ON THE BAHADUR REPRESENTATION OF SAMPLE QUANTILES FOR DEPENDENT SEQUENCES[1]

By Wei Biao Wu

June 2, 2004

University of Chicago

Abstract: We establish the Bahadur representation of sample quantiles for linear and some widely used non-linear processes. Local fluctuations of empirical processes are discussed. Applications to the trimmed and Winsorized means are given. Our results extend previous ones by establishing sharper bounds under milder conditions and thus provide new insight into the theory of empirical processes for dependent random variables.

1 Introduction.

Let $(\varepsilon_k)_{k\in\mathbb{Z}}$ be independent and identically distributed (iid) random elements and let G be a measurable function such that

$$X_n = G(\ldots, \varepsilon_{n-1}, \varepsilon_n) \tag{1}$$

is a well-defined random variable. Clearly X_n represents a huge class of stationary processes. Let $F(x) = \mathbb{P}(X_n \le x)$ be the marginal distribution function of X_n and let f be its density. For $0 < p < 1$, denote by $\xi_p = \inf\{x : F(x) \ge p\}$ the pth quantile of F. Given a sample $X_1, \ldots, X_n$, let $\xi_{n,p}$ be the pth $(0 < p < 1)$ sample quantile and define the empirical distribution function

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{X_i \le x}.$$

For simplicity we also refer to $\xi_{n,p}$ as the pth quantile of F_n. In this paper we are interested in finding asymptotic representations of $\xi_{n,p}$. Assuming that $(X_i)_{k\in\mathbb{Z}}$ are iid and $f(\xi_p) > 0$,

[1] *Mathematical Subject Classification (1991):* Primary 60F05, 60F17; secondary 60G35

Key words and phrases. Long- and short-range dependence, Bahadur representation, nonlinear time series, almost sure convergence, linear process, martingale inequalities, empirical processes, law of iterated logarithm

Bahadur (1966) first established the almost sure result

$$\xi_{n,p} = \xi_p + \frac{p - F_n(\xi_p)}{f(\xi_p)} + O_{\text{a.s.}}[n^{-3/4}(\log n)^{1/2}(\log\log n)^{1/4}], \tag{2}$$

where a sequence of random variables Z_n is said to be $O_{\text{a.s.}}(r_n)$ if Z_n/r_n is almost surely bounded. Refinements of Bahadur's result in the iid setting were provided by Kiefer in a sequence of papers; see Kiefer (1967, 1970a, 1970b). In particular, Kiefer (1967) showed that, if f' is bounded in a neighborhood of ξ_p and $f(\xi_p) > 0$, then

$$\limsup_{n\to\infty} \pm \frac{\xi_{n,p} - \xi_p - [p - F_n(\xi_p)]/f(\xi_p)}{n^{-3/4}(\log\log n)^{3/4}} = \frac{2^{5/4}3^{-3/4}p^{1/2}(1-p)^{1/2}}{f(\xi_p)} \tag{3}$$

almost surely for either choice of sign. Recent contributions can be found in Einmahl (1996) and Deheuvels (1997).

Extensions of the above results to dependent random variables have been pursued in Sen (1968) for m-dependent processes, in Sen (1972) for strongly mixing processes, in Hesse (1990) for short-range dependent (SRD) linear processes and in Ho and Hsing (1996) for long-range dependent (LRD) linear processes. The main objective of this paper is to generalize and refine these results for linear and some nonlinear processes.

Sample quantiles are closely related to empirical processes, and the asymptotic theory of empirical processes is then a natural vehicle for studying their limiting behavior. There is a well-developed theory of empirical processes for iid observations; see for example the excellent treatment by Shorack and Wellner (1986). The celebrated Hungarian construction can be used to obtain asymptotic representations of sample quantiles (cf Chapter 15 in Shorack and Wellner (1986)).

Recently there have been many attempts towards a convergence theory of empirical processes for dependent random variables. Such a theory is needed for the related statistical inference. Ho and Hsing (1996) and Wu (2003a) considered the empirical process theory for LRD sequences and obtained asymptotic expansions, while Doukhan and Surgailis (1998) considered SRD processes. Instantaneous transforms of Gaussian processes are treated in Dehling and Taqqu (1989). Further references on this topic can be found in the recent survey edited by Dehling, Mikosch and Sorensen (2002).

For dependent random variables, powerful tools like the Hungarian construction do not exist in general. To obtain comparable results as in the iid setting, we propose to employ a martingale based method. The main idea is to approximate sums of stationary processes by martingales. Such approximation schemes act as a bridge which connects stationary processes and martingales. One can then leverage several results from martingale theory, such as martingale central limit theorems, martingale inequalities, the martingale law of iterated logarithm, etc to obtain the desired results. Gordin (1969) first applied the martingale approximation method and established a central limit theory for stationary processes; see also Gordin and Lifsic (1978). Wu and Woodroofe (2004) present some recent developments. Several of its applications on various problems are given in Hall and Heyde (1980), Wu and Mielniczuk (2002), Wu (2003a), Wu (2003b), Hsing and Wu (2004).

Historically, many limit theorems for dependent random variables have been established under strongly mixing conditions. On the other hand, although the martingale approximation based approach imposes mild and easily verifiable conditions, it nevertheless may allow one to obtain optimal results, in the sense that they may be as sharp as the corresponding ones in the iid setting.

In this paper, for some SRD linear processes, we obtain the following asymptotic representation of sample quantiles

$$\xi_{n,p} = \xi_p + \frac{p - F_n(\xi_p)}{f(\xi_p)} + O_{\text{a.s.}}[n^{-3/4}(\log\log n)^{3/4}]$$

(cf Theorem 1), which gives an optimal bound $n^{-3/4}(\log\log n)^{3/4}$ in view of Kiefer's result (3) for iid random variables. Sample quantiles for LRD processes and some widely used nonlinear processes are also discussed and similar representations are derived. In establishing such asymptotic representations, we also consider the local and global behavior of empirical processes of dependent random variables.

We next introduce the necessary notation. A random variable ξ is said to be in $\mathcal{L}^q$, $q \geq 1$, if $\|\xi\|_q := [\mathbb{E}(|\xi|^q)]^{1/q} < \infty$. Write $\|\cdot\| = \|\cdot\|_2$. Let the shift process $\mathcal{F}_k = (\ldots, \varepsilon_{k-1}, \varepsilon_k)$ and the projection operator $\mathcal{P}_k\xi = \mathbb{E}(\xi|\mathcal{F}_k) - \mathbb{E}(\xi|\mathcal{F}_{k-1})$, $k \in \mathbb{Z}$. For a sequence of random variables Z_n, we say that $Z_n = o_{\text{a.s.}}(r_n)$ if Z_n/r_n converges to 0 almost surely. Write $a_n \sim b_n$ if $\lim_{n\to\infty} a_n/b_n = 1$.

The rest of the paper is structured as follows. Point-wise and uniform Bahadur's representations for SRD linear processes are presented in Section 2 and proofs are given in Section 6. LRD processes and nonlinear time series are discussed in Sections 3 and 4, respectively. Applications to the trimmed and Winsorized means are given in Section 5. Section 7 contains proofs and some discussions of results presented in Section 3.

2 SRD processes.

A causal (one-sided) linear process is defined by $X_k = \sum_{i=0}^{\infty} a_i \varepsilon_{k-i}$, where ε_k are iid random variables and a_k are real coefficients such that X_k exists almost surely. The almost sure existence of X_n can be checked by the well-known Kolmogorov three-series theorem (cf Chow and Teicher, 1978). Let f_ε and F_ε be the density and distribution functions of ε, respectively. Recall that F and F_n are the distribution and the empirical distribution functions of X_n and ξ_p is the pth quantile of F. Without loss of generality let $a_0 = 1$. Define the truncated process by $X_{n,k} = \sum_{j=n-k}^{\infty} a_j \varepsilon_{n-j}$ and the *conditional empirical distribution function* by

$$F_n^*(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}(\mathbf{1}_{X_i \le x}|\mathcal{F}_{i-1}) = \frac{1}{n}\sum_{i=1}^{n} F_\varepsilon(x - X_{i,i-1}).$$

Throughout the section we assume that

$$\sup_x [f_\varepsilon(x) + |f'_\varepsilon(x)|] < \infty. \tag{4}$$

It is easily seen that (4) implies $\sup_x[f(x) + |f'(x)|] < \infty$ in view of the relation $F(x) = \mathbb{E}[F_\varepsilon(x - \sum_{i=1}^{\infty} a_i \varepsilon_{k-i})]$ and the Lebesgue dominated convergence theorem. Define the function $\ell_q(n) = (\log\log n)^{1/2}$ if $q > 2$ and $\ell_q(n) = (\log n)^{3/2}(\log\log n)$ if $q = 2$.

Theorem 1. *Let* $X_k = \sum_{i=0}^{\infty} a_i \varepsilon_{k-i}$ *and assume (4),* $f(\xi_p) > 0$ *and* $\mathbb{E}(|\varepsilon_k|^\alpha) < \infty$ *for some* $\alpha > 0$. *[a] If*

$$\sum_{i=n}^{\infty} |a_i|^{\min(\alpha/q,1)} = O(\log^{-1/q} n) \tag{5}$$

for some $q > 2$, *then (i) there exists* $C > 0$ *such that* $\delta_{n,q} = C\ell_q(n)/[f(\xi_p)\sqrt{n}]$ *satisfies*

$$F_n(\xi_p + \delta_{n,q}) \ge p \ge F_n(\xi_p - \delta_{n,q}) \text{ almost surely} \tag{6}$$

and $|\xi_{n,p} - \xi_p| \le \delta_{n,q}$ *almost surely and (ii) the Bahadur representation holds:*

$$\xi_{n,p} = \xi_p + \frac{p - F_n(\xi_p)}{f(\xi_p)} + O_{\text{a.s.}}[n^{-3/4}(\log\log n)^{1/2}\ell_q^{1/2}(n)]. \tag{7}$$

[b] If

$$\sum_{i=1}^{\infty} |a_i|^{\min(\alpha/2,1)} < \infty, \tag{8}$$

then (i) and (ii) in [a] hold for $q = 2$.

Remark **1**. If $\alpha = 2$, then the process $(X_k)_{k\in\mathbb{Z}}$ has finite variance, and (8) implies that $(X_k)_{k\in\mathbb{Z}}$ is short-range dependent since its covariances are summable. $\diamondsuit$

Remark **2**. If $\alpha > 2$ and there is a $q > 2$ such that (5) holds, then $\sum_{i=n}^{\infty} |a_i| = O(\log^{-1/\alpha} n)$. The implication is clear if $q < \alpha$. If $q > \alpha$, then $\sum_{i=n}^{\infty} |a_i|^{\alpha/q} \geq (\sum_{i=n}^{\infty} |a_i|)^{\alpha/q}$ and we also have $\sum_{i=n}^{\infty} |a_i| = O(\log^{-1/\alpha} n)$. Therefore, in the case $\alpha > 2$, it suffices to check (5) for the special case $q = \alpha$ instead of verifying it for a whole range of $q > 2$. The condition $\sum_{i=n}^{\infty} |a_i| = O(\log^{-1/\alpha} n)$ is fairly mild for a linear process being short-range dependent. For example, it is satisfied if $a_n = O(n^{-1} \log^{-1-1/\alpha} n)$. $\diamondsuit$

Assuming that $\mathbb{E}(|\varepsilon_k|^\alpha) < \infty$ for some $\alpha > 0$ and that $|a_n| = O(n^{-\kappa})$ with $\kappa > 1 + 2/\alpha$, Hesse (1990) obtained the representation

$$\xi_{n,p} = \xi_p + \frac{p - F_n(\xi_p)}{f(\xi_p)} + O_{\text{a.s.}}(n^{-3/4+\gamma}), \tag{9}$$

where $\gamma > [\alpha^2(8\kappa - 5) + 2\alpha(10\kappa - 9) - 13]/(4\alpha\kappa - 2\alpha - 2)^2$. In comparison to Hesse's result, our condition (5) only requires $\kappa > \max(1, 2/\alpha)$. If $q > 2$, then the error term (7) is $O_{\text{a.s.}}[n^{-3/4}(\log\log n)^{1/2}\ell_q^{1/2}(n)] = O_{\text{a.s.}}[n^{-3/4}(\log\log n)^{3/4}]$, which gives an optimal bound; see Kiefer's relationship (3). The bound is much better than the one in (9). For example, if $\alpha = 1$ and $\kappa = 3.01$, then Hesse's result (9) gives the error bound $O_{\text{a.s.}}(n^{-0.0031\ldots})$. On the other hand, in Hesse's result, ε_i does not need to have a density.

Remark **3**. It is unclear whether Kiefer's law of iterated logarithm (3) can be extended to SRD processes. Our result only provides an upper bound. Kiefer (1967)'s proof involves extremely meticulous analysis and it heavily depends on the iid assumption. It seems that Kiefer's arguments can not be directly applied here. $\diamondsuit$

Example **1**. Suppose that ε_i is symmetric and its distribution function $F_\varepsilon(x) = 1 - L(x)/x^\alpha$, $x > 0$, where $0 < \alpha \leq 2$ and L is slowly varying at ∞. Here a function $L(x)$ is said to be slowly varying at ∞ if, for any $\lambda > 0$, $\lim_{x\to\infty} L(\lambda x)/L(x) = 1$. Notice that ε_i is in the domain of attraction of symmetric $\alpha-$stable distributions.

