A BOOTSTRAP CAUSALITY TEST FOR COVARIANCE STATIONARY PROCESSES*

by

Javier Hidalgo
London School of Economics and Political Science

Contents:

Discussion Paper
No.EM/03/462
October 2003

The Suntory Centre
Suntory and Toyota International Centres for Economics and Related Disciplines
London School of Economics and Political Science
Houghton Street
London WC2A 2AE
Tel.: 020 7955 6698

* This paper is based on research funded by the Economic and Social Research Council (ESRC) reference number: R00023812.

Abstract

This paper examines a nonparametric test for Granger-causality for a vector covariance stationary linear process under, possibly, the presence of long-range dependence. We show that the test converges to a non-distribution free multivariate Gaussian process, say *vec* ($B(\mu)$) indexed by $\mu \in [0,1]$. Because, contrary to the scalar situation, it is not possible, except in very specific cases, to find a time transformation $g(\mu)$ such that *vec* ($B(g(\mu))$) is a vector with independent Brownian motion components, it implies that inferences based on *vec* ($B(\mu)$) will be difficult to implement. To circumvent this problem, we propose bootstrapping the test by two alternative, although similar, algorithms showing their validity and consistency.

Keywords: Causality tests; long-range; bootstrap tests.
JEL No.: C22

Contact address: Professor Javier Hidalgo, Department of Economics, London School of Economics and Political Science, Houghton Street, London WC2A 2AE, UK. Email: f.j.hidalgo@lse.ac.uk

1. INTRODUCTION

In economics and other areas of social sciences, one subject routinely invoked is the concept of causality. This is primarily due to the implication and interpretation that such a concept has on the data. Tests for causality are often performed in the context of unrestricted vector autoregressive ($VAR(P)$) models with P a finite known positive number. See among others, Granger (1969) or Geweke (1982) when the data is short-range dependent, or for variables showing stochastic-trend behaviour, see Sims et al. (1990) or Toda and Phillips (1993). Some extensions are in Hosoya (1991) who analyzes causality for stationary short-range dependent processes which do not necessarily have a VAR representation or Lütkepohl and Poskitt (1996), and references therein, who allow for a $VAR(\infty)$ model.

More recently, Hidalgo (2000) has proposed and examined a test for causality which, unlike the aforementioned papers, covers long-range dependence which has attracted immense attention in recent years. The main attributes of his test are; 1) it is nonparametric, 2) it is consistent, and 3) it has power against $T^{-1/2}$ local alternatives. Thus, he extended previous work in two main directions. First, by allowing a (general) covariance stationary linear process and second, since the test is nonparametric, he avoided the possible danger that a bad specification of the model may induce on the outcome of the test.

Now we briefly describe and discuss the main idea of the test. Consider the $p = p_1 + p_2$ dimensional covariance stationary vector $w_t = (y_t', x_t')'$ admitting the $VAR(\infty)$ representation

$$A(L)\,w_t = \sum_{j=0}^{\infty} A_j w_{t-j} = \varepsilon_t^{(1)}, \quad t = 1, 2, ..., T, \tag{1}$$

where $\varepsilon_t^{(1)}$ is a $p-$dimensional sequence of random variables and A_0 is the identity matrix. The objective of the paper is to test the null hypothesis H_0: $y_t \not\Rightarrow x_t$ (y_t does not cause x_t), against the alternative hypothesis $H_1 : y_t \Rightarrow x_t$ (y_t causes x_t).

Following Sims (1972) or Hosoya (1977), a test for H_0 is equivalent to testing whether the $p_1 \times p_2$ matrices $c(j)$ are simultaneously equal to zero for all $j < 0$ in

$$y_t = \sum_{j=-\infty}^{\infty} c(j)\, x_{t-j} + u_t, \tag{2}$$

where, by construction, $E[u_t \,|x_s, -\infty < s < \infty] = 0$. Alternatively, we can write the null hypothesis H_0 as

$$vec\left(\sum_{j=-\infty}^{0} c(j-1)\cos(\pi j \lambda)\right) = 0 \quad \forall \lambda \in [0,1],$$

or equivalently as

$$S^*(\mu) = \mathrm{Re}\left(\int_0^{\mu} vec\left(\sum_{j=-\infty}^{0} c(j-1)\, e^{-i\pi j\lambda}\right) d\lambda\right) = 0 \quad \forall \mu \in [0,1],$$

where $\mathrm{Re}(a)$ denotes the real part of a complex number (or vector) a. So, we may finally write the hypothesis testing as

$$H_0 : S^*(\mu) = 0 \quad \forall \mu \in [0,1] \text{ against } H_1 : S^*(\mu) \neq 0 \text{ in } \Delta \subset [0,1], \tag{3}$$

where Δ has Lebesgue measure greater than zero.

Given estimates of $c(j)$, for example $\widehat{c}(j)$ in (12), and using Riemann's discrete approximation of integrals by sums, $S^*(\mu)$ can be estimated by

$$S_T(\mu) = \operatorname{Re}\left(\frac{1}{M}\sum_{p=1}^{[M\mu]} vec\left(\sum_{j=2-M}^{0} \widehat{c}(j-1)\, e^{-ij\lambda_{2mp}}\right)\right) \tag{4}$$

where $\lambda_\ell = 2\pi\ell/T$, for integer ℓ, $M = [T/4m]$ with $m = m(T)$ a number which increases slowly with T, that is $m^{-1} + mT^{-1} \to 0$ and $[z]$ denoting the integer part of the number z. The test can thus be based on whether or not $S_T(\mu)$ is significantly different than zero for all $\mu \in [0,1]$ by the implementation of a functional of $S_T(\mu)$, for example a Kolmogorov-Smirnov test.

More specifically, see Theorem 3.2 below, we have that $T^{1/2}S_T(\mu) \overset{weakly}{\Longrightarrow} vec\left(\widetilde{B}(\mu)\right)$, where $\widetilde{B}(\mu)$ is a $p_1 \times p_2$ Gaussian process with covariance structure given by

$$K(\mu_1,\mu_2) = \frac{1}{4\pi}\int_0^{\pi \min(\mu_1,\mu_2)} \left(f_{xx}^{-1}(-\lambda) \otimes f_{uu}(\lambda)\right) d\lambda. \tag{5}$$

However, as it can straightforwardly be observed from (5), the components of the p_1p_2-dimensional Gaussian process $vec\left(\widetilde{B}(\mu)\right)$ are not generally independent. This observation has some consequences regarding the implementation of the test. In particular, it will imply that it is not possible, except in two very specific situations, to find a time transformation, say $g(\mu)$, for which $vec\left(\widetilde{B}(g(\mu))\right)$ becomes a p_1p_2-dimensional vector with independent Brownian motion components. The two cases under which it is possible to find $g(\mu)$ are *a)* $K(\mu_1,\mu_2)$ is a diagonal matrix and *b)* $K(\mu_1,\mu_2) = \min(\mu_1,\mu_2)\,\Omega$ for some positive definite matrix Ω.

The above comments indicate that the results of Theorem 3.2 below may generally be of limited use in order to implement the test for H_0 when p_1 and/or p_2 are greater than one. Although it may be possible to simulate the limiting distribution, this approach can be very demanding and in addition it will require the computation of new critical values everytime a new model/data is under consideration. Therefore, the main objective of the paper is to examine how to circumvent this possible drawback of the test by using a bootstrap algorithm to test H_0. This will justify and permit us to obtain estimates of the critical values of any continuous functional of $T^{1/2}S_T(\mu)$ employed to test H_0.

We finish this section giving the motivation of the bootstrap, whose details are given in Section 4 below, and why we should expect its validity. Suppose for simplicity that $p_1 = p_2 = 1$ and that the model (2) is

$$y_t = \sum_{\ell=-r}^{q} c(\ell)\, x_{t-\ell} + u_t, \quad t = 1,...,T, \tag{6}$$

where both q and r are finite and known a priory. A closer inspection of (6) suggests that the model can be written as

$$w_y(\lambda_j) = \sum_{\ell=-r}^{q} c(\ell)\, w_{x,\ell}(\lambda_j) + w_u(\lambda_j) \quad j = 1,...,T-1, \tag{7}$$

where

$$w_a(\lambda) = \frac{1}{(2\pi T)^{1/2}} \sum_{t=1}^{T} a_t e^{it\lambda}$$

is the *discrete Fourier transform of* $(a_1, ..., a_T)'$ and

$$w_{x,\ell}(\lambda) = \frac{1}{(2\pi T)^{1/2}} \sum_{t=1}^{T} x_{t-\ell} e^{it\lambda},$$

for integer ℓ, is the *discrete Fourier transform of* $(x_{1-\ell}, ..., x_{T-\ell})'$.

It is well known, see Hannan (1970) or Brillinger (1981) among others, that $w_u(\lambda_j)$ is asymptotically uncorrelated although, possibly, heteroscedastic. So, looking at (7), $v_u(\lambda_j) = w_u(\lambda_j) / |w_u(\lambda_j)|$ can be regarded as a sequence of zero mean and asymptotically independent homoscedastic random variables. It is worth noting that it is precisely this observation about the properties of $w_u(\lambda_j)$, which motivated Hannan's (1963) (semiparametric) generalized least squares estimator of the parameters $c(\ell)$ for the model (6) and consequently extended to other useful models in econometrics by Hannan (1965) and Hannan and Terrell (1973) and more recently by Robinson (1991), who also allowed for a data-driven bandwidth. So because the properties of $v_u(\lambda_j)$, it appears that Efron's (1979) resampling scheme should be valid in our framework as it happens to be the case. One consequence of the previous arguments is that even though the errors of the model may be serially correlated with unknown structure, possibly long-range, we are able to avoid the choice of the block length of Künsch's (1989) Moving Block Bootstrap (MBB) which was introduced to handle data with unknown dependence structure or Politis and Romano's (1994) subsampling algorithm. The former was accomplished by writing the model in the frequency domain and then using Efron's (1979) original bootstrap algorithm to the *discrete Fourier transform* of the errors. Finally, it should be mentioned that the structure of the model given in (2) or (7) has some resemblance to Mammen's (1993) model in the sense that as there, we allow the number of regressors in the model to grow to infinity with the sample size.

The remainder of the paper is as follows. Section 2 describes the estimation technique for the matrices $c(j)$. In Section 3, we delimit our statistical framework and present the asymptotic behaviour of $T^{1/2} S_T(\mu)$. In Section 4 we present the bootstrap algorithms and we examine their asymptotic properties. Finally, Section 5 gives the proofs of our results in Sections 3 and 4.

2. THE ESTIMATION OF $c(j)$

In this section we describe the estimation of the matrices $c(j)$ in (2) and discuss why it is more desirable than other approaches, such as least squares estimates (LSE), in the presence of long-range dependence. In the frequency domain, the lag structure given in (2) is described by the frequency response function $C(\lambda) = \sum_{j=-\infty}^{\infty} c(j) e^{-ij\lambda}$ which equals $C(\lambda) = f_{yx}(\lambda) f_{xx}^{-1}(\lambda)$, where $f_{yx}(\lambda)$ and $f_{xx}(\lambda)$ are the indicated elements of the spectral density matrix, $f_{ww}(\lambda)$, of w_t defined from the relationship

$$E\left((w_1 - Ew_1)(w_{j+1} - Ew_1)'\right) = \int_{-\pi}^{\pi} f_{ww}(\lambda) e^{-ij\lambda} d\lambda \quad j = 0, \pm 1, \pm 2, ... \quad (8)$$

So, $c(j)$ may be interpreted as the jth Fourier coefficient of $C(\lambda)$, that is

$$c(j) = \frac{1}{2\pi} \int_0^{2\pi} C(\lambda) e^{ij\lambda} d\lambda. \tag{9}$$

Due to the interpretation of $c(j)$ in (9), Hannan (1963, 1967), see also Brillinger (1981), proposed to estimate $c(j)$ by the sample (discrete) analogue of (9),

$$\breve{c}(j) = \frac{1}{2M} \sum_{p=0}^{2M-1} \widehat{C}_{2mp} e^{ij\lambda_{2mp}}, \tag{10}$$

where $\widehat{C}_{2mp} = \widehat{f}_{yx,2mp} \widehat{f}_{xx,2mp}^{-1}$, and $\widehat{f}_{yx,2mp}$ and $\widehat{f}_{xx,2mp}$ are estimates of $f_{yx,2mp}$ and $f_{xx,2mp}$ respectively given as the indicated elements of (11) below, and where henceforth we shall abbreviate $g(\lambda_p)$ by g_p for a generic function $g(\lambda)$. The estimator $\breve{c}(j)$ in (10) was coined by Sims (1974) as the HI (Hannan's inefficient) estimator.

Our estimator of f_{ww} is given by

$$\widehat{f}_{ww}(\lambda) = \frac{1}{2m+1} \sum_{j=-m}^{m} I_{ww}(\lambda_j + \lambda), \tag{11}$$

where $I_{ww}(\lambda) = (2\pi T)^{-1} \left(\sum_{t=1}^{T} w_t e^{it\lambda} \right) \left(\sum_{t=1}^{T} w_t e^{-it\lambda} \right)'$ is the periodogram of $\{w_t\}$ and where m is as defined in the Introduction.

