ELSEVIER Stochastic Processes and their Applications 60 (1995) 343–353

stochastic processes and their applications

Strong approximation for cross-covariances of linear variables with long-range dependence

Michael A. Kouritzin *

Department of Mathematics and Statistics, Carleton University, Ottawa, Canada K1S 5B6

Received 28 August 1994; revised 2 May 1995

Abstract

Suppose $\{\varepsilon_k, -\infty < k < \infty\}$ is an independent, not necessarily identically distributed sequence of random variables, and $\{c_j\}_{j=0}^{\infty}$, $\{d_j\}_{j=0}^{\infty}$ are sequences of real numbers such that $\sum_j c_j^2 < \infty$, $\sum_j d_j^2 < \infty$. Then, under appropriate moment conditions on $\{\varepsilon_k, -\infty < k < \infty\}$, $y_k \triangleq \sum_{j=0}^{\infty} c_j \varepsilon_{k-j}$, $z_k \triangleq \sum_{j=0}^{\infty} d_j \varepsilon_{k-j}$ exist almost surely and in $\mathcal{L}^4$ and the question of Gaussian approximation to $S_{\lfloor t \rfloor} \triangleq \sum_{k=1}^{\lfloor t \rfloor}(y_k z_k - E\{y_k z_k\})$ becomes of interest. Prior to this work several related central limit theorems and a weak invariance principle were proven under stationary assumptions. In this note, we demonstrate that an almost sure invariance principle for $S_{\lfloor t \rfloor}$, with error bound sharp enough to imply a weak invariance principle, a functional law of the iterated logarithm, and even upper and lower class results, also exists. Moreover, we remove virtually all constraints on ε_k for "time" $k \leqslant 0$, weaken the stationarity assumptions on $\{\varepsilon_k, -\infty < k < \infty\}$, and improve the summability conditions on $\{c_j\}_{j=0}^{\infty}$, $\{d_j\}_{j=0}^{\infty}$ as compared to the existing weak invariance principle. Applications relevant to this work include normal approximation and almost sure fluctuation results in sample covariances (let $d_j = c_{j-m}$ for $j \geqslant m$ and otherwise 0), quadratic forms, Whittle's and Hosoya's estimates, adaptive filtering and stochastic approximation.

Keywords: Almost sure invariance principle; Linear processes; Non-stationary innovations; Covariance process; Law of the iterated logarithm

1. Introduction

Linear process models enjoy widespread use in such diverse fields as engineering, econometrics, and statistics; and the convergence properties of the sample covariance and related processes for such a model is of great interest. Yet, from a mathematical viewpoint linear models are often less desirable than mixing-type assumptions since difficult manipulations usually arise when striving to establish rate of convergence results for the sample covariance of a non-stationary, linear model. Still, the realization that not every linear model yields a strong mixing process and the non-encompassing

* Correspondence address: Institute for Mathematics and its Applications, University of Minnesota, 514 Vincent Hall, 206 Church St. SE, Minneapolis MN 55455-0436, USA.

SSDI 0304-4149(95)00069-0

nature of the known sufficient conditions for the strong mixing property (see Withers, 1981; Gorodetskii, 1977) have resulted in a number of convergence property investigations of sample covariance and related processes for linear models. To deal with the resulting imposition of difficult manipulations, some authors have introduced a martingale approximation to the sample covariance process and then bounded the error term. In this note, we strive to ease the process of bounding the error between a sample covariance or related process and its martingale approximation, when the innovation process of the linear model has finite fourth-order moments, by decomposing this error in a convenient manner. However, of at least equal importance as this decomposition is our main result which is an almost sure invariance principle with error $O(t^{1/2-\lambda})$, $\lambda > 0$ for the sample covariance and related processes. Our result will be established under mild conditions and will imply (see the introduction of Philipp and Stout, 1975) that the weak invariance principle as well as various laws of the iterated logarithm and upper and lower class results hold for these processes.

We let $y_k \triangleq \sum_{j=0}^{\infty} c_j \varepsilon_{k-j}$ and $z_k \triangleq \sum_{j=0}^{\infty} d_j \varepsilon_{k-j}$, $k = 1,2,3,...$ be one-sided linear processes subordinated to the same innovations process and note that for most of the works listed below the independence assumption on $\{\varepsilon_k,\ -\infty < k < \infty\}$ has been replaced with a variety of more general assumptions. Then, the works of Anderson and Walker (1964), and Hannan and Heyde (1972, Theorem 3) resulted in a central limit theorem for the so-called autocorrelation process of $\{y_k,\ k = 1,2,3,...\}$ under strict stationarity and the nearly optimal summability condition that $\sum_j j^{1/2} c_j^2 < \infty$. In fact, Anderson and Walker showed that normal approximation can continue to hold even when ε_k does not have finite fourth-order moments but, as illustrated in Davis and Marengo (1990), this is limited to one-dimensional autocorrelation processes. Later, Hannan (1976) added the autocovariance process to the class of strictly stationary, linear-model-based processes satsifying the central limit theorem under general conditions including finite fourth-order moments for ε_k. Next, in studying the sample covariance process, $(1/N)\sum_{k=1}^{N-s} y_k\, y_{k+s}$, Hosoya and Taniguchi (1982) reduced the strict stationarity assumption to fourth-order stationarity for their central limit theorem. Subsequently, Giraitis and Surgailis (1990, Theorem 3) proved a central limit theorem for the related process $(1/N)\sum_{k=1}^{N} y_k\, z_k$ and allowed both y_k and z_k to be two-sided linear processes. Finally, in a complimentary and very interesting set of results, Davis and Resnick (1985, 1986) have established weak convergence results for sample covariance processes of two-sided linear models to non-normal stable distributions (as well as several other weak convergence results) when ε_k does not have fourth order moments.

