla^d = (1 - B)^d$ is expressed as $\nabla^d = (1 - B)^d = \sum_{j=0}^{\infty} \pi_j B^j$ with $\pi_j = \Gamma(j - d)[\Gamma(j + 1)\Gamma(-d)]^{-1}$. The moving-average process $\{X_n\}$ in this paper covers a class of the fractional ARIMA process with $0 < d < 1/2$ [cf. Granger and Joyeux (1980)].

The proofs of Theorems 2.1 and 2.2 will be given in Section 6. At this point, it might be helpful to explain heuristically why such an expansion exists without providing detailed justifications. Define the truncations

$$X_{n,j} = \sum_{1 \le i \le j} a_i \varepsilon_{n-i}, \qquad \tilde{X}_{n,j} = \sum_{i > j} a_i \varepsilon_{n-i}, \qquad n, j \ge 1. \tag{2.2}$$

Also let

$$F_j(x) = P\{X_{n,j} \le x\}, \qquad \tilde{F}_j(x) = P\{\tilde{X}_{n,j} \le x\}, \qquad n, j \ge 1. \tag{2.3}$$

Let $\tilde{X}_{n,0} = X_n$ and $F_0(x) = I(x \ge 0)$. Write

$$\sum_{n=1}^{N} (I(X_n \le x) - F(x)) = \sum_{n=1}^{N} \sum_{j=1}^{\infty} \left(F_{j-1}\left(x - \tilde{X}_{n,j-1}\right) - F_j\left(x - \tilde{X}_{n,j}\right) \right).$$

For large j, by the Taylor expansion,

$$\begin{aligned} F_{j-1}&\left(x - \tilde{X}_{n,j-1}\right) - F_j\left(x - \tilde{X}_{n,j}\right) \\ &\approx F_j\left(x - \tilde{X}_{n,j-1}\right) - F_j\left(x - \tilde{X}_{n,j}\right) \\ &= F_j\left(x - \tilde{X}_{n,j} - a_j \varepsilon_{n-j}\right) - F_j\left(x - \tilde{X}_{n,j}\right) \\ &\approx -a_j \varepsilon_{n-j} F_j^{(1)}\left(x - \tilde{X}_{n,j}\right) \\ &\approx -a_j \varepsilon_{n-j} F^{(1)}(x). \end{aligned} \tag{2.4}$$

Thus, we obtain the first and second term of the expansion:

$$\sum_{n=1}^{N} I(X_n \le x) = NF(x) - F^{(1)}(x) \sum_{n=1}^{N} \sum_{j=1}^{\infty} a_j \varepsilon_{n-j} + e_N^{(1)}(x), \tag{2.5}$$

where $e_N^{(1)}(x)$ denotes the error part. We came to (2.4) by dropping the quantity

$$-\sum_{n=1}^{N} \sum_{j=1}^{\infty} a_j \varepsilon_{n-j} \Big(F_j^{(1)}\big(x - \tilde{X}_{n,j}\big) - F^{(1)}(x) \Big), \tag{2.6}$$

where the same tactic as before gives

$$F_j^{(1)}\big(x - \tilde{X}_{n,j}\big) - F^{(1)}(x) = \sum_{k=j+1}^{\infty} \Big(F_{k-1}^{(1)}\big(x - \tilde{X}_{n,k-1}\big) - F_k^{(1)}\big(x - \tilde{X}_{n,k}\big) \Big).$$

Again, by Taylor expansion,

$$\begin{aligned} &\sum_{k=j+1}^{\infty} \Big(F_{k-1}^{(1)}\big(x - \tilde{X}_{n,k-1}\big) - F_k^{(1)}\big(x - \tilde{X}_{n,k}\big) \Big) \\ &\qquad \approx -\sum_{k=j+1}^{\infty} a_k \varepsilon_{n-k} F_k^{(2)}\big(x - \tilde{X}_{n,k}\big) \\ &\qquad \approx -F^{(2)}(x) \sum_{k=j+1}^{\infty} a_k \varepsilon_{n-k}, \end{aligned}$$

so that the error $e_N^{(1)}(x)$ of (2.5) contains the term

$$F^{(2)}(x) \sum_{n=1}^{N} \sum_{j=1}^{\infty} \sum_{k=j+1}^{\infty} a_j a_k \varepsilon_{n-j} \varepsilon_{n-k}.$$

Hence, by (2.5) and (2.6),

$$\begin{aligned} \sum_{n=1}^{N} I(X_n \le x) = NF(x) - F^{(1)}(x) \sum_{n=1}^{N} \sum_{j=1}^{\infty} a_j \varepsilon_{n-j} \\ + F^{(2)}(x) \sum_{n=1}^{N} \sum_{j=1}^{\infty} \sum_{k=j+1}^{\infty} a_j a_k \varepsilon_{n-j} \varepsilon_{n-k} + e_N^{(2)}(x). \end{aligned}$$

It is clear that the same story can be told again to obtaina the next term in the expansion and so on. What is left is then a careful analysis of the magnitude of the size of the error $e_N^{(i)}(x)$ in each step. To do that simultaneously for all $x \in \Re$, we use an argument which is similar in spirit to one used by Dehling and Taqqu (1989).

3. Kolmogorov–Smirnov statistic and density estimation. Denote by F_N the empirical process of $X_1, \dots, X_N$; that is,

$$F_N(x) = \frac{1}{N} \sum_{n=1}^{N} I(X_n \le x), \qquad x \in \Re.$$

Assume the conditions of Theorem 2.1 with $p = 1$. Then Theorem 2.2 implies that

$$\frac{N}{\sigma_{N,1}} \sup_{x \in \Re} \left| F_N(x) - F(x) + f(x) N^{-1} Y_{N,1} \right| \to 0 \quad \text{a.s.,} \tag{3.1}$$

where f is the density of F. By Remark 2.2,

$$\sigma_{N,1}^{-1} Y_{N,1} \to_d Z,$$

where Z is standard normal, and hence it follows from (3.1) that

$$\frac{N}{\sigma_{N,1}} \sup_{x \in B} \left| F_N(x) - F(x) \right| \to_d |Z| \sup_{x \in B} f(x) \tag{3.2}$$

for any Borel set B. That is, we derived the limiting distribution of the Kolmogorov–Smirnov statistic.

We note in passing that if B contains only a single point y, then what determines the variance of the empirical distribution is the concentration of probability mass at y, that is, $f(y)$, rather than the accumulation of probability mass up to y, that is, $F(y)$, as in the short-range case. It is also worth noting that for each fixed x_0 the empirical distribution $F_N(x_0)$ can serve as a statistic to test the null hypothesis $f(x_0) = 0$, and, by Theorem 2.2, the critical region will be determined by the distribution of Z_3 provided $3(2\beta - 1) < 1$ (cf. Remark 2.2). Another frequently encountered occasion which needs Theorem 2.2 with $p > 1$ for making statistical inference will be presented later in Section 4.

In order, for example, to find a confidence band for F using normal approximation based on (3.2), a consistent estimate of $\sup_x f(x)$ should be available beforehand. This and some other asymptotic properties related to density estimation are discussed as follows.

Let us consider estimating f using a sample $X_1, \dots, X_N$. The popular kernel density estimator of f is

$$\hat{f}_N(x) = \frac{1}{Nh} \sum_{n=1}^{N} K\left(\frac{x - X_n}{h} \right),$$

where K is a density function unless otherwise specified and h is the bandwidth.

The issues of density estimation for long-memory moving-average processes were first studied by Hall and Hart (1990). They computed the mean integrated squared error of $\hat{f}_N$ and concluded, without providing limit laws, that its rate of convergence cannot be faster than $N^{1-2\beta}$, the rate at which $\overline{X}_N$ converges to 0. We use the maximum absolute deviation criterion and show that the best rate $N^{1-2\beta}$ can be reached under regularity conditions. The maximum absolute deviation criterion for density estimation was used in Bickel and Rosenblatt (1973) in the iid setting.

THEOREM 3.1. *Assume that* K *is of bounded total variation and* $\int |u| K(u)\,du < \infty$, *and that the conditions of Theorem* 2.1 *hold for* $p = 1$. *Let* $h = N^{-\delta}$, *where* $0 < \delta < (2\beta - 1) \wedge 1/2$. *Then, for*

$$\lambda < \left(\beta - \tfrac{1}{2}\right) \wedge \delta \wedge \left[\left((2\beta - 1) \wedge \tfrac{1}{2}\right) - \delta\right], \tag{3.3}$$

we have

$$N^{\lambda} \sup_{x \in \Re} \left| \hat{f}_N(x) - f(x) \right| \to 0 \quad a.s. \tag{3.4}$$

If, additionally, $1/2 < \beta < 5/6$,

$$\tfrac{1}{2}\left(\beta - \tfrac{1}{2}\right) < \delta < \left(\beta - \tfrac{1}{2}\right) \wedge (1 - \beta) \tag{3.5}$$

and

$$\int uK(u)\,du = 0 \quad \text{and} \quad \int u^2 K(u)\,du < \infty,$$

then

$$\frac{N}{\sigma_{N,1}} \sup_{x \in \Re} \left| \hat{f}_N(x) - f(x) \right| \to_d |Z| \sup_{x \in \Re} |f'(x)| \quad \text{as } N \to \infty. \tag{3.6}$$

Relations (3.3) and (3.4) are consistent with the message of Hall and Hart (1990) that the rate of convergence of the kernel density estimator is bounded by the rate of $\overline{X}_N$ converging to 0. On the other hand, (3.6) shows that the optimal rate can be reached for $1/2 < \beta < 5/6$. Note that the condition $\beta < 5/6$ is required to ensure that the choice of δ in (3.5) is possible. This is indeed a troublesome issue, which arises from the bias of the kernel density estimator. The result below serves two purposes. First, it extends (3.4) and (3.6) to include the estimation for the derivatives of f. Second, it shows that the condition $\beta < 5/6$ can be relaxed if one is willing to use higher-order kernels.

