TESTING FOR STATIONARITY OF FUNCTIONAL TIME SERIES IN THE FREQUENCY DOMAIN

BY ALEXANDER AUE[1] AND ANNE VAN DELFT[2]

[1]*Department of Statistics, University of California, aaue@ucdavis.edu*
[2]*Fakultät für Mathematik, Ruhr-Universität Bochum, Anne.vanDelft@rub.de*

Interest in functional time series has spiked in the recent past with papers covering both methodology and applications being published at a much increased pace. This article contributes to the research in this area by proposing a new stationarity test for functional time series based on frequency domain methods. The proposed test statistics is based on joint dimension reduction via functional principal components analysis across the spectral density operators at all Fourier frequencies, explicitly allowing for frequency-dependent levels of truncation to adapt to the dynamics of the underlying functional time series. The properties of the test are derived both under the null hypothesis of stationary functional time series and under the smooth alternative of locally stationary functional time series. The methodology is theoretically justified through asymptotic results. Evidence from simulation studies and an application to annual temperature curves suggests that the test works well in finite samples.

CONTENTS

1. Introduction. The aim of this paper is to provide a new stationarity test for functional time series based on frequency domain methods. Particular attention is given to taking into account alternatives allowing for smooth variation as a source of nonstationarity, even though nonsmooth alternatives are covered within the simulation study. Functional data analysis has seen an upsurge in research contributions for at least one decade. This is reflected in the growing number of monographs in the area. Readers interested in the current state of statistical inference procedures may consult Bosq (2000), Ferraty and Vieu (2010), Horváth and Kokoszka (2012), Hsing and Eubank (2015) and Ramsay and Silverman (2005).

Papers on functional time series have come into focus more recently and now constitute an active area of research. Hörmann and Kokoszka (2010) introduced a general weak dependence concept for stationary functional time series, while van Delft and Eichler (2018) provided a framework for locally stationary functional time series. Antoniadis and Sapatinas (2003), Aue, Dubart Norinho and Hörmann (2015) and Besse, Cardot and Stephenson (2000) constructed prediction methodology that may find application across many areas of science, economics and finance. With the exception of van Delft and Eichler (2018), the above contributions are concerned with procedures in the time domain. Complementing methodology in the frequency domain has been developed in parallel. One should mention Panaretos and Tavakoli (2013), who provided results concerning the Fourier analysis of time series in function spaces, and Hörmann, Kidziński and Hallin (2015) who addressed the problem of dimension reduction for functional time series using dynamic principal components.

The methodology proposed in this paper provides a new frequency domain inference procedure for functional time series. More precisely, tests for second-order stationarity are developed. In the univariate case such tests have a long history, going back at least to the seminal paper Priestley and Subba Rao (1969), who based their method on the evaluation of evolutionary spectra of a given time series. Other contributions building on this work include von Sachs and Neumann (2000), who used local periodograms and wavelet analysis, and Paparoditis (2009), whose test is based on comparing a local estimate of the spectral density to a global estimate. Dette, Preuss and Vetter (2011) and Preuß, Vetter and Dette (2013) developed methods to derive both a measure of and a test for stationarity in locally stationary time series, the latter authors basing their method on empirical process theory. In all papers interest is in smoothly varying alternatives. The same tests, however, also tend to have power against nonsmooth alternatives such as structural breaks or change points. A recent review discussing methodology for structural breaks in time series is Aue and Horváth (2013), while Aue, Rice and Sönmez (2018) is a recent contribution to structural breaks in functional time series.

The proposed test for second-order stationarity of functional time series seeks to exploit that the Discrete Fourier Transform (DFT) of a functional time series evaluated at distinct Fourier frequencies are asymptotically uncorrelated if and only if the series is second-order stationary. Therefore, the proposed method is related to the initial work of Dwivedi and Subba Rao (2011) who put forth similar tests in a univariate framework. Their method has since been generalized to multivariate time series in Jentsch and Subba Rao (2015) as well as to spatial and spatio-temporal data by Bandyopadhyay and Subba Rao (2017) and Bandyopadhyay,

Jentsch and Subba Rao (2017), respectively. A different version of functional stationarity tests, based on time domain methodology involving cumulative sum statistics (Aue and Horváth (2013)), was given in Horváth, Kokoszka and Rice (2014).

The intrinsic variation of a functional time series can be larger than any sample size, and standard results known from univariate and multivariate time series analysis do not directly apply. From a practical perspective this brings to the fore the question of how to compress this infinite-dimensional variation to finite dimension in a meaningful way, as there is a complex interplay between dynamics occurring across frequencies and the function space. This means that dimension reduction has to be done jointly across estimated spectral density operators at all Fourier frequencies, yet separately as the exact level of dimension reduction has to be decided per frequency. The proposed test statistics collect these different sets of projections, obtained via functional principal components analysis, into a quadratic form encapsulating the second-order dynamics. To derive the large-sample behavior of this statistic under both the null hypothesis of a stationary time series and the alternative of a locally stationary functional time series requires new, and perhaps independently interesting, results on distributional convergence of a cross-periodogram operator in the function space where verifying existence of the limit process and tightness are nontrivial tasks. The subsequent proofs of distributional convergence of the test statistics, which require taking into account the pecularities of fPCA estimators, are also complex and new. The main results are derived under the assumption that the curves are observed in their entirety, corresponding to a setting in which functions are sampled on a dense grid rather than a sparse grid. Differences for these two cases have been worked out in Li and Hsing (2010).

The remainder of the paper is organized as follows. Section 2 provides background, gives requisite notations, introduces properties of functional version of the DFT and gives intuition for the test. The exact form of the hypothesis test, model assumptions and the test statistics are introduced in Section 3. The large-sample behavior under the null hypothesis of second-order stationarity and the alternative of local stationarity is established in Sections 4. Empirical aspects are highlighted in Section 5. The proofs are technical and relegated to Appendices A–E. Several further auxiliary results are proved in the Supplementary Material (Aue and van Delft (2019)), henceforth referred to simply as the Online Supplement.

2. Notation and setup. A functional time series $(X_t : t \in \mathbb{Z})$ will be viewed as a sequence of random elements on a probability space $(\Omega, \mathcal{A}, P)$ with paths in a separable Hilbert space. Without loss of generality, we shall focus on processes taking values in $H_{\mathbb{R}} = L^2_{\mathbb{R}}([0, 1])$, the space of equivalence classes of real-valued, square integrable functions on the unit interval $[0, 1]$. Because the methodology introduced in this paper is based on a frequency domain approach, we shall make extensive use of the complex Hilbert space $H = L^2_{\mathbb{C}}([0, 1])$. We briefly introduce notation and relevant properties of this space and associated operators. The complex conjugate of $z \in \mathbb{C}$ is denoted by $\bar{z}$ and the imaginary number by i. For $f, g \in H$, the inner product and the induced L_2-norm on H are, respectively, given by

$$(1) \qquad \langle f, g \rangle = \int_0^1 f(\tau)\overline{g(\tau)} \, d\tau \quad \text{and} \quad \|f\|_2 = \sqrt{\langle f, f \rangle}.$$

Two elements of H are understood to be equal if their difference has vanishing L_2-norm. More generally, for measurable functions $g: [0, 1]^k \to \mathbb{C}$, the L^p-norm shall be denoted by $\|g\|_p$ and the supremum norm by $\|g\|_\infty = \sup_{\tau \in [0, 1]^k} |g(\tau)|$.

Next, some properties of linear operators on H are stated. Denote by $S_\infty(H)$ the Banach space of bounded linear operators $A: H \to H$ equipped with the operator norm $\|A\|_\infty = \sup_{\|g\|_2 \leq 1} \|Ag\|_2$. For all $f, g \in H$, the adjoint operator of A, denoted by $A^\dagger$, is

defined by $\langle Af, g \rangle = \langle f, A^\dagger g \rangle$ and the conjugate operator of A is given by $\overline{A}g = \overline{(A\overline{g})}$. An operator A is called self-adjoint if $\langle Af, g \rangle = \langle f, Ag \rangle$ for all $f, g \in H$ and nonnegative definite if $\langle Ag, g \rangle \geq 0$ for all $g \in H$. For $v \in H$, define the tensor product $f \otimes g \colon H \otimes H \to H$ as the bounded linear operator $(f \otimes g)v = \langle v, g \rangle f$. A compact operator A admits a *singular value decomposition*

$$A = \sum_{n=1}^{\infty} s_n(A) \psi_n \otimes \phi_n, \tag{2}$$

where $(s_n(A) \colon n \in \mathbb{N})$ are the *singular values* of A, and where $(\phi_n \colon n \in \mathbb{N})$ and $(\psi_n \colon n \in \mathbb{N})$ are orthonormal bases of H. The singular values are ordered to form a monotonically decreasing sequence of nonnegative numbers. A compact operator A is said to belong to the *Schatten p-class* $S_p(H)$ if and only if the sequence $s(A) = (s_n(A) \colon n \in \mathbb{N})$ of singular values of A belongs to the sequence space ℓ^p, so if and only if $\|A\|_p = (\sum_{n=1}^{\infty} s_n^p(A))^{1/p} < \infty$, where $\|A\|_p$ is referred to as the *Schatten p-norm*. Relevant here are $S_1(H)$, the space of trace-class operators, and particularly $S_2(H)$, the space of Hilbert–Schmidt operators. The latter is also a Hilbert space with inner product $\langle A, B \rangle_S = \sum_{i=1}^{\infty} \langle A\psi_i, B\psi_i \rangle$, where $A, B \in S_2(H)$ and $(\psi_n \colon n \in \mathbb{N})$ is an ONB of H. The mapping $\mathcal{T} \colon H \otimes H \to S_2(H)$ defined by the linear extension of $\mathcal{T}(f \otimes g) = f \otimes \overline{g}$ is an isometric isomorphism and defines a Hilbert–Schmidt operator with kernel in $H \times H$ given by $(f \otimes g)(\tau, \sigma) = f(\tau)\overline{g}(\sigma)$, $\tau, \sigma \in [0, 1]$. As a consequence, $A \in S_2(H)$ if and only if there exists $a \in H \times H$ such that $\|A\|_2 = \|a\|_2$. Further useful properties needed in the proofs of the various statements of this paper are relegated to Appendices A–E and the Online Supplement.

2.1. *Dependence structure on the function space.* Let $L^2_{\mathbb{C}}(\Omega)$ be the Hilbert space with elements satisfiying $\mathbb{E}[\|X\|_2^2] < \infty$ and denote by $\mathbb{E}[X]$ the mean function of X, where the expectation should be viewed in the sense of a Bochner integral. For $X, Y \in L^2_{\mathbb{C}}(\Omega)$, the covariance operator $\mathcal{C}_{X,Y} \colon H \otimes H \to H$ is defined as $\mathcal{C}_{X,Y} = \mathbb{E}[(X - \mathbb{E}[X]) \otimes (Y - \mathbb{E}[Y])]$ and belongs to $S_1(H)$. A functional time series $X = (X_t \colon t \in \mathbb{Z})$ is called strictly stationary if, for all finite sets of indices $J \subset \mathbb{Z}$, the joint distribution of $(X_{t+j} \colon j \in J)$ does not depend on $t \in \mathbb{Z}$. Similarly, X is weakly stationary if its first- and second-order moments exist and are invariant under translation in time. Without loss of generality, it is assumed throughout that $\mathbb{E}[X_t] = 0$ and that $X_t \in L^2_{\mathbb{R}}(\Omega)$ for all $t \in \mathbb{Z}$. The lag-h covariance operator between X_t and X_{t+h} is denoted by

$$\mathcal{C}_{t,h} = \mathbb{E}[X_{t+h} \otimes X_t]$$

which reduces to $\mathcal{C}_h = \mathbb{E}[X_h \otimes X_0]$ in case of weak stationarity. Note that this object is a nonnegative definite element of $S_1(H_{\mathbb{R}})$ for $h = 0$. The covariance operator $\mathcal{C}_h$ can be shown to form a Fourier pair with a nonnegative Hermitian element of $S_p(H)$. Provided sufficiently fast decay of the second-order structure, the *spectral density operator* $\mathcal{F}_\omega$ is well defined and given by the Fourier transform of $\mathcal{C}_h$,

$$\mathcal{F}_\omega = \frac{1}{2\pi} \sum_{h \in \mathbb{Z}} \mathcal{C}_h e^{-i\omega h}. \tag{3}$$

A sufficient condition for the existence of $\mathcal{F}_\omega$ in $S_p(H)$ is $\sum_{h \in \mathbb{Z}} \|\mathcal{C}_h\|_p < \infty$.

Higher-order dependence among the functional observations is defined through cumulant mixing conditions (Brillinger (1981), Brillinger and Rosenblatt (1967)). For this notion of higher-order cumulant tensors is required; see Appendix B for their definition and a discussion of their properties for nonstationary functional time series.

2.2. *The functional discrete Fourier transform.* The starting point of this paper is the following proposition that characterizes second-order stationary behavior of a functional time series in terms of a spectral representation. Its proof is in Appendix A.

PROPOSITION 2.1. *A zero-mean H-valued stochastic process $(X_t : t \in \mathbb{Z})$ admits the representation*

$$(4) \qquad X_t = \int_{-\pi}^{\pi} e^{it\omega} \, dZ_\omega \quad a.s.,$$

where $(Z_\omega : \omega \in (-\pi, \pi])$ is a right-continuous functional orthogonal-increment process if and only if it is weakly stationary.

If the process is not weakly stationary, then a representation in the frequency domain is not necessarily well defined and certainly not with respect to complex exponential basis functions. However, a time-dependent functional Cramér representation exists if the characteristics of the process are captured by a Bochner-measurable mapping that is an evolutionary operator-valued mapping in time direction (van Delft and Eichler (2018)). Assume that the functions $X_1, \ldots, X_T$ have been observed. If the process is weakly stationary, the *functional Discrete Fourier Transform* (fDFT) evaluated at frequency ω, given by

$$(5) \qquad D_\omega^{(T)} = \frac{1}{\sqrt{2\pi T}} \sum_{t=1}^{T} X_t e^{-i\omega t},$$

can be seen as an estimate of the increment process Z_ω and exists almost surely as an element of H. The functional time series itself can then be represented through the inverse fDFT as

$$(6) \qquad X_t = \sqrt{\frac{2\pi}{T}} \sum_{j=1}^{T} D_{\omega_j}^{(T)} e^{i\omega_j t}.$$

Under regularity conditions, a set of fDFTs evaluated at distinct frequencies yield asymptotically independent Gaussian random elements in H and, for fixed ω, one has $\mathrm{Var}(D_\omega^{(T)}) \to \mathcal{F}_\omega$ (Panaretos and Tavakoli (2013)). The fDFT sequence of a Hilbertian-valued stationary process is, in particular, asymptotically uncorrelated at the canonical frequencies $\omega_j = 2\pi j / T$. Consequently, provided the series is weakly stationary, for $j \neq j'$ or $j \neq T - j'$ we have $\||\mathrm{Cov}(D_{\omega_j}^{(T)}, D_{\omega_{j'}}^{(T)})\||_2 = O(1/T)$. In other words, the lag-h covariance operator of the fDFT converges in norm and hence in the weak operator topology to the zero operator as $T \to \infty$. Similar to the above, the reverse argument (uncorrelatedness of the functional DFT sequence implies weak stationarity) can be shown by means of the inverse fDFT. Using expression (6), the covariance operator $\mathcal{C}_{t,h}$ of X_{t+h} and X_t can be written in terms of the fDFT sequence as

$$\mathcal{C}_{t,h} = \frac{2\pi}{T} \sum_{j,j'=1}^{T} \mathbb{E}[D_{\omega_j}^{(T)} \otimes D_{\omega_{j'}}^{(T)}] e^{i\omega_j h} = \frac{2\pi}{T} \sum_{j=1}^{T} \mathbb{E}[D_{\omega_j}^{(T)} \otimes D_{\omega_j}^{(T)}] e^{i\omega_j h} = \mathcal{C}_h,$$

where the equality holds in an L^2-sense. This demonstrates that the autocovariance kernel of a second-order stationary functional time series is obtained and, hence, that an uncorrelated fDFT sequence implies second-order stationarity up to lag $T - 1$. The fDFT thus captures exactly the defining property of a weakly stationary process and provides a natural starting point for a test of stationarity. It is, however, a nontrivial task to construct a test statistic that optimally extracts the information contained in the infinite-dimensional process to finite dimensions. Not only can the dependence structure and the resulting dynamics of a functional time series be of a complicated nature (see Figure 1 and the example given in Section S8

of the Online Supplement), but the process will vary along both frequency and functional directions. To construct a powerful test, it is therefore crucial to understand how the fDFT's behave when weak stationarity is violated. In accordance with aforementioned time series literature, the theoretical behavior of the fDFT sequence under smooth alternatives is studied. These properties will then be exploited to verify large-sample results for a testing framework for functional stationarity.

3. The functional stationarity testing framework.

This section gives precise formulations of the hypotheses of interest, states the main assumptions of the paper and introduces the test statistics. Throughout, interest is in testing the null hypothesis

$$H_0 \colon (X_t \colon t \in \mathbb{Z}) \text{ is a weakly stationary functional time series}$$

versus the alternative

$$H_A \colon (X_t \colon t \in \mathbb{Z}) \text{ is a locally stationary functional time series,}$$

where locally stationary functional time series are defined as follows.

DEFINITION 1.   A stochastic process $(X_t \colon t \in \mathbb{Z})$ taking values in $H_{\mathbb{R}}$ is said to be *locally stationary* if:

(1) $X_t = X_t^{(T)}$ for $t = 1, \ldots, T$ and $T \in \mathbb{N}$; and
(2) for any rescaled time $u \in [0, 1]$, there is a strictly stationary process $(X_t^{(u)} \colon t \in \mathbb{Z})$ such that

$$\|X_t^{(T)} - X_t^{(u)}\|_2 \leq \left(\left| \frac{t}{T} - u \right| + \frac{1}{T} \right) P_{t,T}^{(u)} \quad \text{a.s.,}$$

where $P_{t,T}^{(u)}$ is a positive, real-valued triangular array of random variables such that, for some $\rho > 0$, $\mathbb{E}[|P_{t,T}^{(u)}|^\rho] < \infty$ for all t and T, uniformly in $u \in [0, 1]$.

Observe that the process constitutes a triangular array of functions under H_A. Inference methods are based on in-fill asymptotics as popularized in Dahlhaus (1997) for univariate time series. This asymptotic theory prescribes that the process is observed on an increasingly finer grid as T increases such that more observations become available at a local level. A rigorous statistical framework for locally stationary functional time series was recently provided in van Delft and Eichler (2018). Note that Definition 1 encompasses the class of weakly stationary processes, for which the asymptotic framework reduces to standard asymptotics.

Based on the observations in Section 2.2, a test for weak stationarity can be set up exploiting the uncorrelatedness of the elements in the sequence $(D_{\omega_j}^{(T)} \colon j = 1, \ldots, T)$. This could be done considering the lag-h sample covariance operator $T^{-1} \sum_{j=1}^T D_{\omega_j}^{(T)} \otimes D_{\omega_{j+h}}^{(T)}$ which should be centered at the zero element of S_2 for all $h = 1, \ldots, T - 1$. Here, two statistics based on the coefficients in the Karhunen–Loève decomposition of the fDFTs are considered. For $j = 1, \ldots, T$, let $(\phi_l^{\omega_j} \colon l \in \mathbb{N})$ be the orthonormal basis of eigenfunctions of $\mathcal{F}_{\omega_j}$ and observe that for this choice of basis $\text{Var}(\langle D_{\omega_j}, \phi_l^{\omega_j} \rangle) = \langle \mathcal{F}_{\omega_j}(\phi_l^{\omega_j}), \phi_l^{\omega_j} \rangle = \lambda_l^{\omega_j}$, where $(\lambda_l^{\omega_j} \colon l \in \mathbb{N}) \in \mathbb{R}_+$ are the eigenvalues of $\mathcal{F}_{\omega_j}$. Then, for any j, j', $(\phi_l^{\omega_j} \otimes \phi_{l'}^{\omega_{j'}} \colon l, l' \in \mathbb{N})$ is an orthonormal basis of $L_{\mathbb{C}}^2([0, 1]^2)$ and, by definition of the Hilbert–Schmidt inner product

on the algebraic tensor product space $H \otimes H$,

$$\frac{1}{T} \sum_{j=1}^{T} D_{\omega j}^{(T)} \otimes D_{\omega_{j+h}}^{(T)}$$

(7)
$$= \frac{1}{T} \sum_{j=1}^{T} \sum_{l=1}^{\infty} \sum_{l'=1}^{\infty} \langle D_{\omega j}^{(T)} \otimes D_{\omega_{j+h}}^{(T)}, \phi_l^{\omega j} \otimes \phi_{l'}^{\omega_{j+h}} \rangle_S \phi_l^{\omega j} \otimes \phi_{l'}^{\omega_{j+h}}$$

$$\approx \frac{1}{T} \sum_{j=1}^{T} \sum_{l=1}^{L} \sum_{l'=1}^{L'} \langle D_{\omega j}^{(T)}, \phi_l^{\omega j} \rangle \overline{\langle D_{\omega_{j+h}}^{(T)}, \phi_{l'}^{\omega_{j+h}} \rangle} \phi_l^{\omega j} \otimes \phi_{l'}^{\omega_{j+h}}$$

for sufficiently large L and L'. The foregoing motivates to set up tests based on the score products

(8)
$$\gamma_{j,h}^{(T)}(l, l') = \langle D_{\omega j}^{(T)}, \phi_l^{\omega j} \rangle \overline{\langle D_{\omega_{j+h}}^{(T)}, \phi_{l'}^{\omega_{j+h}} \rangle}$$

or on the standardized score products

(9)
$$\rho_{j,h}^{(T)}(l, l') = \frac{\gamma_{j,h}^{(T)}(l, l')}{\sqrt{\lambda_l^{\omega j} \lambda_{l'}^{\omega_{j+h}}}}.$$

In practice, the unknown spectral density operators $\mathcal{F}_{\omega j}$ and $\mathcal{F}_{\omega_{j+h}}$ are to be replaced with consistent estimators $\hat{\mathcal{F}}_{\omega j}^{(T)}$ and $\hat{\mathcal{F}}_{\omega_{j+h}}^{(T)}$ which will then yield respective sample eigenvalues $\hat{\lambda}_l^{\omega j}$ and eigenfunctions $\hat{\phi}_l^{\omega j}$. The estimated quantities corresponding to (8) and (9) will be denoted by $\hat{\gamma}_{j,h}^{(T)}(l, l')$ and $\hat{\rho}_{j,h}^{(T)}(l, l')$, respectively. As an estimator of $\mathcal{F}_\omega$, take

(10)
$$\hat{\mathcal{F}}_\omega^{(T)} = \frac{2\pi}{T} \sum_{j=1}^{T} K_b(\omega - \omega_j)(D_{\omega j}^{(T)} \otimes D_{\omega j}^{(T)}),$$

where $K_b(\cdot)$ is a kernel with bandwidth b satisfying the following conditions:

ASSUMPTION 3.1. (a) Let $K: [-\frac{1}{2}, \frac{1}{2}] \to \mathbb{R}_+$ be symmetric with $\int K(x)\,dx = 1$ and $\int K(x)^2\,dx < \infty$.

