MEAN AND AUTOCOVARIANCE FUNCTION ESTIMATION NEAR THE BOUNDARY OF STATIONARITY

By

Liudas Giraitis and Peter C. B. Phillips

January 2009

COWLES FOUNDATION DISCUSSION PAPER NO. 1690

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

http://cowles.econ.yale.edu/

Mean and Autocovariance function Estimation Near the Boundary of Stationarity

Liudas Giraitis*
Queen Mary College, University of London

Peter C.B. Phillips†
Yale University, University of Auckland,
University of York & Singapore Management University

Abstract

We analyze the applicability of standard normal asymptotic theory for linear process models near the boundary of stationarity. The concept of stationarity is refined, allowing for sample size dependence in the array and paying special attention to the rate at which the boundary unit root case is approached using a localizing coefficient around unity. The primary focus of the present paper is on estimation of the 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. An asymptotic result on the estimation of the localizing coefficient is also presented. To assist in the development of the limit theory in the present case, a suitable asymptotic theory for the behavior of quadratic forms in the vicinity of the boundary of stationarity is provided.

JEL classification: C22

Keywords: Asymptotic normality, Integrated periodogram, Linear process, Local to unity, Localizing coefficient, Moderate deviation, Unit root.

*Supported by the ESRC grant RES062230790.
†Partial support from the NSF is acknowledged under Grant No. SES 06-47086.

1 Introduction

The idea of developing asymptotics in near unit root situations is due at various levels of generality to Bobkowski (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 $O\left(n^{-1}\right)$ 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

$$X_t = \rho X_{t-1} + \varepsilon_t, \qquad t = 1, ..., n \tag{1.1}$$

where ε_t is $iid\,(0,1)$ noise and X_0 is some appropriate fixed or 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

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

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

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 ρ taking values in the region

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

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

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

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 (2007). Likewise the region of explosive behavior may be characterized as

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

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

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

In each of these cases we may consider ρ (and hence $v = 1-\rho$) to be functionally dependent on n, or at least confined to an interval that depends on n, thereby making the process X_t in (1.1) an array. This formulation will be understood throughout the paper even though it is seldom made explicit.

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 ρ in (1.1) has a non-Gaussian limit distribution. The size of the stationarity region is determined by the sample size n and ρ, and when $n(1-\rho)$ is large, $\hat{\rho}_n$ has the same asymptotic properties as in the (fixed ρ) stationary case. That is,

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

as shown in Phillips and Magdalinos (2007) 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 ρ and may increase from $\sqrt{n}$ towards the unit root rate n for small $1-\rho$.

It follows from (1.2) that standard asymptotic estimation and inferential theory applies over the whole region of ρ 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 ρ 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 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. Section 5 contains proofs of the supporting asymptotic theory of Section 3. Proofs of the main results of Section 2 are given in Section 6.

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

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

where $0 < \rho < 1$ and Y_t is a linear MA(∞) process with coefficients b_j, where (ε_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 ρ 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

$$f^*(\lambda) = |1 - \rho e^{i\lambda}|^{-2} = \frac{1}{v^2 + 2\rho(1 - \cos(\lambda))}, \quad g(\lambda) = |\sum_{s=0}^{\infty} b_s e^{-i\lambda s}|^2$$

and $|\rho| < 1$. Then

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

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

using $2(1 - \cos(\lambda)) \geq \lambda^2/3$ for $|\lambda| \leq \pi$. We shall assume that $g_0 > 0$, and

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

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

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

It is natural to raise the question of how the closeness of the parameter ρ to 1 impacts the validity of the usual normal approximation of the distribution of the sample mean and second moments. Moreover, if ρ 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, ρ) for which the asymptotic theory corresponding to a stationary model with fixed ρ continues to apply.

We also examine the effect of the closeness of ρ 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 ρ with $|\rho| < 1$ as $n \to \infty$,

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

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

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

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 ρ and n does the approximation implied by the limit theory (2.8)-(2.9) continue to hold? We shall show that, for given (ρ, 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*

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

Then the convergence

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

holds.

The proof of Theorem 2.1 is given in Appendix 2.

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

$$n^{-1/2} << \sqrt{nv^2} \leq \sqrt{n}.$$

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 μ, and estimation of μ dramatically worsens in quality as the unit root model is approached. The sample mean $\bar{X}$ is a consistent estimator of μ only if $v >> n^{-1/2}$, and μ cannot be consistently estimated when $n^{-1} << v << n^{-1/2}$, 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

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

where W_t is the standard Wiener process and ω_Y^2 is the long run variance of 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

$$\gamma_j = Cov(X_j, X_0) = \int_{-\pi}^{\pi} cos(\lambda j) f(\lambda) d\lambda, \quad j \geq 0$$

and the autocorrelation function $\rho_j = \frac{\gamma_j}{\gamma_0}$, $j = 0, 1, 2, ...$ 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 γ_j and ρ_j as $\rho \to 1$. Set

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

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 memory and long memory time series were studied in Hosking (1995). We focus here on stationary short memory time series which approach the unit root region. First, we discuss some asymptotic properties of γ_j and ρ_k as $v \to 0$.

LEMMA 2.1 *For fixed* $k = 0, 1, 2, ..,$ *as* $v \to 0$,

$$\gamma_0 = \frac{g_0}{2v} + \frac{g_0}{4} + \Gamma_0 + o(1), \tag{2.12}$$

$$\gamma_k = \gamma_0 - \Gamma_k + o(1) = \frac{g_0}{2v} + \frac{g_0}{4} + \Gamma_0 - \Gamma_k + o(1), \; k = 1, 2, ... \tag{2.13}$$

and

$$\rho_k = 1 - 2v\Gamma_k + o(v), \quad k = 1, 2, ... \tag{2.14}$$

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

THEOREM 2.2 *Assume that* $(X_1, ..., X_n)$ *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*

$$E|\hat{\gamma}_k - \gamma_k| \leq C \frac{1}{\sqrt{n v_n^3}}, \quad \hat{\gamma}_k = \gamma_k (1 + O_P(\frac{1}{\sqrt{n v_n}})) \tag{2.15}$$

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

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

$$\sqrt{\frac{2nv^3}{g_0^2}}(\hat{\gamma}_k - \gamma_k) \to_d N(0,1). \tag{2.16}$$

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

$$|\hat{\rho}_k - \rho_k| = O_P\left(\frac{1}{nv_n} + \sqrt{\frac{v_n}{n}}\right). \tag{2.17}$$

Moreover, if

$$n\, v_n^3 \to \infty, \quad v_n \to 0,\ n \to \infty \tag{2.18}$$

then

$$\sqrt{\frac{nv_n}{2(1-\rho_k)^2}}(\hat{\rho}_k - \rho_k) \to_d N(0,1), \quad \sqrt{\frac{nv_n}{2(1-\rho_k)^2}} \sim \frac{1}{2\Gamma_k}\sqrt{\frac{n}{2v_n}} \tag{2.19}$$

