A Parametric Bootstrap Test for Cycles*

by

Violetta Dalla
and
Javier Hidalgo
London School of Economics and Political Science

Contents:

The Suntory Centre
Suntory and Toyota International Centres for
Economics and Related Disciplines
London School of Economics and Political Science
Houghton Street
London WC2A 2AE
Tel.: 020 7955 6679

Discussion Paper
No.EM/05/486
February 2005

---

* This article is based on research funded by the Economic and Social Research Council (ESRC), reference number: 4000239936. Also, we thank three referees for their valuable comments which have led to a substantial improvement of the paper.

Abstract

Abstract. The paper proposes a simple test for the hypothesis of strong cycles and as a by-product a test for weak dependence for linear processes. We show that the limit distribution of the test is the maximum of a (semi)Gaussian process $\mathcal{G}(\tau)$, $\tau \in [0; 1]$. Because the covariance structure of $\mathcal{G}(\tau)$ is a complicated function of τ and model dependent, to obtain the critical values (if possible) of $\max_{\tau\in[0;1]} \mathcal{G}(\tau)$ may be difficult. For this reason we propose a bootstrap scheme in the frequency domain to circumvent the problem of obtaining (asymptotically) valid critical values. The proposed bootstrap can be regarded as an alternative procedure to existing bootstrap methods in the time domain such as the residual-based bootstrap. Finally, we illustrate the performance of the bootstrap test by a small Monte Carlo experiment and an empirical example.

Keywords: Cyclical data; strong and weak dependence; spectral density functions; Whittle estimator; bootstrap algorithms.

JEL No.: C15, C22.

Contact details:
London School of Economics, Houghton Street, London WC2A 2AE, United Kingdom. Email: v.dalla@lse.ac.uk and f.j.hidalgo@lse.ac.uk.

1. INTRODUCTION

During the last decade or so, we have seen an increasing interest in the so-called "strong dependent" data. The main motivation for this interest comes from the apparent observation that in many areas, such as hydrology or macroeconomic time series, the data appear to have a cyclic component, although they are not periodic. This is manifested in a number of series whose spectral density estimates peaked sharply around some frequency, indicating a cyclical component, or at say seasonal frequencies. However, when first (seasonal, say) differences are taken, the spectrum tends to exhibit a trough, indicating that the data has been overdifferenced.

One model capable of generating strong dependence in the data is the Gegenbauer model, proposed by Andel (1986) and explored in Gray et al. (1989), which was defined as

$$\left(1-e^{i\lambda^0}L\right)^{d_0}\left(1-e^{-i\lambda^0}L\right)^{d_0}x_t=\varepsilon_t, \tag{1}$$

where L is the backshift operator, $-1/2<d_0<1/2$ if $\lambda^0\neq 0,\pi$ and $-1/4<d_0<1/4$ if $\lambda^0=\{0,\pi\}$. The model (1) was extended by Gray et al. (1989) to the $GARMA$ model where the sequence $\{\varepsilon_t\}_{t\in\mathbb{Z}}$ follows an Autoregressive Moving Average ($ARMA$) model. More generally, model (1) can be extended to allow $\{\varepsilon_t\}_{t\in\mathbb{Z}}$ to follow a stationary $MA(\infty)$ so that $\{x_t\}_{t\in\mathbb{Z}}$ will be characterized by having a spectral density function defined as

$$f(\lambda)=\frac{\sigma_0^2}{2\pi}\left|1-e^{i\left(\lambda-\lambda^0\right)}\right|^{-2d_0}\left|1-e^{i\left(\lambda+\lambda^0\right)}\right|^{-2d_0}h(\lambda)\qquad -\pi<\lambda\leq\pi, \tag{2}$$

where $\sigma_0^2>0$ is the variance of the innovations of the sequence $\{\varepsilon_t\}_{t\in\mathbb{Z}}$ and $h(\lambda)$ is an even, continuous and bounded away from zero function such that $\sigma_0^2h(\lambda)/(2\pi)$ is the spectral density function of $\{\varepsilon_t\}_{t\in\mathbb{Z}}$. When $\lambda^0=0$ and $\{\varepsilon_t\}_{t\in\mathbb{Z}}$ follows an $ARMA(p,q)$ process, (1) becomes the more familiar $FARIMA(p,2d_0,q)$ model, originated by Adenstedt (1974) and further explored and examined by Granger and Joyeux (1980) and Hosking (1981). The coefficient d_0 is the fractional differencing coefficient. One can also sometimes find reference to the coefficient α_0, defined as $\alpha_0=2d_0$, which we shall refer to as the memory parameter. One feature of models such as that given in (2) is that $f(\lambda)$ possesses a pole at λ^0, that is

$$f(\lambda)\sim C\left|\lambda-\lambda^0\right|^{-\alpha_0}\quad\text{as}\quad\lambda\to\lambda^0, \tag{3}$$

where $C\in(0,\infty)$, $0\leq\alpha_0<1$ and "$\sim$" means that the ratio of the left- and right-hand sides tends to 1. The value α_0 given in (3) can be regarded as determining the (local) shape of the spectral density function around λ^0, which can discriminate among different time series. In addition, α_0 gives an indication and summarizes the dependence structure of $\{x_t\}_{t\in\mathbb{Z}}$ in the long run.

A feature of model (2) is that it possesses a stronger and more persistent cyclical pattern than $ARMA$ models, e.g. the $AR(2)$ process

$$\left(1-a_1L-a_2L^2\right)x_t=\varepsilon_t \tag{4}$$

when the roots of the polynomial $\left(1-a_1L-a_2L^2\right)$ are complex, with λ^0 identified as the $arc\cos\left(\frac{a_1}{2\sqrt{-a_2}}\right)$.

So, models (2)/(3) and (4) may have some features similar to those observed with real data. However, the latter two models constitute different descriptions of cyclic behaviour within the stationary class. Therefore, when the practitioner is faced with the problem of choosing between models like (2) and (4), it would be useful to have a testing procedure to discriminate between the aforementioned different cyclical behaviours.

The main objective of the paper is thus to test whether or not the data exhibits strong cyclical components and also to describe a bootstrap method in the frequency domain as an alternative to those based on the time domain. To that end, we describe and examine two different tests. The first one is a Wald (W) type test, whereas the second one is based on the Lagrange Multiplier (LM) principle, which may be computationally more attractive than the former. The tests are based on whether the supremum of a sequence of random variables is significantly greater than zero. In particular, see Section 2, the W and the LM types of tests are based on the supremum of a sequence of estimators of the memory parameter α_0 and the score function respectively when it is believed that the pole of the spectrum is at some particular frequency $\lambda^0 \in [0, \pi]$. One feature of the hypothesis testing is that the null lies at the boundary of the parameter space.

Our tests, as mentioned above, are based on the supremum of a sequence of random variables. It is well known that the rate of convergence of the finite sample distribution to the asymptotic one is very slow, see Hall (1979) for a related statistic. In particular, Hall showed that the rate is logarithmic. In addition, as we show in Theorem 2.1 below, the asymptotic distribution of our tests is nonstandard, so that bootstrap algorithms will allow us to make valid inferences. This motivates us to employ a bootstrap approach to our hypothesis testing. (See also the comments at the end of Section 2 for other motivations to perform a bootstrap algorithm in our context.)

It should be noticed that, as a by-product, our tests provide a way to test for weak dependence against strong dependence in the class of linear models. The concept of strong dependence, sometimes known as long range or long memory dependence, refers to time series data that have an autocovariance function, $\operatorname{cov}(x_t, x_{t+j}) = \gamma(j)$, which is not absolutely summable. However, it should be mentioned that strong dependence refers not only to second moments, although for Gaussian processes is synonymous. This type of processes makes the probabilistic properties of the data and the asymptotic distribution of some relevant statistics/estimates (possibly implicit ones) very different from those of usual "weakly dependent"/mixing processes such as $ARMA$ models, or their properties need to be examined on a case by case basis. The concept of weak dependence draws similarities with that in Doukhan (1994), see also Nze et al. (2002). They define weak dependence as a measure between the covariance between functions of the past and the future. An earlier and similar concept was introduced by McLeish (1975), known as Mixingale (or general near epoch), which measures how fast conditional moments converge to unconditional ones. From a statistical point of view, to know if the data is near epoch dependent with size greater than 1/2 (see, McLeish, 1975) can be important as many statistical results rely heavily on the latter type of dependence. The latter concept of dependence has been argued to be what is really needed for examining the properties of, say, estimates, instead of stronger concepts such as strong-mixing (see Rosenblatt, 1956) or β-mixing (see, Volkonskii and Rozanov, 1956), which are based on the variation norm between the joint probability function and the product of their marginals.

Two points have to be raised concerning our tests. The first one is that the tests are of parametric nature, as it is evident from Condition $C1$ of Section 2. Our tests differ from that of Lobato and Robinson (1998) in two respects. First, contrary to them, we allow λ^0 to be unknown. Secondly, our tests are parametric, so that they are more efficient than the one explored by the former authors, especially when dealing with small or moderate sample sizes. On the other hand, the Lobato and Robinson's (1998) LM test only requires a local knowledge of $f(\lambda)$ around λ^0, so that in this sense, their test can be more attractive than ours since less assumptions

are imposed on the shape of the spectral density. However, in this paper we are taking the view that the practitioner is confident about a correct parameterization of $f(\lambda)$. That is, the data follows a $GARMA$ process, where the order of the $ARMA$ polynomial can be chosen by a criterion function such as the AIC or BIC, see Beran et al. (1998) for their justification in our context. In addition, we should mention that the identification of λ^0 is not always possible, like in the case of a business cycle, and therefore our tests are in that sense more general. Finally, we can cite the test given in Robinson (1994), although his null hypothesis is that $d_0 = 1$, whereas in our case is $d_0 = 0$, apart from the fact that as in Lobato and Robinson's test, it is assumed that λ^0 is known a priori.

The second point is that we restrict our analysis to linear processes. One of the motivations to constrain to linear models is because conditions under which the data satisfy the different concepts of dependence/mixing are quite well understood, depending basically on how fast the coefficients of the MA representation of the $\{x_t\}_{t\in\mathbb{Z}}$ converge to zero. See for instance Gorodetskii (1977), or Pham and Tran (1985) among others. Nevertheless, we are aware that existing results are available for nonlinear (Markovian) models such as nonlinear $AR(p)$, bilinear or $ARCH$ models, see Doukhan (1994) for a review. In particular, given that $\{x_t\}_{t\in\mathbb{Z}}$ follows a linear model as that in Condition $C5$ of Section 2, under some smoothness conditions on the probability density function of the innovations $\{\varepsilon_t\}_{t\in\mathbb{Z}}$, Gorodetskii (1977) and Pham and Tran (1985) gave conditions on the rate of convergence of b_j to zero which are not satisfied for strong dependent processes, suggesting that these processes are neither strong mixing nor β-mixing. Nevertheless, we can see that, they are mixingale, also weak dependent in the sense of Doukhan, although with a size smaller than $1/2$, which is a minimum requirement to guarantee the standard statistical results.

Another motivation to focus on linear models is because they still represent a fairly broad class of models very much employed by practitioners and because the relationship between the rate of decay to zero of b_k and the smoothness of the spectral density function is well established. For instance, for Gaussian processes (see, Ibragimov 1965, 1970), to be strong-mixing it is required that

$$\lim_{\lambda\to\lambda^0} \frac{\log f(\lambda)}{\log\left|\lambda-\lambda^0\right|} = 0, \tag{5}$$

which rules out a spectral density function satisfying (3).

The remainder of the paper is as follows. In the next section, we describe the hypothesis testing and introduce a W and LM type of tests for the null hypothesis of $\alpha_0 = 0$ for all $\lambda^0 \in [0,\pi]$. Because the limit distribution of the tests is nonstandard and model dependent, Section 3 describes and proposes a valid bootstrap scheme which can be regarded as a frequency domain counterpart to the residual-based bootstrap. The proposed bootstrap, contrary to the latter, is far easier to compute in models for which it is difficult to obtain the coefficients of the Moving Average or Autoregressive representation of $\{x_t\}_{t\in\mathbb{Z}}$, as is the case with the Bloomfield's (1973) exponential model. Moreover, our bootstrap algorithm is based on Efron's (1979) naive bootstrap. The proofs of the results in Sections 2 and 3 are given in Section 5, which makes use of a series of Lemmas in Section 6. A small Monte Carlo experiment to examine the performance of our test in small samples is given in Section 4, together with an application of our test to the monthly seasonal adjusted *Industrial Production Index* for the USA. Finally, Section 7 concludes and gives a modification of the bootstrap approach described in Section 3.

2. THE TESTS AND REGULARITY CONDITIONS

Consider $\{x_t\}_{t\in\mathbb{Z}}$ a covariance stationary linear process having mean that it is (without loss of generality) zero and absolute continuous spectral distribution function, so that it has a spectral density function, denoted by $f(\lambda)$, defined from the relation

$$\gamma(j) = E(x_0 x_j) = \int_{-\pi}^{\pi} f(\lambda) e^{ij\lambda} d\lambda, \quad j = 0, \pm 1, \pm 2, ... \tag{6}$$

We shall first describe how our hypothesis testing can be written in terms of some parameters. Suppose that the spectral density f is positive, continuous and known up to a finite set of parameters $\left(\psi_0', \sigma_0^2\right)'$. That is,

$$f(\lambda) = \frac{\sigma_0^2}{2\pi} h^*(\lambda; \psi_0), \tag{7}$$

where $\sigma_0^2 > 0$, $\psi_0 \in \Psi$, a compact set in $\mathbb{R}^{p+1}$, and $h^*(\lambda; \psi_0)$ is a known even function. Under the null hypothesis of no strong cycles, we have that f is a positive and continuous function for all $\lambda \in [0, \pi]$. This means that there exists $0 < K < \infty$ such that

$$K^{-1} < h^*(\lambda; \psi_0) < K \text{ for all } \lambda \in [0, \pi].$$

Hence, our hypothesis of interest can be formulated as

$$H_0 : K^{-1} < h^*(\lambda; \psi_0) < K \text{ for all } \lambda \in [0, \pi]; \tag{8}$$

while the alternative hypothesis becomes

$H_1 : \exists \lambda^0 \in [0, \pi]$ and $0 < \alpha_0 < 1$, such that

$$h^*(\lambda; \psi_0) = g\left(\lambda; \alpha_0, \lambda^0\right) h(\lambda; \theta_0), \tag{9}$$

where $\psi_0 = \left(\alpha_0, \theta_0'\right)'$ and

$$\begin{aligned} K^{-1} &< h(\lambda; \theta_0) < K \text{ for all } \lambda \in [0, \pi] \\ g^{-1}\left(\lambda; \alpha_0, \lambda^0\right) &\sim C\left|\lambda - \lambda^0\right|^{\alpha_0} \to 0 \text{ as } \lambda \to \lambda^0, \end{aligned}$$

where $h(\lambda; \theta_0)$ is a continuous function. In what follows K denotes a positive finite constant.

Observe that since we have assumed that $Ex_t^2 < \infty$, we have that $g\left(\lambda; \alpha_0, \lambda^0\right)$ is an integrable function so that $\alpha_0 < 1$. As we have argued in the introduction, because $h(\lambda; \theta_0)$ is continuous, the case $\alpha_0 = 0$ will refer to weak dependence, whereas the case $0 < \alpha_0 < 1$ to strong dependence.

Before we formally describe the tests, we introduce the following regularity conditions:

C1: The process $\{x_t\}_{t\in\mathbb{Z}}$ has an absolutely continuous spectral distribution function, its spectral density, $f(\lambda) = \frac{\sigma_0^2}{2\pi} f\left(\lambda; \alpha_0, \theta_0, \lambda^0\right)$, being of the form

$$f\left(\lambda; \alpha, \theta, \lambda^0\right) = g\left(\lambda; \alpha, \lambda^0\right) h(\lambda; \theta), \quad \lambda \in (-\pi, \pi], \tag{10}$$

with

$$g\left(\lambda; \alpha, \lambda^0\right) = \left|4 \sin\left(\frac{\lambda + \lambda^0}{2}\right) \sin\left(\frac{\lambda - \lambda^0}{2}\right)\right|^{-\alpha}, \tag{11}$$

where $\theta \in \Theta$, a compact set in $\mathbb{R}^p$, $\alpha \in [0, 1)$ if $\lambda^0 \in (0, \pi)$ and $\alpha \in [0, 1/2)$ if $\lambda^0 = \{0, \pi\}$. For $-\pi < \lambda \leq \pi$, $h(\lambda; \theta)$ is even in λ and bounded away from zero, and the derivatives $\nabla_\theta h(\lambda; \theta)$, $\nabla_\lambda h(\lambda; \theta)$, $\nabla_\theta \nabla_\lambda h(\lambda; \theta)$ and $\nabla_\theta \nabla_\theta' h(\lambda; \theta)$ are continuous.

Henceforth, $\psi = \left(\alpha, \theta'\right)'$ and σ^2 denote any admissible value of $\psi_0 = \left(\alpha_0, \theta_0'\right)'$ and σ_0^2 (the true value of the parameters) respectively.

C2: For all $\theta \in \Theta$, the function $h(\lambda;\theta)$ satisfies that

$$\int_{-\pi}^{\pi} \log\left(h(\lambda;\theta)\right) d\lambda = 0. \tag{12}$$

C3:

$$\inf_{(\alpha,\theta')' \in \Pi \times \Theta} \int_{-\pi}^{\pi} \frac{f\left(\lambda;\alpha_0,\theta_0,\lambda^0\right)}{f\left(\lambda;\alpha,\theta,\lambda^0\right)} d\lambda = 1,$$

where $\Pi = [0,1)$ when $\lambda^0 \in (0,\pi)$ and $\Pi = [0,1/2)$ if $\lambda^0 = \{0,\pi\}$, and the set

$$\left\{\lambda : f\left(\lambda;\alpha_0,\theta_0,\lambda^0\right) \neq f\left(\lambda;\alpha,\theta,\lambda^0\right)\right\} \text{ for } \left(\alpha_0,\theta_0'\right) \neq \left(\alpha,\theta'\right),$$

has positive Lebesgue measure. Also, the matrix

$$\Omega = \frac{1}{4\pi} \int_{-\pi}^{\pi} \nabla_\psi \log f\left(\lambda;\alpha_0,\theta_0,\lambda^0\right) \nabla_{\psi'} \log f\left(\lambda;\alpha_0,\theta_0,\lambda^0\right) d\lambda$$

is positive definite for all $\lambda^0 \in [0,\pi]$.

C4: θ_0 is an interior point of the compact set $\Theta \in \mathbb{R}^p$.

