A test for stationarity for spatio-temporal data

Soutir Bandyopadhyay[*], Carsten Jentsch[†] and Suhasini Subba Rao[‡]

January 27, 2016

Abstract

Many random phenomena in the environmental and geophysical sciences are functions of both space and time; these are usually called spatio-temporal processes. Typically, the spatio-temporal process is observed over discrete equidistant time and at irregularly spaced locations in space. One important aim is to develop statistical models based on what is observed. While doing so a commonly used assumption is that the underlying spatial-temporal process is stationary. If this assumption does not hold, then either the mean or the covariance function is misspecified. This can, for example, lead to inaccurate predictions. In this paper we propose a test for spatio-temporal stationarity. The test is based on the dichotomy that Fourier transforms of stochastic processes are near uncorrelated if the process is second order stationarity but correlated if the process is second order nonstationary. Using this as motivation, a Discrete Fourier transform for spatio-temporal data over discrete equidistant times but on irregularly spaced spatial locations is defined. Two statistics which measure the degree of correlation in the Discrete Fourier transforms are proposed. These statistics are used to test for spatio-temporal stationarity. It is shown that the same statistics can also be adapted to test for the one-way stationarity (either spatial or temporal stationarity). The proposed methodology is illustrated with a small simulation study.

Key words and phrases: Fourier Transforms; Irregular sampling; Nonstationarity; Stationary random fields; Spectral density; Orthogonal samples.

[*]Department of Mathematics, Lehigh University, Bethlehem, PA, U.S.A. sob210@lehigh.edu
[†]Department of Economics, University of Mannheim, Mannheim, Germany. cjentsch@staff.mail.uni-mannheim.de
[‡]Department of Statistics, Texas A&M University, College Station, TX, U.S.A. suhasini@stat.tamu.edu

1 Introduction

Several environmental and geophysical phenomena, such as tropospheric ozone and precipitation levels, are random quantities depending on both space and time. Since, in practice, it is only possible to observe the process on a finite number of locations in space, $\{\boldsymbol{s}_j\}_{j=1}^n$ and typically over discrete equidistant time $t = 1, \ldots, T$, one aim in the geosciences is to develop statistical models based on what is observed. Typically this is done by fitting a parametric space-time covariance function defined on $\{Z_t(\boldsymbol{s}); \boldsymbol{s} \in \mathbb{R}^d, t \in \mathbb{Z}\}$ to the data. Such models can, then, be used for prediction and forecasting at unobserved locations; see Gneiting et al. [2006] and Sherman [2010] for an extensive survey on space-time models. In this context, an assumption that is often used is that the underlying spatio-temporal process $\{Z_t(\boldsymbol{s}); \boldsymbol{s} \in \mathbb{R}^d, t \in \mathbb{Z}\}$ is stationary, in the sense that $\mathrm{E}[Z_t(\boldsymbol{s})] = \mu$ and $\mathrm{cov}[Z_t(\boldsymbol{s}_1), Z_\tau(\boldsymbol{s}_2)] = \kappa_{\tau-t}(\boldsymbol{s}_2 - \boldsymbol{s}_1)$. If this assumption does not hold, then either the mean or the covariance function is misspecified which, for example, can lead to inaccurate predictions. Therefore, in order to understand the underlying structure of the spatio-temporal process correctly we should test for second order stationarity of the spatio-temporal process first. Furthermore, given that often the size of the data sets are extremely large, the test should be computationally feasible. The aim of this paper is to address these issues.

Before we describe the proposed procedure, we start by surveying some of the tests for stationarity that exist in the literature. One of the earliest tests for temporal stationarity is given in Priestley and Subba Rao [1969]. More recently, several tests for temporal stationarity have been proposed; these include von Sachs and Neumann [1999], Paparoditis [2009], Paparoditis [2010], Dette et al. [2011], Dwivedi and Subba Rao [2011], Jentsch [2012], Nason [2013], Lei et al. [2015], Jentsch and Subba Rao [2015], Cho [2014] and Puchstein and Preuss [2016].

For spatial data, Fuentes [2006] generalizes the test proposed in Priestley and Subba Rao [1969] to spatial data defined on a grid and Epharty et al. [2001] proposes a test for spatio-temporal stationarity for data defined on a spatio-temporal grid. However, if the spatial data is defined on irregular locations (typically, a more realistic scenario), then there exists only a few number of tests. As far as we are aware, the first test for spatio-temporal stationarity, where the spatial component of the data is observed at irregular locations is proposed in Jun and Genton [2012]. More recently, Bandyopadhyay and Subba Rao [2016] propose a test for spatial stationarity where the data is observed at irregular locations.

In this paper we develop a test for spatio-temporal stationarity, where time is defined on $\mathbb{Z}$ and the locations are irregular on $\mathbb{R}^d$. Our procedure is heavily motivated by the tests in Epharty et al. [2001], Dwivedi and Subba Rao [2011], Jentsch and Subba Rao [2015] and Bandyopadhyay and Subba Rao [2016], which use a Fourier transform of the data to discriminate between the stationary and nonstationary behavior. To motivate our approach let us consider the Cramér representation of a stationary stochastic process, which states that a second order stationary stochastic process, $\{Z_t(\boldsymbol{s}), \boldsymbol{s} \in \mathbb{R}^d, t \in \mathbb{Z}\}$ can always be represented as

$$Z_t(\boldsymbol{s}) = \frac{1}{2\pi}\int_0^{2\pi}\int_{\mathbb{R}^d} \exp(it\omega)\exp(i\boldsymbol{s}'\boldsymbol{\Omega})dZ(\boldsymbol{\Omega},\omega), \tag{1}$$

where, $Z(\boldsymbol{\Omega},\omega)$ is a stochastic process with orthogonal increments, i.e., $\mathrm{E}[dZ(\boldsymbol{\Omega}_1,\omega_2)\overline{dZ(\boldsymbol{\Omega}_3,\omega_4)}] = 0$ if $\boldsymbol{\Omega}_1 \neq \boldsymbol{\Omega}_3$ or $\omega_2 \neq \omega_4$ and $\mathrm{E}[|dZ(\boldsymbol{\Omega}_1,\omega_2)|^2] = dF(\boldsymbol{\Omega},\omega) = f(\boldsymbol{\Omega},\omega)d\boldsymbol{\Omega}d\omega$, where f denotes the spectral density (and the second equality only holds if the derivative of F exists); see Subba Rao and Terdik [2016]. On the other hand, if the increments are correlated, then the process is not second order stationary (see for example, Gladyshev [1963], Goodman [1965], Yaglom [1987] Lii and Rosenblatt [2002], Hindberg and Olhede [2010], Gorrostieta et al. [2016]). Furthermore, the increment process yields information about the stationarity of the process in particular domains. For example, suppose the process is spatially stationary (but not necessarily temporally stationary), then $\mathrm{E}[dZ(\boldsymbol{\Omega}_1,\omega_2)\overline{dZ(\boldsymbol{\Omega}_3,\omega_2)}] = 0$ if $\boldsymbol{\Omega}_1 \neq \boldsymbol{\Omega}_3$. Conversely, if the process is temporally stationary but not spatially stationary, then $\mathrm{E}[dZ(\boldsymbol{\Omega}_1,\omega_2)\overline{dZ(\boldsymbol{\Omega}_1,\omega_4)}] = 0$ if $\omega_2 \neq \omega_4$.

Of course in practice the increment process is unobserved. However, in time series analysis the Discrete Fourier transform (DFT) of a time series is considered as an estimator of the increments in the increment process and shares many of its properties. In particular, the Discrete Fourier transform of a stationary time series is a 'near uncorrelated' transformation, thus mirroring the properties of the increment process. In Dwivedi and Subba Rao [2011] and Jentsch and Subba Rao [2015] we use the Discrete Fourier transform to test for stationarity. On the other hand, the Fourier transform for spatial data defined on irregular locations is not uniquely defined. However, Matsuda and Yajima [2009] and Bandyopadhyay and Lahiri [2009] define a Fourier transform on spatial data with irregular locations which can be shown to share similar properties as the increment process when the locations are uniformly distributed. In Bandyopadhyay and Subba Rao [2016] we exploit this property to test for spatial stationarity. In this paper we combine both these transformations to define a Discrete Fourier transform for spatio-temporal data that is defined over discrete time but on irregular spatial locations. We show that this space-time Discrete Fourier transform satisfies many of the properties of (1); in particular under stationarity the space-time DFT is asymptotically uncorrelated, whereas under nonstationarity this property does not hold. In this paper we use this dichotomy to define tests for stationarity for spatio-temporal processes.

In Section 2.1 we review the test for temporal stationarity proposed in Dwivedi and Subba Rao [2011] and Jentsch and Subba Rao [2015]. In Section 2.2 we review the test for spatial stationarity proposed in Bandyopadhyay and Subba Rao [2016]. We note that there are some fundamental differences between the testing methodology over time compared to the testing methodology over space. The first is that over discrete time the Fourier transform can only be defined over a compact support, whereas the Fourier transform on space can be defined over $\mathbb{R}^d$ (see the range of the integrals in (1)). This leads to significant differences in the way that the test statistics can be defined. Furthermore, both the test over time and the test over space involve variances which need to be estimated. In the test for stationarity proposed in Jentsch and Subba Rao [2015] we used the stationary bootstrap to estimate the variance, however using a block-type bootstrap for the spatial stationarity test was computationally too intensive. Instead we used the method of orthogonal samples to estimate the variance, which led us to a computationally feasible test statistic. This method can also be applied to testing for temporal stationarity, and in Section 2.3 we use the same method to estimate the variance in the test for temporal stationarity. The results in this section motivate our proposed procedure in the later sections.

In Section 3 we turn to the spatio-temporal data. We define a Fourier transform (to reduce notation we call it a "DFT"), which is over irregular locations in space, but for equidistant discrete time. We obtain the correlation properties of the DFTs in the case of (i) spatial and temporal stationarity, (ii) spatial stationarity (but not necessarily temporally stationary), (iii) temporal stationarity (but not necessarily spatially stationary) and (iv) both temporal and spatial nonstationarity. We show that each case has its own specific characterization in terms of the DFTs. In Section 4 we use the differing behaviors to construct the test statistics. Similar to both the stationarity test over space and the stationarity test over time, the test here involves unknown variances, which are estimated using orthogonal samples. This means the test statistic can be calculated in $O(n^2 T \log T)$ computing operations. In Section 5 we apply the methodology for testing one-way stationarity (stationary in one domain but not necessarily stationary on the other domain). Some results relating to the test statistics are in Section 6. Our proposed methods are illustrated with simulations in Section 7. A rough outline of the proofs are given in the supplementary material.

2 Using the DFT to test for stationarity over time or space

Our test for spatio-temporal stationarity is based on some of the ideas used to develop the temporal and spatial tests in Dwivedi and Subba Rao [2011], Jentsch and Subba Rao [2015] and Bandyopadhyay and Subba Rao [2016]. Therefore in Sections 2.1 and 2.2 we review some pertinent features of these tests. In Section 2.3 we apply some of the methods discussed in Section 2.2 to test for temporal stationarity.

2.1 Testing for temporal stationarity

Let us suppose that $\{X_t\}$ is a stationary time series where $c_h = \mathrm{cov}[X_t, X_{t+h}]$ and $\sum_h |hc_h| < \infty$. Given that we observe $\{X_t\}_{t=1}^T$, we define the DFT of a time series $\{X_t\}_{t=1}^T$ as $J_T(\omega_k) = \frac{1}{\sqrt{2\pi T}} \sum_{t=1}^T X_t e^{it\omega_k}$, where $\omega_k = 2\pi k/T$ are the so called Fourier frequencies. Dwivedi and Subba Rao [2011] and Jentsch and Subba Rao [2015] exploit the property that under stationarity (and the short memory condition stated above) the DFT $\{J_T(\omega_k)\}$ is a 'near uncorrelated sequence' whose variance is approximately equal to the spectral function

$$f(\omega_k) = \frac{1}{2\pi} \sum_{h \in \mathbb{Z}} c_h \exp(-ih\omega_k).$$

We now briefly describe the procedure proposed in Jentsch and Subba Rao [2015] to test for stationarity of a multivariate time series. To understand the pertinent features of the test, we focus on the univariate case. Jentsch and Subba Rao [2015] estimates the spectral density from the data (we denote the estimator, by smoothing the periodogram, as $\widehat{f}_T$), use this to 'standardize' the DFT and define the estimator of the covariance between the DFTs at 'lag' r as

$$\widehat{C}_T(r, \ell) = \frac{1}{T} \sum_{k=1}^T \exp(i\ell\omega_k) \frac{J_T(\omega_k)\overline{J_T(\omega_{k+r})}}{\sqrt{\widehat{f}_T(\omega_k)\widehat{f}_T(\omega_{k+r})}}. \quad (2)$$

If we set $\ell = 0$, then $\{\widehat{C}_T(r, 0)\}_r$ can be viewed as the sample 'autocovariance' of the sequence $\{J_T(\omega_k)/\widehat{f}_T(\omega_k)^{1/2}\}_{k=1}^T$ over frequency. $\widehat{C}_T(r, 0)$ was used as the basis of the test statistic. Jentsch and Subba Rao [2015] showed that the approximate 'variance' (in terms of the limiting distribution) of $\Re\widehat{C}_T(r, \ell)$ and $\Im\widehat{C}_T(r, \ell)$ (where $\Re x$ and $\Im x$ denote the real and imaginary parts of x) is $v_\ell(\omega_r)$, where

$$v_\ell(\omega_r) = \frac{1}{2}[1 + \delta_{\ell,0} + \kappa_\ell(\omega_r)]$$

with

$$\kappa_\ell(\omega_r) = \frac{1}{2\pi} \int_0^{2\pi} \int_0^{2\pi} \frac{f_4(\lambda_1 + \omega_r, \lambda_2, -\lambda_2 - \omega_r)}{\sqrt{f(\lambda_1)f(\lambda_1 + \omega_r)f(\lambda_2)f(\lambda_2 + \omega_r)}} \exp[i\ell(\lambda_1 - \lambda_2)] d\lambda_1 d\lambda_2 \quad (3)$$

and f_4 is the fourth order spectral density, which is defined as

$$f_4(\omega_1, \omega_2, \omega_3) = \frac{1}{(2\pi)^3} \sum_{h_1,h_2,h_3 \in \mathbb{Z}} \kappa_{h_1,h_2,h_3} \exp(-ih_1\omega_1 - ih_2\omega_2 - ih_3\omega_3),$$

where $\kappa_{h_1,h_2,h_3} = \text{cum}[X_0, X_{h_1}, X_{h_2}, X_{h_3}]$. Moreover, for fixed ℓ and m and under suitable mixing conditions we have

$$\sqrt{T}\left[\Re\widehat{C}_T(1,\ell), \Im\widehat{C}_T(1,\ell), \ldots, \Re\widehat{C}_T(m,\ell), \Im\widehat{C}_T(m,\ell)\right] \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, v_\ell(0) I_{2m}\right), \tag{4}$$

as $T \to \infty$, where I_{2m} denotes the identity matrix of order $2m$. Note that $v_\ell(\omega_r) \to v_\ell(0)$ as $T \to \infty$. We observe that in the case where the time series is stationary and Gaussian, we have $\kappa_\ell(0) = 0$ and $v_\ell(0) = \frac{1}{2}(1+\delta_{0,\ell})$; consequently, $\{\widehat{C}_T(r,\ell\}$ is pivotal (does not depend on any nuisance parameters). In contrast, if the time series is stationary but non-Gaussian, the term $\kappa_\ell(0)$ does not vanish (indeed assuming Gaussianity when the process is not Gaussian can lead to inflated type I errors in the test statistic defined below); compare Section 6.2 in Jentsch and Subba Rao [2015].

Based on (4) for a fixed ℓ, the test statistic is then defined as

$$\widetilde{\mathbf{T}}_m = T \sum_{r=1}^{m} \frac{|\widehat{C}_T(r,\ell)|^2}{v_\ell(0)}, \tag{5}$$

which under the null of stationarity, asymptotically has a chi-squared distribution with $2m$ degrees of freedom. In practice $v_\ell(0)$ is unknown. Therefore, Jentsch and Subba Rao [2015] uses the stationary bootstrap, proposed in Politis and Romano [1994], to estimate $v_\ell(0)$ (actually they estimate $v_\ell(\omega_r)$). In Section 2.3 of this paper we describe an alternative method for estimating $v_\ell(0)$, which is a computationally much faster method. Note that in Jentsch and Subba Rao [2015] a more general statistic based on $\{\widehat{C}_T(r,\ell); r = 1,\ldots,m, \ell = 1,\ldots,L\}$ is proposed.

To understand how $\widehat{C}_T(r,\ell)$ behaves in the case the process is nonstationary we assume that the time series 'evolves' slowly over time (a notion that was first introduced in Priestley [1965]). To obtain the asymptotic limit of $\widehat{C}_T(r,\ell)$ we use the rescaling device introduced in Dahlhaus [1997], where it was used to develop and study the class of locally stationary time series. More precisely, we consider the class of locally stationary processes $\{X_{t,T}\}$, whose covariance structure changes slowly over time such that there exist smooth functions $\{\kappa_{r;\cdot}\}_r$ which can approximate the time-varying covariance, i.e., $|\text{cov}(X_{t,T}, X_{t+h,T}) - \kappa_{h;\frac{t}{T}}| \leq T^{-1}\rho_h$, where $\{\rho_h\}$ is such that $\sum_h |h\rho_h| < \infty$ (see Dahlhaus [2012]). Further, we define the time-dependent spectral density $F_u(\omega) = \frac{1}{2\pi}\sum_{h\in\mathbb{Z}} \kappa_{h;u} e^{-ih\omega}$. Under this set-up we have $\widehat{C}_T(r,\ell) \xrightarrow{\mathcal{P}} A(r,\ell)$, where

$$A(r,\ell) = \frac{1}{2\pi}\int_0^{2\pi}\int_0^1 \frac{F_u(\omega)}{f(\omega)} \exp(-i2\pi r u)\exp(i\ell\omega)\,du\,d\omega \tag{6}$$

and $f(\omega) = \int_0^1 F_u(\omega)du$.

2.2 Testing for spatial stationarity

In Bandyopadhyay and Subba Rao [2016] our objective is to test for spatial stationarity for a spatial random process $\{Z(\boldsymbol{s}); \boldsymbol{s} \in \mathbb{R}^d\}$, observed only at a finite number of irregularly spaced locations, denoted as $\{\boldsymbol{s}_j\}_{j=1}^n$, in the region $[-\lambda/2, \lambda/2]^d$, i.e., we observe $\{(\boldsymbol{s}_j, Z(\boldsymbol{s}_j)); j = 1, \ldots, n\}$. Suppose $\{Z(\boldsymbol{s}); \boldsymbol{s} \in \mathbb{R}^d\}$ is spatially stationary and denote $c(\boldsymbol{v}) = \mathrm{cov}[Z(\boldsymbol{s}), Z(\boldsymbol{s} + \boldsymbol{v})]$. Analogous to the stationarity test for a time series described in Section 2.1 we test for spatial stationarity by checking for uncorrelatedness of the Fourier transforms. We note that the DFT of a discrete time time series (as described above) is a linear one-to-one transformation between the time series in the time domain to the frequency domain that can be easily inverted using the inverse DFT. On the other hand, when the locations are irregularly spaced, i.e. they are not on an equidistant grid on $[-\lambda/2, \lambda/2]^d$, there is no unique way to define the Fourier transform. Instead to test for stationarity, we use a suitable Fourier transform for irregularly sampled data which retains the near uncorrelated property. More precisely, we define the Fourier transform as $J_n(\boldsymbol{\Omega}) = \frac{\lambda^{d/2}}{n} \sum_{j=1}^n Z(\boldsymbol{s}_j) \exp(i\boldsymbol{s}_j'\boldsymbol{\Omega})$ where $\boldsymbol{\Omega} \in \mathbb{R}^d$ (this Fourier transform was first defined in Matsuda and Yajima [2009] and Bandyopadhyay and Lahiri [2009]). Note that the factor $\frac{\lambda^{d/2}}{n}$ ensures that the variance of $J_n(\boldsymbol{\Omega})$ is non-degenerate when we let $\lambda \to \infty$. Contrary to the time series case, we use $\boldsymbol{\Omega}$ instead of ω for spatial frequencies as we make use of both notations later for spatio-temporal processes in Section 3.

Under the condition that the locations $\{\boldsymbol{s}_j\}$ are independent and uniformly distributed random variables on $[-\lambda/2, \lambda/2]^d$ and $\{Z(\boldsymbol{s}); \boldsymbol{s} \in \mathbb{R}^d\}$ is a fourth order stationary process (with suitable short memory conditions), Bandyopadhyay and Subba Rao [2016] shows that the Fourier transform at the ordinates $\boldsymbol{\Omega}_{\boldsymbol{k}} = 2\pi(\frac{k_1}{\lambda}, \ldots, \frac{k_d}{\lambda})'$, $\boldsymbol{k} = (k_1, \ldots, k_d) \in \mathbb{Z}^d$, i.e., $\{J_n(\boldsymbol{\Omega}_{\boldsymbol{k}})\}$'s are 'near uncorrelated' random variables. For their variances, we have $\mathrm{var}[J_n(\boldsymbol{\Omega}_k)] = f(\boldsymbol{\Omega}_k) + O(\frac{1}{\lambda} + \frac{\lambda^d}{n})$, where

$$f(\boldsymbol{\Omega}) = \int_{\mathbb{R}^d} c(\boldsymbol{s}) \exp(-i\boldsymbol{s}'\boldsymbol{\Omega}) d\boldsymbol{\Omega}$$

is the spectral density function of the spatial process. So far the results are very similar to those in time series, however, because the spatial process is defined over $\mathbb{R}^d$ and not over $\mathbb{Z}^d$, the spectral density $f(\boldsymbol{\Omega})$ is defined over $\mathbb{R}^d$. For the same reason, f is no longer an infinite sum, but becomes an integral. Furthermore, $|f(\boldsymbol{\Omega})| \to 0$ as $\|\boldsymbol{\Omega}\|_2 \to \infty$, where $\|\cdot\|_2$ denotes the Euclidean norm (since the spatial covariance decays to zero sufficiently fast, $c(\cdot) \in L_2(\mathbb{R}^d)$ and thus by Parseval's inequality $f \in L_2(\mathbb{R}^d)$). Therefore, $1/\sqrt{f(\boldsymbol{\Omega})}$ is not a well defined function for all $\boldsymbol{\Omega} \in \mathbb{R}^d$ and unlike the discrete time series case, the standardized Fourier transform $J_n(\boldsymbol{\Omega}_{\boldsymbol{k}})/\sqrt{f(\boldsymbol{\Omega}_{\boldsymbol{k}})}$ is not a well defined quantity at all frequencies. Instead, to measure the degree of correlation between DFTs, we have to avoid standardization and we define the weighted covariance between the (non-standardized) Fourier transforms as

$$\widehat{A}_\lambda(g;\boldsymbol{r}) = \frac{1}{\lambda^d}\sum_{k_1,\ldots,k_d=-a}^{a} g(\boldsymbol{\Omega}_{\boldsymbol{k}})J_n(\boldsymbol{\Omega}_{\boldsymbol{k}})\overline{J_n(\boldsymbol{\Omega}_{\boldsymbol{k}+\boldsymbol{r}})} - \left[\frac{1}{n^2}\sum_{k_1,\cdots,k_d=-a}^{a} g(\boldsymbol{\Omega}_{\boldsymbol{k}})\sum_{j=1}^{n} Z^2(\boldsymbol{s}_j)\exp(-i\boldsymbol{s}_j'\boldsymbol{\Omega}_{\boldsymbol{r}})\right], \quad (7)$$

where, $\boldsymbol{r} \neq \boldsymbol{0}$, $\boldsymbol{r} = (r_1,\ldots,r_d)' \in \mathbb{Z}^d$ (with $\boldsymbol{k}$ and $\boldsymbol{k}+\boldsymbol{r}$ defined analogously), g is a given Lipschitz continuous function with $\sup_{\boldsymbol{\Omega}\in\mathbb{R}^d}|g(\boldsymbol{\Omega})| < \infty$ and a satisfies $(\lambda a)^d/n^2 \to 0$. In order to avoid the so called 'nugget effect' where the observations are corrupted by measurement error (typically independent noise) we have subtracted the variance-type term in the definition of (7).

We give some examples of functions g below.

Remark 2.1 *Examples of g used in Bandyopadhyay and Subba Rao [2016] are functions of the form $g(\boldsymbol{\Omega}) = e^{i\boldsymbol{v}'\boldsymbol{\Omega}}$, which is geared towards detecting changes in the spatial covariance at lag $\boldsymbol{v}$. However, unlike the case of regularly spaced locations, where we can detect changes at integer lags, it is unclear which lags to use. For this reason in Bandyopadhyay and Subba Rao [2016] we choose $g(\cdot)$ such that it can detect the aggregate change over L lags, namely $g(\boldsymbol{\Omega}) = \sum_{j=1}^{L}\exp(i\boldsymbol{v}_j'\boldsymbol{\Omega})$ (where $\{\boldsymbol{v}_j\}$ is some grid within the main support of the covariance). We should note that $g(\cdot)$ is similar to the weight function $e^{i\ell\omega}[\widehat{f}_T(\omega)\widehat{f}_T(\omega+\omega_r)]^{-1/2}$ used in the definition of $\widehat{C}_T(r,\ell)$ in (2).*

We derive the sampling results under the mixed asymptotic framework, where $\lambda \to \infty$ and $\lambda^d/n \to 0$, i.e., as the spatial domain grows, the number of observations should become denser on the spatial domain (see, Hall and Patil [1994], Lahiri [2003], Matsuda and Yajima [2009], Bandyopadhyay and Lahiri [2009], and Bandyopadhyay et al. [2015]). Under this mixed asymptotic framework, we show in Theorem 3.1 of Bandyopadhyay and Subba Rao [2016], that

$$\mathrm{E}\left[\widehat{A}_\lambda(g;\boldsymbol{r})\right] = \begin{cases} O\left(\frac{1}{\lambda^{d-b}}\prod_{j=1}^{d-b}(\log\lambda + \log|m_j|)\right), & \boldsymbol{r} \in \mathbf{Z}^d/\{\mathbf{0}\} \\ \frac{1}{(2\pi)^d}\int_{\boldsymbol{\Omega}\in\mathbb{R}^d} f(\boldsymbol{\Omega})g(\boldsymbol{\Omega})d\boldsymbol{\Omega} + O\left(\frac{\log\lambda}{\lambda} + \frac{1}{n}\right), & \boldsymbol{r} = \mathbf{0} \end{cases}$$

where $a^d = O(n)$, $a/\lambda \to \infty$ as $n \to \infty$ and $\lambda \to \infty$, $b = b(\boldsymbol{r})$ are the number of zero values in the vector $\boldsymbol{r}$, $\{m_j\}$ are the non-zero values in the vector $\boldsymbol{r}$. From Section 3 (Theorem 3.3 treats the Gaussian case and the non-Gaussian case can be found at the bottom of Section 3), Bandyopadhyay and Subba Rao [2016], we have

$$c_{a,\lambda}^{-1/2}\lambda^{d/2}\left[\ \Re\widehat{A}_\lambda(g;\boldsymbol{r}_1), \Im\widehat{A}_\lambda(g;\boldsymbol{r}_1), \ldots, \Re\widehat{A}_\lambda(g;\boldsymbol{r}_m), \Im\widehat{A}_\lambda(g;\boldsymbol{r}_m)\ \right]' \xrightarrow{\mathcal{D}} \mathcal{N}(\mathbf{0}, I_{2m}) \quad (8)$$

as $\ell_{\lambda,a,n} := \log^2 a\left(\frac{\log a+\log\lambda}{\lambda}\right) + \frac{\lambda^d}{n} + \frac{a^d\lambda^d}{n^2} + \frac{\log^3\lambda}{\lambda} \to 0$, where

$$\begin{aligned} c_{a,\lambda} &= \frac{1}{2(2\pi)^d}\int_{\mathcal{D}} f^2(\mathbf{\Omega})\left(|g(\mathbf{\Omega})|^2 + g(\mathbf{\Omega})\overline{g(-\mathbf{\Omega})}\right)d\mathbf{\Omega} \\ &+ \frac{1}{2(2\pi)^{2d}}\int_{\mathcal{D}^2} f_4(\mathbf{\Omega}_1,\mathbf{\Omega}_2,-\mathbf{\Omega}_2)g(\mathbf{\Omega}_1)\overline{g(\mathbf{\Omega}_2)}d\mathbf{\Omega}_1 d\mathbf{\Omega}_2, \end{aligned}$$

$\mathcal{D} = [-2\pi a/\lambda, 2\pi a/\lambda]^d$ and

$$f_4(\mathbf{\Omega}_1,\mathbf{\Omega}_2,\mathbf{\Omega}_3) = \int_{\mathbb{R}^{3d}} \kappa(\boldsymbol{s}_1,\boldsymbol{s}_2,\boldsymbol{s}_3)e^{-i(\boldsymbol{s}_1'\mathbf{\Omega}_1+\boldsymbol{s}_2'\mathbf{\Omega}_2+\boldsymbol{s}_3'\mathbf{\Omega}_3)}d\boldsymbol{s}_1 d\boldsymbol{s}_2 d\boldsymbol{s}_2$$

is the (spatial) tri-spectral density and $\kappa(\boldsymbol{s}_1,\boldsymbol{s}_2,\boldsymbol{s}_3) = \mathrm{cum}[Z(0),Z(\boldsymbol{s}_1),Z(\boldsymbol{s}_2),Z(\boldsymbol{s}_3)]$ is the fourth order cumulant analogous to κ_{h_1,h_2,h_3} in the time series case. We observe that unlike $\widehat{C}_T(r,\ell)$ (defined in (2)), even in the case that the random field is stationary and Gaussian, $\widehat{A}_\lambda(g;\boldsymbol{r})$ is not asymptotically pivotal. This is because, unlike $\widehat{C}_T(r,\ell)$, in the definition of $\widehat{A}_\lambda(g;\boldsymbol{r})$ we could not standardize the Fourier transform $J_n(\mathbf{\Omega})$ such that $f(\mathbf{\Omega})$ crops up in the asymptotics here. Therefore, even for Gaussian random fields, the variance $c_{a,\lambda}$ needs to be estimated and if the random field is non-Gaussian then $c_{a,\lambda}$ additionally contains a function of the fourth order spectral density.

In the following remark, we present an approach based on orthogonal samples, as proposed by Subba Rao [2015b], to estimate complicated variances.

Remark 2.2 (Using orthogonal samples for variance estimation) *The expression for the variance $\widehat{A}_\lambda(g;0)$ given in the examples above, is rather unwieldy and difficult to estimate directly. For example, in the case that the random field is Gaussian, one can estimate $c_{a,\lambda}$ by replacing the integral with the sum $\sum_{\boldsymbol{k}=-a}^{a}$ and the spectral density function with periodogram $|J_n(\mathbf{\Omega}_{\boldsymbol{k}})|^2$ (see Bandyopadhyay et al. [2015], Lemma 7.5). However, in the case that the process is non-Gaussian this is not possible. Here we describe the method of orthogonal samples, which can be used for both spatial and/or temporal data and it is a simple consistent method for estimating the variance.*

Let us suppose that $\widehat{A}_D(\mathbf{X})$ is an estimator of A where $\mathrm{E}[\widehat{A}_D(\mathbf{X})] \to A$ and $var[\sqrt{D}\widehat{A}_D(\mathbf{X})] = \nu$ (where $D = D(T,\lambda)$ is an appropriate scaling factor such that $var[\sqrt{D}\widehat{A}_D(\mathbf{X})] = O(1)$. For some set $\mathcal{B}$, the sample $\{\sqrt{D}A_D(\mathbf{X};j); j\in\mathcal{B}\}$ (which is not necessarily real-valued) is the orthogonal sample associated with $\widehat{A}_D(\mathbf{X})$ if (i) $\{\sqrt{D}\widehat{A}_D(\mathbf{X};j); j\in\mathcal{B}\}$ and $\widehat{A}_D(\mathbf{X})$ are almost uncorrelated but (ii) $\{\sqrt{D}\widehat{A}_D(\mathbf{X};j); j\in\mathcal{B}\}$ has mean almost zero and $\mathrm{var}[\sqrt{D}\widehat{A}_D(\mathbf{X};r)] = \nu + o(1)$. Based on this we can estimate ν using

$$\widehat{\nu} = \widehat{\sigma}^2(\{\sqrt{D}\widehat{A}_D(\mathbf{X};j); j\in\mathcal{B}\}) = \frac{D}{2|\mathcal{B}|}\sum_{j\in\mathcal{B}}\left[(\Re\widehat{A}_D(\mathbf{X};j)-\bar{A})^2 + (\Im\widehat{A}_D(\mathbf{X};j)-\bar{A})^2\right], \quad (9)$$

and $\bar{A} = \frac{1}{2|\mathcal{B}|}\sum_{j\in\mathcal{B}}[\Re\widehat{A}_D(\mathbf{X};j) + \Im\widehat{A}_D(\mathbf{X};j)]$, *where* $|\mathcal{B}|$ *denotes the cardinality of the set* $\mathcal{B}$. *Furthermore, if* $\sqrt{D}[\widehat{A}_D(\mathbf{X}) - A, \{\Re\widehat{A}_D(\mathbf{X};j), \Im\widehat{A}_D(\mathbf{X};j); j\in\mathcal{B}\}] \xrightarrow{\mathcal{D}} N(0, \nu I_{2|\mathcal{B}|+1})$, *then*

$$\sqrt{D}\frac{[\widehat{A}_D(\mathbf{X}) - A]}{\sqrt{\widehat{\nu}}} \xrightarrow{\mathcal{D}} t_{2|\mathcal{B}|-1}.$$

This method allows us to estimate the variance of an estimator and to quantify the uncertainty in the variance estimator. In the testing procedures described in this paper we make frequent use of this method. We describe below how it is used in the spatial stationarity test.

To implement the test, we define a set $\mathcal{S} \in \mathbb{Z}^d$ that surrounds but does not include zero (examples include $\mathcal{S} = \{(1,0),(1,1),(0,1),(-1,1)\}$) and test for stationarity using the coefficients, $\{\widehat{A}_\lambda(g;\boldsymbol{r}); \boldsymbol{r}\in\mathcal{S}\}$. Of course, the variance $c_{a,\lambda}$ is unknown and needs to be estimated from the data. To estimate the variance, we observe from (8) that (a) $\Re\widehat{A}_\lambda(g;\boldsymbol{r})$ and $\Im\widehat{A}_\lambda(g;\boldsymbol{r})$ have the same variance and (b) for all $\boldsymbol{r}$ 'close' to zero the variance of $\{\Re\widehat{A}_\lambda(g;\boldsymbol{r}), \Im\widehat{A}_\lambda(g;\boldsymbol{r})\}$ is approximately the same which allows us to use the orthogonal sample method described in Remark 2.2 to estimate the variance. Therefore, we define a set $\mathcal{S}' \in \mathbb{Z}^d$ which is relatively 'close' to $\mathcal{S}$, but $\mathcal{S}\cap\mathcal{S}' = \emptyset$. We then estimate $c_{a,\lambda}$ using $\widehat{c}_{a,\lambda}(\mathcal{S}') := \widehat{\sigma}^2(\{\lambda^{d/2}\widehat{A}_\lambda(g;\boldsymbol{r}); \boldsymbol{r}\in\mathcal{S}'\})$ (where $\widehat{\sigma}^2(\cdot)$ is defined in (9)). Using $\widehat{c}_{a,\lambda}(\mathcal{S}')$ and (8) we have

$$\lambda^{d/2}\frac{\Re\widehat{A}_\lambda(g;\boldsymbol{r})}{\sqrt{\widehat{c}_{a,\lambda}(\mathcal{S}')}} \xrightarrow{\mathcal{D}} \frac{Z_{1,\boldsymbol{r}}}{\sqrt{\frac{1}{2|\mathcal{S}'|-1}\chi^2_{2|\mathcal{S}'|-1}}} \sim t_{2|\mathcal{S}'|-1} \text{ and } \lambda^{d/2}\frac{\Im\widehat{A}_\lambda(g;\boldsymbol{r})}{\sqrt{\widehat{c}_{a,\lambda}(\mathcal{S}')}} \xrightarrow{\mathcal{D}} \frac{Z_{2,\boldsymbol{r}}}{\sqrt{\frac{1}{2|\mathcal{S}'|-1}\chi^2_{2|\mathcal{S}'|-1}}} \sim t_{2|\mathcal{S}'|-1}, \tag{10}$$

for $\boldsymbol{r}\in\mathcal{S}$ with $\lambda^d/n \to 0$ as $n\to\infty$ and $\lambda\to\infty$ (so called mixed domain asymptotics), where $\{Z_{1,\boldsymbol{r}}, Z_{2,\boldsymbol{r}}; \boldsymbol{r}\in\mathcal{S}\}$ are iid standard normal random variables and $\chi^2_{2|\mathcal{S}'|-1}$ is a chi-squared distributed random variable (with $2|\mathcal{S}'|-1$ degrees of freedom) which is the same for all $\boldsymbol{r}\in\mathcal{S}$, but independent of $\{Z_{1,\boldsymbol{r}}, Z_{2,\boldsymbol{r}}; \boldsymbol{r}\in\mathcal{S}\}$, and t_q denotes the t-distribution with q degrees of freedom. A test statistic can then be defined as $\max_{\boldsymbol{r}\in\mathcal{S}}[|\widehat{A}_\lambda(g;\boldsymbol{r})|^2/\widehat{c}_{a,\lambda}(\mathcal{S}')]$, whose limiting distribution can easily be obtained from (10). Note that a test statistic based on the sum of squares rather than the maximum is also possible, however in terms of simulations the maximum statistic tends to have slightly better power.

