SFB 823

Comparing spectral densities of stationary time series with unequal sample sizes

Philip Preuß, Thimo Hildebrandt

Nr. 34/2011

Discussion Paper

SFB 823

Comparing spectral densities of stationary time series with unequal sample sizes

Philip Preuß, Thimo Hildebrandt
Ruhr-Universität Bochum
Fakultät für Mathematik
44780 Bochum
Germany
email: philip.preuss@ruhr-uni-bochum.de
thimo.hildebrandt@ruhr-uni-bochum.de
FAX: +49 2 34 32 14 559

September 7, 2011

Abstract

This paper deals with the comparison of several stationary processes with unequal sample sizes. We provide a detailed theoretical framework on the testing problem for equality of spectral densities in the bivariate case, but also present the generalization to the m dimensional case and to other statistical applications like testing for zero correlation or clustering of time series data with different length. We prove asymptotic normality of an appropriately standardized version of the test statistic both under the null and the alternative and investigate the finite sample properties of our method in a comprehensive simulation study. Furthermore we apply our approach to cluster financial time series data with different sample length.

AMS subject classification: 62M10, 62M15, 62G10
Keywords and phrases: spectral density, integrated periodogram, cluster analysis, time series, stationary process, unequal length

1 Introduction

The comparison and clustering of different time series is an important topic in statistical data analysis and has various application in fields like economics, marketing, medicine and physics, among many others. Examples are the grouping of stocks in several categories for portfolio selection in finance or the identification of similar birth and death rates in population studies. One approach to identify similarities or dissimilarities between two stationary processes is to compare the spectral densities of both time series, which directly yields to the testing problem for equality of spectral densities in multivariate time series data. This problem has found considerable interest in the literature [see for example Jenkins (1961) or De Souza and Thomson (1982) for some early results], but in the nonparametric situation nearly all proposed procedures are only reasoned by simulation studys or heuristic proofs, see Coates and Diggle (1986), Pötscher and Reschenhofer (1988), Diggle and Fisher (1991) and Maharaj (2002) among many others. Most recently Eichler (2008), Dette and Paparoditis (2009), Dette et al. (2010) and Dette and Hildebrandt (2011) provided mathematical details for the above testing problem using different L_2-type statistics, but nevertheless in all mentioned articles it is always required that the different time series have the same length, which is typically not the case in practice. Caiado and Pena (2009) considered different metrics for the comparison of time series with unequal sample sizes in a simulation study and Jentsch and Pauly (2011) provided a theoretical result, which however does not yield a consistent test as it was also pointed out by the authors.

This paper generalizes the approach of Dette et al. (2010) to the case of unequal sample sizes and yields a consistent test for the equalness of spectral densities for time series with different length. For the sake of brevity we will focus on the case of two (not necessarily independent) stationary processes, but the results can be easily extended to the case of an m dimensional process [see Remark 2.5]. Basically we want to estimate the L_2-distance

$$D^2 := \frac{1}{4\pi} \int_{-\pi}^{\pi} (f_{11}(\lambda) - f_{22}(\lambda))^2 d\lambda \tag{1.1}$$

where $f_{11}(\lambda)$ and $f_{22}(\lambda)$ are the spectral densities of the first and the second process respectively. Under the null hypothesis

$$\mathrm{H}_0 : \quad f_{11}(\lambda) = f_{22}(\lambda) \tag{1.2}$$

the distance D^2 equals zero while it is strictly positive if $f_{11}(\lambda) \neq f_{22}(\lambda)$ for $\lambda \in A$, where A is a subset of $[-\pi, \pi]$ with positive Lebesgue measure. We will estimate D^2 by sums of the (squared) periodogram, where the sum goes over the Fourier coefficents of the smaller time series. Asymptotic normality both under the null and the alternative will be derived and since the variance terms can be easily estimated also under the alternative, asymptotic confidence intervalls and a precise hypothesis test can be constructed next to the test for (1.2) [see Remark 2.2]. Furthermore our approach has much wider application like testing for no correlation, discriminant analysis or clustering of time series with unequal length [see Remark 2.3] and a simulation study will indicate that some of our assumptions are in fact not necessary (for example our method seems to work also for Long Memory processes).

The remainder of the paper is organized as follows. In section 2 we will introduce the necessary notations and derive the asymptotic distribution of our test statistic. In section 3 we will provide a comprehensive simulation study and an application to real-world data, while all technical details are deferred to an appendix in section 4.

2 The test statistic

Let $n_1, n_2 \in I\!N$ with $n_1 \leq n_2$ and consider the two stationary time series

$$X_t^{(1)} = \sum_{l=-\infty}^{\infty} \psi_l^{(1)} Z_{t-l}^{(1)} \quad t = 1, ..., n_1 \tag{2.1}$$

$$X_t^{(2)} = \sum_{l=-\infty}^{\infty} \psi_l^{(2)} Z_{t-l}^{(2)} \quad t = 1, ..., n_2 \tag{2.2}$$

where the $Z_t^{(j)}$ are independent and identically standard normal distributed for $j = 1, 2$ and

$$\mathbb{E}(Z_{t_1}^{(1)} Z_{t_2}^{(2)}) = \begin{cases} \rho & \text{if } t_2 = \lfloor t_1 q_{n_1,n_2} \rfloor - \lfloor q_{n_1,n_2} - 1 \rfloor \\ 0 & \text{else} \end{cases} \tag{2.3}$$

where $q_{n_1,n_2} = \frac{n_2}{n_1}$ and $\rho \in [0, 1]$. This roughly speaking means that changes in the time series with less observations influence the more frequently observed series but not vice versa, which is for example the case if interest rates and stock returns are compared. Throughout the paper we also assume that the technical condition

$$\sum_{l=-\infty}^{\infty} \psi_l^{(j)} |l|^{\alpha} < \infty \tag{2.4}$$

is satisfied for an $\alpha > 1/2$ ($j = 1, 2$). Note that the assumption of Gaussianity is only imposed to simplify technical arguments and that the results can be easily extended to more general independent and indentically distributed innovations $Z_t^{(j)}$ [see Remark 2.6]. Furthermore innovations with variances different to 1 can be included by choosing other coefficents $\psi_l^{(j)}$.

We define the spectral densities $f_{jj}(\lambda)$ ($j = 1, 2$) by

$$f_{jj}(\lambda) := \frac{1}{2\pi} | \sum_{l=-\infty}^{\infty} \psi_l^{(j)} \exp(-i\lambda l)|$$

and the cross-spectra $f_{12}(\lambda)$ and $f_{21}(\lambda)$ through

$$f_{12}(\lambda) : = \frac{\rho}{2\pi} \sum_{l,m=-\infty}^{\infty} \psi_l^{(1)} \psi_m^{(2)} \exp(-i\lambda(l - m))$$

and

$$f_{21}(\lambda) : = \overline{f_{12}(\lambda)}.$$

An unbiased (but not consistent) estimator for $f_{jj}(\lambda)$ is given by the periodogram

$$I_j(\lambda) := \frac{1}{2\pi n_j}\Big|\sum_{t=1}^{n_j} X_t^{(j)} \exp(-i\lambda t)\Big|^2 \tag{2.5}$$

and although the periodogram does not estimate the spectral density consistently, a Riemann-sum over the Fourier coefficents of an exponentiated periodogram is (up to a constant) a consistent estimator for the corresponding integral over the exponentiated spectral density. For example similar arguments as in the proof of Theorem 2.1 in Dette et al. (2010) yield that

$$\hat{D}_{1,n_1} := \frac{1}{n_1}\sum_{k=1}^{\lfloor \frac{n_1}{2} \rfloor} I_1^2(\lambda_{1,k}) \xrightarrow{P} \frac{1}{2\pi}\int_{-\pi}^{\pi} f_{11}^2(\lambda)d\lambda =: D_1 \tag{2.6}$$

where $\lambda_{1,k} := \frac{2\pi k}{n_1}$ $(k = 1, ..., \lfloor \frac{n_1}{2} \rfloor)$ are the Fourier coefficents of the smaller time series $X_t^{(1)}$. If we can show that

$$\hat{D}_{2,n_1} := \frac{1}{n_1}\sum_{k=1}^{\lfloor \frac{n_1}{2} \rfloor} I_2^2(\lambda_{1,k}) \xrightarrow{P} \frac{1}{2\pi}\int_{-\pi}^{\pi} f_{22}^2(\lambda)d\lambda =: D_2 \tag{2.7}$$

and

$$\hat{D}_{12,n_1} := \frac{1}{n_1}\sum_{k=1}^{\lfloor \frac{n_1}{2} \rfloor -1} I_1(\lambda_{1,k})I_2(\lambda_{1,k+1}) \xrightarrow{P} \frac{1}{4\pi}\int_{-\pi}^{\pi} f_{11}(\lambda)f_{22}(\lambda)d\lambda =: D_{12}, \tag{2.8}$$

we can construct an consistent estimator for D^2 through

$$\hat{D}_{n_1}^2 := \frac{1}{2}(\hat{D}_{1,n_1} + \hat{D}_{2,n_1}) - 2\hat{D}_{12,n_1}. \tag{2.9}$$

Although (2.7) looks very much like (2.6), note that the convergence in (2.7) is different since the coefficents $\lambda_{1,k}$ are not necessarily the Fourier coefficents of the time series $X_t^{(2)}$. Nevertheless the convergence in (2.6) - (2.8) will be implied by the following theorem.