C5: $\{x_t\}_{t\in\mathbb{Z}}$ is a covariance stationary linear process defined as

$$x_t = \sum_{j=0}^{\infty} b_j \varepsilon_{t-j};\ \sum_{j=0}^{\infty} b_j^2 < \infty, \text{ with } b_0 = 1,$$

where $\{\varepsilon_t\}_{t\in\mathbb{Z}}$ is an ergodic process that satisfies (a) $E\left(\varepsilon_t \left|\mathcal{F}_{t-1}\right.\right) = 0$ a.s., (b) $E\left(\varepsilon_t^2 \left|\mathcal{F}_{t-1}\right.\right) = E\left(\varepsilon_t^2\right) = \sigma_0^2$ a.s., (c) $E\left(\left|\varepsilon_t\right|^\ell \left|\mathcal{F}_{t-1}\right.\right) = \mu_\ell < \infty$ for $\ell = 3,...,8$, a.s., where $\mathcal{F}_t$ is the σ-algebra of events generated by $\{\varepsilon_s, s \leq t\}$ and (d)

$$cum\left(\varepsilon_{t_1},\varepsilon_{t_2},\varepsilon_{t_3},\varepsilon_{t_4}\right) = \begin{cases} \kappa, & t_1 = t_2 = t_3 = t_4, \\ 0, & otherwise. \end{cases}$$

We now comment on our conditions. Condition $C1$ covers a wide range of models, including invertible $ARMA$ and Bloomfield (1973) ones, although it allows for models in h, whose autocorrelation coefficients decay to zero much slower than the previous two models. Of course under H_0, $f\left(\lambda;\alpha_0,\theta_0,\lambda^0\right) = h\left(\lambda;\theta_0\right)$. Condition $C2$ is standard and not very strong. In fact, because $\int_{-\pi}^{\pi} \log\left|2\sin\left(\left(\lambda \pm \lambda^0\right)/2\right)\right| d\lambda = 0$, condition (12) implies that σ_0^2 is the one-step mean square linear prediction error, see Hannan (1970, p.121-123). The first part of Condition $C3$ is employed to prove the (strong) consistency of the Whittle estimator defined in (13) below. See for instance Hannan (1973) or Brockwell and Davis (1991, Ch.10). The second part of Condition $C3$ is an identification condition. In Condition $C5$ the normalization $b_0 = 1$ is consistent with (12) and it is similar to others used elsewhere, see Hannan (1973). The finite eighth moments assumed in $C5$ are needed to show the tightness condition of some process indexed by $\tau \in [0,1]$, required to show Theorem 2.1 below. Finally, we should notice that, as in Giraitis et al. (2001), we could have changed Condition $C1$ to

C1': $\{x_t\}_{t\in\mathbb{Z}}$ has a spectral density function given by $f(\lambda) = \frac{\sigma_0^2}{2\pi} f\left(\lambda;\alpha_0,\theta_0,\lambda^0\right)$, where

$$f\left(\lambda;\alpha,\theta,\lambda^0\right) = \begin{cases} \left|\lambda - \lambda^0\right|^{-\alpha} h_1\left(\lambda;\alpha,\theta,\lambda^0\right) & 0 \leq \lambda \leq \pi \\ \left|\lambda + \lambda^0\right|^{-\alpha} h_1\left(\lambda;\alpha,\theta,\lambda^0\right) & -\pi \leq \lambda \leq 0 \end{cases}$$

such that for $-\pi < \lambda \leq \pi$, $\theta \in \Theta$, $\alpha \in [0,1)$, $h_1\left(\lambda;\alpha,\theta,\lambda^0\right)$ is even in λ and bounded and bounded away from zero, the derivatives $\nabla_\psi h_1\left(\lambda;\alpha,\theta,\lambda^0\right)$, $\nabla_\psi^2 h_1\left(\lambda;\alpha,\theta,\lambda^0\right)$ are continuous and bounded, $h_1\left(\lambda;\alpha,\theta,\lambda^0\right)$ and

$\nabla_{(\alpha,\theta')} h_1\left(\lambda;\alpha,\theta,\lambda^0\right)$ satisfy uniform Lipschitz conditions in λ^0 of order greater than $1/2$, while, for $0<|\lambda|<\pi$, $\lambda^0\in[0,\pi]$, $\nabla_\lambda h_1\left(\lambda;\alpha,\theta,\lambda^0\right)$ and $\nabla_\lambda \nabla_{(\alpha,\theta')} h_1\left(\lambda;\alpha,\theta,\lambda^0\right)$ are bounded.

For a motivation and comparison with $C1$ we refer to Giraitis et al. (2001).

Next, we describe the main ideas of the tests. As mentioned in the introduction, the first one is a W type test whereas the second is a LM test. In what follows, we shall abbreviate $\xi\left(\lambda_j,\cdot\right)$ by $\xi_j\left(\cdot\right)$ for a generic function ξ, with $\lambda_j=\left(2\pi j\right)/n$ for integer j and n denoting the sample size.

2.1. **Wald test for $\mathbf{H_0}$.**

Assuming that $f\left(\lambda\right)$ follows model $(10)-(11)$, suppose that for a given λ^0 we have an estimator of $\left(\alpha_0,\theta_0'\right)'$ and denote by λ_s the closest Fourier frequency to λ^0. For example, given a stretch of data $(x_1,...,x_n)$, we can use the Whittle estimator defined as

$$\left(\widehat{\alpha}_s,\widehat{\theta}'\right)'=\underset{(\alpha,\theta')'\in\Pi\times\Theta}{\arg\min}\, Q\left(\alpha,\theta,s\right);\quad \widehat{\sigma}^2=Q\left(\widehat{\alpha}_s,\widehat{\theta},s\right) \tag{13}$$

and

$$Q\left(\alpha,\theta,s\right)=\frac{2\pi}{\tilde{n}}\sum_{j=1}^{\tilde{n}}\frac{I_j}{g_j\left(\alpha,\lambda_s\right)h_j\left(\theta\right)}, \tag{14}$$

where, henceforth $\tilde{n}=[n/2]$ with $[z]$ denoting the integer part of z, and

$$I_j=\frac{1}{2\pi n}\left|\sum_{t=1}^{n}x_t e^{it\lambda_j}\right|^2$$

is the periodogram of $\{x_t\}_{t\in\mathbb{Z}}$.

Next, suppose that $\widehat{\alpha}_s$ is computed for $s=0,1,...,\tilde{n}$. Because under H_0 and suitable regularity conditions, the Whittle estimator is consistent, we expect that $\widehat{\alpha}_s\approx 0$ for all $s=0,1,...,\tilde{n}$, whereas under H_1, there exists an s such that $\widehat{\alpha}_s>0$. So, a test for the hypothesis testing in $(8)-(9)$ can be based on whether $\widehat{\alpha}_s$ is significantly greater than zero for some $s=0,1,...,\tilde{n}$. More precisely, the test statistic for the hypothesis testing for $(8)-(9)$ is given by

$$\mathcal{T}_W=\sup_{s=0,...,\tilde{n}}\widehat{\alpha}_s, \tag{15}$$

rejecting H_0 if $\mathcal{T}_W$ is greater than some critical value.

2.2. **Lagrange Multiplier test.**

The statistic described in (15) involves the estimation of $\alpha_0\left(\lambda_s\right)$ (and any other parameter of the model), for $s=0,...,\tilde{n}$, which can be highly computing intensive as nonlinear optimization algorithms are required. Herewith $\alpha_0\left(\lambda_s\right)=\alpha_0$ if s is such that $\left|\lambda_s-\lambda^0\right|\leq\left|\lambda_j-\lambda^0\right|$ for any j and $\alpha_0\left(\lambda_s\right)=0$, otherwise. Moreover, as the asymptotic distribution of $\mathcal{T}_W$ is not standard, see Theorem 2.1 below, to obtain (asymptotic) valid critical values it will be required to employ Monte-Carlo simulation algorithms such as Bootstrap schemes. Because of that, see Section 3 for another motivation of the bootstrap, the implementation of the test can be a prohibited task in computing time. Thus, we shall introduce a LM type of test which will not require the estimation of $\alpha_s:=\alpha_0\left(\lambda_s\right)$, but only the model under the null. This is computationally simpler and bootstrap algorithms will be easier and more feasible to implement.

To that end, consider the Whittle estimator of θ_0 under H_0, that is

$$\widetilde{\theta}=\underset{\theta\in\Theta}{\arg\min}\frac{1}{\tilde{n}}\sum_{j=1}^{\tilde{n}}\frac{I_j}{h_j\left(\theta\right)};\quad \widetilde{\sigma}^2=\frac{2\pi}{\tilde{n}}\sum_{j=1}^{\tilde{n}}\frac{I_j}{h_j\left(\widetilde{\theta}\right)}. \tag{16}$$

Next, for all s, consider the first derivative of (14) with respect to $\left(\alpha,\theta'\right)'$, and denote

$$q\left(\alpha,\theta,s\right)=\frac{1}{\widetilde{\sigma}^{2}\tilde{n}}\sum_{j=1}^{\tilde{n}}q_{j}\left(\alpha,\theta,s\right)\frac{I_{j}}{h_{j}\left(\theta\right)};\quad q_{j}\left(\alpha,\theta,s\right)=-\frac{\partial\log\left(g_{j}\left(\alpha,\lambda_{s}\right)h_{j}\left(\theta\right)\right)}{\partial\left(\alpha,\theta'\right)'} \tag{17}$$

and

$$V\left(\alpha,\theta,s\right)=\frac{1}{\tilde{n}}\sum_{j=1}^{\tilde{n}}q_{j}\left(\alpha,\theta,s\right)q_{j}'\left(\alpha,\theta,s\right).$$

Note that under standard suitable conditions, $V\left(\alpha,\theta,s\right)/\left(2\pi^{2}\right)$ is an estimate of the asymptotic covariance matrix of $\tilde{n}^{1/2}q\left(\alpha,\theta,s\right)$. Next, suppose that we fix a frequency λ_s for some $s=1,...,\tilde{n}$, say λ_{s^*}. Then the LM test for $\widetilde{H}_0:\alpha_0\left(\lambda_{s^*}\right)=\alpha_{s_*}=0$, becomes whether $\widetilde{q}_{s_*}$ is not significantly different than zero, where

$$\widetilde{q}_{s_*}=\left(V^{11}\left(0,\widetilde{\theta},s_*\right)\right)^{1/2}\tilde{n}^{1/2}q^{(1)}\left(0,\widetilde{\theta},s_*\right) \tag{18}$$

and $\widetilde{\theta}$ is given in (16), $q^{(1)}\left(\cdot\right)$ denotes the first component of the vector $q\left(\cdot\right)$ and $V^{11}\left(\cdot\right)$ is the element $(1,1)$ of $V^{-1}\left(\cdot\right)$.

Because under Conditions $C1-C5$, the Whittle estimator is consistent, then under H_0, we should expect that $\widetilde{q}_s\approx 0$ for all $s=0,1,...,\tilde{n}$. Now, our hypothesis testing is one sided. So, in the same way that for the W test we have rejected the null if $\widehat{\alpha}_s>0$ for some $s=1,...,\tilde{n}$, we need to find the direction of departure from the null hypothesis, that is the sign that $\widetilde{q}_s$ will take under the alternative hypothesis. It is clear that the sign will be that of

$$\frac{1}{\tilde{n}}\sum_{j=1}^{\tilde{n}}q_{j}^{(1)}\left(0,\widetilde{\theta},s\right)\frac{I_{j}}{h_{j}\left(\widetilde{\theta}\right)},$$

where $q_j^{(1)}\left(\cdot\right)$ is the first element of the vector $q_j\left(\cdot\right)$. Notice that because Condition $C5$ implies that $\{x_t\}_{t\in\mathbb{Z}}$ is ergodic, see Stout (1974) Theorem 3.5.8, $\widetilde{\sigma}^{2}\xrightarrow{p}\sigma^{2}>0$ (σ_0^2 under the null hypothesis). Because we are under the alternative, there exists an s for which $\alpha_0\left(\lambda_s\right):=\alpha_0>0$. Suppose for simplicity that $s=0$ and that $h\left(\lambda;\theta_0\right)=1$. Then, we have that the last displayed expression becomes

$$\begin{aligned}&-\frac{1}{\tilde{n}}\sum_{j=1}^{\tilde{n}}\frac{\partial\log g_{j}\left(0,0\right)}{\partial\alpha}I_{j}\\ =\;&-\frac{\sigma_{0}^{2}}{2\pi\tilde{n}}\sum_{j=1}^{\tilde{n}}\frac{\partial\log g_{j}\left(0,0\right)}{\partial\alpha}g_{j}\left(\alpha_{0},0\right)-\frac{1}{\tilde{n}}\sum_{j=1}^{\tilde{n}}\frac{\partial\log g_{j}\left(0,0\right)}{\partial\alpha}g_{j}\left(\alpha_{0},0\right)\left(\frac{I_{j}}{g_{j}\left(\alpha_{0},0\right)}-\frac{\sigma_{0}^{2}}{2\pi}\right).\end{aligned}$$

Now Lemma 6.1 implies that the second term on the right of the last displayed equation is

$$\frac{1}{\tilde{n}}\sum_{j=1}^{\tilde{n}}\frac{\partial\log g_{j}\left(0,0\right)}{\partial\alpha}g_{j}\left(\alpha_{0},0\right)\left(I_{\varepsilon,j}-\frac{\sigma_{0}^{2}}{2\pi}\right)+o_{p}\left(n^{-1/2}\right),$$

where $I_{\varepsilon,j}=\left(2\pi n\right)^{-1}\left|\sum_{t=1}^{n}\varepsilon_{t}e^{-it\lambda_{j}}\right|^{2}$ is the periodogram of $\{\varepsilon_t\}_{t\in\mathbb{Z}}$. However, by standard results on $I_{\varepsilon,j}$, the last expression is $O_p\left(n^{-1/2}\right)$. On the other hand, under $C1$, $g\left(\lambda;\alpha_0,0\right)=\left|2\sin\left(\lambda/2\right)\right|^{-\alpha_0}$, we have from the mean value theorem that

the first term on the right of the last displayed equation is

$$\frac{\sigma_0^2}{2\pi\tilde{n}}\sum_{j=1}^{\tilde{n}}\left|2\sin\left(\frac{\lambda_j}{2}\right)\right|^{-\alpha_0}\log\left|2\sin\left(\frac{\lambda_j}{2}\right)\right| = \frac{\sigma_0^2}{2\pi\tilde{n}}\left\{\sum_{j=1}^{\tilde{n}}\log\left|2\sin\left(\frac{\lambda_j}{2}\right)\right| - \alpha_0\sum_{j=1}^{\tilde{n}}\left|2\sin\left(\frac{\lambda_j}{2}\right)\right|^{-\widetilde{\alpha}_0}\log^2\left|2\sin\left(\frac{\lambda_j}{2}\right)\right|\right\},$$

where $\widetilde{\alpha}_0$ is an intermediate point between 0 and α_0. Next, by an obvious extension of Robinson's (1995*b*) Lemma 2 and that it is well-known that $\int_0^\pi \log|2\sin(\lambda/2)|\,d\lambda = 0$, the first term on the right of the last displayed equality is $O\left(\tilde{n}^{-1}\log\tilde{n}\right)$, whereas the second term is strictly negative since $\alpha_0 > 0$ and $\tilde{n}^{-1}\sum_{j=1}^{\tilde{n}}|2\sin(\lambda_j/2)|^{-\widetilde{\alpha}_0}\log^2|2\sin(\lambda_j/2)|$ is bounded away from zero. So, we conclude that $q^{(1)}(\cdot)\xrightarrow{p} c < 0$ and therefore $\tilde{n}^{1/2}q^{(1)}(\cdot)\xrightarrow{p}-\infty$. Hence the test for (8) will be based on

$$\mathcal{T}_{LM} = \sup_{s=0,1,\ldots,\tilde{n}} -\widetilde{q}_s, \tag{19}$$

rejecting H_0 if $\mathcal{T}_{LM}$ is greater than some critical value.

2.3. **Statistical properties of $\mathcal{T}_W$ and $\mathcal{T}_{LM}$.**

Denote

$$A(\tau) := \begin{pmatrix} A_{11}(\tau) & A_{12}(\tau) \\ A_{21}(\tau) & A_{22}(\tau) \end{pmatrix} = \begin{pmatrix} \int_0^1 \frac{\partial\log|g(\pi x,0,\pi\tau)|}{\partial\alpha}\frac{\partial\log|g(\pi x,0,\pi\tau)|}{\partial\alpha}dx & \int_0^1 \frac{\partial\log|g(\pi x,0,\pi\tau)|}{\partial\alpha}\frac{\partial\log h(\pi x;\theta_0)}{\partial\theta'}dx \\ \int_0^1 \frac{\partial\log h(\pi x;\theta_0)}{\partial\theta}\frac{\partial\log|g(\pi x,0,\pi\tau)|}{\partial\alpha}dx & \int_0^1 \frac{\partial\log h(\pi x;\theta_0)}{\partial\theta}\frac{\partial\log h(\pi x;\theta_0)}{\partial\theta'}dx \end{pmatrix},$$

where $\pi^{-1}\lim_{n\to\infty}\lambda_s = \tau \in [0,1]$, and let

$$A^{-1}(\tau) := \begin{pmatrix} A^{11}(\tau) & A^{12}(\tau) \\ A^{21}(\tau) & A^{22}(\tau) \end{pmatrix}.$$

Also, write

$$\mathcal{C}(\tau_1,\tau_2) = A^{11}(\tau_1)A^{11}(\tau_2)\mathcal{K}(\tau_1,\tau_2) + A^{11}(\tau_2)A^{12}(\tau_1)A_{21}(\tau_2) + A^{11}(\tau_1)A_{12}(\tau_1)A^{21}(\tau_2) + A^{12}(\tau_1)A_{22}A^{21}(\tau_2), \tag{20}$$

where

$$\mathcal{K}(\tau_1,\tau_2) = \int_0^1 \log|g(\pi x,0,\pi\tau_1)|\log|g(\pi x,0,\pi\tau_2)|\,dx. \tag{21}$$

We have the following result.

Theorem 2.1. *Assuming C1-C5, under H_0 given in (8), as $n\to\infty$*

$$(a) \qquad \tilde{n}^{1/2}\mathcal{T}_W \xrightarrow{d} \max_{\tau\in[0,1]}\mathcal{G}(\tau)$$

$$(b) \qquad \tilde{n}^{1/2}\mathcal{T}_{LM} \xrightarrow{d} \max_{\tau\in[0,1]}\mathcal{G}(\tau),$$

where $\mathcal{G}(\tau)$ is a process such that for fixed τ, $\mathcal{G}(\tau)$ is distributed as $X(\tau)\mathcal{I}(X(\tau)\geq 0)$, where $X(\tau)$ is a Gaussian process with covariance structure given by $\mathcal{C}(\tau_1,\tau_2)$ in (20), and $\mathcal{I}(\cdot)$ denotes the indicator function.

Proof. The proof of this result or any other is given in Section 5. □

One basic requirement for any test is its consistency. Also it is useful to learn about its power function against local alternatives to gain some insight about the characteristics of the test. To this end, consider

$$H_a: \quad \exists \lambda^0 \in [0,\pi] \text{ such that } f(\lambda) = \frac{\sigma_0^2}{2\pi} g\left(\lambda; \alpha(n), \lambda^0\right) h(\lambda; \theta_0),$$

where $\alpha(n) = \alpha_0/\tilde{n}^{1/2}$ with $\alpha_0 > 0$. Then, we have the following corollary:

Corollary 2.2. *Assuming C1-C5, under H_a, we have that as $n \to \infty$*

$$(a) \qquad \tilde{n}^{1/2}\mathcal{T}_W \xrightarrow{d} \max_{\tau\in[0,1]} \left(\mathcal{G}(\tau) + \mathcal{C}(\tau,\tau)^{-1/2} \alpha_0(\tau)\right)$$

$$(b) \qquad \tilde{n}^{1/2}\mathcal{T}_{LM} \xrightarrow{d} \max_{\tau\in[0,1]} \left(\mathcal{G}(\tau) + \mathcal{C}(\tau,\tau)^{-1/2} \alpha_0(\tau)\right),$$

where $\alpha_0(\tau) = \alpha_0 \mathcal{I}\left(\tau = \lambda^0/\pi\right)$.

From the results of Corollary 2.2, it is straightforward to observe that the tests will be consistent. This is the case because for fixed alternatives, that is, $\alpha(n) = \alpha_0 > 0$, $\mathcal{T}_W \xrightarrow{p} \mathcal{C}(\tau,\tau)^{-1/2} \alpha_0 > 0$. So we have that for any $z > 0$,

$$\lim_{n\to\infty} \Pr\left\{\tilde{n}^{1/2}\mathcal{T}_W > z\right\} = 1.$$

Similarly, we have that for any $z > 0$,

$$\lim_{n\to\infty} \Pr\left\{\tilde{n}^{1/2}\mathcal{T}_{LM} > z\right\} = 1.$$

Results of Theorem 2.1 give the (asymptotic) justification of the tests. On the other hand, following ideas in Steck (1971) and Noé (1972), the rate of convergence of the finite sample distribution to the asymptotic one seems to be slow, see also Hall (1979) for the rate of convergence of statistics based on the suprema. So, critical values relying on the asymptotic distribution can be a poor approximation to those of the finite sample distribution. One solution could be to employ Edgeworth expansions. However, Hall (1990) has shown that they do not perform well, compared to bootstrap schemes, at the tails of the distribution, which is precisely the most important region when testing. On the other hand, when exploring the properties of the bootstrap for the maximum of the kernel density estimator, Hall (1991) has shown that the bootstrap performs better than Edgeworth expansions in terms of their accuracy to the finite sample distribution of the suprema. In addition, because the asymptotic distribution of $\tilde{n}^{1/2}\mathcal{T}_W$ and $\tilde{n}^{1/2}\mathcal{T}_{LM}$ are nonstandard, it seems necessary to rely on Monte-Carlo algorithms to compute asymptotically valid critical values for the test. For all these reasons, in the next section we propose to use a bootstrap scheme.

3. BOOTSTRAP TESTS FOR $\mathbf{H_0}$

Bootstrap methods, introduced by Efron (1979), have become a routine method for approximating the distribution of a statistical quantity, partly due to the increasing computation power. At a theoretical level, Bootstrap algorithms have attracted considerable effort to their development, as they are capable of approximating the finite sample distribution of statistics more effectively than those based on their asymptotic counterparts, and also because they allow for the computation of valid asymptotic quantiles of the limiting distribution in nonstandard situations. In particular, when the limiting distribution is unknown or if known, the practitioner is unable to compute its quantiles. It is precisely the latter situation that we encounter the limiting process $\max_{\tau\in[0,1]} \mathcal{G}(\tau)$ is nonstandard and model dependent.

