A NONPARAMETRIC TEST FOR WEAK DEPENDENCE AND ITS BOOTSTRAP ANALOGUE*

Javier Hidalgo
London School of Economics

February 13, 2002

Abstract

This paper examines a test for the hypothesis of weak dependence, whose limit distribution, $G(z)$, is a Gumbel distribution which appears as one of the three possible limit distributions in the theory of extreme-values. However, since $G(z)$ may be a poor approximation to the finite sample distribution, the rate of convergence being logarithmic, see Hall (1979), inferences based on $G(z)$ may not be very reliable for moderate sample sizes. For that reason, we describe an approach to bootstrapping the test based on a naive, e.g. Efron (1979), resampling of the data. We show that the bootstrap principle is consistent under very mild regularity conditions. This can be quite surprising since Efron's resampling scheme is generally inconsistent under dependence.

Key words and phrases: Strong and weak dependence. Spectral density function. Spectral density estimation. Extreme-values. Bootstrap tests.

JEL Classification: C22

1. INTRODUCTION

In recent years we have observed an increasing interest on qualitative hypothesis testing, such as testing for monotonicity, convexity or whether a (nonparametric) curve is positive. Examples of the former, in a regression model context, are given in Bowman, Jones and Gijbels (1998), Hall and Heckman (2000), Ghosal, Sen and Van der Vaart (2000), and Dümbgen and Spokoiny (2001), whereas in the context of density functions, Woodroofe and Sun (1999) test uniformity against monotonicity. Testing for these type of hypothesis appears to be a necessary prerequisite to any previous attempt to estimate curves satisfying the restrictions, see for example Mammen (1992) and references therein. In this sense, it is not surprising that the tests can be regarded as goodness-of-fit tests.

*This article is based on research funded by the Economic and Social Research Council (ESRC) reference number: R000238212.

Although the literature has dealt with cross-section data, there is not apparent reason to believe that it may not appear in the context of time series. One leading example is when trying to investigate cycles in macroeconomic or geophysical data. It is obvious that previous to attempt estimating the length of the cycle one needs to test for its presence. The existence of a cycle is shown by a sharp peak of the estimated spectral density function. One model capable to induce peaks is

$$x_t = \mu + \rho_1 \cos\left(\lambda^0 t\right) + \rho_2 \sin\left(\lambda^0 t\right) + \varepsilon_t, \tag{1.1}$$

where ρ_1 and ρ_2 are zero mean uncorrelated random variables with the same variance and $\{\varepsilon_t\}$ is a stationary sequence of random variables independent of ρ_1 and ρ_2. Model (1.1) has enjoyed extensive use and different techniques have been proposed for the estimation of the frequency, amplitude and phase (see Whittle (1952), Grenander and Rosenblatt (1957), Hannan (1961, 1973) and Chen (1988)). Extensions to a model with more than one periodic component have been examined by Quinn (1989) and Kavalieris and Hannan (1994), whose interest was also in testing the number of sinusoidal/cosinusoidal components.

A second statistical model capable of exhibiting peaks in its spectral density function is the autoregressive $AR(2)$ process

$$\left(1 - a_1 L - a_2 L^2\right) x_t = \varepsilon_t \tag{1.2}$$

when the roots of the polynomial $\left(1 - a_1 L - a_2 L^2\right)$ are complex, with λ^0 identified as the $arc\cos\left(\frac{a_1}{2\sqrt{-a_2}}\right)$. Models (1.1) and (1.2) represent two extreme situations explaining cyclical behaviour of the data and the peakness of the spectral density function. Model (1.2) possesses a continuous spectral density function whereas model (1.1) has a spectral distribution function with a jump at the frequency λ^0. The cyclical component of the data remains constant or invariant with time in model (1.1), whereas the cyclical pattern of model (1.2) fades out with time fairly quickly.

Between these two extreme situations, there exists a class of intermediate models in which the spectral density function of x_t exhibits a pole at the frequency λ^0. For that purpose, define the spectral density function of x_t as the function $f(\lambda)$ which satisfies the relationship

$$\gamma(j) = Cov\left(x_t, x_{t+j}\right) = \int_{-\pi}^{\pi} f(\lambda)\cos(j\lambda)\, d\lambda, \quad j = 0, 1, 2, ... \tag{1.3}$$

We say that $f(\lambda)$ has a pole at λ^0 if

$$f(\lambda) \sim C\left|\lambda - \lambda^0\right|^{-\alpha} \quad \text{as} \quad \lambda \to \lambda^0, \tag{1.4}$$

where $C \in (0, \infty)$, $\alpha \in (0, 1)$ is the memory parameter and "$\sim$" means that the ratio of the left- and right-hand sides tends to 1. One of the main objectives of this paper is the estimation of λ^0.

One model capable of generating such a cyclical behaviour in the data has been proposed by Andel (1986) and Gray et al. (1989, 1994), and defined as

$$\left(1 - 2\left(\cos\lambda^0\right) L + L^2\right)^d x_t = \varepsilon_t, \tag{1.5}$$

where L is the backshift operator, $d = \alpha/2$ for $\lambda^0 \in (0, \pi)$ whereas for $\lambda^0 = 0$ or π, $d = \alpha/4$. The model (1.5) was coined by Gray et al. (1989) as the Gegenbauer model, who extended it to the $GARMA$ model where the innovation ε_t follows an autoregressive moving average ($ARMA$) process, and later by Giraitis and Leipus (1995) allowing for more than one pole or cyclical component. The $GARMA$ process is characterized by having a spectral density function defined as

$$f(\lambda) = \frac{\sigma^2}{2\pi} \left|1 - 2\left(\cos\lambda^0\right)e^{i\lambda} + e^{i2\lambda}\right|^{-2d} \left|\frac{a\left(e^{i\lambda};\theta\right)}{b\left(e^{i\lambda};\theta\right)}\right|^2 \qquad -\pi < \lambda \leq \pi \tag{1.6}$$

where $\sigma^2 > 0$, and $a(\cdot)$ and $b(\cdot)$ are polynomials of finite degree, all of whose zeroes lie outside the unit circle. When $\lambda^0 = 0$, we have the more familiar $FARIMA$ model, apparently originated by Adenstedt (1974), and studied by Granger and Joyeux (1980) and Hosking (1981). $GARMA$ models are characterized by a stronger and more persistent cyclical behaviour than $ARMA$ models, e.g. (1.2), but unlike model (1.1), their amplitude does not remain constant over time.

So that the problem of testing for a cycle can be regarded as a qualitative hypothesis testing. That is, does the spectral density function possess a peak or can be regarded as monotonic? When the peak is assumed to be smooth, the problem of testing for a cycle is similar to testing for a mode in the probability density function. However, a cycle explained by a smooth peak implies that this fades out rather quickly with time. On the other hand, it is observed that, if exist, this cycles can be long dependent in time. This could be explained by the so called strong dependence. We can roughly define strong dependence as a process for which there exists at least a frequency $\lambda^0 \in [0, \pi]$ whose spectral density function, $f(\lambda)$, behaves as $C\left|\lambda - \lambda^0\right|^{-\alpha(\lambda^0)}$, with $0 < \alpha\left(\lambda^0\right) < 1$, in a neighbourhood of λ^0. The frequency λ^0 has the interpretation of being the cycle of the data, which is more persistent than those characterize by $ARMA$ models but not deterministic as regression based models.

In this context, the value $\alpha\left(\lambda^0\right)$ can be regarded as determining the (local) shape of the spectral density function around λ^0, whose shape can discriminate among different time series. In addition, it gives indication and summarizes the dependence structure at the long run, albeit it induces quite different statistical/probabilistic properties of estimates (possibly implicit ones) depending on whether $\alpha\left(\lambda^0\right) > 0$ or $= 0$. For example, the estimation of the sample moments of the data, see Taqqu (1975), or $M-$estimation, see Beran (1985) or Koul and Surgailis (1997). Another examples are on the estimation of the probability distribution function via the empirical distribution, see Dehling and Taqqu (1987) or Ho and Hsing (1996), or in the context of estimation of the spectral density function. As Giraitis and Hidalgo (2001) have shown, the asymptotic properties of, possibly, nonlinear transformation of x_t, depends very much on whether or not x_t is strong dependent, even if the transformation of x_t induces a spectral density function which is Lipchitz continuous in $[0, \pi]$. Finally, it is worth mentioning that when $\alpha\left(\lambda^0\right) > 0$, estimators of λ^0 have been provided in parametric and nonparametric frameworks, see Giraitis et al. (2001) and Hidalgo (2001) respectively.

The remainder of the paper is as follows. In the next section, we describe the hypothesis testing and introduce the test for $\alpha\left(\lambda^0\right) = 0$ for all $\lambda^0 \in [0, \pi]$. Section 3 describes a semiparametric estimator of $\alpha\left(\lambda^0\right)$, which is needed to implement the test, as well as we give the statistical properties of the tests. Because the limit distribution of the test is that obtained in the theory of extreme-values, whose finite sample distribution is poorly approximated by its asymptotic counterpart, Section 4 describes and proposes a bootstrap approach based on Efron's (1979) naive scheme, showing its consistency. Finally, the proofs of the results in Sections 3 and 4 are given in Appendix A which makes use of a series of Lemmas in Appendix B.

2. FORMULATION OF THE TEST

The class of models investigated are processes x_t which, after maybe some transformation, are covariance stationary and observed at times $t = 1, ..., n$. We assume that x_t has an absolute continuous spectral distribution, so that it has an spectral density function, $f(\lambda)$, defined by the relation

$$\gamma(j) = E\left((x_0 - Ex_0)(x_j - Ex_j)\right) = \int_{-\pi}^{\pi} f(\lambda)\, e^{ij\lambda} d\lambda, \quad j = 0, \pm 1, \pm 2, ... \tag{2.1}$$

As we mentioned in the introduction, the main objective of the paper is to introduce a test, examining its statistical properties, for the hypothesis of weak dependence. More precisely, we are concerned with the following hypothesis testing

$$\begin{aligned} H_0 &: \quad \exists\, K > 0 \text{ such that } K^{-1} < f(\lambda) < K \quad \forall\, \lambda \in [0, \pi] \\ H_1 &: \quad \exists\, \lambda^0 \in [0, \pi] \text{ such that } f\left(\lambda^0\right) = 0 \text{ or } f^{-1}\left(\lambda^0\right) = 0. \end{aligned} \tag{2.2}$$

An alternative formulation of (2.2) could have been given in terms of the autocovariance function $\gamma(j)$. Specifically whether or not $\sum_{j=0}^{\infty} |\gamma(j)| < \infty$.

Some existing work, see Lobato and Robinson (1997), has focused on testing if $f(0)$ is continuous and positive, that is testing for $I(0)$. However many time series appears to exhibit a periodic behaviour, which it is manifested by a peak of the spectral density estimate. Although knowledge of the location of the (possible) peak can be realistic in many situations, for instance with seasonal data, with no seasonal observations that knowledge is not so clear. An example of the latter is, as we mentioned in the introduction, when the practitioner is interested in estimating the cycle in macroeconomic or geophysical data. So that it seems appropriate to extend Lobato and Robinson (1997) to all frequencies $\lambda \in [0, \pi]$, that is the hypothesis testing given in (2.2). It is worth noting that it can be generalized to say, $\lambda \in [a, b] \subset [0, \pi]$. However, for simplicity of exposition, we will let $a = 0$ and $b = \pi$.

To fix ideas and give the intuition of our strategy to test for H_0, let us suppose first that $f(\lambda)$ is completely known up to a finite set of parameters θ. That is, (2.1) becomes

$$\gamma(j; \theta) = \int_{-\pi}^{\pi} f(\lambda; \theta)\, e^{ij\lambda} d\lambda, \quad j = 0, \pm 1, \pm 2, ... \tag{2.3}$$

Next, assume that for some fixed frequency, say λ^0, we suspect that $f^{-1}\left(\lambda^0; \theta\right)$, or $f\left(\lambda^0; \theta\right)$, is zero. Then, based on estimates of θ, say the Whittle estimator $\widehat{\theta}$, the test for H_0 can be implemented by looking at whether or not $f^{-1}\left(\lambda^0; \widehat{\theta}\right)$, or $f\left(\lambda^0; \widehat{\theta}\right)$, is significantly different than zero. For example, if $f(\lambda)$ is parameterized by a $GARMA$ model, introduced in Gray et al. (1989, 1996), and defined as

$$f(\lambda; \theta) = \left(1 - 2\cos\left(\lambda^0\right) e^{i\lambda} + e^{2i\lambda}\right)^{-\alpha/2} \frac{\left|\Theta\left(e^{i\lambda}; \theta^1\right)\right|^2}{\left|\Phi\left(e^{i\lambda}; \theta^2\right)\right|^2}$$

where $\theta = \left(\alpha, \theta_1', \theta_2'\right)'$ with $-1 < \alpha < 1$ if $\lambda^0 \neq 0, \pi$ and $-1/2 < \alpha < 1/2$ if $\lambda^0 = 0, \pi$, and where $\Theta\left(e^{i\lambda}; \theta^1\right)$ and $\Phi\left(e^{i\lambda}; \theta^2\right)$ are the MA and AR polynomials with no common roots and outside the unit circle. Then given $\widehat{\theta}$, H_0 can be tested by looking at whether or not $\widehat{\alpha}$ is significantly different than zero. Note that when $\lambda^0 = 0$, $f(\lambda; \theta)$ becomes the popular $FARIMA$ model apparently introduced in Adenstedt (1974) and examined by Granger and Joyeux (1980) and Hosking (1981).

Next, let λ^0 be unknown and the Whittle estimator of θ is performed when λ^0 is taken to be the Fourier frequencies $\lambda_j = (2\pi j)/n$, for integer j. Then, for each λ_j we would have an estimate of α, say $\widehat{\alpha}(j)$. From here, it is intuitive that the null hypothesis H_0 in (2.2) can be based on whether or not for all $j = 0, ..., [n/2]$, $\widehat{\alpha}(j)$ are not simultaneously different than zero. That is, we would reject H_0 if $\sup_{j=0,...,[n/2]} |\widehat{\alpha}(j)|$ is greater than some critical value. This is the idea of the test given in Giraitis and Hidalgo (2001), which can be thought as a Wald type test. Since, the estimation of the parameters of $f(\lambda; \theta)$ involves nonlinear optimization procedures, it seems appealing to use an LM test since only the model under the null needs to be estimated. The test can be implemented by looking at the discrepancy between the periodogram and the fitted model under H_0 via a Kolmogorov-Smirnov statistic. That is,

$$\sup_{\mu\in[0,1]} \left| \int_0^{\pi\mu} \left(\frac{\left|\Phi\left(e^{i\lambda}; \widehat{\theta}^1\right)\right|^2}{\left|\Theta\left(e^{i\lambda}; \widehat{\theta}^2\right)\right|^2} I(\lambda) - 1 \right) d\lambda \right|,$$

where $\widehat{\theta}^1$ and $\widehat{\theta}^2$ are the Whittle estimators of θ^1 and θ^2 respectively.

However, the previous approach is very sensitive to a correct specification of the model, which can be difficult to know a priory. More specifically, if the correct specification is, say, a Bloomfield's (1974) exponential instead of an $ARMA$ one, the tests based on a parameterization of f can be misleading and invalid. This is in particular the case if an LM test is implemented. In this case, the null hypothesis would be rejected with probability approaching 1 as $n \to \infty$ even if α were equal to zero, e.g. the data is weakly dependent. Hence, it seems desirable to have a test for (2.2) which does not depend on any particular parameterization of f.

To that end, we first describe how the hypothesis testing in (2.2) can be written in a more standard formulation, e.g. in terms of whether or not a (set of) parameter(s) is equal to some particular value. Suppose first that f is continuous and positive, which implies that for all $\lambda \in [0, \pi]$

$$K^{-1} < f(\lambda) = h(\lambda) < K, \tag{2.4}$$

where henceforth K denotes a positive finite constant.

On the other hand, under H_1, we have that

$$f(\lambda) = h\left(\lambda; \lambda^0\right) \left|\lambda - \lambda^0\right|^{-\alpha} \tag{2.5}$$

for some $\lambda^0 \in [0, \pi]$ with $\alpha \neq 0$. In particular, if $f\left(\lambda^0\right) = 0$ it means that $\alpha < 0$, whereas $f^{-1}\left(\lambda^0\right) = 0$ implies that $\alpha > 0$.

