Testing for stationarity of functional time series in the frequency domain[*][†]

Alexander Aue[‡] Anne van Delft[§]

October 6, 2017

Abstract

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 stationarity tests for functional time series based on frequency domain methods. Setting up the tests requires a delicate understanding of periodogram- and spectral density operators that are the functional counterparts of periodogram- and spectral density matrices in the multivariate world. Two sets of statistics are proposed. One is based on the eigendecomposition of the spectral density operator, the other on a fixed projection basis. Their properties 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 tests work well in finite samples.

Keywords: Frequency domain methods, Functional data analysis, Locally stationary processes, Spectral analysis

MSC 2010: Primary: 62G99, 62H99, Secondary: 62M10, 62M15, 91B84

1 Introduction

The aim of this paper is to provide new stationarity tests 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 non-stationarity, even though non-smooth alternatives can be covered as well. 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

[*]The authors sincerely thank the Associate Editor and two referees for their constructive comments that helped produce a much improved revision of the original paper.

[†]AA was partially supported by NSF grants DMS 1305858 and DMS 1407530. AvD was partially supported by Maastricht University, the contract "Projet d'Actions de Recherche Concertées" No. 12/17-045 of the "Communauté française de Belgique" and by the Collaborative Research Center "Statistical modeling of nonlinear dynamic processes" (SFB 823, Project A1, C1, A7) of the German Research Foundation (DFG).

[‡]Department of Statistics, University of California, Davis, CA 95616, USA, email: aaue@ucdavis.edu

[§]Ruhr-Universität Bochum, Fakultät für Mathematik, 44780 Bochum, Germany, email: Anne.vanDelft@rub.de

may consult Bosq (2000), Ferraty & Vieu (2010), Horváth & Kokoszka (2012), Hsing & Eubank (2015) and Ramsay & Silverman (2005).

Papers on functional time series have come into the focus more recently and constitute now an active area of research. Hörmann & Kokoszka (2010) introduced a general weak dependence concept for stationary functional time series, while van Delft & Eichler (2016) provided a framework for locally stationary functional time series. Antoniadis & Sapatinas (2003), Aue et al. (2015) and Besse et al. (2000) constructed prediction methodology that may find application across many areas of science, economics and finance. With the exception of van Delft & Eichler (2016), 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 & Tavakoli (2013), who provided results concerning the Fourier analysis of time series in function spaces, and Hörmann et al. (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 & 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 & 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 et al. (2011) and Preuß et al. (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 have power against non-smooth alternatives such as structural breaks or change-points. A recent review discussing methodology for structural breaks in time series is Aue & Horváth (2013), while Aue et al. (2017) is a recent contribution to structural breaks in functional time series.

The proposed test for second-order stationarity of functional time series uses the Discrete Fourier Transform (DFT). Its construction seeks to exploit that the DFTs of a functional time series evaluated at distinct Fourier frequencies are asymptotically uncorrelated if and only if the series is second-order stationary. The proposed method is therefore related to the initial work of Dwivedi & Subba Rao (2011), who put forth similar tests in a univariate framework. Their method has since been generalized to multivariate time series in Jentsch & Subba Rao (2015) as well as to spatial and spatio-temporal data by Bandyopadhyay & Subba Rao (2017) and Bandyopadhyay et al. (2017), respectively. A different version of functional stationarity tests, based on time domain methodology involving cumulative sum statistics (Aue & Horváth, 2013), was given in Horváth et al. (2014).

Building on the research summarized in the previous paragraph, the tests introduced here are the first of their kind in the frequency domain analysis of functional time series. A delicate understanding of the functional DFT is needed in order to derive the asymptotic theory given here. In particular, results on the large-sample behavior of a quadratic form test statistic are provided both under the null hypothesis of a stationary functional time series and the alternative of a locally stationary functional time series. For this, a weak convergence result is established that might be of interest in its own right as it is verified using a simplified tightness criterion going back to work of Cremers & Kadelka (1986). 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 & Hsing (2010).

The remainder of the paper is organized as follows. Section 2 provides the background, gives the requisite notations and introduces the functional version of the DFT. 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 the Appendix. Several further auxiliary results are proved in the supplementary document Aue & van Delft (2017), henceforth referred to simply as the Online Supplement.

2 Notation and setup

2.1 The function space

A functional time series $(X_t \colon t \in \mathbb{Z})$ will be viewed in this paper as a sequence of random elements on a probability space $(\Omega, \mathcal{A}, P)$ taking values in the separable Hilbert space of real-valued, square integrable functions on the unit interval $[0,1]$. This Hilbert space will be denoted by $H_{\mathbb{R}} = L^2([0,1], \mathbb{R})$. The functional DFT of $(X_t \colon t \in \mathbb{Z})$, to be introduced in Section 2.3, can then be viewed as an element of $H_{\mathbb{C}} = L^2([0,1], \mathbb{C})$, the complex counterpart of $H_{\mathbb{R}}$. While the interval $[0,1]$ provides a convenient parametrization of the functions, the results of this paper continue to hold for any separable Hilbert space.

The complex conjugate of $z \in \mathbb{C}$ is denoted by $\overline{z}$ and the imaginary number by i. The inner product and the induced norm on $H_{\mathbb{C}}$ are given by

$$\langle f, g \rangle = \int_0^1 f(\tau)\overline{g(\tau)} d\tau \qquad \text{and} \qquad \|f\|_2 = \sqrt{\langle f, f \rangle}, \tag{2.1}$$

respectively, for $f, g \in H_{\mathbb{C}}$. Two elements of $H_{\mathbb{C}}$ are tacitly understood to be equal if their difference has vanishing L_2-norm. More generally, for functions $g \colon [0,1]^k \to \mathbb{C}$, the supremum norm is denoted by

$$\|g\|_\infty = \sup_{\tau_1,\ldots,\tau_k \in [0,1]} |g(\tau_1, \ldots, \tau_k)|$$

and the L^p-norm by

$$\|g\|_p = \left(\int_{[0,1]^k} |g(\tau_1, \ldots, \tau_k)|^p \, d\tau_1 \cdots d\tau_k \right)^{1/p}.$$

In all of the above, the obvious modifications apply to $H_{\mathbb{R}}$, the canonical Hilbert space in the functional data analysis setting.

Let H stand more generally for $H_{\mathbb{C}}$, unless otherwise stated. An operator A on H is said to be compact if its pre-image is compact. A compact operator admits a *singular value decomposition*

$$A = \sum_{n=1}^{\infty} s_n(A)\, \psi_n \otimes \phi_n, \tag{2.2}$$

where $(s_n(A) \colon n \in \mathbb{N})$, are the *singular values* of A, $(\phi_n \colon n \in \mathbb{N})$ and $(\psi_n \colon n \in \mathbb{N})$ orthonormal bases of H and $\otimes$ denoting the tensor product. The singular values are ordered to form a monotonically decreasing sequence of non-negative numbers. Based on the convergence rate to zero, operators on H can be classified into particular *Schatten p-classes*. That is, for $p \geqslant 1$, the *Schatten p-class* $S_p(H)$ is the subspace of all compact operators A on H such that the sequence $s(A) = (s_n(A) \colon n \in \mathbb{N})$ of singular values of A belongs to the sequence space ℓ^p, that is,

$$A \in S_p(H) \qquad \text{if and only if} \qquad \|\!|A|\!\|_p = \left(\sum_{n=1}^{\infty} s_n^p(A) \right)^{1/p} < \infty,$$

where $\|\!|A|\!\|_p$ is referred to as the *Schatten p-norm*. The space $S_p(H)$ together with the norm $\|\!|A|\!\|_p$ forms a Banach space and a Hilbert space in case $p = 2$. By convention, the space $S_\infty(H)$ indicates the space of bounded linear operators equipped with the standard operator norm. For $1 \leqslant p \leqslant q$, the inclusion $S_p(H) \subseteq S_q(H)$ is valid. Two important classes are the Trace-class and the Hilbert-Schmidt operators on H, which are given by $S_1(H)$ and $S_2(H)$, respectively. More properties of Schatten-class operators, in particular of Hilbert–Schmidt operators, are provided in van Delft & Eichler (2016). Finally, the identity operator is denoted by I_H and the zero operator by O_H.

2.2 Dependence structure on the function space

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. The L^2-setup provides a point-wise interpretation to these moments; that is, the mean function m of X can be defined using the parametrization $m(\tau) = \mathbb{E}[X_t(\tau)]$, $\tau \in [0, 1]$, and the autocovariance kernel c_h at lag $h \in \mathbb{Z}$ by

$$c_h(\tau, \tau') = \mathrm{Cov}(X_{t+h}(\tau), X_t(\tau')), \qquad \tau, \tau' \in [0, 1]. \tag{2.3}$$

Both m and c_h are well defined in the L^2-sense if $\mathbb{E}[\|X_0\|_2^2] < \infty$. Each kernel c_h induces a corresponding *autocovariance operator* $\mathcal{C}_h$ on $H_{\mathbb{R}}$ by

$$\mathcal{C}_h\, g(\tau) = \int_0^1 c_h(\tau, \tau')\, g(\tau')\, d\tau' = \mathbb{E}\big[\langle g, X_0 \rangle X_h(\tau)\big], \tag{2.4}$$

for all $g \in H_{\mathbb{R}}$. In analogy to weakly stationary multivariate time series, where the covariance matrix and spectral density matrix form a Fourier pair, the *spectral density operator* $\mathcal{F}_\omega$ is given by the Fourier transform of $\mathcal{C}_h$,

$$\mathcal{F}_\omega = \frac{1}{2\pi} \sum_{h \in \mathbb{Z}} \mathcal{C}_h\, e^{-\mathrm{i}\omega h}. \tag{2.5}$$

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

As in Panaretos & Tavakoli (2013), higher-order dependence among the functional observations is defined through cumulant mixing conditions (Brillinger, 1981; Brillinger & Rosenblatt, 1967). For this, the notion of higher-order cumulant tensors is required; see Appendix B for their definition and a discussion on their properties for nonstationary functional time series. Point-wise, after setting $t_k = 0$ because of stationarity, the k-th order cumulant function of the process X can be defined by

$$c_{t_1,\ldots,t_{k-1}}(\tau_1,\ldots,\tau_k) = \sum_{\nu=(\nu_1,\ldots,\nu_p)} (-1)^{p-1}\,(p-1)! \prod_{l=1}^{p} \mathbb{E}\Big[\prod_{j\in\nu_l} X_{t_j}(\tau_j)\Big], \tag{2.6}$$

where the summation is over all unordered partitions of $\{1,\ldots,k\}$. The quantity in (2.6) will be referred to as the *k-th order cumulant kernel* if it is properly defined in the L^2-sense. A sufficient condition for this to be satisfied is $\mathbb{E}[\|X_0\|_2^k] < \infty$. The cumulant kernel $c_{t_1,\ldots,t_{2k-1}}(\tau_1,\ldots,\tau_{2k})$ induces a *2k-th order cumulant operator* $\mathcal{C}_{t_1,\ldots,t_{2k-1}}$ through right integration:

$$\mathcal{C}_{t_1,\ldots,t_{2k-1}} g(\tau_1,..,\tau_k) = \int_{[0,1]^k} c_{t_1,\ldots,t_{2k-1}}(\tau_1,\ldots,\tau_{2k}) g(\tau_{k+1},\ldots,\tau_{2k})\, d\tau_{k+1}\cdots d\tau_{2k},$$

which maps from $L^2([0,1]^k,\mathbb{R})$ to $L^2([0,1]^k,\mathbb{R})$. Similar to the case $k=2$, this operator will form a Fourier pair with a $2k$-th order cumulant spectral operator given summability with respect to $|||\cdot|||_p$ is satisfied. The $2k$-th order cumulant spectral operator is specified as

$$\mathcal{F}_{\omega_1,\ldots,\omega_{2k-1}} = (2\pi)^{1-2k} \sum_{t_1,\ldots,t_{2k-1}\in\mathbb{Z}} \mathcal{C}_{t_1,\ldots,t_{2k-1}} \exp\Big(-\mathrm{i}\sum_{j=1}^{2k-1}\omega_j\, t_j\Big), \tag{2.7}$$

where the convergence is in $|||\cdot|||_p$. Under suitable regularity conditions, the corresponding kernels also form a Fourier pair. Note that a similar reasoning can be applied to $2(k+1)$-th order cumulant operators and their associated cumulant spectral operators, using more complex notation.

2.3 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.

Proposition 2.1. *A zero-mean, $H_\mathbb{R}$-valued stochastic process $(X_t\colon t\in\mathbb{Z})$ admits the representation*

$$X_t = \int_{-\pi}^{\pi} e^{\mathrm{i}t\omega} dZ_\omega \qquad a.s.\ a.e., \tag{2.8}$$

where $(Z_\omega\colon \omega\in(-\pi,\pi])$ 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. It can be shown that a time-dependent functional Cramér representation holds true if the time-dependent characteristics of the process are captured by a Bochner measurable mapping that is an evolutionary operator-valued mapping in time direction. More details can be found in van Delft & Eichler (2016). In the following the above proposition is utilized to motivate the proposed test procedure.

In practice, the stretch $X_0, \ldots, X_{T-1}$ is observed. If the process is weakly stationary, the *functional Discrete Fourier Transform* (fDFT) $D_\omega^{(T)}$ can be seen as an estimate of the increment process Z_ω. This is in accordance with the spatial and spatio-temporal arguments put forward in Jentsch & Subba Rao (2015) and Bandyopadhyay et al. (2017), respectively. At frequency ω, the fDFT is given by

$$D_\omega^{(T)} = \frac{1}{\sqrt{2\pi T}} \sum_{t=0}^{T-1} X_t e^{-\mathrm{i}\omega t}, \tag{2.9}$$

while the functional time series itself can be represented through the inverse fDFT as

$$X_t = \sqrt{\frac{2\pi}{T}} \sum_{j=0}^{T-1} D_{\omega_j}^{(T)} e^{\mathrm{i}\omega_j t}. \tag{2.10}$$

Under weak dependence conditions expressed through higher-order cumulants, the fDFTs evaluated at distinct frequencies yield asymptotically independent Gaussian random elements in $H_{\mathbb{C}}$ (Panaretos & Tavakoli, 2013). The fDFT sequence of a Hilbert space stationary process is in particular asymptotically uncorrelated at the canonical frequencies $\omega_j = 2\pi j/T$. For weakly stationary processes, it will turn out specifically that, for $j \neq j'$ or $j \neq T - j'$, the covariance of the fDFT satisfies $\langle \mathrm{Cov}(D_{\omega_j}^{(T)}, D_{\omega_{j'}}^{(T)}) g_1, g_2 \rangle = O(1/T)$ for all $g_1, g_2 \in H$. This fundamental result will be extended to locally stationary functional time series in this paper by providing a representation of the cumulant kernel of the fDFT sequence in terms of the (time-varying) spectral density kernel. 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 (2.9), the covariance operator of X_t and $X_{t'}$ in terms of the fDFT sequence can be written as

$$\begin{aligned}
\mathcal{C}_{t,t'} &= \mathbb{E}[X_t \otimes X_{t'}] \\
&= \frac{2\pi}{T} \sum_{j,j'=0}^{T-1} \mathbb{E}[D_{\omega_j}^{(T)} \otimes D_{\omega_{j'}}^{(T)}] e^{\mathrm{i}t\omega_j - \mathrm{i}t'\omega_{j'}} \\
&= \frac{2\pi}{T} \sum_{j=0}^{T-1} \mathbb{E}[I_{\omega_j}^{(T)}] e^{\mathrm{i}\omega_j (t-t')} \qquad (2.11) \\
&= \mathcal{C}_{t-t'},
\end{aligned}$$

where the equality in (2.11) holds in an L^2-sense when $\langle \mathbb{E}[D_{\omega_j}^{(T)} \otimes D_{\omega_{j'}}^{(T)}] g_1, g_2 \rangle = 0$ for all $g_1, g_2 \in H$ with $j \neq j'$ or $j \neq T - j'$ and where $I_{\omega_j}^{(T)} = D_{\omega_{j'}}^{(T)} \otimes D_{\omega_{j'}}^{(T)}$ is the periodogram tensor. This demonstrates that the autocovariance kernel of a second-order stationary functional time series is obtained and, hence, that an uncorrelated DFT sequence implies second-order stationarity up to lag T. Below, the behavior of the fDFT under the smooth alternative of locally stationary functional time series is derived. These properties will then be exploited to set up a testing framework for functional stationarity. Although this work is therefore related to Dwivedi & Subba Rao (2011) — who similarly exploited analogous properties of the DFT of a stationary time series — it should be emphasized that the functional setting is a nontrivial extension.

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

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

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 3.1. *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

$$\big\|X_t^{(T)} - X_t^{(u)}\big\|_2 \leqslant \Big(\Big|\frac{t}{T} - u\Big| + \frac{1}{T}\Big) P_{t,T}^{(u)} \qquad a.s.,$$

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

Note that, under H_A, the process constitutes a triangular array of functions. Inference methods are then based on in-fill asymptotics as popularized in Dahlhaus (1997) for univariate time series. The process is therefore observed on a finer grid as T increases and more observations are available at a local level. A rigorous statistical framework for locally stationary functional time series was recently provided in van Delft & Eichler (2016). These authors established in particular that linear functional time series can be defined by means of a functional time-varying Cramér representation and provided sufficient conditions in the frequency domain for the above definition to be satisfied.

Based on the observations in Section 2.3, a test for weak stationarity can be set up exploiting the uncorrelatedness of the elements in the sequence $(D_{\omega_j}^{(T)})$. Standardizing these quantities is a delicate issue as the spectral density operators $\mathcal{F}_{\omega_j}^{(T)}$ are not only unknown but generally not globally invertible. Here, a statistic based on projections is considered. Let $(\psi_l\colon l \in \mathbb{N})$ be an orthonormal basis of $H_{\mathbb{C}}$. Then, $(\psi_l \otimes \psi_{l'}\colon l, l' \in \mathbb{N})$ is an orthonormal basis of $L^2([0,1]^2, \mathbb{C})$ and, by definition of the Hilbert–Schmidt inner product on the algebraic tensor product space $H \otimes H$,

$$\langle \mathbb{E}\big[D_{\omega_{j_1}}^{(T)} \otimes D_{\omega_{j_2}}^{(T)}\big], \psi_l \otimes \psi_{l'} \rangle_{H \otimes H} = \mathbb{E}[\langle D_{\omega_{j_1}}^{(T)}, \psi_l \rangle \langle D_{\omega_{j_2}}^{(T)}, \psi_{l'} \rangle] = \mathrm{Cov}\big(\langle D_{\omega_{j_1}}^{(T)}, \psi_l \rangle, \overline{\langle D_{\omega_{j_2}}^{(T)}, \psi_{l'} \rangle}\big).$$

This motivates to set up test statistics based on the quantities

$$\gamma_h^{(T)}(l,l') = \frac{1}{T}\sum_{j=1}^{T} \frac{\langle D_{\omega_j}^{(T)}, \psi_l\rangle \overline{\langle D_{\omega_{j+h}}^{(T)}, \psi_{l'}\rangle}}{\sqrt{\langle \mathcal{F}_{\omega_j}(\psi_l), \psi_l\rangle \langle \mathcal{F}_{\omega_{j+h}}^{\dagger}(\psi_{l'}), \psi_{l'}\rangle}}, \qquad h = 1,\ldots,T-1. \tag{3.1}$$

A particularly interesting (random) projection choice is provided by the eigenfunctions $\phi_l^{\omega_j}$ and $\phi_l^{\omega_{j+h}}$ corresponding to the l-th largest eigenvalues of the spectral density operators $\mathcal{F}_{\omega_j}$ and $\mathcal{F}_{\omega_{j+h}}$, respectively. In this case, based on orthogonality relations the statistic can be simplified to

$$\gamma_h^{(T)}(l) = \frac{1}{T}\sum_{j=1}^{T} \frac{\langle D_{\omega_j}^{(T)}, \phi_l^{\omega_j}\rangle \overline{\langle D_{\omega_{j+h}}^{(T)}, \phi_l^{\omega_{j+h}}\rangle}}{\sqrt{\lambda_l^{-\omega_j}\lambda_l^{\omega_{j+h}}}}, \qquad h = 1,\ldots,T-1, \tag{3.2}$$

and thus depends on l only. In the following the notation $\gamma_h^{(T)}$ is used to refer both to a fixed projection basis used for dimension reduction at both frequencies ω_j and ω_{j+h} as in (3.1) as well as for the random projection based method in (3.2) using two separate Karhunen–Loève decompositions at the two frequencies of $\mathcal{F}_{\omega_j}$ and $\mathcal{F}_{\omega_{j+h}}$, respectively. The write-up will generally present theorems for both tests if the formulation allows, but will focus on the latter case if certain aspects have to be treated differently. The counterparts for the former case are then collected in the appendix for completeness.