When analyzing the HI estimator in (10), and similar to technical problems encountered in many other non/semi-parametric estimators, as $\widehat{f}_{xx}(0)$ tries to estimate $f_{xx}(0)$ which, see Condition $C1$ below, may be infinity, proceeding as in Hidalgo (2000) we modify (10) by

$$\widehat{c}(j) = \frac{1}{2M} \sum_{p=1}^{2M-1\,\prime} \widehat{C}_{2mp} e^{ij\lambda_{2mp}}, \tag{12}$$

where $\sum_{p=1}^{2M-1\,\prime} \phi_{2mp} e^{ij\lambda_{2mp}}$ denotes $\sum_{p=1}^{2M-1} \phi_{2mp} e^{ij\lambda_{2mp}} + \phi_{2m}$. Intuitively, we have replaced the estimator of C_0 by that of C_{2m}, that is $\widehat{f}_{yx,2m} \widehat{f}_{xx,2m}^{-1}$.

The motivation of the estimator in (12) is threefold. First, the statistical properties hold the same irrespective of the number of lags specified in (2), which have important consequences when analyzing the properties of $S_T(\mu)$ defined in (4). Second, since there is no gain by exploiting the information on the covariance structure of the errors u_t, as Sims (1974) showed, the HI estimator becomes as efficient as the generalized least squares (GLS) estimator. This motivates the LSE of $c(j)$ given in Robinson (1979), although under stronger assumptions than those we want to impose in this paper.

Finally, the third motivation, which makes the estimate in (12) more appealing when the data may exhibit long-range dependence, is as follows. Assume model (6). When the data is short-range dependent, it is known that, under suitable conditions, the LSE is root-T consistent and asymptotically normal. However, under long-range dependence, as Robinson (1994) observed, when the joint long-range dependence in the regressor x_t and error u_t is sufficiently strong, that is the product of the spectral density functions of x_t and u_t is not integrable, the LSE is no longer root-T consistent nor asymptotically normal.

Motivated by this observation, Robinson and Hidalgo (1997) showed that a class of frequency-domain weighted LSE, including GLS (with parametric error spectral density function) as a special case, is root-T consistent, asymptotically normal and Gauss-Markov efficient in model (6). The intuition why the estimator in Robinson and Hidalgo (1997) is root-T consistent and asymptotically normal is because the weighted function possesses a zero sufficiently strong to compensate for the singularity of the spectral density function induced by the collective long-range dependence of x_t and u_t. So, assuming that $f_{xx}(\lambda)$ has a singularity at the origin, $f_{xx}^{-1}(\lambda)$ will possess a zero at $\lambda = 0$, and we can expect that $\widehat{f}_{xx,2mp}^{-1}$ becomes (asymptotically) a weighted function satisfying the conditions of Robinson and Hidalgo (1997). Indeed, Theorem 3.1 below indicates that the HI estimator given in (12) achieves the root-T consistency and asymptotic normality, so that the HI estimator appears to be a desirable estimator.

3. ASYMPTOTIC PROPERTIES OF (4) AND (12)

Let $z_t = (w_t', u_t')'$ and for $g, h = 1, ..., p+p_1$, denote by $f_{gh}(\lambda)$ the $(g,h)th$ component of the spectral density matrix of z_t, defined as in (8) but with w_t being replaced by z_t. Let us introduce the following regularity conditions:

Condition C1 For all $g = 1, ..., p+p_1$, there exist $C_g \in (0, \infty)$, $d_g \in [0, 1/2)$ and $\alpha \in (0, 2]$, such that

$$f_{gg}(\lambda) = C_g \lambda^{-2d_g}(1 + O(\lambda^{\alpha})) \text{ as } \lambda \to 0+$$

and $|f_{gg}(\lambda)| > 0$ for all $\lambda \in [0, \pi]$.

Defining the coherence between z_{tg} and z_{th} as $R_{gh}(\lambda) = f_{gh}(\lambda) / \left(f_{gg}^{1/2}(\lambda) f_{hh}^{1/2}(\lambda)\right)$, we have

Condition C2 For all $g < h = 2, ..., p+p_1$, $|R_{gh}(\lambda)|$ is twice continuously differentiable in any open set outside the origin and for some $\beta \in (1, 2]$,

$$|R_{gh}(\lambda) - R_{gh}(0)| = O\left(\lambda^{\beta}\right) \text{ as } \lambda \to 0+.$$

Condition C3 $\{w_t\} = \{y_t', x_t'\}'$ and $\{u_t\}$ are covariance stationary linear processes defined as

$$w_t = \sum_{j=0}^{\infty} \Phi_j \varepsilon_{t-j}^{(1)}, \quad \sum_{j=0}^{\infty} \|\Phi_j\|^2 < \infty \quad \text{and } u_t = \sum_{j=0}^{\infty} \Phi_j^u \varepsilon_{t-j}^{(2)}, \quad \sum_{j=0}^{\infty} \left\|\Phi_j^u\right\|^2 < \infty,$$

where Φ_0 and Φ_0^u are the identity matrices and $\|D\|$ denotes the norm of the matrix D. Finally, $\|x_t\|^4$ and $\|u_t\|^4$ are uniformly integrable.

Condition C4 $\left\{\varepsilon_t^{(1)}\right\}$ and $\left\{\varepsilon_t^{(2)}\right\}$ are two mutually independent ergodic sequences with finite fourth moments, where, for $j = 1, 2$, $E\left(\varepsilon_t^{(j)} \middle| \mathcal{F}_{t-1}^{(j)}\right) = 0$, $E\left(\varepsilon_t^{(j)} \varepsilon_t^{(j)\prime} \middle| \mathcal{F}_{t-1}^{(j)}\right) = E\left(\varepsilon_t^{(j)} \varepsilon_t^{(j)\prime}\right) = \Xi^{(j)}$ a.s., (c) $E\left(\varepsilon_{t\ell_1}^{(j)} \varepsilon_{t\ell_2}^{(j)} \varepsilon_{t\ell_3}^{(j)} \middle| \mathcal{F}_{t-1}^{(j)}\right) = \mu_{3,\ell_1\ell_2\ell_3}^{(j)}$ such that $\left|\mu_{3,\ell_1\ell_2\ell_3}^{(j)}\right| <$

∞ for all ℓ_1, ℓ_2, ℓ_3 where $\mathcal{F}_t^{(j)}$ is the σ-algebra of events generated by $\varepsilon_s^{(j)}$, $s \leq t$, and the joint fourth cumulant of $\varepsilon_{t_i \ell_k}^{(j)}$, $\ell_k = 1, ..., p_j$ and $i = 1, ..., 4$ satisfies

$$cum\left(\varepsilon_{t_1\ell_1}^{(j)}, \varepsilon_{t_2\ell_2}^{(j)}, \varepsilon_{t_3\ell_3}^{(j)}, \varepsilon_{t_4\ell_4}^{(j)}\right) = \begin{cases} \kappa_{\ell_1,\ell_2,\ell_3,\ell_4}^{(j)}, & t_1 = t_2 = t_3 = t_4 \\ 0, & \text{otherwise.} \end{cases}$$

Condition C5 $\|(\partial/\partial\lambda)\,\Phi(\lambda)\| = O(\|\Phi(\lambda)\|/\lambda)$ and $\|(\partial/\partial\lambda)\,\Phi^u(\lambda)\| = O(\|\Phi^u(\lambda)\|/\lambda)$ as $\lambda \to 0+$, where

$$\Phi(\lambda) = \sum_{j=0}^{\infty} \Phi_j e^{ij\lambda} \quad \text{and} \quad \Phi^u(\lambda) = \sum_{j=0}^{\infty} \Phi_j^u e^{ij\lambda},$$

such that $\|\Phi(\lambda)\| > 0$ and $\|\Phi^u(\lambda)\| > 0$ for all $\lambda \in [0, \pi]$ and twice continuously differentiable in any open set outside the origin. In addition, for all $g = 1, ..., p+p_1$, $f_{gg}^{-1/2}(\lambda)\,\eta_g(\lambda)$ is a non-zero finite vector, where $\eta_g(\lambda)$ denotes the gth row of $diag(\Phi(\lambda), \Phi^u(\lambda))$.

Condition C6 $\|c(|j|)\| = O\left(|j|^{-3+\tau/2}\right)$ for some $0 < \tau < 1$.

Condition C7 $M^2T^{-1} + M^{\tau-4}T \to 0$ with τ as in C6.

Conditions $C1 - C2$ deal with the behaviour of $f_{zz}(\lambda)$ and they are the same as in Robinson (1995), so his comments apply here. Conditions $C3 - C4$ are restrictive in the linearity they impose, but not otherwise. The requirement of independence between x_t and u_t in $C3 - C4$, as in Robinson and Hidalgo (1997), is necessary for the proof of the asymptotic normality of (12). We believe that it might be possible to relax this condition to some extent, but that will enormously complicate the already technical proof given in Robinson and Hidalgo (1997). Condition $C5$ is similar to that in Robinson (1995). The second part of the condition is not strong, see for instance the comments made after (13) below, once λ^{d_g} is identified as $f_{gg}^{-1/2}$ up to constants there. Condition $C6$ implies that the first derivative of $\|C(\lambda)\|$ is Liptchitz continuous with Liptchitz parameter in the interval $(0, 1 - \tau/2)$. Condition $C7$ gives the admissible values of M. Specifically, the rate of increase of M to infinity cannot be slower than $T^{\delta+1/(4-\tau)}$ or faster than $T^{1/2-\delta}$ for arbitrarily small $\delta > 0$.

Examples of processes satisfying $C1-C5$ are as follows. Let ξ_t be a $p-$dimensional unobservable covariance stationary linear process which possesses a continuous and bounded away from zero spectral density matrix and consider the filter

$$w_t = \sum_{j=0}^{\infty} G(j)\,\xi_{t-j}. \tag{13}$$

Let $G_g(\lambda)$ denote the gth row of the matrix $G(\lambda) = \sum_{j=0}^{\infty} G(j)\,e^{ij\lambda}$ such that $G_g(\lambda)\,\lambda^{d_g}$ tends to a non-zero finite vector as $\lambda \to 0+$, for $g = 1, ..., p$. For instance, let ξ_t be a stationary invertible vector autoregressive moving average ($VARMA$) process with iid innovations and let each w_{tg} be formed by separate fractional integration of the corresponding ξ_t element, so that

$$G(\lambda) = diag\left(\left(1 - e^{i\lambda}\right)^{-d_1}, ..., \left(1 - e^{i\lambda}\right)^{-d_p}\right).$$

Then $C1-C5$ hold. This model is an extension to the vector case of the familiar fractional autoregressive moving average ($ARFIMA$) model, see for instance Granger and Joyeux (1980) or Hosking (1981). Another model which exhibits long-range dependence is the fractional Gaussian noise (fgn) process introduced by Mandelbrot and Van Ness (1968), whose spectral density function, see Sinai (1976), is

$$f(\lambda) = \frac{4\sigma_w^2 \Gamma(2d)}{(2\pi)^{3+2d}} \cos(\pi d) \sin^2(\lambda/2) \sum_{j=-\infty}^{\infty} \left| j + \frac{\lambda}{2\pi} \right|^{-2-2d}$$

where $\sigma_w^2 = E(w_t - E(w_t))^2 < \infty$ and $\Gamma(\cdot)$ denotes the gamma function.

Theorem 3.1. *Assuming C1-C7, for any finite collection* $j_1, ..., j_q$,

(*i*) $T^{1/2} (vec(\widehat{c}(j_1) - c(j_1)), ..., vec(\widehat{c}(j_q) - c(j_q)))' \xrightarrow{d} N\left(0, \Omega = \{\Omega_{j_r j_\ell}\}_{r,\ell=1,..,q}\right)$ *where*

$$\Omega_{j_r j_\ell} = (2\pi)^{-1} \int_{-\pi}^{\pi} \left(f_{xx}^{-1}(-\lambda) \otimes f_{uu}(\lambda)\right) e^{i(j_r - j_\ell)\lambda} d\lambda \tag{14}$$

which corresponds to the asymptotic covariance matrix between $vec(\widehat{c}(j_r))$ *and* $vec(\widehat{c}(j_\ell))$.

(*ii*) *Let* $\widehat{f}_{uu,2mp} = \widehat{f}_{yy,2mp} - \widehat{f}_{yx,2mp} \widehat{f}_{xx,2mp}^{-1} \widehat{f}_{xy,-2mp}$. *A consistent estimator of* $\Omega_{j_r j_\ell}$, $r, \ell = 1, ..., q$, *is*

$$\widehat{\Omega}_{j_r j_\ell} = \frac{1}{2M} \sum_{p=1}^{2M-1} \left(\widehat{f}_{xx,-2mp}^{-1} \otimes \widehat{f}_{uu,2mp}\right) e^{i(j_r - j_\ell)\lambda_{2mp}}.$$

Proof. The proof of this theorem follows by routine extension of that of Hidalgo's (2000) Theorem 1 and so it is omitted. □

So, the results of Theorem 3.1 indicate that the asymptotic properties of the HI estimator given in (12) and established by Hannan (1967) for a finite, possibly of unknown order, distributed lag model and Brillinger (1981) for the infinite distributed lag regression model, when both $f_{xx}(\lambda)$ and $f_{uu}(\lambda)$ are positive and continuous, they also hold true under the possible presence of long-range dependence.

