Analysis of Spatial Point Patterns on Surfaces

a thesis presented for the degree of

Doctor of Philosophy of Imperial College London

and the

Diploma of Imperial College

by

Scott Alexander Ward

Department of Mathematics

Imperial College

180 Queen's Gate, London SW7 2BZ

July 2021

I certify that this thesis, and the research to which it refers, are the product of my own work, and that any ideas or quotations from the work of other people, published or otherwise, are fully acknowledged in accordance with the standard referencing practices of the discipline.

Signed: _____

Copyright

PhD supervisors: Edward A.K. Cohen & Niall Adams Scott Alexander Ward

Analysis of Spatial Point Patterns on Surfaces

Abstract

With the advent of improved data acquisition technologies more complex spatial datasets can be collected at scale meaning theoretical and methodological developments in spatial statistics are imperative in order to analyse and generate meaningful conclusions. Spatial statistics has seen a plethora of applications in life sciences with particular emphasis on ecology, epidemiology and cell microscopy. Applications of these techniques provides researchers with insight on how the locations of objects of interest can be influenced by their *neighbours* and the *environment*. Examples include understanding the spatial distribution of trees observed within some window, and understanding how neighbouring trees and potentially soil contents can influence this.

Whilst the literature for spatial statistics is rich the common assumption is that point processes are usually restricted to some d-dimensional Euclidean space, for example cell locations in a rectangular window of 2-dimensional Euclidean space. As such current theory is not capable of handling patterns which lie on more complex spaces, for example cubes and ellipsoids. Recent efforts have successfully extended methodology from Euclidean space to spheres by using the chordal distance (the shortest distance between any two points on a sphere) in place of the Euclidean distance. In this thesis we build on this work by considering point processes lying on more complex surfaces. Our first significant contribution discusses the construction of *functional summary statistics* for Poisson processes which lie on compact subsets of $\mathbb{R}^d$ which are off lower dimension. We map the process from its original space to the sphere where it is possible to take advantage of rotational symmetries which allow for well-defined summary statistics. These in turn can be used to determine whether an observed point patterns exhibits clustered or regular behaviour.

Partnering this work we also provide a hypothesis testing procedure based on these functional summary statistics to determine whether an observed point pattern is *complete spatially random*. Two test statistics are proposed, one based on the commonly used L-function for planar processes and the other a standardisation of the K-function. These test statistics are compared in an extensive simulation study across ellipsoids of varying dimensions and

processes which display differing levels of aggregation or regularity.

Estimates of first order properties of a point process are extremely important. They can provide a graphical illustration of inhomogeneity and are useful in second order analysis. We demonstrate how kernel estimation can be extended from a Euclidean space to a Riemannian manifold where the Euclidean metric is now substituted for a Riemannian one. Many of the desirable properties for Euclidean kernel estimates carry over to the Riemannian setting. The issue of edge correction is also discussed and two criteria for bandwidth selection are proposed. These two selection criteria are explored through a simulation study.

Finally, an important area of research in spatial statistics is exploring the interaction between different processes, for example how different species of plant spatially interact within some window. Under the framework of marked point processes we show that functional summary statistics for multivariate point patterns can be constructed on the sphere. This is extended to more general convex shapes through an appropriate mapping from the original shape to the sphere. A number of examples highlight that these summary statistics can capture independence, aggregation and repulsion between components of a multivariate process on both the sphere and more general surfaces.

Acknowledgments

First and most importantly I would like to thank my supervisors Edward A.K. Cohen and Niall Adams for their unwavering support and guidance throughout my PhD; it is truly inspiring to work with academics who are incredibly passionate about their research. I am privileged to be supervised by individuals who value my input and have kept me in such high regard.

I am also grateful to the Wellcome Trust Foundation who provided me with the scholarship to be able to complete my studies*. My PhD has been an amazing journey from start to finish and without their financial support none of this would have been possible.

I would like to give thanks Heather Battey for her contributions and interesting discussions surrounding Chapter 6. Thanks also goes to Francisco Cuevas-Pacheco for providing me with the code to simulate Gaussian random fields over spheres used in Chapter 3 which is based on their work Cuevas et al. [2020]. I would like to thank Sandip Kumar and Colin Cleanthous of Oxford University for their interesting discussions which helped focus my work.

During my studies I was fortunate enough to complete an internship at Roche Pharmaceuticals where I broadened my horizons and gained an understanding of how statistics is utilised within the pharmaceutical industry. I would like to thank Robson Machado and Ardo van den Hout for their supervision and to my managers James Black and Ryan Copping for their valuable input.

A special thank you goes to all the friends I have made at Imperial College London. Throughout my time at Imperial I have had many discussions all of which have aided and influenced my work but have also pushed me through more trying times during my studies.

Finally, I would like to thank my family whose love and care throughout these years carried me to the conclusion of my PhD. A special thank you goes to my wonderful fiancée Heidi, whose emotional support and love have been invaluable during trying times throughout my PhD and for which I am eternally grateful.

*This research was funded by the Wellcome Trust [Grant number 203799/Z/16/Z]

List of figures

List of tables

List of Acronyms & Notation

We provide a list of acronyms and common mathematical notation used throughout the remainder of this thesis.

Acronyms

CSR	Completely spatially random
GRF	Gaussian random field
IRWMI	Iteratively reweighted moment isotropic
IRWMS	Iteratively reweighted moment stationary
LGCP	Log Gaussian Cox process
PCF	Pair correlation function
SOIRWI	Second order intensity reweighted isotropic
SOIRWS	Second order intensity reweighted stationary
SORMM	Second order reduced moment measure

Point processes

$\alpha^{(n)}$	n^{th}-order moment factorial intensity measure of a point process
$\mathbb{E}$	Expectation operator of a point process
$\mathbb{E}^{!}_{\mathbf{x}}$	Expectation operator of the reduced Palm process of a point process at some arbitrary point $\mathbf{x}$
$\mathbf{x}$	A singleton of some metric space S
μ	Intensity measure of a point process
$\mu^{(n)}$	n^{th}-order moment intensity measure of a point process
ρ	Intensity function of a point process
$\rho^{(n)}$	n^{th}-order factorial moment intensity function of a point process
ξ_n	n^{th}-order correlation functions of a point process
C	Campbell measure of a point process
$C^{!}$	Reduced Campbell measure of a point process
h	Pair correlation function of a point process
N_X	Random counting measure denoting the cardinality of a point process X in some subset of the metric space it is defined on
n_x	Counting measure denoting the cardinality of a deterministic set x where $x \subseteq S$
P	Probability measure of a point process

$P^!_{\mathbf{x}}$	Probability measure of the reduced Palm process of a point process at some arbitrary point $\mathbf{x}$
X	Point process on some metric space S
x	A subset of S, typically of finite cardinality
$X^!_{\mathbf{x}}$	Reduced Palm Point process of point process X given and arbitrary point $\mathbf{x}$

Sets, Measures & Metrics

$\exp_{\mathbf{p}}$	Exponential map at $\mathbf{p}$ on a Riemannian manifold
λ_S	The Lebesgue measure over some metric space S
$\log_{\mathbf{p}}$	Logarithmic map at $\mathbf{p}$ on a Riemannian manifold
$\mathbb{D}$	Surface of a convex set in $\mathbb{R}^3$
$\mathbb{N}$	The set of natural numbers excluding 0
$\mathbb{N}_0$	The set of natural numbers including 0
$\mathbb{R}$	The real numbers
$\mathbb{R}^d$	The d - dimensional Euclidean plane
$\mathbb{R}_+$	The non-negative real numbers
$\mathcal{B}$	Borel sigma-algebra over S generated by open sets of S
$\mathcal{B}_0$	Bounded elements of $\mathcal{B}$
$\mathcal{M}$	d-dimensional Riemannian manifold
$\mathcal{N}_f$	Sigma-algebra generated by subsets of N_f such that the subsets have finite cardinality for any bounded Borel set of S
$\mathcal{N}_{lf}$	Sigma-algebra generated by subsets of N_{lf} such that the subsets have finite cardinality for any bounded Borel set of S
$A_k(V)$	Alternating k-tensors of a vector space V
$B_S(\mathbf{x}, r)$	Ball of radius r defined by d_S at point $\mathbf{x} \in S$ for a metric space S
C^k	Class of manifolds with k continously differentiable transition maps
$C^\infty_{\mathbf{p}}$	Algebra of germs of smooth functions defined at $\mathbf{p}$ of a differential manifold
$d\text{vol}$	Riemannian volume form
d_S	The metric on a metric space S
g	Level set (or a Riemannian metric in Chapter 6)
N_f	Set of finite subsets of S (when S is bounded)
N_{lf}	Set of locally finite subsets of S (when S is unbounded)
S	A general metric space
S^k	Permutation group of size k
$T\mathcal{M}$	Tangent bundle of a differential manifold
$T_{\mathbf{p}}\mathcal{M}$	Tangent space at $\mathbf{p}$ on a differential manifold
$T^*_{\mathbf{p}}\mathcal{M}$	Dual space to tangent space at $\mathbf{p}$ on a differential manifold

Contents

1

Introduction

Spatial statistics is a collection of techniques used to understand the spatial characteristics of a finite collection of points within a given area. Technology used to acquire such data is becoming ever more advanced with datasets being collected at finer and finer resolutions. Spatial datasets come in many forms, for example Diggle [2003] discusses how to handle data in the form of grid counts. More modern imaging techniques allow researchers to localise objects precisely giving exact coordinates, or at least within some small standard error. Not only is it possible to capture more accurate spatial data but technological advancements have meant that data can be captured at scales ranging from the microscopic [Diggle, 1986, Lagache et al., 2013, Cohen et al., 2019] to the macroscopic [Robeson et al., 2014, Lawrence et al., 2016, Jun et al., 2019]. This advancement in the collection of spatial datasets has fuelled the need for novel methodological development to correctly analyse and draw conclusions from the data.

With spatial datasets being curated at a multitude of scales these point patterns may not necessarily lie on a plane or within some volume, but instead be restricted to some lower dimensional surface embedded within a higher dimensional Euclidean space. This means that current methodologies, which often make the assumption that points lie on a Euclidean space, are not appropriate to analyse such data. Point patterns such as these are an under-researched area with, only recently, extensions being made for point patterns restricted to the sphere [Robeson et al., 2014, Møller and Rubak, 2016, Lawrence et al., 2016, Cuevas-Pacheco and Møller, 2018, Jun et al., 2019] or on linear networks [Ang et al., 2012, Rakshit

et al., 2017, Moradi et al., 2018, Rakshit et al., 2019]. The primary focus of this thesis is to develop theory and methodology for point processes that exist on surfaces embedded within some Euclidean space that do not exhibit symmetric properties like spheres.

Challenges are readily apparent when attempting to extend methodology for a point process lying on a Euclidean or spherical space to a space which exhibits more complex geometry. An immediate issue is the definition of a *distributionally invariant* process. By defining point processes to exist on a d - dimensional Euclidean space or even a $(d-1)$ - dimensional sphere it is possible to define notions of distributional invariance under common isometries, such as translation (on Euclidean space) and rotation (on Euclidean and spherical space). When relaxing the assumption that point processes lie on such symmetric spaces we can no longer assume that there exists an infinitely sized set of isometries for which distributional invariance (i.e. an infinitely large set of transformations T defined on some general space such that for any $t \in T$, a point process X has the same probability measure as $t(X)$) can be well-defined which has subsequent consequences when attempting to define basic properties commonly attributed to point processes.

This thesis makes three contributions to the literature for studying point processes outside of typical symmetric spaces such as the plane and sphere. The first shows that functional summary statistics can be constructed for Poisson processes on convex shapes. We also discuss a formal hypothesis testing procedure to determine whether a pattern exhibits homogeneity. The second discusses how to estimate the intensity of a point process lying on a Riemannian manifold by extending the methodology of kernel estimation for Euclidean Diggle [1985], Berman and Diggle [1989], Møller and Waagepetersen [2003], van Lieshout [2012] and linear network McSwiggan et al. [2017], Moradi et al. [2018], Rakshit et al. [2019] processes. Finally, we demonstrate that functional summary statistics for multivariate processes lying on spheres can be constructed. Using this framework we can also construct summary statistics for multivariate processes on convex shapes. We demonstrate that these summary statistics are capable of determining whether or not components of the multivariate process are independent.

1.1 Structure of thesis

This thesis presents novel methodology for analysing point patterns confined to arbitrary closed convex surfaces within d-dimensional Euclidean, with a focus on the scenario when $d = 3$. The outline of this thesis is as follows:

- Chapter 2 discusses the relevant background on spatial statistics required. Here we establish the notation used for the remainder of the thesis. In this chapter we describe much of the theory under the premise of a general metric space but provide more concrete examples in d-dimensional Euclidean space.

- Chapter 3 extends from Chapter 2 specifically for point processes defined on a $(d-1)$ -dimensional sphere, with a focus on $d = 3$. Point processes on spheres are central to the work coming in later chapters. In this chapter we outline the significant contributions made for the analysis of spherical processes [Robeson et al., 2014, Møller and Rubak, 2016, Lawrence et al., 2016] and also contribute an extension of the *inhomogeneous F, H, and J-functions* [van Lieshout, 2011] from Euclidean to spherical space, utilising rotations as opposed to translations. We also formalise and extend to spheres the proof of White [1979] which derives the infinite series expansion of the F, and H-functions, first when the n^{th}-order moment factorial intensity measure is assumed to exist and when the n^{th}-order factorial moment intensity function is assumed to exist. This is covered in our work, Ward et al. [2021b].

- Chapter 4 describes the first major novel contribution to the spatial statistics literature and is the key concept introduced in our paper, Ward et al. [2021b]. Here we discuss the difficulties imposed by the geometry of surfaces for which our point process exists upon. Based on the Mapping Theorem for Poisson processes [Kingman, 1993] we can circumvent this issue and are able to construct functional summary statistics for such a process. This provides graphical illustrations of how an observed point pattern may deviate away from spatial homogeneity and can guide researchers in modelling their data.

- Chapter 5 builds on the work of Chapter 4. It discusses how to conduct a formal, Monte Carlo based, hypothesis test when the null of the observed point pattern is a homogeneous Poisson process, which is typically conducted early in a point pattern analysis pipeline; this is the final contribution made in our work Ward et al. [2021b]. We propose two test statistics, the first is analagous to the L-function typically used for point process on $\mathbb{R}^d$ and based on the work of Lawrence [2018] whilst the second is based on the standardisation given by Lagache et al. [2013]. The empirical power of these test statistics is explored through simulations on ellipsoids of varying dimensions.

- Chapter 6 discusses the problem of intensity estimation for point processes on compact Riemannian manifolds, which is the focus of our work Ward et al. [2021a]. We focus on kernel estimators and extend theory of nonparametric kernel intensity estimation from

$\mathbb{R}^d$ to arbitrary Riemannian manifolds with a focus on ones embedded within $\mathbb{R}^3$. We discuss theoretical properties of these estimators highlighting the Euclidean analogue and also provide a practical approach to bandwidth selection of point processes on such spaces. We conduct a simulation study over ellipsoids of varying dimensions to compare two bandwidth selection procedures, extended from the Euclidean theory, and based on these are able to provide broad rules of thumbs as to the most appropriate scenarios for each of their applications.

- Chapter 7 develops methodology to detect interactions between components of a multivariate point process lying on convex surfaces. This chapter is related to our work Ward et al. [2021c]. Considering multitype processes as a special case of marked processess, we first discuss marked spheroidal point processes extending the works of van Lieshout [2006], Cronie and van Lieshout [2016]. Utilising the Mapping Theorem [Last and Penrose, 2018], we can map point patterns from a convex shape to the unit sphere which allows for the construction of functional summary statistics. These can be used to illustrate, graphically, whether a multivariate point patterns exhibits attraction or repulsion between its components.

1.2 Publications

Work contained in this thesis has either been published or is in preparation. Chapters 3 - 5 (and Appendices A - C) discuss material in our paper titled 'Testing for complete spatial randomness on three dimensional bounded convex shapes' [Ward et al., 2021b]. Permission and limitations for reuse of this material can be found in the *Author Rights* section at `https://www.elsevier.com/about/policies/copyright`. The website states that this work can be included as part of a thesis or dissertation and a copy of this can be found in Appendix F. Chapters 6 and 7 discuss material for papers currently under preparation and are titled 'Estimation of the intensity function of a spatial point process on a Riemannian manifold' [Ward et al., 2021a] and 'Functional summary statistics for multitype point patterns on three dimensional convex surfaces' [Ward et al., 2021c] respectively.

2

Background

This chapter discusses the necessary spatial statistics background used throughout this thesis. We follow the notation of Møller and Waagepetersen [2003]. Section 2.1 provides a formal definition of a spatial point process, introducing their moment measures and reduced Palm processes. We pay particular emphasis on Poisson processes in Section 2.2 which are a key spatial process used in later chapters. Finally, Section 2.3 discusses functional summary statistics in $\mathbb{R}^d$ typically used in exploratory data analysis, model fitting, and model validation, and provide examples when the process is stationary.

2.1 Point processes

Spatial point processes, or simply point processes, are often considered as a random collection of points located over some given space. For example a random finite set of points over a d-dimensional Euclidean space, $\mathbb{R}^d$, or $(d-1)$-dimensional sphere, $\mathbb{S}^{d-1}$, where $d \in \mathbb{N}$. More commonly we only observe point process within some bounded region of the underlying space, such as the d-dimensional box in $\mathbb{R}^d$ or the upper hemisphere of $\mathbb{S}^{d-1}$. In this chapter we will assume that X lies within some, potentially unbounded, metric space S equipped with a metric d_S (in later chapters this will often be the geodesic or shortest distance between points of the associated metric space).

2.1.1 Definition of a point process

Let X be a point process on a metric space S which is Polish, this is assumed throughout the thesis. Then associated with X is its random counting measure $N_X : S \mapsto \mathbb{N}_0$, where $\mathbb{N}_0$ is the set of natural numbers including 0, which counts the number of elements of X that lie within some subset of S. Analogously, we have a deterministic version of N_X denoted $n_x : S \mapsto \mathbb{N}_0$, where $x \subseteq S$ is a deterministic subset of S. We will also use X_B or x_B to denote $X \cap B$ or $x \cap B$ where $B \subseteq S$ respectively. To denote elements of X we will use the term *event* whilst an arbitrary point in S will be termed a *point*. Define $B_S(\mathbf{x}, r)$ for $\mathbf{x} \in S$ and $r \in \mathbb{R}_+$ to be the ball in S centred at $\mathbf{x}$ with radius r such that $\mathbf{y} \in B_S(\mathbf{x}, r)$ if and only if $d_S(\mathbf{x}, \mathbf{y}) \leq r$. We also define λ_S to be the Hausdorff measure on S, for example we take the Lebesgue measure on $\mathbb{R}^d$, and will define $X \setminus \mathbf{x} \equiv X \setminus \{\mathbf{x}\}$ for ease of notation.

For a formal definition of a point process we define the set of locally finite point configurations of S as,

$$N_{lf} = \{x \subset S : n_x(B) < \infty, B \subseteq S \text{ such that } B \text{ is bounded}\}.$$

In the event that S is bounded we can define the finite point configurations of S as,

$$N_f = \{x \subset S : n_x(B) < \infty, B \subseteq S\}.$$

A point process X then takes values in the space N_{lf} or N_f depending on whether S is unbounded or not. Based on the space of values X can take we may also refer to X as a random locally finite set (or random finite set in the event S is bounded). Equipping S with the Borel sigma-algebra $\mathcal{B}$, the sigma-algebra generated by open subsets of S, and defining $\mathcal{B}_0$ to denote the set of bounded Borel sets we then define $\mathcal{N}_{lf}$ as,

$$\mathcal{N}_{lf} = \sigma(\{x \in N_{lf} : n(x_B) = m\} : B \in \mathcal{B}_0, m \in \mathbb{N}_0),$$

i.e. the sigma-algebra generated by sets constructed from N_{lf} such that they are finite over elements of $\mathcal{B}_0$. We then define a point process as,

Definition 2.1.1. *A point process X which lies on a metric space S is said to be a measurable mapping from some probability space $(\Omega, \mathcal{F}, P)$ to the measurable space $(N_{lf}, \mathcal{N}_{lf})$. We denote $P_X(F) = P(X \in F) = P(\{\omega \in \Omega : X(\omega) \in F\})$ for $F \in \mathcal{N}_{lf}$, to be the probability measure of X.*

In the event that S is bounded a similar definition can be constructed using N_f. A common, mild restriction imposed on point proceses is that they are *simple*. A simple point process is

one such that, almost surely, no two distinct elements of the point process are coincidental, that is the points cannot lie on top of each other. All point processes considered in this thesis will be simple.

Based on Definition 2.1.1 we can characterise a point process in numerous ways. Lemma B.2 of Møller and Waagepetersen [2003] shows that a point process is uniquely determined by the finite dimensional distribution of its random count measure, that is joint distribution of $N_X(B_1), \ldots, N_X(B_m)$ for any $B_1, \ldots, B_m \in \mathcal{B}_0$ and $m \in \mathbb{N}_0$ defines our point process X. This leads to a useful characterisation based on *void probabilities*.

Definition 2.1.2. *Let X be a point process on some metric space S. Then the void probabilities of the point process are,*

$$\nu(B) = P(N_X(B) = 0), \quad \textit{for any } B \in \mathcal{B}_0.$$

Theorem B.1 of Møller and Waagepetersen [2003] shows that X is uniquely determined by its void probabilities. Another useful characterisation is based on so-called *generating functionals*. Generating functionals are useful in Chapters 3 and 4 when extending functional summary statistics defined for stationary processes to nonstationary ones.

Definition 2.1.3. *Let X be a point process on some metric space S. Then the generating functionals of X are defined as,*

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

for functions $u : S \mapsto [0, 1]$ with the set defined by $\{x \in S : u(x) < 1\}$ being bounded.

To see why generating functionals uniquely determine point process, set $u(x) = a^{\mathbb{1}[x \in B]}$ for $0 \leq a \leq 1$, then $G_X(u) = \mathbb{E}\left[t^{N_X(B)}\right]$. This is the moment generating function for $N_X(B)$ which uniquely defines the random counting measure of X. This in turn defines the void probabilities, which by previous discussion, i.e. Theorem B.1 of Møller and Waagepetersen [2003], defines X.

If we restrict our attention to the setting of $S = \mathbb{R}^d$ then common assumptions frequently imposed on a point process X are *stationarity* and *isotropy*.

Definition 2.1.4. *Let X be a point process on $\mathbb{R}^d$. X is said to be stationary if its associated probability measure is invariant under translations, i.e. $X \stackrel{d}{=} X + \mathbf{x}$ for any $\mathbf{x} \in \mathbb{R}^d$, $X + \mathbf{x} = \{\mathbf{z} : \mathbf{z} = \mathbf{y} + \mathbf{x}, \mathbf{y} \in X\}$ and $\stackrel{d}{=}$ means equivalent in distribution. X is also said to be isotropic*

if it is distributionally invariant under rotations around the origin, i.e. $X \stackrel{d}{=} OX$*, where* O *is any rotation around the origin and* $OX = \{\mathbf{y} : \mathbf{y} = O\,\mathbf{x}, \mathbf{x} \in X\}$.

In a few instances it will be convenient to discuss densities of absolutely continuous measures. More precisely we say that a process X is absolutely continuous with respect to another process Y if and only if $P_Y(Y \in F) = 0$ implies $P_X(X \in F) = 0$ for $F \in \mathcal{N}_{lf}$. By the Radon-Nikodyn Theorem [Billingsley, 1986] we then have that if X is absolutely continuous with respect to Y there exists a function $f : N_{lf} \mapsto [0, \infty]$ such that,

$$P_X(X \in F) = \mathbb{E}[\mathbb{1}[Y \in F] f(Y)].$$

We then denote f as the density of X with respect to Y.

2.1.2 Moment measures

In this section we introduce the definitions of moment measures in spatial statistics which are analagous to moments of random variables. We define the n^{th}*-order moment measure* $\alpha^{(n)}$ and the n^{th}*-order factorial moment measure* $\alpha^{(n)}$.

Definition 2.1.5. *The* n^{th}*-order moment measure,* $\mu^{(n)}$*, of a point process* X *is defined as,*

$$\mu^{(n)}(B_1, \ldots, B_n) = \mathbb{E} \sum_{\mathbf{x}_1, \ldots, \mathbf{x}_n \in X} \mathbb{1}[x_1 \in B_1, \ldots, x_n \in B_n],$$

for $B_i \subseteq S, i = 1, \ldots, n$*. Further, when* $n = 1$ *and if we assume that* $\mu^{(1)}$ *is absolutely continuous with respect to the Hausdorff measure over* S *then by the Radon-Nikodyn Theorem [Billingsley, 1986] we have that*

$$\mu(B) \equiv \mu^{(1)}(B) = \int_B \rho(\mathbf{x}) \lambda_S(\mathbf{x}),$$

for $B \subseteq S$*.* $\rho : S \mapsto \mathbb{R}_+$ *where* $\mathbb{R}_+$ *is the non-negative real numbers is denoted the intensity function of* X*, whilst* μ *is denoted the intensity measure.*

Based on Definition 2.1.5 $\rho(\mathbf{x})\lambda_S(d\,\mathbf{x})$ is often considered the probability of observing an event of X lying in some infinitesimal volume $d\,\mathbf{x}$ centred at $\mathbf{x}$. Additionally, notice that when $n = 1$,

$$\mu(B) = \mathbb{E} \sum_{\mathbf{x} \in X} \mathbb{1}[\mathbf{x} \in B] = \mathbb{E}\left[N_X(B)\right],$$

and for $n \in \mathbb{N}$,

$$\mu^{(n)}(B_1, \ldots, B_n) = \mathbb{E}\left[N_X(B_1) \cdots N_X(B_n)\right].$$

An important property of stationary processes is their constant intensity functions when restricting S to $\mathbb{R}^d$ (e.g. see Proposition 8.2 Last and Penrose [2018]). We also say that if a point process has a constant intensity function then it is *homogeneous* otherwise it is *inhomogeneous*. The next definition introduces the n^{th}*-order factorial moment measure*.

Definition 2.1.6. *The n^{th}-order factorial moment measure, $\alpha^{(n)}$, of a point process X is defined as,*

$$\alpha^{(n)}(B_1, \ldots, B_n) = \mathbb{E} \sum_{\mathbf{x}_1, \ldots, \mathbf{x}_n \in X}^{\neq} \mathbb{1}[x_1 \in B_1, \ldots, x_n \in B_n],$$

for $B_i \subseteq S, i = 1, \ldots, n$ and the sum is over pairwise distinct elements, i.e. $\mathbf{x}_i \neq \mathbf{x}_j$ for $i \neq j$. Further, if $\alpha^{(n)}$ is absolutely continuous with respect to the n-fold Hausdorff measure on S^n then by the Radon-Nikodyn Theorem [Billingsley, 1986] we have that,

$$\alpha^{(n)}(B_1, \ldots, B_n) = \int_{B_1} \cdots \int_{B_n} \rho^{(n)}(x_1, \ldots, x_n) \lambda_S(d\,\mathbf{x}_1) \cdots \lambda_S(d\,\mathbf{x}_n), \tag{2.1}$$

where λ_S is the Hausdorff measure on S. The integrand $\rho^{(n)}$ is defined as the n^{th}-order factorial intensity function.

Like the intensity function $\rho^{(n)}(\mathbf{x}_1, \ldots, \mathbf{x}_n)\lambda_S(d\,\mathbf{x}_1) \cdots \lambda_S(d\,\mathbf{x}_n)$ can be considered the probability of observing n events of X each within an infinitesimal area $d\,\mathbf{x}_i$ centred at $\mathbf{x}_i$. Notice that for $n = 1$, $\alpha \equiv \alpha^{(1)} = \mu$. The term *factorial* in n^{th}-order factorial moment measure arises when we consider $B_i = B \subseteq S$ for all $i = 1, \ldots, n$ then,

$$\begin{aligned}
\alpha^{(n)}(B_1, \ldots, B_n) &= \alpha^{(n)}(B, \ldots, B) \\
&= \mathbb{E} \sum_{\mathbf{x}_1, \ldots, \mathbf{x}_n \in X}^{\neq} \prod_{i=1}^{n} \mathbb{1}[\mathbf{x}_i \in B] \\
&= \mathbb{E}\left[N_X(B)(N_X(B) - 1) \cdots (N_X(B) - n + 1)\right].
\end{aligned}$$

Furthermore, it is worthwhile noting the following relation when $n = 2$ (e.g. see Equation 4.1 of Møller and Waagepetersen [2003]),

$$\mu^{(2)}(B_1, B_2) = \alpha^{(2)}(B_1, B_2) + \mu(B_1 \cap B_2),$$

which highlights that the second order properties of N_X are defined by μ and $\alpha^{(2)}$.

The n^{th}-order moment and factorial moment measures can equivalently be defined in terms of Lebesgue integrals of arbitrary non-negative measurable functions (e.g. see Chiu et al. [1995, pp. 110-111]). Equivalence of these definitions holds by standard measure theoretic arguments such as those given by Theorem 16.11 and the subsequent remarks of Billingsley [1986].

2.1.3 Palm Theory

Palm theory plays a central role in point process theory and is crucial in understanding the distribution of a point process X given that we have observed one or more *typical* points of X. These distributions are more commonly referred to as *Palm distributions*. They were first introduced by Palm [1943] for point processes on the real line and then later extended to $\mathbb{R}^d$ and other more abstract spaces (e.g. see Jagers [1973]).

Palm theory is the study a point process X conditional on some *typical* event of that same process. Measure theory is used to give a precise definition of what a typical event of a point process is but heuristically we can think of them as a randomly selected event from X in which each event has an equal probability of being chosen [Chiu et al., 1995]. A typical application of this theory is in the construction of nearest neighbour distance distributions where we are interested in understanding what the closest neighbour is to a typical point of our process is.

In this section we shall follow the tutorial to Palm theory outlined by Coeurjolly et al. [2017] for the general case when X is not absolutely continuous with respect to a unit rate Poisson process (see Section 2.2 for a discussion on Poisson processes). In the event X is absolutely continuous with respect to a unit rate Poisson process then Coeurjolly et al. [2017] also provide a simplified account of Palm distributions. For a more detailed exposition of Palm distributions see Daley and Vere-Jones [2003].

In order to define the Palm process of a point process we require the following definition.

Definition 2.1.7. *Let X be a point process on some metric space S. We then define the Campbell measure, $C : S \times N_{lf} \mapsto \mathbb{R}$, and reduced Campbell measure, $C^! : S \times N_{lf} \mapsto \mathbb{R}$ as,*

$$C(B, F) = \mathbb{E} \sum_{\mathbf{x} \in X} \mathbb{1}[\mathbf{x} \in B, X \in F]$$
$$C^!(B, F) = \mathbb{E} \sum_{\mathbf{x} \in X}^{\neq} \mathbb{1}[\mathbf{x} \in B, X \setminus \mathbf{x} \in F],$$

for $B \subseteq S$ and $F \subseteq N_{lf}$ respectively.

Taking the reduced Campbell measure we notice that $\mu(\cdot) = C^!(\cdot, N_{lf})$ and it is easy to see that $C^!(\cdot, F) \leq C^!(\cdot, N_{lf}) = \mu(\cdot)$ for any $F \in N_{lf}$. Therefore $C^!$ is absolutely continuous with respect to μ and so by the Radon-Nikodyn Theorem [Billingsley, 1986] we have that there exists a density $P^!_{\mathbf{x}}$ such that,

$$C^!(B, F) = \int_B P^!_{\mathbf{x}}(F)\mu(d\,\mathbf{x}), \tag{2.2}$$

where it can be shown that $P^!_{\mathbf{x}}$ is a probability measure for each $\mathbf{x} \in S$ (see Daley and Vere-Jones [2003] for details). We then have the following definition for the reduced Palm distribution,

Definition 2.1.8. *Let X be a point process on a metric space S. Then $P^!_{\mathbf{x}}$ for $\mathbf{x} \in S$ as defined by Equation 2.2 is the reduced Palm distribution of X at point $\mathbf{x}$. We then say that the process defined by probability measure $P^!_{\mathbf{x}}$ is the reduced Palm process of X and denoted $X^!_{\mathbf{x}}$.*

Coeurjolly et al. [2017] extend this definition for multiple typical events. That is given a set of typical events $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ of X then we can equivalently define $P^!_{\mathbf{x}_1,\ldots,\mathbf{x}_n}$ and $X^!_{\mathbf{x}_1,\ldots,\mathbf{x}_n}$ to be the reduced Palm distribution and process given the set of typical points respectively. These definitions are constructed based on the n^{th}-order factorial moment measure and n^{th}-order reduced Campbell measure $C^!_n$,

$$C^!_n(B_1, \ldots, B_n, F) = \mathbb{E} \sum_{\mathbf{x}_1,\ldots,\mathbf{x}_n \in X}^{\neq} \mathbb{1}[\mathbf{x}_1 \in B_1, \ldots, \mathbf{x}_n \in B_n, X \setminus \{\mathbf{x}_1, \ldots, \mathbf{x}_n\} \in F],$$

analogously to the case of an individual typical event of X.

The reduced Palm distribution is often interpreted as the conditional distribution of $X \setminus \mathbf{x}$ given $\mathbf{x}$. To see this in Definition 2.1.8 set $B = B_S(\mathbf{x}, r)$ for any $\mathbf{x} \in S$ and $r \in \mathbb{R}_+$ small, so small that in fact the probability of observing two points of the process X inside this small ball is negligible. Then we have that $\mu(B) \approx P(N_X(B) > 0)$ and $C^!(B, F) \approx P(N_X(B) > 0, X \setminus \mathbf{x} \in F)$ and so by Equation 2.2 we have,

$$\begin{aligned} P^!_{\mathbf{x}}(F) &\approx \frac{C^!(B_S(\mathbf{x}, r), F)}{\mu(B_S(\mathbf{x}, r))} \\ &\approx \frac{P(N_X(B_S(\mathbf{x}, r)) > 0, X \setminus \mathbf{x} \in F)}{P(N_X(B_S(\mathbf{x}, r)) > 0)} \end{aligned}$$

$$= P(X \setminus \mathbf{x} \in F | N_X(B_S(\mathbf{x}, r)) > 0).$$

This leads to the common interpretation of the reduced Palm distribution being a conditional distribution.

By utilising standard measure theoretic arguments (e.g. see Theorem 16.11 of Billingsley [1986]) and Equation 2.2 we can obtain the *Campbell-Mecke Theorem* (e.g. see Equation C.4 of Møller and Waagepetersen [2003]).

Theorem 2.1.9. *(Campbell-Mecke Theorem) Let X be a point process on a metric space S with $X^!_{\mathbf{x}}$ being the reduced Palm process of X given a typical event $\mathbf{x}$. Then it follows that,*

$$\mathbb{E} \sum_{\mathbf{x} \in X} h(\mathbf{x}, X \setminus \mathbf{x}) = \int_S \mathbb{E}\left[h(\mathbf{x}, X^!_{\mathbf{x}})\right] \mu(d\,\mathbf{x})$$

for non-negative functions $h : S \times N_{lf} \mapsto \mathbb{R}$.

A simple corollary to this is the Campbell Theorem (e.g see Equation 4.1.3 of Chiu et al. [1995]).

Corollary 2.1.10. *(Campbell Theorem) Let X be a point process on a metric space S then,*

$$\mathbb{E} \sum_{\mathbf{x} \in X} h(\mathbf{x}) = \int_S h(\mathbf{x}) \mu(d\,\mathbf{x})$$

for non-negative functions $h : S \mapsto \mathbb{R}$.

