View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Research Papers in Economics

ELSEVIER
Available online at www.sciencedirect.com
ScienceDirect
Journal of Multivariate Analysis 99 (2008) 968–1009
Journal of Multivariate Analysis
www.elsevier.com/locate/jmva

Testing nonparametric and semiparametric hypotheses in vector stationary processes

Michael Eichler*

Department of Quantitative Economics, University of Maastricht, P.O. Box 616, 6200 MD Maastricht, The Netherlands

Received 21 August 2006
Available online 16 June 2007

Abstract

We propose a general nonparametric approach for testing hypotheses about the spectral density matrix of multivariate stationary time series based on estimating the integrated deviation from the null hypothesis. This approach covers many important examples from interrelation analysis such as tests for noncorrelation or partial noncorrelation. Based on a central limit theorem for integrated quadratic functionals of the spectral matrix, we derive asymptotic normality of a suitably standardized version of the test statistic under the null hypothesis and under fixed as well as under sequences of local alternatives. The results are extended to cover also parametric and semiparametric hypotheses about spectral density matrices, which includes as examples goodness-of-fit tests and tests for separability.
© 2007 Elsevier Inc. All rights reserved.

AMS 1991 subject classification: primary 62M15; secondary 62G10

Keywords: Frequency domain approach; Nonparametric and semiparametric tests; Spectral density matrix; Partial noncorrelation; Goodness-of-fit test

1. Introduction

Many important hypotheses about the second-order properties of a multivariate stationary time series can be expressed in terms of the spectral density matrix. For instance, two vector processes are uncorrelated if all cross-spectral densities or, equivalently, all spectral coherences between the two processes vanish at all frequencies. More general hypotheses about the interrelation structure in multivariate time series can be formulated in terms of partial coherences; such hypotheses arise, for example, in graphical interaction modelling based on partial correlation graphs [14,13].

* Fax: +31 43 388 4874.
E-mail address: m.eichler@ke.unimaas.nl.

0047-259X/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmva.2007.06.003

The formulation of hypotheses in the frequency domain enjoys the advantage of a general nonparametric framework based on, for example, kernel spectral estimators or the integrated periodogram. The use of the integrated periodogram for testing the goodness-of-fit of a time series model has been described in many papers (e.g., [3,1,7,34]). The type of hypotheses that can be tested with the integrated periodogram, however, is limited, and a more flexible approach is offered by considering estimates for the spectral density. Taniguchi and Kondo [61], Taniguchi et al. [62], and Taniguchi and Kakizawa [60] considered test problems of the form

$$\mathrm{H}_0: \int_\Pi K\big(f(\lambda)\big)\, d\lambda = c \quad \text{against } \mathrm{H}_a: \int_\Pi K\big(f(\lambda)\big)\, d\lambda \neq c,$$

where $f(\lambda)$ is the spectral density matrix, $K(\cdot)$ is an appropriate function, $\Pi = [-\pi, \pi]$, and c is some constant. Although this setting covers a variety of test problems, their results require a nonvanishing first derivative of the function $K(\cdot)$ under the null hypothesis and thus cannot be applied to many interesting hypotheses.

As a motivating example, we consider the problem of testing for partial non-correlation among the components of a multivariate time series. This problem arises, for instance, in the context of so-called partial correlation graphs [13], which generalize the concept of covariance selection models and concentration graphs [16,8] to the multivariate time series case and have become a popular approach for describing the interactions among the components of a multivariate stationary process (e.g., [63,30,26,29,58]).[1] More precisely, let $\{X_V(t)\}$ be a multivariate time series with components $\{X_v(t)\}$, $v \in V$. Then the partial correlation graph of $\{X_V(t)\}$ is defined as the graph G with vertices $v \in V$ and edges a–b for distinct $a, b \in V$ such that a–b is absent in the graph if and only if the corresponding series $\{X_a(t)\}$ and $\{X_b(t)\}$ are uncorrelated after removing the linear effects of the remaining components $\{X_{V\setminus\{a,b\}}(t)\}$. In the frequency domain, this is equivalent to that the partial cross spectrum:

$$f_{ab|V_{ab}}(\lambda) = f_{ab}(\lambda) - f_{aV_{ab}}(\lambda) f_{V_{ab}V_{ab}}(\lambda)^{-1} f_{V_{ab}b}(\lambda)$$

or, after rescaling, the partial spectral coherency

$$R_{ab|V_{ab}}(\lambda) = \frac{f_{ab|V_{ab}}(\lambda)}{\sqrt{f_{aa|V_{ab}}(\lambda) f_{bb|V_{ab}}(\lambda)}}$$

are zero for all frequencies $\lambda \in \Pi$, where we have set $V_{ab} = V\setminus\{a, b\}$ for ease of notation. One straightforward approach for testing for the absence of an edge a–b in the partial correlation graph G is to use the integrated sample partial spectral coherence

$$\int_\Pi |\hat{R}_{ab|V_{ab}}(\lambda)|^2\, d\lambda = \int_\Pi \hat{R}_{ab|V_{ab}}(\lambda)\, \hat{R}_{ba|V_{ab}}(\lambda)\, d\lambda$$

as a test statistic. However, despite its intuitiveness, such a test has not yet been considered in the literature. In particular, we note that although the test problem resembles the type of problems discussed by Taniguchi et al. [62] with $K\big(f(\lambda)\big) = |R_{ab|V_{ab}}(\lambda)|^2$ and $c = 0$, their results are not applicable since the first derivative of K vanishes under the null hypothesis and thus leads

[1] There are other approaches for graphical modelling of multivariate time series, which are more suitable for describing dynamic dependences and causal relationships, see, e.g., Reale and Tunnicliffe Wilson [54], Moneta and Spirtes [48], Eichler [22,25,23,21].

to a zero variance of the asymptotic distribution. An alternative approach for testing for partial noncorrelation has been considered by Dahlhaus et al. [14] based on the maximum partial spectral coherence; for this statistic, however, the exact asymptotic distribution is unknown and only a heuristic approximation has been proposed and used.

In this paper, we consider more generally test problems of the form

$$\mathrm{H}_0 : \psi\big(f(\lambda), \lambda\big) \equiv 0 \quad \text{against } \mathrm{H}_a : \psi\big(f(\lambda), \lambda\big) \not\equiv 0,$$

where $\psi(Z, \lambda)$ is some suitable vector-valued function. Measuring the deviation from the null hypothesis H_0 at frequency λ by the squared Euclidean norm $\big\|\psi\big(f(\lambda), \lambda\big)\big\|^2$, the above test problem equivalently—under some regularity conditions on $f(\lambda)$ and $\psi(Z, \lambda)$—can be expressed as

$$\mathrm{H}_0 : \int_\Pi \big\|\psi\big(f(\lambda), \lambda\big)\big\|^2 \, d\lambda = 0 \quad \text{against } \mathrm{H}_a : \int_\Pi \big\|\psi\big(f(\lambda), \lambda\big)\big\|^2 \, d\lambda \neq 0.$$

For estimation of such nonlinear functionals, we replace the unknown spectral density matrix by a nonparametric spectral estimator $\hat{f}(\lambda)$. We discuss the asymptotic properties of the resulting test statistics under the null hypothesis as well as for fixed alternatives and sequences of local alternatives. Since for nonlinear functionals the quality of estimation depends crucially on the bias of the spectral estimator, we allow the use of data tapers to improve the small sample properties.

As an important generalization, we show that the results can be extended to parametric and semiparametric hypotheses about the spectral density. Assuming that the unknown parameters can be estimated $\sqrt{T}$-consistently, it can be shown that estimation of the parameter is asymptotically negligible and does not affect the asymptotic behaviour of the test statistics. As examples, we discuss goodness-of-fit tests and a test for separability of the covariance structure of a time series.

The paper is organized as follows. Section 2 contains some basic definitions. In Section 3, we derive asymptotic normality for a suitably standardized version of the test statistic. In particular, we discuss hypotheses about partial spectral densities and show their relation to similar hypotheses about spectral densities. In Section 4, these results are extended to parametric and semiparametric hypotheses about the spectral density matrix and general discrepancies to measure the deviation from the null hypothesis. For the investigation of the power of the proposed tests, we derive in Section 5 the limiting distribution of the test statistic under sequences of nonparametric local alternatives. In Section 6, we present the results of a Monte Carlo study, and Section 7 concludes. In the Appendix, we provide a central limit theorem for integrated weighted squared errors, which is central to all asymptotic results in this paper.

2. The test statistic

Let $\{X(t), t \in \mathbb{Z}\}$, $X(t) = \big(X_1(t), \ldots, X_d(t)\big)'$ be a stationary multivariate time series with mean zero and continuous spectral density matrix $f(\lambda) = \big(f_{ab}(\lambda)\big)_{a,b=1,\ldots,d}$. The test problems studied in this paper are of the form

$$\mathrm{H}_0 : \psi\big(f(\lambda), \lambda\big) \equiv 0 \quad \text{against } \mathrm{H}_a : \psi\big(f(\lambda), \lambda\big) \not\equiv 0,$$

where $\psi(Z, \lambda) = \big(\psi_1(Z, \lambda), \ldots, \psi_r(Z, \lambda)\big)'$ is some vector-valued piecewise continuous function. The deviation from the null hypothesis at frequency λ can be measured by $\|\psi\big(f(\lambda), \lambda\big)\|^2$, where $\|\cdot\|$ denotes the Euclidean norm. Since the spectral matrix $f(\lambda)$ is assumed to be continuous, the above null hypothesis is equivalent to

$$\int_{\Pi} \left\| \psi\big(f(\lambda), \lambda\big) \right\|^2 d\lambda = 0.$$

Substituting a nonparametric estimator $\hat{f}(\lambda)$ for the spectral density matrix $f(\lambda)$, we obtain the statistic

$$S_T(\psi) = \int_{\Pi} \left\| \psi\big(\hat{f}(\lambda), \lambda\big) \right\|^2 d\lambda,$$

which then can be used to test nonparametrically for H_0.

The nonparametric estimation of the spectral density matrix $f(\lambda)$ is usually based on the periodogram matrix $I^{(T)}(\lambda)$. In order to reduce the bias of the spectral estimator, we use a tapered version of the periodogram given by

$$I^{(T)}(\lambda) = \big(2\pi H_2^{(T)}(0)\big)^{-1} d^{(T)}(\lambda)\, d^{(T)}(-\lambda)',$$

where

$$d^{(T)}(\lambda) = \sum_{t=1}^{T} h^{(T)}(t)\, X(t) \exp(-\mathrm{i}\lambda t)$$

is the finite Fourier transform of the time series with real-valued data taper $h^{(T)}(t)$ and

$$H_k^{(T)}(\lambda) = \sum_{t=1}^{T} h^{(T)}(t)^k \exp(-\mathrm{i}\lambda t)$$

for $k \in \mathbb{N}$ are the Fourier transforms of the data taper. We assume that the data taper is given by $h^{(T)}(t) = h(t/T)$, where $h : \mathbb{R} \to \mathbb{R}$ is a bounded function of bounded variation that vanishes outside the interval [0, 1]. Although the choice of taper function h does not affect the asymptotic results presented in this paper, it should be smooth with $h(0) = h(1) = 0$ in order to improve the small sample properties of the estimates (e.g., [12]). The effect of tapering when testing for noncorrelation of two univariate time series has been studied in Eichler [24].

Consistent estimators for the spectral density matrix $f(\lambda)$ can be obtained by smoothing of the periodogram matrix, which leads to kernel estimates of the form

$$\hat{f}(\lambda) = \int_{\Pi} w^{(T)}(\lambda - \alpha)\, I^{(T)}(\alpha)\, d\alpha, \tag{1}$$

where the smoothing kernel is given by the 2π-periodic extension of $w^{(T)}(\lambda) = w(\lambda/B_T)/B_T$, $\lambda \in \Pi$, for some real-valued kernel function w and bandwidth B_T. In the sequel, we will implicitly consider 2π-periodic extensions of all functions defined on Π.

3. Asymptotic properties

In this section, we derive the limiting distribution of the test statistic $S_T(\psi)$ under the null hypothesis H_0. Following the approach by Brillinger [5], we impose the following mixing conditions on the process (cf. [5, Assumption 2.6.2]).

Assumption 3.1. $\{X(t)\}$ is a zero mean d vector-valued stationary stochastic process defined on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$. Furthermore, for any $k > 0$, the kth order cumulants of $\{X(t)\}$ satisfy the mixing conditions

$$\sum_{u_1,\dots,u_{k-1}\in\mathbb{Z}} \big(1 + |u_j|^2\big)|c_{a_1,\dots,a_k}(u_1, \dots, u_{k-1})| < \infty$$

for all $j = 1, \dots, k-1$ and $a_1, \dots, a_k = 1, \dots, d$, where $c_{a_1,\dots,a_k}(u_1, \dots, u_{k-1})$ is the joint cumulant of $X_{a_1}(u_1), \dots, X_{a_{k-1}}(u_{k-1}), X_{a_k}(0)$.

The assumption, which requires the existence of the moments of all orders, is satisfied, for instance, for ARMA processes of finite order provided that the innovation process has moments of all orders. Unlike alternative approaches for deriving asymptotic distributions for frequency domain statistics (e.g., [41,60,15]), which usually get by with moments of fourth or eighth order, our approach is not restricted to linear processes. For later use, we also note that the condition on the second-order cumulants implies that the spectral density matrix exists and its entries have bounded and uniformly continuous second derivatives.

In the next assumption and throughout the paper, we make extensive use of matrix notation. As usual, for matrices A and B, $\mathrm{vec}(A)$ denotes the vector resulting from stacking the columns of the matrix A on top of each other and $A \otimes B$ denotes the Kronecker product of A and B. Furthermore, we write A^* for the conjugate transpose of a complex-valued matrix A. Finally, $\|A\| = \mathrm{tr}(A^*A)^{1/2}$ denotes the Euclidean norm of A.

Assumption 3.2. Let $\psi : D \times \Pi \to \mathbb{C}^r$, where D is an open subset of $\mathbb{C}^{d\times d}$ that contains the spectral density matrices $f(\lambda)$, $\lambda \in \Pi$, of the process $\{X(t)\}$.

(i) $\psi(Z, \lambda)$ is holomorphic with respect to Z.
(ii) $\psi(Z, \lambda)$ and its first derivative with respect to $z = \mathrm{vec}(Z)$,

$$\mathsf{D}_z\psi\big(Z, \lambda\big) = \frac{\partial\psi(Z, \lambda)}{\partial z'}$$

are piecewise Lipschitz continuous in $\lambda \in \Pi$.
(iii) There exists a positive constant ρ such that for all $\lambda \in \Pi$ the ball $B_{\rho,\lambda} = \{Z \in \mathbb{C}^{d\times d} \mid \|f(\lambda) - Z\| \leqslant \rho\}$ is contained in D and

$$\sup_{\lambda\in\Pi}\ \sup_{Z\in B_{\rho,\lambda}} \big\|\psi(Z, \lambda)\big\| < \infty.$$

(iv) $\int_\Pi \big\|\mathsf{D}_z\psi\big(f(\lambda), \lambda\big)\big\|\, d\lambda > 0$.

The condition ensures that the function ψ is sufficiently smooth in both its arguments and is defined in the neighbourhood of every value that the spectral density matrix $f(\lambda)$ of the process $\{X(t)\}$ takes. Since usually the domain D comprises the set of all positive-definite Hermitian matrices, a sufficient condition on $f(\lambda)$ is that the eigenvalues of $f(\lambda)$ are bounded and bounded away from zero uniformly for all $\lambda \in \Pi$. We note that in special cases D may also include nonnegative Hermitian matrices. Condition (iv) guarantees that the derived limit distribution of the test statistic $S_T(\psi)$ will have nonzero variance.

Assumption 3.3. Let $\hat{f}(\lambda)$ be the kernel spectral estimator in (1).

(i) The taper function $h : \mathbb{R} \to \mathbb{R}$ is bounded, is of bounded variation, and vanishes outside the interval [0, 1].

(ii) The kernel function $w(\lambda)$ is bounded, symmetric, nonnegative, and Lipschitz continuous with

$$\int_{\mathbb{R}} w(\lambda)\, d\lambda = 1, \qquad \int_{\mathbb{R}} \lambda^2\, w(\lambda)\, d\lambda < \infty, \qquad \text{and} \qquad \limsup_{\lambda\to\infty} \lambda^2\, w(\lambda) = 0.$$

Furthermore, $w(\lambda)$ has continuous Fourier transform $W(u)$ such that

$$C_{w,2} = \int_{\mathbb{R}} W(u)^2\, du < \infty \quad \text{and} \quad C_{w,4} = \int_{\mathbb{R}} W(u)^4\, du < \infty.$$

(iii) Let $(B_T)_{T\in\mathbb{N}}$ be a sequence of kernel bandwidths such that $B_T^{9/2}\, T \to 0$ and $B_T^2\, T \to \infty$ as $T \to \infty$.

The assumptions on the taper function are standard (e.g., [5, Assumption 4.3.1]) and include in particular the nontapered case $h(x) = 1_{(0,1]}(x)$. In contrast, the conditions on the kernel function are more restrictive and exclude some commonly used kernels such as the Daniell or Bartlett window while allowing, for example, the quadratic spectral window or the Parzen window (e.g., [53]). Similar assumptions can be found in Taniguchi and Kakizawa [60, Chapter 6.1]; our additional condition $C_{w,4} < \infty$ is due to the quadratic approximations required for the statistics discussed in this paper. We also note that the simulations in Eichler [24] indicate that violation of the continuity assumption on the kernel function indeed may lead to a serious bias of our test.

Finally, we note that the conditions on the rate by which the bandwidth B_T tends to zero are rather strict compared with those imposed by other authors (e.g., [40,51]). The conditions are a consequence of the generality of our approach, which covers a large variety of nonlinear hypotheses on the spectral matrix. The condition $B_T^{9/2} T \to 0$, for instance, ensures that the bias of the kernel spectral estimator $\hat{f}(\lambda)$ vanishes fast enough in order to not affect the asymptotics based on Taylor expansions about the true spectral density $f(\lambda)$. Our assumptions are slightly weaker than those of Taniguchi and Kondo [61] and Taniguchi et al. [62] (see also [60, Chapter 6.1], who required $B_T = O(T^{-\alpha})$ with $\frac{1}{4} < \alpha < \frac{1}{2}$.

For the formulation of the asymptotic results, we define the matrix-valued function $\Gamma_\psi : \Pi \to \mathbb{C}^{d^2\times d^2}$ with

$$\Gamma_\psi(\lambda) = \mathsf{D}_z\psi\big(f(\lambda), \lambda\big)^* \mathsf{D}_z\psi\big(f(\lambda), \lambda\big) \tag{2}$$

for all $\lambda \in \Pi$, where $\psi(Z, \lambda)$ is a function satisfying Assumption 3.2. We note that for $I = i + (j-1)\, d$ and $K = k + (l-1)\, d$ with $i, j, k, l \in \{1, \ldots, d\}$ the (I, K)th element of $\Gamma_\psi(\lambda)$ is given by

$$\Gamma_{ij,kl;\psi}(\lambda) = \left(\frac{\partial\psi(Z,\lambda)}{\partial Z_{ij}}\right)^* \left(\frac{\partial\psi(Z,\lambda)}{\partial Z_{kl}}\right)\Bigg|_{Z=f(\lambda)}.$$

Furthermore, we set $\tilde{\Gamma}_\psi(\lambda) = K_{dd}\, \Gamma_\psi(\lambda)\, K_{dd}$, where K_{dd} is the $d^2 \times d^2$ commutation matrix satisfying $\mathrm{vec}(A') = K_{dd}\, \mathrm{vec}(A)$ for any $d \times d$ matrix A (e.g., [46]). Thus the (I, K)th entry of the matrix $\tilde{\Gamma}_{\psi}(\lambda)$ is given by

$$\tilde{\Gamma}_{ij,kl;\psi}(\lambda) = \Gamma_{ji,lk;\psi}(\lambda) = \left(\frac{\partial\psi(Z,\lambda)}{\partial Z_{ji}}\right)^{*}\left(\frac{\partial\psi(Z,\lambda)}{\partial Z_{lk}}\right)\Bigg|_{Z=f(\lambda)} \tag{3}$$

for $i, j, k, l = 1, \ldots, d$.

The derivation of the limiting distribution of $S_T(\psi)$ is based on the following result, which shows that $S_T(\psi)$ can be approximated by the integrated weighted squared error with weight function $\Gamma_{\psi}(\lambda)$.

