UNIVERSIDADE DE
SANTIAGO DE COMPOSTELA
DEPARTAMENTO DE
ESTATÍSTICA E INVESTIGACIÓN OPERATIVA

Goodness-of-fit tests for the spatial spectral density

R. M. Crujeiras, R. Fernández-Casal, W. González-Manteiga

Report 07-01

Reports in Statistics and Operations Research

Goodness-of-fit tests for the spatial spectral density

Rosa M. Crujeiras*
Dept. of Statistics and OR
University of Santiago de Compostela

Rubén Fernández-Casal
Dept. of Statistics and OR
University of Vigo

Wenceslao González-Manteiga
Dept. of Statistics and OR
University of Santiago de Compostela

ABSTRACT

We propose in this work two different goodness-of-fit testing techniques for the spatial spectral density. The first approach is based on a smoothed version of the ratio between the periodogram and a parametric estimator of the spectral density. The second one is a generalized likelihood ratio test statistic, based on the log-periodogram representation as the response variable in a regression model. As a particular case, we provide a test for independence. Asymptotic normal distribution of both statistics is obtained, under the null hypothesis. We carry out a simulation study, using resampling techniques to estimate the p-value of the tests. Applications to real data are also provided.

Key words: spatial spectral density; goodness-of-fit tests; Kernel estimators; Local likelihood.

1 Introduction

One of the main problems in spatial statistics is the description of the dependence structure of a data set, both for lattice or geostatistical data. In the geostatistics context, the estimation of the variogram plays a key role since kriging prediction methods depend on the variogram or covariogram estimation. There exists an extensive literature devoted to the estimation problem (see Cressie (1993) for parametric models and García-Soidán *et al.* (2003), García-Soidán *et al.* (2004) for nonparametric estimations of the variogram). Recently, Francisco-Fernandez and Opsomer (2005) proposed a generalized cross-validation (GCV) criteria, suitably adjusted for the presence of spatial correlation, in order to fit a nonparametric regression model to spatial data. The correlation adjustment needed for the modification of the GCV also involves the estimation of the dependence structure. Diblasi and Bowman (2001) propose a goodness-of-fit technique to analyze spatial independence and Maglione and Diblasi (2004) extend this result in order to test whether the spatial

*Corresponding author. Department of Statistics and Operation Research. Faculty of Mathematics, University of Santiago de Compostela. Campus Sur s/n C.P.15782. Santiago de Compostela. Spain. E-mail: rcrujeir@usc.es

dependence structure of a data set can be explained by a certain parametric family. For multidimensional lattice data, there are several classical references on estimation and modelling. Whittle (1954) points out the difference between stationary processes in the plane or in time and derives general estimation equations, which provide the so called Whittle estimates. In the seminal paper from Besag (1974), the author studies different stochastic models which may describe lattice data. Martin (1979) introduces the linear-by-linear processes, a simple subclass of lattice processes in order to represent autocorrelated variables in practical situations. Guyon (1982) is concerned with parameter estimation. The estimation problem in this context has been deeply studied. Our main concern in this paper is to check spatial dependence structures. The theoretical results presented in this work are given for stationary processes observed on a two-dimensional regular grid. The testing methods we propose can be adapted, under suitable modifications, for geostatistical (continuous) data.

In spatial statistics, one could think about solving the crucial problem of modelling the spatial dependence in the spectral domain, instead of working with the variogram or the covariogram in the spatial setting (e.g. Fuentes (2002)). The spatial spectral density is the Fourier transform of the covariogram, so testing a certain covariance structure is equivalent to test a spatial spectral density model. Spectral techniques are a broadly used tool in time series analysis, although its extension to higher dimension problems is not straightforward. From the theoretical point of view, the main advantage of the spectral methodology is that the dependence between observations can be avoided, for a large enough sample, since periodogram values at different Fourier frequencies are asymptotically independent (e.g. Brilinger (1974)).

Besides, the periodogram (the classical non-parametric estimator of the spectral density) can be obtained as the response variable in a multiplicative regression model. In time series context, Paparoditis (2000) proposes a goodness-of-fit test based on a smoothed ratio between the periodogram and a parametric estimator of the spectral density, under the null hypothesis of an underlying parametric model. Equivalently, the log-periodogram can be seen as the exogenous variable in an additive regression model. This idea is considered in Fan and Zhang (2004), in time series context, where the authors apply a generalized likelihood ratio test for regression models (Fan *et al.* (2001)). In order to adapt a regression goodness-of-fit test to the spectral setting, other techniques could be considered. For instance, one could use tests based on the error distribution function, using the empirical process methodology (Stute (1997), Stute *et al.* (1998)). In time series case, Delgado *et al.* (2005) propose a goodness-of-fit test based on empirical processes. Other tests could be based on smoothed estimators of the regression function (Härdle and Mammen (1993), González Manteiga and Cao (1993) and Hart (1997), among others).

We extend in this work, the goodness-of-fit testing techniques proposed in Paparoditis (2000) and Fan and Zhang (2004) to the multidimensional lattice data case.

This paper is organized as follows. In Section 2, we relate the spatial and spectral domains for lattice data. In Section 3, we provide the extensions of different goodness-of-fit tests for regression models to the spectral setting. Section 4 is devoted to the simulation study and real data application. A brief discussion is provided in Section 5.

2 Spectral techniques for spatial processes. Background.

Let Z be a zero-mean second order stationary process observed on a bidimensional regular grid $\mathcal{D} = \{0, \ldots, n_1 - 1\} \times \{0, \ldots, n_2 - 1\}$ and denote by $N = n_1 n_2$, number of observations. The covariance function of the process is defined by:

$$C(\mathbf{u}) = E\left(Z(\mathbf{s}) \cdot Z(\mathbf{s} + \mathbf{u})\right) \quad \mathbf{s}, \mathbf{u} \in \mathbb{Z}^2, \tag{1}$$

Assuming that $\sum_{\mathbf{u}} |C(\mathbf{u})| d\mathbf{u} < \infty$, by Khinchin's theorem (e.g. Grenander (1981), Yaglom (1987)), the covariance funcion of a stationary random process can be written as:

$$C(\mathbf{u}) = \int_{\Pi^2} e^{-i\mathbf{u}^T \boldsymbol{\lambda}} f(\boldsymbol{\lambda}) d\boldsymbol{\lambda}, \quad \Pi^2 = [-\pi, \pi] \times [-\pi, \pi]$$

where f is bounded and continuous and T denotes the transpose operator. This function f is the spectral density of Z.

Assume that Z can be represented as:

$$Z(\mathbf{s}) = \sum_{j=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} a_{jl} \varepsilon(s_1 - j, s_2 - l), \quad \mathbf{s} = (s_1, s_2) \tag{2}$$

where the error variables ε are independent and identically distributed as $N(0, \sigma_\varepsilon^2)$. As a particular case, any Gaussian process with absolutely integrable spectral density can be written in this way. Although representation (2) remind us to time series context, the extension of time series results to spatial processes is not simple, due to the fact that a variable in a time series is influenced only by past values while for spatial processes, this dependence extents in all directions. The spectral density of process (2) at a frequency $\boldsymbol{\lambda}$ is given by:

$$f(\boldsymbol{\lambda}) = \left| \sum_{j=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} a_{jl} e^{-i(j,l)\boldsymbol{\lambda}} \right|^2 \cdot \frac{\sigma_\varepsilon^2}{(2\pi)^2} = |A(\boldsymbol{\lambda})|^2 f_\varepsilon(\boldsymbol{\lambda}), \quad \boldsymbol{\lambda} \in \Pi^2, \tag{3}$$

where

$$A(\boldsymbol{\lambda}) = \sum_{j=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} a_{jl} e^{-i(j,l)\boldsymbol{\lambda}}, \quad \text{and } f_\varepsilon(\boldsymbol{\lambda}) = \frac{\sigma_\varepsilon^2}{(2\pi)^2}$$

with $(j,l)\boldsymbol{\lambda} = j\lambda_1 + l\lambda_2$. The classical nonparametric estimator of the spectral density is the periodogram, which is given by:

$$I(\boldsymbol{\lambda}_\mathbf{k}) = \frac{1}{(2\pi)^2} \cdot \frac{1}{N} \left| \sum_{\mathbf{s} \in \mathcal{D}} Z(\mathbf{s}) e^{-i\mathbf{s}^T \boldsymbol{\lambda}_\mathbf{k}} \right|^2, \tag{4}$$

where $\mathbf{s}^T \boldsymbol{\lambda}_\mathbf{k}$ denotes the scalar product in $\mathbb{R}^2$. The periodogram is usually computed at the set of bidimensional Fourier frequencies $\boldsymbol{\lambda}_\mathbf{k}^T = (\lambda_{k_1}, \lambda_{k_2})$:

$$\begin{aligned} \lambda_{k_1} &= \tfrac{2\pi k_1}{n_1}, \quad k_1 = 0, \pm 1, \ldots, \pm m_1, \quad \text{where } m_1 = [(n_1 - 1)/2], \\ \lambda_{k_2} &= \tfrac{2\pi k_2}{n_2}, \quad k_2 = 0, \pm 1, \ldots, \pm m_2, \quad \text{where } m_2 = [(n_2 - 1)/2] \end{aligned}$$

and denote by $n = (2m_1 + 1)(2m_2 + 1)$ the number of Fourier frequencies. The periodogram can be written in terms of the sample covariances as:

$$I(\boldsymbol{\lambda}_{\mathbf{k}}) = \frac{1}{(2\pi)^2} \sum_{\mathbf{u} \in \mathcal{U}} \hat{C}(\mathbf{u}) e^{-i\mathbf{u}^T \boldsymbol{\lambda}_{\mathbf{k}}}, \tag{5}$$

where $\mathcal{U} = \{\mathbf{u} = (u_1, u_2); u_1 = 1 - n_1, \ldots, n_1 - 1, u_2 = 1 - n_2, \ldots, n_2 - 1\}$ and the sample covariances are given by:

$$\hat{C}(\mathbf{v}) = \frac{1}{N} \sum_{\mathbf{s} \in \mathcal{D}(\mathbf{v})} Z(\mathbf{s}) Z(\mathbf{s} + \mathbf{v}), \quad \mathcal{D}(\mathbf{v}) = \{\mathbf{s} \in \mathcal{D}; \mathbf{s} + \mathbf{v} \in \mathcal{D}\}. \tag{6}$$

As we have seen in (3), f can be written in terms of f_ε, the spectral density of the innovation process. A similar expression is obtained for the periodogram of Z, which can be written in terms of the periodogram of ε:

$$I(\boldsymbol{\lambda}) = |A(\boldsymbol{\lambda})|^2 I_\varepsilon(\boldsymbol{\lambda}) + R_N(\boldsymbol{\lambda}), \tag{7}$$

where, the residual term is uniformly bounded by $\mathcal{O}_{\mathbb{P}}(N^{-1/2} \log N)$, as it is proved later in Lemma 2 in the appendix and I_ε denotes the periodogram for ε and expression (7) can be written as

$$I(\boldsymbol{\lambda}_{\mathbf{k}}) = f(\boldsymbol{\lambda}_{\mathbf{k}}) V_{\mathbf{k}} + R_N(\boldsymbol{\lambda}_{\mathbf{k}}) \tag{8}$$

where $V_{\mathbf{k}}$'s are independent identically distributed random variables with standard exponential distribution. Then, applying logarithms in (8), we have

$$Y_{\mathbf{k}} = m(\boldsymbol{\lambda}_{\mathbf{k}}) + z_{\mathbf{k}} + r_{\mathbf{k}} \tag{9}$$

where $m = \log f$ and

$$r_{\mathbf{k}} = \log\left[1 + \frac{R_N(\boldsymbol{\lambda}_{\mathbf{k}})}{f(\boldsymbol{\lambda}_{\mathbf{k}}) V_{\mathbf{k}}}\right]. \tag{10}$$

The variables $z_{\mathbf{k}}$ are independently and identically distributed with density function:

$$h(x) = e^{-e^x + x}. \tag{11}$$

The mean is the Euler constant $E(z_{\mathbf{k}}) = C_0 = -0.57721$ and the variance is $\mathrm{Var}(z_{\mathbf{k}}) = \pi^2/6$. This is a particular case of the Gumbel distribution, with position and scale parameters 0 and 1, respectively.

The main attractive feature of the periodogram is, as we have already commented, that its values at different Fourier frequencies are asymptotically independent. Besides, it is an asymptotically unbiased estimator of the spatial spectral density, but as it happens in the one-dimensional case, the periodogram is inconsistent (see Fuentes (2002)). In order to overcome this drawback, tapering and smoothing techniques could be used (Brillinger (1981), Robinson (2006)). Properties of spatial periodogram have been investigated by Whittle (1954), Guyon (1982) and Stein (1995).

Despite its lack of consistency as an estimator, the periodogram can be used as a pilot item when trying to fit a parametric model. In the spectral parametric context, Whittle parameter estimation is the most popular method. For a parametric model of the spatial spectral density f_θ with $\theta \in \Theta \subset \mathbb{R}^p$, the Whittle parameter estimator $\hat{\theta}$ is given by:

$$\hat{\theta} = \arg\min_{\theta} L(\theta, I), \tag{12}$$

where $L(\theta, I)$ denotes the Whittle log-likelihood

$$L(\theta, I) = \int_{\Pi^2} \left(\log f_\theta(\boldsymbol{\lambda}) + \frac{I(\boldsymbol{\lambda})}{f_\theta(\boldsymbol{\lambda})} \right) d\boldsymbol{\lambda}. \tag{13}$$

The log-likelihood function (13) can be interpreted as the Kullback-Leibler divergence between I and f_θ. The extension of spectral parametric estimation techniques to higher dimensional settings present some problems. Whittle estimates are studied in Guyon (1982). For dimension $d = 2$, the spatial periodogram bias contributes a bias of order $N^{-1/2}$ in the estimation of θ. In order to obtain a $\sqrt{N}$-consistent estimator of θ, an unbiased version of the periodogram can be used in the Whittle log-likelihood expression (Guyon (1982)). The unbiased periodogram is obtained from (5), replacing the sample covariances $\hat{C}(\mathbf{v})$ by the unbiased sample covariances, namely $\tilde{C}(\mathbf{v})$, with $\mathbf{v}^T = (v_1, v_2)$

$$\tilde{C}(\mathbf{v}) = \sum_{\mathbf{s} \in \mathcal{D}(\mathbf{v})} \frac{1}{(n_1 - s_1 + v_1)(n_1 - s_2 + v_2)} Z(\mathbf{s}) Z(\mathbf{s} + \mathbf{v}). \tag{14}$$

Another alternative is proposed in Dahlhaus and Künsch (1987), who correct this problem with tapering techniques. In this paper, we propose a simple bias correction based on parametric Bootstrap techniques.

3 Testing a model for the spectral density.

Our main goal is testing whether the spectral density for Z belongs to a parametric family $\mathcal{F}_\theta$, with $\theta \in \Theta \subset \mathbb{R}^p$:

$$\begin{array}{llll} H_0: & f & \in & \mathcal{F}_\theta = \{f_\theta; \quad \theta \in \Theta\}, \\ H_a: & f & \notin & \mathcal{F}_\theta = \{f_\theta; \quad \theta \in \Theta\}. \end{array} \tag{15}$$

Considering the log-spectral density, the problem can be written as

$$\begin{array}{llll} \tilde{H}_0: & m & \in & \mathcal{M}_\theta = \{m_\theta; \quad \theta \in \Theta\}, \\ \tilde{H}_a: & m & \notin & \mathcal{M}_\theta = \{m_\theta; \quad \theta \in \Theta\}. \end{array} \tag{16}$$

The periodogram is written in (8) as the exogenous variable in a multiplicative regression model. From equation (9), the log-spectral density function m can be seen as a regression function in a model where the response is given by the log-periodogram (substracting a residual term $r_{\mathbf{k}}$) and the explanatory variables are the corresponding Fourier frequencies (fixed design case).

Provided that $n_1 \to \infty$, $n_2 \to \infty$ and $n_1/n_2 \to c$, for a constant c, the following assumptions on the process, spectral density and bidimensional kernel function, K, are needed.

Assumption 1. Assume the spatial process Z can be represented as in (2), and $\sum_{j,l} |j|^{1/2} |a_{j,l}| < \infty$, $\sum_{j,l} |l|^{1/2} |a_{j,l}| < \infty$ and $\sum_{j,l} |l|^4 |j|^4 |a_{j,l}| < \infty$. Assume also that the error process is such that $E(\varepsilon(\mathbf{s})) = 0$, $E(\varepsilon^2(\mathbf{s})) = \sigma_\varepsilon^2$ and $E(\varepsilon^8(\mathbf{s})) < \infty$.

Assumption 2. The spectral density f is Lipschitz continuous and non vanishing, i.e. $\inf_{\boldsymbol{\lambda} \in [-\pi,\pi] \times [-\pi,\pi]} f(\boldsymbol{\lambda}) > 0$.

Assumption 3. K is symmetric , bounded and non-negative bidimensional kernel with support $\Pi^2 = [-\pi, \pi] \times [-\pi, \pi]$, such that $\int_{\mathbb{R}^2} K(\mathbf{u}) d\mathbf{u} = (2\pi)^2$ and $\int_{\mathbb{R}^2} K^2(\mathbf{u}) d\mathbf{u} < \infty$. The rescaled kernel K_H is defined by $K_H(\mathbf{u}) = |H|^{-1/2} K(H^{-1/2}\mathbf{u})$, following (Ruppert and Wand (1994)). The sequence of bandwidth matrices is such that each entry of H tends to zero and $N|H|^{1/2} \to \infty$. Some further assumptions on the bandwidth matrix are needed in theorem 2.

Assumption 4. The parameter space Θ is an open subset of $\mathbb{R}^p$ and the spectral density f_θ is twice differentiable w.r.t. θ with continuous second derivatives.

3.1 Using the spatial periodogram for hypothesis testing.

It is known that, if **Assumption 1** holds,

$$E\left(\frac{I(\boldsymbol{\lambda}_\mathbf{k})}{f_\theta(\boldsymbol{\lambda}_\mathbf{k})}\right) = \frac{f(\boldsymbol{\lambda}_\mathbf{k})}{f_\theta(\boldsymbol{\lambda}_\mathbf{k})} + \mathcal{O}(N^{-1}\log N), \tag{17}$$

uniformly in $\mathbf{k}$. Equation (17) implies that, under H_0, the asymptotic expected value of this ratio equals one. We consider a squared deviation criterion on a kernel type estimator of the ratio between the periodogram and the spectral density (under H_0), as it is proposed in Paparoditis (2000) for the one-dimensional case.

When testing a composite hypothesis $H_0 : f = f_\theta$ against $H_a : f \neq f_\theta$ as in case (15), the test statistic is given by:

$$T_P = N|H|^{1/4} \int_{\Pi^2} \left(\frac{1}{N|H|^{1/2}} \sum_{\mathbf{k}} K(H^{-1/2}(\boldsymbol{\lambda} - \boldsymbol{\lambda}_\mathbf{k})) \left(\frac{I(\boldsymbol{\lambda}_\mathbf{k})}{f_{\hat{\theta}}(\boldsymbol{\lambda}_\mathbf{k})} - 1 \right) \right)^2 d\boldsymbol{\lambda}, \tag{18}$$

where the sum $\sum_\mathbf{k}$ extends over the Fourier frequencies. Asymptotic normality of this statistic is also obtained.

Theorem 1. *Under assumptions (1)-(4) and under* $H_0 : f_\theta \in \mathcal{F}_\theta$

$$T_P - \mu_H \to N(0, \tau^2) \quad \textit{in distribution},$$

where μ_H *and* τ^2 *are given by:*

$$\mu_H = |H|^{-1/4} \int K^2(\mathbf{s}) d\mathbf{s}, \tag{19}$$

$$\tau^2 = \frac{1}{2\pi^2} \int_{2\Pi^2} \left(\int_{\Pi^2} K(\mathbf{s}) K(\mathbf{s} + \mathbf{u}) d\mathbf{s} \right)^2 d\mathbf{u}, \quad 2\Pi^2 = [-2\pi, 2\pi] \times [-2\pi, 2\pi]. \tag{20}$$

Assume that the true spectral density f lies in $\mathcal{F} - \mathcal{F}_\theta$ and consider θ^* satisfying:

$$\theta^* = \arg\min_\theta L(\theta, f)$$

where

$$L(\theta, f) = \int_{\Pi^2} \left(\log f_\theta(\boldsymbol{\lambda}) + \frac{f(\boldsymbol{\lambda})}{f_\theta(\boldsymbol{\lambda})} \right) d\boldsymbol{\lambda}.$$

where θ^* is the parameter which determines the best fit in $\mathcal{F}_\theta$. The estimator $\hat{\theta}$ given by (12) is an efficient estimator of θ^*. In order to guarantee the necessary conditions for this estimator, a further assumption on the parameter space Θ must be taken.

Assumption 5. $\Theta \subset \mathbb{R}^p$ is compact and f_θ is three times differentiable with respect to θ, with continuous derivatives. Besides, θ^* exists, is unique and lies in the interior of Θ.

Generalizing Theorem 3.2 in Dahlhaus and Wefelmeyer (1996), we see that:

$$\sqrt{N}(\hat{\theta} - \theta^*) = \sqrt{N} \int_{\Pi^2} W(\boldsymbol{\lambda})(I(\boldsymbol{\lambda}) - f(\boldsymbol{\lambda}))d\boldsymbol{\lambda} + o_{\mathbb{P}}(1) \tag{21}$$

where

$$W(\boldsymbol{\lambda}) = -\mathcal{H}^{-1}\nabla f_{\theta^*}^{-1}(\boldsymbol{\lambda}), \qquad \mathcal{H} = \int_{\Pi^2} \nabla^2 G(\theta^*, f, \boldsymbol{\lambda})d\boldsymbol{\lambda},$$

$$G(\theta, f, \boldsymbol{\lambda}) = \log f_\theta(\boldsymbol{\lambda}) + \frac{f(\boldsymbol{\lambda})}{f_\theta(\boldsymbol{\lambda})},$$

and ∇ and ∇^2 denote the first and second derivatives with respect to θ, $\nabla f_{\theta^*}^{-1}(\boldsymbol{\lambda})$ denotes the first derivative w.r.t. θ evaluated in θ^* and $\nabla^2 G(\theta^*, f, \boldsymbol{\lambda})$ denotes the second derivative w.r.t. θ evaluated in θ^*. $L(\theta, f)$ can be interpreted as the Kullback-Leibler divergence between f and f_θ.

Theorem 2. *Consider the problem of testing a composite hypothesis $H_0 : f \in \mathcal{F}_\theta$ vs. $H_a : f \in \mathcal{F} - \mathcal{F}_\theta$. If $f \in \mathcal{F} - \mathcal{F}_\theta$, under assumptions (1)-(3) and (5):*

$$N^{-1}|H|^{-1/4}T_P \to \int_{\Pi^2} \left(\frac{f(\boldsymbol{\lambda})}{f_{\theta^*}(\boldsymbol{\lambda})} - 1 \right)^2 d\boldsymbol{\lambda}$$

in probability.

Remark 1. Both results are a generalization of Theorems 2 and 3 in Paparoditis (2000) for time series context. As in the one-dimensional situation, this result implies the omnibus property of the T_P test, that is, T_P is consistent against any alternative such that $f \notin \mathcal{F}_\theta$. Besides, note that in practice, a discretized version of the T_P statistic is used (see simulation section).

3.2 Using the spatial log-periodogram for hypothesis testing.

In this part, we tackle the testing problem (16). Consider the following regression model:

$$Y_{\mathbf{k}}^{**} = m(\boldsymbol{\lambda}_{\mathbf{k}}) + z_{\mathbf{k}}^*, \tag{22}$$

where we denote by $Y_{\mathbf{k}}^{**} = Y_{\mathbf{k}}^* - r_{\mathbf{k}}$, $Y_{\mathbf{k}}^* = Y_{\mathbf{k}} - C_0$ and $z_{\mathbf{k}}^* = z_{\mathbf{k}} - C_0$. The $Y_{\mathbf{k}}^{**}$ variables are not observed, so we establish the testing procedure in terms of $Y_{\mathbf{k}}$, although the theoretical reasoning takes this fact into account.

Following Fan and Zhang (2004), we introduce the generalized likelihood ratio test statistic based on two likelihood approaches of equation (8). The first approach is given by the loglikelihood maximization under the null hypothesis. The second approach is purely nonparametric, obtained by a local loglikelihood function maximization. The loglikelihood function associated with (9), when $r_{\mathbf{k}}$ has been removed, is

$$\sum_{\mathbf{k}} \left[Y_{\mathbf{k}} - m(\boldsymbol{\lambda}_{\mathbf{k}}) - e^{Y_{\mathbf{k}} - m(\boldsymbol{\lambda}_{\mathbf{k}})} \right]. \tag{23}$$

We will introduce two likelihood-based approaches to obtain the generalized likelihood ratio test statistic. Under the null hypothesis, the maximizer of the loglikelihood function of (9), when ignoring the residual part $r_{\mathbf{k}}$, is the Whittle estimate from equation (12). From a non parametric approach, we consider the estimator obtained for the log-spectral density function m by a multidimensional local linear kernel estimator. For any $\mathbf{x} \in \mathbb{R}^2$, we approximate $m(\boldsymbol{\lambda}_{\mathbf{k}})$ by the plane $a + \mathbf{b}^T(\boldsymbol{\lambda}_{\mathbf{k}} - \mathbf{x})$. Then, we construct the local loglikelihood function

$$\sum_{\mathbf{k}} \left[Y_{\mathbf{k}} - a - \mathbf{b}^T(\boldsymbol{\lambda}_{\mathbf{k}} - \mathbf{x}) - e^{Y_{\mathbf{k}} - a - \mathbf{b}^T(\boldsymbol{\lambda}_{\mathbf{k}} - \mathbf{x})} \right] K_H(\boldsymbol{\lambda}_{\mathbf{k}} - \mathbf{x}), \tag{24}$$

where the function K_H is a reescaled bidimensional kernel, as in **Assumption 3**. The local maximum likelihood estimator $\hat{m}_{LK}(H, x) \equiv \hat{m}_{LK}(\mathbf{x})$ of $m(\mathbf{x})$ is $\hat{a}$ in the maximizer $(\hat{a}, \hat{b})$ of (24). Then, a generalized likelihood test statistic can be constructed as

$$T_{LK} = \sum_{\mathbf{k}} \left[e^{Y_{\mathbf{k}} - m_{\hat{\theta}}(\boldsymbol{\lambda}_{\mathbf{k}})} + m_{\hat{\theta}}(\boldsymbol{\lambda}_{\mathbf{k}}) - e^{Y_{\mathbf{k}} - \hat{m}_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})} - \hat{m}_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}) \right]. \tag{25}$$

The local estimator $\hat{m}_{LK}$ contains biases even under the null hypothesis which affect the distribution under H_0. In the regression context, Härdle and Mammen (1993) in order to compare parametric vs. nonparametric regression fits, propose smoothing the residuals from both approaches. The bias correction technique consists on a reparametrization of the log-periodogram. Let θ denote the true parameter under H_0 and rewrite $m^{BC}(\boldsymbol{\lambda}) = m(\boldsymbol{\lambda}) - m_\theta(\boldsymbol{\lambda})$. Then, the hypothesis testing statement, in terms of m^{BC} is given by:

$$H_0 : m^{BC} = 0,$$
$$H_a : m^{BC} \neq 0.$$

The expression for the test statistic is:

$$T_{LK,BC} = \sum_{\mathbf{k}} \left(e^{\tilde{Y}_{\mathbf{k}}} - e^{\tilde{Y}_{\mathbf{k}} - \hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})} - \hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}) \right),$$

where $\hat{\theta}$ is the Whittle estimator of θ and $\tilde{Y}_{\mathbf{k}} = Y_{\mathbf{k}} - m_{\hat{\theta}}(\boldsymbol{\lambda}_{\mathbf{k}})$ denote the synthetic data. m^*_{LK} is the local linear estimator of m^{BC}, considering these synthetic data. Although asymptotic distribution of the test statistic is also obtained, in practice, we approximate the null distribution of T_{LK} using Monte Carlo simulations. Consider the following decomposition of the test statistic.

$$T_{LK} = T_{LK,1} - T_{LK,2}$$

where

$$T_{LK,1} = \sum_{\mathbf{k}} \left[e^{(Y_{\mathbf{k}} - m_\theta(\boldsymbol{\lambda}_{\mathbf{k}}))} + m_\theta(\boldsymbol{\lambda}_{\mathbf{k}}) - e^{(Y_{\mathbf{k}} - \hat{m}_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}))} - \hat{m}_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}) \right]$$

$$T_{LK,2} = \sum_{\mathbf{k}} \left[e^{(Y_{\mathbf{k}} - m_\theta(\boldsymbol{\lambda}_{\mathbf{k}}))} + m_\theta(\boldsymbol{\lambda}_{\mathbf{k}}) - e^{(Y_{\mathbf{k}} - m_{\hat{\theta}}(\boldsymbol{\lambda}_{\mathbf{k}}))} - m_{\hat{\theta}}(\boldsymbol{\lambda}\mathbf{k}) \right]$$

The test statistic $T_{LK,1}$ is the generalised likelihood ratio test statistic for testing between

$$\begin{aligned}\tilde{H}_0 : m &= m_\theta \\ \tilde{H}_a : m &\neq m_\theta\end{aligned}$$

while $T_{LK,2}$ is the maximum likelihood ratio test statistic for testing between

$$\begin{aligned}\bar{H}_0 : \theta &= \theta_0 \\ \bar{H}_a : \theta &\neq \theta_0\end{aligned}$$

where θ_0 denotes the true parameter in the parametric family of models $\mathcal{M}_\theta$. For simplicity, we will denote the true parameter by θ, instead of θ_0 and the spectral density of Z will be denoted by f_θ. Under certain regularity conditions, the asymptotic null distribution of $T_{LK,2}$ is χ^2_p, where $p = dim(\theta)$. Hence, $T_{LK,2} = O_{\mathbb{P}}(1)$. Therefore, we can simplify the test statistic to $T_{LK,1}$ with a simple null hypothesis test:

$$T_{LK} = \sum_{\mathbf{k}} \left[e^{(Y_{\mathbf{k}} - m_\theta(\boldsymbol{\lambda}_{\mathbf{k}}))} + m_\theta(\boldsymbol{\lambda}_{\mathbf{k}}) - e^{(Y_{\mathbf{k}} - \hat{m}_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}))} - \hat{m}_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}) \right]. \tag{26}$$

In order to study the asymptotic properties of this statistic, we decompose T_{LK} in some addends. We consider T^*_{LK}, which is the same statistic as T_{LK} but replacing $Y_{\mathbf{k}}$ by $Y^{**}_{\mathbf{k}}$ given in equation (22) and $\hat{m}_{LK}$ by $\hat{m}^*_{LK}$. If the observed test statistic is larger than a selected critical value, then we reject the null hypothesis.