Let

$$S_T(\mu) = \text{Re}\left(\frac{1}{M} \sum_{p=1}^{[M\mu]} vec\left(\sum_{j=2-M}^{0} \widehat{c}(j-1) e^{-ij\lambda_{2mp}}\right)\right).$$

Theorem 3.2. *Assuming C1-C7 and* $\widehat{c}(j)$ *given in* (12), *under* H_0,

$$T^{1/2} S_T(\mu) \overset{weakly}{\Longrightarrow} vec\left(\widetilde{B}(\mu)\right)$$

in $\mathcal{D}^{p_1 p_2}[0, 1]$ *endowed with the Skorohod's metric, where* $vec\left(\widetilde{B}(\mu)\right)$ *is a* $p_1 p_2$-*Gaussian process with covariance structure given in* (5).

Now we elaborate on the results of Theorem 3.2. When $p_1 = p_2 = 1$, and because the function $K(\mu, \mu)$ given in (5) is nondecreasing and nonnegative, $\widetilde{B}(\mu)$ admits the representation $B(K(\mu, \mu))$ in distribution, where $B(\mu)$ is the standard Brownian motion in $[0,1]$. This observation, Theorem 3.2 and the continuous mapping theorem yield

$$\sup_{\mu\in[0,1]}\left|T^{1/2}S_T(\mu)\right| \xrightarrow{d} \sup_{\mu\in[0,K(1,1)]}|B(\mu)| = K^{1/2}(1,1)\sup_{\mu\in[0,1]}|B(\mu)|.$$

Let $\widehat{K}(\mu,\mu)$ be the consistent estimate of $K(\mu,\mu)$ defined as

$$\widehat{K}(\mu,\mu) = \frac{1}{4M}\sum_{p=1}^{[M\mu]}\left(\widehat{f}_{xx,-2mp}^{-1}\otimes\widehat{f}_{uu,2mp}\right).$$

Then, for example, the Kolmogorov-Smirnov test based on $T^{1/2}S_T(\mu)$ would reject the null if $\sup\left\{\widehat{K}^{-1/2}(1,1)\left|T^{1/2}S_T(\mu)\right|,\ \mu\in[0,1]\right\}$ exceeded an appropriate critical value obtained from the boundary crossing probabilities of a Brownian motion, which are readily available on the unit interval. More generally, as

$$\widehat{K}^{-1/2}(1,1)\,T^{1/2}S_T\left(\left(\widehat{K}(\mu,\mu)\right)^{-1}(t)\right) \overset{weakly}{\Longrightarrow} B(\mu)$$

where $\left(\widehat{K}(\mu,\mu)\right)^{-1}(t) = \inf\left\{\mu\in[0,1],\ \widehat{K}(\mu,\mu)\geq t\right\}$, the limiting distribution of any continuous functional of $\widehat{K}^{-1/2}(1,1)\,T^{1/2}S_T\left(\left(\widehat{K}(\mu,\mu)\right)^{-1}(t)\right)$ can be obtained from the distribution of the corresponding functional of $B(\mu)$ on $[0,1]$.

However, when p_1 and/or p_2 are greater than one, a time transformation $g(\mu)$ is not generally available for which $vec\left(\widetilde{B}(\mu)\right)$ admits the representation $vec(B(g(\mu)))$, where $B(\mu)$ has independent Brownian motion components. Two situations however where the time transformation $g(\mu)$ is possible were described in the introduction. Namely, 1) when $K(\mu_1,\mu_2)=\min(\mu_1,\mu_2)\,\Omega$ with Ω a positive definite matrix, and 2) when $K(\mu_1,\mu_2)$ is a diagonal matrix. For the latter, see for example Karatzas and Shreve (1991) Theorem 3.4.1 for the construction of such a transformation. Because, these two aforementioned situations are exceptions rather than the rule, it implies that the results of Theorem 3.2 are somehow of limited use for the purpose of statistical inference. In principle the limiting distribution could be simulated, however because it is non-distribution free, it implies that a practitioner will need to compute new critical values everytime a new model/data were under consideration. We thus propose to bootstrap $\varphi\left(T^{1/2}S_T(\mu)\right)$ for any continuous functional $\varphi(z)$, for example

$$\varphi(z(\mu)) = \sup_{\mu}|z(\mu)| \quad \text{or } \varphi(z(\mu)) = \int_0^1 z^2(\mu)\,d\mu,$$

to circumvent the potential problem of how to implement the test in empirical examples.

4. BOOTSTRAP TESTS FOR CAUSALITY

Since Efron's (1979) seminal paper on the bootstrap, an immense effort has been devoted to its development. The primary motivation for this effort is that it has proved to be a very useful statistical tool. We can cite two main reasons. First, bootstrap methods are capable of approximating the finite sample distribution of statistics better than those based on their asymptotic counterparts. Secondly, and perhaps the most important, it allows computing valid asymptotic quantiles of the limiting distribution in situations where 1) the limiting distribution is unknown or 2) even known, the practitioner is unable to compute its quantiles, which is the situation we have in our framework. So, the aim of this section is to propose two bootstrap procedures for $T^{1/2}S_T(\mu)$, and thus for $\varphi\left(T^{1/2}S_T(\mu)\right)$ for any continuous functional $\varphi(\cdot)$.

We now describe the bootstrap approaches. In a time series context, Efron's (1979) approach has been proved to be inconsistent, see Singh (1981). Motivated by the inconsistency of Efron's (1979) bootstrap, Carlstein (1986) proposed to resample from (nonoverlapping) blocks of data. Later Künsch (1989) generalized this idea by using overlapping blocks, known as the Moving Blocks Bootstrap (*MBB*). Another approach is based on subsampling, see Politis and Romano (1994). Both methods, subsampling and moving blocks, are similar in that they utilize blocks of data of size b. The important difference is that subsampling looks upon these blocks as "subseries" whereas moving blocks use the blocks as "building stones" to construct new pseudo-time series. It is worth noting that Efron's (1979) bootstrap is identical to the *MBB* if $b = 1$.

One potential drawback, however, of the *MBB* or the subsampling bootstrap is their implementation in empirical examples, and in particular, the choice of the block-length b. This apparent drawback is motivated by the observation that, more than anything else, their performance depends rather critically on b, especially for moderate sample sizes. Although some automatic or semiautomatic procedures have recently appeared, see Hall et al. (1997) or Loh's (1987) calibration, these methods can be extremely expensive in computing time or they are based on considerations (such as minimum mean squared errors ones), which are not very relevant to the goal of satisfactorily estimating critical values of $\varphi\left(T^{1/2}S_T(\mu)\right)$.

In this paper, however, we are able to circumvent the problem of how to choose the block-length b even when both x_t and u_t are serially correlated which is the situation we have in our setup. The approach is easy to implement and computationally no more expensive than other bootstrap methods valid in the context of regression models where the errors are independent and identically or heteroscedasticaly distributed. The idea is to perform the bootstrap in the frequency domain, by resampling from the *discrete Fourier transform*. Bootstrap procedures in the frequency domain is by all means no new, see for instance Franke and Härdle (1992) or Dahlhaus and Janas (1996) among others, although instead of resampling from the periodogram ordinates we do so from the *discrete Fourier transform.*

The resampling scheme must be such that the conditional distribution of the bootstrap test, say $\varphi\left(S^*_{T,k}(\mu)\right)$ for $k = 1, 2$, given the data consistently estimates the distribution of $\varphi\left(vec\left(\widetilde{B}(\mu)\right)\right)$ under H_0. That is, denoting by $\mathcal{Z}$ our data set, $\varphi\left(S^*_{T,k}(\mu)\right) \xrightarrow{d^*} \varphi\left(vec\left(\widetilde{B}(\mu)\right)\right)$ under H_0, where "$\xrightarrow{d^*}$" denotes

$$\lim_{n\to\infty} \Pr\left[\varphi\left(S^*_{T,k}(\mu)\right) \le z \middle| \mathcal{Z}\right] \xrightarrow{p} G(z),$$

at each continuity point z of $G(z) = \Pr\left(\varphi\left(vec\left(\widetilde{B}(\mu)\right)\right) \le z\right)$ as defined in Giné and Zinn (1990). Moreover, under local alternatives H_a,

$$H_a: \sum_{j=-\infty}^{0} c(j-1)\cos(\pi j\lambda) = T^{-1/2} d(\pi\lambda),$$

where $d(\vartheta)$ is a continuous function in $[0,\pi]$ such that $0 < |d(\vartheta)|$ in a set $\Delta \subset [0,\pi]$ with positive Lebesgue measure, $\varphi\left(S^*_{T,k}(\mu)\right)$ must also converge in bootstrap distribution to $\varphi\left(vec\left(\widetilde{B}(\mu)\right)\right)$, whereas under the alternative H_1 $\varphi\left(S^*_{T,k}(\mu)\right)$ should be bounded in probability.

The bootstrap algorithms consist of five steps differing in the second step below.

STEP 1 Obtain the HI estimator of $c(\ell)$, for $\ell = 1-M, ..., M$ using (12) and the residuals

$$\widehat{u}_t = y_t - \sum_{\ell=1-M}^{M} \widehat{c}(\ell)\, x_{t-\ell}, \quad t = 1, ..., T.$$

STEP 2 Let $\widetilde{v}_{\widehat{u}}(\lambda_j) = \overline{\widehat{f}}^{-1/2}_{\widehat{u}\widehat{u},j}\, w_{\widehat{u}}(\lambda_j)$, $j = 1, ..., [T/2]$, where $\overline{\widehat{f}}_{\widehat{u}\widehat{u},j} = \widehat{f}_{\widehat{u}\widehat{u},2mp}$ for $2mp - m \leq j < 2mp + m$ and $p = 1, ..., M$.

(a) Compute the *discrete Fourier transform* of the residuals $\widehat{u}_t$, denoted $w_{\widehat{u}}(\lambda_j)$, and compute the standardized $v_{\widehat{u}}(\lambda_j)$ defined as

$$v_{\widehat{u}}(\lambda_j) = \widetilde{\Xi}^{-1/2}\left(\widetilde{v}_{\widehat{u}}(\lambda_j) - [T/2]^{-1}\sum_{\ell=1}^{[T/2]} \widetilde{v}_{\widehat{u}}(\lambda_\ell)\right),$$

where

$$\widetilde{\Xi} = \frac{1}{[T/2]}\sum_{\ell=1}^{[T/2]}\left(\widetilde{v}_{\widehat{u}}(\lambda_\ell) - \frac{1}{[T/2]}\sum_{\ell=1}^{[T/2]}\widetilde{v}_{\widehat{u}}(\lambda_\ell)\right)\left(\widetilde{v}_{\widehat{u}}(\lambda_\ell) - \frac{1}{[T/2]}\sum_{\ell=1}^{[T/2]}\widetilde{v}_{\widehat{u}}(\lambda_\ell)\right)'.$$

Draw independent bootstrap residuals $\eta^*_{j,1}$, $j = 1, ..., [T/2]$, from the empirical distribution function of $v_{\widehat{u}}(\lambda_j)$. That is, for all $j = 1, ..., [T/2]$,

$$\Pr\left\{\eta^*_{j,1} = v_{\widehat{u}}(\lambda_k)\right\} = [T/2]^{-1}, \quad k = 1, ..., [T/2].$$

(b) Let $\widetilde{u}^* = (\widetilde{u}^*_1, \widetilde{u}^*_2, ..., \widetilde{u}^*_T)'$ be a random sample with replacement from the standardized residuals

$$\widetilde{u}_t = \widetilde{\Sigma}^{-1/2}_{\widehat{u}}\left(\widehat{u}_t - T^{-1}\sum_{t=1}^{T}\widehat{u}_t\right),\ \widetilde{\Sigma}_{\widehat{u}} = \frac{1}{T}\sum_{t=1}^{T}\left(\widehat{u}_t - \frac{1}{T}\sum_{t=1}^{T}\widehat{u}_t\right)\left(\widehat{u}_t - \frac{1}{T}\sum_{t=1}^{T}\widehat{u}_t\right)',$$

and obtain the "*discrete Fourier transform*" of $\widetilde{u}^*$ as

$$\eta^*_{j,2} = \frac{1}{T^{1/2}}\sum_{t=1}^{T}\widetilde{u}^*_t e^{-it\lambda_j}, \quad j = 1, ..., [T/2]. \tag{15}$$

STEP 3 For $k = 1, 2$, obtain the bootstrap distributed lag regression model

$$w_{y^*,k}(\lambda_j) = \sum_{\ell=0}^{M} \widehat{c}(\ell)\, w_{x,\ell}(\lambda_j) + \overline{\widehat{f}}_{\widehat{u}\widehat{u}}^{1/2}(\lambda_j)\, \eta_{j,k}^* \quad j = 1, ..., [T/2], \tag{16}$$

and $\overline{\widehat{f}}_{\widehat{u}\widehat{u}}^{1/2}(\lambda_j)$ is obtained using the singular value decomposition of $\overline{\widehat{f}}_{\widehat{u}\widehat{u}}(\lambda_j)$, see for example Brillinger (1981, pp.72-75).

