A BOOTSTRAP CAUSALITY TEST FOR COVARIANCE STATIONARY PROCESSES*

J. Hidalgo†
Economics Department
London School of Economics
Houghton Street
London WC2A 2AE
U.K.

January 20, 2000

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 $vec\left(\widetilde{B}(\mu)\right)$ indexed by $\mu \in [0,1]$. Since, in contrast to the scalar situation, it is not easy, if at all possible, to find a time deformation $g(\mu)$ such that $vec\left(\widetilde{B}(g(\mu))\right)$ is a vector with independent Brownian motions components, inferences based on $vec\left(\widetilde{B}(\mu)\right)$ would be difficult to implement. We thus propose, to circumvent this problem, to bootstrap the test by a combination of Wild and Moving Block Bootstraps, showing its validity.

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 (1999) 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)

***Key words and phrases**. Bootstrap, Causality test, Spectral Analysis.

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

it is nonparametric, 2) it is consistent, and 3) it has power against $T^{-1/2}$ contiguous alternatives. Thus, Hidalgo (1999) extended previous work in two main directions. First, by allowing a (general) covariance stationary linear process and secondly, since the test is nonparametric, he avoided the danger that a bad specification of the model may induce on the outcome of the test.

Now we briefly 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

$$A(L)\, w_t = \sum_{j=0}^{\infty} A_j w_{t-j} = \varepsilon_t, \quad t = 1, 2..., T, \tag{1.1}$$

where ε_t is a $p-$dimensional martingale difference sequence (more precise conditions on ε_t are given in $C4$ below) and A_0 is the identity matrix. The objective is to test the null hypothesis H_0: $y_t \nRightarrow x_t$, that is y_t does not cause x_t, against the alternative hypothesis $H_1 : y_t \Longrightarrow x_t$, that is 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

$$y_t = \sum_{j=-\infty}^{\infty} c(j)\, x_{t-j} + u_t, \tag{1.2}$$

where, by construction, $E[u_t \,|x_s, -\infty < s < \infty] = 0$, and x_t and u_t are, possibly, long-range dependent processes. Alternatively, H_0 is equivalent to

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

where $\underset{\sim}{0}$ is the $p_1 p_2-$dimensional null vector, or

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

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

$$H_0 : S^*(\mu) = 0 \quad \forall \mu \in [0,1] \text{ against } H_1 : S^*(\mu) \neq 0 \text{ in } \Delta \subset [0,1] \tag{1.3}$$

where Δ has Lebesgue measure greater than zero.

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

$$S_T(\mu) = \mathrm{Re}\left(\frac{1}{M}\sum_{p=1}^{[M\mu]} vec\left(\sum_{j=-M+1}^{0} \widehat{c}(j-1)\, e^{-ij\lambda_{2mp}}\right)\right) \tag{1.4}$$

where $\lambda_\ell = 2\pi\ell/T$, $\ell = 0, \pm 1, ..., \pm [T/2]$, and $M = [T/4m]$ with $m = m(T)$ a number which increases slowly with T, that is $m^{-1} + mT^{-1} \to 0$. 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, using an extension of Hidalgo's (1999) Corollary 1, we show in Theorem 3.2 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

$$K(\mu_1, \mu_2) = \frac{1}{4\pi} \int_0^{\pi \min(\mu_1, \mu_2)} \left(f_{xx}^{-1}(-\lambda) \otimes f_{uu}(\lambda)\right) d\lambda. \tag{1.5}$$

However, as it can straightforwardly be observed from (1.5), 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 to find a time transformation, say $g(\mu)$, such that $vec\left(\widetilde{B}(g(\mu))\right)$ is a p_1p_2-dimensional vector of independent Brownian processes is not easy, if at all possible. Although, two situations where a transformation $g(\mu)$ can be found 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, these situations are exceptions rather than the rule.

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 is under consideration. Therefore, the main objective of the paper is to examine how to circumvent the problem by using a bootstrap approach to test H_0. In particular, we show that a combination of Wu's (1986) Wild or external Bootstrap and Künsch's (1989) Moving Block Bootstrap (MBB) is consistent. 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.

The remainder of the paper is as follows. In section 2 we describe 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 describe the bootstrap and we examine its asymptotic properties. In section 5 we give the proofs of our results in sections 3 and 4 which use some technical lemmas provided in section 6.

2. THE ESTIMATION OF $c(j)$

In this section we describe the estimation of the matrices $c(j)$ in (1.2) and discuss why it is more desirable than other approaches, such as the least squares (LSE) estimates, in the presence of long-range dependence. In the frequency domain, the lag structure given in (1.2) is described by the frequency response function $C(\lambda) = \sum_{j=-\infty}^{\infty} c(j) e^{-ij\lambda}$. So $c(j)$ is interpreted as the jth Fourier coefficient of $C(\lambda) = f_{yx}(\lambda) f_{xx}^{-1}(\lambda)$, that is,

$$c(j) = (2\pi)^{-1} \int_0^{2\pi} C(\lambda) e^{ij\lambda} d\lambda, \tag{2.1}$$

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

$$E\left((w_1 - Ew_1)(w_{j+1} - Ew_1)'\right) = \int_{-\pi}^{\pi} f_{ww}(\lambda) e^{-ij\lambda} d\lambda \quad j = 0, \pm1, \pm2, ...$$

Due to the interpretation of $c(j)$ in (2.1), Hannan (1963, 1967), see also Brillinger (1981), proposed to estimate $c(j)$ by the sample (discrete) analogue of (2.1),

$$\widetilde{c}(j) = \frac{1}{2M} \sum_{p=0}^{2M-1} \widehat{C}_{2mp} e^{ij\lambda_{2mp}}, \tag{2.2}$$

where $\widehat{C}_{2mp} = \widehat{f}_{yx,2mp} \widehat{f}^{-1}_{xx,2mp}$, 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 (2.3) below, and where henceforth we have abbreviated $g(\lambda_p)$ by g_p for a generic function $g(\lambda)$. The estimator $\widetilde{c}(j)$ in (2.2) was coined by Sims (1974) as the HI (Hannan's inefficient) estimator.

We estimate f_{ww} by

$$\widehat{f}_{ww}(\lambda) = \frac{1}{2m+1} \sum_{j=-m}^{m} I_{ww}(\lambda_j + \lambda), \tag{2.3}$$

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 Section 1.

When analyzing the HI estimator in (2.2), 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 may be infinity, as in Hidalgo (1999) we modify (2.2) by

$$\widehat{c}(j) = \frac{1}{2M} \sum_{p=1}^{2M-1\,'} \widehat{C}_{2mp} e^{ij\lambda_{2mp}}, \tag{2.4}$$

where $\sum_{p=1}^{2M-1\,'} \phi_{2mp} e^{ij\lambda_{2mp}}$ 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 $\widehat{f}_{yx,2m} \widehat{f}^{-1}_{xx,2m}$.

The motivation of the estimator in (2.4) is threefold. First, the statistical properties hold the same irrespective of the number of lags specified in (1.2), which have important consequences when analyzing the properties of $S_T(\mu)$ defined in (1.4). 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 (2.4) more appealing when the data may exhibit long-range dependence, is as follows. Assume, for expositional simplicity, that $p_1 = p_2 = 1$ and that model (1.2) is

$$y_t = \sum_{j=-r}^{q} c(j) x_{t-j} + u_t, \quad t = 1, ..., T, \tag{2.5}$$

where both q and r are finite and known a priory. 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 and 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 (2.5). More generally, their results are also valid when $c(j)$ is known up to a set of parameters θ, that is $c(j) = c(j;\theta)$ for all j, in (1.2). 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 $\widehat{f}_{xx,p}^{-1}$ becomes (asymptotically) a weighted function satisfying the conditions of Robinson and Hidalgo (1997). In fact, in Theorem 3.1 below, we show that the HI estimator in (2.4) achieves the root-T consistency and asymptotic normality, so that the HI estimator is indeed a desirable estimator.