The following theorem considers estimation of the quantity $\sqrt{v}$. Denote the periodogram by $I_n(\lambda) = (2\pi n)^{-1}|\sum_{t=1}^n e^{it\lambda}X_t|^2$, and define

$$\sqrt{\hat{v}_n} = \frac{1}{\sqrt{2}}\frac{\int_{-\pi}^{\pi}\sqrt{\lambda}\, I_n(\lambda)d\lambda}{\int_{-\pi}^{\pi} I_n(\lambda)d\lambda}. \tag{2.20}$$

Theorem 2.3 *Assume that* $(X_1, ..., X_n)$ *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*

$$\sqrt{\hat{v}_n} = \sqrt{v}_n + O_P\left(v_n + \frac{1}{nv_n} + \frac{1}{\sqrt{n}}\right). \tag{2.21}$$

The proofs of Lemma 2.1 and Theorems 2.2-2.3 are given in Appendix 2.

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 $v_n \to 0$, is asymptotically negligible. This representation leads to the requirement of finite $2+\delta$ moments of ε_t in (ii), and second moments in (iii). In the case where v_n is fixed, the convergence (2.16) requires finite fourth moments of ε_t.

(ii) It follows from (2.15) that $\hat{\gamma}_k$ is a consistent estimate of γ_k. The CLT (2.16) is valid with the convergence rate $\sqrt{nv_n^3}$ which depends on the value of v_n and varies in the interval

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

(iii) As v_n decreases, confidence intervals for γ_k will increase. When $nv_n^3 \to 0$ then (2.16) can be written in the form

$$\hat{\gamma}_k \sim \gamma_k(1 + \frac{g_0}{\gamma_k\sqrt{2nv^3}}Z) \sim \gamma_k(1 + \sqrt{\frac{2}{nv}}Z), \quad Z \sim N(0,1).$$

(iv) Theorem 2.2 shows that the sample autocorrelation $\hat{\rho}_k$ is a consistent estimator of ρ_k as long as $nv \to \infty$, and $\mathrm{E}\varepsilon_t^2 < \infty$. 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 $nv^3 \to \infty$.

(v) To apply these results in samples of size n we set $v_n = v = 1 - \rho$ where ρ 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.

Figure 1 shows the ACF ρ_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. Figure 2 shows a realization of the sample ACF, $\hat{\rho}_k$, computed from a sample of $n = 125$ observations. Figures 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 (2.19) that the rate of convergence $\sqrt{n/v}$ of the sample ACF in the near unit root region improves when $v \to 0$, and nv^3 remains large. That condition is not well satisfied when $r = 0.95$, partly explaining the large bias in this case[1].

Figures 5-6 indicate the adequacy of the standard normal approximation (2.19) to the probability density function of the standardized sample ACFs $\hat{t}_n(k) = \sqrt{\frac{n\hat{v}}{2(1-\hat{\rho}_k)^2}}(\hat{\rho}_k - \rho_k)$ 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 $r = 0.8, 0.95$, 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.

Figures 1 - 6 about here

3 Asymptotic theory for quadratic forms

We assume that

(3.1) $$X_j = \Big(\sum_{t=0}^{\infty} a_t L^t\Big) Y_j = \sum_{t=0}^{\infty} a_t Y_{j-t}, \quad j = 0, 1, 2, ..$$

[1]We thank Violetta Dalla for preparing Figures 1-4.

is a linear process where

$$Y_t = \Big(\sum_{s=0}^{\infty} b_s L^s\Big)\varepsilon_s = \sum_{s=0}^{\infty} b_s \varepsilon_{t-s},$$

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

$$X_j = \Big(\sum_{t=0}^{\infty} a_t L^t\Big)\Big(\sum_{s=0}^{\infty} b_s L^s\Big)\varepsilon_j = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}, \quad t = 0, 1, 2, ...$$

with

$$\psi_j = \sum_{k=0}^{j} a_k b_{j-k}, \;\; k = 0, 1, 2, 3, ...$$

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

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

where

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

We impose the following restrictions on a_t and b_s.

Assumption 3.1 *(i) The coefficients a_j satisfy*

$$|a_j| \le C\rho^j, \; j = 1, 2, 3, ... \tag{3.3}$$

for some $0 \le \rho < 1$, where $\rho = \rho_n$ may depend on n.
(ii) The coefficients b_s are such that

$$\sum_{s=j}^{\infty} |b_s| \le Cj^{-1-\alpha}, \; j = 1, 2, 3, ... \tag{3.4}$$

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 ρ. As before, we let

$$v = 1 - \rho. \tag{3.5}$$

Under Assumption 3.1, the spectral density

$$f(\lambda) \le C|\sum_{t=0}^{\infty} \rho^t|^2 \le Cv^{-2} \tag{3.6}$$

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

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

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

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

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

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 ρ 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 ε_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 ε_j. The asymptotic properties of $T_{n,\varepsilon}$ are much easier to analyze then those of $T_{n,X}$, as long as $T_{n,\varepsilon}$ dominates the remainder $T_{n,X} - T_{n,\varepsilon}$. Our objective is to derive a precise upper bound for the remainder term. Then, using asymptotic theory for 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 η_n have the following property.

ASSUMPTION 3.2 *η_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 η_n are bounded but their upper bound k_n might vary with n, for example, η_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 R. Set

$$B_n = \int_{-\pi}^{\pi} h_n^2(x) dx. \tag{3.10}$$

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

$$T_{n,X} = \int_{-\pi}^{\pi} \eta_n(\lambda) f(\lambda) d\lambda + (T_{n,\varepsilon} - E[T_{n,\varepsilon}]) + r_n, \quad E|r_n| \leq C \frac{k_n}{nv^2}. \tag{3.12}$$

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

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

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 ρ, with no restrictions on ρ imposed. The constants k_n play a secondary role. If the functions $\eta_n(\lambda)$ not depend on n, we can set $k_n = 1$. The proof of Theorem 3.1 is given in 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 $\mathrm{Var}(T_{n,\varepsilon})$, 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

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

Then

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

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 assumption.

ASSUMPTION 3.3 *For any $K > 0$,*

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

In Lemma 5.2 in Appendix 1 we show that under Assumption 3.3,

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

Lemma 3.1 below 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, Giraitis and Kokoszka (2007) and Lemma 5.2 below. It takes into account the fact that the upper bound k_n^* in $|\eta_n(\lambda)|f(\lambda) \le k_n^*$ might be smaller than the product $Ck_n \times v^{-2}$ of the upper bounds $|\eta_n(\lambda)| \le k_n$ and $f(\lambda) \le Cv^{-2}$.