Let $f^{(j)}$ be the jth derivative of f. The estimator we consider for $f^{(j)}$ is

$$\hat{f}_N^{(j)}(x) = \frac{1}{Nh^{j+1}} \sum_{n=1}^{N} K^{(j)}\left(\frac{x - X_n}{h}\right).$$

THEOREM 3.2. *Suppose the conditions of Theorem* 2.1 *hold with* $p = 1$. *For a nonnegative integer* q, *let* f *be* $q + 1$ *times differentiable with bounded, continuous and integrable derivatives. Assume that the kernel* K *is* q *times differentiable and* $K^{(q)}$ *is of bounded total variation. Let* $h = N^{-\delta}$ *with* $0 < \delta < [(2\beta - 1) \wedge 1/2]/(q + 1)$. *Then, for*

$$\lambda < \left(\beta - \tfrac{1}{2}\right) \wedge \delta \wedge \left[\left((2\beta - 1) \wedge \tfrac{1}{2}\right) - (q + 1)\delta\right],$$

we have

$$N^{\lambda} \sup_{x \in \Re} \left| \hat{f}_N^{(q)}(x) - f^{(q)}(x) \right| \to 0 \quad a.s. \tag{3.7}$$

Assume, additionally, that for some integer

$$l > (q+1) \vee \frac{(q+1)(\beta - 1/2)}{1-\beta}, \tag{3.8}$$

such that $f^{(q+l)}$ *is bounded and continuous and that* K *is of order* $l-1$, *that is*, $\int u^j K(u)\,du = 0$, $1 \le j \le l-1$, *and satisfies* $\int |u|^l K(u)\,du < \infty$. *If*

$$\frac{\beta - 1/2}{l} < \delta < \frac{(\beta - 1/2) \wedge (1-\beta)}{q+1}, \tag{3.9}$$

then

$$\begin{aligned} &\frac{N}{\sigma_{N,1}} \sup_{x \in \Re} \left| \hat{f}_N^{(q)}(x) - f^{(q)}(x) \right| \\ &\qquad \to_d |Z| \sup_{x \in \Re} \left| f^{(q+1)}(x) \right| \quad \textit{as } N \to \infty. \end{aligned} \tag{3.10}$$

4. Partial sums, empirical characteristic functions, *U* and von Mises statistics. The random variables listed in the heading form a body of important statistics in practice. We now explain how to derive their asymptotic distributions for the sample $X_1, \ldots, X_N$. We consider first partial sums.

The same ideas in the proofs of Theorems 2.1 and 2.2 are still useful for studying partial sums. But the volume of additional details involved is considerable. It therefore seems prudent to leave them to a subsequent paper. Nevertheless, it is useful to describe what kind of theory is possible and the relationship with existing results.

Let K: $\Re \to \Re$ be a measurable function. Define

$$K_\infty(x) = EK(x + X_n),$$

which is assumed to be finite and p-times differentiable at 0, where p is any positive integer less than $(2\beta - 1)^{-1}$. Then, under suitable regularity conditions, we can obtain, with probability 1,

$$\sum_{j=1}^{N} K(X_j) = \sum_{r=0}^{p} K_\infty^{(r)}(0) Y_{N,r} + o\left(\sigma_{N,p} N^{-\lambda}\right) \quad \text{for some } \lambda > 0. \tag{4.1}$$

If $K(u) = I(u \le x)$ for some $x \in \Re$, then (4.1) reduces to the empirical process case for a fixed x. This is related to the weak convergence results of Surgailis (1983) and Avram and Taqqu (1987). Note that here $K_\infty^{(r)}(0)$ plays the role of $EK^{(r)}(X_n)$ in Surgailis (1983) and Avram and Taqqu (1987). While these existing results require that K be smooth, we can bypass that with smoothness on F. We also note that if the unknown parameter $\theta = \int K(x)\,dF(x)$ of interest is one for which $K_\infty^{(r)}(0) = 0$, $1 \le r \le p-1$, with $p < (2\beta - 1)$ and is estimated by $N^{-1}\sum_{n=1}^{N} K(X_n)$, then, by (4.1), the confidence interval for θ should be constructed via Z_p as specified in Remark 2.2. A commonly seen example is $\theta = \mathrm{Var}(X_n)$ and $K(x) = (x - EX_n)^2$ with $p = 2$.

In the investigation of (4.1), we also plan to address a uniformity issue. That is, what is a natural and general class $\mathscr{K}$ for which

$$\sup_{K \in \mathscr{K}} \left| \sum_{j=1}^{N} K(X_n) - \sum_{r=0}^{p} K_{\infty}^{(r)}(0) Y_{N,r} \right| = o(\sigma_{N,p}) \quad \text{a.s.?}$$

See Pollard (1984), Chapter 2. These questions are more general and more difficult to answer than the ones we have dealt with so far in this paper. Our preliminary investigations show that there is a good prospect that we will be able to resolve them to some extent in the near future.

To illustrate the above, we give a complete proof of the following result. Let F_N be the empirical process based on a sample $X_1, \dots, X_N$. Define

$$\phi(t) = \int e^{itx}\, dF(x) \quad \text{and} \quad \phi_N(t) = \int e^{itx}\, dF_N(x),$$

the characteristic function and empirical characteristic function, respectively.

THEOREM 4.1. *Assume that* $\int u\, dG(u) = 0$, $\int u^4\, dG(u) < \infty$ *and* $a_i \in \mathrm{RV}_{-\beta}$, $\beta \in (1/2, 1)$. *Then, for all* $t \in \Re$,

$$\frac{N^{1+\lambda}}{\sigma_{N,1}} \left| \phi_N(t) - \phi(t) - it\phi(t) Y_{N,1} N^{-1} \right| \to 0 \quad a.s.\ \forall\, \lambda < \left(\beta - \frac{1}{2}\right) \wedge (1 - \beta) \tag{4.2}$$

($Y_{N,1}N^{-1}$ *is the sample mean*), *and hence*

$$\frac{N}{\sigma_{N,1}} (\phi_N(t) - \phi(t)) \to_d it\phi(t) Z, \tag{4.3}$$

where $Z \sim N(0,1)$.

The proof, which uses the ideas of Theorem 2.2, is given in Section 6. For the Gaussian case, Beran and Ghosh (1991) obtain the weak convergence of $\sup_{t \in B} |\phi_N(t) - \phi(t)|$ for any bounded interval B. While this can also be done in our setting, we choose not to pursue it here.

Similar to the partial sum case, the basic approach in the proof of Theorems 2.1 and 2.2 can be extended to treat U and von Mises statistics in general. Again, it seems prudent to leave that to a subsequent paper which is solely dedicated to such statistics. To give a flavor of what could be expected, we present the following simplified version, as a direct application of Theorem 2.2. The formulation is adopted from Dehling and Taqqu (1989).

Let $w: \Re^k \to \Re$ be a measurable function satisfying the following: (1) $w(\cdot)$ is invariant under the permutations of its arguments and (2) $w(\cdot)$ is degenerate in the sense that

$$\int |w(x_1, \dots, x_k)|\, dF(x_1) \dots dF(x_k) < \infty$$

and

$$\int w(x_1, \dots, x_k)\, dF(x_1) = 0 \qquad \forall\, x_2, \dots, x_k.$$

Define the nonnormalized U and von Mises statistics $U_N(w)$ and $V_N(w)$ with the kernel w as

$$U_N(w) = \sum_{\substack{1 \le j_1, \dots, j_k \le N \\ j_s \ne j_t,\ s \ne t}} w(X_{j_1}, \dots, X_{j_k})$$

and

$$V_N(w) = \sum_{1 \le j_1, \dots, j_k \le N} w(X_{j_1}, \dots, X_{j_k}).$$

Due to the degeneracy of w, $U_N(w)$ and $V_N(w)$ can be written as

$$U_N(w) = N^k \int_{A^k} w(x_1, \dots, x_k)\, d[F_N(x_1) - F(x_1)] \dots d[F_N(x_k) - F(x_k)],$$

$$V_N(w) = N^k \int_{\Re^k} w(x_1, \dots, x_k)\, d[F_N(x_1) - F(x_1)] \dots d[F_N(x_k) - F(x_k)],$$

where $A^k = \{(x_1, \dots, x_k) \in \Re^k | x_i \ne x_j\ \forall\ i \ne j\}$.

THEOREM 4.2. *Assume that the degenerate kernel w is of bounded total variation and that the conditions of Theorem* 2.1 *hold for $p = 1$. If*

$$\mu_k := (-1)^k \int_{\Re^k} f'(x_1) \cdots f'(x_k) w(x_1, \dots, x_k)\, dx_1 \dots dx_k \ne 0,$$

then $\sigma_{N,1}^{-k} U_N(w)$ and $\sigma_{N,1}^{-k} V_N(w)$ both converge in distribution to $Z^k \mu_k$, where $Z \sim N(0, 1)$.

The proof of Theorem 4.2 follows basically the same line of arguments given by Dehling and Taqqu (1989) and is omitted.

5. A Bahadur-type representation. In a now classic paper, Bahadur (1966) described the asymptotic relationship between the empirical distribution function and the empirical quantile function. Similar results also hold for stationary processes, for example, ϕ-mixing sequences [Sen (1972)], and linear processes, possibly having infinite variance, with coefficients a_j and the iid random variable ε_n satisfying $E|\varepsilon_n|^\alpha < \infty$, $\alpha > 0$, and $|a_j| = O(j^{-q})$, $q > 1 + (2/\alpha)$ [Hesse (1990b)].

Based on the expansion of Theorem 2.2, we show that, using ideas similar to those in Bahadur (1966), it is possible to obtain a "Bahadur-type" representation for the long-memory moving averages discussed here.

Define the quantile function

$$Q(y) = F^{-1}(y) := \inf\{x \colon F(x) \ge y\},$$

and, for a given sample $X_1, X_2, \dots, X_N$, with order statistics $X_{N:1}$, $X_{N:2}, \dots, X_{N:N}$, define the empirical quantile function

$$Q_N(y) = F_N^{-1}(y) := \inf\{x: F_N(x) \geq y\} = X_{N:k} \quad \text{if } \frac{k-1}{N} < y \leq \frac{k}{N},\ k = 1, \dots, N.$$

The following result gives a Bahadur-type representation.

THEOREM 5.1. *Assume the conditions of Theorem* 2.1 *for* $p = 1$. *Let* $0 < a < b < 1$ *be such that* $\inf_{Q(a) < x < Q(b)} f(x) > 0$. *Then*

$$\sup_{a<y<b} \left| Q_N(y) - Q(y) - \frac{Y_{N,1}}{N} \right| = o\left(\frac{\sigma_{N,1}}{N^{1+\lambda}} \right) \quad a.s.\ \forall\, 0 < \lambda < \left(\beta - \frac{1}{2} \right) \wedge (1 - \beta). \tag{5.1}$$

Hence, for all $y \in (a, b)$,

$$\frac{N}{\sigma_{N,1}} (Q_N(y) - Q(y)) \to_d Z, \tag{5.2}$$

where Z *is standard normal.*

It is interesting to note that the first-order behavior of $Q_N(y) - Q(y)$ is independent of y. Is there an expansion for $Q_N(y)$ which mirrors that for $F_N(x)$? This is certainly a topic that needs further investigation.

Beran (1991) observes that the M estimator of location for long-memory Gaussian sequences is as efficient as the sample mean. We show in the next corollary that in the case of symmetric, possibly non-Gaussian $\{X_n\}$, the same holds for the trimmed mean. For $0 < \alpha < 1/2$, define the α-trimmed mean

$$T_{N,\alpha} = \frac{1}{N - 2[N\alpha]} \sum_{n=[N\alpha]+1}^{N-[N\alpha]} X_{N:n}.$$

Clearly,

$$\left| T_{N,\alpha} - \frac{1}{1-2\alpha} \int_\alpha^{1-\alpha} Q_N(y)\, dy \right| = O(N^{-1}) \quad \text{a.s.} \tag{5.3}$$

and

$$\frac{1}{1-2\alpha} \int_\alpha^{1-\alpha} Q(y)\, dy = EX_n. \tag{5.4}$$

The following corollary is evident from (5.1), (5.3) and (5.4).

