

GOODNESS OF FIT FOR LATTICE PROCESSES

JAVIER HIDALGO

ABSTRACT. The paper discusses tests for the correct specification of a model when data is observed in a d -dimensional lattice, extending 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 tests are functionals of a Gaussian sheet process, say $\mathbf{B}(\nu)$, $\nu \in [0, \pi]^d$. Because it is not easy to find a time transformation $h(\nu)$ such that $\mathbf{B}(h(\nu))$ becomes the standard Brownian sheet, a consequence is that the critical values are difficult, if at all possible, to obtain. So, to overcome the problem of its implementation, we propose to employ a bootstrap approach, showing its validity in our context.

JEL Classification: C21, C23.

1. INTRODUCTION

The paper is concerned with testing the goodness of fit of a parametric family of models for data collected in a lattice. More specifically, we are concerned with 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. The key idea of the test is to compare how close is the parametric and nonparametric fits of the data to provide support for the null hypothesis. In the paper, we shall specifically consider data for which $d \leq 3$. 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, or when $d = 3$. An important example of the latter is the spatial-temporal data sets, that is data collected in a lattice during a number of periods. However, we ought to mention that extensions to higher index lattice processes can be adapted under suitable modifications.

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

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

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 $t \neq 0$. Notice that because our model is multilateral, the sequence $\{\varepsilon(t)\}_{t \in \mathbb{Z}^d}$ loses its interpretation as the “prediction” error or that they can be regarded as innovations. Under (1.1), the spectral density function of $\{x(t)\}_{t \in \mathbb{Z}^d}$ can be factorized as

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

where $\Pi = (-\pi, \pi]$ and with

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

Date: 1 February 2008.

Key words and phrases. Goodness of fit tests. Spatial linear processes. Spectral domain. Bootstrap tests.

Henceforth the notation “ $j \cdot \lambda$ ” means the inner product of the d -dimensional vectors s and λ . The function $\Psi(\lambda)$ summarizes the covariogram structure of $\{x(t)\}_{t \in \mathbb{Z}^d}$, which is the main feature to obtain good and accurate prediction/extrapolation and/or interpolation (kriging) in the case of spatial data. Notice that the ultimate aim when modelling data is nothing but to predict the future.

The aim of the paper is then on testing whether the data support the null hypothesis that $\Psi(\lambda)$ belongs to a specific parametric family

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

where $\Theta \subset \mathbb{R}^p$ is a proper compact parameter set. That is, we are interested on the null hypothesis

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

The alternative hypothesis is the negation of H_0 . Alternatively we could have formulated the null hypothesis in terms of the covariogram given by $\{\gamma(s)\}_{s \in \mathbb{Z}^d}$, where $\gamma(s) = \text{Cov}(x(t), x(t+s))$. That is, the null hypothesis is that the covariogram follows a particular parametric family, say $\{\gamma(s)\}_{s \in \mathbb{Z}^d} = \{\gamma_\vartheta(s)\}_{s \in \mathbb{Z}^d}$, where from now on we denote $\vartheta = (\theta', \sigma_\varepsilon^2)'$. This is the case after observing that for any stationary spatial lattice process $\{x(t)\}_{t \in \mathbb{Z}^d}$, the spectral density $f(\lambda)$ and the covariogram $\gamma(s)$ are related through the expression

$$\gamma(s) = \begin{cases} \int_{\Pi^d} f(\lambda) e^{-is \cdot \lambda} d\lambda & ; \quad s = 0, \pm 1, \pm 2, \dots \\ \int_{\Pi^d} f_\vartheta(\lambda) e^{-is \cdot \lambda} d\lambda & \end{cases}$$

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 Π^d), the d -fold Cartesian product of the set \mathbb{Z} (or Π), is referred to as a multi-index of dimension d . Also, we shall write, say, $a = (a[1], \dots, a[d])$ with the square brackets used to denote the components of a .

Some particular parameterizations of (1.1) or (2.12), given in Condition C1 below, are the *ARMA* field model

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

where

$$P(z) = \sum_{j \in \mathbb{Z}^d} \alpha(j) z^j; \quad \alpha(0) = 1$$

$$Q(z) = \sum_{j \in \mathbb{Z}^d} \beta(j) z^j; \quad \beta(0) = 1,$$

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

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

whose spectral density function is

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

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 $\int_{\Pi^d} \log |P(e^{i\lambda})| d\lambda = 0$. So that in general, we can expect that $\int_{\Pi^d} \log |\Psi_\theta(e^{i\lambda})| d\lambda \neq 0$ for any admissible value of θ , see Guyon (1982b). The

latter and comments made after (1.1) become very relevant as when formulating the Whittle estimator $(\hat{\theta}', \hat{\sigma}_\varepsilon^2)'$ given in Section 3 below.

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 $d = 1$. See also Guyon (1982*b*).

Another parametric model of interest is the extension of the classical Bloomfield (1973) exponential model, see also Whittle's (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 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} , $\bar{\mathbb{Z}}^2 = \{(\ell[1], \ell[2]) \in \mathbb{Z}^d : (\ell[1] \leq 0 \wedge \ell[2] = 0) \vee (\ell[1] \leq 0 \wedge \ell[2] < 0)\}$. Observe that if we allowed ℓ in the last displayed equality to belong to \mathbb{Z}^d the model would not be then identified as $\cos(\ell \cdot \lambda) = \cos(-\ell \cdot \lambda)$.

When $d = 1$, the problem of testing a specific dependence structure of the data is very exhaustive and prominent. Different 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, contrary to a "time domain" approach based on the covariance/variogram structure of the data or that of the errors after the model has been fitted.

Our tests falls into the category of goodness of fit tests as we do not specify any particular alternative model or family. 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 for instance Hong (1996) or Paparoditis (2000) among others. However, the latter approach seems to be a mere 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 parameter(s) for which no relevant theory is available. That is, there are not rules available on how to choose the bandwidth parameter with empirical data. In fact, we might

face the strange situation that with the same data set two different practitioners might conclude differently or that if a practitioner chose to optimize the size of the test, that choice would lead to tests which have very poor power properties and viceversa. The latter is clearly not very appealing from both theoretical or applied stand 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 parameter. For that purpose, we rely on the periodogram which although it is not a consistent estimator for $f(\lambda)$, its integral is a consistent estimator of the spectral distribution function as the integral is the most natural smoothing algorithm.

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 θ_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 bootstrap test showing its validity. Section 5 gives the proof of a series of lemmas employed in the proof of our main results in Section 6.

2. TESTS WHEN THE PARAMETERS ARE KNOWN

This section discusses and examines how we can test the null hypothesis H_0 given in (1.4). That is,

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

when the ‘‘true’’ value of θ_0 is known, and where herewith $\tilde{\Pi}^d$ denotes $[0, \pi] \times [-\pi, \pi]^{d-1}$, that is $\lambda \in \tilde{\Pi}^d$ if $\lambda[1] \in [0, \pi]$ and $\lambda[\ell] \in [-\pi, \pi]$ for $\ell = 2, \dots, d$. Before we introduce and describe the test, we notice that the null hypothesis H_0 can be alternatively stated as

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

where

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

with the notation

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

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 $[0, \pi]$. The consequence is that the choice 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

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

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

$$I_v(\lambda) = \frac{1}{N} \left| \sum_{t=1}^n v(t) e^{-it \cdot \lambda} \right|^2; \quad \lambda \in \tilde{\Pi}^d$$

and similarly to the definition of \int_μ^λ , we are employing henceforth the notation

$$(2.3) \quad \sum_{j=[\tilde{n}\nu/\pi]}^{[\tilde{n}\mu/\pi]} = \sum_{j[1]=[\tilde{n}[1]\nu[1]/\pi]_+}^{[\tilde{n}[1]\mu[1]/\pi]} \sum_{j[2]=[\tilde{n}[2]\nu[2]/\pi]}^{[\tilde{n}[2]\mu[2]/\pi]} \cdots \sum_{j[d]=[\tilde{n}[d]\nu[d]/\pi]}^{[\tilde{n}[d]\mu[d]/\pi]},$$

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

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

Unfortunately, as noted by Guyon (1982a), 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 these is for the Whittle estimator of the parameters ϑ , see Guyon (1982a), 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

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

where

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

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 the *taper discrete Fourier transform* $w_v^T(\lambda_j)$ becomes the standard discrete Fourier transform (*DFT*). It is worth mentioning that to alleviate the bias problem, alternative procedures to tapering have been proposed. One of these proposals was due to Guyon (1982a), who replaced the periodogram by

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

where $\hat{\gamma}_v^*(h) = \frac{1}{N-|h|} \sum_{t(h)} v(t) v(t+h)$ and $\mathcal{D} = \{h : -n[\ell] < h[\ell] < n[\ell]; \ell = 1, \dots, d\}$. Notice that the standard periodogram $I_v(\lambda_k)$ replaces $\hat{\gamma}_v^*(h)$ by $\hat{\gamma}_v(h) = \frac{1}{N} \sum_{t(h)} v(t) v(t+h)$. 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 $d \leq 3$, it suffices for our results to hold true to employ the taper periodogram $I_v^T(\lambda_j)$.

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

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

Indeed, denoting the taper Dirichlet kernel by

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

we have that

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

The immediate consequence of property (2.5) is that it 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 the relationship between $D_\ell^T(\lambda[\ell])$ with the standard Dirichlet kernel, $D_\ell(\lambda[\ell]) = \sum_{t[\ell]=1}^{n[\ell]} e^{it[\ell]\lambda[\ell]}$, is

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

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

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

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

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

where

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

It is worth mentioning that similarly we might have employed the U_p - process. The latter is defined as

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

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). Notice that because we have excluded the frequency $\lambda_j = 0$ from the definition of $\sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]}$ and $\alpha_{\theta, N}(\lambda)$ is scale invariant, it is easy to show that a linear transformation of the data does not change the value of $\alpha_{\theta, N}$ and therefore we can assume, without loss of generality, that $\mathbb{E}x(t) = \mu$ is zero and the variance of $\varepsilon(t)$ is the unity.

One rational of the statistic $\alpha_{\theta, N}(\lambda)$ follows from the observation (see Lemma 4 in Section 5) that under H_0 , we have that

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

where “ $a \leq b$ ” means that $a[\ell] \leq b[\ell]$ for all $\ell = 1, \dots, d$ 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 $0 < j[1] \leq \tilde{n}[1]$ whereas $-\tilde{n}[\ell] < j[\ell] \leq \tilde{n}[\ell]$ for $\ell = 2, \dots, d$.

Thus from the previous observation, we can expect that $\alpha_{\theta_0, N}$ will be asymptotically equivalent to Bartlett’s U_p – process for $\{\varepsilon(t)/\sigma_\varepsilon\}_{t \in \mathbb{Z}^d}$, i.e.

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

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 α_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.

We should now comment on 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

$$(2.11) \quad E \{x(t) | x(r) : r \neq t\} = \sum_{j \in \mathbb{Z}^d \wedge \{j \neq 0\}} \delta(j) x(t-j).$$

So, (2.11) gives a motivation why the spectral density function plays a central role, and therefore why we have decided to work in terms of $f(\lambda)$. The equality in (2.11) is related to the well known *CAR* (Conditional Autoregression) model compared to the *SAR* (Simultaneous Autoregression) model in (2.12) below. 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) or Guyon (1982b).

Let us introduce the following regularity conditions.

Condition C1: (a) The process $\{\varepsilon(t)\}_{t \in \mathbb{Z}^d}$ in (1.1) is a zero mean independent identically distributed sequence of random variables with variance σ_ε^2 equal to 1 and finite 4th moments with κ_ε denoting the fourth cumulant of $\{\varepsilon(t)\}_{t \in \mathbb{Z}^d}$.

(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

$$(2.12) \quad \sum_{j \in \mathbb{Z}^d} \xi(j) x(t-j) = \varepsilon(t) \quad \xi(0) = 1,$$

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

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

denoting for multi-indices 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 $\ell = 1, \dots, d$, 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 C3. 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 α_N^0 , for the latter to have finite second moments, we require finite fourth moments for the lattice process $\{\varepsilon(t)\}_{t \in \mathbb{Z}^d}$. Also we have assumed that the true value of σ_ε^2 is 1. The latter follows from 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 σ_ε^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)* is that $\Psi(z)$ be no zero for any $z[\ell]$, $\ell = 1, \dots, d$, which simultaneously satisfy $|z[1]| = 1, \dots, |z[d]| = 1$ at least when the *Moving Average* representation is of finite order. The latter implies that $f(\lambda)$ is a positive function.

Looking at the proof of Proposition 1 below, and then that of Theorem 1, it appears that we do not need to assume finite four moments of the sequence $\{\varepsilon(t)\}_{t \in \mathbb{Z}^d}$. The reason is similar to the work of Anderson and Walker (1964) as the statistic $\alpha_{\theta, N}(\lambda)$ in (2.8). However, as in the more realistic situation when we need to estimate the unknown parameters of the model under H_0 , we require finite fourth moments to obtain the asymptotic properties of the estimates, we have just preferred to leave the condition as it stands.

Condition C2 can be generalized to the case where the rate of convergence to zero of $n^{-1}[\ell]$ differs for different $\ell = 1, \dots, d$. However, for notational simplicity we prefer to leave it as it stands. On the other hand, in 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.

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 Bickel and Wichura, 1971) and convergence in distribution in the corresponding topology will be denoted by “ \Rightarrow ”.

