Tests for comparing time-invariant and time-varying spectra based on the Anderson-Darling statistic

Shibin Zhang
Department of Mathematics, Shanghai Maritime University, 1550 Haigang Ave. In New Harbor City, Shanghai 201306, China.
E-mail: sbzhang@shmtu.edu.cn

Xin M. Tu
Division of Biostatistics and Bioinformatics, Department of Family Medicine and Public Health, UC San Diego, La Jolla, California 92093, USA.
E-mail: x2tu@ucsd.edu

Summary. Based on periodogram-ratios of two univariate time series at different frequency points, two tests are proposed for comparing their spectra. One is an Anderson-Darling-like statistic for testing the equality of two time-invariant spectra. The other is the maximum of Anderson-Darling-like statistics for testing the equality of two spectra no matter that they are time-invariant and time-varying. Both of two tests are applicable for independent or dependent time series. Several simulation examples show that the proposed statistics outperform those that are also based on periodogram-ratios but constructed by the Pearson-like statistics.

Keywords: Goodness-of-fit tests; Locally stationary time series; Anderson-Darling statistic; Periodogram; Spectral density

1. Introduction

Comparison of spectra has wide applications in many fields of research and practice. Research in statistical methodology for comparing spectra has been attracting considerable interest for several decades.

Most existing literature on comparison of spectra assume that spectra are time-invariant, e.g., Diggle and Fisher (1991), Dette et al. (2011a), Decowski and Li (2015) and Jentsch and Pauly (2015). By transforming the problem to the goodness-of-fit (gof) test that the periodogram-ratios of two univariate stationary time series at different frequencies are sampled from $F(2,2)$, Zhang and Tu (2017) recently proposed a Pearson-like statistic to test the equality of two time-invariant spectra, where $F(2,2)$ denotes an $F$ distribution with 2 and 2 degrees of freedom. Using blocking, they further extended the approach to the more general setting of comparing two time-varying spectra of locally stationary time series.

A major limitation of the gof test based on the Pearson statistic is its dependence on the choice of partition sets. In this paper, we propose a new approach to address
this key limitation. The proposed approach is based on Anderson-Darling (A-D) statistics of the form

\[ \hat{A}_n = \hat{A}_n(x_1, \ldots, x_n) = n \int_0^\infty \frac{(\hat{F}_n(x) - F(x))^2}{F(x)(1 - F(x))} dF(x), \] (1.1)

where \( F(x) \) is the cumulative distribution function (cdf), and \( \hat{F}_n(x) \) is the empirical cdf of observations \( x_1, \ldots, x_n \). In our setting of comparing two spectra, the observations are periodogram-ratios of two univariate time series at different frequencies. Our proposed statistics have some advantages. First, the proposed statistics are invariant with respect to the periodogram-ratios, i.e., the value of the statistic is unchanged after exchanging their places in the nominator and denominator. Second, the test is quite powerful. In the paper, the test for comparing spectra becomes the gof test that the periodogram-ratios are sampled from the \( F(2,2) \). Since the \( F(2,2) \) is heavy-tailed and has the reciprocal exchangeability, it is quite sensitive to detecting differences in the tails. It makes the A-D statistic quite powerful to reject the null since differences occur in the tails [Anderson and Darling, 1954]. Our simulation study also confirm the second feature of the approach.

The rest of this article is organized as follows. Sections 2 and 3 present the proposed statistics and their asymptotic properties for comparing time-invariant as well as time-varying spectra. Section 4 reports results for examining performance of the proposed tests using simulated data. Section 5 contains our concluding remarks. Some technical details regarding the asymptotic analysis are relegated to the Appendix. Throughout the paper, \( |a| \) and \( \overline{a} \) denote the complex modulus and conjugate of a complex number \( a \), respectively. For matrix notation, \( A^T \) denotes the transpose of a matrix \( A \), \( A^* \) denotes the conjugate transpose of a complex-value matrix \( A \), and \( \|A\| \) denotes the Euclidean norm of \( A \). For ease of notation, we write \( C \) for any generic positive constant.

Supplementary materials related to this article, including some R programs and a guide for using them, are available online.

### 2. Comparing time-invariant spectra

#### 2.1. Model

Let \( \{x_t, t \in \mathbb{Z}\} \) denote a bivariate stationary process with values in \( \mathbb{R}^2 \) which has a linear representation of the form

\[ x_t = (x_{1,t}, x_{2,t})^T = \sum_{n=-\infty}^\infty \Psi_n Z_{t-n}, \quad \text{where} \quad \sum_{n=-\infty}^\infty \|\Psi_n\| |n|^{1/2} < \infty, \] (2.1)

and \( \{Z_t, t \in \mathbb{Z}\} \) is a sequence of independent identically-distributed (i.i.d.) random variables, with mean \( 0 \), covariance matrix \( \Sigma \) and finite third absolute moment.

We assume that we have a sufficiently large number, \( T \), of observations from the bivariate zero-mean stationary time series in (2.1), \( x_t \). The spectral density matrix is given by

\[ f(\omega) = (f_{ij}(\omega))_{2 \times 2} = (2\pi)^{-1} A(\omega) \Sigma A^*(\omega) \] (2.2)
Comparing time-varying spectra based on A-D statistic

for \( \omega \in [-\pi, \pi] \), where \( A(\omega) = (A_{ij}(\omega))_{2 \times 2} = \sum_{n=-\infty}^{\infty} \Psi_n e^{-in\omega} \) (by e.g. Shumway and Stoffer, 2011, Chapter 4). From (2.2), we know that the spectral density \( f(\omega) \) is continuous with respect to \( \omega \). The discrete Fourier transform of the observation \( \{x_t, t = 1, \cdots , T\} \) is

\[
y_T(\omega) = \frac{1}{\sqrt{2\pi T}} \sum_{t=1}^{T} x_t e^{i\omega t},
\]

where \( \omega_k = 2\pi k/T, k = 0, 1, \cdots , [T/2] \). We extend the definition of \( y(\cdot) \) to a piecewise constant function on \([-\pi, \pi]\) as follows:

\[
y_T(\omega) = \begin{cases} y_T(\omega_k), & \text{if } \omega_k - \pi/T < \omega \leq \omega_k + \pi/T \text{ and } 0 \leq \omega \leq \pi, \\ y_T(-\omega), & \text{if } \omega \in [-\pi, 0). \end{cases}
\]

(2.3)

For any frequency \( \omega \in [-\pi, \pi] \), we define the periodogram by

\[
I_T(\omega) = (I_{T;ij}(\omega))_{2 \times 2} = y_T(\omega)y_T^*(\omega).
\]

Then the periodogram is also a piecewise constant function on \([-\pi, \pi]\), in accordance with the definition of Fuller (1996).

2.2. The test statistic

In this paper, we test the null:

\[
H_0 : f_{11}(\omega) = f_{22}(\omega) \text{ on } (0, \pi).
\]

(2.4)

By Lemma 2.1 of Zhang and Tu (2017), under the null (2.4), the periodogram-ratios

\[
I_{T;11}(l - 1/2\pi)/I_{T;22}(l - 1/2\pi), \quad l = 1, \cdots , L - 1
\]

are expected to behave like i.i.d. \( F(2, 2) \) distribution for a large \( L \) as \( T \) tends to infinity. Then the test for the null (2.4) is transformed to the gof test

\[
\frac{I_{T;11}(l - 1/2\pi)}{I_{T;22}(l - 1/2\pi)}, \quad l = 1, \cdots , L - 1, \quad \text{are sampled from } F(2, 2).
\]

(2.5)

Based on this idea, Zhang and Tu (2017) proposed a Pearson-like statistic, with a limiting chi-square distribution, to test (2.4). However, as it relies on the choice of partition sets, different choices may result in different test results.