Theorem 2.1 *If $f_{11}(\lambda)$, $f_{22}(\lambda)$ and $f_{12}(\lambda)$ are Hölder continuous of order $L > 1/2$ and*

$$\frac{n_2}{n_1} \to Q \tag{2.10}$$

for a $Q \in \mathbb{R}$, then as $n_1 \to \infty$

$$\sqrt{n_1}\begin{pmatrix} \hat{D}_{1,n_1} - D_1 \\ \hat{D}_{12,n_1} - D_{12} \\ \hat{D}_{2,n_1} - D_2 \end{pmatrix} \xrightarrow{D} N(0,\Sigma)$$

with

$$\Sigma = \frac{1}{\pi}\begin{pmatrix} \Sigma_{11} & \Sigma_{12} & \Sigma_{13} \\ \Sigma_{12} & \Sigma_{22} & \Sigma_{23} \\ \Sigma_{13} & \Sigma_{23} & \Sigma_{33} \end{pmatrix}$$

and

$$\begin{aligned}
\Sigma_{11} &= 5\int_{-\pi}^{\pi} f_{11}^4(\lambda)d\lambda \\
\Sigma_{12} &= \int_{-\pi}^{\pi} f_{11}^3(\lambda)f_{22}(\lambda)d\lambda + \int_{-\pi}^{\pi} f_{11}^2(\lambda)|f_{12}(\lambda)|^2 \\
\Sigma_{13} &= \int_{-\pi}^{\pi} f_{12}^2(\lambda)f_{21}^2(\lambda) + 4\int_{-\pi}^{\pi} f_{11}(\lambda)|f_{12}(\lambda)|^2 f_{22}(\lambda) \\
\Sigma_{22} &= \frac{3}{4}\int_{-\pi}^{\pi} f_{11}^2(\lambda)f_{22}^2(\lambda)d\lambda + \frac{1}{2}\int_{-\pi}^{\pi} f_{11}(\lambda)|f_{12}(\lambda)|^2 f_{22}(\lambda)d\lambda \\
\Sigma_{23} &= \int_{-\pi}^{\pi} f_{11}(\lambda)f_{22}^3(\lambda)d\lambda + \int_{-\pi}^{\pi} f_{22}^2(\lambda)|f_{12}(\lambda)|^2 \\
\Sigma_{33} &= 5\int_{-\pi}^{\pi} f_{22}^4(\lambda)d\lambda.
\end{aligned}$$

Although condition (2.10) imposes some restrictions on the growth rate of n_1 and n_2, it is not very restrictive, since in practice there usually occur situations where even $n_2 = Qn_1$ holds for a $Q \in \mathbb{N}$ (if for example daily data are compared with monthly data) and on the other hand this condition needs only to be satisfied in the limit.

From Theorem 2.1 it now follows by a straightforward application of the Delta-Method that

$$\sqrt{n_1}(\hat{D}_{n_1}^2 - D^2) \xrightarrow{D} N(0, \sigma^2) \tag{2.11}$$

where

$$\sigma^2 := \frac{1}{\pi}\left\{\frac{\Sigma_{11} + \Sigma_{33}}{4} + 4\Sigma_{22} + \frac{\Sigma_{13}}{2} - 2\Sigma_{12} - 2\Sigma_{23}\right\},$$

which becomes

$$\sigma_{\mathrm{H}_0}^2 = \frac{3}{2\pi}\int_{-\pi}^{\pi} f_{11}^4(\lambda)d\lambda + \frac{1}{2\pi}\int_{-\pi}^{\pi} |f_{12}|^4 d\lambda$$

under H_0. To obtain a consistent estimator under the null hypothesis we define

$$\begin{aligned}
I_{12}(\lambda) &:= \frac{1}{2\pi\sqrt{n_1 n_2}} \sum_{p_1=1}^{n_1} X_{p_1}^{(1)} \exp(-i\lambda p_1) \sum_{p_2=1}^{n_2} X_{p_2}^{(2)} \exp(i\lambda p_2) \\
I_{21}(\lambda) &:= \overline{I_{12}(\lambda)}
\end{aligned}$$

and analogous to the proof of Theorem 2.1 it can be shown that

$$(2.12) \qquad \frac{1}{n_1}\sum_{k=1}^{\lfloor \frac{n_1}{2} \rfloor}\Big(I_1^4(\lambda_{1,k})+I_2^4(\lambda_{1,k})\Big) \xrightarrow{P} \frac{6}{\pi}\int_{-\pi}^{\pi}(f_{11}^4(\lambda)+f_{22}^4(\lambda))d\lambda$$

and

$$(2.13) \qquad \frac{1}{2n_1}\sum_{k=1}^{\lfloor n_1/2 \rfloor -1} I_{12}^2(\lambda_{1,k})I_{21}^2(\lambda_{1,k+1}) \xrightarrow{P} \frac{1}{2\pi}\int_{-\pi}^{\pi}|f_{12}|^4 d\lambda.$$

Now (2.12) and (2.13) imply that under the null hypothesis

$$\hat{\sigma}^2_{\mathrm{H}_0} := \frac{1}{4n_1}\sum_{k=1}^{\lfloor \frac{n_1}{2} \rfloor}\Big(I_1^4(\lambda_{1,k})+I_2^4(\lambda_{1,k})\Big) + Re\Big(\frac{1}{2n_1}\sum_{k=1}^{\lfloor n_1/2 \rfloor -1} I_{12}^2(\lambda_{1,k})I_{21}^2(\lambda_{1,k+1})\Big) \xrightarrow{P} \sigma^2_{\mathrm{H}_0}.$$

and therefore an asymptotic niveau-α-test for (1.2) is given by: reject (1.2) if

$$(2.14) \qquad \sqrt{n_1}\frac{\hat{D}^2_{n_1}}{\sqrt{\hat{\sigma}^2_{\mathrm{H}_0}}} > u_{1-\alpha},$$

where $u_{1-\alpha}$ denotes the $(1-\alpha)$ quantile of the standard normal distribution. By using (2.11) we obtain that this test has asymptotic power

$$(2.15) \qquad \Phi\Big(\sqrt{n_1}\frac{D^2}{\sigma} - \frac{\sqrt{\hat{\sigma}^2_{\mathrm{H}_0}}}{\sigma}u_{1-\alpha}\Big),$$

where Φ is the distribution function of the standard normal distribution. From (2.15) it follows that the test (2.14) has asymptotic power one for all alternatives with $D^2 > 0$.

Remark 2.2
In order to estimate the variance σ^2 in (2.11) also under the alternative, we define

$$(2.16) \qquad \hat{\Sigma}_{11} = \frac{5\pi}{6n_1}\sum_{k=1}^{\lfloor \frac{n_1}{2} \rfloor} I_1^4(\lambda_{1,k})$$

$$(2.17) \qquad \hat{\Sigma}_{12} = \frac{2\pi}{3n_1}\sum_{k=1}^{\lfloor \frac{n_1}{2} \rfloor -1} I_1^3(\lambda_{1,k})I_2(\lambda_{1,k+1}) + \frac{\pi}{n_1}\sum_{k=1}^{\lfloor \frac{n_1}{2} \rfloor -1} I_1^2(\lambda_{1,k})|I_{12}(\lambda_{1,k+1})|^2$$

$$(2.18) \qquad \hat{\Sigma}_{13} = \frac{\pi}{n_1}\sum_{k=1}^{\lfloor \frac{n_1}{2} \rfloor -1} |I_{12}(\lambda_{1,k})|^2|I_{12}(\lambda_{1,k+1})|^2 + \frac{8\pi}{n_1}\sum_{k=1}^{\lfloor \frac{n_1}{2} \rfloor -2} I_1(\lambda_{1,k})|I_{12}(\lambda_{1,k+1})|^2 I_2(\lambda_{1,k+2})$$

$$\hat{\Sigma}_{22} = \frac{3\pi}{4n_1} \sum_{k=1}^{\lfloor \frac{n_1}{2} \rfloor - 1} I_1^2(\lambda_{1,k}) I_2^2(\lambda_{1,k+1}) + \frac{\pi}{n_1} \sum_{k=1}^{\lfloor \frac{n_1}{2} \rfloor - 2} I_1(\lambda_{1,k}) |I_{12}(\lambda_{1,k+1})|^2 I_2(\lambda_{1,k+2}) \tag{2.19}$$

$$\hat{\Sigma}_{23} = \frac{2\pi}{3n_1} \sum_{k=1}^{\lfloor \frac{n_1}{2} \rfloor - 1} I_1(\lambda_{1,k}) I_2^3(\lambda_{1,k+1}) + \frac{\pi}{n_1} \sum_{k=1}^{\lfloor \frac{n_1}{2} \rfloor - 1} I_2^2(\lambda_{1,k}) |I_{12}(\lambda_{1,k+1})|^2 \tag{2.20}$$

$$\hat{\Sigma}_{33} = \frac{5\pi}{6n_1} \sum_{k=1}^{\lfloor \frac{n_1}{2} \rfloor} I_2^4(\lambda_{1,k}). \tag{2.21}$$

A consistent estimator for σ^2 is now given by

$$\hat{\sigma}^2 := \frac{1}{\pi} \left\{ \frac{\hat{\Sigma}_{11} + \hat{\Sigma}_{33}}{4} + 4\hat{\Sigma}_{22} + \frac{\hat{\Sigma}_{13}}{2} - 2\hat{\Sigma}_{12} - 2\hat{\Sigma}_{23} \right\} \tag{2.22}$$

which enables us to construct asymptotic $(1-\alpha)$ confidence intervals for D^2 through

$$\Big[0, \hat{D}_{n_1}^2 + \sqrt{\frac{\hat{\sigma}^2}{n_1}} u_{1-\alpha}\Big].$$

