ON THE SECOND ORDER PROPERTIES OF THE MULTIDIMENSIONAL PERIODOGRAM FOR REGULARLY SPACED DATA

E. PORCU†, R. CRUJEIRAS‡, J. MATEU†, AND W. GONZALEZ-MANTEIGA‡

Abstract. Recently there is growing interest in improving the level of knowledge of spatial and spatio-temporal processes using spectral techniques. The properties of the estimator of the spectral density, the periodogram, have been broadly studied under different asymptotic assumptions that imply a valuable loss of information about the behavior of the underlying process that is often observed on a grid of small size and with sparse data. In this scheme, neither increasing domain nor shrinking asymptotics applies. The goal of this paper is to study the properties of the multidimensional periodogram, under both cases of tapering and no tapering, and the assumption of finite dimensionality of the regular lattice where the process is observed. We present some theoretical results regarding the second order properties of the multidimensional periodogram. Furthermore, we show that, independent of the tapering procedure, periodogram values present a dependence structure which is not stationary and which particularly depends on weights which are proportional to the Bartlett kernel or the chosen taper.

Key words. periodogram, regular lattice data, spatial and spatio-temporal data, spectral density, tapering

DOI. 10.1137/S0040585X97983663

1. Introduction. Characterizing space-time dependence is essential for a better knowledge of random fields varying along space and time. Atmospheric pollutant concentrations, precipitation fields, income distributions, and mortality fields are only a few examples of attributes that characterize space-time uncertainty and need to be modeled through both epistemic and ontologic approaches.

In the last few years, a growing number of authors have contributed to the study of space-time dependence structures. A typical taxonomy can be identified by referring to two main categories of contributions: the former is devoted to new parametric models of spatial or space-time covariance functions, or equivalently new classes of spectral densities. The latter regards a crucial problem, that of estimation of the spatio-temporal dependence structure.

It is this latter aspect in which we are particularly interested in this paper. Estimation of space-time dependence involves several open problems, as yet unsolved questions and dilemmas of practical and theoretical nature. As far as the former are concerned, the duality and perspective of the practical advantages of working in the spectral domain have been emphasized in [13] and [10], where it is argued that the use of the Fast Fourier transform algorithms for estimating spectral densities can be a good alternative if handling a huge and sparse dataset. Spectral methods have been proposed and argued successfully in several subfields of spatial and space-time statistics. In particular, it is worth citing [9], [10], and [7], where spectral estimation techniques are implemented for both stationary and nonstationary settings.
As far as theoretical aspects are concerned, the study of spatial processes by working in the spectral domain has been encouraged by a number of authors, on the base of rigorous mathematical results. In his *tour de force*, Stein [16] emphasizes the importance of the Matérn model for both spatial and spectral analysis, and justifies the effects of misspecification of the spectral density, working in infilling asymptotics. On the other hand, [8] studies the asymptotic properties of the periodogram via shrinking asymptotics, under the assumptions of either stationarity or nonstationarity. Working on fixed-domain asymptotics, Stein [15] showed that standard asymptotic results for the periodograms do not apply, and the use of the raw data may yield misleading results, so that data tapers are needed. The periodogram is an inconsistent estimator of the spectral density (its variance remains proportional to the square of the spectral density at each frequency). This lack of consistency can be overcome via smoothing techniques as in [14].

A number of proposals regarding estimation techniques and tests for spatial and space-time independence are based on standard asymptotic theoretical results regarding the tapered and untapered periodogram: that is, the periodogram is asymptotically unbiased and the periodogram values are asymptotically independent. This allows the use of least squares techniques for estimation or the implementation of several tests for independence of spatial processes.

