CENTRE FOR **STOCHASTIC GEOMETRY**
AND ADVANCED **BIOIMAGING**

Jesper Møller and Ege Rubak

Determinantal point processes and functional summary statistics on the sphere

No. 02, January 2016

Determinantal point processes and functional summary statistics on the sphere

Jesper Møller and Ege Rubak

Department of Mathematical Sciences, Aalborg University, Denmark
jm@math.aau.dk, rubak@math.aau.dk

Abstract

We discuss the appealing properties of determinantal point process (DPP) models on the d-dimensional unit sphere $\mathbb{S}^d$, considering both the isotropic and the anisotropic case. DPPs are finite point processes exhibiting repulsiveness, but we also use them together with certain dependent thinnings when constructing point process models on $\mathbb{S}^d$ with aggregation on the large scale and regularity on the small scale. Moreover, for general point processes on $\mathbb{S}^d$, we present reduced Palm distributions and functional summary statistics, including nearest neighbour functions, empty space functions, and Ripley's and inhomogeneous K-functions. We conclude with a discussion on future work on statistics for spatial point processes on the sphere.

Keywords: aggregation; empty space function; inhomogeneous K-function; isotropic covariance function; joint intensities; likelihood; nearest neighbour function; Palm distribution; repulsiveness; spectral representation.

1 Introduction

How do we construct spatial point process models on the d-dimensional unit sphere $\mathbb{S}^d \subset \mathbb{R}^{d+1}$ exhibiting regularity/inhibition between the points? Here $d = 1, 2$ are the practically most relevant cases and the regularity/inhibition may be caused by repulsiveness between the points or by some thinning mechanism as specified in the following list of models, usually defined on $\mathbb{R}^d$ but straightforwardly adapted to $\mathbb{S}^d$:

- *Matérn hard core processes* of types I-III can be simulated by their constructions as dependent thinnings of Poisson processes, see [8, 14, 15, 27]. However, for the types I-II, the moments of the process will be tractable while the likelihood (density) will be intractable; and for type III, the opposite is the case.
- *Simple sequential inhibition* and other hard sphere packing models (as reviewed in [3, 9]) can be simulated by their simple constructions of points added one by one, but otherwise they are hard to analyse.

- *Gibbs point processes* offer much more flexibility for modelling inhibition or repulsiveness, but their simulation may be time-consuming and neither the moments nor the likelihood are tractable, see [18, 19] and the references therein.
- *Determinantal point processes (DPPs)* offer relatively flexible models for repulsiveness (although less flexible than Gibbs point processes) and in particular they possess a number of attractive properties, including that they can be easily simulated and the moments and the likelihood are tractable, cf. [7, 10, 11]. DPPs are of interest because of their applications in mathematical physics, combinatorics, random-matrix theory, machine learning, and spatial statistics (see [11] and the references therein).

The first part of the present paper concerns DPPs on the sphere. Simulated examples of DPPs with various degrees of repulsiveness are shown in Figure 1. In comparison with DPPs on $\mathbb{R}^d$, DPPs on $\mathbb{S}^d$ are easier to handle, since they are defined on a compact set and we can more easily deal with certain decompositions of the kernel used in the definition of a DPP (given in Section 2). We restrict attention to DPPs defined on $\mathbb{S}^2$, which apart from a scaling may be considered as an approximation of planet Earth, however, our discussion can easily be extended to DPPs on $\mathbb{S}^1$ and the general case of $\mathbb{S}^d$ may be covered as well. Moreover, partly following [12], we use DPPs together with certain dependent thinnings when constructing point process models with aggregation on the large scale and regularity on the small scale.

The second part of this paper deals with functional summary statistics for first general point processes and second DPPs on the sphere. In the isotropic case of a point process on the sphere, [24] studied Ripley's K-function without providing the mathematical details for the reduced Palm distribution, which is needed in the precise definition of the K-function. We provide this definition without assuming isotropy, so that the inhomogeneous K-function, introduced in [1] for point processes on Euclidean spaces, can be defined on the sphere as well. Moreover, in the isotropic case, we introduce further useful functional summary functions, namely the nearest neighbour function, the empty space function, and the related so-called J-function.

The paper is organized as follows. Section 2, as a supplement to the technical exposition in [16], provides a survey on the attractive properties of DPPs on the sphere, and Section 3 discusses DPP models, both in the isotropic/homogeneous case and in the inhomogeneous case. Section 4 defines Palm distributions and functional summary statistics for general point processes on $\mathbb{S}^2$, in particular nearest neighbour and K-functions which we relate to DPPs, and Section 5 contains our concluding remarks on future work on statistics for spatial point processes on the sphere.

In connection to this paper, the development of software for simulation of DPPs on the sphere and calculation of non-parametric estimators for the functional summary statistics constitutes a substantial amount of work. The software is written in the `R` language [21] and will be available as an extension of the `spatstat` package [2]. Section 4 shows examples of how to use the software and the type of plots it can produce.

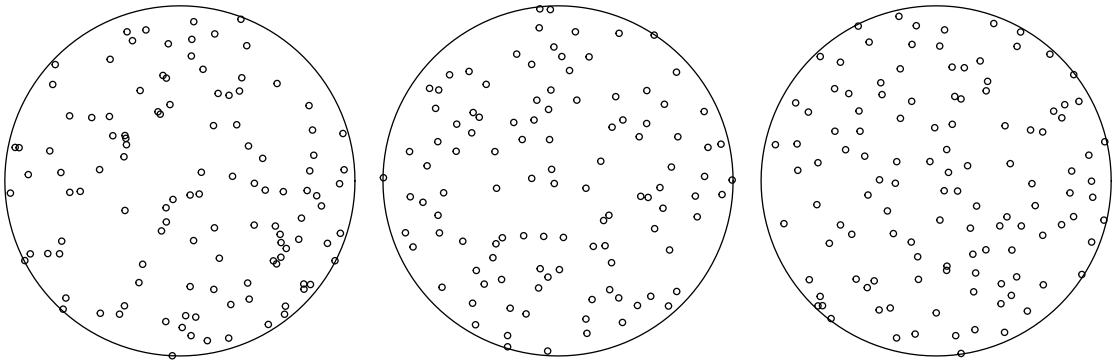

Figure 1: Northern Hemisphere of three spherical point patterns projected to the unit disc with an equal-area azimuthal projection. Each pattern is a simulated realization of a determinantal point process on the sphere with mean number of points 225. Left: Complete spatial randomness (Poisson process). Middle: Multiquadric model with $\tau = 10$ and $\delta = 0.68$. Right: Most repulsive DPP.

2 DPPs on the sphere

Section 2.1 provides the definition of a DPP on $\mathbb{S}^2$, Section 2.2 discusses why it produces regular point patterns, Section 2.3 specifies when it exists, and Section 2.4 deals with its density function with respect to a unit rate Poisson process.

2.1 Definition

Consider a simple finite point process $\boldsymbol{X}$ on $\mathbb{S}^2$, i.e., we can view $\boldsymbol{X}$ as a random finite subset of $\mathbb{S}^2$. Let N be its corresponding counting measure, i.e., $N(A)$ denotes the number of points in $\boldsymbol{X}$ falling in a region $A \subseteq \mathbb{S}^2$. For a given complex function C defined on the product space $\mathbb{S}^2 \times \mathbb{S}^2$, we say that $\boldsymbol{X}$ is a DPP if the factorial moment properties of N can be expressed in terms of certain determinants with entries specified by C as detailed below. An alternative specification in terms of the density for a DPP is given in Section 2.4.

First, we need some notation and a few assumptions. For

$$\boldsymbol{x} = (x_1, x_2, x_3) = (\sin\vartheta\cos\varphi, \sin\vartheta\sin\varphi, \cos\vartheta) \in \mathbb{S}^2 \tag{2.1}$$

where $\vartheta \in [0, \pi]$ is the polar latitude and $\varphi \in [0, 2\pi)$ is the polar longitude, let

$$\mathrm{d}\nu(\boldsymbol{x}) = \sin\vartheta\,\mathrm{d}\varphi\,\mathrm{d}\vartheta \tag{2.2}$$

be the surface measure on $\mathbb{S}^2$. For $n = 1, 2, \ldots$, suppose that $\boldsymbol{X}$ has nth order joint intensity $\rho^{(n)} : (\mathbb{S}^2)^n \mapsto [0, \infty)$ with respect to the n-fold product surface measure $\nu^{(n)} = \nu \times \cdots \times \nu$. That is, for any Borel function $h : (\mathbb{S}^2)^n \mapsto [0, \infty)$,

$$\mathrm{E} \sum_{\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n \in \boldsymbol{X}}^{\neq} h(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) = \int h(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)\rho^{(n)}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)\,\mathrm{d}\nu^{(n)}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) \tag{2.3}$$

where the expectation is with respect to $\boldsymbol{X}$ and $\neq$ over the summation sign means that $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ are pairwise distinct. In particular, $\rho(\boldsymbol{x}) = \rho^{(1)}(\boldsymbol{x})$ is the intensity function (with respect to surface measure) and for any region $A \subseteq \mathbb{S}^2$ the nth order factorial moment of $N(A)$ is

$$\begin{aligned}&\mathrm{E}\left[N(A)(N(A)-1)\cdots(N(A)-n+1)\right]\\ &\quad= \int_A \cdots \int_A \rho^{(n)}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)\,\mathrm{d}\nu(\boldsymbol{x}_1)\cdots\,\mathrm{d}\nu(\boldsymbol{x}_n).\end{aligned}$$

Note that $\rho^{(n)}$ is uniquely determined except on a $\nu^{(n)}$-nullset.

Second, by definition $\boldsymbol{X}$ is a *DPP with kernel* C if for all $n = 1, 2, \ldots$ and all $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n \in \mathbb{S}^2$,

$$\rho^{(n)}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) = \det\left(C(\boldsymbol{x}_i, \boldsymbol{x}_j)_{i,j=1,\ldots,n}\right) \tag{2.4}$$

where $\det\left(C(\boldsymbol{x}_i, \boldsymbol{x}_j)_{i,j=1,\ldots,n}\right)$ is the determinant of the $n \times n$ matrix with (i,j)th entry $C(\boldsymbol{x}_i, \boldsymbol{x}_j)$. Then we write $\boldsymbol{X} \sim \mathrm{DPP}(C)$ and notice the following. The intensity function is the diagonal of the kernel:

$$\rho(\boldsymbol{x}) = C(\boldsymbol{x}, \boldsymbol{x}), \qquad \boldsymbol{x} \in \mathbb{S}^2.$$

The expected number of points is the trace of the kernel:

$$\eta := \mathrm{E}\left[N\left(\mathbb{S}^2\right)\right] = \int C(\boldsymbol{x}, \boldsymbol{x})\,\mathrm{d}\nu(\boldsymbol{x}). \tag{2.5}$$

A Poisson process on $\mathbb{S}^2$ with intensity function ρ is the special case of a DPP where $C(\boldsymbol{x}, \boldsymbol{x}) = \rho(\boldsymbol{x})$ for $\boldsymbol{x} \in \mathbb{S}^2$, and $C(\boldsymbol{x}, \boldsymbol{y}) = 0$ for $\boldsymbol{x} \neq \boldsymbol{y}$.