Although the previously mentioned results are unquestionably important, only weak convergence results are obtained and very strong stationarity assumptions are made. In this note, we envision our linear process models as generalized stable ARMA-type system models where the transients have not fully died out and strive to establish an almost sure invariance principle for $(1/N)\sum_{k=1}^{N} y_k\, z_k$ with a rate sufficient to imply all classical fluctuation results. It is in this light that we refer to the other central limit theorem in Hannan and Heyde (1972, Theorem 2) and the weak invariance principle of Phillips and Solo (1992, Theorem 3.8). These results have obviously suboptimal summability conditions on $\{c_j\}$ and stronger stationarity conditions than we wish to use. However, their method of proof relies on martingale approximations for which there exist general almost sure invariance principles even under non-stationary conditions. Indeed, there are almost sure invariance principles for martingales under much greater non-stationarity than we will make use of here (see Philipp and Stout, 1986; Jain et al., 1975). Still, our motivation stems from situations where mild forms of stationarity prevail asymptotically and the almost sure invariance principle of Eberlein and Philipp (see Philipp, 1986) will prove to be more manageable for our purposes.

The widespread acceptance and use of linear models combined with the lack of almost sure fluctuation results for cross-covariance processes of such models should be motivation enough for our present work. However, we briefly mention below a few situations where normal approximation is of particular interest. Hannan and Heyde (1972) stress the importance of normal approximation of the autocorrelation process for a linear model to the classical inferential theory and apply their central limit theorems to autoregression problems. Hosoya and Taniguchi (1982) apply their result to obtain an asymptotic theory for Hosoya's estimate for selecting a least diverged, as based upon the periodogram of a partial observation, spectral density from a fitted model and then apply this asymptotic theory to estimating parameters of an autoregressive signal when the observed process is corrupted by white noise. Giraitis and Surgailis (1990) utilize their previously mentioned result to obtain a central limit theorem for quadratic forms of strongly dependent linear variables and apply this theorem to prove asymptotic normality of the parameters in Whittle's estimate. Our result and method contribute to these applications respectively a stronger, almost sure form of Gaussian approximation from which all classical fluctuation results for $(1/N)\sum_{k=1}^{N} y_k z_k$ follow and a means (see Remark 4.1) to reduce the stationarity conditions imposed in these works. Finally, we refer to Kouritzin (1994) for the relevance of the present work to the areas of adaptive filtering and stochastic approximation.

2. Notation

Let $\{\varepsilon_k,\ -\infty < k < \infty\}$ be a sequence of zero mean, independent random variables on some probability space $(\Omega, \mathcal{F}, P)$. We will not require that they are identically distributed but rather only the much weaker condition that $\sup_k E|\varepsilon_k|^{4+\delta} < \infty$ for some $\delta > 0$. Obviously, this condition permits the sometimes separately considered situation where $\varepsilon_k \equiv 0$ for $k < 0$. At this point, it is convenient to define the pair of independent σ-algebras:

$$\mathcal{F}_m \triangleq \sigma\{\varepsilon_i,\ -\infty < i \leq m\}, \qquad \mathcal{F}_m^+ \triangleq \sigma\{\varepsilon_i,\ m < i < \infty\} \tag{2.1}$$

for each $m = 0, 1, 2, \ldots$. Then, under the above conditions, it is not difficult to see that the partial summations $\sum_{i=0}^{N} b_i \varepsilon_{k-i}$, $\sum_{j=0}^{N}\sum_{l=0}^{N} a_{l,j}(\varepsilon_{k-j-l}\varepsilon_{k-j} - E\{\varepsilon_{k-j-l}\varepsilon_{k-j}\})$, for any sequences $\{b_i\}$, $\{a_{l,j}\}$ such that $\sum_{i=0}^{\infty} b_i^2 < \infty$ and $\sum_{j,l=0}^{\infty} a_{l,j}^2 < \infty$, converge in $\mathcal{L}^4(\Omega,\ \mathcal{F}_k, P)$ respectively $\mathcal{L}^2(\Omega,\ \mathcal{F}_k, P)$ to limits which we shall denote by $\sum_{i=0}^{\infty} b_i \varepsilon_{k-i}$, $\sum_{j=0}^{\infty}\sum_{l=0}^{\infty} a_{l,j}(\varepsilon_{k-j-l}\varepsilon_{k-j} - E\{\varepsilon_{k-j-l}\varepsilon_{k-j}\})$. (Of course, the Kolmogorov zero-one law ensures that the first series also converges almost surely.)

