# SPECIFICATION FOR LATTICE PROCESSES

JAVIER HIDALGO

Abstract. We describe tests for the correct specification of a model when data is observed in a lattice. We then extend previous work when the data is collected in the real line. As it happens with the latter type of data, the asymptotic distribution of the test is not pivotal and it depends on the model under consideration. On consequence is that its critical values are difficult, if at all possible, to obtain. So, to overcome the problem of its implementation, we propose to employ a martingale transformation, showing its validity in our context.

**JEL Classification:** C21, C23.

## 1. INTRODUCTION

It is agreed that one of the main purposes and aims when working with time series or spatial data is to obtain a correct description of its dynamic structure and more specifically of its autocovariance or covariogram. Its importance is well documented and rely on the fact that those functions play a key role into obtaining good and accurate prediction/extrapolation and/or interpolation (kriging) in the case of spatial data. Given a parametric family of models, say an ARMA model, there is a huge literature devoted to the estimation of the covariogram. However, before we attempt to estimate a specific parametric model, it is important to check its overall validity. The aim of the paper is then to describe tests for the correct specification (or model selection) of the dynamic structure with time series and/or spatial stationary processes $\{x(t)\}_{t\in\mathbb{Z}}$ defined on a $d$-dimensional lattice. In the paper, we shall specifically consider data for which $d \leq 3$, although we ought to mention that extensions to higher index lattice processes can be adapted under suitable modifications. The motivation to focus on the case $d \leq 3$ lies in the fact that the most often type of data available in economics is when $d = 2$, say with agricultural or environmental data, whereas an important example when $d = 3$ is the spatial-temporal data sets, that is data collected in a lattice during a number of periods.

More specifically, we are interested to check whether the covariogram given by $\{\gamma(s)\}_{s\in\mathbb{Z}^d}$, $\gamma(s) = Cov(x(t), x(t+s))$, follows a particular parametric family, say $\{\gamma(s)\}_{s\in\mathbb{Z}^d} = \{\gamma(s;\theta)\}_{s\in\mathbb{Z}^d}$, where $\theta$ is a finite dimensional set of parameters belonging to some (compact) parameter space. Equivalently, to solve the crucial problem of modelling the spatial dependence via the covariogram can be done through its spectral density function, denoted $f(\lambda)$. This is the case after observing that for any stationary spatial lattice process $\{x(t)\}_{t\in\mathbb{Z}^d}$, the spectral density function, $f(\lambda)$, and the covariogram, $\gamma(s)$, are related through the expression

$$\begin{matrix} \gamma(s) \\ \gamma(s;\theta) \end{matrix} = \left\{ \begin{matrix} \int_{\Pi^d} e^{-is\cdot\lambda} f(\lambda)\, d\lambda \\ \int_{\Pi^d} e^{-is\cdot\lambda} f(\lambda;\theta)\, d\lambda \end{matrix} \right. \quad ; \quad s = 0, \pm 1, \pm 2, \ldots,$$

---

*Date*: 26 November 2007.

*Key words and phrases.* Specification test. Spatial processes. Spectral domain. CUSUM.

where "$s \cdot \lambda$" means the inner product of the $d$-dimensional vectors $s$ and $\lambda$ and $\Pi = (-\pi, \pi]$. For the moment, it will be convenient to work in a general $d$-dimensional setting. Herewith, any element $a$ that belongs to $\mathbb{Z}^d$ (or $\Pi^d$), the $d$-fold Cartesian product of the set $\mathbb{Z}$ (or $\Pi$), is referred to as a multi-index of dimension $d$. Also, we shall write, say, $a = (a[1], ..., a[d])$ with the square brackets used to denote the components of $a$.

We will assume all throughout the paper that the (spatial) process $\{x(t)\}_{t\in\mathbb{Z}^d}$ can be represented by the multilateral model

$$x(t) - \mu = \sum_{j\in\mathbb{Z}^d} \psi(j)\,\varepsilon(t-j), \quad \sum_{j\in\mathbb{Z}^d} \psi^2(j) < \infty \quad \psi(0) = 1, \tag{1.1}$$

for some sequence $\{\varepsilon(t)\}_{t\in\mathbb{Z}^d}$ satisfying $\mathbb{E}(\varepsilon(t)) = 0$ and $\mathbb{E}(\varepsilon(0)\,\varepsilon(t)) = \sigma_\varepsilon^2$ if $t = 0$; and $= 0$ for all $t \neq 0$. Under (1.1), the spectral density function of $\{x(t)\}_{t\in\mathbb{Z}^d}$ can be factorized as

$$f(\lambda) = \frac{\sigma_\varepsilon^2}{(2\pi)^d} |\Psi(\lambda)|^2$$

with

$$\Psi(\lambda) = \sum_{j\in\mathbb{Z}^d} \psi(j) \exp(-ij\cdot\lambda), \tag{1.2}$$

so that $\Psi(\lambda)$ summarizes the covariogram structure of $\{x(t)\}_{t\in\mathbb{Z}^d}$.

Much of the existing time series and spatial statistics literature is concerned with parametric modelling, estimation and testing in that $\Psi(\lambda)$ is assumed to belong to a specific parametric family

$$\mathcal{H} = \{\Psi_\theta(\lambda) : \theta \in \Theta\}, \tag{1.3}$$

where $\Theta \subset \mathbb{R}^p$ is a proper compact parameter set. See for instance Cressie (1992) for alternative methods of estimation. That is, it is assumed that $\Psi(\lambda) = \Psi_{\theta_0}(\lambda)$ for some "true" value of the parameter $\theta_0 \in \Theta$. So, the null hypothesis becomes

$$H_0 : \forall\lambda \in [-\pi,\pi]^d \text{ and for some } \theta_0 \in \Theta,\ \ |\Psi(\lambda)|^2 = |\Psi_{\theta_0}(\lambda)|^2, \tag{1.4}$$

where $\Theta \subset \mathbb{R}^p$ is a compact set. The alternative hypothesis is the negation of $H_0$.

When $d = 1$, the problem of testing a specific dependence structure of the data is very exhaustive and prominent. Different types of tests have been formulated either using the spectral density or the autocorrelation functions. Regarding the former, we can cite among others, the pioneer work by Grenander and Rosenblatt (1957) to test for the null hypothesis of white noise dependence. A classical test using the autocorrelation function is the Box and Pierce (1970) statistic. For a latter reference, see Delgado, Hidalgo and Velasco (2005) and references therein. In the paper, we have chosen to employ frequency domain techniques or to base the test in terms of the spectral density function. However we describe or present later in Section 5, the relationship among different tests using the "time domain" approach based on the covariance/variogram structure of the data and tests based on the "spectral domain" approach that we employ in the paper.

Our tests falls into the category of goodness-of-fit tests as we do not specify any particular alternative model or family, although we will see in Section 5 how the tests described in Sections 2 and 3 can be easily modified in such a way that they can target efficiently one (or more) particular set of alternative(s) by employing the so-called directional tests. The tests are based on a direct comparison between two estimates of the spectral density function in a way similar to the well known Hausman-Durbin-Wu tests. That is, they rely on the comparison of two estimates: one which is only consistent under the null, whereas the second (less efficient) estimator is consistent under the maintained hypothesis. Although the literature when

$d > 1$ is not very vast and exhaustive, some work has already been done, see for instance Diblasi and Bowman (2001) or Crujeiras et al. (2006). However, our work differs from theirs in that contrary to Diblasi and Bowman (2001) we do test for general specifications and that contrary to Crujeiras et al. (2006) our test does not involve any bandwidth or smoothing parameter. In fact, the latter approach uses the distance between a smooth estimator of the spectral density function and its parametric estimator under $H_0$. This approach provides asymptotically distribution free tests under a suitable behaviour of the smoothing parameter as the sample size increases, see also Hong (1996) or Paparoditis (2000). However, the latter approach seems an artifact when testing for a particular parametric family and the final outcome of all these tests may depend on the arbitrary choice of the tuning/smoothing parameters for which no relevant theory is available. That is, there are not rules available on how to choose the bandwidth parameter. In fact, we might face the strange situation that with the same data set two different practitioners might conclude differently or that if one chooses to optimize the size of the test, that choice can lead to tests which have very poor power and viceversa. The latter is clearly not very appealing from both theoretical or applied point of view. So, in this context, one of our main motivation is to extend goodness-of-fit tests examined and described when $d = 1$ to $d \geq 1$, where we do not require the choice of any bandwidth. For that purpose, we rely on the periodogram which although it is not a consistent estimator of $f(\lambda)$, its integral is consistent for the spectral distribution function as the latter is the most natural smoothing function.

The remainder of the paper is organized as follows. In the next section, we present the test and examine its asymptotic properties when the true value of the parameter $\theta_0$ is known, whereas Section 3 extends these results to more realistic situations where we need to estimate the parameters of the model. Because, the asymptotic distribution of the test in the latter scenario is not pivotal and model dependent, Section 4 describes the martingale transformation in the spirit of Brown, Durbin and Evans (1975) to obtain test statistics whose asymptotic distribution is pivotal. In fact, we show that the asymptotic distribution, after the transformation, of the tests are functionals of a Brownian sheet. The relationship between different tests in the time domain and those examined in previous sections is described and presented in Section 5. Section 6 gives the proof of a series of lemmas employed in the proof of our main results in Section 7.

## 2. TESTS WHEN THE PARAMETERS ARE KNOWN

This section discusses and examines how we can test the null hypothesis given in (1.4). That is,

$$H_0 : f(\lambda) = \frac{\sigma_\varepsilon^2}{(2\pi)^d} |\Psi_{\theta_0}(\lambda)|^2 \quad \forall \lambda \in \widetilde{\Pi}^d \text{ for some value } \theta_0,$$

when the "true" value of $\theta_0$ is known, and $\widetilde{\Pi}^d$ denotes $[0,\pi] \times [-\pi,\pi]^{d-1}$, that is $\lambda \in \widetilde{\Pi}^d$ if $\lambda[1] \in [0,\pi]$ and $\lambda[\ell] \in [-\pi,\pi]$ for all $\ell = 2, ..., d$. Before we introduce and describe the test, we first observe that the null hypothesis $H_0$ can be stated alternatively as

$$H_0 : \frac{G_{\theta_0}(\lambda)}{G_{\theta_0}(\pi)} = \prod_{\ell=1}^{d} \frac{\lambda[\ell]}{\pi} \quad \text{for all } \lambda \in [0,\pi]^d, \tag{2.1}$$

where

$$G_\theta(\lambda) = 2\int_{-\lambda}^{\lambda} \frac{f(\omega)}{|\Psi_\theta(\omega)|^2} d\omega$$

with

$$\int_{\mu}^{\lambda} = \int_{(\mu[1]\wedge 0)}^{\lambda[1]} \int_{\mu[2]}^{\lambda[2]} \dots \int_{\mu[d]}^{\lambda[d]} . \tag{2.2}$$

Under $H_0$, $G_{\theta_0}(\lambda)$ is the spectral distribution function of the lattice process $\{\varepsilon(t)\}_{t\in\mathbb{Z}}$ and $G_{\theta_0}(\pi) = \sigma_\varepsilon^2$. Notice that by symmetry of $f(\lambda)$ it does not matter which coordinate we take to belong only to $[0,\pi]$. Its consequence is that it would not affect the value of $G_\theta(\lambda)$ and so the value of the test given below.

Given a record $\{x(t)\}_{t=1}^n$ and denoting henceforth $N = \prod_{\ell=1}^d n[\ell]$, a natural estimator of $G_{\theta_0}(\lambda)$ is given by

$$\widetilde{G}_{\theta,N}(\lambda) = 2\frac{1}{N} \sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]} \frac{I_x(\lambda_j)}{|\Psi_\theta(\lambda_j)|^2},$$

where $I_v(\lambda)$ is the *periodogram* of a generic sequence $\{v(t)\}_{t=1}^n$ defined as

$$I_v(\lambda) = \frac{1}{N}\left|\sum_{t=1}^n v(t)\, e^{-it\cdot\lambda}\right|^2$$

and similarly to the definition of $\int_\mu^\lambda$, we are employing henceforth the notation

$$\sum_{j=[\tilde{n}\lambda/\pi]}^{[\tilde{n}\mu/\pi]} = \sum_{j[1]=[\tilde{n}[1]\lambda[1]/\pi]_+}^{[\tilde{n}[1]\mu[1]/\pi]} \sum_{j[2]=[\tilde{n}[2]\lambda[2]/\pi]}^{[\tilde{n}[2]\mu[2]/\pi]} \dots \sum_{j[d]=[\tilde{n}[d]\lambda[d]/\pi]}^{[\tilde{n}[d]\mu[d]/\pi]} - \sum_{j=0}, \tag{2.3}$$

where $[q]_+ = \max\{|q|, 0\}$. Also we have abbreviated $[n[\ell]/2]$ by $\tilde{n}[\ell]$. As usual we have excluded the frequency $\lambda_j = 0$ from the sum $\sum_{j=[\tilde{n}\lambda/\pi]}^{[\tilde{n}\mu/\pi]}$, so that we can take $Ex(t) = 0$ or that $x(t)$ has been centered around its sample mean. It often the case that in applications, to make use of the *fast Fourier transform*, the periodogram is evaluated at the Fourier frequencies, that is $\lambda_k = \left(\lambda_{k[1]}, \dots, \lambda_{k[d]}\right)'$, where

$$\begin{aligned} \lambda_{k[1]} &= \frac{2\pi k[1]}{n[1]}; \quad k[1] = 0, 1, \dots, \tilde{n}[1]; \\ \lambda_{k[\ell]} &= \frac{2\pi k[\ell]}{n[\ell]}; \quad k[\ell] = 0, \pm 1, \dots, \pm\tilde{n}[\ell], \quad \ell = 2, \dots, d. \end{aligned}$$

Unfortunately, as noted by Guyon (1982), due to nonnegligible end effects, the bias of the periodogram does not converge to zero fast enough when $d > 1$, so that it would have unwanted consequences. One of theses cases is for the Whittle estimator of the parameters $\vartheta$, see Guyon (1982), which it does not have the standard asymptotic properties as when $d = 1$. Because of that, in the paper, we shall employ the taper periodogram defined as

$$I_v^T(\lambda_j) = \left|w_v^T(\lambda_j)\right|^2,$$

where

$$w_v^T(\lambda_j) = \frac{1}{\left(\sum_{t=1}^n h^2(t)\right)^{1/2}} \sum_{t=1}^n h(t)\, v(t)\, e^{it\cdot\lambda_j}$$

is the taper discrete Fourier transform of a generic sequence $\{v(t)\}_{t=1}^n$. Tapering is primary a technique employed to reduce the bias of the "standard" periodogram $I_v(\lambda)$. Notice that when $h(t) = 1$, we have that $w_v^T(\lambda_j)$ becomes the standard discrete Fourier transform *(DFT)*. It is worth mentioning that alternative procedures to tapering to alleviate the bias problem have been proposed. One of these proposals was due to Guyon (1982), who replaced the periodogram by

$$I_v^*(\lambda_k) = \frac{1}{(2\pi)^2} \sum_{h \in \mathcal{D}} \widehat{\gamma}_v^*(h) e^{-ih \cdot \lambda_k},$$

where $\widehat{\gamma}_v^*(h) = \frac{1}{N-|h|} \sum_{t(h)} v(t) v(t+h)$ and $\mathcal{D} = \{h : -n[\ell] < h[\ell] < n[\ell];\ \ell = 1, ..., d\}$. Notice that the standard periodogram $I_v(\lambda_k)$ replaces $\widehat{\gamma}_v^*(h)$ by $\widehat{\gamma}_v(h) = \frac{1}{N} \sum_{t(h)} v(t) v(t+h)$. However, Dahlhaus and Künsch (1987) have criticized the use of $I_v^*(\lambda_k)$ on the grounds that when employed to estimate the parameters of the model via a Whittle estimator, see (3.1) below, the estimator loses its minimum distance interpretability and that the objective function possesses several local maxima. The latter implies that to obtain the maximum of the Whittle function becomes more strenuous. Another possibility is that described by Robinson and Vidal-Sanz (2006). The latter proposal will be helpful when $d \geq 4$. However as we consider explicitly only the most common scenario $d \leq 3$, all we need for our results to follow is the taper periodogram $I_v^T(\lambda_k)$.

The benefits of tapering can be seen following the properties of the *cosine-bell (or Hanning)* taper, which is defined as

$$h(t) = \frac{1}{2^d} \prod_{\ell=1}^{d} h_\ell(t[\ell]); \qquad h_\ell(t[\ell]) = \left(1 - \cos\left(\frac{2\pi t[\ell]}{n[\ell]}\right)\right). \tag{2.4}$$

Indeed, denoting the taper Dirichlet kernel by

$$D_\ell^T(\mu[\ell]) = \sum_{t[\ell]=1}^{n[\ell]} h_\ell(t[\ell]) e^{it[\ell]\mu[\ell]},$$

we have that

$$\sup_{n[\ell], \lambda[\ell] > 0} \left| D_\ell^T(\lambda[\ell]) \right| = O\left(\min\left\{n[\ell], n[\ell]^{-2} |\lambda[\ell]|^{-3}\right\}\right). \tag{2.5}$$

The latter will improve the properties with respect to the standard periodogram as the bias is of smaller order of magnitude than with the standard periodogram. Observe that its relation with the standard Dirichlet kernel, $D_\ell(\lambda[\ell]) = \sum_{t[\ell]=1}^{n[\ell]} e^{it[\ell]\lambda[\ell]}$, is

$$D_\ell^T\left(\lambda_{j[\ell]}\right) = \frac{1}{6^{1/2}} \left\{-D_\ell\left(\lambda_{j[\ell]-1}\right) + 2D_\ell\left(\lambda_{j[\ell]}\right) - D_\ell\left(\lambda_{j[\ell]+1}\right)\right\}. \tag{2.6}$$

It is worth observing that the standard *DFT* and the *cosine-bell* taper *DFT* are related by the equality

$$w_x^T(\lambda_j) = \frac{1}{6^{1/2}} \left[-w_x(\lambda_{j-1}) + 2w_x(\lambda_j) - w_x(\lambda_{j+1})\right]. \tag{2.7}$$

In the paper we shall explicitly consider the *cosine-bell*, although the same results follow employing other taper functions such as *Parzen or Kolmogorov's* tapers.

The formulation of $H_0$ given in (2.1) suggests to use the Bartlett's $T_p - process$ as a basis for testing $H_0$. The $T_p - process$ is defined as

$$\alpha_{\theta,N}(\lambda) = 2^{-1/2} N^{1/2} \left[\frac{G_{\theta,N}(\lambda)}{G_{\theta,N}(\pi)} - \prod_{\ell=1}^{d} \left(\frac{\lambda[\ell]}{\pi}\right)\right], \quad \lambda \in [0, \pi]^d, \tag{2.8}$$

where

$$G_{\theta,N}(\lambda) = 2\frac{1}{N} \sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]} \frac{I_x^T(\lambda_j)}{|\Psi_\theta(\lambda_j)|^2}. \tag{2.9}$$

Notice that because we have excluded the frequency $j=0$ from the definition of $\sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]}$, it is easy to show that a linear transformation of the data does not change the value of $\alpha_{\theta,N}$ and therefore it can be assumed that the mean of $x(t)$ is zero and the variance of $\varepsilon(t)$ is the unity. So, without loss of generality, we can assume that in (1.1) $\mu=0$ and the true value of $Var(\varepsilon(t))=1$. It is worth mentioning that similarly we could have employed the $U_p - process$. The latter is defined as

$$U_{\theta,N}(\lambda)=\frac{1}{N}\sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]}\left\{I_x^T(\lambda_j)-\sigma_\varepsilon^2\left|\Psi_\theta(\lambda_j)\right|^2\right\},$$

which is the route followed by Grenander and Rosenblatt (1957). One motivation to employ $\alpha_{\theta,N}(\lambda)$ instead of $U_{\theta,N}(\lambda)$ is that the latter statistic is not invariant to the variance of $\{\varepsilon(t)\}_{t\in\mathbb{Z}^d}$ as is the former statistic $\alpha_{\theta,N}(\lambda)$ in (2.8).

One rational of the statistic $\alpha_{\theta,N}(\lambda)$ follows from the observation (see Lemma 4) that under $H_0$, we have that

$$\max_{(-\tilde{n}\le j\le\tilde{n})\wedge(j\neq 0)}\mathbb{E}\left|\frac{I_x^T(\lambda_j)}{\left|\Psi_{\theta_0}(\lambda_j)\right|^2}-I_\varepsilon^T(\lambda_j)\right|=o(1),$$

where "$a\le b$" means that $a[\ell]\le b[\ell]$ for all $\ell=1,...,d$ and

$$I_\varepsilon^T(\lambda_j)=\frac{1}{\sum_{t=1}^n h^2(t)}\left|\sum_{t=1}^n h(t)\,\varepsilon(t)\,e^{it\cdot\lambda_j}\right|^2.$$

Also, observe that $0<j[1]\le\tilde{n}[1]$ whereas $-\tilde{n}[\ell]<j[\ell]\le\tilde{n}[\ell]$ for $\ell=2,...,d$.