It follows from (2.4) and since $\rho^{(n)}$ is non-negative that C has to be positive semi-definite. Henceforth, as in most other work on DPPs, we restrict attention to the case where the kernel is Hermitian. In other words, C is a complex covariance function. We allow the kernel to be complex, since this becomes convenient when considering simulation of DPPs, cf. [16]. However, isotropy of C implies that it is real, and all specific models for covariance functions considered in this paper will be real. Moreover, we assume that C is of finite trace class or equivalently that $\eta < \infty$. Finally, we assume that C is square integrable with respect to $\nu^{(2)}$.

2.2 Repulsiveness

A Poisson process is the case of no interaction, i.e., it is neither clustered nor repulsive/inhibitive. Comparing a non-Poissonian DPP with a Poisson process with the same intensity function, the DPP is seen to be repulsive as explained below.

By (2.4) and since C is a covariance function, we have

$$\rho^{(n)}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) \leq \rho(\boldsymbol{x}_1)\cdots\rho(\boldsymbol{x}_n)$$

with equality only if $\boldsymbol{X}$ is a Poisson process with intensity function ρ. Intuitively, if $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ are pairwise distinct points on $\mathbb{S}^d$, then $\rho^{(n)}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)\,\mathrm{d}\nu^{(n)}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$ is the probability that $\boldsymbol{X}$ has a point in each of n infinitesimally small regions on $\mathbb{S}^2$

around $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ and of surface measure $\mathrm{d}\nu(\boldsymbol{x}_1), \ldots, \mathrm{d}\nu(\boldsymbol{x}_n)$, respectively, cf. (2.3). Therefore, a DPP is repulsive unless it is a Poisson process.

The intensity function and the so-called *pair correlation function* g for $\boldsymbol{X}$ play a particular role, since they determine the first and second order moment properties of the counts. Letting

$$R(\boldsymbol{x}, \boldsymbol{y}) = \frac{C(\boldsymbol{x}, \boldsymbol{y})}{\sqrt{C(\boldsymbol{x}, \boldsymbol{x})C(\boldsymbol{y}, \boldsymbol{y})}}, \qquad \boldsymbol{x}, \boldsymbol{y} \in \mathbb{S}^2,$$

be the correlation function corresponding to C when $\rho(\boldsymbol{x})\rho(\boldsymbol{y}) > 0$, then

$$g(\boldsymbol{x}, \boldsymbol{y}) := \frac{\rho^{(2)}(\boldsymbol{x}, \boldsymbol{y})}{\rho(\boldsymbol{x})\rho(\boldsymbol{y})} = 1 - |R(\boldsymbol{x}, \boldsymbol{y})|^2 \qquad \text{if } \rho(\boldsymbol{x})\rho(\boldsymbol{y}) > 0$$

and we set $g(\boldsymbol{x}, \boldsymbol{y}) = 0$ if $\rho(\boldsymbol{x})\rho(\boldsymbol{y}) = 0$. Thus $g \leq 1$, again showing that a DPP is repulsive, since for a Poisson process with intensity function ρ, $g(\boldsymbol{x}, \boldsymbol{y}) = 1$ whenever $\rho(\boldsymbol{x})\rho(\boldsymbol{y}) > 0$.

2.3 Existence

By [7, Lemma 4.2.6 and Theorem 4.5.5] DPP(C) exists if and only if the spectrum of C is bounded by 0 and 1, and then DPP(C) is unique. Below we explain what this means.

Consider a complex covariance function $K : \mathbb{S}^d \times \mathbb{S}^d \mapsto \mathbb{C}$ which is of finite trace class and is square integrable (e.g. the kernel C of the DPP considered above satisfies these conditions, and other cases will be considered later). Then, by Mercer's theorem (e.g. [22, Section 98]) and [7, Lemma 4.2.2]), ignoring a $\nu^{(2)}$-nullset, K has a spectral representation

$$K(\boldsymbol{x}, \boldsymbol{y}) = \sum_{n=1}^{\infty} \alpha_n Y_n(\boldsymbol{x})\overline{Y_n(\boldsymbol{y})}, \qquad \boldsymbol{x}, \boldsymbol{y} \in \mathbb{S}^d \tag{2.6}$$

where $Y_1, Y_2, \ldots$ are eigenfunctions which form an orthonormal basis for the space of square integrable complex functions with respect to ν. We call (2.6) the *Mercer representation* of K, the eigenvalues α_i for the *Mercer coefficients*, and $\mathrm{spec}(K) = \{\alpha_1, \alpha_2, \ldots\}$ the spectrum of K.

When $\boldsymbol{X} \sim \mathrm{DPP}(C)$, we denote the Mercer coefficients of C by $\lambda_1, \lambda_2, \ldots$, and to ensure existence we require that $\mathrm{spec}(C) \subset [0, 1]$. Then, by (2.5) and (2.6), the mean number of points is $\eta = \sum_{n=1}^{\infty} \lambda_n$.

2.4 Likelihood

When the spectrum of the kernel for a DPP is strictly bounded by one, we can work with the likelihood/density as given below.

Suppose $\boldsymbol{X} \sim \mathrm{DPP}(C)$ where $\mathrm{spec}(C) \subset [0, 1)$. Let $\tilde{C} : \mathbb{S}^2 \times \mathbb{S}^2 \mapsto \mathbb{C}$ be the complex covariance function with a Mercer representation sharing the same eigenfunctions as C but with Mercer coefficients

$$\tilde{\lambda}_n = \frac{\lambda_n}{1 - \lambda_n}, \qquad n = 1, 2, \ldots.$$

Define

$$D = \sum_{n=1}^{\infty} \log\bigl(1 + \tilde{\lambda}_n\bigr).$$

Then, by [26, Theorem 1.5], DPP(C) is absolutely continuous with respect to the unit rate Poisson process (i.e., the Poisson process on $\mathbb{S}^2$ with intensity measure ν) and has density

$$f(\{\boldsymbol{x}_1, \dots, \boldsymbol{x}_n\}) = \exp(4\pi - D) \det\bigl(\tilde{C}(\boldsymbol{x}_i, \boldsymbol{x}_j)_{i,j=1,\dots,n}\bigr), \tag{2.7}$$

for any point configuration $\{\boldsymbol{x}_1, \dots, \boldsymbol{x}_n\} \subset \mathbb{S}^2$, $n = 0, 1, \dots$

Note that we consider the empty point configuration $\emptyset$ if $n = 0$, so $\exp(-D)$ is the probability that $\boldsymbol{X} = \emptyset$ (since $\exp(-4\pi)$ is the probability that the unit rate Poisson process is empty). Further, since

$$\lambda_n = \frac{\tilde{\lambda}_n}{1 + \tilde{\lambda}_n}, \qquad n = 1, 2, \dots, \tag{2.8}$$

there is a one-to-one correspondence between C and $\tilde{C}$. Thus, in order to construct a DPP we can start by specifying any complex covariance function $\tilde{C}$ which is of finite trace class and is square integrable. Its density is then given by (2.7), and C is determined by the Mercer representation of $\tilde{C}$ and by (2.8). Then $\mathrm{spec}(C) \subset [0, 1)$, and so DPP($C$) is well-defined.

3 DPP models on the sphere

Section 3.1 discusses the construction of isotropic DPPs on $\mathbb{S}^2$, i.e., DPPs those kernels are invariant under rotations. Section 3.2 shows how independent and dependent thinnings of such models can be used to construct anisotropic DPPs and related point processes on the sphere.

3.1 Isotropic/homogeneous DPPs

3.1.1 Characterization of isotropic kernels

Assume that $\boldsymbol{X} \sim \mathrm{DPP}(C)$ where the kernel is isotropic, i.e.,

$$C(\boldsymbol{x}, \boldsymbol{y}) = C_0(s), \quad s = s(\boldsymbol{x}, \boldsymbol{y}) = \arccos(\boldsymbol{x} \cdot \boldsymbol{y}), \qquad \boldsymbol{x}, \boldsymbol{y} \in \mathbb{S}^d.$$

Here s denotes geodesic (or orthodromic or great-circle) distance, and $\cdot$ is the usual inner product on $\mathbb{R}^3$. Further, the assumption that C is a covariance function implies that C_0 is a real function defined on $[0, \pi]$ such that C is positive semi-definite. We follow [5] in calling C_0 the *radial part* of C, and we slightly abuse notation and write $\boldsymbol{X} \sim \mathrm{DPP}(C_0)$. Denote $\mathcal{O}(3)$ the orthogonal group, i.e., the set of all 3×3 matrices O so that $OO^\top = O^\top O = I$, where $O^\top$ is the transpose of O and I is the 3×3 identity matrix. Note that $\boldsymbol{X}$ is invariant in distribution under the action of $\mathcal{O}(3)$ on $\mathbb{S}^2$, and we shall say that $\boldsymbol{X}$ is an *isotropic/homogeneous DPP*. In particular any point in $\boldsymbol{X}$ is uniformly distributed on $\mathbb{S}^2$, the intensity $\rho = C_0(0)$ is constant and

equal to the maximal value of C_0, and $\eta = 4\pi\rho$ is the expected number of points in $\boldsymbol{X}$.

Assume that $\rho > 0$ (otherwise $\boldsymbol{X} = \emptyset$). Then

$$R_0(s) = C_0(s)/\rho \tag{3.1}$$

is the radial part of the correlation function R associated to C. The allowed range of ρ in terms of R_0 is the interval from 0 to

$$\rho_{\max}(R_0) = 1/\|R\| \tag{3.2}$$

where $\|R\| < \infty$ denotes the largest eigenvalue of R. Furthermore, the pair correlation function is isotropic and given by

$$g(\boldsymbol{x}, \boldsymbol{y}) = g_0(s) = 1 - R_0(s)^2. \tag{3.3}$$

This implies that $g_0(0) = 0$ (however, in case of a Poisson process, it is custom to set $g_0(0) = 1$, since $\rho^{(2)}(\boldsymbol{x}, \boldsymbol{y}) = \rho^2$ for $\nu^{(2)}$ almost all $(\boldsymbol{x}, \boldsymbol{y}) \in \mathbb{S}^2 \times \mathbb{S}^2$).