COROLLARY 5.2. *Assume the conditions of Theorem* 2.1 *with* $p = 1$. *Suppose the distribution of* X_n *is symmetric about its mean* μ. *Then, for all* $0 < \alpha < 1/2$,

$$\frac{N}{\sigma_{N,1}}(T_{N,\alpha} - \mu) \to_d Z,$$

provided that $\inf_{Q(\alpha) < x < Q(1-\alpha)} f(x) > 0$.

6. Proofs. Throughout this section, let F_j and $\tilde{X}_{n,j}$ be as defined by (2.2) and (2.3). We first give a few lemmas.

LEMMA 6.1. *Assume that* $a_i = i^{-\beta}L(i)$, *where* $\beta \in (1/2, 1)$ *and* L *is slowly varying at* ∞. *If* $\int u\,dG(u) = 0$ *and* $\int u^2\,dG(u) < \infty$ *and* p *is any positive integer less than* $(2\beta - 1)^{-1}$, *then*

$$\mathrm{Var}(Y_{N,p}) \sim \kappa(\beta, p) N^{2-p(2\beta-1)} L^{2p}(N),$$

where $\kappa(\beta, p)$ *is defined in* (2.1).

PROOF. It is easy to show that

$$\mathrm{Var}(Y_{N,p}) \sim \frac{2}{p!} \sum_{m=1}^{N-1} \sum_{n=1}^{m} \left(\sum_{j=1}^{\infty} a_j a_{n+j} \right)^p.$$

Now, for large n,

$$\left(\sum_{j=1}^{\infty} a_j a_{n+j} \right)^p \sim n^{-p(2\beta-1)} L^{2p}(n) \left(\int_0^{\infty} (x + x^2)^{-\beta}\,dx \right)^p.$$

The rest of the proof is straightforward. □

LEMMA 6.2. *Under the assumption on* G *in Theorem* 2.1, *for each* $j \ge 1$, F_j *is* $p + 3$ *times differentiable with bounded, continuous and integrable derivatives. Further, the integrals* $\int |F_j^{(i)}|$, $1 \le i \le p + 3$, *are nonincreasing in* j. *Similarly,* F *is* $p + 3$ *times differentiable with bounded, continuous and integrable derivatives.*

PROOF. We only show the part for F_j, since the part for F is proved similarly. Clearly, F_1 has the desired properties. Suppose the conclusions hold for $j = 1, \dots, k$. We show that they hold for $j = k + 1$. First,

$$F_{k+1}(x) = \int F_k(x - a_{k+1} y)\,dG(y).$$

Since $F_k^{(1)}$ is integrable and continuous, it follows from Fubini's theorem that

$$F_{k+1}^{(1)}(x) = \int F_k^{(1)}(x - a_{k+1} y)\,dG(y),$$

which is bounded, continuous and integrable, and $\int|F_{k+1}^{(1)}| \le \int|F_k^{(1)}|$. The same argument applies to higher-order derivatives. The proof follows from induction. □

LEMMA 6.3. *For all x, n, as $k \to \infty$,*

$$\sum_{j=1}^{k} \left(F_{j-1}\left(x - \tilde{X}_{n,j-1}\right) - F_j\left(x - \tilde{X}_{n,j}\right)\right) \to I(X_n \le x) - F(x)$$

almost surely and in L_2.

PROOF. Observe that

$$\sum_{j=1}^{k} \left(F_{j-1}\left(x - \tilde{X}_{n,j-1}\right) - F_j\left(x - \tilde{X}_{n,j}\right)\right) = I(X_n \le x) - F_k\left(x - \tilde{X}_{n,k}\right).$$

Since $F_k \to F$ and $\tilde{X}_{n,k} \to 0$ as $k \to \infty$ almost surely and in L_2, the result follows. □

LEMMA 6.4. *For all $x, x' \in \Re$ and $n, n', j, j' \ge 1$ such that $n' - j' \ne n - j$,*

$$\begin{aligned}\operatorname{Cov}\Big(&F_{j-1}\left(x - \tilde{X}_{n,j-1}\right) - F_j\left(x - \tilde{X}_{n,j}\right),\\ &F_{j'-1}\left(x' - \tilde{X}_{n',j'-1}\right) - F_{j'}\left(x' - \tilde{X}_{n',j'}\right)\Big) = 0\end{aligned} \tag{6.1}$$

and

$$\operatorname{Cov}\left(F_{j-1}\left(x - \tilde{X}_{n,j-1}\right) - F_j\left(x - \tilde{X}_{n,j}\right), F_{j'-1}^{(1)}\left(x' - \tilde{X}_{n',j'-1}\right)\varepsilon_{n'-j'}\right) = 0. \tag{6.2}$$

PROOF. Let

$$\xi_n(x) = I(X_n \le x) - F(x)$$

and

$$\xi_{n,j}(x) = F_{j-1}\left(x - \tilde{X}_{n,j-1}\right) - F_j\left(x - \tilde{X}_{n,j}\right), \qquad j \ge 1.$$

Let $\mathscr{F}_j$ be the σ-field generated by ε_i, $i < j$. Clearly,

$$\xi_{n,j}(x) = E\left(\xi_n(x) \middle| \mathscr{F}_{n-j+1}\right) - E\left(\xi_n(x) \middle| \mathscr{F}_{n-j}\right),$$

and hence $\xi_{n,j}(x)$ is measurable with respect to $\mathscr{F}_{n-j+1}$. Assume that $n' - j' < n - j$ so that $\mathscr{F}_{n'-j'+1} \subset \mathscr{F}_{n-j}$. It is clear that

$$\begin{aligned}&E\left(\xi_{n,j}(x)\xi_{n',j'}(x')\right)\\ &\quad= E\left(\xi_{n',j'}(x')E\left(E\left(\xi_n(x)\middle|\mathscr{F}_{n-j+1}\right) - E\left(\xi_n(x)\mathscr{F}_{n-j}\right)\middle|\mathscr{F}_{n'-j'+1}\right)\right)\\ &\quad= E\left(\xi_{n',j'}(x')\left(E\left(\xi_n(x)\middle|\mathscr{F}_{n'-j'+1}\right) - E\left(\xi_n(x)\middle|\mathscr{F}_{n'-j'+1}\right)\right)\right) = 0.\end{aligned}$$

This proves (6.1). A similar argument proves (6.2). □

LEMMA 6.5. *Given constants* $\gamma_1, \dots, \gamma_t > 1/2$, $t \geq 1$, *there exists* $C < \infty$ *such that*, *for all* $l \geq 1$,

$$\sum_{1 \le j_1 < j_2 < \cdots < j_t} \prod_{s=1}^{t} [j_s(l + j_s)]^{-\gamma_s} \le C \sum_{j=1}^{\infty} [j(l+j)]^{-\gamma}$$

$$\le \begin{cases} Cl^{-2\gamma+1}, & \text{if } \gamma \in \left(\frac{1}{2}, 1\right), \\ C\frac{\log l}{l}, & \text{if } \gamma = 1, \\ Cl^{-\gamma}, & \text{if } \gamma > 1, \end{cases} \tag{6.3}$$

where $\gamma = \sum_{s=1}^{t} \gamma_s - (t-1)/2$.

PROOF. The result can be proved by induction, as follows. For $t = 1$, (6.3) obviously holds; for example, if $\gamma_1 \in (1/2, 1)$,

$$\sum_{j=1}^{\infty} [j(l+j)]^{-\gamma_1} < l^{-2\gamma_1+1} \int_0^{\infty} [y(1+y)]^{-\gamma_1}\, dy \le Cl^{-2\gamma_1+1},$$

where the other two cases can be analyzed similarly using elementary arguments. Suppose now (6.3) holds for $t = \tau$. Then, by the Cauchy–Schwarz inequality,

$$\sum_{1 \le j_1 < j_2 < \cdots < j_{\tau+1}} \prod_{s=1}^{\tau+1} [j_s(l+j_s)]^{-\gamma_s}$$

$$\le \sum_{1 \le j_1 < j_2 < \cdots < j_\tau} \left(\prod_{s=1}^{\tau} [j_s(l+j_s)]^{-\gamma_s} \right) \left(\sum_{j > j_\tau} j^{-2\gamma_{\tau+1}} \right)^{1/2} \left(\sum_{j > l + j_\tau} j^{-2\gamma_{\tau+1}} \right)^{1/2}$$

$$\le C \sum_{1 \le j_1 < j_2 < \cdots < j_\tau} \left(\prod_{s=1}^{\tau} [j_s(l+j_s)]^{-\gamma_s} \right) [j_\tau(l+j_\tau)]^{-(\gamma_{\tau+1} - 1/2)}.$$

Thus (6.3) holds for $t = \tau + 1$ by the induction assumption. □

We are now ready to prove Theorem 2.1. To streamline the notation, we shall drop the "p" in $S_{N,p}$ and $\sigma_{N,p}$ and simply write S_N and σ_N. For convenience, if f is a function on $\Re$, let

$$f(x, y) = f(y) - f(x).$$

PROOF OF THEOREM 2.1. Define

$$\Lambda(y) = \sum_{r=0}^{p} \int_{-\infty}^{y} \left| F^{(r+1)}(u) \right| du, \qquad x \in \Re;$$

Λ is continuous, bounded and nondecreasing. The main ingredient of our approach is the so-called "chain argument" used by Dehling and Taqqu (1989). For $k = 1, \dots, K$, where K will be determined later, let

$$y_i(k) = \inf\{y: \Lambda(y) \geq \Lambda(+\infty) i/2^k\}, \qquad i = 0, \dots, 2^k.$$

For $k = 1, \dots, K$ and $x \in \Re$, let $i_k(x)$ be such that

$$y_{i_k(x)}(k) \leq x < y_{i_k(x)+1}(k),$$

whereas $y_{i_0(x)}(0) := -\infty$. Thus,

$$S_N(x) = \sum_{k=0}^{K-1} S_N\big(y_{i_k(x)}(k), y_{i_{k+1}(x)}(k+1)\big) + S_N\big(y_{i_K(x)}(K), x\big).$$

First,

$$\begin{aligned} \frac{1}{\sigma_N}|S_N(x)| \leq{}& \frac{1}{\sigma_N}\left|\sum_{k=0}^{K-1} S_N\big(y_{i_k(x)}(k), y_{i_{k+1}(x)}(k+1)\big)\right| \\ &+ \frac{1}{\sigma_N}\left|S_N\big(y_{i_K(x)}(K), y_{i_K(x)+1}(K)\big)\right| \\ &+ 2\frac{N}{\sigma_N} F\big(y_{i_K(x)}(K), y_{i_K(x)+1}(K)\big) \\ &+ \sum_{r=1}^{p}\Big(\left|F^{(r)}(x) - F^{(r)}\big(y_{i_K(x)}(K)\big)\right| \\ &\qquad + \left|F^{(r)}(x) - F^{(r)}\big(y_{i_K(x)+1}(K)\big)\right|\Big)\frac{1}{\sigma_N}|Y_{N,r}|. \end{aligned} \tag{6.4}$$