Define also the following quantities, related to the asymptotic distribution of the test statistic:

$$\mu_H = \frac{4\pi^2}{|H|^{1/2}} \left(K(\mathbf{0}) - \frac{1}{2} \int K^2(\mathbf{s}) d\mathbf{s} \right), \tag{27}$$

$$b_H = \frac{-|H|^2}{8} \sum_{\mathbf{k}} \frac{1}{f_\theta(\boldsymbol{\lambda}_{\mathbf{k}})} \int\int \mathbf{s}^T H_{m_\theta}(\boldsymbol{\lambda}_{\mathbf{k}}) \mathbf{s} \cdot (\mathbf{s}+\mathbf{u})^T H_{m_\theta}(\boldsymbol{\lambda}_{\mathbf{k}}) (\mathbf{s}+\mathbf{u}) K(\mathbf{s}) K(\mathbf{s}+\mathbf{u}) d\mathbf{s} d\mathbf{u}, \tag{28}$$

$$\sigma^2 = \frac{2\pi^2}{|H|^{1/2}} \int (2K(\mathbf{s}) - K * K(\mathbf{s}))^2 \, d\mathbf{s}. \tag{29}$$

Where $H_{m_\theta}(\boldsymbol{\lambda}_{\mathbf{k}})$ is the Hessian matrix of m_θ.

Theorem 3. *Under assumptions (1)-(4), as $N^{(\zeta-1)/\zeta}|H|^{1/2} \geq c \log^\delta N$, for a constant c and some $\delta > (\zeta - 1)/(\zeta - 2))$, $\zeta > 2$ and provided that H_0 holds,*

$$\sigma^{-1}(T_{LK} - \mu_H + b_H) \to N(0, 1),$$

where μ_H, b_H and σ^2 are given by (27), (28) and (29), respectively.

Remark 2. The former theorem extends Theorem 1 in Fan and Zhang (2004) to the multidimensional setting.

Remark 3. Other goodness-of-fit testing techniques based on smoothed estimators $\hat{m}$ of the log-spectral density could be used. An L^2-approach could be considered:

$$T_C = \sum_{\mathbf{k}} (\hat{m}(\boldsymbol{\lambda}_{\mathbf{k}}) - \hat{m}_{\hat{\theta}}(\boldsymbol{\lambda}_{\mathbf{k}}))^2. \tag{30}$$

This test statistic was studied by González Manteiga and Cao (1993) (and simultaneously by Härdle and Mammen (1993), in a continuous form). Or even Zheng's test (Zheng (1996)), who proposes

$$T_Z = \frac{1}{n}\sum_{\mathbf{k}\neq\mathbf{v}} K_H(\boldsymbol{\lambda}_\mathbf{k} - \boldsymbol{\lambda}_\mathbf{v})(Y_\mathbf{k} - m_{\hat{\theta}}(\boldsymbol{\lambda}_\mathbf{k}))(Y_\mathbf{v} - m_{\hat{\theta}}(\boldsymbol{\lambda}_\mathbf{v})). \tag{31}$$

For the test statistics (30) and (31), asymptotic normal distributions are obtained in the one-dimensional case. Also in the one-dimensional case, Zhang and Dette (2004) give a power comparison between non-parametric regression tests. Similarly, in the spatial context, it would be possible to obtain the normal asymptotic distribution of the extensions of these tests.

Remark 4. Theorems 1 to 3 can be generalized for stationary random fields on $\mathbb{R}^d$, under a similar asymptotic framework. The d-variate kernel K (with support on $\Pi^d = [-\pi,\pi]^d$) and the $d\times d$ bandwidth matrix H must satisfy the corresponding assumption 3. For the T_P test, the expressions for the mean and the variance are given by:

$$\mu_H^{(d)} = \frac{1}{|H|^{1/4}}\int_{\Pi^d} K^2(\mathbf{s})d\mathbf{s},$$

$$\tau^{2(d)} = \frac{1}{2^{d-1}\pi^d}\int_{2\Pi^d}\left(\int_{\Pi^d} K(\mathbf{s})K(\mathbf{s}+\mathbf{u})d\mathbf{s}\right)d\mathbf{u}, \quad 2\Pi^d = [-2\pi, 2\pi]^d.$$

For the T_{LK} test:

$$\mu_H^{(d)} = \frac{(2\pi)^d}{|H|^{1/2}}\left(K(\mathbf{0} - \frac{1}{2}\int_{\Pi^d} K^2(\mathbf{s})d\mathbf{s})\right),$$

$$\sigma^{2(d)} = \frac{2^{d-1}\pi^d}{|H|^{1/2}}\int_{\Pi^d}\left(2K(\mathbf{s}) - K * K(\mathbf{s})\right)^2 d\mathbf{s}.$$

These expressions generalize the results in this section and those provided by Paparoditis (2000) and Fan and Zhang (2004).

3.3 Testing in practice.

Since the rate of convergence of the distributions of T_P and T_{LK} to their Gaussian limit is quite slow, we show an alternative way of estimating the distribution of the test statistic, under H_0, by a Monte Carlo approach. The performance of T_P and T_{LK} tests is shown in a simulation study. We propose the following algorithm, for computing the $p-$value of the test statistics T_P and T_{LK}:

Step 1. Obtain the parametric estimate $\hat{\theta}$.

Step 2. Compute the observed test statistic T^{obs}. For the discrete approximation of the T_P test:

$$T_P^{obs} = N|\hat{H}|^{1/4}\sum_\mathbf{k}\left(N^{-1}|\hat{H}|^{-1/2}\sum_\mathbf{j} K\left(H^{-1/2}(\boldsymbol{\lambda}_\mathbf{k} - \boldsymbol{\lambda}_\mathbf{j})\right)\left(\frac{I(\boldsymbol{\lambda}_\mathbf{j})}{f_{\hat{\theta}}(\boldsymbol{\lambda}_\mathbf{j})} - 1\right)\right)^2$$

and for the T_{LK}, obtain the non-parametric estimate $\hat{m}_{LK}(H,\cdot)$ and:

$$T_{LK}^{obs} = \sum_\mathbf{k}\left\{e^{Y_\mathbf{k} - m_{\hat{\theta}}(\boldsymbol{\lambda}_\mathbf{k})} - e^{Y_\mathbf{k} - \hat{m}_{LK}(\boldsymbol{\lambda}_\mathbf{k})} + m_{\hat{\theta}}(\boldsymbol{\lambda}_\mathbf{k}) - \hat{m}_{LK}(\boldsymbol{\lambda}_\mathbf{k})\right\},$$

Step 3. From $f_{\hat{\theta}}$, generate a random sample of size $N = n_1 \cdot n_2$.

Step 4. Using the generated random sample in *Step 3*, obtain the test statistic T^*.

Step 5. Repeat B times steps 3 and 4 and obtain the bootstrap test statistics $T_1^*, T_2^*, \dots, T_B^*$.

Step 6. Compute the p-value of the test statistic as the percentage of the bootstrap replicates $\{T_1^*, T_2^*, \dots, T_B^*\}$ that exceed T^{obs}.

Both for T_P and T_{LK} non-linear multidimensional optimization problems must be solved. Whittle estimates $\hat{\theta}$ are obtained in *Step 1*, using a discretized version of (13). Newton type methods can be used to solve this problem, although these methods are not suitable for situations where local maximum values are found. In order to guarantee the convergence to a global maximum, genetic algorithms were implemented (see Goldberg (1989)).

In the case of the algorithm for T_{LK}, the computational cost is highly increased in *Step 2* with the non-parametric estimation of the log-spectral density, obtained by local maximum loglikelihood. There is again a non-linear multidimensional optimization problem, which must be solved for every Fourier frequency. For each $\boldsymbol{\lambda}_{\mathbf{k}}$, we take $(Y_{\mathbf{k}}, \mathbf{0})$ as initial values of $(a, \mathbf{b})$ in (24). As it happens for solving *Step 1*, one could think of using genetic algorithms for avoiding convergence problems.

A key problem in nonparametric statistics is the selection of the bandwidth parameter. Optimal bandwidth selection for non parametric testing in high dimensional problems is still an open question. Usually, in practice, the standard approach consists of examining a range of bandwidths.

Automatic bandwidth selection criteria is another alternative. For instance, the bandwidth matrix could be chosen by minimizing the Mean Integrated Square Error of the nonparametric estimator under the null hypothesis that $H_0 : f = f_{\theta_0}$:

$$\hat{H} = \arg\min_H E\left(\int_{\Pi^2} (\hat{m}_{LK}(H, \boldsymbol{\lambda}) - m_{\theta_0}(\boldsymbol{\lambda}))^2 d\boldsymbol{\lambda}\right). \tag{32}$$

Bandwidth estimation can be obtained using a Monte Carlo approach of the MISE error (32), where $\hat{m}_{LK}^j$ denotes the nonparametric estimator in each simulated sample:

$$\hat{H} = \arg\min_H \frac{1}{M} \sum_{j=1}^{M} \int_{\Pi^2} \left(\hat{m}_{LK}^j(H, \boldsymbol{\lambda}) - m_{\theta_0}(\boldsymbol{\lambda})\right)^2 d\boldsymbol{\lambda}, \tag{33}$$

although in practice, the theoretical parameter θ_0 is replaced by an estimator $\hat{\theta}$. However, the computational cost of this approach can be really high in some cases (due to the computation of the local log-likelihood estimator). Since log-periodogram values are asymptotically independent, for a large enough sample, good approximations are expected using a traditional cross-validation criteria. That is, select $\hat{H}$ such that:

$$\hat{H} = \arg\min_H \sum_{\mathbf{k}} \left(\hat{m}_{LK}^{-k}(H, \boldsymbol{\lambda}_{\mathbf{k}}) - m_{\hat{\theta}}(\boldsymbol{\lambda}_{\mathbf{k}})\right)^2, \tag{34}$$

where $\hat{m}_{LK}^{-k}(H, \cdot)$ is the nonparametric estimator of the log-spectral density obtained by maximizing expression (24), deleting the frequency $\boldsymbol{\lambda}_{\mathbf{k}}$.

It is important to note that the bandwidth matrix H plays a different role in both test statistics. In the T_{LK} test, the bandwidth matrix is involved in the nonparametric estimation of the log-spectral density. In the T_P test, the bandwidth matrix is not involved in the estimation procedure. Therefore, it may be expected that this test statistic will be less influenced by the bandwidth parameter.

The algorithm we propose for calibrating the $p-$value of the test statistics needs, in *Step 3*, the generation of a sample of size N, given a parametrically estimated spectral density function $f_{\hat{\theta}}$. For that purpose, we consider a spectral simulation procedure, as it is outlined in (Chilès and Delfiner (1999) pp. 502-503).

Remark 4. If Z is a continous process (geostatistical data), the summation in representation (2) is replaced by an integral (Priestley (1981)) and the spectrum of such a process is defined for all $\boldsymbol{\lambda}$ in $\mathbb{R}^2$. Although asymptotic theory has not been yet obtained in this case, the tests can be applied, with suitable modifications, when the observations are taken on a regular grid. In this case, the spectral density estimators should be modified in order to account for the spacing between data (e.g. Fuentes (2002)).

4 Simulation study

In this section, we study the performance of the testing procedures in terms of size and power. For illustration purposes, we consider the bidimensional autoregressive process (from now on $BAR(1)$):

$$Z(i,j) = \beta_1 Z(i-1,j) + \beta_2 Z(i,j-1) - \beta_1\beta_2 Z(i-1,j-1) + \varepsilon(i,j), \tag{35}$$

where $\varepsilon(i,j)$ are independet identically distributed Gaussian random variables, with zero-mean and variance σ_ε^2. This is the simplest process in the class of linear-by-linear processes, introduced by Martin (1979) and it is also known as the doubly-geometric process.

Parameters β_1 and β_2 belong to $[0,1)$ to guarantee stationarity. The spectral density of this process can be factorized with respect to β_1 and β_2 as

$$f(\omega) = \frac{\sigma_\varepsilon^2}{(2\pi)^2} \cdot \frac{1}{1+\beta_1^2 - 2\beta_1\cos(\omega_1)} \cdot \frac{1}{1+\beta_2^2 - 2\beta_2\cos(\omega_2)}. \tag{36}$$

In order to study the size of the tests, we consider different values for the parameters β_1 and β_2 from 0.0 (which corresponds to the independence case) to 0.9. 1000 simulations of the process are generated on a 20×20 and 50×50 regular grid. Random sample generations of this process are obtained as in (Alonso *et al.* (1996)). Estimators for β_1 and β_2 are obtained from the periodogram of the generated data, using a discretized version of the Whittle log-likelihood (13).

We set the null hypothesis that Z is a doubly-geometric process, considering different parameters. A multiplicative Epanechnikov bidimensional kernel is used along the study. The bandwidth parameter has been chosen using the cross-validation criteria (34). In order to simplify the computations, we consider diagonal bandwidth matrices, with elements proportional to the spacing between frequencies:

$$H = r\cdot diag\left(\frac{2\pi}{n_1}, \frac{2\pi}{n_2}\right). \tag{37}$$

(β_1, β_2)	$\alpha = 0.01$ T_P	T_{LK}	$\alpha = 0.05$ T_P	T_{LK}	$\alpha = 0.1$ T_P	T_{LK}
$(0.0, 0.0)$	0.014	0.009	0.043	0.054	0.090	0.105
$(0.1, 0.1)$	0.014	0.014	0.043	0.045	0.090	0.085
$(0.2, 0.2)$	0.018	0.011	0.051	0.049	0.024	0.088
$(0.3, 0.3)$	0.021	0.080	0.058	0.052	0.112	0.100
$(0.4, 0.4)$	0.020	0.090	0.058	0.053	0.099	0.099
$(0.5, 0.5)$	0.022	0.014	0.058	0.054	0.103	0.105
$(0.6, 0.6)$	0.023	0.015	0.067	0.059	0.113	0.117
$(0.7, 0.7)$	0.044	0.037	0.104	0.097	0.172	0.161
$(0.8, 0.8)$	0.096	0.067	0.210	0.171	0.289	0.225
$(0.9, 0.9)$	0.170	0.189	0.346	0.347	0.443	0.457
$(0.1, 0.9)$	0.088	0.092	0.195	0.186	0.287	0.264

Table 1: Size of the tests. 20×20 grid.

(β_1, β_2)	$\alpha = 0.01$ T_P	T_{LK}	$\alpha = 0.05$ T_P	T_{LK}	$\alpha = 0.1$ T_P	T_{LK}
$(0.7, 0.7)$	0.018	0.023	0.060	0.056	0.125	0.105
$(0.9, 0.9)$	0.097	0.052	0.269	0.114	0.396	0.168

Table 2: Size of the tests, 50×50 grid.

(β_1, β_2)	$\alpha = 0.01$ T_P	T_{LK}	$\alpha = 0.05$ T_P	T_{LK}	$\alpha = 0.1$ T_P	T_{LK}
$(0.7, 0.7)$	0.027	0.019	0.054	0.049	0.097	0.098
$(0.8, 0.8)$	0.034	0.030	0.075	0.070	0.119	0.119
$(0.9, 0.9)$	0.048	0.053	0.107	0.112	0.165	0.169
$(0.1, 0.9)$	0.028	0.031	0.072	0.069	0.131	0.117

Table 3: Size of the tests, 20×20 grid, with bias correction on the parameter estimates.

	$\alpha = 0.01$		$\alpha = 0.05$		$\alpha = 0.1$	
(β_1, β_2)	T_P	T_{LK}	T_P	T_{LK}	T_P	T_{LK}
$(0.01, 0.01)$	0.017	0.009	0.062	0.047	0.110	0.088
$(0.05, 0.05)$	0.036	0.017	0.096	0.079	0.169	0.148
$(0.1, 0.1)$	0.085	0.097	0.192	0.254	0.307	0.374
$(0.2, 0.2)$	0.376	0.713	0.589	0.903	0.720	0.943
$(0.3, 0.3)$	0.882	0.993	0.952	1.000	0.980	1.000

Table 4: Power of the tests.Testing for independence.

The behaviour of the test in size terms is shown in Table 1, at three different significance levels: 0.01, 0.05 and 0.10. The percentage of rejections of both test statitistics are computed from 1000 simulations.The results are quite satisfactory for both test, when the autoregression parameters are smaller than 0.5. For autoregresion parameters near 1, the performance is not so good as in the previous cases. It may happens that, for high dependence parameters, this sample size is too small for hypothesis testing.

As an example, in Table 2, we show the results of applying T_{LK} and T_P for parameters $(0.7, 0.7)$ and $(0.9, 0.9)$, in a 50×50 regular grid. Despite increasing the sample size, the size of the test does not improve as it could be expected. In Figure 1 we observed that, for a 20×20 regular grid, large autoregression parameter estimates from Whittle's likelihood are seriously biased. It seems clear that the bias in the parametric estimation distorts the results in the approximation of the size of the tests.

As we have already commented, Whittle parameter estimates computed from the raw periodogram are biased. We propose a bootstrap correction technique, which can be included in the bootstrap procedure for approximating the test statistic distribution. The modifications in the algorithm described in the section 3.3, in order to include the bias correction technique, are the following:

Step 1. Obtain the parametric estimate $\hat{\theta}$.

1.A. Generate B' random samples of size N fron $f_{\hat{\theta}}$.

1.B. Estimate $\hat{\theta}_i^*$ for each sample.

1.C. $\hat{b}(\theta, \hat{\theta}) = \frac{1}{B'} \sum_i (\hat{\theta} - \hat{\theta}_i^*)$.

1.D. Replace $\hat{\theta}$ by the bias corrected version $\hat{\theta} + \hat{b}(\theta, \hat{\theta})$.

...

Step 5. Using the generated random sample in *Step 4*, obtain the test statistic T^*, correcting the parameter estimator $\hat{\theta}^*$ by $\hat{\theta}^* + \hat{b}(\theta, \hat{\theta})$, and repeat B times steps 3 and 4.

The percentage of rejections of both tests, in a 20×20 grid, when applying the Bootstrap bias correction on the parameter estimates, is shown in Table 3. Significative improvements are observed in all cases, although for parameters near one, the results are not still completely satisfactory.

Behaviour of the test in terms of power is shown in Table 4, when testing for independence, that is $H_0 : f = c$, for some positive constant c. We set as alternatives different parameters

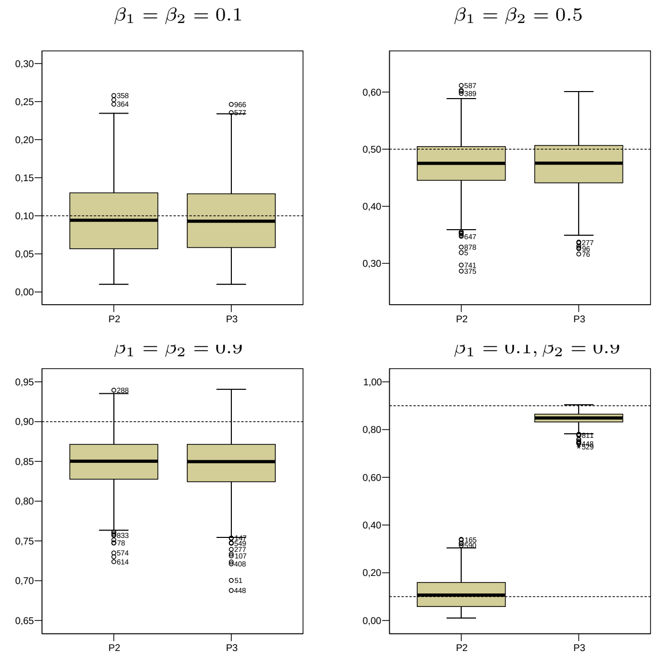

Figure 1: Parameter estimates.

approaching the null hypothesis. It seems that T_{LK} performs better than T_P. This feature may be explained by the fact that the bandwidth matrix approximates the optimal bandwidth for the nonparametric estimation.

5 Real data application. Mercer and Hall wheat-yield data.

In this section, we apply the proposed testing techniques to the well-known wheat data from Mercer and Hall experiment, consisting of a uniformity trial (all the plots received the same treatment) on an area of one acre. The layout is a 20×25 lattice. Although the exact size of the plots from the original data set seems to be unknown, some researchers have used 3.30 meters east to west, and 2.51 meters north to south. This dataset has been broadly studied by different authors (Whittle (1954), Cressie (1993), Young and Young (1998)). Young and Young (1998) conducted an exploratory data analysis on these data and Cressie (1993) shows that data indicate an irregular east-west trend. Whittle (1954) fitted a zero-mean, first-order autoregressive model:

$$Z(\mathbf{s}) = \alpha_1(Z(s_1+1, s_2) + Z(s_1-1, s_2)) + \alpha_2(Z(s_1, s_2+1) + Z(s_1, s_2-1)) + \varepsilon(\mathbf{s}), \quad (38)$$

where $\varepsilon(\mathbf{s})$ are zero-mean independent Gaussian random variables, with variance σ_ε^2. The corresponding spectral density is given by

$$f(\boldsymbol{\omega}) = \frac{\sigma^2}{(2\pi)^2}\left(1 - 2\alpha_1\cos(\omega_1) - 2\alpha_2\cos(\omega_2)\right)^{-2}, \quad \boldsymbol{\omega} \in \Pi^2. \quad (39)$$

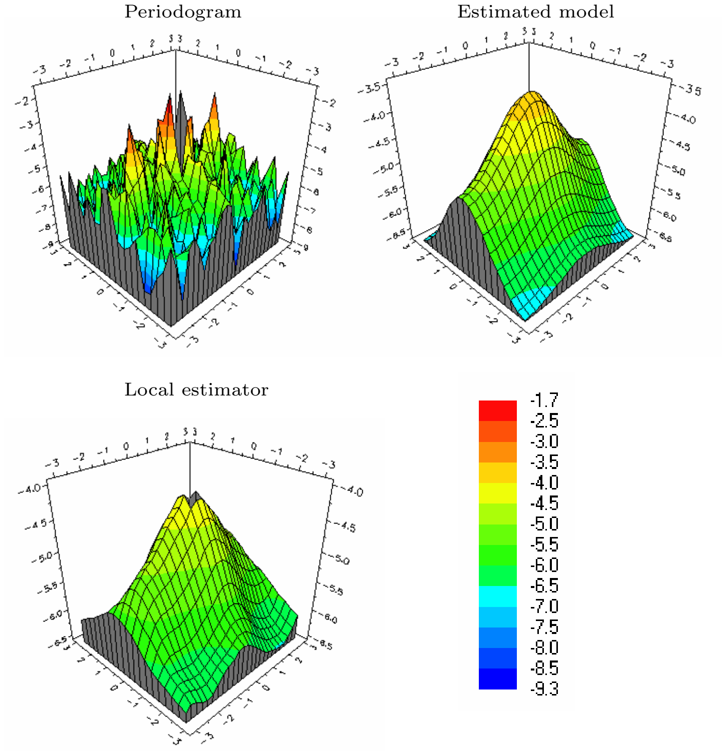

Figure 2: Spectral density estimators for Mercer-Hall data.

We will refer to model (38) as the spatial autoregressive model ($SAR(1)$ model). Whittle obtained $\widehat{\alpha_1} = 0.213$, $\widehat{\alpha_2} = 0.102$. No estimation of the variance σ_ε^2 is given. In Figure 2 we show the plots of the periodogram, the parametric fit (38) and the nonparametric log-likelihood estimator.

As a first approach, we test for independence, using both T_{LK} and T_P test statistics. We examine a range of diagonal bandwidth matrix (37), with r varying from 2.0 to 20.0. In both cases, the hypothesis of independence is rejected (p-values lower than 0.001) along the whole bandwidth range.

Once the independence hypothesis is rejected, we apply T_{LK} and T_P in order to check that model (38) fits the data. We obtain as estimated parameters $\widehat{\alpha_1} = 0.23217$, $\widehat{\alpha_2} = 0.09267$ and variance 0.12452. The $p-$values for different bandwidths are shown in Figure 3. In the horizontal axis, we represent the parameter r from equation (37) varying from 2.0 to 20.0. As it has been commented before, T_P test is less affected by the choice of the bandwidth, and the null hypothesis that the data admit a SAR model fit is accepted. T_{LK} test accepts the null hypothesis, for a significance level $\alpha = 0.05$, in most part of the bandwidth range, as it is shown in Figure 3. In particular, the null hypothesis is accepted for the cross-validation bandwidth.

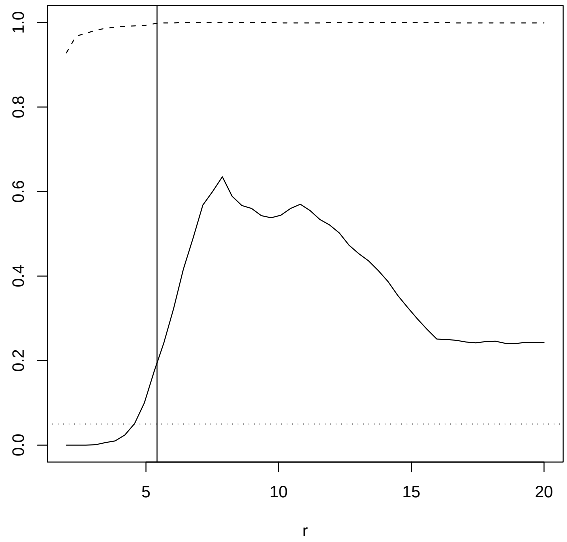

Figure 3: p-values for testing $SAR(1)$ model. Solid line: T_{LK} test $p-$values. Dashed line: T_P test $p-$values. Dotted line: significance level 0.05. Vertical solid line: cross-validation bandwidth.

6 Discussion

We propose in this work two different testing procedures to check whether a spatial spectral density belongs to a parametric family. The first proposal extends the test in (Paparoditis (2000)), for time series case, to the lattice data case in $\mathbb{R}^2$. The second approach is a generalized likelihood ratio test T_{LK}, computed with the bias-reduction approach, as it is proposed in (Fan *et al.* (2001)), also for time series. Both test statistics involve parameter estimation, which is done using the Whittle log-likelihood. For the spatial (or higher dimensional) case, Whittle estimates are not consistent and this feature must be taken into account when computing the test statistics. A Bootstrap bias correction technique is considered, in order to correct the parameter estimates bias. Besides, the T_{LK} test statistic needs also a non-parametric estimation of the log-spectral density. This non-parametric estimation is computed by local log-likelihood maximization.

In practice, the $p-$value of the test statistic, both for T_P and T_{LK} is approximated by Monte Carlo simulations. For both statistics, the parametric estimator must be obtained and for T_{LK}, the non-parametric estimator must be also computed. This non-parametric log-spectral density estimator needs a bandwidth matrix. For simplicity, we have considered diagonal bandwidth matrices, but other choices are also possible. An automatic selection criteria (cross-validation) has been employed to obtain the estimated bandwidth matrices.

