Adv. Appl. Prob. **42,** 137–157 (2010)
Printed in Northern Ireland

COVARIANCES ESTIMATION FOR LONG-MEMORY PROCESSES

WEI BIAO WU * ** AND
YINXIAO HUANG,* *University of Chicago*
WEI ZHENG,*** *University of Illinois at Chicago*

Abstract

For a time series, a plot of sample covariances is a popular way to assess its dependence properties. In this paper we give a systematic characterization of the asymptotic behavior of sample covariances of long-memory linear processes. Central and noncentral limit theorems are obtained for sample covariances with bounded as well as unbounded lags. It is shown that the limiting distribution depends in a very interesting way on the strength of dependence, the heavy-tailedness of the innovations, and the magnitude of the lags.

Keywords: Asymptotic normality; covariance, dichotomy; linear process; long-range dependence; Rosenblatt distribution

2000 Mathematics Subject Classification: Primary 60F05; 62M10
Secondary 60G10

1. Introduction

Auto-covariance functions play a fundamental role in time series analysis and they are used in various inference problems, including parameter estimation and hypothesis testing. They are naturally estimated by sample covariances. Hence, the convergence problem of sample covariances is of critical importance. There is a substantial literature on properties of sample covariance estimates; see, for example, Bartlett (1946), Hannan (1970, pp. 220–229, 326–333), (1976), Anderson (1971, pp. 438–500), Hall and Heyde (1980, pp. 182–194), Porat (1987), Brockwell and Davis (1991, pp. 220–237), Phillips and Solo (1992), Berlinet and Francq (1999), Wu and Min (2005), among others. However, many of the earlier results are for sample covariance estimates with bounded lags. The latter restriction is quite severe. To better understand the dependence structure of a time series, we would like to know the behavior of sample covariances at large lags, namely at lags which increase to infinity with respect to sample sizes. This is especially so in the study of long-memory or long-range dependent processes since for such processes we are particularly interested in covariances at large lags.

The asymptotic problem of sample covariances at large lags is quite challenging. As mentioned in Harris *et al*. (2003), the primary reason for the difficulty is that the standard asymptotic results, such as the functional central limit theorem, stochastic integral convergence, and long-run variance estimation, are not directly applicable since the lag k_n depends on the sample size n in such a way that $k_n \to \infty$. Recently, researchers have made several important breakthroughs and derived central limit theorems for sample covariances at lags

Received 5 February 2009; revision received 18 November 2009.
* Postal address: Department of Statistics, University of Chicago, Chicago, IL 60637, USA.
** Email address: wbwu@galton.uchicago.edu
*** Postal address: Department of Mathematics, Statistics, and Computer Science, University of Illinois at Chicago, 851 S. Morgan Street, Chicago, IL 60607-7045, USA.

k_n with $k_n \to \infty$. Keenan (1997) obtained a central limit theorem for sample covariances at lags k_n with $k_n \to \infty$ under the severe restriction $k_n = o(\log n)$. Harris *et al.* (2003) substantially extended the range of k_n for short-memory linear processes. Wu (2008) obtained a central limit theorem for sample covariances of nonlinear time series with a very wide range of k_n. However, all those results concern short-memory processes in which the covariances are absolutely summable. The techniques therein are not directly applicable to long-memory processes.

For long-memory processes, Hosking (1996) obtained central and noncentral limit theorems for sample covariances with bounded lags. Here the terminology noncentral limit theorem refers to the result that the limiting distribution is not normal, instead, it is the Rosenblatt process (see Rosenblatt (1979)). In Hosking's result, the restriction that the lag k is bounded is quite severe, since in the study of long-memory processes, we often want to study the behavior of sample covariances at large lags. Chung (2002) generalized Hosking's result to multivariate long-memory processes. Again, in Chung's setting the lags are bounded. A result for sample covariances of long-memory processes with unbounded lags is given in Dai (2004), who derived the uniform convergence of sample covariances. However, the latter paper does not provide an asymptotic distributional theory for sample covariances. For an inferential theory, we need to have a distributional theory.

In this paper we shall consider the asymptotic behavior of sample covariances of long-memory linear processes with bounded as well as unbounded lags. Consider the linear process

$$X_k = \mu + \sum_{i=0}^{\infty} a_i \varepsilon_{k-i},$$

where the ε_i, $i \in \mathbb{Z}$, are independent and identically distributed (i.i.d.) innovations with mean 0 and finite variance, μ is the mean, and the a_i are real coefficients of the form

$$a_i = i^{-\beta} \ell(i), \qquad i \in \mathbb{N},$$

where $\frac{1}{2} < \beta < 1$ and ℓ is a slowly varying function (see Bingham *et al.* (1989, pp. 26–28)). By the Karamata theorem in the latter book, we can show that the covariance function $\gamma_k = \operatorname{cov}(X_0, X_k) = \mathrm{E}(\varepsilon_0^2) \sum_{i=0}^{\infty} a_i a_{i+k}$ satisfies

$$\gamma_k \sim C_\beta \, \mathrm{E}(\varepsilon_0^2) \frac{\ell^2(k)}{k^{2\beta-1}}, \quad \text{where} \quad C_\beta = \int_0^\infty (u + u^2)^{-\beta} \, \mathrm{d}u, \tag{1}$$

as $k \to \infty$. Here, for two real sequences (b_k) and (c_k), we write $b_k \sim c_k$ if $\lim_{k\to\infty} b_k / c_k = 1$. Since $\frac{1}{2} < \beta < 1$, the γ_k are not summable, thus meaning long-range dependence or long memory. Given the sample $(X_i)_{i=1}^n$, if μ is known then we can naturally estimate γ_k by

$$\check{\gamma}_k = \frac{1}{n} \sum_{i=k+1}^{n} (X_i - \mu)(X_{i-k} - \mu), \qquad 0 \le k < n,$$

and let $\check{\gamma}_{-k} = \check{\gamma}_k$. If μ is unknown, we can estimate γ_k by the sample covariance

$$\hat{\gamma}_k = \frac{1}{n} \sum_{i=k+1}^{n} (X_i - \bar{X}_n)(X_{i-k} - \bar{X}_n), \quad 0 \le k < n, \quad \text{where} \quad \bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i. \tag{2}$$

Estimation of γ_k allows us to assess the strength of dependence of the process by examining the auto-covariance function plot. Based on (1), we can estimate the long-memory parameter β by performing a linear regression for the model $\log \hat{\gamma}_k \sim \alpha_0 + \alpha_1 \log k$ over $k = l_n, l_n + 1, \ldots, u_n$, where α_0 is the intercept, $\alpha_1 = 1 - 2\beta$, $l_n \to \infty$, and $u_n/n \to 0$. Let $(\hat{\alpha}_0, \hat{\alpha}_1)$ be the least squares estimate. Then β can be estimated by $\hat{\beta} = \frac{1}{2} - \hat{\alpha}_1/2$, and its confidence interval can be constructed if an asymptotic distributional theory of $(\hat{\gamma}_{l_n}, \ldots, \hat{\gamma}_{u_n})$ is available. Long-memory processes have been studied for several decades. However, the asymptotic distributional problem for $\hat{\gamma}_{k_n}$ with large k_n has been rarely touched.

Here we shall present a systematic asymptotic theory for $\check{\gamma}_k$ and $\hat{\gamma}_k$. It is shown that their asymptotic behavior depends in a very interesting way on the strength of dependence, the heavy-tailedness of the innovations, and the magnitude of the lags. The rest of the paper is organized as follows. Our main results are stated in Section 2. Some of the proofs are given in Section 3. In our proofs we have extensively applied the martingale approximation techniques, which in many situations lead to optimal and nearly optimal results.

2. Main results

Before presenting our main results, we shall first introduce some notation. For a random variable Z, write $Z \in \mathcal{L}^p$, $p > 0$, if $\|Z\|_p := (\mathrm{E}|Z|^p)^{1/p} < \infty$ and, for $p = 2$, write $\|Z\| = \|Z\|_2$. Denote by '$\Rightarrow$' the weak convergence and by '$^\top$' the matrix transpose. Let $\mathcal{F}_i = (\ldots, \varepsilon_{i-1}, \varepsilon_i)$, $i \in \mathbb{Z}$, and define the projection operator

$$\mathcal{P}_i \cdot = \mathrm{E}(\cdot \mid \mathcal{F}_i) - \mathrm{E}(\cdot \mid \mathcal{F}_{i-1}). \tag{3}$$

In Theorems 1–6, below, we assume that $\mu = 0$ and deal with $\sum_{i=1}^n X_i X_{i-k}$. As mentioned in Remark 1, below, they also hold for $\sum_{i=1+k}^n X_i X_{i-k} = n\check{\gamma}_k$.

Theorem 1. *Let k be a fixed nonnegative integer, and let $\mathrm{E}(X_i) = 0$; let*

$$Y_i = (X_i, X_{i-1}, \ldots, X_{i-k})^\top \quad \textit{and} \quad \Gamma_k = (\gamma_0, \gamma_1, \ldots, \gamma_k)^\top.$$

Assume that $\varepsilon_i \in \mathcal{L}^4$ and that

$$\sum_{i=1}^{\infty} i^{1/2-2\beta} \ell^4(i) < \infty. \tag{4}$$

Then

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n (X_i Y_i - \Gamma_k) \Rightarrow N[0, \mathrm{E}(D_0 D_0^\top)], \tag{5}$$

where $D_0 = \sum_{i=0}^{\infty} \mathcal{P}_0(X_i Y_i) \in \mathcal{L}^2$ and $\mathcal{P}_0$ is the projection operator (3).

Theorem 1 provides a central limit theorem for sample covariances when the dependence is relatively weak in the sense that (4) holds. Note that, by properties of slowly varying functions, (4) is satisfied if $\frac{3}{4} < \beta < 1$. In the boundary case, $\beta = \frac{3}{4}$, condition (4) becomes $\sum_{i=1}^{\infty} \ell^4(i)/i < \infty$, which is a sharp condition for a $\sqrt{n}$-central limit theorem. Indeed, as indicated by Theorem 3, below, if $\sum_{i=1}^{\infty} \ell^4(i)/i = \infty$, then we no longer have a $\sqrt{n}$-central limit theorem, though the asymptotic normality still holds. Similar results have been obtained in Hosking (1996), Hall and Hyde (1980, pp. 148–153), Wu and Min (2005), among others. However, the results therein are not as sharp and general as Theorem 1. For example, Hosking (1996) required that $\lim_{i\to\infty} \ell(i)$ exists, and Proposition 1 of Wu and Min (2005) required that $\sum_{i=1}^{\infty} i^{1/2} a_i^2 < \infty$, or $\sum_{i=1}^{\infty} \ell^2(i)/i < \infty$, which is stronger than (4) at the boundary case, $\beta = \frac{3}{4}$.