Thus, we can expect that $\alpha_{\theta,N}$ will be asymptotically equivalent to Bartlett's $U_p - process$ for $\left\{\varepsilon(t)/\sigma_\varepsilon^2\right\}_{t\in\mathbb{Z}^d}$,

$$\alpha_N^0(\lambda)=2^{-1/2}N^{1/2}\left[\frac{G_N^0(\lambda)}{G_N^0(\pi)}-\prod_{\ell=1}^d\left(\frac{\lambda[\ell]}{\pi}\right)\right],\quad\lambda\in[0,\pi]^d, \tag{2.10}$$

with

$$G_N^0(\lambda)=2\frac{1}{N}\sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]}I_\varepsilon^T(\lambda_j),\quad\lambda\in[0,\pi]^d.$$

Observe that the $U_p - process$ $\alpha_N^0$ and the $T_p - process$ $\alpha_{\theta_0,N}$ are identical when $\{x(t)\}_{t\in\mathbb{Z}^d}$ is a "white noise" process.

Let us introduce the following regularity conditions.

**Condition C1:** $(a)$ The process $\{\varepsilon(t)\}_{t\in\mathbb{Z}^d}$ in (1.1) are zero mean independent identically distributed with variance $\sigma_\varepsilon^2$ equal to 1 and finite $4th$ moments with $\kappa$ denoting the fourth cumulant.

$(b)$ The multilateral *Moving Average* representation of $\{x(t)\}_{t\in\mathbb{Z}^d}$ in (1.1) can be written (or it has a representation) as a multilateral *Autocorrelation* model

$$\sum_{j\in\mathbb{Z}^d}\xi(j)\{x(t-j)-\mu\}=\varepsilon(t)\qquad\xi(0)=1, \tag{2.11}$$

where $\xi(j)$ is the coefficient of $z^j$ in the Fourier expansion of $\mathcal{L}^{-1}(z)$, where

$$\mathcal{L}(z)=\mathcal{L}(z[1],...,z[d])=\sum_{j\in\mathbb{Z}^d}\psi(j)\,z^j$$

denoting for muti-indeces $z$ and $j$, $z^j=\prod_{\ell=1}^d z[\ell]^{j[\ell]}$ with the convention that $0^0=1$.

**Condition C2:** $N = \prod_{\ell=1}^{d} n[\ell]$, where $n[\ell] \asymp n$ for $\ell = 1, ..., d$, and "$a \asymp b$" means that $0 < C_1 \leq a/b \leq C_2 < \infty$ for some finite positive constants $C_1$ and $C_2$.

**Condition C3:** $\{h(t)\}_{t=1}^{n}$ is the cosine-bell taper function in (2.4).

We now comment on Conditions $C1$ to $C4$. Part $(a)$ of Condition $C1$ seems to be a minimal condition for Proposition 1 below to hold true. Observe that due to the quadratic nature of $\alpha_N^0$, for the latter to have finite second moments, we require finite fourth moments for the lattice process $\{\varepsilon(t)\}_{t\in\mathbb{Z}^d}$. Observe that we have assumed that the true value of $\sigma_\varepsilon^2$ is 1. The latter follows after our comments made after the definition of $G_{\theta,N}(\lambda)$ in (2.9). However, we shall emphasize that we are not saying or suggesting that the true value of $\sigma_\varepsilon^2$ is known, only that it is equal to 1. Sufficient regularity conditions required for the validity of the expansion as an *Autocorrelation in part* $(b)$ formulation is that $\Psi(z)$ be no zero for any $z[\ell]$, $\ell = 1, ...d$, which simultaneously satisfy $|z[1]| = 1, ..., |z[d]| = 1$ at least when the *Moving Average* representation is of finite order. The latter implies that $f(\lambda)$ is a positive function.

Condition $C2$ can be generalized to the case where the rate of convergence to zero of $n^{-1}[\ell]$ differs for different $\ell = 1, ..., d$. However, for notational simplicity we prefer to leave it as it stands. On the other hand, $C3$ the taper function employed for the asymptotics to follow can be more general, as those given by Kolmogorov's or Parzen's tapers. In fact, in situations where the dimension of $d$ is greater than 3, it might be needed for the results of the paper to follow. However, as the most important cases in empirical applications are covered in the paper, we shall leave the cosine-bell taper explicitly as the taper function to be employed.

We should now indicate why the formulation of the test in terms of the spectral density function might be useful. For that purpose, we have just to remember how the spectral density function is related to the conditional distribution at each site on an infinite lattice. Indeed, denoting

$$\frac{\sigma_\varepsilon^2}{f(z)} = \sum_{j\in\mathbb{Z}^d} \delta(j) z^j,$$

we then have that the conditional mean is given by

$$E\{x(t) | x(r) : \ r \neq t\} = \sum_{j\in\mathbb{Z}^d \wedge \{j\neq 0\}} \delta(j) x(t-j). \tag{2.12}$$

The latter expression gives a motivation why the spectral density function plays a central role, and therefore why we have decided to work in terms of the latter function. The equality in (2.12) is related to the well known $CAR$ (Conditional Autoregression) model compared to the $SAR$ (Simultaneous Autoregression) model in (2.11). See, for instance Besag (1974) and Whittle (1954), respectively. It is however known that the class of $CAR$ models is more general than that of $SAR$ models. In fact, as Cressie (1993, Ch.6) observed, any $SAR$ model has a $CAR$ *representation* but not vice versa, see also Besag (1974).

Some particular parameterizations of (1.1) or (2.11) are the $ARMA$ field model

$$P(L)(x(t) - \mu) = Q(L)\varepsilon(t),$$

where

$$\begin{aligned} P(z) &= \sum_{j\in\mathbb{Z}^d} \alpha(j) z^j; \quad \alpha(0) = 1 \\ Q(z) &= \sum_{j\in\mathbb{Z}^d} \beta(j) z^j; \quad \beta(0) = 1, \end{aligned}$$

are finite series in $\mathbb{Z}^d$. That is, only a finite number of the $\alpha(j)$ and $\beta(j)$ are non-zero. For instance the ARMA model given by

$$\sum_{j=-k_1}^{k_2} \alpha(j)(x(t-j)-\mu) = \sum_{j=-\ell_1}^{\ell_2} \beta(j)\varepsilon(t-j) \qquad \alpha(0)=\beta(0)=1$$

whose spectral density function is

$$f(\lambda) = \frac{\sigma_\varepsilon^2}{(2\pi)^d} \frac{\left|\sum_{j=-\ell_1}^{\ell_2} \beta(p) e^{ij\cdot\lambda}\right|^2}{\left|\sum_{j=-k_1}^{k_2} \alpha(p) e^{ij\cdot\lambda}\right|^2}.$$

Notice that the $ARMA$ field model becomes a causal representation if the polynomials $Q(L)$ and $P(L)$ are both unilateral. When say $Q(L)=0$, a condition for the latter is that $\int_{\Pi^d} \log\left|P\left(e^{i\lambda}\right)\right| d\lambda = 0$. So that in general, we can expect that $\int_{\Pi^d} \log\left|\Psi_\theta\left(e^{i\lambda}\right)\right| d\lambda \neq 0$ for any admissible value of $\theta$. The latter becomes very relevant as when formulating the Whittle estimator below in Section 3.

Another parametric model of interest is the extension of the classical Bloomfield (1973) exponential model, see also Whittle (1954) Section 6, to processes in a lattice. These models can be characterized as having a spectral density function defined as

$$f_\vartheta(\lambda) = \sigma_\varepsilon^2 \exp\left\{-\sum_{\ell \prec 0} \alpha(\ell;\theta)\cos(\ell\cdot\lambda)\right\},$$

where from now on $\vartheta = \left(\sigma_\varepsilon^2, \theta'\right)'$ and we denote by "$\prec$" the lexicographical (dictionary) ordering which is defined as

$$j \prec k \Leftrightarrow (\exists \ell > 0)(\forall i < \ell)(j[i] = k[i] \wedge j[\ell] < k[\ell]),$$

that is, if one of the terms $j[\ell] < k[\ell]$ and all the proceeding ones are equal. For instance, when $d=2$, we would then have that, say, $\ell \prec 0$ corresponds to the half plane of $\mathbb{Z}$, $\overline{\mathbb{Z}}^2 = \left\{(\ell[1], \ell[2]) \in \mathbb{Z}^d : (\ell[1] \le 0 \wedge \ell[2] = 0) \vee (\ell[1] \le 0 \wedge \ell[2] < 0)\right\}$. Observe that if we allowed $\ell$ in the last displayed equality to belong to $\mathbb{Z}^d$ the model would not be then identified as $\cos(\ell\cdot\lambda) = \cos(-\ell\cdot\lambda)$.

It is worth mentioning that Whittle (1954) showed that, almost any given stationary bilateral scheme on a plane lattice, there corresponds a unilateral autoregression having the same spectral scheme although not necessarily of finite order as is the case when $d=1$.

The empirical processes $\alpha_N^0(\lambda)$ and $\alpha_{\theta_0,N}(\lambda)$ given in (2.10) and (2.8) respectively are random elements in $D[0,\pi]^d$. The functional space $D[0,\pi]^d$ is endowed with the Skorohod's metric (see e.g. Billingsley, 1968) and convergence in distribution in the corresponding topology will be denoted by "$\Rightarrow$".

**Proposition 1.** *Under $C1-C3$, we have that*

$$\alpha_N^0(\lambda) \Rightarrow \widetilde{\boldsymbol{B}}(\lambda) = \boldsymbol{B}\left(\frac{\lambda}{\pi}\right) - \prod_{\ell=1}^{d}\left(\frac{\lambda[\ell]}{\pi}\right)\boldsymbol{B}(1) \qquad \lambda \in [0,\pi]^d, \tag{2.13}$$

*where $\left\{\boldsymbol{B}(u) : u \in [0,1]^d\right\}$ is a Brownian sheet.*

**Remark 1.** *Recall that the covariance structure of the standard Brownian sheet in $[0,1]^d$ is*

$$Cov(\boldsymbol{B}(u), \boldsymbol{B}(v)) = \prod_{\ell=1}^{d}(u[\ell] \wedge v[\ell]), \quad \text{for } u, v \in [0,1]^d.$$

Proposition 1 extends the results provided in Grenander and Rosenblatt (1957) when $d = 1$, although under stronger conditions than those assumed in the paper. In particular, we do not need to assume eight bounded moments and only finite fourth moments are required. Our main result of this section is the following theorem.

To establish the asymptotic equivalence between $\alpha_{\theta_0,N}$ and $\alpha_N^0$, we introduce the following smoothness assumptions on $\Psi(\lambda)$.

**Condition C4:** $|\Psi(\lambda)|^2$ given in (1.2) is a positive and continuously differentiable function on $[-\pi,\pi]^d$.

Our main result of this section is the following theorem.

**Theorem 1.** *Consider* (1.1) *and assume* $C1 - C4$. *Then, under* $H_0$,

$$\alpha_{\theta_0,N}(\lambda) \Rightarrow \widetilde{\boldsymbol{B}}(\lambda) \qquad \lambda \in [0,\pi]^d .$$

*Proof.* The proof is an immediate consequence of Proposition 1 and Lemma 4 after we observe that Lemma 4 with $\zeta(\lambda) = 1$ there implies that

$$2^{-1/2} N^{1/2} \sup_{\lambda \in [0,\pi]^d} \left| G_{\theta_0,N}(\lambda) - G_N^0(\lambda) \right| = o_p(1)$$

so that

$$2^{-1/2} N^{1/2} \sup_{\lambda \in [0,\pi]^d} \left| \frac{G_{\theta_0,N}(\lambda)}{G_{\theta_0,N}(\pi)} - \frac{G_N^0(\lambda)}{G_N^0(\pi)} \right| = o_p(1)$$

by standard algebra. □

**Remark 2.** *An immediate conclusion that we draw from Theorem 1 and Proposition 1 is that*

$$G_{\theta_0,N}(\pi) - \sigma_\varepsilon^2 = O_p\left(N^{-1/2}\right). \tag{2.14}$$

We now comment on the result of Theorem 1. The theorem indicates that $\alpha_{\theta_0,N}$ is asymptotically pivotal. One consequence is that critical regions of tests based on a continuous functional $\eta : D[0,\pi]^d \mapsto \mathbb{R}^+$ can be easily obtained. Different functionals $\eta$ lead to tests with different power properties. Among them are omnibus, directional and/or Portmanteau-type tests. For example, classical functionals which lead to omnibus tests are the Kolmogorov-Smirnov ($\eta(g) = \sup_{\lambda\in[0,\pi]^d} |g(\lambda)|$) and the Cramér-von Mises ($\eta(g) = \pi^{-d} \int_{[0,\pi]^d} g(\lambda)^2 \, d\lambda$), whereas Portmanteau tests, defined as weighted sums of squared estimated autocorrelations of the errors, and directional tests are obtained by choosing an appropriate functional $\eta$. (See Section 5 for some specific details.) Unfortunately, the results of Theorem 1 are only valid when the "true" value of $\theta_0$ is known, which in practical situations is unrealistic. The question is then how are our previous results affected when $\theta_0$ is estimated? This is the topic of the next section.

## 3. TESTS WHEN THE PARAMETERS ARE UNKNOWN

This section extends the results of Section 2 to the more realistic situation where we need to estimate the parameters $\theta_0$ to implement the test. That is, we replace $\theta_0$ by an estimator, say $\widehat{\theta}$ in (3.1) below, in $\alpha_{\theta,N}(\lambda)$. In this scenario and drawing the terminology from Durbin (1973), our null hypothesis becomes a composite hypothesis.

A popular estimator of $\vartheta_0' = (\theta_0', \sigma_\varepsilon^2)$ is the Whittle (1954) estimator defined as

$$\widehat{\vartheta}_c = \arg \min_{\vartheta \in \Theta \times \mathbb{R}^+} \int_{-\pi}^{\pi} \left\{ \log f_\vartheta(\lambda) + \frac{I_x^T(\lambda)}{(2\pi)^d f_\vartheta(\lambda)} \right\} d\lambda$$

or in its discrete version

$$\widehat{\vartheta} = \arg \min_{\vartheta \in \Theta \times \mathbb{R}^+} \frac{1}{N} \sum_{j=-\tilde{n}}^{\tilde{n}} \left\{ f_\vartheta(\lambda_j) + \frac{I_x^T(\lambda_j)}{(2\pi)^d f_\vartheta(\lambda_j)} \right\}, \tag{3.1}$$

where $f_\vartheta(\lambda_j) = \sigma_\varepsilon^2 |\Psi_\theta(\lambda_j)| / (2\pi)^d$ and $\Theta \subset \mathbb{R}^p$ is a compact set. Recall our notation given in (2.3), and that the true value of the variance of $\varepsilon(t)$ is unknown and therefore we need to estimate it. In this case, the test-statistic becomes

$$\alpha_{\widehat{\theta},N}(\lambda) = 2^{-1/2} N^{1/2} \left[ \frac{G_{\widehat{\theta},N}(\lambda)}{G_{\widehat{\theta},N}(\pi)} - \prod_{\ell=1}^{d} \left( \frac{\lambda[\ell]}{\pi} \right) \right], \quad \lambda \in [0,\pi]^d,$$

where $G_{\theta,N}(\lambda)$ is given in (2.9).

Notice that, contrary to the standard causal models, as Whittle (1954) noticed, the estimator of $\vartheta_0$ obtained by

$$\mathring{\theta} = \arg \min_{\theta \in \Theta} \frac{2}{N} \sum_{j=-\tilde{n}}^{\tilde{n}} \frac{I_{x,j}^T(\lambda_j)}{|\Psi_\theta(\lambda_j)|^2}, \qquad \mathring{\sigma}_\varepsilon^2 = \frac{2}{N} \sum_{j=-\tilde{n}}^{\tilde{n}} \frac{I_{x,j}^T(\lambda_j)}{|\Psi_{\widehat{\theta}}(\lambda_j)|^2}$$

is inconsistent and where now $\{\varepsilon(t)\}_{t \in \mathbb{Z}^d}$ loses its interpretation as the "prediction" error or that they can be regarded as innovations. The main reason for the lack of consistency of $\mathring{\theta}$ is that when the model is not causal then $\int_{-\pi}^{\pi} \varphi_\theta(\lambda) d\lambda \neq 0$, where from now on we write

$$\phi_\vartheta(\lambda) = \frac{\partial}{\partial \vartheta} \log f_\vartheta(\lambda) = \left( \varphi_\theta'(\lambda), \sigma_\varepsilon^{-2} \right)',$$

and

$$\varphi_\theta(\lambda) = \frac{\partial}{\partial \theta} \log |\Psi_\theta(\lambda)|^2. \tag{3.2}$$

Let is introduce the following regularity conditions on $\theta_0$ and on the model (1.1) or (2.11).

**Condition C5:** $\theta_0$ is an interior point of the compact parameter set $\Theta \subset \mathbb{R}^p$ and $\sigma_\varepsilon^2 \in \mathbb{R}^+$.

**Condition C6:** $\Psi_\theta(\lambda)$ is a positive and twice continuously differentiable function in $\theta$ on $[-\pi,\pi]^d$.

**Condition C7:** If $\theta_1 \neq \theta_2$, then $\Psi_{\theta_1}(\lambda) \neq \Psi_{\theta_2}(\lambda)$ in a set $\Delta \subset [-\pi,\pi]^d$ with positive Lebesgue measure.

The conditions imposed on $\Theta$ and the model are standard so that we omit any comment on them. Let

$$\begin{aligned} q_{\vartheta,N} &= \frac{1}{N} \sum_{j=-\tilde{n}}^{\tilde{n}} \phi_\vartheta(\lambda_j) \left\{ \frac{I_x^T(\lambda_j)}{\sigma_\varepsilon^2 |\Psi_\theta(\lambda_j)|^2} - 1 \right\} \\ Q_{\vartheta,N} &= \frac{1}{N} \sum_{j=-\tilde{n}}^{\tilde{n}} \phi_\vartheta(\lambda_j) \phi_\vartheta'(\lambda_j), \end{aligned} \tag{3.3}$$

and also, recalling our notation in (2.2),

$$\Phi_\vartheta = (2\pi)^{-d} \int_{-\pi}^{\pi} \phi_\vartheta(\lambda) d\lambda \quad \text{and} \quad \Lambda_\vartheta = (2\pi)^{-d} \int_{-\pi}^{\pi} \phi_\vartheta(\lambda) \phi_\vartheta'(\lambda) d\lambda.$$

Notice that we write explicitly $\sigma_\varepsilon^2$ as it is a parameter in itself.

**Condition C8:** $\Lambda_{\vartheta_0}$ is a continuous positive definite matrix.

**Theorem 2.** *Under C1-C3 and $C5-C8$, we have that*

$$2^{-1/2}N^{1/2}\left(\widehat{\vartheta}-\vartheta_0\right)\xrightarrow{d}\mathcal{N}\left(0,\Lambda_{\vartheta_0}^{-1}V_{\vartheta_0}\Lambda_{\vartheta_0}^{-1}\right),$$

*where* $V_{\vartheta_0}=2\Lambda_{\vartheta_0}+\kappa\left(\frac{35}{32}\right)^d\Phi_{\vartheta_0}\Phi'_{\vartheta_0}$.