In (2.5) write $\alpha = \alpha\left(\lambda^0\right)$, that is, the rate at which $f(\lambda)$ increases to infinity or decreases to zero at the frequency λ^0. Then, (2.2) can equivalently be written as

$$\begin{aligned} H_0 &: \quad \alpha\left(\lambda^0\right) = 0 \quad \forall \lambda^0 \in [0, \pi] \\ H_1 &: \quad \exists \lambda^0 \in [0, \pi] \quad \text{such that } \alpha\left(\lambda^0\right) \in (-1, 1) / \{0\}. \end{aligned} \tag{2.6}$$

(Recall that since $Ex_t^2 < \infty$ and the process is invertible, $\alpha \in (-1, 1)$.) On the other hand, if we were interested to test the hypothesis $f(\lambda) > K^{-1}$ against $f(\lambda^0) = 0$ for some $\lambda^0 \in [0, \pi]$, (2.6) would be

$$\begin{aligned} H_0 &: \quad \alpha(\lambda^0) = 0 \quad \forall \lambda^0 \in [0, \pi] \\ H_1 &: \quad \exists \lambda^0 \in [0, \pi] \quad \text{such that } \alpha(\lambda^0) < 0. \end{aligned} \tag{2.7}$$

Similarly, a test for $f(\lambda) < K$ against $f^{-1}(\lambda^0) = 0$ for some $\lambda^0 \in [0, \pi]$, the hypothesis testing would be

$$\begin{aligned} H_0 &: \quad \alpha(\lambda^0) = 0 \quad \forall \lambda^0 \in [0, \pi] \\ H_1 &: \quad \exists \lambda^0 \in [0, \pi] \quad \text{such that } \alpha(\lambda^0) > 0. \end{aligned} \tag{2.8}$$

Once we have written the hypothesis testing in terms of parameters, we now describe the main ideas of the test. Assuming that $f(\lambda)$ follows model (2.4), suppose that for a given λ^0, we estimate $\alpha(\lambda^0)$, denoted $\widehat{\alpha}(\lambda^0)$, by some semiparametric method. For instance the log-periodogram and local Whittle estimators given in Robinson ($1995a, b$), or the estimator given in Hidalgo and Yajima (2000). Then, a test for (2.6) can be based on whether $\widehat{\alpha}(\lambda^0)$ is significantly different than zero for all $\lambda^0 \in [0, \pi]$. Similarly, a test for (2.7) and/or (2.8) can be based on whether $\widehat{\alpha}(\lambda^0)$ is significantly greater than zero or less than zero, respectively.

From the above comments, we can now easily formulate the tests. Suppose that the estimate of α, denoted $\widehat{\alpha}(s)$, is computed at frequencies λ_s, $s = 0, 1, ..., [n/2] = \tilde{n}$. Then, under H_0, we should expect that $\widehat{\alpha}(s) \approx 0$ for all $s = 0, 1, ..., \tilde{n}$, whereas under H_1, $\widehat{\alpha}(s) \neq 0$ for some s. Hence, a test for (2.2) can be based on

$$\mathcal{T}_1 = \sup_{s=0,1,\ldots,\tilde{n}} |\widehat{\alpha}(s)| . \tag{2.9}$$

On the other hand, tests for the hypothesis testing in (2.7) or (2.8) can be based on

$$\mathcal{T}_2 = \sup_{s=0,1,\ldots,\tilde{n}} (-\widehat{\alpha}(s)) \tag{2.10}$$

or

$$\mathcal{T}_3 = \sup_{s=0,1,\ldots,\tilde{n}} \widehat{\alpha}(s) , \tag{2.11}$$

respectively. Then, we reject the null hypothesis if $\mathcal{T}_j$, $j = 1, 2, 3$, is greater than some (positive) critical value.

3. THE LIMIT BEHAVIOUR OF THE TEST

To implement the test we need first to provide an estimator of $\alpha(\lambda^0)$, for $\lambda^0 \in [0, \pi]$, which does not depend on any particular parameterization of f, e.g. $h(\lambda)$ given in (2.5). For a specific value λ^0, several semiparametric estimators of the "memory" parameter $\alpha(\lambda^0)$ have been proposed and examined. However, in this paper we employ a modification of an estimator proposed in Parzen (1986). To motivate the estimator, let us assume that the spectral density $f(\lambda)$ satisfies

$$f(\lambda) = C\lambda^{-\alpha(0)} \quad \text{for} \quad \lambda \in (0, \bar{\lambda}) , \tag{3.1}$$

and $\lambda^0 = 0$, without loss of generality. If f follows the model given in (3.1), then after standard calculations we have that

$$\alpha(0) = \bar{\lambda}^{-1} \int_0^{\bar{\lambda}} \log f(\lambda)\, d\lambda - \log f(\bar{\lambda}). \tag{3.2}$$

Suppose now that $\bar{\lambda} = \lambda_k$ with $k = o(n)$. Then, we can expect that the Riemann's discrete approximation of the right side of (3.2), that is

$$\frac{1}{k}\sum_{p=1}^{k} \log f_p - \log f_k$$

to be closed to $\alpha(0)$, where henceforth for a generic function $\vartheta(\cdot)$ we abbreviate $\vartheta(\lambda_j)$ by ϑ_j.

However, $f(\lambda)$ is not known, so to make feasible the last displayed expression, we need to estimate $f(\lambda)$. For that purpose, let us introduce the periodogram of x_t

$$I_\ell = \left| (2\pi n)^{-1/2} \sum_{t=1}^{n} x_t e^{it\lambda_\ell} \right|^2, \quad \ell = 1, ..., \tilde{n} \tag{3.3}$$

with $I_0 = 0$. Note that our definition of I_ℓ entails sample-mean correction. To estimate f_ℓ we employ the weighted periodogram estimator

$$\widehat{f}_\ell = \frac{\pi}{\tilde{n}} \sum_{j=1-\tilde{n}}^{\tilde{n}-1} W_M(\lambda_\ell - \lambda_j)\, I_j \tag{3.4}$$

where $W_M(\lambda) = M\sum_{j\in\mathcal{J}} W(M(\lambda + 2\pi j))$, with $\mathcal{J}$ the set of natural numbers, $M = M(n)$ a number which increases slowly with n, that is $M^{-1} + Mn^{-1} \to 0$, and $W(\lambda)$ a symmetric differentiable nonnegative function which integrates 1 and satisfies

$$\int_{-\infty}^{\infty} \lambda^2 |W(\lambda)|\, d\lambda < \infty.$$

When $W(\lambda)$ lies in a compact set, say $(-\pi, \pi)$, the estimator given in (3.4) becomes

$$\widehat{f}_\ell = \frac{1}{2m+1} \sum_{j=-m}^{m} W_M(\lambda_j)\, I_{j+\ell},$$

where $m = [n/4M]$. That is, $\widehat{f}_\ell$ weights the periodograms of the $2m+1$ closest frequencies to λ_ℓ.

Denote

$$\widehat{f}_p(s) = \frac{1}{2m+1} \sum_{j=-m}^{m} W_j I_{j+p+s}$$

where λ_s is the closest Fourier frequency to λ^0 and $W_j = W(\lambda_j)$. Then, we define our estimator of $\alpha(\lambda_s)$ as

$$\widehat{\alpha}(s) = \frac{1}{2}\phi_w^{-1} \left(\frac{1}{k} \left\{ \sum_{p=1}^{k} + \sum_{p=-k}^{-1} \right\} w_{|p|} \log \widehat{f}_p(s) - \overline{w} \left(\log \widehat{f}_k(s) + \log \widehat{f}_{-k}(s) \right) \right), \tag{3.5}$$

where $w_p = w(p/k)$, $\overline{w} = \sum_{p=1}^{k} w_p$, $0 < \phi_w = -\int_0^1 w(u)(\log u)\,du < \infty$ and for notational simplicity we write $k = m[\log m]$.

It is worth observing two points. Firstly, for $\lambda_s = 0, \pi$, $\widehat{\alpha}(s)$ given in (3.5) collapses to

$$\phi_w^{-1}\frac{1}{k}\sum_{p=1}^{k} w_p\left(\log \widehat{f}_p(s) - \log \widehat{f}_k(s)\right)$$

due to symmetry of f around 0 and π. The second point gives the motivation to use (3.5) instead of

$$\widetilde{\alpha}(s) = \phi_w^{-1}\frac{1}{k}\sum_{p=1}^{k} w_p\left(\log \widehat{f}_p(s) - \log \widehat{f}_k(s)\right).$$

Assume that $\lambda^0 = \lambda_s$ is known for simplicity and suppose that $\alpha(\lambda_s) = \alpha(s)$ is estimated by $\widetilde{\alpha}(s)$. As in other semiparametric estimators, for example Robinson (1995a), one source of the bias of $(2m)^{1/2}(\widetilde{\alpha}(s) - \alpha(s))$ comes from the replacement of $f(\lambda)$ by $h(\lambda_s)\left|\lambda - \lambda^0\right|^{-\alpha(s)}$, see condition C1 below, which in our case is proportional to

$$(2m)^{1/2}\left(\sum_{p=1}^{k} w_p\left(\lambda_p - \lambda_k + O\left(\lambda_p^2 + \lambda_k^2\right)\right)\right) = O\left(n^{-1}m^{3/2}\right).$$

The main reason for this behaviour is that when $\lambda^0 = 0$ or π, by symmetry we have $h'(0) = h'(\pi) = 0$, whereas for $\lambda \neq 0$ or π, $h'(\lambda)$ may not be zero and the approximation of $f(\lambda)$ by $h(\lambda_s)|\lambda - \lambda_s|^{-\alpha}$, e.g. $h^{-1}(\lambda_s)|\lambda - \lambda_s|^{\alpha} f(\lambda)$, is $1 + h^{-1}(\lambda_s)h'(\lambda_s)(\lambda - \lambda_s) + O\left(|\lambda - \lambda_s|^2\right)$ by a Taylor expansion of the function $h(\lambda)$ around λ_s. However, when the estimator $\widehat{\alpha}(s)$ in (3.5) is employed, the contribution of the above approximation (Taylor expansion) to the bias of $(2m)^{1/2}(\widehat{\alpha}(s) - \alpha(s))$ is proportional to

$$\begin{aligned} & (2m)^{1/2}\sum_{p=1}^{k} w_p\left(\lambda_p - \lambda_k + O\left(\lambda_p^2 + \lambda_k^2\right)\right) \\ & + (2m)^{1/2}\sum_{p=1}^{k} w_p\left(\lambda_k - \lambda_p + O\left(\lambda_p^2 + \lambda_k^2\right)\right) \\ = \;& O\left(m^{5/2}n^{-2}\right). \end{aligned}$$

Remark 1. *Alternative estimators of α in case of $\lambda^0 = 0$ were examined by Geweke and Porter-Hudak (1983) and Robinson (1995a, b). In fact, the estimator $\widehat{\alpha}(\lambda_s)$ is very similar to the log-periodogram and local Whittle estimators as we now illustrate. Suppose for simplicity that $\lambda^0 = 0$. The log-periodogram estimator of α is defined as*

$$\begin{aligned} \widehat{\alpha}_{LOG} &= -\frac{\frac{1}{k}\sum_{\ell=1}^{k} \log I_\ell \left(\log \ell - \frac{1}{k}\sum_{j=1}^{k}\log j\right)}{\frac{1}{k}\sum_{\ell=1}^{k} \log \ell \left(\log \ell - \frac{1}{k}\sum_{j=1}^{k}\log j\right)} \\ &\sim -\frac{1}{k}\sum_{\ell=1}^{k} \log I_\ell \left(\log(\ell/k) - \frac{1}{k}\sum_{j=1}^{k}\log(j/k)\right), \end{aligned}$$

*since the denominator approaches 1 as $n \to \infty$. So, if instead of using the periodogram I_ℓ as an "estimator" of f_ℓ we used $\widehat{f}_\ell$, or that provided in (**??**), for n large enough, $\widehat{\alpha}_{LOG}$ would be as*

$$-\frac{1}{k}\sum_{\ell=1}^{k}\log\widehat{f}_\ell\left(\log(\ell/k)-\frac{1}{k}\sum_{j=1}^{k}\log(j/k)\right). \tag{3.6}$$

On the other hand, by standard algebra and that by symmetry $\widetilde{\alpha}_{+p}(\lambda_q)=\widetilde{\alpha}_{-p}(\lambda_q)$, we have that

$$\widehat{\alpha}(0)=\frac{1}{\phi_w\overline{v}}\left(\frac{1}{k}\sum_{\ell=1}^{k}\log\widehat{f}_\ell\left(\frac{1}{k}\sum_{p=\ell}^{k}w\left(\frac{\ell}{p}\right)v\left(\frac{p}{k}\right)\frac{k}{p}-\overline{w}_p v\left(\frac{\ell}{k}\right)\right)\right),$$

*where $\overline{v}=k^{-1}\sum_{p=1}^{k}v_p$. So, we have that (3.6) is of the form $\widehat{\alpha}(0)$, but with a weighting function $w_\ell=(\phi_w\overline{v})^{-1}\left(\sum_{p=\ell}^{k}w(\ell/p)\,v(p/k)\,p^{-1}-\overline{w}_p v(\ell/k)\right)$ replaced by $-\left(\log(\ell/k)-k^{-1}\sum_{j=1}^{k}\log(j/k)\right)$. From the above discussion, we observe that our estimator $\widehat{\alpha}(s)$ of the long memory parameter can be written as $2^{-1}k^{-1}\sum_{\ell=1}^{k}\psi(\ell/k)\log\left(\widehat{f}_{\ell+s}\widehat{f}_{s-\ell}\right)$ such that $\sum_{\ell=1}^{k}\psi(\ell/k)=0$ and $\sum_{\ell=1}^{k}\psi(\ell/k)\log(\ell/k)=-1$. However, we prefer to write the estimator as in (**??**), since as will be discussed below, to achieve asymptotic normality of the estimator of the pole requires some extra regularity conditions on $w(u)$ and $v(u)$ which can be rather complicated in terms of $\psi(u)$.*

With respect to Robinson's (1995b) local Whittle estimator, similar arguments apply. The only difference is to note that $\lambda_\ell^\alpha I_\ell-1$ is the first term in the Taylor expansion of $\log(\lambda_\ell^\alpha I_\ell)$ around 1. So, replacing $\lambda_\ell^\alpha I_\ell-1$ by $\log(\lambda_\ell^\alpha I_\ell)$ in the first order conditions of Robinson's (1995b) estimator, and following the same arguments as above, we readily observe a similar connection between $\widehat{\alpha}(0)$ and the estimator examined in Robinson (1995b) and the connection described between the log-periodogram and $\widehat{\alpha}(0)$ estimators.

Let us introduce the following regularity conditions:

C1

$$f(\lambda)=\begin{cases} h(\lambda)\left|\lambda-\lambda^0\right|^{-\alpha(\lambda^0)}, & \text{if} \quad \lambda\in[0,\pi]/\left\{\lambda^0\right\}, \\ h(\lambda)\left|\lambda+\lambda^0\right|^{-\alpha(\lambda^0)}, & \text{if} \quad \lambda\in[-\pi,0]/\left\{-\lambda^0\right\}, \end{cases}$$

where $\alpha\left(\lambda^0\right)\in(-1,1)$ and where $h(\lambda)$ is a positive bounded symmetric function with two continuous derivatives for $0<\lambda<\pi$.

C2 $\{x_t\}$ is a covariance stationary linear process

$$x_t=\sum_{j=0}^{\infty}\beta_j e_{t-j}, \quad \sum_{j=0}^{\infty}\beta_j^2<\infty \quad \beta_0=1$$

where e_t is a sequence of independent identically distributed random variables such that $E(e_t)=0$; $E\left[e_t^2\right]=1$; $E\left[|e_t|^\ell\right]=\mu_\ell$ for $\ell=3,...,p\geq 8$.

C3 For $\Lambda(\lambda)=\sum_{j=0}^{\infty}\beta_j e^{ij\lambda}$, if $\alpha=\alpha\left(\lambda^0\right)\neq 0$ at some frequency $\lambda^0\in[0,\pi]$

$$\frac{d}{d\lambda}\log\left(\left|\Lambda(\lambda)\right|\right)=O\left(\lambda^{-1}\right) \text{ as } \lambda\to\lambda^0.$$

C4 The weighting function $W(\lambda)$ is an even, nonnegative function on $[-\pi,\pi]$, twice continuously differentiable in $(0,\pi)$, bounded away from zero on $[0,\pi-\varepsilon)$ for some $\varepsilon\in(0,\pi)$ which integrates 1.

