A spectral domain test for stationarity of spatio-temporal data

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

September 9, 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 spatio-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 stationary 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 of 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.

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).

The proposed tests are illustrated with simulations in Section 8 of the supplementary material. And a rough outline of the proofs is also given in the appendix of 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.

2.1 Testing for temporal stationarity

Let us suppose that $\{X_t\}$ is a stationary time series where $c_h = \text{cov}[X_t, X_{t+h}]$ 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. In the case of a second-order stationary time series process $\{X_t\}$ (and the short memory condition stated above), it is well-known that for $1 \leq k_1 \neq k_2 \leq \lfloor T/2 \rfloor$ (to ease notation, from now onwards we assume that T is even) $\text{cov}(J_T(\omega_{k_1}), J_T(\omega_{k_2})) = O(\frac{1}{T})$ holds (uniformly in T, k_1 and k_2).

That is, the DFT transforms a second order stationary time series into a 'near uncorrelated sequence' $\{J_T(\omega_k)\}$ whose variance is approximately equal to the spectral density $f(\omega_k) = \frac{1}{2\pi}\sum_{h\in\mathbb{Z}} c_h \exp(-ih\omega_k)$. In the case of second order nonstationarity, the behavior of $\{J_T(\omega_k)\}$ is indeed much different. This observation has been exploited by Dwivedi and Subba Rao [2011] and Jentsch and Subba Rao [2015] to construct tests for second-order stationarity. 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. In contrast to (second order) stationary time series, the DFT sequence of a nonstationary time series shows a non-vanishing linear dependence structure. Hence, it turns out to be natural to estimate this linear dependence by covariance-type quantities and to construct a test statistic that measures their deviation from zero. Instead of using the 'raw' DFTs, Jentsch and Subba Rao [2015] propose to use the 'standardized' DFTs to define the following estimator of the covariance between the DFTs at 'lag' r by

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

where $\widehat{f}_T$ is the smoothed periodogram to estimate the spectral density f. Note that $J_T(\omega_k)$ has approximately variance $f(\omega_k)$. 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. Under the assumption of fourth order stationarity, Jentsch and Subba Rao [2015] showed that the approximate 'variance' (in terms of the limiting distribution) of both $\Re\widehat{C}_T(r,\ell)$ and $\Im\widehat{C}_T(r,\ell)$ (where $\Re x$ and $\Im x$ denote the real and imaginary parts of x) is

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

with

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

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

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

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 (fourth order) stationary and Gaussian, we have $\kappa_\ell(0) = 0$ and $v_\ell(0)$ simplifies to become $v_\ell(0) = \frac{1}{2}(1 + \delta_{0,\ell})$; consequently, $\{\widehat{C}_T(r, \ell\}$ is pivotal (does not depend on any nuisance parameters). In contrast, if the time series is (fourth order) 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 some fixed (single) ℓ, the test statistic can 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. However, in practice $v_\ell(0)$ is unknown. Therefore, Jentsch and Subba Rao [2015] use the stationary bootstrap, proposed in Politis and Romano [1994], to estimate $v_\ell(0)$ (actually they estimate $v_\ell(\omega_r)$). Note that in Jentsch and Subba Rao [2015] a more general statistic based on the full set$\{\widehat{C}_T(r, \ell); r = 1, \ldots, m, \ell = 1, \ldots, L\}$ (for multivariate time series) is proposed.

After having understood the asymptotic behavior of test statistics as in (5) under the null of second order stationarity (we have to assume fourth order stationarity to establish the asymptotic results above), Jentsch and Subba Rao [2015] assume that the time series 'evolves' slowly over time (a notion that was first introduced in Priestley [1965]) to understand how $\widehat{C}_T(r, \ell)$ behaves in the case of (second order) nonstationarity. 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., $|\mathrm{cov}(X_{t,T}, X_{t+h,T}) - \kappa_{h;\frac{t}{T}}| \leq T^{-1}\rho_h$, 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 $f(\omega) = \int_0^1 F_u(\omega)du$. Note that in the case of stationarity, F_u does not depend on u such that the right hand side of (6) simplifies to get $A(r, \ell) = 0$ for all $r \in \mathbb{Z}, r \neq 0$. This corresponds to $\widetilde{\mathbf{T}}_m$ converging to a chi-squared distribution under the null. Although all theoretical investigations under the alternative have been done for the broad, but still restrictive class of locally stationary processes, tests based on (sums or maxima of) $\{\widehat{C}_T(r, \ell); r = 1, \ldots, m, \ell = 1, \ldots, L\}$ are shown to have non-trivial power also for other types of nonstationarity as e.g. unit root processes or processes with structural breaks; compare Section 6.2 in Jentsch and Subba Rao [2015].

