ECONSTOR

WWW.ECONSTOR.EU

Der Open-Access-Publikationsserver der ZBW – Leibniz-Informationszentrum Wirtschaft
The Open Access Publication Server of the ZBW – Leibniz Information Centre for Economics

Dette, Holger; Paparoditis, Efstathios

Working Paper

Bootstrapping frequency domain tests in multivariate time series with an application to comparing spectral densities

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Mitglied der

Leibniz-Gemeinschaft

BOOTSTRAPPING FREQUENCY DOMAIN TESTS IN MULTIVARIATE TIME SERIES WITH AN APPLICATION TO COMPARING SPECTRAL DENSITIES

HOLGER DETTE AND EFSTATHIOS PAPARODITIS

ABSTRACT. We propose a general bootstrap procedure to approximate the null distribution of nonparametric frequency domain tests about the spectral density matrix of a multivariate time series. Under a set of easy to verify conditions, we establish asymptotic validity of the proposed bootstrap procedure. We apply a version of this procedure together with a new statistic in order to test the hypothesis that the spectral densities of not necessarily independent time series are equal. The test statistic proposed is based on a L_2-distance between the nonparametrically estimated individual spectral densities and an overall, 'pooled' spectral density, the later being obtained using the whole set of m time series considered. The effects of the dependence between the time series on the power behavior of the test are investigated. Some simulations are presented and a real-life data example is discussed.

1. INTRODUCTION

The problem of comparing time series or identifying similarities or dissimilarities in time series data has a long history and arises in several fields, such as economics, marketing, business, finance, medicine, biology, physics, psychology, zoology, and many others. Various statistical tools have been used for this purpose including discriminant and cluster analysis, classification and simple comparisons. Because many important and interesting hypotheses about the multivariate process can be expressed in terms of the spectral density matrix several author use metrics based on the spectral measure in the context of discrimination and classification [see e.g. Darghai-Nourbary and Laycock (1981), Shumway (1982), Zhang and Taniguchi (1994), Kakizawa, Shumway, Taniguchi (1998), Caido, Crato and Peña (2006) among many others]. Cluster and discriminant analysis are closely related to testing problems for the equality of spectral densities in multivariate time series data, which has also found considerable interest in the literature. Jenkins

(1961) was one of the early attempts. De Souza and Thomson (1982) used for this testing problem an autoregressive model-fitting approach. Coates and Diggle (1986) compare the spectral densities of two independent time series using periodogram based test statistics and used this method for analyzing wheat price and British gas data. Swanepoel and van Wyk (1986) consider two independent stationary autoregressive processes and use different test statistics and a parametric, autoregressive bootstrap approach to obtain critical values. Diggle and Fisher (1991) propose graphical devices to compare periodograms and apply Kolmogorov-Smirnov or Cramer-von Mises type test statistics for the analysis of hormonal data. Carmona and Wang (1996) analyzed Lagrangian velocities of drifters at the surface of the ocean by a comparison of spectra. Guo (1999) considers first order autoregressions, Timmer et al. (1999) concentrate on spectral peaks and Maharaj (2002) compares evolutionary spectra of non-stationary processes using randomization tests. A test for homogeneity of autoregressive processes has been also considered by Gómez and Drouiche (2002) and applied to the problem of speech detection in noisy environment. Besides the hypothesis that the spectral densities of the m individual time series (or an appropriately rescaled version thereof) are identical there are numerous other hypotheses of interest, which can be characterized by properties of the spectral measure. Typical examples include the hypothesis of no correlation or no partial correlation between time series after removing the linear effects of some other time series or the hypothesis of separability. To mention only few of the approaches proposed, we refer to Taniguchi and Kondo (1993), Taniguchi et al. (1996), Taniguchi and Kakizawa (2000), Matsuda and Yajima (2004). Recently Eichler (2008) and Yajima and Matsuda (2008) discussed the problem of testing non- and semiparametric hypotheses using spectral analysis in a very general fashion. Following Eichler (2008) we consider a general class of hypotheses in this context, which contains most of the hypotheses mentioned in the previous paragraph and can be stated as

$$(1.1) \qquad H_0 : \int_{-\pi}^{\pi} \|\varphi(f(\lambda), \lambda)\|^2 d\lambda = 0 \quad \text{against} \quad H_1 : \int_{-\pi}^{\pi} \|\varphi(f(\lambda), \lambda)\|^2 d\lambda > 0,$$

where $f(\lambda)$ denotes the spectral density matrix of the underlying (stationary) process, $\varphi(\cdot, \cdot)$ is some suitable vector-valued function specifying the particular null hypothesis of interest and $\|\cdot\|$ denotes the Euclidean norm.

A test statistic for the general pair of hypotheses (1.1) can be obtained by substituting a nonparametric estimator $\hat{f}(\lambda)$ for the spectral density matrix $f(\lambda)$, that is

$$(1.2) \qquad\qquad S_n(\varphi) = \int_{-\pi}^{\pi} \|\varphi(\hat{f}(\lambda), \lambda)\|^2 d\lambda.$$

For test statistics like (1.2), a general asymptotic theory has been developed in Eichler (2008) which enables the approximation of the distribution of $S_n(\varphi)$ under the null and under local alternatives by appropriate Gaussian distributions. For instance, it has been shown that under certain regularity conditions and if H_0 is true, then an appropriate

centering sequence $\mu_n(\varphi)$ and a positive constant $\sigma^2(\varphi)$ exist, such that

$$(1.3) \qquad\qquad n\sqrt{h}S_n(\varphi) - \mu_n(\varphi) \Rightarrow N(0, \sigma^2(\varphi)),$$

as $n \to \infty$ (see Section 2 for details). Here and in the sequel '$\Rightarrow$' stands for weak convergence.

The aim of the present paper is twofold. First and alternatively to the above Gaussian approximation, a simple frequency domain nonparametric bootstrap procedure is proposed which can be applied to approximate correctly the distribution of the test statistic $S_n(\varphi)$ under the null hypothesis. The approach is general enough that enable its applicability to a wide range of testing problems. The motivation for the development of a bootstrap proposal lies in the fact that the quality of the large sample Gaussian approximations for L_2-type functionals like those given in (1.2) is very pure [see e.g. Linton and Fan (2003)]. Furthermore, in many situations the centering sequence $\mu_n(\varphi)$ and the variance $\sigma^2(\varphi)$ of the limiting Gaussian distribution derived under the null hypothesis, depend in a complicated way on unknown and not easy to estimate parameters of the underlying process. The aforementioned concerns make bootstrap approaches an attractive alternative. Notice that nonparametric frequency domain bootstrap methods for stationary time series have been proposed and investigated by many authors in the literature; see Hurvich and Zeger (1987), Franke and Härdle (1992), Dahlhaus and Janas (1996), Paparoditis and Politis (1999) and Kreiss and Paparoditis (2003). For an overview and a discussion of the different approaches to bootstrap time series in the frequency domain see Paparoditis (2000). However, none of these approaches are directly applicable to the testing set-up considered in this paper. The reason for this lies in the fact that for a bootstrap procedure to be successful in a testing context, it should be able to approximate correctly the distribution of the test statistic used under the null hypothesis even if the null hypothesis is not true. This property is important for a good power behavior of a test based on bootstrap critical values. Furthermore, and additionally to satisfying the null hypothesis, the cross-correlation structure of the underlying process should be mimicked correctly in the bootstrap world, since as we will see in the sequel, this affects the validity of the bootstrap procedure in approximating correctly the distribution of interest. The general bootstrap procedure proposed in this paper fulfills these requirements, i.e., it succeeds in generating pseudo-periodogram matrices that mimic correctly all desired properties of the periodogram matrix of the observed process under validity of the null hypothesis. In particular we establish a bootstrap central limit theorem under a set of easy to verify conditions which ensures asymptotic validity of the bootstrap procedure proposed.

The second aim of this paper is to investigate more closely the properties of a test of equality of the spectral densities of a number m of not necessarily independent time series. We propose a nonparametric test for such hypothesis by appropriately specifying the function $\varphi(\cdot)$ and show how the general bootstrap procedure proposed in this paper can be adapted to this particular testing problem. In the case of independent series the asymptotic distribution under the null does not depend on any nuisance parameters. Furthermore, we investigate more closely how the (possible) dependence between the time

series considered affects the distribution of the test statistic and especially the power be-havior of the test under fixed alternatives. Our research is motivated by the fact that all approaches mentioned in the first paragraph of this introduction suffer from at least one of the following three drawbacks: They assume that the time series considered are uncor-related respectively independent, they impose some parametric, commonly autoregressive structure on the underlying process class and the analysis is restricted to bivariate pro-cesses using test statistics generalizations of which to more than two time series are not straightforward. The testing procedure investigated in this paper overcomes all aforemen-tioned drawbacks. The approach proposed is based on an evaluation over all frequencies of the distance between the nonparametrically estimated spectral density of each individ-ual time series and an estimated, pooled spectral density, the later being obtained using the whole set of m time series at hand; see (2.2). Finally, we investigate more closely, both, theoretically and by means of simulations, how the dependence structure between the time series considered affects the power behavior of the test.

The paper is organized as follows. Section 2 states the main assumptions imposed on the m-dimensional process considered, introduces the class of test statistics and presents some illustrative examples. The basic bootstrap procedure and a corresponding consis-tency result are described in Section 3, which ensures asymptotic validity of the bootstrap procedure in approximating correctly the distribution of the test statistic of interest under the null. Section 4 focuses on the problem of testing equality of spectral densities, intro-duces new test statistic and demonstrates how the general bootstrap procedure proposed can be adapted to this particular testing problem. The power behavior of the test for fixed alternatives is investigated and a small simulation study is presented dealing with the behavior of our test in finite sample situations. Furthermore, a real-life data set is analyzed which demonstrates the capability of our testing methodology to detect differ-ences between spectral densities. All proofs are deferred to Section 5.

We finally note that in principle the methods proposed in this paper could be applied to locally stationary processes, which have found considerable interest in the recent litera-ture [see Dahlhaus (1997, 2000) among others]. However, rigorous proofs would be very technical and are deferred to a future research project.

2. Assumptions, test statistic and examples

Suppose that we have n, $n \in \mathbb{N}$, observations $X_1, \ldots, X_n$ of a m-dimensional, zero mean second order stationary stochastic process $\{X_t = (X_{1,t}, X_{2,t}, \ldots, X_{m,t})', t \in \mathbb{Z}\}$ and that

Assumption 1: The random vectors X_t have real components and are generated by the equation

$$X_t = \sum_{j=-\infty}^{\infty} \Psi_j \varepsilon_{t-j},$$

where $\{\Psi_j = (\psi_j(r,s))_{r,s=1,2,\dots,m}, j \in \mathbb{Z}\}$ is a sequence of matrices the components of which satisfy

$$\sum_j |j|^{1/2} |\psi_j(r,s)| < \infty, \quad r, s = 1, 2, \dots, m$$

and $\{\varepsilon_t, \ t \in \mathbb{Z}\}$ are m-dimensional i.i.d. random variables with mean zero, covariance matrix $\Sigma = E[\varepsilon_t \varepsilon_t'] = (\sigma_{k,l})_{k,l=1,\dots,m} > 0$ and $\mathrm{E}[\varepsilon_{r,t}^8] < \infty$, $r = 1, 2, \dots, m.$, where $\varepsilon_t = (\varepsilon_{1,t}, \varepsilon_{2,t}, \dots, \varepsilon_{m,t})'$

Under Assumption 1, the sequence of covariance matrices $\{\Gamma(k), k \in \mathbb{Z}\}$, $\Gamma(k) = E(X_t X_{t+k}')$, has absolutely summable components and the spectral density matrix $f(\lambda) = (f_{r,s}(\lambda))_{r,s=1,2,\dots,m}$, $\lambda \in [-\pi, \pi]$, of the process $\{X_t, t \in \mathbb{Z}\}$ exists and is given by

$$f(\lambda) = \frac{1}{2\pi} \sum_k \Gamma(k) e^{-i\lambda k}.$$

Denote by $f_r(\lambda)$ the spectral density of the r-th component of the m-dimensional process, that is the r-th element $f_{r,r}(\lambda)$ on the main diagonal of the matrix $f(\lambda)$.

The particular null hypothesis of interest and the corresponding test statistic $S_n(\varphi)$ used are determined by means of a function $\varphi : D \times [-\pi, \pi] \to \mathbb{C}^r$, where D is an open subset of $\mathbb{C}^{m \times m}$ that contains the spectral density matrices. Different hypotheses about the spectral density matrix lead to different specifications of the function φ, but there may exist different functions φ which correspond to the same hypothesis as illustrated in the following example.

Example 2.1. Let $m = 2$, i.e., $\{X_t = (X_{1,t}, X_{2,t}), t \in \mathbb{Z}\}$ and suppose that we are interested in testing whether the spectral densities of the two component time series are identical, that is the composite hypothesis $H_0 : f_1(\lambda) = f_2(\lambda)$. Then, the null hypothesis of interest can be tested by specifying the function φ in (1.1) as

$$(2.1) \qquad \varphi_1(f(\lambda), \lambda) = \frac{f_1(\lambda)}{f_2(\lambda)} - 1,$$

see Eichler (2008), Example 3.9. However, the same hypothesis can be alternatively tested using the alternative specification

$$(2.2) \qquad \varphi_2(f(\lambda), \lambda) = \left(\frac{f_1(\lambda)}{w(\lambda)} - 1, \frac{f_2(\lambda)}{w(\lambda)} - 1 \right)$$

of the function φ, where $w(\lambda) = (f_1(\lambda) + f_2(\lambda))/2$ stands for a pooled spectral density, which could be considered as nuisance parameter in this context. Both specifications can be used to test the hypothesis that the spectral densities are equal. Notice that specification $\varphi_1(f(\lambda), \lambda)$ is not symmetric in f_1 and f_2, while $\varphi_2(\cdot)$ seems more appealing because it can be easily generalized to more than two time series and allows for a better understanding of the behavior of the individual spectral densities by measuring their deviation from a pooled version thereof.

Throughout this paper we assume that the function φ satisfies the following conditions; see also Eichler (2008).

Assumption 2:

   (i) $\varphi(Z, \lambda)$ is holomorphic with respect to Z and satisfies $\|\varphi(Z', -\lambda)\| = \|\varphi(Z, \lambda)\|$.

  (ii) $\varphi(Z, \lambda)$ and its first derivative with respect to $z = \text{vec}(Z)$, $D_Z\varphi(Z, \lambda) = \partial\varphi(Z, \lambda)/\partial z'$ are piecewise Lipschitz continuous in λ.

 (iii) There exists a positive constant η such that for all $\lambda \in [-\pi, \pi]$ the ball $B_{\eta,\lambda} = \{Z \in \mathbb{C}^{m \times m} \mid \|f(\lambda) - Z\| \leq \eta\}$ is contained in D and $\sup_{\lambda \in [-\pi,\pi]} \sup_{Z \in B_{\eta,\lambda}} \|\varphi(Z, \lambda)\| < \infty$.