The criterion proposed by Anderson and Darling (1952) covers a broad class of gof test statistics that are constructed by a measure of discrepancy or “distance” between the cdf and its empirical counterpart. Compared to others such as those proposed by Anderson and Darling (1952), the Anderson-Darling statistic (1.1) has two advantages when testing a sample drawn from \( F(2, 2) \). First, when \( F(x) = \frac{x}{1+x} \mathbb{I}_{(0, \infty)}(x) \) is the cdf of \( F(2, 2) \), the simplified expression of \( \hat{A}_n \) is

\[
\hat{A}_n(x_1, \cdots , x_n) = -n - \frac{1}{n} \sum_{i=1}^{n} (2i - 1) \left[ \log(x_{(i)}) - \log(1 + x_{(i)}) - \log(1 + x_{(n-i+1)}) \right],
\]

(2.6)
where \( x(1) < \cdots < x(n) \) is the sorted series of \( x_i, i = 1, \cdots, n \). From (2.6), we find \( A_n(x_1, \cdots, x_n) = A_n(1/x_1, \cdots, 1/x_n) \), keeping consistency for the periodogram-ratios in the sense that the value of \( A_n \) is unchanged after exchanging places of nominator and denominator. Second, since the \( F(2, 2) \) is heavy-tailed and has the reciprocal exchangeability (i.e., if \( X \sim F(2, 2) \) then \( X^{-1} \sim F(2, 2) \)), differences from it is prone to be detected in the tails. Thus, the A-D statistic is quite powerful to reject the null (Anderson and Darling, 1954).

Due to the properties of statistic (1.1), we propose an Anderson-Darling-like statistic

\[
\hat{A}_{T,L} = (L - 1) \int_0^{\infty} \frac{(\hat{F}_{T,L}(x) - F(x))^2}{F(x)(1 - F(x))} \, dF(x), \tag{2.7}
\]

to compare two time-invariant spectra, i.e. test (2.5), and to avoid choosing partitioning sets, where \( L > 1 \) is an integer,

\[
\hat{F}_{T,L}(x) = \frac{1}{L - 1} \sum_{i=1}^{L-1} \mathbf{1}_{(0,x)}(I_{T,11}(\frac{1-1/2}{L \pi}), I_{T,22}(\frac{1}{L \pi})), \tag{2.8}
\]

and \( F(x) = \int_{1/T}^{x} \mathbf{1}_{(0,\infty)}(t) \).

### 2.3. Sampling distribution Under the null

To obtain the asymptotic property of \( \hat{A}_{T,L} \) under the null hypothesis (2.4), we need the following condition.

**Condition 2.1.** As \( T \to \infty \), \( L/\sqrt{T} = O(1) \) or \( \to \infty \), and \( L \log T/T \to 0 \) (For clarity, we omit the dependence of \( L \) on \( T \)).

Then, we have the following theorem.

**Theorem 2.1.** Suppose that \( f_{i,i}(\omega), i = 1, 2 \), are bounded below away from zero and Lipschitz continuous, i.e., \( f_{i,i}(\omega) \geq \delta > 0 \) for all \( \omega \in (0, \pi) \) and \( i = 1, 2 \), and \( |f_{i,i}(\omega_1) - f_{i,i}(\omega_2)| \leq C|\omega_1 - \omega_2|, i = 1, 2 \), hold for any \( \omega_1 \neq \omega_2 \). Then, under Condition 2.1 if the null hypothesis (2.4) holds, we have

\[
\hat{A}_{T,L} \sim \int_0^1 \frac{B_0^2(t)}{t(1-t)} \, dt, \tag{2.9}
\]

where \( B_0(t) \) is a standard Brownian bridge on \([0, 1]\), and “\( \sim \)” denotes convergence in distribution.

**Proof.** For notational brevity, we write \( \ell_i \) and \( \ell_i' \) for \( \frac{l-1/2}{L} \pi \) and \( \frac{1}{L} \pi \), respectively, in the proof.

Let \( \eta_T(\omega) = (\eta_T^{(1)}(\omega), \eta_T^{(2)}(\omega))^T = \mathbf{A}(\omega) \mathbf{z}_T(\omega) \) for \( \omega = 2\pi k/T, k = 0, 1, \cdots, [T/2] \), where \( \mathbf{z}_T(\omega) = \frac{1}{\sqrt{2\pi T}} \sum_{t=1}^{T} Z_t e^{\mathbf{i} \omega t} \). Then we extend the definition of \( \eta_T(\cdot) \) to a piecewise constant function on \([ -\pi, \pi ] \) as in (2.3). For each integer \( L > 1 \), we define

\[
\tilde{A}_{T,L} = (L - 1) \int_0^{\infty} \frac{(\tilde{F}_{L}(x) - F(x))^2}{F(x)(1 - F(x))} \, dF(x)
\]
Comparing time-varying spectra based on A-D statistic

and

\[ \hat{\delta}_{T,L} = (L - 1) \int_0^\infty \frac{\tilde{F}_L(x) - F'(x)^2}{F(x)(1 - F(x))} \, dF(x), \]

where

\[ \tilde{F}_L(x) = \frac{1}{L - 1} \sum_{l=1}^{L-1} \mathbb{I}_{(0,x]} \left( \frac{I_{T,11}(\ell_l)}{I_{T,22}(\ell'_l)} \right) \]

and

\[ \check{F}_L(x) = \frac{1}{L - 1} \sum_{l=1}^{L-1} \mathbb{I}_{(0,x]} \left( \frac{\eta_T^{(1)}(\ell_l)}{\eta_T^{(2)}(\ell'_l)} \right)^2 \]

for \( x \in (0, \infty) \), respectively.

We prove in the Appendix that

\[ \tilde{A}_{T,L} = \check{A}_{T,L} + o_p\left( \sqrt{L \log T} \right) \]  \hspace{1cm} (2.10)

and

\[ \hat{A}_{T,L} = \tilde{A}_{T,L} + o_p\left( \sqrt{L \log T} \right) \]  \hspace{1cm} (2.11)

hold as \( T \) goes to infinity, where \( X_T = o_p(T) \) means \( X_T/T \) converges to zero in probability.

Combining (2.10) with (2.11), we obtain

\[ \hat{A}_{T,L} = \check{A}_{T,L} + o_p\left( \sqrt{L \log T} \right) \]  \hspace{1cm} (2.12)

According to the arguments in the proof of Lemma 2.1 of Zhang and Tu (2017), \( (\eta_T^{(1)}(\ell_l), \eta_T^{(2)}(\ell'_l))^T \) is a function of \( (\eta_T(\ell_l), \eta_T(\ell'_l))^T \), which is an asymptotically zero-mean complex normal random variable with covariance matrix \( \text{diag}(f_{11}(\ell_l), f_{22}(\ell'_l)) \), and independent among \( l = 1, \cdots, L - 1 \). Then we have that

\[ \frac{\eta_T^{(1)}(\ell_l)}{\eta_T^{(2)}(\ell'_l)} \]

is a sequence of i.i.d. random variables with an asymptotic \( F(2,2) \). According to the continuous mapping theorem (e.g. van der Vaart, 1998, pp.6-7) and Theorem 26.1 of DasGupta (2008), we obtain (2.9). This completes the proof.