Since it is more reasonable to estimate a normalized measure, we also defne the alternative distance

$$R^2 := \frac{2D^2}{D_1 + D_2}$$

which can be estimated by

$$\hat{R}_{n_1}^2 := \frac{2\hat{D}_{n_1}^2}{\hat{D}_{1,n_1} + \hat{D}_{2,n_1}}.$$

From Theorem 2.1 and a straightforward application of the delta method, it follows that

$$\sqrt{n_1}(\hat{R}_{n_1}^2 - R^2) \xrightarrow{D} N(0, \sigma_1^2) \tag{2.23}$$

with

$$\sigma_1^2 := \frac{16}{(D_1 + D_2)^2} \left(\frac{D_{12}^2}{(D_1 + D_2)^2} (\Sigma_{11} + 2\Sigma_{13} + \Sigma_{33}) - \frac{2D_{12}}{D_1 + D_2} (\Sigma_{12} + \Sigma_{23}) + \Sigma_{22} \right)$$

and by using (2.16) - (2.22) and Theorem 2.1, a consistent estimator $\hat{\sigma}_1^2$ for σ_1^2 can be easily constructed, which yields the following asymptotic $(1-\alpha)$ confidence intervals for R^2

$$\Big[0, \hat{R}_{n_1}^2 + \sqrt{\frac{\hat{\sigma}_1^2}{n_1}} u_{1-\alpha}\Big].$$

Furthermore (2.23) provides an asymptotic level α test for the so called *precise hypothesis*

$$H_0 : R^2 > \varepsilon \quad \text{versus} \quad H_1 : R^2 \leq \varepsilon\,, \tag{2.24}$$

where $\varepsilon > 0$ [see Berger and Delampady (1987)]. This hypothesis is of interest, because spectral densities of time series in real-world applications are usually never exactly equal and a more realistic question is then to ask, if the processes have approximately the same spectral measure. An asymptotic level α test for (2.24) is obtained by rejecting the null hypothesis, whenever

$$\hat{R}^2_{n_1} - \varepsilon < \frac{\hat{\sigma}_1}{\sqrt{n_1}} u_\alpha . \tag{2.25}$$

Remark 2.3
Theorem 2.1 can also be used for a cluster and a discriminant analysis of time series data with different length, since it yields an estimator for the distance measure $d(f_{11}, f_{22})$, where

$$d(f,g) = \left(1 - \frac{2\int_{-\pi}^{\pi} f(\lambda)g(\lambda)d\lambda}{\int_{-\pi}^{\pi} f^2(\lambda)d\lambda + \int_{-\pi}^{\pi} g^2(\lambda)d\lambda}\right)^{1/2},$$

which can take values between 0 and 1. A value close to 0 indicates some kind of similarities between two processes, whereas a value close to 1 exhibits dissimilarities in the second-order structure. The distance measure $d(f_{11}, f_{22})$ can be estimated by

$$\hat{d}_{12} = \left(\max\left(1 - \frac{2\hat{D}_{12,n_1}}{\hat{D}_{1,n_1} + \hat{D}_{2,n_1}}, 0\right)\right)^{1/2}, \tag{2.26}$$

where the maximum is necessary, because the term $1 - \frac{2\hat{D}_{12,n_1}}{\hat{D}_{1,n_1}+\hat{D}_{2,n_1}}$ can be negative.

Remark 2.4
The main ideas of the proof of Theorem 2.1 can be furthermore employed to construct tests for various other hypothesis. For example a test for zero correlation can be derived by testing for

$$H_0: \quad f_{12} \equiv 0$$

which can be done by estimating $\int_{-\pi}^{\pi} |f_{12}(\lambda)|^2 d\lambda$. An estimator for this quantity is easily derived using the above approach and furthermore the calculation of the variance is straightforward which we omit for the sake of brevity.

Remark 2.5
Although we only considered the bivariate case, our method can be easily extended to an m dimensional process. Let us consider m stationary processes

$$X_t^{(j)} = \sum_{l=-\infty}^{\infty} \psi_l^{(j)} Z_{t-l}^{(j)} \quad t = 1, ..., n_j$$

with $j \in \{1, ..., m\}$ and $n_1 \leq n_2 \leq ... \leq n_m$. We define $I_j(\lambda)$ for $j = 1, ..., m$ and $I_{ij}(\lambda)$ for $i \neq j$ exactly as in the bivariate case, which results in an analogous definition of $\hat{D}_{j,n_1}$ for $j = 1, ..., m$ and

$\hat{D}_{ij,n_1}$ for $i \neq j$. Now an extension of Theorem 2.1 can be proved, which states that a standardized version of $\hat{V}_{n1} := (\hat{D}_{1,n_1}, ..., \hat{D}_{m,n_1}, \hat{D}_{12,n_1}, ..., \hat{D}_{m(m-1),n_1})^T \in I\!R^{\frac{m(m+1)}{2}}$ converges to a multivariate normal distribution with

$$\begin{aligned} \mathbb{E}(\hat{D}_{j,n_1} - D_j) =& o(1/\sqrt{n_1}) \quad \text{ for } j = 1, ..., m \\ \mathbb{E}(\hat{D}_{ij,n_1} - D_{ij}) =& o(1/\sqrt{n_1}) \quad \text{ for } i \neq j \end{aligned}$$

and for the variances and covariances we obtain

$$\begin{aligned} n_1 \mathrm{Cov}(\hat{D}_{j_1,n_1}, \hat{D}_{j_2,n_1}) \to & \frac{1}{\pi} \int_{-\pi}^{\pi} f_{j_1 j_2}^2(\lambda) f_{j_2 j_1}^2(\lambda) d\lambda + \frac{4}{\pi} \int_{-\pi}^{\pi} f_{j_1 j_1}(\lambda) |f_{j_1 j_2}(\lambda)|^2 f_{j_2 j_2}(\lambda) d\lambda \quad \text{ for } j_1, j_2 = 1, ..., m \\ n_1 \mathrm{Cov}(\hat{D}_{i_1 j_1,n_1}, \hat{D}_{i_2 j_2,n_1}) \to & \frac{1}{4\pi} \int_{-\pi}^{\pi} f_{i_1 i_1}(\lambda) f_{i_2 i_2}(\lambda) |f_{j_1 j_2}(\lambda)|^2 d\lambda + \frac{1}{4\pi} \int_{-\pi}^{\pi} f_{j_1 j_1}(\lambda) f_{j_2 j_2}(\lambda) |f_{i_1 i_2}(\lambda)|^2 d\lambda \\ & + \frac{1}{4\pi} \int_{-\pi}^{\pi} |f_{i_1 i_2}(\lambda)|^2 |f_{j_1 j_2}(\lambda)|^2 d\lambda + \frac{1}{4\pi} \int_{-\pi}^{\pi} f_{i_1 i_1}(\lambda) f_{j_2 j_2}(\lambda) |f_{i_2 j_1}(\lambda)|^2 \\ & + \frac{1}{4\pi} \int_{-\pi}^{\pi} f_{j_1 j_1}(\lambda) f_{i_2 i_2}(\lambda) |f_{i_1 j_2}(\lambda)|^2 d\lambda \quad \text{ for } i_1 \neq j_1 \text{ and } i_2 \neq j_2 \end{aligned}$$

and

$$n_1 \mathrm{Cov}(\hat{D}_{i_1,n_1}, \hat{D}_{i_2 j,n_1}) \to \frac{1}{\pi} \int_{-\pi}^{\pi} f_{i_1 i_1}^3(\lambda) f_{jj}(\lambda) + \frac{1}{\pi} \int_{-\pi}^{\pi} f_{i_1 i_1}(\lambda) f_{i_2 i_2}(\lambda) |f_{i_1 j}(\lambda)|^2 \quad \text{ for } i_2 \neq j.$$

Then a test for

$$\mathrm{H}_0 : \quad f_{11}(\lambda) = ... = f_{mm}(\lambda)$$

versus

$$\mathrm{H}_1 : \quad f_{ii}(\lambda) \neq f_{jj}(\lambda) \quad \text{for at least one pair } (i,j) \text{ with } i \neq j.$$

can be easily constructed as in the bivariate case by estimating

$$\tilde{D}^2 := \frac{1}{4\pi} \sum_{1 \leq i < j \leq m} \int_{-\pi}^{\pi} (f_{ii}(\lambda) - f_{jj}(\lambda))^2 d\lambda,$$

which we do not present for the sake of brevity.