 (iv) $\int_{-\pi}^{\pi} \|D_Z\varphi(f(\lambda), \lambda)\| d\lambda > 0$.

Consider now the periodogram matrix $I_n(\lambda) = (I_{n,r,s}(\lambda))_{r,s=1,2,\ldots,m}$ where

$$I_n(\lambda) = J_n(\lambda)\overline{J}_n(\lambda), \text{ and } J_n(\lambda) = \frac{1}{\sqrt{2\pi n}} \sum_{t=1}^{n} X_t e^{-i\lambda t}.$$

Here and in the sequel, $^-$ denotes transposition combined with complex conjugation. $I_n(\lambda)$ is usually calculated at the Fourier frequencies $\lambda_j = 2\pi j/n$, $j = -[(n-1)/2], \ldots, [n/2]$. We write $I_r(\lambda)$ for the r-th element $I_{r,r}(\lambda)$ on the main diagonal of the matrix $I_n(\lambda)$ which corresponds to the periodogram of the r-th time series $X_{r,t}, t = 1, 2, \ldots, n$.

For $\lambda \in [-\pi, \pi]$ consider the kernel estimator $\widehat{f}(\lambda) = (\widehat{f}_{r,s}(\lambda))_{r,s=1,2,\ldots,m}$ of the spectral density matrix $f(\lambda)$ defined by

$$(2.3) \qquad \widehat{f}(\lambda) = \frac{1}{n} \sum_{j \in \mathbb{Z}} K_h(\lambda - \lambda_j) I_n(\lambda_j),$$

where $K_h(\cdot) = h^{-1}K(\cdot/h)$, K is the smoothing kernel and h the smoothing bandwidth.

Assumption 3: K is a bounded, symmetric, Lipschitz continuous and non-negative kernel with compact support $[-\pi, \pi]$ satisfying $(2\pi)^{-1} \int_{-\pi}^{\pi} K(x) dx = 1$.

Assumption 4: $h \to 0$ as $n \to \infty$ such that $h \sim n^{-\nu}$ for some $1/4 < \nu < 1/2$.

Notice that the rate at which the bandwidth h is allowed to converge to zero a $n \to \infty$ ensures that the bias in estimating $f(\lambda)$ vanishes fast enough in order to not affect the asymptotic distribution of the test statistic (1.2). This rate is identical to the rate used in Taniguchi and Kondo (1993), Taniguchi et al. (1996) and Taniguchi and Kakizawa (2000).

Now, a test statistic for the pair of hypotheses (1.1) is obtained by substituting $\widehat{f}(\lambda)$ in $\varphi(f(\lambda), \lambda)$ leading to the test statistic $S_n(\varphi)$ defined in (1.2). Notice that under the assumptions made, the test statistic $S_n(\varphi)$ can be written as

$$(2.4) \qquad S_n(\varphi) = \int_{-\pi}^{\pi} \|\text{vec}(\widehat{f}(\lambda) - f(\lambda))\|_{\Gamma_\varphi(\lambda)}^2 d\lambda + o_P(1),$$

where $\|x\|_A = \overline{x}Ax$, $\Gamma_\varphi(\lambda) = \overline{D_Z\varphi(f(\lambda), \lambda)}D_Z\varphi(f(\lambda), \lambda)$ and

$$D_Z\varphi(f(\lambda), \lambda) = \frac{\partial\varphi(Z, \lambda)}{\partial\mathrm{vec}(Z)}\Big|_{Z=f(\lambda)},$$

see Eichler (2008), Lemma 3.4 for the derivation of such a result. Furthermore, if H_0 is true, then Theorem 3.5 in the same reference yields the weak convergence (1.3), where

$$
\begin{aligned}
\mu_n(\varphi) &= \frac{1}{2\pi\sqrt{h}}\int_{-\pi}^\pi K^2(u)du\int_{-\pi}^\pi \mathrm{tr}[\Gamma_\varphi(\lambda)(f'(\lambda)\otimes f(\lambda))]d\lambda, \\
\sigma^2(\varphi) &= \frac{1}{4\pi^2}\int(K*K)^2(y)dy\int_{-\pi}^\pi \mathrm{tr}\Big[\Gamma_\varphi(\lambda)(f'(\lambda)\otimes f(\lambda)) \\
&\qquad \times\{\Gamma_\varphi(\lambda)+\Gamma'_\varphi(-\lambda)\}(f'(\lambda)\otimes f(\lambda))\Big]d\lambda,
\end{aligned}
$$

and $K*K$ denotes the convolution of the kernel K with itself. Note that Eichler (2008) establishes the weak convergence result (1.3) under a set of conditions on the underlying stochastic process which are different compared to our Assumption 1.

Example 2.2. Consider the situation discussed in Example 2.1. For the statistic obtained from the function φ_1 in (2.1) it follows from Eichler (2008) that

$$
\begin{aligned}
\mu_{n1} &= \frac{1}{\pi\sqrt{h}}\int_{-\pi}^\pi(1-\frac{|f_{12}(\lambda)|^2}{f^2(\lambda)})d\lambda\cdot\int K^2(x)dx \\
\sigma_1^2 &= \frac{2}{\pi^2}\int(K*K)^2(x)dx\int_{-\pi}^\pi(1-\frac{|f_{12}(\lambda)|^2}{f^2(\lambda)})^2d\lambda
\end{aligned}
$$

where we use the notation $f = f_1 = f_2$.

For the function φ_2 defined in (2.2) we obtain

$$D_Z\varphi_2(f(\lambda)) = \frac{2}{(f_1(\lambda)+f_2(\lambda))^2}\begin{pmatrix} f_2(\lambda) & 0 & 0 & -f_1(\lambda) \\ -f_2(\lambda) & 0 & 0 & f_1(\lambda) \end{pmatrix},$$

which gives by a tedious calculation (note that $f = f_1 = f_2$)

$$
\begin{aligned}
\mu_{n2} &= \frac{1}{2\pi\sqrt{h}}\int_{-\pi}^\pi(1-\frac{|f_{12}(\lambda)|^2}{f^2(\lambda)})d\lambda\int K^2(x)dx \\
\sigma_2^2 &= \frac{1}{2\pi^2}\int_{-2\pi}^{2\pi}(K*K)^2(y)dy\int_{-\pi}^\pi(1-\frac{|f_{12}(\lambda)|^2}{f^2(\lambda)})^2d\lambda.
\end{aligned}
$$

3. The Bootstrap procedure

The idea underlying the proposed bootstrap is to generate pseudo-periodogram matrices that satisfy the null hypothesis (1.1) and to approximate the distribution of the corresponding test statistic by the distribution of the bootstrap statistic based on these

pseudo-periodogram matrices. The starting point of our considerations is that for a m-dimensional process satisfying Assumption 1, the periodogram matrix can be expressed as

$$(3.1) \qquad I_n(\lambda_j) = \Psi(\lambda_j) I_{n,\varepsilon}(\lambda_j) \overline{\Psi}(\lambda_j) + R_n(\lambda_j),$$

where $\Psi(\lambda) = \sum_{k=-\infty}^{\infty} \Psi_k \exp\{-i\lambda k\}$, $I_{n,\varepsilon}(\lambda)$ is the periodogram matrix of the i.i.d. series $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$, i.e.,

$$I_{n,\varepsilon}(\lambda) = J_{n,\varepsilon}(\lambda) \overline{J}_{n,\varepsilon}(\lambda), \qquad J_{n,\varepsilon}(\lambda) = \frac{1}{\sqrt{2\pi n}} \sum_{t=1}^{n} \varepsilon_t \exp\{-i\lambda t\},$$

and the components $R_{n,r,s}(\lambda)$, $r, s \in \{1, 2, \ldots, m\}$ of the remainder matrix $R_n(\lambda)$, satisfy

$$\sup_{\lambda} E|R_{n,r,s}(\lambda)|^2 = O(n^{-1}) \,;$$

cf. Brockwell and Davis (1991), Prop. 11.7.4. Now, let

$$f^{1/2}(\lambda) = (2\pi)^{-1/2} \Psi(\lambda) \Sigma^{1/2}$$

and notice that by ignoring $R_n(\lambda)$ the periodogram matrix can be approximately written as

$$(3.2) \qquad I_n(\lambda_j) \approx f^{1/2}(\lambda_j) U_n(\lambda_j) \overline{f}^{1/2}(\lambda_j),$$

where $U_n(\lambda) = 2\pi \Sigma^{-1/2} I_{n,\varepsilon}(\lambda) \Sigma^{-1/2}$ is 2π-times the periodogram matrix of the i.i.d. process $\{e_t = \Sigma^{-1/2} \varepsilon_t, t \in \mathbb{Z}\}$ which has mean zero and covariance matrix the $m \times m$ unit matrix I_m. It is well-known that for any number k, $k \in \mathbb{N}$, of fixed frequencies $0 \leq \lambda_1 < \lambda_2 < \ldots \lambda_k \leq \pi$, the corresponding set of complex valued random matrices $\{U_n(\lambda_1), U_n(\lambda_2), \ldots, U_n(\lambda_k)\}$ are asymptotically independent, complex Wishart $W_m^C(1, I_m)$ distributed if $\lambda \neq 0 (\mathrm{mod}\pi)$ and real Wishart $W_m^R(1, I_m)$ distributed if $\lambda = 0(\mathrm{mod}\pi)$; cf. Brockwell and Davis (1991), Prop. 11.7.3. A straightforward approach to bootstrap in the frequency domain is, therefore, to generate pseudo-periodogram matrices by replacing $f(\lambda)$ by the nonparametric (kernel) estimator

$$\widehat{f}(\lambda) = n^{-1} \sum_{j} K_h(\lambda - \lambda_j) I_n(\lambda_j) \,,$$

and $U_n(\lambda_j)$ by independent draws from an appropriate m-dimensional Wishart distribution. However, such an approach to generate pseudo-periodogram matrices is not appropriate in our set-up. To elaborate on, notice first that application of the bootstrap in the testing set-up considered in this paper, requires approximation of the distribution of the test statistic $S_n(\varphi)$ under validity of the null hypothesis even if the null hypothesis is not true in reality. This requirement on the bootstrap procedure is important for a good power behavior of the test based on bootstrap critical values. Now, to fulfill this requirement, the spectral density matrix used to generate the pseudo-periodogram matrices, should satisfy the null hypothesis. This requirement is, however, not necessarily fulfilled if the nonparametric estimator $\widehat{f}(\lambda)$ is used, because under the assumptions made we have $\widehat{f}(\lambda) \to f(\lambda)$ in probability, and the limiting spectral density matrix $f(\lambda)$ does

not satisfy the null hypothesis if the later is not true.

The idea to develop a valid bootstrap procedure is to use instead of the nonparametric estimator $\widehat{f}(\lambda)$, a version of it, say $\widehat{f}_T(\lambda)$, which satisfies the following requirements.

Condition 1:

(i) $\widehat{f}_T$ is Hermitian and nonnegative definite.

(ii) $\widehat{f}_T$ satisfies $\varphi(\widehat{f}_T(\lambda), \lambda) = 0$ for almost all $\lambda \in (-\pi, \pi]$.

(iii) $\widehat{f}_T(\lambda)$ converges in probability as $n \to \infty$, and the limit, say $f_T(\lambda)$ satisfies for almost all $\lambda \in (-\pi, \pi]$,

$$\Gamma_{\varphi,T}(\lambda)\Big(f'_T(\lambda) \otimes f_T(\lambda)\Big) = \Gamma_\varphi(\lambda)\Big(f'(\lambda) \otimes f(\lambda)\Big),$$

$$\Gamma'_{\varphi,T}(-\lambda)\Big(f'_T(\lambda) \otimes f_T(\lambda)\Big) = \Gamma'_\varphi(-\lambda)\Big(f'(\lambda) \otimes f(\lambda)\Big),$$

where $\Gamma_{\varphi,T}(\lambda) = \overline{D_Z\varphi(f_T(\lambda), \lambda)}D_Z\varphi(f_T(\lambda), \lambda)$

To elaborate on the meaning of the above requirements imposed on the matrix $\widehat{f}_T(\lambda)$ to be used in the bootstrap procedure, notice that (i) ensures that $\widehat{f}_T$ satisfies basic properties of a spectral density matrix. Part (ii) of the above condition requires that the version of the spectral density matrix $\widehat{f}_T$ used to generate the pseudo-periodogram matrices satisfies the null hypothesis of interest. To understand the intuition behind part (iii), recall the limiting distribution of $S_n(\varphi)$ under the null hypothesis given in equation (1.3). Observe that this distribution depends essentially on the function φ and the parameters of the underlying process through the expressions $\Gamma_\varphi(\lambda)(f'(\lambda) \otimes f(\lambda))$ and $\Gamma'_\varphi(-\lambda)(f'(\lambda) \otimes f(\lambda))$. Furthermore, it can be shown that the corresponding expressions for the limiting distribution of the bootstrap statistic are given by $\Gamma_{\varphi,T}(\lambda)(f'_T(\lambda) \otimes f_T(\lambda))$ and $\Gamma'_{\varphi,T}(-\lambda)(f'_T(\lambda) \otimes f_T(\lambda))$ respectively. This implies that for the bootstrap to be successful in approximating correctly the distribution of $S_n(\varphi)$ under the null, the expressions in the bootstrap world should coincide with those in the real world. This is achieved by requirement (iii) of the above condition. Notice that (iii) does not imply the equality of $f(\lambda)$ and $f_T(\lambda)$. Furthermore, $\Gamma_\varphi(\lambda)) \neq \Gamma'_\varphi(-\lambda)$ in general.

The bootstrap procedure proposed to approximate the distribution of $S_n(\varphi)$ under the null hypothesis can now be summarized by the following four steps.

Step 1: Let $\widehat{f}_T(\lambda)$ be a nonparametric estimator of $f(\lambda)$ satisfying Condition 1.

Step 2: Generate $I_n^*(\lambda_j)$ independent random matrices where

$$I_n^*(\lambda_j) \sim W_m^C(1, \widehat{f}_T(\lambda_j)),$$

if $1 \leq j < n/2$,

$$I_n^*(\lambda_j) \sim W_m^R(1, \widehat{f}_T(\lambda_j)),$$

if $j \in \{0, n/2\}$ and $I_n^*(\lambda_{-j}) = (I_n^*(\lambda_j))'$ for $j = 1, 2, ..., n/2$.