Just as in the nonstationary time series case, in order to obtain the limit of $\widehat{A}_\lambda(g;\boldsymbol{r})$ in the nonstationary spatial case, we use rescaled asymptotics. We define a sequence of nonstationary spatial processes $\{Z_\lambda(\boldsymbol{s})\}$ (we use the term 'sequence' loosely, since λ is defined on $\mathbb{R}^+$ and not on $\mathbb{Z}^+$), where for each $\lambda > 0$ and $\boldsymbol{s}\in[-\lambda/2,\lambda/2]^d$ the covariance of $\{Z_\lambda(\boldsymbol{s})\}$ is

$$\mathrm{cov}[Z_\lambda(\boldsymbol{s}), Z_\lambda(\boldsymbol{s}+\boldsymbol{v})] = \kappa\left(\boldsymbol{v}; \frac{\boldsymbol{s}}{\lambda}\right),$$

where $\kappa : \mathbb{R}^d \times [-1/2, 1/2]^d \to \mathbb{R}$ (note that $\boldsymbol{s} \in [-\lambda/2, \lambda/2]^d$) is the location-dependent covariance function. The corresponding location-dependent spectral density function is defined as

$$F\left(\boldsymbol{\Omega}; \frac{\boldsymbol{s}}{\lambda}\right) = \int_{\mathbb{R}^d} \kappa\left(\boldsymbol{v}; \frac{\boldsymbol{s}}{\lambda}\right) \exp(-i2\pi \boldsymbol{v}'\boldsymbol{\Omega}) d\boldsymbol{v}.$$

Under this set-up we have $\widehat{A}_\lambda(g; \boldsymbol{r}) \xrightarrow{\mathcal{P}} A(g; \boldsymbol{r})$ as $\lambda \to \infty$ where

$$A(g; \boldsymbol{r}) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \int_{[-1/2,1/2]^d} F(\boldsymbol{\Omega}; \boldsymbol{u}) \exp\left(-i2\pi \boldsymbol{u}'\boldsymbol{r}\right) g(\boldsymbol{\Omega}) d\boldsymbol{u} d\boldsymbol{\Omega}.$$

We observe that if in the test we let $g(\boldsymbol{\Omega}) = \exp(i\boldsymbol{v}'\boldsymbol{\Omega})$ then $A(e^{i\boldsymbol{v}\cdot}; \boldsymbol{r})$ is the Fourier coefficient of $\int_{[-1/2,1/2]^d} \kappa(\boldsymbol{v}; \boldsymbol{u}) \exp(-i\boldsymbol{r}'\boldsymbol{u}) d\boldsymbol{u}$. Hence the test is geared towards detecting changes in the covariance at lag $\boldsymbol{v}$.

2.3 The test for temporal stationarity using orthogonal samples

Let us return to the temporal stationarity test discussed in Section 2.1. Suppose ℓ is fixed, then by (3) and (4), we test for stationarity with $\{\widehat{C}_T(r, \ell); r \in \mathcal{T}\}$ where $\mathcal{T} = \{1, \ldots, m\}$ and use the L_2-type statistic defined in (5).

In this section we apply the ideas in Section 2.2 to propose a different method for estimating $v_\ell(\omega_r)$. We recall that $\Re\widehat{C}_T(r, \ell)$ and $\Im\widehat{C}_T(r, \ell)$ have 'variance' $v_\ell(\omega_r)$ and that $v_\ell(\omega_r) = v_\ell(0) + O(|r|/T)$. Define the set $\mathcal{T}' = \{m + B_1, \ldots, m + B_2\}$ where $B_1 < B_2$. If B_2 is not 'too large' then for $r \in \mathcal{T}'$ the 'variance' of $\widehat{C}_T(r, \ell)$ is approximately $v_\ell(0)$. Therefore we use

$$\widehat{v}_\ell := \widehat{\sigma}^2(\{T^{1/2}\widehat{C}_T(r, \ell); r \in \mathcal{T}'\}),$$

as an estimator of $v_\ell := v_\ell(0)$, where $\widehat{\sigma}^2(\cdot)$ is defined in (9). Based on this estimator we use the test statistic

$$\mathbf{T}_m = T \sum_{r=1}^{m} \frac{|\widehat{C}_T(r, \ell)|^2}{\widehat{v}_\ell}.$$

It can be shown that $\widehat{v}_\ell \xrightarrow{\mathcal{P}} v_\ell$ as $|\mathcal{T}'| \to \infty$, therefore if m is kept fixed we have that $\mathbf{T}_m \xrightarrow{\mathcal{D}} \chi^2_{2m}$ with $(|\mathcal{T}'| + B_1)/T \to 0$ as $\mathcal{T}' \to \infty$ and $T \to \infty$. However, for finite B_2, i.e., $|\mathcal{T}'| < \infty$, a better finite sample approximation of the distribution of the test statistic under the null uses (4) and that $\widehat{v}_\ell$ is asymptotically independent of $\sum_{r=1}^{m} |\widehat{C}_T(r, \ell)|^2$ to give

$T\sum_{r=1}^{m}|\widehat{C}_T(r,\ell)|^2/v_\ell \xrightarrow{\mathcal{D}} \sum_{j=1}^{2m} Z_j^2$ and $\widehat{v}_\ell/v_\ell \xrightarrow{\mathcal{D}} \frac{1}{2|\mathcal{T}'|-1}\chi^2_{2|\mathcal{T}'|-1}$ such that we obtain

$$\mathbf{T}_m \xrightarrow{\mathcal{D}} \frac{\sum_{j=1}^{2m} Z_j^2}{\frac{1}{2|\mathcal{T}'|-1}\chi^2_{2|\mathcal{T}'|-1}},$$

where $\{Z_j\}_{j=1}^{2m}$ are iid standard normal random variables which are independent of the χ^2 random variable.

We note that to maximize power we require that $|B_1| \to \infty$ in such a way that $|B_1|/T \to 0$ as $T \to \infty$. To understand why, we observe that under the alternative $\widehat{C}_T(r,\ell)$ are estimating Fourier coefficients (see (6)). If B_1 is close to the origin then $\{\widehat{C}_T(r,\ell); r \in \mathcal{T}'\}$ will be estimating 'large' values, thus the sample variance of $\{\widehat{C}_T(r,\ell); r \in \mathcal{T}'\}$ is likely to be large thus reducing $\mathbf{T}_m$ and consequently the power of the test.

3 Properties of spatio-temporal Fourier transforms

We now use some of the ideas discussed in the previous section to test for stationarity of a spatio-temporal process. Let us suppose that $\{Z_t(\boldsymbol{s}); \boldsymbol{s} \in \mathbb{R}^d, t \in \mathbb{Z}\}$ is a spatio-temporal process which is observed at time $t = 1,\ldots,T$ and at locations $\{\boldsymbol{s}_j\}_{j=1}^n$ on the region $[-\lambda/2, \lambda/2]^d$. At any given time point, t, we may not observe all $\{\boldsymbol{s}_j\}_{j=1}^n$ locations, but only a subset $\{\boldsymbol{s}_{t,j}\}_{j=1}^{n_t}$, i.e., the data set we observe is $\{Z_t(\boldsymbol{s}_{t,j}); j = 1,\ldots,n_t, t = 1,\ldots,T\}$.

Throughout this paper we will use the following set of assumptions.

Assumption 3.1

(i) $\{\boldsymbol{s}_j\}$ are iid uniformly distributed random variables on the region $[-\lambda/2, \lambda/2]^d$.

(ii) The number of locations that are observed at each time point is n_t, where for some $0 < c_1 \le c_2 < \infty$ (this does not change with n) we have $c_1 n \le n_t \le c_2 n$.

(iii) The asymptotics are mixed, that is as $\lambda \to \infty$ (spatial domain grows), we have $n \to \infty$ (number of locations grows) such that $\lambda^d/n \to 0$. We also assume that $T \to \infty$.

In much of the discussion below we restrict ourselves to the case $r_2 \in \{0,1,\ldots,T/2-1\}$, but allow $\boldsymbol{r}_1 \in \mathbb{Z}^d$.

Throughout the following, let $\boldsymbol{\Omega_k} = 2\pi(k_1/\lambda,\ldots,k_d/\lambda)$, where $\boldsymbol{k} = (k_1,\ldots,k_d) \in \mathbb{Z}^d$ denote *spatial* frequencies and $\omega_k = 2\pi k/T$ denote *temporal* frequencies. Keeping time or location fixed, respectively, we define the Fourier transform *over space* at time t as

$$J_t(\boldsymbol{\Omega_k}) = \frac{\lambda^{d/2}}{n_t}\sum_{j=1}^{n_t} Z_t(\boldsymbol{s}_{t,j})\exp(i\boldsymbol{s}_{t,j}'\boldsymbol{\Omega_k}),$$

and the Fourier transform *over time* at location $\boldsymbol{s}_j$ as

$$J_{\boldsymbol{s}_j}(\omega_k) = \frac{1}{\sqrt{2\pi T}} \sum_{t=1}^{T} \frac{n}{n_t} \delta_{t,j} Z_t(\boldsymbol{s}_j) e^{it\omega_k},$$

where, $\delta_{t,j} = 0$ if at time t the location $\boldsymbol{s}_j$ is not observed, otherwise $\delta_{t,j} = 1$. Observe that the ratio n/n_t gives a large weight to time points where there are only a few observed locations. We then define the spatio-temporal Fourier transform, i.e., the Fourier transformation *over space and time* as

$$J(\boldsymbol{\Omega}_{\boldsymbol{k}_1}, \omega_{k_2}) = \frac{1}{\sqrt{2\pi T}} \sum_{t=1}^{T} J_t(\boldsymbol{\Omega}_{\boldsymbol{k}_1}) \exp(it\omega_{k_2}) = \frac{\lambda^{d/2}}{n} \sum_{j=1}^{n} J_{\boldsymbol{s}_j}(\omega_{k_2}) \exp(i\boldsymbol{s}_j' \boldsymbol{\Omega}_{\boldsymbol{k}_1}). \tag{11}$$

Our objective is to test for second order stationarity of the spatio-temporal process, in the sense that $\mathrm{cov}[Z_t(\boldsymbol{s}), Z_{t+h}(\boldsymbol{s} + \boldsymbol{v})] = \kappa_h(\boldsymbol{v})$. In the remainder of this section we evaluate the covariance $\mathrm{cov}\left[J(\boldsymbol{\Omega}_{\boldsymbol{k}_1}, \omega_{k_2}), J(\boldsymbol{\Omega}_{\boldsymbol{k}_1+\boldsymbol{r}_1}, \omega_{k_2+r_2})\right]$ under all combinations of temporal and spatial stationarity and nonstationarity, respectively. This will motivate the testing procedures proposed in Section 4.

We first note there is a subtle but important difference between the spectral density over time and space. Under second order stationarity in space and time of $Z_t(\boldsymbol{s})$ the spectral density is

$$f(\boldsymbol{\Omega}, \omega) \quad = \quad \frac{1}{2\pi} \sum_{h \in \mathbb{Z}} e^{-ih\omega} \int_{\mathbb{R}^d} \kappa_h(\boldsymbol{v}) \exp(-i\boldsymbol{v}'\boldsymbol{\Omega}) d\boldsymbol{v},$$

where the equality above is due to $\kappa_h(\boldsymbol{v}) = \kappa_{-h}(-\boldsymbol{v})$. We observe that $f : \mathbb{R}^d \times [0, 2\pi] \to \mathbb{R}$, that is, $f(\cdot, \omega)$ is defined on $\mathbb{R}^d$ (this is because the spatial process is defined over $\mathbb{R}^d$) whereas $f(\boldsymbol{\Omega}, \cdot)$ is a periodic function defined on the interval $[0, 2\pi]$ (because the temporal process is over discrete time $\mathbb{Z}$). The space-time spectral density of the type defined above is studied in detail in Subba Rao and Terdik [2016].

To understand how the correlations between the Fourier transforms behave in the case that the spatio-temporal process is nonstationary we will use the rescaling devise discussed in Sections 2.1 and 2.2. To be able to apply the rescaling devise in *space and time*, we assume that the 'observed' process $Z_t(\boldsymbol{s})$ is an element of a sequence (indexed over λ and T) of nonstationary spatio-temporal processes $\{Z_{t,\lambda,T}(\boldsymbol{s}); t \in \mathbb{Z}, \boldsymbol{s} \in \mathbb{R}^d\}$, i.e., $Z_t(\boldsymbol{s}) = Z_{t,\lambda,T}(\boldsymbol{s})$. Using this formulation we can then place certain regularity conditions on the covariance. To do so, we define the sequence $\{\rho_h\}$, such that $\sum_h |h\rho_h| < \infty$ and function $\beta_\eta(\boldsymbol{v})$, such that for some $\eta > 0$, $\beta_\eta(\boldsymbol{v}) = \prod_{j=1}^{d} \beta_\eta(v_j)$ with

$$\beta_\eta(v_j) = \begin{cases} C & |v_j| \leq 1 \\ C|v_j|^{-\eta} & |v_j| > 1 \end{cases} \tag{12}$$

for some finite constant C. We assume there exists a time and location dependent spatio-temporal covariance, $\kappa_{h;u} : \mathbb{R}^d \times [-1/2, 1/2]^d \to \mathbb{R}$, such that for all $T \in \mathbb{Z}^+$, $\lambda > 0$, $h \in \mathbb{Z}$ and $u \in [0, 1]$, we have

$$\text{cov}[Z_{t,\lambda,T}(\boldsymbol{s}), Z_{t+h,\lambda,T}(\boldsymbol{s} + \boldsymbol{v})] = \kappa_{h;\frac{t}{T}}\left(\boldsymbol{v}; \frac{\boldsymbol{s}}{\lambda}\right) + O\left(\frac{\rho_h \beta_{2+\delta}(\boldsymbol{v})}{T}\right). \tag{13}$$

The function $\kappa_{\cdot}(\cdot)$ satisfies the Lipschitz conditions: (i) $\sup_{u,\boldsymbol{u}} |\kappa_{h;u}(\boldsymbol{v}; \boldsymbol{u})| \leq \rho_h \beta_{2+\delta}(\boldsymbol{v})$, (ii) $\sup_{\boldsymbol{u}} |\kappa_{h;u_1}(\boldsymbol{v}; \boldsymbol{u}) - \kappa_{h;u_2}(\boldsymbol{v}; \boldsymbol{u})| \leq |u_1 - u_2| \rho_h \beta_{2+\delta}(\boldsymbol{v})$ and (iii) $\sup_u |\kappa_{h;u}(\boldsymbol{v}; \boldsymbol{u}_1) - \kappa_{h;u}(\boldsymbol{v}; \boldsymbol{u}_2)| \leq ||\boldsymbol{u}_1 - \boldsymbol{u}_2||_2 \rho_h \beta_{2+\delta}(\boldsymbol{v})$. Note that the index $h; t/T$ refers to covariance at time lag h and rescaled time t/T whereas $(\boldsymbol{v}; \boldsymbol{s}/\lambda)$ refers to spatial covariance "lag" $\boldsymbol{v}$ and rescaled location $\boldsymbol{s}/\lambda$. Using the above definitions we define the *time and location dependent* spectral density as

$$F_u(\boldsymbol{\Omega}, \omega; \boldsymbol{u}) = \frac{1}{2\pi} \sum_{h \in \mathbb{Z}} e^{-ih\omega} \int_{\mathbb{R}^d} \kappa_{h;u}(\boldsymbol{v}; \boldsymbol{u}) e^{-i\boldsymbol{v}'\boldsymbol{\Omega}} d\boldsymbol{v}. \tag{14}$$

In the proposed testing procedure we also consider one-way stationarity tests, where we test for stationarity over one domain without assuming stationarity over the other domain. To understand how these tests behave, we use the following rescaling devises:

- **Spatial stationarity and temporal nonstationarity**

  In this case, we assume that $Z_t(\boldsymbol{s}) = Z_{t,T}(\boldsymbol{s})$ and there exists a κ such that

  $$\text{cov}[Z_t(\boldsymbol{s}), Z_{t+h}(\boldsymbol{s} + \boldsymbol{v})] = \kappa_{h;\frac{t}{T}}(\boldsymbol{v}) + O(\rho_h \beta_{2+\delta}(\boldsymbol{v}) T^{-1}).$$

  The corresponding time-dependent spectral density is $F_{\frac{t}{T}}(\boldsymbol{\Omega}, \omega)$ (defined analogously to (14)).

- **Temporal stationarity and spatial nonstationarity**

  In this case we assume $Z_t(\boldsymbol{s}) = Z_{t,\lambda}(\boldsymbol{s})$ and there exists a κ such that

  $$\text{cov}[Z_t(\boldsymbol{s}), Z_{t+h}(\boldsymbol{s} + \boldsymbol{v})] = \kappa_h\left(\boldsymbol{v}; \frac{\boldsymbol{s}}{\lambda}\right)$$

  with corresponding location dependent spectral density $F(\boldsymbol{\Omega}, \omega; \frac{\boldsymbol{s}}{\lambda})$.

In the following lemma we derive the properties of the DFT for the four different combinations of temporal and spatial stationarity and nonstationarity, respectively.

Lemma 3.1 *Suppose Assumption 3.1 is satisfied. We further assume that under spatial and temporal stationarity* $|\kappa_h(\boldsymbol{v})| \leq \rho_h \beta_{2+\delta}(\boldsymbol{v})$, *temporal stationarity* $\sup_{\boldsymbol{u}} |\kappa_h(\boldsymbol{v};\boldsymbol{u})| \leq \rho_h \beta_{2+\delta}(\boldsymbol{v})$, *spatial stationarity* $\sup_u |\kappa_{h;u}(\boldsymbol{v})| \leq \rho_h \beta_{2+\delta}(\boldsymbol{v})$ *and temporal and spatial nonstationarity* $\sup_{\boldsymbol{u},u} |\kappa_{h;u}(\boldsymbol{v};\boldsymbol{u})| \leq \rho_h \beta_{2+\delta}(\boldsymbol{v})$ *with* $\beta_{2+\delta}(\boldsymbol{v})$ *and* $\{\rho_h\}$ *as defined in* (12). *Let* $b = b(\boldsymbol{r})$ *denote the number of zero entries in the vector* $\boldsymbol{r}$.

(i) If the process $\{Z_t(\boldsymbol{s}); \boldsymbol{s} \in \mathbb{R}^d, t \in \mathbb{Z}\}$ *is spatially and temporally stationary, we have*

$$\begin{aligned}&\operatorname{cov}\left[J(\boldsymbol{\Omega}_{\boldsymbol{k}_1}, \omega_{k_2}), J(\boldsymbol{\Omega}_{\boldsymbol{k}_1+\boldsymbol{r}_1}, \omega_{k_2+r_2})\right] \\ &= \begin{cases} f(\boldsymbol{\Omega}_{\boldsymbol{k}_1}, \omega_{k_2}) + O\left(\frac{1}{T} + \frac{1}{\lambda} + \frac{\lambda^d}{n}\right) & \boldsymbol{r}_1 = 0 \text{ and } r_2 = 0 \\ O\left(\frac{1}{T} + \frac{\lambda^d}{n}\right) & \boldsymbol{r}_1 = 0 \text{ and } r_2 \neq 0 \\ O\left(\frac{1}{\lambda^{d-b}}\right) & \boldsymbol{r}_1 \neq 0 \text{ and } r_2 = 0 \\ O\left(\frac{1}{T\lambda^{d-b}}\right) & \boldsymbol{r}_1 \neq 0 \text{ and } r_2 \neq 0. \end{cases}\end{aligned}$$

(ii) If the process is spatially stationary and temporally nonstationary, we have

$$\begin{aligned}&\operatorname{cov}\left[J(\boldsymbol{\Omega}_{\boldsymbol{k}_1}, \omega_{k_2}), J(\boldsymbol{\Omega}_{\boldsymbol{k}_1+\boldsymbol{r}_1}, \omega_{k_2+r_2})\right] \\ &= \begin{cases} \int_0^1 F_u(\boldsymbol{\Omega}_{\boldsymbol{k}_1}, \omega_{k_2}) \exp(-i2\pi r_2 u) du + O\left(\frac{1}{T} + \frac{\lambda^d}{n}\right) & \boldsymbol{r}_1 = 0 \text{ and } r_2 \in \mathbb{Z} \\ O\left(\frac{1}{\lambda^{d-b}} + \frac{1}{T}\right) & \boldsymbol{r}_1 \neq 0 \text{ and } r_2 \in \mathbb{Z} \end{cases}\end{aligned}$$

(iii) If the process is spatially nonstationary and temporally stationary, we have

$$\begin{aligned}&\operatorname{cov}\left[J(\boldsymbol{\Omega}_{\boldsymbol{k}_1}, \omega_{k_2}), J(\boldsymbol{\Omega}_{\boldsymbol{k}_1+\boldsymbol{r}_1}, \omega_{k_2+r_2})\right] \\ &= \begin{cases} \int_{[-1/2,1/2]^d} \exp(-i2\pi \boldsymbol{r}_1' \boldsymbol{u}) F(\boldsymbol{\Omega}_{\boldsymbol{k}_1}, \omega_{k_2}; \boldsymbol{u}) d\boldsymbol{u} + O\left(\frac{1}{T} + \frac{1}{\lambda}\right) & r_2 = 0 \text{ and } \boldsymbol{r}_1 \in \mathbb{Z}^d \\ O\left(\frac{1}{T}\right) & r_2 \neq 0 \text{ and } \boldsymbol{r}_1 \in \mathbb{Z}^d \end{cases}\end{aligned}$$

(iv) If the process is spatially and temporally nonstationary, we have

$$\begin{aligned}&\operatorname{cov}\left[J(\boldsymbol{\Omega}_{\boldsymbol{k}_1}, \omega_{k_2}), J(\boldsymbol{\Omega}_{\boldsymbol{k}_1+\boldsymbol{r}_1}, \omega_{k_2+r_2})\right] \\ &= \int_0^1 \exp(-i2\pi r_2 u) \int_{[-1/2,1/2]^d} \exp(-i2\pi \boldsymbol{r}_1' \boldsymbol{u}) F_u(\boldsymbol{\Omega}_{\boldsymbol{k}_1}, \omega_{k_2}; \boldsymbol{u}) d\boldsymbol{u} du + O\left(\frac{1}{\lambda} + \frac{1}{T} + \frac{\lambda^d}{n}\right).\end{aligned}$$

In the above lemma we see that if the process is stationary then for non-zero values of $\boldsymbol{r}_1$ or r_2 the covariance between the DFTs is close to zero. On the other hand, when the process is nonstationary the correlation is non-zero. In particular, $\operatorname{cov}\left[J(\boldsymbol{\Omega}_{\boldsymbol{k}_1}, \omega_{k_2}), J(\boldsymbol{\Omega}_{\boldsymbol{k}_1+\boldsymbol{r}_1}, \omega_{k_2+r_2})\right]$

is approximately equal to the Fourier coefficient $b_{r_2}(\boldsymbol{\Omega}_{\boldsymbol{k}_1}, \omega_{k_2}; \boldsymbol{r}_1)$, where

$$b_{r_2}(\boldsymbol{\Omega}, \omega; \boldsymbol{r}_1) = \int_0^1 \exp(-i2\pi r_2 u) \int_{[-1/2,1/2]^d} \exp(-i2\pi \boldsymbol{r}_1' \boldsymbol{u}) F_u(\boldsymbol{\Omega}, \omega; \boldsymbol{u}) d\boldsymbol{u} du. \tag{15}$$

We note that $F_u(\boldsymbol{\Omega}, \omega; \boldsymbol{u}) = \sum_{\boldsymbol{r}_1 \in \mathbb{Z}^d} \sum_{r_2 \in \mathbb{Z}} b_{r_2}(\boldsymbol{\Omega}, \omega; \boldsymbol{r}_1) e^{2\pi i(\boldsymbol{r}_1' \boldsymbol{u} + r_2 u)}$. Therefore, in the case the spatio-temporal process is stationary, for all $\boldsymbol{r}_1 \neq \boldsymbol{0}$ or $r_2 \neq 0$ we have $b_{r_2}(\boldsymbol{\Omega}, \omega; \boldsymbol{r}_1) = 0$ and for all $\boldsymbol{u}$ and u, $F_u(\boldsymbol{\Omega}, \omega; \boldsymbol{u}) = b_0(\boldsymbol{\Omega}, \omega; \boldsymbol{0}) = f(\boldsymbol{\Omega}, \omega)$ holds.

However, in the nonstationary case we have:

- **Spatial stationarity, but temporal nonstationarity**
  For all $\boldsymbol{r}_1 \neq \boldsymbol{0}$, $b_{r_2}(\boldsymbol{\Omega}, \omega; r_2) = 0$. But for at least some $r_2 \neq 0$ and $[\boldsymbol{\Omega}, \omega] \in \mathbb{R}^d \times [0, 2\pi]$ (measure non-zero), $b_{r_2}(\boldsymbol{\Omega}, \omega; 0) \neq 0$. In other words, the temporal nonstationarity is 'seen' on the r_2-axis.

- **Temporal stationarity, but spatial nonstationarity**
  For all $r_2 \neq 0$, $b_{r_2}(\boldsymbol{\Omega}, \omega; r_2) = 0$. But for at least some $\boldsymbol{r}_1 \neq 0$ and $[\boldsymbol{\Omega}, \omega] \in \mathbb{R}^d \times [0, 2\pi]$ (measure non-zero), $b_0(\boldsymbol{\Omega}, \omega; \boldsymbol{r}_1) \neq 0$. In other words, the spatial nonstationarity is 'seen' on the $\boldsymbol{r}_1$-axis.

- **Temporal and spatial nonstationarity**
  For at least some $\boldsymbol{r}_1 \neq 0$ and $r_2 \neq 0$ and $[\boldsymbol{\Omega}, \omega] \in \mathbb{R}^d \times [0, 2\pi]$ (measure non-zero), we have $b_{r_2}(\boldsymbol{\Omega}, \omega; \boldsymbol{r}_1) \neq 0$.

Using this dichotomy between stationary and nonstationary processes, our proposed test for stationarity is based on estimates of $b_{r_2}(\boldsymbol{\Omega}, \omega; \boldsymbol{r}_1)$. However, it is not feasible to test over all $(\boldsymbol{r}_1, r_2) \in \mathbb{Z}^{d+1}$. Instead, we note that since $\int_0^1 \int_{[-1/2,1/2]^d} |F_u(\boldsymbol{\Omega}, \omega; \boldsymbol{u})|^2 d\boldsymbol{u} du < \infty$, we have $\sum_{\boldsymbol{r}_1, r_2} |b_{r_2}(\boldsymbol{\Omega}, \omega; \boldsymbol{r}_1)|^2 < \infty$. Therefore $|b_{r_2}(\boldsymbol{\Omega}, \omega; \boldsymbol{r}_1)| \to 0$ as $\|\boldsymbol{r}_1\| \to \infty$ or $|r_2| \to \infty$. Thus a test based on $b_{r_2}(\boldsymbol{\Omega}, \omega; \boldsymbol{r}_1)$ should use $(\boldsymbol{r}_1, r_2)$ which are close to the origin (where the deviations from zero are likely to be largest, thus leading to maximum power), from now onwards we denote this test set as $\mathcal{P} = \mathcal{S} \times \mathcal{T}$.

For a given $(\boldsymbol{r}_1, r_2)$, one possibility is to simply estimate $b_{r_2}(\boldsymbol{\Omega}, \omega; \boldsymbol{r}_1)$ for *all* $\boldsymbol{\Omega}$ and ω. Therefore, if $b_{r_2}(\boldsymbol{\Omega}, \omega; \boldsymbol{r}_1)$ is non-zero for some values $\boldsymbol{\Omega}, \omega$ of non-zero measure, the test will (asymptotically) have power. However, the drawback of a such an omnipresent test is that it has very little power for small deviations from stationarity (i.e., when $b_{r_2}(\boldsymbol{\Omega}, \omega; \boldsymbol{r}_1)$ is small). Therefore in the following section we propose two different testing approaches. The first estimates a weighted integral of $b_{r_2}(\boldsymbol{\Omega}, \omega; \boldsymbol{r}_1)$, that is

$$A_{g,h}(\boldsymbol{r}_1, r_2) = \langle b_{r_2}(\cdot, \cdot; \boldsymbol{r}_1), g(\cdot)h(\cdot) \rangle = \frac{1}{(2\pi)^d \pi} \int_0^\pi \int_{\mathbb{R}^d} g(\boldsymbol{\Omega}) h(\omega) b_{r_2}(\boldsymbol{\Omega}, \omega; \boldsymbol{r}_1) d\boldsymbol{\Omega} d\omega,$$

for a given set of (weight) functions $g:\mathbb{R}^d\rightarrow\mathbb{R}$ and $h:[0,\pi]\rightarrow\mathbb{R}$. This test has the most power for small deviations from stationarity - but they have to be in a direction that $A_{g,h}$ is non-zero. The second testing method is a compromise, between the omnipresent test and the above test. In this test we estimate

$$D_{g,h,v}(\boldsymbol{r}_1,r_2)=\frac{1}{\pi}\int_0^\pi v(\omega)\left[\frac{h(\omega)}{(2\pi)^d}\int_{\mathbb{R}^d}g(\boldsymbol{\Omega})b_{r_2}(\boldsymbol{\Omega},\omega;\boldsymbol{r}_1)d\boldsymbol{\Omega}\right]^2d\omega \tag{16}$$

for a given set of functions g, h and v. This test uses $g(\boldsymbol{\Omega})$ to set the spatial features it wants to detect, but the sum of squares over all frequencies ω means that it can detect for deviations from temporal stationarity at *all* frequencies ω.

4 The spatio-temporal test for stationarity

In this section we focus on testing for stationarity of a spatio-temporal process. In Section 5 we adapt these methods to testing for one-way stationarity of a spatio-temporal process.

4.1 Measures of correlation in the Fourier transforms

Our aim is to test for second order stationarity by measuring the linear dependence between the Fourier transforms. To do this, we recall that the test for spatial stationarity is a sum of (weighted) sample autocovariances of $\{J(\boldsymbol{\Omega}_{\boldsymbol{k}})\}$ (see (7)). We now define an analogous quantity to test for spatio-temporal stationarity. We start by defining the weighted sample cross-covariance between $\{J(\boldsymbol{\Omega}_{\boldsymbol{k}_1},\omega_{k_2})\}$ and $\{J(\boldsymbol{\Omega}_{\boldsymbol{k}_1+\boldsymbol{r}_1},\omega_{k_2+r_2})\}$ over $\boldsymbol{k}_1$ (but with k_2 kept fixed)

$$\begin{aligned}\widehat{a}_g(\omega_{k_2};\boldsymbol{r}_1,r_2) &= \frac{1}{\lambda^d}\sum_{\boldsymbol{k}_1=-\boldsymbol{a}}^{\boldsymbol{a}}g(\boldsymbol{\Omega}_{\boldsymbol{k}_1})J(\boldsymbol{\Omega}_{k_1},\omega_{k_2})\overline{J(\boldsymbol{\Omega}_{\boldsymbol{k}_1+\boldsymbol{r}_1},\omega_{k_2+r_2})}-N_T\\ &= \frac{1}{n^2}\sum_{\boldsymbol{k}_1=-\boldsymbol{a}}^{\boldsymbol{a}}g(\boldsymbol{\Omega}_{\boldsymbol{k}_1})\sum_{\substack{j_1,j_2=1\\ j_1\neq j_2}}^{n}J_{\boldsymbol{s}_{j_1}}(\omega_{k_2})\overline{J_{\boldsymbol{s}_{j_2}}(\omega_{k_2+r_2})}\exp(i\boldsymbol{s}_{j_1}\boldsymbol{\Omega}_{\boldsymbol{k}_1}-i\boldsymbol{s}_{j_2}\boldsymbol{\Omega}_{\boldsymbol{k}_1+\boldsymbol{r}_1}),\end{aligned} \tag{17}$$

and $g:\mathbb{R}^d\rightarrow\mathbb{R}$ is a user chosen bounded Lipschitz continuous function (see Remark 2.1), a is such that $(a\lambda)^d/n^2\rightarrow 0$, where the last line follows from (11) and

$$\begin{aligned}N_T &= \frac{1}{n^2}\sum_{\boldsymbol{k}_1=-\boldsymbol{a}}^{\boldsymbol{a}}g(\boldsymbol{\Omega}_{\boldsymbol{k}_1})\sum_{j=1}^{n}J_{\boldsymbol{s}_j}(\omega_{k_2})\overline{J_{\boldsymbol{s}_j}(\omega_{k_2+r_2})}\exp(-i\boldsymbol{s}_j\boldsymbol{\Omega}_{\boldsymbol{r}_1}) \\ &= \frac{1}{2\pi T}\sum_{\boldsymbol{k}_1=-a}^{a}g(\boldsymbol{\Omega}_{\boldsymbol{k}_1})\sum_{t,\tau=1}^{T}e^{it\omega_{k_2}-i\tau\omega_{k_2+r_1}}\frac{1}{n_tn_\tau}\sum_{j=1}^{n_{t,\tau}}Z_t(\boldsymbol{s}_{t,\tau,j})Z_\tau(\boldsymbol{s}_{t,\tau,j})\exp(-i\boldsymbol{s}_j\boldsymbol{\Omega}_{\boldsymbol{r}_1}),\end{aligned} \tag{18}$$

where $n_{t,\tau}$ denotes the number of common locations at time t and τ. Our reason for removing the term N_T are two fold; the first is to remove the so called nugget effect which arises due to measurement error in the spatial observations, the second reason is that N_T tends to inflate the variance of $\widehat{a}_g(\cdot)$ (removing such a term is quite common in spatial statistics, see Matsuda and Yajima [2009], Subba Rao [2015a] and Bandyopadhyay et al. [2015]).

Remark 4.1 *An alternative choice of N_T is*

$$N_T = \frac{1}{2\pi T}\sum_{\boldsymbol{k}_1=-\boldsymbol{a}}^{\boldsymbol{a}} g(\boldsymbol{\Omega}_{\boldsymbol{k}_1})\sum_{t=1}^{T}\exp(-it\omega_{r_2})\frac{1}{n_t^2}\sum_{j=1}^{n_t} Z_t^2(\boldsymbol{s}_{t,j})\exp(-i\boldsymbol{s}_j\boldsymbol{\Omega}_{\boldsymbol{r}_1}).$$

Examples of weight functions $g(\cdot)$ are given in Remark 2.1. We will show in Lemma 4.1 that in many ways the sampling properties of $\lambda^{d/2}\widehat{a}_g(\omega_{k_2};\boldsymbol{r}_1,r_2)$ resemble the temporal DFT covariance $J_T(\omega_k)\overline{J_T(\omega_{k+r})}$; compare Section 2.1. To prove this result we require the following assumptions.

Assumption 4.1 *Suppose $\{Z_t(\boldsymbol{u});\boldsymbol{u}\in\mathbb{R}^d, t\in\mathbb{Z}\}$ is a fourth order stationary spatio-temporal process. Let $\kappa_{h_1,h_2,h_3}(\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3)=\mathrm{cum}[Z_t(\boldsymbol{s}),Z_{t+h_1}(\boldsymbol{s}+\boldsymbol{v}_1),Z_{t+h_2}(\boldsymbol{s}+\boldsymbol{v}_2),Z_{t+h_3}(\boldsymbol{s}+\boldsymbol{v}_3)]$ and define the functions*

$$\begin{aligned} f_h(\boldsymbol{\Omega}) &= \int_{\mathbb{R}^d}\kappa_h(\boldsymbol{v})\exp(-i\boldsymbol{v}'\boldsymbol{\Omega})d\boldsymbol{u}, \text{ and} \\ f_{h_1,h_2,h_3}(\boldsymbol{\Omega}_1,\boldsymbol{\Omega}_2,\boldsymbol{\Omega}_3) &= \int_{\mathbb{R}^d}\kappa_{h_1,h_2,h_3}(\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3)\exp(-i\boldsymbol{v}_1'\boldsymbol{\Omega}_1-i\boldsymbol{v}_2'\boldsymbol{\Omega}_2-i\boldsymbol{v}_3'\boldsymbol{\Omega}_3)d\boldsymbol{v}_1d\boldsymbol{v}_2d\boldsymbol{v}_3. \end{aligned}$$

(i) $f_h(\cdot)$ satisfies $\int_{\mathbb{R}^d}|f_h(\boldsymbol{\Omega})|d\boldsymbol{\Omega}\leq\rho_h$, $\int_{\mathbb{R}^d}|f_h(\boldsymbol{\Omega})|^2d\boldsymbol{\Omega}\leq\rho_h$ and $f_h(\boldsymbol{\Omega})\leq\rho_h\beta_{1+\delta}(\boldsymbol{\Omega})$.

(ii) For all $1\leq j\leq d$, the partial derivatives satisfy $|\frac{\partial f_h(\boldsymbol{\Omega})}{\partial\Omega_j}|\leq\rho_h\beta_{1+\delta}(\boldsymbol{\Omega})$, where $\boldsymbol{\Omega}=(\Omega_1,\ldots,\Omega_d)$.

(iii) $|f_{h_1,h_2,h_3}(\boldsymbol{\Omega}_1,\boldsymbol{\Omega}_2,\boldsymbol{\Omega}_3)|<\rho_{h_1}\rho_{h_2}\rho_{h_3}\prod_{j=1}^d\beta_{1+\delta}(\Omega_{1,j})\prod_{j=1}^d\beta_{1+\delta}(\Omega_{2,j})\prod_{j=1}^d\beta_{1+\delta}(\Omega_{3,j})$ and for $1\leq i\leq 3$ and $1\leq j\leq d$,

$$\left|\frac{\partial f_{h_1,h_2,h_3}(\boldsymbol{\Omega}_1,\boldsymbol{\Omega}_2,\boldsymbol{\Omega}_3)}{\partial\boldsymbol{\Omega}_{i,j}}\right|\leq\rho_{h_1}\rho_{h_2}\rho_{h_3}\prod_{j=1}^d\beta_{1+\delta}(\Omega_{1,j})\prod_{j=1}^d\beta_{1+\delta}(\Omega_{2,j})\prod_{j=1}^d\beta_{1+\delta}(\Omega_{3,j}).$$

In the results below we also require the fourth order spectral density

$$f_4(\boldsymbol{\Omega}_1,\omega_1,\boldsymbol{\Omega}_2,\omega_2,\boldsymbol{\Omega}_3,\omega_3)=\frac{1}{(2\pi)^3}\sum_{h_1,h_2,h_3\in\mathbb{Z}}f_{h_1,h_2,h_3}(\boldsymbol{\Omega}_1,\boldsymbol{\Omega}_2,\boldsymbol{\Omega}_3)e^{-ih_1\omega_1-ih_2\omega_3-ih_3\omega_3}.$$

4.1.1 Sampling properties of $\widehat{a}_g(\cdot)$ under stationarity

Below we derive the mean, variance and asymptotic normality of $\widehat{a}_g(\cdot)$ from (17) under the assumption that the spatio-temporal process is fourth order stationary.