The basic idea of the bootstrap is, given a stretch of data $\mathcal{Z}_n = \{z_t, t = 1, ..., n\}$ say, to treat the data as if it were the true population, and to carry out Monte-Carlo experiments in which pseudo-data is drawn from $\mathcal{Z}_n$. Based on the underlying distributional properties of $\mathcal{Z}_n$, different schemes have been adopted and proposed.

In our context, the resampling method must be such that the conditional distribution, given $\underset{\sim}{x} = (x_1, ..., x_n)'$, of the bootstrap statistic, say $n^{1/2}\mathcal{T}_W^*$ the bootstrap analogue of $n^{1/2}\mathcal{T}_W$, consistently estimates the distribution of $\max_{\tau\in[0,1]} \mathcal{G}(\tau)$ under H_0 and the local alternatives H_a. That is, $\tilde{n}^{1/2}\mathcal{T}_W^* \xrightarrow{d^*} \max_{\tau\in[0,1]} \mathcal{G}(\tau)$ in probability under $H_0 \cup H_a$, where "$\xrightarrow{d^*}$" denotes

$$\Pr\left[\tilde{n}^{1/2}\mathcal{T}_W^* \leq z \middle| \underset{\sim}{x}\right] \xrightarrow{p} G(z),$$

at each continuity point z of $G(z) = \Pr\left(\max_{\tau\in[0,1]} \mathcal{G}(\tau) \leq z\right)$, as defined in Giné and Zinn (1990). A second requirement is that under H_1, $\tilde{n}^{1/2}\mathcal{T}_W^*$ must also converge in bootstrap distribution, although possibly to a different one than under H_0. Likewise for the bootstrap analogue of $\mathcal{T}_{LM}$, denoted by $\mathcal{T}_{LM}^*$. To achieve the first requirement, due to the (pseudo)Gaussian behaviour of the limit distribution, one key condition is that the resampling algorithm should preserve the correlation structure of the original data $\underset{\sim}{x}$, whereas the second requirement would be guaranteed if we were capable to bootstrap under the null hypothesis.

To achieve both aims, we propose the following bootstrap algorithm. To that end, denote the "*discrete Fourier transform*" (DFT) of a sequence $\{a_t\}_{t=1}^n$ by

$$w_a(\lambda_j) = \frac{1}{n^{1/2}} \sum_{t=1}^{n} a_t e^{-it\lambda_j}.$$

Suppose that we are under the null H_0, so that $f(\lambda) = \sigma_0^2 h(\lambda;\theta_0)/(2\pi)$. Then, using Condition $C5$, the identity

$$x_t = \frac{1}{n^{1/2}} \sum_{j=1}^{n} e^{it\lambda_j} w_x(\lambda_j) \tag{22}$$

and Bartlett's approximation of $w_x(\lambda_j)$, see Brockwell and Davis's (1991) Theorem 10.3.1, that is $w_x(\lambda_j) \approx B\left(e^{-i\lambda_j}\right) w_\varepsilon(\lambda_j)$ where $B(z) = \sum_{\ell=0}^{\infty} b_\ell z^\ell$, we obtain that x_t in (22) can be approximated by

$$x_t \approx \widetilde{x}_t =: \frac{1}{n^{1/2}} \sum_{j=1}^{n} e^{it\lambda_j} B\left(e^{-i\lambda_j}\right) w_\varepsilon(\lambda_j), \tag{23}$$

where "$\approx$" should be read as "approximately". Observe that (22) is nothing but the discrete (inverse) transformation of $w_x(\lambda_j)$, and that $\left|B\left(e^{-i\lambda}\right)\right|^2 = h(\lambda;\theta_0)$.

That $\widetilde{x}_t$ in (23) will preserve (asymptotically) the covariance structure of x_t can be easily seen by using Brillinger's (1981) Theorem 4.3.2. Indeed, the latter theorem and $C5$ imply that $E(w_\varepsilon(\lambda_j) w_\varepsilon(-\lambda_k)) = \sigma_0^2 \mathcal{I}(j = k)$, so that

$$\begin{aligned} E(\widetilde{x}_t \widetilde{x}_s) &= \frac{\sigma_0^2}{n} \sum_{j=1}^{n} e^{i(t-s)\lambda_j} \left|B\left(e^{-i\lambda_j}\right)\right|^2 \\ &\approx \frac{\sigma_0^2}{2\pi} \int_{-\pi}^{\pi} h(\lambda;\theta_0) e^{i(t-s)\lambda} d\lambda := E(x_t x_s), \end{aligned}$$

because $\left|B\left(e^{-i\lambda_j}\right)\right|^2 = h(\lambda_j;\theta_0)$ and under our conditions in Section 2, we have that $n^{-1}\sum_{j=1}^{n} h(\lambda_j;\theta_0) e^{i\ell\lambda_j} \underset{n\to\infty}{\to} (2\pi)^{-1} \int_0^{2\pi} h(\lambda;\theta_0) e^{i\ell\lambda} d\lambda$ by Brillinger (1981, p.15). Thus, if in the right side of (23), $B\left(e^{-i\lambda_j}\right)$ was replaced by a consistent estimator,

say $\widehat{B}\left(e^{-i\lambda_j}\right)$, the problem of obtaining a bootstrap sample x_t^*, $t=1,...,n$, becomes one of performing a valid bootstrap algorithm for the $DFT's$ $w_\varepsilon(\lambda_j)$, $j=1,...,n$.

The previous arguments suggest the following bootstrap algorithm.

STEP 1: Let $\underset{\sim}{\widetilde{x}}^* = (\widetilde{x}_1^*, \widetilde{x}_2^*, ..., \widetilde{x}_n^*)'$ be a random sample with replacement from the empirical distribution of the standardized residuals

$$\check{x}_t = \widetilde{\sigma}_x^{-1}(x_t - \overline{x}), \ \widetilde{\sigma}_x^2 = \frac{1}{n}\sum_{t=1}^n (x_t - \overline{x})^2; \quad \overline{x} = \frac{1}{n}\sum_{t=1}^n x_t$$

and obtain the DFT of $\underset{\sim}{\widetilde{x}}^*$ as

$$\eta_j^* = \frac{1}{n^{1/2}}\sum_{t=1}^n \widetilde{x}_t^* e^{-it\lambda_j}, \quad j=1,...,n.$$

Remark 3.1. *In the conclusions we mention another procedure to obtain η_j^*, $j=1,...,\tilde{n}$, based on bootstrapping directly from $\widehat{\varepsilon}_t$, $t=1,...,n$, an estimate of the innovations $\{\varepsilon_t\}_{t\in\mathbb{Z}}$ of the process $\{x_t\}_{t\in\mathbb{Z}}$.*

STEP 2: Compute

$$x_t^* = \frac{\widehat{\sigma}}{n^{1/2}}\sum_{j=1}^n e^{it\lambda_j}\widehat{B}\left(e^{-i\lambda_j}\right)\eta_j^*, \quad t=1,...,n,$$

where

$$\widehat{B}\left(e^{-i\lambda_j}\right) = 1 + \widehat{b}_1 e^{-i\lambda_j} + ... + \widehat{b}_n e^{-in\lambda_j},$$

with

$$\widehat{b}_\ell = \frac{1}{2\pi n}\sum_{j=-\tilde{n}+1}^{\tilde{n}-1} h^{1/2}\left(\lambda_j;\widehat{\theta}\right)e^{i\ell\lambda_j}, \quad \ell=1,...,n.$$

STEP 3: For $j=1,...,\tilde{n}$, compute the periodogram of the bootstrap sample x_t^*, $t=1,...,n$,

$$I_j^* = \frac{1}{2\pi n}\left|\sum_{t=1}^n x_t^* e^{it\lambda_j}\right|^2$$

and consider the Whittle objective function

$$Q^*(\alpha,\theta,s) = \frac{2\pi}{\tilde{n}}\sum_{j=1}^{\tilde{n}}\frac{I_j^*}{g_j(\alpha,\lambda_s)h_j(\theta)}. \tag{24}$$

To obtain the bootstrap analogue of $\mathcal{T}_W$, consider for all $s=0,...,\tilde{n}$

$$\left(\widehat{\alpha}_s^*, \widehat{\theta}^*\right)' = \underset{(\alpha,\theta')'\in\Pi\times\Theta}{\arg\min}\ Q^*(\alpha,\theta,s) \text{ and } \widehat{\sigma}^{*2} = Q^*\left(\widehat{\alpha}_s^*, \widehat{\theta}^*, s\right), \tag{25}$$

whereas to obtain the corresponding bootstrap for $\mathcal{T}_{LM}$, let

$$\widetilde{\theta}^* = \underset{\theta\in\Theta}{\arg\min}\frac{1}{\tilde{n}}\sum_{j=1}^{\tilde{n}}\frac{I_j^*}{h_j(\theta)}; \quad \widetilde{\sigma}^{*2} = \frac{2\pi}{\tilde{n}}\sum_{j=1}^{\tilde{n}}\frac{I_j^*}{h_j\left(\widetilde{\theta}^*\right)}.$$

Then, for all s, consider the first derivative of (24) with respect to $\left(\alpha,\theta'\right)'$, and denote

$$q^*(\alpha,\theta,s) = \frac{1}{\widetilde{\sigma}^{*2}\tilde{n}}\sum_{j=1}^{\tilde{n}} q_j(\alpha,\theta,s)\frac{I_j^*}{h_j(\theta)},$$

where $q_j(\alpha,\theta,s)$ is given in (17). From here, for all $s=0,...,\tilde{n}$, compute

$$\widetilde{q}_s^* = \left(V^{11}\left(0,\widetilde{\theta}^*,s\right)\right)^{1/2}\tilde{n}^{1/2}q^{(1)*}\left(0,\widetilde{\theta}^*,s\right), \tag{26}$$

where $q^{(1)*}\left(0,\widetilde{\theta}^{*},s\right)$ is the first element of the vector $q^{*}\left(0,\widetilde{\theta}^{*},s\right)$ and $V^{11}(\cdot)$ is defined in (18).

STEP 4: Compute the bootstrap statistics

$$\mathcal{T}_{W}^{*}=\sup_{s=0,...,\tilde{n}}\widehat{\alpha}_{s}^{*}, \tag{27}$$

$$\mathcal{T}_{LM}^{*}=\sup_{s=0,...,\tilde{n}}-\widetilde{q}_{s}^{*}. \tag{28}$$

The bootstrap scheme described in *STEPS 1 to 4* above is similar to the residual-based bootstrap of Franke and Kreiss (1992), but contrary to them, it is performed in the frequency domain. One particular feature of our bootstrap compared to the latter is its computational simplicity. This is specially true for models where the coefficients b_{ℓ} in $C5$ can be complicated functions, maybe implicit, of the underlying parameters θ_0, such as Bloomfield's exponential (1973) or $ARMA(p,q)$ models with fairly large p and q and complex roots on their autoregressive polynomials. However, if the coefficients b_{ℓ} were easily obtained from the parameters θ, say $b_{\ell}(\theta)$, then $B\left(e^{-i\lambda_j}\right)$ in *STEP 2* could be computed as

$$\widehat{B}\left(e^{-i\lambda_j}\right)=1+b_1\left(\widehat{\theta}\right)e^{-i\lambda_j}+...+b_n\left(\widehat{\theta}\right)e^{-in\lambda_j}.$$

As an example of the latter is when $\{x_t\}_{t\in\mathbb{Z}}$ follows an $AR(1)$ model.

Finally, we note that the difference between the latter and our bootstrap scheme parallels to that existing when we are interested to estimate the spectral density function by (a) approximating the dependence structure of the data by an $AR(p_n)$ model as in Berk (1974) with p_n tending to infinity with n or by (b) the average periodogram, see Brillinger (1981). Finally, it is worth noting that in *STEP 3*, we could have alternatively obtained I_j^* as $I_j^*=h_j\left(\widehat{\theta}\right)\left|\eta_j^*\right|^2\equiv\left|\widehat{B}\left(e^{-i\lambda_j}\right)\right|^2\left|\eta_j^*\right|^2$, see Hidalgo (2003). However, we prefer our method as we are able to approximate the transfer function $B\left(e^{-i\lambda_j}\right)$ instead of its modulus as the latter would do. A similar procedure was proposed by Theiler et al. (1992) and Prichard and Theiler (1994). However, their method amounts to change $\widehat{B}\left(e^{-i\lambda_j}\right)\eta_j^*$ by its module. Although this procedure may be valid for our purposes in this paper, it will not be valid in other simple situations such as bootstrapping the sample mean.

Theorem 3.1. *Assuming C1-C5, under $H_0\cup H_1$, as $n\to\infty$,*

$$(a)\qquad \tilde{n}^{1/2}\mathcal{T}_{W}^{*}\xrightarrow{d^{*}}\max_{\tau\in[0,1]}\mathcal{G}(\tau)\ \textit{in probability,}$$

$$(b)\qquad \tilde{n}^{1/2}\mathcal{T}_{LM}^{*}\xrightarrow{d^{*}}\max_{\tau\in[0,1]}\mathcal{G}(\tau)\ \textit{in probability.}$$

Thus, Theorem 3.1 indicates that the bootstrap statistics given in (27) and (28) are consistent. That is, let $c_{n,1-\beta}^{f,W}\left(c_{n,1-\beta}^{f,LM}\right)$ and $c_{1-\beta}^{a}$ be such that

$$\Pr\left\{\tilde{n}^{1/2}\mathcal{T}_{W}>c_{n,1-\beta}^{f,W}\right\}=\beta;\quad\left(\Pr\left\{\tilde{n}^{1/2}\mathcal{T}_{LM}>c_{n,1-\beta}^{f,LM}\right\}=\beta\right)$$

and

$$\lim_{n\to\infty}\Pr\left\{\tilde{n}^{1/2}\mathcal{T}_{W}>c_{1-\beta}^{a}\right\}=\beta;\quad\left(\lim_{n\to\infty}\Pr\left\{\tilde{n}^{1/2}\mathcal{T}_{LM}>c_{1-\beta}^{a}\right\}=\beta\right),$$

respectively. Hence, Theorems 2.1 and 3.1 indicate that $c_{n,1-\beta}^{f,W}\to c_{1-\beta}^{a}\left(c_{n,1-\beta}^{f,LM}\to c_{1-\beta}^{a}\right)$ and $c_{1-\beta}^{*,W}\xrightarrow{P}c_{1-\beta}^{a}\left(c_{1-\beta}^{*,LM}\xrightarrow{P}c_{1-\beta}^{a}\right)$ respectively, where $c_{1-\beta}^{*,W}\left(c_{1-\beta}^{*,LM}\right)$ are defined respectively by

$$\Pr\left\{\tilde{n}^{1/2}\mathcal{T}_{W}^{*}>c_{1-\beta}^{*,W}\right\}=\beta;\quad\left(\Pr\left\{\tilde{n}^{1/2}\mathcal{T}_{LM}^{*}>c_{1-\beta}^{*,LM}\right\}=\beta\right).$$

Typically, the finite sample distributions of $\tilde{n}^{1/2}\mathcal{T}_W^*$ or $\tilde{n}^{1/2}\mathcal{T}_{LM}^*$ are not available, so that the critical values $c_{1-\beta}^{*,W}$ and/or $c_{1-\beta}^{*,LM}$ are obtained by standard Monte-Carlo simulations, which can be approximated as accurately as desired. To that end, consider the bootstrap samples $\underset{\sim}{\widetilde{x}}^{*(\ell)} = \left(\widetilde{x}_1^{*(\ell)}, ..., \widetilde{x}_n^{*(\ell)}\right)'$, $\ell = 1, ..., B$, and compute $\mathcal{T}_W^{*(\ell)}$ and $\mathcal{T}_{LM}^{*(\ell)}$ as in (27) and (28) for each ℓ. Then, $c_{1-\beta}^{*,W}\left(c_{1-\beta}^{*,LM}\right)$ is approximated by the value $c_{1-\beta}^{*B,W}\left(c_{1-\beta}^{*B,LM}\right)$ satisfying

$$\frac{1}{B}\sum_{\ell=1}^{B}\mathcal{I}\left(\tilde{n}^{1/2}\mathcal{T}_W^{*(\ell)} > c_{1-\beta}^{*B,W}\right) = \beta; \qquad \left(\frac{1}{B}\sum_{\ell=1}^{B}\mathcal{I}\left(\tilde{n}^{1/2}\mathcal{T}_{LM}^{*(\ell)} > c_{1-\beta}^{*B,LM}\right) = \beta\right).$$

4. MONTE-CARLO EXPERIMENT AND AN EMPIRICAL EXAMPLE

4.1. Monte Carlo Experiment.

In order to investigate how well the bootstrap tests $\mathcal{T}_W^* = \sup_{s=0,...,\tilde{n}} \widehat{\alpha}_s^*$ and $\mathcal{T}_{LM}^* = \sup_{s=0,...,\tilde{n}} -\widetilde{q}_s^*$ perform in finite samples, a small Monte Carlo experiment was conducted. All throughout our Monte Carlo experiment we have employed 2000 replications with samples sizes $n = 64$ and 128. To calculate the bootstrap statistics, for all the models and sample sizes considered, 1000 bootstrap samples were employed, that is we have chosen $B = 1000$. To assess the empirical size and power of $\mathcal{T}_W^*$, due to the computationally time consuming of the test, we have only considered the model

$$(1-L)^{d_0} x_t = u_t, \quad (t = 0, \pm 1, ...) \tag{29}$$

where $\{u_t\}_{t\in\mathbb{Z}}$ is a zero mean and unit variance sequence of *iid* Gaussian random variables. In (29) we have chosen $d_0 = 0.0, 0.1, 0.2, 0.3$ and 0.4. When $d_0 = 0$, $\{x_t\}_{t\in\mathbb{Z}}$ is an *iid* Gaussian process and will evaluate the performance of the test in terms of its size, whereas for $d_0 > 0$, we have the well-known $ARFIMA(0, d_0, 0)$ model and will examine the power of our test.

On the other hand, when exploring the performance of the bootstrap test $\mathcal{T}_{LM}^*$, we have also considered, apart from the same set of models considered for the $\mathcal{T}_W^*$, the situation where $\{u_t\}_{t\in\mathbb{Z}}$ in (29) follows an $AR(1)$ or an $MA(1)$ model with parameter $\rho_0 = 0.5$ and $\omega_0 = 0.5$ respectively. That is,

$$\begin{aligned} u_t &= 0.5u_{t-1} + \varepsilon_t, \quad (t = 0, \pm 1, ...) \\ u_t &= \varepsilon_t + 0.5\varepsilon_{t-1}, \end{aligned}$$

where $\{\varepsilon_t\}_{t\in\mathbb{Z}}$ is a zero mean and unit variance sequence of iid Gaussian random variables. Also, to address the power of the test relative to the location of the pole, that is λ^0, we have considered the $GARMA(0, d_0, 0)$, $GARMA(1, d_0, 0)$ and $GARMA(0, d_0, 1)$ models with $\lambda^0 = \pi/2$. Under the null hypothesis the residuals, say e_t, of both AR and MA models are easily obtained. Hence we have also investigated the finite sample performance using the residual-based bootstrap as in Franke and Kreiss (1992) when generating the bootstrap samples and the test. Recall that as we mentioned in Section 3, our bootstrap algorithm can be regarded as an alternative or rival scheme to the latter bootstrap. To that end, and noting that under the null hypothesis $x_t = u_t$, we have followed the following *3 STEPS* (only the situation for the AR model is described, for the MA model the algorithm is similarly done, and so it is omitted).

STEP 1': Let $e_t = x_t - \widehat{\rho} x_{t-1}$, for $t = 2, ..., n$, and $e_1 = \left(1 - \widehat{\rho}^2\right)^{1/2} x_1$, where $\widehat{\rho}$ is the least squares estimator of the parameter ρ_0 in the model

$$x_t = \rho_0 x_{t-1} + \varepsilon_t.$$

Compute the standardized residuals

$$\check{e}_t = \widetilde{\sigma}_\varepsilon^{-1} (e_t - \overline{e}), \quad \widetilde{\sigma}_\varepsilon^2 = \frac{1}{n} \sum_{t=1}^{n} (e_t - \overline{e})^2 \quad \text{and} \quad \overline{e} = \frac{1}{n} \sum_{t=1}^{n} e_t.$$

Let $\underset{\sim}{\widetilde{e}}^* = (\widetilde{e}_1^*, \widetilde{e}_2^*, ..., \widetilde{e}_n^*)'$ be a random sample with replacement from the empirical distribution of $(\check{e}_t)_{t=1}^n$ and obtain the bootstrap observations x_t^*, $t = 1, ..., n$, as

$$\begin{aligned} x_t^* &= \widehat{\rho} x_{t-1}^* + \widetilde{e}_t^* \quad t = 2, ..., n \\ x_1^* &= \left(1 - \widehat{\rho}^2\right)^{-1/2} \widetilde{e}_1^*. \end{aligned}$$

STEP 2': Exactly as *STEP 3* in Section 3, but with x_t^* as generated in *STEP 1'* instead of *STEP 2*, and then, for all $s = 0, ..., \tilde{n}$, compute

$$\widetilde{q}_s^* = \left(V^{11}\left(0, \widetilde{\theta}^*, s\right)\right)^{1/2} \tilde{n}^{1/2} q^{(1)*}\left(0, \widetilde{\theta}^*, s\right).$$

Note that in this model $\theta_0 = \rho_0$.