*Proof.* First, by definition we know that

$$\widehat{\vartheta}-\vartheta_0=\overline{Q}_{\widetilde{\vartheta},N}^{-1}q_{\vartheta_0,N},$$

where $\widetilde{\vartheta}$ is an intermediate point between $\vartheta_0$ and $\widehat{\vartheta}$, $q_{\vartheta,N}$ is given in (3.3) and $\overline{Q}_{\vartheta,N}$ is

$$\begin{aligned} & Q_{\vartheta,N}+\frac{1}{N}\sum_{j=-\tilde{n}}^{\tilde{n}}\left(2\phi_\vartheta(\lambda_j)\phi'_\vartheta(\lambda_j)-\frac{\partial^2 f_\vartheta(\lambda_j)}{\partial\vartheta\partial\vartheta'}\right)\left\{\frac{I_x^T(\lambda_j)}{\sigma_\varepsilon^2\left|\Psi_\theta(\lambda_j)\right|^2}-1\right\} \\ = \; & Q_{\vartheta,N}+o_p(1) \end{aligned}$$

by Lemma 5. On the other hand, by Brillinger (1981, p.15) and standard arguments, since $\widetilde{\vartheta}-\vartheta_0=o_p(1)$ by Lemma 6, we have that $Q_{\widetilde{\vartheta},N}-\Lambda_{\vartheta_0}=o_p(1)$. Next, by Lemma 4 with $\zeta(\lambda)=\phi_{\vartheta_0}(\lambda_j)$ there,

$$q_{\vartheta_0,N}=\frac{1}{N}\sum_{j=-\tilde{n}}^{\tilde{n}}\phi_{\vartheta_0}(\lambda_j)\left\{I_\varepsilon^T(\lambda_j)-1\right\}+o_p(1).$$

From here the proof follows as in Robinson and Vidal-Sanz (2006). □

Looking at the proof of Theorem 2, and denoting in what follows

$$\begin{aligned} \widetilde{\varphi}_\theta(\lambda) &= \varphi_\theta(\lambda)-\frac{2}{(2\pi)^d}\int_{-\pi}^{\pi}\varphi_\theta(\lambda)\,d\lambda, \quad \widetilde{\phi}_\vartheta(\lambda)=\left(\widetilde{\varphi}'_\theta(\lambda),0\right)' \\ \widetilde{\varphi}_{\theta,N}(\lambda_j) &= \varphi_\theta(\lambda_j)-\frac{2}{N}\sum_{j=-\tilde{n}}^{\tilde{n}}\varphi_\theta(\lambda_j), \quad \widetilde{\phi}_{\vartheta,N}(\lambda)=\left(\widetilde{\varphi}'_{\theta,N}(\lambda),0\right)' \end{aligned}$$

with $\varphi_\theta(\lambda)$ given in (3.2), standard algebra establishes that the Whittle estimator $\widehat{\vartheta}$ in (3.1) satisfies the asymptotic linearization is

$$\begin{aligned} \left(\widehat{\vartheta}-\vartheta_0\right) &= Q_{\vartheta_0,N}^{-1}\left\{\int_{-\pi}^{\pi}\widetilde{\phi}_{\theta_0}(\lambda)\,\alpha_{\theta_0,N}(d\lambda)+\int_{-\pi}^{\pi}\phi_{\vartheta_0}(\lambda)\,d\lambda\frac{1}{N}\sum_{j=-\tilde{n}}^{\tilde{n}}\left(\frac{I_x^T(\lambda_j)}{(2\pi)^d f_\vartheta(\lambda_j)}-1\right)\right\} \\ & \quad +o_p\left(N^{-1/2}\right). \end{aligned} \tag{3.4}$$

Then using (3.4) and defining

$$\alpha_\infty(\lambda)=\widetilde{\boldsymbol{B}}(\lambda)-\left(\frac{1}{(2\pi)^d}\int_{-\lambda}^{\lambda}\widetilde{\varphi}'_{\theta_0}(\bar{\lambda})\,d\bar{\lambda}\right)\widetilde{\Lambda}^{-1}(\theta_0)\int_{-\pi}^{\pi}\widetilde{\varphi}_{\theta_0,N}(\bar{\lambda})\,\widetilde{\boldsymbol{B}}(d\bar{\lambda}),$$

where

$$\widetilde{\Lambda}_\theta=\frac{1}{(2\pi)^d}\int_{-\pi}^{\pi}\widetilde{\varphi}_\theta(\bar{\lambda})\,\widetilde{\varphi}'_\theta(\bar{\lambda})\,d\bar{\lambda},$$

we obtain the following result.

**Theorem 3.** *Under $H_0$ and assuming $C1-C3$ and $C5-C8$, uniformly in $\lambda\in[0,\pi]$,*

$$\begin{aligned} (a)\;\; \alpha_{\widehat{\theta},N}(\lambda) &= \alpha_N^0(\lambda)-\left(\frac{1}{N}\sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]}\widetilde{\varphi}'_{\theta_0,N}(\lambda_j)\right)\widetilde{\Lambda}_{\theta_0,N}^{-1}\frac{1}{N}\sum_{j=-\tilde{n}}^{\tilde{n}}\widetilde{\varphi}'_{\theta_0,N}(\lambda_j)\,I_\varepsilon^T(\lambda_j) \\ & \quad +o_p(1), \end{aligned}$$

*where the* $o_p(1)$ *is uniformly in* $\lambda$ *and*

$$\widetilde{\Lambda}_{\theta,N} = \frac{1}{N}\sum_{j=-\tilde{n}}^{\tilde{n}} \widetilde{\varphi}_{\theta,N}(\lambda_j)\,\widetilde{\varphi}'_{\theta,N}(\lambda_j).$$

(*b*) $\alpha_{\widehat{\theta},N} \Rightarrow \alpha_\infty$.

The main conclusion that we draw from Theorem 3 is that the $T_p$-process $\alpha_{\widehat{\theta},N}$ is no longer asymptotically pivotal. The immediate consequence is that tests based on continuous functionals $\eta : D[0,\pi]^d \to \mathbb{R}$, say $\eta\left(\alpha_{\widehat{\theta},N}\right) = \sup_{\lambda\in[0,\pi]^d}\left|\alpha_{\widehat{\theta},N}(\lambda)\right|$ or $\eta\left(\alpha_{\widehat{\theta},N}\right) = (2\pi)^{-d}\int_{-\pi}^{\pi}\alpha_{\widehat{\theta},N}(\lambda)^2\,d\lambda$, are not useful for practical purposes. To overcome the problem of the implementation of $\eta\left(\alpha_{\widehat{\theta},N}\right)$, several approaches have been described and examined to able to compute the critical values of the asymptotic distribution of $\eta\left(\alpha_{\widehat{\theta},N}\right)$. A first approach proposes the usage of bootstrap algorithms. This is the route employed, among others, by Chen and Romano (2000) or Hainz and Dahlhaus (2000) for short-range models using the $U_p-process$ and by Hidalgo and Kreiss (2006), who allow also long-range dependence models using the $T_p-process$. Of course all those works were for $d=1$. A second alternative makes use of a tuning or smoothing parameter that must behave in some required way as the sample size increases. This procedure makes the asymptotic distribution of the tests to be pivotal, so that its critical values are readily available. Among them, the most popular one is the Portmanteau test. Box and Pierce (1970) showed that the partial sum of the residuals squared autocorrelations of a stationary ARMA process is approximately chi-squared distributed assuming that the number of autocorrelations considered diverges to infinity with the sample size at an appropriate rate. Alternatively we could employ a frequency domain approach as in Hong (1996) or Paparoditis (2000), who compared a nonparametric estimator of $f(\lambda)$ and the parametric one. The first downside of the method is that the power of the test is smaller than the one proposed in the paper, that is if we denote by $h_N$ the smoothing parameter, their test has a local power of order $(Nh_N)^{-1/2}$ whereas ours is $N^{-1/2}$. A second potential drawback is that the choice of the bandwidth seems an artifact when testing for a particular parametric family and the final outcome of all these tests may depend on the arbitrary choice of the tuning/smoothing parameters for which no relevant theory is available. That is, there are not rules available on how to choose the bandwidth parameter. A third approach is in the spirit of Durbin, Knott and Taylor (1976) for the classical empirical process, and it was the route followed by Anderson (1997), who proposed to approximate the critical values of the Cramér-von Mises tests for a stationary AR model. The method considers a truncated version of the spectral representation of $\alpha_{\widehat{\theta},N}$ with estimated orthogonal components. The number of estimated orthogonal components must suitably increase with the sample size. Although the latter is for standard time series, e.g. $d=1$, in Section 5, we describe how it can be extended to the case when $d>1$. However, its implementation is very cumbersome even for the rather simpler case when $d=1$. See for instance Anderson (1997) for details.

Finally, a different procedure is that proposed and introduced by Brown, Durbin and Evans (1975) and examined in depth by Khmaladze (1981). This is the topic of the next section.

## 4. MARTINGALE TRANSFORMATION

Theorem 3 indicates that the asymptotic critical values of tests based on $\alpha_{\widehat{\theta},N}$ are model dependent. This implies that to tabulate its critical values can be difficult, if at all possible. However, in this section we will see that we can employ a martingale transformation to the statistic $\alpha_{\widehat{\theta},N}$ such that the transformed process converges in distribution to the standard Brownian sheet. To this end, it is first of all of interest to realize that Theorem 3 (a) provides an asymptotic representation of $\alpha_{\widehat{\theta},N}$ as a scaled cumulative sum (cusum) of the least squares residuals in an artificial regression model. Indeed, the uniform asymptotic expansion in Theorem 3 (a) together with (2.14) indicates that

$$\sup_{\lambda\in[0,\pi]^d}\left|\alpha_{\widehat{\theta},N}(\lambda)-\frac{2\pi}{G_N^0(\pi)}\frac{1}{N^{1/2}}\sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]}u_N(j)\right|=o_p(1),$$

where for $j=-\tilde{n},...,\tilde{n}$

$$u_N(j)=I_\varepsilon^T(\lambda_j)-\widetilde{\varphi}'_{\theta_0,N}(\lambda_j)\left[\sum_{k=-\tilde{n}}^{\tilde{n}}\widetilde{\varphi}_{\theta_0,N}(\lambda_k)\widetilde{\varphi}'_{\theta_0,N}(\lambda_k)\right]^{-1}\sum_{k=-\tilde{n}}^{\tilde{n}}\widetilde{\varphi}_{\theta_0,N}(\lambda_k)I_\varepsilon^T(\lambda_k),$$

are the least squares residuals in an artificial regression model with dependent variable $I_\varepsilon^T(\lambda_j)$ and the vector of explanatory variables $\widetilde{\varphi}_{\theta_0,N}(\lambda_j)$. This fact suggests to employ the cusum of recursive least squares residuals for constructing asymptotically pivotal tests, as they were originally proposed by Brown, Durbin and Evans (1975).

To that end, let us define $p^*=(0,-\tilde{n},...,-\tilde{n},-\tilde{n}+p+1)$

$$A_{\theta,N}(j)=\frac{1}{N}\sum_{k\prec j}\widetilde{\varphi}_{\theta,N}(\lambda_k)\widetilde{\varphi}'_{\theta,N}(\lambda_k),$$

where recall "$\prec$" denotes the lexicographic ordering and let's assume that

**C9:** $A_{\theta_0,N}(p^*)$ is non singular.

Note that $p$ is the dimension of the parameter $\theta_0$. Also observe that $A_{\theta_0,N}$ is a deterministic matrix and thus it can be computed directly from the model. In particular, Condition $C9$ is satisfied for all standard models used with empirical data.

The (scaled) cusum of backward recursive least squares residuals is defined as

$$\beta_N^0(\lambda)=\frac{1}{G_N^0(\pi)}\frac{1}{2^{1/2}N^{1/2}}\sum_{j=-[\tilde{n}\lambda/\pi];p^*\prec j}^{[\tilde{n}\lambda/\pi]}e_N(j),\quad \lambda\in[0,\pi]^d,$$

where

$$e_N(j)=I_\varepsilon^T(\lambda_j)-\widetilde{\varphi}'_{\theta_0,N}(\lambda_j)b_N(j),\quad p^*\prec j,$$

are the backward least squares residuals and

$$b_N(j)=A_{\theta_0,N}^{-1}(j)\frac{1}{N}\sum_{k\prec j}\widetilde{\varphi}_{\theta_0,N}(\lambda_k)I_\varepsilon^T(\lambda_k).$$

It is worth observing that the motivation not to employ the first "$p$" Fourier frequencies to compute the least squares recursive residuals $e_N(j)$ is due to the singularity of $A_{\theta_0,N}(j)$ for all $j\preceq p^*$.

The empirical process $\beta_N^0$ can be written as a linear transformation of $\alpha_N^0$,

$$\beta_N^0(\lambda)=\mathcal{L}_{\theta_0,N}\alpha_N^0(\lambda),\ \lambda\in[0,\pi]^d,$$

where, for any function $g\in D[0,\pi]^d$,

$$\mathcal{L}_{\theta,N}g(\lambda)=g(\lambda)-\frac{1}{N}\sum_{j=-[\tilde{n}\lambda/\pi];p^*\prec j}^{[\tilde{n}\lambda/\pi]}\widetilde{\varphi}'_{\theta_0,N}(\lambda_j)A_{\theta,N}^{-1}(j)\int_{-\lambda_{j+1}}^{\lambda_{j+1}}\widetilde{\varphi}_\theta\left(\widetilde{\lambda}\right)g\left(d\widetilde{\lambda}\right).$$

The transformation $\mathcal{L}_{\theta_0,N}$ has the limiting version $\mathcal{L}^0$, defined as

$$\mathcal{L}^0 g(\lambda) = g(\lambda) - \frac{1}{(2\pi)^d}\int_{-\lambda}^{\lambda} \widetilde{\varphi}_{\theta_0}\left(\overline{\lambda}\right) A_{\theta_0}^{-1}\left(\overline{\lambda}\right) \int_{-\overline{\lambda}}^{\overline{\lambda}} \widetilde{\varphi}_{\theta_0}\left(\widetilde{\lambda}\right) g\left(d\widetilde{\lambda}\right) d\overline{\lambda},$$

where

$$A_{\theta_0}(\lambda) = \int_{-\lambda}^{\lambda} \widetilde{\varphi}_{\theta_0}\left(\widetilde{\lambda}\right) \widetilde{\varphi}_{\theta_0}'\left(\widetilde{\lambda}\right) d\widetilde{\lambda}.$$

Notice that $\mathcal{L}^0\alpha_\infty$ is the martingale innovation of $\alpha_\infty$, see Khmaladze (1981) and that $A_\theta(\lambda) = (2\pi)^d \widetilde{\Lambda}_\theta$.

This type of martingale transformation has been used by Khmaladze (1981) in the standard goodness-of-fit testing problem, by Nikabadze and Stute (1997) for goodness-of-fit of distribution functions under random censorship, by Stute, Thies and Zhu (1998), Koul and Stute (1999) and Khmaladze and Koul (2004) for dynamic regression models, and by Stute and Zhu (2002) for generalized linear models. See also Delgado, Hidalgo and Velasco (2005).

**Theorem 4.** *Under $H_0$ and assuming $C1-C3$ and $C9$,*

$$\beta_N^0(\lambda) \Rightarrow \boldsymbol{B}\left(\frac{\lambda}{\pi}\right) \qquad \lambda \in [0,\pi]^d.$$

Because $\beta_N^0$ cannot be computed in practice, as it depends on $\theta_0$, we suggest to employ $\beta_{\widehat{\theta},N}$, where

$$\begin{aligned}\beta_{\theta,N}(\lambda) &= \mathcal{L}_{\theta,N}\alpha_{\theta,N}(\lambda)\\ &= \frac{1}{G_{\theta,N}(\pi)}\frac{2^{1/2}}{N^{1/2}}\sum_{j=-[\tilde{n}\lambda/\pi];p^*\prec j}^{[\tilde{n}\lambda/\pi]} e_{\theta,N}(j),\ \lambda\in[0,\pi]^d\end{aligned}$$

and

$$e_{\theta,N}(j) = \frac{I_x(\lambda_j)}{|\Psi_\theta(\lambda_j)|} - \widetilde{\varphi}_\theta'(\lambda_j)\, b_{\theta,N}(j),\ j \succ p^*,$$

are the backward recursive least squares residuals in the linear projection of $I_x^T(\lambda_j)/|\Psi_\theta(\lambda_j)|$ on $\widetilde{\varphi}_\theta(\lambda_j)$, and where

$$b_{\theta,N}(j) = A_{\theta,N}^{-1}(j)\frac{1}{N}\sum_{k\prec j}\widetilde{\varphi}_{\theta,N}(\lambda_k)\frac{I_x^T(\lambda_k)}{|\Psi_\theta(\lambda_k)|}.$$

In order to establish the asymptotic equivalence between $\beta_N^0$ and $\beta_{\widehat{\theta},N}$, we also need some extra smoothness assumptions on the model under the null.

**C10:** For some $0<\delta<1$ and all $\lambda\in(0,\pi]^d$, there exists a constant $C<\infty$ such that

$$\sup_{\{\theta:\|\theta-\theta_0\|\le\delta\}}\frac{1}{\|\theta-\theta_0\|^2}\left\|\varphi_\theta(\lambda)-\varphi_{\theta_0}(\lambda)-\dot{\varphi}_\theta(\lambda)(\theta-\theta_0)\right\|\le C,$$

and $\dot{\varphi}_\theta$ is a continuous function in $[0,\pi]^d$.

**Theorem 5.** *Under $H_0$ and assuming $C1-C3$ and $C5-C10$,*

$$\sup_{\lambda\in[0,\pi]^d}\left|\beta_{\widehat{\theta},N}(\lambda)-\beta_N^0(\lambda)\right| = o_p(1).$$

Theorem 5 holds true, mutatis mutandis, with $\widehat{\theta}$ replaced by any $N^{1/2}$-consistent estimator. Also, from a computational point of view, it is worth observing that

$$A_{\theta,N}^{-1}(j^+) = A_{\theta,N}^{-1}(j) - \frac{A_{\theta,N}^{-1}(j)\,\widetilde{\varphi}_{\theta,N}(\lambda_{j^+})\,\widetilde{\varphi}_{\theta,N}'(\lambda_{j^+})\,A_{\theta,N}^{-1}(j)}{N+\widetilde{\varphi}_{\theta,N}'(\lambda_{j^+})\,A_{\theta,N}^{-1}(j)\,\widetilde{\varphi}_{\theta,N}(\lambda_{j^+})}$$

and

$$b_{\theta,N}\left(j^{+}\right)=b_{\theta,N}\left(j\right)+A_{\theta,N}^{-1}\left(j^{+}\right)\widetilde{\varphi}_{\theta,N}\left(\lambda_{j^{+}}\right)\left[\frac{I_x^T\left(\lambda_{j^{+}}\right)}{\left|\Psi_\theta\left(\lambda_{j^{+}}\right)\right|}-\widetilde{\varphi}_{\theta,N}'\left(\lambda_{j^{+}}\right)b_{\theta,N}\left(j\right)\right].$$

see Brown, Durbin and Evans (1975) for similar arguments, where $j^{+}$ denotes $\left(j\left[1\right],j\left[2\right],j\left[3\right]+1\right)$ if $j\left[3\right]<\tilde{n}$, it is $\left(j\left[1\right],j\left[2\right]+1,-\tilde{n}\right)$ if $j\left[3\right]=\tilde{n}$ and $j\left[2\right]<\tilde{n}$ and it is $\left(j\left[1\right]+1,-\tilde{n},-\tilde{n}\right)$ otherwise.

Alternatively to $\beta_{\widehat{\theta},N}$, we could have considered the cusum of forward recursive residuals, i.e.

$$\overline{\beta}_{\widehat{\theta},N}\left(\lambda\right)=\frac{1}{G_{\widehat{\theta},N}\left(\pi\right)}\frac{1}{N^{1/2}}\sum_{j=-\left[\tilde{n}\lambda/\pi\right];j\prec\overline{p}}^{\left[\tilde{n}\lambda/\pi\right]}\overline{e}_{\widehat{\theta},N}\left(j\right),\quad\lambda\in\left[0,\pi\right],$$

where

$$\overline{e}_{\theta,N}\left(j\right)=\frac{I_x^T\left(\lambda_j\right)}{\left|\Psi_\theta\left(\lambda_j\right)\right|}-\widetilde{\varphi}_{\theta,N}'\left(\lambda_j\right)\overline{b}_{\theta,N}\left(j\right),\quad j\prec\overline{p},$$

with $\overline{p}=\left(\tilde{n},...,\tilde{n},\tilde{n}-p\right)$ and

$$\begin{aligned}\overline{b}_{\theta,N}\left(j\right)&=\overline{A}_{\theta,N}^{-1}\left(j\right)\frac{1}{N}\sum_{j\prec k}^{\tilde{n}}\widetilde{\varphi}_{\theta,N}\left(\lambda_k\right)\frac{I_x^T\left(\lambda_k\right)}{\left|\Psi_\theta\left(\lambda_k\right)\right|}\\ \overline{A}_{\theta,N}\left(j\right)&=\frac{1}{N}\sum_{j\prec k}^{\tilde{n}}\widetilde{\varphi}_{\theta,N}\left(\lambda_k\right)\widetilde{\varphi}_{\theta,N}'\left(\lambda_k\right).\end{aligned}$$

In this case, we might take advantage of the computational formulae

$$\overline{A}_{\theta,N}^{-1}\left(j\right)=\overline{A}_{\theta,N}^{-1}\left(j^{-}\right)-\frac{\overline{A}_{\theta,N}^{-1}\left(j^{-}\right)\widetilde{\phi}_{\theta,N}\left(\lambda_j\right)\widetilde{\phi}_{\theta,N}'\left(\lambda_j\right)\overline{A}_{\theta,N}^{-1}\left(j^{-}\right)}{N+\widetilde{\phi}_{\theta,N}'\left(\lambda_j\right)\overline{A}_{\theta,N}^{-1}\left(j^{-}\right)\widetilde{\phi}_{\theta,N}\left(\lambda_j\right)},$$

and

$$\overline{b}_{\theta,N}\left(j\right)=\overline{b}_{\theta,N}\left(j^{-}\right)+\overline{A}_{\theta,N}^{-1}\left(j\right)\widetilde{\phi}_{\theta,N}\left(\lambda_j\right)\left[\frac{I_x^T\left(\lambda_j\right)}{\left|\Psi_\theta\left(\lambda_j\right)\right|}-\widetilde{\phi}_{\theta,N}'\left(\lambda_j\right)b_{\theta,N}\left(j^{-}\right)\right],$$

where $j^{-}$ denotes $\left(j\left[1\right],j\left[2\right],j\left[3\right]-1\right)$ if $j\left[3\right]>-\tilde{n}$, it is $\left(j\left[1\right],j\left[2\right]-1,\tilde{n}\right)$ if $j\left[3\right]=-\tilde{n}$ and $j\left[2\right]>-\tilde{n}$ and it is $\left(j\left[1\right]-1,\tilde{n},\tilde{n}\right)$ otherwise.