C5 $\frac{n\log m}{m^2}+\frac{m^5\log^5 m(\log\log m)}{n^4}\to 0$ as $n\to\infty$, and $k=m[\log m]$.

Condition C1 indicates that under H_0, the "memory" parameter α is equal to 0 for all $\lambda^0\in[0,\pi]$, so that $f(\lambda)=h(\lambda)$, that is $f(\lambda)$ is continuously differentiable in $(0,\pi)$. On the other hand, under H_1, there exists λ^0 for which $\alpha\left(\lambda^0\right)\neq 0$, which implies that f is only twice differentiable outside any open set containing $\pm\lambda^0$. Moreover, it does not allow $f(\lambda)$ to have more than one frequency λ at which $\alpha(\lambda)\neq 0$, although it can be easily relaxed at the expense of unnecessarily complicating the notation and mathematics. Condition C2 is standard and not very strong besides the linearity condition that it implies. Condition C3 was used elsewhere, see Robinson (1995*a*), so his comments apply here. Examples of weighting functions $W(\lambda)$ satisfying Condition C4 are

$$W(\lambda)=\pi^{-1}\left(1-\pi^{-1}|\lambda|\right) \quad\text{and}\quad W(\lambda)=\frac{1}{2}\left(1-\cos(\lambda)\right),\quad \lambda\in[-\pi,\pi].$$

To simplify the notation and arguments, in what follows we take $w_{|p|}=1$ for all $|p|=1,...,k$. Let us introduce

$$\overline{\alpha}(s)=\frac{1}{2}\left(\frac{1}{k}\left\{\sum_{p=1}^{k}+\sum_{p=-k}^{-1}\right\}\log\left(\lambda_{|p|}^{\alpha(s)}\overline{f}_p(s)\right)-\log\left(\lambda_{|k|}^{\alpha(s)}\overline{f}_k(s)\right)-\log\left(\lambda_{|k|}^{\alpha(s)}\overline{f}_{-k}(s)\right)\right),$$

where

$$\overline{f}_p(s)=\frac{1}{2m+1}\sum_{j=-m}^{m}W_j\lambda_{|p+j|+}^{-\alpha(s)},$$

with $|q|_+=\max\{1,|q|\}$.

Proposition 3.1. *Assuming C1-C5, for any finite collection $s_1,...,s_q$ such that $\min_i|s_i-s_{i+1}|>2m$, as $n\to\infty$*

$$(2m)^{1/2}\left(\widehat{\alpha}(s_1)-\widetilde{\alpha}(s_1),...,\widehat{\alpha}(s_q)-\widetilde{\alpha}(s_q)\right)\xrightarrow{d}\mathcal{N}\left(0,\Phi^2 diag(1,...,1)\right)$$

where $\Phi=\left(\int_{-1}^{1}W^2(\lambda)\,d\lambda\right)^{1/2}$ and $\widetilde{\alpha}(s)=\overline{\alpha}(s)+\alpha(s)$.

Proposition 1 forms the basis to perform the test for (2.2) and by extension to the hypotheses testing given in (2.7) and/or (2.8). Note that under H_0, e.g. $\alpha(s)=0$ for all $s=0,...,\tilde{n}$, Proposition 1 indicates that

$$(2m)^{1/2}\left(\widehat{\alpha}(s_1),...,\widehat{\alpha}(s_q)\right)\xrightarrow{d}\mathcal{N}\left(0,\Phi^2 diag(1,...,1)\right).$$

Moreover, an important consequence of Proposition 1 is that the asymptotic distribution does not depend on f, e.g. on the dependence structure of x_t. In other words, whether x_t is *iid* sequence of random variables or follows, for example, an $ARMA$ model, the asymptotic distribution is unaltered. This observation will have relevant consequences when implementing the bootstrap test for the hypothesis testing given in (2.2), and by extension the hypothesis testings in (2.7) or (2.8).

We now state our main results.

Theorem 3.2. *Let $\delta \in (0,1)$ be such that $M = n^{\delta}$. Assuming C1-C5, under H_0 given in (2.7) or (2.8), as $n \to \infty$, for $j = 2, 3$*

$$Prob\left\{\phi_n\left((2m)^{1/2}\, \Phi^{-1}\mathcal{T}_j - b_n\right) \le z\right\} \to \exp\left(-e^{-z}\right),$$

with $\phi_n = (2\log M)^{1/2}$,

$$b_n = \phi_n + \phi_n^{-1}\left\{\log\left(\frac{K_1(W)}{\pi^{1/2}}\right) + \frac{1}{2}\left[\log\delta + \log\log n\right]\right\}$$

where

$$K_1(W) = \frac{W^2(-1) + W^2(1)}{2\int_{-1}^{1} W^2(x)\,dx}$$

if $K_1(W) > 0$, and otherwise

$$b_n = \phi_n + \phi_n^{-1}\log\left[\frac{1}{\pi}\left(\frac{K_2(W)}{2}\right)^{1/2}\right]$$

where $K_2(W) = \left(\int_{-1}^{1} W'(x)^2\,dx\right) / \left(2\int_{-1}^{1} W(x)^2\,dx\right)$ and $W'(x) = \frac{d}{dx}W(x)$.

Remark 2. *The uniform kernel $W(x) = 1/2$, $|x| < 1$ falls under the first case, while the triangular or Hanning-Tukey kernels fall under the second case.*

Remark 3. *The techniques of proof of the result may readily be adapted to prove limit theorems as that of Woodroofe (1967) or Van-Ness and Woodroofe (1967) for the maximum deviation at the points λ_{2mj}, at $j = 0, 1, ..., M$, respectively, where b_n and ϕ_n are as defined in Theorem 3.1.*

Theorem 3.3. *Assuming C1-C5, under H_0 given in (2.2), as $n \to \infty$*

$$Prob\left\{\phi_n\left((2m)^{1/2}\, \Phi^{-1}\mathcal{T}_1 - b_n\right) \le z\right\} \to \exp\left(-2e^{-z}\right),$$

where b_n and ϕ_n are as defined in Theorem 3.1.

Proof. The proof of this theorem is straightforward from Theorem 3.1, after we observe that the limiting distribution of the $\sup_p$ and $\inf_p$ are independent random variables, see for example Theorem 1.8.3 of Leadbetter et al. (1983). □

Once we have obtained under H_0 the limiting distribution of $\mathcal{T}_j$, the next step is to show that $\mathcal{T}_j$ is consistent. The later is a basic feature for any test. Moreover, to gain some idea about the test, it is also convenient to compute the limiting power function. This is obtained by computing the limit distribution of $\mathcal{T}_j$ under a sequence of (contiguous) alternatives that approach the null at some appropriate rate. For that purpose, consider

$$H_a: \quad \exists\lambda^0 \in [0,\pi] \text{ such that } f(\lambda) \sim C\left(\lambda^0\right)\left|\lambda - \lambda^0\right|^{-\alpha_m\left(\lambda^0\right)}$$

where $\alpha_m\left(\lambda^0\right) = \left(m \log^2 M\right)^{-1/2} \alpha$ with $\alpha \neq 0$. Then, we have the following corollary

Corollary 3.4. *Assuming C1-C5, under H_a, as $n \to \infty$*

$$Prob\left\{\phi_n\left((2m)^{1/2}\ \Phi^{-1}\mathcal{T}_1 - b_n\right) \leq z\right\} \to \exp\left(-2e^{-(z-\alpha)}\right),$$

and for $j = 2, 3$

$$Prob\left\{\phi_n\left((2m)^{1/2}\ \Phi^{-1}\mathcal{T}_j - b_n\right) \leq z\right\} \to \exp\left(-e^{-(z-\alpha)}\right),$$

where b_n and ϕ_n are as defined in Theorem 3.1.

From the results of the previous corollary, we can expect the test to be consistent, since for fixed alternatives, that is $\alpha_m\left(\lambda^0\right) = \alpha$, Corollary 1 suggests that

$$\lim_{m\to\infty} \text{Prob}\left\{\phi_n\left(m^{1/2}\Phi^{-1}\mathcal{T}_1 - b_n\right) \leq z\right\} = \lim_{m\to\infty} \exp\left(-2e^{-\left(z-\left(m\log^2 M\right)^{1/2}\alpha\right)}\right) = 0,$$

for example. This is formalized in the next corollary

Corollary 3.5. *Assuming C1-C5, under H_1, as $n \to \infty$*

$$\lim_{m\to\infty} Prob\left\{\phi_n\left((2m)^{1/2}\ \Phi^{-1}\mathcal{T}_j - b_n\right) \leq z\right\} = 0$$

for all $z \in R$ and $j = 1, 2, 3$.

Results of Theorems 3.1 and 3.2 give justification to perform the tests. On the other hand, following results in Hall (1979), it is known that the rate of convergence of the finite sample distribution to the asymptotic one is very slow. In particular, that rate is logarithmic. Moreover, following Konakov and Piterbarg (1984) for other although related problems, this rate appears to be the best possible one. So, critical values relying on the asymptotic distribution can be a poor approximation to the finite sample distribution. One solution could be to employ Edgeworth expansions. But even here the rate is only improved to be $O\left(T^{-\delta}\right)$ for any $\delta > 0$, see for instance Konakov and Piterbarg (1984). So that, the rate is still very slow. In addition, Hall (1990) has shown that Edgeworth corrections do not do a good job, compared to bootstrap schemes, at the tails of the distribution, which is precisely the most interesting region when testing. Because of that, in the next section, we propose to use a bootstrap scheme to obtain a better approximation of the finite sample distribution than that based on it asymptotic counterpart.

4. BOOTSTRAP TEST FOR H_0

In this section we describe and analyze the bootstrap analogue to the statistic in (2.9), and by extension those given in (2.10) and (2.11). The resampling scheme must be such that, given the sample $\chi = \{x_1, ..., x_n\}$, the conditional distribution of the bootstrap statistic $\mathcal{T}_j^*$, $j = 1, 2, 3$, consistently estimates the distribution of $\mathcal{T}_j$ for $j = 1, 2, 3$, under the null hypothesis and local alternatives. That is, $\mathcal{T}_j^* \to_{d^*} \mathcal{T}_j$ in probability, where "$\to_{d^*}$in probability" means convergence in bootstrap distribution according to the following definition, see Giné and Zinn (1990),

Definition 4.1. *Let χ^* denote the bootstrap sample drawn from χ using some given resampling scheme. Let $\mathcal{T}_j^*$ be the test statistic computed from χ^*. We say that $\mathcal{T}_j^*$ converges weakly in bootstrap distribution to the random variable $\mathcal{T}_j$ (with distribution function $G(z)$), and denoted as $\mathcal{T}_j^* \to_{d^*} \mathcal{T}_j$ in probability, whenever the sequence of random variables $\Pr(\mathcal{T}_1^* \le z \mid \chi)$ converges to $G(z)$ in probability for every continuity point z of $G(z)$.*

A second requirement for the bootstrap test to be valid, and thus consistent, is that under the alternative hypothesis H_1, the bootstrap statistic must also converge in bootstrap distribution although, possibly, to a different distribution. In our framework, as Theorem 4.1 below shows, the (asymptotic) distribution of $\mathcal{T}_j^*$ is the same under both the null and alternative hypothesis.

Following our comments made after Proposition 1, since the asymptotic distribution of $\widehat{\alpha}(s)$ is independent of the dependence structure of x_t, it implies that the limit distribution of $\mathcal{T}_j$, $j = 1, 2, 3$, is also independent of the covariance structure of x_t. This observation prompts us to bootstrap the tests $\mathcal{T}_j$ in (2.9) – (2.11) as follows.

STEP 1 Draw independent bootstrap observations x_t^*, $t = 1, ..., n$, from the empirical distribution of $\chi = \{x_1, ..., x_n\}$, that is, for all $t = 1, ..., n$,

$$\Pr\{x_t^* = x_s\} = n^{-1},\ s = 1, ..., n.$$

STEP 2 Compute the bootstrap periodogram of x_t^*, $t = 1, ..., n$,

$$I_\ell^* = \left| (2\pi n)^{-1/2} \sum_{t=1}^{n} x_t^* e^{it\lambda_\ell} \right|^2,\ \ell = 1, ..., \tilde{n},$$

with $I_0^* = 0$, and the bootstrap spectral density estimator

$$\widehat{f}_p^*(s) = \frac{1}{2m+1} \sum_{j=-m}^{m} W_j I_{j+p+s}^*. \tag{4.1}$$

STEP 3 Compute the bootstrap analogue of $\widehat{\alpha}(s)$, that is

$$\widehat{\alpha}^*(s) = \frac{1}{2}\left(\frac{1}{k}\left\{ \sum_{p=1}^{k} + \sum_{p=-k}^{-1} \right\} \log \widehat{f}_p^*(s) - \left(\log \widehat{f}_k^*(s) + \log \widehat{f}_{-k}^*(s) \right) \right), \tag{4.2}$$

and finally,

STEP 4 Compute the bootstrap test for H_0 against H_1 given in (2.7) as

$$\mathcal{T}_1^* = \sup_{s=0,1,...,\tilde{n}} |\widehat{\alpha}^*(s)|, \tag{4.3}$$

whereas for the hypothesis testing in (2.7) and (2.8) is

$$\mathcal{T}_2^* = \sup_{s=0,1,\ldots,\tilde{n}} \widehat{\alpha}^*(s) \quad \text{and } \mathcal{T}_3^* = \sup_{s=0,1,\ldots,\tilde{n}} -\widehat{\alpha}^*(s), \tag{4.4}$$

respectively.

Remark 4. *Observe that the resample was done from the raw observations x_t instead of the rescale one*

$$\widetilde{x}_t = \frac{(x_t - \overline{x})}{\widehat{\sigma}_x}, \quad t = 1, ..., n$$

where $\overline{x} = n^{-1} \sum_{t=1}^{n} x_t$ and $\widehat{\sigma}_x^2 = n^{-1} \sum_{t=1}^{n} (x_t - \overline{x})^2$. The reason is because I_ℓ^* entails sample-mean correction and from the definition of $\widehat{\alpha}^*(s)$ given in (4.2), it is obvious that it is also invariant to multiplicative constants, e.g. $\widehat{\sigma}_x$.

Another possibility to bootstrap $\widehat{\alpha}(s)$ is following Franke and Härdle (1992), see also Dahlhaus and Janas (1996) for similar ideas, as follows:

STEP 1' For $p = 1, ..., \tilde{n}$, compute

$$v_p = \left(\tilde{n}^{-1} \sum_{p=1}^{\tilde{n}} \widehat{f}_p^{-1} I_p \right)^{-1} \left(\widehat{f}_p^{-1} I_p \right).$$

STEP 2' Resample with replacement from the empirical distribution of v_p, $p = 1, ..., \tilde{n}$. That is, obtain the boostrap sample I_p^*, $p = 1, ..., \tilde{n}$, which satisfies for all p

$$\Pr \left\{ I_p^* = v_\ell \right\} = \tilde{n}^{-1}, \; \ell = 1, ..., \tilde{n},$$

with $I_0^* = 0$.

STEP 3' Given the random sample I_p^*, $p = 1, ..., \tilde{n}$, compute the bootstrap analogue to the spectral density estimator

$$\widehat{f}_p^*(s) = \frac{1}{2m+1} \sum_{j=-m}^{m} W_j I_{j+p+s}^* \tag{4.5}$$

and define the bootstrap analogue of $\widehat{\alpha}(s)$ as

$$\widehat{\alpha}^*(s) = \frac{1}{2} \left(\frac{1}{k} \left\{ \sum_{p=1}^{k} + \sum_{p=-k}^{-1} \right\} \log \widehat{f}_p^*(s) - \left(\log \widehat{f}_k^*(s) + \log \widehat{f}_{-k}^*(s) \right) \right), \tag{4.6}$$

and finally,

STEP 4' Compute the bootstrap test as in *STEP 4* but using (4.6) instead of (4.2).