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 $\{(\boldsymbol{s}_j, Z(\boldsymbol{s}_j)); j = 1, \ldots, n\}$. We mention that we do not require that the locations $\boldsymbol{s}_j$ lie on a d-dimensional square centered at zero. The same procedure, described below, applies if the locations $\boldsymbol{s}_j$ are centered about another location $\boldsymbol{u}$; there is no need to centralize the locations. Furthermore, the locations need not lie on a d-dimensional square and a d-dimensional rectangle is sufficient. However, it is not possible to relax this assumption to an irregular domain. This is because on a rectangle domain, the Fourier transform of a constant function is zero at all but the zeroth frequency. This property is fundamental to the testing procedure. But on irregular domains this property will not necessarily hold. If in practice the data lies on an irregular domain the largest rectangular subset must be used when testing for stationarity. Suppose $\{Z(\boldsymbol{s}); \boldsymbol{s} \in \mathbb{R}^d\}$ is spatially (second order) stationary and denote $c(\boldsymbol{v}) = \text{cov}[Z(\boldsymbol{s}), Z(\boldsymbol{s} + \boldsymbol{v})]$. 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 $\text{var}[J_n(\boldsymbol{\Omega}_k)] = f(\boldsymbol{\Omega}_k) + O(\frac{1}{\lambda} + \frac{\lambda^d}{n})$, 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 $J_n(\boldsymbol{\Omega_k})/\sqrt{f(\boldsymbol{\Omega_k})}$ 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

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

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).*

The sampling results are derived 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 and under the null of second order stationarity, we show in Theorem 3.1 of Bandyopadhyay and Subba Rao [2016], that

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

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 $\boldsymbol{r}$. Moreover, from Bandyopadhyay and Subba Rao [2016], Section 3 (Theorem 3.3 treats the Gaussian case and the non-Gaussian case can be found at the bottom of Section 3), we have under the null (and fourth order stationarity) that

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

holds, 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

$$\begin{aligned} c_{a,\lambda} &= \frac{1}{2(2\pi)^d}\int_{\mathcal{D}} f^2(\boldsymbol{\Omega})\left(|g(\boldsymbol{\Omega})|^2+g(\boldsymbol{\Omega})\overline{g(-\boldsymbol{\Omega})}\right)d\boldsymbol{\Omega} \\ &+ \frac{1}{2(2\pi)^{2d}}\int_{\mathcal{D}^2} f_4(\boldsymbol{\Omega}_1,\boldsymbol{\Omega}_2,-\boldsymbol{\Omega}_2)g(\boldsymbol{\Omega}_1)\overline{g(\boldsymbol{\Omega}_2)}d\boldsymbol{\Omega}_1 d\boldsymbol{\Omega}_2, \end{aligned} \tag{9}$$

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

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

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 $J_n(\boldsymbol{\Omega})$ such that $f(\boldsymbol{\Omega})$ 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, we present an approach based on so-called orthogonal samples, as proposed by Subba Rao [2015b] and used in Bandyopadhyay and Subba Rao [2016], to estimate complicated variances. The expression for the variance, $c_{a,\lambda}$, given in (9), 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 $|J_n(\boldsymbol{\Omega}_{\boldsymbol{k}})|^2$ (see Bandyopadhyay et al. [2015], Lemma 7.5). However, in the case that the process is non-Gaussian this is not possible. The following remark describes 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.

Remark 2.2 (Using orthogonal samples for variance estimation) *Suppose that $\widehat{A}_D(\mathbf{X})$ is an estimator of A such that $\mathrm{E}[\widehat{A}_D(\mathbf{X})]\to A$ and $var[\sqrt{D}\widehat{A}_D(\mathbf{X})] = \nu$ holds, where $D = D(T,\lambda)$ is an appropriate scaling factor such that $var[\sqrt{D}\widehat{A}_D(\mathbf{X})] = O(1)$. Further, assume that there exists a non-empty set $\mathcal{B}'$ and a sample $\{\sqrt{D}\widehat{A}_D(\mathbf{X};j); j\in\mathcal{B}'\}$ (which is not necessarily* *real-valued) that satisfies*

(i) $\{\sqrt{D}\Re\widehat{A}_D(\mathbf{X};j); j \in \mathcal{B}'\}$, $\{\sqrt{D}\Im\widehat{A}_D(\mathbf{X};j); j \in \mathcal{B}'\}$ *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* $\mathrm{var}[\sqrt{D}\Re\widehat{A}_D(\mathbf{X};r)] = \nu + o(1)$ *and* $\mathrm{var}[\sqrt{D}\Im\widehat{A}_D(\mathbf{X};r)] = \nu + o(1)$.