This formulation may be useful in small samples when we suspect that the main discrepancy between the null and the alternative is near $\pi$. However, from Theorems 4 and 5, it is easily seen that the empirical processes $\overline{\beta}_{\widehat{\theta},N}$ and $\beta_{\widehat{\theta},N}$ have the same asymptotic behaviour.

Let $\eta:D\left[0,\pi\right]^d\rightarrow\mathbb{R}$ be a continuous functional, under $H_0$ and the conditions in Theorem 5,

$$\eta\left(\beta_{\widehat{\theta},N}\right)\overset{d}{\rightarrow}\varphi\left(\boldsymbol{B}\right),$$

as a consequence of the continuous mapping theorem. For instance,

$$\begin{aligned}\widehat{K}_N&=\sup_{\left(p^*\prec j\right)\cap\left(-\tilde{n}<j<\tilde{n}\right)}\left|\beta_{\widehat{\theta},N}\left(\frac{j\pi}{\tilde{n}}\right)\right|\overset{d}{\rightarrow}\sup_{\lambda\in\left[0,\pi\right]^d}\left|\boldsymbol{B}\left(\frac{\lambda}{\pi}\right)\right|\overset{d}{=}\sup_{\omega\in\left[0,1\right]^d}\left|\boldsymbol{B}\left(\omega\right)\right|,\\ \widehat{C}_N&=\frac{1}{N}\sum_{j=-\tilde{n}\wedge\left(j^*\prec j\right)}^{\tilde{n}}\beta_{\widehat{\theta},N}\left(\frac{j\pi}{\tilde{n}}\right)^2\overset{d}{\rightarrow}\frac{1}{\pi^d}\int_0^\pi\boldsymbol{B}^2\left(\frac{\lambda}{\pi}\right)d\lambda\overset{d}{=}\int_0^1\boldsymbol{B}^2\left(\omega\right)d\omega.\end{aligned}$$

## 5. LOCAL ALTERNATIVES: OMNIBUS, DIRECTIONAL AND PORTMANTEAU TESTS

In this section, we shall show that tests based on $\beta_{\widehat{\theta},N}$ are able to detect local alternatives of the type

$$H_{1N}:\ \Psi(\lambda)=\Psi_{\theta_0}(\lambda)\left(1+\tau\frac{1}{N^{1/2}}l(\lambda)+\frac{1}{N}s_N(\lambda)\right),\ \lambda\in[0,\pi]^d\ \text{ for some } \theta_0\in\Theta,$$

where $\int_{[0,\pi]^d} l(\lambda)\,d\lambda=0$, $l(\lambda)$ satisfies the same properties as $\varphi_{\theta_0}$ in $C10$, $\tau$ is a constant, possibly unknown, and for some finite $N_0$, $\sup_{N>N_0}|s_N(\cdot)|$ is an integrable function. Let us consider a couple of examples.

**Example 1.** *Take $d=2$. We wish to study departures of total independence (the white noise) hypothesis in the direction of a first-order isotropic scheme*

$$\mathbb{E}\{x(t,s)\,|...\}=\frac{\theta}{N^{1/2}}\left(x(t-1,s)+x(t+1,s)+x(t,s-1)+x(t,s+1)\right),$$

*then we have that*

$$\frac{|\Psi(\lambda)|^2}{|\Psi_{\theta_0}(\lambda)|^2}=1-2\frac{\theta}{N^{1/2}}\{\cos(\lambda[1])+\cos(\lambda[2])\},\ \lambda\in[0,\pi]^d.$$

*So, in this case*

$$l(\lambda)=-2\{\cos(\lambda[1])+\cos(\lambda[2])\}\ \text{ and } \tau=\theta,$$

*where the remainder function $s_N(\lambda)$ being equal to zero.*

**Example 2.** *Take $d=2$. Suppose now that we wish to study departures of total independence (the white noise) hypothesis in the direction of a first-order (isotropic) SAR scheme, instead of the CAR scheme of Example 1. That is, we have that under the local alternatives*

$$x(t,s)=\frac{\theta}{N^{1/2}}\left(x(t-1,s)+x(t+1,s)+x(t,s-1)+x(t,s+1)\right)+\varepsilon(t,s).$$

*Then, we have that*

$$\frac{|\Psi(\lambda)|^2}{|\Psi_{\theta_0}(\lambda)|^2}=1-2\frac{\theta}{N^{1/2}}\{\cos(\lambda[1])+\cos(\lambda[2])\}+\frac{\theta}{N}s_N(\lambda),\ \lambda\in[0,\pi]^d.$$

*So, in this case*

$$l(\lambda)=-2\{\cos(\lambda[1])+\cos(\lambda[2])\}\ \text{ and } \tau=\theta,$$

*where the remainder function $s_N(\lambda)$ is a function of* $\cos(2\lambda[1]),\cos(\lambda[1]),\cos(2\lambda[2])$ *and* $\cos(\lambda[2])$ *and being such that $|s_N(\lambda)|<C$.*

For $\lambda\in[0,\pi]^d$, let us define

(5.1)

$$L(\lambda)=\frac{1}{(2\pi)^d}\int_{-\lambda}^{\lambda}\left\{l\left(\overline{\lambda}\right)-\widetilde{\varphi}'_{\theta_0}\left(\overline{\lambda}\right)A^{-1}_{\theta_0}\left(\overline{\lambda}\right)\frac{1}{(2\pi)^d}\int_{-\overline{\lambda}}^{\overline{\lambda}}\widetilde{\varphi}_{\theta_0}\left(\widetilde{\lambda}\right)l\left(\widetilde{\lambda}\right)d\widetilde{\lambda}\right\}d\overline{\lambda}$$

and

$$\boldsymbol{M}(\lambda)=\boldsymbol{B}\left(\frac{\lambda}{\pi}\right)+\tau\cdot\boldsymbol{L}(\lambda),\ \lambda\in[0,\pi]^d.$$

We have the following theorem.

**Theorem 6.** *Assuming the same conditions as in Theorem 5, under $H_{1N}$,*

$$\beta_{\widehat{\theta},N}\Rightarrow\boldsymbol{M}.$$

*Proof.* The proof follows by Theorem 5 and standard arguments, so it is omitted. □

When $d = 1$, we know that using the fact that $\boldsymbol{M}$ and $\boldsymbol{B}$ are identically distributed, except for the deterministic shift $\tau \cdot \boldsymbol{L}$, and taking into account that $2^{1/2}\sin((j-1/2)\lambda)$ and $1/(j-1/2)^2\pi^2$ are the eigenfunctions and eigenvalues in the Kac-Siegert representation of $\boldsymbol{B}(\lambda/\pi)$, the orthogonal components of $\boldsymbol{M}$

$$m(j) = 2^{1/2}\left(j-\frac{1}{2}\right)\int_0^{\pi}\sin\left(\left(j-\frac{1}{2}\right)\lambda\right)\boldsymbol{M}(\lambda)\,d\lambda,\ \ j=1,2,...$$

are independently distributed normal random variables with mean $\tau \cdot \ell(j)$ and variance 1, where

$$\ell(j) = 2^{1/2}\left(j-\frac{1}{2}\right)\int_0^{\pi}\sin\left(\left(j-\frac{1}{2}\right)\lambda\right)\boldsymbol{L}(\lambda)\,d\lambda,\ \ j=1,2,....$$

In the general case $d \geq 1$, the previous formulae becomes

$$m(j) = 2^{d/2}\prod\nolimits_{\ell=1}^{d}\left(j[\ell]-\frac{1}{2}\right)\int_{-\pi}^{\pi}\prod\nolimits_{\ell=1}^{d}\sin\left(\left(j[\ell]-\frac{1}{2}\right)\lambda\right)\boldsymbol{M}(\lambda)\,d\lambda,\quad j=1,2,...$$

which are independently distributed normal random variables with mean $\tau \cdot \ell(j)$ and variance 1, with

$$\ell(j) = 2^{d/2}\prod\nolimits_{\ell=1}^{d}\left(j[\ell]-\frac{1}{2}\right)\int_{-\pi}^{\pi}\prod\nolimits_{\ell=1}^{d}\sin\left(\left(j[\ell]-\frac{1}{2}\right)\lambda\right)\boldsymbol{L}(\lambda)\,d\lambda,\quad j=1,2,....$$

Using, the (asymptotically) orthogonal components of $\beta_{\widehat{\theta},N}$, for $j = 1, 2, ...$,

$$\widetilde{m}_N(j) = 2^{d/2}\prod\nolimits_{\ell=1}^{d}\left(j[\ell]-\frac{1}{2}\right)\int_{-\pi}^{\pi}\prod\nolimits_{\ell=1}^{d}\sin\left(\left(j[\ell]-\frac{1}{2}\right)\lambda\right)\beta_{\widehat{\theta},N}(\lambda)\,d\lambda,\ \ ,$$

we obtain the spectral representation,

$$\beta_{\widehat{\theta},N}(\lambda) = 2^{d/2}\sum_{j=1}^{\infty}\widetilde{m}_N(j)\prod\nolimits_{\ell=1}^{d}\frac{\sin\left(\left(j[\ell]-\frac{1}{2}\right)\lambda\right)}{\pi\left(j[\ell]-\frac{1}{2}\right)},\ \lambda\in[0,\pi]^d.$$

By Theorem 6 and the continuous mapping theorem, finitely many of the $\widetilde{m}_N(j)'s$ converge in distribution to the corresponding $m(j)'s$ under $H_{1N}$. On the other hand, using Parseval's Theorem, we obtain that

$$\widehat{C}_N \xrightarrow{d} \sum_{j=1}^{\infty}\frac{m^2(j)}{\pi^{2d}\prod\nolimits_{\ell=1}^{d}(j[\ell]-\frac{1}{2})^2}.$$

Using similar arguments to those in Eubank and LaRicca (1992) in the context of the standard empirical process with estimated parameters, tests based on

$$\widetilde{W}_{q,N} = \sum_{j=1}^{q}\widetilde{m}_N^2(j),$$

with a reasonable choice of $q \geq 1$, will lead to gains in power, compared to $\widehat{C}_N$, in the direction of alternatives with significant autocorrelations at high lags. These Portmanteau tests are related to Neyman's (1937) smooth tests, a compromise between omnibus and directional tests, and for each $q \geq 1$, under $H_{1N}$, we have that

$$\widetilde{W}_{q,N} \xrightarrow{d} \chi_q^2\left(\tau^2\sum_{j=1}^{q}\ell^2(j)\right).$$

That is, tests based on $\widetilde{W}_{q,N}$ are asymptotically pivotal under $H_0$ ($\tau = 0$) for each choice of $q$, and more importantly, they are able to detect local alternatives converging to the null at the parametric rate $N^{-1/2}$, provided that $\ell(j) \neq 0$ for some

$j=1,...,n$. The latter is in contrast with the classical Portmanteau tests based on

$$\widetilde{Q}_{q_N,N}=\sum_{j=1}^{q_N}\left(N^{1/2}\widetilde{\rho}_N(j)\right)^2, \tag{5.2}$$

where $\widetilde{\rho}_N(j)$ is some estimate of the "$j-th$" autocorrelation of the residuals. It can be shown(as in the case $d=1$) that $\widetilde{Q}_{q_N,N}$ is approximately distributed as a $\chi^2_{q_N-p}$ under $H_0$ specifying a short-range model and assuming that $q_N$ diverges as $N\to\infty$. On the other hand, the resulting test is able to detect alternatives converging to the null at the rate $q_N^{1/4}N^{-1/2}$ (see e.g. Hong, 1996) when $d=1$, which is slower than $N^{-1/2}$.

In practice, it is recommended to use the discrete version

$$\widehat{W}_{q,N}=\sum_{j=1}^{q}\widehat{m}_N^2(j)$$

of $\widetilde{W}_{q,N}$, with

$$\widehat{m}_N(j)=2^{d/2}\left\{\prod_{\ell=1}^{d}\left(j[\ell]-\frac{1}{2}\right)\right\}\frac{\pi^d}{N}\sum_{k=-\tilde{n}}^{\tilde{n}}\prod_{\ell=1}^{d}\sin\left(\left(j[\ell]-\frac{1}{2}\right)\frac{\pi k}{\tilde{n}}\right)\beta_{\widehat{\theta},N}\left(\frac{\pi k}{\tilde{n}}\right).$$

On the other hand, optimal tests of $H_0$ in the direction $H_{1N}$ can be constructed applying results in Grenander (1950) (see also Grenander 1981, and references therein), as was suggested by Stute (1997) in the context of goodness-of-fit testing of a regression function. Asymptotically, testing for $H_0$ in the direction of $H_{1N}$ is equivalent to test $\overline{H}_0:\mathbb{E}(m(j))=0$ for all $j\geq 1$, against $\overline{H}_1:\mathbb{E}(m(j))=\tau\cdot\ell(j)$ for all $j\geq 1$ with $L$ known, but maybe with unknown $\tau$. Under $\overline{H}_0$, the distribution of $\{m(j)\}_{j\geq 1}$ is completely specified, as is also under $\overline{H}_1$ when the parameter $\tau$ is known. Then, the likelihood-ratio for a finite dimensional set $(m(1),...,m(q))$ is

$$\Lambda_q=\exp\left(\tau\cdot\sum_{j=1}^{q}\ell(j)\left(m(j)-\frac{\tau\cdot\ell(j)}{2}\right)\right). \tag{5.3}$$

Grenander (1950) showed that $\Lambda_q\to_p\Lambda_\infty$ as $q\to\infty$, and that the most powerful test at the $\alpha$ significance level has a critical region of the form $\{\Lambda_\infty>k\}$, with $P_0\{\Lambda_\infty>k\}=\alpha$ if $\sum_{j=1}^{\infty}\ell^2(j)<\infty$. The latter condition is satisfied in our context by Parseval's Theorem and A3(c) because $l$ is a square integrable function.

Define

$$\psi=\frac{\sum_{j=1}^{\infty}\ell(j)\,m(j)}{\left(\sum_{j=1}^{\infty}\ell^2(j)\right)^{1/2}}.$$

Then under $H_0$, $\psi\overset{d}{=}N(0,1)$, and in view of (5.3), $\psi$ forms a basis to obtain optimal critical regions. When the sign of $\tau$ is known, the critical region of the uniformly most powerful test at the $\alpha$ significance level is $\{\psi>z_{1-\alpha}\}$ when $\tau>0$ and $\{\psi<-z_{1-\alpha}\}$ when $\tau<0$, where $z_\upsilon$ is the $\upsilon$ quantile of the standard normal. Also, when the sign of $\tau$ is unknown, the most powerful unbiased test at the $\alpha$ significance level has critical region given by $\left\{|\psi|>z_{1-\alpha/2}\right\}$.

These arguments suggest an (asymptotically) optimal Neyman-Pearson test in the direction of $H_{1N}$ based on the first $q$ orthogonal components of $\beta_{\widehat{\theta},N}$, using the test statistic

$$\widehat{\psi}_{q,N}=\frac{\sum_{j=1}^{q}\ell(j)\,\widehat{m}_N(j)}{\left(\sum_{j=1}^{n}\ell^2(j)\right)^{1/2}}.$$

Schoenfeld (1977) proposes the same type of statistic in the standard goodness-of-fit testing context. Under $H_0$ and the assumptions in previous sections, we have that

$$\widehat{\psi}_{q,N} \stackrel{d}{\to} N(0,1) \quad \text{as} \quad N \to \infty \text{ for each fixed } q.$$

Also, arguing as in Schoenfeld's (1977) Theorem 3, it can be shown the convergence in distribution of $\widehat{\psi}_{q_N,N}$ when $q_N$ increases with $N$. Approximately optimal tests for $H_0$ in the direction of $H_{1N}$ reject $H_0$ at the $\alpha$ significance level when $\left|\widehat{\psi}_{q_N,N}\right| > z_{1-\alpha/2}$ if $\tau$ has unknown sign, $\widehat{\psi}_{q_N,N} > z_{1-\alpha}$ when $\tau > 0$ and $\widehat{\psi}_{q_N,N} < -z_{1-\alpha}$ when $\tau < 0$.

## 6. LEMMAS

First, we introduce some notation. We denote the conjugate of a complex number $a$ by $\overline{a}$. Also, for a generic function $\nu(\lambda)$, we abbreviate $\nu(\lambda_j)$ by $\nu_j = \left(\nu_{j[1]}, ..., \nu_{j[d]}\right)'$ and $C$ will denote a generic positive and finite constant.

For the next two lemmas, we shall assume that $\{\xi(t)\}_{t\in\mathbb{Z}^d}$ and $\{\zeta(t)\}_{t\in\mathbb{Z}^d}$ are two stationary spatial processes with a representation as that in (1.1) and whose respective errors satisfy $C1$. Also $f_{\xi\zeta}(\lambda) = (2\pi)^{-d}\sum_{j\in\mathbb{Z}^d}\mathbb{E}(\xi(t)\zeta(t+j))\exp\{-ij\cdot\lambda\}$, the cross-spectral density function, is a twice continuously differentiable function in $\lambda \in \Pi^d$. Denote $\widetilde{\mathbb{Z}}^d = \{j : (-\tilde{n} \prec j \prec \tilde{n}) \wedge (0 < j[1])\}$.

**Lemma 1.** *Consider $j \in \widetilde{\mathbb{Z}}^d$. Then,*

$$(a)\ \mathbb{E}\left(w_{\xi,j}^T\overline{w}_{\zeta,j}^T\right) - f_{\xi\zeta,j} = O\left(n^{-2}\right); \quad (b)\ \mathbb{E}\left(w_{\xi,j}^T w_{\zeta,j}^T\right) = O\left(\prod_{\ell=1}^d j[\ell]^{-3}\right).$$

*Proof.* We begin with part $(a)$. By definition, the left side of the equality in $(a)$ is

$$\int_{\Pi^d}\left(f_{\xi\zeta}(\lambda) - f_{\xi\zeta}(\lambda_j)\right)\prod_{\ell=1}^d K^T\left(\lambda[\ell] - \lambda_{j[\ell]}\right)d\lambda,$$

surprising any reference to $\ell$ in $K_\ell^T$ and/or $D_\ell^T$ for notational simplicity.

Now, because $f_{\xi\zeta}(\lambda)$ is twice continuous differentiability and $\int_\Pi \mu K^T(\mu)\,d\mu = 0$, we have that the last displayed expression is bounded in modulus by

$$\begin{aligned} & C\int_{\Pi^d}\sum_{\ell=1}^d\sum_{p=1}^d\left|\lambda[\ell] - \lambda_{j[\ell]}\right|\left|\lambda[p] - \lambda_{j[p]}\right|\prod_{\ell=1}^d K^T\left(\lambda[\ell] - \lambda_{j[\ell]}\right)d\lambda \\ \leq\ & C\int_{\Pi^d}\sum_{\ell=1}^d\left|\lambda[\ell] - \lambda_{j[\ell]}\right|^2\prod_{\ell=1}^d K^T\left(\lambda[\ell] - \lambda_{j[\ell]}\right)d\lambda \end{aligned}$$

by the Cauchy-Schwarz inequality. Now, using (2.5), that the Fejer's kernel integrates 1, and that $\sum_{t[\ell]=1}^{n[\ell]} h_\ell(t[\ell])^2 \geq Cn[\ell]$, we obtain that the right side of the last displayed inequality is bounded by

$$\frac{C}{N}\int_{\Pi^d}\sum_{\ell=1}^d\left|\lambda[\ell] - \lambda_{j[\ell]}\right|^2\prod_{\ell=1}^d\min\left\{n^2, n^{-4}\left|\lambda[\ell] - \lambda_{j[\ell]}\right|^{-6}\right\}d\lambda = O\left(n^{-2}\right)$$

by $C2$ and standard algebra.