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 $\hat{\mathcal{F}}_{\omega_{j+h}}^{(T)}$. Similarly the population eigenvalues and eigenfunctions need to be replaced by the respective sample eigenvalues and eigenfunctions. The estimated quantity corresponding to (3.1) and (3.2) will be denoted by $\hat{\gamma}_h^{(T)}$. Note that the large-sample distributional properties of tests based on (3.1) and (3.2) are similar. Some additional complexity enters for the statistic in (3.2) because the unknown eigenfunctions can only be identified up to rotation on the unit circle and it has to be ensured that the proposed test statistic is invariant to this rotation; see Section E.2 for the details. As an estimator of $\mathcal{F}_{\omega}^{(T)}$, take

$$\hat{\mathcal{F}}_{\omega}^{(T)} = \frac{2\pi}{T}\sum_{j=1}^{T} K_b(\omega-\omega_j)\big([D_{\omega_j}^{(T)}] \otimes [D_{\omega_j}^{(T)}]^{\dagger}\big), \tag{3.3}$$

where $K_b(\cdot) = \frac{1}{b}K(\frac{\cdot}{b})$ is a window function satisfying the following conditions.

Assumption 3.1. Let $K\colon [-\frac{1}{2},\frac{1}{2}] \to \mathbb{R}$ be a positive, symmetric window function with $\int K(x)dx = 1$ and $\int K(x)^2 dx < \infty$ that is periodically extended, i.e., $K_b(x) = \frac{1}{b}K(\frac{x\pm 2\pi}{b})$.

The periodic extension is to include estimates for frequencies around $\pm\pi$. Further conditions on the bandwidth b are imposed below (Section 4) to determine the large-sample behavior of $\hat{\gamma}_h^{(T)}$ under both the null and the alternative. The replacement of the unknown operators with consistent estimators requires to derive the order of the difference

$$\sqrt{T}\big|\gamma_h^{(T)} - \hat{\gamma}_h^{(T)}\big|. \tag{3.4}$$

It will be shown in the next section that, for appropriate choices of the bandwidth b, this term is negligible under both null and alternative hypothesis.

To set up the test statistic, it now appears reasonable to extract information across a range of directions $l, l' = 1, \ldots, L$ and a selection of lags $h = 1, \ldots, \bar{h}$, where $\bar{h}$ denotes an upper limit. Build therefore first the $L \times L$ matrix $\hat{\Gamma}_h^{(T)} = (\hat{\gamma}_h^{(T)}(l, l') \colon l, l' = 1, \ldots, L)$ and construct the vector $\hat{\boldsymbol{\gamma}}_h^{(T)} = \mathrm{vec}(\hat{\Gamma}_h^{(T)})$ by vectorizing $\hat{\Gamma}_h^{(T)}$ via stacking of its columns. Define

$$\hat{\boldsymbol{\beta}}_h^{(T)} = \boldsymbol{e}^\top \hat{\boldsymbol{\gamma}}_h^{(T)}, \tag{3.5}$$

where $\boldsymbol{e}$ is a vector of dimension L^2 whose elements are all equal to one and $^\top$ denotes transposition. Note that for (3.2), a somewhat simplified expression is obtained, since the off-diagonal terms $l \neq l'$ do not have to be accounted for. Dropping these components reduces the number of terms in the sum (3.5) from L^2 to L, even though the general form of the equation continues to hold. Choose next a collection $h_1, \ldots, h_M$ of lags each of which is upper bounded by $\bar{h}$ to pool information across a number of autocovariances and build the (scaled) vectors

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

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

$$\hat{Q}_M^{(T)} = T(\hat{\boldsymbol{b}}_M^{(T)})^\top \hat{\Sigma}_M^{-1} \hat{\boldsymbol{b}}_M^{(T)}, \tag{3.6}$$

where $\hat{\Sigma}_M$ is an estimator of the asymptotic covariance matrix of the vectors $\boldsymbol{b}_M^{(T)}$ which are defined by replacing $\hat{\gamma}_h^{(T)}$ with $\gamma_h^{(T)}$ in the definition of $\hat{\boldsymbol{b}}_M^{(T)}$. The statistic $\hat{Q}_M^{(T)}$ will be used to test the null of stationarity against the alternative of local stationarity. Note that this quadratic form depends on the tuning parameters L and M. These effects will be briefly discussed in Section 5.

4 Large-sample results

4.1 Properties under the null of stationarity

The following gives the main requirements under stationarity of the functional time series that are needed to establish the asymptotic behavior of the test statistics under the null hypothesis.

Assumption 4.1 (Stationary functional time series). *Let $(X_t \colon t \in \mathbb{Z})$ be a stationary functional time series with values in $H_{\mathbb{R}}$ such that*

(i) $\mathbb{E}[\|X_0\|_2^k] < \infty$,

(ii) $\sum_{t_1,\ldots,t_{k-1}=-\infty}^{\infty} (1 + |t_j|^\ell) \|c_{t_1,\ldots,t_{k-1}}\|_2 < \infty$ *for all* $1 \leqslant j \leqslant k-1$,

for some fixed values of $k, \ell \in \mathbb{N}$.

The conditions of Assumption 4.1 ensure that the k-th order cumulant spectral density kernel

$$f_{\omega_1,\ldots,\omega_{k-1}}(\tau_1, \ldots, \tau_k) = \frac{1}{(2\pi)^{k-1}} \sum_{t_1,\ldots,t_{k-1}=-\infty}^{\infty} c_{t_1,\ldots,t_{k-1}}(\tau_1, \ldots, \tau_k) e^{-\mathrm{i}(\sum_{j=1}^{k} \omega_j t_j)} \tag{4.1}$$

is well-defined in L^2 and is uniformly continuous in ω with respect to $\|\cdot\|_2$. Additionally, the parameter ℓ controls the smoothness of f_ω in the sense that, for all $i \leqslant \ell$,

$$\sup_{\omega} \left\| \frac{\partial^i}{\partial \omega^i} f_\omega \right\|_2 < \infty. \tag{4.2}$$

A proof of these facts can be found in Panaretos & Tavakoli (2013). Through right-integration, the function (4.1) induces a k-th order cumulant spectral density operator $\mathcal{F}_{\omega_1,\ldots,\omega_{k-1}}$ which is Hilbert–Schmidt. The following theorem establishes that the scaled difference between $\gamma_h^{(T)}$ and $\hat{\gamma}_h^{(T)}$ is negligible in large samples.

Theorem 4.1. *Let Assumptions 3.1 and 4.1 be satisfied with $k = 8$ and $\ell = 2$ and assume further that $\inf_\omega \langle \mathcal{F}(\psi), \psi \rangle > 0$ for $\psi \in H$. Then, for any fixed h,*

$$\sqrt{T} \big| \gamma_h^{(T)} - \hat{\gamma}_h^{(T)} \big| = O_p\left(\frac{1}{\sqrt{bT}} + b^2 \right) \qquad (T \to \infty).$$

The proof is given in Section S2 of the Online Supplement. For the eigen-based statistic (3.2), the condition $\inf_\omega \langle \mathcal{F}(\psi), \psi \rangle > 0$ reduces to $\inf_\omega \lambda_L^\omega > 0$. In practice this means that only those eigenfunctions should be included that belong to the l-largest eigenvalues in the quadratic form (3.6). This is discussed in more detail in Section 5, where a criterion is proposed to choose L. Theorem 4.1 shows that the distributional properties of $\hat{\gamma}_h^{(T)}$ are asymptotically the same as those of $\gamma_h^{(T)}$, or both versions (3.1) and (3.2) of the test, provided that the following extra condition on the bandwidth holds.

Assumption 4.2. *The bandwidth b satisfies $b \to 0$ such that $bT \to \infty$ as $T \to \infty$.*

Note that these rates are in fact necessary for the estimator in (3.3) to be consistent, see Panaretos & Tavakoli (2013), and therefore do not impose an additional constraint under H_0. The next theorem derives the second-order structure of $\gamma_h^{(T)}$ for the eigenbased statistics (3.2). It shows that the asymptotic variance is uncorrelated for all lags h and that there is no correlation between the real and imaginary parts.

Theorem 4.2. *Let Assumption 4.1 be satisfied with $k = \{2, 4\}$. Then,*

$$T\,\mathrm{Cov}\Big(\Re\gamma_{h_1}^{(T)}(l_1), \Re\gamma_{h_2}^{(T)}(l_2) \Big) = T\,\mathrm{Cov}\Big(\Im\gamma_{h_1}^{(T)}(l_1), \Im\gamma_{h_2}^{(T)}(l_2) \Big)$$

$$= \begin{cases} \dfrac{\delta_{l_1,l_2}}{2} + \dfrac{1}{4\pi} \displaystyle\int\!\!\int \frac{\langle \mathcal{F}_{\omega,-\omega-\omega_h,-\omega'}(\phi_{l_2}^{\omega'} \otimes \phi_{l_2}^{\omega'+\omega_h'}), \phi_{l_1}^{\omega} \otimes \phi_{l_1}^{\omega+\omega_h} \rangle}{\sqrt{\lambda_{l_1}^{\omega} \lambda_{l_2}^{-\omega'} \lambda_{l_1}^{-\omega-\omega_h} \lambda_{l_2}^{\omega'+\omega_h'}}} d\omega d\omega', & \text{if } h_1 = h_2 = h, \\ O(\frac{1}{T}), & \text{if } h_1 \neq h_2, \end{cases}$$

where $\delta_{i,j} = 1$ if $i = j$ and 0 otherwise. Furthermore,

$$T\,\mathrm{Cov}(\Re\gamma_{h_1}^{(T)}(l_1), \Im\gamma_{h_2}^{(T)}(l_2)) = O\left(\frac{1}{T}\right)$$

uniformly in $h_1, h_2 \in \mathbb{Z}$.

The proof of Theorem 4.2 is given in Appendix D.1, its counterpart for (3.1) is stated as Theorem C.1 in Appendix C. Note that the results in the theorem use at various instances the fact that the k-th order spectral density operator at frequency $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_k)^T \in \mathbb{R}^k$ is equal to the k-th 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 $\hat{Q}_M^{(T)}$ defined in (3.6) can be derived. This is done in the following theorem.

Theorem 4.3. *Let Assumptions 3.1 and 4.2 be satisfied. Let Assumption 4.1 be satisfied with* $k \geqslant 1$ *and* $\ell = 2$ *and assume that* $\inf_\omega \lambda_L^\omega > 0$. *Then,*

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

$$\sqrt{T}\boldsymbol{b}_M^{(T)} \xrightarrow{\mathcal{D}} \mathcal{N}_{2M}(\mathbf{0}, \Sigma_0) \qquad (T \to \infty),$$

where $\xrightarrow{\mathcal{D}}$ *denotes convergence in distribution and* $\mathcal{N}_{2M}(\mathbf{0}, \Sigma_0)$ *a* $2M$*-dimensional normal distribution with mean* $\mathbf{0}$ *and diagonal covariance matrix* $\Sigma_0 = \mathrm{diag}(\sigma_{0,m}^2 \colon m = 1, \ldots, 2M)$ *whose elements are*

$$\sigma_{0,m}^2 = \lim_{T\to\infty} \sum_{l_1,l_2=1}^{L} T\mathrm{Cov}\big(\Re\gamma_{h_m}^{(T)}(l_1), \Re\gamma_{h_m}^{(T)}(l_2)\big), \qquad m = 1, \ldots, M,$$

and $\sigma_{0,M+m}^2 = \sigma_{0,m}^2$. *The explicit form of the limit is given by Theorem 4.2.*

(b) Using the result in (a), it follows that

$$\hat{Q}_M^{(T)} \xrightarrow{\mathcal{D}} \chi_{2M}^2 \qquad (T \to \infty),$$

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

The proof of Theorem 4.3 is provided in Appendix D.1. Part (b) of the theorem can now be used to construct tests with asymptotic level α. Theorem C.2 contains the result for the test based on (3.1). To better understand the power of the test, the next section investigates the behavior under the alternative of local stationarity.

4.2 Properties under the alternative

This section contains the counterparts of the results in Section 4.1 for locally stationary functional time series. The following conditions are essential for the large-sample results to be established here.

Assumption 4.3. Assume $(X_t^{(T)} \colon t \leqslant T, T \in \mathbb{N})$ and $(X_t^{(u)} \colon t \in \mathbb{Z})$ are as in Definition 3.1 and let $\kappa_{k;t_1,\ldots,t_{k-1}}$ be a positive sequence in $L^2([0,1]^k, \mathbb{R})$ independent of T such that, for all $j = 1, \ldots, k-1$ and some $\ell \in \mathbb{N}$,

$$\sum_{t_1,\ldots,t_{k-1}\in\mathbb{Z}} (1 + |t_j|^\ell)\|\kappa_{k;t_1,\ldots,t_{k-1}}\|_2 < \infty. \tag{4.3}$$

Suppose furthermore that there exist representations

$$X_t^{(T)} - X_t^{(t/T)} = Y_t^{(T)} \qquad \text{and} \qquad X_t^{(u)} - X_t^{(v)} = (u-v)Y_t^{(u,v)}, \tag{4.4}$$

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

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

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

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

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

Note that these assumptions are generalizations of the ones in Lee & Subba Rao (2016), who investigated the properties of quadratic forms of stochastic processes in a finite-dimensional setting. For fixed u_0, the process $(X_t^{(u_0)} : t \in \mathbb{Z})$ is stationary and thus the results of van Delft & Eichler (2016) imply that the *local k-th order cumulant spectral kernel*

$$f_{u_0;\omega_1,\ldots,\omega_{k-1}}(\tau_1, \ldots, \tau_k) = \frac{1}{(2\pi)^{k-1}} \sum_{t_1,\ldots,t_{k-1} \in \mathbb{Z}} c_{u_0;t_1,\ldots,t_{k-1}}(\tau_1, \ldots, \tau_k) e^{-\mathrm{i}\sum_{l=1}^{k-1} \omega_l t_l} \tag{4.5}$$

exists, where $\omega_1, \ldots, \omega_{k-1} \in [-\pi, \pi]$ and

$$c_{u_0;t_1,\ldots,t_{k-1}}(\tau_1, \ldots, \tau_k) = \mathrm{cum}\big(X_{t_1}^{(u_0)}(\tau_1), \ldots, X_{t_{k-1}}^{(u_0)}(\tau_{k-1}), X_{t_0}^{(u_0)}(\tau_k)\big) \tag{4.6}$$

is the corresponding local cumulant kernel of order k at time u_0. The quantity $f_{u,\omega}$ will be referred to as the *time-varying spectral density kernel* of the stochastic process $(X_t^{(T)} : t \leqslant T, T \in \mathbb{N})$. Under the given assumptions, this expression is formally justified by Lemma **??**.

Because of the standardization necessary in $\hat{\gamma}^{(T)}$, it is of importance to consider the properties of the estimator (3.3) in case the process is locally stationary. The next theorem shows that it is a consistent estimator of the integrated 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.

Theorem 4.4 (**Consistency and uniform convergence**). *Suppose* $(X_t^{(T)} : t \leqslant T, T \in \mathbb{N})$ *satisfies Assumption 4.3 for* $\ell = 2$ *and consider the estimator* $\hat{\mathcal{F}}_\omega^{(T)}$ *in* (3.3) *with bandwidth fulfilling Assumption 3.1. Then,*

(i) $\mathbb{E}[|||\hat{\mathcal{F}}_\omega^{(T)} - G_\omega|||_2^2] = O((bT)^{-1} + b^4)$;

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

uniformly in $\omega \in [-\pi, \pi]$.

The proof of Theorem 4.4 is given in Section D.2 of the Appendix. Under the conditions of this theorem, the sample eigenelements $(\hat{\lambda}_l^\omega, \hat{\phi}_l^\omega : l \in \mathbb{N})$ of $\hat{\mathcal{F}}_\omega$ are consistent for the eigenelements $(\tilde{\lambda}_l^\omega, \tilde{\phi}_l^\omega : l \in \mathbb{N})$ of G_ω; see Mas & Menneteau (2003).

To obtain the distributional properties of $\hat{\gamma}^{(T)}$, it is necessary to replace the denominator with its deterministic limit. In analogy with Theorem 4.1, the theorem below gives conditions on the bandwidth for which this is justified under the alternative.

Theorem 4.5. *Let Assumption 4.3 be satisfied with $k = 8$ and $\ell = 2$ and assume that $\inf_\omega \langle G_\omega(\psi), \psi \rangle > 0$ for all $\psi \in H_{\mathbb{C}}$. Then,*

$$\sqrt{T} |\gamma_h^{(T)} - \hat{\gamma}_h^{(T)}| = O_p\left(\frac{1}{\sqrt{bT}} + b^2 + \frac{1}{b\sqrt{T}} \right) \qquad (T \to \infty).$$

The proof of Theorem 4.5 is given in Section S2 of the Online Supplement. The theorem shows that for $\hat{\gamma}_h^{(T)}$ to have the same asymptotic sampling properties as $\gamma_h^{(T)}$, additional requirements on the bandwidth b are needed. These are stated next.

Assumption 4.4. *The bandwidth b satisfies $b \to 0$ such that $b\sqrt{T} \to \infty$ as $T \to \infty$.*

The conditions imply that the bandwidth should tend to zero at a slower rate than in the stationary case. The conditions on the bandwidth imposed in this paper are in particular weaker than the ones required by Dwivedi & Subba Rao (2011) in the finite-dimensional context.

The dependence structure of $\gamma_h^{(T)}$ under the alternative is more involved than under the null of stationarity because the mean is nonzero for $h \neq 0 \mod T$. Additionally, the real and imaginary components of the covariance structure are correlated. The following theorem is for the test based on (3.2).

Theorem 4.6. *Let Assumption 4.3 be satisfied with $k = \{2, 4\}$. Then, for $h = 1, \ldots, T-1$,*

$$\mathbb{E}\big[\gamma_h^{(T)}(l)\big] = \frac{1}{2\pi} \int_0^{2\pi} \int_0^1 \frac{\langle \mathcal{F}_{u,\omega} \tilde{\phi}_l^{\omega+\omega_h}, \tilde{\phi}_l^\omega \rangle e^{-\mathrm{i}2\pi u h}}{\sqrt{\tilde{\lambda}_l^\omega \tilde{\lambda}_l^{\omega+\omega_h}}} du d\omega + O\left(\frac{1}{T}\right) = O\left(\frac{1}{h^2}\right) + O\left(\frac{1}{T}\right). \tag{4.7}$$

The covariance structure satisfies

1. $T\,\mathrm{Cov}(\Re\gamma_{h_1}^{(T)}(l_1), \Re\gamma_{h_2}^{(T)}(l_2)) =$
$$\frac{1}{4}\big[\Sigma_{h_1,h_2}^{(T)}(l_1,l_2) + \acute{\Sigma}_{h_1,h_2}^{(T)}(l_1,l_2) + \grave{\Sigma}_{h_1,h_2}^{(T)}(l_1,l_2) + \bar{\Sigma}_{h_1,h_2}^{(T)}(l_1,l_2)\big] + R_T,$$
2. $T\,\mathrm{Cov}(\Re\gamma_{h_1}^{(T)}(l_1,l_2), \Im\gamma_{h_2}^{(T)}(l_3,l_4)) =$
$$\frac{1}{4\mathrm{i}}\big[\Sigma_{h_1,h_2}^{(T)}(l_1,l_2) - \acute{\Sigma}_{h_1,h_2}^{(T)}(l_1,l_2) + \grave{\Sigma}_{h_1,h_2}^{(T)}(l_1,l_2) - \bar{\Sigma}_{h_1,h_2}^{(T)}(l_1,l_2)\big] + R_T,$$
3. $T\,\mathrm{Cov}(\Im\gamma_{h_1}^{(T)}(l_1,l_2), \Im\gamma_{h_2}^{(T)}(l_3,l_4)) =$
$$\frac{1}{4}\big[\Sigma_{h_1,h_2}^{(T)}(l_1,l_2) - \acute{\Sigma}_{h_1,h_2}^{(T)}(l_1,l_2) - \grave{\Sigma}_{h_1,h_2}^{(T)}(l_1,l_2) + \bar{\Sigma}_{h_1,h_2}^{(T)}(l_1,l_2)\big] + R_T,$$

where $\|R_T\|_2 = O(T^{-1})$ *and where* $\Sigma^{(T)}_{h_1,h_2}(l_1,l_2), \acute{\Sigma}^{(T)}_{h_1,h_2}(l_1,l_2)\grave{\Sigma}^{(T)}_{h_1,h_2}(l_1,l_2), \bar{\Sigma}^{(T)}_{h_1,h_2}(l_1,l_2)$ *are derived in the Appendix and general expressions for* (3.1) *are defined in equations (S.5.1)–(S.5.4) of the Online Supplement.*

The proof of Theorem 4.6 is given in Section D.2 of the Appendix. Its companion theorem is stated as Theorem C.3 in Section C. The last result in this section concerns the asymptotic of $\hat{Q}^{(T)}_M$ in (3.6) in the locally stationary setting. Before stating this result, observe that the previous theorem shows that a noncentrality parameter will have to enter, since the mean of $\gamma^{(T)}_h(l)$ is nonzero. Henceforth, the limit of (4.7) shall be denoted by

$$\mu_h(l) = \frac{1}{2\pi}\int_0^{2\pi}\int_0^1 \frac{\langle \mathcal{F}_{u,\omega}\tilde{\phi}^{\omega+\omega_h}_l, \tilde{\phi}^{\omega}_l\rangle e^{-\mathrm{i}2\pi uh}}{\sqrt{\tilde{\lambda}^{\omega}_l\tilde{\lambda}^{\omega+\omega_h}_l}}dud\omega$$

and its vectorization by $\boldsymbol{\mu}_h$. Following Paparoditis (2009) and Dwivedi & Subba Rao (2011), there is an intuitive interpretation of the degree of nonstationarity that can be detected in these functions. For fixed l and small h, they can be seen to approximate the Fourier coefficients of the function $\langle \mathcal{F}_{u,\omega}\tilde{\phi}^{\omega+\omega_h}_l, \tilde{\phi}^{\omega}_l\rangle/(\tilde{\lambda}^{\omega}_l\tilde{\lambda}^{\omega+\omega_h}_l)^{1/2}$. More specifically, for small h and $T\to\infty$, they approximate

$$\vartheta_{h,j}(l) = \frac{1}{2\pi}\int_0^{2\pi}\int_0^1 \frac{\langle \mathcal{F}_{u,\omega}\tilde{\phi}^{\omega+\omega_h}_l, \tilde{\phi}^{\omega}_l\rangle}{\sqrt{\tilde{\lambda}^{\omega}_l\tilde{\lambda}^{\omega+\omega_h}_l}} e^{\mathrm{i}2\pi uh-\mathrm{i}j\omega}dud\omega.$$

Thus, $\mu_h(l)\approx\vartheta_{h,0}(l)$. If the process is weakly stationary, then $\mathcal{F}_{u,\omega}\equiv\mathcal{F}_\omega$ and the eigenelements reduce to $\lambda^{\omega}_l, \lambda^{\omega+h}_l$ and $\phi^{\omega}_l, \phi^{\omega+h}_l$, respectively, and hence the integrand of the coefficients does not depend on u. All Fourier coefficients are zero except $\vartheta_{0,j}(l)$. In particular, $\vartheta_{0,0}(l) = 1$. Observe that, for the statistic (3.1), the above becomes $\vartheta_{h,j}(l,l')$ and, for $h=0$, the off-diagonal elements yield coherence measures. The mean functions can now be seen to reveal long-term non-stationary behavior. Unlike testing methods based on segments in the time domain, the proposed method is consequently able to detect smoothly changing behavior in the temporal dependence structure. A precise formulation of the asymptotic properties of the test statistic in (3.2) under H_A is given in the next theorem.