3. THE ASYMPTOTIC PROPERTIES OF (2.4) AND (1.4)

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 from the relation

$$E\left((z_1 - Ez_1)(z_{j+1} - Ez_1)'\right) = \int_{-\pi}^{\pi} f_{zz}(\lambda)\, e^{-ij\lambda} d\lambda \quad j = 0, \pm 1, \pm 2, ...$$

Let us introduce the following 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

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

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

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

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

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

where Φ_0 and Φ_0^u are the identity matrices and $\|D\|$ stands for the norm of the matrix D. Finally $\{x_t\}$ and $\{u_t\}$ are mutually independent.

Condition C4 $\{\varepsilon_t\}$ is a stochastic process that satisfies $E(\varepsilon_t|\mathcal{F}_{t-1}) = 0$, $E(\varepsilon_t\varepsilon_t'|\mathcal{F}_{t-1}) = E(\varepsilon_t\varepsilon_t') = \Xi$ a.s., $E(\varepsilon_{tj_1}\varepsilon_{tj_2}\varepsilon_{tj_3}|\mathcal{F}_{t-1}) = \mu_{3,j_1,j_2,j_3}$ such that $|\mu_{3,j_1,j_2,j_3}| < \infty$ for all j_1, j_2, j_3 where $\mathcal{F}_t$ is the σ-algebra of events generated by $\varepsilon_s, s \leq t$, and the joint fourth cumulant of $\varepsilon_{t_ij_i}$, $j_i = 1, ..., p$ and $i = 1, ..., 4$ satisfies

$$cum(\varepsilon_{t_1j_1}, \varepsilon_{t_2j_2}, \varepsilon_{t_3j_3}, \varepsilon_{t_4j_4}) = \begin{cases} \kappa_{j_1,j_2,j_3,j_4}, & t_1 = t_2 = t_3 = t_4, \\ 0, & otherwise, \end{cases}$$

with $\kappa = \max_{j_i=1,...,p,i=1,...,4} |\kappa_{j_1,j_2,j_3,j_4}| < \infty$.

Condition C5 $\{\varepsilon_{u,t}\}$ is a stochastic process that satisfies $E(\varepsilon_{u,t}|\mathcal{F}_{t-1}) = 0$, $E(\varepsilon_{u,t}\varepsilon_{u,t}'|\mathcal{F}_{t-1}) = E(\varepsilon_{u,t}\varepsilon_{u,t}') = \Xi_u$ a.s., $E(\varepsilon_{u,tj_1}\varepsilon_{u,tj_2}\varepsilon_{u,tj_3}|\mathcal{F}_{t-1}) = \mu_{u,3,j_1,j_2,j_3}$ such that $|\mu_{u,3,j_1,j_2,j_3}| < \infty$ for all j_1, j_2, j_3 where $\mathcal{F}_t$ is the σ-algebra of events generated by $\varepsilon_{u,s}, s \leq t$, and the joint fourth cumulant of ε_{u,t_ij_i}, $j_i = 1, ..., p_1$ and $i = 1, ..., 4$ satisfies

$$cum(\varepsilon_{u,t_1j_1}, \varepsilon_{u,t_2j_2}, \varepsilon_{u,t_3j_3}, \varepsilon_{u,t_4j_4}) = \begin{cases} \kappa_{u,j_1,j_2,j_3,j_4}, & t_1 = t_2 = t_3 = t_4, \\ 0, & otherwise, \end{cases}$$

with $\kappa_u = \max_{j_i=1,...,p_1,i=1,...,4} |\kappa_{u,j_1,j_2,j_3,j_4}| < \infty$.

Condition C6 $\|(\partial/\partial\lambda)\Phi(\lambda)\| = O(\|\Phi(\lambda)\|/\lambda)$ 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 C7 $\|c(|j|)\| = O\left(|j|^{-3+\tau/2}\right)$ for some $0 < \tau < 1$.

Condition C8 $M^2T^{-1} + M^{\tau-4}T \to 0$ with τ as in C7.

Conditions C1-C2 deal with the behaviour of $f_{zz}(\lambda)$ and they are the same as in Robinson ($1995a, b$), so his comments apply here. Conditions C3-C5 are restrictive in the linearity they impose, but not otherwise. The requirement of independence between x_t and u_t in C3, as in Robinson and Hidalgo (1997), is necessary for the proof of asymptotic normality of (2.4). 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 C6 is similar to that in Robinson ($1995a, b$). The second part of the condition is not strong, see for instance the comments made after (3.1) below, once λ^{d_g} is identified as $f_{gg}^{-1/2}$ up to constants there. Condition C7 implies that the first derivative of $\|C(\lambda)\|$ is Liptchitz continuous with Liptchitz parameter in the interval $(0, 1 - \tau/2)$. Condition C8 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-C6 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

$$w_t = \sum_{j=0}^{\infty} G(j)\, \xi_{t-j}. \tag{3.1}$$

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-C6 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 $\sigma_w^2 = E\left(w_t - E(w_t)\right)^2 < \infty$ and $\Gamma(\cdot)$ denotes the gamma function.

Theorem 3.1. *Assuming C1-C8, for any finite collection $j_1, ..., j_q$,*

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

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

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

(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

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

The asymptotic properties of the estimator given in (2.2) were first 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 bounded and bounded away from zero, whereas Hidalgo (1999) under, possibly, the presence of long-range dependence did the same for that in (2.4).

Let us introduce an additional condition.

Condition C9 $\|x_t\|^4$ and $\|u_t\|^4$ are uniformly integrable.

Observe that a sufficient condition for Condition $C9$ is that for some $\delta > 0$, $\sup_t \left(E \left\| x_t \right\|^{4+\delta} + E \left\| u_t \right\|^{4+\delta} \right) < \infty$.

Let

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

Theorem 3.2. *Assuming C1-C9 and $\widehat{c}(j)$ given in (2.4), 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 metric, where $vec\left(\widetilde{B}(\mu)\right)$ is a $p_1 p_2$-Gaussian process with covariance structure given in (1.5).

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 (1.5) 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

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

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, see for example Koul and Stute (1999), 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 $g(\mu)$ such that $vec\left(\widetilde{B}(\mu)\right)$ admits the representation $vec(B(g(\mu)))$, where $B(\mu)$ has independent Brownian motions components, is not generally available. Two situations where this transformation is possible were described in the introduction. Namely, 1) when $K(\mu_1,\mu_2) = \min(\mu_1,\mu_2)\,\Omega$ where Ω is 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. However, these two situations are exceptions rather than the rule. Thus, the results of Theorem 3.2 are somehow of limited use for the purpose of statistical inference. Although, in principle, the limiting distribution could be simulated, since it is non-distribution free, it would imply that a practictioner would need to compute critical values everytime a new model is under consideration. We propose to bootstrap $T^{1/2}S_T(\mu)$ to circumvent the potential problem of how to implement the test in empirical examples.

4. A BOOTSTRAP APPROACH TO $T^{1/2}S_T(\mu)$

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 three main reasons. First, bootstrap methods are capable of approximating the finite sample distribution of statistics better than those based on their asymptotic counterparts. Second, it is its ability to work in complex situations without imposing strong assumptions on the data mechanism process. And third, 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.

In the present paper we address the latter situation. Following our comments at the end of the previous section, the aim of this section is to propose a bootstrap procedure for $T^{1/2}S_T(\mu)$ based on a combination of Wu's (1986) Wild or External bootstrap and Künsch's (1989) Moving Block Bootstrap (MBB).

We now describe the bootstrap. Let $\widehat{c}_T(j)$ denote the estimator of $c(j)$ in (2.4) and let $n = n(T)$ be a number which increases slowly with T. Consider $L = T - n + 1$ groups of size n, where the ℓth group has observations $\underset{\sim}{w}(\ell) = \left(w'_\ell, ..., w'_{n+\ell-1}\right)'$.

