

EFFICIENT ESTIMATION OF FUNCTIONALS OF THE SPECTRAL DENSITY OF STATIONARY GAUSSIAN FIELDS

CARENNE LUDENA¹

Abstract. Minimax bounds for the risk function of estimators of functionals of the spectral density of Gaussian fields are obtained. This result is a generalization of a previous result of Khas'minskii and Ibragimov on Gaussian processes. Efficient estimators are then constructed for these functionals. In the case of linear functionals these estimators are given for all dimensions. For non-linear integral functionals, these estimators are constructed for the two and three dimensional problems.

Résumé. Nous considérons le problème d'estimation d'une fonctionnelle de la densité spectrale d'un champ gaussien. Nous établissons tout d'abord des bornes inférieures asymptotiques pour le risque minimax. Dans le cas des champs de dimension 2 ou 3, nous construisons ensuite des estimateurs efficaces, c'est à dire atteignant asymptotiquement ces bornes inférieures, pour des fonctionnelles linéaires et des fonctionnelles intégrales non linéaires de la densité spectrale.

AMS Subject Classification. 62G05, 62M40, 62M15.

Received June 7, 1996. Revised September 17, 1998.

1. INTRODUCTION

It is a well known result that if in a parametric setting the log of the likelihood ratio has a certain quadratic expansion in a local asymptotic sense (LAN families) then, for any sequence of estimators, lower bounds to the maximum of a risk function can be found (see, for example, Le Cam *et al.* [18] for a general survey). Here the quadratic expansion is taken with respect to the parameter and only small asymptotic changes in the value of the parameter are considered. A sequence of estimators is efficient if the risk function asymptotically achieves these lower bounds. This was shown by Hajek [14] for a sequence of sufficiently regular i.i.d. observations. Moreover, it has been shown that lower bounds may also be obtained in infinite dimensional settings (see for example Levit [19], Millar [22], Donoho *et al.* [10] and Birgé *et al.* [3]). Also, the sequence of observations need not to be i.i.d. provided it has a certain regularity (see for example Efroimovich [12]).

Our main interest in this article is to find lower bounds for the maximum of the risk function of estimators of sufficiently regular functionals of the spectral density of Gaussian fields. This is an infinite dimensional problem based on dependent observations. However, the non parametric problem can be turned into a parametric one when one approximates it in terms of a first order expansion on the value of the functional. Thus, if Gaussian fields with sufficiently regular spectral densities are considered, then the classical LAN theory may be applied in order to obtain lower bounds. This is what was done by Khas'minskii *et al.* [15] in the one-dimensional case.

Keywords and phrases: Efficient estimation, Gaussian fields, periodogram, tapered periodogram, spectral density, Toeplitz matrices.

¹ Departamento de Matemáticas, IVIC, Caracas, Venezuela; e-mail: cludena@ivic.ivic.ve

It turns out that the regularity conditions on the spectral density that ensure this LAN behavior are related to the asymptotic behavior of the trace of products of certain Toeplitz matrices and their inverses. In Lemma 1 we study this asymptotic behavior under L^2 type conditions. This Lemma is based on a result due to Doukhan *et al.* [11] (see also Avram [1]).

In order to obtain lower bounds that give uniform results over a class of functions, an additional precompactness condition must be introduced. For this it seems necessary to require uniform bounds (in norm) on the functions and their inverses. Let us also mention that the same kind of restrictions appear when a Banach algebra approach is used to deal with this type of problems (see Bouaziz [4]).

Once lower bounds are given, it is necessary to construct efficient estimators to verify that these bounds are indeed attainable. Periodogram based estimators are biased for dimensions greater than or equal to two (see Guyon [13]). Hence, weighted versions of the periodogram must be considered instead. In the case of linear functionals it is possible to construct efficient estimators based on the so called unbiased periodogram. In this case the bias depends on the regularity of the convolution of the spectral density with the function defining the linear functional. The estimator will be efficient if the linear functional is sufficiently smooth. The unbiased periodogram, however, is not positive definite. We deal with this difficulty by considering tapered periodograms (see Dahlhaus [6] and Yao [23]). In this case, restrictions on the dimension of the problem must be introduced.

For non-linear integral functionals, estimators are constructed using a first order expansion of the functional. In order to control the bias, a minimal equicontinuity condition in L_∞ is required of the spectral density. We also indicate how these results can be extended to a more general class of functionals provided one can explicitly control the bilinear form defining the second derivative.

This LAN property for stationary Gaussian fields and the construction of efficient estimators were previously considered in Ludeña [20], under more restrictive conditions on the regularity of the spectral density.

The article is organized as follows. In Section 2 we give the main result which deals with minimax type bounds for risk functions. In Section 3 we construct efficient estimators for linear functionals of the spectral density. In Section 4 we consider efficient estimators of smooth integral non-linear functionals and in Section 5 we include the proofs of the stated results.

2. LOWER BOUNDS

Let $X = \{X_t, t \in \mathbb{Z}^d\}$ be a centered, real, Gaussian stationary field (process if $d = 1$) with spectral density $f_0 \in L^2(\mathcal{T})$ with $\mathcal{T} = (-\pi, \pi]^d$ identified with the d -torus.

For $k, j \in \mathbb{Z}^d$ denote the usual scalar product by $\langle k, j \rangle$. For $n \in \mathbb{Z}^d$ consider L_n the rectangular lattice in \mathbb{Z}^d with vertices (i_1, \dots, i_d) $0 \leq i_j \leq n_j, j = 1, \dots, d$. The cardinality of this subset is $|L_n| = \prod_{i=1}^d n_i$. We assume, for the asymptotic results, that all $n_i, i = 1, \dots, d$ tend to infinity at the same rate, *i.e.*, if $m_n = \min\{n_i\}$ and $M_n = \max\{n_i\}$, then $m_n^{-1}M_n \rightarrow c$, a positive constant. By $L_n - s$ we mean the set of points $t \in \mathbb{Z}^d$ such that $t + s \in L_n$. Finally, for any given function $g \in L^1(\mathcal{T})$, we denote by g_s its s^{th} Fourier coefficient.

Now consider a certain fixed enumeration τ of the elements of L_n . Define, for a given function $g \in L^1(\mathcal{T})$, $T_n^\tau(g)$, the ‘‘Toeplitz’’ matrix of size $|L_n|$ associated to τ , as the $|L_n| \times |L_n|$ matrix whose term (m, l) is given by

$$g_{\tau^{-1}(m) - \tau^{-1}(l)} = \int_{\mathcal{T}} g(\lambda) e^{i\lambda \langle \tau^{-1}(m), \tau^{-1}(l) \rangle} d\lambda.$$

If $d = 1$ this is the usual Toeplitz matrix, and if τ is a row or column enumeration, this is the usual block-Toeplitz matrix. In most cases we shall omit τ in the notation. Indeed, we are mainly interested in the asymptotic behavior of quadratic forms of this kind of matrices or their inverses, and the order of the enumeration (which only changes the order of the eigenvalues) does not alter this behavior.

Let $\mathcal{L}_p, 1 < p \leq \infty$, be the closure of the trigonometric polynomials in the space $L_p(\mathcal{T})$ ($\mathcal{L}_\infty = C$), endowed with the norm $\|\cdot\|_p$.

Assume that we observe X over L_n . We denote the vector of observations by $\mathbf{x}^n = (x_{\tau^{-1}(1)}, \dots, x_{\tau^{-1}(|L_n|)})$, according to the enumeration τ . Let ${}^t\mathbf{x}^n$ stand for the transposed vector. Throughout this article x_s shall denote

the observation taken at site s , regardless of the enumeration τ . Our aim is to estimate a certain functional of the spectral density $\phi(f_0)$ based on these observations.

In order to do this we assume that the (unknown) spectral density f_0 belongs to a closed class $\mathcal{F} \subset L^2(\mathcal{T})$ such that there exist m and M such that

$$\inf_{f \in \mathcal{F}} \inf_{\lambda} f(\lambda) \geq m > 0, \quad (2.1)$$

$$\sup_{f \in \mathcal{F}} \|f\|_2 \leq M < \infty \quad (2.2)$$

and there also exist positive functions w_i satisfying the condition $w_i(h_i) \rightarrow 0$ for $|h_i| \rightarrow 0$, $i = 1, \dots, d$, and

$$\sup_{f \in \mathcal{F}} \|f(\cdot + h) - f(\cdot)\|_2 \leq \sum_{i=1}^d w_i(h_i). \quad (2.3)$$

Consider the class of functionals $\phi : L^2(\mathcal{T}) \rightarrow \mathbb{R}$, which are Gateaux differentiable in an open neighborhood of \mathcal{F} . For fixed f we can identify the derivative of the functional with a function in $L^2(\mathcal{T})$ which we call $D\phi(f)$. Let Φ be the subclass of these functionals that satisfies:

$$\sup_{\phi \in \Phi} \sup_{f \in \mathcal{F}} \|D\phi(f)\| < \infty \quad (2.4)$$

and

$$\inf_{\phi \in \Phi} \inf_{f \in \mathcal{F}} \|D\phi(f)f\| \geq m > 0. \quad (2.5)$$

Here $\|\cdot\|$ stands for the L^2 norm. We assume that $D\phi(f)$ is an even function for all $\phi \in \Phi$ and $f \in \mathcal{F}$.

Now let Φ_n be the set of all estimators ϕ_n of $\phi(f)$ based on \mathbf{x}_n , and let W be the set of all non decreasing symmetric loss functions w over \mathbb{R}^+ , $w(0) = 0$ that satisfy for all $\lambda \in \mathbb{R}$

$$\int_{-\infty}^{+\infty} w(y) \exp\left\{-\frac{1}{2}\lambda^2 y^2\right\} dy < \infty.$$

Our main result is the following

Theorem 2.1. *If conditions (2.1–2.4) and (2.5) are satisfied and if we define for $f \in \mathcal{F}$ and $\phi \in \Phi$*

$$\Delta = \lim_{\delta \rightarrow 0} \liminf_{|L_n| \rightarrow \infty} \inf_{\Phi_n} \sup_{\|g-f\| < \delta} E_g\{w(|L_n|^{1/2}(\phi_n - \phi(g)))\}$$

then,

$$\Delta \geq \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} w\left(\sqrt{2}(2\pi)^{d/2} \|fD\phi(f)\|x\right) e^{-\frac{x^2}{2}} dx. \quad (2.6)$$

This result is a multidimensional extension of an earlier result of Khas'minskii *et al.* [15]. These authors consider (as in Hajek [14]) the problem of efficient estimation in a local asymptotic sense. For completeness, in what follows we recall Hajek's classical definition of LAN families.

Consider the sequence (using Hajek's notation) x^n of observations with a joint distribution given by $P_n(\cdot, \theta)$, where θ belongs to an open set Θ , in R . Let θ^* be the true value of θ . Let us define $P_n = P_n(\cdot, \theta^*)$ and

$P_{n,t} = (\cdot, \theta^* + n^{-1/2}t)$ and let $\frac{dQ}{dP}$ be the Radon-Nikodym derivative of the absolutely continuous part of the measure Q with respect to the measure P . Then given

$$r_n(t, x^n) = \frac{dP_{n,t}}{dP_n} \quad h \in \mathbb{R}$$

and n such that $\theta^* + n^{-1/2}t \in \Theta$, we have

Definition 2.2. $P_{n,t}$ is Locally Asymptotically Normal (LAN) if

$$\log r_n(t, x^n) = t\Delta_{\theta^*} - 1/2t^2\Gamma_{\theta^*} + Z_n(\theta^*, t)$$

where $\Delta_{\theta^*} \xrightarrow{D} N(0, \Gamma_{\theta^*})$ and $Z_n(\theta^*, t) \rightarrow 0$ in P_n probability for all $t \in \mathbb{R}$ and $\Gamma_{\theta^*} > 0$.