STEP 3': Compute the bootstrap statistic

$$\mathcal{T}_{LM}^* = \sup_{s=0,...,\tilde{n}} -\widetilde{q}_s^*. \tag{30}$$

The results of our experiments are given in *TABLES 4.1 to 4.3* below, where the two schemes to bootstrap the LM test will be denoted in the tables as METHOD 2 for that given in (30), whereas the bootstrap based on (28) is denoted as METHOD 1.

TABLES 4.1 TO 4.3 ABOUT HERE

TABLE 4.1 illustrates the size of $\mathcal{T}_W$ and $\mathcal{T}_{LM}$, whereas *TABLE 4.2* illustrate the power of the Wald test and *TABLE 4.3* the power for the LM test for all the different models described above. As *TABLE 4.1* illustrates, in terms of the empirical size, the LM test tends to perform better than the W type test. This is particularly true for small sample sizes, e.g. $n = 64$. When we compare the performance of the test based on (30) and (28), our proposed method appears to behave similarly to the residual-based bootstrap even in a model for which the latter may be preferable due to the simplicity to obtain the residuals e_t under the null hypothesis. This is particularly the situation when $n = 64$, which is the typical size of many macroeconomic data. Regarding the power of the tests, $\mathcal{T}_{LM}$ performs much better than $\mathcal{T}_W$ uniformly for both d_0 and n. However, for all the tests, the power increases with the sample size n and with d_0. The latter is expected as the "distance" between the null and alternative becomes greater as d_0 becomes bigger. On the other hand, when we compare the power performance of the two rival METHODS for the LM test, we observe that once again, our proposed bootstrap tends to perform as good as METHOD 2 for both sample sizes considered. When we compared the power performance of the $\mathcal{T}_{LM}$ test relative to the short-memory and the location of the pole considered we observe the following. First, the power is smaller when the short memory component follows an AR model than when it is a MA model. This is consistent with the empirical observation of the difficulty to discriminate between an AR model with a high value of ρ_0 and an $ARFIMA$ model with positive d_0. This is also the conclusion obtained from the Monte Carlo experiment considered by Lobato and Robinson (1998). However, for MA models

the power is quite high. Second, as the pole moves away from the zero frequency, we see that in that case the influence of the short memory AR and/or MA component seems to be negligible. This may by somehow expected, among other things, as the location of the maximum of the spectral density function for AR model with positive ρ_0 and our $GARMA$ models are very different. Overall, the size and power, for all sample sizes considered in the experiment, are very satisfactory and the frequency-based bootstrap described in Section 3 appears to be a good competitor to the time-domain bootstrap described above, even in situations where the latter is easy to obtain as is the case with an $AR(1)$ model, where the residual-based bootstrap might be preferable or employed a priori. We shall also mention that we have performed the bootstrap test as suggested in the concluding section. However, we did not find any significant difference with the method proposed in Section 3, neither when the distribution of the innovations $\{\varepsilon_t\}_{t\in\mathbb{Z}}$ were a χ^2_1 nor a χ^2_8.

4.2. **Empirical Example.**

A small application of our test has been examined for possible strong cycles to some real data set. We have employed the rate of growth of monthly seasonal adjusted *USA Industrial Production Index*, IPI data, say x_t. The data was obtained from the IFS database of the IMF and expands from February 1957 to October 2003. That is the sample size is $n = 561$. Looking at the graph of the series, see graphs 1 and 2, there is clearly a change of pattern around December 1989/January 1990. Because of that we have implemented our tests for both subsamples. Moreover, graphs 3 and 4 suggests that the serial dependence for the first and second subsample is different. Also, the latter figures suggests for the possibility of a pole away from zero for the first subsample, whereas the second subsample appears to have a clear pole at the zero frequency.

For the first subsample, that is when x_t expands from February 1957 to December 1989, we have used two different specifications for our short memory component. More specifically, the models chosen were

$$\left(1 - 2\left(\cos\lambda^0\right)L + L^2\right)^{d_0} x_t = (1+\theta_1 L)\left(1+\theta_{12}L^{12}+\theta_{24}L^{24}\right)\varepsilon_t \tag{31}$$

$$\left(1-\rho_{12}L^{12}-\rho_{24}L^{24}\right)\left(1 - 2\left(\cos\lambda^0\right)L + L^2\right)^{d_0} x_t = (1+\theta_1 L)\,\varepsilon_t. \tag{32}$$

Also, the previous models were estimated using an $MA(2)$ model instead of the $MA(1)$, say

$$\left(1 - 2\left(\cos\lambda^0\right)L + L^2\right)^{d_0} x_t = (1+\theta_1 L+\theta_2 L)\left(1+\theta_{12}L^{12}+\theta_{24}L^{24}\right)\varepsilon_t.$$

However, the parameter θ_2 was statistically insignificant, so we employed the aforementioned models (31) and (32). We have performed the LM test in both specifications. The results for model (31) were that the null of no strong cycles, e.g. $d_0 = 0$, was rejected at the 1% significant level. When the model (31) was estimated, leaving the location of the pole unknown, we found a strong cycle component at $\lambda = 0.14$, which corresponds to a cycle of 45 months. Proceeding as with the model given in (31), when we considered the model given in (32), we rejected the null hypothesis at the 1% level. When the model (32) was estimated, the estimate of the pole λ^0 was also at the frequency 0.14.

The same analysis was carried out using the second subsample from January 1990 to October 2003. Now, the fitted models were

$$\left(1 - 2\left(\cos\lambda^0\right)L + L^2\right)^{d_0} x_t = (1+\theta_1 L)\,\varepsilon_t \tag{33}$$

$$(1-\rho_1 L)\left(1 - 2\left(\cos\lambda^0\right)L + L^2\right)^{d_0} x_t = \varepsilon_t.$$

We obtained that the null hypothesis was rejected at 1% significant level using the LM test. However, for both models the estimated value of the pole was now at the

zero frequency. The latter suggests that a change of the cyclical component of the data may have occurred around January 1990. Also, when an $MA\,(2)$ model or an $AR\,(2)$ were estimated, they did not provide a better fit when compared to the last displayed two models.

We now comment on the procedure followed to identify the models employed in the empirical example. In a first step, using Brockwell and Davis's ITSM package, an $ARMA\,(4,13)$ and $ARMA\,(1,1)$ model, for the first and second subsamples respectively, were identified. This package chooses the best model among the class of $ARMA\,(27,27)$ models following the AIC criterium. After removing the highly insignificant parameters, the models were estimated and their spectral density functions were then compared with the periodogram of the series in each of the subsamples. We observed a big discrepancy between them. In particular, for the first subsample, the estimated spectral density function was unable to capture neither the observed peak nor the behaviour at high frequencies of the periodogram. Although a formal test comparing the periodogram and estimated spectral density function is available, see Delgado, Hidalgo and Velasco (2004), we have just used this "eyeball" test in our empirical example. Hence, we looked further for other models. Using the AIC criterion, we chose an $ARMA(0,25)$. After removing highly insignificant parameters, we compared again the estimated spectral density function with the periodogram. The new model appears to capture very well the behaviour of the series for high frequencies, but it was still not able to capture the observed peak suggested by the periodogram. All this suggests that a $GARMA$ component maybe adequate to explain the movement of the data. Furthermore, the MA coefficients corresponding to the lags $12, 13, 24$ and 25 were highly significant. This gave us the indication to use the short memory component given in (31) or (32). On the other hand, following the same procedure for the second subsample, the estimated spectral density was able to fit well the high frequencies, but not the observed peak at zero of the periodogram. Given that the $AR(1)$ component is the one that tries to capture the peak, we chose to use also a $MA(1)$ model for the short memory component for the second subsample. Finally, we fitted both models (31) / (32) and (33) using the Whittle estimator in Giraitis, Hidalgo and Robinson (2001). When we fitted those restricted models, we noticed that their corresponding spectral density functions captured well both the peak and the behaviour at the high frequencies of the periodogram in each of the subsamples. Overall, there were not discrepancies between them, and therefore, the eyeball test suggests that the estimated models were able to capture the behaviour of the data very well, at least their most relevant features. Nevertheless, as mentioned in the previous paragraph, the test strongly rejects the null of no strong cycles and the long memory parameter was highly significant. So, to conclude, we have found significant evidence of strong cycles. Hence, the data should be fitted using a $GARMA$ rather than an $ARMA$ model, which gives a poorer fit than the former $GARMA$ models.

5. PROOFS OF THE MAIN RESULTS

5.1. **Proof of Theorem 2.1.** We shall begin showing part (a) first. The proof for the $\mathcal{T}_{LM}$ is similar and we shall only sketch its difference compared to that for the $\mathcal{T}_W$. For given λ_s, that is the closest Fourier frequency to λ^0, under $C1-C5$, Hosoya (1997) or Fox and Taqqu (1987) and Giraitis and Surgailis (1990), among others, have shown that

$$\left(\widehat{\alpha}_s, \left(\widehat{\theta}-\theta_0\right)'\right)' \xrightarrow{P} 0.$$

Next, denote $\ell_{j,n}(\alpha,\theta,s) = g_j^{-1}(\alpha,\lambda_s)\, h_j^{-1}(\theta)\, I_j$ with

$$g_j(\alpha,\lambda_s) = \left|4\sin\left(\frac{\lambda_j+\lambda_s}{2}\right)\sin\left(\frac{\lambda_j-\lambda_s}{2}\right)\right|^{-\alpha}.$$

Then, proceeding as in Hosoya (1997), see also Giraitis et al. (2001), we have that

$$\begin{aligned}\sum_{j=1}^{\tilde{n}}\ell_{j,n}(\alpha,\theta,s) &= \sum_{j=1}^{\tilde{n}}\ell_{j,n}(0,\theta_0,s)+\sum_{j=1}^{\tilde{n}}\left(D\ell_{j,n}(0,\theta_0,s)\right)'\left(\alpha_s,(\theta-\theta_0)'\right)'\\ &\quad+\frac{1}{2}\left(\alpha_s,(\theta-\theta_0)'\right)\sum_{j=1}^{\tilde{n}}D^2\ell_{j,n}(0,\theta_0,s)\left(\alpha_s,(\theta-\theta_0)'\right)'\\ &\quad+R_n(\alpha,\theta,s), \qquad (34)\end{aligned}$$

where for all $s=0,...,\tilde{n}$,

$$\sup_{\left\{n^{1/2}|\alpha|\leq\xi;n^{1/2}\|\theta-\theta_0\|<\xi\right\}}|R_n(\alpha,\theta,s)| = o_p(1)$$

and where D and/or D^2 denote the first and second generalized derivatives respectively, see Andrews (2001) for a definition. On the other hand, Lemma 6.1 implies that

$$\begin{aligned}\sup_{s=0,1,...,\tilde{n}}\left|\frac{1}{\tilde{n}^{1/2}}\sum_{j=1}^{\tilde{n}}\left(D\log g_j(0,\lambda_s),D\log h_j(\theta_0)\right)'\left(\frac{I_j}{h_j(\theta_0)}-I_{\varepsilon,j}\right)\right| &= o_p(1)\\ \sup_{s=0,1,...,\tilde{n}}\left|\frac{1}{\tilde{n}^{1/2}}\sum_{j=1}^{\tilde{n}}\left(|D\log g_j(0,\lambda_s)|^2+\|D\log h_j(\theta_0)\|^2\right)\left(\frac{I_j}{h_j(\theta_0)}-I_{\varepsilon,j}\right)\right| &= o_p(1),\end{aligned}$$

since $|D\log g_j(0,\lambda_s)|^2$ and $\|D\log h_j(\theta_0)\|^2$ satisfy the same conditions of $\zeta(\lambda,s)$ in Lemma 6.1. So, to show that

$$\left(\frac{1}{\tilde{n}^{1/2}}\sum_{j=1}^{\tilde{n}}\left(D\ell_{j,n}(0,\theta_0,[\tilde{n}\tau])\right)',\frac{1}{\tilde{n}}\sum_{j=1}^{\tilde{n}}D^2\ell_{j,n}(0,\theta_0,[\tilde{n}\tau])\right)\Rightarrow\left(X(\tau),A(\tau)\right),\ (\tau\in[0,1]) \qquad (35)$$

it suffices to do so for

$$X_n(\tau) = \frac{1}{\tilde{n}^{1/2}}\sum_{j=1}^{\tilde{n}}\begin{pmatrix}D\log g_j\left(0,\lambda_{[\tilde{n}\tau]}\right)\\ D\log h_j(\theta_0)\end{pmatrix}\left(I_{\varepsilon,j}-\frac{\sigma_0^2}{2\pi}\right) \qquad (36)$$

$$A_n(\tau) = \frac{1}{\tilde{n}}\sum_{j=1}^{\tilde{n}}\begin{pmatrix}D\log g_j\left(0,\lambda_{[\tilde{n}\tau]}\right)\\ D\log h_j(\theta_0)\end{pmatrix}\left(D\log g_j\left(0,\lambda_{[\tilde{n}\tau]}\right),D\log h_j(\theta_0)\right)I_{\varepsilon,j}.$$

To that end, we need to show: (i) That for any finite collection $\tau_{\ell_1},...,\tau_{\ell_p}$,

$$\left(X_n(\tau_{\ell_1}),...,X_n(\tau_{\ell_p})\right)'\xrightarrow{d}\left(X(\tau_{\ell_1}),...,X(\tau_{\ell_p})\right)',$$

where $\left(X(\tau_{\ell_1}),...,X(\tau_{\ell_p})\right)'\simeq\mathcal{N}(0,A)$ and where the $(\tau_{\ell_1},\tau_{\ell_2})\,th$ element of A is

$$\begin{aligned}&\lim_{n\to\infty}\frac{1}{\tilde{n}}\sum_{j=1}^{\tilde{n}}\begin{pmatrix}D\log\left|g_j\left(0,\lambda_{[\pi\tau_{\ell_1}]}\right)\right|\\ D\log h_j(\theta_0)\end{pmatrix}\begin{pmatrix}D\log\left|g_j\left(0,\lambda_{[\pi\tau_{\ell_2}]}\right)\right|\\ D\log h_j(\theta_0)\end{pmatrix}'\\ &= \int_0^1\begin{pmatrix}D\log|g(\pi x,0,\pi\tau_{\ell_1})|\\ D\log h(\pi x;\theta_0)\end{pmatrix}\begin{pmatrix}D\log|g(\pi x,0,\pi\tau_{\ell_2})|\\ D\log h(\pi x;\theta_0)\end{pmatrix}'dx.\end{aligned}$$

(ii) The process $X_n(\tau)$ is tight in the space $\mathbb{D}[0,1]$ with the Skorohod's metric and (iii) that $A_n(\tau)$ converges to $A(\tau)$. Note that as $A_n(\cdot)$ will converge to a

nonstochastic matrix, it suffices to examine the behaviour of $X_n(\tau)$ and $A_n(\tau)$ separately.

First (i) holds true using standard results, see Hosoya (1997) or Giraitis et al. (2001). So, we only need to examine part (ii) to complete the proof of the weak convergence of $X_n(\tau)$ to $X(\tau)$. To prove the tightness of the process $X_n(\tau)$, and since the second component of $X_n(\tau)$ does not depend on τ, it suffices to show the tightness of

$$X_{n1}(\tau) = \frac{1}{\tilde{n}^{1/2}} \sum_{j=1}^{\tilde{n}} D \log g_j\left(0, \lambda_{[\tilde{n}\tau]}\right) \left(I_{\varepsilon,j} - \frac{\sigma_0^2}{2\pi}\right).$$

Because the limit process has continuous paths, see the proof of (41) given in (42) and comments that follow, Billingsley's (1968) Theorems 15.4 and 15.6 imply that it is sufficient to check the Kolmogorov-Chentsov's moment condition

$$E\left[\left|X_{n1}(\tau) - X_{n1}(\tau_1)\right|^2 \left|X_{n1}(\tau_2) - X_{n1}(\tau)\right|^2\right] \leq K\left(F(\tau_2) - F(\tau_1)\right)^{1+\delta} \quad (37)$$

for some $\delta > 0$ where $0 \leq \tau_1 < \tau < \tau_2 \leq 1$ and $F(\tau)$ is a nondecreasing and continuous function.

First, we observe that we can focus in the case $\tilde{n}^{-1} \leq \tau_2 - \tau_1$. If $\tau_2 - \tau_1 < \tilde{n}^{-1}$, then either τ_1 and τ lie in the same subinterval $[(p-1)/\tilde{n}, p/\tilde{n})$, with $p = 1, ..., \tilde{n}$, or else τ and τ_2 do; in either of these cases the left side of (37) vanishes. Because the Schwarz inequality implies that the left side of (37) is bounded by

$$\left(E\left|X_{n1}(\tau) - X_{n1}(\tau_1)\right|^4\right)^{1/2} \left(E\left|X_{n1}(\tau_2) - X_{n1}(\tau)\right|^4\right)^{1/2},$$

and $\left(F(\tau) - F(\tau_1)\right)\left(F(\tau_2) - F(\tau)\right) \leq \left(F(\tau_2) - F(\tau_1)\right)^2$, and hence to show the tightness condition (37), it suffices to show that

$$E\left|X_{n1}(\tau_2) - X_{n1}(\tau_1)\right|^4 \leq K\left(F(\tau_2) - F(\tau_1)\right)^{1+\delta}, \ \delta > 0. \quad (38)$$

Choose $F(\tau) = \tau$ and denote $\eta_j = I_{\varepsilon,j} - (2\pi)^{-1}\sigma_0^2$, then

$$\begin{aligned} E\left|X_{n1}(\tau_2) - X_{n1}(\tau_1)\right|^4 &= \frac{1}{\tilde{n}^2} \sum_{j_1,...,j_4=1}^{\tilde{n}} \prod_{\ell=1}^{4} \left(\log \frac{\left|g_j\left(0, \lambda_{j_\ell} - \lambda_{[\tilde{n}\tau_1]}\right)\right|}{\left|g_j\left(0, \lambda_{j_\ell} - \lambda_{[\tilde{n}\tau_2]}\right)\right|}\right) E\left\{\prod_{\ell=1}^{4} \eta_{j_\ell}\right\} \\ &= \frac{3}{\tilde{n}^2} \sum_{j_1,...,j_4=1}^{\tilde{n}} \prod_{\ell=1}^{4} \left(\log \frac{\left|g_j\left(0, \lambda_{j_\ell} - \lambda_{[\tilde{n}\tau_1]}\right)\right|}{\left|g_j\left(0, \lambda_{j_\ell} - \lambda_{[\tilde{n}\tau_2]}\right)\right|}\right) E\left\{\eta_{j_1}\eta_{j_2}\right\} E\left\{\eta_{j_3}\eta_{j_4}\right\} \end{aligned}$$

$$+\frac{1}{n^2} \sum_{j_1,...,j_4=1}^{\tilde{n}} \prod_{\ell=1}^{4} \left(\log \frac{\left|g_j\left(0, \lambda_{j_\ell} - \lambda_{[\tilde{n}\tau_1]}\right)\right|}{\left|g_j\left(0, \lambda_{j_\ell} - \lambda_{[\tilde{n}\tau_2]}\right)\right|}\right) Cum\left(\eta_{j_1}, \eta_{j_2}, \eta_{j_3}, \eta_{j_4}\right). \quad (39)$$