Assume that $|a_n| = O(n^{-r})$ for some $r > 2/\alpha$. Then for $q \in (2, r\alpha)$, (5) holds. In this case, $\mathbb{E}(|\varepsilon|^\alpha) = 2\alpha \int_0^\infty x^{-1}L(x)dx$ may be infinite. However, there exists a pair (α', q') such that $\mathbb{E}(|\varepsilon|^{\alpha'}) < \infty$ and (5) holds for this pair. Actually, one can simply choose $\alpha' < \alpha$ such that $2 < r\alpha'$ and let $q' = (2 + r\alpha')/2$. Then $\sum_{i=n}^\infty |a_i|^{\min(\alpha'/q',1)} = O(n^{1-r\alpha'/q'})$ with $r\alpha'/q' > 1$ and $\mathbb{E}(|\varepsilon|^{\alpha'}) \le 1 + \int_1^\infty \mathbb{P}(|\varepsilon|^{\alpha'} > u)du = 1 + 2\alpha' \int_1^\infty x^{\alpha'-\alpha-1}L(x)dx < \infty$. By Theorem 1, we have the Bahadur representation (7) with the optimal error bound $O_{\text{a.s.}}[n^{-3/4}(\log\log n)^{3/4}]$. ♢

Theorem 1 establishes Bahadur's representation for a single $p \in (0, 1)$. The uniform behavior of $\xi_{n,p} - \xi_p$ over $p \in [p_0, p_1]$, $0 < p_0 < p_1 < 1$, is addressed in Theorem 2. Such results have applications in the study of the trimmed and Winsorized means; see Section 5. Let $\iota_q(n) = (\log n)^{1/q}(\log\log n)^{2/q}$ if $q > 2$ and $\iota_2(n) = (\log n)^{3/2}(\log\log n)$.

Theorem **2**. *Let* $X_k = \sum_{i=0}^\infty a_i\varepsilon_{k-i}$. *Assume (4),* $\inf_{p_0\le p\le p_1} f(\xi_p) > 0$ *for some* $0 < p_0 < p_1 < 1$ *and*

$$\sup_x |f''_\varepsilon(x)| < \infty. \tag{10}$$

In addition assume that there exist $\alpha > 0$ *and* $q \ge 2$ *such that* $\mathbb{E}(|\varepsilon_k|^\alpha) < \infty$ *and*

$$\sum_{i=1}^\infty |a_i|^{\min(\alpha/q,1)} < \infty. \tag{11}$$

Then (i) $\sup_{p_0\le p\le p_1} |\xi_{n,p} - \xi_p| = o_{\text{a.s.}}[\iota_q(n)/\sqrt{n}]$ *and (ii) the uniform Bahadur representation holds:*

$$\sup_{p_0\le p\le p_1} \left| \xi_{n,p} - \xi_p - \frac{p - F_n(\xi_p)}{f(\xi_p)} \right| = O_{\text{a.s.}}[n^{-3/4}(\iota_q(n)\log n)^{1/2}]. \tag{12}$$

Remark **4**. Generally speaking, (12) cannot be extended to $p_0 = 0$ and/or $p_1 = 1$. The quantity $\xi_{n,p} - \xi_p$ exhibits an erratic behavior as $p \to 0$ or 1. The extremal theory is beyond the scope of the current paper. ♢

Remark **5**. If ε_0 has finite moments of any order, then under the condition $\sum_{i=1}^\infty |a_i| < \infty$, (12) gives the bound $n^{-3/4}(\log n)^{1/2+\eta}$ for any $\eta > 0$. ♢

Remark **6**. The Kiefer-Bahadur theorem asserts that, for iid random variables, the left-hand side of (12) has the optimal order $n^{-3/4}(\log n)^{1/2}(\log\log n)^{1/4}$; see Chapter 15 in Shorack and Wellner (1986). Our bound $n^{-3/4}(\iota_q(n)\log n)^{1/2}$ is not sharp. The reason is that we are unable to obtain a law of iterated logarithm for $\sup_{a\le x\le b}|F_n(x)-F(x)|$; see (54) in the proof of Theorem 2 in Section 6.6 where the weaker result $\sup_{a\le x\le b}|F_n(x)-F(x)| = o_{\text{a.s.}}[\iota_q(n)/\sqrt{n}]$ is proved. On the other hand, in proving Theorem 1, we are able to establish a law of iterated logarithm for $F_n(x)-F(x)$ at a *single point* x (cf. Proposition 1 and (i) of Lemma 10), by which the optimal rate $O_{\text{a.s.}}[n^{-3/4}(\log\log n)^{3/4}]$ in (7) can be derived. $\diamondsuit$

3 LRD processes.

Let the coefficients $a_0 = 1$, $a_n = n^{-\beta}L(n)$, $n \ge 1$, where $1/2 < \beta < 1$ and L is a slowly varying function at infinity; let $X_k = \sum_{i=0}^{\infty} a_i\varepsilon_{k-i}$, where ε_k are iid random variables with mean zero and finite variance. By Karamata's Theorem, the covariances $\gamma(n) = \mathbb{E}(X_0X_n) \sim C_\beta n^{1-2\beta}L^2(n)$, where $C_\beta = \mathbb{E}(\varepsilon_k^2)\int_0^\infty x^{-\beta}(1+x)^{-\beta}dx$, are not summable and the process is said to be long-range dependent. The asymptotic behavior of LRD processes is quite different from that of SRD ones. We shall apply the empirical process theory developed in Wu (2003a) and establish Bahadur's representation for long-range dependent processes.

Let $\Psi_n = \sqrt{n}\sum_{k=1}^n k^{1/2-2\beta}L^2(k)$ and

$$\sigma_{n,1}^2 = \|n\bar{X}_n\|^2 \sim \frac{C_\beta}{(1-\beta)(3-2\beta)} n^{3-2\beta}L^2(n). \tag{13}$$

By Karamata's Theorem, $\Psi_n \sim n^{2-2\beta}L^2(n)/(3/2-2\beta)$ if $\beta < 3/4$, $\Psi_n \sim \sqrt{n}L^*(n)$ if $\beta = 3/4$, where $L^*(n) = \sum_{k=1}^n L^2(k)/k$ is also a slowly varying function, and $\Psi_n \sim \sqrt{n}\sum_{k=1}^\infty k^{1/2-2\beta}L^2(k)$ if $\beta > 3/4$. Let $A_n(\beta) = \Psi_n^2(\log n)(\log\log n)^2$ if $\beta < 3/4$ and $A_n(\beta) = \Psi_n^2(\log n)^3(\log\log n)^2$ if $\beta \ge 3/4$.

Theorem **3**. *Assume* $\inf_{p_0\le p\le p_1} f(\xi_p) > 0$ *for some* $0 < p_0 < p_1 < 1$, $\mathbb{E}(\varepsilon_i^4) < \infty$ *and*

$$\sum_{i=0}^{2}\sup_x |f_\varepsilon^{(i)}(x)| + \int_{\mathbb{R}} |f'_\varepsilon(u)|^2 du < \infty. \tag{14}$$

Let $b_n = \sigma_{n,1}(\log n)^{1/2}(\log\log n)/n$. *Then*

$$\sup_{p_0 \le p \le p_1} \left| \xi_{n,p} - \xi_p - \frac{p - F_n(\xi_p)}{f(\xi_p)} - \frac{\bar{X}_n^2}{2} \frac{f'(\xi_p)}{f(\xi_p)} \right| = O_{\text{a.s.}} \left[b_n^3 + \frac{\sqrt{b_n \log n}}{\sqrt{n}} + \frac{b_n \sqrt{A_n(\beta)}}{n} \right]. \quad (15)$$

The three terms in the $O_{\text{a.s.}}$ bound of (15) have different orders of magnitude for different β and correspondingly the term that dominates the bound is different. If $\beta > 7/10$, since $3/2 - 3\beta < -\beta/2 - 1/4$ and $-\beta < -\beta/2 - 1/4$, it is easily seen that $b_n^3 + b_n\sqrt{A_n(\beta)}/n = o[\sqrt{(b_n \log n)/n}]$ in view of $\Psi_n = O[\sqrt{n}L^*(n) + n^{2-2\beta}L^2(n)]$ and $\sqrt{A_n(\beta)} \le \Psi_n(\log n)^{3/2}(\log\log n)$. Hence the dominant one in the bound of (15) is $O_{\text{a.s.}}[\sqrt{(b_n \log n)/n}]$. On the other hand, if $\beta < 7/10$, then $\sqrt{(b_n \log n)/n} = o[b_n\sqrt{A_n(\beta)}/n]$, $b_n\sqrt{A_n(\beta)}/n \sim C_1 n^{3(1/2-\beta)}L^3(n)(\log n)(\log\log n)^2$ and $b_n^3 \sim C_2 n^{3(1/2-\beta)}L^3(n)(\log n)^{3/2}(\log\log n)^3$ for some $0 < C_1, C_2 < \infty$. So $b_n\sqrt{A_n(\beta)}/n = o(b_n^3)$. For the boundary case $\beta = 7/10$, the situation is more subtle since the growth rate of the slowly varying function L is involved. In summary, noting that $\Psi_n = O[\sqrt{n}L^*(n) + n^{2-2\beta}L^2(n)]$, the error bound of (15) is

$$\begin{aligned} &O\{[n^{3(1/2-\beta)} + n^{(1/2-\beta)/2}/n^{1/2} + n^{1/2-\beta}(n^{1/2} + n^{2-2\beta})/n]L_1(n)\} \\ &= O[n^{\max(-\beta/2-1/4,\ 3/2-3\beta)}L_1(n)] \end{aligned} \quad (16)$$

for some slowly varying function L_1. This bound is less accurate than the one for the SRD or the iid counterparts since $\max(-\beta/2 - 1/4,\ 3/2 - 3\beta) > -3/4$ if $\beta < 1$. If $3/4 < \beta < 1$, then the bound is $O_{\text{a.s.}}[n^{-\beta/2-1/4}L_1(n)]$. See Section 7.2 for more discussion on the sharpness of (15) and (16).

In comparison with Bahadur's representations (2) for iid observations or (7) for short-range dependent processes, (15) has an interesting and different flavor in that it involves the correction term $\frac{1}{2}\bar{X}_n^2 f'(\xi_p)/f(\xi_p)$. More interestingly, this correction term is not needed if $\beta > 5/6$, which includes some LRD processes. Actually, by Lemma 16 in Section 7, $|\bar{X}_n|^2 = o_{\text{a.s.}}(b_n^2)$. Noting that $b_n^2 = o(\sqrt{b_n \log n}/\sqrt{n})$ if $\beta > 5/6$. Then the correction term $\frac{1}{2}\bar{X}_n^2 f'(\xi_p)/f(\xi_p)$ can be absorbed in the bound $\sqrt{b_n \log n}/\sqrt{n}$.

If the dependence of the process is strong enough, then we do need the correction $\frac{1}{2}\bar{X}_n^2 f'(\xi_p)/f(\xi_p)$ for a more accurate representation. Specifically, if $\beta \in (1/2, 5/6)$, then $\sqrt{b_n \log n}/\sqrt{n} = o(\sigma_{n,1}^2/n^2)$, $b_n^3 + b_n\sqrt{A_n(\beta)}/n = o(\sigma_{n,1}^2/n^2)$ and as the central limit theorem $n\bar{X}_n/\sigma_{n,1} \Rightarrow N(0,1)$ holds, the correction term has a non-negligible contribution.

4 Nonlinear time series.

In the case that G may not have a linear form, we assume that G satisfies the *geometric-moment contraction* (GMC) condition. On a possibly richer probability space, define iid random variables $\varepsilon'_j, \varepsilon_{i,k},\ i,j,k \in \mathbb{Z}$, which are identically distributed as ε_0 and are independent of $(\varepsilon_j)_{j\in\mathbb{Z}}$. The process X_n defined in (1) is said to be geometric-moment contracting if there exist $\alpha > 0$, $C = C(\alpha) > 0$ and $0 < r = r(\alpha) < 1$ such that for all $n \geq 0$,

$$\mathbb{E}[|G(\ldots,\varepsilon_{-1},\varepsilon_0,\varepsilon_1,\ldots,\varepsilon_n) - G(\ldots,\varepsilon'_{-1},\varepsilon'_0,\varepsilon_1,\ldots,\varepsilon_n)|^\alpha] \leq Cr^n. \tag{17}$$

The process $X'_n := G(\ldots,\varepsilon'_{-1},\varepsilon'_0,\varepsilon_1,\ldots,\varepsilon_n)$ can be viewed as a coupled version of X_n with the "past" $\mathcal{F}_0 = (\ldots,\varepsilon_{-1},\varepsilon_0)$ replaced by the iid copy $\mathcal{F}'_0 = (\ldots,\varepsilon'_{-1},\varepsilon'_0)$. Here we shall use (17) as our basic assumption for studying the asymptotic behavior of nonlinear time series. Since (17) only imposes the decay rate of the moment of the distance $|X_n - X'_n|$, it is often easily verifiable. In comparison, the classical strong mixing assumptions are typically difficult to be checked. Recently Hsing and Wu (2004) adopted (17) as the underlying assumption and studied the asymptotic behavior of weighted U-statistics for nonlinear time series.