Let us introduce the periodogram of $\underset{\sim}{w}(\ell)$, for $\ell = 1, ..., L$,

$$I_{ww}(\lambda, \ell) = h_w(\lambda, \ell)\, h_w^*(\lambda, \ell),$$

where

$$h_w(\lambda, \ell) = \frac{1}{(2\pi n)^{1/2}} \sum_{t=\ell}^{\ell+n-1} w_t e^{it\lambda}, \tag{4.1}$$

is the discrete Fourier transform and "$*$" means transpose combined with complex conjugation. Then, the spectral density matrix estimator of f_{ww} using $\underset{\sim}{w}(\ell)$, for $\ell = 1, ..., L$, is defined by

$$\widehat{f}_{ww}(\lambda, \ell) = \frac{1}{2m+1} \sum_{j=-m}^{m} I_{ww}(\lambda_j + \lambda, \ell), \tag{4.2}$$

where $m = [n/4M(n)]$ with $M(n)$ a positive number such that $M(n)^{-1} + n^{-1}M(n) \to 0$ and $\lambda_j = (2\pi j)/n$, $j = 0, \pm 1, ..., \pm[n/2]$.

The bootstrap consists of three steps.

STEP 1 Compute $\widehat{c}_n(j; \ell)$ as in (2.4) using $\underset{\sim}{w}(\ell)$, for $\ell = 1, ..., L$. That is,

$$\widehat{c}_n(j; \ell) = \frac{1}{2M(n)} \sum_{p=1}^{2M(n)-1'} \widehat{C}_{2mp}(\ell)\, e^{ij\lambda_{2mp}}, \tag{4.3}$$

where $\widehat{C}_{2mp}(\ell) = \widehat{f}_{yx,2mp}(\ell)\widehat{f}^{-1}_{xx,2mp}(\ell)$, and $\widehat{f}_{yx,2mp}(\ell)$ and $\widehat{f}_{xx,2mp}(\ell)$ denote the estimates of $f_{yx,2mp}$ and $f_{xx,2mp}$ respectively given by the indicated components of (4.2), and where we have abbreviated $g(\lambda_{2mp};\ell)$ by $g_{2mp}(\ell)$ for a generic function $g(\cdot)$.

STEP 2 For all $\ell = 1,...,L$, let us compute

$$\vartheta(\mu,\ell) = \mathrm{Re}\left(\frac{n^{1/2}}{M(n)}\sum_{p=1}^{[M(n)\mu]} vec\left(\sum_{j=-M(n)+1}^{0}\left(\widehat{c}_n(j-1;\ell) - \widehat{c}_T(j-1)\right)e^{-ij\lambda_{2mp}}\right)\right). \tag{4.4}$$

And finally,

STEP 3 Compute the bootstrap statistic as

$$S^*_T(\mu) = \frac{1}{L^{1/2}}\sum_{\ell=1}^{L}\vartheta(\mu,\ell)\eta_\ell, \quad \mu\in[0,1] \tag{4.5}$$

where η_ℓ is an $iid(0,1)$ sequence of bounded random variables.

Remark 1. *It is worth noting that we have used the same random variable η_ℓ for every coordinate of $vec\left(\sum_{j=-M(n)+1}^{0}\left(\widehat{c}_n(j-1;\ell) - \widehat{c}_T(j-1;\ell)\right)e^{-ij\lambda_{2mp}}\right)$. This is crucial for the validity of the bootstrap. The reason comes from the observation that if different random variables η_ℓ were used for each coordinate, unless $vec\left(\widetilde{B}(\mu)\right)$ had independent components, the latter bootstrap approach would destroy the covariance structure of $vec\left(\widetilde{B}(\mu)\right)$ and thus the validity of that bootstrap. In particular, for $\ell = 1,...,L$, let $\eta^*_\ell = (\eta_{\ell,1},...,\eta_{\ell,p})'$ an iid sequence with mutually independent components, and consider*

$$S^{**}_T(\mu) = \frac{1}{L}\sum_{\ell=1}^{L} S_n(\mu,\ell)\circledast\eta^*_\ell, \quad \mu\in[0,1]$$

where for two vectors a and b, $a\circledast b$ denotes multiplication of a and b coordinate by coordinate. Then, proceeding as in the proof of Proposition 6.1, it is easily shown that

$$\begin{aligned} & E\left(L\left(S^{**}_T(\mu)S^{**}_T(\mu)'\right)\Big|(w'_1,...,w'_T)'\right) \\ &= \frac{1}{L}\sum_{\ell=1}^{L}E\left((S_n(\mu,\ell)\circledast\eta^*_\ell)(S_n(\mu,\ell)\circledast\eta^*_\ell)'\Big|(w'_1,...,w'_T)'\right) \\ &= \frac{1}{L}\sum_{\ell=1}^{L}diag\left(S^2_n(\mu,\ell)_1,...,S^2_n(\mu,\ell)_{p_1p_2}\right)\nrightarrow K(\mu,\mu), \end{aligned}$$

where g_q denotes the qth component of a vector g. Moreover, proceeding as in the proof of Theorem 4.1 below, $L^{1/2}S^{**}_T(\mu)\overset{weakly}{\Longrightarrow} vec\left(\overline{B}(\mu)\right)$, where $vec\left(\overline{B}(\mu)\right)$ has independent components, which implies that $vec\left(\overline{B}(\mu)\right)\neq vec\left(\widetilde{B}(\mu)\right)$.

The resampling method must be such that the conditional distribution, given the data, of the bootstrap test, say $\varphi(S_T^*(\mu))$, consistently estimates the distribution of $\varphi\left(vec\left(\widetilde{B}(\mu)\right)\right)$ under H_0. That is, $\varphi(S_T^*(\mu)) \to_{d^*} \varphi\left(vec\left(\widetilde{B}(\mu)\right)\right)$ in probability under H_0, where "$\to_{d^*}$" denotes

$$\lim_{n\to\infty} \Pr\left[\varphi(S_T^*(\mu)) \le z | \underset{\sim}{x}\right] \xrightarrow{p} G(z),$$

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 Gine and Zinn (1990). Moreover, under contiguous alternatives H_a, $\varphi(S_T^*(\mu))$ must also converge, in bootstrap distribution to $\varphi\left(vec\left(\widetilde{B}(\mu)\right)\right)$, whereas under the alternative H_1 $\varphi(S_T^*(\mu))$ should be bounded in probability.

Before we study the statistical properties of the bootstrap statistic $S_T^*(\mu)$ given in (**??**), let us introduce an additional condition.

C.10 $n^{-1} + T^{-1} n \log^2 T + n^{-1} M^2(n) \to 0.$

Theorem 4.1. *let $\varphi(\cdot)$ be a continuous functional. Assuming C.1-C.10, under the maintained hypothesis $H = H_0 \cup H_1$,*

$$\varphi\left(L^{1/2} S_T^*(\mu)\right) \xrightarrow{d^*} \varphi\left(vec\left(\widetilde{B}(\mu)\right)\right) \text{ in probabilty,}$$

where $vec\left(\widetilde{B}(\mu)\right)$ is a $p_1 p_2$-Gaussian process with covariance structure given in (1.5).