By Theorem 2.1, the statistic (2.7) can be used to test the null hypothesis (2.4).

Given the pre-specified level of significance \( \alpha \) (0 < \( \alpha \) < 1), the null hypothesis (2.4) is rejected if

\[ \hat{A}_{T,L} > a_{1 - \alpha}, \]  \hspace{1cm} (2.14)

where \( a_{1 - \alpha} \) denotes the 1 – \( \alpha \) quantile of the A-D test statistic with sample size tending to infinity. Under the null, the rejection probability of test (2.14) converges to \( \alpha \) as \( T \) goes to infinity. The cdf of the null distribution of the A-D statistic can be computed by using the algorithm of Marsaglia and Marsaglia (2004). The quantiles can be computed by root-finding. The functions for computation can be found in the R package gof test (Faraway et al., 2015).
3. An extension for comparing time-varying spectra

3.1. Model

We base our discussion on a multivariate extension (cf. Zhang and Tu, 2017) of the model of Dette et al. (2011b). It covers a more general class of processes than the one introduced by Dahlhaus (1997).

Let \( \{x_{t,T}, t = 1, \cdots, T\} \) be a bivariate locally stationary time series, where each observation \( x_{t,T} \) exhibits a linear representation of the form

\[
   x_{t,T} = \sum_{n=-\infty}^{\infty} \Psi_{t,T,n} Z_{t-n}, \quad t = 1, \cdots, T. \tag{3.1}
\]

We also assume that \( \{Z_t, t \in \mathbb{Z}\} \) is a sequence of i.i.d. random variables, with mean 0, covariance matrix \( \Sigma \) and finite third absolute moment. We further assume the coefficients \( \{\Psi_{t,T,n}\} \) behave like some smooth functions in a neighborhood of time \( t/T \). Therefore we adopt not only the usual condition \( \sum_{n=-\infty}^{\infty} \|\Psi_{t,T,n}\| < \infty \), but impose additionally that there exists a matrix-valued function \( \Psi_n : [0,1] \to \mathbb{R}^{2 \times 2} \) and a constant \( C \) with

\[
   \sum_{n=-\infty}^{\infty} \sup_{t=1,\cdots,T} \|\Psi_{t,T,n} - \Psi_n(t/T)\| \leq C \frac{1}{T}. \tag{3.2}
\]

Furthermore, we assume that

\[
   \sum_{n=-\infty}^{\infty} \sup_{u \in [0,1]} \|\Psi_n(u)\| \|n\|^{1/2} < \infty, \tag{3.3}
\]

and each element of \( \Psi_n \) is a continuously differentiable function of \( u \), and

\[
   \sum_{n=-\infty}^{\infty} \sup_{i,j} \sup_{u \in [0,1]} |d\Psi_n(i,j)(u)| < \infty. \tag{3.4}
\]

Zhang and Tu (2017) gave a detailed discussion on the conditions (3.2)-(3.4). The time-varying spectral density of the locally stationary process \( \{x_{t,T}\} \) is defined in terms of auxiliary function \( \Psi_n \) (cf. Dette et al., 2011b), that is,

\[
   f(u,\omega) = (f_{ij}(u,\omega))_{2 \times 2} = (2\pi)^{-1} A(u,\omega) \Sigma A^*(u,\omega), \quad u \in [0,1], \ \omega \in [-\pi, \pi] \tag{3.5}
\]

where \( A(u,\omega) = \sum_{n=-\infty}^{\infty} \Psi_n(u) e^{-in\omega} \).

Assume without loss of generality that the total sample size \( T \) can be decomposed as \( T = MB \), where \( B \) is an integer and \( M \) is an even integer. The main idea is to split the entire data into \( B \) blocks with \( M \) observations each, from which we define appropriate local periodograms. Specifically, let

\[
   I_M(u,\omega_k) := (I_{M,ij}(u,\omega_k))_{2 \times 2} = y_M(u,\omega_k) y_M^*(u,\omega_k). \tag{3.6}
\]
Comparing time-varying spectra based on A-D statistic

be the usual periodogram around \( u \) computed from \( M \) observations, that is, we set
\[
y_M(u, \omega_k) = \frac{1}{\sqrt{2\pi M}} \sum_{s=1}^{M} x_{[uT]-M/2+s} e^{i\omega_k s} \tag{3.7}
\]
and \( x_{t,T} = 0 \), if \( t \notin \{1, \cdots, T\} \) \cite{Dahlhaus, 1997}, where \( \omega_k = 2\pi k/M \), \( k = 0, 1, \cdots, [M/2] \). For each \( u \), we extend the definition of \( y_M(u, \cdot) \) to a piecewise constant function on \( [0, 1] \times [-\pi, \pi] \), the same extension method as in (2.3). Accordingly, we extend the definition of \( I_M(\cdot, \cdot) \) in (3.6) to the periodograms on \( [0, 1] \times [-\pi, \pi] \).

3.2. The test statistic
In this section, we consider the problem of testing
\[
H_0: f_{11}(u, \omega) = f_{22}(u, \omega) \quad \text{on} \quad \{\omega: \omega \in (0, \pi)\} \quad \text{for each} \quad u \in [0, 1]. \tag{3.8}
\]
We use the notation \( u_k = \frac{(k-1)M+M/2}{T} \) \( (k = 1, \cdots, B) \) for the mid-point of each block, and for each integer \( L > 1 \), define
\[
\hat{F}_{M,L}(x) = \frac{1}{L-1} \sum_{l=1}^{L-1} \mathbb{I}_{[0,x]} \left( \frac{I_{M,11}(u_k, l-1/2\pi)}{I_{M,22}(u_k l/\pi)} \right). \tag{3.9}
\]
Let
\[
\hat{A}_{M,L}^{(k)} = (L-1) \int_0^\infty \frac{(\hat{F}_{M,L}(x) - F(x))^2}{F(x)(1-F(x))} \, dF(x), \tag{3.10}
\]
where \( \hat{F}_{M,L}(x) \) is defined by (3.9), and \( F(x) \) is defined as in (2.8). In the \( k \)-th block, we use (3.10) to compare two local spectra. Intuitively, we should reject (3.8) when at least one of \( \hat{A}_{M,L}^{(k)} \), \( k = 1, \cdots, B \), exceeds a pre-specified threshold. Therefore, we propose a maximum of local A-D statistics
\[
M_{B,M,L} = \max_{k=1, \cdots, B} \{ \hat{A}_{M,L}^{(k)} \} \tag{3.11}
\]
to test (3.8).

3.3. Sampling distribution under the null
To study the asymptotic behavior of \( M_{B,M,L} \) under the null, we need the following condition.

**Condition 3.1.** As \( T \to \infty, M \to \infty, M/\sqrt{T} = O(1) \) or \( \to 0, L/\sqrt{M} = O(1) \) or \( \to \infty \), and \( L \log M/M \to 0 \) (For clarity, we omit the dependence of \( L \) and \( M \) on \( T \)).

