

**MEAN AND AUTOCOVARANCE FUNCTION ESTIMATION  
NEAR THE BOUNDARY OF STATIONARITY**

**By**

**Liudas Giraitis and Peter C.B. Phillips**

**COWLES FOUNDATION PAPER NO. 1363**



**COWLES FOUNDATION FOR RESEARCH IN ECONOMICS  
YALE UNIVERSITY  
Box 208281  
New Haven, Connecticut 06520-8281**

**2012**

**<http://cowles.econ.yale.edu/>**



# Mean and autocovariance function estimation near the boundary of stationarity<sup>☆</sup>

Liudas Giraitis<sup>a,\*</sup>, Peter C.B. Phillips<sup>b,c,d,e</sup>

<sup>a</sup> Queen Mary College, University of London, United Kingdom

<sup>b</sup> Yale University, United States

<sup>c</sup> University of Auckland, New Zealand

<sup>d</sup> University of Southampton, United Kingdom

<sup>e</sup> Singapore Management University, Singapore

## ARTICLE INFO

### Article history:

Available online 6 February 2012

### JEL classification:

C22

### Keywords:

Asymptotic normality  
Integrated periodogram  
Linear process  
Local to unity  
Localizing coefficient

## ABSTRACT

We analyze the applicability of standard normal asymptotic theory for linear process models near the boundary of stationarity. Limit results are given for estimation of the mean, autocovariance and autocorrelation functions within the broad region of stationarity that includes near boundary cases which vary with the sample size. The rate of consistency and the validity of the normal asymptotic approximation for the corresponding estimators is determined both by the sample size  $n$  and a parameter measuring the proximity of the model to the unit root boundary.

© 2012 Elsevier B.V. All rights reserved.

## 1. Introduction

The idea of developing asymptotics in near unit root situations is due at various levels of generality to Bobkoski (1983), Cavanagh (1985), Phillips (1987) and Chan and Wei (1987). These studies consider models in which the dominant autoregressive root is local to unity in the specific sense of  $O(n^{-1})$  departures from unity, thereby making the value of the root sample size dependent. The work has proved useful in studying near integrated processes, in establishing the local asymptotic properties of tests, and in the construction of confidence intervals.

Recent work has shown that it is also useful to provide a broader characterization of the locality of unity, the region of stationarity and the explosive region. In particular, the concept of moderate deviations from unity was suggested and pursued by Phillips and Magdalinos (2007a) and Giraitis and Phillips (2006), which leads to certain new possibilities such as mildly explosive behavior and gives rise to a new limit theory. This broader approach to modeling the region around unity conceptualizes the important practical notion that in finite samples a unit root may be treated as an

interval around unity, whose size is determined by the sample length  $n$  and measured according to units of  $1/n$ . Outside such intervals we have regions that involve certain classifiable types of stationary and explosive behavior, now measured in units of more general functions of  $1/n$ .

The idea is well illustrated in the simple AR(1) model

$$X_t = \rho X_{t-1} + \varepsilon_t, \quad t = 1, \dots, n \quad (1.1)$$

where  $\varepsilon_t$  is i.i.d.  $(0, 1)$  noise and  $X_0$  is some fixed or  $O_p(1)$  random initialization. In this model, the unit root  $\rho = 1$  is conventionally taken to prescribe the boundary case between stationarity and explosive behavior. Accordingly, a model with  $|\rho| < 1$  is stable or stationary, whereas a model with  $|\rho| > 1$  is (non-stationary) explosive. However, from both a practical and theoretical standpoint it has become increasingly clear that in finite samples of data a unit root is effectively an interval of the form

$$\rho \in [1 - a_n, 1 + a_n], \quad a_n = o(1/n),$$

which shrinks to the singular point at unity as  $n \rightarrow \infty$ . Within such intervals the limit theory and statistical tests that rely on that theory cannot distinguish different values of  $\rho$ .

Broadening the interval to include roots that are local to unity in the sense that  $1 - \rho = c/n$ , for some constant  $c$ , gives rise to the class of near integrated processes (Phillips, 1987) with  $\rho$  taking values in the region

$$\rho \in [1 - a_n, 1 + a_n], \quad a_n \sim c/n.$$

<sup>☆</sup> The authors thank the referees for useful remarks and suggestions. Giraitis gratefully acknowledges the financial support from ESRC grant RES062230790. Phillips thanks the NSF for partial research support under Grant No. SES 06-47086.

\* Corresponding author. Tel.: +44 619 699 6808; fax: +44 619 699 6850.

E-mail address: [l.giraitis@qmul.ac.uk](mailto:l.giraitis@qmul.ac.uk) (L. Giraitis).

This class is particularly useful in studying asymptotic local power functions of unit root tests and in constructing confidence intervals for  $\rho$  that allow for limit processes within the diffusion class corresponding to the limits of  $n^{-1/2}X_{[n]}$  for various values of  $c$ .

Based on this classification of unit roots and roots local to unity, the region of stationarity may be described by intervals of the type

$$\rho \in [-1 + a_n, 1 - a_n], \quad a_n n \rightarrow \infty.$$

These intervals of stationarity include moderate deviations from unity of the form  $\rho = 1 - c/k_n$  and  $\rho = -1 + c/k_n$  where  $k_n = o(n)$  and  $c > 0$ , as considered in Phillips and Magdalinos (2007a,b). Likewise the region of explosive behavior may be characterized as

$$\rho \in (-\infty, -1 - a_n] \cup [1 + a_n, \infty), \quad a_n n \rightarrow \infty.$$

In samples of size  $n$  we therefore have the following categories:

- (i) the *unit root* region, described by pairs  $(n, \rho)$  for which  $n(1 - \rho) = o(1)$  is very small;
- (ii) the *near unit root* region, described by pairs  $(n, \rho)$  for which  $n(1 - \rho) = O(1)$  may take moderate values;
- (iii) the *region of stationarity*, described by pairs  $(n, \rho)$  for which  $n(1 - \rho) \rightarrow \infty$  takes large values.

In each of these cases we may consider  $\rho$  (and hence  $v = 1 - \rho$ ) to be functionally dependent on  $n$ , as has often been done in the previous literature, thereby making the process  $X_t$  in (1.1) an array. However, one may consider the formulation from a more general perspective in which the parameter  $\rho$  is confined to an interval that depends on  $n$  and certain limit results may be established that remain valid within the entire region as  $n \rightarrow \infty$ . This set-specific formulation will be understood throughout the paper and it will be made more explicit as it is needed for the limit theory.

The region of stationarity and unit root region are separated by a local to unity region in which the least squares estimator  $\hat{\rho}_n$  of  $\rho$  in (1.1) has a non-Gaussian limit distribution. The size of the stationarity region is determined by the sample size  $n$  and  $\rho$ , and when  $n(1 - \rho)$  is large,  $\hat{\rho}_n$  has the same asymptotic properties as in the (fixed  $\rho$ ) stationary case. That is,

$$\sqrt{\frac{n}{1 - \rho^2}}(\hat{\rho}_n - \rho) \rightarrow_d N(0, 1), \tag{1.2}$$

as shown in Phillips and Magdalinos (2007a,b) and Giraitis and Phillips (2006). The convergence rate behaves as  $\{n/(1 - \rho^2)\}^{1/2} \sim \{n/2(1 - \rho)\}^{1/2}$  when  $1 - \rho$  is small. As the sample size  $n$  increases, the stationarity region approaches the boundaries of the interval  $(-1, 1)$ . Further, the convergence rate  $\{n/2(1 - \rho)\}^{1/2}$  is determined by both  $n$  and  $\rho$  and may increase from  $\sqrt{n}$  towards the unit root rate  $n$  for small  $1 - \rho$ . Thus, (1.2) is a set valued result which holds for all  $\rho$  in a region whose width ultimately depends on the sample size  $n$ .

It follows from (1.2) that standard asymptotic estimation and inferential theory applies over the whole region of  $\rho$  for which (1.2) holds. Similarly, in more general autoregressions than (1.1) and linear regressions where moderate deviations from a unit root occur, asymptotic normality will prevail although the rate of convergence may increase or slow down depending on the value of  $\rho$  and bias effects may emerge because of endogeneity in the regressors (Phillips and Magdalinos, 2007b; Magdalinos and Phillips, 2008).

The present paper seeks to explore generalizations of (1.2) for sample mean, autocovariance and autocorrelation functions near the boundary of stationarity and under a wider class of models that allow for linear process errors. Consistency and limit distribution results are given, as well as conditions for the consistent estimation of the parameter  $v = 1 - \rho$  which measures nearness to the unit root boundary. The practical implication of these set-based

limit results is that the limit theory given here is applicable within certain well defined regions for  $\rho$ , thereby extending the range of conventional theory to a parametric zone that is close to unity as  $n \rightarrow \infty$ . These regions clarify the extent to which conventional tests and confidence intervals are applicable. For instance, in the context of estimating the mean  $\mu = EX_t$ , the usual normal approximation for the sample mean  $\bar{X}$  is applicable provided the parameter  $\rho = 1 - v$  is such that  $nv$  is large. The quantity  $nv$  becomes the effective sample size in validating the normal approximation. The actual rate of convergence to the limiting normal depends on the value of  $v$  (see (2.11)) and therefore varies over an interval, as shown in Theorem 2.1 and (2.12). As we will show, related results hold in the case of autocorrelation estimation.

The paper is organized as follows. Section 2 considers a general class of linear process models, where allowance is made for the presence of roots that deviate moderately from unity. Our main results focus on the sample mean, sample correlation and sample autocovariance function and we establish the rate of consistency and the validity of normal approximations for these sample functions. Section 3 contains asymptotic theory for integrated periodograms (and hence quadratic forms) where the weighting function may depend on  $n$ . These results are discussed in Section 4. Appendix A contains proofs of the supporting asymptotic theory of Section 3. Proofs of the main results of Section 2 are given in Appendix B.

In addition to standard asymptotic notation, it is convenient, given sequences  $a_n, b_n \geq 0$ , to use the notation  $a_n \asymp b_n$  to signify that  $C_1 b_n \leq a_n \leq C_2 b_n$ , holds for  $n \geq 1$  and for some  $C_1, C_2 > 0$ .

## 2. Main results

### 2.1. Model

We consider the model

$$(1 - \rho L)X_t = Y_t, \quad Y_t = \sum_{j=0}^{\infty} b_j \varepsilon_{t-j}, \quad t = 1, 2, \dots, n, \tag{2.1}$$

where  $0 < \rho < 1$  (the closeness of  $\rho$  to unity is later made explicit and depends on  $n$ ) and  $Y_t$  is a linear MA( $\infty$ ) process with coefficients  $b_j$ , where  $(\varepsilon_t)$  is a sequence of i.i.d. random variables with

$$E\varepsilon_t = 0, \quad E\varepsilon_t^2 = 1 \tag{2.2}$$

and  $L$  is the back-shift operator. Our attention will focus on the impact of the closeness of  $\rho$  to 1 (i.e., the smallness of  $v = 1 - \rho$ ) on the validity of the asymptotic normal approximations for the distributions of the sample mean, sample autocovariance and sample autocorrelation.

The spectral density function  $f(\lambda), |\lambda| \leq \pi$ , of  $\{X_t\}$  can be written as

$$f(\lambda) = (2\pi)^{-1} f^*(\lambda) g(\lambda), \quad |\lambda| \leq \pi \tag{2.3}$$

where

$$f^*(\lambda) = |1 - \rho e^{i\lambda}|^{-2} = \frac{1}{v^2 + 2\rho(1 - \cos(\lambda))},$$

$$g(\lambda) = \left| \sum_{s=0}^{\infty} b_s e^{-i\lambda s} \right|^2$$

and  $|\rho| < 1$ . Then

$$f(0) = (2\pi)^{-1} v^{-2} g_0, \quad g_0 = g(0), \tag{2.4}$$

$$f^*(\lambda) \leq (v^2 + \rho\lambda^2/3)^{-1}, \quad |\lambda| \leq \pi, \quad |\rho| < 1, \tag{2.5}$$

using  $2(1 - \cos(\lambda)) \geq \lambda^2/3$  for  $|\lambda| \leq \pi$ . To check this bound,

use the fact that  $2(1 - \cos(\lambda)) = 4 \sin^2(\lambda/2) \geq 4\left(\frac{2}{\pi} \frac{\lambda}{2}\right)^2 = (4/\pi^2)\lambda^2 > (1/3)\lambda^2$  over  $[-\pi, \pi]$ .

We shall assume that  $g_0 > 0$ , and

$$\sum_{s=j}^{\infty} |b_s| \leq Cj^{-1-\alpha}, \quad j \geq 1 \tag{2.6}$$

for some  $\alpha > 2$ . Then, since  $g$  is an even function,

$$|g(\lambda) - g(0)| \leq C\lambda^2, \quad |\lambda| \leq \pi. \tag{2.7}$$

This summability condition (2.6) on the coefficients is easy to interpret and implies the smoothness condition (2.7).