By Brillinger's (1981) Theorems 2.3.1 and 4.3.2, and in particular the equation (4.3.15), and the integrability of $\log^q |x|$, $q = 1, ..., 4$, we have that the second term on the right of (39) is bounded in absolute value by

$$K\tilde{n}^{-2} \log^4 \tilde{n} \leq K\left(\tau_2 - \tau_1\right)^{1+\delta} \leq K\left(\tau_2 - \tau_1\right)^{1+\delta},$$

for any $\frac{1}{2} < \delta < 1$, since $\tilde{n}^{-1} < \tau_2 - \tau_1$. So, to complete the proof we need to examine the first term on the right of (39) which is

$$3\sigma_0^2 \left[\frac{1}{\tilde{n}} \sum_{j=1}^{\tilde{n}} \left(\log \frac{\left|g\left(0, \lambda_j - \lambda_{[\tilde{n}\tau_1]}\right)\right|}{\left|g\left(0, \lambda_j - \lambda_{[\tilde{n}\tau_2]}\right)\right|}\right)^2\right]^2.$$

(Recall that $\tau_j = \lim_{n\to\infty} s_j/\tilde{n}$, for $j = 1, 2$.) Now, by the Cauchy-Schwarz inequality and the definition of $g_j\left(0, \lambda_j - \lambda_{[\tilde{n}\tau_1]}\right)$, the last displayed expression is bounded by

$$12\sigma^2\left[\frac{1}{\tilde{n}}\sum_{j=1}^{\tilde{n}}\left(\log\frac{\left|\sin\left(\frac{\lambda_j-\lambda_{[\tilde{n}\tau_1]}}{2}\right)\right|}{\left|\sin\left(\frac{\lambda_j-\lambda_{[\tilde{n}\tau_2]}}{2}\right)\right|}\right)^2\right]^2 + 12\sigma^2\left[\frac{1}{\tilde{n}}\sum_{j=1}^{\tilde{n}}\left(\log\frac{\left|\sin\left(\frac{\lambda_j+\lambda_{[\tilde{n}\tau_1]}}{2}\right)\right|}{\left|\sin\left(\frac{\lambda_j+\lambda_{[\tilde{n}\tau_2]}}{2}\right)\right|}\right)^2\right]^2.$$

We shall examine the first term, being the second identically handled. This term, except constants, is

$$\frac{1}{\tilde{n}}\sum_{j=1}^{[\tilde{n}\tau_1]}\left(\log\frac{\left|\sin\left(\frac{\lambda_{[\tilde{n}\tau_1]}-\lambda_j}{2}\right)\right|}{\left|\sin\left(\frac{\lambda_{[\tilde{n}\tau_2]}-\lambda_j}{2}\right)\right|}\right)^2 + \frac{1}{\tilde{n}}\sum_{j=[\tilde{n}\tau_2]+1}^{\tilde{n}}\left(\log\frac{\left|\sin\left(\frac{\lambda_j-\lambda_{[\tilde{n}\tau_1]}}{2}\right)\right|}{\left|\sin\left(\frac{\lambda_j-\lambda_{[\tilde{n}\tau_2]}}{2}\right)\right|}\right)^2$$
$$+\frac{2}{\tilde{n}}\sum_{j=[\tilde{n}\tau_1]+1}^{[\tilde{n}\tau_2]-1}\log^2\frac{\left|\sin\left(\frac{\lambda_j-\lambda_{[\tilde{n}\tau_1]}}{2}\right)\right|}{\left|\sin\left(\frac{\lambda_{[\tilde{n}\tau_2]}-\lambda_j}{2}\right)\right|}.$$

Because an obvious extension of Robinson's (1995*b*) Lemma 2 implies that the difference between the Riemann sum and the integral is bounded in absolute value by

$$K\tilde{n}^{-1}\log^2 n \leq K\left(\tau_2-\tau_1\right)^{(1+\delta)/2},$$

for any $\frac{1}{2} < \delta < 1$, the proof is completed if we show that

$$\int_0^{\tau_1}\left(\log\frac{\left|\sin\left(\frac{\pi\tau_1-\pi u}{2}\right)\right|}{\left|\sin\left(\frac{\pi\tau_2-\pi u}{2}\right)\right|}\right)^2 du + \int_{\tau_2}^{1/2}\left(\log\frac{\left|\sin\left(\frac{\pi u-\pi\tau_1}{2}\right)\right|}{\left|\sin\left(\frac{\pi u-\pi\tau_2}{2}\right)\right|}\right)^2 du$$
$$+\int_{\tau_1}^{\tau_2}\log^2\frac{\left|\sin\left(\frac{\pi u-\pi\tau_1}{2}\right)\right|}{\left|\sin\left(\frac{\pi u-\pi\tau_2}{2}\right)\right|}du \qquad (40)$$

is bounded by $K\left(\tau_2-\tau_1\right)^{(1+\delta)/2}$. Recall that we have chosen $F(\tau) = \tau$.

The proof proceeds similarly as that given in Lemma 6.1, cf. (58). First, observing that in the interval $(0, \tau_1)$, we have that

$$\begin{aligned}&\left(\log\left|\sin\left(\frac{\pi\tau_1-\pi u}{2}\right)\right| - \log\left|\sin\left(\frac{\pi\tau_2-\pi u}{2}\right)\right|\right)^2\\ \leq\ & \log^2\left|\sin\left(\frac{\pi\tau_1-\pi u}{2}\right)\right| - \log^2\left|\sin\left(\frac{\pi\tau_2-\pi u}{2}\right)\right|,\end{aligned}$$

then the first term of (40) is bounded by

$$\begin{aligned}&K\left(\int_0^{\tau_1}\log^2\left|\sin\left(\frac{\pi\tau_1-\pi u}{2}\right)\right|du - \int_0^{\tau_1}\log^2\left|\sin\left(\frac{\pi\tau_2-\pi u}{2}\right)\right|du\right)\\ =\ & K\left(H\left(\tau_1\right) - H\left(\tau_2\right) + H\left(\tau_2-\tau_1\right)\right)\\ \leq\ & 2KH\left(\tau_2-\tau_1\right),\end{aligned}$$

where $H(\tau) = \int_0^{\tau}\log^2\left|\sin\left(\frac{\pi u}{2}\right)\right|du$ and because $\log^2\left|\sin\left(\frac{\pi\tau_1-\pi u}{2}\right)\right| - \log^2\left|\sin\left(\frac{\pi\tau_2-\pi u}{2}\right)\right| > 0$ implies that $H\left(\tau_2\right) - H\left(\tau_1\right) < H\left(\tau_2-\tau_1\right)$. But $H(\tau) \leq \tau^{(1+\delta)/2}$ for $0 \leq \tau \leq 1$, so we conclude that the first term of (40) is bounded by

$$K\left(\tau_2-\tau_1\right)^{(1+\delta)/2}.$$

Next, we examine the second term of (40). Using that in the interval $(\tau_2, 1/2)$,

$$\left(\log\left|\sin\left(\frac{\pi u-\pi\tau_1}{2}\right)\right| - \log\left(\frac{\pi u-\pi\tau_2}{2}\right)\right)^2 \leq \log^2\left|\frac{\pi u-\pi\tau_2}{2}\right| - \log^2\left|\sin\left(\frac{\pi u-\pi\tau_1}{2}\right)\right|,$$

we then have that the second term of (40) is bounded by

$$\begin{aligned} 0 &< \int_{\tau_2}^{1/2} \log^2 \left|\sin\left(\frac{\pi\tau_2 - \pi u}{2}\right)\right| du - \int_{\tau_2}^{1/2} \log^2 \left|\sin\left(\frac{\pi\tau_1 - \pi u}{2}\right)\right| du \\ &= H\left(\frac{1}{2} - \tau_2\right) - H\left(\frac{1}{2} - \tau_1\right) + H\left(\tau_2 - \tau_1\right) \\ &\leq 2H\left(\tau_2 - \tau_1\right) \end{aligned}$$

because $H\left(\frac{1}{2} - \tau_1\right) - H\left(\frac{1}{2} - \tau_2\right) \leq H\left(\tau_2 - \tau_1\right)$. So, the second term of (40) is also bounded by $K\left(\tau_2 - \tau_1\right)^{(1+\delta)/2}$. Finally, by a change of variables, the third term of (40) is

$$\int_0^{\tau_2 - \tau_1} \log^2 \left|\sin\left(\frac{\pi u}{2}\right)\right| du = H\left(\tau_2 - \tau_1\right) \leq K\left(\tau_2 - \tau_1\right)^{(1+\delta)/2}.$$

So (38) holds true with $F(\tau) = \tau$.

To complete the proof that (35) holds true, we need to show $\Pr\{X(1-) \neq X(1)\} = 0$ or alternatively that for every positive $\varsigma > 0$,

$$\lim_{\tau \to 1} \Pr\{|X(1) - X(\tau)| > \varsigma\} = 0. \tag{41}$$

But this is the case as we now show. In fact, we shall show that for every positive $\varsigma > 0$

$$\lim_{\zeta \to 0} \sup_{0 \leq \tau \leq 1-\zeta} \zeta^{-1} \Pr\{|X(\tau + \zeta) - X(\tau)| > \varsigma\} = 0. \tag{42}$$

The latter condition will imply that the process $X(\tau)$ belongs to the space $\mathbb{C}[0,1]$ by Problem 15.3 in Billingsley (1968, p. 136). To that end, because Theorem 5.3 in Billingsley (1968) implies that

$$E\left|X(\tau + \zeta) - X(\tau)\right|^4 \leq \liminf_{n \to \infty} E\left|X_{n1}(\tau + \zeta) - X_{n1}(\tau)\right|^4,$$

we have that Markov's inequality implies that (42) holds true if the right side of the last displayed inequality satisfies that

$$\lim_{\zeta \to 0} \sup_{0 \leq \tau \leq 1-\zeta} \zeta^{-1} \liminf_{n \to \infty} E\left|X_{n1}(\tau + \zeta) - X_{n1}(\tau)\right|^4 = 0.$$

But, this is the case because (38) implies that

$$\liminf_{n \to \infty} \zeta^{-1} E\left|X_{n1}(\tau + \zeta) - X_{n1}(\tau)\right|^4 \leq K\zeta^{\xi},$$

with $\xi > 0$. In fact, the continuity is not more than a consequence of Kolmogorov-Chentsov criteria for tightness, which says that a stochastic process has a version with continuous paths if the previous inequality holds true.

To complete the proof of part (a), we need to examine (iii), that is

$$\frac{1}{\tilde{n}} \sum_{j=1}^{\tilde{n}} \left(D \log g_j\left(0, \lambda_{[\tilde{n}\tau]}\right)\right)^2 I_{\varepsilon,j} \tag{43}$$

$$= \frac{\sigma_0^2}{2\pi\tilde{n}} \sum_{j=1}^{\tilde{n}} \left(D \log g_j\left(0, \lambda_{[\tilde{n}\tau]}\right)\right)^2 - \frac{1}{\tilde{n}} \sum_{j=1}^{\tilde{n}} \left(D \log g_j\left(0, \lambda_{[\tilde{n}\tau]}\right)\right)^2 \left(I_{\varepsilon,j} - \frac{\sigma_0^2}{2\pi}\right)$$

converges to $A(\tau)$. The second term on the right of (43) is, proceeding as above, easily shown to be such that

$$\sup_{s=0,\ldots,\tilde{n}} \left| \frac{1}{\tilde{n}} \sum_{j=1}^{\tilde{n}} \left(D \log g_j(0, \lambda_s)\right)^2 \left(I_{\varepsilon,j} - \frac{\sigma_0^2}{2\pi}\right) \right| = o_p(1)$$

since we have already shown that $\tilde{n}^{-1/2}\sum_{j=1}^{\tilde{n}} D\log g_j\left(0,\lambda_{[\tilde{n}\tau]}\right)\left(I_{\varepsilon,j}-\frac{\sigma_0^2}{2\pi}\right)$ converges to a Gaussian process, whereas the first term on the right of (43) satisfies that

$$\sup_{s=0,\ldots,\tilde{n}}\left|\frac{1}{\tilde{n}}\sum_{j=1}^{\tilde{n}}\left(D\log g_j\left(0,\lambda_s\right)\right)^2-\frac{1}{\pi}\int_0^{\pi}\left(D\log g\left(\lambda;0,\lambda_s\right)\right)^2 d\lambda\right|=o\left(1\right)$$

by a straightforward extension of Lemma 2 of Robinson (1995*b*).

The conclusion of part (*a*) follows immediately by the continuous mapping theorem as max is a continuous functional in the space $\mathbb{C}\left[0,1\right]$ and proceeding as in Andrews (2001), the finite limit distributions are those from the minimization of

$$\left(c-X\left(\tau\right)\right)'\mathcal{C}^{-1}\left(\tau,\tau\right)\left(c-X\left(\tau\right)\right)$$

with the constraint that $c\geq 0$.

Part (*b*). Here we shall show the properties of the *LM* test. Arguing as in (*a*), it will suffice to show that

$$\frac{1}{\tilde{n}^{1/2}}\sum_{j=1}^{\tilde{n}}\left(\frac{\partial\log g_j\left(0,\lambda_{[\tilde{n}\tau]}\right)}{\partial\alpha}\right)\frac{I_j}{h_j\left(\widehat{\theta}\right)}\Rightarrow\widetilde{X}\left(\tau\right),\tag{44}$$

where $\widehat{\theta}$ is the restricted Whittle estimator of θ_0 given in (16). Now, Taylor's expansion implies that the left side of (44) is

$$\begin{aligned}&\frac{1}{\tilde{n}^{1/2}}\sum_{j=1}^{\tilde{n}}\left(\frac{\partial\log g_j\left(0,\lambda_{[\tilde{n}\tau]}\right)}{\partial\alpha}\right)\frac{I_j}{h_j\left(\theta_0\right)}\\&-\frac{1}{\tilde{n}^{1/2}}\sum_{j=1}^{\tilde{n}}\left(\frac{\partial\log g_j\left(0,\lambda_{[\tilde{n}\tau]}\right)}{\partial\alpha}\right)\left(\frac{\partial\log h_j\left(\overline{\theta}\right)}{\partial\theta'}\right)\frac{I_j}{h_j\left(\overline{\theta}\right)}\left(\widehat{\theta}-\theta_0\right).\end{aligned}$$

On the other hand, by definition of $\widehat{\theta}$,

$$\begin{aligned}0&=\frac{1}{\tilde{n}^{1/2}}\sum_{j=1}^{\tilde{n}}\left(\frac{\partial\log h_j\left(\widehat{\theta}\right)}{\partial\theta}\right)\frac{I_j}{h_j\left(\widehat{\theta}\right)}\\&=\frac{1}{\tilde{n}^{1/2}}\sum_{j=1}^{\tilde{n}}\left(\frac{\partial\log h_j\left(\theta_0\right)}{\partial\theta}\right)\frac{I_j}{h_j\left(\theta_0\right)}\\&-\frac{1}{\tilde{n}^{1/2}}\sum_{j=1}^{\tilde{n}}\left(\frac{\partial\log h_j\left(\overline{\theta}\right)}{\partial\theta}\right)^2\frac{I_j}{h_j\left(\overline{\theta}\right)}\left(\widehat{\theta}-\theta_0\right),\end{aligned}\tag{45}$$

where $\overline{\theta}$ is an intermediate point between $\widehat{\theta}$ and θ_0. Now

$$\sup_{\theta\in\Theta}\left|\frac{1}{\tilde{n}}\sum_{j=1}^{\tilde{n}}\left(\frac{\partial\log h_j\left(\theta\right)}{\partial\theta}\right)^2\frac{I_j}{h_j\left(\theta\right)}-\frac{\sigma_0^2}{2\pi^2}\int_0^{\pi}\left(\frac{\partial\log h\left(\lambda;\theta\right)}{\partial\theta}\right)^2\frac{h\left(\lambda;\theta_0\right)}{h\left(\lambda;\theta\right)}d\lambda\right|\xrightarrow{p}0\tag{46}$$

as we now show. By the triangle inequality, the left side of (46) is bounded by

$$\begin{aligned}&\frac{\sigma_0^2}{2\pi}\sup_{\theta\in\Theta}\left|\frac{1}{\tilde{n}}\sum_{j=1}^{\tilde{n}}\left(\frac{\partial\log h_j\left(\theta\right)}{\partial\theta}\right)^2\frac{h_j\left(\theta_0\right)}{h_j\left(\theta\right)}-\frac{1}{\pi}\int_0^{\pi}\left(\frac{\partial\log h\left(\lambda;\theta\right)}{\partial\theta}\right)^2\frac{h\left(\lambda;\theta_0\right)}{h\left(\lambda;\theta\right)}d\lambda\right|\\&+\sup_{\theta\in\Theta}\left|\frac{1}{\tilde{n}}\sum_{j=1}^{\tilde{n}}\left(\frac{\partial\log h_j\left(\theta\right)}{\partial\theta}\right)^2\frac{h_j\left(\theta_0\right)}{h_j\left(\theta\right)}\left(\frac{I_j}{h_j\left(\theta_0\right)}-\frac{\sigma_0^2}{2\pi}\right)\right|.\end{aligned}\tag{47}$$

That the first term of (47) converges to zero follows by Brillinger (1981, p.15) since $H(\lambda;\theta) = \left(\frac{\partial \log h(\lambda;\theta)}{\partial \theta}\right)^2 \frac{h(\lambda;\theta_0)}{h(\lambda;\theta)}$ is a continuous differentiable function by $C1$. On the other hand, the second term of (47) converges to zero in probability because the finite dimensional distributions of

$$\frac{1}{\tilde{n}^{1/2}} \sum_{j=1}^{\tilde{n}} \left(\frac{\partial \log h_j(\theta)}{\partial \theta}\right)^2 \frac{h_j(\theta_0)}{h_j(\theta)} \left(\frac{I_j}{h_j(\theta_0)} - \frac{\sigma_0^2}{2\pi}\right)$$

converges to a Normal random variable by standard arguments, whereas that $H(\lambda;\theta)$ is a continuous differentiable function will imply that the Kolmogorov-Chentsov's tightness condition will trivially hold true. From here it is standard to show that the second term of (47) is $o_p(1)$.

Next, because $\overline{\theta} \xrightarrow{p} \theta_0$, it implies that in equation (45), after solving for $\left(\widehat{\theta} - \theta_0\right)$, the left side of (44) is equivalent to

$$\left(1, \frac{1}{\tilde{n}} \sum_{j=1}^{\tilde{n}} \left(\frac{\partial \log g_j(0,\cdot)}{\partial \alpha}\right) \left(\frac{\partial \log h_j(\theta)}{\partial \theta'}\right) \left(\frac{1}{\tilde{n}} \sum_{j=1}^{\tilde{n}} \left(\frac{\partial \log h_j(\theta)}{\partial \theta}\right)^2\right)^{-1}\right)$$
$$\times \begin{pmatrix} \frac{1}{\tilde{n}^{1/2}} \sum_{j=1}^{\tilde{n}} \left(\frac{\partial \log g_j(0,\cdot)}{\partial \alpha}\right) \frac{I_j}{h_j(\theta_0)} \\ \frac{1}{\tilde{n}^{1/2}} \sum_{j=1}^{\tilde{n}} \left(\frac{\partial \log h_j(\theta)}{\partial \theta}\right) \frac{I_j}{h_j(\theta_0)} \end{pmatrix}.$$

So, to prove (44) will be equivalent to show the weak convergence of

$$\frac{1}{\tilde{n}^{1/2}} \sum_{j=1}^{\tilde{n}} \left(\frac{\partial \log g_j(0,\cdot)}{\partial \alpha}\right) \frac{I_j}{h_j(\theta_0)},$$

whose proof proceeds exactly as that given in part (a) and so it is omitted. □

5.2. **Proof of Corollary 2.2.** The proof is immediate after one notice that when examining the finite dimensional distributions of $X_n(\tau)$, $\widehat{\alpha}_s = \widehat{\alpha}_s - \alpha(n) + \alpha(n)$ and $n^{1/2}\alpha(n) = \alpha$. Then proceed as in the proof of theorem 2.1 but with $\widehat{\alpha}_s$ replaced by $\widehat{\alpha}_s - \alpha(n)$. □

5.3. **Proof of Theorem 3.1.** We will only examine part (a), since in view of the proof of Theorem 2.1 part (b) that follows almost immediately from part (a). First, by Lemma 6.2, we have that

$$\left(\widehat{\alpha}_s^*, \left(\widehat{\theta}^* - \widehat{\theta}\right)'\right)' \xrightarrow{P^*} 0.$$

On the other hand, proceeding as in Hidalgo and Kreiss (2003), we have that $\tilde{n}^{1/2}\left(\widehat{\alpha}_s^*, \left(\widehat{\theta}^* - \widehat{\theta}\right)'\right)' = O_{p^*}(1)$ and that we further obtain that a similar expansion to that given in (34) holds. That is, writing $\ell_{j,n}^*(\alpha,\theta,s) = g_j^{-1}(\alpha,\lambda_s) h_j^{-1}(\theta) I_j^*$,

$$\begin{aligned} \sum_{j=1}^{\tilde{n}} \ell_{j,n}^*(\alpha,\theta,s) &= \sum_{j=1}^{\tilde{n}} \ell_{j,n}^*\left(0,\widehat{\theta},s\right) + \sum_{j=1}^{\tilde{n}} \left(D\ell_{j,n}^*\left(0,\widehat{\theta},s\right)\right)' \left(\alpha_s, \left(\theta - \widehat{\theta}\right)'\right)' \\ &\quad + \frac{1}{2}\left(\alpha_s, \left(\theta - \widehat{\theta}\right)'\right) \sum_{j=1}^{\tilde{n}} D^2 \ell_{j,n}^*\left(0,\widehat{\theta},s\right) \left(\alpha_s, \left(\theta - \widehat{\theta}\right)'\right)' \\ &\quad + R_n^*(\alpha,\theta,s), \end{aligned}$$

where for all $s = 0, ..., \tilde{n}$, $\sup_{\left\{\tilde{n}^{1/2}|\alpha| \le \xi; n^{1/2}\left\|\theta - \widehat{\theta}\right\| < \xi\right\}} |R_n^*(\alpha,\theta,s)| = o_{p^*}(1)$.