Remark 2.6

A cumbersome but straightforward examination yields that for general independent and identically distributed innovations with existing moments of all orders, the limit theorem (2.11) still holds with σ^2 replaced by

$$\begin{aligned} \sigma_g^2 := \sigma^2 + & \frac{\kappa_1}{4\pi^2} \left(\int_{-\pi}^{\pi} f_{11}(\lambda)^2 d\lambda - \int_{-\pi}^{\pi} f_{11}(\lambda) f_{22}(\lambda) d\lambda \right)^2 + \frac{\kappa_2}{4\pi^2} \left(\int_{-\pi}^{\pi} f_{22}(\lambda)^2 d\lambda - \int_{-\pi}^{\pi} f_{11}(\lambda) f_{22}(\lambda) d\lambda \right)^2 \\ & + \frac{\kappa_3}{(2\pi\rho)^2} \Bigg(\left(\int_{-\pi}^{\pi} f_{11}(\lambda) f_{22}(\lambda) d\lambda \right)^2 + \int_{-\pi}^{\pi} f_{11}^2(\lambda) d\lambda \int_{-\pi}^{\pi} f_{22}^2(\lambda) d\lambda - \int_{-\pi}^{\pi} f_{11}^2(\lambda) d\lambda \int_{-\pi}^{\pi} f_{11} f_{22}(\lambda) d\lambda \\ & \qquad - \int_{-\pi}^{\pi} f_{22}^2(\lambda) d\lambda \int_{-\pi}^{\pi} f_{11} f_{22}(\lambda) d\lambda \Bigg) \end{aligned}$$

where $\kappa_1 = \mathbb{E}\left(\left(Z_t^{(1)}\right)^4\right) - 3$, $\kappa_2 = \mathbb{E}\left(\left(Z_t^{(2)}\right)^4\right) - 3$ and

$$\kappa_3 := \operatorname{cum}\left(Z_t^{(1)}, Z_t^{(1)}, Z^{(2)}_{\lfloor tq_{n_1,n_2}\rfloor - \lfloor q_{n_1,n_2}-1\rfloor}, Z^{(2)}_{\lfloor tq_{n_1,n_2}\rfloor - \lfloor q_{n_1,n_2}-1\rfloor}\right)$$

are the corresponding fourth order cumulants. Note that σ_g^2 and σ^2 does not differ under the null hypothesis so that the test (2.14) does not change at all in the more general case. A similar phenomenon can be observed for the tests proposed by Eichler (2008), Dette et al. (2010), Dette and Hildebrandt (2011) and Dette et al. (2011).

n_1	n_2	α	$\mathbf{X}^1$	$\mathbf{X}^2$	$\mathbf{X}^3$	$\mathbf{X}^4$	$\mathbf{X}^5$	$\mathbf{X}^6$	$\mathbf{X}^7$
256	256	0.05	0.041	0.038	0.039	0.048	0.045	0.053	0.053
		0.1	0.088	0.109	0.128	0.098	0.106	0.117	0.131
256	384	0.05	0.049	0.031	0.049	0.042	0.034	0.047	0.054
		0.1	0.106	0.104	0.128	0.098	0.099	0.114	0.147
256	512	0.05	0.043	0.030	0.047	0.030	0.026	0.045	0.031
		0.1	0.085	0.105	0.139	0.079	0.069	0.109	0.126
256	640	0.05	0.050	0.040	0.044	0.036	0.030	0.037	0.046
		0.1	0.109	0.125	0.112	0.093	0.081	0.097	0.122
384	384	0.05	0.036	0.036	0.047	0.036	0.037	0.054	0.043
		0.1	0.092	0.099	0.120	0.101	0.103	0.110	0.117
384	512	0.05	0.048	0.038	0.039	0.037	0.045	0.061	0.065
		0.1	0.110	0.091	0.120	0.075	0.091	0.131	0.136
384	640	0.05	0.037	0.033	0.045	0.039	0.046	0.050	0.062
		0.1	0.078	0.091	0.128	0.091	0.096	0.124	0.149
512	512	0.05	0.037	0.034	0.034	0.037	0.048	0.044	0.027
		0.1	0.096	0.105	0.111	0.085	0.106	0.097	0.100
512	640	0.05	0.037	0.035	0.054	0.041	0.046	0.055	0.061
		0.1	0.094	0.093	0.137	0.082	0.103	0.119	0.131
640	640	0.05	0.044	0.034	0.035	0.046	0.037	0.045	0.043
		0.1	0.106	0.092	0.101	0.098	0.089	0.104	0.110

Table 1: *Rejection frequencies of the test* (2.14) *under the null hypothesis for* $\rho = 0$.

3 Finite sample study

3.1 Size and power of the test

In this section we study the size and the power of test (2.14) in the case of finite samples. All simulations are based on 1000 iterations and we consider all different combinations of $n_1, n_2 \in \{256, 384, 512, 640\}$ with $n_1 \leq n_2$. At first we set $\rho = 0$, but it will be demonstrated later that the results do not change with a non-zero correlation of the innovations. To study the approximation of the nominal level, the seven processes

$$
\begin{aligned}
\mathbf{X}^1 : X_t &= Z_t \\
\mathbf{X}^2 : X_t &= 0.8X_{t-1} + Z_t \\
\mathbf{X}^3 : X_t &= -0.8X_{t-1} + Z_t \\
\mathbf{X}^4 : X_t &= Z_t + 0.8Z_{t-1} \\
\mathbf{X}^5 : X_t &= Z_t - 0.8Z_{t-1} \\
\mathbf{X}^6 : X_t &\sim FARIMA(0.45, 0, 0.8) \\
\mathbf{X}^7 : X_t &= Z_t 1_{[t \leq 0.5T]} + 0.8X_{t-1} 1_{[0.5T \leq t \leq 0.75T]} + Z_t 1_{[t \geq 0.75T]} \quad \text{for } t = 1, ..., T.
\end{aligned}
$$

were simulated, where the $FARIMA(0.45, 0, 0.8)$-model corresponds to a LongMemory-process given by

$$(1 - B)^{0.45} X_t = (1 - 0.8B) Z_t$$

with the backshift-operator B (i.e. $B^j X_t = X_{t-j}$) and

$$(1 - B)^d = \sum_{k=0}^{\infty} \binom{d}{k} (-B)^k.$$

Note that the models $\mathbf{X}^6$ and $\mathbf{X}^7$ both do not fit into the theoretical framework considered in section 2, since for the $FARIMA(0.45, 0, 0.8)$-process we obtain

$$\sum_{l=-\infty}^{\infty} |\psi_l| = \infty$$

which contradicts (2.4) and the structural-break model $\mathbf{X}^7$ does not even has a stationary solution. Nevertheless since these models are of great interest in practice, we investigate the performance of our approach in these cases as well. The results are given in Table 1 and it can be seen that the test (2.14) is very robust against different choices of n_1 and n_2. Furthermore our method also seems to work for the models $\mathbf{X}^6$ and $\mathbf{X}^7$ although the convergence seems to be slightly slower.
To study the power of the test we additionally simulated the LongMemory-process

$$\mathbf{X}^8 : X_t \sim FARIMA(0.45, 0, 0.5)$$

n_1	n_2	α	$\mathbf{X}^1\mathbf{X}^5$	$\mathbf{X}^3\mathbf{X}^5$	$\mathbf{X}^4\mathbf{X}^5$	$\mathbf{X}^4\mathbf{X}^6$	$\mathbf{X}^6\mathbf{X}^8$	$\mathbf{X}^1\mathbf{X}^7$
256	256	0.05	0.773	0.628	1	1	0.286	0.523
		0.1	0.894	0.875	1	1	0.509	0.685
256	384	0.05	0.758	0.619	1	1	0.299	0.551
		0.1	0.877	0.841	1	1	0.620	0.719
256	512	0.05	0.776	0.650	1	1	0.255	0.539
		0.1	0.892	0.848	1	1	0.499	0.739
256	640	0.05	0.777	0.636	1	1	0.294	0.563
		0.1	0.904	0.859	1	1	0.565	0.755
384	384	0.05	0.920	0.804	1	1	0.361	0.693
		0.1	0.969	0.936	1	1	0.644	0.814
384	512	0.05	0.895	0.828	1	1	0.384	0.699
		0.1	0.956	0.938	1	1	0.710	0.836
384	640	0.05	0.917	0.788	1	1	0.393	0.702
		0.1	0.968	0.925	1	1	0.721	0.858
512	512	0.05	0.975	0.877	1	1	0.426	0.800
		0.1	0.994	0.967	1	1	0.707	0.889
512	640	0.05	0.971	0.890	1	1	0.506	0.811
		0.1	0.987	0.973	1	1	0.787	0.906
640	640	0.05	0.993	0.959	1	1	0.489	0.906
		0.1	0.999	0.993	1	1	0.759	0.934

Table 2: *Rejection frequencies of the test* (2.14) *for several alternatives for* $\rho = 0$.

to also investigate the finite sample behaviour if different LongMemory-models are compared. We exemplarily present the results of a comparison of $\mathbf{X}^5$ with $\mathbf{X}^j$ for $j \in \{1, 3, 4\}$, $\mathbf{X}^6$ with $\mathbf{X}^j$ for $j \in \{4, 8\}$ and $\mathbf{X}^1$ with $\mathbf{X}^7$ (all other comparison between the eight processes yield better results than the depicted ones). The results are given in Table 2 for the uncorrelated case and again we see that our method also works for LongMemory- and structural-break models although again the power seems to grow more slowly with n_1.
Subsequently we took a look at the case where the innovations have non-zero correlations and for the sake of brevity we again only display some of the results. Under the nullhypothesis we provide the results for $\mathbf{X}^2$ and $\mathbf{X}^4$, while under the alternative only the comparisons between $\mathbf{X}^4$ and $\mathbf{X}^j$ for $j = 1, 2$ are depicted. All other comparison yield similar results and the rejection frequencies are given in Table 3 for the null hypothesis and in Table 4 under the alternative respectively. It can be seen that the power does not seem to change very much (if at all) for different correlations.