(b) Let $b = b_T$ be a bandwidth such that $T^{-1/2} \ll b_T \ll T^{-1/4}$.

(c) Let $K_b(x) = b^{-1} K((2\pi b)^{-1} x)$ and extend the kernel periodically such that $K_b(x) = K_b(x \pm 2\pi)$ in order to include estimates for frequencies around $\pm\pi$.

To set up the test statistics, it now appears reasonable to extract information across a range of directions $l = 1, \ldots, L_j$ and $l' = 1, \ldots, L_{j+h}$ as well as a selection of lags $h = 1, \ldots, \bar{h}$, where $\bar{h}$ denotes an upper limit. The truncation parameters $L_j = L(\omega_j)$ and $L_{j+h} = L(\omega_{j+h})$ are explicitly allowed to depend on the jth and $(j+h)$th Fourier frequencies in order to accommodate heterogeneity in the Karhunen–Loève decompositions across the spectral domain. Set

$$\hat{\beta}_{h,u}^{(T)} = \frac{1}{T} \sum_{j=1}^{T} \sum_{l=1}^{L_j} \sum_{l'=1}^{L_{j+h}} \hat{\gamma}_{j,h}^{(T)}(l, l') \quad \text{and}$$

(11)

$$\hat{\beta}_{h,s}^{(T)} = \frac{1}{T} \sum_{j=1}^{T} \sum_{l=1}^{L_j} \sum_{l'=1}^{L_{j+h}} \hat{\rho}_{j,h}^{(T)}(l, l'),$$

where the subscripts u and s refer to the unstandardized and standardized forms, respectively. In the following, the subscript x will be used to refer to any of these two versions when no confusion can arise.

Choose next a collection $h_1, \ldots, h_M$ of lags, each of which bounded from above by $\bar{h}$, to pool information across a number of autocovariances and build the vectors

$$\hat{\boldsymbol{b}}_{M,x}^{(T)} = (\Re\hat{\beta}_{h_1,x}^{(T)}, \ldots, \Re\hat{\beta}_{h_M,x}^{(T)}, \Im\hat{\beta}_{h_1,x}^{(T)}, \ldots, \Im\hat{\beta}_{h_M,x}^{(T)})^\top,$$

where $\Re$ and $\Im$ denote real and imaginary part, respectively. Finally, set up the quadratic forms

(12)
$$\hat{Q}_{M,x}^{(T)} = T(\hat{\boldsymbol{b}}_{M,x}^{(T)})^\top \hat{\Sigma}_{M,x}^{-1} \hat{\boldsymbol{b}}_{M,x}^{(T)},$$

where $\hat{\Sigma}_{M,x}$ is an estimator of the asymptotic covariance matrix of the vectors $\boldsymbol{b}_{M,x}^{(T)}$ which are defined by replacing $\hat{\gamma}_{j,h}^{(T)}(l, l')$ and $\hat{\rho}_{j,h}^{(T)}(l, l')$ with $\gamma_{j,h}^{(T)}(l, l')$ and $\rho_{j,h}^{(T)}(l, l')$ in (11) and then using the resulting $\beta_{h,x}^{(T)}$ in place of $\hat{\beta}_{h,x}^{(T)}$ in the definition of $\hat{\boldsymbol{b}}_{M,x}^{(T)}$. The foregoing provides the two test statistics $\hat{Q}_{M,u}^{(T)}$ and $\hat{Q}_{M,s}^{(T)}$ that will be used to test the null of stationarity against the alternative of local stationarity. Note that both quadratic forms depend on the tuning parameters L_j, L_{j+h} and M, the selection of which will be evaluated empirically in Section 5.

To facilitate the derivation of large-sample results, the following assumptions are made for the unstandardized respectively standardized test:

Condition C_u: Let $L_j \sim \log T$ and $\lim_l \inf_\omega \lambda_l^\omega > 0$;
Condition C_s: Let $\inf_\omega \lambda_{\bar{L}}^\omega > 0$ for some $\bar{L} \geq \sup_j L_j$.

In keeping with the above arrangement, the respective conditions will be referred to as C_x if no confusion arises. Condition C_u for the unstandardized test allows to send the truncation levels L_j to infinity in a coordinated manner as long as the divergence is slow (here, logarithmic) compared to T; see Fremdt et al. (2014). Condition C_s for the standardized test requires on the other hand a finite truncation level; this to ensure that the smallest eigenvalues of the compact operators $\mathcal{F}_{\omega_j}$ are bounded away from zero, which show up in the denominator of (9).

4. Large-sample results.

4.1. *Assumptions*. The following gives the main requirements under both stationarity and local stationarity in terms of cumulant tensors of the functional time series (Appendix B) that are needed to establish the asymptotic behavior of the test statistics under both hypotheses. Note that the null hypothesis is nested within the alternative. Because of this basic fact, we start with the general assumptions under local stationarity before specializing to the stationary case.

ASSUMPTION I(k, ℓ). Assume $(X_t^{(T)}: t \leq T, T \in \mathbb{N})$ and $(X_t^{(u)}: t \in \mathbb{Z})$ are as in Definition 1. Suppose $\sup_t \mathbb{E}[\|X_t\|_2^{\min(k,12)}] < \infty$ and that there exists a a positive sequence $\kappa_{k;t_1,\ldots,t_{k-1}}$ in $L_\mathbb{R}^2([0, 1]^k)$, independent of T such that, for all $j = 1, \ldots, k - 1$ and some $\ell \in \mathbb{N}$,

(13)
$$\sum_{t_1,\ldots,t_{k-1}\in\mathbb{Z}} (1 + |t_j|^\ell)\|\kappa_{k;t_1,\ldots,t_{k-1}}\|_2 < \infty.$$

Suppose furthermore that there exist representations

(14)
$$X_t^{(T)} - X_t^{(t/T)} = Y_t^{(T)} \quad \text{and} \quad X_t^{(u)} - X_t^{(v)} = (u - v)Y_t^{(u,v)},$$

for some processes $(Y_t^{(T)} : t \leq T, T \in \mathbb{N})$ and $(Y_t^{(u,v)} : t \in \mathbb{Z})$ taking values in $H_{\mathbb{R}}$ whose kth order joint cumulants satisfy:

(i) $\| \operatorname{cum}(X_{t_1}^{(T)}, \ldots, X_{t_{k-1}}^{(T)}, Y_{t_k}^{(T)}) \|_2 \leq \frac{1}{T} \|\kappa_{k; t_1-t_k, \ldots, t_{k-1}-t_k}\|_2$,

(ii) $\| \operatorname{cum}(X_{t_1}^{(u_1)}, \ldots, X_{t_{k-1}}^{(u_{k-1})}, Y_{t_k}^{(u_k, v)}) \|_2 \leq \|\kappa_{k; t_1-t_k, \ldots, t_{k-1}-t_k}\|_2$,

(iii) $\sup_u \| \operatorname{cum}(X_{t_1}^{(u)}, \ldots, X_{t_{k-1}}^{(u)}, X_{t_k}^{(u)}) \|_2 \leq \|\kappa_{k; t_1-t_k, \ldots, t_{k-1}-t_k}\|_2$,

(iv) $\sup_u \| \frac{\partial^\ell}{\partial u^\ell} \operatorname{cum}(X_{t_1}^{(u)}, \ldots, X_{t_{k-1}}^{(u)}, X_{t_k}^{(u)}) \|_2 \leq \|\kappa_{k; t_1-t_k, \ldots, t_{k-1}-t_k}\|_2$.

Assumption I provides Lipschitz conditions that are generalizations of those in Lee and Rao (2017) who investigated the properties of quadratic forms of stochastic processes in a finite-dimensional setting. The above conditions enable to express the behavior of the fDFT's of a kth order locally stationary process in terms of kth order time-varying spectral density tensors (Lemma B.1). This is convenient in order to derive explicit expressions of the distributional properties under the alternative and to understand departures from stationarity. Under H_A, we can uniquely characterize the second-order stucture of the stochastic process $(X_t^{(T)} : t \leq T, T \in \mathbb{N})$ via the *time-varying spectral density operator*

$$(15) \qquad \mathcal{F}_{u, \omega} = \frac{1}{2\pi} \sum_{h \in \mathbb{Z}} \mathcal{C}_{u, h} e^{-\mathrm{i}\omega h},$$

where $\mathcal{C}_{u,h} = \operatorname{cum}(X_h^{(u)}, X_0^{(u)})$ denotes the local cumulant tensor at fixed time u of the stationary approximating process $(X_t^{(u)} : t \in \mathbb{Z})$. Note that the parameter ℓ and (iii)–(iv) in Assumption I influence the smoothness of the operator-valued mapping $(u, \omega) \mapsto \mathcal{F}_{u,\omega}$. Under Assumption I$(2, 2)$, derivative maps are well-defined elements of $S_2(H)$, and $\omega \mapsto \mathcal{F}_{u,\omega}$ is uniformly continuous in ω with respect to $\|\cdot\|_2$. We refer to Lemma S2.2 for details. More generally, under kth order local stationarity, these properties carry over to the *local kth order cumulant spectral density tensor*

$$(16) \qquad \mathcal{F}_{u; \omega_1, \ldots, \omega_{k-1}} = \frac{1}{(2\pi)^{k-1}} \sum_{t_1, \ldots, t_{k-1} \in \mathbb{Z}} \mathcal{C}_{u; t_1, \ldots, t_{k-1}} e^{-\mathrm{i}\sum_{j=1}^{k-1} \omega_j t_j},$$

where $\omega_1, \ldots, \omega_{k-1} \in \mathbb{R}$ and $\mathcal{C}_{u; t_1, \ldots, t_{k-1}} = \operatorname{cum}(X_{t_1}^{(u)}, \ldots, X_{t_{k-1}}^{(u)}, X_{t_0}^{(u)})$ is the corresponding local cumulant kernel tensor of order k at time u_0. Observe that, for $k > 1$, (16) can be viewed as an element of $S_2(H^{\otimes \lfloor (k+1)/2 \rfloor}, H^{\otimes \lfloor k/2 \rfloor})$. Under kth order stationarity the above objects become independent of local time u, so that $\mathcal{F}_{u; \omega_1, \ldots, \omega_{k-1}} \equiv \mathcal{F}_{\omega_1, \ldots, \omega_{k-1}}$, and Assumption I specializes to the following:

ASSUMPTION I$^*(k, \ell)$. Let $(X_t : t \in \mathbb{Z})$ be a kth order stationary functional time series with values in $H_{\mathbb{R}}$ such that (i) $\mathbb{E}[\|X_0\|_2^{\min(k, 12)}] < \infty$ and (ii) $\sum_{t_1, \ldots, t_{k-1} = -\infty}^{\infty} (1 + |t_j|^\ell) \|\mathcal{C}_{t_1, \ldots, t_{k-1}}\|_2 < \infty$ for all $1 \leq j \leq k - 1$.

Because the test statistics require estimators of the eigenelements of $\mathcal{F}_\omega$, it is of importance to consider the properties of the estimator (10) for both null and alternative hypotheses. The next theorem shows that it is a consistent estimator of the integrated (in a Bochner sense) time-varying spectral density operator

$$G_\omega = \int_0^1 \mathcal{F}_{u, \omega} \, du,$$

where the convergence is uniform in $\omega \in [-\pi, \pi]$ with respect to $\|\cdot\|_2$. This, therefore, becomes an operator-valued function in ω that acts on H and is independent of rescaled time u. Under H_0, G_ω thus reduces to $\mathcal{F}_\omega$.

THEOREM 4.1 (Consistency and uniform convergence). *Suppose* $(X_t^{(T)} : t \leq T, T \in \mathbb{N})$ *satisfies Assumption* I(4, 2). *Consider the estimator* $\hat{\mathcal{F}}_\omega^{(T)}$ *in* (10) *with smoothing kernel* K *fulfilling Assumption* 3.1(a) *and* (c). *Then*:

(a) $\mathbb{E}[\||\hat{\mathcal{F}}_\omega^{(T)} - G_\omega\||_2^2] = O((bT)^{-1} + b^4)$, *uniformly in* $\omega \in [-\pi, \pi]$.
(b) *If, in addition, Assumption* 3.1(b) *holds and* K *has bounded derivative on* $(-1/2, 1/2)$ *then*,

$$\sup_{\omega \in [-\pi, \pi]} \||\hat{\mathcal{F}}_\omega^{(T)} - G_\omega\||_2 \xrightarrow{p} 0.$$

The proof of Theorem 4.1 is given in Section C.3 of Appendix C. Since the theorem shows consistency of $\hat{\mathcal{F}}_\omega$, a self-adjoint element of $S_2(H)$, it follows from Mas and Menneteau (2003) that the sample eigenelements $(\hat{\lambda}_l^\omega, \hat{\phi}_l^\omega : l \in \mathbb{N})$ of $\hat{\mathcal{F}}_\omega$ provide consistent estimators for the eigenelements $(\tilde{\lambda}_l^\omega, \tilde{\phi}_l^\omega : l \in \mathbb{N})$ of G_ω. If H_0 is satisfied, then the stated consistency holds for the eigenelements $(\lambda_l^\omega, \phi_l^\omega : l \in \mathbb{N})$ of $\mathcal{F}_\omega$.

4.2. *Properties under the null of stationarity.* The asymptotic results under H_0 are collected in this section. The first theorem establishes that the scaled difference between $\beta_{h,x}^{(T)}$ and $\hat{\beta}_{h,x}^{(T)}$ is negligible in large samples. Note that the assumptions here and for other theorems in this section are formulated imposing stationarity on certain moments for the null hypothesis via Assumption I*. To verify the results, further assumptions on higher-order cumulants are typically required. These are controlled via Assumption I.

THEOREM 4.2. *Let Assumption* 3.1, *Assumption* I(12, 2) *and* C_x *hold. Then, under* H_0, *for any fixed* h,

$$\sqrt{T}|\hat{\beta}_{h,x}^{(T)} - \beta_{h,x}^{(T)}| = O_p\left(\frac{1}{bT} + b^2\right).$$

The proof is given in Section D.2.2 of Appendix D. In view of Assumption 3.1, Theorem 4.2 shows that the distributional properties of $\hat{\beta}_{h,x}^{(T)}$ are asymptotically the same as those of $\beta_{h,x}^{(T)}$ under the stated conditions on the bandwidth rates. Note these rates are necessary for the estimator in (10) to be consistent, as is seen from part (a) of Theorem 4.1, which reduces to the stationary case if the process does not depend on u. They hence do not impose an additional constraint under H_0.

The next theorem derives that, under the additional assumption of fourth-order stationarity, the asymptotic variance is uncorrelated for all lags h and that there is no correlation between the real and imaginary parts. For $n \in \mathbb{N}$, set $[n] = \{1, \ldots, n\}$.

THEOREM 4.3. *Let Assumption* 3.1 *and* C_x *hold. Suppose further that Assumption* I*(4, 2) *is satisfied. Then, for* $h_1 = h_2 = h$,

(a) $T \operatorname{Cov}(\Re\hat{\beta}_{h,u}^{(T)}, \Re\hat{\beta}_{h,u}^{(T)})$

$$= T \operatorname{Cov}(\Im\hat{\beta}_{h,u}^{(T)}, \Im\hat{\beta}_{h,u}^{(T)})$$

$$\to \frac{1}{4\pi} \int \int \sum_{(l,l') \in \mathcal{L} \times \mathcal{L}'} \langle \mathcal{F}_{\omega, -\omega - \omega_h, -\omega'}(\phi_{l_1'}^{\omega'} \otimes \phi_{l_2'}^{\omega' + \omega_h'}), \phi_{l_1}^\omega \otimes \phi_{l_2}^{\omega + \omega_h} \rangle \, d\omega \, d\omega'$$

$$+ \frac{1}{2\pi} \int \sum_{l \in \mathcal{L}} \lambda_{l_1}^\omega \lambda_{l_2}^{\omega + \omega_h} \, d\omega,$$

(b) $T \operatorname{Cov}(\Re \hat{\beta}_{h,s}^{(T)}, \Re \hat{\beta}_{h,s}^{(T)})$

$$= T \operatorname{Cov}(\Im \hat{\beta}_{h,s}^{(T)}, \Im \hat{\beta}_{h,s}^{(T)})$$

$$\to \frac{1}{4\pi} \int \int \sum_{(l,l') \in \mathcal{L} \times \mathcal{L}'} \frac{\langle \mathcal{F}_{\omega, -\omega - \omega_h, -\omega'}(\phi_{l_1'}^{\omega'} \otimes \phi_{l_2'}^{\omega' + \omega_h'}), \phi_{l_1}^{\omega} \otimes \phi_{l_2}^{\omega + \omega_h} \rangle}{\sqrt{\lambda_{l_1}^{\omega} \lambda_{l_2}^{\omega + \omega_h} \lambda_{l_1'}^{\omega'} \lambda_{l_2'}^{\omega' + \omega_h'}}} \, d\omega \, d\omega'$$

$$+ \frac{1}{2\pi} \int \sum_{l \in \mathcal{L}} \delta_{l_1, l_2} \, d\omega,$$

where $l = (l_1, l_2)$, $l' = (l_1', l_2')$, $\mathcal{L} = [L(\omega)] \times [L(\omega + \omega_h)]$, $\mathcal{L}' = [L(\omega')] \times [L(\omega' + \omega_h')]$ and $\delta_{i,j} = 1$ if $i = j$ and 0 otherwise. If $h_1 \neq h_2$, $T \operatorname{Cov}(\Re \hat{\beta}_{h_1,x}^{(T)}, \Re \hat{\beta}_{h_2,x}^{(T)}) \to 0$, $T \operatorname{Cov}(\Im \hat{\beta}_{h_1,x}^{(T)}, \Im \hat{\beta}_{h_2,x}^{(T)}) \to 0$ and $T \operatorname{Cov}(\Re \hat{\beta}_{h_1,x}^{(T)}, \Im \hat{\beta}_{h_2,x}^{(T)}) \to 0$.

The proof of Theorem 4.3 is given in Appendix C.2. Observe that the results in part (b) imply that the standardized test statistics is pivotal if the data is Gaussian. Note also that the results in the theorem use at various instances the fact that the kth order spectral density operator at frequency $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_k)^T \in \mathbb{R}^k$ is equal to the kth order spectral density operator at frequency $-\boldsymbol{\omega}$ in the manifold $\sum_{j=1}^{k} \omega_j \mod 2\pi$.

With the previous results in place, the large-sample behavior of the quadratic form statistics $\hat{Q}_{M,x}^{(T)}$ defined in (12) can be derived. This is done in the following theorem.

THEOREM 4.4. *Let Assumption* 3.1 *and* C_x *hold. Suppose further that Assumption* I(k, 2) *is satisfied for all* $k \geq 3$. *Then, under* H_0:

(a) *For any collection* $h_1, \ldots, h_M$ *bounded by* $\bar{h}$,

$$\sqrt{T} \hat{\boldsymbol{b}}_{M,x}^{(T)} \xrightarrow{\mathcal{D}} \mathcal{N}_{2M}(\boldsymbol{0}, \Sigma_{0,x}) \quad (T \to \infty),$$

where $\xrightarrow{\mathcal{D}}$ *denotes convergence in distribution. Under the additional assumption of fourth-order stationarity,* $\mathcal{N}_{2M}(\boldsymbol{0}, \Sigma_{0,x})$ *is a* $2M$-*dimensional normal distribution with mean* $\boldsymbol{0}$ *and diagonal covariance matrix* $\Sigma_{0,x} = \operatorname{diag}(\sigma_{0,m,x}^2 : m = 1, \ldots, 2M)$ *whose elements are*

$$\sigma_{0,m,x}^2 = \lim_{T \to \infty} T \operatorname{Cov}(\Re \hat{\beta}_{h_m,x}, \Re \hat{\beta}_{h_m,x}), \quad m = 1, \ldots, M,$$

and $\sigma_{0,M+m,x}^2 = \sigma_{0,m,x}^2$. *The explicit form of the limit is determined by Theorem* 4.3. *If fourth-order stationarity is violated, then the limiting normal distribution has a nondiagonal covariance structure.*

(b) *Using the result in* (a), *it follows that for the statistic defined in* (12)

$$\hat{Q}_{M,x}^{(T)} \xrightarrow{\mathcal{D}} \chi_{2M}^2 \quad (T \to \infty),$$

where χ_{2M}^2 *is a* χ^2-*distributed random variable with* $2M$ *degrees of freedom.*

The proof of Theorem 4.4 is provided in Appendix D. Part (b) of the theorem can now be used to construct tests with asymptotic level α. Note that the application of the test requires an estimator of $\hat{\Sigma}_{M,x}$. This will be discussed in Section 4.4.

To explicitly compute the limiting covariance structure in part (a) of Theorem 4.4 under second-order stationarity but fourth-order nonstationarity, the source of nonstationarity needs to be specified. For example, the results put forward in the next two sections allow for the

computation of $\Sigma_{0,x}$ if the process is fourth-order locally stationary. Then, in the covariance structure of the covariance operator of the fDFT's, the fourth-order cumulant tensor component will, for $h_1 \neq h_2$, (quadratically) decay in norm as the distance $|h_1 - h_2|$ increases (see Lemma B.1, Corollary B.1(ii) and equation (26)). As a consequence of this term being present in the covariance structure, the real and imaginary part of the projections are no longer uncorrelated, but the correlation decays with increasing distance $|h_1 - h_2|$. In this scenario a small loss of power is to be expected when the test statistic is built under the assumption of a diagonal covariance structure.

4.3. *Properties under the alternative.* This section contains a generalization of the results in Section 4.2 to locally stationary functional time series. The following theorem is the counterpart to Theorem 4.2 under the null hypothesis.

THEOREM 4.5. *Let Assumption* 3.1, *Assumption* I(12, 2) *and* C_x *hold. Then, under* H_A,

$$\sqrt{T}\mathbb{E}[|\hat{\beta}_{h,x}^{(T)} - \beta_{h,x}^{(T)} - \mathcal{B}_{h,x}^{(T)}|] = O\left(\frac{1}{bT} + b^2 + \frac{1}{b\sqrt{T}} + b^2\sqrt{T}\right),$$

where

$$\mathcal{B}_{h,x}^{(T)} = \frac{1}{T}\sum_{j=1}^{T}\sum_{l\in\mathcal{L}}\zeta_{l,x}\langle\mathbb{E}[D_{\omega_j}\otimes D_{\omega_{j+h}}], \mathbb{E}[\hat{\phi}_l^{\omega_j}\otimes\hat{\phi}_{l'}^{\omega_j+h}] - \hat{\phi}_l^{\omega_j}\otimes\hat{\phi}_{l'}^{\omega_j+h}\rangle_S$$

is a stochastic bias term satisfying $\sqrt{T}\mathcal{B}_{h,x}^{(T)} = O_P(1)$, *and* $\zeta_{l,u} = 1$ *and* $\zeta_{l,s} = (\tilde{\lambda}_l^\omega, \tilde{\lambda}_{l'}^{\omega+\omega_h})^{-1/2}$.