Theorem 4.7. *Let Assumptions 3.1 and 4.4 be satisfied. Let Assumption 4.3 be satisfied* $k\geqslant 1$ *and* $\ell = 2$ *and assume that* $\inf_\omega \lambda^{\omega}_L > 0$. *Then,*

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

$$\sqrt{T}\boldsymbol{b}^{(T)}_M \xrightarrow{\mathcal{D}} \mathcal{N}_{2M}(\boldsymbol{\mu},\Sigma_A) \qquad (T\to\infty),$$

where $\mathcal{N}_{2M}(\boldsymbol{\mu},\Sigma_A)$ *denotes a* $2M$*-dimensional normal distribution with mean vector* $\boldsymbol{\mu}$*, whose first* M *components are* $\Re\boldsymbol{e}^T\boldsymbol{\mu}_{h_m}$ *and last* M *components are* $\Im\boldsymbol{e}^T\boldsymbol{\mu}_{h_m}$*, and block covariance matrix*

$$\Sigma_A = \begin{pmatrix}\Sigma^{(11)}_A & \Sigma^{(12)}_A\\ \Sigma^{(21)}_A & \Sigma^{(22)}_A.\end{pmatrix}$$

whose $M\times M$ *blocks are, for* $m,m'=1,\ldots,M$*, given by*

$$\Sigma^{(11)}_A(m,m') = \lim_{T\to\infty}\sum_{l_1,l_2=1}^{L}\frac{1}{4}\big[\Sigma^{(T)}_{h_1,h_2}(l_1,l_2)+\acute{\Sigma}^{(T)}_{h_1,h_2}(l_1,l_2)+\grave{\Sigma}^{(T)}_{h_1,h_2}(l_1,l_2)+\bar{\Sigma}^{(T)}_{h_1,h_2}(l_1,l_2)\big],$$

$$\Sigma_A^{(12)}(m,m') = \lim_{T\to\infty} \sum_{l_1,l_2=1}^{L} \frac{1}{4\mathrm{i}}\big[\Sigma_{h_1,h_2}^{(T)}(l_1,l_2) - \acute{\Sigma}_{h_1,h_2}^{(T)}(l_1,l_2) + \grave{\Sigma}_{h_1,h_2}^{(T)}(l_1,l_2) - \bar{\Sigma}_{h_1,h_2}^{(T)}(l_1,l_2)\big],$$

$$\Sigma_A^{(22)}(m,m') = \lim_{T\to\infty} \sum_{l_1,l_2=1}^{L} \frac{1}{4}\big[\Sigma_{h_1,h_2}^{(T)}(l_1,l_2) - \acute{\Sigma}_{h_1,h_2}^{(T)}(l_1,l_2) - \grave{\Sigma}_{h_1,h_2}^{(T)}(l_1,l_2) + \bar{\Sigma}_{h_1,h_2}^{(T)}(l_1,l_2)\big],$$

where $\Sigma_{h_m,h_{m'}}^{(T)}(l_1,l_2), \acute{\Sigma}_{h_m,h_{m'}}^{(T)}(l_1,l_2), \grave{\Sigma}_{h_m,h_{m'}}^{(T)}(l_1,l_2), \bar{\Sigma}_{h_m,h_{m'}}^{(T)}(l_1,l_2)$ *are as in Theorem 4.6.*

(b) Using the result in (a), it follows that

$$\hat{Q}_M^{(T)} \xrightarrow{\mathcal{D}} \chi^2_{\mu,2M}, \qquad (T\to\infty),$$

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

The proof of Theorem 4.7 can be found in Appendix D.2. Theorem C.4 contains the result for (3.1).

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 further in future research.

5.1 Simulation setting

To generate functional time series, the general strategy applied, for example in the papers by Aue et al. (2015) and Hörmann et al. (2015), is utilized. For this simulation study, all processes are build 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 $(X_t\colon t\in\mathbb{Z})$ satisfies

$$\begin{aligned}\langle X_t,\psi_l\rangle &= \sum_{l'=1}^{\infty}\sum_{t'=1}^{p}\langle X_{t-t'},\psi_l\rangle\langle A_{t'}(\psi_l),\psi_{l'}\rangle + \langle\varepsilon_t,\psi_l\rangle \\ &\approx \sum_{l'=1}^{L_{\max}}\sum_{t'=1}^{p}\langle X_{t-t'},\psi_l\rangle\langle A_{t'}(\psi_l),\psi_{l'}\rangle + \langle\varepsilon_t,\psi_l\rangle,\end{aligned} \tag{5.1}$$

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 $\mathcal{N}(0,\nu_{l,l'}^{(t')})$ random variables, with the specifications of $\nu_{l,l'}$ given below.

To ensure stationarity or the existence of a causal solution (see Bosq, 2000; van Delft & Eichler, 2016, for the stationary and locally stationary case, respectively), the norms $\kappa_{t'}$ of $\boldsymbol{A}_{t'}$ are required to satisfy certain conditions, for example, $\sum_{t'=1}^{p} \|\boldsymbol{A}_{t'}\|_\infty < 1$. The functional white noise, FWN, process is included in (5.1) setting $p = 0$.

All simulation experiments were implemented by means of the `fda` package 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 (3.6) depends on the tuning parameters L, determining the dimension of the projection spaces, and M, the number of lags to be included in the procedure. In the following a criterion will be set up to choose L, while for M only two values were entertained because the selection is less critical for the performance as long as it is not chosen too large. Note that an optimal choice for L is more cumbersome to obtain in the frequency domain than in the time domain because a compromise over a number of frequencies is to be made, each of which may individually lead to a different optimal L than the one globally selected. To begin with, consider the eigenbased version of the test. The theory provided in the previous two sections indicates that including too small eigenvalues into the procedure would cause instability. On the other hand, the number of eigenvalues used should explain a reasonable proportion of the total variation in the data. These two requirements can be conflicting and a compromise is sought implementing a thresholding approach through the criterion

$$L = L_M(\zeta_1, \zeta_2, \xi) = \max\left\{\ell \colon \zeta_1 < \frac{1}{T}\sum_{j=1}^{T} \frac{\sum_{l=1}^{\ell} \lambda_l^{\omega_j}}{\sum_{l'=1}^{L_{\max}} \lambda_{l'}^{\omega_j}} < \zeta_2 \text{ and } \frac{\inf_j \lambda_\ell^{\omega_j}}{\inf_j \lambda_1^{\omega_j}} > \xi\right\} - \lfloor \log(M) \rfloor, \tag{5.2}$$

where ω_j denotes the j-th Fourier frequency and $\lfloor\cdot\rfloor$ integer part.[1] Extensive simulation studies have shown that the choices $\zeta_1 = 0.70$, $\zeta_2 = 0.90$ and $\xi = 0.15$ work well when the eigenvalues display a moderate decay. Then, the performance of the procedure appears robust against moderate deviations from these presets.

For processes with a very steep decay of the eigenvalues, however, this rule might pick L on the conservative side. To take this into account, the criterion is prefaced with a preliminary check for fast decay that leads to a relaxation of the parameter values ζ_2 and ξ if satisfied. The condition to be checked is

$$C(\xi_1, \xi_2, \xi_3) = \left\{\frac{\inf_j \lambda_2^{\omega_j}}{\inf_j \lambda_1^{\omega_j}} < \xi_1 \text{ and } \frac{\inf_j \lambda_3^{\omega_j}}{\inf_j \lambda_1^{\omega_j}} < \xi_2 \text{ and } \frac{\inf_j \lambda_4^{\omega_j}}{\inf_j \lambda_1^{\omega_j}} < \xi_3\right\}, \tag{5.3}$$

with $\xi_1 = 0.5$, $\xi_2 = 0.25$ and $\xi_3 = 0.125$. Observe that (5.3) provides a quantification of what is termed a 'fast decay' in this paper. If it is satisfied, L is chosen using the criterion (5.2) but now with $\zeta_2 = 0.995$ and $\xi = 0.01$. Condition (5.3) is in particular useful for processes for which the relative decay of the eigenvalues tends to be quadratic, a behavior more commonly observed for nonstationary functional time series. A further automation of the criterion is interesting and may be considered in future research as it is beyond the scope

[1] The rule is $\max\{1, L\}$ in the exceptional cases that (5.2) returns a nonpositive value.

of the present paper. The test statistic $\hat{Q}_M^{(T)}$ in (3.6) is then set up with the above choice of L and with $h_m = m$ and $M = 1$ and 5. A rejection is reported if the simulated test statistic value exceeds the critical level prescribed in part (b) of Theorem 4.3.

Correspondingly, for the fixed projection basis, the test statistic $\hat{Q}_M^{(T)}$ in (3.6) is set up to select those directions $\ell \in \{1, \ldots, L_{\max}\}$ for which $\langle \mathcal{F}_\omega \psi_\ell, \psi_\ell \rangle$ are largest and which explain at least 90 percent of total variation (averaged over frequencies). Furthermore, $h_m = m$ for $m = 1, \ldots, M$ with $M = 1$ and 5. A rejection is reported as above using the critical level prescribed in part (b) of Theorem C.2.

The performance of both tests is evaluated below in a variety of settings. To distinguish between the two approaches, refer to the eigenbased statistic as $\hat{Q}_{M,e}^{(T)}$ and to the fixed projection statistic as $\hat{Q}_{M,f}^{(T)}$. Estimation of the spectral density operator and its eigenelements, needed to compute the two statistics, was achieved using (3.3) with the concave smoothing kernel $K(x) = 6(0.25 - x^2)$ with compact support on $x \in [-1/2, 1/2]$ and bandwidth $b = T^{-1/5}$.

5.3 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 but nonfunctional contexts. Among these contributions are Paparoditis (2009) and Jin et al. (2015), who used the spectral density of the squares, Nason (2013), who worked with locally stationary wavelets, Dwivedi & Subba Rao (2011), who focused on Gaussianity of the observations, and Jentsch & Subba Rao (2015), who employed a stationary bootstrap procedure. A different idea was put forward by Bandyopadhyay & Subba Rao (2017) and Bandyopadhyay et al. (2017). These authors utilized the notion of orthogonal samples to estimate the variance, falling back on a general estimation strategy developed in Subba Rao (2016).

To estimate the tri-spectrum for the eigenbased statistics, the estimator proposed in Brillinger & Rosenblatt (1967) was adopted to the functional context in the following way. Note first that in order to utilize the results of Theorem 4.3, the limiting covariance structure given through discretized population terms of the form

$$\sum_{j_1, j_2 = 1}^{T} \frac{(2\pi)}{T^2} \langle \mathcal{F}_{\omega_{j_1}, -\omega_{j_1+h_1}, -\omega_{j_2}} (\phi_{l_2}^{\omega_{j_2}} \otimes \phi_{l_2}^{\omega_{j_2+h_1}}), \phi_{l_1}^{\omega_{j_1}} \otimes \phi_{l_1}^{\omega_{j_1+h_1}} \rangle + O\bigg(\frac{1}{T^2}\bigg)$$

has to be estimated. To do this, consider the raw estimator

$$\begin{aligned} &I_{\alpha_1,\alpha_2,\alpha_3,\alpha_4}^{(T)}(\hat{\phi}_{l_1}^{\omega_{j_1}}, \hat{\phi}_{l_1}^{\omega_{j_2}}, \hat{\phi}_{l_2}^{\omega_{j_3}}, \hat{\phi}_{l_2}^{\omega_{j_4}}) \\ &\qquad = \frac{1}{(2\pi)^3 T} \sum_{t_1,t_2,t_3,t_4=1}^{T} \langle \mathcal{C}_{t_1,t_2,t_3,t_4} (\hat{\phi}_{l_2}^{\omega_{j_3}} \otimes \hat{\phi}_{l_2}^{\omega_{j_4}}), \hat{\phi}_{l_1}^{\omega_{j_1}} \otimes \hat{\phi}_{l_1}^{\omega_{j_2}} \rangle e^{-\mathrm{i}\sum_{j=1}^{4} \alpha_j}, \end{aligned}$$

and observe that this is the fourth-order periodogram estimator of the random variables $\langle X_t, \hat{\phi}_l^{\omega_j} \rangle$. This estimator is to be evaluated for a combination of frequencies $\alpha_1, \ldots, \alpha_4$ that lie on the principal manifold but not in any proper submanifold (see below) and is an unbiased but inconsistent estimator under the null hypothesis. The verification of this claim is given in the Online Supplement. To construct a consistent version from the raw estimate, smooth over frequencies. The elements of $\hat{\Sigma}_m$ can then be obtained from the more standard second-order estimators above and from fourth-order estimators of the form

$$\frac{(2\pi)^3}{(b_4 T)^3} \sum_{k_1,k_2,k_3,k_4} K_4\Big(\frac{\omega_{j_1}-\alpha_{k_1}}{b_4},\ldots,\frac{\omega_{j_4}-\alpha_{k_4}}{b_4}\Big)\Phi(\alpha_{k_1},\ldots,\alpha_{k_4}) I^{(T)}_{\alpha_{k_1},\ldots,\alpha_{k_4}}(l_1,l_2), \tag{5.4}$$

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 $\sum_{j=1}^4 \alpha_j \equiv 0 \mod 2\pi$ such that $\sum_{j\in J}\alpha_j \not\equiv 0 \mod 2\pi$ where J is any non-empty subset of $\{1,2,3,4\}$ and $\Phi(\alpha_1,\alpha_2,\alpha_3,\alpha_4) = 0$ otherwise. The function Φ therefore controls the selection of frequencies in the estimation, ensuring that no combinations that lie on a proper submanifold are chosen. This is important because, for $k > 2$, the expectation of k-th order periodograms at such submanifolds possibly diverges. This was pointed out in Brillinger & Rosenblatt (1967, Page 163). The estimator in (5.4) is consistent under the null if the bandwidth b_4 satisfies $b_4 \to 0$ but $b_4^{-3}T \to \infty$ as $T \to \infty$. As this estimator is required for various combinations of l_1 and l_2, it is worthwhile to mention that computational complexity increases rapidly. The implementation was therefore partially done with the compiler language `C++` and the `Rcpp`-package in `R`.

The simulations and application to follow below were conducted with $K_4(x_1,\ldots,x_4) = \prod_{j=1}^4 K(x_j)$, where K is as in Section 5.2 and bandwidth $b_4 = T^{-1/6}$. It should be noted that the outcomes were not sensitive with respect to the bandwidth choices.

5.4 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 $\mathrm{Var}(\langle\varepsilon_t,\psi_l\rangle) = \exp((l-1)/10)$;

(b) The FAR(2) variables $X_1,\ldots,X_T$ with operators specified through the respective variances $\nu^{(1)}_{l,l'} = \exp(-l-l')$ and $\nu^{(2)}_{l,l'} = 1/(l+l'^{3/2})$ and Frobenius 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 Frobenius 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 second-order autoregressive dynamics of different persistence, with the process in (c) possessing the stronger temporal dependence.

The empirical rejection levels for the processes (a)–(c) can be found in Table 5.1. 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 rejection levels for the eigenbased statistics tend to be a little conservative for models B and C.

	T	$\hat{Q}^{(T)}_{1,e}$	% level 5	1	$\hat{Q}^{(T)}_{5,e}$	% level 5	1	$\hat{Q}^{(T)}_{1,f}$	% level 5	1	$\hat{Q}^{(T)}_{5,f}$	% level 5	1
(a)	64	1.32	4.00	0.90	8.17	3.40	0.30	1.31	3.60	0.60	8.65	3.90	1.20
	128	1.40	5.20	0.60	8.98	5.00	0.90	1.38	5.00	0.80	9.07	5.10	1.40
	256	1.53	5.40	1.30	8.89	4.50	0.80	1.31	5.60	0.80	9.23	5.30	1.00
	512	1.49	5.50	1.30	9.35	5.50	1.00	1.38	5.20	1.00	9.34	4.90	0.90
	1024	1.44	5.70	1.10	9.17	4.00	0.70	1.36	5.00	1.20	9.51	4.60	0.70
(b)	64	1.23	2.70	0.10	7.94	2.80	0.80	1.41	2.90	0.60	8.59	4.80	1.60
	128	1.40	5.20	0.80	8.80	3.60	0.70	1.41	5.40	0.90	8.91	5.10	1.20
	256	1.41	5.70	1.70	8.89	2.90	0.40	1.34	4.60	0.90	9.21	5.20	1.00
	512	1.29	6.10	1.00	9.11	5.30	0.90	1.34	5.40	1.10	9.36	4.30	0.50
	1024	1.42	4.70	2.00	8.91	3.70	0.70	1.44	4.40	0.90	9.46	4.20	0.70
(c)	64	1.21	3.90	0.80	8.25	2.00	0.30	1.36	3.50	0.60	8.62	3.90	1.20
	128	1.39	5.80	1.40	8.86	5.00	1.30	1.39	5.90	0.90	8.81	4.00	1.20
	256	1.42	4.50	1.50	8.90	3.00	0.50	1.38	4.50	0.90	9.37	4.30	1.20
	512	1.49	5.40	1.10	9.26	3.70	0.60	1.39	6.50	1.10	9.36	5.00	0.70
	1024	1.39	4.50	0.90	9.29	3.90	0.50	1.34	5.90	1.00	9.66	5.70	0.80

Table 5.1: Median of test statistic values and rejection rates of $\hat{Q}^{(T)}_{M,e}$ and $\hat{Q}^{(T)}_{M,f}$ 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	(a)		(b)		(c)		(d)		(e)		(f)	
64	8.20	14.00	8.17	14.01	8.20	14.01	7.99	14.01	5.09	12.45	7.84	14.00
	0.88	0.93	0.88	0.92	0.88	0.92	0.88	0.92	0.93	0.93	0.88	0.93
128	9.89	14.00	9.77	14.01	9.87	14.00	8.95	14.01	5.94	12.35	9.19	14.00
	0.90	0.93	0.89	0.92	0.89	0.92	0.89	0.92	0.94	0.93	0.88	0.93
256	10.52	14.00	10.06	14.01	10.32	14.00	8.88	14.02	6.46	12.48	10.04	14.00
	0.89	0.93	0.88	0.92	0.88	0.92	0.89	0.92	0.93	0.93	0.88	0.93
512	11.00	14.00	11.00	14.02	11.00	14.00	10.00	14.01	6.65	12.37	11.00	14.00
	0.89	0.94	0.89	0.92	0.89	0.93	0.89	0.92	0.94	0.93	0.89	0.93
1024	11.00	14.00	11.00	14.03	11.00	14.00	11.00	14.02	6.44	12.33	11.00	14.00
	0.88	0.94	0.89	0.92	0.88	0.93	0.89	0.92	0.94	0.94	0.88	0.94

Table 5.2: Average choices of L (top entries in each row specified by T) and aTVE (bottom entries) for $M = 1$, various sample sizes and processes (a)–(f). For each process, the left column is for the eigenbased statistics, the right column for the fixed projection statistics.