Lemma 4.1 *Suppose Assumptions 3.1 and 4.1 hold. In addition,* $|\frac{\partial^d f_h(\boldsymbol{\Omega})}{\partial\Omega_1\ldots\partial\Omega_d}| \leq \rho_h\beta_{1+\delta}(\boldsymbol{\Omega})$ *(see the proof of Theorem 3.1, Subba Rao [2015a]). Then*

$$
\begin{aligned}
&\mathrm{E}\left[\widehat{a}_g(\omega_k;\boldsymbol{r}_1,r_2)\right] \\
&= \left\{\begin{array}{ll}
O\left(\frac{1}{T\lambda^{d-b}}\prod_{j=1}^{d-b}(\log\lambda+\log|m_j|)\right) & \boldsymbol{r}_1\in\mathbf{Z}^d/\{\mathbf{0}\} \text{ and } r_2\neq 0\\
O\left(\frac{1}{\lambda^{d-b}}\prod_{j=1}^{d-b}(\log\lambda+\log|m_j|)\right) & \boldsymbol{r}_1\in\mathbf{Z}^d/\{\mathbf{0}\} \text{ and } r_2= 0\\
O\left(\frac{1}{T}\right) & \boldsymbol{r}_1=\mathbf{0} \text{ and } r_2\neq 0\\
\frac{1}{(2\pi)^d}\int_{\boldsymbol{\Omega}\in\mathbb{R}^d} g(\boldsymbol{\Omega})f(\boldsymbol{\Omega},\omega_k)d\boldsymbol{\Omega}+O\left(\frac{\log\lambda}{\lambda}+\frac{1}{n}\right) & \boldsymbol{r}_1=\mathbf{0} \text{ and } r_2= 0
\end{array}\right. ,
\end{aligned}
$$

where $b=b(\boldsymbol{r}_1)$ *denotes the number of zeros in the vector* $\boldsymbol{r}_1$ *and* $\{m_j\}_{j=1}^{d-b}$ *are the non-zero values in* $\boldsymbol{r}_1$*.*

Lemma 4.2 *Suppose Assumptions 3.1 and 4.1 hold and* r_2,r_4 *are such that* $0\leq r_2,r_4\leq T/2$*. Then we have,*

$$
\begin{aligned}
&\lambda^d\mathrm{cov}\left[\Re\widehat{a}_g(\omega_{k_2},\boldsymbol{r}_1,r_2),\Re\widehat{a}_g(\omega_{k_4},\boldsymbol{r}_3,r_4)\right]\\
&= I_{\boldsymbol{r}_1=\boldsymbol{r}_3}I_{r_2=r_4}\left[I_{k_2=k_4}V_g(\omega_{k_2};\boldsymbol{\Omega}_{\boldsymbol{r}_1},\omega_{r_2})+I_{k_4=T-k_2-r_2}V_{g,2}(\omega_{k_2};\boldsymbol{\Omega}_{\boldsymbol{r}_1},\omega_{r_2})+O\left(\frac{1}{T}\right)\right]\\
&\qquad +O\bigg(\ell_{\lambda,a,n}\bigg),
\end{aligned} \tag{19}
$$

where

$$
\begin{aligned}
V_g(\omega;\boldsymbol{\Omega}_{\boldsymbol{r}_1},\omega_{r_2}) &= \frac{1}{2(2\pi)^d}\int_{\mathcal{D}} g(\boldsymbol{\Omega})\overline{g(\boldsymbol{\Omega})}f(\boldsymbol{\Omega},\omega)f(\boldsymbol{\Omega}+\boldsymbol{\Omega}_{\boldsymbol{r}_1},\omega+\omega_{r_2})d\boldsymbol{\Omega},\\
V_{g,2}(\omega;\boldsymbol{\Omega}_{\boldsymbol{r}_1},\omega_{r_2}) &= \frac{1}{2(2\pi)^d}\int_{\mathcal{D}_{\boldsymbol{r}_1}} g(\boldsymbol{\Omega})\overline{g(-\boldsymbol{\Omega}-\boldsymbol{\Omega}_{\boldsymbol{r}_1})}f(\boldsymbol{\Omega},-\omega)f(-\boldsymbol{\Omega}-\boldsymbol{\Omega}_{\boldsymbol{r}_1},\omega+\omega_{r_2})d\boldsymbol{\Omega},
\end{aligned}
$$

and $\int_{\mathcal{D}_{\boldsymbol{r}_1}}=\int_{2\pi\max(-a,-a-r_{1,1})/\lambda}^{2\pi\min(a,a-r_{1,1})/\lambda}\ldots\int_{2\pi\max(-a,-a-r_{1,d})/\lambda}^{2\pi\min(a,a-r_{1,d})/\lambda}$*. Note that* $\ell_{\lambda,a,n}$ *and* $\int_{\mathcal{D}}$ *are defined in Section 2.2. Exactly the same result as in* (19) *holds for* $\lambda^d\mathrm{cov}\left[\Im\widehat{a}_g(\omega_{k_2},\boldsymbol{r}_1,r_2),\Im\widehat{a}_g(\omega_{k_4},\boldsymbol{r}_3,r_4)\right]$*, whereas* $\lambda^d\mathrm{cov}\left[\Re\widehat{a}_g(\omega_{k_2},\boldsymbol{r}_1,r_2),\Im\widehat{a}_g(\omega_{k_4},\boldsymbol{r}_3,r_4)\right]=O(\ell_{\lambda,a,n}+\frac{I_{\boldsymbol{r}_1=\boldsymbol{r}_3}I_{r_2=r_4}}{T})$*.*

Let $\{(k_j,\boldsymbol{r}_1,r_2);1\leq j\leq m,(\boldsymbol{r}_1,r_2)\in\mathcal{P}$ *and* $k_{j_1}\neq T-k_{j_2}-r_2\}$ *be a collection of integer* *vectors. Then under under sufficient mixing conditions of* $\{Z_t(\boldsymbol{s})\}$ *we have*

$$\lambda^{d/2}\left[\left\{\frac{\Re\widehat{a}_g(\omega_{k_j};\boldsymbol{r}_1,r_2)}{\sqrt{V_g(\omega_{k_j};\boldsymbol{\Omega}_{\boldsymbol{r}_1},\omega_{r_2})}},\frac{\Im\widehat{a}_g(\omega_{k_j};\boldsymbol{r}_1,r_2)}{\sqrt{V_g(\omega_{k_j};\boldsymbol{\Omega}_{\boldsymbol{r}_1},\omega_{r_2})}};1\leq j\leq m,(\boldsymbol{r}_1,r_2)\in\mathcal{P}\right\}\right]\stackrel{\mathcal{D}}{\to}\mathcal{N}\left(0,I_{2m|\mathcal{P}|}\right),$$

as $\lambda^d/n\to 0$, $n\to\infty$, $\lambda\to\infty$ *and* $T\to\infty$.

We observe that for $\|\boldsymbol{r}_1\|_2 << \lambda$ and $|r_2| << T$ and by the smoothness of the spectral density f and tri-spectral density f_4, we have

$$V_g(x_1;\boldsymbol{\Omega}_{\boldsymbol{r}_1},\omega_{r_2}) = V_g(x_2) + O\left(|x_1-x_2|+|\boldsymbol{\Omega}_{\boldsymbol{r}_1}|+|\omega_{r_2}|\right), \tag{20}$$

where $V_g(x)=V_g(x;0,0)$. We use these approximations in Sections 4.2.1 and 4.3.1.

The lemmas above show that $\widehat{a}_g(\omega;\boldsymbol{r}_1,r_2)$ is estimating zero in the case that the spatio-temporal process is fourth order stationary. We observe that the variance of $\widehat{a}_g(\omega_k;\boldsymbol{r}_1,r_2)$ does not involve $V_{g,2}(\cdot)$. Therefore, in the definition of the test statistic, in Section 4.2 we average $\widehat{a}_g(\omega_k;\boldsymbol{r}_1,r_2)$ over the frequencies $\{\omega_k\}_{k=1}^{T/2}$. This is to avoid correlations between $\widehat{a}_g(\omega_k;\boldsymbol{r}_1,r_2)$ and $\widehat{a}_g(\omega_{T-k-r_2};\boldsymbol{r}_1,r_2)$ and thus the need to estimate $V_{g,2}$.

In the section below we show that $\widehat{a}_g(\omega;\boldsymbol{r}_1,r_2)$ behaves differently in the case that the spatio-temporal process is nonstationary.

4.1.2 Sampling properties of $a_g(\cdot)$ under nonstationarity

Using the rescaled asymptotic set-up described in Section 3 and the assumptions in Lemma 3.1 we can show that under the alternative of nonstationarity

$$\mathrm{E}\left[\widehat{a}_g(\omega_k;\boldsymbol{r}_1,r_2)\right] = b_{g,r_2}(\omega;\boldsymbol{r}_1) + O\left(\frac{\lambda^d}{n}+\frac{1}{\lambda}+\frac{1}{T}\right),$$

where

$$b_{g,r_2}(\omega;\boldsymbol{r}_1) = \langle g, b_{r_2}(\cdot,\omega;\boldsymbol{r}_1)\rangle = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} g(\boldsymbol{\Omega})b_{r_2}(\boldsymbol{\Omega},\omega;\boldsymbol{r}_1)d\boldsymbol{\Omega} \tag{21}$$

and $b_{r_2}(\cdot,\omega;\boldsymbol{r}_1)$ is defined in (15). Therefore, we see that if the process is nonstationary, $\widehat{a}_g(\omega_k;\boldsymbol{r}_1,r_2)$ is, in some sense, measuring the nonstationarity at frequency $(\boldsymbol{r}_1,r_2)$ in the spectrum.

4.2 Test statistic 1: The average covariance

Motivated by the results above we define the average covariance. To do so, we first note that Lemma 4.2 above shows that there is a 'significant' correlation between $\Re\widehat{a}_g(\omega_k;\boldsymbol{r}_1,r_2)$ and $\Re\widehat{a}_g(\omega_{T-k-r_2};\boldsymbol{r}_1,r_2)$ (and likewise for the imaginary parts). Therefore we restrict the summands below to the frequencies $1, \ldots, T/2$ to ensure the elements of the sum are mostly near uncorrelated. We define the weighted sum as

$$\widehat{A}_{g,h}(\boldsymbol{r}_1, r_2) = \frac{2}{T}\sum_{k=1}^{T/2} h(\omega_k)\widehat{a}_g(\omega_k; \boldsymbol{r}_1, r_2), \tag{22}$$

for given (user-chosen) weight functions g and h. The sampling properties of $\widehat{A}_{g,h}(\boldsymbol{r}_1, r_2)$ are given in Lemma 6.1. Summarizing Lemma 6.1, the variances are

$$\mathrm{var}\left(\sqrt{\frac{\lambda^d T}{2}}\Re\widehat{A}_{g,h}(\boldsymbol{r}_1, r_2)\right) = V_{g,h} + o(1) \text{ and } \mathrm{var}\left(\sqrt{\frac{\lambda^d T}{2}}\Im\widehat{A}_{g,h}(\boldsymbol{r}_1, r_2)\right) = V_{g,h} + o(1) \tag{23}$$

and

$$\sqrt{\frac{\lambda^d T}{2V_{g,h}}}\left(\left\{\Re\widehat{A}_{g,h}(\boldsymbol{r}_1, r_2), \Im\widehat{A}_{g,h}(\boldsymbol{r}_1, r_2); (\boldsymbol{r}_1, r_2) \in \widetilde{\mathcal{P}}\right\}\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0, I_{2|\widetilde{\mathcal{P}}|}) \tag{24}$$

as $\ell_{\lambda,a,n} \to 0$ with $\lambda \to \infty$ and $T \to \infty$, where

$$\begin{aligned} V_{g,h} = {} & \frac{1}{2\pi}\int_0^{\pi} |h(\omega)|^2 V_g(\omega) d\omega \\ & + \frac{2}{(2\pi)^{d+2}}\int_0^{\pi}\int_0^{\pi}\int_{\mathcal{D}^2} g(\boldsymbol{\Omega}_1)\overline{g(\boldsymbol{\Omega}_2)}h(\omega_1)\overline{h(\omega_2)}f_4(\boldsymbol{\Omega}_1, \omega_1, \boldsymbol{\Omega}_2, \omega_2, -\boldsymbol{\Omega}_2, -\omega_2) \\ & \qquad d\boldsymbol{\Omega}_1 d\boldsymbol{\Omega}_2 d\omega_1 d\omega_2. \end{aligned} \tag{25}$$

Therefore, based on the above we can use an L_2 or max norm as the test statistic, i.e.

$$\frac{\lambda^d T}{2V_{g,h}}\sum_{(\boldsymbol{r}_1, r_2)\in\mathcal{P}} |\widehat{A}_{g,h}(\boldsymbol{r}_1, r_2)|^2 \quad \text{or} \quad \frac{\lambda^d T}{2V_{g,h}}\max_{(\boldsymbol{r}_1, r_2)\in\mathcal{P}} |\widehat{A}_{g,h}(\boldsymbol{r}_1, r_2)|^2,$$

which is asymptotically either a chi-square statistic or the maximum of chi-squares with 2 degrees of freedom. However, we stumble across a problem in that the variance $V_{g,h}$ is generally unknown. One solution to this is to use the method of orthogonal samples described in Section 2. We observe from (23) and (24) that if $(\boldsymbol{r}_1, r_2)$ is not too far from the origin, then $\{\widehat{A}_{g,h}(\boldsymbol{r}_1, r_2); (\boldsymbol{r}_1, r_2)\}$'s asymptotically have the same variance and are uncorrelated. Therefore, we can estimate the variance using the elements in the set $\mathcal{P}'$. We describe how to construct the sets $\mathcal{P}$ and $\mathcal{P}'$ below.

Definition 4.1 (The sets $\mathcal{P}$ and $\mathcal{P}'$) *We define the sets $\mathcal{P} = \mathcal{S} \times \mathcal{T}$ and $\mathcal{P}' = \mathcal{S}' \times \mathcal{T}'$, where $\mathcal{S}$, $\mathcal{S}'$, $\mathcal{T}$ and $\mathcal{T}'$ are similar to the sets defined in Section 2.2 and Section 2.3 ($\mathcal{S}$ and $\mathcal{S}'$ contain vectors in $\mathbb{Z}^d$ whereas $\mathcal{T}$ and $\mathcal{T}'$ contain vectors in $\mathbb{Z}$). We place the following* *constraints on the sets;* $\mathbf{0} \notin \mathcal{P}, \mathcal{P}'$, $\mathcal{P} \cap \mathcal{P}' = \emptyset$. *Furthermore if* $(\boldsymbol{r}_1, r_2), (\boldsymbol{r}_3, r_4) \in \mathcal{P}$ *or* $\mathcal{P}'$, *then* $(\boldsymbol{r}_1, r_2) \neq -(\boldsymbol{r}_3, r_4)$. $\mathcal{P}$ *and* $\mathcal{P}'$ *are such that for* $(\boldsymbol{r}_1, r_2) \in \mathcal{P}'$, $\|\boldsymbol{r}_1\|_2 << \lambda$, $|r_2| << T$.

$\mathcal{P}$ *will be the set where we check for zero correlation and conduct the test and* $\mathcal{P}'$ *will be the set which we use to estimate nuisance parameters. In order for the test statistics defined below to be close to the nominal level, under the null of stationarity, the elements of* $\mathcal{P}$ *and* $\mathcal{P}'$ *should be 'close' (in the sense of some distance measure). However, in order for the test to have maximum power (i) the test set* $\mathcal{P}$ *should surround zero and (ii) if* $\mathcal{P}'$ *is too 'close' to* $\mathcal{P}$ *it can result in a loss of power. Further details can be found in Bandyopadhyay and Subba Rao [2016].*

Thus we estimate $V_{g,h}$ with

$$\widehat{V}_{g,h}(\mathcal{P}') = \widehat{\sigma}^2 \left(\sqrt{\frac{T\lambda^d}{2}} \widehat{A}_{g,h}(\boldsymbol{r}_1, r_2); (\boldsymbol{r}_1, r_2) \in \mathcal{P}' \right). \tag{26}$$

and use either the L_2-statistic $\mathbf{T}_{1,g,h}(\mathcal{P}, \mathcal{P}')$ or the maximum statistic $\mathbf{M}_{1,g,h}(\mathcal{P}, \mathcal{P}')$ as the test statistic, where

$$\mathbf{T}_{1,g,h}(\mathcal{P}, \mathcal{P}') = \frac{\lambda^d T}{2} \sum_{(\boldsymbol{r}_1, r_2) \in \mathcal{P}} \frac{|\widehat{A}_{g,h}(\boldsymbol{r}_1, r_2)|^2}{\widehat{V}_{g,h}(\mathcal{P}')} \text{ and } \mathbf{M}_{1,g,h}(\mathcal{P}, \mathcal{P}') = \frac{\lambda^d T}{2} \max_{(\boldsymbol{r}_1, r_2) \in \mathcal{P}} \frac{|\widehat{A}_{g,h}(\boldsymbol{r}_1, r_2)|^2}{\widehat{V}_{g,h}(\mathcal{P}')}. \tag{27}$$

Asymptotically, under the null of stationarity we have

$$\mathbf{T}_{1,g,h}(\mathcal{P}, \mathcal{P}') \xrightarrow{\mathcal{D}} \chi^2_{2|\mathcal{P}|} \quad \text{and} \quad \mathbf{M}_{1,g,h}(\mathcal{P}, \mathcal{P}') \xrightarrow{\mathcal{D}} F_{|\mathcal{P}|},$$

as $|\mathcal{P}'| \to \infty$, $T \to \infty$ and $\lambda \to \infty$, where $F_{|\mathcal{P}|}$ is the distribution function of the maximum of $|\mathcal{P}|$ i.i.d. exponentially distributed random variables with exponential parameter 1/2 (since asymptotically under the null, $(T\lambda^d/2)|\widehat{A}_{g,h}(\boldsymbol{r}_1, r_2)|^2/\widehat{V}_{g,h}(\mathcal{P}')$ limits to an exponential distribution) and is defined as $F_{|\mathcal{P}|}(x) = \frac{|\mathcal{P}|}{2} \exp(-x/2)(1 - \exp(-x/2))^{|\mathcal{P}|-1}$. Using this result we can test for stationarity at the $\alpha \times 100\%$-level with $\alpha \in (0, 1)$.

Remark 4.2 (The test under nonstationarity) *Suppose that* $Z_t(\boldsymbol{s}) = Z_{t,\lambda,T}(\boldsymbol{s})$ *is a nonstationary spatio-temporal process. Then by using the rescaling devise defined in Section 3 we have* $\widehat{A}_{g,h}(\boldsymbol{r}_1, r_2) \xrightarrow{\mathcal{P}} A_{g,h}(\boldsymbol{r}_1, r_2)$ *as* $T \to \infty$, $\lambda^d/n \to 0$, $\lambda \to \infty$ *and* $n \to \infty$, *where*

$$A_{g,h}(\boldsymbol{r}_1, r_2) = \frac{1}{\pi(2\pi)^d} \int_0^\pi h(\omega) \left(\int_{\mathbb{R}^d} g(\boldsymbol{\Omega}) b_{r_2}(\boldsymbol{\Omega}, \omega; \boldsymbol{r}_1) d\omega \right) d\boldsymbol{\Omega},$$

and $b_{r_2}(\boldsymbol{\Omega}, \omega; \boldsymbol{r}_1)$ *is defined in (15).*

We do not give the results of a formal local asymptotic analysis. However, suppose $\{\mu(\boldsymbol{r}_1, r_2)\}$ *is a sequence where* $\sum_{\boldsymbol{r}_1,r_2} |\mu(\boldsymbol{r}_1, r_2)|^2 < \infty$ *and*

$$A_{g,h}(\boldsymbol{r}_1, r_2) = \frac{\mu(\boldsymbol{r}_1, r_2)}{(T\lambda^d)^{1/2}}.$$

If for some $(\boldsymbol{r}_1, r_2) \in \mathcal{P}$, $\mu(\boldsymbol{r}_1, r_2) \neq 0$, *then the test will have power.*

Of course in order to define the test statistic, we need to choose g and h. A reasonable choice of $g(\cdot)$ is given in Remark 2.1. The choice of h is more complex and below we discuss a choice of h that seems to give reasonable results in the simulations.

4.2.1 Choice of h

If we let $h(\omega) = \exp(i\ell\omega)$, then the test is designed to check for nonstationarity only in the spatio-temporal covariance at temporal lag ℓ, i.e., $\kappa_{\ell;u}(\cdot; \boldsymbol{s})$. Instead, we use a weight function similar to the temporal test described in Section 2.1, where we recall that in the construction of the temporal test statistic $J_T(\omega_k)\overline{J_T(\omega_{k+r})}/\sqrt{\widehat{f}_T(\omega_k)\widehat{f}_T(\omega_{k+r})}$'s are near uncorrelated and $\widehat{C}_T(r, \ell)$ is pivotal in the case where the time series is stationary and Gaussian. Similarly, in the construction of $\widehat{A}_{g,h}(\boldsymbol{r}_1, r_2)$ if we let $h(\omega) = V_g(\omega)^{-1/2}$, where $V_g(\omega)$ is defined in (20), we have $\{\Re\widehat{a}_g(\omega_k; \boldsymbol{r}_1, r_2)/\sqrt{V_g(\omega_k)}, \Im\widehat{a}_g(\omega_k; \boldsymbol{r}_1, r_2)/\sqrt{V_g(\omega_k}\}$ are near uncorrelated, asymptotically standard normal random variables. Thus we use $h(\omega) = \sqrt{V_g(\omega)}$ to define $\widehat{A}_{g,V^{-1/2}}(\boldsymbol{r}_1, r_2)$ as

$$\widehat{A}_{g,V^{-1/2}}(\boldsymbol{r}_1, r_2) = \frac{2}{T}\sum_{k=1}^{T/2} \frac{\widehat{a}_g(\omega_k; \boldsymbol{r}_1, r_2)}{\sqrt{V_g(\omega_k)}},$$

which we see from (25) has variance

$$V_{g,V^{-1/2}} = \frac{1}{2} + \frac{2}{(2\pi)^{d+2}} \int_0^\pi \int_0^\pi \int_{\mathcal{D}^2} \frac{g(\boldsymbol{\Omega}_1)\overline{g(\boldsymbol{\Omega}_2)}}{\sqrt{V_g(\omega_1)V_g(\omega_2)}} f_4(\boldsymbol{\Omega}_1, \omega_1, \boldsymbol{\Omega}_2, \omega_2, -\boldsymbol{\Omega}_2, -\omega_2) \, d\boldsymbol{\Omega}_1 d\boldsymbol{\Omega}_2 d\omega_1 d\omega_2.$$

We observe from the above that in the case the spatio-temporal process is stationary *and* Gaussian, $\widehat{A}_{g,V^{-1/2}}(\boldsymbol{r}_1, r_2)$ is asymptotically pivotal; compare with temporal stationarity test described in Section 2.2 where a similar result is true.

However, in general $V_g(\omega)$ is unknown and needs to be estimated. To estimate $V_g(\omega_k)$ we use the orthogonal sample method described in Remark 2.1 and the same set $\mathcal{P}'$ defined in (4.1). Under these conditions we have that the real and imaginary parts of $\{\widehat{a}_g(\omega_{k+i}; \boldsymbol{r}_1, r_2); (\boldsymbol{r}_1, r_2) \in \mathcal{P}', |i| \leq M\}$ for $M << T$, have almost the same variance and are near uncorrelated. Using this we estimate $V_g(\omega_k)$ with

$$\widehat{V}_g(\omega_k;\mathcal{P}') = \widehat{\sigma}^2\left(\{\lambda^{d/2}\widehat{a}_g(\omega_{k+i};\boldsymbol{r}_1,r_2);(\boldsymbol{r}_1,r_2)\in\mathcal{P}',|i|\leq M\}\right), \tag{28}$$

where $\widehat{\sigma}^2(\cdot)$ is defined in (9). We define the *observed* average covariance as

$$\widehat{A}_{g,\widehat{V}^{-1/2}}(\boldsymbol{r}_1,r_2) = \frac{2}{T}\sum_{k=1}^{T/2}\frac{\widehat{a}_g(\omega_k;\boldsymbol{r}_1,r_2)}{\sqrt{\widehat{V}_g(\omega_k;\mathcal{P}')}}.$$

By using the same methods described in Jentsch and Subba Rao [2015], Appendix A.2, we can show that

$$\lambda^{d/2}T^{1/2}\left|\widehat{A}_{g,\widehat{V}^{-1/2}}(\boldsymbol{r}_1,r_2)-\widehat{A}_{g,V^{-1/2}}(\boldsymbol{r}_1,r_2)\right| \stackrel{\mathcal{P}}{\to} 0,$$

with $|M|/T\to 0$ as $M\to\infty$ and $T\to\infty$. Hence $\widehat{A}_{g,\widehat{V}^{-1/2}}(\boldsymbol{r}_1,r_2)$ and $\widehat{A}_{g,V^{-1/2}}(\boldsymbol{r}_1,r_2)$ share the same asymptotic sampling properties. Thus by using (24) we have

$$\sqrt{\frac{\lambda^d T}{2V_{g,V^{-1/2}}}}\left(\left\{\Re\widehat{A}_{g,\widehat{V}^{-1/2}}(\boldsymbol{r}_1,r_2),\Im\widehat{A}_{g,\widehat{V}^{-1/2}}(\boldsymbol{r}_1,r_2);(\boldsymbol{r}_1,r_2)\in\mathcal{P}\cap\mathcal{P}'\right\}\right)\stackrel{\mathcal{D}}{\to}\mathcal{N}(0,I_{2|\widetilde{\mathcal{P}}|}). \tag{29}$$

Since for a given data set, we cannot be sure if the underlying process is Gaussian, we estimate the variance of $V_{g,V^{-1/2}}$ using the method in (33) and use the test statistics $\mathbf{T}_{1,g,\widehat{V}^{-1/2}}(\mathcal{P},\mathcal{P}')$ and $\mathbf{M}_{1,g,\widehat{V}^{-1/2}}(\mathcal{P},\mathcal{P}')$ as defined in (27).

4.3 Test statistic 2: The average squared covariance

In the previous section we considered the average covariance for estimating the linear dependence between the DFTs. As we can see from Remark 4.2 the average covariance is designed to detect the frequency average deviation from stationarity. Of course by considering the frequency average deviation, positive and negative frequency deviations can cancel leading to an average deviation of zero, which would give the misleading impression of stationarity. To address this issue we define a test statistic which estimates the average squared deviation over all frequencies (and thus is designed to detect a wider range of alternatives). More precisely, we group $\{\widehat{a}_g(\omega_k;\boldsymbol{r}_1,r_2)\}_{k=1}^{T/2}$ into blocks of length H and evaluate the local average over each block

$$\widehat{B}_{g,h;H}(\omega_{jH};\boldsymbol{r}_1,r_2) = \frac{1}{H}\sum_{k=1}^{H}h(\omega_{jH+k})\widehat{a}_g(\omega_{jH+k};\boldsymbol{r}_1,r_2),\ \text{for } 0\leq j<T/(2H), \tag{30}$$

where the length of block H is such that H/T as $H \to \infty$ and $T \to \infty$. For ease of notation we assume that H is a multiple of T. This can be considered as a frequency localized version of $\widehat{A}_{g,h}(\boldsymbol{r}_1, r_2)$ defined in the previous section. In Lemma 6.2 we show that

$$\sqrt{\frac{\lambda^d H}{W_{g,h}(\omega_{jH})}} \left(\Re \widehat{B}_{g,h;H}(\omega_{jH}; \boldsymbol{r}_1, r_2), \Im \widehat{B}_{g,h;H}(\omega_{jH}; \boldsymbol{r}_1, r_2) \right) \xrightarrow{\mathcal{D}} \mathcal{N}(0, I_2),$$

where $W_{g,h}(\omega_{jH}) = W_{g,h}(\omega_{jH}; 0, 0)$ with

$$\begin{aligned} & W_{g,h}(\omega_{jH}) \\ & \quad = \frac{T}{2H\pi} \int_{\omega_{jH}}^{\omega_{(j+1)H}} |h(\omega)|^2 V_g(\omega) d\omega + \frac{T}{2H(2\pi)^{2d+2}} \int_{[\omega_{jH}, \omega_{(j+1)H}]^2} \int_{\mathcal{D}^2} g(\boldsymbol{\Omega}_1) \overline{g(\boldsymbol{\Omega}_2)} h(\omega_1) \overline{h(\omega_2)} \\ & \qquad \times f_4(\boldsymbol{\Omega}_1, \omega_1, \boldsymbol{\Omega}_2, \omega_2, -\boldsymbol{\Omega}_2, -\omega_2) d\boldsymbol{\Omega}_1 d\boldsymbol{\Omega}_2 d\omega_1 d\omega_2. \end{aligned}$$

A careful examination of the expression above shows that the term involving the fourth order cumulant is of lower order since it involves a double integral $\int_{[\omega_{jH}, \omega_{(j+1)H}]^2} = O((H/T)^2)$. Thus

$$W_{g,h}(\omega_{jH}) = \frac{T}{2H\pi} \int_{\omega_{jH}}^{\omega_{(j+1)H}} |h(\omega)|^2 V_g(\omega; 0, 0) d\omega + O\left(\frac{H}{T}\right). \tag{31}$$

Furthermore, the correlation *between* each of the blocks $W_{g,h}(\omega_{j_1 H})$ and $W_{g,h}(\omega_{j_2 H})$ is asymptotically negligible. Therefore, heuristically, we can treat the real and imaginary parts of $\{\sqrt{\frac{\lambda^d H}{W_{g,h}(\omega_{jH})}} \widehat{B}_{g,h;H}(\omega_{jH}; \boldsymbol{r}_1, r_2); j = 0, \ldots, \frac{T}{2H} - 1\}$ as 'independent standard normal random variables' and define its mean squared average

$$\widehat{D}_{g,h,W;H}(\boldsymbol{r}_1, r_2) = \frac{2H}{T} \sum_{j=0}^{T/2H-1} \frac{\left| \widehat{B}_{g,h;H}(\omega_{jH}; \boldsymbol{r}_1, r_2) \right|^2}{2W_{g,h}(\omega_{jH})}.$$

Thus, $\mathrm{E}[\widehat{D}_{g,h,W;H}(\boldsymbol{r}_1, r_2)] = \frac{1}{H\lambda^d}$ and analogous to (24) we have

$$\sqrt{\frac{T}{2H}} \left\{ \left[H\lambda^d \widehat{D}_{g,h,W;H}(\boldsymbol{r}_1, r_2) - 1 \right]; (\boldsymbol{r}_1, r_2) \in \widetilde{\mathcal{P}} \right\} \xrightarrow{\mathcal{D}} \mathcal{N}(0, I_{|\widetilde{P}|}), \tag{32}$$

with $H/T \to 0$ as $H, T, \lambda \to \infty$. We define an L_2 or maximum statistic based on the above. However, in practice the variance $W_{g,h}(\omega_{jH})$ is unknown and once again we invoke the method of orthogonal statistics to estimate it. We estimate $W_{g,h}(\omega_{j,H})$ with

$$\widehat{W}_{g,h}(\omega_{jH}; \mathcal{P}') = \widehat{\sigma}^2 \left(\sqrt{\lambda^d H} \widehat{B}_{g,h}(\omega_{jH}; \boldsymbol{r}_1, r_2); (\boldsymbol{r}_1, r_2) \in \mathcal{P}' \right), \tag{33}$$

and the observed $\widehat{D}_{g,h,W;H}(\boldsymbol{r}_1, r_2)$ is defined with W in $\widehat{D}_{g,h,W;H}(\boldsymbol{r}_1, r_2)$ replaced by $\widehat{W}$, that is,

$$\widehat{D}_{g,h,\widehat{W};H}(\boldsymbol{r}_1, r_2) = \frac{2H}{T} \sum_{j=0}^{T/2H-1} \frac{\left|\widehat{B}_{g,h;H}(\omega_{jH}; \boldsymbol{r}_1, r_2)\right|^2}{2\widehat{W}_{g,h}(\omega_{jH})}. \tag{34}$$

The test statistic is constructed using the L_2-sum

$$\mathbf{T}_{2,g,h,\widehat{W}}(\mathcal{P}, \mathcal{P}') = \sqrt{\frac{T}{2H}} \sum_{(\boldsymbol{r}_1, r_2) \in \mathcal{P}} H\lambda^d \widehat{D}_{g,h,\widehat{W};H}(\boldsymbol{r}_1, r_2),$$

and by using (32), under the of stationarity null, we have $\left(\mathbf{T}_{2,g,h,\widehat{W}}(\mathcal{P}, \mathcal{P}') - \sqrt{\frac{T}{2H}}|\mathcal{P}|\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0, |\mathcal{P}|)$ with $H/T \to 0$ and $|\mathcal{P}'| \to \infty$ as $T, \lambda, H \to \infty$. The maximum statistic $\mathbf{M}_{2,g,h,\widehat{W}}$ is defined analogously.

Remark 4.3 (The test under nonstationarity) *Suppose that $\{Z_t(\boldsymbol{s})\}$ is a nonstationary spatio-temporal process. Then by using the rescaling devise described in Section 3 we can show that $\widehat{D}_{g,h,W;H}(\boldsymbol{r}_1, r_2) \xrightarrow{\mathcal{P}} D_{g,h,W;H}(\boldsymbol{r}_1, r_2)$ as $T \to \infty$, $\lambda^d/n \to 0$, $H/T \to 0$, $H \to \infty$, $\lambda \to \infty$ and $n \to \infty$, where $D_{g,h,W;H}(\boldsymbol{r}_1, r_2)$ is defined in (16).*

Again, without conducting a formal local asymptotic analysis, if

$$D_{g,h,W;H}(\boldsymbol{r}_1, r_2) = \frac{\mu(\boldsymbol{r}_1, r_2)}{T^{1/2} H^{1/2} \lambda^d}$$

where $\sum_{\boldsymbol{r}_1, r_2} |\mu(\boldsymbol{r}_1, r_2)|^2 < \infty$ and for some $(\boldsymbol{r}_1, r_2) \in \mathcal{P}$, $\mu(\boldsymbol{r}_1, r_2) \neq 0$, then the test will have power.

4.3.1 Choice of h

Motivated by Section 4.2.1 we let $h(\omega) = \widehat{V}_g(\omega; \mathcal{P}')^{-1/2}$ and define the local average

$$\widehat{B}_{g,\widehat{V}^{-1/2};H}(\omega_{jH}; \boldsymbol{r}_1, r_2) = \frac{1}{H} \sum_{k=1}^{H} \frac{\widehat{a}_g(\omega_{jH+k}; \boldsymbol{r}_1, r_2)}{\sqrt{\widehat{V}_g(\omega_{jH+k}; \mathcal{P}')}}. \tag{35}$$

By using (31) we see its real and imaginary parts have limiting variance

$$W_{g,V^{-1/2}}(\omega_{jH}) = 1 + O\left(\frac{H}{T}\right). \tag{36}$$

Therefore, we observe that by using $h(\omega) = \sqrt{V_g(\omega)}$, $W_{g,V^{-1/2}}(\omega_{jH})$ is asymptotically pivotal (even if the underlying spatio-temporal process is nonstationary). In other words, we can treat the real and imaginary parts of $\{\sqrt{2\lambda^d H}\widehat{B}_{g,h;H}(\omega_{jH};\boldsymbol{r}_1, r_2); j = 0, \ldots, \frac{T}{2H} - 1\}$ as 'independent standard normal random variables' and define its mean squared average

$$\widehat{D}_{g,\widehat{V}^{-1/2},1;H}(\boldsymbol{r}_1, r_2) = \frac{2H}{T} \sum_{j=0}^{T/2H-1} \frac{|\widehat{B}_{g,\widehat{V}^{-1/2};H}(\omega_{jH};\boldsymbol{r}_1, r_2)|^2}{2}. \tag{37}$$

Studying $\widehat{D}$, we have avoided estimating the variance of $\widehat{B}_{g,\widehat{V}^{-1/2};H}(\omega_{jH};\boldsymbol{r}_1, r_2)$ by simply replacing this variance by its limiting variance which is 1. In the simulation study in Section 7 we compare the effect this has on the finite sample properties of the test statistic. Using (32) we have

$$\sqrt{\frac{T}{2H}}\left(H\lambda^d \widehat{D}_{g,\widehat{V}^{-1/2},1;H}(\boldsymbol{r}_1, r_2) - 1\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1) \tag{38}$$

with $H/T \to 0$ as $T \to \infty$, $H \to \infty$ and $\lambda \to \infty$. Therefore we define the test statistic

$$\mathbf{T}_{2,g,\widehat{V}^{-1/2},1}(\mathcal{P},\mathcal{P}') = \sqrt{\frac{T}{2H}} \sum_{(\boldsymbol{r}_1,r_2)\in\mathcal{P}} H\lambda^d \widehat{D}_{g,\widehat{V}^{-1/2},1;H}(\boldsymbol{r}_1, r_2)$$

and that under the null of stationarity $\left(\mathbf{T}_{2,g,\widehat{V}^{-1/2},1}(\mathcal{P},\mathcal{P}') - |\mathcal{P}|\sqrt{T/2H}\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,|\mathcal{P}|)$. Since approximately $T\lambda^d \widehat{D}_{g,\widehat{V}^{-1/2},1;H}(\boldsymbol{r}_1, r_2) \sim \chi^2_{T/H}$, chi-squared with T/H-degrees of freedom, a similar result can be derived for the analogous maximum statistic $\mathbf{M}_{2,g,\widehat{V}^{-1/2},1}(\mathcal{P},\mathcal{P}')$ based on the maximum of chi-squares.

In Section 7 we compare $\mathbf{T}_{2,g,\widehat{V}^{-1/2},\widehat{W}}(\mathcal{P},\mathcal{P}')$ and $\mathbf{M}_{2,g,\widehat{V}^{-1/2},\widehat{W}}(\mathcal{P},\mathcal{P}')$ (when we standardize with sample variance $\widehat{W}$) with $\mathbf{T}_{2,g,\widehat{V}^{-1/2},1}(\mathcal{P},\mathcal{P}')$ and $\mathbf{M}_{2,g,\widehat{V}^{-1/2},1}(\mathcal{P},\mathcal{P}')$.