Proposition 1. *Under C1 – C3, we have that*

$$(2.13) \quad \alpha_N^0(\lambda) \Rightarrow \tilde{\mathbf{B}}(\lambda) = \mathbf{B} \left(\frac{\lambda}{\pi} \right) - \prod_{\ell=1}^d \left(\frac{\lambda[\ell]}{\pi} \right) \mathbf{B}(1) \quad \lambda \in [0, \pi]^d,$$

where $\{\mathbf{B}(u) : u \in [0, 1]^d\}$ is the standard Brownian sheet.

Remark 1. *Recall that the covariance structure of the standard Brownian sheet is*

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

Proposition 1 extends Grenander and Rosenblatt's (1957) results when $d = 1$, although under stronger conditions than those assumed in this paper. In particular, we do not need to assume eight bounded moments and only finite fourth moments are required.

To establish the asymptotic equivalence between $\alpha_{\theta_0, N}$ and α_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 ,*

$$\alpha_{\theta_0, N}(\lambda) \Rightarrow \tilde{\mathbf{B}}(\lambda) \quad \lambda \in [0, \pi]^d.$$

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

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

so that

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

by standard algebra. \square

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

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

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 η 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 ($\eta(g) = \pi^{-d} \int_{-\pi}^{\pi} g(\lambda)^2 d\lambda$).

In fact we have the following corollary.

Corollary 1. *Under H_0 and C1 – C4, we have that for any continuous functional $\eta(\cdot)$,*

$$\eta(\alpha_{\theta_0, N}) \xrightarrow{d} \eta(\tilde{\mathbf{B}}(\lambda)).$$

Proof. The proof follows from Theorem 1 and the continuous mapping theorem. \square

Unfortunately, the results of Theorem 1 and Corollary 1 are only valid when the “true” value of θ_0 is known, which in practical situations is unrealistic. The question is then how are our previous results affected when θ_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 θ_0 to implement the test. That is, we replace θ_0 in $\alpha_{\theta, N}(\lambda)$ by an estimator, say $\hat{\theta}$ in (3.1) below. In this scenario, drawing the terminology from Durbin (1973), we say that our null hypothesis H_0 becomes a composite hypothesis.

A popular estimator of $\vartheta'_0 = (\theta'_0, \sigma_\varepsilon^2)$ is the Whittle (1954) estimator defined as

$$\hat{\vartheta}_c = \arg \min_{\vartheta \in \Theta \times \mathbb{R}^+} \mathcal{Q}^c(\vartheta),$$

where

$$\mathcal{Q}^c(\vartheta) = \int_{-\pi}^{\pi} \left\{ \log f_\vartheta(\lambda) + \frac{I_x^T(\lambda)}{(2\pi)^d f_\vartheta(\lambda)} \right\} d\lambda$$

or in its discrete version

$$(3.1) \quad \hat{\vartheta} = \arg \min_{\vartheta \in \Theta \times \mathbb{R}^+} \mathcal{Q}_N(\vartheta),$$

where

$$(3.2) \quad \mathcal{Q}_N(\vartheta) = \frac{1}{N} \sum_{j=-\tilde{n}}^{\tilde{n}} \left\{ f_\vartheta(\lambda_j) + \frac{I_x^T(\lambda_j)}{(2\pi)^d f_\vartheta(\lambda_j)} \right\}$$

with $f_\vartheta(\lambda_j) = \sigma_\varepsilon^2 |\Psi_\theta(\lambda_j)| / (2\pi)^d$ 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 T_p -process $\alpha_{\theta_0, N}(\lambda)$ becomes

$$(3.3) \quad \alpha_{\hat{\theta}, N}(\lambda) = 2^{-1/2} N^{1/2} \left[\frac{G_{\hat{\theta}, N}(\lambda)}{G_{\hat{\theta}, N}(\pi)} - \prod_{\ell=1}^d \left(\frac{\lambda[\ell]}{\pi} \right) \right], \quad \lambda \in [0, \pi]^d,$$

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 ϑ_0 obtained by

$$\hat{\theta} = \arg \min_{\theta \in \Theta} \frac{2}{N} \sum_{j=-\tilde{n}}^{\tilde{n}} \frac{I_x^T(\lambda_j)}{|\Psi_\theta(\lambda_j)|^2}, \quad \hat{\sigma}_\varepsilon^2 = \frac{2}{N} \sum_{j=-\tilde{n}}^{\tilde{n}} \frac{I_x^T(\lambda_j)}{|\Psi_{\hat{\theta}}(\lambda_j)|^2}$$

is inconsistent. The main reason for the lack of consistency of $\hat{\theta}$ 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

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

and

$$(3.4) \quad \varphi_\theta(\lambda) = \frac{\partial}{\partial \theta} \log |\Psi_\theta(\lambda)|^2.$$

Let's introduce the following regularity conditions on θ_0 and on the model (1.1) or (2.12).

Condition C5: θ_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 θ 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 Θ and the model are standard so that we omit any comment on them. Let

$$(3.5) \quad q_{\vartheta, N} = \frac{1}{N} \sum_{j=-\tilde{n}}^{\tilde{n}} \phi_\vartheta(\lambda_j) \left\{ \frac{I_x^T(\lambda_j)}{\sigma_\varepsilon^2 |\Psi_\theta(\lambda_j)|^2} - 1 \right\}$$

$$Q_{\vartheta, N} = \frac{1}{N} \sum_{j=-\tilde{n}}^{\tilde{n}} \phi_\vartheta(\lambda_j) \phi'_\vartheta(\lambda_j),$$

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 σ_{ε}^2 as it is a parameter in itself.

Condition C8: Λ_{ϑ_0} is a continuous positive definite matrix.

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

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

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

Proof. First, by definition, we know that

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

where $\tilde{\vartheta}$ is an intermediate point between ϑ_0 and $\widehat{\vartheta}$, $q_{\vartheta_0, N}$ is given in (3.5) and $\overline{Q}_{\tilde{\vartheta}, N}$ is given by

$$\begin{aligned} & Q_{\vartheta_0, N} + \frac{1}{N} \sum_{j=-\tilde{n}}^{\tilde{n}} \left(2\phi_{\vartheta_0}(\lambda_j) \phi'_{\vartheta_0}(\lambda_j) - \frac{\partial^2 f_{\vartheta_0}(\lambda_j)}{\partial \vartheta \partial \vartheta'} \right) \left\{ \frac{I_x^T(\lambda_j)}{\sigma_{\varepsilon}^2 |\Psi_{\vartheta_0}(\lambda_j)|^2} - 1 \right\} \\ &= Q_{\vartheta_0, N} + o_p(1) \end{aligned}$$

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

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

From here the proof proceeds as in Robinson and Vidal-Sanz (2006). \square

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

$$\begin{aligned} \tilde{\varphi}_{\vartheta}(\lambda) &= \varphi_{\vartheta}(\lambda) - \frac{2}{(2\pi)^d} \int_{-\pi}^{\pi} \varphi_{\vartheta}(\lambda) d\lambda, \quad \tilde{\phi}_{\vartheta}(\lambda) = (\tilde{\varphi}'_{\vartheta}(\lambda), 0)' \\ \tilde{\varphi}_{\vartheta_0, N}(\lambda_j) &= \varphi_{\vartheta_0}(\lambda_j) - \frac{2}{N} \sum_{j=-\tilde{n}}^{\tilde{n}} \varphi_{\vartheta_0}(\lambda_j), \quad \tilde{\phi}_{\vartheta_0, N}(\lambda) = (\tilde{\varphi}'_{\vartheta_0, N}(\lambda), 0)' \end{aligned}$$

with $\varphi_{\vartheta}(\lambda)$ given in (3.4), standard algebra establishes that the Whittle estimator $\widehat{\vartheta}$ in (3.1) satisfies the asymptotic linearization

$$\begin{aligned} \widehat{\vartheta} - \vartheta_0 &= Q_{\vartheta_0, N}^{-1} \left\{ \int_{-\pi}^{\pi} \tilde{\phi}_{\vartheta_0}(\lambda) \alpha_{\vartheta_0, N}(d\lambda) + \int_{-\pi}^{\pi} \phi_{\vartheta_0}(\lambda) d\lambda \frac{1}{N} \sum_{j=-\tilde{n}}^{\tilde{n}} \left(\frac{I_x^T(\lambda_j)}{(2\pi)^d f_{\vartheta_0}(\lambda_j)} - 1 \right) \right\} \\ &+ o_p(N^{-1/2}). \end{aligned} \tag{3.6}$$

Then using (3.6) and defining

$$\alpha_{\infty}(\lambda) = \tilde{\mathbf{B}}(\lambda) - \left(\frac{1}{(2\pi)^d} \int_{-\lambda}^{\lambda} \tilde{\varphi}'_{\vartheta_0}(\bar{\lambda}) d\bar{\lambda} \right) \tilde{\Lambda}^{-1}(\vartheta_0) \int_{-\pi}^{\pi} \tilde{\varphi}_{\vartheta_0, N}(\bar{\lambda}) \tilde{\mathbf{B}}(d\bar{\lambda}),$$

where

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

we obtain the following result.

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

$$(a) \quad \alpha_{\hat{\theta}, N}(\lambda) = \alpha_N^0(\lambda) - \left(\frac{1}{N} \sum_{j=-\lceil \tilde{n}\lambda/\pi \rceil}^{\lceil \tilde{n}\lambda/\pi \rceil} \tilde{\varphi}'_{\theta_0, N}(\lambda_j) \right) \tilde{\Lambda}_{\theta_0, N}^{-1} \frac{1}{N} \sum_{j=-\tilde{n}}^{\tilde{n}} \tilde{\varphi}'_{\theta_0, N}(\lambda_j) I_\varepsilon^T(\lambda_j) + o_p(1),$$

where the $o_p(1)$ is uniformly in λ and

$$\tilde{\Lambda}_{\theta_0, N} = \frac{1}{N} \sum_{j=-\tilde{n}}^{\tilde{n}} \tilde{\varphi}_{\theta_0, N}(\lambda_j) \tilde{\varphi}'_{\theta_0, N}(\lambda_j).$$

$$(b) \quad \alpha_{\hat{\theta}, N} \Rightarrow \alpha_\infty.$$

From Theorem 3 and continuous mapping theorem we have the following corollary. Let $\eta(\cdot)$ be a continuous functional $\eta : D[0, \pi]^d \rightarrow \mathbb{R}$.

Corollary 2. *Let $\hat{\eta}_N = \eta(\alpha_{\hat{\theta}, N}(\lambda))$. Under H_0 and the same conditions of Theorem 3, we have that*

$$\hat{\eta}_N \xrightarrow{d} \eta(\alpha_\infty(\lambda)).$$

Proof. The proof follows from Theorem 3 and the continuous mapping theorem. \square

The main conclusion that we draw from Theorem 3 is that the T_p -process $\alpha_{\hat{\theta}, N}$ is no longer asymptotically pivotal, so that the immediate consequence is that tests based on η , say $\eta(\alpha_{\hat{\theta}, N}) = \sup_{\lambda \in [0, \pi]^d} |\alpha_{\hat{\theta}, N}(\lambda)|$ or $\eta(\alpha_{\hat{\theta}, N}) = (2\pi)^{-d} \int_{-\pi}^{\pi} \alpha_{\hat{\theta}, N}(\lambda)^2 d\lambda$, are not useful for practical purposes as its asymptotic critical values are difficult, if at all possible, to obtain. To compute the critical values of the asymptotic distribution of $\hat{\eta}_N$, several approaches have been described and examined. A first approach 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 latter method is that the power of the test is smaller than the one proposed in the paper, that is if we denote by b_N the smoothing parameter, their test has a local power of order $(Nb_N)^{-1/2}$ whereas ours is $N^{-1/2}$. A second potential drawback is that the choice of the bandwidth b_N 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 b_N .

A second alternative 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 $\alpha_{\hat{\theta}, N}$ with estimated orthogonal components. The number of estimated orthogonal components must suitably increase with the sample size.

However, its implementation is quite cumbersome even for the rather simpler case when $d = 1$. See for instance Anderson (1997) for details.

So, in view of the preceding arguments, we consider a third approach based on 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 articles were for $d = 1$. Also, we will see that bootstraps employed when $d = 1$ are not valid in our context.

4. BOOTSTRAP TEST FOR THE TEST

Since Efron (1979), bootstrap algorithms have become a common tool in applied work and thus considerable effort has been devoted to its development. The primary motivation for this effort is that they have proved to be a very useful statistical tool. We can cite two main examples/reasons. First, bootstrap methods are capable of approximating the finite sample distribution of statistics better than those based on their asymptotic counterparts. And secondly, and perhaps the most important, they allow computing valid asymptotic quantiles of the limiting distribution in situations when the practitioner is unable to compute its quantiles.