It is natural to raise the question of how the closeness of the parameter  $\rho$  to 1 impacts the validity of the usual normal approximation of the distribution of the sample mean and second moments. Moreover, if  $\rho$  is close to one and may depend on the sample size  $n$  as discussed in the introduction, it is of interest to determine the set of pairs  $(n, \rho)$  for which the asymptotic theory corresponding to a stationary model with fixed  $\rho$  continues to apply.

We also examine the effect of the closeness of  $\rho$  to 1 on the estimation error, the rate of convergence and the length of confidence intervals.

### 2.2. Estimation of the mean

Define the sample mean:

$$\bar{X} = \frac{1}{n} \sum_{t=1}^n X_t.$$

It is well known that for any fixed  $\rho$  with  $|\rho| < 1$  as  $n \rightarrow \infty$ ,

$$\sqrt{\frac{n}{2\pi f(0)}}(\bar{X} - \mu) \rightarrow_d N(0, 1) \tag{2.8}$$

where  $\mu = E[X_t] = 0$  in case of (2.1). Since  $f(0) = (2\pi)^{-1}v^{-2}g_0$ , this implies

$$\sqrt{\frac{nv^2}{g_0}}(\bar{X} - \mu) \rightarrow_d N(0, 1). \tag{2.9}$$

On the other hand, the convergence (2.8)–(2.9) fails to extend smoothly for a unit root model, with  $\rho = 1$ , nor does the model (2.1) itself exist, unless suitable assumptions are made concerning the initialization  $X_0$  to ensure that it is well defined.

The critical question we address is under which restrictions on  $\rho$  and  $n$  does the approximation implied by the limit theory (2.8)–(2.9) continue to hold? We shall show that, for given  $(\rho, n)$ , the normal approximation (2.9) holds if  $nv$  is large. As discussed earlier, we allow for an array formulation of the model in which  $\rho = \rho_n$  and  $v = v_n$  may change with  $n$ .

**Theorem 2.1.** Assume that  $\{X_t\}$  follows the model (2.1), satisfying (2.6) and

$$v_n n \rightarrow \infty, \quad \text{as } n \rightarrow \infty. \tag{2.10}$$

Then the convergence

$$\sqrt{\frac{nv_n^2}{g_0}}(\bar{X} - \mu) \rightarrow_d N(0, 1) \tag{2.11}$$

holds.

The proof of Theorem 2.1 is given in Appendix B.

We conclude that if the parameter  $\rho = 1 - v$  and sample size  $n$  are such that  $nv$  is large, then the normal approximation (2.9) is applicable. The rate of convergence of the normal approximation (2.11) depends on the value of  $v$  and varies in the interval

$$n^{-1/2} \ll \sqrt{nv^2} \leq \sqrt{n}. \tag{2.12}$$

The convergence (2.9) shows that the rate  $\sqrt{nv^2}$  does not exceed  $\sqrt{n}$ . It becomes slow when  $v$  is close to  $n^{-1/2}$  and even tends to 0, when  $v$  approaches  $n^{-1}$ . The value of  $v$  has a strong impact on the length of confidence intervals for  $\mu$ , and estimation of  $\mu$  dramatically worsens in quality as the unit root model is approached. The sample mean  $\bar{X}$  is a consistent estimator of  $\mu$  only if  $v \gg n^{-1/2}$ , and  $\mu$  cannot be consistently estimated when  $n^{-1} \ll v \ll n^{-1/2}$ , although the normal approximation (2.11) with  $\mu = 0$  still holds. Observe, that the lower bound  $n^{-1/2}$  of the rate (2.9) is in line with results in the unit root case  $\rho = 1$ , for which under the initial condition  $X_0 = 0$ , we have  $X_t = \sum_{j=1}^t Y_j$  and

$$n^{-1/2}\bar{X} = n^{-3/2} \sum_{k=1}^n \sum_{j=1}^k Y_j \rightarrow_d \omega_Y \int_0^1 W(u)du,$$

where  $W_u$  is the standard Wiener process and  $\omega_Y^2$  is the long run variance of a stationary sequence  $Y_j$ .

This example demonstrates that the closeness of the model to unit root non-stationarity not only affects the properties of semiparametric estimation but can also have a strong impact on the quality of simple parametric estimation such as the sample mean.

### 2.3. Autocovariance and autocorrelation function estimation

We now consider estimation of the autocovariances

$$\gamma_k = \text{Cov}(X_k, X_0) = \int_{-\pi}^{\pi} \cos(\lambda k) f(\lambda) d\lambda, \quad k \geq 0$$

and the autocorrelation function  $\rho_k = \gamma_k/\gamma_0$ ,  $k = 0, 1, 2, \dots$  using the sample analogues

$$\hat{\gamma}_k = n^{-1} \sum_{t=1}^{n-k} X_{t+k} X_t, \quad \hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0}, \quad k \geq 0.$$

The next lemma describes the asymptotic behavior of  $\gamma_k$  and  $\rho_k$  as  $\rho \rightarrow 1$ . Set

$$\Gamma_k = (2\pi)^{-1} \int_{-\pi}^{\pi} \frac{1 - \cos(k\lambda)}{2(1 - \cos(\lambda))} g(\lambda) d\lambda, \quad k = 1, 2, \dots$$

and

$$\Gamma_0 = (2\pi)^{-1} \int_{-\pi}^{\pi} \frac{g(\lambda) - g_0}{2(1 - \cos(\lambda))} d\lambda.$$

The asymptotic distributions of the sample mean, autocovariances, and autocorrelations for short (and long) memory time series were studied in Hosking (1996). We focus here on stationary short memory time series which approach the unit root region. First, we discuss some asymptotic properties of  $\gamma_k$  and  $\rho_k$  as  $v \rightarrow 0$ .

**Lemma 2.1.** For fixed  $k = 0, 1, 2, \dots$ , as  $v \rightarrow 0$ ,

$$\gamma_0 = \frac{g_0}{2v} + \frac{g_0}{4} + \Gamma_0 + o(1), \tag{2.13}$$

$$\begin{aligned} \gamma_k &= \gamma_0 - \Gamma_k + o(1) \\ &= \frac{g_0}{2v} + \frac{g_0}{4} + \Gamma_0 - \Gamma_k + o(1), \quad k = 1, 2, \dots \end{aligned} \tag{2.14}$$

and

$$\rho_k = 1 - 2v\Gamma_k + o(v), \quad k = 1, 2, \dots \tag{2.15}$$

Our next theorem deals with asymptotic properties of the estimators  $\hat{\gamma}_k$  and  $\hat{\rho}_k$ .

**Theorem 2.2.** Assume that  $(X_1, \dots, X_n)$  is a sample generated by (2.1) which satisfies (2.6) with  $\rho = \rho_n$  and where  $v_n = 1 - \rho_n > 0$  has property (2.10).

(i) If  $E\varepsilon_t^4 < \infty$ , then

$$E|\hat{\gamma}_k - \gamma_k| \leq C \frac{1}{\sqrt{nv_n^3}}, \quad \hat{\gamma}_k = \gamma_k \left( 1 + O_p \left( \frac{1}{\sqrt{nv_n}} \right) \right), \quad (2.16)$$

where  $C$  does not depend on  $n$  and  $v_n$ .

(ii) If  $E\varepsilon_t^{2+\delta} < \infty$ , for some  $\delta > 0$ , and  $v_n \rightarrow 0$ , then

$$\sqrt{\frac{2nv_n^3}{g_0^2}}(\hat{\gamma}_k - \gamma_k) \rightarrow_d N(0, 1). \quad (2.17)$$

(iii) If  $E\varepsilon_t^2 < \infty$  then

$$|\hat{\rho}_k - \rho_k| = O_p \left( \frac{1}{nv_n} + \sqrt{\frac{v_n}{n}} \right). \quad (2.18)$$

Moreover, if

$$nv_n^3 \rightarrow \infty, \quad v_n \rightarrow 0, \quad n \rightarrow \infty \quad (2.19)$$

then

$$\begin{aligned} \sqrt{\frac{nv_n}{2(1-\rho_k)^2}}(\hat{\rho}_k - \rho_k) &\rightarrow_d N(0, 1), \\ \sqrt{\frac{nv_n}{2(1-\rho_k)^2}} &\sim \frac{1}{2\Gamma_k} \sqrt{\frac{n}{2v_n}}. \end{aligned} \quad (2.20)$$

The following theorem considers estimation of the quantity  $\sqrt{v}$  in the neighborhood of 0. Denote the periodogram by  $I_n(\lambda) = (2\pi n)^{-1} |\sum_{j=1}^n e^{ij\lambda} X_j|^2$ , and define

$$\sqrt{\hat{v}_n} = \frac{\int_{-\pi}^{\pi} |\lambda/2|^{1/2} I_n(\lambda) d\lambda}{\int_{-\pi}^{\pi} I_n(\lambda) d\lambda}. \quad (2.21)$$

**Theorem 2.3.** Assume that  $(X_1, \dots, X_n)$  is a sample generated from (2.1) which satisfies (2.6) with  $\rho = \rho_n$  and where  $v_n = 1 - \rho_n > 0$  has property (2.10). If  $E\varepsilon_t^4 < \infty$ , then

$$\sqrt{\hat{v}_n} = \sqrt{v_n} + O_p \left( v_n + \frac{1}{nv_n} + \frac{1}{\sqrt{n}} \right). \quad (2.22)$$

The proofs of Lemma 2.1 and Theorems 2.2–2.3 are given in Appendix B.

**Remarks.**

(i) Estimation of  $\hat{\gamma}_k$  and  $\hat{\rho}_k$  is based on approximation of these statistics by quadratic forms of the form  $\sum_{t,s=1}^n b_n(t-s)\varepsilon_t\varepsilon_s$  with suitable weights  $b_n(t-s)$ . In case of  $\hat{\rho}_k$ , the diagonal elements  $b_n(t-s)$  become 0, whereas in case of  $\hat{\gamma}_k$ , the contribution of the diagonal  $\sum_{t=s=1}^n b_n(t-s)\varepsilon_t^2$ , as  $v_n \rightarrow 0$ , is asymptotically negligible. This representation leads to the requirement of finite  $2 + \delta$  moments of  $\varepsilon_t$  in (ii), and second moments in (iii). In the case where  $v_n$  is fixed, the convergence (2.17) requires finite fourth moments of  $\varepsilon_t$ .

(ii) It follows from (2.16) that  $\hat{\gamma}_k$  is a consistent estimate of  $\gamma_k$ . The CLT (2.17) is valid with the convergence rate  $\sqrt{nv_n^3}$  which depends on the value of  $v_n$  and varies in the interval

$$n^{-1} \ll \sqrt{nv_n^3} \ll \sqrt{n}.$$

(iii) As  $v_n$  decreases, confidence intervals for  $\gamma_k$  will increase. When  $nv_n^3 \rightarrow 0$  then (2.17) can be written in the form

$$\begin{aligned} \hat{\gamma}_k &\sim \gamma_k \left( 1 + \frac{g_0}{\gamma_k \sqrt{2nv_n^3}} Z \right) \sim \gamma_k \left( 1 + \sqrt{\frac{2}{nv_n}} Z \right), \\ Z &\sim N(0, 1). \end{aligned}$$

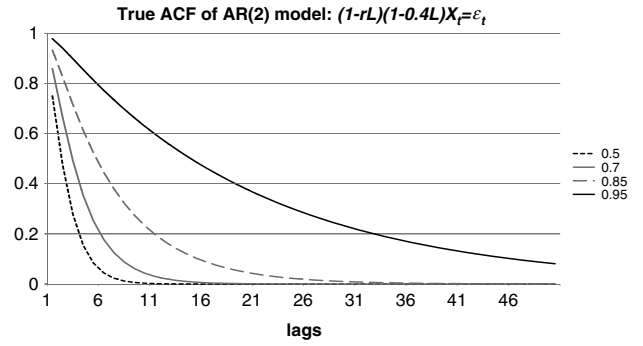


Fig. 1. ACF  $\rho_k$  of AR(2) model with  $r = 0.5, 0.7, 0.85, 0.95, n = 125$ .