In computational cost terms, T_P is less expensive than T_{LK}. This higher computational cost is due to the computation of the log-spectral density non-parametric estimator, which is obtained by local maximum loglikelihood. As we have already commented, in order to solve the non-linear multidimensional optimization problems that arise in both techniques, genetic algorithms could be used. Although computationally more expensive than Newton type methods, these kind of algorithms avoid local extremes.

7 Acknowledgements

The authors acknowledge the financial support of the Ministry of Science and Technology and FEDER, project MTM2005-00820 and grant BES2003-0581.

8 Appendix.

8.1 Proof of Theorem 1.

In order to prove Theorem 1, we must introduce some lemmas. Lemma 6 gives a decomposition of the T_P statistic as a sum of the test statistic when considering a simple null hypothesis plus a negligible term. Lemma 5 gives the asymptotic distribution of the T_P statistic, under $H_0 : \theta = \theta_0$. Lemmas 1 to 5 provide some tools which are needed in Lemma 6 and 7.

Lemma 1. *Assume that $\{\theta_N\}$ is a sequence of estimators of $\theta_0 \in \Theta \in \mathbb{R}^p$ such that $\sqrt{N}(\theta_N - \theta_0) = \mathcal{O}_{\mathbb{P}}(1)$. Assume that the spectral density f_{θ_0} is continuously differentiable w.r.t. θ with bounded derivatives in $\boldsymbol{\lambda} \in \Pi^2 = [-\pi, \pi] \times [-\pi, \pi]$. Then, under the assumptions in Theorem 1:*

$$\sup_{\boldsymbol{\lambda} \in \Pi^2} \left| \frac{f_{\theta_N}(\boldsymbol{\lambda}) - f_{\theta_0}(\boldsymbol{\lambda})}{f_{\theta_N}(\boldsymbol{\lambda})} \right| = \mathcal{O}_{\mathbb{P}}(N^{-1/2}). \tag{40}$$

Proof. Since for any θ_N, the estimated spectral density f_{θ_N} is continuous in Π^2, then

$$\sup_{\boldsymbol{\lambda} \in \Pi^2} \left| \frac{1}{f_{\theta_N}(\boldsymbol{\lambda})} \right| = \mathcal{O}_{\mathbb{P}}(1). \tag{41}$$

Besides, since $\sqrt{N}(\theta_N - \theta_0) = \mathcal{O}_{\mathbb{P}}(1)$, it implies that the difference between the estimator θ_N and the parameter θ_0 can be stochastically bounded by: $\theta_N - \theta_0 = \mathcal{O}_{\mathbb{P}}(N^{-1/2})$. For a fixed $\boldsymbol{\lambda}$, using a Taylor expansion of f_{θ_N} around f_θ and considering the Lagrange remainder, we have:

$$f_{\theta_N}(\boldsymbol{\lambda}) = f_{\theta_0}(\boldsymbol{\lambda}) + (\theta_N - \theta_0)^T \nabla f_{\tilde{\theta}}(\boldsymbol{\lambda}) \le f_{\theta_0}(\boldsymbol{\lambda}) + \sum_{i=1}^{p} |\theta_N^i - \theta_0^i| \sup_{\boldsymbol{\lambda} \in \Pi^2} \left| \frac{\partial}{\partial \theta_i} f_{\tilde{\theta}}(\boldsymbol{\lambda}) \right|$$

for some $\tilde{\theta}$ with $\|\tilde{\theta} - \theta_0\| \le \|\theta_N - \theta_0\|$. Therefore,

$$\sup_{\boldsymbol{\lambda} \in \Pi^2} |f_{\theta_N}(\boldsymbol{\lambda}) - f_{\theta_0}(\boldsymbol{\lambda})| \le \sum_{i=1}^{p} |\theta_N^i - \theta_0^i| \sup_{\boldsymbol{\lambda} \in \Pi^2} \left| \frac{\partial}{\partial \theta_i} f_{\tilde{\theta}}(\boldsymbol{\lambda}) \right| = \mathcal{O}_{\mathbb{P}}(N^{-1/2}). \tag{42}$$

The result is proved combining equations (41) and (42). □

Lemma 2. *Consider Z a spatial process with representation (2) and suppose that assumption (1) holds. Then:*

$$\max_{\boldsymbol{\lambda} \in \Pi^2} E(R_N^4(\boldsymbol{\lambda})) = \mathcal{O}(N^{-2}), \tag{43}$$

$$\max_{\mathbf{k}} |R_N(\boldsymbol{\lambda}_{\mathbf{k}})| = \mathcal{O}_{\mathbb{P}}(N^{-1/2} \log N). \tag{44}$$

Proof. In order to prove (43), the residual term $R_N(\boldsymbol{\lambda})$ can be written as:

$$R_N(\boldsymbol{\lambda}) = A(\boldsymbol{\lambda}) J_\varepsilon(\boldsymbol{\lambda}) Y_N(-\boldsymbol{\lambda}) + A(-\boldsymbol{\lambda}) J_\varepsilon(-\boldsymbol{\lambda}) Y_N(\boldsymbol{\lambda}) + |Y_N(\boldsymbol{\lambda})|^2, \tag{45}$$

where,

$$A(\boldsymbol{\lambda}) = \sum_{j=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} a_{jl} e^{-i(j,l)\boldsymbol{\lambda}}, \tag{46}$$

$$J_{\varepsilon}(\boldsymbol{\lambda}) = \frac{1}{2\pi\sqrt{N}} \sum_{s_1=0}^{n_1-1} \sum_{s_2=0}^{n_2-1} \varepsilon(\mathbf{s}) e^{-i\mathbf{s}^T\boldsymbol{\lambda}}, \tag{47}$$

$$U_{N,j,l}(\boldsymbol{\lambda}) = \frac{1}{2\pi\sqrt{N}} \left\{ \sum_{s_1=-j}^{n_1-1-j} \sum_{s_2=-l}^{n_2-1-l} e^{-i\mathbf{s}^T\boldsymbol{\lambda}} \varepsilon(\mathbf{s}) - \sum_{s_1=0}^{n_1-1} \sum_{s_2=0}^{n_2-1} e^{-i\mathbf{s}^T\boldsymbol{\lambda}} \varepsilon(\mathbf{s}) \right\}, \tag{48}$$

and finally

$$Y_N(\boldsymbol{\lambda}) = \sum_{j=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} a_{jl} e^{-i(j,l)\boldsymbol{\lambda}} U_{N,j,l}(\boldsymbol{\lambda}), \tag{49}$$

just following similar arguments to those in (Brockwell and Davis (1991), pp. 346-347). Therefore, taking expectations on the fourth order moment:

$$E\left(R_N^4(\boldsymbol{\lambda})\right) \leq$$
$$k_1 E\left(|A(\boldsymbol{\lambda})J_{\varepsilon}(\boldsymbol{\lambda})Y_N(-\boldsymbol{\lambda})|^4\right) + k_2 E\left(|A(-\boldsymbol{\lambda})J_{\varepsilon}(-\boldsymbol{\lambda})Y_N(\boldsymbol{\lambda})|^4\right) + k_3 E(|Y_N(\boldsymbol{\lambda})|^8),$$

for some positive constants k_1, $i = 1, 2, 3$. For the first term on the right hand side using Caucy-Schwarz inequality:

$$E\left(|A(\boldsymbol{\lambda})J_{\varepsilon}(\boldsymbol{\lambda})Y_N(-\boldsymbol{\lambda})|^4\right) \leq$$
$$|A(\boldsymbol{\lambda})|^4 \left(E|J_{\varepsilon}(\boldsymbol{\lambda})|^8\right)^{1/2} \left(E|Y_N(-\boldsymbol{\lambda})|^8\right)^{1/2} = \mathcal{O}(1)\sqrt{E(|Y_N(-\boldsymbol{\lambda})|^8)}.$$

For $E|Y_N(-\boldsymbol{\lambda})|^8$, we can get a bound taking into account that, if $|j| < n_1$ and $|l| < n_2$, $2\pi\sqrt{N}U_{N,j,l}$ is a sum of $4|j||l|$ independent random variables. For $|j| \geq n_1$, $|l| \geq n_2$, it is a sum of $4n_1n_2$ independent random variables. In the case $|j| < n_1$, $|l| \geq n_2$, it is a sum of $4|j|n_2$ iid random variables, whereas if $|j| \geq n_1$, $|l| < n_2$, it is a sum of $4|l|n_1$ iid random variables. Then, using the inequality:

$$E\left(\sum_{j=1}^{n} Z_j\right)^8 \leq nEZ_1^8 + 28n^2EZ_1^6EZ_1^2 + 35n^2(EZ_1^4)^2 + 210n^3EZ_1^4(EZ_1^2)^2 + 105(EZ_1^2)^4$$

where Z_j are independent identically distributed random variables, with zero mean and finite eight-order moment, we have:

$$E\left|U_{N,j,l}(\boldsymbol{\lambda})\right|^8 \leq c_1|j||l|E(\varepsilon^8) + c_2|j|^2|l|^2E(\varepsilon^6)E(\varepsilon^2) + c_3|j|^2|l|^2E^2(\varepsilon^4)$$
$$+c_4|j|^3|l|^3E(\varepsilon^4)E^2(\varepsilon^2) + c_5|j|^4|l|^4E^4(\varepsilon^2).$$

By assumption (1), concercing the summability of $\{|j||l|a_{j,l}\}$ and Jensen's inequality, we get $E|Y_N(\boldsymbol{\lambda})|^8 \leq \mathcal{O}(N^{-4})$:

$$E(|Y_N(\boldsymbol{\lambda})|^8) = E\left(\left|\sum_{j,l}|a_{j,l}|e^{-i(j,l)\boldsymbol{\lambda}}U_{N,j,l}(\boldsymbol{\lambda})\right|^8\right) \leq c_6\sum_{j,l}a_{j,l}\left(E\left(|U_{N,j,l}(\boldsymbol{\lambda})|^8\right)\right)$$

$$c_6\left(\frac{1}{N^4}\sum_{j,l}|a_{j,l}|c_1|j||l|E(\varepsilon(\mathbf{s})^8) + \frac{1}{N^4}\sum_{j,l}|a_{j,l}|c_2|j|^2|l|^2E(\varepsilon(\mathbf{s})^6)E(\varepsilon(\mathbf{s})^2)\right.$$

$$+\frac{1}{N^4}\sum_{j,l}|a_{j,l}|c_3|j|^2|l|^2E^2(\varepsilon(\mathbf{s})^4)+\frac{1}{N^4}\sum_{j,l}|a_{j,l}|c_4|j|^3|l|^3E(\varepsilon(\mathbf{s})^4)E^2(\varepsilon(\mathbf{s})^2)$$
$$\left.+\frac{1}{N^4}\sum_{j,l}|a_{j,l}|c_5|j|^4|l|^4E^4(\varepsilon(\mathbf{s})^2)\right)=\mathcal{O}(N^{-4}),$$

and from the expression above, we obtain that $E^{1/2}(|Y_N(\boldsymbol{\lambda})|^8)=\mathcal{O}(N^{-2})$.

The bound for (44) ca be obtained by a straightforward extension of the arguments in (Kooperberg *et al.* (1995)).

Let's prove now (44). Consider the expression of $J_\varepsilon(\lambda)$ given by (47) and split it in its real and imaginary parts. The real part of $J_\varepsilon(\boldsymbol{\lambda})$ is distributed as:

$$Re(J_\varepsilon(\boldsymbol{\lambda}))\sim N\left(0,\frac{A^TA\sigma^2}{(2\pi)^2N}\right),$$

where A is given by

$$A=\begin{pmatrix}1\\ \cos((1,0)\boldsymbol{\lambda})\\ \vdots\\ \cos((1,n_2-1)\boldsymbol{\lambda})\\ \cos((2,1)\boldsymbol{\lambda})\\ \vdots\\ \cos((2,n_2-1)\boldsymbol{\lambda})\\ \vdots\\ \cos((n_1-1,1)\boldsymbol{\lambda})\\ \vdots\\ \cos((n_1-1,n_2-1)\boldsymbol{\lambda})\end{pmatrix}.$$

We prove that the real part is $O_{\mathbb{P}}(\sqrt{\log N})$, where $N=n_1\cdot n_2$. For that purpose, let $\nu\in\mathbb{R}$. We will prove that:

$$\mathbb{P}\left(\frac{1}{2\pi\sqrt{N}}\sum_{\mathbf{s}}\cos(\boldsymbol{\lambda}_{\mathbf{k}}^T\mathbf{s})\varepsilon(\mathbf{s})\geq\nu\sqrt{\log N}\right)\to 0. \tag{50}$$

First, considering the distribution of $Re(J_\varepsilon(\boldsymbol{\lambda}))$, we can write:

$$\mathbb{P}\left(\frac{1}{2\pi\sqrt{N}}\sum_{\mathbf{s}}\cos(\boldsymbol{\lambda}_{\mathbf{k}}^T\mathbf{s})\varepsilon(\mathbf{s})\geq\nu\sqrt{\log N}\right)=\sqrt{\frac{2\pi N}{A^TA\sigma^2}}\int_{\nu\sqrt{\log N}}^{\infty}e^{\left(-\frac{2\pi^2Nx^2}{A^TA\sigma^2}\right)}dx. \tag{51}$$

Applying a change of variable:

$$\frac{2\pi^2Nx^2}{A^TA\sigma^2}=\frac{y^2}{2},$$

we rewrite (51) as:

$$\mathbb{P}\left(\frac{1}{2\pi\sqrt{N}}\sum_{\mathbf{s}}\cos(\boldsymbol{\lambda}_{\mathbf{k}}^T\mathbf{s})\varepsilon(\mathbf{s})\geq\nu\sqrt{\log N}\right)=\frac{1}{\sqrt{2\pi}}\int_{\frac{2\pi\nu}{\sigma\sqrt{A^TA}}\sqrt{N\log N}}^{\infty}e^{-\frac{y^2}{2}}dy. \tag{52}$$

Since the following exponential inequality holds:

$$\int_y^\infty \frac{e^{-x^2/2}}{e^{-y^2/2}}dx \leq \int_y^\infty \frac{x}{y}e^{(-x^2/2}dx = \frac{1}{y}\int_{-y^2/2}^\infty e^{-u}du = \frac{1}{y}e^{-y^2/2},$$

that is

$$\int_y^\infty e^{-x^2/2}dx \leq \frac{1}{y}e^{-y^2/2}, \quad y > 0,$$

expression (52) can be bounded by:

$$\frac{1}{\sqrt{2\pi}}\int_{\frac{2\pi\nu}{\sigma\sqrt{A^TA}}\sqrt{N\log N}}^\infty e^{-\frac{y^2}{2}}dy \leq \frac{\sigma\sqrt{A^TA}}{(2\pi)^{3/2}\nu\sqrt{N\log N}}e^{-\frac{1}{2}\left(\frac{2\pi\nu\sqrt{N\log N}}{\sigma\sqrt{A^TA}}\right)^2}, \tag{53}$$

and since $A^TA = \mathcal{O}(N)$, it is easy to see that the right hand side in (53) tends to zero. Then, (50) is proved. The same result hold for the imaginary part of $J_\varepsilon(\boldsymbol{\lambda})$.

We find a (uniform) bound for $Y_N(\boldsymbol{\lambda})$. We can write the expression as:

$$Y_N(\boldsymbol{\lambda}) = \sum_{j=-\infty}^{\infty}\sum_{l=-\infty}^{\infty} a_{jl}\exp(-i\boldsymbol{\lambda}^T(j,l))\frac{1}{2\pi\sqrt{N}}\left[\sum_{s_1=-j}^{n_1-1-j}\sum_{s_2=-l}^{n_2-1-l} e^{(-i(j,l)^T\boldsymbol{\lambda})\varepsilon(\mathbf{s})}\right.$$
$$\left. - \sum_{s_1=0}^{n_1-1}\sum_{s_2=0}^{n_2-1} e^{(-i(j,l)^T\boldsymbol{\lambda})\varepsilon(\mathbf{s})}\right]$$

Decomposing each addend in real and imaginary part and taking as an example just the one dealing with the cosines (since the same procedure can be applied to the other addends), we have:

$$\frac{1}{2\pi\sqrt{N}}\sum_{j=-\infty}^{\infty}\sum_{l=-\infty}^{\infty}\sum_{s_1=-j}^{n_1-1-j}\sum_{s_2=-l}^{n_2-1-l} a_{jl}\cos(\boldsymbol{\lambda}^T(j+s_1,l+s_2))\varepsilon(\mathbf{s}) =$$
$$\frac{1}{2\pi\sqrt{N}}\sum_{j=-\infty}^{\infty}\sum_{l=-\infty}^{\infty}\sum_{p_1=0}^{n_1-1}\sum_{p_2=0}^{n_2-1} a_{jl}\cos(\boldsymbol{\lambda}^T\mathbf{p})\varepsilon(\mathbf{p}-(j,l)), \quad \mathbf{p}^T = (p_1,p_2).$$

We will see that this term is an $O_{\mathbb{P}}\left(\sqrt{\frac{\log N}{N}}\right)$:

$$\mathbb{P}\left(\frac{1}{2\pi\sqrt{N}}\sum_{j=-\infty}^{\infty}\sum_{l=-\infty}^{\infty}\sum_{p_1=0}^{n_1-1}\sum_{p_2=0}^{n_2-1} a_{jl}\cos(\boldsymbol{\lambda}^T\mathbf{p})\varepsilon(\mathbf{p}-(j,l)) \geq \nu\sqrt{\frac{\log N}{N}}\right) =$$
$$\mathbb{P}\left(\sum_{j=-\infty}^{\infty}\sum_{l=-\infty}^{\infty} a_{jl}N\left(0,\frac{\sigma^2}{(2\pi)^2N}\right) \geq \nu\sqrt{\frac{\log N}{N}}\right) \leq$$
$$\mathbb{P}\left(N\left(0,\frac{\sigma^2}{(2\pi)^2N}\right) \geq \frac{\nu}{S_A}\sqrt{\frac{\log N}{N}}\right) = \int_{\frac{\nu}{S_A}\sqrt{\frac{\log N}{N}}}^\infty \frac{\sqrt{(2\pi)^2N}}{\sigma}e^{\left(-\frac{(2\pi)^2Nx^2}{2\sigma^2}\right)}dx =$$

$$\int_{\frac{\nu 2\pi\sqrt{N}}{\nu S_A}\sqrt{\frac{\log N}{N}}}^{\infty} \frac{\sqrt{(2\pi)^2 N}}{\sigma}\frac{\sigma}{2\pi\sqrt{N}}e^{-\frac{y^2}{2}}dy = \frac{1}{\sqrt{2\pi}}\int_{\frac{2\pi\nu\sqrt{\log N}}{S_A\sigma}}^{\sigma} e^{\left(-\frac{y^2}{2}\right)}dy \leq$$
$$\frac{1}{\sqrt{2\pi}}\frac{1}{\frac{2\pi\nu}{S_A\sigma\sqrt{\log N}}}e^{\left(-\frac{(2\pi)^2\nu^2\log n}{S_A^2\sigma^2}\right)} = \frac{S_A\sigma}{(2\pi)^{3/2}\nu\sqrt{\log N}}e^{\left(-\frac{(2\pi)^2\nu^2\log N}{S_A^1\sigma^2}\right)} =$$
$$\frac{S_A\sigma}{(2\pi)^{3/2}\nu\sqrt{\log N}}e^{\left(-\frac{(2\pi)^2\nu^2}{S_A^2\sigma^2}\log N\right)} = M\frac{e^{-p\log N}}{\sqrt{N}} \to 0$$

The constants involved in the proof are given by:

$$M = \frac{S_A\sigma}{(2\pi)^{3/2}\nu\sqrt{\log N}}, \quad p = \frac{(2\pi)^2\nu^2}{2S_A^2\sigma^2} \quad \text{and} \quad S_A = \sum_{j=-\infty}^{\infty}\sum_{l=-\infty}^{\infty} a_{jl}.$$

So, for all $\boldsymbol{\lambda}_{\mathbf{k}}$, we have the following stochastic convergence rates:

$$J_\varepsilon(\boldsymbol{\lambda}_{\mathbf{k}}) = O_{\mathbb{P}}(\sqrt{\log N}) \quad \text{and} \quad Y_N(\boldsymbol{\lambda}_{\mathbf{k}}) = O_{\mathbb{P}}\left(\sqrt{\frac{\log N}{N}}\right).$$

Then, for the residual term $R_N(\boldsymbol{\lambda}_{\mathbf{k}})$, we obtain that:

$$R_N(\boldsymbol{\lambda}_{\mathbf{k}}) = O_{\mathbb{P}}\left(\frac{\log N}{\sqrt{N}}\right).$$

□

Lemma 3. *Under assumptions (1) and (2), as $N \to \infty$*

$$\frac{|H|^{1/4}}{N}\int_{\Pi^2}\left(\sum_{\mathbf{k}} K_H(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{k}})\frac{R_N(\boldsymbol{\lambda}_{\mathbf{k}})}{f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{k}})}\right)^2 d\boldsymbol{\lambda} \to 0 \text{ in probability.}$$

Proof. The proof of this lemma can be done by similar arguments to those in the proof of Lemma 5 in Paparoditis (2000), with bidimensional kernel function K and bandwidth matrix H.

We have that, from Lemma 2 and using the Cauchy-Schwarz inequality:

$$\begin{aligned}
&E(R_N(\boldsymbol{\lambda}_{\mathbf{k}})R_N(\boldsymbol{\lambda}_{\mathbf{j}})R_N(\boldsymbol{\lambda}_{\mathbf{i}})R_N(\boldsymbol{\lambda}_{\mathbf{m}})) \leq \\
&\left\{E(R_N^2(\boldsymbol{\lambda}_{\mathbf{k}})R_N^2(\boldsymbol{\lambda}_{\mathbf{j}}))\right\}^{1/2}\left\{E(R_N^2(\boldsymbol{\lambda}_{\mathbf{i}})R_N^2(\boldsymbol{\lambda}_{\mathbf{m}}))\right\}^{1/2} \leq \\
&\left\{E(R_N^4(\boldsymbol{\lambda}_{\mathbf{k}}))E(R_N^4(\boldsymbol{\lambda}_{\mathbf{j}}))E(R_N^4(\boldsymbol{\lambda}_{\mathbf{i}}))E(R_N^4(\boldsymbol{\lambda}_{\mathbf{m}}))\right\}^{1/4} = \mathcal{O}(N^{-2}).
\end{aligned}$$

We prove that

$$\frac{|H|^{1/4}}{N}\int_{\Pi^2}\left(\sum_{\mathbf{k}} K_H(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{k}})\frac{R_N(\boldsymbol{\lambda}_{\mathbf{k}})}{f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{k}})}\right)^2 d\boldsymbol{\lambda}$$

tends to zero in L^2 norm.

$$0 \le E\left(\frac{|H|^{1/4}}{N}\int_{\Pi^2}\left(\sum_{\mathbf{k}} K_H(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{k}})\frac{R_N(\boldsymbol{\lambda}_{\mathbf{k}})}{f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{k}})}\right)^2 d\boldsymbol{\lambda}\right)^2 =$$

$$\frac{|H|^{1/2}}{N^2}E\left(\int_{\Pi^2}\sum_{\mathbf{k}}\sum_{\mathbf{j}} K_H(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{k}})K_H(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{j}})\frac{R_N(\boldsymbol{\lambda}_{\mathbf{k}})}{f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{k}})}\frac{R_N(\boldsymbol{\lambda}_{\mathbf{j}})}{f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{j}})}d\boldsymbol{\lambda}\right)^2 =$$

$$\frac{|H|^{1/2}}{N^2}E\left(\sum_{\mathbf{k}}\sum_{\mathbf{j}}\int_{\Pi^2} K_H(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{k}})K_H(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{j}})d\boldsymbol{\lambda}\frac{R_N(\boldsymbol{\lambda}_{\mathbf{k}})}{f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{k}})}\frac{R_N(\boldsymbol{\lambda}_{\mathbf{j}})}{f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{j}})}\right)^2 =$$

$$\frac{|H|^{1/2}}{N^2}E\left(\left[\sum_{\mathbf{k}}\sum_{\mathbf{j}}\int_{\Pi^2} K_H(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{k}})K_H(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{j}})d\boldsymbol{\lambda}\frac{R_N(\boldsymbol{\lambda}_{\mathbf{k}})}{f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{k}})}\frac{R_N(\boldsymbol{\lambda}_{\mathbf{j}})}{f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{j}})}\right]\right.$$

$$\left.\left[\sum_{\mathbf{l}}\sum_{\mathbf{m}}\int_{\Pi^2} K_H(\boldsymbol{\omega}-\boldsymbol{\lambda}_{\mathbf{l}})K_H(\boldsymbol{\omega}-\boldsymbol{\lambda}_{\mathbf{m}})d\boldsymbol{\omega}\frac{R_N(\boldsymbol{\lambda}_{\mathbf{l}})}{f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{l}})}\frac{R_N(\boldsymbol{\lambda}_{\mathbf{m}})}{f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{m}})}\right]\right) \le$$

$$\frac{|H|^{1/2}}{N^2}\sum_{\mathbf{k}}\sum_{\mathbf{j}}\int_{\Pi^2} K_H(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{k}})K_H(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{j}})d\boldsymbol{\lambda}$$

$$\times\sum_{\mathbf{l}}\sum_{\mathbf{m}}\int_{\Pi^2} K_H(\boldsymbol{\omega}-\boldsymbol{\lambda}_{\mathbf{l}})K_H(\boldsymbol{\omega}-\boldsymbol{\lambda}_{\mathbf{m}})d\boldsymbol{\omega}\frac{1}{f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{k}})f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{j}})f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{l}})f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{m}})}$$

$$\times E|R_N(\boldsymbol{\lambda}_{\mathbf{k}})R_N(\boldsymbol{\lambda}_{\mathbf{j}})R_N(\boldsymbol{\lambda}_{\mathbf{l}})R_N(\boldsymbol{\lambda}_{\mathbf{m}})| = \mathcal{O}(|H|^{1/2}),$$

where the last equality follows from the fact that

$$\int_{\Pi^2} K_H(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{k}})K_H(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{j}})d\boldsymbol{\lambda} < \infty.$$

□

Lemma 4. *Consider Z a spatial process with spectral density f and denote $W_{\mathbf{k}} = V_{\mathbf{k}} - 1$, where $V_{\mathbf{k}} \sim Exp(1)$, independent random variables. Under assumptions (1)-(2),*

$$\frac{|H|^{1/4}}{N}\int_{\Pi^2}\sum_{\mathbf{k}}\sum_{\mathbf{j}} K_H(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{k}})K_H(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{j}})W_{\mathbf{k}}\frac{R_N(\boldsymbol{\lambda}_{\mathbf{j}})}{f(\boldsymbol{\lambda}_{\mathbf{j}})}d\boldsymbol{\lambda} \to 0$$

in probability.

Proof. The proof is similar to Lemma 4 in Paparoditis (2000).