Now we let $\{c_j\}_{j=0}^{\infty}$, $\{d_j\}_{j=0}^{\infty}$ be two sequences of real numbers satisfying $\sum_{j=0}^{\infty} c_j^2 < \infty, \sum_{j=0}^{\infty} d_j^2 < \infty$ and summarize our remaining notation:

$$y_k \triangleq \sum_{j=0}^{\infty} c_j \varepsilon_{k-j}, \qquad z_k \triangleq \sum_{j=0}^{\infty} d_j \varepsilon_{k-j} \quad \text{for all } k = 1,2,3,...,$$

$$\sigma_k^2 \triangleq E\varepsilon_k^2,\ \phi_k \triangleq E\varepsilon_k^3,\ \gamma_k \triangleq E\varepsilon_k^4 \quad \text{for all } k \in \mathbb{Z} \text{ (the set of integers)},$$

$$S_{n,m} \triangleq \sum_{k=m+1}^{m+n} (y_k z_k - E\{y_k z_k\}) \quad \text{for all } n,m = 0,1,2,...,$$

$$S_n = S_{n,0} \quad \text{for all } n = 1,2,3,...,$$

$$f_{0,j} \triangleq c_j d_j \quad \text{for all } j = 0,1,2,...,$$

$$f_{l,j} \triangleq c_j d_{j+l} + d_j c_{j+l} \quad \text{for all } l = 1,2,3,...,\ j = 0,1,2,...,$$

$$\tilde{f}_{l,i} \triangleq \sum_{j=i}^{\infty} f_{l,j} \quad \text{for all } l,i = 0,1,2,...,$$

$$\|X\|_p \triangleq (E|X|^p)^{1/p} \quad \text{for } p \geqslant 1,$$

and $a_{n,m} \overset{n,m}{\ll} b_{n,m}$ means that there is a constant c such that $|a_{n,m}| \leqslant c\,|b_{n,m}|$ for all n, m. This is a natural extension to the Vinogradov symbol $\ll$.

3. Results and discussion

Our aim is to produce an almost sure invariance principle for S_n under only mild pseudo-stationarity conditions (which do not require all transients to have died out) and summability conditions on $\{c_j\}$ and $\{d_j\}$ comparable to those required for the central limit theorems mentioned in the introduction to hold (under stronger stationarity conditions). Indeed, we are fortunate to have a powerful theorem due to Eberlein and Philipp from which we can conclude that our desired invariance principle holds provided, for $m \geqslant 0$, $n \geqslant 1$,

(I) $\|E\{S_{n,m}|\mathcal{F}_m\}\|_1 \overset{n,m}{\ll} n^{\frac{1}{2}-\varepsilon}$ for some $\varepsilon > 0$,

(II) There exists an $\alpha^2 \geqslant 0$ such that $\|E\{S_{n,m}^2|\mathcal{F}_m\} - n\alpha^2\|_1 \overset{n,m}{\ll} n^{1-\varepsilon}$ for some $\varepsilon > 0$,

(III) $\sup_{k\geqslant 0} E|y_k|^{4+\delta} < \infty$ and $\sup_{k \geqslant 0} E|z_k|^{4+\delta} < \infty$ for some $0 < \delta < 1$.

It follows from Theorem 1 of Philipp (1986) that under (I)–(III) there exists a process $\{\tilde{S}_n,\ n = 0,1,2,...\}$ and a Brownian motion $\{\tilde{X}(t),\ t \geqslant 0\}$ with incremental variance α^2 on some probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ such that: (a) $\mathcal{L}(\{\tilde{S}_n,\ n = 0,1,2,...\}) = \mathcal{L}(\{S_n,\ n = 0,1,2,...\})$ and (b) for some $\lambda > 0$ we have that $|\tilde{S}_{\lfloor t \rfloor} - \tilde{X}(t)| \overset{t}{\ll} t^{\frac{1}{2}-\lambda}$ for all $t > 0$ a.s. $[\tilde{P}]$.

Remark 3.1. Originally, Serfling (1970, Theorem 4.1) established a central limit theorem under conditions similar to (I)–(III). Later, McLeish (1975, Theorem 2.6) established a weak invariance principle under related conditions. Next, Eberlein (1986, Theorem 1) demonstrated how general these type of conditions are and established an almost sure invariance principle in the vector-valued case. Finally, Philipp (1986, Theorem 1) extended the almost sure invariance principle to Hilbert space valued random variables and concurrently weakened the conditions somewhat. We have stated an $\Re$-valued version of Philipp (1986) with the exception that (I) is more stringent than it has to be. However, we will not be able to use the extra generality afforded by Philipp's version of (I) so our real benefit of Philipp's generalization is being able to treat (II) as one rather than two parts.