Next we show part $(b)$. Again by definition and that $|f_{\xi\zeta}(\lambda)| < C$, we obtain that $\left|\mathbb{E}\left(w_{\xi,j}^T w_{\zeta,j}^T\right)\right|$ is bounded by

$$C\int_{\Pi^d}|f_{\xi\zeta}(\lambda)|\prod_{\ell=1}^d n^{-1}\left|D^T\left(\lambda[\ell] - \lambda_{j[\ell]}\right)D^T\left(\lambda[\ell] + \lambda_{j[\ell]}\right)\right|d\lambda \leq C\prod_{\ell=1}^d j[\ell]^{-3}$$

by standard arguments after using (2.5). □

**Lemma 2.** *Let* $k \prec j \in \widetilde{\mathbb{Z}}^d$ *and* $c_{jk} = \min\left\{\prod_{\ell=1}^{d} |j[\ell] - k[\ell]|_+^{-3}, \frac{\log n}{n}\right\}$, *where* $|q|_+ = \max\{1, |q|\}$. *Then,*

$$(a) \quad \mathbb{E}\left(w_{\xi,j}^T \overline{w}_{\zeta,k}^T\right) = \left(4^{d_2}/6^d\right) f_{\xi\zeta,k} + O\left(c_{jk}\right) \qquad (b) \quad \mathbb{E}\left(w_{\xi,j}^T w_{\zeta,k}^T\right) = O\left(c_{jk}\right),$$

*where* $d_2$ *is given in the proof.*

*Proof.* We shall handle part $(a)$ only, being part $(b)$ identical. By definition,

$$\mathbb{E}\left(w_{\xi,j}^T \overline{w}_{\zeta,k}^T\right) = \int_{\Pi^d} f_{\xi\zeta}(\lambda) \prod_{\ell=1}^{d} n[\ell]^{-1} D^T\left(\lambda[\ell] - \lambda_{j[\ell]}\right) D^T\left(\lambda_{k[\ell]} - \lambda[\ell]\right) d\lambda. \tag{6.1}$$

Because $|f_{\xi\zeta}(\lambda)| < C$, the modulus of the right side of (6.1) is bounded by

$$C \prod_{\ell=1}^{d} n^{-1} \left\{\int_{-\pi}^{\lambda_{(j[\ell]+k[\ell])/2}} + \int_{\lambda_{(j[\ell]+k[\ell])/2}}^{\pi}\right\} \left|D^T\left(\lambda[\ell] - \lambda_{j[\ell]}\right)\right| \left|D^T\left(\lambda_{k[\ell]} - \lambda[\ell]\right)\right| d\lambda[\ell]$$

using $C2$. Now using (2.5), the contribution due to a factor of the type $\int_{-\pi}^{\lambda_{(j[\ell]+k[\ell])/2}}$ is bounded by

$$C\,|j[\ell] - k[\ell]|_+^{-3} \int_{-\pi}^{\lambda_{(j[\ell]+k[\ell])/2}} \left|D^T\left(\lambda_{k[\ell]} - \lambda[\ell]\right)\right| d\lambda[\ell] = O\left(|j[\ell] - k[\ell]|_+^{-3}\right),$$

because $\int_0^\pi \left|D^T(\lambda)\right| d\lambda < C$ if $\lambda_{k[\ell]} < \lambda_{j[\ell]}$ and if $\lambda_{j[\ell]} < \lambda_{k[\ell]}$ by $C\,|j[\ell] - k[\ell]|_+^{-3}$ $\int_{-\pi}^{\lambda_{(j[\ell]+k[\ell])/2}} \left|D^T\left(\lambda[\ell] - \lambda_{j[\ell]}\right)\right| d\lambda[\ell]$, which is also $O\left(|j[\ell] - k[\ell]|_+^{-3}\right)$. Recall that $k \prec j$ we can have that for some $\ell = 1, .., d-1$, $\lambda_{j[\ell]} < \lambda_{k[\ell]}$. Finally, proceeding similarly the contribution due to a factor of the type $\int_{\lambda_{(j[\ell]+k[\ell])/2}[\ell]}^{\pi}$ is also $O\left(|j[\ell] - k[\ell]|_+^{-3}\right)$. Now conclude by Holder's inequality that $\left|\mathbb{E}\left(w_{\xi,j}^T \overline{w}_{\zeta,k}^T\right)\right| = O\left(\prod_{\ell=1}^{d} |j[\ell] - k[\ell]|_+^{-3}\right)$.

On the other hand, using (2.6) we have that the right side of (6.1), except multiplicative constants, is

$$\begin{aligned} &\int_{\Pi^d} f_{\xi\zeta}(\lambda) \prod_{\ell=1}^{d} n^{-1} \left\{D\left(\lambda[\ell] - \lambda_{j[\ell]-1}\right) - 2D\left(\lambda[\ell] - \lambda_{j[\ell]}\right) + D\left(\lambda[\ell] - \lambda_{j[\ell]+1}\right)\right\} \\ &\quad \times \left\{D\left(\lambda_{k[\ell]-1} - \lambda[\ell]\right) - 2D\left(\lambda_{k[\ell]} - \lambda[\ell]\right) + D\left(\lambda_{k[\ell]+1} - \lambda[\ell]\right)\right\} d\lambda. \end{aligned} \tag{6.2}$$

We shall distinguish two cases. $(i)$ when for all $\ell = 1, ..., d$, $|k[\ell] - j[\ell]| < 2$ and $(ii)$ otherwise. Let us examine when $(i)$ holds. Because for $\iota[\ell] \neq \tau[\ell]$,

$$\int_{\Pi} D\left(\lambda[\ell] - \lambda_{\iota[\ell]}\right) D\left(\lambda_{\tau[\ell]} - \lambda[\ell]\right) d\lambda[\ell] = 0,$$

we obtain that a typical term in (6.2), except multiplicative constants, is

$$N^{-1} \int_{\Pi^d} \left(f_{\xi\zeta}(\lambda) - f_{\xi\zeta}(\lambda_j)\right) \prod_{\ell=1}^{d} D\left(\lambda[\ell] - \lambda_{j[\ell]}\right) D\left(\lambda_{k[\ell]} - \lambda[\ell]\right) d\lambda,$$

which, by the mean value theorem, is bounded in absolute value by

$$N^{-1} \int_{\Pi^d} \sum_{\ell=1}^{d} \left|\lambda[\ell] - \lambda_{j[\ell]}\right| \prod_{\ell=1}^{d} \left|D\left(\lambda[\ell] - \lambda_{j[\ell]}\right)\right| \left|D\left(\lambda_{k[\ell]} - \lambda[\ell]\right)\right| d\lambda = O\left(\frac{\log n}{n}\right),$$

because $|\lambda D(\lambda)| < C$, $\int_0^\pi |D(\lambda)| \, d\lambda = O(\log n)$ and the Cauchy-Schwarz inequality imply that $\int_\Pi \left|D\left(\lambda - \lambda_{j[\ell]}\right)\right| \left|D\left(\lambda_{k[\ell]} - \lambda\right)\right| d\lambda = O(n)$.

Next, we examine case $(ii)$. By (2.6), $\prod_{\ell=1}^d D^T\left(\lambda[\ell] - \lambda_{j[\ell]}\right) D^T\left(\lambda_{k[\ell]} - \lambda[\ell]\right)$ is, except constants,

$$
\begin{aligned}
&\prod_{\ell=1}^{d_1} K^T\left(\lambda[\ell] - \lambda_{k[\ell]}\right) \\
&\times \prod_{\ell=d_1+1}^{d_1+d_2} \left\{D\left(\lambda[\ell] - \lambda_{k[\ell]}\right) - 2D\left(\lambda[\ell] - \lambda_{k[\ell]+1}\right) + D\left(\lambda[\ell] - \lambda_{k[\ell]+2}\right)\right\} \\
&\times \left\{D\left(\lambda_{k[\ell]-1} - \lambda[\ell]\right) - 2D\left(\lambda_{k[\ell]} - \lambda[\ell]\right) + D\left(\lambda_{k[\ell]+1} - \lambda[\ell]\right)\right\} \\
&\times \prod_{\ell=d_1+d_2+1}^{d_1+d_2+d_3} \left\{D\left(\lambda[\ell] - \lambda_{k[\ell]+1}\right) - 2D\left(\lambda[\ell] - \lambda_{k[\ell]+2}\right) + D\left(\lambda[\ell] - \lambda_{k[\ell]+3}\right)\right\} \\
&\times \left\{D\left(\lambda_{k[\ell]-1} - \lambda[\ell]\right) - 2D\left(\lambda_{k[\ell]} - \lambda[\ell]\right) + D\left(\lambda_{k[\ell]+1} - \lambda[\ell]\right)\right\} \\
&\times \prod_{\ell=d_1+d_2+d_3+1}^{d} \left\{D\left(\lambda[\ell] - \lambda_{j[\ell]-1}\right) - 2D\left(\lambda[\ell] - \lambda_{j[\ell]}\right) + D\left(\lambda[\ell] - \lambda_{j[\ell]+1}\right)\right\} \\
&\times \left\{D\left(\lambda_{k[\ell]-1} - \lambda[\ell]\right) - 2D\left(\lambda_{k[\ell]} - \lambda[\ell]\right) + D\left(\lambda_{k[\ell]+1} - \lambda[\ell]\right)\right\},
\end{aligned}
$$

where we have assumed without loss of generality that the first $d_1$ "coordinates" are such that $k[\ell] = j[\ell]$, the next $d_2$ are those such that $|k[\ell] - j[\ell]| = 1$, the next $d_3$ are such that $|k[\ell] - j[\ell]| = 2$ and the remaining ones satisfy that $|k[\ell] - j[\ell]| > 2$.

Now, proceeding as with $(i)$, if $d_1+d_2+d_3 < d$, we have that (6.2) is $O\left(n^{-1}\log n\right)$. So, it suffices to examine the situation for which $d_1 + d_2 + d_3 = d$. It is easy to observe that we only need to examine terms of the type

$$
\begin{aligned}
&\prod_{\ell=1}^{d_1} K^T\left(\lambda[\ell] - \lambda_{k[\ell]}\right) \\
&\times 2^{d_2} \prod_{\ell=d_1+1}^{d_1+d_2} \left\{D\left(\lambda[\ell] - \lambda_{k[\ell]}\right) D\left(\lambda_{k[\ell]} - \lambda[\ell]\right) + D\left(\lambda[\ell] - \lambda_{k[\ell]+1}\right) D\left(\lambda_{k[\ell]+1} - \lambda[\ell]\right)\right\} \\
&\times \prod_{\ell=d_1+d_2+1}^{d} D\left(\lambda[\ell] - \lambda_{k[\ell]+1}\right) D\left(\lambda_{k[\ell]+1} - \lambda[\ell]\right),
\end{aligned}
$$

where for simplicity we take $k[\ell] < j[\ell]$. Now recalling that $n^{-1}\left|D\left(\lambda[\ell] - \lambda_{k[\ell]}\right)\right|^2 \asymp K\left(\lambda[\ell] - \lambda_{k[\ell]}\right)$, (6.2) is $2^{d_2}/6^d$ times

$$
\begin{aligned}
&\int_{\Pi^d} f_{\xi\zeta}(\lambda) \left\{\prod_{\ell=1}^{d_1} K^T\left(\lambda[\ell] - \lambda_{k[\ell]}\right) \prod_{\ell=d_1+1}^{d_1+d_2} \left\{K\left(\lambda[\ell] - \lambda_{k[\ell]}\right) + K\left(\lambda[\ell] - \lambda_{k[\ell]+1}\right)\right\}\right. \\
&\left.\prod_{\ell=d_1+d_2+1}^{d} K\left(\lambda[\ell] - \lambda_{k[\ell]+1}\right)\right\} d\lambda + O\left(\frac{1}{n}\right) = 2^{d_2} f_{\xi\zeta,k} + O\left(n^{-1}\right)
\end{aligned}
$$

since by assumption $f_{\xi\zeta,k} - f_{\xi\zeta,k+1} = O\left(n^{-1}\right)$ and then proceeding as with the proof of Lemma 1 part $(a)$. This concludes the proof of the lemma. □

In what follows, we shall abbreviate $w_x^T(\lambda)/\Psi(\lambda)$ and $w_\varepsilon^T(\lambda)$ by $u(\lambda)$ and $v(\lambda)$ respectively for all $\lambda \in \Pi^d$.

**Lemma 3.** *Let $\zeta(\lambda)$ be a continuous differentiable function in $\Pi^d$. Under C1-C4, we have that for all $r \leq s \in \widetilde{\mathbb{Z}}^d$*

$$\mathbb{E}\left|\sum_{j=r}^{s} \zeta_j v_j (u_j - v_j)\right|^2 = O\left(\frac{\log n}{n^{2/3}} \prod_{\ell=1}^{d} |s[\ell] - r[\ell]|_+\right). \tag{6.3}$$

*Proof.* Denote $w_j = u_j - v_j$. By standard arguments, the left side of (6.3) is

$$\begin{aligned} & \sum_{j=r}^{s} \zeta_j^2 \mathbb{E}\{v_j \overline{v}_j \overline{w}_j w_j\} + \sum_{j\neq k=r}^{s} \zeta_j \zeta_k \mathbb{E}\{v_j \overline{v}_k \overline{w}_j w_k\} \\ = & \sum_{j=r}^{s} \zeta_j^2 \{a_{j1} + a_{j2}\} + \sum_{j\neq k=r}^{s} \zeta_j \zeta_k \{b_{jk,1} + b_{jk,2}\}, \end{aligned}$$

where

$$\begin{aligned} a_{j1} &= \mathbb{E}(v_j \overline{v}_j) \mathbb{E}(w_j \overline{w}_j) + |\mathbb{E}(v_j \overline{w}_j)|^2 + |\mathbb{E}(v_j w_j)|^2 \\ a_{j2} &= cum(v_j, \overline{v}_j, \overline{u}_j, u_j) + cum(v_j, \overline{v}_j, \overline{v}_j, v_j) \\ &\quad - cum(v_j, \overline{v}_j, \overline{u}_j, v_j) - cum(v_j, \overline{v}_j, u_j, \overline{v}_j) \\ b_{jk,1} &= \mathbb{E}(v_j \overline{v}_k) \mathbb{E}(\overline{w}_j w_k) + \mathbb{E}(v_j \overline{w}_j) \mathbb{E}(\overline{v}_k w_k) + \mathbb{E}(v_j w_k) \mathbb{E}(\overline{v}_k \overline{w}_j) \\ b_{jk,2} &= cum(v_j, \overline{v}_k, \overline{u}_j, u_k) + cum(v_j, \overline{v}_k, \overline{v}_j, v_k) \\ &\quad - cum(v_j, \overline{v}_k, \overline{u}_j, v_k) - cum(v_j, \overline{v}_k, u_j, \overline{v}_k). \end{aligned}$$

After observing that $\mathbb{E}(v_j \overline{u}_j) = 1 + O(n^{-2})$ and that $\mathbb{E}(v_j \overline{w}_j) = \mathbb{E}(v_j \overline{u}_j) - \mathbb{E}(v_j \overline{v}_j)$ by Lemma 1, we have that $a_{j1} = O(n^{-2})$, whereas Lemma 2 implies that $b_{jk,1} = O(c_{jk})$, with $c_{jk}$ as defined there.

Finally we examine $a_{j2}$ and $b_{jk,2}$. Using formulae in Brillinger [(1975), (2.6.3), page 26, and (2.10.3), page 39], we deduce after standard algebra that

$$\begin{aligned} b_{jk,2} = & \frac{\kappa}{N^2} \int_{\Pi^d} \int_{\Pi^d} \left(\frac{\Psi(\lambda)}{\Psi_j} - 1\right)\left(\frac{\Psi(\mu)}{\Psi_k} - 1\right) D^T(\lambda - \lambda_j) D^T(\mu + \lambda_k) \\ & \times D^T(\lambda_j - \lambda_k - \lambda - \mu) d\lambda d\mu. \end{aligned}$$

By the Cauchy-Schwarz inequality, we have that $|b_{jk,2}|^2$ is bounded by $CN^{-1}$ times

$$\int_{\Pi^d} \left(\frac{\Psi(\lambda)}{\Psi_j} - 1\right)^2 K^T(\lambda - \lambda_j) d\lambda \int_{\Pi^d} \left(\frac{\Psi(\mu)}{\Psi_k} - 1\right)^2 K^T(\mu + \lambda_k) K^T(\lambda_j - \lambda_k - \lambda - \mu) d\lambda d\mu.$$

Proceeding as in Lemma 2 and by $C4$, we then obtain that $b_{jk,2} = O(n^{-2} N^{-1/2})$. Likewise $a_{j2} = O(n^{-2} N^{-1/2})$. From here, the conclusion of the lemma easily follows by observing that $\prod_{\ell=1}^{d} |s[\ell] - r[\ell]|_+ \leq N$. □

**Lemma 4.** *Let $\zeta(\lambda)$ be a function as in Lemma 3. Then, under C1-C4,*

$$\mathbb{E} \sup_{\lambda \in [0,\pi]^d} \left| \sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]} \zeta_j \left\{ \frac{I_{x,j}^T}{|\Psi_j|^2} - I_{\varepsilon,j}^T \right\} \right| = o\left(N^{1/2}\right).$$

*Proof.* We shall consider the proof in the positive quadrant $\sum_{j=1}^{[\tilde{n}\lambda/\pi]}$, being the proof for the remaining $2^{d-1} - 1$ quadrants similarly handled. By Markov's inequality and the triangle inequality, it suffices to show that

$$\mathbb{E} \sup_s \left| \sum_{j=1}^{s} \zeta_j \left\{ \frac{I_{x,j}^T}{|\Psi_j|^2} - I_{\varepsilon,j}^T \right\} \right| \leq \mathbb{E} \sup_s \sum_{j=1}^{s} |\zeta_j| |u_j - v_j|^2 + 2\mathbb{E} \sup_s \left| \sum_{j=1}^{s} \zeta_j v_j (\overline{u}_j - \overline{v}_j) \right| \tag{6.4}$$

is $o\left(N^{1/2}\right)$, where we abbreviate "$\sup_{s=1,\ldots,\tilde{n}}$" by "$\sup_s$".

The first term on the right of (6.4) is bounded by

$$C\sum_{j=1}^{\tilde{n}}\left\{\left(\mathbb{E}\left|u_j\right|^2-1\right)-\left(\mathbb{E}\left(u_j\overline{v}_j\right)-1\right)-\left(\mathbb{E}\left(\overline{u}_jv_j\right)-1\right)+\left(\mathbb{E}\left|v_j\right|^2-1\right)\right\}=o\left(N^{1/2}\right),$$

because $\left|\zeta_j\right|\leq C$, $d<4$ and by Lemma 1, say,

$$\left|\mathbb{E}\left(u_j\left\{\begin{array}{c}\overline{v}_j\\ \overline{u}_j\end{array}\right)-1\right|\leq\sigma_{\varepsilon}^{-2}\left|\Psi_j\right|^{-2}\left(\left|\mathbb{E}\left(w_{x,j}\left\{\begin{array}{c}\overline{w}_{\varepsilon,j}\\ \overline{w}_{x,j}\end{array}\right)-\sigma_{\varepsilon}^2\left\{\begin{array}{c}\Psi_j\\ \left|\Psi_j\right|^2\end{array}\right|\right)=O\left(n^{-2}\right).$$

Next, we examine the second term of (6.4). To that end, let $q=0,\ldots,\left[\tilde{n}^{\varsigma}\right]-1$ for some $0<\varsigma<2/\left(3d\right)$. (Recall that $\left[\tilde{n}^{\psi}\right]=\left(\left[\tilde{n}^{\psi}\left[1\right]\right],\ldots\left[\tilde{n}^{\psi}\left[d\right]\right]\right)$ for any $\psi>0$.) Using the inequality

$$\left(\sup_p\left|c_p\right|\right)^2=\sup_p\left|c_p\right|^2\leq\sum_p\left|c_p\right|^2, \tag{6.5}$$

the triangle inequality implies that the square of the second term on the right of (6.4) is bounded by

$$\mathbb{E}\max_s\left|\left\{\sum_{j=1}^{s}-\sum_{j=1}^{q(s)\tilde{n}/\left[\tilde{n}^{\varsigma}\right]}\right\}\zeta_jv_j\left(\overline{u}_j-\overline{v}_j\right)\right|^2+\mathbb{E}\max_s\left|\sum_{j=1}^{q(s)\tilde{n}/\left[\tilde{n}^{\varsigma}\right]}\zeta_jv_j\left(\overline{u}_j-\overline{v}_j\right)\right|^2, \tag{6.6}$$

where $q(s)$ denotes the value of $q=0,\ldots,\left[\tilde{n}^{\varsigma}\right]-1$ such that $q(s)\tilde{n}/\left[\tilde{n}^{\varsigma}\right]$ is the largest vector $s_1$ such that $s_1\leq s$, and using the convention $\sum_{j=c}^{d}\equiv 0$ if $d<c$. Herewith we denote $\left(\tilde{n}\left[1\right]/\left[\tilde{n}^{\varsigma}\left[1\right]\right],\ldots,\tilde{n}\left[d\right]/\left[\tilde{n}^{\varsigma}\left[d\right]\right]\right)$ by $\tilde{n}/\left[\tilde{n}^{\varsigma}\right]$.

From the definition of $q\left(s\right)$ and (6.5), the second term of (6.6) is bounded by

$$\sum_{q=1}^{\left[\tilde{n}^{\varsigma}\right]-1}\mathbb{E}\left|\sum_{j=1}^{q\tilde{n}/\left[\tilde{n}^{\varsigma}\right]}\zeta_jv_j\left(\overline{u}_j-\overline{v}_j\right)\right|^2=O\left(N^{1+\varsigma}n^{-2/3}\log^2 n\right)$$

which is $o\left(N\right)$ by Lemma 3 and because $\varsigma<2/\left(3d\right)$. To complete the proof we need to show that the first term in (6.6) is $o\left(N\right)$. To that end, we note that it is bounded by

$$\mathbb{E}\max_{q=1,\ldots,\left[\tilde{n}^{\varsigma}\right]-1}\max_{s=1+q\tilde{n}/\left[\tilde{n}^{\varsigma}\right],\ldots,(q+1)\tilde{n}/\left[\tilde{n}^{\varsigma}\right]}\left|\sum_{j=1+q\tilde{n}/\left[\tilde{n}^{\varsigma}\right]}^{s}\zeta_jv_j\left(\overline{u}_j-\overline{v}_j\right)\right|^2$$

which is $O\left(N^{\varsigma}\right)\mathbb{E}\max_{s=1,\ldots,\tilde{n}/\left[\tilde{n}^{\varsigma}\right]}\left|\sum_{j=1}^{s}\zeta_jv_j\left(\overline{u}_j-\overline{v}_j\right)\right|^2$.