STEP 4 For $k = 1, 2$ and $\ell = 1 - M, ..., M$, compute the bootstrap analogue of (12), that is

$$\widehat{c}_k^*(\ell) = \frac{1}{2M} \sum_{p=1}^{2M-1}{}' \widehat{C}_{2mp,k}^* e^{i\ell\lambda_{2mp}}, \tag{17}$$

where

$$\widehat{C}_{2mp,k}^* = \widehat{C}_k^*(\lambda_{2mp}) = \widehat{f}_{y^*x,k}^*(\lambda_{2mp})\, \widehat{f}_{xx}^{-1}(\lambda_{2mp})$$

with

$$\widehat{f}_{y^*x,k}^*(\lambda_{2mp}) = \frac{1}{2m+1} \sum_{j=-m}^{m} w_{y^*,k}(\lambda_{j+2mp})\, \overline{w_x(\lambda_{j+2mp})},$$

and where $\overline{a}$ denotes the complex combined with transposition of a (complex) matrix a.

And finally,

STEP 5 For $k = 1, 2$, compute the bootstrap analogue of $S_T(\mu)$, $S_{T,k}^*(\mu)$, defined as

$$S_{T,k}^*(\mu) = \mathrm{Re}\left(\frac{1}{M} \sum_{p=1}^{[M\mu]} vec\left(\sum_{\ell=2-M}^{0} \widehat{c}_k^*(\ell-1)\, e^{-i\ell\lambda_{2mp}} \right) \right), \quad \mu \in [0,1]. \tag{18}$$

So, the two different bootstrap approaches differ on the form to compute $\eta_{j,k}^*$ for $k = 1, 2$ in *STEP 2*. Namely, in the first bootstrap we compute $v_{\widehat{u}}(\lambda_j)$ for integer j and resample with replacement obtaining $\eta_{j,1}^*$, whereas in the second bootstrap $\eta_{j,2}^*$ is the "*discrete Fourier transform*" from the random sample obtained with replacement from the normalized residuals $\widetilde{u}_t$.

The proposed bootstrap procedures described in *STEPS 1-5* eliminate the need to choose the block-length b of Künsch's (1989) *MBB* or Politis and Romano's (1994) subsampling approach.

Let us introduce the addition condition:

C8 $\int_{-\pi}^{\pi} \left(f_{xx}^{-1}(\lambda) \otimes f_{uu}(\lambda) \right) |\log(|\lambda|)|\, d\lambda < \infty$.

We now comment on the mild Condition C8. C8 is satisfied if for example u_t is weakly dependent, say an *ARMA* model. Moreover, C8 also holds true if u_t is a long-range dependent process whose spectral density $f(\lambda)$ is bounded by $\lambda^{-1} |\log(\lambda)|^{2+\delta}$ for any $\delta > 0$ and $\lambda \in (0, \pi]$. Observe that because $Eu_t^2 < \infty$, it implies that the spectral density matrix f is upper bounded by $K\lambda^{-1} |\log(\lambda)|^{1+\delta}$ for any $\delta > 0$, so that the former bound on f is not much stronger than the latter bound required for u_t to have finite second moments.

Theorem 4.1. *Let $\varphi(\cdot)$ be a continuous functional. Assuming C1-C8, under the maintained hypothesis $H = H_0 \cup H_1$, for $k = 1, 2$,*

$$\varphi\left(T^{1/2} S^*_{T,k}(\mu)\right) \xrightarrow{d^*} \varphi\left(vec\left(\widetilde{B}(\mu)\right)\right) \text{ in probability.}$$

The first conclusion that we can draw from Theorem 4.1 is that the bootstrap converges in probability to the same process whether or not the null hypothesis holds true. This was expected as we were able to bootstrap the model from the null hypothesis. One consequence is that the power function will be greater than if the algorithm chosen had been the MBB or subsampling. In addition, it also indicates that the bootstrap statistic given in (18) is consistent. So, we can now justify the construction of confidence intervals to test H_0.

To that end, let $\varphi(\cdot)$ denote a continuous functional designed to test H_0, and let $c^f_{n,(1-\alpha)}$ and $c^a_{(1-\alpha)}$ be such that

$$\Pr\left\{\left|\varphi\left(T^{1/2} S_T(\mu)\right)\right| > c^f_{n,(1-\alpha)}\right\} = \alpha$$

and

$$\lim_{n\to\infty} \Pr\left\{\left|\varphi\left(T^{1/2} S_T(\mu)\right)\right| > c^a_{(1-\alpha)}\right\} = \alpha,$$

respectively. Theorem 3.2 and the continuous mapping theorem imply that $c^f_{n,(1-\alpha)} \to c^a_{(1-\alpha)}$, whereas Theorem 4.1 indicates that, for $k = 1, 2$, $c^*_{(1-\alpha),k} \xrightarrow{p} c^a_{(1-\alpha)}$ where $c^*_{(1-\alpha),k}$ is defined as the value which satisfies

$$\Pr\left\{\left|\varphi\left(T^{1/2} S^*_{T,k}(\mu)\right)\right| > c^*_{(1-\alpha),k}\right\} = \alpha.$$

Because the finite sample distribution of $\varphi\left(T^{1/2} S^*_{T,k}(\mu)\right)$ is not available or difficult to obtain, $c^*_{(1-\alpha),k}$ is approximated, as accurately as desired, by a standard Monte-Carlo simulation algorithm. That is, for $k = 1, 2$ let $\eta^{(j)}_k = \left(\eta^{(j)}_{1,k}, ..., \eta^{(j)}_{[T/2],k}\right)$ for $j = 1, ..., B$, and for each j, compute $S^{*(j)}_{T,k}$ as in *STEP 5*. Then, $c^*_{(1-\alpha),k}$ is approximated by the value $c^{*B}_{(1-\alpha),k}$ that satisfies

$$\frac{1}{B}\sum_{j=1}^{B} \mathcal{I}\left(\left|\varphi\left(T^{1/2} S^{*(j)}_{T,k}(\mu)\right)\right| \geq c^{*B}_{(1-\alpha),k}\right) = \alpha,$$

where $\mathcal{I}(A)$ denotes the indicator function of the set A.

5. PROOFS

5.1. Proof of Theorem 3.2

Using the change of subindex $-j$ by j, we can write $T^{1/2} S_T(\mu)$ as

$$\operatorname{Re}\left(\sum_{j=0}^{M-2} vec\left(T^{1/2}\widehat{c}(-j-1)\frac{1}{M}\sum_{p=1}^{[M\mu]} e^{-ij\lambda_{2mp}}\right)\right) \qquad (19)$$

$$= vec\left(T^{1/2}\widehat{c}(-1)\mu + \sum_{j=1}^{M-2} T^{1/2}\widehat{c}(-j-1)\frac{\sin(\pi\mu j)}{\pi j}\right)(1+o(1))$$

since $\mathrm{Re}\left(M^{-1}\sum_{p=1}^{[M\mu]} e^{-ij\lambda_{2mp}}\right) M \to \infty \to (\pi j)^{-1}\sin(\pi\mu j)$ uniformly in $\mu \in [0,1]$.

For ℓ such that $1-M \leq \ell \leq M$, writing

$$a(\ell) - c(\ell) = \frac{1}{2M}\sum_{p=1}^{2M-1} \widehat{f}_{ux,2mp} f_{xx,-2mp}^{-1} e^{i\ell\lambda_{2mp}},$$

let $\widehat{c}(\ell) = a(\ell) + H(\ell)$. Thus, the right side of (19) is

$$vec\left(T^{1/2}a(-1)\mu + \sum_{j=1}^{M-2} T^{1/2}a(-j-1)\frac{\sin(\pi\mu j)}{\pi j}\right)(1+o(1)) \tag{20}$$

$$+vec\left(T^{1/2}H(-1)\mu + \sum_{j=1}^{M-2} T^{1/2}H(-j-1)\frac{\sin(\pi\mu j)}{\pi j}\right)(1+o(1)).$$

The second term of (20) is $o_p(1)$ uniformly in μ, as we now show. Proceeding as with the proof of Theorem 1 of Hidalgo (2000), $T^{1/2}H(j) = O_p\left(M^{-1/2}\log M\right)$ uniformly in j. So, by the triangle inequality

$$\sup_{\mu\in[0,1]}\left\|T^{1/2}H(-1)\mu + \sum_{j=1}^{M-2} T^{1/2}H(-j-1)\frac{\sin(\pi\mu j)}{\pi j}\right\| \tag{21}$$

$$\leq O_p\left(\frac{\log M}{M^{1/2}}\right) + K\sum_{j=1}^{M-2}\frac{1}{j}\left\|T^{1/2}H(-j-1)\right\| = O_p\left(\frac{\log^2 M}{M^{1/2}}\right).$$

So, to complete the proof it suffices to show that the first term of (20) converges weakly to $vec\left(\widetilde{B}(\mu)\right)$. Since H_0 implies that $c(\ell) = 0$ for all $\ell < 0$, the proof follows proceeding as that of Hidalgo's (2000) Corollary 1, and thus it is omitted. □

5.2. Proof of Theorem 4.1

The proof is split into three propositions. In Proposition 5.1 we show that the covariance structure of the bootstrap process $S_{T,k}^{*}(\mu)$ converges in probability to (5), for $k = 1, 2$. In Proposition 5.2 we show that the finite dimensional distributions converge to those of $vec\left(\widetilde{B}(\mu)\right)$ in probability, whereas in Proposition 5.3 we show the tightness condition. From theses three propositions the conclusion of the theorem is standard by the continuous mapping theorem.

Henceforth, we shall denote $E^{*}(\cdot)$ as the bootstrap expectation, that is, for any random variable Y, $E^{*}(Y) = E(Y|w_1, ..., w_T)$. Moreover, for notational simplicity we will assume that $p_1 = p_2 = 1$ in the proofs of the propositions. Finally, a word of caution. Since the bootstrap distributed lag regression models given in (16) have been computed under the null, that is as if $\widehat{c}(\ell)$ were equal to 0 for $\ell < 0$, in what follows for $\ell < 0$, $\widehat{c}(\ell)$ should be understood as being 0.

Proposition 5.1. *Assuming C1-C8, for any ℓ_1 and ℓ_2, and $k=1,2$, we have that*

$$TE^*\left(\left(\widehat{c}_k^*\left(\ell_1\right)-\widehat{c}\left(\ell_1\right)\right)\left(\widehat{c}_k^*\left(\ell_2\right)-\widehat{c}\left(\ell_2\right)\right)'\right)\xrightarrow{P}\Omega_{\ell_1\ell_2} \tag{22}$$

where $\Omega_{\ell_1\ell_2}$ was defined in (14).

Proof. From the definitions of $\widehat{c}_k^*(\ell)$ in (17), for $k=1,2$, and $\widehat{c}(\ell)$ in (12) we have that

$$\widehat{c}_k^*(\ell)-\widehat{c}(\ell)=\frac{1}{2M}\sum_{p=1-M}^{M}\widehat{f}_{xx,2mp}^{-1}\left(\sum_{s=0}^{M}\widehat{c}(s)\,\widehat{f}_{xx,2mp}(-s)-\widehat{C}_{2mp}\widehat{f}_{xx,2mp}\right)e^{i\ell\lambda_{2mp}} \tag{23}$$

$$\begin{aligned}&+\frac{1}{2M}\sum_{p=1-M}^{M}\widehat{f}_{xx,2mp}^{-1}\widehat{f}_{x\widehat{u},k}^{*}\left(\lambda_{2mp}\right)e^{i\ell\lambda_{2mp}}\\ =\;&\widehat{\xi}_{1,k}(\ell)+\widehat{\xi}_{2,k}(\ell).\end{aligned}$$

with $\widehat{f}_{xx,p}(-s)$ as the estimator of the cross-spectrum between x_t and x_{t-s} and for $k=1,2$

$$\widehat{f}_{x\widehat{u},k}^{*}\left(\lambda_{2mp}\right)=\frac{1}{2m+1}\sum_{j=-m}^{m}\widehat{f}_{\widehat{u}\widehat{u},2mp+j}^{1/2}\eta_{2mp+j,k}^{*}w_{x,-2mp-j}.$$

We begin showing that $\widehat{\xi}_{1,k}(\ell)=o_p\left(T^{-1/2}\right)$, which does not depend on the resampling scheme. This will imply that when analysing the behaviour of $\widehat{c}_k^*(\ell)$ or $S_{T,k}^*(\mu)$ in Propositions 5.2 and 5.3 we only need to examine $\widehat{\xi}_{2,k}(\ell)$. That is, $T^{1/2}\left(\widehat{c}_k^*(\ell)-\widehat{c}(\ell)\right)$ can be replaced by $T^{1/2}\widehat{\xi}_{2,k}(\ell)$ there. This will be done in the following lemma.