Lemma 3.4. *Suppose that Assumptions* 3.1–3.3 *hold. Then under the null hypothesis* H_0

$$S_T(\psi) = \int_{\Pi} \|\mathrm{vec}\big(\hat{f}(\lambda) - f(\lambda)\big)\|^2_{\Gamma_{\psi}(\lambda)}\, d\lambda + o_P\big((B_T^{1/2}\, T)^{-1}\big), \tag{4}$$

where $\|x\|^2_{\Gamma_{\psi}(\lambda)} = x^{*}\Gamma_{\psi}(\lambda)x$.

Proof. Since the function $\psi(Z,\lambda)$ is holomorphic in D, it has a Taylor expansion about $Z_0 = f(\lambda)$ in an open neighbourhood $U \subseteq D$. Let

$$\hat{\psi}(\lambda) = \psi\big(\hat{f}(\lambda), \lambda\big) - \mathsf{D}_z\psi\big(f(\lambda), \lambda\big)\,\mathrm{vec}\big(\hat{f}(\lambda) - f(\lambda)\big).$$

Using Cauchy's estimate for the derivatives of ψ (e.g., [60, Lemma A.1.3]) and Assumption 3.2(ii), we find constants C_δ and δ such that

$$\big\|\hat{\psi}(\lambda)\big\| \leqslant C_\delta\, \big\|\hat{f}(\lambda) - f(\lambda)\big\|^2$$

uniformly in λ, whenever $\max_{\lambda\in\Pi} \|\hat{f}(\lambda) - f(\lambda)\| \leqslant \delta$. It follows that

$$\left| S_T(\psi) - \int_{\Pi} \|\mathrm{vec}\big(\hat{f}(\lambda) - f(\lambda)\big)\|^2_{\Gamma_{\psi}(\lambda)}\, d\lambda \right| \leqslant C \int_{\Pi} \big\|\hat{f}(\lambda) - f(\lambda)\big\|^3\, d\lambda,$$

whenever $\max_{\lambda\in\Pi} \|\hat{f}(\lambda) - f(\lambda)\| \leqslant \delta$. Under the assumptions, it follows by the convergence of the cumulants of $\sqrt{B_T\, T}\big(\hat{f}(\lambda) - f(\lambda)\big)$ [5, Theorems 7.4.1–7.4.4] and the product theorem for cumulants (e.g., [5, Theorem 2.3.2]) that, for $k \in \mathbb{N}$,

$$\mathbb{E}\left[\prod_{r=1}^{k} \big(\hat{f}_{i_r j_r}(\lambda_r) - f_{i_r j_r}(\lambda_r)\big)\right] = O\big((B_T\, T)^{-k/2}\big)$$

uniformly in $\lambda_1, \ldots, \lambda_k$ and hence that

$$\mathbb{E}\|\hat{f}(\lambda) - f(\lambda)\|^{2k} = O\big((B_T\, T)^{-k}\big)$$

uniformly in λ. Using Cauchy–Schwarz inequality and Fubini's theorem, we finally obtain

$$\int_{\Pi} \|\hat{f}(\lambda) - f(\lambda)\|^k\, d\lambda = O_P\big((B_T\, T)^{-k/2}\big) \tag{5}$$

for $k \in \mathbb{N}$. Since furthermore $\max_{\lambda\in\Pi} \|\hat{f}(\lambda) - f(\lambda)\| = o_P(1)$ (e.g., [60, Eq. (6.1.17)]), there exists for every $\varepsilon > 0$ a constant η_ε such that

$$\mathbb{P}\Big(\Big|S_T(\psi) - \int_\Pi \big\|\mathrm{vec}\big(\hat{f}(\lambda) - f(\lambda)\big)\big\|^2_{\Gamma_\psi(\lambda)}\, d\lambda\Big| \geqslant \eta_\varepsilon\,(B_T\,T)^{-3/2}\Big)$$
$$\leqslant \mathbb{P}\Big(C\int_\Pi \|\hat{f}(\lambda) - f(\lambda)\|^3 d\lambda \geqslant \eta_\varepsilon\,(B_T\,T)^{-3/2}\Big)$$
$$+\mathbb{P}\left(\max_{\lambda\in\Pi} \|\hat{f}(\lambda) - f(\lambda)\| > \delta\right) < \varepsilon.$$

Since $B_T^{-1}\,T^{-1/2} = o(1)$ this proves the lemma. □

In the Appendix, it is shown that the integral on the right-hand side in (4) is asymptotically normally distributed with rate of convergence $B_T^{-1/2}\,T$. Thus we obtain the following central limit theorem for $S_T(\psi)$.

Theorem 3.5. *Suppose that Assumptions* 3.1–3.3 *hold. Then under the null hypothesis* H_0, *we have*

$$B_T^{1/2}\,T\,S_T(\psi) - B_T^{-1/2}\,\mu(\psi) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma^2(\psi)),$$

where

$$\mu(\psi) = C_h\,C_{w,2}\int_\Pi \mathrm{tr}\big[\Gamma_\psi(\lambda)\,\big(f(\lambda)' \otimes f(\lambda)\big)\big]\,d\lambda \tag{6}$$

and

$$\sigma^2(\psi) = 2\pi\,C_h^2\,C_{w,4}\int_\Pi \mathrm{tr}\Big[\Gamma_\psi(\lambda)\big(f(\lambda)' \otimes f(\lambda)\big)\Big\{\Gamma_\psi(\lambda) + \tilde{\Gamma}_\psi(-\lambda)$$
$$+\Gamma_\psi(-\lambda)' + \tilde{\Gamma}_\psi(\lambda)'\Big\}\big(f(\lambda)' \otimes f(\lambda)\big)\Big]\,d\lambda, \tag{7}$$

where $C_h = H_4/H_2^2$ *with* $H_k = \int_0^1 h(t)^k\,dt$, *and* $\Gamma_\psi(\lambda)$ *and* $\tilde{\Gamma}_\psi(\lambda)$ *are given by* (2) *and* (3), *respectively*.

Proof. The result follows from Lemma 3.4 and Theorem B.2. □

We note that the function $\Gamma_\psi(\lambda)$ can also be written as

$$\Gamma_\psi(\lambda) = \sum_{j=1}^r \mathrm{vec}\left(\overline{\frac{\partial\psi_j(Z,\lambda)}{\partial Z}}\right)\mathrm{vec}\left(\frac{\partial\psi_j(Z,\lambda)}{\partial Z}\right)'\Bigg|_{Z=f(\lambda)}.$$

This leads to alternative expressions for the bias $\mu(\psi)$ and variance $\sigma^2(\psi)$, which we give only for scalar-valued $\psi(Z,\lambda)$ as in this case the expressions become particularly simple. For notational simplicity, we write $\frac{\partial\psi}{\partial Z}\big(Z_0,\lambda\big)$ for $\frac{\partial\psi(Z,\lambda)}{\partial Z}\big|_{Z=Z_0}$, where $\frac{\partial\psi(Z,\lambda)}{\partial Z} = \Big(\frac{\partial\psi(Z,\lambda)}{\partial Z_{ij}}\Big)_{i,j=1,\ldots,d}$.

Corollary 3.6. *Suppose that the assumptions of Theorem* 3.5 *hold. Furthermore let* $\psi : D \times \Pi \to \mathbb{C}$. *Then the bias* $\mu(\psi)$ *and the variance* $\sigma^2(\psi)$ *in Theorem* 3.5 *are given by*

$$\mu(\psi) = C_h\, C_{w,2} \int_\Pi \operatorname{tr}\left[\frac{\partial \psi}{\partial Z'}\big(f(\lambda), \lambda\big)\, f(\lambda)\, \overline{\frac{\partial \psi}{\partial Z}\big(f(\lambda), \lambda\big)}\, f(\lambda) \right] d\lambda$$

and

$$\begin{aligned}\sigma^2(\psi) = 2\pi\, C_h^2\, C_{w,4} \int_\Pi \Bigg[&\left| \operatorname{tr}\left\{ \frac{\partial \psi}{\partial Z'}\big(f(\lambda), \lambda\big)\, f(\lambda)\, \overline{\frac{\partial \psi}{\partial Z}\big(f(\lambda), \lambda\big)}\, f(\lambda) \right\} \right|^2 \\ &+ \left| \operatorname{tr}\left\{ \frac{\partial \psi}{\partial Z'}\big(f(\lambda), \lambda\big)\, f(\lambda)\, \overline{\frac{\partial \psi}{\partial Z'}\big(f(-\lambda), -\lambda\big)}\, f(\lambda) \right\} \right|^2 \\ &+ \left| \operatorname{tr}\left\{ \frac{\partial \psi}{\partial Z'}\big(f(\lambda), \lambda\big)\, f(\lambda)\, \frac{\partial \psi}{\partial Z}\big(f(-\lambda), -\lambda\big)\, f(\lambda) \right\} \right|^2 \\ &+ \left| \operatorname{tr}\left\{ \frac{\partial \psi}{\partial Z'}\big(f(\lambda), \lambda\big)\, f(\lambda)\, \frac{\partial \psi}{\partial Z'}\big(f(\lambda), \lambda\big)\, f(\lambda) \right\} \right|^2 \Bigg] d\lambda. \end{aligned} \tag{8}$$

Proof. The result follows directly from Corollary B.8. □

Remark 3.7. The norm $\|\psi\big(f(\lambda), \lambda\big)\|^2$ measures the deviation of the spectral matrix $f(\lambda)$ from the null hypothesis at frequency λ. Since $f(-\lambda) = f(\lambda)'$, every hypothesis on $f(\lambda)$ corresponds to an equivalent hypothesis on $f(-\lambda)$. Consequently, the function $\psi(Z, \lambda)$ often can be chosen such that

$$\|\psi(Z', -\lambda)\| = \|\psi(Z, \lambda)\|.$$

By a Taylor expansion of $\psi(Z, \lambda)$ about Z_0, this leads to the condition

$$\|\operatorname{vec}(Z - Z_0)\|^2_{\tilde{\Gamma}_\psi(-\lambda)} = \|\operatorname{vec}(Z - Z_0)\|^2_{\Gamma_\psi(\lambda)},$$

which implies that $\tilde{\Gamma}_\psi(\lambda) = \Gamma_\psi(-\lambda)$. It follows that for such functions $\psi(Z, \lambda)$ the expression for the asymptotic variance $\sigma^2(\psi)$ in Theorem 3.5 can be simplified to

$$4\pi\, C_h^2\, C_{w,4} \int_\Pi \operatorname{tr}\big\{ \Gamma_\psi(\lambda) \big(f(\lambda)' \otimes f(\lambda)\big) \big(\Gamma_\psi(\lambda) + \tilde{\Gamma}_\psi(\lambda)'\big) \big(f(\lambda)' \otimes f(\lambda)\big) \big\}\, d\lambda.$$

The expression for the variance in Corollary 3.6 adapts accordingly.

We note that the asymptotic distribution of the statistic $S_T(\psi)$ depends only on the second-order spectrum. Thus the bias and the variance can be estimated easily by substituting the consistent estimate $\hat{f}(\lambda)$ for $f(\lambda)$. Then Slutsky's theorem yields that

$$Q_T = \frac{T\, S_T(\psi) - B_T^{-1}\, \hat{\mu}(\psi)}{B_T^{-1/2}\, \hat{\sigma}(\psi)}$$

is asymptotically standard normally distributed.

Although the general expressions for the bias $\mu(\psi)$ and the variance $\sigma^2(\psi)$ look complicated, they often lead to simple expressions for some suitable choice of $\psi(Z, \lambda)$. We give two examples.

Example 3.8 (*Noncorrelation*). The problem of testing whether two multivariate time series $\{X_A(t)\}$ and $\{X_B(t)\}$ are independent has been considered recently by El Himdi and Roy [27], Hallin and Saidi [35], Bouhaddioui and Roy [4], and Saidi [57] in the context of multivariate ARMA processes. Using the general framework presented in this paper, we obtain a nonparametric frequency domain based test for noncorrelation between two multivariate stationary time series. For univariate time series, this test has been considered in Eichler [24].

Two time series $\{X_A(t)\}$ and $\{X_B(t)\}$ are uncorrelated if the spectral coherence $R_{AB}(\lambda) = \big(R_{ab}(\lambda)\big)_{a\in A, b\in B}$ with components

$$R_{ab}(\lambda) = \frac{f_{ab}(\lambda)}{\sqrt{f_{aa}(\lambda)\, f_{bb}(\lambda)}}$$

vanishes at all frequencies $\lambda \in \Pi$. Thus, if $f(\lambda)$ denotes the joint spectral density matrix of the two vector processes, the null hypothesis of noncorrelation can be formulated in the form $\mathrm{H}_0 :\ \psi\big(f(\lambda)\big) \equiv 0$ by setting

$$\psi\big(f(\lambda)\big) = \operatorname{vec}\big(R_{AB}(\lambda)\big).$$

Simple calculations show for $\Gamma_\psi(\lambda)$

$$\Gamma_{ij,kl;\psi}(\lambda) = \begin{cases} \dfrac{\delta_{ik}\delta_{jl}}{f_{ii}(\lambda)\, f_{jj}(\lambda)} & \text{if } i \in A,\ j \in B, \\ 0 & \text{otherwise} \end{cases}$$

and thus

$$\operatorname{tr}\big[\Gamma_\psi(\lambda)\big(f(\lambda)' \otimes f(\lambda)\big)\big] = \sum_{i,j,k,l=1}^{d} \Gamma_{ij,kl;\psi}\, f_{ki}(\lambda)\, f_{jl}(\lambda) = n_A\, n_B,$$

where n_A and n_B are the number of indices in A and B, respectively. Similarly, recalling that $\|R_{AA}(\lambda)\|^2 = \operatorname{tr} R_{AA}(\lambda) R_{AA}(\lambda)^*$, we obtain

$$\operatorname{tr}\big[\Gamma_\psi(\lambda)\big(f(\lambda)' \otimes f(\lambda)\big)\Gamma_\psi(\lambda)\big(f(\lambda)' \otimes f(\lambda)\big)\big] = \|R_{AA}(\lambda)\|^2\, \|R_{BB}(\lambda)\|^2,$$

while

$$\operatorname{tr}\big[\Gamma_\psi(\lambda)\big(f(\lambda)' \otimes f(\lambda)\big)\tilde{\Gamma}_\psi(\lambda)'\big(f(\lambda)' \otimes f(\lambda)\big)\big] = \|R_{AB}(\lambda)\|^2\, \|R_{BA}(\lambda)\|^2 = 0$$

under the null hypothesis of noncorrelation. Since $\Gamma_\psi(-\lambda)' = \Gamma_\psi(\lambda)$, the remaining two terms in the expression for $\sigma^2(\psi)$ lead to similar results and we thus we have

$$\mu(\psi) = 2\pi\, C_h\, C_{w,2}\, n_A\, n_B$$

and

$$\sigma^2(\psi) = 4\pi\, C_h^2\, C_{w,4} \int_\Pi \|R_{AA}(\lambda)\|^2\, \|R_{BB}(\lambda)\|^2\, d\lambda.$$

In the special case of two univariate time series $\{X_a(t)\}$ and $\{X_b(t)\}$, we have $|R_{aa}(\lambda)| = |R_{bb}(\lambda)| = 1$ and the asymptotic variance $\sigma^2(\psi)$ also becomes independent of the spectral matrix $f(\lambda)$ (cf. [24]). In this case, our test has a similar form as the test by Hong [40], which is time-domain based and requires prewhitening of the two series.

In some applications, one is particularly interested whether two vector time series $\{X_A(t)\}$ and $\{X_B(t)\}$ exhibit a relationship over a specific frequency interval. For instance, in the analysis of electroencephalography (EEG) signals, one commonly differentiates various types of waves associated with different frequency bands such as alpha waves with frequency spectrum of 8–13 Hz. In various studies, task-specific differences in the synchronization of brain regions between these frequency bands have been detected by means of EEG coherence analysis (e.g., [66,49,59,67,65]). In Eichler [24], the case of two univariate time series has been considered and illustrated by a neurological application concerning the detection of tremor-related cortical activity. Here, the frequency band of interest has been the range of typical tremor frequencies (1–8 Hz, see, e.g., [63,38]).

Since $|R_{ab}(\lambda)|^2 = |R_{ab}(-\lambda)|^2$, it is sufficient to consider intervals $[\lambda_1, \lambda_2] \subseteq [0, \pi]$. The general test for noncorrelation can be adjusted to the restricted test problem by setting

$$\psi_*\big(f(\lambda), \lambda\big) = \operatorname{vec}\big(R_{AB}(\lambda)\big)\, 1_{[\lambda_1,\lambda_2]}(\lambda),$$

where $1_A(\lambda)$ denotes the indicator function of the set A. Similarly as above we obtain $\Gamma_{\psi_*}(\lambda) = \Gamma_\psi(\lambda)\, 1_{[\lambda_1,\lambda_2]}(\lambda)$, which after some further calculations yields $\mu(\psi_*) = C_h\, C_{w,2}\, n_A\, n_B\, (\lambda_2 - \lambda_1)$ and

$$\sigma^2(\psi_*) = 2\pi\, C_h^2\, C_{w,4} \int_{\lambda_1}^{\lambda_2} \|R_{AA}(\lambda)\|^2\, \|R_{BB}(\lambda)\|^2\, d\lambda.$$

Setting $\lambda_1 = 0$ and $\lambda_2 = \pi$ and noting that $\int_\Pi \|R_{AB}(\lambda)\|^2\, d\lambda = 2 \int_0^\pi \|R_{AB}(\lambda)\|^2\, d\lambda$, we obtain the above expressions for $\mu(\psi)$ and $\sigma^2(\psi)$.

Example 3.9 (*Equality of spectra*). Suppose we are interested in testing whether the spectral densities of two time series $\{X_a(t)\}$ and $\{X_b(t)\}$ are identical. For instance, the two time series could be measurements of one variable in two experiments under different conditions. In that case, the null hypothesis signifies that the variable is not affected by the change in conditions.

Let $f(\lambda)$ be the spectral density matrix of the two time series. Then the null hypothesis of equal spectral densities can be formulated as

$$\mathrm{H}_0 : \psi\big(f(\lambda)\big) = \frac{f_{aa}(\lambda)}{f_{bb}(\lambda)} - 1 \equiv 0.$$

Noting that the first derivative of $\psi(Z, \lambda)$ is given by

$$\left.\frac{\partial \psi(Z)}{\partial Z}\right|_{Z=f(\lambda)} = \left(\frac{\delta_{ai}\,\delta_{aj}}{f_{bb}(\lambda)} - \frac{f_{aa}(\lambda)\,\delta_{bi}\,\delta_{bj}}{f_{bb}(\lambda)^2} \right)_{i,j=1,\ldots,d},$$

we directly obtain

$$\operatorname{tr}\left[\frac{\partial\psi}{\partial Z}\big(f(\lambda)\big)' f(\lambda)\overline{\frac{\partial\psi}{\partial Z}\big(f(\lambda)\big)' f(\lambda)}\right] = 2\,\frac{f_{aa}(\lambda)^2}{f_{bb}(\lambda)^2}\big(1-|R_{ab}(\lambda)|^2\big).$$

Since $\big(\partial\psi/\partial Z\big)\big(f(\lambda)\big)$ is real-valued and symmetric, the four terms in (8) take the same value. Thus the bias $\mu(\psi)$ and the variance $\sigma^2(\psi)$ are given by

$$\mu(\psi) = 2\,C_h\,C_{w,2}\int_\Pi \big(1-|R_{ab}(\lambda)|^2\big)\,d\lambda$$

and

$$\sigma^2(\psi) = 32\,\pi\,C_h^2\,C_{w,4}\int_\Pi \big(1-|R_{ab}(\lambda)|^2\big)^2\,d\lambda.$$

In this case, both parameters depend only on the spectral coherence $|R_{ab}(\lambda)|^2$ between the two processes. In particular, if the two processes are uncorrelated , we have $R_{ab}(\lambda)\equiv 0$ and the mean and variance do not depend on the unknown spectral densities $f_{aa}(\lambda)$ and $f_{bb}(\lambda)$.