In the present paper we face the latter situation. Following our comments at the end of the previous section, the aim of this section is to propose a bootstrap procedure for $\alpha_{\hat{\theta},N}(\lambda)$ given in (3.3) and thus for $\hat{\eta}_N = \eta(\alpha_{\hat{\theta},N})$. The resampling method must be such that the conditional distribution, given $\tilde{x} = \{x_t\}_{t=1}^n$, of the bootstrap statistic, say $\hat{\eta}_N^*$, consistently estimates the distribution of $\eta(\alpha_\infty)$ under H_0 . That is, $\hat{\eta}_N^* \rightarrow_{d^*} \eta(\alpha_\infty)$ in probability under H_0 , where “ \rightarrow_{d^*} ” denotes

$$\Pr \left[\hat{\eta}_N^* \leq z \mid \tilde{x} \right] \xrightarrow{P} \mathcal{G}(z),$$

at each continuity point z of $\mathcal{G}(z) = \Pr(\eta(\alpha_\infty) \leq z)$. Moreover, under local alternatives

$$(4.1) \quad H_a : f_\vartheta(\lambda) \left(1 + \frac{1}{N^{1/2}} g(\lambda) \right) \text{ for some } \vartheta \in \Theta \times \mathbb{R}^+$$

where $g(\lambda)$ is some symmetric, non-constant continuous function in $[0, \pi]$ such that $\frac{1}{N^{1/2}} g(\lambda) > -1$ for all $N \geq 1$, $\hat{\eta}_N^*$ must also converge, in bootstrap distribution to $\eta(\alpha_\infty)$, whereas under the alternative H_1 , we only require that $\hat{\eta}_N^*$ is bounded in probability to have good power properties.

Remark 3. We should point out that H_a could have been written as

$$H_a : f_\vartheta(\lambda) + \frac{1}{N^{1/2}} \tilde{g}(\lambda) \text{ for some } \vartheta \in \Theta \times \mathbb{R}^+$$

where $\tilde{g}(\lambda)$ is a positive integrable function. However, since we are concerned with the relative error of $I^T(\lambda_j)$ compared to $f_\vartheta(\lambda_j)$ ($|\Psi_\vartheta(\lambda_j)|^2$), we found notationally more convenient to write the alternative hypothesis H_a as given in (4.1).

When $d = 1$, Hidalgo and Kreiss (2006) examined a bootstrap algorithm based on an approach in Hidalgo (2003) showing its validity and consistency. This bootstrap consists on the following 3 STEPS.

STEP 1: Let $\tilde{x}(t) = (x(t) - \bar{x}) / \hat{\sigma}_x$, where $\bar{x} = N^{-1} \sum_{t=1}^n x(t)$ and $\hat{\sigma}_x^2 = N^{-1} \sum_{t=1}^N (x(t) - \bar{x})^2$, and a random sample of size N with replacement from the empirical distribution of $\tilde{x}(t)$. Denote that sample as $\tilde{x}^* = \{x^*(t)\}_{t=1}^n$.

STEP 2: For $j = 1, \dots, \tilde{n}$, compute the bootstrap periodogram

$$\tilde{I}_{x^*}^T(\lambda_j) = f_{\hat{\vartheta}}(\lambda_j) \check{I}_{x^*}^T(\lambda_j),$$

where

$$f_{\hat{\vartheta}}(\lambda_j) = \frac{G_{\hat{\theta}, N}(\pi)}{(2\pi)^d} |\Psi_{\hat{\theta}}(\lambda_j)|^2, \quad \check{I}_{x^*}^T(\lambda_j) = \frac{1}{\sum_{t=1}^n h^2(t)} \left| \sum_{t=1}^n h(t) x^*(t) e^{it \cdot \lambda_j} \right|^2$$

and the bootstrap analogue of $\hat{\vartheta}$ by

$$(4.2) \quad \check{\vartheta}^* = \arg \min_{\vartheta \in \Theta \times \mathbb{R}^+} \tilde{\mathcal{Q}}_N^*(\vartheta),$$

where

$$\tilde{\mathcal{Q}}_N^*(\vartheta) = \frac{1}{N} \sum_{j=-\tilde{n}}^{\tilde{n}} \left\{ f_{\vartheta}(\lambda_j) + \frac{\tilde{I}_{x^*}^T(\lambda_j)}{(2\pi)^d f_{\vartheta}(\lambda_j)} \right\}$$

with $f_{\vartheta}(\lambda_j) = \sigma_{\varepsilon}^2 |\Psi_{\theta}(\lambda_j)| / (2\pi)^d$.

STEP 3: Compute the bootstrap T_p - process

$$\alpha_{\check{\vartheta}^*, N}(\lambda) = 2^{-1/2} N^{1/2} \left[\frac{\tilde{G}_{\check{\vartheta}^*, N}^*(\lambda)}{G_{\check{\vartheta}^*, N}(\pi)} - \prod_{\ell=1}^d \left(\frac{\lambda[\ell]}{\pi} \right) \right], \quad \lambda \in [0, \pi]^d,$$

where

$$\tilde{G}_{\check{\vartheta}^*, N}^*(\lambda) = 2 \frac{1}{N} \sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]} \frac{\tilde{I}_{x^*}^T(\lambda_j)}{|\Psi_{\theta}(\lambda_j)|^2}.$$

Some other procedures are possible as that based on that of Franke and Härdle (1992), where the bootstrap periodogram $\tilde{I}_{x^*}^T(\lambda_j) = |\Psi_{\hat{\theta}}(\lambda_j)|^2 \check{I}_{x^*}^T(\lambda_j)$ is replaced by

$$\hat{I}_{x^*}^T(\lambda_j) = f_{\hat{\vartheta}}(\lambda_j) \chi_j,$$

where $\chi_{-\tilde{n}}, \dots, \chi_{\tilde{n}}$ are independent exponential random variables. However, unlike in the case of $d = 1$, the previous bootstrap algorithm will not be valid. The reason is because the bootstrap does not correctly “estimate” the fourth cumulant κ_{ε} . More specifically the asymptotic distribution of the bootstrap estimator $\check{\vartheta}^*$ in (4.2) will not have the same asymptotic variance as that of $\hat{\vartheta}$ in (3.1).

So to overcome this problem, following Hidalgo (2007), see also Hidalgo and Lazarova (2007), we propose in the paper an alternative algorithm, as described in the next 4 STEPS.

STEP 1: We first obtain the residuals

$$\hat{\varepsilon}(t) = (2\pi)^{d/2} \frac{1}{N^{1/2}} \sum_{j=-\tilde{n}}^{\tilde{n}} e^{-it \cdot \lambda_j} \Psi_{\hat{\theta}}^{-1}(e^{i\lambda_j}) w_x(\lambda_j),$$

for $t = 1, \dots, n$. From here as usual, we obtain a random sample of size N with replacement from the empirical distribution function of $\{\hat{\varepsilon}(t)\}_{t=1}^n$. Let's denote the bootstrap sample by $\{\varepsilon^*(t)\}_{t=1}^n$.

Remark 4. (a) Notice that because the test statistic $\hat{\eta}_N = \eta(\alpha_{\hat{\vartheta}, N}(\lambda))$ is asymptotically independent of the mean and variance of $\{\varepsilon(t)\}_{t \in \mathbb{Z}^d}$, we do not need to standardize $\hat{\varepsilon}(t)$ before computing the empirical distribution of $\{\hat{\varepsilon}(t)\}_{t=1}^n$ in order to obtain the bootstrap sample. (b) The motivation to compute the residuals as in STEP 1 comes from the observation that for $t = 1, \dots, n$,

$$\varepsilon(t) \simeq (2\pi)^{d/2} \frac{1}{N^{1/2}} \sum_{j=-\tilde{n}}^{\tilde{n}} e^{-it \cdot \lambda_j} \Psi_{\theta_0}^{-1}(e^{i\lambda_j}) w_x(\lambda_j).$$

STEP 2: For $t = 1, \dots, n$, compute the bootstrap observations

$$(4.3) \quad x^*(t) = (2\pi)^{d/2} \frac{1}{N^{1/2}} \sum_{j=-\tilde{n}}^{\tilde{n}} e^{-it \cdot \lambda_j} \Psi_{\hat{\vartheta}}(e^{i\lambda_j}) w_{\varepsilon^*}(\lambda_j),$$

where $w_{\varepsilon^*}(\lambda_j)$ is the standard *DFT* of $\{\varepsilon^*(t)\}_{t=1}^n$, and the taper periodogram

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

STEP 3: The bootstrap analogue of $\hat{\vartheta}$ is given by

$$(4.4) \quad \hat{\vartheta}^* = \arg \min_{\vartheta \in \Theta \times \mathbb{R}^+} \mathcal{Q}_N^*(\vartheta),$$

where

$$(4.5) \quad \mathcal{Q}_N^*(\vartheta) = \frac{1}{N} \sum_{j=-\tilde{n}}^{\tilde{n}} \left\{ f_{\vartheta}(\lambda_j) + \frac{I_{x^*}^T(\lambda_j)}{(2\pi)^d f_{\vartheta}(\lambda_j)} \right\}.$$

STEP 4: Compute the bootstrap T_p - process

$$(4.6) \quad \alpha_{\hat{\vartheta}^*, N}^*(\lambda) = 2^{-1/2} N^{1/2} \left[\frac{G_{\hat{\vartheta}^*, N}^*(\lambda)}{G_{\hat{\vartheta}^*, N}^*(\pi)} - \prod_{\ell=1}^d \left(\frac{\lambda[\ell]}{\pi} \right) \right], \quad \lambda \in [0, \pi]^d,$$

with

$$(4.7) \quad G_{\hat{\vartheta}, N}^*(\lambda) = 2 \frac{1}{N} \sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]} \frac{I_{x^*}^T(\lambda_j)}{|\Psi_{\hat{\vartheta}}(\lambda_j)|^2}.$$

Theorem 4. Under $C1 - C3$ and $C5 - C8$, we have that

$$N^{1/2} \left(\hat{\vartheta}^* - \hat{\vartheta} \right) \xrightarrow{d^*} \mathcal{N} \left(0, 2\Lambda_{\hat{\vartheta}_0}^{-1} V_{\hat{\vartheta}_0} \Lambda_{\hat{\vartheta}_0}^{-1} \right),$$

where $\xrightarrow{d^*}$ denotes convergence in bootstrap distribution.

As it happens with $\hat{\vartheta}$ in Section 3, we also have that the Whittle bootstrap estimator $\hat{\vartheta}^*$ in (4.4) satisfies the asymptotic linearization

$$(4.8) \quad \begin{aligned} \hat{\vartheta}^* - \hat{\vartheta} &= Q_{\hat{\vartheta}, N}^{-1} \left\{ \int_{-\pi}^{\pi} \tilde{\phi}_{\hat{\vartheta}}(\lambda) \alpha_{\hat{\vartheta}, N}(d\lambda) + \int_{-\pi}^{\pi} \phi_{\hat{\vartheta}}(\lambda) d\lambda \frac{1}{N} \sum_{j=-\tilde{n}}^{\tilde{n}} \left(\frac{I_{x^*}^T(\lambda_j)}{(2\pi)^d f_{\hat{\vartheta}}(\lambda_j)} - 1 \right) \right\} \\ &+ o_{p^*} \left(N^{-1/2} \right). \end{aligned}$$

Let

$$(4.9) \quad \alpha_N^{*0}(\lambda) = 2^{-1/2} N^{1/2} \left[\frac{G_N^{*0}(\lambda)}{G_N^{*0}(\pi)} - \prod_{\ell=1}^d \left(\frac{\lambda[\ell]}{\pi} \right) \right],$$

with

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

Theorem 5. *Under H_0 and assuming C.1 – C.3 and C5 – C8, uniformly in $\lambda \in [\theta, \pi]$,*

$$(a) \alpha_{\hat{\theta}^*, N}^*(\lambda) = \alpha_N^{*0}(\lambda) - \left(\frac{1}{N} \sum_{j=-[\bar{n}\lambda/\pi]}^{[\bar{n}\lambda/\pi]} \tilde{\varphi}'_{\hat{\theta}, N}(\lambda_j) \right) \tilde{\Lambda}_{\hat{\theta}, N}^{-1} \frac{1}{N} \sum_{j=-\bar{n}}^{\bar{n}} \tilde{\varphi}'_{\hat{\theta}, N}(\lambda_j) I_{\varepsilon^*}^T(\lambda_j) + o_{p^*}(1),$$

where the $o_{p^*}(1)$ is uniformly in λ .

$$(b) \alpha_{\hat{\theta}^*, N}^* \xrightarrow{d^*} \alpha_\infty.$$

A conclusion from Theorem 5 is the following corollary.

Corollary 3. *Let $\hat{\eta}_N^* = \eta(\alpha_{\hat{\theta}^*, N}^*(\lambda))$. Under the maintained hypothesis and assuming C1 – C3 and C5 – C8, we have that for any continuous functional,*

$$\hat{\eta}_N^* \xrightarrow{d^*} \eta(\alpha_\infty(\lambda)).$$

Proof. The proof follows from Theorem 5 and the continuous mapping theorem. \square

Thus, Theorem 5 and Corollary 3 indicate that the bootstrap statistic $\hat{\eta}_N^*$ is consistent. That is, let $c_{N, (1-\alpha)}^f$ and $c_{(1-\alpha)}^a$ be such that

$$\Pr \left\{ |\hat{\eta}_N| > c_{n, (1-\alpha)}^f \right\} = \alpha$$

and

$$\lim_{n \rightarrow \infty} \Pr \left\{ |\hat{\eta}_N| > c_{(1-\alpha)}^a \right\} = \alpha,$$

respectively. So, Theorems 3 and 5 indicate that $c_{N, (1-\alpha)}^f \rightarrow c_{(1-\alpha)}^a$ and $c_{(1-\alpha)}^* \xrightarrow{p} c_{(1-\alpha)}^a$, respectively, where $c_{(1-\alpha)}^*$ is defined as

$$\Pr \left\{ |\hat{\eta}_N^*| > c_{(1-\alpha)}^* \right\} = \alpha.$$

Typically, the finite sample distribution of $\hat{\eta}_N^*$ is not available, although the critical values $c_{(1-\alpha)}^*$ can be approximated, as accurately as desired, by standard Monte-Carlo simulation. To that end, consider the bootstrap samples $\left\{ \hat{\varepsilon}^{*\ell}(t) \right\}_{t=1}^n$ for $\ell = 1, \dots, B$, and compute $\alpha_{\hat{\theta}^*, N}^{\ell}(\lambda)$ as in (4.5) for each ℓ . Then, $c_{(1-\alpha)}^*$ is approximated by the value $c_{(1-\alpha)}^{*B}$ that satisfies

$$\frac{1}{B} \sum_{\ell=1}^B \mathcal{I} \left(\eta \left(\alpha_{\hat{\theta}^*, N}^{\ell}(\lambda) \right) > c_{(1-\alpha)}^{*B} \right) = \alpha.$$

Next we study the behaviour of the bootstrap tests under the alternative hypothesis H_1 .