We now state the main result of this note. Later, in Section 4, we will prove this result by establishing (I)–(III) above.

Theorem 1. *Suppose that* $\{\varepsilon_k, -\infty < k < \infty\}$, $\{\tilde{f}_{l,j}\}_{l,j=0}^{\infty}$ *and* $\{S_n\}_{n=0}^{\infty}$ *are as in Section 2 and, specifically, that*

$$\sup_k E|\varepsilon_k|^{4+\delta} < \infty \quad for \quad some \quad \delta > 0\ . \tag{3.1}$$

Moreover, suppose for some $0 < \theta < \frac{1}{2}$ *that*

$$\sum_{i=1}^{\infty}\sum_{l=0}^{\infty}(\tilde{f}_{l,i} - \tilde{f}_{l,i+n})^2 \overset{n}{\ll} n^{1-\theta} \quad for \quad all \quad n = 1,2,3,..., \tag{3.2}$$

$$\sum_{k=1}^{n}\sum_{r=0}^{\infty}\tilde{f}_{r+k,0}^2 \overset{n}{\ll} n^{1-\theta} \quad for \quad all \quad n = 1,2,3,..., \tag{3.3}$$

$$\sum_{k=1}^{n}\sum_{l=0}^{\infty}\tilde{f}_{l,k}^2 \overset{n}{\ll} n^{1-\theta} \quad for \quad all \quad n = 1,2,3,..., \tag{3.4}$$

and there exists some $\alpha^2 \geqslant 0$ *such that*

$$\tilde{f}_{0,0}^2\sum_{j=1}^{n}(\gamma_{m+j} - \sigma_{m+j}^4) + \sum_{j=1}^{n}\sum_{l=1}^{j-1}\tilde{f}_{j-l,0}^2\,\sigma_{m+j}^2\cdot\sigma_{m+l}^2 - n\alpha^2 \overset{n,m}{\ll} n^{1-\theta} \tag{3.5}$$

for all $n,m = 0,1,2,...$ *Then, without changing its distribution, we can redefine the sequence* $\{S_n,\ n = 0,1,2,...\}$ *on a richer probability space on which there exists a Brownian motion* $\{X(t),\ t \geqslant 0\}$ *with variance* $\alpha^2 t$ *such that for some* $\lambda > 0$

$$|S_{\lfloor t\rfloor} - X(t)| \overset{t}{\ll} t^{\frac{1}{2}-\lambda} \quad for \quad all \quad t > 0 \qquad \text{a.s.}$$

Remark 3.2. (i) Clearly, (3.5) does not impose any restriction on the statistics of $\{\varepsilon_k, -\infty < k \leqslant 0\}$. Hence the only restriction on $\{\varepsilon_k, -\infty < k < \infty\}$ at or before "time zero" is through (3.1). Moreover, if $\gamma_k \equiv \gamma$ and $\sigma_k^2 \equiv \sigma^2$ for all $k = 1,2,...$ (or they decay sufficiently fast to these constant values) then (3.5) follows by (3.3) with

$$\alpha^2 \triangleq (\gamma - \sigma^4)\tilde{f}_{0,0}^2 + \sigma^4\sum_{l=1}^{\infty}\tilde{f}_{l,0}^2\ . \tag{3.6}$$

(ii) Next, (3.2)–(3.4) might appear somewhat complicated. However, it follows from (3.6) that summability conditions on our coefficients stronger than $\sum_{l=0}^{\infty}\tilde{f}_{l,0}^2 < \infty$ are required. In fact, it is easily seen that both (3.3) and (3.4) imply $\sum_{l=0}^{\infty}\tilde{f}_{l,0}^2 < \infty$. On the other hand, it is an easy exercise to show constraints like $c_j^2 \overset{j}{\ll} j^{-\frac{3}{2}-\theta}$, $d_j^2 \overset{j}{\ll} j^{-\frac{3}{2}-\theta}$ for all $j \geqslant 1$, which are enough to ensure that $\sum_{l=0}^{\infty}\tilde{f}_{l,0}^2 < \infty$, also imply (3.2)–(3.4).

Of course, these constaints would not be sufficient to ensure that $\sum_{j=0}^{\infty}|c_j| < \infty$ or $\sum_{j=0}^{\infty} j\,c_j^2 < \infty$ as used in the central limit theorem of Hannan and Heyde (1972, Theorem 2) and the weak invariance principle of Phillips and Solo (1992, Theorem 3.8). Finally, Giraitis and Surgailis (1990, Theorem 3) establish a central limit theorem under (stationarity and) conditions somewhat stronger than $\sum_{l=0}^{\infty}\tilde{f}_{l,0}^2 < \infty$. It would be natural to expect that our almost sure invariance principle would require stronger summability conditions than their central limit theorem. However, it appears that their conditions are not completely comparable to (3.2)–(3.4).