Before characterizing the radial part of an isotropic covariance function we need to recall the following:

$$P_\ell(x) = \frac{1}{2^\ell \ell!} \frac{\mathrm{d}^\ell}{\mathrm{d}x^\ell}\{(x^2-1)^\ell\}, \qquad -1 < x < 1,$$

is the Legendre polynomial of degree $\ell = 0, 1, \ldots$, and for $k = 0, \ldots, \ell$, the associated Legendre functions $P_\ell^{(k)}$ and $P_\ell^{(-k)}$ are given by

$$P_\ell^{(k)}(x) = (-1)^k \left(1 - x^2\right)^{k/2} \frac{\mathrm{d}^k}{\mathrm{d}x^k} P_\ell(x), \qquad -1 \le x \le 1,$$

and

$$P_\ell^{(-k)} = (-1)^k \frac{(\ell-k)!}{(\ell+k)!} P_\ell^{(k)}.$$

Moreover, the surface spherical harmonic functions are given by

$$Y_{\ell,k}(\boldsymbol{x}) = Y_{\ell,k}(\vartheta, \varphi) = \sqrt{\frac{2\ell+1}{4\pi} \frac{(\ell-k)!}{(\ell+k)!}}\, P_\ell^{(k)}(\cos\vartheta)\, e^{ik\varphi}, \quad (\vartheta, \varphi) \in [0, \pi] \times [0, 2\pi),$$

for $k = -\ell, \ldots, \ell$, where $\boldsymbol{x} \in \mathbb{S}^2$ is identified by its polar latitude and longitude (ϑ, φ), cf. (2.1), and where $i^2 = -1$. In fact the surface spherical harmonic functions constitute an orthonormal basis for the space of square integrable complex functions with respect to ν.

Now, by a classical result of [25], C_0 being the radial part of a *continuous* isotropic covariance function C is equivalent to assume that

$$C_0(s) = \sum_{\ell=0}^{\infty} \frac{2\ell+1}{4\pi} \alpha_\ell P_\ell(\cos s), \qquad s \in [0, \pi], \tag{3.4}$$

where each $\alpha_\ell \geq 0$ and $\sum_{\ell=0}^{\infty}(2\ell+1)\alpha_\ell < \infty$. Equivalently, by the addition formula for spherical harmonics (see [16]), the Mercer representation becomes

$$C(\boldsymbol{x}_1, \boldsymbol{x}_2) = \sum_{\ell=0}^{\infty} \alpha_\ell \sum_{k=-\ell}^{\ell} Y_{\ell,k}(\boldsymbol{x}_1)\overline{Y_{\ell,k}(\boldsymbol{x}_2)}, \qquad \boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{S}^2, \tag{3.5}$$

i.e., the Mercer coefficients are $\lambda_{\ell,k} = \alpha_\ell$, with $\ell = 0, 1, \ldots$ and $k = -\ell, \ldots, \ell$. Therefore, to ensure that DPP(C_0) is well-defined, we require that the spectrum $\{\alpha_0, \alpha_1, \ldots\}$ is included in $[0, 1]$ and that

$$\eta = \sum_{\ell=0}^{\infty}(2\ell+1)\alpha_\ell$$

is finite.

Some remarks are in order.

- Similarly, if we use the alternative approach of Section 2.4 where we start by specifying $\tilde{C}$: Assuming that $\tilde{C}$ is a *continuous* isotropic covariance function is equivalent to that

$$\tilde{C}(\boldsymbol{x}_1, \boldsymbol{x}_2) = \sum_{\ell=0}^{\infty} \tilde{\alpha}_\ell Y_{\ell,k}(\boldsymbol{x}_1)\overline{Y_{\ell,k}(\boldsymbol{x}_2)}, \qquad \boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{S}^2, \tag{3.6}$$

  where all $\tilde{\alpha}_\ell$ are non-negative and $\sum_{\ell=0}^{\infty}(2\ell+1)\tilde{\alpha}_\ell < \infty$.

- By (3.1) and (3.4),

$$R_0(s) = \frac{1}{4\pi}\sum_{\ell=0}^{\infty} \beta_\ell P_\ell(\cos s), \qquad s \in [0, \pi],$$

  where

$$\beta_\ell = (2\ell+1)\alpha_\ell/\eta, \qquad \ell = 0, 1, \ldots,$$

  is a discrete probability distribution. Conversely, given a continuous correlation function R_0, i.e., given the sequence $\beta_0, \beta_1, \ldots$, (3.2) gives the upper bound on the expected number of points:

$$\eta_{\max}(R_0) = \inf\{(2\ell+1)/\beta_\ell : \ell = 0, 1, \ldots\}.$$

- In [16] we quantify global respective local repulsiveness in terms of the pair correlation function when the intenity is fixed, and we point out that there is a *trade-off between intensity and the degree of repulsiveness*. Loosely speaking the degree of repulsiveness increases as the spectrum of the kernel C tends to a step function which for small indices ℓ is one and for larger indices ℓ is zero. Therefore, for any integer $m \geq 0$, we refer to a DPP with kernel (3.5) such that $\alpha_\ell = 1$ for $\ell \leq m$ and $\alpha_\ell = 0$ for $\ell > m$ as the *most repulsive (isotropic) DPP* with $\eta = (m+1)^2$ (since in this case $\eta = \sum_{\ell=0}^{m}(2\ell+1) = (m+1)^2$; see [16] for a definition when η is any positive number). The Poisson process is another extreme obtained when the spectrum tends to zero (but ρ is still fixed). The right panel of Figure 1 shows a realization of the most repulsive DPP when $\eta = 15^2 = 225$.

3.1.2 Examples

Consider the *multiquadric family* [6] given by

$$C_0 = \rho R_0, \quad R_0(s) = \frac{(1-\delta)^{2\tau}}{(1+\delta^2-2\delta\cos s)^\tau}, \quad 0 < \rho \le \rho_{\max}(R_0),\ \tau > 0,\ 0 < \delta < 1.$$

As detailed in [16], the eigenvalues α_ℓ can easily be calculated numerically, which makes it possible to simulate realizations from this model. In Section 4.3 below we furthermore derive a closed form expression for the K-function for the model, which we will use for statistical tests, and in future work it can also be used for parameter estimation (as discussed later in Section 5). In [16] we show that the model is quite flexible and covers the range from no to intermediate repulsiveness, but in general it does not cover the most repulsive DPP (only when the expected number of points is very low). The middle panel of Figure 1 shows a realization of a multiquadric model where first we fixed $\eta = 225$ and $\tau = 10$, and then $\delta = 0.68$ was chosen as the smallest value such that the model is well-defined. For practical purposes this corresponds to the most repulsive multiquadric model (the degree of repulsiveness grows as τ grows and δ decreases). For comparison a realization of a Poisson process with $\eta = 225$ is shown in the left panel, and it is easy to visually confirm that the multiquadric model has a higher degree of repulsiveness than the Poisson process.

In the special case of $\tau = 1/2$ we obtain the *inverse multiquadric family* where $\beta_\ell = \delta^\ell(1-\delta)$ specifies a geometric distribution and

$$\alpha_\ell = \eta\delta^\ell(1-\delta)/(2\ell+1), \quad \eta_{\max}(R_0) = 1/(1-\delta).$$

To notice the trade-off between the intensity and the degree of repulsiveness, observe that $\eta_{\max}(R_0)$ is a strictly increasing function of δ with range $(1,\infty)$, while since $R_0(s)$ with $s \neq 0$ is a strictly decreasing function of δ, the DPP becomes less repulsive as δ increases. In the limit as $\delta \to 1$ we obtain $R_0(s) = 0$ corresponding to a Poisson process; notice that $\alpha_\ell \to 0$ for $\ell = 0, 1, \ldots$ On the other hand, as $\delta \to 0$ and if $\eta = \eta_{\max}(R_0)$ we obtain the most repulsive DPP but with the mean number of points only equal to 1. As demonstrated in [16], even for $\eta = 10$ the DPP is rather far away from the most repulsive DPP, and for $\eta = 100$ it is rather close to a Poisson process. So the inverse multiquadric family may be of limited interest except for theoretical considerations.

For the inverse multiquadric model both the kernel and the Mercer coefficients are expressible on closed form, while in the general multiquadric model the Mercer coefficients lend themselves to relatively simple numerical evaluation. This is a rather unique case, and in [16] we consider a number of other models and conclude that the most useful approach for obtaining flexible parametric models that cover the full range of possible repulsiveness for DPPs is a direct modelling of the spectrum (i.e., the Mercer coefficients). One example of a flexible model is the case

$$\alpha_\ell = \frac{1}{1+\beta\exp\left((\ell/\alpha)^\kappa\right)}, \qquad \ell = 0, 1, \ldots,$$

where $\alpha > 0$, $\beta > 0$, and $\kappa > 0$ are parameters. Since all $\alpha_\ell \in (0,1)$, the DPP is well defined and has a density specified by (2.7), while η may be evaluated by numerical methods. As demonstrated in [16], the model covers a wide range of repulsive DPPs, including any homogeneous Poisson process and any most repulsive DPP.

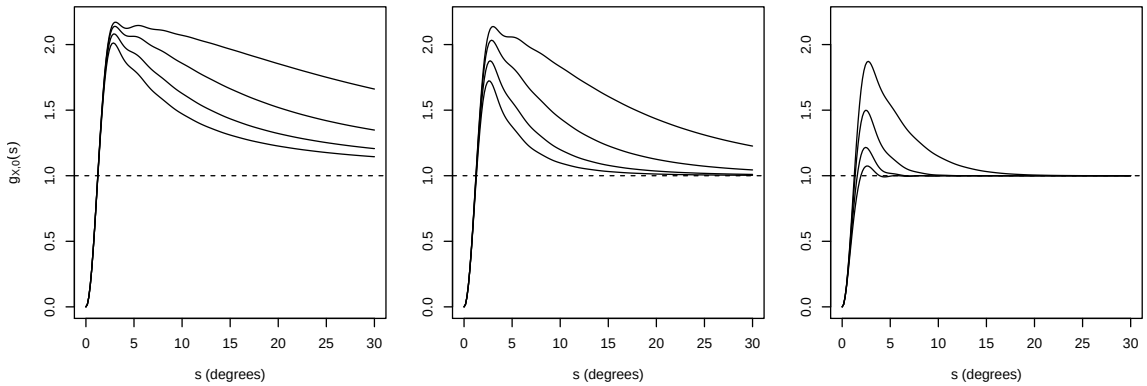

Figure 2: Pair correlation functions for Π-thinnings of a most repulsive DPP with 400 points (see Figure 3 for a realization). In all panels the covariance function for the underlying Gaussian process is multiquadric with variance $\kappa = 8$, while $\tau = 0.25, 1, 10$ from left to right panel, and $\delta = 0.5, 0.7, 0.8, 0.85$ from top to bottom curve within each panel.

