

Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

# Spatial Statistics

journal homepage: [www.elsevier.com/locate/spasta](http://www.elsevier.com/locate/spasta)

## Testing for complete spatial randomness on three dimensional bounded convex shapes<sup>☆</sup>

Scott Ward<sup>a</sup>, Edward A.K. Cohen<sup>a,\*</sup>, Niall Adams<sup>a,b</sup><sup>a</sup> Department of Mathematics, Imperial College London, South Kensington, London, SW7 2AZ, United Kingdom<sup>b</sup> Data Science Institute, Imperial College London, South Kensington, London, SW7 2AZ, United Kingdom

### ARTICLE INFO

#### Article history:

Received 3 February 2020

Received in revised form 30 November 2020

Accepted 18 December 2020

Available online 4 January 2021

#### Keywords:

Complete spatial randomness

Convex shapes

Functional summary statistics

Poisson point processes

### ABSTRACT

There is currently a gap in theory for point patterns that lie on the surface of objects, with researchers focusing on patterns that lie in a Euclidean space, typically planar and spatial data. Methodology for planar and spatial data thus relies on Euclidean geometry and is therefore inappropriate for analysis of point patterns observed in non-Euclidean spaces. Recently, there has been extensions to the analysis of point patterns on a sphere, however, many other shapes are left unexplored. This is in part due to the challenge of defining the notion of *stationarity* for a point process existing on such a space due to the lack of rotational and translational isometries. Here, we construct functional summary statistics for Poisson processes defined on convex shapes in three dimensions. Using the Mapping Theorem, a Poisson process can be transformed from any convex shape to a Poisson process on the unit sphere which has rotational symmetries that allow for functional summary statistics to be constructed. We present the first and second order properties of such summary statistics and demonstrate how they can be used to construct a test statistics to determine whether an observed pattern exhibits complete spatial randomness or spatial preference on the original convex space. We compare this test statistic with one constructed from an analogue  $L$ -function for inhomogeneous point processes on the sphere. A study of the

<sup>☆</sup> This document is the results of the research project funded by Wellcome Trust under the grant 203799/Z/16/Z.

\* Corresponding author.

E-mail addresses: [scott.ward12@imperial.ac.uk](mailto:scott.ward12@imperial.ac.uk) (S. Ward), [e.cohen@imperial.ac.uk](mailto:e.cohen@imperial.ac.uk) (E.A.K. Cohen), [n.adams@imperial.ac.uk](mailto:n.adams@imperial.ac.uk) (N. Adams).

Type I and II errors of our test statistics are explored through simulations on ellipsoids of varying dimensions.

© 2021 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license

(<http://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Research in spatial statistics has predominantly concentrated on the development of theory and methodology for point processes on  $\mathbb{R}^d$ , with a significant focus on planar ( $\mathbb{R}^2$ ) and spatial ( $\mathbb{R}^3$ ) data. Point processes existing on non-Euclidean spaces, however, are still relatively under-explored. Recently, with the advent of spatial data on a global scale, and modelling Earth as a sphere, there have been important developments in the theory and analysis of point processes on the surface of  $d - 1$  dimensional unit spheres,  $\mathbb{S}^{d-1} \subset \mathbb{R}^d$  (Lawrence et al., 2016; Møller and Rubak, 2016; Robeson et al., 2014). Yet patterns can still arise for which these methodologies are inappropriate as they lie on other bounded metric spaces that deviate significantly from  $\mathbb{S}^{d-1}$ . For example, microbiologists are concerned with the spatial arrangement of lipids and proteins on the cellular membranes of microorganisms that are not adequately modelled by spheres. In the case of bacteria, ellipsoids or capsules (a cylinder with two hemispherical caps placed at each end) are far more appropriate candidate surfaces. Recent advances in 3D super-resolution imaging techniques (e.g. Cabriel et al., 2019; Gustavsson et al., 2018) output point patterns of this type, and there is a demand for the correct statistical procedures to analyse them.

Key to the statistical analysis of spatial data is the ability to form functional summary statistics from an observed pattern, primarily for performing exploratory data analysis and testing for complete spatial randomness (CSR). On  $\mathbb{R}^d$  and  $\mathbb{S}^{d-1}$ , there exists an infinite number of isometries, allowing for the notions of stationarity and isotropy to be well defined, which in turn allows for well defined functional summary statistics. However, on the surface of an arbitrary convex shape  $\mathbb{D} \subset \mathbb{R}^3$ , the set of available isometries is finite, and thus defining stationarity, isotropy, and summary statistics directly on  $\mathbb{D}$  is non-trivial. Building on the current literature for spherical point patterns, in particular the discussion of inhomogeneous point processes on a sphere by both Lawrence et al. (2016) and Møller and Rubak (2016), we show that it is possible to construct functional summary statistics for point processes on the surface of arbitrary convex shapes in  $\mathbb{R}^3$ , with our primary interest being to test for CSR.

Our approach is to map the point pattern from an arbitrary convex shape  $\mathbb{D}$  onto  $\mathbb{S}^2$ . For any Poisson process on  $\mathbb{D}$ , the Mapping Theorem (Kingman, 1993) determines that the mapped process on  $\mathbb{S}^2$  remains Poisson with the intensity function dependent on the mapping. By working on  $\mathbb{S}^2$ , we operate on a space that is more amenable to constructing summary statistics. These then allow us to test for CSR on  $\mathbb{D}$ . The functional summary statistics we develop are based on the inhomogeneous counterparts of typical functional summary statistics already established in the spatial statistics literature. In particular, we focus on the inhomogeneous  $K$ -function, first discussed by Baddeley et al. (2000) for  $\mathbb{R}^d$  and later extended to  $\mathbb{S}^{d-1}$  by Lawrence et al. (2016) and Møller and Rubak (2016). Furthermore, we also construct the empty-space function,  $F$ , spherical contact distribution,  $H$ , and  $J$ -function (the ratio of the  $H$ - and  $F$ -functions) for point processes on arbitrary convex shapes by extending the inhomogeneous definitions of van Lieshout (2011) from  $\mathbb{R}^d$  to  $\mathbb{S}^2$ .

Section 2 introduces the notation used throughout this work and formally states the hypothesis for testing CSR on an arbitrary convex shape. Section 3 discusses functional summary statistics on  $\mathbb{S}^2$ , key for the construction of summary statistics on more general bounded subsets of  $\mathbb{R}^3$ , and the impracticalities of attempting to define functional summary statistics directly on  $\mathbb{D}$ . Section 4 extends the inhomogeneous  $F$ -,  $H$ -, and  $J$ -functions from  $\mathbb{R}^d$  (van Lieshout, 2011) to  $\mathbb{S}^2$ . Section 5 describes the construction of functional summary statistics on bounded subsets of  $\mathbb{R}^3$  for Poisson processes, discussing their first and second order properties in the event that the intensity function is known. Section 6 provides two worked examples constructing functional summary statistics for

realisations of a Poisson process observed on a cube and an ellipsoid. Section 7 discusses how regular and cluster processes can be detected based on the deviations of the empirical functional summary statistics. Section 8 describes estimation procedures for the functional summary statistics when the intensity function is unknown and we propose two test statistics for CSR. Finally in Section 9 we conduct empirical power tests using Monte Carlo simulations to explore the properties of our proposed test statistics.

## 2. Preliminaries

In this section we outline the necessary spatial theory and notation used throughout this work. We start by introducing the notion of a bounded convex space in  $\mathbb{R}^3$  and then define what it means for a point process to lie on such a surface. We end with the statement of the problem that this work is primarily focused on.

### 2.1. Notation

Let  $\mathbf{x} \in \mathbb{R}^3$  such that  $\mathbf{x} = (x_1, x_2, x_3)^T$  and define  $\|\mathbf{x}\| = (x_1^2 + x_2^2 + x_3^2)^{1/2}$  to be the Euclidean norm with the origin of  $\mathbb{R}^3$  denoted as  $\mathbf{0} = (0, 0, 0)^T$ . Denote a subset of  $\mathbb{R}^3$  as  $\mathbb{D} = \{\mathbf{x} \in \mathbb{R}^3 : g(\mathbf{x}) = 0\}$ , where  $g : \mathbb{R}^3 \mapsto \mathbb{R}$ . We also suppose that  $\mathbb{D}$  is compact (i.e. closed and bounded) and call  $g$  the level-set function of  $\mathbb{D}$ . Define the set  $\mathbb{D}_{int} = \{\mathbf{x} \in \mathbb{R}^3 : g(\mathbf{x}) < 0\}$ , i.e. the boundary of  $\mathbb{D}_{int}$  is  $\mathbb{D}$  and we refer to  $\mathbb{D}_{int}$  as the interior of  $\mathbb{D}$ . The set  $\mathbb{D}_{int}$  is said to be convex if and only if for all  $\mathbf{x}, \mathbf{y} \in \mathbb{D}_{int}$  such that  $\mathbf{x} \neq \mathbf{y}$  then  $\{\mathbf{z} \in \mathbb{R}^3 : \mathbf{z} = \mathbf{x} + \gamma(\mathbf{y} - \mathbf{x}), \gamma \in (0, 1)\} \in \mathbb{D}_{int}$ . We thus define  $\mathbb{D}$  to be convex if its interior,  $\mathbb{D}_{int}$ , is convex. Examples of bounded convex sets of  $\mathbb{R}^3$  are spheres, ellipsoids, and cubes. Further for any bounded convex set  $\mathbb{D}$  with level-set function  $g$ , we will also define  $\tilde{g}$  which rearranges  $g(\mathbf{x}) = 0$ , such that  $x_3 = \tilde{g}(x_1, x_2)$ , i.e. we write  $x_3$  as a function of  $x_1$  and  $x_2$ . It may not always be possible to find  $\tilde{g}$  explicitly since, as defined previously, it may be the case that the resultant  $\tilde{g}$  is not a proper function. This issue can be rectified by partitioning  $\mathbb{D}$  appropriately. For example take the case of a sphere with radius 1, then  $g(\mathbf{x}) = x_1^2 + x_2^2 + x_3^2 - 1$ , hence  $\tilde{g}(x_1, x_2) = \pm(1 - x_1^2 - x_2^2)^{1/2}$ , which is not a proper function. In this case we partition the region  $\mathbb{D}$  into the regions  $x_3 \geq 0$  and  $x_3 < 0$ . Then for  $x_3 \geq 0$ , we define  $\tilde{g}(x_1, x_2) = +(1 - x_1^2 - x_2^2)^{1/2}$  and  $x_3 < 0$ ,  $\tilde{g}(x_1, x_2) = -(1 - x_1^2 - x_2^2)^{1/2}$ . For any bounded convex sets,  $\mathbb{D}$ , we also define its geodesic as the shortest path between two points  $\mathbf{x}, \mathbf{y} \in \mathbb{D}$  such that every point in the path is also an element of  $\mathbb{D}$  and denote the geodesic distance by  $d : \mathbb{D} \times \mathbb{D} \mapsto \mathbb{R}_+$ , where  $\mathbb{R}_+$  is the positive real line including 0, thus  $(\mathbb{D}, d(\cdot, \cdot))$  is a metric space. Additionally, we will frequently need to evaluate integrals over  $\mathbb{D}$ , which can be done using its infinitesimal area element defined as,

$$d\mathbb{D} = \sqrt{1 + \left(\frac{\partial \tilde{g}}{\partial x_1}\right)^2 + \left(\frac{\partial \tilde{g}}{\partial x_2}\right)^2} dx_1 dx_2.$$

We assume that these convex subsets of  $\mathbb{R}^3$  are defined such that the origin is inside  $\mathbb{D}$ , that is  $\mathbf{0} \in \mathbb{D}_{int}$ , we then say the space  $\mathbb{D}$  is centred. Our methodology can easily be adapted for non-centred spaces by making the appropriate translations to bring the origin inside  $\mathbb{D}$ .

Following the notation of Møller and Waagepetersen (2004), we define  $\lambda_{\mathbb{D}}(\mathbf{x})$  as the Lebesgue measure restricted to the surface of the convex shape  $\mathbb{D}$ . Consider point processes which lie on some bounded convex metric space  $(\mathbb{D}, d(\cdot, \cdot))$ . We define the notation  $A_K = A \cap K, A, K \subseteq \mathbb{D}$ . This nomenclature is often used when the set  $A$  has finite cardinality and  $K$  is any subset of  $\mathbb{D}$ . The cardinality of a set is denoted by  $|\cdot|$ . Further define  $B_{\mathbb{D}}(\mathbf{x}, r) = \{\mathbf{y} \in \mathbb{D}, d(\mathbf{x}, \mathbf{y}) \leq r\}$  and the set  $N_{lf} = \{A \subseteq \mathbb{D} : |A_K| < \infty, K \subseteq \mathbb{D}\}$  where  $K$  is any subset of  $\mathbb{D}$ . In other words,  $N_{lf}$  is the set of subsets of  $\mathbb{D}$  that have finite cardinality. To distinguish between points in a point process  $X$  and any point in the space  $\mathbb{D}$ , we shall refer to elements of our point process  $\mathbf{x} \in X$  as *events* whilst retaining the term *point* for any point in  $\mathbb{D}$ . We consider point processes which are *locally finite* and *simple*. A point process,  $X$ , lying on  $\mathbb{D}$  is said to be locally finite, if for any bounded set  $K \subseteq \mathbb{D}$ , the number of events of  $X$  in  $K$  is finite almost surely, i.e.  $X \in N_{lf}$  almost surely. A simple point process is one in which no coincident events exist almost surely, in other words if  $\mathbf{x}_i, \mathbf{x}_j \in X$  such that  $i \neq j$

then  $\mathbf{x}_i \neq \mathbf{x}_j$ . We also define the counting measure of  $X$  as  $N_X(K) = |X \cap K| = |X_K|$ . We denote the reduced Palm distribution of the point process  $X$  by  $P_{X_K}^\downarrow$  and define  $X_K^\downarrow$  to be the point process following this density, referring to this as the reduced Palm process (Møller and Waagepetersen, 2004). For a point process  $X$  on  $\mathbb{D}$  we define the *intensity measure* as the expected number of events of  $X$  for any  $K \subseteq \mathbb{D}$ , i.e.  $\mu(K) = \mathbb{E}[N_X(K)]$ , whilst the *intensity function*  $\rho(\mathbf{x})$  for all  $\mathbf{x} \in \mathbb{D}$ , if it exists, is given by,

$$\mu(K) = \int_K \rho(\mathbf{x})\lambda_{\mathbb{D}}(d\mathbf{x}),$$

where  $\rho(\mathbf{x})\lambda_{\mathbb{D}}(d\mathbf{x})$  can be interpreted heuristically as the probability of an event of  $X$  being in the infinitesimal area  $\lambda_{\mathbb{D}}(d\mathbf{x})$ .

We also define, for  $n \in \mathbb{N}$ ,  $\alpha^{(n)}(K_1, \dots, K_n) = \mathbb{E} \sum_{\mathbf{x}_1, \dots, \mathbf{x}_n \in X} \mathbb{1}[\mathbf{x}_1 \in K_1, \dots, \mathbf{x}_n \in K_n]$  as the *n*th-order factorial moment measures where the summation is taken over pairwise distinct sets of  $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ . Further we shall assume there exists  $\rho^{(n)}(\mathbf{x}_1, \dots, \mathbf{x}_n)$  such that,

$$\alpha^{(n)}(K_1, \dots, K_n) = \int_{K_1} \dots \int_{K_n} \rho^{(n)}(\mathbf{x}_1, \dots, \mathbf{x}_n)\lambda_{\mathbb{D}}(d\mathbf{x}_1) \dots \lambda_{\mathbb{D}}(d\mathbf{x}_n), \tag{1}$$

where  $K_1, \dots, K_n \subseteq \mathbb{D}$ . We can interpret  $\rho^{(n)}(\mathbf{x}_1, \dots, \mathbf{x}_n)\lambda_{\mathbb{D}}(d\mathbf{x}_1) \dots \lambda_{\mathbb{D}}(d\mathbf{x}_n)$  as the probability that events of  $X$  lie jointly in the infinitesimal areas  $\lambda_{\mathbb{D}}(d\mathbf{x}_i)$ ,  $i = 1, \dots, n$  and call  $\rho^{(n)}$  the *n*th-order factorial moment density. Notice that  $\alpha^{(1)} = \mu$ , and  $\rho^{(1)} = \rho$ . The *pair correlation function* is defined as,

$$h(\mathbf{x}, \mathbf{y}) = \frac{\rho^{(2)}(\mathbf{x}, \mathbf{y})}{\rho(\mathbf{x})\rho(\mathbf{y})},$$

where it is taken that division by 0 results in the pair correlation function equalling 0.