This result follows trivially from the Campbell-Mecke Theorem, by defining $g(\mathbf{x}) = h(\mathbf{x}) \mathbb{1}[X \setminus \mathbf{x} \in N_{lf}]$ where h is as in Corollary 2.1.10. Then using Theorem 2.1.9 with g the Campbell theorem is obtained. Furthermore, these results can be extended from the reduced Palm process $X^!_{\mathbf{x}}$ to the reduced Palm process $X^!_{\mathbf{x}_1,\ldots,\mathbf{x}_n}$ thus giving formulas,

$$\begin{aligned}
&\mathbb{E} \sum_{\mathbf{x}_1,\ldots,\mathbf{x}_n \in X}^{\neq} h(\mathbf{x}_1, \ldots, \mathbf{x}_n, X \setminus \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}) \\
&\qquad = \int_{S^n} \mathbb{E}\left[h(\mathbf{x}_1, \ldots, \mathbf{x}_n, X^!_{\mathbf{x}_1,\ldots,\mathbf{x}_n})\right] \alpha^{(n)}(d\,\mathbf{x}_1, \ldots, d\,\mathbf{x}_n) \qquad (2.3) \\
&\mathbb{E} \sum_{\mathbf{x}_1,\ldots,\mathbf{x}_n \in X}^{\neq} h(\mathbf{x}_1, \ldots, \mathbf{x}_n) = \int_{S^n} h(\mathbf{x}_1, \ldots, \mathbf{x}_n) \alpha^{(n)}(d\,\mathbf{x}_1, \ldots, d\,\mathbf{x}_n)
\end{aligned}$$

which are the *extended* versions of the Campbell-Mecke and Campbell Theorems respectively (e.g. see Equations 12 and 14 of Coeurjolly et al. [2017]).

Both the Campbell-Mecke and Campbell Theorem are useful devices when establishing results involving the expectations of sums over point processes. They convert sums to integrals which can be more easily evaluated, especially if we suppose that $\alpha^{(n)}$ is absolutely continuous with respect to n-dimensional Hausdorff measure of S and hence $\alpha^{(n)}(d\,\mathbf{x}_1, \ldots, d\,\mathbf{x}_n)$ can be replaced with $\rho^{(n)}(\mathbf{x}_1, \ldots, \mathbf{x}_n)\lambda_S(d\,\mathbf{x}_1)\cdots\lambda_S(d\,\mathbf{x}_n)$ in all previous equations.

Suppose now that $\rho^{(n)}$ exists for $n \in \mathbb{N}$ and that X is absolutely continuous with respect to Z, the standard unit rate Poisson (see Section 2.2 for definition of Poisson processes), with density $f(x)$. Then by setting $h(\mathbf{x}_1, \ldots, \mathbf{x}_n, X \setminus \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}) = \mathbb{1}[\mathbf{x}_1 \in B_1, \ldots, \mathbf{x}_n \in B_n, X \setminus \{\mathbf{x}_1, \ldots, \mathbf{x}_n\} \in F]$ for $B_i \in S, i = 1, \ldots, n$ and $F \in \mathcal{N}_{lf}$ we have,

$$
\begin{aligned}
&C_n^!(B_1, \ldots, B_n, F) \\
&= \mathbb{E}\left[f(Z) \sum_{\mathbf{z}_1,\ldots,\mathbf{z}_n \in Z}^{\neq} h(\mathbf{z}_1, \ldots, \mathbf{z}_n, Z \setminus \{\mathbf{z}_1, \ldots, \mathbf{z}_n\}) \right] \\
&= \mathbb{E}\left[\sum_{\mathbf{z}_1,\ldots,\mathbf{z}_n \in Z}^{\neq} f(\mathbf{z}_1, \ldots, \mathbf{z}_n, Z \setminus \{\mathbf{z}_1, \ldots, \mathbf{z}_n\}) h(\mathbf{z}_1, \ldots, \mathbf{z}_n, Z \setminus \{\mathbf{z}_1, \ldots, \mathbf{z}_n\}) \right] \\
&= \int_{B_1} \cdots \int_{B_n} \mathbb{E}\left[f(\mathbf{z}_1, \ldots, \mathbf{z}_n, Z \setminus \{\mathbf{z}_1, \ldots, \mathbf{z}_n\}) \mathbb{1}[Z \in F] \right] \lambda_S(d\,\mathbf{z}_1) \cdots \lambda_S(d\,\mathbf{z}_n) \\
&= \int_{B_1} \cdots \int_{B_n} \mathbb{E}\left[\frac{f(\mathbf{z}_1, \ldots, \mathbf{z}_n, Z \setminus \{\mathbf{z}_1, \ldots, \mathbf{z}_n\})}{\rho^{(n)}(\mathbf{z}_1, \ldots, \mathbf{z}_1)} \mathbb{1}[Z \in F] \right] \\
&\qquad\qquad \rho^{(n)}(\mathbf{z}_1, \ldots, \mathbf{z}_1) \lambda_S(d\,\mathbf{z}_1) \cdots \lambda_S(d\,\mathbf{z}_n),
\end{aligned}
$$

and therefore $X^!_{\mathbf{x}_1,\ldots,\mathbf{x}_n}$ is the reduced Palm process which has density

$$
f_{\mathbf{x}_1,\ldots,\mathbf{x}_n}(x) = \frac{f(\mathbf{x}_1, \ldots, \mathbf{x}_n, x \setminus \{\mathbf{x}_1, \ldots, \mathbf{x}_n\})}{\rho^{(n)}(\mathbf{x}_1, \ldots, \mathbf{x}_1)} \tag{2.4}
$$

with respect to the standard unit rate Poisson, this follows by arguments similar to those in Coeurjolly et al. [2017].

An important result when $S = \mathbb{R}^d$ and our process is stationary is an explicit relationship between $P^!_{\mathbf{x}}$ and $P^!_{\mathbf{o}}$ where $\mathbf{o}$ is the origin of $\mathbb{R}^d$. More precisely, if we define

$$
P^!_{\mathbf{o}}(F) = \frac{1}{\rho\lambda_S(B)} \mathbb{E} \sum_{\mathbf{x} \in X_B} \mathbb{1}[X \setminus \in F + \mathbf{x}], \quad F \subseteq N_{lf},
$$

for some arbitrary set $B \subseteq \mathbb{R}^d$ such that $0 < \lambda_S(B) < \infty$, and

$$P^!_{\mathbf{x}}(F) = P^!_{\mathbf{o}}(F - \mathbf{x}), \quad F \subseteq N_{lf}, \mathbf{x} \in \mathbb{R}^d \tag{2.5}$$

then $P^!_{\mathbf{x}}(F)$ defines the reduced Palm distribution of the stationary process X for any $\mathbf{x} \in \mathbb{R}^d$; this is the result of, for example, Theorem C.1 of Møller and Waagepetersen [2003].

2.2 Poisson processes

Poisson processes are a class of point processes used in the preliminary stages of inference as models for non-interacting events or *complete spatial randomess* (CSR). They provide the foundation to build more complex point processes which may more readily describe data. For the purposes of this section we will assume $0 < \lambda_S(S) < \infty$ and details of extensions to unbounded S can be found in Møller and Waagepetersen [2003].

2.2.1 Definition of a Poisson process

Poisson processes can be defined based on Binomial point processes which describe the distribution of precisely $n \in \mathbb{N}$ independently and identically distributed events over some metric space S.

Definition 2.2.1. *(Binomial point process) Let $n \in \mathbb{N}$ and $B \subseteq S$. Let $f : B \mapsto \mathbb{R}_+$ be a density function such that $\int_B f(\mathbf{x})\lambda_S(d\,\mathbf{x}) = 1$. Then X consisting of n points is said to be a Binomial point process with n points and density f over B if the n points are independently and identically distributed (IID) over B with density f with respect to λ_S.*

The simplest Binomial point process assumes uniform density across B, i.e. $f(\mathbf{x}) = 1/\lambda_S(B)$. Although at first glance the Binomial process looks like a model of no interaction, it is not, in fact if we say that there are $m \leq n$ points in some subset $A \subset B$ then there must be $n-m$ points in $B \setminus A$. In order to build a model of no interaction between points we randomise the total number of points leading to a Poisson process.

Definition 2.2.2. *(Poisson point process) Let $\rho : S \mapsto \mathbb{R}_+$ be integrable over B (i.e. $\int_B \rho(\mathbf{x})\lambda_S(d\,\mathbf{x}) < \infty$) and define,*

$$\mu(B) = \int_B \rho(\mathbf{x})\lambda_S(d\,\mathbf{x}), \quad B \subseteq S.$$

The point process X is said to be Poisson over S with intensity function ρ if:

1. *for any* $B \subseteq S$ *the number of points* $N_X(B)$ *of* X *in* B *is Poisson with mean* $\mu(B)$,

2. *and* $N_X(B) = n$ *these points are distributed as a Binomial process over* B *with density* $f(\mathbf{x}) = \rho(\mathbf{x})/\mu(B)$.

Alternative definitions of a Poisson process can be characterised by the fact that $X = X^!_{\mathbf{x}}$ for all $\mathbf{x} \in S$ [Coeurjolly et al., 2017] or even by independence of the process over disjoint sets [Møller and Waagepetersen, 2003, Proposition 3.2].

If the intensity function of a Poisson process is constant then we say the process is a homogeneous Poisson process (otherwise it is an inhomogeneous Poisson process), in particular we use the homogeneous Poisson process to define a process as being CSR: a homogeneous model of no interactions. In preliminary analysis of spatial data it is common to determine whether or not the point pattern exhibits CSR or not and in the event this hypothesis is rejected then whether it demonstrates attractive or repulsive behaviour. We define the standard unit rate Poisson process to be the Poisson process over S with $\rho(\mathbf{x}) = 1$ for all $\mathbf{x} \in S$ and denote it by Z. Suppose that S is bounded and X is a Poisson process with intensity function $\rho : S \mapsto \mathbb{R}_+$. Then by Proposition 3.8 of Møller and Waagepetersen [2003] we have that X is absolutely continuous with respect to Z with density,

$$f(x) = \exp\left(\lambda_S(S) - \mu(S)\right) \prod_{\mathbf{x} \in x} \rho(\mathbf{x}), \tag{2.6}$$

where $\mu(S) = \int_S \rho(\mathbf{x}) \lambda_S(d\,\mathbf{x})$, f is defined for finite point configurations of S.

2.2.2 Expansion of the Poisson measure

An important property of Poisson processes that will be useful for establishing results involving the mapping of Poisson processes between different metric space is the expansion of the Poisson measure.

Proposition 2.2.3. *Let* X *be a point process over a metric space* S *with intensity function* $\rho : \mathbb{S} \mapsto \mathbb{R}_+$ *that is assumed integrable. Then* X *is Poisson if and only if for all* $B \subseteq S$ *and all* $F \subseteq \mathcal{N}_{lf}$,

$$P(X_B \in F) = \sum_{n=0}^{\infty} \frac{\exp(-\mu(B))}{n!} \int_B \cdots \int_B \mathbb{1}[\{\mathbf{x}_1, \ldots, \mathbf{x}_n\} \in F] \prod_{i=1}^{n} \rho(\mathbf{x}_i) \lambda_S(d\,\mathbf{x}_1) \cdots \lambda_S(d\,\mathbf{x}_n)$$

See, for example, Proposition 3.1 of Møller and Waagepetersen [2003] for a proof; it also

follows directly from Definition 2.2.2.

If we now suppose that we have a point process X that is absolutely continuous with respect to Z, the standard Poisson process, then there must exists a unique non-negative density f such that $P(X \in F) = \mathbb{E}[f(Z)\mathbb{1}[Z \in F]]$, and therefore by Proposition 2.2.3 we have that,

$$P(X \in F) = \sum_{n=0}^{\infty} \frac{\exp(-\mu(S))}{n!} \int_S \cdots \int_S \mathbb{1}[\{\mathbf{x}_1, \ldots, \mathbf{x}_n\} \in F] f(\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}) \lambda_S(d\,\mathbf{x}_1) \cdots \lambda_S(d\,\mathbf{x}_n),$$

for $F \in \mathcal{N}_f$ [Coeurjolly et al., 2017, Equation 3].

2.2.3 Slivnyak-Mecke Theorem

The Slivnyak-Mecke Theorem is another useful way of characterising a Poisson process and can be used to show that if a point process is Poisson then so is its reduced Palm process.

Theorem 2.2.4. *(Slivnyak-Mecke Theorem) Let X be a Poisson process over S with integrable intensity function $\rho : S \mapsto \mathbb{R}_+$. Then for any function $h : S^n \times N_{lf} \mapsto \mathbb{R}_+$,*

$$\mathbb{E} \sum_{\mathbf{x}_1, \ldots, \mathbf{x}_n \in X}^{\neq} h(\mathbf{x}_1, \ldots, \mathbf{x}_n, X \setminus \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}) = \int_{S^n} \mathbb{E}\left[h(\mathbf{x}_1, \ldots, \mathbf{x}_n, X)\right] \prod_{i=1}^{n} \rho(\mathbf{x}_i) \lambda_S(d\,\mathbf{x}_1) \cdots \lambda_S(d\,\mathbf{x}_n)$$

See Theorem 3.3 of Møller and Waagepetersen [2003] for a proof. Immediately from the Slivnyak-Mecke Theorem we have that $\rho^{(n)} = \rho^n$ if our point process is Poisson. This follows by the uniqueness of densities $d\alpha^{(n)}/d\lambda_S^n$.

2.3 Functional summary statistics

Functional summary statistics are popular statistical devices used in the early stages of spatial data analysis. They can be used to fit models through techniques akin to the method of moments estimators commonly employed in traditional statistical analysis; this is known as the method of minimum contrast. They can also be used to validate model fits. During the initial analysis of any spatial data an important application of functional summary statistics

is to determine whether an observed point pattern exhibits CSR. In this section we shall restrict our attention to the case where $S = \mathbb{R}^d$.

2.3.1 Pair correlation function

The *pair correlation function* (PCF), h, is the normalisation of the second order intensity function by the product of the first order intensity functions. More precisely, if both ρ and $\rho^{(2)}$ exists for a process X, then its PCF is,

$$h(\mathbf{x}, \mathbf{y}) = \frac{\rho^{(2)}(\mathbf{x}, \mathbf{y})}{\rho(\mathbf{x})\rho(\mathbf{y})}, \tag{2.7}$$

where we take $a/0 = 0$ for any $a \geq 0$. Immediately notice that, by the Slivnyak-Mecke Theorem we have that $h(\mathbf{x}, \mathbf{y}) = 1$ for all Poisson processes. Then if $h(\mathbf{x}, \mathbf{y}) > 1$ (< 1) this suggests that pairs of points at these locations are more likely (less likely) to occur jointly when compared to a Poisson process with the same intensity function.

A common assumption to make on h is translationally invariant, i.e. $h(\mathbf{x}, \mathbf{y}) = h(\mathbf{x} + \mathbf{z}, \mathbf{y} + \mathbf{z})$. Under this assumption h is then a function of the difference between its arguments i.e. $h(\mathbf{x}, \mathbf{y}) = h(\mathbf{x} - \mathbf{y})$. If in addition to translation invariance we also impose rotational invariance around the origin then we have that h is a function of the distance between the points i.e. $h(\mathbf{x}, \mathbf{y}) = h(d_{\mathbb{R}^d}(\mathbf{x}, \mathbf{y}))$ [Møller and Waagepetersen, 2003].

2.3.2 Stationary processes

We now suppose that our point process X is stationary and therefore has a constant intensity function $\rho > 0$. We define Ripley's K function [Ripley, 1977] and the L-function [Møller and Waagepetersen, 2003].

Definition 2.3.1. *Ripley's K function and the L-function for a stationary process X on $\mathbb{R}^d$ are,*

$$K(r) = \frac{1}{\rho^2 \lambda_S(A)} \mathbb{E} \sum_{\mathbf{x}_1, \mathbf{x}_2 \in X}^{\neq} \mathbb{1}\left[\mathbf{x}_1 \in A, \mathbf{x}_2 - \mathbf{x}_1 \in B_{\mathbb{R}^d}(\mathbf{o}, r)\right],$$

$$L(r) = \left(\frac{K(r)}{\omega_d}\right)^{\frac{1}{d}}, \quad r > 0,$$

where ω_d is the volume of the unit ball in $\mathbb{R}^d$.

By applying the Campbell-Mecke Theorem to the definition of $K(r)$ and using the fact that $P^!_{\mathbf{x}}(F) = P^!_{\mathbf{o}}(F - \mathbf{x})$ for any $F \in \mathcal{N}_{lf}$ (see Equation 2.5) we obtain that $K(r) = \mathbb{E}\left[N_{X^!_{\mathbf{o}}}\left(B_{\mathbb{R}^d}(\mathbf{o}, r)\right)\right]/\rho$ [Møller and Waagepetersen, 2003, Appendix C]. Thus we have a Palm interpretation of $\rho K(r)$ as the expected number of further events of X within a distance r of a given typical event at $\mathbf{o}$. Furthermore, under a Poisson assumption and when $d = 2$ it can be shown that,

$$K(r) = \pi r^2, \quad L(r) = r, r > 0.$$

Thus combined with the Palm interpretation for a point process X whose K-function is such that $K(r) > \pi r^2$ ($< \pi r^2$) or whose L-function is such that $L(r) > r$ ($< r$) then this suggest *aggregation* (*regularity*) at scales less than r when compared to CSR: a model of non-interacting events. When testing for CSR the L-function is more often employed since (a) it is linear in r and (b) under a homogeneous Poisson process the estimators of the L-function are variance stabilised [Besag, 1977], see Section 2.3.4 for nonparametric estimation. Figure 2.1 highlights these properties for estimators of both the K and L-functions.

We next introduce the *empty space function* F, *nearest neighbour function* H, and the J-function [Møller and Waagepetersen, 2003].

Definition 2.3.2. *Let X be a stationary point process over $\mathbb{R}^d$ with constant intensity function $\rho > 0$. Then the empty space function is the probability of the nearest event from a given point, typically the origin, is to the nearest point of X i.e.*

$$F(r) = P\left(X_{B_{\mathbb{R}^d}(\mathbf{o},r)} \neq \emptyset\right), \quad r > 0. \tag{2.8}$$

The nearest neighbour function is,

$$H(r) = \frac{1}{\rho \lambda_{\mathbb{R}^d}(A)} \mathbb{E} \sum_{\mathbf{x} \in X_A} \mathbb{1}[(X \setminus \mathbf{x}) \cap B_{\mathbb{R}^d}(\mathbf{x}, r) \neq \emptyset], \quad r > 0 \tag{2.9}$$

for an arbitrary set $A \subset \mathbb{R}^d$ such that $0 < \lambda_{\mathbb{R}^d}(A) < \infty$. The J-function is then given as,

$$J(r) = \frac{1 - H(r)}{1 - F(r)}, \quad F(r) < 1.$$

By stationarity of X, F is independent of the origin and H is independent of the arbitrary set A. This means that it is possible to form unbiased estimates of these functions can be formulated from an individual observation x of X [Møller and Waagepetersen, 2003].

A useful identity for $F(r)$ can be derived as,

$$F(r) = P\left(X_{B_{\mathbb{R}^d}(\mathbf{o},r)} \neq \emptyset\right) = \frac{\mathbb{E}\left[\lambda_{\mathbb{R}^d}(B \cap X_{\oplus r})\right]}{\lambda_{\mathbb{R}^d}(B)}, \tag{2.10}$$

where $X_{\oplus r} = \{\mathbf{y} : \exists\, \mathbf{x} \in X \text{ s.t. } \mathbf{y} \in B_S(\mathbf{x}, r)\}$ and $r > 0$. This follows since F does not depend on $\mathbf{o}$ by stationarity of X and an interchange of integrals by Fubini's Theorem.

The naming of H-function as the nearest neighbour function becomes apparent when interpreted as a Palm distribution. By using the Campbell-Mecke Theorem and noting that $P^!_{\mathbf{x}}(F) = P^!_{\mathbf{o}}(F - \mathbf{x})$ for any $F \in \mathcal{N}_{lf}$ (see Equation 2.5), in Appendix C of Møller and Waagepetersen [2003] they show that

$$H(r) = P\left(N_{X^!_{\mathbf{o}}}\left(B_{\mathbb{R}^d}(\mathbf{o}, r)\right)\right), \quad r > 0,$$

which leads to the nearest neighbour interpretation.

It can further be shown that under CSR and in $\mathbb{R}^d$, F and H are equivalent and equal to $1 - \exp(-\rho\omega_d r^d)$ for $r > 0$ and ω_d the volume of the unit ball in $\mathbb{R}^d$ [Møller and Waagepetersen, 2003] and so $J(r) = 1$ for a CSR process. This means that, at least for small r, if the process were aggregated then we would expect $H(r) > F(r)$ since we would expect to have greater probability observing two events within a distance r then from a point $\mathbf{x} \in \mathbb{R}^d$ to an event of X. The reverse reasoning holds for more regular processes and thus we have that $J(r) < 1$ if the process exhibits aggregation or clustering and $J(r) > 1$ if the process exhibits regularity. These properties are explored and demonstrated in Figure 2.1.

2.3.3 Nonstationary processes

Before introducing the summary statistics for nonstationary point processes we first introduce the $\mathcal{K}$ measure and the notion of *second order intensity reweighted stationary* (SOIRWS). This allows for the extension of Ripley's K-function to inhomogeneous point processes and is thanks to the work of Baddeley et al. [2000].

Definition 2.3.3. *Let X be a point process over $\mathbb{R}^d$ with locally integrable intensity function $\rho : \mathbb{R}^d \mapsto \mathbb{R}_+$. Suppose that the measure*

$$\mathcal{K}(B) = \frac{1}{\lambda_{\mathbb{R}^d}} \mathbb{E} \sum_{\mathbf{x}_1, \mathbf{x}_2 \in X}^{\neq} \frac{\mathbb{1}[\mathbf{x}_1 \in A, \mathbf{x}_2 - \mathbf{x}_1 \in B]}{\rho(\mathbf{x}_1)\rho(\mathbf{x}_2)}, \quad B \subseteq \mathbb{R}^d \tag{2.11}$$

does not depend on $A \subseteq \mathbb{R}^d$ *with* $0 < \lambda_{\mathbb{R}^d}(A) < \infty$ *and take* $a/0 = 0$ *for any* $a \geq 0$. *Then* X *is said to be second order intensity reweighted stationary and* $\mathcal{K}$ *is referred to as the second order reduced moment measure (SORMM) on* $\mathbb{R}^d$.

Immediately we have that if a process X is stationary then it is also SOIRWS [Møller and Waagepetersen, 2003]. $\mathcal{K}$ can equivalently be expressed using Palm distributions, more precisely in Appendix C of Møller and Waagepetersen [2003] they show that,

$$\mathcal{K}(B) = \int_{N_{lf}} \sum_{\mathbf{x} \in x} \frac{\mathbb{1}[\mathbf{x} - \mathbf{y} \in B]}{\rho(\mathbf{x})} P^!_{\mathbf{y}}(dx), \tag{2.12}$$

for almost all $\mathbf{y} \in \mathbb{R}^d$. This follows by application for the Campbell-Mecke Theorem to Equation 2.11.

Furthermore, we have a useful identity which relates $\mathcal{K}$ to the PCF, h, when it exists and is translational invariant, i.e. $h(\mathbf{x}, \mathbf{y}) = h(\mathbf{x} - \mathbf{y})$, then

$$\mathcal{K}(B) = \int_B h(\mathbf{x}) \lambda_{\mathbb{R}^d}(d\,\mathbf{x}), \tag{2.13}$$

which follows by the Campbell-Mecke Theorem and that the SORMM does not depend on A. Thus h being translationally invariant is sufficient for X to be SOIRWS [Baddeley et al., 2000].

We then define the inhomogeneous version of Ripley's K-function and the L-function as [Møller and Waagepetersen, 2003],

Definition 2.3.4. *Let* X *be a SOIRWS point process with intensity function* $\rho : \mathbb{R}^d \mapsto \mathbb{R}_+$. *Then the inhomogeneous* K *and* L*-functions are given as,*

$$\mathrm{K}_{inhom}(r) = \mathcal{K}\left(B_{\mathbb{R}^d}(\mathbf{o}, r)\right), \quad \mathrm{L}_{inhom}(r) = \left(\frac{\mathrm{K}_{inhom}(r)}{\omega_d}\right)^{\frac{1}{d}}, \quad r > 0$$

where ω_d *is the volume of the unit ball in* $\mathbb{R}^d$.

It is clear from this definition that when X is stationary we recover Ripley's K-function and therefore the L-function of Definition 2.3.1. Additionally, if we suppose that h is isotropic then,

$$\mathrm{K}_{\mathrm{inhom}}(r) = \sigma_d \int_0^r s^{d-1} h(s) ds,$$

where σ_d is the surface area of the unit ball in $\mathbb{R}^d$. This follows from Equation 2.13 and by

switching to polar coordinates in $\mathbb{R}^d$ [Møller and Waagepetersen, 2003].

The F, H, and J-functions have been extended to the inhomogeneous setting by van Lieshout [2011]. We will discuss these in detail in Chapter 3, specifically when $S = \mathbb{S}^d$.

2.3.4 Nonparametric estimation of summary statistics

We discuss estimation of the functional summary statistics detailed in previous sections. In particular we will be focusing on the $K, L, F, H, J, \mathrm{K}_{\mathrm{inhom}}$, and $\mathrm{L}_{\mathrm{inhom}}$-functions, whilst a discussion of how to undergo intensity function estimation is discussed in Chapter 6. Additional details on estimation of the $\mathcal{K}$ and PCFs are given by Møller and Waagepetersen [2003]. We will discuss so-called *border-corrected* [Møller and Waagepetersen, 2003] estimators here. Whilst simple and intuitive to understand border-corrected estimators can inevitably lead to the exclusion of a substantial number of points in an observed pattern when building estimates. Other approaches which attempt to counteract this effect include Ripley's isotropic correction [Ripley, 1991], Hanisch type estimators [Hanisch, 1984, Stoyan, 2006] and Kaplan-Meier type estimators [Baddeley and Gill, 1997]. These will not be discussed here and more details can be found in their respective references. We will suppose that we have observed x an instance of the process X within some region $W \subset \mathbb{R}^d$.

In spatial analysis certain properties of estimators are frequently unobtainable; ones that may even be considered a minimal requirement, such as unbiasedness, for traditional statistical analysis. Instead a common property that is often desirable is *ratio-unbiasedness*. Suppose that we have an estimator of the form $\hat{\theta} = A/B$ estimating a parameter θ, then $\hat{\theta}$ is said to be ratio-unbiased if $\mathbb{E}[A]/\mathbb{E}[B] = \theta$.

Given that we are observing only a portion of our process, precisely X_W, we introduce the fundamental morphological operators: *erosion* and *dilation*. These operators are defined as,

$$A \ominus B = \{a \in A : B_a \subseteq A\}, \quad A \oplus B = \cup_{a \in A} B_a,$$

where B_a is the translation of the set B by a i.e. $B_a = \{c : c = a + b, b \in B\}$ for erosion and dilation respectively. These operators lay the foundation of the theory in mathematical morphology and are the basis for more complex operators, such as *opening* and *closing*, for more details see Chiu et al. [1995]. Erosion is considerably important when considering random finite sets as it is commonly used to construct unbiased/ratio-unbiased estimators for functional summary statistics. In particular we denote $A_{\ominus r} = A \ominus B_S(\mathbf{o}, r)$ where $A \subseteq S$ as this set will be used repeatedly throughout this thesis.

Suppose we observe an instance of X_W as x_W and assume that X is stationary. Then a natural estimator of ρ is $n_x(W)/\lambda_{\mathbb{R}^d}(W)$, which is unbiased and in the event of X being Poisson is also the maximum likelihood estimator of ρ. Furthermore, we will often need estimators of ρ^2 (e.g. see Equation 2.11). In this case we can construct the estimator $\hat{\rho^2} = n_x(W)(n_x(W)-1)/\lambda^2_{\mathbb{R}^d}(W)$ for ρ^2, which by application of the Campbell-Mecke Theorem is unbiased.

Now suppose that momentarily X is no longer stationary and has intensity function $\rho : \mathbb{R}^d \mapsto \mathbb{R}_+$. Then the border-corrected estimator for $\mathrm{K}_{\mathrm{inhom}}$ is,

$$\hat{\mathrm{K}}_{\mathrm{inhom}}(r) = \frac{1}{\lambda_{\mathbb{R}^d}(W)} \sum_{\mathbf{x}_1 \in X_{W\ominus r}} \sum_{\mathbf{x}_2 \in X_W \setminus \mathbf{x}_1} \frac{\mathbb{1}[\mathbf{x}_2 \in B_{\mathbb{R}^d}(\mathbf{x}_1, r)]}{\rho(\mathbf{x}_1)\rho(\mathbf{x}_2)}, \tag{2.14}$$

which, by Equation 2.11, is unbiased. In the more likely event that ρ is unknown a plug-in intensity estimator is often used replacing ρ by $\hat{\rho}$ [Møller and Waagepetersen, 2003]. Specifically in the stationary case we can take $\hat{\rho^2} = n_x(W)(n_x(W)-1)/\lambda^2_{\mathbb{R}^d}(W)$ which, although generally biased, achieves ratio-unbiasedness. Furthermore, we can construct estimates of L as $\hat{L}(r) = (\hat{K}(r)/\omega_d)^{1/d}$. These are demonstrated in Figure 2.1.

Returning to the situation when X is stationary with known constant intensity function ρ then, based on Equation 2.9, we can construct the estimator

$$\hat{H}(r) = \frac{1}{\rho\lambda_{\mathbb{R}^d}(W_{\ominus r})} \sum_{\mathbf{x} \in X_{W\ominus r}} \mathbb{1}\left[X \setminus \mathbf{x} \cap B_{\mathbb{R}^d}(\mathbf{x}, r) = \emptyset\right]. \tag{2.15}$$

Clearly, for known ρ this is unbiased whilst using $\hat{\rho} = N_X(W)/\lambda_{\mathbb{R}}(W)$ in place of ρ incurs bias but attains ratio-unbiasedness. For the F-function we can construct two estimators based on Equations 2.8 and 2.10. Let P be a finite deterministic set of points over W then construct estimators,

$$\hat{F}_1(r) = 1 - \frac{\sum_{\mathbf{i} \in P_{W\ominus r}} \prod_{\mathbf{x} \in X} (1 - \mathbb{1}[d_{\mathbb{R}^d}(\mathbf{i}, \mathbf{x}) \leq r])}{n_P(W_{\ominus r})} \tag{2.16}$$

$$\hat{F}_2(r) = \frac{\lambda_{\mathbb{R}^d}(W_{\ominus r} \cap X_{\oplus r})}{\lambda_{\mathbb{R}^d}(W_{\ominus r})}, \tag{2.17}$$

where $\hat{F}_1$ is based on Equation 2.8 and $\hat{F}_2$ on Equation 2.10. Furthermore, it can be seen that both these estimators are unbiased. To construct an estimator of the J-function we can

use estimators of the F and H-functions as,

$$\hat{J}(r) = \frac{1 - \hat{H}(r)}{1 - \hat{F}_i(r)}, \quad i \in \{1, 2\}, \tag{2.18}$$

which by unbiasedness of $\hat{F}_i$ and ratio-unbiasedness of $\hat{H}$ means that $\hat{J}$ is ratio-unbiased. Construction of $\hat{J}$ is demonstrated in Figure 2.1 and how this compares when the assumption of CSR is violated by the observed point pattern.

2.3.5 Monte Carlo hypothesis testing

Since spatial point processes are complex objects it is often challenging if not impossible to construct analytic results and thus it is common to employ Monte Carlo based methods since simulation of point processes can often be achieved; this is especially true when conducting hypothesis testing. Let us suppose that we have a simple hypothesis where our null is completely characterised by P_{null}. We then consider some test statistic $T_0 = T(X)$, for example $T_0 = \hat{L}_X(r)$ for a given $r > 0$, where the subscript X indicates that $\hat{L}$ was constructed using a realisation of X. We can then construct a p-value of T_0 using Monte Carlo simulates under the null. More precisely we can simulate point patterns $X_1, \ldots, X_n$ for $n \in \mathbb{N}$ under the null and then construct $T_i = T(X_i)$ statistics for each i. These replicates can then be used to construct p-values for T_0 as

$$p_0 = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}[T_0 \leq T_i],$$

in the event of a one-sided test. The precision of these p-value estimates can be made arbitrarily more accurate by considering larger and larger values of n.

Since it is common to use functional summary statistics as test statistics it is useful to construct so-called *simulation envelopes* which can be used to determine behaviour at differing scales. Define $T_0(r) = T(X, r)$ to be the test statistics from the observed data at scale $r > 0$ and $T_i(r)$ to be the same test statistic constructed from null simulate X_i, $i = 1, \ldots, n$. We can then take,

$$T_{\min}(r) = \min\{T_i(r) : i = 1, \ldots, n\}, \quad T_{\max}(r) = \max\{T_i(r) : i = 1, \ldots, n\}.$$

Then since each X_i are IID, we have exchangeability. Further, assuming $T_i(r)$ for $i = 0, \ldots, n$

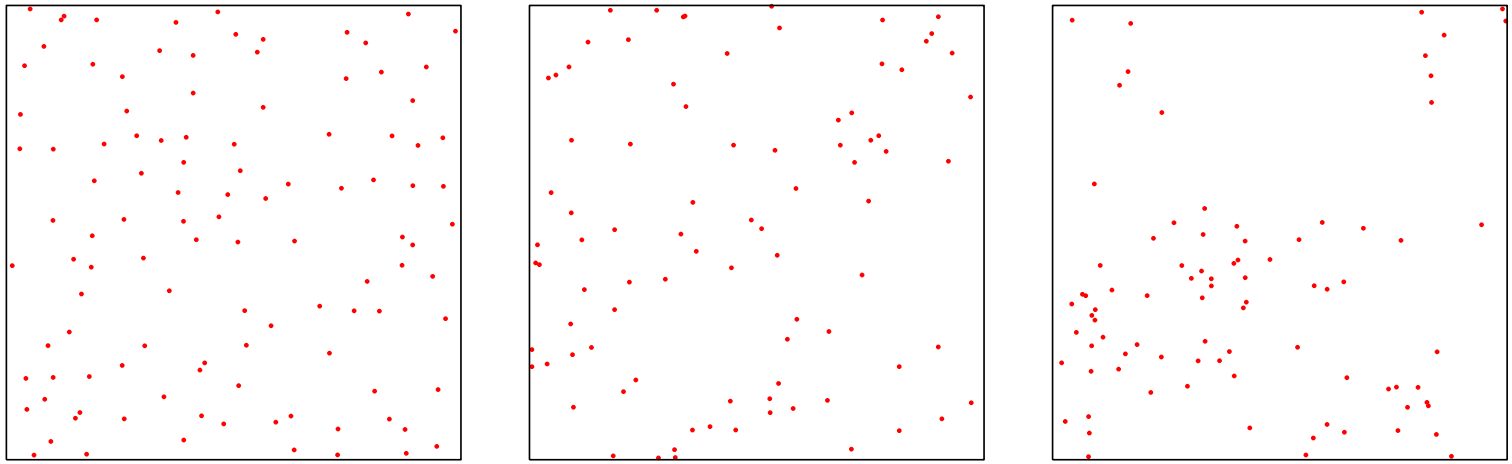
are almost surely different then,

$$P(T_0(r) < T_{\min}(r)) = P(T_0(r) > T_{\min}(r)) = \frac{1}{n+1}.$$

The bounds $T_{\min}(r)$ and $T_{\max}(r)$ are denoted the $100/(n+1)$ lower and $100n/(n+1)$ upper simulation envelopes at distance r. We need not only consider the most extreme values but also the k^{th} smallest and largest values of $T_i(r)$ and thus can construct the $100k/(n+1)$ lower and $100(n-k+1)/(n+1)$ upper simulation envelopes. We can then compare plots of $T_0(r)$ with its lower and upper simulation envelopes to determine whether there is statistical evidence to support rejecting the null. A cautionary word is necessary when comparing our observed test statistic $T_0(r)$ with $T_{\min}(r)$ and $T_{\max}(r)$; although for fixed r $T_i(r)$ are independent and hence exchangeable, when comparing across different values of r the random vectors $(T_1(r), \ldots, T_n(r))$ are no longer independent i.e. for $r \neq r'$ $(T_1(r), \ldots, T_n(r))$ and $(T_1(r'), \ldots, T_n(r'))$ are not independent. Plots of these functional summary statistics are demonstrated in Figure 2.1 under a CSR null hypothesis for various stationary processes.