For the eigenbased statistics set up with $M = 1$, Table 5.2 displays the average choices of L according to (5.2) and the corresponding average total variation explained (aTVE) $T^{-1}\sum_{j=1}^{T}\big(\sum_{l=1}^{L}\lambda_l^{\omega_j}/\sum_{l'=1}^{L_{\max}}\lambda_{l'}^{\omega_j}\big)$. It can be seen that larger sample sizes lead to the selection of larger L, as more degrees of freedom become available. Moreover, aTVE tends to be close to 0.90 for the models under consideration. Some evidence on closeness between empirical and limit densities for the statistics $\hat{Q}^{(T)}_{5,e}$ and $\hat{Q}^{(T)}_{5,f}$ are provided in Figure 5.1. Naturally, the numbers for the fixed projection based statistics are more homogeneous.

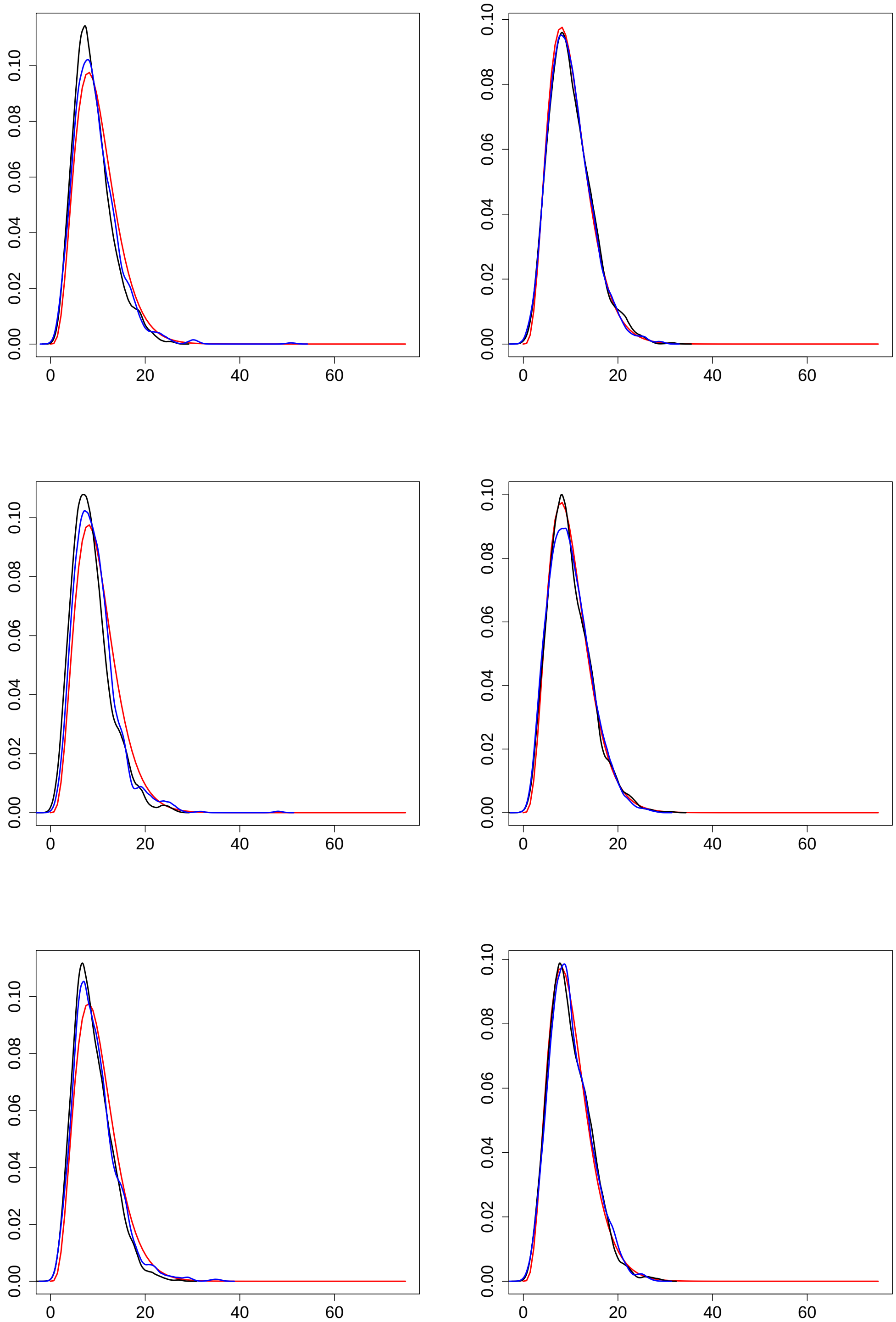

Figure 5.1: black: Empirical density of $\hat{Q}_{5,e}^{(T)}$ (black) and $\hat{Q}_{5,f}^{(T)}$ (blue) for $T = 64$ (left panel) and $T = 512$ (right panel) for DGPs (a)–(c) (top to bottom). Red: The corresponding chi-squared densities predicted under the null.

5.5 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 $\nu^{(1)}_{l,l'} = \exp(-l - l')$ and Frobenius norm $\kappa_1 = 0.8$, and innovations given by (a) with added multiplicative time-varying variance
$$\sigma^2(t) = \frac{1}{2} + \cos\left(\frac{2\pi t}{1024}\right) + 0.3 \sin\left(\frac{2\pi t}{1024}\right);$$

(e) The tvFAR(2) variables $X_1, \ldots, X_T$ with both operators as in (d) but with time-varying Frobenius norm
$$\kappa_{1,t} = 1.8 \cos\left(1.5 - \cos\left(\frac{4\pi t}{T}\right)\right),$$
constant Frobenius 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 $t \leqslant 3T/8$, the operators are as in (b) but with Frobenius 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 Frobenius 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 5.4. 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 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 5.3. Power results are roughly similar across the selected values of M for both statistics. For DGP (d), power is low for the small sample sizes $T = 64$ for the eigenbased statistics and even lower for the fixed projection statistics. It is at 100% for all T larger or equal to 256. The low power is explained by the form of the time-varying variance which takes 1024 observations to complete a full cycle of the sine and cosine components. In the situation of the smallest sample size, this slowly varying variance appears more stationary, explaining why rejections of the null are less common. Generally, the eigenbased statistics performs better than its fixed projection counterpart for this process.

DGP (e) shows lower power throughout. The reason for this is that the form of local stationarity under consideration here is more difficult to separate from stationary behavior expected under the null hypothesis.

	T	$\hat{Q}_{1,e}^{(T)}$	% level 5	% level 1	$\hat{Q}_{5,e}^{(T)}$	% level 5	% level 1	$\hat{Q}_{1,f}^{(T)}$	% level 5	% level 1	$\hat{Q}_{5,f}^{(T)}$	% level 5	% level 1
(d)	64	6.39	54.30	30.00	14.33	29.40	12.30	2.68	17.20	4.00	11.08	13.20	4.80
	128	66.97	99.90	99.80	86.94	99.80	99.40	18.94	98.80	93.30	35.79	98.10	90.20
	256	611.13	100.00	100.00	747.09	100.00	100.00	130.23	100.00	100.00	206.05	100.00	100.00
	512	1462.97	100.00	100.00	1506.20	100.00	100.00	269.31	100.00	100.00	347.91	100.00	100.00
	1024	1550.04	100.00	100.00	1627.24	100.00	100.00	195.85	100.00	100.00	251.64	100.00	100.00
(e)	64	5.29	48.90	46.30	15.78	47.20	45.40	3.93	44.20	39.40	14.74	44.40	40.50
	128	6.05	50.20	46.30	20.74	52.00	48.40	4.20	45.30	42.10	15.47	46.50	43.20
	256	5.67	49.50	45.80	19.71	52.80	47.90	4.04	43.70	40.40	15.67	46.10	43.10
	512	6.35	50.65	47.55	25.80	56.96	51.45	3.91	44.40	41.80	17.37	48.50	44.30
	1024	9.80	52.06	50.35	42.40	65.83	59.50	4.64	46.80	43.80	22.10	54.80	49.00
(f)	64	23.84	97.00	93.10	30.79	89.70	79.60	6.19	52.50	23.50	15.81	33.60	14.60
	128	30.37	97.00	93.20	40.66	90.70	83.10	11.97	90.30	69.90	23.45	80.70	51.60
	256	80.27	100.00	100.00	87.25	100.00	100.00	23.10	99.90	98.60	38.49	99.80	98.10
	512	288.79	100.00	100.00	298.20	100.00	100.00	46.99	100.00	100.00	68.60	100.00	100.00
	1024	781.13	100.00	100.00	824.65	100.00	100.00	94.85	100.00	100.00	127.87	100.00	100.00

Table 5.3: Median of test statistic values and rejection rates of $\hat{Q}_{M,e}^{(T)}$ and $\hat{Q}_{M,f}^{(T)}$ 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.

This is corroborated in Figure 5.2, where the center of the empirical version of the non-central generalized chi-squared limit under the alternative is more closely aligned with that of the standard chi-squared limit expected under the null hypothesis than for the cases (d) and (f). Both test statistics behave similarly across the board.

The results for DGP (f) indicate that the proposed statistics have power against structural break alternatives. This statement is supported by Figure 5.2. Here the eigenbased version picks up power quicker than the fixed projection counterpart as the sample size T increases. All statistics work well once T reaches 256.

Going back to Table 5.2, it further seen that L is under H_A generally chosen in a similar way as under H_0, with one notable exception being the process under (e), where L tends to be significantly smaller. Under the alternative, aTVE tends to be marginally larger for models (d)–(f) than for the models (a)–(c) considered under the null hypothesis.

5.6 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_{10}-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. The sample sizes considered were $T = 64$ and $T = 128$, since these are most

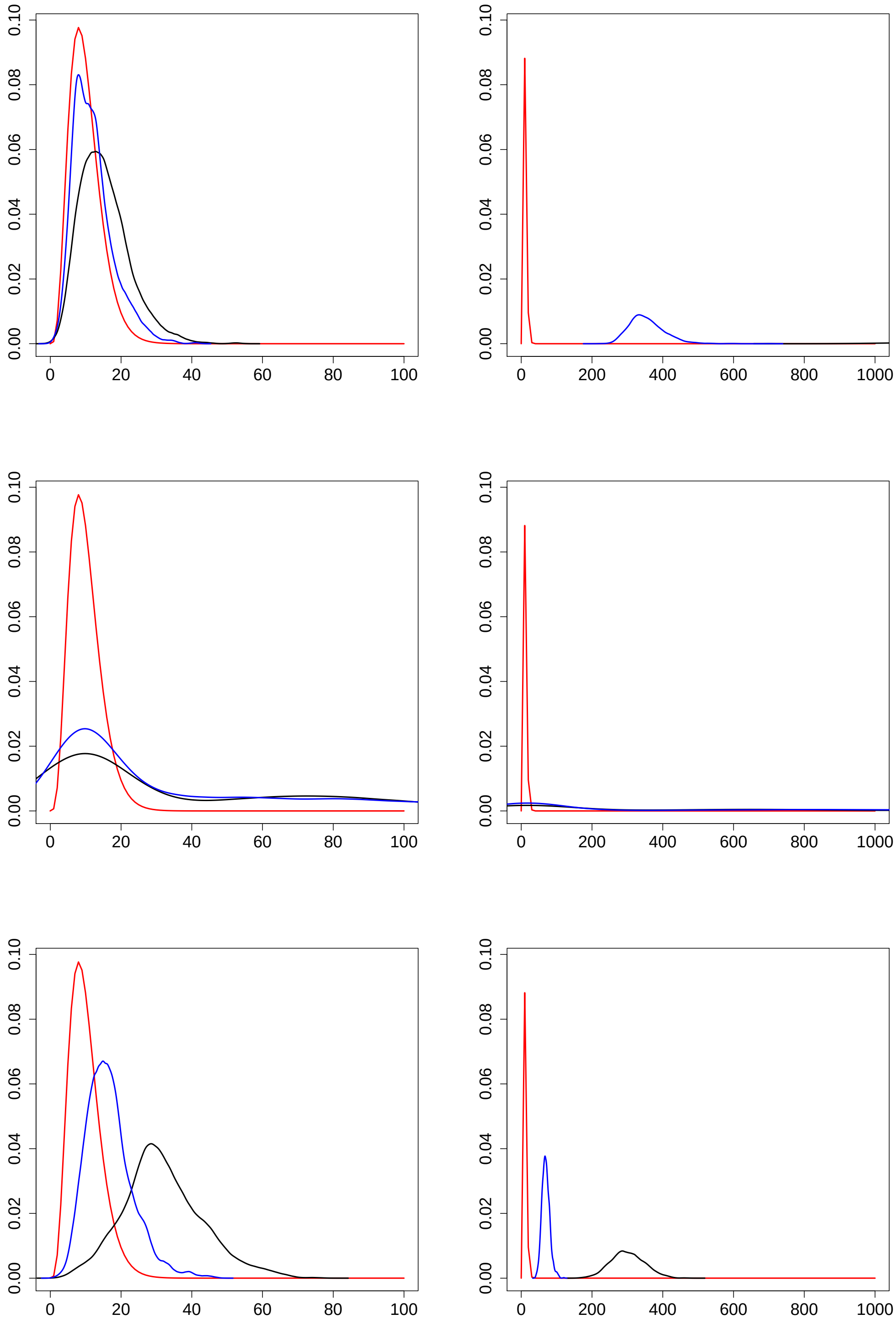

Figure 5.2: Empirical density of $\hat{Q}^{(T)}_{5,e}$ (black) and $\hat{Q}^{(T)}_{5,f}$ (blue) for $T = 64$ (left panel) and $T = 512$ (right panel) for DGPs (d)–(f) (top to bottom). Note that 2.5% outliers have been removed for Model (e). Red: The corresponding chi-squared densities predicted under the null.

similar to the ones observed for the temperature data discussed in Section 5.7, and $M = 1, 3$ and 5. All other aspects are as detailed in Section 5.4. These 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 5.4 and can be readily compared with corresponding outcomes for the Gaussian processes (b) and (d) in Tables 5.1–.5.3.

						% level			% level	
		T	L	aTVE	$\hat{Q}^{(T)}_{1,e}$	5	1	$\hat{Q}^{(T)}_{5,e}$	5	1
(g)	t	64	8.13	0.88	1.33	3.30	0.30	8.24	2.40	0.70
		128	9.73	0.89	1.22	4.90	1.40	8.80	3.90	1.00
	β	64	8.18	0.88	1.16	2.90	0.60	7.61	2.40	0.40
		128	9.80	0.89	1.34	4.60	0.90	8.97	4.00	0.60
(h)	t	64	4.32	0.98	7.44	59.70	40.10	16.55	42.80	26.80
		128	4.57	0.98	85.85	100.00	100.00	98.84	99.80	99.50
	β	64	4.42	0.98	7.07	58.60	37.80	16.14	39.30	22.50
		128	4.65	0.98	88.78	100.00	100.00	98.57	99.90	99.80

Table 5.4: Median of test statistic values and rejection rates of $\hat{Q}^{(T)}_{M,e}$ 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. The values for L and aTVE correspond to $M = 1$. All table entries are generated from 1000 repetitions.

It can be seen from the results in Table 5.4 that the proposed procedures perform roughly as expected. First, under the null hypothesis for process (g) and for $T = 64$, levels tend to be conservative for all M. For $T = 128$, levels are well adjusted for $M = 1$ and still decent for $M = 5$. Both t_{19} and $\beta(6, 6)$ variables tend to be produce levels that differ only little from their Gaussian counterparts in Table 5.1. Second, under the alternative for process (h), powers align roughly as for the Gaussian case in Table 5.3. Comparing to Table 5.2, it can be seen that for process (h), the chosen L are lower than in the Gaussian case. Overall, the simulation results reveal that the estimation of the fourth-order spectrum does not lead to a marked decay in performance.

5.7 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 5.5, and the annual temperature profiles recorded at the Gayndah station are displayed for illustration in the left panel of Figure 5.3. 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. Focusing here on the eigenbased version, the test statistic $\hat{Q}_{M,e}^{(T)}$ in (3.6) was applied with L chosen according to (5.2) and $h_m = m$ for $m = 1, \ldots, M$, where $M = 1$ and 5.

The testing results are summarized in Table 5.5. It can be seen that stationarity is rejected in favor of the alternative at the 1% significance level at all measuring stations for $\hat{Q}_{5,e}^{(T)}$ and for all but two measuring stations for $\hat{Q}_{1,e}^{(T)}$, the exceptions being Melbourne and Sydney, where the test is rejected at the 5% level with a p-value of 0.014, and Sydney, where the p-value is approximately 0.072. The clearest rejection of the null hypothesis was found for the measuring station at Gunnedah Pool for both choices of M. The values of L chosen with (5.2) range from 4 to 8 and are similar to the values observed in the simulation study. The right-hand side of Figure 5.3 shows the decay of eigenvalues associated with the different measuring stations. Moreover, for $M = 1$ ($M = 5$) aTVE ranges from 0.762 (0.705) in Hobart to 0.882 (0.847) in Melbourne.

Station	T	L	aTVE	$\hat{Q}_{1,e}^{(T)}$	L	aTVE	$\hat{Q}_{5,e}^{(T)}$
Boulia	120	7	0.864	11.623	6	0.823	65.463
Robe	129	8	0.876	13.262	7	0.837	39.339
Cape Otway	149	5	0.844	35.382	4	0.798	75.876
Gayndah	117	6	0.840	14.638	5	0.792	55.967
Gunnedah	133	4	0.825	41.494	3	0.760	76.805
Hobart	121	5	0.762	20.405	4	0.705	38.288
Melbourne	158	8	0.882	5.274	7	0.847	54.879
Sydney	154	8	0.873	8.591	7	0.837	28.359

Table 5.5: Summary of results for eight Australian measuring stations. The column labeled T reports the sample size, L gives the value chosen by (5.2), aTVE is the average total variation explained as used in this criterion.

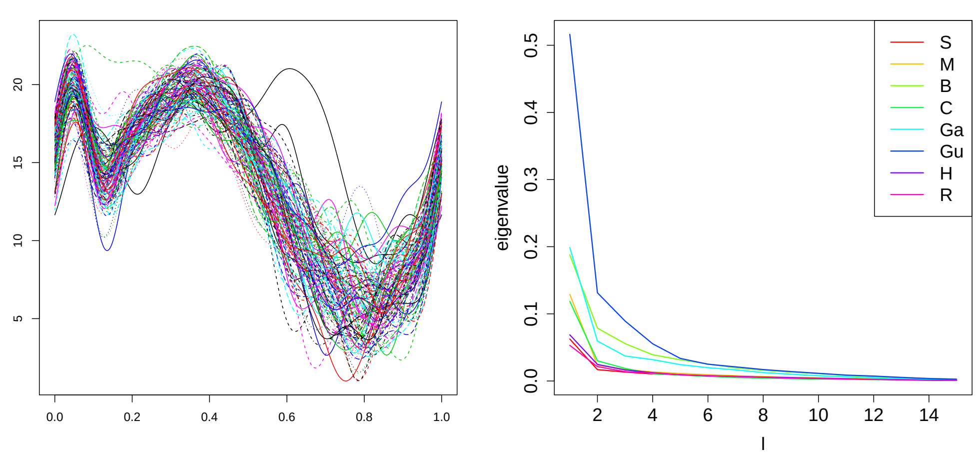

Figure 5.3: Plot of annual temperature curves at Gayndah station (left) and of eigenvalue decay across different measuring stations (right).

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, where deviations from stationarity were detected.

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 $m = 1, \ldots, M$. This can be approached through a more refined analysis of the size of the various $\hat{\boldsymbol{\beta}}_{h_m}^{(T)}$ in (3.5) whose real and imaginary part make up the vector $\hat{\mathbf{b}}_M^{(T)}$ in the test statistics $\hat{Q}_M^{(T)}$.

Another promising route of research is as follows. Figure 6.1 provides contour plots of squared modulus of $\hat{\gamma}_1^{(T)}$ for model (b) under the null and for models (d)–(f) under the alternative. The contours are obtained from averaging (over the simulation runs) the aggregated contributions $\hat{\gamma}_1^{(T)}$ and projecting these onto a Fourier basis of dimension 15. It can be seen that the magnitude of the contours provides another indicator for how easy or hard it may be to reject the null hypothesis. The top row in the figure is for the stationary DGP (b). For any of the sample sizes considered, the magnitude across $[0, 1]^2$ remains small, as expected under the null. The behavior under the alternative is markedly different, but the specifics depend on the type of alternative. For the time-varying noise process (d), the contribution of non-stationarity is at the diagonal, with the magnitude along this ridge depending on the sample size. For DGP (e), the form of non-stationarity creates very different contours. The structural break process (f) induces non-stationarity in the contours in a similar way as DGP (d), with most concentration occurring at the diagonal for all sample sizes. Any future refinement of the tests will have to take these features into account.