With arguments similar to the proof of Theorem 3.1 of Zhang and Tu \(2017\), we obtain the asymptotic properties of (3.11) under the null as follows.
Theorem 3.1. Suppose that the locally stationary time series \( \{x_{t,T}, t = 1, \cdots, T\} \) \((T \in \mathbb{N})\) satisfies conditions (3.2)-(3.4), the diagonal elements of \( f(u, \omega) \) are bounded blow away from 0 for all \((u, \omega) \in [0, 1] \times (0, \pi)\), and \( f_{11}(u, \omega) \) and \( f_{22}(u, \omega) \) are uniformly Lipschitz continuous with respect to \( \omega \) for \( u \in [0, 1] \), i.e. \( |f_{ii}(u, \omega_1) - f_{ii}(u, \omega_2)| \leq C|\omega_1 - \omega_2|, i = 1, 2, \) hold for any \( \omega_1 \neq \omega_2 \) and some positive constant \( C \) independent of \( u \) and \( \omega \). Then, under Condition 3.1 if the null hypothesis (3.8) holds, we have
\[
M_{B,M,L} = \max_{k=1,\cdots,B} \{\tilde{A}_{M,L}^{(k)}\} + o_{p}\left(\frac{\sqrt{L \log M}}{\sqrt{M}}\right),
\]
where \( \tilde{A}_{M,L}^{(k)} \), \( k = 1, \cdots, B \), are i.i.d. random variables for each fixed \( B \). Moreover, for each fixed \( k \), \( \tilde{A}_{M,L}^{(k)} \sim \int_0^1 \frac{B_i^2(t)}{t(1-t)} \, dt \) as \( T \) goes to infinity.

We propose the statistic (3.11) to test the null hypothesis (3.8). According to Theorem 3.1, the cumulative distribution function (cdf) of \( M_{B,M,L} \) is approximately equal to that of \( \max_{k=1,\cdots,B} \{A_k\} \) in distribution, where \( A_k \), \( k = 1, \cdots, B \), are i.i.d. random variables. The cdf of \( \max_{k=1,\cdots,B} \{A_k\} \) is \( F_M(x) = F_A(x)^B \), where \( F_A(x) \) is the cdf function of \( \int_0^1 \frac{B_i^2(t)}{t(1-t)} \, dt \). For a pre-specified level of significance \( \alpha \) (\( 0 < \alpha < 1 \)), the null hypothesis (3.8) is rejected if
\[
M_{B,M,L} > \kappa_{1-\alpha}(B),
\]
where \( \kappa_{1-\alpha}(B) \) denotes the \( 1 - \alpha \) quantile of the distribution of \( \max_{k=1,\cdots,B} \{A_k\} \). Under the null hypothesis, the rejection probability of test (3.13) is approximately equal to \( \alpha \).

The test (3.13) is convenient to use since \( \kappa_{1-\alpha}(B) \) can be easily obtained by solving \( F_A(x)^B = 1 - \alpha \).

4. Numerical examples

In this section, we use simulations to evaluate small-sample performance of the proposed statistics (2.7) and (3.11) when applied to test the null hypotheses (2.4) and (3.8). All results are based on 1,000 Monte Carlo replications unless stated otherwise.

In all the simulation examples of Section 4.1 we set \( L = \min\{[T/4], [T^{3/4}]\} \). In the examples of Section 4.2 we set \( B = [\sqrt{T}/5], M = [T/B] \), and \( L = \min\{[M/4], [M^{3/4}]\} \). These settings all satisfy Condition 2.1 or 3.1.

4.1. Comparing time-invariant spectra

To show the performance of test (3.13), we consider nine stationary time series \( x_t = (x_{1,t}, x_{2,t})' \), \( t = 1, \cdots, T \) in this section. Their components are given as follows.

Model A (Copied MA(1) model.) \( x_{i,t} = Z_{i,t} - 0.8Z_{i,t-1}, i = 1, 2 \).

Model B (Copied MA(2) model.) \( x_{i,t} = Z_{i,t} - 0.8Z_{i,t-1} - 0.5Z_{i,t-2}, i = 1, 2 \).
Comparing time-varying spectra based on A-D statistic

Model C (Copied AR(1) model.) \( x_{i,t} = 0.5x_{i,t-1} + \sqrt{0.75}Z_{i,t}, i = 1, 2. \)
Model D (Copied ARMA(1,1) model.) \( x_{i,t} = 0.5x_{i,t-1} + Z_{i,t} - 0.5Z_{i,t-1}, i = 1, 2. \)
Model E (Copied AR(2) model.) \( x_{i,t} = 0.5x_{i,t-1} - 0.5x_{i,t-2} + (1/\sqrt{1.5})Z_{i,t}, i = 1, 2. \)
Model F (MA(2)-MA(1) model.) \( x_{1,t} = Z_{1,t} - 0.8Z_{1,t-1} - 0.5Z_{1,t-2} \) and \( x_{2,t} = Z_{2,t} - 0.8Z_{2,t-1}. \)
Model G (AR(1)-ARMA(1,1) model.) \( x_{1,t} = 0.5x_{1,t-1} + \sqrt{0.75}Z_{1,t} \) and \( x_{2,t} = 0.5x_{2,t-1} + Z_{2,t} - 0.5Z_{2,t-1}. \)
Model H (AR(1)-AR(2) model.) \( x_{1,t} = 0.5x_{1,t-1} + \sqrt{0.75}Z_{1,t} \) and \( x_{2,t} = 0.5x_{2,t-1} - 0.5x_{2,t-2} + (1/\sqrt{1.5})Z_{2,t}. \)
Model I (AR(2)-AR(2) model.) \( x_{1,t} = 0.5x_{1,t-1} - 0.5x_{2,t-2} + (1/\sqrt{1.5})Z_{1,t} \) and \( x_{2,t} = 0.6x_{2,t-1} - 0.6x_{2,t-2} + \sqrt{0.55}Z_{2,t}. \)

In Models A-I, \( Z_t = (Z_{1,t}, Z_{2,t})^T \) is an independent centered stationary Gaussian process with covariance matrix \( \Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \). In each model, we consider two different choices of \( \rho \), 0.1 and 0.5. For each of Models C-E and G-I, we have standardized both components such that both of them have variance 1.

For each sample size \( T = 128, 256, 512, 1024 \) and each nominal size \( \alpha = 5\%, 10\%, 15\% \), Tables 1 and 2 present empirical rejection probabilities of tests (2.14) for Models A-E and F-I, respectively.

Results in Table 1 show that each empirical type I error rate is very close to the pre-specified nominal value. We conclude from Table 2 as follows. First, as \( T \) increases, the performance of test (2.14) improves significantly. Second, the empirical power of test (2.14) has no clear relationship with the correlation \( \rho \) of the process \( Z_t \). The test statistic (2.7) has efficiently stripped the dependence between components of the time series. Third, the proposed test outperforms the competitive tests based on the Pearson statistic proposed by Zhang and Tu (2017). Models F-I are also employed in the simulation studies of Zhang and Tu (2017), with identical settings. By comparing Table 2 with Tables 1 and S.1 of Zhang and Tu (2017), we find for each parameter setting, the test (2.14) has higher empirical power than those in Zhang and Tu (2017). By our simulation experiments that are not reported here for space consideration, we also find that the empirical power of test (2.14) is quite robust when exchanging the order of components of the time series, i.e., replacing \( (x_{1,t}, x_{2,t})^T \) with \( (x_{2,t}, x_{1,t})^T \).

4.2. Comparing time-varying spectra

To demonstrate small-sample performance of test (3.13), we consider nine non-stationary time series \( x_{i,T} = (x_{1,t,T}, x_{2,t,T})^T, t = 1, \ldots, T, \) in this section. Their components are as follows.