Then we call $\{\sqrt{D}\widehat{A}_D(\mathbf{X};j); j \in \mathcal{B}'\}$ *an orthogonal sample associated with* $\widehat{A}_D(\mathbf{X})$*. Based on this, a strategy to estimate* ν *is to define*

$$\widehat{\nu} = \widehat{\sigma}^2(\{\sqrt{D}\widehat{A}_D(\mathbf{X};j); j \in \mathcal{B}'\}) = \frac{D}{2|\mathcal{B}'|}\sum_{j\in\mathcal{B}'}\left[(\Re\widehat{A}_D(\mathbf{X};j) - \bar{A})^2 + (\Im\widehat{A}_D(\mathbf{X};j) - \bar{A})^2\right], \quad (10)$$

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 we have joint asymptotic normality (and independence)* $\sqrt{D}[\widehat{A}_D(\mathbf{X}) - A, \{\Re\widehat{A}_D(\mathbf{X};j), \Im\widehat{A}_D(\mathbf{X};j); j \in \mathcal{B}'\}] \overset{\mathcal{D}}{\to} N(0, \nu I_{2|\mathcal{B}'|+1})$*, then we have*

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

where t_q *denotes the t-distribution with* q *degrees of freedom. Hence, if an orthogonal sample in the sense of above is available, this general method allows 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 the method of orthogonal samples. In the following, we describe how it is used in the spatial set-up. To implement a test for spatial stationarity, 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. This property allows us to use the orthogonal sample method described in Remark 2.2 to estimate the variance. 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 note that for each element in $\{\Re\widehat{A}_\lambda(g;\boldsymbol{r}), \Im\widehat{A}_\lambda(g;\boldsymbol{r}); \boldsymbol{r} \in \mathcal{S}\}$ the set $\sqrt{\lambda^d}\{\Re\widehat{A}_\lambda(g;\boldsymbol{r}), \Im\widehat{A}_\lambda(g;\boldsymbol{r}); \boldsymbol{r} \in \mathcal{S}'\}$ can be considered as its orthogonal sample (where we set $\mathcal{S}' = \mathcal{B}'$), since conditions (i) and (ii) in Remark 2.2 are satisfied. Thus we 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 (10). Using $\widehat{c}_{a,\lambda}(\mathcal{S}')$ and (8) we have

$$\lambda^{d/2}\frac{\Re\widehat{A}_{\lambda}(g;\boldsymbol{r})}{\sqrt{\widehat{c}_{a,\lambda}(\mathcal{S}')}} \overset{\mathcal{D}}{\to} \frac{Z_{1,\boldsymbol{r}}}{\sqrt{\frac{1}{2|\mathcal{S}'|-1}\chi^2_{2|\mathcal{S}'|-1}}} \sim t_{2|\mathcal{S}'|-1} \text{ and } \lambda^{d/2}\frac{\Im\widehat{A}_{\lambda}(g;\boldsymbol{r})}{\sqrt{\widehat{c}_{a,\lambda}(\mathcal{S}')}} \overset{\mathcal{D}}{\to} \frac{Z_{2,\boldsymbol{r}}}{\sqrt{\frac{1}{2|\mathcal{S}'|-1}\chi^2_{2|\mathcal{S}'|-1}}} \sim t_{2|\mathcal{S}'|-1}, \tag{11}$$

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 $\{Z_{1,\boldsymbol{r}}, Z_{2,\boldsymbol{r}}; \boldsymbol{r} \in \mathcal{S}\}$. 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 (11). 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 $\widehat{A}_{\lambda}(g;\boldsymbol{r}) \overset{\mathcal{P}}{\to} A(g;\boldsymbol{r})$ 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 $\boldsymbol{v}$.

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* $0 < c_1 \leq c_2 < \infty$ *(this does not change with* n*) we have* $c_1 n \leq n_t \leq c_2 n$.

(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 $\boldsymbol{\Omega}_{\boldsymbol{k}} = 2\pi(k_1/\lambda, \ldots, k_d/\lambda)$, 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

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

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

$$J(\boldsymbol{\Omega}_{\boldsymbol{k}_1}, \omega_{k_2}) = \frac{1}{\sqrt{2\pi T}} \sum_{t=1}^{T} J_t(\boldsymbol{\Omega}_{\boldsymbol{k}_1}) \exp(it\omega_{k_2}) = \frac{\lambda^{d/2}}{n} \sum_{j=1}^{n} J_{\boldsymbol{s}_j}(\omega_{k_2}) \exp(i\boldsymbol{s}_j' \boldsymbol{\Omega}_{\boldsymbol{k}_1}). \tag{12}$$

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

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

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

$$\beta_\eta(v_j) = \left\{\begin{array}{cc} C & |v_j|\le 1 \\ C|v_j|^{-\eta} & |v_j|>1 \end{array}\right. \tag{13}$$

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

$$\mathrm{cov}[Z_{t,\lambda,T}(\boldsymbol{s}), Z_{t+h,\lambda,T}(\boldsymbol{s}+\boldsymbol{v})] = \kappa_{h;\frac{t}{T}}\left(\boldsymbol{v};\frac{\boldsymbol{s}}{\lambda}\right) + O\left(\frac{\rho_h\beta_{2+\delta}(\boldsymbol{v})}{T}\right). \tag{14}$$