Step 3: Motivated by (2.4), calculate

$$(3.3) \qquad S_n^*(\varphi) = \int_{-\pi}^{\pi} \|\text{vec}(\widehat{f}^*(\lambda) - \widehat{f}_T(\lambda))\|_{\widehat{\Gamma}_{\varphi,T}(\lambda)}^2 d\lambda,$$

where for $\lambda \in (-\pi, \pi]$,

$$\widehat{f}^*(\lambda) = \frac{1}{n} \sum_j K_h(\lambda - \lambda_j) I_n^*(\lambda_j),$$

$$\widehat{\Gamma}_{\varphi,T}(\lambda) = \overline{D_Z \varphi(\widehat{f}_T(\lambda), \lambda)} D_Z \varphi(\widehat{f}_T(\lambda), \lambda) .$$

Step 4: Approximate the distribution of $Z_n(\varphi) = (n\sqrt{h}S_n(\varphi) - \mu_n(\varphi))/\sigma(\varphi)$ under the null by the conditional distribution of $Z_n^*(\varphi) = (n\sqrt{h}S_n^*(\varphi) - \widehat{\mu}_n(\varphi))/\widehat{\sigma}(\varphi)$ given the sample $X_t, t = 1, 2, \ldots, n$, where $\widehat{\mu}_n(\varphi)$ and $\widehat{\sigma}(\varphi)$ are obtained by replacing $f(\lambda)$ and $\Gamma_\varphi(\lambda)$ by $\widehat{f}_T(\lambda)$ and $\widehat{\Gamma}_{\varphi,T}(\lambda)$ respectively in the expressions for $\mu_n(\varphi)$ and $\sigma(\varphi)$.

As it is obvious from the above considerations, the crucial point of the above bootstrap procedure is the determination of the spectral density estimate $\widehat{f}_T(\lambda)$ used in the bootstrap procedure. This determination depends on the specific null hypothesis of interest, i.e., on the particular specification of the function φ considered. Before we continue we present additional examples in order to demonstrate the general applicability of Condition 1.

Example 3.1. Suppose that $\varphi(f(\lambda), \lambda) = (f_{i,j}(\lambda), i, j \in \{1, 2, \ldots, m\}, j > i)$, i.e., that the null hypothesis is that the m time series are not correlated. In this case $\widehat{f}_T(\lambda)$ satisfying Condition 1 is easily determined as

$$\widehat{f}_T(\lambda) = \text{diag} < \widehat{f}_1(\lambda), \widehat{f}_2(\lambda), \ldots, \widehat{f}_m(\lambda) > .$$

If we are interested in testing that two subsets of vector time series $X_{A,t} = (X_{1,t}, X_{2,t}, \ldots, X_{m_1,t})'$ and $X_{B,t} = (X_{m_1+1,t}, X_{m_1+2,t}, \ldots, X_{m,t})'$ are uncorrelated, where $1 \le m_1 < m$, then $\widehat{f}_T(\lambda)$ can be determined as

$$(3.4) \qquad \widehat{f}_T(\lambda) = \begin{pmatrix} \widehat{f}_A(\lambda) & 0 \\ 0 & \widehat{f}_B(\lambda) \end{pmatrix},$$

where for any nonempty set $C \subset \{1, 2, \ldots, m\}$, $\widehat{f}_C(\lambda) = (\widehat{f}_{i,j}(\lambda))_{i,j \in C}$.
Similarly, define $i, j \in A$, $i \ne j$, $\kappa_{i,j|B}^2(\lambda)$ as the partial squared coherency between time series $X_{i,t}$ and $X_{j,t}$ after removing the linear effects of the time series $X_{B,t}$, i.e.,

$$\kappa_{i,j|B}^2 = |f_{i,j|B}(\lambda)|^2/(f_{i,i|B}(\lambda)f_{j,j|B}(\lambda)),$$

where $f_{i,j|B}(\lambda) = f_{i,j}(\lambda) - f_{i,B}(\lambda)f_B^{-1}(\lambda)f_{B,j}(\lambda)$. Suppose that we are interested in testing that the partial coherencies $\kappa_{i,j|B}^2(\lambda)$ are zero, that is $f_{i,j|B}(\lambda) = 0$ for all $i, j \in A$, $i \ne j$. Equivalently, this can be stated as $f_{i,j}(\lambda) = f_{i,B}(\lambda)f_B^{-1}(\lambda)f_{B,j}(\lambda)$. The corresponding specification of the function φ to test this hypothesis is then given by $\varphi(f(\lambda), \lambda) =$

$(f_{i,j|B}(\lambda)|\ i,j \in A, j > i)$ and the matrix $\widehat{f}_T(\lambda)$ used in the bootstrap procedure is set equal to

$$(3.5) \qquad \widehat{f}_T(\lambda) = \begin{pmatrix} \widetilde{f}_A(\lambda) & \widehat{f}_{A,B}(\lambda) \\ \widehat{f}_{B,A}(\lambda) & \widehat{f}_B(\lambda) \end{pmatrix},$$

where $\widehat{f}_{AB}(\lambda) = (\widehat{f}_{i,j}(\lambda))_{i \in A, j \in B}$ and $\widetilde{f}_A(\lambda) = (\widetilde{f}_{i,j}(\lambda))_{i,j \in A}$ with

$$\widetilde{f}_{i,j}(\lambda) = \begin{cases} \widehat{f}_{i,i}(\lambda) & \text{if } i = j \\[2mm] \widehat{f}_{i,B}(\lambda)\widehat{f}_B^{-1}(\lambda)\widehat{f}_{B,j}(\lambda) & \text{if } i \neq j. \end{cases}$$

A straightforward algebra shows that the specifications of $\widehat{f}_T(\lambda)$ given in (3.4) and (3.5) satisfy all requirements stated in Condition 1. A more involved example on how to apply the bootstrap procedure proposed and which is related to testing equality of spectral densities is discussed in the next section.

The following theorem establishes asymptotic validity of the bootstrap procedure, i.e., it shows that Kolmogorov's distance between the distribution of $S_n(\varphi)$ under the null hypothesis and the distribution of $S_n^*(\varphi)$ given the sample $X_1, X_2, \ldots, X_n$, converges to zero in probability as $n \to \infty$. As a careful read of the proof of this theorem shows, the essential assumption needed to establish consistency of the bootstrap is that the matrix $\widehat{f}_T(\lambda)$ used in Step 1 satisfies Condition 1. This implies that to establish validity of the bootstrap procedure in a particular testing situation that fits into the framework (1.1) it suffices to prove that the matrix $\widehat{f}_T(\lambda)$ used in Step 2 of the bootstrap algorithm to generate the pseudo-periodogram matrices $I_n^*(\lambda_j)$ satisfies this condition.

Theorem 3.1. *Suppose that Assumptions 1-4 are satisfied and that the matrix $\widehat{f}_T(\lambda)$ fulfills Condition 1. Then, as $n \to \infty$,*

$$\sup_{x \in \mathbb{R}} \left| P_{H_0}(Z_n(\varphi) \leq x) - P(Z_n^*(\varphi) \leq x | X_1, X_2, \ldots, X_n) \right| \ \to \ 0,$$

in probability, where $P_{H_0}(Z_n(\varphi) \leq \cdot)$ denotes the distribution function of $Z_n(\varphi)$ under the assumption that the null hypothesis is true.

As pointed out by a referee it is of some interest to investigate if it is possible to prove that the proposed bootstrap distribution is a better approximation compared to the asymptotic normal distribution using Edgeworth expansions. Because such a technical investigation would be beyond the scope of the present paper we defer it to a future research project. The complexity of such results is indicated by a recent paper of Linton and Yao (2003) who derived an Edgeworth expansion for an L^2-type statistic in the context of testing for a parametric form of a regression based on an i.i.d. sample.

4. Testing Equality of Spectral Densities

4.1. **Test Statistic.** Consider the specific problem of testing

$$H_0: \quad f_1 = f_2 = \cdots = f_m, \quad \text{a.e. in } [-\pi, \pi],$$

(4.1)   vs.

$$H_1: \quad f_r \neq f_s \quad \text{for at least one pair } (r, s), \quad r \neq s, \quad \text{and on a set of}$$

$$\text{frequencies } \Lambda \subset [-\pi, \pi] \text{ with positive Lebesque measure.}$$

Assume that the spectral densities $f_r(\lambda)$ fulfill

Assumption 5: $\min_{1 \leq r \leq m} \inf_{-\pi \leq \lambda \leq \pi} f_r(\lambda) > 0$. Note that this condition implies that the underlying process is invertible.

Let $N = mn$ and consider the pooled kernel estimator

$$\widehat{w}(\lambda) = \frac{1}{N} \sum_{r=1}^{m} \sum_{j \in \mathbb{Z}} K_h(\lambda - \lambda_j) I_r(\lambda_j). \tag{4.2}$$

Standard calculations yield under Assumptions 1, 3 and 4 that

$$\mathrm{E}[\widehat{w}(\lambda)] = \frac{1}{N} \sum_{r=1}^{m} \sum_{j \in \mathbb{Z}} K_h(\lambda - \lambda_j)(f_r(\lambda_j) + O(\log(n)n^{-1}))$$

$$= \frac{1}{m} \sum_{r=1}^{m} f_r(\lambda) + O(h^2 + \log(n)n^{-1}) \ \to \ \frac{1}{m} \sum_{r=1}^{m} f_r(\lambda)$$

and

$$\mathrm{Var}[\widehat{w}(\lambda)] = \frac{1}{m^2 n^2} \sum_{r_1, r_2} \sum_{j_1, j_2} K_h(\lambda - \lambda_{j_1}) K_h(\lambda - \lambda_{j_2}) \mathrm{Cov}(I_{r_1}(\lambda_{j_1}), I_{r_2}(\lambda_{j_2}))$$

$$= O(n^{-1}h^{-1}) \to 0.$$

Thus, $\widehat{w}(\lambda)$ is a mean square consistent estimator of the pooled spectral density $w(\lambda) = m^{-1} \sum_{r=1}^{m} f_r(\lambda)$, which could be considered as a nuisance parameter in this context. The statistic we propose to test hypotheses (4.1) is now given by

$$T_n = \frac{1}{m} \sum_{r=1}^{m} \int_{-\pi}^{\pi} \left(\frac{\widehat{f}_r(\lambda)}{\widehat{w}(\lambda)} - 1 \right)^2 d\lambda. \tag{4.3}$$

Notice that T_n is a special case of $S_n(\varphi)$ where the function φ is specified as $\varphi(f(\lambda), \lambda) = (mf_j(\lambda)/\sum_{s=1}^{m} f_s(\lambda) - 1,)_{j=1,2,\dots,m}$. T_n has a nice interpretation since it is an average of the L_2-distances between each estimated individual spectral density $\widehat{f}_r(\cdot)$ and the estimated pooled spectral density $\widehat{w}(\cdot)$.

The limiting distribution of T_n under the null hypothesis is summarized in the following

proposition which shows precisely how this distribution is affected by the dependence structure of the underlying process.

Proposition 4.1. *Let Assumption 1 and Assumptions 3-5 be satisfied and suppose that the null hypothesis H_0 in (4.1) is true. Then, as $n \to \infty$,*

$$N\sqrt{h}\, T_n \; - \; \mu_n \; \Rightarrow \; N(0, \tau^2),$$

where

$$\mu_n = \frac{1}{2\pi\sqrt{h}} \int K^2(x)dx \sum_{r=1}^{m} \sum_{s_1=1}^{m} \sum_{s_2=1}^{m} g_{r,s_1} g_{r,s_2} \int_{-\pi}^{\pi} \kappa_{s_1,s_2}^2(\lambda)d\lambda,$$

$$\tau^2 = \frac{1}{2\pi^2} \int (K*K)^2(y)dy \sum_{r_1=1}^{m} \sum_{r_2=1}^{m} \int_{-\pi}^{\pi} \Big(\sum_{s_1=1}^{m} \sum_{s_2=1}^{m} g_{r_1,s_1} g_{r_2,s_2} \kappa_{s_1,s_2}^2(\lambda) \Big)^2 d\lambda,$$

$g_{r,s} = (\delta_{r,s} - m^{-1})$ *with* $\delta_{r,s}$ *Kronecker's delta, i.e.,* $\delta_{r,s} = 1$ *if* $r = s$ *and* $\delta_{r,s} = 0$ *otherwise, and*

$$\kappa_{s_1,s_2}^2(\lambda) = |f_{s_1,s_2}(\lambda)|^2 / (f_{s_1}(\lambda) f_{s_2}(\lambda))$$

the squared coherency between the component processes $\{X_{s_1,t}\}$ *and* $\{X_{s_2,t}\}$ *respectively.*

By the above proposition, the limiting distribution of T_n under the null hypothesis depends on the entire cross-correlation structure between the individual components of the m-dimensional stochastic process $\{X_t, t \in \mathbb{Z}\}$ as this is expressed by the squared coherencies $\kappa_{s_1,s_2}^2(\lambda)$ appearing in the centering sequence μ_n and the variance τ^2 of the limiting Gaussian distribution. Note that this distribution is not affected by the intra-individual autocorrelation structure of each component series; see also Corollary 4.1 below.

In applications it might be computationally more convenient to use instead of T_n the discretized version

$$T_{D,n} = \frac{2\pi}{N} \sum_{r=1}^{m} \sum_{j=-\nu}^{\nu} \Big(\frac{\widehat{f}_r(\lambda_j)}{\widehat{w}(\lambda_j)} - 1 \Big)^2.$$

It is easily seen that under the assumptions made

$$N\sqrt{h}T_n - \mu_n = N\sqrt{h}T_{D,n} - \mu_n + O_P(\sqrt{h}),$$

which implies that the discretized statistic $T_{D,n}$ has asymptotically the same distribution as the statistic T_n.

An interesting special case of the testing problem (4.1) appears if the m time series considered are uncorrelated. Recall that $\kappa_{s,s}^2(\cdot) \equiv 1$, while for $s_1 \neq s_2$ and $\{X_{s_1,t}\}$ and $\{X_{s_2,t}\}$ uncorrelated processes, $\kappa_{s_1,s_2}^2(\cdot) = 0$. Furthermore, straightforward algebra yields

$$\sum_{r=1}^{m} \sum_{s=1}^{m} g_{r,s}^2 = m - 1 \quad \text{and} \quad \sum_{r_1=1}^{m} \sum_{r_2=1}^{m} \Big(\sum_{s=1}^{m} g_{r_1,s} g_{r_2,s} \Big)^2 = m - 1.$$

These observations lead to the following useful corollary of Proposition 4.1.

Corollary 4.1. *Under the conditions of Proposition 4.1 and if $\{X_t\}$ consists of m un-correlated processes, then as $n \to \infty$,*

$$N\sqrt{h}\, T_n \; - \; \widetilde{\mu}_n \; \Rightarrow \; N(0, \widetilde{\tau}^2),$$

where

$$\widetilde{\mu}_n = \frac{m-1}{\sqrt{h}} \int K^2(x)dx \quad and \quad \widetilde{\tau}^2 = \frac{m-1}{\pi} \int (K * K)^2(y)dy.$$

In certain situations it might be of interest to test whether instead of the autocovariance structure, the autocorrelation structure of the m individual processes is the same, i.e., to test instead of (4.1) the modified null hypothesis

$$(4.4) \qquad\qquad H_0: \; f_1 = c_2 f_2 = \cdots = c_m f_m, \quad \text{a.e. in } [-\pi, \pi],$$

where the (unknown) positive real constants c_r, $r = 2, 3, \ldots, m$ are not all identical. The above hypothesis allows for the stationary variances of the m component process to be different, requires however, that all component processes have the same autocorrelation structure.