A useful alternative to the *n*th order product intensities are the *n*th order correlation functions (van Lieshout, 2011). They are recursively defined for  $n \in \mathbb{N}$ , based on product densities, with  $\xi_1 = 1$  and

$$\frac{\rho^{(n)}(\mathbf{x}_1, \dots, \mathbf{x}_n)}{\rho(\mathbf{x}_1) \dots \rho(\mathbf{x}_n)} = \sum_{k=1}^n \sum_{D_1, \dots, D_k} \xi_{|D_1|}(\mathbf{x}_{D_1}) \dots \xi_{|D_k|}(\mathbf{x}_{D_k}),$$

where the final sum ranges over all partitions  $\{D_1, \dots, D_k\}$  of  $\{1, \dots, n\}$  in  $k$  non-empty, disjoint sets,  $\mathbf{x}_{D_j} = \{\mathbf{x}_i : i \in D_j\}$ ,  $j = 1, \dots, k$ , and  $\mathbf{x}_i \in \mathbb{S}^2$  (van Lieshout, 2011). Further, we define the *generating functional* (Møller and Waagepetersen, 2004) of a point process  $X$  as

$$G_X(u) = \mathbb{E} \prod_{\mathbf{x} \in X} u(\mathbf{x}),$$

for a function  $u : \mathbb{D} \mapsto [0, 1]$ , which are also useful when discussing the *F*-, *H*-, and *J*-functions in Section 4.

We define a Poisson process on  $\mathbb{D}$  identically to one on  $\mathbb{R}^2$ . Let  $X$  be a point process on  $\mathbb{D}$  such that  $N_X(\mathbb{D}) \sim \text{Poisson}(\mu(\mathbb{D}))$  where,

$$\mu(B) = \int_B \rho(\mathbf{x})\lambda_{\mathbb{D}}(d\mathbf{x}), \quad \text{for } B \subseteq \mathbb{D},$$

where  $\mu(\mathbb{D}) < \infty$ . Then given  $N_X(\mathbb{D}) = n$ ,  $\mathbf{x}_i \in X$ ,  $i = 1, \dots, n$  are independent and identically distributed across  $\mathbb{D}$  with density proportional to  $\rho(\mathbf{x})$ . We say that  $X$  is a Poisson process on  $\mathbb{D}$  with intensity function  $\rho : \mathbb{D} \mapsto \mathbb{R}_+$ . When  $\rho \in \mathbb{R}_+$  is constant we say the process is homogeneous Poisson or CSR.

### 2.2. Statement of the problem

We are now in a position to formally state the hypothesis of CSR we are interested in and for which this work provides an approach to testing.

Let  $X$  be a spatial point process, such that  $g(X) = 0$  where  $g(X)$  is a notational convenience for  $g(\mathbf{x}) = 0$ , for all  $\mathbf{x} \in X$  and  $g$  is the level set for the convex shape  $\mathbb{D}$ . From a realisation of  $X$  we wish to conduct the following hypothesis test,

$$H_0 : X \text{ is CSR on } \mathbb{D} \quad \text{vs.} \quad H_1 : X \text{ is not CSR on } \mathbb{D}.$$

### 3. Summary statistics on $\mathbb{S}^2$ and the impracticalities of defining functional summary statistics on general convex shapes

To analyse point patterns on a convex shape  $\mathbb{D}$ , we will be required to map the point pattern onto the unit sphere  $\mathbb{S}^2$ ; we use the term spheroidal point process to denote a point process that resides on  $\mathbb{S}^2$ . Hence it is necessary to discuss functional summary statistics in this space. We also explain why it is non trivial to construct analogous functional summary statistics directly on  $\mathbb{D}$ , motivating a need for new methodology in order to test patterns which arise on such surfaces.

#### 3.1. Summary statistics on $\mathbb{S}^2$

Mentioned in passing by Ripley (1977), spherical point patterns only garnered interest over the past five years (Lawrence et al., 2016; Møller and Rubak, 2016; Robeson et al., 2014). Robeson et al. (2014) show that, on a sphere of radius  $R$ , the spherical  $K$ -function for a homogeneous Poisson process is  $K(r) = 2\pi R^2(1 - \cos(r/R))$ , where  $r$  is the geodesic distance from an arbitrary event of the process. Building on this, both Lawrence et al. (2016) and Møller and Rubak (2016) define a range of typical functional summary statistics for isotropic (rotationally invariant) point processes, including the empty-space and nearest-neighbour distributions. They also extend the  $K$ -function to the class of inhomogeneous point processes that have rotationally invariant pair correlation functions. This is analogous to the inhomogeneous extension given by Baddeley et al. (2000) for point processes on  $\mathbb{R}^d$ .

The geodesic on  $\mathbb{S}^2$  is commonly referred to as the great circle distance and for two points  $\mathbf{x}, \mathbf{y} \in \mathbb{S}^2$  has analytic form  $d(\mathbf{x}, \mathbf{y}) = \cos^{-1}(\mathbf{x} \cdot \mathbf{y})$ , where  $\mathbf{x} \cdot \mathbf{y}$  is the dot product between vectors  $\mathbf{x}, \mathbf{y} \in \mathbb{S}^2$ . We define a point process  $X$  on  $\mathbb{S}^2$  to be isotropic if its distribution is invariant under rotations, i.e.  $X \stackrel{d}{=} OX$ , for all  $O \in \mathcal{O}(3)$ , where  $\mathcal{O}(3)$  is the set of orthogonal  $3 \times 3$  matrices (Møller and Rubak, 2016). Here we use  $\stackrel{d}{=}$  to denote *equal in distribution*. Such an isotropic process has constant intensity function  $\rho \in \mathbb{R}_+$ .

Functional summary statistics are frequently employed for both exploratory data analysis and model fitting, playing a pivotal role in the early stages of any in depth investigation of an observed point pattern. In the homogeneous case, let  $X$  be an isotropic spheroidal point process with constant intensity function  $\rho \in \mathbb{R}_+$  then the  $F$ -,  $H$ -,  $J$ -, and  $K$ -functions are defined as,

$$F(r) = P(X_{B(\mathbf{o}, r)} \neq \emptyset) \tag{2}$$

$$H(r) = P(X_{\mathbf{o}, B(\mathbf{o}, r)}^1 \neq \emptyset) \tag{3}$$

$$J(r) = \frac{1 - H(r)}{1 - F(r)} \tag{4}$$

$$K(r) = \frac{1}{\rho} \mathbb{E} \sum_{\mathbf{x} \in X_{\mathbf{o}}^1} \mathbb{1}[d(\mathbf{x}, \mathbf{o}) \leq r], \tag{5}$$

where  $r \in [0, \pi]$  and  $\mathbf{o} = (0, 0, 1)^T$  is defined as the origin of  $\mathbb{S}^2$ . Estimators of these functional summary statistics can be used to determine whether the underlying process of an observed point pattern follows a specific distribution. In particular, they can be used to test whether a pattern arises from a CSR process or whether the underlying process exhibits regularity or clustering. A treatment of the standard isotropic functional summary statistics on  $\mathbb{S}^2$  is given in Lawrence et al. (2016) and Møller and Rubak (2016).

It will be necessary for us to consider the inhomogeneous extensions of these functional summary statistics as they will form the foundation for our functional summary statistics for point

process on convex bounded shapes in  $\mathbb{R}^3$ . We begin by reviewing the inhomogeneous  $K$ -function, originally attributed to [Baddeley et al. \(2000\)](#) for a class of inhomogeneous point processes in  $\mathbb{R}^d$ , and then extended to non-isotropic point processes on  $\mathbb{S}^2$  by [Lawrence et al. \(2016\)](#) and [Møller and Rubak \(2016\)](#). In Section 4, we will construct the inhomogeneous  $F$ -,  $H$ -, and  $J$ -functions for non-isotropic point processes on  $\mathbb{S}^2$ . This builds on the formulation of [van Lieshout \(2011\)](#) for non-stationary point processes in  $\mathbb{R}^d$ .

### 3.2. Inhomogeneous $K$ -function

For the extension of the  $K$ -function to inhomogeneous processes on  $\mathbb{R}^{2,3}$ , [Baddeley et al. \(2000\)](#) introduce the notion of a point process being *second order intensity reweighted stationary* (SOIRWS). Here we focus on the extension for  $\mathbb{S}^2$  where [Lawrence et al. \(2016\)](#) and [Møller and Rubak \(2016\)](#) define the notion of *second order intensity reweighted isotropic* (SOIRWI). A point process  $X$  on  $\mathbb{S}^2$  is said to be SOIRWI if its pair correlation function,  $h$ , is rotationally invariant, that is  $h(\mathbf{x}, \mathbf{y}) = h(d(\mathbf{x}, \mathbf{y}))$ , where  $d$  is the great circle distance on  $\mathbb{S}^2$ . [Lawrence et al. \(2016\)](#) and [Møller and Rubak \(2016\)](#) define the inhomogeneous  $K$ -function for a SOIRWI process as,

$$K_{\text{inhom}}(r) = \frac{1}{\lambda_{\mathbb{S}^2}(A)} \mathbb{E} \sum_{\mathbf{x}, \mathbf{y} \in X}^{\neq} \frac{\mathbb{1}[\mathbf{x} \in A, O_{\mathbf{x}}(\mathbf{y}) \in B_{\mathbb{S}^2}(\mathbf{o}, r)]}{\rho(\mathbf{x})\rho(\mathbf{y})}, \quad 0 \leq r \leq \pi,$$

where  $O_{\mathbf{x}} : \mathbb{S}^2 \mapsto \mathbb{S}^2$  is a rotation that takes  $\mathbf{x}$  to  $\mathbf{o}$ . Here,  $K_{\text{inhom}}$  is independent of the choice of  $A \subseteq \mathbb{S}^2$  with  $\lambda_{\mathbb{S}^2}(A) > 0$ , and by convention  $a/0 = 0$ , for  $a \geq 0$ . For a Poisson process it is easy to show that  $K_{\text{inhom}}(r) = 2\pi(1 - \cos(r))$ . Therefore the  $K$ -function is the same for all Poisson processes regardless of whether the intensity function is constant or not ([Møller and Rubak, 2016](#)).

Both [Lawrence et al. \(2016\)](#) and [Møller and Rubak \(2016\)](#) propose the following estimator for  $K_{\text{inhom}}$  for a fully observed point pattern on  $\mathbb{S}^2$ ,

$$\hat{K}_{\text{inhom}}(r) = \frac{1}{4\pi} \sum_{\mathbf{x}, \mathbf{y} \in X}^{\neq} \frac{\mathbb{1}[d(\mathbf{x}, \mathbf{y}) \leq r]}{\rho(\mathbf{x})\rho(\mathbf{y})}, \tag{6}$$

which is unbiased if  $\rho$  is known. In the more likely event that  $\rho$  is unknown, [Lawrence et al. \(2016\)](#) and [Møller and Rubak \(2016\)](#) suggest using a plugin estimator for  $\rho(\mathbf{x})\rho(\mathbf{y})$ .

### 3.3. Impracticalities of defining functional summary statistics directly on $\mathbb{D}$

We now explain the subtle reasoning as to why constructing functional summary statistics directly on  $\mathbb{D}$  is not a trivial extension from  $\mathbb{S}^2$ . The definitions given by Eqs. (2)–(5) are well defined when considering stationary or isotropic point processes on  $\mathbb{R}^d$  or  $\mathbb{S}^2$  respectively. This is because the symmetries of the space admit well defined notions of stationarity/isotropy based on translations and rotations. Since an arbitrary convex space  $\mathbb{D}$  does not, in general, have isometries these notions of stationarity/isotropy cannot be well defined. Therefore, defining functional summary statistics analogous to (2)–(5) is not possible.

Further, we also argue that we cannot define a point process to be SOIRWI on  $\mathbb{D}$ . On  $\mathbb{S}^2$  being SOIRWI is equivalent to having a rotationally invariant pair correlation function. We may be tempted to equivalently define a point process to be SOIRWI on  $\mathbb{D}$  if it has an invariant form for its pair correlation function. In particular this would make sense for a Poisson process on  $\mathbb{D}$  as it would have pair correlation function,  $h(\mathbf{x}) = 1$  for all  $\mathbf{x} \in \mathbb{D}$ . Closer inspection though leads us to conclude that this is not an appropriate definition for SOIRWI on  $\mathbb{D}$ . Based on [Møller and Waagepetersen \(2004, Definition 4.5, p. 32\)](#) we can take a point process  $X$  with intensity function  $\rho : \mathbb{S}^2 \mapsto \mathbb{R}_+$  as being SOIRWI on  $\mathbb{S}^2$  if the measure,

$$\kappa(B) = \frac{1}{\lambda_{\mathbb{S}^2}(A)} \mathbb{E} \sum_{\mathbf{x}, \mathbf{y} \in X}^{\neq} \frac{\mathbb{1}[\mathbf{x} \in A, O_{\mathbf{x}}(\mathbf{y}) \in B]}{\rho(\mathbf{x})\rho(\mathbf{y})}, \quad B \subseteq \mathbb{S}^2, \tag{7}$$

does not depend on the choice of  $A \subseteq \mathbb{S}^2$  for  $0 < |A| < \infty$ , where we take  $a/0 = 0$ .  $\mathcal{K}$  is then called the *second order reduced moment measure*. If the pair correlation function exists and is invariant under rotations, then by the Campbell–Mecke Theorem (Møller and Waagepetersen, 2004) it follows that

$$\mathcal{K}(B) = \int_B h(d(\mathbf{o}, \mathbf{x})) \lambda_{\mathbb{S}^2}(d\mathbf{x}), \quad B \subseteq \mathbb{S}^2.$$

Thus on  $\mathbb{S}^2$  a point process is SOIRWI if  $h$  is invariant under rotations. Eq. (7) implicitly depends on rotations  $O_{\mathbf{x}}(\mathbf{y})$ . If we now consider a point process on  $\mathbb{D}$ , we cannot construct the second order reduced moment measure as, in general, we do not have an analogous isometry. This in turn means that we cannot define SOIRWI directly on  $\mathbb{D}$  based on an invariance of the pair correlation function.

Moreover, for a point process on  $\mathbb{S}^2$ , consider the more specific case when  $B = B_{\mathbb{S}^2}(\mathbf{o}, r)$ ,  $r > 0$  in (7). This is identically the inhomogeneous  $K$ -function. The indicator function of (7) is still well-defined in the case of  $\mathbb{S}^2$  such that we are counting the events of  $X \setminus \{\mathbf{x}\}$  that are at most a distance  $r$  from  $\mathbf{x} \in X$ . This same intuition could not equivalently be applied to point processes on a convex shape as the ball of radius  $r$  from a point  $\mathbf{x}$  on  $\mathbb{D}$  also depends on  $\mathbf{x}$ , i.e.  $B_{\mathbb{D}}(\mathbf{x}, r) \subset \mathbb{D}$  is different for each  $\mathbf{x} \in \mathbb{D}$ . Thus it is not possible to directly define an inhomogeneous  $K$ -function on  $\mathbb{D}$ .

#### 4. Extending the inhomogeneous $F$ -, $H$ -, and $J$ -functions to $\mathbb{S}^2$

On  $\mathbb{R}^d$ , in the stationary case, it can be shown that the  $F$ -, and  $H$ -functions have infinite series representations (White, 1979), and further work by van Lieshout (2006) also gives an infinite series representation for the  $J$ -function based on the  $n$ th-order correlation functions. Theorem 1 gives an infinite series representation when the  $n$ th-order reduced factorial moment measure of all  $n$  exists, similar to (White, 1979) but where the underlying space is  $\mathbb{S}^2$ .

**Theorem 1.** *Let  $X$  be an isotropic spheroidal point process with constant intensity function  $\rho$ . Further we assume the existence of all  $n$ th-order factorial moment measures for both  $X$  and its reduced Palm process,  $X_{\mathbf{x}}^1$ . Then the  $F$ - and  $H$ -functions have the following series representation,*

$$F(r) = - \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \alpha^{(n)}(B_{\mathbb{S}^2}(\mathbf{o}, r), \dots, B_{\mathbb{S}^2}(\mathbf{o}, r))$$

$$H(r) = - \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \alpha_{\mathbf{o}}^{!(n)}(B_{\mathbb{S}^2}(\mathbf{o}, r) \dots, B_{\mathbb{S}^2}(\mathbf{o}, r))$$

where  $\alpha^{(n)}$  and  $\alpha_{\mathbf{x}}^{!(n)}$  are the factorial moment measure for  $X$  and  $X_{\mathbf{x}}^1$  and  $B_{\mathbb{S}^2}(\mathbf{o}, r)$  is the spherical cap of radius  $r$  at the origin  $\mathbf{o} \in \mathbb{S}^2$ . These representations hold provided the series is absolutely convergent, that is if  $\lim_{n \rightarrow \infty} |a_{n+1}/a_n| < 1$  or  $\limsup_{n \rightarrow \infty} (|a_n|)^{1/n} < 1$ , where  $a_n = ((-1)^n/n!) \alpha^{(n)}(B_{\mathbb{S}^2}(\mathbf{o}, r), \dots, B_{\mathbb{S}^2}(\mathbf{o}, r))$  for the  $F$ -function or  $a_n = ((-1)^n/n!) \alpha_{\mathbf{o}}^{!(n)}(B_{\mathbb{S}^2}(\mathbf{o}, r) \dots, B_{\mathbb{S}^2}(\mathbf{o}, r))$  for the  $H$ -function.

