Biometrika (2026), **113**, 2, asaf089 https://doi.org/10.1093/biomet/asaf089
Advance Access publication 18 December 2025

Spectral estimation for point processes and random fields

By J. P. GRAINGER

Institute of Mathematics, École Polytechnique Fédérale de Lausanne, Station 8, 1015 Lausanne, Switzerland
jake.grainger@epfl.ch

T. A. RAJALA

Natural Resources Institute Finland, Latokartanonkaari 9, 00790 Helsinki, Finland
tuomas.rajala@luke.fi

D. J. MURRELL

Research Department of Genetics, Evolution and Environment, Centre for Biodiversity and Environment Research, University College London, Gower Street, London WC1E 6BT, U.K.
d.murrell@ucl.ac.uk

and S. C. OLHEDE

Institute of Mathematics, École Polytechnique Fédérale de Lausanne, Station 8, 1015 Lausanne, Switzerland
sofia.olhede@epfl.ch

Summary

Spatial variables can be observed in many different forms, such as regularly sampled random fields (lattice data), point processes and randomly sampled spatial processes. Joint analysis of such collections of observations is clearly desirable, but complicated by the lack of an easily implementable analysis framework. We fill this gap by providing a multitaper analysis framework using coupled discrete and continuous data tapers, combined with the discrete Fourier transform for inference. Using this set of tools is important, as it forms the backbone for practical spectral analysis. In higher dimensions it is especially important not to be constrained to Cartesian product domains, and so we develop the methodology for spectral analysis using irregular domain data tapers and the tapered discrete Fourier transform. We discuss its fast implementation, as well as the asymptotic and large finite-domain properties. Estimators of partial association between different spatial processes are provided, as are principled methods to determine their significance, and we demonstrate their practical utility using a large-scale ecological dataset.

Some key words: Coherence; Partial coherence; Random field; Spatial multitapering; Spatial point pattern; Spectral representation.

1. Introduction

Collections of spatial variables studied in geostatistics, ecology and other spatial sciences involve complex interactions between a variety of different components. Often we need to jointly analyse data of different types, such as spatial point patterns, marked point patterns and realizations of random fields. We therefore need a common framework to include all data types in our analysis. Spectral analysis provides a convenient way to construct notions of correlation and partial correlation between these different types of processes. In this paper, we develop methodology to estimate such quantities for any combination of point patterns, marked point patterns and realizations of random fields, when the processes may be recorded using differing sampling methods, and when the observational region is not necessarily rectangular, but is common to all observed spatial variables. Existing methodology, with the exception of that for univariate Gaussian random fields (Andén & Romero, 2020), cannot handle arbitrary observational regions or different sampling mechanisms (including Rajala et al. 2023), and the existing spectral estimation methodology for marked point processes can be seen to be biased. All of these issues are resolved by the novel methodology that we propose in this paper.

Probabilistically, a spectral representation for general multivariate random measures is available (Brillinger, 1972; Daley & Vere-Jones, 2003), providing the theoretical background for our work. However, statistical estimation of the spectral density matrix function for such processes has not yet been developed. Our introduction of multitapering is a necessary step to develop statistical methodology for the spectral analysis of multivariate spatial data. We show in simulations that the large sample theory developed in this paper is applicable to data that is similar to data of practical interest, on which we also illustrate our methodology. This enables us to use principled thresholds to determine significance, and gives confidence in the quality of the proposed methodology.

Although methodology for spectral estimation exists for spatial point processes (Bartlett, 1964; Diggle et al., 1987; Mugglestone & Renshaw, 1996a,b; Rajala et al., 2023) and random fields (Bandyopadhyay & Lahiri, 2009; Matsuda & Yajima, 2009) separately, the extension to multivariate spatial data is more challenging. Kanaan et al. (2008) proposed an estimator for the cross-spectral density function between a random field and an unmarked point pattern, which is limited to random fields continuously sampled in a rectangular region. Eckardt & Mateu (2019b) proposed a periodogram when the random field is recorded on an integer grid within a rectangular domain, but did not study its properties or discuss smoothing or tapering. In reality, random fields can never be sampled continuously, and often we have multiple random fields recorded on different grids. Handling this is not trivial, and getting it wrong can result in substantial bias in the estimated spectra. We also consider marked point processes, where each point is associated with an additional random variable called a mark (e.g., size of a tree). A periodogram estimator has been proposed in the marked setting (Renshaw, 2002; Eckardt & Mateu, 2019a); however, this estimator is biased (see § 8.1 in the Supplementary Material for details). Hence, several key issues must be resolved to develop a practical, unified framework for spectral analysis of multivariate spatial data.

Whilst Andén & Romero (2020) considered spectral estimation for univariate Gaussian random fields on nonrectangular domains, performing such analyses in the case of multivariate random measures is more challenging. In particular, one first needs to construct continuous families of taper functions in order to analyse the point processes, and then build

discrete families of taper sequences that are appropriate for each grid used to record the random fields, but are related to the continuous tapers (otherwise, estimates of cross-statistics will be biased). We start from the discrete tapers of Simons & Wang (2011) and use these to generate continuous tapers, using the low wave-number concentration already required to retain the same desirable properties. We then construct discrete taper sequences from these continuous tapers, again exploiting the low wave-number concentration. As a result, we can combine all of these different kinds of data, with no need for aggregation or interpolation, and requiring no additional tuning parameters beyond the single bandwidth parameter already used for multitaper estimation in the case of time series and random fields (Walden, 2000). This results in a methodology that neatly handles all of these additional complexities presented by the richer class of spatial processes.

Understanding dependence between spatial processes is challenging, making it valuable to have diverse tools for analysing observations. However, existing exploratory methods may fail to clearly reveal underlying processes. A common approach is to model point processes, given some covariates, and then examine pairwise residual dependence, but this ignores interactions with other point processes in the system. Partial coherence (see § 2.3) offers a measure of dependence that accounts for all other processes, not just covariates.

Consider an inhomogeneous Poisson process and two cluster processes that share the Poisson process as a parent, but are otherwise independent. Suppose that the Poisson process intensity (Λ) is observed as a covariate (see the Appendix for details). Figure 1 illustrates a sample process on an irregular domain (see § 4.1), along with the inhomogeneous cross-L function (Baddeley et al., 2000), and the partial coherence between processes. These statistics are standardized to zero when no interaction remains after accounting for other processes. Pointwise extrema over 100 replications are shown as bands; the lines correspond to the realization in the top row. The inhomogeneous cross-L function incorrectly suggests interactions among all processes, while partial coherence (which we estimate in this paper) correctly identifies the independence of the cluster processes once the Poisson process is accounted for. Partial coherence also captures the association between the covariate (Λ) and the Poisson process, but not with the cluster processes, as expected. This toy example highlights how partial coherence complements existing spatial statistics tools, and why its accurate estimation is essential.

In this paper, we provide a single unified multitaper framework for estimation of the spectra and cross-spectra of random fields, point processes and marked point processes. The methodology is discussed in detail, and we provide theoretical justifications for the proposed estimators. We verify the properties of the proposed estimator through simulation studies, the first of their kind for spectral analysis of multivariate spatial data, and apply the methodology to forest ecology data from Barro Colorado Island (Condit et al., 2019). Several outstanding issues for spectral analysis in this setting are resolved, including the extension to nonrectangular regions, the bias in the marked case and handling different kinds of grid sampling. We then discuss how to use this framework to compute estimates of coherence and partial coherence between the different processes, with principled significance thresholds. We further demonstrate the promise of spectral analysis to summarize dependence in complex systems.

For readers from a point process background, Percival & Walden (2020) provided an introduction to spectral analysis for time series and, in particular, the multitaper method. Additionally, Walden (2000) discusses multitaper methods for multivariate time series, while Andén & Romero (2020) discussed multitaper spectral estimation for random fields on irregular domains. For readers from a time series background, Illian et al. (2008) provided an

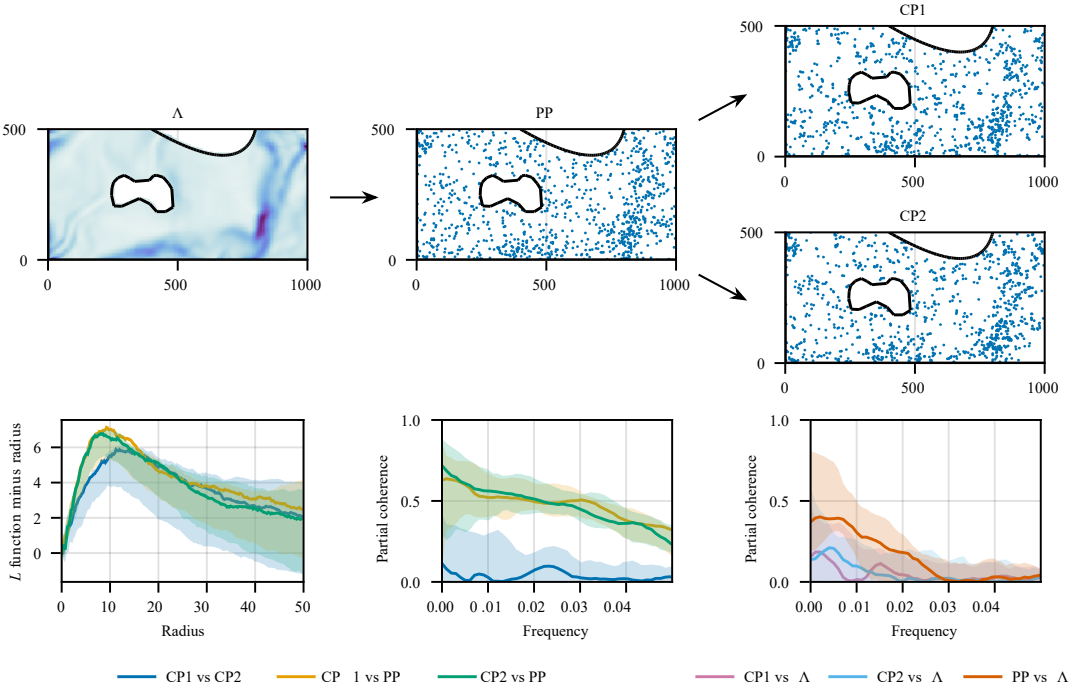

Fig. 1. Simulated patterns (top), the inhomogeneous cross-L function (bottom left) and the partial coherence between the point processes (bottom middle) and between the covariate and the processes (bottom right).

introduction to point processes. In either case, Chapter 8 of Daley & Vere-Jones (2003) provides a detailed treatment of the spectra of random measures, which is the most relevant theoretical background for this paper.