For the analysis of interrelationships in multivariate time series, it is often of interest to distinguish direct interactions among a couple of time series from indirect relationships that involve other time series. In the frequency domain, such a distinction can be accomplished by partialization analysis based on the concepts of partial spectra and partial spectral coherencies (e.g., [31,5,14,56,44,64]). This leads to nonparametric hypotheses that can be formulated in terms of partial spectra or partial spectral matrices. More precisely, suppose that we are interested in testing the null hypothesis

$$\mathrm{H}_0:\ \psi\big(f_{AA|B}(\lambda),\lambda\big)\equiv 0, \tag{9}$$

where $f_{AA|B}(\lambda)$ is the partial spectral matrix of a process $\{X_A(t)\}$ after removing the linear effects of another process $\{X_B(t)\}$. Denoting the spectral matrix of the joint process by $f(\lambda)$, we define

$$\pi\big(f(\lambda)\big) = f_{AA|B}(\lambda) = f_{AA}(\lambda) - f_{AB}(\lambda)f_{BB}(\lambda)^{-1}f_{BA}(\lambda) \tag{10}$$

and $\psi_\pi\big(f(\lambda),\lambda\big) = \psi\big[\pi\big(f(\lambda)\big),\lambda\big]$. Then the null hypothesis can also be written as $\mathrm{H}_0:\ \psi_\pi\big(f(\lambda),\lambda\big)\equiv 0$, and we obtain as a test statistic

$$S_T(\psi_\pi) = \int_\Pi \big\|\psi_\pi\big(\hat f(\lambda),\lambda\big)\big\|^2\,d\lambda = \int_\Pi \big\|\psi\big(\hat f_{AA|B}(\lambda),\lambda\big)\big\|^2\,d\lambda.$$

The following theorem shows that the asymptotic mean $\mu(\psi_\pi)$ and variance $\sigma^2(\psi_\pi)$ can be derived from the expressions for $\mu(\psi)$ and $\sigma^2(\psi)$ by use of the following lemma. For simplicity, the theorem is formulated only for functions ψ such that $\tilde\Gamma_\psi(\lambda) = \Gamma_\psi(-\lambda)$. The extension to the general case is straightforward.

Lemma 3.10. *Let π be defined as in* (10). *Then*

$$\mathsf{D}_z \operatorname{vec}\pi\big(f(\lambda)\big)\big(f(\lambda)'\otimes f(\lambda)\big)\mathsf{D}_z\operatorname{vec}\pi\big(f(\lambda)\big)^* = f_{AA|B}(\lambda)'\otimes f_{AA|B}(\lambda), \tag{11}$$

where $\mathsf{D}_z\operatorname{vec}\pi(Z)$ *denotes the vector of first derivatives of* $\operatorname{vec}\pi(Z)$ *with respect to* $z = \operatorname{vec}(Z)$.

Proof. For $a_1, a_2 \in A$ let $\pi_{a_1a_2}(Z) = Z_{a_1a_2} - Z_{a_1B}(Z_{BB})^{-1}Z_{Ba_2}$. Then the first derivatives of $\pi_{ab}(Z)$ are given by

$$\frac{\partial\pi_{a_1a_2}(Z)}{\partial Z_{a_1a_2}} = 1, \quad \frac{\partial\pi_{a_1a_2}(Z)}{\partial Z_{a_1B}} = -Z_{a_2B}(Z_{BB})^{-1},$$

$$\frac{\partial\pi_{a_1a_2}(Z)}{\partial Z_{Ba_2}} = -(Z_{BB})^{-1}Z_{Ba_1}, \quad \frac{\partial\pi_{a_1a_2}(Z)}{\partial Z_{BB}} = (Z_{BB})^{-1}Z_{Ba_1}Z_{a_2B}(Z_{BB})^{-1},$$

while all other derivatives are zero. Inserting these expressions in (11), we obtain the right-hand side by straightforward calculations. □

Theorem 3.11. *Suppose that the assumptions of Theorem* 3.5 *hold. Furthermore*, *suppose that* $\tilde{\Gamma}_\psi(-\lambda) = \Gamma_\psi(\lambda)$. *Then under the null hypothesis* (9), *we have*

$$B_T^{1/2}\, T\, S_T(\psi_\pi) - B_T^{-1/2}\, \mu(\psi_\pi) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma^2(\psi_\pi)),$$

where

$$\mu(\psi_\pi) = C_h\, C_{w,2} \int_\Pi \mathrm{tr}\big[\Gamma_\psi(\lambda)\, \big(f_{AA|B}(\lambda)' \otimes f_{AA|B}(\lambda)\big)\big]\, d\lambda$$

and

$$\sigma^2(\psi_\pi) = 4\pi\, C_h^2\, C_{w,4} \int_\Pi \mathrm{tr}\big[\Gamma_\psi(\lambda)\big(f_{AA|B}(\lambda)' \otimes f_{AA|B}(\lambda)\big)$$
$$\times\big(\Gamma_\psi(\lambda) + \tilde{\Gamma}_\psi(\lambda)'\big)\big(f_{AA|B}(\lambda)' \otimes f_{AA|B}(\lambda)\big)\big]\, d\lambda.$$

Proof. Since $\mathrm{D}_z\psi_\pi(Z,\lambda) = \mathrm{D}_z\psi\big(\pi(Z),\lambda\big)\,\mathrm{D}_z\,\mathrm{vec}\,\pi(Z)$, we have

$$\Gamma_{\psi_\pi}(\lambda) = \big[\mathrm{D}_z\,\mathrm{vec}\,\pi\big(f(\lambda)\big)\big]^*\,\Gamma_\psi(\lambda)\,\mathrm{D}_z\,\mathrm{vec}\,\pi\big(f(\lambda)\big).$$

Furthermore, the derivatives of π satisfy

$$\frac{\partial\pi_{ab}(Z)}{\partial Z_{ij}} = \overline{\frac{\partial\pi_{ba}(Z)}{\partial Z_{ji}}},$$

from which it follows that

$$\tilde{\Gamma}_{ij,kl;\psi_\pi}(\lambda) = \Gamma_{ji,lk;\psi_\pi}(\lambda) = \sum_{i',j',k',l'=1}^{d} \overline{\frac{\partial\pi_{i'j'}}{\partial Z_{ji}}}\big(f(\lambda)\big)\,\frac{\partial\pi_{k'l'}}{\partial Z_{lk}}\big(f(\lambda)\big)\,\Gamma_{i'j',k'l';\psi}(\lambda)$$
$$= \sum_{i',j',k',l'=1}^{d} \frac{\partial\pi_{j'i'}}{\partial Z_{ij}}\big(f(\lambda)\big)\,\overline{\frac{\partial\pi_{l'k'}}{\partial Z_{kl}}}\big(f(\lambda)\big)\,\tilde{\Gamma}_{j'i',l'k';\psi}(\lambda)$$

and thus

$$\tilde{\Gamma}_{\psi_\pi}(\lambda)' = \big[\mathrm{D}_z\mathrm{vec}\,\pi\big(f(\lambda)\big)\big]^*\,\tilde{\Gamma}_\psi(\lambda)'\,\mathrm{D}_z\mathrm{vec}\,\pi\big(f(\lambda)\big).$$

The result now follows from Theorem 3.5 and Lemma 3.10. □

Example 3.12 (*Partial noncorrelation*). We first consider the problem of testing whether two multivariate time series $\{X_A(t)\}$ and $\{X_B(t)\}$ are uncorrelated after the linear effects of a third time series $\{X_C(t)\}$ have been removed. Let $f(\lambda)$ be the joint spectral density matrix of the three processes. Then the null hypothesis of interest can be formulated as

$$\mathrm{H}_0:\ \psi_\pi\big(f(\lambda),\lambda\big) = \mathrm{vec}\big(R_{AB|C}(\lambda)\big) \equiv 0.$$

From Theorem 3.11 and Example 3.8, we find that for the statistic

$$S_T(\psi_\pi) = \int_\Pi \big\| \hat{R}_{AB|C}(\lambda) \big\|^2 \, d\lambda$$

the constants $\mu(\psi_\pi)$ and $\sigma^2(\psi_\pi)$ are given by

$$\mu(\psi_\pi) = 2\pi\, C_h\, C_{w,2}\, n_A\, n_B$$

and

$$\sigma^2(\psi_\pi) = 4\pi\, C_h^2\, C_{w,4} \int_\Pi \big\| R_{AA|C}(\lambda) \big\|^2 \big\| R_{BB|C}(\lambda) \big\|^2 \, d\lambda.$$

In the special case of two univariate time series $\{X_a(t)\}$ and $\{X_b(t)\}$, the limiting distribution of the statistic $S_T(\psi_\pi)$ becomes independent of the unknown spectral matrix $f(\lambda)$. We obtain as a test statistic

$$Q_T(\psi_\pi) = \frac{T\, S_T(\psi_\pi) - 2\pi\, C_h\, C_{w,2}/B_T}{2\pi\, C_h \sqrt{2\, C_{w,4}/B_T}},$$

which does not depend on the unknown spectral matrix and is asymptotically standard normally distributed.

Example 3.13 (*Equality of partial spectra*). Next, let $\{X_A(t)\}$ and $\{X_B(t)\}$ be two independent multivariate time series with joint spectral density matrix $f(\lambda)$. Similarly as in Example 3.9, equality of the two partial spectral densities $f_{aa|A'}(\lambda)$ and $f_{bb|B'}(\lambda)$, where $A' = A\backslash\{a\}$ and $B' = B\backslash\{b\}$, can be tested using the statistic

$$S_T(\psi_\pi) = \int_\Pi \bigg\| \frac{\hat{f}_{aa|A'}(\lambda)}{\hat{f}_{bb|B'}(\lambda)} - 1 \bigg\|^2 \, d\lambda.$$

Note, that the independence assumption implies that $f_{aa|A'B'}(\lambda) = f_{aa|A'}(\lambda)$ and $f_{bb|A'B'}(\lambda) = f_{bb|B'}(\lambda)$. From Theorem 3.11 and Example 3.9, we now find that the standardized version

$$Q_T(\psi_\pi) = \frac{T\, S_T - 4\pi\, C_h\, C_{w,2}/B_T}{8\pi\, C_h \sqrt{C_{w,4}/B_T}},$$

is asymptotically standard normally distributed.

4. Extensions

In this section, we generalize the nonparametric approach to include also parametric and semi-parametric hypotheses about the spectral matrix. Furthermore, it is shown that the results are still valid if the Euclidian norm used to measure departures from the null hypothesis is replaced by general discrepancies.

4.1. Parametric and semi-parametric hypotheses

First, we consider hypotheses about the spectral matrix that also depend on an unknown parameter θ_0 from a finite-dimensional parameter set Θ. More precisely, we are interested in testing the null hypothesis

$$\mathrm{H}_0 : \psi(f(\lambda), \lambda, \theta_0) \equiv 0 \quad \text{for some } \theta_0 \in \Theta \tag{12a}$$

against the alternative

$$\mathrm{H}_a : \psi(f(\lambda), \lambda, \theta) \not\equiv 0 \quad \text{for all } \theta \in \Theta, \tag{12b}$$

where $\Theta \subseteq \mathbb{R}^p$. We assume that the unknown parameter θ_0 is uniquely determined—under the null hypothesis and the alternative—by the distribution of the process and can be estimated by $\hat{\theta}$. Thus, replacing θ_0 by its estimator $\hat{\theta}$, we can use

$$S_T(\psi) = \int_\Pi \left\|\psi\big(\hat{f}(\lambda), \lambda, \hat{\theta}\big)\right\|^2 d\lambda$$

as a test statistic for the above test problem. We impose the following conditions on the estimator $\hat{\theta}$ and on the function $\psi(Z, \lambda, \theta)$.

Assumption 4.1. Let $\psi : D \times \Pi \times \Theta \to \mathbb{C}^r$, where $\Theta \subseteq \mathbb{R}^p$ and D is an open subset of $\mathbb{C}^{d\times d}$.

(i) $\psi(Z, \lambda, \theta)$ is holomorphic in Z. There exists a positive constant ρ such that for all $\lambda \in \Pi$ the ball $B_{\rho,\lambda} = \big\{Z \in \mathbb{C}^{d\times d} \,\big|\, \|f(\lambda) - Z\| \leqslant \rho\big\}$ is contained in D and

$$\sup_{\|\theta-\theta_0\|\leqslant\rho} \ \sup_{\lambda\in\Pi} \ \sup_{Z\in B_{\rho,\lambda}} \big\|\psi(Z, \lambda, \theta)\big\| < \infty.$$

(ii) $\psi\big(f(\lambda), \lambda, \theta\big)$ and $\mathsf{D}_z\psi\big(f(\lambda), \lambda, \theta\big)$ are piecewise Lipschitz continuous in $\lambda \in \Pi$.

(iii) $\psi\big(f(\lambda), \lambda, \theta\big)$ and $\mathsf{D}_z\psi\big(f(\lambda), \lambda, \theta\big)$ are twice, respectively, once continuously differentiable in a neighbourhood of θ_0. The partial derivatives

$$\left.\frac{\partial\psi_i(Z, \lambda, \theta)}{\partial\theta_k}\right|_{Z=f(\lambda)}, \quad \left.\frac{\partial^2\psi_i(Z, \lambda, \theta)}{\partial\theta_k\partial\theta_l}\right|_{Z=f(\lambda)} \quad \text{and} \quad \left.\frac{\partial^2\psi_i(Z, \lambda, \theta)}{\partial\theta_k\partial Z_{ab}}\right|_{Z=f(\lambda)}$$

are uniformly bounded for all $\lambda \in \Pi$ and $\|\theta - \theta_0\| \leqslant \rho$.

(iv) $\int_\Pi \big\|\mathsf{D}_z\psi\big(f(\lambda), \lambda, \theta\big)\big\|^2 d\lambda > 0$.

Assumption 4.2. The parameter θ_0 lies in the interior of Θ and its estimator $\hat{\theta}$ satisfies $\|\hat{\theta}-\theta_0\| = O_P(T^{-1/2})$.

The rate of convergence of $\hat{\theta}$ guarantees that the error due to estimation of θ_0 is asymptotically negligible. More precisely, we have the following version of Lemma 3.4.

Lemma 4.3. *Suppose Assumptions* 3.1, 3.3, 4.1, *and* 4.2 *hold. Then under the null hypothesis* (12)

$$S_T(\psi) = \int_\Pi \big\|\mathrm{vec}\big(\hat{f}(\lambda) - f(\lambda)\big)\big\|^2_{\Gamma_\psi(\lambda)} d\lambda + o_P\big((B_T^{1/2}\,T)^{-1}\big),$$

where $\Gamma_\psi(\lambda) = \mathsf{D}_z\psi\big(f(\lambda), \lambda, \theta_0\big)^*\mathsf{D}_z\psi\big(f(\lambda), \lambda, \theta_0\big)$.

Proof. Using Cauchy's estimate for the derivatives of $\psi(Z, \lambda, \theta)$ with respect to Z, we find that there exists $\delta > 0$ such that

$$\psi\big(\hat{f}(\lambda), \lambda, \hat{\theta}\big) = \psi\big(f(\lambda), \lambda, \hat{\theta}\big) + \mathsf{D}_z\psi\big(f(\lambda), \lambda, \hat{\theta}\big)\big(\hat{f}(\lambda) - f(\lambda)\big) + R_1(\lambda)$$

with $\|R_1(\lambda)\| \leqslant C\|\hat{f}(\lambda) - f(\lambda)\|^2$, whenever

$$\|\hat{\theta} - \theta_0\| \leqslant \delta \quad \text{and} \quad \max_{\lambda\in\Pi} \|\hat{f}(\lambda) - f(\lambda)\| \leqslant \delta. \tag{13}$$

Furthermore, by Assumption 4.2(iii), Taylor approximation of the first two terms about θ_0 yields

$$\psi\big(\hat{f}(\lambda), \lambda, \hat{\theta}\big) = \mathsf{D}_\theta\psi\big(f(\lambda), \lambda, \theta_0\big)(\hat{\theta} - \theta_0)$$
$$+\mathsf{D}_z\psi\big(f(\lambda), \lambda, \theta_0\big)\big(\hat{f}(\lambda) - f(\lambda)\big) + R_2(\lambda) + R_1(\lambda),$$

with $\|R_2(\lambda)\| \leqslant C\|\hat{\theta} - \theta_0\|^2 + C\|\hat{\theta} - \theta_0\| \cdot \|\hat{f}(\lambda) - f(\lambda)\|$, whenever (13) holds (possibly with smaller δ). It follows that, still under condition (13) ,

$$S_T(\psi) = \int_\Pi \|\mathrm{vec}\big(\hat{f}(\lambda) - f(\lambda)\big)\|^2_{\Gamma_\psi(\lambda)}\, d\lambda$$
$$+(\hat{\theta} - \theta_0)'\left[\int_\Pi \big[\mathsf{D}_\theta\psi\big(f(\lambda), \lambda, \theta_0\big)\big]^* \mathsf{D}_\theta\psi\big(f(\lambda), \lambda, \theta_0\big)\, d\lambda\right](\hat{\theta} - \theta_0)$$
$$+(\hat{\theta} - \theta_0)'\left[\int_\Pi \big[\mathsf{D}_\theta\psi\big(f(\lambda), \lambda, \theta_0\big)\big]^* \mathsf{D}_z\psi\big(f(\lambda), \lambda, \theta_0\big)\big(\hat{f}(\lambda) - f(\lambda)\big)\, d\lambda\right] + R_3,$$

where the remainder term R_3 is of order $O_P\big((B_T T)^{-3/2}\big)$ by the $\sqrt{T}$-consistency of $\hat{\theta}$ and (5). We note that the linear functional in the third term on the right-hand side is of order $O_P(T^{-1/2})$ (e.g., [60, Theorem 6.1.2]), which together with the $\sqrt{T}$-consistency of $\hat{\theta}$ yields $O_P(T^{-1})$ for the second and the third terms. By a similar argument as in the proof of Lemma 3.4, it now follows that the difference of $S_T(\psi)$ and the first term on the right-hand side is of order $o_P\big((B_T^{1/2} T)^{-1}\big)$. □

From the lemma and Theorem B.2, it follows that Theorem 3.5 remains valid for test statistics $S_T(\psi)$ with parameter estimates substituted for any unknown parameters. We give two examples.

Example 4.4 (*Separability*). Recently, Matsuda and Yajima [47] discussed a test for separability of the correlation structure of multivariate Gaussian time series. Such models have been considered in various fields (e.g., [9,33]). The correlation structure of a stationary multivariate time series $\{X(t)\}$ is said to be separable if

$$\mathrm{cov}\big(X_a(t), X_b(s)\big) = \sigma_a\, \sigma_b\, \rho_1(a, b)\, \rho_2(t - s),$$

with $\rho_1(a, a) = 1$ and $\rho_2(0) = 1$ for all $a, b = 1, \ldots, d$ and $t, s \in \mathbb{Z}$ or, equivalently, if the spectral matrix $f(\lambda)$ is of the form

$$f(\lambda) = \Sigma\, f_0(\lambda),$$

where Σ is a $d \times d$ positive definite matrix and $f_0(\lambda)$ is a scalar-valued function. Without loss of generality, we can choose Σ as the variance $\operatorname{var}\bigl(X(t)\bigr)$. It follows that

$$f_0(\lambda) = \frac{1}{d} \operatorname{tr}\bigl[f(\lambda)\, V_\Sigma^{-1}\bigr],$$

where $V_\Sigma = \operatorname{diag}(\sigma_1^2, \ldots, \sigma_d^2)$ is the diagonal matrix of variances σ_a^2. Therefore, the null hypothesis of separability of the correlation structure can be formulated as

$$\mathrm{H}_0 : \psi\bigl(f(\lambda), \Sigma\bigr) = \operatorname{vec}\left(f(\lambda) \left[\frac{1}{d} \Sigma \operatorname{tr}\bigl(f(\lambda) V_\Sigma^{-1}\bigr) \right]^{-1} - \mathbb{1}_d \right) = 0,$$

where $\mathbb{1}_d$ denotes the $d \times d$ identity matrix. In this case, the variance Σ is an unknown parameter, which can be estimated by

$$\hat{\Sigma} = \frac{1}{T} \sum_{t=1}^{T} X(t) X(t)',$$

since $\{X(t)\}$ is assumed to have zero mean. Similarly, V_Σ can be estimated by $\hat{V}_\Sigma = V_{\hat{\Sigma}} = \operatorname{diag}(\hat{\sigma}_1^2, \ldots, \hat{\sigma}_d^2)$, where $\hat{\sigma}_1^2, \ldots, \hat{\sigma}_d^2$ are the diagonal entries in $\hat{\Sigma}$. This suggests to use

$$S_T(\psi) = \int_\Pi \left\| \operatorname{vec}\Bigl(\hat{f}(\lambda) \Bigl[\frac{1}{d} \hat{\Sigma} \operatorname{tr}\bigl(\hat{f}(\lambda) V_{\hat{\Sigma}}^{-1}\bigr) \Bigr]^{-1} - \mathbb{1}_d \Bigr) \right\|^2 d\lambda$$

for a test for separability. For the derivation of the limiting distribution, we first note that the function $\psi(Z, \Sigma)$ can be written as

$$\psi\bigl(Z, \Sigma\bigr) = \frac{d}{\operatorname{tr}\bigl(Z\, V_\Sigma^{-1}\bigr)} (\Sigma^{-1} \otimes \mathbb{1}_d) \operatorname{vec}\bigl(Z\bigr) - \operatorname{vec}(\mathbb{1}_d),$$

which, under the null hypothesis, leads to

$$\mathsf{D}_z \psi\bigl(f(\lambda), \Sigma\bigr) = f_0(\lambda)^{-1} \left(\Sigma^{-1} \otimes \mathbb{1}_d - \frac{1}{d} \operatorname{vec}(\mathbb{1}_d) \operatorname{vec}(V_\Sigma^{-1})' \right).$$

Furthermore, we have $f(\lambda)' \otimes f(\lambda) = f_0(\lambda)^2 \bigl(\Sigma \otimes \Sigma\bigr)$. Simple calculations then show that the constants $\mu(\psi)$ and $\sigma^2(\psi)$ in Theorem 3.5 are given by

$$\mu(\psi) = 2\pi\, C_h\, C_{w,2} \left(d^2 - 2 + \frac{\rho}{d} \right)$$

and

$$\sigma^2(\psi) = 4\pi\, C_h^2\, C_{w,4} \left(d^2 - 2\, \frac{\rho}{d} + \frac{\rho^2}{d^2} \right),$$

with

$$\rho = \operatorname{tr}\bigl(\Sigma\, V_\Sigma^{-1}\, \Sigma\, V_\Sigma^{-1}\bigr) = \sum_{i,j=1}^{d} \frac{\sigma_{ij}^2}{\sigma_i^2\, \sigma_j^2},$$

where σ_{ij} are the nondiagonal entries of Σ. Note that both constants $\mu(\psi)$ and $\sigma^2(\psi)$ depend only on the parameter Σ and thus can be estimated efficiently by substituting the estimator $\hat{\Sigma}$ for Σ. It follows from Slutsky's theorem that the test statistic

$$Q_T(\psi) = \frac{T\, S_T(\psi) - \hat{\mu}(\psi)/B_T}{\hat{\sigma}(\psi)/\sqrt{B_T}}$$

is asymptotically standard normally distributed.