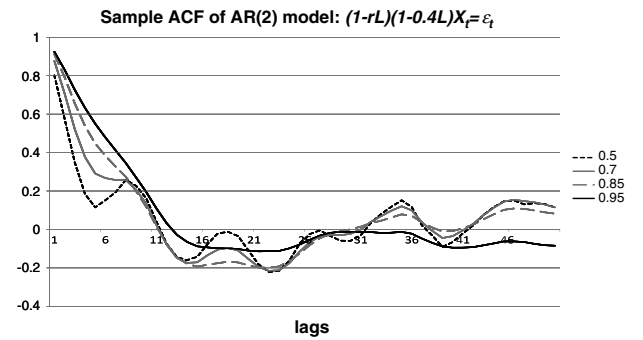


Fig. 2. Example of realizations of Sample ACF  $\hat{\rho}_k$  of AR(2) model with  $r = 0.5, 0.7, 0.85, 0.95, n = 125$ .

(iv) Theorem 2.2 shows that the sample autocorrelation  $\hat{\rho}_k$  is a consistent estimator of  $\rho_k$  as long as  $nv \rightarrow \infty$ , and  $E\varepsilon_t^2 < \infty$ . The proof indicates that  $\hat{\rho}_k - \rho_k$  can be decomposed into a bias term of order  $O_p((nv_n)^{-1})$  and the stochastic CLT term  $\sqrt{\frac{v_n}{n}} N(0, 1)$  which dominates the bias under the condition  $nv_n^3 \rightarrow \infty$ .

(v) To apply these results in samples of size  $n$  we set  $v_n = v = 1 - \rho$  where  $\rho$  is the parameter of the data generating process. The parameter  $v$  can be consistently estimated as shown in Theorem 2.3.

The proof of Theorem 2.2 is based on central limit theory for certain quadratic forms and this theory is developed in the following section.

Fig. 1 shows the ACF  $\rho_k$  of the AR(2) model  $(1-rL)(1-0.4L)X_t = \varepsilon_t$  for the parameter values  $r = 0.5, 0.7, 0.85$  and  $0.95$ . Fig. 2 shows a realization of the sample ACF,  $\hat{\rho}_k$ , computed from a sample of  $n = 125$  observations. Figs. 3 and 4 show the bias  $\hat{\rho}_k - \rho_k$  and the relative bias  $(\hat{\rho}_k - \rho_k)/\rho_k$  corresponding to these realizations. The figures confirm the theory based on (2.20) that the rate of convergence  $\sqrt{n/v}$  of the sample ACF in the near unit root region improves when  $v \rightarrow 0$ , and  $nv^3$  remains large. That condition is not well satisfied when  $r = 0.95$ , partly explaining the large bias in this case.<sup>1</sup>

Figs. 5–6 indicate the adequacy of the standard normal approximation (2.20) to the probability density function of the standardized sample ACFs  $\hat{t}_n(k) = \sqrt{\frac{nv}{2(1-\rho_k)^2}}(\hat{\rho}_k - \rho_k)$  for lags  $k = 5, 25, 45$  in the same AR(2) model and with  $n = 2000$ . The probability density of  $\hat{t}_n(5)$  was estimated using a kernel estimator based on 50,000 replications. The figures indicate that the density is generally well fitted by the standard normal for  $r = 0.8, 0.95$ ,

<sup>1</sup> We thank Violetta Dalla for preparing Figs. 1–4.

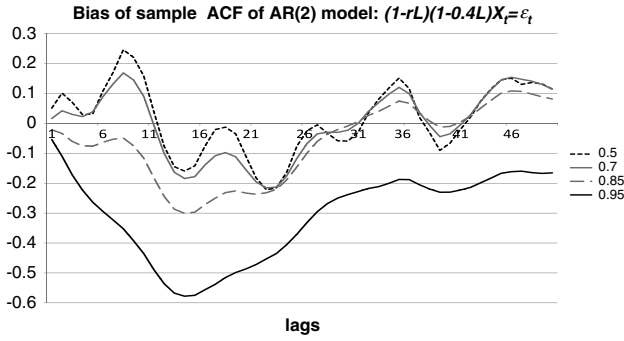


Fig. 3. Bias of Sample ACF of AR(2) model with  $r = 0.5, 0.7, 0.85, 0.95, n = 125$ .

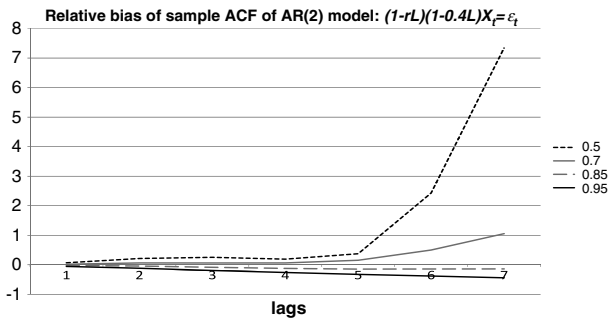


Fig. 4. Relative bias  $(\hat{\rho}_k - \rho_k)/\rho_k$  of Sample ACF of AR(2) model with  $r = 0.5, 0.7, 0.85, 0.95, n = 125$ .

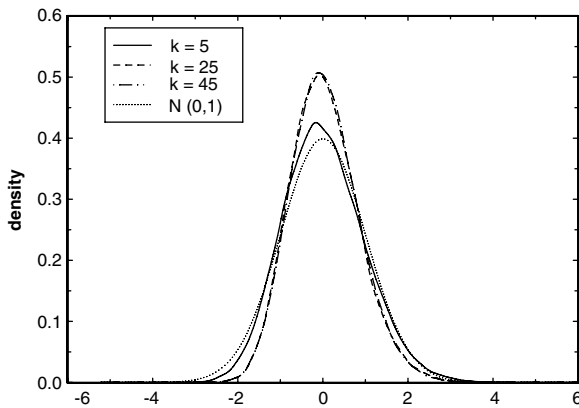


Fig. 5. Densities of  $\hat{t}_n(k) : k = 5, 25, 45$  versus the standard normal for  $r = 0.8$  and  $n = 2000$ .

corresponding to the near unit root case with  $v = 0.2$  and  $0.05$ , respectively, although we note that the departure from the standard normal is greater for larger lag values.

### 3. Asymptotic theory for quadratic forms

This section provides some supportive asymptotic theory for quadratic forms and their approximations by quadratic forms in i.i.d. variables. The methodology developed in [Bhansali et al. \(2007\)](#) that we follow cannot be applied directly since the bounds for remainder terms in case of a near unit root parameter  $\rho_n$  have to be derived uniformly in  $\rho = \rho_n$ , corresponding to the closeness of the model to a unit root model—see [Assumption 3.1](#). The results of this Section are applied to the sample autocovariances and autocorrelations of Section 2.

We assume that

$$X_t = \left( \sum_{s=0}^{\infty} a_s L^s \right) Y_t = \sum_{s=0}^{\infty} a_s Y_{t-s}, \quad t = 0, 1, 2, \dots \quad (3.1)$$

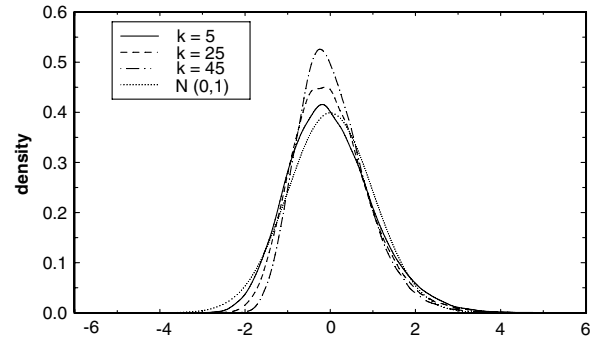


Fig. 6. Densities of  $\hat{t}_n(k) : k = 5, 25, 45$  versus the standard normal for  $r = 0.95$  and  $n = 2000$ .

is a linear process where

$$Y_t = \left( \sum_{s=0}^{\infty} b_s L^s \right) \varepsilon_t = \sum_{s=0}^{\infty} b_s \varepsilon_{t-s},$$

$(\varepsilon_t)$  is a sequence of i.i.d. random variables with  $E\varepsilon_t = 0, E\varepsilon_t^2 = 1$ , and the real coefficients  $a_s, b_s$  are absolutely summable. We can write  $X_t$  as

$$X_t = \left( \sum_{j=0}^{\infty} a_j L^j \right) \left( \sum_{s=0}^{\infty} b_s L^s \right) \varepsilon_t = \sum_{u=0}^{\infty} \psi_u \varepsilon_{t-u}, \quad t = 0, 1, 2, \dots$$

with

$$\psi_u = \sum_{k=0}^u a_k b_{u-k}, \quad u = 0, 1, 2, \dots$$

The spectral density function  $f(\lambda), |\lambda| \leq \pi$ , of  $\{X_t\}$  can be written as

$$f(\lambda) = (2\pi)^{-1} |\Psi(\lambda)|^2, \quad \Psi(\lambda) = \Psi_a(\lambda) \Psi_b(\lambda) \quad (3.2)$$

where

$$\Psi_a(\lambda) = \sum_{s=0}^{\infty} a_s e^{-i\lambda s}, \quad \Psi_b(\lambda) = \sum_{s=0}^{\infty} b_s e^{-i\lambda s}.$$

We impose the following restrictions on  $a_s$  and  $b_s$ .

**Assumption 3.1.** (i) The coefficients  $a_j$  satisfy

$$|a_j| \leq C \rho^j, \quad j = 1, 2, 3, \dots \quad (3.3)$$

for some  $0 \leq \rho < 1$ , where  $\rho = \rho_n$  may depend on  $n$ .

(ii) The coefficients  $b_s$  are such that

$$\sum_{s=j}^{\infty} |b_s| \leq C j^{-1-\alpha}, \quad j = 1, 2, 3, \dots \quad (3.4)$$

for some  $\alpha > 1/2$ , and the  $b_s$  do not vary when  $n$  changes.

$C$  here and below denotes a generic positive constant which may change from line to line but does not depend on  $n$  and  $\rho$ . The parameter  $\rho$  in [Assumption 3.1](#) characterizes the closeness of the process  $X_t$  to a unit root process and plays the same role as the AR(1) parameter  $\rho$  in model (2.1). We keep the same notation, to relate the results to the applications of Section 2. As before, we let  $v = 1 - \rho$ .

Under [Assumption 3.1](#), the spectral density

$$f(\lambda) \leq C \left| \sum_{j=0}^{\infty} \rho^j \right|^2 \leq C v^{-2} \quad (3.6)$$

is bounded by a constant times  $v^{-2}$  which increases to  $\infty$  as  $\rho$  tends to 1. For example, the model

$$X_t = (1 - \rho L)^{-1} Y_t = \left( \sum_{j=0}^{\infty} \rho^j L^j \right) Y_t = \sum_{j=0}^{\infty} \rho^j Y_{t-j},$$

where  $0 < \rho < 1$  and  $Y_t$  is an ARMA( $p, q$ ) model that has properties (3.3) and (3.4).

In effect, we consider data that takes the form of a triangular array

$$(X_1, \dots, X_n) = (X_1^{(n)}, \dots, X_n^{(n)}), \quad n = 1, 2, \dots$$

generated by model (3.1) where, as  $n$  increases, the coefficient  $\rho = \rho_n$  in (3.3) may change with  $n$ , e.g. they may approach unity, whereas the coefficients  $b_j$  remain the same and satisfy condition (3.4) with the same  $C$  and  $\rho$  for all  $n$ .

Denote by

$$I_n(\lambda) = \frac{1}{2\pi n} \left| \sum_{j=1}^n X_j e^{i\lambda j} \right|^2, \quad I_{n,\varepsilon}(\lambda) = \frac{1}{2\pi n} \left| \sum_{j=1}^n \varepsilon_j e^{i\lambda j} \right|^2$$

the periodograms of the observed variable  $X_t$  and the noise variable  $\varepsilon_t$ . A number of useful statistics can be written in the form of functionals of the integrated periodogram

$$T_{n,X} = \int_{-\pi}^{\pi} \eta_n(\lambda) I_n(\lambda) d\lambda,$$

where  $\eta_n(\lambda)$  is a real even function. The well-known Bartlett (1955) decomposition

$$I_n(\lambda) = 2\pi f(\lambda) I_{n,\varepsilon}(\lambda) + L_n(\lambda) \tag{3.7}$$

divides the periodogram  $I_n(\lambda)$  into the weighted periodogram  $2\pi f(\lambda) I_{n,\varepsilon}(\lambda)$  of the noise and the remainder  $L_n(\lambda)$ . The expression suggests that  $T_{n,X}$  can be similarly decomposed as

$$T_{n,X} = T_{n,\varepsilon} + \text{“small term”}$$

where

$$T_{n,\varepsilon} = 2\pi \int_{-\pi}^{\pi} \eta_n(\lambda) f(\lambda) I_{n,\varepsilon}(\lambda) d\lambda,$$

is a quadratic form of the i.i.d. variables  $\varepsilon_j$ . If the remainder above is dominated by  $T_{n,\varepsilon}$ , then one can analyze  $T_{n,X}$  via  $T_{n,\varepsilon}$ . Our objective is to derive a precise upper bound for the remainder term. Then, using asymptotic theory for the quadratic forms  $T_{n,\varepsilon}$  in i.i.d. variables, we derive the asymptotic distribution of  $T_{n,X}$ . We shall assume that the functions  $\eta_n$  have the following property.