Condition (17) is actually very mild as well. Consider the important special class of *iterated random functions* (Elton, 1990), which is recursively defined by

$$X_n = G(X_{n-1},\varepsilon_n), \tag{18}$$

where $G(\cdot,\cdot)$ is a bivariate measurable function with the Lipschitz constant

$$L_\varepsilon = \sup_{x'\neq x} \frac{|G(x,\varepsilon) - G(x',\varepsilon)|}{|x - x'|} \leq \infty \tag{19}$$

satisfying

$$\mathbb{E}(\log L_\varepsilon) < 0 \text{ and } \mathbb{E}[L_\varepsilon^\alpha + |x_0 - G(x_0,\varepsilon)|^\alpha] < \infty \tag{20}$$

for some $\alpha > 0$ and x_0. Diaconis and Freedman (1999) showed that, under (20), the Markov chain (18) admits a unique stationary distribution. Wu and Woodroofe (2000) further argued that (20) also implies the geometric-moment contraction (17); see Lemma

3 in the latter paper. Some recent improvements are presented in Wu and Shao (2004). Under suitable conditions on model parameters, many popular nonlinear time series models such as TAR, RCA and ARCH etc satisfy (20). Our main result is given next.

THEOREM 4. *Assume (17),* $\sup_x[f(x) + |f'(x)|] < \infty$ *and* $\inf_{p_0 \le p \le p_1} f(\xi_p) > 0$ *for some* $0 < p_0 < p_1 < 1$. *Then*

$$\sup_{p_0 \le p \le p_1} \left| \xi_{n,p} - \xi_p - \frac{p - F_n(\xi_p)}{f(\xi_p)} \right| = O_{\text{a.s.}}(n^{-3/4} \log^{3/2} n). \tag{21}$$

PROOF. For a fixed $\tau > 2$ let $m = \lfloor \omega \log n \rfloor$, where $\omega = \omega_\tau$ is given in Lemma 1 and $\lfloor t \rfloor$ denotes the integer part of t; let

$$\tilde{X}_k = G(\ldots, \varepsilon_{k-m-2,k}, \varepsilon_{k-m-1,k}, \varepsilon_{k-m,k}, \varepsilon_{k-m+1}, \varepsilon_{k-m+2}, \ldots, \varepsilon_{k-1}, \varepsilon_k). \tag{22}$$

Our strategy is to replace the "past" $\mathcal{F}_{k-m} = (\ldots, \varepsilon_{k-m-1}, \varepsilon_{k-m})$ in X_k by the iid copies $(\ldots, \varepsilon_{k-m-2,k}, \varepsilon_{k-m-1,k}, \varepsilon_{k-m,k})$ so that $(\tilde{X}_k)_{k \in \mathbb{Z}}$ is an m-dependent process. When X_n is a linear process, Hesse (1990) adopted a truncation argument which *forgets* the past $\mathcal{F}_{k-m}$ and approximates X_k by $G_n(\varepsilon_{k-m+1}, \ldots, \varepsilon_{k-1}, \varepsilon_k)$ for some measurable function G_n. Clearly the distribution function of $G_n(\varepsilon_{k-m+1}, \ldots, \varepsilon_{k-1}, \varepsilon_k)$ may be different from F. Our coupling argument has the advantage that the marginal distribution function of $\tilde{X}_k$ is still F. For $j = 1, 2, \ldots, m$, let

$$\tilde{F}_{n,j}(x) = \frac{1}{1 + A_n(j)} \sum_{i=0}^{A_n(j)} \mathbf{1}_{\tilde{X}_{j+im} \le x} \text{ and } \tilde{F}_n(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{\tilde{X}_i \le x}, \tag{23}$$

where $A_n(j) = \lfloor n/m \rfloor$ for $1 \le j \le n - m\lfloor n/m \rfloor$ and $A_n(j) = \lfloor n/m \rfloor - 1$ for $1 + n - m\lfloor n/m \rfloor \le j \le m$. Let $A = A_n = n/m$ and $b_n = c\sqrt{\log A}/\sqrt{A}$, where the constant c will be determined later. Let $\tilde{M}_{n,j}(x) = \tilde{F}_{n,j}(x) - F(x)$ and $\tilde{M}_n(x) = \tilde{F}_n(x) - F(x)$. Since $|\tilde{M}_n(x) - \tilde{M}_n(y)| \le \max_{1 \le j \le m} |\tilde{M}_{n,j}(x) - \tilde{M}_{n,j}(y)|$, by Lemma 2, there is a $\delta_\tau > 0$ such that

$$\begin{aligned} &\mathbb{P}\left[\sup_{|x-y| \le b_n} |\tilde{M}_n(x) - \tilde{M}_n(y)| > \frac{\delta_\tau (b_n \log A)^{1/2}}{A^{1/2}}\right] \\ &\le \sum_{j=1}^{m} \mathbb{P}\left[\sup_{|u| \le b_n} |\tilde{M}_{n,j}(x) - \tilde{M}_{n,j}(y)| > \frac{\delta_\tau (b_n \log A)^{1/2}}{A^{1/2}}\right] = mO(A^{-\tau}) \end{aligned} \tag{24}$$

and similarly $\mathbb{P}[\sup_x |\tilde{M}_n(x)| > \delta_\tau \sqrt{\log A}/\sqrt{A}] = mO(A^{-\tau})$. Since $\tau > 2$, $mA^{-\tau} = O[n^{-\tau}(\log n)^{\tau+1}]$ is summable over n. By Lemma 1 and the Borel-Cantelli lemma, we have

$$
\begin{aligned}
\sup_{|x-y|\le b_n} |[F_n(x)-F(x)]-[F_n(y)-F(y)]| &\le \sup_{|x-y|\le b_n} |\tilde{M}_n(x)-\tilde{M}_n(y)| + \frac{2C_\tau \log n}{n} \\
&= \frac{\delta_\tau (b_n \log A)^{1/2}}{A^{1/2}} + \frac{2C_\tau \log n}{n} \qquad (25)
\end{aligned}
$$

and $\sup_x |F_n(x)-F(x)| \le \delta_\tau \sqrt{\log A}/\sqrt{A} + C_\tau(n^{-1}\log n)$ almost surely. Now in $b_n = c\sqrt{\log A}/\sqrt{A}$ we choose $c = (2+\delta_\tau)/[\inf_{p_0\le p\le p_1} f(\xi_p)]$. Then we have

$$
\begin{aligned}
\inf_{p_0\le p\le p_1} [F_n(\xi_p+b_n)-p] &\ge \inf_{p_0\le p\le p_1} [F(\xi_p+b_n)-p] - \sup_{p_0\le p\le p_1} |F_n(\xi_p)-p| \\
&\quad - \sup_{|x-y|\le b_n} |[F_n(x)-F(x)]-[F_n(y)-F(y)]| \\
&\ge b_n \inf_{p_0\le p\le p_1} f(\xi_p) + O(b_n^2) - [\delta_\tau \sqrt{\log A}/\sqrt{A} + C_\tau(n^{-1}\log n)] \\
&\quad -[\delta_\tau \sqrt{b_n \log A}/\sqrt{A} + 2C_\tau(n^{-1}\log n)] > \sqrt{\log A}/\sqrt{A}
\end{aligned}
$$

almost surely. Similarly $\sup_{p_0\le p\le p_1}[F_n(\xi_p-b_n)-p] < 0$ almost surely. Hence for $\Delta_{n,p} = \xi_{n,p}-\xi_p$, $\sup_{p_0\le p\le p_1} |\Delta_{n,p}| \le b_n$ almost surely since F_n is non-decreasing. Since $|F_n(\xi_{n,p}) - p| \le 1/n$, by (25),

$$
\begin{aligned}
&\sup_{p_0\le p\le p_1} |[F_n(\xi_{n,p})-F(\xi_{n,p})]-[F_n(\xi_p)-F(\xi_p)]| \\
&= \sup_{p_0\le p\le p_1} |[p-F(\xi_{n,p})]-[F_n(\xi_p)-F(\xi_p)]| + O(1/n) \\
&= O_{\text{a.s.}}\left[\frac{\delta_\tau (b_n \log A)^{1/2}}{A^{1/2}} + \frac{2C_\tau \log n}{n}\right] + O(1/n) = O_{\text{a.s.}}(n^{-3/4}\log^{3/2} n).
\end{aligned}
$$

which entails (21) in view of $\inf_{p_0\le p\le p_1} f(\xi_p) > 0$ and, by Taylor's expansion, $F(\xi_{n,p}) - F(\xi_p) = \Delta_{n,p} f(\xi_p) + O(\Delta_{n,p}^2)$ since $\sup_x |f'(x)| < \infty$. ◊

LEMMA 1. *Assume (17) and $\sup_x f(x) < \infty$. Then for any $\tau > 1$, there exist $\omega_\tau, C_\tau > 0$ such that for $m = \lfloor \omega_\tau \log n \rfloor$, we have*

$$
\mathbb{P}\left[\sup_x |\tilde{F}_n(x) - F_n(x)| \ge C_\tau n^{-1} \log n\right] = O(n^{-\tau}). \qquad (26)
$$

PROOF. Let $\rho = r^{1/(2\alpha)}$, $\omega_\tau = -(1+\alpha^{-1})(\tau+2)/\log\rho$ and $C_\tau = 1-(1+\alpha^{-1})(\tau+1)/\log\rho$; let R_n be the set $\cap_{i=1}^n \{|X_i - \tilde{X}_i| \le \rho^m\}$ and R_n' be its complement. Then

$$
\begin{aligned}
\mathbb{P}(R_n') &\le n\mathbb{P}(|X_i-\tilde{X}_i| \ge \rho^m) \le n\rho^{-\alpha m}\mathbb{E}(|X_i-\tilde{X}_i|^\alpha) \\
&\le n\rho^{-\alpha m} C r^m = nC\rho^{\alpha m} = o(n^{-\tau}).
\end{aligned}
$$

Let $K = C_\tau - 1$. By the triangle inequality, to establish (26), it suffices to show that

$$\mathbb{P}\left[\sup_x |\tilde{F}_n(x) - F_n(x)|\mathbf{1}_{R_n} > Kn^{-1}\log n\right] = O(n^{-\tau}). \tag{27}$$

Notice that $\sup_x |\tilde{F}_n(x) - F_n(x)|\mathbf{1}_{R_n} \le \sup_x[F_n(x+\rho^m) - F_n(x-\rho^m)]$. Clearly, the event $\{\sup_x[F_n(x+\rho^m) - F_n(x-\rho^m)] > Kn^{-1}\log n\}$ implies that there exist two indices i and j with $j - i \ge \lfloor K\log n\rfloor$ such that both X_i and X_j are in the interval $[x-\rho^m, x+\rho^m]$ for some $x \in \mathbb{R}$. Therefore,

$$\begin{aligned}
&\mathbb{P}\left[\sup_x[F_n(x+\rho^m) - F_n(x-\rho^m)] > Kn^{-1}\log n\right] \\
&\le \mathbb{P}\left[\cup_{i=1}^{n-\lfloor K\log n\rfloor} \cup_{j=i+\lfloor K\log n\rfloor}^{n} \{|X_i - X_j| \le 2\rho^m\}\right] \\
&\le \sum_{i=1}^{n-\lfloor K\log n\rfloor} \sum_{j=i+\lfloor K\log n\rfloor}^{n} \mathbb{P}(|X_i - X_j| \le 2\rho^m) \\
&\le n \sum_{j=\lfloor K\log n\rfloor}^{n} \mathbb{P}(|X_0 - X_j| \le 2\rho^m).
\end{aligned}$$

Recall (17) for $X'_j = G(\ldots, \varepsilon'_{-1}, \varepsilon'_0, \varepsilon_1, \ldots, \varepsilon_j)$. Then

$$\begin{aligned}
\mathbb{P}(|X_0 - X_j| \le 2\rho^m) &\le \mathbb{P}(|X_0 - X_j| \le 2\rho^m, |X_j - X'_j| \le \rho^j) + \mathbb{P}(|X_j - X'_j| > \rho^j) \\
&\le \mathbb{P}(|X_0 - X'_j| \le 2\rho^m + \rho^j) + \rho^{-\alpha j} C r^j.
\end{aligned}$$

Observe that X_0 and X'_j are iid, $\mathbb{P}(|X_0 - X'_j| \le \delta) = \mathbb{E}[\mathbb{P}(|X_0 - X'_j| \le \delta | X'_j)] \le 2c\delta$, where $c = \sup_x f(x) < \infty$. Thus

$$\begin{aligned}
\mathbb{P}\left[\sup_x |\tilde{F}_n(x) - F(x)|\mathbf{1}_{R_n} > Kn^{-1}\log n\right] &\le n \sum_{j=\lfloor K\log n\rfloor}^{n} [2c(2\rho^m + \rho^j) + \rho^{\alpha j} C] \\
&= nO(n\rho^m + \rho^{K\log n}) + nO(\rho^{\alpha K\log n}),
\end{aligned}$$

which ensures (27) by the choice of K and ω_τ. ♢

Lemma **2**. *Let $(Z_k)_{k\in\mathbb{Z}}$ be iid random variables with distribution and density functions F_Z and f_Z for which $\sup_z f_Z(z) < \infty$; let $F_{n,Z}(z) = \frac{1}{n}\sum_{i=1}^n \mathbf{1}_{Z_i \le z}$. Then for all $\tau > 1$, there exists $C_\tau > 0$ such that*

$$\mathbb{P}\left[\sup_x |F_{n,Z}(x) - F_Z(x)| > \frac{C_\tau(\log n)^{1/2}}{n^{1/2}}\right] = O(n^{-\tau}) \tag{28}$$

and

$$\mathbb{P}\left[\sup_{|x-y|\le b_n} |F_{n,Z}(x) - F_Z(x) - \{F_{n,Z}(y) - F_Z(y)\}| > \frac{C_\tau (b_n \log n)^{1/2}}{n^{1/2}}\right] = O(n^{-\tau}), \quad (29)$$

where $(b_n)_{n\ge1}$ *is a positive, bounded sequence of real numbers such that* $\log n = o(nb_n)$.