Remark 5. *In the definition of $\widehat{f}_p^*(s)$ given in (4.5), we could have defined it as*

$$\check{f}_p^*(s) = \frac{\widehat{f}_p}{2m+1} \sum_{j=-m}^{m} W_j I_{j+p+s}^*. \tag{4.7}$$

However, this is irrelevant for our purposes. Indeed, if (4.7) *were used, from the definition of* $\widehat{\alpha}^*(s)$ *in* (4.6), *the bootstrap analogue of* $\widehat{\alpha}(s)$ *would become*

$$\overline{\alpha}^*(s) \overset{def}{=} \widehat{\alpha}^*(s) + \widehat{\alpha}(s),$$

which implies that if (4.7) *were chosen, the definition of the bootstrap test would be given by*

$$\overline{\mathcal{T}}_j^* \overset{def}{=} \mathcal{T}_j^* - \widehat{\alpha}(s)$$

where $\mathcal{T}_j^*$, $j = 1, 2, 3$, *is given in* (4.3) *and* (4.4). *That is, the difference between* $\mathcal{T}_j^*$ *and* $\overline{\mathcal{T}}_j^*$ *would be that the latter had to be centered around its "sample mean"* $\widehat{\alpha}(s)$.

Theorem 4.2. *Assuming C1-C5, under the maintain hypothesis, e.g.* $H_0 \cup H_1$, *as* $n \to \infty$

$$Prob\left\{ \phi_n \left((2m)^{1/2}\, \Phi^{-1} \mathcal{T}_1^* - b_n \right) \le z \middle| \chi \right\} \xrightarrow{P} \exp\left(-2e^{-z}\right),$$

and for $j = 2, 3$

$$Prob\left\{ \phi_n \left((2m)^{1/2}\, \Phi^{-1} \mathcal{T}_j^* - b_n \right) \le z \middle| \chi \right\} \xrightarrow{P} \exp\left(-e^{-z}\right),$$

where ϕ_n *and* b_n *are as in Theorem 3.1.*

Theorem 4.1 gives the theoretical justification for the bootstrap test. However, since the conditional distribution of $\mathcal{T}_j^*$ given χ is computationally very difficult, if possible, to obtain, it implies that the bootstrap critical value of the test at, say, β significance level $c^*_{n(1-\beta)}$ can not be obtained. So, the value $c^*_{n(1-\beta)}$, which satisfies $\Pr\left(\mathcal{T}_j^* > c^*_{n(1-\beta)} \middle| \chi \right) = \beta$, has to be approximated via Monte Carlo, as accurate as desired. To that end, we can use the quantiles obtained from the empirical distribution of the Monte Carlo sample of $\mathcal{T}_j^*$ as estimators of the corresponding quantiles of $\mathcal{T}_1$. That is, consider B bootstrap samples of size n, $\underset{\sim}{x}^{*(k)} = \left(x_1^{*(k)}, x_2^{*(k)}, ..., x_n^{*(k)} \right)'$, $k = 1, ..., B$, and compute for each of the samples the corresponding test statistic value $\mathcal{T}_j^{*(k)}$ as in *STEP 3*. Then, the critical value $c^*_{n(1-\beta)}$ is approximated by $c^*_{n(1-\beta),B}$, where $c^*_{n(1-\beta),B}$ satisfies

$$B^{-1} \sum_{k=1}^{B} \mathcal{I}\left(\mathcal{T}_j^{*(k)} > c^*_{n(1-\beta),B} \right) = \beta.$$

That is, $c^*_{n(1-\beta),B}$ is the $(1-\beta)\, th$ quantile of the Monte Carlo sample $\left(\mathcal{T}_j^{*(k)}, k = 1, ..., B \right)$, so that the null hypothesis is rejected when $\mathcal{T}_j > c^*_{n(1-\beta),B}$. Finally, if the second bootstrap approach were implemented, then the procedure would be the same as above but with $\underset{\sim}{x}^{*(k)}$ and *STEP 3* being replaced by I_p^* and *STEP 3* respectively.

APPENDIX A

Proof of Proposition 1 Let us introduce,

$$\widetilde{f}_p(s) = \frac{1}{2m+1} \sum_{j=-m}^{m} W_j \left(f_{j+p+s} \mathcal{I}(j + p \neq 0) + f_{s+1} \mathcal{I}(j + p = 0) \right).$$

From the definition of $\widehat{\alpha}(s)$ and $\widetilde{\alpha}(s)$, $\widehat{\alpha}(s)-\widetilde{\alpha}(s) := 2^{-1}\sum_{q=1}^{5} a_q(s)$ where

$$\begin{aligned}
a_1(s) &= \frac{1}{k}\sum_{p=1}^{k}\left(\log \widehat{f}_p(s)/\widetilde{f}_p(s)\right)-\left(\log \widehat{f}_k(s)/\widetilde{f}_k(s)\right)\\
a_2(s) &= \frac{1}{k}\sum_{p=1}^{k}\left(\log \widehat{f}_{-p}(s)/\widetilde{f}_{-p}(s)\right)-\left(\log \widehat{f}_{-k}(s)/\widetilde{f}_{-k}(s)\right)\\
a_3(s) &= \frac{1}{k}\sum_{p=1}^{k}\left(\log \widetilde{f}_p(s)/\overline{f}_p(s)\right)-\left(\log \widetilde{f}_k(s)/\overline{f}_k(s)\right) \qquad \text{(A.1)}\\
a_4(s) &= \frac{1}{k}\sum_{p=1}^{k}\left(\log \widetilde{f}_{-p}(s)/\overline{f}_{-p}(s)\right)-\left(\log \widetilde{f}_{-k}(s)/\overline{f}_{-k}(s)\right)\\
a_5(s) &= -2\alpha(s)\left(\frac{1}{k}\sum_{p=1}^{k}\log(p/k)-1\right).
\end{aligned}$$

Because by Lemma 2 of Robinson (1995*b*), $k^{-1}\sum_{p=1}^{k}\log(p/k)+1 = O\left(k^{-1}\log k\right)$, the proposition is shown if $a_3(s)+a_4(s) = o\left(m^{-1/2}\right)$ and for any finite set of constants $\psi_1,...,\psi_q$

$$(2m)^{1/2}\sum_{\ell=1}^{q}\psi_\ell\left(a_1(s_\ell)+a_2(s_\ell)\right)\xrightarrow{d} N\left(0,\Phi^2\sum_{\ell=1}^{q}\psi_\ell^2\right). \qquad \text{(A.2)}$$

Next, we examine $a_3(s)+a_4(s) := a_{31}(s)+a_{32}(s)+a_{41}(s)+a_{42}(s)$, where

$$\begin{aligned}
a_{31}(s) &= \frac{1}{k}\sum_{p=1}^{m}\left(\log \widetilde{f}_p(s)/\overline{f}_p(s)\right) \text{ and}\\
a_{32}(s) &= \frac{1}{k}\sum_{p=m+1}^{k}\left(\log \widetilde{f}_p(s)/\overline{f}_p(s)\right)-\left(\log \widetilde{f}_k(s)/\overline{f}_k(s)\right),
\end{aligned}$$

and likewise $a_{41}(s)$ and $a_{42}(s)$. We first estimate $a_{32}(s)+a_{42}(s)$. By C1,

$$\begin{aligned}
\overline{f}_p^{-1}(s)\widetilde{f}_p(s)-1 &= \overline{f}_p^{-1}(s)\frac{1}{2m+1}\sum_{j=-m}^{m}\lambda_{j+p}^{-\alpha(s)}\left(g_{j+p+s}-g_s\right)\\
&= \overline{f}_p^{-1}(s)\left(\frac{1}{2m+1}\sum_{j=-m}^{m}\lambda_{j+p}^{-\alpha(s)}\left(\lambda_{j+p}g'(\lambda_s)+2^{-1}\lambda_{j+p}^2 g''\left(\overline{\lambda}\right)\right)\right)\\
&= \overline{f}_p^{-1}(s)\left(\frac{1}{2m+1}\sum_{j=-m}^{m}\lambda_{j+p}^{-\alpha(s)}\lambda_{j+p}g'(\lambda_s)\right)+O\left(\lambda_{2k}^2\right),
\end{aligned}$$

where in the second equality we have employ Taylor expansion, being $\overline{\lambda}$ is an intermediate point between λ_s and λ_{s+p+j}, and for the equality that $\overline{f}_p^{-1}(s)(2m+1)^{-1}\sum_{j=-m}^{m}\lambda_{j+p}^{-\alpha(s)+2} = O\left(\lambda_{2k}^2\right)$. Likewise,

$$\overline{f}_{-p}^{-1}(s)\widetilde{f}_{-p}(s)-1 = \overline{f}_{-p}^{-1}(s)\frac{1}{2m+1}\sum_{j=-m}^{m}\lambda_{j+p}^{-\alpha(s)}\left(g_{s-j-p}-g_s\right)$$

$$= \overline{f}_p^{-1}(s)\left(\frac{1}{2m+1}\sum_{j=-m}^{m}\lambda_{j+p}^{-\alpha(s)}\left(-\lambda_{j+p}g'(\lambda_s)+2^{-1}\lambda_{j+p}^2 g''\left(\overline{\overline{\lambda}}\right)\right)\right)$$

$$= -\overline{f}_p^{-1}(s)\left(\frac{1}{2m+1}\sum_{j=-m}^{m}\lambda_{j+p}^{-\alpha(s)}\lambda_{j+p}g'(\lambda_s)\right)+O\left(\lambda_{2k}^2\right),$$

where $\overline{\overline{\lambda}}$ is an intermediate point between λ_s and λ_{s-p-j} and because $\overline{f}_{-p}(s)=\overline{f}_p(s)$. Thus, by the mean value theorem, for some $\xi=\xi(p)$ with $|\xi|<1$,

$$\left|\log\left(\frac{\widetilde{f}_p(s)}{\overline{f}_p(s)}-1+1\right)+\log\left(\frac{\widetilde{f}_{-p}(s)}{\overline{f}_{-p}(s)}-1+1\right)\right|$$

$$\leq \left|\left(\frac{\widetilde{f}_p(s)}{\overline{f}_p(s)}-1\right)+\left(\frac{\widetilde{f}_{-p}(s)}{\overline{f}_{-p}(s)}-1\right)+\left(\frac{\widetilde{f}_p(s)}{\overline{f}_p(s)}-1\right)\left(\frac{\widetilde{f}_{-p}(s)}{\overline{f}_{-p}(s)}-1\right)\right|\left|\frac{1}{1+\xi\left(\overline{f}_p^{-1}(s)\widetilde{f}_p(s)-1\right)}\right|$$

because $|\log(ce-1+1)|\leq K|ce-1|$ and $(ce-1)=(c-1)+(e-1)+(c-1)(e-1)$, which implies that $(2m)^{1/2}(a_{32}(s)+a_{42}(s))=O\left(m^{1/2}\lambda_{2k}^2\right)=o(1)$ by C5. Proceeding similarly, we obtain that $(2m)^{1/2}(a_{31}(s)+a_{41}(s))=O\left(m^{3/2}\lambda_m^2/k\right)=o\left(m^{1/2}\lambda_m^2\right)$ by C5 and that $(2m)^{1/2}(a_3(s)+a_4(s))=o(1)$.

To complete the proof we need to show $(A.2)$. To this end, and denoting $g_p(s)=\widetilde{f}_p^{-1}(s)\widehat{f}_p(s)-1$, it suffices to show that

$$(2m)^{1/2}\sum_{\ell=1}^{q}\psi_\ell\left(b_k(s_\ell)+b_{-k}(s_\ell)\right)\xrightarrow{d}N\left(0,\Phi^2\sum_{\ell=1}^{q}\psi_\ell^2\right) \tag{A.3}$$

where $b_k(s_\ell)=k^{-1}\sum_{p=1}^{k}g_p(s_\ell)-g_k(s_\ell)$ and $b_{-k}(s_\ell)=k^{-1}\sum_{p=1}^{k}g_{-p}(s_\ell)-g_{-k}(s_\ell)$, and for all $\ell=1,...,q$,

$$(2m)^{1/2}\left\{(a_1(s_\ell)-b_k(s_\ell))+(a_2(s_\ell)-b_{-k}(s_\ell))\right\}=o_p(1). \tag{A.4}$$

We begin showing $(A.3)$. Write $c_k(s_\ell)=\left((2m)^{1/2}/k\right)\sum_{p=1}^{k}g_p(s_\ell)$. First,

$$E(c_k(s_\ell))=(2m)^{1/2}\frac{1}{k}\left\{\sum_{p=1}^{m}E(g_p(s_\ell))+\sum_{p=1+m}^{2m-1}E(g_p(s_\ell))+\sum_{p=2m}^{k}E(g_p(s_\ell))\right\}.$$

By Proposition A.1 part (a) of Hidalgo and Robinson (2001), the first term on the right of the last displayed equation is

$$O\left(\frac{1}{k}\sum_{p=1}^{m}\left(m^{\alpha(s)-1/2}\frac{\log m}{p^{\alpha(s)}}\mathcal{I}(\alpha(s)>0)+\frac{\log^2 m}{m^{1/2}}\mathcal{I}(\alpha(s)\leq 0)\right)\right)=O\left(\frac{m^{1/2}\log^2 m}{k}\right)=o(1)$$

by C5, whereas by Proposition A.1 part (b) of Hidalgo and Robinson (2001), the second term is bounded by

$$K\frac{m^{1/2}}{k}\sum_{p=1+m}^{2m-1}\left(\frac{\log^2 m}{m}\mathcal{I}(\alpha(s)\leq 0)+\frac{\mathcal{I}(\alpha(s)>0)\log m}{m^{1-\alpha(s)}(2p-m)^{\alpha(s)}}\right)=O\left(k^{-1}m^{1/2}\log^2 m\right).$$

Finally, the third term is $O\left(m^{-1/2}\log m\right) = o(1)$ by Proposition A.1 part (a) of Hidalgo and Robinson (2001) and C5.

Next, the variance of $c_k(s_\ell)$ is bounded by

$$mK\left(Var\left(\frac{1}{k}\sum_{p=1}^{m} g_p(s_\ell)\right) + Var\left(\frac{1}{k}\sum_{p=1+m}^{2m-1} g_p(s_\ell)\right) + Var\left(\frac{1}{k}\sum_{p=2m}^{k} g_p(s_\ell)\right)\right). \tag{A.5}$$

The third term of $(A.5)$ is

$$\begin{aligned} & m\left\{\frac{K}{k^2}\sum_{p=2m}^{k} Var\left(g_p(s_\ell)\right) + \frac{2K}{k^2}\sum_{2m\le p<q}^{k} Cov\left(g_p(s_\ell), g_q(s_\ell)\right)\right\} \\ = \quad & O\left(\frac{1}{k}\right) + \frac{Km}{k^2}\sum_{2m\le p<q}^{k} Cov\left(g_p(s_\ell), g_q(s_\ell)\right) \end{aligned}$$

because Hidalgo and Robinson's (2001) Proposition A.2 implies that $mVar\left(g_p(s_\ell)\right) = O(1)$ for $p \ge m$. The second term on the right of the last displayed equation is bounded in absolute value by

$$\frac{Km}{k^2}\left(\sum_{\mathcal{J}_{p,q}} Var^{1/2}\left(g_p(s_\ell)\right) Var^{1/2}\left(g_q(s_\ell)\right) + \sum_{\mathcal{J}_{p,q}^c}\left|Cov\left(g_p(s_\ell), g_q(s_\ell)\right)\right|\right) = o(1),$$

where $\mathcal{J}_{p,q} = \{2m \le p < q \le k, |q-p| \le m\}$ and $\mathcal{J}_{p,q}^c = \{2m \le p < q \le k, |q-p| > m\}$, as we now show. The second term on the left is $O\left(k^{-1}m\sum_{p=2m}^{k}\left(n^{-1}+m^{-3/2}\right)\right) = o(1)$ because by Proposition 2 part (a) of Hidalgo and Yajima (2000), because $p, q > 2m$ implies $p^{-1}q^{-1/2} < m^{-3/2}$, whereas the first term is also $o(1)$ because the sum $\sum_{\mathcal{J}_{p,q}}$ has at most km terms and by Proposition 2 part (a) of Hidalgo and Yajima (2000), $Var\left(g_p(s_\ell)\right) = O\left(m^{-1}\right)$ and by C5 $m/k \to 0$. So, we conclude that the third term of $(A.5)$ is $o(1)$.