Proceeding as with the proof of Theorem 2.1, and following its arguments, it suffices to show that

$$X_n^*(\tau) = \frac{1}{\tilde{n}^{1/2}} \sum_{j=1}^{\tilde{n}} D \log g_j\left(0, \lambda_{[\tilde{n}\tau]}\right) \left(\frac{I_j^*}{h_j\left(\widehat{\theta}\right)} - \frac{\widehat{\sigma}^2}{2\pi} \right)$$

$$A_n^*(\tau) = \frac{1}{\tilde{n}} \sum_{j=1}^{\tilde{n}} \begin{pmatrix} D \log g_j\left(0, \lambda_{[\tilde{n}\tau]}\right) \\ D \log h_j\left(\widehat{\theta}\right) \end{pmatrix} \left(D \log g_j\left(0, \lambda_{[\tilde{n}\tau]}\right), D \log h_j\left(\widehat{\theta}\right) \right) \frac{I_j^*}{h_j\left(\widehat{\theta}\right)}$$

converge in bootstrap sense to the same processes as

$$X_n(\tau) = \frac{1}{\tilde{n}^{1/2}} \sum_{j=1}^{\tilde{n}} D \log g_j\left(0, \lambda_{[\tilde{n}\tau]}\right) \left(\frac{I_j}{h_j(\theta_0) g_j\left(0, \lambda_{[\tilde{n}\tau]}\right)} - \frac{\sigma_0^2}{2\pi} \right) \text{ and } A_n(\tau)$$

respectively.

We begin with the proof that $X_n^*(\tau)$ converges to the same process as $X_n(\tau)$. To that end, we split the proof into three propositions. Proposition 5.1 shows that $X_{n,\ell}^*(\tau)$ has a covariance structure, conditional on $\underset{\sim}{x}$, that converges in probability to $\mathcal{K}(\tau_1, \tau_2)$ given in (21). Proposition 5.2 shows that the finite dimensional limiting distribution of $X_n^*(\tau)$ is Gaussian centered at zero. Finally, Proposition 5.3 shows tightness of $X_n^*(\tau)$. Thus, combining Propositions 5.1 to 5.3, $X_n^*(\tau) \overset{weakly}{\Longrightarrow} X(\tau)$ in $\mathbb{D}[0,1]$ in probability, as defined by Giné and Zinn (1990).

In what follows for a random variable z, $E^*(z) = E\left(z | \underset{\sim}{x}\right)$, that is the bootstrap expectation of the random variable z.

Proposition 5.1. *Assuming C1-C5,*

$$E^*\left(X_n^*(\tau_1) \;\; X_n^*(\tau_2)\right) \xrightarrow{P} \mathcal{K}(\tau_1, \tau_2). \tag{48}$$

Proof. Let us abbreviate $\frac{\partial}{\partial \alpha} \log g_j(0; \lambda_s)$ by $\psi_j(s)$. First, note that

$$\sum_{t=1}^{n} x_t^* e^{-it\lambda_j} = \widehat{B}\left(e^{-i\lambda_j}\right) \sum_{t=1}^{n} \widetilde{x}_t^* e^{-it\lambda_j}, \tag{49}$$

because by definition of x_t^* and that $\sum_{p=1}^{n} e^{ip\lambda_j} = n\mathcal{I}(j = 0, 2n, ...)$, the left side of (49) is

$$\frac{\widehat{\sigma}^2}{n} \sum_{t=1}^{n} \left(\sum_{\ell=1}^{n} e^{it\lambda_\ell} \widehat{B}\left(e^{-i\lambda_j}\right) \sum_{p=1}^{n} \widetilde{x}_p^* e^{-ip\lambda_j} \right) e^{-it\lambda_j} = \widehat{\sigma}^2 \sum_{q=1}^{n} \widehat{b}_q e^{-iq\lambda_j} \sum_{p=1}^{n} \widetilde{x}_p^* e^{-ip\lambda_j}.$$

Now using (49), $E^*\left(X_n^*(\tau_1) \;\; X_n^*(\tau_2)\right)$ is

$$\frac{1}{\tilde{n}} \sum_{j_1, j_2 = 1}^{\tilde{n}} \psi_{j_1}(\pi\tau_1) \psi_{j_2}(\pi\tau_2) E^* \left\{ \left(\frac{I_{j_1}^* - E^* I_{j_1}^*}{h_{j_1}\left(\widehat{\theta}\right)} \right) \left(\frac{I_{j_2}^* - E^* I_{j_2}^*}{h_{j_2}\left(\widehat{\theta}\right)} \right) \right\}$$

$$- \frac{1}{\tilde{n}} \sum_{j_1, j_2 = 1}^{\tilde{n}} \psi_{j_1}(\pi\tau_1) \psi_{j_2}(\pi\tau_2) \left\{ \left(\frac{E^* I_{j_1}^*}{h_{j_1}\left(\widehat{\theta}\right)} - \frac{\widehat{\sigma}^2}{2\pi} \right) \left(\frac{E^* I_{j2}^*}{h_{j_2}\left(\widehat{\theta}\right)} - \frac{\widehat{\sigma}^2}{2\pi} \right) \right\}. \tag{50}$$

Because $\widetilde{x}_t^*$ is an $iid(0,1)$ sequence of random variables, then the second term of (50) is

$$\frac{1}{\tilde{n}} \sum_{j_1, j_2 = 1}^{\tilde{n}} \psi_{j_1}(\pi\tau_1) \psi_{j_2}(\pi\tau_2) \left\{ \left(\frac{\left|\widehat{B}\left(e^{-i\lambda_{j_1}}\right)\right|^2}{h_{j_1}\left(\widehat{\theta}\right)} - \frac{\widehat{\sigma}^2}{2\pi} \right) \left(\frac{\left|\widehat{B}\left(e^{-i\lambda_{j_2}}\right)\right|^2}{h_{j_2}\left(\widehat{\theta}\right)} - \frac{\widehat{\sigma}^2}{2\pi} \right) \right\}$$

which is equal to zero because

$$\begin{aligned}\sum_{q=1}^{n}\widehat{b}_q e^{-iq\lambda_j} &= \frac{1}{n}\sum_{q=1}^{n}\sum_{\ell=1}^{n} e^{i\ell\lambda_j} h^{1/2}\left(e^{-i\lambda_\ell};\widehat{\theta}\right) e^{-iq\lambda_j} \\ &= \sum_{\ell=1}^{n} h^{1/2}\left(e^{-i\lambda_\ell};\widehat{\theta}\right)\frac{1}{n}\sum_{q=1}^{n} e^{iq\lambda_{\ell-j}} \qquad (51)\\ &= h^{1/2}\left(e^{-i\lambda_j};\widehat{\theta}\right).\end{aligned}$$

Hence, to conclude the proof, we need to show that the first term of (50) converges in probability to $\mathcal{K}(\tau_1,\tau_2)$. But because $\widetilde{x}_t^*$ is an $iid(0,1)$ sequence, by Brockwell and Davis's (1991) Proposition 10.3.2, this term is

$$\begin{aligned}&\frac{1}{n}\sum_{j=1}^{\tilde{n}}\psi_j(\pi\tau_1)\psi_j(\pi\tau_2) + \frac{\widehat{\kappa}_4}{\tilde{n}^2}\sum_{j_1,j_2=1}^{\tilde{n}}\psi_{j_1}(\pi\tau_1)\psi_{j_2}(\pi\tau_2)\\ =\;& \mathcal{K}(\tau_1,\tau_2)\left(1+O\left(\tilde{n}^{-1}\log\tilde{n}\right)\right) + \frac{\widehat{\kappa}_4}{\tilde{n}^2}\sum_{j_1,j_2=1}^{\tilde{n}}\psi_{j_1}(\pi\tau_1)\psi_{j_2}(\pi\tau_2),\end{aligned}$$

by a straightforward modification of Lemma 2 of Robinson (1995*b*). However, because by a well known argument (see Stout's 1974, Theorem 3.5.8) $C1$ implies that x_t is ergodic, we have that

$$\widehat{\kappa}_4 = \frac{1}{n}\sum_{t=1}^{n} x_t^4 - 3\left(\frac{1}{n}\sum_{t=1}^{n} x_t^2\right)^2 \xrightarrow{p} \kappa_4.$$

From here, the conclusion follows because

$$\frac{1}{\tilde{n}}\sum_{j=1}^{\tilde{n}}\psi_j(\pi\tau) - \frac{1}{\pi}\int_0^{\pi}\log\left|4\sin\left(\frac{\lambda-\pi\tau}{2}\right)\sin\left(\frac{\lambda+\pi\tau}{2}\right)\right| d\lambda = O\left(\tilde{n}^{-1}\log\tilde{n}\right)$$

by a straightforward modification of Lemma 2 of Robinson (1995*b*) and then that condition (12) implies that $\int_0^{\pi}\log\left|4\sin\left(\frac{\lambda-\pi\tau}{2}\right)\sin\left(\frac{\lambda+\pi\tau}{2}\right)\right| d\lambda = 0$ for all τ. □

Proposition 5.2. *Under the same conditions of Proposition 5.1, the finite dimensional distributions of R_n^* converge in bootstrap law to those of a centered Gaussian process.*

Proof. Fix $\tau_1,...,\tau_q$ and constants $a_1,...,a_q$. By Cramér-Rao device, it suffices to examine the limit distribution of

$$\sum_{p=1}^{q} a_p\left(\frac{1}{\tilde{n}^{1/2}}\sum_{j=1}^{\tilde{n}}\psi_j(\tau_p)\left(\frac{2\pi I_j^*}{h_j\left(\widehat{\theta}\right)} - \widehat{\sigma}^2\right)\right). \qquad (52)$$

By Proposition 5.1, the (bootstrap) second moment of (52) converges in probability to

$$\sum_{p_1,p_2=1}^{q} a_{p_1}a_{p_2}\mathcal{K}(\tau_{p_1},\tau_{p_2}).$$

So, to complete the proof it remains to verify that (52) satisfies the Lindeberg's condition, that is $\forall\delta>0$,

$$\frac{1}{\tilde{n}}\sum_{j=1}^{\tilde{n}} E^*\left|\psi_j(\tau_p)\left(\frac{2\pi I_j^*}{h_j\left(\widehat{\theta}\right)} - \widehat{\sigma}^2\right)\right|^2 \mathcal{I}\left(\frac{1}{\tilde{n}}\left|\psi_j(\tau_p)\left(\frac{2\pi I_j^*}{h_j\left(\widehat{\theta}\right)} - \widehat{\sigma}^2\right)\right|^2 > \delta\right) \xrightarrow{P} 0,$$

or the sufficient condition

$$\tilde{n}^{-1}E^*\left|\left(\frac{2\pi I_j^*}{h_j\left(\widehat{\theta}\right)}-\widehat{\sigma}^2\right)\right|^4\frac{1}{\tilde{n}}\sum_{j=1}^{\tilde{n}}\left|\psi_j\left(\tau_p\right)\right|^4\xrightarrow{P}0.$$

But this is the case since proceeding as above, the left side of the last displayed expression is

$$\tilde{n}^{-1}\log^4\tilde{n}E^*\left|\left(\frac{1}{2\pi n}\sum_{t=1}^{n}\widetilde{x}_t^*e^{-it\lambda_j}-1\right)\right|^4\widehat{\sigma}^2=O_p\left(\sum_{p=1}^{n}\widetilde{x}_p^*e^{-ip\lambda_j}\right)$$

by Brillinger (1981), since $\widetilde{x}_t^*$ is an $iid\left(0,1\right)$ sequence of random variables and x_t is ergodic in that $n^{-1}\sum_{t=1}^{n}\left|x_t\right|^r-E\left|x_t\right|^r=o_p\left(1\right)$ for $r=1,...,8$, and $\widehat{\sigma}^2=O_p\left(1\right)$. □

Proposition 5.3. *Under the same conditions of Proposition 5.1, conditional on* $\underset{\sim}{x}$, $R_n^*\left(\tau\right)$ *is tight.*

Proof. Denote $\eta_j^*=n^{-1}\sum_{t=1}^{n}\widetilde{x}_t^*e^{-it\lambda_j}-1$. Proceeding as with the proof of Theorem 2.1, we only need to check the Kolmogorov-Chentsov's condition. That is

$$E^*\left|\frac{1}{\tilde{n}}\sum_{j=1}^{\tilde{n}}\left(\psi_j\left(\tau_1\right)-\psi_j\left(\tau_2\right)\right)\left(\frac{I_j^*}{h_j\left(\widehat{\theta}\right)}-\frac{\widehat{\sigma}^2}{2\pi}\right)\right|^4\leq G_n\left(\tau_1,\tau_2\right)\left|\tau_1-\tau_2\right|^{1+\delta},\quad(53)$$

for some $\delta>0$ and where $G_n\left(\tau_1,\tau_2\right)$ is bounded in probability. Now, by definition of I_j^*, the left side of (53) is

$$\begin{aligned}&\left(\frac{\widehat{\sigma}^2}{2\pi}\right)^4E^*\left|\frac{1}{\tilde{n}}\sum_{j=1}^{\tilde{n}}\left(\psi_j\left(\tau_1\right)-\psi_j\left(\tau_2\right)\right)\eta_j^*\right|^4\\=&\left(\frac{\widehat{\sigma}^2}{2\pi}\right)^4\frac{1}{n^2}\sum_{j_1,...,j_4=1}^{\tilde{n}}\prod_{\ell=1}^{4}\left(\log\frac{\left|\sin\left(\frac{\lambda_{j_\ell}-\lambda_{[\pi\tau_1]}}{2}\right)\right|}{\left|\sin\left(\frac{\lambda_{j_\ell}-\lambda_{[\pi\tau_2]}}{2}\right)\right|}\right)E\left\{\prod_{\ell=1}^{4}\eta_{j_\ell}^*\right\}\\&+\left(\frac{\widehat{\sigma}^2}{2\pi}\right)^4\frac{1}{n^2}\sum_{j_1,...,j_4=1}^{\tilde{n}}\prod_{\ell=1}^{4}\left(\log\frac{\left|\sin\left(\frac{\lambda_{j_\ell}+\lambda_{[\pi\tau_1]}}{2}\right)\right|}{\left|\sin\left(\frac{\lambda_{j_\ell}+\lambda_{[\pi\tau_2]}}{2}\right)\right|}\right)E\left\{\prod_{\ell=1}^{4}\eta_{j_\ell}^*\right\}.\end{aligned}$$

We examine only the first term on the right, the second being identically handled. That term is

$$\begin{aligned}&=\left(\frac{\widehat{\sigma}^2}{2\pi}\right)^4\frac{3}{n^2}\sum_{j_1,...,j_4=1}^{\tilde{n}}\prod_{\ell=1}^{4}\left(\log\frac{\left|\sin\left(\frac{\lambda_{j_\ell}-\lambda_{[\pi\tau_1]}}{2}\right)\right|}{\left|\sin\left(\frac{\lambda_{j_\ell}-\lambda_{[\pi\tau_2]}}{2}\right)\right|}\right)E\left\{\eta_{j_1}^*\eta_{j_2}^*\right\}E\left\{\eta_{j_3}^*\eta_{j_4}^*\right\}\\&+\left(\frac{\widehat{\sigma}^2}{2\pi}\right)^4\frac{1}{n^2}\sum_{j_1,...,j_4=1}^{\tilde{n}}\prod_{\ell=1}^{4}\left(\log\frac{\left|\sin\left(\frac{\lambda_{j_\ell}-\lambda_{[\pi\tau_1]}}{2}\right)\right|}{\left|\sin\left(\frac{\lambda_{j_\ell}-\lambda_{[\pi\tau_2]}}{2}\right)\right|}\right)Cum^*\left(\eta_{j_1}^*,\eta_{j_2}^*,\eta_{j_3}^*,\eta_{j_4}^*\right).\quad(54)\end{aligned}$$

By Brillinger's (1981) Theorems 2.3.1 and 4.3.2, and in particular equation (4.3.15), and the integrability of $\log^q\left|x\right|$, $q=1,...,4$, we obtain that the second term on the right of (54) is bounded in absolute value by

$$K\tilde{n}^{-2}\log^2\tilde{n}\left(\frac{\widehat{\sigma}^2}{2\pi}\right)^4\leq K\left(\frac{\widehat{\sigma}^2}{2\pi}\right)^4\left(\tau_2-\tau_1\right)^{1+\delta}\leq G_n\left(\tau_1,\tau_2\right)\left(\tau_2-\tau_1\right)^{1+\delta},$$

for any $0<\delta<1$. So, to complete the proof we need to examine the first term on the right of (54), which is

$$3\left(\frac{\widehat{\sigma}^2}{2\pi}\right)^4\left[\frac{1}{\tilde{n}}\sum_{j=1}^{\tilde{n}}\left(\log\frac{\left|\sin\left(\frac{\lambda_j-\lambda_{[\pi\tau_1]}}{2}\right)\right|}{\left|\sin\left(\frac{\lambda_j-\lambda_{[\pi\tau_2]}}{2}\right)\right|}\right)^2\right]^2\leq G_n\left(\tau_1,\tau_2\right)\left|\tau_1-\tau_2\right|^{1+\delta}$$

proceeding as in Theorem 2.1 and that $\widehat{\sigma}^2$ is bounded in probability.

So, we have completed the proof that $X_n^*(\tau)$ converges in bootstrap to the same process as that $X_n(\tau)$. Next, we shall show that $A_n^*(\tau)$ converges in bootstrap to $A(\tau)$, but the proof is immediate following the ideas in Proposition 5.1 and those in Theorem 2.1. So, the proof is omitted. From here the conclusion of the Theorem part (a) follows by continuous mapping theorem and that the finite limit distributions are those from the minimization of

$$\left(c-X\left(\tau\right)\right)'\mathcal{C}^{-1}\left(\tau,\tau\right)\left(c-X\left(\tau\right)\right)$$

with the constraint that $c\geq 0$.

The proof under H_1 is exactly the same as that of Theorem 3.1 with the only difference that instead of writing $\widehat{\theta}-\theta_0=o_p(1)$ we write $\widehat{\theta}-\theta_1=o_p(1)$ and θ_1 instead of θ_0. □

6. TECHNICAL LEMMAS

To simplify the notation, henceforth, we assume without loss of generality that $\sigma_0^2=1$.