Lemma 2 easily follows from classical results for iid uniform random variables under quantile transformations; see the Dvoretzky-Kiefer-Wolfowitz inequality and Inequality 14.0.9 in Shorack and Wellner (1986). The lemma is needed in the proof of Theorem 4 and it is a special case of Lemma 7 in Section 6.2. We purposefully state Lemma 2 here also for the sake of comparison: the martingale based method may yield comparable results as those obtained under the iid assumption.

5 Trimmed and Winsorized means.

Let $\xi_{n,1/n} \le \xi_{n,2/n} \ldots \le \xi_{n,1}$ be the order statistics of $X_1, \ldots, X_n$. Then the trimmed and Winsorized means are of the forms $\sum_{i=\alpha(n)+1}^{\beta(n)} \xi_{n,i/n}/[\beta(n) - \alpha(n)]$ and $n^{-1}[\alpha(n)\xi_{n,\alpha(n)/n} + (n-\beta(n))\xi_{n,\beta(n)/n+1/n} + \sum_{i=\alpha(n)+1}^{\beta(n)} \xi_{n,i/n}]$ respectively, where $\alpha(n) = \lfloor np_0 \rfloor$ and $\beta(n) = \lfloor np_1 \rfloor$.

Stigler (1973) studied the asymptotic behavior of trimmed means for iid random variables. Here we shall apply Theorems 2 and 4 to obtain a central limit theorem for some dependent random variables. SRD linear processes and casual processes satisfying (17) are considered in (i) and (ii) of Theorem 5 respectively. Denote by $N(\mu, \sigma^2)$ a normal distribution with mean μ and variance σ^2.

Theorem **5**. *(i) Let* $q = 2$ *and assume that the conditions of Theorem 2 are satisfied. Then there is a* $\sigma < \infty$ *such that*

$$\sqrt{n}\left[\frac{\sum_{i=\alpha(n)+1}^{\beta(n)} \xi_{n,i/n}}{\beta(n) - \alpha(n)} - \frac{1}{p_1 - p_0}\int_{p_0}^{p_1} \xi_u du\right] \Rightarrow N(0, \sigma^2). \quad (30)$$

(ii) Assume that the conditions of Theorem 4 are satisfied. Then the central limit theorem (30) holds.

Proof. (i) Since $\xi_{n,u}$ is non-decreasing in u, $n\int_{(i-1)/n}^{i/n}\xi_{n,u}du \le \xi_{n,i/n} \le n\int_{i/n}^{(i+1)/n}\xi_{n,u}du$ holds for $1 < i < n-1$. Hence

$$n\int_{\alpha(n)/n}^{\beta(n)/n}\xi_{n,u}du \le \sum_{i=\alpha(n)+1}^{\beta(n)}\xi_{n,i/n} \le n\int_{[1+\alpha(n)]/n}^{[1+\beta(n)]/n}\xi_{n,u}du.$$

It is easily seen that, under the conditions of Theorem 2, (12) also holds over the expanded interval $[p_0-\tau, p_1+\tau]$ for some sufficiently small $\tau > 0$. Therefore, we have $\sup_{\alpha(n)/n\le u\le[1+\beta(n)]/n}|\xi_{n,u}| = O_{\text{a.s.}}(1)$ and consequently

$$\sum_{i=\alpha(n)+1}^{\beta(n)}\xi_{n,i/n} - n\int_{p_0}^{p_1}\xi_{n,u}du = O_{\text{a.s.}}(1). \tag{31}$$

By (12) of Theorem 2,

$$\int_{p_0}^{p_1}\xi_{n,u}du - \int_{p_0}^{p_1}\xi_u du - \int_{p_0}^{p_1}\frac{u-F_n(\xi_u)}{f(\xi_u)}du = O_{\text{a.s.}}[n^{-3/4}(\iota_2(n)\log n)^{1/2}]. \tag{32}$$

Lemma 11 in Section 6.4 asserts that $\{\sqrt{n}[F_n(x)-F(x)],\ \xi_{p_0}\le x\le\xi_{p_1}\} \Rightarrow \{W(x),\ \xi_{p_0}\le x\le\xi_{p_1}\}$ for some centered Gaussian process W in the Skorokhod space $D[\xi_{p_0},\xi_{p_1}]$ (Billingsley, 1968). By the continuous mapping theorem, (30) follows from (31) and (32).

(ii) By Theorem 4 in Wu and Shao (2004), under the conditions (17) and $\sup_x f(x) < \infty$, we also have the functional central limit theorem $\{\sqrt{n}[F_n(x)-F(x)],\ \xi_{p_0}\le x\le\xi_{p_1}\} \Rightarrow \{W(x),\ \xi_{p_0}\le x\le\xi_{p_1}\}$ for some Gaussian process W. So (30) holds in view of the argument in (i). $\diamondsuit$

Remark 7. Using the same argument, it is easily seen that for the Winsorized mean $n^{-1}[\alpha(n)\xi_{n,\alpha(n)/n} + (n-\beta(n))\xi_{n,\beta(n)/n+1/n} + \sum_{i=\alpha(n)+1}^{\beta(n)}\xi_{n,i/n}]$, we also have the central limit theorem (30) with the asymptotic mean $(p_1-p_0)^{-1}\int_{p_0}^{p_1}\xi_u du$ replaced by $p_0\xi_{p_0}+(1-p_1)\xi_{p_1}+\int_{p_0}^{p_1}\xi_u du$. Other forms of linear functions of order statistics can be similarly handled. $\diamondsuit$

6 Proofs of Theorems 1 and 2.

We first introduce our method. Recall $\mathcal{F}_k = (\ldots,\varepsilon_{k-1},\varepsilon_k)$ and $F_n^*(x) = \sum_{i=1}^n F_\varepsilon(x - X_{i,i-1})/n$. Write $F_n(x)-F(x) = M_n(x)+N_n(x)$, where $M_n(x) = F_n(x)-F_n^*(x)$ and $N_n(x) = F_n^*(x)-F(x)$.

Notice that under (4), the conditional empirical distribution function F_n^* is differentiable with the uniformly bounded derivative $f_n^*(x) = n^{-1}\sum_{i=1}^n f_\varepsilon(x - X_{i,i-1})$ and hence $dN_n(x)/dx = f_n^*(x) - f(x)$ is also uniformly bounded. The differentiability property greatly facilitates the related analysis. In comparison, F_n is a step function and hence discontinuous. On the other hand, $nM_n(x)$ forms a martingale with bounded, stationary and ergodic increments $\mathbf{1}_{X_i \le x} - \mathbb{E}(\mathbf{1}_{X_i \le x}|\mathcal{F}_{i-1})$. Therefore, results from martingale theory are applicable.

The martingale part M_n and the differentiable part N_n are treated in Sections 6.2 and 6.3 respectively. Section 6.4 discusses the oscillatory behavior and some asymptotic properties of empirical processes, which are needed for the derivation of Bahadur's representations. Proofs of Theorems 1 and 2 are given in Sections 6.5 and 6.6 respectively.

6.1 Some useful results.

The following Proposition 1 is needed in proving Theorems 1 and 2. See Wu (2004) for a proof.

PROPOSITION **1**. *Let $S_n(g) = \sum_{i=1}^n g(\mathcal{F}_i)$, where g is a measurable function such that $g(\mathcal{F}_0) \in \mathcal{L}^q$ for some $q \ge 2$, $\mathbb{E}[g(\mathcal{F}_0)] = 0$ and*

$$\Theta_{0,q} := \sum_{i=0}^{\infty} \|\mathcal{P}_0 g(\mathcal{F}_i)\|_q < \infty. \tag{33}$$

Let $B_q = 18q^{3/2}(q-1)^{1/2}$ if $q > 2$ and $B_q = 1$ if $q = 2$. Then

$$\|S_n(g)\|_q \le B_q \sqrt{n}\,\Theta_{0,q}. \tag{34}$$

Furthermore, if $\Theta_{m,q} := \sum_{i=m}^\infty \|\mathcal{P}_0 g(\mathcal{F}_i)\|_q = O[(\log m)^{-1/q}]$ for some $q > 2$, then

$$\limsup_{n\to\infty} \pm \frac{S_n(g)}{\sqrt{2n \log\log n}} = \sigma \tag{35}$$

almost surely for either choice of sign, where $\sigma = \|\sum_{i=0}^\infty \mathcal{P}_0 g(\mathcal{F}_i)\| < \infty$.

In order to apply Proposition 1 to $S_n(g) = n[F_n(x) - F(x)]$ or $n[f_n^*(x) - f(x)]$, one needs to estimate $\|\mathcal{P}_0 \mathbf{1}_{X_i \le x}\|$ or $\|\mathcal{P}_0 f_\varepsilon(x - X_{i,i-1})\|$. The following Lemma 3 provides a simple upper bound if the innovation ε_0 satisfies certain moment conditions. In particular, the innovations are allowed to have infinite variance.

Lemma 3. *Let* $X_k = \sum_{i=0}^\infty a_i \varepsilon_{k-i}$, *where* ε_k *are iid with* $\mathbb{E}(|\varepsilon_k|^\alpha) < \infty$ *for some* $\alpha > 0$. *Then under (4),* $\|\mathcal{P}_0 g(\mathcal{F}_n)\|_q = O[|a_n|^{\min(\alpha/q,1)}]$ *holds for* $g(\mathcal{F}_n) = \mathbf{1}_{X_n \le x}$ *and* $g(\mathcal{F}_n) = f_\varepsilon(x - X_{n,n-1})$. *If additionally (10) is satisfied, then the same bound also holds for* $g(\mathcal{F}_n) = f'_\varepsilon(x - X_{n,n-1})$.

Proof. Let $(\varepsilon'_i)_{i\in\mathbb{Z}}$ be an iid copy of $(\varepsilon_i)_{i\in\mathbb{Z}}$ and $X_n^* = X_n - a_n\varepsilon_0 + a_n\varepsilon'_0$; let G_n be the distribution function of $X_n - X_{n,0} = \sum_{j=0}^{n-1} a_j \varepsilon_{n-j}$. Since $c = \sup_x f_\varepsilon(x) < \infty$, it is easily seen that the density $g_n(x) = G'_n(x)$ is also bounded by c. Observe that $\mathbb{P}(X_n^* \le x|\mathcal{F}_0) = \mathbb{P}(X_n \le x|\mathcal{F}_{-1})$, by Jensen's inequality,

$$\begin{aligned}
\|\mathcal{P}_0 g(\mathcal{F}_n)\|_q &\le \|\mathbb{P}(X_n \le x|\mathcal{F}_0) - \mathbb{P}(X_n^* \le x|\mathcal{F}_0)\|_q \\
&= \|\mathbb{E}[G_n(x - X_{n,0}) - G_n(x - X_{n,0} + a_n\varepsilon_0 - a_n\varepsilon'_0)|\mathcal{F}_0]\|_q \\
&\le \|G_n(x - X_{n,0}) - G_n(x - X_{n,0} + a_n\varepsilon_0 - a_n\varepsilon'_0)\|_q \\
&\le \|\min(c|a_n\varepsilon_0 - a_n\varepsilon'_0|,\ 1)\|_q \\
&\le [\mathbb{E}(c|a_n\varepsilon_0 - a_n\varepsilon'_0|)^{\min(\alpha,q)}]^{1/q} \\
&= O[|a_n|^{\min(\alpha/q,1)}].
\end{aligned}$$

Here the elementary inequality $[\min(|b|, 1)]^q \le |b|^{\min(\alpha,q)}$ is applied. The other cases $g(\mathcal{F}_n) = f_\varepsilon(x - X_{n,n-1})$ and $g(\mathcal{F}_n) = f'_\varepsilon(x - X_{n,n-1})$ can be similarly proved. ◇

To establish a uniform Bahadur representation for $\xi_{n,p} - \xi_p$ over $p \in [p_0, p_1]$, $0 < p_0 < p_1 < 1$, we need the following version of maximal inequality, which will be used to obtain an almost sure upper bound of $\sup_{\xi_{p_0} \le x \le \xi_{p_1}} |F_n(x) - F(x)|$. Similar versions appeared in Billingsley (1968), Serfling (1970), Moricz (1976) and Lai and Stout (1980). For a proof of Lemma 4 see Wu (2004).

Lemma 4. *Let* $(Y_{k,\theta},\ k \in \mathbb{Z})_{\theta\in\Theta}$ *be a class of centered stationary processes in* $\mathcal{L}^q, q > 1$. *Namely for each* $\theta \in \Theta$, $(Y_{k,\theta})_{k\in\mathbb{Z}}$ *is a stationary process in* $\mathcal{L}^q$ *and* $\mathbb{E}(Y_{k,\theta}) = 0$. *Let* $S_{n,\theta} = Y_{1,\theta} + \cdots + Y_{n,\theta}$ *and let* $d = d(n)$ *be an integer such that* $2^{d-1} < n \le 2^d$. *Then*

$$\left\{\mathbb{E}^*\left[\max_{k\le n}\sup_{\theta\in\Theta}|S_{k,\theta}|^q\right]\right\}^{1/q} \le \sum_{j=0}^{d} 2^{(d-j)/q}\left\{\mathbb{E}^*\left[\sup_{\theta\in\Theta}|S_{2^j,\theta}|^q\right]\right\}^{1/q}, \tag{36}$$

where $\mathbb{E}^*$ *is the outer expectation:* $\mathbb{E}^* Z = \inf\{\mathbb{E}X :\ X \ge Z, X$ *is a random variable*$\}$.