Next, the second term of $(A.5)$ is bounded by

$$K\frac{m^2}{k^2}\sum_{p=1+m}^{2m-1} Var\left(g_p(s_\ell)\right) = o(1),$$

since by Proposition A.2 part (b) of Hidalgo and Robinson (2001), the left side is $O\left(k^{-2}\left(m^2 + m^{1+\alpha(s)}\log^2 m\right)\right) = o(1)$ if $\alpha(s) \in (-1, 1/2)$, for $\alpha(s) = 1/2$ using that $\left|m^{-1}\sum_{p=1}^{3m/2}\log\left((p+m)/p\right)\right| \to \left|\int_0^{3/2}\left(\log(v+1) - \log(v)\right)dv\right| \le K$, and for $\alpha(s) \in (1/2, 1)$, because

$$O\left(\frac{m^2}{k^2}\sum_{p=1}^{3m/2}\frac{m^{2\alpha(s)-2}}{p^{2\alpha(s)-1}} + \frac{m^{1+\alpha(s)}\log^2 m}{k^2}\right) = O\left(\frac{m^2 + m^{1+\alpha(s)}\log^2 m}{k^2}\right) = o(1),$$

using that $m^{-1}\sum_{p=1}^{m-1}(p/m)^{1-2\alpha(s)} \le K$. Finally, the first term of $(A.5)$ is also $o(1)$ using Proposition 2 part (c) of Hidalgo and Yajima (2000). Thus, we obtain that

$c_k(s_\ell) = \left((2m)^{1/2}/k\right) \sum_{p=1}^k g_p(s_\ell) = o_p(1)$. Likewise, $c_{-k}(s_\ell) = \left((2m)^{1/2}/k\right) \sum_{p=1}^k g_{-p}(s_\ell) = o_p(1)$, which implies that

$$(2m)^{1/2}(b_k(s_\ell) + b_{-k}(s_\ell)) = -(2m)^{1/2}(g_k(s_\ell) + g_{-k}(s_\ell)) + o_p(1)$$

which converges in distribution to $\mathcal{N}\left(0, \Phi^2\right)$ by Lemma 1 in Appendix B. Form here $(A.3)$ holds true.

To complete the proof, we need to show $(A.4)$. First, for any arbitrary $\varepsilon, \eta > 0$,

$$\begin{aligned}\Pr\left\{\left|(2m)^{1/2} a_1(s) - b_k(s)\right| > \varepsilon\right\} &= \Pr\left\{\left|(2m)^{1/2} a_1(s) - b_k(s)\right| > \varepsilon;\ \sup_p |g_p(s)| \geq \eta\right\} \\ &\quad + \Pr\left\{\left|(2m)^{1/2} a_1(s) - b_k(s)\right| > \varepsilon;\ \sup_p |g_p(s)| < \eta\right\}.\end{aligned}$$

The first term on the right converges to zero since $\sup_p |g_p(s)| = o_p(1)$ by Proposition A.3 of Hidalgo and Robinson (2001), with $r = 2$ there. Since $|\log x - (x-1)| \leq 2^{-1}(x-1)^2$ for x close to 1, by Markov inequality, the first term on the right also converges to zero if

$$\begin{aligned}&\sum_{p=1}^k E\left(g_p^2(s) + g_k^2(s)\right) = \sum_{p=1}^k (Var(g_p(s)) + Var(g_k(s))) + \sum_{p=1}^k \left(E^2(g_p(s)) + E^2(g_k(s))\right) \\ &= o\left(\frac{k}{m^{1/2}}\right),\end{aligned}$$

which is the case by Hidalgo and Robinson's (2001) Propositions A.2 and A.1 respectively. Similarly,

$$\Pr\left\{\left|(2m)^{1/2} a_2(s) - b_{-k}(s)\right| > \varepsilon\right\} \to 0.$$

This concludes the proof of the theorem. □

Proof of Theorem 3.1 Using $(A.1)$,

$$\widehat{\alpha}(s) - \alpha_m(s) = \sum_{q=1}^5 a_q(s) + \overline{\alpha}(s).$$

Proceeding as with the proof of Proposition 1 and by Lemma 4, $\sup_{s=0,\ldots,\tilde{n}} \sum_{j=3}^5 |a_j(s)| = o\left(m^{-1/2}\log^{-1} n\right)$ and $\sup_{s=0,\ldots,\tilde{n}} |\widetilde{\alpha}(s)| = o\left(m^{-1/2}\log^{-1} n\right)$, respectively. So, we are left to examine the behaviour of $a_1(s)$. Let $h_1(s) = a_1(s) - c_1(s)$ and $h_2(s) = a_2(s) - c_2(s)$, where

$$\begin{aligned}c_1(s) &= \frac{1}{k}\sum_{p=1}^k \left\{\left(\frac{\widehat{f}_p(s) - \widetilde{f}_p(s)}{\widetilde{f}_p(s)}\right) - \left(\frac{\widehat{f}_k(s) - \widetilde{f}_k(s)}{\widetilde{f}_k(s)}\right)\right\}, \\ c_2(s) &= \frac{1}{k}\sum_{p=1}^k \left\{\left(\frac{\widehat{f}_{-p}(s) - \widetilde{f}_{-p}(s)}{\widetilde{f}_{-p}(s)}\right) - \left(\frac{\widehat{f}_{-k}(s) - \widetilde{f}_{-k}(s)}{\widetilde{f}_{-k}(s)}\right)\right\}.\end{aligned}$$

We show first that

$$(2m)^{1/2} \sup_{s=0,\ldots,\tilde{n}} |h_j(s)| = o_p(1), \quad j = 1, 2. \tag{A.6}$$

Since the proof for $j = 1$ and $j = 2$ are similar, we only handled the case $j = 1$. For any arbitrary $\varepsilon, \eta > 0$,

$$\begin{aligned}
\Pr\left\{\sup_{q=0,\dots,\tilde{n}} \left|2m^{1/2} h_1(s)\right| > \varepsilon\right\} &= \Pr\left\{\sup_{s=0,\dots,\tilde{n}} \left|(2m)^{1/2} h_1(s)\right| > \varepsilon;\ \sup_{q=0,\dots,\tilde{n}} \left|\xi_q\right| \geq \eta\right\} \\
&\quad + \Pr\left\{\sup_{s=0,\dots,\tilde{n}} \left|(2m)^{1/2} h_1(s)\right| > \varepsilon;\ \sup_{q=0,\dots,\tilde{n}} \left|\xi_q\right| < \eta\right\},
\end{aligned}$$

where $\xi_q = \left(\widetilde{f}_q^{-1} \widehat{f}_q - 1\right)$. The first term on the right of the last displayed equation converges to zero since Propositions A.1 and A.3 of Hidalgo and Robinson (2001) imply that $\sup_q \left|\xi_q\right| = o_p(1)$. So, using the definition of $a_1(s)$ and that $\left|\log x - (x-1)\right| \leq 2^{-1}(x-1)^2$ for x close to 1, we obtain that

$$\begin{aligned}
(2m)^{1/2} \sup_{s=0,\dots,\tilde{n}} \left|h_1(s)\right| &\leq K \sup_{s=0,\dots,\tilde{n}} \frac{(2m)^{1/2}}{k} \sum_{q=1}^{k} \left\{g_q^2(s) + g_k^2(s)\right\} \\
&\leq K (2m)^{1/2} \sup_{p=0,\dots,\tilde{n}} \xi_p^2.
\end{aligned}$$

Therefore, to conclude the proof of $(A.6)$, it suffices to show that

$$(2m)^{1/2} \sup_{p=0,\dots,\tilde{n}} \left|\xi_p^2\right| = o_p\left(\log^{-1} n\right). \tag{A.7}$$

Denoting

$$\begin{aligned}
\psi_{1p} &= \frac{1}{2m+1} \sum_{j=-m, j+p \neq s}^{m} W_j f_{j+p} \left(\frac{I_{j+p}}{f_{j+p}} - (2\pi) \frac{I_{e,j+p}}{\sigma_e^2}\right) \\
\psi_{2p} &= \frac{1}{2m+1} \sum_{j=-m, j+p \neq s}^{m} W_j f_{j+p} \left(1 - \frac{E\left(I_{j+p}\right)}{f_{j+p}}\right) \\
\psi_{3p} &= \frac{1}{2m+1} \sum_{j=-m, j+p \neq s}^{m} W_j f_{j+p} \left((2\pi) \frac{I_{e,j+p}}{\sigma_e^2} - 1\right),
\end{aligned}$$

we obtain that

$$\xi_p = \widetilde{f}_p^{-1} \left(\psi_{1p} + \psi_{2p} + \psi_{3p} + \frac{\mathcal{I}\left(|p-s| < m\right)}{2m+1} I_s W_{s-p}\right) - \left(E g_p - 1\right). \tag{A.8}$$

Since $\sup_p \xi_p^2 = \left(\sup_p \left|\xi_p\right|\right)^2$, the contribution of the last term on the right of $(A.8)$ into the left side of $(A.7)$ is $o\left(\log^{-1} n\right)$ by $C5$ and because by Proposition A1 of Hidalgo and Robinson (2001) and that $K^{-1} < \left|j^{-\alpha_m(s)}\right| < K$ for $j = 1, \dots, \tilde{n}$, $\left|E\xi_p - 1\right| = O\left(n^{-1}m + m^{-1}\log m\right)$. Next, Hidalgo and Robinson's (2001) Proposition A.3 implies that the contribution due to $\psi_{1p}(s) + \psi_{2p}(s) + \frac{\mathcal{I}(|p-s|<m)}{2m+1} I_s W_{s-p}$ into the left side of $(A.7)$ is $O_p\left(m^{-1/2} \log^2 n\right) = o_p\left(\log^{-1} n\right)$. So, it remains to show that the contribution due to $\psi_{3p}(s)$ is also $o_p\left(\log^{-1} n\right)$, which follows by Lemma 3 in Appendix B. So, $(A.7)$ holds true.

Thus, to complete the proof of the theorem, it suffices to examine the limit distribution of $\phi_n \left(\sup_{s=0,\dots,\tilde{n}} (2m)^{1/2} \Phi^{-1} (c_1(s) + c_2(s)) - b_n \right)$. To that end, denote

$$
\begin{aligned}
c_{1,e}(s) &= \frac{1}{k} \sum_{p=1}^{k} \left\{ \widetilde{f}_p^{-1}(s) \left(\widehat{f}_e(\lambda_p + \lambda_s) - 1 \right) - \widetilde{f}_k^{-1}(s) \left(\widehat{f}_e(\lambda_k + \lambda_s) - 1 \right) \right\} \\
c_{2,e}(s) &= \frac{1}{k} \sum_{p=1}^{k} \left\{ \widetilde{f}_{-p}^{-1}(s) \left(\widehat{f}_e(\lambda_{-p} + \lambda_s) - 1 \right) - \widetilde{f}_{-k}^{-1}(s) \left(\widehat{f}_e(\lambda_{-k} + \lambda_s) - 1 \right) \right\}
\end{aligned}
$$

where

$$
\widehat{f}_e(\lambda) = \frac{1}{2m+1} \sum_{j=-m, j+p \neq s}^{m} W_j f(\lambda_j + \lambda) \frac{(2\pi) I_e(\lambda_j + \lambda)}{\sigma_\varepsilon^2}.
$$

We first show that

$$
\sup_{s=1,\dots,\tilde{n}} |c_j(s) - c_{j,e}(s)| = o_p\left(m^{-1/2} \log^{-1} n \right), \quad j = 1, 2. \tag{A.9}
$$

We deal only with the proof for $j = 1$, that for $j = 2$ being similarly handled. Proceeding as with the proof that $\sup_{s=0,\dots,\tilde{n}} \left| (2m)^{1/2} h_1(s) \right| = o_p\left(\log^{-1} n \right)$, $(A.9)$ holds true if

$$
\sup_{s=0,\dots,\tilde{n}} \left| \widetilde{f}_1^{-1}(s) \sum_{j=-m, j+p \neq s}^{m} W_j f_{j+1+s} \left(\frac{I_{j+1+s}}{f_{j+1+s}} - \frac{(2\pi)}{\sigma_\varepsilon^2} I_{e,j+1+s} \right) \right| = o_p\left(m^{1/2} \log^{-1} n \right).
$$

Writing $\sigma_\varepsilon^2 = 1$, without loss of generality, and observing that under $H_0 \cup H_a$, $K^{-1} < \widetilde{f}_1(s) < K$ for all $\lambda_s \in [0, \pi]$, we have that the left side of the last displayed equation is bounded by

$$
K \left\{ \sup_{s=0,\dots,m} + \sup_{s=m+1,\dots,\tilde{n}} \right\} \left| \sum_{j=-m, j+p \neq s}^{m} W_j f_{j+1+s} \left(\frac{I_{j+1+s}}{f_{j+1+s}} - (2\pi) I_{e,j+1+s} \right) \right|. \tag{A.10}
$$

Since $K^{-1} < f_j < K$ for all $j \neq s$ and $W(\lambda)$ is bounded, the expectation of the first term of $(A.10)$ is bounded by

$$
\sum_{s=1}^{2m} E \left| \frac{I_s}{f_s} - (2\pi) I_{e,s} \right| = O\left(mn^{-1} \right) = o\left(m^{1/2} \log^{-1} n \right)
$$

by C5, proceeding as in the proof of equation (4.8) of Robinson (1995*b*), but using Lemma 4 of Appendix B instead of Theorem 2 of Robinson (1995*a*) there. On the other hand, the second term of $(A.10)$ is $o_p\left(m^{1/2} \log^{-1} n \right)$ by Proposition A.3 of Hidalgo and Robinson (2001), which completes the proof of $(A.9)$.

To finish the proof we need to examine $\psi_n = \phi_n \left((2m)^{1/2} \Phi^{-1} \sup_{s=0,\dots,\tilde{n}} (c_{1,e}(s) + c_{2,e}(s)) - b_n \right)$ under H_0. Note that under H_0, $\alpha_m(s)$ and $\overline{\alpha}(s)$ are zero. We first give an intuition/heuristic explanation about what we should expect regarding the behaviour of ψ_n. Because $(2m)^{1/2} c_{j,e}(s)$, $j = 1, 2$, behaves as independent random variables when evaluated at $\lambda_{2m\ell}$ for $\ell = 1, \dots, M$, whereas for any λ^1 and λ^2 such that

$\left|\lambda^1-\lambda^2\right| \leq 4\pi m/n$, $(2m)^{1/2}\left(c_{1,e}\left(\lambda^1\right)+c_{2,e}\left(\lambda^1\right)\right)$ and $(2m)^{1/2}\left(c_{1,e}\left(\lambda^2\right)+c_{2,e}\left(\lambda^2\right)\right)$ has a correlation structure which satisfies condition (v) of Theorem A.1 of Bickel and Rosenblatt (1973), see also Pickands (1969) equation (1.2). Hence, we should expect that the limiting distribution of ψ_n corresponds to that given in Theorem 1 of Bickel and Rosenblatt (1973), see also Theorem 8.2.7 of Leadbetter et al. (1983).

So, to show that

$$\operatorname{Prob}\left\{\phi_n\left((2m)^{1/2}\Phi^{-1}\sup_{s=0,\ldots,\tilde{n}}\left(c_{1,e}(s)+c_{2,e}(s)\right)-b_n\right)\leq z\right\}\to\exp\left(-e^{-z}\right), \tag{A.12}$$

the strategy will be to show that $(2m)^{1/2}c_e(s)$ converges to a Gaussian process, say $B(u)$, in $\mathbb{D}[0,\infty)$ with a covariance structure satisfying conditions (v) and (vi) of Theorem A1 of Bickel and Rosenblatt (1973) with $\alpha=1$ or 2 there depending on $W(\lambda)$, see also Pickands (1969) equations (1.2) and (2.1). It worth noting that the proof of

$$\operatorname{Prob}\left\{\phi_n\left((2m)^{1/2}\Phi^{-1}\sup_{s=0,\ldots,\tilde{n}}-\left(c_{1,e}(s)+c_{2,e}(s)\right)-b_n\right)\leq z\right\}\to\exp\left(-e^{-z}\right)$$

follows similarly and thus is omitted.