n_1	n_2	α	$\mathbf{X}^2$				$\mathbf{X}^4$			
			ρ				ρ			
			-0.7	-0.3	0.3	0.7	-0.7	-0.3	0.3	0.7
256	256	0.05	0.038	0.038	0.042	0.039	0.030	0.030	0.036	0.043
		0.1	0.123	0.116	0.116	0.111	0.082	0.087	0.092	0.104
256	384	0.05	0.040	0.045	0.037	0.033	0.035	0.033	0.041	0.043
		0.1	0.118	0.113	0.115	0.111	0.091	0.086	0.099	0.103
256	512	0.05	0.041	0.041	0.030	0.040	0.035	0.030	0.035	0.034
		0.1	0.118	0.119	0.116	0.122	0.092	0.086	0.084	0.090
256	640	0.05	0.028	0.043	0.032	0.031	0.031	0.038	0.040	0.038
		0.1	0.110	0.118	0.102	0.098	0.086	0.094	0.093	0.083
384	384	0.05	0.034	0.034	0.031	0.036	0.039	0.032	0.034	0.036
		0.1	0.110	0.091	0.099	0.097	0.085	0.082	0.084	0.090
384	512	0.05	0.035	0.041	0.035	0.032	0.035	0.018	0.048	0.029
		0.1	0.102	0.107	0.103	0.096	0.084	0.062	0.104	0.080
384	640	0.05	0.033	0.033	0.035	0.034	0.034	0.031	0.028	0.032
		0.1	0.103	0.111	0.097	0.107	0.093	0.083	0.087	0.082
512	512	0.05	0.029	0.046	0.042	0.038	0.035	0.039	0.030	0.040
		0.1	0.080	0.108	0.115	0.094	0.086	0.094	0.078	0.097
512	640	0.05	0.031	0.030	0.044	0.039	0.031	0.036	0.031	0.049
		0.1	0.092	0.098	0.089	0.091	0.089	0.077	0.084	0.106
640	640	0.05	0.030	0.036	0.033	0.035	0.050	0.034	0.039	0.057
		0.1	0.092	0.088	0.101	0.100	0.099	0.087	0.090	0.100

Table 3: *Rejection frequencies of the test* (2.14) *under the null hypothesis for different* ρ.

Finally we considered two time series with non-Gaussian innovations a non-linear GARCH process. These examined models are given by

$$
\begin{aligned}
\mathbf{X}_a^9 : X_t &= aX_{t-1} + U_t \quad \text{with} \quad U_t \sim U[-1,1] \\
\mathbf{X}_a^{10} : X_t &= aX_{t-1} + \sqrt{5/48}\,(P_t - \mathbb{E}(P_t)) \quad \text{with} \quad P_t \sim Pareto(1,5) \\
\mathbf{X}_a^{11} : X_t &\sim GARCH_{(1,1)}(0.3, a, 0.85),
\end{aligned}
$$

n_1	n_2	α	$\mathbf{X}^4\mathbf{X}^1$				$\mathbf{X}^4\mathbf{X}^2$			
			ρ				ρ			
			-0.7	-0.3	0.3	0.7	-0.7	-0.3	0.3	0.7
256	256	0.05	0.771	0.799	0.807	0.784	0.668	0.640	0.613	0.638
		0.1	0.893	0.901	0.904	0.893	0.888	0.858	0.849	0.888
256	384	0.05	0.755	0.748	0.771	0.781	0.658	0.613	0.620	0.630
		0.1	0.888	0.862	0.880	0.887	0.872	0.830	0.835	0.858
256	512	0.05	0.798	0.759	0.749	0.764	0.634	0.614	0.603	0.673
		0.1	0.902	0.870	0.885	0.884	0.860	0.831	0.841	0.871
256	640	0.05	0.759	0.769	0.755	0.769	0.591	0.630	0.586	0.636
		0.1	0.895	0.896	0.885	0.892	0.846	0.837	0.848	0.884
384	384	0.05	0.921	0.901	0.910	0.920	0.843	0.788	0.805	0.852
		0.1	0.969	0.965	0.968	0.973	0.958	0.929	0.942	0.965
384	512	0.05	0.922	0.906	0.890	0.908	0.848	0.815	0.801	0.830
		0.1	0.971	0.957	0.955	0.961	0.959	0.932	0.938	0.948
384	640	0.05	0.904	0.904	0.898	0.901	0.818	0.803	0.815	0.794
		0.1	0.966	0.968	0.960	0.960	0.943	0.927	0.926	0.947
512	512	0.05	0.968	0.964	0.966	0.984	0.935	0.916	0.890	0.937
		0.1	0.994	0.992	0.990	0.995	0.991	0.982	0.977	0.992
512	640	0.05	0.967	0.959	0.972	0.974	0.921	0.906	0.915	0.921
		0.1	0.991	0.979	0.988	0.994	0.983	0.979	0.975	0.980
640	640	0.05	0.977	0.993	0.990	0.994	0.977	0.966	0.961	0.970
		0.1	0.998	0.996	0.997	1.000	0.997	0.995	0.991	0.995

Table 4: *Rejection frequencies of the test* (2.14) *for several alternatives for different* ρ.

where a $GARCH_{(p,q)}(\alpha_0, \alpha_1, ..., \alpha_q, \beta_1, ..., \beta_p)$ process is defined through the equations

$$\begin{aligned} X_t &= \sigma_t Z_t \\ \sigma_t^2 &= \alpha_0 + \sum_{k=1}^{q} \alpha_k X_{t-k}^2 + \sum_{l=1}^{p} \beta_l \sigma_{t-l}^2 \\ Z_t &\sim N(0,1). \end{aligned}$$

Note that in the case $p = q = 1$ the GARCH model possesses a stationary solution if $\alpha_1 + \beta_1 < 1$. For these processes we obtain four possible null hypothesises and three alternatives, which are depicted in Table 5, and it can be seen that our method works very well also in this cases.

n_1	n_2	α	$\mathbf{X}^9_{0.7}$	$\mathbf{X}^{10}_{0.7}$	$\mathbf{X}^{11}_{0.05}$	$\mathbf{X}^9_{0.7}\mathbf{X}^{10}_{0.7}$	$\mathbf{X}^9_{0.7}\mathbf{X}^{10}_{0.5}$	$\mathbf{X}^{10}_{0.7}\mathbf{X}^{10}_{0.5}$	$\mathbf{X}^{11}_{0.05}\mathbf{X}^{11}_{0.1}$
256	256	0.05	0.041	0.063	0.058	0.042	0.193	0.178	0.478
		0.1	0.110	0.134	0.113	0.118	0.366	0.332	0.613
256	384	0.05	0.034	0.044	0.047	0.038	0.170	0.181	0.518
		0.1	0.084	0.127	0.106	0.122	0.327	0.335	0.647
256	512	0.05	0.040	0.041	0.042	0.041	0.174	0.182	0.499
		0.1	0.101	0.128	0.090	0.117	0.340	0.319	0.628
256	640	0.05	0.030	0.042	0.050	0.042	0.140	0.163	0.525
		0.1	0.098	0.113	0.099	0.110	0.306	0.323	0.659
384	384	0.05	0.034	0.045	0.051	0.031	0.246	0.241	0.594
		0.1	0.091	0.129	0.109	0.108	0.429	0.391	0.695
384	512	0.05	0.035	0.032	0.065	0.045	0.247	0.241	0.620
		0.1	0.103	0.094	0.121	0.101	0.433	0.398	0.721
384	640	0.05	0.037	0.047	0.062	0.034	0.225	0.218	0.619
		0.1	0.091	0.125	0.121	0.104	0.404	0.376	0.726
512	512	0.05	0.026	0.032	0.049	0.047	0.295	0.292	0.673
		0.1	0.089	0.096	0.104	0.120	0.482	0.484	0.771
512	640	0.05	0.040	0.033	0.056	0.044	0.307	0.287	0.701
		0.1	0.093	0.101	0.109	0.103	0.491	0.463	0.794
640	640	0.05	0.031	0.039	0.053	0.042	0.371	0.350	0.764
		0.1	0.101	0.092	0.118	0.114	0.559	0.491	0.843

Table 5: *Rejection frequencies of the test* (2.14) *under the null hypothesis and for several alternatives in non-Gaussian and non-linear models [for* $\rho = 0$*].*

3.2 Real world data

In this section we investigate how the clustering-method described in Remark 2.3 performs if it is applied to real world data. Therefore we took three log-returns of stock prices from the financial sector, three log-returns from the health sector and two key interest rates. Exemplarily for the finance sector we choosed the stocks of Barclays, Deutsche Bank and Goldman Sachs and the health sector is represented by GlaxoSmithKline, Novartis and Pfizer. The key interest rates were taken from Great Britain and the EU and all time series data were recorded between March 1st, 2003 and July 29th, 2011. While the interest rates data were observed monthly, the stock prices were recorded daily or weekly . However, even if two stock prices were observed daily they might differ in length, since they are for example traded on different stock exchanges with different trading days. The result of our cluster analysis using (2.26) is presented in the dendrogram given in Figure 1. As expected we get three different groups which correspond to the finance sector, the health sector and the key interes rates.

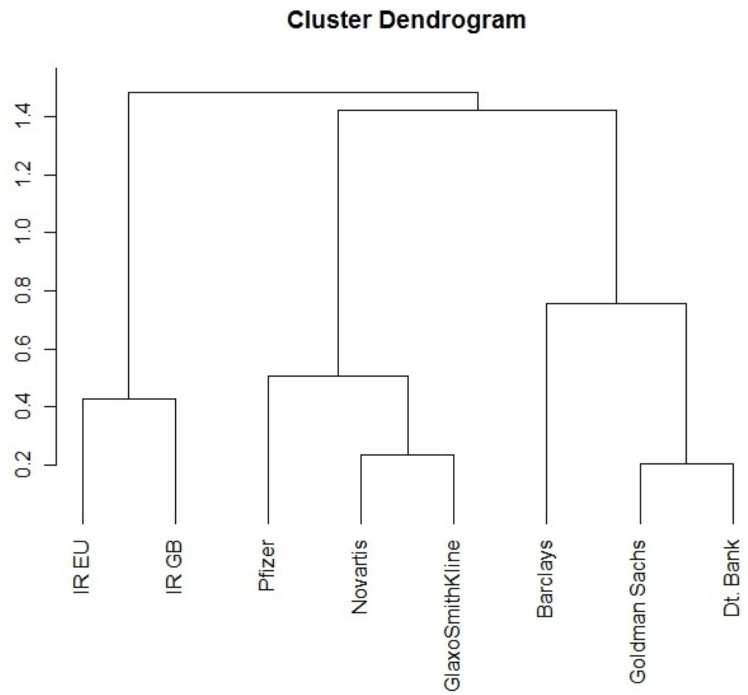

Figure 1: Clustering of financial time series data.