Lemma 5.1 *For $k=1,2$, $\widehat{\xi}_{1,k}(\ell)=O_p\left(T^{-1/2}\log^{-2}M\right)$ uniformly in ℓ.*

Proof. Denoting $\widetilde{c}(s)=\widehat{c}(s)-c(s)$ and $\widetilde{C}_{2mp}=\sum_{\ell=0}^{M}\widetilde{c}(\ell)\,e^{-i\ell\lambda_{2mp}}$, $\widehat{\xi}_{1,k}(\ell)$ is

$$\frac{1}{2M}\sum_{p=1-M}^{M}\widehat{f}_{xx,2mp}^{-1}\left(\sum_{s=0}^{M}c(s)\,\widehat{f}_{xx,2mp}(-s)-C_{2mp}\widehat{f}_{xx,2mp}\right)e^{i\ell\lambda_{2mp}} \tag{24}$$

$$+\frac{1}{2M}\sum_{p=1-M}^{M}\widehat{f}_{xx,2mp}^{-1}\left(\sum_{s=0}^{M}\widetilde{c}(s)\,\widehat{f}_{xx,2mp}(-s)-\widetilde{C}_{2mp}\widehat{f}_{xx,2mp}\right)e^{i\ell\lambda_{2mp}}.$$

Proceeding as with the proof of Hidalgo's (2000) Theorem 1, in particular his expression $(A.6)$ there, the first term of (24) is $O_p\left(T^{-1/2}M^{(\tau-1)/2}\right)=O_p\left(T^{-1/2}\log^{-2}M\right)$ since $C7$ implies that $\tau<1$. So, we are left to show that the second term of (24) is also $O_p\left(T^{-1/2}\log^{-2}M\right)$. That term is

$$\frac{1}{2M}\sum_{p=1-M}^{M}\left(\frac{1}{\widehat{f}_{xx,2mp}}-\frac{1}{E\left(\widehat{f}_{xx,2mp}\right)}\right)\left(\sum_{s=0}^{M}\widetilde{c}(s)\,\widehat{f}_{xx,2mp}(-s)-\widetilde{C}_{2mp}\widehat{f}_{xx,2mp}\right)e^{i\ell\lambda_{2mp}}$$

$$+\frac{1}{2M}\sum_{p=1-M}^{M}\left(E\left(\widehat{f}_{xx,2mp}\right)\right)^{-1}\left(\sum_{s=0}^{M}\widetilde{c}\left(s\right)\widehat{f}_{xx,2mp}\left(-s\right)-\widetilde{C}_{2mp}\widehat{f}_{xx,2mp}\right)e^{i\ell\lambda_{2mp}}. \tag{25}$$

Because by Hidalgo's (2000) Proposition 3, $\sup_p f_{xx,2mp}^{-1}\left|\widehat{f}_{xx,2mp}-E\left(\widehat{f}_{xx,2mp}\right)\right| = O_p\left(T^{-1/2}M\right)=o_p\left(1\right)$ by $C7$, it is clear that the order of magnitude of (25) will be dominated by that of the second term of (25) which we now examine.

For any $s=0,...,M$, writing

$$\overline{f}_{xx,2mp}\left(-s\right)=\frac{1}{2m+1}\sum_{j=-m}^{m}f_{xx,j+2mp}\left(-s\right),$$

and $\widehat{f}_{xx,2mp}$ as $\widehat{f}_{xx,2mp}\left(0\right)$, from the definition of $\widetilde{C}_{2mp}$ we obtain that

$$\sum_{s=0}^{M}\widetilde{c}\left(s\right)\widehat{f}_{xx,2mp}\left(-s\right)-\widetilde{C}_{2mp}\widehat{f}_{xx,2mp}=\sum_{s=0}^{M}\widetilde{c}\left(s\right)\left(\overline{f}_{xx,2mp}\left(-s\right)e^{is\lambda_{2mp}}-\overline{f}_{xx,2mp}\left(0\right)\right)e^{-is\lambda_{2mp}}$$

$$+\sum_{s=0}^{M}\widetilde{c}\left(s\right)\widetilde{f}_{xx,2mp}\left(-s\right)-\widetilde{C}_{2mp}\widetilde{f}_{xx,2mp}\left(0\right) \tag{26}$$

where $\widetilde{f}_{xx,2mp}\left(-s\right)=\widehat{f}_{xx,2mp}\left(-s\right)-\overline{f}_{xx,2mp}\left(-s\right)$ for all $s=0,...,M$. Since by Theorem 3.1 $T^{1/2}\widetilde{c}\left(s\right)$ behaves as a process with spectral density $f_{xx}^{-1}\left(\lambda\right)f_{uu}\left(\lambda\right)$ and by $C3$ x_t^4 and u_t^4 are uniformly integrable, it implies that $E\left(T^{1/2}\widetilde{c}\left(s_1\right)T^{1/2}\widetilde{c}\left(s_2\right)\right)= O\left(\left|s_1-s_2\right|^{2(d_u-d_x)-1}\right)$ for T large enough by Lemma 4 of Fox and Taqqu (1986) and Serfling's (1980, p.14) Theorem A. So, the contribution of the second moment of the first term on the right of (26) into the second term of (25) is bounded by

$$\frac{K}{TM^2}\sum_{s_1,s_2=0}^{M}\left|s_1-s_2\right|^{2(d_u-d_x)-1}\left(h_{2mp}\left(-s_1\right)-h_{2mp}\left(0\right)\right)\left(h_{2mp}\left(-s_2\right)-h_{2mp}\left(0\right)\right) \tag{27}$$

where

$$h_{2mp}\left(-s\right)=\sum_{p=1-M}^{M}\left(E\left(\widehat{f}_{xx,2mp}\right)\right)^{-1}\left(\left(\overline{f}_{xx,2mp}\left(-s\right)-f_{xx,2mp}\left(-s\right)\right)e^{is\lambda_{2mp}}\right)e^{i(\ell-s)\lambda_{2mp}}$$

because by definition $f_{xx,2mp}\left(-s\right)=f_{xx,2mp}\left(0\right)e^{-is\lambda_{2mp}}$, e.g. $f_{xx,2mp}\left(-s\right)e^{is\lambda_{2mp}}- f_{xx,2mp}\left(0\right)=0$. But because by Lemma 1 and Proposition 1 of Hidalgo (2000), we have that

$$\overline{f}_{xx,2mp}\left(-s\right)-f_{xx,2mp}\left(-s\right)=O\left(f_{xx,2mp}\left(-s\right)p^{-1}\right),\ K^{-1}<f_{xx,2mp}^{-1}E\widehat{f}_{xx,2mp}<K$$

respectively, we conclude that (27) is $O\left(T^{-1}M^{2(d_u-d_x)-1}\log^2 M\right)=O\left(T^{-1}\log^{-4}M\right)$. Then, by Markov's inequality we conclude that the contribution of the first term on the right of (26) into the second term of (25) is $O_p\left(T^{-1/2}\log^{-2}M\right)$ for all ℓ.

To complete the proof of the lemma, we need to prove that the contribution from the second term on the right of (26) to the second term of (25) is also $o_p\left(T^{-1/2}\right)$. But this is the case as we now show. First observe that using expressions $(A.14)$ and $(A.15)$ in Hidalgo (2000), we obtain that the second term on the right of (26) is

$$\sum_{s=0}^{M} \widetilde{c}(s)\, e^{-is\lambda_{2mp}} \varphi_{2mp}(s) \tag{28}$$

$$+\sum_{s=0}^{M} \widetilde{c}(s)\, e^{-is\lambda_{2mp}} \frac{1}{2m+1} \sum_{j=-m}^{m} \left(e^{-is\lambda_j}-1\right)\left(E\left(I_{xx,j+2mp}\right)-\overline{f}_{xx,2mp}(0)\right),$$

where $\widetilde{I}_{xx,r}=\left(I_{xx,r}-E\left(I_{xx,r}\right)\right)$ and $\varphi_{2mp}(s)=(2m+1)^{-1}\sum_{j=-m}^{m}\left(e^{-is\lambda_j}-1\right)\widetilde{I}_{xx,j+2mp}$. Proceeding as with the second term on the right of (26), the contribution of the second term of (28) into the second term of (25) is $O_p\left(T^{-1/2}\log^{-2}M\right)$ since Hidalgo's (2000) Proposition 1 implies that

$$\frac{1}{2m+1}\sum_{j=-m}^{m}\left(e^{-is\lambda_j}-1\right)\left(E\left(I_{xx,j+2mp}\right)-\overline{f}_{xx,2mp}(0)\right)=O\left(M^{-1}p^{-1}\log M\right). \tag{29}$$

On the other hand, proceeding as with the proof of $(A.9)$ in Hidalgo (2000),

$$\widetilde{c}(s) = L_1(s) + L_2(s)$$

$$= L_1(s) + O_p\left(T^{-1/2}M^{-1/2}\log M + M^{-2+\tau/2}\right) \tag{30}$$

uniformly in s, where

$$L_1(s) = \frac{1}{2M}\sum_{q=1-M}^{M} f_{xx,2mq}^{-1}\widehat{f}_{xu,2mq}e^{is\lambda_{2mq}}.$$

In addition, by a routine extension of Hidalgo's (2000) Proposition 2, we obtain that

$$E\left|\varphi_{2mp}(s)\right|^2 = O_p\left(T^{-1}Mf_{xx,2mp}^2\right). \tag{31}$$

So, combining (29) to (31) it implies that the contribution from the first term of (28) to the second term of (25) is

$$\begin{aligned}
& \frac{1}{2M}\sum_{p=1-M}^{M}\frac{1}{E\left(\widehat{f}_{xx,2mp}\right)}e^{i\ell\lambda_{2mp}}\sum_{s=0}^{M-1}L_1(s)\,e_{2mp}^{-is\lambda_{2mp}}\varphi_{2mp}(s) \\
& +O_p\left(T^{-1/2}\left(T^{-1/2}M^{-1/2}\log M + M^{(\tau-1)/2}\right)\right) \\
= \; & \frac{1}{2M}\sum_{p=1-M}^{M}\frac{1}{E\left(\widehat{f}_{xx,2mp}\right)}e^{i\ell\lambda_{2mp}}\sum_{s=0}^{M-1}L_1(s)\,e_{2mp}^{-is\lambda_{2mp}}\varphi_{2mp}(s) \\
& +O_p\left(T^{-1/2}\log^{-2}M\right)
\end{aligned}$$

by $C7$.

Next, because $E\left(\widehat{f}_{xx,2mp}\right)-f_{xx,2mp}=O\left(p^{-1}f_{xx,2mp}\right)$ by Proposition 1 of Hidalgo (2000), we have that the stochastic order of magnitude of the first term on the right of the last displayed equation is bounded by that of

$$\frac{1}{2M}\sum_{p=1-M}^{M} f_{xx,2mp}^{-1}e^{i\ell\lambda_{2mp}}\sum_{s=0}^{M-1} L_1\left(s\right)e_{2mp}^{-is\lambda_{2mp}}\varphi_{2mp}\left(s\right). \tag{32}$$

Since $T^{1/2}L_1\left(s\right)$ converges to a normal random variable and by $C3$ x_t^2 and u_t^2 are uniformly integrable, then Serfling's (1980, p.14) Theorem A implies that $E\left(TL_1^2\left(s\right)\right)$ is bounded, and hence that by the Cauchy-Schwarz inequality,

$$\begin{aligned}
&E\left|\frac{1}{2M}\sum_{p=1-M}^{M} f_{xx,2mp}^{-1}e^{i\ell\lambda_{2mp}}\sum_{s=0}^{\left[M^{1/2}/\log^2 M\right]} L_1\left(s\right)e^{-is\lambda_{2mp}}\varphi_{2mp}\left(s\right)\right| \\
\leq\ &\frac{1}{2M}\sum_{p=1-M}^{M} f_{xx,2mp}^{-1}\sum_{s=0}^{\left[M^{1/2}/\log^2 M\right]}\left(E\left(L_1^2\left(s\right)\right)E\left|\varphi_{2mp}\left(s\right)\right|^2\right)^{1/2} \\
=\ &O\left(T^{-1/2}\log^{-2}M\right)
\end{aligned}$$

by (31) and then $C7$. Note that the convergence is uniform in ℓ.