The function $\kappa_\cdot(\cdot)$ satisfies the Lipschitz conditions: (i) $\sup_{u,\boldsymbol{u}}|\kappa_{h;u}(\boldsymbol{v};\boldsymbol{u})| \le \rho_h\beta_{2+\delta}(\boldsymbol{v})$, (ii) $\sup_{\boldsymbol{u}}|\kappa_{h;u_1}(\boldsymbol{v};\boldsymbol{u}) - \kappa_{h;u_2}(\boldsymbol{v};\boldsymbol{u})| \le |u_1-u_2|\rho_h\beta_{2+\delta}(\boldsymbol{v})$ and (iii) $\sup_u|\kappa_{h;u}(\boldsymbol{v};\boldsymbol{u}_1) - \kappa_{h;u}(\boldsymbol{v};\boldsymbol{u}_2)| \le ||\boldsymbol{u}_1-\boldsymbol{u}_2||_2\rho_h\beta_{2+\delta}(\boldsymbol{v})$. 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

$$F_u(\boldsymbol{\Omega},\omega;\boldsymbol{u}) = \frac{1}{2\pi}\sum_{h\in\mathbb{Z}} e^{-ih\omega}\int_{\mathbb{R}^d}\kappa_{h;u}(\boldsymbol{v};\boldsymbol{u})e^{-i\boldsymbol{v}'\boldsymbol{\Omega}}d\boldsymbol{v}. \tag{15}$$

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

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

  The corresponding time-dependent spectral density is $F_{\frac{t}{T}}(\boldsymbol{\Omega}, \omega)$ (defined analogously to (15)).

- **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

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

  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* $\sup_{\boldsymbol{u},u} |\kappa_{h;u}(\boldsymbol{v};\boldsymbol{u}$ $\rho_h \beta_{2+\delta}(\boldsymbol{v})$ *with* $\beta_{2+\delta}(\boldsymbol{v})$ *and* $\{\rho_h\}$ *as defined in* (13)*. 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*

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

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

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

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

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

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

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

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, $\mathrm{cov}\left[J(\boldsymbol{\Omega}_{\boldsymbol{k}_1}, \omega_{k_2}), J(\boldsymbol{\Omega}_{\boldsymbol{k}_1+\boldsymbol{r}_1}, \omega_{k_2+r_2})\right]$ is approximately equal to the Fourier coefficient $b_{r_2}(\boldsymbol{\Omega}_{\boldsymbol{k}_1}, \omega_{k_2}; \boldsymbol{r}_1)$, where

$$b_{r_2}(\boldsymbol{\Omega}, \omega; \boldsymbol{r}_1) = \int_0^1 \exp(-i2\pi r_2 u) \int_{[-1/2,1/2]^d} \exp(-i2\pi \boldsymbol{r}_1' \boldsymbol{u}) F_u(\boldsymbol{\Omega}, \omega; \boldsymbol{u}) d\boldsymbol{u} du. \tag{16}$$

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 $g : \mathbb{R}^d \to \mathbb{R}$ and $h : [0, \pi] \to \mathbb{R}$. 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

$$D_{g,h,v}(\boldsymbol{r}_1, r_2) = \frac{1}{\pi} \int_0^\pi v(\omega) \left[\frac{h(\omega)}{(2\pi)^d} \int_{\mathbb{R}^d} g(\boldsymbol{\Omega}) b_{r_2}(\boldsymbol{\Omega}, \omega; \boldsymbol{r}_1) d\boldsymbol{\Omega} \right]^2 d\omega \tag{17}$$

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)

$$\begin{aligned}\widehat{a}_g(\omega_{k_2}; \boldsymbol{r}_1, r_2) &= \frac{1}{\lambda^d} \sum_{\boldsymbol{k}_1=-\boldsymbol{a}}^{\boldsymbol{a}} g(\boldsymbol{\Omega}_{\boldsymbol{k}_1}) J(\boldsymbol{\Omega}_{\boldsymbol{k}_1}, \omega_{k_2}) \overline{J(\boldsymbol{\Omega}_{\boldsymbol{k}_1+\boldsymbol{r}_1}, \omega_{k_2+r_2})} - N_T \\ &= \frac{1}{n^2} \sum_{\boldsymbol{k}_1=-\boldsymbol{a}}^{\boldsymbol{a}} g(\boldsymbol{\Omega}_{\boldsymbol{k}_1}) \sum_{\substack{j_1,j_2=1 \\ j_1\neq j_2}}^{n} J_{\boldsymbol{s}_{j_1}}(\omega_{k_2}) \overline{J_{\boldsymbol{s}_{j_2}}(\omega_{k_2+r_2})} \exp(i\boldsymbol{s}_{j_1}\boldsymbol{\Omega}_{\boldsymbol{k}_1} - i\boldsymbol{s}_{j_2}\boldsymbol{\Omega}_{\boldsymbol{k}_1+\boldsymbol{r}_1}), \quad (18)\end{aligned}$$