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

Case (c1):

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

Case (c2):

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

Case (c1) corresponds to the case where E_n has zero diagonal, whereas case (c2) corresponds to the case of an asymptotically vanishing diagonal.

LEMMA 3.1 *Suppose that h_n satisfies Assumption 3.3,*

$$|h_n(\lambda)| \le k_n^*, \ n \ge 1,$$

and

$$\frac{k_n^*}{\sqrt{nB_n}} \to 0, \quad n \to \infty. \tag{3.21}$$

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

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

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

$$\sqrt{\frac{n}{4\pi B_n}}\Big(T_{n,\varepsilon} - \int_{-\pi}^{\pi} h_n(\lambda)d\lambda\Big) \xrightarrow{d} N(0,1). \tag{3.23}$$

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) below 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, 3.2 and 3.3 are satisfied and, as $n \to \infty$,*

$$\frac{k_n/v^2}{\sqrt{nB_n}} \to 0. \tag{3.24}$$

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

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

and

$$\mathrm{Var}(T_{n,X}) \sim \mathrm{Var}(T_{n,\varepsilon}) \asymp \frac{B_n}{n}. \tag{3.26}$$

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

$$\sqrt{\frac{n}{4\pi B_n}}\Big(T_{n,X} - \int_{-\pi}^{\pi} \eta_n(\lambda)f(\lambda)d\lambda\Big) \xrightarrow{d} N(0,1). \tag{3.27}$$

4 Discussion

The idea of approximation results similar to those in Theorem 3.1 goes back to the work of Hannan and Heyde (1972) and Hannan (1973). 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 stable ARMA process. Brockwell and Davis (1991), Proposition 10.8.5, showed that

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

Recently, Bhansali, Giraitis and Kokoszka (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

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

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

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. The Theorem simplifies the derivation of asymptotics for statistics which can be written in the form of functionals of the integrated periodogram $T_{n,X}$.

5 Appendix 1

Proof of Theorem 3.1. Set

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

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 $R_n = \sum_{j=0}^{\infty} |\psi_j|$. First we prove the following result.

Lemma 5.1 *Assume that $\eta_n(\lambda)$ satisfies Assumption 3.2. Then*

$$(5.1)\; E|T_n - T_{n,\varepsilon}| \le Cn^{-1} k_n \Big(V_n + nd_n + V_n^{1/2} R + d_0^{1/2} \sum_{k=0}^{n} v_k^{1/2} + nd_n^{1/2} R + nd_n^{1/2} d_0^{1/2}\Big)$$

where C does not depend on n and ρ.

Proof of Lemma 5.1. Define

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

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:

$$\nu_n(k,t) = \int_{-\pi}^{\pi} e^{i(t-k)\lambda} d_k(\lambda)\overline{d_t(\lambda)}|\eta_n(\lambda)|d\lambda, \quad \beta_n(k,t) = \int_{-\pi}^{\pi} e^{i(t-k)\lambda} d_k(\lambda)\Psi(\lambda)\eta_n(\lambda)d\lambda,$$

$$\mu_n(k,t) = \int_{-\pi}^{\pi} e^{i(t-k)\lambda}|c_n(\lambda)|^2|\eta_n(\lambda)|d\lambda, \quad \zeta_n(k,t) = \int_{-\pi}^{\pi} e^{i(t-k)\lambda} c_n(\lambda)\Psi(\lambda)\eta_n(\lambda)d\lambda.$$

Observe that

$$v_k = \int_{-\pi}^{\pi} |d_k(\lambda)|^2 d\lambda; \quad d_n := \int_{-\pi}^{\pi} |c_n(\lambda)|^2 d\lambda; \quad d_0 := \int_{-\pi}^{\pi} |\Psi(\lambda)|^2 d\lambda$$

$$v_k = C\sum_{j=1-k}^{n-k} \psi_j^2 \text{ for } k \le 0;\ C\sum_{j=n-k+1}^{n} \psi_j^2 \text{ for } 1 \le k \le n;$$

$$d_n = C\sum_{j=n+1}^{\infty} \psi_j^2.$$

To derive the bound (5.1), we shall use the estimate (5.31) of BGL(2007)

$$E|T_n - T_{n,\varepsilon}| \le Cn^{-1}(E|Y_n| + E|V_{n,1}| + E|V_{n,2}|) \tag{5.2}$$

where it was shown that

$$E|Y_n| \le C\Big(\sum_{k=-\infty}^{n} \nu_n(k,k) + \sum_{k=1}^{n} \mu_n(k,k)\Big) =: C[s_{n,1} + s_{n,2}];$$

$$E|V_{n,1}| \le C\Big[\Big(\sum_{k=-\infty}^{n}\sum_{t=1:t\ne k}^{n} |\beta_n(k,t)|^2\Big)^{1/2} + \sum_{k=1}^{n} |\beta_n(k,k)|\Big] =: C[s_{n,3}^{1/2} + s_{n,4}]$$

$$E|V_{n,2}| \le C\Big[\Big(\sum_{k=1}^{n}\sum_{t=1:t\ne k}^{n} |\zeta_n(k,t)|^2\Big)^{1/2} + \sum_{k=1}^{n} |\zeta_n(k,k)|\Big] =: C[s_{n,5}^{1/2} + s_{n,6}].$$

Recall that $|\eta_n(\lambda)| \le k_n$. Hence, by (5.2),

$$|s_{n,1}| \le Ck_n \sum_{k=-\infty}^{n} \int_{-\pi}^{\pi} |d_k(\lambda)|^2 d\lambda = Ck_n \sum_{k=-\infty}^{n} v_k = Ck_n V_n;$$

$$|s_{n,2}| \le Ck_n n \int_{-\pi}^{\pi} |c_n(\lambda)|^2 d\lambda = Ck_n n d_n;$$

$$|s_{n,4}| \le Ck_n \sum_{k=1}^{n} \int_{-\pi}^{\pi} |d_k(\lambda)||\Psi(\lambda)|d\lambda$$

$$\le Ck_n \sum_{k=1}^{n} \Big(\int_{-\pi}^{\pi} |d_k(\lambda)|^2 d\lambda\Big)^{1/2}\Big(\int_{-\pi}^{\pi} |\Psi(\lambda)|^2 d\lambda\Big)^{1/2} \le Ck_n \sum_{k=1}^{n} v_k^{1/2} d_0^{1/2},$$

$$|s_{n,6}| \le Ck_n n \int_{-\pi}^{\pi} |c_n(\lambda)||\Psi(\lambda)|d\lambda \le k_n n d_n^{1/2} d_0^{1/2}.$$

The estimates (5.26)-(5.27) and (5.28)-(5.29) of BGK(2007) imply that

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

Since $|\eta_n(\lambda| \le k_n$ and

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

it follows that

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

Hence

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

The above bounds for $s_{n,j}$, $j = 1, ..., 6$ prove (5.1). ∎

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

$$\psi_j = \sum_{s=0}^{j} a_s b_{j-s}, \ s \ge 0$$

Recall that $|a_j| \le C\rho_j$, $j = 1, 2, ...$ and $\sum_{j=k}^{\infty} |b_j| \le C|k|^{-1-\alpha}$, $k \ge 1$, where $\alpha > 1/2$.