For testing hypothesis (4.4) we can proceed as in the construction of the test statistic T_n but our considerations are now based on the rescaled time series $\widetilde{X}_t = \widehat{C}^{-1/2} X_t$ where $\widehat{C}^{-1/2}$ is the diagonal matrix $\widehat{C}^{-1/2} = \text{diag}(\widehat{\gamma}_1(0)^{-1/2}, \widehat{\gamma}_2(0)^{-1/2}, \ldots, \widehat{\gamma}_m(0)^{-1/2})$, $\widehat{\gamma}_r(0) = n^{-1}\sum_{t=1}^{n}(X_{r,t} - \overline{X}_r)^2$ and $\overline{X}_r = n^{-1}\sum_{t=1}^{n} X_{r,t}$. Rescaling by $\widehat{C}^{-1/2}$ forces all time series to have the same sample variance so that possible differences between the corresponding individual spectral densities are attributed to differences in the autocorrelation structure of the component processes.

Let $\widetilde{I}_r(\lambda)$ be the periodogram of the rth rescaled series $\widetilde{X}_{r,t}$, $t = 1, 2, \ldots, n$ and denote by $\widehat{g}_r(\lambda)$ the kernel estimator $\widehat{g}_r(\lambda) = n^{-1}\sum_j K_h(\lambda - \lambda_j)\widetilde{I}_r(\lambda_j)$. Notice that $\widehat{g}_r(\lambda)$ is a consistent estimator of the rescaled individual spectral density $g_r(\lambda) = f_r(\lambda)/\gamma_r(0)$. Furthermore, let $\widehat{v}(\lambda) = m^{-1}\sum_{r=1}^{m} \widehat{g}_r(\lambda)$ which is a kernel estimator of the pooled rescaled spectral density $v(\lambda) = m^{-1}\sum_{r=1}^{m} g_r(\lambda)$. Analogously to (4.3), a useful statistic to test hypothesis (4.4) is then given by

$$(4.5) \qquad\qquad W_n = \frac{1}{m}\sum_{r=1}^{m} \int_{-\pi}^{\pi} \left(\frac{\widehat{g}_r(\lambda)}{\widehat{v}(\lambda)} - 1\right)^2 d\lambda.$$

Let $\widetilde{g}_r(\lambda)$ be the same kernel estimator as $\widehat{g}_r$ but based on the rescaled series $X_{r,t}/\sqrt{\gamma_r(0)}$, where $\gamma_r(0) = \text{Var}(X_{r,t})$. Since $\widehat{\gamma}_r(0) = \gamma_r(0) + O_P(n^{-1/2})$, cf. Brockwell and Davis (1991), Proposition 7.3.1, we get $\widehat{f}_r(\lambda) = f_r(\lambda)/\gamma_r(0) + O_P(n^{-1/2})$ which yields

$$(4.6) \qquad N\sqrt{h}\, W_n - \mu_n = n\sqrt{h}\sum_{r=1}^{m} \int_{-\pi}^{\pi} \left(\frac{\widetilde{g}_r(\lambda)}{\widetilde{v}(\lambda)} - 1\right)^2 d\lambda - \mu_n + O_P(\sqrt{h})$$

with $\widetilde{v}(\lambda) = m^{-1}\sum_{r=1}^{m}\widetilde{g}_r(\lambda)$. By equation (4.6) and the fact that the process $\{C^{-1/2}X_t, t \in \mathbb{Z}\}$ with $C^{-1/2} = \mathrm{diag}(\gamma_1^{-1/2}(0), \gamma_2^{-1/2}(0), \dots, \gamma_m^{-1/2}(0))$ has the same cross-correlation structure as $\{X_t, t \in \mathbb{Z}\}$, we immediately get the following result.

Corollary 4.2. *Let Assumption 1 and Assumptions 3-5 be satisfied and suppose that H_0 in (4.4) is true. Then, as $n \to \infty$, $N\sqrt{h}W_n - \mu_n \Rightarrow N(0, \tau^2)$ where μ_n and τ^2 are given in Proposition 4.1.*

4.2. **Bootstrapping the Test Statistic.** To obtain the spectral density matrix $\widehat{f}_T(\lambda)$ to be used in the bootstrap procedure applied to approximate the distribution of T_n under the null, recall that this matrix should satisfy Condition 1. Part (ii) of this condition requires that $\widehat{f}_T(\lambda)$ satisfies the null hypothesis, i.e., the components on the main diagonal of this matrix should all be equal. Although this requirement can be fulfilled by simple replacing the elements on the main diagonal of $\widehat{f}(\lambda)$ by the pooled estimator $\widehat{w}(\lambda)$, the resulting matrix does not necessarily satisfy part (iii) of the same condition since such a replacement of the elements on the main diagonal of $\widehat{f}(\lambda)$ affects the coherencies, i.e., cross-correlation structure of the m-dimensional process. Recall that by Proposition 4.1 the parameters of the limiting Gaussian distribution of T_n under the null are affected by the cross-correlation structure of the underlying m-dimensional stochastic process. Thus retaining the cross-correlation structure of the observed process is essential for bootstrap consistency. To elaborate on the dependency of the limiting distribution on the cross-correlation structure of the underlying process, notice that for the specification $\varphi(f(\lambda), \lambda) = (\varphi_1, \varphi_2, \dots, \varphi_m)(Z, \lambda)$ with $\varphi_k(f(\lambda), \lambda) = mf_k(\lambda)/\sum_{s=1}^{m}f_s(\lambda) - 1$ and which leads to the test statistic (4.3), we have that under validity of the null hypothesis,

$$\frac{\partial \varphi_k(Z, \lambda)}{\partial z_{i,j}}\bigg|_{Z=f(\lambda)} = \begin{cases} (m-1)(mf_1(\lambda))^{-1} & \text{if } i = j = k \\ -(mf_1(\lambda))^{-1} & \text{if } i = j \neq k \\ 0 & \text{else.} \end{cases}$$

Hence for this specification of φ we have $\Gamma_\varphi(\lambda) = \Gamma_\varphi(-\lambda) = \Gamma_\varphi'(\lambda)$ and we get by straightforward calculations that

$$\mathrm{tr}\{\Gamma_\varphi(\lambda)(f'(\lambda) \otimes f(\lambda))\} = \sum_{r=1}^{m}\sum_{s_1=1}^{m}\sum_{s_2=1}^{m} g_{r,s_1}g_{r,s_2}\int_{-\pi}^{\pi}\kappa_{s_1,s_2}^{2}(\lambda)d\lambda$$

which is a function only of the coherencies $\kappa_{s_1,s_2}^{2}(\lambda)$; see also Proposition 4.1. Clearly, the same problem regarding the alternation of the coherence structure appears if the matrix $\widehat{f}_T(\lambda)$ used in the bootstrap procedure is set equal to a diagonal matrix with all diagonal elements equal to $\widehat{w}(\lambda)$. Also in this case we have that part (i) and (ii) of Condition 1 are satisfied but not part (iii).

The idea to obtain a matrix $\widehat{f}_T(\lambda)$ which satisfies all requirements of Condition 1 lies in the following result. Let $f(\lambda) = (f_{r,s}(\lambda))_{r,s=1,2,\dots,m}$ be a spectral density matrix satisfying Assumption 5 and $D(\lambda)$ the diagonal matrix defined by

$$D(\lambda) = \mathrm{diag}(w(\lambda)f_1^{-1}(\lambda), w(\lambda)f_2^{-1}(\lambda), \dots, w(\lambda)f_m^{-1}(\lambda)).$$

Consider the matrix $f_T(\lambda) = (f_{r,s,T}(\lambda))_{r,s=1,2,\dots,m}$ defined by

$$(4.7) \qquad\qquad f_T(\lambda) = D^{1/2}(\lambda)f(\lambda)D^{1/2}(\lambda),$$

and verify by straightforward calculations that this matrix has the following properties:

(a) f_T is a spectral density matrix, and the elements on its main diagonal are all equal and equal to $w(\lambda) = m^{-1}\sum_{r=1}^m f_r(\lambda)$.
(b) $\widetilde{\kappa}^2_{r,s}(\lambda) = \kappa^2_{r,s}(\lambda)$ where $\widetilde{\kappa}^2_{r,s}(\lambda) = |f_{r,s,T}(\lambda)|^2/(f_{r,T}(\lambda)f_{s,T}(\lambda))$.

Motivated by the above considerations, the spectral density matrix used in the bootstrap procedure is set equal to

$$\widehat{f}_T(\lambda) = \widehat{D}^{1/2}(\lambda)\widehat{f}(\lambda)\widehat{D}^{1/2}(\lambda),$$

where

$$\widehat{D}(\lambda) = \mathrm{diag}(\widehat{w}(\lambda)\widehat{f}_1^{-1}(\lambda), \widehat{w}(\lambda)\widehat{f}_2^{-1}(\lambda), \dots, \widehat{w}(\lambda)\widehat{f}_m^{-1}(\lambda)).$$

Now, let

$$T_n^* = \frac{1}{m}\sum_{r=1}^m \int_{-\pi}^{\pi} \left(\frac{\widehat{f}_r^*(\lambda)}{\widehat{w}^*(\lambda)} - 1\right)^2 d\lambda,$$

where

$$\widehat{f}_r^*(\lambda) = \frac{1}{n}\sum_j K_h(\lambda - \lambda_j)I_r^*(\lambda_j),$$

$\widehat{w}^*(\lambda) = m^{-1}\sum_{r=1}^m \widehat{f}_r^*(\lambda)$ and $I_r^*(\lambda_j)$ is the r-th element on the main diagonal of the pseudo-periodogram matrix $I_n^*(\lambda_j)$. To approximate the distribution of $Z_n = (N\sqrt{h}T_n - \mu_n)/\tau$ under the null, the conditional distribution of $Z_n^* = (N\sqrt{h}T_n^* - \widehat{\mu}_n)/\widehat{\tau}$ given the sample $X_t, t = 1, 2, \dots, n$ is used. Recall that $\widehat{\mu}_n$ and $\widehat{\tau}$ are obtained by replacing $\kappa^2_{s_1,s_2}(\lambda)$ in μ_n and in $\tau = \sqrt{\tau^2}$ by the nonparametric estimator

$$\widehat{\kappa}^2_{s_1,s_2}(\lambda) = \left|n^{-1}\sum_j K_h(\lambda - \lambda_j)I_{s_1,s_2}(\lambda_j)\right|^2 \left(\widehat{f}_{s_1}(\lambda)\widehat{f}_{s_2}(\lambda)\right)^{-1}.$$

Notice that by property (a) transformation (4.7) produces a spectral density matrix which satisfies part (i) and (ii) of Condition 1. Furthermore, by property (b), transformation (4.7) preserves at the same time the cross-correlation structure of the underlying m-dimensional process, i.e., the coherencies corresponding to the transformed spectral density matrix $f_T(\lambda)$ are identical to those of the original spectral density matrix $f(\lambda)$. This ensures that transformation (4.7) fulfills also part (iii) of Condition 1. Thus and because this transformation (4.7) satisfies all requirements of Condition 1, the following result immediately appears.

Proposition 4.2. *Suppose that Assumption 1 and Assumptions 3-5 are satisfied. Then, as $n \to \infty$,*

$$\sup_{x \in \mathbb{R}} \left|P_{H_0}(Z_n \leq x) - P(Z_n^* \leq x|X_1, X_2, \dots, X_n)\right| \;\to\; 0,$$

in probability, where $P_{H_0}(Z_n \leq \cdot)$ denotes the distribution function of Z_n when the null hypothesis is true.

Remark 4.1. An asymptotic level α test is obtained comparing the statistic Z_n with the $(1 - \alpha)$-quantile of the (simulated) distribution of Z_n^*. Note that the test does not provide any information which densities are different from others. This question is closely related to the problem of discriminant and clustering analysis [see for example Kakizawa, Shumway and Taniguchi (1998)] or to a further analysis based on pairwise comparisons. The introduced bootstrap methodology is general enough to be used in such problems. For example pairwise comparisons can be performed using the appropriate function φ proposed in Example 2.1 and the corresponding bootstrap in the frequency domain. Similarly the question where the frequencies differ is not answered by a global test and to solve problems of this type one could apply a local version of our test as proposed by Opsomer and Francisco-Fernández (2008) in the context of comparing regression curves. An alternative approach to localize the differences between the spectral densities is briefly mentioned in Section 5.1.

4.3. **Power Behavior under Fixed Alternatives.** In deriving the power properties of the proposed test, it is important to investigate its behavior under fixed alternatives, that is for the case where the spectral densities of the underlying m time series are not equal. Notice that the power behavior of the general test statistic (1.2) for particular sequences of local alternatives has been investigated by Eichler (2008). The following theorem establishes the limiting distribution of T_n in the case of fixed alternatives.

Theorem 4.1. *Let Assumption 1 and Assumptions 3-5 be satisfied and suppose that the alternative H_1 in (4.1) is true. Then, as $n \to \infty$,*

$$\sqrt{N}\{T_n - M^2 - b_h\} \Rightarrow N(0, \tau_1^2),$$

where

$$(4.8) \qquad M^2 = \frac{1}{m} \sum_{r=1}^{m} \int_{-\pi}^{\pi} \left(\frac{f_r(\lambda)}{w(\lambda)} - 1 \right)^2 d\lambda$$

$$b_h = \frac{2}{\sqrt{m}} \sum_{r=1}^{m} \int_{-\pi}^{\pi} \frac{h_r(\lambda)}{w(\lambda)} \left[\frac{1}{2\pi} \int K_h(\lambda - x) f_r(x) dx - f_r(\lambda) \right] d\lambda$$

$$\tau_1^2 = \frac{16\pi}{m} \int_{-\pi}^{\pi} \left\{ \sum_{r=1}^{m} \frac{f_r(x)}{w(x)} \left(\frac{f_r(x)}{w(x)} - 1 \right) \right\}^2 dx,$$

and $h_r(\lambda) = f_r(\lambda)/w(\lambda) - 1$.

Note that under fixed alternatives asymptotic normality is still valid but with a different standardization. In particular the rate of convergence is $\sqrt{N}$. The result of Theorem 4.1 can be used for several purposes, which we briefly discuss in the following.

