Statistica Sinica **12**(2002), 635-649

CENTRAL LIMIT THEOREMS FOR FUNCTIONALS OF LINEAR PROCESSES AND THEIR APPLICATIONS

Wei Biao Wu

University of Chicago

Abstract: This paper establishes central limit theorems and invariance principles for functionals of one-sided linear processes. These results are applied to long-range dependent sequences whose covariances are summable but not absolutely summable. We also consider empirical processes and 0-crossings for linear processes whose innovations may have infinite variance. Comparisons with earlier results are indicated.

Key words and phrases: Central limit theorem, invariance principle, long- and short-range dependence, linear process, 0-crossings, empirical process, martingale.

1. Introduction

In this paper, we consider central limit theorems for additive functionals of a moving-average process $X_n = \sum_{i=0}^{\infty} a_i \varepsilon_{n-i}, a_i \in \mathbb{R}, n \in \mathbb{Z}$, where $\{\varepsilon, \varepsilon_i, i \in \mathbb{Z}\}$ are i.i.d. random variables. The classical ARMA models with causality are typical examples of such processes. Consider a univariate function or instantaneous transformation $K : \mathbb{R} \mapsto \mathbb{R}$ satisfying $\mathbb{E}[K(X_0)] = 0$, and define the partial sum $S_n(K) = \sum_{i=1}^{n} K(X_i)$. In statistical inference for time series, it is of critical importance to know the asymptotic behavior of $S_n(K)$. This problem has been considered by many authors and we only provide a brief account of some of them. Earlier papers mainly deal with the case where the sequence $\{X_n\}$ is Gaussian. A well known one is by Taqqu (1975), in which central and non-central limit theorems are proved by using the Hermite expansion of the function $K(\cdot)$. For the case of non-normal innovations ε_i, prior works focus on some special forms of $K(\cdot)$. For example, Davydov (1970) considers the special case $K(x) = x$, which is also discussed in Phillips and Solo (1992), while Giraitis and Surgailis (1986) consider Appell polynomials. The recent work of Ho and Hsing (1996, 1997) represents the first attempt to deal with general univariate functions. Moreover, in Hsing (1999) the case where the innovations obey stable laws is analyzed.

An important feature of Ho and Hsing's method is briefly presented next. The authors discovered an interesting martingale structure for the one-sided linear process:

$$K(X_0) - \mathbb{E}[K(X_0)] = \sum_{i=0}^{\infty} \{\mathbb{E}[K(X_0)|\mathbf{E}_{-i}] - \mathbb{E}[K(X_0)|\mathbf{E}_{-i-1}]\},$$

where $\mathbf{E}_i = (\ldots, \varepsilon_{i-1}, \varepsilon_i)$. It is clear that this decomposition induces a martingale difference sequence which, surprisingly, had not received much attention in the earlier literature. However, their martingale structure does not seem to work well for non-instantaneous transformations.

Let the additive functional $K(\cdot)$ be l-variate for some fixed $l \in \mathbb{N}$, and define $S_{n,l}(K) = \sum_{i=1}^n K(X_{i-l+1}, \ldots, X_{i-1}, X_i)$. For example, the 0-crossings of $\{X_k, 1 \le k \le n\}$ require $l = 2$, corresponding to the kernel $K(x, y) = \mathbf{1}_{[xy \le 0]} - \mathbb{P}(X_0 X_1 \le 0)$ on the plane, where $\mathbf{1}_E$ is the indicator function of event E. Furthermore, sample covariances involve multivariate K. Ho and Sun (1987) obtain central limit theorems for $S_{n,l}(K)$ when $\{X_n\}$ is stationary Gaussian.

The recent work of Maxwell and Woodroofe (2000) provides new insight into asymptotic normality for additive functionals of general Markov chains. It is shown that the sufficient conditions proposed therein are almost necessary for asymptotic normality with $\sqrt{n}$-norming. We apply their result to handle $S_{n,l}(K)$ by taking advantage of the linear structure of moving-average processes. In the special case $l = 1$, the results are comparable to those by Ho and Hsing (1997) when the norming sequence is $\sqrt{n}$. However, we are able to obtain limit theorems under conditions which appear to be simpler than theirs, although they have different ranges of applications.

Our general result allows us to consider long-range dependent (LRD) sequences with spectral density having multiple singularities away from the origin, which extends the concept of the so-called *cyclic* fractionally integrated autoregressive moving-average (FARIMA) models (See Robinson (1997) and Gray, Zhang and Woodward (1989)). In this case, we derive central limit theorems with a $\sqrt{n}$-norming sequence.

The paper is organized as follows. Notation and main results are given in Section 2. These results are applied in Section 3 to LRD sequences whose covariances are summable, but not absolutely summable. Section 4 focuses on empirical processes and 0-crossings, where asymptotic normality and invariance principles are derived under some smoothness assumptions on the characteristic function of ε. The proofs of results in Section 2 are given in Section 5. Section 6 discusses further developments and conjectures.

2. Notation and Main results

Let $X_n = \sum_{i=0}^\infty a_i \varepsilon_{n-i}$ be a one-sided linear process, where $\{\varepsilon_i, i \in \mathbb{Z}\}$ are i.i.d. innovations. We assume throughout the paper that $\mathbb{E}\varepsilon = 0$, $\mathbb{E}\varepsilon^2 = 1$ except in Section 4, where infinite variance is allowed. By Kolmogorov's Three Series Theorem, X_n exists almost surely if and only if the sequence $\{a_i, i = 0, 1, \ldots\}$ satisfies $\sum_{i=0}^\infty a_i^2 < \infty$. We denote its tail by $A_t = \sum_{i=t}^\infty a_i^2$. When $n \ge 1$, define the truncated processes $X_{n,+}$ and $X_{n,-}$ of X_n by

$$X_n = \sum_{i=0}^{n-1} a_i \varepsilon_{n-i} + \sum_{i=n}^{\infty} a_i \varepsilon_{n-i} =: X_{n,+} + X_{n,-}. \qquad (1)$$

Denote the Lebesgue shift process by $\mathbf{E}_n = (\ldots, \varepsilon_{n-1}, \varepsilon_n)$; then $X_{n,-} = \mathbb{E}[X_n|\mathbf{E}_0]$. Let $\pi_l, l \in \mathbb{N}$, be the joint distribution of $(X_1, \ldots, X_l)$. Furthermore, for $k \in \mathbb{Z}$, define the l-dimensional random vectors $W_k = (X_{k-l+1}, \ldots, X_k)$. The dimension l is assumed to be fixed throughout the paper. Next let $W_{k,+} = (X_{k-l+1,+}, \ldots, X_{k,+})$ and $W_{k,-} = (X_{k-l+1,-}, \ldots, X_{k,-})$ be the corresponding truncated vectors. In this paper, we are interested in the asymptotic distributions of $S_n(K) = K(W_1) + \ldots + K(W_n)$, where $K \in \mathcal{L}_0^2(\pi_l)$ and

$$\mathcal{L}_0^2(\pi_l) = \Big\{K(\cdot) : \mathbb{E}[K(W_0)] = \int_{\mathbb{R}^l} K(w)\pi_l\{dw\} = 0, \int_{\mathbb{R}^l} K^2(w)\pi_l\{dw\} < \infty\Big\}.$$