4. Proof of Theorem 1

4.1. Martingale approximation and Condition (I)

The Conditions (I) and (II) of Section 3 are suggestive to making martingale approximations. Hence, we follow previous developments (see e.g. Phillips and Solo, 1992, pp. 972 and 979) somewhat and introduce the $\mathcal{L}^2$, $\{\mathcal{F}_m\}$-martingale:

$$M_m \triangleq \sum_{k=0}^{m}\left\{\tilde{f}_{0,0}\left(\varepsilon_k^2-\sigma_k^2\right)+\varepsilon_k\sum_{l=1}^{\infty}\tilde{f}_{l,0}\,\varepsilon_{k-l}\right\} \quad \text{for all} \quad m=0,1,2,\ldots \tag{4.1}$$

(Using (3.3), we have $\sum_{l=1}^{\infty}\tilde{f}_{l,0}^2 < \infty$ so $\sum_{l=1}^{\infty}\tilde{f}_{l,0}\,\varepsilon_{k-l} \in \mathcal{L}^4(\Omega,\mathcal{F}_{k-1},P)$ and the martingale property follows easily.)

As for the error

$$R_{n,m} \triangleq S_{n,m} - M_{n+m} + M_m, \quad \text{for all} \quad n,m=0,1,2,\ldots, \tag{4.2}$$

we make use of the following lemma:

Lemma 2. *Under the conditions of Theorem* 1,

$$R_{p,m} = Q_p + P_p + O_p \quad \textit{for all } m,p=0,1,2,\ldots \quad \text{a.s.}, \tag{4.3}$$

where

$$Q_p = Q_{p,m} \triangleq \sum_{i=0}^{p-1}\sum_{l=0}^{p-i-1}\tilde{f}_{l,i+1}(\sigma_{m-p-i}^2\,1_{\{l=0\}} - \varepsilon_{m+p-i}\,\varepsilon_{m+p-i-l}), \tag{4.4}$$

$$P_p = P_{p,m} \triangleq \sum_{i=0}^{\infty}\sum_{l=0}^{\infty}\left(\tilde{f}_{l,i+1} - \tilde{f}_{l,i+p+1}\right)(\varepsilon_{m-i}\,\varepsilon_{m-i-l} - \sigma_{m-i}^2\,1_{\{l=0\}}), \tag{4.5}$$

$$O_p = O_{p,m} \triangleq -\sum_{i=0}^{p-1}\sum_{k=p}^{\infty}\tilde{f}_{k-i,i+1}\,\varepsilon_{m+p-k}\,\varepsilon_{m+p-i}, \tag{4.6}$$

for all $m,p=0,1,2,\ldots$ *Moreover,*

$$\max\{EQ_p^2,\ EP_p^2,\ EO_p^2\} \overset{m,p}{\ll} p^{1-\theta} \quad \textit{for all} \quad m,p=0,1,2,\ldots, \tag{4.7}$$

where $0<\theta<\frac{1}{2}$ *is the constant of* (3.2)–(3.4) *in Section* 3.

Remark 4.1. (i) It will become obvious during the course of the proof of Lemma 2 that Q_p and P_p are the projections of $S_{p,m}-M_{m+p}+M_m$ onto the closed linear spans of respectively $\{\varepsilon_r\,\varepsilon_s-\sigma_r^2 1_{\{r=s\}},\ m<r,s\leqslant m+p\}$ and $\{\varepsilon_r\,\varepsilon_s-\sigma_r^2 1_{\{r=s\}},\ -\infty<r,s\leqslant m\}$. (ii) In light of (3.4), (3.2), and the discussion in Section 2, it is easy to see that O_p and P_p are well defined. (iii) This decomposition and, in fact, all the development in Subsections 4.1 and 4.2 require only uniform fourth-order moment conditions on $\{\varepsilon_k,\ -\infty<k<\infty\}$. Hence, one could use our method and the result of Serfling (1970, Theorem 4.1) to establish a central limit theorem under our pseudo-stationarity conditions but with weaker moment bounds.

Proof of Lemma 2. We will first prove (4.7). Fix integers $m,p\geqslant 0$. Then, by (3.1) and (3.4) it follows that

$$EQ_p^2=\sum_{i=0}^{p-1}\tilde f^2_{0,i+1}(\gamma_{m+p-i}-\sigma^4_{m+p-i})+\sum_{i=0}^{p-1}\sum_{l=1}^{p-i-1}\tilde f^2_{l,i+1}\,\sigma^2_{m+p-i}\sigma^2_{m+p-i-l}$$
$$\overset{m,p}{\ll}\sum_{i=1}^{p}\sum_{l=0}^{p-i}\tilde f^2_{l,i}\overset{m,p}{\ll}p^{1-\theta}. \tag{4.8}$$