Nevertheless, in many applications (see [6] and [11]), although working on regular grids, one only disposes of very few data. This key point motivated our research. In particular, a natural question would be how to quantify the magnitude of the multidimensional periodogram bias when working on regular and finite lattices. Another very important aspect would be to assess the dependence between periodogram values when working on finite grids, so as to establish if this dependence can be actually considered as negligible, as supported by the asymptotic theory. Finally, it would be desirable to compare the finite properties of the classic versus tapered periodogram. In this paper, we tackle these problems from a theoretical perspective. To do this, we consider the multidimensional classic and tapered periodograms, and multidimensional processes defined on a lattice, so that the temporal component can be considered as an additional coordinate of a hypercube defined on $\mathbb{Z}^d$. This kind of setting is justified by the fact that we are working with grids of finite dimensions, while for the asymptotic theory one must be cautious as the spatial and the temporal component cannot be considered with the same approach, as emphasized in [5]. This procedure allows us to compare asymptotic with finite results, on the one hand, and finite results between them, on the other hand.

Thus, the plan of the paper is the following. After necessary material about random fields, their spectral representation, and multidimensional periodogram are shown in section 2, in section 3 we focus on the theoretical results that allow us to reach some interesting conclusions about the previously mentioned key points. Section 4 concludes the paper with a critical discussion and opens new points of research for the future.

2. Background and setup. The present section is largely expository and includes material about random fields, their spectral representation and multidimensional periodograms, with and without taper.

For the remainder of the paper, we shall denote a real-valued weakly stationary random field $\{Z(s), \ s \in \mathbb{R}^d\}$ with the condensed notation $Z \sim SRF(\mu, C(\cdot))$, where $\mu$ is the mean of the process, that is constant in the whole domain, and where $C: \mathbb{R}^d \to \mathbb{R}$ is the covariance function defined as

\[ C(u) = \text{cov} (Z(s), Z(s + u)), \quad u, s \in \mathbb{R}^d, \]

and that depends only on the separation or lag vector $u \in \mathbb{R}^d$.

Any weakly stationary process can be represented in the form of a Fourier–Stieltjes integral (see [17]):

\[ Z(s) = \int_{\mathbb{R}^d} e^{i s \cdot \lambda} \, dY(\lambda), \]

(1)
where \( Y \) is a random function with uncorrelated increments, and where \( s' \lambda \) denotes the usual inner scalar product between \( s \) and \( \lambda \). The celebrated Bochner’s theorem [1] establishes a one-to-one correspondence between the class of continuous covariance functions and that of positive definite functions in \( \mathbb{R}^d \). Besides, the positive bounded symmetric measure \( \mathcal{F} : \mathbb{R}^d \to \mathbb{R}_+ \) defined by \( \mathcal{F}(d\lambda) = \mathbb{E}|Y(d\lambda)|^2 \) provides the following representation for the covariance of the process:

\[
(2) \quad C(u) = \int_{\mathbb{R}^d} e^{i\langle u, \lambda \rangle} \mathcal{F}(d\lambda).
\]

Additionally, if \( F \) is absolutely continuous with respect to the Lebesgue measure, then representation (2) can be rewritten in terms of the so-called spectral density \( f \) where the one-dimensional tapers

\[
D_j = \left( \prod_{j=1}^d d\lambda \right) = \int_{\mathbb{R}^d} e^{i\langle u, \lambda \rangle} du, \quad \lambda \in \Pi^d = [-\pi, \pi]^d.
\]

Now, let us assume that \( F \) is observed on a \( d \)-dimensional lattice of \( N = \prod_{j=1}^d n_j \) points \( D = \prod_{j=1}^d \{0, \ldots, n_j-1\} = \prod_{j=1}^d D_j \). In this context, the classical nonparametric estimator of the spectral density is the multidimensional periodogram, which is defined by

\[
(3) \quad I_N(\lambda) = \frac{1}{(2\pi)^d N} \sum_{s \in D} Z(s) e^{-i\lambda's},
\]

The periodogram is usually evaluated at the set of Fourier frequencies \( \lambda = (\lambda_1, \ldots, \lambda_d) \), where each component, for \( j = 1, \ldots, d \), is given by \( \lambda_j = 2\pi k_j n_j^{-1} \), with \( k_j = 0, \pm 1, \ldots, \pm m_j \), \( m_j = [(n_j - 1)/2] \).