We denote by $\|\cdot\|_p, p \geq 1$, the norm in $\mathcal{L}^p$, i.e., $\|X\|_p = [\mathbb{E}(|X|^p)]^{1/p}$, and $\|\cdot\| = \|\cdot\|_2$. Let $|x - y|$ be the usual Euclidean distance. Clearly, there exists a measurable function g on $\mathbb{R}^{\mathbb{M}}$, where $\mathbb{M} = (\ldots, -1, 0)$, such that $g(\mathbf{E}_n) = K(W_n)$. We introduce the transition operator

$$(Q^n g)(\mathbf{e}) = \mathbb{E}[K(W_n)|\mathbf{E}_0 = \mathbf{e}] = \mathbb{E}[g(\mathbf{E}_n)|\mathbf{E}_0 = \mathbf{e}], \quad \mathbf{e} \in \mathbb{R}^{\mathbb{M}}, \qquad (2)$$

where $\{Q^i, i = 0, 1, \ldots\}$ forms a semi-group due to the Markov property of $\mathbf{E}_n$. The partial potential operator V_n is given by

$$(V_n g)(\mathbf{e}) = \sum_{k=0}^{n-1} (Q^k g)(\mathbf{e}) = \mathbb{E}[S_n(K)|\mathbf{E}_1 = \mathbf{e}]. \qquad (3)$$

When $l = 1$, similarly to Ho and Hsing (1997) we define $K_n(x) = \mathbb{E}[K(X_{n,+}+x)]$, and then $(Q^n g)(\mathbf{E}_0)$ is nothing but $K_n(X_{n,-})$. A similar definition exists for the multivariate case:

$$K_n(y_1, \ldots, y_l) = \mathbb{E}[K(X_{n-l+1,+} + y_1, X_{n-l+2,+} + y_2, \ldots, X_{n,+} + y_l)] \qquad (4)$$

with $(Q^n g)(\mathbf{E}_0) = K_n(X_{n-l+1,-}, X_{n-l+2,-}, \ldots, X_{n,-})$.

Let $\Delta(F, G) = \inf\{\epsilon > 0 : G(x - \epsilon) - \epsilon \leq F(x) \leq G(x + \epsilon) + \epsilon \text{ for all } x \in \mathbb{R}\}$ be the Lévy distance between two distribution functions. Denote by $\rho(\cdot, \cdot)$ the Prokhorov metric between probability measures on the space $D[0, 1]$ (see Billingsley (1968) for definitions). Let $F_n(\mathbf{e}; z)$ be the conditional probability $\mathbb{P}[n^{-1/2}S_n(K) \leq z|\mathbf{E}_0 = \mathbf{e}]$, and $F_n^{\#}(\mathbf{e}; t), t \in [0, 1)$, be the distribution of the partial sum process $\mathbb{B}_n^{\mathbf{e}}(t) = n^{-1/2}S_{\lceil nt \rceil}(K)$ given $\mathbf{E}_0 = \mathbf{e}$, where $\mathbb{B}_n^{\mathbf{e}}(1) = \mathbb{B}_n^{\mathbf{e}}(1-)$, and where $\lceil z \rceil$ denotes the smallest integer not less than z. We adopt the standard notation $I\!B$ for a standard Brownian motion on $[0,1]$ and $N(0,\sigma^2)$ for a normal random variable with mean zero and variance σ^2. Maxwell and Woodroofe (2000) derive the following theorem, it is a significant improvement over the classical result by Gordin and Lifsic (1978).

Theorem 1. (Maxwell and Woodroofe). *Let $\{\mathbf{E}_n, n \in \mathbb{Z}\}$ be a stationary ergodic Markov chain, and define $S_n(g) = \sum_{i=1}^n g(\mathbf{E}_i)$, where $\mathbb{E}[g(\mathbf{E}_1)] = 0, \mathbb{E}[g^2(\mathbf{E}_1)] < \infty$. If*

$$\sum_{n=1}^{\infty} n^{-\frac{3}{2}} \|(V_n g)(\mathbf{E}_0)\| < \infty, \tag{5}$$

then $\sigma^2 = \lim_{n\to\infty} n^{-1}\mathbb{E}[S_n^2(g)]$ exists and is finite, and

$$\lim_{n\to\infty} \mathbb{E}\{\Delta[N(0,\sigma^2), F_n(\mathbf{E}_0;\cdot)]\} = 0. \tag{6}$$

Moreover, if there exist $p > 2$ and $\kappa < 1/2$ such that $\mathbb{E}[|g(\mathbf{E}_1)|^p] < \infty$ and $\|(V_n g)(\mathbf{E}_0)\| = O(n^\kappa)$, then

$$\lim_{n\to\infty} \mathbb{E}\{\rho[\sigma I\!B, F_n^{\#}(\mathbf{E}_0)]\} = 0. \tag{7}$$

Theorem 1 in conjunction with (4) readily yields the corollary below. When $l = 1$, the corollary and the central limit theorem for a short-range dependent process $\{X_n\}$ given in Ho and Hsing (1997) have different ranges of applications. The latter paper assumes the existence of the derivatives of $K_n(\cdot)$ as well as the finiteness of the fourth moments of local maxima of $|dK_n(t)/dt|$, while our corollary requires no differentiability. However only $\sum_{n=1}^\infty |a_n| < \infty$ is imposed in their paper. This condition is weaker than $\sum_{n=1}^\infty n^{-1/2}\sqrt{A_n} < \infty$ by Lemmas 1 and 2 if $K_n(\cdot)$ satisfies (9) with $\alpha = \beta = 1$. The difference between the two conditions on the summability of a_n is minor in view of Remark 1 below.

Corollary 1. *If $\sum_{n=1}^\infty n^{-1/2}\|K_n(W_{n,-})\| < \infty$, then we have (6) and hence $S_n(K)/\sqrt{n} \Rightarrow N(0,\sigma_K^2)$. If in addition $\mathbb{E}[|K_n(W_1)|^p] < \infty$ and $\|K_n(W_{n,-})\| = \mathcal{O}(n^{\kappa-1})$ for some $p > 2$ and $\kappa < 1/2$, then (7) holds and $\{n^{-1/2}S_{\lceil nt\rceil}(K), 0 \le t \le 1\} \Rightarrow \{\sigma_K I\!B(t), 0 \le t \le 1\}$.*

Lemma 1. *If $a_n \ge 0$ for all $n \in \mathbb{N}$, then*

$$\sum_{n=1}^{\infty} a_n \le 3\sum_{n=1}^{\infty} n^{-1/2}\sqrt{A_n}. \tag{8}$$

Remark 1. If a_n is of the form $n^{-\gamma}L(n), \gamma \ge 1$, for some slowly varying function $L(\cdot)$, then $\sum_{n=1}^\infty |a_n| < \infty$ implies $\sum_{n=1}^\infty n^{-1/2}\sqrt{A_n} < \infty$. This is obviously true for $\gamma > 1$. For $\gamma = 1$, by Karamata's theorem (see Bingham, Goldie and Teugels (1987)), $A_n \sim L^2(n)n^{-1}$ and then the equivalence is clear again.