Theorem 1 requires k to be bounded. It turns out that, interestingly, under the same condition (4), we can also have asymptotic normality under the natural and mild condition on k_n: $k_n \to \infty$ and $k_n/n \to 0$. More interestingly, in Theorem 2, below, the limiting distribution $N(0, \Sigma_h)$ in (6) does not depend on the speed of k_n growing to infinity. This interesting property has been discovered in Theorem 2 of Wu (2009) which concerns short-range-dependent processes.

Theorem 2. *Let* $W_i = (X_i, X_{i-1}, \dots, X_{i-h+1})^\top$, *where* $h \in \mathbb{N}$ *is fixed. Let* $k_n \to \infty$, $k_n/n \to 0$, $\mathrm{E}(\varepsilon_i) = 0$, *and* $\varepsilon_i \in \mathcal{L}^4$, *and assume that (4) holds. Then we have*

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} [X_i W_{i-k_n} - \mathrm{E}(X_{k_n} W_0)] \Rightarrow N(0, \Sigma_h), \tag{6}$$

where Σ_h *is an* $h \times h$ *matrix with entries*

$$\sigma_{ab} = \sum_{j \in \mathbb{Z}} \gamma_{j+a} \gamma_{j+b} = \sum_{j \in \mathbb{Z}} \gamma_j \gamma_{j+b-a} =: \sigma_{a-b}, \qquad 1 \le a, b \le h.$$

A key step in proving Theorems 1 and 2 is that we approximate $\sum_{i=1}^{n} (X_i X_{i-k} - \gamma_k)$ by the martingale

$$M_{n,k} = \sum_{l=1}^{n} D_{l,k}, \quad \text{where} \quad D_{l,k} = \varepsilon_l \sum_{j=-\infty}^{-1} (\gamma_{k+j} + \gamma_{k-j}) \varepsilon_{l+j} + \gamma_k (\varepsilon_l^2 - \mathrm{E}\, \varepsilon_l^2).$$

See (17) and Lemma 1, below, for more details. Note that $D_{1,k}, D_{2,k}, \dots,$ are martingale differences. The above martingale approximation provides an interesting insight into the Bartlett formula for asymptotic distributions of sample covariance functions (see, for example, Proposition 7.3.1 of Brockwell and Davis (1991)) by noting that

$$\mathrm{E}(D_{l,k} D_{l,k'}) = \sum_{j=-\infty}^{-1} (\gamma_{k+j} + \gamma_{k-j})(\gamma_{k'+j} + \gamma_{k'-j}) \|\varepsilon_0\|^4 + \gamma_k \gamma_{k'} \kappa_4,$$

where $\kappa_4 = \|\varepsilon_0^2 - \mathrm{E}\, \varepsilon_0^2\|^2$. In other words, $D_{l,k}$ provides a probabilistic representation for the Bartlett formula.

Theorem 3, below, concerns the boundary case, $\beta = \frac{3}{4}$, while (4) is violated. Together with Theorem 1, they give a complete characterization of the asymptotic behavior of $\hat{\gamma}_k$ with bounded k at the boundary $\beta = \frac{3}{4}$. A special case of Theorem 3 gives Hosking's (1996) Theorem 4(ii), where in his setting the ε_i are i.i.d. Gaussian and $a_i \sim c i^{-3/4}$ with some positive constant c. In the latter case $\lim_{i\to\infty} \ell(i) = c$ and $\bar{\ell}(n) = \sum_{i=1}^{n} \ell^4(i)/i \sim c^4 \log n$. In Theorem 3, we recall (1) for C_β and Theorem 1 for Y_i and Γ_k, $k \ge 0$. Then $C_{3/4} = 5.244\,115\dots$. For $h \in \mathbb{N}$, let $\mathbb{I}_h = (1, \dots, 1)^\top$ be the column vector of h 1s.

Theorem 3. *Assume that* $\mathrm{E}(\varepsilon_i) = 0$, $\varepsilon_i \in \mathcal{L}^4$, $\beta = \frac{3}{4}$, *and* $\bar{\ell}(n) = \sum_{i=1}^{n} \ell^4(i)/i \to \infty$. *Let* $\mathcal{G}$ *be a standard normal random variable. Then, for fixed* $k \ge 0$, *we have*

$$\frac{1}{\sqrt{n \bar{\ell}(n)}} \sum_{i=1}^{n} (X_i Y_i - \Gamma_k) \Rightarrow 2 C_{3/4} \|\varepsilon_0\|^2 \mathcal{G} \mathbb{I}_{k+1}. \tag{7}$$

In Theorem 3 it is assumed that k is bounded. It is unclear what is the asymptotic distribution of $\sum_{i=1}^{n}(X_i X_{i-k} - \gamma_k)$ if $k = k_n \to \infty$ with $k_n = o(n)$. We conjecture that it is still asymptotically normal and pose it as an open problem.

If the dependence is strong enough such that $\beta < \frac{3}{4}$ then we can have a noncentral limit theorem in that the limiting distribution is the Rosenblatt distribution which is non-Gaussian. Noncentral limit theorems have a long history; see Rosenblatt (1979), Taqqu (1979), Avram and Taqqu (1987), Ho and Hsing (1997), among others. To define the Rosenblatt distribution, let $\mathbb{B}(s)$, $s \in \mathbb{R}$, be a standard Brownian motion. For $a \in \mathbb{R}$, let $a^+ = \max(a, 0)$ be the nonnegative part of a. For $r \in \mathbb{N}$ and $\beta < \frac{1}{2} + 1/(2r)$, define the multiple Wiener–Itô (MWI) integral

$$R_{r,\beta} = c_{r,\beta} \int_{\mathcal{S}_r} \left\{ \int_0^1 \left[\prod_{i=1}^{r} (v - u_i)^+ \right]^{-\beta} \mathrm{d}v \right\} \mathrm{d}\mathbb{B}(u_1) \cdots \mathrm{d}\mathbb{B}(u_r),$$

where $\mathcal{S}_r = \{(u_1, \ldots, u_r) : -\infty < u_1 < \cdots < u_r < 1\}$ is a simplex and $c_{r,\beta}$ is a norming constant such that $\|R_{r,\beta}\| = 1$. For $r = 2$ and $\frac{1}{2} < \beta < \frac{3}{4}$, we call $R_{r,\beta}$ the Rosenblatt distribution. Note that $R_{1,\beta}$ is Gaussian and, for all $r > 1$, $R_{r,\beta}$ is non-Gaussian (see Taqqu (1979)). For a review of the MWI integral, see Giraitis and Taqqu (1999) and Major (1981, pp. 22–37). For $r \in \mathbb{N}$ with $r < 1/(2\beta - 1)$, define

$$\sigma_{n,r}^2 = n^{2-r(2\beta-1)} \ell^{2r}(n) \|\varepsilon_0\|^{2r} \frac{[\int_0^\infty (x + x^2)^{-\beta} \,\mathrm{d}x]^r}{r!\,[1 - r(\beta - 1/2)][1 - r(2\beta - 1)]}. \tag{8}$$

Recall Theorem 2 for W_i.

Theorem 4. *Assume that* $\mathrm{E}(\varepsilon_i) = 0$, $\varepsilon_i \in \mathcal{L}^4$, $\frac{1}{2} < \beta < \frac{3}{4}$, $\ell(i+1)/\ell(i) - 1 = O(1/i)$, *and* $k_n/n \to 0$. *Then*

$$\frac{1}{\sigma_{n,2}} \sum_{i=1}^{n} [X_i W_{i-k_n} - \mathrm{E}(X_{k_n} W_0)] \Rightarrow 2R_{2,\beta}\mathbb{I}_h. \tag{9}$$

Theorem 4 allows for a very wide range of k_n, which can be bounded as well as unbounded. An interesting feature of this theorem is that the limiting distribution $R_{2,\beta}$ does not depend on k_n, regardless of whether it is bounded or not. Chung (2002) pointed out that, in the situation that the lag is bounded, the limiting distribution does not depend on the lag. The phenomenon in (9) is interestingly different from Theorems 1 and 2, the mild long-memory case. The latter two theorems assert different limiting distributions in the sense that the asymptotic variances are different, depending on whether k_n is bounded or not.

In Theorems 1–4, we assume that $\varepsilon_i \in \mathcal{L}^4$. If ε_i does not have a finite fourth moment then we may have weak convergence to stable distributions. Recently, Horváth and Kokoszka (2008) obtained various types of convergence rates and limiting distributions, depending on the heaviness of tails and the strength of dependence. In their treatment, however, they assumed that k was bounded. For Theorem 5, below, we assume that $\varepsilon_i^2 - \mathrm{E}\,\varepsilon_i^2$ is in the domain of attraction of a stable distribution Z_α with index $\alpha \in (1, 2)$ (see Chow and Teicher (1988, pp. 448–457)), namely there exists a slowly varying function $\ell_0(\cdot)$ such that

$$\frac{\sum_{i=1}^{n}(\varepsilon_i^2 - \mathrm{E}\,\varepsilon_i^2)}{n^{1/\alpha}\ell_0(n)} \Rightarrow Z_\alpha. \tag{10}$$

In this case the asymptotic behavior of $\check{\gamma}_k$ depends in a very interesting way on the heavy tail index α, the long memory index β, and the lag index λ. Here we let the lag k_n be of the form $n^\lambda \ell_1(n)$, where $\lambda \in (0, 1)$ and ℓ_1 is a slowly varying function.