Corollary 4. *Assuming C.1-C.8, under H_1 ,*

$$\hat{\eta}_N^* \xrightarrow{d^*} \eta(\tilde{\alpha}_\infty(\lambda)) \text{ in probability,}$$

where $\tilde{\alpha}_\infty(\lambda)$ is a centered Gaussian process with covariance structure as $\alpha_\infty(\lambda)$ but with θ_0 replaced by $\theta_1 = \text{plim } \hat{\theta}$.

Proof. The proof proceeds exactly the same as that of Theorem 5 and then Corollary 3 but instead of writing $\hat{\theta} - \theta_0 = o_p(1)$ we write $\hat{\theta} - \theta_1 = o_p(1)$ and θ_1 instead of θ_0 . \square

5. LEMMAS

First, we introduce some notation. We denote the conjugate of a complex number a by \bar{a} . Also, for a generic function $\nu(\lambda)$, we abbreviate $\nu(\lambda_j)$ by $\nu_j = (\nu_{j[1]}, \dots, \nu_{j[d]})'$ 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 $\tilde{\mathbb{Z}}^d = \{j : (-\tilde{n} \prec j \prec \tilde{n}) \wedge (0 < j[1])\}$.

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

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

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

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

surprising any reference to ℓ in K_ℓ^T and/or D_ℓ^T for notational simplicity.

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

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

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

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

by C2 and standard algebra.

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

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

by standard arguments after using (2.5). \square

Lemma 2. *Let $k \prec j \in \tilde{\mathbb{Z}}^d$ 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,*

$$\begin{aligned} (a) \quad \mathbb{E}(w_{\xi,j}^T \bar{w}_{\zeta,k}^T) &= f_{\xi\zeta,j} \mathcal{I}(|j[\ell] - k[\ell]| = 2, \ell = 1, \dots, d) + O(c_{jk}) \\ (b) \quad \mathbb{E}(w_{\xi,j}^T w_{\zeta,k}^T) &= O(c_{jk}). \end{aligned}$$

Proof. We shall handle part (a) only, being part (b) identical. By definition, (5.1)

$$\mathbb{E} (w_{\xi,j}^T \bar{w}_{\zeta,k}^T) = 6^{-d} \int_{\Pi^d} f_{\xi\zeta}(\lambda) \prod_{\ell=1}^d n^{[\ell]-1} D^T(\lambda^{[\ell]} - \lambda_{j^{[\ell]}}) D^T(\lambda_{k^{[\ell]}} - \lambda^{[\ell]}) d\lambda.$$

Because $|f_{\xi\zeta}(\lambda)| < C$, the modulus of the right side of (5.1) is bounded by

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

using C2. Now using (2.5) and because $\int_0^\pi |D^T(\lambda)| d\lambda < C$, the contribution due to a factor of the type $\int_{-\pi}^{\lambda_{(j^{[\ell]}+k^{[\ell]})/2}}$ is bounded by

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

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

$$(5.2) \quad \int_{\Pi^d} \prod_{\ell=1}^d D^T(\lambda^{[\ell]} - \lambda_{j^{[\ell]}}) D^T(\lambda_{k^{[\ell]}} - \lambda^{[\ell]}) d\lambda = 0,$$

when $|j^{[\ell]} - k^{[\ell]}| \neq 2$ for some $\ell = 1, \dots, d$, we have that, in this case, except multiplicative constants, the left side of (5.1) is

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

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

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

because $|\lambda D(\lambda)| < C$, $\int_0^\pi |D(\lambda)| d\lambda = O(\log n)$ and the Cauchy-Schwarz inequality imply that $\int_{\Pi} |D(\lambda - \lambda_{j^{[\ell]}})| |D(\lambda_{k^{[\ell]}} - \lambda)| d\lambda = O(n)$. Now, when for all $\ell = 1, \dots, d$, $|j^{[\ell]} - k^{[\ell]}| = 2$, because the left side of (5.2) is 1, we have that proceeding as above, the left side of (5.1) is $f_{\xi\zeta,j} + O(n^{-1} \log n)$. This concludes the proof. \square

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 Π^d . Under C1-C4, we have that for all $r \leq s \in \tilde{\mathbb{Z}}^d$*

$$(5.3) \quad \mathbb{E} \left| \sum_{j=r}^s \zeta_j v_j (u_j - v_j) \right|^2 = O\left(\frac{\log n}{n} \prod_{\ell=1}^d |s^{[\ell]} - r^{[\ell]}|_+\right).$$

Proof. Denote $\varrho_j = u_j - v_j$. By standard arguments, the left side of (5.3) is

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

where

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

After observing that $\mathbb{E}(v_j \bar{u}_j) = 1 + O(n^{-2})$ and $\mathbb{E}(v_j \bar{\varrho}_j) = \mathbb{E}(v_j \bar{u}_j) - \mathbb{E}(v_j \bar{v}_j)$, we have that Lemma 1 implies that $a_{j1} = O(n^{-2})$, whereas Lemmas 1 and 2 imply that $b_{jk,1} = O\left(c_{jk}^2 + n^{-1} \log n \mathcal{I}(|j[\ell] - k[\ell]| = 2, \ell = 1, \dots, d)\right)$, with c_{jk} as defined there. From here it is immediate to conclude that the contribution due to a_{j1} and $b_{jk,1}$ into the left of (5.3) is its right side.

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

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

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 $\prod_{\ell=1}^d |s[\ell] - r[\ell]|_+ \leq N$. \square

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

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

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

$$(5.4) \quad \mathbb{E} \sup_s \left| \sum_{j=1}^s \zeta_j \left\{ \frac{I_{x,j}^T}{|\Psi_j|^2} - I_{\varepsilon,j}^T \right\} \right| \leq \mathbb{E} \sup_s \sum_{j=1}^s |\zeta_j| |\varrho_j|^2 + 2 \mathbb{E} \sup_s \left| \sum_{j=1}^s \zeta_j v_j \bar{\varrho}_j \right|$$

is $o(N^{1/2})$, where we abbreviate “ $\sup_{s=1, \dots, \bar{n}}$ ” by “ \sup_s ” and $\varrho_j = u_j - v_j$.

The first term on the right of (5.4) is bounded by

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

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

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

Next, we examine the second term of (5.4). To that end, let $q = 0, \dots, [\tilde{n}^\varsigma] - 1$ for some $0 < \varsigma < 1/d$. (Recall that $[\tilde{n}^\psi] = ([\tilde{n}^\psi [1]], \dots, [\tilde{n}^\psi [d]])$ for any $\psi > 0$.) Standard inequalities imply that the square of the second term on the right of (5.4) is bounded by

$$(5.5) \quad \mathbb{E} \max_s \left| \left\{ \sum_{j=1}^s - \sum_{j=1}^{q(s)[\tilde{n}^{1-\varsigma}]} \right\} \zeta_j v_j \bar{\varrho}_j \right|^2 + \mathbb{E} \max_s \left| \sum_{j=1}^{q(s)[\tilde{n}^{1-\varsigma}]} \zeta_j v_j \bar{\varrho}_j \right|^2,$$

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

From the definition of $q(s)$ and

$$(5.6) \quad (\sup_p |c_p|)^2 = \sup_p |c_p|^2 \leq \sum_p |c_p|^2,$$

the second term of (5.5) is bounded by

$$\sum_{q=1}^{[\tilde{n}^\varsigma]-1} \mathbb{E} \left| \sum_{j=1}^{q[\tilde{n}^{1-\varsigma}]} \zeta_j v_j \bar{\varrho}_j \right|^2 = O(N^{1+\varsigma} n^{-1} \log^2 n) = o(N)$$

by Lemma 3 and because $\varsigma < 1/d$. To complete the proof we need to show that the first term in (5.5) is $o(N)$. To that end, we note that it is bounded by

$$\mathbb{E} \max_{q=1, \dots, [\tilde{n}^\varsigma]-1} \max_{s=1+q[\tilde{n}^{1-\varsigma}], \dots, (q+1)[\tilde{n}^{1-\varsigma}]} \left| \sum_{j=1+q[\tilde{n}^{1-\varsigma}]}^s \zeta_j v_j \bar{\varrho}_j \right|^2$$

which is $O(N^\varsigma) \mathbb{E} \max_{s=1, \dots, [\tilde{n}^{1-\varsigma}]} \left| \sum_{j=1}^s \zeta_j v_j \bar{\varrho}_j \right|^2$.

So, we have that the square of the second term on the right of (5.4) is

$$O(N^{1+\varsigma} n^{-1} \log^2 n) + O(N^\varsigma) \mathbb{E} \max_{s=1, \dots, [\tilde{n}^{1-\varsigma}]} \left| \sum_{j=1}^s \zeta_j v_j \bar{\varrho}_j \right|^2.$$

Observe that the second factor of the second term of the last displayed expression is similar to the second term on the right of (5.4) but with $s = 1, \dots, [\tilde{n}^{1-\varsigma}]$ instead of $s = 1, \dots, \tilde{n}$. So, repeating the same steps, the last displayed expression, and so

the square of the second term on the right of (5.4), is

$$\begin{aligned} & O \left(\left\{ N^{1+\varsigma} + \sum_{q=0}^{\iota-1} N^{(1+\varsigma)(1-\varsigma)^{q+1} + \varsigma \sum_{p=0}^q (1-\varsigma)^p} \right\} n^{-1} \log^2 n \right) \\ & + N^{\varsigma \sum_{p=0}^{\iota-1} (1-\varsigma)^p} \mathbb{E} \max_{s=1, \dots, (\tilde{n}^{1-\varsigma})^\iota} \left| \sum_{j=1}^s \zeta_j v_j \bar{\varrho}_j \right|^2 \\ & = O \left(N^{1+\varsigma} n^{-1} \log^2 n \right) + O \left(N^{\varsigma \sum_{p=0}^{\iota-1} (1-\varsigma)^p} \right) \sum_{s=1}^{\tilde{n}^{(1-\varsigma)^\iota}} \mathbb{E} \left| \sum_{j=1}^s \zeta_j v_j \bar{\varrho}_j \right|^2 = o(N) \end{aligned}$$

after choosing ι large enough because $\varsigma < 1/d$. This completes the proof. \square

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

$$(5.7) \quad \frac{1}{N} \sup_{\vartheta \in \Theta \times \mathbb{R}^+} \left| \sum_{j=-\tilde{n}}^{\tilde{n}} \zeta_j(\vartheta) \left(\frac{I_{x,j}^T}{|\Psi_{\theta_0,j}|^2} - 1 \right) \right| = o_p(1).$$

Proof. By the triangle inequality, the left side of (5.7) is bounded by

$$(5.8) \quad \frac{C}{N} \sup_{\vartheta \in \Theta \times \mathbb{R}^+} \left| \sum_{j=-\tilde{n}}^{\tilde{n}} \zeta_j(\vartheta) \left(\frac{I_{x,j}^T}{|\Psi_{\theta_0,j}|^2} - I_{\varepsilon,j}^T \right) \right| + \frac{C}{N} \sup_{\vartheta \in \Theta \times \mathbb{R}^+} \left| \sum_{j=-\tilde{n}}^{\tilde{n}} \zeta_j(\vartheta) (I_{\varepsilon,j}^T - 1) \right|.$$

Now, because by assumption $|\zeta(\lambda; \vartheta)| < C$, the first term of (5.8) is bounded by

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

by Markov's inequality because by the Cauchy-Schwarz inequality $\mathbb{E} |v_j \bar{\varrho}_j|^2 \leq \mathbb{E} |v_j|^2 \mathbb{E} |\bar{\varrho}_j|^2$ and then proceeding as in Lemma 3. Next, we show that the second term of (5.8) is $o_p(1)$. But this follows by standard arguments (see also Lemma 15) and because $\zeta_j(\vartheta)$ is continuously differentiable in ϑ . \square

Lemma 6. Assume C1 – C3 and C5 – C8. Then,

$$\hat{\vartheta} - \vartheta_0 = o_p(1).$$

Proof. The proof follows very easily using Lemma 5. Indeed, (3.2) is

$$(5.9) \quad \frac{1}{N} \sum_{j=-\tilde{n}}^{\tilde{n}} \frac{f_{\vartheta_0,j}}{f_{\vartheta,j}} \left(\frac{I_{x,j}^T}{(2\pi)^d f_{\vartheta_0,j}} - 1 \right) + \frac{1}{N} \left\{ \sum_{j=-\tilde{n}}^{\tilde{n}} \frac{f_{\vartheta_0,j}}{f_{\vartheta,j}} - \log \frac{f_{\vartheta_0,j}}{f_{\vartheta,j}} + \log f_{\vartheta_0,j} \right\}.$$