So, we have that the square of the second term on the right of (6.4) is

$$O\left(N^{1+\varsigma}n^{-2/3}\right)+O\left(N^{\varsigma}\right)\mathbb{E}\max_{s=1,\ldots,\tilde{n}/\left[\tilde{n}^{\varsigma}\right]}\left|\sum_{j=1}^{s}\zeta_jv_j\left(\overline{u}_j-\overline{v}_j\right)\right|^2.$$

Observe that the second factor of the second term of the last displayed expression is similar to the second term on the right of (6.4) but with $s=1,\ldots,\tilde{n}/\left[\tilde{n}^{\varsigma}\right]$ instead of $s=1,\ldots,\tilde{n}$. So, repeating the same steps, the last displayed expression, so is the square of the second term on the right of (6.4), is

$$O\left(\left\{N^{1+\varsigma}+\sum_{q=0}^{\iota-1}N^{(1+\varsigma)(1-\varsigma)^{q+1}+\varsigma\sum_{p=0}^{q}(1-\varsigma)^p}\right\}n^{-2}\right)$$

$$+N^{\varsigma\sum_{p=0}^{q}(1-\varsigma)^p}\mathbb{E}\max_{s=1,\ldots,(\tilde{n}/[\tilde{n}^{\varsigma}])^{\iota}}\left|\sum_{j=1}^{s}\zeta_j v_j\left(\overline{u}_j-\overline{v}_j\right)\right|^2$$

$$= O\left(N^{1+\varsigma}n^{-2/3}\right)+O\left(N\right)\sum_{s=1}^{\tilde{n}^{(1-\varsigma)^{\iota}}}\mathbb{E}\left|\sum_{j=1}^{s}\zeta_j v_j\left(\overline{u}_j-\overline{v}_j\right)\right|^2=o\left(N\right)$$

after choosing $\iota$ large enough because $\varsigma<2/(3d)$. This completes the proof. □

**Lemma 5.** *Let $\zeta(\lambda;\vartheta)$ be is as in Lemma 3 for all $\vartheta\in\Theta\times\mathbb{R}^+$, and continuously differentiable in $\vartheta$ for all $\lambda$. Assuming $C1-C4$, then,*

$$\frac{1}{N}\sup_{\vartheta\in\Theta\times\mathbb{R}^+}\left|\sum_{j=-\tilde{n}}^{\tilde{n}}\zeta_j\left(\vartheta\right)\left(\frac{I_{x,j}^T}{\left|\Psi_{\theta_0,j}\right|^2}-1\right)\right|=o_p\left(1\right).$$

*Proof.* By the triangle inequality, the left side of the last displayed equality is bounded by

(6.7)

$$\frac{C}{N}\sup_{\vartheta\in\Theta\times\mathbb{R}^+}\left|\sum_{j=-\tilde{n}}^{\tilde{n}}\zeta_j\left(\vartheta\right)\left(\frac{I_{x,j}^T}{\left|\Psi_{\theta_0,j}\right|^2}-I_{\varepsilon,j}^T\right)\right|+\frac{C}{N}\sup_{\vartheta\in\Theta\times\mathbb{R}^+}\left|\sum_{j=-\tilde{n}}^{\tilde{n}}\zeta_j\left(\vartheta\right)\left(I_{\varepsilon,j}^T-1\right)\right|.$$

Now, because by assumption $\left|\zeta\left(\lambda;\vartheta\right)\right|<C$, the first term of (6.7) is bounded by

$$\frac{C}{N}\sum_{j=-\tilde{n}}^{\tilde{n}}\left|\frac{I_{x,j}^T}{\left|\Psi_{\theta_0,j}\right|^2}-I_{\varepsilon,j}^T\right|\leq\frac{C}{N}\sum_{j=-\tilde{n}}^{\tilde{n}}\left|u_j-v_j\right|^2+\frac{C}{N}\sum_{j=-\tilde{n}}^{\tilde{n}}\left|v_j\left(\overline{u}_j-\overline{v}_j\right)\right|=o_p\left(1\right)$$

by Markov's inequality because by the Cauchy-Schwarz inequality $\mathbb{E}\left|v_j\left(\overline{u}_j-\overline{v}_j\right)\right|^2\leq\mathbb{E}\left|v_j\right|^2\mathbb{E}\left|\overline{u}_j-\overline{v}_j\right|^2$ and then proceeding as in Lemma 3. Next, we show that the second term of (6.7) is $o_p(1)$. But this follows by standard arguments because $\zeta_j(\vartheta)$ is continuously differentiable in $\vartheta$, and thus it is omitted. □

**Lemma 6.** *Assume $C1-C3$ and $C5-C8$. Then,*

$$\widehat{\vartheta}-\vartheta_0=o_p\left(1\right).$$

*Proof.* The proof follows very easily using Lemma 5. Indeed,

$$\frac{1}{N}\sum_{j=-\tilde{n}}^{\tilde{n}}\left\{\log f_{\vartheta,j}+\frac{I_{x,j}^T}{(2\pi)^d f_{\vartheta,j}}\right\} = \frac{1}{N}\sum_{j=-\tilde{n}}^{\tilde{n}}\frac{f_{\vartheta_0,j}}{f_{\vartheta,j}}\left(\frac{I_{x,j}^T}{(2\pi)^d f_{\vartheta_0,j}}-1\right)$$

$$+\frac{1}{N}\left\{\sum_{j=-\tilde{n}}^{\tilde{n}}\frac{f_{\vartheta_0,j}}{f_{\vartheta,j}}-\log\frac{f_{\vartheta_0,j}}{f_{\vartheta,j}}+\log f_{\vartheta_0,j}\right\}.$$

However the second term on the right converges using Brillinger (1981, *p*.15) to

$$\int_{\pi}^{\pi}\left\{\frac{f_{\vartheta_0}(\lambda)}{f_{\vartheta}(\lambda)}-\log\left(\frac{f_{\vartheta_0}(\lambda)}{f_{\vartheta}(\lambda)}\right)\right\}d\lambda+\int_{\pi}^{\pi}f_{\vartheta_0}(\lambda)\,d\lambda\geq\frac{(2\pi)^d}{2}+\int_{\pi}^{\pi}f_{\vartheta_0}(\lambda)\,d\lambda$$

with equality when $f_{\vartheta_0}(\lambda)=f_{\vartheta}(\lambda)$ which is the case only if $\vartheta=\vartheta_0$ by $C7$. On the other hand, the first term converges to zero uniformly in $\vartheta$ by Lemma 5 because $f_{\vartheta,j}^{-1}f_{\vartheta_0,j}$ satisfies the same conditions as $\zeta(\lambda;\vartheta)$ there by $C6$. From here the conclusion of the lemma is standard and so we omit its details. □

**Lemma 7.** *Assume $C1 - C8$. Under $H_0$,*

$$\frac{1}{N^{1/2}} \sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]} \zeta_j \left( \frac{I^T_{x,j}}{\left|\Psi_{\widehat{\theta},j}\right|^2} - 1 \right) = \frac{1}{N^{1/2}} \sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]} \zeta_j \left( I^T_{\varepsilon,j} - 1 \right)$$

$$- \left( \frac{1}{N^{1/2}} \sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]} \zeta_j \varphi'_{\theta_0,j} \right) N^{1/2} \left( \widehat{\theta} - \theta_0 \right) + o_p(1), \tag{6.8}$$

*where the $o_p(1)$ is uniform in $\lambda \in [0,\pi]^d$ and where $\zeta(\lambda)$ is as in Lemma 3.*

*Proof.* The difference between the left side of (6.8) and the first term on its right side is

(6.9)

$$\frac{1}{N^{1/2}} \sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]} \zeta_j \frac{I^T_{x,j}}{\left|\Psi_{\theta_0,j}\right|^2} \left[ \frac{\left|\Psi_{\theta_0,j}\right|^2}{\left|\Psi_{\widehat{\theta},j}\right|^2} - 1 + \phi'_{\theta_0,j} \left( \widehat{\theta} - \theta_0 \right) \right]$$

$$+ \frac{1}{N^{1/2}} \sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]} \zeta_j \left( \frac{I^T_{x,j}}{\left|\Psi_{\theta_0,j}\right|^2} - I^T_{\varepsilon,j} \right) - \frac{1}{N^{1/2}} \sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]} \zeta_j \phi'_{\theta_0,j} \frac{I^T_{x,j}}{\left|\Psi_{\theta_0,j}\right|^2} \left( \widehat{\theta} - \theta_0 \right).$$

First, because each component of the vector $\zeta(\lambda)\phi_{\theta_0}(\lambda)$ satisfies the same conditions of $\zeta(\lambda)$ in Lemma 3, Markov's inequality implies that

$$\frac{1}{N^{1/2}} \sum_{j=-\tilde{n}}^{\tilde{n}} \zeta_j \phi_{\theta_0,j} \left( \frac{I^T_{x,j}}{\left|\Psi_{\theta_0,j}\right|^2} - I^T_{\varepsilon,j} \right) = o_p(1), \tag{6.10}$$

whereas proceeding as in Robinson and Vidal-Sanz (2006),

$$\frac{1}{N^{1/2}} \sum_{j=-\tilde{n}}^{\tilde{n}} \zeta_j \phi_{\theta_0,j} \left( I^T_{\varepsilon,j} - 1 \right) = O_p(1). \tag{6.11}$$

Hence the third term of (6.9) is $N^{-1} \sum_{j=1}^{[\tilde{n}\lambda/\pi]} \zeta_j \phi'_{\theta_0,j} N^{1/2} \left( \widehat{\theta} - \theta_0 \right) + o_p(1)$ using $(6.10)-(6.11)$, whereas the second term of (6.9) is $o_p(1)$ by Lemma 4 and Markov's inequality. Finally by the mean value theorem, the norm of the first term of (6.9) is bounded by

$$CN^{1/2} \left\| \widehat{\theta} - \theta_0 \right\|^2 \frac{1}{N^{1/2}} \sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]} \frac{I^T_{x,j}}{\left|\Psi_{\theta_0,j}\right|^2} = O_p\left( N^{-1/2} \right), \tag{6.12}$$

by Theorem 2 and $(6.10) - (6.11)$. This concludes the proof. □

We now introduce the following notation. For $v_1 < v_2 \in [0,\pi]^d$,

$$\mathcal{E}_{1,N}(v_1, v_2) = \left( \frac{1}{N} \sum_{j=[\tilde{n}v_1/\pi]}^{[\tilde{n}v_2/\pi]} \zeta_j \right) \left( \frac{N^{1/2}}{\sum_{t=1}^{n} h^2(t)} \sum_{t=1}^{n} \left( \varepsilon^T(t)^2 - 1 \right) \right) \tag{6.13}$$

$$\mathcal{E}_{2,N}(v_1, v_2) = \frac{1}{N} \sum_{j=[\tilde{n}v_1/\pi]}^{[\tilde{n}v_2/\pi]} \zeta_j \frac{N^{1/2}}{\sum_{t=1}^{n} h^2(t)} \sum_{t_1 \neq t_2 = 1}^{n} \varepsilon^T(t_1) \varepsilon^T(t_2) e^{i(t_1 - t_2)\cdot\lambda_j}, \tag{6.14}$$

where by definition $\mathcal{E}_{1,N}(v_1, v_2) + \mathcal{E}_{2,N}(v_1, v_2) = N^{-1/2} \sum_{j=[\tilde{n}v_1/\pi]+1}^{[\tilde{n}v_2/\pi]} \zeta_j \left( I^T_{\varepsilon,j} - 1 \right)$ and $\varepsilon^T(t) = h(t)\varepsilon(t)$.

**Lemma 8.** *Let $v_1 < v < v_2 \in [0,\pi]^d$. Then, assuming $C1 - C3$, and for some $\beta > 0$, with $\zeta(\lambda)$ as in Lemma 3,*

$$\mathbb{E}\left(|\mathcal{E}_{j,N}(v_1,v)|^{\beta}|\mathcal{E}_{j,N}(v,v_2)|^{\beta}\right) \le C\prod_{\ell=1}^{d}(v_2[\ell]-v_1[\ell])^2,\ j=1,2. \tag{6.15}$$

*Proof.* The proof follows proceeding as that of Lemma 6 of Delgado et al. (2005) and observing that because $\zeta(\lambda)$ is a continuous function, then $\left|N^{-1}\sum_{p=[\tilde{n}v_1/\pi]+1}^{[\tilde{n}v_2/\pi]}\zeta_p^q\right| \le C\prod_{\ell=1}^{d}(v_2[\ell]-v_1[\ell])$ for any $q \ge 1$. □

Next we will show that the processes $\left(\prod_{\ell=1}^{d}\lambda[\ell]^{-\upsilon}\right)(\mathcal{E}_{1,N}(0,\lambda)$ and $\mathcal{E}_{2,N}(0,\lambda))$ are tight for some values of $\upsilon > 0$. From Bickel and Wichura (1971) it suffices to show the following lemma.

**Lemma 9.** *Assuming $C1$, for any $0 \le \upsilon < 1/4$, we have that*

(6.16)

$$(a)\ \mathbb{E}\prod_{\ell=1}^{d}\left(\frac{\mathcal{E}_{1,N}^{(\ell)}(0,\lambda_{1[\ell]})}{\lambda_{1[\ell]}^{\upsilon}}-\frac{\mathcal{E}_{1,N}^{(\ell)}(0,\lambda_{2[\ell]})}{\lambda_{2[\ell]}^{\upsilon}}\right)^2 \le C\prod_{\ell=1}^{d}\left(\lambda_{2[\ell]}-\lambda_{1[\ell]}\right)^{2-2\upsilon}$$

(6.17)

$$(b)\ \mathbb{E}\prod_{\ell=1}^{d}\left(\frac{\mathcal{E}_{2,N}^{(\ell)}(0,\lambda_{1[\ell]})}{\lambda_{1[\ell]}^{\upsilon}}-\frac{\mathcal{E}_{2,N}^{(\ell)}(0,\lambda_{2[\ell]})}{\lambda_{2[\ell]}^{\upsilon}}\right)^4 \le C\prod_{\ell=1}^{d}\left(\lambda_{2[\ell]}-\lambda_{1[\ell]}\right)^{2-4\upsilon}$$

*for all $\lambda_{1[\ell]} < \lambda_{2[\ell]} \in [0,\pi]$, $\ell = 1,...,d$, and where, say,*

$$\mathcal{E}_{2,N}^{(\ell)}\left(\lambda_{1[\ell]},\lambda_{2[\ell]}\right) = \frac{1}{n}\sum_{j[\ell]=\left[\tilde{n}\lambda_{1[\ell]}/\pi\right]}^{\left[\tilde{n}\lambda_{2[\ell]}/\pi\right]}\zeta_j\frac{N^{1/2}}{\sum_{t=1}^{n}h^2(t)}\sum_{t_1\neq t_2=1}^{n}\varepsilon^T(t_1)\,\varepsilon^T(t_2)\,e^{i(t_1-t_2)\cdot\omega_j}$$

$$\mathcal{E}_{1,N}^{(\ell)}\left(\lambda_{1[\ell]},\lambda_{2[\ell]}\right) = \left(\frac{1}{n}\sum_{j[\ell]=\left[\tilde{n}\lambda_{1[\ell]}/\pi\right]}^{\left[\tilde{n}\lambda_{2[\ell]}/\pi\right]}\zeta_j\right)\left(\frac{N^{1/2}}{\sum_{t=1}^{n}h^2(t)}\sum_{t=1}^{n}\left(\varepsilon^T(t)^2-1\right)\right).$$

*Proof.* We begin with $(b)$. By standard inequalities and then by Lemma 8, the left side of (6.17) is bounded by

$$C\mathbb{E}\left(\prod_{\ell=1}^{d}\frac{\mathcal{E}_{2,N}^{(\ell)}\left(\lambda_{1[\ell]},\lambda_{2[\ell]}\right)}{\lambda_{2[\ell]}^{\upsilon}}\right)^4 + C\prod_{\ell=1}^{d}\left(\frac{1}{\lambda_{1[\ell]}^{\upsilon}}-\frac{1}{\lambda_{2[\ell]}^{\upsilon}}\right)^4\mathbb{E}\left(\prod_{\ell=1}^{d}\mathcal{E}_{2,N}^{(\ell)}\left(0,\lambda_{1[\ell]}\right)\right)^4$$

$$\le C\prod_{\ell=1}^{d}\frac{\left(\lambda_{2[\ell]}-\lambda_{1[\ell]}\right)^2}{\lambda_{2[\ell]}^{4\upsilon}} + C\prod_{\ell=1}^{d}\left(\frac{1}{\lambda_{1[\ell]}^{\upsilon}}-\frac{1}{\lambda_{2[\ell]}^{\upsilon}}\right)^4\lambda_{1[\ell]}^2. \tag{6.18}$$

Consider the case that $\lambda_{2[\ell]} - \lambda_{1[\ell]} \le 2^{-1}\lambda_{1[\ell]}$ first. By mean value theorem, a typical factor on the right of (6.18) is

$$\begin{aligned} & C\frac{\left(\lambda_{2[\ell]}-\lambda_{1[\ell]}\right)^2}{\lambda_{2[\ell]}^{4\upsilon}} + \frac{C}{\lambda_{2[\ell]}^{4\upsilon}\lambda_{1[\ell]}^{4\upsilon-2}}\frac{\upsilon^4\left(\lambda_{2[\ell]}-\lambda_{1[\ell]}\right)^4}{\left(\beta\lambda_{1[\ell]}+(1-\beta)\lambda_{2[\ell]}\right)^{4-4\upsilon}} \\ \le\ & C\left\{\left(\lambda_{2[\ell]}-\lambda_{1[\ell]}\right)^{2-4\upsilon} + \lambda_{1[\ell]}^{-4\upsilon-2}\left(\lambda_{2[\ell]}-\lambda_{1[\ell]}\right)^4\right\}\end{aligned}$$

where $\beta = \beta\left(\lambda_{1[\ell]},\lambda_{2[\ell]}\right) \in (0,1)$, and then because $\lambda_{2[\ell]} > \lambda_{[\ell]1} \ge 2\left(\lambda_{2[\ell]}-\lambda_{1[\ell]}\right)$. But the right side of the last displayed inequality is bounded by $C\left(\lambda_{2[\ell]}-\lambda_{1[\ell]}\right)^{2-4\upsilon}$ since $\lambda_{2[\ell]} - \lambda_{1[\ell]} \le 2^{-1}\lambda_{1[\ell]}$.

Next, consider the case for which $2^{-1}\lambda_{1[\ell]} < \lambda_{2[\ell]} - \lambda_{1[\ell]}$. Using the inequality $a^{\varsigma} - b^{\varsigma} \leq (a-b)^{\varsigma}$ for any $0 < \varsigma < 1$ and $a \geq b$, we have that (6.18) is bounded by

$$C\prod_{\ell=1}^{d}\left\{\left(\lambda_{2[\ell]} - \lambda_{1[\ell]}\right)^{2-4\upsilon} + \frac{\left(\lambda_{2[\ell]} - \lambda_{1[\ell]}\right)^{4\upsilon}\lambda_{1[\ell]}^{2}}{\lambda_{1[\ell]}^{4\upsilon}\lambda_{2[\ell]}^{4\upsilon}}\right\} \leq C\prod_{\ell=1}^{d}\left(\lambda_2[\ell] - \lambda_{1[\ell]}\right)^{2-4\upsilon},$$

where we have used that $\lambda_{2[\ell]} - \lambda_{1[\ell]} \leq \lambda_{2[\ell]}$ and $\lambda_{1[\ell]} < 2\left(\lambda_2[\ell] - \lambda_{1[\ell]}\right)$. This completes the proof of part $(b)$.

Next part $(a)$. By definition and $C1$, the left side of (6.16) is bounded by

$$\begin{aligned}&\frac{C}{\prod_{\ell=1}^{d}\lambda_{2[\ell]}^{2\upsilon}}\left(\frac{1}{N}\sum_{j=[\tilde{n}\lambda_1/\pi]}^{[\tilde{n}\lambda_2/\pi]}\zeta_j\right)^2 + C\prod_{\ell=1}^{d}\left(\frac{\lambda_{2[\ell]}^{\upsilon} - \lambda_{1[\ell]}^{\upsilon}}{\lambda_1^{\upsilon}[\ell]\,\lambda_{2[\ell]}^{\upsilon}}\right)^2\left(\frac{1}{N}\sum_{j=1}^{[\tilde{n}\lambda_1/\pi]}\zeta_j\right)^2\\ \leq\ & C\prod_{\ell=1}^{d}\left(\lambda_{2[\ell]} - \lambda_{1[\ell]}\right)^{2-2\upsilon}\end{aligned}$$

by continuity of $\zeta(\lambda)$ and proceeding as in part $(b)$. □

In what follows $H_{\theta,N}(k) = \widetilde{\varphi}_{\theta,k}' A_{\theta,N}^{-1}(k)$ and $\overline{k} = \#\{j \prec k\}$. Also, $\lambda_0 = \left(0, -\pi, , ..., -\pi, \mathring{\lambda}\right)$.