**Assumption 3.2.**  $\eta_n$  is a real even function such that

$$|\eta_n(\lambda)| \leq k_n, \quad \lambda \in [-\pi, \pi], \quad n \geq 1. \tag{3.8}$$

Thus, the functions  $\eta_n$  are bounded but their upper bound  $k_n$  might vary with  $n$ , for example,  $\eta_n$  may be a kernel function.

Let

$$h_n(\lambda) = \eta_n(\lambda) f(\lambda). \tag{3.9}$$

Then  $T_{n,\varepsilon} = 2\pi \int_{-\pi}^{\pi} h_n(\lambda) I_{n,\varepsilon}(\lambda) d\lambda$ . We shall assume that  $h_n(u)$  is periodically extended to the real line  $R$ . Set

$$\|h_n\|^2 := \int_{-\pi}^{\pi} h_n^2(\lambda) d\lambda. \tag{3.10}$$

**Theorem 3.1.** Suppose that Assumptions 3.1 and 3.2 hold and the noise  $\{\varepsilon_t\}$  has finite second moment.

Then, for  $n \geq 1$ ,

$$E|T_{n,X} - T_{n,\varepsilon}| \leq C \frac{k_n}{nv^2}, \tag{3.11}$$

and

$$T_{n,X} = \int_{-\pi}^{\pi} \eta_n(\lambda) f(\lambda) d\lambda + (T_{n,\varepsilon} - E[T_{n,\varepsilon}]) + r_n, \tag{3.12}$$

$$E|r_n| \leq C \frac{k_n}{nv^2}.$$

If  $\varepsilon_t^4 < \infty$ , then

$$E \left| T_{n,X} - \int_{-\pi}^{\pi} \eta_n(\lambda) f(\lambda) d\lambda \right| \leq C \left( \frac{k_n}{nv^2} + \sqrt{\frac{\|h_n\|^2}{n}} \right) \leq C \left( \frac{k_n}{nv^2} + \frac{k_n}{\sqrt{nv^3}} \right) \tag{3.13}$$

where  $C$  does not depend on  $n$  and  $v = 1 - \rho$ .

Theorem 3.1 provides sharp upper bounds for the remainder term which reflects the interplay of  $n$  and  $\rho$ , with no restrictions on  $\rho$  imposed. The bounds (3.11)–(3.13) hold uniformly in  $0 \leq \rho < 1$ , but they show that the remainder is negligible, only if  $nv^2$  is large. The constants  $k_n$  play a secondary role. If the functions  $\eta_n(\lambda)$  do not depend on  $n$ , we can set  $k_n = 1$ . The proof of Theorem 3.1 is given in the Appendix.

Next we derive the CLT for the term  $T_{n,\varepsilon} - E[T_{n,\varepsilon}]$  in (3.12) and describe conditions under which it dominates the remainder  $r_n$ .

First, to evaluate  $\text{Var}(T_{n,\varepsilon})$ , we introduce the matrix  $E_n = (e_n(t - k))_{t,k=1,\dots,n}$  with the entries

$$e_n(t) = 2\pi \int_{-\pi}^{\pi} h_n(\lambda) e^{i\lambda t} d\lambda, \tag{3.14}$$

and denote by  $\|E_n\| = (\sum_{t,k=1}^n e_n^2(t - k))^{1/2}$  its Euclidean norm. Observe that

$$(2\pi n)^2 \text{Var}(T_{n,\varepsilon}) = 2 \sum_{t,k=1:t \neq k}^n e_n^2(t - k) + \text{Var}(\varepsilon_0^2) e_n^2(0) n. \tag{3.15}$$

Then

$$\text{Var}(T_{n,\varepsilon}) \asymp \frac{1}{n^2} \|E_n\|^2. \tag{3.16}$$

If  $e_n^2(0) = 0$ , then  $E_n$  has zero diagonal, and

$$\text{Var}(T_{n,\varepsilon}) = \frac{2}{(2\pi n)^2} \|E_n\|^2. \tag{3.17}$$

To derive the asymptotic behavior of  $\|E_n\|^2$  we introduce the following  $L_2$  continuity assumption.

**Assumption 3.3.** For any  $K > 0$ ,

$$\sup_{|u| \leq K/n} \int_{-\pi}^{\pi} |h_n(u - x) - h_n(x)|^2 dx / \|h_n\|^2 \rightarrow 0, \quad n \rightarrow \infty. \tag{3.18}$$

Assumption (3.18) means that, for  $u = O(n^{-1})$ , the normalized function  $h_n(u - \cdot)$  is close to  $h_n(\cdot)$  in the  $L_2$  norm. It is satisfied for example, if  $h_n \equiv h$  does not depend on  $n$  and  $\|h\| < \infty$ . In Lemma A.2 in Appendix A we show that under Assumption 3.3,

$$\|E_n\|^2 \sim (2\pi)^3 n \|h_n\|^2.$$

Lemma 3.1 provides the central limit theorem for the quadratic form  $T_{n,\varepsilon}$  in i.i.d. variables, and is a direct consequence of Theorems 4.1 and 4.2 in Bhansali et al. (2007) and Lemma A.2 below. It takes into account the fact that the upper bound  $k_n^*$  in  $|\eta_n(\lambda) f(\lambda)| \leq k_n^*$  might be smaller than the product  $Ck_n \times v^{-2}$  of the upper bounds  $|\eta_n(\lambda)| \leq k_n$  and  $f(\lambda) \leq Cv^{-2}$ .

We shall distinguish two cases, (c1) and (c2), when the CLT does not require finite fourth moment of the noise  $\varepsilon_t$ .

Case (c1):

$$E\varepsilon_t^{2+\delta} < \infty, \quad \text{and} \quad \int_{-\pi}^{\pi} h_n(\lambda) d\lambda = 0. \quad (3.19)$$

Case (c2):

$$E\varepsilon_t^{2+\delta} < \infty \quad \text{for some } \delta > 0, \quad \text{and} \quad \int_{-\pi}^{\pi} h_n(\lambda) d\lambda = o\left(\left(\int_{-\pi}^{\pi} h_n(\lambda)^2 d\lambda\right)^{1/2}\right). \quad (3.20)$$

Case (c1) corresponds to the case where  $E_n$  has zero diagonal, whereas case (c2) corresponds to the case of an asymptotically vanishing diagonal, a situation that often occurs, for instance, in the case of quadratic forms in kernel estimation.

**Lemma 3.1.** *Suppose that  $h_n$  satisfies Assumption 3.3,*

$$|h_n(\lambda)| \leq k_n^*, \quad n \geq 1,$$

and

$$\frac{k_n^*}{\sqrt{n \|h_n\|^2}} \rightarrow 0, \quad n \rightarrow \infty. \quad (3.21)$$

(i) *If  $E\varepsilon_t^4 < \infty$ , then*

$$[\text{Var}(T_{n,\varepsilon})]^{-1/2} \left( T_{n,\varepsilon} - \int_{-\pi}^{\pi} h_n(\lambda) d\lambda \right) \xrightarrow{d} N(0, 1), \quad (3.22)$$

$$\text{Var}(T_{n,\varepsilon}) \asymp \frac{\|h_n\|^2}{n}.$$

(ii) *If (c1) or (c2) hold, then*

$$\sqrt{\frac{n}{4\pi \|h_n\|^2}} \left( T_{n,\varepsilon} - \int_{-\pi}^{\pi} h_n(\lambda) d\lambda \right) \xrightarrow{d} N(0, 1). \quad (3.23)$$

Lemma 3.1 remains valid also for any sequence of real even functions  $h_n(\lambda)$  without assuming (3.9).

Applying Lemma 3.1 to the asymptotic expansion (3.12) in Theorem 3.1, we obtain the CLT for  $T_{n,X}$ . Condition (3.24) assures that the main term  $T_{n,\varepsilon} - E[T_{n,\varepsilon}]$  satisfies the CLT and dominates the remainder term  $r_n$ .

**Theorem 3.2.** *Suppose that Assumptions 3.1–3.3 are satisfied and, as  $n \rightarrow \infty$ ,*

$$\frac{k_n/v^2}{\sqrt{n \|h_n\|^2}} \rightarrow 0. \quad (3.24)$$

(i) *If  $E\varepsilon_t^4 < \infty$  then*

$$[\text{Var}(T_{n,X})]^{-1/2} \left( T_{n,X} - \int_{-\pi}^{\pi} \eta_n(\lambda) f(\lambda) d\lambda \right) \xrightarrow{d} N(0, 1) \quad (3.25)$$

and

$$\text{Var}(T_{n,X}) \sim \text{Var}(T_{n,\varepsilon}) \asymp \frac{\|h_n\|^2}{n}. \quad (3.26)$$

(ii) *If (c2) or (c3) hold, then*

$$\sqrt{\frac{n}{4\pi \|h_n\|^2}} \left( T_{n,X} - \int_{-\pi}^{\pi} \eta_n(\lambda) f(\lambda) d\lambda \right) \xrightarrow{d} N(0, 1). \quad (3.27)$$

#### 4. Discussion

The idea of approximation results similar to those in Theorem 3.1 goes back to the work of Hannan (1973) and Hannan and Heyde (1972). Classical results in the time series literature cover

the case where the function  $\eta_n(\lambda) = \eta(\lambda)$  is continuous and does not depend on  $n$ , and  $\{X_t\}$  is a stationary ARMA process. Brockwell and Davis (1991), Proposition 10.8.5, showed that

$$E|T_{n,X} - T_{n,\varepsilon}| = o(n^{-1/2}). \quad (4.1)$$

Bhansali et al. (2007) extended this type of approximation to the class of linear processes  $\{X_t\}$  allowing for both weak and strong dependence as well as antipersistence, and allowing  $\eta_n(\lambda)$  to depend on  $n$ . The bound (4.1) was improved to the sharper bound

$$E|T_{n,X} - T_{n,\varepsilon}| = O(n^{-1}).$$

In the present paper, Theorem 3.1 provides the approximating bounds

$$E|T_{n,X} - T_{n,\varepsilon}| = O((nv^2)^{-1}),$$

$$E \left| T_{n,X} - \int_{-\pi}^{\pi} \eta_n(\lambda) f(\lambda) d\lambda \right| \leq C \left( \frac{k_n}{nv^2} + \frac{k_n}{\sqrt{nv^3}} \right),$$

that hold uniformly over  $n$  and the parameter  $v = 1 - \rho$  characterizing closeness of the model to the boundary of the stationary region.

The conditions of the CLT of Theorem 3.2 are easy to check. This theorem simplifies the derivation of asymptotics for statistics which can be written in the form of functionals of the integrated periodogram  $T_{n,X}$ . For instance, these approximations can be used to derive the asymptotic distributions of sample autocovariances and spectral density estimates based on kernel methods.

#### Appendix A

**Proof of Theorem 3.1.** Set

$$v_k = \begin{cases} 2\pi \sum_{j=1-k}^{n-k} \psi_j^2, & \text{for } k \leq 0, \\ 2\pi \sum_{j=n-k+1}^n \psi_j^2, & \text{for } 1 \leq k \leq n \end{cases}$$

and

$$d_n = 2\pi \sum_{j=n+1}^{\infty} \psi_j^2, \quad n = 0, 1, 2, \dots$$

Let  $V_n = \sum_{k=-\infty}^n v_k$  and  $R = \sum_{j=0}^{\infty} |\psi_j|$ . First we prove the following technical result.

**Lemma A.1.** *Assume that  $\eta_n(\lambda)$  satisfies Assumption 3.2. Then*

$$E|T_n - T_{n,\varepsilon}| \leq Cn^{-1}k_n \left( V_n + nd_n + V_n^{1/2}R + d_0^{1/2} \sum_{k=0}^n v_k^{1/2} + n^{1/2}d_n^{1/2}R + nd_n^{1/2}d_0^{1/2} \right) \quad (A.1)$$

where  $C$  does not depend on  $n$  and  $\rho$ .