3.2 Anisotropic models

This section focuses on *anisotropic/inhomogeneous DPP* $\boldsymbol{X}$ constructed by independent thinning of an isotropic/homogeneous DPP $\boldsymbol{Y}$ on $\mathbb{S}^2$ with kernel C_Y and nth order product intensity $\rho_Y^{(n)}$. We also follow [12] in considering a doubly stochastic construction where $\boldsymbol{X}$ is obtained by a dependent thinning of $\boldsymbol{Y}$. Thereby we can model regularity on the small scale and clustering on the large scale.

Suppose

$$\boldsymbol{X} = \{\boldsymbol{x} \in \boldsymbol{Y} : \Pi(\boldsymbol{x}) \geq U(\boldsymbol{x})\}$$

where $\boldsymbol{\Pi} = \{\Pi(\boldsymbol{x}) : \boldsymbol{x} \in \mathbb{S}^2\}$ is a random process of 'selection probabilities' $\Pi(\boldsymbol{x})$, $\mathbf{U} = \{U(\boldsymbol{x}) : \boldsymbol{x} \in \mathbb{S}^2\}$ is a process of mutually independent random variables $U(\boldsymbol{x})$ which are uniformly distributed on $[0, 1]$, and $\boldsymbol{Y}, \boldsymbol{\Pi}, \mathbf{U}$ are mutually independent. If $\boldsymbol{\Pi}$ is deterministic, then $\boldsymbol{X}$ is an *independent thinning* of $\boldsymbol{Y}$, having nth order product intensity

$$\rho_X^{(n)}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) = \Pi(\boldsymbol{x}_1) \cdots \Pi(\boldsymbol{x}_n) \rho_Y^{(n)}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$$

and so $\boldsymbol{X}$ is seen to be a DPP with kernel

$$C_X(\boldsymbol{x}, \boldsymbol{y}) = \Pi(\boldsymbol{x}) \Pi(\boldsymbol{y}) C_Y(\boldsymbol{x}, \boldsymbol{y}).$$

If $\boldsymbol{\Pi}$ is random, then $\boldsymbol{X}$ is in general a *dependent thinning* of $\boldsymbol{Y}$, with

$$\rho_X^{(n)}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) = \mathrm{E}\left[\Pi(\boldsymbol{x}_1) \cdots \Pi(\boldsymbol{x}_n)\right] \rho_Y^{(n)}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n).$$

In the latter case, $\boldsymbol{X}$ is not a DPP unless the selection probabilities are independent.

In particular assume that $\boldsymbol{Y}$ is homogeneous with intensity ρ_Y and pair correlation function $g_{0,Y}(s)$, and the distribution of $\boldsymbol{\Pi}$ is invariant under the action of $\mathcal{O}(3)$ on $\mathbb{S}^2$. Then $\boldsymbol{X}$ is homogeneous, with intensity and pair correlation function

$$\rho_X = q\rho_Y, \qquad g_{X,0}(s) = M_0(s) g_{Y,0}(s), \qquad s \in [0, \pi],$$

where $q = \mathrm{E}[\Pi(\boldsymbol{x})]$ is the mean selection probability and, setting $0/0 = 0$,

$$M_0(s) = M(\boldsymbol{x}, \boldsymbol{y}) = \frac{\mathrm{E}[\Pi(\boldsymbol{x})\Pi(\boldsymbol{y})]}{\mathrm{E}[\Pi(\boldsymbol{x})]\mathrm{E}[\Pi(\boldsymbol{y})]}, \qquad \boldsymbol{x}, \boldsymbol{y} \in \mathbb{S}^2,$$

depends only on $s = s(\boldsymbol{x}, \boldsymbol{y})$. For instance, assume that $-\log \boldsymbol{\Pi}$ is the χ^2-process given by

$$\Pi(\boldsymbol{x}) = \exp\left(-\mathbf{Z}(\boldsymbol{x})^2/2\right), \qquad \boldsymbol{x} \in \mathbb{S}^2, \tag{3.7}$$

where $\mathbf{Z}$ is a zero-mean Gaussian process with isotropic covariance function K. Denoting K_0 the radial part of $\mathbf{K}_0$ and assuming the variance $\kappa = K_0(0)$ is positive, we have

$$q = (1+\kappa)^{-1/2}, \quad M_0(s) = \left[1 - \frac{R_0(s)^2}{(1+1/\kappa)^2}\right]^{-1/2},$$

where $R_0 = K_0/\kappa$ is the (radial part of the) correlation function of $\mathbf{Z}$. Note that $g_{Y,0} \leq 1$ while $M_0(s) \geq 1$ is typically a decreasing function of s. In fact it is possible to obtain that $g_{X,0}(s) \leq 1$ for small values of s and $g_{X,0}(s) \geq 1$ for large values of s, reflecting regularity on the small scale and clustering on the large scale.

This is illustrated in Figure 2, where the original process $\mathbf{Y}$ is a most repulsive DPP with 400 points and the underlying Gaussian process $\mathbf{Z}$ has a multiquadric covariance function with variance $\kappa = 8$ such that the mean selection probability is $q = 1/3$. Thus the expected number of points of the thinned process $\mathbf{X}$ is $400/3$. As can be seen from the figure, both τ and δ influence the range of the positive association between points on the longer scale: For both parameters smaller values yield long range dependence while the dependence dies out quicker for larger values. Similar figures (not shown here) show that changing the original DPP to a multiquadric model effectively shifts the curves such that the value of s where $g_{X,0}(s)$ crosses the Poisson reference value 1 shifts to the left, which is to be expected since the original DPP is less repulsive in this case. Note that the geodesic distance in this and subsequent figures is given in terms of the angle between points on the sphere measured in degrees from 0 to 180 (as we expect the reader to relate more easily to these than distances which are effectively in radians).

Note that simulation of $\mathbf{Z}$ is easy, if we assume that K has a Mercer representation as in (3.5), with eigenvalues α_ℓ^Z. Then we generate independent standard normally distributed random variables $W_{\ell,k}^{(1)}$ and $W_{\ell,k}^{(2)}$ for $\ell = 0, 1, \ldots$ and $k = -\ell, \ldots, \ell$, and observe that

$$\sum_{\ell=0}^{\infty} \sqrt{\alpha_\ell^Z} \sum_{k=-\ell}^{\ell} \left\{ W_{\ell,k}^{(1)} \mathrm{Re}\left[Y_{\ell,k}(\boldsymbol{x})\right] + W_{\ell,k}^{(2)} \mathrm{Im}\left[Y_{\ell,k}(\boldsymbol{x})\right] \right\}, \qquad \boldsymbol{x} \in \mathbb{S}^2, \tag{3.8}$$

is a zero-mean Gaussian process with covariance function K (this follows from a straightforward calculation, using (3.5) and the fact that K is real). In practice a truncation of the infinite series in (3.8) has to be used. From (3.4) we have $\kappa = K_0(0) = \sum_0^\infty \alpha_\ell^Z (2\ell+1)/(4\pi)$, and we choose the truncation such that the truncated series equals 99% of the given value of κ. Figure 3 shows a realization of the original unthinned DPP $\mathbf{Y}$ while Figure 4 shows the result after Π-thinning.

Finally, we notice another construction, namely by applying a one-to-one smooth transformation $\mathbb{S}^2 \mapsto \mathbb{S}^2$ on $\boldsymbol{Y}$ to obtain $\boldsymbol{X}$. This result again in that $\boldsymbol{X}$ is a DPP

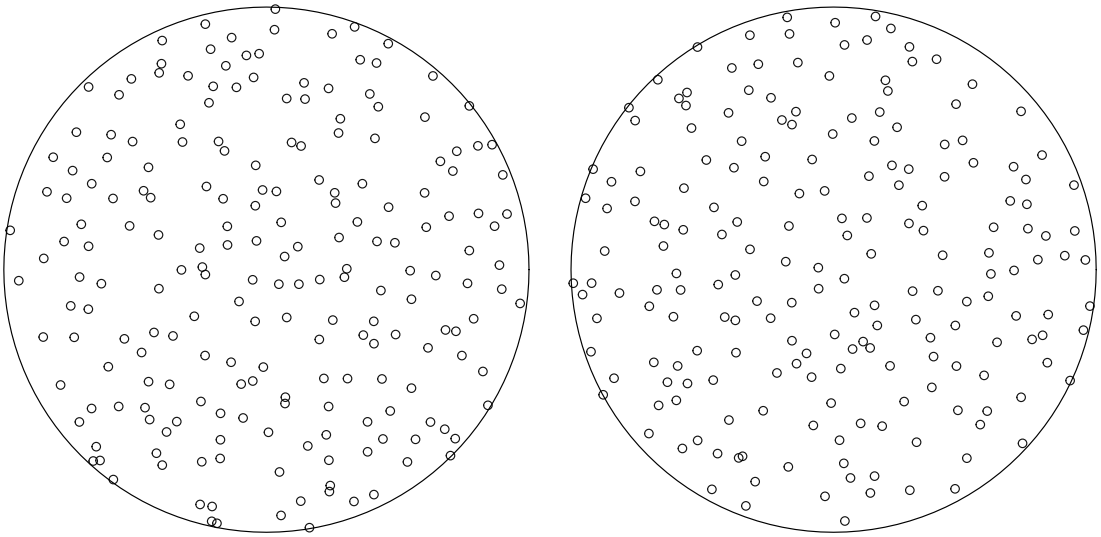

Figure 3: Northern and Southern Hemispheres of a most repulsive DPP with 400 points used as the original process Y before Π-thinning (see text). The hemispheres are projected to unit discs with an equal-area azimuthal projection.