To that end, we will show that $c_e(\lambda_p)$ converges to a Gaussian process in $\mathbb{D}[0,\infty)$ with Stone's (1963) extension of Skorohod's J_1 topology. Following Pollard (1980), it suffices the convergence in $\mathbb{D}[0,K]$ for any $K>0$. First, because under H_0, $f_p^{-1}\widetilde{f}_p-1=O\left(m^{-1}\right)$ and $\sup_p\left|f_p-f_{p+j}\right|=o\left(mn^{-1}\right)$, the distribution of ψ_n is governed by that of $\phi_n\left((2m)^{1/2}\Phi^{-1}\sup_{s=0,\ldots,\tilde{n}}\left(\widetilde{c}_{1,e}(s)+\widetilde{c}_{2,e}(s)\right)-b_n\right)$ where

$$\begin{aligned}\widetilde{c}_{1,e}(s) &= \frac{1}{k}\sum_{p=1}^{k}\left\{\left(\breve{f}_e(\lambda_p+\lambda_s)-1\right)-\left(\breve{f}_e(\lambda_k+\lambda_s)-1\right)\right\}\\ \widetilde{c}_{2,e}(s) &= \frac{1}{k}\sum_{p=1}^{k}\left\{\left(\breve{f}_e(\lambda_{-p}+\lambda_s)-1\right)-\left(\breve{f}_e(\lambda_{-k}+\lambda_s)-1\right)\right\}\end{aligned}$$

and

$$\breve{f}_e(\lambda)=\frac{2\pi}{2m+1}\sum_{j=-m,j+p\neq s}^{m}W_jI_e(\lambda_j+\lambda).$$

First, Proposition 1 and Wold's device imply that the finite dimensional distributions of $(2m)^{1/2}\widetilde{c}_e(s)$ converge to those of $B(u)$. Next, we examine the correlation structure of $(2m)^{1/2}\left(\widetilde{c}_{1,e}(s)+\widetilde{c}_{2,e}(s)\right)$. Since $k^{-1}\sum_{p=1}^{k}\left(\breve{f}_e(\lambda_p+\lambda_s)-1\right)=o_p\left(m^{-1/2}\log^{-1}n\right)$, say, we only need to examine the correlation structure of $\left((2\pi)\breve{f}_e(\lambda_k+\lambda_\ell)-1\right)+\left((2\pi)\breve{f}_e(\lambda_{-k}+\lambda_\ell)-1\right)$. Let $M(\lambda_{\ell_1}-\lambda_{\ell_2})=(2\pi)mb(\ell_1,\ell_2)$, where $b(\cdot)$ does not depend on m and without loss of generality greater than or equal to zero. From definition of $\widehat{f}_e(\lambda_k+\lambda_\ell)$ it is well known that

$$Cov\left(\breve{f}_e(\lambda_k+\lambda_{\ell_1})+\breve{f}_e(\lambda_{-k}+\lambda_{\ell_1}),\breve{f}_e(\lambda_k+\lambda_{\ell_2})+\breve{f}_e(\lambda_{-k}+\lambda_{\ell_2})\right)=o\left(m^{-1}\right)$$

if $b(\ell_1, \ell_2) \geq 1$, whereas if $0 \leq b(\ell_1, \ell_2) < 1$

$$\begin{aligned}
& (2\pi)^2 \lim_{m \to \infty} 2mCov\left(\breve{f}_e(\lambda_k + \lambda_{\ell_1}) + \breve{f}_e(\lambda_{-k} + \lambda_{\ell_1}), \breve{f}_e(\lambda_k + \lambda_{\ell_2}) + \breve{f}_e(\lambda_{-k} + \lambda_{\ell_2})\right) \\
= \quad & r(b) = \frac{\int_{-\pi}^{\pi} W(u) W(u+b) \, du}{\int_{-\pi}^{\pi} W^2(u) \, du},
\end{aligned}$$

where for notational simplicity we have denoted $b(\ell_1, \ell_2)$ by the constant b. However as $b \to 0$, by Bickel and Rosenblatt's (1973) Theorem B.1 and C3, the right side of the last displayed equation satisfies

$$\begin{aligned}
r(b) &= 1 - \frac{K}{2}|b| + o(|b|) \quad \text{if } W(1) \neq 0 \\
r(b) &= 1 - \frac{K}{2}|b|^2 + o\left(|b|^2\right) \quad \text{otherwise.}
\end{aligned}$$

So, if the process $\left((2\pi)\widehat{f}_e(\lambda_k + \lambda_\ell) - 1\right) + \left((2\pi)\widehat{f}_e(\lambda_{-k} + \lambda_\ell) - 1\right)$ were tight, applying Theorem 3.1 of Bickel and Rosenblatt's (1973), see also Stadtmüller's (1986) Theorem 8, we would conclude the proof of the theorem. So, it remains to prove the tightness condition to complete the proof.

To that end, denote

$$X_n(k) = \frac{1}{(2m+1)^{1/2}} \sum_{j=-m}^{m} W_j \left(I_{e,j+k} - E\left(I_{e,j+k}\right)\right), \quad k = 1, ..., \tilde{n}$$

and let $k/2m = 1/2m, ..., M = [n/4m]$. So, $X_n(k)$ is a process in $\mathbb{D}[0, M]$ equipped with Skorohod's metric. Extend $\mathbb{D}[0, M]$ to $\mathbb{D}[0, \infty)$ by writing $X_n(\infty) = X_n(M)$. By Pollard (1981, *Ch.V*), we need to show tightness in $\mathbb{D}[0, K]$ for any finite $K > 0$. To that end, by Brillinger's (1968) Theorem 15.6, it suffices to show the moment condition

$$E\left(|X_n(k_2) - X_n(k)|^2 |X_n(k) - X_n(k_1)|^2\right) \leq K \left|\frac{k_2 - k}{m}\right|^{(1+\delta)/2} \left|\frac{k - k_1}{m}\right|^{(1+\delta)/2}$$

for some $\delta > 0$ and where $0 \leq k_1/2m < k/2m < k_2/2m \leq K$. Because for any $0 \leq a < b < c \leq K$, $|c - b||b - a| \leq |c - a|^2$, by Schwarz's inequality, the last displayed inequality holds true if

$$E|X_n(k_2) - X_n(k_1)|^4 \leq K \left|\frac{k_2 - k_1}{m}\right|^{(1+\delta)}, \tag{A.13}$$

which is the case as we now show. First, by Brillinger's (1981) Theorem 5.2.3., if $|k_2 - k_1| < 2m$

$$\begin{aligned}
& E\left|\frac{1}{(2m+1)^{1/2}} \sum_{j=-m}^{m} W_j \left\{\left(I_{e,j+k_2} - E\left(I_{e,j+k_2}\right)\right) - \left(I_{e,j+k_1} - E\left(I_{e,j+k_1}\right)\right)\right\}\right|^4 \\
\leq \quad & K\left|\frac{k_2 - k_1}{m^2}\right| \leq K\left|\frac{k_2 - k_1}{m}\right|^2,
\end{aligned}$$

whereas if $|k_2 - k_1| \geq 2m$,

$$E\left|\frac{1}{(2m+1)^{1/2}}\sum_{j=-m}^{m} W_j \left\{\left(I_{e,j+k_2} - E\left(I_{e,j+k_2}\right)\right) - \left(I_{e,j+k_1} - E\left(I_{e,j+k_1}\right)\right)\right\}\right|^4$$
$$\leq \frac{K}{m} \leq K\left|\frac{k_2-k_1}{m}\right|^2.$$

So $(A.13)$ holds with $\delta = 1$, which completes the proof of the tightness condition of $X_n(k)$. By the same arguments it is clear that $k^{-1}\sum_{p=1}^{k}\left(\breve{f}_e(\lambda_p + \lambda_s) - 1\right)$ is also tight. Similarly, the processes $X_n(-k)$ and $k^{-1}\sum_{p=1}^{k}\left(\breve{f}_e(\lambda_{-p} + \lambda_s) - 1\right)$ are also tight. This concludes the proof of the theorem. □

Proof of Corollary 1 Following the arguments of Theorem 3.1 and 3.2, we only need to examine the behaviour of

$$\phi_n\left((2m)^{1/2}\Phi^{-1}\sup_{s=0,\ldots,\tilde{n}}\widehat{\alpha}(s) - b_n\right).$$

However, from the proof of Theorem 3.1 and since under H_a $\phi_n(2m)^{1/2}\alpha_m(s) \rightarrow \alpha$ and $\overline{\alpha}(s) = o\left(m^{-1/2}\log^{-1} n\right)$ by Lemma 4, we only need to show that

$$\text{Prob}\left\{\phi_n\left((2m)^{1/2}\Phi^{-1}\sup_{s=0,\ldots,\tilde{n}}\left(c_{1,e}(s) + c_{2,e}(s)\right) - b_n\right) \leq z - \alpha\right\} \rightarrow \exp\left(-e^{-z}\right). \tag{A.14}$$

But using the same arguments as in Theorem 3.1, $(A.14)$ holds true if

(a) The finite dimensional distributions of $(2m)^{1/2}\left(c_{1,e}(s) + c_{2,e}(s)\right)$ converges to those of a Gaussian process with correlation structure $r(b)$.

(b) Tightness condition of $(2m)^{1/2}\left(c_{1,e}(s) + c_{2,e}(s)\right)$.

However, assuming for simplicity that $f_p = \lambda_p^{-\alpha_m(s)}$, because for all p $(2\pi) f_p - 1 = \exp\left(-m^{-1/2}\log(n/p)\right) - 1 = O\left(m^{-1/2}\log^{-1} n\right)$, we have that the distribution of $(A.14)$ is governed by that of $\phi_n\left((2m)^{1/2}\Phi^{-1}\sup_{s=0,\ldots,\tilde{n}}\left(\widetilde{c}_{1,e}(s) + \widetilde{c}_{2,e}(s)\right) - b_n\right)$. From here, the proof of (a) and (b) follows proceeding as with the proof of Theorem 3.1. □

Proof of Corollary 2 The proof follows after observing that Proposition 1 implies that $\widehat{\alpha}(s) \xrightarrow{p} \alpha(s)$. So,

$$(2m)^{1/2}\widehat{\alpha}(s) = (2m)^{1/2}\left(\alpha(s) + o_p(1)\right).$$

Now denote by λ_{s^*} the closest Fourier frequency to λ^0. Then, the right side of the last displayed equation implies that

$$\sup_{s=0,\ldots,\tilde{n}}\left|(2m)^{1/2}\widehat{\alpha}(s)\right| - b_n \geq (2m)^{1/2}\left|\alpha(s^*)\right| - b_n \geq K(2m)^{1/2}\left|\alpha(s^*)\right| \rightarrow \infty,$$

which completes the proof. □

Proof of Theorem 4.1 Denote $\widetilde{f}_p(s) = (2m+1)^{-1}\sum_{j=-m}^{m} W_j$. Since $k^{-1}\sum_{p=1}^{k}\log\widetilde{f}_p(s) - \log\widetilde{f}_k(s) = 0$, it implies that

$$\begin{aligned}\widehat{\alpha}^*(s) &= \frac{1}{2k}\sum_{p=1}^{k}\left(\log\widehat{f}_p^*(s)/\widetilde{f}_p(s)\right) - \left(\log\widehat{f}_k^*(s)/\widetilde{f}_k(s)\right)\\ &\quad + \frac{1}{2k}\sum_{p=1}^{k}\left(\log\widehat{f}_{-p}^*(s)/\widetilde{f}_{-p}(s)\right) - \left(\log\widehat{f}_{-k}^*(s)/\widetilde{f}_{-k}(s)\right).\end{aligned}$$

We first show that the finite dimensional distributions of $\widehat{\alpha}^*(s)$ converge to those of a Gaussian process. To that end, it suffices to show that for any finite set of constants $\psi_1, ..., \psi_q$

$$(2m)^{1/2}\sum_{\ell=1}^{q}\psi_\ell\widehat{\alpha}^*(s_\ell) \xrightarrow{d^*} N\left(0, \Phi^{-1}\sum_{\ell=1}^{q}\psi_\ell^2\right).$$

For that purpose, and denoting $g_p^*(s) = \widetilde{f}_p^{-1}(s)\widehat{f}_p^*(s) - 1$, it suffices to show that

$$(2m)^{1/2}\sum_{\ell=1}^{q}\psi_\ell\left(b_k^*(s_\ell) + b_{-k}^*(s_\ell)\right) \xrightarrow{d^*} N\left(0, \Phi^{-1}\sum_{\ell=1}^{q}\psi_\ell^2\right) \tag{A.15}$$

where $b_k^*(s_\ell) = k^{-1}\sum_{p=1}^{k} g_p^*(s_\ell) - g_k^*(s_\ell)$ and $b_{-k}^*(s_\ell) = k^{-1}\sum_{p=1}^{k} g_{-p}^*(s_\ell) - g_{-k}^*(s_\ell)$, and for all $\ell = 1, ..., q$,

$$(2m)^{1/2}\left(\widehat{\alpha}^*(s_\ell) - \left(b_k^*(s_\ell) + b_{-k}^*(s_\ell)\right)\right) = o_{p^*}(1). \tag{A.16}$$

We show $(A.15)$ first. Write $c_k^*(s_\ell) = \left((2m)^{1/2}/k\right)\sum_{p=1}^{k} g_p^*(s_\ell)$. Because $E^* I_j^* = (2\pi)^{-1}$, $E\left(c_k^*(s_\ell)\right) = 0$. Next,

$$\begin{aligned}E^*\left(c_k^*(s_\ell)\right)^2 &= \frac{2m}{k^2}\left(\sum_{p=1}^{k} E^* g_p^{*2}(s_\ell) + \sum_{p_1,p_2=1;p_1\neq p_2}^{k} E^*\left(g_{p_1}^*(s_\ell)\, g_{p_2}^*(s_\ell)\right)\right)\\ &= \frac{2m}{k^2}\left(\sum_{p=1}^{k}\left(\frac{\widehat{\mu}_4}{n} + \frac{2\widehat{\sigma}^4}{m}\right) + \sum_{p_1,p_2=1;p_1\neq p_2}^{k}\frac{\widehat{\mu}_4}{n} + \sum_{|p_1-p_2|<m;p_1\neq p_2}^{k}\frac{2\widehat{\sigma}^4}{m}\right)\end{aligned}$$

using Proposition 10.3.2. of Brockwell and Davies (1991), where $\widehat{\mu}_4$ and $\widehat{\sigma}^2$ are the sample estimators of the fourth order cumulant and variance of x_t. But the right side of the last displayed equation is

$$\frac{m}{n}\widehat{\mu}_4 + \frac{2}{k}\widehat{\sigma}^4 = o_{p^*}(1),$$

which implies that $c_k^*(s_\ell) = o_{p^*}(1)$. Likewise $c_{-k}^*(s_\ell) = \left((2m)^{1/2}/k\right)\sum_{p=1}^{k} g_{-p}^*(s_\ell) = o_{p^*}(1)$. Thus, we conclude that

$$(2m)^{1/2}\left(b_k^*(s_\ell) + b_{-k}^*(s_\ell)\right) = -(2m)^{1/2}\left(g_k^*(s_\ell) + g_{-k}^*(s_\ell)\right) + o_{p^*}(1)$$

so that $(A.15)$ is shown if $(2m)^{1/2}\sum_{\ell=1}^{q}\psi_\ell\left(g_k^*(s_\ell)+g_{-k}^*(s_\ell)\right)\xrightarrow{d^*}N\left(0,\Phi^{-1}\sum_{\ell=1}^{q}\psi_\ell^2\right)$. Proceeding as in the proof of $E^*\left(c_k^*(s_\ell)\right)^2\xrightarrow{p}0$, it is easily observed that $2mE^*\left(\sum_{\ell=1}^{q}\psi_\ell\left(g_k^*(s_\ell)+g_{-k}^*(s_\ell)\right)\right)^2\xrightarrow{p}\Phi^{-1}\sum_{\ell=1}^{q}\psi_\ell^2$. So, to complete the proof of $(A.15)$, we need to show the Lindeberg's condition, that is

$$E^*\left(\frac{1}{2m}\sum_{j=-m}^{m}I_{k+s_\ell}^{*2}\mathcal{I}\left((2m)^{-1}I_{k+s_\ell}^{*2}>\delta\right)\right)\xrightarrow{P}0.$$

A sufficient condition is that $E^*m^{-2}\sum_{j=-m}^{m}I_{k+s_\ell}^{*4}=o_p(1)$, which is easily shown to hold true.