Theorem 5. *Assume that (10) holds with* $1 < \alpha < 2$ *and* $\frac{3}{4} < \beta < 1$. *Let* $k_n = n^\lambda \ell_1(n)$, *where* $\lambda \in (0, 1)$ *and* ℓ_1 *is a slowly varying function.*

(i) *If* $\lambda > (\alpha^{-1} - 2^{-1})/(2\beta - 1)$ *then (6) holds.*

(ii) *If* $\lambda < (\alpha^{-1} - 2^{-1})/(2\beta - 1)$ *then*

$$\frac{1}{\gamma_{k_n} n^{1/\alpha} \ell_0(n)} \sum_{i=1}^{n} [X_i W_{i-k_n} - \mathrm{E}(X_{k_n} W_0)] \Rightarrow Z_\alpha \mathbb{I}_h. \tag{11}$$

In Theorem 5, cases (i) and (ii) suggest the dichotomy phenomenon: for small λ, we have the weak convergence to stable distributions, while, for large λ, we still have the conventional central limit theorem. A similar phenomenon has been discovered in Csörgő and Mielniczuk (2000) for kernel estimation of long-memory processes. They showed that large and small bandwidths correspond to different asymptotic distributions of the kernel estimates. See also Surgailis (2004), Sly and Heyde (2008), Mikosch *et al.* (2002), and Hsieh *et al.* (2007) for similar observations under different settings. In Theorem 5, the lag parameter k_n plays a similar role. Theorem 5 does not cover the boundary case $\lambda = (\alpha^{-1} - 2^{-1})/(2\beta - 1)$. In this case the situation is more subtle since the growth rates of the slowly varying functions $\ell(\cdot)$, $\ell_0(\cdot)$, and $\ell_1(\cdot)$ will be involved in the limiting distribution. We decide not to pursue the boundary case since the involved manipulations seem quite tedious.

If the dependence of (X_i) is sufficiently strong such that $\frac{1}{2} < \beta < \frac{3}{4}$, then we have a different type of dichotomy. As asserted by Theorem 6, below, the limiting distributions for large and small lags are Rosenblatt and stable distributions, respectively.

Theorem 6. *Assume that (10) holds with* $1 < \alpha < 2$, $\frac{1}{2} < \beta < \frac{3}{4}$, *and* $\ell(i+1)/\ell(i) - 1 = O(1/i)$. *Let* $k_n = n^\lambda \ell_1(n)$, *where* $\lambda \in (0, 1)$ *and* ℓ_1 *is a slowly varying function.*

(i) *If* $2 - 2\beta > \lambda(1 - 2\beta) + \alpha^{-1}$ *then (9) holds.*

(ii) *If* $2 - 2\beta < \lambda(1 - 2\beta) + \alpha^{-1}$ *then (11) holds.*

Remark 1. It is easily seen that Theorems 1–6 are still valid if the sums $\sum_{i=1}^n$ therein are replaced by $\sum_{i=1+k_n}^n$ under the condition that $k_n = o(n)$. For example, let us consider (9) of Theorem 4. Define $n_* = n - k_n$. By (9) and stationarity,

$$\frac{1}{\sigma_{n_*,2}} \sum_{i=1+k_n}^{n} (X_i X_{i-k_n} - \gamma_{k_n}) \Rightarrow 2R_{2,\beta}.$$

Since $n_*/n \to 1$, we have $n_*^{2-2\beta} \ell^2(n_*)/[n^{2-2\beta} \ell^2(n)] \to 1$ by properties of slowly varying functions and, hence,

$$\frac{n\check{\gamma}_{k_n} - (n - k_n)\gamma_{k_n}}{\sigma_{n,2}} = \frac{1}{\sigma_{n,2}} \sum_{i=1+k_n}^{n} (X_i X_{i-k_n} - \gamma_{k_n}) \Rightarrow 2R_{2,\beta}. \tag{12}$$

Similar claims can be made for other theorems. Additionally, the term $(n - k_n)\gamma_{k_n}$ in (12) can be replaced by $n\gamma_{k_n}$ since $k_n \gamma_{k_n} = O[k_n^{2-2\beta} \ell^2(k_n)] = o(\sqrt{n})$ if $\frac{3}{4} < \beta < 1$, $k_n \gamma_{k_n} = o[\sqrt{n} \ell^2(n)] = o[\sqrt{n \bar{\ell}(n)}]$ if $\beta = \frac{3}{4}$, and $k_n \gamma_{k_n} = o(\sigma_{n,2})$ if $\frac{3}{4} > \beta > \frac{1}{2}$.

Remark 2. Under the dependence condition (4), the sample covariance estimator (2) is asymptotically close to $\check{\gamma}_k := n^{-1}\sum_{i=k+1}^{n} X_i X_{i-k}$ since

$$\begin{aligned} n\,\mathrm{E}\,|\hat{\gamma}_k - \check{\gamma}_k| &\le \mathrm{E}\left|\bar{X}_n \sum_{i=k+1}^{n} X_i\right| + \mathrm{E}\left|\bar{X}_n \sum_{i=k+1}^{n} X_{i-k}\right| + \mathrm{E}|(n-k)\bar{X}_n^2| \\ &\le 2\|\bar{X}_n\|\sigma_{n-k,1} + n\|\bar{X}_n\|^2 \\ &= O[n^{2-2\beta}\ell^2(n)], \end{aligned}$$

and $n^{2-2\beta}\ell^2(n) = o(\sqrt{n})$ if $\frac{3}{4} < \beta < 1$ and $n^{2-2\beta}\ell^2(n) = o(\sqrt{n\bar{\ell}(n)})$ if $\beta = \frac{3}{4}$. With simple manipulations, we conclude that Theorems 1–3 and 5 continue to hold if X_i therein is replaced by $X_i - \bar{X}_n$.

If $\frac{1}{2} < \beta < \frac{3}{4}$ then the difference between $\hat{\gamma}_k$ and $\check{\gamma}_k$ is no longer negligible; see Hosking (1996), Dehling and Taqqu (1991), and Yajima (1992). Corollary 1, below, provides the asymptotic distribution of $\hat{\gamma}_k$.

Corollary 1. *Let $\frac{1}{2} < \beta < \frac{3}{4}$. Then, under the conditions of Theorem 4 or Theorem 6(i), we have*

$$\frac{1}{\sigma_{n,2}} \sum_{i=1+k_n}^{n} [(X_i - \bar{X}_n)(X_{i-k_n} - \bar{X}_n) - \gamma_{k_n}] \Rightarrow 2R_{2,\beta} - \frac{2(3-4\beta)^{1/2}}{(1-\beta)^{1/2}(3-2\beta)} R_{1,\beta}^2. \qquad (13)$$

Under Theorem 6(ii), (11) still holds if X_i therein is replaced by $X_i - \bar{X}_n$.

3. Proofs

This section provides proofs for the results in Section 2. Without loss of generality, we assume that $\|\varepsilon_0\| = 1$ throughout the proofs. Let $\kappa_4 = \|\varepsilon_i^2 - 1\|^2$ if $\varepsilon_i \in \mathcal{L}^4$. Define $a_i = 0$ if $i < 0$, and let $A_i = \sum_{j=i}^{\infty} a_j^2$. By Karamata's theorem, $A_n \sim \ell^2(n) n^{1-2\beta}/(2\beta - 1) = O(na_n^2)$. Let

$$\dot{\gamma}_h = \sum_{i\in\mathbb{Z}} |a_i a_{i-h}|.$$

Then, again by Karamata's theorem, as in (1), both γ_h and $\dot{\gamma}_h \sim |h|^{1-2\beta}\ell^2(|h|)C_\beta$ as $|h| \to \infty$. Note that $\dot{\gamma}_h = \gamma_h$ if all $a_i \ge 0$.

3.1. Proofs of Theorems 1 and 2

To prove Theorems 1 and 2, we need the following lemma. With this lemma, we shall first prove Theorem 2 and then prove Theorem 1.

Lemma 1. *Let $i, k \ge 0$. Assume that $\varepsilon_i \in \mathcal{L}^4$. Then*

$$\|\mathcal{P}_0(X_i X_{i-k})\| \le |a_i| A_{i-k+1}^{1/2} + |a_{i-k}| A_{i+1}^{1/2} + |a_i a_{i-k}| \|\varepsilon_0^2 - 1\|. \qquad (14)$$

Note that the above bound is $|a_i| A_0^{1/2}$ if $i < k$. Additionally, under (4), we have

$$\sup_{i_1,i_2,k} \left\| \sum_{i=i_1}^{i_2} \mathcal{P}_0(X_i X_{i-k}) \right\| = O(1) \qquad (15)$$

and

$$\lim_{g\to\infty}\sup_k\left\|\sum_{i=g}^{\infty}\mathcal{P}_0(X_iX_{i-k})\right\|=0. \tag{16}$$

For $l\in\mathbb{Z}$, let $D_{l,k}=\sum_{i\in\mathbb{Z}}\mathcal{P}_l(X_iX_{i-k})$. Then

$$D_{l,k}=\varepsilon_l\sum_{j=-\infty}^{-1}(\gamma_{k+j}+\gamma_{k-j})\varepsilon_{l+j}+\gamma_k(\varepsilon_l^2-1). \tag{17}$$

Proof. Observe that $\mathcal{P}_0(\varepsilon_j\varepsilon_{j'})=0$ if $jj'\neq 0$, and $\mathcal{P}_0\varepsilon_0^2=\varepsilon_0^2-1$. Then

$$\begin{aligned}\mathcal{P}_0(X_iX_{i-k}) &= \sum_{j,j'\in\mathbb{Z}}a_{i-j}a_{i-k-j'}\mathcal{P}_0(\varepsilon_j\varepsilon_{j'})\\ &= \varepsilon_0\sum_{j'=-\infty}^{-1}a_ia_{i-k-j'}\varepsilon_{j'}+\varepsilon_0\sum_{j=-\infty}^{-1}a_{i-j}a_{i-k}\varepsilon_j+a_ia_{i-k}(\varepsilon_0^2-1),\end{aligned} \tag{18}$$

which implies (14). Since $\dot\gamma_h=\sum_{i\in\mathbb{Z}}|a_ia_{i-h}|\sim|h|^{1-2\beta}\ell^2(|h|)C_\beta$ as $|h|\to\infty$, we have

$$\sum_{j'=-\infty}^{-1}\left(\sum_{i=i_1}^{i_2}a_ia_{i-k-j'}\right)^2\leq\sum_{j'=-\infty}^{-1}\dot\gamma_{k+j'}^2=O(1).$$

By (18), (15) follows from a similar argument for $\sum_{i=i_1}^{i_2}a_{i-j}a_{i-k}$. We now prove (16). By Schwarz's inequality, $(\sum_{i=g}^{\infty}a_ia_{i-k-j'})^2\leq A_gA_0\to 0$ as $g\to\infty$. By Lebesgue's dominated convergence theorem, as $g\to\infty$,

$$\sup_k\sum_{j'=-\infty}^{-1}\left(\sum_{i=g}^{\infty}a_ia_{i-k-j'}\right)^2\leq\sup_k\sum_{j'=-\infty}^{\infty}\min(\dot\gamma_{k+j'}^2,A_gA_0)\to 0.$$

With a similar treatment for $\sum_{i=g}^{\infty}a_{i-j}a_{i-k}$, we have (16) since $(\sum_{i=g}^{\infty}a_ia_{i-k})^2\leq A_gA_0$. Since $\gamma_h=\sum_{i\in\mathbb{Z}}a_ia_{i+h}$, (18) implies (17) with $l=0$. The case in which $l\neq 0$ follows similarly.