The proof of Theorem 4.5 is given in Section D.2.2 of Appendix D. In view of Assumption 3.1, the theorem shows that $\hat{\beta}_{h,x}^{(T)}$ has the same asymptotic sampling properties as $\beta_{h,x}^{(T)}$ up to a stochastically bounded bias term (after scaling with $\sqrt{T}$). Note that $|\hat{\beta}_{h,x}^{(T)} - \mathbb{E}[\beta_{h,x}^{(T)}]| \xrightarrow{P} 0$, where

$$(17) \qquad \mathbb{E}[\beta_{h,x}^{(T)}] \to \frac{1}{2\pi}\int_0^{2\pi}\int_0^1\sum_{l\in\mathcal{L}}\zeta_{l,x}\langle\mathcal{F}_{u;\omega}e^{-\imath 2\pi uh}, \tilde{\phi}_l^\omega\otimes\tilde{\phi}_{l'}^{\omega+\omega_h}\rangle_S\,du\,d\omega = \mu_{h,x}$$

is an noncentrality parameter (see Appendix C.1) that will have to enter the limit distribution of $\hat{Q}_{M,x}^{(T)}$ as a consequence of the violation of weak stationarity. We discuss this term in some more detail below.

A precise formulation of the asymptotic properties under H_A is given in the next theorem.

THEOREM 4.6. *Let Assumption* 3.1 *and* C_x *hold. Suppose further that Assumption* I(k, 2) *is satisfied for all* $k \geq 2$. *Then, under* H_A:

(a) *For any collection* $h_1, \ldots, h_M$ *bounded by* $\bar{h}$,

$$\sqrt{T}\hat{b}_{M,x}^{(T)} \xrightarrow{\mathcal{D}} \mathcal{N}_{2M}(\boldsymbol{\mu}_x, \Sigma_{A,x}) \quad (T \to \infty),$$

where $\mathcal{N}_{2M}(\boldsymbol{\mu}_x, \Sigma_{A,x})$ *denotes a 2M-dimensional normal distribution with mean vector* $\boldsymbol{\mu}_x$ *whose first M components are* $\Re\mu_{h_m,x}$ *and last M components are* $\Im\mu_{h_m,x}$, *where* $\mu_{h_m,x}$ *is defined through* (17), *and nondiagonal block covariance matrix*

$$\Sigma_{A,x} = \begin{pmatrix} \Sigma_{A,x}^{(11)} & \Sigma_{A,x}^{(12)} \\ \Sigma_{A,x}^{(21)} & \Sigma_{A,x}^{(22)} \end{pmatrix}$$

whose $M \times M$ *blocks are determined by the results in Appendix E and Section S6.2 of the Online Supplement.*

(b) *Using the result in* (a), *it follows that for the statistic defined in* (12)

$$\hat{Q}_{M,x}^{(T)} \overset{\mathcal{D}}{\to} \chi^2_{\mu_x, 2M} \quad (T \to \infty),$$

where $\chi^2_{\mu_x, 2M}$ denotes a generalized noncentral χ^2-distributed random variable with noncentrality parameter $\mu_x = \|\boldsymbol{\mu}_x\|_2^2$ and $2M$ degrees of freedom.

The proof of Theorem 4.6 can be found in Appendix E. Observe that the limiting noncentrality parameter μ_x of the statistic $\hat{Q}_{M,x}^{(T)}$ measures the aggregation of the functions in (17). Under H_A, the operator in (7) no longer converges in norm to the zero operator but instead to the operator $\frac{1}{2\pi} \int_0^{2\pi} \int_0^1 \mathcal{F}_{u,\omega} e^{-i2\pi uh} \, du \, d\omega$. The properties of the latter, which are extracted to finite dimension via $\mu_{h,x}$, carry some meaningful information on the behavior of the test under the alternative. First, denote a general term in the limiting expansion of $\mu_{h,x}$ by

$$\mu_{h,x}(\boldsymbol{l}) = \frac{1}{2\pi} \int_0^{2\pi} \int_0^1 \zeta_{\boldsymbol{l},x} \langle \mathcal{F}_{u;\omega} e^{-i2\pi uh}, \tilde{\phi}_l^\omega \otimes \tilde{\phi}_{l'}^{\omega+\omega_h} \rangle_S \, du \, d\omega.$$

For fixed directions $\boldsymbol{l} = (l, l')$, this function can be seen to approximate the $(h, 0)$th Fourier coefficients of the function $(u, \omega) \mapsto \zeta_{\boldsymbol{l},x} \langle \mathcal{F}_{u,\omega}(\tilde{\phi}_{l'}^{\omega+\omega_h}), \tilde{\phi}_l^\omega \rangle$, that is, for small h and $T \to \infty$ they approximate

$$\vartheta_{h,j,x}(\boldsymbol{l}) = \frac{1}{2\pi} \int_0^{2\pi} \int_0^1 \zeta_{\boldsymbol{l},x} \langle \mathcal{F}_{u,\omega} \tilde{\phi}_{l'}^{\omega+\omega_h}, \tilde{\phi}_l^\omega \rangle e^{i2\pi uh - ij\omega} \, du \, d\omega$$

with $j = 0$. In other words, $\mu_{h,x}(\boldsymbol{l}) \approx \vartheta_{h,0,x}(\boldsymbol{l})$. If the process is weakly stationary, then the integrand of the coefficient does not depend on u and all Fourier coefficients are zero except $\vartheta_{0,j,x}(\boldsymbol{l})$. In particular, $\vartheta_{0,0,s}(\boldsymbol{l}) = 1$. Following Paparoditis (2009) and Dwivedi and Subba Rao (2011), the mean functions can thus be seen to reveal long-term nonstationary behavior. Unlike testing methods based on segments in the time domain, the proposed method is therefore able to detect smoothly changing behavior in the temporal dependence structure.

Secondly, the operator $\int_0^1 \mathcal{F}_{u,\omega} e^{-i2\pi uh} \, du$ can be viewed as the hth Fourier coefficient of the operator-valued function $(u) \mapsto \mathcal{F}_{u,\omega}$ for fixed ω (Lemma B.1) which exhibits a quadratic decay in norm as a function of h such that the sum of the norms of these coefficients is finite (Corollary B.1). Since this behavior carries over to the projections, the contribution to μ_x of the functions $\mu_{h,x}$ in (17) for larger values of h will become negligible. Intuitively, utilizing large values of M in the statistic $\hat{Q}_{M,x}^{(T)}$ is hence expected to increase the likelihood of a type II error; see also Section 5.

The results in this and the previous section require an understanding of the estimator $\hat{\Sigma}_{M,x}$ used in the definition of the test statistics $\hat{Q}_{M,x}^{(T)}$ in (12). The corresponding results are part of the next subsection.

4.4. *Estimating the fourth-order spectrum.* The estimation of the matrix Σ_M is a necessary ingredient in the application of the proposed stationarity test. Generally, the estimation of the sample (co)variance can influence the power of tests, as has been observed in a number of previous works set in similar albeit nonfunctional contexts. Among the contributions more closely related to this paper are Paparoditis (2009), who used the spectral density of the squares, Dwivedi and Subba Rao (2011), who focused on Gaussianity of the observations, and Jentsch and Subba Rao (2015), who employed a stationary bootstrap procedure. A different idea was put forward by Bandyopadhyay and Subba Rao (2017) and Bandyopadhyay, Jentsch and Subba Rao (2017). These authors utilized the notion of orthogonal samples to estimate the variance, falling back on a general estimation strategy developed in Subba Rao (2018).

In order to utilize the results of Theorem 4.4, we require an estimator of the trispectral density operator $\mathcal{F}_{\omega,-\omega-\omega_h,-\omega'}$ which can then, subsequently, be projected onto the (standardized) empirical eigenfunctions and integrated over ω, ω'. As an estimator, consider

$$
\hat{\mathcal{F}}_{\omega_{j_1},\ldots,\omega_{j_4}} = \frac{(2\pi)^3}{(b_4 T)^3} \sum_{k_1,k_2,k_3} K_4\left(\frac{\omega_{j_1}-\omega_{k_1}}{b_4},\ldots,\frac{\omega_{j_4}-\omega_{k_4}}{b_4}\right)
$$
$$
\times \Phi(\omega_{k_1},\ldots,\omega_{k_4}) I^{(T)}_{\omega_{k_1},\ldots,\omega_{k_4}},
$$
(18)

where

$$
I^{(T)}_{\omega_{k_1},\omega_{k_2},\omega_{k_3},\omega_{k_4}} = \frac{T}{2\pi} D_{\omega_{k_1}} \otimes D_{\omega_{k_2}} \otimes D_{\omega_{k_3}} \otimes D_{\omega_{k_4}}
$$

denotes the triperiodogram tensor and where $K_4(x_1,\ldots,x_4)$ is a smoothing kernel with compact support on $\mathbb{R}^4$ and where

$$
\Phi(\alpha_1, \alpha_2, \alpha_3, \alpha_4) = 1
$$

if $\sum_{k=1}^4 \alpha_k \equiv 0 \mod 2\pi$ such that $\sum_{k \in J} \alpha_k \not\equiv 0 \mod 2\pi$ where J is any nonempty subset of $\{1, 2, 3, 4\}$ and equals 0 otherwise. This function therefore controls that we are only working with those combinations of frequencies that lie on the principal manifold but do not lie in any proper submanifold. The reason for this is that, for $k > 2$, the expectation of kth order periodogram tensors evaluated at such submanifolds possibly diverges (see also Brillinger and Rosenblatt (1967), for the Euclidean case). As the next theorem shows, the estimator in (18) can be shown to be consistent if the bandwidth b_4 satisfies $b_4 \to 0$ but $b_4^{-3} T \to \infty$ as $T \to \infty$.

THEOREM 4.7.   *Suppose Assumption* $\mathrm{I}^*(4, 2)$ *and Assumption* $\mathrm{I}(8, 2)$ *hold. Then, the estimator* (18) *of the trispectral density operator satisfies*

$$
\mathbb{E}\left[\left\|\left\|\frac{(2\pi)^2}{T^2} \sum_{j_1,j_2=1}^T \hat{\mathcal{F}}_{\omega_{j_1},-\omega_{j_1+h},-\omega_{j_2}} - \int\int \mathcal{F}_{\omega,-\omega+\omega_h,-\omega'}\, d\omega\, d\omega'\right\|\right\|_2^2\right]
$$
(19)
$$
= O\left(\frac{1}{b_4^3 T} + b_4^4\right).
$$

The section is rounded out with large-sample behavior under the alternative.

THEOREM 4.8.   *Suppose Assumption* $\mathrm{I}(8, 2)$ *holds. Then,*

(a) $\left\|\left\|\mathbb{E}\left[\frac{(2\pi)^2}{T^2} \sum_{j_1,j_2=1}^T \hat{\mathcal{F}}_{\omega_{j_1},-\omega_{j_1+h},-\omega_{j_2}}\right] - \int\int G_{\omega,-\omega+\omega_h,-\omega'}\, d\omega\, d\omega' - \mathcal{Z}_h\right\|\right\|_2$

$$
= O\left(\frac{1}{b_4 T} + b_4\right),
$$

(b) $\left\|\left\|\mathrm{Cov}(\hat{\mathcal{F}}_{\omega_{j_1},\omega_{j_2},\omega_{j_3}}, \hat{\mathcal{F}}_{\omega_{j_1},\omega_{j_2},\omega_{j_3}})\right\|\right\|_2^2 = O\left(\frac{1}{b_4^3 T}\right),$

where $G_{\omega,-\omega+\omega_h,-\omega'}$ *denotes the time-integrated trispectral operator and where* $\mathcal{Z}_h \in S_2(H \otimes H)$ *is a bias term of order* $O(\|\|\mathcal{Z}_h\|\|_2) = 1$.

The proofs of Theorems 4.7 and 4.8 are given in Section S7 of the Online Supplement. Using continuity of the inner-product, Theorem 4.1(a) and the continuous mapping theorem imply projecting onto the empirical eigenfunctions will not affect the rates.

5. Empirical results. This section reports the results of an illustrative simulation study designed to verify that the large-sample theory is useful for applications to finite samples. The test is subsequently applied to annual temperature curves data. The findings provide guidelines for a further fine-tuning of the test procedures to be investigated in future research.

5.1. *Simulation setting.* To generate functional time series, the general strategy applied, for example, in Aue, Dubart Norinho and Hörmann (2015) and Hörmann, Kidziński and Hallin (2015), is utilized. For this simulation study all processes are built on a Fourier basis representation on the unit interval $[0, 1]$ with basis functions $\psi_1, \ldots, \psi_{15}$. Note that the lth Fourier coefficient of a pth-order functional autoregressive, FAR(p), process $(X_t : t \in \mathbb{Z})$ satisfies

(20)
$$\langle X_t, \psi_l \rangle = \sum_{l'=1}^{\infty} \sum_{t'=1}^{p} \langle X_{t-t'}, \psi_l \rangle \langle A_{t'}(\psi_l), \psi_{l'} \rangle + \langle \varepsilon_t, \psi_l \rangle$$

$$\approx \sum_{l'=1}^{L_{\max}} \sum_{t'=1}^{p} \langle X_{t-t'}, \psi_l \rangle \langle A_{t'}(\psi_l), \psi_{l'} \rangle + \langle \varepsilon_t, \psi_l \rangle,$$

the quality of the approximation depending on the choice of $L_{\max}$. The vector of the first $L_{\max}$ Fourier coefficients $\mathbf{X}_t = (\langle X_t, \psi_1 \rangle, \ldots, \langle X_t, \psi_{L_{\max}} \rangle)^\top$ can thus be generated using the pth-order vector autoregressive, VAR(p), equations

$$\mathbf{X}_t = \sum_{t'=1}^{p} \mathbf{A}_{t'} \mathbf{X}_{t-t'} + \boldsymbol{\varepsilon}_t,$$

where the (l, l') element of $\mathbf{A}_{t'}$ is given by $\langle A_{t'}(\psi_l), \psi_{l'} \rangle$ and $\boldsymbol{\varepsilon}_t = (\langle \varepsilon_t, \psi_1 \rangle, \ldots, \langle \varepsilon_t, \psi_{L_{\max}} \rangle)^\top$. The entries of the matrices $\mathbf{A}_{t'}$ are generated as $\mathcal{N}(0, v_{l,l'}^{(t')})$ random variables with the specifications of $v_{l,l'}$ given below. To ensure stationarity or the existence of a causal solution, the norms $\kappa_{t'}$ of $A_{t'}$ are required to satisfy certain conditions, for example, $\sum_{t'=1}^{p} \|\|A_{t'}\|\|_\infty < 1$, which might be of more complicated nature (see Bosq (2000), van Delft and Eichler (2018), for the stationary and locally stationary case, respectively). The functional white noise, FWN, process is included in (20) setting $p = 0$. All simulation experiments were implement in R and any result reported in the remainder of this section is based on 1000 simulation runs.

5.2. *Specification of tuning parameters.* The test statistics in (12) depends on the tuning parameters $L_j = L(\omega_j)$, determining the dimension of the projection spaces, and M, the number of frequency lags to be included in the procedure. In the following, a criterion will be set up to choose L_j, while for M only a limited number of values were entertained because the selection is less critical for the performance as long as it is not chosen too large. Figure 1 shows that it can well be of interest in practice to choose L_j in a frequency-dependent way, as the eigenvalue decay might vary significantly between different ω_j. The left part of the figure shows the situation for a functional white noise sequence. The spectral density operators are constant operator-valued functions of frequency and consequently their spectral decompositions coincide, producing relatively straight lines in the sample eigenvalue plots. In this case one would not necessarily have to resort to determining the various truncation levels L_j individually. However, the right part of the figure shows a time series with a significant level of dependence, in fact DGP (b) introduced in Section 5.3 below. The functional variation of this second-order autoregressive process receives drastically different contributions from different frequency bands, yielding large differences also in the spectral decompositions: sample eigenvalues plotted against frequency are far from constant. Note also how the plot of the top
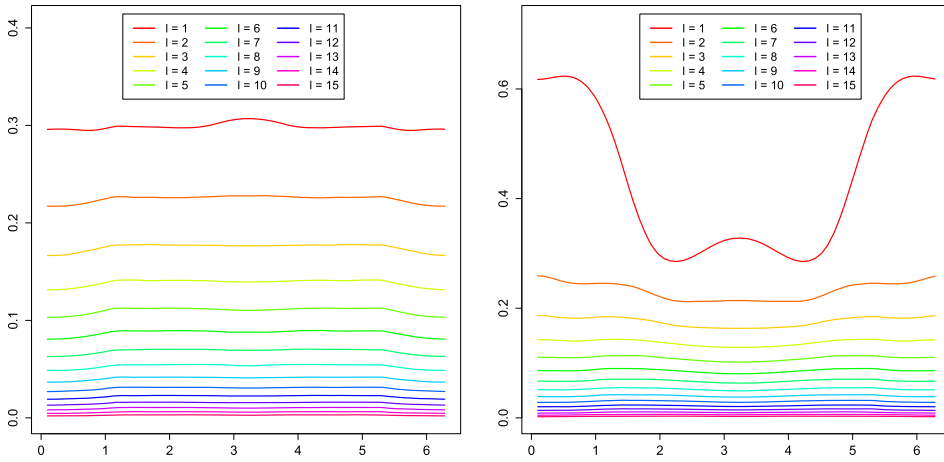
FIG. 1.   *Plot of sample eigenvalues $\lambda_l^{\omega_j}$ across Fourier frequencies ω_j for $l = 1, \ldots, 15$ for a functional white noise process (left) and a second-order functional autoregression (right).*

sample eigenvalue resembles the univariate spectral density of a scalar second-order autoregression with levels of dependence determined by the operator norm $\|\cdot\|_\infty$. Both plots taken together highlight that some flexibility in choosing the L_j is desirable.

To accommodate the previous observation, the following arrangements were made for the standardized test based on $\hat{Q}_{M,s}^{(T)}$. Firstly, a reasonable level of variation explained at each frequency ω_j is ensured through requiring that $0.5 < \text{TVE}_j < 0.9$ for all j. Secondly, the procedure adapts to different eigenvalue decays by choosing

$$L_j = \max\left\{ l : \frac{\lambda_l^{\omega_j}}{\lambda_1^{\omega_j}} > 0.2 - \frac{1}{\sqrt{bT}} \right\}$$

subject to the TVE criterion being satisfied. If no such L_j exists, choose $L_j = 1$. The unstandardized test statistics is very stable in practice and does not require the specification of tuning parameters.

Estimation of the spectral density operator and its eigenelements, needed to compute the two statistics, was achieved using (10) with the concave smoothing kernel $K(x) = 6(0.25 - x^2)$ with compact support on $x \in [-1/2, 1/2]$ and bandwidth $b = T^{-0.26}$. The fourth-order estimation is done with $K_4(x_1, \ldots, x_4) = \prod_{j=1}^4 K(x_j)$, where K is same as before, and bandwidth $b_4 = T^{-1/5}$. It should be noted that the outcomes were not overly sensitive with respect to bandwidth choices for b respecting Assumption 3.1. It is worthwhile to mention that the computational complexity of the fourth-order estimator is considerable for larger sample sizes. The implementation was therefore partially done with the compiler language C++ and the Rcpp-package in R.

5.3. *Finite sample performance under the null.*   Under the null hypothesis of stationarity the following data generating processes, DGPs, were studied:

(a) The Gaussian FWN variables $\varepsilon_1, \ldots, \varepsilon_T$ with coefficient variances $\text{Var}(\langle \varepsilon_t, \psi_l \rangle) = \exp(-(l-1)/10)$;

(b) The FAR(2) variables $X_1, \ldots, X_T$ with operators specified through the respective variances $v_{l,l'}^{(1)} = \exp(-l - l')$ and $v_{l,l'}^{(2)} = 1/(l + l'^{3/2})$ and operator norms $\kappa_1 = 0.75$ and $\kappa_2 = -0.4$, and innovations $\varepsilon_1, \ldots, \varepsilon_T$ as in (a);

(c) The FAR(2) variables $X_1, \ldots, X_T$ as in (b) but with operator norms $\kappa_1 = 0.4$ and $\kappa_2 = 0.45$.

The sample sizes under consideration are $T = 2^n$ for $n = 6, \ldots, 10$, so that the smallest sample size consists of 64 functions and the largest of 1024. The processes in (a)–(c) comprise a range of stationary scenarios. DGP (a) is the simplest model, specifying an independent FWN process. DGPs (b) and (c) exhibit significant second-order autoregressive dynamics of different persistence.

The empirical rejection levels for the processes (a)–(c) can be found in Table 1. It can be seen that the empirical levels for both statistics with $M = 1$ are generally well adjusted with slight deviations in a few cases. The performance of the statistics with $M = 5$ is similar, although the empirical rejection levels tend toward the nominal ones with increasing sample size. Some evidence on closeness between empirical and limit densities for the statistics $\hat{Q}_{5,u}^{(T)}$ and $\hat{Q}_{5,s}^{(T)}$ are provided in Figure 2.

Figure 3 shows the average choices of L over the 1000 repetitions for the various DGPs for the sample sizes $T = 64$ and $T = 1024$. First, one can see that the average L increases with the sample size, as more degrees of freedom become available. For the small sample size $T = 64$, choices of L_j under the null hypothesis are more similar both across frequencies and across the three DGPs because the form of dependence is not yet entirely evident. With increasing sample size, the average L_j increases uniformly for DGP (a), while for DGPs (b) and (c) L_j in certain frequency bands are accentuated while others are attenuated according to their contributions to the spectral analysis of variance of the underlying functional time series. For DGP (b) the shape of the curve $\omega_j \mapsto L_j$ might also be compared to the shape of the curve $\omega_j \mapsto \lambda_1^{\omega_j}$ in the right panel of Figure 1.

5.4. *Finite sample performance under the alternative.* Under the alternative the following data generating processes are considered:

(d) The tvFAR(1) variables $X_1, \ldots, X_T$ with operator specified through the variances $v_{l,l'}^{(1)} = \exp(-l - l')$ and operator norm $\kappa_1 = 0.8$ and innovations given by (a) with added multiplicative time-varying variance

$$\sigma^2(t) = \cos\left(\frac{1}{2} + \cos\left(\frac{2\pi t}{T}\right) + 0.3 \sin\left(\frac{2\pi t}{T}\right)\right);$$

(e) The tvFAR(2) variables $X_1, \ldots, X_T$ with both operators as in (d) but with time-varying operator norm

$$\kappa_{1,t} = 1.8 \cos\left(1.5 - \cos\left(\frac{4\pi t}{T}\right)\right),$$

constant operator norm $\kappa_2 = -0.81$ and innovations as in (a);

(f) The structural break FAR(2) variables $X_1, \ldots, X_T$ given in the following way:

– For $t \le 3T/8$, the operators are as in (b) but with operator norms $\kappa_1 = 0.7$ and $\kappa_2 = 0.2$, and innovations as in (a);
– For $t > 3T/8$, the operators are as in (b) but with operator norms $\kappa_1 = 0$ and $\kappa_2 = -0.2$, and innovations as in (a) but with variances $\mathrm{Var}(\langle \varepsilon_t, \psi_l \rangle) = 2 \exp(-(l-1)/10)$.