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. In addition, it also indicates that the bootstrap statistic given in (4.5) 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. Then, Theorem 3.2 and the continuous mapping theorem indicate that $c^f_{n,(1-\alpha)} \to c^a_{(1-\alpha)}$, whereas Theorem 4.1 indicates that $c^*_{(1-\alpha)}$ satisfy $c^*_{(1-\alpha)} \xrightarrow{p} c^a_{(1-\alpha)}$ where

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

Since the finite sample distribution of $\varphi\left(L^{1/2} S_T^*(\mu)\right)$ is not available, $c^*_{(1-\alpha)}$ is approximated, as accurately as desired, by a standard Monte-Carlo simulation algorithm. That is, let $\eta^{(j)} = \left(\eta_1^{(j)}, ..., \eta_L^{(j)}\right)$ for $j = 1, ..., B$, and for each j, compute $S_T^{*(j)}$ as in *STEP 3*. Then, $c^*_{(1-\alpha)}$ is approximated by the value $c^{*B}_{(1-\alpha)}$ that satisfies

$$\frac{1}{B} \sum_{j=1}^{B} 1\left(\left|\varphi\left(L^{1/2} S_T^{*(j)}(\mu)\right)\right| \ge c^{*B}_{(1-\alpha)}\right) = \alpha.$$

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

$$\begin{aligned} & \operatorname{Re}\left(\sum_{j=0}^{M-1} vec\left(T^{1/2}\widehat{c}(-j-1)\frac{1}{M}\sum_{p=1}^{[M\mu]} e^{ij\lambda_{2mp}}\right)\right) \qquad (5.1) \\ = \quad & vec\left(T^{1/2}\widehat{c}(-1)\mu + \sum_{j=1}^{M-1} T^{1/2}\widehat{c}(-j-1)\frac{\sin(\pi\mu j)}{\pi j}\right)(1+o(1)) \end{aligned}$$

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

Writing

$$a_j - c(j) = \frac{1}{2M}\sum_{p=1}^{2M-1} \widehat{f}_{ux,2mp} f_{xx,-2mp}^{-1} e^{ij\lambda_{2mp}}, \quad j = 0, \pm, ..., \pm M,$$

let $\widehat{c}(j) = a_j + H(j)$. So the right side of (5.1) is

$$\begin{aligned} & vec\left(T^{1/2}a_{-1}\mu + \sum_{j=1}^{M-1} T^{1/2}a_{-j-1}\frac{\sin(\pi\mu j)}{\pi j}\right)(1+o(1)) \qquad (5.2) \\ & + vec\left(T^{1/2}H(-1)\mu + \sum_{j=1}^{M-1} T^{1/2}H(-j-1)\frac{\sin(\pi\mu j)}{\pi j}\right)(1+o(1)). \end{aligned}$$

The second term of (5.2) is $o_p(1)$ uniformly in μ, as we now show. From the proof of Theorem 1 of Hidalgo (1999), uniformly in j, $T^{1/2}H(j) = O_p\left(M^{-1/2}\log M\right)$. So, by the triangle inequality

$$\begin{aligned} & \sup_{\mu\in[0,1]}\left\| T^{1/2}H(-1) + \sum_{j=1}^{M-1} T^{1/2}H(-j-1)\frac{\sin(\pi\mu j)}{\pi j}\right\| \qquad (5.3) \\ \leq \quad & O_p\left(\frac{\log M}{M^{1/2}}\right) + K\sum_{j=1}^{M-1}\frac{1}{j}\left\|T^{1/2}H(-j-1)\right\| = O_p\left(M^{-1/2}\log^2 M\right). \end{aligned}$$

So, to complete the proof it suffices to show that the first term of (5.2), that is,

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

converges weakly to $vec\left(\widetilde{B}(\mu)\right)$. But, since under H_0, $c(j) = 0$ for all $j < 0$, the proof follows proceeding as that of Hidalgo's (1999) Corollary 1, and thus 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 $S_T^*(\mu)$ converges in probability to (1.5). In Proposition 5.2 we show that the finite dimensional distributions converge to those of $vec\left(\widetilde{B}(\mu)\right)$, whereas in Proposition 5.3 we show the tightness condition. Then apply the continuous mapping theorem to conclude.

Henceforth, we shall denote $E^*(\cdot)$ as the bootstrap expectation, that is, for any random variable Y, $E^*(Y) = E(Y|w_1, ..., w_T)$.

Proposition 5.1. *Under C1-C10, for all* $0 \leq \mu_1 \leq \mu_2 \leq 1$,

$$E^*\left(\frac{1}{L^{1/2}}\sum_{\ell=1}^{L}\vartheta(\mu_1,\ell)\eta_\ell\frac{1}{L^{1/2}}\sum_{\ell=1}^{L}\vartheta(\mu_2,\ell)'\eta_\ell\right)\overset{P}{\to}K(\mu_1,\mu_2). \tag{5.4}$$

Proof. Because η_ℓ is a sequence of $iid(0,1)$ random variables, the left side of (5.4) is

$$\frac{1}{L}\sum_{\ell=1}^{L}\vartheta(\mu_1,\ell)\vartheta(\mu_2,\ell)'. \tag{5.5}$$

So, it suffices to show that (5.5) converges in probability to $K(\mu_1, \mu_2)$.

First, proceeding as in the proof of Hidalgo's (1999) Theorem 1,

$$\begin{aligned}(\widehat{c}_n(-j-1;\ell)-\widehat{c}_T(-j-1)) &= \widehat{a}_n(-j-1;\ell)-\widehat{a}_T(-j-1) \\ &\quad +(H_n(-j-1;\ell)-H_T(-j-1)),\end{aligned} \tag{5.6}$$

where

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

$$\widehat{a}_T(-j-1)-c(-j-1)=\frac{1}{2M}\sum_{p=1}^{2M-1}\widehat{f}_{ux}(\pi p/M)f_{xx}^{-1}(\pi p/M)e^{ij\pi p/M}$$

and such that uniformly j,

$$n^{1/2}(H_n(-j-1;\ell)-H_T(-j-1))=O_p\left(M^{-1/2}(n)\log M(n)\right). \tag{5.7}$$

Using (5.6), (5.7) and the convention that $x^{-1}\sin(\mu x)=\mu$ for $x=0$, since uniformly in $\mu\in[0,1]$, $\mathrm{Re}\left(M^{-1}(n)\sum_{p=1}^{[M(n)\mu]}e^{ij\lambda_{2mp}}\right)\underset{M(n)\to\infty}{\to}(\pi j)^{-1}\sin(\pi j\mu)$, we conclude that the right side of (4.4), proceeding as in the proof of the second term of (5.2), is

$$\begin{aligned}&\sum_{j=0}^{M(n)-1}n^{1/2}vec(\widehat{a}_n(-j-1;\ell)-\widehat{a}_T(-j-1))\frac{\sin(\pi j\mu)}{\pi j}+o_p(1)\\ =&\sum_{j=0}^{M(n)-1}n^{1/2}vec(\widehat{a}_n(-j-1;\ell)-c(-j-1))\frac{\sin(\pi j\mu)}{\pi j}+o_p(1) \qquad (5.8)\\ &-(n/T)^{1/2}\sum_{j=0}^{M(n)-1}T^{1/2}vec(\widehat{a}_T(-j-1)-c(-j-1))\frac{\sin(\pi j\mu)}{\pi j}+o_p(1).\end{aligned}$$

The second term on the right of (5.8) is $o_p(1)$ uniformly in μ as we now show. First, proceeding as in the proof of Theorem 1 of Hidalgo (1999) and the uniform integrability of $\|x_t\|^4$ and $\|u_t\|^4$ we get that $E\left\|T^{1/2}\left(\widehat{a}_T(-j-1)-c(-j-1)\right)\right\| < D$ by Theorem A of Serfling (1981, p.32), where henceforth, D denotes a finite positive constant. So the first absolute moment of that term is bounded by

$$D\frac{n^{1/2}}{T^{1/2}}\sum_{j=0}^{M(n)-1} j^{-1} \leq D\frac{n^{1/2}}{T^{1/2}}\log M(n) = o(1)$$

by $C10$. On the other hand, proceeding as in the proof of Corollary 1 of Hidalgo (1999), the first term on the right of (5.8) is

$$\begin{aligned}
&\sum_{j=0}^{k-1} n^{1/2} vec\left(\widehat{a}_n(-j-1;\ell)-c(-j-1)\right)\frac{\sin(\pi j\mu)}{\pi j} \\
&\quad+\sum_{j=k}^{M(n)-1} n^{1/2} vec\left(\widehat{a}_n(-j-1;\ell)-c(-j-1)\right)\frac{\sin(\pi j\mu)}{\pi j} \\
=&\sum_{j=0}^{k-1} n^{1/2} vec\left(\widehat{a}_n(-j-1;\ell)-c(-j-1)\right)\frac{\sin(\pi j\mu)}{\pi j}+o_p(1)
\end{aligned}$$

where the $o_p(1)$ is uniform in μ. (In fact, it was shown there that

$$E\sup_{\mu}\left|\sum_{j=k}^{M(n)-1} n^{1/2} vec\left(\widehat{a}_n(-j-1;\ell)-c(-j-1)\right)_s\frac{\sin(\pi j\mu)}{\pi j}\right| = o(1),$$

for $s=1,...,p_1p_2$, where g_s denotes the sth coordinate of a vector g.)