2. Background

2.1. *Basic notation*

We write $\mathbb{N} = \{1, 2, \ldots\}$ and, for $n \in \mathbb{N}$, $[n] = \{1, \ldots, n\}$. Given a set A and a binary operator on A, say $*$, then, for $a \in A$, we write $A * a = \{b * a \mid b \in A\}$ for the right coset and $a * A$ for the left coset. For two vectors a and b, we write $a \circ b$ for the elementwise (Hadamard) product, $a \oslash b$ for elementwise division and $a \cdot b$ for the dot product. Furthermore, we write $\prod(a) = \prod_{j=1}^{d} a_j$ for the product of elements of $a = (a_1, \ldots, a_d)^{\mathrm{T}}$.

2.2. *Random measures and spectral density functions*

A *point process* is a random set of locations in some space, say $\mathbb{R}^d$, such that there are finitely many points within a given bounded Borel set (Møller & Waagepetersen, 2003). A point process is said to be *simple* if points do not occur in the same location almost surely. A *marked point process* is a point process with additional information at each of the points, called a mark, which we take to be real valued and nonnegative. For a marked point process, the point process without the marks is often referred to as the *ground process*. The unmarked case is recovered when the marks are independent and the conditional distribution of the mark, given a point present at a location, the *mark kernel*, is just a point mass at 1. Let X be the set of random point locations, and let $W(x)$ be the mark at x, the point location.

The *mark-sum measure*, or *count measure* in the unmarked case (setting $W(x) = 1$), is

$$\xi(A) = \sum_{x \in X \cap A} W(x), \qquad A \in \mathcal{B}(\mathbb{R}^d),$$

where $\mathcal{B}(\mathbb{R}^d)$ denotes the Borel sets of $\mathbb{R}^d$ (Daley & Vere-Jones, 2003).

A *random field* is a random function, say Y, on $\mathbb{R}^d$ (Adler, 2010). Assuming that the random field is almost surely continuous and nonnegative, we can define a random measure from Y by

$$\xi(A) = \int_A Y(u)\,\mathrm{d}u, \qquad A \in \mathcal{B}(\mathbb{R}^d).$$

The two preceding equations show that random measures provide a unified framework in which to study these different spatial processes (Daley & Vere-Jones, 2003).

Assume that we have P such processes, and augment our previous notation by writing ξ_p for the pth random measure. We assume that these processes are homogeneous (Daley & Vere-Jones, 2007, Ch. 12) and that their second-order moment measures exist and are finite (the equivalent of finite variance for time series). From Daley & Vere-Jones (2003), for $1 \leqslant p, q \leqslant P$, the first and second moment measures are given for $A, B \in \mathcal{B}(\mathbb{R}^d)$ by

$$M_p(A) = E\{\xi_p(A)\}, \qquad M_{p,q}(A \times B) = E\{\xi_p(A)\xi_q(B)\},$$

with the appropriate extension of $M_{p,q}$ to $\mathcal{B}(\mathbb{R}^{2d})$. Under stationarity, these moment measures have reduced forms such that, for any $A \in \mathcal{B}(\mathbb{R}^d)$ and any g, a bounded measurable function of bounded support,

$$M_p(A) = \lambda_p \ell(A), \qquad \int_{\mathbb{R}^{2d}} g(x,y) M_{p,q}(\mathrm{d}x \times \mathrm{d}y) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} g(s+u,s)\ell(\mathrm{d}s)\breve{M}_{p,q}(\mathrm{d}u),$$

where λ_p is called the mean density, $\breve{M}_{p,q}$ is the reduced second-order moment measure and ℓ is the Lebesgue measure (Daley & Vere-Jones, 2003, Ch. 8). We write the lag in the first index to match the usual time series convention, in contrast to Daley & Vere-Jones (2003).

If ξ_p is a simple point process then λ_p is the intensity, which describes the average number of points per unit area. If ξ_p is a random field, λ_p is the mean at any given location. If ξ_p is a marked point process with a simple ground process then λ_p is the product of the mean mark and the intensity of the ground process (Illian et al., 2008, Equation 5.1.19).

The reduced covariance (signed) measure between the pth and qth processes is

$$\breve{C}_{p,q}(A) = \breve{M}_{p,q}(A) - \lambda_p \lambda_q \ell(A), \qquad A \in \mathcal{B}(\mathbb{R}^d).$$

Whilst the reduced covariance measures do not necessarily have densities, we write the reduced covariance density, $\breve{c}_{p,q}$, as a generalized density (which may have point masses), satisfying

$$\breve{C}_{p,q}(A) = \int_A \breve{c}_{p,q}(u)\,\mathrm{d}u, \qquad A \in \mathcal{B}(\mathbb{R}^d).$$

If $p = q$ and the process in question is a simple point process with a reduced factorial moment measure that admits a density, which we call $\rho_{p,p}$, then, $\breve{c}_{p,p}(\cdot) = \rho_{p,p}(\cdot) - \lambda_p^2 + \lambda_p \delta(\cdot)$,

where $\delta(\cdot)$ is the Dirac delta function. This is referred to as the complete covariance function by Bartlett (1963), who first introduced the spectra of point processes. If the processes in question are both random fields then $\breve{c}_{p,q}(\cdot) = \mathrm{cov}\{Y_p(\cdot), Y_q(0)\}$, the usual autocovariance function.

The notion of spectra exists in a more general form than that given here (Daley & Vere-Jones, 2003), but we are interested in processes for which the spectral density function exits. To ensure this existence, we require that the reduced covariance measure $\breve{C}_{p,q}$ is totally finite. In the random field case, this corresponds to the standard assumption that $\breve{c}_{p,q}$ is integrable, e.g., Brillinger (1974). In the point process case, this corresponds to assuming that $\rho_{p,q}(\cdot) - \lambda_p\lambda_q$ is integrable, as in Rajala et al. (2023) for example.

The (cross-)spectral density function between the pth and qth processes is defined as

$$f_{p,q}(k) = \int_{\mathbb{R}^d} \mathrm{e}^{-2\pi i k\cdot u}\breve{C}_{p,q}(\mathrm{d}u), \qquad k \in \mathbb{R}^d.$$

We call the matrix-valued function $f(\cdot) = [f_{p,q}(\cdot)]_{1\leqslant p,q\leqslant P}$ the spectral density matrix function. At a given wave number, $f(k)$ plays the role of a wave-number domain covariance matrix (Brillinger, 1972; Daley & Vere-Jones, 2003).

This definition generalizes the usual notion of the spectral density matrix function from time series and random fields, as well as including the point process case, as introduced by Bartlett (1963). However, in the marked case this differs slightly from the definition introduced by Renshaw (2002). In particular, Renshaw (2002) defined the spectral density to be proportional to the Fourier transform of the reduced factorial moment density of the mark-sum measure, whereas we define it to be the Fourier transform of the reduced covariance measure. Importantly, our definition corresponds to a special case of the definition for random measures, so we inherit all the properties of the spectral density matrix function listed by Daley & Vere-Jones (2003), such as positive semidefiniteness. In addition, setting the mark kernel to a point mass at one recovers the unmarked case, in that the spectral densities as we define them are the same. For a more thorough discussion, see § 8.2 of the Supplementary Material.

In order to handle sampling of the random fields, we need an additional assumption on the decay of their covariance function and spectral density function.

Assumption 1 (*Covariance decay*). The reduced covariance measure $\breve{C}_{p,q}$ is totally finite. In the case where both processes are random fields, the covariance density (function) $\breve{c}_{p,q}$ is continuous and there exist $C, \delta > 0$ such that, for all $x \in \mathbb{R}^d$, $|\breve{c}_{p,q}(x)| + |f_{p,q}(x)| \leqslant C(1 + \|x\|_2)^{-d-\delta}$.

This latter condition ensures that we have an aliasing relation between the spectral density function of the continuous process and its sampled counterpart. For example, random fields with Matérn (cross-)covariance functions satisfy Assumption 1.

So far, we only considered almost-surely nonnegative random fields; however, this condition can be relaxed by viewing everything as random-signed measures. Not all of the theory necessarily follows; see Daley & Vere-Jones (2003, § 8.4) for details. However, if these random signed measures do have reduced covariance measures satisfying Assumption 1 then we can also perform estimation with our framework (e.g., for many Gaussian processes).

2.3. *Coherence and partial coherence*

Typically, we standardise the cross-spectral density functions to complex *coherence*, which is the wave-number domain correlation between the two processes. Because coherence is complex valued, we take the magnitude, called magnitude coherence, and the argument, called phase (Carter, 1987). In particular, for $k \in \mathbb{R}^d$, define the coherence and phase as

$$r_{p,q}(k) = \frac{|f_{p,q}(k)|}{\{f_{p,p}(k) f_{q,q}(k)\}^{1/2}}, \qquad \theta_{p,q}(k) = \arg f_{p,q}(k).$$

Coherence has both benefits and limitations in this setting. One benefit is that the coherence between a random field and a point process is natural to define. In contrast, when considering spatial-domain cross-statistics between a random field and a point process, the random field needs to be made into a point process (or vice versa), after which standard spatial cross-statistics for point processes (or random fields) can be computed (Illian et al., 2008, § 6.11.2). This requires choices around how we convert one process into the other, which is not necessary with the wave-number domain approach. However, there are some clear limitations. In particular, the coherence assumes that the processes in question are pairwise homogeneous, and does not account for potential confounding from other observed processes.