First we show that

$$|\psi_j| \le C(j^{-1-\alpha} + \rho^{j/2}), \tag{5.3}$$

where C does not depend on ρ and $j = 1, 2, 3, ...$ 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 ψ_j^- we have $j - t \ge j/2$. Therefore

$$\begin{aligned} |\psi_j^-| &\le C \sum_{t=0}^{j/2} |b_{j-t}| \le C \sum_{v=j/2}^{\infty} |b_v| \le C|j|^{-1-\alpha}, \quad j = 1, 2, ... \\ |\psi_j^+| &\le C \sum_{t=j/2+1}^{j} \rho^t b_{j-t} \le C\rho^{j/2} \sum_{v=0}^{\infty} |b_v| \le C\rho^{j/2}. \end{aligned}$$

Applying (5.3), it follows that for $k \ge 1$,

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

Using this bound in (5.2), it follows that

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

and

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

$$\int_{-\pi}^{\pi} f(\lambda) d\lambda = C \sum_{j=0}^{\infty} \psi_j^2 = C d_0 \le C v^{-1}. \tag{5.4}$$

Then

$$V_n = \sum_{k=-\infty}^{n} v_k \le C \sum_{k \le 0} ((-k+1)^{-1-2\alpha} + \rho^{-k} v^{-1}) + \sum_{k=1}^{n} ((n-k+1)^{-1-2\alpha} + \rho^{n-k+1} v^{-1})$$

$$\le C(\sum_{k=0}^{\infty} (k+1)^{-1-2\alpha} + \sum_{k=0}^{\infty} \rho^k v^{-1}) \le C v^{-2},$$

and

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

Moreover, since

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

and $\alpha > 1/2$, then

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

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

$$\begin{aligned} n\rho^n &= n \exp(n \log \rho) \le n \exp(-n(1-\rho)) \le 1/(1-\rho) = 1/v, \\ n\rho^{n/2} &\le 1/(1-\sqrt{\rho}) \le C/v. \end{aligned}$$

Now we use these bound to estimate the terms on the right hand side of (5.1):

$$\begin{aligned} V_n \le C v^{-2}, \quad n d_n &\le C((n^{-2\alpha} + n\rho^n v^{-1}) \le C v^{-2}, \\ V_n^{1/2} R &\le C v^{-2}, \quad d_0^{1/2} \sum_{k=1}^{n} v_k^{1/2} \le C v^{-2} \\ n d_n^{1/2} R &\le C n (n^{-1/2-\alpha} + \rho^{n/2}) v^{-1} \le C v^{-2}, \\ n d_n^{1/2} d_0^{1/2} &\le C n (n^{-1/2-\alpha} + \rho^{n/2} v^{-1/2}) v^{-1/2} \le C v^{-2}, \end{aligned}$$

we obtain

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

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),

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

by (3.6) and (5.4), which together with (3.11) prove (3.13) ∎

Lemma 5.2 *If function h_n satisfies Assumption 3.3 then, as $n \to \infty$,*

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

Proof of Lemma 5.2. By definition,

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

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

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

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

(5.6) $$|I_n| = o(nB_n).$$

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

$$I_{n,1} = \int_{K/n \le |u| \le \pi} |D_n(u)|^2 (B_n(u) - B_n(0)) du, \;\; I_{n,2} = \int_{|u| \le K/n} |D_n(u)|^2 (B_n(u) - B_n(0)) du.$$

By the Cauchy inequality

$$|B_n(u)| \le (\int_{-\pi}^{\pi} h_n^2(u-x) dx)^{1/2} (\int_{-\pi}^{\pi} h_n^2(x) dx)^{1/2} = B_n$$

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

$$(5.7) \qquad |D_n(x)| \leq C \frac{n}{1+n|x|}, \;\; |x| \leq \pi.$$

So, for any fixed $K > 0$,

$$I_{n,1} \leq CB_n \int_{K/n \leq |u| \leq \pi} \frac{n^2}{(1+n|x|)^2} dx \leq CB_n n \delta_K$$

where

$$\delta_K = \int_{|u|>K} \frac{1}{(1+|x|)^2} dx \to 0, \;\; K \to \infty.$$

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

$$\sup_{|u| \leq K/n} |B_n(u) - B_n(0)| \leq \sup_{|u| \leq K/n} (\int_{-\pi}^{\pi} |h_n(u-x) - h(x)|^2 dx)^{1/2} (\int_{-\pi}^{\pi} h_n^2(x) dx)^{1/2} = o(B_n)$$

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

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

for any which proves (5.6). ∎

6 Appendix 2

Proof of Theorem 2.1. The general idea of the proof is similar to that of Theorem 18.6.5 in Ibragimov and Linnink (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

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

We shall show that as $n \to \infty$,

$$(6.1) \qquad \sigma_n^2 \equiv Var(S_n) \sim n g_0 v^{-2};$$

$$(6.2) \qquad S_{n,1} := \sigma_n^{-1} \sum_{s=1}^{n} c_{n,s} \varepsilon_s \to_d N(0,1);$$

$$(6.3) \qquad S_{n,2} := \sigma_n^{-1} \sum_{s=-\infty}^{0} c_{n,s} \varepsilon_s \to_P 0,$$

which proves (2.11).

Observe that

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

where

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

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

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

Then (6.1) follows if

$$\int_{-\pi}^{\pi} |f(\lambda) - f(0)||D_n(\lambda)|^2 d\lambda = o(nv^{-2}). \tag{6.4}$$

By (2.3)

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

since

$$|f^*(\lambda) - f^*(0)| = |(v^2 + 2\rho(1 - \cos(\lambda))^{-1} - v^{-2}| \le C\lambda^2 f^*(\lambda)v^{-2},$$

and, by (2.7), $|g(\lambda) - g(0)| \le C\lambda^2$. Since $|D_n(\lambda)|^2\lambda^2 \le C$, the left hand side of (6.4) is bounded by

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

because $nv \to \infty$ by (2.10), and using the bound (2.5) for f^*.