4.4 Asymptotic 'finite sample' approximations of the distribution of the test statistics under the null

We recall that in Section 2.2 real and imaginary parts of the estimator $\frac{\widehat{A}_\lambda(g;\boldsymbol{r})}{\sqrt{\widehat{c}_{a,\lambda}(\mathcal{S}')}}$ converge to a standard normal distribution under the null of stationarity as $\lambda \to \infty$ and $|\mathcal{S}'| \to \infty$. However, in reality $|\mathcal{S}'|$ is finite and not that large. Therefore, in (10) we estimate it with the t-distribution, which can be considered as the 'asymptotic finite sample distribution' of this ratio. In this section we use an analogous method to approximate the distributions of

$$\frac{|\widehat{A}_{g,\widehat{V}^{-1/2}}(\boldsymbol{r}_1,r_2)|^2}{\widehat{V}_g(\mathcal{P}')}, \quad \widehat{D}_{g,\widehat{V}^{-1/2},\widehat{W};H}(\boldsymbol{r}_1,r_2) \text{ and } \widehat{D}_{g,\widehat{V}^{-1/2},1;H}(\boldsymbol{r}_1,r_2).$$

Let $\{Z_{R,j}(\boldsymbol{r}_1,r_2),Z_{I,j}(\boldsymbol{r}_1,r_2);j=1,\ldots,T/2\}$ and $\{Z_{j,k};j=1,\ldots,2|\mathcal{P}'|,k=1,\ldots,T/2\}$ denote iid standard Gaussian random variables (we use a double index because it simplifies some of the notations later on). We recall from the definition of $\widehat{V}_g(\omega_k;\mathcal{P}')$ in (28) that it is composed of $(2M+1)$-local average of $\Re\widehat{a}_g(\cdot)$ and $\Im\widehat{a}_g(\cdot)$, each term being asymptotically normal. Therefore we replace all the $\Re\widehat{a}_g(\cdot)$ and $\Im\widehat{a}_g(\cdot)$ in the definition of $\widehat{V}_g(\omega_k;\mathcal{P}')$ with standard normal distributions to give

$$\lambda^{d/2}\frac{\Re\widehat{a}_g(\omega_k;\boldsymbol{r}_1,r_2)}{\sqrt{\widehat{V}_g(\omega_k;\mathcal{P}')}} \sim t_{R,k}(\boldsymbol{r}_1,r_2) \quad \lambda^{d/2}\frac{\Im\widehat{a}_g(\omega_k;\boldsymbol{r}_1,r_2)}{\sqrt{\widehat{V}_g(\omega_k;\mathcal{P}')}} \sim t_{I,k}(\boldsymbol{r}_1,r_2),$$

where

$$\begin{aligned} t_{R,k}(\boldsymbol{r}_1,r_2) &= \lambda^{d/2}\frac{\Re\widehat{a}_g(\omega_k;\boldsymbol{r}_1,r_2)}{\sqrt{\widehat{V}_g(\omega_k;\mathcal{P}')}} \sim \frac{Z_{R,k}(\boldsymbol{r}_1,r_2)}{\sqrt{\frac{1}{2(2M+1)|\mathcal{P}'|-1}\sum_{i=-M}^{M}\sum_{j=1}^{2|\mathcal{P}'|}(Z_{j,k+i}-\bar{Z}_k)^2}} \\ t_{I,k}(\boldsymbol{r}_1,r_2) &= \lambda^{d/2}\frac{\Im\widehat{a}_g(\omega_k;\boldsymbol{r}_1,r_2)}{\sqrt{\widehat{V}_g(\omega_k;\mathcal{P}')}} \sim \frac{Z_{I,k}(\boldsymbol{r}_1,r_2)}{\sqrt{\frac{1}{2(2M+1)|\mathcal{P}'|-1}\sum_{i=-M}^{M}\sum_{j=1}^{2|\mathcal{P}'|}(Z_{j,k+i}-\bar{Z}_k)^2}} \end{aligned}$$

and $\bar{Z}_k = \frac{1}{2(2M+1)|\mathcal{P}'|}\sum_{i=-M}^{M}\sum_{j=1}^{2|\mathcal{P}'|}Z_{j,k+i}$ is the local average.

Noting that the test statistic is in terms of $\frac{\lambda^d T}{2}|\widehat{A}_{g,\widehat{V}^{-1/2}}(\boldsymbol{r}_1,r_2)|^2$, we replace the real and imaginary parts of $\widehat{a}_g(\omega_k;\boldsymbol{r}_1,r_2)/\sqrt{\widehat{V}_g(\omega_k;\mathcal{P}')}$ with the above to give

$$\frac{T\lambda^d}{2}\frac{|\widehat{A}_{g,\widehat{V}^{-1/2}}(\boldsymbol{r}_1,r_2)|^2}{V_{g,V^{-1/2}}} \sim X_{\boldsymbol{r}_1,r_2}$$

where

$$X_{\boldsymbol{r}_1,r_2} = \frac{T}{2}\left|\sum_{k=1}^{T/2}t_{R,k}(\boldsymbol{r}_1,r_2)\right|^2 + \frac{T}{2}\left|\sum_{k=1}^{T/2}t_{I,k}(\boldsymbol{r}_1,r_2)\right|^2,$$

for other $(\boldsymbol{r}_1,r_2)$, we use independent $\{Z_{k,R}(\boldsymbol{r}_1,r_2),Z_{k,I}(\boldsymbol{r}_1,r_2)\}$ but the same $\{Z_{j,k}\}$. We also recall that we estimate the variance $V_{g,V^{-1/2}}$, therefore we approximate its distribution with a weighted chi-squared with $(2|\mathcal{P}'|-1)$ degrees of freedom. Since the orthogonal sample which was used to estimate it contained $2|\mathcal{P}'|$ terms

$$\frac{\widehat{V}_{g,\widehat{V}^{-1/2}}}{V_{g,V^{-1/2}}} \sim \frac{1}{2|\mathcal{P}'|-1}\chi^2_{2|\mathcal{P}'|-1}.$$

Therefore, based on the above, the following distribution is used

$$\mathbf{T}_{1,g,\widehat{V}^{-1/2}}(\mathcal{P},\mathcal{P}') \sim \frac{1}{\frac{1}{2|\mathcal{P}'|-1}\chi^2_{2|\mathcal{P}'|-1}} \sum_{(\boldsymbol{r}_1,r_2)\in\mathcal{S}_1} X_{\boldsymbol{r}_1,r_2}$$

to approximate the 'asymptotic finite sample properties' of $\mathbf{T}_{1,g,\widehat{V}^{-1/2}}(\mathcal{P},\mathcal{P}')$, under the null. Using the same method we can obtain the 'asymptotic finite sample properties' for $\mathbf{M}_{1,g,\widehat{V}^{-1/2}}(\mathcal{P},\mathcal{P}')$.

Next we consider the local average DFTs, $\widehat{B}_{g,\widehat{V}^{-1/2}}(\omega_{jH};\boldsymbol{r}_1,r_2)$, $\widehat{D}_{g,\widehat{V}^{-1/2},\widehat{W};H}(\boldsymbol{r}_1,r_2)$ and $\widehat{D}_{g,\widehat{V}^{-1/2},1;H}(\boldsymbol{r}_1,r_2)$, which lead to the test statistics $\mathbf{T}_{2,g,\widehat{V}^{-1/2},\widehat{W}}$ and $\mathbf{T}_{2,g,\widehat{V}^{-1/2},1}$. Using the arguments given above we have,

$$H\lambda^d \left|\widehat{B}_{g,\widehat{V}^{-1/2};H}(\omega_{jH};\boldsymbol{r}_1,r_2)\right|^2 \sim Y_{jH}(\boldsymbol{r}_1,r_2),$$

where,

$$Y_{jH}(\boldsymbol{r}_1,r_2) = \left|\frac{1}{\sqrt{H}}\sum_{k=1}^{H} t_{R,jH+k}(\boldsymbol{r}_1,r_2)\right|^2 + \left|\frac{1}{\sqrt{H}}\sum_{k=1}^{H} t_{I,jH+k}(\boldsymbol{r}_1,r_2)\right|^2.$$

Using the above, we estimate the distribution of $\widehat{D}_{g,\widehat{V}^{-1/2},1;H}(\boldsymbol{r}_1,r_2)$ with

$$H\lambda^d \widehat{D}_{g,\widehat{V}^{-1/2},1;H}(\boldsymbol{r}_1,r_2) \sim \frac{2H}{T}\sum_{j=1}^{T/2H} Y_{jH}(\boldsymbol{r}_1,r_2).$$

This gives us the 'asymptotic finite sample' distributions of $\mathbf{T}_{2,g,\widehat{V}^{-1/2},1}(\mathcal{P},\mathcal{P}')$ and $\mathbf{M}_{2,g,\widehat{V}^{-1/2},1}(\mathcal{P},\mathcal{P}')$.

In order to derive the sampling properties of $\mathbf{T}_{2,g,\widehat{V}^{-1/2},\widehat{W}}(\mathcal{P},\mathcal{P}')$ and $\mathbf{M}_{2,g,\widehat{V}^{-1/2},\widehat{W}}(\mathcal{P},\mathcal{P}')$, we recall that $\widehat{D}_{g,\widehat{V}^{-1/2},\widehat{W};H}(\boldsymbol{r}_1,r_2)$ involves $\widehat{W}_{g,\widehat{V}^{-1/2}}(\omega_{jk};\mathcal{P}')$ and we approximate this distribution by independent chi-squares $\{\frac{1}{2|\mathcal{P}'|-1}\chi^2_{2|\mathcal{P}'|-1,k}\}_{k=1}^{T/2H}$. This gives

$$\lambda^d H\widehat{D}_{g,\widehat{V}^{-1/2},\widehat{W};H}(\boldsymbol{r}_1,r_2) \sim \frac{2H}{T}\sum_{j=1}^{T/2H} \frac{Y_{jH}(\boldsymbol{r}_1,r_2)}{\frac{1}{2|\mathcal{P}'|-1}\chi^2_{2|\mathcal{P}'|-1,k}}.$$

Using this we can obtain the distributions of $\mathbf{T}_{2,g,\widehat{V}^{-1/2},\widehat{W}}(\mathcal{P},\mathcal{P}')$ and $\mathbf{M}_{2,g,\widehat{V}^{-1/2},\widehat{W}}(\mathcal{P},\mathcal{P}')$. Note that the same $\{\frac{1}{2|\mathcal{P}'|-1}\chi^2_{2|\mathcal{P}'|-1,k}\}_{k=1}^{T/2H}$ is used for all $\{\widehat{D}_{g,\widehat{V}^{-1/2},\widehat{W};H}(\boldsymbol{r}_1,r_2);(\boldsymbol{r}_1,r_2)\in\mathcal{P}\}$.

The 'asymptotic finite sample' distribution derived above are used in all the simulations below.

5 Testing for one-way stationarity

In this section we gear the procedure to specifically test for stationarity over one domain, without necessarily assuming stationarity over the other domain. For the one-way stationarity test we use the same test statistics defined in Section 4, however we use observations made in Section 3 in order to define the test set $\mathcal{P}$ over which the test statistic is defined. For $\widehat{a}_g(\cdot;\boldsymbol{r}_1,r_2)$ as defined in (17) we observe:

- **Spatial stationarity, but not necessarily temporal stationarity**
  If $\boldsymbol{r}_1 \neq \boldsymbol{0}$ for all ω, we have $\widehat{a}_g(\omega;\boldsymbol{r}_1,0) = o_p(1)$. On the other hand, if the process is spatially nonstationary then the latter is not necessarily true. Therefore the test set is $\mathcal{P} = \mathcal{S} \times \{0\}$.

- **Temporal stationarity, but not necessarily spatial stationarity**
  If $r_2 \neq 0$ for all ω we have $\widehat{a}_g(\omega;\boldsymbol{0},r_2) = o_p(1)$. On the other hand, if the process is temporally nonstationary then the latter is not necessarily true. Therefore the test set is $\mathcal{P} = \{\boldsymbol{0}\} \times \mathcal{T}$.

We recall that in order to ensure the test statistics defined in Section 4 are asymptotically pivotal we used the method of orthogonal samples to estimate the variance for various parts of the test statistic. Therefore, we need to ensure the set $\mathcal{P}'$ over which the orthogonal sample is defined is such that it consistently estimates the variance. To do this we derive expressions for the covariances of $\widehat{a}_g(\omega_{k_1};\boldsymbol{r}_1,r_2)$ and $\widehat{A}_{g,h}(\boldsymbol{r}_1,r_2)$, respectively, under the general nonstationary setting. In the following sections we consider the specific cases of spatial *or* temporal stationarity.

By using (15) we can show that under temporal and spatial nonstationarity the covariance of $\widehat{a}_g(\omega_{k_2};\boldsymbol{r}_1,r_2)$ is

$$\begin{aligned}
&\lambda^d \mathrm{cov}\left[\Re\widehat{a}_g(\omega_{k_2};\boldsymbol{r}_1,r_2), \Re\widehat{a}_g(\omega_{k_4};\boldsymbol{r}_3,r_4)\right] \\
=\ &\frac{1}{2}\Re\left[b^{(1)}_{r_2,r_4,k_2,k_4}(\omega_{k_2},\omega_{k_4};\boldsymbol{r}_1,\boldsymbol{r}_3) + b^{(2)}_{r_2,r_4,k_2,k_4}(\omega_{k_2},\omega_{k_4};\boldsymbol{r}_1,\boldsymbol{r}_3)\right] + O\left(\frac{1}{T}+\frac{1}{\lambda}+\ell_{\lambda,a,n}\right),
\end{aligned} \tag{39}$$

$$\begin{aligned}
&\lambda^d \mathrm{cov}\left[\Im\widehat{a}_g(\omega_{k_2};\boldsymbol{r}_1,r_2), \Im\widehat{a}_g(\omega_{k_4};\boldsymbol{r}_3,r_4)\right] \\
=\ &\frac{1}{2}\Re\left[b^{(1)}_{r_2,r_4,k_2,k_4}(\omega_{k_2},\omega_{k_4};\boldsymbol{r}_1,\boldsymbol{r}_3) - b^{(2)}_{r_2,r_4,k_2,k_4}(\omega_{k_2},\omega_{k_4};\boldsymbol{r}_1,\boldsymbol{r}_3)\right] + O\left(\frac{1}{T}+\frac{1}{\lambda}+\ell_{\lambda,a,n}\right),
\end{aligned}$$

and

$$\begin{aligned}
&\lambda^d \text{cov}\left[\Re \widehat{a}_g(\omega_{k_2};\boldsymbol{r}_1,r_2), \Im \widehat{a}_g(\omega_{k_4};\boldsymbol{r}_3,r_4)\right] \\
= \quad & -\frac{1}{2}\Im\left[b^{(1)}_{r_2,r_4,k_2,k_4}(\omega_{k_2},\omega_{k_4};\boldsymbol{r}_1,\boldsymbol{r}_3) - b^{(2)}_{r_2,r_4,k_2,k_4}(\omega_{k_2},\omega_{k_4};\boldsymbol{r}_1,\boldsymbol{r}_3)\right] + O\left(\frac{1}{T}+\frac{1}{\lambda}+\ell_{\lambda,a,n}\right),
\end{aligned}$$

where

$$\begin{aligned}
&b^{(1)}_{r_2,r_4,k_2,k_4}(\omega_{k_2},\omega_{k_4};\boldsymbol{r}_1,\boldsymbol{r}_3) = \frac{1}{\lambda^d}\sum_{\boldsymbol{k}_1,\boldsymbol{k}_3=-a}^{a} g(\boldsymbol{\Omega}_{\boldsymbol{k}_1})\overline{g(\boldsymbol{\Omega}_{\boldsymbol{k}_3})} \times \\
&\Big[b_{k_4-k_2}(\boldsymbol{\Omega}_{\boldsymbol{k}_1},\omega_{k_2};\boldsymbol{k}_3-\boldsymbol{k}_1)b_{k_4+r_4-k_2-r_2}(-\boldsymbol{\Omega}_{\boldsymbol{k}_1+\boldsymbol{r}_1},-\omega_{k_2+r_2};\boldsymbol{k}_3+\boldsymbol{r}_3-\boldsymbol{k}_1-\boldsymbol{r}_1) \\
&+b_{-k_4-k_2-r_4}(\boldsymbol{\Omega}_{\boldsymbol{k}_1},\omega_{k_2};-\boldsymbol{k}_3-\boldsymbol{k}_1-\boldsymbol{r}_3)b_{k_4+k_2+r_2}(-\boldsymbol{\Omega}_{\boldsymbol{k}_1+\boldsymbol{r}_1},-\omega_{k_2+r_2};\boldsymbol{k}_1+\boldsymbol{k}_3+\boldsymbol{r}_1)\Big],
\end{aligned}$$

and

$$\begin{aligned}
&b^{(2)}_{r_2,r_4,k_2,k_4}(\omega_{k_2},\omega_{k_4};\boldsymbol{r}_1,\boldsymbol{r}_3) = \frac{1}{\lambda^d}\sum_{\boldsymbol{k}_1,\boldsymbol{k}_3=-a}^{a} g(\boldsymbol{\Omega}_{\boldsymbol{k}_1})g(\boldsymbol{\Omega}_{\boldsymbol{k}_3}) \\
&\Big[b_{-k_4-k_2}(\boldsymbol{\Omega}_{\boldsymbol{k}_1},\omega_{k_2};-\boldsymbol{k}_3-\boldsymbol{k}_1)b_{k_4+k_2+r_4+k_2}(-\boldsymbol{\Omega}_{\boldsymbol{k}_1+\boldsymbol{r}_1},-\omega_{k_2+r_2};\boldsymbol{k}_3+\boldsymbol{r}_3+\boldsymbol{k}_1+\boldsymbol{r}_1) \\
&+b_{k_4-k_2+r_4}(\boldsymbol{\Omega}_{\boldsymbol{k}_1},\omega_{k_2};\boldsymbol{k}_3-\boldsymbol{k}_1+\boldsymbol{r}_3)b_{-k_4+k_2+r_2}(-\boldsymbol{\Omega}_{\boldsymbol{k}_1+\boldsymbol{r}_1},-\omega_{k_2+r_2};-\boldsymbol{k}_3+\boldsymbol{k}_1+\boldsymbol{r}_1)\Big].
\end{aligned}$$

Similar expressions for $\widehat{A}_{g,h}(\boldsymbol{r}_1,r_2)$ can be found in Section 6.2.

The above expressions are cumbersome, however, under one-way stationarity simplifications can be made. We recall from the definition in (15) that

$$b_{r_2}(\boldsymbol{\Omega},\omega;\boldsymbol{r}_1) = \begin{cases} 0 & \boldsymbol{r}_1 \neq 0 \text{ and spatial stationarity} \\ 0 & r_2 \neq 0 \text{ and temporal stationarity} \end{cases} \tag{40}$$

We use these results to simplify the expressions for $\text{cov}[\widehat{a}_g(\omega_{k_2};\boldsymbol{r}_1,r_2),\widehat{a}_g(\omega_{k_4};\boldsymbol{r}_3,r_4)]$ in the case of one-way stationarity.

5.1 Testing for spatial stationarity

In this section we adapt the test to testing for spatial stationarity. By using (39) and (40), under the null that $\{Z_t(\boldsymbol{s}); t \in \mathbb{Z}, \boldsymbol{s} \in \mathbb{R}^d\}$ is spatially stationary but not necessarily temporally stationary we have

$$\lambda^d \text{cov}[\Re\widehat{a}_g(\omega_{k_2};\boldsymbol{r}_1,r_2),\Re\widehat{a}_g(\omega_{k_4};\boldsymbol{r}_3,r_4)] = \begin{cases} b^{(1)}_{r_2,r_4,k_2,k_4}(\omega_{k_2},\omega_{k_4};\boldsymbol{r}_1,\boldsymbol{r}_1) + O(\frac{1}{T}+\ell_{\lambda,a,n}) & \boldsymbol{r}_1=\boldsymbol{r}_3 \\ b^{(2)}_{r_2,r_4,k_2,k_4}(\omega_{k_2},\omega_{k_4};\boldsymbol{r}_1,-\boldsymbol{r}_1) + O(\frac{1}{T}+\ell_{\lambda,a,n}) & \boldsymbol{r}_1=-\boldsymbol{r}_3 \\ O(\frac{1}{T}+\ell_{\lambda,a,n}) & \text{otherwise} \end{cases}$$

where

$$\begin{aligned} & b^{(1)}_{r_2,r_4,k_2,k_4}(\omega_{k_2},\omega_{k_4};\boldsymbol{r}_1,\boldsymbol{r}_1) \\ &= \frac{1}{2\lambda^d}\sum_{\boldsymbol{k}_1=-\boldsymbol{a}}^{\boldsymbol{a}} |g(\boldsymbol{\Omega}_{\boldsymbol{k}_1})|^2 b_{k_4-k_2}\left(\boldsymbol{\Omega}_{\boldsymbol{k}_1},\omega_{k_2};0\right) b_{k_4+r_4-k_2-r_2}\left(-\boldsymbol{\Omega}_{\boldsymbol{k}_1+\boldsymbol{r}_1},-\omega_{k_2+r_2};0\right) \\ &+ \frac{1}{2\lambda^d}\sum_{\boldsymbol{k}_1=\max(-\boldsymbol{a},-\boldsymbol{a}-\boldsymbol{r}_1)}^{\min(\boldsymbol{a},\boldsymbol{a}+\boldsymbol{r}_1)} g(\boldsymbol{\Omega}_{\boldsymbol{k}_1})\overline{g(-\boldsymbol{\Omega}_{\boldsymbol{k}_1+\boldsymbol{r}_1})} b_{-k_2-k_4-r_4}\left(\boldsymbol{\Omega}_{\boldsymbol{k}_1},\omega_{k_2};0\right) b_{k_2+k_4+r_4}\left(-\boldsymbol{\Omega}_{\boldsymbol{k}_1+\boldsymbol{r}_1},-\omega_{k_2+r_2};0\right), \end{aligned}$$

and $b^{(2)}_{r_2,r_4,k_2,k_4}(\omega_{k_2},\omega_{k_4};\boldsymbol{r}_1,-\boldsymbol{r}_1)$ is defined similarly. The same result can be shown for $\lambda^d\text{cov}[\Re\widehat{a}_g(\omega_{k_2};\boldsymbol{r}_1,r_2)\Re\widehat{a}_g(\omega_{k_4};\boldsymbol{r}_3,r_4)]$. Furthermore, $\lambda^d\text{cov}[\Re\widehat{a}_g(\omega_{k_2};\boldsymbol{r}_1,r_2),\Im\widehat{a}_g(\omega_{k_4};\boldsymbol{r}_3,r_4)] = o(1)$.

In order to test for spatial stationarity, without the temporal effect influencing the result, we focus on $r_2=0$. In this case, the above reduces to

$$\lambda^d \text{cov}[\Re\widehat{a}_g(\omega_{k_2};\boldsymbol{r}_1,0),\Re\widehat{a}_g(\omega_{k_4};\boldsymbol{r}_3,0)] = \begin{cases} b^{(1)}_{k_2,k_4}(\omega_{k_2},\omega_{k_4};\boldsymbol{r}_1) + O(\frac{\|\boldsymbol{r}_1\|_1}{\lambda}+\frac{1}{T}+\ell_{\lambda,a,n}) & \boldsymbol{r}_1=\boldsymbol{r}_3 \\ b^{(2)}_{k_2,k_4}(\omega_{k_2},\omega_{k_4};\boldsymbol{r}_1) + O(\frac{\|\boldsymbol{r}_1\|_1}{\lambda}+\frac{1}{T}+\ell_{\lambda,a,n}) & \boldsymbol{r}_1=-\boldsymbol{r}_3 \\ O(\frac{1}{T}+\ell_{\lambda,a,n}) & \text{otherwise} \end{cases}$$

where

$$\begin{aligned} & b^{(1)}_{k_2,k_4}(\omega_{k_2},\omega_{k_4};\boldsymbol{r}_1) \\ &= \frac{1}{2\lambda^d}\sum_{\boldsymbol{k}_1=-\boldsymbol{a}}^{\boldsymbol{a}} |g(\boldsymbol{\Omega}_{\boldsymbol{k}_1})|^2 b_{k_4-k_2}\left(\boldsymbol{\Omega}_{\boldsymbol{k}_1},\omega_{k_2};0\right) b_{k_4+k_2}\left(-\boldsymbol{\Omega}_{\boldsymbol{k}_1},-\omega_{k_2};0\right) \\ &+ \frac{1}{2\lambda^d}\sum_{\boldsymbol{k}_1=\max(-\boldsymbol{a},-\boldsymbol{a}-\boldsymbol{r}_1)}^{\min(\boldsymbol{a},\boldsymbol{a}+\boldsymbol{r}_1)} g(\boldsymbol{\Omega}_{\boldsymbol{k}_1})\overline{g(-\boldsymbol{\Omega}_{\boldsymbol{k}_1})} b_{-k_2-k_4}\left(\boldsymbol{\Omega}_{\boldsymbol{k}_1},\omega_{k_2};0\right) b_{k_2+k_4}\left(-\boldsymbol{\Omega}_{\boldsymbol{k}_1},-\omega_{k_2};0\right) \end{aligned}$$

Furthermore, defining $\widehat{A}_{g,h}(\boldsymbol{r}_1,0)$ as in (22) we have

$$\begin{aligned} & \frac{T\lambda^d}{2}\text{cov}[\Re\widehat{A}_{g,h}(\boldsymbol{r}_1,0),\Re\widehat{A}_{g,h}(\boldsymbol{r}_3,0)] \\ \approx & \begin{cases} \frac{1}{2}b + O\left(\frac{\|\boldsymbol{r}_1\|_1}{\lambda}+\frac{1}{T}+\ell_{\lambda,a,n}\right) & \boldsymbol{r}_1=\boldsymbol{r}_3 \\ O\left(\frac{1}{T}+\ell_{\lambda,a,n}\right) & \text{otherwise, except when } \boldsymbol{r}_1=-\boldsymbol{r}_3. \end{cases} \end{aligned}$$

where,

$$
\begin{aligned}
b &= \frac{2}{T}\sum_{k_2,k_4=1}^{T/2} h(\omega_{k_2})\overline{h(\omega_{k_4})}b^{(1)}_{k_2,k_4}(\omega_{k_2},\omega_{k_4};\boldsymbol{r}_1) + \\
&\quad \frac{4}{T^2\lambda^{2d}}\sum_{k_2,k_4=1}^{T/2}\sum_{\boldsymbol{k}_1,\boldsymbol{k}_3=-a}^{a} h(\omega_{k_2})\overline{h(\omega_{k_4})}g(\boldsymbol{\Omega}_{\boldsymbol{k}_1})\overline{g(\boldsymbol{\Omega}_{\boldsymbol{k}_3})}b_{0,4}(\boldsymbol{\Omega}_{\boldsymbol{k}_1},\omega_{k_2},\boldsymbol{\Omega}_{\boldsymbol{k}_3},\omega_{k_3},\omega_{k_2},-\boldsymbol{\Omega}_{\boldsymbol{k}_3},-\omega_{k_3};0)
\end{aligned}
$$

and $b_{0,4}$ is defined in Section 6.2.

Based on the above observations we define the test set $\mathcal{P} = \mathcal{S}\times\{0\}$ (where $\mathcal{S}$ surrounds zero, but is such that if $\boldsymbol{r}_1,\boldsymbol{r}_3 \in \mathcal{S}$ then $\boldsymbol{r}_1 \neq -\boldsymbol{r}_3$). The set over which the orthogonal statistics are defined is $\mathcal{P}' = \mathcal{S}'\times\{0\}$, with $\mathcal{S}\cap\mathcal{S}' = \emptyset$. The DFT covariance is defined as

$$
\widehat{A}_{g,\widehat{V}^{-1/2}}(\boldsymbol{r}_1,0) = \frac{2}{T}\sum_{k=1}^{T/2}\frac{\widehat{a}_g(\omega_k;\boldsymbol{r}_1,0)}{\sqrt{\widehat{V}_g(\omega_k;\mathcal{P}')}}, \text{ where } \widehat{V}_g(\omega_k;\mathcal{P}') = \widehat{\sigma}^2\left(\{\lambda^{d/2}\widehat{a}_g(\omega_k;\boldsymbol{r}_1,0);\boldsymbol{r}_1\in\mathcal{S}'\}\right).
$$

We observe that unlike the spatio-temporal test described in Section 4, in the definition of $\widehat{V}_g(\omega_k;\mathcal{P}')$ we only use frequency ω_k (i.e., we should let $M=0$).

We use $\widehat{A}_{g,\widehat{V}^{-1/2}}(\boldsymbol{r}_1,0)$, defined above, to define the test statistics $\mathbf{T}_{1,g,h}(\mathcal{P},\mathcal{P}')$ and $\mathbf{M}_{1,g,h}(\mathcal{P},\mathcal{P}')$ (see Section 4.2). Note that when testing for spatial stationarity, we have to be careful about using the test statistics $\mathbf{T}_{2,g,\widehat{V}^{-1/2},\widehat{W}}(\mathcal{P},\mathcal{P}')$ and $\mathbf{T}_{2,g,\widehat{V}^{-1/2},1}(\mathcal{P},\mathcal{P}')$. This is because when the process is temporally nonstationary the local average $\widehat{B}_{g,h;H}(\omega_{jH};\boldsymbol{r}_1,r_2)$ is dependent over j.

5.2 Testing for temporal stationarity

Next we consider how to adapt the procedure to test for temporal stationarity. Under the null that $\{Z_t(\boldsymbol{s})\}$ is temporally stationary but not necessarily spatially stationary and using (39) and (40) we have,

$$
\begin{aligned}
&\lambda^d\mathrm{cov}[\Re\widehat{a}_g(\omega_{k_2};\boldsymbol{r}_1,r_2),\Re\widehat{a}_g(\omega_{k_4};\boldsymbol{r}_3,r_4)] \\
&= \begin{cases} b^{(1)}_{r_2,r_2,k_2,k_2}(\omega_{k_2},\omega_{k_2};\boldsymbol{r}_1,\boldsymbol{r}_3) + O\left(\frac{1}{T}+\ell_{\lambda,a,n}\right) & r_2=r_4, k_2=k_4 \\ O\left(\frac{1}{T}+\ell_{\lambda,a,n}\right) & \text{otherwise} \end{cases},
\end{aligned}
$$

where, r_2,r_4,k_2 and k_4 is constrained such that $1\le r_2,r_4,k_2,k_4\le T/2$ and

$$
\begin{aligned}
&b^{(1)}_{r_2,r_2,k_2,k_4}(\omega_{k_2},\omega_{k_2};\boldsymbol{r}_1,\boldsymbol{r}_3) \\
&= \frac{1}{\lambda^d}\sum_{\boldsymbol{k}_1,\boldsymbol{k}_3=-a}^{a} g(\boldsymbol{\Omega}_{\boldsymbol{k}_1})\overline{g(\boldsymbol{\Omega}_{\boldsymbol{k}_3})}b_0(\boldsymbol{\Omega}_{\boldsymbol{k}_1},\omega_{k_2};\boldsymbol{k}_3-\boldsymbol{k}_1)b_0(-\boldsymbol{\Omega}_{\boldsymbol{k}_1+\boldsymbol{r}_1},-\omega_{k_2+r_2};\boldsymbol{k}_3+\boldsymbol{r}_3-\boldsymbol{k}_1-\boldsymbol{r}_1)
\end{aligned}
$$

A similar result holds for $\lambda^d \text{cov}[\Im\widehat{a}_g(\omega_{k_2};\boldsymbol{r}_1,r_2),\Im\widehat{a}_g(\omega_{k_4};\boldsymbol{r}_3,r_4)]$ and cross-covariance term $\lambda^d \text{cov}[\Re\widehat{a}_g(\omega_{k_2};\boldsymbol{r}_1,r_2),\Im\widehat{a}_g(\omega_{k_4};\boldsymbol{r}_3,r_4)]$ is asymptotically zero.

In order to test for temporal stationarity and to avoid the influence of the spatial component we focus on $\boldsymbol{r}_1 = 0$. In this case the above reduces to

$$\lambda^d \text{cov}[\Re\widehat{a}_g(\omega_{k_2};0,r_2),\Re\widehat{a}_g(\omega_{k_4};0,r_4)] = \begin{cases} \frac{1}{2}b^{(1)}(\omega_{k_2}) + O\left(\frac{1+|r_2|}{T} + \ell_{\lambda,a,n}\right) & r_2 = r_4, k_2 = k_4 \\ O\left(\frac{1}{T} + \ell_{\lambda,a,n}\right) & \text{otherwise} \end{cases},$$

where,

$$b^{(1)}(\omega_{k_2}) = \frac{1}{\lambda^d}\sum_{\boldsymbol{k}_1,\boldsymbol{k}_3=-a}^{a} g(\boldsymbol{\Omega}_{\boldsymbol{k}_1})\overline{g(\boldsymbol{\Omega}_{\boldsymbol{k}_3})}|b_0(\boldsymbol{\Omega}_{\boldsymbol{k}_1},\omega_{k_2};\boldsymbol{k}_3-\boldsymbol{k}_1)|^2.$$

And similarly for $\lambda^d \text{cov}[\Im\widehat{a}_g(\omega_{k_2};\boldsymbol{r}_1,r_2),\Im\widehat{a}_g(\omega_{k_4};\boldsymbol{r}_3,r_4)]$. Furthermore, defining $\widehat{A}_{g,h}(0,r_2)$ as in (22) we have,

$$\frac{T\lambda^d}{2}\text{cov}[\Re\widehat{A}_{g,h}(0,r_2),\Re\widehat{A}_{g,h}(0,r_4)] = \begin{cases} \frac{1}{2}c + O\left(\frac{1+|r_2|}{T} + \ell_{\lambda,a,n}\right) & r_2 = r_4 \\ O\left(\frac{1}{T} + \ell_{\lambda,a,n}\right) & \text{otherwise} \end{cases},$$

where,

$$c = \frac{2}{T}\sum_{k_2=1}^{T/2}|h(\omega_{k_2})|^2|b^{(1)}(\omega_{k_2})|^2 + \frac{4}{T^2\lambda^{2d}}\sum_{k_2,k_4=1}^{T/2}\sum_{\boldsymbol{k}_1,\boldsymbol{k}_3=-a}^{a} h(\omega_{k_2})\overline{h(\omega_{k_4})}g(\boldsymbol{\Omega}_{\boldsymbol{k}_1})\overline{g(\boldsymbol{\Omega}_{\boldsymbol{k}_3})} \times b_{0,4}(\boldsymbol{\Omega}_{\boldsymbol{k}_1},\omega_{k_2},\boldsymbol{\Omega}_{\boldsymbol{k}_3},\omega_{k_3},\omega_{k_2},-\boldsymbol{\Omega}_{\boldsymbol{k}_3},-\omega_{k_3};0).$$

Using these observations we use the same test statistics as those described in Section 4. The only differences are that we set $\boldsymbol{r}_1 = \boldsymbol{0}$ when we test for spatial stationarity and use the set $\mathcal{P} = \{\boldsymbol{0}\}\times\mathcal{T}$ (where $\mathcal{T}\subset\mathbb{Z}^+$). We do the same in order to estimate the nuisance parameters $V_g(\omega)$ and $V_{g,\widehat{V}^{-1/2}}$ and $W_{g,\widehat{V}^{-1/2}}(\omega)$ (where $\mathcal{T}'\subset\mathbb{Z}^+$).

6 Auxiliary Results

6.1 Results in the case of stationarity

We first consider the sampling properties of $\widehat{A}_{g,h}(\boldsymbol{r}_1,r_2)$, which is used to define the test statistics $\mathbf{T}_{1,g,\widehat{V}^{-1/2}}(\mathcal{P},\mathcal{P}')$ and $\mathbf{M}_{1,g,\widehat{V}^{-1/2}}(\mathcal{P},\mathcal{P}')$.