One way to address both limitations is to consider partial coherence, which is the wave-number domain equivalent of partial correlation. This is equivalent to computing the coherence of residual processes, after first removing the linear effect of the other processes (Eichler et al., 2003). In other words, this approach can account for conditional inhomogeneity in the mean of the processes, as is also the case with the intensity reweighted stationary approaches introduced by Baddeley et al. (2000), and the parametric approach proposed by Waagepetersen et al. (2016).

More formally, define $\mathcal{V}_{p,q} = [P] \setminus \{p, q\}$ and let $f_{p,q \bullet \mathcal{V}_{p,q}}$ be the cross-spectral density function of two residual processes formed by producing the best linear prediction of ξ_p and ξ_q from the other processes $\{\xi_r \mid r \in \mathcal{V}_{p,q}\}$. As with partial correlation, the magnitude partial coherence and partial phase for $k \in \mathbb{R}^d$ are

$$r_{p,q \bullet \mathcal{V}_{p,q}}(k) = \frac{|f_{p,q \bullet \mathcal{V}_{p,q}}(k)|}{\{f_{p,p \bullet \mathcal{V}_{p,q}}(k) f_{q,q \bullet \mathcal{V}_{p,q}}(k)\}^{1/2}}, \qquad \theta_{p,q \bullet \mathcal{V}_{p,q}}(k) = \arg f_{p,q \bullet \mathcal{V}_{p,q}}(k),$$

respectively (Dahlhaus, 2000; Eckardt & Mateu, 2019a). Importantly, the magnitude partial coherence (and phase) can be computed efficiently by inverting the spectral matrix and making appropriate transformations (Dahlhaus, 2000). This approach is then able to account for conditional inhomogeneity in the processes, which is not possible with coherence alone. As demonstrated in Fig. 1, it is easy to account for the other observed point processes, as well as the random fields.

3. Estimation

3.1. *Multitapering*

Assume that we observe the random measures $\xi_1, \ldots, \xi_P$ on some bounded region $\mathcal{R} \subset \mathbb{R}^d$. Given a single realization, we aim to construct estimators of the spectral density matrix function. Of course, when the processes are random fields, we cannot observe them everywhere in $\mathcal{R}$, but must instead sample them discretely in space. In § 3.4 below, we discuss

irregular sampling approaches; however, here we assume that the random fields have been recorded on a regular grid. In particular, if the pth process is a random field, assume that it is observed on the intersection between $\mathcal{R}$ and the grid $\mathcal{G}_p = \mathbb{Z}^d \circ \Delta_p + v_p$, where $\Delta_p \in \mathbb{Q}^d_{>0}$ denotes the sampling interval and $v_p \in \mathbb{Q}^d$ the offset of the grid.

To estimate the spectral density function, we first construct multiple tapered Fourier transforms, whose covariance matrix is asymptotically $f(k)$, and then compute their sample covariance matrix. All of the processes in which we are interested can be thought of generally as random measures (Daley & Vere-Jones, 2003). Similarly, all of the tapered Fourier transforms can be seen as special cases of the tapered Fourier transform of a general random measure.

Consider a function $h\colon \mathbb{R}^d \to \mathbb{R}$, bounded in magnitude and zero outside $\mathcal{R}$, which we call a taper. Then the tapered Fourier transform of ξ_p is defined as

$$J_p(k; h) = \int_{\mathbb{R}^d} h(s)\, \mathrm{e}^{-2\pi i k\cdot s} \xi_p^0(\mathrm{d}s), \qquad k \in \mathbb{R}^d,$$

where ξ_p^0 denotes the centred random measure, i.e., $\xi_p^0(A) = \xi_p(A) - \lambda_p \ell(A)$. Because h is only nonzero inside $\mathcal{R}$, it is possible to compute the tapered Fourier transform from our observations.

If the pth process is a (marked) point process, with locations X_p and marks W_p, then

$$J_p(k; h) = \sum_{x \in X_p} h(x) W_p(x)\, \mathrm{e}^{-2\pi i k\cdot s} - \lambda_p H(k), \qquad k \in \mathbb{R}^d,$$

where, for $k \in \mathbb{R}^d$, $H(k) = \int_{\mathbb{R}^d} h(s)\, \mathrm{e}^{-2\pi i k\cdot s}\, \mathrm{d}s$ is the Fourier transform of h. When the process is not marked ($W_p(x) = 1$ almost surely), this is the tapered Fourier transform proposed by Rajala et al. (2023). When the process is marked, our approach represents a generalization to the marked setting.

If the pth process is a random field, with the field denoted by Y_p, then

$$J_p(k; h) = \int_{\mathbb{R}^d} h(s)\{Y_p(s) - \lambda_p\}\, \mathrm{e}^{-2\pi i k\cdot s}\, \mathrm{d}s, \qquad k \in \mathbb{R}^d.$$

However, computing this would require us to record the random field everywhere in $\mathcal{R}$. For some function g, let the grid-sampled function $g^{(\mathcal{G}_p)}$ be $g^{(\mathcal{G}_p)}(s) = g(s) \prod(\Delta_p) \sum_{z \in \mathcal{G}_p} \delta(s - z)$ for $s \in \mathbb{R}^d$, where $\delta(\cdot)$ is the Dirac delta function. Then, if we instead consider

$$J_p(k; h^{(\mathcal{G}_p)}) = \prod(\Delta_p) \sum_{s \in \mathcal{G}_p} h(s)\{Y_p(s) - \lambda_p\}\, \mathrm{e}^{-2\pi i k\cdot s}, \qquad k \in \mathbb{R}^d,$$

we see that this is essentially the usual tapered Fourier transform for random fields (see, e.g., Andén & Romero, 2020). It is important that the tapers h and $h^{(\mathcal{G}_p)}$ behave similarly (see Assumption 3 below), and this is easiest to achieve when the tapers used for the random fields are subsampled from the continuous base taper (with appropriate rescaling).

The tapered periodogram between processes p and q is

$$I_{p,q}(k; h_p, h_q) = J_p(k; h_p)\overline{J_q(k; h_q)}, \qquad k \in \mathbb{R}^d,$$

for the appropriate choices of h_p and h_q. The periodogram is the sample variance of one observation (with known mean zero). The rationale for this choice is that the tapered Fourier transform has a variance related to f. However, the clear issue is that the sample variance of one observation is a very poor estimate of the population variance. One technique to resolve this problem is multitapering, first proposed by Thomson (1982) for the time series setting, which constructs multiple different tapered Fourier transforms and then computes the sample variance of this collection. In particular, consider a family of tapers for each process $\{h_{p;m}\}_{m\in[M]}$, and define the shorthand notation

$$J_{p;m}(k) = J_p(k; h_{p;m}), \qquad I_{p,q;m}(k) = I_{p,q}(k; h_{p;m}, h_{q;m}), \qquad k \in \mathbb{R}^d,$$

where $h_{p;m}$ is the appropriate taper for the given sampling regime; that is, if the pth process is a (marked) point process, $h_{p;m} = h_m$, and if the pth process is a random field sampled on the grid $\mathcal{G}_p$ then we have $h_{p;m} = h_m^{(\mathcal{G}_p)}$. For a given m, the taper is always the same for two point processes, but for random fields, it will depend on the sampling grid. Then the multitaper estimator is

$$\hat{f}_{p,q}(k) = \frac{1}{M}\sum_{m=1}^{M} I_{p,q;m}(k), \qquad k \in \mathbb{R}^d.$$

The estimator performs well if the family of vectors of tapered Fourier transforms $J_m(k) = [J_{p;m}(k)]_{p\in[P]}$ is independent and identically distributed across m, which is true asymptotically (see Theorem 3 below).

3.2. *A general characterization for growing observational regions*

Whilst we can make stronger statements about specific region shapes and asymptotic regimes, we start with the generic setting in which we have a sequence of increasing regions, and ask what properties we would need tapers to satisfy in order to construct consistent estimators. Such properties will place some implicit constraints on the kinds of sequences of regions for which we can use multitapering. In § 2.2 of the Supplementary Material, we give explicit forms of growing domain and taper choice that satisfy all of the assumptions we introduce here. In general, these properties are useful, firstly to abstract some of the proofs, and secondly because they give a good indication of the desirable taper properties in practice, when given a single region.

Consider a sequence of regions $\{\mathcal{R}_n\}_{n\in\mathbb{N}}$ such that, for $n \in \mathbb{N}$, $\mathcal{R}_n \subset \mathcal{R}_{n+1}$ and $\ell(\bigcup_{n=1}^{\infty}\mathcal{R}_n) = \infty$. We aim to construct a sequence of families of tapers $\{h_{m,n}\}_{m\in[M_n]}$, where the number of tapers M_n may depend on n. For consistency of the multitaper estimate (Theorem 2), we will need M_n to grow with n, but for the asymptotic normality of the tapered Fourier transform (Theorem 3 below), we need M_n to be fixed. This will be made clear in the relevant results.

Assumption 2 (*Finite sample taper properties*). The family of tapers $\{h_{m,n}\}_{m\in[M_n]}$ is such that, for all $n \in \mathbb{N}$ and all $m \in [M_n]$, $h_{m,n}\colon \mathbb{R}^d \to \mathbb{R}$ is bounded, continuous, supported on a subset of $\mathcal{R}_n$ with $\|h_{m,n}\|_2^2 = 1$ and $\|H_{m,n}\|_1 < \infty$, where $H_{m,n}$ is the Fourier transform of $h_{m,n}$.

Boundedness and continuity ensure that the tapered Fourier transforms are well defined. The L^2-norm assumption ensures that the tapers are normalized. The finite L^1 norm allows

us to rearrange the order of integration in the proof of some results and to invert various Fourier transforms, and is satisfied by all taper families constructed in this paper. Using no taper (a scaled indicator function), we would not satisfy this assumption. However, due to the substantial leakage bias (Percival & Walden, 2020), we always use a taper.