3.1.1. *Proof of Theorem 2.* Recall (17) of Lemma 1 for $D_{l,k}$. Let $M_{n,k}=\sum_{l=1}^nD_{l,k}$ and $S_{n,k}=\sum_{l=1}^nX_lX_{l-k}-n\gamma_k$. Due to the orthogonality of $\mathcal{P}_r,\ r\in\mathbb{Z}$, we have

$$\|S_{n,k_n}-M_{n,k_n}\|^2=\left(\sum_{r=-\infty}^{0}+\sum_{r=1}^{n}\right)\|\mathcal{P}_r(S_{n,k_n}-M_{n,k_n})\|^2. \tag{19}$$

If $r\leq-3k_n$ and $1\leq i\leq n$, by (14) of Lemma 1 and since $A_j=O(ja_j^2)$ as $j\to\infty$,

$$\|\mathcal{P}_r(X_iX_{i-k_n})\|=O(\sqrt{i-r-k_n}|a_{i-r}a_{i-r-k_n}|)=O(b_{i-r}),$$

where $b_j=j^{1/2-2\beta}\ell^2(j),\ j\in\mathbb{N}$. For $r\leq-3k_n$, we have $\mathcal{P}_rM_{n,k_n}=0$ and

$$\|\mathcal{P}_r(S_{n,k_n}-M_{n,k_n})\|\leq\sum_{i=1}^n\|\mathcal{P}_rX_iX_{i-k_n}\|=\sum_{i=1}^nO(b_{i-r}). \tag{20}$$

Let $p \in (1, (2\beta - 1)^{-1})$ and $q = p/(p-1)$. By Hölder's inequality, if $3k_n \le r \le n$, $\sum_{i=1}^{n} b_{i+r} \le (\sum_{i=1}^{n} b_{i+r}^{p})^{1/p} n^{1/q}$. By Karamata's theorem, $\sum_{i=r}^{\infty} b_i^p = O(r b_r^p)$ since $p(\frac{1}{2} - 2\beta) < -1$. Hence, since $2(1/p + \frac{1}{2} - 2\beta) > -1$, again by Karamata's theorem,

$$\begin{aligned}
\sum_{r=3k_n}^{n} \left(\sum_{i=1}^{n} b_{i+r} \right)^2 &= \sum_{r=1}^{n} O[(r^{1/p} b_r n^{1/q})^2] \\
&= nO[(n^{1/p} b_n n^{1/q})^2] \\
&= O[n^{4-4\beta} \ell^4(n)] \\
&= o(n). \qquad (21)
\end{aligned}$$

If $r > n$ then $\sum_{i=1}^{n} b_{i+r} = O(nb_r)$. Since $1 - 4\beta < -1$, by Karamata's theorem,

$$\sum_{r=1+n}^{\infty} \left(\sum_{i=1}^{n} b_{i+r} \right)^2 = \sum_{r=1+n}^{\infty} O(n^2 b_r^2) = n^3 O(b_n^2) = o(n). \qquad (22)$$

If $1 \le r \le n$, $\mathcal{P}_r(S_{n,k_n} - M_{n,k_n}) = -\sum_{i=n+1}^{\infty} \mathcal{P}_r(X_i X_{i-k_n})$. By stationarity and (16),

$$\sum_{r=1}^{n-3k_n} \|\mathcal{P}_r(S_{n,k_n} - M_{n,k_n})\|^2 = \sum_{g=1+3k_n}^{n} \left\| \sum_{i=g}^{\infty} \mathcal{P}_0(X_i X_{i-k_n}) \right\|^2 = o(n). \qquad (23)$$

Hence, by (15) of Lemma 1, since $k_n = o(n)$, we have, by (19) and (20)–(23),

$$\|S_{n,k_n} - M_{n,k_n}\|^2 = o(n). \qquad (24)$$

It remains to show the central limit theorem for M_{n,k_n}. For a fixed $m \in \mathbb{N}$, let

$$\tilde{M}_{n,k} = \sum_{l=1}^{n} \tilde{D}_{l,k}, \quad \text{where} \quad \tilde{D}_{l,k} = \varepsilon_l \sum_{j=-k-m}^{-k+m} \gamma_{k+j} \varepsilon_{l+j}.$$

Since $D_{l,k} - \tilde{D}_{l,k}$, $l = 1, 2, \ldots$, are martingale differences,

$$\begin{aligned}
\frac{\|\tilde{M}_{n,k} - M_{n,k}\|}{\sqrt{n}} &= \|D_{0,k} - \tilde{D}_{0,k}\| \\
&\le |\gamma_k| \|\varepsilon_0^2 - 1\| + \left(\sum_{j=-\infty}^{-1} \gamma_{k-j}^2 \right)^{1/2} + \left(\sum_{j=-\infty}^{-1} \gamma_{k+j}^2 \mathbf{1}_{\{|j+k|>m\}} \right)^{1/2}.
\end{aligned}$$

Since $\gamma_k \to 0$ as $k \to \infty$ and $\sum_{g \in \mathbb{Z}} \gamma_g^2 < \infty$, we have

$$\limsup_{m\to\infty} \limsup_{n\to\infty} \frac{\|\tilde{M}_{n,k} - M_{n,k}\|}{\sqrt{n}} = 0. \qquad (25)$$

We shall now apply the martingale central limit theorem for $\tilde{M}_{n,k_n}/\sqrt{n}$. By the mean ergodic theorem, since m is fixed, we have

$$\frac{1}{n} \sum_{l=1}^{n} \mathrm{E}(\tilde{D}_{l,k}^2 \mid \mathcal{F}_{l-1}) = \frac{1}{n} \sum_{l=1}^{n} \left(\sum_{j=-k-m}^{-k+m} \gamma_{k+j} \varepsilon_{l+j} \right)^2 \to \sum_{j=-m}^{m} \gamma_j^2$$

in probability. Let $\eta = \varepsilon_0 \sum_{j=-m}^{m} \gamma_j \varepsilon_{j-m-1}$. For any $\lambda > 0$, since

$$\lim_{n\to\infty} \mathrm{E}(\tilde{D}_{l,k}^2 \mathbf{1}_{\{|\tilde{D}_{l,k}|\geq\lambda\sqrt{n}\}}) = \lim_{n\to\infty} \mathrm{E}(\eta^2 \mathbf{1}_{\{|\eta|\geq\lambda\sqrt{n}\}}) = 0,$$

which implies the Lindeberg condition. Hence, $\tilde{M}_{n,k}/\sqrt{n} \Rightarrow N(0, \sum_{j=-m}^{m} \gamma_j^2)$, and the theorem follows from (24) and (25).

3.1.2. *Proof of Theorem 1.* A careful check of the proof of Theorem 2 reveals that, under (4), (24) still holds if k_n is bounded. Namely, for fixed k, we have $\|S_{n,k} - M_{n,k}\|^2 = o(n)$. Then we can just apply the classical martingale central limit theorem and obtain $M_{n,k}/\sqrt{n} \Rightarrow N(0, \|D_{0,k}\|^2)$. Then (5) easily follows from the Cramer–Wold device.

3.2. Proof of Theorem 3

The treatment of the boundary case, $\beta = \frac{3}{4}$, is very intricate. Here we will apply the martingale approximation technique (see Wu and Woodroofe (2004)). We first deal with the case in which $k = 0$. Let

$$V_j = X_j^2 - \gamma_0 - \sum_{l=0}^{\infty} a_l^2 (\varepsilon_{j-l}^2 - 1). \tag{26}$$

We shall approximate $\sum_{j=1}^{n} V_j$ by $\sum_{j=1}^{n} D_{j,n}$, where

$$D_{j,n} = \varepsilon_j \sum_{h=1}^{\infty} 2c_{n,h}\varepsilon_{j-h}, \quad \text{where} \quad c_{n,h} = \sum_{i=0}^{n-1} a_i a_{i+h}. \tag{27}$$

Note that $D_{1,n}, D_{2,n}, \ldots, D_{n,n}$ are martingale differences. Let

$$R_n = \sum_{j=1}^{n} (V_j - D_{j,n}).$$

Next we shall control $\|R_n\|$. Since the $\mathcal{P}_h,\ h \in \mathbb{Z}$, are orthogonal,

$$\|R_n\|^2 = \left(\sum_{h=-\infty}^{-n} + \sum_{h=1-n}^{0} + \sum_{h=1}^{n}\right) \|\mathcal{P}_h R_n\|^2. \tag{28}$$

If $h \leq -n$ then $\mathcal{P}_h R_n = \sum_{i=1}^{n} \mathcal{P}_h V_i$, and, by independence,

$$\begin{aligned}
\sum_{h=-\infty}^{-n} \|\mathcal{P}_h R_n\|^2 &= \sum_{h=-\infty}^{-n} \left\| \sum_{i=1}^{n} 2a_{i-h}\varepsilon_h \sum_{j=1}^{\infty} a_{i-h+j}\varepsilon_{h-j} \right\|^2 \\
&\leq 4 \sum_{h=-\infty}^{-n} \sum_{j=1}^{\infty} \left(\sum_{i=1}^{n} a_{i-h} a_{i-h+j} \right)^2 \\
&= \sum_{h=-\infty}^{-n} \sum_{j=1}^{\infty} O(n|a_{-h} a_{j-h}|)^2 \\
&= \sum_{h=-\infty}^{-n} O(n^2 a_{-h}^2 |h| a_{-h}^2|) \\
&= o(n\bar{\ell}(n)).
\end{aligned} \tag{29}$$

In (29) we have applied Karamata's theorem by noting the fact that, if $h \le -n$ and $1 \le i \le n$, then $a_{i-h} = O(a_{-h})$. By Lemma 4 of Wu and Min (2005), $\ell^4(n) = o(\bar{\ell}(n))$.