Since the ε_t are i.i.d. variables with zero mean and unit variance, to prove (6.2) it suffices to check validity of Lindeberg condition, i.e. to show that for any $\delta > 0$,

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

First we show that for all $s = 1, .., n$,

$$|c_{n,s}| = o(\sigma_n). \tag{6.5}$$

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

$$c_{n,s} = \sum_{t=\max(1,s)}^{n} \psi_{t-s} = (2\pi)^{-1}\int_{-\pi}^{\pi}\sum_{t=s}^{n} e^{i(t-s)x}\Psi(x)dx. \tag{6.6}$$

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

$$|c_{n,s}| \le Cv^{-1}\int_{-\pi}^{\pi}|\sum_{t=s}^{n} e^{i(t-s)x}|dx \le Cv^{-1}\int_{-\pi}^{\pi}(|D_{n-s}(x)|+1)dx \le Cv^{-1}\log n$$

for all $s = 1, \ldots, n$ using (5.7). Since $\sigma_n \sim Cn^{1/2}v^{-1}$, this proves (6.5).

Fix $K > 0$. Then $\theta_K := E[\varepsilon_s^2 1_{|\varepsilon_s|>K}] \to 0$, as $K \to \infty$. Therefore, in view of (6.5),

$$E[c_{n,s}^2\varepsilon_s^2 1_{|c_{n,s}\varepsilon_s|\ge\delta\sigma_n}] \le (\delta\sigma_n)^{-2}E[c_{n,s}^4\varepsilon_s^4 1_{|\varepsilon_s|\le K}] + c_{n,s}^2 E[\varepsilon_s^2 1_{|\varepsilon_s|>K}] = c_{n,s}^2(o(1)+\theta_K).$$

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

$$i_n \le \sigma_n^{-2}\sum_{s=1}^{n} c_{n,s}^2(o(1)+\theta_K) = o(1)+\theta_K \to 0, \quad n, K \to \infty,$$

which completes proof of (6.2).

To show (6.3), note that

$$ES_{n,2}^2 = \sigma_n^{-2}\sum_{s=-\infty}^{0} c_{n,s}^2. \tag{6.7}$$

Note that for $s \le 0$, by (6.6),

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

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

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

Observe that

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

We have that

$$|\Psi_\rho(x)-\Psi_\rho(0)| \le |(1-\rho e^{-ix})^{-1}-(1-\rho)^{-1}| \le 2|x|v^{-1}|\Psi_\rho(x)|,$$

whereas by (2.6),

$$|\Psi_b(x) - \Psi_b(0)| \le C|x|.$$

Then

$$|\Psi(x) - \Psi(0)| \le |\Psi_\rho(x) - \Psi_\rho(0)||\Psi_b(x)| + |\Psi_\rho(0)||\Psi_b(x) - \Psi_b(0)| \le C(|x|v^{-1}|\Psi_\rho(x)| + |x|v^{-1}).$$

So,

$$ES_{n,2}^2 \le C\sigma_n^{-2} \int_{-\pi}^{\pi} |D_n(x)|^2 (|x|v^{-1}|\Psi_\rho(x)| + |x|v^{-1})^2 dx$$

$$\le C\sigma_n^{-2} \int_{-\pi}^{\pi} (v^{-2}|\Psi_\rho(x)|^2 + v^{-2}) dx \le C\sigma_n^{-2} v^{-3} \le C\frac{v^2}{nv^3} = \frac{1}{nv} \to 0$$

by assumption (2.10), which proves (6.3). ∎

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

$$s_1 = (2\pi)^{-1} \int_{-\pi}^{\pi} f^*(\lambda) g_0 d\lambda, \quad s_2 = (2\pi)^{-1} \int_{-\pi}^{\pi} f^*(\lambda)(g(\lambda) - g_0) d\lambda =: g_1 + g_2.$$

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),

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

for any v. Since for each λ, $f^*(\lambda) \to (2(1-\cos(\lambda)))^{-1}$ as $v \to 0$, then by theorem of dominating convergence,

$$s_2 \to \Gamma_0 = (2\pi)^{-1} \int_{-\pi}^{\pi} \frac{g(\lambda) - g_0}{2(1-\cos(\lambda))} d\lambda, \quad v \to 0 \tag{6.8}$$

which completes the proof of (2.12).

To prove (2.13), write

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

Since γ_0 satisfies (2.12), and $|\cos(k\lambda) - 1| \le C\lambda^2$, then by the same argument as used in (6.8), it follows

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

to prove (2.13).

Finally, by (2.13) and (2.12),

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{\gamma_0 - \Gamma_k + o(1)}{\gamma_0} = 1 - \frac{\Gamma_k + o(1)}{\gamma_0} = 1 - \frac{\Gamma_k + o(1)}{(2v)^{-1}(1 + o(1))} = 1 - 2v\Gamma_k + o(1).$$

∎

Proof of Theorem 2.2. *Proof of* (2.15)-(2.16). Write

$$\hat{\gamma}_k \equiv T_{n,X} = \int_{\pi}^{\pi} \cos(k\lambda) I(\lambda) d\lambda. \tag{6.9}$$

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

$$|\hat{\gamma}_k - \gamma_k| \le C(\frac{1}{nv^2} + \frac{1}{\sqrt{nv^3}}) \le C\frac{1}{\sqrt{nv^3}}$$

since $1/(nv^2) \le C/\sqrt{nv^3}$ under (2.10). Next, by (2.13), $1/\gamma_k \le Cv$, and therefore

$$\hat{\gamma}_k = \gamma_k(1 + O_P(\frac{1}{\gamma_k\sqrt{nv^3}})) = \gamma_k(1 + O_P(\frac{1}{\sqrt{nv}}))$$

proving (2.15).

To prove (2.16), we shall show that assumptions of (ii) (c2) of Theorem 3.2 are satisfied. Set $h_n(\lambda) = \cos(k\lambda) f(\lambda)$. Set

$$B_n = \int_{-\pi}^{\pi} h_n^2(\lambda) d\lambda, \quad J_n(u) = \int_{-\pi}^{\pi} |h_n(x+u) - h_n(x)|^2 dx. \tag{6.10}$$

For simplicity, we write below $v = v_n$. By Lemma 6.1 (ii) below,

$$B_n \sim \frac{1}{8\pi} g_0^2 v^{-3} \tag{6.11}$$

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

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

in view of (6.11), since $vn \to \infty$. Next

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

because of (6.11). Finally, since $k_n = 1$

$$\frac{k_n/v^2}{\sqrt{nB_n}} \sim C\frac{1/v^2}{\sqrt{nv^{-3}}} = C\frac{1}{\sqrt{nv}} \to 0$$

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

$$\sqrt{\frac{n}{4\pi B_n}}\Big(\hat{\gamma}_k-\gamma_k\Big)\xrightarrow{d}N(0,1),$$

where $\sqrt{\frac{n}{4\pi B_n}}=\sqrt{\frac{2nv^3}{g_0^2}}$, proving (2.16).