Various sufficient conditions ensuring (5) in Theorem 1 are presented below. Specifically, Lemma 2 gives bounds for $\|(Q^n g)(\mathbf{E}_0)\|$, which leads to a central limit theorem for short-range dependent (SRD) sequences (*cf.* Corollary 2). Theorems 2 and 3 provide bounds on $\|(V_n g)(\mathbf{E}_0)\|$ with applications to some special LRD sequences (*cf.* Section 3). The proofs of Theorems 2 and 3 are deferred to Section 5. Recall (4) for the definition of K_n.

Lemma 2. *Suppose that there exist* $0 < \alpha \le 1 \le \beta < \infty$ *with* $\mathbb{E}[|\varepsilon|^{2\beta}] < \infty$. *Further suppose that either* (a)

$$C_n(\alpha, \beta) := \sup_{x \neq 0} \frac{|K_n(x) - K_n(0)|}{|x|^\alpha + |x|^\beta} < M < \infty \tag{9}$$

holds for all sufficiently large n, *or* (b)

$$\mathbb{E}[M^2_{\alpha,\beta}(W_1)] < \infty \tag{10}$$

holds, where $M_{\alpha,\beta}(x) = \sup_{y \neq x} |K(x) - K(y)|/(|x-y|^\alpha + |x-y|^\beta)$. *Then as* $n \to \infty$, $\|(Q^n g)(\mathbf{E}_0)\| = \mathcal{O}(A^{\alpha/2}_{n-l})$.

Corollary 2. *Assume that* (9) *or* (10) *holds with* $\alpha = 1$ *and* $\mathbb{E}[|\varepsilon|^{2\beta}] < \infty$ *for some* $\beta \ge 1$. *Let* $a_n = n^{-\gamma} L(n)$ *for some* $\gamma > 1$. *Then we have* (6), *and hence* $n^{-1/2} S_n(K) \Rightarrow N(0, \sigma_K^2)$ *for some* $\sigma_K^2 \in [0, \infty)$. *If in addition,* $\mathbb{E}[|K(W_1)|^p] < \infty$ *for some* $p > 2$, *then* (7) *holds and* $\{n^{-1/2} S_{\lceil nt \rceil}(K), 0 \le t \le 1\} \Rightarrow \{\sigma_K I\!\!B(t), 0 \le t \le 1\}$.

Let $\{\varepsilon_i', i \in \mathbb{Z}\}$ be another sequence of innovations such that $\{\varepsilon_i, \varepsilon_i', i \in \mathbb{Z}\}$ are i.i.d., and the coupled version $X_n' = X_{n,+} + X_{n,-}' := \sum_{i=0}^{n-1} a_i \varepsilon_{n-i} + \sum_{i=n}^{\infty} a_i \varepsilon_{n-i}'$; let $W_{k,l}' = (X_{k-l+1}', \ldots, X_k')$. Then $W_{k,l}'$ is independent of $\mathbf{E}_0$. Let the gradient $\nabla K(x_1, \ldots, x_l) = (\partial K/\partial x_1, \ldots, \partial K/\partial x_l)^{\mathbf{T}}$, $^{\mathbf{T}}$ the transpose of a vector. Write $s_n = \sum_{i=0}^n a_i$.

Theorem 2. *Suppose that* K *has gradient* ∇K *satisfying*

$$\frac{\|K(W_k) - K(W_k') - (W_k - W_k')\nabla K(W_k)\|}{\|W_k - W_k'\|^2} \le M < \infty \tag{11}$$

for all $k \ge n_0$. *Then*

$$\|(V_n g)(\mathbf{E}_0)\| = \mathcal{O}\Big[M \sum_{k=1}^n A_k\Big] + \mathcal{O}[\lambda_n |\mathbb{E}\nabla K(W_1)|], \tag{12}$$

$$\lambda_n := \Big[\sum_{i=0}^\infty (s_{n+i} - s_i)^2\Big]^{\frac{1}{2}}. \tag{13}$$

Remark 2. If K is linear, then M in (11) is 0 and the first term on the right side of (12) vanishes. If K is quadratic, then $\mathbb{E}\nabla K(W_1)$ in (12) becomes 0.

Theorem 3. *Suppose that* K *has gradient* ∇K *with*

$$M(x) := \sup_{y\neq x} \frac{|\nabla K(x) - \nabla K(y)|}{|x-y|^{\alpha} + |x-y|^{\beta}} \tag{14}$$

satisfying $\mathbb{E}[M^2(W_1)] < \infty$ *for some* $0 < \alpha \le 1 \le \beta < \infty$. *If* $\mathbb{E}[|\varepsilon|^{2+2\beta}] < \infty$, *then*

$$\|(V_n g)(\mathbf{E}_0)\| = \mathcal{O}\Big[\sum_{k=1}^{n} A_k^{(1+\alpha)/2}\Big] + \mathcal{O}[\lambda_n |\mathbb{E}\nabla K(W_1)|], \tag{15}$$

where λ_n *is defined in* (13).

3. LRD Processes with Summable Covariances

A particularly interesting case is when the λ_n defined in (13) satisfies

$$\lambda := \sup_{n\in\mathbb{N}} \lambda_n < \infty. \tag{16}$$

Then the second term in (12) or (15) contributes only $\mathcal{O}(1)$ to $\|(V_n g)(\mathbf{E}_0)\|$. Proposition 1 provides an equivalent statement of the finiteness of λ defined in (16). Proposition 2 asserts that (16) implies the summability of the covariances. Some special sequences are constructed in Propositions 3 and 4 without absolute summability of the covariances.

Proposition 1. *The quantity* λ *in* (16) *is finite if and only if the sum* $\sum_{i=1}^{\infty} a_i$ *exists and*

$$\sum_{i=1}^{\infty} \Big[\sum_{j=i}^{\infty} a_j\Big]^2 < \infty. \tag{17}$$

Remark 3. Condition (17) appears as inequality (3.57) in Hall and Heyde (1980, page 146) for a one-sided linear process.

Recall that ε_n are i.i.d. with mean 0 and variance 1. Let $a_t = 0$ for $t < 0$. Then $\Gamma(k) = \sum_{t=-\infty}^{\infty} a_t a_{t+k}$ is the covariance function of X_n.

Proposition 2. *Suppose that* (16) *holds. Then* $\lim_{k\to\infty} \sum_{t=-k}^{k} \Gamma(k) = (\sum_{i=0}^{\infty} a_i)^2$.