Now, for a given $b > 0$, choose

$$K = \log_2\left(\frac{N^\alpha}{b}\right) \tag{6.5}$$

for some $\alpha > 0$ which is large enough to guarantee

$$\frac{N^{1-\alpha}}{\sigma_N} \leq \frac{1}{4} \quad \text{and} \quad \sum_{r=1}^{p} \frac{\sigma_{N,r}^2}{N^{2\alpha}\sigma_N^2} = o\big(N^{-(2\beta-1)}\big). \tag{6.6}$$

Thus,

$$\sup_{x\in\Re} 2\frac{N}{\sigma_N} F\big(y_{i_K(x)}(K), y_{i_K(x)+1}(K)\big) \leq \frac{b}{2}, \tag{6.7}$$

$$\begin{aligned} \max_{1\leq r\leq p}\sup_{x\in\Re}&\Big(\left|F^{(r)}(x) - F^{(r)}\big(y_{i_K(x)}(K)\big)\right| \\ &+ \left|F^{(r)}(x) - F^{(r)}\big(y_{i_K(x)+1}(K)\big)\right|\Big) \\ \leq{}& 2\cdot 2^{-K} = 2N^{-\alpha}b. \end{aligned} \tag{6.8}$$

By (6.7), (6.4) and Boole's inequality,

$$
\begin{aligned}
P\Big\{\sigma_N^{-1}\sup_x |S_N(x)| > b\Big\}
&\le P\Bigg\{\sigma_N^{-1}\max_x\Bigg(\bigg|\sum_{k=0}^{K-1} S_N\big(y_{i_k(x)}(k), y_{i_{k+1}(x)}(k+1)\big)\bigg| \\
&\qquad + \Big|S_N\big(y_{i_K(x)}(K), y_{i_K(x)+1}(K)\big)\Big|\Bigg) > \frac{b}{4}\Bigg\} \\
&\quad + \sum_{r=1}^{p} P\Bigg\{\sup_x\Big(\Big|F^{(r)}(x) - F^{(r)}\big(y_{i_K(x)}(K)\big)\Big| \\
&\qquad + \Big|F^{(r)}(x) - F^{(r)}\big(y_{i_K(x)+1}(K)\big)\Big|\Big)\sigma_N^{-1}|Y_{N,r}| > \frac{b}{4p}\Bigg\}.
\end{aligned}
$$

By (6.8), Lemma 6.6 and Chebyshev's inequality, for any $\zeta > 0$ there exists $C < \infty$ such that the sum on the right is bounded by

$$
Cb^{-2}(1 \vee b^{-\zeta})N^{-\gamma(\beta,p)+\zeta} + (8p)^2 \sum_{r=1}^{p} \frac{\sigma_{N,r}^2}{N^{2\alpha}\sigma_N^2}.
$$

The conclusion of the theorem follows from (6.6). $\square$

To complete the proof of Theorem 2.1, it suffices to prove the following lemma.

LEMMA 6.6. *Assume the conditions of Theorem* 2.1 *and let K be chosen by* (6.5). *Then, for any $\zeta > 0$, there exists a constant $C < \infty$ such that, for all $b > 0$,*

$$
\begin{aligned}
&P\Bigg\{\max_x\Bigg(\bigg|\sum_{k=0}^{K-1} S_N\big(y_{i_k(x)}(k), y_{i_{k+1}(x)}(k+1)\big)\bigg| \\
&\qquad + \Big|S_N\big(y_{i_K(x)}(K), y_{i_K(x)+1}(K)\big)\Big|\Bigg) > b\sigma_N\Bigg\} \\
&\quad \le Cb^{-2}(1 \vee b^{-\zeta})N^{-\gamma(\beta,p)+\zeta}.
\end{aligned}
$$

PROOF. In the following, B and C are generic constants whose values may change from line to line. Write

$$
\begin{aligned}
T_{N,1}(x) = {} & \sum_{n=1}^{N}\big(I(X_n \le x) - F(x)\big) \\
& - \sum_{r=1}^{p-1}(-1)^r F^{(r)}(x)\sum_{n=1}^{N}\sum_{2\le j_1<\cdots<j_r}\prod_{s=1}^{r} a_{j_s}\varepsilon_{n-j_s} \\
& - (-1)^p \sum_{n=1}^{N}\sum_{2\le j_1<\cdots<j_p}\left(\prod_{s=1}^{p} a_{j_s}\varepsilon_{n-j_s}\right)F_{j_p-1}^{(p)}\big(x - \tilde{X}_{n,j_p}\big)
\end{aligned}
$$

and

$$T_{N,2}(x) = (-1)^{p-1} \sum_{n=1}^{N} \sum_{2 \le j_1 < \cdots < j_p} \left(\prod_{s=1}^{p} a_{j_s} \varepsilon_{n-j_s} \right) \left(F^{(p)}(x) - F^{(p)}_{j_p - 1}\left(x - \tilde{X}_{n, j_p}\right) \right)$$

and

$$T_{N,3}(x) = -\sum_{r=1}^{p} (-1)^r F^{(r)}(x) \sum_{n=1}^{N} \sum_{1 < j_2 < \cdots < j_r} a_1 \varepsilon_{n-1} \prod_{s=2}^{r} a_{j_s} \varepsilon_{n-j_s},$$

so that

$$S_N(x) = T_{N,1}(x) + T_{N,2}(x) + T_{N,3}(x),$$

where we use the convention $\sum_{r=1}^{0} \cdot = 0$ and $\sum_{1 < j_2 < \cdots < j_1} \prod_{s=2}^{1} \cdot = 1$. It clearly suffices to show the claim in the lemma with S_N replaced by $T_{N,1}$, $T_{N,2}$ and $T_{N,3}$.

By Boole's inequality and the fact that $\sum_{k=0}^{\infty} (k+3)^{-2} \le 1/2$,

$$\begin{aligned} P\Bigg\{ \max_x \Bigg(\Bigg| \sum_{k=0}^{K-1} T_{N,m}\big(y_{i_k(x)}(k), y_{i_{k+1}(x)}(k+1)\big) \Bigg| \\ + \Big| T_{N,m}\big(y_{i_K(x)}(K), y_{i_K(x)+1}(K)\big) \Big| \Bigg) > b \Bigg\} \\ \le \sum_{k=0}^{K-1} P\left\{ \max_x \Big| T_{N,m}\big(y_{i_k(x)}(k), y_{i_{k+1}(x)}(k+1)\big) \Big| > \frac{b}{(k+3)^2} \right\} \\ + P\left\{ \max_x \Big| T_{N,m}\big(y_{i_K(x)}(K), y_{i_K(x)+1}(K)\big) \Big| > \frac{b}{(K+3)^2} \right\}. \end{aligned} \tag{6.9}$$

For $k < K$, observe that $y_{i_k(x)}(k)$ and $y_{i_{k+1}(x)}(k+1)$ are neighboring points in

$$y_0(k+1), \dots, y_{2^{k+1}}(k+1).$$

By Boole's and Chebyshev's inequalities,

$$\begin{aligned} P\left\{ \max_x \Big| T_{N,m}\big(y_{i_k(x)}(k), y_{i_{k+1}(x)}(k+1)\big) \Big| > \frac{b}{(k+3)^2} \right\} \\ \le \sum_{i=0}^{2^{k+1}-1} P\left\{ \Big| T_{N,m}\big(y_i(k+1), y_{i+1}(k+1)\big) \Big| > \frac{b}{(k+3)^2} \right\} \\ \le \frac{(k+3)^4}{b^2} \sum_{i=0}^{2^{k+1}-1} \operatorname{Var}\Big(T_{N,m}\big(y_i(k+1), y_{i+1}(k+1)\big)\Big). \end{aligned} \tag{6.10}$$

The term $P\{\max_x |T_{N,m}(y_{i_K(x)}(K), y_{i_{K+1}(x)}(K))| > b/(K+3)^2\}$ is handled in the same way. By (6.9), (6.10) and the choice of K [cf. (6.5)], for any $\zeta > 0$ there exists $C < \infty$ such that

$$P\left\{\max_x \left(\left| \sum_{k=0}^{K-1} \frac{1}{\sigma_N} T_{N,m}\big(y_{i_k(x)}(k), y_{i_{k+1}(x)}(k+1)\big) \right| + \frac{1}{\sigma_N} \left| T_{N,m}\big(y_{i_K(x)}(K), y_{i_K(x)+1}(K)\big) \right| \right) > b \right\}$$

$$\begin{aligned} &\le \frac{2}{\sigma_N^2 b^2} \sum_{k=0}^{K-1} (k+3)^4 \sum_{i=0}^{2^{k+1}-1} \operatorname{Var}\big(T_{N,m}(y_i(k+1), y_{i+1}(k+1))\big) \qquad (6.11)\\ &\le C\sigma_N^{-2} b^{-2} (1 \vee b^{-\zeta}) N^{\zeta} \\ &\quad \times \max_{0 \le k \le K-1} \sum_{i=0}^{2^{k+1}-1} \operatorname{Var}\big(T_{N,m}(y_i(k+1), y_{i+1}(k+1))\big). \end{aligned}$$

From this point on, we will drop the "$k+1$" in $y_i(k+1)$ and $y_{i+1}(k+1)$ to simplify the notation.

Our plan is to show that, for some universal constant $B < \infty$,

$$\sum_{i=0}^{2^{k+1}-1} \operatorname{Var}\big(T_{N,1}(y_i, y_{i+1})\big) \le BN, \qquad (6.12)$$

$$\sum_{i=0}^{2^{k+1}-1} \operatorname{Var}\big(T_{N,3}(y_i, y_{i+1})\big) \le BN, \qquad (6.13)$$

and, for any given small $\zeta > 0$, there exists $C < \infty$, independent of k, such that

$$\sum_{i=0}^{2^{k+1}-1} \operatorname{Var}\big(T_{N,2}(y_i, y_{i+1})\big) \le C\big(N \vee N^{2-(p+1)(2\beta-1)+\zeta}\big). \qquad (6.14)$$

Since $\sigma_N^2 \in \mathrm{RV}_{2-p(2\beta-1)}$ (cf. Lemma 6.1), the result follows from (6.11), (6.12), (6.13) and (6.14).