**Proof.** See Theorem 1, Section S1 of the Supplementary Material.  $\square$

The following corollary reduces the representations for the  $F$ -, and  $H$ -function for when the  $n$ th-order product density exist. These representations are those used by van Lieshout (2011).

**Corollary 1 (White, 1979).** *Under the same assumptions as Theorem 1, let  $X$  be an isotropic spheroidal point process with constant intensity function  $\rho$ . Further we assume the existence of all  $n$ th-order product intensities for both  $X$  and its reduced Palm process,  $X_{\mathbf{x}}^1$ . Then the  $F$ - and  $H$ -functions have the following series representation,*

$$F(r) = - \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \int_{B_{\mathbb{S}^2}(\mathbf{o}, r)} \dots \int_{B_{\mathbb{S}^2}(\mathbf{o}, r)} \rho^{(n)}(\mathbf{x}_1, \dots, \mathbf{x}_n) \lambda_{\mathbb{S}^2}(d\mathbf{x}_1) \dots \lambda_{\mathbb{S}^2}(d\mathbf{x}_n)$$

$$H(r) = - \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \int_{B_{\mathbb{S}^2}(\mathbf{o}, r)} \cdots \int_{B_{\mathbb{S}^2}(\mathbf{o}, r)} \frac{\rho^{(n+1)}(\mathbf{o}, \mathbf{x}_1, \dots, \mathbf{x}_n)}{\rho} \lambda_{\mathbb{S}^2}(d\mathbf{x}_1) \cdots \lambda_{\mathbb{S}^2}(d\mathbf{x}_n)$$

provided the series is absolutely convergent, where  $B_{\mathbb{S}^2}(\mathbf{o}, r)$  is the spherical cap of radius  $r$  at the origin  $\mathbf{o} \in \mathbb{S}^2$ .

**Proof.** See Corollary 1, Section S1 of the Supplementary Material.  $\square$

Adapting the work of van Lieshout (2011), the  $J$ -function for an isotropic spheroidal point process, based on the series for the  $F$ -, and  $H$ -function given by Theorem 1, has the following infinite series representation

$$J(r) = 1 + \sum_{n=1}^{\infty} \frac{(-\rho)^n}{n!} J_n(r), \quad 0 \leq r \leq \pi,$$

where  $J_n(r) = \int_{B_{\mathbb{S}^2}(\mathbf{o}, r)} \cdots \int_{B_{\mathbb{S}^2}(\mathbf{o}, r)} \xi^{(n+1)}(\mathbf{o}, \mathbf{x}_1, \dots, \mathbf{x}_n) \lambda_{\mathbb{S}^2}(d\mathbf{x}_1) \cdots \lambda_{\mathbb{S}^2}(d\mathbf{x}_n)$ , and  $B_{\mathbb{S}^2}(\mathbf{o}, r)$  is the spherical cap at the origin  $\mathbf{o} \in \mathbb{S}^2$ .

In order to define the inhomogeneous  $F$ -,  $H$ -, and  $J$ -function we first define the notion of iterative reweighted moment isotropic (IRWMI) for a class of non-isotropic spheroidal point processes, similar to the notion of iterative reweighted moment stationary in  $\mathbb{R}^{2,3}$  (van Lieshout, 2011).

**Definition 1.** A spheroidal point process  $X$  is said to be IRWMI if, for all  $n \in \mathbb{N}$ , the  $n$ th-order correlation functions are rotationally invariant. That is  $\xi_n(\mathbf{x}_1, \dots, \mathbf{x}_n) = \xi_n(O\mathbf{x}_1, \dots, O\mathbf{x}_n)$  for all  $n \in \mathbb{N}$  and  $O \in \mathcal{O}(3)$ .

Identically to the inhomogeneous  $J$ -function in  $\mathbb{R}^{2,3}$  (van Lieshout, 2011), we define the inhomogeneous  $J$ -function on  $\mathbb{S}^2$ .

**Definition 2.** For an IRWMI point process  $X$  with intensity function  $\rho : \mathbb{S}^2 \mapsto \mathbb{R}_+$  such that  $\bar{\rho} \equiv \inf_{\mathbf{x} \in \mathbb{S}^2} \rho(\mathbf{x}) > 0$ ,

$$J_{\text{inhom}}(r) = 1 + \sum_{n=1}^{\infty} \frac{(-\bar{\rho})^n}{n!} J_n(r), \quad 0 \leq r \leq \pi$$

where  $J_n(r) = \int_{B_{\mathbb{S}^2}(\mathbf{o}, r)} \cdots \int_{B_{\mathbb{S}^2}(\mathbf{o}, r)} \xi_{n+1}(\mathbf{o}, \mathbf{x}_1, \dots, \mathbf{x}_n) \lambda_{\mathbb{S}^2}(d\mathbf{x}_1) \cdots \lambda_{\mathbb{S}^2}(d\mathbf{x}_n)$  and the series is absolutely convergent.

Notice that since the point process is IRWMI then the  $J$ -function does not depend on the origin  $\mathbf{o}$ , and furthermore when the point process is isotropic  $J_{\text{inhom}}$  collapses down to  $J$  since the intensity function is constant. In the context of  $\mathbb{R}^d$ , van Lieshout (2011) shows that the inhomogeneous  $J$ -function can be written as the ratio of generating functionals of the point process. Here we easily adapt the theorem for IRWMI point processes on  $\mathbb{S}^2$ .

**Theorem.** For all  $r \in [0, \pi]$  and  $\mathbf{y} \in \mathbb{S}^2$ ,

$$u_r^{\mathbf{y}}(\mathbf{x}) = \frac{\bar{\rho} \mathbb{1}[O_{\mathbf{y}}(\mathbf{x}) \in B_{\mathbb{S}^2}(\mathbf{o}, r)]}{\rho(\mathbf{x})}, \quad \mathbf{x} \in \mathbb{S}^2,$$

where  $O_{\mathbf{y}} : \mathbb{S}^2 \mapsto \mathbb{S}^2$  is a rotation that maps  $\mathbf{y}$  to  $\mathbf{o}$ . Assuming that the series  $\sum_{n=1}^{\infty} \frac{\bar{\rho}^n}{n!} \int_{B_{\mathbb{S}^2}(\mathbf{o}, r)} \cdots \int_{B_{\mathbb{S}^2}(\mathbf{o}, r)} \frac{\rho^{(n)}(\mathbf{x}_1, \dots, \mathbf{x}_n)}{\rho(\mathbf{x}_1) \cdots \rho(\mathbf{x}_n)} \lambda_{\mathbb{S}^2}(d\mathbf{x}_1) \cdots \lambda_{\mathbb{S}^2}(d\mathbf{x}_n)$  is absolutely convergent. Then under the further assumptions associated with the inhomogeneous  $J$ -function and the existence of all  $n$ th-order intensity function  $\rho_{\mathbf{y}}^{(n)}$  for the reduced Palm distribution  $X_{\mathbf{y}}^!$ ,  $\forall \mathbf{y} \in \mathbb{S}^2$ ,

$$J_{\text{inhom}}(r) = \frac{G_{\mathbf{y}}^!(1 - u_r^{\mathbf{y}})}{G(1 - u_r^{\mathbf{y}})}, \quad 0 \leq r \leq \pi,$$

for when  $G(1 - u_r^{\mathbf{y}}) > 0$ , where  $G_{\mathbf{y}}^!$  and  $G$  are the generating functionals for  $X_{\mathbf{y}}^!$  and  $X$  respectively.



**Proof.** See Theorem 1 of van Lieshout (2011). □

From the proof given by van Lieshout (2011), it can be shown that the numerator and denominator do not depend on the arbitrary event or point  $\mathbf{y}$  respectively. Further, in the case of an isotropic point process the numerator can be shown to be  $G_{\mathbf{y}}^{\downarrow}(1 - u_r^{\mathbf{y}}) = 1 - H(r)$ , whilst the denominator is  $G(1 - u_r^{\mathbf{y}}) = 1 - F(r)$ , and so the  $F$ -, and  $H$ -functions can be extended to the inhomogeneous case,

$$\begin{aligned} F_{\text{inhom}}(r) &= 1 - G(1 - u_r^{\mathbf{y}}) \\ H_{\text{inhom}}(r) &= 1 - G_{\mathbf{y}}^{\downarrow}(1 - u_r^{\mathbf{y}}), \end{aligned}$$

where the functions do not depend on the arbitrary event  $\mathbf{y}$  of the point process.

Similar to van Lieshout (2011) in  $\mathbb{R}^{2,3}$ , we propose the following estimators for the inhomogeneous  $F$ -, and  $H$ -functions for spheroidal IRWMI point processes as,

$$\hat{F}_{\text{inhom}}(r) = 1 - \frac{\sum_{\mathbf{p} \in P} \prod_{\mathbf{x} \in X \cap B_{\mathbb{S}^2}(\mathbf{p}, r)} \left(1 - \frac{\hat{\rho}}{\rho(\mathbf{x})}\right)}{|P|} \tag{8}$$

$$\hat{H}_{\text{inhom}}(r) = 1 - \frac{\sum_{\mathbf{x} \in X} \prod_{\mathbf{y} \in (X \setminus \{\mathbf{x}\}) \cap B_{\mathbb{S}^2}(\mathbf{x}, r)} \left(1 - \frac{\hat{\rho}}{\rho(\mathbf{y})}\right)}{N_X(\mathbb{S}^2)}, \tag{9}$$

where  $P \subseteq \mathbb{S}^2$  is a finite grid of points. The properties of the  $\hat{F}_{\text{inhom}}$ -function are independent of the choice of  $P$  (van Lieshout, 2011). In this work we choose  $P$  such that the points on  $\mathbb{S}^2$  are equidistant. van Lieshout (2011) shows that  $\hat{F}_{\text{inhom}}(r)$  is unbiased whilst  $\hat{H}_{\text{inhom}}(r)$  is ratio-unbiased. Then since  $\hat{F}_{\text{inhom}}$  is unbiased and  $\hat{H}_{\text{inhom}}$  is ratio-unbiased, constructing  $\hat{J}_{\text{inhom}}$  as

$$\hat{J}_{\text{inhom}}(r) = \frac{1 - \hat{H}_{\text{inhom}}(r)}{1 - \hat{F}_{\text{inhom}}(r)}, \tag{10}$$

gives a ratio-unbiased estimator for  $J_{\text{inhom}}(r)$ .

### 5. Summary statistics for Poisson processes on convex shapes

Here, we construct summary statistics for Poisson processes on general convex shapes. We show that a Poisson process on a general convex shape,  $\mathbb{D}$ , can be mapped to a Poisson process on a sphere, and then define functional summary statistics for such processes. We discuss properties of these functional summary statistics in the more general setting of inhomogeneous Poisson processes on  $\mathbb{S}^2$ .

#### 5.1. Mapping from $\mathbb{D}$ to $\mathbb{S}^2$

To circumvent the geometrical restrictions of  $\mathbb{D}$  we show, in this section, that we can map Poisson processes from  $\mathbb{D}$  to  $\mathbb{S}^2$  and construct functional summary statistics in this space. Theorem 2 shows that a Poisson process on  $\mathbb{D}$  can be transformed to a Poisson process on a sphere where we can take advantage of the rotational symmetries. The invariance of Poisson processes between metric spaces is known as the Mapping Theorem (Kingman, 1993). We use the function  $f(\mathbf{x}) = \mathbf{x}/\|\mathbf{x}\|$  to map point patterns from  $\mathbb{D}$  to  $\mathbb{S}^2$ . Lemma 1 shows that this function is bijective and hence measurable.

**Lemma 1.** *Let  $\mathbb{D}$  be a convex subset of  $\mathbb{R}^3$  such that the origin in  $\mathbb{R}^3$  is in the interior of  $\mathbb{D}$ , i.e.  $\mathbf{o} \in \mathbb{D}_{\text{int}}$ . Then the function  $f(\mathbf{x}) = \mathbf{x}/\|\mathbf{x}\|, f : \mathbb{D} \mapsto \mathbb{S}^2$  is bijective.*

**Proof.** See Lemma 1, Section S2 of the Supplementary Material. □

Rather than using the Mapping Theorem (Kingman, 1993), we utilise Proposition 3.1 of Møller and Waagepetersen (2004) to show that mapping a Poisson process from  $\mathbb{D}$  to  $\mathbb{S}^2$  results in a new Poisson process on  $\mathbb{S}^2$  and also derive the intensity function of the mapped process on  $\mathbb{S}^2$ .

**Theorem 2.** Let  $X$  be a Poisson process on an arbitrary bounded convex shape  $\mathbb{D} \subset \mathbb{R}^3$  with intensity function  $\rho : \mathbb{D} \mapsto \mathbb{R}_+$ . We assume that  $\mathbb{D} = \{\mathbf{x} \in \mathbb{R}^3 : g(\mathbf{x}) = 0\}$  where  $g(\mathbf{x}) = 0$  is the level-set function and is defined as,

$$g(\mathbf{x}) = \begin{cases} g_1(\mathbf{x}) = 0, & \mathbf{x} \in \mathbb{D}_1 \\ \vdots \\ g_n(\mathbf{x}) = 0, & \mathbf{x} \in \mathbb{D}_n \end{cases}$$

such that  $\cup_{i=1}^n \mathbb{D}_i = \mathbb{D}$  and  $\mathbb{D}_i \cap \mathbb{D}_j = \emptyset, \forall i \neq j$ . Let  $Y = f(X)$ , where  $f(\mathbf{x}) = \mathbf{x}/\|\mathbf{x}\|$  with  $f(X) = \{\mathbf{y} \in \mathbb{S}^2 : \mathbf{y} = \mathbf{x}/\|\mathbf{x}\|, \mathbf{x} \in X\}$ . Then  $Y$  is a Poisson process on  $\mathbb{S}^2$ , with intensity function,

$$\rho^*(\mathbf{x}) = \begin{cases} \rho(f^{-1}(\mathbf{x}))l_1(f^{-1}(\mathbf{x}))J_{(1,f^*)}(\mathbf{x})\sqrt{1-x_1^2-x_2^2}, & \mathbf{x} \in f(\mathbb{D}_1) \\ \vdots \\ \rho(f^{-1}(\mathbf{x}))l_n(f^{-1}(\mathbf{x}))J_{(n,f^*)}(\mathbf{x})\sqrt{1-x_1^2-x_2^2}, & \mathbf{x} \in f(\mathbb{D}_n) \end{cases} \tag{11}$$

where,

$$\begin{aligned} x_3 &= \tilde{g}_i(x_1, x_2) \\ l_i(\mathbf{x}) &= \left[ 1 + \left( \frac{\partial \tilde{g}_i}{\partial x_1} \right)^2 + \left( \frac{\partial \tilde{g}_i}{\partial x_2} \right)^2 \right]^{\frac{1}{2}} \\ J_{(i,f^{*-1})}(\mathbf{x}) &= \frac{1}{(x_1^2 + x_2^2 + \tilde{g}_i^2(x_1, x_2))^3} \\ \det &\left[ \begin{pmatrix} x_2^2 + \tilde{g}_i^2(x_1, x_2) - x_1 \tilde{g}_i(x_1, x_2) \frac{\partial \tilde{g}_i}{\partial x_1} & -x_1 \left( x_2 + \tilde{g}_i(x_1, x_2) \frac{\partial \tilde{g}_i}{\partial x_2} \right) \\ -x_2 \left( x_1 + \tilde{g}_i(x_1, x_2) \frac{\partial \tilde{g}_i}{\partial x_1} \right) & x_1^2 + \tilde{g}_i^2(x_1, x_2) - x_2 \tilde{g}_i(x_1, x_2) \frac{\partial \tilde{g}_i}{\partial x_2} \end{pmatrix} \right] \\ J_{(i,f^*)}(\mathbf{x}) &= \frac{1}{J_{(i,f^{*-1})}(f^{-1}(\mathbf{x}))}, \end{aligned}$$

where  $f^{-1}$  is the inverse of  $f$ ,  $\det(\cdot)$  is the determinant operator, and  $f^* : \mathbb{R}^2 \mapsto \mathbb{R}^2$  is the function which maps  $x_1 \mapsto x_1/\|\mathbf{x}\|$  and  $x_2 \mapsto x_2/\|\mathbf{x}\|$ .