Consider the following notation, in order to make the proof more brief:

$$K_H(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{k}}) = K_H^{\mathbf{k}}(\boldsymbol{\lambda}), \quad f(\boldsymbol{\lambda}_{\mathbf{k}}) = f_{\mathbf{k}} \quad \text{and} \quad R_N(\boldsymbol{\lambda}_{\mathbf{k}}) = R_N^{\mathbf{k}}.$$

We will prove L^2-consistency:

$$E\left(\frac{|H|^{1/4}}{N}\int_{\Pi^2}\sum_{\mathbf{k}}\sum_{\mathbf{j}} K_H(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{k}})K_H(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{j}})(V_{\mathbf{k}}-1)\frac{R_N(\boldsymbol{\lambda}_{\mathbf{j}})}{f(\boldsymbol{\lambda}_{\mathbf{j}})}d\boldsymbol{\lambda}\right)^2 \tag{54}$$

$$= \frac{|H|^{1/2}}{N^2}\int\int\sum_{\mathbf{k},\mathbf{j},\mathbf{l},\mathbf{m}} K_H^{\mathbf{k}}(\boldsymbol{\lambda})K_H^{\mathbf{j}}(\boldsymbol{\lambda})K_H^{\mathbf{l}}(\boldsymbol{\omega})K_H^{\mathbf{m}}(\boldsymbol{\omega})\frac{E\left(W_{\mathbf{k}}R_N^{\mathbf{j}}W_{\mathbf{l}}R_N^{\mathbf{m}}\right)}{f_{\mathbf{k}}f_{\mathbf{j}}}d\boldsymbol{\lambda}d\boldsymbol{\omega} \tag{55}$$

In order to find a bound for this term, consider that $\mathbf{k}$, $\mathbf{j}$, $\mathbf{l}$ and $\mathbf{m}$ are all different indexes. From Theorem 2.3.2 in (Brillinger (1981)):

$$\begin{aligned}
&E(W_{\mathbf{k}}R_N^{\mathbf{j}}W_{\mathbf{l}}R_N^{\mathbf{m}}) = \\
&\mathrm{cum}(W_{\mathbf{k}}R_N^{\mathbf{j}})\mathrm{cum}(W_{\mathbf{l}}R_N^{\mathbf{m}}) + \mathrm{cum}(W_{\mathbf{k}}R_N^{\mathbf{m}})\mathrm{cum}(W_{\mathbf{l}}R_N^{\mathbf{j}}) + \mathrm{cum}(R_N^{\mathbf{j}})\mathrm{cum}(W_{\mathbf{k}}W_{\mathbf{l}}R_N^{\mathbf{m}}) \\
&+\mathrm{cum}(R_N^{\mathbf{l}})\mathrm{cum}(W_{\mathbf{k}}W_{\mathbf{l}}R_N^{\mathbf{j}}) + \mathrm{cum}(W_{\mathbf{k}})\mathrm{cum}(R_N^{\mathbf{j}}W_{\mathbf{l}}R_N^{\mathbf{m}}) + \mathrm{cum}(W_{\mathbf{l}})\mathrm{cum}(W_{\mathbf{k}}R_N^{\mathbf{j}}R_N^{\mathbf{m}}).
\end{aligned}$$

Since $\mathrm{cum}(W_{\mathbf{k}}) = E(W_{\mathbf{k}}) = 0$ and $\mathrm{cum}(W_{\mathbf{k}}W_{\mathbf{l}}) = E(W_{\mathbf{k}}W_{\mathbf{l}}) = 0$, and applying Theorem 2.3.2 of (Brillinger (1981)) on the three term cumulants, the expression above can be simplified:

$$E(W_{\mathbf{k}}R_N^{\mathbf{j}}W_{\mathbf{l}}R_N^{\mathbf{m}}) = E(W_{\mathbf{k}}R_N^{\mathbf{j}})E(W_{\mathbf{k}}R_N^{\mathbf{m}}) + E(W_{\mathbf{l}}R_N^{\mathbf{j}})E(W_{\mathbf{l}}R_N^{\mathbf{m}}) = \mathcal{O}(N^{-2}),$$

where the last equality is obtained recalling the expression for $R_N^{\mathbf{j}}$ in (45), and from a straightforward extension of Lemma 2 in Paparoditis (2000). Then, (54) is $\mathcal{O}(|H|^{1/2})$. Consider the case $\mathbf{k} = \mathbf{j} \neq \mathbf{l} = \mathbf{m}$. By the Cauchy-Schwarz inequality and Lemma 2,

$$|E(W_{\mathbf{k}}R_N^{\mathbf{j}}W_{\mathbf{l}}R_N^{\mathbf{m}})| \leq \sqrt{E(W_{\mathbf{k}}R_N^{\mathbf{j}})^2 E(W_{\mathbf{l}}R_N^{\mathbf{m}})^2} \leq \mathcal{O}(N^{-1}).$$

Then, (54) is $\mathcal{O}(N^{-1}|H|^{1/2})$. For the case $\mathbf{k} \neq \mathbf{j} \neq \mathbf{l} = \mathbf{m}$, using the same arguments, (54) is also $\mathcal{O}(N^{-1}|H|^{1/2})$. □

Lemma 5. *Assume that assumption (2) is fulfilled and consider $U_{\mathbf{k}}$ independent identically distributed random variables with $E(U_{\mathbf{k}}) = 1$, $Var(U_{\mathbf{k}}) = 1$ and $E(U_{\mathbf{k}}^4) < \infty$. Then,*

$$\frac{|H|^{1/4}}{N}\int_{\Pi^2}\left(\sum_{\mathbf{k}} K_H(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{k}})(U_{\mathbf{k}}-1)\right)^2 d\boldsymbol{\lambda} - \mu_H \to N(0,\tau^2),$$

where μ_H and τ^2 are given in (19) and (20), respectively and the sum $\sum_{\mathbf{k}}$ extends over the set of Fourier frequencies.

Proof. Let $Z_{\mathbf{k}} = U_{\mathbf{k}} - 1$.

$$\begin{aligned}
&\frac{|H|^{1/4}}{N}\int_{\Pi^2}\left(\sum_{\mathbf{k}} K_H(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{k}})(U_{\mathbf{k}}-1)\right)^2 d\boldsymbol{\lambda} - \mu_H = \frac{|H|^{1/4}}{N}\int_{\Pi^2}\sum_{\mathbf{k}} K_H^2(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{k}})Z_{\mathbf{k}}^2 d\boldsymbol{\lambda} \\
&-|H|^{1/4}\int_{\Pi^2} K^2(\mathbf{u})d\mathbf{u} + \frac{|H|^{1/4}}{N}\int_{\Pi^2}\sum_{\mathbf{k}\neq\mathbf{j}} K_H(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{k}})K_H(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{j}})Z_{\mathbf{k}}Z_{\mathbf{j}}d\boldsymbol{\lambda} \\
&= T_1 - \mu_H + T_2.
\end{aligned}$$

Note that, as $N \to \infty$:

$$|E(T_1) - \mu_H| = |H|^{-1/4}\left|\int_{\Pi^2}\frac{1}{N|H|^{1/2}}\sum_{\mathbf{k}} K^2(H^{-1/2}(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{k}}))d\boldsymbol{\lambda} - \int_{\Pi^2} K^2(\mathbf{u})d\mathbf{u}\right| \to 0.$$

For the variance of this first term T_1, since the $Z_{\mathbf{k}}$ are independent zero-mean variables:

$$Var(T_1) = Var\left(|H|^{1/4}\int_{\Pi^2}\frac{1}{N|H|^{1/2}}\sum_{\mathbf{k}} K^2(H^{-1/2}(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{k}}))Z_{\mathbf{k}}d\boldsymbol{\lambda}\right)$$

$$= \frac{1}{N^2|H|^{3/2}} \left(\sum_{\mathbf{k}} \int_{\Pi^2} K^2(H^{-1/2}(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{k}}))d\boldsymbol{\lambda} \right)^2 Var(Z_{\mathbf{0}}^2) = \mathcal{O}\left(N^{-1}|H|^{-1/2}\right).$$

using the same arguments as above. Let's analyze T_2. Define, for $\mathbf{j} \neq \mathbf{k}$

$$a(\mathbf{k}, \mathbf{j}) = a(k_1, k_2, j_1, j_2) = \frac{|H|^{1/4}}{N} \int_{\Pi^2} K_H(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{k}}) K_H(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{j}}) d\boldsymbol{\lambda}$$

and define $a(\mathbf{k}, \mathbf{k}) = 0$; then, T_2 can be decomposed as follows:

$$T_2 = \frac{|H|^{1/4}}{N} \int_{\Pi^2} \sum_{\mathbf{k} \neq \mathbf{j}} K_H(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{k}}) K_H(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{j}}) Z_{\mathbf{k}} Z_{\mathbf{j}} d\boldsymbol{\lambda} = \sum_{\mathbf{k}} \sum_{\mathbf{j}} a(\mathbf{k}, \mathbf{j}) Z_{\mathbf{k}} Z_{\mathbf{j}}.$$

Define $b(\mathbf{k}, \mathbf{j}) = b(k_1, k_2, j_1, j_2)$ as:

$$b(k_1, k_2, j_1, j_2) = a(k_1, k_2, j_1, j_2) + a(k_1, -k_2, j_1, j_2) + a(k_1, k_2, -j_1, j_2) + a(k_1, k_2, j_1, -j_2) +$$
$$a(-k_1, k_2, j_1, j_2) + a(-k_1, -k_2, j_1, j_2) + a(-k_1, k_2, -j_1, j_2) + a(-k_1, k_2, j_1, -j_2) +$$
$$a(-k_1, -k_2, -j_1, j_2) + a(-k_1, -k_2, j_1, -j_2) + a(-k_1, -k_2, -j_1, j_2) + a(k_1, -k_2, -j_1, -j_2) +$$
$$a(k_1, -k_2, -j_1, -j_2) + a(k_1, k_2, -j_1, -j_2) + a(-k_1, k_2, -j_1, -j_2) + a(-k_1, -k_2, -j_1, -j_2).$$

Then, T_2 can be written as $T_2 = Q_N + T_3$ where

$$Q_N = \sum_{k_1=1}^{m_1} \sum_{k_2=1}^{m_2} \sum_{j_1=1}^{m_1} \sum_{j_2=1}^{m_2} b(\mathbf{k}, \mathbf{j}) Z_{\mathbf{k}} Z_{\mathbf{j}}$$

and

$$T_3 = \sum_{k_2=-m_2}^{m_2} \sum_{j_1=-m_1}^{m_1} \sum_{j_2=-m_2}^{m_2} a(k_1, k_2, j_1, j_2) Z_{\mathbf{k}} Z_{\mathbf{j}} \delta_{\mathbf{kj}},$$

where the function

$$\delta_{\mathbf{kj}} = \begin{cases} 1 & \text{if} \quad 0 < \mathbf{1}_{(k_1=0)} + \mathbf{1}_{(k_2=0)} + \mathbf{1}_{(j_1=0)} + \mathbf{1}_{(j_2=0)}, \\ 0 & \text{otherwise,} \end{cases}$$

and $\mathbf{1}$ is the indicator function. Consider any of the addends in the expression of T_3, for instance:

$$\sum_{k_2=-m_2}^{m_2} \sum_{j_1=-m_1}^{m_1} \sum_{j_2=-m_2}^{m_2} a(0, k_2, j_1, j_2) Z_{\mathbf{k}} Z_{\mathbf{j}}.$$

Since the $Z_{\mathbf{k}}$ are independent zero-mean random variables, in order to obtain a non-null expectation term, $\mathbf{k}$ must be equal to $\mathbf{j}$. For $\mathbf{k}_0 = (0, k_2)$:

$$E\left(\sum_{k_2=-m_2}^{m_2} \sum_{j_1=-m_1}^{m_1} \sum_{j_2=-m_2}^{m_2} a(0, k_2, j_1, j_2) Z_{\mathbf{k}} Z_{\mathbf{j}} \right) =$$

$$\frac{|H|^{1/4}}{N} \sum_{k_2=-m_2}^{m_2} \int_{\Pi^2} K_H^2(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{k}_0}) d\boldsymbol{\lambda} E(Z_{\mathbf{k}_0}^2) = \mathcal{O}(n_1^{-1}|H|^{-1/4}).$$

Besides:

$$E\left(\sum_{k_2=-m_2}^{m_2}\sum_{j_1=-m_1}^{m_1}\sum_{j_2=-m_2}^{m_2} a(0,k_2,j_1,j_2)Z_{\mathbf{k}}Z_{\mathbf{j}}\right)^2 =$$

$$\frac{|H|^{1/2}}{N^2}\sum_{k_2=-m_2}^{m_2}\sum_{j_1=-m_1}^{m_1}\sum_{j_2=-m_2}^{m_2}\int_{\Pi^2}\int_{\Pi^2} K_H(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{k}_0})K_H(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{j}})$$

$$\times K_H(\boldsymbol{\omega}-\boldsymbol{\lambda}_{\mathbf{k}_0})K_H(\boldsymbol{\omega}-\boldsymbol{\lambda}_{\mathbf{j}})d\boldsymbol{\lambda}d\boldsymbol{\omega}E\left(Z_{\mathbf{k}_0}Z_{\mathbf{j}}\right)^2$$

and then

$$E\left(\sum_{k_2=-m_2}^{m_2}\sum_{j_1=-m_1}^{m_1}\sum_{j_2=-m_2}^{m_2} a(0,k_2,j_1,j_2)Z_{\mathbf{k}}Z_{\mathbf{j}}\right)^2 = \mathcal{O}(n_1^{-1}).$$

Analogous expressions are obtained for the other addends. Therefore

$$T_2 = \sum_{k_1=1}^{m_1}\sum_{k_2=1}^{m_2}\sum_{j_1=1}^{m_1}\sum_{j_2=1}^{m_2} b(\mathbf{k},\mathbf{j})Z_{\mathbf{k}}Z_{\mathbf{j}} + o_{\mathbb{P}}(1) = Q_N + o_{\mathbb{P}}(1).$$

In order to prove the asymptotic normal distribution of Q_N, we will apply Theorem 5.2 in de Jong (1987). For that purpose, we must write Q_N as a quadratic form, namely $Q_N = \sum_{i,j} c_{i,j}Z_iZ_j$, where i and j are one-dimensional indexes and Z_i are i.i.d. random variables with zero mean and unit variance.

First, define a new subindex for the Fourier frequencies $\boldsymbol{\lambda}_{\mathbf{k}}$, with $\mathbf{k} = (k_1,k_2)$ and $k_l = 0,\pm1,\ldots,\pm m_l$, for $l=1,2$. Consider $\boldsymbol{\lambda}_{\mathbf{k}} = \boldsymbol{\lambda}_{\mathbf{k}'}$ where $\mathbf{k}' = (k_1',k_2')$, with $k_l' = 1,\ldots,m_l' = 2m_l+1$, in such a way that $k_l' = k_l + m_l + 1$ for $l=1,2$. Let $M = m_1' \times m_2'$ and denote by $\mathcal{M}_{M\times M}$ the space of square matrices with size M, that is, with M rows and M columns.

The new coefficients, with one dimensional indexes, are given by the following matrix:

$$A = (c_{ij}), \quad A \in \mathcal{M}_{M\times M},$$

and each entry of this matrix is defined by $c_{ij} = b_{\mathbf{ij}}$ and $c_{ii} = 0$, where the bidimensional indexes $\mathbf{i}$ determine unidimensional indexes i such that:

$$\mathbf{i} = (i_1,i_2), \quad \text{if } (i_1-1)m_2' \le i \le i_1 m_2' \quad \text{and } i = (i_1-1)m_2' + i_2, \tag{56}$$

Now, define the variables:

$$Z_i = Z_{\mathbf{i}}, \quad \text{where } i = (i_1-1)m_2' + i_2, \quad i = 1,\ldots,M.$$

With this definitions, Q_N can be written as a quadratic form with one-dimensional indexes:

$$Q_N = \sum_{i,j} c_{i,j}Z_iZ_j.$$

In order to apply Theorem 5.2 (de Jong (1987)) on the quadratic form Q_N, we must prove that, as $N\to\infty$:

1. There exists a sequence $k(n_1, n_2) \to \infty$ such that

$$k(n_1, n_2)^4 \frac{1}{Var(Q_N)} \max_i \sum_j c_{ij}^2 \to 0.$$

Taking into account that n_1 and n_2 tend to infinity at the same rate, it holds that $E(T_3) = \mathcal{O}(n_1^{-1}|H|^{-1/4})$ and $Var(T_3) \leq E(T_3^2) = \mathcal{O}(n_1^{-1})$. Then, applying that $Var(Q_N) = Var(T_2) + Var(T_3) - 2Cov(T_2, T_3)$, the variance of the quadratic form can be approximated by $Var(T_2)$, since $Var(T_3) \leq \mathcal{O}(n_1^{-1})$ and $|Cov(T_2, T_3)| \leq \sqrt{Var(T_2)Var(T_3)} = \mathcal{O}(n_1^{-1})$, using the Cauchy-Schwarz inequality.

We prove that

$$Var(T_2) = Var\left(\sum_{\mathbf{k}}\sum_{\mathbf{j}} a(\mathbf{k},\mathbf{j})Z_{\mathbf{k}}Z_{\mathbf{j}}\right) = \sum_{\mathbf{k}}\sum_{\mathbf{j}}\sum_{\mathbf{l}}\sum_{\mathbf{m}} a(\mathbf{k},\mathbf{j})a(\mathbf{l},\mathbf{m})E(Z_{\mathbf{k}}Z_{\mathbf{j}}Z_{\mathbf{l}}Z_{\mathbf{m}}) \tag{57}$$

The non-vanishing terms correspond to $\mathbf{k} \neq \mathbf{j}$ and $\mathbf{l} \neq \mathbf{m}$.Then, (57) can be written as:

$$\begin{aligned}
&Var(T_2) = \\
&\frac{|H|^{1/2}}{N^2}\sum_{\mathbf{k}\neq\mathbf{j}}\sum_{\mathbf{l}\neq\mathbf{m}}\int_{\Pi^2} K_H(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{k}})K_H(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{j}})d\boldsymbol{\lambda} \\
&\qquad\qquad \times \int_{\Pi^2} K_H(\boldsymbol{\omega}-\boldsymbol{\lambda}_{\mathbf{l}})K_H(\boldsymbol{\omega}-\boldsymbol{\lambda}_{\mathbf{m}})d\boldsymbol{\omega}E(Z_{\mathbf{k}}Z_{\mathbf{j}}Z_{\mathbf{l}}Z_{\mathbf{m}}) = \\
&\frac{2|H|^{1/2}}{N^2}\sum_{\mathbf{k}}\sum_{\mathbf{j}}\int_{\Pi^2} K_H(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{k}})K_H(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{k}})d\boldsymbol{\lambda}\int_{\Pi^2} K_H(\boldsymbol{\omega}-\boldsymbol{\lambda}_{\mathbf{j}})K_H(\boldsymbol{\omega}-\boldsymbol{\lambda}_{\mathbf{j}})d\boldsymbol{\omega} \\
&-\frac{2|H|^{1/2}}{N^2}\sum_{\mathbf{k}}\int_{\Pi^2} K_H^2(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{k}})d\boldsymbol{\lambda}\int_{\Pi^2} K_H^2(\boldsymbol{\omega}-\boldsymbol{\lambda}_{\mathbf{k}})d\boldsymbol{\omega},
\end{aligned}$$

where $E(Z_{\mathbf{k}}Z_{\mathbf{j}}Z_{\mathbf{l}}Z_{\mathbf{m}}) = 1$ if an only if $\mathbf{k} = \mathbf{j} \neq \mathbf{l} = \mathbf{m}$ or $\mathbf{k} = \mathbf{m} \neq \mathbf{j} = \mathbf{l}$. The second addend is $\mathcal{O}(N^{-1}|H|^{-1/2})$:

$$\begin{aligned}
&\frac{2|H|^{1/2}}{N^2}\sum_{\mathbf{k}}\int_{\Pi^2}\int_{\Pi^2} K_H^2(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{k}})K_H^2(\boldsymbol{\omega}-\boldsymbol{\lambda}_{\mathbf{k}})d\boldsymbol{\lambda}d\boldsymbol{\omega} \\
&= \frac{2}{N^2|H|^{3/2}}\sum_{\mathbf{k}}\left(\int_{\Pi^2} K^2(H^{-1/2}(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{k}}))d\boldsymbol{\lambda}\right)^2 \\
&= \frac{2}{N^2|H|^{1/2}}\sum_{\mathbf{k}}\left(\int_{\Pi^2} K^2(\boldsymbol{\omega}-\boldsymbol{\lambda}_{\mathbf{k}})d\boldsymbol{\omega}\right)^2 = \mathcal{O}(N^{-1}|H|^{-1/2}),
\end{aligned}$$

whereas, for the first term we obtain:

$$\frac{2|H|^{1/2}}{N^2}\sum_{\mathbf{k}}\sum_{\mathbf{j}}\int\int K_H(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{k}})K_H(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{j}})K_H(\boldsymbol{\omega}-\boldsymbol{\lambda}_{\mathbf{k}})K_H(\boldsymbol{\omega}-\boldsymbol{\lambda}_{\mathbf{j}})d\boldsymbol{\lambda}d\boldsymbol{\omega}$$
$$=\frac{2|H|^{1/2}}{N^2}\sum_{\mathbf{k}}\sum_{\mathbf{j}}\left(\int K_H(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{k}})K_H(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{j}})d\boldsymbol{\lambda}\right)^2$$
$$=\frac{2}{N^2|H|^{1/2}}\sum_{\mathbf{k}}\sum_{\mathbf{j}}\left(\int K(\mathbf{u})K\left(\mathbf{u}+H^{-1/2}2\pi\left(\frac{k_1-j_1}{n_1},\frac{k_2-j_2}{n_2}\right)^T\right)d\mathbf{u}\right)^2$$
$$=\frac{2}{N|H|^{1/2}}\sum_{k_1=-2m_1}^{2m_1}\sum_{k_2=-2m_2}^{2m_2}c(\mathbf{k},N)\left(\int K(\mathbf{u})K\left(\mathbf{u}+H^{-1/2}2\pi\left(\frac{k_1}{n_1},\frac{k_2}{n_2}\right)^T\right)d\mathbf{u}\right)^2$$
$$\to\frac{1}{2\pi^2}\int_{2\Pi^2}\left(\int_{\Pi^2}K(\mathbf{u})K(\mathbf{u}+\mathbf{x})d\mathbf{u}\right)^2d\mathbf{x},$$

where $2\Pi^2=[-2\pi,2\pi]\times[-2\pi,2\pi]$ and $c(\mathbf{k},N)=\frac{2m_1+1-|k_1|}{n_1}\frac{2m_2+1-|k_2|}{n_2}$. Therefore, in order to prove the required condition, since c_{ij}^2 is a squared sum of $a(\mathbf{i},\mathbf{j})$ terms, we prove the condition for one of the addends, that is, for $a^2(\mathbf{i},\mathbf{j})$. Besides, using that $K_H(\cdot)\le|H|^{-1/2}C$, for $0<C<\infty$:

$$k^4(n_1,n_2)\max_{\mathbf{k}}\sum_{\mathbf{j}}a^2(\mathbf{k},\mathbf{j})=$$
$$k^4(n_1,n_2)\max_{\mathbf{k}}\sum_{\mathbf{j}}\left(\frac{|H|^{1/4}}{N}\int K_H(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{k}})K_H(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{j}})d\boldsymbol{\lambda}\right)^2=$$
$$=\mathcal{O}(k^4(n_1,n_2)N^{-1}|H|^{-1/2}).$$

So, this condition is satisfied for all $k(n_1,n_2)\to\infty$ such that $\dfrac{k^4(n_1,n_2)}{n_1n_2|H|^{1/2}}\to 0$.

2. We also have to check that $\max_{\mathbf{k}}E(Z_{\mathbf{k}}^2)\mathbf{1}_{\{|Z_{\mathbf{k}}|>k(n_1,n_2)\}}\to 0$, but this assertion follows just taking into account that $Z_{\mathbf{k}}$ are identically distributed with $E(Z_{\mathbf{k}}^2)=1$.

3. It remains to show that $\dfrac{\max_i\mu_i^2}{Var(Q_N)}\to 0$ where μ_i, $i=1,2,\ldots,M$ are the eigenvalues of the matrix $A=(c_{ij})$ define above.

   The matrix A is symmetric, because the c_{ij} entries are defined in terms of the $a(\mathbf{i},\mathbf{j})$ terms defined above. Besides, the $a(\mathbf{i},\mathbf{j})$ satisfy that $a(\mathbf{i},\mathbf{j})=a(\mathbf{j},\mathbf{i})$ and

$$\sum_{\mathbf{j}}|a(\mathbf{i},\mathbf{j})|=\mathcal{O}(|H|^{1/4}).$$

   Thus, the same condition applies on the c_{ij} terms.

Now, to prove the required condition, since A is a symmetric $M\times M$ matrix, there exists an ortogonal matrix U such that $U^{-1}AU$ is diagonal. This result implies that B

is diagonalizable with real eigenvalues, $\{\mu_i\}$, with $i = 1, \ldots, M$, with $M = m'_1 \times m'_2$. The $\|\cdot\|_\infty$ norm of the matrix B is given by:

$$\|A\|_\infty = \max_i \sum_j |c_{ij}|,$$

where the maximum is taken over $i \in \{1, \ldots, M\}$ and consider the spectral ratio of the matrix:

$$\rho(A) = \max_i |\mu_i|.$$

The spectral ratio of the matrix can be bounded by any norm in the matrix space $\mathcal{M}_{M\times M}$; therefore, for the particular case of the supremun norm $\|\cdot\|_\infty$:

$$\max_i |\mu_i| \le \max_i \sum_j |c_{ij}|.$$

To prove the result, just take into account that:

$$\max_i \mu_i^2 \le \left(\max_i |\mu_i|\right)^2 \le \left(\max_i \sum_j |c_{ij}|\right)^2.$$

Then, since $|H|^{1/2} \to 0$ and $Var(Q_N) \to \tau^2$:

$$\frac{\max_i \mu_i^2}{Var(Q_N)} \le \frac{\left(\max_i \sum_j |c_{ij}|\right)^2}{Var(Q_N)} = \frac{\mathcal{O}(|H|^{1/2})}{Var(Q_N)} \to 0.$$

□

Lemma 6. *Let T_P^0 denote the test statistic in (18) assuming that the true parameter is given by θ_0. Then, under assumptions in Theorem 1:*

$$T_P = T_P^0 + o_{\mathbb{P}}(1).$$

Proof. The test statistic T_P^0 is given by

$$T_P^0 = N|H|^{1/4} \int_{\Pi^2} \left(\frac{1}{N|H|^{1/2}} \sum_{\mathbf{k}} K(H^{-1/2}(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{k}})) \left(\frac{I(\boldsymbol{\lambda}_{\mathbf{k}})}{f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{k}})} - 1 \right) \right)^2 d\boldsymbol{\lambda}. \tag{58}$$

Note that:

$$\frac{I(\boldsymbol{\lambda}_{\mathbf{k}})}{f_{\hat\theta}(\boldsymbol{\lambda}_{\mathbf{k}})} - 1 = \left(\frac{I(\boldsymbol{\lambda}_{\mathbf{k}})}{f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{k}})} - 1 \right) - \left(\frac{f_{\hat\theta}(\boldsymbol{\lambda}_{\mathbf{k}}) - f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{k}})}{f_{\hat\theta}(\boldsymbol{\lambda}_{\mathbf{k}})} \right) \frac{I(\boldsymbol{\lambda}_{\mathbf{k}})}{f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{k}})}. \tag{59}$$

Therefore,

$$T_P = T_P^0 + N|H|^{1/4} \int_{\Pi^2} \left(\frac{1}{N} \sum_{\mathbf{k}} K_H(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{k}}) \frac{I(\boldsymbol{\lambda}_{\mathbf{k}})}{f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{k}})} \left(\frac{f_{\hat\theta}(\boldsymbol{\lambda}_{\mathbf{k}}) - f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{k}})}{f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{k}})} \right) \right)^2 d\boldsymbol{\lambda}$$

$$-\frac{2|H|^{1/4}}{N} \int_{\Pi^2} \sum_{\mathbf{k}} \sum_{\mathbf{j}} K_H(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{k}}) K_H(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{j}}) \left(\frac{I(\boldsymbol{\lambda}_{\mathbf{k}})}{f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{k}})} - 1 \right) \frac{f_{\hat\theta}(\boldsymbol{\lambda}_{\mathbf{j}}) - f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{j}})}{f_{\hat\theta}(\boldsymbol{\lambda}_{\mathbf{j}})} \frac{I(\boldsymbol{\lambda}_{\mathbf{j}})}{f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{j}})} d\boldsymbol{\lambda}$$

For the second addend, using Lemma 1 and the fact that

$$\int_{\Pi^2} \left(\frac{1}{N} \sum_{\mathbf{k}} K_H(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{k}}) \frac{I(\boldsymbol{\lambda}_{\mathbf{k}})}{f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{k}})} \right)^2 d\boldsymbol{\lambda} = \mathcal{O}_{\mathbb{P}}(1)$$

we have

$$\begin{aligned}
&N|H|^{1/4} \int_{\Pi^2} \left(\frac{1}{N} \sum_{\mathbf{k}} K_H(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{k}}) \frac{I(\boldsymbol{\lambda}_{\mathbf{k}})}{f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{k}})} \left(\frac{f_{\hat{\theta}}(\boldsymbol{\lambda}_{\mathbf{k}}) - f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{k}})}{f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{k}})} \right) \right)^2 d\boldsymbol{\lambda} \\
&\leq N|H|^{1/4} \left(\sup_{\mathbf{k}} \left| \frac{f_{\hat{\theta}}(\boldsymbol{\lambda}_{\mathbf{k}}) - f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{k}})}{f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{k}})} \right| \right)^2 \int_{\Pi^2} \left(\frac{1}{N} \sum_{\mathbf{k}} K_H(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{k}}) \frac{I(\boldsymbol{\lambda}_{\mathbf{k}})}{f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{k}})} \right)^2 d\boldsymbol{\lambda} \\
&= \mathcal{O}_{\mathbb{P}}(|H|^{1/4}).
\end{aligned}$$