Example 4.5 (*Goodness-of-fit*). As a second example, suppose that we are interested in testing the fit of a parametric model. More precisely, we consider the composite hypothesis

$$\mathrm{H}_0 : f \in \mathscr{F}_\Theta \quad \text{against } \mathrm{H}_a : f \notin \mathscr{F}_\Theta, \tag{14}$$

where $\mathscr{F}_\Theta = \{f_\theta | \theta \in \Theta\}$ denotes a parametric class of spectral matrix models and Θ is a parameter set. For univariate time series, such goodness-of-fit tests have a long history (e.g., [53,1,39,50]); more recently, the case of multivariate stationary and cointegrated processes has been considered by Duchesne and Roy [18], Paparoditis [51], and Duchesne [17]. The general framework discussed in this paper provides an alternative test for arbitrary multivariate stationary models.

The test problem (14) can be formulated in the form (12) by setting

$$\begin{aligned}\psi\big(f(\lambda), \theta_0\big) &= \mathrm{vec}\big(f_{\theta_0}(\lambda)^{-1/2} f(\lambda)(f_{\theta_0}(\lambda)^{-1/2})^* - \mathbb{1}_d\big)\\ &= \big(f_{\theta_0}(-\lambda)^{-1/2} \otimes f_{\theta_0}(\lambda)^{-1/2}\big)\mathrm{vec}\big(f(\lambda)\big) - \mathrm{vec}(\mathbb{1}_d),\end{aligned}$$

where $f_{\theta_0}(\lambda) = f_{\theta_0}(\lambda)^{1/2}(f_{\theta_0}(\lambda)^{1/2})^*$ is the Cholesky decomposition of the Hermitian matrix $f_{\theta_0}(\lambda)$ and $\theta_0 \in \Theta$ is the unknown parameter of the best fitting model. Estimating θ_0 by some $\sqrt{T}$-consistent estimator $\hat{\theta}$, the model fit can be evaluated by the statistic

$$S_T(\psi) = \int_\Pi \left\| \mathrm{vec}\big[f_{\hat{\theta}}(\lambda)^{-1/2} \hat{f}(\lambda)(f_{\hat{\theta}}(-\lambda)^{-1/2})' - \mathbb{1}_d\big] \right\|^2 d\lambda,$$

which by Lemma 4.3 and Theorem 3.5 is asymptotically normally distributed. For the calculation of the corresponding bias $\mu(\psi)$ and variance $\sigma^2(\psi)$, we note that $\mathsf{D}_z\psi(f(\lambda), \theta_0) = f_{\theta_0}(-\lambda)^{-1/2} \otimes f_{\theta_0}(-\lambda)^{-1/2}$, which leads to

$$\Gamma_\psi(\lambda) = f_{\theta_0}(-\lambda)^{-1} \otimes f_{\theta_0}(\lambda)^{-1} = \big(f_{\theta_0}(\lambda)^{-1}\big)' \otimes f_{\theta_0}(\lambda)^{-1}. \tag{15}$$

Since under the null hypothesis $f(\lambda) = f_{\theta_0}(\lambda)$, we obtain for the bias and variance

$$\mu(\psi) = 2\pi\, C_h\, C_{w,2}\, d^2 \quad \text{and} \quad \sigma^2(\psi) = 16\pi^2\, C_h^2\, C_{w,4}\, d^2.$$

Consequently, the standardized test statistic is given by

$$Q_T(\psi) = \frac{T\, S_T(\psi) - 2\pi\, C_h\, C_{w,2}\, d^2/B_T}{4\pi\, C_h\, d\sqrt{C_{w,4}/B_T}}$$

and does not depend on the spectral matrix $f(\lambda)$ or the parameter θ_0.

Example 4.6 (*Goodness-of-fit for partial spectra*). While partialization analysis allows identification of the direct interactions among the components of a process $\{X_A(t)\}$ by removing the linear effects of a second process $\{X_B(t)\}$, it is sometimes of interest to investigate the nature of these direct interactions further. For instance, one might be interested in whether the process $\{X_A(t)\}$ is white noise or whether one component of $\{X_A(t)\}$ Granger-causes another component after removal of the linear effects of $\{X_B(t)\}$ (for the concept of Granger causality, see [32]). This can be accomplished by fitting appropriate parametric models to the partialized process $\{X_{A|B}(t)\}$—that is, to the residual process obtained by removing the linear effects of $\{X_B(t)\}$—and applying the above goodness-of-fit test to the corresponding spectral matrix, namely the partial spectral matrix $f_{AA|B}(\lambda)$.

In the case of testing whether $\{X_{A|B}(t)\}$ is white noise, this leads to the null hypothesis

$$\mathrm{H}_0 :\ \mathrm{vec}\big(\Sigma_{AA|B}^{-1/2}\, f_{AA|B}(\lambda)\, \Sigma_{AA|B}^{-1/2} - \mathbb{1}_d\big) \equiv 0,$$

where $\Sigma_{AA|B}$ is the covariance matrix of the partialized process $X_{A|B}(t)$ and $A^{1/2}(A^{1/2})'$ denotes the Cholesky decomposition of a symmetric matrix A. Setting $\psi\big(f(\lambda), \Sigma\big) = \mathrm{vec}\big(\Sigma^{-1/2} f(\lambda)\, \Sigma^{-1/2} - \mathbb{1}_d\big)$, it follows from the previous example and Theorem 3.11 that $\mu(\psi_\pi) = 2\pi C_h C_{w,2} d^2$ and $\sigma^2(\psi_\pi) = 16\pi^2 C_h^2 C_{w,4} d^2$.

4.2. General discrepancies

For the construction of test statistics, the deviation from the null hypothesis so far has been measured by the Euclidean norm, which led to test statistics of the form

$$S_T(\psi) = \int_\Pi \big\|\psi\big(\hat{f}(\lambda), \lambda, \hat{\theta}\big)\big\|^2\, d\lambda,$$

where $\|\cdot\|$ denotes the Euclidean norm. More generally, other distance measures may be used. For instance, many authors have considered Whittle's log-likelihood for estimating the parameters of a time series model (e.g., [20,19,41,11]), which suggests to assess the goodness-of-fit by a similar discrepancy measure.

Taniguchi and Kakizawa [60] studied parameter estimation for multivariate processes based on general discrepancy functions of the form

$$D(f, \theta) = \int_\Pi K\big(f(\lambda), \lambda, \theta\big)\, d\lambda,$$

where $K(Z, \lambda, \theta)$ is a complex-valued function that is real-valued and nonnegative for all nonnegative definite Hermitian matrices Z. To extend our approach to this general setting, we consider the problem of testing the null hypothesis

$$\mathrm{H}_0 : K\big(f(\lambda), \lambda, \theta_0\big) \equiv 0 \quad \text{for some } \theta_0 \in \Theta \tag{16a}$$

against the alternative

$$\mathrm{H}_a : K\big(f(\lambda), \lambda, \theta\big) \not\equiv 0 \quad \text{for all } \theta \in \Theta, \tag{16b}$$

where $K(Z, \lambda, \theta)$ has similar properties as above. Integration of $K(Z, \lambda, \theta)$ then leads to the test statistic

$$S_T(\psi, D) = \int_\Pi K\big(\hat{f}(\lambda), \lambda, \hat{\theta}\big)\, d\lambda. \tag{17}$$

In the following, we show that the results of the previous section can be generalized to this larger class of test statistics. To this end, we impose the following assumptions on the function K.

Assumption 4.7. Let $K : D \times \Pi \times \Theta \to \mathbb{C}$, where D is an open subset of $\mathbb{C}^{d\times d}$ and $\Theta \subseteq \mathbb{R}^p$.

(i) $K(Z, \lambda, \theta)$ is real-valued and nonnegative for all nonnegative definite Hermitian matrices $Z \in \mathbb{C}^{d\times d}$, $\lambda \in \Pi$, and $\theta \in \Theta$.

(ii) $K(Z, \lambda, \theta)$ is holomorphic with respect to Z. There exists a positive constant ρ such that for all $\lambda \in \Pi$ the ball $B_{\rho,\lambda} = \{Z \in \mathbb{C}^{d\times d} \,|\, \|f(\lambda) - Z\| \leqslant \rho\}$ is contained in D and

$$\sup_{\|\theta-\theta_0\|\leqslant\rho} \sup_{\lambda\in\Pi} \sup_{Z\in B_{\rho,\lambda}} \left|K(Z, \lambda, \theta)\right| < \infty.$$

(iii) $K(Z, \lambda, \theta)$ and its second derivative with respect to $z = \mathrm{vec}(Z)$,

$$\mathsf{H}_{zz} K(Z, \lambda, \theta) = \frac{\partial^2}{\partial z \partial z'} K(Z, \lambda, \theta),$$

are Lipschitz continuous in λ except for possibly finitely many points.

(iv) $K\big(f(\lambda), \lambda, \theta\big)$ and $\mathsf{D}_z K\big(f(\lambda), \lambda, \theta\big)$ are twice continuously differentiable with respect to θ in a neighbourhood of θ_0 with

$$\sup_{\lambda\in\Pi} \left\|\mathsf{D}_\theta K\big(f(\lambda), \lambda, \theta_0\big)\right\| < \infty, \qquad \sup_{\lambda\in\Pi} \left\|\mathsf{H}_{\theta z} K\big(f(\lambda), \lambda, \theta_0\big)\right\| < \infty$$

and

$$\sup_{\|\theta-\theta_0\|<\rho} \sup_{\lambda\in\Pi} \left|\frac{\partial^2 K\big(f(\lambda), \lambda, \theta\big)}{\partial\theta_i \partial\theta_j}\right| < \infty, \qquad \sup_{\|\theta-\theta_0\|<\rho} \sup_{\lambda\in\Pi} \left|\frac{\partial^3 K\big(f(\lambda), \lambda, \theta\big)}{\partial\theta_i \partial\theta_j \partial Z_{ab}}\right| < \infty$$

for some $\rho > 0$.

(v) $\mathsf{H}_{zz} K\big(f(\lambda), \lambda, \theta\big)$ is continuously differentiable with respect to θ in a neighbourhood of θ_0 with

$$\sup_{\|\theta-\theta_0\|<\rho} \sup_{\lambda\in\Pi} \left|\frac{\partial^3 K\big(f(\lambda), \lambda, \theta\big)}{\partial\theta_i \partial Z_{ab} \partial Z_{cd}}\right| < \infty$$

for some $\rho > 0$.

(v) $\int_\Pi \left\|\mathsf{H}_{zz} K\big(f(\lambda), \lambda, \theta\big)\right\| d\lambda > 0$.

We note that the assumptions imply that $K\big(Z, \lambda, \theta\big) \geqslant K\big(f(\lambda), \lambda, \theta_0\big)$ for every nonnegative Hermitian matrix Z and $\theta \in \Theta$. It follows that the directional derivatives

$$\mathsf{D}_z K\big(f(\lambda), \lambda, \theta_0\big) \mathrm{vec}(Z)$$

for any nonnegative Hermitian matrix Z with $\|Z\| = 1$ and

$$\mathsf{D}_\theta K\big(f(\lambda), \lambda, \theta_0\big)\, \theta$$

for any $\theta \in \mathbb{R}^p$ with $\|\theta\| = 1$ both vanish. Furthermore, since $K(Z, \lambda, \theta)$ takes only real values for all nonnegative definite Hermitian matrices Z, a power series expansion of $K(Z, \lambda, \theta)$ about Z shows that

$$\frac{\partial K(Z, \lambda, \theta)}{\partial Z_{ij}} = \overline{\frac{\partial K(Z, \lambda, \theta)}{\partial Z_{ji}}} \quad \text{and} \quad \frac{\partial^2 K(Z, \lambda, \theta)}{\partial Z_{ij} \partial Z_{kl}} = \overline{\frac{\partial^2 K(Z, \lambda, \theta)}{\partial Z_{ji} \partial Z_{lk}}}.$$

Consequently, we have $\mathsf{H}_{zz} K\big(Z, \lambda, \theta\big) = K_{dd} \big[\mathsf{H}_{zz} K\big(Z, \lambda, \theta\big)\big]^* K_{dd}$, which implies that the matrix

$$\Gamma_K(\lambda) = \tfrac{1}{2} K_{dd} \, \mathsf{H}_{zz} K\big(f(\lambda), \lambda, \theta_0\big)$$

is Hermitian. Thus, the conditions on the function K ensure that for nonnegative definite Hermitian matrices Z close to $f(\lambda)$, the function $K(Z, \lambda, \theta_0)$ can be approximated by a quadratic norm $\|\mathrm{vec}(Z - f(\lambda))\|^2_{\Gamma_K(\lambda)}$. The following lemma shows that this approximation of the discrepancy can be applied to the general test statistic in (17).

Lemma 4.8. *Suppose that Assumptions* 3.1, 3.3, 4.2, *and* 4.7 *hold. Then under the null hypothesis* (16)

$$S_T(K) = \int_\Pi \big\|\mathrm{vec}\big(\hat{f}(\lambda) - f(\lambda)\big)\big\|^2_{\Gamma_K(\lambda)} \, d\lambda + o_P\big((B_T^{1/2} T)^{-1}\big),$$

where

$$\Gamma_K(\lambda) = \tfrac{1}{2} K_{dd} \, \mathsf{H}_{zz} K\big(f(\lambda), \lambda, \theta_0\big). \tag{18}$$

Proof. Expanding $K\big(\hat{f}(\lambda), \lambda, \hat{\theta}\big)$ about $Z = f(\lambda)$, we obtain

$$\begin{aligned} &K\big(\hat{f}(\lambda), \lambda, \hat{\theta}\big) \\ &\quad = K\big(f(\lambda), \lambda, \hat{\theta}\big) + \mathsf{D}_z K\big(f(\lambda), \lambda, \hat{\theta}\big)\mathrm{vec}\big(\hat{f}(\lambda) - f(\lambda)\big) \\ &\qquad + \tfrac{1}{2}\,\mathrm{vec}\big(\hat{f}(\lambda) - f(\lambda)\big)' \mathsf{H}_{zz} K\big(f(\lambda), \lambda, \hat{\theta}\big)\mathrm{vec}\big(\hat{f}(\lambda) - f(\lambda)\big) + R_1(\lambda). \end{aligned} \tag{19}$$

Using Cauchy's estimate for the derivatives with respect to $z = \mathrm{vec}(Z)$ we find by Assumption 4.7(iii) that $|R_1(\lambda)| \leqslant C \|\hat{f}(\lambda) - f(\lambda)\|^3$ whenever

$$\|\hat{\theta} - \theta_0\| \leqslant \delta \quad \text{and} \quad \max_{\lambda \in \Pi} \|\hat{f}(\lambda) - f(\lambda)\| \leqslant \delta. \tag{20}$$

Furthermore, conditions (iv) and (v) of the same assumption imply that

$$K\big(f(\lambda), \lambda, \hat{\theta}\big) = K\big(f(\lambda), \lambda, \theta_0\big) + \mathsf{D}_\theta K\big(f(\lambda), \lambda, \theta_0\big)(\hat{\theta} - \theta_0) + R_2(\lambda)$$

and

$$\mathsf{D}_z K\big(f(\lambda), \lambda, \hat{\theta}\big) = \mathsf{D}_z K\big(f(\lambda), \lambda, \theta_0\big) + (\hat{\theta} - \theta_0)' \mathsf{H}_{\theta z} K\big(f(\lambda), \lambda, \theta_0\big) + R_3(\lambda),$$

where both $|R_2(\lambda)|$ and $\|R_3(\lambda)\|$ are bounded by $C\|\hat{\theta}-\theta_0\|^2$, while

$$\begin{aligned}&\tfrac{1}{2}\,\mathrm{vec}\big(\hat{f}(\lambda)-f(\lambda)\big)'\mathsf{H}_{zz}K\big(f(\lambda),\lambda,\hat{\theta}\big)\,\mathrm{vec}\big(\hat{f}(\lambda)-f(\lambda)\big)\\ &\quad=\big\|\mathrm{vec}\big(\hat{f}(\lambda)-f(\lambda)\big)\big\|^2_{\Gamma_K(\lambda)}+R_4(\lambda),\end{aligned}$$

with $|R_4(\lambda)|\leqslant C\|\hat{\theta}-\theta_0\|\,\|\hat{f}(\lambda)-f(\lambda)\|^2$, where we have used that $K_{dd}\mathrm{vec}(Z)=\mathrm{vec}(\overline{Z})$ for all Hermitian $d\times d$ matrices Z. Substituting these expressions into (19) and integrating over λ, we get by (5) and the $\sqrt{T}$ consistency of $\hat{\theta}$

$$\begin{aligned}S_T(K)=&\int_\Pi\big\|\mathrm{vec}\big(\hat{f}(\lambda)-f(\lambda)\big)\big\|^2_{\Gamma_K(\lambda)}\,d\lambda+\int_\Pi \mathsf{D}_\theta K\big(f(\lambda),\lambda,\theta_0\big)(\hat{\theta}-\theta_0)\,d\lambda\\ &+\int_\Pi \mathsf{D}_z K\big(f(\lambda),\lambda,\theta_0\big)\,\mathrm{vec}\big(\hat{f}(\lambda)-f(\lambda)\big)\,d\lambda\\ &+(\hat{\theta}-\theta_0)'\left[\int_\Pi \mathsf{H}_{\theta z}K\big(f(\lambda),\lambda,\hat{\theta}\big)\,\mathrm{vec}\big(\hat{f}(\lambda)-f(\lambda)\big)\,d\lambda\right]+R,\end{aligned}$$

where the remainder R is of order $o_P\big((B_T^{1/2}\,T)^{-1}\big)$ whenever (20) holds. As remarked before, by Assumption 4.7(i) the directional derivatives in the second and third term vanish. Furthermore, the integral in the fourth term on the right-hand side is of order $O_P(T^{-1/2})$. The statement of the lemma follows now by a similar argument as in the proof of Lemma 3.4. □

With the previous lemma and Theorem B.2, we can establish the following result, which extends Theorem 3.5 to the general case.