2.4 Discussion

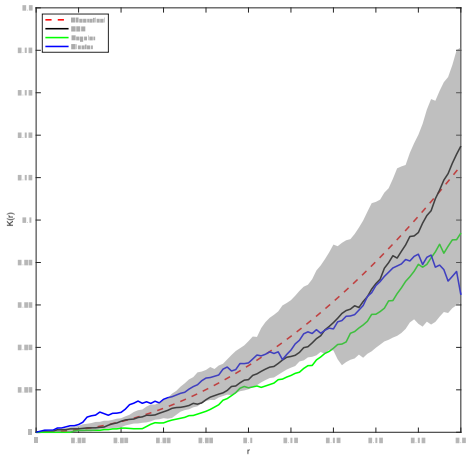
This chapter discusses the necessary spatial point pattern theory required for this thesis. We started by defining a point process on an arbitrary metric space and introduced many of their important properties such as moment measures, intensity functions, and Palm processes. We define a Poisson process as N IID points distributed over some window such that N is a Poisson random variable and discuss a handful of properties related to such processes. To end the chapter we introduce functional summary statistics for stationary and nonstationary processes, showing how they can be estimated and subsequently compared to a null model using simulation envelopes in order to determine whether the null should be rejected. Poisson processes will come to play a key role in Chapters 3 and 4 where they are assumed to be the null model, whilst functional summary statistics are used to determine appropriateness of such a hypothesis. In the next chapter we discuss point process theory specifically on spheres. We provide details of functional summary statistics currently in the literature whilst extending others from $\mathbb{R}^d$ to the sphere.



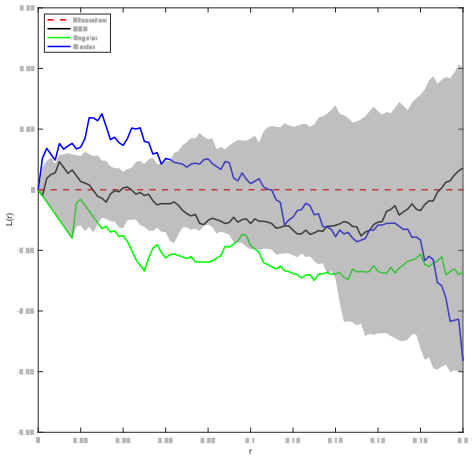
(a) Realisation of a regular process

(b) Realisation of a CSR process

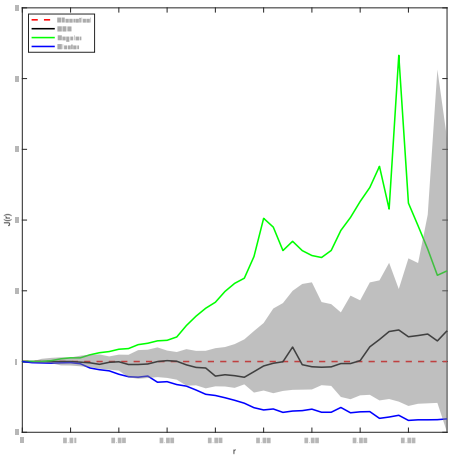
(c) Realisation of a cluster process



(d) Example K-function plot



(e) Example L-function plot



(f) Example J-function plot

Figure 2.1: Plots of functional summary statistics for typical planar point patterns. All processes shown are over the unit square and have an expected number of 100 over the unit square. **Top row:** realisation of a Matérn II process with hard-core distance 0.05 (*left*), realisation of a CSR process (*middle*), and realisation of a Thomas process with an expected number of 10 offspring per parent and each offspring distributed with a Gaussian kernel of bandwidth 0.01 (*right*). **Bottom row:** plots of functional summary statistics of Matérn II (*green*), CSR (*black*), and Thomas process (*blue*), with the theoretical K-function for a CSR process (*dashed red*) and simulation envelopes constructed from 39 realisations of a CSR process.

3

Point processes on $\mathbb{S}^{d-1}$

This chapter provides an overview of the current literature on spherical point processes, making particular reference to the recent works of Robeson et al. [2014], Møller and Rubak [2016] and Lawrence et al. [2016] which lay the foundation of spherical point process theory. We discuss these papers and others in Section 3.1 providing an introduction to the area and functional summary statistics for spherical point processes. We outline some basic notation needed for the geometry imposed by working on a spherical space in Section 3.2. Section 3.3 revisits the functional summary statistics for rotationally invariant point processes on $\mathbb{S}^{d-1}$ whilst Section 3.4 discusses the inhomogeneous extension of the K-function. In Sections 3.5 and 3.6 we build on this work by providing modest extensions from $\mathbb{R}^d$ to $\mathbb{S}^{d-1}$ of the inhomogeneous F, H, and J-functions in Section 3.5. We provide a similar infinite series expansion for the F and H-functions given by White [1979] for processes on $\mathbb{S}^{d-1}$ and also extend the inhomogeneous F, H and J-functions first discussed by van Lieshout [2011] for processes in $\mathbb{R}^d$. Properties of these functional summary statistics are discussed in Section 3.6 when we assume they are Poisson.

3.1 Literature review

Analysis of spatial point processes residing on $\mathbb{S}^{d-1}$ is a relatively new area of research. The first discussion was given in passing by Ripley [1977] with interest in the area only recently been reignited with the works of Robeson et al. [2014], Lawrence et al. [2016] and Møller and Rubak [2016].

Robeson et al. [2014] discusses the spherical analogue to Ripley's K-function [Ripley, 1977]. This is achieved by exchanging the Euclidean distance typically used to define K-functions

for stationary processes on $\mathbb{R}^d$ with the *great circle distance*. They show that a homogeneous Poisson process on a sphere of radius R in three dimensions has K-function given by $K(r) = 2\pi R^2(1 - \cos(r/R))$, where r is taken as the great circle distance. They also provide a nonparametric estimator for the spherical K-function of homogeneous processes which again follows the standard estimator on $\mathbb{R}^d$ replacing $\mathbb{1}[\mathbf{y} \in B_{\mathbb{R}^d}(\mathbf{x}, r)]$ with $\mathbb{1}[\mathbf{y} \in B_{\mathbb{S}^2_R}(\mathbf{x}, r)]$ in Equation 2.14 where $\mathbb{S}^2_R$ is the two dimensional sphere of radius R. It is also worthwhile noting that Robeson et al. [2014] highlight that, unlike the planar case where we can only observe the process through some window of finite area, in the spherical setting we need not handle edge-effects for a completely observed process. Robeson et al. [2014] demonstrate their spherical K-functions on three datasets: (1) a deterministic pattern of longitude and latitude points, (2) points of a hexagonal grid constructed to have the same area as the sphere, and (3) locations of Global Climate Observing System Upper-Air Network stations used to track local climate patterns. Although the first two are somewhat contrived, deterministic datasets they highlight similar properties of the K-function for planar data but in the spherical setting. That is more repulsive processes tend to display K-functions that lie below the theoretical CSR K-function whilst for attractive processes the opposite phenomena occurs.

Although an important first step is made by Robeson et al. [2014] the methodology presented is limited to homogeneous spherical point processes and only focuses on one functional summary statistics, namely Ripley's K-function which may only capture a subset of behaviour in an observed point pattern [Baddeley and Silverman, 1984]. This leads to the work of Lawrence et al. [2016] and Møller and Rubak [2016]. Møller and Rubak [2016] notes that these two works were conducted independently of one another although exploring very similar topics. That being said Møller and Rubak [2016] do highlight the differences between the works and that they supplement each other providing a much more holistic view of spherical point processes than if either were considered in isolation of the other. Starting with Lawrence et al. [2016] they provide three key extensions to that of Robeson et al. [2014]: (1) other functional summary statistics for rotationally invariant spherical processes, (2) functional summary statistics for inhomogeneous point processes, and (3) dealing with partially observed point patterns on $\mathbb{S}^{d-1}$. Lawrence et al. [2016] provide spherical counterparts for the F, H and J-functions for rotationally invariant spherical processes whilst focusing on the inhomogeneous K-function for processes with non-constant intensity functions. In Section 3.3 we formalise the extension of the inhomogeneous F, H, and J-functions to the setting of $\mathbb{S}^{d-1}$ from $\mathbb{R}^d$ [van Lieshout, 2011]. To deal with partial observation Lawrence et al. [2016] propose border-corrected estimates of the functional summary statistics whilst their later work, Lawrence [2018], provides details of alternative edge-correction approaches such as Ripley's isotropic correction [Ripley, 1991] and Hanisch-type estimators [Hanisch,

1984]. Lawrence et al. [2016] also discuss Neyman-Scott processes, more precisely giving a precise definition of a Thomas process on $\mathbb{S}^{d-1}$ based on the Von-Mises Fisher distribution and showing that model inference can be achieved for this class of processes through the method of minimum contrast. Møller and Rubak [2016] also discuss the inhomogeneous extension of the K-function for spherical point processes as well the spherical counterparts for the F, H, and J-functions. Møller and Rubak [2016] diverges from Lawrence et al. [2016] by discussing the theoretical details of spherical determinantal point processes and their usefulness in modelling repulsive processes.

These three papers can be considered to have laid the foundation for the theory of modern spherical point process theory. Further research has been conducted in this field. Cuevas-Pacheco and Møller [2018] discuss the theoretical considerations of log Gaussian Cox processes on $\mathbb{S}^{d-1}$ whilst Møller et al. [2018] build on Møller and Rubak [2016] discussion of determinantal point processes providing additional theoretical details. Møller et al. [2021] discusses point processes on $\mathbb{R}^d \times \mathbb{S}^k$ extending the inhomogeneous K-function from $\mathbb{R}^d$. Jun et al. [2019] discuss multivariate extensions on the sphere, more precisely modelling three different rain-types simultaneously under the framework of a trivariate log Gaussian Cox process. An overview of recent developments in spatial statistics including those on $\mathbb{S}^{d-1}$ can be found in Møller and Waagepetersen [2017].

For the remainder of this chapter we will discuss spherical point patterns when $d = 3$ but easily extends to any $d \in \mathbb{N}$. Methodology also extends easily when considering spheres not of unit radius.

3.2 Notation on $\mathbb{S}^2$

We consider the metric space $(\mathbb{S}^2, d_{\mathbb{S}^2})$ where $d_{\mathbb{S}^2} : \mathbb{S}^2 \times \mathbb{S}^2 \mapsto \mathbb{R}_+$ is defined to be the shortest distance between two points in $\mathbb{S}^2$, this is also known as the great circle distance and commonly denoted the *geodesic distance* within the framework of manifolds, whilst the *geodesic* is a continuous path between the two points which attains the geodesic distance. On the unit sphere the geodesic distance has a closed form,

$$d_{\mathbb{S}^2}(\mathbf{x}, \mathbf{y}) = \cos^{-1}(\mathbf{x}^T \mathbf{y}), \quad \mathbf{x}, \mathbf{y} \in \mathbb{S}^2.$$

Further, the area of the unit sphere is $\lambda_{\mathbb{S}^2}(\mathbb{S}^2) = 4\pi$ whilst the area of a spherical cap of angle $r \in [0, \pi]$ is given by $\lambda_{\mathbb{S}^2}(B_{\mathbb{S}^2}(\mathbf{x}, r)) = 2\pi(1 - \cos(r))$ for any $\mathbf{x} \in \mathbb{S}^2$, see Figure 3.1. We will take the origin on the unit sphere to be $\mathbf{o} = (0, 0, 1)^T$. Let $\mathcal{O}(3)$ be the set of all

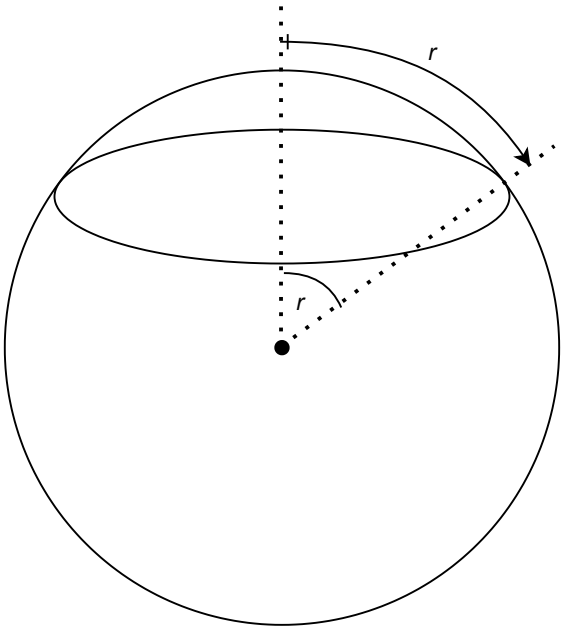


Figure 3.1: Diagram of spherical cap on unit sphere with radius r.

three dimensional rotations around the origin of $\mathbb{R}^3$ then we define $O_{\mathbf{x}} \in \mathcal{O}(3)$ to be the rotation which takes $\mathbf{x}$ to $\mathbf{o}$ (the origin in $\mathbb{S}^2$) i.e. the rotation matrix through the angle $\alpha = \cos^{-1}(\mathbf{o}^T\,\mathbf{x})$ anticlockwise (noting that $||\,\mathbf{x}\,|| = ||\,\mathbf{o}\,|| = 1$) around the axis $\mathbf{x} \times \mathbf{o}$ where $\times$ is the cross product of vectors.

We define the $d-1$-dimensional spherical point process X to follow Definition 2.1.1 with $S = \mathbb{S}^{d-1}$. We say that X is isotropic in accordance with Definition 2.1.4 and note that since $\mathbb{S}^{d-1}$ is a compact set notions of stationarity do not make sense on such a metric space since a translation of the points does not map to the same space i.e. if $T_{\mathbf{y}}(\mathbf{x}) = \mathbf{x} + \mathbf{y}$ is the translation of $\mathbf{x}$ by a vector $\mathbf{y} \in \mathbb{R}^d$ then $T : \mathbb{S}^{d-1} \not\to \mathbb{S}^{d-1}$. Further note that an isotropic spheroidal point process also has constant intensity function $\rho \in \mathbb{R}_+$ [Møller and Rubak, 2016].

3.3 Isotropic functional summary statistics

The functional summary statistics for isotropic spheroidal point processes were first discussed, as noted in the literature review, by Robeson et al. [2014], Lawrence et al. [2016] and Møller and Rubak [2016]. More precisely the empty-space, nearest-neighbour, J, and

Ripley's K-functions are defined as

$$F(r) = P(X_{B_{\mathbb{S}^2}(\boldsymbol{o},r)} \neq \emptyset) \tag{3.1}$$

$$H(r) = P(X^!_{\boldsymbol{o},B_{\mathbb{S}^2}(\boldsymbol{o},r)} \neq \emptyset) \tag{3.2}$$

$$J(r) = \frac{1 - H(r)}{1 - F(r)} \tag{3.3}$$

$$K(r) = \frac{1}{\rho}\mathbb{E} \sum_{\mathbf{x} \in X^!_{\boldsymbol{o}}} \mathbb{1}[d_{\mathbb{S}^2}(\mathbf{x}, \mathbf{o}) \leq r], \tag{3.4}$$

for $r \in [0, \pi]$ and where ρ is the constant intensity respectively for rotationally invariant point processes on $\mathbb{S}^2$. By isotropy these functional summary statistics, as in the stationary Euclidean case, are independent of the origin $\mathbf{o}$. The definitions given above of the H and K-functions are based on the Palm interpretation and can equivalently be defined as

$$H(r) = \frac{1}{\rho\lambda_{\mathbb{S}^2}(A)}\,\mathbb{E} \sum_{\mathbf{x} \in X_A} \mathbb{1}[(X \setminus \mathbf{x}) \cap B_{\mathbb{S}^2}(\mathbf{x}, r) \neq \emptyset]$$

$$K(r) = \frac{1}{\rho^2\lambda_{\mathbb{S}^2}(A)}\mathbb{E} \sum_{\mathbf{x},\mathbf{y} \in X}^{\neq} \mathbb{1}[\mathbf{x} \in A, d_{\mathbb{S}^2}(\mathbf{x}, \mathbf{y}) \leq r],$$

where A is an arbitrary subset of $\mathbb{S}^2$. This identity follows from the Campbell-Mecke Theorem (see Theorem 2.1.9). Lawrence et al. [2016] and Møller and Rubak [2016] then suggest border-corrected estimates analogous to Equation 2.14 with constant intensity function for Ripley's K-function, Equation 2.15 for the nearest-neighbour function, Equations 2.16 and 2.17 for the empty-space function and Equation 2.18 for the J-function.

A Poisson process over the unit sphere is defined by Definition 2.2.2 with $S = \mathbb{S}^2$. It can then be shown that for a spherical CSR process with constant intensity function ρ we have,

$$\begin{aligned} F(r) &= 1 - \exp(-\rho 2\pi(1 - \cos(r))) \\ H(r) &= 1 - \exp(-\rho 2\pi(1 - \cos(r))) \\ J(r) &= 1 \\ K(r) &= 2\pi(1 - \cos(r)), \end{aligned}$$

for $r \in [0, \pi]$, where we note that $\lambda_{\mathbb{S}^2}(B_{\mathbb{S}^2}(\mathbf{x}, r)) = 2\pi(1 - \cos(r))$ for any $\mathbf{x} \in \mathbb{S}^2$.

Additionally, the L-function was discussed by Lawrence [2018] for spherical processes. Originally the L-function arose as the result of Besag [1977] discussion on Ripley [1977] work

in which they noted that for a CSR process on $\mathbb{R}^2$, $\sqrt{K(r)}$ is approximately variance stabilised. Furthermore by introducing a factor of $1/\sqrt{\pi}$ Besag [1977] introduces the L-function which stabilises for variance over different scales and also linearises the function, i.e. $L(r) = \sqrt{K(r)/\pi} = r$. As discussed by Lawrence [2018] $\sqrt{K(r)}$ is still a variance stabilising transformation for isotropic spherical processes but a simple factor, as in the planar case, is not trivially introduced in order for the transformation to also be linearised and therefore they define $K^{\text{stab}}(r) = \sqrt{K(r)}$ so as to avoid confusion with the more typical L-function. Lawrence [2018] does suggest centralising functional summary statistics by subtracting their theoretical counterpart. Figure 3.2 demonstrates these isotropic functional summary statistics comparing them to a null Poisson hypothesis, where the middle plot corresponds to the function $P(r) = K^{\text{stab}}(r) - \sqrt{2\pi(1 - \cos(r))}$.

3.4 Inhomogeneous K-function

For inhomogeneous point processes Møller and Rubak [2016] and Lawrence et al. [2016] introduce the inhomogeneous K-function on $\mathbb{S}^2$. To properly define the inhomogeneous K-function we introduce the notion of *second order intensity reweighted isotropic* (SOIRWI) which extends SOIRWS Baddeley et al. [2000] from $\mathbb{R}^d$ to $\mathbb{S}^{d-1}$, first discussed in Møller and Rubak [2016] and Lawrence et al. [2016]. The following definition does just this extending Definition 2.1.5 to the sphere.

Definition 3.4.1. *Let X be a point process over $\mathbb{S}^2$ with locally integrable intensity function $\rho : \mathbb{S}^2 \mapsto \mathbb{R}_+$. Suppose that the measure*

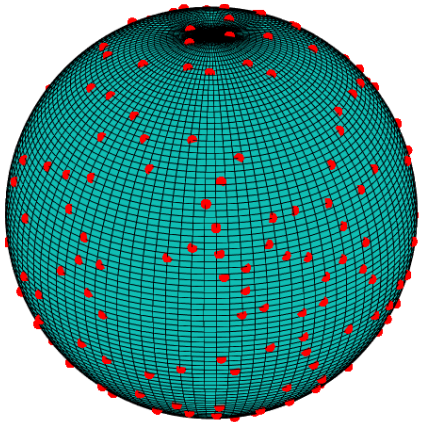
$$\mathcal{K}(B) = \frac{1}{\lambda_{\mathbb{S}^2}(A)} \mathbb{E} \sum_{\mathbf{x},\mathbf{y} \in X}^{\neq} \frac{\mathbb{1}[\mathbf{x} \in A, O_{\mathbf{x}}\, \mathbf{y} \in B]}{\rho(\mathbf{x})\rho(\mathbf{y})}, \quad B \subseteq \mathbb{S}^2 \tag{3.5}$$

does not depend on $A \subseteq \mathbb{S}^2$ with $0 < \lambda_{\mathbb{S}^2}(A) < \infty$ and take $a/0 = 0$ for any $a \geq 0$. Then X is said to be second order intensity reweighted isotropic and $\mathcal{K}$ is referred to as the second order reduced moment measure on $\mathbb{S}^2$.

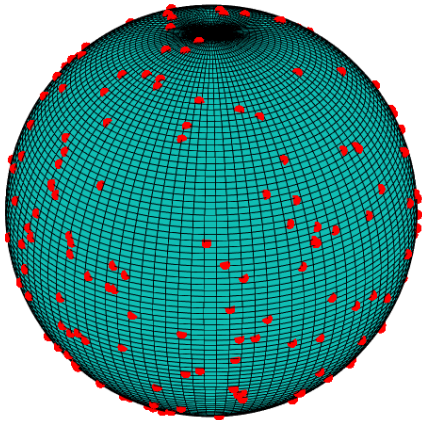
Similar to Equation 2.12 we also have the Palm interpretation,

$$\mathcal{K}(B) = \int_{N_{lf}} \sum_{\mathbf{x} \in x} \frac{\mathbb{1}[O_{\mathbf{y}}\, \mathbf{x} \in B]}{\rho(\mathbf{x})} P^!_{\mathbf{y}}(dx), \tag{3.6}$$

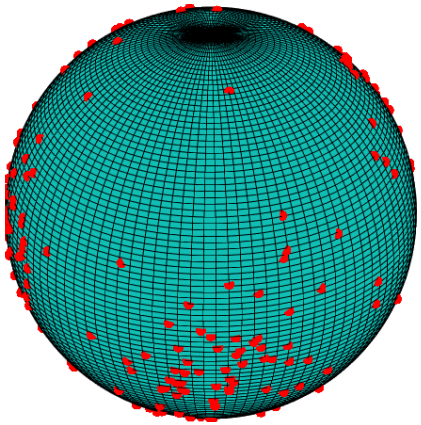
for almost all $\mathbf{y} \in \mathbb{S}^2$. Like in the Euclidean case this follows by the Campbell-Mecke Theorem



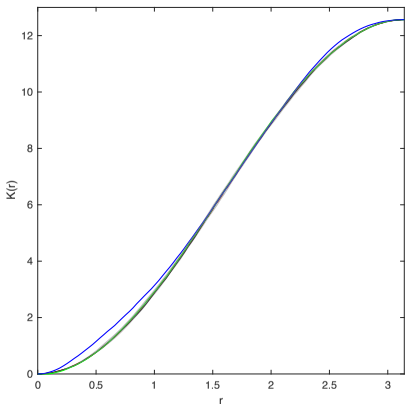
(a) Realisation of a regular process



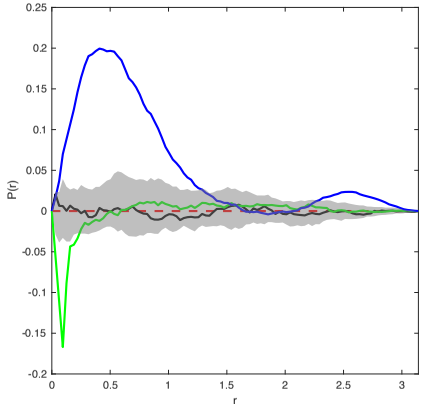
(b) Realisation of a CSR process



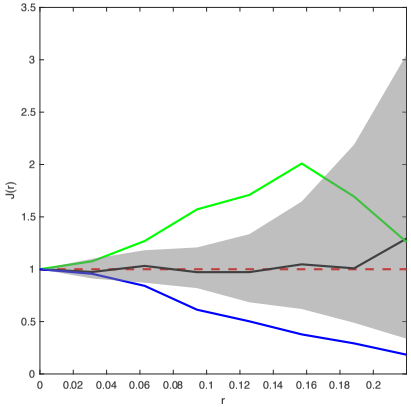
(c) Realisation of a cluster process



(d) Example K-function plot



(e) Example P-function plot



(f) Example J-function plot

Figure 3.2: Plots of functional summary statistics for typical isotropic spherical point patterns. All processes shown are over $\mathbb{S}^2$ and have an expected number of 200 over the unit square. **Top row:** realisation of a Matérn II process with hard-core distance 0.1 (*left*), realisation of a CSR process (*middle*), and realisation of a Neymann-Scott process with an expected number of 20 offspring per parent and the offspring are distributed with density proportional to $\exp(-\kappa d_{\mathbb{S}^2}(\mathbf{x}, \mathbf{c}))$ where $\kappa = 5$ and $\mathbf{c}$ corresponds to the parent (*right*). **Bottom row:** plots of functional summary statistics of Matérn II (*green*), CSR (*black*), and Neymann-Scott process (*blue*), with the theoretical K-function for a CSR process (*dashed red*) and simulation envelopes constructed from 99 realisations of a CSR process.

(see Theorem 2.1.9). We can define the inhomogeneous K-function as,

$$\mathrm{K}_{\mathrm{inhom}}(r) = \mathcal{K}(B_{\mathbb{S}^2}(\mathbf{o}, r)) = \frac{1}{\lambda_{\mathbb{S}^2}(A)} \mathbb{E} \sum_{\mathbf{x},\mathbf{y} \in X}^{\neq} \frac{\mathbb{1}[\mathbf{x} \in A, O_{\mathbf{x}}\, \mathbf{y} \in B_{\mathbb{S}^2}(\mathbf{o}, r)]}{\rho(\mathbf{x})\rho(\mathbf{y})}, \tag{3.7}$$

for $r \in [0, \pi]$ and any $A \subseteq \mathbb{S}^2$ such that $0 < \lambda_{\mathbb{S}^2}(A)$, which by Equation 3.6 has an equivalent Palm interpretation with B replaced by $B_{\mathbb{S}^2}(\mathbf{o}, r)$. In the event that the process is completely observed over all $\mathbb{S}^2$ then edge correction is not needed and in the event that the process is only observed over a portion of $\mathbb{S}^2$ both Møller and Rubak [2016] and Lawrence et al. [2016] suggest border corrected versions of the estimator. Additionally, they both also suggest to replace ρ with some estimator $\hat{\rho}$ in the more typical scenario when ρ is unknown which generally leads to a biased estimator for $\mathrm{K}_{\mathrm{inhom}}$, although it may be possible to achieve ratio-unbiasedness in specific cases.

Defining $h : \mathbb{S}^2 \times \mathbb{S}^2 \mapsto \mathbb{R}$ to be our PCF and suppose that it is isotropic, that is to say for any rotation $O \in \mathcal{O}(3)$ $h(\mathbf{x}, \mathbf{y}) = h(O\,\mathbf{x}, O\,\mathbf{y})$ which is in turn equivalent to $h(\mathbf{x}, \mathbf{y}) = h(d_{\mathbb{S}^2}(\mathbf{x}, \mathbf{y}))$ depending only on the distance between $\mathbf{x}$ and $\mathbf{y}$, then it can be shown that,

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

where $h(\mathbf{x}) = h(d_{\mathbb{S}^2}(\mathbf{o}, \mathbf{x}))$. This follows by the Campbell-Mecke Theorem (see Theorem 2.1.9). In this scenario and when focusing on the inhomogeneous K-function we can rewrite $\mathrm{K}_{\mathrm{inhom}}$ using spherical coordinates as,

$$\mathrm{K}_{\mathrm{inhom}}(r) = 2\pi \int_0^r h(t) \sin(t) dt, \quad r \in [0, \pi] \tag{3.8}$$

as shown by Møller and Rubak [2016].

3.5 Extensions of inhomogeneous F, H, and J-functions to $\mathbb{S}^2$

Having reviewed the existing functional summary statistics for both isotropic and non-isotropic point processes we now discuss how to extend the inhomogeneous F, H, and J-functions from $\mathbb{R}^d$, first discussed by van Lieshout [2011], to $\mathbb{S}^{d-1}$, where we primarily focus on $d = 3$. The argument follows analogously for other values of d. Many of the proofs in this section follow similarly to those found in van Lieshout [2011] with certain caveats to account for the different geometry imposed by being on a sphere as opposed to a Euclidean space.

The inhomogeneous F, H, and J-functions are constructed based on an infinite series representation of the isotropic F and H-functions first discussed by White [1979] in $\mathbb{R}^d$. Corollary 3.5.2 gives the same representation as White [1979] but for spherical instead of Euclidean point processess.

Theorem 3.5.1. *Let X be an isotropic spheroidal point process with constant intensity function ρ. Further we assume the existence of all n^{th}-order factorial moment measures for both X and its reduced Palm process, $X^!_{\mathbf{x}}$. Then the F and H-functions have the following series representation,*

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

where $\alpha^{(n)}$ and $\alpha^{!(n)}_{\mathbf{x}}$ are the factorial moment measure for X and $X^!_{\mathbf{x}}$ 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, $\lim_{n\to\infty} |a_{n+1}/a_n| < 1$ or $\limsup_{n\to\infty}(|a_n|)^{1/n} < 1$, where we set $a_n = ((-1)^n/n!)\ \alpha^{(n)}(B_{\mathbb{S}^2}(\mathbf{o}, r), \ldots, B_{\mathbb{S}^2}(\mathbf{o}, r))$ for the F-function and set $a_n = ((-1)^n/n!)\ \alpha^{!(n)}_{\mathbf{o}}(B_{\mathbb{S}^2}(\mathbf{o}, r) \ldots, B_{\mathbb{S}^2}(\mathbf{o}, r))$ for the H-function.

Proof. See Appendix A.1. □

The following corollary provides an infinite series representations of the F, and H-function from Theorem 3.5.1 in terms of the n^{th}-order product intensities, assuming they exists for all $n \in \mathbb{N}$. This representation is similar to the one given by White [1979] the only difference being that they worked on the Euclidean plane and here we are on a sphere.

Corollary 3.5.2. *(White [1979]) Let X be an isotropic spheroidal point process with constant intensity function ρ. Further we assume the existence of all n^{th}-order product intensities for both X and its reduced Palm process, $X^!_{\mathbf{x}}$. Then the F and H-functions have the following series representation,*

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

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 Appendix A.2. For original proof on $\mathbb{R}^d$ see White [1979]. □

Before introducing the infinite series representation for the isotropic J-function we need to introduce the n^{th}*-order correlation functions*. These are a useful alternative to the n^{th}-order product intensities [van Lieshout, 2011] and are commonly used in the astrophysics literature [Peebles, 1980].

Definition 3.5.3. *Let* X *be a spherical point process for which all its* n^{th}*-order product intensities exists. Then the* n^{th}*-order correlation functions are recursively defined for* $n \in \mathbb{N}$*, based on* n^{th}*-order product intensities, with* $\xi_1 = 1$ *and*

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

where the final sum ranges over all partitions $\{D_1, \ldots, D_k\}$ *of* $\{1, \ldots, n\}$ *in* k *non-empty, disjoint sets,* $\mathbf{x}_{D_j} = \{\mathbf{x}_i : i \in D_j\}$, $j = 1, \ldots, k$, *and* $\mathbf{x}_i \in \mathbb{S}^2$.

The following theorem provides the infinite series representation of the isotropic J-function which was first discussed by van Lieshout [2006] for marked point processes on a Euclidean space and is based on the n^{th}-order correlation functions. The following theorem is adapted from Proposition 4.2 of van Lieshout [2006] where the underlying metric space is now $\mathbb{S}^2$ rather than $\mathbb{R}^d$ and focuses on unmarked point processes rather than marked as considered originally by van Lieshout [2006].

Theorem 3.5.4. *Let* X *be an isotropic, spherical point process with constant intensity function* ρ*. Further, assume that the* n^{th}*-order intensity functions exists for all* $n \in \mathbb{N}$*. Then the* J*-function has the following series representation,*

$$J(r) = 1 + \sum_{n=1}^{\infty} \frac{(-\rho)^n}{n!} J_n(r), \quad 0 \leq r \leq \pi \tag{3.9}$$

for which $F(r) < 1$, $J_n(t) = \int_{B_{\mathbb{S}^2}(\mathbf{o},r)} \cdots \int_{B_{\mathbb{S}^2}(\mathbf{o},r)} \xi^{(n+1)}(0, \mathbf{x}_1, \ldots, \mathbf{x}_n) \lambda_{\mathbb{S}^2}(d\mathbf{x}_1) \cdots \lambda_{\mathbb{S}^2}(d\mathbf{x}_n)$*, and* $B_{\mathbb{S}^2}(\mathbf{o}, r)$ *is the spherical cap at the origin* $\mathbf{o} \in \mathbb{S}^2$.

Proof. See Appendix 3.5.4. The original proof for marked point processes on $\mathbb{R}^d$ is given in Proposition 4.2 of van Lieshout [2006]. □

Next we introduce the notion of a spherical point process being *intensity reweighted moment isotropic* (IRWMI), which extends to *intensity reweighted moment stationary* (IRWMS) for Euclidean processes defined by van Lieshout [2011] to the sphere by basing invariance on rotations as opposed to translations.

Definition 3.5.5. *Let X be a spheroidal point process. Then X is said to be intensity reweighted moment isotropic if, for all $n \in \mathbb{N}$, the n^{th}-order correlation functions are rotationally invariant. That is $\xi_n(\mathbf{x}_1, \ldots, \mathbf{x}_n) = \xi_n(O\mathbf{x}_1, \ldots, O\mathbf{x}_n)$ for all $n \in \mathbb{N}$ and $O \in \mathcal{O}(3)$.*

We note a few immediate results from this definitions similar to those identified by van Lieshout [2011]. First an isotropic process X is immediatedly IRWMI. Additionally if we suppose that X is IRWMI then for $n = 2$ and using Definition 3.5.3,

$$h(\mathbf{x}, \mathbf{y}) = \frac{\rho^{(2)}(\mathbf{x}, \mathbf{y})}{\rho(\mathbf{x})\rho(\mathbf{y})} = \xi_2(\mathbf{x}, \mathbf{y}) + 1,$$

where h is the PCF of X. Then by rotational invariance of ξ_2 we have that the pair correlation function is isotropic which, from our discussion on the inhomogeneous K-function, means that X is also SOIRWI. This also highlights that IRWMI is a stronger condition then SOIRWI. We are now in a position to define the inhomogeneous J-function on a sphere.

Definition 3.5.6. *For an IRWMI point process, X, with intensity function $\rho : \mathbb{S}^2 \mapsto \mathbb{R}$ such that $\bar{\rho} \equiv \inf_{\mathbf{x} \in \mathbb{S}^2} \rho(\mathbf{x}) > 0$,*

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

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

Based on this definition and Theorem 3.5.4 in the event that the process is isotropic with constant intensity function ρ, then $\bar{\rho} = \rho$ and the definition of the inhomogeneous J-function aligns with the isotropic one. As identified by van Lieshout [2011] $\mathrm{J}_{\text{inhom}}(r) > 1$ suggests inhibition at range r whilst $\mathrm{J}_{\text{inhom}}(r) < 1$ suggests aggregation or clustering at range r.

Similar to Theorem 1 of van Lieshout [2011] the following theorem shows that the inhomogeneous J-function can be represented as the ratio of generating functionals of the reduced Palm process $X^!_{\mathbf{x}}$ and the original process X.

Theorem 3.5.7. *For all $r \in [0, \pi]$ and $\mathbf{y} \in \mathbb{S}^2$,*

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

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

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

is absolutely convergent. Then under the further assumptions associated with Definition 3.5.6 of the inhomogeneous J-function and the existence of all n^{th}-order intensity function $\rho_{\mathbf{y}}^{!(n)}$ for the reduced Palm distribution $X_{\mathbf{y}}^!$, $\forall \mathbf{y} \in \mathbb{S}^2$,

$$J_{\text{inhom}}(r) = \frac{G_{X_{\mathbf{y}}^!}(1 - u_r^{\mathbf{y}})}{G_X(1 - u_r^{\mathbf{y}})}, \quad r \in [0, \pi]$$

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

Proof. See Appendix A.4. See Theorem 1 van Lieshout [2011] for the original proof for point processes on $\mathbb{R}^d$. □

By an identical argument that van Lieshout [2011] makes immediately after stating Theorem 1 in their work, we note that in the isotropic case $u_r^{\mathbf{y}}(\mathbf{x}) = \mathbb{1}[\mathbf{x} \in B_{\mathbb{S}^2}(\mathbf{y}, r)]$ and so,

$$\begin{aligned} G_X(1 - u_r^{\mathbf{y}}) &= P(X_{B_{\mathbb{S}^2}(\mathbf{y},r)} = \emptyset) = P(X_{B_{\mathbb{S}^2}(\mathbf{o},r)} = \emptyset) = 1 - F(r) \\ G_{X_{\mathbf{y}}^!}(1 - u_r^{\mathbf{y}}) &= P(X^!_{\mathbf{y},B_{\mathbb{S}^2}(\mathbf{y},r)} = \emptyset) = P(X^!_{\mathbf{o},B_{\mathbb{S}^2}(\mathbf{o},r)} = \emptyset) = 1 - H(r), \end{aligned}$$

thus in the isotropic setting, the numerator and denominator of $\mathrm{J}_{\text{inhom}}$ as given by Theorem 3.5.7 collapse to their isotropic counterparts. This naturally extends the isotropic F and H-functions to the inhomogeneous setting as,

$$\begin{aligned} \mathrm{F}_{\text{inhom}}(r) &= 1 - G_X(1 - u_r^{\mathbf{y}}) \\ \mathrm{H}_{\text{inhom}}(r) &= 1 - G_{X_{\mathbf{y}}^!}(1 - u_r^{\mathbf{y}}), \end{aligned}$$

where by the proof of Theorem 3.5.7 this holds for almost all $\mathbf{y} \in \mathbb{S}^2$.