**Proof.** See Theorem 2, Section S2 of the Supplementary Material.  $\square$

To solidify the notation used to describe the space  $\mathbb{D}$  in Theorem 1, we demonstrate it with a clear example. Let us suppose that  $\mathbb{D}$  is an ellipsoid with semi-major axis lengths  $a, b$ , and  $c$  along the  $x, y$  and  $z$  - axes. Then we define  $\mathbb{D}_1$  to be the  $\mathbb{D} \cap \{\mathbf{x} \in \mathbb{R}^3 : \mathbf{x} = (x_1, x_2, x_3)^T \text{ and } x_3 \geq 0\}$  i.e. the elements of  $\mathbb{D}$  with non-negative  $x_3$  component. Similarly we define  $\mathbb{D}_2 = \mathbb{D} \cap \{\mathbf{x} \in \mathbb{R}^3 : \mathbf{x} = (x_1, x_2, x_3)^T \text{ and } x_3 < 0\}$ . Then using the notation outlined in Section 2.1 we take  $\tilde{g}_1(\mathbf{x}) = +c\sqrt{1 - (x_1/a)^2 - (x_2/b)^2}$  and  $\tilde{g}_2(\mathbf{x}) = -c\sqrt{1 - (x_1/a)^2 - (x_2/b)^2}$ .

**Remark 1.** A notion of bijectivity arises from this theorem. Consider the set of all Poisson processes on  $\mathbb{D}$  such that their intensity functions exist, label this set  $T_{\mathbb{D}}$ . Also define  $T_{\mathbb{S}^2}$  as all the Poisson processes on  $\mathbb{S}^2$  such that their intensity functions exist. Then for any  $X \in T_{\mathbb{D}}$  implies that  $f(X) \in T_{\mathbb{S}^2}$ . Similarly by considering the inverse operation  $f^{-1}$ , which exists by Lemma 1, for all  $Y \in T_{\mathbb{S}^2}$  implies that  $f(Y) \in T_{\mathbb{D}}$ . Hence the mapping  $f : T_{\mathbb{D}} \mapsto T_{\mathbb{S}^2}$  is surjective. By Theorem 2 if  $X, Y$  are Poisson processes on  $\mathbb{D}$  with intensity function  $\rho_X$  and  $\rho_Y$  respectively then  $f(X)$  and  $f(Y)$  are the same Poisson process if and only if  $\rho_X = \rho_Y$  and so the mapping is also injective, and hence bijective. This means that analysis of a Poisson process,  $X$ , on  $\mathbb{D}$  is equivalent to the analysis of  $f(X)$  on  $\mathbb{S}^2$ .

**Remark 2.** Further, another useful result which follows directly from [Theorem 2](#) is the construction of approximate Poisson processes on  $\mathbb{S}^2$ . More precisely consider a convex surface  $\mathbb{D}$  for which instead of having a level-set function,  $g$ , we have an approximation to the space, for example consider we have a finite piecewise planar approximation to  $\mathbb{D}$ . Then  $\mathbb{D}$  can be approximated by  $\cup_{i=1}^n \mathbb{D}_i$ , where each  $\mathbb{D}_i$  is a planar piece and there are  $n \in \mathbb{N}$  pieces to the approximation. For each  $\mathbb{D}_i$  the level-set function is  $g_i(\mathbf{x}) = a_i x_1 + b_i x_2 + c_i x_3 + d_i = 0$ , and we can then use this approximation of  $\mathbb{D}$  to map a Poisson process on  $\mathbb{D}$  to  $\mathbb{S}^2$ .

### 5.2. Construction of functional summary statistics

We are now in a position to construct functional summary statistics for a Poisson process which lies on some bounded convex space  $\mathbb{D}$ . Since all Poisson processes on  $\mathbb{S}^2$  are SOIRWI ([Møller and Rubak, 2016](#)) and IRWMI ([van Lieshout, 2011](#)), the estimators for  $F_{\text{inhom}}$ ,  $H_{\text{inhom}}$ ,  $J_{\text{inhom}}$ , and  $K_{\text{inhom}}$  (see Eqs. (8)–(10) and (6) respectively) ([van Lieshout, 2011](#); [Møller and Rubak, 2016](#)) can be combined with the mapped intensity function from [Theorem 2](#) to construct estimators as follows,

$$\hat{F}_{\text{inhom},\mathbb{D}}(r) = 1 - \frac{\sum_{\mathbf{p} \in P} \prod_{\mathbf{x} \in Y \cap B_{\mathbb{S}^2}(\mathbf{p}, r)} \left(1 - \frac{\bar{\rho}^*}{\rho^*(\mathbf{x})}\right)}{|P|} \tag{12}$$

$$\hat{H}_{\text{inhom},\mathbb{D}}(r) = 1 - \frac{\sum_{\mathbf{x} \in Y} \prod_{\mathbf{y} \in (Y \setminus \{\mathbf{x}\}) \cap B_{\mathbb{S}^2}(\mathbf{x}, r)} \left(1 - \frac{\bar{\rho}^*}{\rho^*(\mathbf{y})}\right)}{N_Y(\mathbb{S}^2)} \tag{13}$$

$$\hat{J}_{\text{inhom},\mathbb{D}}(r) = \frac{1 - \hat{H}_{\text{inhom},\mathbb{D}}(r)}{1 - \hat{F}_{\text{inhom},\mathbb{D}}(r)} \tag{14}$$

$$\hat{K}_{\text{inhom},\mathbb{D}}(r) = \frac{1}{4\pi} \sum_{\mathbf{x}, \mathbf{y} \in Y}^{\neq} \frac{\mathbb{1}[d(\mathbf{x}, \mathbf{y}) \leq r]}{\rho^*(\mathbf{x})\rho^*(\mathbf{y})}, \tag{15}$$

where  $X$  is a Poisson process on  $\mathbb{D}$  with intensity function  $\rho$ ,  $Y = f(X)$  is the mapped Poisson process onto  $\mathbb{S}^2$ ,  $\rho^*$  is given by (11) and  $\bar{\rho}^* = \inf_{\mathbf{x} \in \mathbb{S}^2} \rho^*(\mathbf{x})$ . In the event that  $\rho : \mathbb{D} \mapsto \mathbb{R}_+$  is unknown and therefore  $\rho^*$  is unknown, nonparametric plug-in estimates of  $\rho^*$  can be constructed on  $\mathbb{S}^2$  ([Lawrence et al., 2016](#); [Møller and Rubak, 2016](#)).

### 5.3. Properties of functional summary statistics

Consider the general case of all Poisson processes on  $\mathbb{S}^2$ . [Theorem 3](#) gives the expectations of  $\hat{F}_{\text{inhom}}(r)$ ,  $\hat{H}_{\text{inhom}}(r)$ , and  $\hat{K}_{\text{inhom}}(r)$ . We restate the mean of  $\hat{K}_{\text{inhom}}$  ([Lawrence et al., 2016](#); [Møller and Rubak, 2016](#)) and adapt the proof to Proposition 1 in [van Lieshout \(2011\)](#) for  $\mathbb{R}^d$ , to show that  $\hat{F}_{\text{inhom}}$  is unbiased and  $\hat{H}_{\text{inhom}}$  is ratio unbiased for  $\mathbb{S}^2$ . In addition we also provide the expectation of  $\hat{H}_{\text{inhom}}(r)$ .

**Theorem 3.** *Let  $X$  be a spherical Poisson process on  $\mathbb{S}^2$  with known intensity function  $\rho : \mathbb{S}^2 \mapsto \mathbb{R}_+$ , such that  $\bar{\rho} = \inf_{\mathbf{x} \in \mathbb{S}^2} \rho(\mathbf{x}) > 0$ . Then the estimators for  $\hat{F}_{\text{inhom}}(r)$ , and  $\hat{K}_{\text{inhom}}(r)$  are unbiased whilst  $\hat{H}_{\text{inhom}}(r)$  is ratio-unbiased. More precisely,*

$$\begin{aligned} \mathbb{E}[\hat{F}_{\text{inhom}}(r)] &= 1 - \exp(-\bar{\rho}2\pi(1 - \cos r)) \\ \mathbb{E}[\hat{H}_{\text{inhom}}(r)] &= 1 - \frac{\exp(-\bar{\rho}2\pi(1 - \cos r)) - \exp(-\mu(\mathbb{S}^2))}{1 - \frac{\bar{\rho}2\pi(1 - \cos r)}{\mu(\mathbb{S}^2)}} \\ \mathbb{E}[\hat{K}_{\text{inhom}}(r)] &= 2\pi(1 - \cos r), \end{aligned}$$

where  $r \in [0, \pi]$ , and  $\bar{\rho} = \inf_{\mathbf{x} \in \mathbb{S}^2} \rho(\mathbf{x}) > 0$ . Further by unbiasedness and ratio-unbiasedness of  $\hat{F}_{\text{inhom}}(r)$  and  $\hat{H}_{\text{inhom}}(r)$ , respectively, we immediately have ratio-unbiasedness of  $\hat{J}_{\text{inhom}}(r)$ .

**Proof.** See Lawrence et al. (2016) for treatment of  $\hat{K}_{\text{inhom}}(r)$ . Results for  $\hat{F}_{\text{inhom}}(r)$  and  $\hat{H}_{\text{inhom}}(r)$  follow from a trivial adaptation of the proof for Proposition 1 in van Lieshout (2011). For the expectation of  $\hat{H}_{\text{inhom}}(r)$  see Theorem 3, Section S3 of the Supplementary Material.  $\square$

Theorem 3 shows that  $\hat{H}_{\text{inhom}}(r)$  is a biased estimator for  $H_{\text{inhom}}(r)$ . Although biased it can be bounded.

**Corollary 2.** With the same assumptions as Theorem 3, let  $X$  be a spherical Poisson process on  $\mathbb{S}^2$  with intensity function  $\rho : \mathbb{S}^2 \mapsto \mathbb{R}_+$ . Defining  $\bar{\rho} = \inf_{\mathbf{x} \in \mathbb{S}^2} \rho(\mathbf{x})$ , the bias of the estimator  $\hat{H}_{\text{inhom}}(r)$  is bounded by

$$|\text{Bias}(\hat{H}_{\text{inhom}}(r))| \leq \exp(-\mu(\mathbb{S}^2)) \leq \exp(-4\pi \bar{\rho}),$$

for all  $r \in [0, \pi]$ .

**Proof.** See Corollary 2, Section S3 of the Supplementary Material.  $\square$

Corollary 2 shows that, depending on the intensity function and hence  $\bar{\rho} = \inf_{\mathbf{x} \in \mathbb{S}^2} \rho(\mathbf{x})$ , the bias can be considered negligible. In the examples to come we set the expected number of events in the process to be large enough for the bias to be considered negligible. Next we provide the variance of the estimators of the functional summary statistics.

**Theorem 4.** Let  $X$  be a spherical Poisson process on  $\mathbb{S}^2$  with known intensity function  $\rho : \mathbb{S}^2 \mapsto \mathbb{R}_+$ , such that  $\bar{\rho} = \inf_{\mathbf{x} \in \mathbb{S}^2} \rho(\mathbf{x}) > 0$ . Then the estimators  $\hat{K}_{\text{inhom}}(r)$ ,  $\hat{F}_{\text{inhom}}(r)$ , and  $H_{\text{inhom}}(r)$  have variance,

$$\text{Var}(\hat{K}_{\text{inhom}}(r)) = \frac{1}{8\pi^2} \int_{\mathbb{S}^2} \int_{\mathbb{S}^2} \frac{\mathbb{1}[d(\mathbf{x}, \mathbf{y}) \leq r]}{\rho(\mathbf{x})\rho(\mathbf{y})} \lambda_{\mathbb{S}^2}(d\mathbf{x})\lambda_{\mathbb{S}^2}(d\mathbf{y}) + (1 - \cos r)^2 \int_{\mathbb{S}^2} \frac{1}{\rho(\mathbf{x})} \lambda_{\mathbb{S}^2}(d\mathbf{x}),$$

$$\text{Var}(\hat{F}_{\text{inhom}}(r)) = \frac{\exp(-2\bar{\rho}\lambda_{\mathbb{S}^2}(B_{\mathbb{S}^2}(\mathbf{o}, r)))}{|P|^2}$$

$$\sum_{\mathbf{p} \in P} \sum_{\mathbf{p}' \in P} \exp\left(\int_{B_{\mathbb{S}^2}(\mathbf{p}, r) \cap B_{\mathbb{S}^2}(\mathbf{p}', r)} \frac{\bar{\rho}^2}{\rho(\mathbf{x})} \lambda_{\mathbb{S}^2}(d\mathbf{x})\right) - \exp(-2\bar{\rho}\lambda_{\mathbb{S}^2}(B_{\mathbb{S}^2}(\mathbf{o}, r))),$$

$$\text{Var}(\hat{H}_{\text{inhom}}(r))$$

$$= \frac{1}{\mu^2(\mathbb{S}^2)} \int_{\mathbb{S}^2} \int_{\mathbb{S}^2} (\rho(\mathbf{x}) - \bar{\rho}\mathbb{1}[\mathbf{x} \in B_{\mathbb{S}^2}(\mathbf{y}, r)]) (\rho(\mathbf{y}) - \bar{\rho}\mathbb{1}[\mathbf{y} \in B_{\mathbb{S}^2}(\mathbf{x}, r)])$$

$$\frac{e^{-\mu(\mathbb{S}^2)}}{A_1^2(\mathbf{x}, \mathbf{y})} \left( e^{\mu(\mathbb{S}^2)A_1(\mathbf{x}, \mathbf{y})} - 1 - \text{Ei}(\mu(\mathbb{S}^2)A_1(\mathbf{x}, \mathbf{y})) + \gamma + \log(\mu(\mathbb{S}^2)A_1(\mathbf{x}, \mathbf{y})) \right) \lambda_{\mathbb{S}^2}(d\mathbf{x})\lambda_{\mathbb{S}^2}(d\mathbf{y})$$

$$+ \frac{1}{\mu(\mathbb{S}^2)} \int_{\mathbb{S}^2} \frac{e^{-\mu(\mathbb{S}^2)}}{A_2(\mathbf{x})} (\gamma + \log(\mu(\mathbb{S}^2)A_2(\mathbf{x})) - \text{Ei}(\mu(\mathbb{S}^2)A_2(\mathbf{x}))) \rho(\mathbf{y})\lambda_{\mathbb{S}^2}(d\mathbf{y})$$

$$- \frac{e^{-2\mu(\mathbb{S}^2)}}{\left(1 - \frac{\bar{\rho}}{\mu(\mathbb{S}^2)} 2\pi(1 - \cos r)\right)^2} \left( e^{\mu(\mathbb{S}^2)\left(1 - \frac{\bar{\rho}}{\mu(\mathbb{S}^2)} 2\pi(1 - \cos r)\right)} - 1 \right)^2$$

where,

$$A_1(\mathbf{x}, \mathbf{y}) = 1 - \frac{2\bar{\rho}}{\mu(\mathbb{S}^2)} 2\pi(1 - \cos r) + \frac{\bar{\rho}^2}{\mu(\mathbb{S}^2)} \int_{B_{\mathbb{S}^2}(\mathbf{x}, r) \cap B_{\mathbb{S}^2}(\mathbf{y}, r)} \frac{1}{\rho(\mathbf{z})} \lambda_{\mathbb{S}^2}(d\mathbf{z})$$

$$A_2(\mathbf{x}) = 1 - \frac{2\bar{\rho}}{\mu(\mathbb{S}^2)} 2\pi(1 - \cos r) + \frac{\bar{\rho}^2}{\mu(\mathbb{S}^2)} \int_{B_{\mathbb{S}^2}(\mathbf{x}, r)} \frac{1}{\rho(\mathbf{y})} \lambda_{\mathbb{S}^2}(\mathbf{d}\mathbf{y})$$

$$\text{Ei}(x) = - \int_{-x}^{\infty} \frac{e^{-t}}{t} dt$$

and  $\text{Ei}(x)$  is the exponential integral and  $r \in [0, \pi]$ .

**Proof.** See [Theorem 4](#), Section S4 of the Supplementary Material.  $\square$

Due to the complexity of the estimator for the  $\hat{J}_{\text{inhom}}$ -function, its mean and variance are extremely complex and although can be derived in terms of integrals over  $\mathbb{S}^2$ , we instead give an approximation based on the Taylor series expansion of the function  $f(x, y) = x/y$  around the means of the numerator and denominator. We first provide conditions for which the first two moments of  $\hat{J}_{\text{inhom}}(r)$  exist and then proceed to show how it can be approximated.