So we conclude that (5.5) is

$$\frac{1}{L}\sum_{\ell=1}^{L} b_\ell(\mu_1)\, b_\ell(\mu_2)' + o_p(1), \tag{5.9}$$

where

$$b_\ell(\mu)=\sum_{j=0}^{k-1} n^{1/2} vec\left(\widehat{a}_n(-j-1;\ell)-c(-j-1)\right)\frac{\sin(\pi j\mu)}{\pi j}.$$

To complete the proof, it suffices to show that the first term of (5.9) converges in probability to $K(\mu_1,\mu_2)$. To that end, without loss of generality, we focus on one of the components of $b_\ell(\mu_1)\, b_\ell(\mu_2)'$, for example the $(1,1)\,th$.

Let $K_{1,1}(\mu_1,\mu_2)$ denote the $(1,1)\,th$ element of $K(\mu_1,\mu_2)$ and $b_{\ell,1}(\mu)$ the first element of $b_\ell(\mu)$. First, since $n^{1/2}\left(\widehat{a}_n(-j-1;\ell)-c(-j-1)\right)$ converges in distribution by Theorem 1 of Hidalgo (1999), and by C9 is a uniformly integrable sequence then $E\left\|n^{1/2} vec\left(\widehat{a}_n(-j-1;\ell)-c(-j-1)\right)\right\|^2 < D$ by Theorem A of Serfling (1981, p.32). Thus proceeding as in Corollary 1 of Hidalgo (1999),

$$E\left(b_{\ell,1}(\mu_1)\, b_{\ell,1}(\mu_2)\right) \to K_{1,1}(\mu_1,\mu_2),$$

which implies by stationarity of $(x_t', u_t')'$ that

$$E\left(\frac{1}{L}\sum_{\ell=1}^{L} b_{\ell,1}(\mu_1)\, b_{\ell,1}(\mu_2)\right) \to K_{1,1}(\mu_1,\mu_2).$$

Next,

$$E\left(\frac{1}{L}\sum_{\ell=1}^{L} b_{\ell,1}(\mu_1)\, b_{\ell,1}(\mu_2)\right)^2 = \frac{1}{L^2}\sum_{\ell=1}^{L} E\left(b_{\ell,1}^2(\mu_1)\, b_{\ell,1}^2(\mu_2)\right) \tag{5.10}$$
$$+\frac{2}{L^2}\sum_{\ell_1=1<\ell_2}^{L} E\left(b_{\ell_1,1}(\mu_1)\, b_{\ell_1,1}(\mu_2)\, b_{\ell_2,1}(\mu_1)\, b_{\ell_2,1}(\mu_2)\right).$$

The first term on the right of (5.10) is

$$\frac{1}{L^2}\sum_{\ell=1}^{L}\Big\{2\left(E\left(b_{\ell,1}(\mu_1)\, b_{\ell,1}(\mu_2)\right)\right)^2 + E\left(b_{\ell,1}^2(\mu_1)\right) E\left(b_{\ell,1}^2(\mu_2)\right)$$
$$+cum\left(b_{\ell,1}(\mu_1), b_{\ell,1}(\mu_1), b_{\ell,1}(\mu_2), b_{\ell,1}(\mu_2)\right)\Big\} = O\left(L^{-1}\right)$$

by Cauchy inequality, and because $\left|E\left(b_{\ell,1}^2(\mu)\right)\right| \le D$ and $\left|cum\left(b_{\ell,1}(\mu_1), b_{\ell,1}(\mu_1), b_{\ell,1}(\mu_2), b_{\ell,1}(\mu_2)\right)\right| \to 0$ since by the uniform integrability of $\|x_t\|^4$ and $\|u_t\|^4$ by C9 $E\left\|n^{1/2}\left(\widehat{a}_n(-j-1;\ell) - c(-j-1)\right)\right\|^4$ converges to that of its limiting distribution, (a normal random variable), by Theorem A of Serfling (1981, p.32).

To finish the proof of the proposition, it thus remains to show that the second term on the right of (5.10) converges to $K_{1,1}^2(\mu_1,\mu_2)$. But this term is

$$\frac{2}{L^2}\sum_{\ell_1=1<\ell_2}^{L}\Big\{E\left(b_{\ell_1,1}(\mu_1)\, b_{\ell_1,1}(\mu_2)\right) E\left(b_{\ell_2,1}(\mu_1)\, b_{\ell_2,1}(\mu_2)\right)$$
$$+E\left(b_{\ell_1,1}(\mu_1)\, b_{\ell_2,1}(\mu_2)\right) E\left(b_{\ell_1,1}(\mu_2)\, b_{\ell_2,1}(\mu_1)\right) \tag{5.11}$$
$$+E\left(b_{\ell_1,1}(\mu_1)\, b_{\ell_2,1}(\mu_1)\right) E\left(b_{\ell_1,1}(\mu_2)\, b_{\ell_2,1}(\mu_2)\right)$$
$$+cum\left(b_{\ell_1,1}(\mu_1), b_{\ell_2,1}(\mu_1), b_{\ell_1,1}(\mu_2), b_{\ell_2,1}(\mu_2)\right)\Big\}.$$

The first term of (5.11) is

$$\left(\frac{1}{L}\sum_{\ell=1}^{L} E\left(b_{\ell,1}(\mu_1)\, b_{\ell,1}(\mu_2)\right)\right)^2 - \frac{1}{L^2}\sum_{\ell=1}^{L}\left(E\left(b_{\ell,1}(\mu_1)\, b_{\ell,1}(\mu_2)\right)\right)^2 \to K_{1,1}^2(\mu_1,\mu_2),$$

since by stationarity of x_t and u_t,

$$\frac{1}{L}\sum_{\ell=1}^{L} E\left(b_{\ell,1}(\mu_1)\, b_{\ell,1}(\mu_2)\right) = E\left(b_{1,1}(\mu_1)\, b_{1,1}(\mu_2)\right) \to K_{1,1}(\mu_1,\mu_2)$$

proceeding as in the proof of Corollary 1 of Hidalgo (1999). The remaining three terms of (5.11) are $o(1)$ since by Lemmas 6.2 and 6.3 $\left|E\left(b_{\ell_1,1}(\mu_1)\, b_{\ell_2,1}(\mu_2)\right)\right| \to 0$ if $|\ell_1-\ell_2| > n$ and $cum\left(b_{\ell_1,1}(\mu_1), b_{\ell_2,1}(\mu_1), b_{\ell_1,1}(\mu_2), b_{\ell_2,1}(\mu_2)\right) \to 0$ for all ℓ_1 and

ℓ_2, using the same arguments as in the proof of the first term on the right of (5.10). Thus, the second moment of the first term of (5.9) converges to the square of its first moment, which implies that (5.9) converges in probability to $K_{1,1}(\mu_1, \mu_2)$. This concludes the proof of the proposition. ■

Proposition 5.2. *Under C1-C10, the finite dimensional distributions of $S_T^*(\mu)$ given in (4.5) converge to those of vec $\left(\widetilde{B}(\mu)\right)$.*