For the last addend:

$$\begin{aligned}
&\frac{|H|^{1/4}}{N} \int_{\Pi^2} \sum_{\mathbf{k}} \sum_{\mathbf{j}} K_H(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{k}}) K_H(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{j}}) \left(\frac{I(\boldsymbol{\lambda}_{\mathbf{k}})}{f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{k}})} - 1 \right) \frac{f_{\hat{\theta}}(\boldsymbol{\lambda}_{\mathbf{j}}) - f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{j}})}{f_{\hat{\theta}}(\boldsymbol{\lambda}_{\mathbf{j}})} \frac{I(\boldsymbol{\lambda}_{\mathbf{j}})}{f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{j}})} d\boldsymbol{\lambda} = \\
&M_1 + M_2,
\end{aligned}$$

where

$$\begin{aligned}
M_1 = \frac{|H|^{1/4}}{N} \int_{\Pi^2} \sum_{\mathbf{k}} \sum_{\mathbf{j}} K_H(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{k}}) K_H(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{j}}) \left(\frac{I(\boldsymbol{\lambda}_{\mathbf{k}})}{f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{k}})} - 1 \right) \\
\times \frac{f_{\hat{\theta}}(\boldsymbol{\lambda}_{\mathbf{j}}) - f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{j}})}{f_{\hat{\theta}}(\boldsymbol{\lambda}_{\mathbf{j}})} \left(\frac{I(\boldsymbol{\lambda}_{\mathbf{j}})}{f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{j}})} - 1 \right) d\boldsymbol{\lambda}
\end{aligned}$$

and

$$M_2 = \frac{|H|^{1/4}}{N} \int_{\Pi^2} \sum_{\mathbf{k}} \sum_{\mathbf{j}} K_H(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{k}}) K_H(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{j}}) \left(\frac{I(\boldsymbol{\lambda}_{\mathbf{k}})}{f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{k}})} - 1 \right) \frac{f_{\hat{\theta}}(\boldsymbol{\lambda}_{\mathbf{j}}) - f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{j}})}{f_{\hat{\theta}}(\boldsymbol{\lambda}_{\mathbf{j}})} d\boldsymbol{\lambda}.$$

We will prove that $M_1 = o_{\mathbb{P}}(1)$. Recall that

$$\frac{I(\boldsymbol{\lambda}_{\mathbf{k}})}{f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{k}})} - 1 = W_{\mathbf{k}} + \frac{R_N(\boldsymbol{\lambda}_{\mathbf{k}})}{f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{k}})},$$

where $W_{\mathbf{k}} = V_{\mathbf{k}} - 1$, and the $V_{\mathbf{k}}$ are independent identically distributed $Exp(1)$. Then,

$$\begin{aligned}
M_1 &= \frac{|H|^{1/4}}{N} \int_{\Pi^2} \sum_{\mathbf{k}} \sum_{\mathbf{j}} K_H(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{k}}) K_H(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{j}}) d\boldsymbol{\lambda} \\
&\times \left\{ W_{\mathbf{k}} W_{\mathbf{j}} \left(\frac{f_{\hat{\theta}}(\boldsymbol{\lambda}_{\mathbf{j}}) - f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{j}})}{f_{\hat{\theta}}(\boldsymbol{\lambda}_{\mathbf{j}})} \right) + W_{\mathbf{k}} \frac{R_N(\boldsymbol{\lambda}_{\mathbf{j}})}{f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{j}})} \left(\frac{f_{\hat{\theta}}(\boldsymbol{\lambda}_{\mathbf{j}}) - f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{j}})}{f_{\hat{\theta}}(\boldsymbol{\lambda}_{\mathbf{j}})} \right) \right. \\
&+ W_{\mathbf{j}} \frac{R_N(\boldsymbol{\lambda}_{\mathbf{k}})}{f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{k}})} \left(\frac{f_{\hat{\theta}}(\boldsymbol{\lambda}_{\mathbf{j}}) - f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{j}})}{f_{\hat{\theta}}(\boldsymbol{\lambda}_{\mathbf{j}})} \right) \\
&\left. + \frac{R_N(\boldsymbol{\lambda}_{\mathbf{j}})}{f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{j}})} \frac{R_N(\boldsymbol{\lambda}_{\mathbf{k}})}{f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{k}})} \left(\frac{f_{\hat{\theta}}(\boldsymbol{\lambda}_{\mathbf{j}}) - f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{j}})}{f_{\hat{\theta}}(\boldsymbol{\lambda}_{\mathbf{j}})} \right) \right\} = C_1 + C_2 + C_3 + C_4.
\end{aligned}$$

In order to prove the bounds for C_j, $j = 1, 2, 3, 4$, we have to consider the Taylor expansion of $f_{\hat{\theta}}(\boldsymbol{\lambda})$ around $f_{\theta_0}(\boldsymbol{\lambda})$, for a fixed $\boldsymbol{\lambda}$:

$$f_{\hat{\theta}}(\boldsymbol{\lambda}) = f_{\theta_0}(\boldsymbol{\lambda}) + (\hat{\theta} - \theta_0)^T \nabla f_{\theta_0}(\boldsymbol{\lambda}) + \frac{1}{2}(\hat{\theta} - \theta_0)^T \nabla^2 f_{\tilde{\theta}}(\boldsymbol{\lambda})(\hat{\theta} - \theta_0),$$

where $\|\tilde{\theta} - \theta_0\| \leq \|\hat{\theta} - \theta_0\|$. By similar arguments to those in Lemma 1, for $\boldsymbol{\lambda}_{\mathbf{j}}$

$$\frac{f_{\hat{\theta}}(\boldsymbol{\lambda}_j) - f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{j}})}{f_{\hat{\theta}}(\boldsymbol{\lambda}_{\mathbf{j}})} = \mathcal{O}_{\mathbb{P}}(1) \left((\hat{\theta} - \theta_0)^T \nabla f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{j}}) + \frac{1}{2}(\hat{\theta} - \theta_0)^T \nabla^2 f_{\tilde{\theta}}(\boldsymbol{\lambda}_{\mathbf{j}})(\hat{\theta} - \theta_0) \right), \quad (60)$$

and the $\mathcal{O}_{\mathbb{P}}(1)$ factor is uniform in $\mathbf{j}$. We will see that $C_1 = \mathcal{O}_{\mathbb{P}}(N^{-1/2}) + \mathcal{O}_{\mathbb{P}}(|H|^{1/4})$. Taking into account (60), C_1 can be written as:

$$C_1 = \mathcal{O}_{\mathbb{P}}(1)(\hat{\theta} - \theta_0)^T \frac{|H|^{1/4}}{N} \int_{\Pi^2} \sum_{\mathbf{k}} \sum_{\mathbf{j}} K_H(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{k}}) K_H(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{j}}) W_{\mathbf{k}} W_{\mathbf{j}} \nabla f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{j}}) d\boldsymbol{\lambda}$$

$$+\mathcal{O}_{\mathbb{P}}(1)(\hat{\theta} - \theta_0)^T \frac{|H|^{1/4}}{2N} \int_{\Pi^2} \sum_{\mathbf{k}} \sum_{\mathbf{j}} K_H(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{k}}) K_H(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{j}}) W_{\mathbf{k}} W_{\mathbf{j}} \nabla^2 f_{\tilde{\theta}}(\boldsymbol{\lambda}_{\mathbf{j}}) d\boldsymbol{\lambda}(\hat{\theta} - \theta_0).$$

Since

$$\frac{|H|^{1/4}}{N} \int_{\Pi^2} \sum_{\mathbf{k}} \sum_{\mathbf{j}} K_H(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{k}}) K_H(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{j}}) W_{\mathbf{k}} W_{\mathbf{j}} d\boldsymbol{\lambda} = \mathcal{O}_{\mathbb{P}}(1)$$

and the derivatives of f_θ are uniformly bounded, the first addend in C_1 is $\mathcal{O}_{\mathbb{P}}(N^{-1/2})$. Taking into account that $(\hat{\theta} - \theta_0) = \mathcal{O}_{\mathbb{P}}(N^{-1/2})$, the second addend is $\mathcal{O}_{\mathbb{P}}(|H|^{1/4})$. In order to obtain a bound for C_2, one should consider the results in Lemma 4. From Taylor expansion (60), C_2 can be written as:

$$C_2 = \mathcal{O}_{\mathbb{P}}(1)(\hat{\theta} - \theta_0)^T \frac{|H|^{1/4}}{N} \int_{\Pi^2} \sum_{\mathbf{k}} \sum_{\mathbf{j}} K_H(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{k}}) K_H(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{j}}) W_{\mathbf{k}} \frac{R_N(\boldsymbol{\lambda}_{\mathbf{j}})}{f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{j}})} \nabla f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{j}}) d\boldsymbol{\lambda}$$

$$+\mathcal{O}_{\mathbb{P}}(1)(\hat{\theta} - \theta_0)^T \frac{|H|^{1/4}}{2N} \int_{\Pi^2} \sum_{\mathbf{k}} \sum_{\mathbf{j}} K_H(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{k}}) K_H(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{j}}) W_{\mathbf{k}} \frac{R_N(\boldsymbol{\lambda}_{\mathbf{j}})}{f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{j}})} \nabla^2 f_{\tilde{\theta}}(\boldsymbol{\lambda}_{\mathbf{j}})(\hat{\theta} - \theta_0).$$

From Lemma 4, we have that:

$$\frac{|H|^{1/4}}{2N} \int_{\Pi^2} \sum_{\mathbf{k}} \sum_{\mathbf{j}} K_H(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{k}}) K_H(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{j}}) W_{\mathbf{k}} \frac{R_N(\boldsymbol{\lambda}_{\mathbf{j}})}{f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{j}})} d\boldsymbol{\lambda} = o_{\mathbb{P}}(1).$$

Then, the first addend in C_2 is $\mathcal{O}_{\mathbb{P}}(N^{-1/2}) o_{\mathbb{P}}(1)$. For the second addend, one should note that $|R_N(\boldsymbol{\lambda}_{\mathbf{j}})| = \mathcal{O}_{\mathbb{P}}(N^{-1/2})$, from Lemma 2. Then:

$$\mathcal{O}_{\mathbb{P}}(1)(\hat{\theta} - \theta_0)^T \frac{|H|^{1/4}}{2N} \int_{\Pi^2} \sum_{\mathbf{k}} \sum_{\mathbf{j}} K_H(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{k}}) K_H(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{j}}) W_{\mathbf{k}} \frac{R_N(\boldsymbol{\lambda}_{\mathbf{j}})}{f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{j}})} \nabla^2 f_{\tilde{\theta}}(\boldsymbol{\lambda}_{\mathbf{j}})(\hat{\theta} - \theta_0) d\boldsymbol{\lambda} =$$

$$= \mathcal{O}_{\mathbb{P}}(N^{-1/2} |H|^{1/4}).$$

The third addend C_3 can be bounded using the same arguments as in the proof for C_2. For the last addend C_4, and taking also into account Lemma 2:

$$|C_4| \leq \sum_{\mathbf{j}} \left| \frac{f_{\hat{\theta}}(\boldsymbol{\lambda}_{\mathbf{j}}) - f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{j}})}{f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{j}})} \right| \frac{|H|^{1/4}}{N} \int_{\Pi^2} \sum_{\mathbf{k}} \sum_{\mathbf{j}} K_H(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{k}}) K_H(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{j}}) \frac{|R_N(\boldsymbol{\lambda}_{\mathbf{j}})|}{f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{j}})} \frac{|R_N(\boldsymbol{\lambda}_{\mathbf{k}})|}{f_{\theta_0}(\boldsymbol{\lambda}_{\mathbf{k}})} d\boldsymbol{\lambda}$$

$$= \mathcal{O}_{\mathbb{P}}(N^{-1/2})\mathcal{O}_{\mathbb{P}}(|H|^{1/4}N)\mathcal{O}_{\mathbb{P}}(N^{-1}) = \mathcal{O}_{\mathbb{P}}(|H|^{1/4}N^{-1/2}).$$

$M_2 = o_{\mathbb{P}}(1)$ can be proved using similar arguments. □

Lemma 7. *If $\theta = \theta_0$ is the true parameter, under assumptions (1)-(4):*

$$T_P^0 - \mu_H \to N(0, \tau^2),$$

as $N \to \infty$, where μ_H and τ^2 are given in (19) and (20), respectively and T_P^0 is given in (58).

Proof. Recall the expression for the periodogram

$$I(\boldsymbol{\lambda}_\mathbf{k}) = f(\boldsymbol{\lambda}_\mathbf{k})V_\mathbf{k} + R_N(\boldsymbol{\lambda}_\mathbf{k}), \tag{61}$$

where $\{\boldsymbol{\lambda}_\mathbf{k}\}$ denote the Fourier frequencies and recall the notation $W_\mathbf{k} = 1 - V_\mathbf{k}$ (where $V_\mathbf{k}$ are independent identically distributed random variables with $Exp(1)$ distribution) introduced in Lemma 6. Then:

$$\frac{I(\boldsymbol{\lambda}_\mathbf{k})}{f_{\theta_0}(\boldsymbol{\lambda}_\mathbf{k})} - 1 = W_\mathbf{k} + \frac{R_N(\boldsymbol{\lambda}_\mathbf{k})}{f_{\theta_0}(\boldsymbol{\lambda}_\mathbf{k})}.$$

The statistic T_P^0 can be decomposed in three addends in the following way:

$$T_P^0 - \mu_H = \frac{|H|^{1/4}}{N}\int_{\Pi^2}\left(\sum_\mathbf{k} K_H(\boldsymbol{\lambda} - \boldsymbol{\lambda}_\mathbf{k})W_\mathbf{k}\right)^2 d\boldsymbol{\lambda} - \mu_H \tag{62}$$

$$+\frac{|H|^{1/4}}{N}\int_{\Pi^2}\left(\sum_\mathbf{k} K_H(\boldsymbol{\lambda} - \boldsymbol{\lambda}_\mathbf{k})\frac{R_N(\boldsymbol{\lambda}_\mathbf{k})}{f_{\theta_0}(\boldsymbol{\lambda}_\mathbf{k})}\right)^2 d\boldsymbol{\lambda} \tag{63}$$

$$+\frac{2|H|^{1/4}}{N}\int_{\Pi^2}\sum_\mathbf{k}\sum_\mathbf{j} K_H(\boldsymbol{\lambda} - \boldsymbol{\lambda}_\mathbf{k})K_H(\boldsymbol{\lambda} - \boldsymbol{\lambda}_\mathbf{j})W_\mathbf{k}\frac{R_N(\boldsymbol{\lambda}_\mathbf{j})}{f_{\theta_0}(\boldsymbol{\lambda}_\mathbf{j})}d\boldsymbol{\lambda} \tag{64}$$

From Lemma 3, (63) tends to zero in probability. Also, from Lemma 4, (64) tends to zero in probability. This lemma is proved considering Lemma 5. □

Proof. **Proof of Theorem 1.** Theorem 1 is proved combining the results in Lemma 6 and Lemma 7. □

8.2 Proof of theorem 2.

Before proving the theorem, we must verify that (21) holds. The prove of the following lemma is obtained generalizing Theorem 3.2 in Dahlhaus and Wefelmeyer (1996).

Lemma 8. *Under assumptions (2) and (5), if f is bounded and bounded away from cero, then*

$$\sqrt{N}(\hat{\theta} - \theta^*) - \sqrt{N}\int_{\Pi^2} W(\boldsymbol{\lambda})(I(\boldsymbol{\lambda}) - f(\boldsymbol{\lambda}))d\boldsymbol{\lambda} \to 0$$

in probability, where

$$W(\boldsymbol{\lambda}) = -\mathcal{H}^{-1}\nabla f_\theta^{-1}(\boldsymbol{\lambda})|_{\theta=\theta^*}, \qquad \mathcal{H} = \int_{\Pi^2}\nabla^2 G(\theta^*, f, \boldsymbol{\lambda})d\boldsymbol{\lambda},$$

$$G(\theta, f, \boldsymbol{\lambda}) = \log f_\theta(\boldsymbol{\lambda}) + \frac{f(\boldsymbol{\lambda})}{f_\theta(\boldsymbol{\lambda})}.$$

Proof. Let's write the Kullback-Lebiler discrepancy between f and f_θ

$$L(\theta, f) = \int_{\Pi^2} \left(\log f_\theta(\boldsymbol{\lambda}) + \frac{f(\boldsymbol{\lambda})}{f_\theta(\boldsymbol{\lambda})} \right) d\boldsymbol{\lambda}, \tag{65}$$

in a more general form as:

$$L(\theta, f) = \int_{\Pi^2} G(\theta, f, \boldsymbol{\lambda}) d\boldsymbol{\lambda}.$$

In our particular case, the function in the integrand is given by:

$$G(\theta, f, \boldsymbol{\lambda}) = a_\theta(\boldsymbol{\lambda}) + b_\theta(\boldsymbol{\lambda}) f(\boldsymbol{\lambda}) \quad \text{where } a_\theta(\boldsymbol{\lambda}) = \log f_\theta(\boldsymbol{\lambda}), \quad b_\theta(\boldsymbol{\lambda}) = f_\theta^{-1}(\boldsymbol{\lambda}).$$

Then, the $\mathcal{H}$ matrix can be written as:

$$\mathcal{H} = \int_{\Pi^2} \left(\nabla^2 a_{\theta^*}(\boldsymbol{\lambda}) + \nabla^2 b_{\theta^*}(\boldsymbol{\lambda}) f(\boldsymbol{\lambda}) \right) d\boldsymbol{\lambda},$$

where θ^* gives the best fit in $\mathcal{F}_\theta$. Considering $L(\theta, I)$ the analogous expression to (65), but replacing f by the periodogram I, it is straightforward to see that $L(\hat{\theta}, I) \leq L(\theta^*, I)$ and $L(\theta^*, f) \leq L(\hat{\theta}, f)$, only recalling the definitions of $\hat{\theta}$ and θ^*:

$$\hat{\theta} = \arg\min_\theta L(\theta, I) \text{ and } \theta^* = \arg\min_\theta L(\theta, f).$$

Since

$$\sup_\theta |L(\theta, I) - L(\theta, f)| \to 0 \tag{66}$$

in probability (see Dahlhaus and Wefelmeyer (1996), Lemma A.7), then the Kullback-Leibler discrepancy $L(\hat{\theta}, f)$ converges to $L(\theta^*, f)$ in probability. This result is proved by the convergence of Cesaro sums of the Fourier transform of $f_\theta^{-1}(\boldsymbol{\lambda})$.

Therefore, $\hat{\theta}$ tends to θ^* in probability. The result follows from a Taylor expansion of $\nabla L(\hat{\theta}, I)$ around $\nabla L(\theta^*, I)$. Note that $\nabla L(\hat{\theta}, I) = 0$, then:

$$0 = \nabla L(\theta^*, I) + \nabla^2 L(\tilde{\theta}, I)(\hat{\theta} - \theta^*). \tag{67}$$

for $\tilde{\theta}$ such that $\|\tilde{\theta} - \theta^*\| \leq \|\hat{\theta} - \theta^*\|$. For the first addend:

$$\begin{aligned}
\nabla L(\theta^*, I) &= \int_{\Pi^2} \nabla G(\theta^*, I, \boldsymbol{\lambda}) d\boldsymbol{\lambda} = \int_{\Pi^2} \left(\nabla a_{\theta^*}(\boldsymbol{\lambda}) + \nabla b_{\theta^*}(\boldsymbol{\lambda}) I(\boldsymbol{\lambda}) \right) d\boldsymbol{\lambda} \\
&= \int_{\Pi^2} \nabla G(\theta^*, f, \boldsymbol{\lambda}) d\boldsymbol{\lambda} + \int_{\Pi^2} \nabla b_{\theta^*}(\boldsymbol{\lambda})(I(\boldsymbol{\lambda}) - f(\boldsymbol{\lambda})) d\boldsymbol{\lambda} \\
&= \int_{\Pi^2} \nabla b_{\theta^*}(\boldsymbol{\lambda})(I(\boldsymbol{\lambda}) - f(\boldsymbol{\lambda})) d\boldsymbol{\lambda},
\end{aligned}$$

since the first term is zero. For the second addend in (67), it can be seen that, for $\tilde{\theta}$ such that $\|\tilde{\theta} - \theta^*\| \leq \|\hat{\theta} - \theta^*\|$:

$$\nabla^2 L(\tilde{\theta}, I) = \int_{\Pi^2} \nabla^2 G(\tilde{\theta}, f, \boldsymbol{\lambda}) d\boldsymbol{\lambda} + \int_{\Pi^2} \nabla^2 b_{\tilde{\theta}}(\boldsymbol{\lambda})(I(\boldsymbol{\lambda}) - f(\boldsymbol{\lambda})) d\boldsymbol{\lambda}.$$

Then, (67) can be written as:

$$-\int_{\Pi^2} \nabla b_{\theta^*}(\boldsymbol{\lambda})(I(\boldsymbol{\lambda}) - f(\boldsymbol{\lambda})) d\boldsymbol{\lambda} = \int_{\Pi^2} \nabla^2 G(\tilde{\theta}, f, \boldsymbol{\lambda}) d\boldsymbol{\lambda} + \int_{\Pi^2} \nabla^2 b_{\tilde{\theta}}(\boldsymbol{\lambda})(I(\boldsymbol{\lambda}) - f(\boldsymbol{\lambda})) d\boldsymbol{\lambda}.$$

Provided that $\mathcal{H}$ is non-singular:

$$-\sqrt{N}\mathcal{H}^{-1}\int_{\Pi^2}\nabla b_{\theta^*}(\boldsymbol{\lambda})(I(\boldsymbol{\lambda})-f(\boldsymbol{\lambda}))d\boldsymbol{\lambda} = \tag{68}$$

$$\sqrt{N}\mathcal{H}^{-1}\left(\int_{\Pi^2}\nabla^2 G(\tilde{\theta},f,\boldsymbol{\lambda})d\boldsymbol{\lambda}+\int_{\Pi^2}\nabla^2 b_{\tilde{\theta}}(\boldsymbol{\lambda})(I(\boldsymbol{\lambda})-f(\boldsymbol{\lambda}))d\boldsymbol{\lambda}\right)(\hat{\theta}-\theta^*). \tag{69}$$

By the smoothness of G,

$$\int_{\Pi^2}\nabla^2 G(\tilde{\theta},f,\boldsymbol{\lambda})d\boldsymbol{\lambda}\to\mathcal{H} \tag{70}$$

in probability, and by Lemma A.7 in (Dahlhaus and Wefelmeyer (1996)):

$$\int_{\Pi^2}\nabla^2 b_{\tilde{\theta}}(\boldsymbol{\lambda})(I(\boldsymbol{\lambda})-f(\boldsymbol{\lambda}))d\boldsymbol{\lambda}\to 0 \tag{71}$$

also in probability.

The result is proved replacing (70) and (71) in (68)-(69). □

Proof. **Proof of Theorem 2.** Once we have obtained the $\sqrt{N}$-consistency of $\hat{\theta}$ as an estimator of θ^*, the proof of the theorem is analogous as the proof of Theorem 3 in Paparoditis (2000). Note that:

$$\frac{I(\boldsymbol{\lambda}_{\mathbf{k}})}{f_{\hat{\theta}}(\boldsymbol{\lambda}_{\mathbf{k}})}-1=\left(\frac{I(\boldsymbol{\lambda}_{\mathbf{k}})}{f(\boldsymbol{\lambda}_{\mathbf{k}})}-1\right)+\frac{I(\boldsymbol{\lambda}_{\mathbf{k}})}{f_{\theta^*}(\boldsymbol{\lambda}_{\mathbf{k}})}\left(\frac{f_{\theta^*}(\boldsymbol{\lambda}_{\mathbf{k}})}{f_{\hat{\theta}}(\boldsymbol{\lambda}_{\mathbf{k}})}-1\right)+\frac{I(\boldsymbol{\lambda}_{\mathbf{k}})}{f(\boldsymbol{\lambda}_{\mathbf{k}})}\left(\frac{f(\boldsymbol{\lambda}_{\mathbf{k}})}{f_{\theta^*}(\boldsymbol{\lambda}_{\mathbf{k}})}-1\right),$$

and recall that $I(\boldsymbol{\lambda}_{\mathbf{k}})/f(\boldsymbol{\lambda}_{\mathbf{k}})=V_{\mathbf{k}}+R_N(\boldsymbol{\lambda}_{\mathbf{k}})/f(\boldsymbol{\lambda}_{\mathbf{k}})$. Then,

$$N^{-1}|H|^{-1/4}T_P=\int_{\Pi^2}\left(\frac{1}{N|H|^{1/2}}\sum_{\mathbf{k}}K(H^{-1/2}(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{k}}))\frac{I(\boldsymbol{\lambda}_{\mathbf{k}})}{f(\boldsymbol{\lambda}_{\mathbf{k}})}\left(\frac{f(\boldsymbol{\lambda}_{\mathbf{k}})}{f_{\theta^*}(\boldsymbol{\lambda}_{\mathbf{k}})}-1\right)\right)^2 d\boldsymbol{\lambda}+o_{\mathbb{P}}(1).$$

The first addend can be decomposed in two terms:

$$\int_{\Pi^2}\left(N^{-1}\sum_{\mathbf{k}}K_H(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{k}})\left(\frac{f(\boldsymbol{\lambda}_{\mathbf{k}})}{f_{\theta^*}(\boldsymbol{\lambda}_{\mathbf{k}})}-1\right)\right)^2 d\boldsymbol{\lambda}$$
$$+\int_{\Pi^2}\left(N^{-1}\sum_{\mathbf{k}}K_H(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{k}})\left(W_{\mathbf{k}}+\frac{R_N(\boldsymbol{\lambda}_{\mathbf{k}})}{f(\boldsymbol{\lambda}_{\mathbf{k}})}\right)\left(\frac{f(\boldsymbol{\lambda}_{\mathbf{k}})}{f_{\theta^*}(\boldsymbol{\lambda}_{\mathbf{k}})}-1\right)\right)^2 d\boldsymbol{\lambda},$$

where $W_{\mathbf{k}}=V_{\mathbf{k}}-1$. From Lemma 4:

$$N^{-2}\int_{\Pi^2}\sum_{\mathbf{k}}\sum_{\mathbf{j}}K_H(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{k}})K_H(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{j}})f(\boldsymbol{\lambda}_{\mathbf{k}})/f_{\theta^*}(\boldsymbol{\lambda}_{\mathbf{k}})f(\boldsymbol{\lambda}_{\mathbf{j}})/f_{\theta^*}(\boldsymbol{\lambda}_{\mathbf{j}})W_{\mathbf{k}}\frac{R_N(\boldsymbol{\lambda}_{\mathbf{j}})}{f(\boldsymbol{\lambda}_{\mathbf{j}})}d\boldsymbol{\lambda}=o_{\mathbb{P}}(1).$$

From Lemma 3, we have:

$$N^{-2}\int_{\Pi^2}\left(\sum_{\mathbf{k}}K_H(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{k}})f(\boldsymbol{\lambda}_{\mathbf{k}})/f_{\theta^*}(\boldsymbol{\lambda}_{\mathbf{k}})\frac{R_N(\boldsymbol{\lambda}_{\mathbf{k}})}{f(\boldsymbol{\lambda}_{\mathbf{k}})}\right)^2 d\boldsymbol{\lambda}=o_{\mathbb{P}}(1).$$

Besides,

$$N^{-2}\int_{\Pi^2}\left(\sum_{\mathbf{k}}K_H(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{k}})W_{\mathbf{k}}\right)^2 d\boldsymbol{\lambda}\to 0.$$

Then,

$$N^{-1}|H|^{-1/4}T_P = \int_{\Pi^2} \left(N^{-1} \sum_{\mathbf{k}} K_H(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{k}}) \left(\frac{f(\boldsymbol{\lambda}_{\mathbf{k}})}{f_{\theta^*}(\boldsymbol{\lambda}_{\mathbf{k}})} - 1 \right) \right)^2 d\boldsymbol{\lambda} + o_{\mathbb{P}}(1). \tag{72}$$

Besides, this uniform convergence holds:

$$N^{-1} \sum_{\mathbf{k}} K_H(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{k}}) \frac{f(\boldsymbol{\lambda}_{\mathbf{k}})}{f_{\theta^*}(\boldsymbol{\lambda}_{\mathbf{k}})} \to \frac{f(\boldsymbol{\lambda})}{f_{\theta^*}(\boldsymbol{\lambda})} \tag{73}$$

The result is concluded from (72) and (73).