Theorem 4.9. *Suppose that Assumptions* 3.1, 3.3, 4.2, *and* 4.7 *hold. Then*

$$B_T^{1/2}\,T\,S_T(K)-B_T^{-1/2}\,\mu(K)\xrightarrow{\mathcal{D}}\mathcal{N}(0,\sigma^2(K)),$$

where $\mu(K)$ and $\sigma^2(K)$ are given by (6) *and* (7), *respectively, with $\Gamma_K(\lambda)$ substituted for $\Gamma_\psi(\lambda)$.*

Example 4.10. We consider again the problem of assessing the fit of a parametric model. Alternatively to the L^2 distance discussed in Example 4.5, the difference between the estimated spectral matrix $\hat{f}(\lambda)$ and the fitted spectral matrix $f_{\hat{\theta}}(\lambda)$ can be measured by the Kullback–Leibler discrepancy [43]

$$D(\hat{f},f_{\hat{\theta}})=\int_\Pi\left[\log\left(\frac{\det f_{\hat{\theta}}(\lambda)}{\det \hat{f}(\lambda)}\right)+\mathrm{tr}\big(\hat{f}(\lambda)f_{\hat{\theta}}(\lambda)^{-1}-\mathbb{1}_d\big)\right]d\lambda.$$

Setting

$$K(Z,\lambda,\theta)=\log\left(\frac{\det f_\theta(\lambda)}{\det Z}\right)+\mathrm{tr}\big(Z\,f_\theta(\lambda)^{-1}-\mathbb{1}_d\big),$$

we find for the second derivatives with respect to Z

$$\left.\frac{\partial^2 K(Z,\lambda,\theta_0)}{\partial Z_{ij}\partial Z_{kl}}\right|_{Z=f_{\theta_0}(\lambda)}=\big(f_{\theta_0}(\lambda)^{-1}\big)_{jk}\big(f_{\theta_0}(\lambda)^{-1}\big)_{li}.$$

It follows that

$$\Gamma_K(\lambda) = \tfrac{1}{2}\, K_{dd}\, \mathsf{H}_{zz} K\big(f_{\theta_0}(\lambda), \lambda, \theta_0\big) = \big(f_{\theta_0}(\lambda)^{-1}\big)' \otimes f_{\theta_0}(\lambda)^{-1},$$

which is—up to a factor 1/2—the same as $\Gamma_\psi(\lambda)$ in (15). Thus we obtain for $\mu(K)$ and $\sigma^2(K)$ similar expressions as in Example 4.5.

5. Asymptotic global and local power

In this section, we evaluate the asymptotic behaviour of the nonparametric test discussed in this paper for fixed alternatives and under a class of local alternatives. As in the previous section, we consider the general test problem (16) and the corresponding test statistic

$$S_T(K) = \int_\Pi K\big(\hat{f}(\lambda), \lambda, \hat{\theta}\big)\, d\lambda.$$

Let $Q_T(K)$ be the standardized version of $S_T(K)$. We first investigate the power of the test based on $Q_T(K)$ under fixed alternatives.

Theorem 5.1. *Suppose that Assumptions* 3.1, 3.3, 4.2, *and* 4.7 *hold. Then if the null hypothesis is false*,

$$\frac{1}{B_T^{1/2}\, T}\, Q_T(K) \xrightarrow{P} \frac{1}{\sigma(K)} \int_\Pi K\big(f(\lambda), \lambda, \theta_0\big)\, d\lambda \tag{21}$$

as T tends to infinity.

Proof. By the assumptions on the function K, we obtain

$$\left| S_T(K) - \int_\Pi K\big(f(\lambda), \lambda, \theta_0\big)\, d\lambda \right| \leqslant C \left[\max_{\lambda\in\Pi} \big\| \hat{f}(\lambda) - f(\lambda) \big\| + \big\| \hat{\theta} - \theta_0 \big\| \right],$$

whenever $\max_{\lambda\in\Pi} \big\| \hat{f}(\lambda) - f(\lambda) \big\| < \delta$ and $\|\hat{\theta} - \theta_0\| < \delta$. The result follows by similar arguments as in the proof of Lemma 3.4. □

We note that for any spectral density matrix $f(\lambda)$ under the alternative H_a the integrated discrepancy $K\big(f(\lambda), \lambda, \theta_0\big)$ on the right-hand side in (21) is nonzero since $K\big(f(\lambda), \lambda, \theta_0\big)$ being a continuous function of λ differs from zero on a set of strictly positive measure. Thus the above theorem shows that the test based on $Q_T(K)$ is consistent, that is, under the alternative H_a the null hypothesis is rejected asymptotically with probability one. Furthermore, the power of the test is monotonically increasing in the integrated discrepancy, which measures the total deviation from the null hypothesis. This means that, although the proposed test is an omnibus test that is able to detect any alternative for large enough sample size, it may be designed to focus on certain directions of departures from the null hypothesis by appropriate choice of the discrepancy K.

Next, we investigate the asymptotic properties of $Q_T(K)$ for sequences of local alternatives. More precisely, we consider sequences of processes $\{X^{(T)}(t)\}$, $T \in \mathbb{N}$, with spectral matrices

$$f^{(T)}(\lambda) = f(\lambda) + c_T\, g(\lambda), \tag{22}$$

where $(c_T)_{T\in\mathbb{N}}$ is a real-valued sequence such that $c_T \to 0$ as $T \to \infty$. We impose the following conditions on the processes $\{X^{(T)}(t)\}$.

Assumption 5.2. Let $\{X^{(T)}(t)\}$, $T \in \mathbb{N}$, be a sequence of stationary processes with mean zero and spectral matrices (22).

(i) The matrices $f(\lambda)$, $\lambda \in \Pi$, are nonnegative definite and Hermitian. Furthermore, $f(\lambda)$ and $g(\lambda)$ are twice continuously differentiable in λ, and $f(\lambda)$ satisfies the null hypothesis H_0 in (16) for $\theta_0 \in \Theta$.

(ii) The parameters $\theta^{(T)} \in \Theta$, $T \in \mathbb{N}$, associated with the processes $\{X^{(T)}(t)\}$ satisfy $\|\theta^{(T)} - \theta_0\| = O(T^{-1/2})$.

(iii) Let $c^{(T)}_{a_1,\dots,a_k}(u_1,\dots,u_{k-1})$ be the joint cumulant of $X^{(T)}_{a_1}(u_1),\dots,X^{(T)}_{a_{k-1}}(u_{k-1}),\ X^{(T)}_{a_k}(0)$. For all $k \geqslant 2$ there exists a constant $C_k > 0$ not depending on T such that

$$\sum_{u_1,\dots,u_{k-1}\in\mathbb{Z}} \bigl(1+|u_j|^2\bigr)|c^{(T)}_{a_1,\dots,a_k}(u_1,\dots,u_{k-1})| < C_k$$

for all $j = 1,\dots,k-1$, $a_1,\dots,a_k = 1,\dots,d$, and $T \in \mathbb{N}$.

The last assumption implies that the cumulant spectra of the processes $\{X^{(T)}(t)\}$ are uniformly bounded in $T \in \mathbb{N}$. Thus we have

$$\mathrm{cum}\{d^{(T)}_{a_1}(\lambda_1), d^{(T)}_{a_2}(\lambda_2)\} = 2\pi\,\mathrm{H}^{(T)}_2(\lambda_1+\lambda_2)\, f^{(T)}_{a_1a_2}(\lambda_1) + O(1)$$

uniformly in T while for all cumulants of higher order we obtain

$$|\,\mathrm{cum}\{d^{(T)}_{a_1}(\lambda_1),\dots,d^{(T)}_{a_k}(\lambda_k)\}| \leqslant C\, L^{(T)}(\lambda_1+\dots+\lambda_k)$$

with a constant C independent of T.

Lemma 5.3. *Suppose that Assumptions* 3.3, 4.2, 4.7, *and* 5.2 *hold. Then for* $c_T = B_T^{-1/4}\,T^{-1/2}$, *$S_T(K)$ can be approximated by*

$$S_T(K) = c_T^2 \int_\Pi \bigl\|\mathrm{vec}\bigl(g(\lambda)\bigr)\bigr\|^2_{\Gamma_K(\lambda)}\, d\lambda + \int_\Pi \bigl\|\mathrm{vec}\bigl(\hat f(\lambda) - f^{(T)}(\lambda)\bigr)\bigr\|^2_{\Gamma_K(\lambda)}\, d\lambda + o_P\bigl((B_T^{1/2}\,T)^{-1}\bigr),$$

where $\Gamma_K(\lambda)$ is defined by (18).

Proof. First, we note that the expansion of $K\bigl(\hat f(\lambda),\lambda,\hat\theta\bigr)$ about $Z = f(\lambda)$ and $\theta = \theta_0$ in the proof of Lemma 4.8 remains valid in the case of local alternatives. Thus we have

$$K\bigl(\hat f(\lambda),\lambda,\hat\theta\bigr) = \bigl\|\mathrm{vec}\bigl(\hat f(\lambda) - f(\lambda)\bigr)\bigr\|^2_{\Gamma_K(\lambda)} + R(\lambda),$$

with $\int_\Pi |R(\lambda)|\,d\lambda = o_P\bigl((B_T^{1/2}T)^{-1}\bigr)$, whenever $\max_{\lambda\in\Pi}\|\hat f(\lambda) - f(\lambda)\| \leqslant \delta$ and $\|\hat\theta - \theta_0\| \leqslant \delta$. Since

$$\max_{\lambda\in\Pi}\|\hat f(\lambda) - f(\lambda)\| \leqslant \max_{\lambda\in\Pi}\|\hat f(\lambda) - f^{(T)}(\lambda)\| + c_T \max_{\lambda\in\Pi}\|g(\lambda)\| = o_P(1),$$

a similar argument as in the proof of Lemma 3.4 shows that

$$\begin{aligned} S_T(K) &= \int_\Pi \big\|\mathrm{vec}\big(\hat{f}(\lambda) - f(\lambda)\big)\big\|^2_{\Gamma_K(\lambda)}\, d\lambda + o_P\big((B_T^{1/2}T)^{-1}\big) \\ &= \int_\Pi \big\|\mathrm{vec}\big(\hat{f}(\lambda) - f^{(T)}(\lambda)\big)\big\|^2_{\Gamma_K(\lambda)}\, d\lambda + c_T^2 \int_\Pi \big\|\mathrm{vec}\big(g(\lambda)\big)\big\|^2_{\Gamma_K(\lambda)}\, d\lambda \\ &\quad + 2\, c_T \int_\Pi \mathrm{vec}\big(g(\lambda)\big)^* \Gamma_K(\lambda)\, \mathrm{vec}\big(\hat{f}(\lambda) - f^{(T)}(\lambda)\big)\, d\lambda + o_P\big((B_T^{1/2}T)^{-1}\big). \end{aligned}$$

Here, the integral in the third term on the right-hand side is of order $O_P(T^{-1/2})$; hence the whole term also becomes $o_P\big((B_T^{1/2}T)^{-1}\big)$, which concludes the proof. □

The above lemma now allows us to establish the following result on the asymptotic distribution of the test statistic under sequences of local alternatives that converge to the null hypothesis with rate $c_T = B_T^{-1/4}\, T^{-1/2}$.

Theorem 5.4. *Suppose that Assumptions* 3.3, 4.2, 4.7, *and* 5.2 *hold. Then for* $c_T = B_T^{-1/4}\, T^{-1/2}$

$$B_T^{1/2}\, T\, S_T(K) - B_T^{-1/2}\, \mu(K) \xrightarrow{\mathcal{D}} \mathcal{N}(\nu(K), \sigma^2(K))$$

with

$$\nu(K) = \int_\Pi \big\|\mathrm{vec}\big(g(\lambda)\big)\big\|^2_{\Gamma_K(\lambda)}\, d\lambda,$$

and $\mu(K)$, $\sigma^2(K)$, *and* $\Gamma_K(\lambda)$ *defined as in Lemma* 4.8 *and Theorem* 4.9.

Proof. Let

$$\Psi_T^*\big(\Gamma_K\big) = \int_\Pi \big\|\mathrm{vec}\big(\hat{f}(\lambda) - f^{(T)}(\lambda)\big)\big\|^2_{\Gamma_K(\lambda)}\, d\lambda.$$

Under the conditions stated in Assumption 5.2, the bounds for the cumulants of $\Psi_T(\Gamma_K)$ derived in Lemmas B.4, B.5, and B.7 also hold for $\Psi_T^*(\Gamma_K)$ with constants independent of *T*. Thus the standardized statistic satisfies

$$\frac{B_T^{1/2}\, T\, \Psi_T^*\big(\Gamma_K\big) - B_T^{-1/2}\, \mu_T(K)}{\sigma_T(K)} \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1),$$

where $\mu_T(K)$ and $\sigma_T^2(K)$ are given by (6) and (7) with $f^{(T)}(\lambda)$ substituted for $f(\lambda)$. By a Taylor expansion of $\mu_T(K)$ and $\sigma_T^2(K)$ about $f(\lambda)$, we obtain bias $\mu(K)$ and variance $\sigma^2(K)$ with extra terms of lower order. Thus $\Psi_T^*(\Gamma_K)$ has the same asymptotic normal distribution as $\Psi_T(\Gamma_K)$ under the null hypothesis. The result follows now from Lemma 5.3. □

From the above theorem it follows that for the sequence of processes $\{X^{(T)}(t)\}$ the standardized test statistic

$$Q_T(K) = \frac{T\, S_T(K) - \mu(K)/B_T}{\sqrt{\sigma^2(K)/B_T}}$$

is asymptotically normally distributed with mean $v(K)/\sigma(K)$ and variance 1. Thus the ability of the test $Q_T(K)$ to detect the sequence of local alternatives can be assessed similarly as in Taniguchi et al. [62] by the efficacy of $Q_T(K)$,

$$\mathrm{eff}\big(Q_T(K)\big) = \lim_{T\to\infty} \mathbb{E}_{f^{(T)}}\big(Q_T(K)\big) = \frac{v(K)}{\sqrt{\sigma^2(K)}}.$$

We note that with $v(K)$ and $\sigma^2(K)$ the efficacy of $Q_T(K)$ also depends only on the second-order spectrum $f(\lambda)$, that is, the asymptotic power of the test is the same for Gaussian and non-Gaussian processes. Taniguchi et al. [62] called such tests "non-Gaussian robust".

The relative asymptotic performance of different tests can be evaluated by Pitman's asymptotic relative efficiency (ARE) [52, Chapter 7]. More precisely, let Q_T^* be another test for the test problem (16) and suppose that $B_T = O(T^{-\delta})$ with $\frac{2}{9} < \delta < \frac{1}{2}$. Then the ARE of $Q_T(K)$ with respect to Q_T^* is given by

$$\mathrm{ARE_P}\big(Q_T(K), Q_T^*\big) = \left(\frac{\mathrm{eff}\big(Q_T(K)\big)}{\mathrm{eff}\big(Q_T^*\big)}\right)^{\frac{2}{2-\delta}}$$

and compares the asymptotic power of the two tests $\big(Q_T(K)\big)$ and Q_T^* to detect local alternatives of the form (22).

Example 5.5. In order to test for a relationship between two time series $\{X_a(t)\}$ and $\{X_b(t)\}$, we consider the two statistics

$$S_T^{(1)} = \frac{1}{2\pi}\int_\Pi \big|\hat{R}_{ab}(\lambda)\big|^2\, d\lambda,$$

and

$$S_T^{(2)} = \frac{1}{\lambda_2-\lambda_1}\int_{\lambda_1}^{\lambda_2} \big|\hat{R}_{ab}(\lambda)\big|^2\, d\lambda.$$

Here, the latter statistic is typically used if one is particularly interested in dependencies over the frequency range $[\lambda_1, \lambda_2]$. For assessment of the power to detect deviations from the null hypothesis of noncorrelation in the frequency range $[\lambda_1, \lambda_2]$, suppose that $g(\lambda)$ in (22) vanishes outside $[\lambda_1, \lambda_2]$ for $\lambda \geqslant 0$. Then it follows from Example 3.8 that

$$\mathrm{ARE_P}\big(Q_T^{(1)}, Q_T^{(2)}\big) = \Big(\frac{\lambda_2-\lambda_1}{\pi}\Big)^{\frac{1}{2-\delta}} \leqslant 1.$$

In other words, the global test based on $Q_T^{(1)}$ is asymptotically less powerful in detecting dependencies restricted to the frequency range $[\lambda_1, \lambda_2]$ than the local test based on $Q_T^{(2)}$. This is the price we have to pay for achieving consistency against a larger class of alternatives.

Example 5.6. The asymptotic power of the test $Q_T(\psi, D)$ depends on the chosen kernel function w through the variance $\sigma^2(\psi, D)$. Let $Q_T^{(1)}$ and $Q_T^{(2)}$ be the tests obtained for kernel functions w_1 and w_2, respectively. Then the ARE of $Q_T^{(1)}$ with respect to $Q_T^{(2)}$ is given by

$$\mathrm{ARE_P}\big(Q_T^{(1)}, Q_T^{(2)}\big) = \left(\frac{C_{w_2,4}}{C_{w_1,4}}\right)^{\frac{1}{2-\delta}}. \tag{23}$$

For instance, if w_1 and w_2 are the quadratic spectral kernel and the Parzen window, respectively, we obtain $\mathrm{ARE}_{\mathrm{P}}\big(Q_T^{(1)}, Q_T^{(2)}\big) \approx 1.05$, that is, the quadratic spectral kernel leads to a slightly more powerful test than the Parzen window. The same ARE as in (23) has been obtained by Hong [39,40], who considered tests for serial correlation and tests for noncorrelation; notice, however, that the optimal kernel determined by Hong [39,40], namely the rectangular kernel (or Daniell window), is not admissible for our test as it does not satisfy Assumption 3.3.

Alternatively to Pitman's ARE, which is based on local power analysis, relative efficiencies may also be investigated under fixed alternatives on the basis of Bahadur's asymptotic slope criterion [2]. For details, we refer to Hong [39,40], who obtained similar results as in Theorems 5.1 and 5.4 in the special case of tests for serial correlation and tests for noncorrelation; see also Paparoditis [50] for the case of goodness-of-fit tests. We note that those results have been established under weaker conditions on the bandwidth B_T, namely $B_T = O(T^{-\delta})$ with $0 < \delta < 1$, which are partly possible since the bias introduced by smoothing vanishes in those special cases.

6. Simulations

The test statistics for the non- and semi-parametric hypotheses discussed in this paper are based on L^2 distances or discrepancies that are well approximated by L^2 distances in the neighbourhood of the null hypothesis. This poses a possible problem for the finite sample case as it is well known that even in the univariate case with i.i.d. data such L^2-type statistics may converge very slowly to their asymptotic limits (e.g., [36]). Therefore, in order to investigate the small sample properties of the proposed tests, we carried out a Monte Carlo study, in which the components of a multivariate process were tested for partial noncorrelation.

For the simulations, we considered a three-dimensional ARMA(2,1) process $\{X(t)\}$ given by

$$(\mathbb{1}_3 - \Phi_1 B - \Phi_2 B^2) X(t) = (\mathbb{1}_3 + \Psi B)\varepsilon(t), \tag{24}$$

where B is the backward shift operator and the innovations $\varepsilon(t)$ are independent and normally distributed with mean zero and covariance matrix $\Sigma = \mathbb{1}_3$. The coefficient matrices are given by

$$\Phi_1 = \begin{pmatrix} 1.3 & 0.2 & 0 \\ 0 & -1.4 & 0.1 \\ 0 & 0 & 0.7 \end{pmatrix}, \quad \Phi_2 = \begin{pmatrix} -0.9 & -0.1 & 0 \\ 0 & -0.9 & 0.1 \\ 0 & 0 & -0.9 \end{pmatrix}, \quad \Psi = \begin{pmatrix} 0.5 & 0 & 0 \\ 0 & 0.5 & 0 \\ 0 & 0 & -0.5 \end{pmatrix}.$$

The spectral densities and partial spectral coherences for this process are shown in Fig. 1. Here, only the partial spectral coherence between processes X_1 and X_3 given X_2 is identical to zero, that is, the two processes X_1 and X_3 are partially uncorrelated given the third process X_2, denoted as $X_1 \perp X_3 \mid X_2$, whereas the same is not true for the other two pairs of processes.

For this model, samples of size $T = 200, 500, 1000, 2000$, and 5000 were generated. Estimation of the spectral density matrix was based on the tapered periodogram using a 20% cosine taper and smoothing with a quadratic kernel. To study the effects of the smoothing bandwidth B_T, the estimates were computed for various bandwidths. As in Example 3.12, we used the integrated partial spectral coherence

$$S_T = \int_\Pi |\hat{R}_{ab|c}(\lambda)|^2 \, d\lambda$$

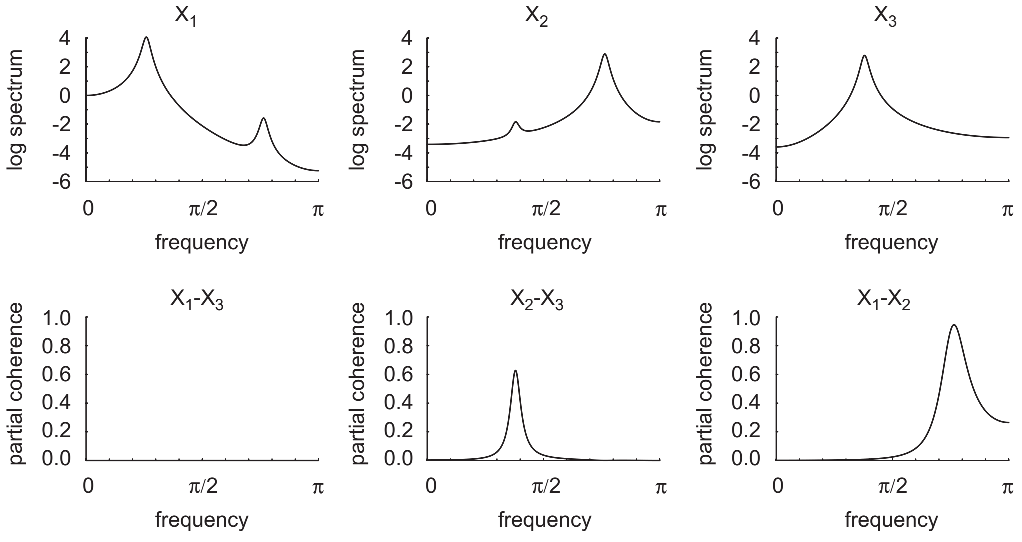

Fig. 1. Theoretical spectral densities and partial spectral coherencies for the model in (24).

to test whether two components X_a and X_b are partially uncorrelated given the third component X_c. The corresponding normalized test statistic is given by

$$Q_T = \frac{T\ S_T - 2\pi C_h C_{w,2}/B_T}{2\pi C_h \sqrt{2C_{w,4}/B_T}}$$

and does not depend on the unknown spectral density matrix of the process.