Hajek showed in this context the existence of lower bounds for the maximum risk in a local asymptotic sense for any sequence of estimators.

The proof of the existence of lower bounds given by Khas'minskii and Ibragimov is based on the fact that the infinite dimensional problem of estimating in $L^2(\mathcal{T})$ can be transformed into a parametric one, based on the value of the functional in the true spectral density, which satisfies classical LAN conditions.

In order to extend these results to fields, making use of Toeplitz matrices techniques, the following lemma is required.

Lemma 2.3. *If $f, g \in \mathcal{F}$, and if $\|f\|_\infty \leq M$, $\|g\|_\infty \leq M$, then for $p \geq 2$*

$$\text{Tr}[(T_n(f)T_n(g))^p - T_n((fg)^p)] = o(|L_n|M^p) \quad (2.7)$$

and

$$\text{Tr}[(T_n^{-1}(f)T_n(g))^p - T_n\left(\left(\frac{g}{f}\right)^p\right)] = o(|L_n|M^p). \quad (2.8)$$

In equations (2.7) and (2.8), the term $o(|L_n|M^p)$ is uniform in $f, g \in \mathcal{F}$ which satisfy $\|f\|_\infty \leq M$, $\|g\|_\infty \leq M$. Also note that $\frac{1}{|L_n|}\text{Tr}(T_n(h)) = \int_{\mathcal{T}} h(\lambda) d\lambda$ for any integrable function h .

For $d = 1$ the original proof of Theorem 1 in Khas'minskii *et al.* [15] required (2.1) and the existence of $\beta > 1/2$ such that

$$\sup_{f \in \mathcal{F}} \sum_k (1 + |k|^{2\beta}) f_k^2 < \infty.$$

However, Bouaziz [4] has shown using Banach algebra techniques that Lemma 2.3 holds under (2.1) and an alternative set of conditions of the type

$$\sup_{f \in \mathcal{F}} \sum_k (1 + |k|) f_k^2 \leq M < \infty, \text{ and } \sup_{f \in \mathcal{F}} \|f\|_\infty \leq M < \infty. \quad (2.9)$$

Using this result, it can be shown that for $d = 1$ under (2.1) and (2.9) Theorem 2.1 is still true.

3. ESTIMATION OF LINEAR FUNCTIONALS

In this section we consider estimators of linear functionals of the spectral density f which is assumed to belong to the class \mathcal{F} described above. We are interested in estimators which are efficient for the quadratic loss function in the sense described in the previous section, namely, estimators which satisfy a CLT with $|L_n|^{1/2}$ rate, and have minimum variance.

By the Riesz Representation Theorem the continuous linear on $L^2(\mathcal{T})$ functional ϕ can be identified with a $L^2(\mathcal{T})$ function $\phi(\lambda)$, such that $\phi(f) = \int_{\mathcal{T}} \phi(\lambda) f(\lambda) d\lambda$. Clearly, $D\phi(f)(\lambda) = \phi(\lambda)$ for all $f \in \mathcal{F}$. In what follows we shall use this integral notation which we hope will not be confusing.

Consider the following estimator of $\phi(f)$,

$$I_n(\phi) = \int_{\mathcal{T}} \phi(\lambda) I_n(\lambda) d\lambda$$

where

$$I_n(\lambda) = \frac{1}{n(2\pi)^d} \sum_{s \in L_n} e^{i\langle s, \lambda \rangle} \sum_{t \in L_n - s} x_s x_{s+t},$$

is the usual periodogram.

It was shown by Guyon [13] that the bias of the periodogram cannot be controlled at the CLT rates for dimensions greater than one. He proposed estimators based on the unbiased periodogram defined as

$$I_n^*(\phi) = \int_{\mathcal{T}} \phi(\lambda) I_n^*(\lambda) d\lambda.$$

Here

$$I_n^*(\lambda) = \frac{1}{2\pi^d} \sum_{s \in L_n} \frac{e^{i\langle s, \lambda \rangle}}{c_s} \sum_{t \in L_n - s} x_s x_{s+t}$$

with

$$c_s = \prod_{k \in L_n} (n_k - |s_k|).$$

Assume $\phi(\lambda)$ belongs to a class Φ , such that

$$\sup_{\phi \in \Phi} \sum_{s \in \mathbb{Z}^d} \phi_s^2 < \infty \text{ and} \tag{3.1}$$

$$\sup_{\phi \in \Phi} \sum_{s \in \mathbb{Z}^d} |s|^d |\phi_s| < \infty. \tag{3.2}$$

We then have

Theorem 3.1. *Assume that \mathcal{F} and Φ satisfy conditions (2.3, 2.5) of Section 2 and conditions (3.1) and (3.2) then*

$$V_n^* = |L_n|^{1/2} (I_n^*(\phi) - EI_n^*(\phi))$$

converges in distribution to a centered Gaussian r.v. with variance $2(2\pi)^d \|f\phi\|^2$ for each $f \in \mathcal{F}$ and $\phi \in \Phi$. Also

$$\sup_{\phi \in \Phi} \sup_{f \in \mathcal{F}} E[|L_n|^{1/2} (I_n^*(\phi) - \phi(f))] = o(1).$$

As a Corollary, $I_n^*(\phi)$ is asymptotically efficient for the quadratic loss function.

The unbiased periodogram is unfortunately not positive definite. One possible solution is to consider estimation based on the tapered periodogram [5, 6, 8, 23]. We assume w.l.o.g. that $n = n \mathbf{1}$, where $\mathbf{1}$ is the vector of ones in \mathbb{Z}^d (L_n is a hypercube of edge length n).

Define

$$h_\epsilon(u) = \begin{cases} u/\epsilon & \text{if } u < \epsilon \\ 1 & \text{if } \epsilon \leq u \leq 1 - \epsilon \\ h_\epsilon(1 - u) & \text{if } u > 1 - \epsilon \end{cases} \quad (3.3)$$

with $\epsilon = n^{-\gamma}$, $\gamma > 0$.

Let $j = (j_1, \dots, j_d)$ and $\lambda = (\lambda_1, \dots, \lambda_d)$. Define

$$\underline{d}^T(\lambda) = \sum_{j \in L_n} \underline{h}_\epsilon(j) x_j e^{-i \langle j, \lambda \rangle} \quad \text{with } \underline{h}_\epsilon(j) = \prod_{l=1}^d h_\epsilon(j_l/n).$$

Now define for each $1 \leq l \leq d$

$$H_r(\lambda_l) = \sum_{j_l=1}^n h_\epsilon^r(j_l/n) e^{-ij_l \lambda_l}$$

and

$$\underline{H}_r(\lambda) = \sum_{j \in L_n} \underline{h}_\epsilon(j) e^{-i \langle j, \lambda \rangle} = \prod_{l=1}^d H_r(\lambda_l).$$

The tapered periodogram is given by

$$I_n^T(\lambda) = \left| \frac{\underline{d}^T(\lambda)}{\underline{H}_2(0)} \right|^2. \quad (3.4)$$

For $\phi \in \Phi$ and $f \in \mathcal{F}$ consider the convolution $t = \phi * f$. Assume that there exist $\beta > 0$ and B such that

$$\sup_{\Phi} \sup_{\mathcal{F}} \sup_{x \in \mathcal{T}} \left| \frac{\partial}{\partial x_k} t(x+h) - \frac{\partial}{\partial x_k} t(x) \right| \leq B \sum_{l=1}^d |h_l|^\beta \quad (3.5)$$

for $k = 1, \dots, d$. We have the following result,

Theorem 3.2. *If \mathcal{F} and Φ satisfy (2.3, 2.5) and (3.5) then for $f \in \mathcal{F}$, $\phi \in \Phi$ and $\gamma < 1 + \frac{1}{\beta} (1 - \frac{d}{2})$ we have*

$$\sup_{\phi \in \Phi} \sup_{f \in \mathcal{F}} |L_n|^{1/2} E(I_n^T(\phi) - \phi(f)) = o(1) \quad (3.6)$$

$$|L_n|^{1/2} (I_n^T(\phi) - E I_n^T(\phi)) \xrightarrow{\mathcal{D}} \mathcal{N}(0, 2(2\pi)^d \int_{\mathcal{T}} \phi^2(\lambda) f^2(\lambda) d\lambda). \quad (3.7)$$

Furthermore if condition (2.2) is satisfied, then for all $z \in \mathbb{R}$,

$$\sup_{\phi \in \Phi} \sup_{f \in \mathcal{F}} \lim_{|L_n| \rightarrow \infty} |P \left\{ (|L_n|^{1/2} (I_n^T(\phi) - \phi(f))) < z \right\} - P \{ \xi < z \} | = 0 \quad (3.8)$$

for $\xi \sim \mathcal{N}(0, 2(2\pi)^d \int_{\mathcal{T}} \phi^2(\lambda) f^2(\lambda) d\lambda)$.

As a result of the uniform convergence in (3.8), $I_n^T(\phi)$ is asymptotically efficient for any $w \in W$.

The condition imposed on γ yields that $d < 2(1 + \beta)$. For $d = 2$ it is enough that $\beta > 0$. For $d = 3$ the condition $\beta > 1/2$ must be satisfied. It is not possible to have $d > 3$.

4. NON LINEAR FUNCTIONALS

The results in this section deal with estimators of integral non linear functionals for Gaussian fields under certain regularity conditions for the spectral density and the non linear functional. The regularity conditions on the spectral density are used to control the bias in the multiparametric setting. There are also certain restrictions that appear in order to control the remainder of the first order expansion of the functional. If higher order expansions are possible, less regular spectral densities could be considered.

Throughout this section we require instead of (2.3) that there exist $\beta > 0$ and B such that

$$\sup_{\mathcal{F}} \sup_{x \in \mathcal{T}} \left| \frac{\partial}{\partial x_k} f(x+h) - \frac{\partial}{\partial x_k} f(x) \right| \leq B \sum_{l=1}^d |h_l|^\beta \quad (4.1)$$

for $k = 1, \dots, d$.