□

8.3 Proof of Theorem 3.

From now on, note that $N = n_1 n_2$ denotes the number of data points whereas n denotes the number of Fourier frequencies and $\|\cdot\|$ is the L^2-norm. We will drop the subindex 0 and denote by θ the true parameter under the null hypothesis. Define

$$q_1(m(\boldsymbol{\lambda}_{\mathbf{k}}), Y_{\mathbf{k}}) = +e^{Y_{\mathbf{k}} - m(\boldsymbol{\lambda}_{\mathbf{k}})} - 1, \quad q_2(m(\boldsymbol{\lambda}_{\mathbf{k}}), Y_{\mathbf{k}}) = -e^{Y_{\mathbf{k}} - m(\boldsymbol{\lambda}_{\mathbf{k}})},$$

where m can be replaced by m_θ or by $\hat{m}_{LK}$. Assume m_θ is the log-likelihood under the null hypothesis and denote

$$\varepsilon_{\mathbf{k}} = q_1(m_\theta(\boldsymbol{\lambda}_{\mathbf{k}}), Y_{\mathbf{k}}^{**}) = e^{Y_{\mathbf{k}}^{**} - m_\theta(\boldsymbol{\lambda}_{\mathbf{k}})} - 1, \quad q_2^{\mathbf{k}} = q_2(m_\theta(\boldsymbol{\lambda}_{\mathbf{k}}), Y_{\mathbf{k}}^{**}) = -e^{Y_{\mathbf{k}}^{**} - m_\theta(\boldsymbol{\lambda}_{\mathbf{k}})},$$

where $Y_{\mathbf{k}}^{**}$ is given by (22). Define, also

$$\Gamma(\boldsymbol{\lambda}_{\mathbf{k}}) = -E\left(q_2(m_\theta(\boldsymbol{\lambda}), Y_{\mathbf{k}}^{**})\right) \frac{1}{4\pi^2}.$$

Some other constants and vectors that will appear in our computations are

$$\boldsymbol{\beta}^T = \boldsymbol{\beta}(\boldsymbol{\lambda})^T = \left(m_\theta(\boldsymbol{\lambda}), |H|^{1/2} \nabla^T m_\theta(\boldsymbol{\lambda}) \right) \in \mathbb{R}^3,$$

where $\nabla m_\theta(\boldsymbol{\lambda})$ denotes the gradient vector $\nabla m_\theta(\boldsymbol{\lambda}) = \left(\frac{\partial m_\theta}{\partial x}(\boldsymbol{\lambda}), \frac{\partial m_\theta}{\partial y}(\boldsymbol{\lambda}) \right)$ and $\frac{\partial}{\partial x}$, $\frac{\partial}{\partial y}$ denote the derivatives with respect to the first and second components. Besides, define

$$\boldsymbol{\beta}_2^T = \boldsymbol{\beta}_2(\boldsymbol{\lambda})^T = \sqrt{n|H|^{1/2}} (a - m(\boldsymbol{\lambda}), |H|^{1/2}(\mathbf{b} - \nabla m(\boldsymbol{\lambda}))^T) \in \mathbb{R}^3,$$

where $(a, \mathbf{b})$ are the parameters in the non-parametric model (24),

$$W_{\mathbf{k}} = W_{\mathbf{k}}(\boldsymbol{\lambda}) = (1, |H|^{-1/2}(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{k}})) \in \mathbb{R}^3,$$

$$r_N^2 = \frac{1}{N|H|^{1/2}}, \quad \text{and } \bar{m}_{\mathbf{k}} = \bar{m}_{\mathbf{k}}(\boldsymbol{\lambda}) = m_\theta(\boldsymbol{\lambda}) + \nabla^T m_\theta(\boldsymbol{\lambda})(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{k}}).$$

Besides,

$$\Phi_{n,j} = \sup_{\boldsymbol{\lambda} \in \Pi^2, \|\alpha\| = c_1 r_N} \left| q_2(\boldsymbol{\beta}_*^T W_{\mathbf{k}} + \alpha^T W_{\mathbf{k}}, Y_{\mathbf{k}}) |H|^{(j-1)/2} \|(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{k}})\|^{(j-1)} K_H(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{k}}) \right|,$$

where $\boldsymbol{\beta}_*$ denotes $\boldsymbol{\beta}$ or $\boldsymbol{\beta}_2$ and assume that

$$E(\Phi_{n,j})^\zeta = \mathcal{O}(1), j = 1, 2, 3.$$

Lemma 9. *The Generalized Likelihood Ratio Test statistic*

$$T_{LK} = \sum_{\mathbf{k}} \left[e^{Y_{\mathbf{k}} - m_\theta(\boldsymbol{\lambda}_{\mathbf{k}})} + m_\theta(\boldsymbol{\lambda}_{\mathbf{k}}) - e^{Y_{\mathbf{k}} - \hat{m}_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})} - \hat{m}_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}) \right] \tag{74}$$

admits the following decomposition

$$T_{LK} = T^*_{LK} + B_1 + B_2 - B_3 \tag{75}$$

where T^*_{LK} is the same as T_{LK} but replacing $Y_{\mathbf{k}}$ by $Y^{**}_{\mathbf{k}}$ and $\hat{m}_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})$ by $\hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})$ and,

$$B_1 = \sum_{\mathbf{k}} \left\{ 1 - e^{Y_{\mathbf{k}} - \hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})} \right\} \left(\hat{m}_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}) - \hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}) \right),$$

$$B_2 = \sum_{\mathbf{k}} e^{Y_{\mathbf{k}} - \hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})} \left(\hat{m}_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}) - \hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}) \right)^2,$$

$$B_3 = \sum_{\mathbf{k}} \frac{R_N(\boldsymbol{\lambda}_{\mathbf{k}})}{f_\theta(\boldsymbol{\lambda}_{\mathbf{k}})} \left\{ e^{m_\theta(\boldsymbol{\lambda}_{\mathbf{k}}) - \hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})} - 1 \right\}.$$

Proof.

$$\begin{aligned}
&T_{LK} - T^*_{LK} = \\
&\sum_{\mathbf{k}} \Big[\hat{m}^*_{LK}(\boldsymbol{\lambda}_k) - \hat{m}_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}) + e^{Y_{\mathbf{k}} - m_\theta(\boldsymbol{\lambda}_{\mathbf{k}})} - e^{Y^{**}_{\mathbf{k}} - m_\theta(\boldsymbol{\lambda}_{\mathbf{k}})} + e^{Y^{**}_{\mathbf{k}} - \hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})} - e^{Y_{\mathbf{k}} - \hat{m}_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})} \Big] = \\
&\sum_{\mathbf{k}} \Big[\hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}) - \hat{m}_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}) + e^{Y_{\mathbf{k}} - m_\theta(\boldsymbol{\lambda}_{\mathbf{k}})} - e^{Y^{**}_{\mathbf{k}} - m_\theta(\boldsymbol{\lambda}_{\mathbf{k}})} + e^{Y_{\mathbf{k}} - \hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})} \\
&- e^{Y_{\mathbf{k}} - \hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})} + e^{Y^{**}_{\mathbf{k}} - \hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})} - e^{Y_{\mathbf{k}} - \hat{m}_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})} \Big]
\end{aligned}$$

By Taylor's expansion of $h_{\mathbf{k}}(x) = e^{Y_{\mathbf{k}} - x}$ evaluated at $\hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})$, and doing the expansion around $\hat{m}_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})$:

$$\begin{aligned}
e^{Y_{\mathbf{k}} - \hat{m}_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})} &= e^{Y_{\mathbf{k}} - \hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})} - (\hat{m}_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}) - \hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})) e^{Y_{\mathbf{k}} - \hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})} \\
&+ \frac{1}{2} (\hat{m}_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}) - \hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}))^2 e^{Y_{\mathbf{k}} - \hat{m}_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})} - \frac{1}{3!} (\hat{m}_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}) - \hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}))^3 e^{Y_{\mathbf{k}} - z_{\mathbf{k}}}
\end{aligned}$$

where $z_{\mathbf{k}}$ is such that $|\hat{m}_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}) - \hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})| \geq |z_{\mathbf{k}} - \hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})|$. The last addend is given in Lagrange's remainder form, and it can be bounded by:

$$\begin{aligned}
&(\hat{m}_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}) - \hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}))^3 e^{Y_{\mathbf{k}} - z_{\mathbf{k}}} \\
&= (\hat{m}_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}) - \hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}))^3 \left(1 + (\hat{m}_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}) - z_{\mathbf{k}}) + \frac{1}{2} (\hat{m}_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}) - z_{\mathbf{k}})^2 + \dots \right) \\
&= \mathcal{O}_{\mathbb{P}}(N^{-3/2} \log^3 N),
\end{aligned}$$

applying Lemma 2:

$$|Y_{\mathbf{k}} - z_{\mathbf{k}}| \leq |\hat{m}_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}) - z_{\mathbf{k}}| \leq \mathcal{O}_{\mathbb{P}}(N^{-1/2} \log N).$$

Then, $T_{LK} - T^*_{LK}$ can be written as:

$$\sum_{\mathbf{k}} \Big[e^{Y_{\mathbf{k}} - m_\theta(\boldsymbol{\lambda}_{\mathbf{k}})} - e^{Y^{**}_{\mathbf{k}} - m_\theta(\boldsymbol{\lambda}_{\mathbf{k}})} + e^{Y^{**}_{\mathbf{k}} - \hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})} - e^{Y_{\mathbf{k}} - \hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})} +$$
$$\hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}) - \hat{m}_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}) + e^{Y_{\mathbf{k}} - \hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})}(\hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}) - \hat{m}_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})$$
$$+\frac{1}{2} e^{Y_{\mathbf{k}} - \hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})}(\hat{m}_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}) - \hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}))^2 + \mathcal{O}_{\mathbb{P}}(N^{-3/2} \log^3 N) \Big]$$
$$= \sum_{\mathbf{k}} \Big[e^{Y_{\mathbf{k}} - m_\theta(\boldsymbol{\lambda}_{\mathbf{k}})} - e^{Y^{**}_{\mathbf{k}} - m_\theta(\boldsymbol{\lambda}_{\mathbf{k}})} + e^{Y^{**}_{\mathbf{k}} - \hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})} - e^{Y_{\mathbf{k}} - \hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})}$$
$$+(1 - e^{Y_{\mathbf{k}} - \hat{m}_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})})(\hat{m}_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}) - \hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}))$$
$$+\frac{1}{2} e^{Y_{\mathbf{k}} - \hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})}(\hat{m}_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}) - \hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}))^2 + \mathcal{O}_{\mathbb{P}}(N^{-3/2} \log^3 N) \Big] =$$
$$\sum_{\mathbf{k}} \Big[(1 - e^{Y_{\mathbf{k}} - \hat{m}_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})})(\hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}) - \hat{m}_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})) \Big] +$$
$$\sum_{\mathbf{k}} \left[\frac{1}{2} e^{Y_{\mathbf{k}} - \hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})}(\hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}) - \hat{m}_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}))^2 \right] +$$
$$\sum_{\mathbf{k}} \Big[e^{Y_{\mathbf{k}} - m_\theta(\boldsymbol{\lambda}_{\mathbf{k}})} - e^{Y^{**}_{\mathbf{k}} - m_\theta(\boldsymbol{\lambda}_{\mathbf{k}})} + e^{Y^{**}_{\mathbf{k}} - \hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})} - e^{Y_{\mathbf{k}} - \hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})} \Big]$$
$$+\mathcal{O}_{\mathbb{P}}(N^{-3/2} \log^3 N)$$

The first two addends correspond to B_1 and B_2. To obtain the third part, B_3, we should recall the following relations:

$$I(\boldsymbol{\lambda}_{\mathbf{k}}) = f_\theta(\boldsymbol{\lambda}_{\mathbf{k}}) V_{\mathbf{k}} + R_N(\boldsymbol{\lambda}_{\mathbf{k}}), \quad Y_{\mathbf{k}} = m_\theta(\boldsymbol{\lambda}_{\mathbf{k}}) + z_{\mathbf{k}} + r_{\mathbf{k}} \tag{76}$$

$$e^{Y_{\mathbf{k}}} = f_\theta(\boldsymbol{\lambda}_{\mathbf{k}}) V_{\mathbf{k}} e^{r_{\mathbf{k}}}, \quad e^{Y^{**}_{\mathbf{k}}} = f_\theta(\boldsymbol{\lambda}_{\mathbf{k}}) V_{\mathbf{k}}, \quad e^{Y_{\mathbf{k}} - Y^{**}_{\mathbf{k}}} = f_\theta(\boldsymbol{\lambda}_{\mathbf{k}}) V_{\mathbf{k}} (e^{r_{\mathbf{k}}} - 1). \tag{77}$$

And recall also that $e^{m_\theta(\boldsymbol{\lambda})} = f_\theta(\boldsymbol{\lambda})$. In order to derive the final expression for $B3$, we must consider the following Taylor's expansion:

$$e^{Y_{\mathbf{k}} - m_\theta(\boldsymbol{\lambda}_{\mathbf{k}})} = e^{Y^{**}_{\mathbf{k}} - m_\theta(\boldsymbol{\lambda}_{\mathbf{k}})} + (Y_{\mathbf{k}} - Y^{**}_{\mathbf{k}}) e^{Y^{**}_{\mathbf{k}} - m_\theta(\boldsymbol{\lambda}_{\mathbf{k}})} + \frac{1}{2}(Y_{\mathbf{k}} - Y^{**}_{\mathbf{k}})^2 e^{c_{\mathbf{k}} - m_\theta(\boldsymbol{\lambda}_{\mathbf{k}})}, \tag{78}$$

where $c_{\mathbf{k}}$ is such that $|Y_{\mathbf{k}} - Y^{**}_{\mathbf{k}}| \geq |c_{\mathbf{k}} - Y^{**}_{\mathbf{k}}|$. Besides, the difference between $Y_{\mathbf{k}}$ and $Y^{**}_{\mathbf{k}}$ is bounded by:

$$|Y_{\mathbf{k}} - Y^{**}_{\mathbf{k}}| = |r_{\mathbf{k}} - C_0|, \quad \text{where } C_0 \text{ is the Euler constant.}$$

From the expression for $r_{\mathbf{k}}$:

$$r_{\mathbf{k}} = \log \left(1 + \frac{R_N(\boldsymbol{\lambda}_{\mathbf{k}})}{f(\boldsymbol{\lambda}_{\mathbf{k}}) V_{\mathbf{k}}} \right),$$

and taking into account that $R_N(\boldsymbol{\lambda}_{\mathbf{k}})$ is uniformly bounded by:

$$\max_{\mathbf{k}} |R_N(\boldsymbol{\lambda}_{\mathbf{k}})| = \mathcal{O}_{\mathbb{P}}(N^{-1/2} \log N), \tag{79}$$

the remainder in Taylor's expansion can be bounded by:

$$|Y_{\mathbf{k}} - Y^{**}_{\mathbf{k}}| = \mathcal{O}_{\mathbb{P}}(\log N^{-1/2} \log \log N),$$

and this bound is uniform in $\boldsymbol{\lambda}_{\mathbf{k}}$. Using similar arguments, we have:

$$e^{Y_{\mathbf{k}}-\hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})} = e^{Y^{**}_{\mathbf{k}}-\hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})} + (Y_{\mathbf{k}} - Y^{**}_{\mathbf{k}})e^{Y_{\mathbf{k}}-\hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})} + \mathcal{O}_{\mathbb{P}}(\log N^{-1/2} + \log\log N). \quad (80)$$

Applying Taylor's expansion in $B3$ (for two groups of addends) we have

$$B3 = \sum_{\mathbf{k}} \Big(e^{Y_{\mathbf{k}}-m_\theta(\boldsymbol{\lambda}_{\mathbf{k}})}(Y_{\mathbf{k}} - Y^{**}_{\mathbf{k}}) - e^{Y_{\mathbf{k}}-\hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})}(Y_{\mathbf{k}} - Y^{**}_{\mathbf{k}}) + \mathcal{O}_{\mathbb{P}}(\log N^{-1/2} + \log\log N) \Big) =$$

$$\sum_{\mathbf{k}} e^{Y_{\mathbf{k}}}(Y_{\mathbf{k}} - Y^{**}_{\mathbf{k}})(e^{-m_\theta(\boldsymbol{\lambda}_{\mathbf{k}})} - e^{-\hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})}) + \mathcal{O}_{\mathbb{P}}(N\log N^{-1/2}\log\log N) =$$

$$\sum_{\mathbf{k}} e^{Y_{\mathbf{k}}}(Y_{\mathbf{k}} - Y^{**}_{\mathbf{k}})\frac{1}{f_\theta(\boldsymbol{\lambda}_{\mathbf{k}})}\Big(1 - e^{m_\theta(\boldsymbol{\lambda}_{\mathbf{k}})-\hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})}\Big) + \mathcal{O}_{\mathbb{P}}(N\log N^{-1/2} + N\log\log N)$$

And, with another Taylor's expansion on $e^{Y_{\mathbf{k}}}$ around $e^{Y^{**}_{\mathbf{k}}}$:

$$B3 = \sum_{\mathbf{k}}(e^{Y_{\mathbf{k}}} - e^{Y^{**}_{\mathbf{k}}})\frac{1}{f_\theta(\boldsymbol{\lambda}_{\mathbf{k}})}(1 - e^{m_\theta(\boldsymbol{\lambda}_{\mathbf{k}})-\hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})}) + \mathcal{O}_{\mathbb{P}}(N\log N^{-1/2} + N\log\log N) =$$

$$\sum_{\mathbf{k}} r_N(\boldsymbol{\lambda}_{\mathbf{k}})\frac{1}{f_\theta(\boldsymbol{\lambda}_{\mathbf{k}})}(1 - e^{m_\theta(\boldsymbol{\lambda}_{\mathbf{k}})-\hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})}) + \mathcal{O}_{\mathbb{P}}(N\log N^{-1/2} + N\log\log N).$$

□

We prove now that T^*_{LK} follows an asymptotically normal distribution.

Proof. **Proof of Theorem 3.**
The regression model (22) under the null hypothesis

$$Y^{**}_{\mathbf{k}} = m_\theta(\boldsymbol{\lambda}_{\mathbf{k}}) + z_k \quad (81)$$

can be seen regresion model with non-Gaussian error variables with density (11). The asymptotic distribution of T^*_{LK} is obtained as a particular case of Theorem 10 in (Fan *et al.* (2001)), generalized to dimension $d = 2$. The loglikelihood associated with model (81):

$$f(Y^{**}_{\mathbf{k}}, m_\theta(\boldsymbol{\lambda}_{\mathbf{k}})) = Y^{**}_{\mathbf{k}} - m_\theta(\boldsymbol{\lambda}_{\mathbf{k}}) - e^{Y^{**}_{\mathbf{k}}-m_\theta(\boldsymbol{\lambda}_{\mathbf{k}})} \quad (82)$$

and the generalized likelihood ratio test statistic is given by

$$T^*_{LK} = \sum_{\mathbf{k}} \left\{ e^{Y^{**}_{\mathbf{k}}-m_\theta(\boldsymbol{\lambda}_{\mathbf{k}})} + m_\theta(\boldsymbol{\lambda}_{\mathbf{k}}) - e^{Y^{**}_{\mathbf{k}}-\hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})} - \hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}) \right\}.$$

Using Taylor's expansion of the loglikelihood function, with the notation introduced above:

$$T^*_{LK} = \sum_{\mathbf{k}} \left\{ \varepsilon_{\mathbf{k}}(\hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}) - m_\theta(\boldsymbol{\lambda}_{\mathbf{k}})) + q_2^{\mathbf{k}}\frac{1}{2}(\hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}) - m_\theta(\boldsymbol{\lambda}_{\mathbf{k}}))^2 + \mathcal{O}_{\mathbb{P}}(N^{-3/2}\log^3 N) \right\} \quad (83)$$

For the sake of simplicity, we will drop the residual part. Now, using the asymptotic representation for the nonparametric estimator given in Lemma 11, and the expression for $H_N(\boldsymbol{\lambda})$

in (84), the non-negligible part of (83) can be written as:

$$
\begin{aligned}
T^*_{LK} &= \sum_{\mathbf{k}} \Bigg\{ \varepsilon_{\mathbf{k}} r_N^2 \Gamma(\boldsymbol{\lambda}_{\mathbf{k}})^{-1} \sum_{\mathbf{i}} \varepsilon_{\mathbf{i}} K(H^{-1/2}(\boldsymbol{\lambda}_{\mathbf{i}} - \boldsymbol{\lambda}_{\mathbf{k}}))(1 + o_{\mathbb{P}}(1)) + \varepsilon_{\mathbf{k}} H_N(\boldsymbol{\lambda}_{\mathbf{k}}) \\
&+ \frac{1}{2} q_2^{\mathbf{k}} \left[r_N^2 \Gamma(\boldsymbol{\lambda}_{\mathbf{k}})^{-1} \sum_{\mathbf{i}} K(H^{-1/2}(\boldsymbol{\lambda}_{\mathbf{i}} - \boldsymbol{\lambda}_{\mathbf{k}}))(1 + o_{\mathbb{P}}(1)) + H_N(\boldsymbol{\lambda}_{\mathbf{k}}) \right]^2 \Bigg\} = \\
& r_N^2 \sum_{\mathbf{k}} \sum_{i} \varepsilon_{\mathbf{k}} \varepsilon_i \Gamma(\boldsymbol{\lambda}_{\mathbf{k}})^{-1} K(H^{-1/2}(\boldsymbol{\lambda}_{\mathbf{i}} - \boldsymbol{\lambda}_{\mathbf{k}}))(1 + o_{\mathbb{P}}(1)) + \sum_{\mathbf{k}} \varepsilon_{\mathbf{k}} H_N(\boldsymbol{\lambda}_{\mathbf{k}}) \\
&+ \frac{1}{2} \sum_{\mathbf{k}} q_2^{\mathbf{k}} \Bigg[r_N^4 \Gamma(\boldsymbol{\lambda}_{\mathbf{k}})^{-2} \sum_{\mathbf{i}} \sum_{\mathbf{j}} \varepsilon_{\mathbf{i}} \varepsilon_{\mathbf{j}} K(H^{-1/2}(\boldsymbol{\lambda}_{\mathbf{i}} - \boldsymbol{\lambda}_{\mathbf{k}})) K(H^{-1/2}(\boldsymbol{\lambda}_{\mathbf{j}} - \boldsymbol{\lambda}_{\mathbf{k}}))(1 + o_{\mathbb{P}}(1))^2 \\
&+ H_N^2(\boldsymbol{\lambda}_{\mathbf{k}}) + 2 r_N^2 \Gamma(\boldsymbol{\lambda}_{\mathbf{k}})^{-1} \sum_{\mathbf{i}} \varepsilon_{\mathbf{i}} K(H^{-1/2}(\boldsymbol{\lambda}_{\mathbf{i}} - \boldsymbol{\lambda}_{\mathbf{k}}))(1 + o_{\mathbb{P}}(1)) H_N(\boldsymbol{\lambda}_{\mathbf{k}}) \Bigg] = \\
& r_N^2 \sum_{\mathbf{k}} \sum_{\mathbf{i}} \varepsilon_{\mathbf{k}} \varepsilon_i \Gamma(\boldsymbol{\lambda}_{\mathbf{k}})^{-1} K(H^{-1/2}(\boldsymbol{\lambda}_{\mathbf{i}} - \boldsymbol{\lambda}_{\mathbf{k}}))(1 + o_{\mathbb{P}}(1)) + \sum_{\mathbf{k}} \varepsilon_{\mathbf{k}} H_N(\boldsymbol{\lambda}_{\mathbf{k}}) \\
&+ \frac{r_N^4}{2} \sum_{\mathbf{k}} q_2^{\mathbf{k}} \Gamma(\boldsymbol{\lambda}_{\mathbf{k}})^{-2} \sum_{\mathbf{i}} \sum_{\mathbf{j}} \varepsilon_{\mathbf{i}} \varepsilon_{\mathbf{j}} K(H^{-1/2}(\boldsymbol{\lambda}_{\mathbf{i}} - \boldsymbol{\lambda}_{\mathbf{k}})) K(H^{-1/2}(\boldsymbol{\lambda}_{\mathbf{j}} - \boldsymbol{\lambda}_{\mathbf{k}}))(1 + o_{\mathbb{P}}(1))^2 \\
&+ \frac{1}{2} \sum_{\mathbf{k}} q_2^{\mathbf{k}} H_N^2(\boldsymbol{\lambda}_{\mathbf{k}}) - r_N^2 \sum_{\mathbf{k}} q_2^{\mathbf{k}} \Gamma(\boldsymbol{\lambda}_{\mathbf{k}})^{-1} \sum_{\mathbf{i}} \varepsilon_{\mathbf{i}} K(H^{-1/2}(\boldsymbol{\lambda}_{\mathbf{i}} - \boldsymbol{\lambda}_{\mathbf{k}}))(1 + o_{\mathbb{P}}(1)) H_N(\boldsymbol{\lambda}_{\mathbf{k}}) = \\
& S_{1N} + S_{2N} + R_{1N} + R_{2N} + R_{3N}
\end{aligned}
$$

where

$$
H_N(\boldsymbol{\lambda}) = r_N^2 \Gamma(\boldsymbol{\lambda})^{-1} \sum_{\mathbf{k}} \left[q_1(\beta(\boldsymbol{\lambda})^T W_{\mathbf{k}}, Y^{**}_{\mathbf{k}}) - \varepsilon_{\mathbf{k}} \right] K(H^{-1/2}(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{k}}))(1 + o_{\mathbb{P}}(1)) \tag{84}
$$

and the residual terms are given by:

$$
\begin{aligned}
R_{1N} &= \sum_{\mathbf{k}} \varepsilon_{\mathbf{k}} H_N(\boldsymbol{\lambda}_{\mathbf{k}}), \\
R_{2N} &= \frac{-1}{2} \sum_{\mathbf{k}} q_2^{\mathbf{k}} H_N^2(\boldsymbol{\lambda}_{\mathbf{k}}), \\
R_{3N} &= -r_N^2 \sum_{\mathbf{k}} q_2^{\mathbf{k}} \Gamma(\boldsymbol{\lambda}_{\mathbf{k}})^{-1} \sum_{\mathbf{i}} \varepsilon_{\mathbf{i}} K(H^{-1/2}(\boldsymbol{\lambda}_{\mathbf{i}} - \boldsymbol{\lambda}_{\mathbf{k}})) H_N(\boldsymbol{\lambda}_{\mathbf{k}}).
\end{aligned}
$$

The leading terms are:

$$
S_{1N} = r_N^2 \sum_{\mathbf{k}} \sum_{\mathbf{i}} \varepsilon_{\mathbf{k}} \varepsilon_{\mathbf{i}} \Gamma(\boldsymbol{\lambda}_{\mathbf{k}})^{-1} K(H^{-1/2}(\boldsymbol{\lambda}_{\mathbf{i}} - \boldsymbol{\lambda}_{\mathbf{k}})) \tag{85}
$$

and

$$
S_{2N} = \frac{r_N^4}{2} \sum_{\mathbf{k}} q_2^{\mathbf{k}} \Gamma(\boldsymbol{\lambda}_{\mathbf{k}})^{-2} \sum_{\mathbf{i}} \sum_{\mathbf{j}} \varepsilon_{\mathbf{i}} \varepsilon_{\mathbf{j}} K(H^{-1/2}(\boldsymbol{\lambda}_{\mathbf{i}} - \boldsymbol{\lambda}_{\mathbf{k}})) K(H^{-1/2}(\boldsymbol{\lambda}_{\mathbf{j}} - \boldsymbol{\lambda}_{\mathbf{k}})). \tag{86}
$$

S_{1N} can be decomposed as:

$$\begin{aligned}
&S_{1N} \\
&= r_N^2 \sum_{\mathbf{k}} \varepsilon_{\mathbf{k}}^2 \Gamma(\boldsymbol{\lambda}_{\mathbf{k}})^{-1} K(H^{-1/2}(\boldsymbol{\lambda}_{\mathbf{k}} - \boldsymbol{\lambda}_{\mathbf{k}})) + r_N^2 \sum_{\mathbf{k}\neq\mathbf{i}} \varepsilon_{\mathbf{k}} \varepsilon_i \Gamma(\boldsymbol{\lambda}_{\mathbf{k}})^{-1} K(H^{-1/2}(\boldsymbol{\lambda}_{\mathbf{i}} - \boldsymbol{\lambda}_{\mathbf{k}})) \\
&= r_N^2 \sum_{\mathbf{k}} \varepsilon_{\mathbf{k}}^2 \Gamma(\boldsymbol{\lambda}_{\mathbf{k}})^{-1} K(H^{-1/2}\mathbf{0}) + r_N^2 \sum_{\mathbf{k}\neq\mathbf{i}} \varepsilon_{\mathbf{i}} \varepsilon_{\mathbf{k}} \Gamma(\boldsymbol{\lambda}_{\mathbf{k}})^{-1} K(H^{-1/2}(\boldsymbol{\lambda}_{\mathbf{i}} - \boldsymbol{\lambda}_{\mathbf{k}})) \\
&= S_{1N}^1 + S_{1N}^2.
\end{aligned}$$

For the first addend,

$$\begin{aligned}
S_{1N}^1 &= r_N^2 \sum_{\mathbf{k}} \varepsilon_{\mathbf{k}}^2 \Gamma(\boldsymbol{\lambda}_{\mathbf{k}})^{-1} K(H^{-1/2}\mathbf{0}) = \\
&= \frac{1}{N|H|^{1/2}} \sum_{\mathbf{k}} 4\pi^2 E^{-1}(q_2(m_\theta(\boldsymbol{\lambda}_{\mathbf{k}}), Y_{\mathbf{k}}^{**})) K(H^{-1/2}\mathbf{0}) \to_{\mathbb{P}} \frac{4\pi^2}{|H|^{1/2}} K(\mathbf{0}).
\end{aligned}$$