All other aspects of the simulations are as in Section 5.3. The processes studied under the alternative provide intuition for the behavior of the proposed tests under different deviations from the null hypothesis. DGP (d) is time varying only through the innovation structure, in the form of a slowly varying variance component. The first-order autoregressive structure is independent of time. DGP (e) is a time-varying second-order FAR process for which the first autoregressive operator varies with time. The final DGP in (f) models a structural break, a different type of alternative. Here, the process is not locally stationary as prescribed under

TABLE 1
Median of test statistic values and rejection rates of $\hat{Q}_{M,u}^{(T)}$ and $\hat{Q}_{M,s}^{(T)}$ at the 1% and 5% asymptotic level for the processes (a)–(c) *for various choices of M and T. All table entries are generated from* 1000 *repetitions*

| | T | $\hat{Q}_{1,u}^{(T)}$ | % level | | $\hat{Q}_{5,u}^{(T)}$ | % level | | $\hat{Q}_{1,s}^{(T)}$ | % level | | $\hat{Q}_{5,s}^{(T)}$ | % level | |
			5	1		5	1		5	1		5	1
(a)	64	1.33	5.80	1.40	8.93	9.10	2.60	1.29	4.30	1.50	8.26	7.80	2.70
	128	1.41	5.90	1.20	9.03	7.20	2.10	1.36	5.70	1.00	8.96	5.70	2.10
	256	1.26	5.10	0.90	9.15	5.30	1.70	1.27	5.20	1.40	9.02	5.10	1.00
	512	1.37	4.80	1.30	9.27	6.80	1.40	1.40	4.60	1.30	9.16	6.30	1.30
	1024	1.32	4.70	1.20	9.19	5.20	1.50	1.33	5.40	0.60	9.33	4.60	1.10
(b)	64	1.58	6.00	1.50	9.50	9.40	3.50	1.35	5.70	1.40	8.65	6.10	2.70
	128	1.44	5.70	1.60	9.35	8.90	2.80	1.30	4.70	1.50	8.72	6.30	1.70
	256	1.28	4.20	0.90	9.11	6.20	2.30	1.32	4.70	0.60	8.78	7.00	1.70
	512	1.32	5.00	1.70	9.42	6.70	1.90	1.26	4.70	0.90	9.11	6.10	0.90
	1024	1.44	4.40	0.80	9.26	5.40	1.10	1.32	4.70	0.50	8.87	4.80	0.90
(c)	64	1.42	5.60	1.90	8.50	7.60	3.30	1.20	5.70	0.90	8.36	8.20	2.60
	124	1.31	5.20	1.00	9.05	6.20	2.50	1.29	4.00	0.50	8.77	5.70	2.00
	256	1.48	6.10	1.20	9.19	6.70	1.90	1.42	5.20	1.70	8.90	6.10	1.30
	512	1.35	5.60	0.70	9.48	4.90	1.00	1.41	4.50	0.60	8.99	5.30	1.40
	1024	1.34	6.90	1.60	9.26	5.70	1.30	1.35	4.60	1.10	9.10	4.40	0.90
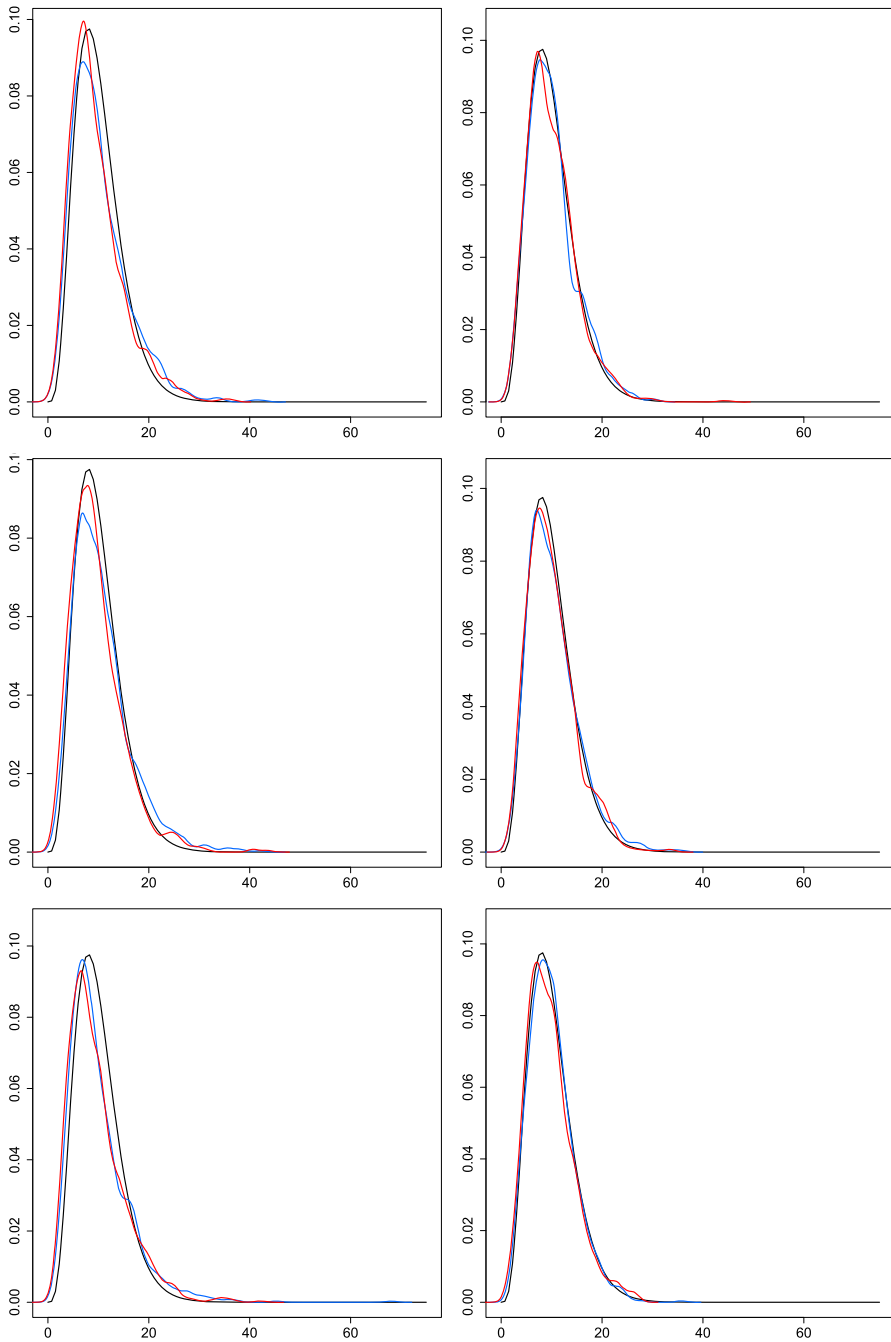
FIG. 2. *Empirical density of* $\hat{Q}_{5,u}^{(T)}$ *(black) and* $\hat{Q}_{5,s}^{(T)}$ *(blue) for* $T = 64$ *(left panel) and* $T = 512$ *(right panel) for DGPs* (a)–(c) *(top to bottom). Red: The corresponding chi-squared densities predicted under the null.*

the alternative in this paper but piecewise stationary with the two pieces being specified as two distinct FAR(2) processes.

The empirical power of the various test statistics for the processes in (d)–(f) are in Table 2. Power results are roughly similar across the selected values of M for both statistics. For DGP (f) and to some extent for DGP (d), power is low for the small sample sizes $T = 64$. It reaches 100% for all T larger or equal to 256 for all DGPs but (f) where close to perfect detection is reached for $T = 512$. Generally, the standardized statistic is slightly more unstable than its
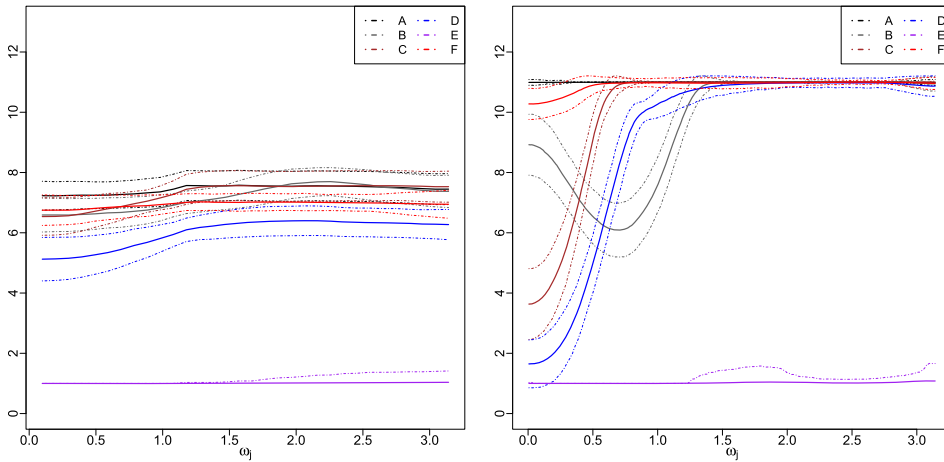
FIG. 3. *Average choice of truncation level L_j against frequency ω_j for the six DGPs (solid lines) with respective standard deviations (dashed lines) for $T = 64$ (left) and $T = 1024$ (right).*

unstandardized counterpart for DGP (e), while both statistics behave remarkably similar for the other processes. The results for DGP (f) indicate that the proposed statistics have power against structural break alternatives. This is intuitive since the second-order structure is in this case not invariant under translations of time and, hence, induces a nonzero mean in the test statistics.

Figure 4 exhibits exemplary the empirical densities for DGP (d). It can be seen that the deviation from the chi-squared distribution predicted under the null hypothesis grows with increasing sample size. Figure 3 contains the average choice of L_j for DGPs (d)–(f) under the alternative. While processes (d) and (f) display behavior more similar to the null DGPs, process (e) is significantly different, as almost always only one principal component is chosen at each frequency for both the small and the large sample size.

5.5. *Finite sample performance under non-Gaussian observations.* In this section the behavior of the eigenbased test under non-Gaussianity is further investigated through the following processes:

(g) The FAR(2) variables $X_1, \ldots, X_T$ as in (b) but with both independent t_{19}-distributed FWN and independent $\beta(6, 6)$-distributed FWN;

(h) The tvFAR(1) variables $X_1, \ldots, X_T$ as in (d) but with independent t_{19}-distributed FWN and independent $\beta(6, 6)$-distributed FWN.

For direct comparison both t_{19}- and $\beta(6, 6)$-distributions were standardized to conform to zero mean and unit variance as the standard normal. All other aspects are as detailed in Section 5.3. The additional simulations were designed to shed further light on the effect of estimating the fourth-order spectrum in situations deviating from the standard Gaussian setting. Note in particular that the t_{19}-distribution serves as an example for leptokurtosis (the excess kurtosis is 0.4) and the $\beta(6, 6)$ distribution for platykurtosis (the excess kurtosis is −0.4). Process (g) showcases the behavior under the null, while process (h) highlights the performance under the alternative. The corresponding results are given in Table 3 and can be readily compared with corresponding outcomes for the Gaussian processes (b) and (d) in Tables 1–2.

It can be seen from the results in Table 3 that the proposed procedures perform roughly as expected. First, under the null hypothesis the levels for both types of innovations, both sets

TABLE 2
Median of test statistic values and rejection rates of $\hat{Q}_{M,u}^{(T)}$ and $\hat{Q}_{M,s}^{(T)}$ at the 1% and 5% asymptotic level for the processes (d)–(f) *for various choices of M and T. All table entries are generated from* 1000 *repetitions*

|	T	$\hat{Q}_{1,u}^{(T)}$	% level 5	1	$\hat{Q}_{5,u}^{(T)}$	% level 5	1	$\hat{Q}_{1,s}^{(T)}$	% level 5	1	$\hat{Q}_{5,s}^{(T)}$	% level 5	1
(d)	64	9.84	77.80	54.30	20.33	57.30	39.80	8.61	71.30	46.70	17.92	48.80	30.70
	128	19.55	99.00	94.40	33.34	94.10	84.10	18.26	98.20	91.40	30.44	90.20	76.30
	256	36.70	100.00	100.00	54.07	99.90	99.70	34.40	100.00	100.00	50.27	99.80	99.40
	512	69.49	100.00	100.00	94.47	100.00	100.00	62.90	100.00	100.00	84.75	100.00	100.00
	1024	140.53	100.00	100.00	179.75	100.00	100.00	118.18	100.00	100.00	152.12	100.00	100.00
(e)	64	33.38	100.00	100.00	131.80	100.00	98.10	33.46	99.50	99.20	100.13	99.30	99.20
	128	49.04	100.00	100.00	118.13	100.00	100.00	66.48	99.70	99.30	172.30	99.80	99.70
	256	98.43	100.00	100.00	393.65	100.00	100.00	151.44	99.70	99.60	568.55	99.90	99.80
	512	173.35	100.00	100.00	763.11	100.00	100.00	302.51	100.00	100.00	1257.93	100.00	100.00
	1024	286.54	99.90	99.90	1311.08	100.00	100.00	579.00	99.80	99.80	2484.54	100.00	99.90
(f)	64	5.64	46.50	25.40	15.02	33.70	19.90	4.38	35.20	16.50	12.36	24.40	12.70
	128	10.90	82.80	60.90	21.65	64.30	43.20	8.93	83.10	48.40	18.37	50.40	29.30
	256	18.29	98.20	90.50	30.40	90.00	77.50	15.71	95.70	85.20	27.03	84.70	66.00
	512	31.81	100.00	100.00	47.49	99.90	99.20	30.71	99.90	99.80	45.71	99.80	98.50
	1024	62.72	100.00	100.00	83.82	100.00	100.00	62.29	100.00	100.00	83.18	100.00	100.00
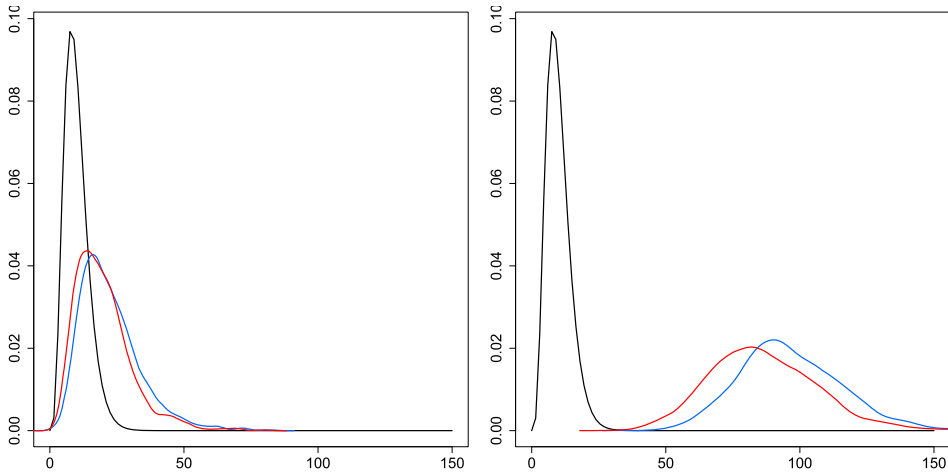
FIG. 4. *Empirical density of $\hat{Q}_{5,u}^{(T)}$ (black) and $\hat{Q}_{5,s}^{(T)}$ (blue) for $T = 64$ (left panel) and $T = 512$ (right panel) for DGP (d).*

of tests and both choices of M are well adjusted and observe similar patterns as their normal counterparts in DGP (b) in Table 1. Second, under the alternative for process (h), powers align roughly as for the Gaussian case in Table 2. Overall, the simulation results reveal that the estimation of the fourth-order spectrum does not lead to a marked decay in performance.

5.6. *Application to annual temperature curves.*   To give an instructive data example, the proposed method was applied to annual temperature curves recorded at several measuring stations across Australia over the last century and a half. The exact locations and lengths of the functional time series are reported in Table 4, and the annual temperature profiles recorded at the Gayndah station are displayed for illustration in the left panel of Figure 5. To test whether these annual temperature profiles constitute stationary functional time series or not, the proposed testing method was utilized, using specifications similar to those in the simulation study. To get an idea of the spectral structure of these different temperature curves, the left-hand side of Figure 6 shows the averaged eigenvalue decay standardized with respect to the largest eigenvalue at each frequency. More precisely, $\frac{1}{T}\sum_{j=1}^{T} \lambda_l^{\omega_j}/\lambda_1^{\omega_j}$ is plotted against l. Figure 7 displays, in addition, the plots of the l-largest sample eigenvalues $\lambda_l^{\omega_j}$ against j for $l = 1, \ldots, 15$. It can be seen that frequency-specific contributions are heterogeneous for each of the four stations. There are also substantial differences in the eigenvalue plots across different stations. The choices of L_j across frequency ω_j as used by the standardized test procedure are shown in the right-hand side of Figure 6.

The p-values for the standardized test statistics are essentially zero for all stations and all $M = 1, \ldots, 5$. The testing results for the unstandardized statistics are summarized in Table 4. Stationarity is rejected in favor of the alternative at the 1% significance level at all measuring stations for $\hat{Q}_{M,u}^{(T)}$ with all specifications of M, with one notable exception; no choice of M leads to a rejection of the null hypothesis at Boulia station. Additionally, rejection at Melbourne and Sydney stations is not possible at the smallest significance levels for several M. At all other measuring stations rejection of the null is very strong. Note that Boulia station showed the slowest eigenvalue decay in Figure 6 and the spectral behavior most different from the other stations in Figure 7. It is particularly interesting that around frequency π there is little to no separation between first and second sample eigenvalues. The lack of estimation accuracy in the case of tied eigenvalues might help explain why Boulia station delivers results

TABLE 3
Median of test statistic values and rejection rates of $\hat{Q}_{M,u}^{(T)}$ and $\hat{Q}_{M,s}^{(T)}$ at the 1% and 5% asymptotic level for the processes (g) *and* (h), *where t and β indicate t_{19}- and $\beta(6,6)$-distributed innovations, respectively. All table entries are generated from* 1000 *repetitions*

		% level			% level			% level			% level		
	T	$\hat{Q}_{1,u}^{(T)}$	5	1	$\hat{Q}_{5,u}^{(T)}$	5	1	$\hat{Q}_{1,s}^{(T)}$	5	1	$\hat{Q}_{5,s}^{(T)}$	5	1
(g), t	64	1.58	5.60	0.70	9.46	7.70	2.70	1.40	2.90	0.60	8.65	6.20	1.40
	128	1.42	4.50	1.30	9.37	6.90	2.00	1.30	3.60	0.40	8.81	4.80	1.30
	256	1.40	4.40	0.80	9.17	5.50	1.70	1.29	4.90	0.90	8.89	5.70	1.20
	512	1.47	4.70	0.70	9.33	4.70	1.30	1.44	4.10	0.50	9.32	4.30	1.20
	1024	1.53	5.90	0.40	9.52	5.00	1.00	1.43	5.40	0.90	8.92	4.70	1.10
(g), β	64	1.31	3.10	0.80	9.00	6.70	1.60	1.29	3.40	0.80	8.66	5.50	1.10
	128	1.37	4.80	1.10	9.13	6.10	1.90	1.25	3.70	0.60	8.89	4.50	0.90
	256	1.39	4.70	1.00	9.13	4.10	1.30	1.32	4.70	1.10	8.57	3.40	0.60
	512	1.30	3.90	0.70	9.22	4.50	1.00	1.32	4.50	0.90	9.13	5.80	1.40
	1024	1.43	5.00	0.90	9.57	4.40	0.80	1.34	4.20	1.20	9.37	4.90	0.70
(h), t	64	9.07	77.10	49.00	18.95	53.10	29.10	8.16	69.30	41.30	16.82	43.50	22.10
	128	17.21	98.30	92.80	28.78	91.10	74.30	16.47	98.00	89.70	26.52	87.00	66.70
	256	31.12	100.00	99.90	45.94	100.00	99.70	30.12	100.00	99.70	43.54	99.70	98.60
	512	57.81	100.00	100.00	78.95	100.00	100.00	53.14	100.00	100.00	71.69	100.00	100.00
	1024	112.95	100.00	100.00	146.21	100.00	100.00	98.88	100.00	100.00	127.16	100.00	100.00
(h), β	64	9.17	77.80	49.60	18.30	50.00	28.10	8.20	69.40	40.20	16.86	42.60	21.60
	128	17.49	98.40	91.40	29.06	91.30	74.10	16.31	97.20	88.50	27.21	87.00	67.70
	256	31.05	100.00	100.00	46.75	99.90	99.30	29.58	100.00	100.00	44.06	99.90	98.60
	512	57.90	100.00	100.00	78.97	100.00	100.00	52.97	100.00	100.00	71.40	100.00	100.00
	1024	114.13	100.00	100.00	146.75	100.00	100.00	100.95	100.00	100.00	128.45	100.00	100.00

TABLE 4
Summary of results for eight Australian measuring stations. The column labeled T reports the sample size, the other columns report p-values for the given choices of M for $\hat{Q}_{M,u}^{(T)}$

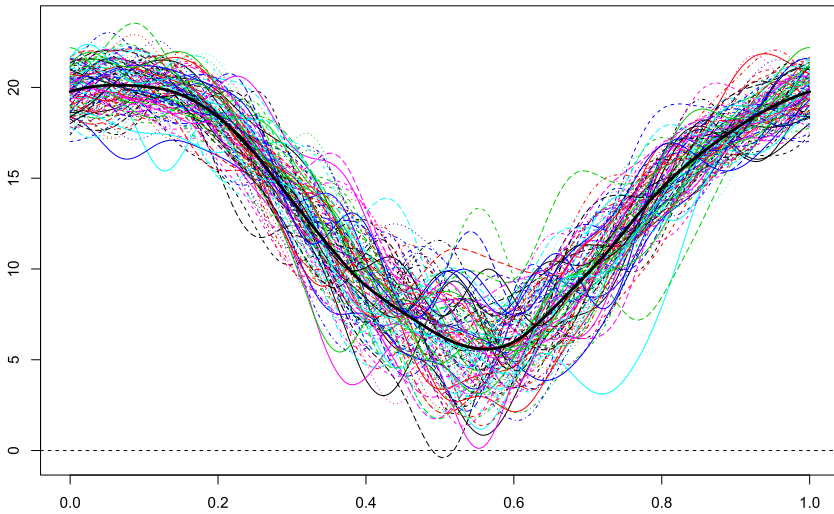| Station	T	M = 1	M = 2	M = 3	M = 4	M = 5
Boulia	120	0.71	0.17	0.20	0.36	0.44
Robe	130	0.01	0.00	0.00	0.00	0.00
Cape Otway	150	0.00	0.00	0.00	0.00	0.00
Gayndah	118	0.00	0.00	0.00	0.00	0.00
Gunnedah	134	0.00	0.00	0.00	0.00	0.00
Hobart	122	0.00	0.00	0.00	0.00	0.00
Melbourne	158	0.03	0.04	0.02	0.01	0.01
Sydney	154	0.15	0.01	0.00	0.00	0.00

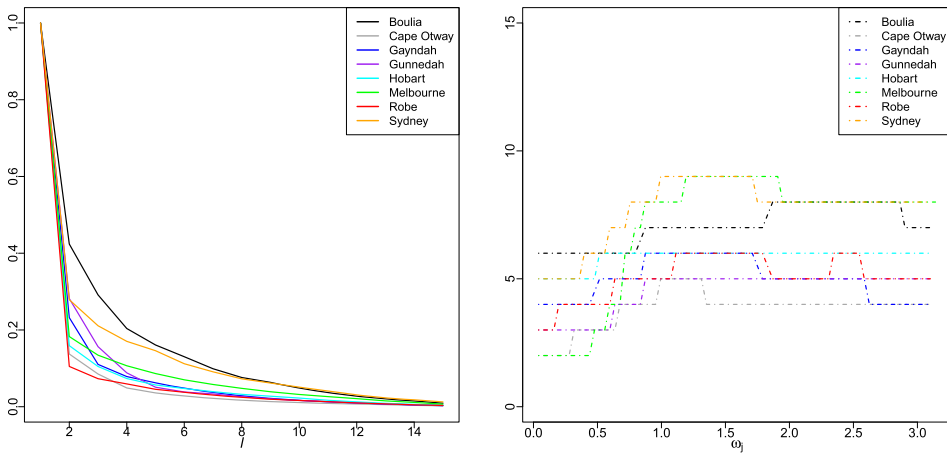FIG. 5. *Annual temperature curves at Gayndah station.*