Proof of (2.17). We have

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

where

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

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

(6.12) $$|\eta_n(\lambda)|\le C|\cos(k\lambda)-1|+|1-\rho_k|\le C(\lambda^2+v)\le C$$

by (2.14). Then by (3.11) of Theorem 3.1,

(6.13) $$T_{n,X}=T_{n,\varepsilon}-E[T_{n,\varepsilon}]+r_k,\quad E|r_k|\le C(nv^2)^{-1}$$

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 5.2 it follows that

$$\mathrm{Var}(T_{n,\varepsilon})=2(2\pi n)^{-2}||E_n||^2\le CB_n/n,\quad B_n=\int_{-\pi}^{\pi}\eta_n^2(\lambda)f^2(\lambda)d\lambda.$$

This implies

$$|EJ_n|\le C(\frac{1}{nv^2}+\sqrt{\frac{B_n}{n}}).$$

Estimating $\eta_n(\lambda)$ by (6.12), and noting that for small v, (2.5)implies $f(\lambda)\le C(v^2+\lambda^2)^{-1}$, we obtain

$$B_n\le C\int_{-\pi}^{\pi}\frac{(\lambda^2+v)^2}{(\lambda^2+v^2)^2}d\lambda\le Cv^{-1}\int_{-\infty}^{\infty}\frac{(\lambda^2+1)^2}{(\lambda^2+1)^2}d\lambda\le Cv^{-1}.$$

Thus

$$|EJ_n|\le C(\frac{1}{nv^2}+\frac{1}{\sqrt{nv}}),\quad J_n=O_P(\frac{1}{nv^2}+\frac{1}{\sqrt{nv}}).$$

We show below that

(6.14) $$\hat{\gamma}_0=\frac{g_0}{2v}(1+o_P(1))$$

as $v \to 0$, which implies

$$\hat{\rho}_k - \rho_k = O_P(\frac{1}{nv^2} + \frac{1}{v\sqrt{nv}})v = O_P(\frac{1}{nv} + \sqrt{\frac{v}{n}}),$$

to prove (2.17).

In addition we show

$$\sqrt{\frac{2nv^3}{(1-\rho_k)^2 g_0^2}}(T_{n,\varepsilon} - E[T_{n,\varepsilon}]) \to_d N(0,1) \tag{6.15}$$

which together with (6.14) and (6.13) implies (2.19), since $(nv_n^2)^{-1} = o(1/\sqrt{nv_n})$ when $nv_n^3 \to \infty$.

Proof of (6.14). 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 (2.12) of Lemma 2.1. Using the matrix E_n with entries defined as in (3.14), we can write

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

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

$$\mathrm{Var}(Q_{n,1}) \le Cn^{-2}||E_n||^2 \le Cn^{-1}\int_{-\pi}^{\pi} f^2(x)dx \le C(nv^3)^{-1} = o(v^{-2})$$

by Lemma 5.2 and (6.18), using assumption $nv \to \infty$. 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

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

Therefore $Q_{n,2} = o_P(v^{-1})$ which proves (6.14).

Proof of (6.15). The proof of (6.15) will be based on part (ii) of Theorem 3.2 and assumption (c1). For that we need to evaluate quantities B_n and $J_n(u)$ in (6.10).

Note that

$$\begin{aligned} h_n(x) &= (\cos(k\lambda) - \rho_k)f(x) = (\cos(kx) - 1)f(x) + (1-\rho_k)f(x) \\ &= O(1) + (1-\rho_k)f(x) \end{aligned} \tag{6.16}$$

since

$$|(\cos(kx) - 1)f(x)| \le Cx^2 f^*(x) \le C, \quad |x| \le \pi.$$

By (2.14), $1-\rho_k \sim 2v\Gamma_k$, $v \to 0$. Hence

$$h_n^2(x) = (O(1) + (1-\rho_k)f(\lambda))^2 = O(1) + O(v)f(\lambda) + (1-\rho_k)^2 f^2(\lambda)$$

and

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

By (6.11), $\int_{-\pi}^{\pi} f^2(x)dx \sim \frac{1}{8\pi} g_0^2 v^{-3}$ which implies

$$B_n \sim (1-\rho_k)^2 \frac{1}{8\pi} g_0^2 v^{-3}. \tag{6.17}$$

To estimate $J_n(u)$, note that by (6.16)

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

Hence

$$J_n(u) \le C\Big(1 + (1-\rho_k)^2 \int_{-\pi}^{\pi} |f(x+u) - f(x)|^2 dx\Big) = C + (1-\rho_k^2)O(u^2 v^{-5})$$

in view of (6.19). So for $|u| \le K/n$, where K is a fixed constant,

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

because of (6.17) and (2.14). Hence h_n satisfies Assumption 3.3.

It remains to show (3.27). By (6.12) and (2.5),

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

Then

$$\frac{k_n^*}{\sqrt{nB_n}} \le C\frac{1/v}{\sqrt{nv^{-1}}} = C\frac{1}{\sqrt{nv}} \to 0.$$

Therefore, by (3.27),

$$\sqrt{\frac{n}{4\pi B_n}}\Big(T_{n,\varepsilon} - E[T_{n,\varepsilon}]\Big) \xrightarrow{d} N(0,1),$$

where $\sqrt{\frac{n}{4\pi B_n}} \sim \sqrt{\frac{2nv^3}{(1-\rho_k)^2 g_0^2}} \sim c\sqrt{nv}$ which proves (6.15). ∎

Lemma 6.1 *(i) Under assumption* (2.3) *and* (3.27), *as* $v \to 0$,

$$\int_{-\pi}^{\pi} f^2(x)dx \sim \frac{1}{8\pi} g_0^2 v^{-3} \tag{6.18}$$

and

$$V(u) := \int_{-\pi}^{\pi} |f(x+u) - f(x)|^2 dx \le Cu^2 v^{-5} \tag{6.19}$$

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

(ii) Estimates (6.18) *and* (6.19) *remain valid when* $f(x)$ *is replaced by* $\cos(kx)f(x)$.

Proof of Lemma 6.1. First we show (6.18). Note that $f = (2\pi)^{-1} f^* g$ where $|g(x) - g_0| \leq Cx^2$, and $x^2 f^*(x) \leq C$. Hence,

$$f(x) = (2\pi)^{-1} f^*(x) g(x) = (2\pi)^{-1} f^*(x) g_0 + O(1),$$

and

$$f^2(x) = (2\pi)^{-2} g_0^2 f^*(x)^2 + O(1) f^*(x) + O(1) = (2\pi)^{-2} g_0^2 f^*(x)^2 + O(v^{-2}).$$

Observe that, as $v \to 0$,

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

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

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

to prove (6.18)

To show (6.19), note that

$$\begin{aligned}|f^*(x+u) - f^*(x)| &\leq 2|\cos(x+u) - \cos(x)| f^*(x+u) f^*(x) \\ &\leq Cu(|x| + |x+u|) f^*(x+u) f^*(x)\end{aligned}$$

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

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

Hence

$$V(u) \leq Cu^2 v^{-2} \int_{-\pi}^{\pi} (f^*(x) + f^*(x+u))^2 dx \leq Cu^2 v^{-2} \int_{-\pi}^{\pi} f^*(x)^2 dx \leq Cu^2 v^{-5}$$

by (6.18), which proves (6.19).