Proposition 3. *Let* $\{b_n, n \in \mathbb{N}\}$ *be a square summable sequence that converges non-increasingly to* 0 *when* $n \ge n_0$ *for some* $n_0 \in \mathbb{N}$. *Then the sequence* $a_n = b_n \cos(\theta_1 + n\omega_1)$, *where* $\omega_1 \ne 0$ *(mod* 2π*), is summable and satisfies* (17), *hence* λ *is finite.*

Proposition 4. *Suppose that* $a_n = n^{-\gamma}L(n)\cos(\theta_1 + n\omega_1)$ *for* $n \in \mathbb{N}$, *where* $1/2 < \gamma < 1, \omega_1 \neq 0$ $(mod\, 2\pi)$ *and that*

$$\sum_{j=n}^{\infty} \frac{|L(j+1) - L(j)|}{j^\gamma} = \mathcal{O}\left[\frac{L_1(n)}{n^\gamma}\right] \tag{18}$$

for some slowly varying function $L_1(n)$. *Then* λ *defined in* (16) *is finite.*

We call a function $L(\cdot)$ *very slowly varying* if (18) holds for all $1/2 < \gamma < 1$. Clearly, if $L(\cdot)$ is monotone, then it is very slowly varying. A simple example for a not very slowly varying function is given by $L(n) = 1 + (-1)^n/\log n, n \geq 2$. Let $L(x)$ be positive for sufficiently large x. Recall that $L(x)$ is *normalized* or in the *Zygmund Class* if for all $\epsilon > 0$, $x^\epsilon L(x)$ is ultimately increasing and $x^{-\epsilon}L(x)$ is ultimately decreasing (see Bingham *et al* (1987, page 24) for a definition and basic properties). The following lemma shows that a normalized slowly varying function is very slowly varying.

Lemma 3. *If there exists an* $\epsilon > 0$ *and an* $x_0 > 0$ *such that* $x^\epsilon L(x)$ *is increasing and* $x^{-\epsilon}L(x)$ *is decreasing when* $x > x_0$, *then* (18) *holds for all* $1/2 < \gamma < 1$.

The proof of Lemma 3 is straightforward. Let $n > x_0 + 1$. By the monotonicity of L we get

$$|L(n+1) - L(n)| \leq |L(n)| \times \max\left[\left(\frac{n+1}{n}\right)^\epsilon - 1, 1 - \left(\frac{n}{n+1}\right)^\epsilon\right] = \mathcal{O}\left[\frac{L(n)}{n}\right],$$

which yields (18) by Karamata's theorem. Proposition 4 and Theorem 2 together yield.

Corollary 3. *Let* a_n *be the sequence defined in Proposition* 4 *with* $3/4 < \gamma < 1$. *Suppose that* K *satisfies* (11). *Then we have* (6), *and hence* $n^{-1/2}S_n(K) \Rightarrow N(0, \sigma_K^2)$ *for some* $\sigma_K^2 \in [0, \infty)$. *If, in addition,* $\mathbb{E}[|K(W_1)|^p] < \infty$ *for some* $p > 2$, *then* (7) *holds and* $\{n^{-1/2}S_{\lceil nt \rceil}(K), 0 \leq t \leq 1\} \Rightarrow \{\sigma_K I\!B(t), 0 \leq t \leq 1\}$.

3.1. Examples

If $\omega_i \neq 0$ (mod 2π), $1 \leq i \leq I$, then the sequence $a_n = \sum_{i=1}^I b_{n,i}\cos(\theta_i + n\omega_i)$ satisfies (16) when $b_{n,i}$ is either eventually monotone as in Proposition 3, or is of the form $n^{-\gamma}L(n)$ with a slowly varying function $L(n)$ satisfying (18).

Let $b_{n,i} = n^{-\gamma_i}, 1 \leq i \leq I$, where $1/2 < \gamma_i < 1$, and $a_n = \sum_{i=1}^I b_{n,i}\cos(\theta_i + n\omega_i)$, where $\omega_i \neq 0$ (mod 2π). Then the associated spectral density function $f(\omega) = |\sum_{t=0}^\infty a_t \exp(\sqrt{-1}\omega t)|^2/(2\pi)$ has poles at $\omega = \omega_i$. This linear process has multiple singularities away from 0. By Proposition 2, the covariances of such processes are summable. However, they are not absolutely summable. To see this, take $I = 1$ and $a_0 = 1$, $a_n = n^{-\gamma}\cos(n\omega), 1/2 < \gamma < 1$. Then $\Gamma(k) =$

$c_k k^{1-2\gamma}\cos(k\omega)$, where $c_k \to c \neq 0$. Hence the $\Gamma(k)$ are oscillatory and not absolutely summable. So X_k is LRD. Theorem 2 can be applied to guarantee the asymptotic normality of $\sum_{t=1}^n X_t/\sqrt{n}$ by taking $K(x) = x$.

The spectral density of a cyclic FARIMA (Robinson (1997)) has one pole away from 0.

4. Empirical Processes and 0-crossings

Let $\phi(\cdot)$ be the characteristic function of ε, i.e., $\phi(t) = \mathbb{E}[\exp(t\varepsilon\sqrt{-1})], t \in \mathbb{R}$. In this section, ε is allowed to have infinite variance. Suppose that there exists $0 < \delta \leq 2$ for which $\mathbb{E}(|\varepsilon|^\delta) < \infty$ and $\mathbb{E}\varepsilon = 0$ when $1 \leq \delta \leq 2$. Then by the Kolmogorov Three Series Theorem (see Corollary 5.1.3 in Chow and Teicher (1988)), $X_n = \sum_{i=0}^\infty a_i\varepsilon_{n-i}$ exists almost surely if $\sum_{t=0}^\infty |a_t|^\delta < \infty$. Set $A_k(\delta) := \sum_{t=k}^\infty |a_t|^\delta$. For any fixed $s_1, \ldots, s_l \in \mathbb{R}$, let $S_n(K) = \sum_{t=1}^n K(W_t)$, where $K_{s_1,\ldots,s_l}(x_1, \ldots, x_l) := \mathbf{1}_{[x_1\leq s_1,\ldots,x_l\leq s_l]} - \mathbb{P}(X_1 \leq s_1, \ldots, X_l \leq s_l)$.

Empirical processes of linear processes have been discussed by several authors. Here we only mention a few recent results. Ho and Hsing (1996) derive asymptotic expansions of the empirical process of long range dependent linear processes, while Giraitis, Koul, and Surgailis (1996) obtain functional non-central limit theorems. Using the martingale difference decomposition presented in Section 1, Giraitis and Surgailis (1999) recently establish central limit theorems for the empirical processes. Hsing (1999) discusses general functionals K and symmetric α stable (SαS) innovations. All these papers deal with the univariate case. The following theorem examines the limiting behavior for multivariate empirical processes.

Theorem 4. *Suppose there exists an $r \in \mathbb{N}$ such that $\int_{-\infty}^\infty |\phi(t)|^r dt < \infty$, and that $\#\{i : a_i \neq 0\} = \infty$. Then $\|K_n(W_{n,-})\|^2 = \mathcal{O}[A_{n-l}(\delta)]$. Hence* (a) *if $\sum_{n=1}^\infty [A_n(\delta)/n]^{1/2} < \infty$, then $S_n/n^{1/2} \Rightarrow N(0, \sigma_K^2)$ for some $\sigma_K = \sigma_K(s_1, \ldots, s_l) \geq 0$;* (b) *if in addition $A_n(\delta) = \mathcal{O}(n^{-q})$ for some $q > 1$, then $\{S_{\lceil nt\rceil}/\sqrt{n}, 0 \leq t \leq 1\} \Rightarrow \sigma_K \mathbb{B}$.*

Before proving Theorem 4, let us discuss its conditions.