In order to develop estimators of $\mathrm{F}_{\text{inhom}}$ and $\mathrm{H}_{\text{inhom}}$ we follow van Lieshout [2011] who

suggests border-corrected nonparametric estimators. Thus suppose we observe X within some window $W \subseteq \mathbb{S}^2$ and let P be a deterministic, finite point configuration over W. Then we estimate $\mathrm{F}_{\mathrm{inhom}}$ and $\mathrm{H}_{\mathrm{inhom}}$ as,

$$\hat{F}_{\mathrm{inhom}}(r) = 1 - \frac{\sum_{\mathbf{p}\in P\cap W_{\ominus r}} \prod_{\mathbf{x}\in X\cap B(\mathbf{p},r)} \left(1 - \frac{\bar{\rho}}{\rho(\mathbf{x})}\right)}{|P \cap W_{\ominus r}|} \tag{3.10}$$

$$\hat{H}_{\mathrm{inhom}}(r) = 1 - \frac{\sum_{\mathbf{x}\in X\cap W_{\ominus r}} \prod_{\mathbf{y}\in (X\setminus\mathbf{x})\cap B(\mathbf{x},r)} \left(1 - \frac{\bar{\rho}}{\rho(\mathbf{y})}\right)}{N_X(W_{\ominus r})}, \tag{3.11}$$

and therefore noting Theorem 3.5.7 we estimate the inhomgeneous J-function as,

$$\hat{\mathrm{J}}_{\mathrm{inhom}}(r) = \frac{1 - \hat{\mathrm{H}}_{\mathrm{inhom}}(r)}{1 - \hat{\mathrm{F}}_{\mathrm{inhom}}(r)}. \tag{3.12}$$

To select P we use the algorithm of Deserno [2004] to generate equidistant points on the unit sphere: we set the number of points to 99. The following proposition, like Proposition 1 van Lieshout [2011], shows unbiasedness of $\hat{\mathrm{F}}_{\mathrm{inhom}}$ and ratio-unbiasedness of $\hat{\mathrm{H}}_{\mathrm{inhom}}$ and $\hat{\mathrm{J}}_{\mathrm{inhom}}$.

Proposition 3.5.8. *The $\hat{\mathrm{F}}_{inhom}$ defined by Equation 3.10 is unbiased. $\hat{\mathrm{H}}_{inhom}$ and $\hat{\mathrm{J}}_{inhom}$ as defined by Equations 3.11 and 3.12 are ratio-unbiased.*

Proof. See Appendix A.5. See Proposition 1 van Lieshout [2011] for proof of Euclidean processes. □

3.6 Examples of inhomogeneous functional summary statistics

In this section we shall discuss classes of spherical point processes for which their inhomogeneous functional summary statistics are well defined. This builds on the works of Lawrence et al. [2016] who describes the inhomogeneous K-function for inhomogeneous Neyman-Scott processes that have a location dependent expected number of offspring per parent and van Lieshout [2011] who describes the inhomogeneous F, H, and J-functions for Poisson processes, Log Gaussian Cox processes (LGCP), and location dependent thinned processes. These examples follow similar resasoning as those given by van Lieshout [2011] on $\mathbb{R}^d$.

3.6.1 Poisson processes

Suppose X is a spherical Poisson process over $\mathbb{S}^2$ with 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 we also know that $\rho^{(n)}(\mathbf{x}_1, \ldots, \mathbf{x}_n) = \rho(\mathbf{x}_1) \cdots \rho(\mathbf{x}_n)$, this follows for example for the Slivynak-Mecke Theorem (see Theorem 2.2.4). Then

$$\frac{\rho^{(n)}(\mathbf{x}_1, \ldots, \mathbf{x}_n)}{\rho(\mathbf{x}_1) \cdots \rho(\mathbf{x}_n)} = 1,$$

and so the Poisson process is clearly IRWMI and hence also SOIRWI. Therefore the $\mathrm{K}_{\mathrm{inhom}}$, $\mathrm{F}_{\mathrm{inhom}}, \mathrm{H}_{\mathrm{inhom}}$, and $\mathrm{J}_{\mathrm{inhom}}$ are all well defined. Furthermore by, for example, Proposition 3.3 of Møller and Waagepetersen [2003] we have that the generating functional for a Poisson process is given by,

$$G_X(u) = \exp\left(-\int_{\mathbb{S}^2} (1 - u(\mathbf{x}))\rho(\mathbf{x})\lambda_{\mathbb{S}^2}(d\,\mathbf{x})\right), \tag{3.13}$$

for functions $u : \mathbb{S}^2 \mapsto [0, 1]$, which can be used to calculate $\mathrm{F}_{\mathrm{inhom}}$ and $\mathrm{H}_{\mathrm{inhom}}$. The $\mathrm{K}_{\mathrm{inhom}}, \mathrm{F}_{\mathrm{inhom}}, \mathrm{H}_{\mathrm{inhom}}$, and $\mathrm{J}_{\mathrm{inhom}}$ for a Poisson process are given by,

$$\begin{aligned}
\mathrm{K}_{\mathrm{inhom}}(r) &= 2\pi(1 - \cos(r)) \\
\mathrm{F}_{\mathrm{inhom}}(r) &= \mathrm{H}_{\mathrm{inhom}}(r) = 1 - \exp(-\bar{\rho}2\pi(1 - \cos(r))) \\
\mathrm{J}_{\mathrm{inhom}}(r) &= 1,
\end{aligned}$$

where we used Equation 3.13 for $\mathrm{F}_{\mathrm{inhom}}$ and $\mathrm{H}_{\mathrm{inhom}}$ and note that by the Slivnyak-Mecke Theorem $X^!_{\mathbf{x}} \stackrel{d}{=} X$ for $\mathrm{H}_{\mathrm{inhom}}$. In order to demonstrate these we now consider two examples of inhomogeneous Poisson processes with the following intensity functions,

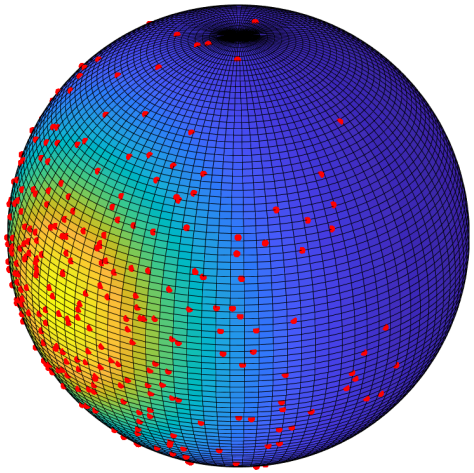
$$\rho_1(\mathbf{x}) = \exp(-4x + 2) \tag{3.14}$$

$$\rho_2(\mathbf{x}) = \frac{300}{4\pi} \cdot \frac{2}{5}\left(\sin(-\pi x) + \frac{3}{2}\right), \tag{3.15}$$

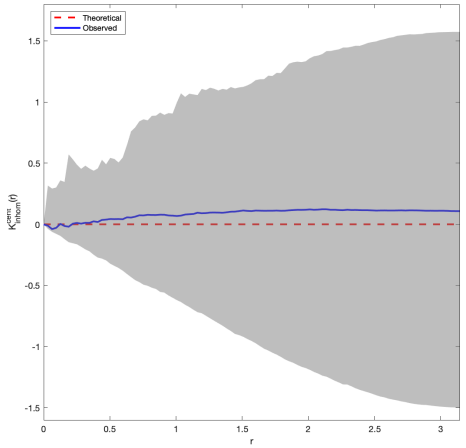
where $\mathbf{x} = (x, y, z)^T \in \mathbb{S}^2$. Examples of their functional summary statistics are given in Figure 3.3.

3.6.2 LGCPs

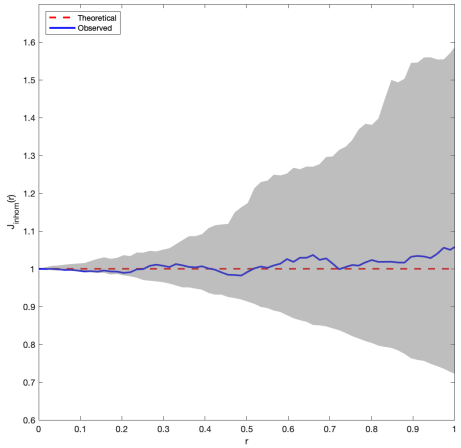
LGCPs are a popular class of models frequently used for aggregation. Recently Cuevas-Pacheco and Møller [2018] discussed the theoretical considerations of LGCPs on spheres,



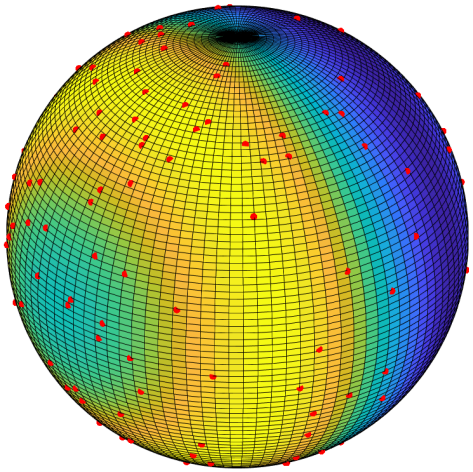
(a) Realisation of an inhomogeneous Poisson process with intensity ρ_1 (see Equation 3.14)



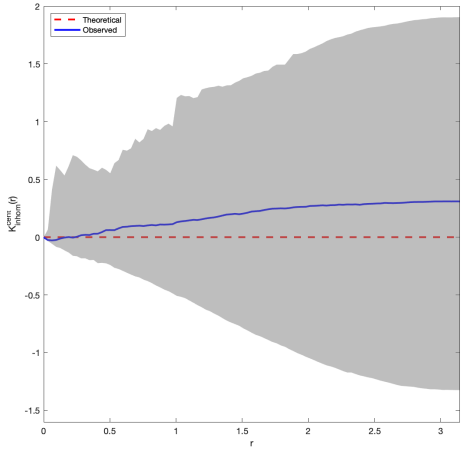
(b) Plot of $K^{\text{cent}}_{\text{inhom}}(r)$



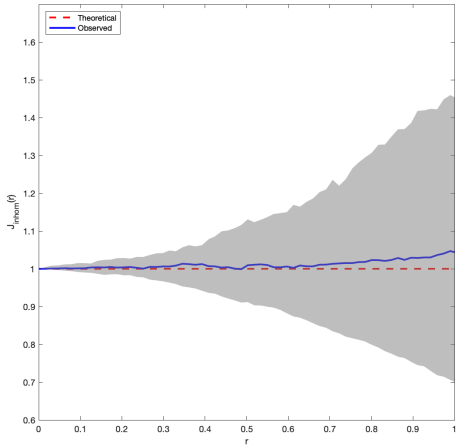
(c) Plot of $J_{\text{inhom}}(r)$



(d) Realisation of an inhomogeneous Poisson process with intensity ρ_2 (see Equation 3.15)



(e) Plot of $K^{\text{cent}}_{\text{inhom}}(r)$



(f) Plot of $J_{\text{inhom}}(r)$

Figure 3.3: Examples of inhomogeneous functional summary statistics for Poisson processes. **Top row:** Realisation of a Poisson process with intensity ρ_1 (*left*), plot of $K^{\text{cent}}_{\text{inhom}}(r)$ (*middle*), and plot of $J_{\text{inhom}}(r)$ (*right*. **Bottom row:** Realisation of a Poisson process with intensity ρ_2 (*left*), plot of $K^{\text{cent}}_{\text{inhom}}(r)$ (*middle*), and plot of $J_{\text{inhom}}(r)$ (*right*. The simulation envelopes are constructed from 99 inhomogeneous Poisson processes with intensity function given by Equations 3.14 and 3.15 for the top and bottom row respectively.

in particular providing results of their existence based on the work of Lang et al. [2016] who discuss continuity of random fields on manifolds. There have also been discussions in the multivariate setting on spheres by Jun et al. [2019] who use LGCPs to model rainfall patterns on a global scale.

To define a LGCP we first consider the more general definition of a Cox process which we will take from Chiu et al. [1995]. Let us define $[\mathbb{M}, \mathcal{M}]$ to be the measurable space over non-negative finite measures on $\mathbb{S}^2$ and let Ψ be a random measure with distribution $\mathbb{P}_\Psi$ over $[\mathbb{M}, \mathcal{M}]$. We then say X is a Cox process with driving random measure Ψ if, given Ψ, X is a Poisson process with intensity measure Ψ. Throughout this thesis we shall assume that almost surely Ψ is absolutely continuous with respect $\lambda_{\mathbb{S}^2}$ and so its Radon-Nikodyn derivative almost surely exists and is thus the conditional intensity function of X given Ψ; we shall use $\psi(\mathbf{x}) = \Psi(d\,\mathbf{x})/\lambda_{\mathbb{S}^2}(d\,\mathbf{x}) = d\Psi/d\lambda_{\mathbb{S}^2}$ to denote the Radon-Nikodyn derivative where $\Psi(B) = \int_B \psi(\mathbf{x})\lambda_{\mathbb{S}^2}(d\,\mathbf{x})$. We shall also assume that all moments of ψ exists, that is $\mathbb{E}\prod_{i=1}^n \psi(\mathbf{x}_i)$, for $\mathbf{x}_i \in \mathbb{S}^2, i = 1, \ldots, n$ $n \in \mathbb{N}$ exist. It can easily be shown that the n^{th}-order intensity function of the Cox process is,

$$\rho^{(n)}(\mathbf{x}_1, \ldots, \mathbf{x}_n) = \mathbb{E}\prod_{i=1}^n \psi(\mathbf{x}_i),$$

by taking iterated expectations conditioning on the random measure, Ψ [Møller and Waagepetersen, 1998].

We now briefly discuss Palm processes of Cox processes following a similar discussion to that of Coeurjolly et al. [2017]. Since we assume that Ψ is absolutely continuous with respect $\lambda_{\mathbb{S}^2}$, X, on $\mathbb{S}^2$ is then absolutely continuous with respect to the standard unit rate Poisson over $\mathbb{S}^2$. Consider $W \subseteq \mathbb{S}^2$, then conditional on Ψ, X_W has density,

$$f(x|\Psi) = \exp\left(\lambda_{\mathbb{S}^2}(W) - \int_W \psi(\mathbf{x})\lambda_{\mathbb{S}^2}(d\,\mathbf{x})\right)\prod_{\mathbf{x}\in x}\psi(\mathbf{x}),$$

for finite point configurations $x \subseteq W$ which follows from Equation 2.6. Then unconditionally X_W has density $\mathbb{E}[f(x|\Psi)]$ and the reduced Palm process given typical points $\mathbf{x}_1, \ldots, \mathbf{x}_n \in W$ has density,

$$f_{\mathbf{x}_1,\ldots,\mathbf{x}_n}(x) = \mathbb{E}\left[f(x|\Psi)\frac{\prod_{i=1}^n \psi(\mathbf{x}_i)}{\rho^{(n)}(\mathbf{x}_1, \ldots, \mathbf{x}_n)}\right], \tag{3.16}$$

which follows from Equation 2.4 [Coeurjolly et al., 2017]. If we then define $\Psi_{\mathbf{x}_1,\ldots,\mathbf{x}_n}$ to be the random measure that is absolutely continuous with respect to Ψ with density $\left(\prod_{i=1}^n \psi(\mathbf{x}_i)\right) / \rho^{(n)}(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ then by Equation 3.16 we can see that $X^!_{W,\mathbf{x}_1,\ldots,\mathbf{x}_n}$ is also a Cox process with

driving random measure $\Psi_{\mathbf{x}_1,\ldots,\mathbf{x}_n}$ [Coeurjolly et al., 2017]. For generalisations when Ψ is not necessarily absolutely continuous with respect to $\lambda_{\mathbb{S}^2}$ see, for example, Chiu et al. [1995, p. 156].

As the name suggests a LGCP is defined by an underlying Gaussian object, more precisely a Gaussian random field (GRF). Let $U : \mathbb{S}^2 \mapsto \mathbb{R}$ be a random function over the unit sphere, then U is said to be a GRF if for any $n \in \mathbb{N}$, $\mathbf{x}_1, \ldots, \mathbf{x}_n \in \mathbb{S}^2$ and $a_1, \ldots, a_n \in \mathbb{R}$ then $\sum_{i=1}^n a_i U(\mathbf{x}_i)$ is a univariate Gaussian random variable. By this definition we know that the GRF is then defined by its mean function $\mu : \mathbb{S}^2 \mapsto \mathbb{R}$ and covariance function $c : \mathbb{S}^2 \times \mathbb{S}^2 \mapsto \mathbb{R}$ such that,

$$\begin{aligned} \mu(\mathbf{x}) &= \mathbb{E}[U(\mathbf{x})] \\ c(\mathbf{x}, \mathbf{y}) &= \mathbb{E}[(U(\mathbf{x}) - \mu(\mathbf{x}))(U(\mathbf{y}) - \mu(\mathbf{y}))]. \end{aligned}$$

A LGCP is then defined through the Radon-Nikodyn derivative, ψ, of a Cox process. More precisely,

$$\psi(\mathbf{x}) = \mathrm{e}^{U(\mathbf{x})},$$

where U is a GRF with mean function $\mu : \mathbb{S}^2 \mapsto \mathbb{R}$ and covariance function $c : \mathbb{S}^2 \times \mathbb{S}^2 \mapsto \mathbb{R}$. In order for the resulting point process X to be well defined we require that the ψ is almost surely integrable in order for Ψ to be a well defined finite random measure over $\mathbb{S}^2$. This is discussed by Cuevas-Pacheco and Møller [2018] who argue that, given μ is continuous and bounded, almost sure integrability of ψ is satisfied if it is almost surely continuous. This in turn is satisfied if the zero mean field, $U_0(\mathbf{x}) = U(\mathbf{x}) - \mu(\mathbf{x})$, which has the same covariance function c as U, is almost surely so called *locally sample Hölder continuous of some order* $k > 0$. That is, a function $f : \mathbb{S}^2 \mapsto \mathbb{R}$ is said to be locally sample Hölder continuous of order k if, for every $\mathbf{z} \in \mathbb{S}^2$, there exists a non-empty neighbourhood of $\mathbf{z}$ denoted V and constant $C_{V,k}$ such that,

$$\sup_{\mathbf{x},\mathbf{y} \in V, \mathbf{x} \neq \mathbf{y}} \left| \frac{f(\mathbf{x}) - f(\mathbf{y})}{d_{\mathbb{S}^2}(\mathbf{x}, \mathbf{y})} \right| \leq C_{V,k}.$$

The following proposition, which is Proposition 1 of Cuevas-Pacheco and Møller [2018], provides a simple condition for a mean zero GRF being locally sample Hölder continuous of some order $k > 0$.

Proposition 3.6.1. *[Cuevas-Pacheco and Møller, 2018, Proposition 1] Let U_0 be a zero*

mean GRF with covariance function c. Define the variogram of U_0 as,

$$\gamma(\mathbf{x}, \mathbf{y}) = \frac{1}{2}\mathbb{E}\left[(U_0(\mathbf{x}) - U_0(\mathbf{y}))^2\right].$$

Suppose there exists $s \in (0, 1], l \in (0, 1)$, and $m > 0$ such that,

$$\gamma(\mathbf{x}, \mathbf{y}) \leq m(d_{\mathbb{S}^2}(\mathbf{x}, \mathbf{y}))^{l/2}, \quad \text{whenever } d_{\mathbb{S}^2}(\mathbf{x}, \mathbf{y}) < s. \tag{3.17}$$

Then, for any $k \in (0, l/2), U_0$ is almost surely locally sample Hölder continuous of order k.

This proposition is a special case of Corollary 4.5 of Lang et al. [2016] when the space under consideration is a sphere.

A common assumption on the covariance function is to suppose that it is isotropic; we will suppose this is the case from now on. That is $c(\mathbf{x}, \mathbf{y}) = \sigma^2 s(d_{\mathbb{S}^2}(\mathbf{x}, \mathbf{y})) = \sigma^2 s(r)$, for $\sigma^2 > 0$ and $r \in [0, \pi]$ where s is denoted the correlation function and only depends on the geodesic distance $d_{\mathbb{S}^2}(\mathbf{x}, \mathbf{y}) = r$ between points $\mathbf{x}, \mathbf{y} \in \mathbb{S}^2$ with $s(0) = 1$. Under this assumption Cuevas-Pacheco and Møller [2018] demonstrate that, for a number of correlation functions, it is possible to construct many well-defined LGCPs over $\mathbb{S}^2$.

Furthermore by arguments similar to those made by Møller and Waagepetersen [1998] we have that a LGCP defined by a GRF with mean function μ and covariance function c then,

$$\begin{aligned}
\rho^{(n)}(\mathbf{x}_1, \ldots, \mathbf{x}_n) &= \mathbb{E}\prod_{i=1}^{n}\psi(\mathbf{x}_i) \\
&= \mathbb{E}\prod_{i=1}^{n} e^{U(\mathbf{x}_i)} \\
&= \mathbb{E}\left[e^{\sum_{i=1}^{n} U(\mathbf{x}_i)}\right] \\
&= \left(\prod_{i=1}^{n}\exp\left(\mu(\mathbf{x}_i) + \frac{c(\mathbf{x}_i, \mathbf{x}_i)}{2}\right)\right)\left(\prod_{1\leq i<j\leq n}\exp\left(c(\mathbf{x}_i, \mathbf{x}_j)\right)\right),
\end{aligned}$$

which follows by using the moment generating function for a multivariate normal vector. Thus we have that,

$$\frac{\rho^{(n)}(\mathbf{x}_1, \ldots, \mathbf{x}_n)}{\prod_{i=1}^{n}\rho(\mathbf{x}_i)} = \prod_{1\leq i<j\leq n}\exp\left(c(\mathbf{x}_i, \mathbf{x}_j)\right),$$

and so LGCPs with isotropic covariance functions are IRWMI, and hence also SORWI. Therefore, the inhomogeneous K, F, H, and J-functions are all well-defined.

To calculate the theoretical inhomogeneous K-function for a LGCP with isotropic correlation function Equation 3.8 can be used where the pair correlation function h is,

$$h(\mathbf{x},\mathbf{y}) = \frac{\rho^{(2)}(\mathbf{x},\mathbf{y})}{\rho(\mathbf{x})\rho(\mathbf{y})} = \exp\left(c(\mathbf{x},\mathbf{y})\right) = \exp\left(\sigma^2 s(d_{\mathbb{S}^2}(\mathbf{x},\mathbf{y}))\right),$$

and so the inhomogeneous K-function is,

$$\mathrm{K}_{\mathrm{inhom}}(r) = 2\pi \int_0^r \exp\left(\sigma^2 s(t)\right)\sin(t)dt, \quad r \in [0,\pi].$$

For the inhomogeneous F and H-functions we need to calculate generating functionals for LGCPs. As such we use the following standard result for a Cox process, X, to do so (e.g. see Equation 5.2 of Møller and Waagepetersen [2003]),

$$G_X(u) = \mathbb{E}\left[\exp\left(-\int_{\mathbb{S}^2}(1-u(\mathbf{x}))\psi(\mathbf{x})\lambda_{\mathbb{S}^2}(d\,\mathbf{x})\right)\right].$$

Thus, like van Lieshout [2011], we can calculate the inhomogeneous F-function. First denote $\bar{\mu} = \inf_{\mathbf{x}\in\mathbb{S}^2} \mathrm{e}^{\mu(\mathbf{x})}$ then,

$$\begin{aligned}
1 - \mathrm{F}_{\mathrm{inhom}}(r) &= G_X(1-u_r^{\mathbf{y}}) \\
&= G_X(1-u_r^{\mathbf{o}}) \\
&= \mathbb{E}\left[\exp\left(-\int_{\mathbb{S}^2}\frac{\bar{\rho}\mathbb{1}[\mathbf{x}\in B_{\mathbb{S}^2}(\mathbf{o},r)]}{\rho(\mathbf{x})}\mathrm{e}^{U(\mathbf{x})}\lambda_{\mathbb{S}^2}(d\,\mathbf{x})\right)\right] \\
&= \mathbb{E}\left[\exp\left(-\bar{\mu}\int_{B_{\mathbb{S}^2}(\mathbf{o},r)}\mathrm{e}^{U(\mathbf{x})-\mu(\mathbf{x})}\lambda_{\mathbb{S}^2}(d\,\mathbf{x})\right)\right],
\end{aligned}$$

where the second equality holds since an LGCP is IRWMI so it does not depend on $\mathbf{y}$ and so we are free to choose it, the last line follows since $\rho(\mathbf{x}) = \exp(\mu(\mathbf{x}) + \sigma^2/2)$ where we have assumed an isotropic covariance function. The inhomogeneous H-function for LGCPs follows from our discussion of Palm processes of Cox processes (similar to van Lieshout [2011]), more specifically using Equation 3.16 we have,

$$\begin{aligned}
1 - \mathrm{H}_{\mathrm{inhom}}(r) &= G_{X^!_{\mathbf{y}}}(1-u_r^{\mathbf{y}}) \\
&= G_{X^!_{\mathbf{o}}}(1-u_r^{\mathbf{o}}) \\
&= \mathbb{E}\left[\prod_{\mathbf{x}\in X^!_{\mathbf{o}}}(1-u_r^{\mathbf{o}}(\mathbf{x}))\right]
\end{aligned}$$

$$= \mathbb{E}_{\Psi_{\mathbf{o}}}\left[\mathbb{E}_{X^!_{\mathbf{o}}|\Psi_{\mathbf{o}}}\left[\prod_{\mathbf{x}\in X^!_{\mathbf{o}}}(1-u_r^{\mathbf{o}}(\mathbf{x}))\middle|\Psi_{\mathbf{o}}\right]\right]$$
$$= \mathbb{E}_{\Psi}\left[\mathbb{E}_{X^!_{\mathbf{o}}|\Psi}\left[\prod_{\mathbf{x}\in X^!_{\mathbf{o}}}(1-u_r^{\mathbf{o}}(\mathbf{x}))\middle|\Psi\right]\frac{\psi(\mathbf{o})}{\rho(\mathbf{o})}\right],$$

$\mathbb{E}_{X^!_{\mathbf{o}}|\Psi}\left[\prod_{\mathbf{x}\in X^!_{\mathbf{o}}}(1-u_r^{\mathbf{o}}(\mathbf{x}))\middle|\Psi\right] = \exp\left(-\bar{\mu}\int_{B_{\mathbb{S}^2}(\mathbf{o},r)}\mathrm{e}^{U(\mathbf{x})-\mu(\mathbf{x})}\lambda_{\mathbb{S}^2}(d\,\mathbf{x})\right)$ since $X^!_{\mathbf{o}}$ is also a Cox process then we have as van Lieshout [2011] for LGCPs on $\mathbb{R}^d$,

$$= \mathbb{E}\left[\frac{\mathrm{e}^{U(\mathbf{o})}}{\mathrm{e}^{\mu(\mathbf{o})+\sigma^2/2}}\exp\left(-\bar{\mu}\int_{B_{\mathbb{S}^2}(\mathbf{o},r)}\mathrm{e}^{U(\mathbf{x})-\mu(\mathbf{x})}\lambda_{\mathbb{S}^2}(d\,\mathbf{x})\right)\right]$$
$$= \mathbb{E}\left[\frac{\mathrm{e}^{U(\mathbf{o})-\mu(\mathbf{o})}}{\mathrm{e}^{\sigma^2/2}}\exp\left(-\bar{\mu}\int_{B_{\mathbb{S}^2}(\mathbf{o},r)}\mathrm{e}^{U(\mathbf{x})-\mu(\mathbf{x})}\lambda_{\mathbb{S}^2}(d\,\mathbf{x})\right)\right].$$

Note that $\mathbb{E}\left[\mathrm{e}^{U(\mathbf{o})-\mu(\mathbf{o})}\right] = e^{\sigma^2/2}$ (which follows by the moment generating function for a normal random variable), then we can write, as van Lieshout [2011] does, the inhomogeneous J-function as,

$$\mathrm{J}_{\mathrm{inhom}}(r) = \frac{\mathbb{E}\left[\mathrm{e}^{U(\mathbf{o})-\mu(\mathbf{o})}\exp\left(-\bar{\mu}\int_{B_{\mathbb{S}^2}(\mathbf{o},r)}\mathrm{e}^{U(\mathbf{x})-\mu(\mathbf{x})}\lambda_{\mathbb{S}^2}(d\,\mathbf{x})\right)\right]}{\mathbb{E}\left[\mathrm{e}^{U(\mathbf{o})-\mu(\mathbf{o})}\right]\mathbb{E}\left[\exp\left(-\bar{\mu}\int_{B_{\mathbb{S}^2}(\mathbf{o},r)}\mathrm{e}^{U(\mathbf{x})-\mu(\mathbf{x})}\lambda_{\mathbb{S}^2}(d\,\mathbf{x})\right)\right]}.$$

Using an identical argument to that outlined in van Lieshout [2011], we can show that LGCPs with mean and positive correlation function defined to ensure almost surely continuous paths of the underlying Gaussian random field (see Proposition 3.6.1) results in $J_{\mathrm{inhom}}(r) \leq 1$ for all $r \in [0, \pi]$.

Proposition 3.6.2. *Let X be a LGCP such that the assumptions of Proposition 3.6.1 hold and the correlation function is nonnegative, then $J_{inhom}(r) \leq 1$ for all $r \in [0, \pi]$.*

Proof. See Appendix A.6. For a proof in $\mathbb{R}^d$ see discussion given by van Lieshout [2011] at the end of Section 5.4. □

In Figure 3.4 we consider three different LGCPs where the underlying Gaussian random field is simulated using the approach outlined in Cuevas et al. [2020]. We also thank the authors of Cuevas et al. [2020] for providing the code to simulate the GRFs. For each of them the underlying GRF has one of the following correlation functions,

1. The powered exponential correlation (Figure 3.4a) function,

$$s(r) = \exp\left(-\left(\frac{r}{\phi}\right)^{\alpha}\right), \quad \phi > 0, 0 < \alpha \leq 1.$$

2. The Matérn correlation (Figure 3.4b) function,

$$s(r) = \frac{2}{\Gamma(\alpha)}\left(\frac{r}{2\phi}\right)^{\alpha} K_{\alpha}\left(\frac{r}{\phi}\right), \quad \phi > 0, 0 < \alpha \leq \frac{1}{2}.$$

3. The generalised Cauchy correlation (Figure 3.4c) function,

$$s(r) = \left(1 + \left(\frac{r}{\phi}\right)^{\alpha}\right)^{-\frac{\tau}{\alpha}}, \quad \phi > 0, 0 < \alpha \leq 1, \tau > 0.$$

For the Matérn correlation function, Γ is the gamma function and K_{α} is the modified Bessel function of the second kind defined as,

$$\Gamma(x) = \int_0^{\infty} t^{x-1} \mathrm{e}^{-x} dx, \quad x > 0$$
$$K_{\alpha}(x) = \frac{\pi}{2} \frac{I_{-\alpha}(x) - I_{\alpha}(x)}{\sin(\alpha\pi)},$$

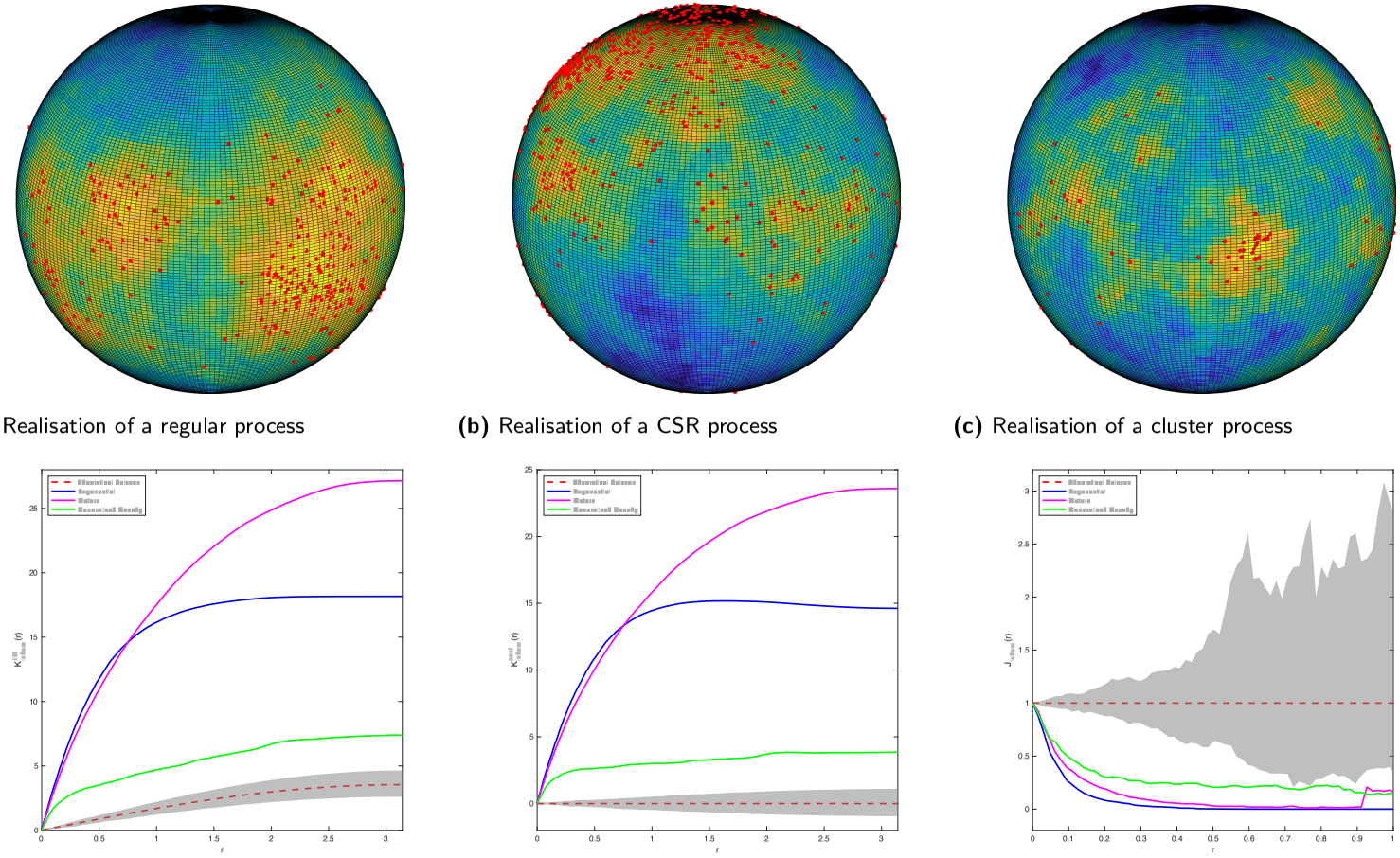
respectively, where

$$I_{\alpha}(x) = \sum_{n=0}^{\infty} \frac{1}{n!\Gamma(n + \alpha + 1)} \left(\frac{x}{2}\right)^{2n+\alpha}$$

is the modified Bessel function of the first kind. The mean function for each GRF is $\mu(\mathbf{x}) = x$, where $\mathbf{x} = (x, y, z)^T$ in Figure 3.4. Proposition 4 of Cuevas-Pacheco and Møller [2018] guarantees that Inequality 3.17 holds and hence the resulting LGCPs are well defined, i.e. $\psi(\mathbf{x}) = \mathrm{e}^{U(\mathbf{x})}$ is almost surely integrable where U is our GRF.

3.6.3 Location dependent thinning

The previous two examples have considered non-interacting inhomogeneous point processes (inhomogeneous Poisson processes) and clustered inhomogeneous point processes (LGCPs). We know consider functional summary statistics for inhomogeneous processes arising from *location dependent thinning* and can be used to exemplify inhomogeneous processes which exhibit regularity. The example considered here on the sphere is similar to that considered by van Lieshout [2011] for thinned point processes on $\mathbb{R}^d$.