In case (ii), the estimates (6.18)-(6.19) follow using the same argument. ∎

PROOF OF THEOREM 2.3. By (2.6) and (2.12) we have that

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

We shall show that

$$t_n \equiv \int_{-\pi}^{\pi} \sqrt{|x|} I_n(x)dx = \frac{g_0}{\sqrt{2v}} + O_P(1 + \frac{1}{nv^2} + \frac{1}{v\sqrt{n}}) \tag{6.20}$$

which implies (2.21). By (3.13) of Theorem 3.1,

$$t_n = \int_{-\pi}^{\pi} \sqrt{|x|} f(x)dx + O_P(\frac{1}{nv^2} + \sqrt{\frac{B_n}{n}})$$

where, using (2.5),

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

To prove (6.20) it remains to show that

$$i_n := \int_{-\pi}^{\pi} \sqrt{|x|} f(x)dx = (2v)^{-1/2} g_0 + O(1). \tag{6.21}$$

Write

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

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

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

Observe that that

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

since

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

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

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

which together with estimates above implies (6.21).
∎

References

Bhansali, R.J., Giraitis, L. and 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.

Cavanagh, C. (1985), "Roots Local to Unity", manuscript, Department of Economics, Harvard University.

Chan, N. H. and C. Z. Wei (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. and Heyde, C. C. (1972). On limit theorems for quadratic functions of discrete time series. *The Annals of Mathematical Statistics* **43**, 2058-2066.

Hannan, E. J. (1973). The asymptotic theory of linear time series models. *J. Appl. Prob.* **10**, 130-145.

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. and Linnik, Yu. V. *Independent and Stationary Sequences of Random Variables.* Wolters–Noordhoff, Groningen, 1971.

Jeffrey, A. Handbook of mathematical formulas and integrals. Academic Press, London, 1995.

Magdalinos, T. and P. C. B. Phillips (2008). "Limit theory for cointegrated systems with moderately integrated and moderately explosive regressors", Working paper, Yale university.

Phillips, P. C. B. (1987) Towards an unified asymptotic theory for autoregression. *Biometrika* **74**, 535–547.

Phillips, P. C. B. and Magdalinos, T. (2007a) Limit theory for moderate deviations from a unit root. *Journal of Econometrics* **136**, 115 – 130.

Phillips, P. C. B., and T. Magdalinos (2007b), "Limit Theory for Moderate Deviations from a Unit Root Under Weak Dependence," in G. D. A. Phillips and E. Tzavalis (Eds.) *The Refinement of Econometric Estimation and Test Procedures: Finite Sample and Asymptotic Analysis.* Cambridge: Cambridge University Press.

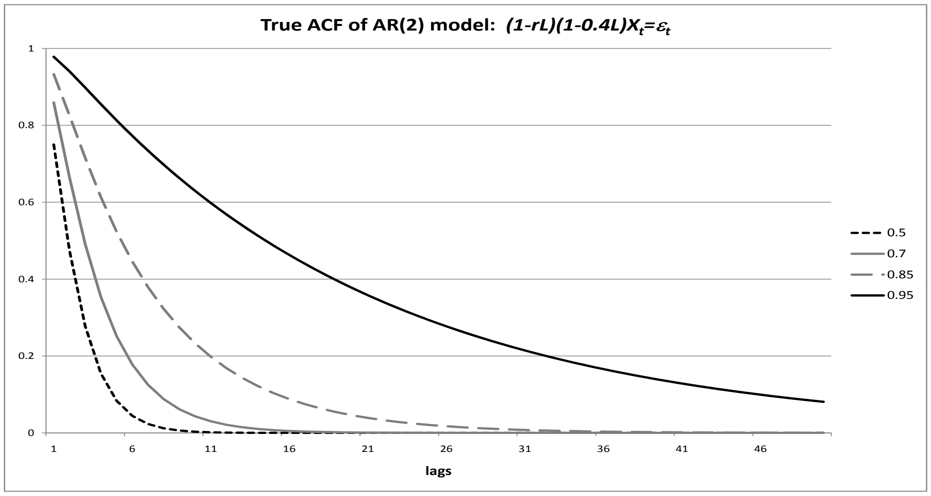

Figure 1: ACF ρ_k of AR(2) model with $r = 0.5, 0.7, 0.85, 0.95$, $n = 125$

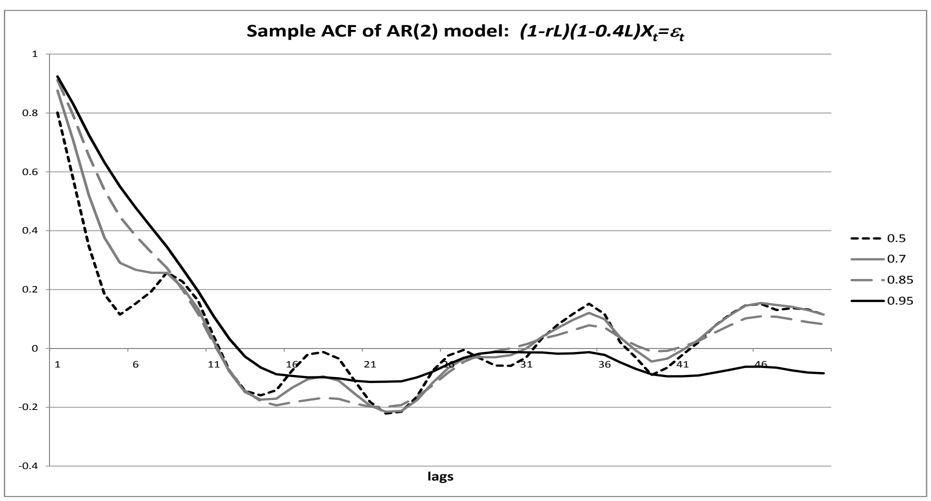

Figure 2: Example of realizations of Sample ACF $\hat{\rho}_k$ of AR(2) model with $r = 0.5, 0.7, 0.85, 0.95$, $n = 125$