**Proof of Lemma A.1.** Define

$$d_k(\lambda) = \begin{cases} \sum_{j=1-k}^{n-k} \psi_j e^{-i\lambda j}, & \text{for } k \leq 0, \\ - \sum_{j=n-k+1}^n \psi_j e^{-i\lambda j}, & \text{for } 1 \leq k \leq n \end{cases}$$

and

$$c_n(\lambda) = \sum_{j=n+1}^{\infty} \psi_j e^{-i\lambda j}, \quad n = 0, 1, 2, \dots$$



For integers  $k$  and  $t$ , we introduce the coefficients:

$$\begin{aligned} v_n(k, t) &:= \int_{-\pi}^{\pi} e^{i(t-k)\lambda} d_k(\lambda) \overline{d_t(\lambda)} |\eta_n(\lambda)| d\lambda, \\ \beta_n(k, t) &:= \int_{-\pi}^{\pi} e^{i(t-k)\lambda} d_k(\lambda) \Psi(\lambda) \eta_n(\lambda) d\lambda, \\ \mu_n(k, t) &:= \int_{-\pi}^{\pi} e^{i(t-k)\lambda} |c_n(\lambda)|^2 |\eta_n(\lambda)| d\lambda, \\ \zeta_n(k, t) &:= \int_{-\pi}^{\pi} e^{i(t-k)\lambda} c_n(\lambda) \Psi(\lambda) \eta_n(\lambda) d\lambda. \end{aligned}$$

Observe that

$$\begin{aligned} v_k &:= \int_{-\pi}^{\pi} |d_k(\lambda)|^2 d\lambda, & d_n &:= \int_{-\pi}^{\pi} |c_n(\lambda)|^2 d\lambda, \\ d_0 &:= \int_{-\pi}^{\pi} |\Psi(\lambda)|^2 d\lambda, \\ v_k &:= C \sum_{j=1-k}^{n-k} \psi_j^2 \quad \text{for } k \leq 0, \\ v_k &:= C \sum_{j=n-k+1}^n \psi_j^2 \quad \text{for } 1 \leq k \leq n, \\ d_n &:= C \sum_{j=n+1}^{\infty} \psi_j^2. \end{aligned}$$

To derive the bound (A.1), we shall use the estimate (5.31) of Bhansali et al. (2007)

$$E|T_n - T_{n,\varepsilon}| \leq Cn^{-1}(E|Y_n| + E|V_{n,1}| + E|V_{n,2}|) \tag{A.2}$$

where it was shown that

$$\begin{aligned} E|Y_n| &\leq C \left( \sum_{k=-\infty}^n v_n(k, k) + \sum_{k=1}^n \mu_n(k, k) \right) =: C[s_{n,1} + s_{n,2}], \\ E|V_{n,1}| &\leq C \left[ \left( \sum_{k=-\infty}^n \sum_{t=1:t \neq k}^n |\beta_n(k, t)|^2 \right)^{1/2} + \sum_{k=1}^n |\beta_n(k, k)| \right] \\ &=: C[s_{n,3} + s_{n,4}], \\ E|V_{n,2}| &\leq C \left[ \left( \sum_{k=1}^n \sum_{t=1:t \neq k}^n |\zeta_n(k, t)|^2 \right)^{1/2} + \sum_{k=1}^n |\zeta_n(k, k)| \right] \\ &=: C[s_{n,5} + s_{n,6}]. \end{aligned}$$

Recall that  $|\eta_n(\lambda)| \leq k_n$ . Hence, by (A.2),

$$\begin{aligned} |s_{n,1}| &\leq Ck_n \sum_{k=-\infty}^n \int_{-\pi}^{\pi} |d_k(\lambda)|^2 d\lambda = Ck_n \sum_{k=-\infty}^n v_k = Ck_n V_n, \\ |s_{n,2}| &\leq Ck_n n \int_{-\pi}^{\pi} |c_n(\lambda)|^2 d\lambda = Ck_n n d_n, \\ |s_{n,4}| &\leq Ck_n \sum_{k=1}^n \int_{-\pi}^{\pi} |d_k(\lambda)| |\Psi(\lambda)| d\lambda, \\ &\leq Ck_n \sum_{k=1}^n \left( \int_{-\pi}^{\pi} |d_k(\lambda)|^2 d\lambda \right)^{1/2} \left( \int_{-\pi}^{\pi} |\Psi(\lambda)|^2 d\lambda \right)^{1/2} \\ &\leq Ck_n \sum_{k=1}^n v_k^{1/2} d_0^{1/2}, \\ |s_{n,6}| &\leq Ck_n n \int_{-\pi}^{\pi} |c_n(\lambda)| |\Psi(\lambda)| d\lambda \leq k_n n d_n^{1/2} d_0^{1/2}. \end{aligned}$$

The estimates (5.26)–(5.27) and (5.28)–(5.29) of Bhansali et al. (2007) imply that

$$\begin{aligned} s_{n,3} &\leq C \sum_{k=-\infty}^n \int_{-\pi}^{\pi} |d_k(\lambda) \Psi(\lambda) \eta_n(\lambda)|^2 d\lambda, \\ s_{n,5} &\leq Cn \int_{-\pi}^{\pi} |c_n(\lambda) \Psi(\lambda) \eta_n(\lambda)|^2 d\lambda. \end{aligned}$$

Since  $|\eta_n(\lambda)| \leq k_n$  and

$$|\Psi(\lambda)| \leq C \sum_{j=0}^{\infty} |\psi_j| \leq CR,$$

it follows that

$$\begin{aligned} s_{n,3} &\leq Ck_n^2 R^2 \sum_{k=-\infty}^n \int_{-\pi}^{\pi} |d_k(\lambda)|^2 d\lambda = Ck_n^2 R^2 \sum_{k=-\infty}^n v_k = Ck_n^2 R^2 V_n, \\ s_{n,5} &\leq Ck_n^2 n R^2 \int_{-\pi}^{\pi} |c_n(\lambda)|^2 d\lambda = Ck_n^2 n R^2 d_n. \end{aligned}$$

Hence

$$s_{n,3}^{1/2} \leq Ck_n V_n^{1/2} R, \quad s_{n,5}^{1/2} \leq Ck_n n^{1/2} d_n^{1/2} R.$$

The above bounds for  $s_{n,j}$ ,  $j = 1, \dots, 6$  prove (A.1).  $\square$

Now, using Assumption 3.1 we estimated quantities  $V_n$ ,  $d_n$ ,  $d_0$  and  $R$ . We have

$$\psi_j = \sum_{s=0}^j a_s b_{j-s}, \quad j \geq 0.$$

Recall that  $|a_j| \leq C\rho^j$ ,  $j = 1, 2, \dots$  and  $\sum_{j=k}^{\infty} |b_j| \leq C|k|^{-1-\alpha}$ ,  $k \geq 1$ , where  $\alpha > 1/2$ .

First we show that

$$|\psi_j| \leq C(j^{-1-\alpha} + \rho^{j/2}), \tag{A.3}$$

where  $C$  does not depend on  $\rho$  and  $j = 1, 2, 3, \dots$ . Write  $\psi_j = \psi_j^- + \psi_j^+$  where

$$\psi_j^- = (2\pi)^{-1} \sum_{t=0}^{j/2} a_t b_{j-t}, \quad \psi_j^+ = (2\pi)^{-1} \sum_{t=j/2+1}^j a_t b_{j-t}.$$

In the sum in  $\psi_j^-$  we have  $j - t \geq j/2$ . Therefore

$$|\psi_j^-| \leq C \sum_{t=0}^{j/2} |b_{j-t}| \leq C \sum_{v=j/2}^{\infty} |b_v| \leq C|j|^{-1-\alpha}, \quad j = 1, 2, \dots$$

$$|\psi_j^+| \leq C \sum_{t=j/2+1}^j \rho^t b_{j-t} \leq C\rho^{j/2} \sum_{v=0}^{\infty} |b_v| \leq C\rho^{j/2}.$$

Applying (A.3), it follows that for  $k \geq 1$ ,

$$\begin{aligned} \sum_{j=k}^{\infty} \psi_j^2 &\leq C \sum_{j=k}^{\infty} (j^{-2-2\alpha} + \rho^j) \\ &\leq C \left( k^{-1-2\alpha} + \rho^k \sum_{j=0}^{\infty} \rho^j \right) \leq C(k^{-1-2\alpha} + \rho^k v^{-1}). \end{aligned}$$

Using this bound, it follows that

$$v_k \leq C \begin{cases} |k-1|^{-1-2\alpha} + \rho^{-k} v^{-1}, & \text{for } k \leq 0, \\ (n-k+1)^{-1-2\alpha} + \rho^{n-k+1} v^{-1}, & \text{for } 1 \leq k \leq n \end{cases}$$

and

$$d_0 \leq Cv^{-1}, \quad d_n \leq C(n^{-1-2\alpha} + \rho^n v^{-1}), \quad n = 1, 2, \dots,$$

$$\int_{-\pi}^{\pi} f(\lambda) d\lambda = C \sum_{j=0}^{\infty} \psi_j^2 = Cd_0 \leq Cv^{-1}. \tag{A.4}$$

Then

$$V_n = \sum_{k=-\infty}^n v_k \leq C \sum_{k \leq 0} \left( (-k+1)^{-1-2\alpha} + \rho^{-k} v^{-1} \right) + \sum_{k=1}^n \left( (n-k+1)^{-1-2\alpha} + \rho^{n-k+1} v^{-1} \right) \leq C \left( \sum_{k=0}^{\infty} (k+1)^{-1-2\alpha} + \sum_{k=0}^{\infty} \rho^k v^{-1} \right) \leq Cv^{-2},$$

and

$$R \leq \sum_{k=0}^{\infty} |\psi_k| \leq C \sum_{k=0}^{\infty} \left( (k+1)^{-1-\alpha} + \rho^{k/2} \right) \leq Cv^{-1}.$$

Moreover, since for  $k = 1, \dots, n$

$$v_k^{1/2} \leq C \left( (n-k+1)^{-1-2\alpha} + \rho^{n-k+1} v^{-1} \right)^{1/2} \leq C \left( (n-k+1)^{-1/2-\alpha} + \rho^{(n-k+1)/2} v^{-1/2} \right)$$

and  $\alpha > 1/2$ , then

$$\sum_{k=1}^n v_k^{1/2} \leq C \sum_{k=1}^n \left( (n-k+1)^{-1/2-\alpha} + \rho^{(n-k+1)/2} v^{-1/2} \right) \leq Cv^{-3/2}.$$

Note that  $\log \rho \leq -(1-\rho)$  for  $0 < \rho < 1$  implies

$$n\rho^n = n \exp(n \log \rho) \leq n \exp(-n(1-\rho)) \leq 1/(1-\rho) = 1/v,$$

$$n\rho^{n/2} \leq 1/(1-\sqrt{\rho}) \leq C/v.$$

Now we use these bound to estimate the terms on the right hand side of (A.1):

$$V_n \leq Cv^{-2}, \quad nd_n \leq C(n^{-2\alpha} + n\rho^n v^{-1}) \leq Cv^{-2},$$

$$V_n^{1/2} R \leq Cv^{-2}, \quad d_0^{1/2} \sum_{k=1}^n v_k^{1/2} \leq Cv^{-2},$$

$$n^{1/2} d_n^{1/2} R \leq Cn^{1/2} (n^{-1/2-2\alpha} + \rho^{n/2} v^{-1/2}) v^{-1} \leq Cv^{-2},$$

$$nd_n^{1/2} d_0^{1/2} \leq Cn(n^{-1/2-2\alpha} + \rho^{n/2} v^{-1/2}) v^{-1/2} \leq Cv^{-2},$$

we obtain

$$E|T_n - T_{n,\varepsilon}| \leq Cn^{-1} k_n v^{-2}$$

which proves (3.11).