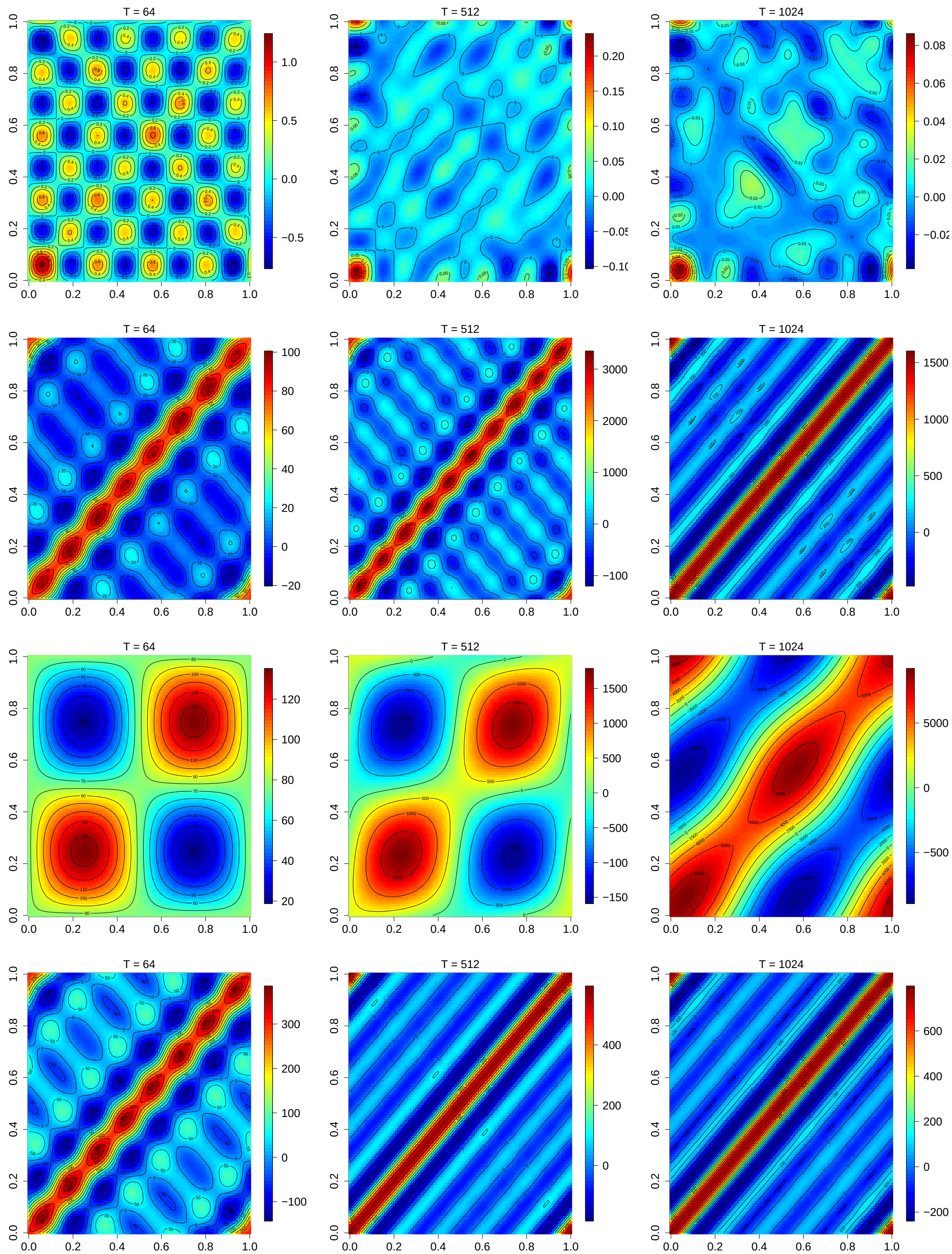

Figure 6.1: Contour plots of $\hat{\gamma}_1^{(T)}$ for $L^\star$ for various sample sizes T and DGPs (b) and (d)–(f) (top to bottom).

A A functional Cramér representation

Proof of Proposition 2.1. Let $(X_t \colon t \in \mathbb{Z})$ be a zero-mean, $H_{\mathbb{R}}$-valued stochastic process that is weakly stationary. It has been shown (Panaretos & Tavakoli, 2013) that for processes with $\sum_{h\in\mathbb{Z}} |||\mathcal{C}_h|||_1 < \infty$, there exists an isomorphic mapping between the subspaces $\overline{\text{sp}}(X_t \colon t \in \mathbb{Z})$ of $\mathbb{H}$ and $\overline{\text{sp}}(e^{\mathrm{i}t\cdot} \colon t \in \mathbb{Z})$ of $L^2([-\pi,\pi], \mathscr{B}, |||\mathcal{F}_\omega|||_1 d\omega)$. As a consequence, X admits the representation

$$X_t = \int_{-\pi}^{\pi} e^{\mathrm{i}t\lambda} dZ_\lambda \qquad \text{a.s. a.e.,} \tag{A.1}$$

where $(Z_\lambda \colon \lambda \in (-\pi,\pi])$ is a right-continuous, functional orthogonal-increment process. Conversely, for any process with the above representation, orthogonality of the increments and Theorem S2.2 of van Delft & Eichler (2016) imply

$$\begin{aligned}\text{Cov}(X_t, X_s) &= \mathbb{E}\Big[\int_{-\pi}^{\pi} e^{\mathrm{i}t\lambda_1} dZ_{\lambda_1} \otimes \int_{-\pi}^{\pi} e^{\mathrm{i}s\lambda_2} dZ_{\lambda_2}\Big] \\ &= \int_{-\pi}^{\pi} e^{\mathrm{i}(t-s)\lambda} \mathcal{F}_\lambda d\lambda = \mathcal{C}_{t-s},\end{aligned}$$

showing that a process that admits representation (A.1) must be weakly stationary. □

B Properties of functional cumulants

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

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

where the elementary tensors $(\psi_{l_1} \otimes \cdots \otimes \psi_{l_k} \colon l_1, \dots, l_k \in \mathbb{N})$ form an orthonormal basis in the tensor product space $\bigotimes_{j=1}^{k} H$. The latter follows since $(\psi_l \colon l \in \mathbb{N})$ is an orthornormal basis of the separable Hilbert space H. Similarly, define the *k-th order cumulant tensor* by

$$\text{cum}\big(X_1, \dots, X_k\big) = \sum_{l_1,\dots l_k \in \mathbb{N}} \text{cum}\Big(\prod_{t=1}^{k} \langle X_t, \psi_{l_t}\rangle\Big)(\psi_{l_1} \otimes \cdots \otimes \psi_{l_k}), \tag{B.1}$$

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

$$\text{cum}\big(\langle X_1, \psi_{l_1}\rangle, \dots, \langle X_k, \psi_{l_k}\rangle\big) = \sum_{\nu=(\nu_1,\dots,\nu_p)} (-1)^{p-1}\,(p-1)!\,\prod_{r=1}^{p} \mathbb{E}\Big[\prod_{t\in\nu_r} \langle X_t, \psi_{l_t}\rangle\Big],$$

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

Theorem B.1. *Consider the tensor $X_t = \otimes_{j=1}^{J_t} X_{tj}$ for random elements X_{tj} in H with $j = 1, \dots, J_t$ and $t = 1, \dots, k$. Let $\nu = \{\nu_1, \dots \nu_p\}$ be a partition of $\{1, \dots, k\}$. The joint cumulant tensor $\text{cum}(X_1, \dots, X_k)$ is given by*

$$\text{cum}(X_1, \dots, X_k) = \sum_{r_{11},\dots,r_{kJ_t}} \sum_{\nu=(\nu_1,\dots,\nu_p)} \prod_{n=1}^{p} \text{cum}\big(\langle X_{tj}, \psi_{r_{tj}}\rangle | (t,j) \in \nu_n\big)\, \psi_{r_{11}} \otimes \cdots \otimes \psi_{r_{kJ_t}},$$

where the summation extends over all indecomposable partitions $\nu = (\nu_1, \ldots, \nu_p)$ 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

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

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

Next, results for cumulants of the fDFT are stated under both stationarity and local stationarity regimes.

Lemma B.1 (**Cumulants of the fDFT under stationarity**). *Let $(X_t \colon t \in \mathbb{Z})$ be a k-th order stationary sequence taking values in $H_{\mathbb{R}}$ that satisfies Assumption 4.1. The cumulant operator of the fDFT then satisfies*

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

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

The proof of this lemma can be found in Panaretos & Tavakoli (2013). To give the analog of Lemma B.1 in the locally stationary case, a number of auxiliary statements must be derived first. Because of space constraints, these statements and their proofs are relegated to the Online supplement. Lemma **??** and Lemma **??** allow to derive the that the cumulant tensors of the local fDFT can be expressed in terms of the time-varying spectral operator. At the Fourier frequencies, the time-varying spectral operator can in turn be shown to possess a well-defined Fourier transform. The properties of the resulting Fourier coefficients make 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 coefficients additionally provide an upper bound on the norm of the cumulant operator. This is summarized in the next lemma, which is the locally stationary version of Lemma B.1.

Lemma B.2 (**Cumulants of the fDFT under local stationarity**). *Let $(X_{t,T} \colon t \leqslant T, T \in \mathbb{N})$ be a k-th order locally stationary process in H satisfying Assumption 4.3. The cumulant operator of the local fDFT satisfies*

$$\begin{aligned} \mathrm{cum}\big(D^{(T)}_{\omega_{j_1}}, \ldots, D^{(T)}_{\omega_{j_k}}\big) &= \frac{(2\pi)^{k/2-1}}{T^{k/2}} \sum_{t=0}^{T-1} \mathcal{F}_{t/T;\omega_{j_1},\ldots,\omega_{j_{k-1}}} e^{-\mathrm{i}\sum_{l=1}^{k} t\omega_{j_l}} + R_{k,T} \\ &= \frac{(2\pi)^{k/2-1}}{T^{k/2-1}} \tilde{\mathcal{F}}_{j_1+\ldots+j_k;\omega_{j_1},\ldots,\omega_{j_{k-1}}} + R_{k,T}, \end{aligned} \tag{B.3}$$

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

$$\tilde{\mathcal{F}}_{s;\omega_{j_1},\ldots,\omega_{j_{k-1}}} = \int_0^1 \mathcal{F}_{u;\omega_{j_1},\ldots,\omega_{j_{k-1}}} e^{-\mathrm{i}2\pi s u} du \tag{B.4}$$

denotes the s-th Fourier coefficient of $\mathcal{F}_{u;\omega_{j_1},\ldots,\omega_{j_{k-1}}}$ and is Hilbert–Schmidt. for some constant $C > 0$.

In case the process does not depend on u, we have $\tilde{\mathcal{F}}_{s;\omega_{j_1},\ldots,\omega_{j_{k-1}}} = O_H$ 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 $s \neq 0$. The corollary below is a direct consequence of Lemma B.2.

Corollary B.1. *If Assumption 4.3 holds with $\ell = 2$, then*

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

(ii) $\sup_\omega \sum_{s\in\mathbb{Z}} |\!|\!|\tilde{\mathcal{F}}_{s;\omega}|\!|\!|_2 \leqslant \infty.$

C Companion results for the test defined through (3.1)

Theorem C.1. *Let Assumption 4.1 be satisfied with $k = \{2,4\}$. Then,*

$$\begin{aligned}
&T\,\mathrm{Cov}\Big(\Re\gamma^{(T)}_{h_1}(l_1,l_1'),\Re\gamma^{(T)}_{h_2}(l_2,l_2')\Big)\\
&\quad= T\,\mathrm{Cov}\Big(\Im\gamma^{(T)}_{h_1}(l_1,l_1'),\Im\gamma^{(T)}_{h_2}(l_2,l_2')\Big)\\
&\quad=\begin{cases}
\dfrac{1}{4\pi}\Bigg(\displaystyle\int \frac{\mathcal{F}^{(l_1,l_2)}_{\omega}\mathcal{F}^{(l_1',l_2')}_{-\omega-\omega_h}}{\sqrt{\mathcal{F}^{(l_1,l_1)}_{\omega}\mathcal{F}^{(l_2,l_2)}_{-\omega}\mathcal{F}^{(l_1',l_1')}_{-\omega-\omega_h}\mathcal{F}^{(l_2',l_2')}_{\omega+\omega_h}}}d\omega+\int \frac{\mathcal{F}^{(l_1,l_2')}_{\omega}\mathcal{F}^{(l_1',l_2)}_{-\omega-\omega_h}}{\sqrt{\mathcal{F}^{(l_1,l_1)}_{\omega}\mathcal{F}^{(l_2',l_2')}_{-\omega}\mathcal{F}^{(l_1',l_1')}_{-\omega-\omega_h}\mathcal{F}^{(l_2,l_2)}_{\omega+\omega_h}}}d\omega\\
\qquad+\displaystyle\iint \frac{\mathcal{F}^{(l_1\,l_1',l_2\,l_2')}_{\omega,-\omega-\omega_h,-\omega'}}{\sqrt{\mathcal{F}^{(l_1,l_1)}_{\omega}\mathcal{F}^{(l_1',l_1')}_{-\omega-\omega_h}\mathcal{F}^{(l_2,l_2)}_{-\omega'}\mathcal{F}^{(l_2',l_2')}_{\omega'+\omega_h}}}d\omega d\omega'\Bigg), & \text{if } h_1=h_2=h.\\
O(\frac{1}{T}), & \text{if } h_1\neq h_2.
\end{cases}
\end{aligned}$$

where $\mathcal{F}^{(l,l')}_{\omega} = \langle \mathcal{F}_\omega(\psi_{l'}),\psi_l\rangle$, $\mathcal{F}^{(l_1l_1',l_2l_2')}_{\omega_1,\omega_2,\omega_3} = \langle \mathcal{F}_{\omega_1,\omega_2,\omega_3}(\psi_{l_2l_2'}),\psi_{l_1l_1'}\rangle$ *and* $\psi_{ll'} = \psi_l\otimes\psi_{l'}$. *Furthermore,*

$$T\,\mathrm{Cov}(\Re\gamma^{(T)}_{h_1}(l_1,l_1'),\Im\gamma^{(T)}_{h_2}(l_2,l_2')) = O\Big(\frac{1}{T}\Big)$$

uniformly in $h_1,h_2\in\mathbb{Z}$.

Theorem C.2. *If the condition $\inf_\omega \lambda^\omega_L > 0$ is replaced with $\inf_\omega\langle\mathcal{F}(\psi),\psi\rangle > 0$ for all $\psi\in H_{\mathbb{C}}$, then parts (a) and (b) of Theorem 4.3 are retained using*

$$\sigma^2_{0,m} = \lim_{T\to\infty}\sum_{l_1,l_1',l_2,l_2'=1}^{L} T\mathrm{Cov}\big(\Re\gamma^{(T)}_{h_m}(l_1,l_1'),\Re\gamma^{(T)}_{h_m}(l_2,l_2')\big),\qquad m=1,\ldots,M,$$

where the explicit form of the limit is given by Theorem C.1.

Theorem C.3. *Let Assumption 4.3 be satisfied with $k = \{2,4\}$. Then, for $h = 1,\ldots,T-1$,*

$$\mathbb{E}\big[\gamma^{(T)}_h(l,l')\big] = \frac{1}{2\pi}\int_0^{2\pi}\int_0^1 \frac{\mathcal{F}^{(l,l')}_{u,\omega}e^{-\mathrm{i}2\pi uh}}{(G^{(l,l)}_{\omega}G^{(l',l')}_{\omega+\omega_h})^{1/2}}dud\omega + O\Big(\frac{1}{T}\Big) = O\Big(\frac{1}{h^2}\Big)+O\Big(\frac{1}{T}\Big). \tag{C.1}$$

The covariance structure satisfies

1. $T\,\mathrm{Cov}(\Re\gamma^{(T)}_{h_1}(l_1,l_2),\Re\gamma^{(T)}_{h_2}(l_3,l_4)) =$
$$\frac{1}{4}\big[\Sigma^{(T)}_{h_1,h_2}(\boldsymbol{l}_4)+\acute{\Sigma}^{(T)}_{h_1,h_2}(\boldsymbol{l}_4)+\grave{\Sigma}^{(T)}_{h_1,h_2}(\boldsymbol{l}_4)+\bar{\Sigma}^{(T)}_{h_1,h_2}(\boldsymbol{l}_4)\big]+R_T,$$

2. $T\,\mathrm{Cov}(\Re\gamma^{(T)}_{h_1}(l_1,l_2),\Im\gamma^{(T)}_{h_2}(l_3,l_4)) =$
$$\frac{1}{4\mathrm{i}}\big[\Sigma^{(T)}_{h_1,h_2}(\boldsymbol{l}_4)-\acute{\Sigma}^{(T)}_{h_1,h_2}(\boldsymbol{l}_4)+\grave{\Sigma}^{(T)}_{h_1,h_2}(\boldsymbol{l}_4)-\bar{\Sigma}^{(T)}_{h_1,h_2}(\boldsymbol{l}_4)\big]+R_T,$$

3. $T\,\mathrm{Cov}(\Im\gamma^{(T)}_{h_1}(l_1,l_2),\Im\gamma^{(T)}_{h_2}(l_3,l_4)) =$
$$\frac{1}{4}\big[\Sigma^{(T)}_{h_1,h_2}(\boldsymbol{l}_4)-\acute{\Sigma}^{(T)}_{h_1,h_2}(\boldsymbol{l}_4)-\grave{\Sigma}^{(T)}_{h_1,h_2}(\boldsymbol{l}_4)+\bar{\Sigma}^{(T)}_{h_1,h_2}(\boldsymbol{l}_4)\big]+R_T,$$

where $\Sigma^{(T)}_{h_1,h_2}(\boldsymbol{l}_4),\acute{\Sigma}^{(T)}_{h_1,h_2}(\boldsymbol{l}_4),\grave{\Sigma}^{(T)}_{h_1,h_2}(\boldsymbol{l}_4),\bar{\Sigma}^{(T)}_{h_1,h_2}(\boldsymbol{l}_4)$ *are defined in equations (S.5.1)–(S.5.4) of the Online Supplement. and* $\|R_T\|_2 = O(T^{-1})$.