Now then the second term of (5.9) converges using Brillinger (1981, p.15) to

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

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 of (5.9) converges to zero uniformly in ϑ 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, so we omit its details. \square

Lemma 7. *Assume C1 – C3 and C5 – C8. Under H_0 ,*

$$(5.10) \quad \frac{1}{N^{1/2}} \sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]} \zeta_j \left(\frac{I_{x,j}^T}{|\Psi_{\hat{\theta},j}|^2} - 1 \right) = \frac{1}{N^{1/2}} \sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]} \zeta_j (I_{\varepsilon,j}^T - 1) \\ - \left(\frac{1}{N} \sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]} \zeta_j \varphi'_{\theta_0,j} \right) N^{1/2} (\hat{\theta} - \theta_0) + o_p(1),$$

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 (5.10) and the first term on its right side is

$$(5.11) \quad \frac{1}{N^{1/2}} \sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]} \zeta_j \frac{I_{x,j}^T}{|\Psi_{\theta_0,j}|^2} \left[\frac{|\Psi_{\theta_0,j}|^2}{|\Psi_{\hat{\theta},j}|^2} - 1 + \varphi'_{\theta_0,j} (\hat{\theta} - \theta_0) \right] \\ + \frac{1}{N^{1/2}} \sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]} \zeta_j \left(\frac{I_{x,j}^T}{|\Psi_{\theta_0,j}|^2} - I_{\varepsilon,j}^T \right) - \frac{1}{N^{1/2}} \sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]} \zeta_j \varphi'_{\theta_0,j} \frac{I_{x,j}^T}{|\Psi_{\theta_0,j}|^2} (\hat{\theta} - \theta_0).$$

First, because each component of the vector $\zeta(\lambda) \varphi_{\theta_0}(\lambda)$ satisfies the same conditions of $\zeta(\lambda)$ in Lemma 4, Markov's inequality implies that the second term of (5.11) is $o_p(1)$, whereas the third term is $N^{-1} \sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]} \zeta_j \varphi'_{\theta_0,j} N^{1/2} (\hat{\theta} - \theta_0) + o_p(1)$ by Lemma 4 and because proceeding as in Robinson and Vidal-Sanz (2006)

$$\frac{1}{N^{1/2}} \sum_{j=-\tilde{n}}^{\tilde{n}} \zeta_j \varphi_{\theta_0,j} (I_{\varepsilon,j}^T - 1) = O_p(1)$$

Finally, by mean value theorem, the norm of the first term of (5.11) is bounded by

$$(5.12) \quad CN^{1/2} \|\hat{\theta} - \theta_0\|^2 \frac{1}{N} \sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]} \frac{I_{x,j}^T}{|\Psi_{\theta_0,j}|^2} = O_p(N^{-1/2}),$$

by Theorem 2 and proceeding as with the third term of (5.11). This concludes the proof. \square

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

$$(5.13) \quad \mathcal{E}_{1,N}(v_1, v_2) = \left(\frac{1}{N} \sum_{j=[\tilde{n}v_1/\pi]}^{[\tilde{n}v_2/\pi]} \zeta_j \right) \left(\frac{N^{1/2}}{\sum_{t=1}^n h^2(t)} \sum_{t=1}^n (\varepsilon^T(t)^2 - 1) \right)$$

$$(5.14) \quad \mathcal{E}_{2,N}(v_1, v_2) = \frac{1}{N} \sum_{j=[\tilde{n}v_1/\pi]}^{[\tilde{n}v_2/\pi]} \zeta_j \frac{N^{1/2}}{\sum_{t=1}^n h^2(t)} \sum_{t_1 \neq t_2=1}^n \varepsilon^T(t_1) \varepsilon^T(t_2) e^{i(t_1-t_2) \cdot \lambda_j},$$

where $\varepsilon^T(t) = h(t) \varepsilon(t)$. Notice that $\mathcal{E}_{1,N}(v_1, v_2) + \mathcal{E}_{2,N}(v_1, v_2) = N^{-1/2} \sum_{j=[\tilde{n}v_1/\pi]}^{[\tilde{n}v_2/\pi]} \zeta_j (I_{\varepsilon,j}^T - 1)$.

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

$$(5.15) \quad \mathbb{E} \left(|\mathcal{E}_{j,N}(v_1, v)|^\beta |\mathcal{E}_{j,N}(v, v_2)|^\beta \right) \leq C \prod_{\ell=1}^d (v_2[\ell] - v_1[\ell])^2, \quad j = 1, 2.$$

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 $\left| N^{-1} \sum_{p=\lceil \tilde{n}v_1/\pi \rceil}^{\lceil \tilde{n}v_2/\pi \rceil} \zeta_p^q \right| \leq C \prod_{\ell=1}^d (v_2[\ell] - v_1[\ell])$ for any $q \geq 1$. \square

Next we will show that the processes $\mathcal{E}_{1,N}(0, \lambda)$ and $\mathcal{E}_{2,N}(0, \lambda)$ are tight. From Bickel and Wichura (1971) it suffices to show the following lemma.

Lemma 9. *Assuming C1, we have that*

$$(5.16) \quad (a) \quad \mathbb{E} \prod_{\ell=1}^d \left(\mathcal{E}_{1,N}^{(\ell)}(0, \lambda_{1[\ell]}) - \mathcal{E}_{1,N}^{(\ell)}(0, \lambda_{2[\ell]}) \right)^2 \leq C \prod_{\ell=1}^d (\lambda_{2[\ell]} - \lambda_{1[\ell]})^2$$

$$(5.17) \quad (b) \quad \mathbb{E} \prod_{\ell=1}^d \left(\mathcal{E}_{2,N}^{(\ell)}(0, \lambda_{1[\ell]}) - \mathcal{E}_{2,N}^{(\ell)}(0, \lambda_{2[\ell]}) \right)^4 \leq C \prod_{\ell=1}^d (\lambda_{2[\ell]} - \lambda_{1[\ell]})^2$$

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

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

Proof. The proof follows after observing that $\mathcal{E}_{j,N}^{(\ell)}(0, \lambda_{1[\ell]}) - \mathcal{E}_{j,N}^{(\ell)}(0, \lambda_{2[\ell]}) = \mathcal{E}_{j,N}^{(\ell)}(\lambda_{1[\ell]}, \lambda_{2[\ell]})$ for $j = 1, 2$ and then by Lemma 8. \square

Lemma 10. *Under C1 – C3 and C5 – C8,*

$$\hat{\sigma}_\varepsilon^2 = \frac{1}{N} \sum_{t=1}^n \hat{\varepsilon}^2(t) \quad \text{and} \quad \hat{\mu}_{4,\varepsilon} = \frac{1}{N} \sum_{t=1}^n \hat{\varepsilon}^4(t)$$

are consistent estimators of σ_ε^2 and $\mu_{4,\varepsilon}$, respectively.

Proof. See Theorem 4 of Robinson and Vidal-Sanz (2006). \square

In what follows, we shall abbreviate $w_{x^*}^T(\lambda) / \Psi_{\hat{\theta}}(\lambda)$ and $w_{\varepsilon^*}^T(\lambda)$ by $u^*(\lambda)$ and $v^*(\lambda)$ respectively for all $\lambda \in \Pi^d$.

Lemma 11. *Consider $j \in \tilde{\mathbb{Z}}^d$. Then, for $\zeta_{x^*,j}^T$ and $\xi_{x^*,j}^T$ equal to u_j^* or v_j^* ,*

$$(a) \quad \mathbb{E}^* \left(\zeta_{x^*,j}^T \bar{\xi}_{x^*,j}^T \right) - \hat{\sigma}_\varepsilon^2 = O_p(n^{-2}); \quad (b) \quad \mathbb{E}^* \left(\zeta_{x^*,j}^T \xi_{x^*,j}^T \right) = 0.$$

Proof. We shall handle the case when both $\zeta_{x^*,j}^T$ and $\xi_{x^*,j}^T$ are u_j^* , being the other cases identically handled. We begin with part (a). By definition and using (2.7), it is easy to show that

$$w_{x^*,j}^T = \frac{1}{\mathfrak{G}^{d/2}} \sum_k \Psi_{\hat{\theta},k} w_{\varepsilon^*,k} \prod_{\ell=1}^d \bar{\mathcal{I}}_\ell(j, k),$$

where $\bar{\mathcal{I}}_\ell(j, k) = 2\mathcal{I}(j[\ell] = k[\ell]) - \mathcal{I}(j[\ell] - 1 = k[\ell]) - \mathcal{I}(j[\ell] + 1 = k[\ell])$. So, because $\mathbb{E}^*(w_{\varepsilon^*,j} \bar{w}_{\varepsilon^*,k}) = \hat{\sigma}_\varepsilon^2 \mathcal{I}(j = k)$, the left side of the equality in (a) is

$$\hat{\sigma}_\varepsilon^2 \left(\sum_k \left| \Psi_{\hat{\theta},j} \right|^{-2} \left| \Psi_{\hat{\theta},k} \right|^2 \left\{ \prod_{\ell=1}^d \bar{\mathcal{I}}_\ell(j, k) \right\}^2 - 1 \right)$$

From here the conclusion is standard because $|\Psi_\theta(\lambda)|^2$ is twice differentiable uniformly in $\theta \in \Theta$ for all $\lambda \in \Pi^d$ and that $\hat{\theta} - \theta_0 = o_p(1)$. Next we show part (b). That follows immediately because, say, $\mathbb{E}^*(w_{\varepsilon^*,k} w_{\varepsilon^*,k}) = 0$. \square

Lemma 12. *Let $k \prec j \in \widetilde{\mathbb{Z}}^d$. Then,*

$$(a) \quad \mathbb{E}^* \left(\zeta_{x^*,j}^T \bar{\zeta}_{x^*,k}^T \right) = \widehat{\sigma}_\varepsilon^2 (1 + O_p(n^{-1})) \mathcal{I}(|j[\ell] - k[\ell]| = 2, \ell = 1, \dots, d)$$

$$(b) \quad \mathbb{E}^* \left(\zeta_{x^*,j}^T \zeta_{x^*,k}^T \right) = 0.$$

Proof. Again, we shall handle the case when $\zeta_{x^*,j}^T$ and $\bar{\zeta}_{x^*,j}^T$ are u_j^* . We shall examine part (a) only, being part (b) identical. Proceeding as with the proof of part (a) of the previous lemma, the left side of the equality in (a) is

$$\widehat{\sigma}_\varepsilon^2 \sum_p \left| \Psi_{\widehat{\theta},j} \right|^{-2} \left| \Psi_{\widehat{\theta},p} \right|^2 \left\{ \prod_{\ell=1}^d \bar{\mathcal{I}}_\ell(j,p) \bar{\mathcal{I}}_\ell(k,p) \right\}.$$

From here, we see that the last expression is zero except when $\mathcal{I}|j[\ell] - k[\ell]| = 2$, for all $\ell = 1, \dots, d$, in which case is $\left| \Psi_{\widehat{\theta},j \pm 1} \right|^2 / \left| \Psi_{\widehat{\theta},j} \right|^2 = 1 + O_p(n^{-1})$ as $|\Psi_\theta(\lambda)|^2$ is twice differentiable uniformly in $\theta \in \Theta$ for all $\lambda \in \Pi^d$ and $\widehat{\theta} - \theta_0 = o_p(1)$. \square

Lemma 13. *Let $\zeta(\lambda)$ be as in Lemma 3. Under C1 – C4, we have that for all $r \leq s \in \widetilde{\mathbb{Z}}^d$*

$$(5.18) \quad \mathbb{E}^* \left| \sum_{j=r}^s \zeta_j v_j^* (u_j^* - v_j^*) \right|^2 = O_p \left(\frac{1}{n} \prod_{\ell=1}^d |s[\ell] - r[\ell]|_+ \right).$$

Proof. Denote $\varrho_j^* = u_j^* - v_j^*$. By standard arguments, the left side of (5.18) is

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

where, by Lemmas 11 and 12 part (b),

$$\begin{aligned} a_{j1}^* &= \mathbb{E}^* (v_j^* \bar{v}_j^*) \mathbb{E}^* (\varrho_j^* \bar{\varrho}_j^*) + \left| \mathbb{E}^* (v_j^* \bar{\varrho}_j^*) \right|^2; \\ a_{j2}^* &= \text{cum}^* (v_j^*, \bar{v}_j^*, \bar{u}_j^*, u_j^*) + \text{cum}^* (v_j^*, \bar{v}_j^*, \bar{v}_j^*, v_j^*) \\ &\quad - \text{cum}^* (v_j^*, \bar{v}_j^*, \bar{u}_j^*, v_j^*) - \text{cum}^* (v_j^*, \bar{v}_j^*, u_j^*, \bar{v}_j^*) \\ b_{jk,1}^* &= \mathbb{E}^* (v_j^* \bar{v}_k^*) \mathbb{E}^* (\varrho_j^* \bar{\varrho}_k^*) + \mathbb{E}^* (v_j^* \varrho_j^*) \mathbb{E}^* (\bar{v}_k^* \bar{\varrho}_k^*) \\ b_{jk,2}^* &= \text{cum}^* (v_j^*, \bar{v}_k^*, \bar{u}_j^*, u_k^*) + \text{cum}^* (v_j^*, \bar{v}_k^*, \bar{v}_j^*, v_k^*) \\ &\quad - \text{cum}^* (v_j^*, \bar{v}_k^*, \bar{u}_j^*, v_k^*) - \text{cum}^* (v_j^*, \bar{v}_k^*, u_j^*, \bar{v}_k^*). \end{aligned}$$

After observing that $\mathbb{E}^* (v_j^* \bar{u}_j^*) = \widehat{\sigma}_\varepsilon^2 + O_p(n^{-2})$ and $\mathbb{E}^* (v_j^* \bar{\varrho}_j^*) = \mathbb{E}^* (v_j^* \bar{u}_j^*) - \mathbb{E}^* (v_j^* \bar{v}_j^*)$, we have that Lemma 11 implies that $a_{j1}^* = O_p(n^{-2})$, whereas Lemmas 11 and 12 imply that $b_{jk,1}^* = O_p(n^{-2} + n^{-1} \mathcal{I}(|j[\ell] - k[\ell]| = 2, \ell = 1, \dots, d))$. From here it is immediate to conclude that the contribution due to a_{j1}^* and $b_{jk,1}^*$ into the left of (5.18) is its right side.