We assume that L_n is a hypercube of edge length n . Also that we can write $\phi(f) = \int_{\mathcal{T}} \phi(f(\lambda), \lambda) d\lambda$ in a rather abusive notation, which we hope will not be confusing. We require $\phi(g)$, $g \in L^2(\mathcal{T})$ to be continuously differentiable in the Fréchet sense. We shall call $D\phi(g)$ the derivative at point g , as before. We remark that for a given g , $D\phi(g) : \mathcal{T} \rightarrow \mathbb{R}$ can be identified with a function of $L^2(\mathcal{T})$. Because of the integral form of the functional, we have that the derivative can be written as $D\phi(g)(x) = \partial/\partial u \phi(u, x)|_{(g(x), x)}$. For $g \in L^2(\mathcal{T})$, the second derivative, $D^2\phi(g)$ can also be identified with a $L^2(\mathcal{T})$ function, which because of the integral form of the functional can again be written as $D^2\phi(g)(x) = \partial^2/\partial u^2 \phi(u, x)|_{(g(x), x)}$. Now, assume there exist k and $0 < \delta < 1$, such that

$$\left\| \frac{\partial^2}{\partial u^2} \phi \right\|_\infty < k \text{ and } \left\| \frac{\partial^2}{\partial u^2} \phi(u_1, \cdot) - \frac{\partial^2}{\partial u^2} \phi(u_2, \cdot) \right\|_\infty \leq k |u_1 - u_2|^\delta. \quad (4.2)$$

Actually, we deal with integral functionals in order to control the bilinear form which defines the second derivative of the functional (in order to control the remainder of the Taylor series). More generally, we could consider functionals of the spectral density $\phi(f)$ such that the bilinear form which defines its second derivative can be written as

$$D^2\phi(f)(h, h) = \int_{\mathcal{T}} \int_{\mathcal{T}} K_{\phi(f)}(x, y) h(y) h(x) dx dy \quad (4.3)$$

for any function $h \in L^2$ and such that if $h \in C(\mathcal{T})$, then

$$\sup_{f \in \mathcal{F}} \left\| \int_{\mathcal{T}} K_{\phi(f)}(\cdot, y) h(y) dy \right\|_\infty < M \|h\|_\infty.$$

Assuming additionally the Hölder condition for the second derivative

$$\|(D^2\phi(f) - D^2\phi(g))(h, h)\| \leq k \|f - g\|^\delta \|h\|^2,$$

it is not difficult to show that in this more general setting the proof of Lemma 6 (Sect. 5) is essentially the same. We have preferred to give the proof only for integral functionals for simplicity in the notation. We remark that condition (4.3) is trivially true for functionals for which the bilinear form of its second derivative is defined by a bounded kernel $K_{\phi(f)}(x, y)$. In the case of integral functionals we would formally have $D^2\phi(f)(x, y) = D^2\phi(f)(x)\delta_0(x - y)$, where δ_0 is the usual Dirac delta.

Let $K : \mathbb{R} \rightarrow \mathbb{R}$ be a symmetric, non negative and bounded density function supported over \mathcal{T} . Then define $\underline{K}(u) : \mathbb{R}^d \rightarrow \mathbb{R}$ as

$$\underline{K}(u) = \prod_{l=1}^d K(u_l).$$

Consider the following kernel estimator of f ,

$$f_h(x) = \frac{1}{h^d} \int_{\mathcal{T}} \underline{K}\left(\frac{u-x}{h}\right) I_n^T(u) du, \quad (4.4)$$

$h = n^{-\zeta}$ and I_n^T defined in (3.4).

Based on this pointwise estimator, we construct an estimator of $\phi(f)$, improving the evaluation of $\phi(f_h)$ with a linear term based on the derivative of the functional. Define,

$$\phi_n(f) = \phi(f_h) + \int_{\mathcal{T}} D\phi(f_h)(\lambda)(I_n^T - f_h)(\lambda) d\lambda. \quad (4.5)$$

We have the following result,

Theorem 4.1. *Assume that \mathcal{F} and Φ satisfy (2.4, 2.5, 4.1) and (4.2). Assume that γ defined in (3.3) and ζ defined in (4.4) satisfy*

$$\gamma < 1 + \frac{1}{\beta} \left(1 - \frac{d}{2(1+\delta)}\right) \quad (4.6)$$

$$\frac{d}{2(1+\delta)(1+\beta)} < \zeta < \frac{2}{5} \quad (4.7)$$

and that β defined in (4.1) and δ defined in (4.2) satisfy

$$\frac{d}{2} < 2(1+\delta)(1+\beta)/5. \quad (4.8)$$

Then for $f \in \mathcal{F}$ and $\phi \in \Phi$,

$$\sup_{\phi \in \Phi} \sup_{f \in \mathcal{F}} |L_n|^{1/2} E(\phi_n(f) - \phi(f)) = o(1) \text{ and} \quad (4.9)$$

$$|L_n|^{1/2} (\phi_n(f) - E\phi_n(f)) \xrightarrow{\mathcal{D}} \mathcal{N}(0, 2(2\pi)^d \int_{\mathcal{T}} (D\phi(f)f)^2(\lambda) d\lambda). \quad (4.10)$$

Remark 4.2. As a consequence of the above set of restrictions we have $d \leq 3$ as in Theorem 3.2. For $d = 2$, β has to be greater than $1/4$ and for $d = 3$, $\beta > 7/8$ (if $\delta = 1$).

Remark 4.3. For $d = 1$ the proof of Theorem 4.1 yields that the most severe restriction on the regularity of the spectral density is given by the error of a first order expansion. This error is given by a term of the type $\|f - f_h\|^2$, with f_h defined in equation (4.4) (as $d = 1$ we may use $\gamma = 1$ in Eq. (3.3)). This suggests the convenience of studying estimators based on a second order expansion of the nonlinear functional ϕ .

As a corollary of Theorem 4.1 it follows from the results in Section 3 that $\phi_n(f)$ is an asymptotically efficient estimator of $\phi(f)$ for the quadratic loss function.

Requiring from ϕ to be Fréchet differentiable on the entire L^2 can be a restrictive condition for certain interesting examples. Consider, for example, the functionals ϕ_B and ϕ_S , corresponding to the Burg entropy

and the Shannon entropy respectively, which are not Fréchet differentiable on the entire L^2 but only are so on the set \mathcal{F} of spectral densities satisfying (2.1). In both cases a slight modification to the proof of Theorem 4.1 yields the asymptotic minimum variance Gaussian behavior of the estimator defined in Theorem 4.1. Consider the Burg entropy functional

$$\phi_B(f) = \int_{\mathcal{T}} \log f(\lambda) d\lambda$$

and the Shannon entropy functional

$$\phi_S(f) = \int_{\mathcal{T}} f(\lambda) \log f(\lambda) d\lambda.$$

Then

$$D\phi_B(f)(\lambda) = \frac{1}{f(\lambda)}$$

which is well defined and bounded over \mathcal{F} . Consider the truncated functional

$$\phi_B^N(f) = \int_{\mathcal{T}} \log(f(\lambda) \vee \frac{1}{N}) d\lambda.$$

It follows that ϕ_B^N satisfies

$$\|D^2\phi_B^N(f) - D^2\phi_B^N(g)\| \leq CN^2\|f\|_{\infty}\|g\|_{\infty}\|f - g\| \quad (4.11)$$

with C independent of N . Considering the similar truncation $\phi_S^N(f)$ of $\phi_S(f)$, we have the following corollary of Theorem 4.1

Corollary 4.4. *If \mathcal{F} satisfies (2.1) and (4.1) then there exists $\eta = \eta(\delta, \gamma, \zeta, \beta)$, $\eta > 0$ such that for $N = O(n^\eta)$*

$$n^{d/2}E(\phi_B^N(I_n^T) - \phi_B(f)) = o(1) \quad (4.12)$$

$$n^{d/2}(\phi_B^N(I_n^T) - E\phi_B^N(I_n^T)) \xrightarrow{\mathcal{D}} \mathcal{N}(0, 2(2\pi)^{2d}) \quad (4.13)$$

and

$$n^{d/2}E(\phi_S^N(I_n^T) - \phi_S(f)) = o(1), \quad (4.14)$$

$$n^{d/2}(\phi_S^N(I_n^T) - E\phi_S^N(I_n^T)) \xrightarrow{\mathcal{D}} \mathcal{N}(0, 2(2\pi)^d \int_{\mathcal{T}} [\log(f(\lambda)) + 1]^2 f^2(\lambda) d\lambda). \quad (4.15)$$

5. PROOFS

Proof of Lemma 2.3: In order to prove this Lemma we will follow the ideas of the proof of Theorem 1 in Avram [1] and those of Theorem 5.1 in Dahlhaus [7] on the convergence of the trace of products of Toeplitz matrices and their inverses.

We also need a result obtained by Doukhan *et al.* [11]. Actually, they deal with tapered Toeplitz matrices (tapering functions are introduced in Sect. 3), so that non tapered matrices are just a particular case. Let $T_n(f)$ be the Toeplitz matrix associated with a function f . We have that $T_n(f)$ is a Hermitian matrix, and if f is real valued and even, then $T_n(f)$ is real and symmetric.

Let A be a linear operator on a finite-dimensional euclidean space K of dimension d_K . Assume A is symmetric, and let λ_j , $j = 1, \dots, d_K$ be its eigenvalues. The p -Schatten norms (for symmetric operators) are defined as follows:

$$\begin{aligned} \|A\|_p &= \left[\sum_{j=1}^{d_K} (\lambda_j)^p \right]^{1/p} \text{ for } 1 \leq p < \infty \\ \|A\|_\infty &= \max_{1 \leq j \leq d_K} \lambda_j \text{ for } p = \infty. \end{aligned}$$

We have the following result,

Proposition 5.1. (*Doukhan et al. [11]*) *Let f belong to $L_p(\mathcal{T})$, $1 \leq p \leq \infty$. Then,*

$$\|T_n(f)\|_p \leq |L_n|^{1/p} \|f\|_p. \quad (5.1)$$

Proof of (2.7): Actually (2.7) is a particular case of the more general result,

Lemma 5.2. *Let $\{f^k\}_{1 \leq k \leq p}$ be a collection of functions that satisfy the conditions of Lemma 1. Then,*

$$\frac{1}{|L_n|} \text{Tr} \left[\prod_{k=1}^p T_n(f^k) \right] - \int_{\mathcal{T}} \prod_{k=1}^p f^k(x) dx = o(M^p). \quad (5.2)$$

Following Avram [1], the proof of (5.2) is done by induction on the number $s \leq p$ of those f^k which are non-polynomials. First assume f^k , $k = 1, \dots, p$ are polynomials ($s=0$), then (5.2) holds by multilinearity. Now, for $l \in \mathbb{Z}^d$ let $F_l * f^k$ be the Féjer polynomial of f^k . By the induction hypothesis,

$$\frac{1}{|L_n|} \text{Tr} \left[T_n(F_l * f^1) \prod_{k=2}^s T_n(f^k) \right] - \int_{\mathcal{T}} F_l * f(x) \prod_{k=2}^s f^k(x) dx = o(M^p). \quad (5.3)$$

Consider the first term of the r.h.s. of (5.3). Given $\varepsilon > 0$, there exists $l \in \mathbb{Z}^d$ such that $\|F_l * f - f\|_2 < \varepsilon$ for all $f \in \mathcal{F}$. Then it follows by Proposition 5.1 that

$$\begin{aligned} \frac{1}{|L_n|} \text{Tr} \left[T_n(F_l * f^1 - f^1) \prod_{k=2}^s T_n(f^k) \right] &\leq \frac{1}{|L_n|} \left\| \left[T_n(F_l * f^1 - f^1) \prod_{k=2}^s T_n(f^k) \right] \right\|_1 \\ &\leq \frac{1}{|L_n|} \|T_n(F_l * f^1 - f^1)\|_2 \left\| \prod_{k=2}^s T_n(f^k) \right\|_2 \leq \varepsilon (2\pi)^{d(p-1)} M^{p-1}. \end{aligned}$$

In our case we have used $p_k = \infty$ because the functions are bounded, and have approximated by the Féjer polynomials in L_2 . As a consequence of the equicontinuity condition in (2.3), since \mathcal{F} is a closed, uniformly bounded (by condition (2.2)) and equicontinuous class of functions from L^2 , by the Fréchet-Kolmogorov theorem, there exists a uniform bound over all this class.