1) By Theorem 3.1 and Proposition 4.1 an asymptotic level α test for the hypothesis of equal spectral densities is obtained. It now follows from Theorem 4.1 that the power of this test can be approximated by

(4.9) $$P(H_0 \text{ rejected} \mid H_1 \text{ is true}) \approx 1 - \Phi \left(-\frac{\sqrt{N}(M^2 + b_n)}{\tau_1} + \frac{\mu_n + \tau_0 z_\alpha}{\tau_1 \sqrt{Nh}} \right).$$

Note that in principle the power for a fixed alternative could be also approximated by simulation as proposed in Beran (1986) in the context of i.i.d. observations. This author suggested to generate bootstrap samples under a fixed alternative and to calculate critical values from theses samples by a further nested bootstrap procedure. A proof of the consistency of this concept in the context of a nonparametric null hypothesis or a stationary time series as considered in this paper is still an open problem.

Moreover, formula (4.9) provides additional information on the behavior of the power function, which depends (asymptotically) on the particular alternative only through the three quantities b_h, M^2 and τ_1^2.

2) Note that the quantity M^2 defined in (4.8) can be interpreted as a measure of equality between the spectral densities of the m time series considered. From Theorem 4.1 we obtain

$$T_n - b_h + \frac{\hat{\tau}_1 z_{1-\alpha}}{\sqrt{N}}$$

as an upper (asymptotic) $(1 - \alpha)$ confidence bound for the parameter M^2, where $\hat{\tau}_1^2$ is an appropriate (consistent) estimator of the asymptotic variance given in Theorem 4.1. Such an estimator is obtained, for instance, if $f_r(x)$ and $w(x)$ are replaced by their kernel estimators $\widehat{f}_r(x)$ and $\widehat{w}(x)$, respectively.

3) A further important application of Theorem 4.1 arises from the fact that in practice the second order behavior of the m time series will usually never be precisely identical. The more realistic question in this context is, if the different time series show approximately the same second order behaviour. Therefore we propose to investigate the so called *precise hypotheses* [see Berger and Delampady (1987)]

(4.10) $$H_0 : M^2 > \varepsilon \quad \text{versus} \quad H_1 : M^2 \leq \varepsilon \,,$$

where M^2 is the measure defined by (4.8) and $\varepsilon > 0$ is a prespecified constant for which die statistician agrees to analyse the data under the additional assumption of equal spectral densities. An asymptotic α-level test for the hypothesis (4.10) is obtained by rejecting the null hypothesis, whenever

$$\sqrt{N}(T_n - \varepsilon - b_h) < \hat{\tau}_1 z_{1-\alpha} \,.$$

4) Equation (4.9) is important also because it demonstrates how the correlation structure between the individual series considered, affects the power behavior of the test. In particular, the cross-correlation structure of the m-dimensional process enters the (approximative) power function of the test through the term

$(\mu_n + \tau_0 z_\alpha)/(\tau_1 \sqrt{Nh})$ only, which given the level α of the test and the smoothing quantities h and K, it is determined by the coherencies $\kappa^2_{s_1,s_2}(\cdot)$. Now, this term, although asymptotically negligible because it is of order $N^{-1/2}h^{-1}$, it may affect the power of the test in finite sample situations depending on the values of $\kappa^2_{s_1,s_2}(\cdot)$ and the resulting values of μ_n and τ_0. In particular, given the quantities $f_r(\cdot)/w(\cdot) - 1$, $r = 1, 2, \ldots, m$, i.e., given the deviations between the individual spectral densities f_r and the overall spectral density w, the power of the test is the larger (smaller) the smaller (larger) is the quantity $(\mu_n + \tau_0 z_\alpha)/(\tau_1 \sqrt{Nh})$; see Section 5.1 for an illustration of this point.

5. Numerical Examples

5.1. **Simulations.** To investigate the behavior of our testing procedure in finite sample situations we have conducted a small Monte Carlo experiment where the test statistic $T_{D,n}$ proposed as well as the bootstrap procedure used to obtain critical values have been studied empirically. In this context, observations $X_1, X_2, \ldots, X_n$ have been generated from the simple bivariate process $\{X_t = (X_{1,t}, X_{2,t})', t \in \mathbb{Z}\}$, where

$$(5.1) \qquad X_{1,t} = \phi X_{1,t-1} + \delta X_{1,t-2} + \varepsilon_{1,t}$$
$$X_{2,t} = \phi X_{2,t-1} + \varepsilon_{2,t},$$

$\phi = 0.8$ and $\varepsilon_t = (\varepsilon_{1,t}, \varepsilon_{2,t})' \sim N((0,0)', \Sigma)$, with $\Sigma = (\sigma_{r,s})_{r,s=1,2}$, $\sigma_{1,1} = \sigma_{2,2} = 1$ and $\sigma_{1,2} = \rho \in (-1, 1)$. Different values of δ and ρ have been considered. Notice that ρ controls the degree of dependence between the two processes ($\rho = 0$ corresponds to the case where the two processes are independent) while δ controls the degree of deviation between the spectral densities of the individual series $X_{1,t}$ and $X_{2,t}$. In particular, for $\delta = 0$ the null hypothesis of equal spectral densities is true while for $\delta \neq 0$ we are in the case where the alternative is true. Recall that causality of X_t requires that $\delta \in (-1, 0.2)$ [see Brockwell and Davis (1991), Theorem 11.3.1].

To investigate empirically the size and power behavior of the test $T_{D,n}$, 500 replications of the bivariate process (5.1) have been generated for different sample sizes n and different values of the dependence parameter ρ and the deviation parameter δ. The nonparametric estimators involved in our testing procedure have been calculated using Bartlett-Priestley's kernel (see Priestley (1981), p. 448) and different values of the smoothing bandwidth h. Furthermore, to obtain the critical points of the test using the bootstrap procedure proposed, 1000 bootstrap replications have been generated. The results obtained for $\alpha = 0.05$ are reported in Table 1.

Please insert Table 1 here

As Table 1 shows, although the test leads to some over rejection for the smallest sample size considered, the situation improves rapidly as the time series length n increases with the test achieving the desired size behavior. This behavior is not surprising since due

to the allowed dependence between the individual time series, implementation of the test requires nonparametric, frequency domain estimation of the entire cross-correlation structure of the underlying m-dimensional process which is a difficult task. Concerning the power behavior of the test, we observe that the test leads to high rejection rates even for small differences between the two spectral densities, like those considered in the Monte Carlo experiment ($\delta = \pm 0.1$). Interestingly in model (5.1) detecting differences between the spectral densities under independence ($\rho = 0$) appears to be more difficult than under dependence ($\rho \neq 0$). The explanation for this is given by formula (4.9) of the power function. Notice that for the particular bivariate process (5.1) considered, it is easily seen that $\kappa_{s_1,s_2}^2(\lambda) = \rho^2$ for all $\lambda \in [0, \pi]$, which by straightforward calculations yields

$$\mu_n = (1 - \rho^2)\frac{1}{\sqrt{h}} \int K^2(x)dx, \quad \text{and} \quad \tau_0^2 = (1 - \rho^2)^2 \frac{1}{\pi} \int \left(\int K(x)K(x+y)dx \right)^2 dy.$$

Now, other things being equal, if $\rho^2 = \kappa_{s_1,s_2}^2(\cdot) = 0$, i.e., if the two processes are independent, then μ_n and τ_0^2 achieve their maximal value leading to a large value of $(\mu_n + \tau_0 z_\alpha)/(\tau_1\sqrt{Nh})$ and, consequently, to a drop of power. On the other hand as $\rho^2 = \kappa_{s_1,s_2}^2(\cdot)$ increases, i.e. as the cross-correlation between the two processes becomes stronger, then μ_n and τ_0^2 decrease, leading to a lower value of $(\mu_n + \tau_0 z_\alpha)/(\tau_1\sqrt{Nh})$ and, therefore, to an increase of power.

We next compare the performance of the $T_{D,n}$ test with that of two other tests of equality of spectral densities, namely

$$S_{D,n}^{(1)} = \frac{2\pi}{N} \sum_{j=-\nu}^{\nu} \left(\frac{\widehat{f}_1(\lambda_j)}{\widehat{f}_2(\lambda_j)} - 1 \right)^2 \quad \text{and} \quad S_{D,n}^{(2)} = \frac{2\pi}{N} \sum_{j=-\nu}^{\nu} \left(\frac{\widehat{f}_2(\lambda_j)}{\widehat{f}_1(\lambda_j)} - 1 \right)^2.$$

Notice that $S_{D,n}^{(1)}$ is a discretized version of the test proposed by Eichler (2008). We run several simulations for the three tests using the same testing parameters, the same number of repetitions and bootstrap replications to those used to obtain the results reported in Table 1. Table 2 presents the results obtained for the case of $n = 256$ observations the quality of which is the same to those obtained using other sample sizes and parameter constellations.

Please insert Table 2 here

As this table shows, all tests have approximately a similar size behavior. Recall from the discussion of the results presented in Table 1, that the size behavior of the tests improves considerably as n increases. Now, compared to the two tests $S_{D,n}^{(1)}$ and $S_{D,n}^{(2)}$, the test $T_{D,n}$ is more stable with respects to its power behavior and its ability to detect differences between the spectral densities considered. In particular, the tests $S_{D,n}^{(1)}$ and $S_{D,n}^{(2)}$ show a different power behavior with the test $S_{D,n}^{(1)}$ being better for deviations corresponding to the case $\delta = 0.1$ and the test $S_{D,n}^{(2)}$ being better for the case $\delta = -0.1$. At the same time, both tests have the lowest power in detecting differences between the spectral densities for some of the parameter constellations considered. Notice that while the test $S_{D,n}^{(1)}$ has

slightly more power than the test $T_{D,n}$ for the case $\delta = 0.1$, it has much less power than the same test for $\delta = -0.1$. A similar behavior with different signs occurs for the test $S_{D,n}^{(2)}$.

5.2. **Analysis of grain price data.** The data set considered consists of monthly averages of grain prices for corn, wheat and rye in the United States of America for the period January 1961 to October 1972. It has been discussed in Ahn and Reinsel (1988) and a complete description is given in Reinsel (2003). The original three-variate series is shown in Figure 1. We investigate the question wether the market forces for grain products lead to different price behavior of corn, wheat and rye. Does the price for corn, for wheat and for rye evolve differently over time following a different dependence pattern? In terms of second order properties, this question refers to the question whether the three price series obey a similar autocovariance structure. To test the hypothesis that the spectral densities of the three price series are equal we use the discretized statistic $T_{D,n}$ together with Bartlett-Priestley's smoothing kernel and the bandwidth $h = 0.1$ obtained by means of a cross-validation criterion [Beltrão and Bloomfield (1987)] and applied to the pooled spectral density estimator $\widehat{w}(\lambda)$. For this choice of the smoothing parameters the value of the test statistic is equal to $T_{D,n} = 2.005$, which compared with the upper 5% critical point 0.5057 obtained using $B = 1000$ bootstrap replications, leads to a rejection of the null hypothesis that the autocovariance structure of the three series is identical. Figure 2a) shows on a log scale, the estimated individual spectral densities together with the estimated pooled spectral density $\widehat{w}(\lambda)$.

To get a deeper insight into the reasons leading to the above rejection of the hypothesis of equal spectral densities, and to investigate more closely were the differences between the individual spectral densities lie, we consider the statistic $Q_{r,n}^2(\lambda_j) = (\widehat{f}_r(\lambda_j)/\widehat{w}(\lambda_j) - 1)^2$ calculated for $\lambda_j = 2\pi j/n$, $j = 0, 1, \ldots, [n/2]$. Notice that $Q_{r,n}^2(\lambda_j)$ describes for every frequency λ_j, the squared difference between the estimated rth individual spectral density $\widehat{f}_r(\lambda_j)$ and the pooled spectral density $\widehat{w}(\lambda_j)$ and that the test statistic T_n can be approximately written as $T_n \approx 2\pi m^{-1} n^{-1} \sum_{r=1}^m \sum_{j=-\nu}^{\nu} Q_{r,n}^2(\lambda_j)$, $\nu = [(n-1)/2]$. Large values of $Q_{r,n}^2$ pinpoint, therefore, to frequencies where the spectral density of the rth series deviates from the pooled spectral density. A plot of the statistic $Q_{r,n}^2(\lambda_j)$ for different frequencies and for each of the three price series considered is given in Figure 2b). To better evaluate the plots shown we include in the same figure an estimate of the upper 5%-percentage point of the distribution of the maximum statistic $M_n = \max_{1 \le r \le m} \max_{0 \le \lambda_j \le \pi} Q_{r,n}^2(\lambda_j)$, under the hypothesis that all spectral densities are equal. To estimate the upper 5% percentage-point of this distribution we use the bootstrap procedure described in Section 3 to generate $B = 1000$ replications of $M_n^* = \max_{1 \le r \le m} \max_{0 \le \lambda_j \le \pi} Q_{r,n}^{*2}(\lambda_j)$, where $Q_{r,n}^{*2}(\lambda_j) = (\widehat{f}_r^*(\lambda_j)/\widehat{w}^*(\lambda_j) - 1)^2$ and $\widehat{f}_r^*(\lambda)$ and $\widehat{w}^*(\lambda)$ are defined in Step 3 of the aforementioned bootstrap algorithm.

Please insert Figure 1 and Figure 2 about here

As Figure 2 shows, the autocovariance structure of corn and ray prices seem to be very similar and different to that of wheat prices. The differences lie not only in the fact that wheat prices have a larger variance compared to the other two prices, but also that the spectral density of wheat prices show a moderate peak at frequency $\lambda = 0.796$ which corresponds to a cyclical component of approximately 8 months and which is not apparent in corn and rye prices; cf. Figure 2b). It is worth mentioning here, that these findings are in contrast to what could be expected by a simple inspection of the time series plots of the three series shown in Figure 1. Such an inspection suggests namely that corn and wheat prices behave similar and differently to ray prices.