Lemma 6.1. *Let X_t be a linear process satisfying C1, C2 and C5, and let $\zeta(s):[0,\pi]\to\mathbb{R}^p$ be a continuous differentiable function for $|\lambda-\lambda_s|>0$ and as $\lambda\to\lambda_s$ $|\zeta(\lambda,s)|\leq K|\log|\lambda-\lambda_s||$. Then, as $n\to\infty$,*

$$\sup_{s=0,\ldots,\tilde{n}}\left|\frac{2\pi}{\tilde{n}^{1/2}}\sum_{j=1}^{\tilde{n}}\zeta_j(s)\left\{\frac{I_j}{h_j}-I_{\varepsilon,j}\right\}\right|=o_p(1).$$

Proof. Let $u_j:=h(\lambda_j;\theta_0)^{-1/2}w_x(\lambda_j)$, $v_j:=w_\varepsilon(\lambda_j)$, where for a generic sequence z_t $w_z(\lambda_j):=(2\pi n)^{-1/2}\sum_{t=1}^n z_t e^{it\lambda_j}$. The lemma is shown if (i) the finite dimensional distributions converge to zero in probability and (ii)

$$\frac{1}{\tilde{n}^{1/2}}\sum_{j=1}^{\tilde{n}}\zeta_j(s)\left\{\frac{I_j}{h_j}-I_{\varepsilon,j}\right\}\tag{55}$$

is tight. We shall begin with (i). Proceeding as in Robinson (1995*b*), the second moment of the last displayed expression is

$$E\left(\frac{1}{\tilde{n}^{1/2}}\sum_{j=1}^{\tilde{n}}\zeta_j(s)\left\{\frac{I_j}{h_j}-I_{\varepsilon,j}\right\}\right)^2=\frac{2\pi}{\tilde{n}}\left(a_1+a_2+b_1+b_2\right),\tag{56}$$

where

$$\begin{aligned}a_1 &= \sum_{j=1}^n\zeta_j^2(s)\left\{2\left(E|u_j|^2\right)^2+\left|E\left(u_j^2\right)\right|^2-2\left|E\left(u_jv_j\right)\right|^2-2\left|E\left(u_j\overline{v}_j\right)\right|^2\right.\\ &\quad\left.-2E|u_j|^2E|v_j|^2+2\left(E|v_j|^2\right)^2+\left|E\left(v_j^2\right)\right|^2\right\},\end{aligned}$$

$$a_2=\sum_{j=1}^n\zeta_j^2(s)\left\{cum\left(u_j,u_j,\overline{u}_j,\overline{u}_j\right)-2cum\left(u_j,v_j,\overline{u}_j,\overline{v}_j\right)+cum\left(v_j,v_j,\overline{v}_j,\overline{v}_j\right)\right\},$$

$$\begin{aligned} b_1 &= 2\sum_{1=j<k}^{n}\sum \zeta_j(s)\,\zeta_k(s)\Big\{E|u_j|^2 E|u_k|^2 + E|u_j u_k|^2 + E|u_j\overline{u}_k|^2 - E|u_j|^2 E|v_k|^2 \\ &\quad - |E(u_j v_k)|^2 - |Eu_j\overline{v}_k|^2 - E|u_k|^2 E|v_j|^2 - |E(u_k v_j)|^2 \\ &\quad |E(u_k\overline{v}_j)|^2 + E|v_j|^2 E|v_k|^2 + |E(v_j v_k)|^2 + |E(v_j\overline{v}_k)|\Big\}, \end{aligned}$$

$$\begin{aligned} b_2 &= 2\sum_{1=j<k}^{n}\sum \zeta_j(s)\,\zeta_k(s)\{cum(u_j,u_k,\overline{u}_j,\overline{u}_k) - cum(u_j,v_k,\overline{u}_j,\overline{v}_k) \\ &\quad -cum(v_j,u_k,\overline{v}_j,\overline{u}_k) + cum(v_j,v_k,\overline{v}_j,\overline{v}_k)\}, \end{aligned}$$

where $cum(\cdot,\cdot,\cdot,\cdot)$ is the joint cumulant of the argument random variables.

Because $E|v_j|^2 = 1$,

$$\begin{aligned} a_1 &= \sum_{j=1}^{n}\zeta_j^2(s)\Big\{2\left(E|u_j|^2-1\right)^2 + 2\left(E|u_j|^2-1\right) + \left|E\left(u_j^2\right)\right|^2 - 2|E(u_j v_j)|^2 \\ &\quad -2|E(u_j\overline{v}_j)-1|^2 - 2\left(E(u_j\overline{v}_j)-1\right) - 2\left(E(\overline{u}_j v_j - 1) + \left|E\left(v_j^2\right)\right|^2\right)\Big\} \\ &\leq K\frac{1}{\tilde{n}}\sum_{j=1}^{n}\zeta_j^2(s) = O(1) \end{aligned}$$

$$\begin{aligned} b_1 &= 2\sum_{1=j<k}^{n}\sum \zeta_j(s)\,\zeta_k(s)\Big\{\left(E|u_j|^2-1\right)\left(E|u_k|^2-1\right) + |E(u_j u_k)|^2 + |E(u_j\overline{u}_k)|^2 \\ &\quad - |E(u_j v_k)|^2 - |E(u_j\overline{v}_k)|^2 - |E(u_k v_j)|^2 - |E(u_k\overline{v}_j)|^2 \\ &\quad + |E(v_j v_k)|^2 + |E(v_j\overline{v}_k)|^2\Big\} \\ &\leq K\frac{1}{\tilde{n}^2}\sum_{1=j<k}^{n}\sum \left|\zeta_j(s)\,\zeta_k(s)\right| \end{aligned}$$

as $n\to\infty$ by Brillinger's (1981) Theorem 4.3.2. With regard to a_2 and b_2, proceeding as in Robinson (1995*b*, p.1649) but noting that $h(\lambda)$ is continuously differentiable, we can conclude that

$$a_2 \leq K\frac{1}{\tilde{n}}\sum_{j=1}^{n}\zeta_j^2(s); \quad b_2 \leq K\frac{1}{\tilde{n}^2}\sum_{1=j<k}^{n}\sum \left|\zeta_j(s)\,\zeta_k(s)\right|.$$

Hence, using (56) we have that

$$E\left|\frac{1}{\tilde{n}^{1/2}}\sum_{j=1}^{\tilde{n}}\zeta_j(s)\left\{\frac{I_j}{h_j} - I_{\varepsilon,j}\right\}\right|^2 \leq \frac{K}{\tilde{n}}\left(\frac{1}{\tilde{n}}\sum_{j=1}^{n}\zeta_j^2(s) + \frac{1}{\tilde{n}^2}\sum_{1=j<k}^{n}\sum\left|\zeta_j(s)\,\zeta_k(s)\right|\right),$$

so that by Markov's inequality, the finite dimensional distributions of (55) converge to zero in probability. So, we are left to show (ii), that is the tightness condition. But this follows after observing that after replacing $\zeta_j(s)$ by $\zeta_j(s_1)-\zeta_j(s_2)$, we have that

$$E\left|\frac{1}{\tilde{n}^{1/2}}\sum_{j=1}^{\tilde{n}}\left(\zeta_j(s_1)-\zeta_j(s_2)\right)\left\{\frac{I_j}{h_j} - I_{\varepsilon,j}\right\}\right|^2 \tag{57}$$

$$\leq \frac{K}{\tilde{n}^2}\left(\sum_{j=1}^{n}\left(\zeta_j(s_1)-\zeta_j(s_2)\right)^2 + \frac{1}{\tilde{n}}\sum_{1=j<k}^{n}\sum\left|\left(\zeta_j(s_1)-\zeta_j(s_2)\right)\left(\zeta_j(s_1)-\zeta_j(s_2)\right)\right|\right).$$

Now, since we can assume, without loss of generality that $\tilde{n}^{-1} \leq \tau_2 - \tau_1$, where $\tau_j = \lim_{n\to\infty} s_j/\tilde{n}$,

$$\frac{1}{\tilde{n}}\sum_{j=1}^{s_1-1}\left(\log\frac{\left|\sin\left(\frac{\lambda_{s_1}-\lambda_j}{2}\right)\right|}{\left|\sin\left(\frac{\lambda_{s_2}-\lambda_j}{2}\right)\right|}\right)^2+\frac{1}{\tilde{n}}\sum_{j=s_2+1}^{\tilde{n}}\left(\log\frac{\left|\sin\left(\frac{\lambda_j-\lambda_{s_1}}{2}\right)\right|}{\left|\sin\left(\frac{\lambda_j-\lambda_{s_2}}{2}\right)\right|}\right)^2 \tag{58}$$

$$+\frac{2}{\tilde{n}}\sum_{j=s_1+1}^{s_2-1}\left(\log\frac{\left|\sin\left(\frac{\lambda_j-\lambda_{s_1}}{2}\right)\right|}{\left|\sin\left(\frac{\lambda_{s_2}-\lambda_j}{2}\right)\right|}\right)^2.$$

First, observing that in the interval $(0,\lambda_{s_1})$, we have that

$$\begin{aligned}&\left(\log\left|\sin\left(\frac{\lambda_{s_1}-\lambda_j}{2}\right)\right|-\log\left|\sin\left(\frac{\lambda_{s_2}-\lambda_j}{2}\right)\right|\right)^2\\ \leq\ &\log^2\left|\sin\left(\frac{\lambda_{s_1}-\lambda_j}{2}\right)\right|-\log^2\left|\sin\left(\frac{\lambda_{s_2}-\lambda_j}{2}\right)\right|,\end{aligned}$$

we obtain that the first term of (58) is bounded by

$$\begin{aligned}&\frac{1}{\tilde{n}}\sum_{j=1}^{s_1-1}\log^2\left|\sin\left(\frac{\lambda_{s_1}-\lambda_j}{2}\right)\right|du-\int_0^{\tau_1}\log^2\left|\sin\left(\frac{\lambda_{s_2}-\lambda_j}{2}\right)\right|\\ =\ &H(s_1)-H(s_2)+H(s_2-s_1)\\ \leq\ &2H(s_2-s_1),\end{aligned}$$

where $H(s)=\tilde{n}^{-1}\sum_{j=1}^{s-1}\log^2|\sin(\lambda_j/2)|$, and because $\log^2|\sin((\lambda_{s_1}-\lambda_j)/2)|-\log^2|\sin((\lambda_{s_2}-\lambda_j)/2)|>0$ implies that $H(s_2)-H(s_1)<H(s_2-s_1)$. But $H(s)\leq K\tau^{(1+\delta)/2}$ for $0\leq\tau\leq 1$, so we conclude that the first term of (58) is bounded by

$$K(\tau_2-\tau_1)^{(1+\delta)/2}.$$

Next, we examine the second term of (58). Using that in the interval $(\tau_2,1/2)$,

$$\left(\log\left|\sin\left(\frac{\lambda_j-\lambda_{s_1}}{2}\right)\right|-\log\left(\frac{\lambda_j-\lambda_{s_2}}{2}\right)\right)^2\leq\log^2\left|\frac{\lambda_j-\lambda_{s_2}}{2}\right|-\log^2\left|\sin\left(\frac{\lambda_j-\lambda_{s_1}}{2}\right)\right|,$$

we have that the second term of (58) is bounded by

$$\begin{aligned}0\ <\ &\frac{1}{\tilde{n}}\sum_{j=s_2+1}^{\tilde{n}}\log^2\left|\sin\left(\frac{\lambda_j-\lambda_{s_2}}{2}\right)\right|-\frac{1}{\tilde{n}}\sum_{j=s_2+1}^{\tilde{n}}\log^2\left|\sin\left(\frac{\lambda_j-\lambda_{s_1}}{2}\right)\right|\\ =\ &H(\tilde{n}-s_2)-H(\tilde{n}-s_1)+H(s_2-s_1)\\ \leq\ &2H(s_2-s_1)\end{aligned}$$

because $H(\tilde{n}-s_1)-H(\tilde{n}-s_2)\leq H(s_2-s_1)$. So, the second term of (58) is also bounded by $K(\tau_2-\tau_1)^{(1+\delta)/2}$. Finally, by a change of subindexes, the third term of (58) is

$$\frac{2}{\tilde{n}}\sum_{j=1}^{s_2-s_1-1}\left(\log\left|\sin\left(\frac{\lambda_j}{2}\right)\right|\right)^2=H(s_2-s_1)\leq K(\tau_2-\tau_1)^{(1+\delta)/2}.$$

Hence, we have shown that (57) is bounded by

$$K(\tau_2-\tau_1)^{\frac{1+\delta}{2}+1}+K(\tau_2-\tau_1)^{1+\delta},$$

which implies that the Kolmogorov-Chentsov's inequality holds true and hence that (55) is tight. From here the conclusion of the lemma is standard. □

Let θ_1 be the value which minimizes

$$\int_{-\pi}^{\pi} \frac{f\left(\lambda;\alpha_0,\theta_0,\lambda^0\right)}{f\left(\lambda;0,\theta,\lambda^0\right)}d\lambda.$$

Then, we have that

Lemma 6.2. *Let $\widehat{\theta}$ be such that it converges almost surely to $\theta_1 \in \Theta$. Then*

$$\theta^* - \widehat{\theta} = o_{p^*}(1).$$

Proof. Because by construction, conditional on the sample $\underset{\sim}{x}$, x_t^* is a zero mean iid sequence of random variables with unit variance, then it is an ergodic sequence in a quadratic mean sense. Then, proceeding as in the proof of Lemma 1 of Hannan (1973), we have that uniformly in $\theta \in \Theta$,

$$\frac{1}{n}\sum_{j=1-\tilde{n}}^{\tilde{n}} \frac{I_{n,j}^*}{f_j(\theta)} - \int_{-\pi}^{\pi} \frac{f\left(\lambda;\widehat{\theta}\right)}{f(\lambda;\theta)} d\lambda \xrightarrow{P} 0.$$

Now, proceeding as in the proof of Hannan's (1973) Theorem 1, we conclude that $\theta^* - \widehat{\theta} = o_{p^*}(1)$ because

$$\widehat{\theta} = \arg\min_{\theta\in\Theta} \int_{-\pi}^{\pi} \frac{f\left(\lambda;\widehat{\theta}\right)}{f(\lambda;\theta)} d\lambda,$$

which concludes the proof. □

7. CONCLUSIONS

In the paper we have described two tests for the hypothesis of the presence of strong cycles and as a by-product for weak dependence in linear models. The first one is a Wald type test whereas the second one has similarities to the LM principle. Because of the nonstandard limiting distribution of our tests, we have explored a bootstrap scheme in the frequency domain. Our bootstrap algorithm can be regarded as a frequency domain counterpart to residual-based bootstraps. We have also described some possible advantages over the latter method in terms of computation and simplicity. One possible drawback or criticism, when compared to residual-based bootstraps, is that we do our resampling from the standardized original data x_t and not from the residuals/innovations ε_t. Because of that, we envisage that we can modify our algorithm to allow bootstrapping from ε_t, or $\widehat{\varepsilon}_t$ (an estimate of ε_t), as follows.

For $t = 1, ..., n$, compute

$$\widehat{\varepsilon}_t = \frac{1}{n^{1/2}} \sum_{j=1}^{n} e^{it\lambda_j} \widehat{B}\left(e^{i\lambda_j}\right) \left|\widehat{B}\left(e^{-i\lambda_j}\right)\right|^{-2} w_x(\lambda_j),$$

and obtain the standardized residuals

$$\widetilde{\varepsilon}_t = \widetilde{\sigma}_{\varepsilon}^{-1}\left(\widehat{\varepsilon}_t - \overline{\widehat{\varepsilon}}\right), \quad \widetilde{\sigma}_{\varepsilon}^2 = \frac{1}{n}\sum_{t=1}^{n}\left(\widehat{\varepsilon}_t - \overline{\widehat{\varepsilon}}\right)^2, \quad \overline{\widehat{\varepsilon}} = \frac{1}{n}\sum_{t=1}^{n}\widehat{\varepsilon}_t.$$

Then in *STEP 1*, instead of resampling with replacement from $\check{x}_t$ to obtain $\underset{\sim}{\widetilde{x}}^* = (\widetilde{x}_1^*, \widetilde{x}_2^*, ..., \widetilde{x}_n^*)'$, we shall do the same but with $\widetilde{\varepsilon}_t$ obtaining a bootstrap sample $\underset{\sim}{\widetilde{\varepsilon}}^* = (\widetilde{\varepsilon}_1^*, \widetilde{\varepsilon}_2^*, ..., \widetilde{\varepsilon}_n^*)'$. Then, η_j^* in *STEP 1* is replaced by

$$\eta_j^* = \frac{1}{n^{1/2}} \sum_{t=1}^{n} \widetilde{\varepsilon}_t^* e^{-it\lambda_j}, \quad j = 1, ..., n.$$

Hereafter, proceed as in *STEPS 2 to 4*. This method to bootstrap the data x_t, $t = 1, ..., n$, may be preferable over that given in Section 3 as it may capture higher order moment properties of $\{x_t\}_{t\in\mathbb{Z}}$ than the bootstrap described in the paper.

Although we have only considered the situation when x_t is observed, it appears that the same results should hold true when $\{x_t\}_{t\in\mathbb{Z}}$ are the errors of a regression model. That is, consider the following linear regression model

$$y_t = \beta' z_t + x_t, \quad (t = 1, ..., n)$$

where $\{x_t\}_{t\in\mathbb{Z}}$ follows a $FARIMA(p, d_0, q)$. When $d_0 = 0$, that is weak dependence, it is well known that under suitable regularity conditions, the GLS estimator of β, say $\widehat{\beta}$, and the Whittle estimator of the parameters of the $ARMA(p, q)$ process of $\{x_t\}_{t\in\mathbb{Z}}$ are (asymptotically) independent. More recently, Robinson and Hidalgo (1997) have shown that the same holds when $\{x_t\}_{t\in\mathbb{Z}}$ exhibits strong dependence. So, this result leads us to think that the results obtained in the paper would hold true when $\widehat{x}_t = y_t - \widehat{\beta}' z_t$ is used instead of the unobserved x_t.

We finish this section mentioning two issues. Although we have only considered stationary alternatives, that is $\alpha_0 < 1$, following results by Velasco and Robinson (2000), our tests should detect also nonstationary, $\alpha_0 > 1$, alternatives. The reason comes from the observation that using the taper periodogram we have by Velasco and Robinson (2000), that $n^{1/2}(\widehat{\alpha} - \alpha_0)$ converges in distribution to a normal random variable. So, this result indicates that the same type of results obtained in Section 2 will hold under nonstationary alternatives. A second issue is what would happen with our Wald and LM tests when we allow for asymmetric strong dependence, that is when

$$f(\lambda) \sim \begin{cases} K_1 \left(\lambda - \lambda^0\right)^{-\alpha_1} & \lambda \to \lambda^0+ \\ K_2 \left(\lambda^0 - \lambda\right)^{-\alpha_2} & \lambda \to \lambda^0 - , \end{cases}$$

with α_1 possibly different than α_2. Unfortunately, we are unaware of any result for the Whittle estimator of the parameters for the latter type of dependence. Definitely this is an open problem. However, we should observe that under the null $\alpha_1 = \alpha_2 = 0$. So, we may conjecture that our tests may deliver the correct answer under the null hypothesis. Note that all what we need for the consistency is that, say, $\widehat{\alpha}$ converges to a positive quantity under the alternative.

TABLE 4.1

SIZE OF THE WALD TEST True model White noise				SIZE OF THE LM TEST True model White noise			
	$d_0 = 0.0$				$d_0 = 0.0$		
	β				β		
n	10%	5%	1%		10%	5%	1%
n=64	10.65	10.10	10.10	n=64	10.60	5.00	1.20
n=128	11.00	5.15	4.15	n=128	9.65	5.10	1.40

TABLE 4.1 (Cont.)

SIZE OF THE LM TEST True model AR(1) METHOD 1				SIZE OF THE LM TEST True model AR(1) METHOD 2				SIZE OF THE LM TEST True model MA(1) METHOD 1				SIZE OF THE LM TEST True model MA(1) METHOD 2			
	$d_0 = 0.0$				$d_0 = 0.0$				$d_0 = 0.0$				$d_0 = 0.0$		
	β				β				β				β		
	10%	5%	1%		10%	5%	1%		10%	5%	1%		10%	5%	1%
n=64	10.05	5.40	1.35	n=64	10.05	5.55	1.25	n=64	9.50	5.55	1.10	n=64	9.85	5.35	1.05
n=128	9.65	4.30	1.45	n=128	10.20	4.60	1.50	n=128	10.00	4.80	1.60	n=128	9.80	4.70	1.60

TABLE 4.2

POWER OF THE WALD TEST.

True model ARFIMA(0,d_0,0))

$d_0 = 0.10$				$d_0 = 0.20$				$d_0 = 0.30$				$d_0 = 0.40$			
	β				β				β				β		
	10%	5%	1%		10%	5%	1%	n	10%	5%	1%	n	10%	5%	1%
n=64	11.8	11.4	11.4	n=64	26.95	15.20	3.55	64	55.20	39.30	14.20	64	81.50	68.45	39.30
n=128	15.1	7.95	6.45	n=128	49.80	34.40	11.50	128	86.60	75.30	47.65	128	98.75	96.60	86.25

TABLE 4.3

POWER OF THE LM TEST.
True model ARFIMA(0,d_0,0))

$d_0 = 0.10$				$d_0 = 0.20$				$d_0 = 0.30$				$d_0 = 0.40$			
	β				β				β				β		
n	10%	5%	1%	n	10%	5%	1%	n	10%	5%	1%	n	10%	5%	1%
64	14.60	8.05	2.00	64	38.10	26.75	12.90	64	66.70	57.75	38.50	64	86.25	81.15	66.60
128	22.90	15.05	5.45	128	68.45	57.55	38.00	128	94.45	91.70	81.10	128	99.35	98.95	97.50

TABLE 4.3

POWER OF THE LM TEST.
True model is a GARMA(0,d_0,0))

$d_0 = 0.10$				$d_0 = 0.20$				$d_0 = 0.30$				$d_0 = 0.40$			
	β				β				β				β		
n	10%	5%	1%	n	10%	5%	1%	n	10%	5%	1%	n	10%	5%	1%
64	18.10	9.60	2.20	64	47.45	36.85	17.90	64	80.65	73.75	56.90	64	96.85	94.40	88.35
128	28.55	20.00	7.30	128	76.55	67.85	49.35	128	97.55	96.05	90.90	128	99.90	99.90	99.55

TABLE 4.3 (Cont.)