Similarly, the second term of the r.h.s. of equation (5.3) tends to $\int_{\mathcal{T}} f^1(x) \prod_{k=2}^s f^k(x) dx$ at the stated rates, for all $f \in \mathcal{F}$. \square

Proof of (2.8): For the proof of the second part of Lemma 2.3 we shall use a less general version of Theorem 5.1 in Dahlhaus [7]. In our case, we require the uniform continuity condition given by (2.3).

Lemma 5.3. *Assume that f belongs to \mathcal{F} where the latter satisfies (2.1) and (2.3). Assume that $g \in L_\infty$. Then*

$$|\text{Tr}[(T_n^{-1}(f)T_n(g))^p - T_n\left(\left(\frac{g}{f}\right)^p\right)]| \leq o(|L_n|). \quad (5.4)$$

Proof of Lemma 5.3: We follow the same line of proof as Dahlhaus [7] and shall divide the proof into several lemmas. Assume without loss of generality that L_n is a square of edge length n . Given a matrix A , consider its norm $\|A\|_\infty$. We have

$$\|A\| = \sup_{x \in \mathbb{C}^{nd}} \left(\frac{x' A' A x}{x' x} \right)$$

where A' is the conjugate transpose of A . Then,

Lemma 5.4. *Under conditions (2.1, 2.2) and (2.3) if $\|f\|_\infty < M$ we have*

$$D_\alpha(f) = \|I - T_n(f)^{1/2} T_n(((2\pi)^{2d} f)^{-1}) T_n(f)^{1/2}\|_2^2 = o(|L_n| M m). \quad (5.5)$$

Proof of Lemma 5.4: We have that the l.h.s. of (5.5) is equal to

$$|L_n| - 2\text{Tr}[T_n(f) T_n(((2\pi)^{2d} f)^{-1})] + \text{Tr}[(T_n(f) T_n(((2\pi)^{2d} f)^{-1}))^2].$$

The proof follows from (2.7), using conditions (2.1–2.3) and $\|f\|_\infty < M$. We have used that

$$\int_{\mathcal{T}} f(x) (2\pi)^{2d} f^{-1}(x) dx = \int_{\mathcal{T}} (f(x) (2\pi)^{2d} f^{-1}(x))^2 dx = |L_n|.$$

□

Lemma 5.5. *Let f and g be two positive, even functions with $\|g\|_\infty < M$ and $\|f^{-1}\|_\infty < m$, then*

$$\|T_n(f)^{-1/2} T_n(g)^{1/2}\|_\infty = O(Mm).$$

Proof of Lemma 5.5: As $T_n(f)$ is invertible, we have

$$\|T_n(f)^{-1/2} T_n(g)^{1/2}\|_\infty = \|T_n(g)^{1/2} T_n(f)^{-1/2}\|_\infty = \sup_{|x|=1} \frac{x^t T_n(g) x}{x^t T_n(f) x} \leq \frac{\int_{\mathcal{T}} |g(\lambda)| |\sum_j x_j e^{-ij\lambda}|^2}{\int_{\mathcal{T}} |f(\lambda)| |\sum_j x_j e^{-ij\lambda}|^2} \leq Mm.$$

□
□

Now, continue with the proof of Lemma 5.3 as in the proof of Theorem 5.1 in Dahlhaus [7].

Proof of Theorem 2.1: It is essentially the same as in Khas'minskii *et al.* [15]. Let

$$\tilde{g}_N = \begin{cases} g & \text{if } |g| \leq N \\ 0 & \text{if } |g| > N \end{cases}$$

and let g_N be an approximation to \tilde{g}_N satisfying (2.3). Define

$$\psi_N(\lambda) = (f(\lambda) D\phi(f)(\lambda))_N \| (f D\phi(f))_N \|^2.$$

Let $\theta = \phi(f)$ and

$$f_\tau(\lambda) = f(\lambda) + (\tau - \theta) f_N(\lambda) \psi_N(\lambda).$$

For large enough N , $\|(f D\phi(f))_N\| > 0$.

If $|\tau - \theta|$ is small then $f_\tau(\lambda)$ is a spectral density,

$$\|f_\tau - f\| = |\tau - \theta| \|\psi_N f\| \leq |\tau - \theta| KN \|f\| = O(|\tau - \theta|), \left| \frac{1}{f} - \frac{1}{f_\tau} \right| \leq \frac{KN|\tau - \theta|}{f_\tau} = O(|\tau - \theta|)$$

and

$$\frac{df_\tau}{d\tau} = f_N(\lambda)\psi_N(\lambda)$$

where this last expression does not depend on τ and satisfies (2.3).

Also, because of ϕ 's differentiability

$$\phi(f_\tau) = \phi(f) + (\tau - \theta) \int_{\mathcal{T}} f(\lambda) D\phi(f)(\lambda) \psi_N(\lambda) d\lambda + o(|\tau - \theta|) = \tau + o(|\tau - \theta|)$$

since $\int_{\mathcal{T}} f D\phi(f) \psi_N = 1 + o(1)$.

Estimating $\phi(f)$, when f belongs to the family f_τ is equivalent to estimating τ . Let $P_{n,\tau}$ be the distribution of x^n with spectral density f_τ .

Using results of Davies (the proof can be found in Appendix II in Davis [9]) and Lemma 2.3 we have the following lemma:

Lemma 5.6. *If \mathcal{F} satisfies conditions (2.1) and (2.3), if $\sup_{f \in \mathcal{F}} \|df_\tau/d\tau\|_\infty \leq CN$ and $|\tau - \theta| = O(|L_n|^{-1/2})$, then $P_{n,\tau}$ is LAN with asymptotic variance*

$$\frac{\Gamma_\theta}{2} = \frac{1}{2(2\pi)^d} \int_{\mathcal{T}} f^{-2} \frac{df_\tau}{d\tau} \Big|_{\tau=\theta} \text{ and } \Delta_\theta = P_{n,\theta}^{-1} \frac{1}{2n^{d/2}} \frac{d}{d\tau} P_{n,\tau} \Big|_{\tau=\theta}.$$

The proof of Lemma 5.6 is given at the end of the proof of the theorem.

Following the definition of Δ it is clear that

$$\Delta \geq \lim_{\delta \rightarrow 0} \liminf_{|L_n| \rightarrow \infty} \inf_{\phi_n} \sup_{|\tau - \theta| < \delta} E_{f_\tau} \{w(n^{1/2}(\phi_n - \phi(f_\tau)))\}$$

and now, because of Section 4 in Hajek [14], the above expression is bounded from below by

$$2\pi^{-1/2} \int_{-\infty}^{+\infty} w(y\Gamma_\theta^{-1/2}/2) \exp\{-1/2y^2\} dy = 2\pi^{-1/2} \int_{-\infty}^{+\infty} w\left(y(2\pi)^{d/2}\sqrt{2} \left[\int_{\mathcal{T}} \frac{f_N}{f^2} \psi_N^2 d\lambda\right]^{-1/2}\right) \exp\{-1/2y^2\} dy.$$

Letting $N \rightarrow \infty$ the proposition is thus proved. \square

Proof of Lemma 5.6: Define $\theta = \phi(f_0)$. Let $\tau = \theta + t|L_n|^{-1/2}$, for some $t \in \mathbb{R}$. Call

$$\Delta_\theta = \frac{1}{2} \frac{1}{|L_n|^{1/2}} \frac{dP_{n,\tau}}{d\tau} \Big|_{\tau=\theta} P_{n,\theta}^{-1}. \quad (5.6)$$

From the definition of the multivariate Gaussian density, we have

$$\frac{dP_{n,\tau}}{d\tau} = \frac{d}{d\tau} \int_{\mathbb{R}^{|L_n|}} e^{-i^t x^n y} e^{-1/2^t y \Sigma_\tau y} dy \quad (5.7)$$

where

$$(\Sigma_\tau)_{kj} = \int_{\mathcal{T}} f_\tau e^{i(j-k)\lambda} d\lambda.$$

In the multidimensional case, k and j are multi-indices so that matrix Σ may be considered (w.l.o.g.) by columns. That is, $(\Sigma)_{kj}$ refers to the term in column $m = (j_1 - 1)n + (j_2 - 1)n + \dots + j_d$ and row $l = (k_1 - 1)n + (k_2 - 1)n + \dots + k_d$.

Since

$$\int_{\mathbb{R}^{|L_n|}} y_k y_j e^{-i^t x^n y} e^{-1/2^t y \Sigma_\tau y} dy = \frac{\partial^2}{\partial x_j \partial x_k} \int_{\mathbb{R}^{|L_n|}} e^{-i^t x^n y} e^{-1/2^t y \Sigma_\tau y} dy,$$

we have

$$\frac{dP_{n,\tau}}{d\tau} = P_{n,\tau} \sum_{k \in L_n} \sum_{j \in L_n} \int_{\mathcal{T}} \frac{df_{\tau}}{d\tau} e^{i\langle k-j, \lambda \rangle} d\lambda \left[-a_{jk} + \sum_{l,m} x_l x_m a_{lk} a_{mj} \right] \quad (5.8)$$

where $\Sigma_{\tau}^{-1} = (a_{jk})$. Set

$$(\Sigma'_{\tau})_{kj} = \int_{\mathcal{T}} \frac{df_{\tau}}{d\tau} e^{i(j-k)\lambda} d\lambda.$$

With this notation,

$$\Delta_{\theta} = \frac{1}{2} \frac{1}{|L_n|^{1/2}} \left[{}^t \mathbf{x}^n \Sigma_{\theta}^{-1} \Sigma'_{\theta} \Sigma_{\theta}^{-1} \mathbf{x}^n - \text{Tr}(\Sigma_{\theta}^{-1} \Sigma'_{\theta}) \right]. \quad (5.9)$$

We have,

1. $E(\Delta_{\theta}) = 0$.
2. $\text{Var}(\Delta_{\theta}) = \frac{1}{2} \frac{1}{|L_n|} \sum \varsigma_i^2$ where ς_i are the eigenvalues of $\Sigma_{\theta}^{-1} \Sigma'_{\theta}$.
3. $\text{cum}_m(\Delta_{\theta}) = \frac{1}{2} \frac{1}{|L_n|^{m/2}} \text{Tr}(\Sigma_{\theta}^{-1} \Sigma'_{\theta})^m$.

In order to prove the first part of the lemma, we must show that

$$\text{Var}(\Delta_{\theta}) \rightarrow \frac{\Gamma_{\theta}}{2} = \frac{1}{2(2\pi)^d} \int_{\pi} f^{-2} \frac{df_{\tau}}{d\tau} \Big|_{\tau=\theta} \quad (5.10)$$

and that

$$\lim_{|L_n| \rightarrow \infty} \text{cum}_m(\Delta_{\theta}) = 0. \quad (5.11)$$

In order to complete the proof, we must verify that the sum of remaining terms of the logarithm of the likelihood ratio minus Δ_{θ} tends to zero in probability. The proof can be found in Davis [9] (*cf.* Th. 4.2). We shall give a sketch.