There are two forms of bias present in spectral estimation, which we refer to as leakage bias and aliasing bias. The leakage bias results from the boundary effects, and is importantly a property of the taper that we control. The aliasing bias comes from the sampling of the random fields, and is a property of both the sampling grid and the underlying random field. Intuitively, aliasing results from only being able to sample a function on some regular grid. In this case, we are thinking of this function as the autocovariance of a random field (or between two random fields). The impact of such sampling in space is clear: we cannot observe the covariance at all possible lags, but only on a subset of them. In the wave-number domain, grid sampling results in aliasing, which is a more complicated phenomenon, where the Fourier transform we observe is the Fourier transform we care about, plus some erroneous values from higher wave numbers.

The bias effects of tapers are most easily understood in terms of a smoothing in the wave-number domain. To understand this, we need to consider the Fourier transform of the taper, and the aliased spectral density function. In particular, write $H_{p;m,n}$ for the Fourier transform of the taper, so that if the pth process is a point process,

$$H_{p;m,n}(k) = H_{m,n}(k) = \int_{\mathbb{R}^d} h_{m,n}(s)\, \mathrm{e}^{-2\pi\, ik\cdot s}\, \mathrm{d}s,$$

and, for a random field,

$$H_{p;m,n}(k) = H^{(\mathcal{G}_p)}_{m,n}(k) = \prod\left(\Delta_p\right) \sum_{s\in\mathcal{G}_p} h_{m,n}(s)\, \mathrm{e}^{-2\pi\, ik\cdot s},$$

for any $k \in \mathbb{R}^d$. In the latter case, $|H_{p;m,n}|$ is periodic on $K_p = [-1/2, 1/2]^d \oslash \Delta_p$, which we call the Nyquist box. For convenience, if the process is sampled continuously, we set $K_p = \mathbb{R}^d$. Aliasing for processes on a single grid (Percival & Walden, 2020) is simpler compared to multiple processes sampled in different ways. Therefore, we need a more general notion of aliased spectra.

Definition 1 (The aliased spectral density). *Introduce a set of aliasing wave numbers Ψ_p and a phase adjustment function $w_p\colon \mathbb{R}^d \to \mathbb{C}$, so that if the pth process is sampled on a grid then $\Psi_p = \mathbb{Z}^d \oslash \Delta_p$ and $w_p(x) = e^{-2\pi\, iv_p\cdot x}$; otherwise, $\Psi_p = \{0\}$ and $w_p(x) = 1$. Let the aliased spectral density function be*

$$\tilde{f}_{p,q}(k) = \sum_{\psi\in\Psi_p\cap\Psi_q} f_{p,q}(k+\psi) w_p(\psi)\overline{w_q(\psi)}, \qquad k \in \mathbb{R}^d.$$

Aliasing is not always present, but it is notationally convenient to define the aliased spectral density function in all cases. Since zero is always in the set $\Psi_p \cap \Psi_q$, we have $f_{p,q}(k) = \tilde{f}_{p,q}(k)$ if there is no aliasing, in particular, if one of the processes is sampled continuously in space. When both processes p and q are sampled on a grid, $|\tilde{f}_{p,q}|$ is periodic on the set $K_{p,q} = [-1/2, 1/2]^d \oslash \Delta_{p,q}$, where $\Delta_{p,q}$ is elementwise the largest number such that both $\Delta_p \oslash \Delta_{p,q} \in \mathbb{N}^d$ and $\Delta_q \oslash \Delta_{p,q} \in \mathbb{N}^d$. For example, if $\Delta_p = (4, 1/3)^{\mathrm{T}}$ and

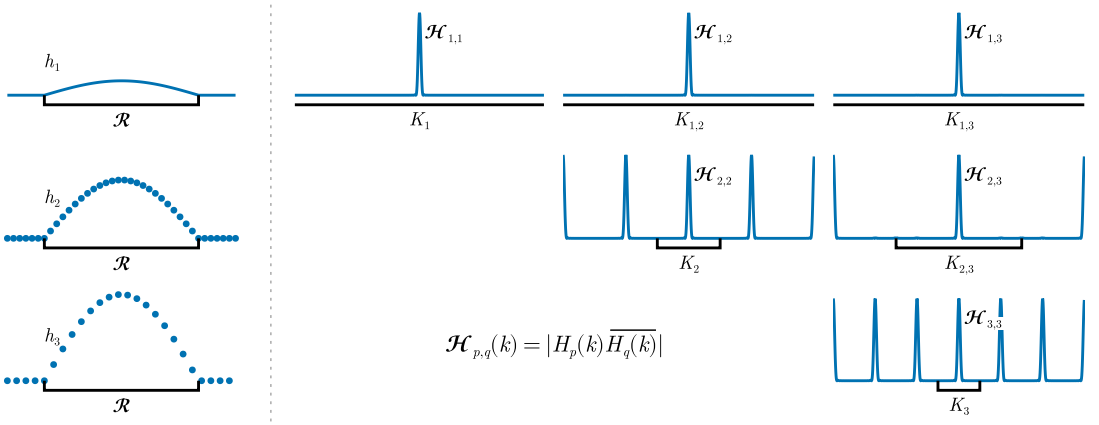

Fig. 2. An illustration of the aliasing effects in one dimension, assuming that the first process was recorded continuously, and that the second and third processes were recorded on grids with sampling intervals of 4 and 6, respectively. The taper in question is a minimum bias taper (Riedel & Sidorenko, 1995).

$\Delta_q = (6, 1/2)^{\mathrm{T}}$, $\Delta_{p,q} = (2, 1/6)^{\mathrm{T}}$. When one or both of the processes are sampled continuously in space, then we write $K_{p,q} = \mathbb{R}^d$. The product $H_{p;m,n}(k)\overline{H_{q;m,n}(k)}$, which is important for leakage bias, is also periodic in magnitude on $K_{p,q}$.

The left-hand side of Fig. 2 shows an example taper and its discrete counterparts on two different grids. The right-hand side of Fig. 2 shows the products of the different Fourier transforms of the taper and its discrete counterparts. In particular, we see from the diagonal terms that the discrete version repeats at different intervals. In addition, from the off-diagonals, we see that pairs including the continuous taper do not repeat, and the cross term between the two discrete tapers is periodic with a longer period than either of the originals, specifically $K_{p,q} \supseteq K_p$ for all $p, q \in [P]$.

Unless specified otherwise, we assume for simplicity that the intensity/mean λ_p is known, which we refer to as the 'oracle case'. Estimating the intensity has a negligible effect, except near wave number zero. However, the finite-sample equations become more complicated, which we discuss in detail in § 6 of the Supplementary Material. We can now obtain a general expression for the effect of the tapers and grid sampling on the bias of the periodogram. This generalizes the standard result for multivariate time series (Walden, 2000), random fields (Guillaumin et al., 2022) and univariate point processes (Rajala et al., 2023).

Proposition 1 (Expectation of the periodogram). *Given that the processes satisfy Assumption 1 and the tapers satisfy Assumption 2, the expectation of the periodogram is*

$$E\left\{I_{p,q;m,n}(k)\right\} = \int_{K_{p,q}} H_{p;m,n}(k')\overline{H_{q;m,n}(k')}\tilde{f}_{p,q}(k - k')\,\mathrm{d}k'$$

for all n, all $m \in [M_n]$ *and* $k \in \mathbb{R}^d$.

Proof. See the Supplementary Material. □

The only way to reduce aliasing bias is to sample the process at a higher resolution, or to sample it randomly and treat it as a marked point process (see § 9 of the Supplementary Material). Since we typically do not control the sampling mechanism, we cannot control the aliasing bias, and should just be aware that it exists. However, since the expectation is

a convolution between the aliased spectral density function and a property of the taper, we can control leakage bias by constraining the behaviour of the taper.

The underlying idea is to construct the tapers such that their Fourier transforms concentrate within a bandwidth around zero, so that we keep the smoothing effects local. In finite samples, we use the methodology proposed by Simons & Wang (2011) in order to obtain tapers with such a property on arbitrary domains. Asymptotically, under the assumption that the concentration increases appropriately with the growing domain and with the continuity of the spectral density function, we obtain asymptotic unbiasedness (up to an aliasing effect).

Assumption 3 (*Taper concentration*). There exists a sequence of bandwidths b_n such that $b_n \to 0$ as $n \to \infty$. Furthermore, for all $p, q \in [P]$,

$$\max_{m\in[M_n]} \left|1 - \int_{B_{b_n}} H_{p;m,n}(k)\overline{H_{q;m,n}(k)}\,\mathrm{d}k\right| \to 0, \qquad \max_{m\in[M_n]} \int_{K_p\setminus B_{b_n}} \left|H_{p;m,n}(k)\right|^2 \mathrm{d}k \to 0,$$

as $n \to \infty$, where B_{b_n} is a ball centred at zero of radius b_n.

In the case when every process is sampled continuously, the latter condition is not necessary, as $\|H_{m,n}\|_2 = \|h_{m,n}\|_2 = 1$. However, for the processes sampled on a grid, it is necessary. If the former condition holds then, for the grid-sampled cases, the latter condition is equivalent to $\max_{m\in[M_n]} |1 - \prod(\Delta_p)\sum_{u\in\mathcal{G}_p} h_{m,n}(u)^2| \to 0$ as $n \to \infty$.

Theorem 1 (Asymptotic bias of the periodogram). *Let the processes satisfy Assumption 1 and the tapers satisfy Assumptions 2 and 3 for some fixed number of tapers M. For all $k \in \mathbb{R}^d$, $p, q \in [P]$ and all fixed $m \in [M]$, $E\{I_{p,q;m,n}(k)\} \to \tilde{f}_{p,q}(k)$, as $n \to \infty$.*

Proof. See the Supplementary Material. □

Theorem 1 shows that the only asymptotic bias of the periodogram is due to aliasing. Importantly, in the case where at least one of the processes is not sampled on a grid, the periodogram is asymptotically unbiased. This includes marked point processes, meaning that we have resolved the bias present in the current state-of-the-art estimator (Renshaw, 2002). The two main results we will establish are consistency of the multitaper periodogram under a growing number of tapers, and asymptotic normality of the tapered discrete Fourier transform under a finite number of tapers. In both cases, we need the tapers to be asymptotically orthogonal.