Lemma 6.1 *Suppose Assumptions 3.1 and 4.1 hold,* $0 \le r_2, r_4 \le T/2-1$, $\boldsymbol{r}_1 \neq -\boldsymbol{r}_3$ *and*

$h : [0, \pi] \to \mathbb{R}$ is a Lipschitz continuous function. Then

$$\frac{\lambda^d T}{2}\text{cov}\left[\Re\widehat{A}_{g,h}(\boldsymbol{r}_1, r_2), \Re\widehat{A}_{g,h}(\boldsymbol{r}_3, r_4)\right] = I_{\boldsymbol{r}_1=\boldsymbol{r}_3} I_{r_2=r_4} V_{g,h}(\boldsymbol{\Omega}_{\boldsymbol{r}_1}, \omega_{r_2}) + O\bigg(\ell_{\lambda,a,n} + T^{-1}\bigg) \quad (41)$$

$$\frac{\lambda^d T}{2}\text{cov}\left[\Im\widehat{A}_{g,h}(\boldsymbol{r}_1, r_2), \Im\widehat{A}_{g,h}(\boldsymbol{r}_3, r_4)\right] = I_{\boldsymbol{r}_1=\boldsymbol{r}_3} I_{r_2=r_4} V_{g,h}(\boldsymbol{\Omega}_{\boldsymbol{r}_1}, \omega_{r_2}) + O\bigg(\ell_{\lambda,a,n} + T^{-1}\bigg)$$

and $\lambda^d T\text{cov}[\Re\widehat{A}_{g,h}(\boldsymbol{r}_1, r_2), \Im\widehat{A}_{g,h}(\boldsymbol{r}_3, r_4)] = O(\ell_{\lambda,a,n} + T^{-1})$ where,

$$V_{g,h}(\boldsymbol{\Omega}_{\boldsymbol{r}_1}, \omega_{r_2}) = \frac{1}{\pi}\int_0^{\pi} |h(\omega)|^2 V_g(\omega; \boldsymbol{\Omega}_{\boldsymbol{r}_1}, \omega_{r_2}) d\omega + \frac{2}{(2\pi)^d \pi^2}\int_0^{\pi}\int_0^{\pi}\int_{\mathcal{D}^2} g(\boldsymbol{\Omega}_1)\overline{g(\boldsymbol{\Omega}_2)} h(\omega_1)\overline{h(\omega_2)} \times f_4(\boldsymbol{\Omega}_1 + \boldsymbol{\Omega}_{\boldsymbol{r}_1}, \omega_1 + \omega_{r_2}, \boldsymbol{\Omega}_2, \omega_2, -\boldsymbol{\Omega}_2 - \boldsymbol{\Omega}_{\boldsymbol{r}_1}, -\omega_2 - \omega_{r_2}) d\boldsymbol{\Omega}_1 d\boldsymbol{\Omega}_2 d\omega_1 d\omega_2.$$

Let $\{(\boldsymbol{r}_j, r_j); 1 \leq j \leq m\}$ be a collection of integer vectors constrained such that $0 \leq r_j \leq T/2 - 1$ and $\boldsymbol{r}_{j_1} \neq -\boldsymbol{r}_{j_2}$. Then under stationarity of $\{Z_t(\boldsymbol{s})\}$ and sufficient mixing conditions we have,

$$\sqrt{\frac{\lambda^d T}{2}}\left[\frac{\Re\widehat{A}_{g,h}(\boldsymbol{r}_{j_1}, r_{j_1})}{V_{g,h}(\boldsymbol{\Omega}_{\boldsymbol{r}_{j_1}}, \omega_{r_{j_1}})^{1/2}}, \frac{\Im\widehat{A}_{g,h}(\boldsymbol{r}_{j_1}, r_{j_1})}{V_{g,h}(\boldsymbol{\Omega}_{\boldsymbol{r}_{j_1}}, \omega_{r_{j_1}})^{1/2}}, \ldots, \frac{\Re\widehat{A}_{g,h}(\boldsymbol{r}_{j_m}, r_{j_m})}{V_{g,h}(\boldsymbol{\Omega}_{\boldsymbol{r}_{j_m}}, \omega_{r_{j_m}})^{1/2}}, \frac{\Im\widehat{A}_{g,h}(\boldsymbol{r}_{j_m}, r_{j_m})}{V_{g,h}(\boldsymbol{\Omega}_{\boldsymbol{r}_{j_m}}, \omega_{r_{j_m}})^{1/2}}\right] \xrightarrow{\mathcal{D}} \mathcal{N}(0, I_{2m}).$$

We note that when $\|\boldsymbol{r}_1\|_1 << \lambda$ and $|r_2| << T$ that the variances above approximate to

$$V_{g,h}(\boldsymbol{\Omega}_{\boldsymbol{r}_1}, \omega_{r_2}) = V_{g,h}(0, 0) + O\left(\frac{\|\boldsymbol{r}_1\|_1}{\lambda} + \frac{|r_2|}{T}\right).$$

We now consider the sampling properties of $\widehat{B}_{g,h;H}(\boldsymbol{r}_1, r_2)$ and $\widehat{D}_{g,h,v;H}(\boldsymbol{r}_1, r_2)$, which are used to define the test statistics $\mathbf{T}_{2,g,\widehat{V}^{-1/2},W}(\mathcal{P}, \mathcal{P}')$ and $\mathbf{M}_{2,g,\widehat{V}^{-1/2},W}(\mathcal{P}, \mathcal{P}')$. We start by studying $\widehat{B}_{g,h;H}(\boldsymbol{r}_1, r_2)$.

Lemma 6.2 *Suppose Assumptions 3.1 and 4.1 hold, $0 \leq r_2, r_4 \leq T/2 - 1$, $\boldsymbol{r}_1 \neq -\boldsymbol{r}_3$ and $h : [0, \pi] \to \mathbb{R}$ is a Lipschitz continuous function. Then*

$$\lambda^d H\text{cov}\left[\Re\widehat{B}_{g,h;H}(\omega_{jH}; \boldsymbol{r}_1, r_2), \Re\widehat{B}_{g,h;H}(\omega_{jH}; \boldsymbol{r}_3, r_4)\right] = I_{\boldsymbol{r}_1=\boldsymbol{r}_3} I_{r_2=r_4} W_{g,h}(\omega_{jH}; \boldsymbol{\Omega}_{\boldsymbol{r}_1}, \omega_{r_2}) + O\bigg(\ell_{\lambda,a,n} + H^{-1}\bigg). \quad (42)$$

Exactly the same result holds for $\lambda^d H\text{cov}\left[\Im\widehat{B}_{g,h;H}(\omega_{jH};\boldsymbol{r}_1,r_2),\Im\widehat{B}_{g,h;H}(\omega_{jH};\boldsymbol{r}_3,r_4)\right]$, *where*

$$
\begin{aligned}
&2W_{g,h}(\omega_{jH};\boldsymbol{\Omega}_{\boldsymbol{r}_1},\omega_{r_2}) \\
&= \frac{T}{H\pi}\int_{\omega_{jH}}^{\omega_{(j+1)H}}|h(\omega)|^2 V_g(\omega;\boldsymbol{\Omega}_{\boldsymbol{r}_1},\omega_{r_2})d\omega + \frac{T}{H(2\pi)^{2d+2}}\int_{\omega_{jH}}^{\omega_{(j+1)H}}\int_{\omega_{jH}}^{\omega_{(j+1)H}}\int_{\mathcal{D}^2} g(\boldsymbol{\Omega}_1)\overline{g(\boldsymbol{\Omega}_2)} \\
&\times h(\omega_1)\overline{h(\omega_2)}f_4(\boldsymbol{\Omega}_1+\boldsymbol{\Omega}_{r_1},\omega_1+\omega_{r_2},\boldsymbol{\Omega}_2,\omega_2,-\boldsymbol{\Omega}_2-\boldsymbol{\Omega}_{\boldsymbol{r}_1},-\omega_2-\omega_{r_2})d\boldsymbol{\Omega}_1 d\boldsymbol{\Omega}_2 d\omega_1 d\omega_2,
\end{aligned}
$$

noting that the first (covariance) term in $W_{g,h}=O(1)$, *whereas the second term of* $W_{g,h}$ *which is the fourth order cumulant term is of order* $O(H/T)$ *since the cumulant term involves a double integral which is of order* $O((H/T)^2)$. *On the other hand, if* $(\boldsymbol{r}_1,r_2),(\boldsymbol{r}_3,r_4)\neq 0$, *then (with* $0\leq r_2,r_4<T/2$ *and* $(\boldsymbol{r}_1,r_2)\neq-(\boldsymbol{r}_3,r_4)$*) we have* $\lambda^d H\text{cov}\left[\Re\widehat{B}_{g,h;H}(\omega_{jH};\boldsymbol{r}_1,r_2),\Im\widehat{B}_{g,h;H}(\omega_{jH};\boldsymbol{r}_3,r_4)\right]$ $O(\ell_{\lambda,a,n})$ *and for* $j_1\neq j_2$,

$$
\lambda^d H\text{cov}\left[\Re\widehat{B}_{g,h;H}(\omega_{j_1H};\boldsymbol{r}_1,r_2),\Re\widehat{B}_{g,h;H}(\omega_{j_2H};\boldsymbol{r}_3,r_4)\right] = O\left(\ell_{\lambda,a,n}+H^{-1}+\frac{H}{T}\right), \quad (43)
$$

where the same holds for $\lambda^d H\text{cov}\left[\Im\widehat{B}_{g,h;H}(\omega_{j_1H};\boldsymbol{r}_1,r_2),\Im\widehat{B}_{g,h;H}(\omega_{j_2H};\boldsymbol{r}_3,r_4)\right]$ *and also for* $\lambda^d H\text{cov}\left[\Re\widehat{B}_{g,h;H}(\omega_{j_1H};\boldsymbol{r}_1,r_2),\Im\widehat{B}_{g,h;H}(\omega_{j_2H};\boldsymbol{r}_3,r_4)\right]$.

Let $\{(k_j,\boldsymbol{r}_{1,i},r_{2,i});1\leq j\leq m,(\boldsymbol{r}_1,r_2)\in\mathcal{P}\}$ *be a collection of integer vectors constrained such that,* $1\leq k_j\leq T/2$, $1\leq r_j\leq T/2$ *and* $\boldsymbol{r}_{j_1}\neq-\boldsymbol{r}_{j_2}$. *Then under stationarity of* $\{Z_t(\boldsymbol{s})\}$ *and sufficient mixing conditions we have,*

$$
\begin{aligned}
&\sqrt{\lambda^d H}\left[\frac{\Re\widehat{B}_{g,h}(\omega_{k_jH};\boldsymbol{r}_{1,i},r_{2,i})}{W_{g,h}(\omega_{k_jH})^{1/2}},\frac{\Im\widehat{B}_{g,h}(\omega_{k_jH};\boldsymbol{r}_{1,i},r_{2,i})}{W_{g,h}(\omega_{k_jH})^{1/2}},1\leq j\leq m,(\boldsymbol{r}_{1,i},r_{2,i})\in\mathcal{P}\right] \\
&\qquad\overset{\mathcal{D}}{\to}\mathcal{N}\left(0,I_{2m|\mathcal{P}|}\right),
\end{aligned}
$$

where $W_{g,h}(\omega)=W_{g,h}(\omega;0,0)$.

In the following lemma we consider the sampling properties of $\widehat{D}_{g,h,v;H}(\boldsymbol{r}_1,r_2)$. Note that we consider general functions v, whereas in Section 4.3 we set v to be the variance of $W_{g,h}(\omega)$, which means the mean of $\widehat{D}$ is asymptotically pivotal.

Lemma 6.3 *Suppose the assumptions in Lemma 6.2 hold and* $h:[0,\pi]\to\mathbb{R}$ *is a Lipschitz continuous function. Then we have*

$$
\begin{aligned}
&\text{E}[H\lambda^d\widehat{D}_{g,h,v;H}(\boldsymbol{r}_1,r_2)] \\
=\;& E_{g,h,v}(\boldsymbol{\Omega}_{\boldsymbol{r}_1},\omega_{r_2})+O\left(\ell_{\lambda,a,n}+\frac{1}{H}+\frac{H}{T}+\frac{\lambda^d H[\prod_{j=1}^{d-b}(\log\lambda+\log|m_j|)]^2}{(T^{I_{r_2-r_4\neq 0}}\lambda^{d-b})^2}\right) \quad (44)
\end{aligned}
$$

and

$$
\begin{aligned}
& \frac{T}{2H}\text{cov}\left[\lambda^d H\widehat{D}_{g,h,v;H}(\boldsymbol{r}_1,r_2),\lambda^d H\widehat{D}_{g,h,v;H}(\boldsymbol{r}_3,r_4)\right] \\
&= \begin{cases} U_{g,h,v}(\boldsymbol{\Omega}_{\boldsymbol{r}_1},\omega_{r_2})+O\left(\frac{H}{T}+\ell_{\lambda,a,n}\right) & \boldsymbol{r}_1=\boldsymbol{r}_3 \text{ and } r_2=r_4 \\ O\left(\frac{H}{T}+\ell_{\lambda,a,n}\right) & \text{otherwise} \end{cases} \qquad (45)
\end{aligned}
$$

where,

$$
\begin{aligned}
E_{g,h,v}(\boldsymbol{\Omega}_{\boldsymbol{r}_1},\omega_{r_2}) &= \frac{1}{\pi}\int_0^{\pi}\frac{W_{g,h}(\omega;\boldsymbol{\Omega}_{\boldsymbol{r}_1},\omega_{r_2})}{v(\omega)}d\omega \\
\text{and } U_{g,h,v}(\boldsymbol{\Omega}_{\boldsymbol{r}_1},\omega_{r_2}) &= \frac{1}{\pi}\int_0^{\pi}\frac{|W_{g,h}(\omega;\boldsymbol{\Omega}_{\boldsymbol{r}_1},\omega_{r_2})|^2}{|v(\omega)|^2}d\omega.
\end{aligned}
$$

Let $\{(\boldsymbol{r}_j,r_j);1\leq j\leq m\}$ *be a collection of integer vectors constrained such that* $1\leq k_j\leq T/2$, $1\leq r_j\leq T/2$ *and* $\boldsymbol{r}_{j_1}\neq-\boldsymbol{r}_{j_2}$, *then we have*

$$
\sqrt{\frac{T}{2HU_{g,h,v}}}\begin{pmatrix} \lambda^d H\widehat{D}_{g,h,v;H}(\boldsymbol{r}_{j_1},r_{j_1})-E_{g,h,v} \\ \vdots \\ \lambda^d H\widehat{D}_{g,h,v;H}(\boldsymbol{r}_{j_m},r_{j_m})-E_{g,h,v} \end{pmatrix} \overset{\mathcal{D}}{\to}\mathcal{N}(0,I_{2m})
$$

where $U_{g,h,v}=U_{g,h,v}(0,0)$ *and* $E_{g,h,v}=E_{g,h,v}(0,0)$.

6.2 Results in the case of nonstationarity

We first generalize (13) from covariances to fourth order cumulants. We assume there exists a function κ such that

$$
\begin{aligned}
& \text{cov}[Z_{t,\lambda,T}(\boldsymbol{s}),Z_{t+h_1,\lambda,T}(\boldsymbol{s}+\boldsymbol{v}_1),Z_{t+h_2,\lambda,T}(\boldsymbol{s}+\boldsymbol{v}_2),Z_{t+h_3,\lambda,T}(\boldsymbol{s}+\boldsymbol{v}_3)] \\
&= \kappa_{h_1,h_2,h_3;\frac{t}{T}}\left(\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3;\frac{\boldsymbol{s}}{\lambda}\right)+O\left(\frac{\prod_{i=1}^3\beta_{2+\delta}(\boldsymbol{v}_i)\rho_{h_i}}{T}\right), \qquad (46)
\end{aligned}
$$

where, $\sup_{u,\boldsymbol{s}}|\kappa_{h_1,h_2,h_3;u}(\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3;\boldsymbol{s})|\leq\prod_{i=1}^3\rho_{h_i}\beta_{2+\delta}(\boldsymbol{v}_i)$,

$$
\begin{aligned}
|\kappa_{h_1,h_2,h_3;u_1}(\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3;\boldsymbol{s})-\kappa_{h_1,h_2,h_3;u_2}(\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3;\boldsymbol{s})| &\leq |u_1-u_2|\prod_{i=1}^3\beta_{2+\delta}(\boldsymbol{v}_i)\rho_{h_i}, \\
|\kappa_{h_1,h_2,h_3;u}(\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3;\boldsymbol{s}_1)-\kappa_{h_1,h_2,h_3;u}(\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3;\boldsymbol{s}_2)| &\leq |\boldsymbol{s}_1-\boldsymbol{s}_2|_1\prod_{i=1}^3\beta_{2+\delta}(\boldsymbol{v}_i)\rho_{h_i}.
\end{aligned}
$$

Using the above, we define the location and time dependent fourth order spectral density as

$$F_{u,4}\left(\boldsymbol{\Omega}_1,\omega_1,\boldsymbol{\Omega}_2,\omega_2,\boldsymbol{\Omega}_3,\omega_3;\boldsymbol{s}\right) = \frac{1}{(2\pi)^3}\sum_{h_1,h_2,h_3\in\mathbb{Z}} e^{-i(h_1\omega_1+h_2\omega_2+h_3\omega_3)}\int_{\mathbb{R}^{3d}}\kappa_{h_1,h_2,h_3;u}\left(\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3;\boldsymbol{s}\right) \times e^{-i(\boldsymbol{v}_1'\boldsymbol{\Omega}_1+\boldsymbol{v}_2'\boldsymbol{\Omega}_2+\boldsymbol{v}_3'\boldsymbol{\Omega}_3)}d\boldsymbol{v}_1 d\boldsymbol{v}_2 d\boldsymbol{v}_3. \tag{47}$$

In order to obtain the expressions below we start by generalizing the covariance result in (15) to fourth order cumulants. By using Lee and Subba Rao [2015] and similar methods to those used in Bandyopadhyay and Subba Rao [2016] it can be shown that

$$\begin{aligned}&\mathrm{cum}[J(\boldsymbol{\Omega}_{\boldsymbol{k}_1},\omega_{k_2}),\overline{J(\boldsymbol{\Omega}_{\boldsymbol{k}_1+\boldsymbol{r}_1},\omega_{k_2+r_2})},\overline{J(\boldsymbol{\Omega}_{\boldsymbol{k}_3},\omega_{k_4})},J(\boldsymbol{\Omega}_{\boldsymbol{k}_3+\boldsymbol{r}_3},\omega_{k_4+r_4})]\\&=\frac{1}{T\lambda^d}b_{r_2-r_4,4}(\boldsymbol{\Omega}_{\boldsymbol{k}_1+\boldsymbol{r}_1},\omega_{k_2+r_2},\boldsymbol{\Omega}_{\boldsymbol{k}_3},\omega_{k_4},\boldsymbol{\Omega}_{-\boldsymbol{k}_3-\boldsymbol{r}_3},\omega_{-k_4-r_4};\boldsymbol{r}_1-\boldsymbol{r}_3)+O\left(\frac{1}{T^2\lambda^d}+\frac{1}{T\lambda^{d+1}}\right)\end{aligned}$$

where,

$$b_{r_2,4}(\boldsymbol{\Omega}_1,\omega_1,\boldsymbol{\Omega}_2,\omega_2,\boldsymbol{\Omega}_3,\omega_3;\boldsymbol{r}_1) = \int_0^1\int_{[-1/2,1/2]^d}F_{u,4}\left(\boldsymbol{\Omega}_1,\omega_1,\boldsymbol{\Omega}_2,\omega_2,\boldsymbol{\Omega}_3,\omega_3;\boldsymbol{s}\right)e^{-2\pi i\boldsymbol{r}_1'\boldsymbol{s}}e^{-2\pi i r_2 u}d\boldsymbol{s}du,$$

and $F_{u,4}$ is defined in (47).

Lemma 6.4 *Suppose the assumptions in Assumptions 3.1 and 4.1 (generalized to the non-stationary set-up) and (46) are satisfied. Then we have,*

$$\lambda^d\mathrm{cov}\left[\widehat{a}_g(\omega_{k_2};\boldsymbol{r}_1,r_2),\widehat{a}_g(\omega_{k_4};\boldsymbol{r}_3,r_4)\right] = b^{(1)}_{r_2,r_4,k_2,k_4}(\omega_{k_2},\omega_{k_4};\boldsymbol{r}_1,\boldsymbol{r}_3)+O\left(\frac{1}{T}+\frac{1}{\lambda}\right),$$

and

$$\lambda^d\mathrm{cov}\left[\widehat{a}_g(\omega_{k_2};\boldsymbol{r}_1,r_2),\overline{\widehat{a}_g(\omega_{k_4};\boldsymbol{r}_3,r_4)}\right] = b^{(2)}_{r_2,r_4,k_2,k_4}(\omega_{k_2},\omega_{k_4};\boldsymbol{r}_1,\boldsymbol{r}_3)+O\left(\frac{1}{T}+\frac{1}{\lambda}+\ell_{\lambda,a,n}\right).$$

We also have,

$$\frac{T\lambda^d}{2}\mathrm{cov}\left[\widehat{A}_{g,h}(\boldsymbol{r}_1,r_2),\widehat{A}_{g,h}(\boldsymbol{r}_3,r_4)\right]$$

$$= \frac{2}{T\lambda^d}\sum_{k_2,k_4=1}^{T/2} h(\omega_{k_2})\overline{h(\omega_{k_4})}b^{(1)}_{r_2,r_4,k_2,k_4}(\omega_{k_2},\omega_{k_4};\boldsymbol{r}_1,\boldsymbol{r}_3)$$

$$+\frac{2}{T^2\lambda^{2d}}\sum_{k_2,k_4=1}^{T/2}\sum_{\boldsymbol{k}_1,\boldsymbol{k}_3=-a}^{a} h(\omega_{k_2})\overline{h(\omega_{k_4})}g(\boldsymbol{\Omega}_{\boldsymbol{k}_1})\overline{g(\boldsymbol{\Omega}_{\boldsymbol{k}_3})}$$

$$\times b_{r_2-r_4,4}(\boldsymbol{\Omega}_{\boldsymbol{k}_1+\boldsymbol{r}_1},\omega_{k_2+r_2},\boldsymbol{\Omega}_{\boldsymbol{k}_3},\omega_{k_4},\boldsymbol{\Omega}_{-\boldsymbol{k}_3-\boldsymbol{r}_2},\omega_{-k_4-r_4};\boldsymbol{r}_1-\boldsymbol{r}_3)+O\left(\frac{1}{T}+\frac{1}{\lambda}+\ell_{\lambda,a,n}\right),$$

and

$$\frac{T\lambda^d}{2}\mathrm{cov}\left[\widehat{A}_{g,h}(\boldsymbol{r}_1,r_2),\overline{\widehat{A}_{g,h}(\boldsymbol{r}_3,r_4)}\right]$$

$$= \frac{2}{T\lambda^d}\sum_{k_2,k_4=1}^{T/2} h(\omega_{k_2})h(\omega_{k_4})b^{(2)}_{r_2,r_4,k_2,k_4}(\omega_{k_2},\omega_{k_4};\boldsymbol{r}_1,\boldsymbol{r}_3)$$

$$+\frac{2}{T^2\lambda^{2d}}\sum_{k_2,k_4=1}^{T/2}\sum_{\boldsymbol{k}_1,\boldsymbol{k}_3=-a}^{a} h(\omega_{k_2})h(\omega_{k_4})g(\boldsymbol{\Omega}_{\boldsymbol{k}_1})g(\boldsymbol{\Omega}_{\boldsymbol{k}_3})$$

$$\times b_{r_2+r_4,4}(\boldsymbol{\Omega}_{\boldsymbol{k}_1+\boldsymbol{r}_1},\omega_{k_2+r_2},\boldsymbol{\Omega}_{-\boldsymbol{k}_3},\omega_{-k_3},\boldsymbol{\Omega}_{\boldsymbol{k}_3+\boldsymbol{r}_3},\omega_{k_4+r_4};\boldsymbol{r}_1+\boldsymbol{r}_3)+O\left(\frac{1}{T}+\frac{1}{\lambda}+\ell_{\lambda,a,n}\right).$$

7 Simulations

7.1 Set-up

We now assess the finite sample performances of the test statistics described above through simulations. In all cases we consider mean zero spatio-temporal processes, where $T = 200$ and at each time point we observe $n = 100$ or 500 locations (the locations are drawn from a uniform distribution defined on $[-\lambda/2,\lambda/2]^2$ and we use the same set of locations at each time point). All tests are done at the 5% level and all results are based on 300 replications. Further, we investigate the performances of the tests when the coefficients $\widehat{a}_g(\omega_k;\boldsymbol{r}_1,r_2)$, as defined in Section 4.1, are being calculated while both removing and keeping the 'nugget effect' N_T (in Tables 1-5 the rejection rates for the test statistics without removing N_T are reported in the parentheses). All simulations are done for spatial dimension $d = 2$.

Next, we briefly discuss the implementation issues.

1. *Choice of set $\mathcal{P}$ and $\mathcal{P}'$*: All test statistics depend on the choice of $\mathcal{P}$ and $\mathcal{P}'$. In all simulations described in this section we use $\mathcal{P} = \{(1,0),(1,1),(0,1),(-1,1)\}\times$

$\{1,2\}$ and $\mathcal{P}' = \{(1,0),(1,1),(0,1),(-1,1)\} \times \{4,5\}$ to calculate empirical type I errors and overall powers. Further, to test for stationarity over space, we take $\mathcal{P} = \{(1,0),(1,1),(0,1),(-1,1)\} \times \{0\}$ and $\mathcal{P}' = \{(2,0),(2,1),(2,2),(1,2),(0,2),(-1,2),(-2,2),(-2,1)\} \times \{0\}$ and to test for stationarity over time, we take $\mathcal{P} = \{(0,0)\} \times \{1,2\}$ and $\mathcal{P}' = \{(0,0)\} \times \{4,5\}$.

2. *Choice of* $g(\cdot)$: Based on the discussion in Remark 2.1 we use the weight function $g(\mathbf{\Omega}) = \sum_{j=1}^{L} e^{i\boldsymbol{v}_j'\mathbf{\Omega}}$. The choice of $\boldsymbol{v}_j$'s should depend on the density of the sampling region. Following the same rationale as described in Bandyopadhyay and Subba Rao [2016], in all simulations we define the v grid as $\mathbb{V} = \{\boldsymbol{v}_j = (v_{j1}, v_{j2})' \in \mathbb{R}^d : v_{jk} = -s, -s/2, 0, s/2, s, \text{ for } k = 1,2\}$ such that $\boldsymbol{v}_j + \boldsymbol{v}_{j'} \neq \mathbf{0}$ for $\boldsymbol{v}_j, \boldsymbol{v}_{j'} \in \mathbb{V}$, where $s = \lambda/n^{1/d}$ is the 'average spacing' between the observations on each axis. We should mention that if the support of the empirical covariance of the data appears far greater than $s = \lambda/n^{1/d}$, then using a wider v grid is appropriate. If changes in the spatial covariance function happen mainly at lags much smaller than $s = \lambda/n^{1/d}$, then data is not available to detect changes in the spatial covariance structure.

3. *Choice of frequency grid*: In all simulations we use $a = \sqrt{n}$ in the definition of $\widehat{a}_g(\omega_k; \boldsymbol{r}_1, r_2)$.

4. *Choice of* H *in the definition of* $\mathbf{T}_2$ *and* $\mathbf{M}_2$: For all simulations we use $H = 10$ and $H = 20$.

5. *Choice of* M *to calculate the local averages*: In order to estimate $V_g(\omega_k)$ we use the estimator $\widehat{V}_g(\omega_k; \mathcal{P}')$ (defined in (28)) with $M = 2$ (thus taking a local average of 5).

To obtain the critical values of the tests we use the asymptotic 'finite sample' approximations of the distributions of the test statistics as described in Section 4.4. For ease of discussion below we refer to (i) $\mathbf{T}_{1,g,\widehat{V}^{-1/2}}$ and $\mathbf{M}_{1,g,\widehat{V}^{-1/2}}$ as the average covariance test statistics (ii) $\mathbf{T}_{2,g,\widehat{V}^{-1/2},1}$ and $\mathbf{M}_{2,g,\widehat{V}^{-1/2},1}$ as the average squared covariance test statistics and (iii) $\mathbf{T}_{2,g,\widehat{V}^{-1/2},\widehat{W}}$ and $\mathbf{M}_{2,g,\widehat{V}^{-1/2},\widehat{W}}$ as the variance adjusted average squared covariance test statistics.

7.2 Simulations under the null

7.2.1 Models

In order to define the spatio-temporal models, we start by defining the 'innovations' process. Let $\{\varepsilon_t(\boldsymbol{s}); \boldsymbol{s} \in \mathbb{R}^2, t \in \mathbb{Z}\}$ denote a spatio-temporal stationary Gaussian random field which is independent over time with spatial exponential covariance $\text{cov}[\varepsilon_t(\boldsymbol{s}_1), \varepsilon_t(\boldsymbol{s}_2)] = \kappa_0(\|\boldsymbol{s}_2 - \boldsymbol{s}_1\|_2) = \exp(-\|\boldsymbol{s}_1 - \boldsymbol{s}_2\|_2/\rho)$, where ρ is the 'range parameter'. We do all the simulations under the null with $\rho = 0.5$, $\rho = 1$ and $\lambda = 5$ (in the case that $\rho = 1$ the range of dependence for the innovations is 20%, whereas for $\rho = 0.5$ the range reduces to 10%; see Figure 1). We mention that the decorrelation property of Fourier Transforms, given in Lemma 3.1 implicitly depends on the range of dependence with respect to λ. If the range of dependence is too large with respect to the observed random field then the degree of correlation in the Fourier transforms will be non-negligible (leading to false rejection of the null).

(S1) *Spatially and temporally stationary Gaussian random field*: We define a spatio-temporal model with the temporal AR(1) structure $Z_t(\boldsymbol{s}) = 0.5Z_{t-1}(\boldsymbol{s}) + \varepsilon_t(\boldsymbol{s})$.

(S2) *Spatially and temporally stationary non-Gaussian random field:* To induce non-linearity and non-Gaussianity in the random field we use a Bilinear model of the form

$$Z_t(\boldsymbol{s}) = 0.5Z_{t-1}(\boldsymbol{s}) + 0.4Z_{t-1}(\boldsymbol{s})\varepsilon_{t-1}(\boldsymbol{s}) + \varepsilon_t(\boldsymbol{s}).$$

We note that the nonlinear term $0.4Z_{t-1}(\boldsymbol{s})\varepsilon_{t-1}(\boldsymbol{s})$ induces sporadic bursts in the spatio-temporal process. The coefficients 0.5 and 0.4 are chosen to ensure that the process has a finite second moment (see Subba Rao and Gabr [1984] for details).

7.2.2 Discussion

The results for model S1 and S2 are given Table 1.

We first consider the stationary Gaussian model (S1). The results for all the tests are relatively good for both $\rho = 0.5$ and $\rho = 1$. However, for the average squared statistics (without variance adjustment) for $H = 10$ and $\rho = 1$ there are some inflations in the type I error. This is probably because without the variance adjustment the average squared statistics depend on the asymptotic result $W_{g,\widehat{V}^{-1/2}}(\omega_{jH}) \xrightarrow{\mathcal{P}} 1$ (see (36)) which depends on the range parameter ρ and the degree of non-Gaussianity. Model S1 is Gaussian, and it seems the error in this approximation seems only to mildly impact the case $H = 10$ and $\rho = 1$.

The results from the simulations for the stationary but non-Gaussian model (S2) are very different. The average covariance test results keep close to the nominal level (for both $\rho = 0.5$ and 1) however there is a substantial inflation in the type I error (between 70-90%) for the average squared statistic without variance adjustment (for both $H = 10$ and $H = 20$). This is likely due to the non-Gaussianity of the process which seems to greatly impact the rate that $W_{g,\widehat{V}^{-1/2}}(\omega_{jH}) \xrightarrow{\mathcal{P}} 1$. However, the variance adjusted average squared covariance test statistics appear to keep close to the nominal level for both $\rho = 0.5$ and 1 and $H = 10$ and 20. This demonstrates that $W_{g,\widehat{V}^{-1/2}}(\omega_{jH}) \xrightarrow{\mathcal{P}} 1$ is an *asymptotic* result and for finite samples it is important to estimate the variance.

In all cases, both removing and keeping the nugget term N_T give comparable results.

Our results in the simulation study demonstrate that both the average covariance and the variance adjusted average squared covariance test statistics perform well under the null, but caution needs to be taken when interpreting the results of the non-variance adjusted average squared covariance tests.

		Model S1				Model S2			
		ρ				ρ			
		0.5		1		0.5		1	
	n	100	500	100	500	100	500	100	500
	$\mathbf{T}_{1,g,\widehat{V}^{-1/2}}$	0.08 (0.08)	0.08 (0.07)	0.07 (0.07)	0.04 (0.04)	0.09 (0.09)	0.08 (0.09)	0.09 (0.08)	0.08 (0.07)
	$\mathbf{M}_{1,g,\widehat{V}^{-1/2}}$	0.04 (0.07)	0.05 (0.06)	0.05 (0.06)	0.02 (0.02)	0.06 (0.04)	0.05 (0.06)	0.07 (0.07)	0.06 (0.08)
H=20	$\mathbf{T}_{2,g,\widehat{V}^{-1/2},1}$	0.04 (0.03)	0.01 (0.01)	0.06 (0.04)	0.03 (0.02)	0.45 (0.70)	0.86 (0.91)	0.62 (0.72)	0.94 (0.98)
	$\mathbf{M}_{2,g,\widehat{V}^{-1/2},1}$	0.07 (0.08)	0.06 (0.05)	0.07 (0.08)	0.02 (0.03)	0.48 (0.67)	0.80 (0.88)	0.60 (0.70)	0.86 (0.88)
	$\mathbf{T}_{2,g,\widehat{V}^{-1/2},\widehat{W}}$	0.05 (0.05)	0.02 (0.01)	0.07 (0.03)	0.02 (0.01)	0.06 (0.07)	0.05 (0.04)	0.05 (0.05)	0.06 (0.05)
	$\mathbf{M}_{2,g,\widehat{V}^{-1/2},\widehat{W}}$	0.05 (0.06)	0.03 (0.03)	0.05 (0.05)	0.06 (0.05)	0.04 (0.07)	0.06 (0.07)	0.10 (0.09)	0.08 (0.08)
H=10	$\mathbf{T}_{2,g,\widehat{V}^{-1/2},1}$	0.10 (0.10)	0.08 (0.07)	0.12 (0.12)	0.13 (0.14)	0.50 (0.71)	0.85 (0.90)	0.67 (0.78)	0.85 (0.88)
	$\mathbf{M}_{2,g,\widehat{V}^{-1/2},1}$	0.08 (0.08)	0.09 (0.08)	0.11 (0.12)	0.16 (0.15)	0.42 (0.58)	0.65 (0.77)	0.53 (0.67)	0.79 (0.83)
	$\mathbf{T}_{2,g,\widehat{V}^{-1/2},\widehat{W}}$	0.10 (0.08)	0.04 (0.04)	0.06 (0.04)	0.04 (0.04)	0.09 (0.10)	0.06 (0.06)	0.05 (0.06)	0.08 (0.10)
	$\mathbf{M}_{2,g,\widehat{V}^{-1/2},\widehat{W}}$	0.05 (0.06)	0.03 (0.04)	0.06 (0.05)	0.05 (0.05)	0.08 (0.10)	0.09 (0.09)	0.10 (0.13)	0.11 (0.12)

Table 1: Empirical type I errors at 5% level based on different tests with $\lambda = 5$ for Gaussian and non-Gaussian stationary data with innovations coming from a Gaussian random field with exponential covariance functions. Rejection rate without removing N_T (see (18)) are in the parentheses.

7.3 Simulations under the alternative

7.3.1 Models

In order to induce spatial nonstationarity in the models (NS2) and (NS3) (defined below) we define the Gaussian innovations process $\{\eta_t(\boldsymbol{s}); \boldsymbol{s} \in [-\lambda/2, \lambda/2]^2\}$, which is independent over time with nonstationary covariance $\mathrm{cov}[\eta_t(\boldsymbol{s}_1), \eta_t(\boldsymbol{s}_2)] = c_\lambda(\boldsymbol{s}_1, \boldsymbol{s}_2) = \kappa_0(\boldsymbol{s}_2 - \boldsymbol{s}_1; \boldsymbol{s}_1)$ where,

$$c_\lambda(\boldsymbol{s}_1, \boldsymbol{s}_2) = |\Sigma\left(\frac{\boldsymbol{s}_1}{\lambda}\right)|^{1/4}|\Sigma\left(\frac{\boldsymbol{s}_2}{\lambda}\right)|^{1/4}\left|\frac{\Sigma(\frac{\boldsymbol{s}_1}{\lambda}) + \Sigma(\frac{\boldsymbol{s}_2}{\lambda})}{2}\right|^{-1/2} \exp[-\sqrt{Q_\lambda(\boldsymbol{s}_1, \boldsymbol{s}_2)}],$$

$|\cdot|$ denotes the determinant of a matrix, $Q_\lambda(\boldsymbol{s}_1, \boldsymbol{s}_2) = 2(\boldsymbol{s}_1 - \boldsymbol{s}_2)'[\Sigma(\frac{\boldsymbol{s}_1}{\lambda}) + \Sigma(\frac{\boldsymbol{s}_2}{\lambda})]^{-1}(\boldsymbol{s}_1 - \boldsymbol{s}_2)$ and $\Sigma(\frac{\boldsymbol{s}}{\lambda}) = \Gamma(\frac{\boldsymbol{s}}{\lambda})\Lambda\Gamma(\frac{\boldsymbol{s}}{\lambda})'$, where

$$\Gamma\left(\frac{\boldsymbol{s}}{\lambda}\right) = \begin{bmatrix} \gamma_1(\boldsymbol{s}/\lambda) & -\gamma_2(\boldsymbol{s}/\lambda) \\ \gamma_2(\boldsymbol{s}/\lambda) & \gamma_1(\boldsymbol{s}/\lambda) \end{bmatrix}, \ \Lambda = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{bmatrix},$$

with $\gamma_1(\boldsymbol{s}/\lambda) = \log(s_x/\lambda + 0.75)$, $\gamma_2(\boldsymbol{s}/\lambda) = (s_x/\lambda)^2 + (s_y/\lambda)^2$, and $\boldsymbol{s} = (s_x, s_y)'$ (see Paciorek and Schervish [2006] and Jun and Genton [2012] for the details on this process). Note that the variance of this process is constant over the spatial random field and it is simply the correlation structure that varies over space.

(NS1) *Temporally nonstationary but spatially stationary Gaussian random field:* $Z_t(\boldsymbol{s}) = 0.5Z_{t-1}(\boldsymbol{s}) + \left(1.3 + \sin\left(\frac{2\pi t}{400}\right)\right)\varepsilon_t(\boldsymbol{s})$, where $\{\varepsilon_t(\boldsymbol{s})\}$ is defined in Section 7.2.1. We use $\rho = 0.5$, $\rho = 1$ and $\lambda = 5$.

(NS2) *Temporally stationary but spatially nonstationary Gaussian random field:* The spatio-temporal process is defined with an AR(1) model $Z_t(\boldsymbol{s}) = 0.5Z_{t-1}(\boldsymbol{s}) + \eta_t(\boldsymbol{s})$. Following a similar set-up as in Bandyopadhyay and Subba Rao [2016] we use $\lambda = 20$. This process has a constant variance over space and time.

(NS3) *Both temporally and spatially nonstationary Gaussian random field:* The spatio-temporal process is defined using an AR(1) model with time-dependent innovations

$$Z_t(\boldsymbol{s}) = 0.5Z_{t-1}(\boldsymbol{s}) + \left(1.3 + \sin\left(\frac{2\pi t}{400}\right)\right)\eta_t(\boldsymbol{s}).$$

For the simulations we use $\lambda = 20$.

7.3.2 Discussion

The empirical powers based on Models NS1 - NS3 are given in Table 2-5.