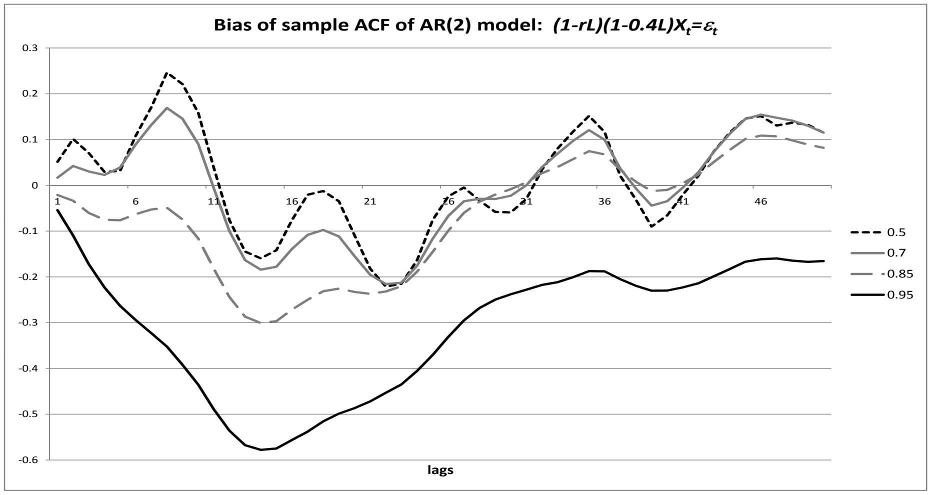

Figure 3: Bias of Sample ACF of AR(2) model with $r = 0.5, 0.7, 0.85, 0.95$, $n = 125$

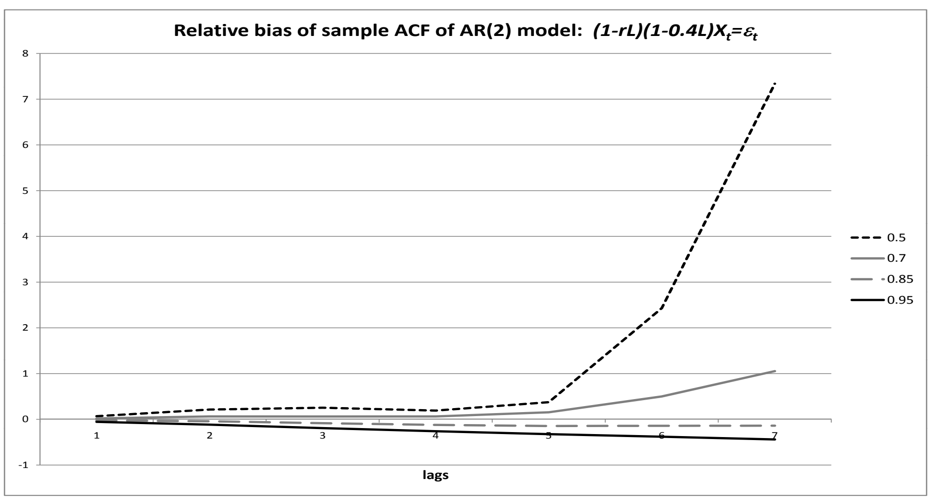

Figure 4: Relative bias $(\hat{\rho}_k - \rho_k)/\rho_k$ of Sample ACF of AR(2) model with $r = 0.5, 0.7, 0.85, 0.95$, $n = 125$

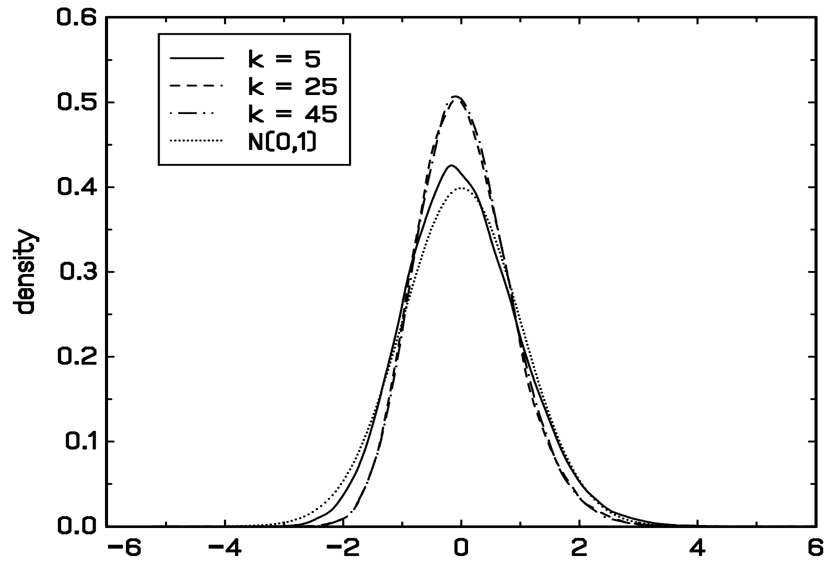

Figure 5: Densities of $\hat{t}_n(k) : k = 5, 25, 45$ versus the standard normal for $r = 0.8$ and $n = 2000$.

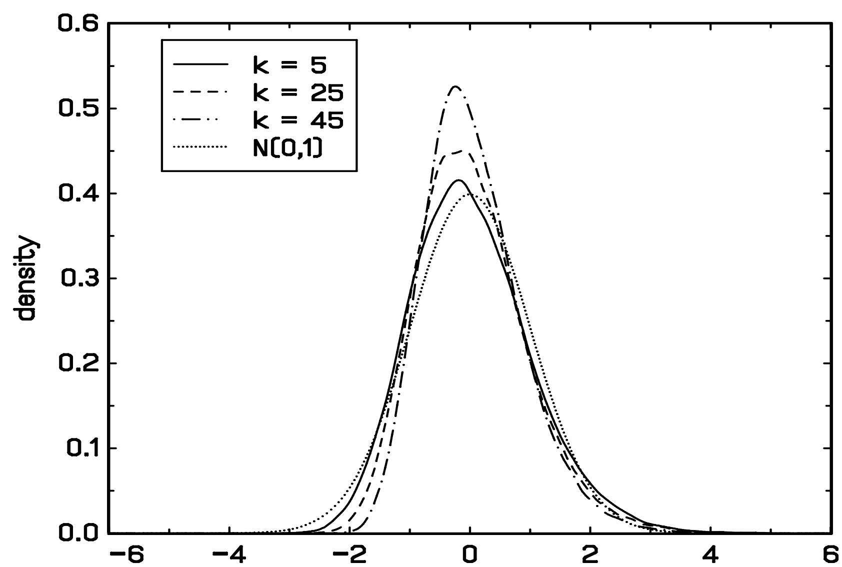

Figure 6: Densities of $\hat{t}_n(k) : k = 5, 25, 45$ versus the standard normal for $r = 0.95$ and $n = 2000$.