6. Proofs

Proof of Theorem 3.1: For $I = i + (j-1)m$ and $R = r + (l-1)m$ with $i, j, r, l \in \{1, 2, \ldots, m\}$ we denote by $\widehat{\Gamma}_{ij,rl,T}(\lambda)$ the (I, R) element of $\widehat{\Gamma}_{\varphi,T}(\lambda)$ which is given by

$$\widehat{\Gamma}_{ij,rl,T}(\lambda) = \overline{\left(\frac{\partial \varphi(Z, \lambda)}{\partial z_{ij}}\right)}\left(\frac{\partial \varphi(Z, \lambda)}{\partial z_{rl}}\right)\Big|_{Z = \widehat{f}_T(\lambda)}.$$

Similarly, denote by $\Gamma_{ij,rl,T}(\lambda)$ the corresponding element of $\Gamma_{\varphi,T}(\lambda)$. Let $\widehat{f}_{i,j,T}(\lambda)$ and $f_{i,j,T}(\lambda)$ be the (i, j) elements of $\widehat{f}_T(\lambda)$ and $f_T(\lambda)$ respectively and define

$$(6.1) \qquad\qquad Y_{i,j}^*(\lambda_s) = e_i' \widehat{f}_T^{1/2}(\lambda_s)(U_s^* - I_m)\widehat{f}_T^{1/2}(\lambda_s)e_j,$$

where e_k is the m-dimensional vector $e_k = (0, \ldots, 0, 1, 0, \ldots, 0)'$ with the "1" appearing in the $k-th$ position, $\widehat{f}_T(\lambda) = \widehat{f}_T^{1/2}(\lambda)\widehat{f}_T^{1/2}(\lambda)$, I_m is the $m \times m$ unit matrix and $U_s^* = U_n^*(\lambda_s)$ are independent, complex Wishart $W_m^C(1, I_m)$ distributed if $\lambda_s \neq 0(\mod \pi)$ and real Wishart $W_m^R(1, I_m)$ distributed if $\lambda_s = 0(\mod \pi)$. Denote by * expectations with respect to the bootstrap distribution and notice that for all λ_s,

$$(6.2) \qquad\qquad E^*(Y_{i,j}^*(\lambda_s)) = 0,$$

while

$$(6.3) \qquad Cov^*(Y_{i,j}^*(\lambda_{s_1}), Y_{r,l}^*(\lambda_{s_2})) = \begin{cases} 0 & \text{if } s_1 \neq s_2 \\[2mm] \widehat{f}_{i,r,T}(\lambda_s)\overline{\widehat{f}_{j,l,T}}(\lambda_s) & \text{if } s_1 = s_2 = s. \end{cases}$$

Using the above notation and since $E^*(\widehat{f}(\lambda)^* - \widehat{f}_T(\lambda)) = O_P(h^2)$ we get

$$
\begin{aligned}
n\sqrt{h}S_n^*(\varphi) =& \frac{\sqrt{h}}{n} \sum_{i,j,r,l=1}^{m} \sum_{s_1} \sum_{s_2} \int_{-\pi}^{\pi} K_h(\lambda - \lambda_{s_1}) K_h(\lambda - \lambda_{s_2}) \widehat{\Gamma}_{ij,rl,T}(\lambda) d\lambda \\
& \times Y_{j,i}^*(\lambda_{s_1}) Y_{r,l}^*(\lambda_{s_1}) \ + \ O_P(\sqrt{n}h^2) \\
=& \frac{\sqrt{h}}{n} \sum_{i,j,r,l=1}^{m} \sum_{s} \int_{-\pi}^{\pi} K_h^2(\lambda - \lambda_s) \widehat{\Gamma}_{ij,rl,T}(\lambda) d\lambda Y_{j,i}^*(\lambda_s) Y_{r,l}^*(\lambda_s) \\
& + \frac{\sqrt{h}}{n} \sum_{i,j,r,l=1}^{m} \sum_{\substack{s_1,s_2 \\ s_1 \neq s_2}} \int_{-\pi}^{\pi} K_h(\lambda - \lambda_{s_1}) K_h(\lambda - \lambda_{s_2}) \widehat{\Gamma}_{ij,rl,T}(\lambda) d\lambda \\
& \times Y_{j,i}^*(\lambda_{s_1}) Y_{r,l}^*(\lambda_{s_1}) + O_P(\sqrt{n}h^2) \\
=& M_n^* + L_n^* + o_P(1)
\end{aligned}
$$

with an obvious notation for M_n^* and L_n^* and where the $o_P(1)$ term is due to Assumption 4. We show that

$$(6.4) \qquad\qquad\qquad M_n^* - \widehat{\mu}_n(\varphi) \to 0,$$

in probability, and that

$$(6.5) \qquad\qquad\qquad L_n^* \ \Rightarrow \ N(0, \sigma^2(\varphi)).$$

To establish (6.4) notice that because of (6.3) we have

$$
\begin{aligned}
E^*(M_n^*) =& \frac{\sqrt{h}}{n} \sum_{i,j,r,l=1}^{m} \sum_{s} \int_{-\pi}^{\pi} K_h^2(\lambda - \lambda_s) \widehat{\Gamma}_{ij,rl,T}(\lambda) d\lambda \widehat{f}_{j,r,T}(\lambda_s) \overline{\widehat{f}}_{i,l}(\lambda_s) \\
=& \frac{\sqrt{h}}{2\pi} \sum_{i,j,r,l=1}^{m} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} K_h^2(\lambda - x) \widehat{\Gamma}_{ij,rl}(\lambda) \widehat{f}_{j,r,T}(x) \overline{\widehat{f}}_{i,l}(x) dx d\lambda + O_P(\sqrt{h}).
\end{aligned}
$$

Using the substitution $(\lambda - x)/h = u$ it is easily seen that

$$
E^*(M_n^*) = \frac{1}{2\pi\sqrt{h}} \int_{-\pi}^{\pi} K^2(u) du \sum_{i,j,r,l=1}^{m} \int_{-\pi}^{\pi} \widehat{\Gamma}_{ij,rl,T}(\lambda) \widehat{f}_{j,r,T}(\lambda) \overline{\widehat{f}}_{i,l,T}(\lambda) d\lambda + O_P(\sqrt{h}).
$$

Furthermore,

$$
\begin{aligned}
Var^*(M_n^*) =& \frac{h}{n^2} \sum_{i_1,j_1,r_1,l_1} \sum_{i_2,j_2,r_2,l_2} \sum_{s_1} \sum_{s_2} \int \int K_h^2(\lambda_1 - \lambda_{s_1}) K_h^2(\lambda_1 - \lambda_{s_2}) \widehat{\Gamma}_{i_1 j_1, r_1 l_1, T}(\lambda_1) \\
& \times \widehat{\Gamma}_{i_2 j_2, r_2 l_2, T}(\lambda_2) d\lambda_1 d\lambda 2 \\
& \times \Big\{ Cov^*(Y_{j_1,i_1}^*(\lambda_{s_1}), Y_{j_2,i_2}^*(\lambda_{s_2})) Cov^*(Y_{r_1,l_1}^*(\lambda_{s_1}), Y_{r_2,l_2}^*(\lambda_{s_2})) \\
& + Cov^*(Y_{j_1,i_1}^*(\lambda_{s_1}), Y_{r_2,l_2}^*(\lambda_{s_2})) Cov^*(Y_{r_1,l_1}^*(\lambda_{s_1}), Y_{j_2,i_2}^*(\lambda_{s_2})) \\
& + cum^*(Y_{j_1,i_1}^*(\lambda_{s_1}), Y_{r_1,l_1}^*(\lambda_{s_1}), Y_{j_2,i_2}^*(\lambda_{s_2}), Y_{r_2,l_2}^*(\lambda_{s_2})) \Big\}
\end{aligned}
$$

from which we get by straightforward calculations and using (6.3), that $Var^*(M_n^*) = O_P(n^{-1}h^{-1}) \to 0$. Thus and because

$$(6.6) \qquad \sum_{i,j,r,l=1}^{m} \widehat{\Gamma}_{ij,rl,T}(\lambda)\widehat{f}_{j,r,T}(\lambda)\overline{\widehat{f}}_{i,l,T}(\lambda) = \mathrm{tr}\{\widehat{\Gamma}_{\varphi,T}(\lambda)(\widehat{f}_T'(\lambda) \otimes \widehat{f}_T(\lambda))\} = \widehat{\mu}_n(\varphi),$$

we have (6.4).

We next establish (6.5). For this notice first that

$$Var^*(L_n^*) = \frac{h}{n^2} \int\int \sum_{i_1,j_1,r_1,l_1} \sum_{i_2,j_2,r_2,l_2} \sum_{\substack{s_1,s_2 \\ s_1 \neq s_2}} \sum_{\substack{q_1,q_2 \\ q_1 \neq q_2}} K_h(\lambda_1 - \lambda_{s_1})K_h(\lambda_1 - \lambda_{s_2})$$

$$\times K_h(\lambda_2 - \lambda_{q_1})K_h(\lambda_2 - \lambda_{q_2})\widehat{\Gamma}_{i_1j_1,r_1l_1,T}(\lambda_1)\widehat{\Gamma}_{i_2j_2,r_2l_2,T}(\lambda_2)d\lambda_1 d\lambda_2$$

$$\times \Big\{Cov^*(Y^*_{j_1,i_1}(\lambda_{s_1}), Y^*_{j_2,i_2}(\lambda_{q_1}))Cov^*(Y^*_{r_1,l_1}(\lambda_{s_2}), Y^*_{r_2,l_2}(\lambda_{q_2}))$$

$$+ Cov^*(Y^*_{j_1,i_1}(\lambda_{s_1}), Y^*_{r_2,l_2}(\lambda_{q_2}))Cov^*(Y^*_{r_1,l_1}(\lambda_{s_2}), Y^*_{j_2,i_2}(\lambda_{q_1}))$$

$$+ cum^*(Y^*_{j_1,i_1}(\lambda_{s_1}), Y^*_{r_1,l_1}(\lambda_{s_2}), Y^*_{j_2,i_2}(\lambda_{q_1}), Y^*_{r_2,l_2}(\lambda_{q_2}))\Big\}$$

$$= \frac{h}{n^2} \int\int \sum_{i_1,j_1,r_1,l_1} \sum_{i_2,j_2,r_2,l_2} \sum_{\substack{s,q \\ s \neq q}} K_h(\lambda_1 - \lambda_{s_1})K_h(\lambda_1 - \lambda_{s_2})$$

$$\times K_h(\lambda_2 - \lambda_{q_1})K_h(\lambda_2 - \lambda_{q_2})\widehat{\Gamma}_{i_1j_1,r_1l_1,T}(\lambda_1)\widehat{\Gamma}_{i_2j_2,r_2l_2,T}(\lambda_2)d\lambda_1 d\lambda_2$$

$$\times \Big\{\widehat{f}_{j_1,j_2,T}(\lambda_s)\overline{\widehat{f}}_{i_1,i_2,T}(\lambda_s)\widehat{f}_{r_1,r_2,T}(\lambda_q)\overline{\widehat{f}}_{l_1,l_2,T}(\lambda_q)$$

$$+ \quad \widehat{f}_{j_1,r_2,T}(\lambda_s)\overline{\widehat{f}}_{i_1,l_2,T}(\lambda_s)\widehat{f}_{r_1,j_2,T}(\lambda_q)\overline{\widehat{f}}_{l_1,i_2,T}(\lambda_q)\Big\} + o_P(1),$$

from which we get by straightforward algebra and because $\widehat{f}_T(\lambda) \to f_T(\lambda)$ and $\widehat{\Gamma}_{\varphi,T}(\lambda) \to \Gamma_{\varphi,T}(\lambda)$,

$$Var^*(L_n^*) \to \frac{1}{4\pi^2} \int \left(\int K(u)K(u+y)du\right)^2 dy \int \sum_{i_1,j_1,r_1,l_1} \sum_{i_2,j_2,r_2,l_2} \Gamma_{i_1j_1,r_1l_1,T}(\lambda)\Gamma_{i_2j_2,r_2l_2,T}(\lambda)$$

$$\times \Big\{f_{j_1,j_2,T}(\lambda)\overline{f}_{i_1,i_2,T}(\lambda)f_{r_1,r_2,T}(\lambda)\overline{f}_{l_1,l_2,T}(\lambda) + f_{j_1,r_2,T}(\lambda)\overline{f}_{i_1,l_2,T}(\lambda)$$

$$\times f_{r_1,j_2,T}(\lambda)\overline{f}_{l_1,i_2,T}(\lambda)\Big\}d\lambda.$$

Denote by $\sigma_T^2(\varphi)$ the right hand side of the last expression, then $\sigma^2(\varphi) = \sigma_T^2(\varphi)$ follows by Condition 1(iii) and relations similar to (6.6). To proceed with the proof of (6.5) let

$$W_n^*(\lambda_{s_1}, \lambda_{s_2}) = \frac{\sqrt{h}}{n} \sum_{i,j,r,l=1}^{m} \int K_h(\lambda - \lambda_{s_1})K_h(\lambda - \lambda_{s_2})\widehat{\Gamma}_{ij,rl,T}(\lambda)d\lambda Y^*_{j,i}(\lambda_{s_1})Y^*_{j,i}(\lambda_{s_2}),$$

and notice that

$$L_n^* = \sum_{1 \le s_1 < s_2 \le N} V_n^*(\lambda_{s_1}, \lambda_{s_2}) + \sum_{s \ne 0} W_n^*(\lambda_s, 0) + \sum_{s \ne 0} W_n^*(0, \lambda_s)$$

$$= L_{1,n}^* + L_{2,n}^* + L_{3,n}^*,$$

where $N = [n/2]$, $V_n^*(\lambda_{s_1}, \lambda_{s_2}) = W_n^*(\lambda_{s_1}, \lambda_{s_2}) + W_n^*(\lambda_{-s_1}, \lambda_{s_2}) + W_n^*(\lambda_{s_1}, \lambda_{-s_2}) + W_n^*(\lambda_{-s_1}, \lambda_{-s_2})$
and an obvious notation for $L_{j,n}^*$. Now, $L_{2,n}^* \to 0$ and $L_{3,n}^* \to 0$ in probability, since

$$E^*|L_{2,n}^*| \le \frac{\sqrt{h}}{n} \sum_{i,j,r,l} \sum_s \int K_h(\lambda - \lambda_s) |\widehat{\Gamma}_{ij,rl,T}(\lambda)| E^* |Y_{j,i}^*(\lambda_s) Y_{r,l}^*(0)|$$

$$= O_P(\sqrt{h})$$

and the same argument apply to $L_{3,n}^*$. To conclude the proof of (6.5) it remains to show
that $L_{1,n}^* \Rightarrow N(0, \sigma^2(\varphi))$. For this and by Theorem 2.1 of deJong (1987), it suffices to
show that

a) $\left(\max_{1 \le s_1 \le N} \sum_{s_2=1}^N E^*(V_n^*(\lambda_{s_1}, \lambda_{s_2}))^2 \right)/\sigma_n^2 \to 0$, and
b) $E^*(L_{1,n}^*)^4 / \sigma_n^4 \to 3$,

as $n \to \infty$, where $\sigma_n^2 = Var(L_{1,n}^*)$. Notice that $\sigma_n \to \sigma_T^2(\varphi) = O(1)$ by the same
arguments as those used to handle $Var^*(L_n^*)$.