Set

$$A_n = 1/2 \left[\log \det(\Sigma_{\theta}) - \log \det(\Sigma_{\tau}) + {}^t \mathbf{x}^n \Sigma_{\theta}^{-1} \mathbf{x}^n - {}^t \mathbf{x}^n \Sigma_{\tau}^{-1} \mathbf{x}^n - t |L_n|^{-1/2} [{}^t \mathbf{x}^n \Sigma_{\theta}^{-1} \Sigma'_{\theta} \Sigma_{\theta}^{-1} \mathbf{x}^n - \text{Tr}(\Sigma_{\theta}^{-1} \Sigma'_{\theta})] + 1/2 t^2 \Gamma_{\theta} \right]. \quad (5.12)$$

We want to show that A_n tends to zero in probability. It will be shown that both the expectation and the variance tend to zero. First, we calculate the expectation.

$$E(A_n) = 1/2 \left[\log \det(\Sigma_{\theta}) - \log \det(\Sigma_{\tau}) + \text{Tr}(I - \Sigma_{\theta} \Sigma_{\tau}^{-1}) + \frac{t^2}{2|L_n|} \text{Tr}(I - \Sigma_{\theta}'^2 \Sigma_{\theta}^{-2}) \right].$$

From linear algebra it is known that,

$$\|T_n(f)\|_{\infty} \leq \|f\|_{\infty} \text{ and } \|T_n(f)^{-1}\|_{\infty} \leq \|f^{-1}\|_{\infty}, \quad (5.13)$$

$$\text{Tr}(AB) \leq \|A\|_2 \|B\|_2, \quad (5.14)$$

$$\|AB\|_2 \leq \|A\|_2 \|B\|_{\infty}, \quad (5.15)$$

$$\|A\|_{\infty} \leq \|A\|_2 \leq |L_n|^{1/2} \|A\|_{\infty}. \quad (5.16)$$

Also if $\|A\|_\infty < 1$ then

$$\log |I + A| - \text{Tr}(A) + \frac{1}{2}\text{Tr}(A^2) \leq \frac{1}{3}\|A\|_\infty \|A\|_2^2 / (1 - \|A\|_\infty)^3 \quad (5.17)$$

and if $A = A(\theta)$ and $B = B(\phi)$ are differentiable, then

$$\text{Tr}(A(\theta)B(\phi)) = \text{Tr} \left(\left[A(0) + \theta \frac{d}{d\theta} A(\theta_1) \right] \left[B(0) + \phi \frac{d}{d\phi} B(\phi_1) \right] \right) \quad (5.18)$$

for $0 < \theta_1 < \theta$ and $0 < \phi_1 < \phi$.

Call $A = \Sigma_\theta \Sigma_\tau^{-1} - I$. We have

$$E(A_n) \leq 1/6 \|A\|_\infty \|A\|_2^2 (1 - \|A\|_\infty)^3 + 1/4 [t^2 \text{Tr}(\Sigma_\theta^{-2} (\Sigma'_\theta)^2) - \text{Tr}(A^2)]$$

since

$$\|A\|_\infty \leq \|\Sigma_\theta - \Sigma_\tau\|_\infty \|\Sigma_\tau^{-1}\|_\infty \leq Kt |L_n|^{-1/2} < 1 \quad (5.19)$$

by the first expression in (5.13). Also, $\|f_\tau - f\|_\infty \leq KNt |L_n|^{-1/2}$ and $\|\Sigma_\tau^{-1}\|_\infty < \infty$ again by (5.13).

Making use of the bounds obtained for $\|A\|_\infty$ and the inequality stated in (5.15), it follows that the first summand in $E(A_n)$ is $O(|L_n|^{-1/2})$. Defining $A(\tau) = B(\tau) = \Sigma_\tau^{-1}(\Sigma_\theta - \Sigma_\tau)$, using the relations $|\tau - \theta| = t|L_n|^{-1/2}$, $A(\theta) = 0$ and $\frac{d}{d\tau} A(\tau) = \Sigma_\tau^{-1} \Sigma'_\tau$, in view of equation (5.18) we have that the second term is equal to

$$1/4 \frac{t^2}{|L_n|} \text{Tr}(\Sigma_\theta^{-2} (\Sigma'_\theta)^2 - \Sigma_{\theta_1}^{-1} \Sigma'_\theta \Sigma_{\theta_2}^{-1} \Sigma'_\theta). \quad (5.20)$$

Again by (5.13) we have that

$$\|\Sigma_\theta^{-1} - \Sigma_{\theta_1}^{-1}\|_\infty \leq |f^{-1} - f_{\theta_1}^{-1}| = O(|L_n|^{-1/2}). \quad (5.21)$$

Rewrite (5.20) as

$$1/4 \frac{t^2}{|L_n|} \text{Tr}(\Sigma_\theta^{-2} (\Sigma'_\theta)^2 - \Sigma_{\theta_1}^{-1} \Sigma'_\theta \Sigma_{\theta_1}^{-1} \Sigma'_\theta) + 1/4 \frac{t^2}{|L_n|} \text{Tr}(\Sigma_{\theta_1}^{-1} \Sigma'_\theta \Sigma_{\theta_2}^{-1} \Sigma'_\theta - \Sigma_{\theta_1}^{-1} \Sigma'_\theta \Sigma_{\theta_2}^{-1} \Sigma'_\theta). \quad (5.22)$$

The first term in (5.22), following (5.14, 5.16, 5.17) and (5.21) is bounded by

$$\|\Sigma_\theta\|_\infty^2 \|\Sigma_\theta^{-1}\|_\infty O(|L_n|^{-1/2}). \quad (5.23)$$

By (5.13) this expression is bounded as $\frac{df_\tau(\cdot)}{d\tau}$ and $f^{-1}(\cdot)$ are bounded. The second term in (5.22) can be controlled in an equivalent fashion.

Now consider the variance. We have $\text{Var}(A_n) = 2 \sum_{i \in L_n} \varsigma_i^2$ with ς_i the eigenvalues of

$$\Upsilon_\tau = 1/2 \left[\left(\Sigma_\theta^{-1} \Sigma_\tau^{-1} - \frac{t}{|L_n|^{1/2}} \Sigma_\theta^{-1} \Sigma'_\theta \Sigma_{\theta_1}^{-1} \right) \Sigma_\theta \right].$$

Using the same arguments as for the expectation,

$$\text{Var}(A_n) \leq t^2 \|\Sigma'_\theta\|_\infty \|\Sigma_{\theta_1}^{-1} - \Sigma_\theta^{-1}\|_\infty \|\Sigma_{\theta_2}^{-1} - \Sigma_{\theta_1}^{-1}\|_\infty = O\left(\frac{1}{|L_n|}\right). \quad (5.24)$$

We remark that the above bounds are based on inequalities in (5.13) as applied to the difference $|f - f_\tau|$.

To complete the proof of Lemma 5.6, the convergence in distribution of Δ_θ is a consequence of Lemma 2.3. This is just because its p -cumulants with $p > 2$ are given by

$$2^{p-1}(p-1)! \frac{\text{Tr}(T_n(f)T_n(\phi))^p}{|L_n|^{p/2}}.$$

□

Proof of Theorem 3.1: We have

$$\begin{aligned} E(I_n(\phi)) &= 2\pi^d \sum_{|s_k| < n_k} \frac{c_s \phi_s f_s}{|L_n|} = \frac{1}{|L_n|} \text{Tr}(T_n(\phi)T_n((2\pi)^d f)) \\ &= \frac{1}{|L_n|} \text{Tr}(T_n(\phi(2\pi)^d f)) + O\left(\frac{|L_n|^{1-1/d}}{|L_n|}\right) \\ E(I_n^*(\phi)) &= 2\pi^d \sum_{|s_k| < n_k} \phi_s f_s = 2\pi^d \sum_{|s_k| < n_k} \phi_s f_s + \int_{\mathcal{T}} \phi f - 2\pi^d \sum_s f_s \phi_s \\ &= \int_{\mathcal{T}} \phi f - 2\pi^d \sum_{s \notin \alpha} \phi_s f_s \leq \int_{\mathcal{T}} \frac{2\pi^d}{|L_n|} \sum_s |s| \phi_s f_s = \int_{\mathcal{T}} \phi f + O(|L_n|^{-1}) \end{aligned}$$

where $O(|L_n|^{-1})$ is uniform for $f \in \mathcal{F}$ and $\phi \in \Phi$. Also, we have

$$\text{Var}(I_n(\phi)) = \sum_s \sum_t \frac{\phi_s \phi_t}{|L_n|^2} E \left(\sum_{l \in L_n - s} x_l x_{l+s} \sum_{j \in L_n - t} x_j x_{t+j} - (2\pi)^{2d} c_s c_t f_s f_t \right)$$

and

$$\text{Var}(I_n^*(\phi)) = \sum_s \sum_t \frac{\phi_s \phi_t}{c_s c_t} E \left(\sum_{l \in L_n - s} x_l x_{l+s} \sum_{j \in L_n - t} x_j x_{t+j} - (2\pi)^{2d} c_s c_t f_s f_t \right).$$

It follows that,

$$\begin{aligned} |\text{Var}(I_n^*(\phi)) - \text{Var}(I_n(\phi))| &\leq (2\pi)^{2d} \sum_s \sum_t \left[\frac{1}{c_s c_t} - \frac{1}{|L_n|^2} \right] \phi_s \phi_t K |L_n| \\ &= \frac{(2\pi)^{2d}}{|L_n|} \sum_s \sum_t \phi_s \phi_t \left[\frac{|L_n|^2 - c_s c_t}{c_s c_t} \right] \\ &= \frac{(2\pi)^{2d}}{|L_n|} \sum_s \sum_t \phi_s \phi_t |s|^2 |t|^2 \left[\frac{|L_n|^2 - c_s c_t}{c_s c_t |s|^2 |t|^2} \right] \end{aligned}$$

with

$$\left[\frac{|L_n|^2 - c_s c_t}{c_s c_t |s|^2 |t|^2} \right] = O\left(\frac{1}{m_n}\right).$$

Let τ_k be the eigenvalues of $T_n((2\pi)^d f)T_n(\phi)$. We have

$$\text{Var}(|L_n|^{1/2} I_n(\phi)) = \frac{2}{|L_n|} \sum_{k \in L_n} \tau_k^2.$$

So that it follows from Lemma 2.3 that

$$\text{Var}(|L_n|^{1/2}I_n(\phi)) = \frac{2}{|L_n|} \text{Tr}(T_n((2\pi)^{2d}f^2\phi^2)) + O(m_n^{-1}) = 2(2\pi)^d \int_{\mathcal{T}} f^2\phi^2 + O(m_n^{-1}).$$

Now we shall use the fact that V_n is asymptotically Gaussian to show that V_n^* also is so. Let ϕ_N be a trigonometric polynomial of degree N , $N = (N_1, \dots, N_d)$. In this case

$$I_n(\phi_N) = \sum_{|s_1| \leq N_1} \dots \sum_{|s_d| \leq N_d} \phi_s \hat{r}_s$$

where $s = (s_1, \dots, s_d)$, and

$$\hat{r}_s = \frac{1}{|L_n|} \sum_{l \in L_{n-s}} x_l x_{l+s}, \quad r_s = E(x_l x_{l+s}).$$

Since V_n is Gaussian, then

$$|L_n|^{1/2} \left(\hat{r}_s - \frac{c_s}{|L_n|} r_s \right)$$

is also Gaussian for any s such that $|s_k| \leq N_k$, $k = 1, \dots, d$.