First, we show (6.12). For $t = 1, \dots, p$, define

$$\begin{aligned} T_{N,1}^{(t)}(x) = &\sum_{n=1}^{N} \big(I(X_n \le x) - F(x)\big) \\ &- \sum_{r=1}^{t-1} (-1)^r F^{(r)}(x) \sum_{n=1}^{N} \sum_{2 \le j_1 < \cdots < j_r} \prod_{s=1}^{r} a_{j_s} \varepsilon_{n-j_s} \\ &- (-1)^t \sum_{n=1}^{N} \sum_{2 \le j_1 < \cdots < j_t} \left(\prod_{s=1}^{t} a_{j_s} \varepsilon_{n-j_s} \right) F_{j_t-1}^{(t)}\big(x - \tilde{X}_{n,j_t}\big). \end{aligned}$$

We shall show inductively that, for $t = 1, \dots, p$,

$$\sum_{i=0}^{2^{k+1}-1} \operatorname{Var}\Big(T_{N,1}^{(t)}(y_i, y_{i+1})\Big) \le BN. \tag{6.15}$$

First, we show (6.15) for $t = 1$. By Lemma 6.3,

$$I(y_i < X_n \le y_{i+1}) - F(y_i, y_{i+1}) = \sum_{j=1}^{\infty} \Big(F_{j-1}\Big(y_i - \tilde{X}_{n,j-1}, y_{i+1} - \tilde{X}_{n,j-1}\Big) - F_j\Big(y_i - \tilde{X}_{n,j}, y_{i+1} - \tilde{X}_{n,j}\Big)\Big),$$

where the sum converges a.s. and in L_2. Hence,

$$\begin{aligned} T_{N,1}^{(1)}(y_i, y_{i+1}) &= \sum_{n=1}^{N} \big(I(y_i < X_n \le y_{i+1}) - F(y_i, y_{i+1})\big) \\ &\quad + \sum_{n=1}^{N} \sum_{j=2}^{\infty} a_j \varepsilon_{n-j} F_{j-1}^{(1)}\Big(y_i - \tilde{X}_{n,j}, y_{i+1} - \tilde{X}_{n,j}\Big) \\ &= \sum_{n=1}^{N} \sum_{j=1}^{\infty} R_{n,j}(y_i, y_{i+1}), \end{aligned}$$

where

$$R_{n,j}(x) = F_{j-1}\Big(x - \tilde{X}_{n,j-1}\Big) - F_j\Big(x - \tilde{X}_{n,j}\Big) + I(j \ge 2) F_{j-1}^{(1)}\Big(x - \tilde{X}_{n,j}\Big) a_j \varepsilon_{n-j}.$$

By Lemma 6.4, $R_{n,j}(y_i, y_{i+1})$ and $R_{n',j'}(y_i, y_{i+1})$ are uncorrelated if $n - j \ne n' - j'$. Letting $j' = n' - n + j$, we get

$$\begin{aligned} &\sum_{i=0}^{2^{k+1}-1} \operatorname{Var}\Big(T_{N,1}^{(1)}(y_i, y_{i+1})\Big) \\ &\quad \le 2 \sum_{i=0}^{2^{k+1}-1} \sum_{n=1}^{N} \sum_{n'=n}^{N} \sum_{j=1}^{\infty} \operatorname{Cov}\Big(R_{n,j}(y_i, y_{i+1}), R_{n',j'}(y_i, y_{i+1})\Big). \end{aligned} \tag{6.16}$$

We first explain the main idea in handling the covariance. For the moment, let us focus on the case $j \ge 2$. By the Taylor expansion, taking into account that $\int u\, dG(u) = 0$,

$$\begin{aligned} &F_{j-1}\Big(x - \tilde{X}_{n,j-1}\Big) - F_j\Big(x - \tilde{X}_{n,j}\Big) \\ &\quad = \int \Big(F_{j-1}\Big(x - \tilde{X}_{n,j-1}\Big) - F_{j-1}\Big(x - \tilde{X}_{n,j-1} + a_j(\varepsilon_{n-j} - u)\Big)\Big)\, dG(u) \\ &\quad = -a_j \varepsilon_{n-j} F_{j-1}^{(1)}\Big(x - \tilde{X}_{n,j-1}\Big) \\ &\qquad + \frac{a_j^2}{2} \int (\varepsilon_{n-j} - u)^2 F_{j-1}^{(2)}\Big(x - \tilde{X}_{n,j-1} + \delta(u)\Big)\, dG(u), \end{aligned}$$

where $|\delta(u)| \le |a_j(\varepsilon_{n-j} - u)|$. Another Taylor expansion shows that

$$F_{j-1}^{(1)}\big(x - \tilde{X}_{n,j-1}\big) - F_{j-1}^{(1)}\big(x - \tilde{X}_{n,j}\big) = -a_j\varepsilon_{n-j}F_{j-1}^{(2)}\big(x - \tilde{X}_{n,j-1} + \eta\big),$$

where $|\eta| \le |a_j\varepsilon_{n-j}|$, and hence

$$\begin{aligned} R_{n,j}(x) &= F_{j-1}\big(x - \tilde{X}_{n,j-1}\big) - F_j\big(x - \tilde{X}_{n,j}\big) + a_j\varepsilon_{n-j}F_{j-1}^{(1)}\big(x - \tilde{X}_{n,j}\big) \\ &= a_j^2\varepsilon_{n-j}^2F_{j-1}^{(2)}\big(x - \tilde{X}_{n,j-1} + \eta\big) \\ &\quad + \frac{a_j^2}{2}\int(\varepsilon_{n-j} - u)^2F_{j-1}^{(2)}\big(x - \tilde{X}_{n,j-1} + \delta(u)\big)\,dG(u). \end{aligned} \tag{6.17}$$

Using essentially this idea, we obtain, for $j \ge 2$ and $n' \ge n$,

$$\begin{aligned} &\sum_{i=0}^{2^{k+1}-1}\operatorname{Cov}\big(R_{n,j}(y_i, y_{i+1}), R_{n',j'}(y_i, y_{i+1})\big) \\ &= \sum_{i=0}^{2^{k+1}-1}E\Bigg[a_j^2\varepsilon_{n-j}^2F_{j-1}^{(2)}\big(y_i - \tilde{X}_{n,j-1} + \eta_{n,j}, y_{i+1} - \tilde{X}_{n,j-1} + \eta_{n,j}\big) \\ &\qquad + \frac{a_j^2}{2}\int(\varepsilon_{n-j} - u)^2F_{j-1}^{(2)}\big(y_i - \tilde{X}_{n,j-1} + \delta_{n,j}(u), \\ &\qquad\qquad y_{i+1} - \tilde{X}_{n,j-1} + \delta_{n,j}(u)\big)\,dG(u)\Bigg] \\ &\quad\times\Bigg[a_{j'}^2\varepsilon_{n-j}^2F_{j'-1}^{(2)}\big(y_i - \tilde{X}_{n',j'-1} + \eta'_{n',j'}, y_{i+1} - \tilde{X}_{n',j'-1} + \eta'_{n',j'}\big) \\ &\qquad + \frac{a_{j'}^2}{2}\int(\varepsilon_{n-j} - u')^2F_{j'-1}^{(2)}\big(y_i - \tilde{X}_{n',j'-1} + \delta'_{n',j'}(u'), \\ &\qquad\qquad y_{i+1} - \tilde{X}_{n',j'-1} + \delta'_{n',j'}(u')\big)\,dG(u')\Bigg], \end{aligned}$$

where $|\eta_{n,j}| \le |a_j\varepsilon_{n-j}|$, $|\eta'_{n',j'}| \le |a_{j'}\varepsilon_{n-j}|$, $|\delta_{n,j}(u)| \le |a_j(\varepsilon_{n-j} - u)|$ and $|\delta'_{n',j'}(u)| \le |a_{j'}(\varepsilon_{n-j} - u)|$. It is worth noting that we applied the Taylor expansion to the whole sum, *not* to the summands individually, so that $\eta_{n,j}, \eta'_{n',j'}, \delta_{n,j}, \delta'_{n',j'}$ are free of i. This is a crucial point in this approach. Recall that $\int u^4\,dG(u) < \infty$, $F_{j-1}^{(2)}$ is bounded and

$$\sup_{y\in\Re}\sum_{i=0}^{2^{k+1}-1}\big|F_{j-1}^{(2)}(y_i + y, y_{i+1} + y)\big| \le \int|F_{j-1}^{(3)}| < \infty.$$

Therefore, we conclude that, for $j \ge 2$ and $n' \ge n$ (i.e., $j' \ge j \ge 2$),

$$\left|\sum_{i=0}^{2^{k+1}-1}\operatorname{Cov}\big(R_{n,j}(y_i, y_{i+1}), R_{n',j'}(y_i, y_{i+1})\big)\right| \le Ba_j^2a_{j'}^2,$$

where B is a universal constant. Now, for the case $j = 1$ and $n' = n$ (i.e., $j' = j = 1$), we use the boundedness and monotonicity of the distribution function to conclude

$$\left| \sum_{i=0}^{2^{k+1}-1} \operatorname{Cov}\left(R_{n,1}(y_i, y_{i+1}), R_{n,1}(y_i, y_{i+1})\right) \right| \le 4,$$

and, similarly, for the case $j = 1$ and $n' \ge n + 1$ (i.e., $j' > j = 1$), using (6.17), we obtain

$$\left| \sum_{i=0}^{2^{k+1}-1} \operatorname{Cov}\left(R_{n,1}(y_i, y_{i+1}), R_{n',j'}(y_i, y_{i+1})\right) \right| \le B a_{j'}^2.$$

Combining these three cases and making use of the fact that $\beta > 1/2$, we conclude that

$$\left| \sum_{i=0}^{2^{k+1}-1} \sum_{n=1}^{N} \sum_{n'=n}^{N} \sum_{j=1}^{\infty} \operatorname{Cov}\left(R_{n,j}(y_i, y_{i+1}), R_{n',j'}(y_i, y_{i+1})\right) \right| \le BN.$$

Thus, (6.15) with $t = 1$ follows from (6.16).

Suppose now (6.15) holds for t equals some τ satisfying $1 \le \tau < p$. We show that (6.15) holds for $t = \tau + 1$, and then mathematical induction gives (6.12). Define

$$\begin{aligned} Z_N(x) &= T_{N,1}^{(\tau+1)}(x) - T_{N,1}^{(\tau)}(x) \\ &= (-1)^{\tau} \sum_{n=1}^{N} \sum_{2 \le j_1 < \cdots < j_\tau} \left(\prod_{s=1}^{\tau} a_{j_s} \varepsilon_{n-j_s} \right) \left(F_{j_\tau - 1}^{(\tau)}\left(x - \tilde{X}_{n,j_\tau}\right) - F^{(\tau)}(x) \right) \\ &\quad + (-1)^{\tau} \sum_{n=1}^{N} \sum_{2 \le j_1 < \cdots < j_{\tau+1}} \left(\prod_{s=1}^{\tau+1} a_{j_s} \varepsilon_{n-j_s} \right) F_{j_{\tau+1}-1}^{(\tau+1)}\left(x - \tilde{X}_{n,j_{\tau+1}}\right). \end{aligned}$$

By the induction assumption and the Cauchy–Schwarz inequality, it suffices to show

$$\sum_{i=0}^{2^{k+1}-1} \operatorname{Var}\left(Z_n(y_i, y_{i+1})\right) \le BN. \tag{6.18}$$

Using arguments similar to those in Lemma 6.3, we can write

$$Z_N(x) = (-1)^{\tau} \sum_{n=1}^{N} \sum_{2 \le j_1 < \cdots < j_{\tau+1}} \left(\prod_{s=1}^{\tau} a_{j_R} \varepsilon_{n-j_s} \right) R_{n,j_{\tau+1}}(x),$$

where, in this instance,

$$R_{n,j}(x) = F_{j-2}^{(\tau)}\left(x - \tilde{X}_{n,j-1}\right) - F_{j-1}^{(\tau)}\left(x - \tilde{X}_{n,j}\right) + a_j \varepsilon_{n-j} F_{j-1}^{(\tau+1)}\left(x - \tilde{X}_{n,j}\right).$$