Assumption 4 (*Asymptotic orthogonality*). For all $p, q \in [P]$, as $n \to \infty$,

$$\max_{m\in[M_n]} \max_{m'\in[M_n]\setminus\{m\}} \left|\int_{B_{b_n}} H_{p;m,n}(k)\overline{H_{q;m',n}(k)}\,\mathrm{d}k\right| \to 0.$$

In general, we construct tapers that are close to being orthogonal, but not quite, in particular, when constructing tapers for unusual region shapes. In practice, Assumptions 3 and 4 can be checked numerically, as will be discussed in § 3.3 below. As well as yielding variance reduction for consistency, this assumption also ensures that the different tapered Fourier transforms are asymptotically uncorrelated when we study asymptotic normality.

We also assume that the L^1 norms of the Fourier transforms of the tapers shrink to zero as the observational region grows.

Assumption 5 (*Shrinking taper Fourier transform*). For all $p \in [P]$, there exists a $C_p < \infty$ such that

$$\max_{m\in[M_n]} \int_{K_p} \left|H_{p;m,n}(k)\right| \,\mathrm{d}k \to 0, \qquad \max_{m\in[M_n]} \sum_{\psi\in\Psi_p} \left(\int_{K_p+\psi} \left|H_{m,n}(k)\right|^2 \,\mathrm{d}k \right)^{1/2} \to C_p,$$

as $n \to \infty$.

The first part of Assumption 5 is commonly used when dealing with spectral estimation on unusual domains or more abstract settings (Brillinger, 1982). The second part is a technical condition to handle the grid-sampled processes. See § 2.2 of the Supplementary Material for examples of constructions where both conditions are satisfied. In particular, both conditions are implied if we assume that the Fourier transform of the taper is bounded by a decreasing function with appropriate tail decay.

Finally, we need to make an assumption about the higher-order cumulants of the processes. At a high level, we just need to ensure that fourth-order terms that appear in the variance of the periodogram (due to the periodogram being a product of two tapered Fourier transforms) are negligible. For definitions of the sampled process ζ_p and cumulant measures, see the Appendix.

Assumption 6 (*Fourth-order cumulants*). For $p, q \in [P]$, the first four joint moment measures of ζ_p, ζ_q exist and are finite, and the reduced cumulant measure $\breve{C}^{(\zeta)}_{p,p,q,q}$ is totally finite.

Assumption 6 corresponds to the standard integrability/absolute summability of the higher-order cumulant function in the time series case (Brillinger, 1965), analogously to Assumption 1.

Theorem 2 (Consistency of the multitaper periodogram). *If the processes satisfy Assumptions 1 and 6, the tapers satisfy Assumptions 2 to 5 and $M_n \to \infty$ as $n \to \infty$, then, for all $k \in \mathbb{R}^d$ and all $p, q \in [P]$, $E\{\hat{f}_{p,q;n}(k)\} \to \tilde{f}_{p,q}(k)$ and $\mathrm{var}\{\hat{f}_{p,q;n}(k)\} \to 0$ as $n \to \infty$.*

Proof. See the Supplementary Material. □

As a result, we see that if we grow the number of tapers with appropriate control on their behaviour then we obtain an estimator that is mean-square consistent, which is not true when using a single taper.

There are two main approaches to establishing asymptotic normality of the tapered Fourier transform. The first is to use Brillinger mixing conditions (Brillinger, 1982), which assume that all the higher-order cumulant measures exist and are totally finite. The second is to use an α-mixing condition (Yang & Guan, 2025, for example), typically with the assumption that at least finitely many moments exist and are finite (see, e.g., Biscio & Coeurjolly, 2016 for a discussion of the two assumptions in the point process setting). In § 2.1 of the Supplementary Material, we give a result using the general α-mixing central limit theorem from Biscio & Waagepetersen (2019). However, we use Brillinger mixing here as we already introduced higher-order cumulants.

Assumption 7 (*Brillinger mixing*). For all $r \in \mathbb{N}$ and all $p_1, \ldots, p_r \in [P]$, the joint moment measures of $\zeta_{p_1}, \ldots, \zeta_{p_r}$ up to order r exist and are finite, and $\breve{C}^{(\zeta)}_{p_1,\ldots,p_r}$ is totally finite.

Assumption 7 will be used for asymptotic normality of the tapered Fourier transforms. Since the tapered Fourier transforms are complex valued, convergence will be to a complex normal distribution. We write $\mathcal{CN}(\mu, S, R)$ for a multivariate complex-normal distribution mean vector μ, covariance matrix S and relation matrix R. Furthermore, since we transform real-valued processes, the tapered Fourier transforms are symmetric. In addition, for certain wave numbers, the tapered Fourier transforms are real valued (at zero and multiples of the Nyquist wave numbers). This is the same as the time series case (see Brillinger, 1974), with a slightly more complicated formulation due to the different kind of sampling. For this reason, we restrict attention to wave numbers where these issues do not occur.

THEOREM 3 (ASYMPTOTIC NORMALITY OF TAPERED FOURIER TRANSFORMS). *Let the processes satisfy Assumptions 1 and 7 and the tapers satisfy Assumptions 2 to 5. Consider* $k_1, \ldots, k_r \in \mathbb{R}^d$ *such that, for all* $i, j \in [r]$ *with* $i \neq j$, $2k_i, 2k_j, k_i \pm k_j \notin \bigcup_{p\in[P]} \Psi_p$. *Then, for all* $i \in [r]$,

$$J_{m,n}(k_i) = [J_{p;m,n}(k_i)]_{1\leqslant p\leqslant P} \xrightarrow{\mathrm{D}} \mathcal{CN}\{0, \tilde{f}(k_i), 0\}$$

as $n \to \infty$. *In addition, for any two distinct* m, m', *the vectors* $J_{m,n}(k_i)$ *and* $J_{m',n}(k_i)$ *are asymptotically independent, so the matrix* $J_n(k_i) = [J_{m,n}(k_i)^{\mathrm{T}}]_{1\leqslant m\leqslant M}$ *is asymptotically complex normal with uncorrelated rows. Finally, for* $i \neq j$, $J_n(k_i)$ *and* $J_n(k_j)$ *are asymptotically independent.*

Proof. See the Supplementary Material. □

3.3. *Practical construction of tapers on a region of interest*

For a general region that we observe (including nonconvex regions), we can construct a family of tapers that satisfy the finite-sample conditions in Assumption 2, and for which we can check how well asymptotic conditions in the remaining assumptions are satisfied. We do this by first taking (multi-dimensional) Slepian sequences on some grid that can be constructed numerically using the methodology of Simons & Wang (2011). Then we multilinearly interpolate these taper sequences to construct a family of continuous tapers on the region of interest.

PROPOSITION 2 *Let* $g_1, \ldots, g_M$ *be a family of discrete space tapers on a grid* $\mathcal{G} = \Delta \circ \mathbb{Z}^d + v$, *nonzero only on the bounded region* $\tilde{\mathcal{R}} = \{s \in \mathcal{R} \mid s + \Delta \circ [-1, 1]^d \subseteq \mathcal{R}\}$. *Then, if we construct a family of tapers via multilinear interpolation of* $g_1, \ldots, g_m$, *the interpolated family satisfies Assumption 2, provided they are correctly normalized. Furthermore, their Fourier transform is*

$$H_m(k) = G_m^{(\mathcal{G})}(k) \prod_{j=1}^{d} \mathrm{sinc}^2(\pi\, \Delta_j k_j),$$

where k_j *and* Δ_j *are the* j*th components of* k *and* Δ, *respectively, and* $\mathrm{sinc}(x) = \sin(x)/x$.

Proof. See § 3.6 of the Supplementary Material. □

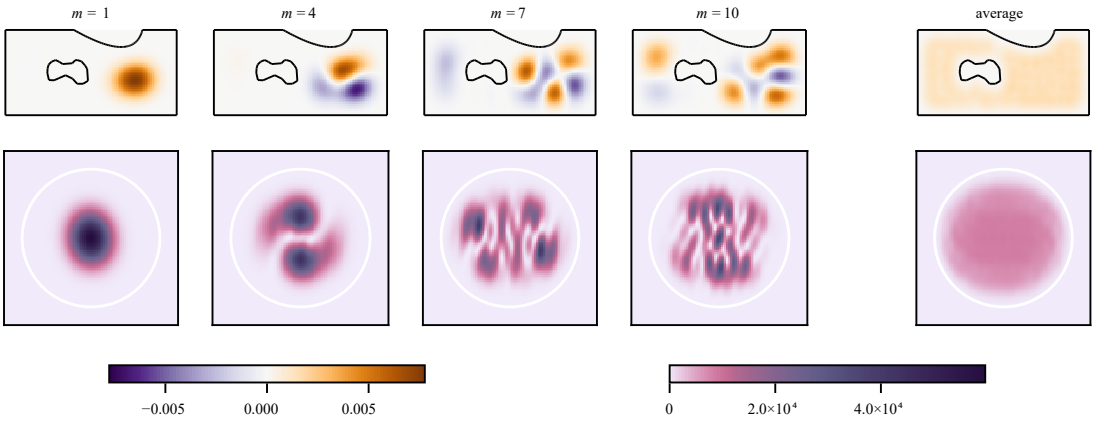

Fig. 3. Four of the tapers and the total average absolute taper for an irregular region (top), and the corresponding spectral windows (bottom). The spatial data have axes ranging from 0 to 1000 and 0 to 500, and the wave-number domain plots range from −0.01 to 0.01 in each dimension.

The construction on the smaller region $\tilde{\mathcal{R}}$ is to ensure that the interpolated tapers are zero outside $\mathcal{R}$, avoiding linear interpolation across holes not fully covered by the grid for example. In order to correctly normalize the tapers, we need their L^2 norm, which can be computed exactly with finite sums using the results in § 7.2 of the Supplementary Material. We can then check the conditions of Assumptions 3 and 4 numerically.