Let

$$\tilde{r}_s = \frac{1}{c_s} \sum_{l \in L_{n-s}} x_l x_{l+s}.$$

If we prove that

$$|L_n|^{1/2}(\tilde{r}_s - r_s) \tag{5.25}$$

is asymptotically Gaussian, then we would have proved this property for $|L_n|^{1/2}V_n^*$ in the case of trigonometric polynomials since

$$|L_n|^{1/2}V_n^*(\phi_N) = |L_n|^{1/2} \sum_{|s_k| \leq N} \phi_s(\tilde{r}_s - r_s),$$

where the r.h.s. is a linear combination of terms given by (5.25).

In order to show the asymptotic normality of (5.25), we notice that we can rewrite this expression as

$$\frac{|L_n|^{3/2}}{c_s} \left(\hat{r}_s - \frac{c_s}{|L_n|} r_s \right)$$

and that $\frac{|L_n|}{c_s}$ tends to one because N is fixed. Thus, because both V_n^* and V_n have equivalent variances, we can write

$$|L_n|^{1/2}V_n^*(\phi_N) \rightarrow \mathcal{N}(0, \sigma_N^2), \quad \sigma_N^2 = \int_{\mathcal{T}} \phi_N^2 f^2$$

or equivalently,

$$E \left(e^{iu|L_n|^{1/2}V_n^*(\phi_N)} \right) = e^{-u^2/2\sigma_N^2}. \tag{5.26}$$

Asymptotic normality in the general case is obtained by approximating $\phi \in L^2(\mathcal{T})$ by trigonometric polynomials, and letting $N \rightarrow \infty$ in (5.25). \square

Proof of Theorem 3.2: The convergence in distribution of the l.h.s. of (3.7) is a consequence of Proposition 5.1, given above, and has been proved in Doukhan *et al.* [11].

To show the convergence to zero of the l.h.s of equation (3.6), we will verify that

$$n^{d/2} E \int_{\mathcal{T}} \phi(\lambda)(I_n^T - f)(\lambda) = o(1), \quad (5.27)$$

and that $o(1)$ is uniform for any $f \in \mathcal{F}$ and $\phi \in \Phi$. To verify (5.27) assume first that $d = 1$. Then

$$\begin{aligned} & \int_{-\pi}^{\pi} \frac{|u|^{\beta+1}}{\sin^2 u/2} \left| \sum_{j=0}^{n-1} h\left(\frac{j-1}{n}\right) - h\left(\frac{j}{n}\right) e^{iuj} \right|^2 du \\ &= \int_{-\pi}^{\pi} \frac{|u|^{\beta+1}}{\sin^2 u/2} \left| \sum_{j=0}^{n^{1-\gamma}} n^{\gamma-1} e^{iuj} + \sum_{j=n^{1-\gamma}}^n n^{\gamma-1} e^{iuj} \right|^2 du \\ &= (n^{\gamma-1})^2 \int_{-\pi}^{\pi} \frac{|u|^{\beta+1}}{\sin^2 u/2} \left| \sum_{j=0}^{n^{1-\gamma}} e^{iuj} + \sum_{j=n^{1-\gamma}}^n e^{iuj} \right|^2 du \\ &= (n^{\gamma-1})^2 \int_{-\pi}^{\pi} \frac{|u|^{\beta+1}}{\sin^2 u/2} \left| \frac{\sin(2n^{1-\gamma}-1)u/2}{\sin u/2} (e^{-i(n-n^{1-\gamma})} + 1) \right|^2 du \leq 4n^{\gamma-1} \int_{-\pi}^{\pi} \frac{|u|^{\beta+1}}{\sin^2 u/2} F_{n^{1-\gamma}}(u) du \end{aligned} \quad (5.28)$$

where $F_n(t)$ is the Féjer kernel. Now we use the fact that the Fourier coefficients of $b(u) = \frac{|u|^{\beta+1}}{\sin^2 u/2}$ are of order $b_k = O(|k|^{-\beta})$. It follows that

$$n^{\gamma-1} \int_{-\pi}^{\pi} \frac{|u|^{\beta+1}}{\sin^2 u/2} F_{n^{1-\gamma}}(u) du = O(n^{\gamma-1}(n^{1-\gamma})^{1-\beta}). \quad (5.29)$$

We have $t(x) - t(0) = x \int_0^1 Dt(\theta x) d\theta$. Thus, the l.h.s. of (5.27) can be written as

$$\begin{aligned} & \frac{1}{H_2(0)} \int_{-\pi}^{\pi} (t(u) - t(0)) \frac{1}{\sin^2 u/2} \left| \sum_{j=0}^{n-1} \left(h\left(\frac{j+1}{n}\right) - h\left(\frac{j}{n}\right) \right) e^{-iuj} \right|^2 du \\ &= \frac{1}{H_2(0)} \int_{-\pi}^{\pi} \frac{u}{\sin^2 u/2} \int_0^1 (Dt(\theta u) - Dt(0)) d\theta \times \left| \sum_{j=0}^{n-1} \left(h\left(\frac{j+1}{n}\right) - h\left(\frac{j}{n}\right) \right) e^{-iuj} \right|^2 du \\ &\leq \frac{K}{H_2(0)} \int_{-\pi}^{\pi} \frac{|u|^{1+\beta}}{\sin^2 u/2} \int_0^1 \theta^\beta d\theta \times \left| \sum_{j=0}^{n-1} \left(h\left(\frac{j+1}{n}\right) - h\left(\frac{j}{n}\right) \right) e^{-iuj} \right|^2 du. \end{aligned} \quad (5.30)$$

The second equality is true because

$$\frac{u}{\sin^2 u/2} \left| \sum_{j=0}^{n-1} \left(h\left(\frac{j+1}{n}\right) - h\left(\frac{j}{n}\right) \right) \exp -iuj \right|^2$$

is an odd function and $Dt(0)$ is a constant, so that we are integrating an odd function over \mathcal{T} . The inequality is a consequence of (3.5). So under the theorem's conditions over β and γ , (5.27) follows from (5.29). Finally the multivariate result is a consequence of the multiplicative structure of $\underline{H}_r(u)$ and the multidimensional Taylor expansion of $t(x)$. Notice that the bounds in equation (5.29) only depend on condition (3.5).

The proof of (3.8) proceeds in the same manner as the proof of Lemma 2.3 in Khas'minskii *et al.* [15]. \square

Proof of Theorem 4.1: Let us first notice that

$$\begin{aligned}\phi_n(f) - \phi(f) &= \phi(f_h) - \phi(f) - \int_{\mathcal{T}} D\phi(f_h)(f_h - f)(\lambda) d\lambda \\ &\quad + \int_{\mathcal{T}} D\phi(f)(I_n^T - f)(\lambda) d\lambda + \int_{\mathcal{T}} (D\phi(f_h) - D\phi(f))(I_n^T - f)(\lambda) d\lambda.\end{aligned}$$

Because of the regularity conditions over ϕ given in (4.2)

$$|\phi(f_h) - \phi(f) - \int_{\mathcal{T}} D\phi(f_h)(f_h - f)(\lambda) d\lambda| \leq \|f_h - f\| \sup_{0 < \theta < 1} \|D\phi(f_h + \theta(f_h - f)) - D\phi(f_h)\| \leq K\|f_h - f\|^2.$$

It then follows that we can write $\phi_n(f) - \phi(f) = A_n + B_n + C_n$ where $|A_n| \leq K\|f_h - f\|^2$ and $n^{d/2}B_n$ is asymptotically Gaussian (for all $f \in \mathcal{F}$ and all $\phi \in \Phi$) as seen in Theorem 3.2 for $d \leq 3$. Note that in this case (3.5) is a consequence of (4.1). On the other hand,

$$\begin{aligned}C_n &= \int_{\mathcal{T}} (D\phi(f_h) - D\phi(f))(\lambda) (I_n^T - f)(\lambda) d\lambda = \int_{\mathcal{T}} D^2\phi(f)(f_h - f)(I_n^T - f)(\lambda) d\lambda \\ &\quad + \int_{\mathcal{T}} (f_h - f)(I_n^T - f)(\lambda) \int_0^1 (D^2\phi(f + \theta(f_h - f)) - D^2\phi(f))(\lambda) d\theta d\lambda = J + L.\end{aligned}$$

Because of (4.2) we have

$$L \leq \int_0^1 \theta^\delta d\theta \|(f_h - f)(I_n^T - f)\| \|f_h - f\|^\delta \leq \|(f_h - f)\|_4 \|(I_n^T - f)\|_4 \|f_h - f\|^\delta.$$

To prove the convergence in distribution of the l.h.s. of equation (4.10), we must verify that $n^{d/2}A_n$ and $n^{d/2}C_n$ tend to zero in probability for all $f \in \mathcal{F}$ and $\phi \in \Phi$.

The proof is completed by the following lemma:

Lemma 5.7. *Assume that \mathcal{F} and Φ satisfy (4.1) and (4.2). Then, under the conditions of Theorem 4.1 for γ and ζ we have that $n^{d/2}A_n$ and $n^{d/2}C_n$ tend to zero in probability for all $f \in \mathcal{F}$ and $\phi \in \Phi$.*

Remark 5.8. If the bilinear form which defines the second derivative of the functional ϕ (see Eq. (4.3)) is defined by a two variable kernel $K_{\phi(f)}$, bounded uniformly for $f \in \mathcal{F}$, the term C_n above is actually bounded by $K\|f_h - f\|^{1+\delta} \|I_n^T - f\|$, which can be controlled in the same manner as $\|f_h - f\|^2$. Then Lemma 5.7 follows using essentially the same arguments. The more general case described in (4.3) can also be controlled in a similar fashion.

To prove Lemma 5.7 we will show that for all $f \in \mathcal{F}$ and $\phi \in \Phi$,

$$n^{d/2}E\|f_h - f\|^2 \rightarrow 0, \tag{5.31}$$

$$E \left(n^{d/2} \int_{\mathcal{T}} D^2\phi(f)(f_h - f)(\lambda)(I_n^T - f)(\lambda) d\lambda \right)^2 \rightarrow 0 \tag{5.32}$$

and

$$n^{d/2}E\|f_h - f\|_4 \|I_n^T - f\|_4 \|f_h - f\|^\delta \rightarrow 0. \tag{5.33}$$

In order to do this we need to give satisfactory bounds for $E\|f_h - f\|_p^p$, with $p = 2, 4$.

It can be shown that (see Dahlhaus [5]):

- For γ as defined above

$$EI_n^T(\lambda) = f(\lambda) + O\left(\frac{1}{n}^{1+\beta(1-\gamma)}\right)$$

•

$$\text{Cov}(I_n^T(\mu), I_n^T(\lambda)) = C_0 \left(\left| \frac{\underline{H}_2(\lambda + \mu)}{\underline{H}_2(0)} \right|^2 + \left| \frac{\underline{H}_2(\lambda - \mu)}{\underline{H}_2(0)} \right|^2 \right) (1 + o(1)).$$

- There exists a family of summable functions $L_s^n(\lambda)$ such that $|\underline{H}_1(\lambda)| \leq KL_0^n(\lambda)$ (see Dahlhaus [5], Eq. (6)); $L_s^n(\lambda) \leq \log^s n L_0^n(\lambda)$ and $\int_{\mathcal{T}} L_p(\lambda - \mu) L_q(-\lambda + \nu) d\lambda \leq L_{p+q+1}(\nu - \mu)$ (see Dahlhaus [5], Lem. 2). Although Dahlhaus' proofs are one-dimensional, the multiplicative structure of functions $\underline{H}_r(\lambda)$ allows for a direct generalization to higher dimensions. In this case n should be changed to n^d and $\log n$ to $\log^d n$.