It is important to stress that this setting allows us to treat equivalently spatial or spatio-temporal random fields defined on a lattice, as, when working on grids of finite dimension, the temporal component can be treated as a mere additional coordinate. We shall refer to this setting for the subsequent theoretical results.

For what follows, it will be useful to recall that \( I_N(\lambda) = d(\lambda) \overline{d(\lambda)} \), where \( d(\lambda) \) denotes the discrete Fourier transform (DFT) of the data, at a frequency \( \lambda \in \Pi^d \), namely,

\[
d(\lambda) = \frac{1}{(2\pi)^{d/2} \sqrt{N}} \sum_{s \in D} Z(s) e^{-i\lambda's},
\]

and \( \overline{()} \) denotes the complex conjugate.

**Bias reduction via tapering.** It should be stressed that, for a finite sample, the periodogram presents biases due to leakage. This edge-effect bias can be reduced through the use of data tapers [2] that are required to be measurable functions, bounded with bounded support, \( L^2 \)-integrable, and Lipschitz continuous. Examples of one-dimensional data tapers can be seen in [12]. Usually, multidimensional tapers are obtained as the tensor product of one-dimensional tapers, that is,

\[
h(s) = \prod_{j=1}^d h_j(s_j), \quad s \in \mathbb{R}^d, \quad s_j \in \mathbb{R},
\]

where the one-dimensional tapers \( h_j \) may be all equal or different. This is of practical interest in the spatio-temporal context, where \( h \) is usually obtained as the tensor product of a spatial taper, \( h_S \), and a temporal one, \( h_T \). Data tapers usually involve a maximum of \( 1 \) at the origin and then decrease to \( 0 \) as \( \|s\| \) increases.

Consider \( r \in \mathbb{N} \) and denote by \( H_r \) the series expansion

\[
H_r(\lambda) = \sum_{s \in D} h_r(s) e^{-i\lambda's}, \quad \lambda \in \mathbb{R}^d,
\]

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.
which is well defined given that $h$ is of bounded and compact support. By the Riemann–Lebesgue theorem, $H_t(\lambda) \to 0$ as $\|\lambda\|_2 \to \infty$. Consider now the DFT of the tapered observed values:

\[
d^T(\lambda) = \frac{1}{(2\pi)^d/2\sqrt{H_2(0)}} \sum_{s \in D} h(s) Z(s) e^{-is^T\lambda}.
\]

Then the multidimensional tapered periodogram is defined as

\[
I^T_N(\lambda) = d^T(\lambda) d^T(\lambda) = \frac{1}{(2\pi)^d/2\sqrt{H_2(0)}} \left| \sum_{s \in D} h(s) Z(s) e^{-is^T\lambda} \right|^2.
\]

Besides, the DFT of the tapered values can be written as the convolution of the spectral window $H$ with the spectral random measure $dY$:

\[
d^T(\lambda) = \int_{\mathbb{R}^d} H(\lambda - \omega) dY(\omega).
\]

In practice, we wish $d^T(\lambda)$ not to be too different from $dY(\lambda)$. Therefore, $h$ should be chosen so that $H$ is a smooth weight function with most of its mass concentrated near the origin. Some data tapers for spatial processes can be found in [2], and recently Fuentes [10] proposed a rounded taper which gives more tapering to corner observations. This rounded taper presents discontinuities when weighting the borders, which may result unnaturally. Approximate large sample distributions for $d^T(\lambda)$ (and $d(\lambda)$ as a particular case) have been obtained in [3].

3. Main results. This section is devoted to some theoretical results concerning the expectation and the covariance of the multidimensional periodogram.