Additional specific details of the taper construction and the effect of linear interpolation are given in § 7 of the Supplementary Material. In the case of random fields sampled on a regular grid, the tapered Fourier transforms can be computed efficiently using fast Fourier transforms (Cooley & Tukey, 1965), as is well known. For point patterns, we can use nonuniform fast Fourier transforms, the properties of which were studied by Dutt & Rokhlin (1993). We use the NUFFT implementation from Barnett et al. (2019).

A subset of the tapers we use for the Barro Colorado Island data (see § 4.1 below) is shown in Fig. 3. Whilst any given taper concentrates on a certain region of the spatial domain, the total weight is approximately uniform across the spatial domain, except at the borders (Fig. 3, top row). Similarly, in the wave-number domain we see that the spectral windows are concentrated on different regions within the bandwidth of zero (denoted by the white circle), but together they cover this region evenly (Fig. 3, bottom row). This points to the major benefit of multitapering: if we were to use a single taper and kernel smoothing, we would focus on a small part of the observational region, whereas multitapering uses most of the region evenly.

3.4. *Irregular sampling*

So far, we have only considered the case where the processes are sampled on a regular grid. However, we may also want to handle situations in which they are not. This specific case has been considered previously in the literature, for example, Masry (1978, 2003) and Matsuda & Yajima (2009). In fact, this also fits into the framework we introduce in this paper. Suppose that we have a mean-zero random field Y that we sample at the now-random locations given

by the point process X. Then we can regard this sampled process as a mark-sum measure, say ξ, corresponding to the marked point process with marks $Y(x)$ for $x \in X$. Then, if the marks are independent of the locations, from Daley & Vere-Jones (2003, p. 338), we have

$$f_{\xi\xi}(k) = \int_{\mathbb{R}^d} f_{YY}(k-k')f_{XX}(k')\,\mathrm{d}k' + \lambda_X f_{YY}(k), \qquad k \in \mathbb{R}^d,$$

where $f_{\xi\xi}$ is the spectral density function of the marked process, f_{YY} is the spectral density function of the random field, f_{XX} is the spectral density function of the point process and λ_X is the intensity of the point process. If the point process used for sampling is Poisson then

$$f_{\xi\xi}(k) = \lambda_X \mathrm{var}\{Y(0)\} + \lambda_X^2 f_{YY}(k), \qquad k \in \mathbb{R}^d.$$

The spectra of the random field f_{YY} can then be estimated by plug-in estimation and appropriate rearrangement. That estimator would be identical to the one we would obtain if the sample locations were uniform over the region and we used the methodology proposed by Matsuda & Yajima (2009). For more details, see § 9 of the Supplementary Material.

3.5. *Coherence and partial coherence*

As usual for spectral estimation, we use a plug-in estimator for the magnitude coherence, phase and their partial counterparts. This requires the estimated spectral density matrix at a given frequency to be invertible, so we need to have more tapers than processes, a necessary, but not sufficient condition (Walden, 2000). In particular, we cannot use the periodogram as a plug-in estimator for the magnitude partial coherence as the periodogram is not invertible (Walden, 2000).

The probability density function of the absolute value of the sample correlation of a bivariate complex proper normal random vector with absolute correlation ρ and N observations is

$$f(r;\rho,N) = 2(N-1)(1-\rho^2)^N r(1-r^2)^{N-2}\,{}_2F_1(N,N,1;\rho^2 r^2) \tag{1}$$

if $0 < r < 1$ and zero otherwise (Miller, 1980). Provided the assumptions of Theorem 3 hold, matrix $J_n(k)$ is asymptotically complex normal. Therefore, the plug-in estimator for $r_{p,q}$, which is the absolute value of the sample correlation of matrix $J_n(k)$, has an asymptotic distribution with probability density function given by (1), with $\rho = \tilde{r}_{p,q}$ and $N = M$, as in the case of segment averaging (Goodman, 1957; Carter et al., 1973). The same applies to the partial coherence estimator, provided that $M \gg P$, which is the case here, and is easier to achieve in the spatial setting compared to time series, due to the typically large number of tapers that can be constructed. When this condition fails, better approximations of the null distribution (Medkour et al., 2009) or shrinkage approaches (Schneider-Luftman & Walden, 2016) may be required. These extensions work mutatis mutandis, given our developments.

4. Application and simulation study

4.1. *Barro Colorado Island data*

The Barro Colorado Island study records the locations of every individual tree of a trunk diameter of at least 10 mm within a 1000×500 m^2 rectangle of tropical rainforest, along

with additional measurements, such as the diameter and species name of each tree (Condit et al., 2019). In addition, they record topological features and the concentration of various soil chemicals, all sampled on regular grids.

Whilst the study records data on a rectangular domain, it is known that the forest is not homogeneous. In particular, there is a region near the northern border of the plot known to consist of a much younger forest dominated by the pioneer tree species *Gustavia superba* (Hubbell & Foster, 1983, 1986), which we remove to make the assumption of homogeneity more reasonable. Similarly, we also remove a swamp region in the centre of the plot where the soil has much higher moisture content and the associated flora is more typical of a riparian habitat (Harms et al., 2001).

There are around 200 species present in the plot, along with many additional covariates. To fully analyse such data, we would need to develop methodology to handle that high-dimensional setting. This is beyond the scope of this paper, so, for the sake of illustration, we consider a small subset of the available species in our analysis. In particular, Harms et al. (2001) listed five species that are known to be associated with the gradient of the terrain. We take the three most abundant species from this list, *Beilschmiedia tovarensis*, *Poulsenia armata* and *Unonopsis pittieri*, and consider the gradient of the terrain as the fourth process.

To test the significance of magnitude (partial) coherence, we use the distribution in (1) and a Bonferroni correction (Bonferroni, 1936). In particular, we correct as if we were performing $K_n P(P-1)/2$ tests, where $P = 4$ is the number of processes and K_n is the approximate number of wave numbers on which the multitaper estimate is uncorrelated. Approximately, the multitaper estimate is uncorrelated when wave numbers are spaced two bandwidths apart and when we use only wave numbers where the first coordinate is nonnegative (as there is a symmetry); see § 2.3 of the Supplementary Material for details. We display more than K_n wave numbers for visualization purposes, but the threshold is based on K_n, because correcting for all displayed wave numbers would be overly conservative. It would be preferable to utilize envelope approaches, e.g., Mrkvička & Myllymäki (2023), but these require resampling from the null distribution, which is difficult when the region shape is irregular, and are beyond the scope of this paper.

Figure 4 shows the coherence and partial coherence between each of the four processes from the Barro Colorado Island dataset, alongside the observed data. We have thresholded the coherence and partial coherence based on the marginal asymptotic null distribution with the correction described above. As expected, we can see that there is (statistically) significant coherence between each of the three species' locations and the gradient (Fig. 4, plot matrix upper-right triangle). In addition, there is significant coherence between each of the pairs of species. When we consider partial coherence, thus correcting to some extent for the other processes, we see that there is still residual coherence between the gradient and each of the species' locations. (Fig. 4, plot matrix lower-left triangle). However, *B. tovarensis* and *P. armata* do not have significant partial coherence, whereas *U. pittieri* has significant partial coherence with both of the other species. From the phase, shown in § 12 of the Supplementary Material, we see that all significant associations have a positive sign. Although this example is a preliminary analysis that could change when all other processes and covariates are considered, this example highlights the utility of partial coherence and how it is likely to lead to new inferences compared to existing methods.

4.2. *Simulation study*

To investigate the asymptotic properties of the proposed method, we conduct a simulation study, examining the distribution of magnitude coherence and partial coherence

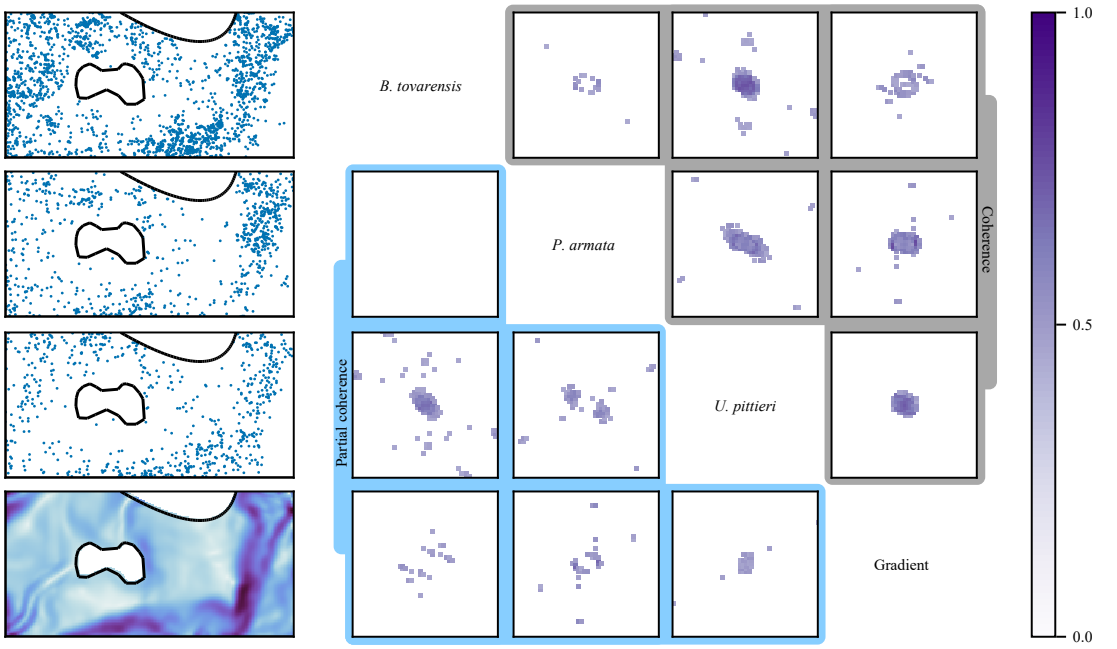

Fig. 4. The first column corresponds to the three point processes *B. tovarensis*, *P. armata*, *U. pittieri* and the gradient of the terrain. The 4 by 4 panel subplot on the right shows the magnitude coherence (upper triangle) and magnitude partial coherence (lower triangle) between the four processes. The spectral plots range from to 0.05 on each axis, and the data are shown with axes ranging from 0 to 1000 and 0 to 500.

between two point processes and a random field, and comparing it to the asymptotic distribution in (1). In the following models, we always have two point processes (PP1 and PP2) and a random field (RF). In each case, the random field is a log-Gaussian process, where the Gaussian process has mean -6 and a Matérn covariance function with length scale 20 m, smoothness 3 and variance 1, and is recorded on a grid with spacing every 5 m starting from (0 m, 0 m), as in the case of the gradient data from Barro Colorado Island. The specific details of the three models are as follows.