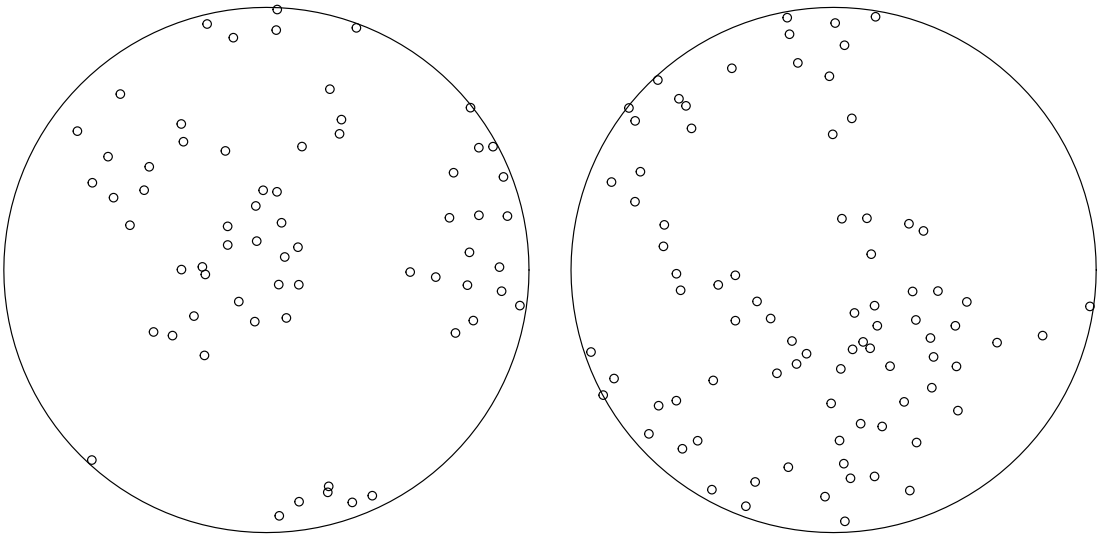

Figure 4: Northern and Southern Hemispheres of a Π-thinning of the most repulsive DPP with 400 points shown in Figure 3. The Gaussian process underlying the Π-thinning has multiquadric covariance function with $\kappa = 8$, $\tau = 1$, and $\delta = 0.5$. The hemispheres are projected to unit discs with an equal-area azimuthal projection and the pattern contains 148 points in total.

those kernel can be specified in terms of C_Y and the derivative of the transformation. We skip the details here, but see [10, 11] for the result in the case of transformed DPPs on $\mathbb{R}^d$.

4 Palm distributions and functional summary statistics

The most popular functional summary statistic for a stationary point process on $\mathbb{R}^d$ with intensity ρ is Ripley's K-function [23], where $\rho K(t)$ is interpreted as the mean number of further points within distance t of a typical point in the process. The formal definition of K requires Palm measure theory, and its definition can be extended to the inhomogeneous K-function for so-called second order intensity reweighted stationary point processes [1, 4]. The adaption of Ripley's K-function to a general isotropic point process on the sphere is given in [24], without explicitly specifying the reduced Palm distribution which is needed in the precise definition. Section 4.1 gives a definition of this distribution without assuming isotropy and Section 4.2 provides the formal definition of Ripley's and the inhomogeneous K-function and the nearest-neighbour function on the sphere, together with various useful interpretations and results for non-parametric estimation. Moreover, Section 4.3 relates all this to DPPs.

4.1 Palm distribution for a general point process on the sphere

Suppose $\boldsymbol{X}$ is a general point process on $\mathbb{S}^2$ with an integrable intensity function ρ. The so-called Campbell-Mecke formula gives that for any $\boldsymbol{x} \in \mathbb{S}^2$ there exists a finite point process $\boldsymbol{X}_{\boldsymbol{x}}^!$ so that

$$\mathrm{E} \sum_{\boldsymbol{x} \in \boldsymbol{X}} h(\boldsymbol{x}, \boldsymbol{X} \setminus \{\boldsymbol{x}\}) = \int \mathrm{E} h(\boldsymbol{x}, \boldsymbol{X}_{\boldsymbol{x}}^!) \rho(\boldsymbol{x}) \, \mathrm{d}\nu(\boldsymbol{x}) \tag{4.1}$$

for any non-negative Borel function h. Moreover, the distribution of $\boldsymbol{X}_{\boldsymbol{x}}^!$ is unique for ν almost all $\boldsymbol{x}$ with $\rho(\boldsymbol{x}) > 0$, and it is called the *reduced Palm distribution* at the point $\boldsymbol{x}$. If $\rho(\boldsymbol{x}) = 0$, this distribution can be chosen to be arbitrary, since this case will play no role in this paper.

Intuitively, $\boldsymbol{X}_{\boldsymbol{x}}^!$ follows the conditional distribution of $\boldsymbol{X} \setminus \{\boldsymbol{x}\}$ given that $\boldsymbol{x} \in \boldsymbol{X}$. This interpretation follows from (4.1) or perhaps more easily by assuming that $\boldsymbol{X}$ has a density f (with respect to the unit rate Poisson process on $\mathbb{S}^2$) since, for $\rho(\boldsymbol{x}) > 0$, $\boldsymbol{X}_{\boldsymbol{x}}^!$ has density $f_{\boldsymbol{x}}(\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}) = f(\{\boldsymbol{x}, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\})/\rho(\boldsymbol{x})$.

The situation simplifies if the distribution of $\boldsymbol{X}$ is isotropic. Then $\rho(\boldsymbol{x}) = \rho$ is constant. Denote $\mathcal{SO}(3)$ the 3D rotation group, i.e., $O \in \mathcal{SO}(3)$ if and only if $O \in \mathcal{O}(3)$ and $\det O = 1$. Let $\mathbf{e} = (0, 0, 1)$ be the North Pole (for the following considerations, any other fixed point on the sphere can be used instead). For $\boldsymbol{x} \in \mathbb{S}^2 \setminus \{\mathbf{e}\}$, there is a unique $R_{\boldsymbol{x}} \in \mathcal{SO}(3)$ so that $R_{\boldsymbol{x}}\mathbf{e} = \boldsymbol{x}$ and its axis of rotation is

orthogonal to the circle on $\mathbb{S}^2$ which contains $\boldsymbol{x}$ and $\mathbf{e}$. Define $R_{\mathbf{e}} = I$. Assume $\rho > 0$ and let $\boldsymbol{Y}_{\mathbf{e}}^!$ be a point process on $\mathbb{S}^2$ with distribution

$$\mathrm{P}\left(\boldsymbol{Y}_{\mathbf{e}}^! \in F\right) = \frac{1}{4\pi\rho}\,\mathrm{E}\sum_{\boldsymbol{x}\in\boldsymbol{X}} 1\left[R_{\boldsymbol{x}}^\top(\boldsymbol{X}\setminus\{\boldsymbol{x}\}) \in F\right], \tag{4.2}$$

where $1[\cdot]$ denotes the indicator function. Thus

$$\mathrm{E}k\left(\boldsymbol{Y}_{\mathbf{e}}^!\right) = \frac{1}{4\pi\rho}\,\mathrm{E}\sum_{\boldsymbol{x}\in\boldsymbol{X}} k\left(R_{\boldsymbol{x}}^\top(\boldsymbol{X}\setminus\{\boldsymbol{x}\})\right)$$

for non-negative measurable functions k. Suppose h is a non-negative measurable function such that for any $O \in \mathcal{SO}(3)$, with probability one,

$$h(\boldsymbol{x}, \boldsymbol{X}\setminus\{\boldsymbol{x}\}) = h(O\boldsymbol{x}, O\boldsymbol{X}\setminus\{O\boldsymbol{x}\}) \qquad \text{for all } \boldsymbol{x}\in\boldsymbol{X} \tag{4.3}$$

and $\mu(\mathbb{S}^2) < \infty$, where

$$\mu(A) := \mathrm{E}\sum_{\boldsymbol{x}\in\boldsymbol{X}\cap A} h(\boldsymbol{x}, \boldsymbol{X}\setminus\{\boldsymbol{x}\}), \qquad A\subseteq\mathbb{S}^2.$$

Then it is straightforwardly verified that μ is a rotation invariant measure, and hence

$$\mathrm{E}\sum_{\boldsymbol{x}\in\boldsymbol{X}\cap A} h(\boldsymbol{x}, \boldsymbol{X}\setminus\{\boldsymbol{x}\}) = \frac{\nu(A)}{4\pi}\,\mathrm{E}\sum_{\boldsymbol{x}\in\boldsymbol{X}} h(\boldsymbol{x}, \boldsymbol{X}\setminus\{\boldsymbol{x}\}). \tag{4.4}$$

In particular, suppose k is a non-negative measurable function such that for any $O \in \mathcal{SO}(3)$, with probability one,

$$k\left(R_{\boldsymbol{x}}^\top\left(\boldsymbol{X}\setminus\{\boldsymbol{x}\}\right)\right) = k\left(R_{O\boldsymbol{x}}^\top\left(O\boldsymbol{X}\setminus\{O\boldsymbol{x}\}\right)\right) \qquad \text{for all } \boldsymbol{x}\in\boldsymbol{X}. \tag{4.5}$$

Then $h(\boldsymbol{x}, \boldsymbol{X}\setminus\{\boldsymbol{x}\}) = k(R_{\boldsymbol{x}}^\top(\boldsymbol{X}\setminus\{\boldsymbol{x}\}))$ satisfies (4.3), and so it follows from (4.2) and (4.4) that

$$\rho\nu(A)\mathrm{E}k\left(\boldsymbol{Y}_{\mathbf{e}}^!\right) = \mathrm{E}\sum_{\boldsymbol{x}\in\boldsymbol{X}\cap A} k\left(R_{\boldsymbol{x}}^\top(\boldsymbol{X}\setminus\{\boldsymbol{x}\})\right). \tag{4.6}$$

Finally, combining (4.1) and (4.6), we see that for ν almost all $\boldsymbol{x}\in\mathbb{S}^2$,

$$\mathrm{E}k\left(R_{\boldsymbol{x}}^\top\boldsymbol{X}_{\boldsymbol{x}}^!\right) = \mathrm{E}k\left(\boldsymbol{Y}_{\mathrm{e}}^!\right). \tag{4.7}$$

Therefore, when restricting attention to the case (4.5), we can let $\boldsymbol{X}_{\mathbf{e}}^!$ be distributed as $\boldsymbol{Y}_{\mathrm{e}}^!$ and let $\boldsymbol{X}_{\boldsymbol{x}}^!$ be distributed as $R_{\boldsymbol{x}}\boldsymbol{X}_{\mathbf{e}}^!$.

In light of the conditions and results (4.3)-(4.7) it is natural to ask if the definition (4.2) could be replaced by

$$\mathrm{P}\left(\boldsymbol{Y}_{\mathbf{e}}^! \in F\right) = \frac{1}{\nu(A)\rho}\,\mathrm{E}\sum_{\boldsymbol{x}\in\boldsymbol{X}\cap A} 1\left[R_{\boldsymbol{x}}^\top(\boldsymbol{X}\setminus\{\boldsymbol{x}\}) \in F\right] \tag{4.8}$$

for an arbitrary set $A\subseteq\mathbb{S}^2$ with $\nu(A) > 0$. Clearly, this is not the case: Let e.g. $F = \{\{\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n\}\subset\mathbb{S}^2 : s(\boldsymbol{x},\mathbf{e}) \le 1, n\ge 0\}$ and $A = \{\boldsymbol{x}\in\mathbb{S}^2 : s(\boldsymbol{x},\mathbf{e}) > 1\}$. Then in (4.8) the left hand side may be positive while the right hand side is 0. Similarly, in general, we cannot let $\boldsymbol{X}_{\mathbf{e}}^!$ be distributed as $\boldsymbol{Y}_{\mathrm{e}}^!$ and let $\boldsymbol{X}_{\boldsymbol{x}}^!$ be distributed as $R_{\boldsymbol{x}}\boldsymbol{X}_{\mathbf{e}}^!$, since (4.1) with $h(\boldsymbol{x}, \boldsymbol{X}\setminus\{\boldsymbol{x}\}) = 1[\boldsymbol{x}\in A, \boldsymbol{X}\setminus\in F]$ would then lead to a contradiction.