Proposition 1. Let $Z \sim SRF(\mu, C(\cdot))$ be observed on a regular grid $D = \prod_{j=1}^d D_j$ of size $N = \prod_{j=1}^d n_j$. Let $I_N$ be the multidimensional periodogram with no taper as in (3). Then

\[
E I_N(\lambda) = \frac{1}{(2\pi)^d N} \int_{\mathbb{R}^d} \prod_{j=1}^d \left( \frac{\sin(n_j \lambda_j)/2}{\sin \lambda_j/2} \right)^2 f(\omega) d\omega + \frac{1}{(2\pi)^d N} \prod_{j=1}^d \left( \frac{\sin(n_j \lambda_j)/2}{\sin \lambda_j/2} \right)^2 \mu^2.
\]

Proof. This statement is proved by just recalling that

\[
E(d(\lambda) d(\lambda)) = \text{cov}(d(\lambda), d(\lambda)) + E(d(\lambda)) E(d(\lambda))
\]

and noting that the Bartlett kernel can be obtained from

\[
\sum_{u \in U} \prod_{j=1}^d \left( 1 - \frac{|u_j|}{n_j} \right) e^{-iu^T\lambda} = \frac{1}{N} \prod_{j=1}^d \left( \frac{\sin(n_j \lambda_j)/2}{\sin \lambda_j/2} \right)^2,
\]

where the index set $U$ is given by $U = \prod_{j=1}^d U_j$, $U_j = \{1 - n_j, \ldots, n_j - 1\}$.

Proposition 2. Let $Z \sim SRF(\mu, C(\cdot))$ be observed on a regular grid $D = \prod_{j=1}^d D_j$ of size $N = \prod_{j=1}^d n_j$. Let $I^T_N$ be the tapered multidimensional periodogram as in (5). Assume that $\sum_{u \in U} |C(u)| < \infty$, with $U = \prod_{j=1}^d U_j$, where $U_j = \{1 - n_j, \ldots, n_j - 1\}$. Then

\[
E I^T_N(\lambda) = \frac{1}{(2\pi)^d/2\sqrt{H_2(0)}} \int_{\mathbb{R}^d} \left| H(\lambda - \omega) \right|^2 f(\omega) d\omega + \frac{|H(\lambda)|^2 \mu^2}{(2\pi)^d/2\sqrt{H_2(0)}}.
\]

Proof. Considering the expression of the multidimensional tapered periodogram in terms of the DFT of the tapered data (5), we have that

\[
E I^T_N(\lambda) = \frac{1}{(2\pi)^d/2\sqrt{H_2(0)}} \left( \text{cov}(d^T(\lambda), d^T(\lambda)) + E d^T(\lambda) \right).
\]
For the addend concerning the covariance operator, we obtain
\[
\frac{1}{(2\pi)^d H_2(0)} \text{cov} \left( d^T (\lambda), d^T (\lambda) \right) = \frac{1}{(2\pi)^d H_2(0)} \sum_{s \in D} \sum_{x \in D} h(s) h(x) e^{-i(s-x)' \lambda} C(s-x)
\]
\[
= \frac{1}{(2\pi)^d H_2(0)} \int_{\mathbb{R}^d} \left( \sum_{s \in D} h(s) e^{-i(s-x)'(\lambda - \omega)} \sum_{x \in D} h(x) e^{i\omega' (\lambda - \omega)} \right) f(\omega) \, d\omega
\]
\[
= \frac{1}{(2\pi)^d H_2(0)} \int_{\mathbb{R}^d} |H(\lambda - \omega)|^2 f(\omega) \, d\omega.
\]
For the second addend, it is immediate to see that it can be written as
\[
\frac{1}{(2\pi)^d H_2(0)} |H(\lambda)|^2 \mu^2
\]
which completes the proof.

Before introducing the results on the dependence structure of the multidimensional periodogram, we will give some brief notes on cumulants, which provide a tool for proving the next statements. More information on this topic can be found in [4], and a complete description of cumulants for spatial series is given in [2].

In the same way that the generating function of a random variable provides its moments, the logarithm of the generating function generates a sequence of numbers called cumulants. Cumulants are symmetric and multilinear in their arguments [4], and if any subset of \( \{X_1, \ldots, X_r\} \) is statistically independent of the remaining set, then \( \text{cum}(X_1, \ldots, X_r) = 0 \). From this property, cumulants may be used to measure the statistical dependence of variables.