Model 1. The two point processes are independent log-Gaussian Cox processes with the Gaussian process described above driving them (independently of the recorded random field).

Model 2. The two point processes are independently Cox processes, with the random intensity driven by the field (i.e., they are log-Gaussian Cox processes; see Møller et al., 1998).

Model 3. The second point process (PP2) is a Cox process with the random intensity driven by the field, and the first point process (PP1) is the result of random clustering around the second point process, where, for each point in the first process, there are Poisson(1) offspring, who are placed randomly around the parent with an $N(0, 5^2)$ distribution in each direction.

For each case, we generate 1000 replications, and compute estimates of the spectral density matrix function. All simulations use the region from the Barro Colorado Island data. It is not possible to visualize all of the results, so we focus on two wave numbers that illustrate different areas of the spectrum. The first is at a low wave number on the slope (so

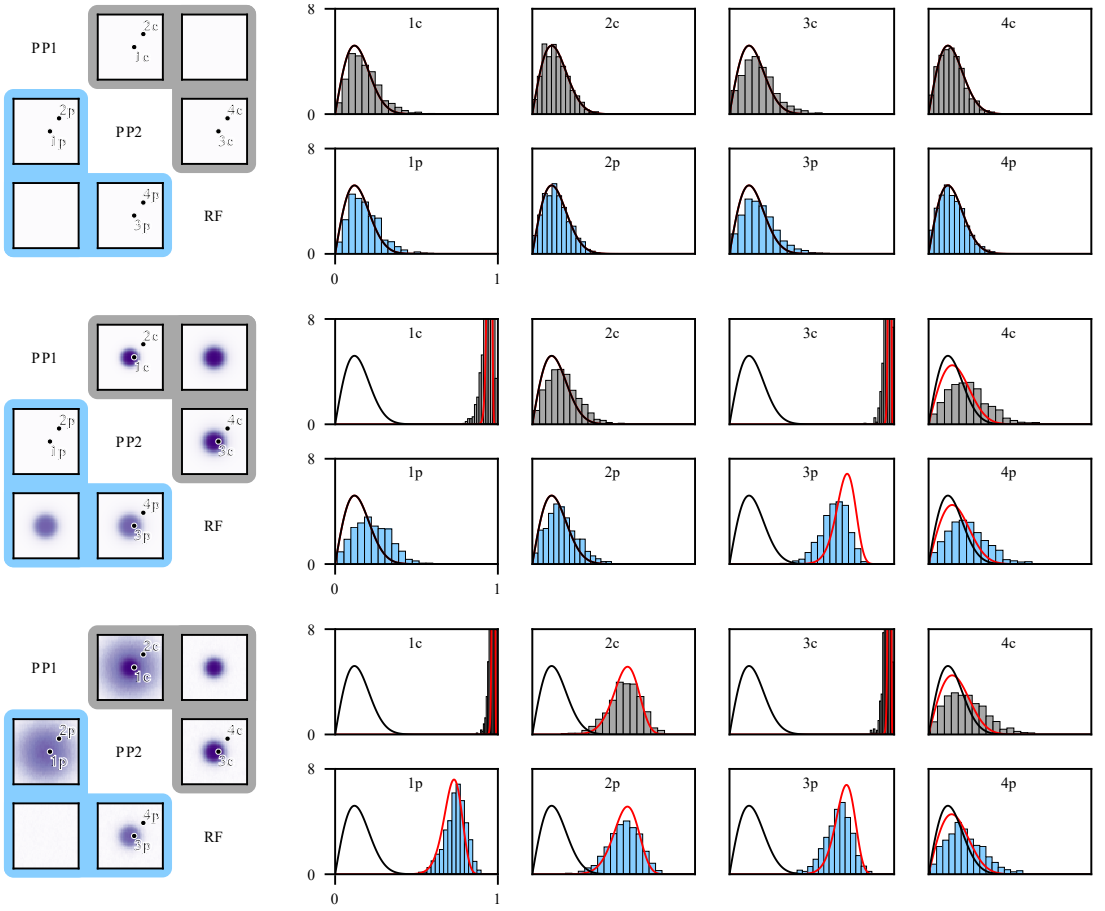

Fig. 5. Simulation results for the sampling distribution of the magnitude coherence. The left column shows, for each model, the true (partial) coherence in the upper (lower) triangle, where both wave-number components k_1 and k_2 range from -0.05 to 0.05. The right-hand column shows the estimated (partial) coherence on top (bottom), along with the asymptotic distribution (red line) and null distribution (black line). The rows correspond to models 1–3, from top to bottom. The result is optimal when the asymptotic (red line) matches the simulation-based histogram.

we expect smoothing effects to degrade the quality of the approximation); the second is at a high wave number, exploring the effects in the tails. At each wave number, we show the empirical and asymptotic distributions of magnitude coherence and magnitude partial coherence between the point processes (processes 1 and 2) and between one of the point processes and the field (processes 2 and 3). The left-hand column of Fig. 5 shows the true spectral matrix for each model, with the magnitude (partial) coherence in the upper (lower) triangle, respectively. The specific wave numbers for each comparison are marked with a point and a label corresponding to the appropriate panel in the right-hand column. For model 1, the empirical distribution of the magnitude coherence and magnitude partial coherence align well with the asymptotic distribution (Fig. 5, top section). In models 2 and 3, the asymptotic distributions begin to break down in some of the nonnull cases (Fig. 5, middle and bottom sections). However, we would still detect the effect as being significant when compared to the null asymptotic distribution. Whilst not perfect, the results are promising and suggest that the asymptotic distribution is a good approximation in many cases.

ACKNOWLEDGEMENT

S.C.O. would like to thank the European Research Council (CoG 2015-682172NETS), within the Seventh European Union Framework Program. We thank the editorial team and the reviewers for their comments and suggestions.

SUPPLEMENTARY MATERIAL

The Supplementary Material includes proofs, additional results and discussion, and further simulation studies. The code is available from https://doi.org/10.5281/zenodo.17780751, and the methodology is implemented in the `SpatialMultitaper.jl` Julia package, available from https://doi.org/10.5281/zenodo.17779549.

APPENDIX

Simulation details

For the example in § 1, each cluster process was generated by placing a cluster at every point in the parent process, where the number of children N and location relative to parent X were given by $N \sim \text{Po}(1)$, $X \sim N(0, 5^2 I_2)$, where I_2 is the identity. The intensity for the inhomogeneous Poisson was specified by $\log\{\lambda(x)\} = 9Y(x) - 7$, where Y is the gradient of the Barro Colorado Island terrain. The coefficients were based on fitting a model for the intensity of *P. armata* from the gradient.

Higher-order cumulants and reduced cumulant measures

Let $\mathcal{G} = \Delta \circ \mathbb{Z}^d$ be a grid that contains all of the other grids we consider. Weighted sums on the original grid can always be constructed by multiplying the desired weights by $\mathbb{1}_{\mathcal{G}_p}$. Define the sampled random measures ζ_p for $A \in \mathcal{B}(\mathbb{R}^d)$ by $\zeta_p(A) = \xi_p(A)$ if the pth process is sampled continuously, and by $\zeta_p(A) = \sum_{u \in \mathcal{G} \cap A} Y_p(u)$ if the pth process is sampled on a grid. Let $\mathscr{C}[X_1, \ldots, X_r]$ denote the rth-order cumulant of some random variables $X_1, \ldots, X_r$. Define the rth-order cumulant measure of the sampled process as $C^{(\zeta)}_{p_1,\ldots,p_r}(A_1 \times \cdots \times A_r) = \mathscr{C}[\zeta_{p_1}(A_1), \ldots, \zeta_{p_r}(A_r)]$ for $A_1, \ldots, A_r \in \mathcal{B}(\mathbb{R}^d)$, appropriately extended to $\mathcal{B}(\mathbb{R}^{rd})$ (Daley & Vere-Jones, 2007). We need to extend Proposition 12.6.III of Daley & Vere-Jones (2007) to include grid sampling.

LEMMA A1 (REDUCED CUMULANT MEASURE). *For $p_1, \ldots, p_r \in [P]$, assume that the rth joint moment measures of $\xi_{p_1}, \ldots, \xi_{p_r}$ exist and are finite. Suppose that either all of the processes are sampled continuously or they are enumerated so that at least ξ_{p_r} is sampled on a grid. Let ℓ_{p_r} be the Lebesgue measure on $\mathbb{R}^d$ (counting measure on $\mathcal{G}$) if ξ_{p_r} is sampled continuously (on a grid). For any bounded measurable function of bounded support $g\colon \mathbb{R}^{rd} \to \mathbb{R}$, there exists a reduced cumulant measure $\breve{C}^{(\zeta)}_{p_1,\ldots,p_r}$ such that*

$$\int_{\mathbb{R}^{rd}} g(x_1, \ldots, x_r) C^{(\zeta)}_{p_1,\ldots,p_r}(\mathrm{d}x_1 \times \cdots \times \mathrm{d}x_r)$$
$$= \int_{\mathbb{R}^d} \int_{\mathbb{R}^{(r-1)d}} g(x + u_1, \ldots, x + u_{r-1}, x) \breve{C}^{(\zeta)}_{p_1,\ldots,p_r}(\mathrm{d}u_1 \times \cdots \times \mathrm{d}u_{r-1}) \ell_{p_r}(\mathrm{d}x).$$

Proof. See § 3.3 of the Supplementary Material. □

If all of the processes are sampled on a grid, the fourth-order case becomes

$$\breve{C}^{(\zeta)}_{p,q,r,s}(A_1 \times A_2 \times A_3) = \sum_{u_1 \in \mathcal{G} \cap A_1} \sum_{u_2 \in \mathcal{G} \cap A_2} \sum_{u_3 \in \mathcal{G} \cap A_3} \mathscr{C}[Y_p(u_1), Y_q(u_2), Y_r(u_3), Y_s(0)],$$

so that, in particular, the summand is the usual cumulant function of a stationary multivariate time series or random field; see, for example, Brillinger (1974).