First we consider Model NS1, which is temporarily nonstationary, but stationary over space. The results of the general spatio-temporal test using the test set $\mathcal{P} = \{(1,0),(1,1),(0,1),(-1,1)\}\times\{1,2\}$ and orthogonal estimates set $\mathcal{P}' = \{(1,0),(1,1),(0,1),(-1,1)\}\times\{4,5\}$ (described in Section 4) are given in Table 2. Before discussing the results we note that over the test set $\mathcal{P}$ the Fourier transforms are near uncorrelated. However, the temporal nonstationarity means that the orthogonal estimators $\widehat{A}_{g,h}(\boldsymbol{r}_1, r_2)$ and $\widehat{B}_{g,h;H}(\omega_{jH};\boldsymbol{r}_1, r_2)$ for $(\boldsymbol{r}_1, r_2) \in \mathcal{P}'$ do not necessarily share the same variance. Furthermore, there is correlation between the terms. These conflicting behaves (decorrelation of DFTs but inability to capture the true variance) helps explain why the power in the overall test varies between 27%-80% in the case $\rho = 0.5$ and 21% - 80% in the case $\rho = 1$ (excluding the non-variance adjusted tests). The results of the one-way temporal stationary and one-way spatial stationary tests (described in Section 5) are given in Table 3. The power in the one-way temporal tests are close to 100% for all the test statistics (as we would expect since the process is temporally nonstationary) for both $\rho = 0.5$ and $\rho = 1$. The power for the one-way spatial tests drops considerably (as expected because $NS1$ is spatially stationary) for the average covariance test and variance adjusted average squared covariance test. In the case of the variance adjusted average squared tests the proportion of rejection is least in the case $\rho = 0.5$ and $H = 10$.

Next we consider Model NS2, which is temporarily stationary, but spatially nonstationary. The results are reported in Table 4. In this case the general spatio-temporal test using the test set $\mathcal{P} = \{(1,0),(1,1),(0,1),(-1,1)\} \times \{1,2\}$ and orthogonal estimates set $\mathcal{P}' = \{(1,0),(1,1),(0,1),(-1,1)\} \times \{4,5\}$ gives very little power. As we would expect, in the one-way test for temporal stationarity the proportion of rejections is close to the nominal level (with the exception of the variance adjusted test average squared test with $H = 20$ when the proportion of rejection is about 12%). However, the test does seem to have some power in the one-way test for spatial stationarity. In the case that $n = 500$ all the tests (excluding the non-variance adjusted tests) have power between 8-21%. This level of power is not high but it is higher than the case $n = 100$. The overall low power is because the number of observations is relatively sparse on the random field ($n = 500$ and $\lambda = 20$). Therefore most of the observations are unlikely to be highly correlated and thus contains very little information about the nonstationary structure (recall the variance of the spatio-temporal process is constant). It is likely if a larger n were used in the simulations, the power would increase (compare with the simulations in Bandyopadhyay and Subba Rao [2016]).

Lastly, we consider Model NS3, which is both temporal and spatial nonstationarity. The results are presented in Table 5. For the general spatio-temporal tests we get higher powers than for Model NS1 across all the tests. The power increases to 100% for the one-way temporal stationary test. For the one-way spatial stationarity tests the power is more than for the same tests using model NS2.

We mention that for all the models (NS1-NS3) the power for the average squared covariance test without variance adjustment is very high. However, we have to be cautious about interpreting the result of these tests as the simulations under the null of stationarity show that the these test statistics are unable to keep the nominal level when the process is not Gaussian.

Comparing the rejection rates with and without the nugget term removed (the values outside and insides the parentheses), we observe that for models NS1 and NS2 the rejection rates with and without the nugget term are about the same. However, for NS3 the power is slightly more after removing the nugget term.

Dedication

SSR was very fortunate to attend a course on nonparametric statistics given by Professor M. B. Priestley when she was an undergraduate student. His classes were a joy to attend.

During the 1960's, Professor M. B. Priestley was one the first researchers to study nonstationary time series, without his fundamental contributions this paper would not have been possible. Therefore, this paper is dedicated to the memory of Professor M. B. Priestley whose kind nature and encouragement was an inspiration to all.

		Model NS1: Overall Power			
		ρ			
		0.5		1	
	n	100	500	100	500
	$\mathbf{T}_{1,g,\widehat{V}^{-1/2}}$	0.73 (0.80)	0.60 (0.59)	0.76 (0.74)	0.57 (0.49)
	$\mathbf{M}_{1,g,\widehat{V}^{-1/2}}$	0.74 (0.78)	0.61 (0.59)	0.80 (0.75)	0.59 (0.51)
H=20	$\mathbf{T}_{2,g,\widehat{V}^{-1/2},1}$	0.99 (0.99)	0.99 (0.99)	0.99 (1.00)	0.97 (0.96)
	$\mathbf{M}_{2,g,\widehat{V}^{-1/2},1}$	0.97 (0.99)	0.99 (0.99)	0.98 (0.98)	0.93 (0.92)
	$\mathbf{T}_{2,g,\widehat{V}^{-1/2},\widehat{W}}$	0.44 (0.56)	0.33 (0.27)	0.45 (0.45)	0.22 (0.21)
	$\mathbf{M}_{2,g,\widehat{V}^{-1/2},\widehat{W}}$	0.51 (0.64)	0.47 (0.45)	0.60 (0.57)	0.34 (0.27)
H=10	$\mathbf{T}_{2,g,\widehat{V}^{-1/2},1}$	1.00 (1.00)	1.00 (0.99)	0.99 (1.00)	1.00 (1.00)
	$\mathbf{M}_{2,g,\widehat{V}^{-1/2},1}$	0.98 (0.99)	0.99 (0.98)	0.99 (0.98)	0.98 (0.97)
	$\mathbf{T}_{2,g,\widehat{V}^{-1/2},\widehat{W}}$	0.53 (0.56)	0.39 (0.37)	0.55 (0.50)	0.30 (0.27)
	$\mathbf{M}_{2,g,\widehat{V}^{-1/2},\widehat{W}}$	0.52 (0.56)	0.46 (0.44)	0.52 (0.48)	0.33 (0.28)

Table 2: Overall empirical power at 5% level based on different tests with $\lambda = 5$ for non-stationary data generated from the model NS1 with innovations coming from a Gaussian random field with exponential covariance functions. Rejection rate without removing N_T (see (18)) are in the parentheses.

		Model NS1							
		Temporal Power				Spatial Power			
		ρ				ρ			
		0.5		1		0.5		1	
	n	100	500	100	500	100	500	100	500
	$\mathbf{T}_{1,g,\widehat{V}^{-1/2}}$	0.97 (0.99)	1.00 (1.00)	0.97 (0.99)	1.00 (1.00)	0.01 (0.01)	0.01 (0.01)	0.04 (0.03)	0.04 (0.04)
	$\mathbf{M}_{1,g,\widehat{V}^{-1/2}}$	0.99 (1.00)	1.00 (1.00)	0.99 (0.99)	1.00 (1.00)	0.01 (0.01)	0.02 (0.01)	0.02 (0.02)	0.03 (0.01)
H=20	$\mathbf{T}_{2,g,\widehat{V}^{-1/2},1}$	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	0.99 (0.99)	0.99 (0.99)	1.00 (1.00)	1.00 (1.00)
	$\mathbf{M}_{2,g,\widehat{V}^{-1/2},1}$	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	0.99 (0.98)	0.99 (0.99)	1.00 (1.00)	1.00 (1.00)
	$\mathbf{T}_{2,g,\widehat{V}^{-1/2},\widehat{W}}$	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	0.05 (0.04)	0.10 (0.08)	0.15 (0.14)	0.28 (0.31)
	$\mathbf{M}_{2,g,\widehat{V}^{-1/2},\widehat{W}}$	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	0.15 (0.15)	0.18 (0.16)	0.23 (0.24)	0.31 (0.32)
H=10	$\mathbf{T}_{2,g,\widehat{V}^{-1/2},1}$	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	0.99 (0.99)	0.99 (0.99)	1.00 (1.00)	1.00 (1.00)
	$\mathbf{M}_{2,g,\widehat{V}^{-1/2},1}$	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	0.99 (0.99)	0.99 (0.99)	1.00 (1.00)	1.00 (1.00)
	$\mathbf{T}_{2,g,\widehat{V}^{-1/2},\widehat{W}}$	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	0.02 (0.02)	0.06 (0.05)	0.10 (0.11)	0.26 (0.28)
	$\mathbf{M}_{2,g,\widehat{V}^{-1/2},\widehat{W}}$	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	0.06 (0.06)	0.09 (0.07)	0.14 (0.14)	0.24 (0.26)

Table 3: One-way empirical powers at 5% level based on different tests with $\lambda = 5$ for nonstationary data generated from the model NS1 with innovations coming from a Gaussian random field with exponential covariance functions. Rejection rate without removing N_T (see (18)) are in the parentheses.

Acknowledgments

SB's and SSR's work has been partially supported by the National Science Foundation, DMS-1406622 and DMS-1513647 respectively. CJ has been supported by the Research Center 884 "Political Economy of Reforms" (Project B6) funded by the German Science Foundation (DFG). The authors also wish to thank Tilmann Gneiting and Tata Subba Rao who encouraged us to consider this problem.

		Model NS2					
		Overall Power		Temporal Power		Spatial Power	
	n	100	500	100	500	100	500
	$\mathbf{T}_{1,g,\widehat{V}^{-1/2}}$	0.09 (0.11)	0.11 (0.11)	0.04 (0.05)	0.05 (0.06)	0.06 (0.07)	0.17 (0.17)
	$\mathbf{M}_{1,g,\widehat{V}^{-1/2}}$	0.08 (0.08)	0.10 (0.09)	0.03 (0.03)	0.05 (0.06)	0.03 (0.02)	0.15 (0.15)
H=20	$\mathbf{T}_{2,g,\widehat{V}^{-1/2},1}$	0.04 (0.06)	0.06 (0.06)	0.04 (0.05)	0.05 (0.06)	0.24 (0.26)	0.28 (0.31)
	$\mathbf{M}_{2,g,\widehat{V}^{-1/2},1}$	0.07 (0.09)	0.07 (0.07)	0.08 (0.08)	0.03 (0.04)	0.12 (0.15)	0.21 (0.25)
	$\mathbf{T}_{2,g,\widehat{V}^{-1/2},\widehat{W}}$	0.07 (0.06)	0.05 (0.04)	0.11 (0.12)	0.12 (0.12)	0.05 (0.06)	0.18 (0.19)
	$\mathbf{M}_{2,g,\widehat{V}^{-1/2},\widehat{W}}$	0.08 (0.06)	0.08 (0.09)	0.11 (0.12)	0.13 (0.12)	0.07 (0.10)	0.21 (0.20)
H=10	$\mathbf{T}_{2,g,\widehat{V}^{-1/2},1}$	0.15 (0.18)	0.15 (0.15)	0.06 (0.07)	0.06 (0.07)	0.38 (0.47)	0.56 (0.59)
	$\mathbf{M}_{2,g,\widehat{V}^{-1/2},1}$	0.12 (0.10)	0.13 (0.15)	0.05 (0.06)	0.06 (0.06)	0.32 (0.34)	0.48 (0.50)
	$\mathbf{T}_{2,g,\widehat{V}^{-1/2},\widehat{W}}$	0.08 (0.10)	0.06 (0.05)	0.05 (0.04)	0.04 (0.04)	0.01 (0.02)	0.09 (0.10)
	$\mathbf{M}_{2,g,\widehat{V}^{-1/2},\widehat{W}}$	0.13 (0.14)	0.10 (0.10)	0.06 (0.03)	0.04 (0.05)	0.01 (0.01)	0.08 (0.09)

Table 4: Empirical powers at 5% level based on different tests with $\lambda = 20$ for nonstationary data generated from the model NS2. Rejection rate without removing N_T (see (18)) are in the parentheses.

		Model NS3					
		Overall Power		Temporal Power		Spatial Power	
	n	100	500	100	500	100	500
	$\mathbf{T}_{1,g,\widehat{V}^{-1/2}}$	0.83 (0.80)	0.98 (0.92)	0.92 (0.99)	1.00 (1.00)	0.11 (0.07)	0.33 (0.19)
	$\mathbf{M}_{1,g,\widehat{V}^{-1/2}}$	0.92 (0.88)	0.99 (0.97)	0.95 (1.00)	1.00 (1.00)	0.18 (0.08)	0.54 (0.25)
H=20	$\mathbf{T}_{2,g,\widehat{V}^{-1/2},1}$	0.99 (0.99)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	0.99 (1.00)	1.00 (1.00)
	$\mathbf{M}_{2,g,\widehat{V}^{-1/2},1}$	0.99 (0.98)	1.00 (0.99)	1.00 (1.00)	1.00 (1.00)	0.99 (1.00)	1.00 (1.00)
	$\mathbf{T}_{2,g,\widehat{V}^{-1/2},\widehat{W}}$	0.65 (0.66)	0.85 (0.77)	1.00 (1.00)	1.00 (1.00)	0.34 (0.22)	0.74 (0.50)
	$\mathbf{M}_{2,g,\widehat{V}^{-1/2},\widehat{W}}$	0.82 (0.80)	0.98 (0.87)	1.00 (1.00)	1.00 (1.00)	0.52 (0.34)	0.90 (0.69)
H=10	$\mathbf{T}_{2,g,\widehat{V}^{-1/2},1}$	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)
	$\mathbf{M}_{2,g,\widehat{V}^{-1/2},1}$	0.99 (0.99)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)	1.00 (1.00)
	$\mathbf{T}_{2,g,\widehat{V}^{-1/2},\widehat{W}}$	0.65 (0.63)	0.81 (0.77)	1.00 (1.00)	1.00 (1.00)	0.13 (0.08)	0.45 (0.35)
	$\mathbf{M}_{2,g,\widehat{V}^{-1/2},\widehat{W}}$	0.79 (0.70)	0.94 (0.84)	1.00 (1.00)	1.00 (1.00)	0.32 (0.19)	0.79 (0.58)

Table 5: Empirical powers at 5% level based on different tests with $\lambda = 20$ for nonstationary data generated from the model NS3. Rejection rate without removing N_T (see (18)) are in the parentheses.

References

S. Bandyopadhyay and S. N. Lahiri. Asymptotic properties of Discrete Fourier Transforms for spatial data. *Sankhya, Series A*, 71:221–259, 2009.

S. Bandyopadhyay and S. Subba Rao. A test for spatial stationarity with irregular spaced locations. *To appear in Journal of the Royal Statistical Society (B)*, 2016.

S. Bandyopadhyay, S. N. Lahiri, and D. Nordman. A frequency domain empirical likelihood method for irregular spaced spatial data. *Annals of Statistics*, 43(2):519–545, 2015.

H. Cho. A randomized test for stationarity. *Preprint*, 2014.

R. Dahlhaus. Fitting time series models to nonstationary processes. *Annals of Statistics*, 16: 1–37, 1997.

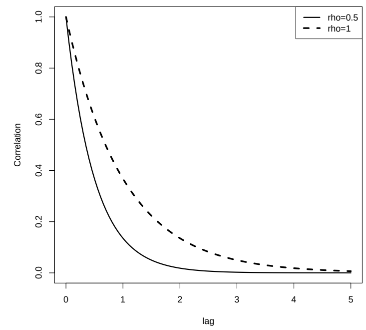

Figure 1: Plot of the Exponential correlation function with the range parameters $\rho = 0.5$ and $\rho = 1$.

R. Dahlhaus. Locally stationary processes. *Handbook of Statistics, Time Series Analysis: Methods and Applications*, 30:351–413, 2012.

H. Dette, P. Preuss, and M. Vetter. A measure of stationarity in locally stationary processes with applications to testing. *Journal of the Americal Statistical Association*, 106:1113–1124, 2011.

Y. Dwivedi and S. Subba Rao. A test for second order stationarity based on the Discrete Fourier Transform. *Journal of Time Series Analysis*, 32:68–91, 2011.

M. Eichler. Testing nonparametric and semiparametric hypotheses in vector stationary processes. *Journal of Multivariate Analysis*, 99(5):968–1009, 2008.

A. Epharty, J. Tabrikian, and H. Messer. Underwater source detection using a spatial stationary test. *Journal of the Acoustical Society of America*, 109:1053–1063, 2001.

M. Fuentes. Testing for separability of a spatial temporal covariance functions. *Journal of Statistical Planning and Inference*, 136:447–466, 2006.

E. G. Gladyshev. Periodically and almost periodically correlated random process with a continuous parameter. *Theory of Probability and Applications*, 8:184–189, 1963.

T. Gneiting, M. G. Genton, and P. Guttorp. Geostatistical space-time models, stationarity, separability, and full symmetry. *Monographs On Statistics and Applied Probability*, 107: 151, 2006.

N. R. Goodman. Statistical tests for stationarity within the framework of harmonizable processes. Technical report, Office of Navel Research: Research Report No. 65-28, 1965.

C. Gorrostieta, H. Ombao, and von Sachs. R. Time-dependent dual frequency coherency in multivariate non-stationary time series. *In Submission*, 2016.

P. Hall and P. Patil. Properties of nonparametric estimators of autocovariance for stationary random fields. *Probability Theory and Related Fields*, 99:399–424, 1994.

H. Hindberg and S. C. Olhede. Estimation of ambiguity functions with limited spread. *IEEE Transactions on Signal Processing*, 58:2383–2388, 2010.

T. Hsing and W.B. Wu. On weighted U-statistics for stationary processes. *Annals of Probability*, pages 1600–1631, 2004.

C. Jentsch. A new frequency domain approach of testing for covariance stationarity and for periodic stationarity in multivariate linear processes. *Journal of Time Series Analysis*, 33: 177–192, 2012.

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

M. Jun and M. Genton. A test for stationarity of spatio-temporal random fields on planar and spherical domains. *Statistica Sinica*, 22:1737–1764, 2012.

S. N. Lahiri. Central limit theorems for weighted sums of a spatial process under a class of stochastic and fixed designs. *Sankhya (A)*, 65:356–388, 2003.

J. Lee and S. Subba Rao. A note on general quadratic forms of nonstationary processes. *Preprint*, 2015.

J. Lei, H. Wang, and S. Wang. A new nonparametric test for stationarity in the time domain. *Journal of the Royal Statistical Society (B)*, 2015.

A. Leucht. Degenerate U- and V-statistics under weak dependence: Asymptotic theory and bootstrap consistency. *Bernoulli*, 18:552–585, 2012.

K. S. Lii and M. Rosenblatt. Spectral analysis for harmonizable processes. *Annals of Statistics*, 30:258–297, 2002.

Y. Matsuda and Y. Yajima. Fourier analysis of irregularly spaced data on $\mathbb{R}^d$. *Journal of the Royal Statistical Society (B)*, 71:191–217, 2009.

G. Nason. 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, 2013.

C. J. Paciorek and M. J. Schervish. Spatial modelling using a new class of nonstationary covariance functions. *Environmetrics*, 17(5):483–506, 2006.

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

E. Paparoditis. Validitating stationarity assumptions in time series by rolling local periodograms. *Journal of the American Statistical Association*, 105:839–851, 2010.

D.N. Politis and J. P. Romano. The stationary bootstrap. *Journal of the American Statistical Association*, 89:1303–1313, 1994.

M. B. Priestley. Evolutionary spectral and nonstationary processes. *Journal of the Royal Statistical Society (B)*, 27:385–392, 1965.

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

R. Puchstein and P. Preuss. Testing for stationarity in multivariate locally stationary processes. *Journal of Time Series Analysis*, 37:3–29, 2016.

M. Sherman. *Spatial Statistics and Spatio-Temporal Data: Covariance Functions and Directional Properties*. Wiley, New York, 2010.

S. Subba Rao. Fourier based statistics for irregular spaced spatial data. *Submitted*, 2015a.

S. Subba Rao. Orthogonal samples: A simple method for variance estimation in time series. *Preprint*, 2015b.

T. Subba Rao and M. M. Gabr. *An Introduction to Bispectral Analysis and Bilinear Time Series Models*. Lecture Notes in Statistics (24). Springer, New York, 1984.

T. Subba Rao and G. Terdik. On the frequency variogram and on frequency domain methods for the analysis of spatio-temporal data. *Submitted*, 2016.

R. von Sachs and M. H. Neumann. A wavelet-based test for stationarity. *Journal of Time Series Analysis*, 21:597–613, 1999.

A. M. Yaglom. *Correlation Theory of Stationary and Related Random functions*. Springer, New York, 1987.

A Supplementary material

A.1 Proof of Lemma 3.1

To prove the result we start by expanding $\text{cov}\left[J(\boldsymbol{\Omega}_{\boldsymbol{k}_1}, \omega_{k_2}), J(\boldsymbol{\Omega}_{\boldsymbol{k}_1+\boldsymbol{r}_1}, \omega_{k_2+r_2})\right]$.

$$\begin{aligned}
&\text{cov}\left[J(\boldsymbol{\Omega}_{\boldsymbol{k}_1}, \omega_{k_2}), J(\boldsymbol{\Omega}_{\boldsymbol{k}_1+\boldsymbol{r}_1}, \omega_{k_2+r_2})\right] \\
&= \frac{1}{2\pi} \sum_{h=-(T-1)}^{T-1} e^{-ih\omega_{k_2}} \frac{1}{T} \sum_{t=1-\min(0,h)}^{T-\max(0,h)} \text{cov}[J_t(\boldsymbol{\Omega}_{k_1}), J_{t+h}(\boldsymbol{\Omega}_{\boldsymbol{k}_1+\boldsymbol{r}_1})] e^{-it\omega_{r_2}} \\
&= M + R, \qquad (48)
\end{aligned}$$

where M is the main term

$$M = \frac{1}{2\pi} \sum_{h=-(T-1)}^{T-1} e^{-ih\omega_{k_2}} \frac{1}{T} \sum_{t=1}^{T} \text{cov}[J_t(\boldsymbol{\Omega}_{k_1}), J_{t+h}(\boldsymbol{\Omega}_{\boldsymbol{k}_1+\boldsymbol{r}_1})] e^{-it\omega_{r_2}},$$

and R is the remainder

$$\begin{aligned}
R &= \frac{1}{2\pi} \sum_{h=0}^{T-1} e^{-ih\omega_{k_2}} \frac{1}{T} \sum_{t=T-h+1}^{T} \text{cov}[J_t(\boldsymbol{\Omega}_{k_1}), J_{t+h}(\boldsymbol{\Omega}_{\boldsymbol{k}_1+\boldsymbol{r}_1})] e^{-it\omega_{r_2}} \\
&\quad + \frac{1}{2\pi} \sum_{h=-(T-1)}^{-1} e^{-ih\omega_{k_2}} \frac{1}{T} \sum_{t=1}^{|h|} \text{cov}[J_t(\boldsymbol{\Omega}_{k_1}), J_{t+h}(\boldsymbol{\Omega}_{\boldsymbol{k}_1+\boldsymbol{r}_1})] e^{-it\omega_{r_2}}.
\end{aligned}$$

The expansions above are valid in the general case. Below we obtain expressions for M (the main term) and bounds for R in the case that the spatio-temporal process is stationary and nonstationary.

- **Spatially stationary**

  By using the same proof used to prove Theorem 2.1(i), Bandyopadhyay and Subba Rao [2016], and the rescaling devise over time, under spatial stationary we have, for $\boldsymbol{r}_1 \neq \boldsymbol{0}$,

$$\begin{aligned}
&\text{cov}[J_t(\boldsymbol{\Omega}_{\boldsymbol{k}_1}), J_{t+h}(\boldsymbol{\Omega}_{\boldsymbol{k}_1+\boldsymbol{r}_1})] \\
&= \underbrace{\int_{[-\lambda/2,\lambda/2]^d} e^{-i\boldsymbol{\Omega}_{\boldsymbol{r}}'\boldsymbol{s}_2} \int_{[-\lambda/2-\boldsymbol{s}_1,-\lambda/2]^d} \kappa_{h;\frac{t}{T}}(\boldsymbol{s}_1) e^{i\boldsymbol{\Omega}_{\boldsymbol{k}_1}'\boldsymbol{s}_1} d\boldsymbol{s}_1 d\boldsymbol{s}_2}_{O(\frac{\rho_h}{\lambda^{d-b}})} \\
&+ \underbrace{\int_{[-\lambda/2,\lambda/2]^d} e^{-i\boldsymbol{\Omega}_{\boldsymbol{r}}'\boldsymbol{s}_2} \int_{[\lambda/2,\lambda/2+\boldsymbol{s}_1]^d} \kappa_{h;\frac{t}{T}}(\boldsymbol{s}_1) e^{i\boldsymbol{\Omega}_{\boldsymbol{k}_1}'\boldsymbol{s}_1} d\boldsymbol{s}_1 d\boldsymbol{s}_2}_{O(\frac{\rho_h}{\lambda^{d-b}})} + O\left(\frac{\rho_h}{T\lambda^{d-b}} I_{\text{Time=NS}}\right),
\end{aligned}$$

and for $\boldsymbol{r}_1 = \boldsymbol{0}$,

$$\begin{aligned}
&\mathrm{cov}[J_t(\boldsymbol{\Omega}_{\boldsymbol{k}_1}), J_{t+h}(\boldsymbol{\Omega}_{\boldsymbol{k}_1+\boldsymbol{r}_1})] \\
&= \frac{c_{t,t+h}}{\lambda^d}\int_{[-\lambda/2,\lambda/2]^d} \kappa_{h;\frac{t}{T}}(\boldsymbol{v})\exp(-i\boldsymbol{v}'\boldsymbol{\Omega}_{\boldsymbol{k}_1})d\boldsymbol{v} + \frac{\lambda^d n_{t,t+h}}{n_t n_{t+h}}\kappa_{h,\frac{t}{T}}(0) + O\left(\frac{\rho_h I_{\text{Time=NS}}}{T}\right),
\end{aligned}$$

where, $b = b(\boldsymbol{r}_1)$ is the number of zeros in $\boldsymbol{r}_1$, $c_{t,t+h} = (n_t n_{t+h} - n_{t,t+h})/n_t n_{t+h}$, $n_{t,t+h} = |\{\boldsymbol{s}_{t,j}\}_{j=1}^{n_t}\} \cap \{\boldsymbol{s}_{t+h,j}\}_{j=1}^{n_{t+h}}|$ and $I_{\text{Time=NS}}$ denotes the indicator variable for temporal nonstationarity. Note that we use the notation $[-\lambda/2 - \boldsymbol{s}_1, -\lambda/2]^d = [-\lambda/2 - s_{11}] \times \ldots \times [-\lambda/2 - s_{1d}]$. Substituting the above into the remainder R we see that $|R| = O([T^{-1} + \lambda^d/n]I(\boldsymbol{r}_1 = 0) + \frac{1}{\lambda^{d-b}T}I(\boldsymbol{r}_1 \neq 0))$. Now we derive expression for M for the temporally stationary and nonstationary separately.

(a) **Temporally stationary** (i.e., $\kappa_{h,\frac{t}{T}}(\boldsymbol{v}) = \kappa_h(\boldsymbol{v})$) First we look at the case $\boldsymbol{r}_1 \neq \boldsymbol{0}$. In the case that $\boldsymbol{r}_1 \neq \boldsymbol{0}$ and $r_2 \neq 0$, we take the summand $\sum_{t=1}^{T} e^{-it\omega_{r_2}}$ in M separate of κ_h giving $M = 0$. Therefore, $\mathrm{cov}\left[J(\boldsymbol{\Omega}_{\boldsymbol{k}_1}, \omega_{k_2}), J(\boldsymbol{\Omega}_{\boldsymbol{k}_1+\boldsymbol{r}_1}, \omega_{k_2+r_2})\right] = O(\lambda^{-(d-b)}T^{-1})$. In the case that $\boldsymbol{r}_1 \neq \boldsymbol{0}$ but $r_2 = 0$, we get $M = O(\lambda^{-(b-d)})$, and thus $\mathrm{cov}\left[J(\boldsymbol{\Omega}_{\boldsymbol{k}_1}, \omega_{k_2}), J(\boldsymbol{\Omega}_{\boldsymbol{k}_1+\boldsymbol{r}_1}, \omega_{k_2+r_2})\right] = O(\lambda^{-(d-b)})$.

Now we consider the case $\boldsymbol{r}_1 = \boldsymbol{0}$. In the case that $\boldsymbol{r}_1 = \boldsymbol{0}$ but $r_2 \neq 0$, we use Assumption 3.1(ii), where $c_1 n \leq n_t \leq c_2 n$, which implies that $|c_{t,t+h} - 1| \leq \frac{c_2}{c_1 n}$ and immediately gives $M = O(T^{-1} + \lambda^d/n)$ and $\mathrm{cov}\left[J(\boldsymbol{\Omega}_{\boldsymbol{k}_1}, \omega_{k_2}), J(\boldsymbol{\Omega}_{\boldsymbol{k}_1}, \omega_{k_2+r_2})\right] = O(\lambda^d/n + T^{-1})$. On the other hand, when $\boldsymbol{r}_1 = \boldsymbol{0}$ *and* $r_2 = 0$ we have $M = f(\boldsymbol{\Omega}_{\boldsymbol{k}_1}, \omega_{k_2}) + O(T^{-1} + \lambda^{-1} + \lambda^d/n)$, which immediately leads us to $\mathrm{cov}\left[J(\boldsymbol{\Omega}_{\boldsymbol{k}_1}, \omega_{k_2}), J(\boldsymbol{\Omega}_{\boldsymbol{k}_1}, \omega_{k_2})\right] = f(\boldsymbol{\Omega}_{\boldsymbol{k}_1}, \omega_{k_2}) + O(T^{-1} + \lambda^{-1} + \lambda^d/n)$.

(b) **Temporally nonstationary** Again it is immediately clear that when $\boldsymbol{r}_1 \neq \boldsymbol{0}$ ($r_2 \in \mathbb{Z}$) we have $M = O\left(\lambda^{-(d-b)}\right)$, which gives $\mathrm{cov}\left[J(\boldsymbol{\Omega}_{\boldsymbol{k}_1}, \omega_{k_2}), J(\boldsymbol{\Omega}_{\boldsymbol{k}_1+\boldsymbol{r}_1}, \omega_{k_2+r_2})\right] = O(\lambda^{-(d-b)} + T^{-1})$. However, when $\boldsymbol{r}_1 = \boldsymbol{0}$ ($r_2 \in \mathbb{Z}$) (and using Assumption 3.1(ii)) it is clear that

$$M = \frac{1}{2\pi}\sum_{h=-(T-1)}^{T-1} e^{-ih\omega_{k_2}}\frac{1}{T}\sum_{t=1}^{T}\frac{1}{\lambda^d}\int_{[-\lambda/2,\lambda/2]^d}\kappa_{h;\frac{t}{T}}(\boldsymbol{v})\exp(-i\boldsymbol{v}'\boldsymbol{\Omega}_{\boldsymbol{k}_1})d\boldsymbol{v} + O\left(\frac{\lambda^d}{n} + \frac{1}{T}\right),$$

which gives the desired result.

- **Spatially nonstationary** If the spatio-temporal process is spatially nonstationary, using the same proof to prove Theorem 2.1(ii), Bandyopadhyay and Subba Rao [2016] and the rescaling devise over time and space we have,

$$
\begin{aligned}
&\mathrm{cov}[J_t(\boldsymbol{\Omega}_{\boldsymbol{k}_1}), J_{t+h}(\boldsymbol{\Omega}_{\boldsymbol{k}_1+\boldsymbol{r}_1})] \\
&= \frac{c_{t,t+h}}{\lambda^d}\int_{[-\lambda/2,\lambda/2]^{2d}} \kappa_{h;\frac{t}{T}}\left(\boldsymbol{v};\frac{\boldsymbol{s}}{\lambda}\right)\exp(-i\boldsymbol{v}'\boldsymbol{\Omega}_{\boldsymbol{k}_1})\exp(-i\boldsymbol{s}'\boldsymbol{\Omega}_{\boldsymbol{r}_1})d\boldsymbol{v}d\boldsymbol{s} \\
&+ \underbrace{\int_{[-\lambda/2,\lambda/2]^d} e^{-i\boldsymbol{\Omega}_{\boldsymbol{r}}'\boldsymbol{s}}\int_{[-\lambda/2-\boldsymbol{s},-\lambda/2]^d}\kappa_{h;\frac{t}{T}}\left(\boldsymbol{v};\frac{\boldsymbol{s}}{\lambda}\right)e^{i\boldsymbol{\Omega}_{\boldsymbol{k}_1}'\boldsymbol{v}}d\boldsymbol{v}d\boldsymbol{s}}_{O(\rho_h/\lambda)} \\
&+ \underbrace{\int_{[-\lambda/2,\lambda/2]^d} e^{-i\boldsymbol{\Omega}_{\boldsymbol{r}}'\boldsymbol{s}}\int_{[\lambda/2,\lambda/2+\boldsymbol{s}]^d}\kappa_{h;\frac{t}{T}}\left(\boldsymbol{v};\frac{\boldsymbol{s}}{\lambda}\right)e^{i\boldsymbol{\Omega}_{\boldsymbol{k}_1}'\boldsymbol{v}}d\boldsymbol{v}d\boldsymbol{s}}_{O(\rho_h/\lambda)} \\
&+\frac{n_{t,t+h}}{n_t n_{t+h}}\int_{[-\lambda/2,\lambda/2]^d}\kappa_{h;\frac{t}{T}}\left(0;\frac{\boldsymbol{s}}{\lambda}\right)\exp(-i\boldsymbol{s}'\boldsymbol{\Omega}_{\boldsymbol{r}_1})d\boldsymbol{s} + O\left(\frac{\rho_h I_{\mathrm{Time=NS}}}{T}\right).
\end{aligned}
$$

Using the above result it is straightforward to show that $R = O([1+\lambda^d/n]T^{-1})$.

(a) **Temporally stationary** (i.e., $\kappa_{h,\frac{t}{T}}(\boldsymbol{v},\boldsymbol{s}) = \kappa_h(\boldsymbol{v},\boldsymbol{s})$). Since the process is spatially nonstationary, we consider $\boldsymbol{r}_1 = \boldsymbol{0}$ and $\boldsymbol{r}_1 \neq 0$ together. In the case that $r_2 \neq 0$ $\sum_{t=1}^{T} e^{-itr_2}$ is separate of κ_h, thus $M = 0$ and $\mathrm{cov}\left[J(\boldsymbol{\Omega}_{\boldsymbol{k}_1},\omega_{k_2}), J(\boldsymbol{\Omega}_{\boldsymbol{k}_1},\omega_{k_2})\right] = O(T^{-1})$. If $r_2 = 0$ we have,

$$
M = \frac{1}{2\pi}\sum_{h=-(T-1)}^{T-1} e^{-ih\omega_{k_2}}\frac{1}{\lambda^d}\int_{[-\lambda/2,\lambda/2]^d}\kappa_h\left(\boldsymbol{v};\frac{\boldsymbol{s}}{\lambda}\right)\exp(-i\boldsymbol{v}'\boldsymbol{\Omega}_{\boldsymbol{k}_1})\exp(-i\boldsymbol{s}'\boldsymbol{\Omega}_{\boldsymbol{r}_1})d\boldsymbol{v}d\boldsymbol{s}+O\left(\frac{1}{\lambda}\right),
$$

which immediately leads to the desired result.

(b) **Temporally nonstationary** In this case using Assumption 3.1(ii) we have,

$$
\begin{aligned}
M &= \frac{1}{2\pi}\sum_{h=-(T-1)}^{T-1} e^{-ih\omega_{k_2}}\frac{1}{T}\sum_{t=1}^{T}e^{-it\omega_{r_2}} \\
&\times\frac{c_{t,t+h}}{\lambda^d}\int_{[-\lambda/2,\lambda/2]^{2d}}\kappa_{h;\frac{t}{T}}\left(\boldsymbol{v};\frac{\boldsymbol{s}}{\lambda}\right)\exp(-i\boldsymbol{v}'\boldsymbol{\Omega}_{\boldsymbol{k}_1})\exp(-i\boldsymbol{s}'\boldsymbol{\Omega}_{\boldsymbol{r}_1})d\boldsymbol{v}d\boldsymbol{s} + O\left(\frac{\lambda^d}{n}+\frac{1}{\lambda}\right),
\end{aligned}
$$

thus leading to the desired result.

A.2 Proof of results for stationary spatio-temporal processes

PROOF of Lemma 4.1 The proof of this lemma is identical to the proof of Lemma 3.1 in Bandyopadhyay and Subba Rao [2016] and hence omitted. □

To prove the remainder of the results in Section 4 we use the following notation

$$f_h(\boldsymbol{\Omega}) = \int_{\mathbb{R}^d} \kappa_h(\boldsymbol{s}) \exp(-i\boldsymbol{\Omega}'\boldsymbol{s}) d\boldsymbol{s},$$

$$f(\boldsymbol{\Omega}, \omega) = \sum_{h\in\mathbb{Z}} \exp(-ih\omega) \int_{\mathbb{R}^d} \kappa_h(\boldsymbol{s}) \exp(-i\boldsymbol{\Omega}'\boldsymbol{s}) d\boldsymbol{s},$$

and

$$f_{h_1,h_2,h_3}(\boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2, \boldsymbol{\Omega}_3) = \int_{\mathbb{R}^{3d}} \kappa_{h_1,h_2,h_3}(\boldsymbol{s}_1, \boldsymbol{s}_2, \boldsymbol{s}_3) \exp(-i(\boldsymbol{s}_1'\boldsymbol{\Omega}_1 + \boldsymbol{s}_2'\boldsymbol{\Omega}_2 + \boldsymbol{s}_3'\boldsymbol{\Omega}_3)) d\boldsymbol{s}_1 d\boldsymbol{s}_2 d\boldsymbol{s}_3.$$

Note that in this section we do not prove any central limit theorems. However, we conjecture that by combining Bandyopadhyay et al. [2015], which give a CLT for mixing spatial processes and the CLT for quadratic forms of a time series (see, for example, Hsing and Wu [2004], Leucht [2012], Lee and Subba Rao [2015]) asymptotic normality of spatio-temporal quadratic forms can be proved.