It remains to show (3.13). We have  $T_{n,\varepsilon} = 2\pi \int_{-\pi}^{\pi} h_n(\lambda) I_{n,\varepsilon}(\lambda)$ . By (3.16),

$$\begin{aligned} \text{Var}(T_{n,\varepsilon}) &\leq Cn^{-2} \|E_n\|^2 \leq Cn^{-2} \sum_{t,s=1}^n e_n^2(t-s) \\ &\leq Cn^{-1} \sum_{v=-\infty}^{\infty} e_n^2(v) = Cn^{-1} \int_{-\pi}^{\pi} |\eta_n(\lambda) f(\lambda)|^2 d\lambda \\ &\leq Ck_n^2 n^{-1} v^{-2} \int_{-\pi}^{\pi} f(\lambda) d\lambda \leq Cn^{-1} k_n^2 v^{-3} \end{aligned}$$

by (3.6) and (A.4), which together with (3.11) prove (3.13).  $\square$

**Lemma A.2.** If function  $h_n$  satisfies Assumption 3.3 then, as  $n \rightarrow \infty$ ,

$$\|E_n\|^2 \sim (2\pi)^3 n \|h_n\|^2. \tag{A.5}$$

**Proof of Lemma A.2.** By definition,

$$\begin{aligned} \|E_n\|^2 &= \sum_{t,s=1}^n e_n(t-s)^2 \\ &= (2\pi)^2 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \sum_{t,s=1}^n e^{i(t-s)(x+y)} h_n(x) h_n(y) dx dy \\ &= (2\pi)^2 \int_{-\pi}^{\pi} |D_n(u)|^2 s_n(u) du \end{aligned}$$

where  $D_n(u) = \sum_{t=1}^n e^{itu}$  and  $s_n(u) = \int_{-\pi}^{\pi} h_n(u-x) h_n(x) dx$ ,  $|u| \leq \pi$ . Write

$$\int_{-\pi}^{\pi} |D_n(u)|^2 s_n(u) du = s_n(0) \int_{-\pi}^{\pi} |D_n(u)|^2 du + I_n$$

where  $I_n = \int_{-\pi}^{\pi} |D_n(u)|^2 (s_n(u) - s_n(0)) du$ . Since  $s_n(0) = \|h_n\|^2$  and  $\int_{-\pi}^{\pi} |D_n(u)|^2 du = 2\pi n$ , it suffices to show that

$$|I_n| = o(n \|h_n\|^2). \tag{A.6}$$

For  $K > 0$ , write  $I_n = I_{n,1} + I_{n,2}$  where

$$I_{n,1} = \int_{K/n \leq |u| \leq \pi} |D_n(u)|^2 (s_n(u) - s_n(0)) du,$$

$$I_{n,2} = \int_{|u| \leq K/n} |D_n(u)|^2 (s_n(u) - s_n(0)) du.$$

By the Cauchy inequality

$$|s_n(u)| \leq \left( \int_{-\pi}^{\pi} h_n^2(u-x) dx \right)^{1/2} \left( \int_{-\pi}^{\pi} h_n^2(x) dx \right)^{1/2} = \|h_n\|^2$$

since  $h_n$  is periodically extended to  $R$ . Moreover, for  $n \geq 1$ ,

$$|D_n(u)| \leq C \frac{n}{1+n|u|}, \quad |u| \leq \pi. \tag{A.7}$$

So, for any fixed  $K > 0$ ,

$$I_{n,1} \leq C \|h_n\|^2 \int_{K/n \leq |u| \leq \pi} \frac{n^2}{(1+n|u|)^2} du \leq C \|h_n\|^2 n \delta_K,$$

$$\delta_K := \int_{|u| > K} \frac{1}{(1+|u|)^2} du \rightarrow 0, \quad K \rightarrow \infty.$$

To estimate  $I_{n,2}$ , note that

$$\begin{aligned} \sup_{|u| \leq K/n} |s_n(u) - s_n(0)| &\leq \sup_{|u| \leq K/n} \left( \int_{-\pi}^{\pi} |h_n(u-x) - h_n(x)|^2 dx \right)^{1/2} \\ &\quad \times \left( \int_{-\pi}^{\pi} h_n^2(x) dx \right)^{1/2} = o(\|h_n\|^2) \end{aligned}$$

by Assumption 3.3. Then for fixed  $K > 0$ ,

$$I_{n,2} \leq \sup_{|u| \leq K/n} |s_n(u) - s_n(0)| \int_{|u| \leq \pi} |D_n(u)|^2 du = o(\|h_n\|^2 n)$$

which completes proof of (A.6).  $\square$

**Appendix B**

**Proof of Theorem 2.1.** The general idea of the proof is similar to that of Theorem 18.6.5 in Ibragimov and Linnik (1971). We provide a detailed proof. Start by writing the linear process  $X_t$ , given in (2.1), in the form  $X_t = \sum_{s=-\infty}^t \psi_{t-s} \varepsilon_s$ . Then

$$S_n := \sum_{t=1}^n X_t = \sum_{s=-\infty}^n c_{n,s} \varepsilon_s, \quad c_{n,s} = \sum_{t=\max(1,s)}^n \psi_{t-s}.$$

Write,  $S_n = S_{n,1} + S_{n,2}$  where  $S_{n,1} := \sum_{s=1}^n c_{n,s}$ ,  $S_{n,2} := \sum_{s=-\infty}^0 c_{n,s} \varepsilon_s$ . We shall show that as  $n \rightarrow \infty$ ,

$$\sigma_n^2 \equiv \text{Var}(S_n) \sim n g_0 v^{-2}, \tag{B.1}$$

$$\sigma_n^{-1} S_{n,1} \rightarrow_d N(0, 1), \tag{B.2}$$

$$\sigma_n^{-1} S_{n,2} \rightarrow_p 0, \tag{B.3}$$

which proves (2.11). Observe that

$$\sigma_n^2 = \text{Var}(S_n) = \int_{-\pi}^{\pi} f(\lambda) |D_n(\lambda)|^2 d\lambda$$

where

$$|D_n(\lambda)|^2 = \left| \sum_{t=1}^n e^{it\lambda} \right|^2 = \left| \frac{\sin(n\lambda/2)}{\sin(\lambda/2)} \right|^2.$$

Since by (2.4)  $f(0) = (2\pi)^{-1} v^{-2} g_0$ , then

$$\int_{-\pi}^{\pi} f(0) |D_n(\lambda)|^2 d\lambda = n v^{-2} g_0.$$

Then (B.1) follows if

$$\int_{-\pi}^{\pi} |f(\lambda) - f(0)| |D_n(\lambda)|^2 d\lambda = o(n v^{-2}). \tag{B.4}$$

By (2.3)

$$\begin{aligned} |f(\lambda) - f(0)| &= (2\pi)^{-1} |f^*(\lambda)g(\lambda) - v^{-2}g_0| \\ &\leq C \left( |f^*(\lambda) - f^*(0)|g(\lambda) + f^*(0)|g(\lambda) - g(0)| \right) \\ &\leq C (\lambda^2 f^*(\lambda) v^{-2} + v^{-2} \lambda^2), \end{aligned}$$

since  $f^*(0) \leq C v^{-2}$  and

$$\begin{aligned} |f^*(\lambda) - f^*(0)| &= |v^2 + 2\rho(1 - \cos(\lambda))^{-1} - v^{-2}| \\ &\leq C \lambda^2 f^*(\lambda) v^{-2}, \end{aligned}$$

and, by (2.7),  $|g(\lambda) - g(0)| \leq C \lambda^2$ . Since  $|D_n(\lambda)|^2 \lambda^2 \leq C$ , the left hand side of (B.4) is bounded by

$$\begin{aligned} C \int_{-\pi}^{\pi} v^{-2} (\lambda^2 f^*(\lambda) + \lambda^2) |D_n(\lambda)|^2 d\lambda &\leq C \int_{-\pi}^{\pi} v^{-2} (f^*(\lambda) + 1) d\lambda \\ &\leq C \int_{-\pi}^{\pi} v^{-2} [(v^2 + \rho \lambda^2/3)^{-1} + 1] d\lambda \\ &\leq C v^{-3} = o(n v^{-2}), \end{aligned}$$

because  $n v \rightarrow \infty$  by (2.10), and using the bound (2.5) for  $f^*$ .

Since the  $\varepsilon_t$  are i.i.d. variables with zero mean and unit variance, to prove (B.2) it suffices to check validity of Lindeberg condition, i.e. to show that for any  $\delta > 0$ ,

$$i_n := \sigma_n^{-2} \sum_{s=1}^n E[c_{n,s}^2 \varepsilon_s^2 1_{|c_{n,s} \varepsilon_s| \geq \sigma_n \delta}] \rightarrow 0, \quad n \rightarrow \infty.$$

First we show that,

$$\max_{s=1, \dots, n} |c_{n,s}| = o(\sigma_n). \tag{B.5}$$

Using the notation  $\Psi$  from (3.2), we can write  $\psi_s = (2\pi)^{-1} \int_{-\pi}^{\pi} e^{isx} \Psi(x) dx$ ,  $s = 0, \pm 1, \dots$ . Then, for  $1 \leq s \leq n$ ,

$$c_{n,s} = \sum_{t=\max(1,s)}^n \psi_{t-s} = (2\pi)^{-1} \int_{-\pi}^{\pi} \sum_{t=s}^n e^{i(t-s)x} \Psi(x) dx. \tag{B.6}$$

Using the bound  $|\Psi(x)| \leq C f(x)^{1/2} \leq C f^*(x)^{1/2} \leq C v^{-1}$  which follows from (3.2), (2.3) and (2.4), we obtain

$$\begin{aligned} |c_{n,s}| &\leq C v^{-1} \int_{-\pi}^{\pi} \left| \sum_{t=s}^n e^{i(t-s)x} \right| dx \\ &\leq C v^{-1} \int_{-\pi}^{\pi} |D_{n-s}(x)| dx \leq C v^{-1} \log n \end{aligned}$$

for all  $s = 1, \dots, n$  using (A.7). Since  $\sigma_n \sim C n^{1/2} v^{-1}$ , this proves (B.5).

Fix  $K > 0$ . Then  $\theta_K := E[\varepsilon_s^2 1_{|\varepsilon_s| > K}] \rightarrow 0$ , as  $K \rightarrow \infty$ . Therefore, in view of (B.5),

$$\begin{aligned} E[c_{n,s}^2 \varepsilon_s^2 1_{|c_{n,s} \varepsilon_s| \geq \delta \sigma_n}] &\leq (\delta \sigma_n)^{-2} E[c_{n,s}^4 \varepsilon_s^4 1_{|\varepsilon_s| \leq K}] + c_{n,s}^2 E[\varepsilon_s^2 1_{|\varepsilon_s| > K}] \\ &= c_{n,s}^2 (o(1) + \theta_K). \end{aligned}$$

Then, using  $\sigma_n^2 = \sum_{s=-\infty}^n c_{n,s}^2$ ,

$$\begin{aligned} i_n &\leq \sigma_n^{-2} \sum_{s=1}^n c_{n,s}^2 (o(1) + \theta_K) \\ &= o(1) + \theta_K \rightarrow 0, \quad n \rightarrow \infty, \quad K \rightarrow \infty, \end{aligned}$$

which completes proof of (B.2).

To show (B.3), note that

$$ES_{n,2}^2 = \sum_{s=-\infty}^0 c_{n,s}^2. \tag{B.7}$$

Note that for  $s \leq 0$ , by (B.6),

$$\begin{aligned} c_{n,s} &= (2\pi)^{-1} \int_{-\pi}^{\pi} e^{-isx} \sum_{t=1}^n e^{itx} \Psi(x) dx \\ &= (2\pi)^{-1} \int_{-\pi}^{\pi} e^{-isx} \sum_{t=1}^n e^{itx} (\Psi(x) - \Psi(0)) dx \end{aligned}$$

since  $|\Psi(0)| < \infty$  and  $\int_{-\pi}^{\pi} e^{-isx} \sum_{t=1}^n e^{itx} \Psi(0) dx = 0$  for  $s \leq 0$ . Then by Parseval's identity,

$$ES_{n,2}^2 \leq C \int_{-\pi}^{\pi} \left| \sum_{t=1}^n e^{itx} \right|^2 |\Psi(x) - \Psi(0)|^2 dx.$$

Observe that

$$\Psi(x) = \left( \sum_{t=0}^{\infty} \rho^t e^{-itx} \right) \left( \sum_{s=0}^{\infty} b_s e^{-isx} \right) =: \Psi_{\rho}(x) \Psi_b(x).$$

We have that

$$\begin{aligned} |\Psi_{\rho}(x) - \Psi_{\rho}(0)| &\leq |(1 - \rho e^{-ix})^{-1} - (1 - \rho)^{-1}| \\ &\leq 2|x|v^{-1} |\Psi_{\rho}(x)|, \end{aligned}$$

whereas by (2.6),  $|\Psi_b(x) - \Psi_b(0)| \leq C|x|$ . Then

$$\begin{aligned} |\Psi(x) - \Psi(0)| &\leq |\Psi_{\rho}(x) - \Psi_{\rho}(0)| |\Psi_b(x)| \\ &\quad + |\Psi_{\rho}(0)| |\Psi_b(x) - \Psi_b(0)| \\ &\leq C(|x|v^{-1} |\Psi_{\rho}(x)| + |x|v^{-1}). \end{aligned}$$

So,

$$\begin{aligned} \sigma_n^{-2} ES_{n,2}^2 &\leq C\sigma_n^{-2} \int_{-\pi}^{\pi} |D_n(x)|^2 (|x|v^{-1}|\Psi_\rho(x)| + |x|v^{-1})^2 dx \\ &\leq C\sigma_n^{-2} \int_{-\pi}^{\pi} (v^{-2}|\Psi_\rho(x)|^2 + v^{-2}) dx \leq C\sigma_n^{-2} v^{-3} \\ &\leq C \frac{v^2}{nv^3} = \frac{1}{nv} \rightarrow 0 \end{aligned}$$

by assumption (2.10), which proves (B.3).  $\square$

**Proof of Lemma 2.1.** Using  $f(\lambda) = (2\pi)^{-1}f^*(\lambda)g(\lambda)$ , write

$$\gamma_0 = \int_{-\pi}^{\pi} f(\lambda) d\lambda = s_1 + s_2$$

where

$$\begin{aligned} s_1 &= (2\pi)^{-1} \int_{-\pi}^{\pi} f^*(\lambda)g_0 d\lambda, \\ s_2 &= (2\pi)^{-1} \int_{-\pi}^{\pi} f^*(\lambda)(g(\lambda) - g_0) d\lambda. \end{aligned}$$

Then

$$s_1 = g_0(1 - \rho^2)^{-1} = \frac{g_0}{v(2-v)} = \frac{g_0}{2v} + \frac{g_0}{4} + o(1).$$

By (2.7) and (2.5),

$$f^*(\lambda)|g(\lambda) - g_0| \leq C\lambda^2 f^*(\lambda) \leq C$$

for any  $v$ . Since for each  $\lambda$ ,  $f^*(\lambda) \rightarrow 2(1 - \cos(\lambda))^{-1}$  as  $v \rightarrow 0$ , then by the dominated convergence theorem,

$$s_2 = (2\pi)^{-1} \int_{-\pi}^{\pi} \frac{g(\lambda) - g_0}{2(1 - \cos(\lambda))} d\lambda, \quad v \rightarrow 0 \tag{B.8}$$

which completes the proof of (2.13).