Remark 4. It is easy to see that $\int_{-\infty}^\infty |\phi(t)|^r dt < \infty$ for $r > 2/\delta + 2$ if there exist constants $C, \delta > 0$ such that $|\phi(t)| \leq C/(1 + |t|)^\delta$ for all $t \in \mathbb{R}$, as in Giraitis *et al.* (1996) and Giraitis *et al.* (1999). The aforementioned inequality places a rather weak restriction on the smoothness of the distribution function of ε.

Remark 5. If $\#\{i : a_i \neq 0\} < \infty$, then X_n is an m-dependent sequence (see Hoeffding and Robbins (1948)). Hence the central limit theorems become a direct consequence.

Proof of Theorem 4. Choose $i(1) < i(2) < \ldots < i(r)$ such that $a_* = \min\{|a_{i(j)}| : 1 \le j \le r\} > 0$. Then by the inversion formula, for $n > i(r)$, the density function $f_n(x)$ of $X_{n,+}$ satisfies

$$\begin{aligned}\sup_{x\in\mathbb{R}} f_n(x) &\le \frac{1}{2\pi}\int_{\mathbb{R}}\prod_{i=0}^{n}|\phi(a_i t)|dt \le \frac{1}{2\pi}\int_{\mathbb{R}}\prod_{j=1}^{r}|\phi(a_{i(j)}t)|dt \\ &\le \frac{1}{2\pi}\prod_{j=1}^{r}\left\{\int_{\mathbb{R}}|\phi[a_{i(j)}t]|^r dt\right\}^{1/r} \le \frac{1}{2\pi a_*}\left[\int_{\mathbb{R}}|\phi(t)|^r dt\right]^{1/r} < \infty.\end{aligned}$$

Take $n > i(r) + l$ and let $F_{n,l}(\cdot)$ be the joint distribution function of $W_{n,+}$ with density function $f_{n,l}(\cdot)$. Then by the form of $K(\cdot)$, we have that

$$K_n(x_1, \ldots, x_l) = F_{n,l}(s_1 - x_1, \ldots, s_l - x_l) - \mathbb{P}(X_1 \le s_1, \ldots, X_l \le s_l).$$

Define the event $E_i(x_i) = \{X_{n-l+i,+} \le s_i - x_i\}$. Then

$$\begin{aligned}&|K_n(x_1, \ldots, x_l) - K_n(0, \ldots, 0)| \\ &\le \mathbb{E}|\mathbf{1}_{E_1(x_1)\cap\ldots E_l(x_l)} - \mathbf{1}_{E_1(0)\cap\ldots E_l(0)}| \\ &\le \sum_{i=1}^{l}\mathbb{E}|\mathbf{1}_{E_i(x_i)} - \mathbf{1}_{E_i(0)}| \le \sum_{i=1}^{l}\int_{s_i-|x_i|}^{s_i+|x_i|} f_{n-l+i}(x)dx = \mathcal{O}(|x|)\end{aligned}$$

since $X_{n-l+i,+}$ has a density uniformly bounded for all sufficiently large n. Observe that K is bounded by 1, $|K_n(W_{n,-}) - K_n(0, \ldots, 0)| = \mathcal{O}[\min(1, |W_{n,-}|)]$. Recall that $W'_{n,-}$ and $W_{n,-}$ are i.i.d., and $|K_n(W_{n,-}) - K_n(W'_{n,-})| \le |K_n(W_{n,-}) - K_n(0, \ldots, 0)| + |K_n(W'_{n,-}) - K_n(0, \ldots, 0)|$, then $\|K_n(W_{n,-})\|^2 = \|K_n(W_{n,-}) - K_n(W'_{n,-})\|^2/2 = \mathcal{O}(1)\,\mathbb{E}\{[\min(1, |W_{n,-}|)]^2\} = \mathcal{O}(1)\,\mathbb{E}[|W_{n,-}|^\delta] = \mathcal{O}[A_{n-l}(\delta)]$, where the last step is due to the claim $\mathbb{E}[|X_{n,-}|^\delta] = \mathcal{O}[A_n(\delta)]$. This claim is obvious when $0 < \delta \le 1$. Now assume $1 < \delta \le 2$. Then by the Burkholder-Davis-Gundy inequality (*cf.* Theorem 11.3.1 in Chow and Teicher (1988)), there exists $C > 0$ such that $\mathbb{E}[|X_{n,-}|^\delta] \le C\mathbb{E}\{[\sum_{t=n}^{\infty}|a_t\varepsilon_{n-t}|^2]^{\delta/2}\} \le C\sum_{t=n}^{\infty}\mathbb{E}[|a_t\varepsilon_{n-t}|^2]^{\delta/2} = \mathcal{O}[A_n(\delta)]$. The rest follows from Corollary 1.

The next corollary deals with the 0-crossings of a SRD sequence $\{X_n\}$. Let $K(x, y) = \mathbf{1}_{[xy\le 0]} - \mathbb{P}(X_0X_1 \le 0)$ and $N_n(K)$ be the number of times the sequence $\{X_k, 1 \le k \le n\}$ crosses 0. Then $S_n(K) = N_n(K) - n\mathbb{P}(X_0X_1 \le 0)$. There is a substantial history of 0-crossings with various applications to signal processing, biomedical engineering, seismology, etc. For example, Niederjohn and Castelaz (1978) discuss 0-crossing analysis (ZCA) method for speech sound classification and show its effectiveness for the characterization of speech sounds. Further references to this problem can be found in Slud (1994). Kedem (1994) suggests that the 0-crossing analysis can provide an alternative approach in time series analysis to the popular time-domain approach based on the auto-covariance function and frequency-domain approach based on spectral density. Malevich (1969) and Cuzick (1976) derive central limit theorems. All these articles deal with stationary Gaussian sequences X_n. Corollary 4 may have application to the statistical inference based on 0-crossing analysis.

Corollary 4. *Assume the conditions of Theorem* 4 *hold. Then there exists* $\sigma_K \geq 0$ *such that* $\{S_{\lceil nt\rceil}/\sqrt{n}, 0 \leq t \leq 1\} \Rightarrow \sigma_K I\!B$.

Proof of Corollary 4. We compute again, for all $x, y \in \mathbb{R}$,

$$\begin{aligned} K_n(x,y) &= \mathbb{E}\mathbf{1}_{[(X_{n-1,+}+x)(X_{n,+}+y)\leq 0]} - \mathbb{P}(X_0X_1 \leq 0) \\ &= F_{n-1}(-x) + F_n(-y) - 2F_{n,2}(-x,-y) - \mathbb{P}(X_0X_1 \leq 0). \end{aligned}$$

Then $|K_n(x,y)-K_n(0,0)| \leq |F_{n-1}(-x)-F_{n-1}(0)|+|F_n(-y)-F_n(0)|+2|F_{n,2}(-x, -y) - F_{n,2}(0,0)|$, which satisfies (9) by the first assertion of Theorem 4.