Acknowledgements This work has been supported in part by the Collaborative Research Center "Statistical modeling of nonlinear dynamic processes" (SFB 823, Teilprojekt C1, C4) of the German Research Foundation (DFG).

4 Appendix: Technical details

Proof of Theorem 2.1: By using the Cramer-Wold device, we have to show that

$$c^T \sqrt{n_1} \Big\{ (\hat{D}_{1,n_1}, \hat{D}_{12,n_1}, \hat{D}_{2,n_1})^T - (D_1, D_{12}, D_2)^T \Big\} \xrightarrow{D} N(0, c^T \Sigma c)$$

for all vectors $c \in I\!R^3$. For the sake of brevity, we restrict ourselve to the case $c = (0,1,0)^T$ since the more general follows with exactly the same arguments. Therefore we show

$$\hat{T}_{n_1} := \sqrt{n_1}(\hat{D}_{12,n_1} - D_{12}) \xrightarrow{D} N(0, \Sigma_{22}) \tag{4.1}$$

and we do that by using the method of cumulants, which is described in chapter 2.3. of Brillinger (1981) (and whose notations we will make heavy use of), i.e. in the following it is proved that

$$\text{cum}_l(\hat{T}_{n_1}) = o(1) \tag{4.2}$$

for $l = 1$ and $l \geq 3$ and that

$$\text{(4.3)} \qquad \text{cum}_2(\hat{T}_{n_1}) \xrightarrow{n_1 \to \infty} \Sigma_{22},$$

which will yield the assertion.

Proof of (4.2) **for the case** $l = 1$**:** Because of the symmetry of the periodogram, it is

$$\mathbb{E}(\hat{D}_{12,n_1}) = \frac{1}{2n_1} \sum_{k=-\lfloor \frac{n_1-1}{2} \rfloor}^{\lfloor \frac{n_1}{2} \rfloor} \frac{1}{(2\pi)^2 n_1 n_2} \sum_{p_1,q_1=1}^{n_1} \sum_{p_2,q_2=1}^{n_2} \sum_{l_1,m_1=-\infty}^{\infty} \sum_{l_2,m_2=-\infty}^{\infty} \psi_{l_1}^{(1)} \psi_{m_1}^{(1)} \psi_{l_2}^{(2)} \psi_{m_2}^{(2)}$$
$$E(Z_{p_1-l_1}^{(1)} Z_{q_1-m_1}^{(1)} Z_{p_2-l_2}^{(2)} Z_{q_2-m_2}^{(2)}) e^{-i\lambda_{1,k}(p_1-q_1)-i\lambda_{1,k+1}(p_2-q_2)} + O(1/n_1)$$

and because of the standard normality of the innovations we obtain

$$\mathbb{E}(Z_{p_1-l_1}^{(1)} Z_{q_1-m_1}^{(1)} Z_{p_2-l_2}^{(2)} Z_{q_2-m_2}^{(2)}) = \mathbb{E}(Z_{p_1-l_1}^{(1)} Z_{q_1-m_1}^{(1)}) \mathbb{E}(Z_{p_2-l_2}^{(2)} Z_{q_2-m_2}^{(2)}) + \mathbb{E}(Z_{p_1-l_1}^{(1)} Z_{q_2-m_2}^{(2)}) \mathbb{E}(Z_{q_1-m_1}^{(1)} Z_{p_2-l_2}^{(2)})$$
$$+ \mathbb{E}(Z_{p_1-l_1}^{(1)} Z_{p_2-l_2}^{(2)}) \mathbb{E}(Z_{q_1-m_1}^{(1)} Z_{q_2-m_2}^{(2)})$$

which yields that $\mathbb{E}(\hat{D}_{12,n_1})$ (without the $O(1/n_1)$-term) can be divided into the sums of three terms which are called A, B and C respectively. For the first term we obtain the conditions

$$p_1 = q_1 + l_1 - m_1$$
$$p_2 = q_2 + l_2 - m_2$$

(all others cases are equal to zero) which results in

$$A = \frac{1}{2n_1} \sum_{k=-\lfloor \frac{n_1-1}{2} \rfloor}^{\lfloor \frac{n_1}{2} \rfloor} \frac{1}{(2\pi)^2 n_1 n_2} \sum_{l_1,l_2,m_1,m_2=-\infty}^{\infty} \sum_{\substack{q_1=1 \\ 1 \leq q_1+l_1-m_1 \leq n_1}}^{n_1} \sum_{\substack{q_2=1 \\ 1 \leq q_2+l_2-m_2 \leq n_2}}^{n_2} \psi_{l_1}^{(1)} \dots \psi_{m_2}^{(2)} e^{-i\lambda_{1,k}(l_1-m_1)-i\lambda_{1,k+1}(l_2-m_2)}$$
$$= \frac{1}{2n_1} \sum_{k=-\lfloor \frac{n_1-1}{2} \rfloor}^{\lfloor \frac{n_1}{2} \rfloor} \frac{1}{(2\pi)^2 n_1 n_2} \sum_{q_1=1}^{n_1} \sum_{q_2=1}^{n_2} \sum_{l_1,l_2,m_1,m_2=-\infty}^{\infty} \psi_{l_1}^{(1)} \dots \psi_{m_2}^{(2)} e^{-i\lambda_{1,k}(l_1-m_1)-i\lambda_{1,k+1}(l_2-m_2)} + o\left(\frac{1}{\sqrt{n}}\right),$$

where the last equality follows from

$$\text{(4.4)} \qquad \frac{1}{n_j} \sum_{l:|l|<Mn_j} \psi_l^{(j)} |l| = \frac{1}{n_j} \sum_{l:|l|<Mn_j} \psi_l^{(j)} |l|^{\alpha} |l|^{1-\alpha} = o(1/n_j^{\alpha})$$

with $M \in I\!R$, where (2.4) was used. It now follows by the Hölder continuity condition that A equals

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} f_{11}(\lambda) f_{22}(\lambda) d\lambda + o\left(\frac{1}{\sqrt{n}}\right).$$

If we consider the summand B, we obtain the conditions

$$q_1 = \lfloor (p_2 - l_2) q_{n_1,n_2} \rfloor + m_1 - \lfloor q_{n_1,n_2} - 1 \rfloor$$
$$q_2 = \lfloor (p_1 - l_1) q_{n_1,n_2} \rfloor + m_2 - \lfloor q_{n_1,n_2} - 1 \rfloor$$

which yields

$$B = \frac{\rho^2}{2n_1} \sum_{k=-\lfloor \frac{n_1-1}{2} \rfloor}^{\lfloor \frac{n_1}{2} \rfloor} \frac{1}{(2\pi)^2 n_1 n_2} \sum_{l_1,m_1,l_2,m_2=-\infty}^{\infty} \sum_{\substack{p_1=1 \\ 1 \le \lfloor (p_1-l_1) q_{n_1,n_2} \rfloor + m_2 - \lfloor q_{n_1,n_2} - 1 \rfloor \le n_2}}^{n_1} \sum_{\substack{p_2=1 \\ 1 \le \lfloor (p_2-l_2) q_{n_1,n_2} \rfloor + m_1 - \lfloor q_{n_1,n_2} - 1 \rfloor \le n_1}}^{n_2}$$
$$\psi_{l_1}^{(1)} ... \psi_{m_2}^{(2)} e^{-i\lambda_{1,k}(p_1 - \lfloor (p_2-l_2) q_{n_1,n_2} \rfloor + m_1 - \lfloor q_{n_1,n_2} - 1 \rfloor)} e^{-i\lambda_{1,k+1}(p_2 - \lfloor (p_1-l_1) q_{n_1,n_2} \rfloor + m_2 - \lfloor q_{n_1,n_2} - 1 \rfloor)}.$$

If we now employ the identity

$$\text{(4.5)} \qquad \frac{1}{n_1} \sum_{k=-\lfloor \frac{n_1-1}{2} \rfloor}^{\lfloor \frac{n_1}{2} \rfloor} e^{-i\lambda_{1,k} p} = \begin{cases} 1 & \text{if } p = 0, \pm n_1, \pm 2n_1, ... \\ 0 & \text{else} \end{cases},$$

it follows with (2.10) that if p_1 is chosen there are only finitely many p_2 which yields a non-zero summand. Therefore we obtain that $B = o(1/\sqrt{n_1})$ and with the same arguments it can be shown that $C = o(1/\sqrt{n_1})$.