(a) Realisation of a regular process

(b) Realisation of a CSR process

(c) Realisation of a cluster process

(d) Example K-function plot

(e) Example K^{cent}-function plot

(f) Example J-function plot

Figure 3.4: Plots of functional summary statistics for typical inhomogeneous spherical LGCPs, where all the underlying Gaussian random fields have mean function $\mu(\mathbf{x}) = x$ for $\mathbf{x} = (x, y, z)^T \in \mathbb{S}^2$. All processes shown are over $\mathbb{S}^2$ and have an intensity function of $\rho(\mathbf{x}) = \exp(x + \sigma^2/2)$, for $\mathbf{x} = (x, y, z)^T \in \mathbb{S}^2$. The intensities given by the colour are from the realisation of the Gaussian random field, i.e. the intensities shown are realisations of U. **Top row:** Realisations of LGCPs with different correlation functions, powered exponential with $(\sigma^2, \phi, \alpha) = (4, 1, 1)$ (*left*), Matérn with $(\sigma^2, \phi, \alpha) = (4, 1, 0.5)$, and generalised Cauchy with $(\sigma^2, \phi, \alpha, \tau) = (4, 1, 1, 3)$. **Bottom row:** plots of functional summary statistics of for the LGCPs simulated in the top row with the theoretical K-function for a Poisson process (*dashed red*) and simulation envelopes constructed from 99 realisations of inhomogeneous Poisson processes with intensity function $\rho(\mathbf{x}) = \exp(x + \sigma^2/2)$.

Thinning is a common operation applied to point processes in which points of the process are excluded based on a certain set of rules; this can lead to alternative models. Let X be a point process and suppose we have some rule $\Phi : N_{lf} \mapsto N_{lf}$ which dictates how thinning is applied to X to give X_{th}, the thinned process, i.e. $X_{th} = \Phi(X)$. Then Φ is often classified into one of the following two classes:

1. *Independent thinning*: the point $\mathbf{x} \in X$ is included/excluded in X_{th} independentally of any other point of $\mathbf{y} \in X$, or
2. *Dependent thinning*: the inclusion/exclusion of the point $\mathbf{x} \in X$ in X_{th} may depend on other points of X.

In this section we will consider functional summary statistics for independently thinned point processes over $\mathbb{S}^2$. We shall use dependent thinning to define a regular process and use this to exemplify how functional summary statistics can be constructed for independently thinned processes.

We will consider X_{th} to be the thinning of X by the rule Φ defined as,

$$\Phi(X) = \{\mathbf{x} \in X | R(\mathbf{x}) \leq p(\mathbf{x}), R(\mathbf{x}) \sim \text{Uniform}([0,1])\},$$

where $R(\mathbf{x})$ for any $\mathbf{x} \in \mathbb{S}^2$ are mutually independent and independent of X and $p : \mathbb{S}^2 \mapsto [0,1]$ is a known function with each $p(\mathbf{x})$ denoted the retention probability at $\mathbf{x}$. It can be shown that if X is SOIRWI or IRWMI then so is X_{th} respectively. This follows by standard results on the n^{th}-order factorial moment measures; Møller and Waagepetersen [2003] shows this for SOIRWS Euclidean processes in Proposition 4.3 and van Lieshout [2011] discusses this for IRWMS Euclidean processes. More explicitly, let X_{th} be the thinning of X and suppose that the intensity functions of all orders of X exists. Then consider the n^{th}-order factorial moment measure of X_{th}, $\alpha_{th}^{(n)}$,

$$\begin{aligned}
\alpha_{th}^{(n)}(A_1 \times \cdots \times A_n) &= \mathbb{E} \sum_{\mathbf{x}_1,\ldots,\mathbf{x}_n \in X_{th}}^{\neq} \mathbb{1}[x_1 \in A_1, \ldots, x_n \in A_n] \\
&= \mathbb{E} \sum_{\mathbf{x}_1,\ldots,\mathbf{x}_n \in X}^{\neq} \mathbb{1}[x_1 \in A_1, R(\mathbf{x}_1) \leq p(\mathbf{x}_1) \ldots, x_n \in A_n, R(\mathbf{x}_n) \leq p(\mathbf{x}_n)] \\
&= \int_{A_1} \cdots \int_{A_n} P(R(\mathbf{x}_1) \leq p(\mathbf{x}_1), \ldots, R(\mathbf{x}_n) \leq p(\mathbf{x}_n)) \rho(\mathbf{x}_1, \ldots, \mathbf{x}_n) \\
&\qquad \lambda_{\mathbb{S}^2}(d\,\mathbf{x}_1) \cdots \lambda_{\mathbb{S}^2}(d\,\mathbf{x}_n)
\end{aligned}$$

$$= \int_{A_1} \cdots \int_{A_n} \left(\prod_{i=1}^{n} P(R(\mathbf{x}_i) \leq p(\mathbf{x}_i)) \right) \rho(\mathbf{x}_1, \ldots, \mathbf{x}_n) \\ \lambda_{\mathbb{S}^2}(d\,\mathbf{x}_1) \cdots \lambda_{\mathbb{S}^2}(d\,\mathbf{x}_n)$$
$$= \int_{A_1} \cdots \int_{A_n} \left(\prod_{i=1}^{n} p(\mathbf{x}_i) \right) \rho(\mathbf{x}_1, \ldots, \mathbf{x}_n) \lambda_{\mathbb{S}^2}(d\,\mathbf{x}_1) \cdots \lambda_{\mathbb{S}^2}(d\,\mathbf{x}_n),$$

and so the n^{th}-order intensity function of X_{th} is,

$$\rho_{th}^{(n)}(\mathbf{x}_1, \ldots, \mathbf{x}_n) = p(\mathbf{x}_1) \cdots p(\mathbf{x}_n) \rho^{(n)}(\mathbf{x}_1, \ldots, \mathbf{x}_n).$$

Therefore

$$\frac{\rho_{th}^{(n)}(\mathbf{x}_1, \ldots, \mathbf{x}_n)}{\rho_{th}(\mathbf{x}_1) \cdots \rho_{th}(\mathbf{x}_n)} = \frac{\rho^{(n)}(\mathbf{x}_1, \ldots, \mathbf{x}_n)}{\rho(\mathbf{x}_1) \cdots \rho(\mathbf{x}_n)}.$$

So if X is IRWMI then X_{th} is also IRWMI by Definition 3.5.5 and if X is SOIWI then so is X_{th} when considering $n = 2$ by Definition 3.4.1. An immediate consequence of this is that the pair correlation functions for X and X_{th} are identical meaning that the inhomogeneous K-functions are also identical, whilst for inhomogeneous J-function an adjustment is needed for the infimum of the intensity function for the thinned process. In particular, let $\bar{\rho}_{th} = \inf_{\mathbf{x} \in \mathbb{S}^2} \rho_{th}(\mathbf{x})$ and suppose greater then 0, then by Definition 3.5.6 we have that,

$$J_{\text{inhom},th}(r) = 1 + \sum_{n=1}^{\infty} \frac{(-\bar{\rho}_{th})^n}{n!} J_n(t), \quad r \in [0, \pi]$$

where $J_n(t)$ is the same as that defined for X since the n^{th}-order correlation functions for X and X_{th} are identical.

In order to get the ratio representation of $J_{\text{inhom},th}$ we need to calculate $F_{\text{inhom},th}$ and $H_{\text{inhom},th}$: the inhomogeneous F and H functions for the thinned process. For $F_{\text{inhom},th}$ we will use the fact that $G_{X_{th}}(u) = G_X(up + 1 - p)$ where G_X and $G_{X_{th}}$ are the generating functionals for X and X_{th} respectively [Chiu et al., 1995, p. 147]. Defining $u_{r,th}^{\mathbf{y}}(\mathbf{x}) = (\bar{\rho}_{th} \mathbb{1}[\mathbf{x} \in B_{\mathbb{S}^2}(\mathbf{y}, r)])/(p(\mathbf{x})\rho(\mathbf{x}))$ we have,

$$\begin{aligned} 1 - F_{\text{inhom},th}(r) &= G_{X_{th}}(1 - u_{r,th}^{\mathbf{y}}) \\ &= G_X(1 - u_{r,th}^{\mathbf{o}} p) \\ &= \mathbb{E} \prod_{\mathbf{x} \in X} \left(1 - \frac{\bar{\rho}_{th} \mathbb{1}[\mathbf{x} \in B_{\mathbb{S}^2}(\mathbf{o}, r)]}{\rho(\mathbf{x})} \right) \end{aligned}$$

To calculate $H_{\text{inhom},th}$ for a location thinned process it can be shown that the reduced Palm process of X_{th} is identical to a location dependent thinning by the same rule Φ of the reduced Palm process of X [van Lieshout, 2011]. Starting from the reduced Campbell measure for the thinned process (see Definition 2.1.7), let $F \in \mathcal{N}_{lf}$ be an element of the sigma-algebra such that for any $B \in \mathcal{B}_0$, no elements of a point process lie in B, i.e. $X \in F$ if and only if $N_X(B) = 0$ then,

$$
\begin{aligned}
C^!_{th}(B,F) &= \mathbb{E} \sum_{\mathbf{x}\in X_{th}} \mathbb{1}[\mathbf{x} \in B, X_{th} \setminus \mathbf{x} \in F] \\
&= \mathbb{E} \sum_{\mathbf{x}\in X_{th}} \mathbb{1}[\mathbf{x} \in B, N_{X_{th}\setminus\mathbf{x}}(B) = 0] \\
&= \mathbb{E} \sum_{\mathbf{x}\in X_{th}} \mathbb{1}[\mathbf{x} \in B] \left(\prod_{\mathbf{y}\in X_{th}\setminus\mathbf{x}} \mathbb{1}[\mathbf{y} \notin B] \right) \\
&= \mathbb{E} \sum_{\mathbf{x}\in X} \mathbb{1}[\mathbf{x} \in B, R(\mathbf{x}) \le p(\mathbf{x})] \left(\prod_{\mathbf{y}\in X\setminus\mathbf{x}} \mathbb{1}[\mathbf{y} \notin B, R(\mathbf{y}) \le p(\mathbf{y})] \right) \\
&= \mathbb{E} \sum_{\mathbf{x}\in X} \mathbb{1}[\mathbf{x} \in B, R(\mathbf{x}) \le p(\mathbf{x})] \left(\left(\prod_{\mathbf{y}\in X\setminus\mathbf{x}} \mathbb{1}[\mathbf{y} \notin B] \right) \left(\prod_{\mathbf{y}\in X\setminus\mathbf{x}} \mathbb{1}[R(\mathbf{y}) \le p(\mathbf{y})] \right) \right) \\
&= \mathbb{E} \sum_{\mathbf{x}\in X} \mathbb{1}[\mathbf{x} \in B, R(\mathbf{x}) \le p(\mathbf{x}), N_{X\setminus\mathbf{x}}(B) = 0] \left(\prod_{\mathbf{y}\in X\setminus\mathbf{x}} \mathbb{1}[R(\mathbf{y}) \le p(\mathbf{y})] \right) \\
&= \int_B \mathbb{E} \left[\mathbb{1}[R(\mathbf{x}) \le p(\mathbf{x}), N_{X^!_{\mathbf{x}}}(B) = 0] \left(\prod_{\mathbf{y}\in X^!_{\mathbf{x}}} \mathbb{1}[R(\mathbf{y}) \le p(\mathbf{y})] \right) \right] \rho(\mathbf{x}) \lambda_{\mathbb{S}^2}(d\,\mathbf{x}) \\
&= \int_B \mathbb{E}[\mathbb{1}[R(\mathbf{x}) \le p(\mathbf{x})]] \, \mathbb{E} \left[\mathbb{1}[N_{X^!_{\mathbf{x}}}(B) = 0] \left(\prod_{\mathbf{y}\in X^!_{\mathbf{x}}} \mathbb{1}[R(\mathbf{y}) \le p(\mathbf{y})] \right) \right] \rho(\mathbf{x}) \lambda_{\mathbb{S}^2}(d\,\mathbf{x}) \\
&= \int_B \mathbb{E} \left[\mathbb{1}[N_{X^!_{\mathbf{x}}}(B) = 0] \left(\prod_{\mathbf{y}\in X^!_{\mathbf{x}}} \mathbb{1}[R(\mathbf{y}) \le p(\mathbf{y})] \right) \right] p(\mathbf{x}) \rho(\mathbf{x}) \lambda_{\mathbb{S}^2}(d\,\mathbf{x}) \\
&= \int_B \mathbb{E} \left[\mathbb{1}[N_{X^!_{\mathbf{x}}}(B) = 0] \left(\prod_{\mathbf{y}\in X^!_{\mathbf{x}}} \mathbb{1}[R(\mathbf{y}) \le p(\mathbf{y})] \right) \right] \rho_{th}(\mathbf{x}) \lambda_{\mathbb{S}^2}(d\,\mathbf{x}),
\end{aligned}
$$

where $\rho_{th}(\mathbf{x}) = p(\mathbf{x})\rho(\mathbf{x})$ is the intensity function of the thinned process. By uniqueness of the Radon-Nikodyn derivative $P^!_{X_{th}} = dC^!_{th}/d\mu_{th}$, where μ_{th} is the intensity measure of the

thinned process we have that,

$$P^!_{X_{th}}(F) = P(N_{X^!_{\mathbf{x}}}(B) = 0) = \mathbb{E}\left[\mathbb{1}[N_{X^!_{\mathbf{x}}}(B) = 0]\left(\prod_{\mathbf{y}\in X^!_{\mathbf{x}}} \mathbb{1}[R(\mathbf{y}) \leq p(\mathbf{y})]\right)\right],$$

where the right hand side is exactly the void probability associated with the location dependent thinning of $X^!_{\mathbf{x}}$ with retention probabilities $p(\mathbf{x})$. To see this consider the point process $Y^!_{\mathbf{x}}$ derived from $X^!_{\mathbf{x}}$ by thinning with retention probability $p(\mathbf{x})$ then,

$$\begin{aligned}
P(N_{Y^!_{\mathbf{x}}}(B) = 0) &= \mathbb{E}[\mathbb{1}[N_{Y^!_{\mathbf{x}}}(B) = 0]] \\
&= \mathbb{E}\prod_{\mathbf{y}\in Y^!_{\mathbf{x}}} \mathbb{1}[\mathbf{x} \notin B] \\
&= \mathbb{E}\prod_{\mathbf{y}\in X^!_{\mathbf{x}}} \mathbb{1}[\mathbf{x} \notin B, R(\mathbf{x}) \leq p(\mathbf{x})] \\
&= \mathbb{E}\left[\left(\prod_{\mathbf{y}\in X^!_{\mathbf{x}}} \mathbb{1}[\mathbf{x} \notin B]\right)\left(\prod_{\mathbf{y}\in X^!_{\mathbf{x}}} \mathbb{1}[R(\mathbf{y}) \leq p(\mathbf{y})]\right)\right] \\
&= \mathbb{E}\left[\mathbb{1}[N_{X^!_{\mathbf{x}}}(B) = 0]\left(\prod_{\mathbf{y}\in X^!_{\mathbf{x}}} \mathbb{1}[R(\mathbf{y}) \leq p(\mathbf{y})]\right)\right].
\end{aligned}$$

Thus since $P^!_{X_{th}}$ and $P_{Y^!_{\mathbf{x}}}$ are identical on the void probabilities and the fact that void probabilities uniquely characterise a point process this must mean that for any $F \in \mathcal{N}_{lf}$ $P^!_{X_{th}}(F) = P_{Y^!_{\mathbf{x}}}(F)$ and so reduced Palm process of X_{th} can be considered the thinning of the reduced Palm process of X by the same retention probability. Therefore, using the same result as for the generating functionals of thinned processes [Chiu et al., 1995, p. 147] we have,

$$1 - H_{\text{inhom},th}(r) = \mathbb{E}\prod_{\mathbf{x}\in X^!_{\mathbf{o}}}\left(1 - \frac{\bar{\rho}_{th}\mathbb{1}[\mathbf{x} \in B_{\mathbb{S}^2}(\mathbf{o}, r)]}{\rho(\mathbf{x})}\right).$$

Thus we can represent the inhomogeneous J-function as,

$$J_{\text{inhom},th}(r) = \frac{\mathbb{E}\prod_{\mathbf{x}\in X^!_{\mathbf{o}}}\left(1 - \frac{\bar{\rho}_{th}\mathbb{1}[\mathbf{x}\in B_{\mathbb{S}^2}(\mathbf{o},r)]}{\rho(\mathbf{x})}\right)}{\mathbb{E}\prod_{\mathbf{x}\in X}\left(1 - \frac{\bar{\rho}_{th}\mathbb{1}[\mathbf{x}\in B_{\mathbb{S}^2}(\mathbf{o},r)]}{\rho(\mathbf{x})}\right)},$$

and if X, the unthinned process, is isotropic then $\bar{\rho}_{th} = \bar{p}\rho$, where $\bar{p} = \inf_{\mathbf{x}\in\mathbb{S}^2}$ and ρ is the

constant intensity. Thus

$$J_{\text{inhom},th}(r) = \frac{\mathbb{E}\left[(1-\bar{p})^{N_{X^!_{\mathbf{o}}}(B_{\mathbb{S}^2}(\mathbf{o},r))}\right]}{\mathbb{E}\left[(1-\bar{p})^{N_X(B_{\mathbb{S}^2}(\mathbf{o},r))}\right]}.$$

To exemplify functional summary statistics for thinned point processes on $\mathbb{S}^2$ we will consider the so-called *Matérn* I and II hardcore processes, first introduced Bertil Matérn in his seminal work Matérn [2013] and later extended by Stoyan and Stoyan [1985]. These processes both arise from dependent thinning of homogeneous Poisson processes. They are defined as follows on $\mathbb{S}^2$.

Definition 3.6.3. *Let X be a homogeneous Poisson process on $\mathbb{S}^2$ with intensity $\rho \in \mathbb{R}_+$. Fix $R \in [0,\pi]$, 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 $d_{\mathbb{S}^2}(\mathbf{x},\mathbf{y}) < R$, otherwise retain $\mathbf{x}$. The resulting thinned process is then defined as a Matérn I inhibition process on $\mathbb{S}^2$.*

Definition 3.6.4. *Let X be a homogeneous Poisson process on $\mathbb{S}^2$ with intensity $\rho \in \mathbb{R}_+$. Fix $R \in [0,\pi]$, 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 points 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 $d_{\mathbb{S}^2}(\mathbf{x},\mathbf{y}) < R$ 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 $\mathbb{S}^2$.*

Since a homogeneous Poisson process is isotropic over $\mathbb{S}^2$ and the thinning rule is not a function of location the resulting Matérn I and II processes are also isotropic [Teichmann et al., 2013]. The following proposition, which is Proposition 3.9 and 3.11 of Lawrence [2018], gives the intensity functions for the Matérn processes.

Proposition 3.6.5. *The spheroidal Matérn I and II processes with hardcore distance $R \in [0,\pi]$, underlying isotropic Poisson process with intensity function $\rho > 0$ and independent mark random variable $M_{\mathbf{x}}$ for the latter have constant intensity functions*

$$\rho_1 = \rho \exp(-\rho 2\pi(1-\cos(R)))$$
$$\rho_2 = \frac{1-\exp(-\rho 2\pi(1-\cos(R)))}{2\pi(1-cos(R))}$$

respectively.

Proof. This follows directly from Proposition 4.6.4 when $\mathbb{D} = \mathbb{S}^2$. Also see Proposition 3.9

and 3.11 of Lawrence [2018]. □

Since the Matérn processes are isotropic then they are also IRWMI and SOIRWI, thus we can independently thinning these processes with some set of retention probabilities $p : \mathbb{S}^2 \mapsto [0, 1]$ and the new process will have well defined inhomogeneous K, F, H, and J-functions.

In Figure 3.5 we consider examples of Matérn I and II processes that have been independently thinned with retention probabilities,

$$p(\mathbf{x}) = \frac{2}{5}\left(\sin(-\pi x) + \frac{3}{2}\right), \quad \mathbf{x} = (x, y, z)^T \in \mathbb{S}^2.$$

We simulate both underlying Matérn processes with hardcore parameter $R = 0.07$ and expected number of points across the sphere being $E[N_X(\mathbb{S}^2)] = 300$; that is we set the intensity, ρ, of the underlying isotropic Poisson point process set such that $4\pi\rho_i = 300$ for $i = 1, 2$ where ρ_i is taken from Proposition 3.6.5. Whilst for the Matérn II process we simulate each mark independently with a unit rate exponential distribution. By isotropy its intensity function is constant and $300/(4\pi)$. This means that the intensity functions of both the Matérn processes are identical and given by

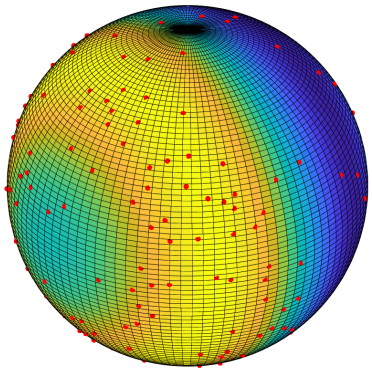
$$\rho(\mathbf{x}) = \frac{300}{4\pi} \cdot \frac{2}{5}\left(\sin(-\pi x) + \frac{3}{2}\right). \tag{3.18}$$

This intensity function is plotted over the sphere in Figure 3.5a and 3.5d where large intensities are indicated by yellow regions and low intensities by blue regions.

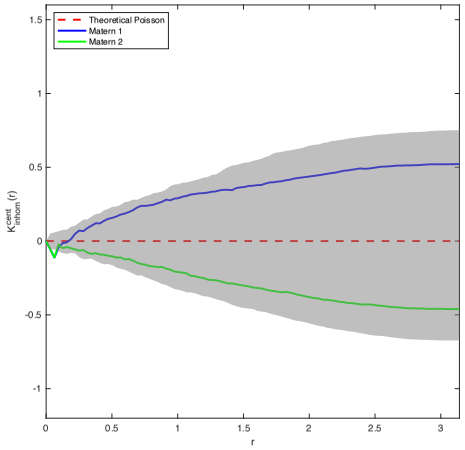
As expected due to the hardcore distance imposed by both processes we see this highlighted in both Figures 3.5b and 3.5e where there is a steep descent for the former whilst a flat line for the latter until, approximately, the hardcore distance of $R = 0.07$ is reached. This also translate to the plot of the inhomogeneous J-function in Figure 3.5f where in the enlarged portion the estimates for the Matérn I and II are non-decreasing functions until the hardcore distance is reached.

3.7 Discussion

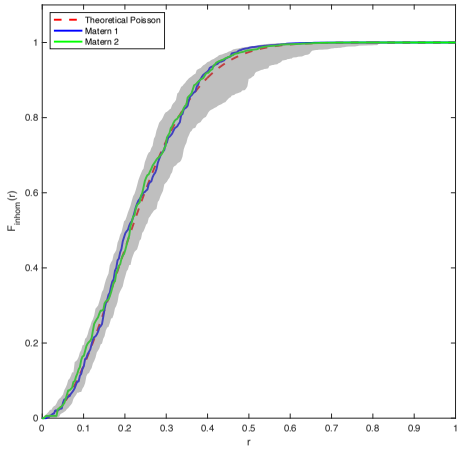
Spheroidal point process theory remains a relatively new area of research and remains relatively underexplored. We began with a literature review discussing contributions made by Robeson et al. [2014], Møller and Rubak [2016], Lawrence et al. [2016] which introduce key definitions for spheroidal processes and accompanying functional summary statistics, in par-



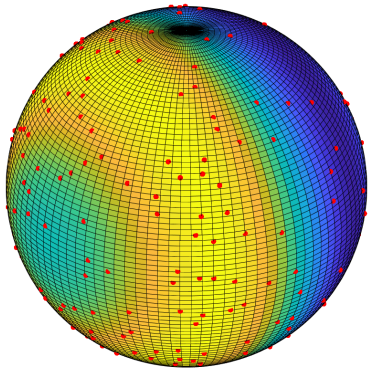
(a) Realisation of an independently thinned Matérn I process



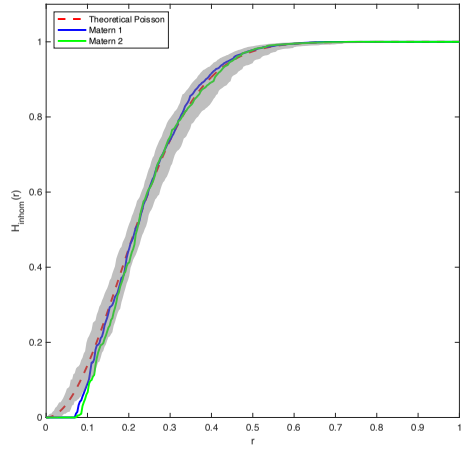
(b) Plot of $K^{\text{cent}}_{\text{inhom}}(r)$



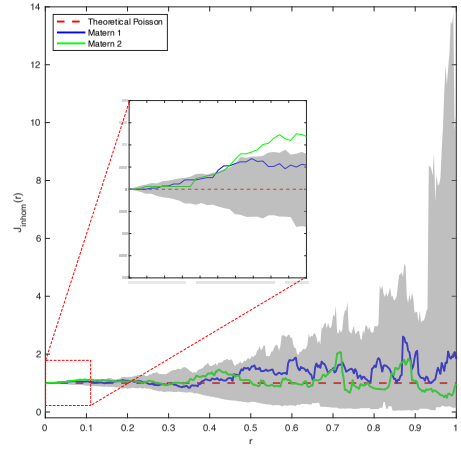
(c) Plot of $F_{\text{inhom}}(r)$



(d) Realisation of an independently thinned Matérn II process



(e) Plot of $H_{\text{inhom}}(r)$



(f) Plot of $J_{\text{inhom}}(r)$

Figure 3.5: Examples of inhomogeneous functional summary statistics for location thinned Matérn I and II processes with intensity functions are given by Equation 3.18. The simulation envelopes are constructed from 99 inhomogeneous Poisson processes with intensity function given by Equation 3.18.

ticular Møller and Rubak [2016], Lawrence et al. [2016] both introduce the inhomogeneous K-function. We provided extensions of the inhomogeneous F, H, and J-functions from $\mathbb{R}^{2,3}$ to $\mathbb{S}^2$ [van Lieshout, 2011] and demonstrated these on various spheroidal point processes. The functional summary statistics developed here come to play key roles in Chapters 3 and 4 where they are used to determine if a point pattern is CSR or not based on simulation envelopes.

4

Summary statistics for Poisson processes on convex shapes

This chapter discusses our first major contribution made to the area of spatial statistics based on our work Ward et al. [2021b]. In this chapter we discuss the difficulties imposed when observing point processes outside of symmetric spaces such as $\mathbb{R}^d$ and $\mathbb{S}^{d-1}$. With a focus on $d = 3$, we show that Poisson processes on bounded convex shapes can be mappped to a process on $\mathbb{S}^2$ where rotational symmetries provide us the flexibility to construct well-defined functional summary statistics. We demonstrate these using simulation envelopes for Poisson processes on cubes and ellipsoid. Using these functional summary statistics we are also able to examine whether an observed point pattern exhibits evidence of aggregation and/or repulsion and explore this through simulated examples of Thomas and Matérn processes on ellipsoids.

4.1 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 $\mathbf{0} = (0, 0, 0)^T$. We equip $\mathbb{R}^3$ with its Lebesgue measure $\lambda_{\mathbb{R}^3}$.

Define $\bar{\mathbb{D}}$ be a connected set of $\mathbb{R}^3$, that is $\bar{\mathbb{D}}$ cannot be divided into two distinct non-empty open sets, and $\lambda_{\mathbb{R}^3}(\mathbb{D})$ finite. We also suppose that $\bar{\mathbb{D}}$ is compact (i.e. closed and bounded). The set $\bar{\mathbb{D}}$ is said to be convex if and only if for all $\mathbf{x}, \mathbf{y} \in \bar{\mathbb{D}}$ such that $\mathbf{x} \neq \mathbf{y}$ then $\{\mathbf{z} \in \mathbb{R}^3 : \mathbf{z} = \mathbf{x} + \gamma(\mathbf{y} - \mathbf{x}),\ \gamma \in (0, 1)\} \in \bar{\mathbb{D}}$. Let $\mathbb{D}$ be the boundary of $\bar{\mathbb{D}}$ and we say that $\mathbb{D}$ is convex if $\bar{\mathbb{D}}$ is convex. Examples of bounded convex surfaces of $\mathbb{R}^3$ are spheres, ellipsoids,

and cubes. Throughout this chapter we shall focus on such convex surfaces where $\mathbb{D}$ will be used to denote the boundary of a convex subset of $\mathbb{R}^3$.

Further for any bounded convex set $\mathbb{D}$ we suppose that the surface can be parametrised. More precisely, suppose that $\mathbb{D}$ can be partitioned into $p > 0$ sets, i.e. there exists sets $\mathbb{D}_i$ such that $\cup_{i=1}^{p} \mathbb{D}_i = \mathbb{D}$ and $\mathbb{D}_i \neq \mathbb{D}_j$ for $i \neq j$. Then for each $\mathbb{D}_i$ there exists $\tilde{g}_i$,

$$x_o = \tilde{g}_i(x_m, x_n) \text{ such that } (x_r, x_s, x_t)^T \in \mathbb{D}_i \tag{4.1}$$

for some $m, n, o \in \{1, 2, 3\}, m \neq n \neq o \neq m$ and $r = \min(m, n, o), t = \max(m, n, o), s \in \{m, n, o\} \setminus \{r, t\}$. For example take the case of a sphere with radius 1 centred at the origin, then for the upper hemisphere we can parametrise x_3 as $x_3 = \tilde{g}(x_1, x_2) = (1 - x_1^2 - x_2^2)^{1/2}$ whilst for the lower hemisphere we can parametrise x_3 as $x_3 = \tilde{g}(x_1, x_2) = -(1 - x_1^2 - x_2^2)^{1/2}$. For any bounded convex sets, $\mathbb{D}$, we also define its geodesic as a 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 $d_{\mathbb{D}} : \mathbb{D} \times \mathbb{D} \mapsto \mathbb{R}_+$, where $\mathbb{R}_+$ is the positive real line including 0, thus $(\mathbb{D}, d_{\mathbb{D}})$ is a metric space. Additionally, we will frequently need to evaluate integrals over $\mathbb{D}$, which can be done using the infinitesimal area element over each of the $\mathbb{D}_i$ defined as,

$$d\mathbb{D}_i = \sqrt{1 + \left(\frac{\partial \tilde{g}_i}{\partial x_m}\right)^2 + \left(\frac{\partial \tilde{g}_i}{\partial x_n}\right)^2} dx_m dx_n,$$

where $\tilde{g}_i$ is differentiable for all $i = 1, \ldots, p$. We assume that these convex subsets of $\mathbb{R}^3$ are defined such that the origin is inside $\mathbb{D}$, that is $\mathbf{0} \in \bar{\mathbb{D}}$, 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 $\mathbb{D}$.

4.2 Defining summary statistics 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 Equations (3.1)-(3.4) 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 (3.1)-(3.4) 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 $\mathbb{D}$. Using instead Definition 3.4.1 (see also Møller and Waagepetersen [2003, Definition 4.5, p. 32]) to define SOIRWI we notice that Equation (3.5) 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 (3.5). This is identically the inhomogeneous K-function. The indicator function of (3.5) 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 $\mathbb{D}$.

4.3 Mapping from $\mathbb{D}$ to $\mathbb{S}^2$

Having discussed the impracticalities of defining functional summary statistics directly on $\mathbb{D}$ we consider an alternative construction of 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 $\mathbb{S}^2$.

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 4.3.2 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 4.3.1 shows that this function is bijective and hence measurable.

Lemma 4.3.1. *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. $\mathbf{o} \in \bar{\mathbb{D}}$. Then the function $f(\mathbf{x}) = \mathbf{x}/||\mathbf{x}||, f : \mathbb{D} \mapsto \mathbb{S}^2$ is bijective.*

Proof. See Appendix B.1. □

Rather than using the Mapping Theorem [Kingman, 1993], we utilise Proposition 3.1 of Møller and Waagepetersen [2003] 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 4.3.2. *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 $\mathbb{D}$ can be partitioned into $p > 0$ sets $\mathbb{D}_i$ each with an associated parametrisation $x_{o_i} = \tilde{g}_i(x_{m_i}, x_{n_i})$ for some $m_i, n_i, o_i \in \{1, 2, 3\}, m_i \neq n_i \neq o_i \neq m_i$ as defined by Equation 4.1. We assume for all $i = 1, \ldots, p$ that $\tilde{g}_i$ are differentiable. 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,*

$$\rho^*(\mathbf{x}) = \begin{cases} \rho(f^{-1}(\mathbf{x}))l_1(f^{-1}(\mathbf{x}))J_{(1,f^*)}(\mathbf{x})\sqrt{1 - x_{m_1}^2 - x_{n_1}^2}, & \mathbf{x} = (x_{r_1}, x_{s_1}, x_{t_1})^T \in f(\mathbb{D}_1) \\ \vdots & \\ \rho(f^{-1}(\mathbf{x}))l_p(f^{-1}(\mathbf{x}))J_{(p,f^*)}(\mathbf{x})\sqrt{1 - x_{m_p}^2 - x_{n_p}^2}, & \mathbf{x} = (x_{r_p}, x_{s_p}, x_{t_p})^T \in f(\mathbb{D}_p) \end{cases} \tag{4.2}$$

where,

$$x_{o_i} = \tilde{g}_i(x_{m_i}, x_{n_i})$$
$$r_i = \min(m_i, n_i, o_i), \quad t_i = \max(m_i, n_i, o_i), \quad s_i \in \{m_i, n_i, o_i\} \setminus \{r_i, t_i\}$$
$$l_i(\mathbf{x}) = \left[1 + \left(\frac{\partial \tilde{g}_i}{\partial x_{m_i}}\right)^2 + \left(\frac{\partial \tilde{g}_i}{\partial x_{n_i}}\right)^2\right]^{\frac{1}{2}}$$
$$J_{(i,f^{*-1})}(\mathbf{x}) = \frac{1}{(x_{m_i}^2 + x_{n_i}^2 + \tilde{g}_i^2(x_{m_i}, x_{n_i}))^3}$$
$$\det\left[\begin{pmatrix} x_{n_i}^2 + \tilde{g}_i^2(x_{m_i}, x_{n_i}) - x_{m_i}\tilde{g}_i(x_{m_i}, x_{n_i})\frac{\partial \tilde{g}_i}{\partial x_{m_i}} & -x_{m_i}\left(x_{n_i} + \tilde{g}_i(x_{m_i}, x_{n_i})\frac{\partial \tilde{g}_i}{\partial x_{n_i}}\right) \\ -x_{n_i}\left(x_{m_i} + \tilde{g}_i(x_{m_i}, x_{n_i})\frac{\partial \tilde{g}_i}{\partial x_{m_i}}\right) & x_{m_i}^2 + \tilde{g}_i^2(x_{m_i}, x_{n_i}) - x_{n_i}\tilde{g}_i(x_{m_i}, x_{n_i})\frac{\partial \tilde{g}_i}{\partial x_{n_i}} \end{pmatrix}\right]$$
$$J_{(i,f^*)}(\mathbf{x}) = \frac{1}{J_{(i,f^{*-1})}(f^{-1}(\mathbf{x}))},$$

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

$$f^*(x_{m_i}, x_{n_i}) = \left(\frac{x_{m_i}}{||(x_{r_i}, x_{s_i}, x_{t_i})||}, \frac{x_{n_i}}{||(x_{r_i}, x_{s_i}, x_{t_i})||} \right).$$

Proof. See Appendix B.2. □

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 4.3.3. *A notion of bijectivety 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 4.3.1, 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 4.3.2 if X, Y are Poisson processes on $\mathbb{D}$ with intensity function ρ_X and ρ_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 $\mathbb{S}^2$. This is the result of the following corollary.*

Corollary 4.3.4. *Suppose that X_1 and X_2 are two Poisson processes on $\mathbb{D}$ with intensity functions ρ_1 and ρ_2 respectively, such that $\rho_1(\mathbf{x}) \neq \rho_2(\mathbf{x})$ for $\mathbf{x} \in B \subseteq \mathbb{D}$ such that $\lambda_{\mathbb{D}}(B) > 0$. Define the transformed processes $Y_1 = f(X_1)$ and $Y_2 = f(X_2)$ from $\mathbb{D}$ to $\mathbb{S}^2$ where $f(\mathbf{x}) = \mathbf{x}/||\mathbf{x}||$. Then the Y_1 and Y_2 are Poisson processes with intensity functions ρ_1^* and ρ_2^* respectively such that $\rho_1^*(\mathbf{x}) \neq \rho_2^*(\mathbf{x})$ for $\mathbf{x}^* \in B^*$ where $B^* = f(B) \subseteq \mathbb{S}^2$.*

Proof. See Appendix B.3. □

Remark 4.3.5. *Further, another useful result which follows directly from Theorem 4.3.2 is the construction of approximate Poisson processes on $\mathbb{S}^2$. Suppose that we do not have a parametrisation of the surface or the parametrisation is complex then we can approximate the space, by considering a finite piecewise planar approximation of $\mathbb{D}$. For example this piecewise planar approximation could be constructed by considering the boundary of the convex hull of*

a finite set of points $U \subset \mathbb{D}$, denote this approximation $\tilde{\mathbb{D}}$. Then each face of the $\tilde{\mathbb{D}}$ is a plane for which $\tilde{g}_i$ exists. We can then use this approximation of $\mathbb{D}$ to map a Poisson process on $\mathbb{D}$ to $\mathbb{S}^2$.