Theorem C.4. *If the condition* $\inf_\omega \lambda^\omega_L > 0$ *is replaced with* $\inf_\omega\langle\mathcal{F}(\psi),\psi\rangle > 0$ *for all* $\psi \in H_{\mathbb{C}}$*, then parts (a) and (b) of Theorem C.4 are retained using* Σ_A *defined through*

$$\Sigma^{(11)}_A(m,m') = \lim_{T\to\infty}\sum_{l_1,l_2,l_3,l_4=1}^{L}\frac{1}{4}\big[\Sigma^{(T)}_{h_m,h_{m'}}(\boldsymbol{l}_4)+\acute{\Sigma}^{(T)}_{h_m,h_{m'}}(\boldsymbol{l}_4)+\grave{\Sigma}^{(T)}_{h_m,h_{m'}}(\boldsymbol{l}_4)+\bar{\Sigma}^{(T)}_{h_m,h_{m'}}(\boldsymbol{l}_4)\big],$$
$$\Sigma^{(12)}_A(m,m') = \lim_{T\to\infty}\sum_{l_1,l_2,l_3,l_4=1}^{L}\frac{1}{4\bar{\mathrm{i}}}\big[\Sigma^{(T)}_{h_m,h_{m'}}(\boldsymbol{l}_4)-\acute{\Sigma}^{(T)}_{h_m,h_{m'}}(\boldsymbol{l}_4)+\grave{\Sigma}^{(T)}_{h_m,h_{m'}}(\boldsymbol{l}_4)-\bar{\Sigma}^{(T)}_{h_m,h_{m'}}(\boldsymbol{l}_4)\big],$$
$$\Sigma^{(22)}_A(m,m') = \lim_{T\to\infty}\sum_{l_1,l_2,l_3,l_4=1}^{L}\frac{1}{4}\big[\Sigma^{(T)}_{h_m,h_{m'}}(\boldsymbol{l}_4)-\acute{\Sigma}^{(T)}_{h_m,h_{m'}}(\boldsymbol{l}_4)-\grave{\Sigma}^{(T)}_{h_m,h_{m'}}(\boldsymbol{l}_4)+\bar{\Sigma}^{(T)}_{h_m,h_{m'}}(\boldsymbol{l}_4)\big],$$

where $\Sigma^{(T)}_{h_m,h_{m'}}(\boldsymbol{l}_4),\acute{\Sigma}^{(T)}_{h_m,h_{m'}}(\boldsymbol{l}_4),\grave{\Sigma}^{(T)}_{h_m,h_{m'}}(\boldsymbol{l}_4),\bar{\Sigma}^{(T)}_{h_m,h_{m'}}(\boldsymbol{l}_4)$ *are as in Theorem C.3.*

D First and second order dependence structure

D.1 Under the null hypothesis of stationarity

Proof of Theorem 4.2. Using Lemma B.1, for $h = 1,\dots,T-1$, and

$$\langle x,y\rangle\overline{\langle z,w\rangle} = \langle (x\otimes z)w,y\rangle = \langle x\otimes z, y\otimes w\rangle_{HS} \qquad x,y,z,w\in H,$$

where $\langle\cdot,\cdot\rangle_{HS}$ denotes the Hilbert–Schmidt inner product, the expectation of (3.2) is given by

$$\mathbb{E}\big[\sqrt{T}\gamma^{(T)}_h(l)\big] = \frac{1}{\sqrt{T}}\sum_{j=1}^{T}\left[\frac{1}{T}\Delta^{(h)}_T\frac{\langle\mathcal{F}_{\omega_j},\phi^{\omega_j}_l\otimes\phi^{\omega_{j+h}}_l\rangle_{HS}}{\sqrt{\lambda^{\omega_j}_l\lambda^{\omega_{j+h}}_l}}+O\Big(\frac{1}{T}\Big)\right] = O\Big(\frac{1}{\sqrt{T}}\Big). \tag{D.1}$$

Additionally, Theorem B.1 implies that the covariance structure of the fDFT's is given by

$$\begin{aligned}\mathrm{Cov}\big(D^{(T)}_{\omega_{j_1}}\otimes D^{(T)}_{\omega_{j_1+h_1}},D^{(T)}_{\omega_{j_2}}\otimes D^{(T)}_{\omega_{j_2+h_2}}\big) &= \mathrm{cum}(D^{(T)}_{\omega_{j_1}},D^{(T)}_{-\omega_{j_1+h_1}},D^{(T)}_{-\omega_{j_2}},D^{(T)}_{\omega_{j_2+h_2}})\\ &+S_{1324}\big(\mathrm{cum}(D^{(T)}_{\omega_{j_1}},D^{(T)}_{-\omega_{j_2}})\otimes\mathrm{cum}(D^{(T)}_{-\omega_{j_1+h_1}},D^{(T)}_{\omega_{j_2+h_2}})\big)\end{aligned}$$

$$+ S_{1423}\big(\mathrm{cum}(D^{(T)}_{\omega_{j_1}}, D^{(T)}_{\omega_{j_2+h_2}}) \otimes \mathrm{cum}(D^{(T)}_{-\omega_{j_1+h_1}}, D^{(T)}_{-\omega_{j_2}})\big), \quad \text{(D.2)}$$

where S_{ijkl} denotes the permutation operator on $\otimes_{i=1}^{4} L^2_{\mathbb{C}}([0,1])$ that permutes the components of a tensor according to the permutation $(1,2,3,4) \mapsto (i,j,k,l)$, that is, $S_{ijkl}(x_1 \otimes \cdots \otimes x_4) = x_i \otimes \cdots \otimes x_l$. Denote then the elementary tensor $\psi_{ll'} = \psi_l \otimes \psi_{l'}$. Write $D^{(l)}_\omega = \langle D^{(T)}_\omega, \psi_l \rangle$, $\mathcal{F}^{(l,l')}_\omega = \langle \mathcal{F}_\omega(\psi_{l'}), \psi_l \rangle$ and $\mathcal{F}^{(lm,l'm')}_{\omega_{j_1},\omega_{j_2},\omega_{j_3}} = \langle \mathcal{F}_{\omega_{j_1},\omega_{j_2},\omega_{j_3}}(\psi_{l'm'}), \psi_{lm} \rangle$. Under the conditions of Theorem 4.2, Lemma B.2 implies for the denominator in (3.1) that

$$\frac{1}{T}\sum_{j_1,j_2=1}^{T}\Bigg(\frac{(2\pi)}{T^2}\mathcal{F}^{(l_1\,l_2,l_3\,l_4)}_{\omega_{j_1},-\omega_{j_1+h_1},-\omega_{j_2}}\Delta_T^{(\omega_{h_2}-\omega_{h_1})} + O\Big(\frac{1}{T^2}\Big)$$
$$+\left[\mathcal{F}^{(l_1,l_3)}_{\omega_{j_1}}\frac{1}{T}\Delta_T^{(\omega_{j_1}-\omega_{j_2})} + O\Big(\frac{1}{T}\Big)\right]\left[\mathcal{F}^{(l_2,l_4)}_{-\omega_{j_1+h_1}}\frac{1}{T}\Delta_T^{(\omega_{j_1+h_1}-\omega_{j_2+h_2})} + O\Big(\frac{1}{T}\Big)\right]$$
$$+\left[\mathcal{F}^{(l_1,l_4)}_{\omega_{j_1}}\frac{1}{T}\Delta_T^{(\omega_{j_1}+\omega_{j_2+h_2})} + O\Big(\frac{1}{T}\Big)\right]\left[\mathcal{F}^{(l_2,l_3)}_{-\omega_{j_1+h_1}}\frac{1}{T}\Delta_T^{(-\omega_{j_1+h_1}-\omega_{j_2})} + O\Big(\frac{1}{T}\Big)\right]\Bigg).$$

For the statistic (3.2) this structure becomes

$$\frac{1}{T}\sum_{j_1,j_2=1}^{T}\Bigg(\frac{(2\pi)}{T^2}\langle\mathcal{F}_{\omega_{j_1},-\omega_{j_1+h_1},-\omega_{j_2}}(\phi^{\omega_{j_2}}_{l_2}\otimes\phi^{\omega_{j_2+h_2}}_{l_2}), \phi^{\omega_{j_1}}_{l_1}\otimes\phi^{\omega_{j_1+h_1}}_{l_1}\rangle\Delta_T^{(\omega_{h_2}-\omega_{h_1})} + O\big(\frac{1}{T^2}\big)$$
$$+\left[\langle\mathcal{F}_{\omega_{j_1}}\phi^{\omega_{j_2}}_{l_2}, \phi^{\omega_{j_1}}_{l_1}\rangle\frac{1}{T}\Delta_T^{(\omega_{j_1}-\omega_{j_2})} + O\big(\frac{1}{T}\big)\right]\left[\langle\mathcal{F}_{-\omega_{j_1+h_1}}\phi^{-\omega_{j_2+h_2}}_{l_2}, \phi^{-\omega_{j_1+h_1}}_{l_1}\rangle\frac{1}{T}\Delta_T^{(\omega_{j_1+h_1}-\omega_{j_2+h_2})} + O\big(\frac{1}{T}\big)\right]$$
$$+\big[\langle\mathcal{F}_{\omega_{j_1}}\phi^{-\omega_{j_2+h_2}}_{l_2}, \phi^{\omega_{j_1}}_{l_1}\rangle\frac{1}{T}\Delta_T^{(\omega_{j_1}+\omega_{j_2+h_2})} + O\big(\frac{1}{T}\big)\Big]\Big[\langle(\mathcal{F}_{-\omega_{j_1+h_1}}\phi^{\omega_{j_2}}_{l_2}, \phi^{-\omega_{j_1+h_1}}_{l_1}\rangle\frac{1}{T}\Delta_T^{(-\omega_{j_1+h_1}-\omega_{j_2})} + O\big(\frac{1}{T}\big)\Big]\Bigg).$$

Using that the spectral density operator is self-adjoint and the projection basis are its eigenfunctions, it follows

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

The covariance strucure of the numerator of (3.2) then reduces to

$$= \frac{1}{T}\sum_{j_1,j_2=1}^{T}\Bigg(\frac{(2\pi)}{T^2}\langle\mathcal{F}_{\omega_{j_1},-\omega_{j_1+h_1},-\omega_{j_2}}(\phi^{\omega_{j_2}}_{l_2}\otimes\phi^{\omega_{j_2+h_2}}_{l_2}), \phi^{\omega_{j_1}}_{l_1}\otimes\phi^{\omega_{j_1+h_1}}_{l_1}\rangle\Delta_T^{(\omega_{h_2}-\omega_{h_1})} + O\big(\frac{1}{T^2}\big)$$
$$+\left[\lambda^{\omega_{j_1}}_{l_1}\langle\phi^{\omega_{j_2}}_{l_2}, \phi^{\omega_{j_1}}_{l_1}\rangle\frac{1}{T}\Delta_T^{(\omega_{j_1}-\omega_{j_2})} + O\big(\frac{1}{T}\big)\right]\left[\lambda^{-\omega_{j_1+h_1}}_{l_1}\langle\phi^{-\omega_{j_2+h_2}}_{l_2}, \phi^{-\omega_{j_1+h_1}}_{l_1}\rangle\frac{1}{T}\Delta_T^{(\omega_{j_1+h_1}-\omega_{j_2+h_2})} + O\big(\frac{1}{T}\big)\right]$$
$$+\big[\lambda^{\omega_{j_1}}_{l_1}\langle\phi^{-\omega_{j_2+h_2}}_{l_2}, \phi^{\omega_{j_1}}_{l_1}\rangle\frac{1}{T}\Delta_T^{(\omega_{j_1}+\omega_{j_2+h_2})} + O\big(\frac{1}{T}\big)\Big]\Big[\lambda^{-\omega_{j_1+h_1}}_{l_1}\langle\phi^{\omega_{j_2}}_{l_2}, \phi^{-\omega_{j_1+h_1}}_{l_1}\rangle\frac{1}{T}\Delta_T^{(-\omega_{j_1+h_1}-\omega_{j_2})} + O\big(\frac{1}{T}\big)\Big].$$

In case $h_1 \neq h_2$, the first line is of order $O(T^{-1})$. In the second and third line, it can be seen that the cross terms will be of order $O(T^{-1})$ uniformly in ω_j. The first product term in the second line will only be of order $O(1)$ if $j_1 = j_2$ and $j_1 + h_1 = j_2 + h_2$, while in the third line this requires $j_1 = -j_2 - h_2$ and $j_2 = -j_1 - h_1$. It can then be derived that $T\mathrm{Cov}(\gamma^{(T)}_{h_1}(l_1), \overline{\gamma^{(T)}_{h_2}(l_2)}) = O(T^{-1})$ for $h_2 \neq T - h_1$. Since

$$\Re\gamma^{(T)}_{h_1}(l_1) = \frac{1}{2}(\gamma^{(T)}_{h_1}(l_1) + \overline{\gamma^{(T)}_{h_1}(l_1)})$$

and

$$\Im\gamma_{h_1}^{(T)}(l_1) = \frac{1}{2\mathrm{i}}\Big(\gamma_{h_1}^{(T)}(l_1) - \overline{\gamma_{h_1}^{(T)}(l_1)}\Big),$$

it follows therefore

$$T\mathrm{Cov}\big(\Re\gamma_{h_1}^{(T)}(l_1, \Im\gamma_{h_2}^{(T)}(l_2)\big) = O(T^{-1})$$

uniformly in h_1, h_2. All together, the above derivation yields

$$\begin{aligned} T\,\mathrm{Cov}\big(\Re\gamma_{h_1}^{(T)}(l_1), \Re\gamma_{h_2}^{(T)}(l_2)\big) &= T\,\mathrm{Cov}\big(\Im\gamma_{h_1}^{(T)}(l_1), \Im\gamma_{h_2}^{(T)}(l_2)\big) \\ &= \frac{T}{2}\mathrm{Cov}\big(\gamma_{h_1}^{(T)}(l_1), \gamma_{h_2}^{(T)}(l_2)\big), \end{aligned}$$

and thus Lipschitz-continuity and orthogonality of the eigenfunctions imply that the limiting covariance structure of (3.2) is given by

$$\begin{aligned} T\,\mathrm{Cov}\Big(\Re\gamma_{h_1}^{(T)}(l_1), \Re\gamma_{h_2}^{(T)}(l_2)\Big) &= T\,\mathrm{Cov}\Big(\Im\gamma_{h_1}^{(T)}(l_1), \Im\gamma_{h_2}^{(T)}(l_2)\Big) \\ = \frac{\delta_{l_1,l_2}}{2} + \frac{1}{4\pi}\int\int &\frac{\langle \mathcal{F}_{\omega,-\omega-\omega_h,-\omega'}(\phi_{l_2}^{\omega'} \otimes \phi_{l_2}^{\omega'+\omega'_h}), \phi_{l_1}^{\omega} \otimes \phi_{l_1}^{\omega+\omega_h}\rangle}{\sqrt{\lambda_{l_1}^{\omega}\lambda_{l_2}^{-\omega'}\lambda_{l_1}^{-\omega-\omega_h}\lambda_{l_2}^{\omega'+\omega'_h}}} d\omega d\omega'. \end{aligned}$$

This completes the proof. □

D.2 Under the alternative hypothesis of local stationarity

Proof of Theorem 4.4. (i) In order to prove the first assertion of the theorem, introduce the bias-variance decomposition

$$\begin{aligned} &\mathbb{E}\Big[\big\|\hat{\mathcal{F}}_{\omega}^{(T)} - \mathbb{E}\big[\hat{\mathcal{F}}_{\omega}^{(T)}\big] + \mathbb{E}\big[\hat{\mathcal{F}}_{\omega}^{(T)}\big] - G_{\omega}\big\|_2^2\Big] \\ &\quad = \mathbb{E}\Big[\big\|\hat{\mathcal{F}}_{\omega}^{(T)} - \mathbb{E}\big[\hat{G}_{\omega}\big]\big\|_2^2\Big] + \mathbb{E}\Big[\big\|\mathbb{E}\big[\hat{\mathcal{F}}_{\omega}^{(T)}\big] - G_{\omega}\big\|_2^2\Big]. \end{aligned} \tag{D.3}$$

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 Lemma B.1,

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

where $\|R_{T,2}\|_2 = O(T^{-1})$. Convolution of the cumulant tensor with the smoothing kernel and a subsequent Taylor expansion give

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

where $\|\epsilon_{b,T}\|_2 = O(b^2 + (bT))^{-1}$. The interchange of summations is justified by Fubini's Theorem, since $\sup_{\omega,u}\|f_{u,\omega}\|_2 < \infty$ and $\sup_{\omega,u}\|\frac{\partial^2}{\partial\omega_2}f_{u,\omega}\|_2 < \infty$. Here, the error term $(bT)^{-1}$ follows from discretization of the window function (see, for example, Lemma P5.1 of Brillinger, 1981). Note that the integral approximation in time direction does not change the error term because of Lipschitz continuity in u. Thus, the second term in (D.3) satisfies

$$\mathbb{E}\big[\|\!|\mathbb{E}\hat{\mathcal{F}}_{\omega}^{(T)} - G_{\omega}\|\!|_2^2\big] = O\Big(b^2 + \frac{1}{bT}\Big)^2.$$

To bound the first term of the right-hand side in (D.3), observe that, for $\omega \neq \omega'$,

$$\begin{aligned}
\operatorname{cum}(D_{\omega}^{(T)}, D_{\omega'}^{(T)}) &= \frac{1}{2\pi T}\sum_{t,t'=1}^{T}\operatorname{cum}(X_t^{(T)}, X_{t'}^{(T)})e^{-\mathrm{i}(t\omega+t'\omega')} \\
&= \frac{1}{2\pi T}\sum_{t,t'=1}^{T}\mathcal{C}_{t/T;t'-t}(\tau,\tau')e^{-\mathrm{i}(t'-t)\omega'-\mathrm{i}t(\omega+\omega')} + \epsilon_{t-t',T} \\
&= \frac{1}{2\pi T}\sum_{t=1}^{T}\sum_{|h|\leqslant T-t}\operatorname{cum}(X_t^{(T)}, X_{t+h}^{(T)})e^{-\mathrm{i}(h\omega')-\mathrm{i}t(\omega+\omega')} + R_{T,2} \\
&= \frac{1}{2\pi T}\sum_{t=1}^{T}\sum_{|h|\leqslant T-t}\mathcal{F}_{t/T,h}e^{-\mathrm{i}(h\omega')-\mathrm{i}t(\omega+\omega')} + R_{T,2}, \qquad \text{(D.4)}
\end{aligned}$$

where Lemma **??** was applied to obtain the second equality sign in combination with

$$\|\epsilon_T\|_2 = \frac{1}{2\pi T}\sum_{t_1=1}^{T}\frac{1+|t_1-t_2|}{T}\|\kappa_{2;t_1-t_2}\|_2 = O\Big(\frac{1}{T}\Big)$$

by (4.3). Decompose the corresponding local autocovariance operator of (D.4) as

$$\frac{1}{2\pi T}\Bigg(\sum_{h=0}^{T-1}\sum_{t=1}^{T-h}\mathcal{C}_{t/T;h}e^{-\mathrm{i}(h\omega')-\mathrm{i}t(\omega+\omega')} + \sum_{h=-T+1}^{-1}\sum_{t=1}^{T-|h|}\mathcal{C}_{t/T;h}e^{-\mathrm{i}(h\omega')-\mathrm{i}t(\omega+\omega')}\Bigg) + R_{T,2}. \qquad \text{(D.5)}$$

Under Assumption 4.3,

$$\begin{aligned}
\Bigg\|\!\Bigg|\frac{1}{2\pi T}\sum_{h=0}^{T-1}\sum_{t=1}^{T-h}\mathcal{C}_{t/T,h}e^{-\mathrm{i}(h\omega')-\mathrm{i}t(\omega+\omega')}\Bigg\|\!\Bigg|_2 &\leqslant \frac{1}{2\pi T}\sum_{h=0}^{T-1}\Bigg|\sum_{t=1}^{T-h}e^{-\mathrm{i}t(\omega+\omega')}\Bigg|\|\!|\mathcal{C}_{t/T,h}\|\!|_2 \\
\leqslant \frac{1}{2\pi T}\sum_{h=0}^{T-1}\big|\Delta_{T-h}^{(\omega+\omega')}\big|\|\!|\mathcal{C}_{t/T,h}\|\!|_2 &\leqslant \frac{C}{T}\sum_{h\in\mathbb{Z}}|h|\|\kappa_{2;h}\|_2 = O\Big(\frac{1}{T}\Big)
\end{aligned}$$

for some constant C. A similar derivation shows the same bound holds for the second term of (D.5). It can therefore be concluded that $\|\!|\operatorname{cum}(D_{\omega}^{(T)}, D_{\lambda}^{(T)})\|\!|_2 = O(T^{-1})$ uniformly in $\omega \neq \omega', 0 \leqslant \omega, \omega' < \pi$. Furthermore, Lemma B.1 and Minkowski's Inequality yield

$$\begin{aligned}
\Big\|\!\Big|\operatorname{cum}(D_{\omega}^{(T)}, D_{-\omega}^{(T)}, D_{\omega'}^{(T)}, D_{-\omega'}^{(T)})\Big\|\!\Big|_2 &\leqslant \frac{1}{T}\Bigg\|\!\Bigg|\frac{1}{T}\sum_{t=0}^{T-1}\mathcal{F}_{\frac{t}{T},\omega,-\omega,\omega'}\Bigg\|\!\Bigg|_2 + O\Big(\frac{1}{T^2}\Big) \\
&= \frac{1}{T}\Big\|\!\Big|G_{\omega,-\omega,\omega'}^{(T)}\Big\|\!\Big|_2 + O\Big(\frac{1}{T^2}\Big) = O\Big(\frac{1}{T}\Big).
\end{aligned}$$

The last equality follows since $\sup_{u,\omega}\|\!|\mathcal{F}_{t/T,\omega,-\omega,\omega'}\|\!|_2 \leqslant \sum_{h_1,h_2,h_3\in\mathbb{Z}}\|\kappa_{3;h_1,h_2,h_3}\|_2 = O(1)$ by Assumption 4.3. Therefore the product theorem for cumulant tensors (Theorem B.1) implies that

$$\operatorname{Cov}\big(I_{\omega}^{(T)}, I_{\omega'}^{(T)}\big) = \operatorname{cum}\big(D_{\omega}^{(T)}, D_{-\omega}^{(T)}, D_{-\omega'}^{(T)}, D_{\omega'}^{(T)}\big)$$

$$+ S_{1324}\Big(\mathrm{cum}\big(D_\omega^{(T)}, D_{-\omega'}^{(T)}\big) \otimes \mathrm{cum}\big(D_{-\omega}^{(T)}, D_{\omega'}^{(T)}\big)\Big)$$
$$+ S_{1423}\Big(\mathrm{cum}\big(D_\omega^{(T)}, D_{\omega'}^{(T)}\big) \otimes \mathrm{cum}\big(D_{-\omega}^{(T)}, D_{-\omega'}^{(T)}\big)\Big), \tag{D.6}$$

where S_{ijkl} denotes the permutation operator on $\otimes_{i=1}^4 L^2_{\mathbb{C}}([0,1])$ that permutes the components of a tensor according to the permutation $(1,2,3,4) \mapsto (i,j,k,l)$, that is, $S_{ijkl}(x_1 \otimes \cdots \otimes x_4) = x_i \otimes \cdots \otimes x_l$. It is clear from (D.6) that $|||\mathrm{Cov}(I_\omega^{(T)}, I_{\omega'}^{(T)})|||_2 = O(T^{-1})$ for $\omega' \neq \omega, 0 \leqslant \omega, \omega' < \pi$, while for $\omega' = \omega$ it follows that

$$\big|\big|\big|\mathrm{Cov}\big(I_\omega^{(T)}, I_\omega^{(T)}\big)\big|\big|\big|_2 \leqslant \Big|\Big|\Big| S_{1324}\Big(\mathrm{cum}\big(D_\omega^{(T)}, D_{-\omega}^{(T)}\big) \otimes \mathrm{cum}\big(D_{-\omega}^{(T)}, D_{\omega}^{(T)}\big)\Big)\Big|\Big|\Big|_2$$
$$+ \Big|\Big|\Big| S_{1423}\Big(\mathrm{cum}\big(D_\omega^{(T)}, D_{\omega}^{(T)}\big) \otimes \mathrm{cum}\big(D_{-\omega}^{(T)}, D_{-\omega}^{(T)}\big)\Big)\Big|\Big|\Big|_2 + R_{T,2} = O(1). \tag{D.7}$$