**Theorem 5.** Let  $X$  be a spheroidal Poisson process with intensity function  $\rho : \mathbb{S}^2 \mapsto \mathbb{R}_+$  such that  $\bar{\rho} \equiv \inf_{\mathbf{x} \in \mathbb{S}^2} \rho(\mathbf{x}) > 0$ . Let  $P$  be any finite grid on  $\mathbb{S}^2$  and define  $r_{\max} = \sup\{r \in [0, \pi] : \text{there exists } \mathbf{p} \in P \text{ such that } \rho(\mathbf{x}) \neq \bar{\rho} \text{ for all } \mathbf{x} \in B_{\mathbb{S}^2}(\mathbf{p}, r)\}$ . Then for any given  $r \in [0, r_{\max}]$  both  $\mathbb{E}[\hat{J}_{\text{inhom}}(r)]$  and  $\text{Var}(\hat{J}_{\text{inhom}}(r))$  exist.

**Proof.** See [Theorem 5](#), Section S5 of the Supplementary Material.  $\square$

**Proposition 1.** Let  $X$  be a spheroidal Poisson process with known intensity function  $\rho : \mathbb{S}^2 \mapsto \mathbb{R}_+$ . Then the covariance between  $1 - \hat{H}_{\text{inhom}}(r)$  and  $1 - \hat{F}_{\text{inhom}}(r)$  for  $r \in [0, \pi]$  is,

$$\begin{aligned} & \text{Cov}(1 - \hat{H}_{\text{inhom}}(r), 1 - \hat{F}_{\text{inhom}}(r)) \\ &= \frac{1}{|P|} \sum_{\mathbf{p} \in P} \int_{\mathbb{S}^2} \left( 1 - \frac{\bar{\rho} \mathbb{1}[\mathbf{x} \in B_{\mathbb{S}^2}(\mathbf{p}, r)]}{\rho(\mathbf{x})} \right) \\ & \quad \frac{\exp \left\{ -2\bar{\rho} 2\pi(1 - \cos r) - \int_{B_{\mathbb{S}^2}(\mathbf{x}, r) \cap B_{\mathbb{S}^2}(\mathbf{p}, r)} \frac{\bar{\rho}^2}{\rho(\mathbf{y})} \lambda_{\mathbb{S}^2}(\mathbf{d}\mathbf{y}) \right\}}{A(\mathbf{x}, \mathbf{p})} \frac{\rho(\mathbf{x})}{\mu(\mathbb{S}^2)} \lambda_{\mathbb{S}^2}(\mathbf{d}\mathbf{x}) \\ & \quad - \exp(-2\pi(1 - \cos r)\bar{\rho}) \left( \exp(-2\pi(1 - \cos r)\bar{\rho}) - \exp(-\mu(\mathbb{S}^2)) \right) \frac{\mu(\mathbb{S}^2)}{\mu(\mathbb{S}^2) - 2\pi(1 - \cos r)\bar{\rho}}, \end{aligned}$$

where  $P$  is a finite grid of points on  $\mathbb{S}^2$  and,

$$A(\mathbf{x}, \mathbf{p}) = 1 - \frac{2\bar{\rho}}{\mu(\mathbb{S}^2)} 2\pi(1 - \cos r) + \frac{1}{\mu(\mathbb{S}^2)} \int_{B_{\mathbb{S}^2}(\mathbf{x}, r) \cap B_{\mathbb{S}^2}(\mathbf{p}, r)} \frac{\bar{\rho}^2}{\rho(\mathbf{y})} \lambda_{\mathbb{S}^2}(\mathbf{d}\mathbf{y}).$$

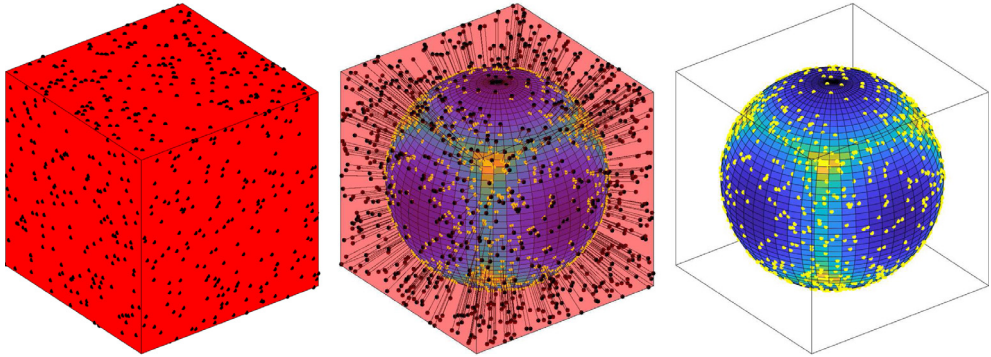
**Proof.** See [Proposition 1](#), Section S5 of the Supplementary Material.  $\square$

Using a Taylor series expansion (see Section S5.2 of the Supplementary Material), we can approximate the expectation and variance of  $\hat{J}_{\text{inhom}}(r)$  as

$$\mathbb{E} \left[ \frac{X}{Y} \right] \approx \frac{\mu_X}{\mu_Y} - \frac{\text{Cov}(X, Y)}{\mu_Y^2} + \frac{\text{Var}(Y)\mu_X}{\mu_Y^3} \tag{16}$$

$$\text{Var} \left( \frac{X}{Y} \right) \approx \frac{\mu_X}{\mu_Y} \left[ \frac{\text{Var}(X)}{\mu_X^2} - 2 \frac{\text{Cov}(X, Y)}{\mu_X \mu_Y} + \frac{\text{Var}(Y)}{\mu_Y^2} \right], \tag{17}$$

where  $X = 1 - \hat{H}_{\text{inhom}}(r)$  and  $Y = 1 - \hat{F}_{\text{inhom}}(r)$ . The terms in Eqs. (16) and (17) are given in [Theorems 3 and 4](#), and [Proposition 1](#).



**Fig. 1.** Example of simulating and mapping a CSR process on the cube to the sphere. *Left:* example of a CSR process on a cube with  $l = 1$  and constant intensity function 50. *Middle:* mapping of events from cube to the sphere by the function  $f(\mathbf{x}) = \mathbf{x}/\|\mathbf{x}\|$ . *Right:* mapped point pattern on the sphere with the new intensity function indicated by the colour on the sphere. High intensity is indicated by light areas whilst low intensity is indicated by dark areas.

### 6. Examples

We now look at two examples where we simulate homogeneous Poisson processes on their surfaces and construct the previously described functional summary statistics.

#### 6.1. Cube

We define a centred cube over each of the six faces with a side length  $2l$ , where  $l = 1$ . The level-set function for a cube is,

$$g(\mathbf{x}) = \begin{cases} x_3 - l, & \text{for } -l \leq x_1, x_2 \leq l \\ x_3 + l, & \text{for } -l \leq x_1, x_2 \leq l \\ x_2 - l, & \text{for } -l \leq x_1, x_3 \leq l \\ x_2 + l, & \text{for } -l \leq x_1, x_3 \leq l \\ x_1 - l, & \text{for } -l \leq x_2, x_3 \leq l \\ x_1 + l, & \text{for } -l \leq x_2, x_3 \leq l. \end{cases}$$

Using [Theorem 2](#) we can derive the intensity function for the point process that is mapped to the sphere. By symmetry we need only consider one of the faces of the cube and by rotation we will be able to derive the intensity function on the sphere. Consider the bottom face, i.e.  $x_3 = -l$ , and in the notation of [Theorem 2](#) label this  $\mathbb{D}_1$ . Then,

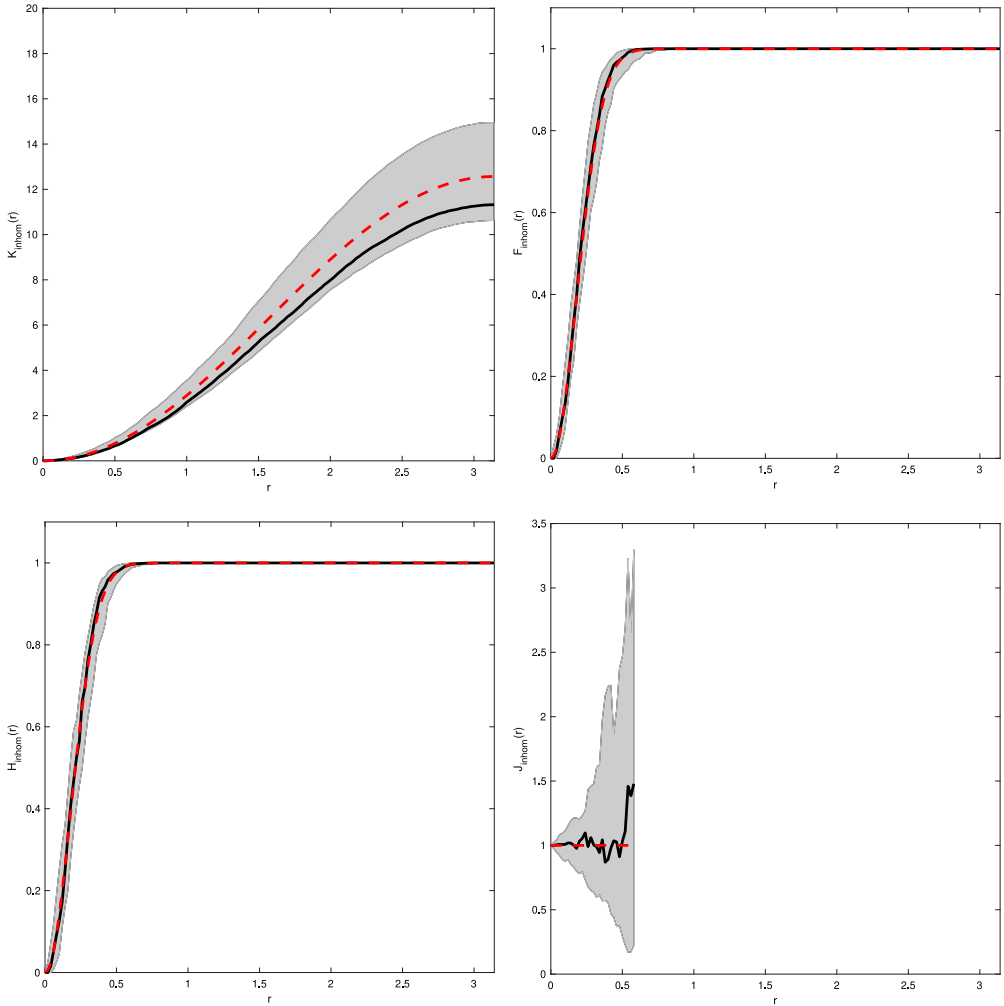
$$l_1(\mathbf{x}) = 1 \\ J_{(1,f^*)}(\mathbf{x}) = (1 + x_1 + x_2)^2,$$

and so the intensity function over  $f(\mathbb{D}_1)$  is,

$$\rho_1^*(\mathbf{x}) = \rho(1 + (f_1^{*-1}(x_1))^2 + (f_2^{*-1}(x_2))^2)(1 - x_1^2 - x_2^2)^{\frac{1}{2}},$$

where  $f_i(x_i) = x_i/\|\mathbf{x}\|$ . Thus by the appropriate rotations the intensity function over the entire sphere is,

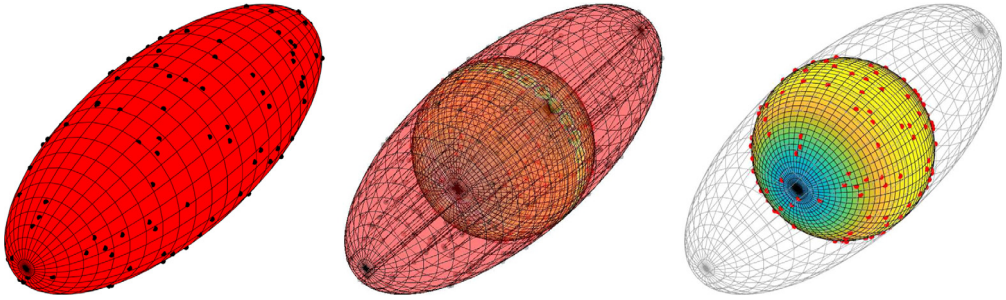
$$\rho^*(\mathbf{x}) = \begin{cases} \rho(1 + (f_1^{*-1}(x_1))^2 + (f_2^{*-1}(x_2))^2)(1 - x_1^2 - x_2^2)^{\frac{1}{2}}, & \mathbf{x} \in f(\mathbb{D}_1) \cup f(\mathbb{D}_2) \\ \rho(1 + (f_1^{*-1}(x_1))^2 + (f_3^{*-1}(x_3))^2)(1 - x_1^2 - x_3^2)^{\frac{1}{2}}, & \mathbf{x} \in f(\mathbb{D}_3) \cup f(\mathbb{D}_4) \\ \rho(1 + (f_2^{*-1}(x_2))^2 + (f_3^{*-1}(x_3))^2)(1 - x_2^2 - x_3^2)^{\frac{1}{2}}, & \mathbf{x} \in f(\mathbb{D}_5) \cup f(\mathbb{D}_6), \end{cases}$$



**Fig. 2.** Examples of  $K_{\text{inhom}}$ - (top left),  $F_{\text{inhom}}$ - (top right),  $H_{\text{inhom}}$ - (bottom left), and  $J_{\text{inhom}}$ - (bottom right) functions for CSR patterns on a cube with  $l = 1$  and  $\rho = 5$ . Solid line is the estimated functional summary statistic for our observed data, dashed line is the theoretical functional summary statistic for a Poisson process, and the grey shaded area represents the simulation envelope from 99 Monte Carlo simulations of Poisson processes fitted to the observed data.

where  $\mathbb{D}_1, \mathbb{D}_2, \mathbb{D}_3, \mathbb{D}_4, \mathbb{D}_5$ , and  $\mathbb{D}_6$  are the faces such that  $z = -1, z = 1, y = -1, y = 1, x = -1$ , and  $x = 1$  respectively. Fig. 1 demonstrates mapping from a cube with  $l = 1$  and  $\rho = 50$  to the unit sphere where the shading over the sphere indicates areas of low (dark) and high (light) intensity. The figure also shows an example of a CSR pattern over the cube and how this pattern changes under the mapping.

In order to be able to construct the inhomogeneous  $F$ -, and  $H$ -function we need to determine  $\inf_{\mathbf{x} \in \mathbb{S}^2} \rho^*(\mathbf{x})$ . By the nature of the function  $f(\mathbf{x}) = \mathbf{x}/\|\mathbf{x}\|$  and assuming that  $l \geq 1$ , then mapping events from the cube to the sphere causes events to be more concentrated on the sphere compared to the cube, thus increasing the corresponding intensity on the sphere. Therefore, the lowest achievable intensity occurs at the centre of each face of the cube, i.e. for the bottom face it occurs



**Fig. 3.** Example of simulating and mapping a CSR process on a prolate ellipsoid to the sphere. *Left:* example of a CSR process on a prolate ellipsoid with  $a = b = 1, c = 3$  and  $\rho = 5$ . *Middle:* mapping of events from prolate ellipsoid to the sphere by the function  $f(\mathbf{x}) = (x_1/a, x_2/b, x_3/c)^T$ . *Right:* mapped point pattern on the sphere with the new intensity function indicated by the colour on the sphere. High intensity is indicated by light areas whilst low intensity is indicated by dark areas.

when  $x_1 = x_2 = 0$ , giving  $\inf_{\mathbf{x} \in \mathbb{S}^2} \rho^*(\mathbf{x}) = \rho$ . Fig. 2 gives examples of the inhomogeneous  $K$ -,  $F$ -,  $H$ -, and  $J$ -functions where  $l = 1$  and  $\rho = 5$  and are typical when the observed process is CSR.

### 6.2. Ellipsoid

An ellipsoid is defined by its semi-major axis lengths  $a, b, c \in \mathbb{R}$  along  $x$ -,  $y$ -, and  $z$ -axis respectively. Again we also assume that the ellipsoid is centred at the origin. The level-set function for ellipsoids is given by  $g(\mathbf{x}) = x_1^2/a^2 + x_2^2/b^2 + x_3^2/c^2 - 1$ , in this current form  $\tilde{g}$  is not well defined in which case we shall use the following equivalent representation

$$g(\mathbf{x}) = \begin{cases} x_1^2/a^2 + x_2^2/b^2 + x_3^2/c^2 - 1, & \text{for } x_3 \geq 0 \\ x_1^2/a^2 + x_2^2/b^2 + x_3^2/c^2 - 1, & \text{for } x_3 < 0. \end{cases}$$

This representation then allows for  $\tilde{g}$  to be well defined for each partition of the ellipsoid.