Proof. Fixed $\phi_1, ..., \phi_q$ and $\mu_1, ..., \mu_q$. By Cràmer-Wold device it suffices to examine the limiting distribution of

$$\sum_{p=1}^{q} \phi_p \frac{1}{L^{1/2}} \sum_{\ell=1}^{L} \xi' \vartheta(\mu_p, \ell) \eta_\ell = \frac{1}{L^{1/2}} \sum_{\ell=1}^{L} \left(\sum_{p=1}^{q} \phi_p \xi' \vartheta(\mu_p, \ell) \right) \eta_\ell,$$

where ξ is a $p_1 p_2-$dimensional finite vector. Because, conditional on the data, $\mathcal{W}_n(\ell) = \left(\sum_{p=1}^{q} \phi_p \xi' \vartheta(\mu_p, \ell) \right) \eta_\ell$ is a sequence of zero mean independent random variables, by Proposition 5.1

$$E^* \left(\frac{1}{L^{1/2}} \sum_{\ell=1}^{L} \left(\sum_{p=1}^{q} \phi_p \xi' \vartheta(\mu_p, \ell) \right) \eta_\ell \right)^2 = \frac{1}{L} \sum_{\ell=1}^{L} \left(\sum_{p=1}^{q} \phi_p \xi' \vartheta(\mu_p, \ell) \right)^2$$
$$\xrightarrow{P} \sum_{p_1, p_2=1}^{q} \phi_{p_1} \left(\xi' K(\mu_{p_1}, \mu_{p_2}) \xi \right) \phi_{p_2}.$$

So to complete the proof, we need to verify the Lindeberg's condition, that is, for all $\delta > 0$

$$\frac{1}{L} \sum_{\ell=1}^{L} E^* \left(\mathcal{W}_n^2(\ell) \mathcal{I}\left(|\mathcal{W}_n(\ell)| > \delta L^{1/2} \right) \right) \xrightarrow{P} 0. \tag{5.12}$$

Assuming that $|\eta_\ell| \leq D$, the left side of (5.12) is bounded by

$$\frac{D^2}{L} \sum_{\ell=1}^{L} \left| \sum_{p=1}^{q} \phi_p \xi' \vartheta(\mu_p, \ell) \right|^2 \mathcal{I}\left(\left| \sum_{p=1}^{q} \phi_p \xi' \vartheta(\mu_p, \ell) \right| > D^{-1} \delta L^{1/2} \right)$$
$$\leq \frac{D^2}{L} \sum_{\ell=1}^{L} \left| \sum_{p=1}^{q} \phi_p \xi' \vartheta(\mu_p, \ell) \right|^2 \mathcal{I}\left(\left| \sum_{p=1}^{q} \phi_p \xi' \vartheta(\mu_p, \ell) \right| > k \right)$$
$$\xrightarrow{P} E \left[\left| \sum_{p=1}^{q} \phi_p \xi' \vartheta(\mu_p, \ell) \right|^2 \mathcal{I}\left(\left| \sum_{p=1}^{q} \phi_p \xi' \vartheta(\mu_p, \ell) \right| > k \right) \right]$$

proceeding as in the proof of Proposition 5.1. Then, let $k \to \infty$ to complete the proof. ■

Proposition 5.3. *Under the same conditions of Proposition 5.2, the process*

$$\mathcal{Z}_n(\mu) = \frac{1}{L^{1/2}} \sum_{\ell=1}^{L} \xi' \vartheta(\mu, \ell) \eta_\ell$$

is tight in probability.

Proof. Because the process belongs to the space $\mathcal{D}[0,1]$, to show tightness it suffices to check the moment condition

$$E^*\left(|\mathcal{Z}_n(\mu_2) - \mathcal{Z}_n(\mu_1)|^4\right) \leq [G_n(\mu_2, \mu_1)] + o_p(1)$$

where

$$G_n(\mu_2, \mu_1) = (G(\mu_2) - G(\mu_1))^2 + o_p(1)$$

where $G(\mu)$ are nondecreasing functions on $[0,1]$ and where $0 \leq \mu_1 < \mu_2 \leq 1$.

Because η_ℓ is a sequence of bounded iid random variables with mean zero and variance 1

$$E^*\left(|\mathcal{Z}_n(\mu_2) - \mathcal{Z}_n(\mu_1)|^4\right) = \frac{1}{L^2} \sum_{\ell=1}^{L} \left(\left(\xi'(\vartheta(\mu_2, \ell) - \vartheta(\mu_1, \ell))\right)\right)^4 \tag{5.13}$$
$$+\frac{1}{L^2} \sum_{\ell_1 \neq \ell_2}^{L} \left(\left(\xi'(\vartheta(\mu_2, \ell_1) - \vartheta(\mu_1, \ell_1))\right)\right)^2 \left(\left(\xi'(\vartheta(\mu_2, \ell_2) - \vartheta(\mu_1, \ell_2))\right)\right)^2.$$

The first term on the right of (5.13) is $o_p(1)$ uniformly in $\mu \in [0,1]$ as we now show. First following (5.2) and (5.3), it is

$$\frac{1}{L^2} \sum_{\ell=1}^{L} \left(\sum_{j=0}^{M(n)-1} n^{1/2} vec(\widehat{a}_n(-j-1; \ell) - c(-j-1)) \frac{(\sin(\pi j \mu_2) - \sin(\pi j \mu_1))}{\pi j} \right)^4 + o_p(1)$$

where the $o_p(1)$ is uniform in μ. Next, proceeding as in the proof of Proposition 5.1 in the usual way,

$$E\left(\sum_{j=0}^{M(n)-1} n^{1/2} vec(\widehat{a}_n(-j-1; \ell) - c(-j-1)) \frac{(\sin(\pi j \mu_2) - \sin(\pi j \mu_1))}{\pi j} \right)^4 \leq D$$

so that, as we can focus on μ_1 and μ_2 of the type $qM^{-1}(n)$, $q = 0, 1, ..., M(n)$, from the definition of $\vartheta(\mu, \ell)$,

$$\begin{aligned}
& E \sup_{\mu_1, \mu_2} \frac{1}{L^2} \sum_{\ell=1}^{L} \left(\sum_{j=0}^{M(n)-1} n^{1/2} vec(\widehat{a}_n(-j-1; \ell) - c(-j-1)) \frac{(\sin(\pi j \mu_2) - \sin(\pi j \mu_1))}{\pi j} \right)^4 \\
\leq \; & E \sum_{i,k=1}^{M(n)} \frac{1}{L^2} \sum_{\ell=1}^{L} \left(\sum_{j=0}^{M(n)-1} n^{1/2} vec(\widehat{a}_n(-j-1; \ell) - c(-j-1)) \frac{(\sin(\pi j \mu_2) - \sin(\pi j \mu_1))}{\pi j} \right)^4 \\
\leq \; & M(n)^2 L^{-1} D = o(1)
\end{aligned}$$

since by C8 $M^2(n)n^{-1} = o(1)$ and $Ln^{-1} \to 1$. So, the first term on the right of (5.13) is $o_p(1)$ uniformly in $\mu \in [0,1]$.