FIG. 6. *Average eigenvalue decay standardized with respect to the largest eigenvalue at each frequency,* $\frac{1}{T}\sum_{j=1}^{T}\lambda_l^{\omega_j}/\lambda_1^{\omega_j}$ *(left) and truncation level* L_j *across* ω_j *(right) across different measuring stations.*
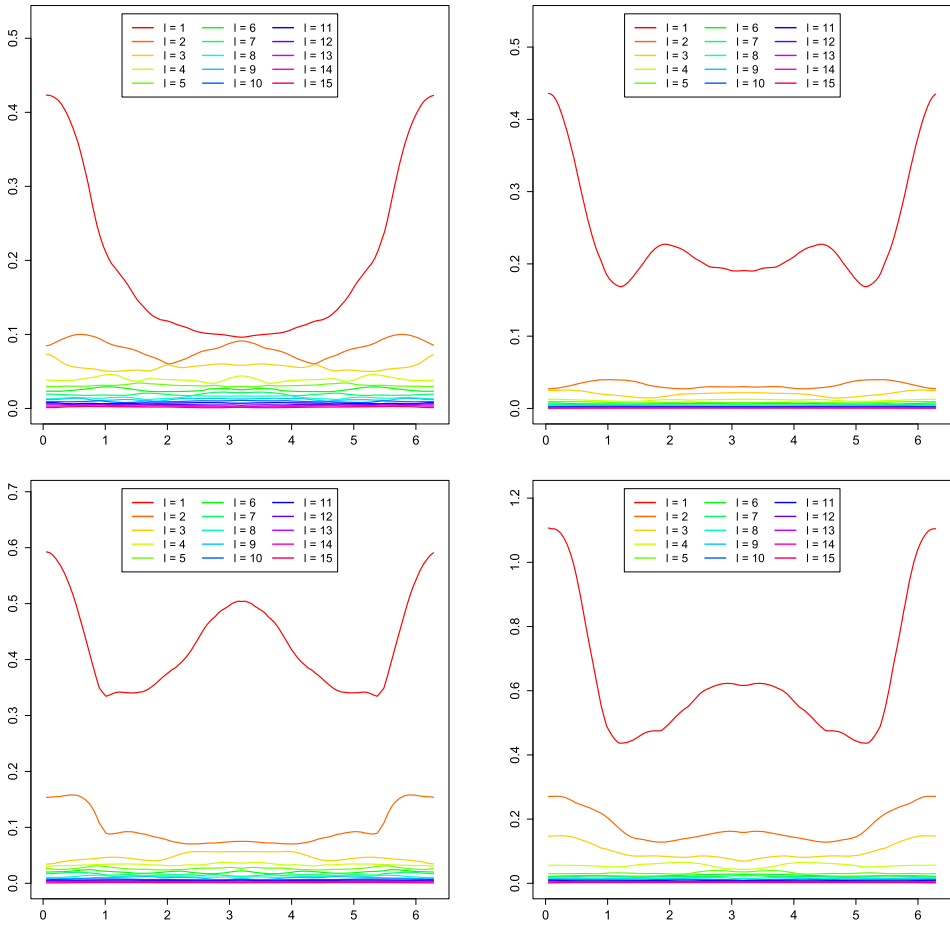
FIG. 7. *Plots of the* 15 *largest sample sample eigenvalues across Fourier frequencies at Boulia (top left), Cape Otway (top right), Gayndah (bottom left) and Gunnedah (bottom right) stations.*

at odds with the findings at the other stations. In the future it might be worthwhile looking into running the stationarity tests only in certain frequency bands, excluding those frequencies for which separation of sample eigenvalues is not sufficiently large. This is, however, beyond the scope of the current paper.

6. Conclusions and future work. In this paper methodology for testing the stationarity of a functional time series is put forward. The tests are based on frequency domain analysis and exploit that fDFTs at different canonical frequencies are uncorrelated if and only if the underlying functional time series are stationary. The limit distribution of the quadratic form-type test statistics has been determined under the null hypothesis as well as under the alternative of local stationarity. Finite sample properties were highlighted in simulation experiments with various data-generating processes and an application to annual temperature profiles.

The empirical results show promise for further applications to real data, but future research has to be devoted to a further fine-tuning of the proposed method; for example, an automated selection of frequencies h_m outside of the standard choice $h_m = m$ for all $m = 1, \dots, M$. This can be approached through a more refined analysis of the size of the various $\hat{\beta}_{h_m,x}^{(T)}$ in (11) whose real and imaginary part make up the vector $\hat{\mathbf{b}}_{M,x}^{(T)}$ in the test statistics $\hat{Q}_{M,x}^{(T)}$.

APPENDIX A: A FUNCTIONAL CRAMÉR REPRESENTATION

PROOF OF PROPOSITION 2.1. Let $X = (X_t : t \in \mathbb{Z})$ be a zero-mean, weakly stationary H-valued stochastic process. Then X admits the functional Cramér representation in (4) (see van Delft and Eichler (2019), Theorem 4.5). Conversely,

$$\mathrm{Cov}(X_t, X_s) = \mathbb{E}\left[\int_{-\pi}^{\pi} e^{it\lambda_1} dZ_{\lambda_1} \otimes \int_{-\pi}^{\pi} e^{is\lambda_2} dZ_{\lambda_2}\right] = \int_{-\pi}^{\pi} e^{i(t-s)\lambda} dF(\lambda) = \mathcal{C}_{t-s},$$

where F denotes the operator-valued measure induced by the sequence of covariance operators of X (see van Delft and Eichler (2019), Theorem 4.4), which shows that a process that admits representation (4) must be weakly stationary. $\square$

APPENDIX B: PROPERTIES OF FUNCTIONAL CUMULANTS

For random elements $X_1, \ldots, X_k$ in a Hilbert space H, the *moment tensor of order k* can be defined as

$$\mathbb{E}[X_1 \otimes \cdots \otimes X_k] = \sum_{l_1, \ldots, l_k \in \mathbb{N}} \mathbb{E}\left[\prod_{t=1}^{k} \langle X_t, \psi_{l_t} \rangle\right](\psi_{l_1} \otimes \cdots \otimes \psi_{l_k}),$$

where the elementary tensors $(\psi_{l_1} \otimes \cdots \otimes \psi_{l_k} : l_1, \ldots, l_k \in \mathbb{N})$ form an orthonormal basis in the tensor product space $\bigotimes_{j=1}^{k} H$ if $(\psi_l : l \in \mathbb{N})$ is an orthornormal basis of the separable Hilbert space H. Similarly, define the *kth order cumulant tensor* by

$$
\begin{aligned}
& \mathrm{cum}(X_1, \ldots, X_k) \\
(21) \quad & = \sum_{l_1, \ldots, l_k \in \mathbb{N}} \mathrm{cum}(\langle X_1, \psi_{l_1} \rangle, \ldots, \langle X_k, \psi_{l_k} \rangle)(\psi_{l_1} \otimes \cdots \otimes \psi_{l_k}),
\end{aligned}
$$

where the cumulants on the right-hand side are as usual given by

$$
\begin{aligned}
& \mathrm{cum}(\langle X_1, \psi_{l_1} \rangle, \ldots, \langle X_k, \psi_{l_k} \rangle) \\
& = \sum_{v = (v_1, \ldots, v_p)} (-1)^{p-1}(p-1)! \prod_{r=1}^{p} \mathbb{E}\left[\prod_{t \in v_r} \langle X_t, \psi_{l_t} \rangle\right],
\end{aligned}
$$

the summation extending over all unordered partitions v of $\{1, \ldots, k\}$. The following is a generalization of the product theorem for cumulants (Brillinger (1981), Theorem 2.3.2):

THEOREM B.1. *Consider the tensor $X_t = \bigotimes_{j=1}^{J_t} X_{tj}$ for random elements X_{tj} in H with $j = 1, \ldots, J_t$ and $t = 1, \ldots, k$. Let $v = \{v_1, \ldots, v_p\}$ be a partition of $\{1, \ldots, k\}$. The joint cumulant tensor is given by*

$$
\begin{aligned}
& \mathrm{cum}(X_1, \ldots, X_k) \\
& = \sum_{r_{11}, \ldots, r_{kJ_t}} \sum_{v = (v_1, \ldots, v_p)} \prod_{n=1}^{p} \mathrm{cum}(\langle X_{tj}, \psi_{r_{tj}} \rangle | (t, j) \in v_n)\psi_{r_{11}} \otimes \cdots \otimes \psi_{r_{kJ_t}},
\end{aligned}
$$

where the summation extends over all indecomposable partitions $v = (v_1, \ldots, v_p)$ of the table

$$
\begin{matrix}
(1, 1) & \cdots & (1, J_1) \\
\vdots & \ddots & \vdots \\
(k, 1) & \cdots & (k, J_t).
\end{matrix}
$$

Formally, abbreviate this by

$$\mathrm{cum}(X_1,\ldots,X_k) = \sum_{\nu=(\nu_1,\ldots,\nu_p)} S_\nu\left(\bigotimes_{n=1}^{p} \mathrm{cum}(X_{tj}|(t,j)\in\nu_n)\right),$$

where S_ν is the permutation that maps the components of the tensor back into the original order, that is, $S_\nu(\bigotimes_{r=1}^{p}\bigotimes_{(t,j)\in\nu_r} X_{tj}) = X_{11}\otimes\cdots\otimes X_{kJ_t}$.

Next, expressions and bounds for cumulants of the fDFT are given in both locally stationary and stationary regimes.

LEMMA B.1 (Cumulants of the fDFT under local stationarity). *Let $(X_{t,T}: t\leq T, T\in\mathbb{N})$ be a kth order locally stationary process in H satisfying Assumption* $\mathrm{I}(k,1)$ *for arbitrary fixed k. The cumulant tensor of the local fDFT satisfies*

$$\mathrm{cum}(D_{\omega_{j_1}}^{(T)},\ldots,D_{\omega_{j_k}}^{(T)})$$

(22)
$$= \frac{(2\pi)^{k/2-1}}{T^{k/2}}\sum_{t=0}^{T-1}\mathcal{F}_{t/T;\omega_{j_1},\ldots,\omega_{j_{k-1}}}e^{-\mathrm{i}\sum_{l=1}^{k}t\omega_{j_l}} + R_{k,T}$$

$$= \frac{(2\pi)^{k/2-1}}{T^{k/2-1}}\tilde{\mathcal{F}}_{j_1+\cdots+j_k;\omega_{j_1},\ldots,\omega_{j_{k-1}}} + R_{k,T},$$

where $\|R_{k,T}\|_2 = O(T^{-k/2})$ and the operator

(23)
$$\tilde{\mathcal{F}}_{s;\omega_{j_1},\ldots,\omega_{j_{k-1}}} = \int_0^1 \mathcal{F}_{u;\omega_{j_1},\ldots,\omega_{j_{k-1}}}e^{-\mathrm{i}2\pi su}\,du$$

denotes the sth Fourier coefficient of $\mathcal{F}_{u;\omega_{j_1},\ldots,\omega_{j_{k-1}}}$ and belongs to S_2.

The proof can be found in Section S2 of the Online Supplement. Lemma B.1 provides a relation between the kth order cumulant tensor of the local fDFT and the Fourier coefficients of the kth order time-varying spectral density tensors which induce Hilbert–Schmidt operators. The proof of (23) makes apparent that the dependence structure of the local fDFT behaves in a very specific manner that is based on the distance of the frequencies. The Fourier coefficients additionally provide an upper bound on the norm of the cumulant operator.

COROLLARY B.1. *If Assumption* $\mathrm{I}(k,2)$ *holds for arbitrary fixed k, then:*

(i) $\|\mathrm{cum}(D_{\omega_{j_1}}^{(T)},\ldots,D_{\omega_{j_k}}^{(T)})\|_2 \leq \dfrac{C}{T^{k/2-1}|j_1+\cdots+j_k|^2} + O\left(\dfrac{1}{T^{k/2}}\right);$

(ii) $\sup\limits_{\omega}\sum\limits_{s\in\mathbb{Z}}\|\tilde{\mathcal{F}}_{s;\omega}\|_2 \leq \infty.$

Note that if $\sum_{l=1}^{k}\omega_{j_l} = 0 \mod 2\pi$, then (22) yields approximately a time average of the kth order time-varying spectral density tensor. In case the process does not depend on time u, $\tilde{\mathcal{F}}_{s;\omega_{j_1},\ldots,\omega_{j_{2k-1}}} = O_H$ for $s\neq 0$. That is, the operator $\tilde{\mathcal{F}}_{s;\omega_{j_1},\ldots,\omega_{j_{2k-1}}}$ maps any $\psi\in L^2([0,1]^k,\mathbb{C})$ to the origin for $s\neq 0$. Consequently, under kth order stationarity the following corollary holds.

COROLLARY B.2 (Cumulants of the fDFT under stationarity). *Let $(X_t: t\in\mathbb{Z})$ be a kth order stationary sequence taking values in $H_\mathbb{R}$ that satisfies Assumption* $\mathrm{I}^*(k,1)$ *for arbitrary fixed k. Then, the cumulant tensor of the fDFT satisfies*

(24)
$$\mathrm{cum}(D_{\omega_{j_1}}^{(T)},\ldots,D_{\omega_{j_k}}^{(T)}) = \frac{(2\pi)^{k/2-1}}{T^{k/2}}\Delta_T^{(\sum_{l=1}^{k}\omega_{j_l})}\mathcal{F}_{\omega_{j_1},\ldots,\omega_{j_{k-1}}} + R_{T,k},$$

where the function $\Delta_T^{(\omega)} = T$ for $\omega \equiv 0 \mod 2\pi$, $\Delta_T^{(\omega_j)} = 0$ for $j \not\equiv 0 \mod T$ and the remainder satisfies $\|R_{T,k}\|_2 = O(T^{-k/2})$.

APPENDIX C: FIRST- AND SECOND-ORDER DEPENDENCE STRUCTURE

C.1. Expectation. From Lemma B.1, for $h \neq 0 \mod T$,

$$\left\| \frac{1}{T} \sum_{j=1}^{T} \mathbb{E}(D_{\omega_j}^{(T)} \otimes D_{\omega_{j+h}}^{(T)}) \right\|_2 = \left\| \frac{1}{T} \sum_{j=1}^{T} \frac{1}{T} \sum_{t=1}^{T} \mathcal{F}_{t/T;\omega_j} e^{-it\omega_h} + R_{T,2} \right\|_2$$

$$= \begin{cases} O(T^{-1}) & \text{under } H_0, \\ O(h^{-2} + T^{-1}) & \text{under } H_A. \end{cases}$$

In particular, using that the operator-valued functions $(u, \omega) \mapsto \mathcal{F}_{u,\omega}$ are Lipschitz continuous in (u, ω), yields that, under H_A,

$$\frac{1}{T} \sum_{j=1}^{T} \frac{1}{T} \sum_{t=1}^{T} \mathcal{F}_{t/T;\omega_j} e^{-it\omega_h} \to \frac{1}{2\pi} \int_0^{2\pi} \int_0^1 \mathcal{F}_{u,\omega} e^{-i2\pi uh} \, du \, d\omega,$$

where the convergence is in $S_2(H)$. Since $\mathbb{E}[\|D_\omega^{(T)}\|_2^2] < \infty$, the Cauchy–Schwarz inequality implies Fubini's theorem can be applied. Together with the above, it follows that the expectation of $\beta_{h,u}^{(T)}$ satisfies

$$\mathbb{E}[\beta_{h,u}^{(T)}] = \frac{1}{T} \sum_{j=1}^{T} \sum_{l=1}^{L(\omega_j)} \sum_{l'=1}^{L(\omega_{j+h})} \langle \mathbb{E}(D_{\omega_j}^{(T)} \otimes D_{\omega_{j+h}}^{(T)}), \tilde{\phi}_l^{\omega_j} \otimes \tilde{\phi}_{l'}^{\omega_{j+h}} \rangle_S$$

$$= \frac{1}{T} \sum_{j=1}^{T} \sum_{l=1}^{L(\omega_j)} \sum_{l'=1}^{L(\omega_{j+h})} \left\langle \frac{1}{T} \sum_{t=0}^{T-1} \mathcal{F}_{t/T,\omega_j} e^{-it\omega_h} + R_{2,T}, \tilde{\phi}_l^{\omega_j} \otimes \tilde{\phi}_{l'}^{\omega_{j+h}} \right\rangle_S$$

$$= O\left(\frac{1}{h^2}\right) + O\left(\frac{1}{T}\right)$$

$$\to \frac{1}{2\pi} \int_0^{2\pi} \int_0^1 \sum_{l=1}^{L(\omega)} \sum_{l'=1}^{L(\omega+\omega_h)} \langle \mathcal{F}_{u,\omega} e^{-i2\pi uh}, \tilde{\phi}_l^{\omega} \otimes \tilde{\phi}_{l'}^{\omega+\omega_h} \rangle_S \, du \, d\omega,$$

where the stated order $O(\cdot)$ for the projections follows from the previously stated convergence in norm. Similarly,

$$\mathbb{E}[\beta_{h,s}^{(T)}] \to \frac{1}{2\pi} \int_0^{2\pi} \int_0^1 \sum_{l=1}^{L(\omega)} \sum_{l'=1}^{L(\omega+\omega_h)} \frac{\langle \mathcal{F}_{u,\omega}(\tilde{\phi}_{l'}^{\omega+\omega_h}), \tilde{\phi}_l^{\omega} \rangle e^{-i2\pi uh}}{\sqrt{\tilde{\lambda}_l^{\omega} \tilde{\lambda}_{l'}^{\omega+\omega_h}}} \, du \, d\omega,$$

under Condition C_s.

C.2. Covariance structure. Theorem B.1 implies that the covariance structure of the cross-periodogram operators is given by

$$\text{Cov}(D_{\omega_{j_1}}^{(T)} \otimes D_{\omega_{j_1+h_1}}^{(T)}, D_{\omega_{j_2}}^{(T)} \otimes D_{\omega_{j_2+h_2}}^{(T)})$$

(25)
$$= \text{cum}(D_{\omega_{j_1}}^{(T)}, D_{-\omega_{j_1+h_1}}^{(T)}, D_{-\omega_{j_2}}^{(T)}, D_{\omega_{j_2+h_2}}^{(T)})$$

$$+ S_{1324}(\text{cum}(D_{\omega_{j_1}}^{(T)}, D_{-\omega_{j_2}}^{(T)}) \otimes \text{cum}(D_{-\omega_{j_1+h_1}}^{(T)}, D_{\omega_{j_2+h_2}}^{(T)}))$$

$$+ S_{1423}(\text{cum}(D_{\omega_{j_1}}^{(T)}, D_{\omega_{j_2+h_2}}^{(T)}) \otimes \text{cum}(D_{-\omega_{j_1+h_1}}^{(T)}, D_{-\omega_{j_2}}^{(T)})),$$

where S_{ijkl} denotes the permutation operator on $\bigotimes_{i=1}^4 L_{\mathbb{C}}^2([0,1])$ that permutes the components of a tensor according to the permutation $(1,2,3,4) \mapsto (i,j,k,l)$, that is, $S_{ijkl}(x_1 \otimes \cdots \otimes x_4) = x_i \otimes \cdots \otimes x_l$. Under Assumption I(4, 2) we obtain from Lemma B.1,

$$
\begin{aligned}
&\mathrm{Cov}\left(\frac{1}{\sqrt{T}} \sum_{j_1}^T D_{\omega_{j_1}}^{(T)} \otimes D_{\omega_{j_1+h_1}}^{(T)}, \frac{1}{\sqrt{T}} \sum_{j_2}^T D_{\omega_{j_2}}^{(T)} \otimes D_{\omega_{j_2+h_2}}^{(T)}\right) \\
&= \frac{1}{T} \sum_{j_1,j_2}^T \left(\frac{(2\pi)}{T^2} \sum_t \mathcal{F}_{t/T:\omega_{j_1},-\omega_{j_1+h_1},-\omega_{j_2}} e^{-it(\omega_{h_2}-\omega_{h_1})} + R_{T,4}\right) \\
&\quad + S_{1324}\left(\left(\frac{1}{T} \sum_t \mathcal{F}_{t/T:\omega_{j_1}} e^{-it(\omega_{j_1}-\omega_{j_2})} + R_{T,2}\right) \right. \\
&\qquad\qquad \left. \otimes \left(\frac{1}{T} \sum_t \mathcal{F}_{t/T:-\omega_{j_1+h_1}} e^{-it(-\omega_{j_1+h_1}+\omega_{j_2+h_2})} + R_{T,2}\right)\right) \\
&\quad + S_{1423}\left(\left(\frac{1}{T} \sum_t \mathcal{F}_{t/T:\omega_{j_1}} e^{-it(\omega_{j_1}+\omega_{j_2+h_2})} + R_{T,2}\right) \right. \\
&\qquad\qquad \left. \otimes \left(\frac{1}{T} \sum_t \mathcal{F}_{t/T:-\omega_{j_1+h_1}} e^{-it(-\omega_{j_1+h_1}-\omega_{j_2})} + R_{T,2}\right)\right).
\end{aligned}
$$
(26)

Using Minkowski's inequality and Corollary B.1(ii), it follows that, for all T, h_1, h_2,

$$
(27) \qquad \left\|\mathrm{Cov}\left(\frac{1}{\sqrt{T}} \sum_{j_1}^T D_{\omega_{j_1}}^{(T)} \otimes D_{\omega_{j_1+h_1}}^{(T)}, \frac{1}{\sqrt{T}} \sum_{j_2}^T D_{\omega_{j_2}}^{(T)} \otimes D_{\omega_{j_2+h_2}}^{(T)}\right)\right\|_2 = O(1),
$$

both under H_A and H_0. The focus is here on the covariance structure under fourth-order stationarity. The more general expression is derived in Section S6 of the Online Supplement.