POWER OF THE LM TEST.
True model ARFIMA(1,d_0,0).
METHOD 1

$d_0 = 0.10$				$d_0 = 0.20$				$d_0 = 0.30$				$d_0 = 0.40$			
	β				β				β				β		
n	10%	5%	1%	n	10%	5%	1%	n	10%	5%	1%	n	10%	5%	1%
64	11.30	6.20	1.10	64	11.85	6.35	1.20	64	13.65	7.45	1.45	64	15.55	9.00	2.05
128	11.75	6.15	1.05	128	11.95	6.75	1.15	128	12.85	7.30	1.45	128	17.60	10.20	2.60

TABLE 4.3 (Cont.)

POWER OF THE LM TEST.
True model ARFIMA(1,d_0,0).
METHOD 2

$d_0 = 0.10$				$d_0 = 0.20$				$d_0 = 0.30$				$d_0 = 0.40$			
	β				β				β				β		
n	10%	5%	1%	n	10%	5%	1%	n	10%	5%	1%	n	10%	5%	1%
64	11.60	6.05	1.20	64	12.20	6.70	1.25	64	13.60	7.65	1.35	64	15.90	9.75	1.95
128	11.65	6.30	0.95	128	12.05	6.95	1.05	128	13.35	7.50	1.50	128	17.85	10.40	2.85

TABLE 4.3 (Cont.)

POWER OF THE LM TEST.
True model GARMA(1,d_0,0).
METHOD 1

$d_0 = 0.10$				$d_0 = 0.20$				$d_0 = 0.30$				$d_0 = 0.40$			
	β				β				β				β		
n	10%	5%	1%	n	10%	5%	1%	n	10%	5%	1%	n	10%	5%	1%
64	19.70	10.85	2.80	64	46.85	34.50	16.55	64	79.15	71.60	51.75	64	95.90	93.65	84.35
128	27.05	18.10	5.95	128	72.55	62.65	43.85	128	96.55	94.95	88.05	128	99.90	99.85	99.15

TABLE 4.3 (Cont.)

POWER OF THE LM TEST.
True model GARMA(1,d_0,0)
METHOD 2

$d_0 = 0.10$				$d_0 = 0.20$				$d_0 = 0.30$				$d_0 = 0.40$			
	β				β				β				β		
n	10%	5%	1%	n	10%	5%	1%	n	10%	5%	1%	n	10%	5%	1%
64	19.65	10.70	2.70	64	47.00	34.80	16.00	64	79.00	71.15	52.20	64	96.05	93.40	84.10
128	26.50	17.75	6.15	128	72.55	62.70	43.40	128	96.55	94.90	87.70	128	99.90	99.90	99.20

TABLE 4.3 (Cont.)

POWER OF THE LM TEST.
True model ARFIMA(0,d_0,1).
METHOD 1

$d_0 = 0.10$				$d_0 = 0.20$				$d_0 = 0.30$				$d_0 = 0.40$			
	β				β				β				β		
n	10%	5%	1%	n	10%	5%	1%	n	10%	5%	1%	n	10%	5%	1%
64	15.30	9.05	2.45	64	33.30	23.25	11.25	64	57.60	47.95	31.60	64	79.10	72.15	57.25
128	19.25	11.65	4.95	128	56.45	45.60	27.05	128	87.05	81.55	68.65	128	98.60	97.55	92.00

TABLE 4.3 (Cont.)

POWER OF THE LM TEST.
True model ARFIMA(0,d_0,1)
METHOD 2

$d_0 = 0.10$				$d_0 = 0.20$				$d_0 = 0.30$				$d_0 = 0.40$			
	β				β				β				β		
n	10%	5%	1%	n	10%	5%	1%	n	10%	5%	1%	n	10%	5%	1%
64	15.10	8.65	2.85	64	33.75	23.50	11.60	64	57.85	48.75	31.25	64	79.20	72.70	57.50
128	19.35	11.45	4.85	128	56.45	45.60	27.20	128	87.50	81.80	69.25	128	98.55	97.50	92.15

TABLE 4.3 (Cont.)

POWER OF THE LM TEST.
True model GARMA(0,d_0,1).
METHOD 1

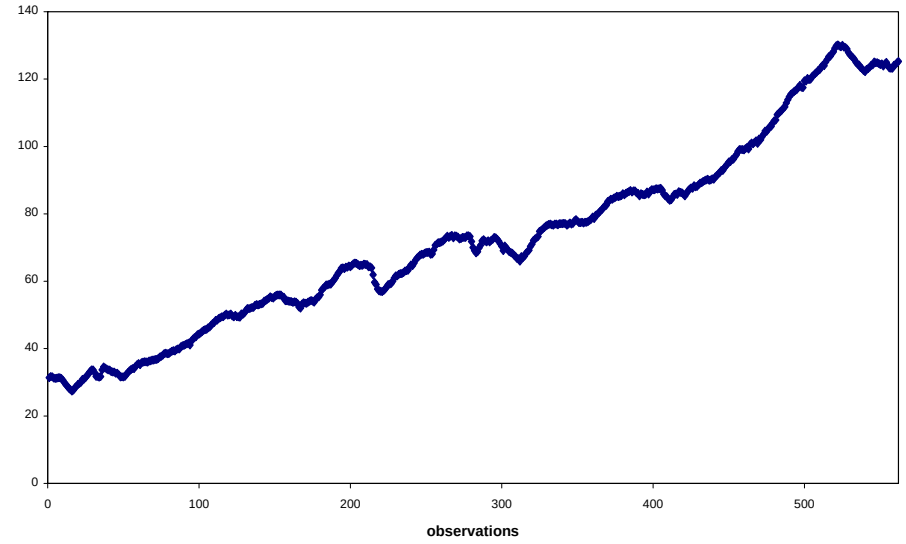
$d_0 = 0.10$				$d_0 = 0.20$				$d_0 = 0.30$				$d_0 = 0.40$			
	β				β				β				β		
n	10%	5%	1%	n	10%	5%	1%	n	10%	5%	1%	n	10%	5%	1%
64	21.00	13.20	3.55	64	51.75	39.80	21.10	64	80.65	74.35	57.30	64	96.35	93.40	87.90
128	30.65	21.25	8.40	128	75.35	67.30	48.00	128	97.00	95.00	88.85	128	99.90	99.85	99.15

TABLE 4.3 (Cont.)

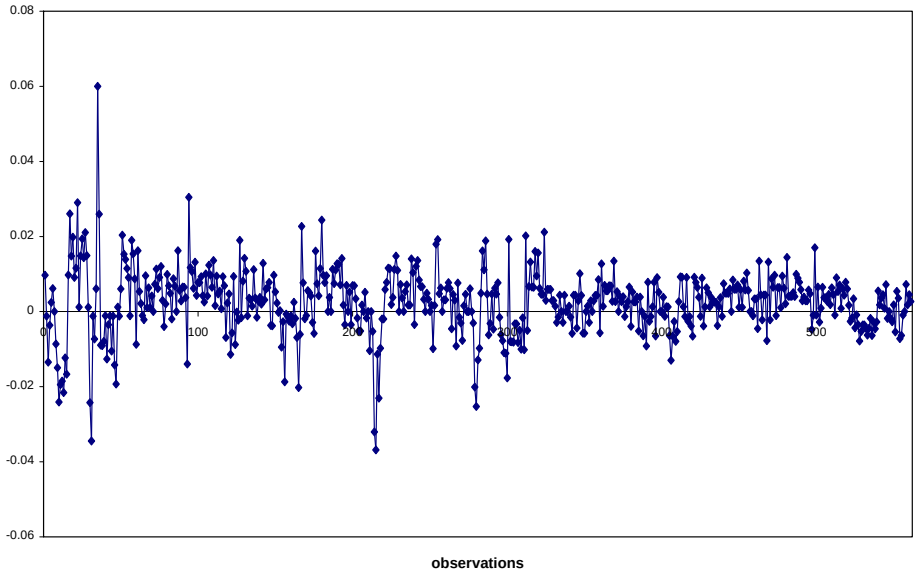
POWER OF THE LM TEST.
True model GARMA(0,d_0,1)
METHOD 2

$d_0 = 0.10$				$d_0 = 0.20$				$d_0 = 0.30$				$d_0 = 0.40$			
	β				β				β				β		
n	10%	5%	1%	n	10%	5%	1%	n	10%	5%	1%	n	10%	5%	1%
64	21.05	13.35	3.85	64	52.45	40.05	21.05	64	80.55	74.20	58.15	64	96.35	93.75	88.70
128	30.20	20.95	8.40	128	75.50	67.55	48.25	128	96.95	95.05	89.10	128	99.90	99.85	99.30

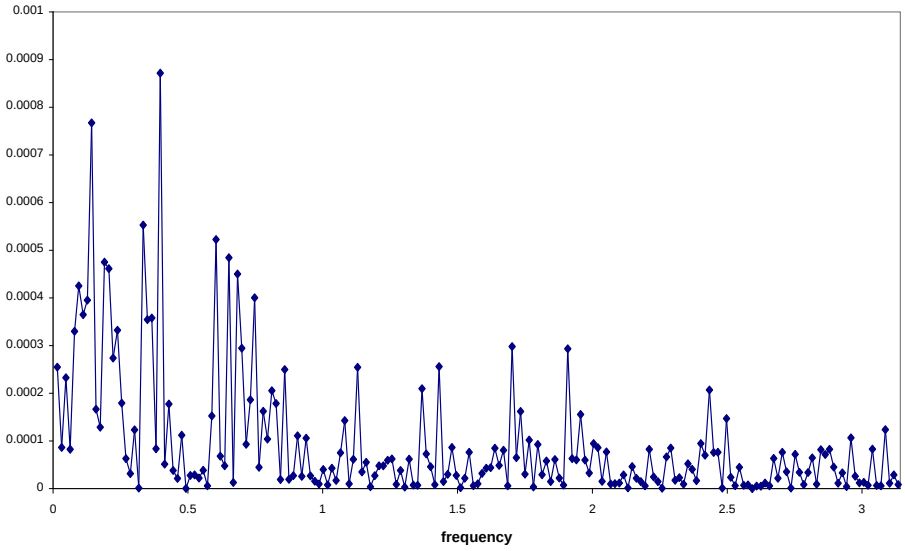
Graph 1 : IPI series



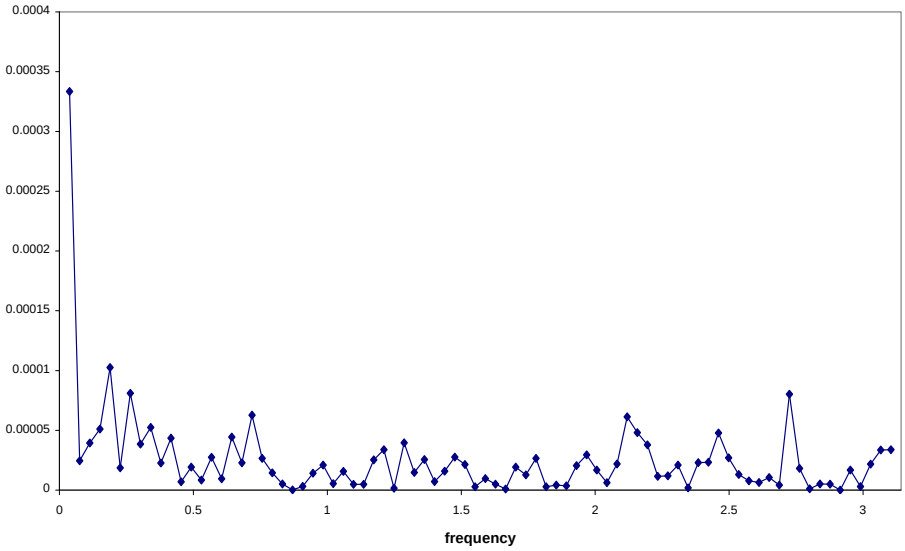
Graph 2 : Rate of growth of IPI series



Graph 3 : Periodogram of rate of growth of IPI series
January 1957-December 1989



Graph 4 : Periodogram of rate of growth of IPI series
January 1990-October 2003



References

[1] Adenstedt, I., 1974. On large-sample estimation for the mean of a stationary random sequence. *Annals of Mathematical Statistics* 2, 1095-1107.

[2] Andel, J., 1986. Long-memory time series models," *Kybernetika* 22, 105-123.

[3] Andrews, D.W.K., 2001. Testing when a parameter is on the boundary of the maintained hypothesis. *Econometrica* 69, 683-734.

[4] Arteche, J. and Robinson, P.M., 2000. Semiparametric inference in seasonal and cyclical long memory processes. *Journal of Time Series Analysis* 21, 1-25.

[5] Beran, J., Bhansali, R.J. and Ocker, D., 1998. On unified model selection for stationary and nonstationary short- and long-memory autoregressive processes. *Biometrika* 85, 921-934.

[6] Berk, K.N., 1974. Consistent autoregressive spectral estimates. *Annals of Statistics* 2, 489-502.

[7] Bickel, P. and Bühlmann, P., 1995. Mixing property and functional central limit theorems for a sieve bootstrap in time series. *Technical report* 440, Dept. Stat., U.C. Berkeley.

[8] Billingsley, P., 1968. Convergence of probability measures. John Wiley: New York.

[9] Bloomfield, P., 1973. An Exponential Model for the Spectrum of a Scalar Time Series. *Biometrika* 60, 217-226.

[10] Brillinger, D.R., 1981. Time Series: Data analysis and theory. McGraw-Hill.

[11] Brockwell, P.J. and Davis, R.A., 1991. Time Series: Theory and Methods. Springer-Verlag: Berlin.

[12] Davies, R.B. and Harte, D.S., 1987. Tests for Hurst effect. *Biometrika* 74, 95-101.

[13] Delgado, M., Hidalgo, J. and Velasco, C., 2003. Distribution free goodness-of-fit tests for linear processes. *Preprint.*

[14] Doukhan, P., 1994. Mixing: Properties and Examples. *Lecture Notes in Statistics* 85. Springer-Verlag: Berlin.

[15] Doukhan, P., 2003. Models, inequalities, and limit theorems for stationary sequences. In Theory and Applications of Long-Range dependence. Ed. P. Doukhan, G. Oppenheim and M.S. Taqqu. Birkhäuser, Berlin.

[16] Efron, B., 1979. Bootstrap Methods: Another Look at the Jackknife. *Annals of Statistics* 7, 1-26.

[17] Fox, R. and Taqqu, M.S., 1986. Large-sample properties of parameter estimates for strongly dependent stationary Gaussian time series. *Annals of Statistics* 14, 517-532.

[18] Franke, J. and Kreiss, J.-P., 1992. Bootstrapping stationary autoregressive moving average models. *Journal of Time Series Analysis* 13, 297-319.

[19] Giné, E. and Zinn, S., 1990. Bootstrapping general empirical measures. *Annals of Probability* 18, 851-869.

[20] Giraitis, L., Hidalgo, J. and Robinson, P.M., 2001. Gaussian estimation of parametric spectral densities with unknown pole. *Annals of Statistics* 29, 987-1023.

[21] Giraitis, L. and Surgailis, D., 1990. A Central Limit Theorem for Quadratic Forms in Strongly Dependent Linear Variables and its Application to Asymptotic Normality of Whittle's Estimate. *Probability Theory and Related Fields* 86, 87-104.

[22] Gorodetskii, V.V., 1977. On the strong mixing property for linear sequences. Theory of Probability and its Applications 22, 411-413.

[23] Granger, C.W.J. and Joyeux,R., 1980. An Introduction to Long Memory Time Series and Fractional Differencing. *Journal of Time Series Analysis* 1, 424-438.

[24] Gray, Zhang, N-F. and Woodward,W.A., 1989. On Generalized Fractional Processes. *Journal of Time Series Analysis* 10, 233-257.

[25] Hall, P., 1979. On the rate of convergence of normal extremes. *Journal of Applied Probability* 16, 433-439.

[26] Hall, P., 1990. On the relative performance of bootstrap and Edgeworth approximations of a distribution function. *Journal of Multivariate Analysis* 35, 108-129.

[27] Hall, P., 1991. On convergence rates of suprema. *Probability Theory and Related Fields,* 89, 447-455.

[28] Hannan, E.J., 1970. Multiple Time Series. *Wiley and Sons*: New York.

[29] Hannan, E.J., 1973. The asymptotic theory of linear time series models. *Journal of Applied Probability* 10, 130-145.

[30] Hidalgo, J., 2003. An alternative bootstrap to moving blocks for time series regression models. *Journal of Econometrics* 117, 363-395.

[31] Hidalgo, J. and Kreiss, J.P., 2003. Bootstrap specification tests for linear covariance stationary processes. *Forthcoming in Journal of Econometrics.*

[32] Hosking, J., 1981. Fractional differencing. *Biometrika* 68, 165-176.

[33] Hosoya, Y., 1997. A limit theory with long-range dependence and statistical inference on related models. *Annals of Statistics* 25, 105-137.

[34] Ibragimov, I.A., 1965. On the spectrum of stationary Gaussian sequences satisfying the strong mixing condition. I. Necessary conditions. *Theory of Probability and its Applications* 10, 85-106.

[35] Ibragimov, I.A., 1970. On the spectrum of stationary Gaussian sequences satisfying the strong mixing condition. II. Sufficient conditions. *Theory of Probability and its Applications* 15, 23-36.

[36] Lobato, I.N. and Robinson, P.M., 1998. A nonparametric test for I(0). *Review of Economic Studies* 65, 475-495.

[37] McLeish, D.L., 1975. Invariance principles for dependent variables. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* 32, 165-178.

[38] Noé, M., 1972. The calculation of distributions of two-sided Kolmogorov-Smirnov type statistics. *Annals of Mathematical Statistics* 43, 58-64.

[39] Nze, P.A., Bühlmann, P. and Doukhan, P., 2002. Weak dependence beyond mixing and asymptotics for parametric regression. *Annals of Statistics* 30, 397-430.

[40] Pham, T.D. and Tran, L.T., 1985. Some mixing properties of time series models. *Stochastic Processes and its Applications* 19, 297-303.

[41] Prichard, D. and Theiler, J., 1994. Generating surrogate data for time series with several simultaneously measured variables. Working paper #94-04-023. Santa Fe Institute.

[42] Robinson, P.M., 1994. Efficient tests of nonstationary hypothesis. *Journal of the American Statistical Association* 89, 1420-1437.

[43] Robinson, P.M., 1995a. Log-periodogram regression of time series with long range dependence. *Annals of Statistics* 23, 1048-1072.

[44] Robinson, P.M., 1995b. Gaussian semiparametric estimation of long range dependence. *Annals of Statistics* 23, 1630-1661.

[45] Robinson, P.M. and Hidalgo, F.J., 1997. Time series regression with long range dependence. *Annals of Statistics* 25, 77-104.

[46] Rosenblatt, M., 1956. A central limit theorem and a strong mixing condition. *Proceedings National Academy of Science U.S.A.* 42, 43-47.

[47] Steck, G.P., 1971. "Rectangle probabilities for uniform order statistics and the probability that the empirical distribution lies between two distributions functions. *Annals of Mathematical Statistics* 42, 1-11.

[48] Stout, W.F., 1974. Almost sure convergence. *Academic Press*, New York.

[49] Theiler, J., Galdrikian, B., Longtin, A., Eubank, S. and Farmer, J.D., 1992. Using surrogate data to detect nonlinearity in time series. A proceedings volume in the Santa Fe Institute Studies in the Sciences of Complexity. Nonlinear Modeling and Forecasting. Ed: Casdagli, M. and Eubank, S., 163-188.

[50] Velasco, C. and Robinson, P.M., 2000. Whittle pseudo-maximum likelihood estimates for non-stationary time series. *Journal of the American Statistical Association* 95, 1229-1243.

[51] Volkonskii, V.A. and Rozanov, Y.A., 1959. Some limit theorems for random functions, Part I. *Theory of Probability and its Applications* 4, 178-197.

Economics Department, London School of Economics, Houghton Street, London WC2A 2AE, U.K.