□

Proof of (4.3)**:** It is

$$\text{cum}_2(\sqrt{n_1} \hat{D}_{12,n_1}) = \frac{1}{n_1} \sum_{k=1}^{\lfloor \frac{n_1}{2} \rfloor - 1} \text{cum}_2(I_1(\lambda_{1,k}) I_2(\lambda_{1,k+1})) + \frac{1}{n_1} \sum_{\substack{k_1,k_2=1 \\ k_1 \neq k_2}}^{\lfloor \frac{n_1}{2} \rfloor - 1} \text{cum}(I_1(\lambda_{1,k_1}) I_2(\lambda_{1,k_1+1}), I_1(\lambda_{1,k_2}) I_2(\lambda_{1,k_2+1}))$$

and the assertion follows if we show that

$$\text{(4.6)} \qquad \frac{1}{n_1} \sum_{k=1}^{\lfloor \frac{n_1}{2} \rfloor - 1} \text{cum}_2(I_1(\lambda_{1,k}) I_2(\lambda_{1,k+1})) \xrightarrow{n_1 \to \infty} \frac{3}{4\pi} \int_{-\pi}^{\pi} f_{11}^2(\lambda) f_{22}^2(\lambda) d\lambda$$

and

$$\text{(4.7)} \qquad \frac{1}{n_1} \sum_{\substack{k_1,k_2=1 \\ k_1 \neq k_2}}^{\lfloor \frac{n_1}{2} \rfloor - 1} \text{cum}(I_1(\lambda_{1,k_1}) I_2(\lambda_{1,k_1+1}), I_1(\lambda_{1,k_2}) I_2(\lambda_{1,k_2+1})) \xrightarrow{n_1 \to \infty} \frac{1}{2\pi} \int_{-\pi}^{\pi} f_{11}(\lambda) |f_{12}(\lambda)|^2 f_{22}(\lambda) d\lambda.$$

We present a detailed proof of (4.6) and then comment briefly on (4.7) since it is proved analogously. Employing the symmetry of the periodogram again, we get

$$\begin{aligned}
\frac{1}{n_1}\sum_{k=1}^{\lfloor\frac{n_1}{2}\rfloor-1}\mathrm{cum}_2(I_1(\lambda_{1,k})I_2(\lambda_{1,k+1})) &= \frac{1}{2n_1}\sum_{k=-\lfloor\frac{n_1-1}{2}\rfloor}^{\lfloor\frac{n_1}{2}\rfloor}\mathrm{cum}_2(I_1(\lambda_{1,k})I_2(\lambda_{1,k+1}))+O(1/n_1)\\
&= \frac{1}{2n_1}\sum_{k=-\lfloor\frac{n_1-1}{2}\rfloor}^{\lfloor\frac{n_1}{2}\rfloor}\frac{1}{(2\pi)^4n_1^2n_2^2}\sum_{j=1}^{2}\sum_{p_j,q_j,r_j,s_j=1}^{n_j}\sum_{a_j,b_j,c_j,d_j=-\infty}^{\infty}\psi_{a_1}^{(1)}...\psi_{d_2}^{(2)}\\
&\quad e^{-i\lambda_{1,k}(p_1-q_1+r_1-s_1)-i\lambda_{1,k+1}(p_2-q_2+r_2-s_2)}\\
&\quad \mathrm{cum}(Z^{(1)}_{p_1-a_1}Z^{(1)}_{q_1-b_1}Z^{(2)}_{p_2-a_2}Z^{(2)}_{q_2-b_2},Z^{(1)}_{r_1-c_1}Z^{(1)}_{s_1-d_1}Z^{(2)}_{r_2-c_2}Z^{(2)}_{s_2-d_2})+O(1/n_1)\\
&= \sum_{\nu}\frac{1}{2n_1}\sum_{k=-\lfloor\frac{n_1-1}{2}\rfloor}^{\lfloor\frac{n_1}{2}\rfloor}\frac{1}{(2\pi)^4n_1^2n_2^2}\sum_{j=1}^{2}\sum_{p_j,q_j,r_j,s_j=1}^{n_j}\sum_{a_j,b_j,c_j,d_j=-\infty}^{\infty}\psi_{a_1}^{(1)}...\psi_{d_2}^{(2)}\\
&\quad e^{-i\lambda_{1,k}(p_1-q_1+r_1-s_1)-i\lambda_{1,k+1}(p_2-q_2+r_2-s_2)}\\
&\quad \mathrm{cum}(Z_i^{(j)};(i,j)\in\nu_1)\cdots\mathrm{cum}(Z_i^{(j)};(i,j)\in\nu_4)+O(1/n_1)
\end{aligned}\tag{4.8}$$

where the sum goes over all indecomposable partitions $\nu=\nu_1\cup...\cup\nu_4$ of

$$\begin{matrix} Z^{(1)}_{p_1} & Z^{(1)}_{q_1} & Z^{(1)}_{r_1} & Z^{(1)}_{s_1}\\ Z^{(2)}_{p_2} & Z^{(2)}_{q_2} & Z^{(2)}_{r_2} & Z^{(2)}_{s_2}\end{matrix}$$

with $|\nu_i|=2\ \forall i=1,...,4$ (we only have to consider partitions with two elements in each set, because of the Gaussianity of the innovations). Every chosen partition will imply conditions for the choice of p_j,q_j,r_j,s_j as in the calculation of the expectation. For some partitions there will not be a p_j,q_j,r_j,s_j in the exponent of e after the substitution of the conditions and for other partitions there will still remain one. Let us take an example of the latter one and consider the partitions which corresponds to

$$\mathrm{cum}(Z^{(1)}_{p_1-a_1},Z^{(1)}_{s_1-d_1})\mathrm{cum}(Z^{(1)}_{q_1-b_1},Z^{(1)}_{r_1-c_1})\mathrm{cum}(Z^{(2)}_{p_2-a_2},Z^{(2)}_{r_2-c_2})\mathrm{cum}(Z^{(2)}_{q_2-b_2},Z^{(2)}_{s_2-d_2}).$$

We name the corresponding term of this partition in (4.8) with V_2 and obtain the conditions

$$\begin{aligned} p_1&=s_1+a_1-d_1\\ q_1&=r_1+b_1-c_1\\ p_2&=r_2+a_2-c_2\\ q_2&=s_2+b_2-d_2\end{aligned}$$

which yields

$$\begin{aligned} V_2 &= \frac{1}{2n_1}\sum_{k=-\lfloor\frac{n_1-1}{2}\rfloor}^{\lfloor\frac{n_1}{2}\rfloor}\frac{1}{(2\pi)^4 n_1^2 n_2^2}\sum_{j=1}^{2}\sum_{a_j,b_j,c_j,d_j=-\infty}^{\infty}\sum_{\substack{s_1,r_1=1\\ 1\le s_1+a_1-d_1\le n_1\\ 1\le r_1+b_1-c_1\le n_1}}^{n_1}\sum_{\substack{r_2,s_2=1\\ 1\le r_2+a_2-c_2\le n_2\\ 1\le s_2+b_2-d_2\le n_2}}^{n_2}\psi_{a_1}^{(1)}...\psi_{d_2}^{(2)} \\ &\quad e^{-i\lambda_{1,k}(a_1-d_1+c_1-b_1)-i\lambda_{1,k+1}(2r_2-2s_2+a_2-c_2+d_2-b_2)} \\ &= \frac{1}{2n_1}\sum_{k=-\lfloor\frac{n_1-1}{2}\rfloor}^{\lfloor\frac{n_1}{2}\rfloor}\frac{1}{(2\pi)^4 n_1^2 n_2^2}\sum_{j=1}^{2}\sum_{a_j,b_j,c_j,d_j=-\infty}^{\infty}\sum_{s_1,r_1=1}^{n_1}\sum_{r_2,s_2=1}^{n_2}\psi_{a_1}^{(1)}...\psi_{d_2}^{(2)} \\ &\quad e^{-i\lambda_{1,k+1}(2r_2-2s_2)}e^{-i\lambda_{1,k}(a_1-d_1+c_1-b_1)-i\lambda_{1,k+1}(a_2-c_2+d_2-b_2)}+o(1/\sqrt{n_1}), \end{aligned}$$

where the last equality again follows with (4.4). Now as in the handling of B in the calculation of the expectation, (4.5) implies that $V_2 = o(1)$.

Every other indecomposable partition is treated in exactly the same way and there are only three partitions which corresponding term in (4.8) does not vanish in the limit. These partitions correspond to one of the following three terms:

$$\begin{aligned} &1)\quad \mathrm{cum}(Z_{p_1-a_1}^{(1)},Z_{q_1-b_1}^{(1)})\mathrm{cum}(Z_{r_1-c_1}^{(1)},Z_{s_1-d_1}^{(1)})\mathrm{cum}(Z_{p_2-a_2}^{(2)},Z_{s_2-d_2}^{(2)})\mathrm{cum}(Z_{q_2-b_2}^{(2)},Z_{r_2-c_2}^{(2)}) \\ &2)\quad \mathrm{cum}(Z_{p_1-a_1}^{(1)},Z_{s_1-d_1}^{(1)})\mathrm{cum}(Z_{q_1-b_1}^{(1)},Z_{r_1-c_1}^{(1)})\mathrm{cum}(Z_{p_2-a_2}^{(2)},Z_{q_2-b_2}^{(2)})\mathrm{cum}(Z_{r_2-c_2}^{(2)},Z_{s_2-d_2}^{(2)}) \\ &3)\quad \mathrm{cum}(Z_{p_1-a_1}^{(1)},Z_{s_1-d_1}^{(1)})\mathrm{cum}(Z_{q_1-b_1}^{(1)},Z_{r_1-c_1}^{(1)})\mathrm{cum}(Z_{p_2-a_2}^{(2)},Z_{s_2-d_2}^{(2)})\mathrm{cum}(Z_{q_2-b_2}^{(2)},Z_{r_2-c_2}^{(2)}) \end{aligned}$$