PROOF OF THEOREM 4.3. Under Assumption I*(4, 2), Corollary B.2 implies that (26) becomes

$$
\begin{aligned}
&\mathrm{Cov}\left(\frac{1}{\sqrt{T}} \sum_{j_1}^T D_{\omega_{j_1}}^{(T)} \otimes D_{\omega_{j_1+h_1}}^{(T)}, \frac{1}{\sqrt{T}} \sum_{j_2}^T D_{\omega_{j_2}}^{(T)} \otimes D_{\omega_{j_2+h_2}}^{(T)}\right) \\
&= \frac{1}{T} \sum_{j_1,j_2}^T \left(\frac{(2\pi)}{T^2} \mathcal{F}_{\omega_{j_1},-\omega_{j_1+h_1},-\omega_{j_2}} \Delta_T^{(\omega_{h_2}-\omega_{h_1})} + R_{T,4}\right) \\
&\quad + S_{1324}\left(\left(\mathcal{F}_{\omega_{j_1}} \frac{1}{T} \Delta_T^{(\omega_{j_1}-\omega_{j_2})} + R_{T,2}\right) \right. \\
&\qquad\qquad \left. \otimes \left(\mathcal{F}_{-\omega_{j_1+h_1}} \frac{1}{T} \Delta_T^{(-\omega_{j_1+h_1}+\omega_{j_2+h_2})} + R_{T,2}\right)\right) \\
&\quad + S_{1423}\left(\left(\mathcal{F}_{\omega_{j_1}} \frac{1}{T} \Delta_T^{(\omega_{j_1}+\omega_{j_2+h_2})} + R_{T,2}\right) \right. \\
&\qquad\qquad \left. \otimes \left(\mathcal{F}_{-\omega_{j_1+h_1}} \frac{1}{T} \Delta_T^{(-\omega_{j_1+h_1}-\omega_{j_2})} + R_{T,2}\right)\right).
\end{aligned}
$$

By the properties of $\Delta_T^{(\cdot)}$, the term on the second line is of lower order unless $h_1 - h_2 = 0$ mod T, while the third line requires $j_1 - j_2 = 0$ mod T and $h_1 - h_2 = 0$ mod T. For the fourth line to not be of lower order we require $j_1 + j_2 + h_2 = 0$ mod T and $-j_1 - h_1 - j_2 = 0$

mod T which give the constraints $j_1 + j_2 = T - h_2$ and $j_1 + j_2 = T - h_1$, implying we must have $j_1 + j_2 = T - h$. It follows therefore that the covariance is of order $O(T^{-1})$ in Hilbert–Schmidt norm if $h_1 - h_2 \neq 0 \mod T$. If $h_1 - h_2 = 0 \mod T$, then

(28)
$$
\begin{aligned}
\operatorname{Cov}&\left(\frac{1}{\sqrt{T}} \sum_{j_1}^{T} D_{\omega_{j_1}}^{(T)} \otimes D_{\omega_{j_1+h}}^{(T)}, \frac{1}{\sqrt{T}} \sum_{j_2}^{T} D_{\omega_{j_2}}^{(T)} \otimes D_{\omega_{j_2+h}}^{(T)} \right) \\
&= \frac{1}{T} \sum_{j_1, j_2}^{T} \frac{(2\pi)}{T} \mathcal{F}_{\omega_{j_1}, -\omega_{j_1+h}, -\omega_{j_2}} + R_{T,2} \\
&\quad + \frac{1}{T} \sum_{j_1}^{T} \left((\mathcal{F}_{\omega_{j_1}} + R_{T,2}) \tilde{\otimes} (\mathcal{F}_{\omega_{j_1+h}} + R_{T,2}) \right. \\
&\qquad \left. + (\mathcal{F}_{\omega_{j_1}} + R_{T,2}) \tilde{\otimes}_T (\mathcal{F}_{\omega_{j_1+h_1}} + R_{T,2}) \right),
\end{aligned}
$$

where Definition S1.1 was used. Thus, as $T \to \infty$, this converges in norm to

$$
\frac{1}{4\pi} \int \int \mathcal{F}_{\omega, -\omega - \omega_h, -\omega'} \, d\omega \, d\omega' + \int \mathcal{F}_\omega \tilde{\otimes} \mathcal{F}_{\omega + \omega_h} + \mathcal{F}_\omega \tilde{\otimes}_T \mathcal{F}_{\omega + \omega_h} \, d\omega.
$$

Consider then the covariance structure of $\hat{\boldsymbol{\beta}}_{h,u}^{(T)}$ which is obtained by projecting the fDFT onto the eigenfunctions of $\mathcal{F}_\omega$. Write this covariance structure as

$$
\begin{aligned}
\operatorname{Cov}&(\sqrt{T} \boldsymbol{\beta}_{h_1,u}^{(T)}, \sqrt{T} \boldsymbol{\beta}_{h_2,u}^{(T)}) \\
&= \frac{1}{T} \sum_{j_1, j_2}^{T} \sum_{\substack{l_1 \in [L(\omega_{j_1})], l_2 \in [L(\omega_{j_1+h_1})], \\ l_3 \in [L(\omega_{j_2})], l_4 \in [L(\omega_{j_2+h_2})]}} \langle \operatorname{Cov}(D_{\omega_{j_1}}^{(T)} \otimes D_{\omega_{j_1+h}}^{(T)}, D_{\omega_{j_2}}^{(T)} \otimes D_{\omega_{j_2+h}}^{(T)}) \\
&\quad \times (\phi_{l_3}^{\omega_{j_2}} \otimes \phi_{l_4}^{\omega_{j_2+h_2}}), \phi_{l_1}^{\omega_{j_1}} \otimes \phi_{l_2}^{\omega_{j_1+h_1}} \rangle.
\end{aligned}
$$

Under the conditions of Theorem 4.3, (28) yields that the summand of the above expression reduces to

$$
\begin{aligned}
&= \frac{(2\pi)}{T^2} \langle \mathcal{F}_{\omega_{j_1}, -\omega_{j_1+h_1}, -\omega_{j_2}} (\phi_{l_3}^{\omega_{j_2}} \otimes \phi_{l_4}^{\omega_{j_2+h_2}}), \phi_{l_1}^{\omega_{j_1}} \otimes \phi_{l_2}^{\omega_{j_1+h_1}} \rangle \Delta_T^{(\omega_{h_2} - \omega_{h_1})} \\
&\quad + O\left(\frac{1}{T^2} \right) + \left[\lambda_{l_1}^{\omega_{j_1}} \langle \phi_{l_3}^{\omega_{j_2}}, \phi_{l_1}^{\omega_{j_1}} \rangle \frac{1}{T} \Delta_T^{(\omega_{j_1} - \omega_{j_2})} + O\left(\frac{1}{T} \right) \right] \\
&\quad \times \left[\lambda_{l_2}^{-\omega_{j_1+h_1}} \langle \phi_{l_4}^{-\omega_{j_2+h_2}}, \phi_{l_2}^{-\omega_{j_1+h_1}} \rangle \frac{1}{T} \Delta_T^{(\omega_{j_1+h_1} - \omega_{j_2+h_2})} + O\left(\frac{1}{T} \right) \right] \\
&\quad + \left[\lambda_{l_1}^{\omega_{j_1}} \langle \phi_{l_4}^{-\omega_{j_2+h_2}}, \phi_{l_1}^{\omega_{j_1}} \rangle \frac{1}{T} \Delta_T^{(\omega_{j_1} + \omega_{j_2+h_2})} + O\left(\frac{1}{T} \right) \right] \\
&\quad \times \left[\lambda_{l_2}^{-\omega_{j_1+h_1}} \langle \phi_{l_3}^{\omega_{j_2}}, \phi_{l_2}^{-\omega_{j_1+h_1}} \rangle \frac{1}{T} \Delta_T^{(-\omega_{j_1+h_1} - \omega_{j_2})} + O\left(\frac{1}{T} \right) \right],
\end{aligned}
$$

where self-adjointness of the spectral density operator gave

$$
\langle \mathcal{F}_{\omega_{j_1}} (\phi_{l_2}^{\omega_{j_2}}), \phi_{l_1}^{\omega_{j_1}} \rangle = \langle \phi_{l_2}^{\omega_{j_2}}, \mathcal{F}_{\omega_{j_1}} (\phi_{l_1}^{\omega_{j_1}}) \rangle = \lambda_{l_1}^{\omega_{j_1}} \langle \phi_{l_2}^{\omega_{j_2}}, \phi_{l_1}^{\omega_{j_1}} \rangle.
$$

Self-adjointness of $\mathcal{F}_\omega$, orthogonality of the eigenfunctions and 2π-periodicity of the eigenelements imply that

$$
\mathrm{Cov}\big(\sqrt{T}\boldsymbol{\beta}_{h,u}^{(T)}, \sqrt{T}\boldsymbol{\beta}_{h,u}^{(T)}\big)
$$

$$
= \frac{2\pi}{T^2} \sum_{j_1,j_2=1}^{T} \sum_{\substack{l_1\in[L(\omega_{j_1})],l_2\in[L(\omega_{j_1+h})],\\ l_3\in[L(\omega_{j_2})],l_4\in[L(\omega_{j_2+h})]}} \big\langle \mathcal{F}_{\omega_{j_1},-\omega_{j_1+h},-\omega_{j_2}}\big(\phi_{l_3}^{\omega_{j_2}}\otimes\phi_{l_4}^{\omega_{j_2+h}}\big),
$$

$$
\phi_{l_1}^{\omega_{j_1}}\otimes\phi_{l_2}^{\omega_{j_1+h}}\big\rangle
$$

$$
+ \frac{2}{T}\sum_{j_1}\sum_{l_1\in[L(\omega_{j_1})],l_2\in[L(\omega_{j_1+h})]} \lambda_{l_1}^{\omega_{j_1}}\lambda_{l_2}^{\omega_{j_1+h}} + O\Big(\frac{1}{T}\Big),
$$

in case $h_1 = h_2 = h$ and $\mathrm{Cov}(\sqrt{T}\boldsymbol{\beta}_{h_1,u}^{(T)}, \sqrt{T}\boldsymbol{\beta}_{h_2,u}^{(T)}) = O(T^{-1})$ if $h_2 \neq h_1 \mod T$. It can then be derived, similarly, that $T\,\mathrm{Cov}(\boldsymbol{\beta}_{h_1,u}^{(T)}, \boldsymbol{\beta}_{h_2,u}^{(T)}) = O(T^{-1})$ for $h_2 \neq T - h_1 \mod T$. Since

$$
\Re\boldsymbol{\beta}_{h_1,u}^{(T)} = \frac{1}{2}\big(\boldsymbol{\beta}_{h_1,u}^{(T)} + \overline{\boldsymbol{\beta}_{h_1,u}^{(T)}}\big) \quad \text{and} \quad \Im\boldsymbol{\beta}_{h_1,u}^{(T)} = \frac{1}{2\mathrm{i}}\big(\boldsymbol{\beta}_{h_1,u}^{(T)} - \overline{\boldsymbol{\beta}_{h_1,u}^{(T)}}\big),
$$

it follows therefore that

$$
T\,\mathrm{Cov}\big(\Re\boldsymbol{\beta}_{h_1,u}^{(T)}, \Im\boldsymbol{\beta}_{h_2,u}^{(T)}\big) = O(T^{-1})
$$

uniformly in h_1, h_2 and, thus, $T\,\mathrm{Cov}(\Re\boldsymbol{\beta}_{h_1,u}^{(T)}, \Re\boldsymbol{\beta}_{h_2,u}^{(T)}) = T\,\mathrm{Cov}(\Im\boldsymbol{\beta}_{h_1,u}^{(T)}, \Im\boldsymbol{\beta}_{h_2,u}^{(T)}) = \frac{T}{2} \times \mathrm{Cov}(\boldsymbol{\beta}_{h_1,u}^{(T)}, \boldsymbol{\beta}_{h_2,u}^{(T)})$. Finally, using Lipschitz continuity of $\omega \mapsto \mathcal{F}_\omega$ and of its eigenelements to replace the Riemann approximations with their limits completes the proof. □

C.3. Proof of Theorem 4.1.

PROOF OF THEOREM 4.1. (i) In order to prove the first assertion of the theorem, introduce the bias-variance decomposition

$$
(29) \quad
\begin{aligned}
&\mathbb{E}\big[\big\|\|\hat{\mathcal{F}}_\omega^{(T)} - \mathbb{E}[\hat{\mathcal{F}}_\omega^{(T)}] + \mathbb{E}[\hat{\mathcal{F}}_\omega^{(T)}] - G_\omega\|_2^2\big] \\
&\quad = \mathbb{E}\big[\big\|\|\hat{\mathcal{F}}_\omega^{(T)} - \mathbb{E}[\hat{G}_\omega]\|_2^2\big] + \mathbb{E}\big[\big\|\|\mathbb{E}[\hat{\mathcal{F}}_\omega^{(T)}] - G_\omega\|_2^2\big].
\end{aligned}
$$

The cross terms cancel because $\mathbb{E}[\langle\hat{\mathcal{F}}_\omega^{(T)} - \mathbb{E}[\hat{\mathcal{F}}_\omega^{(T)}], \mathbb{E}[\hat{\mathcal{F}}_\omega^{(T)}] - G_\omega\rangle_{H\otimes H}]$ and $\mathbb{E}[\hat{\mathcal{F}}_\omega^{(T)} - \mathbb{E}[\hat{\mathcal{F}}_\omega^{(T)}]] = O_H$. Now, by Corollary B.2,

$$
\mathrm{cum}\big(D_\omega^{(T)}, D_{-\omega}^{(T)}\big) = \frac{1}{T}\sum_{t=0}^{T-1}\mathcal{F}_{t/T,\omega} + R_{T,2} = G_\omega^{(T)} + R_{T,2},
$$

where $\|R_{T,2}\|_2 = O(T^{-1})$. Note that the integral approximation in time direction does not change the error term because of Lipschitz continuity of the mapping $(u, \omega) \mapsto \mathcal{F}_{u,\omega}$ in u. Convolution of the cumulant tensor with the smoothing kernel, replacing the integral approximation with the limit and a change of variables give

$$
\mathbb{E}[\hat{\mathcal{F}}_\omega^{(T)}] = \frac{2\pi}{bT}\sum_{j=1}^{T} K_b(\omega - \omega_j)\,\mathrm{cum}\big(D_{\omega_j}^{(T)}, D_{-\omega_j}^{(T)}\big)
$$

$$
= \int K(x)G_{\omega-xb}\,dx + R_{b,T} + R_{T,2},
$$

where $\|R_{b,T}\|_2 = O(bT^{-1})$. Since $\sup_{\omega,u} \||\mathcal{F}_{u,\omega}\||_2 < \infty$ and $\sup_{\omega,u} \||\frac{\partial^2}{\partial\omega_2}\mathcal{F}_{u,\omega}\||_2 < \infty$, the mapping $\omega \mapsto G_\omega$ is twice differentiable and $\sup_\omega \||\frac{\partial^2}{\partial\omega_2}G_\omega\||_2 < \infty$. Therefore, a Taylor expansion around ω and symmetry of the kernel then lead to

$$\mathbb{E}[\hat{\mathcal{F}}_\omega^{(T)}] = \int K(x)G_{\omega-xb}\,dx = G_\omega + \sum_{i=1}^2 \frac{1}{i!}(b)^i \frac{\partial^i G_v}{\partial v^i}\bigg|_{v=\omega} \int x^i K(x)\,dx$$

$$= G_\omega + \epsilon_{b,T},$$

where $\|\epsilon_{b,T}\|_2 = O(b^2 + (bT))^{-1}$. Thus, the second term on the right-hand side of (29) satisfies

$$(30) \qquad \mathbb{E}[\||\mathbb{E}\hat{\mathcal{F}}_\omega^{(T)} - G_\omega\||_2^2] = O\left(b^2 + \frac{1}{bT}\right)^2$$

uniformly in $\omega \in [-\pi, \pi]$. To bound the first term of the right-hand side in (29), observe that, for $j_1 + j_2 \equiv 0 \mod T$, Lemma B.1 with $k = 2$ yields

$$\operatorname{cum}(D_{\omega_{j_1}}^{(T)}, D_{\omega_{j_2}}^{(T)}) = \frac{1}{T}\sum_{t=1}^T \mathcal{F}_{t/T}e^{-i(\omega_{j_1}+\omega_{j_2})t} + R_{2,T}$$

$$\to \int_0^1 \mathcal{F}_{u;\omega_{j_1}}e^{-i2\pi u(j_1+j_2)}\,du = \tilde{\mathcal{F}}_{j_1+j_2;\omega_{j_1}}.$$

Furthermore, from Corollary B.2 and Minkowski's inequality,

$$\||\operatorname{cum}(D_\omega^{(T)}, D_{-\omega}^{(T)}, D_{\omega'}^{(T)}, D_{-\omega'}^{(T)})\||_2$$

$$(31) \qquad \leq \frac{1}{T}\left\||\frac{1}{T}\sum_{t=0}^{T-1}\mathcal{F}_{\frac{t}{T},\omega,-\omega,\omega'}\right\||_2 + O\left(\frac{1}{T^2}\right)$$

$$= \frac{1}{T}\||G_{\omega,-\omega,\omega'}^{(T)}\||_2 + O\left(\frac{1}{T^2}\right) = O\left(\frac{1}{T}\right).$$

The last equality follows since

$$\sup_{u,\omega,\omega'}\||\mathcal{F}_{t/T,\omega,-\omega,\omega'}\||_2 \leq \sum_{h_1,h_2,h_3\in\mathbb{Z}}\|\kappa_{3;h_1,h_2,h_3}\|_2 = O(1)$$

by Assumption I. Hence, Theorem B.1 implies that

$$\operatorname{Cov}(\hat{\mathcal{F}}_\omega, \hat{\mathcal{F}}_\omega)$$

$$= \frac{1}{(bT)^2}\sum_{k_1,k_2=1}^T K\left(\frac{\omega-\omega_{k_1}}{b}\right)K\left(\frac{\omega-\omega_{k_2}}{b}\right)[\operatorname{cum}(D_{\omega_{k_1}}^{(T)}, D_{-\omega_{k_1}}^{(T)}, D_{-\omega_{k_2}}^{(T)}, D_{\omega_{k_2}}^{(T)})$$

$$+ S_{1324}(\operatorname{cum}(D_{\omega_{k_1}}^{(T)}, D_{-\omega_{k_2}}^{(T)}) \otimes \operatorname{cum}(D_{-\omega_{k_1}}^{(T)}, D_{\omega_{k_2}}^{(T)}))$$

$$+ S_{1423}(\operatorname{cum}(D_{\omega_{k_1}}^{(T)}, D_{\omega_{k_2}}^{(T)}) \otimes \operatorname{cum}(D_{-\omega_{k_1}}^{(T)}, D_{-\omega_{k_2}}^{(T)}))].$$

Using Lemma B.1, this equals

$$\frac{1}{(bT)^2}\sum_{k_1,k_2=1}^T K\left(\frac{\omega-\omega_{k_1}}{b}\right)K\left(\frac{\omega-\omega_{k_2}}{b}\right)$$

$$(32)$$

$$\times\big[S_{1324}(\tilde{\mathcal{F}}_{k_1-k_2;\omega_{k_1}} \otimes \tilde{\mathcal{F}}_{-k_1+k_2;-\omega_{k_1}})$$

$$+ S_{1423}(\tilde{\mathcal{F}}_{k_1+k_2;\omega_{k_1}} \otimes \tilde{\mathcal{F}}_{-k_1-k_2;-\omega_{k_1}}) + R_{2,T}\big],$$

where we used (31) is of order $O(\frac{1}{T})$ in S_2 uniformly in $-\pi \leq \omega, \omega' \leq \pi$. Using a change of variables, the properties of the smoothing kernel, Hölder's inequality and Corollary B.1, it follows that

$$
\left\|\!\left\| \frac{1}{(bT)^2} \sum_{k_1,k_2=1}^{T} K\!\left(\frac{\omega - \omega_{k_1}}{b}\right) K\!\left(\frac{\omega - \omega_{k_2}}{b}\right) \tilde{\mathcal{F}}_{k_1-k_2;\omega_{k_1}} \widetilde{\otimes} \tilde{\mathcal{F}}_{-k_1+k_2;-\omega_{k_1}} \right\|\!\right\|_2
$$

$$
\leq \left\|\!\left\| \frac{1}{(bT)^2} \sum_{k_1}^{T} K\!\left(\frac{\omega - \omega_{k_1}}{b}\right)^2 \sum_s \tilde{\mathcal{F}}_{s;\omega_{k_1}} \otimes \tilde{\mathcal{F}}_{-s;-\omega_{k_1}} \right\|\!\right\|_2
$$

$$
\leq \sup_\omega \sum_{s \in \mathbb{Z}} \|\!\|\tilde{\mathcal{F}}_{s;\omega}\|\!\|_2^2 \left| \frac{1}{(bT)^2} \sum_{k_1}^{T} K\!\left(\frac{\omega_{j_1} - \omega_{k_1}}{b}\right)^2 \right| = O\!\left(\frac{1}{bT}\right).
$$

A similar argument holds for the remaining term of (32). Hence,

$$
\|\!\|\mathrm{Cov}(\hat{\mathcal{F}}_\omega^{(T)}, \hat{\mathcal{F}}_\omega^{(T)})\|\!\|_2
$$

$$
= \left\|\!\left\| \left(\frac{2\pi}{T}\right)^2 \sum_{j,j'=1}^{T} K_b(\omega - \omega_j) K_b(\omega - \omega_{j'}) \mathrm{Cov}(I_{\omega_j}^{(T)}, I_{\omega_{j'}}^{(T)}) \right\|\!\right\|_2
$$

$$
= O\!\left(\frac{1}{bT}\right).
$$

Fubini's theorem together with the above implies that the first term of (29) satisfies

$$
\mathbb{E}[\|\!\|\hat{\mathcal{F}}_\omega^{(T)} - \mathbb{E}\hat{\mathcal{F}}_\omega^{(T)}\|\!\|_2^2] = \mathrm{trace}(\mathrm{Var}(\hat{\mathcal{F}}_\omega^{(T)})) = O\!\left(\frac{1}{bT}\right)
$$

uniformly in $\omega \in [-\pi, \pi]$. This establishes (i).

(ii) This part of the proof requires the following lemma verified in Section S1 of the Online Supplement.