**Lemma 10.** *Assuming $C1-C3, C5-C10$, for all $\varepsilon > 0$,*

(6.19)
$$\lim_{\mathring{\lambda}\to-\pi}\limsup_{N\to\infty}\Pr\left\{\sup_{\lambda\preceq\lambda_0}\left|\frac{1}{N}\sum_{k=[\tilde{n}\lambda/\pi];p^*\prec k}^{[\tilde{n}\lambda_0/\pi]}\frac{H_{\theta_0,N}(k)}{N^{1/2}}\sum_{j\prec k}\widetilde{\varphi}_{\theta_0,j}\left(\frac{I_{x,j}^T}{\left|\Psi_{\widehat{\theta},j}\right|^2} - 1\right)\right| > \varepsilon\right\} = 0.$$

*Proof.* Denote $\left|\Psi_{\widehat{\theta},j}\right|^{-2} I_{x,j}^T - I_{\varepsilon,j}^T$ by $\varkappa_j$ and let $\mathring{\lambda} < -\pi/2$ without loss of generality. Because $\left|\Psi_{\widehat{\theta},j}\right|^{-2} I_{x,j}^T - 1 = \varkappa_j + \eta_j$, where $\eta_j = I_{\varepsilon,j}^T - 1$, we have that

(6.20)
$$\begin{aligned}&\sup_{\lambda\preceq\lambda_0}\left|\frac{1}{N}\sum_{k=[\tilde{n}\lambda/\pi];p^*\prec k}^{[\tilde{n}\lambda_0/\pi]}\frac{H_{\theta_0,N}(k)}{N^{1/2}}\sum_{j\prec k}\widetilde{\varphi}_{\theta_0,j}\left(\varkappa_j + \eta_j\right)\right|\\ &\leq \frac{C}{N}\sum_{k=[\tilde{n}\lambda/\pi];p^*\prec k}^{[\tilde{n}\lambda_0/\pi]}\left\|H_{\theta_0,N}(k)\right\|\left(1+\frac{k}{\tilde{n}}\right)^{\frac{\delta}{2}}\left\{\sup_{p^*\prec k\leq[\tilde{n}\lambda_0/\pi]}\left\|\frac{\left(1+\frac{k}{\tilde{n}}\right)^{-\frac{\delta}{2}}}{N^{1/2}}\sum_{j\prec k}\widetilde{\varphi}_{\theta_0,j}\varkappa_j\right\|\right.\\ &\left.+\sup_{p^*\prec k\leq[\tilde{n}\lambda_0/\pi]}\left\|\frac{\left(1+\frac{k}{\tilde{n}}\right)^{-\frac{\delta}{2}}}{N^{1/2}}\sum_{j\prec k}\widetilde{\varphi}_{\theta_0,j}\eta_j\right\|\right\},\end{aligned}$$

for any $0 < \delta < 1$. The first factor on the right of (6.20) is bounded by

$$C\left|\frac{1}{n^{\delta/2}}\sum_{k=-\tilde{n}}^{[\tilde{n}\lambda_0/\pi]}\left\|\widetilde{\varphi}_{\theta_0,k}\right\|\left(\overline{k}+\tilde{n}\right)^{\frac{\delta}{2}-1}\right| \leq C\left(\pi + \mathring{\lambda}\right)^{\delta/2},$$

using that $\left\|A_{\theta_0,N}^{-1}(k)\right\| \leq C\left(\overline{k}/N\right)^{-1}$, because $\left\|A_{\theta_0}(\lambda)\right\| \geq C\prod_{\ell=1}^{d}\lambda[\ell]$ by $C9$ and because continuity of $\varphi_\theta(\lambda)$ implies that $\sup_{p^*\prec k\leq[\tilde{n}\lambda_0/\pi]}\left\|A_{\theta_0,N}(k) - A_{\theta_0}([k\pi/\tilde{n}])\right\| = O\left(n^{-1}\right)$.

Next, by Lemma 9, the second term inside the braces on the right of (6.20) is $O_p(1)$ for $\delta > 0$ small enough, whereas Lemma 7, Theorem 2 and continuity of

$\varphi_{\theta_0}(\lambda)$ imply that the first term is bounded by

$$O_p\left(\sup_{-\tilde{n}<k\le[\tilde{n}\lambda_0/\pi]}\left\|\left(1+\frac{k}{\tilde{n}}\right)^{1-\frac{\delta}{2}}\right\|\right)+O_p\left(\sup_{-\tilde{n}<k\le[\tilde{n}\lambda_0/\pi]}\frac{\left(1+\frac{k}{\tilde{n}}\right)^{-\frac{\delta}{2}}}{\tilde{n}^{\delta}}\right)=O_p\left(\left(\pi+\mathring{\lambda}\right)^{\delta/2}\right)$$

because $\tilde{n}^{-1}\le\inf_{-\tilde{n}<k\le[\tilde{n}\lambda_0/\pi]}(1+k/\tilde{n})$, $0<\delta<1$. So, (6.20) is $O_p\left(\left(\pi+\mathring{\lambda}\right)^{\delta}\right)$, which implies that (6.19) holds true because $\delta>0$. □

**Lemma 11.** *Assuming* $C1-C3, C5-C10$,

$$\sup_{\lambda\in[0,\pi]}\left|\frac{1}{N^{1/2}}\sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]}\left(\varphi_{\widehat{\theta},j}-\varphi_{\theta_0,j}\right)\left(\frac{I^T_{x,j}}{\left|\Psi_{\widehat{\theta},j}\right|^2}-1\right)\right|=O_p\left(\frac{1}{N^{1/2}}\right). \tag{6.21}$$

*Proof.* The expression inside the norm on the left of (6.21) is

$$\frac{1}{N^{1/2}}\sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]}\dot{\varphi}'_{\theta_0,j}\left(\frac{I^T_{x,j}}{\left|\Psi_{\widehat{\theta},j}\right|^2}-1\right)\left(\widehat{\theta}-\theta_0\right)$$

$$+\frac{1}{N^{1/2}}\sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]}\left(\varphi_{\widehat{\theta},j}-\varphi_{\theta_0,j}-\dot{\varphi}'_{\theta_0,j}\left(\widehat{\theta}-\theta_0\right)\right)\left(\frac{I^T_{x,j}}{\left|\Psi_{\widehat{\theta},j}\right|^2}-1\right). \tag{6.22}$$

By $C10$ and then noting that $|a-b|\le(a-b)+2b$ for $a>0$ and $b>0$, the norm of the second term of (6.22) is bounded by

$$C\frac{\left\|\widehat{\theta}-\theta_0\right\|^2}{N^{1/2}}\left\{\sum_{j=-\tilde{n}}^{\tilde{n}}\left(\frac{I^T_{x,j}}{\left|\Psi_{\widehat{\theta},j}\right|^2}-1\right)+2\right\}=O_p\left(\frac{1}{N^{1/2}}\right)$$

by Theorem 2, Proposition 1 and Lemma 7 with $\zeta(\lambda)=1$. So, uniformly in $\lambda$ the second term of (6.22) is $o_p(1)$. Likewise, the first term of (6.22) is $O_p\left(N^{-1/2}\right)$ uniformly in $\lambda$ using Lemma 7 with $\zeta(\lambda)=\dot{\varphi}_{\theta_0}(\lambda)$ and Theorem 2. □

**Lemma 12.** *Assuming* $C1-C3, C5-C10$, *for all* $\varepsilon>0$,

$$\lim_{\mathring{\lambda}\to-\pi}\limsup_{N\to\infty}\Pr\left\{\sup_{\lambda\preceq\lambda_0}\left|\frac{1}{N}\sum_{k=[\tilde{n}\lambda/\pi];p^*\prec k}^{[\tilde{n}\lambda_0/\pi]}\frac{H_{\widehat{\theta},N}(k)}{N^{1/2}}\sum_{j\prec k}\widetilde{\varphi}_{\widehat{\theta},j}\left(\frac{I^T_{x,j}}{\left|\Psi_{\widehat{\theta},j}\right|^2}-1\right)\right|>\varepsilon\right\}=0. \tag{6.23}$$

*Proof.* Notice that Theorem 2 implies that it suffices to show (6.23) in the set $\left\{\left\|\widehat{\theta}-\theta_0\right\|<CN^{-1/2}m_N^{-1}\right\}$, where $m_N+m_N^{-1}N^{-1/2}\to0$. On the other hand, Lemma 11 and then Lemma 7 imply that, uniformly in $k$,

$$\frac{1}{N^{1/2}}\sum_{j\prec k}\widetilde{\varphi}_{\widehat{\theta},j}\varkappa_j = \left(\frac{1}{N}\sum_{j\prec k}\widetilde{\varphi}_{\theta_0,j}\widetilde{\varphi}'_{\theta_0,j}\right)N^{1/2}\left(\theta_0-\widehat{\theta}\right)+o_p(1)$$

$$\frac{1}{N^{1/2}}\sum_{j\prec k}\widetilde{\varphi}_{\widehat{\theta},j}\eta_j = \frac{1}{N^{1/2}}\sum_{j\prec k}\widetilde{\varphi}_{\theta_0,j}\eta_j+O_p\left(N^{-1/2}\right) \tag{6.24}$$

proceeding as in the proof of (6.21) but with $\varkappa_j+\eta_j$ replaced by $\eta_j$ there. Observe that we can take $\mathring{\lambda}<-\pi/2$. Next, uniformly in $k$, $C9$ implies that

$$\sup_k\left(\frac{\overline{k}}{N}\right)^{-1}\left\|A_{\widehat{\theta},N}(k)-A_{\theta_0,N}(k)\right\|=O_p\left(\left\|\widehat{\theta}-\theta_0\right\|\right)$$

which will imply that, with probability approaching one, as $N \to \infty$,

$$\left\| A^{-1}_{\widehat{\theta},N}(k) \right\| \le \left\| A^{-1}_{\theta_0,N}(k) \right\| \left( 1 + CN^{-1/2} m_N^{-1} \right) \le C \left( \overline{k}/N \right)^{-1}.$$

So, we have that for $0 < \delta < 1/2$,

$$\text{(6.25)} \quad \sup_{\lambda \preceq \lambda_0} \left\| \frac{1}{N} \sum_{k=[\tilde{n}\lambda/\pi];p^* \prec k}^{[\tilde{n}\lambda_0/\pi]} \frac{H_{\widehat{\theta},N}(k)}{N^{1/2}} \sum_{j \prec k} \phi_{\widehat{\theta},j} \left( \frac{I^T_{x,j}}{\left| \Psi_{\widehat{\theta},j} \right|^2} - 1 \right) \right\|$$

$$\le C \sup_{\lambda \preceq \lambda_0} \left| \frac{1}{N} \sum_{k=-\tilde{n}}^{[\tilde{n}\lambda_0/\pi]} \left\| \phi_{\theta_0,k} \right\| \left( 1 + \frac{k}{\tilde{n}} \right)^{-1+\delta/2} \right|$$

$$\times \left\{ \sup_{-\tilde{n} < k \le [\tilde{n}\lambda_0/\pi]} \left\| \left( 1 + \frac{k}{\tilde{n}} \right)^{-\delta/2} \frac{1}{N^{1/2}} \sum_{j \prec k} \phi_{\theta_0,j} \eta_j \right\| + O_p \left( \left( \pi + \mathring{\lambda} \right)^{\delta/2} \right) \right\},$$

by (6.24) and because $\tilde{n}^{-1} \le \inf_{-\tilde{n} < k \le [n\lambda_0/\pi]} (1 + k/\tilde{n})$. But Lemma 8 implies that $\sup_{-\tilde{n} < k \le [n\lambda_0/\pi]} \left\| \left( 1 + \frac{k}{\tilde{n}} \right)^{-\delta/2} N^{-1/2} \sum_{j \prec k} \varphi_{\theta_0,j} \eta_j \right\| = O_p(1)$, and by $C4$,

$$\sup_{\lambda \preceq \lambda_0} \frac{1}{N} \sum_{k=-\tilde{n}}^{[\tilde{n}\lambda_0/\pi]} \left\| \phi_{\theta_0,k} \right\| \left( 1 + \frac{k}{\tilde{n}} \right)^{-1+\delta/2} \le C \left( \pi + \mathring{\lambda} \right)^{\delta/2},$$

and hence the left side of (6.25) is $O_p \left( \left( \pi + \mathring{\lambda} \right)^{\delta/2} \right)$. From here we conclude that (6.23) holds true because $\delta > 0$. □ □

## 7. PROOFS

### 7.1. Proof of Proposition 1.

We shall be a bit more general. Indeed, for a vector function $\zeta(\lambda)$ as in Lemma 3, we will show that

$$\boldsymbol{S}_N(\mu) = \frac{1}{N} \sum_{j=-[\tilde{n}\mu/\pi]}^{[\tilde{n}\mu/\pi]} \zeta_j \left( I^T_{\varepsilon,j} - 1 \right) \Rightarrow \boldsymbol{B}_\zeta(\mu), \qquad \mu \in [0,\pi]^d$$

where for $\mu \le v \in [0,\pi]^d$, $Cov\left( \boldsymbol{B}_\zeta(\mu), \boldsymbol{B}_\zeta(v) \right) = \left( 2 + \kappa \left( \frac{35}{32} \right)^d \right) \int_{-\mu}^{\mu} \zeta(\lambda) \zeta'(\lambda) \, d\lambda$

To that end, it suffices to show that $(a)$ for all $\mu \in [0,\pi]^d$,

$$\text{(7.1)} \quad \frac{1}{N^{1/2}} \sum_{j=-[\tilde{n}\mu/\pi]}^{[\tilde{n}\mu/\pi]} \zeta_j \left( I^T_{\varepsilon,j} - 1 \right) \xrightarrow{d} \mathcal{N} \left( 0, \left( 2 + \kappa \left( \frac{35}{32} \right)^d \right) \int_{-\mu}^{\mu} \zeta(\lambda) \zeta'(\lambda) \, d\lambda \right).$$

$(b)$ The covariance structure satisfies

$$\frac{1}{N} E \left\{ \sum_{j=-[\tilde{n}\mu/\pi]}^{[\tilde{n}\mu/\pi]} \zeta_j \left( I^T_{\varepsilon,j} - 1 \right) \sum_{j=-[\tilde{n}v/\pi]}^{[\tilde{n}v/\pi]} \zeta_j \left( I^T_{\varepsilon,j} - 1 \right) \right\} \to \left( 2 + \kappa \left( \frac{35}{32} \right)^d \right) \int_{-\mu}^{\mu} \zeta(\lambda) \zeta'(\lambda) \, d\lambda$$

for $\mu \le v$, and $(c)$ the process $\left\{ \boldsymbol{S}_N(\mu) : \mu \in [0,\pi]^d \right\}$ is tight.

We begin with $(a)$. Its proof follows directly by that in Robinson and Vidal-Sanz (2006) and observing that because $\zeta(\lambda)$ is continuously differentiable, then by Brillinger (1981, $p$.15), $N^{-1} \sum_{j=-[\tilde{n}v/\pi]}^{[\tilde{n}v/\pi]} \zeta_j \zeta'_j - (2\pi)^{-d} \int_{-v}^{v} \zeta(\lambda) \zeta(\lambda) \, d\lambda = O_p\left( n^{-1} \right)$ and thus it is omitted.

Part $(b)$ follows after observing that $\mathbb{E}I^T_{\varepsilon,j}=1$ by $C1$ and $\mathbb{E}\left(I^T_{\varepsilon,j}I^T_{\varepsilon,k}\right)$ is

$$\begin{aligned}&\frac{1}{\left(\sum_{t=1}^{n}h^2(t)\right)^2}\sum_{t,s,r,u=1}^{n}\mathbb{E}\left\{\varepsilon^T(t)\,\varepsilon^T(s)\,\varepsilon^T(r)\,\varepsilon^T(s)\right\}e^{-i(t-s)\cdot\lambda_j+i(r-u)\cdot\lambda_k}\\ =\;& 2\mathcal{I}(j-k=0,n)+\mathcal{I}(j+k=0,n)+\frac{\kappa}{N}\left(2+\kappa\left(\frac{35}{32}\right)^d\right)\end{aligned}$$

using that, say $\sum_{p[\ell]=1}^{n[\ell]}h_\ell(t[\ell])\,e^{-ip[\ell]\left(\lambda_{j[\ell]\pm k[\ell]}\right)}=n[\ell]\,\mathcal{I}(j[\ell]\pm k[\ell]=0,n[\ell])$ and that by Brillinger (1981, p.15) we have that $N^{-1}\sum_{t=1}^{\tilde n}h^p(t)\to\int_{[0,1]^d}h^p(u)\,du$ for all $p\geq 0$. Finally, part $(c)$ follows by Lemma 9 with $v=0$ there.

### 7.2. **Proof of Theorem 3.**

Part $(a)$. By Lemma 7 with $\zeta(\lambda)=1$ there and the definitions of $G_{\theta,N}(\lambda)$ and $G^0_N(\lambda)$, we have that

$$\begin{aligned}N^{1/2}\left(G_{\widehat\theta,N}(\lambda)-G^0_N(\lambda)\right)&=-\left(\frac{1}{N}\sum_{j=-[\tilde n\lambda/\pi]}^{[\tilde n\lambda/\pi]}\widetilde\varphi'_{\theta_0,N}(j)\right)N^{1/2}\left(\widehat\theta-\theta_0\right)+o_p(1)\\ &=-\left(\frac{1}{N}\sum_{j=-[\tilde n\lambda/\pi]}^{[\tilde n\lambda/\pi]}\widetilde\varphi'_{\theta_0,N}(j)\right)\widetilde\Lambda^{-1}_{\theta_0,N}\frac{1}{N^{1/2}}\sum_{k=-\tilde n}^{\tilde n}\widetilde\varphi_{\theta_0,N}(k)\frac{I^T_{x,k}}{\left|\Psi_{\theta_0,k}\right|^2}\qquad(7.2)\\ &\quad+o_p(1)\,,\end{aligned}$$

by Theorem 2, where the $o_p(1)$ is uniform in $\lambda\in[0,\pi]^d$, where $\widetilde\varphi_{\theta,N}(k)=\widetilde\varphi_{\theta,N}(\lambda_k)$.

But using (2.14) and that $\left|G_{\theta_0,N}(\pi)-G^0_N(\pi)\right|=o_p\left(N^{-1/2}\right)$ by Lemma 4, then by (7.2) and (6.10), uniformly in $\lambda$ we obtain that

$$\begin{aligned}\alpha_{\widehat\theta,N}(\lambda)&=\alpha^0_N(\lambda)+\frac{N^{1/2}\left(G_{\widehat\theta,N}(\lambda)-G^0_N(\lambda)\right)}{G^0_N(\pi)}\\ &\quad+G_{\widehat\theta,N}(\lambda)\,N^{1/2}\left(\frac{1}{G_{\widehat\theta,N}(\pi)}-\frac{1}{G^0_N(\pi)}\right)\qquad(7.3)\\ &=\alpha^0_N(\lambda)-\frac{1}{N}\sum_{j=-[\tilde n\lambda/\pi]}^{[\tilde n\lambda/\pi]}\left(\widetilde\varphi'_{\theta_0,N}(j)\,\widetilde\Lambda^{-1}_{\theta_0,N}\frac{1}{N^{1/2}}\sum_{k=-\tilde n}^{\tilde n}\widetilde\varphi_{\theta_0,N}(k)\,I^T_{\varepsilon,k}\right)+o_p(1)\,,\end{aligned}$$

which concludes the proof of part $(a)$, because $G^0_N(\pi)-1=o_p(1)$.

Next part $(b)$. Taking into account part $(a)$, part $(b)$ follows because Theorem 2 guarantees the fidi's convergence of $\alpha^0_N$ and its tightness. Tightness of the second term on the right of (7.3) follows by (6.11) and because $\varphi_{\theta_0}(u)$ is continuously differentiable. This concludes the proof. □

### 7.3. **Proof of Theorem 4.** Using (2.14), we obtain that

(7.4)
$$\beta^0_N(\lambda)=\frac{1}{N^{1/2}}\sum_{j=-[\tilde n\lambda/\pi];p^*\prec j}^{[\tilde n\lambda/\pi]}\left(\left(I^T_{\varepsilon,j}-1\right)-H_{\theta_0,N}(j)\,\frac{1}{N}\sum_{k\prec j}\widetilde\varphi_{\theta_0,N}(k)\left(I^T_{\varepsilon,k}-1\right)\right)+o_p(1)\,,$$

where the $o_p(1)$ is uniform in $\lambda\in[0,\pi]^d$. Recall that $H_{\theta,N}(j)=\widetilde\varphi'_{\theta,N}(j)\,A^{-1}_{\theta,N}(j)$.