Using arguments similar to those in Lemma 6.4, it is straightforward to show that

$$\sum_{i=0}^{2^{k+1}-1} \mathrm{Var}(Z_N(y_i, y_{i+1})) \le 2 \sum_{i=0}^{2^{k+1}-1} \sum_{n=1}^{N} \sum_{n'=n}^{N} \sum_{2 \le j_1 < \cdots < j_{\tau+1}} \left(\prod_{s=1}^{\tau} a_{j_s} a_{j'_s} \right) \times E\left(R_{n, j_{\tau+1}}(y_i, y_{i+1}) R_{n', j'_{\tau+1}}(y_i, y_{i+1})\right),$$

with $j'_s = n' - n + j_s$ [cf. (6.16)] where, again, the Taylor expansion gives

$$\begin{aligned} R_{n,j}(x) &= -a_j \varepsilon_{n-j}\left(F_{j-2}^{(\tau+1)}\left(x - \tilde{X}_{n,j-1}\right) - F_{j-1}^{(\tau+1)}\left(x - \tilde{X}_{n,j}\right)\right) \\ &\quad + \int \frac{(a_j \varepsilon_{n-j} - a_{j-1} u)^2}{2} F_{j-2}^{(\tau+2)}\left(x - \tilde{X}_{n,j-1} + \delta(u)\right) dG(u) \\ &= (a_j \varepsilon_{n-j})^2 F_{j-2}^{(\tau+2)}\left(x - \tilde{X}_{n,j-1}\right) \\ &\quad + \int \frac{(a_j \varepsilon_{n-j} - a_{j-1} u)^2}{2} F_{j-2}^{(\tau+2)}\left(x - \tilde{X}_{n,j-1} + \delta(u)\right) dG(u) \\ &\quad - a_j \varepsilon_{n-j} \int \frac{(a_j \varepsilon_{n-j} - a_{j-1} u)^2}{2} F_{j-2}^{(\tau+3)}\left(x - \tilde{X}_{n,j-1} + \eta(u)\right) dG(u), \end{aligned}$$

with $|\delta(u)| \le |a_j \varepsilon_{n-j} - a_{j-1} u|$ and $|\eta(u)| \le |a_j \varepsilon_{n-j} - a_{j-1} u|$. Using this plus the assumptions $\sup_{j,x} |F_j^{(\tau+2)}(x)| < \infty$, $\sup_{j,x} |F_j^{(\tau+3)}(x)| < \infty$, $\sup_j \int |F_j^{(\tau+3)}| < \infty$ and $\sup_j \int |F_j^{(\tau+4)}| < \infty$, we obtain, as in the previous step of the induction,

$$\left| \sum_{i=0}^{2^{k+1}-1} E\left(R_{n, j_{\tau+1}}(y_i, y_{i+1}) R_{n', j'_{\tau+1}}(y_i, y_{i+1})\right) \right| \le B a_{j_{\tau+1}}^2 a_{j'_{\tau+1}}^2,$$

and therefore

$$\sum_{i=0}^{2^{k+1}-1} \mathrm{Var}(Z_N(y_i, y_{i+1})) \le B \sum_{n=1}^{N} \sum_{n'=n}^{N} \sum_{2 \le j_1 < \cdots < j_\tau} \left(\prod_{s=1}^{\tau} |a_{j_s} a_{j'_s}| \right) \sum_{j_{\tau+1} > j_\tau} a_{j_{\tau+1}}^2 a_{j'_{\tau+1}}^2, \tag{6.19}$$

where B is a universal constant. For any $\zeta > 0$, there exists $C < \infty$ such that

$$|a_j| \le C_j^{-\beta+\zeta}.$$

Choose $0 < \zeta < \beta - 1/2$. Simple algebra shows that the right-hand side of (6.19) is bounded by

$$CN \sum_{n=0}^{N-1} \sum_{2 \le j_1 < \cdots < j_\tau} \left(\prod_{s=1}^{\tau} [j_s(n+j_s)]^{-\beta+\zeta} \right) \sum_{j_{\tau+1} > j_\tau} [j_{\tau+1}(n+j_{\tau+1})]^{-2\beta+\zeta}$$

for some $C < \infty$. By Lemma 6.5, for $n \ge 1$,

$$\begin{aligned} &\sum_{2 \le j_1 < j_2 < \cdots < j_\tau} \left(\prod_{s=1}^{\tau} [j_s(n+j_s)]^{-\beta+\zeta} \right) \sum_{j_{\tau+1} > j_\tau} [j_{\tau+1}(n+j_{\tau+1})]^{-2\beta+\zeta} \\ &\quad \le C \sum_{j=2}^{\infty} [j(n+j)]^{-[(\tau+1)(\beta-\zeta-1/2)+\beta+1/2]} \\ &\quad \le C \sum_{j=2}^{\infty} [j(n+j)]^{-(\beta+1/2)} \le Cn^{-(\beta+1/2)}, \end{aligned}$$

which is summable. It is then clear that (6.18) holds. We have shown (6.15) for $t = \tau + 1$ and this concludes the proof of (6.12).

To show (6.13) for $p > 1$, simply note that

$$\operatorname{Var}(T_{N,3}(y_i, y_{i+1})) = \sum_{r=1}^{p} \left(F^{(r)}(y_i, y_{i+1})\right)^2 N a_1^2 E\varepsilon_n^2 \sum_{1 < j_2 < \cdots < j_r} \prod_{s=2}^{r} a_{j_s}^2$$

from which (6.13) is immediate.

Finally, we show (6.14). As before,

$$\operatorname{Cov}\left(\left(\prod_{s=1}^{p} a_{j_s} \varepsilon_{n-j_s} \right) R_{n,j_p}(x), \left(\prod_{s=1}^{p} a_{j'_s} \varepsilon_{n'-j'_s} \right) R_{n',j'_p}(x) \right) = 0$$

if $n - j_s \ne n' - j'_s$ for some $1 \le s \le p$, where now

$$R_{n,j}(x) = F^{(p)}(x) - F_{j-1}^{(p)}\left(x - \tilde{X}_{n,j}\right).$$

Thus,

$$\begin{aligned} &\sum_{i=0}^{2^{k+1}-1} \operatorname{Var}(T_{N,2}(y_i, y_{i+1})) \\ &\quad \le 2 \sum_{i=0}^{2^{k+1}-1} \sum_{n=1}^{N} \sum_{n'=n}^{N} \sum_{2 \le j_1 < \cdots < j_p} \left(\prod_{s=1}^{p} a_{j_s} a_{j'_s} \right) \\ &\qquad\qquad \times \operatorname{Cov}\left(R_{n,j_p}(y_i, y_{i+1}), R_{n',j'_p}(y_i, y_{i+1})\right), \end{aligned} \tag{6.20}$$

where $j'_s = n' - n + j_s$ [cf. (6.16)]. Note that

$$\begin{aligned} R_{n,j}(x) &= \int \left(F_{j-1}^{(p)}(x-u) - F_{j-1}^{(p)}\left(x - \tilde{X}_{n,j}\right) \right) d\tilde{F}_{j-1}(u) \\ &= \int \left(\tilde{X}_{n,j} - u \right) F_{j-1}^{(p+1)}(x - \delta(u))\, d\tilde{F}_{j-1}(u), \end{aligned}$$

where $\delta(u)$ is between u and $\tilde{X}_{n,j}$. Using this, the assumption $\sup_j \int |F_j^{(p+2)}| < \infty$ and an argument used twice already, we obtain

$$\left| \sum_{i=0}^{2^{k+1}-1} \operatorname{Cov}\Big(R_{n,j_p}(y_i, y_{i+1}), R_{n',j'_p}(y_i, y_{i+1})\Big) \right| \tag{6.21}$$
$$\leq C\Big(E|\tilde{X}_{n,j_p}\tilde{X}_{n',j'_p}| + E|\tilde{X}_{n,j_p}|E|\tilde{X}_{n',j'_p-1}| + E|\tilde{X}_{n,j_p-1}|E|\tilde{X}_{n',j'_p}| \Big)$$

for some $C < \infty$. By the Cauchy–Schwarz inequality, for any $0 < \zeta < \beta - 1/2$, there exists C such that, for all $j, j' \geq 2$,

$$E|\tilde{X}_{n,j}|E|\tilde{X}_{n',j'}| \vee E|\tilde{X}_{n,j}\tilde{X}_{n',j'}| \tag{6.22}$$
$$\leq \Big(E\tilde{X}^2_{n,j}\Big)^{1/2}\Big(E\tilde{X}^2_{n',j'}\Big)^{1/2} \leq C(jj')^{-(\beta-1/2)+\zeta}.$$

By (6.20), (6.21) and (6.22),

$$\sum_{i=0}^{2^{k+1}-1} \operatorname{Var}(T_{N,2}(y_i, y_{i+1}))$$
$$\leq C \sum_{n=1}^{N} \sum_{n'=n}^{N} \sum_{2 \leq j_1 < \cdots < j_p} \left(\prod_{s=1}^{p} |a_{j_s} a_{j'_s}| \right) (j_p j'_p)^{-(\beta-1/2)+\zeta}$$
$$\leq CN \sum_{n=0}^{N-1} \sum_{2 \leq j_1 < \cdots < j_p} \left(\prod_{s=1}^{p} [j_s(n+j_s)]^{-\beta+\zeta} \right) \big[j_p(n+j_p)\big]^{-(\beta-1/2)+\zeta}.$$

By Lemma 6.5, for $n \geq 1$,

$$\sum_{2 \leq j_1 < \cdots < j_p} \left(\prod_{s=1}^{p} [j_s(n+j_s)]^{-\beta+\zeta} \right) \big[j_p(n+j_p)\big]^{-(\beta-1/2)+\zeta}$$
$$\leq C \sum_{j=2}^{\infty} [j(n+j)]^{-[(p+1)(\beta-\zeta-1/2)+1/2]}.$$

Note that there are two possibilities: for small $\zeta > 0$, $(p+1)(\beta-\zeta-1/2)+1/2$ is greater than 1 or less than 1, where, by Lemma 6.5, in the first case the right-hand side of the preceding inequality is summable in n and in the second case it is bounded by $Cn^{-2(p+1)(\beta-\zeta-1/2)}$. Thus, for some $C < \infty$,

$$\sum_{i=0}^{2^{k+1}-1} \operatorname{Var}(T_{N,2}(y_i, y_{i+1})) \leq C\big(N \vee N^{2-2(p+1)(\beta-\zeta-1/2)}\big).$$

This shows (6.14) and completes the proof. □

PROOF OF THEOREM 2.2. Define

$$M(N) = \sup_{x \in \Re} |S_N(x)| \quad \text{and} \quad M(N_1, N_2) = M(N_2) - M(N_1).$$