4.4 Constructing 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_{inhom} (see Equations 3.10-3.12 and 3.7 respectively) [van Lieshout, 2011, Møller and Rubak, 2016] can be combined with the mapped intensity function from Theorem 4.3.2 to construct estimators as follows,

$$\hat{F}_{\text{inhom},\mathbb{D}}(r) = 1 - \frac{\sum_{\mathbf{p}\in P} \prod_{\mathbf{x}\in Y\cap B_{\mathbb{S}^2}(\mathbf{p},r)} \left(1 - \frac{\bar{\rho^*}}{\rho^*(\mathbf{x})}\right)}{|P|} \tag{4.3}$$

$$\hat{H}_{\text{inhom},\mathbb{D}}(r) = 1 - \frac{\sum_{\mathbf{x}\in Y} \prod_{\mathbf{y}\in(Y\setminus\{\mathbf{x}\})\cap B_{\mathbb{S}^2}(\mathbf{x},r)} \left(1 - \frac{\bar{\rho^*}}{\rho^*(\mathbf{y})}\right)}{N_Y(\mathbb{S}^2)} \tag{4.4}$$

$$\hat{J}_{\text{inhom},\mathbb{D}}(r) = \frac{1 - \hat{H}_{\text{inhom},\mathbb{D}}(r)}{1 - \hat{F}_{\text{inhom},\mathbb{D}}(r)} \tag{4.5}$$

$$\hat{K}_{\text{inhom},\mathbb{D}}(r) = \frac{1}{4\pi} \sum_{\mathbf{x},\mathbf{y}\in Y}^{\neq} \frac{\mathbb{1}[d(\mathbf{x},\mathbf{y}) \leq r]}{\rho^*(\mathbf{x})\rho^*(\mathbf{y})}, \tag{4.6}$$

where P is a finite grid on $\mathbb{S}^2$, X is a Poisson process on $\mathbb{D}$ with intensity function ρ, $Y = f(X)$ is the mapped Poisson process onto $\mathbb{S}^2$, ρ^* is given by (4.2) 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 ρ^* is unknown, nonparametric plug-in estimates of ρ^* can be constructed [Lawrence et al., 2016, Møller and Rubak, 2016]. In Chapter 6 we consider this problem when the underlying space is a Riemannian manifold.

4.4.1 Properties of functional summary statistics

Consider the general case of all Poisson processes on $\mathbb{S}^2$. Theorem 4.4.1 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 4.4.1. *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 $\hat{F}_{inhom}(r)$, and $\hat{K}_{inhom}(r)$ are unbiased whilst $\hat{H}_{inhom}(r)$ is ratio-unbiased. More precisely,*

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

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 $\hat{F}_{inhom}(r)$ and $\hat{H}_{inhom}(r)$, respectively, we immediately have ratio-unbiasedness of $\hat{J}_{inhom}(r)$.

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 Appendix B.4. □

Theorem 4.4.1 shows that $\hat{H}_{\text{inhom}}(r)$ is a biased estimator for $H_{\text{inhom}}(r)$. Although biased it can be bounded.

Corollary 4.4.2. *With the same assumptions as Theorem 4.4.1, 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 $\hat{H}_{inhom}(r)$ is bounded by*

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

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

Proof. See Appendix B.5. □

Corollary 4.4.2 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 points of 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 4.4.3. *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 $\hat{K}_{\text{inhom}}(r)$, $\hat{F}_{\text{inhom}}(r)$, and*

$\hat{H}_{\text{inhom}}(r)$ *have variance,*

$$\begin{aligned}
\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}) + \\
&\quad (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\left(-2\bar{\rho}\lambda_{\mathbb{S}^2}(B_{\mathbb{S}^2}(\mathbf{o},r))\right)}{|P|^2} \\
&\quad \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) \\
&\quad -\exp\left(-2\bar{\rho}\lambda_{\mathbb{S}^2}(B_{\mathbb{S}^2}(\mathbf{o},r))\right), \\
&\text{Var}(\hat{H}_{\text{inhom}}(r)) \\
&= \frac{1}{\mu^2(\mathbb{S}^2)}\int_{\mathbb{S}^2}\int_{\mathbb{S}^2}\left(\rho(\mathbf{x})-\bar{\rho}\mathbb{1}[\mathbf{x}\in B_{\mathbb{S}^2}(\mathbf{y},r)]\right)\left(\rho(\mathbf{y})-\bar{\rho}\mathbb{1}[\mathbf{y}\in B_{\mathbb{S}^2}(\mathbf{x},r)]\right) \\
&\quad \frac{\mathrm{e}^{-\mu(\mathbb{S}^2)}}{A_1^2(\mathbf{x},\mathbf{y})}\left(\mathrm{e}^{\mu(\mathbb{S}^2)A_1(\mathbf{x},\mathbf{y})}-1-\text{Ei}(\mu(\mathbb{S}^2)\text{A}_1(\mathbf{x},\mathbf{y}))+\gamma+\log(\mu(\mathbb{S}^2)\text{A}_1(\mathbf{x},\mathbf{y}))\right) \\
&\quad \lambda_{\mathbb{S}^2}(d\mathbf{x})\lambda_{\mathbb{S}^2}(d\mathbf{y}) \\
&\quad + \frac{1}{\mu(\mathbb{S}^2)}\int_{\mathbb{S}^2}\frac{\mathrm{e}^{-\mu(\mathbb{S}^2)}}{A_2(\mathbf{x})}\left(\gamma+\log(\mu(\mathbb{S}^2)A_2(\mathbf{x}))-\text{Ei}(\mu(\mathbb{S}^2)\text{A}_2(\mathbf{x}))\right)\rho(\mathbf{y})\lambda_{\mathbb{S}^2}(d\mathbf{y}) \\
&\quad - \frac{\mathrm{e}^{-2\mu(\mathbb{S}^2)}}{\left(1-\frac{\bar{\rho}}{\mu(\mathbb{S}^2)}2\pi(1-\cos r)\right)^2}\left(\mathrm{e}^{\mu(\mathbb{S}^2)\left(1-\frac{\bar{\rho}}{\mu(\mathbb{S}^2)}2\pi(1-\cos r)\right)}-1\right)^2
\end{aligned}$$

where,

$$\begin{aligned}
A_1(\mathbf{x},\mathbf{y}) &= 1-\frac{2\bar{\rho}}{\mu(\mathbb{S}^2)}2\pi(1-\cos r)+\frac{\bar{\rho}^2}{\mu(\mathbb{S}^2)}\int_{B_{\mathbb{S}^2}(\mathbf{x},r)\cap B_{\mathbb{S}^2}(\mathbf{y},r)}\frac{1}{\rho(\mathbf{z})}\lambda_{\mathbb{S}^2}(d\mathbf{z}) \\
A_2(\mathbf{x}) &= 1-\frac{2\bar{\rho}}{\mu(\mathbb{S}^2)}2\pi(1-\cos r)+\frac{\bar{\rho}^2}{\mu(\mathbb{S}^2)}\int_{B_{\mathbb{S}^2}(\mathbf{x},r)}\frac{1}{\rho(\mathbf{y})}\lambda_{\mathbb{S}^2}(d\mathbf{y}) \\
\text{Ei(x)} &= -\int_{-x}^{\infty}\frac{\mathrm{e}^{-t}}{t}dt
\end{aligned}$$

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

Proof. See Appendix B.6. □

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 4.4.4. *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 $r_{\max} = \sup\{r \in [0, \pi] :$ there exists $\mathbf{p} \in P$ such that $\rho(\mathbf{x}) \neq \bar{\rho}$ 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 Appendix B.7. □

Proposition 4.4.5. *Let X be a spheroidal Poisson process with known intensity function $\rho : \mathbb{S}^2 \mapsto \mathbb{R}$. Then the covariance between $1 - \hat{H}_{inhom}(r)$ and $1 - \hat{F}_{inhom}(r)$ for $r \in [0, \pi]$ is,*

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

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

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

Proof. See Appendix B.8. □

Using a Taylor series expansion (see Appendix B.9), we can approximate the expectation and variance of $\hat{J}_{\text{inhom}}(r)$ as

$$
\mathbb{E}\left[\frac{X}{Y}\right] \approx \frac{\mu_X}{\mu_Y} - \frac{\text{Cov}(X, Y)}{\mu_Y^2} + \frac{\text{Var}(Y)\mu_X}{\mu_Y^3} \tag{4.7}
$$

$$
\text{Var}\left(\frac{X}{Y}\right) \approx \frac{\mu_X}{\mu_Y}\left[\frac{\text{Var}(X)}{\mu_X^2} - 2\frac{\text{Cov}(X, Y)}{\mu_X\mu_Y} + \frac{\text{Var}(Y)}{\mu_Y^2}\right], \tag{4.8}
$$

where $X = 1 - \hat{H}_{\text{inhom}}(r)$ and $Y = 1 - \hat{F}_{\text{inhom}}(r)$. The terms in Equations (4.7) and (4.8) are given in Theorems 4.4.1 and 4.4.3, and Proposition 4.4.5.

4.5 Examples

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

4.5.1 Cube

We define a centred cube over each of the six faces with a side length $2l$, where $l = 1$. We have the six following equations parametrising each face,

$$\begin{aligned}
x_3 \equiv \tilde{g}_1(x_1, x_2) &= -l, \quad \text{for } -l \leq x_1, x_2 \leq l \\
x_3 \equiv \tilde{g}_2(x_1, x_2) &= l, \quad \text{for } -l \leq x_1, x_2 \leq l \\
x_2 \equiv \tilde{g}_3(x_1, x_3) &= -l, \quad \text{for } -l \leq x_1, x_3 \leq l \\
x_2 \equiv \tilde{g}_4(x_1, x_3) &= l, \quad \text{for } -l \leq x_1, x_3 \leq l \\
x_1 \equiv \tilde{g}_5(x_2, x_3) &= -l, \quad \text{for } -l \leq x_2, x_3 \leq l \\
x_1 \equiv \tilde{g}_6(x_2, x_3) &= l, \quad \text{for } -l \leq x_2, x_3 \leq l.
\end{aligned}$$

Using Theorem 4.3.2 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. $x_3 = -1$, and in the notation of Theorem 4.3.2 label this $\mathbb{D}_1$. Then,

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

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

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

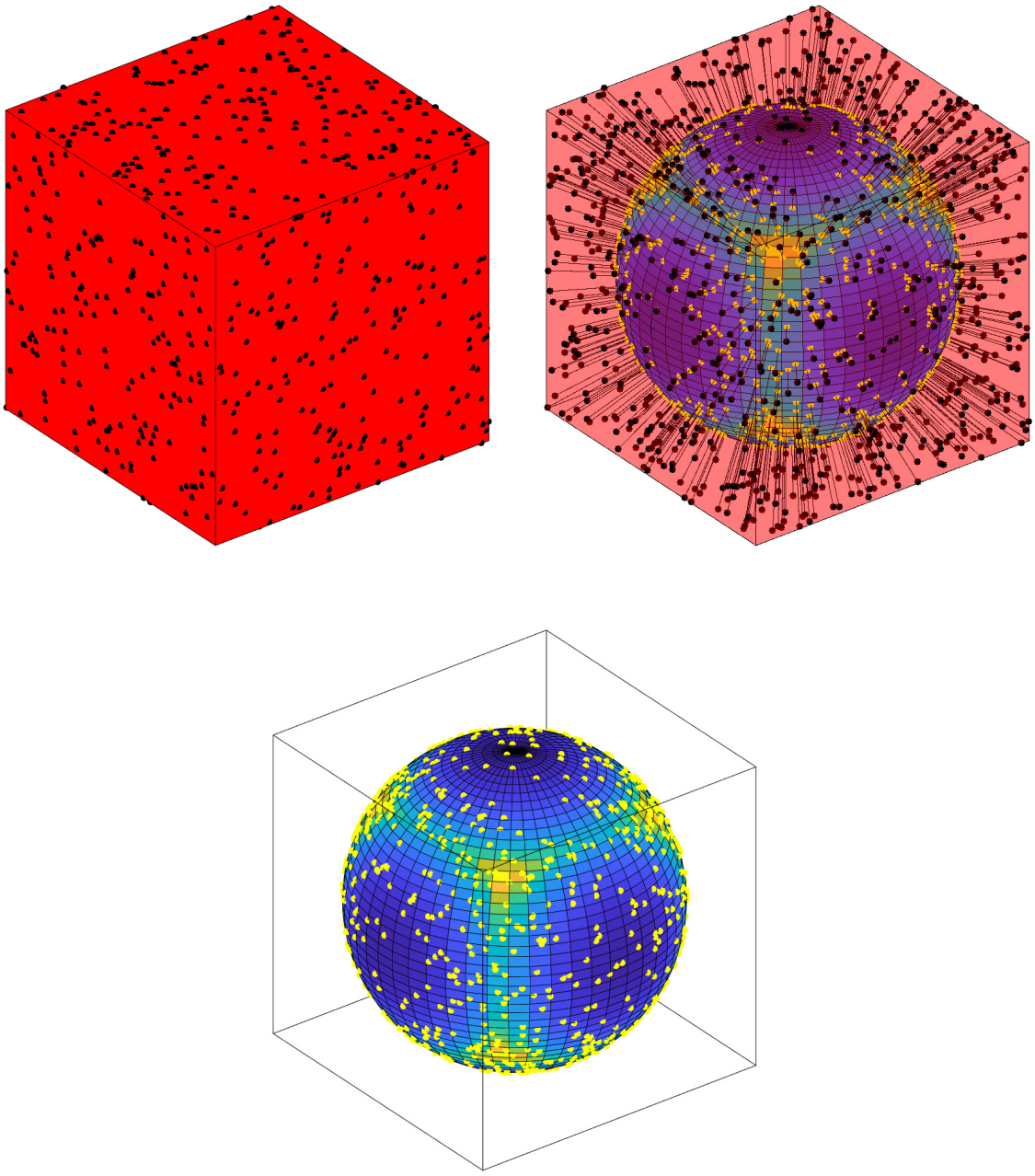


Figure 4.1: Example of simulating and mapping a CSR process on the cube to the sphere. *Top left:* example of a CSR process on a cube with $l = 1$ and constant intensity function 50. *Top right:* mapping of points from cube to the sphere by the function $f(\mathbf{x}) = \mathbf{x}/||\mathbf{x}||$. *Bottom:* 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.

thus by the appropriate rotations the intensity function over the entire sphere is,

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

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. Figure 4.1 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 when $x_1 = x_2 = 0$, giving $\inf_{\mathbf{x}\in\mathbb{S}^2} \rho^*(\mathbf{x}) = \rho$. Figure 4.2 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.

4.5.2 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. We can parametrise an ellipsoid separately on the upper and lower hemiellipsoids as,

$$x_3 \equiv \tilde{g}_+(x_1, x_2) = c\left(1 - \frac{x_1^2}{a^2} - \frac{x_2^2}{b^2}\right)^{1/2}, \quad \text{for} \quad \frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} \leq 1$$

$$x_3 \equiv \tilde{g}_+(x_1, x_2) = -c\left(1 - \frac{x_1^2}{a^2} - \frac{x_2^2}{b^2}\right)^{1/2}, \quad \text{for} \quad \frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} \leq 1.$$

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 $f(\mathbf{x}) =$

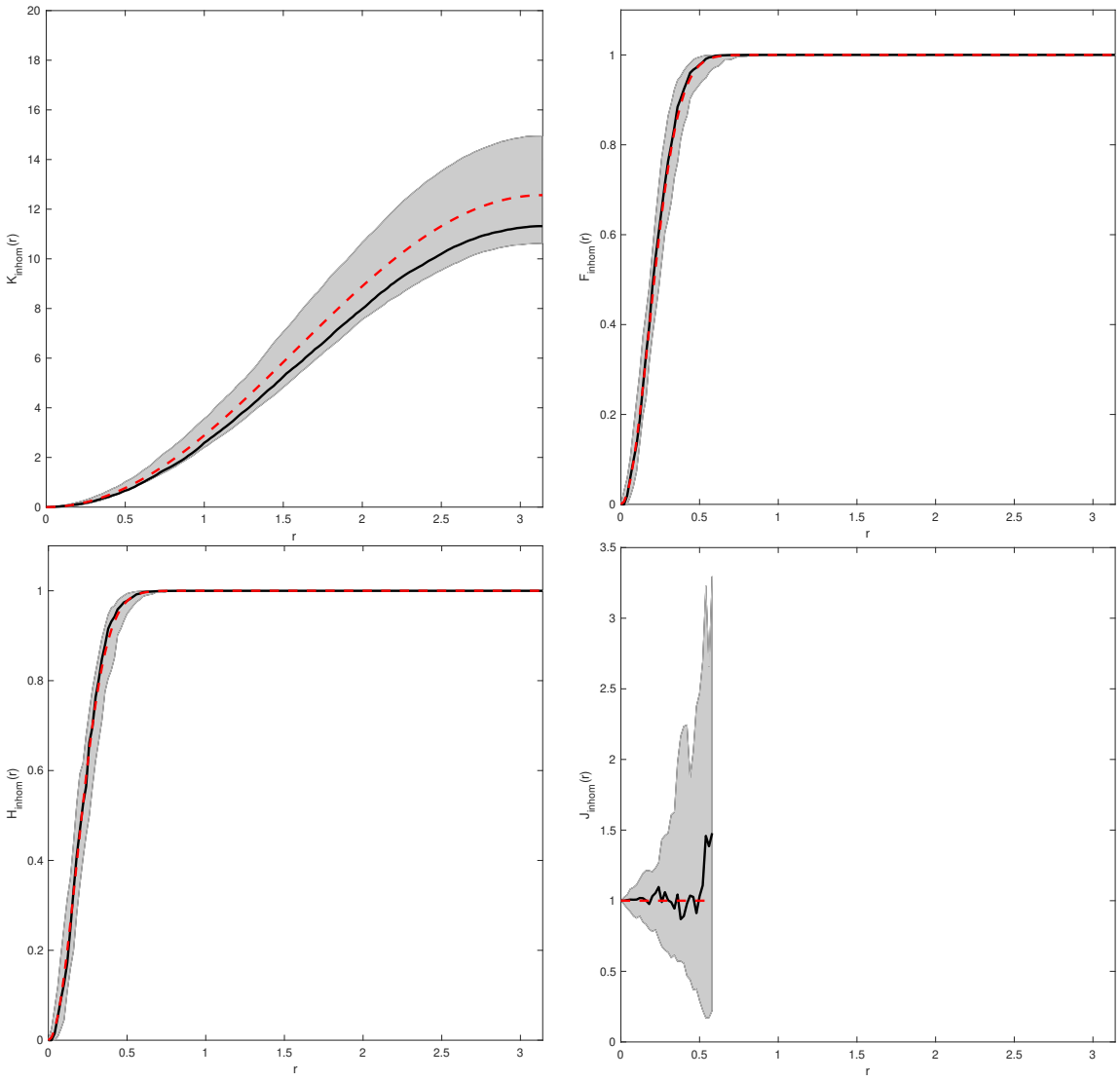


Figure 4.2: Examples of K_{inhom}- (*top left*), F_{inhom}- (*top right*), H_{inhom}- (*bottom left*), and J_{inhom}- (*bottom right*) functions for CSR patterns on a cube with $l = 1$ and $\rho = 5$. Solid line is the estimated functional summary statistic for our observed data, dashed line is the theoretical 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.

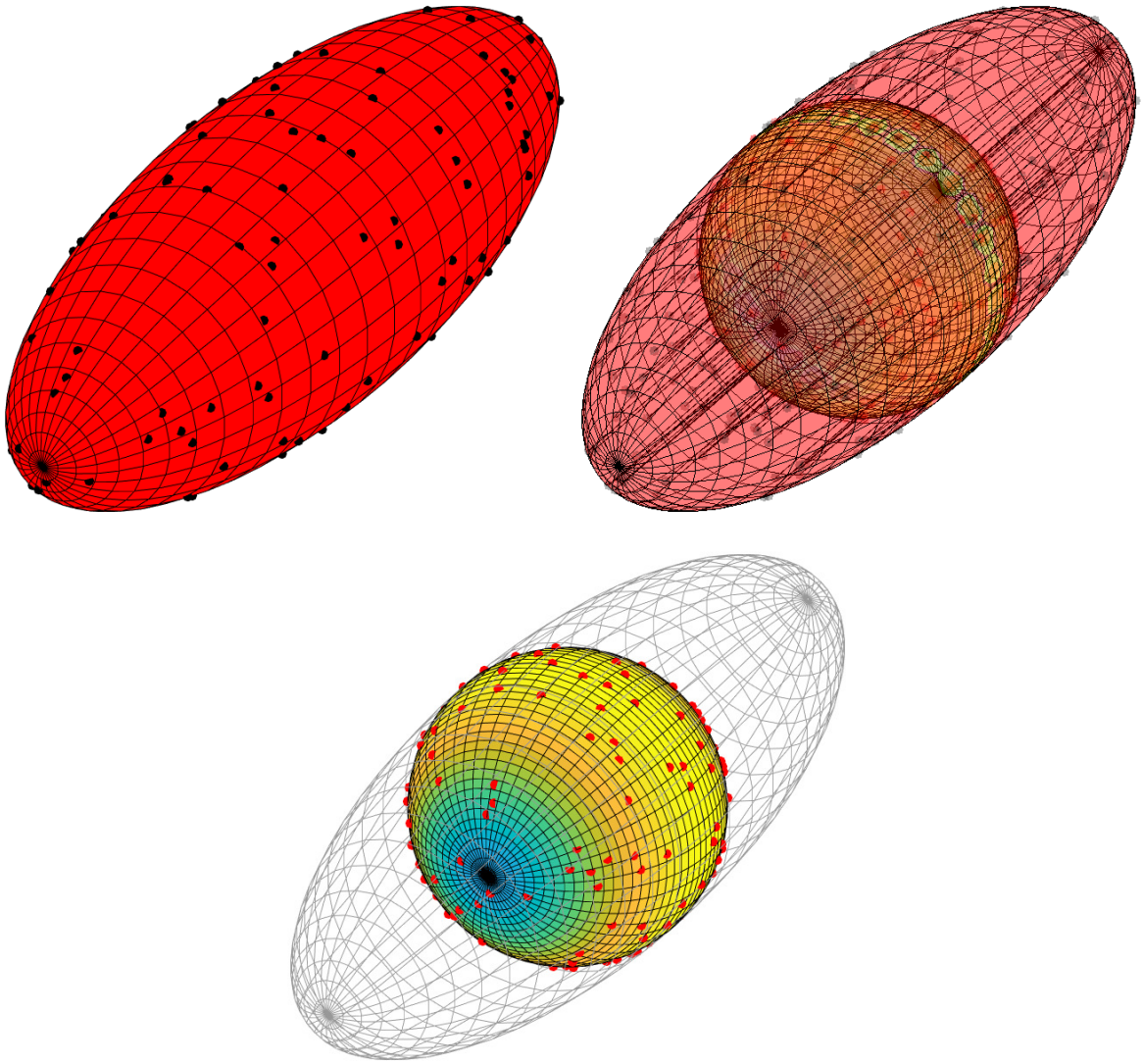


Figure 4.3: Example of simulating and mapping a CSR process on a prolate ellipsoid to the sphere. *Top left:* example of a CSR process on a prolate ellipsoid with $a = b = 1, c = 3$ and $\rho = 5$. *Top right:* mapping of points from prolate ellipsoid to the sphere by the function $f(\mathbf{x}) = (x_1/a, x_2/b, x_3/c)^T$. *Bottom:* 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.

$\mathbf{x}/||\mathbf{x}||$ 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 4.3.2 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

$$\rho^*(\mathbf{x}) = \rho ab \sqrt{1 - \left(1 - \frac{c^2}{a^2}\right) x_1^2 - \left(1 - \frac{c^2}{b^2}\right) x_2^2}. \tag{4.9}$$

Again we need to calculate $\inf_{\mathbf{x}\in\mathbb{S}^2} \rho^*(\mathbf{x})$. Noting that $c \geq a = b$, thus $-\left(1 - c^2/a^2\right) \geq 0$ and $-\left(1 - c^2/b^2\right) \geq 0$, then the square root term is minimised when x_1 and x_2 are 0, hence $\inf_{\mathbf{x}\in\mathbb{S}^2} \rho^*(\mathbf{x}) = \rho ab$. Using this we can construct the estimators of the inhomogeneous functional summary statistics given by Equations (4.3)-(4.6). Examples are given in Figure 4.4. These figures are typical for CSR with the estimated functional summary statistics lying well within the simulation envelopes.

4.6 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. [1995] 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 also provided.

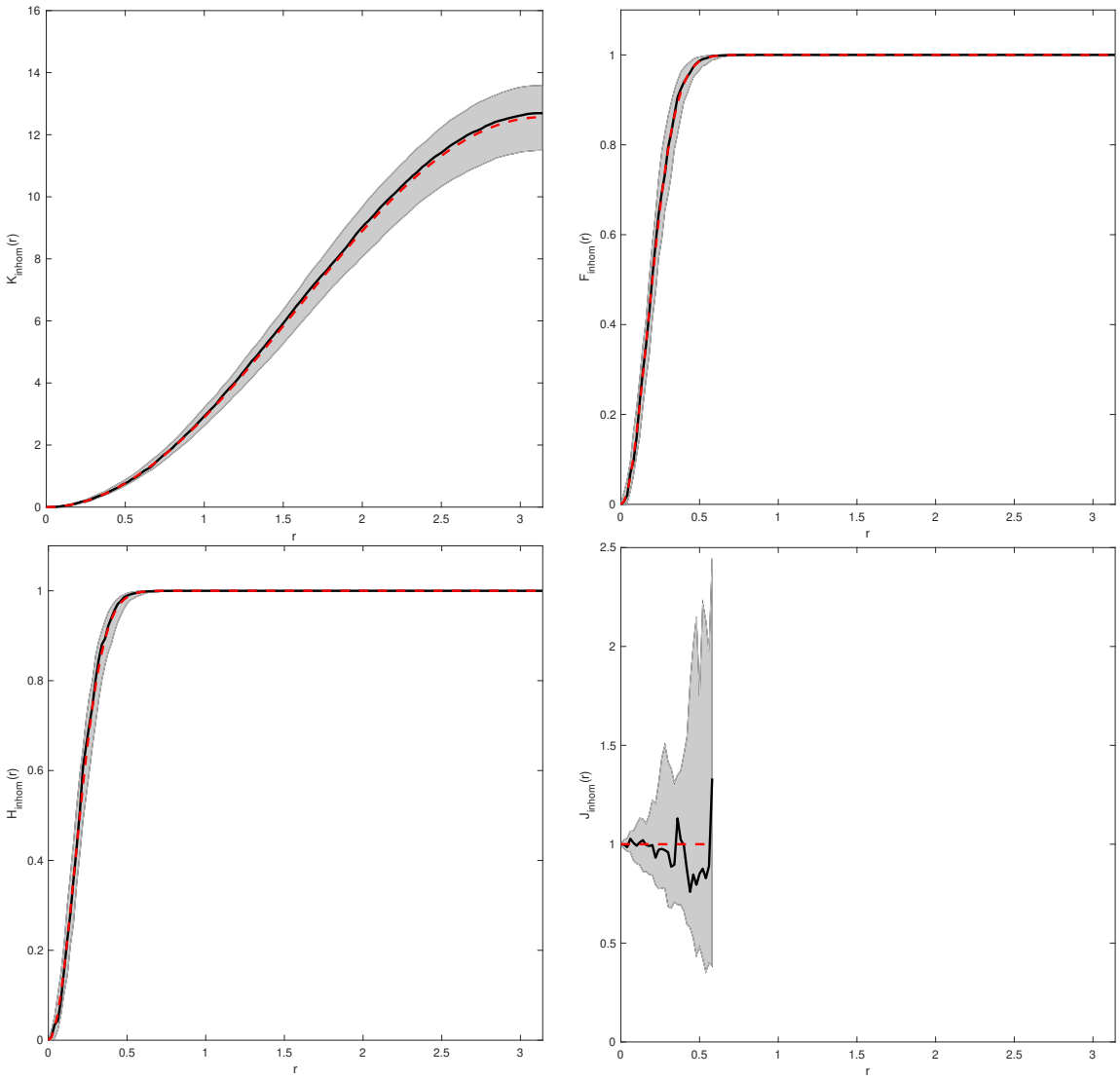


Figure 4.4: Examples of K_{inhom}- (*top left*), F_{inhom}- (*top right*), H_{inhom}- (*bottom left*), and J_{inhom}- (*bottom right*) functions for CSR patterns on a prolate ellipsoid with $a = b = 1, c = 3$, and $\rho = 5$. Solid line is the estimated functional summary statistic for our observed data, 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.

4.6.1 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 point 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 $d_{\mathbb{D}}(\mathbf{x}, \mathbf{y})$, $\mathbf{x}, \mathbf{y} \in \mathbb{D}$.

Definition 4.6.1. *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 $d_{\mathbb{D}}(\mathbf{x}, \mathbf{y}) < R$, otherwise retain $\mathbf{x}$. The resulting thinned process is then defined as a Matérn I inhibition process on $\mathbb{D}$.*

Definition 4.6.2. *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 points 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 $d_{\mathbb{D}}(\mathbf{x}, \mathbf{y}) < R$ 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 4.6.3. *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 $\alpha > 0$, $k : \mathbb{D} \times \mathbb{D} \mapsto \mathbb{R}$ such that for fixed $\mathbf{y} \in \mathbb{D}$ $k(\cdot, \mathbf{y})$ 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 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, 2003], 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,

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

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

4.6.2 Properties of regular & cluster processes

The following proposition gives the expected number of points for Matérn I and II processes in any subset of $\mathbb{D}$.

Proposition 4.6.4. *Let X_1 and X_2 be a Matérn I and II inhibition processes respectively. In the Matérn II case we also define a mark distribution $P_{M_{\mathbf{x}}}$ such that the mark is independent not only of all other points $\mathbf{y} \in X_2 \setminus \{\mathbf{x}\}$ and marks $M_{\mathbf{y}}$, for $\mathbf{y} \in X_2 \setminus \{\mathbf{x}\}$ but also of $\mathbf{x}$, the point associated to the mark $M_{\mathbf{x}}$. Define N_1 and N_2 to be the random counting measures of X_1 and X_2. Then the expectations of N_{X_1} and N_{X_2} are given as,*

$$\mathbb{E}[N_{X_1}(B)] = \rho \int_B \mathrm{e}^{-\rho\lambda_{\mathbb{D}}(B_{\mathbb{D}}(\mathbf{x},R))} \lambda_{\mathbb{D}}(d\mathbf{x})$$
$$\mathbb{E}[N_{X_2}(B)] = \int_B \frac{1 - \mathrm{e}^{-\rho\lambda_{\mathbb{D}}(B_{\mathbb{D}}(\mathbf{x},R))}}{\lambda_{\mathbb{D}}(B_{\mathbb{D}}(\mathbf{x}, R))} \lambda_{\mathbb{D}}(d\mathbf{x}),$$

where $B \subseteq \mathbb{D}$.

Proof. See Appendix B.10. □

As we will be running simulations based on Matérn II inhibition processes and given that it is a regular process there is a finite maximum number of points that can arise on the surface $\mathbb{D}$. The following corollary to the previous proposition gives the maximum expected value of $N_{X_2}(\mathbb{D})$ for a fixed hard-core distance.

Corollary 4.6.5. *Let X be a Matérn II process over $\mathbb{D}$ with hard-core distance R and defined by a Poisson process with constant intensity function ρ. Then,*

$$\sup_{\rho \in \mathbb{R}^+} \mathbb{E}[N_X(\mathbb{D})] = \int_{\mathbb{D}} \frac{1}{\lambda_{\mathbb{D}}(B_{\mathbb{D}}(\mathbf{x}, R))} \lambda_{\mathbb{D}}(d\mathbf{x}).$$

Proof. See Appendix B.11. □

The following proposition gives the expected number of points for a Thomas-type process on $\mathbb{D}$.

Proposition 4.6.6. *Let X be a Thomas-type process on $\mathbb{D}$ with concentration parameter κ and a constant expected number of offspring, λ, for each parent point. Then,*

$$\mathbb{E}[N_X(\mathbb{D})] = \lambda_{\mathbb{D}}(\mathbb{D})\rho\lambda.$$

Proof. See Appendix B.12. □

4.6.3 Simulation of regular and cluster processes

We briefly discuss how to simulate these point processes on convex shapes, much of which is based on simulation of homogeneous Poisson processes on convex shapes. To simulate homogeneous Poisson processes we can first simulate the number of points in the pattern as $\rho\lambda_{\mathbb{D}}(\mathbb{D})$ and then distribute them uniformly across $\mathbb{D}$. In order to distribute the uniformly across $\mathbb{D}$ we can use the rejection sampler outlined by Kopytov and Mityushov [2013].

Simulation of Matérn I and II then depend upon removing events in an underlying homogeneous Poisson process based on their distance between the events. This depends on being able to calculate the geodesic distance on a given surface. Assuming this can be achieved, then it is simple to simulate Matérn I and II process using their definitions. For ellipsoidal Matérn I and II processes we use the `geographiclib` [Karney, 2017] available as a `MATLAB` toolbox.

To simulate a Thomas process this can easily be achieved by rejection sampling again. Simulation of the parents is a homogeneous Poisson process. Then, for each parent, we simulate its random number of offspring and then construct a rejection sampler to sample from

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

where $\mathbf{x}_p$ is in the parent process, σ is a bandwidth parameter and $\chi(\mathbf{x}_p, \sigma^2) = \int_{\mathbb{D}} \exp(-d^2(\mathbf{x}_p, \mathbf{x}) /2\sigma^2)\lambda_{\mathbb{D}}(d\mathbf{x})$. Again for an ellipsoidal Thomas process we use `geographiclib` Karney [2017] to calculate geodesic distances on the surface of an ellipsoid.