So, to complete the proof of the lemma it remains to show that uniformly in ℓ

$$\begin{aligned}
&\frac{1}{2M}\sum_{p=1-M}^{M} f_{xx,2mp}^{-1}e^{i\ell\lambda_{2mp}}\sum_{s=1+\left[M^{1/2}/\log^2 M\right]}^{M} L_1\left(s\right)e^{-is\lambda_{2mp}}\varphi_{2mp}\left(s\right) \\
=\ &O_p\left(T^{-1/2}\log^{-2}M\right).
\end{aligned}$$

But by summation by parts, the left side of the last displayed equation is

$$\begin{aligned}
&\frac{1}{2M}\sum_{p=1-M}^{M} f_{xx,2mp}^{-1}e^{i\ell\lambda_{2mp}}\sum_{s=1+\left[M^{1/2}/\log^2 M\right]}^{M}\left\{\frac{\left(2\pi s\right)^{1/2}}{2m+1}\sum_{j=-m}^{m} e^{-is\lambda_j}\left(e^{-i\lambda_j}-1\right)\widetilde{I}_{xx,j+2mp}\right. \\
&\qquad\left.\times\frac{1}{\left(2\pi s\right)^{1/2}}\sum_{v=1+\left[M^{1/2}/\log^2 M\right]}^{s} L_1\left(v\right)e^{-iv\lambda_{2mp}}\right\}.
\end{aligned} \tag{33}$$

Observing that $T^{1/2}L_1\left(v\right)$ behaves as a Gaussian process with spectral density function $g\left(\lambda\right)=f_{xx}^{-1}\left(\lambda\right)f_{uu}\left(\lambda\right)$ by Theorem 3.1, Theorem 2 of Robinson (1995) implies that

$$E\left|\frac{1}{\left(2\pi s\right)^{1/2}}\sum_{v=1}^{s} T^{1/2}L_1\left(v\right)e^{-iv\lambda_{2mp}}\right|^2-g_{2mp}=O\left(p^{-1}g_{2mp}\log\left(2mp\right)\right).$$

So, Markov's inequality implies that (33) is, uniformly in ℓ, $O_p\left(T^{-1}M\log M\right) = O_p\left(T^{-1/2}\log^{-2}M\right)$ by C7 and that

$$E\left|\frac{1}{2m+1}\sum_{j=-m}^{m}e^{-is\lambda_j}\left(e^{-i\lambda_j}-1\right)\widetilde{I}_{xx,j+2mp}\right|^2 = O\left(T^{-1}M^{-1}f^2_{xx,2mp}\right)$$

because $M\left|e^{-i\lambda_j}-1\right|$ is bounded for $|j| \leq m$ and $\sum_{s=1}^{M} s^{1/2} = O\left(M^{3/2}\right)$. This completes the proof that the expression (32) is $O_p\left(T^{-1/2}\log^{-2}M\right)$ and the lemma.□

To complete the proof of the proposition, it suffices to prove that

$$TE^*\left(\widehat{\xi}_{2,k}(\ell_1)\,\widehat{\xi}_{2,k}(\ell_2)'\right) \overset{P}{\to} \Omega_{\ell_1\ell_2}. \tag{34}$$

First, from the definition of $\widehat{\xi}_{2,k}(\ell)$ in (23),

$$T^{1/2}\widehat{\xi}_{2,k}(\ell) = \frac{T^{1/2}}{2M}\sum_{p=1-M}^{M}\left(\widehat{f}^{-1}_{xx,2mp}-f^{-1}_{xx,2mp}\right)\widehat{f}^{*}_{x\widehat{u},k}(\lambda_{2mp})\,e^{-i\ell\lambda_{2mp}}$$

$$+\frac{T^{1/2}}{2M}\sum_{p=1-M}^{M}f^{-1}_{xx,2mp}\widehat{f}^{*}_{x\widehat{u},k}(\lambda_{2mp})\,e^{-i\ell\lambda_{2mp}}. \tag{35}$$

Since the first term on the right of (35) is of smaller order of magnitude than the second term on the right, (34) is shown if

$$\frac{T}{4M^2}E^*\left(\left(\sum_{p=1-M}^{M}f^{-1}_{xx,2mp}\widehat{f}^{*}_{x\widehat{u},k}(\lambda_{2mp})\,e^{-i\ell_1\lambda_{2mp}}\right)\left(\sum_{p=1-M}^{M}f^{-1}_{xx,2mp}\widehat{f}^{*}_{x\widehat{u},k}(-\lambda_{2mp})\,e^{i\ell_2\lambda_{2mp}}\right)\right) \tag{36}$$

converges in probability to $\Omega_{\ell_1\ell_2}$. The last claim follows because, for $k=1,2$, $\eta^*_{r,k}$ are $iid\,(0,1)$ random variables and $T=4mM$, so that the second moment, in bootstrap sense, of the first term on the right of (35) is

$$\frac{1}{2M}\sum_{p=1-M}^{M}\left(\widehat{f}^{-1}_{xx,2mp}-f^{-1}_{xx,2mp}\right)^2\widehat{f}_{\widehat{u}\widehat{u},2mp}\frac{1}{2m+1}\sum_{j=-m}^{m}I_{xx,j+2mp}e^{-i(\ell_1-\ell_2)\lambda_{2mp}}$$

and Proposition 3 of Hidalgo (2000) implies that both $f^{-1}_{xx,2mp}\left(\widehat{f}_{xx,2mp}-f_{xx,2mp}\right)$ and $f^{-1}_{uu,2mp}\left(\widehat{f}_{\widehat{u}\widehat{u},2mp}-f_{uu,2mp}\right)$ are $o_p(1)$ uniformly in p and the integrability of $f^{-1}_{xx}(\lambda)\,f_{uu}(\lambda)$. Again using the independence of $\eta^*_{r,k}$ for $k=1,2$, (36) is

$$\frac{1}{2M}\sum_{p=1-M}^{M}f^{-2}_{xx,2mp}\widehat{f}_{\widehat{u}\widehat{u},2mp}\frac{1}{2m+1}\sum_{j=-m}^{m}I_{xx,j+2mp}e^{-i(\ell_1-\ell_2)\lambda_{2mp}}$$

$$= \frac{1}{2M}\sum_{p=1-M}^{M}f^{-1}_{xx,2mp}f_{uu,2mp}\left(\left[f^{-1}_{xx,2mp}\widehat{f}_{xx,2mp}-1\right]+1\right)e^{-i(\ell_1-\ell_2)\lambda_{2mp}}\left(1+o_p(1)\right)$$

because uniformly in p, $\widehat{f}_{\widehat{u}\widehat{u},2mp} = f_{uu,2mp} + o_p(f_{uu,2mp})$. But by Proposition 3 of Hidalgo (2000), the right side of the last displayed equation is

$$\frac{1}{2M}\sum_{p=1-M}^{M} f_{xx,2mp}^{-1} f_{uu,2mp} e^{-i(\ell_1-\ell_2)\lambda_{2mp}} \left(1+o_p(1)\right) \xrightarrow{P} \Omega_{\ell_1\ell_2},$$

which completes the proof of the proposition. □

Proposition 5.2. *Assuming C1-C8, for any finite collection $\ell_1, ..., \ell_r$, for $k = 1, 2$*

$$T^{1/2}\left(\widehat{c}_k^*(\ell_1) - \widehat{c}(\ell_1), ..., \widehat{c}_k^*(\ell_r) - \widehat{c}(\ell_r)\right) \xrightarrow{d^*} N\left(0, \Omega = \left(\Omega_{\ell_h\ell_j}\right)_{h,j=1,...,r}\right)$$

where $\Omega_{\ell_h\ell_j}$ is defined in (14) and denotes the indicated element of Ω, which corresponds to the asymptotic covariance matrix between $\widehat{c}_k^*(\ell_h) - \widehat{c}(\ell_h)$ and $\widehat{c}_k^*(\ell_j) - \widehat{c}(\ell_j)$.

Proof. Following our comment before Lemma 5.1, it suffices to show that

$$T^{1/2}\left(\widehat{\xi}_{2,k}(\ell_1), ..., \widehat{\xi}_{2,k}(\ell_r)\right) \xrightarrow{d^*} N\left(0, \Omega = \left(\Omega_{\ell_h\ell_j}\right)_{h,j=1,...,r}\right).$$

Furthermore, from Lemma 5.1 and proceeding as with the proof of (35), it is clear that we only need to examine the above when $\widehat{\xi}_{2,k}(\ell)$ is replaced by the second term on the right of (35), that is

$$\frac{1}{2M}\sum_{p=1-M}^{M} f_{xx,2mp}^{-1} \widehat{f}_{x\widehat{u},k}^*(\lambda_{2mp}) e^{-i\ell\lambda_{2mp}}.$$

Next, from the definition of $\widehat{f}_{x\widehat{u},k}^*(\lambda_{2mp})$, standard algebra yields that the last displayed expression is

$$\widehat{\vartheta}(\ell) = \frac{1}{T}\sum_{q=1-[T/2]}^{[T/2]} h_{xx,q}^{-1} \widehat{h}_{\widehat{u}\widehat{u},q}^{1/2} \eta_{q,k}^* w_{x,-q} e^{-i\ell\lambda_q} \tag{37}$$

where $h_{aa,q}$ is a step function with jumps at the points $2mp$, $1 - M \le p \le M$.

So, by Wold device it suffices to show that for all set of constants a_ℓ, $\ell = 1, ..., r$ such that $\sum_{\ell=1}^{r} a_\ell^2 = 1$,

$$(a)\;\; E^*\left|T^{1/2}\sum_{\ell=1}^{r} a_\ell \widehat{\vartheta}(\ell)\right|^2 \xrightarrow{P} \sum_{\ell_1,\ell_2=1}^{r} a_{\ell_1}\Omega_{\ell_1\ell_2}a_{\ell_2} \tag{38}$$

and denoting $\zeta_{q,T} = \sum_{\ell=1}^{r} a_\ell T^{-1/2} h_{xx,q}^{-1} \widehat{h}_{\widehat{u}\widehat{u},q}^{1/2} \eta_{q,k}^* w_{x,-q} e^{-i\ell\lambda_q}$,

$$(b)\;\; \sum_{q=1-[T/2]}^{[T/2]} E^*\left|\zeta_{q,T}\mathcal{I}\left(\left|\zeta_{q,T}\right|^2 > \psi\right)\right|^2 \xrightarrow{P} 0 \tag{39}$$

for all $\psi > 0$. (38) indicates that the second bootstrap moments converge in probability to those of the asymptotic distribution of the HI estimator given in (12), whereas (39) is simply the Lindeberg's condition.

Part (a) follows by direct application of Proposition 5.1.

We now show part (b). We first examine $\sup_{q=1,\ldots,[T/2]} T^{-1} \left\| I_{xx,q} \widehat{h}_{\widehat{u}\widehat{u},q} \right\|$ which is

$$\sup_{q=1,\ldots,[T/2]} T^{-1} f_{xx,q} h_{uu,q} \left| f_{xx,q}^{-1/2} w_{x,q} \right|^2 h_{uu,q}^{-1} \widehat{h}_{\widehat{u}\widehat{u},q}. \tag{40}$$

Next, denoting by $w_\xi(\lambda)$ the *discrete Fourier transform* of the innovations ξ_t in the Wold decomposition of x_t and $\sigma_\xi^2 = E\xi_t^2$, by Robinson's (1995) Theorem 2 and an obvious extension to all $1 \le j \le [T/2]$, see Giraitis et al.'s (2001) Lemma 4.4,

$$E \left| f_{xx,q}^{-1/2} w_{x,q} - (2\pi)^{1/2} \frac{w_{\xi,q}}{\sigma_\xi} \right|^2 = O\left(\frac{\log q}{q} \right)$$

(observe that $w_{x,q}$ is normalized by $f_{xx,q}^{-1/2}$ instead of its approximation given in $C1$).