References

Adler, R. J. (2010). *The Geometry of Random Fields*. Philadelphia, PA: Society for Industrial and Applied Mathematics.

Andén, J. & Romero, J. L. (2020). Multitaper estimation on arbitrary domains. *SIAM J. Imag. Sci.* **13**, 1565–94.

Baddeley, A. J., Møller, J. & Waagepetersen, R. (2000). Non-and semi-parametric estimation of interaction in inhomogeneous point patterns. *Statist. Neer.* **54**, 329–50.

Bandyopadhyay, S. & Lahiri, S. N. (2009). Asymptotic properties of discrete Fourier transforms for spatial data. *Sankhya* **71**, 221–59.

Barnett, A. H., Magland, J. & af Klinteberg, L. (2019). A parallel nonuniform fast Fourier transform library based on an 'exponential of semicircle' kernel. *SIAM J. Sci. Comp.* **41**, C479–504.

Bartlett, M. S. (1963). The spectral analysis of point processes. *J. R. Statist. Soc. B* **25**, 264–81.

Bartlett, M. S. (1964). The spectral analysis of two-dimensional point processes. *Biometrika* **51**, 299–311.

Biscio, C. A. N. & Coeurjolly, J.-F. (2016). Standard and robust intensity parameter estimation for stationary determinantal point processes. *Spatial Statist.* **18**, 24–39.

Biscio, C. A. N. & Waagepetersen, R. (2019). A general central limit theorem and a subsampling variance estimator for α-mixing point processes. *Scand. J. Statist.* **46**, 1168–90.

Bonferroni, C. (1936). Teoria statistica delle classi e calcolo delle probabilita. *Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commericiali di Firenze* **8**, 3–62.

Brillinger, D. R. (1965). An introduction to polyspectra. *Ann. Math. Statist.* **36**, 1351–74.

Brillinger, D. R. (1972). The spectral analysis of stationary interval functions. In *Proc. 6th Berkeley Symp. Math. Statist. Prob.*, Ed. L. M. Le Cam, J. Neyman and E. L. Scott, pp. 483–513. Berkeley, CA: University of California Press.

Brillinger, D. R. (1974). *Time Series: Data Analysis and Theory*. New York: Holt, Rinehart and Winston.

Brillinger, D. R. (1982). Asymptotic normality of finite Fourier transforms of stationary generalized processes. *J. Mult. Anal.* **12**, 64–71.

Carter, G. C. (1987). Coherence and time delay estimation. *Proc. IEEE* **75**, 236–55.

Carter, G. C., Knapp, C. & Nuttall, A. (1973). Statistics of the estimate of the magnitude-coherence function. *IEEE Trans. Audio Electroacoust.* **21**, 388–9.

Condit, R., Pérez, R., Aguilar, S., Lao, S., Foster, R. & Hubbell, S. (2019). Complete data from the Barro Colorado 50-ha plot: 423617 trees, 35 years. https://doi.org/10.15146/5xcp-0d46.

Cooley, J. W. & Tukey, J. W. (1965). An algorithm for the machine calculation of complex Fourier series. *Math. Comp.* **19**, 297–301.

Dahlhaus, R. (2000). Graphical interaction models for multivariate time series. *Metrika* **51**, 157–72.

Daley, D. J. & Vere-Jones, D. (2003). *An Introduction to the Theory of Point Processes: Volume I: Elementary Theory and Methods*. New York: Springer.

Daley, D. J. & Vere-Jones, D. (2007). *An Introduction to the Theory of Point Processes: Volume II: General Theory and Structure*. New York: Springer.

Diggle, P. J., Gates, D. J. & Stibbard, A. (1987). A nonparametric estimator for pairwise-interaction point processes. *Biometrika* **74**, 763–70.

Dutt, A. & Rokhlin, V. (1993). Fast Fourier transforms for nonequispaced data. *SIAM J. Sci. Comp.* **14**, 1368–93.

Eckardt, M. & Mateu, J. (2019a). Partial characteristics for marked spatial point processes. *Environmetrics* **30**, e2565.

Eckardt, M. & Mateu, J. (2019b). A spatial dependence graph model for multivariate spatial hybrid processes. *arXiv:* 1906.07798v1.

Eichler, M., Dahlhaus, R. & Sandkühler, J. (2003). Partial correlation analysis for the identification of synaptic connections. *Biol. Cybern.* **89**, 289–302.

Goodman, N. R. (1957). *On the Joint Estimation of the Spectra, Cospectrum and Quadrature Spectrum of a Two-Dimensional Stationary Gaussian Process*. Princeton, NJ: Princeton University Press.

Guillaumin, A. P., Sykulski, A. M., Olhede, S. C. & Simons, F. J. (2022). The debiased spatial Whittle likelihood. *J. R. Statist. Soc. B* **84**, 1526–57.

Harms, K. E., Condit, R., Hubbell, S. P. & Foster, R. B. (2001). Habitat associations of trees and shrubs in a 50-ha neotropical forest plot. *J. Ecol.* **89**, 947–99.

Hubbell, S. P. & Foster, R. B. (1983). Diversity of canopy trees in a neotropical forest and implications for conservation. In *Tropical Rain Forest: Ecology and Management*, Ed. T. Whitmore, A. Chadwick and A. Sutton, pp. 25–41. Oxford: The British Ecological Society.

Hubbell, S. P. & Foster, R. B. (1986). Commonness and rarity in a neotropical forest: implications for tropical tree conservation. In *Conservation Biology: The Science of Scarcity and Diversity*, Ed. M. Soule, pp. 205–31. Sunderland, MA: Sinauer Associates.

Illian, J., Penttinen, A., Stoyan, H. & Stoyan, D. (2008). *Statistical Analysis and Modelling of Spatial Point Patterns*. New York: John Wiley and Sons.

Kanaan, M., Taylor, P. C. & Mugglestone, M. (2008). Cross-spectral properties of a spatial point-lattice process. *Statist. Prob. Lett*. **78**, 3238–43.

Masry, E. (1978). Poisson sampling and spectral estimation of continuous-time processes. *IEEE Trans. Info. Theory* **24**, 173–83.

Masry, E. (2003). Alias-free sampling: an alternative conceptualization and its applications. *IEEE Trans. Info. Theory* **24**, 317–24.

Matsuda, Y. & Yajima, Y. (2009). Fourier analysis of irregularly spaced data on $\mathbb{R}^d$. *J. R. Statist. Soc. B* **71**, 191–217.

Medkour, T., Walden, A. T. & Burgess, A. (2009). Graphical modelling for brain connectivity via partial coherence. *J. Neurosci. Methods* **180**, 374–83.

Miller, K. S. (1980). *Hypothesis Testing with Complex Distributions*. Huntington, NY: R. E. Krieger Publishing.

Møller, J., Syversveen, A. R. & Waagepetersen, R. P. (1998). Log Gaussian Cox processes. *Scand. J. Statist*. **25**, 451–82.

Møller, J. & Waagepetersen, R. P. (2003). *Statistical Inference and Simulation for Spatial Point Processes*. Boca Raton, FL: CRC Press.

Mrkvička, T. & Myllymäki, M. (2023). False discovery rate envelopes. *Statist. Comp*. **33**, 109.

Mugglestone, M. A. & Renshaw, E. (1996a). The exploratory analysis of bivariate spatial point patterns using cross-spectra. *Environmetrics* **7**, 361–77.

Mugglestone, M. A. & Renshaw, E. (1996b). A practical guide to the spectral analysis of spatial point processes. *Comp. Statist. Data Anal*. **21**, 43–65.

Percival, D. B. & Walden, A. T. (2020). *Spectral Analysis for Univariate Time Series (Camb. Ser. Statist. Prob. Math. 51)*. Cambridge: Cambridge University Press.

Rajala, T. A., Olhede, S. C., Grainger, J. P. & Murrell, D. J. (2023). What is the Fourier transform of a spatial point process? *IEEE Trans. Info. Theory* **69**, 5219–52.

Renshaw, E. (2002). Two-dimensional spectral analysis for marked point processes. *Biomet. J*. **44**, 718–45.

Riedel, K. S. & Sidorenko, A. (1995). Minimum bias multiple taper spectral estimation. *IEEE Trans. Sig. Proces*. **43**, 188–95.

Schneider-Luftman, D. & Walden, A. T. (2016). Partial coherence estimation via spectral matrix shrinkage under quadratic loss. *IEEE Trans. Sig. Proces*. **64**, 5767–77.

Simons, F. J. & Wang, D. V. (2011). Spatiospectral concentration in the Cartesian plane. *GEM - Int. J. Geomath.* **2**, 1–36.

Thomson, D. (1982). Spectrum estimation and harmonic analysis. *Proc. IEEE* **70**, 1055–96.

Waagepetersen, R., Guan, Y., Jalilian, A. & Mateu, J. (2016). Analysis of multispecies point patterns by using multivariate log-Gaussian Cox processes. *Appl. Statist*. **65**, 77–96.

Walden, A. T. (2000). A unified view of multitaper multivariate spectral estimation. *Biometrika* **87**, 767–88.

Yang, J. & Guan, Y. (2025). Fourier analysis of spatial point processes. *arXiv:* 2401.06403v3.

[*Received on* 25 *September* 2024. *Accepted on* 21 *November* 2025]