Table 1 reports the empirical rejection rates of the test at significance levels 5% and 10%, based on 10,000 replications. For the hypothesis $X_1 \perp X_3 \mid X_2$ the values give the empirical sizes of the test while for the other hypotheses $X_1 \perp X_2 \mid X_3$ and $X_2 \perp X_3 \mid X_1$ the values yield the empirical power of the test to detect the alternatives. The results show that the test in general performs well with empirical sizes that are reasonably close to the nominal ones for a broad range of bandwidths and high empirical power to detect the two alternatives for samples of size $T = 500$ and larger. A closer inspection of the results reveals some important features. Firstly, we note that the empirical size and power of the test depend on the bandwidth B_T: the use of smoother spectral estimates on the one hand increases the power of the test but on the other hand may lead to a violation of the nominal significance level, whereas too small bandwidths—and hence too rough spectral estimates—result in over-rejection of the null hypothesis. Secondly, for sample size $T = 200$, the test violates the nominal significance levels regardless of the choice of bandwidth. Additional simulations have shown that these size distortions become more serious for smaller sample sizes. These observations show that due to the nonlinearity of the hypothesis on the spectral density matrix the test requires spectral estimates that are close to the true spectral densities; this is in line with the theoretical results reflected in the condition in Assumption 3.3(iii) on the rate by which the bandwidth should converge to zero. Thirdly, comparing the results at significance levels 5% and 10%, we find that the empirical distribution of the test statistic deviates slightly from the asymptotic normal distribution even for large sample sizes. This indicates that convergence to the asymptotic limits is indeed slow.

Table 1
Rejection rates out of 10,000 replications for the three tests on partial noncorrelation at significance 5% and 10% for various bandwidths

T	B_T	$X_1 \perp X_3 \mid X_2$		$X_2 \perp X_3 \mid X_1$		$X_1 \perp X_2 \mid X_3$	
		5%	10%	5%	10%	5%	10%
200	0.08	4.18	9.81	21.83	37.09	95.91	98.69
200	0.10	4.44	10.42	28.39	44.33	98.94	99.70
200	0.15	4.62	9.56	41.05	55.72	99.94	100.00
200	0.20	4.07	8.11	49.02	62.12	100.00	100.00
200	0.25	4.43	8.93	59.20	70.03	100.00	100.00
200	0.30	3.75	7.82	63.39	73.25	100.00	100.00
200	0.35	3.73	7.81	67.91	76.92	100.00	100.00
500	0.06	7.75	15.38	74.12	84.74	100.00	100.00
500	0.08	5.91	11.64	80.58	88.71	100.00	100.00
500	0.10	6.70	12.38	88.03	93.35	100.00	100.00
500	0.12	5.43	10.36	91.01	94.98	100.00	100.00
500	0.15	5.47	10.01	94.49	97.04	100.00	100.00
500	0.20	5.01	9.19	97.14	98.53	100.00	100.00
500	0.25	4.93	8.60	98.30	99.07	100.00	100.00
500	0.30	4.68	8.30	98.93	99.37	100.00	100.00
1000	0.04	7.74	14.54	97.04	98.92	100.00	100.00
1000	0.06	6.69	12.36	99.41	99.78	100.00	100.00
1000	0.08	6.19	11.06	99.83	99.99	100.00	100.00
1000	0.10	5.82	10.50	99.98	100.00	100.00	100.00
1000	0.15	5.29	8.97	100.00	100.00	100.00	100.00
1000	0.20	4.93	8.69	100.00	100.00	100.00	100.00
1000	0.25	4.66	8.24	100.00	100.00	100.00	100.00
2000	0.04	6.92	13.15	100.00	100.00	100.00	100.00
2000	0.06	5.87	11.39	100.00	100.00	100.00	100.00
2000	0.08	5.15	10.29	100.00	100.00	100.00	100.00
2000	0.10	5.18	9.72	100.00	100.00	100.00	100.00
2000	0.15	4.88	8.94	100.00	100.00	100.00	100.00
2000	0.20	4.74	8.74	100.00	100.00	100.00	100.00
2000	0.25	4.68	8.54	100.00	100.00	100.00	100.00
5000	0.02	8.40	15.12	100.00	100.00	100.00	100.00
5000	0.04	6.22	11.17	100.00	100.00	100.00	100.00
5000	0.06	5.90	10.53	100.00	100.00	100.00	100.00
5000	0.08	5.33	9.93	100.00	100.00	100.00	100.00
5000	0.10	5.05	8.97	100.00	100.00	100.00	100.00
5000	0.15	4.76	8.75	100.00	100.00	100.00	100.00
5000	0.20	4.86	8.46	100.00	100.00	100.00	100.00

The simulation demonstrates that the selection of an appropriate bandwidth is an important step in the application of the proposed tests since the size and power of the test is sensitive to under- as well as to over-smoothing. Consequently, a data-driven method for choosing an optimal bandwidth is required. In the literature on nonparametric spectral density estimation, a number of criteria for bandwidth selection have been proposed; a partial overview and comparison is given, for example, in Fortin and Kuzmics [28]. In the following, we will consider two methods: a global version of the iterative procedure (ITP) suggested by Bühlmann [6], and a method by Lee [45] that combines plug-in and unbiased risk estimation (PURE) ideas. As both methods

Table 2
Rejection rates out of 10,000 replications for the three tests on partial noncorrelation at significance 5% and 10% with data-driven choice of bandwidth

Criterion	T	$X_1 \perp X_3 \mid X_2$		$X_2 \perp X_3 \mid X_1$		$X_1 \perp X_2 \mid X_3$	
		5%	10%	5%	10%	5%	10%
PURE	200	8.25	14.02	30.10	45.10	96.22	98.17
	500	8.54	14.68	80.49	88.47	99.90	99.93
	1000	8.04	14.02	98.46	99.22	100.00	100.00
	2000	7.67	13.55	99.96	99.99	100.00	100.00
	5000	7.55	12.70	100.00	100.00	100.00	100.00
	10000	6.66	11.88	100.00	100.00	100.00	100.00
	20000	5.89	10.75	100.00	100.00	100.00	100.00
ITP	200	2.49	4.42	75.71	81.76	100.00	100.00
	500	4.41	8.28	98.65	99.24	100.00	100.00
	1000	5.40	9.98	99.99	100.00	100.00	100.00
	2000	5.74	10.34	100.00	100.00	100.00	100.00
	5000	5.91	10.58	100.00	100.00	100.00	100.00
	10000	5.58	10.62	100.00	100.00	100.00	100.00
	20000	5.78	11.00	100.00	100.00	100.00	100.00

originally have been developed only for univariate time series, we used a simple modification by basically adding the risk functions for each component. We note that Robinson [55] suggested a multivariate approach for bandwidth selection based on cross validation.

Table 2 reports the empirical sizes and power of the test for partial noncorrelation when the bandwidth is selected by one of the two methods. For sample sizes $T \geqslant 500$, the ITP procedure leads to a good performance of the test both under the null hypothesis and the alternative although the null hypothesis is slightly over-rejected even for very large sample sizes. For sample size $T = 200$, on the other hand, the bandwidth selection by the ITP procedure seems to perform quite poorly since the test violates the nominal significance level severely. We note that for even smaller sample sizes, the ITP procedure failed to converge in most cases, which demonstrates that this method is not well suited for small sample sizes; similar results have been reported in Eichler [24]. In comparison, the PURE method leads to a far more conservative but—for small sample sizes—more stable behaviour of the test.

It should be noted that there are theoretical concerns about the use of common bandwidth selection methods such as the ITP procedure or the PURE method to our nonparametric test. Firstly, the asymptotic results in this paper have been derived only for nonstochastic sequences of bandwidths $(B_T)_{T\in\mathbb{N}}$. Secondly, the optimal bandwidth for spectral density estimation is of order $T^{-1/5}$ and thus does not satisfy the restrictions of Assumption 3.3. Therefore, the use of optimal bandwidths is expected to lead to an additional bias. Nevertheless, the above results indicate that in practice such methods work well—at least for moderate sample sizes—when used in conjunction with our test procedure.

7. Concluding remarks

We have presented a general approach for testing non- and semi-parametric hypotheses on the spectral density matrix of a vector-valued stationary process. The tests are based on integrated (squared) deviation measures that accumulate the pointwise departures of the estimated spectral density matrix—obtained by nonparametric kernel smoothers—from the null hypothesis. This approach covers not only many test problems that have been considered previously in the literature such as goodness-of-fit tests or tests for separability but also extends to new important problems; for instance, it allows considering hypotheses on partial spectral matrices and, in particular, yields a test for partial noncorrelation which is of interest in the context of graphical interaction models for multivariate time series.

The frequency domain approach for hypothesis testing presented in this paper offers a number of advantages. Firstly, the approach allows a completely nonparametric testing of hypotheses and thus avoids the problem of model misspecification, which typically arises for tests in the time domain. For example, time domain based tests for noncorrelation between two time series that have been proposed in the literature ([37,40,42,27,35,4,57]) all require fitting of a univariate or multivariate autoregressive (AR) or autoregressive moving average (ARMA) models. Secondly, the limiting distribution of the test statistic depends only on the second-order spectrum and thus can be easily evaluated. Moreover, in many important situation, the bias and the variance of the test statistic become also independent of the spectral densities. In contrast, the limiting distribution of the test statistics considered by Taniguchi et al. [62] typically involve the fourth-order cumulant spectrum. Thirdly, under any fixed alternative, the standardized test statistic tends to infinity at a rate faster than the parametric rate $\sqrt{T}$. This implies that the test has asymptotically high power for detecting any alternatives. On the other hand, tests based on the empirical spectral distribution function have nontrivial asymptotic power against local alternatives converging to the null hypothesis with the parametric rate $T^{-1/2}$ whereas our test requires a strictly larger rate $B_T^{-1/4}T^{-1/2}$, where B_T is the bandwidth of the kernel estimator for the spectral matrix. However, the type of hypotheses discussed in this paper are generally expressed in terms of non-linear functions of the spectral density and, thus, cannot be expressed directly in terms of the spectral distribution function.

In simulations, we have shown that the asymptotic results derived in this paper provide reasonable approximations for the distribution of the test statistics for medium and large sample sizes. Furthermore, the results indicated that one important issue in the practical application of the tests is the selection of the bandwidth for the kernel spectral estimates, which might have severe effects on the performance of the test. We have found that an appropriate choice seems the ITP by Bühlmann [6], which in our simulations led to empirical sizes close to the nominal ones and achieved high power under the alternatives.

Acknowledgments

The author is grateful for the comments and suggestions of two anonymous referees, which helped to improve upon earlier versions of the paper. Furthermore, it is acknowledged that most of the research presented here has been carried out while the author was at the Institute of Applied Mathematics, University of Heidelberg, Germany.

Appendix A. Data taper

The use of data tapers often improves the small sample properties of spectral estimates (e.g., [12,24]). For the discussion of the asymptotic properties of frequency domain statistics based on tapered data, we use the following function, which has been introduced by Dahlhaus [10]. Let $L^{(T)}: \mathbb{R} \to \mathbb{R}$ be the periodic extension (with period 2π) of

$$L^{(T)}(\lambda) = \begin{cases} T, & |\lambda| \leqslant 1/T, \\ \dfrac{1}{|\lambda|}, & 1/T < |\lambda| \leqslant \pi. \end{cases} \tag{A.1}$$

The properties of these functions are summarized by the following lemma. The proofs are straightforward and can be found in Dahlhaus [10,12].

Lemma A.1. *Let $L^{(T)}(\lambda)$ be defined as in* (A.1), *and let $\alpha, \beta, \gamma \in \mathbb{R}$ and $r, s \in \mathbb{N}$ with $r - s \geqslant 2$. We obtain with a constant C independent of T and S*

(i) $L^{(T)}(\alpha)$ *is monotonically increasing in $T \in \mathbb{R}^+$ and decreasing in $\alpha \in [0, \pi]$.*
(ii) $L^{(T)}(c\alpha) \leqslant c^{-1} L^{(T)}(\alpha)$ *for all $c \in (0, 1]$;*
(iii) $\displaystyle\int_\Pi L^{(T)}(\alpha)\, d\alpha \leqslant C \log(T)$;
(iv) $\displaystyle\int_\Pi L^{(T)}(\beta + \alpha) L^{(S)}(\gamma - \alpha)\, d\alpha \leqslant C \max\{\log(T), \log(S)\} L^{(\min\{T,S\})}(\beta + \gamma)$;
(v) $\displaystyle\int_\Pi |\alpha|^s L^{(T)}(\alpha)^r\, d\alpha \leqslant C T^{r-s-1}$;
(vi) $\displaystyle\int_\Pi L^{(T)}(\beta + \alpha)^r L^{(S)}(\gamma - \alpha)^r\, d\alpha \leqslant C \max\{T^{r-1}, S^{r-1}\} L^{(\min\{T,S\})}(\beta + \gamma)^r$.

Let $h^{(T)}$ be a data taper and let $H_k^{(T)}(\lambda)$ be its Fourier transform. If the taper function satisfies Assumption 3.3(i), then $H_k^{(T)}(\lambda)$ can be bounded by

$$\left|H_k^{(T)}(\lambda)\right| \leqslant C\, L^{(T)}(\lambda), \tag{A.2}$$

with a constant $C \in \mathbb{R}$ independent of T and λ. Similarly, we obtain by Assumption 3.3(ii) for the kernel function

$$w^{(T)}(\lambda) \leqslant \frac{C}{M_T} L^{(M_T)}(\lambda)^2,$$

where $M_T = 1/B_T$ for bandwidth B_T and C is again some constant independent of λ and T. For notational convenience, we use $M_T = 1/B_T$ throughout the appendix. Finally, let $\{\Phi_2^{(T)}\}_{T \in \mathbb{N}}$ be the sequence of function given by

$$\Phi_2^{(T)}(\lambda) = \frac{|H_2^{(T)}(\lambda)|^2}{2\pi H_4^{(T)}(0)}. \tag{A.3}$$

By Lemma A.1 and the upper bound in (A.2), it can be shown that $\{\Phi_2^{(T)}\}_{T \in \mathbb{N}}$ is a Dirac sequence (e.g., [10]).

Appendix B. A central limit theorem

In this appendix, we derive the limiting distribution of integrated squared differences between $\hat{f}(\lambda)$ and $f(\lambda)$ of the form

$$\Psi_T(\Gamma) = \int_\Pi \left\| \operatorname{vec}\big(\hat{f}(\lambda) - f(\lambda)\big) \right\|^2_{\Gamma(\lambda)} d\lambda,$$

where $\|x\|^2_{\Gamma(\lambda)} = x^*\Gamma(\lambda)x$ and $\Gamma(\lambda)$ is a Hermitian nonnegative definite matrix for all $\lambda \in \Pi$. Such differences play a central role in the discussion of the asymptotic properties of the class of nonlinear functionals that has been considered in this paper.

Assumption B.1. $\Gamma : \Pi \to \mathbb{C}^{d^2\times d^2}$ is Hermitian nonnegative function that is Lipschitz continuous except for possibly finitely many points $\lambda \in \Pi$.

In the following, we write $\Gamma(\lambda) = \big(\Gamma_{ij,kl}(\lambda)\big)_{i,j,k,l=1,\ldots,d}$, where the indices are such that

$$\left\| \operatorname{vec}\big(\hat{f}(\lambda) - f(\lambda)\big) \right\|^2_{\Gamma(\lambda)} = \sum_{i,j,k,l=1}^{d} \big(\hat{f}_{ji}(\lambda) - f_{ji}(\lambda)\big)\, \Gamma_{ij,kl}(\lambda)\, \big(\hat{f}_{kl}(\lambda) - f_{kl}(\lambda)\big),$$

that is, for $I = i+(j-1)\,d$ and $K = k+(l-1)\,d$ with $i, j, k, l \in \{1, \ldots, d\}$ the (I, K)th element of $\Gamma(\lambda)$ is given by $\Gamma_{ij,kl}(\lambda)$. Furthermore, we define the $d^2 \times d^2$ matrix $\tilde{\Gamma}(\lambda) = K_{dd}\, \Gamma(\lambda)\, K_{dd}$, where K_{dd} is the $d^2 \times d^2$ commutation matrix (e.g., [46]). Using the same indexation as for $\Gamma(\lambda)$, the (I, K)th entry of $\tilde{\Gamma}(\lambda)$ is given by

$$\tilde{\Gamma}_{ij,kl}(\lambda) = \Gamma_{ji,lk}(\lambda) \tag{B.1}$$

for $i, j, k, l = 1, \ldots, d$.

Theorem B.2. *Suppose that Assumptions* 3.1, 3.3, *and* B.1 *hold and let* $M_T = 1/B_T$. *Then*

$$M_T^{-1/2}\, T\, \Psi_T(\Gamma) - M_T^{1/2}\, \mu_\Psi \xrightarrow{\mathcal{D}} \mathcal{N}\big(0, \sigma^2_\Psi\big),$$

where

$$\mu_\Psi = C_h\, C_{w,2} \int_\Pi \operatorname{tr}\big[\Gamma(\lambda)\, \big(f(\lambda)' \otimes f(\lambda)\big)\big]\, d\lambda$$

and

$$\begin{aligned}\sigma^2_\Psi = 2\pi\, C_h^2\, C_{w,4} \int_\Pi \Big[&\operatorname{tr}\big\{\Gamma(\lambda)\big(f(\lambda)' \otimes f(\lambda)\big)\big(\Gamma(-\lambda) + \tilde{\Gamma}(\lambda)\big)'\big(f(\lambda)' \otimes f(\lambda)\big)\big\}\\ &+\operatorname{tr}\big\{\Gamma(\lambda)\big(f(\lambda)' \otimes f(\lambda)\big)\big(\Gamma(\lambda) + \tilde{\Gamma}(-\lambda)\big)\big(f(\lambda)' \otimes f(\lambda)\big)\big\}\Big]\, d\lambda.\end{aligned}$$

Remark B.3. We note that the property that $\Gamma(\lambda)$ is a nonnegative Hermitian matrix is not required for the proof of the result. Therefore, the theorem extends easily to statistics of the form

$$\begin{aligned}\Psi_T(\Gamma) &= \int_\Pi \operatorname{vec}\big(\hat{f}(\lambda) - f(\lambda)\big)'\, \Gamma(\lambda)\, \operatorname{vec}\big(\hat{f}(\lambda) - f(\lambda)\big)\, d\lambda\\ &= \int_\Pi \operatorname{vec}\big(\hat{f}(\lambda) - f(\lambda)\big)^*\, \big[K_{dd}\Gamma(\lambda)\big]\, \operatorname{vec}\big(\hat{f}(\lambda) - f(\lambda)\big)\, d\lambda,\end{aligned}$$

where the corresponding mean μ_Ψ and variance σ_Ψ^2 are obtained by replacing $\Gamma(\lambda)$ and $\tilde{\Gamma}(\lambda)$ by $K_{dd}\Gamma(\lambda)$ and $K_{dd}\tilde{\Gamma}(\lambda) = \Gamma(\lambda)K_{dd}$, respectively.