5. Proofs

This section provides the proofs of results given in Section 2. We first need the following simple lemma.

Lemma 4. *Assume that* $\mathbb{E}[\varepsilon] = 0, \mathbb{E}[\varepsilon^2] < \infty$, *and* $\mathbb{E}[|\varepsilon|^{2p}] < \infty, p > 0$. *Then* $\mathbb{E}[|X_{n,-}|^{2p}] = \mathcal{O}(A_n^p)$.

Proof of Lemma 4. If $0 < p \leq 1$, by Hölder's inequality we have $\mathbb{E}[|X_{n,-}|^{2p}] \leq \{\mathbb{E}[|X_{n,-}|^2]\}^p = \mathcal{O}(A_n^p)$. If $p > 1$, the Marcinkiewicz-Zygmund inequality (see Corollary 10.3.3, Chow and Teicher (1988)) yields $\mathbb{E}(|X_{n,-}|^{2p}) = \mathcal{O}\{\mathbb{E}[\sum_{i=0}^\infty (a_{n+i}\varepsilon_{-i})^2]^p\}$, which is $\mathcal{O}(A_n^p)$ via Minkowski's inequality $\|\sum_{i=0}^\infty (a_{n+i}\varepsilon_{-i})^2\|_p \leq \sum_{i=0}^\infty \|(a_{n+i}\varepsilon_{-i})^2\|_p = \mathcal{O}(A_n)$.

Proof of Lemma 1. We assume without loss of generality that the sequence $a_n, n \in \mathbb{N}$ is non-increasing, since if $a_{n_0} < a_{n_0+1}$ for some $n_0 \in \mathbb{N}$, then the right hand side of (8) will be smaller while the left remains the same by exchanging a_{n_0} with a_{n_0+1}. Now (8) follows immediately from $A_n \geq na_{2n}^2$ and $\sum_{n=1}^\infty n^{-1/2}\sqrt{A_n} \geq \sum_{n=1}^\infty a_{2n} \geq \frac{1}{2}\sum_{n=2}^\infty a_n$.

Proof of Lemma 2. We apply coupling to prove this. Assume $n > l$. (a) First observe that $\mathbb{E}[K(W_n')|\mathbf{E}_0] = \mathbb{E}[K(W_n')] = 0$ since W_n' and $\mathbf{E}_0$ are independent. Then using $(a+b)^2/2 \leq a^2 + b^2$ and $\mathbb{E}(|Z|^p) \leq \|Z\|^p$ for $p \leq 2$, we have

$$\begin{aligned} \|(Q^n g)(\mathbf{E}_0)\|^2 &= \frac{1}{2}\mathbb{E}[|K_n(W_{n,-}) - K_n(W_{n,-}')|^2] \\ &\leq \mathbb{E}[|K_n(W_{n,-}) - K_n(0)|^2] + \mathbb{E}[|K_n(W_{n,-}') - K_n(0)|^2] \\ &\leq 4C_n^2(\alpha,\beta)\mathbb{E}[|W_{n,-}|^{2\alpha} + |W_{n,-}|^{2\beta}] \\ &\leq \mathcal{O}[\|W_{n,-}\|^{2\alpha} + \mathbb{E}(|W_{n,-}|^{2\beta})] \\ &= \mathcal{O}(A_{n-l}^\alpha) + \mathcal{O}(A_{n-l}^\beta) = \mathcal{O}(A_{n-l}^\alpha). \end{aligned}$$

The last step is due to Lemma 4. (b) By (10) and Cauchy's inequality,

$$\begin{aligned}
&|(Q^n g)(\mathbf{E}_0)|^2 \\
&= |\mathbb{E}[K(W_n) - K(W_n')|\mathbf{E}_0]|^2 \\
&\le \{\mathbb{E}[M_{\alpha,\beta}(W_n') \times (|W_{n,-} - W_{n,-}'|^\alpha + |W_{n,-} - W_{n,-}'|^\beta)|\mathbf{E}_0]\}^2 \\
&\le 2\{\mathbb{E}[M_{\alpha,\beta}^2(W_n')|\mathbf{E}_0]\}\{\mathbb{E}[|W_{n,-} - W_{n,-}'|^{2\alpha} + |W_{n,-} - W_{n,-}'|^{2\beta}|\mathbf{E}_0]\} \\
&= 2\|M_{\alpha,\beta}\|^2[\mathcal{O}(\|W_{n,-}'\|^2) + 2^{2\alpha}|W_{n,-}|^{2\alpha} + 2^{2\beta}|W_{n,-}|^{2\beta}]
\end{aligned}$$

which, in view of Lemma 4, yields $\|(Q^n g)(\mathbf{E}_0)\|^2 = \mathcal{O}(A_{n-l}^\alpha)$.

Proof of Theorem 2. Let $R_k = K(W_k) - K(W_k') - (W_{k,-} - W_{k,-}')\nabla K(W_k')$, and $\mathbf{c} = (c_1, \ldots, c_l)^{\mathbf{T}} = \mathbb{E}[\nabla K(W_1)]$. Then for $k \ge l_0 = l + n_0 + 1$,

$$(Q^k g)(\mathbf{E}_0) = \mathbb{E}[R_k|\mathbf{E}_0] + W_{k,-}\mathbf{c} - \mathbb{E}[W_{k,-}'\nabla K(W_k')] \tag{19}$$

since $\mathbb{E}[K(W_k')|\mathbf{E}_0] = 0$. By (11), $\|R_k\| \le M\|W_k - W_k'\|^2 = \mathcal{O}(MA_{k-l})$ as $k \to \infty$. So

$$\begin{aligned}
\|(V_n g)(\mathbf{E}_0)\| &\le \Big\| \sum_{k=0}^{l_0-1} Q^k g \Big\| + \Big\| \sum_{k=l_0}^{n-1} Q^k g \Big\| \\
&\le \mathcal{O}(1) + \Big\| \sum_{k=l_0}^{n-1} \{\mathbb{E}[R_k|\mathbf{E}_0] + W_{k,-}\mathbf{c} - \mathbb{E}[W_{k,-}'\nabla K(W_k')]\} \Big\| \\
&\le \mathcal{O}(1) + \sum_{k=l_0}^{n-1} A_{k-l} M\mathcal{O}(1) + \Big\| \sum_{k=l_0}^{n-1} W_{k,-}\mathbf{c} \Big\| + \sum_{k=l_0}^{n-1} |\mathbb{E}[W_{k,-}'\nabla K(W_k')]|.
\end{aligned}$$