Let $\delta \in (0, \frac{1}{2})$. For $1 + \lfloor n\delta \rfloor \le h \le 2n$ and $1 \le j \le n$, we have

$$\sum_{l=h}^{2n} |a_l a_{l+j}| = \sum_{l=1}^{2n} O(a_n^2) = O(n^{-1/2}\ell^2(n)) = n^{-1/2} o[\bar{\ell}(n)^{1/2}]. \tag{30}$$

Therefore, since $\dot{\gamma}_j \sim j^{-1/2}\ell^2(j)C_\beta$, we have $\sum_{j=1}^{n} \dot{\gamma}_j^2 \sim \bar{\ell}(n)C_\beta^2$. By (30),

$$\begin{aligned}
\limsup_{n\to\infty} \frac{\sum_{h=1}^{n}\sum_{j=1}^{n}(\sum_{l=h}^{2n}|a_l a_{l+j}|)^2}{n\bar{\ell}(n)} &\le \limsup_{n\to\infty} \frac{\sum_{h=1}^{\lfloor n\delta\rfloor}\sum_{j=1}^{n}(\sum_{l=h}^{2n}|a_l a_{l+j}|)^2}{n\bar{\ell}(n)} \\
&\quad + \limsup_{n\to\infty} \frac{\sum_{h=1+\lfloor n\delta\rfloor}^{n}\sum_{j=1}^{n}(\sum_{l=h}^{2n}|a_l a_{l+j}|)^2}{n\bar{\ell}(n)} \\
&\le \limsup_{n\to\infty} \frac{\sum_{h=1}^{\lfloor n\delta\rfloor}\sum_{j=1}^{n}\dot{\gamma}_j^2}{n\bar{\ell}(n)} \\
&= \delta C_\beta^2.
\end{aligned} \tag{31}$$

Since $\delta > 0$ can be arbitrarily small, $\sum_{h=1}^{n}\sum_{j=1}^{n}(\sum_{l=h}^{2n}|a_l a_{l+j}|)^2 = o[n\bar{\ell}(n)]$. Next,

$$\begin{aligned}
\frac{\sum_{h=1}^{n}\sum_{j=1+n}^{\infty}(\sum_{l=h}^{2n}|a_l a_{l+j}|)^2}{n\bar{\ell}(n)} &= \frac{\sum_{h=1}^{n}\sum_{j=1+n}^{\infty}O(a_j^2)(\sum_{l=h}^{2n}|a_l|)^2}{n\bar{\ell}(n)} \\
&= \frac{nO(na_n^2)(\sum_{l=1}^{2n}|a_l|)^2}{n\bar{\ell}(n)} \\
&= \frac{O(\ell^4(n))}{\bar{\ell}(n)} \\
&= o(1).
\end{aligned} \tag{32}$$

We now deal with the sums $\sum_{h=1-n}^{0}$ and $\sum_{h=1}^{n}$ in (28). By (31) and (32),

$$\begin{aligned}
\sum_{h=1-n}^{0} \|\mathcal{P}_h R_n\|^2 &= \sum_{h=1-n}^{0} \left\| \sum_{i=1}^{n} 2a_{i-h}\varepsilon_h \sum_{j=1}^{\infty} a_{i-h+j}\varepsilon_{h-j} \right\|^2 \\
&\le 4 \sum_{h=1-n}^{0}\sum_{j=1}^{\infty}\left(\sum_{i=1}^{n} a_{i-h}a_{i-h+j}\right)^2 \\
&= o[n\bar{\ell}(n)].
\end{aligned} \tag{33}$$

For $1 \le h \le n$, we have

$$\mathcal{P}_h R_n = \sum_{i=h}^{n} \mathcal{P}_h V_i - D_{h,n} = \sum_{i=n+1}^{n+h-1} \mathcal{P}_h V_i.$$

By stationarity, (31), and (32),

$$\begin{aligned}\sum_{h=1}^{n}\|\mathcal{P}_h R_n\|^2 &= \sum_{h=1}^{n}\left\|\sum_{i=n+1}^{n+h-1}\mathcal{P}_h V_i\right\|^2\\ &= \sum_{h=1}^{n}\left\|\sum_{l=n+1-h}^{n-1}\mathcal{P}_0 V_l\right\|^2\\ &\le 4\sum_{h'=1}^{n}\sum_{j=1}^{\infty}\left(\sum_{l=h'}^{n-1}a_l a_{l+j}\right)^2\\ &= o[n\bar{\ell}(n)]. \end{aligned}\tag{34}$$

Therefore, by (28) we have

$$\|R_n\|^2 = o[n\bar{\ell}(n)]. \tag{35}$$

We now further approximate $\sum_{i=1}^n D_{i,n}$ by

$$\sum_{i=1}^{n} H_i, \quad \text{where} \quad H_i = H_{i,n} = \varepsilon_i \sum_{j=1}^{n} 2\gamma_j \varepsilon_{i-j}. \tag{36}$$

Note that $\|H_i\|^2 = 4\sum_{j=1}^n \gamma_j^2 \sim 4C_\beta^2 \sum_{j=1}^n \ell^4(j)/j$. Since $\ell^4(n) = o(\bar{\ell}(n))$, we have

$$\left\|\sum_{i=1}^{n} H_i - \sum_{i=1}^{n} D_{i,n}\right\|^2 = o(n\bar{\ell}(n)) \tag{37}$$

in view of

$$\sum_{j=1}^{n}(c_{n,j}-\gamma_j)^2 \le \sum_{j=1}^{n}\left(\sum_{i=n}^{\infty}|a_i a_{i+j}|\right)^2 = \sum_{j=1}^{n} O(nn^{-2\beta}\ell^2(n))^2 = o(\bar{\ell}(n))$$

since $\beta = \frac{3}{4}$. It remains to show that

$$\frac{\sum_{i=1}^n H_i}{(n\bar{\ell}(n))^{1/2}} \Rightarrow N(0, 4C_\beta^2). \tag{38}$$

To this end, we shall apply the martingale central limit theorem. The Lindeberg condition trivially holds since

$$\frac{\mathrm{E}\, H_i^4}{\bar{\ell}(n)^2} = \frac{\mathrm{E}(\sum_{j=1}^n \gamma_j \varepsilon_{-j})^4}{\bar{\ell}(n)^2} \le C\frac{(\sum_{j=1}^n \gamma_j^2)^2}{\bar{\ell}(n)^2} = O(1)$$

for some constant $C > 0$ in view of Rosenthal's inequality (see Hall and Heyde (1980, p. 23)). It then suffices to verify the following convergence of conditional variances:

$$\frac{1}{n\bar{\ell}(n)}\sum_{i=1}^{n}\mathrm{E}(H_i^2 \mid \mathcal{F}_{i-1}) = \frac{4}{n\bar{\ell}(n)}\sum_{i=1}^{n}\left(\sum_{j=1}^{n}\gamma_j\varepsilon_{i-j}\right)^2 \to 4C_\beta^2 \tag{39}$$

in probability. By the mean ergodic theorem, $\mathrm{E}|\sum_{j=1}^{n} \varepsilon_j^2 - n| = o(n)$. Hence,

$$\mathrm{E}\left|\sum_{i=1}^{n}\sum_{j=1}^{n} \gamma_j^2(\varepsilon_{i-j}^2 - 1)\right| = \sum_{j=1}^{n} \gamma_j^2 o(n) = o(n\bar{\ell}(n)).$$

Hence, for (39), it remains to deal with the cross product terms

$$\sum_{i=1}^{n} \sum_{1\le j\ne j'\le n} \gamma_j\gamma_{j'}\varepsilon_{i-j}\varepsilon_{i-j'} = \sum_{1-n\le l\ne l'\le n-1} \varepsilon_l\varepsilon_{l'} f_{l,l'},$$

where the coefficients

$$f_{l,l'} = \sum_{i=1+\max(l,l',0)}^{n+\min(l,l',0)} \gamma_{i-l}\gamma_{i-l'}.$$

Note that $|f_{l,l'}| \le \sum_{i=1}^{n} |\gamma_i\gamma_{i+|l-l'|}| =: \mu_{|l-l'|}$. By independence,

$$\left\|\sum_{1-n\le l\ne l'\le n-1} \varepsilon_l\varepsilon_{l'} f_{l,l'}\right\|^2 \le \sum_{1-n\le l\ne l'\le n-1} 2f_{l,l'}^2 \le 8n\sum_{i=1}^{2n} \mu_i^2.$$

Let $0 < \delta < \frac{1}{2}$ and $l \ge \delta n$. Then

$$\mu_l \le \sum_{i=1}^{n} |\gamma_i| O(n^{1-2\beta}\ell^2(n)) = O(\ell^4(n)).$$

So

$$\limsup_{n\to\infty} \frac{n(\sum_{i=1}^{n\delta} + \sum_{i=1+n\delta}^{2n})\mu_i^2}{n^2\bar{\ell}^2(n)} \le \limsup_{n\to\infty} \frac{n^2\delta\mu_0^2}{n^2\bar{\ell}^2(n)} = C_\beta^2\delta.$$

Let $\delta \to 0$. Then $n\sum_{i=1}^{2n} \mu_i^2 = o(n^2\bar{\ell}^2(n))$ and, hence, (39) follows. By the expression of V_j in (26), since $\varepsilon_l^2 - 1 \in \mathcal{L}^2$, we have $\|\sum_{j=1}^{n}(\varepsilon_j^2 - 1)\|^2 \le \kappa_4 n$ and

$$\left\|\sum_{j=1}^{n}(X_j^2 - \gamma_0 - V_j)\right\| \le \sum_{l=0}^{\infty} a_l^2 \left\|\sum_{j=1}^{n}(\varepsilon_{j-l}^2 - 1)\right\| = O(\sqrt{n}). \tag{40}$$

So, if $k = 0$, since $\bar{\ell}(n) \to \infty$, (7) with $k = 0$ follows from (35), (37), (38), and (40). For the general case with finite $k > 0$, we replace V_j in (26) by

$$V_{j,k} = X_jX_{j-k} - \gamma_k - \sum_{l=0}^{\infty} a_la_{l+k}(\varepsilon_{j-k-l}^2 - 1).$$

If we replace $2c_{n,h}$ in (27) by $\sum_{j=0}^{n-1}(a_{h+j}a_{j-k} + a_ja_{j+h-k})$ and H_j in (36) by

$$H_i^{(k)} := \varepsilon_i \sum_{j=1}^{n}(\gamma_{j+k} + \gamma_{j-k})\varepsilon_{i-j},$$

using the argument for $k = 0$, we similarly have

$$\left\| \sum_{j=1}^{n} (X_j X_{j-k} - \gamma_k) - \sum_{i=1}^{n} H_i^{(k)} \right\|^2 = o(n\bar{\ell}(n)).$$

So (7) follows if $\| \sum_{i=1}^{n} (H_i - H_i^{(k)}) \|^2 = o(n\bar{\ell}(n))$, which is equivalent to

$$\|H_0 - H_0^{(k)}\|^2 = \sum_{j=1}^{n} (2\gamma_j - \gamma_{j+k} - \gamma_{j-k})^2 = o(\bar{\ell}(n)).$$

By (1), as $j \to \infty$, $\gamma_{j+k}/\gamma_j \to 1$. So the above relation holds since $\bar{\ell}(n) \to \infty$ and $\sum_{j=1}^{n} [(\gamma_j - \gamma_{j+k})^2 + (\gamma_j - \gamma_{j-k})^2] = \sum_{j=1}^{n} o(\gamma_j^2) = o(\bar{\ell}(n))$.