Therefore:

$$S_{1N} \approx \frac{4\pi^2}{|H|^{1/2}} K(\mathbf{0}) + S_{1N}^2,$$

with

$$S_{1N}^2 = r_N^2 \sum_{\mathbf{k}\neq\mathbf{i}} \varepsilon_{\mathbf{k}} \varepsilon_{\mathbf{i}} \Gamma(\boldsymbol{\lambda}_{\mathbf{k}})^{-1} K(H^{-1/2}(\boldsymbol{\lambda}_{\mathbf{i}} - \boldsymbol{\lambda}_{\mathbf{k}})).$$

Note that:

$$E(\varepsilon_{\mathbf{k}}^2 / \boldsymbol{\lambda} = \boldsymbol{\lambda}_{\mathbf{k}}) = -E(q_2(m_\theta(\boldsymbol{\lambda}_{\mathbf{k}}), Y_{\mathbf{k}}^{**}),$$

and consider the following decomposition for (86):

$$\begin{aligned}
S_{2N} &= \frac{r_N^4}{2} \sum_{\mathbf{k}} q_2^{\mathbf{k}} \Gamma(\boldsymbol{\lambda}_{\mathbf{k}})^{-2} \sum_{\mathbf{i}=\mathbf{j}} \varepsilon_{\mathbf{i}} \varepsilon_{\mathbf{j}} K(H^{-1/2}(\boldsymbol{\lambda}_{\mathbf{i}} - \boldsymbol{\lambda}_{\mathbf{k}})) K(H^{-1/2}(\boldsymbol{\lambda}_{\mathbf{j}} - \boldsymbol{\lambda}_{\mathbf{k}})) \\
&+ \frac{r_N^4}{2} \sum_{\mathbf{k}} q_2^{\mathbf{k}} \Gamma(\boldsymbol{\lambda}_{\mathbf{k}})^{-2} \sum_{\mathbf{i}\neq\mathbf{j}} \varepsilon_{\mathbf{i}} \varepsilon_{\mathbf{j}} K(H^{-1/2}(\boldsymbol{\lambda}_{\mathbf{i}} - \boldsymbol{\lambda}_{\mathbf{k}})) K(H^{-1/2}(\boldsymbol{\lambda}_{\mathbf{j}} - \boldsymbol{\lambda}_{\mathbf{k}})) = \\
&S_{2N}^1 + S_{2N}^2.
\end{aligned}$$

The first part S_{2N}^1 converges in probability to:

$$S_{2N}^1 = \frac{r_N^4}{2} \sum_{\mathbf{k}} q_2^{\mathbf{k}} \Gamma(\boldsymbol{\lambda}_{\mathbf{k}})^{-1} \sum_{\mathbf{i}} \varepsilon_{\mathbf{i}}^2 K^2(H^{-1/2}(\boldsymbol{\lambda}_{\mathbf{i}} - \boldsymbol{\lambda}_{\mathbf{k}})) \to_{\mathbb{P}} \frac{4\pi^2}{2|H|^{1/2}} \int K^2(\mathbf{u}) d\mathbf{u}.$$

The addend S_{2N}^2 can be decomposed in two parts:

$$\begin{aligned}
S_{2N}^2 &= \frac{r_N^4}{2} \sum_{\mathbf{k}} q_2^{\mathbf{k}} \Gamma(\boldsymbol{\lambda}_{\mathbf{k}})^{-2} \sum_{\mathbf{i}\neq\mathbf{j}} \varepsilon_{\mathbf{i}} \varepsilon_{\mathbf{j}} K(H^{-1/2}(\boldsymbol{\lambda}_{\mathbf{i}} - \boldsymbol{\lambda}_{\mathbf{k}})) K(H^{-1/2}(\boldsymbol{\lambda}_{\mathbf{j}} - \boldsymbol{\lambda}_{\mathbf{k}})) \\
&= \frac{r_N^4}{2} K(\mathbf{0}) \sum_{\mathbf{k}} \sum_{\mathbf{j}} \varepsilon_{\mathbf{k}} \varepsilon_{\mathbf{j}} q_2^{\mathbf{k}} \Gamma(\boldsymbol{\lambda}_{\mathbf{k}})^{-2} K(H^{-1/2}(\boldsymbol{\lambda}_{\mathbf{j}} - \boldsymbol{\lambda}_{\mathbf{k}})) + \\
&\frac{r_N^4}{2} \sum_{\mathbf{i},\mathbf{j}} \varepsilon_{\mathbf{i}} \varepsilon_{\mathbf{j}} \sum_{\mathbf{k}\neq\mathbf{i},\mathbf{k}\neq\mathbf{j}} q_2^{\mathbf{k}} \Gamma(\boldsymbol{\lambda}_{\mathbf{k}})^{-2} K(H^{-1/2}(\boldsymbol{\lambda}_{\mathbf{i}} - \boldsymbol{\lambda}_{\mathbf{k}})) K(H^{-1/2}(\boldsymbol{\lambda}_{\mathbf{j}} - \boldsymbol{\lambda}_{\mathbf{k}})) \\
&= S_{2N}^{21} + S_{2N}^{22}.
\end{aligned}$$

The variance of the first addend can be bounded by:

$$Var(S_{2N}^{21}) = Var\left(\frac{r_N^4}{2}K(\mathbf{0})\sum_{\mathbf{k}}\sum_{\mathbf{j}}\varepsilon_{\mathbf{k}}\varepsilon_{\mathbf{j}}q_2^{\mathbf{k}}\Gamma(\boldsymbol{\lambda}_{\mathbf{k}})^{-2}K(H^{-1/2}(\boldsymbol{\lambda}_{\mathbf{j}}-\boldsymbol{\lambda}_{\mathbf{k}}))\right)$$
$$= \mathcal{O}(N^{-2}|H|^{-3/2}) = o(|H|^{-1/2}),$$

therefore, this addend is

$$S_{2N}^{21} = o_{\mathbb{P}}(|H|^{-1/4}).$$

Then, in the expression of T_{LK}^* we have:

$$T_{LK}^* \approx \mu_H + R_{1N} + R_{2N} + R_{3N} + S_{1N}^2 + S_{2N}^{22} = \mu_H + R_{1N} + R_{2N} + R_{3N} + \frac{1}{2}W_N|H|^{-1/4},$$

where

$$W_N = \frac{|H|^{1/4}}{N}\sum_{\mathbf{i}\neq\mathbf{j}}\varepsilon_{\mathbf{i}}\varepsilon_{\mathbf{j}}\Gamma(\boldsymbol{\lambda}_{\mathbf{j}})^{-1}\left(2K_H(\boldsymbol{\lambda}_{\mathbf{i}}-\boldsymbol{\lambda}_{\mathbf{j}}) - K_H * K_H(\boldsymbol{\lambda}_{\mathbf{i}}-\boldsymbol{\lambda}_{\mathbf{j}})\right).$$

Besides, if we define, for $\mathbf{i}\neq\mathbf{j}$:

$$b(\mathbf{i},\mathbf{j}) = \frac{|H|^{1/4}}{N}(2K_H(\boldsymbol{\lambda}_{\mathbf{i}}-\boldsymbol{\lambda}_{\mathbf{j}}) - K_H * K_H(\boldsymbol{\lambda}_{\mathbf{i}}-\boldsymbol{\lambda}_{\mathbf{j}}))\Gamma(\boldsymbol{\lambda}_{\mathbf{j}})^{-1},$$

and $b(\mathbf{i},\mathbf{i}) = 0$. Then, W_N can be written as:

$$W_N = \sum_{\mathbf{i}}\sum_{\mathbf{j}} b(\mathbf{i},\mathbf{j})\varepsilon_{\mathbf{i}}\varepsilon_{\mathbf{j}}.$$

In order to prove the asymptotic normal distribution of W_N, we will apply Proposition 3.2 in de Jong (1987). For that purpose, we must write W_N as a quadratic form of indpendent random variables, namely $W_N = \sum_{i<j} c_{i,j}\varepsilon_i\varepsilon_j$, where i and j are one-dimensional indexes.

As it is done in the proof of Theorem 1, define a new subindex for the Fourier frequencies $\boldsymbol{\lambda}_{\mathbf{k}}$, with $\mathbf{k} = (k_1,k_2)$ and $k_l = 0,\pm1,\ldots,\pm m_l$, for $l = 1,2$. Consider $\boldsymbol{\lambda}_{\mathbf{k}} = \boldsymbol{\lambda}_{\mathbf{k}'}$ where $\mathbf{k}' = (k_1',k_2')$, with $k_l' = 1,\ldots,m_l' = 2m_l+1$, in such a way that $k_l' = k_l + m_l + 1$ for $l = 1,2$. Let $M = m_1'\times m_2'$. The new coefficients, with one dimensional indexes, are given by the following matrix:

$$A = (a_{ij}), \quad A \in \mathcal{M}_{M\times M},$$

and each entry of this matrix is defined by $a_{ij} = b_{\mathbf{ij}}$ and $a_{ii} = 0$, where the bidimensional indexes $\mathbf{i}$ and $\mathbf{j}$ are given by:

$$\mathbf{i} = (i_1,i_2) = (k,k_0), \quad \text{if } (k-1)m_2' \le i \le km_2' \quad \text{and } i = (k-1)m_2' + k_0, \tag{87}$$

$$\mathbf{j} = (j_1,j_2) = (l,l_0), \quad \text{if } (l-1)m_2' \le j \le lm_2' \quad \text{and } j = (l-1)m_2' + l_0. \tag{88}$$

With this definitions, W_N can be written as a quadratic form with one-dimensional indexes:

$$W_N = \sum_{i,j} a_{i,j}\varepsilon_i\varepsilon_j.$$

For a_{ij} we have that:

$$a_{ij} = \frac{|H|^{1/4}}{N}2K_H(\boldsymbol{\lambda}_i-\boldsymbol{\lambda}_j)\Gamma(\boldsymbol{\lambda}_j)^{-1} - \frac{|H|^{1/4}}{N}K_H * K_H(\boldsymbol{\lambda}_i-\boldsymbol{\lambda}_j)\Gamma(\boldsymbol{\lambda}_j)^{-1} = a_{i,j}^1 - a_{i,j}^2,$$

where

$$a^1_{ij} = \frac{|H|^{1/4}}{N} 2K_H(\boldsymbol{\lambda}_i - \boldsymbol{\lambda}_j)\Gamma(\boldsymbol{\lambda}_j)^{-1}, \quad a^2_{ij} = \frac{|H|^{1/4}}{N} K_H * K_H(\boldsymbol{\lambda}_i - \boldsymbol{\lambda}_j)\Gamma(\boldsymbol{\lambda}_j)^{-1}.$$

Now, if we define:

$$c^1_{ij} = a^1_{ij}, \quad c^2_{ij} = c^1_{ji},$$
$$c^3_{ij} = a^2_{ij}, \quad c^4_{ij} = c^3_{ji}.$$

Define also, $W_{ij} = \left(c^1_{ij} + c^2_{ij} - c^3_{ij} - c^4_{ij}\right)\varepsilon_i\varepsilon_j$. Then, W_N can be wrriten as:

$$W_N = \sum_{i<j} W_{ij}.$$

The variance of this form is given by(29). In order to apply Proposition 3.2 in (de Jong (1987)), we must check some conditions on W_N. The first one is the W_N is clean, but this is clear, by definition (see definition 2.1 in de Jong (1987)). Consider:

$$\begin{aligned}
G_I &= \sum_{i<j} W^4_{ij}, \\
G_{II} &= \sum_{i<j<k} \left\{E\left(W^2_{ij}W^2_{ik}\right) + E\left(W^2_{ji}W^2_{jk}\right) + E\left(W^2_{ki}W^2_{kj}\right)\right\}, \\
G_{III} &= \sum_{i<j<k<l} \left\{E\left(W_{ij}W_{ik}W_{lj}W_{lk}\right) + E\left(W_{ij}W_{il}W_{kj}W_{kl}\right) + E\left(W_{ik}W_{il}W_{jk}W_{jl}\right)\right\}.
\end{aligned}$$

We must check that G_I, G_{II} and G_{III} are of smaller order than $Var(W_N)$, which is given by (29). It is easy to see that $G_I = \mathcal{O}(N^{-2}|H|^{-1/2})$, just taking into account that:

$$E\left(a^1_{ij}\varepsilon_i\varepsilon_j\right)^4 = \mathcal{O}\left(\frac{|H|}{N^4}16K^4_H(\boldsymbol{\lambda}_\mathbf{i} - \boldsymbol{\lambda}_\mathbf{j})\Gamma(\boldsymbol{\lambda}_\mathbf{j})^{-4}\right) = \mathcal{O}\left(\frac{1}{N^4|H|^{1/2}}\right),$$

and $E(a^3_{ij}\varepsilon_i\varepsilon_j)^4 = \mathcal{O}(N^{-4})$. G_{II} is $\mathcal{O}(N^{-1}|H|^{-1/2})$, since:

$$E\left((a^1_{ij})^2(a^2_{ij})^2\right) = \mathcal{O}\left(\frac{1}{N^4|H|^{1/2}}\right).$$

Similar computations lead to $G_{III} = \mathcal{O}\left(|H|^{1/2}\right)$. Then, we have that $W_N \to N(0,\sigma^2)$. Finally, we must find a bound for R_{1N}, R_{2N} and R_{3N} in (85), (85) and (85), respectively, with $H_N(\boldsymbol{\lambda})$ is given by (84). We can see that both R_{1N} and R_{3N} are stochastically bounded. In fact $R_{1N} = N^{1/2}|H|R_{1N0}$, where

$$R_{1N0} = \frac{1}{\sqrt{N}}\sum_{\mathbf{k}} \varepsilon_{\mathbf{k}} \int \mathbf{u}^T H_{m_\theta}(\boldsymbol{\lambda}_\mathbf{k})\mathbf{u}K(\mathbf{u})d\mathbf{u}$$

and $R_{3N} = N^{1/2}|H|R_{3N0}$, where

$$R_{3N0} = \frac{-1}{\sqrt{N}}\sum_{\mathbf{k}} \varepsilon_{\mathbf{k}}\Gamma(\boldsymbol{\lambda}_\mathbf{k})^{-1} \int\int \mathbf{s}^T H_{m_\theta}(\boldsymbol{\lambda}_\mathbf{k})\mathbf{s}\mathbf{u}^T H_{m_\theta}(\boldsymbol{\lambda}_\mathbf{k})\mathbf{u}K(\mathbf{s})K(\mathbf{u})d\mathbf{s}d\mathbf{u}.$$

Both R_{1N0} and R_{3N0} are asymptotically normal, and therefore, stochastically bounded. The remaining residual term, R_{2N} admits the following asymptotic expression:

$$R_{2N} = \frac{-|H|^2}{8}\sum_{\mathbf{k}} \frac{1}{f_\theta(\boldsymbol{\lambda}_\mathbf{k})} \int\int \mathbf{s}^T H_{m_\theta}(\boldsymbol{\lambda}_\mathbf{k})\mathbf{s}(\mathbf{s}+\mathbf{u})^T H_{m_\theta}(\boldsymbol{\lambda}_\mathbf{k})(\mathbf{s}+\mathbf{u})K(\mathbf{s})K(\mathbf{u})d\mathbf{s}d\mathbf{u}.$$

An additional term of the bias, b_H is obtained from R_{2N}, as $N^{1/2}|H| \to \infty$. $\square$

The following lemmas are needed for bounding B_1, B_2 and B_3 in Lemma 9.

Lemma 10. *Define*

$$\Psi_N(\boldsymbol{\lambda}) = |H|^{1/2}(N|H|^{1/2})^{-1/2}\sum_{\mathbf{k}}(e^{Y^{**}_{\mathbf{k}}-\bar{m}_{\mathbf{k}}}-1)W_{\mathbf{k}}K_H(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{k}}) \tag{89}$$

By Taylor's expansion and conditions in Theorem 3, the following hold also uniformly in $\boldsymbol{\lambda}$:

$$|H|^{1/2}l(\boldsymbol{\beta}) = \Psi_N(\boldsymbol{\lambda})^T\boldsymbol{\beta} + \frac{1}{2}\boldsymbol{\beta}^T A\boldsymbol{\beta} + \Delta_1(\boldsymbol{\beta}),$$

where $l(\boldsymbol{\beta})$ is given by:

$$l(\boldsymbol{\beta}) = \sum_{\mathbf{k}}\left[-(N|H|^{1/2})^{-1/2}\boldsymbol{\beta}^T W_{\mathbf{k}} - e^{Y^{**}_{\mathbf{k}}-\bar{m}_{\mathbf{k}}-(N|H|^{1/2})^{-1/2}\boldsymbol{\beta}^T W_{\mathbf{k}}} + e^{Y^{**}_{\mathbf{k}}-\bar{m}_{\mathbf{k}}}\right]K_H(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{k}}).$$

Proof. The expression for $l(\beta)$ is given by:

$$\begin{aligned}&|H|^{1/2}l(\boldsymbol{\beta}) = \\ &|H|^{1/2}\sum_{\mathbf{k}}\left\{-(N|H|^{1/2})^{-1/2}\boldsymbol{\beta}^T W_{\mathbf{k}} - e^{Y^{**}_{\mathbf{k}}-(N|H|^{1/2})^{-1/2}\boldsymbol{\beta}^T W_{\mathbf{k}}} + e^{Y^{**}_{\mathbf{k}}-\bar{m}_k}\right\}K_H(\boldsymbol{\lambda}_k-\boldsymbol{\lambda}),\end{aligned}$$

and by Taylor's expansion, we can write:

$$\begin{aligned}&e^{Y^{**}_{\mathbf{k}}-(N|H|^{-1/2})^{-1/2}\boldsymbol{\beta}^T W_{\mathbf{k}}} = e^{Y^{**}_{\mathbf{k}}-\bar{m}_{\mathbf{k}}} + (N|H|^{-1/2})^{-1/2}\boldsymbol{\beta}^T W_{\mathbf{k}}e^{Y^{**}_{\mathbf{k}}-\bar{m}_{\mathbf{k}}} + \\ &\frac{1}{2}(N|H|^{-1/2})^{-1}\boldsymbol{\beta}^T W_{\mathbf{k}}W_{\mathbf{k}}^T\boldsymbol{\beta}e^{Y^{**}_{\mathbf{k}}-c_{\mathbf{k}}},\end{aligned}$$

where $c_{\mathbf{k}}$ is such that $|(N|H|^{-1/2})^{-1/2}\boldsymbol{\beta}^T W_{\mathbf{k}} - \bar{m}_{\mathbf{k}}| \geq |c_{\mathbf{k}} - \bar{m}_{\mathbf{k}}|$. Then, $|H|^{1/2}l(\boldsymbol{\beta})$ is given by:

$$\begin{aligned}&|H|^{1/2}\sum_{\mathbf{k}}\Big\{-(N|H|^{1/2})^{-1/2}\boldsymbol{\beta}^T W_{\mathbf{k}} - (N|H|^{-1/2})^{-1/2}\boldsymbol{\beta}^T W_{\mathbf{k}}e^{Y^{**}_{\mathbf{k}}-\bar{m}_k} \\ &-\frac{1}{2}(N|H|^{1/2})^{-1}\boldsymbol{\beta}^T W_{\mathbf{k}}W_{\mathbf{k}}^T\boldsymbol{\beta}(e^{Y^{**}_{\mathbf{k}}-c_{\mathbf{k}}}-1+1)\Big\}K_H(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{k}}) \\ &= |H|^{1/2}(N|H|^{1/2})^{-1/2}\sum_{\mathbf{k}}(e^{Y^{**}_{\mathbf{k}}-\bar{m}_k}-1)W_{\mathbf{k}}^T K_H(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{k}})\boldsymbol{\beta} \\ &-\frac{1}{2}|H|^{1/2}(N|H|^{1/2})^{-1}\sum_{\mathbf{k}}\boldsymbol{\beta}^T W_{\mathbf{k}}W_{\mathbf{k}}^T\boldsymbol{\beta}K_H(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{k}}) \\ &-\frac{1}{2}|H|^{1/2}(N|H|^{1/2})^{-1}\sum_{\mathbf{k}}\boldsymbol{\beta}^T W_{\mathbf{k}}W_{\mathbf{k}}^T\boldsymbol{\beta}(e^{Y^{**}_{\mathbf{k}}-c_{\mathbf{k}}}-1)K_H(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{k}}) \\ &= \Psi_N(\boldsymbol{\lambda})^T\boldsymbol{\beta} + \frac{1}{2N}\sum_{\mathbf{k}}\boldsymbol{\beta}^T W_{\mathbf{k}}W_{\mathbf{k}}^T\boldsymbol{\beta}K_H(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{k}}) + \Delta_1(\boldsymbol{\beta}) \\ &= \Psi_N(\boldsymbol{\lambda})^T\boldsymbol{\beta} + \frac{1}{2N}\sum_{\mathbf{k}}\boldsymbol{\beta}^T A_{\mathbf{k}}\boldsymbol{\beta}K_H(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{k}}) + \Delta_1(\boldsymbol{\beta}).\end{aligned}$$

The matrix $A_{\mathbf{k}}$ is given by $W_{\mathbf{k}}W_{\mathbf{k}}^T$:

$$A_{\mathbf{k}} = W_{\mathbf{k}}W_{\mathbf{k}}^T = \begin{pmatrix} 1 & |H|^{-1/2}(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{k}})^T \\ |H|^{-1/2}(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{k}})^T & |H|^{-1/4}(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{k}})(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{k}})^T \end{pmatrix}$$

and it converges in probability to

$$\frac{1}{N}\sum_{\mathbf{k}} A_{\mathbf{k}} K_H(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{k}}) \to_{\mathbb{P}} A,$$

where A is given by

$$A = \frac{-1}{4\pi^2}\begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \int \mathbf{u}\mathbf{u}^T K(\mathbf{u})d\mathbf{u} \end{pmatrix}.$$

The residual term $\Delta_1(\boldsymbol{\beta})$ is $\mathcal{O}_{\mathbb{P}}(1)$, provided that $N^{(\zeta-1)/\zeta}|H|^{1/2} \geq c_0 \log N$:

$$\Delta_1(\boldsymbol{\beta}) = \frac{-|H|^{1/2}}{2}(N|H|^{1/2})^{-1}\sum_{\mathbf{k}} \boldsymbol{\beta}^T W_{\mathbf{k}} W_{\mathbf{k}}^T \boldsymbol{\beta}(e^{Y_{\mathbf{k}}^{**}-c_{\mathbf{k}}}-1)K_H(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{k}}).$$

Since $|r_N\boldsymbol{\beta}^T W_{\mathbf{k}} - \bar{m}_{\mathbf{k}}| \geq |c_{\mathbf{k}} - \bar{m}_{\mathbf{k}}|$, $c_{\mathbf{k}}$ can be written as $c_{\mathbf{k}} = \boldsymbol{\beta}^T W_{\mathbf{k}} + \alpha^T W_{\mathbf{k}}$, where $\|\alpha\| \leq c_1 r_N$, for some $c_1 > 0$. Then, $\Delta_1(\boldsymbol{\beta})$ can be decomposed in two addends:

$$\Delta_1^1(\boldsymbol{\beta}) = \frac{-|H|^{1/2}}{2}(N|H|^{1/2})^{-1}\sum_{\mathbf{k}} q_2(\boldsymbol{\beta}^T W_{\mathbf{k}} + \alpha^T W_{\mathbf{k}}, Y_{\mathbf{k}}^{**})\boldsymbol{\beta}^T W_{\mathbf{k}} W_{\mathbf{k}}^T \boldsymbol{\beta} K_H(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{k}}),$$

$$\Delta_1^2(\boldsymbol{\beta}) = \frac{|H|^{1/2}}{2}(N|H|^{1/2})^{-1}\sum_{\mathbf{k}} \boldsymbol{\beta}^T W_{\mathbf{k}} W_{\mathbf{k}}^T \boldsymbol{\beta} K_H(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{k}}).$$

Uniform bound $\mathcal{O}_{\mathbb{P}}(1)$ for both addends is obtained from condition on $\Phi_{n,j}$, for the particular case of $j = 1$ and $j = 3$. Then, the expression for $l(\boldsymbol{\beta})$ is proved. □

Lemma 11. *If f_θ is twice differentiable, we have the following representation for the difference between the non parametric estimation $\hat{m}^*_{LK}$ and the log-spectral density under the null hypothesis m_θ, in a frequency $\boldsymbol{\lambda}$ and under conditions in Theorem 3:*

$$\hat{m}^*_{LK}(\boldsymbol{\lambda}) - m_\theta(\boldsymbol{\lambda}) = r_N^2 \Gamma(\boldsymbol{\lambda})^{-1}\sum_{\mathbf{k}} \varepsilon_{\mathbf{k}} K\left(H^{-1/2}(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{k}})\right)(1 + o_{\mathbb{P}}(1)) + H_N(\boldsymbol{\lambda}), \qquad (90)$$

where H_N is given by (84).