Therefore, to complete the proof of the convergence of the finite dimensional distributions of $\widehat{\alpha}^*(s)$, we need to show $(A.16)$. First, for any arbitrary $\varepsilon,\eta>0$,

$$\begin{aligned}
&\Pr\left\{\left|(2m)^{1/2}\widehat{\alpha}^*(s)-\left(b_k^*(s_\ell)+b_{-k}^*(s_\ell)\right)\right|>\varepsilon\,|\chi\right\}\\
=\;&\Pr\left\{\left|(2m)^{1/2}\widehat{\alpha}^*(s)-\left(b_k^*(s_\ell)+b_{-k}^*(s_\ell)\right)\right|>\varepsilon;\ \sup_p\left|g_p^*(s)\right|\geq\eta\,|\chi\right\}\\
&+\Pr\left\{\left|(2m)^{1/2}\widehat{\alpha}^*(s)-\left(b_k^*(s_\ell)+b_{-k}^*(s_\ell)\right)\right|>\varepsilon;\ \sup_p\left|g_p^*(s)\right|<\eta\,|\chi\right\}.
\end{aligned}$$

The first term on the right converges to zero since $\sup_p\left|g_p^*(s)\right|=o_{p^*}(1)$ by an obvious extension of Hidalgo and Robinson's (2001) Proposition A3, with $r=2$ there. Since $|\log x-(x-1)|\leq 2^{-1}(x-1)^2$ for x close to 1, by Markov inequality, the first term on the right also converges to zero if

$$\frac{(2m)^{1/2}}{k}\sum_{p=1}^{k}E^*\left(g_p^{*2}(s)+g_k^{*2}(s)+g_{-p}^{*2}(s)+g_{-k}^{*2}(s)\right)\xrightarrow{P}0,$$

which is the case by Hidalgo and Robinson's (2001) Propositions A2 and A1 respectively. This concludes the proof of the finite dimensional distributions of $\widehat{\alpha}^*(s)$.

Next we show that

$$\sup_{s=0,\ldots,\tilde{n}}\left|(2m)^{1/2}h_j^*(s)\right|=o_{p^*}(1),\quad j=1,2 \tag{A.17}$$

where $h^*(s)=\widehat{\alpha}^*(s)-c_1^*(s)-c_2^*(s)$ with

$$\begin{aligned}
c_1^*(s)&=\frac{1}{k}\sum_{p=1}^{k}\left\{\left(\frac{\widehat{f}_p^*(s)-\widetilde{f}_p(s)}{\widetilde{f}_p(s)}\right)-\left(\frac{\widehat{f}_k^*(s)-\widetilde{f}_k(s)}{\widetilde{f}_k(s)}\right)\right\}\\
c_2^*(s)&=\frac{1}{k}\sum_{p=1}^{k}\left\{\left(\frac{\widehat{f}_{-p}^*(s)-\widetilde{f}_{-p}(s)}{\widetilde{f}_{-p}(s)}\right)-\left(\frac{\widehat{f}_{-k}^*(s)-\widetilde{f}_{-k}(s)}{\widetilde{f}_{-k}(s)}\right)\right\}.
\end{aligned}$$

Indeed, first for any arbitrary $\varepsilon,\eta>0$,

$$\begin{aligned}
\Pr\left\{\sup_{q=0,\ldots,\tilde{n}}\left|(2m)^{1/2}h^*(s)\right|>\varepsilon\,|\chi\right\}&=\Pr\left\{\sup_{s=0,\ldots,\tilde{n}}\left|(2m)^{1/2}h^*(s)\right|>\varepsilon;\ \sup_{q=0,\ldots,\tilde{n}}\left|\xi_q^*\right|\geq\eta\,|\chi\right\}\\
&\quad+\Pr\left\{\sup_{s=0,\ldots,\tilde{n}}\left|(2m)^{1/2}h^*(s)\right|>\varepsilon;\ \sup_{q=0,\ldots,\tilde{n}}\left|\xi_q^*\right|<\eta\,|\chi\right\}.
\end{aligned}$$

where $\xi_q^* = \left(\widetilde{f}_q^{-1}\widehat{f}_q^* - 1\right)$. The first term on the right of the last displayed equation converges to zero since Propositions A.1 and A.3 of Hidalgo and Robinson (2001) imply that $\sup_q \left|\xi_q^*\right| = o_{p^*}(1)$. So, using the definition of $\widehat{\alpha}^*(s)$ and that $|\log x - (x-1)| \leq 2^{-1}(x-1)^2$ for x close to 1, we have that

$$\begin{aligned}
(2m)^{1/2} \sup_{s=0,\ldots,\tilde{n}} h^*(s) &\leq K \sup_{s=0,\ldots,\tilde{n}} \frac{(2m)^{1/2}}{k} \sum_{q=1}^{k} \left\{g_q^{*2}(s) + g_k^{*2}(s) + g_{-q}^{*2}(s) + g_{-k}^{*2}(s)\right\} \\
&\leq K (2m)^{1/2} \sup_{p=0,\ldots,\tilde{n}} \xi_p^{*2}.
\end{aligned}$$

Therefore, the proof of $(A.17)$ is completed if

$$(2m)^{1/2} \sup_{p=0,\ldots,\tilde{n}} \xi_p^{*2} = o_{p^*}\left(\log^{-1} n\right),$$

which is the case since following the same steps of Lemma 3 in Appendix B, it implies that it is $o_{p^*}\left(m^{-1/4}\right) = o_{p^*}\left(\log^{-1} n\right)$.

Thus, to complete the proof of the theorem it suffices to examine the limit distribution of $\psi_n^* = \phi_n\left(\sup_{s=0,\ldots,\tilde{n}} (2m)^{1/2}\Phi^{-1}(c_1^*(s) + c_2^*(s)) - b_n\right)$. To that end, we need to show that

$$\text{Prob}\left\{\psi_n^* \leq z\,|\chi\right\} \xrightarrow{P} \exp\left(-e^{-z}\right).$$

However, since $\widetilde{f}_p(s)$ is constant and $\widetilde{f}_p(s) - \int_{-1}^{1} W(u)\,du = O\left(m^{-1}\right)$, it suffices to show that

$$\text{Prob}\left\{\phi_n\left((2m)^{1/2}\Phi^{-1}\sup_{s=0,\ldots,\tilde{n}}\left(\widetilde{c}_1^*(s) + \widetilde{c}_2^*(s)\right) - b_n\right) \leq z\,|\chi\right\} \xrightarrow{P} \exp\left(-e^{-z}\right), \tag{A.18}$$

where

$$\begin{aligned}
\widetilde{c}_1^*(s) &= \frac{1}{k}\sum_{p=1}^{k}\left\{\left(\widehat{f}_p^*(s) - (2\pi)^{-1}\right) - \left(\widehat{f}_k^*(s) - (2\pi)^{-1}\right)\right\} \\
\widetilde{c}_2^*(s) &= \frac{1}{k}\sum_{p=1}^{k}\left\{\left(\widehat{f}_{-p}^*(s) - (2\pi)^{-1}\right) - \left(\widehat{f}_{-k}^*(s) - (2\pi)^{-1}\right)\right\}.
\end{aligned}$$

The strategy to show $(A.18)$ is similar to that of $(A.12)$. In particular, we will show that $\widetilde{c}^*(s)$ converges, in bootstrap, to a Gaussian process in $\mathbb{D}[0,\infty)$ with Stone's (1963) extension of Skorohod's J_1 topology. Following Pollard (1980), it suffices the convergence in $\mathbb{D}[0,K]$ for any $K > 0$. First, Proposition 1 and Wold's device imply that the finite dimensional distributions of $\widetilde{c}^*(s)$ converge in bootstrap to those of $B(u)$. Next, we examine the correlation structure of $(2m)^{1/2}\widetilde{c}^*(s)$. Since $E^*\left(k^{-1}\sum_{p=1}^{k}\left(\widehat{f}_p^*(s) - 1\right)\right)^2 = o_p\left(m^{-1}\log^{-2} n\right)$, we only need to examine the correlation structure of $\left((2\pi)\widehat{f}_k^*(s) - 1\right) + \left((2\pi)\widehat{f}_{-k}^*(s) - 1\right)$. Let $M(\lambda_{\ell_1} - \lambda_{\ell_2}) = (2\pi)\,mb(\ell_1,\ell_2)$, where $b(\cdot)$ does not depend on m and without loss of generality greater than or equal to zero. From definition of $\widehat{f}_{k+\ell}^*(s)$ it is well known that

$$Cov^*\left(\widehat{f}_{k+\ell_1}^*(s), \widehat{f}_{k+\ell_2}^*(s)\right) = n^{-1}\widehat{\mu}_4 + m^{-1}2\widehat{\sigma}^4 = o_p\left(m^{-1}\right)$$

if $b(\ell_1,\ell_2) \geq 1$, whereas if $0 \leq b(\ell_1,\ell_2) < 1$

$$(2\pi)^2\, 2mCov^*\left(\widehat{f}^*_{k+\ell_1}(s) + \widehat{f}^*_{-k+\ell_1}(s), \widehat{f}^*_{k+\ell_2}(s) + \widehat{f}^*_{-k+\ell_2}(s)\right) \xrightarrow{P} r(b) = \frac{\int_{-\pi}^{\pi} W(u)\, W(u+b)\, du}{\int_{-\pi}^{\pi} W^2(u)\, du}.$$

Now proceeds as in the proof of Theorem 3.1. So, it remains to prove the tightness condition to finish the proof.

Denote

$$X_n^*(k) = \frac{1}{(2m+1)^{1/2}} \sum_{j=-m}^{m} W_j \left(I^*_{j+k} - (2\pi)^{-1}\right), \quad k = 1, ..., \tilde{n}$$

and let $k/2m = 1/2m, ..., M = [n/4m]$. So, $X_n^*(k)$ is a process in $\mathbb{D}[0, M]$ equipped with Skorohod's metric. Extend $\mathbb{D}[0, M]$ to $\mathbb{D}[0, \infty)$ by writing $X_n(\infty) = X_n(M)$. By Pollard (1981, $Ch.V$), we need to show tightness in $\mathbb{D}[0, K]$ for any finite $K > 0$. To that end, by Brillinger's (1968) Theorem 15.6, it suffices to show the moment condition

$$E^*\left(|X_n^*(k_2) - X_n^*(k)|^2 |X_n(k) - X_n(k_1)|^2\right) \leq KH_n(k_2,k_1)\left|\frac{k_2-k}{m}\right|^{(1+\delta)/2}\left|\frac{k-k_1}{m}\right|^{(1+\delta)/2}$$

for some $\delta > 0$ with $H_n(k_2,k_1)$ being bounded in probability, and where $0 \leq k_1/2m < k/2m < k_2/2m \leq K$. Because for any $0 \leq a < b < c \leq K$, $|c-b|\,|b-a| \leq |c-a|^2$, by Schwarz's inequality, the last displayed inequality holds true if

$$E^*|X_n^*(k_2) - X_n^*(k_1)|^4 \leq KH_n(k_2,k_1)\left|\frac{k_2-k_1}{m}\right|^{(1+\delta)}, \tag{A.19}$$

which is the case as we now show. First, by Brillinger's (1981) Theorem 5.2.3., if $|k_2 - k_1| < 2m$

$$\begin{aligned} &E^*\left|\frac{1}{(2m+1)^{1/2}} \sum_{j=-m}^{m} W_j \left\{\left(I^*_{j+k_2} - (2\pi)^{-1}\right) - \left(I^*_{j+k_1} - (2\pi)^{-1}\right)\right\}\right|^4 \\ \leq\ & K\left|\frac{k_2-k_1}{m^2}\right| \leq K\left|\frac{k_2-k_1}{m}\right|^2, \end{aligned}$$

whereas if $|k_2 - k_1| \geq 2m$,

$$E^*\left|\frac{1}{(2m+1)^{1/2}} \sum_{j=-m}^{m} W_j \left\{\left(I^*_{j+k_2} - (2\pi)^{-1}\right) - \left(I^*_{j+k_1} - (2\pi)^{-1}\right)\right\}\right|^4 \leq \frac{K}{m} \leq K\left|\frac{k_2-k_1}{m}\right|^2.$$

So $(A.19)$ holds with $\delta = 1$ and $H_n(k_2,k_1) = 1$. By the same arguments it is clear that $(2m)^{1/2} k^{-1} \sum_{p=1}^{k} \left(\breve{f}^*_p(s) - (2\pi)^{-1}\right)$ is also tight. Similarly, the processes $X_n(-k)$ and $k^{-1}\sum_{p=1}^{k}\left(\breve{f}^*_{-p}(s) - 1\right)$ are also tight. This concludes the proof of the theorem. □

APPENDIX B

Lemma 4.3. *Let $p = p(n)$ be such that $p^{-1} + p^{-1}m + n^{-1}p \to 0$. Assuming C1-C5, as $n \to \infty$,*

$$(2m)^{1/2} \left(g_p(s) + g_{-p}(s)\right) \xrightarrow{d} N(0, 1),$$

where $\xrightarrow{d}$ means convergence in distribution.

Proof. The proof follows from Theorem 1 of Hidalgo and Yajima (2000). □

Lemma 4.4. *Assuming C1-C5, as $n \to \infty$*

$$\sup_{p=1,\dots,\tilde{n}} \left|\psi_{3p}\right| = o_p(1).$$

Proof. Denote $r = qm^{\beta}$, for $q = 1, ..., m^{-\beta}\tilde{n}$, where $0 < \beta < 1$. Then

$$\sup_{p=1,\dots,\tilde{n}} \left|\psi_{3p}\right| = \sup_q \sup_p |a_{3p}|$$

where $\sup_q$ and $\sup_p$ denote $\sup_{q=1,\dots,m^{-\beta}\tilde{n}}$ and $\sup_{p=1+r-m^{\beta},\dots,r}$. The triangle inequality implies that the left side of the last displayed equation is bounded by

$$2\sup_q \left|\frac{1}{2m+1}\sum_{j=-m}^{m} W_j f_{j+r}\left((2\pi)\frac{I_{e,j+r}}{\sigma_e^2} - 1\right)\right| \tag{B.1}$$
$$+2\sup_q \sup_p \left|\frac{1}{2m+1}\sum_{j=-m}^{m} W_j \left(f_{j+r}\left((2\pi)\frac{I_{e,j+r}}{\sigma_e^2} - 1\right) - f_{j+p}\left((2\pi)\frac{I_{e,j+p}}{\sigma_e^2} - 1\right)\right)\right|.$$

Because $\sup_p |c_p| \le \left(\sum_p |c_p|^4\right)^{1/4}$, Brillinger's (1981) Theorem 7.2.2 implies that the expectation of the first term of $(B.1)$ is bounded by

$$\begin{aligned} K\left(\sum_{q=1}^{m^{-\beta}\tilde{n}} E\left|\frac{1}{2m+1}\sum_{j=-m}^{m} W_j f_{j+r}\left((2\pi)\frac{I_{e,j+r}}{\sigma_e^2} - 1\right)\right|^4\right)^{1/4} &= O\left(m^{-1/2-\beta/4}n^{1/4}\right) \\ &= o(1), \end{aligned}$$

by C4. Then Markov's inequality implies that the first term of $(B.1)$ is $o_p(1)$. The proof is thus completed if the second term of $(B.1)$ is also $o_p(1)$. By triangle inequality, the second term of $(B.1)$ is bounded by

$$2\sup_q \sup_p \left|\frac{1}{2m+1}\sum_{j=-m}^{m} W_j \left(f_{j+r} - f_{j+p}\right)\left((2\pi)\frac{I_{e,j+r}}{\sigma_e^2} - 1\right)\right| \tag{B.2}$$
$$+2\sup_q \sup_p \left|\frac{1}{2m+1}\sum_{j=-m}^{m} W_j f_{j+p}\left(\left((2\pi)\frac{I_{e,j+r}}{\sigma_e^2} - 1\right) - \left((2\pi)\frac{I_{e,j+p}}{\sigma_e^2} - 1\right)\right)\right|.$$

Since $\widetilde{\sum_j} W_j \left((2\pi) \frac{I_{e,j+r}}{\sigma_e^2} - 1 \right) = O_p \left(m^{1/2} \right)$ and the continuous differentiability of f implies that $\left| f_{j+r} - f_{j+p} \right| \leq K (r-p)/n$ it is straightforward to show that the first term of $(B.2)$ is $o_p(1)$. Finally the second term of $(B.2)$ is, except constants,

$$\sup_q \sup_p \left| \frac{1}{2m+1} \sum_{j=-m}^{m} W_j f_{j+p} \left(I_{e,j+r} - I_{e,j+p} \right) \right|.$$

The fourth power of the last displayed expression is bounded by

$$\sup_q \sup_p \left| \frac{1}{2m+1} \sum_{j=p+m}^{q\left[m^\beta\right]+m} W_j f_{j+p} \left(I_{e,j+p} - EI_{e,j+p} \right) \right|^4$$
$$+ \sup_q \sup_p \left| \frac{1}{2m+1} \sum_{j=p-m}^{q\left[m^\beta\right]-m} W_j f_{j+p} \left(I_{e,j+p} - EI_{e,j+p} \right) \right|^4$$

whose expectation is bounded by

$$K \frac{1}{m^4} \sum_{q=1}^{m^{-\beta}[n/2]} \sum_{p=1+q[m^\beta]-m}^{q\left[m^\beta\right]} \left(q \left[m^\beta \right] - p \right)^2 = O \left(\frac{n}{m^{4-2\beta}} \right) = o(1)$$

by C4 and that from the proof of Brillinger's (1981) Theorem 7.7.4., $E \left| \sum_{j=\ell_1+1}^{\ell_2} W_j f_{j+p} \left(I_{e,j+p} - EI_{e,j+p} \right) \right|^4 < K \left(\ell_2 - \ell_1 \right)^2$. This completes the proof of the Lemma. □

Denote the Fourier Discrete Transform of a sequence of random variables a_t by

$$w_a(\lambda_j) = \frac{1}{2\pi n} \sum_{t=1}^{n} a_t e^{it\lambda_j},$$

and $v_a(\lambda_j) = w_a(\lambda_j) / f_a^{1/2}(\lambda_j)$, where $f_a(\lambda)$ is the spectral density function of a_t.