3.3. Proof of Theorem 4

By Lemmas 2 and 3, we have

$$\left\| \sum_{i=1}^{n} (X_i^2 - X_i X_{i-k_n} - \gamma_0 + \gamma_{k_n}) \right\|^2 = O(n k_n^{3-4\beta} \ell^4(k_n)). \tag{41}$$

By properties of slowly varying functions we have $k_n^{3-4\beta} \ell^4(k_n) = o(n^{3-4\beta} \ell^4(n))$ under $k_n = o(n)$. It is well known that (see, for example, Avram and Taqqu (1987)), for $\frac{1}{2} < \beta < \frac{3}{4}$, we have

$$\frac{\sum_{i=1}^{n} (X_i^2 - \gamma_0)}{\sigma_{n,2}} \Rightarrow 2R_{2,\beta}.$$

Hence, Theorem 4 follows.

Lemma 2. *Assume that* $\varepsilon_i \in \mathcal{L}^4$, $\frac{1}{2} < \beta < \frac{3}{4}$, *and* $k_n/n \to 0$. *Then*

$$\left\| \sum_{i=1}^{n} [X_i X_{i-k_n} - \mathrm{E}(X_i X_{i-k_n} \mid \mathcal{F}_{i-k_n})] \right\|^2 = O(n k_n^{3-4\beta} \ell^4(k_n)) \tag{42}$$

and

$$\left\| \sum_{i=1}^{n} [X_i^2 - \mathrm{E}(X_i^2 \mid \mathcal{F}_{i-k_n})] \right\|^2 = O(n k_n^{3-4\beta} \ell^4(k_n)). \tag{43}$$

Proof. Let $X_i^* = X_i - \mathrm{E}(X_i \mid \mathcal{F}_{i-k_n})$. Since $X_i X_{i-k_n} - \mathrm{E}(X_i X_{i-k_n} \mid \mathcal{F}_{i-k_n}) = X_{i-k_n} X_i^*$,

$$\begin{aligned} \sum_{i=1}^{n} [X_i X_{i-k_n} - \mathrm{E}(X_i X_{i-k_n} \mid \mathcal{F}_{i-k_n})] &= \sum_{i=1}^{n} X_{i-k_n} X_i^* \\ &= \sum_{j=2-k_n}^{n} \varepsilon_j \sum_{i=\max(j,1)}^{\min(n,j+k_n-1)} a_{i-j} X_{i-k_n}. \end{aligned} \tag{44}$$

Since the ε_i are i.i.d., (42) follows from the fact that, for $-k_n \le j \le n$,

$$\begin{aligned}\left\|\sum_{i=\max(j,1)}^{\min(n,j+k_n-1)} a_{i-j} X_{i-k_n}\right\|^2 &= \sum_{i,i'=\max(j,1)}^{\min(n,j+k_n-1)} a_{i-j} a_{i'-j} \,\mathrm{E}(X_{i-k_n} X_{i'-k_n}) \\ &\le \sum_{m=1-k_n}^{k_n-1} |\gamma_m \dot{\gamma}_m| \\ &= O(k_n^{3-4\beta} \ell^4(k_n)). \end{aligned} \tag{45}$$

We now prove (43). Since $X_i^* = \sum_{j=i-k_n+1}^{i} a_{i-j}\varepsilon_j$, we have

$$X_i^2 - \mathrm{E}(X_i^2 \mid \mathcal{F}_{i-k_n}) = (X_i^*)^2 - \mathrm{E}(X_i^*)^2 + 2X_i^* \sum_{g=k_n}^{\infty} a_g \varepsilon_{i-g}.$$

Similarly as the argument in (44) and (45) for (42), we have

$$\left\|\sum_{i=1}^{n} X_i^* \sum_{g=k_n}^{\infty} a_g \varepsilon_{i-g}\right\|^2 = O(n k_n^{3-4\beta} \ell^4(k_n)).$$

It therefore remains to verify that

$$\left\|\sum_{i=1}^{n} [(X_i^*)^2 - \mathrm{E}(X_i^*)^2]\right\|^2 = \sum_{h=-k_n}^{n} \left\|\sum_{i=1}^{n} \mathcal{P}_h (X_i^*)^2\right\|^2 = O(n k_n^{3-4\beta} \ell^4(k_n)). \tag{46}$$

To this end, uniformly over $h = -k_n, 1-k_n, \ldots, n$, we have

$$\begin{aligned}\left\|\sum_{i=1}^{n} \mathcal{P}_h (X_i^*)^2\right\|^2 &= \left\|\sum_{i=\max(h,1)}^{\min(n,h+k_n-1)} \left[a_{i-h}^2(\varepsilon_h^2 - 1) + 2a_{i-h}\varepsilon_h \sum_{j=i-k_n+1}^{h-1} a_{i-j}\varepsilon_j\right]\right\|^2 \\ &\le 2\gamma_0^2 \|\varepsilon_0^2 - 1\|^2 + 8\left\|\sum_{j=\max(h,1)-k_n+1}^{h-1} \varepsilon_j \sum_{i=\max(h,1)}^{\min(n,h+k_n-1,j+k_n-1)} a_{i-j} a_{i-h}\right\|^2 \\ &\le 2\gamma_0^2 \|\varepsilon_0^2 - 1\|^2 + 8 \sum_{j=\max(h,1)-k_n+1}^{h-1} \dot{\gamma}_{j-h}^2 \\ &\le 2\gamma_0^2 \|\varepsilon_0^2 - 1\|^2 + 8 \sum_{m=1}^{k_n-1} \dot{\gamma}_m^2 \\ &= O(k_n^{3-4\beta} \ell^4(k_n)). \end{aligned}$$

So (46) holds and the proof of Lemma 2 is now complete.

Lemma 3. *Under the conditions of Theorem 4, we have*

$$\left\|\sum_{i=1}^{n} [\mathrm{E}(X_i^2 - X_i X_{i-k_n} \mid \mathcal{F}_{i-k_n}) - \gamma_0 + \gamma_{k_n}]\right\|^2 = O(n k_n^{3-4\beta} \ell^4(k_n)). \tag{47}$$

Proof. Let $d_i = a_i - a_{i-k_n}$. For $i \geq k_n$, we have $\sum_{j=i}^{\infty} d_j^2 \leq 4A_{i-k_n}$, and

$$\begin{aligned}
\|\mathcal{P}_0(X_i^2 - X_i X_{i-k_n})\| &= \left\| a_i \varepsilon_0 \sum_{j=i+1}^{\infty} d_j \varepsilon_{i-j} + d_i \varepsilon_0 \sum_{j=i+1}^{\infty} a_j \varepsilon_{i-j} + a_i d_i (\varepsilon_0^2 - 1) \right\| \\
&\leq |a_i| \left(\sum_{j=i+1}^{\infty} d_j^2 \right)^{1/2} + |d_i| A_{i+1}^{1/2} + \|a_i d_i (\varepsilon_0^2 - 1)\| \\
&\leq 2|a_i| A_{i+1-k_n}^{1/2} + |d_i| A_{i+1}^{1/2} + \|a_i d_i (\varepsilon_0^2 - 1)\|. \qquad (48)
\end{aligned}$$

If $i \geq 2k_n$, since $\ell(i+1)/\ell(i) - 1 = O(1/i)$, we have $a_{i+1} - a_i = O(a_i/i)$ and $d_i = O(|a_i| k_n / i)$. By Karamata's theorem, since $\sqrt{A_i} = O(\sqrt{i}|a_i|)$, we have, by elementary calculations,

$$\sum_{i=2k_n}^{\infty} |d_i| A_{i+1}^{1/2} = \sum_{i=2k_n}^{\infty} O\left(\frac{|a_i| k_n}{i} \right) O(\sqrt{i}|a_i|) = O(k_n^{3/2-2\beta} \ell^2(k_n)), \qquad (49)$$

$$\sum_{i=2k_n}^{\infty} |a_i| A_{i+1-k_n}^{1/2} = \sum_{i=2k_n}^{\infty} O[|a_i a_{i+1-k_n}| (i+1-k_n)^{1/2}] = O(k_n^{3/2-2\beta} \ell^2(k_n)) \qquad (50)$$

since, for $i \geq 2k_n$, $a_i a_{i+1-k_n} = O(a_i^2)$, and

$$\sum_{i=2k_n}^{\infty} |a_i d_i| = \sum_{i=2k_n}^{\infty} O\left(\frac{a_i^2 k_n}{i} \right) = O(k_n^{1-2\beta} \ell^2(k_n)). \qquad (51)$$

For $k_n \leq i < 2k_n$, since $a_i = O(a_{k_n})$, we have

$$\sum_{i=k_n}^{2k_n-1} |a_i| A_{i+1-k_n}^{1/2} = O(a_{k_n}) \sum_{i=k_n}^{2k_n-1} A_{i+1-k_n}^{1/2} = O(k_n^{3/2-2\beta} \ell^2(k_n)), \qquad (52)$$

and, since $\sum_{i=k_n}^{2k_n-1} |d_i| \leq 2 \sum_{i=0}^{2k_n-1} |a_i| = O(k_n a_{k_n})$ and $A_{i+1} = O(k_n a_{k_n}^2)$,

$$\sum_{i=k_n}^{2k_n-1} |d_i| A_{i+1}^{1/2} = O(k_n^{1/2} |a_{k_n}|) \sum_{i=k_n}^{2k_n-1} |d_i| = O(k_n^{3/2-2\beta} \ell^2(k_n)). \qquad (53)$$

By Theorem 1 of Wu (2007) we have

$$\begin{aligned}
\left\| \sum_{i=1}^{n} [\mathrm{E}(X_i^2 - X_i X_{i-k_n} \mid \mathcal{F}_{i-k_n}) - \gamma_0 + \gamma_{k_n}] \right\| &\leq \sqrt{n} \sum_{i=0}^{\infty} \|\mathcal{P}_0 \, \mathrm{E}(X_i^2 - X_i X_{i-k_n} \mid \mathcal{F}_{i-k_n})]\| \\
&= \sqrt{n} \sum_{i=k_n}^{\infty} \|\mathcal{P}_0 (X_i^2 - X_i X_{i-k_n})\|,
\end{aligned}$$

which, by inequalities (48)–(53), implies (47).