6.2 The martingale part M_n.

LEMMA **5**. *Let $(b_n)_{n\geq 1}$ be a positive, bounded sequence of real numbers such that $\log^3 n = o(nb_n)$. Assume $\sup_x f_\varepsilon(x) < \infty$. Then for any $\tau > 1$, there exists a constant $C_\tau > 0$ such that*

$$\mathbb{P}\left\{\sup_{|u|\leq b_{2^k}} \max_{2^{k-1}<n\leq 2^k} n|M_n(x+u) - M_n(x)| > C_\tau \sqrt{2^k b_{2^k} \log k}\right\} = O(k^{-\tau}). \tag{37}$$

PROOF. Let $c = \sup_x f_\varepsilon(x) < \infty$. For a given $u > 0$, since $\mathbb{P}(x < X_i \leq x+u|\mathcal{F}_{i-1}) \leq cu$, we have

$$\sum_{i=1}^{n}[\mathbb{E}(\mathbf{1}_{x<X_i\leq x+u}|\mathcal{F}_{i-1}) - \mathbb{E}^2(\mathbf{1}_{x<X_i\leq x+u}|\mathcal{F}_{i-1})] \leq ncu.$$

Here without loss of generality we restrict u to be nonnegative. Let $t_k = \sqrt{2^k b_{2^k}\log k}$. Since $\mathbf{1}_{x<X_i\leq x+u} - \mathbb{E}(\mathbf{1}_{x<X_i\leq x+u}|\mathcal{F}_{i-1})$, $1 \leq i \leq n$, form bounded martingale differences, by Freedman's inequality (cf. Theorem 1.6 in Freedman (1975)) we get that

$$\mathbb{P}\left\{\max_{2^{k-1}<n\leq 2^k} n|M_n(x+u) - M_n(x)| > Ct_k\right\} \leq 2\exp[-C^2t_k^2/(2Ct_k + 2\times 2^k cu)] \tag{38}$$

for all $C > 0$. Let $\alpha_k = b_{2^k}/k$, $u_i = i\alpha_k$, $i = 0, 1, \ldots, k-1$ and $v_m = mb_{2^k}/(k2^k)$, $m = 0, 1, \ldots, 2^k - 1$. Since $t_k = o(2^k b_{2^k}/k)$, we have for sufficiently large k that

$$\begin{aligned}
&\mathbb{P}\left\{\max_{0\leq i\leq k-1}\max_{2^{k-1}<n\leq 2^k} n|M_n(x+u_i) - M_n(x)| > Ct_k\right\} \\
&\leq \sum_{i=0}^{k-1}\mathbb{P}\left\{\max_{2^{k-1}<n\leq 2^k} n|M_n(x+u_i) - M_n(x)| > Ct_k\right\} \\
&\leq 2k\exp\left[\frac{-C^2\log k}{2c+1}\right]
\end{aligned} \tag{39}$$

and similarly

$$\begin{aligned}
&\mathbb{P}\left\{\max_{0\leq m\leq 2^k-1}\max_{2^{k-1}<n\leq 2^k} n|M_n(x+v_m) - M_n(x)| > Ct_k\right\} \\
&\leq \sum_{m=0}^{2^k-1} 2\exp[-C^2t_k^2/(2Ct_k + 2\times 2^k cv_m)] \\
&\leq \sum_{m=0}^{2^k-1} 2\exp[-C^2t_k^2/(2Ct_k + 2\times 2^k cv_{2^k})]
\end{aligned}$$

$$\leq 2^{k+1} \exp\left[\frac{-C^2 k \log k}{2c+1}\right]. \tag{40}$$

For any v, $v_m < v \leq v_{m+1}$, observe that $0 \leq F_n^*(x+v_{m+1}) - F_n^*(x+v_m) \leq cb_{2^k}/(k2^k)$,

$$M_n(x+v) - M_n(x) \leq M_n(x+v_{m+1}) - M_n(x) + cb_{2^k}/(k2^k)$$

and similarly, $M_n(x+v) - M_n(x) \geq M_n(x+v_m) - M_n(x) - cb_{2^k}/(k2^k)$. So (40) yields

$$\mathbb{P}\left\{\sup_{0\leq v\leq \alpha_k} \max_{2^{k-1}<n\leq 2^k} n|M_n(x+v) - M_n(x)| > (C+1)t_k\right\} \leq 2^{k+1} \exp\left[\frac{-C^2 k \log k}{2c+1}\right]. \tag{41}$$

Since (41) holds for all $x \in \mathbb{R}$, by the triangle inequality, (41) together with (39) imply

$$\begin{aligned}
&\mathbb{P}\left\{\max_{0\leq u\leq b_{2^k}} \max_{2^{k-1}<n\leq 2^k} n|M_n(x+u) - M_n(x)| > (2C+1)t_k\right\} \\
&\leq \mathbb{P}\left\{\max_{0\leq i\leq k-1} \max_{2^{k-1}<n\leq 2^k} n|M_n(x+u_i) - M_n(x)| > Ct_k\right\} \\
&+ \sum_{i=0}^{k-1} \mathbb{P}\left\{\sup_{0\leq v\leq \alpha_k} \max_{2^{k-1}<n\leq 2^k} n|M_n(x+v+u_i) - M_n(x+u_i)| > (C+1)t_k\right\} \\
&\leq 2k \exp\left[\frac{-C^2 \log k}{2c+1}\right] + k \times 2^{k+1} \exp\left[\frac{-C^2 k \log k}{2c+1}\right].
\end{aligned}$$

Therefore (37) follows by letting $C_\tau = 1 + 2(\tau+1)^{1/2}(2c+1)^{1/2}$. ♢

LEMMA **6**. *Assume that the conditions of Lemma 5 are satisfied and in addition assume that there is a $\rho \geq 1$ such that for all sufficiently large n we have that*

$$\frac{b_{2n}}{\rho} \leq \min_{n\leq j\leq 2n} b_j \leq \max_{n\leq j\leq 2n} b_j \leq \rho b_{2n}. \tag{42}$$

Then for each fixed $x \in \mathbb{R}$,

$$\sup_{|u|\leq b_n} |M_n(x+u) - M_n(x)| = O_{\text{a.s.}}\left[\frac{\sqrt{b_n \log\log n}}{\sqrt{n}}\right]. \tag{43}$$

PROOF. Observe that due to (42), for all sufficiently large n, we have

$$\begin{aligned}
&\max_{2^{k-1}<n\leq 2^k} \frac{\sqrt{n}\sup_{|u|\leq b_n} |M_n(x+u) - M_n(x)|}{\sqrt{b_n \log\log n}} \\
&\leq \sup_{|u|\leq \rho b_{2^k}} \max_{2^{k-1}<n\leq 2^k} \frac{n|M_n(x+u) - M_n(x)|}{\sqrt{nb_n \log\log n}} \\
&\leq \sup_{|u|\leq \rho b_{2^k}} \max_{2^{k-1}<n\leq 2^k} \frac{n|M_n(x+u) - M_n(x)|}{\sqrt{2^{k-1}\rho^{-1} b_{2^k} \log\log 2^{k-1}}}
\end{aligned}$$

Hence (43) follows from Lemma 5 via the Borel-Cantelli lemma. ♢

LEMMA **7**. *Assume (4) and that $\mathbb{E}(|X_1|^\alpha) < \infty$ for some $\alpha > 0$. Then for all $\tau > 1$, there exists $C_\tau > 0$ such that*

$$\mathbb{P}\left[\sup_x |M_n(x)| > \frac{C_\tau (\log n)^{1/2}}{n^{1/2}}\right] = O(n^{-\tau}) \tag{44}$$

and

$$\mathbb{P}\left[\sup_{|x-y|\le b_n} |M_n(y) - M_n(x)| > C_\tau \frac{b_n^{1/2}(\log n)^{1/2}}{n^{1/2}}\right] = O(n^{-\tau}), \tag{45}$$

where $(b_n)_{n\ge 1}$ is a positive, bounded sequence of real numbers such that $\log n = o(nb_n)$.

PROOF. We only prove (45) since (44) can be similarly proved. Let $c = \sup_x f_\varepsilon(x) < \infty$, $v_n = \sqrt{nb_n \log n}$, $t_n = v_n/n$, $J = n^{(\tau+4)/\alpha}$ and $Y_i(x) = \mathbf{1}_{X_i \le x} - \mathbb{E}(\mathbf{1}_{X_i\le x}|\mathcal{F}_{i-1})$. Then

$$\begin{aligned}
I_n &:= \mathbb{P}\left[\sup_{|x-y|\le b_n, x\le -J} |M_n(y) - M_n(x)| > Ct_n\right] \\
&\le n\mathbb{P}\left[\sup_{|x-y|\le b_n, x\le -J} |Y_1(y) - Y_1(x)| > Ct_n\right] \\
&\le n(Ct_n)^{-1}\mathbb{E}\left[\sup_{|x-y|\le b_n, x\le -J} |Y_1(y) - Y_1(x)|\right] \\
&= O(nt_n^{-1})\mathbb{E}\left[\sup_{x\le -J+b_n} |Y_1(x)|\right] \\
&= O(n^2 v_n^{-1})(J-b_n)^{-\alpha}\mathbb{E}(|X_1|^\alpha) = O(n^{-1-\tau}),
\end{aligned}$$

where Markov's inequality is used in the second inequality. Similarly,

$$III_n := \mathbb{P}\left[\sup_{|x-y|\le b_n, x\ge J} |M_n(y) - M_n(x)| > Ct_n\right] = O(n^{-1-\tau}).$$

Let $x_i = ib_n/n$, $i = -N-1, \ldots, N+1$, where $N = \lfloor Jn/b_n \rfloor$, and

$$II_n := \mathbb{P}\left[\sup_{|x-y|\le b_n, -J<x<J} |M_n(y) - M_n(x)| > Ct_n\right].$$

Again by Freedman's inequality, for $|x-y| \le b_n$ and sufficiently large n,

$$\mathbb{P}[n|M_n(y) - M_n(x)| > Cv_n] \le 2\exp[-C^2 v_n^2/(2Cv_n + 2ncb_n)] \le 2n^{-C^2/(2c+1)}.$$

Thus

$$\mathbb{P}\left[\max_{i,j=-N-1,\ldots,N+1:\ |x_i-x_j|\le b_n} n|M_n(x_i)-M_n(x_j)|>Cv_n\right]=O(N^2)n^{-C^2/(2c+1)}.$$

For any x, y with $|x-y|\le b_n$, $|x|\le J$ and $|y|\le J$, choose i and j such that $x_i\le x<x_{i+1}$ and $x_j\le y<x_{j+1}$. Then

$$n[M_n(x_j)-M_n(x_{i+1})]-2cb_n\le n[M_n(y)-M_n(x)]\le n[M_n(x_{j+1})-M_n(x_i)]+2cb_n.$$

Therefore (45) follows by choosing $C_\tau^2=(2c+1)[(8+2\tau)/\alpha+\tau+5]$, given that

$$\mathbb{P}\left\{\sup_{|x-y|\le b_n}|M_n(y)-M_n(x)|>(C_\tau+1)t_n\right\}\le I_n+II_n+III_n$$

by the triangle inequality. ♢

REMARK 8. In Lemmas 5-7 it is not required that $b_n\to 0$. We shall use this fact to derive (54), which is a key step in proving Theorem 2. ♢

REMARK 9. It is worth noting that Lemmas 5-7 also apply to LRD processes. In Section 7 we will use them to prove the Bahadur representation for LRD processes. For iid random variables, the increments of the empirical and quantile processes are discussed in great detail in Deheuvels and Mason (1992). ♢

6.3 The differentiable part N_n.

LEMMA 8. *Let $b_n\to 0$. Assume (4) and $\mathbb{E}(|\varepsilon_k|^\alpha)<\infty$ for some $\alpha>0$. Further assume (5) if $q>2$ or (8) if $q=2$. Then*

$$\sup_{|t|\le b_n}|N_n(x+t)-N_n(x)|=\frac{\ell_q(n)}{\sqrt{n}}O_{\text{a.s.}}(b_n)+O_{\text{a.s.}}(b_n^2). \tag{46}$$

PROOF. Let $c_0=\sup_x|f_\varepsilon'(x)|$ and recall $f_n^*(x)=dF_n^*(x)/dx$. Clearly $|f'(x)|\le c_0$ since $f'(x)=\mathbb{E}[f_\varepsilon'(x-X_{i,i-1})]$. Using Taylor's expansion, we get

$$\sup_{|t|\le b_n}|N_n(x+t)-N_n(x)-t[f_n^*(x)-f(x)]|\le\frac{b_n^2}{2}\sup_x|d[f_n^*(x)-f(x)]/dx|\le b_n^2c_0.$$

Let $S_n(x)=n[f_n^*(x)-f(x)]$. If $q>2$, by (50) of Lemma 10, there exists $C<\infty$ such that $\limsup_{n\to\infty}|S_n(x)|/\sqrt{2n\log\log n}\le C<\infty$ almost surely. Hence (46) follows. The case that $q=2$ similarly follows from (ii) of Lemma 10. ♢

Lemma **9**. *Assume (4), (10) and (11). Then for any* $-\infty < l < u < \infty$, *we have*

$$\mathbb{E}\left[\max_{l\le x\le u}(|N_n(x)|^q + |N'_n(x)|^q)\right] = O(n^{-q/2}) \tag{47}$$

and

$$\sup_{x\in[l,u]}[|N_n(x)| + |N'_n(x)|] = o_{\text{a.s.}}[\iota_q(n)/\sqrt{n}]. \tag{48}$$