Lemma 4.5. *Assuming C1-C3, as $n \to \infty$*

$$\begin{aligned}
(a1)\ \ E\left(v_a(\lambda_j)\overline{v}_b(\lambda_j)\right) - R_{ab}(\lambda_j) &= O\left(n^{-1}\log n\right), \\
(a2)\ \ E\left(v_a(\lambda_j)v_b(\lambda_j)\right) &= O\left(n^{-1}\log n\right), \\
(a3)\ \ E\left(v_a(\lambda_j)\overline{v}_b(\lambda_k)\right) &= O\left(n^{-1}\log n\right), \\
(a4)\ \ E\left(v_a(\lambda_j)v_b(\lambda_k)\right) &= O\left(n^{-1}\log n\right)
\end{aligned}$$

under the null hypothesis, whereas under local alternatives, that is $\alpha(m) = K/m$,

$$\begin{aligned}
(b1)\ \ E\left(v_a(\lambda_j)\overline{v}_b(\lambda_j)\right) - R_{ab}(\lambda_j) &= O\left(\min\left(\left(m\log^2 M\right)^{-1/2}, j^{-1}\log n\right)\right) \\
(b2)\ \ E\left(v_a(\lambda_j)v_b(\lambda_j)\right) &= O\left(\min\left(\left(m\log^2 M\right)^{-1/2}, j^{-1}\log n\right)\right) \\
(b3)\ \ E\left(v_a(\lambda_j)\overline{v}_b(\lambda_k)\right) &= O\left(\min\left(\left(m\log^2 M\right)^{-1/2}, j^{-1}\log n\right)\right) \\
(b4)\ \ E\left(v_a(\lambda_j)v_b(\lambda_k)\right) &= O\left(\min\left(\left(m\log^2 M\right)^{-1/2}, j^{-1}\log n\right)\right)
\end{aligned}$$

where a_t and b_t are x_t and/or ε_t.

Proof. The proof of $(a1)-(a4)$ follows from Theorem 4.3.2. of Brillinger (1981). Next, $(b1)$. For notational simplicity, we will consider the case where $a_t = b_t = x_t$, the case when a_t and/or b_t is ε_t follows similarly if not easier. From definition of $v_x(\lambda_j)$ and the triangle inequality, $|E(v_x(\lambda_j)\overline{v}_x(\lambda_j))-1|$ is bounded by

$$\left| f_1(\lambda_j)\int_{-\pi}^{\pi}(f_2(\lambda)-f_2(\lambda_j))\,G(\lambda-\lambda_j)\,d\lambda \right| \tag{B.3}$$

$$+\left| \int_{-\pi}^{\pi} f_2(\lambda)(f_1(\lambda)-f_1(\lambda_j))\,G(\lambda-\lambda_j)\,d\lambda \right|, \tag{B.4}$$

because under local alternatives, for all n $K^{-1} < f(\lambda_j) < K$ for $j=1,...,\tilde{n}$, where $G(u)$ is the Fejer's Kernel and where we have written $f(\lambda) = f_1(\lambda) f_2(\lambda)$ such that $f_2(\lambda)$ is continuously differentiable and $f_1(\lambda)$ is continuously differentiable outside any open set outside the origin and $f_1(\lambda) = \lambda^{-\alpha/\left(m\log^2 M\right)^{1/2}}$ as $\lambda \to 0+$.

By Theorem 4.3.2. of Brillinger (1981), $(B.3)$ is $O\left(n^{-1}\log n\right)$. So, it remains to examine $(B.4)$. C1 implies that we can pick δ so small that the component of $(B.4)$ due to integration over $(-\pi,-\delta)\cup(\delta,\pi)$ is $O\left(n^{-1}\log n\right)$ proceeding as in Robinson (1995a, page 1061). Next, the contribution of the integral $\int_{-\delta}^{-\lambda_j/2}$ to $(B.4)$ is bounded by

$$\begin{aligned}
&\left|\int_{-\delta}^{-\lambda_j/2}(f_1(\lambda)-f_1(\lambda_j))\,f_2(\lambda)\,G(\lambda-\lambda_j)\,d\lambda\right| \\
\le\ &\left|\int_{-\delta}^{-\lambda_j/2}(f_1(\lambda)-1)\,f_2(\lambda)\,G(\lambda-\lambda_j)\,d\lambda\right| \\
&+\left|\int_{-\delta}^{-\lambda_j/2}(f_1(\lambda_j)-1)\,f_2(\lambda)\,G(\lambda-\lambda_j)\,d\lambda\right| \\
=\ &O\left(m^{-1}j^{-1}\right),
\end{aligned}$$

because for $\lambda>0$, $\left|\lambda^{-\alpha/\left(m\log^2 M\right)^{1/2}}-1\right| = O\left(\left(m\log^2 M\right)^{-1/2}\right)$ and $\int_{-\delta}^{-\lambda_j/2}G(\lambda-\lambda_j)\,d\lambda = O\left(j^{-1}\right)$. By an identical argument, $\left|\int_{2\lambda_j}^{\delta}\right| = O\left(\left(m\log^2 M\right)^{-1/2}j^{-1}\right)$ and the contribution of $\left|\int_{\lambda_j/2}^{2\lambda_j}\right|$ is $O\left(\left(m\log^2 M\right)^{-1/2}\right)$. To complete the proof that $(B.4)$ is $O\left(\min(\left(m\log^2 M\right)^{-1/2}, j^{-1}\log n)\right)$, we have to examine the contribution of

$$\begin{aligned}
&\left|\int_{-\lambda_j/2}^{\lambda_j/2} f_2(\lambda)(f_1(\lambda)-f_1(\lambda_j))\,G(\lambda-\lambda_j)\,d\lambda\right| \\
\le\ &\left|\int_{(-\lambda_j/2,\lambda_j/2)/(-n^{-2},n^{-2})} f_2(\lambda)(f_1(\lambda)-f_1(\lambda_j))\,G(\lambda-\lambda_j)\,d\lambda\right| \\
&+\left|\int_{-n^{-2}}^{n^{-2}} f_2(\lambda)(f_1(\lambda)-f_1(\lambda_j))\,G(\lambda-\lambda_j)\,d\lambda\right|.
\end{aligned}$$

The first term on the right of the last displayed inequality is $O\left(\left(m\log^2 M\right)^{-1/2}\right)$ proceeding as before, whereas the second term is bounded by

$$K\max_{|\lambda|\le n^{-2}} G\left(\lambda-\lambda_j\right)\int_{-n^{-2}}^{n^{-2}}\left(f_1\left(\lambda\right)+f_1\left(\lambda_j\right)\right)d\lambda = O\left(n^{-1}\right).$$

This concludes the proof of $(b1)$. The proof of $(b2)-(b3)$ proceeds as in the proof of Robinson's (1995a) Theorem 2 and the arguments given in the proof of $(b1)$. $\square$

Lemma 4.6. *Under the local alternative H_a, $\overline{\alpha}\left(s\right)=o\left(m^{-1/2}\log^{-1}n\right)$.*

Proof. Taking $W_j=1$ for notional simplicity, we have that

$$\begin{aligned}\frac{1}{2m+1}\sum_{j=-m}^{m}\left(\lambda_p^{\alpha_m(s)}\lambda_{|p+j|_+}^{-\alpha_m(s)}-1\right) &= \frac{1}{2m+1}\sum_{j=-m}^{m}\left(\left|p/\left(p+j\right)\right|_+^{\alpha_m(s)}-1\right)\\ &= \frac{1}{2m+1}\sum_{j=-m}^{m}\left(\exp\left(\alpha_m\left(s\right)\log\left|p/\left(p+j\right)\right|_+\right)-1\right)\\ &= \alpha_m\left(s\right)\log\left(\left|p/\left(p+j\right)\right|_+\right)+o\left(\alpha_m\left(s\right)^2\log n\right)\end{aligned}$$

by Taylor expansion of e^x around $x=0$. So, from the definition of $\overline{\alpha}\left(s\right)$ and Taylor expansion of $\log z$ around $z=1$, we obtain that

$$\begin{aligned}\overline{\alpha}\left(s\right) &= \alpha_m\left(s\right)\left(\frac{1}{k}\sum_{p=1}^{k}\left(\frac{1}{2m+1}\sum_{j=-m}^{m}\log\left|\frac{p}{p+j}\right|_+\right)-\frac{1}{2m+1}\sum_{j=-m}^{m}\log\left|\frac{k}{k+j}\right|_+\right)\\ &\quad +o\left(\alpha_m\left(s\right)^2\log n\right)\\ &= o\left(\alpha_m\left(s\right)\right)\end{aligned}$$

since $\left(2m+1\right)^{-1}\sum_{j=-m}^{m}\log\left|p/\left(p+j\right)\right|_+=O\left(1\right)$ by integrability of $\log\left(z/\left(1+z\right)\right)$ and that $mk^{-1}=o\left(1\right)$ by C5. Now use $\alpha_m\left(s\right)=O\left(m^{-1/2}\log^{-1}n\right)$ to conclude. $\square$

References

[1] Adenstedt, R.K. (1974): "On Large-Sample Estimation of the Mean of a Stationary Random Sequence," *Annals of Statistics*, 2, 1095-1107.

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

[3] Bickel, P.J. and Rosenblatt, M. (1973): "On some global measures of the deviations of density function estimates," *Annals of Statistics,* 1, 1071-1095.

[4] Bickel, P.J. and Rosenblatt, M. (1975): "On some global measures of the deviations of density function estimates - corrections," *Annals of Statistics,* 3, 1370.

[5] Billingsley, P. (1968): *Convergence of probability measures.* Wiley, New York.

[6] Bowman, A.W., Jones, M.C. and Gijbels, I. (1998): "Testing monotonicity of regressions," *Journal of Computation Statistics*, 7, 489-500.

[7] Brillinger, D.R. (1981): *Time Series: Data analysis and theory.* McGraw-Hill.

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

[9] Dümbgen, L. and Spokoiny, V.G. (2001): "Multiscale testing of qualitative hypotheses," *Annals of Statistics,* 29, 124-152.

[10] Dehling, H. and Taqqu, M.S. (1989): "The empirical process of some long-range dependent sequences with an application to U-statistics," *Annals of Statistics,* 17, 1767-1783.

[11] Efron, B. (1979): "Bootstrap methods: Another look at the Jackknife," *Annals of Statistics,* 7, 1-26.

[12] Franke, J. and Härdle, W. (1992): "On bootstrap kernel spectral estimates," *Annals of Statistics,* 20, 121-145.

[13] Gil-Alaña, L.A. and Robinson P.M. (1996): "Testing for unit roots and other nonstationary hypothesis in macroeconomic time series," *Journal of Econometrics*, 80, 241-268.

[14] Giné, E. and Zinn, J. (1989): "Necessary conditions for the bootstrap of the mean," *Annals of Statistics*, 12, 1130-1135.

[15] Giraitis, L., Hidalgo, J. and Robinson P.M. (2001): "Gaussian estimation of parametric spectral density with unknown pole," Forthcoming in *Annals of Statistics*.

[16] Giraitis, L., and Hidalgo, J. (2001): "Spectral Estimation of Nonlinear Transformation of a Long Range Process". Preprint.

[17] Giraitis, L., and Hidalgo, J. (2001): "A Parametric Test for Weak Dependence and its Bootstrap Approach". Preprint.

[18] Goshal, S. Sen, A. and Van Der Vaart, A.W. (2000): "Testing monotonicity of regression," *Annals of Statistics*, 28, 1054-1082.

[19] 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.

[20] Gray, H.L., Zhang, N-F. and Woodward, W.A. (1989): "On generalized fractional processes," *Journal of Time Series Analysis*, 10, 233-257.

[21] Gray, H.L., Zhang, N-F. and Woodward, W.A. (1994): "On generalized fractional processes - a correction," *Journal of Time Series Analysis*, 15, 561-562.

[22] Hall, P. (1979): "On the rate of convergence of normal extremes," *Journal of Applied Probability*, 16, 433-439.

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

[24] Hall, P. and Heckman, N.E. (2000): "Testing for monotonicity of a regression mean by calibrating for linear functions," *Annals of Statistics,* 28, 20-39.

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

[26] Hidalgo, J. (2001): "Semiparametric estimation of the location of the pole," Preprint.

[27] Hidalgo, J. and Robinson, P.M. (2001): "Adapting to unknown disturbance autocorrelation in regression with long memory". Forthcoming in *Econometrica*.

[28] Hidalgo, J. and Yajima, Y. (2000): "Semiparametric estimation of the long-range parameter," Preprint.

[29] Ho, H.C. and Hsing, T. (1996): "On the asymptotic expansion of the empirical process of a long memory linear sequence," *Annals of Statistics*, 24, 992-1024.

[30] Konakov, V.D. and Piterbarg, V.I. (1984): "On the convergence rate of maximal deviation distribution for kernel regression estimates," *Journal of Multivariate Analysis,* 15, 279-294.

[31] Hosking, J. (1981): "Fractional differencing," *Biometrika*, 68, 165-176.

[32] Koul, H. and Surgailis, D. (1997): "Asymptotic expansion of M-estimators with long-memory errors," *Annals of Statistics,*25, 818-850.

[33] Leadbetter, M.R., Lindgren, G. and Rootzén, H. (1983): Extremes and related properties of random sequences and processes. Springer-Verlag. New York.

[34] Lobato, I. and Robinson P.M. (1997): "A nonparametric test for $I(0)$," *Review of Economic Studies*.

[35] Mammen, E. (1991): "Estimating a smooth monotone regression function," *Annals of Statistics,* 19, 724-740.

[36] Mandelbrot, B.B. and Van Ness J.W. (1968): "Fractional Brownian motions, fractional noises and applications," *SIAM Review*, 10, 422-437.

[37] Parzen, E. (1986): "Quantile spectral analysis and long-memory time series," *Journal of Applied Probability,* 41-54.

[38] Pickands, J. (1969): "Asymptotic properties of the maximum in a stationary gaussian process," *Transactions of the American Mathematical Society*, 145, 75-86.

[39] Pollard, D. (1984): *Convergence of Stochastic Processes*. Springer-Verlag. Berlin.

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

[41] Robinson, P.M. (1995b): "Gaussian semiparametric estimation of long range dependence," *Annals of Statistics,* 23, 1630-1661.

[42] Stadtmüller, U. (1986): "Asymptotic properties of nonparametric curve estimates," *Periodica Mathematica Hungarica*, 17, 83-108.

[43] Stout, W.F. (1974): *Almost sure convergence*. Academic Press, New York.

[44] Taqqu, M.B. (1975): "Weak convergence to fractional Brownian to the Rosenblatt process," *Zeitschrift. für Wahrscheinlichkeitstheorie und Verwandte Gebiete,* 31, 203-238.

[45] Woodroofe, M.B. (1967): "On the maximum deviation of the sample density," *Annals of the Mathematical Statistics*, 38, 475-481.

[46] Woodroofe, M.B. and Van Ness, J.W. (1967): "The maximum deviation of sample spectral densities," *Annals of the Mathematical Statistics*, 38, 1558-1569.

[47] Woodroofe, M.B. and Sun, J. (1999): "Testing uniformity versus a monotone density," *Annals of Statistics*, 27, 338-360.