We are ready to calculate $E\|f_h - f\|^2$. First of all,

$$E\|f_h - f\|^2 \leq C (E\|f_h - Ef_h\|^2 + E\|Ef_h - f\|^2).$$

Call

$$\underline{w}(x) = \frac{1}{h^d} \underline{K}\left(\frac{x}{h}\right).$$

With this notation

$$f_h = \int \underline{w}(x - \mu) I_n^T(\mu) d\mu.$$

Since for small enough h

$$\int \underline{w}(x - \mu) d\mu = 1,$$

then

$$f(x) = \int \underline{w}(x - \mu) f(\mu) d\mu \quad x \in \mathcal{T}.$$

It follows that

$$Ef_h - f = \int \underline{w}(x - \mu) (f(\mu) - f(x)) d\mu + O\left(\frac{1}{n^{1+\beta(1-\gamma)}}\right),$$

and, hence, we obtain

$$\begin{aligned} \|Ef_h - f\|^2 &= \int \int \underline{w}(x - \mu) (f(\mu) - f(x)) d\mu \int \underline{w}(x - \lambda) (f(\lambda) - f(x)) d\lambda dx \\ &+ O\left(\frac{1}{n^{1+\beta(1-\gamma)}}\right) \times \int \underline{w}(x - \lambda) (f(\lambda) - f(x)) d\lambda dx + O\left(\frac{1}{n^{1+\beta(1-\gamma)}}\right). \end{aligned}$$

Expanding $f(\lambda)$ in a Taylor series, by (4.1) we have

$$\begin{aligned} \int \underline{w}(x - \mu)(f(\mu) - f(x))d\mu &= \sum_{l=1}^d \int \underline{w}(x - \mu)(x_l - \mu_l) \\ &\quad \int_0^1 \frac{\partial f}{\partial x_l}(x_l + \theta(x_l - \mu_l)) - \frac{\partial f}{\partial x_l}(x_l) d\theta d\mu \\ &\leq \sum_{l=1}^d B \int w(x_j - \mu_j)(x_j - \mu_j)^{1+\beta} = O(h^{1+\beta}) = O\left(\frac{1}{n^{(1+\beta)\zeta}}\right). \end{aligned} \quad (5.34)$$

The latter bound is proved using the change of variables $z_l = \frac{x_l - \mu_l}{h}$ in each of the above integrals. Thus

$$\|Ef_h - f\|^2 = O\left(\frac{1}{n^{1+\beta(1-\gamma)}} + \frac{1}{n^{(1+\beta)\zeta}}\right)^2$$

or more generally, for the case of p integrals we have

$$\|Ef_h - f\|_p^p = O\left(\left[\frac{1}{n^{1+\beta(1-\gamma)}} + \frac{1}{n^{(1+\beta)\zeta}}\right]^p\right). \quad (5.35)$$

Second,

$$E\|f_h - Ef_h\|^2 = E \int \int \int \underline{w}(x - \mu)\underline{w}(x - \lambda)Cov(I_n^T(\lambda), I_n^T(\mu))d\lambda d\mu dx.$$

We see that

$$\begin{aligned} \int \int w(x - \mu)w(x - \lambda)|H_2(\lambda + \mu)|^2 d\lambda d\mu &= \sum_j \sum_k h^2(j/n)h^2(k/n) \left(\int e^{-i\lambda(j-k)} w(x - \lambda) d\lambda\right)^2 \\ &= \sum_j \sum_k h^2(j/n)h^2(k/n)w_{j-k}^2 e^{-2ix(j-k)}. \end{aligned}$$

It then follows that

$$\int \int \int w(x - \mu)w(x - \lambda)|H_2(\lambda + \mu)|^2 d\lambda d\mu dx = \sum_j h^4(j/n)w_0^2 = O(n).$$

On the other hand,

$$\int \int w(x - \mu)w(x - \lambda)|H_2(\lambda - \mu)|^2 d\lambda d\mu = \sum_j \sum_k h^2(j/n)h^2(k/n)w_{j-k}^2$$

and

$$\int \int \int w(x - \mu)w(x - \lambda)|H_2(\lambda - \mu)|^2 d\lambda d\mu dx = O(n^{d(1+\zeta)}).$$

Factorizing $\underline{w}(\cdot)$ and $\underline{H}_2(\cdot)$ we obtain

$$E\|f_h - Ef_h\|^2 = C_0 \frac{1}{\underline{H}_2^2(0)} O(n^{d(1+\zeta)})(1 + o(1))$$

and since $H_2(0) = O(n)$

$$E\|f_h - Ef_h\|^2 = O\left(\frac{1 + o(1)}{n^{d(1-\zeta)}}\right). \quad (5.36)$$

Third, because of (5.35) we see that in order to study $E\|f_h - f\|_p^p$ for even p , it is enough to consider $E\|f_h - Ef_h\|_p^p$. For this purpose we consider the centered even moments of order greater than or equal to 2 of the tapered periodogram.

We have that

$$\left| \int E(f_h - Ef_h)^p(x) dx \right| \quad (5.37)$$

is bounded by

$$K \int \left| \int_{\mathcal{T}^{pd}} E \prod_{j=1}^p \underline{w}(x - \alpha_j) (I_n^T - EI_n^T)(\alpha_j) d\alpha \right| dx. \quad (5.38)$$

As in the proof of Lemma 7 in Dahlhaus [5], we want to control the expectation of products of the centered periodogram. As the underlying process is Gaussian, the expectation in equation (5.38) is the sum over all possible ways of coupling a vector of size $2p$. From the definition of \underline{d}_n^T , and as the periodogram is centered, this is equivalent to consider all couples $(\alpha_j, -\alpha_k)$ for $1 \leq j < k \leq p$.

In order to formalize this construction, consider the set of all partitions $\mathcal{P} = \{P = (P_1, P_2, \dots, P_p)\}$ with $|P_j| = 2$, $j = 1, \dots, p$ of the table

$$\begin{array}{cc} a_1 & b_1 \\ a_2 & b_2 \\ \vdots & \vdots \\ a_p & b_p. \end{array}$$

Here a_i and b_i stand for the corresponding position of α_i and $-\alpha_i$ in equation (5.38). Define now variables β_{a_i} and β_{b_i} belonging to the set $\mathcal{C} = \{\beta_1, -\beta_1, \dots, \beta_p, -\beta_p\}$, for each partition $P = (P_1, P_2, \dots, P_p)$, taking into account the P_j to which a_i and b_i belong. The sign depends on the order (given by that of the array) of appearance in the corresponding P_j . We have $\sum_{i=1}^p \beta_{a_i} + \beta_{b_i} = 0$. We say we have a subtable if $\sum_{i=q_1}^{q_2} \beta_{a_i} + \beta_{b_i} = 0$, for some $1 \leq q_1 < q_2 \leq p$. Since couples of the form $(\alpha_j, -\alpha_j)$ do not occur, we can decompose the table into at most $p/2$ separate subtables of size 2.

Consider a partition which divides the table into $p/2$ separate subtables. For this partition the corresponding term in equation (5.38) is,

$$n^{-pd} \int \left| \int_{\mathcal{T}^p} \prod_{j=1}^p \underline{w}(x - \alpha_j) \left(\int_{\mathcal{T}^2} \prod_{j=1}^2 f(\beta_{a_j}) \times \prod_{j=1}^2 \underline{H}_1(\alpha_j - \beta_{a_j}) \prod_{j=1}^2 \underline{H}_1(-\alpha_j - \beta_{b_j}) d\beta \right)^{p/2} d\alpha \right| dx. \quad (5.39)$$

Recalling the definition of $\underline{w}(x)$,

$$\int \prod_{j=1}^p |\underline{w}(x - \alpha_j)| dx = O(h^{-(p-1)d}) = O(n^{d\zeta(p-1)}).$$

So, as a consequence of equation (6) and Lemma 2 (a) in Dahlhaus [5] which give bounds for $H(x)$ and its integrals we obtain the upper bound for (5.39)

$$n^{-pd} \left(\int_{\mathcal{T}^2} f(\lambda_1) f(\lambda_2) L_2^T(\lambda_1 - \lambda_2) \right)^{p/2} \times O(n^{d(p-1)\zeta}) = O(n^{-d(p/2-(p-1)\zeta)} \log^2 n). \quad (5.40)$$

In obtaining the latter bound we have used the fact that $f(\lambda)$ is bounded and have taken into account the properties of the tapered periodogram (*cf.* Dahlhaus' Lemma 2 (b)). Notice that (5.36) is just (5.40) when $p = 2$.

It is not difficult to show that an upper bound for each of the terms in equation (5.38) is given when we consider partitions which lead to $p/2$ separate subtables. As the total number of partitions depends only on p , the upper bound found in equation (5.40) yields an upper bound for equation (5.38).

Proof of Lemma 5.7: We have to show the convergence to zero of the l.h.s. of equations (5.31, 5.32) and (5.33). Under (4.6, 4.7) and (4.8) the convergence to zero of the l.h.s. of equation (5.31) follows from (5.35) and (5.36). On the other hand, since $\delta \leq 1$ we have

$$n^{d/2} E \|f_h - f\|_4 \|I_n^T - f\|_4 \|f_h - f\|^\delta \leq n^{d/2} (E \|f_h - f\|_4^4 E \|f_h - f\|^{4\delta})^{1/4} (E \|I_n^T - f\|_4^2)^{1/2}.$$

Now because of (5.35) and (5.40), we have asymptotically

$$\begin{aligned} E \|f_h - f\|_4^4 &= O \left(\frac{1}{n^{d(2-3\zeta)}} + \frac{1}{n^{4(1+\beta)\zeta}} + \frac{1}{n^{2(1+\beta(1-\gamma))}} \right) \\ E \|f_h - f\|^{4\delta} &= O \left(\frac{1}{n^{d(1-\zeta)2\delta}} + \frac{1}{n^{4\delta(1+\beta)\zeta}} + \frac{1}{n^{4\delta(1+\beta(1-\gamma))}} \right). \end{aligned}$$

Under the lemma's assumptions we have

$$\begin{aligned} d(2-3\zeta) + d(1-\zeta)2\delta &> 2d \\ 4(1+\beta)\zeta + 4\delta(1+\beta)\zeta &> 2d \\ 4(1+\beta(1-\gamma)) + 4\delta(1+\beta(1-\gamma)) &> 2d \end{aligned}$$

which asserts that the l.h.s. of equation (5.33) converges to zero provided we show that $E \|I_n^T - f\|_4^4 = O(1)$. Following the same line of proof as that used for equation (5.35) we can show that $\|E I_n^T - f\|_4^4 \rightarrow 0$. Thus it is enough to verify that $E \|I_n^T - E I_n^T\|_4^4$ is bounded. As before, we have that

$$\begin{aligned} E \int_{\mathcal{T}} (I_n^T - E I_n^T)^2(\mu) d\mu &\leq \frac{1}{n^2} \left| \int_{\mathcal{T}} \int_{\mathcal{T}^2} f(\lambda_1) f(\lambda_2) \underline{H}_1(\lambda_1 - \mu) \underline{H}_1(-\lambda_1 + \mu) \right. \\ &\quad \times \left. \underline{H}_1(\lambda_2 - \mu) \underline{H}_1(-\lambda_2 + \mu) d\lambda_1 d\lambda_2 d\mu \right| \\ &= \frac{1}{n^2} \int_{\mathcal{T}} \left(\int_{\mathcal{T}} f(\lambda) |\underline{H}_1(\mu - \lambda)|^2 d\lambda \right)^2 d\mu = O(1). \end{aligned}$$