Proof. We only consider $q > 2$ since the case $q = 2$ can be similarly handled. By Lemma 3 and (34) of Proposition 1, (11) entails $\max_{l\le x\le u}\|N_n(x)\|_q = O(1/\sqrt{n})$ and $\max_{l\le x\le u}\|N'_n(x)\|_q = O(1/\sqrt{n})$. Since $N_n(x) = N_n(l) + \int_l^x N'_n(t)dt$,

$$\begin{aligned}\mathbb{E}\left[\max_{l\le x\le u}|N_n(x)|^q\right] &= O\{\mathbb{E}[|N_n(l)|^q]\} + O\left\{\mathbb{E}\left[\int_l^u |N'_n(x)|dx\right]^q\right\}\\ &= O(n^{-q/2}) + O\left\{(u-l)\mathbb{E}\int_l^u |N'_n(x)|^q dx\right\} = O(n^{-q/2}).\end{aligned}$$

Similarly, $\mathbb{E}[\max_{l\le x\le u}|N'_n(x)|^q] = O(n^{-q/2})$. Then (47) follows. Let $G_n(x) = n[F^*_n(x) - F(x)]$. By Lemma 4, (47) implies that

$$\begin{aligned}\sum_{k=4}^{\infty}\frac{\mathbb{E}\left[\max_{n\le 2^k}\max_{l\le x\le u}|G_n(x)|^q\right]}{2^{qk/2}\iota_q^q(2^k)} &= \sum_{k=4}^{\infty}\frac{O[\sum_{j=0}^k 2^{(k-j)/q}2^{j/2}]^q}{2^{qk/2}\iota_q^q(2^k)}\\ &= \sum_{k=4}^{\infty}O\{[\iota_q(2^k)]^{-q}\} = \sum_{k=4}^{\infty}\frac{O(1)}{k(\log k)^2} < \infty,\end{aligned}$$

Then by the Borel-Cantelli lemma, $\max_{l\le x\le u}|G_n(x)| = o_{\text{a.s.}}[\iota_q(n)\sqrt{n}]$, which in conjunction with the similar claim $\max_{l\le x\le u}|G'_n(x)| = o_{\text{a.s.}}[\iota_q(n)\sqrt{n}]$ entail (48). ◊

6.4 Limit theorems for $F_n - F$.

Lemma **10**. *(i) Assume (4) and (5) for some* $q > 2$. *Then for every* x, *there exist* $0 \le \sigma_1, \sigma_2 < \infty$ *such that*

$$\limsup_{n\to\infty} \pm\frac{\sqrt{n}[F_n(x) - F(x)]}{\sqrt{2\log\log n}} = \sigma_1 \tag{49}$$

and

$$\limsup_{n\to\infty} \pm\frac{\sqrt{n}[f^*_n(x) - f(x)]}{\sqrt{2\log\log n}} = \sigma_2 \tag{50}$$

almost surely for either choice of sign. (ii) Assume (4) and (8). Then for every x,

$$|F_n(x) - F(x)| + |f_n^*(x) - f(x)| = o_{\text{a.s.}}[\ell_2(n)/\sqrt{n}]. \tag{51}$$

PROOF. (i) It is a direct consequence of (35) of Proposition 1 and Lemma 3. (ii) Let $R_n(x) = n[F_n(x) - F(x)]$. By (8) and (34) of Proposition 1, $\|R_n(x)\| = O(\sqrt{n})$. Then by Lemma 4,

$$\sum_{k=4}^{\infty} \frac{\mathbb{E}[\max_{n\le 2^k} |R_n(x)|^2]}{2^k \ell_2^2(2^k)} = \sum_{k=4}^{\infty} \frac{O[\sum_{j=0}^{k} 2^{(k-j)/2} 2^{j/2}]^2}{2^k k^3 (\log k)^2} = \sum_{k=4}^{\infty} \frac{O(k^2 2^k)}{2^k k^3 (\log k)^2} < \infty,$$

which entails $|R_n(x)| = o_{\text{a.s.}}[\ell_2(n)\sqrt{n}]$ by the Borel-Cantelli lemma. That $|f_n^*(x) - f(x)| = o_{\text{a.s.}}[\ell_2(n)/\sqrt{n}]$ similarly follows. ♢

LEMMA **11**. *Let $q = 2$ and assume that the conditions of Theorem 2 are satisfied. Then $\{\sqrt{n}[F_n(x) - F(x)],\ \xi_{p_0} \le x \le \xi_{p_1}\} \Rightarrow \{W(x),\ \xi_{p_0} \le x \le \xi_{p_1}\}$ for some centered Gaussian process W in the Skorokhod space $D[\xi_{p_0}, \xi_{p_1}]$.*

PROOF. It suffices to verify the finite dimensional convergence and the tightness (Billingsley, 1968). By Lemma 3, $\|\mathcal{P}_0 \mathbf{1}_{X_n \le x}\| = O[|a_n|^{\min(\alpha/2,1)}]$, which is summable in view of (11) since $q = 2$. Then by the Cramér-Wold device, the finite dimensional convergence easily follows from Lemma 3 in Wu (2003a).

Write $l = \xi_{p_0}$ and $u = \xi_{p_1}$. Recall $F_n(x) - F(x) = M_n(x) + N_n(x)$. To show the tightness of $\{\sqrt{n}[F_n(x) - F(x)],\ l \le x \le u\}$, it suffices to show that both $\{\sqrt{n}N_n(x),\ l \le x \le u\}$ and $\{\sqrt{n}M_n(x),\ l \le x \le u\}$ are tight. The former easily follows from

$$\mathbb{E}\left[\sup_{|x-y|\le\delta,\ l\le x,y\le u} n|N_n(x) - N_n(y)|^2\right] \le \delta^2 n \mathbb{E}\left[\sup_{l\le\theta\le u} |f_n^*(\theta) - f(\theta)|^2\right] \le C\delta^2$$

in view of (47) of Lemma 9 with $q = 2$. For the latter, let $d_i = \mathbf{1}_{x<X_i\le y} - \mathbb{E}(\mathbf{1}_{x<X_i\le y}|\mathcal{F}_{i-1})$, $l \le x < y \le u$. Then by (4), $\mathbb{E}(d_i^2|\mathcal{F}_{i-1}) \le C(y-x)$. Here C denotes a constant which does not depend on n, x and y and it may vary from line to line. By Burkholder's inequality (Chow and Teicher, 1988),

$$\mathbb{E}[n^2|M_n(x) - M_n(y)|^4] \le \frac{C}{n^2}\left\|\sum_{i=1}^{n} d_i^2\right\|^2$$

$$
\begin{aligned}
&\leq \frac{C}{n^2}\left\|\sum_{i=1}^{n}(d_i^2-\mathbb{E}(d_i^2|\mathcal{F}_{i-1})\right\|^2+\frac{C}{n^2}\left\|\mathbb{E}(d_i^2|\mathcal{F}_{i-1})\right\|^2 \\
&\leq \frac{C}{n}\|d_1^2-\mathbb{E}(d_1^2|\mathcal{F}_0)\|^2+C(y-x)^2 \\
&\leq \frac{C}{n}(y-x)+C(y-x)^2.
\end{aligned}
$$

See inequality (48) in Wu (2003a) for a similar claim. Therefore, by the argument of Theorem 22.1 in Billingsley (pages 197-199, 1968), the process $\{\sqrt{n}M_n(x),\ l \leq x \leq u\}$ is tight. $\diamondsuit$

REMARK **10**. Under conditions of the type given in (8), Wu (2003b) obtained a central limit theorem for $S_n(K)/\sqrt{n}$, where $S_n(K)=\sum_{i=1}^n[K(X_i)-\mathbb{E}K(X_i)]$, K is a measurable function and ε_i may have infinite variance. $\diamondsuit$

LEMMA **12**. *Let $X_k=\sum_{i=0}^\infty a_i\varepsilon_{k-i}$ and assume (4) and $\mathbb{E}(|\varepsilon_k|^\alpha)<\infty$ for some $\alpha>0$. Further assume (42) and $\log^3 n=o(nb_n)$. (i) If (5) holds with $q>2$, then for every fixed x,*

$$
\begin{aligned}
&\sup_{|u|\leq b_n}|F_n(x+u)-F(x+u)-[F_n(x)-F(x)]| \\
&=\frac{O_{\text{a.s.}}(\sqrt{b_n\log\log n})}{\sqrt{n}}+\frac{O_{\text{a.s.}}[b_n\ell_q(n)]}{\sqrt{n}}+O_{\text{a.s.}}(b_n^2) \qquad (52)
\end{aligned}
$$

(ii) If (8) holds, then we have (52) with $q=2$.

REMARK **11**. The second term $O_{\text{a.s.}}[b_n\ell_q(n)]/\sqrt{n}$ in the bound of (52) is needed only when $q=2$. $\diamondsuit$

Lemma 12 follows from Lemmas 6 and 8 and it provides a local fluctuation rate of empirical processes for linear processes. The last two terms of (52) are due to the presence of dependence, in the sense that they disappear if X_k are iid. Actually, if X_i are iid, then $F_n^*\equiv F$ and hence $N_n\equiv 0$.

LEMMA **13**. *Assume (4), (10) and (11). Then under the conditions of Lemma 7, we have for any $-\infty<l<u<\infty$ that*

$$
\sup_{|x-y|\leq b_n,\ x,y\in[l,u]}|[F_n(x)-F(x)]-[F_n(y)-F(y)]|=O_{\text{a.s.}}\left[\frac{\sqrt{b_n\log n}}{\sqrt{n}}+\frac{b_n\ell_q(n)}{\sqrt{n}}\right]. \qquad (53)
$$

Proof. By Lemma 7, it suffices to show that

$$\begin{aligned}\sup_{|x-y|\le b_n,\ x,y\in[l,u]} |[F_n^*(x)-F(x)]-[F_n^*(y)-F(y)]| &\le b_n \sup_{\theta\in[l,u]} |f_n^*(\theta)-f(\theta)| \\ &= b_n o_{\mathrm{a.s.}}[\iota_q(n)/\sqrt{n}],\end{aligned}$$

which is an easy consequence of Lemma 9. $\diamondsuit$

6.5 Proof of Theorem 1.

We only consider $q > 2$ since the case $q = 2$ follows along similar lines. (i) Let $b_n = \delta_{n,q}$. Then (42) holds. By Lemma 12, there exists a constant $C_1 < \infty$ such that

$$|[F_n(\xi_p + b_n) - F(\xi_p + b_n)] - [F_n(\xi_p) - F(\xi_p)]| \le C_1\sqrt{(b_n \log\log n)/n}.$$

almost surely. Observe that $F(\xi_p+b_n) = F(\xi_p)+b_n f(\xi_p)+O(b_n^2)$ in view of (4) via Taylor's expansion. By (i) of Lemma 10, there exists a constant $C_2 < \infty$ such that $n|F_n(\xi_p) - F(\xi_p)| \le C_2\sqrt{n}\ell_q(n)$ almost surely. Choose $C > 0$ such that $C - C_2 - C_1\sqrt{C/f(\xi_p)} \ge 1$, namely $C \ge [C_1/\sqrt{f(\xi_p)} + \sqrt{C_1^2/f(\xi_p) + 4(1 + C_2)}]^2/4$. Then for $b_n = C\ell_q(n)/[f(\xi_p)\sqrt{n}]$, $F_n(\xi_p + b_n) > p$ holds almost surely. The other statement that $p > F_n(\xi_p - b_n)$ almost surely similarly follows. Let $\Delta_n = \xi_{n,p} - \xi_p$. Since F_n is non-decreasing, by (6), $|\Delta_n| \le b_n$ almost surely.