and $g : \mathbb{R}^d \to \mathbb{R}$ is a user chosen bounded Lipschitz continuous function (see Remark 2.1), a is such that $(a\lambda)^d/n^2 \to 0$, where the last line follows from (12) and

$$\begin{aligned}N_T &= \frac{1}{n^2} \sum_{\boldsymbol{k}_1=-\boldsymbol{a}}^{\boldsymbol{a}} g(\boldsymbol{\Omega}_{\boldsymbol{k}_1}) \sum_{j=1}^{n} J_{\boldsymbol{s}_j}(\omega_{k_2}) \overline{J_{\boldsymbol{s}_j}(\omega_{k_2+r_2})} \exp(-i\boldsymbol{s}_j\boldsymbol{\Omega}_{\boldsymbol{r}_1}) \quad (19) \\ &= \frac{1}{2\pi T} \sum_{\boldsymbol{k}_1=-a}^{a} g(\boldsymbol{\Omega}_{\boldsymbol{k}_1}) \sum_{t,\tau=1}^{T} e^{it\omega_{k_2} - i\tau\omega_{k_2+r_1}} \frac{1}{n_t n_\tau} \sum_{j=1}^{n} \delta_{t,j}\delta_{\tau,j} Z_t(\boldsymbol{s}_j) Z_\tau(\boldsymbol{s}_j) \exp(-i\boldsymbol{s}_j\boldsymbol{\Omega}_{\boldsymbol{r}_1}).\end{aligned}$$

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*

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

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*

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

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

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

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

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

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

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

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 (18) 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,* $|\frac{\partial^d f_h(\mathbf{\Omega})}{\partial \Omega_1 \ldots \partial \Omega_d}| \leq \rho_h \beta_{1+\delta}(\mathbf{\Omega})$ *(see the proof of Theorem 3.1, Subba Rao [2015a]). Then*

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

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,*

$$\begin{aligned} &\lambda^d \mathrm{cov}\left[\Re \widehat{a}_g(\omega_{k_2}, \boldsymbol{r}_1, r_2), \Re \widehat{a}_g(\omega_{k_4}, \boldsymbol{r}_3, r_4)\right] \\ &= I_{\boldsymbol{r}_1 = \boldsymbol{r}_3} I_{r_2 = r_4} \left[I_{k_2 = k_4} V_g(\omega_{k_2}; \mathbf{\Omega}_{\boldsymbol{r}_1}, \omega_{r_2}) + I_{k_4 = T - k_2 - r_2} V_{g,2}(\omega_{k_2}; \mathbf{\Omega}_{\boldsymbol{r}_1}, \omega_{r_2}) + O\left(\frac{1}{T}\right)\right] \\ &\qquad + O\left(\ell_{\lambda,a,n}\right), \end{aligned} \tag{20}$$

where

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

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

and $\int_{\mathcal{D}_{\boldsymbol{r}_1}} = \int_{2\pi\max(-a,-a-r_{1,1})/\lambda}^{2\pi\min(a,a-r_{1,1})/\lambda}\cdots\int_{2\pi\max(-a,-a-r_{1,d})/\lambda}^{2\pi\min(a,a-r_{1,d})/\lambda}$. *Note that* $\ell_{\lambda,a,n}$ *and* $\int_{\mathcal{D}}$ *are defined in Section 2.2. Exactly the same result as in* (20) *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 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

$$V_g(x_1;\boldsymbol{\Omega}_{\boldsymbol{r}_1},\omega_{r_2}) = V_g(x_2)+O\left(|x_1-x_2|+|\boldsymbol{\Omega}_{\boldsymbol{r}_1}|+|\omega_{r_2}|\right), \tag{21}$$

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

$$b_{g,r_2}(\omega;\boldsymbol{r}_1) = \langle g, b_{r_2}(\cdot,\omega;\boldsymbol{r}_1)\rangle = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} g(\boldsymbol{\Omega})b_{r_2}(\boldsymbol{\Omega},\omega;\boldsymbol{r}_1)d\boldsymbol{\Omega} \tag{22}$$

and $b_{r_2}(\cdot,\omega;\boldsymbol{r}_1)$ is defined in (16). 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

$$\widehat{A}_{g,h}(\boldsymbol{r}_1,r_2)=\frac{2}{T}\sum_{k=1}^{T/2}h(\omega_k)\widehat{a}_g(\omega_k;\boldsymbol{r}_1,r_2), \tag{23}$$