4.2 Functional summary statistics for isotropic and second order intensity reweighted isotropic point processes on the sphere

4.2.1 The homogeneous case

Assume $\boldsymbol{X}$ is an isotropic point process on $\mathbb{S}^2$ with finite intensity $\rho > 0$. Let $s(A, B) = \inf_{\boldsymbol{x}\in A, \boldsymbol{y}\in B} s(\boldsymbol{x}, \boldsymbol{y})$ be the shortest geodesic distance between $A, B \subset \mathbb{S}^2$ and define the *nearest neighbour function* by

$$G(t) = \mathrm{P}\left(s(\mathbf{Y}^!_{\mathbf{e}}, \mathbf{e}) \leq t\right), \qquad 0 \leq t \leq \pi.$$

Since

$$k\left(R^\top_{\boldsymbol{x}}\left(\boldsymbol{X} \setminus \{\boldsymbol{x}\}\right)\right) = 1\left[s\left(R^\top_{\boldsymbol{x}}(\boldsymbol{X} \setminus \{\boldsymbol{x}\}), \mathbf{e}\right) \leq t\right] = 1\left[s(\boldsymbol{X} \setminus \{\boldsymbol{x}\}, \boldsymbol{x}) \leq t\right]$$

satisfies (4.5), (4.7) gives

$$G(t) = \mathrm{P}\left(s(\boldsymbol{X}^!_{\boldsymbol{x}}, \boldsymbol{x}) \leq t\right),$$

i.e., G is the distribution function for the distance from a typical point to the nearest other point in $\boldsymbol{X}$. Furthermore, for an arbitrary point $\boldsymbol{x} \in \mathbb{S}^2$, we define the *empty space function* by

$$F(t) = \mathrm{P}\left(s(\boldsymbol{X}, \boldsymbol{x}) \leq t\right), \qquad 0 \leq t \leq \pi,$$

and following [13], we define the J-function by

$$J(t) = \frac{1 - G(t)}{1 - F(t)} \qquad \text{for } F(t) < 1.$$

Since $\boldsymbol{X}$ is isotropic, F does not depend on the choice of $\boldsymbol{x}$. If $\boldsymbol{X}$ is a homogeneous Poisson process, then $F_{\mathrm{Pois}}(t) = G_{\mathrm{Pois}}(t) = \exp(-2\pi\rho(1 - \cos t))$ and $J_{\mathrm{Pois}} = 1$.

Note that (4.6) yields

$$\rho\nu(A)G(t) = \mathrm{E}\sum_{\boldsymbol{x}\in\boldsymbol{X}\cap A} 1\left[s(\boldsymbol{X} \setminus \{\boldsymbol{x}\}, \boldsymbol{x}) \leq t\right].$$

Thinking of A as an observation window, this is a useful result when deriving non-parametric estimates: Let $\mathcal{G} \subset \mathbb{S}^2$ be a finite grid of $m > 0$ points. If $\boldsymbol{X}$ is fully observed on $\mathbb{S}^2$, i.e., $A = \mathbb{S}^2$, then natural estimates are

$$\widehat{F}(t) = \frac{1}{m}\sum_{\boldsymbol{x}\in\mathcal{G}} 1\left[s(\boldsymbol{X}, \boldsymbol{x}) \leq t\right]$$

and

$$\widehat{G}(t) = \frac{1}{n}\sum_{\boldsymbol{x}\in\boldsymbol{X}} 1\left[s(\boldsymbol{X} \setminus \{\boldsymbol{x}\}, \boldsymbol{x}) \leq t\right]$$

provided $N(\mathbb{S}^2) = n > 0$. In case the observation window A is a proper subset of $\mathbb{S}^2$, minus sampling may be used: Let $A_{\ominus t} = \{\boldsymbol{x} \in A : s(\mathbb{S}^2 \setminus A, \boldsymbol{x}) > t\}$ be the set of

those points in A with geodesic distance at least t to any point outside A. Then minus sampling gives the estimate

$$\widehat{G}(t) = \frac{1}{N(A_{\ominus t})} \sum_{\boldsymbol{x} \in \boldsymbol{X} \cap A_{\ominus t}} 1[s(\boldsymbol{X}, \boldsymbol{x}) \le t]$$

provided $N(A_{\ominus t}) > 0$. Moreover, for $\widehat{F}$ we choose the grid so that $\mathcal{G} \subset A_{\ominus t}$.

Now, we define the *K-function* by

$$K(t) = \frac{1}{\rho} \mathrm{E} \sum_{\boldsymbol{x} \in \boldsymbol{Y}_{\mathbf{e}}^{!}} 1\left[s(\mathbf{e}, \boldsymbol{x}) \le t\right], \qquad 0 \le t \le \pi. \tag{4.9}$$

Since

$$k\left(R_{\boldsymbol{x}}^{\top}\left(\boldsymbol{X} \setminus \{\boldsymbol{x}\}\right)\right) = \sum_{\boldsymbol{y} \in R_{\boldsymbol{x}}^{\top}(\boldsymbol{X} \setminus \{\boldsymbol{x}\})} 1[s(\mathbf{e}, \boldsymbol{y}) \le t] = \sum_{\boldsymbol{y} \in \boldsymbol{X} \setminus \{\boldsymbol{x}\}} 1[s(\boldsymbol{x}, \boldsymbol{y}) \le t]$$

satisfies (4.5), (4.7) implies that

$$\rho K(t) = \mathrm{E} \sum_{\boldsymbol{y} \in \boldsymbol{X}_{\boldsymbol{x}}^{!}} 1[s(\boldsymbol{x}, \boldsymbol{y}) \le t] \tag{4.10}$$

is the mean number of further points within geodetic distance t of a typical point in the process. Furthermore, by (4.6) and (4.9), we obtain

$$\rho^2 \nu(A) K(t) = \mathrm{E} \sum_{\boldsymbol{x} \in \boldsymbol{X} \cap A} \sum_{\boldsymbol{y} \in \boldsymbol{X} \setminus \{\boldsymbol{x}\}} 1\left[s(\boldsymbol{x}, \boldsymbol{y}) \le t\right], \tag{4.11}$$

which is another useful formula for deriving non-parametric estimates. For example, if $\boldsymbol{X}$ is fully observed on $\mathbb{S}^2$,

$$\widehat{K}(t) = \frac{4\pi}{N(N-1)} \sum_{\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X}}^{\neq} 1\left[s(\boldsymbol{x}, \boldsymbol{y}) \le t\right] \tag{4.12}$$

is a natural estimate. If instead the observation window A is a proper subset of $\mathbb{S}^2$, minus sampling gives

$$\widehat{K}(t) = \frac{\nu(A)}{N(A)(N(A)-1)} \sum_{\substack{\boldsymbol{x} \in \boldsymbol{X} \cap A \\ \boldsymbol{y} \in \boldsymbol{X} \cap A_{\ominus t}}}^{\neq} 1\left[s(\boldsymbol{x}, \boldsymbol{y}) \le t\right]$$

provided $N(A) > 1$.

The estimate (4.12) was also suggested in [24], where plots for all values of $t \in [0, \pi]$ were considered. Apart from the case of Poisson models we warn against such plots for the following reason. When $\boldsymbol{X}$ has pair correlation function g_0, we have

$$K(t) = 2\pi \int_0^t g_0(s) \sin s \, \mathrm{d}s, \qquad 0 \le t \le \pi, \tag{4.13}$$

cf. (2.2)-(2.3), (3.3), and (4.11). For an isotropic/homogeneous Poisson process, the pair correlation function is $g_{\text{Pois}} = 1$, so the K-function is $K_{\text{Pois}}(t) = 2\pi(1 - \cos t)$. Thus using (4.12) gives $\widehat{K}(\pi) = K_{\text{Pois}}(\pi) = 4\pi$, but for non-Poissonian models $\widehat{K}(\pi)$ may be seriously biased, since K is an accumulative function of g_0, cf. (4.13). For example, if the pair correlation function for $\boldsymbol{X}$ is smaller than one (as in the case of a DPP), we may have $\widehat{K}(t) \gg K(t)$ for large values of t; we illustrate this in Section 4.2.3 below. Therefore we recommend only interpreting plots of $\widehat{K}(t)$ for smaller values of t: If for smaller or modest values of t, $\widehat{K}(t)$ is below (above) $K_{\text{Pois}}(t)$, then we interpret this as inhibition or repulsiveness (aggregation or clustering) between nearby points in $\boldsymbol{X}$. This interpretation is just like in the case of planar point processes. Incidentally, a second order Taylor approximation around $t = 0$ gives $K_{\text{Pois}}(t) \approx \pi t^2$, where πt^2 is the K-function for a planar Poisson process. Similar, when interpreting plots of non-parametric estimates of $F(t), G(t), J(t)$, we focus on the behavior for small and modest values of t. We refer to Section 4.3 for examples of how to interpret plots of the functional summary statistics. In case of the G-function as compared to G_{Pois}, the interpretation is similar to that of the K-function.

4.2.2 The inhomogeneous case

Assume $\boldsymbol{X}$ is an inhomogeneous point process on $\mathbb{S}^2$ with finite intensity function $\rho(\boldsymbol{x})$ and an isotropic pair correlation function, i.e., it is of the form (3.3). Then, in accordance with [1] we say that $\boldsymbol{X}$ is *second order intensity reweighted isotropic (or pseudo/correlation isotropic)* and define the *inhomogeneous K-function* by (4.13). So this definition of K is in accordance with the isotropic case. In particular, if $\boldsymbol{X}$ is a Poisson process, then it is second order intensity reweighted isotropic and we still have $K_{\text{Pois}}(t) = 2\pi(1 - \cos t)$.