Proof of Theorem B.2. In Lemmas B.4, B.5, and B.7 we prove the convergence of the cumulants of first, second, and higher order of Ψ_T to the corresponding cumulants of the limit distribution. $\square$

Lemma B.4. *Suppose that the assumptions of Theorem* B.2 *hold. Then we have for the mean of* $\Psi_T(\Gamma)$

$$\mathbb{E}\big(\Psi_T(\Gamma)\big) = C_h\, C_{w,2}\, \frac{M_T}{T} \int_\Pi \mathrm{tr}\big[\Gamma(\lambda)\left(f(\lambda)' \otimes f(\lambda)\right)\big]\, d\lambda + o\left(\frac{M_T^{1/2}}{T}\right). \tag{B.2}$$

Proof. Since $\overline{f(\lambda)} = f(\lambda)'$, we can write $\Psi_T(\Gamma)$ as

$$\Psi_T(\Gamma) = \sum_{i,j,k,l=1}^{d} \int_\Pi \Gamma_{ji,kl}(\lambda)\, \big(\hat{f}_{ij}(\lambda) - f_{ij}(\lambda)\big)\big(\hat{f}_{kl}(\lambda) - f_{kl}(\lambda)\big)\, d\lambda.$$

From the product theorem for cumulants (cf. [5, Theorem 2.3.2]) and

$$\mathbb{E}\big(\hat{f}_{ij}(\lambda) - f_{ij}(\lambda)\big) = O(M_T^{-2}), \tag{B.3}$$

it follows that

$$\begin{aligned}
&\mathbb{E}\left[\int_\Pi \Gamma_{ji,kl}(\lambda)\, \big(\hat{f}_{ij}(\lambda) - f_{ij}(\lambda)\big)\big(\hat{f}_{kl}(\lambda) - f_{kl}(\lambda)\big)\, d\lambda\right] \\
&\quad = \int_\Pi \Gamma_{ji,kl}(\lambda)\ \mathrm{cum}\{\hat{f}_{ij}(\lambda),\, \hat{f}_{kl}(\lambda)\}\, d\lambda + O(M_T^{-4}) \\
&\quad = \frac{1}{(2\pi H_2^{(T)}(0))^2} \int_{\Pi^3} \Gamma_{ji,kl}(\lambda)\, w^{(T)}(\lambda-\alpha)\, w^{(T)}(\lambda-\beta) \\
&\qquad \times \mathrm{cum}\{d_i^{(T)}(\alpha)\, d_j^{(T)}(-\alpha),\, d_k^{(T)}(\beta)\, d_l^{(T)}(-\beta)\}\, d\alpha\, d\beta\, d\lambda + O(M_T^{-4}).
\end{aligned}$$

Applying again the product theorem, we find for the main term

$$\begin{aligned}
&\frac{1}{(2\pi H_2^{(T)}(0))^2} \int_{\Pi^3} \Gamma_{ji,kl}(\lambda)\, w^{(T)}(\lambda-\alpha)\, w^{(T)}(\lambda-\beta) \\
&\quad \times \Big[\mathrm{cum}\{d_i^{(T)}(\alpha),\, d_j^{(T)}(-\alpha),\, d_k^{(T)}(\beta),\, d_l^{(T)}(-\beta)\} \\
&\quad + \mathrm{cum}\{d_i^{(T)}(\alpha),\, d_k^{(T)}(\beta)\}\, \mathrm{cum}\{d_j^{(T)}(-\alpha),\, d_l^{(T)}(-\beta)\} \\
&\quad + \mathrm{cum}\{d_i^{(T)}(\alpha),\, d_l^{(T)}(-\beta)\}\, \mathrm{cum}\{d_j^{(T)}(-\alpha),\, d_k^{(T)}(\beta)\}\Big]\, d\alpha\, d\beta\, d\lambda.
\end{aligned} \tag{B.4}$$

By Theorem 4.3.2 of Brillinger [5], we have

$$\begin{aligned}
&\mathrm{cum}\{d_{a_1}^{(T)}(\alpha_1), \ldots, d_{a_k}^{(T)}(\alpha_k)\} \\
&\quad = (2\pi)^{k-1} H_k^{(T)}(\alpha_1 + \cdots + \alpha_k)\, f_{a_1\ldots a_k}(\alpha_1, \ldots, \alpha_{k-1}) + O(1)
\end{aligned} \tag{B.5}$$

uniformly in $\alpha_1, \ldots, \alpha_k$ and T. Substituting into (B.4), we find that the first term is of order $O(T^{-1})$ while the second term can be bounded by

$$\frac{C}{T^2 M_T^2} \int_{\Pi^3} L^{(M_T)}(\lambda-\alpha)^2\, L^{(M_T)}(\lambda-\beta)^2\, L^{(T)}(\alpha+\beta)^2\, d\alpha\, d\beta\, d\lambda$$
$$\leqslant \frac{C}{T M_T^2} \int_{\Pi^2} L^{(M_T)}(\lambda+\alpha)^2 L^{(M_T)}(\lambda-\alpha)^2 d\alpha\, d\lambda = O\big(T^{-1}\big).$$

The last term in (B.4) can be rewritten as

$$\frac{2\pi H_4^{(T)}(0)}{H_2^{(T)}(0)^2} \int_{\Pi^3} \xi(\lambda,\alpha)\, w^{(T)}(\alpha)\, w^{(T)}(\alpha-\beta)\, \Phi_2^{(T)}(\beta)\, d\alpha\, d\beta\, d\lambda + O\big(T^{-2}\big)$$

with $\xi(\lambda,\alpha) = \Gamma_{ji,kl}(\lambda)\, f_{il}(\lambda-\alpha)\, f_{kj}(\lambda-\alpha)$. Noting that $w(\alpha)$ and $\xi(\lambda,\alpha)$ are piecewise Lipschitz continuous in α, we obtain

$$\int_{\Pi^3} |\xi(\lambda,\alpha)|\, |w^{(T)}(\alpha) w^{(T)}(\alpha+\beta) - w^{(T)}(\alpha)^2|\, \Phi_2^{(T)}(\beta)\, d\alpha\, d\beta\, d\lambda$$
$$\leqslant C M_T^2 \int_{\Pi} |\beta|\, \Phi_2^{(T)}(\beta)\, d\beta \leqslant \frac{C M_T^2}{T} \int_{\Pi} L^{(T)}(\beta)\, d\beta \leqslant \frac{C\, M_T^2\, \log(T)}{T}$$

and

$$\int_{\Pi^2} |\xi(\lambda,\alpha) - \xi(\lambda,0)|\, w^{(T)}(\alpha)^2\, d\alpha\, d\lambda \leqslant \frac{C}{M_T^2} \int_{\Pi} |\alpha|\, L^{(M_T)}(\alpha)^4\, d\alpha \leqslant C,$$

which, after summation over i, j, k, l, shows the convergence of the last term in (B.4) to the term on the right side in (B.2). $\square$

For the cumulants of second and higher order, we have

$$\operatorname{cum}_r\big\{\Psi_T(\Gamma)\big\} = \sum_{\substack{a_{11},a_{12},\ldots,a_{r1},a_{r2}=1\\ b_{11},b_{12},\ldots,b_{r1},b_{r2}=1}}^{d} \int_{\Pi^r} \prod_{j=1}^{r} \Gamma_{b_{j1}a_{j1},a_{j2}b_{j2}}(\lambda_j)$$
$$\times \operatorname{cum}\left\{ \prod_{i=1}^{2} \big(\hat f_{a_{ji}b_{ji}}(\lambda_j) - f_{a_{ji}b_{ji}}(\lambda_j)\big),\ j=1,\ldots,r \right\} d\lambda_1 \cdots d\lambda_r. \tag{B.6}$$

Using the product theorem for cumulants, we can write the cumulants as

$$\sum_{i.p.} \prod_{l=1}^{n} \operatorname{cum}\big\{\big(\hat f_{a_{ji}b_{ji}}(\lambda_j) - f_{a_{ji}b_{ji}}(\lambda_j)\big),\ (j,i)\in Q_l\big\}, \tag{B.7}$$

where the sum $\sum_{i.p.}$ extends over all indecomposable partitions $\{Q_1, \ldots, Q_n\}$ (e.g., [5]) of the table

$$\begin{matrix} (1,1) & (1,2) \\ \vdots & \vdots \\ (r,1) & (r,2). \end{matrix} \tag{B.8}$$

Suppose the partition $\{Q_1, \ldots, Q_n\}$ contains u sets, $Q_{n-u+1}, \ldots, Q_n$ say, with only one element. Then (B.7) can be written as

$$\sum_{i.p.} \prod_{l=1}^{n-u} \operatorname{cum}\big\{\hat{f}_{a_{ji}b_{ji}}(\lambda_j),\ (j,i) \in Q_l\big\} \prod_{l=n-u+1}^{n} \operatorname{cum}\big\{\hat{f}_{a_{j_l i_l} b_{j_l i_l}}(\lambda_{j_l}) - f_{a_{j_l i_l} b_{j_l i_l}}(\lambda_{j_l})\big\}.$$

Substituting (1) for $\hat{f}$, we obtain for the right-hand side in (B.6) by application of the product theorem and (B.5)

$$\begin{aligned}
&\sum_{\substack{a_{11},a_{12},\ldots,a_{r1},a_{r2}=1\\ b_{11},b_{12},\ldots,b_{r1},b_{r2}=1}}^{d} \sum_{i.p.} \sum_{i.p.*} \big(2\pi H_2^{(T)}(0)\big)^{u-2r} \int_{\Pi^{3r-u}} \prod_{j=1}^{r} \Gamma_{b_{j1}a_{j1},a_{j2}b_{j2}}(\lambda_j)\\
&\times \prod_{l=n-u+1}^{n} \operatorname{cum}\big\{\hat{f}_{a_{j_l i_l} b_{j_l i_l}}(\lambda_{j_l}) - f_{a_{j_l i_l} b_{j_l i_l}}(\lambda_{j_l})\big\} \prod_{l=1}^{n-u} \Bigg[\prod_{(j,i)\in Q_l} w^{(T)}(\lambda_j - \alpha_{ji})\\
&\times \prod_{k=1}^{m_l} (2\pi)^{p_{lk}-1} H_{p_{lk}}^{(T)}(\bar{\gamma}_{lk})\, f_{c_{lk,1},\ldots,c_{lk,p_{lk}}}(\gamma_{lk,1}, \ldots, \gamma_{lk,p_{lk}-1})\\
&\times \prod_{(j,i)\in Q_l} d\alpha_{ji}\Bigg]\, d\lambda_1 \cdots d\lambda_r + \text{lower order terms},
\end{aligned} \tag{B.9}$$

where, for each partition $\{Q_1, \ldots, Q_n\}$, the sum $\sum_{i.p.*}$ extends over all indecomposable partitions $\{P_{l1}, \ldots, P_{lm_l}\}$ of the tables

$$\begin{array}{cc}
a_{j_{l1}i_{l1}} & b_{j_{l1}i_{l1}}\\
\vdots & \vdots\\
a_{j_{lq_l}i_{lq_l}} & b_{j_{lq_l}i_{lq_l}}
\end{array} \tag{B.10}$$

for all sets $Q_l = \{(j_{l1}, i_{l1}), \ldots, (j_{lq_l}, i_{lq_l})\}$. Furthermore, for $P_{lk} = \{c_{lk,1}, \ldots, c_{lk,p_{lk}}\}$ we have set $\gamma_{lk,p} = \alpha_{ij}$ and $\gamma_{lk,p} = -\alpha_{ij}$ if $c_{lk,p} = a_{ij}$ and $c_{lk,p} = b_{ij}$, respectively, and $\bar{\gamma}_{lk} = \gamma_{lk,1} + \cdots + \gamma_{lk,p_{lk}}$.

Lemma B.5. *Suppose that the assumptions of Theorem* B.2 *hold. Then*

$$\begin{aligned}
&\lim_{T\to\infty} \frac{T^2}{M_T} \operatorname{var}\big(\Psi_T(\Gamma)\big)\\
&= 2\pi\, C_h^2\, C_{w,4} \int_\Pi \Big[\operatorname{tr}\big\{\Gamma(\lambda)\big(f(\lambda)' \otimes f(\lambda)\big)\big(\Gamma(-\lambda) + \tilde{\Gamma}(\lambda)\big)'\big(f(\lambda)' \otimes f(\lambda)\big)\big\}\\
&\quad + \operatorname{tr}\big\{\Gamma(\lambda)\big(f(\lambda)' \otimes f(\lambda)\big)\big(\Gamma(\lambda) + \tilde{\Gamma}(-\lambda)\big)\big(f(\lambda)' \otimes f(\lambda)\big)\big\}\Big]\, d\lambda.
\end{aligned}$$

Proof. We evaluate the terms for the different partitions in (B.9) separately. We start by considering $Q_l = \{(1, l), (2, \sigma_l)\}$ with $P_{l1} = \{a_{1l}, a_{2\sigma_l}\}$ and $P_{l2} = \{b_{1l}, b_{2\sigma_l}\}$ for a permutation (σ_1, σ_2)

of (1, 2). Then the corresponding term in (B.9) is given by

$$\frac{(2\pi)^2\, H_4^{(T)}(0)^2}{H_2^{(T)}(0)^4} \int_{\Pi^6} \prod_{j=1}^{2} \Big[\Gamma_{b_{j1}a_{j1},a_{j2}b_{j2}}(\lambda_j) w^{(T)}(\lambda_j - \alpha_{j1}) w^{(T)}(\lambda_j - \alpha_{j2}) \Big]$$

$$\times \prod_{l=1}^{2} \Big[\Phi_2^{(T)}(\alpha_{1l} + \alpha_{2\sigma_l}) f_{a_{1l}a_{2\sigma_l}}(\alpha_{1l}) f_{b_{2\sigma_l}b_{1l}}(\alpha_{1l}) \Big]\, d\alpha_{11} \cdots d\alpha_{22}\, d\lambda_1\, d\lambda_2. \tag{B.11}$$

Define

$$\tilde{\Phi}_4^{(T)}(\alpha) = w^{(T)}(\alpha_1) \cdots w^{(T)}(\alpha_4)\, \Phi_2^{(T)}(\alpha_5)\, \Phi_2^{(T)}(\alpha_1 + \alpha_2 - \alpha_3 - \alpha_4 + \alpha_5)$$

and $\phi_T = \int_{\Pi^5} \tilde{\Phi}_4^{(T)}(\alpha)\, d\alpha_1 \cdots d\alpha_5$. Then it can be shown that $2\pi\, \phi_T / M_T$ converges to $C_{w,4}$. Furthermore, $\Phi_4^{(T)}(\alpha) = \tilde{\Phi}_4^{(T)}(\alpha)/\phi_T$ is a Dirac sequence. Hence (B.11) converges to

$$2\pi\, C_h^2\, C_{w,4}\, \frac{M_T}{T^2} \int_{\Pi} \Gamma_{b_{11}a_{11},a_{12}b_{12}}(\lambda)\, \Gamma_{b_{21}a_{21},a_{22}b_{22}}(-\lambda)$$
$$\times f_{a_{11}a_{2\sigma_1}}(\lambda)\, f_{b_{2\sigma_1}b_{11}}(\lambda)\, f_{a_{12}a_{2\sigma_2}}(\lambda)\, f_{b_{2\sigma_2}b_{12}}(\lambda)\, d\lambda,$$

where we have used that $H_k^{(T)}(0)/T \to H_k$ as $T \to \infty$. For $(\sigma_1, \sigma_2) = (1, 2)$ summation over all indices a_{ij} and b_{ij} with $i, j \in \{1, 2\}$ yields for this term

$$2\pi\, C_h^2\, C_{w,4}\, \frac{M_T}{T^2} \int_{\Pi} \operatorname{tr}\big[\Gamma(\lambda) \big(f(\lambda)' \otimes f(\lambda) \big) \Gamma(-\lambda)' \big(f(\lambda)' \otimes f(\lambda) \big) \big]\, d\lambda,$$

whereas for $(\sigma_1, \sigma_2) = (2, 1)$ we obtain

$$2\pi\, C_h^2\, C_{w,4}\, \frac{M_T}{T^2} \int_{\Pi} \operatorname{tr}\big[\Gamma(\lambda) \big(f(\lambda)' \otimes f(\lambda) \big) \tilde{\Gamma}(-\lambda) \big(f(\lambda)' \otimes f(\lambda) \big) \big]\, d\lambda,$$

where $\tilde{\Gamma}(\lambda)$ is given by (B.1). Next, we consider for the same partition $\{Q_1, Q_2\}$ the partitions defined by $P_{l1} = \{a_{1l}, b_{2\sigma_l}\}$ and $P_{l2} = \{b_{1l}, a_{2\sigma_l}\}$, for which the corresponding term in (B.9) is given by

$$\frac{(2\pi)^2\, H_4^{(T)}(0)^2}{H_2^{(T)}(0)^4} \int_{\Pi^6} \prod_{j=1}^{2} \Big[\Gamma_{b_{j1}a_{j1},a_{j2}b_{j2}}(\lambda_j) w^{(T)}(\lambda_j - \alpha_{j1}) w^{(T)}(\lambda_j - \alpha_{j2}) \Big]$$

$$\times \prod_{l=1}^{2} \Big[\Phi_2^{(T)}(\alpha_{1l} + \alpha_{2\sigma_l}) f_{a_{1l}a_{2\sigma_l}}(\alpha_{1l}) f_{b_{2\sigma_l}b_{1l}}(\alpha_{1l}) \Big]\, d\alpha_{11} \cdots d\alpha_{22}\, d\lambda_1\, d\lambda_2.$$

By a similar argument as above, it can be shown that this converges to

$$2\pi\, C_h^2\, C_{w,4}\, \frac{M_T}{T^2} \int_{\Pi} \Gamma_{b_{11}a_{11},a_{12}b_{12}}(\lambda)\, \Gamma_{b_{21}a_{21},a_{22}b_{22}}(\lambda)$$
$$\times f_{a_{11}b_{2\sigma_1}}(\lambda)\, f_{a_{2\sigma_1}b_{11}}(\lambda)\, f_{a_{12}b_{2\sigma_2}}(\lambda)\, f_{a_{2\sigma_2}b_{12}}(\lambda)\, d\lambda.$$

For $(\sigma_1, \sigma_2) = (1, 2)$ this yields

$$2\pi\, C_h^2\, C_{w,4}\, \frac{M_T}{T^2} \int_\Pi \mathrm{tr}\big[\Gamma(\lambda)\big(f(\lambda)' \otimes f(\lambda)\big)\Gamma(\lambda)'\big(f(\lambda)' \otimes f(\lambda)\big)\big]\, d\lambda,$$

while for $(\sigma_1, \sigma_2) = (2, 1)$ we obtain

$$2\pi\, C_h^2\, C_{w,4}\, \frac{M_T}{T^2} \int_\Pi \mathrm{tr}\big[\Gamma(\lambda)\big(f(\lambda)' \otimes f(\lambda)\big)\Gamma(\lambda)\big(f(\lambda)' \otimes f(\lambda)\big)\big]\, d\lambda.$$

Combining these results, we obtain the variance stated in the lemma.

To finish the proof, we need to show that for all other partitions the corresponding term in (B.9) is of lower order. First, let $Q_l = \{l, \sigma_l\}$ and $P_{11} = \{a_{11}, a_{2\sigma_1}\}$, $P_{12} = \{b_{11}, b_{2\sigma_1}\}$, $P_{21} = \{a_{12}, b_{2\sigma_2}\}$, $P_{22} = \{b_{12}, a_{2\sigma_2}\}$. By the bounds for $H_k^{(T)}(\lambda)$ and $w^{(T)}(\lambda)$, the corresponding term is dominated by

$$\frac{C}{T^4\, M_T^4} \int_{\Pi^4} \prod_{j=1}^{2} \Big[L^{(M_T)}(\lambda_j - \alpha_{j1})^2\, L^{(M_T)}(\lambda_j - \alpha_{j2})^2\Big]$$
$$\times L^{(T)}(\alpha_{11} + \alpha_{2\sigma_1})^2\, L^{(T)}(\alpha_{12} - \alpha_{2\sigma_2})^2\, d\alpha_{11} \cdots d\alpha_{22}\, d\lambda_1\, d\lambda_2.$$

Repeated application of Lemma A.1(v) and (vi) shows that this is of order $O(T^{-2})$. All remaining partitions with $u = 0$ and $n = 2$ can be bounded similarly.

For all other partitions, we note that there is only one set Q_1 with $q_1 > 1$. Therefore with (B.3), the terms are bounded by

$$\frac{C}{T^{4-u}\, M_T^{4+u}} \int_{\Pi^{6-u}} \prod_{(j,i)\in Q_1} L^{(M_T)}(\lambda_j - \alpha_{ji})^2 \prod_{k=1}^{m} L^{(T)}(\bar{\gamma}_{1k}) \prod_{(j,i)\in Q_1} d\alpha_{ji}\, d\lambda_1\, d\lambda_2.$$

Since the partitions are indecomposable, the set Q_1 covers all rows in table (B.8). Thus integrating over λ_1 and λ_2 and using $L^{(M_T)}(\lambda) \leqslant M_T^2$, we obtain

$$\frac{C}{T^{4-u}\, M_T^{3u-2}} \int_{\Pi^{4-u}} \prod_{k=1}^{m} L^{(T)}(\bar{\gamma}_{1k}) \prod_{(j,i)\in Q_1} d\alpha_{ji} \leqslant \frac{C\, \log(T)^\rho}{T^{3-u}\, M_T^{3u-2}},$$

since the partition $\{P_{11}, \ldots, P_{1m}\}$ of table (B.10) is indecomposable. By Assumption 3.3(iii) this is of order $o(M_T/T^2)$ which concludes the proof. □

For the result on the cumulants of higher order, we need the following lemma, which has been proved in Eichler [24, Lemma 2].