Model J (Copied smoothly-varying (SV) MA(1) model.) \( x_{i,t,T} = Z_{i,t} - \beta_1(t/T)Z_{i,t-1}, i = 1, 2, \) where \( \beta_1(u) = 0.8(1 + \sin(\pi u/2)) \) for \( u \in [0, 1] \).
Model K (Copied SV AR(1) model.) \( x_{i,t,T} = \phi(t/T)x_{i,t-1,T} + Z_{i,t}, i = 1, 2, \) where \( \phi(u) = 0.6\sin(4\pi u) \) for \( u \in [0, 1] \).
Table 1. Rejection probabilities of tests (2.14) for the hypothesis of equal spectral density in models A-E from simulated data.

| $T$  | $\rho = 0.1$ |       |       | $\rho = 0.5$ |       |       |
|------|--------------|-------|-------|--------------|-------|-------|
|      | 5%           | 10%   | 15%   | 5%           | 10%   | 15%   |
| Model A |              |       |       |              |       |       |
| 128   | 0.044        | 0.081 | 0.133 | 0.041        | 0.077 | 0.133 |
| 256   | 0.067        | 0.131 | 0.181 | 0.075        | 0.120 | 0.168 |
| 512   | 0.067        | 0.120 | 0.171 | 0.060        | 0.125 | 0.156 |
| 1024  | 0.046        | 0.094 | 0.155 | 0.039        | 0.090 | 0.140 |
| Model B |              |       |       |              |       |       |
| 128   | 0.056        | 0.104 | 0.150 | 0.050        | 0.088 | 0.148 |
| 256   | 0.056        | 0.111 | 0.168 | 0.045        | 0.103 | 0.158 |
| 512   | 0.066        | 0.106 | 0.155 | 0.073        | 0.118 | 0.172 |
| 1024  | 0.056        | 0.101 | 0.164 | 0.050        | 0.097 | 0.145 |
| Model C |              |       |       |              |       |       |
| 128   | 0.050        | 0.098 | 0.140 | 0.042        | 0.089 | 0.135 |
| 256   | 0.065        | 0.127 | 0.176 | 0.057        | 0.119 | 0.165 |
| 512   | 0.061        | 0.110 | 0.162 | 0.054        | 0.109 | 0.150 |
| 1024  | 0.044        | 0.096 | 0.148 | 0.044        | 0.098 | 0.137 |
| Model D |              |       |       |              |       |       |
| 128   | 0.046        | 0.095 | 0.130 | 0.041        | 0.080 | 0.122 |
| 256   | 0.058        | 0.121 | 0.172 | 0.058        | 0.113 | 0.162 |
| 512   | 0.060        | 0.121 | 0.160 | 0.053        | 0.105 | 0.155 |
| 1024  | 0.045        | 0.098 | 0.148 | 0.037        | 0.091 | 0.137 |
| Model E |              |       |       |              |       |       |
| 128   | 0.056        | 0.118 | 0.193 | 0.053        | 0.112 | 0.167 |
| 256   | 0.056        | 0.093 | 0.148 | 0.057        | 0.106 | 0.149 |
| 512   | 0.030        | 0.077 | 0.121 | 0.038        | 0.093 | 0.134 |
| 1024  | 0.050        | 0.112 | 0.164 | 0.064        | 0.111 | 0.166 |
Table 2. Rejection probabilities of tests (2.14) for the hypothesis of equal spectral density in models F-I from simulated data.

|    | $T$ | $\rho = 0.1$ | $\rho = 0.5$ |
|----|-----|--------------|--------------|
|    | 5%  | 10%          | 15%          | 5%  | 10%          | 15%          |
|    |     |              |              |     |              |              |
| Model F |     |              |              |     |              |              |
| 128 | 0.183 | 0.279 | 0.358 | 0.197 | 0.279 | 0.347 |
| 256 | 0.425 | 0.541 | 0.614 | 0.417 | 0.533 | 0.606 |
| 512 | 0.617 | 0.724 | 0.782 | 0.614 | 0.727 | 0.783 |
| 1024 | 0.835 | 0.899 | 0.929 | 0.829 | 0.895 | 0.922 |
| Model G |     |              |              |     |              |              |
| 128 | 0.128 | 0.228 | 0.299 | 0.130 | 0.206 | 0.295 |
| 256 | 0.285 | 0.410 | 0.488 | 0.273 | 0.384 | 0.447 |
| 512 | 0.443 | 0.566 | 0.628 | 0.440 | 0.567 | 0.628 |
| 1024 | 0.639 | 0.765 | 0.828 | 0.642 | 0.750 | 0.809 |
| Model H |     |              |              |     |              |              |
| 128 | 0.089 | 0.162 | 0.224 | 0.072 | 0.153 | 0.223 |
| 256 | 0.117 | 0.209 | 0.286 | 0.111 | 0.201 | 0.280 |
| 512 | 0.138 | 0.235 | 0.321 | 0.142 | 0.245 | 0.330 |
| 1024 | 0.260 | 0.388 | 0.480 | 0.256 | 0.396 | 0.504 |
| Model I |     |              |              |     |              |              |
| 128 | 0.129 | 0.200 | 0.254 | 0.113 | 0.183 | 0.244 |
| 256 | 0.166 | 0.242 | 0.304 | 0.147 | 0.240 | 0.315 |
| 512 | 0.182 | 0.281 | 0.354 | 0.174 | 0.269 | 0.350 |
| 1024 | 0.288 | 0.419 | 0.488 | 0.320 | 0.428 | 0.507 |
Model L (Copied wavelet process.) \( x_{i,t,T} = \sum_{k=0}^{T-1} w_1(k/T)\psi_{1,k-t}Z_{i,k}, \ i = 1, 2, \)
where \( w_1(u) = \cos(\pi u/2) \) for \( u \in [0,1] \) and \( \psi_{1,k} = \frac{1}{\sqrt{2}}(I_{0}(k) - I_{1}(k)) \) for \( k \in \mathbb{Z}. \)
This type of process was also considered by van Bellegem and von Sachs (2008).

Model M (Copied Cholesky-decomposition model.) The model is given by

\[
\mathbf{x}_{t,T} = \sum_{k=1}^{T} \Phi(t/T, k/T) \exp(2\pi kti/T)\mathbf{\epsilon}_k,
\]
\( k = 1, \cdots, n \) are independent. Moreover, for \( k/T \neq 0, 0.5, 1, \) \( \mathbf{\epsilon}_k \)
follows a bivariate normal distribution with mean zero and covariance \( n^{-1}\mathbf{I}_2, \) and \( \mathbf{\epsilon}_k = \mathbf{\epsilon}_{n-k}, \) for \( k/T = 0, 0.5, 1, \) \( \mathbf{\epsilon}_k \)
follows a bivariate normal distribution with mean zero and covariance \( n^{-1}\mathbf{I}_2. \) The process (4.1) was constructed with the pre-specified spectral matrix \( f_u, \) \( \alpha = \Phi(u, \omega)\{\Phi(u, \omega)\}^* \) for \( u \in [0,1] \) and \( \omega \in (0, \pi) \) (Guo and Dai, 2006, Theorem 3.1). In model (4.1), We set \( \Phi(u,v) = (\psi_{ij}(u,v), \sum_{\omega=0}^{2}\sin(2\pi u)), \) with \( \psi_{ij}(u,v) = \psi_{11}(u,v) = \psi_{12}(u,v) = \{1.2\cos(2\pi u) + 0.3\sin(2\pi u) + 0.7\} \)
for \( u \in [0,1]. \)