Consider a). Since $s_1 \ne s_2$ and by the independence of the $Y_{ij,rl}^*(\lambda_s)$ we get by straight-
forward calculations that

$$\sum_{s_2=1}^N E^*(W_n^*(\lambda_{s_1}, \lambda_{s_2}))^2 = \frac{h}{n^2} \sum_{s_2} \sum_{i_1,j_1,r_1,l_1} \sum_{i_2,j_2,r_2,l_2} \int \int K_h(\lambda_1 - \lambda_{s_1}) K_h(\lambda_1 - \lambda_{s_2})$$

$$\times K_h(\lambda_2 - \lambda_{s_1}) K_h(\lambda_2 - \lambda_{s_2}) \widehat{\Gamma}_{i_1 j_1, r_1 l_1, T}(\lambda_1) \widehat{\Gamma}_{i_2 j_2, r_2 l_2, T}(\lambda_2) d\lambda_1 d\lambda_2$$

$$\times \widehat{f}_{j_1,i_1,T}(\lambda_{s_1}) \widehat{f}_{j_2,i_2,T}(\lambda_{s_1}) \widehat{f}_{r_1,l_1}(\lambda_{s_2}) \widehat{f}_{r_2,l_2}(\lambda_{s_2})$$

$$= O_P(n^{-1}h^{-1}),$$

which implies because $E^*(V_n^*(\lambda_{s_1}, \lambda_{s_2}))^2 \le 4 \sum_{g_1 \in \{\lambda_{s_1}, \lambda_{-s_1}\}} \sum_{g_2 \in \{\lambda_{s_2}, \lambda_{-s_2}\}} E^*(W_n^*(g_1, g_2))^2$
that $(\max_{1 \le s_1 \le N} \sum_{s_2=1}^N E^*(V_n^*(\lambda_{s_1}, \lambda_{s_2}))^2)/\sigma_n^2 = O_P(n^{-1}h^{-1}) \to 0$ in probability as $n \to \infty$.

To establish b) notice that by ignoring the asymptotically vanishing fourth order cumulant
term and using $E^*(V_n^*(\lambda_{s_1}, \lambda_{s_2}) V_n^*(\lambda_{\nu_1}, \lambda_{\nu_2})) = \delta_{s_1,\nu_1} \delta_{s_2,\nu_2} E(V_n^*(\lambda_{s_1}, \lambda_{s_2}))^2$ for $s_1 < s_2$ and
$\nu_1 < \nu_2$, that the fourth moment of $\sum_{1 \le s_1 < s_2 \le N} V_n^*(\lambda_{s_1}, \lambda_{s_2})$ equals

$$E^*\left(\sum_{1 \le s_1 < s_2 \le N} V_n^*(\lambda_{s_1}, \lambda_{s_2}) \right)^4 = 3 \sum_{1 \le s_1 < s_2 \le N} \sum_{1 \le \nu_1 < \nu_2 \le N} E^*(V_n^*(\lambda_{s_1}, \lambda_{s_2}))^2 E^*(V_n^*(\lambda_{\nu_1}, \lambda_{\nu_2}))^2$$

$$+ o(1)$$

$$= 3(\sigma_n^2)^2 + o(1),$$

which leads to the desired assertion.

$\square$

Proof of Theorem 4.1: Define

$$(6.7) \qquad h_r(\lambda) = \frac{f_r(\lambda)}{w(\lambda)} - 1; \quad \tilde{h}_r(\lambda) = \frac{\hat{f}_r(\lambda)}{w(\lambda)} - 1,$$

then we obtain from the proof of Theorem 3.1

$$(6.8) \quad \sqrt{N}\Big(T_N - \frac{1}{m}\sum_{r=1}^{m}\int h_r^2(\lambda)d\lambda\Big) = \sqrt{\frac{n}{m}}\sum_{r=1}^{m}\Big\{\int \tilde{h}_r^2(\lambda)d\lambda - \int h_r^2(\lambda)d\lambda\Big\}$$

$$\times (1 + o_p(1))$$

$$= \sqrt{\frac{n}{m}}\sum_{r=1}^{m}\Big\{\int \Big(\frac{\hat{f}_r(\lambda) - f_r(\lambda)}{w(\lambda)}\Big)^2 d\lambda$$

$$+ 2\int h_r(\lambda)\frac{\hat{f}_r(\lambda) - f_r(\lambda)}{w(\lambda)}d\lambda\Big\}(1 + o_p(1))$$

$$= (A_{n1} + A_{n2})(1 + o_p(1))$$

with an obvious definition of the quantities A_{n1}, A_{n2}. The term A_{n1} can be treated by similar methods as used in the proof of Theorem 3.1, which yield

$$(6.9) \qquad A_{n1} = O_p\Big(\frac{1}{\sqrt{nh}}\Big) = o_p(1).$$

The analysis of the term A_{n2} is more difficult and we obtain

$$(6.10) \qquad \frac{\sqrt{m}}{2}A_{n2} = \sqrt{n}\sum_{r=1}^{m}\int \frac{h_r(\lambda)}{w(\lambda)}\Big(\hat{f}_r(\lambda) - f_r(\lambda)\Big)d\lambda = B_{1n} + B_{2n},$$

where

$$(6.11) \qquad B_{1n} = \frac{1}{\sqrt{n}}\sum_{j}\int K_h(\lambda - \lambda_j)\sum_{r=1}^{m}\frac{h_r(\lambda)}{w(\lambda)}\Big(I_r(\lambda_j) - f_r(\lambda_j)\Big)d\lambda,$$

$$B_{2n} = \sqrt{n}\Big\{\frac{1}{n}\sum_{j}\int K_h(\lambda - \lambda_j)\sum_{r=1}^{m}\frac{h_r(\lambda)}{w(\lambda)}f_r(\lambda_j)d\lambda - \sum_{r=1}^{m}\int \frac{h_r(\lambda)}{w(\lambda)}f_r(\lambda)d\lambda\Big\}$$

$$(6.12)$$

$$= \sqrt{n}\Big\{\frac{1}{2\pi}\int\int K_h(\lambda - x)\sum_{r=1}^{m}\frac{h_r(\lambda)}{w(\lambda)}f_r(x)dxd\lambda - \sum_{r=1}^{m}\int \frac{h_r(\lambda)}{w(\lambda)}f_r(\lambda)d\lambda\Big\}$$

$$+ o(1)$$

$$= \sqrt{n}\sum_{r=1}^{m}\Big\{\int \frac{h_r(\lambda)}{w(\lambda)}\Big[\frac{1}{2\pi}\int K_h(\lambda - x)f_r(x)dx - f_r(\lambda)\Big]d\lambda\Big\} + o(1)$$

$$= \sqrt{n}\tilde{b}_h + o(1)$$

with an obvious definition of $\tilde{b}_h$. From (6.8), (6.9), (6.10), (6.11), (6.12) and the notation $b_h = 2\tilde{b}_h/sqrtm$ we therefore obtain

$$(6.13) \qquad \sqrt{N}\Big(T_n - \frac{1}{m}\sum_{r=1}^{m}\int h_r^2(\lambda)d\lambda - b_h\Big) = \frac{2}{\sqrt{m}}B_{1n} + o_p(1) ,$$

and it remains to consider the asymptotic distribution of the statistic B_{1n}. For this recall that by (3.3), $I_n(\lambda_j) = f^{1/2}(\lambda_j)U_n(\lambda_j)\overline{f^{1/2}}(\lambda_j) + R_{n,j}$, where the remainder is of order $O_p(1/\sqrt{n})$ uniformly with respect to j. Recall that $U_n(\lambda_j)$ has asymptotically a complex $W_c(1, I_m)$ Wishart distribution, and consequently

$$(6.14) \qquad I_n(\lambda_j) = W_n(\lambda_j) + R_{nj} ,$$

where the random variables $W_n(\lambda_j)$ are asymptotically $W_c(1, f(\lambda_j))$ distributed and independent [see Brockwell and Davis (1991), Proposition 11.7.3]. From Muirhead (1982) p. 90 and a similar argument as given in the Proof of Proposition 10.3.2 in Brockwell and Davis (1991) we have

$$E[e_r^T W_n(\lambda_j)e_r] = f_r(\lambda_j)(1 + o(1)) = f_r(\lambda_j)(1 + o(1)) ,$$

$$(6.15)$$

$$\mathrm{Cov}[e_r^T W_n(\lambda_i)e_r, e_s^T W_n(\lambda_j)e_s] = 2\delta_{ij}f_r(\lambda_j)f_s(\lambda_j)(1 + o(1)) ,$$

where $\delta_{ij} = 1$ if $i = j$ and $\delta_{i,j} = 0$ for $i \neq j$. For the moment we ignore the remainder and obtain with the notation $W_r(\lambda_j) = e_r^T W_n(\lambda_j)e_r$,

$$(6.16) \qquad \tilde{B}_{1n} = \frac{1}{\sqrt{n}}\sum_{j}\int K_h(\lambda - \lambda_j)\sum_{r=1}^{m}\frac{h_r(\lambda)}{w(\lambda)}\Big(W_r(\lambda_j) - f_r(\lambda_j)\Big)$$

the estimates

$$
\begin{aligned}
E[\tilde{B}_{1n}] &= o(1) \\
E[\tilde{B}_{1n}^2] &= \frac{1}{n}E\Big[\sum_{j}\int\int K_h(\lambda - \lambda_j)K_h(\mu - \lambda_j) \\
&\qquad \sum_{r,s=1}^{m}\Big[W_r(\lambda_j) - f_r(\lambda_j)\Big]\Big[W_s(\lambda_j) - f_r(\lambda_j)\Big]\frac{h_r(\lambda)h_s(\mu)}{w(\lambda)w(\mu)}d\mu d\lambda\Big] \cdot (1 + o(1)) \\
&= \frac{1}{2\pi}\sum_{r,s=1}^{m}\int\int\int K_h(\lambda - x)K_h(\mu - x)\frac{h_r(\lambda)h_s(\mu)}{w(\lambda)w(\mu)}2f_r(x)f_s(x)dxd\mu d\lambda \cdot (1 + o(1)) \\
&= 4\pi\int\frac{1}{w^2(x)}\Big(\sum_{r=1}^{m}h_r(x)f_r(x)\Big)^2 dx \cdot (1 + o(1)) \\
&= 4\pi\int\Big[\sum_{r=1}^{m}\frac{f_r(x)}{w(x)}\Big(\frac{f_r(x)}{w(x)} - 1\Big)\Big]^2 dx \cdot (1 + o(1)) \\
&= \alpha^2 \cdot (1 + o(1))
\end{aligned}
$$

with an obvious definition of α^2. The asymptotic normality of $\tilde{B}_{1n}$ now follows along the lines given in the proof of Proposition 10.3.2 and 11.7.3 in Brockwell and Davis (1991), that is

$$(6.17) \qquad\qquad\qquad \tilde{B}_{1n} \Rightarrow N(0, \alpha^2)$$

We finally show that B_{1n} and $\tilde{B}_{1n}$ are asymptotically equivalent, that is

$$(6.18) \qquad\qquad\qquad B_{1n} - \tilde{B}_{1n} = o_p(1).$$

For this we note that Lemma 3 in Dette and Spreckelsen (2003) holds also in the multivariate case considered here (this follows by the same arguments given by these authors observing that the assumption of normally distributed innovations is in fact not needed to establish the result). More precisely, if $R_{n,r,j} = e_r^T R_{n,j} e_r$, then we have

$$(6.19) \qquad\qquad E[R_{n,r,j}] = o(n^{-1})$$

$$\mathrm{Cov}(R_{n,r,i}, R_{n,r,j}) = \begin{cases} O(n^{-1}) & \text{if} \quad \lambda_j = \mp\lambda_i \\ o(n^{-1}) & \text{if} \quad \lambda_j \neq \mp\lambda_i \end{cases}$$

Observing (6.20) we obtain

$$
\begin{aligned}
E[(B_{n1} - \tilde{B}_{n1})^2] &= \frac{1}{n} E\Big[\Big(\sum_j \int K_h(\lambda - \lambda_j) \sum_{r=1}^{m} \frac{h_r(\lambda)}{w(\lambda)} R_{n,r,j}\Big)^2\Big] \\
&= \Big|\frac{1}{n} \sum_{i,j} \sum_{r,s=1}^{m} \int\int K_h(\lambda - \lambda_j) K_h(\mu - \lambda_i) \frac{h_r(\lambda) h_s(\mu)}{w(\lambda) w(\mu)} E[R_{n,r,i} R_{n,s,j}] d\lambda d\mu\Big| \\
&\leq \frac{1}{n} \sum_i \sum_{r,s=1}^{m} \int K_h(\lambda - \lambda_i) K_h(\mu - \lambda_i)\Big|\frac{h_r(\lambda) h_s(\mu)}{w(\lambda) w(\mu)}\Big| d\lambda d\mu \cdot O(\frac{1}{n}) \\
&\quad + \frac{1}{n^2} \sum_{i,j} \sum_{r,s=1}^{m} \int K_h(\lambda - \lambda_i) K_h(\lambda - \lambda_j) \frac{h_r(\lambda) h_s(\mu)}{w(\lambda) w(\mu)} d\lambda d\mu \cdot o(1) \\
&= O\Big(\frac{1}{nh}\Big) + o(1) = o(1),
\end{aligned}
$$

and (6.18) is a consequence of Markov's inequality. Consequently, the assertion of the theorem follows from (6.13), (6.17) and (6.18). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

References

[1] Ahn, S. K. and Reinsel, G. C. (1988). Nested reduced-rank autoregressive models for multiple time series. *Journal of the American Statistical Association*, 83, 849-856.

[2] Beltrão, K. I. and Bloomfield, P. (1987). Determining the bandwidth of a kernel spectrum estimate. *Journal of Time Series Analysis*, 8, 21-38.

[3] Beran, R. (1986). Simulated power functions. *Annals of Statistics*, 14, 151-173.

[4] Berger, J.O. and Delampady, M. (1987). Testing precise hypotheses. *Statist. Sci.*, 2, 317-352.

[5] Brockwell, P. and Davis, R.A. (1991). *Time Series: Theory and Methods, 2nd Ed.*, Springer, New York.

[6] Caido, J., Crato, N. and Peña, D. (2006) A periodogram-based metric for time series classification. *Comput. Statist. Data Anal.*, 50, 2668-2684.

[7] Carmona, R.A. and Wang, A. (1996). Comparison tests for the spectra of dependent multivariate time series. *Stochastic Modeling in Physical Oceanography*, (ed. R.J. Adler, P. Müller, B.L. Rozovskii), 69-88. Birkhäuser Boston Inc.

[8] Coates, D. S. and Diggle P. J. (1986). Test for computing two estimated spectral densities. *Journal of Time Series Analysis.* 7, 7-20.

[9] Dahlhaus R. (1997). Fitting time series models to nonstationary processes. *The Annals of Statistics* **25**, 1-37.

[10] Dahlhaus R. (2000). *A likelihood approximation for locally stationary processes*, The Annals of Statistics **28**, 1762-1794.

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

[12] Dargahi-Noubary, G.R. and Laycock, P.J. (1981). Spectral ratio discriminants and information theory. *J. Time Series Anal.*, 2, 70-86.

[13] Dette, H. and Spreckelsen, I. (2003). A note on a specification test for time series models based on spectral density estimation. *Scandinavian Journal of Statistics*, 30, 481–491.

[14] De Souza, P. and Thomson, P. J. (1982). LPC distance measures and statistical tests with particular reference to likelihood ratio. *IEEE Transactions Accoustics Speech Signal Process*, 30, 304-315.

[15] Jenkins, G. M. (1961). General considerations in the analysis of spectra. *Technometrics*, 3, 133-166.