(ii) The argument for Theorem 4 can be applied here. Applying Lemma 12 with $x = \xi_p$, we have

$$|F_n(\xi_{n,p}) - F(\xi_p + \Delta_n) - [F_n(\xi_p) - F(\xi_p)]| = O_{\mathrm{a.s.}}[\sqrt{(b_n \log\log n)/n}].$$

Notice that $|F_n(\xi_{n,p}) - p| \le 1/n$ and, by Taylor's expansion, $F(\xi_p + \Delta_n) = p + \Delta_n f(\xi_p) + O(\Delta_n^2)$ since $\sup_x |f'(x)| < \infty$. Then

$$\Delta_n f(\xi_p) = p - F_n(\xi_p) + O_{\mathrm{a.s.}}[\sqrt{(b_n \log\log n)/n}]$$

and it entails (7). $\diamondsuit$

6.6 Proof of Theorem 2.

Let $l = \xi_{p_0}$ and $u = \xi_{p_1}$. By Lemma 6 and (48) of Lemma 9, we have

$$\begin{aligned} \sup_{x\in[l,u]} |F_n(x) - F(x)| &\le \sup_{x\in[l,u]} |F_n(x) - F_n^*(x)| + \sup_{x\in[l,u]} |F_n^*(x) - F(x)| \\ &= O_{\text{a.s.}}\left[\frac{\sqrt{\log\log n}}{\sqrt{n}}\right] + o_{\text{a.s.}}\left[\frac{\iota_q(n)}{\sqrt{n}}\right] = o_{\text{a.s.}}\left[\frac{\iota_q(n)}{\sqrt{n}}\right]. \end{aligned} \tag{54}$$

Let $b_n = \iota_q(n)/\sqrt{n}$. (i) By Lemma 13,

$$\begin{aligned} &\inf_{l\le x\le u}[F_n(x+b_n) - F(x)] \ge \inf_{l\le x\le u}[F(x+b_n) - F(x)] \\ &- \sup_{l\le x\le u} |F_n(x) - F(x)| - \sup_{|x-y|\le b_n,\ l\le x,y\le u} |[F_n(x) - F(x)] - [F_n(y) - F(y)]| \\ &\ge b_n \inf_{p_0\le p u} f(\xi_p) + O(b_n^2) + o_{\text{a.s.}}(b_n) + O_{\text{a.s.}}[\sqrt{b_n(\log n)/n} + b_n\iota_q(n)/\sqrt{n}]. \end{aligned}$$

Hence $\inf\{F_n(x+b_n) - F(x) :\ l \le x \le u\} > 0$ almost surely, which implies (i) together with a similar claim that $\sup\{F_n(x-b_n) - F(x) :\ l \le x \le u\} < 0$ almost surely. The representation (12) then follows from Lemma 13 by using the same argument as in the proof of (ii) of Theorem 1. ♢

7 Proof and the sharpness of Theorem 3.

In the study of LRD processes, the asymptotic expansion of empirical processes plays an important role [Ho and Hsing (1996), Wu (2003a)]. Let $U_{n,r} = \sum_{0\le j_1<\ldots<j_r} \prod_{s=1}^r a_{j_s}\varepsilon_{n-j_s}$, $U_{n,0} = 1$. For a nonnegative integer ρ, similarly to equation (4) in Wu (2003a) let

$$S_n(y;\rho) = \sum_{i=1}^n \left[\mathbf{1}(X_i \le y) - \sum_{r=0}^{\rho}(-1)^r F^{(r)}(y)U_{i,r}\right];$$

see also Ho and Hsing (1996). The quantity $S_n(y;\rho)$ can be viewed as the remainder of the ρth order expansion of $F_n(y)$. In our derivation of Bahadur's representation for LRD processes, we only deal with $\rho = 1$ and do not pursue the higher order case $\rho \ge 2$ since it involves some really cumbersome manipulations.

As in Wu (2003a), let $\theta_n = |a_{n-1}|[|a_{n-1}| + (\sum_{i=n-1}^\infty a_i^2)^{1/2} + (\sum_{i=n-1}^\infty a_i^4)^{\rho/2}]$, $\Theta_n = \sum_{i=1}^n \theta_i$, $\Xi_n = n\Theta_n^2 + \sum_{i=1}^\infty(\Theta_{n+i} - \Theta_i)^2$. Since $\rho = 1$, $\theta_n = O[|a_{n-1}|(\sum_{i=n-1}^\infty a_i^2)^{1/2}]$. Recall that $\Psi_n = \sqrt{n}\sum_{k=1}^n k^{1/2-2\beta}L^2(k)$, $A_n(\beta) = \Psi_n^2(\log n)(\log\log n)^2$ if $\beta < 3/4$ and

$A_n(\beta) = \Psi_n^2(\log n)^3(\log\log n)^2$ if $\beta \geq 3/4$. Let $H_n(y) = n[F_n^*(y) - F(y) + f(y)\bar{X}_n]$ and $h_n(y) = dH_n(y)/dy$.

LEMMA **14**. *Assume* $\mathbb{E}(\varepsilon_i^4) < \infty$ *and*

$$\sup_x |f_\varepsilon(x)| + \sup_x |f'_\varepsilon(x)| + \int_{\mathbb{R}} |f'_\varepsilon(u)|^2 du < \infty. \tag{55}$$

Then

$$\left\| \sup_y |H_n(y)| \right\| + \left\| \sup_y |h_n(y)| \right\| = O(\Psi_n). \tag{56}$$

PROOF. Let $I = \int_R |f'_\varepsilon(u)|^2 du$ and $K_\theta(x) = [f_\varepsilon(\theta - x) - f_\varepsilon(\theta)]/\sqrt{I}$. Then $k_\theta(x) = \partial K_\theta(x)/\partial x = -f'_\varepsilon(\theta - x)/\sqrt{I}$ satisfies $\int_{\mathbb{R}} k_\theta^2(x)dx = 1$. Hence for all θ,

$$K_\theta \in \mathcal{K}(0) := \left\{ K(x) = \int_0^x g(t)dt : \quad \int_{\mathbb{R}} g^2(t) \leq 1 \right\};$$

see Wu (2003a) for the definition of the class $\mathcal{K}$. By Theorem 1 in Wu (2003a), for

$$S_n(K_\theta, 1) = \frac{1}{\sqrt{I}} \sum_{i=1}^n [f_\varepsilon(\theta - X_{i,i-1}) - f(\theta) + f'(\theta)X_{i,i-1}],$$

we have

$$\mathbb{E}\left[\sup_{\theta\in\mathbb{R}} S_n^2(K_\theta, 1)\right] = \mathcal{O}(\Xi_n).$$

Notice that $S_n(K_\theta, 1)\sqrt{I} - h_n(\theta) = -f'(\theta)\sum_{i=1}^n \varepsilon_i$. Then $\| \sup_y |h_n(y)|\| = O(\Xi_n^{1/2})$ since $\sup_\theta |f'(\theta)| < \infty$ and $\| \sum_{i=1}^n X_{i,i-1} - n\bar{X}_n\| = O(\sqrt{n})$. By Karamata's Theorem, it is easily seen that $\Xi_n = O(\Psi_n^2)$ (cf Lemma 5 in Wu (2003a)). Similarly, $\| \sup_y |H_n(y)|\| = O(\Xi_n^{1/2})$ holds under the condition $\int_{\mathbb{R}} f_\varepsilon^2(u)du < \infty$. The last inequality trivially holds since $\sup_u f_\varepsilon(u) < \infty$. ♢

LEMMA **15**. *Assume* $\mathbb{E}(\varepsilon_i^4) < \infty$ *and (55). (i) Let* $(\delta_n)_{n\geq 1}$ *be a positive, bounded sequence such that* $\log n = o(n\delta_n)$. *Then*

$$\sup_{|x-y|\leq\delta_n} |S_n(y;1) - S_n(x;1)| = O_{\text{a.s.}}[\sqrt{n\delta_n \log n} + \delta_n A_n^{1/2}(\beta)]. \tag{57}$$

(ii) For any $-\infty < l < u < \infty$, $\sup_{l\leq y\leq u} |S_n(y;1)| = o_{\text{a.s.}}[A_n^{1/2}(\beta)]$.

PROOF. (i) By Lemma 7, since $\mathbb{E}(X_1^2) < \infty$, $\sqrt{n}\sup_{|x-y|\le\delta_n}|M_n(y)-M_n(x)| = O_{\text{a.s.}}(\sqrt{\delta_n\log n})$. To show (57), notice that $\sup_{|x-y|\le\delta_n}|H_n(y)-H_n(x)| \le \delta_n\sup_\theta|h_n(\theta)|$, it suffices to verify that $\sup_\theta|h_n(\theta)| = o_{\text{a.s.}}[A_n^{1/2}(\beta)]$ in view of

$$S_n(y;1) - S_n(x;1) = n[M_n(y)-M_n(x)] + [H_n(y)-H_n(x)]. \tag{58}$$

By Karamata's Theorem, $\sum_{j=0}^d 2^{(d-j)/2}\Psi_{2^j} = O(\Psi_{2^d})$ if $\beta < 3/4$ and $\sum_{j=0}^d 2^{(d-j)/2}\Psi_{2^j} = O(d\Psi_{2^d})$ if $\beta \ge 3/4$. So it follows from Lemma 14 that

$$\sum_{j=0}^d 2^{(d-j)/2}\left\|\sup_y|h_{2^j}(y)|\right\| = \sum_{j=0}^d 2^{(d-j)/2}O(\Psi_{2^j}) = \frac{O[A_{2^d}^{1/2}(\beta)]}{d^{1/2}\log d},$$

which in conjunction with Lemma 4 implies

$$\sum_{d=3}^\infty \frac{1}{A_{2^d}(\beta)}\mathbb{E}\left[\max_{j\le 2^d}\sup_y|h_j(y)|^2\right] = \sum_{d=3}^\infty O(d^{-1}\log^{-2}d) < \infty.$$

Hence $\sup_y|h_n(y)| = o_{\text{a.s.}}[\sqrt{A_n(\beta)}]$ via the Borel-Cantelli lemma.

(ii) Notice that $S_n(y;1) = nM_n(y) + H_n(y)$. By Lemma 7, $\sqrt{n}\sup_{l\le y\le u}|M_n(y)| = O_{\text{a.s.}}(\sqrt{\log n})$. Using the argument in (i), (56) implies $\sup_y|H_n(y)| = o_{\text{a.s.}}[\sqrt{A_n(\beta)}]$. Hence (ii) follows in view of $\sqrt{n} = O(\Psi_n)$ and $\sqrt{n\log n} = o[\sqrt{A_n(\beta)}]$. ♢

LEMMA **16**. *Assume $\mathbb{E}(\varepsilon_i^4) < \infty$ and (55). Let $B_n = \sigma_{n,1}(\log n)^{1/2}(\log\log n)$, $b_n = B_n/n$ and $\Delta_{n,p} = \xi_{n,p} - \xi_p$. Then (i) $\bar{X}_n = o_{\text{a.s.}}(b_n)$ and (ii) if, in addition, $\inf_{p_0\le p\le p_1} f(\xi_p) > 0$ for some $0 < p_0 < p_1 < 1$, we have*

$$\sup_{p_0\le p\le p_1}|\Delta_{n,p}| = o_{\text{a.s.}}(b_n) \tag{59}$$

and

$$\sup_{p_0\le p\le p_1}|\Delta_{n,p} - \bar{X}_n| = o_{\text{a.s.}}(b_n^2) + o_{\text{a.s.}}[n^{-1}A_n^{1/2}(\beta)]. \tag{60}$$

PROOF. (i) Let $S_n = \sum_{i=1}^n X_i$. Since $\sigma_{n,1} = \|S_n\| \sim Cn^{3/2-\beta}L(n)$, by Lemma 4,

$$B_{2^d}^{-2}\left\|\max_{i\le 2^d}|S_i|\right\|^2 \le B_{2^d}^{-2}\left[\sum_{r=0}^d 2^{(d-r)/2}\sigma_{2^r,1}\right]^2 = O(d^{-1}\log^{-2}d).$$

Again by the Borel-Cantelli lemma, $\bar{X}_n = o_{\text{a.s.}}(b_n)$.

(ii) Similarly as in the proof of Theorem 1, it suffices to show that, due to the monotonicity of F_n, $\inf_{p_0 \le p \le p_1}[F_n(\xi_p + b_n) - p] > 0$ holds almost surely since the other inequality $\sup_{p_0 \le p \le p_1}[F_n(\xi_p - b_n) - p] < 0$ can be similarly derived. By Lemma 15,

$$\begin{aligned}
\inf_{p_0 \le p \le p_1}[F_n(\xi_p + b_n) - p] &\ge \inf_{p_0 \le p \le p_1}[F(\xi_p + b_n) - f(\xi_p + b_n)\bar{X}_n - p + S_n(\xi_p; 1)/n] \\
&\quad - \sup_{|x-y| \le b_n} |S_n(y; 1) - S_n(x; 1)|/n =: I_n + II_n
\end{aligned}$$

Since $\sup_x |f'_\varepsilon(x)| < \infty$, by Taylor's expansion, $\sup_p |F(\xi_p + b_n) - p - b_n f(\xi_p)| = O(b_n^2)$ and $\sup_p |f(\xi_p + b_n) - f(\xi_p)| = O(b_n)$. Let $l = \xi_{p_0}$ and $u = \xi_{p_1}$. By (ii) of Lemma 15, $\sup_{l \le x \le u} |S_n(x; 1)| = o_{\text{a.s.}}[A_n^{1/2}(\beta)]$. By (i), $\bar{X}_n = o_{\text{a.s.}}(b_n)$. Therefore,

$$\begin{aligned}
I_n &= \inf_{p_0 \le p \le p_1} f(\xi_p)(b_n - \bar{X}_n) + O(b_n^2 + b_n|\bar{X}_n|) + o_{\text{a.s.}}[A_n^{1/2}(\beta)/n] \\
&\ge \frac{1}{2} \inf_{p_0 \le p \le p_1} f(\xi_p) b_n
\end{aligned}$$

almost surely. By (57) of Lemma 15, $II_n = o_{\text{a.s.}}(b_n)$ and hence (59) holds. Relation (60) follows by letting $y = \xi_{n,p} = \xi_p + \Delta_{n,p}$ in (ii) of Lemma 15 in view of

$$\sup_{p_0 \le p \le p_1} |F(\xi_p + \Delta_{n,p}) - p - f(\xi_p)\Delta_{n,p}| \le \frac{\sup_x |f'(x)|}{2} \sup_{p_0 \le p \le p_1} \Delta_{n,p}^2 = o_{\text{a.s.}}(b_n^2)$$

and $\sup_{p_0 \le p \le p_1} |f(\xi_p + \Delta_{n,p}) - f(\xi_p)| = o_{\text{a.s.}}(b_n)$. ♢

Remark **12**. Under the stronger condition that f_ε is 4 times differentiable with bounded, continuous and integrable derivatives, Ho and Hsing (1996) obtained

$$\sup_{p_0 \le p \le p_1} |\Delta_{n,p} - \bar{X}_n| = o_{\text{a.s.}}(n^{-1-\lambda}\sigma_{n,1}) \tag{61}$$

for all $0 < \lambda < \min(1 - \beta, \beta - 1/2)$; see Theorem 5.1 therein. The result (61) is very interesting in the sense that $\Delta_{n,p}$ can be approximated by $\bar{X}_n$, which *does not* depend on p. Consequently the asymptotic distribution of the trimmed and Winsorized means easily follows from that of $\bar{X}_n$. After elementary calculations, it is easily seen that our bound (60) is slightly sharper. ♢