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 7.1. Summarizing Lemma 7.1, the variances are

$$\mathrm{var}\left(\sqrt{\frac{\lambda^d T}{2}}\Re\widehat{A}_{g,h}(\boldsymbol{r}_1,r_2)\right)=V_{g,h}+o(1) \text{ and } \mathrm{var}\left(\sqrt{\frac{\lambda^d T}{2}}\Im\widehat{A}_{g,h}(\boldsymbol{r}_1,r_2)\right)=V_{g,h}+o(1) \tag{24}$$

and

$$\sqrt{\frac{\lambda^d T}{2V_{g,h}}}\left(\left\{\Re\widehat{A}_{g,h}(\boldsymbol{r}_1,r_2),\Im\widehat{A}_{g,h}(\boldsymbol{r}_1,r_2);(\boldsymbol{r}_1,r_2)\in\widetilde{\mathcal{P}}\right\}\right)\overset{\mathcal{D}}{\to}\mathcal{N}(0,I_{2|\widetilde{\mathcal{P}}|}) \tag{25}$$

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

$$\begin{aligned}V_{g,h}=&\frac{1}{2\pi}\int_0^{\pi}|h(\omega)|^2V_g(\omega)d\omega\\&+\frac{2}{(2\pi)^{d+2}}\int_0^{\pi}\int_0^{\pi}\int_{\mathcal{D}^2}g(\boldsymbol{\Omega}_1)\overline{g(\boldsymbol{\Omega}_2)}h(\omega_1)\overline{h(\omega_2)}f_4(\boldsymbol{\Omega}_1,\omega_1,\boldsymbol{\Omega}_2,\omega_2,-\boldsymbol{\Omega}_2,-\omega_2)\\&\qquad d\boldsymbol{\Omega}_1d\boldsymbol{\Omega}_2d\omega_1d\omega_2.\end{aligned} \tag{26}$$

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 (24) and (25) 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 a suitable set $\mathcal{P}'$. We describe how to construct sets $\mathcal{P}$ and $\mathcal{P}'$ below.

Definition 4.1 (The sets $\mathcal{P}$ and $\mathcal{P}'$) *Let $\mathcal{S}$ and $\mathcal{S}'$ be similar to the sets defined in Section 2.2 ($\mathcal{S}$ and $\mathcal{S}'$ contain vectors in $\mathbb{Z}^d$). We define the temporal sets $\mathcal{T} = \{1, \ldots, m\}$ and $\mathcal{T}' = \{m + B_1, \ldots, m + B_2\}$ where $B_1 < B_2$ ($\mathcal{T}$ and $\mathcal{T}'$ contain vectors in $\mathbb{Z}$). Since $\widehat{A}_{g,h}(\boldsymbol{r}_1, r_2)$ is a function of $\boldsymbol{r}_1$ and r_2 we define the sets $\mathcal{P} = \mathcal{S} \times \mathcal{T}$ and $\mathcal{P}' = \mathcal{S}' \times \mathcal{T}'$. 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].

From the definition of $\mathcal{P}$ and $\mathcal{P}'$ given above we see that $\{\widehat{A}_{g,h}(\boldsymbol{r}_1, r_2); (\boldsymbol{r}_1, r_2); (\boldsymbol{r}_1, r_2) \in \mathcal{P}'\}$ satisfies the conditions of an orthogonal sample given in Remark 2.2. Thus we estimate $V_{g,h}$ with

$$\widehat{V}_{g,h}(\mathcal{P}') = \widehat{\sigma}^2 \left(\sqrt{\frac{T\lambda^d}{2}} \widehat{A}_{g,h}(\boldsymbol{r}_1, r_2); (\boldsymbol{r}_1, r_2) \in \mathcal{P}' \right). \tag{27}$$

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

$$\mathbf{T}_{1,g,h}(\mathcal{P}, \mathcal{P}') = \frac{\lambda^d T}{2} \sum_{(\boldsymbol{r}_1, r_2) \in \mathcal{P}} \frac{|\widehat{A}_{g,h}(\boldsymbol{r}_1, r_2)|^2}{\widehat{V}_{g,h}(\mathcal{P}')} \text{ and } \mathbf{M}_{1,g,h}(\mathcal{P}, \mathcal{P}') = \frac{\lambda^d T}{2} \max_{(\boldsymbol{r}_1, r_2) \in \mathcal{P}} \frac{|\widehat{A}_{g,h}(\boldsymbol{r}_1, r_2)|^2}{\widehat{V}_{g,h}(\mathcal{P}')}. \tag{28}$$

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 device defined in Section 3 we have $\widehat{A}_{g,h}(\boldsymbol{r}_1, r_2) \stackrel{\mathcal{P}}{\to} A_{g,h}(\boldsymbol{r}_1, r_2)$ as $T \to \infty$, $\lambda^d/n \to 0$, $\lambda \to \infty$ and $n \to \infty$, where*

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

and $b_{r_2}(\boldsymbol{\Omega}, \omega; \boldsymbol{r}_1)$ is defined in (16).