LEMMA C.1. *Let Y_v, $v \in [a, b]$ be a zero-mean $L^2([0, 1]^k)$-valued stochastic process of which the derivative mapping $v \mapsto \frac{\partial}{\partial v} Y_v$ is well defined in $L^2([0, 1]^k)$ for any $v \in [a, b]$. If $\mathbb{E}\|Y_v\|_2^2 < \infty$ and $\mathbb{E}\|\frac{\partial Y_v}{\partial v}\|_2^2 < \infty$, then*

$$
2\mathbb{E} \sup_{a \leq v \leq b} \|Y_v\|_2^2 \leq \mathbb{E}\|Y_a\|_2^2 + \mathbb{E}\|Y_b\|_2^2 + \int_a^b \sqrt{\mathbb{E}\left\|\frac{\partial}{\partial \alpha} Y_\alpha\right\|_2^2} \sqrt{\mathbb{E}\|\overline{Y_\alpha}\|_2^2}
$$

$$
+ \int_a^b \sqrt{\mathbb{E}\|Y_\alpha\|_2^2} \, d\alpha \sqrt{\mathbb{E}\left\|\frac{\partial}{\partial \alpha} \overline{Y_\alpha}\right\|_2^2} \, d\alpha.
$$

Lemma C.1 with $k = 2$ applied to the spectral density kernel function $\hat{f}_\omega$ implies—due to the norm equivalence with the operator $\hat{\mathcal{F}}_\omega$—that

$$
\mathbb{E} \sup_{0 \leq \omega \leq \pi} 2\|\!\|\hat{\mathcal{F}}_\omega - \mathbb{E}\hat{\mathcal{F}}_\omega\|\!\|_2^2
$$

$$
\leq \mathbb{E}\|\!\|\hat{\mathcal{F}}_0 - \mathbb{E}\hat{\mathcal{F}}_0\|\!\|_2^2 + \mathbb{E}\|\!\|\hat{\mathcal{F}}_\pi - \mathbb{E}\hat{\mathcal{F}}_\pi\|\!\|_2^2
$$

$$
(33) \qquad + 2 \int_0^\pi \sqrt{\mathbb{E}\|\!\|\hat{\mathcal{F}}_\omega - \mathbb{E}\hat{\mathcal{F}}_\omega\|\!\|_2^2} \sqrt{\mathbb{E}\left\|\!\left\|\frac{\partial}{\partial \omega}(\hat{\mathcal{F}}_\omega - \mathbb{E}\hat{\mathcal{F}}_\omega)\right\|\!\right\|_2^2} \, d\omega
$$

$$= \text{Tr Var}(\hat{\mathcal{F}}_0) + \text{Tr Var}(\hat{\mathcal{F}}_\pi) + 2 \int_0^\pi \sqrt{\text{Tr Var}(\hat{\mathcal{F}}_\omega)} \sqrt{\text{Tr Var}\left(\frac{\partial}{\partial \omega} \hat{\mathcal{F}}_\omega\right)} \, d\omega$$

$$= O\left(\frac{1}{bT}\right) + O\left(\frac{1}{\sqrt{bT}\sqrt{b^2 T}}\right) = O\left(\frac{1}{b^2 T}\right),$$

where the latter follows from part (i). The rate for the covariance structure of the operator-valued function $\omega \mapsto \frac{\partial}{\partial \omega} \hat{\mathcal{F}}_\omega$ follow as before, noting that an application of the chain rule of the derivative will lead to an extra $O(\frac{1}{b^2})$ term in $S_2(H)$ in comparison to the covariance of $\hat{\mathcal{F}}_\omega$. Minkowski's inequality therefore implies

$$\mathbb{P}\left(\sup_{\omega \in [-\pi,\pi]} \|\hat{\mathcal{F}}_\omega^{(T)} - G_\omega\|_2 > \epsilon\right) \leq \mathbb{P}\left(\sup_{\omega \in [-\pi,\pi]} \|\hat{\mathcal{F}}_\omega^{(T)} - \mathbb{E}\hat{\mathcal{F}}_\omega\|_2 > \frac{\epsilon}{2}\right)$$

$$+ \mathbb{P}\left(\sup_{\omega \in [-\pi,\pi]} \|\mathbb{E}\hat{\mathcal{F}}_\omega^{(T)} - G_\omega\|_2 > \frac{\epsilon}{2}\right).$$

Using Markov's inequality together with (33), for any $\epsilon > 0$,

$$\mathbb{P}\left(\sup_{\omega \in [-\pi,\pi]} \|\hat{\mathcal{F}}_\omega^{(T)} - \mathbb{E}\hat{\mathcal{F}}_\omega\|_2 > \frac{\epsilon}{2}\right) \leq O\left(\frac{1}{\epsilon^2 b^2 T}\right) \to 0$$

as $b^2 T \to \infty$. Similary, Markov's inequality together with (30) yields

$$\mathbb{P}\left(\sup_{\omega \in [-\pi,\pi]} \|\mathbb{E}\hat{\mathcal{F}}_\omega^{(T)} - G_\omega\|_2 > \frac{\epsilon}{2}\right) \leq O\left(\frac{1}{\epsilon^2}\left(b^2 + \frac{1}{bT}\right)^2\right) \to 0$$

as $bT \to \infty$ and $b \to 0$ as $T \to \infty$. The result therefore holds provided Assumption 3.1 is satisfied. □

APPENDIX D: WEAK CONVERGENCE

The proof of the distributional properties of $\hat{\boldsymbol{\beta}}_{h,x}^{(T)}$, as stated in Theorem 4.4 and 4.6, are established in this section. The proof consists of several steps. First, the distributional properties are derived for $\boldsymbol{\beta}_{h,x}^{(T)}$, when spectral density operators and their corresponding eigenelements are known. For this, we investigate the distributional properties of the operator

$$(34) \qquad w_h^{(T)} = \frac{1}{T} \sum_{j=1}^T D_{\omega_j}^{(T)} \otimes D_{\omega_{j+h}}^{(T)}, \quad h = 1, \dots, T-1.$$

Theorem D.1 below shows that $\sqrt{T}(w_h^{(T)} - \mathbb{E}w_h^{(T)})$ converges weakly to a functional Gaussian process both under the null and the alternative. The distributional properties of $\boldsymbol{\beta}_{h,x}^{(T)}$ immediately follow from this result and thus converge weakly to a Gaussian process under both hypotheses. Focus is finally on $\hat{\boldsymbol{\beta}}_{h,x}^{(T)}$, where the effect of replacing the eigenelements with their empirical counterparts on the distributional properties is clarified. In particular, Theorems 4.2 and 4.5 are established as well as the orders of

$$\mathbb{E}\sqrt{T}|\hat{\boldsymbol{\beta}}_{h,u}^{(T)} - \boldsymbol{\beta}_{h,u}^{(T)}| \quad \text{and} \quad \mathbb{E}\sqrt{T}|\hat{\boldsymbol{\beta}}_{h,s}^{(T)} - \boldsymbol{\beta}_{h,s}^{(T)}|.$$

D.1. Weak convergence on the function space. To demonstrate weak convergence of (34), the following result by Cremers and Kadelka (1986) is used which considerably simplifies the verification of the usual tightness condition often invoked in weak convergence proofs of Banach space-valued random variables.

LEMMA D.1. *Let $(\mathscr{T}, \mathcal{A}, \mu)$ be a measure space, let $(B, |\cdot|)$ be a Banach space, and let $X = (X_n : n \in \mathbb{N})$ be a sequence of random elements in $L^p_B(\mathscr{T}, \mu)$ such that:*

(i) the finite-dimensional distributions of X converge weakly to those of a random element X_0 in $L^p_B(\mathscr{T}, \mu)$;

(ii) $\limsup_{n \to \infty} \mathbb{E}[\|X_n\|^p_p] \leq \mathbb{E}[\|X_0\|^p_p] < \infty$.

Then, X converges weakly to X_0 in $L^p_B(\mathscr{T}, \mu)$.

To apply Lemma D.1 in the present context, consider the sequence $(\hat{E}^{(T)}_h : T \in \mathbb{N})$ of random elements in $L^2_{\mathbb{C}}([0,1]^2)$, for $h = 1, \ldots, T-1$ defined through

$$\hat{E}^{(T)}_h = \sqrt{T}(w^{(T)}_h - \mathbb{E}[w^{(T)}_h]) = \sum_{l,l'=1}^{\infty} \langle \hat{E}^{(T)}_h, \psi_{ll'} \rangle \psi_{ll'},$$

where the second equality uses a representation with respect to an $L^2_{\mathbb{C}}([0,1]^2)$ orthonormal basis $\psi_{ll'} = \psi_l \otimes \psi'_l$. From this representation it is easily seen that the finite-dimensional distributions of the basis coefficients provide a complete characterization of the distributional properties of $\hat{E}^{(T)}_h$. To formalize this, we put the functional $\hat{E}^{(T)}_h$ in duality with $(\hat{E}^{(T)}_h)^* \in L^2_{\mathbb{C}}([0,1]^2)^*$ through the pairing $\hat{E}^{(T)}_h(\phi) = \langle \hat{E}^{(T)}_h, \phi \rangle$ for all $\phi \in L^2_{\mathbb{C}}([0,1]^2)^*$. This leads to the following result, which is stated under the more general Assumption I, which encompasses the stationary case.

THEOREM D.1 (Weak convergence). *Let $(X_t : t \in \mathbb{Z})$ be a stochastic process taking values in $H_{\mathbb{R}}$ satisfying Assumption I with $\ell = 2$. Then,*

$$(35) \qquad (\Re \hat{E}^{(T)}_{h_i}, \Im \hat{E}^{(T)}_{h_i} : i = 1, \ldots, k) \xrightarrow{d} (\mathcal{R}_{h_i}, \mathcal{I}_{h_i} : i = 1, \ldots, k),$$

where $\mathcal{R}_h, \mathcal{I}_{h'}, h, h' \in \{1, \ldots, T-1\}$ are jointly Gaussian elements in $L^2_{\mathbb{C}}([0,1]^2)$ with means $\mathbb{E}[\mathcal{R}_h(\psi_{ll'})] = \mathbb{E}[\mathcal{I}_{h'}(\psi_{ll'})] = 0$ and covariance structure:

1. $\mathrm{Cov}(\mathcal{R}_h(\psi_{l_1 l'_1}), \mathcal{R}_{h'}(\psi_{l_2 l'_2}))$

$$= \frac{1}{4}[\Upsilon_{h,h'}(\psi_{l_1 l'_1 l_2 l'_2}) + \acute{\Upsilon}_{h,h'}(\psi_{l_1 l'_1 l_2 l'_2})$$
$$+ \grave{\Upsilon}_{h,h'}(\psi_{l_1 l'_1 l_2 l'_2}) + \bar{\Upsilon}_{h,h'}(\psi_{l_1 l'_1 l_2 l'_2})],$$

2. $\mathrm{Cov}(\mathcal{I}_h(\psi_{l_1 l'_1}), \mathcal{R}_{h'}(\psi_{l_2 l'_2}))$

$$= \frac{1}{4\mathrm{i}}[\Upsilon_{h,h'}(\psi_{l_1 l'_1 l_2 l'_2}) + \acute{\Upsilon}_{h,h'}(\psi_{l_1 l'_1 l_2 l'_2})$$
$$- \grave{\Upsilon}_{h,h'}(\psi_{l_1 l'_1 l_2 l'_2}) - \bar{\Upsilon}_{h,h'}(\psi_{l_1 l'_1 l_2 l'_2})],$$

3. $\mathrm{Cov}(\mathcal{R}_h(\psi_{l_1 l'_1}), \mathcal{I}_{h'}(\psi_{l_2 l'_2}))$

$$= \frac{1}{4\mathrm{i}}[\Upsilon_{h,h'}(\psi_{l_1 l'_1 l_2 l'_2}) - \acute{\Upsilon}_{h,h'}(\psi_{l_1 l'_1 l_2 l'_2})$$
$$+ \grave{\Upsilon}_{h,h'}(\psi_{l_1 l'_1 l_2 l'_2}) - \bar{\Upsilon}_{h,h'}(\psi_{l_1 l'_1 l_2 l'_2})],$$

4. $\mathrm{Cov}\big(\mathcal{I}_h(\psi_{l_1 l_1'}), \mathcal{I}_{h'}(\psi_{l_2 l_2'})\big)$

$$= \frac{1}{4}\big[\Upsilon_{h,h'}(\psi_{l_1 l_1' l_2 l_2'}) - \acute{\Upsilon}_{h,h'}(\psi_{l_1 l_1' l_2 l_2'})$$

$$- \grave{\Upsilon}_{h,h'}(\psi_{l_1 l_1' l_2 l_2'}) + \bar{\Upsilon}_{h,h'}(\psi_{l_1 l_1' l_2 l_2'})\big]$$

for all h, h' *and* l_1, l_1', l_2, l_2', *and where* $\Upsilon_{h,h'}$, $\acute{\Upsilon}_{h,h'}$, $\grave{\Upsilon}_{h,h'}$ *and* $\bar{\Upsilon}_{h,h'}$ *are given in* (S6.3)–(S6.5).

PROOF. It remains to verifiy the conditions of Lemma D.1. For the first, the following theorem establishes that the finite-dimensional distributions converge weakly to a Gaussian process both under the null and the alternative.

THEOREM D.2. *Under the conditions of Theorem* D.1, *we have for all* $l_i, l_i' \in \mathbb{N}$, $h_i = 1, \ldots, T-1$, $i = 1, \ldots, k$ *and* $k \geq 3$,

$$\mathrm{cum}\big(\hat{E}_{h_1}^{(T)}(\psi_{l_1 l_1'}), \ldots, \hat{E}_{h_k}^{(T)}(\psi_{l_k l_k'})\big) = o(1) \quad (T \to \infty).$$

The proof of Theorem D.2 can be found in Section S3 of the Online Supplement. Note that, for the second condition of Lemma D.1, Parseval's identity and the monotone convergence theorem yield

(36)
$$\mathbb{E}\big[\|\hat{E}_h^{(T)}\|_2^2\big] = \sum_{l,l'=1}^{\infty} \mathbb{E}\big[|\hat{E}_h^{(T)}(\psi_{ll'})|^2\big]$$

$$\to \sum_{l,l'=1}^{\infty} \mathbb{E}\big[|E_h(\psi_{ll'})|^2\big] = \mathbb{E}\big[\|E_h\|_2^2\big] \quad (T \to \infty),$$

with E_h denoting the limiting process. Observe that, from (27) and the Cauchy–Schwarz inequality, the terms $\Upsilon_{h,h'}$, $\acute{\Upsilon}_{h,h'}$, $\grave{\Upsilon}_{h,h'}$ and $\bar{\Upsilon}_{h,h'}$ are finite. Condition (ii) of Lemma D.1 is then satisfied, since

$$\mathbb{E}\big[\|\hat{E}_h^{(T)}\|_2^2\big] = \int_{[0,1]^2} \mathrm{Var}(\hat{E}_h^{(T)}(\tau, \tau')) \, d\tau \, d\tau' = T\mathbb{E}\|w_h^{(T)}\|_2^2$$

$$\to \mathrm{Tr}\big(\mathrm{Var}(\mathcal{R}_h)\big) + \mathrm{Tr}\big(\mathrm{Var}(\mathcal{I}_h)\big) < \infty,$$

where Tonelli's theorem was applied to obtain the first equality. This completes the proof. □

D.2. Replacing eigenelements with estimates.

D.2.1. *Invariance under rotation.* We now focus on replacing the projection basis with estimates of the eigenfunctions of the spectral density operators. It can be shown (Mas and Menneteau (2003)) that for rates of the bandwidth b for which the estimated spectral density operator is a consistent estimator of the true spectral density operator, the corresponding estimated eigenprojectors $\hat{\Pi}_l^{\omega} = \hat{\phi}_l^{\omega} \otimes \hat{\phi}_l^{\omega}$ are consistent for the eigenprojectors Π_l^{ω}. However, the estimated eigenfunctions are not unique and only identified up to rotation on the unit circle. In order to show that replacing the eigenfunctions with estimates does not affect the limiting distribution, the issue of rotation has to be considered first. More specifically, when estimating, a version $\hat{z}_l \hat{\phi}_l^{\omega_j}$, where $\hat{z}_l \in \mathbb{C}$ with modulus $|\hat{z}_l| = 1$, is obtained which cannot

be guaranteed to be close to the true eigenfunction $\phi_l^{\omega_j}$. It is therefore essential that the test statistic is invariant under rotations. To show this, write

$$\Psi_h(j, l, l') = \langle D_{\omega_j}^{(T)}, \hat{\phi}_l^{\omega_j} \rangle \overline{\langle D_{\omega_{j+h}}^{(T)}, \hat{\phi}_{l'}^{\omega_{j+h}} \rangle}$$

and let $\boldsymbol{\Psi}(h) = \text{vec}(\Psi_h(j, l, l',))$ be the stacked vector of dimension $\prod_{j=1}^{T-h} L(\omega_j) L(\omega_{j+h})$. Note that then $\hat{\boldsymbol{\beta}}_{h,u}^{(T)} = e^\top \boldsymbol{\Psi}(h)$. Construct the diagonal matrix

$$Z_{L(\omega_j)}^j = \begin{pmatrix} \hat{z}_1^j & \cdots & & & \\ \vdots & \hat{z}_2^j & & & \\ & & \ddots & & \\ & & & & \hat{z}_{L(\omega_j)}^j \end{pmatrix},$$

the block diagonal matrix $Z_{L(\omega_j)}^{1:T} = \text{diag}(Z_{L(\omega_j)}^j : j = 1, \ldots, T)$ and the Kronecker product $\boldsymbol{Z}(h) = Z_{L(\omega_j)}^{1:T-h} \otimes Z_{L(\omega_j)}^{h:T}$. This is a diagonal object of dimension $(\prod_{j=1}^{T-h} L(\omega_j) L(\omega_{j+h}))^2$, whose diagonal elements are given by $\{\hat{z}_l^j \overline{\hat{z}}_{l'}^{j+h}\}$. Rotating the eigenfunctions on the unit circle yields versions

$$\hat{\boldsymbol{\beta}}_{h,u}^{(T)} = e^\top \boldsymbol{Z}(h) \boldsymbol{\Psi}(h).$$

For these versions, write $\sqrt{T} \mathcal{Z}_M \hat{\boldsymbol{b}}_{M,u}^{(T)} = \sqrt{T} (\Re\hat{\boldsymbol{\beta}}_{h_1,u}^{(T)}, \ldots, \Im\hat{\boldsymbol{\beta}}_{h_M,u}^{(T)}, \Im\hat{\boldsymbol{\beta}}_{h_1,u}^{(T)}, \ldots, \Im\hat{\boldsymbol{\beta}}_{h_M,u}^{(T)})^\top$, where the block diagonal matrix is given by $\mathcal{Z}_M = \text{diag}(\Re\boldsymbol{Z}(h_1), \ldots, \Re\boldsymbol{Z}(h_M), \Im\boldsymbol{Z}(h_1), \ldots, \Im\boldsymbol{Z}(h_M))^\top$. The same rotation, however, also implies that $\hat{\Sigma}_{M,u}$ becomes $\mathcal{Z}_M \hat{\Sigma}_{M,u} \mathcal{Z}_M^\top$ and hence

$$T(\hat{\boldsymbol{b}}_{M,u}^{(T)})^\top (\mathcal{Z}_M)^\top [\mathcal{Z}_M \hat{\Sigma}_{M,u} \mathcal{Z}_M^\top]^{-1} \mathcal{Z}_M \hat{\boldsymbol{b}}_{M,u}^{(T)} = \hat{Q}_{M,u}^{(T)},$$

thereby showing that the value of the test statistic is not affected by rotation of the estimated eigenfunctions. In the rest of the proof, focus is therefore only on estimates $\hat{\phi}_l^{\omega_{j+h}}$ and $\hat{\phi}_{l'}^{\omega_{j+h}}$, and their respective unknown rotations $\hat{z}_l^j$ and $\overline{\hat{z}}_{l'}^{j+h}$ are ignored.

D.2.2. *Limiting distributions of $\hat{\boldsymbol{\beta}}_{h,u}^{(T)}$ and $\hat{\boldsymbol{\beta}}_{h,s}^{(T)}$.* We now investigate the rate of convergence of the statistic, when the eigenfunctions as a basis are replaced with their empirical counterparts, and prove Theorems 4.2 and 4.5. For this, it is sufficient to derive the order of the difference

$$\tag{37} \sqrt{T} \mathbb{E} |\hat{\boldsymbol{\beta}}_{h,x}^{(T)} - \boldsymbol{\beta}_{h,x}^{(T)}|.$$

In the following we shall focus on $\hat{\boldsymbol{\beta}}_{h,u}^{(T)}$ and postpone the derivation for $\hat{\boldsymbol{\beta}}_{h,s}^{(T)}$ to Section S5.2. In order to bound (37), relate $\hat{\phi}_l^{\omega_j} \otimes \hat{\phi}_{l'}^{\omega_{j+h}} - \phi_l^{\omega_j} \otimes \phi_{l'}^{\omega_{j+h}}$ with $\hat{\mathcal{F}}_{\omega_{j_1}} \tilde{\otimes} \hat{\mathcal{F}}_{\omega_{j_1+h}} - \mathcal{F}_{\omega_{j_1}} \tilde{\otimes} \mathcal{F}_{\omega_{j_1+h}}$ from noting that

$$\begin{aligned} &(\mathcal{F}_{\omega_j} \tilde{\otimes} \mathcal{F}_{\omega_{j+h}})(\phi_l^{\omega_j} \otimes \phi_{l'}^{\omega_{j+h}}) \\ &= \sum_{m,m'} \lambda_m^{\omega_j} \lambda_{m'}^{\omega_{j+h}} \phi_m^{\omega_j} \otimes \phi_m^{\omega_j} (\phi_l^{\omega_j} \otimes \phi_{l'}^{\omega_{j+h}}) \phi_m^{\omega_{j+h}} \otimes \phi_{m'}^{\omega_{j+h}} \\ &= \lambda_l^{\omega_j} \lambda_{l'}^{\omega_{j+h}} (\phi_l^{\omega_j} \otimes \phi_{l'}^{\omega_{j+h}}), \end{aligned}$$

where we used Definition S1.1(i). Similarly, $(\hat{\mathcal{F}}_{\omega_j} \tilde{\otimes} \hat{\mathcal{F}}_{\omega_{j+h}})(\hat{\phi}_l^{\omega_j} \otimes \hat{\phi}_{l'}^{\omega_{j+h}}) = \hat{\lambda}_l^{\omega_j} \hat{\lambda}_{l'}^{\omega_{j+h}}(\hat{\phi}_l^{\omega_j} \otimes \hat{\phi}_{l'}^{\omega_{j+h}})$. A first-order Taylor expansion of the eigenvalue–eigenvector equation yields (e.g., Hall and Hosseini-Nasab (2006))

$$
(38) \quad
\begin{aligned}
&\hat{\phi}_l^{\omega_j} \otimes \hat{\phi}_{l'}^{\omega_{j+h}} - (\phi_l^{\omega_j} \otimes \phi_{l'}^{\omega_{j+h}}) \\
&= \sum_{\substack{m \neq l \\ m' \neq l'}} \frac{1}{\lambda_l^{\omega_j} \lambda_{l'}^{\omega_{j+h}} - \lambda_m^{\omega_j} \lambda_{m'}^{\omega_{j+h}}} \langle (\hat{\mathcal{F}}_{\omega_{j_1}} \tilde{\otimes} \hat{\mathcal{F}}_{\omega_{j+h}} - \mathcal{F}_{\omega_j} \tilde{\otimes} \mathcal{F}_{\omega_{j+h}}) \\
&\quad \times (\phi_l^{\omega_j} \otimes \phi_{l'}^{\omega_{j+h}}), \phi_m^{\omega_j} \otimes \phi_{m'}^{\omega_{j+h}} \rangle \phi_m^{\omega_j} \otimes \phi_{m'}^{\omega_{j+h}} + R,
\end{aligned}
$$

where the remainder R is of order $\|R\|_2 = O_p(\|\hat{\mathcal{F}}_{\omega_{j_1}} \tilde{\otimes} \hat{\mathcal{F}}_{\omega_{j_1+h}} - \mathcal{F}_{\omega_{j_1}} \tilde{\otimes} \mathcal{F}_{\omega_{j_1+h}}\|_2^2)$ and will be of smaller order than the first term on the right-hand side of (38). In the proof we require thus that

$$
(39) \quad \lambda_l^{\omega_j} \lambda_{l'}^{\omega_{j+h}} - \lambda_m^{\omega_j} \lambda_{m'}^{\omega_{j+h}} = \lambda_l^{\omega_j}(\lambda_{l'}^{\omega_{j+h}} - \lambda_{m'}^{\omega_{j+h}}) + (\lambda_l^{\omega_j} - \lambda_m^{\omega_j})\lambda_{m'}^{\omega_{j+h}} > 0,
$$

which implies no multiplicity of eigenvalues. It is also required that the spectral density operators are strictly positive definite, a condition needed to ensure that their eigenfunctions form a complete orthonormal basis of H. Note, however, that the assumption of no multiplicity is without loss of generality as one can group multiple eigenelement pairs into blocks and apply the same techniques over these blocks (e.g., Wilkinson (1965)). Given (39) holds true, linearity and continuity of the inner product imply that the error can be rewritten as