We now demonstrate our methodology on an ellipsoid with semi-major axis lengths  $a = 1, b = 1$ , and  $c = 3$  along the  $x$ -,  $y$ -, and  $z$ -axis respectively. Instead of using the function  $f(\mathbf{x}) = \mathbf{x}/\|\mathbf{x}\|$  to map from the ellipsoid to the sphere, we can use a simpler mapping function which makes calculation of the determinant  $J_{(i,f^*)}(\mathbf{x})$  in Theorem 2 significantly easier. We can simply scale along the axis directions, i.e. use the mapping  $f(\mathbf{x}) = (x_1/a, x_2/b, x_3/c)^T$ . Using this mapping function, as opposed to dividing each vector by the norm of itself, and focusing on the bottom hemiellipsoid (indicated by the minus superscript), then

$$L_-(\mathbf{x}) = \sqrt{\frac{1 - \left(1 - \frac{c^2}{a^2}\right)x_1^2 - \left(1 - \frac{c^2}{b^2}\right)x_2^2}{1 - x_1^2 - x_2^2}}, \quad J_{(-f^*)}(\mathbf{x}) = ab,$$

and so on the lower hemisphere the intensity function takes the form

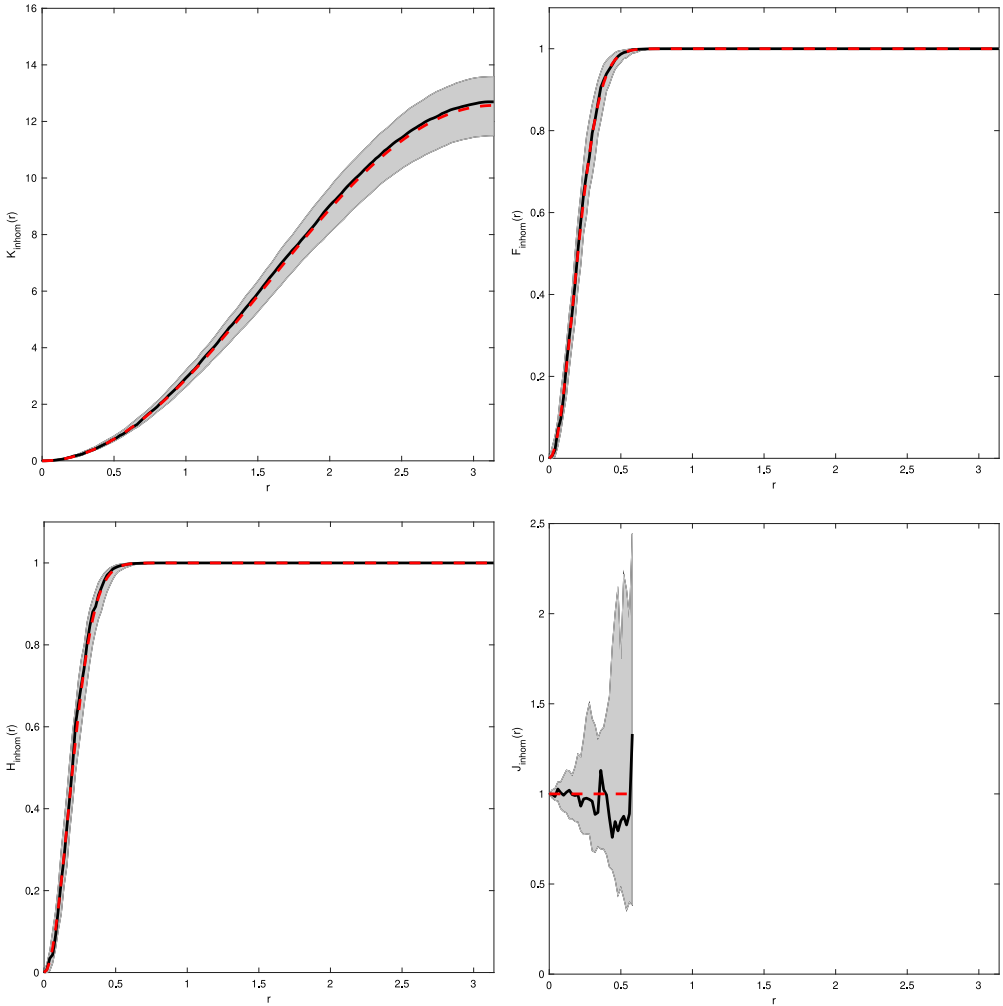
$$\rho_-^*(\mathbf{x}) = \rho ab \sqrt{1 - \left(1 - \frac{c^2}{a^2}\right)x_1^2 - \left(1 - \frac{c^2}{b^2}\right)x_2^2}.$$

By symmetry the mapped intensity function over the whole sphere is then

$$\rho^*(\mathbf{x}) = \rho ab \sqrt{1 - \left(1 - \frac{c^2}{a^2}\right)x_1^2 - \left(1 - \frac{c^2}{b^2}\right)x_2^2}.$$

Again we need to calculate  $\inf_{\mathbf{x} \in \mathbb{S}^2} \rho^*(\mathbf{x})$ . Noting that  $c \geq a = b$ , thus  $-(1 - c^2/a^2) \geq 0$  and  $-(1 - c^2/b^2) \geq 0$ , then the square root term is minimised when  $x_1$  and  $x_2$  are 0, hence





**Fig. 4.** Examples of  $K_{\text{inhom}}$ - (top left),  $F_{\text{inhom}}$ - (top right),  $H_{\text{inhom}}$ - (bottom left), and  $J_{\text{inhom}}$ - (bottom right) functions for CSR patterns on a prolate ellipsoid with  $a = b = 1, c = 3$ , and  $\rho = 5$ . Solid line is the estimated functional summary statistic for our observed data, dashed line is the theoretical functional summary statistic for a Poisson process, and the grey shaded area represents the simulation envelope from 99 Monte Carlo simulations of Poisson processes fitted to the observed data.

$\inf_{\mathbf{x} \in \mathbb{S}^2} \rho^*(\mathbf{x}) = \rho ab$ . An example of this mapping is given in Fig. 3. Using this we can construct the estimators of the inhomogeneous functional summary statistics given by Eqs. (12)–(15). Examples are given in Fig. 4. These figures are typical for CSR with the estimated functional summary statistics lying well within the simulation envelopes.

**7. Regular & cluster processes on  $\mathbb{D}$**

We examine some regular and cluster processes on  $\mathbb{D}$ . In particular, we examine how functional summary statistics constructed under the Poisson hypothesis deviate when the underlying process is in fact not Poisson. We shall be using the Matérn I and II inhibition processes (Chiu et al., 2013)

as examples of regular processes, and Thomas processes as a cluster example. Definitions for the Matérn I, II and Thomas processes on convex shapes will also be presented, whilst properties of such processes are given in Section S6 of the Supplementary Material.

7.1. Examples of regular and cluster processes on convex shapes

A common way of defining a regular process is using a minimum distance  $R$ , known as the hardcore distance, for which no events in the process has a nearest neighbour closer than  $R$ . In typical applications  $R$  is usually the Euclidean distance (in  $\mathbb{R}^d$ ) or the great circle distance (in  $\mathbb{S}^2$ ), but on an arbitrary three dimensional convex shape,  $\mathbb{D}$ , this distance is taken as the geodesic distance defined by the surface. The following definitions extend the Matérn I and II processes to a convex shape with geodesic distance  $d(\mathbf{x}, \mathbf{y})$ ,  $\mathbf{x}, \mathbf{y} \in \mathbb{D}$ .

**Definition 3.** Let  $X$  be a homogeneous Poisson process on  $\mathbb{D}$  with intensity  $\rho \in \mathbb{R}_+$ . Fix  $R > 0$ , and thin  $X$  according to the following rule: delete events  $\mathbf{x} \in X$  if there exists  $\mathbf{y} \in X \setminus \{\mathbf{x}\}$  such that  $d(\mathbf{x}, \mathbf{y}) < R$ , otherwise retain  $\mathbf{x}$ . The resulting thinned process is then defined as a Matérn I inhibition process on  $\mathbb{D}$ .

**Definition 4.** Let  $X$  be a homogeneous Poisson process on  $\mathbb{D}$  with intensity  $\rho \in \mathbb{R}_+$ . Fix  $R > 0$ , and let each  $\mathbf{x} \in X$  have an associated mark,  $M_{\mathbf{x}}$  drawn from some mark density  $P_M$  independently of all other marks and events in  $X$ . Thin  $X$  according to the following rule: delete the event  $\mathbf{x} \in X$  if there exists  $\mathbf{y} \in X \setminus \{\mathbf{x}\}$  such that  $d(\mathbf{x}, \mathbf{y}) < R$  and  $M_{\mathbf{y}} < M_{\mathbf{x}}$ , otherwise retain  $\mathbf{x}$ . The resulting thinned process is then defined as a Matérn II inhibition process on  $\mathbb{D}$ .

We also extend the Neyman–Scott process, a class of cluster processes, to arbitrary convex shapes.

**Definition 5.** Let  $X_p$  be a homogeneous Poisson process on  $\mathbb{D}$  with intensity  $\rho \in \mathbb{R}_+$ . Then for each  $\mathbf{c} \in X_p$  define  $X_{\mathbf{c}}$  to the point process with intensity function  $\rho_{\mathbf{c}}(\mathbf{x}) = \alpha k(\mathbf{x}, \mathbf{c})$ , where  $\alpha > 0$  and  $k : \mathbb{D} \times \mathbb{D} \mapsto \mathbb{R}$  is a density function and  $N_{X_{\mathbf{c}}}(\mathbb{D})$  can be any random counting measure associated to  $X_{\mathbf{c}}$ . The point process  $X = \cup_{\mathbf{c} \in X_p} X_{\mathbf{c}}$  is a Neyman–Scott process.

A Thomas process is a specific Neyman–Scott process where the density function  $k(\cdot, \cdot)$  has a specific form. In  $\mathbb{R}^2$ ,  $k$  is taken to be an isotropic bivariate Gaussian distribution (Møller and Waagepetersen, 2004), whilst on  $\mathbb{S}^2$  it is taken as the Von-Mises Fisher distribution (Lawrence et al., 2016). We define a Thomas process on  $\mathbb{D}$  to be a Neyman–Scott process with density function  $k$  of the form,

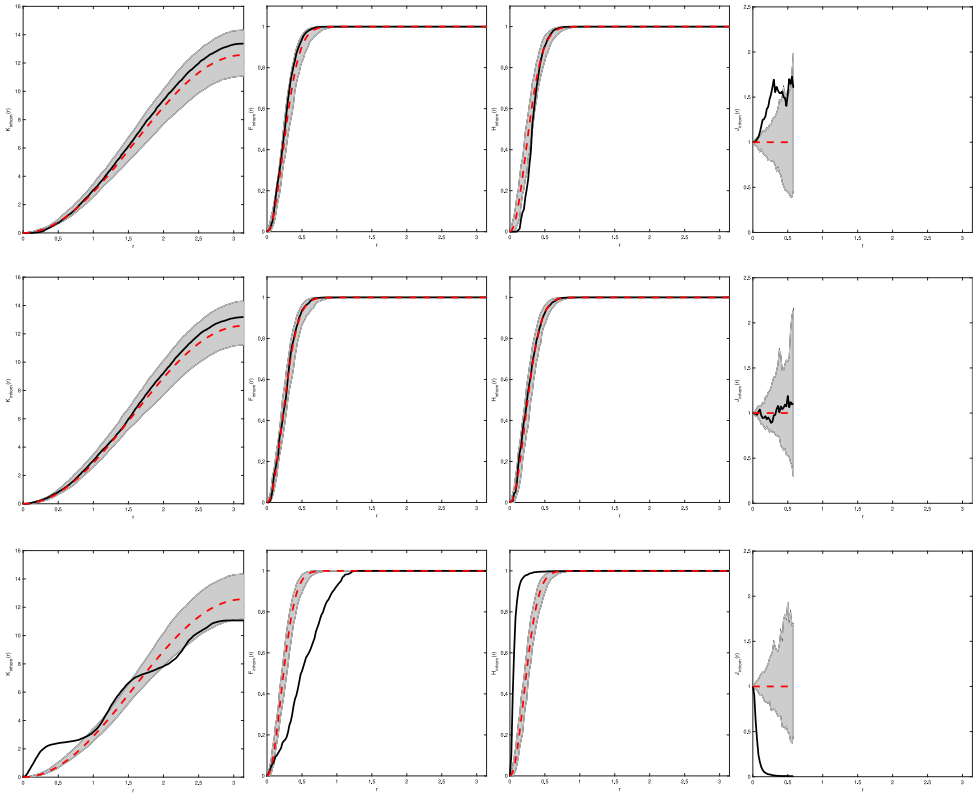
$$k(\mathbf{x}, \mathbf{y}) = \frac{1}{\chi(\kappa^2)} \exp\left(-\frac{d^2(\mathbf{x}, \mathbf{y})}{2\kappa^2}\right),$$

where  $\kappa$  is a bandwidth parameter and  $\chi(\kappa) = \int_{\mathbb{D}} \exp(-d(\mathbf{x}, \mathbf{y})/2\kappa^2) \lambda_{\mathbb{D}}(\mathbf{y})$ . This is known as the Riemannian Gaussian distribution (Said et al., 2017), where on the plane this would reduce to an isotropic bivariate Gaussian.

7.2. Functional summary statistics assuming a homogeneous Poisson process

We simulate Matérn II, and Thomas processes and construct estimates of their functional summary statistics under the assumption that they are CSR. The inhomogeneous functional summary statistics are displayed in Fig. 5. Comparing Fig. 5 to typical functional summary statistics for regular and cluster processes in  $\mathbb{R}^2$ , we see the same types of deviations away from CSR. In particular, we see for regular processes with small  $r$  that there are negative deviations, whilst the cluster process has large positive deviations for the  $\tilde{K}_{\text{inhom}}$ -function. Furthermore, the  $\hat{J}_{\text{inhom}}$ -function shows significant positive deviations for regular processes whilst negative ones are observed for cluster processes.

Fig. 5 highlights the importance of considering many different functional summary statistics when attempting to draw conclusions from the data (Diggle, 2003). More precisely consider the top



**Fig. 5.** Example of (from left to right)  $K_{\text{inhom}}$ ,  $J_{\text{inhom}}$ ,  $F_{\text{inhom}}$ , and  $H_{\text{inhom}}$ -functions for a Matérn II with parameters  $R = 0.3$  and exponential mark distribution with rate  $\lambda = 1$  (top row), Poisson process (middle row), and Thomas process with parameters  $\kappa = 0.1$  and offspring expectation 15 (bottom row) on a prolate spheroid with dimensions  $(a, b, c) = (1, 1, 3)$  all with expectation 100. Solid line is the estimated functional summary statistics for our observed data, dashed line is the theoretical functional summary statistic for a Poisson process, and the grey shaded area is the simulation envelopes from 99 Monte Carlo simulations of Poisson processes fitted to the observed data.

row of Fig. 5 which refers to a Matérn II process with hardcore distance  $R = 0.3$  and  $\mathbb{E}[N_X(\mathbb{D})] = 100$ . If we were to only consult the  $K_{\text{inhom}}$  simulation envelope (top left figure) then we may be hesitant to reject CSR, but when we examine the  $H_{\text{inhom}}$  and  $J_{\text{inhom}}$  simulation envelope plots there is strong evidence to suggest that this process is not CSR and is in fact regular. Thus, for this specific setting, the simulation envelope plots of the  $J_{\text{inhom}}$  function provide greater power compared to those produced with the  $K_{\text{inhom}}$  function, especially for smaller  $r$  (Baddeley et al., 2015, p. 235). This effect is also seen in Figure 1 of van Lieshout (2011) for Log Gaussian Cox processes on  $\mathbb{R}^2$ , where for small values of  $r$  the  $K_{\text{inhom}}$  is unable to provide evidence against the process being inhomogeneous Poisson but examining both the  $F_{\text{inhom}}$  and  $H_{\text{inhom}}$  provides greater evidence against this hypothesis.

This phenomena was first discussed by Baddeley and Silverman (1984), where they show that it is possible to construct planar processes which are not homogeneous Poisson but have the same  $K$ -function. This is also discussed by Baddeley et al. (2000) in Section 2.4 in the inhomogeneous setting. Similar arguments to those in Baddeley et al. (2000) and Baddeley and Silverman (1984) can be used to construct spheroidal processes that are not Poisson but exhibit the same  $K$ -function as any given Poisson process with the same intensity. This discussion serves as a precautionary warning to consulting only one individual functional summary statistic and it is therefore important to consider many to avoid drawing improper conclusions from data.

### 8. Testing for CSR on convex shapes in $\mathbb{R}^3$

Exploratory data analysis for spatial point patterns in  $\mathbb{R}^2$  typically begins with testing whether the observed point pattern exhibits CSR where test statistics are frequently based on the  $L$ -function,  $L(r) = \sqrt{K(r)/\pi}$ . On  $\mathbb{R}^2$  and under CSR the  $L$ -function is linear in  $r$  and variance stabilised (Besag, 1977) whilst Lawrence (2018) discusses the analogue  $L$ -function in  $\mathbb{S}^2$  where again it is variance stabilised when the underlying process is CSR. As we are working with inhomogeneous Poisson processes on  $\mathbb{S}^2$  an equivalent transformation for the  $L$  function has not been discussed previously. We propose two test statistics:

1. an extension of the analogue  $L$  - function proposed by Lawrence (2018) to the inhomogeneous setting, and
2. a standardisation of the inhomogeneous  $K$  - function inspired by Lagache et al. (2013).