Having established an expression for the mean of $\widehat{a}_g(\cdot)$ under stationarity, the main focus is obtaining expressions for the variance and covariance of $\widehat{a}_g(\cdot)$ and the corresponding test statistics. To do this we define the related quantity $\widetilde{a}_g(\cdot)$ such that

$$\widetilde{a}_g(\omega_{k_2}; \boldsymbol{r}_1, r_2) = \widehat{a}_g(\omega_{k_2}; \boldsymbol{r}_1, r_2) + N_T,$$

where,

$$N_T = \frac{1}{2\pi T} \sum_{t,\tau=1}^{T} e^{it\omega_{k_2} - i\tau\omega_{k_2+r_2}} \frac{1}{\lambda^d} \sum_{\boldsymbol{k}_1=-\boldsymbol{a}}^{\boldsymbol{a}} g(\boldsymbol{\Omega}_{\boldsymbol{k}_1}) \frac{1}{n_t n_\tau} \sum_{j=1}^{n} \delta_{t,j}\delta_{\tau,j} Z_t(\boldsymbol{s}_j) Z_\tau(\boldsymbol{s}_j) e^{-i\boldsymbol{s}_j\boldsymbol{\Omega}_{\boldsymbol{r}_1}}.$$

More precisely, we have,

$$\begin{aligned}
&\widetilde{a}_g(\omega_{k_2}; \boldsymbol{r}_1, r_2) \\
= & \frac{1}{\lambda^d} \sum_{\boldsymbol{k}_1=-\boldsymbol{a}}^{\boldsymbol{a}} g(\boldsymbol{\Omega}_{\boldsymbol{k}_1}) J(\boldsymbol{\Omega}_{k_1}, \omega_{k_2}) \overline{J(\boldsymbol{\Omega}_{\boldsymbol{k}_1+\boldsymbol{r}_1}, \omega_{k_2+r_2})} \\
= & \frac{1}{2\pi T} \sum_{t,\tau=1}^{T} e^{it\omega_{k_2} - i\tau\omega_{k_2+r_2}} \frac{1}{\lambda^d} \sum_{\boldsymbol{k}_1=-\boldsymbol{a}}^{\boldsymbol{a}} g(\boldsymbol{\Omega}_{\boldsymbol{k}_1}) J_t(\boldsymbol{\Omega}_{k_1}) \overline{J_\tau(\boldsymbol{\Omega}_{\boldsymbol{k}_1+\boldsymbol{r}_1})} \\
= & \frac{1}{2\pi T} \sum_{t,\tau=1}^{T} e^{it\omega_{k_2} - i\tau\omega_{k_2+r_2}} \frac{1}{\lambda^d} \sum_{\boldsymbol{k}_1=-\boldsymbol{a}}^{\boldsymbol{a}} g(\boldsymbol{\Omega}_{\boldsymbol{k}_1}) \frac{1}{n_t n_\tau} \sum_{j_1,j_2=1}^{n} \delta_{t,j_1}\delta_{\tau,j_2} Z_t(\boldsymbol{s}_{j_1}) Z_\tau(\boldsymbol{s}_{j_2}) \\
& \times e^{i\boldsymbol{s}_{j_1}\boldsymbol{\Omega}_{\boldsymbol{k}_1} - i\boldsymbol{s}_{j_2}\boldsymbol{\Omega}_{\boldsymbol{k}_1+\boldsymbol{r}_1}},
\end{aligned}$$

where the second equation follows by expanding $J(\boldsymbol{\Omega}_{k_1}, \omega_{k_2})$. To understand the role N_T plays, consider the expectation of N_T for the case $\boldsymbol{r}_1 = \boldsymbol{0}$ and $r_2 = 0$; not a case included in the text, but useful in understanding its role. Taking expectation of N_T (under stationarity) we have

$$\begin{aligned} \mathrm{E}[N_T] &= \frac{1}{2\pi T} \sum_{\boldsymbol{k}=-\boldsymbol{a}}^{\boldsymbol{a}} g(\boldsymbol{\Omega}_{\boldsymbol{k}_1}) \sum_{t,\tau=1}^{T} \exp(-i(\tau - t)\omega_{k_2}) \frac{n_{t,\tau}}{n_t n_\tau} \\ &\approx \frac{\lambda^d}{n} \int_{2\pi[-a/\lambda, a/\lambda]^d} g(\boldsymbol{\Omega}) d\boldsymbol{\Omega} \times \frac{1}{2\pi} \sum_{h\in\mathbb{Z}} \exp(-ih\omega_{k_2}) \kappa_h(0) = O\left(\frac{a^d}{n}\right). \end{aligned}$$

In the case that we constrain the frequency grid $\{\boldsymbol{\Omega}_{\boldsymbol{k}}; \boldsymbol{k} = (k_1, \ldots, k_d), -a \leq k_j \leq a\}$ to be bounded, i.e., $a/\lambda \to c < \infty$ as $a, \lambda \to \infty$, then it is clear that $\mathrm{E}[N_T] = O(\lambda^d/n) = o(1)$. Furthermore, using similar arguments it can be shown that the variance of N_T is asymptotically negligible and $\lambda^d \mathrm{var}[\widehat{a}_g(\omega_{k_2}; \boldsymbol{r}_1, r_2)] = \lambda^d \mathrm{var}[\widetilde{a}_g(\omega_{k_2}; \boldsymbol{r}_1, r_2)] + o(1)$ when the frequency grid is bounded. On the other hand, if the frequency grid is not bounded and $a/\lambda \to \infty$ as $\lambda \to \infty$ then we can show that for $\boldsymbol{r}_1 = \boldsymbol{0}$ and $r_2 = 0$ we have $\mathrm{E}[N_T] = O(a^d/n)$ and for general $\boldsymbol{r}_1$ and r_2 $\lambda^d \mathrm{var}[N_T] = (a^{2d}/n^2)$. Therefore, if the frequency grid is not bounded, $\widehat{a}_g(\omega_{k_2}; \boldsymbol{r}_1, r_2)$ and $\widetilde{a}_g(\omega_{k_2}; \boldsymbol{r}_1, r_2)$ are not asymptotically equivalent. However, $\widetilde{a}_g(\omega_{k_2}; \boldsymbol{r}_1, r_2)$ does play an important role in understanding the covariance of $\widehat{a}_g(\omega_{k_2}; \boldsymbol{r}_1, r_2)$, and we come back to this later on.

Returning to $\widehat{a}_g(\omega_{k_2}; \boldsymbol{r}_1, r_2)$, we see from the definition of $\widehat{a}_g(\cdot)$ that in order to obtain the covariance of $\widehat{a}_g(\cdot)$ we require the expansion

$$\begin{aligned} &\lambda^d \mathrm{cov}\left[\frac{1}{\lambda^d} \sum_{\boldsymbol{k}_1=-a}^{a} g(\boldsymbol{\Omega}_{\boldsymbol{k}_1}) \frac{1}{n_{t_1} n_{t_2}} \sum_{j_1 \neq j_2} \delta_{t_1,j_1} \delta_{t_2,j_2} Z_{t_1}(\boldsymbol{s}_{j_1}) Z_{t_2}(\boldsymbol{s}_{j_2}) e^{i\boldsymbol{s}_{j_1}\boldsymbol{\Omega}_{\boldsymbol{k}_1} - i\boldsymbol{s}_{j_2}\boldsymbol{\Omega}_{\boldsymbol{k}_1+\boldsymbol{r}_1}}, \right. \\ &\quad \left. \frac{1}{\lambda^d} \sum_{\boldsymbol{k}_3=-a}^{a} g(\boldsymbol{\Omega}_{\boldsymbol{k}_3}) \frac{1}{n_{t_3} n_{t_4}} \sum_{j_3 \neq j_4} \delta_{t_3,j_3} \delta_{t_4,j_4} Z_{t_4}(\boldsymbol{s}_{j_3}) Z_{t_4}(\boldsymbol{s}_{j_4}) e^{i\boldsymbol{s}_{j_3}\boldsymbol{\Omega}_{\boldsymbol{k}_3} - i\boldsymbol{s}_{j_4}\boldsymbol{\Omega}_{\boldsymbol{k}_3+\boldsymbol{r}_3}} \right] \\ &\quad = \widehat{A} + \widehat{B} + \widehat{C} \qquad (49) \end{aligned}$$

where,

$$\begin{aligned} \widehat{A} &= \frac{1}{\lambda^d} \sum_{\boldsymbol{k}_1,\boldsymbol{k}_3=-a}^{a} g(\boldsymbol{\Omega}_{\boldsymbol{k}_1}) \overline{g(\boldsymbol{\Omega}_{\boldsymbol{k}_3})} \frac{1}{\prod_{j=1}^{4} n_{t_j}} \sum_{\substack{j_1 \neq j_2 \\ j_3 \neq j_4}} \mathrm{cov}\left[\delta_{t_1,j_1} Z_{t_1}(\boldsymbol{s}_{j_1}) e^{i\boldsymbol{s}_{j_1}\boldsymbol{\Omega}_{\boldsymbol{k}_1}}, \delta_{t_3,j_1} Z_{t_3}(\boldsymbol{s}_{j_1}) e^{i\boldsymbol{s}_{j_3}\boldsymbol{\Omega}_{\boldsymbol{k}_3}}\right] \\ &\quad \times \mathrm{cov}\left[\delta_{t_2,j_2} Z_{t_2}(\boldsymbol{s}_{j_2}) e^{-i\boldsymbol{s}_{j_2}\boldsymbol{\Omega}_{\boldsymbol{k}_1+\boldsymbol{r}_1}}, \delta_{t_4,j_4} Z_{t_4}(\boldsymbol{s}_{j_4}) e^{-i\boldsymbol{s}_{j_4}\boldsymbol{\Omega}_{\boldsymbol{k}_3+\boldsymbol{r}_3}}\right] \end{aligned}$$

$$\widehat{B} = \frac{1}{\lambda^d}\sum_{\boldsymbol{k}_1,\boldsymbol{k}_3=-a}^{a} g(\boldsymbol{\Omega}_{\boldsymbol{k}_1})\overline{g(\boldsymbol{\Omega}_{\boldsymbol{k}_3})}\frac{1}{\prod_{j=1}^4 n_j}\sum_{\substack{j_1\neq j_2\\ j_3\neq j_4}} \mathrm{cov}\left[\delta_{t_1,j_1}Z_{t_1}(\boldsymbol{s}_{j_1})e^{i\boldsymbol{s}_{j_1}\boldsymbol{\Omega}_{\boldsymbol{k}_1}}, \delta_{t_4,j_4}Z_{t_4}(\boldsymbol{s}_{j_4})e^{-i\boldsymbol{s}_{j_4}\boldsymbol{\Omega}_{\boldsymbol{k}_3+\boldsymbol{r}_3}}\right]$$
$$\times\mathrm{cov}\left[\delta_{t_2,j_2}Z_{t_2}(\boldsymbol{s}_{j_2})e^{-i\boldsymbol{s}_{j_2}\boldsymbol{\Omega}_{\boldsymbol{k}_1+\boldsymbol{r}_1}}, \delta_{t_3,j_1}Z_{t_3}(\boldsymbol{s}_{j_1})e^{i\boldsymbol{s}_{j_3}\boldsymbol{\Omega}_{\boldsymbol{k}_3}}\right]$$

$$\widehat{C} = \frac{1}{\lambda^d}\sum_{\boldsymbol{k}_1,\boldsymbol{k}_3=-a}^{a} g(\boldsymbol{\Omega}_{\boldsymbol{k}_1})\overline{g(\boldsymbol{\Omega}_{\boldsymbol{k}_3})}\frac{1}{\prod_{j=1}^4 n_j}\sum_{\substack{j_1\neq j_2\\ j_3\neq j_4}} \mathrm{cum}\Big[\delta_{t_1,j_1}Z_{t_1}(\boldsymbol{s}_{j_1})e^{i\boldsymbol{s}_{j_1}\boldsymbol{\Omega}_{\boldsymbol{k}_1}}, \delta_{t_2,j_2}Z_{t_2}(\boldsymbol{s}_{j_2})e^{-i\boldsymbol{s}_{j_2}\boldsymbol{\Omega}_{\boldsymbol{k}_1+\boldsymbol{r}_1}},$$
$$\delta_{t_3,j_1}Z_{t_3}(\boldsymbol{s}_{j_1})e^{-i\boldsymbol{s}_{j_3}\boldsymbol{\Omega}_{\boldsymbol{k}_3}}, \delta_{t_4,j_4}Z_{t_4}(\boldsymbol{s}_{j_4})e^{i\boldsymbol{s}_{j_4}\boldsymbol{\Omega}_{\boldsymbol{k}_3+\boldsymbol{r}_3}}\Big].$$

Simplifications for these terms can be obtained by using the methods developed in Subba Rao [2015a]. Using this we can show

$$\widehat{A} = \frac{I_{\boldsymbol{r}_1=\boldsymbol{r}_3}}{(2\pi)^d}\int_{\mathcal{D}} |g(\boldsymbol{\Omega})|^2 f_{t_3-t_1}(\boldsymbol{\Omega})\overline{f_{t_4-t_2}(\boldsymbol{\Omega}+\boldsymbol{\Omega}_{\boldsymbol{r}_1})}d\boldsymbol{\Omega} + R_{1,t_3-t_1,t_4-t_2},$$

$$\widehat{B} = \frac{I_{\boldsymbol{r}_1=\boldsymbol{r}_3}}{(2\pi)^d}\int_{\mathcal{D}_{\boldsymbol{r}_1}} g(\boldsymbol{\Omega})\overline{g(-\boldsymbol{\Omega}-\boldsymbol{\Omega}_{\boldsymbol{r}_1})}f_{t_4-t_1}(\boldsymbol{\Omega})\overline{f_{t_3-t_2}(\boldsymbol{\Omega}+\boldsymbol{\Omega}_{\boldsymbol{r}})}d\boldsymbol{\Omega} + R_{2,t_4-t_1,t_3-t_2}$$

and,

$$\widehat{C} = \frac{I_{\boldsymbol{r}_1=\boldsymbol{r}_3}}{(2\pi)^{2d}}\int_{\mathcal{D}^2} g(\boldsymbol{\Omega}_1)\overline{g(\boldsymbol{\Omega}_2)}f_{t_2-t_1,t_3-t_1,t_4-t_1}(\boldsymbol{\Omega}_1+\boldsymbol{\Omega}_{\boldsymbol{r}_1},\boldsymbol{\Omega}_2,-\boldsymbol{\Omega}_2-\boldsymbol{\Omega}_{\boldsymbol{r}_1})d\boldsymbol{\Omega}_1 d\boldsymbol{\Omega}_2 + R_{3,t_2-t_1,t_3-t_1,t_4-t_1},$$

where,

$$\begin{aligned}
|R_{1,t_3-t_1,t_4-t_2}| &= O(\rho_{t_3-t_1}\rho_{t_4-t_2}\ell_{\lambda,a,n}),\\
|R_{2,t_4-t_1,t_3-t_2}| &= O(\rho_{t_4-t_1}\rho_{t_3-t_2}\ell_{\lambda,a,n}), \text{ and}\\
R_{3,t_2-t_1,t_3-t_1,t_4-t_1} &= O\left(\rho_{t_2-t_1}\rho_{t_3-t_1}\rho_{t_4-t_1}\left[\ell_{\lambda,a,n}+\frac{(a\lambda)^d}{n^2}\right]\right).
\end{aligned}$$

We further observe that use of the expansions given in (49) to obtain an expression for $\mathrm{var}[\widehat{a}_g(\omega_k;\boldsymbol{r}_1,r_2)]$ can make the notations extremely cumbersome and difficult to follow. Proofs which only involve DFTs can substantially reduce cumbersome notations. However, a DFT based proof requires the frequency grid to be bounded, and as mentioned in the discussion at the start of this section, $\widehat{a}_g(\omega_{k_2};\boldsymbol{r}_1,r_2)$ and $\widetilde{a}_g(\omega_{k_2};\boldsymbol{r}_1,r_2)$ are only asymptotically equivalent if the frequency grid is bounded. Therefore to simplify notations, for the remainder of this section we focus on the case that the frequency grid is bounded. However, we mention that exactly the same bounds apply to the case when the frequency grid is unbounded.

We observe that in order to obtain an expression for $\lambda^d \text{cov}[\widehat{a}_g(\omega_{k_1};\boldsymbol{r}_1,r_2),\widehat{a}_g(\omega_{k_2};\boldsymbol{r}_3,r_4)]$ (in the case that the frequency grid is bounded) we require the expansion

$$\lambda^d \text{cov}\left[\sum_{\boldsymbol{k}_1=-\boldsymbol{a}}^{\boldsymbol{a}} g(\boldsymbol{\Omega}_{\boldsymbol{k}_1})J_{t_1}(\boldsymbol{\Omega}_{\boldsymbol{k}_1})\overline{J_{t_2}(\boldsymbol{\Omega}_{\boldsymbol{k}_1+\boldsymbol{r}_1})},\sum_{\boldsymbol{k}_3=-\boldsymbol{a}}^{\boldsymbol{a}} g(\boldsymbol{\Omega}_{\boldsymbol{k}_3})J_{t_2}(\boldsymbol{\Omega}_{\boldsymbol{k}_3})\overline{J_{t_4}(\boldsymbol{\Omega}_{\boldsymbol{k}_3+\boldsymbol{r}_3})}\right] = \widetilde{A}+\widetilde{B}+\widetilde{C},$$

where

$$\begin{aligned}
\widetilde{A} &= \frac{1}{\lambda^d}\sum_{\boldsymbol{k}_1,\boldsymbol{k}_3=-a}^{a} g(\boldsymbol{\Omega}_{\boldsymbol{k}_1})\overline{g(\boldsymbol{\Omega}_{\boldsymbol{k}_3})}\text{cov}\left[J_{t_1}(\boldsymbol{\Omega}_{\boldsymbol{k}_1}),J_{t_3}(\boldsymbol{\Omega}_{\boldsymbol{k}_3})\right]\text{cov}\left[\overline{J_{t_2}(\boldsymbol{\Omega}_{\boldsymbol{k}_1+\boldsymbol{r}_1})},\overline{J_{t_4}(\boldsymbol{\Omega}_{\boldsymbol{k}_3+\boldsymbol{r}_3})}\right],\\
\widetilde{B} &= \frac{1}{\lambda^d}\sum_{\boldsymbol{k}_1,\boldsymbol{k}_3=-a}^{a} g(\boldsymbol{\Omega}_{\boldsymbol{k}_1})\overline{g(\boldsymbol{\Omega}_{\boldsymbol{k}_3})}\text{cov}\left[J_{t_1}(\boldsymbol{\Omega}_{\boldsymbol{k}_1}),\overline{J_{t_4}(\boldsymbol{\Omega}_{\boldsymbol{k}_3+\boldsymbol{r}_3})}\right]\text{cov}\left[\overline{J_{t_2}(\boldsymbol{\Omega}_{\boldsymbol{k}_1+\boldsymbol{r}_1})},J_{t_3}(\boldsymbol{\Omega}_{\boldsymbol{k}_3})\right],\\
\widetilde{C} &= \frac{1}{\lambda^d}\sum_{\boldsymbol{k}_1,\boldsymbol{k}_3=-a}^{a} g(\boldsymbol{\Omega}_{\boldsymbol{k}_1})\overline{g(\boldsymbol{\Omega}_{\boldsymbol{k}_3})}\text{cum}\left[J_{t_1}(\boldsymbol{\Omega}_{\boldsymbol{k}_1}),\overline{J_{t_2}(\boldsymbol{\Omega}_{\boldsymbol{k}_1+\boldsymbol{r}_1})},\overline{J_{t_3}(\boldsymbol{\Omega}_{\boldsymbol{k}_3})},J_{t_4}(\boldsymbol{\Omega}_{\boldsymbol{k}_3+\boldsymbol{r}_3})\right].
\end{aligned}$$

Now we obtain simplified expressions for $\widetilde{A}$, $\widetilde{B}$ and $\widetilde{C}$.

$$\widetilde{A} = \frac{I_{\boldsymbol{r}_1=\boldsymbol{r}_3}}{(2\pi)^d}\int_{\mathcal{D}}|g(\boldsymbol{\Omega})|^2 f_{t_3-t_1}(\boldsymbol{\Omega})\overline{f_{t_4-t_2}(\boldsymbol{\Omega}+\boldsymbol{\Omega}_{\boldsymbol{r}_1})}d\boldsymbol{\Omega}+R_{1,t_3-t_1,t_4-t_2}, \tag{50}$$

$$\widetilde{B} = \frac{I_{\boldsymbol{r}_1=\boldsymbol{r}_3}}{(2\pi)^d}\int_{\mathcal{D}_{\boldsymbol{r}_1}} g(\boldsymbol{\Omega})\overline{g(-\boldsymbol{\Omega}-\boldsymbol{\Omega}_{\boldsymbol{r}_1})}f_{t_4-t_1}(\boldsymbol{\Omega})\overline{f_{t_3-t_2}(\boldsymbol{\Omega}+\boldsymbol{\Omega}_{\boldsymbol{r}})}d\boldsymbol{\Omega}+R_{2,t_4-t_1,t_3-t_2}, \tag{51}$$

$$\begin{aligned}
\widetilde{C} = \frac{I_{\boldsymbol{r}_1=\boldsymbol{r}_3}}{(2\pi)^{2d}}\int_{\mathcal{D}^2} & g(\boldsymbol{\Omega}_1)\overline{g(\boldsymbol{\Omega}_2)}f_{t_2-t_1,t_3-t_1,t_4-t_1}(\boldsymbol{\Omega}_1+\boldsymbol{\Omega}_{\boldsymbol{r}_1},\boldsymbol{\Omega}_2,-\boldsymbol{\Omega}_2-\boldsymbol{\Omega}_{\boldsymbol{r}_1})d\boldsymbol{\Omega}_1 d\boldsymbol{\Omega}_2\\
&+R_{3,t_2-t_1,t_3-t_1,t_4-t_1}.
\end{aligned} \tag{52}$$

Comparing the above with (49), when the frequency grid is unbounded, see that the expressions are identical. We use the above to prove Lemma 4.2.

PROOF of Lemma 4.2 By decomposing the covariance we have

$$\lambda^d \text{cov}\left[\widehat{a}_g(\omega_{k_2};\boldsymbol{r}_1,r_2),\widehat{a}_g(\omega_{k_4};\boldsymbol{r}_3,r_4)\right] = I_{k_2,k_4}+II_{k_2,k_4}+III_{k_2,k_4},$$

where,

$$\begin{aligned}
I_{k_2,k_4} &= \frac{1}{\lambda^d}\sum_{\boldsymbol{k}_1,\boldsymbol{k}_3=-a}^{a} g(\boldsymbol{\Omega}_{\boldsymbol{k}_1})\overline{g(\boldsymbol{\Omega}_{\boldsymbol{k}_3})}\mathrm{cov}\left[J(\boldsymbol{\Omega}_{\boldsymbol{k}_1},\omega_{k_2}), J(\boldsymbol{\Omega}_{\boldsymbol{k}_3},\omega_{k_4})\right] \\
&\qquad \times\mathrm{cov}\left[\overline{J(\boldsymbol{\Omega}_{\boldsymbol{k}_1+\boldsymbol{r}_1},\omega_{k_2+r_2})}, \overline{J(\boldsymbol{\Omega}_{\boldsymbol{k}_3+\boldsymbol{r}_3},\omega_{k_4+r_4})}\right],
\end{aligned}$$

$$\begin{aligned}
II_{k_2,k_4} &= \frac{1}{\lambda^d}\sum_{\boldsymbol{k}_1,\boldsymbol{k}_3=-a}^{a} g(\boldsymbol{\Omega}_{\boldsymbol{k}_1})\overline{g(\boldsymbol{\Omega}_{\boldsymbol{k}_3})}\mathrm{cov}\left[J(\boldsymbol{\Omega}_{\boldsymbol{k}_1},\omega_{k_2}), \overline{J(\boldsymbol{\Omega}_{\boldsymbol{k}_3+\boldsymbol{r}_3},\omega_{k_4+r_4})}\right] \\
&\qquad \times\mathrm{cov}\left[\overline{J(\boldsymbol{\Omega}_{\boldsymbol{k}_1+\boldsymbol{r}_1},\omega_{k_2+r_2})}, J(\boldsymbol{\Omega}_{\boldsymbol{k}_3},\omega_{k_4})\right],
\end{aligned}$$

and

$$\begin{aligned}
III_{k_2,k_4} &= \frac{1}{\lambda^d}\sum_{\boldsymbol{k}_1,\boldsymbol{k}_3=-a}^{a} g(\boldsymbol{\Omega}_{\boldsymbol{k}_1})\overline{g(\boldsymbol{\Omega}_{\boldsymbol{k}_3})} \\
&\qquad \times\mathrm{cum}\left[J(\boldsymbol{\Omega}_{\boldsymbol{k}_1},\omega_{k_2}), \overline{J(\boldsymbol{\Omega}_{\boldsymbol{k}_1+\boldsymbol{r}_1},\omega_{k_2+r_2})}, \overline{J(\boldsymbol{\Omega}_{\boldsymbol{k}_3},\omega_{k_4})}, J(\boldsymbol{\Omega}_{\boldsymbol{k}_3+\boldsymbol{r}_3},\omega_{k_4+r_4})\right].
\end{aligned}$$

By using (50)-(51) we obtain expressions for the I_{k_2,k_4}, II_{k_2,k_4} and III_{k_2,k_4}. We first consider I_{k_2,k_4}. Using (50) we have,

$$\begin{aligned}
I_{k_2,k_4} &= \frac{1}{\lambda^d}\sum_{\boldsymbol{k}_1,\boldsymbol{k}_3=-a}^{a} g(\boldsymbol{\Omega}_{\boldsymbol{k}_1})\overline{g(\boldsymbol{\Omega}_{\boldsymbol{k}_3})}\mathrm{cov}\left[J(\boldsymbol{\Omega}_{\boldsymbol{k}_1},\omega_{k_2}), J(\boldsymbol{\Omega}_{\boldsymbol{k}_3},\omega_{k_4})\right] \\
&\qquad \times\mathrm{cov}\left[\overline{J(\boldsymbol{\Omega}_{\boldsymbol{k}_1+\boldsymbol{r}_1},\omega_{k_2+r_2})}, \overline{J(\boldsymbol{\Omega}_{\boldsymbol{k}_3+\boldsymbol{r}_3},\omega_{k_4+r_4})}\right] \\
&= I_{k_1,k_2,M} + I_{k_1,k_2,R}, \qquad (53)
\end{aligned}$$

where,

$$\begin{aligned}
I_{k_1,k_2,M} &= \frac{I_{\boldsymbol{r}_1=\boldsymbol{r}_3}}{(2\pi)^{d+2}T^2}\sum_{t_1,t_2,t_3,t_4=1}^{T}\int_{\mathcal{D}} |g(\boldsymbol{\Omega})|^2 f_{t_3-t_1}(\boldsymbol{\Omega})\overline{f_{t_4-t_2}(\boldsymbol{\Omega}+\boldsymbol{\Omega}_{\boldsymbol{r}})} \\
&\qquad \times e^{it_1\omega_{k_2}-it_2\omega_{k_2+r_2}-it_3\omega_{k_3}+it_4\omega_{k_4+r_4}}d\boldsymbol{\Omega}, \\
I_{k_1,k_2,R} &= \frac{1}{(2\pi)^{d+2}T^2}\sum_{t_1,t_2,t_3,t_4=1}^{T} R_{1,t_3-t_1,t_4-t_2}e^{it_1\omega_{k_2}-it_2\omega_{k_2+r_2}-it_3\omega_{k_3}+it_4\omega_{k_4+r_4}}.
\end{aligned}$$

We first find an expression for $I_{k_1,k_2,M}$

$$
\begin{aligned}
I_{k_1,k_2,M} &= \frac{I_{\boldsymbol{r}_1=\boldsymbol{r}_3}}{(2\pi)^{d+2}T^2}\int_{\mathcal{D}}|g(\boldsymbol{\Omega})|^2\left(\sum_{s_1=-(T-1)}^{T-1} f_{s_1}(\boldsymbol{\Omega})e^{-is_1\omega_{k_4}}\sum_{t_1=1}^{T-|s_1|}e^{it_1(\omega_{k_4}-\omega_{k_2})}\right)\times \\
&\quad \left(\sum_{s_2=-(T-1)}^{T-1}\overline{f_{s_2}(\boldsymbol{\Omega}+\boldsymbol{\Omega}_{\boldsymbol{r}})}e^{is_2\omega_{k_4+r_2}}\sum_{t_2=1}^{T-|s_2|}e^{it_2(\omega_{k_2+r_2}-\omega_{k_4+r_4})}\right)d\boldsymbol{\Omega} \\
&= \frac{I_{\boldsymbol{r}_1=\boldsymbol{r}_3}I_{k_2=k_4}I_{r_2=r_4}}{(2\pi)^d}\int_{\mathcal{D}}|g(\boldsymbol{\Omega})|^2 f(\boldsymbol{\Omega},\omega_{k_2})\overline{f(\boldsymbol{\Omega}+\boldsymbol{\Omega}_{\boldsymbol{r}_1},\omega_{k_2+r_2})}d\boldsymbol{\Omega}+O\left(\frac{I_{\boldsymbol{r}_1=\boldsymbol{r}_3}}{T}+\ell_{\lambda,a,n}\right) \\
&= \frac{I_{\boldsymbol{r}_1=\boldsymbol{r}_3}I_{k_2=k_4}I_{r_2=r_4}}{(2\pi)^d}\int_{\mathcal{D}}|g(\boldsymbol{\Omega})|^2 f(\boldsymbol{\Omega},\omega_{k_2})f(\boldsymbol{\Omega}+\boldsymbol{\Omega}_{\boldsymbol{r}_1},\omega_{k_2+r_2})d\boldsymbol{\Omega}+O\left(\frac{I_{\boldsymbol{r}_1=\boldsymbol{r}_3}}{T}+\ell_{\lambda,a,n}\right).
\end{aligned}
$$

It is straightforward to show that $I_{k_1,k_2,R}=O(\ell_{\lambda,a,n})$. Therefore we have

$$
I_{k_1,k_2} = \frac{I_{\boldsymbol{r}_1=\boldsymbol{r}_3}I_{k_2=k_4}I_{r_2=r_4}}{(2\pi)^d}\int_{\mathcal{D}}|g(\boldsymbol{\Omega})|^2 f(\boldsymbol{\Omega},\omega_{k_2})f(\boldsymbol{\Omega}+\boldsymbol{\Omega}_{\boldsymbol{r}_1},\omega_{k_2+r_2})d\boldsymbol{\Omega}+O\left(\frac{I_{\boldsymbol{r}_1=\boldsymbol{r}_3}}{T}+\ell_{\lambda,a,n}\right).
$$

Using the same arguments and (51)

$$
\begin{aligned}
II_{k_2,k_4} &= \frac{1}{\lambda^d}\sum_{\boldsymbol{k}_1,\boldsymbol{k}_3=-a}^{a} g(\boldsymbol{\Omega}_{\boldsymbol{k}_1})\overline{g(\boldsymbol{\Omega}_{\boldsymbol{k}_2})}\sum_{t_1,\ldots,t_4=1}^{T}\exp(it_1\omega_{k_2}+it_4\omega_{k_4+r_4}-it_2\omega_{k_2+r_2}-it_3\omega_{k_4}) \\
&\quad \mathrm{cov}\left[J_{t_1}(\boldsymbol{\Omega}_{\boldsymbol{k}_1}),\overline{J_{t_4}(\boldsymbol{\Omega}_{\boldsymbol{k}_3+\boldsymbol{r}_3})}\right]\mathrm{cov}\left[\overline{J_{t_2}(\boldsymbol{\Omega}_{\boldsymbol{k}_1+\boldsymbol{r}_1})},J_{t_3}(\boldsymbol{\Omega}_{\boldsymbol{k}_3})\right] \\
&= \frac{I_{\boldsymbol{r}_1=\boldsymbol{r}_3}}{(2\pi)^{d+2}T^2}\int_{\mathcal{D}_{\boldsymbol{r}_1}} g(\boldsymbol{\Omega})\overline{g(-\boldsymbol{\Omega}-\boldsymbol{\Omega}_{\boldsymbol{r}_1})}\left(\sum_{s_1=-(T-1)}^{T-1} f_{s_1}(\boldsymbol{\Omega})e^{is_1\omega_{k_2}}\sum_{t_1=1}^{T-|s_1|}e^{it_1(\omega_{k_4}+\omega_{k_2}+\omega_{r_2})}\right) \\
&\quad \times\left(\sum_{s_2=-(T-1)}^{T-1}\overline{f_{s_2}(\boldsymbol{\Omega}+\boldsymbol{\Omega}_{\boldsymbol{r}_1})}e^{-is_2\omega_{k_2+r_2}}\sum_{t_3=1}^{T-|s_2|}e^{it_3(\omega_{k_2}+\omega_{r_4}+\omega_{k_4})}\right)d\boldsymbol{\Omega}+O(\ell_{\lambda,a,n}) \\
&= \frac{I_{\boldsymbol{r}_1=\boldsymbol{r}_3}I_{k_4=T-k_2-r_2}I_{r_2=r_4}}{(2\pi)^d}\int_{\mathcal{D}_{\boldsymbol{r}_1}} g(\boldsymbol{\Omega})\overline{g(-\boldsymbol{\Omega}-\boldsymbol{\Omega}_{\boldsymbol{r}_1})}f(\boldsymbol{\Omega},-\omega_{k_2})f(-\boldsymbol{\Omega}-\boldsymbol{\Omega}_{\boldsymbol{r}_1},\omega_{k_2+r_2})d\boldsymbol{\Omega} \\
&\quad +O\left(\frac{I_{\boldsymbol{r}_1=\boldsymbol{r}_3}}{T}+\ell_{\lambda,a,n}\right). \qquad (54)
\end{aligned}
$$

Using (52) (see the proof of Theorem 4.1, Jentsch and Subba Rao [2015] for details) we have

$$
\begin{aligned}
& III_{k_2,k_4} \\
&= \frac{I_{\boldsymbol{r}_1=\boldsymbol{r}_3}}{(2\pi)^{2d}T^2}\int_{\mathcal{D}^2} g(\boldsymbol{\Omega}_1)\overline{g(\boldsymbol{\Omega}_2)} \sum_{t_1,t_2,t_3,t_4=1}^{T} f_{t_2-t_1,t_3-t_1,t_4-t_1}(\boldsymbol{\Omega}_1+\boldsymbol{\Omega}_{\boldsymbol{r}_1},\boldsymbol{\Omega}_2,-\boldsymbol{\Omega}_2-\boldsymbol{\Omega}_{\boldsymbol{r}_2}) \\
&\quad \times e^{it_1\omega_{k_2}-it_2\omega_{k_2+r_2}-it_3\omega_{k_4}+it_4\omega_{k_4+r_4}} d\boldsymbol{\Omega}_1 d\boldsymbol{\Omega}_2 \\
&\qquad + \frac{1}{(2\pi)^d T^2}\sum_{t_1,t_2,t_3,t_4=1}^{T} R_{3,t_2-t_1,t_3-t_1,t_4-t_1} e^{it_1\omega_{k_2}-it_2\omega_{k_2+r_2}-it_3\omega_{k_4}+it_4\omega_{k_4+r_4}} \\
&= \frac{I_{\boldsymbol{r}_1=\boldsymbol{r}_3}}{(2\pi)^{2d}T^2}\int_{\mathcal{D}^2} g(\boldsymbol{\Omega}_1)\overline{g(\boldsymbol{\Omega}_2)} \sum_{s_1,s_2,s_2=-(T-1)}^{T-1} f_{s_1,s_2,s_3}(\boldsymbol{\Omega}_1+\boldsymbol{\Omega}_{\boldsymbol{r}_1},\boldsymbol{\Omega}_2,-\boldsymbol{\Omega}_2-\boldsymbol{\Omega}_{\boldsymbol{r}_1}) \\
&\quad \times e^{is_1\omega_{k_2+r_2}+is_2\omega_{k_4}-is_3\omega_{k_4+r_4}} \sum_{t=|\min(s_i,0)|+1}^{T-|\max(s_i,0)|} e^{it(\omega_{k_2}-\omega_{k_2+r_2}-\omega_{k_4}+\omega_{k_4+r_4})} d\boldsymbol{\Omega}_1 d\boldsymbol{\Omega}_2 \\
&\quad + \frac{1}{(2\pi)^{2d}T^2}\sum_{s_1,s_2,s_2=-(T-1)}^{T-1} R_{3,s_1,s_2,s_3} e^{is_1\omega_{k_2+r_2}+is_2\omega_{k_4}-is_3\omega_{k_4+r_4}} \\
&\quad \times \sum_{t=|\min(s_i,0)|+1}^{T-|\max(s_i,0)|} e^{it(\omega_{k_2}-\omega_{k_2+r_2}-\omega_{k_4}+\omega_{k_4+r_4})}.
\end{aligned}
$$

By changing the limits of the sum we have

$$
\begin{aligned}
& III_{k_2,k_4} \\
&= \frac{I_{\boldsymbol{r}_1=\boldsymbol{r}_3}}{(2\pi)^{2d}T^2}\int_{\mathcal{D}^2} g(\boldsymbol{\Omega}_1)\overline{g(\boldsymbol{\Omega}_2)} \sum_{s_1,s_2,s_2=-(T-1)}^{T-1} f_{s_1,s_2,s_3}(\boldsymbol{\Omega}_1+\boldsymbol{\Omega}_{\boldsymbol{r}_1},\boldsymbol{\Omega}_2,-\boldsymbol{\Omega}_2-\boldsymbol{\Omega}_{\boldsymbol{r}_1}) \\
&\quad \times e^{-is_1\omega_{k_2+r_2}-is_2\omega_{k_4}+is_3\omega_{k_4+r_4}} \sum_{t=1}^{T} e^{it(\omega_{k_2}-\omega_{k_2+r_2}-\omega_{k_4}+\omega_{k_4+r_4})} d\boldsymbol{\Omega}_1 d\boldsymbol{\Omega}_2 + O\left(\ell_{\lambda,a,n}+\frac{1}{T^2}\right) \\
&= \frac{I_{\boldsymbol{r}_1=\boldsymbol{r}_3} I_{r_2=r_4}}{(2\pi)^{2d}T}\int_{\mathcal{D}^2} g(\boldsymbol{\Omega}_1)\overline{g(\boldsymbol{\Omega}_2)} f(\boldsymbol{\Omega}_1+\boldsymbol{\Omega}_{\boldsymbol{r}_1},\omega_{k_2}+\omega_{r_2},\boldsymbol{\Omega}_1,\omega_{k_2},-\boldsymbol{\Omega}_2-\boldsymbol{\Omega}_{\boldsymbol{r}_1},-\omega_{k_4}-\omega_{r_2}) \\
&\quad d\boldsymbol{\Omega}_1 d\boldsymbol{\Omega}_2 + O\left(\frac{\ell_{\lambda,a,n}}{T}+\frac{I_{\boldsymbol{r}_1=\boldsymbol{r}_3} I_{r_2=r_4}}{T^2}\right) \qquad (55)
\end{aligned}
$$

The above results imply

$$
\begin{aligned}
&\lambda^{d}\mathrm{cov}[\widehat{a}_{g}(\omega_{k_2};\boldsymbol{r}_1,r_2),\widehat{a}_{g}(\omega_{k_4};\boldsymbol{r}_3,r_4)] \\
&= \frac{I_{\boldsymbol{r}_1=\boldsymbol{r}_3}I_{r_2=r_4}}{(2\pi)^{d}}\Bigg(I_{k_2=k_4}\int_{\mathcal{D}} g(\boldsymbol{\Omega})\overline{g(\boldsymbol{\Omega})}f(\boldsymbol{\Omega}+\boldsymbol{\Omega}_{\boldsymbol{r}_1},\omega_{k_2+r_2})f(\boldsymbol{\Omega},\omega_{k_2})d\boldsymbol{\Omega} \\
&\qquad +I_{k_4=T-k_2-r_2}\int_{\mathcal{D}_{\boldsymbol{r}}} g(\boldsymbol{\Omega})\overline{g(-\boldsymbol{\Omega}-\boldsymbol{\Omega}_{\boldsymbol{r}_1})}f(\boldsymbol{\Omega},-\omega_{k_2})f(-\boldsymbol{\Omega}-\boldsymbol{\Omega}_{\boldsymbol{r}_1},\omega_{k_2+r_2})d\boldsymbol{\Omega}\Bigg) \\
&\qquad\quad +O\left(\ell_{\lambda,a,n}+\frac{1}{T}\right).
\end{aligned}
$$

By using the well known identities

$$
\begin{aligned}
\mathrm{cov}(\Re A,\Re B) &= \frac{1}{2}\big(\Re\mathrm{cov}(A,B)+\Re\mathrm{cov}(A,\bar{B})\big) \\
\mathrm{cov}(\Im A,\Im B) &= \frac{1}{2}\big(\Re\mathrm{cov}(A,B)-\Re\mathrm{cov}(A,\bar{B})\big), \\
\mathrm{cov}(\Re A,\Im B) &= \frac{-1}{2}\big(\Im\mathrm{cov}(A,B)-\Im\mathrm{cov}(A,\bar{B})\big),
\end{aligned} \tag{56}
$$

we immediately obtain (19).