To prove (2.14), write

$$\begin{aligned} \gamma_k &= \int_{-\pi}^{\pi} \cos(k\lambda)f(\lambda) d\lambda = \int_{-\pi}^{\pi} f(\lambda) d\lambda \\ &\quad + \int_{-\pi}^{\pi} (\cos(k\lambda) - 1)f(\lambda) d\lambda =: \gamma_0 + R_k. \end{aligned}$$

Since  $\gamma_0$  satisfies (2.13), and  $|\cos(k\lambda) - 1| \leq C\lambda^2$ , then by the same argument as used in (B.8), it follows

$$R_k \rightarrow -\Gamma_k, \quad v \rightarrow 0,$$

to prove (2.14).

Finally, by (2.14) and (2.13),

$$\begin{aligned} \rho_k &= \frac{\gamma_k}{\gamma_0} = \frac{\gamma_0 - \Gamma_k + o(1)}{\gamma_0} = 1 - \frac{\Gamma_k + o(1)}{\gamma_0} \\ &= 1 - \frac{\Gamma_k + o(1)}{(2v)^{-1}(1 + o(1))} = 1 - 2v\Gamma_k + o(1). \quad \square \end{aligned}$$

**Proof of Theorem 2.2.**

**Proof of (2.16)–(2.17).** Write

$$\hat{\gamma}_k \equiv T_{n,X} := \int_{-\pi}^{\pi} \cos(k\lambda)I(\lambda) d\lambda. \tag{B.9}$$

Applying (3.13) of Theorem 3.1 with  $\eta_n(\lambda) = \cos(k\lambda)$  and  $k_n = 1$ , it follows that

$$|\hat{\gamma}_k - \gamma_k| \leq C \left( \frac{1}{nv_n^2} + \frac{1}{\sqrt{nv_n^3}} \right) \leq C \frac{1}{\sqrt{nv_n^3}},$$

since  $1/(nv_n^2) \leq C/\sqrt{nv_n^3}$  under (2.10). Next, by (2.14),  $1/\gamma_k \leq Cv_n$ , and therefore

$$\hat{\gamma}_k = \gamma_k \left( 1 + O_p \left( \frac{1}{\gamma_k \sqrt{nv_n^3}} \right) \right) = \gamma_k \left( 1 + O_p \left( \frac{1}{\sqrt{nv_n}} \right) \right),$$

proving (2.16).

To prove (2.17), we shall show that assumptions of (ii) imply that (c2) of Theorem 3.2 is satisfied. Set  $h_n(\lambda) = \cos(k\lambda)f(\lambda)$ . Let  $v = v_n$  and

$$\|h_n\|^2 = \int_{-\pi}^{\pi} h_n^2(\lambda) d\lambda, \tag{B.10}$$

$$J_n(u) = \int_{-\pi}^{\pi} |h_n(x+u) - h_n(x)|^2 dx.$$

By Lemma B.1(ii) below,

$$\|h_n\|^2 \sim \frac{1}{8\pi} g_0^2 v^{-3} \tag{B.11}$$

and  $J_n(u) \leq Cu^2 v^{-5}$ . Therefore, for any fixed  $K > 0$ ,

$$\sup_{|u| \leq K/n} J_n(u) \leq C(nv)^{-2} v^{-3} = o(\|h_n\|^2)$$

in view of (B.11), since  $vn \rightarrow \infty$ . Next

$$\int_{-\pi}^{\pi} |h_n(\lambda)| d\lambda \leq C \int_{-\pi}^{\pi} f^*(\lambda) d\lambda \leq Cv^{-1} = o(\|h_n\|^2),$$

because of (B.11). Finally, since  $k_n = 1$

$$\frac{k_n/v^2}{\sqrt{n} \|h_n\|^2} \sim C \frac{1/v^2}{\sqrt{nv^{-3}}} = C \frac{1}{\sqrt{nv}} \rightarrow 0$$

showing that condition (3.24) of Theorem 3.2 is satisfied. Therefore, by (3.27),

$$\sqrt{\frac{n}{4\pi \|h_n\|^2}} (\hat{\gamma}_k - \gamma_k) \xrightarrow{d} N(0, 1),$$

where  $n/(4\pi \|h_n\|^2) \sim 2nv^3/g_0^2$ , proving (2.17).  $\square$

**Proof of (2.18).** We have

$$\hat{\rho}_k - \rho_k = \frac{\int_{-\pi}^{\pi} \cos(k\lambda)I_n(\lambda) d\lambda}{\int_{-\pi}^{\pi} I_n(\lambda) d\lambda} - \frac{\gamma_k}{\gamma_0} = \frac{T_{n,X}}{\hat{\gamma}_0}$$

where

$$T_{n,X} := \int_{-\pi}^{\pi} \eta_n(\lambda)I_n(\lambda) d\lambda, \quad \eta_n(\lambda) = \cos(k\lambda) - \rho_k.$$

Observe that  $\int_{-\pi}^{\pi} \eta_n(\lambda)f(\lambda) d\lambda = 0$  and

$$|\eta_n(\lambda)| \leq C|\cos(k\lambda) - 1| + |1 - \rho_k| \leq C(\lambda^2 + v) \leq C \tag{B.12}$$

by (2.15). Then by (3.11) of Theorem 3.1,

$$T_{n,X} = T_{n,\varepsilon} - E[T_{n,\varepsilon}] + r_k, \quad E|r_k| \leq C(nv^2)^{-1} \tag{B.13}$$

where

$$T_{n,\varepsilon} = 2\pi \int_{\pi}^{\pi} h_n(\lambda)I_{n,\varepsilon} d\lambda, \quad h_n(\lambda) = (\cos(k\lambda) - \rho_k)f(\lambda).$$

Since  $\int_{-\pi}^{\pi} h_n(\lambda) d\lambda = 0$ , from (3.17) and Lemma A.2 it follows that

$$\text{Var}(T_{n,\varepsilon}) = 2(2\pi n)^{-2} \|E_n\|^2 \leq C \|h_n\|^2/n,$$

$$\|h_n\|^2 \equiv \int_{-\pi}^{\pi} \eta_n^2(\lambda)f^2(\lambda) d\lambda.$$

This implies

$$|ET_{n,X}| \leq C \left( \frac{1}{nv^2} + \sqrt{\frac{\|h_n\|^2}{n}} \right).$$

Estimating  $\eta_n(\lambda)$  by (B.12), and noting that for small  $v$ , (2.5) implies  $f(\lambda) \leq C(v^2 + \lambda^2)^{-1}$ , we obtain

$$\begin{aligned} \|h_n\|^2 &\leq C \int_{-\pi}^{\pi} \frac{(\lambda^2 + v)^2}{(\lambda^2 + v^2)^2} d\lambda \leq C + C \int_{-1}^1 \frac{(|\lambda| + v)^2}{(\lambda^2 + v^2)^2} d\lambda \\ &\leq C + Cv^{-1} \int_{-\infty}^{\infty} \frac{(|\lambda| + 1)^2}{(\lambda^2 + 1)^2} d\lambda \leq Cv^{-1}. \end{aligned}$$

Thus

$$|ET_{n,X}| \leq C \left( \frac{1}{nv^2} + \frac{1}{\sqrt{nv}} \right), \quad T_{n,X} = O_p \left( \frac{1}{nv^2} + \frac{1}{\sqrt{nv}} \right).$$

We show below that

$$\hat{\gamma}_0 = \frac{g_0}{2v} \left( 1 + O(v) + o_p(1) \right) \tag{B.14}$$

as  $v \rightarrow 0$ , which implies  $\hat{\rho}_k - \rho_k = O_p(\frac{1}{nv} + \sqrt{\frac{v}{n}})$ , to prove (2.18). In addition we show

$$\sqrt{\frac{2nv^3}{(1 - \rho_k)^2 g_0^2}} (T_{n,\varepsilon} - E[T_{n,\varepsilon}]) \rightarrow_d N(0, 1) \tag{B.15}$$

which together with (B.14) and (B.13) implies (2.20), since  $(nv_n^2)^{-1} = o(1/\sqrt{nv_n})$  when  $nv_n^3 \rightarrow \infty$ .  $\square$

**Proof of (B.14).** Write  $\hat{\gamma}_0 = \int_{-\pi}^{\pi} I_n(\lambda) d\lambda$ . By (3.12) of Theorem 3.1,

$$\hat{\gamma}_0 = \gamma_0 + Q_n + O_p((nv^2)^{-1}), \quad Q_n = T_{n,\varepsilon} - E[T_{n,\varepsilon}].$$

Note that  $\gamma_0 = \frac{g_0}{2v} (1 + O(v))$  by (2.13) of Lemma 2.1. Using the matrix  $E_n$  with entries defined as in (3.14), we can write

$$\begin{aligned} Q_n &= n^{-1} \sum_{t,s=1:t \neq s}^n e_n(t-s) \varepsilon_t \varepsilon_s + e_n(0) n^{-1} \sum_{t=1}^n (\varepsilon_t^2 - E\varepsilon_t^2) \\ &= Q_{n,1} + Q_{n,2}. \end{aligned}$$

Under assumption  $E\varepsilon_t^2 < \infty$ ,

$$\begin{aligned} \text{Var}(Q_{n,1}) &\leq Cn^{-2} \|E_n\|^2 \leq Cn^{-1} \int_{-\pi}^{\pi} f^2(x) dx \\ &\leq C(nv^3)^{-1} = o(v^{-2}) \end{aligned}$$

by Lemma A.2 and (B.18), using assumption  $nv \rightarrow \infty$ . Hence  $Q_{n,1} = o_p(v^{-1})$ . On the other hand, by ergodicity,  $n^{-1} \sum_{t=1}^n (\varepsilon_t^2 - E\varepsilon_t^2) = o_p(1)$ , and

$$e_n(0) = (2\pi)^{-1} \int_{-\pi}^{\pi} f(\lambda) d\lambda \leq C \int_{-\pi}^{\pi} f^*(\lambda) d\lambda \leq Cv^{-1}.$$

Therefore  $Q_{n,2} = o_p(v^{-1})$  which proves (B.14).  $\square$

**Proof of (B.15).** The proof of this fact is based on part (ii) of Lemma 3.1, noting that  $T_{n,\varepsilon}$  satisfies assumption (c1). For that we need to evaluate quantities  $\|h_n\|^2$  and  $J_n(u)$  in (B.10).