Suppose, to be shown later, that the convergence in $\lambda_0\preceq\lambda$ holds true for any $\mathring\lambda\in(-\pi,\pi)$, where $\lambda_0=\left(0,-\pi,,...,-\pi,\mathring\lambda\right)$. Then, because $\boldsymbol{B}(\lambda/\pi)$ and the limit of the process $N^{-1/2}\sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]}\left(I_{\varepsilon,j}^{T}-1\right)$ is continuous in $\widetilde{\Pi}$, Billingsley's (1968) Theorem 4.2 implies that it suffices to show that for all $\varepsilon>0$,

$$\lim_{\mathring{\lambda}\to-\pi}\limsup_{N\to\infty}\Pr\left\{\sup_{\lambda\preceq\lambda_0}\left|\frac{1}{N}\sum_{j=[\tilde{n}\lambda/\pi];p^*\prec j}^{[\tilde{n}\lambda_0/\pi]}\frac{H_{\theta_0,N}(j)}{N^{1/2}}\sum_{k\prec j}\widetilde{\varphi}_{\theta_0,N}(k)\left(I_{\varepsilon,k}^{T}-1\right)\right|>\varepsilon\right\}=0,$$

which follows by Lemma 10, cf. the second term on the right of (6.20).

So, to complete the proof we need to show that, for any $\mathring{\lambda}\in(-\pi,\pi)$,

(7.5)
$$\frac{1}{N^{1/2}}\sum_{j=-[\tilde{n}\lambda/\pi];p^*\prec j}^{[\tilde{n}\lambda/\pi]}\left(\left(I_{\varepsilon,j}^{T}-1\right)-H_{\theta_0,N}(j)\frac{1}{N}\sum_{k\prec j}\widetilde{\varphi}_{\theta_0,N}(k)\left(I_{\varepsilon,k}^{T}-1\right)\right)\Rightarrow\frac{1}{\pi^d}\boldsymbol{B}\left(\frac{\lambda}{\pi}\right),$$

in $\lambda_0\preceq\lambda$. Fidi's convergence follows by Theorem 2 after we note that the second term on the right of (7.4) is

$$\frac{1}{N^{1/2}}\sum_{k=-\tilde{n}}^{\tilde{n}}\left(\frac{1}{N}\sum_{j=-[\tilde{n}\lambda/\pi]}^{k\wedge[\tilde{n}\lambda/\pi]}H_{\theta_0,N}(j)\right)\widetilde{\varphi}_{\theta_0,N}(k)\left(I_{\varepsilon,k}^{T}-1\right)$$

and $\left(N^{-1}\sum_{j=-[\tilde{n}\lambda/\pi]}^{k\wedge[\tilde{n}\lambda/\pi]}H_{\theta_0,N}(j)\right)\widetilde{\varphi}_{\theta_0,N}(k)$ satisfies the same conditions of $h_n(\lambda)$ in Giraitis, Hidalgo and Robinson's (2001) Theorem 4.2. Then, it suffices to prove tightness. Since $\alpha_N^0$ is tight, we only need to show the tightness condition of

$$F_N(\lambda)=\frac{1}{N}\sum_{j=0}^{[\tilde{n}\lambda/\pi]}H_{\theta_0,N}(j)\left(\frac{1}{N^{1/2}}\sum_{k\prec j}\widetilde{\varphi}_{\theta_0,N}(k)\left(I_{\varepsilon,k}^{T}-1\right)\right). \tag{7.6}$$

Denoting $F_N^{(\ell)}(\lambda[\ell])=n^{-1}\sum_{j[\ell]=0}^{[\tilde{n}\lambda[\ell]/\pi]}H_{\theta_0,N}(j)\left(\frac{1}{N^{1/2}}\sum_{k\prec j}\widetilde{\varphi}_{\theta_0,N}(k)\left(I_{\varepsilon,k}^{T}-1\right)\right)$, by Bickel and Wichura (1968), it suffices to show that

(7.7)
$$\mathbb{E}\left(\prod_{\ell=1}^{d}\left|F_N^{(\ell)}(\vartheta[\ell])-F_N^{(\ell)}(\mu[\ell])\right|\left|F_N^{(\ell)}(\lambda[\ell])-F_N^{(\ell)}(\vartheta[\ell])\right|\right)\leq C\prod_{\ell=1}^{d}|\lambda[\ell]-\mu[\ell]|^2$$

for all $-\pi\leq\mu[\ell]<\vartheta[\ell]<\lambda[\ell]\leq\pi$. Observe that we can take $N^{-1}<\prod_{\ell=1}^{d}|\lambda[\ell]-\mu[\ell]|$ since otherwise the last inequality is trivial. Because $(\lambda[\ell]-\vartheta[\ell])(\vartheta[\ell]-\mu[\ell])<(\lambda[\ell]-\mu[\ell])^2$, by the Cauchy-Schwarz's inequality, it suffices to show that (7.7) holds for $\mathbb{E}\prod_{\ell=1}^{d}\left|F_N^{(\ell)}(\lambda[\ell])-F_N^{(\ell)}(\mu[\ell])\right|^2$ which is

$$\frac{1}{N^3}\sum_{j,k=[\tilde{n}\mu/\pi]}^{[\tilde{n}\lambda/\pi]}H_{\theta_0,N}(j)\sum_{\ell_1\prec j}\sum_{\ell_2\prec k}\widetilde{\varphi}_{\theta_0,N}(\ell_1)\widetilde{\varphi}'_{\theta_0,N}(\ell_2)\,\mathbb{E}\left[\left(I_{\varepsilon,\ell_1}^{T}-1\right)\left(I_{\varepsilon,\ell_2}^{T}-1\right)\right]H'_{\theta_0,N}(k)$$

$$\leq\frac{C}{N^2}\sum_{j,k=[\tilde{n}\mu/\pi]}^{[\tilde{n}\lambda/\pi]}\|H_{\theta_0,N}(j)\|\,\|H_{\theta_0,N}(k)\|\leq C\left(\left|\widetilde{H}(\lambda,\mu)\right|^2+N^{-2}\right),$$

where $\widetilde{H}(\lambda,\mu):=(2\pi)^{-d}\int_{\mu}^{\lambda}H_{\theta_0}(x)\,dx$ and $\left\|\widetilde{H}_N(\lambda,\mu)\right\|\leq CN^{-1}$, where $\widetilde{H}_N(\lambda,\mu):=N^{-1}\sum_{j=-[\tilde{n}\mu/\pi]}^{[\tilde{n}\lambda/\pi]}\|H_{\theta_0,N}(j)\|$. Now conclude by Bickel and Wichura (1971), because $\left|\widetilde{H}(\lambda,\mu)\right|\leq C\prod_{\ell=1}^{d}|\lambda[\ell]-\mu[\ell]|$ and $N^{-1}\leq\prod_{\ell=1}^{d}|\lambda[\ell]-\mu[\ell]|$.□

7.4. **Proof of Theorem 5.** By definition of $\beta_{\widehat{\theta},N}$ and $\beta_N^0$, it suffices to show that

(7.8)
$$\left|\frac{1}{N^{1/2}}\sum_{k=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]}\left(\frac{I_{x,j}^T}{\left|\Psi_{\widehat{\theta},k}\right|^2}-I_{\varepsilon,k}^T\right)-H_{\theta_0,N}(k)\frac{1}{N}\sum_{j\prec k}\phi_{\theta_0,N}(j)\left(\frac{I_{x,j}^T}{\left|\Psi_{\widehat{\theta},j}\right|^2}-I_{\varepsilon,j}^T\right)\right|$$

$$\frac{1}{G_{\widehat{\theta},N}(\pi)}\left\{\left(\frac{1}{N}\sum_{k=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]}H_{\theta_0,N}(k)\frac{1}{N^{1/2}}\sum_{j\prec k}\phi_{\theta_0,N}(j)\left(\frac{I_{x,j}^T}{\left|\Psi_{\widehat{\theta},j}\right|^2}-G_{\widehat{\theta},N}(\pi)\right)\right)\right.$$

$$(7.9)\quad\left.\left(\frac{1}{N}\sum_{k=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]}H_{\widehat{\theta},N}(k)\frac{1}{N^{1/2}}\sum_{j\prec k}\phi_{\widehat{\theta},N}(j)\left(\frac{I_{x,j}^T}{\left|\Psi_{\widehat{\theta},j}\right|^2}-G_{\widehat{\theta},N}(\pi)\right)\right)\right\}$$

converge to zero uniformly in $\lambda\in\widetilde{\Pi}$. Expression (7.8) is $o_p(1)$, uniformly in $\lambda\in\widetilde{\Pi}$, because the contribution due to the term in brackets in the last line of (7.3), that is $-\widetilde{\varphi}'_{\theta_0,N}(j)\left(\widetilde{\Lambda}_{\theta_0,N}^{-1}N^{-1/2}\sum_{k=-\tilde{n}}^{\tilde{n}}\widetilde{\varphi}_{\theta_0,N}(k)\,I_{\varepsilon,k}^T\right)$ is easily seen to be zero. Next, because

$$\frac{1}{N^2}\sum_{k=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]}\left\|\phi_{\theta_0,N}(k)\right\|\left\|A_{\theta_0,N}^{-1}(k)\right\|\sum_{j\prec k}\left\|\phi_{\theta_0,N}(j)\right\|$$
$$\leq\ C\frac{1}{N}\sum_{k=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]}\left\|\phi_{\theta_0,N}(k)\right\|\left\|A_{\theta_0,N}^{-1}(k)\left(\frac{\overline{k}}{N}\right)\right\|\leq C\frac{1}{N}\sum_{k=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]}\left\|\phi_{\theta_0,N}(k)\right\|\leq C$$

by integrability of $\phi_{\theta_0}(\lambda)$ and that $\left\|A_{\theta_0,N}(k)\left(\overline{k}/N\right)^{-1}\right\|>0$ by $C9$, it implies that the contribution into (7.8) due to the term $o_p(1)$ on the right of (7.3) is negligible.

Next we examine (7.9). Because $G_{\widehat{\theta},N}(\cdot)-G_N^0(\cdot)=\left(N^{-1/2}\right)$ and Remark 2, it suffices to show that

(7.10)
$$\frac{1}{N}\sum_{k=-[\tilde{n}\lambda/\pi];p^*\prec k}^{[\tilde{n}\lambda/\pi]}\left\{\frac{H_{\theta_0,N}(k)}{N^{1/2}}\sum_{j\prec k}\phi_{\theta_0,N}(j)\left(\frac{I_{x,j}^T}{\left|\Psi_{\widehat{\theta},j}\right|^2}-1\right)-\frac{H_{\widehat{\theta},N}(k)}{N^{1/2}}\sum_{j\prec k}\phi_{\widehat{\theta},N}(j)\left(\frac{I_{x,j}^T}{\left|\Psi_{\widehat{\theta},j}\right|^2}-1\right)\right\}$$

converges to zero uniformly in $\lambda\in\widetilde{\Pi}$, after observing that

$$\sup_{\lambda\in[0,\pi]^d}\left|\sum_{k=-[\tilde{n}\lambda/\pi];p^*\prec k}^{[\tilde{n}\lambda/\pi]}H_{\widehat{\theta},N}(k)\sum_{j\prec k}\varphi_{\widehat{\theta},N}(j)-\sum_{k=-[\tilde{n}\lambda/\pi];p^*\prec k}^{[\tilde{n}\lambda/\pi]}H_{\theta_0,N}(k)\sum_{j\prec k}\varphi_{\theta_0,N}(j)\right|=0.$$

First, we observe that Lemmas 9 and 11 imply that it suffices to show the uniform convergence in $\lambda_0\preceq\lambda$ for any $\mathring{\lambda}\in(-\pi,\pi)$. But (7.10) is equal to

(7.11)
$$\frac{1}{N}\sum_{k=-[\tilde{n}\lambda/\pi];p^*\prec k}^{[\tilde{n}\lambda/\pi]}H_{\widehat{\theta},N}(k)\frac{1}{N^{1/2}}\sum_{j\prec k}\left(\phi_{\theta_0,N}(j)-\phi_{\widehat{\theta},N}(j)\right)\left(\frac{I_{x,j}^T}{\left|\Psi_{\widehat{\theta},j}\right|^2}-1\right)$$

(7.12)
$$+\frac{1}{N}\sum_{k=-[\tilde{n}\lambda/\pi];p^*\prec k}^{[\tilde{n}\lambda/\pi]}\left(H_{\theta_0,N}(k)-H_{\widehat{\theta},N}(k)\right)\frac{1}{N^{1/2}}\sum_{j\prec k}\phi_{\theta_0,N}(j)\left(\frac{I_{x,j}^T}{\left|\Psi_{\widehat{\theta},j}\right|^2}-1\right).$$

So, the theorem follows if (7.11) and (7.12) are $o_p(1)$ uniformly in $\lambda_0 \preceq \lambda$.

To that end, we first show that

$$\sup_{\lambda_0 \preceq \lambda} \frac{1}{N} \sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]} \left\| \varphi_{\theta_0,j} - \varphi_{\widehat{\theta},j} \right\| = o_p(1), \tag{7.13}$$

$$\sup_{\lambda_0 \preceq \lambda} \left\| A^{-1}_{\theta_0,N}(\lambda) - A^{-1}_{\theta_0}(\lambda) \right\| = o(1), \tag{7.14}$$

$$\sup_{\lambda_0 \preceq \lambda} \left\| A^{-1}_{\widehat{\theta},N}(\lambda) - A^{-1}_{\theta_0,N}(\lambda) \right\| = o_p(1). \tag{7.15}$$

(7.13) follows proceeding as with the proof of (6.21) in Lemma 10 but without the factor $\left|\Psi_{\widehat{\theta},j}\right|^{-2} I^T_{x,j} - 1$, (7.14) follows because $C9$ implies that $A_{\theta_0}(\lambda_0) > 0$ and by $C6$, $\left\|\phi_{\theta_0}(\lambda)\phi'_{\theta_0}(\lambda)\right\|$ is continuously differentiable, so

$$\sup_{\lambda_0 \preceq \lambda} \left\| A_{\theta_0}(\lambda) - A_{\theta_0,N}(\lambda) \right\| = O\left(N^{-1}\right),$$

whereas (7.15) follows proceeding as with the proof of (7.13) and (7.14).

Now we show that (7.11) is $o_p(1)$ uniformly in $\lambda_0 \preceq \lambda$ which follows by Lemma 11 and (7.13) – (7.15) noting that $\left(\phi'_{\theta_0,j} - \phi'_{\widehat{\theta},j}\right) = \left(\sigma_\varepsilon^{-2} - \widehat{\sigma}_\varepsilon^{-2}, \varphi'_{\theta_0,j} - \varphi'_{\widehat{\theta},j}\right)$, so does (7.12) by (7.13) and (7.15) and that

$$\sup_{\lambda \in \widetilde{\Pi}} \left| \frac{1}{N^{1/2}} \sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]} \phi_{\theta_0,j} \left( \frac{I^T_{x,j}}{\left|\Psi_{\widehat{\theta},j}\right|^2} - 1 \right) \right| = O_p(1)$$

by Lemma 8 and Proposition 1with $\zeta(\lambda) = \gamma_{\theta_0}(\lambda)$ there, Theorem 2 and that $N^{-1} \sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]} \varphi_{\theta_0,j}\phi'_{\theta_0,j} \to (2\pi)^{-1} \int_{-\lambda}^{\lambda} \varphi_{\theta_0}(x)\phi'_{\theta_0}(x)\,dx$. □

## References


[1] Anderson, T.W. (1997): "Goodness-of-fit for autoregressive processes," *Journal of Time Series Analysis,* **18**, 321-339.

[2] Besag, J. (1982): "Spatial interaction and the statistical analysis of lattice systems (with discussion)," *Journal of the Royal Statistical Society, Ser. B.*, **36**, 192-236.

[3] Bickel, P.J. and Wichura, M.J. (1971): "Convergence criteria for multiparameter stochastic processes and its applications," *Annals of Mathematical Statistics*, **42**, 1656-1670.

[4] Billingsley, P. (1968): Convergence of probability measures. Wiley and Sons. New York.

[5] Bloomfield, P. (1973): "An exponential model for the spectrum of a scalar time series," *Biometrika*, **61**, 217-226.

[6] Box, G.E.P. and Pierce, D.A. (1970): "Distribution of residual autocorrelations in autoregressive-integrated moving average time series models," *Journal of the American Statistical Association,* **65**, 1509-1526.

[7] Brillinger, D.R. (1981): Time series, data analysis and theory. Holden-Day. San Francisco.

[8] Chen, H. and Romano J.P. (1999): "Bootstrap-assisted goodness-of-fit tests in the frequency domain," *Journal of Time Series Analysis,* **20,** 619-654.

[9] Brown, Durbin, J. and Evans, (1975): "Components of Cramér-von Mises statistics II," *Journal of the Royal Statistical Society. Ser. B.,* ***37,*** 216-237.

[10] Cressie, N.A.C. (1993): Statistics for spatial data. Wiley. New York.

[11] Crujeiras, R.M., Fernandez-Casal, R. and Gonzalez-manteiga, W. (2006): "Gooodness-of-fit tests for the spatial spectral density," *Preprint.* Universidad de Santiago de Compostela.

[12] Dahlhaus, R., and Künsch, H. (1986): "Edge effect and efficient parameter estimation for stationary random fields," *Biometrika*, **74**, 877-882.

[13] Delgado, M.A., Hidalgo, J. and Velasco, C. (2005): "Distribution free goodness-of-fit tests for linear processes," *Annals of Statistics*, ***33,*** *2568-2609.*

[14] Diblasi, A. and Bowman, A.W. (2001): "On the use of the variogram in checking the independence in spatial data," *Biometrics*, **57**, 211-218.

[15] Durbin, J. (1973): "Weak convergence of the sample distribution function when parameters are estimated," *Annals of Statistics*, 1, 274-290.

[16] Durbin, J., Knott, M. and Taylor, C.C. (1975): "Components of Cramér-von Mises statistics II," *Journal of the Royal Statistical Society, Ser. B*, **37**, 216-237.

[17] Eubank, R.L. and LaRicca, V. (1992): "Asymptotic comparison of Cramér-von Mises and nonparametric techniques for testing goodness-of-fit," *Annals of Statistics*, **20,** 2071-2086.

[18] Giraitis, L., Hidalgo, J. and Robinson, P.M. (2001): "Gaussian estimation of parametric spectral density with unknown pole," *Annals of Statistics*, **29**, 987-1023.

[19] Grenander, U. (1950): " Stochastic processes and statistical inference," *Ark. Mat.,* **1,** 195-277.

[20] Grenander, U. and Rosenblatt, M. (1957). *Statistical Analysis of Stationary Time Series.* John Wiley, New York.

[21] Guyon, X. (1982): "Parameter estimation for a stationary process on a $d$-dimension lattice," *Biometrika*, **69**, 95-105.

[22] Guyon, X. (1982): Random fields on a network. Springer-Verlag. Berlin.

[23] Hainz, G. and Dahlhaus, R. (2000): "Spectral domain bootstrap tests for stationary time series." Preprint.

[24] Hidalgo, J. and Kreiss, J.P. (2006): "Bootstrap specification tests for linear covariance stationary processes," *Journal of Econometrics*, **133**, 807-839.

[25] Hong, Y. (1996). "Consistent testing for serial-correlation of unknown form," *Econometrica*, **64,** 837-864.

[26] Khmaladze, E.V. (1981): "Martingale approach in the theory of Goodness-of-fit tests," *Theory of Probability and Applications,* **26**, 240-257.

[27] Khmaladze, E.V. and Koul, H. (2004): "Martingale transforms goodness-of-fit tests in regression models," *Annals of Statistics„* **32**, 995-1034.

[28] Koul, H.L. and Stute, W. (1999): "Nonparametric model checks for time series," *Annals of Statistics,* **27**, 204-236.

[29] Ljung, G.M. and Box G.E.P. (1978): "On a measure of lack of fit in time series models," *Biometrika*, **65**, 297-303.

[30] Neyman, J. (1937): "Smooth test for goodness of fit," *Skan. Aktuarietidskr.,* **20**, 150-199.

[31] Nikabadze, A. and Stute, W. (1997): "Model checks under random censorship," *Statistics and Probability Letters,* **32**, 249-259.

[32] Paparoditis, E. (2000): "Spectral density based goodness-of-fit tests for time series models," *Scandinavian Journal of Statistics,* **27**, 143-176.

[33] Robinson, P.M. and Vidal-Sanz, J. (2006): "Modified Whittle estimation of multilateral models on a lattice," *Journal of Multivariate Analysis*, **60**, 234-349.

[34] Schoenfeld, D.A. (1977): "Asymptotic properties of tests based on linear combinations of the Cramér-von Mises statistic," *Annals of Statistics*, **5,** 1017-1026.

[35] Stute, W. (1997): "Nonparametric model checks for regression," *Annals of Statistics*, **25**, 613-641.

[36] Stute, W., Thies, S. and Zhu, L. (1998): "Model checks for regression: An innovation process approach," *Annals of Statistics,* **26**, 1916-1934.

[37] Whittle, P. (1954): "On stationary processes in the plane," *Biometrika*, **41**, 434-449.

Economics Department, London School of Economics, Houghton Street, London WC2A 2AE, U.K.

*E-mail address*: f.j.hidalgo@lse.ac.uk