So, using the inequality $\sup_t |\phi_t|^2 \le \sum_t |\phi_t|^2$, we have that

$$E \left(\sup_{q=1,\ldots,[T/2]} \left| f_{xx,q}^{-1/2} w_{x,q} - (2\pi)^{1/2} \frac{w_{\xi,q}}{\sigma_\xi} \right|^2 \right)$$

$$\le \sum_{q=1}^{[T/2]} E \left| f_{xx,q}^{-1/2} w_{x,q} - (2\pi)^{1/2} \frac{w_{\xi,q}}{\sigma_\xi} \right|^2 = O\left(\log^2 T\right). \tag{41}$$

On the other hand, by An et al. (1983),

$$\sup_{q=1,\ldots,[T/2]} \left((2\pi) \sigma_\xi^{-2} \log^{-1} T \left| w_{\xi,q} \right|^2 \right) \le 1 \quad \text{a.s.}. \tag{42}$$

Then, combining (41), (42) and that Proposition 3 of Hidalgo (2000) and Theorem 3.1 imply that

$$h_{uu,q}^{-1} h_{\widehat{u}\widehat{u},q}^{-1} - 1 = O_p\left(MT^{-1} + q^{-1} \log q \right),$$

we have that (40) is bounded by

$$\begin{aligned} & K \sup_{q=1,\ldots,[T/2]} T^{-1} f_{xx,q}^{-1} f_{uu,q} \sup_{q=1,\ldots,[T/2]} \left| f_{xx,q}^{-1/2} w_{x,q} \right|^2 \left| h_{uu,q}^{-1} h_{\widehat{u}\widehat{u},q}^{-1} \right| \\ \le \; & K \sup_{q=1,\ldots,[T/2]} T^{-1} f_{xx,q}^{-1} f_{uu,q} \log^2 T \le D \log^{-\delta} T \end{aligned}$$

in probability since integrability of $\left\| f_{xx}^{-1}(\lambda) f_{uu}(\lambda) \right\| |\log(|\lambda|)|$ implies that for any arbitrarily small $\delta > 0$, $\left\| f_{xx,q}^{-1} f_{uu,q} \right\| |\log \lambda_q| \le K \lambda_q^{-1} |\log \lambda_q|^{-1-\delta}$, so that $\left\| f_{xx,q}^{-1} f_{uu,q} \right\| \le K \lambda_q^{-1} |\log \lambda_q|^{-2-\delta}$. Now, for $k = 1$, we can conclude that the left side of (39) is bounded by

$$\begin{aligned} & \frac{K}{T} \sum_{q=1-[T/2]}^{[T/2]} f_{xx,q}^{-2} I_{xx,q} f_{uu,q} E^* \left| \eta_{q,1}^* \right|^2 \mathcal{I}\left(\left| \eta_{q,1}^* \right|^2 > \psi \log^\delta T \right) \\ = \; & K E^* \left(\left| \eta_{1,1}^* \right|^2 \mathcal{I}\left(\left| \eta_{1,1}^* \right|^2 > \psi \log^\delta T \right) \right) \frac{1}{T} \sum_{q=1-[T/2]}^{[T/2]} f_{xx,q}^{-2} I_{xx,q} f_{uu,q}, \end{aligned}$$

since $\eta^*_{q,1}$ are *iid* random variables with zero mean and unit variance. On the other hand, for $k=2$, the left side of (39) is bounded by

$$
\begin{aligned}
& \frac{K}{T} \sum_{q=1-[T/2]}^{[T/2]} f_{xx,q}^{-2} I_{xx,q} I_{uu,q} E^* \left|\eta^*_{q,2}\right|^2 \mathcal{I}\left(\left|\eta^*_{q,2}\right|^2 > \psi \log^\delta T\right) \\
= \;& K \sup_T \sup_q E^* \left(\left|\eta^*_{q,2}\right|^2 \mathcal{I}\left(\left|\eta^*_{q,2}\right|^2 > \psi \log^\delta T\right)\right) \frac{1}{T} \sum_{q=1-[T/2]}^{[T/2]} f_{xx,q}^{-2} I_{xx,q} f_{uu,q},
\end{aligned}
$$

But the second factor on the right of the last displayed equation converges in probability to $(2\pi)^{-1} \Omega_{1,1}$, whereas the first factor on the right of the last equation converges to zero because $\psi > 0$ and

$$
\begin{aligned}
E^* \left(\left|\eta^*_{q,2}\right|^2 \mathcal{I}\left(\left|\eta^*_{q,2}\right|^2 > \psi \log^\delta T\right)\right) &\leq \frac{D}{\psi^2 \log^{2\delta} T} E^* \left|\eta^*_{q,2}\right|^4 \\
&\leq \frac{D}{\psi^2 \log^{2\delta} T} \left\{ \frac{1}{T} \sum_{t=1}^{T} \widehat{u}_t^4 + \left(\frac{1}{T} \sum_{t=1}^{T} \widehat{u}_t^2 \right)^2 \right\}.
\end{aligned}
$$

But, by Theorem 3.1 the right side of the last displayed inequality is

$$
\frac{D}{\psi^2 \log^{2\delta} T} \left\{ \frac{1}{T} \sum_{t=1}^{T} u_t^4 + \left(\frac{1}{T} \sum_{t=1}^{T} u_t^2 \right)^2 \right\} (1 + o_p(1)).
$$

But, by a well-know argument, see Stout's (1974) Theorem 3.5.8, Condition $C4$ implies that u_t is also ergodic. So, the last displayed expression converges in probability to zero because $\delta > 0$. □

Proposition 5.3. *Assuming C1-C8, for k = 1, 2*

$$
\varphi\left(S^*_{T,k}(\mu)\right) \xrightarrow{d^*} \varphi\left(\widetilde{B}(\mu)\right)
$$

for any continuous functional $\varphi(\cdot)$.

Proof. Because in (16) $\widehat{c}(j) = 0$ for all $j < 0$, we have that $T^{1/2}\widehat{c}^*_k(j) = T^{1/2}\widehat{\xi}_{1,k}(j) + \widehat{\xi}_{2,k}(j)$ where

$$
\widehat{\xi}_{2,k}(j) = \frac{T^{1/2}}{2M} \sum_{p=1-M}^{M} f_{xx,2mp}^{-1} \widehat{f}^*_{x\widehat{u},k}(\lambda_{2mp}) e^{ij\lambda_{2mp}}
$$

suppressing reference to T in $\widehat{\xi}_{2,k}(j)$ or $\widehat{\xi}_{1,k}(j)$. Thus, denoting $\widehat{\xi}_{2,k}(j)$ by $a^*_k(j)$, with the change of subindices $-j$ by j

$$
T^{1/2} S^*_{T,k}(\mu) = \sum_{j=0}^{M-2} \left(a^*_k(-j-1) + T^{1/2}\widehat{\xi}_{1,k}(-j-1)\right) \frac{1}{M} \sum_{p=1}^{[M\mu]} e^{ij\lambda_{2mp}} \tag{43}
$$

$$
= \left(a^*_k(-1) + T^{1/2}\widehat{\xi}_{1,k}(-1)\right)\mu + \sum_{j=1}^{M-2} \left(a^*_k(-j-1) + T^{1/2}\widehat{\xi}_{1,k}(-j-1)\right) g_j(\mu)
$$

where $g_j(\mu) = M^{-1} \sum_{p=1}^{[M\mu]} e^{ij\lambda_{2mp}}$, which satisfies that

$$\operatorname{Re}(g_j(\mu)) = \frac{\sin(\pi j \mu)}{\pi j} + O\left(\frac{1}{M}\right) \tag{44}$$

uniformly in $\mu \in [0,1]$ by Brillinger (1981, p.15). Now,

$$\sup_{\mu} \left| \sum_{j=1}^{M-2} T^{1/2} \widehat{\xi}_{1,k}(-j-1) \operatorname{Re}(g_j(\mu)) \right| = o_{p^*}(1)$$

because by Lemma 5.1, $E^* \left| T^{1/2}\widehat{\xi}_{1,k}(-j-1) \right|^2 = \left| T^{1/2}\widehat{\xi}_{1,k}(-j-1) \right|^2 = O_p\left(\log^{-4} M\right)$ uniformly in j and $\sum_{j=1}^{M-1} j^{-1} = O(\log M)$.

So, using (44) the behaviour of the real part of (43) is that of

$$a_k^*(-1)\mu + \sum_{j=1}^{b-1} a_k^*(-j-1) \operatorname{Re}(g_j(\mu)) + \sum_{j=b}^{M-2} a_k^*(-j-1) \operatorname{Re}(g_j(\mu))$$

where b is a fixed but large constant. The proof is completed if, for $k = 1, 2$,

$$G_b(\mu) = \sum_{j=0}^{b-1} a_k^*(-j) \operatorname{Re}(g_j(\mu)) \text{ converges to a Gaussian process indexed by } \mu, \tag{45}$$

$$\left| E^* G_b^2(\mu) - \frac{1}{2} \int_0^{2\pi\mu} f_{xx}^{-1}(\lambda) f_{uu}(\lambda) d\lambda \right| = o_p(\upsilon(b)), \tag{46}$$

where $\upsilon(b) \to 0$ as $b \to \infty$, and

$$\sum_{j=b}^{M-2} a_k^*(-j) \operatorname{Re}(g_j(\mu)) \text{ is small uniformly in } M \text{ and } \mu, \tag{47}$$

where $(a_k^*(-j))_{j=0,\ldots,b-1} \xrightarrow{d^*} N(0,\Omega)$ by Proposition 5.2 with the $(j_1, j_2)\,th$ element of Ω denoted by $\Omega_{|j_1 - j_2|}$.

We begin with the assertion (47). Because by $C3$, $E^*(a_k^*(-j))^4 < K$, Proposition 5.2 and Serfling's (1980, p.14) Theorem A imply that $E^* |a_k^*(-j_1) a_k^*(-j_2) a_k^*(-j_3) a_k^*(-j_4)|$ converges to that of the limit distribution of $a_k^*(-j)$.

Using that for b large enough (44) implies that $|jg_j(\mu)| < K$, for $0 < s_1 < s_2 \leq M-1$, we have that

$$\left| \sum_{j=s_1}^{s_2} a_k^*(-j) g_j(\mu) \right|^2 \leq \left| \sum_{j=s_1}^{s_2} a_k^*(-j) \frac{e^{2\pi j \mu}}{j} \right|^2 \leq 2 \sum_{v=1}^{s_2 - s_1} \left| \sum_{j=s_1}^{s_2 - v} \frac{a_k^*(-j) a_k^*(-j-v)}{j(j+v)} \right|. \tag{48}$$

Next,

$$E^* \left| \sum_{j=s_1}^{s_2-v} \frac{a_k^*(-j) a_k^*(-j-v)}{j(j+v)} \right|^2 = \sum_{j_1, j_2 = s_1}^{s_2 - v} E^* \left(\frac{a_k^*(-j_1) a_k^*(-j_2) a_k^*(-j_3) a_k^*(-j_4)}{j_1(j_1+v) j_2(j_2+v)} \right)$$

whose expectation is

$$\sum_{j_1,j_2=s_1}^{s_2-v}\left|\frac{\Omega_v^2+\Omega_{|j_1-j_2|}^2+\Omega_{|j_1-j_2-v|}\Omega_{|j_1-j_2+v|}}{j_1\left(j_1+v\right)j_2\left(j_2+v\right)}\right|\left(1+o\left(1\right)\right)$$

because $a_k^*\left(-j\right)\overset{d^*}{\to}\mathcal{N}\left(0,\Omega\right)$ and the uniform integrability of x_t^4 and u_t^4 by $C3$ imply that by Theorem A of Serfling (1980, $p.$14), the fourth moments converge to those of the limit distribution whose fourth cumulant is 0 by normality.

Next, let $s_1=2^\ell$ and $s_2=2^{\ell+1}$. Then by the Cauchy-Schwarz inequality, the left side of (48) has expectation bounded by

$$\begin{aligned}
& K\sum_{v=1}^{2^\ell}\left|\sum_{j_1,j_2=2^\ell}^{2^{\ell+1}-v}\frac{\Omega_v^2+\Omega_{|j_1-j_2|}^2+\Omega_{|j_1-j_2-v|}\Omega_{|j_1-j_2+v|}}{j_1\left(j_1+v\right)j_2\left(j_2+v\right)}\right|^{1/2}\\
\leq\ & K\sum_{v=1}^{2^\ell}\left|\sum_{j_1=2^\ell}^{2^{\ell+1}-v}\frac{1}{j_1^2\left(j_1+v\right)^2}\right|^{1/2}+K\sum_{v=1}^{2^\ell}\left|\sum_{2^\ell=j_1<j_2}^{2^{\ell+1}-v}\frac{\Omega_v^2+\Omega_{|j_1-j_2|}^2+\Omega_{|j_1-j_2-v|}\Omega_{|j_1-j_2+v|}}{j_1\left(j_1+v\right)j_2\left(j_2+v\right)}\right|^{1/2},
\end{aligned}$$

by the triangle inequality. But Lemma 4.1 of Fox and Taqqu (1986) implies that $\Omega_v=O\left(v^{-1+\alpha}\right)$, where $\alpha=d_x-d_u<1$, since Ω_v is the vth Fourier coefficient of $f_{xx}^{-1}\left(\lambda\right)f_{uu}\left(\lambda\right)$. So, after standard calculations, we have that the left side of the last displayed inequality is bounded by

$$K\sum_{v=1}^{2^\ell}\left|\sum_{j_1=2^\ell}^{2^{\ell+1}-v}\frac{1}{j_1^2\left(j_1+v\right)^2}\right|^{1/2}+K\sum_{v=1}^{2^\ell}v^{-1+\alpha}\left|\sum_{j_1=2^\ell}^{2^{\ell+1}-v}\frac{1}{j_1\left(j_1+v\right)}\right|\leq K\left(2^{-\ell/2}+2^{-\ell(1-\alpha)}\right),$$

and then, using that $\operatorname{Re}\left(g_j\left(\mu\right)\right)=2^{-1}\left(g_j\left(\mu\right)+g_j\left(-\mu\right)\right)$, we conclude that

$$\begin{aligned}
& E\left(E^*\left|\max_{0\leq\mu\leq1}\left|\sum_{j=2^\ell}^{2^{\ell+1}}a_k^*\left(-j\right)\operatorname{Re}\left(g_j\left(\mu\right)\right)\right|\right|\right)\\
\leq\ & K\sum_{v=1}^{2^\ell}\left(\left|\sum_{j_1=2^\ell}^{2^{\ell+1}-v}\frac{1}{j_1^2\left(j_1+v\right)^2}\right|^{1/2}+v^{-1+\alpha}\left|\sum_{j_1=2^\ell}^{2^{\ell+1}-v}\frac{1}{j_1\left(j_1+v\right)}\right|\right)\\
\leq\ & K\left(2^{-\ell/2}+2^{-\ell(1-\alpha)}\right).
\end{aligned}$$

Now choose $b=2^n$ to conclude that with probability greater than $1-K\left(2^{-\ell/2}+2^{-\ell(1-\alpha)}\right)$,

$$Z_\ell^*=\max_{0\leq\mu\leq1}\left|\sum_{j=2^\ell}^{2^{\ell+1}}a_k^*\left(-j\right)\operatorname{Re}\left(g_j\left(\mu\right)\right)\right|\leq K\left(2^{-\ell/2}+2^{-\ell(1-\alpha)}\right),$$

which implies that the expression in (47) is, in absolute value, bounded by

$$\sum_{\ell=n}^{[\log(M)]+1}Z_\ell^*\leq K\left(2^{-n(1-\alpha)}+2^{-n/2}\right).$$

From here we conclude the proof of (47) by choosing n large enough.