Finally we examine a_{j2}^* and $b_{jk,2}^*$. By definition of, say, $w_{x^*}^T(\lambda)$ and that $\text{cum}^*(\varepsilon^*(t_1), \dots, \varepsilon^*(t_4)) = \widehat{\kappa}_\varepsilon \mathcal{I}(t_1 = \dots = t_4)$, it is obvious that $b_{jk,2}^* = O_p(n^{-1} \mathcal{I}(|j[\ell] - k[\ell]| = 2, \ell = 1, \dots, d))$. Notice that $\widehat{\kappa}_\varepsilon = N^{-1} \sum_{t=1}^n \widehat{\varepsilon}^4(t) - 3 \left(N^{-1} \sum_{t=1}^n \widehat{\varepsilon}^2(t) \right)^2 = O_p(1)$ by Lemma 10. \square

Lemma 14. *Let $\zeta(\lambda)$ be a function as in Lemma 3. Then, under $C1 - C4$,*

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

Proof. We shall consider the proof in the positive quadrant $\sum_{j=1}^{\lceil \tilde{n}\lambda/\pi \rceil}$, being the proof for the remaining $2^{d-1} - 1$ quadrants similarly handled. By the Cauchy-Schwarz and triangle inequalities, it suffices to show that

$$(5.19) \quad \mathbb{E}^* \sup_s \left| \sum_{j=1}^s \zeta_j \left\{ \frac{I_{x^*,j}^T}{|\Psi_{\hat{\theta},j}|^2} - I_{\varepsilon^*,j}^T \right\} \right| \leq \mathbb{E}^* \sup_s \sum_{j=1}^s |\zeta_j| |\varrho_j^*|^2 + 2\mathbb{E}^* \sup_s \left| \sum_{j=1}^s \zeta_j v_j^* \bar{\varrho}_j^* \right|$$

is $o_p(N^{1/2})$, where we abbreviate “ $\sup_{s=1, \dots, \tilde{n}}$ ” by “ \sup_s ” and $\varrho_j^* = u_j^* - v_j^*$.

The first term on the right of (5.19) is bounded by

$$C \sum_{j=1}^{\tilde{n}} \left\{ \left(\mathbb{E}^* |u_j^*|^2 - \hat{\sigma}_\varepsilon^2 \right) - \left(\mathbb{E}^* (u_j^* \bar{v}_j^*) - \hat{\sigma}_\varepsilon^2 \right) - \left(\mathbb{E}^* (\bar{u}_j^* v_j^*) - \hat{\sigma}_\varepsilon^2 \right) + \left(\mathbb{E}^* |v_j^*|^2 - \hat{\sigma}_\varepsilon^2 \right) \right\} = o_p(N^{1/2}),$$

because $|\zeta_j| \leq C$, $d < 4$ and by Lemma 11, say,

$$\left| \mathbb{E}^* \left(u_j^* \begin{pmatrix} \bar{v}_j^* \\ \bar{u}_j^* \end{pmatrix} \right) - \hat{\sigma}_\varepsilon^2 \right| \leq \hat{\sigma}_\varepsilon^{-2} |\Psi_{\hat{\theta},j}|^{-2} \left(\left| \mathbb{E}^* \left(w_{x^*,j} \begin{pmatrix} \bar{w}_{\varepsilon^*,j} \\ \bar{w}_{x^*,j} \end{pmatrix} \right) - \hat{\sigma}_\varepsilon^2 \right| \left| \frac{\Psi_{\hat{\theta},j}}{|\Psi_{\hat{\theta},j}|^2} \right| \right) = O_p(n^{-2}).$$

Next, we examine the second term of (5.19). With same notation as in Lemma 4, the square of the second term on the right of (5.19) is bounded by

$$(5.20) \quad \mathbb{E}^* \max_s \left| \left\{ \sum_{j=1}^s - \sum_{j=1}^{q(s)[\tilde{n}^{1-\varsigma}]} \right\} \zeta_j v_j^* \bar{\varrho}_j^* \right|^2 + \mathbb{E}^* \max_s \left| \sum_{j=1}^{q(s)[\tilde{n}^{1-\varsigma}]} \zeta_j v_j^* \bar{\varrho}_j^* \right|^2.$$

From the definition of $q(s)$ and (5.6), the second term of (5.20) is bounded by

$$\sum_{q=1}^{\lceil \tilde{n}^\varsigma \rceil - 1} \mathbb{E}^* \left| \sum_{j=1}^{q[\tilde{n}^{1-\varsigma}]} \zeta_j v_j^* \bar{\varrho}_j^* \right|^2 = O_p(N^{1+\varsigma} n^{-1} \log^2 n) = o_p(N)$$

by Lemma 13 and because $\varsigma < 1/d$. To complete the proof we need to show that the first term in (5.20) is $o_p(N)$. To that end, we note that it is bounded by

$$\mathbb{E}^* \max_{q=1, \dots, \lceil \tilde{n}^\varsigma \rceil - 1} \max_{s=1+q[\tilde{n}^{1-\varsigma}], \dots, (q+1)[\tilde{n}^{1-\varsigma}]} \left| \sum_{j=1+q[\tilde{n}^{1-\varsigma}]}^s \zeta_j v_j^* \bar{\varrho}_j^* \right|^2$$

which is $O_p(N^\varsigma) \mathbb{E}^* \max_{s=1, \dots, \lceil \tilde{n}^{1-\varsigma} \rceil} \left| \sum_{j=1}^s \zeta_j v_j^* \bar{\varrho}_j^* \right|^2$.

So, we have that the square of the second term on the right of (5.19) is

$$O_p(N^{1+\varsigma} n^{-1}) + O_p(N^\varsigma) \mathbb{E}^* \max_{s=1, \dots, \lceil \tilde{n}^{1-\varsigma} \rceil} \left| \sum_{j=1}^s \zeta_j v_j^* \bar{\varrho}_j^* \right|^2.$$

Now proceeding as in Lemma 4, the square of the second term on the right of (5.19), is

$$\begin{aligned} & O_p \left(\left\{ N^{1+\varsigma} + \sum_{q=0}^{\iota-1} N^{(1+\varsigma)(1-\varsigma)^{q+1} + \varsigma \sum_{p=0}^q (1-\varsigma)^p} \right\} n^{-1} \right) \\ & + N^{\varsigma \sum_{p=0}^{\iota-1} (1-\varsigma)^p} \mathbb{E}^* \max_{s=1, \dots, (\tilde{n}^{1-\varsigma})^\iota} \left| \sum_{j=1}^s \zeta_j v_j^* \bar{\varrho}_j^* \right|^2 \\ & = O_p(N^{1+\varsigma} n^{-1}) + O_p \left(N^{\varsigma \sum_{p=0}^{\iota-1} (1-\varsigma)^p} \sum_{s=1}^{\tilde{n}^{(1-\varsigma)^\iota}} \mathbb{E}^* \left| \sum_{j=1}^s \zeta_j v_j^* \bar{\varrho}_j^* \right|^2 \right) = o_p(N) \end{aligned}$$

after choosing ι large enough because $\varsigma < 1/d$. This completes the proof. \square

Lemma 15. *Let $\zeta(\lambda; \vartheta)$ be as in Lemma 5. Assuming C1 – C4,*

$$(5.21) \quad \frac{1}{N} \sup_{\vartheta \in \Theta \times \mathbb{R}^+} \left| \sum_{j=-\tilde{n}}^{\tilde{n}} \zeta_j(\vartheta) \left(\frac{I_{x^*,j}^T}{|\Psi_{\hat{\theta},j}|^2} - \hat{\sigma}_\varepsilon^2 \right) \right| = o_{p^*}(1).$$

Proof. By the triangle inequality, the left side of (5.21) is bounded by

$$(5.22) \quad \frac{C}{N} \sup_{\vartheta \in \Theta \times \mathbb{R}^+} \left| \sum_{j=-\tilde{n}}^{\tilde{n}} \zeta_j(\vartheta) \left(\frac{I_{x^*,j}^T}{|\Psi_{\hat{\theta},j}|^2} - I_{\varepsilon^*,j}^T \right) \right| + \frac{C}{N} \sup_{\vartheta \in \Theta \times \mathbb{R}^+} \left| \sum_{j=-\tilde{n}}^{\tilde{n}} \zeta_j(\vartheta) \left(I_{\varepsilon^*,j}^T - \hat{\sigma}_\varepsilon^2 \right) \right|.$$

Now, because by assumption $|\zeta(\lambda; \vartheta)| < C$, the first term of (5.22) is bounded by

$$\frac{C}{N} \sum_{j=-\tilde{n}}^{\tilde{n}} \left| \frac{I_{x^*,j}^T}{|\Psi_{\hat{\theta},j}|^2} - I_{\varepsilon^*,j}^T \right| \leq \frac{C}{N} \sum_{j=-\tilde{n}}^{\tilde{n}} |u_j^* - v_j^*|^2 + \frac{C}{N} \sum_{j=-\tilde{n}}^{\tilde{n}} |v_j^* (\bar{u}_j^* - \bar{v}_j^*)| = o_{p^*}(1)$$

by Markov's inequality because by the Cauchy-Schwarz inequality $\mathbb{E}^* |v_j^* (\bar{u}_j^* - \bar{v}_j^*)|^2 \leq \mathbb{E}^* |v_j^*|^2 \mathbb{E}^* |\bar{u}_j^* - \bar{v}_j^*|^2$ and then proceeding as in Lemma 13. Next, the second term of (5.22) is $o_{p^*}(1)$. First, the finite dimensional distributions of $S^*(\vartheta) = N^{-1} \sum_{j=-\tilde{n}}^{\tilde{n}} \zeta_j(\vartheta) \left(I_{\varepsilon^*,j}^T - \hat{\sigma}_\varepsilon^2 \right)$ converge to zero in probability. Indeed, the second bootstrap moment is

$$\frac{1}{N^2} \sum_{j,k=-\tilde{n}}^{\tilde{n}} \zeta_j(\vartheta) \zeta_k(\vartheta) \mathbb{E}^* \left\{ \left(I_{\varepsilon^*,j}^T - \hat{\sigma}_\varepsilon^2 \right) \left(I_{\varepsilon^*,k}^T - \hat{\sigma}_\varepsilon^2 \right) \right\} = O_p(N^{-1})$$

by standard algebra after observing that $\mathbb{E}^*(w_{\varepsilon^*,j} \bar{w}_{\varepsilon^*,k}) = \hat{\sigma}_\varepsilon^2 \mathcal{I}(j=k)$, $\mathbb{E}^*(w_{\varepsilon^*,j} w_{\varepsilon^*,k}) = 0$ and $\text{cum}^*(w_{\varepsilon^*,j}, \bar{w}_{\varepsilon^*,j}, w_{\varepsilon^*,k}, \bar{w}_{\varepsilon^*,k}) = O(N^{-1}) \hat{\kappa}_\varepsilon \mathcal{I}(j=k)$, where $\hat{\kappa}_\varepsilon$ is the standard estimator of the fourth cumulant κ_ε . To finish, we need to show the tightness of the process $S^*(\vartheta)$. But this is immediate because

$$\begin{aligned} N^{-2} \mathbb{E}^* |S^*(\vartheta_2) - S^*(\vartheta_1)|^2 &= N^{-1} \sum_{j=-\tilde{n}}^{\tilde{n}} |\zeta_j(\vartheta_2) - \zeta_j(\vartheta_1)|^2 \mathbb{E}^* \left| I_{\varepsilon^*,k}^T - \hat{\sigma}_\varepsilon^2 \right|^2 \\ &= |\vartheta_2 - \vartheta_1|^2 O_p(1) \end{aligned}$$

by continuous differentiability of $\zeta(\lambda; \vartheta)$ for all $\lambda \in \Pi^d$. \square