7.1 Proof of Theorem 3.

By (59), $\sup_{p_0 \le p \le p_1} |\Delta_{n,p}| = o_{\text{a.s.}}(b_n)$. Applying Lemma 15 with $x = \xi_p$ and $y = \xi_{n,p}$, $p_0 \le p \le p_1$, we have

$$\begin{aligned} & n \sup_{p_0 \le p \le p_1} |p - F(\xi_p + \Delta_{n,p}) + f(\xi_{n,p})\bar{X}_n - [F_n(\xi_p) - F(\xi_p) + f(\xi_p)\bar{X}_n]| \\ & = O_{\text{a.s.}}[\sqrt{nb_n \log n} + b_n A_n^{1/2}(\beta)]. \end{aligned} \tag{62}$$

Since $\sup_x[|f'(x)| + |f''(x)|] < \infty$, by Taylor's expansion,

$$\sup_{p_0 \le p \le p_1} |F(\xi_p + \Delta_{n,p}) - p - \Delta_{n,p} f(\xi_p) - \Delta_{n,p}^2 f'(\xi_p)/2| = o_{\text{a.s.}}(b_n^3)$$

and

$$\sup_{p_0 \le p \le p_1} |f(\xi_p + \Delta_{n,p}) - f(\xi_p) - \Delta_{n,p} f'(\xi_p)| = o_{\text{a.s.}}(b_n^2).$$

After some elementary calculations, (62) implies

$$\begin{aligned} & \sup_{p_0 \le p \le p_1} \left| f(\xi_p)\Delta_{n,p} + \frac{f'(\xi_p)}{2}(\Delta_{n,p} - \bar{X}_n)^2 - \frac{1}{2} f'(\xi_p)\bar{X}_n^2 - [p - F_n(\xi_p)] \right| \\ & = o_{\text{a.s.}}(b_n^3) + n^{-1} O_{\text{a.s.}}[\sqrt{nb_n \log n} + b_n A_n^{1/2}(\beta)]. \end{aligned}$$

Observe that $\Psi_n = O[\sqrt{n}L^*(n) + n^{2-2\beta}L^2(n)]$ and $A_n^{1/2}(\beta) \le \Psi_n(\log n)^{3/2}(\log\log n) = o(nb_n)$. Thus (15) follows from (60) and

$$\begin{aligned} \sup_{p_0 \le p \le p_1} (\Delta_{n,p} - \bar{X}_n)^2 &= o_{\text{a.s.}}[b_n^2 + A_n^{1/2}(\beta)/n]^2 \\ &= o_{\text{a.s.}}[b_n^4 + A_n(\beta)/n^2] = o_{\text{a.s.}}[b_n^3 + b_n A_n^{1/2}(\beta)/n]. \end{aligned}$$

$\diamondsuit$

7.2 The sharpness of Theorem 3.

It is challenging to obtain a sharp bound for the left hand side of (15) in Theorem 3. We now comment on the sharpness of Lemma 15, which describes the oscillations of $F_n(x) - F(x) + f(x)\bar{X}_n$. Recall that in the SRD case, the sharp oscillation rate of $F_n(x) - F(x)$ at a fixed x in Lemma 12 leads to the Bahadur representation with optimal bound by letting $b_n = c\sqrt{(\log\log n)/n}$ for some $c > 0$. Here we claim that, the bound in (57) of Lemma 15, which is a key ingredient for the derivation of (15), is optimal up to a multiplicative slowly varying function.

LEMMA **17**. *Assume* $\mathbb{E}(\varepsilon_i^4) < \infty$, *(14) and* $\int_{\mathbb{R}} |f_\varepsilon''(u)|^2 du < \infty$. *Let* $\delta_n = n^\gamma L_2(n)$ *for some slowly varying function* L_2, $-1 < \gamma < 0$, *and* $\sigma_{n,2} = n^{2-2\beta} L^2(n)$. *(i) If* $4\beta - 3 > \gamma$, *then* $[S_n(x+\delta_n;1) - S_n(x;1)]/\sqrt{n\delta_n} \Rightarrow N[0, f(x)]$. *(ii) If* $4\beta - 3 < \gamma$, *then*

$$\frac{S_n(x+\delta_n;1) - S_n(x;1)}{\sigma_{n,2}\delta_n} \Rightarrow$$
$$f'(x) C_\beta \int_{u_1<u_2<1} \int_0^1 [(v-u_1)_+ (v-u_2)_+]^{-\beta} dv d\mathbb{B}(u_1) d\mathbb{B}(u_2) \quad (63)$$

for some constant $C_\beta > 0$, *where* $\mathbb{B}$ *is a standard two-sided Brownian motion and* $z_+ = \max(z, 0)$. *In particular, if* $\gamma = 1/2 - \beta$, *then (i) [resp. (ii)] holds if* $7/10 < \beta < 1$ *[resp.* $1/2 < \beta < 7/10$*].*

REMARK **13**. The limiting distribution in (63) is called the Rosenblatt distribution, a special case of multiple Wiener-Itô integrals [Major (1981)]. ♢

PROOF OF LEMMA 17. Observe that

$$S_n(x+\delta_n;1) - S_n(x;1) = n[M_n(x+\delta_n) - M_n(x)] + [H_n(x+\delta_n) - H_n(x)].$$

Write $n[F_n(x+\delta_n) - F_n(x)] = \sum_{i=1}^n K[(x - X_i)/\delta_n]$, where the kernel $K(u) = \mathbf{1}_{-1\le u\le 0}$. By Lemma 2 in Wu and Mielniczuk (2002), $n[M_n(x+\delta_n) - M_n(x)]/\sqrt{n\delta_n} \Rightarrow N[0, \sigma^2(x)]$ with $\sigma^2(x) = f(x)\int_{\mathbb{R}} K^2(u)du = f(x)$. By Lemma 14,

$$\begin{aligned} \|H_n(x+\delta_n) - H_n(x)\| &\le \delta_n \left\| \sup_y |h_n(y)| \right\| \\ &= O(\delta_n \Psi_n) = \delta_n O[\sqrt{n} L^*(n) + n^{2-2\beta} L^2(n)]. \end{aligned}$$

If $4\beta - 3 > \gamma$, then $\delta_n \Psi_n = o(\sqrt{n\delta_n})$ and (i) follows.

On the other hand, if $4\beta - 3 < \gamma$, then $\beta \in (1/2, 3/4)$ and $h_n(x)/\sigma_{n,2}$ converges to the Rosenblatt distribution in (63); see Lemma 4 in Wu and Mielniczuk (2002) and Corollary 3 in Wu (2003a). Under the conditions (14) and $\int_{\mathbb{R}} |f_\varepsilon''(u)|^2 du < \infty$, by the argument of Lemma 14, we have $\|\sup_u |h_n'(u)|\| = O(\Psi_n)$. Then $\|H_n(x+\delta_n) - H_n(x) - \delta_n h_n(x)\| \le \frac{1}{2}\delta_n^2 \|\sup_u |h_n'(u)|\| = O(\delta_n^2 \Psi_n)$ and (ii) follows in view of $\sqrt{n\delta_n} + \delta_n^2 \Psi_n = o(\delta_n \sigma_{n,2})$.

If $\gamma = 1/2 - \beta$, then $4\beta - 3 > \gamma$ if and only if $7/10 < \beta$. ♢

Lemma 17 asserts the dichotomous convergence of $S_n(x+\delta_n;1) - S_n(x;1)$ at a fixed point x. Notice that $\bar{X}_n/[n^{1/2-\beta}L(n)] \Rightarrow N(0, \sigma^2)$ and, by (60), $(\xi_{n,p} - \xi_p)/[n^{1/2-\beta}L(n)] \Rightarrow$

$N(0,\sigma^2)$ for some $\sigma^2 < \infty$. For $\delta_n = n^{1/2-\beta}L_2(n)$, Lemma 17 shows that, up to a multiplicative slowly varying function, the optimal bound of $[S_n(x+\delta_n;1) - S_n(x;1)]/n$ is $n^{\max(-\beta/2-1/4,\ 3/2-3\beta)}$. This bound indicates that (15) or (16) is optimal up to a multiplicative slowly varying function. It also explains why there is a boundary $\beta = 7/10$ in (15) or (16); see the discussion of the three terms in the $O_{\text{a.s.}}$ bound of (15) in Section 3.

Acknowledgments. I am grateful to the referees and an Associate Editor for their many helpful comments. I would also like to thank Jan Mielniczuk, George Michailidis and Xiaofeng Shao for valuable suggestions.

References

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

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

Chow, Y. S. and Teicher, H. (1988). *Probability Theory*, 2nd ed. Springer, New York.

Deheuvels, P. (1997). Strong laws for local quantile processes. *Ann. Probab.* **25** 2007-2054.

Deheuvels P. and Mason, D. M. (1992). Functional laws of the iterated logarithm for the increments of empirical and quantile processes. *Ann. Probab.* **20** 1248-1287.

Dehling, H. and Taqqu, M. (1989). The empirical process of some long-range dependent sequences with an application to U-statistics. *Ann. Statist.* **17** 1767-1783.

Dehling, H., T. Mikosch and M. Sorensen (eds), (2002). *Empirical Process Techniques for Dependent Data*, Boston: Birkhäuser.

Diaconis, P. and Freedman. D. (1999). Iterated random functions. *SIAM Rev.* 41, 41–76.

Doob, J. (1953). *Stochastic Processes.* Wiley.

Doukhan, P. and Surgailis, D. (1998). Functional central limit theorem for the empirical process of short memory linear processes. *C. R. Acad. Sci. Paris Ser. I Math.* **326** 87–92.

Einmahl, J. H. J. (1996). A short and elementary proof of the main Bahadur-Kiefer theorem. *Ann. Probab.* **24** 526-531.

Elton, J. H. (1990). A multiplicative ergodic theorem for Lipschitz maps. *Stochastic Process. Appl.* **34** 39–47.

Freedman, D. A. (1975). On tail probabilities for martingales. *Ann. Probab.* **3** 100–118.

Gordin, M. I. (1969). The central limit theorem for stationary processes. *Dokl. Akad. Nauk SSSR* **188** 739741.

Gordin, M. I. and Lifsic, B. (1978). The central limit theorem for stationary Markov processes. *Soviet. Math. Dokl.* **19** 392394.

Hall, P. and Heyde, C. C. (1980). *Martingale Limit Theory and its Applications.* Academic Press, New York.

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

Ho, H. C. and Hsing, T. (1996). On the asymptotic expansion of the empirical process of long-memory moving averages. *Ann. Statist.* **24** 992–1024.

Hsing, T. and Wu, W. B. (2004). On weighted U-statistics for stationary processes. *Ann. Probab.* **32** 1600–1631.

Kiefer, J. (1967). On Bahadur's representation of sample quantiles. *Ann. Math. Statist.* **38** 1323–1342.

Kiefer, J. (1970a). Deviations between the sample quantile process and the sample df. In *Nonparametric Techniques in Statistical Inference* (M. L. Puri, ed.) 299–319. Cambridge Univ. Press, London.

Kiefer, J. (1970b). Old and new methods for studying order statistics and sample quantiles. In *Nonparametric Techniques in Statistical Inference* (M. L. Puri, ed.) 349–357. Cambridge Univ. Press, London.

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.* Berlin; New York: Springer.

Moricz, F. (1976). Moment inequalities and strong laws of large numbers, *Z. Wahrsch. Verw. Gebiete* **35** 299–314.

Ruymgaart, F. H. (1994). Conditional empirical, quantile and difference processes for a large class of time series with applications, *J. Statist. Plann. Inference*, **40** 15–31.

Sen, P. K. (1968). Asymptotic normality of sample quantiles for m-dependent processes. *Ann. Math. Statist.* **39** 1724–1730.

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

Serfling, R. J. (1970). Moment inequalities for the maximum cumulative sum. *Ann. Math. Statist.* **41** 1227–1234.

Shorack, G. R. and Wellner, J. A. (1986). *Empirical processes with applications to statistics.* John Wiley & Sons, New York.

Stigler, S. (1973). The asymptotic distribution of the trimmed mean. *Ann. Statist.* **1** 472–477.

Wu, W. B. (2003a). Empirical processes of long-memory sequences. *Bernoulli* **9** 809–831.

Wu, W. B. (2003b). Additive functionals of infinite-variance moving averages. *Statist. Sinica* **13** 1259–1267.

Wu, W. B. and Mielniczuk, J. (2002). Kernel density estimation for linear processes. *Ann. Statist.* **30** 1441–1459.

Wu, W. B. and Woodroofe, M., (2000). A central limit theorem for iterated random functions. *J. Appl. Probab.* **37** 748–755.

Wu, W. B. and Woodroofe, M. (2004). Martingale approximations for sums of stationary processes. *Ann. Probab.* **32** 1674–1690.

Wu, W. B. and Shao, X. (2004). Limit theorems for iterated random functions. *J. Appl. Probab.* **41** 425–436.

Wu, W. B. (2004). On strong convergence for sums of stationary processes. *Preprint.*

Department of Statistics,

The University of Chicago,

5734 S. University Avenue, Chicago, IL 60637

wbwu@galton.uchicago.edu