Model N (SV MA(2)-SV MA(1) model.) \( x_{1,1,T} = Z_{1,1,t} - \beta_{1}(t/T)Z_{1,1,t-1} - \beta_{2}(t/T)Z_{1,1,t-2} \)
and \( x_{2,1,T} = Z_{2,1,t} - \beta_{1}(t/T)Z_{2,1,t-1}, \) where \( \beta_{1}(u) \) is given as in Model J and \( \beta_{2}(u) = 0.5(1 - \cos(\pi u)) \) for \( u \in [0,1]. \)

Model O (Abruptly-varying MA(2)-MA(2) model.) \( x_{1,1,T} = Z_{1,1,t} - 0.8Z_{1,1,t-1} - (0.5 - \gamma(t/T))Z_{1,1,t-2} \) and \( x_{2,1,T} = Z_{2,1,t} - 0.8Z_{2,1,t-1} - 0.5Z_{1,1,t-2}, \) where \( \gamma(u) = \mathbb{I}_{[0,0.5]}(u) \) for \( u \in [0,1]. \)

Model P (SV AR(1)-SV AR(1) model.) \( x_{1,1,T} = \phi(t/T)x_{1,1,t-1,T} + 1.5Z_{1,1,t} \) and \( x_{2,1,T} = \phi(t/T)x_{2,1,t-1,T} + Z_{2,1,t}, \) where \( \phi(u) \) is set as in Model K.

Model Q (Wavelet-wavelet process.) \( x_{1,1,T} = \sum_{k=0}^{T-1} w_1(k/T)\psi_{1,k-t}Z_{1,k} \)
\[ + \sum_{k=0}^{T-1} w_2(k/T)\psi_{2,k-t}Z_{2,k}, \]
where \( w_1(u) = \sum_{\omega=0}^{2}\sin(2\pi u) + 0.3w_2(u) = 0.3u^2 \) for \( u \in [0,1] \) and \( \psi_{1,k} = \frac{1}{2}(I_{0}(k) - I_{1}(k)) \)
for \( k \in \mathbb{Z}. \)

Model R (Cholesky-decomposition model.) The model is given by (4.1) but we set \( \psi_{11}(u,v) \) as in Model M and \( \psi_{22}(u,v) = \{1.2\cos(2\pi u)^2 + 0.6\sin(2\pi u) + 0.7\} \)
for \( u \in [0,1]. \) This choice makes \( f_{22}(u, \omega) \) changes more rapidly over time than \( f_{11}(u, \omega). \)

In Models J-L and N-Q, \( Z_t = (Z_{1,t}, Z_{2,t})^T \) and its covariance matrix are defined the same way as in Models A-H. In each model, we only consider the case of \( \rho = 0.5. \)
For each model in Models J-M, the two components have equal time-varying spectral densities, while for each model in Models N-R, the two marginal spectral densities are unequal. For each model of Models N and P-R, both marginal spectra change smoothly over time; while for Model O, one changes smoothly and the other changes abruptly.

To illustrate that the test (3.13) still works well when applied to testing the null (2.24), we consider stationary Models A, B and F, again. In each of these three models, we only consider the case of \( \rho = 0.5. \)

For each sample size \( T = 128, 256, 512, 1024 \) and each nominal size \( \alpha = 5\%, 10\%, 15\%, \) Table 3 presents empirical rejection probabilities of test (3.13) for Models J-M and A-B, and Table 4 for Models N-R and F.
Comparing time-varying spectra based on A-D statistic

Table 3. Rejection probabilities of test (3.13) for the hypothesis of equal time-varying spectral density in Models J-M and A-B from simulated data.

| $T$ | $B$ | 5% | 10% | 15% | 5% | 10% | 15% |
|-----|-----|-----|-----|-----|-----|-----|-----|
|     | Model J |     |     |     | Model K |     |     |     |
| 128 | 2   | 0.057 | 0.119 | 0.176 | 0.055 | 0.094 | 0.135 |
| 256 | 3   | 0.062 | 0.116 | 0.174 | 0.055 | 0.106 | 0.158 |
| 512 | 4   | 0.066 | 0.132 | 0.180 | 0.059 | 0.096 | 0.135 |
| 1024| 6   | 0.054 | 0.109 | 0.161 | 0.045 | 0.102 | 0.162 |
|     | Model L |     |     |     | Model M |     |     |     |
| 128 | 2   | 0.066 | 0.125 | 0.186 | 0.048 | 0.100 | 0.166 |
| 256 | 3   | 0.106 | 0.163 | 0.223 | 0.055 | 0.099 | 0.149 |
| 512 | 4   | 0.078 | 0.141 | 0.206 | 0.055 | 0.102 | 0.153 |
| 1024| 6   | 0.083 | 0.152 | 0.218 | 0.038 | 0.093 | 0.154 |
|     | Model A |     |     |     | Model B |     |     |     |
| 128 | 2   | 0.053 | 0.119 | 0.177 | 0.064 | 0.118 | 0.159 |
| 256 | 3   | 0.062 | 0.126 | 0.190 | 0.045 | 0.090 | 0.147 |
| 512 | 4   | 0.071 | 0.130 | 0.186 | 0.051 | 0.105 | 0.145 |
| 1024| 6   | 0.049 | 0.099 | 0.156 | 0.057 | 0.123 | 0.171 |

Table 4. Rejection probabilities of test (3.13) for the hypothesis of equal time-varying spectral density in Models N-R and F from simulated data.

| $T$ | $B$ | 5% | 10% | 15% | 5% | 10% | 15% |
|-----|-----|-----|-----|-----|-----|-----|-----|
|     | Model N |     |     |     | Model O |     |     |     |
| 128 | 2   | 0.114 | 0.204 | 0.277 | 0.182 | 0.301 | 0.389 |
| 256 | 3   | 0.205 | 0.328 | 0.404 | 0.168 | 0.288 | 0.366 |
| 512 | 4   | 0.353 | 0.492 | 0.590 | 0.332 | 0.466 | 0.571 |
| 1024| 6   | 0.514 | 0.662 | 0.753 | 0.466 | 0.618 | 0.715 |
|     | Model P |     |     |     | Model Q |     |     |     |
| 128 | 2   | 0.529 | 0.658 | 0.732 | 0.099 | 0.171 | 0.230 |
| 256 | 3   | 0.766 | 0.862 | 0.901 | 0.249 | 0.360 | 0.443 |
| 512 | 4   | 0.955 | 0.988 | 0.991 | 0.764 | 0.851 | 0.879 |
| 1024| 6   | 0.999 | 1.000 | 1.000 | 0.999 | 1.000 | 1.000 |
|     | Model R |     |     |     | Model F |     |     |     |
| 128 | 2   | 0.122 | 0.209 | 0.287 | 0.117 | 0.185 | 0.261 |
| 256 | 3   | 0.163 | 0.265 | 0.344 | 0.173 | 0.272 | 0.379 |
| 512 | 4   | 0.261 | 0.380 | 0.497 | 0.307 | 0.442 | 0.533 |
| 1024| 6   | 0.714 | 0.812 | 0.864 | 0.408 | 0.586 | 0.692 |
From Table 3, we can see that the empirical size is close to the nominal level for all models considered. From Table 4, we can also observe that all deviations from equality of time-varying spectra are detected with reasonably large probabilities. Moreover, the proposed test outperforms the competitions based on the maximum of Pearson statistics proposed by Zhang and Tu (2017). Models N-O and Q-R were also employed in the simulation studies of Zhang and Tu (2017), with identical settings. By comparing Table 4 with Tables 3-4 of Zhang and Tu (2017), we find for each parameter setting, the test (2.14) has higher empirical power than its counterpart test proposed by Zhang and Tu (2017). When applied to testing the null (2.4), the statistic (3.11) also outperforms the statistic (3.1) of Zhang and Tu (2017).