Lemma 16. *Assume C1 – C3 and C5 – C8. Then,*

$$\hat{\vartheta}^* - \hat{\vartheta} = o_{p^*}(1).$$

Proof. The proof follows very easily using Lemma 15. Indeed, (4.5) is

$$(5.23) \quad \frac{1}{N} \sum_{j=-\tilde{n}}^{\tilde{n}} \frac{f_{\hat{\vartheta},j}}{f_{\vartheta,j}} \left(\frac{I_{x^*,j}^T}{(2\pi)^d f_{\hat{\vartheta},j}} - 1 \right) + \frac{1}{N} \left\{ \sum_{j=-\tilde{n}}^{\tilde{n}} \frac{f_{\hat{\vartheta},j}}{f_{\vartheta,j}} - \log \frac{f_{\hat{\vartheta},j}}{f_{\vartheta,j}} + \log f_{\hat{\vartheta},j} \right\}.$$

Now, the difference between the second term of (5.23) and

$$\int_{\pi}^{\pi} \left\{ \frac{f_{\hat{\vartheta}}(\lambda)}{f_{\vartheta}(\lambda)} - \log \left(\frac{f_{\hat{\vartheta}}(\lambda)}{f_{\vartheta}(\lambda)} \right) \right\} d\lambda + \int_{\pi}^{\pi} \log f_{\hat{\vartheta}}(\lambda) d\lambda$$

converges to zero in probability using Brillinger (1981, p.15) and that uniformly in λ , $|f_{\hat{\vartheta}}(\lambda) - f_{\vartheta_0}(\lambda)| = o_p(1)$. Moreover, the last displayed expression is greater than or equal to $\frac{(2\pi)^d}{2} + \int_{\pi}^{\pi} \log f_{\hat{\vartheta}}(\lambda) d\lambda$ with equality when $f_{\hat{\vartheta}}(\lambda) = f_{\vartheta}(\lambda)$ which is the case only if $\vartheta = \hat{\vartheta}$ by C7. On the other hand, the first term of (5.23) converges to zero uniformly in ϑ by Lemma 15 because $f_{\vartheta,j}^{-1} f_{\hat{\vartheta},j}$ satisfies the same conditions as $\zeta(\lambda; \vartheta)$ there by C6. From here the conclusion of the lemma is standard proceeding as in Theorem 1 of Hannan (1973), so we omit its details. \square

Lemma 17. *Assume C1 – C8. Under H_0 ,*

$$(5.24) \quad \frac{1}{N^{1/2}} \sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]} \zeta_j \left(\frac{I_{x^*,j}^T}{|\Psi_{\hat{\theta}^*,j}|^2} - \hat{\sigma}_{\varepsilon}^2 \right) = \frac{1}{N^{1/2}} \sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]} \zeta_j \left(I_{\varepsilon^*,j}^T - \hat{\sigma}_{\varepsilon}^2 \right) - \left(\frac{1}{N} \sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]} \zeta_j \varphi'_{\hat{\theta},j} \right) N^{1/2} (\hat{\theta}^* - \hat{\theta}) + o_{p^*}(1),$$

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 (5.24) and the first term on its right side is

$$(5.25) \quad \frac{1}{N^{1/2}} \sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]} \zeta_j \frac{I_{x^*,j}^T}{|\Psi_{\hat{\theta},j}|^2} \left[\frac{|\Psi_{\hat{\theta},j}|^2}{|\Psi_{\hat{\theta}^*,j}|^2} - 1 + \varphi'_{\hat{\theta},j} (\hat{\theta}^* - \hat{\theta}) \right] + \frac{1}{N^{1/2}} \sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]} \zeta_j \left(\frac{I_{x^*,j}^T}{|\Psi_{\hat{\theta},j}|^2} - I_{\varepsilon^*,j}^T \right) - \frac{1}{N^{1/2}} \sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]} \zeta_j \varphi'_{\hat{\theta},j} \frac{I_{x^*,j}^T}{|\Psi_{\hat{\theta},j}|^2} (\hat{\theta}^* - \hat{\theta}).$$

First, because each component of the vector $\zeta(\lambda) \varphi_{\hat{\theta}}(\lambda)$ satisfies the same conditions of $\zeta(\lambda)$ in Lemma 14, Markov's inequality implies that the second term of (5.25) is $o_{p^*}(1)$, whereas the third term is $N^{-1} \sum_{j=1}^{[\tilde{n}\lambda/\pi]} \zeta_j \varphi'_{\hat{\theta},j} N^{1/2} (\hat{\theta}^* - \hat{\theta}) + o_{p^*}(1)$ by Lemma 14 and because proceeding as in the proof of Theorem 4

$$\frac{1}{N^{1/2}} \sum_{j=-\tilde{n}}^{\tilde{n}} \zeta_j \varphi_{\hat{\theta},j} \left(I_{\varepsilon^*,j}^T - \hat{\sigma}_{\varepsilon}^2 \right) = O_{p^*}(1)$$

Finally, by mean value theorem, the norm of the first term of (5.25) is bounded by

$$CN^{1/2} \left\| \hat{\theta}^* - \hat{\theta} \right\|^2 \frac{1}{N} \sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]} \frac{I_{x^*,j}^T}{|\Psi_{\hat{\theta}^*,j}|^2} = O_{p^*}(N^{-1/2}),$$

by Theorem 4 and proceeding as with the second term of (5.25). This concludes the proof. \square

We now introduce the following notation. For $v_1, v_2 \in [0, \pi]^d$, denote $\mathcal{E}_{1,N}^*(v_1, v_2)$ and $\mathcal{E}_{2,N}^*(v_1, v_2)$ as in (5.13) and (5.14) but with $\varepsilon^T(t)$ there replaced by $\varepsilon^{*T}(t) = h(t)\varepsilon^*(t)$ and also let $H_N = H_N(v_1, v_2)$ a sequence bounded in probability.

Lemma 18. *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,*

$$\mathbb{E}^* \left(|\mathcal{E}_{j,N}^*(v_1, v)|^\beta |\mathcal{E}_{j,N}^*(v, v_2)|^\beta \right) \leq H_N(v_1, v_2) \prod_{\ell=1}^d (v_2[\ell] - v_1[\ell])^2, \quad j = 1, 2.$$

Proof. The proof proceeds as in Lemma 8 but instead of using Delgado et al.' (2005) Lemma 5 we use Lemma 7.3 of Hidalgo and Kreiss (2006). \square

Next we will show that the processes $\mathcal{E}_{1,N}^*(0, \lambda)$ and $\mathcal{E}_{2,N}^*(0, \lambda)$ are tight. To that end, it suffices to show the following lemma.

Lemma 19. *Assuming C1 we have that*

(5.26)

$$(a) \quad \mathbb{E}^* \prod_{\ell=1}^d \left(\mathcal{E}_{1,N}^{(\ell)*}(0, \lambda_{1[\ell]}) - \mathcal{E}_{1,N}^{(\ell)*}(0, \lambda_{2[\ell]}) \right)^2 = H_N(\lambda_1, \lambda_2) \prod_{\ell=1}^d (\lambda_{2[\ell]} - \lambda_{1[\ell]})^2$$

(5.27)

$$(b) \quad \mathbb{E}^* \prod_{\ell=1}^d \left(\mathcal{E}_{2,N}^{(\ell)*}(0, \lambda_{1[\ell]}) - \mathcal{E}_{2,N}^{(\ell)*}(0, \lambda_{2[\ell]}) \right)^4 = H_N(\lambda_1, \lambda_2) \prod_{\ell=1}^d (\lambda_{2[\ell]} - \lambda_{1[\ell]})^2$$

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

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

Proof. The proof follows after observing that $\mathcal{E}_{j,N}^{(\ell)*}(0, \lambda_{1[\ell]}) - \mathcal{E}_{j,N}^{(\ell)*}(0, \lambda_{2[\ell]}) = \mathcal{E}_{j,N}^{(\ell)*}(\lambda_{1[\ell]}, \lambda_{2[\ell]})$ for $j = 1, 2$ and then by Lemma 18. \square

6. PROOFS

6.1. Proof of Proposition 1.

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

$$\mathbf{S}_N(\mu) = \frac{1}{N^{1/2}} \sum_{j=-[\tilde{n}\mu/\pi]}^{[\tilde{n}\mu/\pi]} \zeta_j (I_{\varepsilon,j}^T - 1) \Rightarrow \mathbf{B}_\zeta(\mu), \quad \mu \in [0, \pi]^d$$

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

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

$$(6.1) \quad \frac{1}{N^{1/2}} \sum_{j=-[\tilde{n}\mu/\pi]}^{[\tilde{n}\mu/\pi]} \zeta_j (I_{\varepsilon,j}^T - 1) \xrightarrow{d} \mathcal{N} \left(0, \left(2 + \kappa_\varepsilon \left(\frac{35}{32} \right)^d \right) \int_{-\mu}^\mu \zeta(\lambda) \zeta'(\lambda) d\lambda \right).$$

(b) For $\mu \leq v$, the covariance structure is such that

$$\frac{1}{N} E \left\{ \sum_{j=-[\tilde{n}\mu/\pi]}^{[\tilde{n}\mu/\pi]} \zeta_j (I_{\varepsilon,j}^T - 1) \sum_{j=-[\tilde{n}v/\pi]}^{[\tilde{n}v/\pi]} \zeta_j (I_{\varepsilon,j}^T - 1) \right\} \rightarrow \left(2 + \kappa_\varepsilon \left(\frac{35}{32} \right)^d \right) \int_{-\mu}^\mu \zeta(\lambda) \zeta'(\lambda) d\lambda.$$

(c) the process $\{\mathbf{S}_N(\mu) : \mu \in [0, \pi]^d\}$ 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), $N^{-1} \sum_{j=-[\tilde{n}v/\pi]}^{[\tilde{n}v/\pi]} \zeta_j \zeta'_j - (2\pi)^{-d} \int_{-v}^v \zeta(\lambda) \zeta'(\lambda) d\lambda = O_p(n^{-1})$, and thus it is omitted.

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

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

using that, say $\sum_{p[\ell]=1}^{n[\ell]} h_\ell(t[\ell]) e^{-ip[\ell](\lambda_{j[\ell]} \pm k[\ell])} = n[\ell] \mathcal{I}(j[\ell] \pm k[\ell] = 0, n[\ell])$ and that by Brillinger (1981, p.15) we have that $N^{-1} \sum_{t=1}^{\tilde{n}} h^p(t) \rightarrow 2^d \int_{[0,1]^d} h^p(u) du$ for all $p \geq 0$. Finally, part (c) follows by Lemma 9.

6.2. Proof of Theorem 3.

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

$$\begin{aligned} (6.2) \quad N^{1/2} \left(G_{\hat{\theta},N}(\lambda) - G_N^0(\lambda) \right) &= - \left(\frac{1}{N} \sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]} \tilde{\varphi}'_{\theta_0,N}(j) \right) N^{1/2} (\hat{\theta} - \theta_0) + o_p(1) \\ &= - \left(\frac{1}{N} \sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]} \tilde{\varphi}'_{\theta_0,N}(j) \right) \tilde{\Lambda}_{\theta_0,N}^{-1} \frac{1}{N^{1/2}} \sum_{k=-\tilde{n}}^{\tilde{n}} \tilde{\varphi}_{\theta_0,N}(k) \frac{I_{x,k}^T}{|\Psi_{\theta_0,k}|^2} \\ &\quad + o_p(1), \end{aligned}$$

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

But using (2.14) and that $|G_{\theta_0,N}(\pi) - G_N^0(\pi)| = o_p(N^{-1/2})$ by Lemma 4, then by (6.2), uniformly in λ , we obtain that

$$\begin{aligned} (6.3) \quad \alpha_{\hat{\theta},N}(\lambda) &= \alpha_N^0(\lambda) + \frac{N^{1/2} \left(G_{\hat{\theta},N}(\lambda) - G_N^0(\lambda) \right)}{G_N^0(\pi)} \\ &\quad + G_{\hat{\theta},N}(\lambda) N^{1/2} \left(\frac{1}{G_{\hat{\theta},N}(\pi)} - \frac{1}{G_N^0(\pi)} \right) \\ &= \alpha_N^0(\lambda) - \left(\frac{1}{N} \sum_{j=-[\tilde{n}\lambda/\pi]}^{[\tilde{n}\lambda/\pi]} \tilde{\varphi}'_{\theta_0,N}(j) \right) \tilde{\Lambda}_{\theta_0,N}^{-1} \frac{1}{N^{1/2}} \sum_{k=-\tilde{n}}^{\tilde{n}} \tilde{\varphi}_{\theta_0,N}(k) I_{\varepsilon,k}^T + o_p(1), \end{aligned}$$

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

Next part (b). Taking into account part (a), part (b) follows because Proposition 1 guarantees the fidi's convergence of α_N^0 and its tightness. Tightness of the second term on the right of (6.3) follows by Lemmas 8 and 9 because $\varphi_{\theta_0}(u)$ is continuously differentiable in θ and u . This concludes the proof. \square