Next we prove assertion (45). From Proposition 5.2, the finite dimensional distributions of $G_k(\mu)$ converges to those of a normal random variable. To finish, we need to check the tightness condition, which by Billingsley's (1968) Theorem 15.6, it suffices to check the sufficient moment condition

$$E^*\left[(G_b(\mu)-G_b(\mu_1))^2(G_b(\mu_2)-G_b(\mu))^2\right]\le KH_T(\mu_2,\mu_1)(\mu_2-\mu_1)^2.$$

where $H_T(\mu_2,\mu_1)=K(1+o_p(1))$ with $0\le\mu_1<\mu<\mu_2\le 1$. First observe that we can take $M^{-1}<|\mu_2-\mu_1|$ since otherwise either μ_1 and μ lie in the same subinterval $\left[M^{-1}q, M^{-1}(q+1)\right]$ or else μ and μ_2 do; in either of these cases the left side of the last displayed inequality vanishes. But since $(\mu-\mu_1)(\mu_2-\mu)\le(\mu_2-\mu_1)^2$, it implies that a sufficient condition for the last displayed inequality to hold true is

$$E^*\left[(G_b(\mu_2)-G_b(\mu_1))^4\right]\le KH_T(\mu_2,\mu_1)(\mu_2-\mu_1)^2, \tag{49}$$

with $H_T(\mu_2,\mu_1)$ being bounded in probability. But (49) holds true because for all j,

$$g_j(\mu_2)-g_j(\mu_1)=\frac{1}{M}\sum_{p=[M\mu_1]+1}^{[M\mu_2]}e^{-ij\lambda_{2mp}}=e^{i[M\mu_1]\lambda_{2mp}}g_j(\mu_2-\mu_1) \tag{50}$$

and

$$|\mathrm{Re}(g_j(\upsilon))|^2\le K\frac{\sin^2(2\pi j\upsilon)}{j^2}\le K\frac{\upsilon}{j}$$

by (44) and Taylor expansion. The last displayed inequality and (50) imply that

$$\begin{aligned}&E^*\left|a_k^*(-j)\,\mathrm{Re}(g_j(\mu_2)-g_j(\mu_1))\right|^4\\ \le\;& K\left\{\frac{1}{T^2}\sum_{q=1-[T/2]}^{[T/2]}\left|h_{xx,q}^{-1}\widehat{h}_{\widehat{u}\widehat{u},q}^{1/2}w_{x,q}\right|^4+\left(\frac{1}{T}\sum_{q=1-[T/2]}^{[T/2]}\left|h_{xx,q}^{-1}\widehat{h}_{\widehat{u}\widehat{u},q}^{1/2}w_{x,q}\right|^2\right)^2\right\}(\mu_2-\mu_1)^2.\end{aligned}$$

Then, denoting by $H_T(\mu_2,\mu_1)$ the term inside the braces of the last displayed inequality, we conclude (49), since proceeding as in the proof of Theorem 1 of Hidalgo (2000), it is obvious that $H_T(\mu_2,\mu_1)=K(1+o_p(1))$. That concludes the proof of (45).

Finally (46). First, observe that

$$\begin{aligned}&E^*\left[\sum_{q_1,q_2=0}^{M-2}a_k^*(-q_1)\,g_{q_1}(\mu)\,a_k^*(-q_2)\,g_{q_2}(\upsilon)\right]\\ =\;&\sum_{q_1,q_2=0}^{M-2}\left(\frac{1}{T}\sum_{s=1-[T/2]}^{[T/2]}f_{xx,s}^{-1}f_{uu,s}\frac{I_{xx,s}}{f_{xx,s}}e^{-i(q_1-q_2)\lambda_s}\right)\frac{1}{M^2}\sum_{j_1=1}^{[M\mu]}\sum_{j_2=1}^{[M\upsilon]}e^{iq_1\lambda_{j_1}}e^{-iq_2\lambda_{j_2}}\\ =\;&\frac{1}{T}\sum_{s=1-[T/2]}^{[T/2]}f_{xx,s}^{-1}f_{uu,s}\frac{I_{xx,s}}{f_{xx,s}}\frac{1}{M^2}\sum_{j_1=1}^{[M\mu]}\sum_{j_2=1}^{[M\upsilon]}\sum_{q_1,q_2=0}^{M-2}e^{i(j_1-s)\lambda_{q_1}}e^{-i(j_2-s)\lambda_{q_2}},\end{aligned} \tag{51}$$

whose real part is

$$\frac{1}{T}\sum_{s=1-[T/2]}^{[T/2]} f_{xx,s}^{-1} f_{uu,s} \frac{I_{xx,s}}{f_{xx,s}} \left(\mathcal{I}\left(j_1 = j_2 = s\right) + O\left(\frac{1}{M}\right)\right) \tag{52}$$

because

$$\operatorname{Re}\left(\frac{1}{M^2}\sum_{j_1=1}^{[M\mu_1]}\sum_{j_2=1}^{[M\mu_2]}\sum_{q_1,q_2=0}^{M-2} e^{i(s-j_1)\lambda_{q_1}} e^{-i(s-j_2)\lambda_{q_2}}\right) = \mathcal{I}\left(j_1 = j_2 = s\right) + O\left(\frac{1}{M}\right),$$

observing that the sum in q is from $q = 0$ to $M - 2$. Then, since from the proof of (47), the sum from $q = b$ to $M - 2$ is negligible for b large enough, it implies that using (51) – (52), the left side of (46) is

$$\begin{aligned} & \frac{1}{T}\sum_{s=1}^{\min\{[M\mu],[Mv]\}} f_{xx,s}^{-1} f_{uu,s}\left(\frac{I_{xx,s}}{f_{xx,s}} - 1\right)\{1 + o(1)\} + \frac{1}{T}\sum_{s=1}^{\min\{[M\mu],[Mv]\}} f_{xx,s}^{-1} f_{uu,s} \\ = & \int_0^{\min\{\mu,v\}} f_{xx}^{-1}(\lambda) f_{uu}(\lambda)\, d\lambda \left(1 + o_p(1)\right), \end{aligned}$$

by Markov's inequality since Theorem 2 of Robinson (1995) implies that $E\left|f_{xx,s}^{-1} I_{xx,s} - 1\right| = O\left(s^{-1}\log s\right)$. This completes the proof that $T^{1/2} S_T^*(\mu)$ converges in bootstrap weakly to the Gaussian process $\widetilde{B}(\mu)$. From here the conclusion of the theorem is standard by the continuous mapping theorem. □

References

[1] An, H.-Z., Chen, Z.-G. and Hannan, E.J., 1983. The maximum of the periodogram, Journal of Multivariate Analysis, 13, 383-400.

[2] Billingsley, P., 1968. Convergence of probability measures. (John Wiley, New York).

[3] Brillinger, D.R., 1981. Time series, data analysis and theory, (Holden-Day, San Francisco).

[4] Carlstein, E., 1986. The use of subseries values for estimating the variance of a general statistic from a stationary sequence, Annals of Statistics, 14, 1171-1179.

[5] Dahlhaus, R. and Janas, D., 1996. A frequency domain bootstrap for ratio statistics in time series analysis, Annals of Statistics, 24, 1934-1963.

[6] Efron, B., 1979. Bootstrap methods: another look at the jackknife, Annals of Statistics, 7, 1-26.

[7] Franke, J. and Härdle, W., 1992. On bootstrapping kernel spectral estimates, Annals of Statistics, 20, 121-145.

[8] Hall, P., Horowitz, J.L. and Jing, B.-Y., 1996. On blocking rules for the bootstrap with dependent data, Biometrika, 50, 561-574.

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

[10] Geweke, J., 1982. Measurement for linear dependence and feedback between multiple time series, Journal of the American Statistical Association, 77, 303-323.

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

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

[13] Granger, C.W.J., 1969. Investigating causal relations by econometric models and cross-spectral methods, Econometrica, 37, 424-438.

[14] Granger, C.W.J. and Joyeux, R., 1980. An introduction to long memory time series and fractional differencing, Journal of Time Series Analysis, 1, 15-30.

[15] Hannan, E.J., 1963. Regression for time series. In M. Rosenblatt (ed.), Time Series Analysis, 17-37. (John Wiley, New York).

[16] Hannan, E.J., 1965. The estimation of relations involving distributed lags, Econometrica, 33, 206-224.

[17] Hannan, E.J., 1967. The estimation of a lagged regression relation, Biometrika, 54, 409-418.

[18] Hannan, E.J., 1970. Multiple Time Series. (Wiley and Sons, New York).

[19] Hannan, E.J. and Terrell, R.D., 1973. Multiple equation systems with stationary errors, Econometrica, 41, 299-320.

[20] Hidalgo, J., 2000. Nonparametric test for causality with long-range dependence, Econometrica, 68, 1465-1490.

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

[22] Hosoya, Y., 1977. On the Granger condition for non-causality, Econometrica, 45, 1735-1736.

[23] Hosoya, Y., 1991. The decomposition and measurement of the interdependency between second-order stationary processes, Probability Theory and Related Fields, 88, 429-444.

[24] Karatzas, I. and Shreve, S.E., 1991. Brownian motion and stochastic calculus. (Springer Verlag, Berlin).

[25] Künsch, H.-R., 1989. The jackknife and the bootstrap for general stationary observations, Annals of Statistics, 17, 1217-1241.

[26] Loh, W.Y., 1988. Calibrating confidence coefficients, Journal of the American Statistical Association, 82, 155-162.

[27] Lütkepohl, H. and Poskitt, D.S., 1996. Testing for causation using infinite order vector autoregressive processes, Econometric Theory, 12, 61-87.

[28] Mammen, E., 1993. Bootstrap and wild bootstrap for high dimensional linear models, Annals of Statistics, 21, 255-285.

[29] Mandelbrot, B.B. and Van Ness, J.W., 1968. Fractional Brownian motions, fractional noises and applications, SIAM Review, 10, 422-437.

[30] Politis, D.N., Romano, J.P., 1994. Large sample confidence regions based on subsamples under minimal assumptions, Annals of Statistics, 22, 2031-2050.

[31] Robinson, P.M., 1979. Distributed lag approximation to linear-invariant systems, Annals of Statistics, 7, 507-515.

[32] Robinson, P.M., 1991. Automatic frequency domain inference on semiparametric and nonparametric models, Econometrica, 59, 1329-1363.

[33] Robinson, P.M., 1994. Time series with strong dependence. In C.A. Sims, ed., Advances in Econometrics: Sixth World Congress, Vol. 1, (Cambridge University Press, Cambridge). 47-95.

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

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

[36] Serfling, R.J., 1980. Approximation theorems of mathematical statistics. (John Wiley and Sons, New York).

[37] Sims, C.A., 1972. Money, income and causality, American Economic Review, 62, 540-552.

[38] Sims, C.A., 1974. Distributed lags, In Frontiers of Quantitative Economics, Vol. II (M.D. Intrilligator and D.A. Kendrick, eds.). (North Holland, Amsterdam).

[39] Sims, C.A., Stock, J.H. and Watson, M.W., 1990. Inference in linear time series models with some unit roots, Econometrica, 58, 113-144.

[40] Sinai, Y.G., 1976. Self-similar probability distribution, Theory of Probability and its Applications, 21, 64-80.

[41] Singh, K., 1981. On the asymptotic accuracy of Efron's bootstrap, Annals of Statistics, 9, 1187-1195.

[42] Stout, W.F., 1974. Almost Sure Convergence, (Academic Press, New York).

[43] Toda, H.Y. and Phillips, P.C.B., 1993. Vector autoregression and causality, Econometrica, 61, 1367-1393.