Lemma B.6. *Let $\{P_1, \ldots, P_m\}$ be an indecomposable partition of the table*

$$\begin{array}{cc} \alpha_1 & -\alpha_1 \\ \vdots & \vdots \\ \alpha_n & -\alpha_n. \end{array}$$

If $m = n$ then for any $n-2$ variables $\alpha_{i_1}, \ldots, \alpha_{i_{n-2}}$ we obtain

$$\int_{\Pi^{n-2}} \prod_{j=1}^{n} L^{(T)}(\bar{\gamma}_j)\, d\alpha_{i_1} \cdots d\alpha_{i_{n-2}} \leqslant C\, L^{(T)}(\alpha_{i_{n-1}} \pm \alpha_{i_n})^2 \log(T)^{n-2}.$$

If $m < n$ then there exist $n-2$ variables $\alpha_{i_1}, \ldots, \alpha_{i_{n-2}}$ such that

$$\int_{\Pi^{n-2}} \prod_{j=1}^{m} L^{(T)}(\bar{\gamma}_j)\, d\alpha_{i_1} \cdots d\alpha_{i_{n-2}} \leqslant CT \log(T)^{n-2}.$$

Lemma B.7. *Suppose that the assumption of Theorem* B.2 *hold. Then for all $r \geqslant 3$*

$$\mathrm{cum}_r\{\Psi_T(\Gamma)\} = o\left(\frac{M_T^{r/2}}{T^r}\right).$$

Proof. From (B.9) and (B.3), it follows that $|\,\mathrm{cum}_r\{\Psi_T(\Gamma)\}|$ is dominated by

$$\sum_{i.p.} \sum_{i.p.^*} \frac{C}{T^{2r-u}\, M_T^{2r+u}} \int_{\Pi^{3k-u}} \prod_{j=1}^{r} L^{(M_T)}(\lambda_j - \alpha_{j1})^2 \prod_{j=1}^{r-u} L^{(M_T)}(\lambda_j - \alpha_{j2})^2$$
$$\times \prod_{l=1}^{n-u} \prod_{k=1}^{m_l} L^{(T)}(\bar{\gamma}_{lk})\, d\alpha_{11} \cdots d\alpha_{r1}\, d\alpha_{12} \cdots d\alpha_{(r-u)2}\, d\lambda_1 \cdots d\lambda_r. \tag{B.12}$$

It suffices to show the stated rate of convergence for fixed partitions $Q_1, \ldots, Q_n$ and $P_{l1}, \ldots, P_{lm_l}$ of tables (B.8) and (B.10), respectively.

Suppose there are v sets Q_l such that $|P_{lk}| = 2$ for all $k = 1, \ldots, m_l$. Because the partition $\{Q_l\}$ is indecomposable, we can select subsets $Q'_l = \{(j_{l1}, i_{l1}), (j_{l2}, i_{l2})\} \subseteq Q_l$ for $l = 1, \ldots, v$ such that $Q'_1, \ldots, Q'_v$ together with the complement of their union, Q_{CU} say, is an indecomposable partition of table (B.8). Similarly, for each of the sets $Q_{v+1}, \ldots Q_{n-u}$, we can choose two elements such that conditions of Lemma B.6 are satisfied. Then the $L^{(M_T)}(\lambda_j - \alpha_{ji})^2$ for the remaining $2r - 2n + u$ indices in table (B.8) can be bounded by M_T^2. Integrating over the corresponding variables α_{ji}, we thus obtain by Lemma B.6

$$C \log(T)^{\rho} \frac{M_T^{2r-4n+u}}{T^{2r-n+v}} \int_{\Pi^{r+2(n-u)}} \prod_{l=1}^{n-u} \left[L^{(M_T)}(\lambda_{j_{l1}} - \alpha_{j_{l1}i_{l1}})^2 L^{(M_T)}(\lambda_{j_{l2}} - \alpha_{j_{l2}i_{l2}})^2\right]$$
$$\times \prod_{l=1}^{v} L^{(T)}(\alpha_{j_{l1}i_{l1}} \pm \alpha_{j_{l2}i_{l2}})^2\, d\alpha_{j_{11}i_{11}} \cdots d\alpha_{j_{(n-u)2}i_{(n-u)2}}\, d\lambda_1 \cdots d\lambda_r$$

for some $\rho \geqslant 0$. Integrating first over $\alpha_{j_{(v+1)1}i_{(v+1)1}}, \ldots, \alpha_{j_{(n-u)2}i_{(s-u)2}}$ and then over the remaining variables α_{ji}, we have

$$C \log(T)^{\rho} \frac{M_T^{2(r-n)-v}}{T^{2r-n}\, M_T^{u}} \int_{\Pi^r} \prod_{l=1}^{v} L^{(M_T)}(\lambda_{j_{j2}} \mp \lambda_{j_{l1}})^2\, d\lambda_1 \cdots d\lambda_r.$$

Because of the indecomposability of the chosen partition $\{Q'_1, \ldots, Q'_v, Q_{\mathrm{CU}}\}$, integration over $\lambda_1, \ldots, \lambda_r$ yields an additional factor of order $O(M_T^{v+1})$. Thus for fixed partitions $\{Q_l\}$ and $\{P_{lk}\}$, the corresponding summand in (B.12) is bounded by

$$C\, \frac{M_T^{r/2}}{T^r} \left(\frac{M_T^2}{T}\right)^{r-n} \frac{M_T \log(T)^{\rho}}{M_T^{r/2+u}}. \tag{B.13}$$

If $n \leqslant r$, this is of order $o(M_T^{r/2}/T^r)$ since $M_T^2/T \to 0$. Otherwise, we have $n = r + 1$, which implies that $u \geqslant 2$. We first consider the case $u = 2$. Then all sets Q_l contain exactly two elements, which implies that $\rho = 0$ as application of Lemma B.6 is not required. Consequently, the term in (B.13) is of order $o(M_T^{r/2}/T^r)$ since $T/M_T^{9/2} \to 0$ as $T \to \infty$. Otherwise, if $u \geqslant 3$, the same conclusion follows directly from $T/M_T^5 \to 0$ and $\log(T)^\rho/M_T^{1/2} \to 0$. $\square$

For the special case of quadratic functionals of the form

$$\Psi_T = \int_\Pi \left|\mathrm{tr}\left[\psi(\lambda)'\left(\hat{f}(\lambda) - f(\lambda)\right)\right]\right|^2 d\lambda,$$

where $\psi : \Pi \to \mathbb{C}^{d\times d}$ is some matrix-valued function, we obtain the following result.

Corollary B.8. *Suppose that the assumptions of Theorem* B.2 *hold. Then*

$$M_T^{-1/2}\, T\, \Psi_T - M_T^{1/2}\, \mu_\Psi \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \sigma_\Psi^2\right),$$

where

$$\mu_\Psi = C_h\, C_{w,2} \int_\Pi \mathrm{tr}\left[\psi(\lambda)'\, f(\lambda)\, \overline{\psi(\lambda)}\, f(\lambda)\right] d\lambda$$

and

$$\sigma_\Psi^2 = 2\pi\, C_h^2\, C_{w,4} \int_\Pi \Big[\left|\mathrm{tr}\left\{\psi(\lambda)' f(\lambda)\overline{\psi(\lambda)} f(\lambda)\right\}\right|^2 + \left|\mathrm{tr}\left\{\psi(\lambda)' f(\lambda)\overline{\psi(-\lambda)'} f(\lambda)\right\}\right|^2$$
$$+ \left|\mathrm{tr}\left\{\psi(\lambda)' f(\lambda)\psi(-\lambda) f(\lambda)\right\}\right|^2 + \left|\mathrm{tr}\left\{\psi(\lambda)' f(\lambda)\psi(\lambda)' f(\lambda)\right\}\right|^2\Big]\, d\lambda.$$

Proof. Note that Ψ_T can be written as

$$\Psi_T = \int_\Pi \left|\mathrm{vec}\left(\psi(\lambda)\right)'\mathrm{vec}\left(\hat{f}(\lambda) - f(\lambda)\right)\right|^2 d\lambda = \Psi_T(\Gamma),$$

where $\Gamma(\lambda) = \mathrm{vec}\left(\overline{\psi(\lambda)}\right)\mathrm{vec}\left(\psi(\lambda)\right)'$. Thus the result follows directly from Theorem B.2 and the relation $\mathrm{tr}(ABCD) = \mathrm{vec}(A')'(D' \otimes B)\mathrm{vec}(C)$. $\square$

Remark B.9. Since the spectral matrix satisfies $f(-\lambda) = f(\lambda)'$, it seems plausible to consider weight functions Γ such that

$$\left\|\mathrm{vec}\left(\hat{f}(-\lambda) - f(-\lambda)\right)\right\|_{\Gamma(-\lambda)}^2 = \left\|\mathrm{vec}\left(\hat{f}(\lambda) - f(\lambda)\right)\right\|_{\Gamma(\lambda)}^2.$$

This implies that $\Gamma_{ij,kl}(-\lambda) = \Gamma_{ji,lk}(\lambda)$ and hence $\tilde{\Gamma}(\lambda) = \Gamma(-\lambda)$. For such weight functions Γ, the expression for the asymptotic variance σ_Ψ of $\Psi_T(\Gamma)$ in Theorem B.2 can be simplified to

$$2\pi\, C_h^2\, C_{w,4} \int_\Pi \mathrm{tr}\left\{\Gamma(\lambda)\left(f(\lambda)' \otimes f(\lambda)\right)\left(\Gamma(\lambda) + \Gamma(-\lambda)'\right)\left(f(\lambda)' \otimes f(\lambda)\right)\right\} d\lambda.$$

References

[1] T.W. Anderson, Goodness of fit tests for spectral distributions, Ann. Statist. 21 (1993) 830–847.
[2] R.R. Bahadur, Stochastic comparison of tests, Ann. Math. Statist. 31 (1960) 276–295.
[3] M.S. Bartlett, An Introduction to Stochastic Processes with Special Reference to Methods and Applications, second ed, Cambridge University Press, Cambridge, MA, 1966.

[4] C. Bouhaddioui, R. Roy, A generalized Portmanteau test for independence of two infinite-order vector autoregressive series, J. Time Ser. Anal. 27 (2006) 505–544.
[5] D.R. Brillinger, Time Series: Data Analysis and Theory, McGraw Hill, New York, 1981.
[6] P. Bühlmann, Locally adaptive lag-window spectral estimation, J. Time Ser. Anal. 17 (1996) 247–270.
[7] H. Chen, J.P. Romano, Bootstrap-assisted goodness-of-fit tests in the frequency domain, J. Time Ser. Anal. 20 (1997) 619–654.
[8] D.R. Cox, N. Wermuth, Multivariate Dependencies—Models, Analysis and Interpretation, Chapman & Hall, London, 1996.
[9] N. Cressie, Statistics for Spatial Data, Wiley, New York, 1991.
[10] R. Dahlhaus, Spectral analysis with tapered data, J. Time Ser. Anal. 4 (1983) 163–175.
[11] R. Dahlhaus, Small sample effects in time series analysis: a new asymptotic theory and a new estimate, Ann. Statist. 18 (1988) 808–841.
[12] R. Dahlhaus, Nonparametric high resolution spectral estimation, Probab. Theory Related Fields 85 (1990) 147–180.
[13] R. Dahlhaus, Graphical interaction models for multivariate time series, Metrika 51 (2000) 157–172.
[14] R. Dahlhaus, M. Eichler, J. Sandkühler, Identification of synaptic connections in neural ensembles by graphical models, J. Neurosci. Methods 77 (1997) 93–107.
[15] M.A. Delgado, J. Hidalgo, C. Velasco, Distribution free goodness-of-fit tests for linear processes, Ann. Statist. 33 (2005) 2568–2609.
[16] A.P. Dempster, Covariance selection, Biometrics 28 (1972) 157–175.
[17] P. Duchesne, Testing for serial correlation of unknown form in cointegrated time series models, Ann. Inst. Statist. Math. 57 (2005) 575–595.
[18] P. Duchesne, R. Roy, On consistent testing for serial correlation of unknown form in vector time series models, J. Multivariate Anal. 89 (2004) 148–180.
[19] W. Dunsmuir, A central limit theorem for parameter estimation in stationary vector time series and its application to models for a signal observed with noise, Ann. Statist. 7 (1979) 490–506.
[20] W. Dunsmuir, E.J. Hannan, Vector linear time series models, Adv. Appl. Prob. 8 (1976) 339–364.
[21] M. Eichler, Graphical modelling of multivariate time series, Technical Report, Universität Heidelberg, 2001. (`arXiv:math.ST/0610654`).
[22] M. Eichler, A graphical approach for evaluating effective connectivity in neural systems, Philos. Trans. Roy. Soc. B 360 (2005) 953–967.
[23] M. Eichler, Graphical modelling of dynamic relationships in multivariate time series, in: M. Winterhalder, B. Schelter, J. Timmer (Eds.), Handbook of Time Series Analysis, Wiley-VCH, 2006, pp. 335–372.
[24] M. Eichler, A frequency-domain based test for independence between stationary time series, Metrika 65 (2007) 133–157.
[25] M. Eichler, Granger causality and path diagrams for multivariate time series, J. Econometrics 137 (2007) 334–353.
[26] M. Eichler, R. Dahlhaus, J. Sandkühler, Partial correlation analysis for the identification of synaptic connections, Biol. Cybernet. 89 (2003) 289–302.
[27] K. El Himdi, R. Roy, Tests for noncorrelation of two multivariate ARMA time series, Canad. J. Statist. 25 (1997) 233–256.
[28] I. Fortin, C. Kuzmics, Optimal window width choice in spectral density estimation, J. Statist. Comput. Simul. 67 (2000) 109–131.
[29] R. Fried, V. Didelez, Decomposability and selection of graphical models for time series, Biometrika 90 (2003) 251–267.
[30] U. Gather, M. Imhoff, R. Fried, Graphical models for multivariate time series from intensive care monitoring, Statist. Med. 21 (2002) 2685–2701.
[31] W. Gersch, G.V. Goddard, Epileptic focus location: spectral analysis method, Science 169 (1970) 701–702.
[32] C.W.J. Granger, Investigating causal relations by econometric models and cross-spectral methods, Econometrica 37 (1969) 424–438.
[33] X. Guyon, Random fields on a network, modeling, statistics, and applications, Springer, New York, 1995.
[34] G. Hainz, R. Dahlhaus, Spectral domain bootstrap tests for stationary time series, Beiträge zur Statistik 61, Universität Heidelberg, 1999.
[35] M. Hallin, A. Saidi, Testing non-correlation and non-causality between multivariate ARMA time series, J. Time Ser. Anal. 26 (2005) 83–105.
[36] W. Härdle, E. Mammen, Comparing nonparametric versus parametric regression fits, Ann. Statist. 21 (1993) 1926–1947.
[37] L.D. Haugh, Checking the independence of two covariance-stationary time series: a univariate residual covariance approach, J. Amer. Statist. Assoc. 71 (1976) 378–385.

[38] B. Hellwig, S. Häußler, B. Schelter, M. Lauk, B. Guschlbaur, J. Timmer, C.H. Lücking, Tremor correlated cortical activity in essential tremor, Lancet 357 (2001) 519–523.
[39] Y. Hong, Consistent testing for serial correlation of unknown form, Econometrica 64 (1996) 837–864.
[40] Y. Hong, Testing for independence between two covariance stationary time series, Biometrika 83 (1996) 615–625.
[41] Y. Hosoya, M. Taniguchi, A central limit theorem for stationary processes and the parameter estimation of linear processes, Ann. Statist. 10 (1982) 132–153 Correction: 21 (1993) 1115–1117.
[42] P.D. Koch, S.S. Yang, A method for testing the independence of two time series that accounts for a potential pattern in the cross-correlation function, J. Amer. Statist. Assoc. 81 (1986) 533–544.
[43] S. Kullback, R.A. Leibler, On information and sufficiency, Ann. Math. Statist. 22 (1951) 79–86.
[44] P.D. Larsen, C.D. Lewis, G.L. Gebber, S. Zhong, Partial spectral analysis of caridac-related sympathetic nerve discharge, J. Neurophysiology 84 (2000) 1168–1179.
[45] T.C.M. Lee, A stabilized bandwidth selection method for kernel smoothing of the periodogram, Signal Process. 81 (2001) 419–430.
[46] H. Lütkepohl, Introduction to Multiple Time Series Analysis, Springer, New York, 1993.
[47] Y. Matsuda, Y. Yajima, On testing for separable correlations of multivariate time series, J. Time Ser. Anal. 25 (2004) 501–528.
[48] A. Moneta, P. Spirtes, Graph-based search procedure for vector autoregressive models, LEM Working Paper 2005/14, Sant'Anna School of Advanced Studies, Pisa, 2005.
[49] H.M. Müller, S. Weiss, P. Rappelsberger, EEG coherence analysis of auditory sentence processing, in: H. Witte, U. Zwiener, B. Schack, A. Doering (Eds.), Quantitative and Topological EEG and MEG Analysis, Durckhaus Mayer, Jena, 1997, pp. 403–405.
[50] E. Paparoditis, Spectral density based goodness-of-fit tests for time series models, Scand. J. Statist. 27 (2000) 143–176.
[51] E. Paparoditis, Testing the fit of a vector autoregressive moving average model, J. Time Ser. Anal. 26 (2005) 543–568.
[52] E.J.G. Pitman, Some Basic Theory for Statistical Inference, Chapman & Hall, London, 1979.
[53] M.B. Priestley, Spectral Analysis and Time Series, vol. 1, Academic Press, London, 1981.
[54] M. Reale, G. Tunnicliffe Wilson, Identification of vector AR models with recursive structural errors using conditional independence graphs, Statist. Methods Appl. 10 (2001) 49–65.
[55] P.M. Robinson, Automatic frequency domain inference on semiparametric and nonparametric models, Econometrica 59 (1991) 1329–1363.
[56] J.R. Rosenberg, D.M. Halliday, P. Breeze, B.A. Conway, Identification of patterns of neural connectivity—partial spectra, partial coherence, and neural interactions, J. Neurosci. Methods 83 (1998) 57–72.
[57] A. Saidi, Consistent testing for non-correlation of two cointegrated ARMA time series, Canad. J. Statist. 35 (2007) 169–188.
[58] R. Salvador, J. Suckling, C. Schwarzbauer, E. Bullmure, Undirected graphs of frequency-dependent functional connectivity in whole brain networks, Philos. Trans. Roy. Soc. B 360 (2005) 937–946.
[59] B. Schack, P. Rappelsberger, S. Weiss, E. Möller, Adaptive phase estimation and its application in EEG analysis of word processing, J. Neurosci. Methods 93 (1999) 49–59.
[60] M. Taniguchi, Y. Kakizawa, Asymptotic Theory of Statistical Inference for Time Series, Springer, New York, 2000.
[61] M. Taniguchi, M. Kondo, Non-parametric approach in time series analysis, J. Time Ser. Anal. 14 (1993) 397–408.
[62] M. Taniguchi, M.L. Puri, M. Kondo, Nonparametric approach for non-gaussian vector stationary processes, J. Multivariate Anal. 56 (1996) 259–283.
[63] J. Timmer, M. Lauk, B. Köster, B. Hellwig, S. Häußler, B. Guschlbauer, V. Radt, M. Eichler, G. Deuschl, C.H. Lücking, Cross-spectral analysis of tremor time series, Internat. J. Bifurcation and Chaos 10 (2000) 2595–2610.
[64] L. Timmermann, J. Gross, M. Dirks, J. Volkmann, H.-J. Freund, A. Schnitzler, The cerebral oscillatory network of parkinsonian resting tremor, Brain 126 (2003) 199–212.
[65] S. Weiss, H.M. Müller, P. Rappelsberger, Theta synchronisation predicts efficient memory encoding of concrete and abstract nouns, NeuroReport 11 (2000) 2357–2361.
[66] S. Weiss, P. Rappelsberger, EEG coherence within the 13–18 Hz band as a correlate of a distinct lexical organisation of concrete and abstract nouns in humans, Neurosci. Lett. 209 (1996) 17–20.
[67] S. Weiss, P. Rappelsberger, Long-range EEG synchronization during word encoding correlates with successful memory performance, Cognitive Brain Res. 9 (2000) 299–312.