[16] deJong, P. (1987). A central limit theorem for generalized qudratic forms, *Probability Theory and Related Fields*, 75, 261-277.

[17] Diggle, P. J. and Fisher, N. I. (1991). Nonparametric comparison of cumulative periodograms. *Applied Statistics*, 40, 423-434.

[18] Eichler, M. (2008). Testing nonparametric and semiparametric hypotheses in vector stationary processes. *Journal of Multivariate Analysis*, 99, 968-1009.

[19] Franke, J. and Härdle, W. (1992). On bootstrapping kernel spectral estimates. *The Annals of Statistics*, **20**, 121-145.

[20] Gómez, R. M. P. and Drouiche, K. (2002). A test of homogeneity for autoregressive processes. *International Journal of Adaptive Control and Sgnal Processing*, 16, 231-242.

[21] Guo, J. H. (1999). A nonparametric test for the parallelism of two first-order autoregressive processes. *Australian and New Zeland Journal of Statistics*, 41, 59-65.

[22] Hurvich, C. M. and Zeger, S. L. (1987). Frequency domain bootstrap methods for time series. Technical Report 87-115, Graduate School of Business Administration, New York Univ.

[23] Jenkins, G. M., (1961). General considerations in the analysis of spectra. *Technometrics*, 3, 133-166.

[24] Kakizawa, Y., Shumway, R.H. and Taniguchi, M. (1998). Discrimination and clustering for multivariate time series. *J. Amer. Statist. Assoc.*, 93, 328-340.

[25] Linton, O. and Fan, Y. (2003). Some higher order theory for a consistent nonparametric model specification test. *Journal of Statistical Planning and Inference*, 109, 125-154.

[26] Marahaj, E. A. (2002). Comparison of non-stationary time series in the frequency domain. *Computational Statistics and Data Analysis*, 40, 131-141.

[27] Matsuda, Y. and Yahima, Y. (2004). On testing for separable correlations of multivariate time series. *J. Time Series Anal.*, 25, 501-528.

[28] Kreiss, J.-P. and Paparoditis, E. (2003). Autoregressive Aided Periodogram Bootstrap for Time Series. *The Annals of Statistics*, 31, 1923-1955.

[29] Muirhead, R. J. (1982). *Aspects of multivariate statistical theory*. Wiley Series in Probability and Mathematical Statistics. John Wiley: New York.

[30] Opsomer, J.D. and Francisco-Fernández, M. (2008). Finding local departures from a parametric model using nonparametric regression. *Statistical Papers*, to appear.

[31] Paparoditis, E. and Politis, D. N. (1999). The local bootstarp for periodogram statistics. *Journal of Time Series Analysis*, 20, 193-222.

[32] Paparoditis, E. (2002). Frequency domain bootstrap for time series. In: *Empirical Process Techniques for Dependent Data* (Ed. by H. Dehling, T. Mikosch and M. Sørensen), Birkhäuser, Boston, 365-381.

[33] Priestley, M. P. (1981). *Spectral Analysis and Time Series*. Academic Press, New-York.

[34] Reinsel, G. C. (2003). *Elements of Multivariate Time Series*. Springer, New-York.

[35] Shumway, R. H. (1982). Discriminant analysis for time series. Classification, pattern recognition and reduction of dimensionality. *Hanbook of Statistics*, 2. North-Holland, Amsterdam.

[36] Swanepoel J. W. H. and van Wyk, J. W. J. (1986). The comparison of two spectral density functions using the bootstrap. *Journal of Statistical Computation and Simulation*, 24, 271-282.

[37] Taniguchi, M. and Kakizawa, Y. (2000). *Asymptotic Theory of Statistical Inference for Time Series*, Springer-Verla: New York.

[38] Taniguchi, M. and Kondo, M. (1993). Non-parametric approach in time series analysis. *Journal of Time Series Analysis*, 14, 397-408.

[39] Taniguchi, M., Puri, M. L. and Kondo, M. (1996). Nonparametric approach for nongaussian vector stationary processes. *Journal of Multivariate Analysis* 56, 259-283.

[40] Timmer, J., Lauk, M., Vach, W. and Lucking, C. H. (1999). A test for a difference between spectral peak frequencies. *Computational Statistics and Data Analysis*, 30, 45-55.

[41] Yajima, Y. and Matsuda, Y. (2008). On nonparametric and semiparametric testing for multivariate linear time series. RePEc [http://oai.repec.openlib.org] (Germany)

[42] Zhang, G. and Taniguchi, M. (1994). Discriminant analysis for stationary vector time series. *J. Time Series Anal.*, 15, 117-126.

ρ	δ	$n = 128$		$n = 512$		$n = 1024$	
		h= 0.15	h=0.20	h=0.10	h=0.14	h=0.10	h= 0.12
0.9	0.0	0.065	0.078	0.052	0.056	0.055	0.055
	0.1	0.674	0.702	0.996	0.998	1.000	1.000
	-0.1	0.496	0.512	0.990	0.991	1.000	1.000
0.5	0.0	0.092	0.098	0.055	0.058	0.055	0.058
	0.1	0.302	0.364	0.766	0.890	0.966	0.970
	-0.1	0.201	0.216	0.546	0.583	0.888	0.894
0.0	0.0	0.093	0.094	0.058	0.065	0.054	0.055
	0.1	0.274	0.320	0.648	0.724	0.912	0.928
	-0.1	0.191	0.188	0.422	0.446	0.760	0.782
-0.5	0.0	0.094	0.093	0.058	0.062	0.056	0.057
	0.1	0.298	0.366	0.754	0.806	0.972	0.973
	-0.1	0.226	0.221	0.560	0.588	0.880	0.896
-0.9	0.0	0.079	0.081	0.046	0.048	0.054	0.054
	0.1	0.686	0.722	0.998	0.999	1.000	1.000
	-0.1	0.504	0.508	0.982	0.985	1.000	1.000

Table 1: Empirical rejection probabilities ($\alpha = 0.05$) of the test $T_{D,n}$ over 500 replications of the bivariate process (5.1) for different sample sizes n, values of the bandwidth h and of the process parameters ρ and δ.

.

| ρ	δ	h= 0.14			h=0.18		
		$T_{D,n}$	$S_{D,n}^{(1)}$	$S_{D,n}^{(2)}$	$T_{D,n}$	$S_{D,n}^{(1)}$	$S_{D,n}^{(2)}$
0.9	0.0	0.062	0.070	0.052	0.063	0.086	0.060
	0.1	0.940	0.962	0.847	0.964	0.972	0.899
	-0.1	0.844	0.698	0.843	0.852	0.726	0.848
0.5	0.0	0.068	0.072	0.071	0.071	0.071	0.075
	0.1	0.567	0.639	0.165	0.618	0.675	0.208
	-0.1	0.295	0.132	0.317	0.318	0.133	0.331
0.0	0.0	0.069	0.068	0.085	0.070	0.071	0.074
	0.1	0.486	0.537	0.140	0.530	0.593	0.140
	-0.1	0.242	0.108	0.268	0.628	0.108	0.289
-0.5	0.0	0.072	0.080	0.082	0.081	0.082	0.076
	0.1	0.550	0.629	0.164	0.598	0.670	0.194
	-0.1	0.302	0.144	0.345	0.144	0.148	0.361
-0.9	0.0	0.056	0.056	0.064	0.061	0.066	0.075
	0.1	0.956	0.967	0.839	0.964	0.972	0.898
	-0.1	0.858	0.688	0.868	0.852	0.731	0.874

Table 2: Empirical rejection probabilities ($\alpha = 0.05$) of the tests $T_{D,n}$, $S_{D,n}^{(1)}$ and $S_{D,n}^{(2)}$ over 500 replications of the bivariate process (5.1) for $n = 256$ and different sample sizes, values of the bandwidth h and of the process parameters ρ and δ.

RUHR-UNIVERSITÄT BOCHUM, FAKULTÄT FÜR MATHEMATIK, 44780 BOCHUM, GERMANY
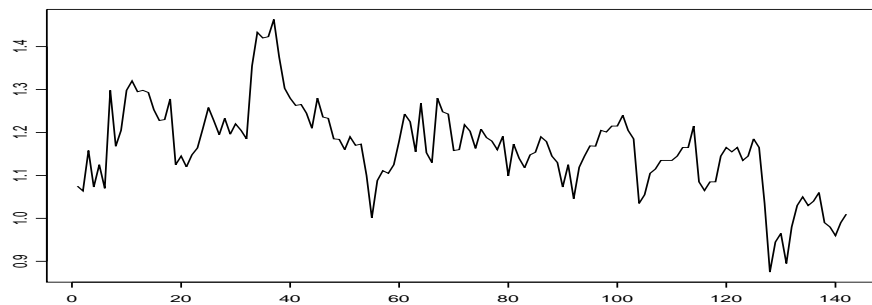
UNIVERSITY OF CYPRUS, DEPARTMENT OF MATHEMATICS AND STATISTICS, P.O.BOX 20537, CY-1678 NICOSIA, CYPRUS
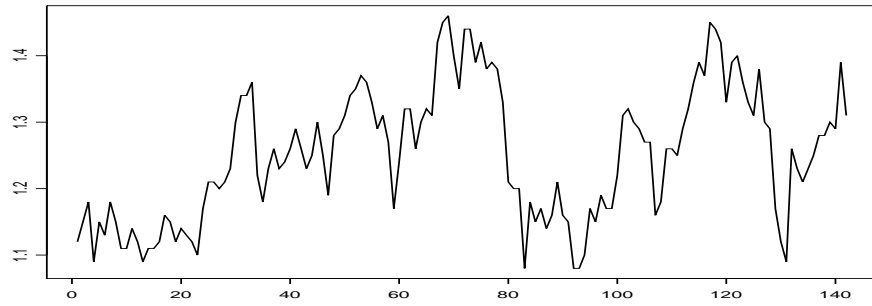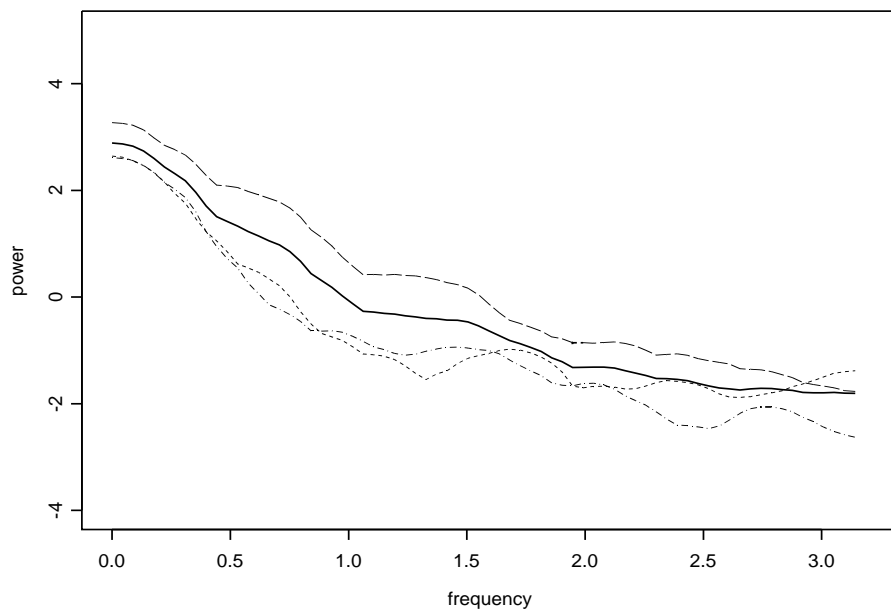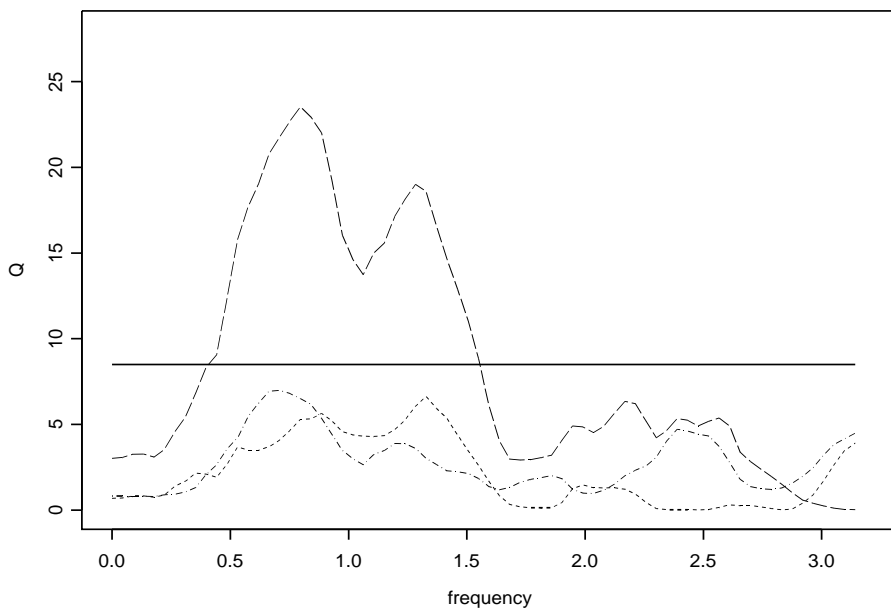
(a)
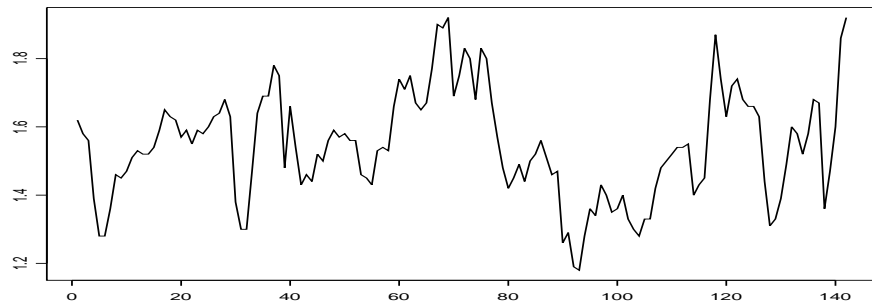
(b)

(c)

FIGURE 1. Plot of US grain price data: (a) corn prices, (b) wheat prices, (c) ray prices.

(a)

(b)

FIGURE 2. (a) Plot of estimated spectral densities (log-scale) for the grain
price data set and (b) plot of the statistic $Q^2_{r,n}(\lambda_j)$. The dashed and dotted
line refers to corn prices, the dashed line to wheat prices and the dotted
line to ray prices. The solid line in part (a) refers to the estimated pooled
spectral density and in part (b) to the bootstrap estimate of the upper 5%-
percentage point of the distribution of the statistic $M_n =
\max_{1 \le r \le m} \max_{0 \le \lambda_j \le \pi} Q^2_{r,n}(\lambda_j)$.