In order to construct the latter we must derive first and second order properties of the estimated functional summary statistics. Section 7 discusses derivations for any spherical Poisson process when  $\rho$  is known. In this section we consider the scenario when we have a homogeneous Poisson process on  $\mathbb{D}$  with unknown intensity  $\rho \in \mathbb{R}_+$ . Furthermore we shall only focus on the inhomogeneous  $K$ -function as standardisation of the remaining functional summary statistics follow identically.

#### 8.1. Test statistic for CSR

Given a homogeneous Poisson process on  $\mathbb{D}$  with intensity  $\rho \in \mathbb{R}_+$ , we map this to  $\mathbb{S}^2$  giving a new Poisson process on the sphere with inhomogeneous intensity function given by Theorem 2 as

$$\rho^*(\mathbf{x}) = \begin{cases} \rho l_1(f^{-1}(\mathbf{x}))J_{(1,f^*)}(\mathbf{x})\sqrt{1-x_1^2-x_2^2}, & \mathbf{x} \in f(\mathbb{D}_1) \\ \vdots \\ \rho l_n(f^{-1}(\mathbf{x}))J_{(n,f^*)}(\mathbf{x})\sqrt{1-x_1^2-x_2^2}, & \mathbf{x} \in f(\mathbb{D}_n). \end{cases} \tag{18}$$

Using Theorems 3 and 4 we can calculate the mean and variances of the inhomogeneous  $K$ -function when  $\rho$  is known. When  $\rho$  is unknown we use estimators of  $\rho$  when constructing functional summary statistics. In particular we use

$$\hat{\rho} = \frac{N_X(\mathbb{D})}{\lambda_{\mathbb{D}}(\mathbb{D})}, \quad \hat{\rho}^2 = \frac{N_X(\mathbb{D})(N_X(\mathbb{D}) - 1)}{\lambda_{\mathbb{D}}^2(\mathbb{D})},$$

which are both unbiased for  $\rho$  and  $\rho^2$  respectively by application of the Campbell–Mecke Theorem (Møller and Waagepetersen, 2004). Thus our estimator for  $K_{\text{inhom}}(r)$  when  $\rho$  is unknown takes the following form,

$$\tilde{K}_{\text{inhom}}(r) = \begin{cases} \frac{\lambda_{\mathbb{D}}^2(\mathbb{D})}{4\pi N_Y(\mathbb{S}^2)(N_Y(\mathbb{S}^2)-1)} \sum_{\mathbf{x} \in Y} \sum_{\mathbf{y} \in Y \setminus \{\mathbf{x}\}} \frac{\mathbb{1}_{\{d(\mathbf{x}, \mathbf{y}) \leq r\}}}{\tilde{\rho}(\mathbf{x})\tilde{\rho}(\mathbf{y})}, & \text{if } N_Y(\mathbb{S}^2) > 1 \\ 0, & \text{otherwise,} \end{cases} \tag{19}$$

where  $Y = f(X)$ ,  $f$  is our mapping from the ellipsoid to the sphere, and  $\tilde{\rho}(\mathbf{x})$  is given by,

$$\tilde{\rho}(\mathbf{x}) = \begin{cases} l_1(f^{-1}(\mathbf{x}))J_{(1,f^*)}(\mathbf{x})\sqrt{1-x_1^2-x_2^2}, & \mathbf{x} \in f(\mathbb{D}_1) \\ \vdots \\ l_n(f^{-1}(\mathbf{x}))J_{(n,f^*)}(\mathbf{x})\sqrt{1-x_1^2-x_2^2}, & \mathbf{x} \in f(\mathbb{D}_n). \end{cases} \tag{20}$$

Note that  $N_Y(\mathbb{S}^2) = N_X(\mathbb{D}^2)$ .

An analogue  $L$  - function for isotropic spherical point processes was introduced by Lawrence (2018) in which the square root of Ripley’s spherical  $K$  - function is taken. This transformation benefits from approximate variance stabilisation in the same sense as the  $L$  - function does in  $\mathbb{R}^d$  (Besag, 1977) but is not linearised. In the planar setting a multiplicative factor of  $1/\sqrt{\pi}$  can

be used such that  $L(r) = r$  but due to the more complex form of  $K$  on  $\mathbb{S}^2$  a simple linearising transformation is not intuitive. Therefore Lawrence (2018) suggests subtracting the theoretical value in order for the summary statistic to be zero in the event a process is Poisson. Following this line of thought we then propose, in the inhomogeneous setting, the following functional summary statistic  $P_{\text{inhom}}(r) = \sqrt{\tilde{K}_{\text{inhom}}(r)} - \sqrt{2\pi(1 - \cos(r))}$ , where we use  $P_{\text{inhom}}$  rather than  $L_{\text{inhom}}$  to avoid confusion with the Euclidean  $L$ -function as  $P_{\text{inhom}}$  takes a very different form to its Euclidean counterpart. In the event of a homogeneous Poisson process over  $\mathbb{D}$  (and hence a inhomogeneous Poisson process over  $\mathbb{S}^2$ ) we can estimate  $P_{\text{inhom}}$  as,

$$\tilde{P}_{\text{inhom}}(r) = \sqrt{\tilde{K}_{\text{inhom}}(r)} - \sqrt{2\pi(1 - \cos(r))}.$$

Diggle (2003) proposes using the maximum absolute value between the theoretical and the estimated functional summary statistics to test for CSR. Based on this we propose the following two test statistics,

$$T_1 = \sup_{r \in [0, \pi]} |\tilde{P}_{\text{inhom}}(r)|, \quad T_2 = \sup_{r \in [0, \pi]} \left| \frac{\tilde{K}_{\text{inhom}}(r) - 2\pi(1 - \cos(r))}{\sqrt{\widehat{\text{Var}}(\tilde{K}_{\text{inhom}}(r))}} \right|, \tag{21}$$

the first based on the work of Lawrence (2018) and the second on the work of Lagache et al. (2013). In order to be able to construct the test statistic  $T_2$ , an estimate of the variance of the empirical functional summary statistics are required. Further, we need show that the bias of  $\tilde{K}_{\text{inhom}}(r)$  is negligible and hence  $\mathbb{E}[\tilde{K}_{\text{inhom}}(r)] \approx 2\pi(1 - \cos(r))$  for Poisson processes, validating its use in (21). By using estimators for  $\rho$  and  $\rho^2$  we alter the first and second order properties given by Theorems 3 and 4. In the following we consider the first and second order moments of  $\tilde{K}_{\text{inhom}}$ .

### 8.2. Estimating moments of $\tilde{K}_{\text{inhom}}(r)$ on $\mathbb{S}^2$ for CSR process on $\mathbb{D}$

**Theorem 6.** *The bias and variance of  $\tilde{K}_{\text{inhom}}(r)$  are,*

$$\text{Bias}(\tilde{K}_{\text{inhom}}(r)) = -P(N_Y(\mathbb{S}^2) \leq 1)2\pi(1 - \cos r),$$

and,

$$\begin{aligned} \text{Var}(\tilde{K}_{\text{inhom}}(r)) &= 4\pi^2(1 - \cos r)^2(1 - P(N_Y(\mathbb{S}^2) \leq 1))P(N_Y(\mathbb{S}^2) \leq 1) \\ &+ \rho^3 \lambda_{\mathbb{D}}^4(\mathbb{D})(1 - \cos r)^2 \left( \int_{\mathbb{S}^2} \frac{1}{\tilde{\rho}(\mathbf{x})} \lambda_{\mathbb{S}^2}(d\mathbf{x}) - \frac{16\pi^2}{\lambda_{\mathbb{D}}(\mathbb{D})} \right) \mathbb{E} \left[ \frac{1}{(N_Y(\mathbb{S}^2) + 3)^2(N_Y(\mathbb{S}^2) + 2)^2} \right] \\ &+ \frac{\rho^2 \lambda_{\mathbb{D}}^4(\mathbb{D})}{8\pi^2} \left( \int_{\mathbb{S}^2} \int_{\mathbb{S}^2} \frac{\mathbb{1}[d(\mathbf{x}_1, \mathbf{x}_2) \leq r]}{\tilde{\rho}(\mathbf{x}_1)\tilde{\rho}(\mathbf{x}_2)} \lambda_{\mathbb{S}^2}(d\mathbf{x}_1)\lambda_{\mathbb{S}^2}(d\mathbf{x}_2) - \frac{64\pi^4(1 - \cos r)^2}{\lambda_{\mathbb{D}}^2(\mathbb{D})} \right) \\ &\times \mathbb{E} \left[ \frac{1}{(N_Y(\mathbb{S}^2) + 2)^2(N_Y(\mathbb{S}^2) + 1)^2} \right], \end{aligned} \tag{22}$$

where  $\tilde{\rho}(\mathbf{x})$  is given by Eq. (20).

**Proof.** See Theorem 6, Section S7 of the Supplementary Material.  $\square$

The form of the variance derived in Theorem 6 is near identical to that derived by Lang and Marcon (2013) except that our derivations consider inhomogeneous Poisson processes, does not require corrections for edge effects, and the space is  $\mathbb{S}^2$  instead of  $\mathbb{R}^2$ . Further we can bound the absolute value of the bias as follows

$$\begin{aligned} |\text{Bias}(\tilde{K}_{\text{inhom}}(r))| &= P(N_Y(\mathbb{S}^2) \leq 1)2\pi(1 - \cos r) \\ &= \exp(-\mu_Y(\mathbb{S}^2))(1 + \mu_Y(\mathbb{S}^2))2\pi(1 - \cos r) \\ &\leq 4\pi(1 + \mu_Y(\mathbb{S}^2))\exp(-\mu_Y(\mathbb{S}^2)) \end{aligned} \tag{23}$$

$$\leq 4\pi(1 + \mu_Y(\mathbb{S}^2))\mu_Y(\mathbb{S}^2)^{-e} = \mathcal{O}(\mu_Y^{1-e}(\mathbb{S}^2)), \tag{24}$$

where  $\mu_Y$  is the intensity measure of  $Y = f(X)$ , the inequality in (23) is attained by setting  $r = \pi$ , and (24) follows from  $e^x \geq x^e$ . Thus, for shapes considered in this work, the bias will be negligible.

From Theorem 6 it is possible to construct a ratio-unbiased estimator for the variance. In particular by the Campbell–Mecke Theorem, and defining the estimator  $\hat{\rho}_k = N_Y(\mathbb{S}^2)(N_Y(\mathbb{S}^2) - 1) \cdots (N_Y(\mathbb{S}^2) - k - 1)/\lambda_{\mathbb{D}}^k(\mathbb{D})$ , then  $\mathbb{E}[\hat{\rho}_k] = \rho^k$  and so  $\hat{\rho}_k$  is unbiased for  $\rho^k$ . We can substitute the expectations in (22) with their corresponding observed values, for example we substitute  $(N_Y(\mathbb{S}^2) + 3)^{-2}(N_Y(\mathbb{S}^2) + 2)^{-2}$  for  $\mathbb{E}[(N_Y(\mathbb{S}^2) + 3)^{-2}(N_Y(\mathbb{S}^2) + 2)^{-2}]$ . Additionally, the following lemma helps derive a ratio unbiased estimator for  $P(N_Y(\mathbb{S}^2) < 1)$ .

**Lemma 2.** *Let  $N \sim \text{Poisson}(\lambda)$ ,  $k \in \mathbb{N}$  and  $p \in \mathbb{R}_+$ . Define the following random variable,*

$$R = \frac{N!e^{N-k}}{(N - k)!(e + p)^N}. \tag{25}$$

Then  $R$  is ratio-unbiased for  $\lambda^k e^{-p\lambda}$ .

**Proof.** See Lemma 2, Section S7 of the Supplementary Material.  $\square$

Using Lemma 2 we can construct a ratio-unbiased estimator for  $(1 - P(N_Y(\mathbb{S}^2) < 1))P(N_Y(\mathbb{S}^2) < 1)$ . Defining  $\lambda = \rho\lambda_{\mathbb{D}}(\mathbb{L})$ ,

$$\begin{aligned} (1 - P(N_Y(\mathbb{S}^2) < 1))P(N_Y(\mathbb{S}^2) < 1) &= (1 - e^{-\lambda} - \lambda e^{-\lambda})(e^{-\lambda} + \lambda e^{-\lambda}) \\ &= e^{-\lambda} + \lambda e^{-\lambda} - e^{-2\lambda} - 2\lambda e^{-2\lambda} - \lambda^2 e^{-2\lambda}, \end{aligned}$$

and so a ratio-unbiased estimator for  $(1 - P(N_Y(\mathbb{S}^2) < 1))P(N_Y(\mathbb{S}^2) < 1)$  is

$$\begin{aligned} &\frac{e^{N_Y(\mathbb{S}^2)}}{(e + 1)^{N_Y(\mathbb{S}^2)}} + \frac{N_Y(\mathbb{S}^2)e^{N_Y(\mathbb{S}^2)-1}}{(e + 1)^{N_Y(\mathbb{S}^2)}} + \frac{e^{N_Y(\mathbb{S}^2)}}{(e + 2)^{N_Y(\mathbb{S}^2)}} - \frac{2N_Y(\mathbb{S}^2)e^{N_Y(\mathbb{S}^2)-1}}{(e + 2)^{N_Y(\mathbb{S}^2)}} \\ &- \frac{N_Y(\mathbb{S}^2)(N_Y(\mathbb{S}^2) - 1)e^{N_Y(\mathbb{S}^2)-2}}{(e + 2)^{N_Y(\mathbb{S}^2)}}. \end{aligned}$$

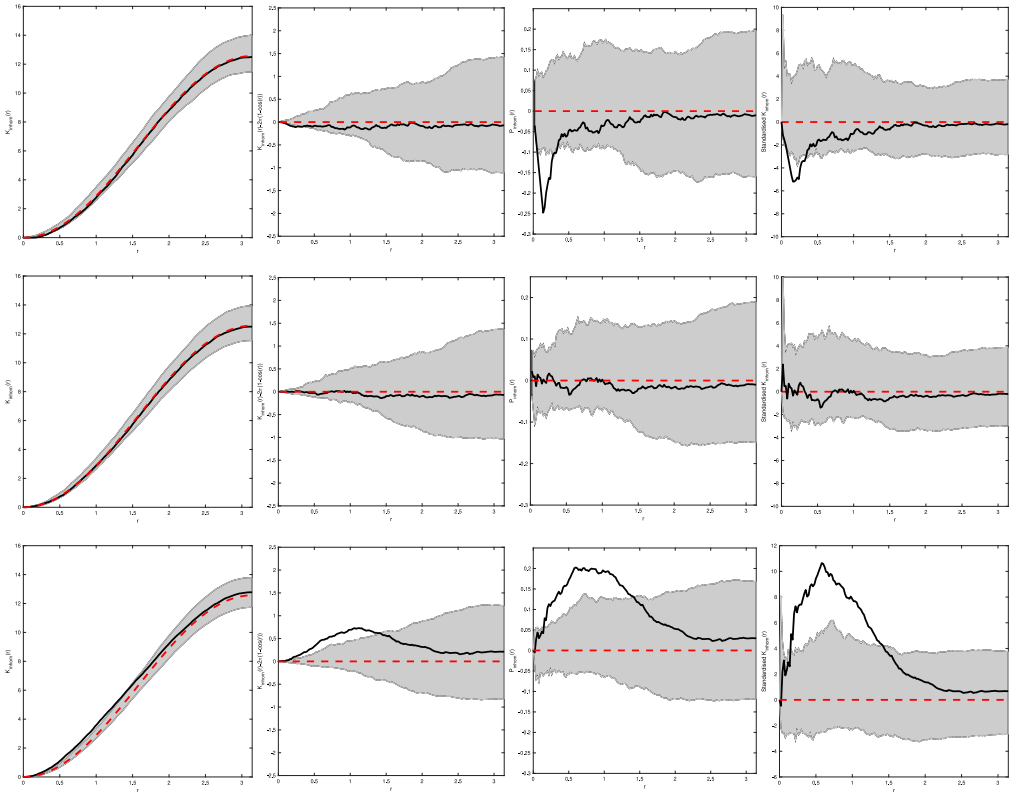
Plugging the given estimators for  $(1 - P(N_Y(\mathbb{S}^2) < 1))P(N_Y(\mathbb{S}^2) < 1)$ ,  $\rho^2$ ,  $\rho^3$ ,  $\mathbb{E}[(N_Y(\mathbb{S}^2) + 3)^{-2}(N_Y(\mathbb{S}^2) + 2)^{-2}]$  and  $\mathbb{E}[(N_Y(\mathbb{S}^2) + 2)^{-2}(N_Y(\mathbb{S}^2) + 1)^{-2}]$  into (22) gives a ratio unbiased estimator for  $\text{Var}(\tilde{K}_{\text{inhom}}(r))$ , which in turn allows for the construction of the test statistic  $T_2$  in (21).