For an \( r \)-variate random variable \( \{X_1, \ldots, X_r\} \) with \( \mathbb{E}[|X_j|^r] \leq \infty, j = 1, \ldots, r \) (\( X_j \) real or complex), the \( r \)th order joint cumulant \( \text{cum}(X_1, \ldots, X_r) \) of \( \{X_1, \ldots, X_r\} \) is defined as
\[
\text{cum}(X_1, \ldots, X_r) = \sum (-1)^{p-1} (p-1)! \mathbb{E} \left( \prod_{j \in \nu_1} X_j \right) \cdots \mathbb{E} \left( \prod_{j \in \nu_p} X_j \right),
\]
where the sum and products extend over all partitions \( \{\nu_1, \ldots, \nu_p\}, p = 1, \ldots, r \) of \( \{1, \ldots, r\} \). Note that the cumulant of a single variable is its expectation and the covariance between \( X \) and \( Y \) is the joint cumulant of \( \{X, Y\} \). In general, for a stationary random process \( Z \) such that its span of dependence satisfies
\[
\sum_{u \in \mathbb{R}^d} |\text{cum}(u_1, \ldots, u_{k-1})| < \infty \quad \text{for} \quad k = 2, 3, \ldots,
\]
the cumulant spectra of order \( k \) is defined by
\[
f_k(\lambda_1, \ldots, \lambda_{k-1}) = \frac{1}{(2\pi)^{dk(k-1)/2}} \sum_{u \in \mathbb{R}^d} \text{cum}(u_1, \ldots, u_{k-1}) \exp \left\{ -\frac{i}{2} \sum_{l=1}^{k-1} u_l^1 \lambda_l \right\},
\]
where \( \lambda_l \in \mathbb{R}^d \).

In order to show the subsequent results, we need the following assumption.

**Assumption (A1).** Given a stationary process \( Z \), there is an \( l \geq 0 \) such that
\[
\sum_{u \in \mathbb{R}^d} (1 + |u|^l) |C(u)| < \infty.
\]
This assumption implies that the spectral density has bounded and uniformly continuous derivatives of order less than or equal to \( l \).
Proposition 3. Let \( Z \sim SRF(0, C(\cdot)) \) be observed on a regular grid \( D = \prod_{j=1}^{d} D_j \) of size \( N = \prod_{j=1}^{d} n_j \). Let \( I_N \) be the multidimensional periodogram with no taper as in (3). Suppose that the covariance \( C \) satisfies Assumption (A1). Then, the covariance structure associated with the periodogram \( I_N \) is given by

\[
\text{cov} (I_N(\lambda), I_N(\omega)) = \left( \prod_{j=1}^{d} \left( \frac{\sin(n_j(\lambda_j + \omega_j))/2}{n_j \sin(\lambda_j + \omega_j)/2} \right)^2 \right) f^2(\lambda) + O(N^{-1}).
\]

Proof. Consider the following cumulants-based equality:

\[
\text{cov} (I_N(\lambda), I_N(\omega)) = \text{cum}(d(\lambda) d(\lambda), d(\omega) d(\omega)),
\]

and, for simplicity, denote

\[
\Delta(\lambda) = \sum_{s \in D} e^{-i s^T \lambda}.
\]

Since the cumulant of product variables can be expressed in terms of the sum of cumulants, we have that

\[
\text{cov} (I_N(\lambda), I_N(\omega)) = ((2\pi)^d)^3 f_4(\lambda, -\lambda, \omega) + O(1)
\]

\[
+ (\Delta(\lambda) f_1 + O(1) \Delta(-\lambda) f_2(\lambda, -\lambda, \omega) + O(1))
\]

\[
+ \text{three similar terms}
\]

\[
+ (\Delta(\lambda) f_1 + O(1) \Delta(\omega) f_1 + O(1) \Delta(-\lambda - \omega) f_2(-\lambda, -\lambda, \omega) + O(1))
\]

\[
+ \text{three similar terms}
\]

\[
+ ((2\pi)^d \Delta(\lambda + \omega) f_2(\lambda) + O(1) \Delta(-\lambda - \omega) f_2(-\lambda) + O(1))
\]

\[
+ ((2\pi)^d \Delta(\lambda - \omega) f_2(\lambda) + O(1) \Delta(-\lambda + \omega) f_2(-\lambda) + O(1))
\]