Furthermore,

$$\mathrm{Cov}\big(\hat{\mathcal{F}}_\omega^{(T)}, \hat{\mathcal{F}}_\omega^{(T)}\big) = \left(\frac{2\pi}{T}\right)^2 \sum_{j,j'=1}^{T} K_b(\omega - \omega_j) K_b(\omega - \omega_{j'})$$
$$\times \frac{1}{T^2}\Bigg(\sum_{t,t'=0}^{T-1} \mathcal{F}_{t/T,\omega_j} \otimes \mathcal{F}_{t/T,-\omega_j} e^{-\mathrm{i}(t-t')(\omega_j - \omega_{j'})}$$
$$+ \sum_{t,t'=0}^{T-1} \mathcal{F}_{t/T,\omega_j} \otimes \mathcal{F}_{t/T,-\omega_j} e^{-\mathrm{i}(t-t')(\omega_j + \omega_{j'})}\Bigg) + R_{T,2}. \tag{D.8}$$

Hence,

$$\Big|\Big|\Big|\mathrm{Cov}(\hat{\mathcal{F}}_\omega^{(T)}, \hat{\mathcal{F}}_\omega^{(T)})\Big|\Big|\Big|_2 \leqslant \sup_{u,\omega} |||\mathcal{F}_{u,\omega}|||_2^2 \left(\frac{2\pi}{T}\right)^2 \sum_{j,j'=1}^{T} K_b(\omega - \omega_j) K_b(\omega - \omega_{j'}) \frac{|\Delta_T^{(\omega_j - \omega_{j'})}|^2}{T^2}$$
$$+ \sup_{u,\omega} |||\mathcal{F}_{u,\omega}|||_2^2 \left(\frac{2\pi}{T}\right)^2 \sum_{j,j'=1}^{T} K_b(\omega - \omega_j) K_b(\omega - \omega_{j'}) \frac{|\Delta_T^{(\omega_j + \omega_{j'})}|^2}{T^2} + O\left(\frac{1}{T}\right)$$
$$= O\left(\frac{1}{bT}\right). \tag{D.9}$$

Together with the equivalence of the Hilbert–Schmidt norm of the operator and the L_2–norm of its kernel, the above implies then that the second term of (D.3) satisfies

$$\mathbb{E}\big[|||\hat{\mathcal{F}}_\omega^{(T)} - \mathbb{E}\hat{\mathcal{F}}_\omega^{(T)}|||_2^2\big] = \int_{[0,1]^2} \mathrm{Var}(\hat{f}_\omega^{(T)}(\tau,\tau'))d\tau d\tau' = O\left(\frac{1}{bT}\right)$$

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

(ii) The second part of the proof proceeds along similar lines as Paparoditis (2009). An application of Minkowski's inequality yields

$$|||\hat{\mathcal{F}}_\omega^{(T)} - G_\omega|||_2 \leqslant \Bigg|\Bigg|\Bigg| \frac{2\pi}{T} \sum_{j=1}^{T} K_b(\omega_{j'} - \omega_j)\Big[\mathrm{cum}(D_{\omega_j}^{(T}, D_{-\omega_j}^{(T)}) - \frac{1}{T}\sum_{t=1} \mathcal{F}_{t/T,\omega_j}\Big]\Bigg|\Bigg|\Bigg|_2$$
$$+ \Bigg|\Bigg|\Bigg| \frac{2\pi}{T} \sum_{j=1}^{T} K_b(\omega_{j'} - \omega_j) \frac{1}{T}\sum_{t=1} \big[f_{t/T,\omega_j} - \mathcal{F}_{t/T,\omega}\big]\Bigg|\Bigg|\Bigg|_2$$

$$+\left\|\left(\frac{2\pi}{T}\sum_{j=1}^{T}K_b(\omega_{j'}-\omega_j)-1\right)\frac{1}{T}\sum_{t=1}\mathcal{F}_{t/T,\omega}\right\|_2.$$

Markov's inequality together with (i), which is not affected by the discretization of the integral, imply the first term tends to zero. Since the spectral operator is Lipschitz continuous in ω, the second term is bounded by

$$\left|1-\int_{-\pi}^{\pi}K_b(\omega_k-\omega')d\omega'\right||b|=O\left(1+\frac{1}{bT}\right)O(b)=O(b).$$

Finally, the third term is seen to be of order $O((bT)^{-1})$. □

Proof of Theorem 4.6. By Lemma B.2, the expectation of $\gamma_h^{(T)}(l,l')$ satisfies

$$\mathbb{E}\big[\gamma_h^{(T)}(l,l')\big]=\frac{1}{T}\sum_{j=1}^{T}\left[\frac{1}{T}\sum_{t=0}^{T-1}\frac{\mathcal{F}^{(l,l')}_{t/T,\omega_j}}{(G^{(l,l)}_{\omega_j}G^{(l',l')}_{\omega_{j+h}})^{1/2}}e^{-\mathrm{i}t\omega_h}+O\left(\frac{1}{T}\right)\right]=O\left(\frac{1}{h^2}+\frac{1}{T}\right),$$

for all $h=1,\ldots,T-1$. Using then Lipschitz continuity of the spectral operators taking as basis the eigenfunctions of spectral density operator, expression (C.1) becomes

$$\mathbb{E}\big[\gamma_h^{(T)}(l)\big]=\frac{1}{2\pi}\int_0^{2\pi}\int_0^1\frac{\langle\mathcal{F}_{u,\omega}\tilde{\phi}_l^{\omega+\omega_h},\tilde{\phi}_l^{\omega}\rangle e^{-\mathrm{i}2\pi uh}}{(\tilde{\lambda}_l^{\omega}\tilde{\lambda}_l^{\omega+\omega_h})^{1/2}}du d\omega+O\left(\frac{1}{T}\right)=O\left(\frac{1}{h^2}\right)+O\left(\frac{1}{T}\right).$$

Note once more that

$$\Re\gamma_h^{(T)}(l)=\frac{1}{2}\left(\gamma_h^{(T)}(l)+\overline{\gamma_h^{(T)}(l)}\right)\quad\text{and}\quad\Im\gamma_h^{(T)}(l)=\frac{1}{2\mathrm{i}}\left(\gamma_h^{(T)}(l)-\overline{\gamma_h^{(T)}(l)}\right).$$

Under the alternative, these are in fact correlated and four separate cases will have to be considered:

$$\text{(i)}\quad \mathrm{Cov}\big(\Re\gamma_{h_1}^{(T)}(l_1),\Re\gamma_{h_2}^{(T)}(l_2)\big)=\frac{1}{4}\Big[\mathrm{Cov}\big(\gamma_{h_1}^{(T)}(l_1),\gamma_{h_2}^{(T)}(l_2)\big)+\mathrm{Cov}\big(\gamma_{h_1}^{(T)}(l_1),\overline{\gamma_{h_2}^{(T)}(l_2)}\big)+\mathrm{Cov}\big(\overline{\gamma_{h_1}^{(T)}(l_1)},\gamma_{h_2}^{(T)}(l_2)\big)+\mathrm{Cov}\big(\overline{\gamma_{h_1}^{(T)}(l_1)},\overline{\gamma_{h_2}^{(T)}(l_2)}\big)\Big],$$

$$\text{(ii)}\quad \mathrm{Cov}\big(\Re\gamma_{h_1}^{(T)}(l_1),\Im\gamma_{h_2}^{(T)}(l_2)\big)=\frac{1}{4\bar{\mathrm{i}}}\Big[\mathrm{Cov}\big(\gamma_{h_1}^{(T)}(l_1),\gamma_{h_2}^{(T)}(l_2)\big)-\mathrm{Cov}\big(\gamma_{h_1}^{(T)}(l_1),\overline{\gamma_{h_2}^{(T)}(l_2)}\big)+\mathrm{Cov}\big(\overline{\gamma_{h_1}^{(T)}(l_1)},\gamma_{h_2}^{(T)}(l_2)\big)-\mathrm{Cov}\big(\overline{\gamma_{h_1}^{(T)}(l_1)},\overline{\gamma_{h_2}^{(T)}(l_2)}\big)\Big],$$

$$\text{(iii)}\quad \mathrm{Cov}\big(\Im\gamma_{h_1}^{(T)}(l_1),\Re\gamma_{h_2}^{(T)}(l_2)\big)=\frac{1}{4\bar{\mathrm{i}}}\Big[\mathrm{Cov}\big(\gamma_{h_1}^{(T)}(l_1),\gamma_{h_2}^{(T)}(l_2)\big)+\mathrm{Cov}\big(\gamma_{h_1}^{(T)}(l_1),\overline{\gamma_{h_2}^{(T)}(l_2)}\big)-\mathrm{Cov}\big(\overline{\gamma_{h_1}^{(T)}(l_1)},\gamma_{h_2}^{(T)}(l_2)\big)-\mathrm{Cov}\big(\overline{\gamma_{h_1}^{(T)}(l_1)},\overline{\gamma_{h_2}^{(T)}(l_2)}\big)\Big],$$

$$\text{(iv)}\quad \mathrm{Cov}\big(\Im\gamma_{h_1}^{(T)}(l_1),\Im\gamma_{h_2}^{(T)}(l_2)\big)=\frac{1}{4}\Big[\mathrm{Cov}\big(\gamma_{h_1}^{(T)}(l_1),\gamma_{h_2}^{(T)}(l_2)\big)-\mathrm{Cov}\big(\gamma_{h_1}^{(T)}(l_1),\overline{\gamma_{h_2}^{(T)}(l_2)}\big)-\mathrm{Cov}\big(\overline{\gamma_{h_1}^{(T)}(l_1)},\gamma_{h_2}^{(T)}(l_2)\big)+\mathrm{Cov}\big(\overline{\gamma_{h_1}^{(T)}(l_1)},\overline{\gamma_{h_2}^{(T)}(l_2)}\big)\Big].$$

The expressions for the covariance structure of $\sqrt{T}\gamma_h^{(T)}$ and its conjugate can now be derived using Lemma B.2. For example,

$$\mathrm{Cov}\big(\langle D_{\omega_{j_1}},\phi_{l_1}^{\omega_{j_1}}\rangle\overline{\langle D_{\omega_{j_1+h_1}},\phi_{l_1}^{\omega_{j_1+h_1}}\rangle},\langle D_{\omega_{j_2}},\phi_{l_2}^{\omega_{j_2}}\rangle\overline{\langle D_{\omega_{j_2+h_2}},\phi_{l_2}^{\omega_{j_2+h_2}}\rangle}\big)$$

$$=\Big[\langle \tilde{\mathfrak{F}}_{j_1-j_2;\omega_{j_1}} \phi_{l_2}^{\omega_{j_2}}, \phi_{l_1}^{\omega_{j_1}} \rangle + O\Big(\frac{1}{T}\Big)\Big]\Big[\langle \tilde{\mathfrak{F}}_{-j_1-h_1+j_2+h_2;-\omega_{j_1+h_1}} \phi_{l_2}^{-\omega_{j_2+h_2}}, \phi_{l_1}^{-\omega_{j_1+h_1}} \rangle + O\Big(\frac{1}{T}\Big)\Big]$$
$$+ \Big[\langle \tilde{\mathfrak{F}}_{j_1+j_2+h_2;\omega_{j_1}} \phi_{l_2}^{-\omega_{j_2+h_2}}, \phi_{l_1}^{\omega_{j_1}} \rangle + O\Big(\frac{1}{T}\Big)\Big]\Big[\langle \tilde{\mathfrak{F}}_{-j_1-h_1-j_2;-\omega_{j_1+h_1}} \phi_{l_2}^{\omega_{j_2}}, \phi_{l_1}^{-\omega_{j_1+h_1}} \rangle + O\Big(\frac{1}{T}\Big)\Big]$$
$$+ \frac{2\pi}{T}\langle \tilde{\mathfrak{F}}_{-h_1+h_2;\omega_{j_1},-\omega_{j_1+h_1},-\omega_{j_2}} (\phi_{l_2}^{\omega_{j_2}} \otimes \phi_{l_2}^{\omega_{j_2+h_2}}), \phi_{l_1}^{\omega_{j_1}} \otimes \phi_{l_1}^{\omega_{j_1+h_1}} \rangle + O\Big(\frac{1}{T^2}\Big)$$

and therefore

$$\begin{aligned}
&\mathrm{Cov}\big(\sqrt{T}\gamma_{h_1}^{(T)}(l_1), \sqrt{T}\gamma_{h_2}^{(T)}(l_2)\big) \\
&= \frac{1}{T}\sum_{j_1,j_2=1}^{T}\Bigg\{\frac{2\pi}{T}\frac{\langle \tilde{\mathfrak{F}}_{-h_1+h_2;\omega_{j_1},-\omega_{j_1+h_1},-\omega_{j_2}} (\phi_{l_2}^{\omega_{j_2}} \otimes \phi_{l_2}^{\omega_{j_2+h_2}}), \phi_{l_1}^{\omega_{j_1}} \otimes \phi_{l_1}^{\omega_{j_1+h_1}} \rangle}{(\lambda_{l_1}^{\omega_{j_1}}\lambda_{l_1}^{-\omega_{j_1+h_1}}\lambda_{l_2}^{-\omega_{j_2}}\lambda_{l_2}^{\omega_{j_2+h_2}})^{1/2}} \\
&+ \frac{\langle \tilde{\mathfrak{F}}_{j_1-j_2;\omega_{j_1}} \phi_{l_2}^{\omega_{j_2}}, \phi_{l_1}^{\omega_{j_1}} \rangle\langle \tilde{\mathfrak{F}}_{-j_1-h_1+j_2+h_2;-\omega_{j_1+h_1}} \phi_{l_2}^{-\omega_{j_2+h_2}}, \phi_{l_1}^{-\omega_{j_1+h_1}} \rangle}{(\lambda_{l_1}^{\omega_{j_1}}\lambda_{l_1}^{-\omega_{j_1+h_1}}\lambda_{l_2}^{-\omega_{j_2}}\lambda_{l_2}^{\omega_{j_2+h_2}})^{1/2}} \\
&+ \frac{\langle \tilde{\mathfrak{F}}_{j_1+j_2+h_2;\omega_{j_1}} \phi_{l_2}^{-\omega_{j_2+h_2}}, \phi_{l_1}^{\omega_{j_1}} \rangle\langle \tilde{\mathfrak{F}}_{-j_1-h_1-j_2;-\omega_{j_1+h_1}} \phi_{l_2}^{\omega_{j_2}}, \phi_{l_1}^{-\omega_{j_1+h_1}} \rangle}{(\lambda_{l_1}^{\omega_{j_1}}\lambda_{l_1}^{-\omega_{j_1+h_1}}\lambda_{l_2}^{-\omega_{j_2}}\lambda_{l_2}^{\omega_{j_2+h_2}})^{1/2}} \\
&+ O\Bigg(\frac{1}{T}\Bigg[\frac{\langle \tilde{\mathfrak{F}}_{j_1-j_2;\omega_{j_1}} \phi_{l_2}^{\omega_{j_2}}, \phi_{l_1}^{\omega_{j_1}} \rangle}{(\lambda_{l_1}^{\omega_{j_1}}\lambda_{l_2}^{-\omega_{j_2}})^{1/2}} + \frac{\langle \tilde{\mathfrak{F}}_{-j_1-h_1+j_2+h_2;-\omega_{j_1+h_1}} \phi_{l_2}^{-\omega_{j_2+h_2}}, \phi_{l_1}^{-\omega_{j_1+h_1}} \rangle}{(\lambda_{l_1}^{-\omega_{j_1+h_1}}\lambda_{l_2}^{\omega_{j_2+h_2}})^{1/2}} \\
&+ \frac{\langle \tilde{\mathfrak{F}}_{j_1+j_2+h_2;\omega_{j_1}} \phi_{l_2}^{-\omega_{j_2+h_2}}, \phi_{l_1}^{\omega_{j_1}} \rangle}{(\lambda_{l_1}^{\omega_{j_1}}\lambda_{l_2}^{\omega_{j_2+h_2}})^{1/2}} + \frac{\langle \tilde{\mathfrak{F}}_{-j_1-h_1-j_2;-\omega_{j_1+h_1}} \phi_{l_2}^{\omega_{j_2}}, \phi_{l_1}^{-\omega_{j_1+h_1}} \rangle}{(\lambda_{l_1}^{-\omega_{j_1+h_1}}\lambda_{l_2}^{-\omega_{j_2}})^{1/2}}\Bigg] + \frac{1}{T^2}\Bigg)\Bigg\}.
\end{aligned}$$

The more general expressions are provided in the Online Supplement. □

E Weak convergence

The proof of the distributional properties of $\hat{\gamma}^{(T)}$ as stated in Theorem 4.3 and 4.7 are established in this section. The proof consists of a few steps. First, we derive the properties of $\gamma^{(T)}$, i.e., in which the spectral density operators and the corresponding eigenelements are known. For this, we investigate the distributional properties of the operator

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

Given appropriate rates of the bandwidths, replacing the denominator of (3.1) with, respectively, consistent estimates of the spectral density operators and its eigenvalues will follow from theorems 4.1 and 4.5. The proofs of these theorems can be found in the Online Supplement. For the empirical counterpart of (3.2), it then remains to show that we can replace the projection basis with the sample versions of the eigenfunctions. In particular, we shall below that

$$\frac{1}{\sqrt{T}}\sum_{j=1}^{T}\big\langle D_{\omega_j}^{(T)} \otimes D_{\omega_{j+h}}^{(T)}, \hat{\phi}_{l,T}^{\omega_j} \otimes \hat{\phi}_{l,T}^{\omega_{j+h}}\big\rangle_{HS} \xrightarrow{D} \frac{1}{\sqrt{T}}\sum_{j=1}^{T}\big\langle D_{\omega_j}^{(T)} \otimes D_{\omega_{j+h}}^{(T)}, \phi_{l}^{\omega_j} \otimes \phi_{l}^{\omega_{j+h}}\big\rangle_{HS}.$$

E.1 Weak convergence on the function space

To demonstrate weak convergence of (E.1), we shall make use of a result (Cremers & Kadelka, 1986), which considerably simplifies the verification of the usual tightness condition often invoked in weak convergence proofs. In particular, the following lemma indicates that weak convergence of the functional process will almost directly follow from the weak convergence of the finite dimensional distributions once it is weakly tight in the following sense.

Lemma E.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) $\limsup_{n\to\infty} \mathbb{E}[\|X_n\|_p^p] \leqslant \mathbb{E}[\|X_0\|_p^p]$.

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

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

$$\hat{E}^{(T)}_h = \sqrt{T}\Big(w^{(T)}_h - \mathbb{E}\big[w^{(T)}_h\big]\Big).$$

For $\psi_{ll'} = \psi_l \otimes \psi'_l$ an orthonormal basis of $L^2([0,1]^2, \mathbb{C})$, we can represent the process as

$$\hat{E}^{(T)}_h = \sum_{l,l'=1}^{\infty} \langle \hat{E}^{(T)}_h, \psi_{ll'} \rangle \psi_{ll'}$$

from which it is easily seen the finite-dimensional distributions of the basis coefficients provide a complete characterization of the distributional properties of $\hat{E}^{(T)}_h$. Thus, weak convergence of $(\langle \hat{E}^{(T)}_h, \psi_{ll'} \rangle : l, l' \in \mathbb{N})$ in the sequence space $\ell^2_{\mathbb{C}}$ will imply weak convergence of the process $(\hat{E}^{(T)}_h : T \in \mathbb{N})$. To formalize this, we put the functional $\hat{E}^{(T)}_h$ in duality with $(\hat{E}^{(T)}_h)^* \in L^2([0,1]^2, \mathbb{C})^*$ through the pairing

$$\hat{E}^{(T)}_h(\phi) = \langle \hat{E}^{(T)}_h, \phi \rangle$$

for all $\phi \in L^2([0,1]^2, \mathbb{C})^*$. The conditions of Lemma E.1 can now be verified. 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 E.1. *Under Assumption 4.1 or Assumption 4.3, for all $l_i, l'_i \in \mathbb{N}$, $h_i = 1, \ldots, T-1$, $i = 1, \ldots, k$ and $k \geqslant 3$,*

$$\operatorname{cum}\Big(\hat{E}^{(T)}_{h_1}(\psi_{l_1 l'_1}), \ldots, \hat{E}^{(T)}_{h_k}(\psi_{l_k l'_k})\Big) = o(1) \qquad (T \to \infty).$$

Its proof can be found in Section **??** of the Online Supplement. The second condition of Lemma E.1 will be satisfied if

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

with E_h denoting the limiting process. Using Theorem (E.1) and (E.2) weak convergence of the functional process can now be determined, distinguishing between the real and imaginary parts.