So to conclude it suffices to show that the second term on the right of (5.13) $= G_n(\mu_2, \mu_1) = (G(\mu_2) - G(\mu_1))^2 + o_p(1)$. But this is the case since that term is bounded by

$$\left(\frac{1}{L} \sum_{\ell=1}^{L} \left((\xi' (\vartheta(\mu_2, \ell) - \vartheta(\mu_1, \ell))) \right)^2 \right)^2 .$$

By Proposition 5.1,

$$\begin{aligned} E\frac{1}{L} \sum_{\ell=1}^{L} \left((\xi' (\vartheta(\mu_2, \ell) - \vartheta(\mu_1, \ell))) \right)^2 &\to \xi' \left(\frac{1}{4\pi} \int_{\mu_1 \pi}^{\mu_2 \pi} \left(f_{xx}^{-1}(-\lambda) \otimes f_{uu}(\lambda) \right) d\lambda \right) \xi \\ &= G(\mu_2) - G(\mu_1) \end{aligned}$$

and it is straightforward to show that

$$\left(\frac{1}{L} \sum_{\ell=1}^{L} \left((\xi' (\vartheta(\mu_2, \ell) - \vartheta(\mu_1, \ell))) \right)^2 - (G(\mu_2) - G(\mu_1)) \right)^2 = O_p\left(L^{-1}\right)$$

so that proceeding as with the proof of the first term on the right of (5.13) since by C10 $M^2(n)n^{-1} = o(1)$ and $Ln^{-1} \to 1$, we conclude that the second term on the right of (5.13) is

$$(G(\mu_2) - G(\mu_1))^2 + o_p(1),$$

where the $o_p(1)$ is uniform in μ. The proof now follows since by $C6$, $f_{xx}^{-1}(-\lambda) \otimes f_{uu}(\lambda)$ is an Hermitian matrix, so $G(\mu)$ is nondecreasing. ■

6. AUXILIARY RESULTS

Lemma 6.1. *Let $h_{w_p,j}(\ell)$ denote the pth element of $h_{w,j}(\ell)$ given in (4.1). Assuming C1-C10, for $p, q = 1, ..., p_1 p_2$,*

$$(a) \quad \text{For } \ell > n, \; E\left(h_{w_p,j}(1)\, h_{w_q,-j}(\ell)\right) = O\left(\frac{\log j}{j} f_{w_p w_p,j}^{1/2} f_{w_q w_q,j}^{1/2}\right)$$

$$(b) \quad \text{For } \ell \leq n, \; E\left(h_{w_p,j}(1)\, h_{w_q,-j}(\ell)\right) = Df_{w_p w_q,j} + O\left(\frac{\log j}{j} f_{w_p w_p,j}^{1/2} f_{w_q w_q,j}^{1/2}\right)$$

and for $k < j$,

$$(c) \quad E\left(h_{w_p,j}(1)\, h_{w_q,-k}(\ell)\right) = O\left(f_{w_p w_p,j}^{1/2} f_{w_q w_q,k}^{1/2} \log j \left(\frac{\mathcal{I}(2k < j)}{j} + \frac{\mathcal{I}(k < j \leq 2k)}{k} \right)\right).$$

Proof. Let $K(\lambda) = (2\pi n)^{-1} \left| \sum_{t=1}^{n} e^{it\lambda} \right|^2$ be the Fèjer kernel. Since $\int_{-\pi}^{\pi} e^{i\ell\lambda} K(\lambda - \lambda_j)\, d\lambda = 0$ for $\ell > n$, then

$$\begin{aligned} E\left(h_{w_p,j}(1)\, h_{w_q,-j}(\ell)\right) &= \int_{-\pi}^{\pi} f_{w_p w_q}(\lambda) e^{i\ell\lambda} K(\lambda - \lambda_j)\, d\lambda \\ &= \int_{-\pi}^{\pi} \left(f_{w_p w_q}(\lambda) - f_{w_p w_q}(\lambda_j) \right) e^{i\ell\lambda} K(\lambda - \lambda_j)\, d\lambda. \end{aligned}$$

Now proceed as in the proof of Theorem 2 part (a) of Robinson (1995) to conclude the proof of part (a). Observe that the term $(j/T)^{\alpha}$ which appears in that theorem is due to the approximation given in $C1$ of, for instance, $f_{w_p w_p}(\lambda)$ by $C_{w_p}\lambda^{-2d_{w_p}}$, which is not our case.

Parts (b) and (c) follow straightforwardly, proceeding as in the proof of Theorem 2 parts (a) and (c) of Robinson (1995) respectively. ■

Remark 2. *Observe that part (a) of Lemma 6.1 examines the behaviour of the finite discrete Fourier transform for two samples with no common observations.*

For notational simplicity, in the next two lemmas we will assume that both x_t and u_t are scalars.

Lemma 6.2. *Assuming C1-C6, and C8*

$$\frac{n^{1/2}}{M(n)}\sum_{p=-M(n)}^{M(n)-1}\widehat{f}_{ux,2mp}f_{xx,2mp}^{-1}e^{ij\lambda_{2mp}}=\frac{1}{n^{1/2}}\sum_{s=-[n/2]}^{[n/2]-1}I_{ux,s}f_{xx,s}^{-1}e^{ij\lambda_s}+O_p\left(\left(\frac{m}{n}\right)^{\frac{1}{2}+(d_x-d_u)}\right). \tag{6.1}$$

Proof. By definition of $\widehat{f}_{ux,2mp}$, after straightforward calculations, the left side of (6.1) is

$$\begin{aligned}
&\frac{1}{n^{1/2}}\sum_{s=-[n/2]}^{[n/2]-1}I_{ux,s}f_{xx}^{-1}\left(\lambda_{2m([s/2m]+1)}\right)e^{ij\lambda_{2m([s/2m]+1)}}\\
=\;&\frac{1}{n^{1/2}}\sum_{s=-[n/2]}^{[n/2]-1}I_{ux,s}f_{xx,s}^{-1}e^{ij\lambda_s} \qquad (6.2)\\
&-\frac{1}{n^{1/2}}\sum_{s=-[n/2]}^{[n/2]-1}I_{ux,s}f_{xx,s}^{-1}e^{ij\lambda_s}\left(1-f_{xx,s}f_{xx}^{-1}\left(\lambda_{2m([s/2m]+1)}\right)e^{ij\left(\lambda_{2m([s/2m]+1)}-\lambda_s\right)}\right),
\end{aligned}$$

where $[a]$ denotes the integer part of the number a.

So it suffices to show that the second term on the right of (6.2) is $O_p\left((m/n)^{1/2+(d_x-d_u)}\right)$. First observe that $f_{xx}^{-1}\left(\lambda_{2m([s/2m]+1)}\right)$ is a constant function for those s such that $pm+1<s\leq(p+1)m$ and for each $p=-M(n),...,M(n)-1$. Now, by symmetry, the second term on the right of (6.2) is proportional to

$$\frac{1}{n^{1/2}}\sum_{p=0}^{[n/m]-1}\left(\sum_{s=pm+1}^{(p+1)m}I_{ux,s}f_{xx,s}^{-1}e^{ij\lambda_s}g_s\right),$$

where $g_s=f_{xx,s}f_{xx}^{-1}\left(\lambda_{2m([s/2m]+1)}\right)e^{ij\left(\lambda_{2m([s/2m]+1)}-\lambda_s\right)}-1$. The second moment of the last displayed expression is

$$\begin{aligned}
&\frac{1}{n}\sum_{p=0}^{[n/m]-1}E\left|\sum_{s=pm+1}^{(p+1)m}I_{ux,s}f_{xx,s}^{-1}e^{ij\lambda_s}g_s\right|^2 \qquad (6.3)\\
&+\frac{2}{n}\sum_{p_1=0<p_2}^{[n/m]-1}E\left(\sum_{s=p_1m+1}^{(p_1+1)m}I_{xu,s}f_{xx,s}^{-1}e^{ij\lambda_s}g_s\times\sum_{s=p_2m+1}^{(p_2+1)m}I_{ux,-s}f_{xx,-s}^{-1}e^{-ij\lambda_s}g_{-s}\right).
\end{aligned}$$

By Lemma 6.1, the first term of (6.3) is bounded by n^{-1} times

$$D\sum_{p=1}^{[n/m]}\left\{\sum_{s=pm+1}^{(p+1)m} f_{uu,s}f_{xx,s}^{-1}\left|g_s\right|^2+\sum_{s_1=pm+1<s_2}^{(p+1)m} f_{xx,s_1}^{-1/2}f_{uu,s_1}^{1/2}f_{uu,s_2}^{1/2}f_{xx,s_2}^{-1/2}g_{s_1}g_{-s_2}\frac{\log^2 s_2}{s_1^2}\right\}$$

$$\leq \quad D\log^2 n\sum_{p=1}^{[n/m]}\frac{m^{1+2(d_x-d_u)}}{p^{2+2(d_u-d_x)}}=O\left(m\left(m/n\right)^{2(d_x-d_u)}\log^2 n\right)$$

since by C1 and C6 $|g_s|\leq D\left|1-\left(s/m\left(p+1\right)\right)\right|\leq D\left(p+1\right)^{-1}$ for $pm+1<|s|\leq (p+1)\,m$, and $D^{-1}\lambda_s^{-2d_x}\leq f_{xx,s}\leq D\lambda_s^{-2d_x}$.