Moreover, by continuity of inner product, (3.1), and (3.2), we find that

$$EP_p^2=\sum_{i=0}^{\infty}\left(\tilde f_{0,i+1}-\tilde f_{0,i+p+1}\right)^2(\gamma_{m-i}-\sigma^4_{m-i})$$
$$+\sum_{i=0}^{\infty}\sum_{l=1}^{\infty}\left(\tilde f_{l,i+1}-\tilde f_{l,i+p+1}\right)^2\sigma^2_{m-i}\sigma^2_{m-i-l}$$
$$\overset{m,p}{\ll}\sum_{i=1}^{\infty}\sum_{l=0}^{\infty}\left(\tilde f_{l,i}-\tilde f_{l,i+p}\right)^2\overset{m,p}{\ll}p^{1-\theta}. \tag{4.9}$$

Next, we use Fatou's lemma along with (3.1) and (3.4) to conclude that

$$EO_p^2\leqslant\liminf_{N\to\infty}\sum_{i=0}^{p-1}\sum_{k=p}^{N}\tilde f^2_{k-i,i+1}\,\sigma^2_{m+p-i}\,\sigma^2_{m+p-k}$$
$$\overset{m,p}{\ll}\sum_{i=0}^{p-1}\sum_{k=p}^{\infty}\tilde f^2_{k-i,i+1}\leqslant\sum_{i=1}^{p}\sum_{l=1}^{\infty}\tilde f^2_{l,i}\overset{m,p}{\ll}p^{1-\theta}. \tag{4.10}$$

Now, to show (4.3), we again fix integers $p,m\geqslant 0$ and define $\mathcal{L}=\overline{\operatorname{span}}\{\varepsilon_r\,\varepsilon_s-\sigma_r^2\,1_{\{r=s\}}:\ -\infty<r,s\leqslant m+p\}$ (the $\mathcal{L}^2(\Omega,\mathcal{F},P)$-closed linear span of $\{\varepsilon_r\,\varepsilon_s-\sigma_r^2 1_{\{r=s\}}: -\infty<r,s\leqslant m+p\}$). One can easily show, using only elementary methods, that $S_{p,m}$, M_{m+p}, M_m, Q_p, P_p, $O_p\in\mathcal{L}$. Hence, the lemma will follow by linearity and continuity of inner product as well as the zero-mean nature of $S_{p,m}-M_{m+p}+M_m-Q_p-P_p-O_p$ if we show that

$$E\{(S_{p,m}-M_{m+p}+M_m-Q_p-P_p)\varepsilon_r\,\varepsilon_s\}=E\{O_p\,\varepsilon_r\,\varepsilon_s\} \tag{4.11}$$

for all $-\infty < r \leqslant s \leqslant m+p$. However, the right-hand side of (4.11) is clearly 0 except when $-\infty < r \leqslant m < s \leqslant m+p$. In this case it is $-\sigma_r^2 \sigma_s^2 \tilde{f}_{s-r,m+p-s+1}$. Noting that

$$S_{p,m} - M_{m+p} + M_m$$
$$= \mathcal{L}^2 - \lim_{N\to\infty} \sum_{k=m+1}^{m+p} \left\{ \sum_{j=0}^{N} f_{0,j}(\varepsilon_{k-j}^2 - \sigma_{k-j}^2 - \varepsilon_k^2 + \sigma_k^2) \right.$$
$$\left. + \sum_{j=0}^{N} \sum_{l=1}^{N-j} f_{l,j}\varepsilon_{k-j-l}\varepsilon_{k-j} - \sum_{l=1}^{N} \tilde{f}_{l,0}\varepsilon_{k-l}\varepsilon_k \right\} \quad \text{a.s.} \tag{4.12}$$

by simple manipulation, one can easily validate (4.11). □

Of course, Condition (I) is immediately verified through Lemma 2 and (4.2). However, this lemma will also be found useful while investigating Condition (II).

4.2. Condition (II): conditional variance

Now in preparation of verifying Condition (II), we fix integers $m, n \geqslant 0$ and note by the martingale property that

$$\|E\left\{(M_{m+n} - M_m)^2 | \mathcal{F}_m\right\} - n\alpha^2\|_1 = \| \sum_{j=1}^{n} \left[E\{(M_{m+j} - M_{m+j-1})^2 | \mathcal{F}_m\} - \alpha^2\right] \|_1. \tag{4.13}$$

Next, defining $L_0 = K_0 = 0$ and

$$L_j \triangleq \sum_{k=m+1}^{m+j} \left\{ (\varepsilon_k^2 - \sigma_k^2)\tilde{f}_{0,0} + \sum_{l=1}^{k-m-1} \tilde{f}_{l,0}\, \varepsilon_k\, \varepsilon_{k-l} \right\}, \quad K_j \triangleq M_{m+j} - M_m - L_j, \tag{4.14}$$

for $j = 1, 2, ..., n$, we find that

$$\| \sum_{j=1}^{n} \left[E\{(M_{m+j} - M_{m+j-1})^2 | \mathcal{F}_m\} - \alpha^2\right] \|_1$$
$$\leqslant | \sum_{j=1}^{n} \left[E\{(L_j - L_{j-1})^2\} - \alpha^2\right] | + 2\| \sum_{j=1}^{n} E\{(L_j - L_{j-1})(K_j - K_{j-1}) | \mathcal{F}_m\}\|_1$$
$$+ \| \sum_{j=1}^{n} E\{(K_j - K_{j-1})^2 | \mathcal{F}_m\}\|_1 \; . \tag{4.15}$$