In analogy with (4.11), we have

$$\nu(A)K(t) = \mathrm{E} \sum_{\boldsymbol{x} \in \boldsymbol{X} \cap A} \sum_{\boldsymbol{y} \in \boldsymbol{X} \setminus \{\boldsymbol{x}\}} \frac{1\left[s(\boldsymbol{x}, \boldsymbol{y}) \le t\right]}{\rho(\boldsymbol{x})\rho(\boldsymbol{y})}. \tag{4.14}$$

If the intensity function is known or estimated and $\boldsymbol{X}$ is fully observed on $\mathbb{S}^2$, this suggests the non-parametric estimate

$$\widehat{K}(t) = \sum_{\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X}}^{\neq} \frac{1\left[s(\boldsymbol{x}, \boldsymbol{y}) \le t\right]}{4\pi\rho(\boldsymbol{x})\rho(\boldsymbol{y})},$$

cf. [1]. If instead the observation window A is a proper subset of $\mathbb{S}^2$, minus sampling gives

$$\widehat{K}(t) = \sum_{\boldsymbol{x} \in \boldsymbol{X} \cap A,\, \boldsymbol{y} \in \boldsymbol{X} \cap A_{\ominus t}}^{\neq} \frac{1\left[s(\boldsymbol{x}, \boldsymbol{y}) \le t\right]}{\nu(A)\rho(\boldsymbol{x})\rho(\boldsymbol{y})}.$$

Finally, by (4.1) and (4.14), for ν almost all $\boldsymbol{x} \in \mathbb{S}^2$ with $\rho(\boldsymbol{x}) > 0$,

$$K(t) = \mathrm{E} \sum_{\boldsymbol{y} \in \boldsymbol{X}_{\boldsymbol{x}}^!} \frac{1[s(\boldsymbol{x}, \boldsymbol{y}) \le t]}{\rho(\boldsymbol{y})}.$$

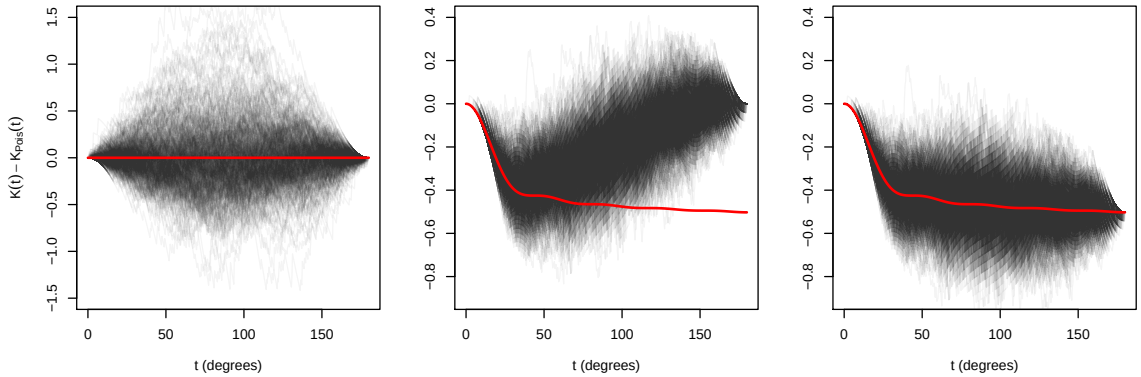

Figure 5: Each panel shows $\hat{K}(t) - K_{\text{Pois}}(t)$ for 500 simulated point patterns on S^2 with 25 points on average together with the theoretical value of $K(t) - K_{\text{Pois}}(t)$ for the model (red line). Left: Poisson model and usual non-parametric estimator (4.12). Middle: Most repulsive DPP and usual non-parametric estimator (4.12). Right: Most repulsive DPP and modified non-parametric estimator (4.15).

Hence, if $\rho(\boldsymbol{y})$ is close to $\rho(\boldsymbol{x})$ for $s(\boldsymbol{x}, \boldsymbol{y}) \leq t$,

$$\rho(\boldsymbol{x})K(t) \approx \mathrm{E} \sum_{\boldsymbol{y} \in \boldsymbol{X}^!_{\boldsymbol{x}}} 1[s(\boldsymbol{x}, \boldsymbol{y}) \leq t],$$

which is a local version of the interpretation of $K(t)$ in the isotropic case, cf. (4.10).

4.2.3 Normalization of $\widehat{K}$

The non-parametric estimate given in (4.12) effectively corresponds to estimating ρ^2 by $N(N-1)/(4\pi)^2$. As previously mentioned, this implies $\widehat{K}(\pi) = K_{\text{Pois}}(\pi) = 4\pi$ making it the natural choice for Poisson models. In general it is easy to show that $K(\pi) \geq 4\pi - 1/\rho$, with equality for models with a fixed number of points such as a most repulsive DPP. The lower bound value $K(\pi) = 4\pi - 1/\rho$ would be obtained by the non-parametric estimator if ρ^2 was estimated by $(N/(4\pi))^2$ instead such that

$$\widehat{K}(t) = \frac{4\pi}{N^2} \sum_{\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X}}^{\neq} 1\left[s(\boldsymbol{x}, \boldsymbol{y}) \leq t\right]. \tag{4.15}$$

Thus, in the special case of a model with a non-random number of points this estimator may be better, but in general as the true model is unknown we prefer to use (4.12).

Figure 5 illustrates the potential bias for a most repulsive DPP with 25 points (the low number of points helps to emphasize the bias since the error in this case is $1/\rho = 1/(25/(4\pi)) \approx 0.5$). Each panel shows the difference between a non-parametric estimate of K and the theoretical value for a Poisson process K_{Pois} based on 500 simulated point patterns on S^2. For reference the left panel shows the perfectly unbiased result obtained when simulating a Poisson process with 25 points on average and using (4.12) to estimate K. The middle panel shows the bias when (4.12) is used in the case of a most repulsive DPP with 25 points. The right panel shows how the bias is removed if (4.15) is used instead.

In the case of a DPP the bias problem of the non-parametric estimator for large distances is best illustrated in the somewhat special case of very low intensity, but we believe the critique and bias problem remains valid for for many other model classes. In particular we expect that models for clustering may attain values of $K(\pi)$ much larger than 4π and suffer from much larger bias, but it is left as an open problem to investigate this further.

4.3 The case of DPPs on the sphere

Let $\boldsymbol{X}$ be an isotropic DPP with an explicit model for the kernel $C_0(s) = \rho R_0(s)$. Then the pair correlation function is given by (3.3) and the K-function is given by (4.13). In particular for the multiquadric model of Section 3.1.2 we can easily derive a closed-form expression for the K-function using integration by substitution:

$$K_{\mathrm{mq}}(s) = K_{\mathrm{Pois}}(s) - 2\pi \frac{(1+\delta)(1-\delta)}{2\delta(1-2\tau)} \Big(\Big(\frac{1+\delta^2-2\delta}{1+\delta^2-2\delta\cos(s)} \Big)^{2\tau-1} - 1 \Big), \quad \tau \neq 1/2, \tag{4.16}$$

while for the inverse multiquadric family we get

$$K_{\mathrm{imq}}(s) = K_{\mathrm{Pois}}(s) - 2\pi \frac{(1+\delta)(1-\delta)}{2\delta} \ln \left(\frac{1+\delta^2-2\delta\cos(s)}{1+\delta^2-2\delta} \right), \quad \tau = 1/2. \tag{4.17}$$

The formulae (4.16)-(4.17) also hold for the inhomogeneous K-function when X is a correlation isotropic DPP obtained by independent thinning of an isotropic DPP with multiquadric kernel.

For DPP models where we have explicit expressions for the Mercer coefficients but not for the kernel, we can use (3.3)-(3.4) to calculate g_0 numerically, and then use numerical integration to calculate the K-function, cf. (4.13). For instance, this approach has to be used both for the flexible model mentioned in Section 3.1.2 and for the most repulsive DPP, and it was used to produce the theoretical (red) curve in the middle and right panels of Figure 5.

The three point patterns in Figure 1 are realizations of DPPs with different degrees of repulsiveness: From left to right, there is none (Poisson DPP), intermediate (multiquadric DPP), and strong (most repulsive DPP) interaction. In the following we will discuss the corresponding non-parametric estimates of K and G and to what extend these can be used to discriminate between the three cases.

Figure 6 shows the non-parametric estimates of K and G for all three patterns along with the theoretical curve for a Poisson process, and as expected the estimates generally have smaller values for the more repulsive models. To produce this figure with the developed software, e.g. for the multiquadric DPP, we simulate a realization of the model and estimate K by using the following commands:

```
mqmodel <- dppMQ(lambda = 225/(4*pi), delta = 0.68, tau = 10)
Xmq     <- simulate(mqmodel)
Kmq     <- Kspp(Xmq, rmax = 10, angle = TRUE)
plot(Kmq)
```

A simple way to assess the difference between the summary statistics is to use pointwise envelopes simulated under the null model, which is a technique with a

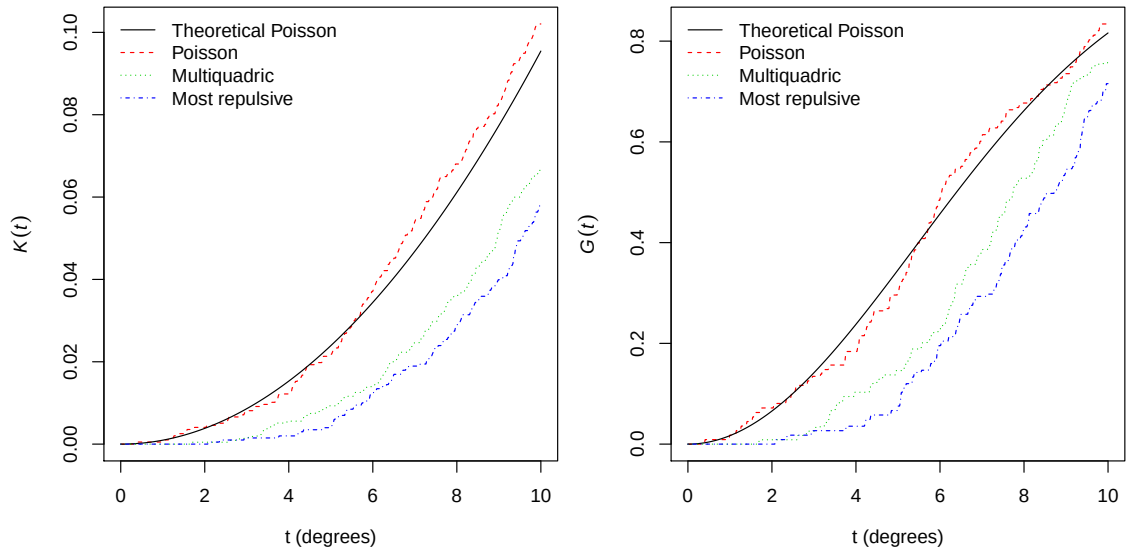

Figure 6: Non-parametric estimates of K (left) and G (right) for the three point patterns in Figure 1 and the theoretical curves for a Poisson process as reference. The abscissa is the angle in degrees between pairs of points on the sphere.

long history for point patterns in Euclidean space (see e.g. [2] for an accessible account). For example, in order to generate pointwise envelopes for the K-function, with significance level 1% for the realization of a multiquadric DPP generated above under a Poisson null model, and plot the results (not shown here), we use the commands

```
envmq <- envelope(Xmq, Kspp, nsim = 199)
plot(envmq)
```