$$
\begin{aligned}
&\frac{1}{\sqrt{T}} \sum_{j=1}^{T} \sum_{l,l'} \langle D_{\omega_j} \otimes D_{\omega_{j+h}}, \phi_l^{\omega_j} \otimes \phi_{l'}^{\omega_{j+h}} - (\hat{\phi}_l^{\omega_j} \otimes \hat{\phi}_{l'}^{\omega_{j+h}}) \rangle_S, \\
&= O_p\left(\frac{1}{\sqrt{T}} \sum_{j=1}^{T} \sum_{l,l'} \sum_{\substack{m \neq l \\ m' \neq l'}} \langle D_{\omega_j} \otimes D_{\omega_{j+h}}, \phi_m^{\omega_j} \otimes \phi_{m'}^{\omega_{j+h}} \rangle_S \right. \\
(40) \quad &\qquad \left. \times \langle (\hat{\mathcal{F}}_{\omega_j} \tilde{\otimes} \hat{\mathcal{F}}_{\omega_{j+h}} - \mathcal{F}_{\omega_j} \tilde{\otimes} \mathcal{F}_{\omega_{j+h}})(\phi_l^{\omega_j} \otimes \phi_{l'}^{\omega_{j+h}}), \phi_m^{\omega_j} \otimes \phi_{m'}^{\omega_{j+h}} \rangle_S \right) \\
&= O_p\left(\frac{1}{\sqrt{T}} \sum_{j=1}^{T} \sum_{l,l'} \langle D_{\omega_j} \otimes D_{\omega_{j+h}}, \right. \\
&\qquad \left. (\hat{\mathcal{F}}_{-\omega_j} \tilde{\otimes} \hat{\mathcal{F}}_{-\omega_{j+h}} - \mathcal{F}_{-\omega_j} \tilde{\otimes} \mathcal{F}_{-\omega_{j+h}})(\phi_l^{-\omega_j} \otimes \phi_{l'}^{-\omega_{j+h}}) \rangle_S \right),
\end{aligned}
$$

using that $\langle A, B \rangle_S = \sum_{m \in \mathbb{N}} \langle A, \psi_{mm'} \rangle_S \langle \psi_{mm'}, B \rangle_S$ for any orthonormal basis $\{\psi_{mm'}\}_{m,m' \in \mathbb{N}}$ of S_2. In other words, the order of the difference is completely determined by the order of the difference when replacing the Kronecker products of the estimated spectral density operators with their empirical counterparts. This finding can be utilized to determine the order of (37) by decomposing it as follows and considering each of the terms separately:

$$
(41) \quad
\begin{aligned}
J_1 &= \frac{1}{\sqrt{T}} \sum_{j=1}^{T} \sum_{l,l'} \langle D_{\omega_j} \otimes D_{\omega_{j+h}} - \mathbb{E}(D_{\omega_j} \otimes D_{\omega_{j+h}}), \\
&\qquad \phi_l^{\omega_j} \otimes \phi_{l'}^{\omega_{j+h}} - \mathbb{E}(\hat{\phi}_l^{\omega_j} \otimes \hat{\phi}_{l'}^{\omega_{j+h}}) \rangle_S, \\
J_2 &= \frac{1}{\sqrt{T}} \sum_{j=1}^{T} \sum_{l,l'} \langle D_{\omega_j} \otimes D_{\omega_{j+h}} - \mathbb{E}(D_{\omega_j} \otimes D_{\omega_{j+h}}),
\end{aligned}
$$

(42)
$$\mathbb{E}(\hat{\phi}_l^{\omega_j} \otimes \hat{\phi}_{l'}^{\omega_{j+h}}) - \hat{\phi}_l^{\omega_j} \otimes \hat{\phi}_{l'}^{\omega_{j+h}}\rangle_S,$$

(43)
$$J_3 = \frac{1}{\sqrt{T}} \sum_{j=1}^T \sum_{l,l'} \langle \mathbb{E}(D_{\omega_j} \otimes D_{\omega_{j+h}}), \phi_l^{\omega_j} \otimes \phi_{l'}^{\omega_{j+h}} - \mathbb{E}(\hat{\phi}_l^{\omega_j} \otimes \hat{\phi}_{l'}^{\omega_{j+h}})\rangle_S,$$

(44)
$$J_4 = \frac{1}{\sqrt{T}} \sum_{j=1}^T \sum_{l,l'} \mathbb{E}(\langle D_{\omega_j} \otimes D_{\omega_{j+h}}), \mathbb{E}(\hat{\phi}_l^{\omega_j} \otimes \hat{\phi}_{l'}^{\omega_{j+h}}) - \hat{\phi}_l^{\omega_j} \otimes \hat{\phi}_{l'}^{\omega_{j+h}}\rangle_S.$$

The following lemma contains the order of these four terms:

LEMMA D.2. *Under Assumption* I(12, 2),

(45)
$$\mathbb{E}|J_1| = O\left(\frac{1}{bT} + b^2\right),$$

(46)
$$\mathbb{E}|J_2| = \begin{cases} O\left(\dfrac{1}{bT\sqrt{T}}\right) + O\left(\dfrac{1}{bT}\right) & under\ H_0, \\[2ex] O\left(\dfrac{1}{b\sqrt{T}}\right) + O\left(\dfrac{1}{bT}\right) & under\ H_A, \end{cases}$$

(47)
$$\mathbb{E}|J_3| = \begin{cases} O\left(\dfrac{1}{bT\sqrt{T}} + \dfrac{b^2}{\sqrt{T}}\right) & under\ H_0, \\[2ex] O\left(\dfrac{1}{b\sqrt{T}} + b^2\sqrt{T}\right) & under\ H_A, \end{cases}$$

(48)
$$\mathbb{E}|J_4| = \begin{cases} O\left(\dfrac{1}{\sqrt{b}T}\right) & under\ H_0, \\[2ex] O(1) & under\ H_A. \end{cases}$$

The proof is relegated to Section S5 of the Online Supplement.

APPENDIX E: LIMITING DISTRIBUTION UNDER H_A

THEOREM E.1. *Under the conditions of Theorem* 4.6, *we have, for all* $h_i, h_j \in \mathbb{Z}$ *with* $i, j = 1, \ldots, k,$

(49)
$$T^{k/2} \operatorname{cum}_{n,r}(\boldsymbol{\beta}_{h_i}^{(T)}, \mathcal{B}_{h_j}^{(T)}) = o(1) \quad (T \to \infty),$$

where $\operatorname{cum}_{n,r}(\boldsymbol{\beta}_h^{(T)}, \mathcal{B}_{h'}^{(T)})$ *denotes the joint cumulant*

$$\operatorname{cum}(\underbrace{\boldsymbol{\beta}_h^{(T)}, \ldots, \boldsymbol{\beta}_h^{(T)}}_{n\ times}, \underbrace{\mathcal{B}_{h'}^{(T)}, \ldots, \mathcal{B}_{h'}^{(T)}}_{r\ times})$$

with $0 \le n, r \le k$ *such that* $n + r = k.$

PROOF. We will show that $\sqrt{T}\boldsymbol{\beta}_h^{(T)}$ and $\sqrt{T}\mathcal{B}_h^{(T)}$ are jointly normal. Using (40) and hence that the order of $\mathcal{B}_h^{(T)}$ is determined by the order of

$$\mathcal{V}_h^{(T)} = \frac{1}{T} \sum_{j=1}^T \langle \mathbb{E}(D_{\omega_j} \otimes D_{\omega_{j+h}}), (\hat{\mathcal{F}}_{\omega_{j_1}} \tilde{\otimes} \hat{\mathcal{F}}_{\omega_{j+h}} - \mathbb{E}\hat{\mathcal{F}}_{\omega_j} \tilde{\otimes} \hat{\mathcal{F}}_{\omega_{j+h}})(\phi_l^{\omega_j} \otimes \phi_{l'}^{\omega_{j+h}})\rangle_S,$$

we will show that, for $k > 2$,

$$T^{k/2} \operatorname{cum}_{n,r}(\boldsymbol{\beta}_h^{(T)}, \mathcal{V}_h^{(T)}) = \operatorname{cum}(\underbrace{\boldsymbol{\beta}_h^{(T)}, \dots, \boldsymbol{\beta}_h^{(T)}}_{n \text{ times}}, \underbrace{\mathcal{V}_h^{(T)}, \dots, \mathcal{V}_h^{(T)}}_{r \text{ times}}) = o(1),$$

where $0 \le n$, $r \le n$ such that $n + r = k$. First, note that the operator $\mathbb{E}[D_{\omega_j} \otimes D_{\omega_{j+h}}]$ is compact and therefore separable. Without loss of generality, in order to ease notation, write therefore $D_{\omega_{jk}}^{(l)} = \langle D_{\omega_{jk}}^{(T)}, \psi_l \rangle$ and $\hat{\mathcal{F}}_{\omega_j}^{(lm)} = \langle \hat{\mathcal{F}}_{\omega_j}(\psi_m), \psi_l \rangle$, where $\{\psi_l\}_{l \in \mathbb{N}}$ forms a basis of H. Using then Theorem B.1

$$T^{k/2} \operatorname{cum}_{n,r}(\boldsymbol{\beta}_h^{(T)}, \mathcal{V}_h^{(T)})$$

$$= T^{k/2} \sum_{j_1, \dots, j_k} \operatorname{cum}(D_{\omega_{j_1}}^{(l_1)} D_{-\omega_{j_1+h_1}}^{(l_1')}, \dots, D_{\omega_{j_n}}^{(l_n)} D_{-\omega_{j_n+h_n}}^{(l_n')},$$

$$\hat{\mathcal{F}}_{\omega_{j_{n+1}}}^{(l_{n+1} m_{n+1})} \overline{\hat{\mathcal{F}}_{\omega_{j_{n+1}+h_{n+1}}}^{(l_{n+1}' m_{n+1}')}}, \dots, \hat{\mathcal{F}}_{\omega_{j_k}}^{(l_k m_k)} \overline{\hat{\mathcal{F}}_{\omega_{j_k+h_k}}^{(l_k' m_k')}})$$

$$= T^{-k/2} \sum_{j_1, \dots, j_k} \left(\frac{2\pi}{bT}\right)^{2r} \prod_{d=n+1}^{k} \sum_{q_d=1}^{T} K\left(\frac{\omega_{j_d} - \omega_{q_{2(d-n)-1}}}{b}\right)$$

$$\times K\left(\frac{\omega_{j_d+h_d} - \omega_{q_{2(d-n)}}}{b}\right)$$

$$\times \sum_{i.p.} \operatorname{cum}(D_{\lambda_{j_s}}^{(v_s)} : s \in P_1) \cdots \operatorname{cum}(D_{\omega_{j_s}}^{(v_s)} : s \in P_Q),$$

where the summation extends over all indecomposable partitions $P = \{P_1, \dots, P_Q\}$ of the array

(50)

$$\begin{array}{llll}
(1, 1) & (1, 2) & & \\
\vdots & \vdots & & \\
(n, 1) & (n, 2) & & \\
(n+1, 1) & (n+1, 2) & (n+1, 3) & (n+1, 4) \\
\vdots & \vdots & \vdots & \\
(k, 1) & (k, 2) & (k, 3) & (k, 4),
\end{array}$$

using similar notation as in the proof of Theorem D.1. In particular, the value $s = ii'$ corresponds to entry (i, i') of (50). For a partition $P = \{P_1, \dots, P_Q\}$, the elements of a set P_v will be denoted by $s_{v1}, \dots, s_{v|P_v|}$, with $|P_v|$ being the number of elements in P_v. In this case we associate with entry s the frequency index $j_s = j_{ii'} = (-1)^{i'-1}(j_i + h_i^{i'-1})$ for $i \le n$; for $i > n$ we associate the frequency index $j_s = q_{ii'} = (-1)^{i'-1} q_{2(i-n)-1+\lfloor i'/3 \rfloor}$ such that $\lambda_{j_s} = \frac{2\pi j_s}{T}$ and the basis function index $v_s = v_{ii'} = l_i^{2-i'} l_i'^{i'-1}$ for $i = 1, \dots, k$ and $i' = \{1, 2\}$, while for $i' = \{3, 4\}$ we set $v_s = v_{ii'} = m_i^{4-i'} m_i'^{i'-3}$.

For the array to be indecomposable, the rows must hook. Since interest is only in a bound for the partition of highest order, only partitions have to be considered for which each set satisfies $|P_v| = 2$, since all other partitions will be of lower order. Without loss of generality, consider that row i hooks with $i+1$ for $i = 2, \dots, k-1$, and let the first and the last row hook. In particular, a partition of highest order would be one for which $P_i = \{(i, 2) \cup (i+1, 1)\}$ for $i = 2, \dots, k$ and $P_1 = \{(1, 1) \cup (k, 2)\}$ and where the $2r$ variables in the third and fourth columns of the last r rows are decomposable, meaning that $P_{n+i} = \{(n+i, 3) \cup (n+i, 4)\}$

for $i = 1, \ldots, r$, s that these latter r sets form proper submanifolds of the frequency manifold. Using Lemma B.1, such a partition can be written as

$$T^{-k/2} \sum_{j_1, \ldots, j_k} \left(\frac{2\pi}{bT}\right)^{2r} \prod_{d=n+1}^{k} \sum_{q_d=1}^{T} K\left(\frac{\omega_{j_d} - \omega_{q_{2(d-n)-1}}}{b}\right) K\left(\frac{\omega_{j_d+h_d} - \omega_{q_{2(d-n)}}}{b}\right)$$

$$\times \prod_{v=1}^{k} \left[(\tilde{\mathcal{F}}_{\sum_s j_s; \lambda_{j_s}}^{(v_s)} : s \in P_v) + O\left(\frac{1}{T}\right)\right] \prod_{v=n+1}^{n+r} \left[(\tilde{\mathcal{F}}_{0; \lambda_{j_s}}^{(v_s)} : s \in P_v) + O\left(\frac{1}{T}\right)\right].$$

In exactly k sets of the partition there are exactly $k - 1$ equations of the form $y_s = \sum_s j_s$. In the above partition the first k sets yield the following set of equations:

$$y_i = (-1)^i (j_i + h_i - j_{i+1}), \quad i = 1, \ldots, n - 1,$$
$$\tilde{y}_i = (-1)^i (q_{2i-1} - q_{2i+1}), \quad i = 1, \ldots, r - 2,$$
$$\tilde{y}_{r-1} = (j_n - q_1),$$
$$\tilde{y}_r = (j_1 - q_{2r-1}).$$

By Corollary B.1 these equations correspond to $k - 1$ summations out of the total $k + 2r$ summations that are bounded. It can be verified that the above set of equations and an iterative change of variables shows that the other $2r + 1$ free variables are interrelated via the $2r$ kernel functions. These means that $2r$ sums can at most be of order bT, while one of them can be of order T. Consequently,

$$T^{k/2} \operatorname{cum}_{n,r}(\boldsymbol{\beta}_h^{(T)}, \mathcal{V}_h^{(T))}) = O(T^{-k/2}(bT)^{-2r}(bT)^{2r}T) = O(T^{-k/2+1}),$$

which converges to zero for $k > 2$ as $T \to \infty$, for any choice of n and r such that $n + r = k$. $\square$

SUPPLEMENTARY MATERIAL

Supplementary Material for "Testing for stationarity of functional time series in the frequency domain" (DOI: 10.1214/19-AOS1895SUPP; .pdf). The supplement contains additional simulation results and proof details.

REFERENCES

ANTONIADIS, A. and SAPATINAS, T. (2003). Wavelet methods for continuous-time prediction using Hilbert-valued autoregressive processes. *J. Multivariate Anal.* **87** 133–158. MR2007265 https://doi.org/10.1016/S0047-259X(03)00028-9

AUE, A. and HORVÁTH, L. (2013). Structural breaks in time series. *J. Time Series Anal.* **34** 1–16. MR3008012 https://doi.org/10.1111/j.1467-9892.2012.00819.x

AUE, A., DUBART NORINHO, D. D. and HÖRMANN, S. (2015). On the prediction of stationary functional time series. *J. Amer. Statist. Assoc.* **110** 378–392. MR3338510 https://doi.org/10.1080/01621459.2014.909317

AUE, A., RICE, G. and SÖNMEZ, O. (2018). Detecting and dating structural breaks in functional data without dimension reduction. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **80** 509–529. MR3798876 https://doi.org/10.1111/rssb.12257

AUE, A. and VAN DELFT, A. (2019). Supplement to "Testing for stationarity of functional time series in the frequency domain." https://doi.org/10.1214/19-AOS1895SUPP.

BANDYOPADHYAY, S., JENTSCH, C. and SUBBA RAO, S. (2017). A spectral domain test for stationarity of spatio-temporal data. *J. Time Series Anal.* **38** 326–351. MR3611747 https://doi.org/10.1111/jtsa.12222

BANDYOPADHYAY, S. and SUBBA RAO, S. (2017). A test for stationarity for irregularly spaced spatial data. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **79** 95–123. MR3597966 https://doi.org/10.1111/rssb.12161

BESSE, P., CARDOT, H. and STEPHENSON, D. (2000). Autoregressive forecasting of some functional climatic variations. *Scand. J. Stat.* **27** 673–687.

BOSQ, D. (2000). *Linear Processes in Function Spaces*: *Theory and Applications*. *Lecture Notes in Statistics* **149**. Springer, New York. MR1783138 https://doi.org/10.1007/978-1-4612-1154-9

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

BRILLINGER, D. R. and ROSENBLATT, M. (1967). Asymptotic theory of estimates of k-th order spectra. In *Spectral Analysis Time Series* (*Proc. Advanced Sem.*, *Madison*, *Wis.*, 1966) 153–188. John Wiley, New York. MR0211566

CREMERS, H. and KADELKA, D. (1986). On weak convergence of integral functionals of stochastic processes with applications to processes taking paths in L_p^E. *Stochastic Process. Appl.* **21** 305–317. MR0833957 https://doi.org/10.1016/0304-4149(86)90102-X

DAHLHAUS, R. (1997). Fitting time series models to nonstationary processes. *Ann. Statist.* **25** 1–37. MR1429916 https://doi.org/10.1214/aos/1034276620

DETTE, H., PREUSS, P. and VETTER, M. (2011). A measure of stationarity in locally stationary processes with applications to testing. *J. Amer. Statist. Assoc.* **106** 1113–1124. MR2894768 https://doi.org/10.1198/jasa.2011.tm10811

DWIVEDI, Y. and SUBBA RAO, S. (2011). A test for second-order stationarity of a time series based on the discrete Fourier transform. *J. Time Series Anal.* **32** 68–91. MR2790673 https://doi.org/10.1111/j.1467-9892.2010.00685.x

FERRATY, F. and VIEU, P. (2010). *Nonparametric Functional Data Analysis*: *Theory and Practice*. *Springer Series in Statistics*. Springer, New York. MR2229687

FREMDT, S., HORVÁTH, L., KOKOSZKA, P. and STEINEBACH, J. G. (2014). Functional data analysis with increasing number of projections. *J. Multivariate Anal.* **124** 313–332. MR3147328 https://doi.org/10.1016/j.jmva.2013.11.009

HALL, P. and HOSSEINI-NASAB, M. (2006). On properties of functional principal components analysis. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **68** 109–126. MR2212577 https://doi.org/10.1111/j.1467-9868.2005.00535.x

HÖRMANN, S., KIDZIŃSKI, Ł. and HALLIN, M. (2015). Dynamic functional principal components. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **77** 319–348. MR3310529 https://doi.org/10.1111/rssb.12076

HÖRMANN, S. and KOKOSZKA, P. (2010). Weakly dependent functional data. *Ann. Statist.* **38** 1845–1884. MR2662361 https://doi.org/10.1214/09-AOS768

HORVÁTH, L. and KOKOSZKA, P. (2012). *Inference for Functional Data with Applications*. *Springer Series in Statistics*. Springer, New York. MR2920735 https://doi.org/10.1007/978-1-4614-3655-3

HORVÁTH, L., KOKOSZKA, P. and RICE, G. (2014). Testing stationarity of functional time series. *J. Econometrics* **179** 66–82. MR3153649 https://doi.org/10.1016/j.jeconom.2013.11.002

HSING, T. and EUBANK, R. (2015). *Theoretical Foundations of Functional Data Analysis*, *with an Introduction to Linear Operators*. *Wiley Series in Probability and Statistics*. Wiley, Chichester. MR3379106 https://doi.org/10.1002/9781118762547

JENTSCH, C. and SUBBA RAO, S. (2015). A test for second order stationarity of a multivariate time series. *J. Econometrics* **185** 124–161. MR3300340 https://doi.org/10.1016/j.jeconom.2014.09.010

LEE, J. and RAO, S. S. (2017). A note on general quadratic forms of nonstationary stochastic processes. *Statistics* **51** 949–968. MR3698495 https://doi.org/10.1080/02331888.2017.1318880

LI, Y. and HSING, T. (2010). Uniform convergence rates for nonparametric regression and principal component analysis in functional/longitudinal data. *Ann. Statist.* **38** 3321–3351. MR2766854 https://doi.org/10.1214/10-AOS813

MAS, A. and MENNETEAU, L. (2003). Perturbation approach applied to the asymptotic study of random operators. In *High Dimensional Probability*, *III* (*Sandjberg*, 2002). *Progress in Probability* **55** 127–134. Birkhäuser, Basel. MR2033885

PANARETOS, V. M. and TAVAKOLI, S. (2013). Fourier analysis of stationary time series in function space. *Ann. Statist.* **41** 568–603. MR3099114 https://doi.org/10.1214/13-AOS1086

PAPARODITIS, E. (2009). Testing temporal constancy of the spectral structure of a time series. *Bernoulli* **15** 1190–1221. MR2597589 https://doi.org/10.3150/08-BEJ179

PREUSS, P., VETTER, M. and DETTE, H. (2013). A test for stationarity based on empirical processes. *Bernoulli* **19** 2715–2749. MR3160569 https://doi.org/10.3150/12-BEJ472

PRIESTLEY, M. B. and SUBBA RAO, T. (1969). A test for non-stationarity of time-series. *J. Roy. Statist. Soc. Ser. B* **31** 140–149. MR0269062

RAMSAY, J. O. and SILVERMAN, B. W. (2005). *Functional Data Analysis*, 2nd ed. *Springer Series in Statistics*. Springer, New York. MR2168993

SUBBA RAO, S. (2018). Orthogonal samples for estimators in time series. *J. Time Series Anal.* **39** 313–337. MR3796522 https://doi.org/10.1111/jtsa.12269

VAN DELFT, A. and EICHLER, M. (2018). Locally stationary functional time series. *Electron. J. Stat.* **12** 107–170. MR3746979 https://doi.org/10.1214/17-EJS1384

VAN DELFT, A. and EICHLER, M. (2019) A note on Herglotz's theorem for time series on function spaces. https://doi.org/10.1016/j.spa.2019.10.006, forthcoming.

VON SACHS, R. and NEUMANN, M. H. (2000). A wavelet-based test for stationarity. *J. Time Series Anal.* **21** 597–613. MR1794489 https://doi.org/10.1111/1467-9892.00200

WILKINSON, H. (1965). *The Algebraic Eigenvalue Problem*. Clarendon Press, Oxford.