However, by (4.14) and (3.5) we have that

$$| \sum_{j=1}^{n} E(L_j - L_{j-1})^2 - n\alpha^2 | \overset{n,m}{\ll} n^{1-\theta} \; . \tag{4.16}$$

Moreover, clearly $K_j - K_{j-1} = \varepsilon_{m+j} \sum_{l=j}^{\infty} \tilde{f}_{l,0}\, \varepsilon_{m+j-l}$ a.s. and $\sum_{l=j}^{\infty} \tilde{f}_{l,0}\, \varepsilon_{m+j-l} \in \mathcal{L}^4 (\Omega, \mathcal{F}_m, P)$ so

$$E\{(K_j - K_{j-1})^2 | \mathcal{F}_m\} = \sigma_{m+j}^2 \left(\sum_{l=j}^{\infty} \tilde{f}_{l,0}\, \varepsilon_{m+j-l} \right)^2 \qquad \text{a.s.} \tag{4.17}$$

and by (4.14)

$$E\{(L_j - L_{j-1})(K_j - K_{j-1})|\mathcal{F}_m\} = \tilde{f}_{0,0}\,\phi_{m+j}\sum_{l=j}^{\infty}\tilde{f}_{l,0}\,\varepsilon_{m+j-l} \qquad \text{a.s.} \tag{4.18}$$

Hence, by (4.15)–(4.18), Cauchy–Schwarz and (3.1) it follows that

$$\left\|\sum_{j=1}^{n}\left[E\{(M_{m+j}-M_{m+j-1})^2|\mathcal{F}_m\}-\alpha^2\right]\right\|_1 \overset{m,n}{\ll} n^{1-\theta} + \sum_{j=1}^{n}\sum_{l=j}^{\infty}\tilde{f}_{l,0}^2 + \sum_{j=1}^{n}\sqrt{\sum_{l=j}^{\infty}\tilde{f}_{l,0}^2}\,. \tag{4.19}$$

Finally, because

$$\left(\frac{1}{n}\sum_{j=1}^{n}\sqrt{\sum_{l=j}^{\infty}\tilde{f}_{l,0}^2}\right)^2 \leqslant \frac{1}{n}\sum_{j=1}^{n}\sum_{l=j}^{\infty}\tilde{f}_{l,0}^2 \tag{4.20}$$

by Jensen's inequality, we have by (4.13), (4.19) and (3.3) that

$$\left\|E\{(M_{m+n}-M_m)^2|\,\mathcal{F}_m\}-n\alpha^2\right\|_1 \overset{n,m}{\ll} n^{1-\frac{\theta}{2}}. \tag{4.21}$$

Now, it follows from Lemma 2 that

$$\|E\{R_{n,m}^2|\,\mathcal{F}_m\}\|_1 \overset{n,m}{\ll} n^{1-\theta} \tag{4.22}$$

and from conditional Cauchy–Schwarz, (4.21), and (4.22) that

$$\begin{aligned}\|E\{S_{n,m}^2|\mathcal{F}_m\}-n\alpha^2\|_1 &\leqslant \|E\{R_{n,m}^2|\,\mathcal{F}_m\}\|_1 \\ &\quad + 2\left\|\sqrt{E\{R_{n,m}^2|\,\mathcal{F}_m\}}\sqrt{E\{(M_{m+n}-M_m)^2|\,\mathcal{F}_m\}}\right\|_1 \\ &\quad + \|E\{(M_{m+n}-M_m)^2|\,\mathcal{F}_m\}-n\alpha^2\|_1 \\ &\overset{n,m}{\ll} n^{1-\theta} + n^{\frac{1}{2}-\frac{\theta}{2}}n^{\frac{1}{2}} + n^{1-\frac{\theta}{2}} \overset{n,m}{\ll} n^{1-\frac{\theta}{2}}.\end{aligned} \tag{4.23}$$

4.3. Condition (III): moment condition

We assume without loss of generality that $\delta < 2$, let $\chi = \delta/2$ and note (see Longnecker and Serfling, 1978, Lemma 3.1) that

$$\left\{\sum_{j=0}^{\infty}|c_j|^{2(2+\chi)}\right\}^{\frac{1}{2+\chi}} \leqslant \sum_{j=0}^{\infty}c_j^2 < \infty, \tag{4.24}$$

so it follows by (4.24) and (3.1) that

$$\sup_k \max\{E\left|\sum_{j=0}^{\infty}c_j^2\,\varepsilon_{k-j}^2\right|^2, \sum_{j=0}^{\infty}E(c_j^2\,\varepsilon_{k-j}^2)^{2+\chi}, \sum_{j=0}^{\infty}c_j^2\,E(\varepsilon_{k-j}^2)\} < \infty. \tag{4.25}$$