5. Conclusion

In this paper, we proposed two test statistics, (2.7) and (3.11), for comparing time-invariant and time-varying spectra. As in Zhang and Tu (2017), the test problems are transformed to the setting of goodness-of-fit tests. The test statistic (2.7) is constructed with the form as the A-D statistic. The test statistic (3.11) is constructed by first computing local A-D statistics and then maximizing them. Like the tests proposed by Zhang and Tu (2017), they have two advantages. First, it is easy to program and quite computationally efficient. Second, the proposed test statistic (3.11) has a wide range of applicability in that it is applicable to stationary and locally stationary time series, with either independent or dependent components. Moreover, the proposed tests are independent of the choice of partitioning sets since they are based on the A-D statistic. In our simulation examples, we also find the proposed tests outperform those based on the Pearson statistic proposed by Zhang and Tu (2017).

Our recommended guidelines for implementing the proposed tests, such as the choice of \( L \) and \( B \), and extensions to multiple spectra, are all the same as those discussed in Zhang and Tu (2017).

Appendix: Technical Details

This appendix contains details of the proofs of (2.10) and (2.11).

Note that

\[
\tilde{A}_{T,L} = (L - 1) \int_0^\infty \frac{\tilde{F}_L(x) - \hat{F}_L(x)^2}{F(x)(1 - F(x))} dF(x) \\
+ 2 \int_0^\infty \frac{(\sqrt{L} - 1)(\tilde{F}_L(x) - \hat{F}_L(x))(\sqrt{L} - 1)(\tilde{F}_L(x) - F(x))}{F(x)(1 - F(x))} dF(x) \\
+ \hat{A}_{T,L}.
\]

(A.1)

To prove (2.10), it suffices to verify

\[
\sup_{x > 0} \sqrt{L - 1} |\tilde{F}_L(x) - \hat{F}_L(x)| = o_p\left( \frac{\sqrt{L \log T}}{\sqrt{T}} \right)
\]

(A.2)
Comparing time-varying spectra based on A-D statistic

and

$$\int_0^\infty \frac{\sqrt{L - 1}(\tilde{F}_L(x) - \hat{F}_L(x))}{F(x)(1 - F(x))} \, dF(x) = o_p\left(\frac{\sqrt{L \log T}}{\sqrt{T}}\right).$$  \hspace{1cm} (A.3)

If \((A.2)\) and \((A.3)\) hold, we have

$$(L - 1) \int_0^\infty \frac{(\tilde{F}_L(x) - \hat{F}_L(x))^2}{F(x)(1 - F(x))} \, dF(x) = \int_0^\infty \frac{(\sqrt{L - 1}(\tilde{F}_L(x) - \hat{F}_L(x)))^2}{F(x)(1 - F(x))} \, dF(x) \leq o_p\left(\frac{\sqrt{L \log T}}{\sqrt{T}}\right).$$  \hspace{1cm} (A.4)

Moreover, by the Cauchy inequality, we have

$$\left| \int_0^\infty \frac{(\sqrt{L - 1}(\tilde{F}_L(x) - \hat{F}_L(x)))^2}{F(x)(1 - F(x))} \, dF(x) \right| \leq \sqrt{(L - 1) \int_0^\infty \frac{(\tilde{F}_L(x) - \hat{F}_L(x))^2}{F(x)(1 - F(x))} \, dF(x)} \cdot \hat{A}_{T,L} = o_p\left(\frac{\sqrt{L \log T}}{\sqrt{T}}\right).$$  \hspace{1cm} (A.5)

Combining \((A.1)\), \((A.4)\) and \((A.5)\), we obtain \((2.10)\).

Similarly, to prove \((2.11)\), it suffices to verify

$$\sup_{x > 0} \sqrt{L - 1} |\tilde{F}_L(x) - \hat{F}_L(x)| = o_p\left(\frac{\sqrt{L \log T}}{\sqrt{T}}\right)$$  \hspace{1cm} (A.6)

and

$$\int_0^\infty \frac{\sqrt{L - 1}(\tilde{F}_L(x) - \hat{F}_L(x))}{F(x)(1 - F(x))} \, dF(x) = o_p\left(\frac{\sqrt{L \log T}}{\sqrt{T}}\right).$$  \hspace{1cm} (A.7)

The rest of this appendix pertains to verification of \((A.2)\), \((A.3)\), \((A.6)\) and \((A.7)\).

**Proof of \((A.2)\) and \((A.6)\).** It suffices to check

$$\max_{l=1,\ldots,L-1} \mathbb{E} \left[ \mathbb{I}_{[0,x]} \left( \frac{I_{T,11}(\ell_l)/f_{11}(\ell_l)}{I_{T,22}(\ell_l')/f_{22}(\ell_l')} \right) - \mathbb{I}_{[0,x]} \left( \frac{\gamma_{11}(\ell_l)^2}{\gamma_{22}(\ell_l')^2} \right) \right] \leq C \frac{1}{\sqrt{T}}.$$  \hspace{1cm} (A.8)

and

$$\max_{l=1,\ldots,L-1} \mathbb{E} \left[ \mathbb{I}_{[0,x]} \left( \frac{I_{T,11}(\ell_l)/f_{11}(\ell_l)}{I_{T,22}(\ell_l')/f_{22}(\ell_l')} \right) - \mathbb{I}_{[0,x]} \left( \frac{I_{T,11}(\ell_l)}{I_{T,22}(\ell_l')} \right) \right] \leq C \frac{1}{\sqrt{T}}.$$  \hspace{1cm} (A.9)

hold for all \(x > 0\).

For instance, if the inequality \((A.8)\) holds, by the Markov inequality and the triangular inequality, we have for all \(\epsilon > 0\),

$$\mathbb{P}\left( \frac{\sqrt{L - 1}\tilde{F}_L(x) - \sqrt{L - 1}\hat{F}_L(x)}{\sqrt{L \log T/\sqrt{T}}} > \epsilon \right) \leq \frac{\epsilon\sqrt{T}}{\sqrt{L \log T}} \frac{1}{\sqrt{L - 1}} \sum_{l=1}^{L-1} \mathbb{E} \left[ \mathbb{I}_{[0,x]} \left( \frac{I_{T,11}(\ell_l)/f_{11}(\ell_l)}{I_{T,22}(\ell_l')/f_{22}(\ell_l')} \right) - \mathbb{I}_{[0,x]} \left( \frac{\gamma_{11}(\ell_l)^2}{\gamma_{22}(\ell_l')^2} \right) \right] \leq C \frac{\sqrt{T}}{\epsilon \sqrt{L \log T}} \sqrt{L - 1} \frac{1}{\sqrt{T}} \rightarrow 0.$$
as \( T \) goes to infinity. This proves \((A.2)\).

Note that both inequalities of \((A.8)\) and \((A.9)\) hold, since they are the special cases of inequalities \((A.9)\) and \((A.10)\) of Zhang and Tu (2017) when \( a = 0 \) and \( b = x \), respectively. This proves \((A.8)\) and \((A.9)\).