Note that

$$\begin{aligned} h_n(x) &= (\cos(kx) - \rho_k) f(x) \\ &= (\cos(kx) - 1) f(x) + (1 - \rho_k) f(x) \\ &= O(1) + (1 - \rho_k) f(x) \end{aligned} \tag{B.16}$$

since  $|(\cos(kx) - 1) f(x)| \leq Cx^2 f^*(x) \leq C, |x| \leq \pi$ . By (2.15),  $1 - \rho_k \sim 2v\Gamma_k, v \rightarrow 0$ . Hence

$$\begin{aligned} h_n^2(x) &= (O(1) + (1 - \rho_k) f(x))^2 \\ &= O(1) + O(v) f(x) + (1 - \rho_k)^2 f^2(x) \end{aligned}$$

and

$$\|h_n\|^2 = \int_{-\pi}^{\pi} h_n^2(x) dx = O(1) + (1 - \rho_k)^2 \int_{-\pi}^{\pi} f(x)^2 dx.$$

By (B.11),  $\int_{-\pi}^{\pi} f^2(x) dx \sim g_0^2 v^{-3} / (8\pi)$  which implies

$$\|h_n\|^2 \sim (1 - \rho_k)^2 \frac{1}{8\pi} g_0^2 v^{-3}. \tag{B.17}$$

To estimate  $J_n(u)$ , note that by (B.16)

$$|(h(x+u) - h(x))| = |O(1) + (1 - \rho_k)(f(x+u) - f(x))|$$

Hence

$$\begin{aligned} J_n(u) &\leq C \left( 1 + (1 - \rho_k)^2 \int_{-\pi}^{\pi} |f(x+u) - f(x)|^2 dx \right) \\ &= C + (1 - \rho_k)^2 O(u^2 v^{-5}) \end{aligned}$$

in view of (B.19). So for  $|u| \leq K/n$ , where  $K$  is a fixed constant,

$$|J_n(u)| \leq C + (1 - \rho_k)^2 O((nv)^{-2} v^{-3}) = o(\|h_n\|^2)$$

because of (B.17) as  $nv \rightarrow \infty$ . Hence  $h_n$  satisfies Assumption 3.3.

It remains to show validity of assumption (3.21) of Lemma 3.1. By (B.12) and (2.5),

$$|h_n(x)| = |\eta_n(x) f(x)| \leq C(x^2 + v)/(v^2 + x^2) \leq Cv^{-1} = k_n^*.$$

Then

$$k_n^*/\sqrt{n \|h_n\|^2} \leq C(1/v)/\sqrt{nv^{-1}} = C/\sqrt{nv} \rightarrow 0.$$

Therefore, by (3.23),

$$\sqrt{n/(4\pi \|h_n\|^2)} (T_{n,\varepsilon} - E[T_{n,\varepsilon}]) \xrightarrow{d} N(0, 1),$$

where  $n/(4\pi \|h_n\|^2) \sim (2nv^3)/((1 - \rho_k)^2 g_0^2) \sim cnv$  which proves (B.15).  $\square$

**Lemma B.1.** (i) Let  $f$  be as in (2.3) and (2.6) holds. Then, as  $v \rightarrow 0$ ,

$$\int_{-\pi}^{\pi} f^2(x) dx \sim \frac{1}{8\pi} g_0^2 v^{-3} \tag{B.18}$$

and

$$J(u) := \int_{-\pi}^{\pi} |f(x+u) - f(x)|^2 dx \leq Cu^2 v^{-5} \tag{B.19}$$

where  $C$  does not depend on  $u$  and  $v$ .

(ii) Estimates (B.18) and (B.19) remain valid when  $f(x)$  is replaced by  $\cos(kx)f(x)$ .

**Proof of Lemma B.1.** First we show (B.18). Note that  $f = (2\pi)^{-1} f^* g$  where  $|g(x) - g_0| \leq Cx^2$ , and  $x^2 f^*(x) \leq C$ . Hence,

$$\begin{aligned} f(x) &= (2\pi)^{-1} f^*(x) g(x) = (2\pi)^{-1} f^*(x) g_0 + O(1), \\ f^2(x) &= (2\pi)^{-2} g_0^2 f^{*2}(x) + O(1) f^*(x) + O(1) \\ &= (2\pi)^{-2} g_0^2 f^{*2}(x) + O(v^{-2}). \end{aligned}$$

Observe that, as  $v \rightarrow 0$ ,

$$\begin{aligned} \int_{-\pi}^{\pi} f^{*2}(x) dx &= \int_{-\pi}^{\pi} v^2 + 2\rho(1 - \cos(x))^{-2} dx \\ &\sim v^{-3} \int_{-\infty}^{\infty} (1 + x^2)^{-2} dx \sim \frac{\pi}{2} v^{-3}, \end{aligned}$$

since  $\int_{-\infty}^{\infty} (1+x^2)^{-2} dx = \pi/2$ , see Jeffrey (1995, 15.1.1(16)). Hence

$$\begin{aligned} \int_{-\pi}^{\pi} f^2(x) dx &= (2\pi)^{-2} g_0^2 \int_{-\pi}^{\pi} (f^*(x)^2 + O(v^{-2})) dx \\ &= \frac{1}{8\pi} g_0^2 v^{-3} (1 + o(1)), \end{aligned}$$

to prove (B.18)

To show (B.19), note that

$$|f^*(x+u) - f^*(x)| \leq 2|\cos(x+u) - \cos(x)|f^*(x+u)f^*(x) \leq Cu(|x| + |x+u|)f^*(x+u)f^*(x)$$

since

$$|\cos(x+u) - \cos(x)| \leq |u| \sup_{\xi \in [x, x+u]} |\sin(\xi)| \leq |u|(|x| + |x+u|).$$

Since  $f^*(x) \leq Cv^{-2}$  and  $|x|\sqrt{f^*(x)} \leq C$ , then  $|x|f^*(x) \leq Cv^{-1}$ , and

$$|f^*(x+u) - f^*(x)| \leq C|u|v^{-1}(f^*(x) + f^*(x+u)).$$

Under assumption (2.6),  $|g(x+u) - g(x)| \leq C|u|$ . Therefore

$$\begin{aligned} |f(x+u) - f(x)| &= |f^*(x+u)g(x+u) - f^*(x)g(x)| \\ &\leq C|f^*(x+u) - f^*(x)| \\ &\quad + f^*(x)|g(x) - g(x+u)| \\ &\leq C|u|v^{-1}(f^*(x) + f^*(x+u)). \end{aligned}$$

Hence

$$\begin{aligned} J(u) &\leq Cu^2v^{-2} \int_{-\pi}^{\pi} (f^*(x) + f^*(x+u))^2 dx \\ &\leq Cu^2v^{-2} \int_{-\pi}^{\pi} f^*(x)^2 dx \leq Cu^2v^{-5} \end{aligned}$$

by (B.18), which proves (B.19).

In case (ii), the estimates (B.18)–(B.19) follow using the same argument.  $\square$

**Proof of Theorem 2.3.** By (B.14), we have that

$$2v\hat{\gamma}_0 = g_0 + O(v) + o_p(1).$$

We shall show that

$$t_n \equiv \int_{-\pi}^{\pi} |x|^{1/2} I_n(x) dx = \frac{g_0}{\sqrt{2v}} + O_p\left(1 + \frac{1}{nv^2} + \frac{1}{v\sqrt{n}}\right) \quad (\text{B.20})$$

which implies (2.22). By (3.13) of Theorem 3.1,

$$t_n = \int_{-\pi}^{\pi} |x|^{1/2} f(x) dx + O_p\left(\frac{1}{nv^2} + \sqrt{\frac{\|h_n\|^2}{n}}\right)$$

where, using (2.5),

$$\|h_n\|^2 = \int_{-\pi}^{\pi} |x|f^2(x) dx \leq C \int_{-\pi}^{\pi} |x|(v^2 + x^2)^{-2} dx \leq Cv^{-2}.$$

To prove (B.20) it remains to show that

$$i_n := \int_{-\pi}^{\pi} |x|^{1/2} f(x) dx = (2v)^{-1/2} g_0 + O(1). \quad (\text{B.21})$$

Write

$$\begin{aligned} i_n &:= (2\pi)^{-1} \int_{-\pi}^{\pi} |x|^{1/2} f^*(x) g_0 dx \\ &\quad + (2\pi)^{-1} \int_{-\pi}^{\pi} |x|^{1/2} f^*(x) (g(x) - g_0) dx \\ &= i_{n,1} + i_{n,2}. \end{aligned}$$

Since  $|f^*(x)(g(x) - g_0)| \leq Cf^*(x)x^2 \leq C$ , then  $i_{n,2} \leq C$ . To estimate  $i_{n,1}$ , write  $i_{n,1} = j_{n,1} + j_{n,2}$ , where

$$j_{n,1} = g_0(2\pi)^{-1} \int_{-\pi}^{\pi} |x|^{1/2} (v^2 + x^2)^{-1} dx,$$

$$j_{n,2} = (2\pi)^{-1} \int_{-\pi}^{\pi} |x|^{1/2} (f^*(x) - (v^2 + x^2)^{-1}) dx.$$

Observe that

$$\begin{aligned} |f^*(x) - (v^2 + x^2)^{-1}| &\leq |x^2 - 2\rho(1 - \cos(x))|f^*(x)(v^2 + x^2)^{-1} \\ &\leq C(vx^2 + x^4)(v^2 + x^2)^{-2} \\ &\leq C(v(v^2 + x^2)^{-1} + 1) \end{aligned}$$

since

$$|x^2 - 2\rho(1 - \cos(x))| = |x^2 - \rho(x^2 + O((x^4)))| = x^2v + O(x^4)$$

and  $f(x) \leq C(v^2 + x^2)^{-1}$  by (2.5), as  $v \rightarrow 0$ . So,

$$|j_{n,2}| \leq C \int_{-\pi}^{\pi} |x|^{1/2} (v(v^2 + x^2)^{-1} + 1) dx \leq C.$$

Next, observe that

$$\begin{aligned} A &:= \int_{-\infty}^{\infty} |x|^{1/2} (1 + x^2)^{-1} dx = \int_0^{\infty} |y|^{-1/4} (1 + y)^{-1} dy \\ &= \frac{\pi}{\sin(3\pi/4)} = \sqrt{2}\pi, \end{aligned}$$

using formula 15.1.1(2) from Jeffrey (1995):

$$\int_0^{\infty} \frac{y^{p-1}}{1+y} dy = \frac{\pi}{\sin(p\pi)}, \quad 0 < p < 1.$$

Therefore, changing variables we obtain,

$$\begin{aligned} j_{n,1} &= g_0v^{-1/2}(2\pi)^{-1} \int_{-\pi/v}^{\pi/v} |x|^{1/2} (1 + x^2)^{-1} dx \\ &= g_0v^{-1/2}(2\pi)^{-1}A + O(1) = g_0(2v)^{-1/2} + O(1) \end{aligned}$$

which together with estimates above implies (B.21).  $\square$

## References

- Bartlett, M.S., 1955. An Introduction to Stochastic Processes, with Special Reference to Methods and Applications. Cambridge University Press.
- Bhansali, R.J., Giraitis, L., Kokoszka, P., 2007. Decomposition and asymptotic properties of quadratic forms in linear variables. *Stochastic Processes and their Applications* 117, 71–95.
- Bobkoski, M.J., 1983. Hypothesis testing in nonstationary time series. Ph.D Thesis. Department of Statistics, University of Wisconsin.
- Brockwell, P.J., Davis, R.A., 1991. *Time Series: Theory and Methods*, 2nd ed. Springer-Verlag, New York.
- Cavanagh, C., 1985. Roots local to unity. Manuscript. Department of Economics, Harvard University.
- Chan, N.H., Wei, C.Z., 1987. Asymptotic inference for nearly nonstationary AR(1) processes. *Annals of Statistics* 15, 1050–1063.
- Giraitis, L., Phillips, P.C.B., 2006. Uniform limit theory for stationary autoregression. *Journal of Time Series Analysis* 26, 51–60.
- Hannan, E.J., 1973. The asymptotic theory of linear time series models. *Journal of Applied Probability* 10, 130–145.
- Hannan, E.J., Heyde, C.C., 1972. On limit theorems for quadratic functions of discrete time series. *The Annals of Mathematical Statistics* 43, 2058–2066.
- Hosking, J.R.M., 1996. Asymptotic distributions of the sample mean, autocovariances, and autocorrelations of long-memory time series. *Journal of Econometrics* 73, 261–284.
- Ibragimov, I.A., Linnik, Yu.V., 1971. *Independent and Stationary Sequences of Random Variables*. Wolters-Noordhoff, Groningen.
- Jeffrey, A., 1995. *Handbook of Mathematical Formulas and Integrals*. Academic Press, London.
- Magdalinos, T., Phillips, P.C.B., 2008. Limit theory for cointegrated systems with moderately integrated and moderately explosive regressors. Working Paper. Yale University.
- Phillips, P.C.B., 1987. Towards a unified asymptotic theory for autoregression. *Biometrika* 74, 535–547.
- Phillips, P.C.B., Magdalinos, T., 2007a. Limit theory for moderate deviations from a unit root. *Journal of Econometrics* 136, 115–130.
- Phillips, P.C.B., Magdalinos, T., 2007b. Limit theory for moderate deviations from a unit root under weak dependence. In: Phillips, G.D.A., Tzavalis, E. (Eds.), *The Refinement of Econometric Estimation and Test Procedures: Finite Sample and Asymptotic Analysis*. Cambridge University Press, Cambridge.