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 (21), 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 (26) 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

$$\widehat{V}_g(\omega_k;\mathcal{P}')=\widehat{\sigma}^2\left(\{\lambda^{d/2}\widehat{a}_g(\omega_{k+i};\boldsymbol{r}_1,r_2);(\boldsymbol{r}_1,r_2)\in\mathcal{P}',|i|\leq M\}\right), \tag{29}$$

where $\widehat{\sigma}^2(\cdot)$ is defined in (10). 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

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

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 (25) we have

$$\sqrt{\frac{\lambda^dT}{2V_{g,V^{-1/2}}}}\left(\left\{\Re\widehat{A}_{g,\widehat{V}^{-1/2}}(\boldsymbol{r}_1,r_2),\Im\widehat{A}_{g,\widehat{V}^{-1/2}}(\boldsymbol{r}_1,r_2);(\boldsymbol{r}_1,r_2)\in\mathcal{P}\cap\mathcal{P}'\right\}\right)\overset{\mathcal{D}}{\to}\mathcal{N}(0,I_{2|\widetilde{\mathcal{P}}|}). \tag{30}$$

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 (34) 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 (28).

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

$$\widehat{B}_{g,h;H}(\omega_{jH};\boldsymbol{r}_1,r_2) = \frac{1}{H}\sum_{k=1}^{H} h(\omega_{jH+k})\widehat{a}_g(\omega_{jH+k};\boldsymbol{r}_1,r_2), \text{ for } 0 \leq j < T/(2H), \tag{31}$$

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 7.2 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

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

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

$$W_{g,h}(\omega_{jH}) = \frac{T}{2H\pi}\int_{\omega_{jH}}^{\omega_{(j+1)H}} |h(\omega)|^2 V_g(\omega;0,0)d\omega + O\left(\frac{H}{T}\right). \tag{32}$$

Furthermore, the correlation *between* each of the blocks $W_{g,h}(\omega_{j_1H})$ and $W_{g,h}(\omega_{j_2H})$ 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 (25) we have

$$\sqrt{\frac{T}{2H}}\left\{\left[H\lambda^d\widehat{D}_{g,h,W;H}(\boldsymbol{r}_1,r_2)-1\right];(\boldsymbol{r}_1,r_2)\in\widetilde{\mathcal{P}}\right\} \xrightarrow{\mathcal{D}} \mathcal{N}(0,I_{|\widetilde{\mathcal{P}}|}), \tag{33}$$

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. Since $\{\widehat{B}_{g,h}(\omega_{jH};\boldsymbol{r}_1,r_2);(\boldsymbol{r}_1,r_2)\in\mathcal{P}'\}$ satisfies the conditions of an orthogonal sample (see Remark 2.2), we estimate $W_{g,h}(\omega_{j,H})$ with

$$\widehat{W}_{g,h}(\omega_{jH};\mathcal{P}') = \widehat{\sigma}^2\left(\sqrt{\lambda^d H}\widehat{B}_{g,h}(\omega_{jH};\boldsymbol{r}_1,r_2);(\boldsymbol{r}_1,r_2)\in\mathcal{P}'\right), \tag{34}$$

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,

$$\widehat{D}_{g,h,\widehat{W};H}(\boldsymbol{r}_1,r_2) = \frac{2H}{T}\sum_{j=0}^{T/2H-1}\frac{\left|\widehat{B}_{g,h;H}(\omega_{jH};\boldsymbol{r}_1,r_2)\right|^2}{2\widehat{W}_{g,h}(\omega_{jH})}. \tag{35}$$

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 (33), 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 device 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 (17).*

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

$$\widehat{B}_{g,\widehat{V}^{-1/2};H}(\omega_{jH};\boldsymbol{r}_1,r_2) = \frac{1}{H}\sum_{k=1}^{H}\frac{\widehat{a}_g(\omega_{jH+k};\boldsymbol{r}_1,r_2)}{\sqrt{\widehat{V}_g(\omega_{jH+k};\mathcal{P}')}}. \tag{36}$$

By using (32) we see its real and imaginary parts have limiting variance

$$W_{g,V^{-1/2}}(\omega_{jH}) \;=\; 1+O\left(\frac{H}{T}\right). \tag{37}$$

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

$$\widehat{D}_{g,\widehat{V}^{-1/2},1;H}(\boldsymbol{r}_1,r_2) = \frac{2H}{T}\sum_{j=0}^{T/2H-1}\frac{|\widehat{B}_{g,\widehat{V}^{-1/2};H}(\omega_{jH};\boldsymbol{r}_1,r_2)|^2}{2}. \tag{38}$$

Studying (38), 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 8 we compare the effect this has on the finite sample properties of the test statistic. Using (33) we have

$$\sqrt{\frac{T}{2H}}\left(H\lambda^d\widehat{D}_{g,\widehat{V}^{-1/2},1;H}(\boldsymbol{r}_1,r_2)-1\right)\xrightarrow{\mathcal{D}}\mathcal{N}(0,1) \tag{39}$$

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 8 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' distribution approximations

In the previous section we derive the asymptotic distribution of the three test statistics. However, these results are asymptotic and do not take into account the extra uncertainty caused by the variance estimation. In this section we "correct" for this additional uncertainty to obtain better approximations of distributions of the test statistics under the null.