### 8.3. Standardised inhomogeneous $K$ -function plots

Fig. 6 highlights how the empirical  $K$ -function estimates deviate when the underlying process is not CSR. For the regular processes we notice considerable negative deviations for small  $r$  whilst for cluster processes positive deviations are observed, highlighted in the right column of Fig. 6.

Intuitively, this is to be expected, with a near identical reasoning to what is observed for the  $K$ -function in  $\mathbb{R}^{2,3}$ . Since the regular process has a hard-core distance between events, we observe estimates for  $K_{\text{inhom}}(r)$  that are close to zero for small  $r$ , thus resulting in the large negative deviation observed in Fig. 6. On the other hand, for the Thomas cluster process, we observe events in closer proximity than would be expected for a CSR process, thus the estimated  $K_{\text{inhom}}(r)$  function has large positive deviations away from CSR.

Further to this, the second row of Fig. 6 corresponding to a Poisson process highlights the importance of applying a variance stabilising transform. If we based a test statistic on  $K_{\text{inhom}}(r) - 2\pi(1 - \cos(r))$  (or even  $K_{\text{inhom}}(r)$ ) then these plots suggest worst power compared to the  $P_{\text{inhom}}$ - and the standardised  $K_{\text{inhom}}$ -functions.



**Fig. 6.** Example of (from left to right)  $K_{\text{inhom}}(\cdot)$ ,  $K_{\text{inhom}}(\cdot) - 2\pi(1 - \cos(r))$ ,  $P_{\text{inhom}}(\cdot)$ , and the standardised  $K_{\text{inhom}}(\cdot)$ -functions for a Matérn II with  $R = 0.2$ , an exponential mark distribution with rate  $\lambda = 1$  and expectation  $100$  (top row), Poisson process with expectation  $40\pi$  (middle row), and Thomas process with parameters  $\kappa = 0.5$ , expectation of  $150$  and offspring mean of  $20$  (bottom row) on a prolate spheroid with  $a = b = 0.8000$ ,  $c = 1.43983$  (dimensions chosen so that the area of the ellipsoid is  $4\pi$ ). Solid line is the estimated functional summary statistics for our observed data, dashed line is the theoretical functional summary statistic for a Poisson process, and the grey shaded area is the simulation envelopes from 999 Monte Carlo simulations of Poisson processes fitted to the observed data.

### 9. Simulation study

We conduct empirical Type I and II error studies to evaluate the effectiveness of the proposed test statistic in determining whether or not a point process on a convex shape exhibits CSR. We consider different prolate ellipsoids such that the area of the ellipsoid is the same across differing semi-major axis lengths. This will allow us to determine how the power of our test changes as the space under consideration deforms further away from the unit sphere.

#### 9.1. Design of simulations

In order to best understand the properties of our testing procedure we will consider CSR, Matérn II and Thomas processes on different prolate spheroids. We design the experiments such that the expected number of events is similar across all experiments. For both the CSR and Thomas process simulations this is easily controlled. For a Poisson process, the expected number on  $\mathbb{D}$  is  $\rho\lambda_{\mathbb{D}}(\mathbb{D})$  whilst for a Thomas process it is given by Proposition S3, Section S6 in the Supplementary Materials.

On the other hand, the Matérn II process requires a little more attention since Corollary S4 (Section S6 of the Supplementary Materials) limits the maximum expected number of possible

**Table 1**

Results when the *observed* data is CSR. The semi-major axis length along the  $x$ -axis,  $a$ , and  $y$ -axis,  $b$ , are equivalent and the semi-major axis length along the  $z$ -axis is determined such that the area of the ellipsoid is  $4\pi$ .

Experiment No.	Expectation	$a$	$\rho$	Reject $H_0 : T_1$	Reject $H_0 : T_2$
1a	$40\pi$	1	10	0.0250	0.0480
1b	$40\pi$	0.8	10	0.0390	0.0390
1c	$40\pi$	0.6	10	0.0440	0.0430
1d	$40\pi$	0.4	10	0.0560	0.0560

events for a given space  $\mathbb{D}$ . Thus, for a given expected number  $\mu$  that is less than or equal to the one prescribed by Corollary S4 (Section 6 of the Supplementary Materials) we fix the hard-core distance  $R$  and solve the following equation for  $\rho$

$$\int_B \frac{1 - e^{-\rho \lambda_{\mathbb{D}}(B_{\mathbb{D}}(\mathbf{x}, R))}}{\lambda_{\mathbb{D}}(B_{\mathbb{D}}(\mathbf{x}, R))} \lambda_{\mathbb{D}}(d\mathbf{x}) = \mu. \tag{26}$$

A full outline of all the experiments and the parameters chosen are given in Tables 1, 2, and 3 for CSR, regular, and cluster process simulations respectively. Note that when  $R = 0$  for the Matérn II process and when  $\kappa = \infty$  for the Thomas process both processes are CSR.

9.2. Test statistics

Calculating  $|\tilde{P}_{\text{inhom}}(r)|$  and  $|(\tilde{K}_{\text{inhom}}(r) - 2\pi(1 - \cos r))/(\widehat{\text{Var}}(\tilde{K}_{\text{inhom}}(r)))^{1/2}|$  over all  $r \in [0, \pi]$  is computationally infeasible and so we instead calculate it for  $r \in \mathcal{R} = \{r_1, \dots, r_m\}$ ,  $m \in \mathbb{N}$ , where  $\mathcal{R}$  is a finite set of distinct, evenly spaced points such that  $r_i \in [0, \pi]$ ,  $i = 1, \dots, m$ , for the purposes of our simulation studies. We then take our test statistic as

$$T_1 = \max_{r \in \mathcal{R}} |\tilde{P}_{\text{inhom}}(r)|, \quad T_2 = \max_{r \in \mathcal{R}} \left| \frac{\tilde{K}_{\text{inhom}}(r) - 2\pi(1 - \cos(r))}{\sqrt{\widehat{\text{Var}}(\tilde{K}_{\text{inhom}}(r))}} \right|.$$

where for these simulation studies we set  $\mathcal{R} = \{0, 0.02, 0.04, \dots, \pi\}$ . These simulations are tested at a 5% significance level. Each experiment is repeated 1000 times, and for each experiment we simulate 999 Poisson processes to approximate the critical values of the hypothesis test.

9.3. Results

Tables 1–3 outline the parameter selection and results of our simulations. By the nature of Monte Carlo simulations the CSR results given in Table 1 are as to be expected with an empirical rejection rate close to 0.05. Expectedly, we see that for the same ellipsoid i.e.  $a$  kept constant, that when the Matérn II parameter  $R$  increases and the Thomas process parameter  $\kappa$  decreases (each representing an increased departure from CSR), the power of our test improves. In the Section S6.3 of the Supplementary Material we discuss a potential reason for the power of our test decreasing as  $a$  decreases (hence  $c$  increases), for both regular and cluster processes, for the same  $R$  and  $\kappa$  respectively. Additionally, Fig. 5 suggests we may gain power by considering a two sided test.

We see that  $T_1$  achieves greater empirical power compared to  $T_2$  over the majority of experiments considered in our simulation study. This is clearly the case when  $a = 1$  or 0.8 for the Matérn II process whilst when  $a = 0.6$  or 0.4 the distinction is not as clear. In particular,  $T_2$  achieves considerably greater empirical power for Experiment (2civ). For the Thomas process we see that  $T_1$  out performs  $T_2$  in nearly all experiments for all values of  $a$ . These result suggest the following considerations in practice: (1) consider  $T_1$  before  $T_2$  since computing  $T_1$  is simpler, (2) if  $T_1$  does not provide sufficient evidence against the null then  $T_2$  may provide additional information or even considering another functional summary statistic that has been developed in this work. Although a formal hypothesis test may be sought, these results emphasise that this should not be without a detailed examination of simulation envelope plots which can potentially provide further information. In particular, simulation envelope plots can give indications as to whether the point pattern exhibits more regular or clustered behaviour (Diggle, 2003).



**Table 2**

Results when the *observed* data is a Matérn II process, with independent mark being exponential with rate 1. The semi-major axis length along the  $x$ -axis,  $a$ , and  $y$ -axis,  $b$ , are equivalent and the semi-major axis length along the  $z$ -axis is determined such that the area of the ellipsoid is  $4\pi$ . Fixing the expectation,  $\mu$ , and hard-core distance,  $R$ , we use Eq. (26) to calculate  $\rho$  for the underlying constant Poisson process intensity function. When  $R = 0$  a Matérn II process collapses to a CSR process.

Experiment No.	Expectation	$a$	$R$	Reject $H_0 : T_1$	Reject $H_0 : T_2$
2ai	100	1	0	0.0450	0.0750
2aai	100	1	0.05	0.2520	0.0270
2aiii	100	1	0.1	1.0000	0.4550
2aiv	100	1	0.2	1.0000	1.0000
2bi	100	0.8	0	0.0440	0.0550
2bii	100	0.8	0.05	0.0460	0.0030
2biii	100	0.8	0.1	0.8460	0.0370
2biv	100	0.8	0.2	1.0000	1.0000
2ci	100	0.6	0	0.0520	0.0510
2cii	100	0.6	0.05	0.0540	0.0060
2ciii	100	0.6	0.1	0.0260	0.0010
2civ	100	0.6	0.2	0.2990	0.7790
2di	100	0.4	0	0.0440	0.0410
2dii	100	0.4	0.05	0.0400	0.0100
2diii	100	0.4	0.1	0.0270	0.0000
2div	100	0.4	0.2	0.0020	0.0020

**Table 3**

Results when the *observed* data is an ellipsoidal Thomas process. The expected number of offspring per parent is  $\lambda = 20$  and the underlying Poisson parent process has constant intensity function  $\rho = \mu/(4\pi\lambda)$ , where  $\mu$  is the expectation. The semi-major axis length along the  $x$ -axis,  $a$ , and  $y$ -axis,  $b$ , are equivalent and the semi-major axis length along the  $z$ -axis is determined such that the area of the ellipsoid is  $4\pi$ . When  $\kappa = \infty$  an ellipsoidal Thomas process collapses to a CSR process.

Experiment No.	Expectation	$a$	$\kappa$	Reject $H_0 : T_1$	Reject $H_0 : T_2$
3ai	150	1	$\infty$	0.0290	0.0440
3aai	150	1	5	0.0330	0.0470
3aiii	150	1	1	0.4620	0.5630
3aiv	150	1	0.5	0.9840	0.9830
3bi	150	0.8	$\infty$	0.0460	0.0540
3bii	150	0.8	5	0.0530	0.0570
3biii	150	0.8	1	0.2950	0.2120
3biv	150	0.8	0.5	0.9260	0.9340
3ci	150	0.6	$\infty$	0.0490	0.0460
3cii	150	0.6	5	0.0570	0.0610
3ciii	150	0.6	1	0.3730	0.1400
3civ	150	0.6	0.5	0.7480	0.7800
3di	150	0.4	$\infty$	0.0670	0.0600
3dii	150	0.4	5	0.0530	0.0360
3diii	150	0.4	1	0.4460	0.2020
3div	150	0.4	0.5	0.6460	0.6350

## 10. Discussion & conclusion

In this work we have discussed point patterns observed on arbitrary, bounded convex shapes in  $\mathbb{R}^3$ , motivated by the need for such exploratory analyses in the area of microbiology. We have highlighted the challenge of handling such spaces due to the lack of isometries for such objects. Using the invariance of Poisson processes (Kingman, 1993), we can circumvent this lack of isometries in the original space by mapping to the sphere which has rotational symmetries. By doing so we propose a set of functional summary statistics for the class of Poisson processes. Further to this we have also proposed functional summary statistics for CSR processes on the convex space

and explored their properties. Using this we have, in turn, been able to construct test statistics which can be used to reject the hypothesis of CSR for observed point patterns. We have also conducted simulation studies to investigate the effectiveness of the proposed test statistics in rejecting the null hypothesis when the observed data is either regular or clustered.

Interesting extensions to this work would include relaxing the need for convexity of the shape of interest. This presents a significant challenge as how one constructs the required mapping is not obvious. Another consideration is how to construct an estimator of the intensity function on  $\mathbb{D}$ . One approach might be to construct it on  $\mathbb{S}^2$  and inverse map to  $\mathbb{D}$ . There is, of course, the open question of how one forms summary statistics for multivariate point processes on convex shapes. Answering this would have immediate impact in bioimaging applications where experimentalists are regularly interested in spatial dependencies that exist between two or more different types of molecules.

## Acknowledgements

We gratefully acknowledge the input made by the Associate Editor and two anonymous referees; their constructive comments led to a significantly improved article. Scott Ward is supported by the Wellcome Trust under grant 203799/Z/16/Z.

## Appendix A. Supplementary data

Supplementary material related to this article can be found online at <https://doi.org/10.1016/j.spasta.2020.100489>.

## References

- Baddeley, A.J., Møller, J., Waagepetersen, R.P., 2000. Non- and semi-parametric estimation of interaction in inhomogeneous point patterns. *Stat. Neerl.* 54 (3), 329–350.
- Baddeley, A.J., Rubak, E., Turner, R., 2015. *Spatial Point Patterns: Methodology and Applications with R*. CRC press.
- Baddeley, A.J., Silverman, B.W., 1984. A cautionary example on the use of second-order methods for analyzing point patterns. *Biometrics* 1089–1093.
- Besag, J.E., 1977. Discussion on Dr Ripley's Paper. *J. R. Statist. Soc. B* 39, 193–195.
- Cabriel, C., Bourg, N., Jouchet, P., Dupuis, G., Letierrier, C., Baron, A., Badet-Denisot, M.-A., Vauzeilles, B., Fort, E., Lévêque-Fort, S., 2019. Combining 3D single molecule localization strategies for reproducible bioimaging. *Nature Commun.* 10 (1), 1–10.
- Chiu, S.N., Stoyan, D., Kendall, W.S., Mecke, J., 2013. *Stochastic Geometry and Its Applications*. John Wiley & Sons.
- Diggle, P.J., 2003. *Statistical Analysis of Spatial Point Patterns*, second ed. Arnold, London.
- Gustavsson, A.-K., Petrov, P.N., Lee, M.Y., Shechtman, Y., Moerner, W.E., 2018. 3D Single-molecule super-resolution microscopy with a tilted light sheet. *Nature Commun.* 9 (1), 1–8.
- Kingman, J.F.C., 1993. *Poisson Processes*. Clarendon Press, Oxford.
- Lagache, T., Lang, G., Sauvonnnet, N., Olivo-Marin, J.C., 2013. Analysis of the spatial organization of molecules with robust statistics. *PLoS One* 8 (12), e80914.
- Lang, G., Marcon, E., 2013. Testing randomness of spatial point patterns with the ripley statistic. *ESAIM Probab. Stat.* 17, 767–788.
- Lawrence, T.J., 2018. *Point Pattern Analysis on a Sphere* (Master's thesis). The University of Western Australia.
- Lawrence, T.J., Baddeley, A.J., Milne, R.K., Nair, G., 2016. Point pattern analysis on a region of a sphere. *Stat* 5 (1), 144–157.
- van Lieshout, M.N., 2006. A J-function for marked point patterns. *Ann. Inst. Statist. Math.* 52 (2), 235–259.
- van Lieshout, M.N., 2011. A J-function for inhomogeneous point processes. *Stat. Neerl.* 65 (2), 183–201.
- Møller, J., Rubak, E., 2016. Functional summary statistics for point processes on the sphere with an application to determinantal point processes. *Spat. Statist.* 18, 4–23.
- Møller, J., Waagepetersen, R.P., 2004. *Statistical Inference and Simulation for Spatial Point Processes*. Monographs on statistics and applied probability 100, CRC Press, Boca Raton, Florida.
- Ripley, B.D., 1977. Modelling spatial patterns. *J. R. Stat. Soc. Ser. B Stat. Methodol.* 39 (2), 172–192.
- Robeson, S.M., Li, A., Huang, C., 2014. Point-pattern analysis on the sphere. *Spat. Statist.* 10, 76–86.
- Said, S., Bombrun, L., Berthoumieu, Y., Manton, J.H., 2017. Riemannian Gaussian distributions on the space of symmetric positive definite matrices. *IEEE Trans. Inform. Theory* 63 (4), 2153–2170.
- White, S.D., 1979. The hierarchy of correlation functions and its relation to other measures of galaxy clustering. *Mon. Not. R. Astron. Soc.* 186 (2), 145–154.