4.6.4 Functional summary statistics assuming CSR

We simulate Matérn II, and Thomas processes and construct estimates of their functional summary statistics under the assumption that they are CSR. The inhomogenous functional summary statistics are displayed in Figure 4.5. Comparing Figure 4.5 to typical functional

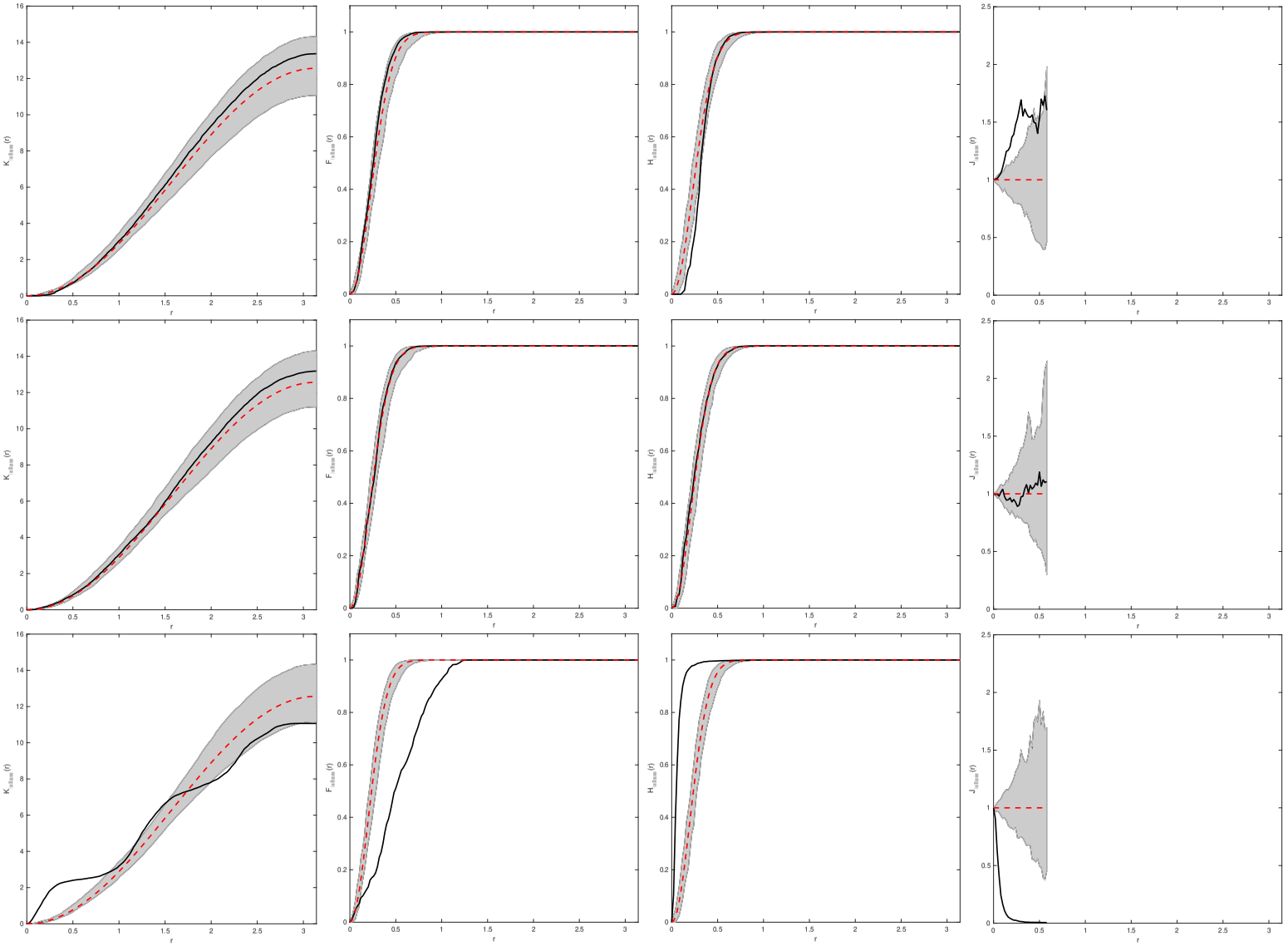


Figure 4.5: Example of (from left to right) K_{inhom}-, J_{inhom}-, F_{inhom}-, and H_{inhom}-functions for a Matérn II with parameters $R = 0.3$ and exponential mark distribution with rate $\lambda = 1$ (*top row*), Poisson process (*middle row*), and Thomas process with parameters $\kappa = 0.1$ and 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.

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.

Figure 4.5 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 row of Figure 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_{inhom} simulation envelope (top left figure) then we may be hesitant to reject CSR, but when we examine the H_{inhom} and J_{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_{inhom} function provide greater power compared to those produced with the K_{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_{inhom} is unable to provide evidence against a the process being inhomogeneous Poisson but examining both the F_{inhom} and H_{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 of their work in the inhomogeneous setting. Similar arguments to those in Baddeley et al. [2000], 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.

4.7 Discussion

In this chapter we have highlighted that construction of functional summary statistics outside of symmetric spaces such as $\mathbb{R}^d$ and $\mathbb{S}^{d-1}$ is non-trivial. We have discussed Poisson point processes on convex shapes and demonstrated, using the Mapping theorem [Kingman, 1993], that it is possible to construct functional summary statistics. This is achieved by switching analysis of point processes from the original space to the sphere where rotational symmetries

can be utilised to facilitate well-defined functional summary statistics. Using simulation envelopes, it is possible to determine whether a point pattern exhibits CSR or not and if the latter holds then, based on how the observed functional summary statistics deviates, we can determine if the point pattern displays regular or clustered behaviour. The functional summary statistics developed in this chapter lay the foundation for the following chapter, in particular the inhomogeneous K-function, where we discuss how formal hypothesis testing can be conducted using Monte Carlo techniques.

5

Testing for CSR on convex shapes

This chapter extends on the foundations made in Chapter 4 and are the test statistics constructed and explored in Ward et al. [2021b] for testing CSR. We begin by discussing the inhomogeneous K-function for a CSR processes on a convex shape $\mathbb{D}$ mapped to $\mathbb{S}^2$ when the intensity is unknown, providing derivations of the first two moments. Based on these derivations we construct an approximately variance stabilised version of the inhomogeneous K-function and use this to suggest a test statistics for the hypothesis of CSR. We also consider an analogue inhomogeneous L-function originally proposed by Lawrence [2018] in the isotropic setting and compare these two test statistics in an extensive simulation study across ellipsoids of varying dimensions.

5.1 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. Chapter 4 discusses derivations for any spherical Poisson process when ρ is known. In this section we consider the scenario when we have a CSR 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.

5.1.1 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 chapter discusses 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,
>
> $$H_0 : X \text{ is CSR on } \mathbb{D} \qquad \text{vs.} \qquad H_1 : X \text{ is not CSR on } \mathbb{D}.$$

5.1.2 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 4.3.2 as

$$\rho^*(\mathbf{x}) = \begin{cases} \rho l_1(f^{-1}(\mathbf{x}))J_{(1,f^*)}(\mathbf{x})\sqrt{1-x_1^2-x_2^2}, & \mathbf{x} \in f(\mathbb{D}_1) \\ \vdots & \\ \rho l_n(f^{-1}(\mathbf{x}))J_{(n,f^*)}(\mathbf{x})\sqrt{1-x_1^2-x_2^2}, & \mathbf{x} \in f(\mathbb{D}_n). \end{cases} \tag{5.1}$$

Using Theorems 4.4.1 and 4.4.3 we can calculate the mean and variances of the inhomogeneous K-function when ρ is known. When ρ is unknown we use estimators of ρ when constructing functional summary statistics. In particular we use

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

which are both unbiased for ρ and ρ^2 respectively by application of the Campbell-Mecke Theorem [Møller and Waagepetersen, 2003]. Thus our estimator for $K_{\text{inhom}}(r)$ when ρ is unknown takes the following form,

$$\tilde{K}_{\text{inhom}}(r) = \begin{cases} \frac{\lambda^2_{\mathbb{D}}(\mathbb{D})}{4\pi N_Y(\mathbb{S}^2)(N_Y(\mathbb{S}^2)-1)} \sum_{\mathbf{x}\in Y} \sum_{\mathbf{y}\in Y\setminus\{\mathbf{x}\}} \frac{\mathbb{1}[d(\mathbf{x},\mathbf{y})\leq r]}{\tilde{\rho}(\mathbf{x})\tilde{\rho}(\mathbf{y})}, & \text{if } N_Y(\mathbb{S}^2) > 1 \\ 0, & \text{otherwise,} \end{cases} \tag{5.2}$$

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

$$\tilde{\rho}(\mathbf{x}) = \begin{cases} l_1(f^{-1}(\mathbf{x}))J_{(1,f^*)}(\mathbf{x})\sqrt{1-x_1^2-x_2^2}, & \mathbf{x} \in f(\mathbb{D}_1) \\ \vdots & \\ l_n(f^{-1}(\mathbf{x}))J_{(n,f^*)}(\mathbf{x})\sqrt{1-x_1^2-x_2^2}, & \mathbf{x} \in f(\mathbb{D}_n). \end{cases} \tag{5.3}$$

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] suggest 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 $P_{\text{inhom}}(r) = \sqrt{K_{\text{inhom}}(r)} - \sqrt{2\pi(1-\cos(r))}$, where we use P_{inhom} rather than L_{inhom} to avoid confusion with the Euclidean L-function as P_{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_{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,

$$T_1 = \sup_{r\in[0,\pi]} \left|\tilde{P}_{\text{inhom}}(r)\right|, \qquad T_2 = \sup_{r\in[0,\pi]} \left|\frac{\tilde{K}_{\text{inhom}}(r) - 2\pi(1-\cos(r))}{\sqrt{\widehat{\text{Var}}(\tilde{K}_{\text{inhom}}(r))}}\right|, \tag{5.4}$$

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 (5.4). By using estimators for ρ and ρ^2 we alter the first and second order properties given by Theorems (4.4.1) and (4.4.3). In the following we consider the first and second order moments of $\tilde{K}_{\text{inhom}}$.

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

Theorem 5.1.1. *The bias and variance of $\tilde{K}_{inhom}(r)$ are,*

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

and,

$$\begin{aligned}
&Var(\tilde{K}_{inhom}(r)) = 4\pi^2(1-\cos r)^2(1-P(N_Y(\mathbb{S}^2)\leq 1))P(N_Y(\mathbb{S}^2)\leq 1)\\
&+\rho^3\lambda^4_{\mathbb{D}}(\mathbb{D})(1-\cos r)^2\left(\int_{\mathbb{S}^2}\frac{1}{\tilde{\rho}(\mathbf{x})}\lambda_{\mathbb{S}^2}(d\mathbf{x}) - \frac{16\pi^2}{\lambda_{\mathbb{D}}(\mathbb{D})}\right)\\
&\mathbb{E}\left[\frac{1}{(N_Y(\mathbb{S}^2)+3)^2(N_Y(\mathbb{S}^2)+2)^2}\right]\\
&+\frac{\rho^2\lambda^4_{\mathbb{D}}(\mathbb{D})}{8\pi^2}\left(\int_{\mathbb{S}^2}\int_{\mathbb{S}^2}\frac{\mathbb{1}[d(\mathbf{x}_1,\mathbf{x}_2)\leq r]}{\tilde{\rho}(\mathbf{x}_1)\tilde{\rho}(\mathbf{x}_2)}\lambda_{\mathbb{S}^2}(d\mathbf{x}_1)\lambda_{\mathbb{S}^2}(d\mathbf{x}_2) - \frac{64\pi^4(1-\cos r)^2}{\lambda^2_{\mathbb{D}}(\mathbb{D})}\right)\\
&\mathbb{E}\left[\frac{1}{(N_Y(\mathbb{S}^2)+2)^2(N_Y(\mathbb{S}^2)+1)^2}\right],
\end{aligned} \tag{5.5}$$

where $\tilde{\rho}(\mathbf{x})$ is given by Equation (5.3).

Proof. See Appendix C.1. □

The form of the variance derived in Theorem 5.1.1 is near identical to that derived by Lang and Marcon [2012] except that our derivations considers 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

$$\begin{aligned}
|\text{Bias}(\tilde{K}_{\text{inhom}}(r))| &= P(N_Y(\mathbb{S}^2) \leq 1)2\pi(1-\cos r) \\
&= \exp(-\mu_Y(\mathbb{S}^2))(1+\mu_Y(\mathbb{S}^2))2\pi(1-\cos r) \\
&\leq 4\pi(1+\mu_Y(\mathbb{S}^2))\exp(-\mu_Y(\mathbb{S}^2)) \qquad (5.6) \\
&\leq 4\pi(1+\mu_Y(\mathbb{S}^2))\mu_Y(\mathbb{S}^2)^{-\mathrm{e}} = O\left(\mu_Y^{1-\mathrm{e}}(\mathbb{S}^2)\right), \qquad (5.7)
\end{aligned}$$

where μ_Y is the intensity measure of $Y = f(X)$, the inequality in (5.6) is attained by setting $r = \pi$, and (5.7) follows from $\mathrm{e}^x \geq x^{\mathrm{e}}$. Thus, for shapes considered in this chapter, the bias will be negligible.

From Theorem 5.1.1 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 ρ^k. We can substitute the expectations in (5.5) 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 5.1.2. *Let $N \sim$ Poisson(λ), $k \in \mathbb{N}$ and $p \in \mathbb{R}_+$. Define the following random variable,*

$$R = \frac{N!e^{N-k}}{(N-k)!(e+p)^N}. \qquad (5.8)$$

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

Proof. See Appendix C.2. □

Using Lemma 5.1.2 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

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

Plugging the given estimators for $(1 - P(N_Y(\mathbb{S}^2) < 1))P(N_Y(\mathbb{S}^2) < 1)$, $\mathbb{E}[(N_Y(\mathbb{S}^2)+3)^{-2}(N_Y(\mathbb{S}^2)+2)^{-2}]$, $\mathbb{E}\left[(N_Y(\mathbb{S}^2)+2)^{-2}(N_Y(\mathbb{S}^2)+1)^{-2}\right]$, ρ^2 and ρ^3 into (5.5) gives a ratio unbiased estimator for $\mathrm{Var}(\tilde{K}_{\mathrm{inhom}}(r))$, which in turn allows for the construction of the test statistic T_2 in (5.4).

5.1.4 Standardised inhomogeneous K-function plots

Figure 5.1 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 Figure 5.1.

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 $K_{\mathrm{inhom}}(r)$ that are close to zero for small r, thus resulting in the large negative deviation observed in Figure 5.1. 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 $K_{\mathrm{inhom}}(r)$ function has large positive deviations away from CSR.

Further to this, the second row of Figure 5.1 corresponding to a Poisson process highlights the importance of applying a variance stabilising transform. If we based a test statistic on $K_{\mathrm{inhom}}(r) - 2\pi(1 - \cos(r))$ (or even $K_{\mathrm{inhom}}(r)$) then these plots suggests worst power compared to the P_{inhom}- and the standardised K_{inhom}-functions.

5.2 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

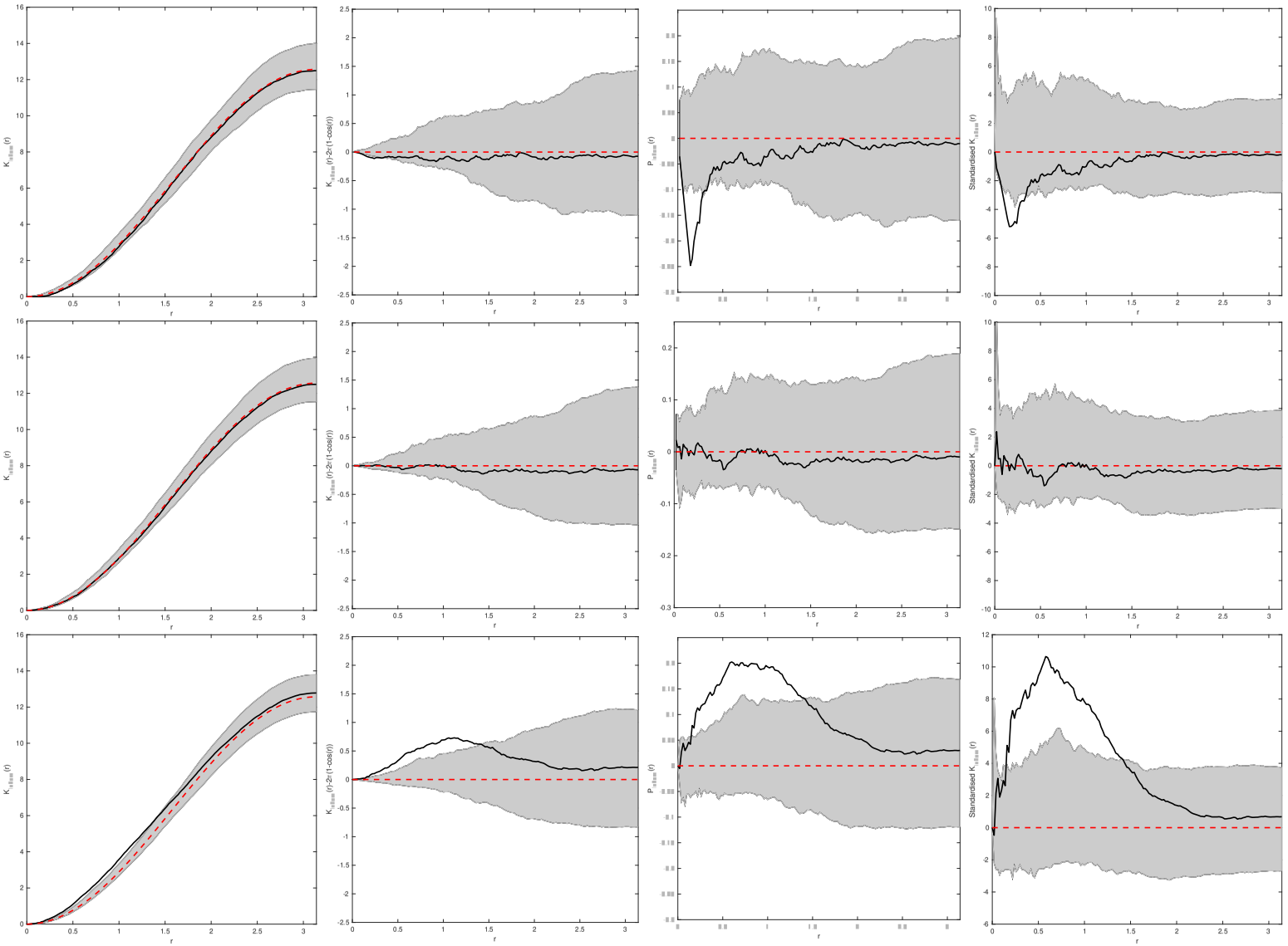


Figure 5.1: Example of (from left to right) K_{inhom}-, $K_{\text{inhom}} - 2\pi(1 - \cos(r))$-, P_{inhom}-, and the standardised K_{inhom}-functions for a Matérn II with $R = 0.2$ and expectation 100 (*top row*), Poisson process with expectation 40π , and Thomas process with parameters $\kappa = 0.5$, expectation of 150, an exponential mark distribution with rate $\lambda = 1$ and offspring mean of 20, exponential mark distribution with rate $\lambda = 1$ and offspring expectation 15 (*bottom row*) on a prolate spheroid with $a = b = 0.8000, c = 1.43983$ (dimensions chosen so that the area of the ellipsoid is 4π). 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.

changes as the space under consideration deforms further away from the unit sphere.

5.2.1 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 4.6.6.

On the other hand, the Matérn II process requires a little more attention since Corollary 4.6.5 limits the maximum expected number of possible events for a given space $\mathbb{D}$. Thus, for a given expected number μ that is less than or equal to the one prescribed by Corollary 4.6.5 we fix the hard-core distance R and solve the following equation for ρ

$$\int_B \frac{1-\mathrm{e}^{-\rho\lambda_{\mathbb{D}}(B_{\mathbb{D}}(\mathbf{x},R))}}{\lambda_{\mathbb{D}}(B_{\mathbb{D}}(\mathbf{x},R))}\lambda_{\mathbb{D}}(d\mathbf{x}) = \mu. \tag{5.9}$$

A full outline of all the experiments and the parameters chosen are given in Tables 5.1, 5.2, and 5.3 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.

5.2.2 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 $r \in [0,\pi]$ could be computationally infeasible and so we instead calculate it for $r \in \mathcal{R} = \{r_1,\ldots,r_m\}$, $m \in \mathbb{N}$, where $\mathcal{R}$ is a finite set of distinct, evenly spaced points such that $r_i \in [0,\pi], i = 1,\ldots,m$, for the purposes of our simulation studies. We then take our test statistic as

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

where for these simulation studies we set $\mathcal{R} = \{0, 0.02, 0.04, \ldots, \pi\}$. 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.

5.2.3 Results

Tables 5.1, 5.2 and 5.3 outline the parameter selection and results of our simulations. By the nature of Monte Carlo simulations the CSR results given in Table 5.1 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 κ decreases (each representing an increased departure from CSR), the power of our test improves. In Appendix C.3 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 κ respectively. Additionally, Figure 4.5 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 Chapter 4. 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].

5.3 Discussion

This chapter discusses the final contribution made in Ward et al. [2021b]. Two test statistics were proposed, one based on the work of Lawrence [2018] and the other based on a standardisation inspired by Lagache et al. [2013]. We derived properties of the latter and investigated both for their ability in determining whether a point pattern exhibits CSR through an extensive simulation study which explored ellipsoids of varying dimensions as well as processes that exhibit differing levels of aggregative and repulsive behaviour. Following testing of CSR, we would typically be interested in estimating first order properties, in particular the intensity function and is the discussion of the next chapter.

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

Table 5.1: 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 4π.

Experiment No.	Expectation	a	R	Reject $H_0 : T_1$	Reject $H_0 : T_2$
2ai	100	1	0	0.0450	0.0750
2aii	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 5.2: 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 4π. Fixing the expectation, μ, and hard-core distance, R, we use Equation 26 to calculate ρ for the underlying constant Poisson process intensity function. When $R = 0$ a Matérn II process collapses to a CSR process.

Experiment No.	Expectation	a	κ	Reject $H_0 : T_1$	Reject $H_0 : T_2$
3ai	150	1	∞	0.0290	0.0440
3aii	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	∞	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	∞	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	∞	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

Table 5.3: 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 μ 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 4π. When $\kappa = \infty$ an ellipsoidal Thomas process collapses to a CSR process.

6

Intensity estimation on Riemannian manifolds

Intensity estimation is an important step in analysing any spatial data and is the focal point of this chapter. Not only can it provide a graphical description of inhomogeneity in the data but they are key statistics that can be utilised for second order analysis of the observed pattern, e.g. construction of inhomogeneous functional summary statistics. Extensive research has been conducted for intensity estimation of two and three dimensional point patterns whilst contributions have also been made for point patterns on linear networks. Outside of such spaces research is limited. In response to this we extend the current theory for kernel estimation of intensities for point patterns observed on d-dimensional Euclidean spaces to d-dimensional Riemannian manifolds. This is the primary focus of our work Ward et al. [2021a]. We propose natural extensions of kernel intensity estimators from the Euclidean case based on the geodesic distance defined by the Riemannian metric and show that it enjoys similar statistical properties to its Euclidean counterparts. A discussion of different bandwidth selection procedures is given and an extensive simulation study is conducted to determine the effectiveness of these procedures on various types of point processes on ellipsoids of varying dimensions endowed with their canonical metric.

6.1 Introduction

Point patterns restricted to two dimensional surfaces are a relatively new form of spatial data and encourages the development of new theory and methodology as existing frameworks typically assume the process lies on some d-dimensional Euclidean space, typically two or three dimensions, and thus implicitly rely on its geometry for notions such as sta-

tionarity and isotropy. Such data has arisen in super resolution microscopy Cabriel et al. [2019], Gustavsson et al. [2018] where these experimental techniques are capable of producing spatial data restricted to the surfaces of microbes and thus there is need for proper statistical procedures to facilitate correct analysis. This chapter attempts to bridge this gap constructing nonparametric kernel estimates for the intensity function of point processes lying on Riemannian manifolds. We note that the unification of spatial point processes and differential geometry is discussed scarcely in the literature with contributions including: Alonso-Ruiz and Spodarev [2017] who discuss marked point processes where the point process lies in $\mathbb{R}^d$ whilst its marks exists on some Riemannian manifold and Jensen and Nielsen [2001] who discuss Markov point processes on Riemmanian manifolds embedded within $\mathbb{R}^d$. We extend the theory further discussing the construction of an estimator for the intensity function based on edge corrections given by Diggle [1985] and Jones [1993], van Lieshout [2012] for Euclidean processes. These estimators can help garner the level of inhomogeneity in the data but also be a useful tool in subsequent analysis, for example the construction of inhomogeneous K-functions [Baddeley et al., 2000], which in turn can help elucidate second order properties of the underlying point process.

The geometrical restriction imposed when considering point processes lying on arbitrary surfaces incurs significant challenges compared to typical two or three dimensional Euclidean processes. In particular this is discussed by Ward et al. [2021b] when constructing functional summary statistics for point processes on convex spaces highlighting that notions of stationarity and isotropy are not well-defined on such surfaces and instead advocating mapping the process onto a sphere where rotational symmetries facilitate well-defined functional summary statistics for Poisson processes. These challenges are also apparent when analysing point processes on linear networks: Ang et al. [2012], Rakshit et al. [2017], Cronie et al. [2020] discusses geometrical corrections to typical summary statistics such as Ripley's K-function Ripley [1977] and the J-function van Lieshout and Baddeley [1996], whilst McSwiggan et al. [2017], Moradi et al. [2018], Rakshit et al. [2019] discuss approaches to estimate the intensity function using kernels which respect the geometry of the linear network.

Taking inspiration from the recent advancement in the statistical analysis of linear networks, this work proposes kernel based intensity estimators for point processes residing on d-dimensional manifolds with a focus on two dimensional manifolds embedded within a three dimensional Euclidean space. In particular, like Moradi et al. [2018], we propose natural extensions of the kernel estimates used for typical two dimensional spatial data by replacing the Euclidean metric with the Riemannian metric. Even under an alternative metric we show that we can benefit from many of the same statistical properties as the Euclidean

counterpart.

Section 6.2 outlines the preliminary material needed and discusses how kernel estimation of the intensity function is conducted for two dimensional spatial data. 6.3 extends kernel estimates for planar data to point patterns residing on Riemannian manifolds, discussing a number of statistical properties. Two bandwidth selection procedures are then considered in Section 6.4 where we extend cross-validation under a Poisson hypothesis [Baddeley et al., 2015] and the nonparametric approach of Cronie and Van Lieshout [2018] to the Riemannian manifold setting. We conclude this cahpter with an in depth simulation study exploring different point processes over ellipsoids of varying dimensions in Section 6.5.

6.2 Preliminaries

In this section we will introduce the background and notation for point processes used throughout the remainder of this chapter as well as the types of surfaces we shall be working on. The notation utilised in this chapter is specific to this chapter only due to the conventions used in differential geometry. To help solidify the ideas developed later in this chapter we will briefly discuss how kernel estimation of the intensity function is currently achieved on d-dimensional Euclidean spaces to the end of this section.

6.2.1 Differential geometry

A d-dimensional manifold, $\mathcal{M}$, is a topological space that can be locally approximated at any point $\mathbf{p} \in \mathcal{M}$ by a subset of $\mathbb{R}^d$. To formally define a manifold we require the definition of a local chart. For any $U \subset \mathcal{M}$ a neighbourhood of $\mathbf{p}$ and function $\psi : U \mapsto \mathbb{R}^d$ a homeomorphism (a continuous, bijective map whose inverse exists and is continuous), (U, ψ) is said to be a *local chart of* $\mathcal{M}$. Let $\mathcal{M}$ be a topological space, and define a collection of local charts $\{(U_i, \psi_i)\}_{i \in I}$ to be an *atlas* of $\mathcal{M}$ if the set $\cup_{i \in I} U_i = \mathcal{M}$, i.e. the set of U_i covers $\mathcal{M}$ for some index set I. If for all ψ they map to the d-dimensional Euclidean plane, then $\mathcal{M}$ is said to be a manifold. We define the transition maps of $\mathcal{M}$ as $\psi_i \circ \psi_j^{-1}$ for all $i, j \in I$ such that $U_i \cap U_j \neq \emptyset$ and say that $\mathcal{M}$ is of class C^k if all transition maps are C^k $k \in \mathbb{N}$, that is the all transition maps are k-times continuously differentiable. Furthermore, the manifold is said to be *orientated* if the Jacobian determinants of the transition maps are all positive. This is a key property to define differential forms on $\mathcal{M}$ (although it is possible to do so on non-orientated manifolds through densities [Carmo, 2016]). In this chapter we will assume that our manifold is C^∞ differentiable (or smooth) and orientated. We will

write $\psi(\mathbf{p}) = (x^1(\mathbf{p}), \ldots, x^n(\mathbf{p}))$ to denote the local coordinate representation of $\mathbf{p} \in U$ in the local chart (U, ψ).

A tangent vector at $\mathbf{p}$ can be consider either as the instantaneous velocity of a curve passing through $\mathbf{p}$ in $\mathcal{M}$ or as a derivation at $\mathbf{p}$ acting on the algbera of germs of smooth functions $f : \mathcal{M} \mapsto \mathbb{R}$ (denoted $C^\infty_\mathbf{p}$) defined at $\mathbf{p}$ [Tu, 2011]. We follow the latter definition which relaxes the need for a local coordinate system as would be required by the former. A derivation at $\mathbf{p} \in \mathcal{M}$ is defined to be a linear map $D : C^\infty_\mathbf{p} \mapsto \mathbb{R}$ such that,

$$D(fg) = (Df)g(\mathbf{p}) + f(\mathbf{p})(Dg).$$

D is then a tangent at $\mathbf{p}$ [Tu, 2011]. Let $r^1, \ldots, r^n$ be the standard coordinates on $\mathbb{R}^d$ and (U, ψ) be a local chart, then $x^i = r^i \circ \psi : U \mapsto \mathbb{R}$ and then if f is a smooth function defined within some neighbourhood of $\mathbf{p}$ we define,

$$\left.\frac{\partial}{\partial x^i}\right|_\mathbf{p} f = \left.\frac{\partial}{\partial r^i}\right|_{\phi(\mathbf{p})} (f \circ \psi^{-1}),$$

and so $\partial/\partial x^i|_\mathbf{p}$ satisfies this condition and hence is a tangent. We denote the set of tangents at $\mathbf{p}$ to be $T_\mathbf{p}\mathcal{M}$ and it can be shown that the set $\partial/\partial x^1|_\mathbf{p}, \ldots, \partial/\partial x^n|_\mathbf{p}$ forms a basis of $T_\mathbf{p}\mathcal{M}$ [Tu, 2011]. We will frequently denote a tangent of $T_\mathbf{p}\mathcal{M}$ as $X_\mathbf{p}$ whilst the function X over $\mathcal{M}$ assigns an element of $T_\mathbf{p}\mathcal{M}$ for each $\mathbf{p} \in \mathcal{M}$ and is called a *vector field*. $T\mathcal{M} = \cup_{\mathbf{p}\in\mathcal{M}} T_\mathbf{p}\mathcal{M}$ is denoted the tangent bundle of $\mathcal{M}$. We let $T^*_\mathbf{p}\mathcal{M}$ be the dual space of $T_\mathbf{p}\mathcal{M}$, that is the space of all linear functions from $T_\mathbf{p}\mathcal{M}$ to $\mathbb{R}$. We define the *differential* of a smooth function f on $\mathcal{M}$ as df, which assigns for each point $\mathbf{p}$ in $\mathcal{M}$ with,

$$(df)_\mathbf{p}(X_\mathbf{p}) = X_\mathbf{p} f, \quad X_\mathbf{p} \in T_\mathbf{p}\mathcal{M}.$$

Then for a local cordinate system $(U, \psi) = (U, x^1, \ldots, x^n)$ the elements $(dx^1)_\mathbf{p}, \ldots, (dx^n)_\mathbf{p}$ are all elements of $T^*_\mathbf{p}\mathcal{M}$ and form a basis dual to the basis $\partial/\partial x^1|_\mathbf{p}, \ldots, \partial/\partial x^n|_\mathbf{p}$ of $T_\mathbf{p}\mathcal{M}$ [Tu, 2011]. We call an element $(df)_\mathbf{p} \in T^*_\mathbf{p}\mathcal{M}$ a *covector* whilst df is called a *covector field* or *1-form*, i.e. for each $\mathbf{p} \in \mathcal{M}$ it assigns an element of $T^*_\mathbf{p}\mathcal{M}$.

Consider a function on the k Cartesian product of a vector space V, i.e. $f : V^k \mapsto \mathbb{R}$. We say that f is a *k-tensor* if it is linear in each of its k arguments and the vector space of all $k-$tensors functions on V is denoted by $L_k(V)$. Furthermore, the function f is then said to be *alternating* if for all permutations $\sigma \in S_k$ (where S_k is the permuation group of size $k \in \mathbb{N}$), $f(\mathbf{v}_{\sigma(1)}, \ldots, \mathbf{v}_{\sigma(k)}) = \text{sgn}(\sigma) f(\mathbf{v}_1, \ldots, \mathbf{v}_k)$ for $\mathbf{v}_1, \ldots, \mathbf{v}_k \in V$ and we denote the space of all alternating k-tensors by $A_k(V)$. Moreover, we refer to elements of $A_k(V)$ as

k-covectors. We apply the functor $A_k(\cdot)$ to $T_{\mathbf{p}}\mathcal{M}$ and thus define $\omega_{\mathbf{p}} \in A_k(T_{\mathbf{p}}\mathcal{M})$ to be a k-covector on $\mathcal{M}$. A *k-covector field*, ω, is a function over $\mathcal{M}$ that assigns a k-covector, $\omega_{\mathbf{p}}$, for each element of $\mathbf{p} \in \mathcal{M}$. A k-covector field is also called a *k-form* and has *degree* k.

We now introduce the idea of *pushforwards* and *pullbacks* for some smooth map, F, between two manifolds $\mathcal{N}$ and $\mathcal{M}$, i.e. $F : \mathcal{N} \mapsto \mathcal{M}$. The map F induces a pushforward at each point $\mathbf{p} \in \mathcal{N}$, denoted $F_{*,\mathbf{p}}$ as,

$$(F_{*,\mathbf{p}}X_{\mathbf{p}})(f) = X_{\mathbf{p}}(f \circ F), \quad X_{\mathbf{p}} \in T_{\mathbf{p}}\mathcal{N}, f \in C^{\infty}_{F(\mathbf{p})}(\mathcal{M}).$$

Thus $F_{*,\mathbf{p}}$ maps points from $T_{\mathbf{p}}\mathcal{N}$ to $T_{F(\mathbf{p})}\mathcal{M}$. It can shown that if $f : \mathcal{M} \mapsto \mathbb{R}$ is a C^{∞} function, then for $\mathbf{p} \in \mathcal{M}$ and $X_{\mathbf{p}} \in T_{\mathbf{p}}\mathcal{M}$ we have that,

$$f_{*,\mathbf{p}}(X_{\mathbf{p}}) = (df)_{\mathbf{p}}(X_{\mathbf{p}}) \left.\frac{d}{dt}\right|_{f(\mathbf{p})},$$

(e.g. see [Tu, 2011, Proposition 17.2]) and hence $f_{*,\mathbf{p}}$ is the same as $(df)_{\mathbf{p}}$ and thus we can also refer to the differential as a pushforward and vice versa. The pullback of F, denoted $(F_{*,\mathbf{p}})^*$ is the dual notion to its pushforward, mapping $A_k(T_{F(\mathbf{p})}\mathcal{M})$ to $A_k(T_{\mathbf{p}}\mathcal{N})$. Letting $\omega_{F(\mathbf{p})}$ be a k-covector at $F(\mathbf{p})$ in $\mathcal{M}$ then $(F_{*,\mathbf{p}})^*(\omega_{F(\mathbf{p})})$ is a k-covector at $\mathbf{p}$ in $\mathcal{N}$ defined as,

$$(F_{*,\mathbf{p}})^*(\omega_{F(\mathbf{p})})(\mathbf{v}_1, \ldots, \mathbf{v}_k) = \omega_{F(\mathbf{p})}(F_{*,\mathbf{p}}\mathbf{v}_1, \ldots, F_{*,\mathbf{p}}\mathbf{v}_k), \quad \mathbf{v}, \ldots, \mathbf{v}_k \in T_{\mathbf{p}}\mathcal{N}.$$

This notion can be cumbersome and we shall denote $(F^*\omega)_{\mathbf{p}} = (F_{*,\mathbf{p}})^*(\omega_{F(\mathbf{p})})$. By dropping the subscript $\mathbf{p}$ and considering $(F^*\omega)$ as a function over $\mathbf{p} \in \mathcal{N}$ then $(F^*\omega)$ is a k-form on $\mathcal{N}$.