3.4. Proof of Theorem 5

As (26), we define

$$V_{j,k} = X_j X_{j-k} - \gamma_k - \sum_{l=0}^{\infty} a_l a_{l+k} (\varepsilon_{j-k-l}^2 - 1). \tag{54}$$

A careful check of the proof of Theorem 2 implies that (6) holds if $X_i X_{i-k} - \gamma_k$ therein is replaced by $V_{i,k}$. Indeed, if $X_i X_{i-k}$ in Lemma 1 is replaced by $V_{i,k}$, then (14) becomes

$$\|\mathcal{P}_0 V_{i,k}\| \le |a_i| A_{i-k+1}^{1/2} + |a_{i-k}| A_{i+1}^{1/2}$$

under the condition $\varepsilon_i \in \mathcal{L}^2$ and we do not need to impose $\varepsilon_i \in \mathcal{L}^4$. Also, (15) and (16) hold with $X_i X_{i-k}$ therein being replaced by $V_{i,k}$, and the approximating martingale differences $D_{l,k}$ in (17) now become

$$D_{l,k}^{\circ} = \varepsilon_l \sum_{j=-\infty}^{-1} (\gamma_{k+j} + \gamma_{k-j}) \varepsilon_{l+j}.$$

The proof of Theorem 2 is still valid if we replace $M_{n,k}$ by $M_{n,k}^{\circ} = \sum_{l=1}^{n} D_{l,k}^{\circ}$.

Let p satisfy $\alpha > p > \max(1, \alpha\lambda)$ and $(2\beta - 1)(1 - \lambda) + \alpha^{-1} > p^{-1}$. Since $\beta > \frac{1}{2}$ and $\lambda \in (0, 1)$, such a p always exists. Since $\varepsilon_i^2 - 1$ satisfies (10) and $p < \alpha$, $\mathrm{E}|\varepsilon_i^2 - 1|^p < \infty$.

(i) By the argument above, it suffices to show that

$$Q_n := \sum_{j=1}^{n} \sum_{l=0}^{\infty} a_l a_{l+k_n} (\varepsilon_{j-k_n-l}^2 - 1) = \sum_{g \in \mathbb{Z}} \sum_{j=1}^{n} a_{j-g} a_{j-g+k_n} (\varepsilon_{g-k_n}^2 - 1)$$

satisfies $\|Q_n\|_p = o(\sqrt{n})$. By Burkholder's and Minkowski's inequalities,

$$\begin{aligned}\|Q_n\|_p^p &\le C_p \sum_{g \in \mathbb{Z}} \left| \sum_{j=1}^{n} a_{j-g} a_{j-g+k_n} \right|^p \|\varepsilon_0^2 - 1\|_p^p \\ &= O(1) \sum_{g=-\infty}^{0} \left| \sum_{j=1}^{n} a_{j-g} a_{j-g+k_n} \right|^p + O(n) \left| \sum_{j=0}^{n} |a_j a_{j+k_n}| \right|^p.\end{aligned}$$

Since $\lambda > (\alpha^{-1} - 2^{-1})/(2\beta - 1)$, we can choose a $p < \alpha$ such that $p^{-1} + \lambda(1 - 2\beta) < 2^{-1}$. So

$$n \left| \sum_{j=0}^{n} |a_j a_{j+k_n}| \right|^p = O(n \dot{\gamma}_{k_n}^p) = O\{[n^{1/p} k_n^{1-2\beta} \ell^2(k_n)]^p\} = o(\sqrt{n}^p),$$

since $k_n = n^{\lambda} \ell_1(n)$ and ℓ_1 is a slowly varying function. Hence, similarly,

$$\sum_{g=1-n}^{0} \left| \sum_{j=1}^{n} a_{j-g} a_{j-g+k_n} \right|^p \le n \left| \sum_{j=1}^{2n} |a_j a_{j+k_n}| \right|^p = o(\sqrt{n}^p).$$

If $g \le -n$, by properties of slowly varying functions, for $1 \le j \le n$ and $k_n < n$, $a_{j-g} a_{j-g+k_n} = O(a_{-g}^2)$. Hence,

$$\sum_{g=-\infty}^{-n} \left| \sum_{j=1}^{n} a_{j-g} a_{j-g+k_n} \right|^p = \sum_{g=-\infty}^{-n} O[(n a_{-g}^2)^p] = O(n^{p+1} a_n^{2p}) = o(\sqrt{n}^p)$$

in view of $2^{-1} + p^{-1} < 2\beta$ since $1 < p < 2$ and $\beta > \frac{3}{4}$.

(ii) We first show that (11) holds with $h = 1$. Introduce

$$T_n = T_{n,k_n} = \sum_{j=1}^{n}\sum_{l=0}^{\infty} a_l a_{l+k_n}(\varepsilon_{j-k_n-l}^2 - 1) - \gamma_{k_n}\sum_{j=1}^{n}(\varepsilon_{j-k_n}^2 - 1).$$

Under $2^{-1} < \lambda(1-2\beta) + \alpha^{-1}$, we have $n^{1/2} = o[\gamma_{k_n} n^{1/\alpha}\ell_0(n)]$. Since (6) holds with $X_i X_{i-k_n} - \gamma_{k_n}$ therein being replaced by V_{i,k_n}, by (10), it suffices to show that

$$\|T_n^\circ\|_p = o[\gamma_{k_n} n^{1/\alpha}\ell_0(n)], \tag{55}$$

where

$$T_n^\circ := \sum_{j=1}^{n}\sum_{l=0}^{\infty} a_l a_{l+k_n}(\varepsilon_{j-l}^2 - 1) - \gamma_{k_n}\sum_{j=1}^{n}(\varepsilon_j^2 - 1)$$

has the same distribution as T_n. To this end, note that $\mathcal{P}_l T_n^\circ$, $l = -\infty, \dots, n-1, n$, are martingale differences, we have, by Burkholder's and Minkowski's inequalities,

$$\begin{aligned}\|T_n^\circ\|_p^p &\le C_p\left(\sum_{l=-\infty}^{0} + \sum_{l=1}^{n}\right)\|\mathcal{P}_l T_n^\circ\|_p^p \\ &\le C_p\|\varepsilon_0^2 - 1\|_p^p\left(\sum_{l=-\infty}^{0}\left|\sum_{j=1}^{n} a_{j-l}a_{j-l+k_n}\right|^p + \sum_{l=1}^{n}\left|\sum_{j=1}^{n} a_{j-l}a_{j-l+k_n} - \gamma_{k_n}\right|^p\right).\end{aligned} \tag{56}$$

We shall apply the technique in (28)–(35). Clearly,

$$\sum_{l=1}^{n}\left|\sum_{j=1}^{n} a_{j-l}a_{j-l+k_n} - \gamma_{k_n}\right|^p = \sum_{l'=1}^{n}\left|\sum_{j=l'}^{\infty} a_j a_{j+k_n}\right|^p.$$

If $j \ge k_n$ then $a_j a_{j+k_n} = O(a_j^2)$. Hence,

$$\sum_{l=k_n}^{n}\left|\sum_{j=l}^{\infty} a_j a_{j+k_n}\right|^p = \sum_{l=k_n}^{n}\left|\sum_{j=l}^{\infty} O(a_j^2)\right|^p = \sum_{l=k_n}^{n} O[(la_l^2)^p] = \sum_{l=k_n}^{n} O[l^{p(1-2\beta)}\ell^{2p}(l)]. \tag{57}$$

If $p(1-2\beta) > -1$ then, by Karamata's theorem, the above term is $O[n^{1+p(1-2\beta)}\ell^{2p}(n)]$, which is $o[k_n^{1-2\beta}\ell^2(k_n)n^{1/\alpha}\ell_0(n)] = o[\gamma_{k_n} n^{1/\alpha}\ell_0(n)]$ since $1+p(1-2\beta) < \lambda(1-2\beta)+\alpha^{-1}$. If $p(1-2\beta) \le -1$, it is easily seen that the above term is $o(\sqrt{n})$, which is $o[\gamma_{k_n} n^{1/\alpha}\ell_0(n)]$ since $2^{-1} < \alpha^{-1} + \lambda(1-2\beta)$.