Clearly, for $N_1 < N_2$,

$$|M(N_1, N_2)| \le \sup_{x \in \Re} |S_{N_2}(x) - S_{N_1}(x)|$$
$$=_d \sup_{x \in \Re} |S_{N_2 - N_1}(x)| = M(N_2 - N_1)$$

by stationarity. Fix $\lambda < \gamma(\beta, p)/2$. Set

$$N_k = 2^k, \qquad k = 1, 2, \dots .$$

By Theorem 2.1 and the Borel–Cantelli lemma, it is easy to show that

$$\frac{N_k^{\lambda}}{\sigma_{N_k}} M(N_k) \to 0 \quad \text{a.s. as } k \to \infty. \tag{6.23}$$

By (6.23) and stationarity, it suffices to show that

$$\frac{N_k^{\lambda}}{\sigma_{N_k}} \max_{0 \le N < 2^k} |M(N)| \to 0 \quad \text{a.s. as } k \to \infty.$$

Using the method of dyadic expansion [cf. Dehling and Taqqu (1989), the proof of Theorem 3.1] we get, for each N, $0 \le N < 2^k$,

$$P\left\{ \frac{N_k^{\lambda}}{\sigma_{N_k}} \max_{0 \le N < 2^k} |M(N)| > b \right\} \le \sum_{i=0}^{k-1} 2^{k-i-1} P\left\{ \frac{N_k^{\lambda}}{\sigma_{N_k}} |M(2^i)| > \frac{b}{k} \right\}.$$

Fix $b > 0$ and pick $\zeta, \zeta' > 0$ to satisfy

$$\gamma(\beta, p) - 2\lambda - \zeta(\lambda + 1) - 2\zeta' > 0, \tag{6.24}$$

which is possible since $\gamma(\beta, p) > 2\lambda$. By Theorem 2.1, for the ζ we chose there exists $C < \infty$ such that

$$P\left\{ \frac{N_k^{\lambda}}{\sigma_{N_k}} |M(2^i)| > \frac{b}{k} \right\} = P\left\{ \sigma_{2^i}^{-1} |M(2^i)| > \frac{b\sigma_{2^k}}{k 2^{k\lambda} \sigma_{2^i}} \right\}$$
$$\le C \frac{\sigma_{2^i}^2}{\sigma_{2^k}^2} k^{2+\zeta} 2^{k\lambda(2+\zeta)} 2^{-i(\gamma(\beta, p) - \zeta)}.$$

Since $\sigma_N^2 \in \mathrm{RV}_H$, by Lemma 2 on page 277 of Feller (1971), there exists $C < \infty$ such that the rightmost term of the previous inequality is bounded by

$$Ck^{2+\zeta} 2^{i(H+\zeta')} 2^{-k(H-\zeta')} 2^{k\lambda(2+\zeta)} 2^{-i(\gamma(\beta, p) - \zeta)}$$
$$= Ck^{2+\zeta} 2^{-k[H - \lambda(2+\zeta) - \zeta']} 2^{i(H - \gamma(\beta, p) + \zeta + \zeta')}.$$

Hence,

$$\sum_{i=0}^{k-1} 2^{k-i-1} P\left\{ \frac{N_k^{\lambda}}{\sigma_{N_k}} |M(2^i)| > \frac{b}{k} \right\}$$
$$\le Ck^{2+\zeta} 2^{-k[H - \lambda(2+\zeta) - \zeta' - 1]} 2^{k(H - \gamma(\beta, p) - 1 + \zeta + \zeta')}$$
$$= Ck^{2+\zeta} 2^{-k[\gamma(\beta, p) - 2\lambda - \zeta(\lambda+1) - 2\zeta']}.$$

By (6.24), the last expression is summable (in k), and hence the result follows from the Borel–Cantelli lemma. □

This completes the proofs of Theorems 2.1 and 2.2. We proceed to prove Theorems 3.1 and 3.2.

PROOF OF THEOREM 3.1. Write

$$\hat{f}_N(x) - E\hat{f}_N(x) = \frac{1}{h}\int K\left(\frac{x-u}{h}\right) d[F_N(u) - F(u)],$$

$$E\hat{f}_N(x) - f(x) = \int [f(x-hu) - f(x)] K(u)\, du.$$

Then integration by parts gives

$$\begin{aligned}\sup_{x\in\Re}\left|\hat{f}_N(x) - f(x)\right| \\ = \sup_{x\in\Re}\bigg|\frac{Y_{N,1}}{N}\int f'(x-hu)K(u)\,du \\ + \frac{1}{Nh}\int S_{N,1}(x-hu)\,dK(u) + h\int f'(u^*)uK(u)\,du\bigg|,\end{aligned} \tag{6.25}$$

where $Y_{N,1}/N$ is the sample mean $\tilde{X}_N$, and $|u^* - x| < |hu|$. By Lai and Stout (1980), Theorem 7,

$$N^{\lambda}\overline{X}_N \to 0 \quad \text{a.s. for } \lambda < \beta - \tfrac{1}{2}.$$

Since $\sigma_{N,1} \in \mathrm{RV}_{3/2-\beta}$, it follows from Theorem 2.2 that

$$\sup_{x\in\Re}\frac{N^{\lambda}}{Nh}\left|\int S_{N,1}(x-hu)\,dK(u)\right| \to 0 \quad \text{a.s. for } \lambda < \frac{\gamma(\beta,1)}{2} + \beta - \frac{1}{2} - \delta;$$

(3.4) follows using these and (6.25). Rewrite (6.25) as

$$\begin{aligned}\frac{N}{\sigma_{N,1}}\sup_{x\in\Re}\left|\hat{f}_N(x) - f(x)\right| \\ = \sup_{x\in\Re}\bigg|\frac{Y_{N,1}}{\sigma_{N,1}}\int f'(x-hu)K(u)\,du + \frac{1}{\sigma_{N,1}h}\int S_{N,1}(x-hu)\,dK(u) \\ - \frac{Nh^2}{2\sigma_{N,1}}\int f''(u^*)u^2K(u)\,du\bigg|,\end{aligned}$$

where $|u^* - x| \le |hu|$. Clearly (3.6) holds. The proof is complete. □

PROOF OF THEOREM 3.2. The same arguments leading to (6.25) and repeated integration by parts give

$$\begin{aligned}
&\sup_{x\in\Re}\left|\hat f_N^{(q)}(x)-f^{(q)}(x)\right| \\
&\quad=\sup_{x\in\Re}\left|\frac{Y_{N,1}}{N}\int f^{(q+1)}(x-hu)K(u)\,du\right. \\
&\qquad+\frac{1}{Nh^{q+1}}\int S_{N,1}(x-hu)\,dK^{(q)}(u) \\
&\qquad\left.+h\int f^{(q+1)}(u^*)uK(u)\,du\right|
\end{aligned}\tag{6.26}$$

for some u^* satisfying $|u^*-x|<|hu|$. Thus, (3.7) follows as (3.4). Now rewrite (6.26) as

$$\begin{aligned}
&\frac{N}{\sigma_{N,1}}\sup_{x\in\Re}\left|\hat f_N^{(q)}(x)-f^{(q)}(x)\right| \\
&\quad=\sup_{x\in\Re}\left|\frac{Y_{N,1}}{\sigma_{N,1}}\int f^{(q)}(x-hu)K(u)\,du\right. \\
&\qquad+\frac{1}{\sigma_{N,1}h^{q+1}}\int S_{N,1}(x-hu)\,dK^{(q)}(u) \\
&\qquad\left.+\frac{N\cdot(-h)^l}{l!\,\sigma_{N,1}}\int f^{(q+l)}(u^*)u^lK(u)\,du\right|,
\end{aligned}\tag{6.27}$$

where $|u^*-x|<|hu|$. Note that (3.8) ensures that the inequality (3.9) is not vacant, and then (3.10) follows from verifying that the second and third term of the right-hand side of (6.27) tends to 0 a.s. □

Next we prove Theorem 4.1.

PROOF OF THEOREM 4.1. The proof is a simplified version of those of Theorems 2.1 and 2.2. For a fixed t, let $K(x)=\cos(tx)$ [or $\sin(tx)$]. Recall that $K_\infty(x)=EK(x+X_n)$ and $F_j(x)=P\{X_{n,j}\le x\}$. Consider the quantity

$$R_N=\sum_{n=1}^N\left(K(X_n)-EK(X_n)\right)-K_\infty'(0)\sum_{n=1}^N Y_{N,1}.$$

Define

$$K_j(x)=EK(x+X_{n,j})=\int K(x+y)\,dF_j(y),\qquad j=0,1,\dots.$$

Thus,

$$R_N = \sum_{n=1}^{N} \sum_{j=1}^{\infty} \Big(K_{j-1}\big(\tilde{X}_{n,j-1}\big) - K_j\big(\tilde{X}_{n,j}\big) - a_j \varepsilon_{n-j} K'_\infty(0) \Big)$$
$$=: \sum_{n=1}^{N} \sum_{j=1}^{\infty} R_{n,j},$$

where the sum converges both a.s. and in L_2. Clearly,

$$\operatorname{Var}(R_N) \le 2 \sum_{n=1}^{N} \sum_{n'=n}^{N} \sum_{j=1}^{\infty} \operatorname{Cov}(R_{n,j}, R_{n',j'}),$$

where $j' = n' - n + j$. By the Taylor expansion,

$$R_{n,j} = \frac{a_j^2}{2} \int (\varepsilon_{n-j} - u)^2 K''_{j-1}\big(\tilde{X}_{n,j-1} + \delta(u)\big)\, dG(u)$$
$$+ a_j \varepsilon_{n-j} \Big(K'_{j-1}\big(\tilde{X}_{n,j-1}\big) - K'_\infty(0) \Big),$$

where $|\delta(u)| \le |a_j(\varepsilon_{n-j} - u)|$. Another Taylor expansion gives

$$\Big| K'_{j-1}\big(\tilde{X}_{n,j-1}\big) - K'_\infty(0) \Big| \le C\Big(|\tilde{X}_{n,j-1}| + E|\tilde{X}_{n,j-1}| \Big).$$

Thus, by the Cauchy–Schwarz inequality, for any $\zeta > 0$ there exists $C < \infty$ such that

$$\sum_{n=1}^{N} \sum_{n'=n}^{N} \sum_{j=1}^{\infty} \operatorname{Cov}(R_{n,j}, R_{n',j'}) \le C \sum_{n=1}^{N} \sum_{n'=n}^{N} \sum_{j=1}^{\infty} \Big(a_j^2 a_{j'}^2 + |a_j a_{j'}| (jj')^{-(\beta - 1/2) + \zeta} \Big)$$
$$\le C\big(N \vee N^{2 - 2(2\beta - 1 - 2\zeta)} \big)$$

[cf. (6.14) with $p = 1$]. Hence, for any $\zeta > 0$ there exists $C < \infty$ such that

$$\operatorname{Var}\big(\sigma_{N,1}^{-1} R_N \big) \le C N^{-\gamma(\beta,1) + \zeta}.$$

The same proof of Theorem 2.2 then shows that

$$\frac{N^\lambda}{\sigma_{N,1}} \left(\sum_{n=1}^{N} (K(X_n) - NEK(X_n)) - K'_\infty(0) Y_{N,1} \right) \to 0 \quad \text{a.s. } \forall \lambda < \frac{\gamma(\beta, 1)}{2}.$$