This means that if we fix a distance a priori and reject the null hypothesis if the non-parametric estimate of the summary statistic for the data is outside the envelopes at this distance, then this is a test with significance level 1%. However, the main drawback is that in practice it is very hard to only do a pointwise test when the envelopes show the test results at many scales at once. This problem has been well-known for decades and a recent account can be found in Chapter 10 of [2]. As an alternative to this approach so-called rank count envelopes were developed in [20] which have an interpretation as a global test, while still providing a graphical output that can be used to infer the spatial scales where the data significantly deviates from the null model. An extra advantage of this approach is that several functional summary statistics can be combined in one test to give an overall correct significance level and thereby avoid any multiple testing problems. This test was performed on the K- and G-function simultaneously for the multiquadric point pattern with a Poisson null model which based on 2499 simulations yielded a highly significant p-value of 0.0004, which is the lowest possible p-value based on 2500 summary functions (2499 simulated and 1 data). The corresponding graphical test is shown in Figure 7, where we have separated the values related to K and G into separate plots even though the calculation of the envelopes are based on concatenating the values of K and G into one long vector. Notice that more significant departures from the null

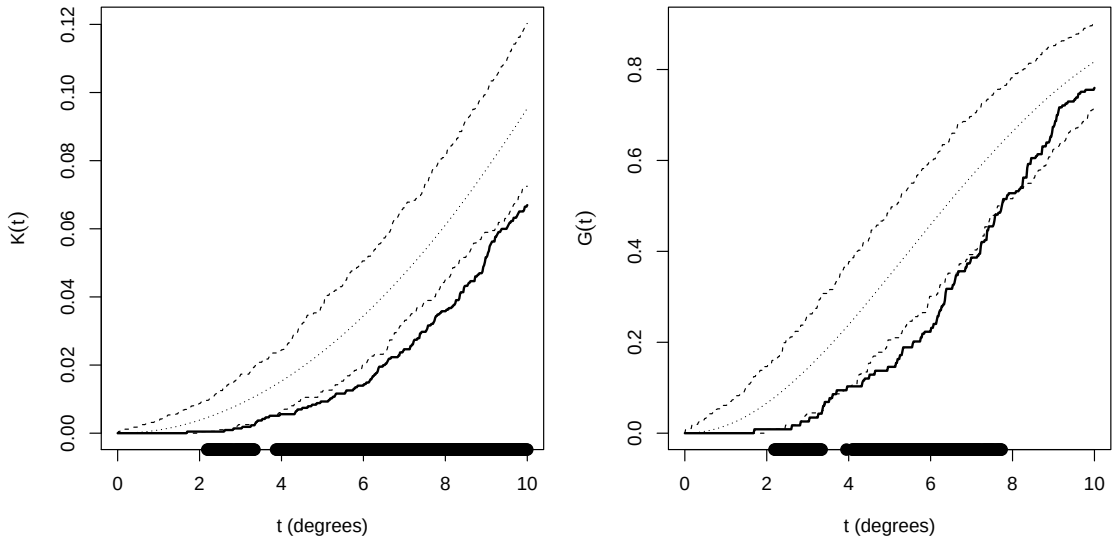

Figure 7: Simultaneous rank count envelopes for K (left) and G (right) for the multiquadric point pattern in the middle panel of Figure 1 based on 2499 simulations from the Poisson null model. Filled circles on the abscissa correspond to distances (in terms of angles in degrees) where the data curve exits the envelopes indicating significant departure from the Poisson null model.

model are detected by the K-function which appears to provide a more powerful test in this case (and in our experience this also applies to the other examples in this paper). While it is useful to know the spatial scales leading to rejection of the null hypothesis, we should be very cautious when interpreting the K-function due to the cumulative nature of the function, cf. Section 4.2.

If we use the same test against the most repulsive DPP as the null model based on 999 simulations, the p-value is 0.001 (which is the lowest possible value based on 1000 summary functions). Finally, if we test the most repulsive pattern (right panel in Figure 1) against the multiquadric DPP model with $\eta = 225$, $\delta = 0.68$, and $\tau = 10$, we get a p-value of 0.008. If instead we use K respective G only for the rank count test, we obtain a p-value of 0.005 respective 0.106, which again indicates that K is the more powerful of the two.

5 Future work

In this paper we have considered examples of simulated point patterns on the sphere under various DPP models. Indeed it would be interesting to analyze real point pattern data sets on the sphere using parametric DPP models. Here we expect that inhomogeneous/anisotropic DPPs will be of more relevance than homogeneous/isotropic DPPs. As in [10, 11] parameter estimation may be done by either maximum likelihood or a composite likelihood or minimum contrast method based on the intensity and pair correlation functions. In [10, 11] we noticed that the latter methods work quite well in comparison with maximum likelihood.

Other point process models than DPPs on the sphere may of course be of rel-

evance for applications. For instance, the spectral representation (2.6) allows us to construct and simulate Gaussian processes, cf. (3.8). Thus we can also deal with log Gaussian Cox processes (LGCPs) on the sphere, where all the statistical methodology for LGCPs on Euclidean spaces [17, 18, 19] can be easily adapted to the sphere.

In Section 4.2, for a second order intensity reweighted isotropic point process on the sphere, we provided non-parametric estimates of the F, G and K-functions. In the literature for point processes defined on Euclidean spaces there is considerable discussion of edge correction factors, which account for the edge effects that arise when estimating functional summary statistics near the boundaries of an observation window. In Section 4.2, we exemplified this only in the case when the process is fully observed or when minus sampling is used, while we leave it for another paper to derive further edge correction factors. In the planar case [2] mentions that "So long as some kind of edge correction is performed ..., the particular choice of edge correction technique is usually not critical." We expect the situation to be similar for point patterns on the sphere.

Finally, we notice that space-time point process models on the sphere, whether being DPPs or LGCPs or of another type, might be worth studying, where of course the direction of time should be taken into consideration.

Acknowledgment

Supported by the Danish Council for Independent Research | Natural Sciences, grant 12-124675, "Mathematical and Statistical Analysis of Spatial Data", and by the "Centre for Stochastic Geometry and Advanced Bioimaging", funded by grant 8721 from the Villum Foundation.

References

[1] A. Baddeley, J. Møller, and R. Waagepetersen. Non- and semi-parametric estimation of interaction in inhomogeneous point patterns. *Statistica Neerlandica*, 54:329–350, 2000.

[2] A. Baddeley, E. Rubak, and R. Turner. *Spatial Point Patterns: Methodology and Applications with R*. Chapman & Hall/CRC Press, Boca Raton, 2015.

[3] S. N. Chiu, D. Stoyan, W. S. Kendall, and J. Mecke. *Stochastic Geometry and Its Applications*. Wiley, Chichester, third edition, 2013.

[4] J.-F. Coeurjolly, J. Møller, and R. Waagepetersen. Conditioning in spatial point processes. Technical report, Centre for Stochastic Geometry and Advanced Bioimaging. Submitted for journal publication, 2015.

[5] D. Daley and E. Porcu. Dimension walks through Schoenberg spectral measures. *Proceedings of the American Mathematical Society*, 42:1813–1824, 2013.

[6] T. Gneiting. Strictly and non-strictly positive definite functions on spheres. *Bernoulli*, 19:1327–1349, 2013.

[7] J. B. Hough, M. Krishnapur, Y. Peres, and B. Viràg. *Zeros of Gaussian Analytic Functions and Determinantal Point Processes*. American Mathematical Society, Providence, 2009.

[8] M. L. Huber and R. L. Wolpert. Likelihood-based inference for Matérn type-III repulsive point processes. *Advances in Applied Probability*, 41:958–977, 2009.

[9] J. Illian, A. Penttinen, H. Stoyan, and D. Stoyan. *Statistical Analysis and Modelling of Spatial Point Patterns*. John Wiley and Sons, Chichester, 2008.

[10] F. Lavancier, J. Møller, and E. Rubak. Determinantal point process models and statistical inference: Extended version. Technical report, available at arXiv:1205.4818, 2014.

[11] F. Lavancier, J. Møller, and E. Rubak. Determinantal point process models and statistical inference. *Journal of Royal Statistical Society: Series B (Statistical Methodology)*, 77:853–877, 2015.

[12] F. Lavencier and J. Møller. Modelling aggregation on the large scale and regularity on the small scale in spatial point pattern datasets. *Scandinavian Journal of Statistics*, 2016. To appear.

[13] M. N. M. van Lieshout and A. J. Baddeley. A nonparametric measure of spatial interaction in point patterns. *Statistica Neerlandica*, 50:344–361, 1996.

[14] B. Matérn. *Spatial Variation*. Lecture Notes in Statistics 36, Springer-Verlag, Berlin, 1986.

[15] J. Møller, M. L. Huber, and R. L. Wolpert. Perfect simulation and moment properties for the Matérn type III process. *Stochastic Processes and their Applications*, 120:2142–2158, 2010.

[16] J. Møller, M. Nielsen, E. Porcu, and E. Rubak. Determinantal point processes on the sphere. Technical report, Centre for Stochastic Geometry and Advanced Bioimaging. Submitted for journal publication, 2015.

[17] J. Møller, A. R. Syversveen, and R. P. Waagepetersen. Log Gaussian Cox processes. *Scandinavian Journal of Statistics*, 25:451–482, 1998.

[18] J. Møller and R. P. Waagepetersen. *Statistical Inference and Simulation for Spatial Point Processes*. Chapman and Hall/CRC, Boca Raton, 2004.

[19] J. Møller and R. P. Waagepetersen. Modern spatial point process modelling and inference (with discussion). *Scandinavian Journal of Statistics*, 34:643–711, 2007.

[20] M. Myllymäki, T. Mrkvicka, P. Grabarnik, H. Seijo, and U. Hahn. Global envelope tests for spatial processes, 2015. Preprint on arxiv:1307.0239v4.

[21] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2015.

[22] F. Riesz and B. Sz.-Nagy. *Functional Analysis*. Dover Publications, New York, 1990.

[23] B. D. Ripley. The second-order analysis of stationary point processes. *Journal of Applied Probability*, 13:255–266, 1976.

[24] S. M. Robeson, A. Li, and C. Huang. Point-pattern analysis on the sphere. *Spatial Statistics*, 10:76–86, 2014.

[25] I. J. Schoenberg. Positive definite functions on spheres. *Duke Mathematical Journal*, 9:96–108, 1942.

[26] T. Shirai and Y. Takahashi. Random point fields associated with certain Fredholm determinants I: fermion, Poisson and boson point processes. *Journal of Functional Analysis*, 205:414–463, 2003.

[27] J. Teichmann, F. Ballani, and K. van den Boogaart. Generalizations of Matérn's hard-core point processes. *Spatial Statistics*, 3:33–53, 2013.