Therefore (12) follows immediately from the inequality

$$\begin{aligned}
|\mathbb{E}[W'_{k,-}\nabla K(W_k')]| &= \frac{1}{2}|\mathbb{E}\{(W_{k,-} - W_{k,-}')[\nabla K(W_k) - \nabla K(W_k')]\}| \\
&\le \frac{1}{2}\|(W_{k,-} - W_{k,-}')[\nabla K(W_k) - \nabla K(W_k')]\| \\
&\le \frac{1}{2}[\|K(W_k) - K(W_k') - (W_k' - W_k)\nabla K(W_k)\| \\
&\qquad + \|K(W_k') - K(W_k) - (W_k - W_k')\nabla K(W_k')\|] \\
&\le \frac{1}{2} \times 2M\|W_k - W_k'\|^2 = \mathcal{O}(MA_{k-l})
\end{aligned}$$

in view of (11), and from

$$\Big\| \sum_{k=l_0}^{n-1} W_{k,-}\mathbf{c} \Big\| = \Big\| \sum_{k=l_0}^{n-1} \sum_{j=1}^{l} X_{k-l+j,-} c_j \Big\|$$

$$\leq \sum_{j=1}^{l} |c_j| \times \Big\| \sum_{k=l_0}^{n-1} X_{k-l+j,-} \Big\| = |\mathbf{c}| \times \mathcal{O}[\lambda_n + \mathcal{O}(1)],$$

thus proving the theorem.

Proof of Theorem 3. The proof of Theorem 2 can be easily adapted to this case. Let R_k be defined as in the previous proof. By the Mean Value Theorem, there exists a ξ with $|\xi - W'_k| \leq |W_k - W'_k| = |W_{k,-} - W'_{k,-}|$ such that $R_k = (W_{k,-} - W'_{k,-})[\nabla K(\xi) - \nabla K(W'_k)]$. Whence by Cauchy's inequality and the independence between W'_k and $\mathbf{E}_0$, we have

$$\begin{aligned}\mathbb{E}[|R_k||\mathbf{E}_0] &\leq \mathbb{E}[|W_{k,-} - W'_{k,-}|M(W'_k)(|\xi - W'_k|^{\alpha} + |\xi - W'_k|^{\beta})|\mathbf{E}_0] \\ &\leq \{\mathbb{E}[M^2(W'_k)]\}^{1/2}\{\mathbb{E}[(|W_{k,-}-W'_{k,-}|^{1+\alpha}+|W_{k,-}-W'_{k,-}|^{1+\beta})^2|\mathbf{E}_0]\}^{1/2},\end{aligned}$$

which ensures $\|\mathbb{E}[|R_k||\mathbf{E}_0]\|^2 = \mathcal{O}(A_{k-l}^{1+\alpha})$ via Lemma 4. Similarly we have

$$|\mathbb{E}[W'_{k,-}\nabla K(W'_k)]| = |\mathbb{E}\{W'_{k,-}[\nabla K(W_k) - \nabla K(W'_k)]\}| = \mathcal{O}(A_{k-l}^{(1+\alpha)/2})$$

by the same argument, which completes the proof in view of the proof of Theorem 2.

Proof of Proposition 1. Observe that $\lambda < \infty$ is tantamount to

$$\sup_{n\geq 1} \sum_{i=1}^{\infty} \Big[\sum_{j=i}^{n+i-1} a_j \Big]^2 < \infty. \tag{20}$$

This is implied by (17) since

$$\Big[\sum_{j=i}^{n+i-1} a_j \Big]^2 \leq 2\Big[\sum_{j=i}^{\infty} a_j \Big]^2 + 2\Big[\sum_{j=n+i}^{\infty} a_j \Big]^2,$$

and then

$$\lambda \leq 2\sum_{i=1}^{\infty} \Big[\sum_{j=i}^{\infty} a_j \Big]^2 + 2\sup_{n\geq 1}\sum_{i=1}^{\infty} \Big[\sum_{j=n+i}^{\infty} a_j \Big]^2 < \infty.$$

Conversely, recalling $s_n = \sum_{i=0}^{n} a_i$, $\sup\{|s_n| : n \in \mathbb{N}\} < \infty$ by (20). Hence there exists a subsequence $n' \subset \mathbb{N}$ and a real number s_* such that $s_{n'} \to s_*$ as $n' \to \infty$. So for each fixed $m \in \mathbb{N}$, $s_{n'+m} \to s_*$ since $\sum'(s_{n'+m} - s_{n'})^2 < \infty$ due to (20), where $\sum'$ sums along the subsequence n'. Therefore, for any fixed $I \in \mathbb{N}$, again by (20) we get

$$\lambda \geq \sup_{n'\geq 1} \sum_{i=1}^{I} [s_{n'+i-1} - s_{i-1}]^2 \geq \sum_{i=1}^{I} [s_* - s_{i-1}]^2,$$

which entails $\lim_{n\to\infty} s_n = s_*$ and then (17), since I is arbitrarily chosen.

Proof of Proposition 2. By Proposition 1, the tail $\tau_k := \sum_{t=k}^{\infty} a_t$ exists. For fixed $k \in \mathbb{N}$, $\sum_{t=-k}^{k} \Gamma(t) = \sum_{|i-j|\le k} a_i a_j$ is clearly absolutely summable. So we can rewrite this sum as $\tau_0 s_k - \eta_k + \xi_k$, where $\eta_k = \sum_{t=0}^{\infty} a_t \tau_{k+1+t}$ and $\xi_k = \sum_{t=1}^{\infty} a_{k+t}\tau_t$. By Cauchy's inequality and Proposition 1, as $k \to \infty$, we have $|\eta_k|^2 \le (\sum_{t=0}^{\infty} a_t^2)(\sum_{t=0}^{\infty} \tau_{k+1+t}^2) \to 0$ and $|\xi_k|^2 \le (\sum_{t=1}^{\infty} a_{k+t}^2)(\sum_{t=1}^{\infty} \tau_t^2) \to 0$, which proves the proposition as $s_k \to \tau_0$.

Proof of Proposition 3. This proof is quite similar to the Proof of Proposition 4. Observe that $h_n = \sum_{k=1}^{n} \cos(\theta_1 + k\omega_1)$ satisfies $M := \sup_{n\in\mathbb{N}} |h_n| < \infty$ since $\omega_1 \ne 0 \pmod{2\pi}$. Then for $m > n > n_0$, by the triangle inequality,

$$\Big|\sum_{k=n}^{m} a_k\Big| = \Big| - b_n h_{n-1} + b_m h_m + \sum_{k=n}^{m-1} (b_k - b_{k+1}) h_k\Big| \le 2Mb_n,$$

where the last step is due to the monotonicity of b_n when $n > n_0$. Therefore a_n is summable and Proposition 1 completes the proof since $\sum_{n=1}^{\infty} b_n^2 < \infty$.