6.3. Proof of Theorem 4.

First, by standard algebra

$$\hat{\vartheta}^* - \hat{\vartheta} = \overline{Q_{\hat{\vartheta},N}^*}^{-1} q_{\hat{\vartheta},N}^*,$$

where $\widehat{\vartheta}^*$ is an intermediate point between $\widehat{\vartheta}$ and $\widehat{\vartheta}^*$,

$$q_{\widehat{\vartheta},N}^* = \frac{1}{N} \sum_{j=-\widehat{n}}^{\widehat{n}} \phi_{\widehat{\vartheta},j} \left\{ \frac{I_{x^*,j}^T}{\sigma_\varepsilon^2 |\Psi_{\theta,j}|^2} - 1 \right\}$$

and $\overline{Q}_{\widehat{\vartheta},N}^*$ is as defined in the proof of Theorem 2 but with $I_{x,j}^T$ replaced by $I_{x^*,j}^T$ there. Now, because $f_{\widehat{\vartheta},j}$ is twice continuously differentiable by C6, and that $\widehat{\vartheta}^* - \widehat{\vartheta} = o_{p^*}(1)$ by Lemma 16, we easily conclude that

$$\left| \overline{Q}_{\widehat{\vartheta}^*,N}^* - Q_{\widehat{\vartheta}^*,N} \right| = o_{p^*}(1),$$

and that $\left| Q_{\widehat{\vartheta}^*,N} - Q_{\widehat{\vartheta}_0,N} \right| = o_{p^*}(1)$. On the other hand, by Lemma 14 with $\zeta_j = \phi_{\widehat{\vartheta},j}$ there,

$$q_{\widehat{\vartheta},N}^* = N^{-1/2} \sum_{j=-\widehat{n}}^{\widehat{n}} \phi_{\widehat{\vartheta},j} \left\{ I_{\varepsilon^*,j}^T / \widehat{\sigma}_\varepsilon^2 - 1 \right\} + o_{p^*}(1).$$

So, to complete the proof it suffices to show

$$(a) \quad \frac{1}{N} \mathbb{E}^* \left\{ \sum_{j=-\widehat{n}}^{\widehat{n}} \phi_{\widehat{\vartheta},j} \left\{ \frac{I_{\varepsilon^*,j}^T}{\widehat{\sigma}_\varepsilon^2} - 1 \right\} \sum_{j=-\widehat{n}}^{\widehat{n}} \phi'_{\widehat{\vartheta},j} \left\{ \frac{I_{\varepsilon^*,j}^T}{\widehat{\sigma}_\varepsilon^2} - 1 \right\} \right\} \\ \xrightarrow{P} 2 + \kappa_\varepsilon \left(\frac{35}{32} \right)^d \int_{-\mu}^{\mu} \phi_{\widehat{\vartheta}_0}(\lambda) \phi'_{\widehat{\vartheta}_0}(\lambda) d\lambda,$$

and then (b) the Lindeberg's condition

$$\mathbb{E}^* \left\{ \frac{1}{N} \sum_{j=-\widehat{n}}^{\widehat{n}} \left| \frac{I_{\varepsilon^*,j}^T}{\widehat{\sigma}_\varepsilon^2} - 1 \right|^2 \mathcal{I} \left(\left| \frac{I_{\varepsilon^*,j}^T}{\widehat{\sigma}_\varepsilon^2} - 1 \right| > \delta N^{-1} \right) \right\} \xrightarrow{P} 0.$$

We begin with (a). By Lemmas 11 and 12, we have that

$$\mathbb{E}^* \left\{ \left(\frac{I_{\varepsilon^*,j}^T}{\widehat{\sigma}_\varepsilon^2} - 1 \right) \left(\frac{I_{\varepsilon^*,k}^T}{\widehat{\sigma}_\varepsilon^2} - 1 \right) \right\} = \mathcal{I}(j = k \pm 2) + cum^*(w_{\varepsilon^*,j}, \overline{w}_{\varepsilon^*,j}, w_{\varepsilon^*,k}, \overline{w}_{\varepsilon^*,k}).$$

Now, because $cum^*(\varepsilon^*(t_1), \varepsilon^*(t_2), \varepsilon^*(t_3), \varepsilon^*(t_4)) = \widehat{\kappa}_\varepsilon \mathcal{I}(t_1 = \dots = t_4)$, we have that the left side in (a) is

$$\frac{1}{N} \sum_{j=-\widehat{n}}^{\widehat{n}} \phi_{\widehat{\vartheta},j} \phi'_{\widehat{\vartheta},j} \left(2 + \widehat{\kappa}_\varepsilon \left(\frac{35}{32} \right) \right).$$

From here we conclude (a) by Lemma 10 and that $\phi_{\widehat{\vartheta}}(\lambda)$ is continuous by C6.

We now show (b). By standard inequalities, the left side is bounded by

$$\frac{1}{N^2} \sum_{j=-\widehat{n}}^{\widehat{n}} \mathbb{E}^* \left| \frac{I_{\varepsilon^*,j}^T}{\widehat{\sigma}_\varepsilon^2} - 1 \right|^4 \leq C \frac{\log^2 N}{N^2} \sum_{j=-\widehat{n}}^{\widehat{n}} \mathbb{E}^* \left| \frac{I_{\varepsilon^*,j}^T}{\widehat{\sigma}_\varepsilon^2} \right|^2$$

by An et al. (1983) because $\{\varepsilon^*(t)\}_{t=1}^n$ is a random sample. Now conclude part (b) since $\mathbb{E}^* \left| \frac{I_{\varepsilon^*,j}^T}{\widehat{\sigma}_\varepsilon^2} \right|^2 = O_p(1)$ by Lemma 10. This concludes the proof of the theorem. \square

6.4. Proof of Theorem 5.

Part (a). By Lemma 17 with $\zeta(\lambda) = 1$ there and the definitions of $G_{\hat{\theta},N}^*(\lambda)$ and $G_N^{*0}(\lambda)$, we have that

$$\begin{aligned}
 N^{1/2} \left(G_{\hat{\theta},N}^*(\lambda) - G_N^{*0}(\lambda) \right) &= - \left(\frac{1}{N} \sum_{j=-[\bar{n}\lambda/\pi]}^{[\bar{n}\lambda/\pi]} \tilde{\varphi}'_{\hat{\theta},N}(j) \right) N^{1/2} \left(\hat{\theta}^* - \hat{\theta} \right) + o_{p^*}(1) \\
 (6.4) \qquad \qquad \qquad &= - \left(\frac{1}{N} \sum_{j=-[\bar{n}\lambda/\pi]}^{[\bar{n}\lambda/\pi]} \tilde{\varphi}'_{\hat{\theta},N}(j) \right) \tilde{\Lambda}_{\hat{\theta},N}^{-1} \frac{1}{N^{1/2}} \sum_{k=-\bar{n}}^{\bar{n}} \tilde{\varphi}_{\hat{\theta},N}(k) \frac{I_{x^*,k}^T}{\left| \Psi_{\hat{\theta},k} \right|^2} \\
 &\quad + o_{p^*}(1),
 \end{aligned}$$

by Theorem 4, where the $o_{p^*}(1)$ is uniform in $\lambda \in [0, \pi]^d$.

Now because $\left| G_{\hat{\theta},N}^*(\pi) - G_N^{*0}(\pi) \right| = o_{p^*}(N^{-1/2})$ by Lemma 14, by (6.4), we obtain that

$$\begin{aligned}
 \alpha_{\hat{\theta},N}^*(\lambda) &= \alpha_N^{*0}(\lambda) + \frac{N^{1/2} \left(G_{\hat{\theta},N}^*(\lambda) - G_N^{*0}(\lambda) \right)}{G_N^{*0}(\pi)} \\
 (6.5) \qquad &\quad + G_{\hat{\theta},N}^*(\lambda) N^{1/2} \left(\frac{1}{G_{\hat{\theta},N}^*(\pi)} - \frac{1}{G_N^{*0}(\pi)} \right) \\
 &= \alpha_N^{*0}(\lambda) - \left(\frac{1}{N} \sum_{j=-[\bar{n}\lambda/\pi]}^{[\bar{n}\lambda/\pi]} \tilde{\varphi}'_{\hat{\theta},N}(j) \right) \tilde{\Lambda}_{\hat{\theta},N}^{-1} \frac{1}{N^{1/2}} \sum_{k=-\bar{n}}^{\bar{n}} \tilde{\varphi}_{\hat{\theta},N}(k) I_{\varepsilon^*,k}^T + o_{p^*}(1),
 \end{aligned}$$

uniformly in λ . Now conclude the proof of part (a) by observing that Lemma 10 implies that $\mathbb{E}^* \left| G_N^{*0}(\pi) - \hat{\sigma}_\varepsilon^2 \right| = o_p(1)$ and then by Markov's inequality.

Next part (b). Taking into account part (a), part (b) follows because Theorem 4 guarantees the fidi's convergence of α_N^{*0} , whereas its tightness follows by Lemma 19. Tightness of the second term on the right of (6.5) follows by Lemmas 18 and 19 because $\varphi_\theta(u)$ is continuously differentiable in θ and u . This concludes the proof. \square

REFERENCES

- [1] AN, H.-Z., CHEN, Z.G. AND HANNAN, E.J. (1983): "The maximum of the periodogram," *Journal of Multivariate Analysis*, **13**, 383-400.
- [2] ANDERSON, T.W. (1997): "Goodness-of-fit for autoregressive processes," *Journal of Time Series Analysis*, **18**, 321-339.
- [3] ANDERSON, T.W. AND WALKER A.M. (1964): "On the asymptotic distribution of the autocorrelations of a sample from a linear stochastic process," *Annals of Mathematical Statistics*, **35**, 1296-1303.
- [4] 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.
- [5] BICKEL, P.J. AND WICHURA, M.J. (1971): "Convergence criteria for multiparameter stochastic processes and its applications," *Annals of Mathematical Statistics*, **42**, 1656-1670.
- [6] BILLINGSLEY, P. (1968): *Convergence of probability measures*. Wiley and Sons. New York.
- [7] BLOOMFIELD, P. (1973): "An exponential model for the spectrum of a scalar time series," *Biometrika*, **61**, 217-226.
- [8] 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.
- [9] BRILLINGER, D.R. (1981): *Time series, data analysis and theory*. Holden-Day. San Francisco.
- [10] 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.
- [11] CRESSIE, N.A.C. (1993): *Statistics for spatial data*. Wiley. New York.

- [12] CRUJEIRAS, R.M., FERNANDEZ-CASAL, R. AND GONZALEZ-MANTEIGA, W. (2006): "Goodness-of-fit tests for the spatial spectral density," *Preprint*. Universidad de Santiago de Compostela.
- [13] DAHLHAUS, R., AND KÜNSCH, H. (1986): "Edge effect and efficient parameter estimation for stationary random fields," *Biometrika*, **74**, 877-882.
- [14] DELGADO, M.A., HIDALGO, J. AND VELASCO, C. (2005): "Distribution free goodness-of-fit tests for linear processes," *Annals of Statistics*, **33**, 2568-2609.
- [15] DIBLASI, A. AND BOWMAN, A.W. (2001): "On the use of the variogram in checking the independence in spatial data," *Biometrics*, **57**, 211-218.
- [16] DURBIN, J. (1973): "Weak convergence of the sample distribution function when parameters are estimated," *Annals of Statistics*, **1**, 274-290.
- [17] 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.
- [18] 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.
- [19] FRANKE, J. AND HÄRDLE, W. (1992): "On bootstrapping kernel spectral estimate," *Annals of Statistics*, **20**, 121-145.
- [20] GIRAITIS, L., HIDALGO, J. AND ROBINSON, P.M. (2001): "Gaussian estimation of parametric spectral density with unknown pole," *Annals of Statistics*, **29**, 987-1023.
- [21] GRENANDER, U. (1950): "Stochastic processes and statistical inference," *Ark. Mat.*, **1**, 195-277.
- [22] GRENANDER, U. AND ROSENBLATT, M. (1957): *Statistical Analysis of Stationary Time Series*. John Wiley, New York.
- [23] GUYON, X. (1982a): "Parameter estimation for a stationary process on a d -dimension lattice," *Biometrika*, **69**, 95-105.
- [24] GUYON, X. (1982b): *Random fields on a network*. Springer-Verlag. Berlin.
- [25] HANNAN, E.J. (1973): "The asymptotic theory of linear time series models," *Journal of Applied Probability*, **10**, 130-145.
- [26] HAINZ, G. AND DAHLHAUS, R. (2000): "Spectral domain bootstrap tests for stationary time series." Preprint.
- [27] HIDALGO, J. (2003): "An alternative bootstrap to moving blocks for time series regression models," *Journal of Econometrics*, **117**, 369-399.
- [28] HIDALGO, J. (2007): "Uniform confidence bands for kernel regression estimates with dependent data," Forthcoming in *Journal of Econometrics*.
- [29] HIDALGO, J. AND LAZAROVA, S. (2007): "Bootstrap in time series regression models," Preprint.
- [30] HIDALGO, J. AND KREISS, J.P. (2006): "Bootstrap specification tests for linear covariance stationary processes," *Journal of Econometrics*, **133**, 807-839.
- [31] HONG, Y. (1996). "Consistent testing for serial-correlation of unknown form," *Econometrica*, **64**, 837-864.
- [32] LJUNG, G.M. AND BOX G.E.P. (1978): "On a measure of lack of fit in time series models," *Biometrika*, **65**, 297-303.
- [33] NEYMAN, J. (1937): "Smooth test for goodness of fit," *Skand. Aktuarietidskr.*, **20**, 150-199.
- [34] PAPANODITIS, E. (2000): "Spectral density based goodness-of-fit tests for time series models," *Scandinavian Journal of Statistics*, **27**, 143-176.
- [35] ROBINSON, P.M. AND VIDAL-SANZ, J. (2006): "Modified Whittle estimation of multilateral models on a lattice," *Journal of Multivariate Analysis*, **60**, 234-349.
- [36] STUTE, W. (1997): "Nonparametric model checks for regression," *Annals of Statistics*, **25**, 613-641.
- [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