We will exemplarily present the calculation concerning the 1) partition and denote the corresponding sum in (4.8) with V_1. We get

$$\begin{aligned} V_1 &= \frac{1}{2n_1}\sum_{k=-\lfloor\frac{n_1-1}{2}\rfloor}^{\lfloor\frac{n_1}{2}\rfloor}\frac{1}{(2\pi)^4 n_1^2 n_2^2}\sum_{j=1}^{2}\sum_{a_j,b_j,c_j,d_j=-\infty}^{\infty}\sum_{\substack{q_1,s_1=1\\ 1\le q_1+a_1-b_1\le n_1\\ 1\le s_1+c_1-d_1\le n_1}}^{n_1}\sum_{\substack{r_2,s_2=1\\ 1\le s_2+a_2-d_2\le n_2\\ 1\le r_2+b_2-c_2\le n_2}}^{n_2}\psi_{a_1}^{(1)}...\psi_{d_2}^{(2)} \\ &\quad e^{-i\lambda_{1,k}(a_1-b_1+c_1-d_1)-i\lambda_{1,k+1}(a_2-d_2+c_2-b_2)} \\ &= \frac{1}{2n_1}\sum_{k=-\lfloor\frac{n_1-1}{2}\rfloor}^{\lfloor\frac{n_1}{2}\rfloor}\frac{1}{(2\pi)^4 n_1^2 n_2^2}\sum_{j=1}^{2}\sum_{a_j,b_j,c_j,d_j=-\infty}^{\infty}\sum_{q_1,s_1=1}^{n_1}\sum_{r_2,s_2=1}^{n_2}\psi_{a_1}^{(1)}...\psi_{d_2}^{(2)} \\ &\quad e^{-i\lambda_{1,k}(a_1-b_1+c_1-d_1)-i\lambda_{1,k+1}(a_2-d_2+c_2-b_2)}+o(1/\sqrt{n_1}) \end{aligned}$$

by using (4.4). Now the Hölder continuity condition implies

$$V_1 = \frac{1}{4\pi}\int_{-\pi}^{\pi} f_{11}(\lambda)^2 f_{22}(\lambda)^2 d\lambda + o(1/\sqrt{n_1})$$

and since the partitions 2) and 3) yield the same result, we have shown (4.6).

With the same arguments as in the proof of (4.6) it is shown that

$$\frac{1}{n_1}\sum_{\substack{k_1,k_2=1\\k_1\neq k_2}}^{\lfloor\frac{n_1}{2}\rfloor-1}\mathrm{cum}(I_1(\lambda_{1,k_1})I_2(\lambda_{1,k_1+1}),I_1(\lambda_{1,k_2})I_2(\lambda_{1,k_2+1}))$$
$$=\frac{1}{n_1}\sum_{k_1=1}^{\lfloor\frac{n_1}{2}\rfloor-1}\mathrm{cum}(I_1(\lambda_{1,k_1})I_2(\lambda_{1,k_1+1}),I_1(\lambda_{1,k_1+1})I_2(\lambda_{1,k_1+2}))$$
$$+\frac{1}{n_1}\sum_{k_1=1}^{\lfloor\frac{n_1}{2}\rfloor-1}\mathrm{cum}(I_1(\lambda_{1,k_1})I_2(\lambda_{1,k_1+1}),I_1(\lambda_{1,k_1-1})I_2(\lambda_{1,k_1}))+o(1)$$

and it is shown completely analogously to the proof of (4.6) that

$$\frac{1}{n_1}\sum_{k_1=1}^{\lfloor\frac{n_1}{2}\rfloor-1}\mathrm{cum}(I_1(\lambda_{1,k_1})I_2(\lambda_{1,k_1+1}),I_1(\lambda_{1,k_1+1})I_2(\lambda_{1,k_1+2}))$$

and

$$\frac{1}{n_1}\sum_{k_1=1}^{\lfloor\frac{n_1}{2}\rfloor-1}\mathrm{cum}(I_1(\lambda_{1,k_1})I_2(\lambda_{1,k_1+1}),I_1(\lambda_{1,k_1-1})I_2(\lambda_{1,k_1}))$$

both converge to

$$\frac{1}{4\pi}\int_{-\pi}^{\pi}f_{11}(\lambda)|f_{12}(\lambda)|^2f_{22}(\lambda)d\lambda.$$

which yields (4.7).

□

Proof of (4.2) **for the case** $l\geq 3$**:** Since the proof is done by combining standard cumulants methods with the arguments that are used in the calculation of the expectation and the variance, we will restrict ourselve to a brief explanation of the main ideas [a more detailed discussion in a similar situation can be found in Dette et al. (2011)]. We obtain

$$\mathrm{cum}_l(\sqrt{n_1}D_{12,n_1})=\frac{1}{(2n_1)^{l/2}}\sum_{j_1=1}^{l}\sum_{k_{j_1}=-\lfloor\frac{n_1-1}{2}\rfloor}^{\lfloor\frac{n_1}{2}\rfloor}\frac{1}{(2\pi)^{2l}n_1^ln_2^l}\sum_{j_2=1}^{2}\sum_{a_{j_1,j_2},b_{j_1,j_2}=-\infty}^{\infty}\sum_{p_{j_1,j_2},q_{j_1,j_2}=1}^{n_{j_2}}\psi_{a_{1,1}}^{(1)}\cdots\psi_{b_{l,2}}^{(2)}$$
$$\exp(-i\lambda_{1,k}(p_{11}-q_{11})-i\lambda_{1,k+1}(p_{12}-q_{12}))\cdots\exp(-i\lambda_{1,k}(p_{l1}-q_{l1})-i\lambda_{1,k+1}(p_{l2}-q_{l2}))$$
$$\mathrm{cum}(Z^{(1)}_{p_{11}-a_{11}}Z^{(1)}_{q_{11}-b_{11}}Z^{(2)}_{p_{12}-a_{12}}Z^{(2)}_{q_{12}-b_{12}},\ldots,Z^{(1)}_{p_{l1}-a_{l1}}Z^{(1)}_{q_{l1}-b_{l1}}Z^{(2)}_{p_{l2}-a_{l2}}Z^{(2)}_{q_{l2}-b_{l2}})$$

and if we now take a indecomposable partition of

$$\begin{matrix}Z^{(1)}_{p_{11}-a_{11}} & Z^{(1)}_{q_{11}-b_{11}} & Z^{(2)}_{p_{12}-a_{12}} & Z^{(2)}_{q_{12}-b_{12}}\\ \vdots & \vdots & \vdots & \vdots\\ Z^{(1)}_{p_{l1}-a_{l1}} & Z^{(1)}_{q_{l1}-b_{l1}} & Z^{(2)}_{p_{l2}-a_{l2}} & Z^{(2)}_{q_{l2}-b_{l2}}\end{matrix}$$

which consists only of sets with two elements (again this suffices because of the Gaussianity of the innovations), it follows directly that at most $2l$ of the $4l$ variables p_{j_1,j_2}, q_{j_1,j_2} $(j_1 = 1, ..., l,\ j_2 = 1, 2)$ are free to choose. By using the same arguments as in the calculation of the variance and the expectation it then follows by the indecomposability of the partition that in fact only $l + 1$ of the remaining $2l$ variables p_{j_1,j_2}, q_{j_1,j_2} are free to choose. This implies that

$$\mathrm{cum}_l(\sqrt{n_1} D_{12,n_1}) = O(n_1^{1-l/2})$$

which yields the assertion.

□

References

Berger, J. O. and Delampady, M. (1987). Testing precise hypotheses. *Statistical Science*, 2:317–335.

Brillinger, D. R. (1981). *Time Series: Data Analysis and Theory.* McGraw Hill, New York.

Caiado, J., C. N. and Pena, D. (2009). Comparison of times series with unequal length in the frequency domain. *Communications in Statistics - Simulation and Computation*, 38:527–540.

Coates, D. and Diggle, P. (1986). Tests for comparing two estimated spectral densities. *Journal of Time Series Analysis*, 7:7–20.

De Souza, P. and Thomson, P. (1982). Lpc distance measures and statistical tests with particular reference to likelihood ratio. *IEEE Transactions Accoustics Speech Signal Process*, 30:304–315.

Dette, H. and Hildebrandt, T. (2011). A note on testing hypothesis for stationary processes in the frequency domain. *to appear Journal of Multivariate Analysis.*

Dette, H., Kinsvater, T., and Vetter, M. (2010). Testing nonparametric hypotheses for stationary processes by estimating minimal distances. *to appear in Journal of Time Series Analysis.*

Dette, H. and Paparoditis, E. (2009). Bootstrapping frequency domain tests in multivariate time series with an application to comparing spectral densities. *Journal of the Royal Statistical Society B*, 71:831–8571.

Dette, H., Preuß, P., and Vetter, M. (2011). A measure of stationarity in locally stationary processes with applications to testing. *to appear in Journal of the American Statistical Association.*

Diggle, P. and Fisher, N. (1991). Nonparametric comparison of cumulative periodograms. *Journal of the Royal Statistical Society C*, 40:423–434.

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

Jenkins, G. M. (1961). General considerations in the analysis of spectra. *Technometrics*, 8:133–166.

Jentsch, C. and Pauly, M. (2011). A note on using periodogram-based distances for comparing spectral densities. Preprint.

Maharaj, E. (2002). Comparison of non-stationary time series in the frequency domain. *Computational Statistical Data Analysis*, 40:131–141.

Pötscher, B. and Reschenhofer, E. (1988). Discriminating between two spectral densities in case of replicated observations. *Journal of Time Series Analysis*, 9:221–224.