We recall from Remark 2.2 and Section 2.2, that the 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 (11) we estimate it with the t-distribution, which can be considered as the 'asymptotic finite sample distribution' of this ratio. Using a similar argument we can obtain the 'asymptotic finite sample distribution' of $|\widehat{A}_{g,\widehat{V}^{-1/2}}(\boldsymbol{r}_1,r_2)|^2/\widehat{V}_g(\mathcal{P}')$, $\widehat{D}_{g,\widehat{V}^{-1/2},\widehat{W};H}(\boldsymbol{r}_1,r_2)$ and $\widehat{D}_{g,\widehat{V}^{-1/2},1;H}(\boldsymbol{r}_1,r_2)$. Using these results we can obtain finite sample asymptotic approximations to the distribution of the test statistics $\mathbf{T}_{1,g,\widehat{V}^{-1/2}}(\mathcal{P},\mathcal{P}')$, $\mathbf{T}_{2,g,\widehat{V}^{-1/2}}(\mathcal{P},\mathcal{P}')$ and $\mathbf{M}_{2,g,\widehat{V}^{-1/2}}(\mathcal{P},\mathcal{P}')$ under the null of stationarity. The details can be found in Section 6 and these approximations are used in the simulations in Section 8 (both these sections are in the supplementary material).

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. Many of the results in this section depend on the auxillary results stated in Section 7 of the supplementary material. For $\widehat{a}_g(\cdot;\boldsymbol{r}_1,r_2)$ as defined in (18) we observe:

- **Spatial stationarity, but not necessarily temporal stationarity**
  If $\boldsymbol{r}_1 \neq \mathbf{0}$ 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 $\widehat{a}_g(\omega;\mathbf{0},r_2) = o_p(1)$. On the other hand, if the process is temporally nonstationary then the latter is not necessarily true. Therefore the test set is $\mathcal{P} = \{\mathbf{0}\} \times \mathcal{T}$.

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

$$\begin{aligned}&\lambda^d \mathrm{cov}\left[\Re\widehat{a}_g(\omega_{k_2};\boldsymbol{r}_1,r_2),\Re\widehat{a}_g(\omega_{k_4};\boldsymbol{r}_3,r_4)\right]\\ =\;& \frac{1}{2}\Re\left[b^{(1)}_{r_2,r_4,k_2,k_4}(\omega_{k_2},\omega_{k_4};\boldsymbol{r}_1,\boldsymbol{r}_3)+b^{(2)}_{r_2,r_4,k_2,k_4}(\omega_{k_2},\omega_{k_4};\boldsymbol{r}_1,\boldsymbol{r}_3)\right]+O\left(\frac{1}{T}+\frac{1}{\lambda}+\ell_{\lambda,a,n}\right),\end{aligned} \quad (40)$$

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

and

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

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 7.2.

The above expressions are cumbersome, however, under one-way stationarity simplifications can be made. We recall from the definition in (16) that

$$
b_{r_2}(\boldsymbol{\Omega},\omega;\boldsymbol{r}_1)=\begin{cases}0 & \boldsymbol{r}_1\neq 0 \text{ and spatial stationarity}\\ 0 & r_2\neq 0 \text{ and temporal stationarity}\end{cases}\tag{41}
$$

We use these results to simplify the expressions for $\mathrm{cov}[\widehat{a}_g(\omega_{k_2};\boldsymbol{r}_1,r_2),\widehat{a}_g(\omega_{k_4};\boldsymbol{r}_3,r_4)]$ 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 (40) and (41), 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

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

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 $\lambda^d \text{cov}[\Re\widehat{a}_g(\omega_{k_2};\boldsymbol{r}_1,$ 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

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

Furthermore, defining $\widehat{A}_{g,h}(\boldsymbol{r}_1,0)$ as in (23) 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,

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

and $b_{0,4}$ is defined in Section 7.2.

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 (40) and (41) 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 (23) 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 $\mathcal{T}' \subset \mathbb{Z}^+$).

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 non-stationary 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.

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. The authors also wish to thank the Editor and Guest Editor and two referees for making several useful suggestions which improved many parts of the paper.

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.

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)*, 77:893–922, 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. Statistical inference for spatial statistics defined in the Fourier domain. *Submitted*, 2015a.

S. Subba Rao. Orthogonal samples for estimators 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.