Next, the second term of (6.3). By Lemma 6.1 part (a) and independence between x_t and u_t by C3, that term is bounded in absolute value by n^{-1} times

$$\sum_{p_1=1<p_2}^{[n/m]}\frac{1}{p_1p_2}\sum_{s_1=p_1m+1}^{(p_1+1)m}\sum_{s_2=p_2m+1}^{(p_2+1)m} f_{xx,s_1}^{-1/2}f_{uu,s_1}^{1/2}f_{xx,s_2}^{-1/2}f_{uu,s_2}^{1/2}\frac{\log^2 s_2}{s_1^2}$$

$$= \quad n^{2(d_u-d_x)}\sum_{p_1=1<p_2}^{[n/m]}\frac{1}{p_1p_2}\sum_{s_1=p_1m+1}^{(p_1+1)m}\sum_{s_2=p_2m+1}^{(p_2+1)m}\frac{1}{s_1^{2+d_u-d_x}}\frac{1}{s_2^{d_u-d_x}}$$

$$= \quad O\left(m^{-1}\left(n/m\right)^{1+(d_u-d_x)}\right)=o\left(n\left(n/m\right)^{1/2+(d_u-d_x)}\right),$$

using $C1$ and $C6$ as above, and since $n/m^3=o\left(1\right)$ by C8. This concludes the proof that the second term on the right of (6.2) is $O\left(\left(m/n\right)^{1/2+(d_x-d_u)}\right)$, and thus the lemma. ■

Lemma 6.3. *Assuming C1-C6, for $\ell>n$,*

$$E\left(\sum_{s=1}^{[n/2]} f_{xx,s}^{-1}I_{xu,s}\left(1\right)e^{ij\lambda_s}\sum_{s=1}^{[n/2]} f_{xx,-s}^{-1}I_{ux,-s}\left(\ell\right)e^{-ik\lambda_s}\right)$$

$$=O\left(n^{2(d_u-d_x)}\mathcal{I}\left(d_u>d_x\right)+\mathcal{I}\left(d_u\leq d_x\right)\log^2 n\right).$$

Proof. The left side of the last displayed equality is

$$\sum_{s=1}^{[n/2]} f_{xx,s}^{-2}E\left(I_{xu,s}\left(1\right)I_{ux,-s}\left(\ell\right)\right)e^{i(j-k)\lambda_s} \tag{6.4}$$

$$+2\sum_{s_1=1<s_2}^{[n/2]} f_{xx,s_1}^{-1}E\left(I_{xu,s_1}\left(1\right)I_{ux,-s_2}\left(\ell\right)\right)f_{xx,s_2}^{-1}e^{ij\lambda_{s_1}-ik\lambda_{s_2}}.$$

By Lemma 6.1 part (a), $C1$ and because $\{x_t\}$ and $\{u_t\}$ are mutually independent by $C3$, the first term of (6.4) is bounded in absolute value by

$$\sum_{s=1}^{[n/2]} f_{xx,s}^{-1}f_{uu,s}\frac{\log^2 s}{s^2}\leq Kn^{2(d_u-d_x)}\sum_{s=1}^{[n/2]}\frac{\log^2 s}{s^{2+2(d_u-d_x)}}\leq Kn^{2(d_u-d_x)}$$

since $d_x < 1/2$. The second term of (6.4) is by Lemma 6.1 part (c), bounded in absolute value by

$$\begin{aligned} & \sum_{s_1 < s_2 \leq 2s_1} f_{xx,s_1}^{-1/2} f_{xx,s_2}^{-1/2} f_{uu,s_1}^{1/2} f_{uu,s_2}^{1/2} \frac{\log^2 s_2}{s_1^2} + \sum_{2s_1 < s_2} f_{xx,s_1}^{-1/2} f_{xx,s_2}^{-1/2} f_{uu,s_1}^{1/2} f_{uu,s_2}^{1/2} \frac{\log^2 s_2}{s_2^2} \\ \leq & \; K n^{2(d_u - d_x)} \mathcal{I}\left(d_u > d_x\right) + \mathcal{I}\left(d_u \leq d_x\right) \log^2 n \end{aligned}$$

since $f_{xx,s_1}^{-1/2} \leq K f_{xx,s_2}^{-1/2}$ and $f_{uu,s_2}^{1/2} \leq K f_{uu,s_1}^{1/2}$ by $C1$ and $C6$. ■

Remark 3. *Lemmas 6.2 and 6.3 implies that estimators of* $c\left(j\right)$ *given in* (4.3) *are asymptotically independent if they employ two groups of observations with no common elements.*

Lemma 6.4. *Assuming C1-C10, for* $k < n < \ell$,

$$\begin{aligned} & E\left\{ \left(\sum_{j=0}^{k-1} \left(\widehat{a}_n\left(-j-1;1\right) - c\left(-j-1\right)\right) \frac{\sin\left(\pi j \mu_1\right)}{\pi j} \right) \right. \\ & \left. \times \left(\sum_{j=0}^{k-1} \left(\widehat{a}_n\left(-j-1;\ell\right) - c\left(-j-1\right)\right) \frac{\sin\left(\pi j \mu_2\right)}{\pi j} \right) \right\} = o\left(1\right). \end{aligned}$$

Proof. The proof is immediate, since by Lemmas 6.2 and67.3

$$\begin{aligned} & E\left[\left(\widehat{a}_n\left(-j-1;1\right) - c\left(-j-1\right)\right)\left(\widehat{a}_n\left(-j-1;\ell\right) - c\left(-j-1\right)\right)\right] \\ = & \; O\left(n^{2(d_u - d_x)-1} \log^2 n \mathcal{I}\left(d_u > d_x\right) + n^{-1} \log^4 n \mathcal{I}\left(d_u \leq d_x\right)\right), \end{aligned}$$

and $\sum_{j=1}^{k-1} j^{-1} \leq D \log k \leq D \log n$ and $0 \leq d_u, d_x < 1/2$. ■

Remark 4. *Again, as in part* (a) *of Lemma 6.1, the last lemma indicates that the* HI *estimators of* $c\left(j\right)$ *using two samples with no common observations are asymptotically uncorrelated.*

References

[1] Billingsley, P. (1968), Convergence of Probability Measures. John Wiley, New York.

[2] Brillinger, D.R. (1981), Time Series, Data Analysis and Theory. San Francisco: Holden-Day.

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

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

[5] Gine, E. and S. Zinn (1990), Bootstrapping general empirical measures. *Annals of Probability*, 18, 851-869.

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

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

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

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

[10] Hidalgo, J. (1999), Nonparametric test for causality with long-range dependence. Forthcoming in Econometrica.

[11] Hosking, J. (1981), Fractional differencing. Biometrika, 68, pp. 165-176.

[12] Hosoya, Y. (1977), On the Granger condition for non-causality. Econometrica, 45, pp. 1735-1736.

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

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

[15] Koul, H.L. and W. Stute (1999), Nonparametric model checks for time series. Annals of Statistics, 27, pp.204-236.

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

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

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

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

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

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

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

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

[24] Sims, C.A. (1972), Money, income and causality. American Economic Review, 62, pp. 540-552.

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

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

[27] Sinai, Y.G. (1976), Self-similar probability distribution. Theo. Probability and its Applications, 21, pp. 64-80.

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

[29] Wu, C.F.J. (1986), Jackknife, bootstrap and other resampling methods in regression analysis. Annals of Statistics, 14, pp. 1261-1295.