From the last two addends in the expression above, we obtain

\[
(2\pi)^d |\Delta(\lambda + \omega)|^2 f_2^2(\lambda) + (2\pi)^d |\Delta(\lambda - \omega)|^2 f_2^2(\lambda),
\]

which is equal to

\[
\left( \prod_{j=1}^{d} \left( \frac{\sin(n_j(\lambda_j + \omega_j))/2}{\sin(\lambda_j + \omega_j)/2} \right)^2 \right) f_2^2(\lambda),
\]

the leading term in (6). The \( O(N^{-1}) \) comes from some computations on the first addend, writing \( f_4 \) in terms of cumulants. Besides, assuming that \( f_1 = \mathbf{E} Z(s) = 0 \), the other terms in the sum vanish. Thus, the proof is completed.

Proposition 4. Let \( Z \sim SRF(0, C(\cdot)) \) be observed on a regular grid \( D = \prod_{j=1}^{d} D_j \) of size \( N = \prod_{j=1}^{d} n_j \). Let \( I_N^T \) be the tapered multidimensional periodogram as in (5). Suppose that the covariance \( C \) satisfies Assumption (A1). Then the covariance structure associated with the periodogram \( I_N^T \) is given by

\[
\text{cov} (I_N^T(\lambda), I_N^T(\omega)) = \left( |H_2(\lambda - \omega)|^2 + |H_2(\lambda + \omega)|^2 \right) \frac{f^2(\lambda)}{|H_2(0)|^2} + O(N^{-1}).
\]
Proof. For the complete expression of the tapered multidimensional periodogram, we must consider the fractional factors, where $H_2(0)$ is involved. Thus,

$$\text{cum}(d^T(\lambda), d^T(\hat{\lambda})) = ((2\pi)^d H_4(0) f_4(\lambda, -\lambda, \omega) + O(1)$$

$+ ((2\pi)^d H_2(\lambda + \omega) f_2(\lambda) + O(1))(2\pi)^d H_2(-\lambda - \omega) + O(1))$

$+ ((2\pi)^d H_2(\lambda - \omega) f_2(\lambda) + O(1))(2\pi)^d H_2(-\lambda + \omega) + O(1))$. 

The expression of the covariance is obtained following arguments similar to those in the proof of Proposition 3.

4. Conclusions and discussion. The goal of this work was to study the properties of the spatial and spatio-temporal periodogram (in both cases of tapering and no tapering) under the assumption of finite dimensionality of the regular grid where the process is observed. The work was motivated by the fact that, even if asymptotic assumptions allow one to work with periodograms meeting some nice features (independence, unbiasedness), on the other hand important information is lost about the behavior of the process.

The obtained results highlight several important aspects involving the bias of the multidimensional periodogram and the dependence between the periodogram values. For both cases, we argue that they are nonnegligible, despite what is shown by the asymptotic theory, and thus one must be cautious when implementing tests for independence or weighted least squares techniques if working on grids of small dimensions or with sparse data.

Another important aspect is that of separability in spatio-temporal processes. We showed in Propositions 3 and 4 that the covariance depends on the factorization of the weights, which are of the same type in the case of no tapering. If we consider the periodogram as a process to be studied, as we did in this work, we obtain that the weights of the squared spectral density always work in a separable way (in the sense of [5]), so we are not able to see whether the covariance of the periodogram is separable or not. It is generally difficult to give a dependence characterization of the periodogram values in terms of separability.

Acknowledgments. We are grateful to the referee and editor for comments that have clearly improved an earlier version of the manuscript.

REFERENCES