Since

$$\phi_N(x) = \sum_{n=1}^{N} \cos(tX_n) + i \sum_{n=1}^{N} \sin(tX_n)$$

and

$$K'_\infty(0) = \begin{cases} -tE\big(\sin(tX_n)\big), & \text{if } K(x) = \cos(tx), \\ tE\big(\cos(tX_n)\big), & \text{if } K(x) = \sin(tx), \end{cases}$$

(4.2) follows. Clearly, (4.3) follows from (4.2) and Remark 2.2. □

We next prove Theorem 5.1. We need some notation and a few lemmas. Set

$$\|F_N - F\|_u^v = \sup_{u<x<v} |F_N(x) - F(x)|$$

and

$$\|Q_N - Q\|_a^b = \sup_{a<x<b} |Q_N(y) - Q(y)|.$$

LEMMA 6.7. *Assume the conditions of Theorem* 2.1 *with* $p = 1$. *Then, for each* $\delta > 0$,

$$\frac{N\|F_N - F\|_{-\infty}^{\infty}}{\sigma_{N,1}(\log)^{(1+\delta)/2}} \to 0 \quad a.s.$$

PROOF. The proof follows from a direct application of Theorem 2.2 and Lai and Stout (1980), Theorem 7. □

For any fixed $0 < a < b < 1$, define $f_{a,b} = \inf_{Q(a)<x<Q(b)} f(x)$. Also, define

$$Z_n = F(X_n),$$

$$E_N(y) = N^{-1} \sum_{n=1}^{N} I(Z_n \le y), \qquad 0 \le y \le 1,$$

$$E_N^{-1}(y) = \inf\{u: E_N(u) \ge y\} = \begin{cases} 0, & y = 0, \\ F(X_{N:k}), & \dfrac{k-1}{N} < y \le \dfrac{k}{N}, 1 \le k \le N. \end{cases}$$

LEMMA 6.8. *Assume that* $f_{a,b} > 0$. *Then*

$$\|Q_N - Q\|_a^b \le f_{a,b}^{-1} \|F_N - F\|_{Q(a)}^{Q(b)}.$$

PROOF. For all $(k-1)/N < y \le k/N$,

$$\begin{aligned} Q_N(y) - Q(y) &= X_{N:k} - F^{-1}(y) \\ &= F^{-1}(F(X_{N:k})) - F^{-1}(y) = F^{-1}(E_N^{-1}(y)) - F^{-1}(y). \end{aligned}$$

Hence, by the Taylor expansion,

$$\|Q_N - Q\|_a^b \le f_{a,b}^{-1} \sup_{a<y<b} |E_N^{-1}(y) - y| = f_{a,b}^{-1} \|F_N - F\|_{Q(a)}^{Q(b)}. \qquad \square$$

Define

$$R_n(y) = \{F_N(Q_N(y)) - F_N(Q(y))\} - \{F(Q_N(y)) - F(Q(y))\},$$

$$d_N = N^{-1}\sigma_{N,1}(\log N)^{(1+\delta)/2},$$

and the interval

$$I_N(y) = [Q(y) - d_N, Q(y) + d_N].$$

LEMMA 6.9. *With probability* 1,

$$\sup_{a<y<b} R_N(y) = o\left(d_N^2 \vee N^{-(1+\lambda)}\sigma_{N,1} d_N^{1/2}\right) \quad \forall\, 0 < \lambda < \left(\beta - \tfrac{1}{2}\right) \wedge (1-\beta).$$

PROOF. By Lemmas 6.7 and 6.8, with probability 1 for all large N,

$$|R_N(y)| \le \sup_{x \in I_N(y)} \left|\{F_N(x) - F_N(Q(y))\} - \{F(x) - F(Q(y))\}\right|,$$

which is equal to

$$\sup_{x \in I_N(y)} \left|A_N(x, Q(y)) + B_N(x, Q(y))\right|,$$

where

$$A_N(x, z) = (f(z) - f(x)) Y_{N,1} N^{-1}$$

and

$$B_N(x, z) = (S_{N,1}(x) - S_{N,1}(z)) N^{-1}.$$

Fix $-\infty < u < v < \infty$. By Lemma 6.7, we conclude readily

$$\sup_{u<x<v} \sup_{|\delta|<d_N} A_N(x, x+\delta) = o\left(d_N^2\right) \quad \text{a.s.}$$

For each $\delta > 0$, a slightly modified chain argument in the proof of Theorem 2.1 shows that for any $\zeta > 0$ there exists $C < \infty$ such that

$$P\left\{\sup_{u<x<v} B_N(x, x+\delta) > b\right\} \le C b^{-2}(1 \vee b^{-\zeta})\, \delta N^{-\gamma(\beta,p)+\zeta} \qquad \forall\, b, \delta > 0.$$

Another chain argument on δ then gives

$$P\left\{\sup_{u<x<v} \sup_{|\delta|<d_N} B_N(x, x+\delta) > b\right\} \le C b^{-2}(1 \vee b^{-\zeta})\, d_N N^{-\gamma(\beta,p)+\zeta} \quad \forall\, b > 0.$$

Finally, an argument similar to the proof of Theorem 2.2 gives

$$\sup_{u<x<v} \sup_{|\delta|<d_N} B_N(x, x+\delta) = o\left(d_N^{1/2} N^{(1+\lambda)} \sigma_{N,1}\right) \quad \text{a.s.}$$

The details of these derivations do not contain new ideas and are therefore omitted. □

We are now ready to prove Theorem 5.1.

PROOF OF THEOREM 5.1. Note that

$$\sup_{0<y<1} \left|F_N(Q_N(y)) - y\right| = N^{-1}.$$

Applying a Taylor expansion on R_N, we have, with probability 1 for all $y \in (a, b)$,

$$\begin{aligned} Q_N(y) - Q(y) &= \frac{y - F_N(Q(y))}{f(Q(y))} + O\big(R_N(y) + N^{-1}\big) + O\Big(\big(Q_N(y) - Q(y)\big)^2\Big) \\ &= \frac{Y_{N,1}}{N} + O\big(S_{N,1}(Q(y))N^{-1} + R_N(y) + N^{-1}\big) \\ &\quad + O\Big(\big(Q_N(y) - Q(y)\big)^2\Big), \end{aligned}$$

where we used the fact that $f(Q(y))$ is bounded away from 0. By Theorem 2.2 and Lemma 6.9,

$$\frac{N^{1+\lambda}}{\sigma_{N,1}} \sup_{a<y<b} \big|S_{N,1}(Q(y))N^{-1} + R_N(y) + N^{-1}\big| \to 0 \quad \text{a.s.}$$

This completes the proof. □

Acknowledgments. The authors thank the referees for their comments, and in particular for pointing out an error previously in Lemma 6.5. T. Hsing is pleased to acknowledge the hospitality of Chii-Ruey Hwang of the Academia Sinica during his stay in Taipei during 1993 and 1994.

REFERENCES

AVRAM, F. and TAQQU, M. S. (1987). Noncentral limit theorems and Appell polynomials. *Ann. Probab.* **15** 767–775.

BAHADUR, R. R. (1966). A note on quantiles in large samples. *Ann. Math. Statist.* **37** 577–580.

BERAN, J. (1991). *M*-estimator of location for data with slowly decaying serial correlations. *J. Amer. Statist. Assoc.* **86** 704–708.

BERAN, J. (1992). Statistical methods for data with long-range dependence. *Statist. Sci.* **7** 404–420.

BERAN, J. and GHOSH, S. (1991). Slowly decaying correlation, testing normality, nuisance parameters. *J. Amer. Statist. Assoc.* **86** 785–791.

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

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

BROCKWELL, P. J. and DAVIS, R. A. (1987). *Time Series: Theory and Methods*. Springer, New York.

CHANDA, K. C. and RUYMGAART, F. H. (1990). General linear processes: a property of the empirical process applied to density and mode estimation. *J. Time Ser. Anal.* **11** 185–199.

DAVYDOV, Y. A. (1970). The invariance principle for stationary processes. *Theory Probab. Appl.* **15** 487–498.

DEHLING, H. and TAQQU, M. S. (1989). The empirical processes of some long-range dependent sequences with an application to *U*-statistics. *Ann. Statist.* **17** 1767–1783.

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

FELLER, W. (1971). *An Introduction to Probability Theory and Its Applications* **2**, 2nd ed. Wiley, New York.

GRANGER, C. W. and JOYEUX, R. (1980). An introduction to long-memory time series models and fractional differencing. *J. Time Ser. Anal.* **1** 15–29.

HALL, P. and HART, J. D. (1990). Convergence rates in density estimation for data from infinite-order moving average processes. *Probab. Theory Related Fields* **87** 253–274.

HESSE, C. H. (1990a). Rates of convergence for the empirical distribution function and the empirical characteristic function of a broad class of linear processes. *J. Multivariate Anal.* **35** 186–202.

HESSE, C. H. (1990b). A Bahadur-type representation for empirical quantiles of a large class of stationary, possibly infinite-variance, linear processes. *Ann. Statist.* **18** 1188–1202.

HOSKING, J. R. M. (1981). Fractional differencing. *Biometrika* **68** 165–176.

HURST, H. E. (1951). Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers* **116** 770–808.

KÜNSCH, H. (1986). Statistical aspects of self-similar processes. *Proceedings of the First World Congress of the Bernoulli Society* **1** 67–74.

LAI, T. L. and STOUT, W. (1980). Limit theorems for sums of dependent random variables. *Z. Wahrsch. Verw. Gebiete* **51** 1–14.

MAJOR, P. (1981). *Multiple Wiener–Itô Integrals: With Applications to Limit Theorems. Lecture Notes in Math.* **849**. Springer, New York.

MANDELBROT, B. B. and TAQQU, M. S. (1979). Robust R/S analysis of long run serial correlation. *Proceedings of the 42nd Session of the ISI* **2** 69–100.

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

ROBINSON, P. M. (1994). Time series with strong dependence. *Sixth World Congress of the Econometric Society* 47–96. Cambridge Univ. Press.

ROSENBLATT, M. (1961). Independence and dependence. *Proc. Fourth Berkeley Symp. Math. Statist. Probab.* **2** 431–443. Univ. California Press, Berkeley.

SEN, P. K. (1972). On the Bahadur representation of sample quantiles for sequences of ϕ-mixing random variables. *J. Multivariate Anal.* **2** 77–95.

SURGAILIS, D. (1983). Zones of attraction of self-similar multiple integrals. *Lithuanian Math. J.* **22** 327–340.

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

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

TAQQU, M. S. (1985). A bibliographic guide to self-similar processes and long-range dependence. In *Dependence in Probability and Statistics* (E. Eberlein and M. S. Taqqu, eds.) 137–165. Birkhäuser, Boston.

INSTITUTE OF STATISTICAL SCIENCE
ACADEMIA SINICA
TAIPEI, TAIWAN 115
REPUBLIC OF CHINA

DEPARTMENT OF STATISTICS
TEXAS A & M UNIVERSITY
COLLEGE STATION, TEXAS 77843-3143
E-MAIL: thsing@stat.tamu.edu