Finally, we have

$$\begin{aligned}
E \int_{\mathcal{T}} (I_n^T - EI_n^T)^4(\mu) \, d\mu &\leq n^{-4} \left| \int_{\mathcal{T}} \left(\int_{\mathcal{T}^2} f(\lambda_1) f(\lambda_2) \underline{H}_1(\lambda_1 - \mu) \underline{H}_1(-\lambda_1 + \mu) \right. \right. \\
&\quad \left. \left. \times \underline{H}_1(\lambda_2 - \mu) \underline{H}_1(-\lambda_2 + \mu) \, d\lambda_1 d\lambda_2 \right)^2 \, d\mu \right| \\
&= n^{-4d} \int_{\mathcal{T}} \left(\int_{\mathcal{T}} f(\lambda) |\underline{H}_1(\mu - \lambda)|^2 \, d\lambda \right)^4 \, d\mu = O(1).
\end{aligned}$$

Now consider the l.h.s. of equation (5.32). We rewrite this expression as

$$\begin{aligned}
&n^d E \left(\int_{\mathcal{T}} D^2 \phi(f)(I_n^T - EI_n^T)(f_h - Ef_h)(x) \, dx \right)^2 + n^d 2E \int_{\mathcal{T}^2} D^2 \phi(f)(x_1) D^2 \phi(f)(x_2) (EI_n^T - f)(x_1) \\
&\times (I_n^T - EI_n^T)(x_2) (f_h - Ef_h)(x_1) (f_h - Ef_h)(x_2) \, dx + n^d E \int_{\mathcal{T}^2} D^2 \phi(f)(x_1) D^2 \phi(f)(x_2) (EI_n^T - f)(x_1) \\
&\times (EI_n^T - f)(x_2) (f_h - Ef_h)(x_1) (f_h - Ef_h)(x_2) \, dx + n^d 2E \int_{\mathcal{T}^2} D^2 \phi(f)(x_1) D^2 \phi(f)(x_2) (I_n^T - EI_n^T)(x_1) \\
&\times (I_n^T - EI_n^T)(x_2) (Ef_h - f)(x_1) (f_h - Ef_h)(x_2) \, dx + n^d E \int_{\mathcal{T}^2} D^2 \phi(f)(x_1) D^2 \phi(f)(x_2) (I_n^T - EI_n^T)(x_1) \\
&\times (I_n^T - EI_n^T)(x_2) (Ef_h - f)(x_1) (Ef_h - f)(x_2) \, dx + n^d 4E \int_{\mathcal{T}^2} D^2 \phi(f)(x_1) D^2 \phi(f)(x_2) (I_n^T - EI_n^T)(x_1) \\
&\times (EI_n^T - f)(x_2) (Ef_h - f)(x_1) (f_h - Ef_h)(x_2) \, dx + n^d E \int_{\mathcal{T}^2} D^2 \phi(f)(x_1) D^2 \phi(f)(x_2) (EI_n^T - f)(x_1) \\
&\times (EI_n^T - f)(x_2) (Ef_h - f)(x_1) (Ef_h - f)(x_2) \, dx = A_1 + A_2 + A_3 + A_4 + A_5 + A_6 + A_7.
\end{aligned}$$

First, remark

$$\begin{aligned}
&n^d E \left(\int_{\mathcal{T}} D^2 \phi(f)(I_n^T - EI_n^T)(f_h - Ef_h)(x) \, dx \right)^2 \\
&= n^d \int_{\mathcal{T}^4} D^2 \phi(f)(x_2) D^2 \phi(f)(x_4) w(x_1 - x_2) w(x_3 - x_4) \\
&\quad \times E \prod_{j=1}^4 (I_n^T - EI_n^T)(x_j) \, dx \leq O(n^{d(2\zeta-1)}). \tag{5.41}
\end{aligned}$$

On the other hand, as in the proof of the convergence to zero of the l.h.s. of (5.33) it can be shown that for odd p

$$\int_{\mathcal{T}^p} |E(I_n^T - EI_n^T)(x_1) \dots (I_n^T - EI_n^T)(x_p)| \, dx = O(n^{-dp/2}). \tag{5.42}$$

Thus, we have

$$\begin{aligned}
A_2 &= 2n^d E \int_{\mathcal{T}^4} D^2 \phi(f)(x_1) D^2 \phi(f)(x_2) (EI_n^T - f)(x_1) w(x_1 - x_2) \\
&\quad \times w(x_3 - x_4) (I_n^T - EI_n^T)(x_2) (I_n^T - EI_n^T)(x_3) (I_n^T - EI_n^T)(x_4) \\
&\leq n^d \|w\|_\infty \|EI_n^T - f\|_\infty \|D^2 \phi(f)\|_\infty^2 \int_{\mathcal{T}} w(x_1 - x_2) dx_1 \\
&\quad \times \int_{\mathcal{T}^3} |E(I_n^T - EI_n^T)(x_2)(I_n^T - EI_n^T)(x_3)(I_n^T - EI_n^T)(x_4)| \leq O\left(\frac{n^{d(5/2+\zeta)}}{n^{3d+1+\beta(1-\gamma)}}\right).
\end{aligned}$$

The bound in equation (5.43) is a consequence of the properties of the tapered periodogram, the definition of $w(\cdot)$ and equation (5.42).

Using (5.35) and going through essentially the same steps as in the case of A_2 we have,

$$A_4 = O\left(\frac{n^{d(5/2+\zeta)}}{n^{3d+(1+\beta)\zeta}}\right). \quad (5.43)$$

Again using the properties of the tapered periodogram, equation (5.35) and going through the steps of the proof of equation (5.33) we have

$$\begin{aligned}
A_3 &= O(n^{-2(1+\beta(1-\gamma))}) \\
A_5 &= O(n^{-2(1+\beta)\zeta}) \\
A_6 &= O(n^{-(1+\beta(1-\zeta)+(1+\beta)\zeta)}) \\
A_7 &= O(n^{d-2(1+\beta(1-\gamma)+(1+\beta)\zeta)}).
\end{aligned}$$

This completes the proof, under the lemma's restrictions on γ, ζ and β . \square

Proof of Corollary 4.4: The proof of (4.12) follows from

$$E(\phi_B^N(f) - \phi_B(f)) = E(\phi_B^N(f) - \phi_B^N(f)) + (\phi_B^N(f) - \phi_B(f)).$$

The first term is $o(n^{d/2})$ as ϕ_B^N satisfies the conditions of Theorem 4.1. The second term is zero for large enough N , since f is strictly positive. The convergence in distribution stated in (4.13) is a consequence of Theorem 4.1 for fixed N . The result follows by choosing η small enough so that (5.31, 5.32) and (5.33) tend to zero in probability, as $n \rightarrow \infty$. The asymptotic variance in (4.13) follows because $D\phi_B^N = \frac{1}{f\sqrt{N}}$ which tends to $\frac{1}{f}$ as $N \rightarrow \infty$.

The proof of statements (4.14) and (4.15) are obtained in a similar fashion. \square

I would like to thank the anonymous referees, whose very helpful suggestions have improved the presentation considerably.

REFERENCES

- [1] F. Avram, On bilinear forms in Gaussian random variables and Toeplitz matrices. *Prob. Th. Rel. Fields* **79** (1988) 37–45.
- [2] R. Azencott and D. Dacunha-Castelle, Series of Irregular Observations. *Forecasting and Model Building*, Masson Éditions, Paris (1984).
- [3] L. Birgé and P. Massart, Estimation of integral functionals of a density. *Ann. Statist.* **23** (1995) 11–29.
- [4] M. Bouaziz, Inégalités de trace pour des matrices de Toeplitz et applications à des vraisemblances gaussiennes. *Probability Math. Statistics* **13** (1992) 253–267.
- [5] R. Dahlhaus, Spectral analysis with tapered data. *J. Time Ser. Anal.* **4** (1983) 163–175.
- [6] R. Dahlhaus, Parameter estimation of stationary processes with spectra containing strong peaks. Robust and nonlinear time series analysis. J. Franke, W. Härdle and D. Martin Eds. *Lecture Notes in Statistics* **26**, Springer Verlag (1983).

- [7] R. Dahlhaus, Efficient parameter estimation for self similar processes. *Ann. Statist.* **17** (1989) 1749–1766.
- [8] R. Dahlhaus and H. Künsch, Edge effects and efficient parameter estimation for stationary random fields. *Biometrika* **74** (1987) 877–882.
- [9] R. Davies, Asymptotic inference in stationary Gaussian time series. *Adv. Appl. Prob.* **5** (1973) 469–497.
- [10] D. Donoho and R. Liu, Geometrizing rates of convergence II. *Ann. Statist.* **19** (1991) 633–667.
- [11] P. Doukhan, J. León and P. Soulier, Central and non-central limit theorems for quadratic forms of a strongly dependent stationary Gaussian field. *Rebrape* **10** (1996) 205–223.
- [12] S. Yu. Efrimovich, Local asymptotic normality for dependent observations. *Translated from Problemy Pederachi Informatsii* **14** (1978) 73–84.
- [13] X. Guyon, Parameter estimation for a stationary process on a d dimensional lattice. *Biometrika* **69** (1982) 95–105.
- [14] J. Hajek, Local asymptotic minimax and admissibility in estimation. *Sixth Berkeley Symposium* (1972) 175–194.
- [15] R.Z. Khas'minskii and I.A. Ibragimov, Asymptotically efficient nonparametric estimation of functionals of a spectral density function. *Prob. Th. Rel. Fields* **73** (1986) 447–461.
- [16] Yu.A. Koshevnik and B.Ya. Levit, On a nonparametric analogue of the information matrix. *Theor. Prob. Appl.* **21** (1976) 738–759.
- [17] L. Le Cam, On the assumptions used to prove the asymptotic normality of maximum likelihood estimates. *Ann. Math. Statist.* **41** (1970) 802–828.
- [18] L. Le Cam and G. Lo Yang, *Asymptotics in Statistics* Springer Series in Statistics (1990).
- [19] B. Levit, Infinite dimensional information lower bounds. *Theor. Prob. Appl.* **23** (1978) 388–394.
- [20] C. Ludeña, Estimación eficiente de funcionales de la densidad espectral de procesos gaussianos multiparamétricos. *Proceedings IV CLAPEM* **4** (1990) 140–153.
- [21] C. Ludeña, *Estimation des fonctionnelles de la densité spectrale des processus gaussiens dans différents cadres de dépendance*. Thèse (Docteur en Sciences), Université de Paris Sud, Centre d'Orsay, France (1996).
- [22] P.W. Millar, Nonparametric applications of an infinite dimensional convolution theorem. *Z. Wahrsch. verw. Gebiete* **68** (1985) 545–556.
- [23] Y.F. Yao, *Méthodes bayésiennes en segmentation d'image et estimation par rabotage des modèles spatiaux*. Thèse (Docteur en Sciences), Université de Paris Sud, Centre d'Orsay, France (1990).