Since $\lambda < p/\alpha$, we have

$$\sum_{l=0}^{k_n-1}\left|\sum_{j=l}^{\infty} a_j a_{j+k_n}\right|^p = O(k_n\gamma_{k_n}^p) = o[\gamma_{k_n} n^{1/\alpha}\ell_0(n)]^p. \tag{58}$$

If $l \ge n$ and $1 \le j \le n$, then $(\sum_{j=1}^{n}|a_{j+l}a_{j+l+k_n}|)^p = O[(na_l^2)^p]$. By Karamata's theorem,

$$\sum_{l=n}^{\infty}\left|\sum_{j=1}^{n} a_{j+l}a_{j+l+k_n}\right|^p = \sum_{l=n}^{\infty} O[(na_l^2)^p] = O(n^{p+1}a_n^{2p}) = o[\gamma_{k_n} n^{1/\alpha}\ell_0(n)]^p, \tag{59}$$

since $1 + p(1 - 2\beta) < p\lambda(1 - 2\beta) + p/\alpha$. Hence, by (56)–(59) we have

$$\|T_n^\circ\|_p^p \le C_p \sum_{l=0}^{\infty} \left| \sum_{j=1}^{n} a_{j+l} a_{j+l+k_n} \right|^p + C_p \sum_{l'=1}^{n} \left| \sum_{j=l'}^{\infty} a_j a_{j+k_n} \right|^p = o[\gamma_{k_n} n^{1/\alpha} \ell_0(n)]^p,$$

which implies (55) and, hence, case (ii) with $h = 1$. For the case with $h > 1$, let

$$U_k = \gamma_k \sum_{j=1}^{n} (\varepsilon_{j-k}^2 - 1).$$

By (1), $\gamma_{k_n} - \gamma_{k_n+h} = o(\gamma_{k_n})$. So (11) follows from (55) and

$$U_{k_n} - U_{k_n+h} = (\gamma_{k_n} - \gamma_{k_n+h}) \sum_{j=1}^{n} (\varepsilon_{j-k_n}^2 - 1) + \gamma_{k_n+h} O_{\mathbb{P}}(1) = o_{\mathbb{P}}(U_{k_n}) + \gamma_{k_n+h} O_{\mathbb{P}}(1).$$

3.5. Proof of Theorem 6

The argument in the proof of Theorem 5 can be easily modified to prove Theorem 6. For $V_{j,k}$ defined in (54), under $\frac{1}{2} < \beta < \frac{3}{4}$, we can similarly have the noncentral limit theorem $\sum_{j=1}^{n} V_{j,k}/\sigma_{n,2} \Rightarrow 2R_{2,\beta}$. Then we need to compare the magnitudes of $n^{2-2\beta}\ell^2(n)$ and $\gamma_{k_n} n^{1/\alpha} \ell_0(n)$. Under (i), the former is larger, and we have the noncentral limit theorem (9); under (ii), we have the convergence in stable distribution (11). The details are omitted since there will be no essential extra difficulties involved.

3.6. Proof of Corollary 1

By Lemma 4, $\|\sum_{i=1}^{m} X_i\| \sim \sigma_{m,1}$. Since $\check{\gamma}_k = n^{-1} \sum_{i=k+1}^{n} X_i X_{i-k}$, by simple algebra,

$$\begin{aligned} \mathrm{E}[|n(\hat{\gamma}_{k_n} - \check{\gamma}_{k_n} + \bar{X}_n^2)|] &= \mathrm{E}\left| \bar{X}_n \sum_{i=n-k_n+1}^{n} X_i + \bar{X}_n \sum_{i=1}^{k_n} X_i - k_n \bar{X}_n^2 \right| \\ &\le \frac{2\sigma_{n,1}\sigma_{k_n,1}}{n} + \frac{k_n \sigma_{n,1}^2}{n^2} \\ &= o[n^{2-2\beta}\ell^2(n)], \end{aligned} \tag{60}$$

in view of $k_n = o(n)$. Let $Y_{n,r}$ be as given in (61), below. Then $Y_{n,1} = n\bar{X}_n$ and

$$n\check{\gamma}_0 = \sum_{i=1}^{n} X_i^2 = 2Y_{n,2} + \sum_{i=-\infty}^{n} \left(\sum_{t=1}^{n} a_{t-i}^2 \right) \varepsilon_i^2.$$

By Lemma 4, below, we have the joint convergence $(Y_{n,1}/\sigma_{n,1}, Y_{n,2}/\sigma_{n,2}) \Rightarrow (R_{1,\beta}, R_{2,\beta})$. Hence, by (60), we have (13) in view of (41), and, by elementary calculations,

$$\frac{\sigma_{n,1}^2}{n\sigma_{n,2}} \to \frac{2(3 - 4\beta)^{1/2}}{(1 - \beta)^{1/2}(3 - 2\beta)}.$$

Under (ii) of Theorem 6, since $n^{2-2\beta}\ell^2(n) = o(\gamma_{k_n} n^{1/\alpha} \ell_0(n))$, it is easily seen that (11) still holds if X_i therein is replaced by $X_i - \bar{X}_n$.

Lemma 4. *Assume that* $\mathrm{E}(\varepsilon_i) = 0$ *and* $\varepsilon_i \in \mathcal{L}^2$. *Recall (8) for* $\sigma_{n,r}$. *Let*

$$Y_{n,r} = \sum_{t=1}^{n} \sum_{0 \le j_1 < \cdots < j_r} \prod_{s=1}^{r} a_{j_s} \varepsilon_{t-j_s}, \quad r \ge 1, \qquad Y_{n,0} = n. \tag{61}$$

For $r \in \mathbb{N}$ *with* $r(2\beta - 1) < 1$, *we have* $\mathrm{E}(Y_{n,r}^2) \sim \sigma_{n,r}^2$ *and the joint convergence*

$$\left(\frac{Y_{n,1}}{\sigma_{n,1}}, \ldots, \frac{Y_{n,r}}{\sigma_{n,r}} \right) \Rightarrow (R_{1,\beta}, \ldots, R_{r,\beta}). \tag{62}$$

Lemma 4 can be proved by using the *same* argument as that of Lemma 5 in Surgailis (1982). A careful check of the proof of his Lemma 5 suggests that the moment condition $\varepsilon_i \in \mathcal{L}^2$ suffices and the joint convergence (62) holds. We omit the details of the derivation.

Acknowledgements

We are grateful to two anonymous referees for their helpful comments.

References

Anderson, T. W. (1971). *The Statistical Analysis of Time Series*. John Wiley, New York.

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

Avram, F. and Taqqu, M. S. (1987). Noncentral limit theorems and Appell polynomials. *Ann. Prob.* **15,** 767–775.

Bartlett, M. S. (1946). On the theoretical specification and sampling properties of autocorrelated time-series. *J. R. Statist. Soc.* **8,** 27–41.

Berlinet, A. and Francq, C. (1999). Estimation of the asymptotic behavior of sample autocovariances and empirical autocorrelations of multivalued processes. *Canad. J. Statist.* **27,** 525–546.

Bingham, N. H., Goldie, C. M. and Teugels, J. L. (1989). *Regular Variation.* Cambridge University Press.

Brockwell, P. J. and Davis, R. A. (1991). *Time Series: Theory and Methods*, 2nd edn. Springer, New York.

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

Chung, C. F. (2002). Sample means, sample autocovariances, and linear regression of stationary multivariate long memory processes. *Econometric Theory* **18,** 51–78.

Csörgő, S. and Mielniczuk, J. (2000). The smoothing dichotomy in random-design regression with long-memory errors based on moving averages. *Statistica Sinica* **10,** 771–787.

Dai, W. (2004). Asymptotics of the sample mean and sample covariance of long-range-dependent series. In *Stochastic Methods and Their Applications* (J. Appl. Prob. Spec. Vol. **41A**), eds J. Gani and E. Seneta, Applied Probability Trust, Sheffield, pp. 383–392.

Dehling, H. and Taqqu, M. (1991). Bivariate symmetric statistics of long-range dependent observations. *J. Statist. Planning Infer.* **28,** 153–165.

Giraitis, L. and Taqqu, M. S. (1999). Convergence of normalized quadratic forms. *J. Statist. Planning Infer.* **80,** 15–35.

Hall, P. and Heyde, C. C. (1980). *Martingale Limit Theorem and Its Application*. Academic Press, New York.

Hannan, E. J. (1970). *Multiple Time Series*. John Wiley, New York.

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

Harris, D., McCabe, B. and Leybourne, S. (2003). Some limit theory for autocovariances whose order depends on sample size. *Econometric Theory* **19,** 829–864.

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

Horváth, L. and Kokoszka, P. (2008). Sample autocovariances of long-memory time series. *Bernoulli* **14,** 405–418.

Hosking, J. R. M. (1996). Asymptotic distributions of the sample mean, autocovariances, and autocorrelations of long-memory time series. *J. Econometrics* **73,** 261–284.

Hsieh, M.-C., Hurvich, C. M. and Soulier, P. (2007). Asymptotics for duration-driven long range dependent processes. *J. Econometrics* **141,** 913–949.

Keenan, D. M. (1997). A central limit theorem for $m(n)$ autocovariances. *J. Time Ser. Anal.* **18,** 61–78.

Major, P. (1981). *Multiple Wiener-Itô Integrals* (Lecture Notes Math. **849**). Springer, New York.

Mikosch, T., Resnick, S., Rootzen, H. and Stegeman, A. (2002). Is network traffic approximated by stable Lévy motion or fractional Brownian motion? *Ann. Appl. Prob.* **12,** 23–68.

PHILLIPS, P. C. B. AND SOLO, V. (1992). Asymptotics for linear processes. *Ann. Statist.* **20,** 971–1001.

PORAT, B. (1987). Some asymptotic properties of the sample covariances of Gaussian autoregressive moving average processes. *J. Time Ser. Anal.* **8,** 205–220.

ROSENBLATT, M. (1979). Some limit theorems for partial sums of quadratic forms in stationary Gaussian variables. *Z. Wahrscheinlichkeitsth.* **49,** 125–132.

SLY, A. AND HEYDE, C. (2008). Nonstandard limit theorem for infinite variance functionals. *Ann. Prob.* **36,** 796–805.

SURGAILIS, D. (1982). Zones of attraction of self-similar multiple integrals. *Lithuanian Math. J.* **22,** 327–340.

SURGAILIS, D. (2004). Stable limits of sums of bounded functions of long-memory moving averages with finite variance. *Bernoulli* **10,** 327–355.

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

WU, W. B. (2007). Strong invariance principles for dependent random variables. *Ann. Prob.* **35,** 2294–2320.

WU, W. B. (2009). An asymptotic theory for sample covariances of Bernoulli shifts. *Stoch. Process. Appl.* **119,** 453–467.

WU, W. B. AND MIN, W. (2005). On linear processes with dependent innovations. *Stoch. Process. Appl.* **115,** 939–958.

WU, W. B. AND WOODROOFE, M. (2004). Martingale approximations for sums of stationary processes. *Ann. Prob.* **32,** 1674–1690

YAJIMA, Y. (1993). Asymptotic properties of estimates in incorrect ARMA models for long-memory time series. In *New Directions in Time Series Analysis. Part II* (IMA Vol. Math. Appl. **46**), eds D. Brillinger *et al*., Springer, New York, pp. 375–382.