Asymptotic normality is proved using sufficient mixing assumptions. □

A.2.1 PROOF of results in Section 6.1 (used in Section 4.2)

We start by analyzing the sampling properties of the first test statistic $\widehat{A}_{g,h}(\boldsymbol{r}_1,r_2)$.

PROOF of Lemma 6.1 We first note that

$$
\frac{\lambda^{d}T}{2}\mathrm{cov}\left[\widehat{A}_{g,h}(\boldsymbol{r}_1,r_2),\widehat{A}_{g,h}(\boldsymbol{r}_3,r_4)\right]=I+II+III,
$$

where,

$$
I = \frac{2}{T}\sum_{k_2,k_4=1}^{T/2}I_{k_2,k_4}\quad II=\frac{2}{T}\sum_{k_2,k_4=1}^{T/2}II_{k_2,k_4}\quad III=\frac{2}{T}\sum_{k_2,k_4=1}^{T/2}III_{k_2,k_4}
$$

and I_{k_2,k_4}, II_{k_2,k_4} and III_{k_2,k_4} are defined in the proof of Lemma 4.2. We now obtain expressions for these terms. By substituting the expression for I_{k_2,k_4} in (53) into I we have

$$\begin{aligned}
I &= \frac{2I_{\boldsymbol{r}_1=\boldsymbol{r}_3}}{(2\pi)^{d+2}T}\sum_{t_1,t_2,t_3,t_4=1}^{T}\int_{\mathcal{D}}|g(\boldsymbol{\Omega})|^2 f_{t_3-t_1}(\boldsymbol{\Omega})\overline{f_{t_4-t_2}(\boldsymbol{\Omega}+\boldsymbol{\Omega}_{\boldsymbol{r}})}d\boldsymbol{\Omega}\\
&\quad\times e^{-it_2\omega_{r_2}+it_4\omega_{r_4}}\left(\frac{1}{T^2}\sum_{k_2,k_4=1}^{T/2}h(\omega_{k_2})\overline{h(\omega_{k_4})}e^{i\omega_{k_2}(t_1-t_2)}e^{-i\omega_{k_4}(t_3-t_4)}\right)\\
&\quad+\frac{2}{T^3}\sum_{k_2,k_4=1}^{T/2}h(\omega_{k_2})\overline{h(\omega_{k_4})}\sum_{t_1,t_2,t_3,t_4=1}R_{1,t_3-t_1,t_4-t_2}e^{it_1\omega_{k_2}-it_2\omega_{k_2+r_2}-it_3\omega_{k_4}it_4\omega_{k_4+r_4}}\\
&= I_M+I_R.
\end{aligned}$$

We first obtain a neat expression for the leading term I_M. Using that the function $h:[0,\pi]\to\mathbb{R}$ is piecewise Lipschitz continuous and the integral approximation of the Riemann sum, we have

$$\frac{2}{T}\sum_{k=1}^{T/2}h(\omega_k)e^{ij\omega_k}=h_j+O(T^{-1})$$

where $h_j=\frac{1}{\pi}\int_0^\pi h(\omega)e^{ij\omega}d\omega$ and the Fourier coefficients decay at the rate $|h_j|\leq C|j|^{-1}I(j\neq 0)$. This approximation gives

$$\begin{aligned}
&\frac{4}{(2\pi)^2T^2}\sum_{k_2,k_4=1}^{T/2}h(\omega_{k_2})\overline{h(\omega_{k_4})}\exp(i\omega_{k_2}(t_1-t_2))\exp(-i\omega_{k_4}(t_3-t_4))\\
&\qquad= h_{t_1-t_2}\overline{h_{t_3-t_4}}+O(h_{t_1-t_2}T^{-1}+h_{t_3-t_4}T^{-1}+T^{-2}).
\end{aligned}$$

Substituting this into I_M and using that $|h_j|\leq C|j|^{-1}I(j\neq 0)$ gives

$$\begin{aligned}
I_M &= \frac{2I_{\boldsymbol{r}_1=\boldsymbol{r}_3}}{(2\pi)^{d+2}T}\sum_{t_1,t_2,t_3,t_4=1}^{T}h_{t_1-t_2}\overline{h_{t_3-t_4}}e^{-it_2\omega_{r_2}+it_4\omega_{r_4}}\int_{\mathcal{D}}|g(\boldsymbol{\Omega})|^2 f_{t_3-t_1}(\boldsymbol{\Omega})\overline{f_{t_4-t_2}(\boldsymbol{\Omega}+\boldsymbol{\Omega}_{\boldsymbol{r}})}d\boldsymbol{\Omega}\\
&\quad+O((\log T)T^{-1}).
\end{aligned}$$

By making the following change of variables, $s_1=t_3-t_1$, $s_2=t_4-t_2$ and $s_3=t_1-t_2$ (so

$t_3 - t_4 = s_1 - s_2 + s_3$) we have

$$\begin{aligned} I_M &= \frac{I_{\boldsymbol{r}_1=\boldsymbol{r}_3}}{(2\pi)^{d+2}T}\sum_{s_1,s_2,s_3,t_2} h_{s_3}\overline{h_{s_1-s_2+s_3}}e^{is_2\omega_{r_4}}e^{-it_2(\omega_{r_2}-\omega_{r_4})}\int_{\mathcal{D}}|g(\boldsymbol{\Omega})|^2 f_{s_1}(\boldsymbol{\Omega})\overline{f_{s_2}(\boldsymbol{\Omega}+\boldsymbol{\Omega}_{\boldsymbol{r}_1})}d\boldsymbol{\Omega} \\ &\quad +O((\log T)T^{-1}) \\ &= \frac{I_{\boldsymbol{r}_1=\boldsymbol{r}_3}I_{r_2=r_4}}{(2\pi)^{d+2}}\sum_{s_1,s_2,s_3} h_{s_3}\overline{h_{s_1-s_2+s_3}}e^{is_2\omega_{r_2}}\int_{\mathcal{D}}|g(\boldsymbol{\Omega})|^2 f_{s_1}(\boldsymbol{\Omega})\overline{f_{s_2}(\boldsymbol{\Omega}+\boldsymbol{\Omega}_{\boldsymbol{r}_1})}d\boldsymbol{\Omega} \\ &\quad +O((\log T)T^{-1}), \end{aligned}$$

where in the last term we have used that $T^{-1}\sum_{t=1}^{T} e^{-it_2(\omega_{r_2}-\omega_{r_4})} = I(r_1 = r_2)$. Next we use that $\sum_{s_3} h_{s_3}\overline{h_{s_3+(s_1-s_2)}} = \frac{1}{\pi}\int |h(\omega)|^2 \exp(-i\omega(s_1-s_2))d\omega$ to give

$$\begin{aligned} I_M &= \frac{I_{\boldsymbol{r}_1=\boldsymbol{r}_3}I_{r_2=r_4}}{\pi(2\pi)^{d+2}}\int_0^{\pi}|h(\omega)|^2\sum_{s_1,s_2} e^{-i\omega(s_1-s_2)}e^{is_2\omega_{r_2}}\int_{\mathcal{D}}|g(\boldsymbol{\Omega})|^2 f_{s_1}(\boldsymbol{\Omega})\overline{f_{s_2}(\boldsymbol{\Omega}+\boldsymbol{\Omega}_{\boldsymbol{r}_1})}d\boldsymbol{\Omega}d\omega \\ &\quad +O((\log T)T^{-1}) \\ &= \frac{I_{\boldsymbol{r}_1=\boldsymbol{r}_3}I_{r_2=r_4}}{\pi(2\pi)^{d}}\int_0^{\pi}\int_{\mathcal{D}}|h(\omega)|^2|g(\boldsymbol{\Omega})|^2 f(\boldsymbol{\Omega},\omega)\overline{f(\boldsymbol{\Omega}+\boldsymbol{\Omega}_{\boldsymbol{r}_1},\omega+\omega_{r_2})}d\boldsymbol{\Omega},d\omega \\ &\quad +O((\log T)T^{-1}) \end{aligned}$$

By using a similar method we can show that $|I_R| = O(\ell_{\lambda,a,n})$. Altogether (using that f is real) we get

$$\begin{aligned} I &= \frac{I_{\boldsymbol{r}_1=\boldsymbol{r}_3}I_{r_2=r_4}}{(2\pi)^{d}}\int_0^{\pi}\int_{\mathcal{D}}|h(\omega)|^2|g(\boldsymbol{\Omega})|^2 f(\boldsymbol{\Omega},\omega)f(\boldsymbol{\Omega}+\boldsymbol{\Omega}_{\boldsymbol{r}_1},\omega+\omega_{r_2})d\boldsymbol{\Omega}d\omega \\ &\quad +O((\log T)T^{-1}+\ell_{\lambda,a,n}) \\ &= I_{\boldsymbol{r}_1=\boldsymbol{r}_3}I_{r_2=r_4}\int_0^{\pi}|h(\omega)|^2 V_g(\omega;\boldsymbol{\Omega}_{\boldsymbol{r}_1},\omega_{r_2})d\omega. \end{aligned}$$

Using similar arguments we can show that

$$\begin{aligned} II &= \frac{I_{\boldsymbol{r}_1=\boldsymbol{r}_3}}{(2\pi)^{d+2}T}\sum_{t_1,t_2,t_3,t_4=1}^{T}\int_{\mathcal{D}_{\boldsymbol{r}_1}} g(\boldsymbol{\Omega})\overline{g(-\boldsymbol{\Omega}-\boldsymbol{\Omega}_{\boldsymbol{r}})}f_{t_4-t_1}(\boldsymbol{\Omega})\overline{f_{t_3-t_2}(\boldsymbol{\Omega}+\boldsymbol{\Omega}_{\boldsymbol{r}})}d\boldsymbol{\Omega} \\ &\quad \times e^{it_2\omega_{r_2}+it_4\omega_{r_4}}\left(\frac{1}{T^2}\sum_{k_2,k_4=1}^{T/2} h(\omega_{k_2})\overline{h(\omega_{k_4})}e^{i\omega_{k_2}(t_1-t_2)}e^{-i\omega_{k_4}(t_3-t_4)}\right)+O(\ell_{\lambda,a,n}). \end{aligned}$$

We set $s_1 = t_4 - t_1$, $s_2 = t_3 - t_2$, $s_3 = t_1 - t_2$ (and $t_3 - t_4 = s_2 - s_1 - s_3$) to give

$$\begin{aligned} II &= \frac{I_{\boldsymbol{r}_1=\boldsymbol{r}_3}}{(2\pi)^{d+2}T} \sum_{s_1,s_2,s_3,t_2} h_{s_3}\overline{h_{s_2-s_1-s_3}} \int_{\mathcal{D}_{\boldsymbol{r}_1}} g(\boldsymbol{\Omega})\overline{g(-\boldsymbol{\Omega}-\boldsymbol{\Omega}_{\boldsymbol{r}})} f_{s_1}(\boldsymbol{\Omega}) f_{s_2}(\boldsymbol{\Omega}+\boldsymbol{\Omega}_{\boldsymbol{r}}) d\boldsymbol{\Omega} \\ &\quad \times e^{it_2\omega_{r_2}+i(s_1+s_3+t_2)\omega_{r_4}} + O(\ell_{\lambda,a,n} + (\log T)T^{-1}). \end{aligned}$$

By changing the limits of the sum over t_2 we have

$$\begin{aligned} II &= \frac{I_{\boldsymbol{r}_1=\boldsymbol{r}_3}}{(2\pi)^{d+2}T} \sum_{s_1,s_2,s_3} h_{s_3}\overline{h_{s_2-s_1-s_3}} \int_{\mathcal{D}_{\boldsymbol{r}_1}} g(\boldsymbol{\Omega})\overline{g(-\boldsymbol{\Omega}-\boldsymbol{\Omega}_{\boldsymbol{r}})} f_{s_1}(\boldsymbol{\Omega}) f_{s_2}(\boldsymbol{\Omega}+\boldsymbol{\Omega}_{\boldsymbol{r}}) d\boldsymbol{\Omega} \\ &\quad \times e^{i(s_1+s_3)\omega_{r_4}} \underbrace{\sum_{t_2=1}^{T} e^{it_2(\omega_{r_1}+\omega_{r_2})}}_{r_1=T-r_2} + O(\ell_{\lambda,a,n} + (\log T)T^{-1}) = O(\ell_{\lambda,a,n} + (\log T)T^{-1}), \end{aligned}$$

where the last line follows from the fact that r_1 and r_2 are constrained such that $0 \leq r_1 \leq r_2 < T/2$. The following expression for III follows immediately from (55).

$$\begin{aligned} III &= \frac{I_{\boldsymbol{r}_1=\boldsymbol{r}_3} I_{r_2=r_4}}{\pi^2(2\pi)^{2d}} \int_0^\pi \int_0^\pi \int_{\mathcal{D}^2} g(\boldsymbol{\Omega}_1)\overline{g(\boldsymbol{\Omega}_2)} h(\omega_1)\overline{h(\omega_2)} \\ &\quad \times f_4(\boldsymbol{\Omega}_1+\boldsymbol{\Omega}_{\boldsymbol{r}_1}, \omega_1+\omega_{r_2}, \boldsymbol{\Omega}_2, \omega_2, -\boldsymbol{\Omega}_2-\boldsymbol{\Omega}_{\boldsymbol{r}_1}, -\omega_2-\omega_{r_2}) d\boldsymbol{\Omega}_1 d\boldsymbol{\Omega}_2 d\omega_1 d\omega_2 \\ &\qquad + O((\log T)T^{-1} + \ell_{\lambda,a,n}). \end{aligned}$$

This gives us

$$\begin{aligned} &\frac{\lambda^d T}{2} \mathrm{cov}\left[\widehat{A}_{g,h}(\boldsymbol{r}_1, r_2), \widehat{A}_{g,h}(\boldsymbol{r}_3, r_4)\right] \\ &= I_{\boldsymbol{r}_1=\boldsymbol{r}_3} I_{r_2=r_4} \Bigg(\frac{1}{\pi} \int_0^\pi |h(\omega)|^2 V_g(\omega; \boldsymbol{\Omega}_{\boldsymbol{r}_1}, \omega_{r_2}) d\omega + \frac{1}{(2\pi)^{2d}\pi^2} \int_0^\pi \int_0^\pi \int_{\mathcal{D}^2} g(\boldsymbol{\Omega}_1)\overline{g(\boldsymbol{\Omega}_2)} \\ &\quad \times h(\omega_1)\overline{h(\omega_2)} f_4(\boldsymbol{\Omega}_1+\boldsymbol{\Omega}_{\boldsymbol{r}_1}, \omega_1+\omega_{r_2}, \boldsymbol{\Omega}_2, \omega_2, -\boldsymbol{\Omega}_2-\boldsymbol{\Omega}_{\boldsymbol{r}_1}, -\omega_2-\omega_{r_2}) d\boldsymbol{\Omega}_1 d\boldsymbol{\Omega}_2 d\omega_1 d\omega_2 \Bigg) \\ &\quad + O((\log T)T^{-1} + \ell_{\lambda,a,n}). \end{aligned}$$

Note that $1/[(2\pi)^d\pi^2] = 4/(2\pi)^{2d+2}$ gives the fourth order cumulant term in (41).

By using the expressions for I, II and III and (56), we obtain (41).

By using mixing-type arguments the CLT can be proved. □

A.2.2 Proof of results in Section 6.1 (used in Section 4.3)

PROOF of Lemma 6.2 equation (42) Expanding $\mathrm{cov}\left[B_{g,h;H}(\omega_{j_1H}; \boldsymbol{r}_1, r_2), B_{g,h;H}(\omega_{j_2H}; \boldsymbol{r}_3, r_4)\right]$

gives

$$\lambda^{d} H \text{cov}\left[B_{g,h;H}(\omega_{j_1 H}; \boldsymbol{r}_1, r_2), B_{g,h;H}(\omega_{j_2 H}; \boldsymbol{r}_3, r_4)\right] = I_H + II_H + III_H,$$

where

$$\begin{aligned}
I_H &= \frac{1}{H}\sum_{k_2,k_4=1}^{H} I_{j_1 H+k_2, j_2 H+k_4}, \\
II_H &= \frac{1}{H}\sum_{k_2,k_4=1}^{H} II_{j_1 H+k_2, j_2 H+k_4}, \\
III_H &= \frac{1}{H}\sum_{k_2,k_4=1}^{H} III_{j_1 H+k_2, j_2 H+k_4},
\end{aligned}$$

and I_{k_2,k_4}, II_{k_2,k_4} and III_{k_2,k_4} are defined in the proof of Lemma 4.2. We now find expressions for these terms, first focusing on the case $j_1 = j_2 = j$. By using (54) we have,

$$I_H = I_{H,M} + I_{H,R},$$

where

$$\begin{aligned}
I_{H,M} &= \frac{H I_{\boldsymbol{r}_1=\boldsymbol{r}_3}}{(2\pi)^{d+2}T^2}\sum_{t_1,t_2,t_3,t_4=1}^{T}\int_{\mathcal{D}}|g(\boldsymbol{\Omega})|^2 f_{t_3-t_1}(\boldsymbol{\Omega})\overline{f_{t_4-t_2}(\boldsymbol{\Omega}+\boldsymbol{\Omega}_{\boldsymbol{r}})}d\boldsymbol{\Omega}\times e^{-it_2\omega_{r_2}+it_4\omega_{r_4}} \\
&\quad \times\left(\frac{1}{H^2}\sum_{k_2,k_4=1}^{H} h(\omega_{jH}+\omega_{k_2})\overline{h(\omega_{jH}+\omega_{k_4})}e^{i(\omega_{jH}+\omega_{k_2})(t_1-t_2)}e^{-i(\omega_{jH}+\omega_{k_4})(t_3-t_4)}\right) \\
I_{H,R} &= \frac{1}{T^2 H}\sum_{k_2,k_4=1}^{H} h(\omega_{jH+k_2})\overline{h(\omega_{jH+k_4})} \\
&\qquad \sum_{t_1,t_2,t_3,t_4=1} R_{1,t_3-t_1,t_4-t_2}e^{it_1\omega_{jH+k_2}-it_2\omega_{jH+k_2+r_2}-it_3\omega_{jH+k_4}+it_4\omega_{jH+k_4+r_4}}.
\end{aligned}$$

We first bound the inner sum in $I_{H,M}$. Using the approximation of the Riemann sum by an integral we have,

$$\frac{1}{H}\sum_{k=1}^{H} h(\omega_{jH}+\omega_k)e^{is\omega_k} = \frac{T}{H}\int_{\omega_{jH}}^{\omega_{(j+1)H}} h(\omega)e^{is\omega}d\omega + O(H^{-1}) = h_{s,H}(\omega_{jH}) + O(H^{-1}). \tag{57}$$

Applying the above to the following product gives

$$\begin{aligned} & \frac{1}{H^2}\sum_{k_2,k_4=1}^{H} h(\omega_{jH}+\omega_{k_2})\overline{h(\omega_{jH}+\omega_{k_4})}e^{i(\omega_{jH}+\omega_{k_2})(t_1-t_2)}e^{-i(\omega_{jH}+\omega_{k_4})(t_3-t_4)} \\ = \; & h_{t_1-t_2,H}(\omega_{jH})\overline{h_{t_3-t_4,H}(\omega_{jH})}+O(h_{t_1-t_2,H}(\omega_{jH})H^{-1}+h_{t_3-t_4,H}(\omega_{jH})H^{-1}+H^{-2}). \end{aligned}$$

Substituting the above into $I_{H,M}$, using that

$$\frac{H}{(2\pi)^{d+2}T^2}\sum_{t_1,t_2,t_3,t_4=1}^{T}\int_{\mathcal{D}}|g(\boldsymbol{\Omega})|^2 f_{t_3-t_1}(\boldsymbol{\Omega})\overline{f_{t_4-t_2}(\boldsymbol{\Omega}+\boldsymbol{\Omega}_{\boldsymbol{r}})}d\boldsymbol{\Omega}=O(H)$$

and the same arguments used to bound I_M in the proof of Lemma 6.1 we have,

$$\begin{aligned} I_{H,M} \;=\; & \frac{I_{\boldsymbol{r}_1=\boldsymbol{r}_3}I_{r_2=r_4}T}{(2\pi)^{d+2}H}\int_{2\pi\omega_{jH}}^{2\pi\omega_{(j+1)H}}\int_{\mathcal{D}}|h(\omega)|^2|g(\boldsymbol{\Omega})|^2 f(\boldsymbol{\Omega},\omega)f(\boldsymbol{\Omega}+\boldsymbol{\Omega}_{\boldsymbol{r}},\omega+\omega_{r_2})d\boldsymbol{\Omega}d\omega \\ & +O(H^{-1}+(\log T)T^{-1}). \end{aligned}$$

Using the same argument we can show that $I_{H,R}=O(\ell_{\lambda,a,n})$, which gives altogether

$$\begin{aligned} I_H \;=\; & \frac{I_{\boldsymbol{r}_1=\boldsymbol{r}_3}I_{r_2=r_4}T}{(2\pi)^{d+2}H}\int_{\omega_{jH}}^{\omega_{(j+1)H}}\int_{\mathcal{D}}|h(\omega)|^2|g(\boldsymbol{\Omega})|^2 f(\boldsymbol{\Omega},\omega)f(\boldsymbol{\Omega}+\boldsymbol{\Omega}_{\boldsymbol{r}},\omega+\omega_{r_2})d\boldsymbol{\Omega}d\omega \\ & +O(H^{-1}+(\log T)T^{-1}+\ell_{\lambda,a,n}). \end{aligned}$$

Using the same methods, we can show that $II_H = O(H^{-1}+(\log T)T^{-1}+\ell_{\lambda,a,n})$ (since $\leq r_2, r_4 \leq T/2$). Finally to bound III_H we substitute (55) into III_H to give

$$\begin{aligned} III_H \;=\; & \frac{I_{\boldsymbol{r}_1=\boldsymbol{r}_3}I_{r_2=r_4}}{(2\pi)^{2d}TH}\sum_{k_2,k_4=1}^{H}\int_{\mathcal{D}^2} g(\boldsymbol{\Omega}_1)\overline{g(\boldsymbol{\Omega}_2)}\times \\ & f(\boldsymbol{\Omega}_1,\omega_{jH+k_2},-\boldsymbol{\Omega}_1-\boldsymbol{\Omega}_{\boldsymbol{r}_1},-\omega_{jH+k_2+r_2},-\boldsymbol{\Omega}_2,-\omega_{jH+k_4})d\boldsymbol{\Omega}_1 d\boldsymbol{\Omega}_2 \\ & \quad +O\left(\frac{H\ell_{\lambda,a,n}}{T}+\frac{HI_{\boldsymbol{r}_1=\boldsymbol{r}_3}I_{r_2=r_4}}{T^2}\right) \end{aligned}$$

By using (57) we have

$$\begin{aligned} III_H \;=\; & \frac{TI_{\boldsymbol{r}_1=\boldsymbol{r}_3}I_{r_2=r_4}}{H(2\pi)^{2d+2}}\int_{\omega_{jH}}^{\omega_{(j+1)H}}\int_{\omega_{jH}}^{\omega_{(j+1)H}}\int_{\mathcal{D}^2} g(\boldsymbol{\Omega}_1)\overline{g(\boldsymbol{\Omega}_2)}h(\omega_1)\overline{h(\omega_2)} \\ & \times f_4(\boldsymbol{\Omega}_1+\boldsymbol{\Omega}_{\boldsymbol{r}_1},\boldsymbol{\Omega}_2,\omega_2,-\boldsymbol{\Omega}_2-\boldsymbol{\Omega}_{\boldsymbol{r}_1},-\omega_2-\omega_{r_2})d\boldsymbol{\Omega}_1 d\boldsymbol{\Omega}_2 d\omega_1 d\omega_2 \\ & \quad +O((\log T)T^{-1}+\ell_{\lambda,a,n}+H^{-1}). \end{aligned}$$

We observe that $III_H = O(H/T)$. Thus by using (56) we obtain (42) and a similar expression for the imaginary parts. $\square$

PROOF of Lemma 6.2 equation (43) The proof of (43) follows immediately from (41). $\square$

Finally we consider the sampling properties of $\widehat{D}_{g,h,v;H}(\boldsymbol{r}_1, r_2)$.

PROOF of Lemma 6.3. To prove (44) we expand the expectation squared in terms of covariance and expectations to give

$$\mathrm{E}[\lambda^d D_{g,h,v;H}(\boldsymbol{r}_1, r_2)] = I + II$$

where

$$\begin{aligned} I &= \frac{2H}{2T}\sum_{j=0}^{(T/2H)-1} \mathrm{var}[\sqrt{H\lambda^d} B_{g,h;H}(\omega_{j_1H}; \boldsymbol{r}_1, r_2)], \text{ and} \\ II &= \frac{2H}{2T}\sum_{j=0}^{(T/2H)-1} \left| \mathrm{E}[\sqrt{\lambda^d H} B_{g,h;H}(\omega_{j_1H}; \boldsymbol{r}_1, r_2)] \right|^2. \end{aligned}$$

Using (42) we have

$$\begin{aligned} I &= \frac{2H}{T}\sum_{j=0}^{T/(2H)-1} \frac{2}{2v(\omega_{j_1H})}\Big(\mathrm{var}[\sqrt{H\lambda^d}\Re B_{g,h;H}(\omega_{j_1H}; \boldsymbol{r}_1, r_2)] \\ &\qquad + \mathrm{var}[\sqrt{H\lambda^d}\Im B_{g,h;H}(\omega_{j_1H}; \boldsymbol{r}_1, r_2)]\Big) + O(\ell_{\lambda,a,n}) \\ &= \frac{2H}{T}\sum_{j=0}^{T/(2H)-1} \frac{1}{v(\omega_{j_1H})} W_{g,h}(\omega_{jH}; \boldsymbol{\Omega}_{\boldsymbol{r}_1}, \omega_{r_2}) + O\left(\ell_{\lambda,a,n} + \frac{1}{H}\right) \\ &= \frac{1}{\pi}\int_0^{\pi} \frac{W_{g,h}(\omega; \boldsymbol{\Omega}_{\boldsymbol{r}_1}, \omega_{r_2})}{v(\omega)} d\omega + O\left(\ell_{\lambda,a,n} + \frac{1}{H} + \frac{H}{T}\right) \\ &= E_{g,h,v}(\boldsymbol{\Omega}_{\boldsymbol{r}_1}, \omega_{r_2}) + O\left(\ell_{\lambda,a,n} + \frac{1}{H} + \frac{H}{T}\right). \end{aligned}$$

Next we consider the second term II. First considering the expectation we note that

$$\mathrm{E}[\sqrt{\lambda^d H} B_{g,h;H}(\omega_{j_1H}; \boldsymbol{r}_1, r_2)] = \frac{\sqrt{\lambda^{d/2}}}{\sqrt{H}} \sum_{k=1}^{H} h(\omega_{jH+k}) \mathrm{E}[\widehat{a}_g(\omega_{jH+k}; \boldsymbol{r}_1, r_2)].$$

By using Lemma 4.1 we obtain bounds on $\mathrm{E}[\widehat{a}_g(\omega_{jH+k}; \boldsymbol{r}_1, r_2)]$, however, these rely on the number of zeros in $\boldsymbol{r}_1$ and whether r_2 is zero or not. More precisely,

$$\left|\frac{\sqrt{\lambda^{d/2}}}{\sqrt{H}}\sum_{k=1}^{H}h(\omega_{jH+k})\mathrm{E}[\widehat{a}_g(\omega_{jH+k};\boldsymbol{r}_1,r_2)]\right| = O\left(\frac{\lambda^{d/2}H^{1/2}\prod_{j=1}^{d-b}(\log\lambda+\log|m_j|)}{T^{I_{r_2-r_4\neq 0}}\lambda^{d-b}}\right).$$

Therefore,

$$II = O\left(\frac{\lambda^d H[\prod_{j=1}^{d-b}(\log\lambda+\log|m_j|)]^2}{(T^{I_{r_2-r_4\neq 0}}\lambda^{d-b})^2}\right) = o(1).$$

This proves (44).

To prove (45) we expand the covariance in terms of products cumulants to give

$$\begin{aligned}
&\frac{T}{2M}\mathrm{cov}\left[\lambda^d D_{g,h,v;H}(\boldsymbol{r}_1,r_2),\lambda^d D_{g,h,v;H}(\boldsymbol{r}_3,r_4)\right]\\
&= \frac{2\lambda^{2d}H}{T}\sum_{j_1,j_2=0}^{(T/2H)-1}\frac{1}{v(\omega_{j_1H})\overline{v(\omega_{j_2H})}}\Bigg(|\mathrm{cov}[B_{g,h;H}(\omega_{j_1H};\boldsymbol{r}_1,r_2),B_{g,h;H}(\omega_{j_2H};\boldsymbol{r}_1,r_2)]|^2\\
&\quad+\left|\mathrm{cov}[B_{g,h;H}(\omega_{j_1H};\boldsymbol{r}_1,r_2),\overline{B_{g,h;H}(\omega_{j_2H};\boldsymbol{r}_1,r_2)}]\right|^2\\
&\quad+\mathrm{cum}\left[B_{g,h;H}(\omega_{j_1H};\boldsymbol{r}_1,r_2),\overline{B_{g,h;H}(\omega_{j_1H};\boldsymbol{r}_1,r_2)},B_{g,h;H}(\omega_{j_2H};\boldsymbol{r}_1,r_2),\overline{B_{g,h;H}(\omega_{j_2H};\boldsymbol{r}_1,r_2)}\right]\\
&\quad+\mathrm{E}[B_{g,h;H}(\omega_{j_1H};\boldsymbol{r}_1,r_2)]\mathrm{cum}\left[\overline{B_{g,h;H}(\omega_{j_1H};\boldsymbol{r}_1,r_2)},B_{g,h;H}(\omega_{j_2H};\boldsymbol{r}_1,r_2),\overline{B_{g,h;H}(\omega_{j_2H};\boldsymbol{r}_1,r_2)}\right]\\
&\quad+\text{ similar terms involving the product of third and first order cumulants}\Bigg).
\end{aligned}$$

By using that

$$\begin{aligned}
&\frac{1}{\lambda^{3d}}\sum_{\boldsymbol{k}_1,\boldsymbol{k}_3,\boldsymbol{k}_5=-a}^{a}g(\boldsymbol{\Omega}_{\boldsymbol{k}_1})g(\boldsymbol{\Omega}_{\boldsymbol{k}_3})g(\boldsymbol{\Omega}_{\boldsymbol{k}_5})\mathrm{cum}\Big[J_{t_1}(\boldsymbol{\Omega}_{\boldsymbol{k}_1})\overline{J_{t_2}(\boldsymbol{\Omega}_{\boldsymbol{k}_1+\boldsymbol{r}_1})},J_{t_3}(\boldsymbol{\Omega}_{\boldsymbol{k}_3})\overline{J_{t_4}(\boldsymbol{\Omega}_{\boldsymbol{k}_3+\boldsymbol{r}_3})},\\
&J_{t_5}(\boldsymbol{\Omega}_{\boldsymbol{k}_5})\overline{J_{t_6}(\boldsymbol{\Omega}_{\boldsymbol{k}_5+\boldsymbol{r}_5})}\Big] = O\left(\sum_{\mathcal{B}_3}\prod_{(t_i,t_j)\in\mathcal{B}_3}\rho_{t_i-t_j}\frac{\log^{3d}(a)}{\lambda^{2d}}\right)
\end{aligned}\tag{58}$$

and

$$\begin{aligned}
&\frac{1}{\lambda^{4d}}\sum_{\boldsymbol{k}_1,\boldsymbol{k}_3,\boldsymbol{k}_5,\boldsymbol{k}_7=-a}^{a}g(\boldsymbol{\Omega}_{\boldsymbol{k}_1})\overline{g(\boldsymbol{\Omega}_{\boldsymbol{k}_3})g(\boldsymbol{\Omega}_{\boldsymbol{k}_5})}g(\boldsymbol{\Omega}_{\boldsymbol{k}_7})\mathrm{cum}\Big[J_{t_1}(\boldsymbol{\Omega}_{\boldsymbol{k}_1})\overline{J_{t_2}(\boldsymbol{\Omega}_{\boldsymbol{k}_1+\boldsymbol{r}_1})},\overline{J_{t_3}(\boldsymbol{\Omega}_{\boldsymbol{k}_3})}J_{t_4}(\boldsymbol{\Omega}_{\boldsymbol{k}_3+\boldsymbol{r}_3}),\\
&J_{t_5}(\boldsymbol{\Omega}_{\boldsymbol{k}_5})\overline{J_{t_6}(\boldsymbol{\Omega}_{\boldsymbol{k}_5+\boldsymbol{r}_1})},\overline{J_{t_7}(\boldsymbol{\Omega}_{\boldsymbol{k}_3})}J_{t_8}(\boldsymbol{\Omega}_{\boldsymbol{k}_7+\boldsymbol{r}_3})\Big] = O\left(\sum_{\mathcal{B}_4}\prod_{(t_i,t_j)\in\mathcal{B}_4}\rho_{t_i-t_j}\frac{\log^{4d}(a)}{\lambda^{3d}}\right),
\end{aligned}\tag{59}$$

where $\mathcal{B}_3$ and $\mathcal{B}_4$ denotes the set of all pairwise indecomposable partitions of the sets $\{1,2,3\}\times\{4,5,6\}$ and $\{1,2,3,4\}\times\{5,4,6,7\}$ (for example, it contains the element $(1,4),(3,6)$, $(5,8),(2,7)$) respectively, we can show that

$$\lambda^{3d/2}\mathrm{cum}\left[a_g(\omega_{k_2};\boldsymbol{r}_1,r_2),\overline{a_g(\omega_{k_2};\boldsymbol{r}_1,r_2)},\overline{a_g(\omega_{k_4};\boldsymbol{r}_1,r_2)}\right] = O\left(\frac{\log^{3d}(a)}{\lambda^{d/2}}\right)$$

$$\lambda^{2d}\mathrm{cum}\left[a_g(\omega_{k_2};\boldsymbol{r}_1,r_2),\overline{a_g(\omega_{k_2};\boldsymbol{r}_1,r_2)},\overline{a_g(\omega_{k_4};\boldsymbol{r}_1,r_2)},a_g(\omega_{k_4};\boldsymbol{r}_1,r_2)\right] = O\left(\frac{\log^{4d}(a)}{\lambda^{d}}\right).$$

From this we expect (by using the methods detailed in the proof of Lemma B.5, Eichler [2008]), though a formal proof is not given, that the terms involving cumulants of order three and above are asymptotically negligible. Moreover that $\left|\mathrm{cov}[B_{g,h;H}(\omega_{j_1H};\boldsymbol{r}_1,r_2),\overline{B_{g,h;H}(\omega_{j_2H};\boldsymbol{r}_1,r_2)}]\right|^2$ is asymptotically negligible for $j_1\neq j_2$. Using this we have

$$\begin{aligned}&\frac{T}{2M}\mathrm{cov}\left[\lambda^d D_{g,h,v;H}(\boldsymbol{r}_1,r_2),\lambda^d D_{g,h,v;H}(\boldsymbol{r}_3,r_4)\right]\\ =\;&\frac{2\lambda^{2d}H}{T}\sum_{j_1,j_2=0}^{(T/2H)-1}\frac{1}{v(\omega_{j_1H})\overline{v(\omega_{j_2H})}}\left|\mathrm{cov}[B_{g,h;H}(\omega_{j_1H};\boldsymbol{r}_1,r_2),B_{g,h;H}(\omega_{j_2H};\boldsymbol{r}_1,r_2)]\right|^2.\end{aligned}$$

Substituting (42) into the above gives (45). □