Theorem E.2 (Weak convergence under the null). *Let $(X_t\colon t \in \mathbb{Z})$ be a stochastic process taking values in $H_{\mathbb{R}}$ satisfying Assumption 4.1 with $\ell = 2$. Then,*

$$\big(\Re\hat{E}_{h_i}^{(T)}, \Im\hat{E}_{h_i}^{(T)}\colon i = 1,\dots,k\big) \xrightarrow{D} (\mathcal{R}_{h_i}, \mathcal{I}_{h_i}\colon i = 1,\dots,k), \tag{E.3}$$

where $\mathcal{R}_{h_1}, \mathcal{I}_{h_2}$, $h_1, h_2 \in \{1,\dots,T-1\}$, are jointly Gaussian elements in $L^2([0,1]^2,\mathbb{C})$ with means $\mathbb{E}[\mathcal{R}_{h_1}(\psi_{ll'})] = \mathbb{E}[\mathcal{I}_{h_2}(\psi_{ll'})] = 0$ and covariances

$$\begin{aligned}
\mathrm{Cov}\big(\mathcal{R}_{h_1}(\psi_{l_1l_1'}),&\mathcal{R}_{h_2}(\psi_{l_2l_2'})\big) && \text{(E.4)}\\
&=\mathrm{Cov}\big(\mathcal{I}_{h_1}(\psi_{l_1l_1'}), \mathcal{I}_{h_2}(\psi_{l_2l_2'})\big)\\
&= \frac{1}{4\pi}\int_0^{2\pi} \big\langle \mathcal{F}_\omega(\psi_{l_2}), \psi_{l_1}\big\rangle \big\langle \mathcal{F}_{-\omega-\omega_h}(\psi_{l_2'}), \psi_{l_1}\big\rangle d\omega && \text{(E.5)}\\
&\quad + \frac{1}{4\pi}\int_0^{2\pi} \big\langle \mathcal{F}_\omega(\psi_{l_2'}), \psi_{l_1}\big\rangle \big\langle \mathcal{F}_{-\omega-\omega_h}(\psi_{l_2}), \psi_{l_1'}\big\rangle d\omega\\
&\quad + \frac{1}{4\pi}\int_0^{2\pi}\int_0^{2\pi} \big\langle \mathcal{F}_{\omega,-\omega-\omega_h,-\omega'}(\psi_{l_2l_2'}), \psi_{l_1l_1'}\big\rangle d\omega d\omega'
\end{aligned}$$

for all $h_1 = h_2$ and l_1, l_1', l_2, l_2', and 0 otherwise. In addition,

$$\mathrm{Cov}(\mathcal{R}_{h_1}(\psi_{l_1l_1'}), \mathcal{I}_{h_2}(\psi_{l_2l_2'})) = 0$$

uniformly in h_1, h_2 and l_1, l_1', l_2, l_2'.

Proof. The covariance structure is the more general version of Theorem 4.2 as derived in Appendix D.1 while the convergence of the finite–dimensional distributions from Theorem E.1. It then remains to verify that the condition (ii) of Lemma E.1 is satisfied. This follows from the covariance structure since

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

This completes the proof. □

Under the alternative, a similar result is obtained.

Theorem E.3 (Weak convergence under the alternative). *Let $(X_t\colon t \in \mathbb{Z})$ be a stochastic process taking values in $H_{\mathbb{R}}$ satisfying Assumption 4.3 with $\ell = 2$. Then,*

$$\big(\Re\hat{E}_{h_i}^{(T)}, \Im\hat{E}_{h_i}^{(T)}\colon i = 1,\dots,k\big) \xrightarrow{d} (\mathcal{R}_{h_i}, \mathcal{I}_{h_i}\colon i = 1,\dots,k), \tag{E.6}$$

where $\mathcal{R}_{h_1}, \mathcal{I}_{h_2}$, $h_1, h_2 \in \{1, \ldots, T-1\}$, are jointly Gaussian elements in $L^2([0,1]^2, \mathbb{C})$ with means $\mathbb{E}[\mathcal{R}_{h_1}(\psi_{ll'})] = \mathbb{E}[\mathcal{I}_{h_2}(\psi_{ll'})] = 0$ and covariance structure

1. $\mathrm{Cov}(\mathcal{R}_{h_1}(\psi_{l_1 l_1'}), \mathcal{R}_{h_2}(\psi_{l_2 l_2'})) =$
$$\frac{1}{4}\big[\Upsilon_{h_1,h_2}(\psi_{l_1 l_1'\, l_2 l_2'}) + \acute{\Upsilon}_{h_1,h_2}(\psi_{l_1 l_1'\, l_2 l_2'}) + \grave{\Upsilon}_{h_1,h_2}(\psi_{l_1 l_1'\, l_2 l_2'}) + \bar{\Upsilon}_{h_1,h_2}(\psi_{l_1 l_1'\, l_2 l_2'})\big]$$
2. $\mathrm{Cov}(\mathcal{R}_{h_1}(\psi_{l_1 l_1'}), \mathcal{I}_{h_2}(\psi_{l_2 l_2'})) =$
$$\frac{1}{4\mathrm{i}}\big[\Upsilon_{h_1,h_2}(\psi_{l_1 l_1'\, l_2 l_2'}) - \acute{\Upsilon}_{h_1,h_2}(\psi_{l_1 l_1'\, l_2 l_2'}) + \grave{\Upsilon}_{h_1,h_2}(\psi_{l_1 l_1'\, l_2 l_2'}) - \bar{\Upsilon}_{h_1,h_2}(\psi_{l_1 l_1'\, l_2 l_2'})\big]$$
3. $\mathrm{Cov}(\mathcal{I}_{h_1}(\psi_{l_1 l_1'}), \mathcal{I}_{h_2}(\psi_{l_2 l_2'})) =$
$$\frac{1}{4}\big[\Upsilon_{h_1,h_2}(\psi_{l_1 l_1'\, l_2 l_2'}) - \acute{\Upsilon}_{h_1,h_2}(\psi_{l_1 l_1'\, l_2 l_2'}) - \grave{\Upsilon}_{h_1,h_2}(\psi_{l_1 l_1'\, l_2 l_2'}) + \bar{\Upsilon}_{h_1,h_2}(\psi_{l_1 l_1'\, l_2 l_2'})\big]$$

for all h_1, h_2 and l_1, l_1', l_2, l_2', and where Υ_{h_1,h_2}, $\acute{\Upsilon}_{h_1,h_2}$, $\grave{\Upsilon}_{h_1,h_2}$ and $\bar{\Upsilon}_{h_1,h_2}$ are given in (**??**)–(**??**).

Proof. The covariance structure is fully derived in the Online Supplement while the convergence of the finite–dimensional distributions follows again from Theorem E.1. Condition (ii) of Lemma E.1 is satisfied since

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

which completes the proof. □

E.2 Replace projection basis with estimates

We now focus on replacing the projection basis with estimates of the eigenfunctions of the spectral density operators. It can be shown (Mas & Menneteau, 2003) that for appropriate 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 $\Pi_{l,T}^{\omega}$. However, the estimated eigenfunctions are not unique and are 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, we therefore first have to consider the issue of rotation. More specifically, when we estimate $\phi_l^{\omega_j}$, 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. We can therefore not guarantee that an estimated version $\hat{z}_l\hat{\phi}_{l,T}^{\omega_j}$ is close to the true eigenfunction $\hat{\phi}_l^{\omega_j}$. It is therefore essential that our test statistic is invariant to this rotation. To show this, write

$$\Psi_h(l,j) = \langle D_{\omega_j}^{(T)}, \hat{\phi}_l^{\omega_j}\rangle\overline{\langle D_{\omega_{j+h}}^{(T)}, \hat{\phi}_l^{\omega_{j+h}}\rangle}$$

and let $\mathbf{\Psi}(h,L) = \mathrm{vec}(\Psi_h(l,j))$ for the stacked vector of dimension $L(T-h) \times 1$, then (3.5) can be written as

$$\hat{\boldsymbol{\beta}}_h^{(T)} = e^{\top}\mathbf{\Psi}(h,L)$$

Let us then construct the matrix

$$Z_L^j = \begin{pmatrix} \hat{z}_1^j & \cdots & & \\ \vdots & \hat{z}_2^j & & \\ & & \ddots & \\ & & & \hat{z}_L^j \end{pmatrix}$$

and the block diagonal matrix $Z_L^{1:T} = diag(Z_L^j : j = 1, \dots, T)$ and additionally the kronecker product

$$\boldsymbol{Z}(h, L) = Z_L^{1:T-h} \otimes Z_L^{h:T}.$$

This object is then of dimension $L(T-h) \times L(T-h)$ and is diagonal with elements $\hat{z}_l^j \overline{\hat{z}}_l^{j+h}, l = 1, \dots, L; j = 1, \dots, T-h$. If we rotate the eigenfunctions on the unit circle, we obtain a version

$$\hat{\boldsymbol{\beta}}_h^{(T)} = e^\top \boldsymbol{Z}(h, L)\boldsymbol{\Psi}(h, L),$$

in which case we can write

$$\sqrt{T}\mathcal{Z}_{L,M}\hat{\boldsymbol{b}}_M^{(T)} = \sqrt{T}\big(\Re\hat{\boldsymbol{\beta}}_{h_1}^{(T)}, \dots, \Im\hat{\boldsymbol{\beta}}_{h_M}^{(T)}, \Im\hat{\boldsymbol{\beta}}_{h_1}^{(T)}, \dots, \Im\hat{\boldsymbol{\beta}}_{h_M}^{(T)}\big)^\top,$$

where the block diagonal matrix is given by

$$\mathcal{Z}_L = diag(\boldsymbol{\Re Z}(h_1, L), \dots, \boldsymbol{\Re Z}(h_M, L), \boldsymbol{\Im Z}(h_1, L), \dots \boldsymbol{\Im Z}(h_M, L))^\top.$$

However this rotation then also implies that $\hat{\Sigma}_M$ becomes $\mathcal{Z}_{L,M}\hat{\Sigma}_M\mathcal{Z}_{L,M}^\top$ and therefore

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

showing our statistic does not depend on rotation of the estimated eigenfunctions. In the rest of the proof, we can therefore focus on estimates $\hat{\phi}_{l,T}^{\omega_{j+h}}$ and $\hat{\phi}_{l,T}^{\omega_{j+h}}$ and ignore their respective unknown rotations $\hat{z}_l^j$ and $\overline{\hat{z}}_l^{j+h}$.

Given the estimates of the spectral density operators are consistent, note that the consistency of the estimated eigenprojectors together with continuity of the tensor product imply that

$$\sup_\omega \|\widehat{\Pi}_{l,T,h}^\omega - \Pi_{l,h}^\omega\|_2^2 \xrightarrow{p} 0 \qquad T \to \infty, \tag{E.7}$$

where the projectors are given by $\Pi_{l,h}^\omega := \phi_l^{\omega_j} \otimes \phi_l^{\omega_{j+h}}$. In the previous section, we moreover showed

$$\frac{1}{\sqrt{T}}\sum_{j=1}^T D_{\omega_j}^{(T)} \otimes D_{\omega_{j+h}}^{(T)} \xrightarrow{\mathcal{D}} w_h^{(T)}$$

Using Portmanteau's Theorem, this means there exists a large M_δ s.t.

$$\limsup_{T\to\infty} \mathbb{P}\Big(\Big\|\frac{1}{\sqrt{T}}\sum_{j=1}^T D_{\omega_j}^{(T)} \otimes D_{\omega_{j+h}}^{(T)}\Big\|_2 \geqslant M_\delta\Big) \leqslant \mathbb{P}\Big(\|w_h^{(T)}\|_2 \geqslant M_\delta\Big) \leqslant \delta, \delta > 0.$$

The Cauchy-Schwarz inequality and the definition of the Hilbert-Schmidt inner product yield

$$\limsup_{T\to\infty} \mathbb{P}\Big(\Big|\frac{1}{\sqrt{T}} \sum_{j=1}^{T} \big\langle D^{(T)}_{\omega_j} \otimes D^{(T)}_{\omega_{j+h}}, \hat{\phi}^{\omega_j}_{l,T} \otimes \hat{\phi}^{\omega_{j+h}}_{l,T} - \phi^{\omega_j}_{l} \otimes \phi^{\omega_{j+h}}_{l} \big\rangle_{HS}\Big| > \epsilon\Big)$$
$$\leqslant \limsup_{T\to\infty} \mathbb{P}\Big(\int\int \Big|\frac{1}{\sqrt{T}} \sum_{j=1}^{T} D^{(T)}_{\omega_j}(\tau) D^{(T)}_{-\omega_{j+h}}(\tau')\big[\hat{\phi}^{-\omega_j}_{l,T}(\tau)\hat{\phi}^{\omega_{j+h}}_{l,T}(\tau') - \phi^{-\omega_j}_{l}(\tau)\phi^{\omega_{j+h}}_{l}(\tau')\big]\Big| d\tau d\tau' > \epsilon\Big)$$

and using (E.7)

$$\leqslant \limsup_{T\to\infty} \mathbb{P}\Big(\sup_{\omega} \|\hat{\phi}^{\omega}_{l,T} \otimes \hat{\phi}^{\omega_h}_{l,T} - \phi^{\omega}_{l} \otimes \phi^{\omega_h}_{l}\|_2 \Big\|\frac{1}{\sqrt{T}} \sum_{j=1}^{T} D^{(T)}_{\omega_j} \otimes D^{(T)}_{\omega_{j+h}}\Big\|_2 > \epsilon\Big)$$
$$\leqslant \limsup_{T\to\infty} \mathbb{P}(\sup_{\omega} \|\hat{\phi}^{\omega}_{l,T} \otimes \hat{\phi}^{\omega_h}_{l,T} - \phi^{\omega}_{l} \otimes \phi^{\omega_h}\|_2 > \epsilon/M_\delta\Big) + \limsup_{T\to\infty} \mathbb{P}\Big(\Big\|\frac{1}{\sqrt{T}} \sum_{j=1}^{T} D^{(T)}_{\omega_j} \otimes D^{(T)}_{\omega_{j+h}}\Big\|_2 \geqslant M_\delta\Big)$$
$$\leqslant 0 + \delta$$

for every $\delta > 0$. Given the bandwidth condition (Assumption 4.2) is satisfied, we therefore finally obtain

$$\frac{1}{\sqrt{T}} \sum_{j=1}^{T} \big\langle D^{(T)}_{\omega_j} \otimes D^{(T)}_{\omega_{j+h}}, \hat{\phi}^{\omega_{j+h}}_{l} \otimes \hat{\phi}^{\omega_j}_{l} \big\rangle_{HS} \xrightarrow{\mathcal{D}} \frac{1}{\sqrt{T}} \sum_{j=1}^{T} \big\langle D^{(T)}_{\omega_j} \otimes D^{(T)}_{\omega_{j+h}}, \phi^{\omega_{j+h}}_{l} \otimes \phi^{\omega_j}_{l} \big\rangle_{HS} \tag{E.8}$$

as $T \to \infty$.

References

Antoniadis, A. & T. Sapatinas (2003). Wavelet methods for continuous time prediction using Hilbert-valued autoregressive processes. *Journal of Multivariate Analysis* **87**, 133–158.

Aue, A., Dubart Nourinho, D. & S. Hörmann (2015). On the prediction of stationary functional time series. *Journal of the American Statistical Association* **110**, 378–392.

Aue, A. & L. Horváth (2013). Structural breaks in time series. *Journal of Time Series Analysis* **34**, 1–16.

Aue, A., Rice, G. & O. Sönmez (2017). Detecting and dating structural breaks in functional data without dimension reduction. *Journal of the Royal Statistical Society, Series B*, forthcoming.

Aue, A. & A. van Delft (2017). Online supplement to "Testing for stationarity of functional time series in the frequency domain".

Bandyopadhyay, S. & S. Subba Rao (2017). A test for stationarity for irregularly spaced spatial data. *Journal of the Royal Statistical Society, Series B* **79**, 95–123.

Bandyopadhyay, S., Jentsch, C. & S. Subba Rao (2017). A spectral domain test for stationarity of spatio-temporal data. *Journal of Time Series Analysis* **38**, 326–351.

Besse, P., Cardot, H. & D. Stephenson (2000). Autoregressive forecasting of some functional climatic variations. *Scandinavian Journal of Statistics* **27**, 673–687.

Bosq, D. (2000). *Linear Processes in Function Spaces.* Springer-Verlag, New York.

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

Brillinger, D. & M. Rosenblatt (1967). Asymptotic theory of estimates of k-th order spectra. In *Spectral Analysis of Time Series (Ed. B. Harris)*, Wiley, New York, pages 153–188.

Cremers, H. & D. Kadelka (1986). On weak convergence of integral functions of stochastic processes with applications to processes taking paths in L_p^E. *Stochastic Processes and their Applications* **21**, 305–317.

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

Dette, H., Preuß, P. & M. Vetter (2011). A measure of stationarity in locally stationary processes with applications to testing. *Journal of the American Statistical Association* **106**, 1113–1124.

Dwivedi, Y. & Subba Rao, S. (2011). A test for second-order stationarity of a time series based on the discrete Fourier transform. *Journal of Time Series Analysis* **32**, 68–91.

Ferraty, F. & Vieu, P. (2010). *Nonparametric Functional Data Analysis.* Springer-Verlag, New York.

Hörmann, S., Kidziński, Ł. & M. Hallin (2015). Dynamic functional principal components. *Journal of the Royal Statistical Society, Series B* **77**, 319–348.

Hörmann, S. & P. Kokoszka (2010). Weakly dependent functional data. *The Annals of Statistics* **38**, 1845–1884.

Horváth, L. & P. Kokoszka (2012). *Inference for Functional Data with Applications.* Springer-Verlag, New York.

Horváth, L. Kokoszka, P. & G. Rice (2014). Testing stationarity of functional time series. *Journal of Econometrics* **179**, 66–82.

Hsing, T. & R. Eubank (2015). *Theoretical Foundations of Functional Data Analysis, with an Introduction to Linear Operators.* Wiley, New York.

Jentsch, C. & S. Subba Rao (2015). A test for second order stationarity of a multivariate time series. *Journal of Econometrics* **185**, 124–161.

Jin, L., Wang, S. & H. Wang (2015). A new non-parametric stationarity test of time series in the time domain. *Journal of the Royal Statistical Society, Series B* **77**, 893–922.

Lee, J. & S. Subba Rao (2016). A note on general quadratic forms of nonstationary stochastic processes. Technical Report, Texas A&M University.

Li, Y. & T. Hsing (2010). Uniform convergence rates for nonparametric regression and principal component analysis in functional/longitudinal data. *The Annals of Statistics* **38**, 3321–3351.

Mas, A. & L. Menneteau (2003). Perturbation approach applied to the asymptotic study of random operators. In: *Hoffmann-Jørgensen et al. (eds.). High dimensional probability III.* Birkhäuser, Boston, pages 127–134.

Nason, G. (2013). A test for second-order stationarity and approximate confidence intervals for localized autocovariances for locally stationary time series. *Journal of the Royal Statistical Society, Series B* **75**, 879–904.

Panaretos, V. & S. Tavakoli (2013). Fourier analysis of stationary time series in function space. *The Annals of Statistics* **41**, 568–603.

Paparoditis, E. (2009). Testing temporal constancy of the spectral structure of a time series. *Bernoulli* **15**, 1190–1221.

Preuß, P., Vetter, M. & H. Dette (2013). A test for stationarity based on empirical processes. *Bernoulli* **19**, 2715–2749.

Priestley, M.B. & T. Subba Rao (1969). A test for non-stationarity of time-series. *Journal of the Royal Statistical Society, Series B* **31**, 140–149.

Ramsay, J.O. & B.W. Silverman (2005). *Functional Data Analysis (2nd ed.).* Springer-Verlag, New York.

Subba Rao, S. (2016). Orthogonal samples for estimators in time series. Preprint available at `https://arxiv.org/abs/1611.00398`.

Van Delft, A. & M. Eichler (2016). Locally stationary functional time series. Preprint, available at `https://arxiv.org/pdf/1602.05125v2.pdf`.

Von Sachs, R. & M.H. Neumann (1999). A wavelet-based test for stationarity. *Journal of Time Series Analysis* **21**, 597–613.