Proof of Proposition 4. For simplicity, we prove it in the complex domain. Without loss of generality, we assume $a_0 = 1, a_n = n^{-\gamma} L(n) \exp[(\theta_1 + n\omega_1)\sqrt{-1}], n \in \mathbb{N}$, where $L(\cdot)$ satisfies (18) and $\omega_1 \ne 0 \pmod{2\pi}$. Hence $h(n) = \sum_{k=1}^{n} \exp[(\theta_1 + k\omega_1)\sqrt{-1}]$ is bounded. For $i, n \in \mathbb{N}$,

$$\begin{aligned}
\Big|\sum_{k=1}^{n} a_{k+i}\Big| &= \Big|\sum_{k=1}^{n} \frac{L(k+i)}{(k+i)^{\gamma}} \times [h(k+i) - h(k+i-1)]\Big| \\
&\le \mathcal{O}\Big[\frac{L(1+i)}{(1+i)^{\gamma}} + \frac{L(n+i)}{(n+i)^{\gamma}}\Big] + \sum_{k=1}^{n-1} \Big|\Big[\frac{L(k+i)}{(k+i)^{\gamma}} - \frac{L(k+i+1)}{(k+i+1)^{\gamma}}\Big] h(k+i)\Big| \\
&\le \mathcal{O}\Big[\frac{L(1+i)}{(1+i)^{\gamma}} + \frac{L(n+i)}{(n+i)^{\gamma}}\Big] + \mathcal{O}\Big[\sum_{k=1}^{n-1} \frac{|L(k+i) - L(k+i+1)|}{(k+i)^{\gamma}}\Big] \\
&\quad + \mathcal{O}\Big\{\sum_{k=1}^{n-1} \Big| L(k+i+1)\Big[\frac{1}{(k+i)^{\gamma}} - \frac{1}{(k+i+1)^{\gamma}}\Big]\Big|\Big\} \\
&\le \mathcal{O}\Big[\frac{L(1+i)}{(1+i)^{\gamma}} + \frac{L(n+i)}{(n+i)^{\gamma}}\Big] + \mathcal{O}\Big[\frac{L_1(i)}{i^{\gamma}}\Big] + \mathcal{O}\Big[\frac{L_2(i)}{i^{\gamma}}\Big],
\end{aligned}$$

where $L_2(\cdot)$ is another slowly varying function. The last inequality follows from (18) and elementary properties of slowly varying functions. So the sequence $\{a_n\}$ is summable. Then, letting $n \to \infty$, the proposition follows in view of Proposition 1 and the fact that $\gamma > 1/2$.

6. Discussion and Further Study

It is clear that the finite dimensional convergence holds for the empirical process discussed in Section 3. The tightness argument could perhaps be made possible by computations of the type performed in Doukhan and Surgailis (1998) in the space $D[0,1]$.

For the model with a spectral density having multiple singularities presented in Section 3.1, we have to assume $\gamma_i > 3/4$ as in Corollary 3 to ensure the invariance principle in view of the term $\mathcal{O}[M\sum_{k=1}^{n} A_k]$ in (12). We conjecture that the constraint $\gamma_i > 3/4$ cannot be removed in the following sense. If $1/2 < \gamma_i < 3/4$ for some i, then we could possibly have some non-central limit theorems where the limiting distribution is expressed in terms of multiple Wiener-Itô integrals (see, for example, Theorem 3.1 along with Corollary 3.3 in Ho and Hsing (1997)). See Rosenblatt (1981) for some relevant results for Gaussian processes.

Edgeworth expansions would naturally be the next topic to study, once central limit theorems have been established. To derive an Edgeworth expansion, Götze and Hipp (1994) discuss linear processes where the coefficients a_n vanish exponentially fast, which is a consequence of their ARMA model. We believe that extensions to general ARIMA models are worth pursuing.

Acknowledgement

This research supported by the U.S. Army Research Office. The author is grateful to Michael Woodroofe for his supervision, and to Professor Jan Mielniczuk for introducing him to the topic. Suggestions of Professor Sándor Csörgő and George Michailidis led to a significant improvement of the paper. The comments by the Editor, an associated editor and a referee are greatly appreciated.

References

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

Bingham, N. H., Goldie, C. M. and Teugels, J. L. (1987). *Regular variation.* Cambridge University Press, Cambridge.

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

Cuzick, J. (1976). A central limit theorem for the number of zeros of a stationary Gaussian process. *Ann. Probab.* **4**, 547-556.

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

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.

Giraitis, L. and Surgailis, D. (1986). Multivariate Appell polynomials and the Central Limit Theorem. In *Dependence in Probability and Statistics: A Survey of Recent Results*, 21-71. Birkhauser, Boston, MA.

Giraitis, L. and Surgailis, D. (1999). Central limit theorem for the empirical process of a linear sequence with long memory. *J. Statist. Plann. Inference* **80**, 81-93.

Giraitis, L., Koul, H. L. and Surgailis, D. (1996). Asymptotic normality of regression estimators with long memory errors. *Statist. Probab. Lett.* **29**, 317-335.

Gordin, M. I. and Lifsic, B. (1978). The central limit theorem for stationary Markov processes. *Doklady* **19**, 392-394.

Götze, F. and Hipp, C. (1994). Asymptotic distribution of statistics in time series. *Ann. Statist.* **22**, 2062-2088.

Gray, H. L., Zhang, N.-F. and Woodward, W. A. (1989). On generalized fractional processes. *J. Time Ser. Anal.* **10**, 233-257.

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

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.

Ho, H. C. and Hsing, T. (1997). Limit theorems for functionals of moving averages. *Ann. Probab.* **25**, 1636-1669.

Ho, H. C. and Sun, T. C. (1987). A central limit theorem for non-instantaneous filters of a stationary Gaussian process. *J. Multivariate Anal.* **22**, 144-155.

Hoeffding, W. and Robbins, H. (1948). The central limit theorem for dependent random variables. *Duke Math. J.* **15**, 773-780.

Hsing, T. (1999). On the asymptotic distributions of partial sums of functionals of infinite-variance moving averages. *Ann. Probab.* **27**, 1579-1599.

Kedem, B. (1994). *Time series analysis by higher order crossings.* IEEE Press, Piscataway, NJ.

Malevich, T. (1969). Asymptotic normality of crossings of level zero by a Gaussian process. *Theory Probab. Appl.* **14**, 287-295.

Maxwell, M. and Woodroofe, M. (2000). Central limit theorems for additive functionals of Markov chains. *Ann. Probab.* **28**, 713-724.

Niederjohn, R. J. and Castelaz, P. F. (1978). Zero-crossing analysis methods for speech recognition. In *Proceedings of the* 1978 *Conference on Pattern Recognition and Image Processing*, volume CH1318, 507-513. IEEE, New York.

Phillips, P. C. B. and Solo, V. (1992). Asymptotics for linear processes. *Ann. Statist.* **20**, 971-1001.

Robinson, P. M. (1997). Large-sample inference for nonparametric regression with dependent errors. *Ann. Statist.* **25**, 2054-2083.

Rosenblatt, M. (1981). Limit theorems for Fourier transforms of functionals of Gaussian sequences. *Z. Wahrsch. Verw. Gebiete* **55**, 123-132.

Slud, E. V. (1994). MWI representation of the number of curve-crossings by a differentiable Gaussian process, with applications. *Ann. Probab.* **22**, 1355-1380.

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

Department of Statistics, The University of Chicago, IL 60637, U.S.A.

E-mail: wbwu@galton.uchicago.edu

(Received March 2000; accepted May 2001)