Proof. Using the expression for $l(\boldsymbol{\beta})$ obtained in Lemma 10 and applying the convexity lemma of Pollard (1991) we obtain the maximizer $\hat{\boldsymbol{\beta}}$ of the expression for $l(\boldsymbol{\beta})$ is given by

$$\hat{\boldsymbol{\beta}} = B^{-1}\Psi_N(\boldsymbol{\lambda}) + o_{\mathbb{P}}(1).$$

The inverse of matrix B is given by:

$$B^{-1} = -\pi^2 \begin{pmatrix} 1 & 0 & 0 \\ 0 & b_2 & -b_{12} \\ 0 & -b_{12} & b_1 \end{pmatrix},$$

$$b_j = \frac{\int u_j^2 K(\mathbf{u})d\mathbf{u}}{b_1}, j = 1, 2, \quad b_{12} = \frac{-\int u_1 u_2 K(\mathbf{u})d\mathbf{u}}{a_1},$$

$$b_1 = \int u_1^2 K(\mathbf{u})d\mathbf{u} \int u_2^2 K(\mathbf{u})d\mathbf{u} - \left(\int u_1 u_2 K(\mathbf{u})d\mathbf{u}\right)^2,$$

where u_1 and u_2 denote the first and the second components of vector $\mathbf{u} \in \mathbb{R}^2$. The first component of $\boldsymbol{\beta}$ is

$$\hat{\beta}_{(1)} = (N|H|^{1/2})^{-1/2}(\hat{m}_{LK}(\boldsymbol{\lambda}) - m_\theta(\boldsymbol{\lambda})).$$

We obtain, from the expression for $\Psi_n(\boldsymbol{\lambda})$ in (89):

$$\begin{aligned}
&\Psi_N(\boldsymbol{\lambda}) = |H|^{1/2}(N|H|^{1/2})^{-1/2}\sum_{\mathbf{k}}(e^{Y^{**}_{\mathbf{k}} - \bar{m}_{\mathbf{k}}} - 1)W_{\mathbf{k}}K_H(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{k}}) = \\
&|H|^{1/2}(N|H|^{1/2})^{-1/2}\sum_{\mathbf{k}}(e^{Y^{**}_{\mathbf{k}} - \bar{m}_{\mathbf{k}}} - 1)W_{\mathbf{k}}K_H(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{k}}) = \\
&|H|^{1/2}(N|H|^{1/2})^{-1/2}\sum_{\mathbf{k}}(e^{Y^{**}_{\mathbf{k}} - \boldsymbol{\beta}^T W_{\mathbf{k}}} - 1)W_{\mathbf{k}}K_H(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{k}}) = \\
&|H|^{1/2}(n|H|^{1/2})^{-1/2}\sum_{\mathbf{k}}(e^{Y^{**}_{\mathbf{k}} - \boldsymbol{\beta}^T W_{\mathbf{k}}} + e^{Y^{**}_{\mathbf{k}} - m_\theta(\boldsymbol{\lambda}_{\mathbf{k}})} - e^{Y^{**}_{\mathbf{k}} - m_\theta(\boldsymbol{\lambda}_{\mathbf{k}})} - 1)W_{\mathbf{k}}K_H(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{k}}) = \\
&|H|^{1/2}(N|H|^{1/2})^{-1/2}\sum_{\mathbf{k}}(e^{Y^{**}_{\mathbf{k}} - m_\theta(\boldsymbol{\lambda}_{\mathbf{k}})} - 1)W_{\mathbf{k}}K_H(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{k}}) + \\
&|H|^{1/2}(N|H|^{1/2})^{-1/2}\sum_{\mathbf{k}}(e^{Y^{**}_{\mathbf{k}} - \boldsymbol{\beta}^T W_{\mathbf{k}}} - e^{Y^{**}_{\mathbf{k}} - m_\theta(\boldsymbol{\lambda}_{\mathbf{k}})})W_{\mathbf{k}}K_H(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{k}}).
\end{aligned}$$

The result is proved just considering the first component of $\Psi_n(\boldsymbol{\lambda})$.

$$\begin{aligned}
&\Psi^{(1)}_N(\boldsymbol{\lambda}) = |H|^{1/2}(N|H|^{1/2})^{-1/2}\sum_{\mathbf{k}}(e^{Y^{**}_{\mathbf{k}} - m_\theta(\boldsymbol{\lambda}_{\mathbf{k}})} - 1)K_H(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{k}}) + \\
&|H|^{1/2}(N|H|^{1/2})^{-1/2}\sum_{\mathbf{k}}(e^{Y^{**}_{\mathbf{k}} - \boldsymbol{\beta}^T W_{\mathbf{k}}} - e^{Y^{**}_{\mathbf{k}} - m_\theta(\boldsymbol{\lambda}_{\mathbf{k}})})K_H(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{k}}) = \\
&(N|H|^{1/2})^{-1/2}\sum_{\mathbf{k}}\varepsilon_{\mathbf{k}}K(H^{-1/2}(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{k}})) + \\
&(N|H|^{1/2})^{-1/2}\sum_{\mathbf{k}}(q_1(\boldsymbol{\beta}^T W_{\mathbf{k}}, Y^{**}_{\mathbf{k}}) - \varepsilon_k)K(H^{-1/2}(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{k}})).
\end{aligned}$$

□

Lemma 12. *Under assumption (1)-(3), we have*

$$\sup_{\boldsymbol{\lambda} \in [0,\pi]^2} |\hat{m}_{LK}(\boldsymbol{\lambda}) - \hat{m}^*_{LK}(\boldsymbol{\lambda})| = O_{\mathbb{P}}(N^{-1/2}\log N)$$

Proof. The proof of this Lemma is obatined using similar arguments as that for the proof of Lemma 11. Recall the expression for the local loglikelihood given by:

$$\sum_{\mathbf{k}}\left[Y_{\mathbf{k}} - a - b^T(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{k}}) - e^{Y_{\mathbf{k}} - a - b^T(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{k}})}\right]K_H(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{k}}). \tag{91}$$

This expression can be written in terms of the vector $\boldsymbol{\beta}_2$ as

$$L(\boldsymbol{\beta}_2) = \sum_{\mathbf{k}}\left[Y_{\mathbf{k}} - \bar{m}_{\mathbf{k}} - (N|H|^{1/2})^{-1/2}\boldsymbol{\beta}_2^T W_{\mathbf{k}} - e^{Y_{\mathbf{k}} - \bar{m}_{\mathbf{k}} - (N|H|^{1/2})^{-1/2}\boldsymbol{\beta}_2^T W_{\mathbf{k}}}\right]K_H(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{k}})$$

and the difference:

$$L(\boldsymbol{\beta}_2) - L(0) = \sum_{\mathbf{k}}\left[-(N|H|^{1/2})^{-1/2}\boldsymbol{\beta}_2^T W_{\mathbf{k}} - e^{Y_{\mathbf{k}} - \bar{m}_{\mathbf{k}} - (N|H|^{1/2})^{-1/2}\boldsymbol{\beta}_2^T W_{\mathbf{k}}}\right]K_H(\boldsymbol{\lambda} - \boldsymbol{\lambda}_{\mathbf{k}}).$$

If we set

$$U_n(\boldsymbol{\beta}_2) = \sum_{\mathbf{k}} r_N(\boldsymbol{\lambda}_{\mathbf{k}}) \left[e^{-\bar{m}_{\mathbf{k}}-(n|H|^{1/2})^{-1/2}\boldsymbol{\beta}_2^T W_{\mathbf{k}}} - e^{-\bar{m}_{\mathbf{k}}}\right] K_H(\boldsymbol{\lambda}-\boldsymbol{\lambda}_{\mathbf{k}})$$

Then,

$$L(\boldsymbol{\beta}_2) - L(0) = l(\boldsymbol{\beta}_2) - U_N(\boldsymbol{\beta}_2)$$

An uniform bound for $U_N(\boldsymbol{\beta}_2)$ is easily found just by Taylor's expansion and using the bound $\max_{\mathbf{k}} |r_N(\boldsymbol{\lambda}_{\mathbf{k}})| = \mathcal{O}_{\mathbb{P}}(N^{-1/2}\log N)$:

$$|H|^{1/2} \sup_{\boldsymbol{\lambda}\in[0,\pi]^2} U_N(\boldsymbol{\beta}_2) = O_{\mathbb{P}}(|H|^{-1/4}\log N)$$

With the same arguments as in Lemma 11, we show that the following bounds also hold uniformly in $\boldsymbol{\lambda}$:

$$|H|^{1/2}(l(\boldsymbol{\beta}_2) - U_N(\boldsymbol{\beta}_2)) = |H|^{1/2}(L(\boldsymbol{\beta}_2) - L(\mathbf{0})) = \Psi_N(\boldsymbol{\lambda})^T\boldsymbol{\beta}_2 + \frac{1}{2}\boldsymbol{\beta}_2^T A\boldsymbol{\beta}_2 + \Delta_2(\boldsymbol{\beta}_2)$$

$$\Delta_1(\boldsymbol{\beta}_2) = O_{\mathbb{P}}(1), \qquad \Delta_2(\boldsymbol{\beta}_2) = O_{\mathbb{P}}(1)$$
$$\nabla\Delta_1(\boldsymbol{\beta}_2) = O_{\mathbb{P}}((N|H|^{1/2})^{-1/2}\log|H|^{1/2}\alpha_N + |H|)$$
$$\Delta_2(\boldsymbol{\beta}_2) = \nabla\Delta_1(\boldsymbol{\beta}_2) + O_{\mathbb{P}}(|H|^{1/4}\log N)$$

where $\alpha_N \to \infty$. Using the same arguments as that for the proof of Theorem 2 in Carroll *et al.* (1997) and the proof of the quadratic approximation lemma in Fan and Gijbels (1995) we obtain:

$$(N|H|^{1/2})^{1/2}\{\hat{m}_{LK}(\boldsymbol{\lambda}) - m(\boldsymbol{\lambda})\} = (\pi^2, 0, 0)\Psi_N(\boldsymbol{\lambda}) + O_{\mathbb{P}}\left(|H|^{1/4}\log N\right), \tag{92}$$

and

$$(N|H|^{1/2})^{1/2}\{\hat{m}^{**}_{LK}(\boldsymbol{\lambda}) - m(\boldsymbol{\lambda})\} = (\pi^2, 0, 0)\Psi_N(\boldsymbol{\lambda}) + O_{\mathbb{P}}\left(\frac{\log|H|^{1/2}}{\sqrt{N|H|^{1/2}}}\alpha_N + |H|\right). \tag{93}$$

□

Lemma 13. *Assume that $\varepsilon_1, \ldots, \varepsilon_N$ are independent identically distributed random variables, with $E(\varepsilon_1) = 0$ and $E(|\varepsilon_1|^s) < \infty$, for some $s > 2$. Assume that $\mathbf{x}_1, \ldots, \mathbf{x}_N$ are fixed design points in $[0,1]^2 \subset \mathbb{R}^2$ such that $\mathbf{x}_i \in A_i \subset \mathbb{R}^2$, $\cup_{i=1}^N A_i = [0,1]^2$, $A_i \cap A_j = \oslash$, where A_i is Jordan-measurable, with $\max_i \mu(A_i) = \mathcal{O}(N^{-1})$, where μ is the Jordan measure and*

$$\max_i d(A_i) = \mathcal{O}(N^{-1}),$$

where $d(B) = \sup_{\mathbf{x},\mathbf{y}\in B}\|\mathbf{x}-\mathbf{y}\|$, $\|\cdot\|$ is the L^2-norm. Assume that W is a weight function satisfying a Lipchitz condition and:

$$\max_i |W_i(\mathbf{x}))| \geq c_0 N^{-1},$$

uniformly in $\mathbf{x} \in [0,1]^2$, for a constant c_0. Finally assume that there is a sequence $\alpha_N \to 0$ and constants $\eta \in (0, s-2)$, $c > 1/2$ such that, for all $\mathbf{x} \in [0,1]^2$:

$$N^{2/(s-\eta)} \max_i |W_i(\mathbf{x})| logN \leq \alpha_N c, \quad \textit{and} \quad \left(\sum_i W_i(\mathbf{x}) \log N\right)^2 \leq \alpha_N c. \tag{94}$$

Then:

$$\sup_{\mathbf{x}\in[0,1]^2}\left|\sum_i W_i(\mathbf{x})\varepsilon_i\right| = \mathcal{O}(\alpha_N).$$

Proof. This lemma is a straightforward extension of Theorem 11.2 in Müller (1988). The proof is similar, since the stochastic part is not affected by the dimension. □

Lemma 14. *Assume conditions in Lemma 13 hold and suppose that the weight functions are kernel weights:*

$$W_i(\mathbf{x}) = |H|^{-1/2}K(H^{-1/2}(\mathbf{x}-\mathbf{x}_i)).$$

Then,

$$\sup_{\mathbf{x}\in[0,1]^2}\frac{1}{N}\left|\sum_i K_H(\mathbf{x}-\mathbf{x}_i)\varepsilon_i\right| = o((N|H|^{1/2})^{-1/2}(-\log|H|^{1/2})\beta_N),$$

where the sequence $\beta_N \to \infty$ *and provided that there exists* $s > 2$*,* $\eta \in (0, s-2)$ *such that* $N^{2/(s-\eta)}|H|^{-1/2}\log N \to C$*, for some constant* C*.*

Proof. The proof is inmediate from Lemma 13. The condition on s and η is obtained from the restriction (94) on the kernel weights. □

Lemma 15. *The terms* B_1*,* B_2 *and* B_3 *in Lemma 9 are bounded by:*

$$B_1 = \mathcal{O}_{\mathbb{P}}\left(\frac{\log N}{\sqrt{N}}|H|^{-1/2}\log|H|^{1/2}\alpha_N\right), +\mathcal{O}_{\mathbb{P}}(\log^2 N)$$
$$B_2 = \mathcal{O}_{\mathbb{P}}(\log^2 N),$$
$$B_3 = \mathcal{O}_{\mathbb{P}}(|H|^{-1/4}\log N(-\log|H|^{1/2})\alpha_N),$$

where $\alpha_N \to \infty$*.*

Proof. Recall the expression for B_2 is given by:

$$B_2 = \sum_{\mathbf{k}} e^{Y_{\mathbf{k}}-\hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})}\left(\hat{m}_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}) - \hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})\right)^2.$$

By Taylor's expansion of $e^{\mathbf{x}}$ and Lemma 12:

$$\begin{aligned}
B_2 &= \sum_{\mathbf{k}}[1 + Y_{\mathbf{k}} - \hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}) + \mathcal{O}_{\mathbb{P}}((Y_{\mathbf{k}} - \hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}))^2)]\cdot(\hat{m}_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}) - \hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}))^2 \approx \\
&\sum_{\mathbf{k}}(\hat{m}_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}) - \hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}))^2 + \sum_{\mathbf{k}}(Y_{\mathbf{k}} - \hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}))(\hat{m}_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}) - \hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}))^2 \leq \\
&\leq N\sup_{\mathbf{k}}|\hat{m}_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}) - \hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})|^2 = \mathcal{O}_{\mathbb{P}}(\log^2 N).
\end{aligned}$$

Just taking into account that:

$$e^{Y_{\mathbf{k}}} = I(\boldsymbol{\lambda}_{\mathbf{k}}) = f_\theta(\boldsymbol{\lambda}_{\mathbf{k}})V_{\mathbf{k}} + r_N(\boldsymbol{\lambda}_{\mathbf{k}}) = V_{\mathbf{k}}e^{m_\theta(\boldsymbol{\lambda}_{\mathbf{k}})} + r_N(\boldsymbol{\lambda}_{\mathbf{k}})$$

the term B_1 is decomposed in two addends:

$$\begin{aligned}
B_1 &= \sum_{\mathbf{k}}(1 - e^{Y_{\mathbf{k}}-\hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})})\left(\hat{m}_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}) - \hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})\right) \\
&= \sum_{\mathbf{k}}(1 - V_{\mathbf{k}}e^{m_\theta(\boldsymbol{\lambda}_{\mathbf{k}})-\hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})})(\hat{m}_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}) - \hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})) + \\
&\sum_{\mathbf{k}} r_N(\boldsymbol{\lambda}_{\mathbf{k}})e^{-\hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})}(\hat{m}_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}) - \hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})) = \\
&B_{1,1} + B_{1,2}.
\end{aligned}$$

The second addend, $B_{1,2}$ can be bounded by:

$$
\begin{aligned}
&B_{1,2} \leq \sum_{\mathbf{k}} \sup_{\mathbf{k}} |r_N(\boldsymbol{\lambda}_{\mathbf{k}})| \cdot |\hat{m}_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}) - \hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})| \leq \\
&\sum_{\mathbf{k}} \mathcal{O}_{\mathbb{P}}(N^{-1/2}\log N) \sup_{\mathbf{k}} |\hat{m}_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}) - \hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})| = \mathcal{O}_{\mathbb{P}}\left(\log^2 N\right).
\end{aligned}
$$

For the first addedn, $B_{1,1}$, applying Taylor's expansion on $e^{m_\theta(\boldsymbol{\lambda}_{\mathbf{k}})-\hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})}$ around the origin, we have:

$$
\begin{aligned}
B_{1,1} &= \sum_{\mathbf{k}} \left(1 - V_{\mathbf{k}} e^{m_\theta(\boldsymbol{\lambda}_{\mathbf{k}})-\hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})}\right)(\hat{m}_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}) - \hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})) \\
&= \sum_{\mathbf{k}} (1 - V_{\mathbf{k}} - V_{\mathbf{k}}(m_\theta(\boldsymbol{\lambda}_{\mathbf{k}}) - \hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}))e^{c_{\mathbf{k}}})(\hat{m}_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}) - \hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})),
\end{aligned}
$$

where $c_{\mathbf{k}}$ satisfies that $|m_\theta(\boldsymbol{\lambda}_{\mathbf{k}}) - \hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})| \geq |c_{\mathbf{k}}|$. Then,

$$
\begin{aligned}
&B_{1,1} = \sum_{\mathbf{k}} (V_{\mathbf{k}} - 1)(\hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}) - \hat{m}_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})) \\
&- \sum_{\mathbf{k}} V_{\mathbf{k}}(m_\theta(\boldsymbol{\lambda}_{\mathbf{k}}) - \hat{m}_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}))(\hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}) - \hat{m}_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}))e^{c_{\mathbf{k}}}.
\end{aligned}
$$

The variables $V_{\mathbf{k}}$ are zero-mean, so the first term in $B_{1,2}$ is $o_{\mathbb{P}}(1)$. For the first addend, applying Lemma 12:

$$
\begin{aligned}
&B_{1,1} = \sum_{\mathbf{k}} V_{\mathbf{k}}(m_\theta(\boldsymbol{\lambda}_{\mathbf{k}}) - \hat{m}^*_{LK}\boldsymbol{\lambda}_{\mathbf{k}})(\hat{m}_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}) - \hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}}))e^{c_{\mathbf{k}}} \\
&\leq \mathcal{O}_{\mathbb{P}}(N^{-1/2}\log N) \sum_{\mathbf{k}} (m_\theta(\boldsymbol{\lambda}_{\mathbf{k}}) - \hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})) \\
&= \mathcal{O}_{\mathbb{P}}(N^{-1/2}\log N) \\
&\times \sum_{\mathbf{k}} \left((\pi^2, 0, 0)\Psi_N(\boldsymbol{\lambda}_{\mathbf{k}}) + \mathcal{O}_{\mathbb{P}}((N|H|^{1/2})^{-1/2}) \log |H|^{1/2}\alpha_N + |H|\right)(N|H|^{1/2})^{-1/2} \\
&= \mathcal{O}_{\mathbb{P}}\left(\frac{\log N}{\sqrt{N}}\right) \\
&\times \sum_{\mathbf{k}} \left(\frac{|H|^{1/4}}{\sqrt{N}} \sum_{\mathbf{j}} (e^{Y^{**}_{\mathbf{j}} - \bar{m}_{\mathbf{j}}} - 1)K_H(\boldsymbol{\lambda}_{\mathbf{k}} - \boldsymbol{\lambda}_{\mathbf{j}}) + \mathcal{O}_{\mathbb{P}}((N|H|^{1/2})^{-1/2} \log |H|^{1/2}\alpha_N + |H|)\right),
\end{aligned}
$$

applying the expressions (92) and (93). By a Taylor's expansion on $e^{Y^{**}_{\mathbf{j}} - \bar{m}_{\mathbf{j}}}$ around the origin, Lemma 14 can be applied on the sum in the first addend. Now, applying Lemmas 2 and 12,

$$
B_{1,2} = \mathcal{O}_{\mathbb{P}}\left(\frac{\log N}{\sqrt{N}} |H|^{-1/2} \log |H|^{1/2}\alpha_N\right),
$$

For the last term, B_3, also using Lemmas 2, 12, and Lemma 14 in a Taylor's expansion for the expression of $\Psi_N(\boldsymbol{\lambda}_{\mathbf{k}})$ we derive:

$$
\begin{aligned}
&B_3 = \sum_{\mathbf{k}} \frac{R_N(\boldsymbol{\lambda}_{\mathbf{k}})}{f_\theta(\boldsymbol{\lambda}_{\mathbf{k}})} \left\{ e^{m_\theta(\boldsymbol{\lambda}_{\mathbf{k}}) - \hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})} - 1 \right\} \leq \\
&N \max_{\mathbf{k}} \frac{R_N(\boldsymbol{\lambda}_{\mathbf{k}})}{f_\theta(\boldsymbol{\lambda}_{\mathbf{k}})} \sup_{\mathbf{k}} \left| e^{m_\theta(\boldsymbol{\lambda}_{\mathbf{k}}) - \hat{m}^*_{LK}(\boldsymbol{\lambda}_{\mathbf{k}})} - 1 \right| \leq \\
&N \max_{\mathbf{k}} \frac{R_N(\boldsymbol{\lambda}_{\mathbf{k}})}{f_\theta(\boldsymbol{\lambda}_{\mathbf{k}})} \left(N|H|^{1/2} \right)^{-1/2} \\
&\qquad\qquad \times \sup_{\mathbf{k}} \left| (\pi^2, 0, 0)\Psi_N(\boldsymbol{\lambda}_{\mathbf{k}}) + \mathcal{O}_{\mathbb{P}} \left((N|H|^{1/2})^{-1/2} \log |H|^{1/2} \alpha_N + |H| \right) \right| = \\
&\mathcal{O}_{\mathbb{P}} \left((N|H|^{1/2})^{-1/2} \log N \log |H|^{1/2} \right) + \mathcal{O}_{\mathbb{P}} \left(|H|^{-1/4} \log N (-\log |H|^{1/2}) \alpha_N \right) + \\
&\mathcal{O}_{\mathbb{P}} \left(\log N |H|^{3/4} \right) = \mathcal{O}_{\mathbb{P}}(|H|^{-1/4} \log N (-\log |H|^{1/2}) \alpha_N).
\end{aligned}
$$

□

References

Alonso, F. J.; Angulo, J. M.; and Bueso, M. C. (1996). Estimation and smoothing from incomplete data for a class of lattice processes. *Comm. Statist. Theory Methods*, volume 25, no. 3:pp. 667–681.

Besag, J. (1974). Spatial interaction and the statistical analysis of lattice systems. *J. Roy. Statist. Soc. Ser. B*, volume 36:pp. 192–236.

Brilinger, D. R. (1974). Fourier analysis of stationary processes. In *Proceedings of the IEEE*, volume 62, pp. 1628–1643.

Brillinger, D. R. (1981). *Time series*. Oakland, Calif.: Holden-Day Inc., second edition.

Brockwell, P. J. and Davis, R. A. (1991). *Time series: theory and methods*. Springer Series in Statistics. New York: Springer-Verlag, second edition.

Carroll, R. J.; Fan, J.; Gijbels, I.; and Wand, M. P. (1997). Generalized partially linear single-index models. *J. Amer. Statist. Assoc.*, volume 92, no. 438:pp. 477–489.

Chilès, J.-P. and Delfiner, P. (1999). *Geostatistics*. Wiley Series in Probability and Statistics: Applied Probability and Statistics. New York: John Wiley & Sons Inc.

Cressie, N. A. C. (1993). *Statistics for spatial data*. Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics. New York: John Wiley & Sons Inc.

Dahlhaus, R. and Künsch, H. (1987). Edge effects and efficient parameter estimation for stationary random fields. *Biometrika*, volume 74, no. 4:pp. 877–882.

Dahlhaus, R. and Wefelmeyer, W. (1996). Asimptotically optimal estimation in misspecified times series models. *Annals of Statistics*.

de Jong, P. (1987). A central limit theorem for generalized quadratic forms. *Probab. Theory Related Fields*, volume 75, no. 2:pp. 261–277.

DELGADO, M.; HIDALGO, J.; and VELASCO, C. (2005). Distribucin free goodness-of-fit tests for linear processes. *Ann. Statist.*, volume 33, no. 6.

DIBLASI, A. and BOWMAN, A. W. (2001). On the use of the variogram in checking for independence in spatial data. *Biometrics*, volume 57, no. 1:pp. 211–218.

FAN, J. and GIJBELS, I. (1995). Data-driven bandwidth selection in local polynomial fitting: variable bandwidth and spatial adaptation. *J. Roy. Statist. Soc. Ser. B*, volume 57, no. 2:pp. 371–394.

FAN, J.; ZHANG, C.; and ZHANG, J. (2001). Generalized likelihood ratio statistics and Wilks phenomenon. *Ann. Statist.*, volume 29, no. 1:pp. 153–193.

FAN, J. and ZHANG, W. (2004). Generalised likelihood ratio tests for spectral density. *Biometrika*, volume 91, no. 1:pp. 195–209.

FRANCISCO-FERNANDEZ, M. and OPSOMER, J. D. (2005). Smoothing parameter selection methods for nonparametric regression with spatially correlated errors. *Canad. J. Statist.*, volume 33, no. 2:pp. 279–295.

FUENTES, M. (2002). Spectral methods for nonstationary spatial processes. *Biometrika*, volume 89, no. 1:pp. 197–210.

GARCÍA-SOIDÁN, P. H.; FEBRERO-BANDE, M.; and GONZÁLEZ-MANTEIGA, W. (2004). Nonparametric kernel estimation of an isotropic variogram. *J. Statist. Plann. Inference*, volume 121, no. 1:pp. 65–92.

GARCÍA-SOIDÁN, P. H.; GONZÁLEZ-MANTEIGA, W.; and FEBRERO-BANDE, M. (2003). Local linear regression estimation of the variogram. *Statist. Probab. Lett.*, volume 64, no. 2:pp. 169–179.

GOLDBERG, D. (1989). *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley.

GONZÁLEZ MANTEIGA, W. and CAO, R. (1993). Testing the hypothesis of a general linear model using nonparametric regression estimation. *Test*, volume 2, no. 1-2:pp. 161–188.

GRENANDER, U. (1981). *Abstract Inference*. Wiley.

GUYON, X. (1982). Parameter estimation for a stationary process on a d-dimensional lattice. *Biometrika*, volume 69, no. 1:pp. 95–105.

HÄRDLE, W. and MAMMEN, E. (1993). Comparing nonparametric versus parametric regression fits. *Ann. Statist.*, volume 21, no. 4:pp. 1926–1947.

HART, J. D. (1997). *Nonparametric smoothing and lack-of-fit tests*. Springer Series in Statistics. New York: Springer-Verlag.

KOOPERBERG, C.; STONE, C. J.; and TRUONG, Y. K. (1995). Rate of convergence for logspline spectral density estimation. *J. Time Ser. Anal.*, volume 16, no. 4:pp. 389–401.

MAGLIONE, D. and DIBLASI, A. (2004). Exploring a valid model for the variogram of an isotropic spatial process. *Stoch. Envir. Res. and Risk Ass.*, volume 18:pp. 366–376.

MARTIN, R. J. (1979). A subclass of lattice processes applied to a problem in planar sampling. *Biometrika*, volume 66, no. 2:pp. 209–217.

MÜLLER, H.-G. (1988). *Nonparametric regression analysis of longitudinal data*, volume 46 of *Lecture Notes in Statistics*. Berlin: Springer-Verlag.

PAPARODITIS, E. (2000). Spectral density based goodness-of-fit tests for time series models. *Scand. J. Statist.*, volume 27, no. 1:pp. 143–176.

PRIESTLEY, M. B. (1981). *Spectral analysis and time series. Vol. 1.* London: Academic Press Inc. [Harcourt Brace Jovanovich Publishers].

ROBINSON, P. (2006). Nonparametric spectrum estimation for spatial data. *to appear in Journal of Statistical Planning and Inference.*

RUPPERT, D. and WAND, M. P. (1994). Multivariate locally weighted least squares regression. *Ann. Statist.*, volume 22, no. 3:pp. 1346–1370.

STEIN, M. L. (1995). Fixed domain asymptotics for spatial periodograms. *J. Am. Statist. Assoc.*, volume 90:pp. 1277–88.

STUTE, W. (1997). Nonparametric model checks for regression. *Ann. Statist.*, volume 25, no. 2:pp. 613–641.

STUTE, W.; GONZÁLEZ MANTEIGA, W.; and PRESEDO QUINDIMIL, M. (1998). Bootstrap approximations in model checks for regression. *J. Amer. Statist. Assoc.*, volume 93, no. 441:pp. 141–149.

WHITTLE, P. (1954). On stationary processes in the plane. *Biometrika*, volume 41:pp. 434–449.

YAGLOM, A. M. (1987). *Correlation theory of stationary and related random functions. Vol. II.* Springer Series in Statistics. New York: Springer-Verlag.

YOUNG, L. and YOUNG, J. (1998). *Statistical Ecology.* Kluwer Academic Press.

ZHANG, C. and DETTE, H. (2004). A power comparison between nonparametric regression tests. *Statist. Probab. Lett.*, volume 66, no. 3:pp. 289–301.

ZHENG, J. X. (1996). A consistent test of functional form via nonparametric estimation techniques. *J. Econometrics*, volume 75, no. 2:pp. 263–289.

Reports in Statistics and Operations Research

2004

04-01 Goodness of fit test for linear regression models with missing response data. *González Manteiga, W., Pérez González, A.*
Canadian Journal of Statistics (to appear).

04-02 Boosting for Real and Functional Samples. An Application to an Environmental Problem. *B. M. Fernández de Castro and W. González Manteiga*.

04-03 Nonparametric classification of time series: Application to the bank share prices in Spanish stock market. *Juan M. Vilar, José A. Vilar and Sonia Pértega*.

04-04 Boosting and Neural Networks for Prediction of Heteroskedatic Time Series. *J. M. Matías, M. Febrero, W. González Manteiga and J. C. Reboredo.*

04-05 Partially Linear Regression Models with Farima-Garch Errors. An Application to the Forward Exchange Market. *G. Aneiros Pérez, W. González Manteiga and J. C. Reboredo Nogueira.*

04-06 A Flexible Method to Measure Synchrony in Neuronal Firing. *C. Faes, H. Geys, G. Molenberghs, M. Aerts, C. Cadarso-Suárez, C. Acuña and M. Cano.*

04-07 Testing for factor-by-curve interactions in generalized additive models: an application to neuronal activity in the prefrontal cortex during a discrimination task. *J. Roca-Pardiñas, C. Cadarso-Suárez, V. Nacher and C. Acuña.*

04-08 Bootstrap Estimation of the Mean Squared Error of an EBLUP in Mixed Linear Models for Small Areas. *W. González Manteiga, M. J. Lombardía, I. Molina, D. Morales and L. Santamaría.*

04-09 Set estimation under convexity type assumptions. *A. Rodríguez Casal.*

2005

05-01 SiZer Map for Evaluating a Bootstrap Local Bandwidth Selector in Nonparametric Additive Models. *M. D. Martínez-Miranda, R. Raya-Miranda, W. González-Manteiga and A. González-Carmona.*

05-02 The Role of Commitment in Repeated Games. *I. García Jurado, Julio González Díaz*.

05-03 Project Games. *A. Estévez Fernández, P. Borm, H. Hamers*

05-04 Semiparametric Inference in Generalized Mixed Effects Models. *M. J. Lombardía, S. Sperlich*

2006

06-01 A unifying model for contests: effort-prize games. J. González Díaz

06-02 The Harsanyi paradox and the "right to talk" in bargaining among coalitions. J. J. Vidal Puga

06-03 A functional analysis of NOx levels: location and scale estimation and outlier detection. M. Febrero, P. Galeano, W. González-Manteiga

06-04 Comparing spatial dependence structures. R. M. Crujeiras, R. Fernández-Casal, W. González-Manteiga

06-05 On the spectral simulation of spatial dependence structures. R. M. Crujeiras, R. Fernández-Casal

06-06 An L_2-test for comparing spatial spectral densities. R. M. Crujeiras, R. Fernández-Casal, W. González-Manteiga.

2007

07-06 Goodness-of-fit tests for the spatial spectral density. R. M. Crujeiras, R. Fernández-Casal, W. González-Manteiga.

Previous issues (2001 – 2003):
http://eio.usc.es/pub/reports.html