**Proof of \((A.3)\).** By \((A.6)\) and \((A.8)\) of Zhang and Tu (2017), we have

\[
\vartheta_T^{(1)}(\ell_l) = o_p\left(\frac{\sqrt{\log T}}{\sqrt{T}}\right) \quad \text{and} \quad \vartheta_T^{(2)}(\ell_l) = o_p\left(\frac{\sqrt{\log T}}{\sqrt{T}}\right) \quad (A.10)
\]

hold for all \( l = 1, \cdots, L \). Note that

\[
\int_0^\infty \frac{\sqrt{L-1}(\widetilde{F}_L(x) - \bar{F}_L(x))}{F(x)(1 - F(x))} \, dF(x) = \sqrt{L-1} \int_0^\infty (\widetilde{F}_L(x) - \bar{F}_L(x)) x^{-1} \, dx
\]

\[
= \frac{1}{\sqrt{L-1}} \sum_{l=1}^{L-1} \int_0^\infty \frac{\log \left( \frac{I_{T,11}(\ell_l)/f_{11}(\ell_l)}{I_{T,22}(\ell_l)/f_{22}(\ell_l)} \right)}{x} \, dx - \int_0^\infty \frac{\log \left( \frac{[\eta_T^{(1)}(\ell_l)]^2/f_{11}(\ell_l)}{[\eta_T^{(2)}(\ell_l)]^2/f_{22}(\ell_l)} \right)}{x} \, dx \quad (A.11)
\]

Since \(|\log(1 + x)| = o(x)\) as \( x \) goes to zero, it follows from \((A.10)\) that

\[
\left| \int_0^\infty \frac{\log \left( \frac{I_{T,11}(\ell_l)/f_{11}(\ell_l)}{I_{T,22}(\ell_l)/f_{22}(\ell_l)} \right)}{x} \, dx - \int_0^\infty \frac{\log \left( \frac{[\eta_T^{(1)}(\ell_l)]^2/f_{11}(\ell_l)}{[\eta_T^{(2)}(\ell_l)]^2/f_{22}(\ell_l)} \right)}{x} \, dx \right|
\]

\[
= \left| \log \left( \frac{I_{T,11}(\ell_l)/f_{11}(\ell_l)}{I_{T,22}(\ell_l)/f_{22}(\ell_l)} \right) \right| - \left| \log \left( \frac{[\eta_T^{(1)}(\ell_l)]^2/f_{11}(\ell_l)}{[\eta_T^{(2)}(\ell_l)]^2/f_{22}(\ell_l)} \right) \right|
\]

\[
= \left| \log \left( \frac{1 + \vartheta_T^{(1)}(\ell_l)}{[\eta_T^{(1)}(\ell_l)]^2/f_{11}(\ell_l)} \right) \right| - \left| \log \left( \frac{1 + \vartheta_T^{(2)}(\ell_l)}{[\eta_T^{(2)}(\ell_l)]^2/f_{22}(\ell_l)} \right) \right|
\]

\[
\leq \left| \log \left( \frac{1 + \vartheta_T^{(1)}(\ell_l)}{[\eta_T^{(1)}(\ell_l)]^2/f_{11}(\ell_l)} \right) \right| + \left| \log \left( \frac{1 + \vartheta_T^{(2)}(\ell_l)}{[\eta_T^{(2)}(\ell_l)]^2/f_{22}(\ell_l)} \right) \right|
\]

\[
= o_p\left(\frac{\sqrt{\log T}}{\sqrt{T}}\right) \quad (A.12)
\]

holds for all \( x > 0 \) and \( l = 1, \cdots, L \). Combining \((A.11)\) and \((A.12)\), we obtain \((A.3)\).

**Proof of \((A.7)\).** By the Lipschitz continuity, we have \(|f_{22}(\ell_l) - f_{22}(\ell'_l)| \leq C \frac{1}{T} \leq C \frac{1}{\sqrt{T}} = o\left(\frac{\sqrt{\log T}}{\sqrt{T}}\right)\). Then, if the null hypothesis \((2.4)\) is true, we obtain

\[
\left| \log \left( \frac{I_{T,11}(\ell_l)/f_{11}(\ell_l)}{I_{T,22}(\ell_l)/f_{22}(\ell_l)} \right) - \log \left( \frac{I_{T,11}(\ell_l)}{I_{T,22}(\ell_l)} \right) \right| = \left| \log \left( \frac{f_{22}(\ell'_l) - f_{22}(\ell_l)}{f_{22}(\ell_l)} \right) \right|
\]

\[
= \left| \log \left( 1 + \frac{f_{22}(\ell'_l) - f_{22}(\ell_l)}{f_{22}(\ell_l)} \right) \right| = o\left(\frac{\sqrt{\log T}}{\sqrt{T}}\right)
\]

holds for all \( x > 0 \) and \( l = 1, \cdots, L \), since \( f_{22}(\omega) \) is lower bounded. With arguments similar to the proof of \((A.3)\), we obtain \((A.7)\).
Acknowledgements

This work was supported in part by the National Natural Science Foundation of China (grant number: 11671416).

References

Anderson, T. W. and Darling, D. A. (1952) Asymptotic theory of certain ‘goodness of fit’ criteria based on stochastic processes. *Ann. Math. Statist.* 23, 193–212.

Anderson, T. W. and Darling, D. A. (1954) A test of goodness of fit. *J. Amer. Statist. Assoc.* 49, 765–769.

Dahlhaus, R. (1997) Fitting time series models to nonstationary processes. *Ann. Statist.* 25, 1–37.

DasGupta, A. (2008) *Asymptotic Theory of Statistics and Probability*. New York: Springer.

Decowski, J. and Li, L. (2015) Wavelet-based tests for comparing two time series analysis. *J. Time Ser. Anal.* 36, 189–208.

Dette, H., Kinsvater, T. and Vetter, M. (2011a) Testing non-parametric hypotheses for stationary processes by estimating minimal distances. *J. Time Ser. Anal.* 32, 447–461.

Dette, H., Preuss, P. and Vetter, M. (2011b) A measure of stationarity in locally stationary processes with applications to testing. *J. Amer. Statist. Assoc.* 106, 1113–1124.

Diggle, P. J. and Fisher, N. I. (1991) Nonparametric comparison of cumulative periodograms. *J. R. Stat. Soc. Ser. B Stat. Methodol.* 71, 831–857.

Faraway, J., Marsaglia, G., Marsaglia, J. and Baddeley, A. (2015) gof: Classical goodness-of-fit tests for univariate distributions. URL https://CRAN.R-project.org/package=gof.

Fuller, W. A. (1996) *Introduction to Statistical Time Series*, 2nd edn. New York: John Wiley.

Guo, W. and Dai, M. (2006) Multivariate time-dependent spectral analysis using Cholesky decomposition. *Statist. Sinica* 16, 825–845.

Jentsch, C. and Pauly, M. (2015) Testing equality of spectral densities using randomization techniques. *Bernoulli* 21, 697–739.

Marsaglia, G. and Marsaglia, J. (2004) Evaluating the Anderson-Darling distribution. *J. Stat. Softw.* 9(2), 1–5.

Shumway, R. H. and Stoffer, D. S. (2011) *Time Series Analysis and its Applications with R Examples*, 3rd edn. New York: Springer.
van Bellegem, S. and von Sachs, R. (2008) Locally adaptive estimation of evolutionary wavelet spectra. *Ann. Statist.* 36, 1879–1924.

van der Vaart, A. W. (1998) *Asymptotic Statistics*. Cambridge: Cambridge University Press.

Zhang, S. and Tu, X. M. (2017) Comparing time-invariant and time-varying spectra based on the Pearson statistic. Submitted to *J. Time Ser. Anal.* for publication.