Therefore, following the proof of Marcinkiewicz–Zygmund in Chow and Teicher (1988, pp.368–369) and applying (4.25) and Jensen's inequality, we find

$$
\begin{aligned}
E\left|\sum_{j=0}^{\infty} c_j \varepsilon_{k-j}\right|^{4+\delta} &\overset{k}{\ll} E\left|\sum_{j=0}^{\infty} c_j^2 \varepsilon_{k-j}^2\right|^{2+\chi} \\
&\overset{k}{\ll} \sum_{j=0}^{\infty}\left[E\left\{c_j^2 \varepsilon_{k-j}^2 \left|c_j^2 \varepsilon_{k-j}^2\right|^{1+\chi}\right\} + E\left\{c_j^2 \varepsilon_{k-j}^2 \left|\sum_{n\neq j} c_n^2 \varepsilon_{k-n}^2\right|^{1+\chi}\right\}\right] \\
&\overset{k}{\ll} \sum_{j=0}^{\infty} E\left|c_j^2 \varepsilon_{k-j}^2\right|^{2+\chi} + \sum_{j=0}^{\infty} c_j^2 E\varepsilon_{k-j}^2 \cdot \left\{E\left|\sum_{n=0}^{\infty} c_n^2 \varepsilon_{k-n}^2\right|^2\right\}^{\frac{1+\chi}{2}} \\
&\overset{k}{\ll} 1\,.
\end{aligned}
\qquad (4.26)
$$

References

T.W. Anderson and A.M. Walker, On the asymptotic distribution of the autocorrelations of a sample from a linear process, Ann. Math. Statist. 35 (1964) 1296–1303.

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

R. A. Davis and J.E. Marengo, Limit theory for sample covariance and correlation matrix functions of a class of multivariate linear processes, Comm. Statist. Stochastic Models 6 (1990) 483–497.

R. A. Davis and S. I. Resnick, Limit theory for moving averages of random variables with regularly varying tail probabilities, Ann. Probab. 13 (1985) 179–195.

R. A. Davis and S. I. Resnick, Limit theory for sample covariance and correlation functions of moving averages, Ann. Statist. 14 (1986) 533–558.

E. Eberlein, On strong invariance principles under dependence assumptions, Ann. Probab. 14(1) (1986) 260–270.

L. Giraitis and D. Surgailis, A central limit theorem for quadratic forms in strongly dependent linear variables and its application to asymptotical normality of Whittle's estimate, Probab. Theory Related Fields 86 (1990) 87–104.

V.V. Gorodetskii, On the strong mixing property of linear stochastic processes, Theory. Probab. Appl. 22 (1977) 411–413.

E.J. Hannan, The asymptotic distribution of serial covariances, Ann. Statist. 4 (1976) 396–399.

E.J. Hannan and C.C. Heyde, On limit theorems for quadratic functions of discrete time series, Ann. Math. Statist. 43 (1972) 2058–2066.

Y. Hosoya and M. Taniguchi, A central limit theorem for stationary processes and the parameter estimation of linear processes, Ann. Statist. 10 (1982) 132–153.

N.C. Jain, K. Jogdeo and W.F. Stout, Upper and lower functions for martingales and mixing processes, Ann. Probab. 3 (1975) 119–145.

M. A. Kouritzin, On the interrelation of almost sure invariance principles for certain stochastic adaptive algorithms and for partial sums of random variables, Tech. Rept. No. 249, Dept. of Mathematics and Statistics, Carleton Univ. (Ottawa, Canada, 1994).

M. Longnecker and R.J Serfling, Moment inequalities for S_n under general dependence restrictions, with applications, Z. Wahrsch. Verw. Gebiete 43 (1978) 1–21.

D. L. McLeish, Invariance principles for dependent variables, Z. Wahrsch. Verw. Gebiete 32 (1975) 165–178.

W. Philipp, A note on the almost sure approximation of weakly dependent random variables, Monatsch. Math. 102 (1986) 227–236.

W. Philipp and W.F. Stout, Almost sure invariance principles for partial sums of weakly dependent random variables, Mem. Amer. Math. Soc. 161 (1975) 1–140.

W. Philipp and W.F. Stout, Invariance principles for martingales and sums of independent random variables, Math. Z. 192 (1986) 253–264.

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

R.J. Serfling, Convergence properties of S_n under moment restrictions, Ann. Math. Statist. 41(4) (1970) 1235–1248.

C.S. Withers, Conditions for linear processes to be strong mixing, Z. Wahrsch. Verw. Gebiete 57 (1981) 477–480.