6.2.2 Riemannian manifolds

A Riemannian metric, g, on a smooth manifold $\mathcal{M}$ is a function over $\mathcal{M}$ which assigns to each point $\mathbf{p}$ an inner product on $T_{\mathbf{p}}\mathcal{M}$, i.e. $g_{\mathbf{p}} : T_{\mathbf{p}}\mathcal{M} \times T_{\mathbf{p}}\mathcal{M} \mapsto \mathbb{R}$, that varies smoothly with $\mathbf{p}$. $(\mathcal{M}, g)$ is then a *Riemannian manifold* and we note that g is a 2-tensor of $\mathcal{M}$. Suppose we have a local chart (U, ψ) and let $\mathbf{p} \in U \subset \mathcal{M}$ then as a function of $\mathbf{p}$ the Riemannian metric can be written locally as an $n \times n$ matrix with (i, j) elements given by [Lee, 2018],

$$g_{ij}(\mathbf{p}) = g\left(\left.\frac{\partial}{\partial x^i}\right|_{\mathbf{p}}, \left.\frac{\partial}{\partial x^j}\right|_{\mathbf{p}}\right).$$

The dependence on $\mathbf{p}$ will often be omitted, i.e. $g_{ij} = g_{ij}(\mathbf{p})$ and the associated norm will be denoted $||\cdot|| := \sqrt{g_{\mathbf{p}}(\cdot,\cdot)} = \sqrt{g(\cdot,\cdot)}$. The distance between two points, $\mathbf{p}$ and $\mathbf{q}$, on a Riemannian manifolds is then defined as the *length* of the shortest piecewise continuously-differentiable curve between $\mathbf{p}$ and $\mathbf{q}$. More precisely, let $\gamma : [a,b] \mapsto \mathcal{M}$ be a piecewise continuously-differentiable function, then the geodesic between $\gamma(a)$ and $\gamma(b)$ is the curve which minimises

$$L(\gamma) = \int_a^b \sqrt{g_{\gamma(t)}(\gamma'(t),\gamma'(t))}dt,$$

where $\gamma'(t)$ is the velocity (i.e. a derivation at $\gamma(t)$) of the curve and hence an element of $T_{\gamma(t)}\mathcal{M}$. If the geodesic between $\gamma(a)$ and $\gamma(b)$ is $\tilde{\gamma}$ then the *geodesic distance* is given by $L(\tilde{\gamma})$ and we denote it $d_g : \mathcal{M} \times \mathcal{M} \mapsto \mathbb{R}_+$. In addition to this $(\mathcal{M}, d_g)$ defines a metric space over $\mathcal{M}$.

Having defined geodesics on $\mathcal{M}$ we can now define the *exponential* and *logarithm* maps at a point $\mathbf{p}$ in the manifold. More precisely, the exponential map, $\exp_{\mathbf{p}}(\mathbf{v})$, takes elements of the tangent space at $\mathbf{p}$ and maps them to $\exp_{\mathbf{p}}(\mathbf{v}) = \gamma_{\mathbf{p},\mathbf{v}}(1) \in \mathcal{M}$ where γ is the geodesic with $\gamma(0) = \mathbf{p}$ and $\gamma'(0) = \mathbf{v} \in T_{\mathbf{p}}\mathcal{M}$. Moreover the exponential mapping is diffeomorphic within a neighbourhood of $\mathbf{0} \in T_{\mathbf{p}\mathcal{M}}$ and we define the set of all vectors $\mathbf{v}$ in $T_{\mathbf{p}\mathcal{M}}$ for which $\gamma_{\mathbf{p},\mathbf{v}}(t) = \exp_{\mathbf{p}}(t\mathbf{v})$ is the minimising geodesic for $t \in [0,1]$ but is no longer minimising for any $t = 1 + \epsilon$ $\epsilon > 0$ to be the *cut locus of* $\mathbf{p}$ *in the tangent space*. The *cut locus of* $\mathbf{p}$ is then defined as the exponential map of tangent vectors $\mathbf{v}$ at $\mathbf{p}$ that are in the cut locus of $\mathbf{p}$ in the tangent space. The infimum of the distances from $\mathbf{p}$ to elements of its cut locus is then denoted the *injectivity radius at* $\mathbf{p}$. We define the *global injectivity radius of* $\mathcal{M}$ as the infimum of the injectivity radius at $\mathbf{p}$ over all points in $\mathcal{M}$ and denote it r^*. By assuming that the global injectivity radius is positive then for any $\mathbf{p}$ in $\mathcal{M}$ we can find some ball of radius $r > 0$ for which $B_{T_{\mathbf{p}}\mathcal{M}}(\mathbf{0}, r) \subset T_{\mathbf{p}}\mathcal{M}$ is diffeomorphic into a neighbourhood U of $\mathbf{p}$. Conversely, the logarithm map at $\mathbf{p}$ is defined as the inverse of the exponential map, i.e. $\log_{\mathbf{p}}(\mathbf{q}) = \exp_{\mathbf{p}}^{-1}(\mathbf{q})$ for $\mathbf{q} \in \mathcal{M}$.

Instead of integrating functions we must instead integrate forms over a manifold [Tu, 2011]. We define the *top* or *maximal forms* to be the forms of degree d on a d-dimensional manifold. These are the only forms which can be integrated on a manifold of dimension d. Let (U, ψ) be a local chart and ω an d-form on $\mathcal{M}$ with support U, then the pullback $(\psi^{-1}) * \omega$ is a d-form on $\mathbb{R}^d$. By, for example [Tu, 2011, Proposition 3.29], a d-form on $\mathbb{R}^d$ can be written as $f(\mathbf{x})dx^1 \wedge \cdots \wedge dx^d$ for some $f : \mathbb{R}^d \mapsto \mathbb{R}$ where $dx^1 \wedge \cdots \wedge dx^d$ is the volume element on $\mathbb{R}^d$ written in terms of the exterior product of the local coordinate system, for short we will

write $d\mathbf{x} = dx^1 \wedge \cdots \wedge dx^d$. Then we define the integral of the d-form ω on U as,

$$\int_U \omega = \int_{\psi(U)} (\psi^{-1})^* \omega = \int_{\psi(U)} f(\mathbf{x}) d\mathbf{x},$$

where the integral does not depend on the choice of local chart (e.g. see [Tu, 2011, Section 23.4]). Notice that every top form also defines a measure on $\mathcal{M}$. An important d-form is the *Riemannian volume form*, denoted by $d\text{vol}$ and is given by,

$$d\text{vol}(\mathbf{x}) = \sqrt{\det(g_{ij})} d\mathbf{x} = \sqrt{\det(g_{ij})} dx^1 \wedge \cdots \wedge dx^d$$

in local coordinates. Moreover, let K be any compact subset of U, then we say that the *volume of* K is,

$$\text{Vol}(K) = \int_K d\text{vol}(\mathbf{x}).$$

Using the atlas of the manifold combined with its corresponding *partition of unity* it is also possible to define the volume of K when K is a compact set of $\mathcal{M}$ instead of just U, an individual local chart of the manifold, see for example Lee [2018] for this generalisation. The Riemannian volume form will take the place of the typical d-dimensional Lebesgue measure when considering point processes over a manifold.

The brief exposition of differential and Riemannian geometry provided here is general and based on Tu [2011], Lee [2018]; a more in depth discussion of these topics is can be found in these references. In particular we do not specify the Riemannian metric tensor g. Given the nature of spatial point processes and that, typically in applications, they are nearly always realisations in $\mathbb{R}^d$ it is therefore natural to consider Riemmanian manifolds that are embedded within some higher dimensional Euclidean space, for example the $d-1$ dimensional sphere, $\mathbb{S}^{d-1}$, is embedded within $\mathbb{R}^d$. In addition to this we can impose the canonical Riemannian metric tensor, that is the manifold inherits the metric tensor of the space its embedded in, for example returning to the $d-1$ dimensional sphere, $\mathbb{S}^{d-1}$ would inherit the Euclidean metric of $\mathbb{R}^d$.

6.2.3 Point processes on Riemannian manifolds

Throughout this chapter we denote X to be a point process over a compact d-dimensional Riemannian manifold $(\mathcal{M}, g)$ with the induced distance metric d_g. We define N_{lf} to be the set of locally finite point configurations in $\mathcal{M}$ and we assume that, almost surely, $X \in N_{lf}$. We suppose that X is observed in some compact window W of $\mathcal{M}$. Define $N_X(B)$ to be the

random counting measure associated to X which denotes the number of points of X residing in $B \subset \mathcal{M}$ and that the intensity measure is given by $\mu(B) = \mathbb{E}[N_X(B)]$, i.e. the expected number of points of X in B. If μ is absolutely continuous with respects to the Riemannian volume form, then there exists a function $\rho : \mathcal{M} \mapsto \mathbb{R}$, such that,

$$\mu(B) = \int_B \rho(\mathbf{x})\, d\text{vol}(\mathbf{x}).$$

The function ρ is the Radon-Nikodyn derivative of the intensity measure with respects to the Riemannian volume form and is called the intensity function of X. Heuristically, $\rho(\mathbf{x})\, d\text{vol}(\mathbf{x})$ can be interpreted as probability of an event of X being in an infinitesimal volume $d\text{vol}(\mathbf{x})$. If ρ is constant then we say that X is homogeneous, otherwise X is inhomogeneous. We assume that the intensity function exists and then by the Campbell theorem [Daley and Vere-Jones, 2003] we have that for any measurable, nonnegative function $f : \mathcal{M} \mapsto \mathbb{R}_+$,

$$\mathbb{E} \sum_{\mathbf{x} \in X} f(\mathbf{x}) = \int_{\mathcal{M}} f(\mathbf{x}) \rho(\mathbf{x})\, d\text{vol}(\mathbf{x}). \tag{6.1}$$

Similar to point processes in the $\mathbb{R}^d$ we define a Poisson process analogously: X is said to be a Poisson process with intensity function ρ if for any compact subset M of $\mathcal{M}$ then $N_X(M)$ is Poisson with mean $\mu(M)$ and given $N_X(M) = n$, each n event is independently and identically distributed with density $\rho(\mathbf{x})/\mu(M)$ over M. In the event that ρ is constant the resulting Poisson process is said to be CSR.

6.2.4 Kernel estimation on $\mathbb{R}^d$

Before describing kernel estimation on a Riemannian manifold we first review kernel estimation of intensity functions for point processes on $\mathbb{R}^d$ that have been observed through some bounded window $W \subseteq \mathbb{R}^d$. As defined by Silverman [1986] a kernel function $k : \mathbb{R}^d \mapsto \mathbb{R}_+$ is typically an d-dimensional symmetric probability density function and given a bandwidth parameter $h > 0$, the intensity function $\rho : \mathbb{R}^d$ may be estimated as,

$$\hat{\rho}_h(\mathbf{x}) = \sum_{\mathbf{y} \in X \cap W} \frac{c_h(\mathbf{x}, \mathbf{y})^{-1}}{h^d} k\left(\frac{\mathbf{x} - \mathbf{y}}{h}\right), \tag{6.2}$$

for $\mathbf{x} \in W$ and c_h is some edge correction. Two popular edge correction factors used for Euclidean point processes are the global [Diggle, 1985, Berman and Diggle, 1989] and local

[Jones, 1993, van Lieshout, 2012] edge corrections which are defined as,

$$c_h(\mathbf{x}, \mathbf{y}) = \frac{1}{h^d} \int_W k\left(\frac{\mathbf{x} - \mathbf{z}}{h}\right) d\mathbf{z} \quad \text{and} \tag{6.3}$$

$$c_h(\mathbf{x}, \mathbf{y}) = \frac{1}{h^d} \int_W k\left(\frac{\mathbf{z} - \mathbf{y}}{h}\right) d\mathbf{z} \tag{6.4}$$

respectively. Both edge corrections are in general biased for the true intensity function but in the event of a homogeneous point process the global correction becomes unbiased whilst the local correction remains biased. On the other hand van Lieshout [2012] demonstrated that the local correction benefits from the following mass preservation property,

$$\int_W \hat{\rho}_h(\mathbf{x})\, d\mathbf{x} = N_X(W),$$

and that taking expectations of the previous equation shows that global unbiasedness is achieved, i.e. $\int_W \hat{\rho}_h(\mathbf{x})\, d\mathbf{x}$ is unbiased for $\mu(W)$. Rakshit et al. [2019] also consider similar edge correction to these for point processes on linear networks, with a focus on rapid computation of the kernel estimate using fast Fourier transforms.

The choice of kernel k is considered to be less important than the correct choice of bandwidth h which balances close and distant interactions [Møller and Waagepetersen, 2003]. Common choices of kernels in spatial analysis include the Epanechnikov, box and Gaussian kernels. The integrated squared error (ISE) can be used to assess *quality* of a kernel estimate and is defined as,

$$\text{ISE}(\hat{\rho}_h) = \int_W \left(\hat{\rho}_h(\mathbf{x}) - \rho(\mathbf{x})\right)^2 d\mathbf{x}, \tag{6.5}$$

whilst the mean integrated squared error (MISE) is defined as the expectation of the ISE. Moreover the MISE can be decomposed as,

$$\text{MISE}(\hat{\rho}_h) = \int_W \text{Var}(\hat{\rho}_h(\mathbf{x})) + \text{Bias}^2(\hat{\rho}_h(\mathbf{x}))\, d\mathbf{x}. \tag{6.6}$$

In practice though both these measures of performance are limited since the true intensity function is rarely known but is a useful metric when conducting simulation studies as the intensity functions are indeed known in this scenario and thus can be used to measure performance of a statistical procedure.

6.3 Kernel estimation on Riemmanian manifolds

Following similar arguments made by Moradi et al. [2018] to extend kernel estimation to linear networks we propose extensions of Equation 6.2 with the edge corrections given by Equations 6.3 [Diggle, 1985, Berman and Diggle, 1989] and 6.4 [van Lieshout, 2012] when the underlying space is a compact Riemannian manifold of dimension d with metric tensor g (and hence induced geodesic distance d_g). We propose that the intensity function be estimated by,

$$\hat{\rho}_h(\mathbf{x}) = \sum_{\mathbf{y} \in X \cap W} \frac{c_h(\mathbf{x}, \mathbf{y})^{-1}}{h^d} k\left(\frac{d_g(\mathbf{x}, \mathbf{y})}{h}\right) \tag{6.7}$$

where W is some bounded window of $\mathcal{M}$, $c_h(\mathbf{x}, \mathbf{y})$ is an edge correction factor, the global and local corrections being extended from $\mathbb{R}^d$ as

$$c_h(\mathbf{x}, \mathbf{y}) = \frac{1}{h^d} \int_W k\left(\frac{d_g(\mathbf{x}, \mathbf{z})}{h}\right) d\text{vol}(\mathbf{z}) \tag{6.8}$$

$$c_h(\mathbf{x}, \mathbf{y}) = \frac{1}{h^d} \int_W k\left(\frac{d_g(\mathbf{z}, \mathbf{y})}{h}\right) d\text{vol}(\mathbf{z}) \tag{6.9}$$

respectively [Diggle, 1985, Berman and Diggle, 1989, van Lieshout, 2012], and we assume that the functions, $k : \mathbb{R} \mapsto \mathbb{R}_+$ is a one-dimensional symmetric probability distribution so $(1/h^d)k(\cdot/h)$ defines an d-dimensional kernel function. In contrast with Pelletier [2005] definition of a kernel, which requires five conditions (see Equations (41a)-(41e) of Le Brigant and Puechmorel [2019]) we require a single constraint on k. Our proposed estimator is in fact derived identically to that of Moradi et al. [2018] but where the underlying space is a Riemannian manifold rather then a linear network. We refer to $\hat{\rho}_h^{(1)}$ and $\hat{\rho}_h^{(2)}$ as the kernel estimates for ρ of X using Equations 6.8 and 6.9 for edge correction respectively. Given that previous work highlights that the choice of k is less critical than h in $\mathbb{R}^d$ [Møller and Waagepetersen, 2003] in this work we shall consider the Gaussian kernel,

$$k(d_g(\,\cdot\,, \mathbf{y})) \propto \exp\left(-\frac{d_g^2(\,\cdot\,, \mathbf{y})}{2}\right), \tag{6.10}$$

where the normalising constant results in the kernel integrating to 1 over $\mathcal{M}$. On $\mathbb{R}^d$ under the standard Euclidean metric this kernel collapses to a standard multivariate Gaussian distribution and hence the typical Gaussian kernel in d-dimensions. See also Pennec [2006] for a discussion of probability distributions on manifolds, in particular the normal distribution. It is also plausible to use other kernels in order to construct estimates of the intensity

function.

6.3.1 Properties

The statistical properties of the kernel estimates given by using the edge corrections of Equations 6.3 and 6.4 have been well studied [Diggle, 1985, Berman and Diggle, 1989, van Lieshout, 2012] and many of these properties carry over to the setting of point processes on Riemannian manifolds. In particular, the bias and variance of $\hat{\rho}_h^{(1)}$ and $\hat{\rho}_h^{(2)}$ are explored and we also provide some asymptotic results showing that convergence in MISE can be achieved for the *shape* of the true intensity function following similar arguments made Cucala [2006, 2008] for point processes that are observed with and without some error.

Bias

Using Campbell's theorem it can easily be shown that the pointwise bias is given by,

$$\mathbb{E}\left[\hat{\rho}_h^{(i)}(\mathbf{x})\right] = \int_W \frac{c_h^{-1}(\mathbf{x},\mathbf{y})}{h^d} k\left(\frac{d_g(\mathbf{x},\mathbf{y})}{h}\right) \rho(\mathbf{y})\, d\text{vol}(\mathbf{y}),$$

for $i = 1, 2$ and W a bounded window of $\mathcal{M}$. If we then suppose that ρ is constant over $\mathcal{M}$ then for $\hat{\rho}_h^{(1)}$ $c_h(\mathbf{x},\mathbf{y})$ does depend on $\mathbf{x}$ but not on $\mathbf{y}$ whilst the opposite holds for $\hat{\rho}_h^{(2)}$ and so we have that,

$$\begin{aligned}
\mathbb{E}\left[\hat{\rho}_h^{(1)}(\mathbf{x})\right] &= \frac{\rho}{c_h(\mathbf{x},\,\cdot\,)} \int_W \frac{1}{h^d} k\left(\frac{d_g(\mathbf{x},\mathbf{y})}{h}\right) d\text{vol}(\mathbf{y}) = \frac{\rho}{c_h(\mathbf{x},\,\cdot\,)} c_h(\mathbf{x},\,\cdot\,) = \rho \\
\mathbb{E}\left[\hat{\rho}_h^{(2)}(\mathbf{x})\right] &= \rho \int_W \frac{c_h^{-1}(\,\cdot\,,\mathbf{y})}{h^d} k\left(\frac{d_g(\mathbf{x},\mathbf{y})}{h}\right) d\text{vol}(\mathbf{y}).
\end{aligned}$$

Therefore $\hat{\rho}_h^{(1)}$ is unbiased and $\hat{\rho}_h^{(2)}$ biased when the process is homogeneous. Although $\hat{\rho}_h^{(2)}$ may be biased in the homogeneous setting it does maintain the mass preservation property given by van Lieshout [2012] for both homogeneous and inhomogeneous processes,

$$\int_W \hat{\rho}_h^{(2)}(\mathbf{x})\, d\text{vol}(\mathbf{x}) = N_X(W)$$

which is easily shown by interchanging integration and summation in the previous equation. Further, to this we also obtain that $\int_W \hat{\rho}_h^{(2)}(\mathbf{x})\, d\text{vol}(\mathbf{x})$ is unbiased for $\mu(W)$.

Variance

Now suppose that our process X is Poisson on $\mathcal{M}$ with intensity function ρ. It can be shown that for any measurable nonnegative function $f : \mathcal{M} \mapsto \mathbb{R}_+$ we have that $\text{Var}(\sum_{\mathbf{x}\in X} f(\mathbf{x})) = \int_W f^2(\mathbf{x})\rho(\mathbf{x})\, d\text{vol}(\mathbf{x})$ [Baddeley et al., 2015]. As such, the pointwise variances for $\rho_h^{(1)}$ and $\rho_h^{(2)}$ under a Poisson assumption are

$$\text{Var}\left(\hat{\rho}_h^{(1)}(\mathbf{x})\right) = c_h^{-2}(\mathbf{x}, \cdot) \int_W \left(\frac{1}{h^d} k\left(\frac{d_g(\mathbf{x}, \mathbf{y})}{h}\right)\right)^2 \rho(\mathbf{y})\, d\text{vol}(\mathbf{y})$$
$$\text{Var}\left(\hat{\rho}_h^{(2)}(\mathbf{x})\right) = \int_W c_h^{-2}(\cdot, \mathbf{y}) \left(\frac{1}{h^d} k\left(\frac{d_g(\mathbf{x}, \mathbf{y})}{h}\right)\right)^2 \rho(\mathbf{y})\, d\text{vol}(\mathbf{y}).$$

Similar to Rakshit et al. [2019], if the Poisson process is homogeneous with constant intensity ρ then

$$\text{Var}\left(\hat{\rho}_h^{(1)}(\mathbf{x})\right) = \rho c_h^{-2}(\mathbf{x}, \cdot) \int_W \left(\frac{1}{h^d} k\left(\frac{d_g(\mathbf{x}, \mathbf{y})}{h}\right)\right)^2 d\text{vol}(\mathbf{y})$$
$$\text{Var}\left(\hat{\rho}_h^{(2)}(\mathbf{x})\right) = \rho \int_W c_h^{-2}(\cdot, \mathbf{y}) \left(\frac{1}{h^d} k\left(\frac{d_g(\mathbf{x}, \mathbf{y})}{h}\right)\right)^2 d\text{vol}(\mathbf{y}),$$

highlighting that variance is not constant across $\mathcal{M}$ even if the true intensity function is constant. Unbiased estimates of the variance can easily be obtained using

$$\widehat{\text{Var}}\left(\hat{\rho}_h^{(i)}(\mathbf{x})\right) = \sum_{\mathbf{y}\in X\cap W} \left(\frac{c_h^{-1}(\mathbf{x}, \mathbf{y})}{h^d} k\left(\frac{d_g(\mathbf{x}, \mathbf{y})}{h}\right)\right)^2,$$

where unbiasedness follows by applying Campbell's theorem. Again these formulas are similar to those derived by Rakshit et al. [2019] for point processes on linear networks.

Asymptotics

We now look at the asymptotics of the bias and variance under a Poisson assumption we also assume, for this section, that our kernel is Gaussian, see Equation 6.10. When typically dealing with kernel density estimation all points are assumed IID and the total number of points is deterministic, where the asymptotic scenario is to let N, the number of observations, go to infinity whilst h goes to zero. In the point process setting in order to determine asymptotic characteristics the increasing-domain asymptotic framework is often employed [Cressie, 1993], where the observation window is allowed to increase such that the expected

number of points goes to infinity. There are two issues with applying this framework in the current situation:

1. As noted by Cucala [2008], as the bandwidth tends to 0 the estimated intensity at each point then depends on an expected number of events which tends to 0.

2. Since we are working on a compact manifold the assumption of an infinitely increasing domain is not a feasible.

Instead we can let the expected number of points go to infinity over $W \subseteq \mathcal{M}$ in such a way that the *shape* of ρ remains unchanged, for example we could suppose that $\rho(\mathbf{x}) = \nu\rho_1(\mathbf{x})$ such that ρ_1 integrates to 1 over W then $\mu_X(W) = \nu$ and as ν (and hence $\mu_X(W)$) goes to infinity ρ_1 remains unchanged. This is the setting considered by Cucala [2006, 2008] for point processes on $\mathbb{R}^d$ with and without observational noise and Rakshit et al. [2019] for linear networks. We thus consider the asymptotic properties of

$$\hat{\rho}^{(i)}_{h,1}(\mathbf{y}) = \frac{\mathbb{1}[N_X(W) \neq 0]}{N_X(W)} \sum_{\mathbf{x} \in X \cap W} \frac{1}{h^d} k\left(\frac{-d_g(\mathbf{x}, \mathbf{y})}{h}\right) c_h(\mathbf{x}, \mathbf{y})^{-1}, \tag{6.11}$$

for $i = 1, 2$. We show that it is possible to achieve pointwise unbiasedness and consistency and thus convergence in integrated means squared error to $\rho_1(\mathbf{y}) = \rho(\mathbf{y})/\mu(\mathcal{M})$. Clearly, based on Equation 6.11 we have the relationship $\hat{\rho}^{(i)}_h(\mathbf{y}) = N_X(W)\hat{\rho}^{(i)}_{h,1}(\mathbf{y})$, with our original kernel density estimator. The following proposition gives our pointwise asymptotic properties for $\hat{\rho}^{(i)}_{h,1}$ for $i = 1, 2$.

Proposition 6.3.1. *Let $(\mathcal{M}, g)$ be a Riemannian manifold with Riemannian metric tensor g. We observe our process through some bounded window $W \subseteq \mathcal{M}$ and suppose that it is Poisson with continuous intensity ρ such that $\rho_1(\mathbf{x}) = \rho(\mathbf{x})/\mu(W) = O(1)$ as $\mu(W)$ goes to ∞ and is smooth. Then,*

$$\mathbb{E}\left[\hat{\rho}^{(i)}_{h,1}(\mathbf{x})\right] \to \rho_1(\mathbf{x}) \quad \textit{and},$$
$$\textit{Var}\left(\hat{\rho}^{(i)}_{h,1}(\mathbf{x})\right) \to 0,$$

for $i = 1, 2$ with Gaussian kernel, $\mathbf{x} \in W \subseteq \mathcal{M}$ as $h \to 0$, and $\mu(W) \to \infty$, such that $A(\mu(W))/h^d \to 0$ where $A(\mu(W)) = \mathbb{E}[\mathbb{1}[N_X(W) \neq 0]/N_X(W)]$.

Proof. See Appendix D.2. For a proof in $\mathbb{R}^d$ see Cucala [2006, pp. 105-107]. □

In Appendix D.2 we show that under a Poisson hypothesis,

$$A(\mu(W)) \leq \frac{2}{\mu(W)},$$

and hence we can replace the condition $A(\mu(W))/h^d \to 0$ with $1/(\mu(W)h^d) \to 0$. If we further suppose that our intensity is of the form $\rho(\mathbf{x}) = \nu\rho_1(\mathbf{x})$ such that ρ_1 integrates to 1 over W, then we have that $\mu(W) = \nu$ and so our condition can be written as $1/(\nu h^d) \to 0$, which is precisely the condition found for kernel density estimators for ν IID points distributed with density ρ_1, see for example Wand and Jones [1994, Equation 4.8].

An immediate consequence from Propositon 6.3.1 is pointwise consistency of our kernel estimator with a Gaussian kernel. We can also achieve convergence in MISE.

Corollary 6.3.2. *Let the same conditions on X hold as for Proposition 6.3.1 and $W \subseteq \mathcal{M}$ be a bounded window of our Riemannian manifold. Then we have that,*

$$\mathbb{E}\left[\int_W (\hat{\rho}^{(i)}_{h,1}(\mathbf{x}) - \rho_1(\mathbf{x}))^2 dvol(\mathbf{x})\right] \to 0,$$

as $h \to 0$ and $\mu(W) \to \infty$ such that $A(\mu(W))/h^d \to 0$ where $A(\mu(W)) = \mathbb{E}[\mathbb{1}[N_X(W) \neq 0]/N_X(W)]$.

Proof. Follows immediately from Proposition 6.3.1 and using Equation 6.6, but adapted for Riemmanian manifolds instead of $\mathbb{R}^d$. □

6.4 Bandwidth selection

To select h it has been suggested that a critical inspection of intensity plots is necessary in order to balance local and global features in the data [Møller and Waagepetersen, 2003], whilst other approaches appeal to optimisation criteria. Baddeley et al. [2015] suggest selecting bandwidths based on optimising the cross-validation Poisson log likelihood given by

$$\ell_{\text{cv}} = \sum_{\mathbf{x}\in X} \log\left(\hat{\rho}_h^{-\mathbf{x}}(\mathbf{x})\right) - \int_{\mathcal{M}} \hat{\rho}_h(\mathbf{x})\, d\text{vol}(\mathbf{x}), \tag{6.12}$$

where $\hat{\rho}_h^{-\mathbf{z}}(\mathbf{x}) = h^{-d}\sum_{\mathbf{y}\in X\setminus\{\mathbf{z}\}} k\left(-d_g(\mathbf{x},\mathbf{y})/h\right) c_h^{-1}(\mathbf{x},\mathbf{y})$ is then an estimate of ρ without the observation $\mathbf{z} \in X$. It is easy to show that under a Poisson assumption and using the Cambell-Mecke Theorem, that the cross validation log likelihood is unbiased for the log

likelihood function

$$\ell = \log L = \sum_{\mathbf{x}\in X} \log(\rho(\mathbf{x})) - \int_{\mathcal{M}} \rho(\mathbf{x})\, d\text{vol}(\mathbf{x}).$$

A bandwidth selection technique that requires special consideration here is *state estimation for isotropic Cox processes* [Diggle, 2014, Section 5.3], in particular this approach cannot be trivially adapted to Riemmanian manifolds. Starting on $\mathbb{R}^d$, Diggle [1985] assumes that the underlying process is a stationary, isotropic Cox process such that the driving random field, Λ, has constant mean μ and covariance function $c(\mathbf{x}, \mathbf{y}), \mathbf{x}, \mathbf{y} \in \mathbb{R}^d$ such that it only depends on $||\mathbf{x} - \mathbf{y}||$, i.e. the Euclidean distance between $\mathbf{x}$ and $\mathbf{y}$. By assuming a stationary and istropic Cox process the first order intensity function, ρ is constant, and the second order intensity function of X also only depends on the distance between the points i.e. $\rho^{(2)}(\mathbf{x}, \mathbf{y}) = \rho^{(2)}(||\mathbf{x} - \mathbf{y}||)$. Then Diggle [1985] suggest minimising the mean squared error between the kernel density estimate and the realisation of Λ i.e. minimise,

$$\text{MISE}(h) = \mathbb{E}\left[(\hat{\rho}_h(\mathbf{x}) - \Lambda(\mathbf{x}))^2\right], \tag{6.13}$$

where the expectation is with respect to both Λ and the process X. By the assumption of stationarity MISE(h) does not depend on $\mathbf{x}$ and so without loss of generality $\mathbf{x}$ can be taken to be $\mathbf{0}$ [Diggle, 2014]. Additionally under a box kernel, i.e. $\hat{\rho}_h(\mathbf{x}) = N_X(B_{\mathbb{R}^d}(\mathbf{x}, r))/(\pi h^2)$ it can be shown that [Diggle, 2014],

$$\text{MISE}(h) = \rho^{(2)}(0) + \frac{\rho(1 - 2\rho K(h))}{\pi h^2} + \frac{1}{\pi^2 h^4} \int_{B_{\mathbb{R}^d}(\mathbf{0},h)} \int_{B_{\mathbb{R}^d}(\mathbf{0},h)} \rho^{(2)}(||\mathbf{x} - \mathbf{y}||) d\mathbf{y} d\mathbf{x}, \tag{6.14}$$

where the double integral can be converted into a single integral using the substitution $\mathbf{z} = \mathbf{x} - \mathbf{y}$, and converting to polar coordinates (see for example Diggle [2014], Cronie and Van Lieshout [2018]) to give

$$\text{MISE}(h) = \rho^{(2)}(0) + \frac{\rho(1 - 2\rho K(h))}{\pi h^2} + \frac{\rho^2}{\pi^2 h^4} \int_0^{2h} \left(2h^2 \cos^{-1}\left(\frac{s}{2h}\right) - \frac{s}{2}\left(4h^2 t^2\right)^{1/2}\right) K(dt).$$

Since the first term does not depend on h it can be ignored and thus in order to optimise the MISE we need estimates of ρ and Ripley's K-function.

Now if we are working on a Riemmanian manifold this approach does not easily translate, primarily due to the fact that notions of stationarity and isotropy are ill-defined on such surfaces [Ward et al., 2021b]. This means that the assumption of a stationary and isotropic Cox process, which the MISE in $\mathbb{R}^d$ is built on is inapproriate for point processes on $\mathcal{M}$. Importantly this means that the MISE in Equation 6.13 is not independent of $\mathbf{x}$, and so

$\mathbf{x} = \mathbf{0}$ cannot be used without loss of generality. Instead one could use

$$\text{MISE}(h) = \int_{\mathcal{M}} \mathbb{E}\left[(\hat{\rho}_h(\mathbf{x}) - \Lambda(\mathbf{x}))^2\right] \, d\text{vol}(\mathbf{x}). \tag{6.15}$$

If we suppose that X is Cox process which has constant mean function and covariance function which only depends through $d_g(\mathbf{x}, \mathbf{y})$ then again X has constant intensity function and a second order intensity function which only depends on the geodesic distance between the points. If we then take the integrand of Equation 6.15 it can be shown, using standard conditioning arguments that,

$$\begin{aligned}\mathbb{E}\big[(\hat{\rho}_h(\mathbf{x}) - \Lambda(\mathbf{x})\,)^2\big] &= \rho^{(2)}(0) + \frac{\rho}{\text{Vol}(B_{\mathcal{M}}(\mathbf{x}, h))} \\ &+ \frac{2}{\text{Vol}(B_{\mathcal{M}}(\mathbf{x}, h))} \int_{B_{\mathcal{M}}(\mathbf{x},h)} \rho^{(2)}(d_g(\mathbf{x}, \mathbf{y})) \, d\text{vol}(\mathbf{y}) \\ &+ \frac{1}{\text{Vol}^2(B_{\mathcal{M}}(\mathbf{x}, h))} \int_{B_{\mathcal{M}}(\mathbf{x},h)} \int_{B_{\mathcal{M}}(\mathbf{x},h)} \rho^{(2)}(d_g(\mathbf{y}, \mathbf{z})) \, d\text{vol}(\mathbf{y}) \, d\text{vol}(\mathbf{z}).\end{aligned} \tag{6.16}$$

Unlike the Euclidean counterpart given by Equation 6.14, simplifications for the two integrals are not easily obtainable due to the difficulties imposed by the geometry of $\mathcal{M}$. This means that Equation 6.16 cannot be easily be simplified into a function depending on some type of Ripley's K function as in the Euclidean case, moreover from the discussion given by Ward et al. [2021b], functional summary statistics defined directly on $\mathcal{M}$ are not easily obtainable. One could potentially impose structure on $\rho^{(2)}$ by assuming some parametric model for the Cox process, for example supposing it is a Log Gaussian Cox process defined by some parameters $\boldsymbol{\theta}$. Although a potential solution it imposes further restrictions which may not be appropriate in practice. Moreover the minimisation of Equation 6.15 would then also need to be made over both h and $\boldsymbol{\theta}$ adding to the required computational resources. Irrespective of how to deal with $\rho^{(2)}$, there is an additional integral over $\mathcal{M}$ which needs calculating, compare Equations 6.14 and 6.15, adding further to the complexity of using state estimation for bandwidth selection on Riemannian manifolds. For these reasons state estimation is not considered further.

Another approach to bandwidth selection is given by Cronie and Van Lieshout [2018] who utilise Campbell's Theorem in order to select an appropriate bandwidth. Unlike the two previous methods discussed, Cronie and Van Lieshout [2018]'s approach does not rely on any model assumptions and the optimisation criteria derived by Cronie and Van Lieshout [2018] is free of any integrals and thus benefits from faster computation. Translated into the Riemannian setting, Cronie and Van Lieshout [2018] noticed that by setting $f(\mathbf{x}) = 1/\rho(\mathbf{x})$

in Campbell's Theorem, Equation 6.1, and assuming that the intensity function is positive everywhere on $\mathcal{M}$ then $\sum_{\mathbf{x}\in X} \rho^{-1}(\mathbf{x})$, also known as the Stoyan & Grabarnik statistic [Stoyan and Grabarnik, 1991], is an unbiased estimator of $\text{Vol}(W)$ in the event that ρ is known. More often the not, ρ is not available and instead Cronie and Van Lieshout [2018] replace ρ with $\hat{\rho}_h$ and select the value of h which minimises

$$F(h) = \left(T(\hat{\rho}_h) - \text{Vol}(W)\right)^2, \tag{6.17}$$

where $T(\hat{\rho}_h) = \sum_{\mathbf{x}\in X} \hat{\rho}_h^{-1}(\mathbf{x})$. From Equation 6.17 it is clear that there are computational advantages over minimising Equation 6.12: (1) no integration is required and, (2) a comparable number of calculations are conducted to evaluate the first term of both equations since they both involve a double sum over a similar number of elements.

Cronie and Van Lieshout [2018] discuss continuity and limit properties of T on $\mathbb{R}^d$ which are important in order to optimise F. Before extending this to Riemannian manifolds we first discuss some properties of the edge correction factor given by Equations 6.8 and 6.9. First we show that it is continuous in h.

Lemma 6.4.1. *For the Gaussian kernel (see Equation 6.10), the edge correction factors as defined by Equations 6.8 and 6.9 are continuous for all $h \in (0, \infty)$.*

Proof. See Appendix D.3. See Cronie and Van Lieshout [2018] for a proof in $\mathbb{R}^d$. □

As an immediate result of Lemma 6.4.1 it can easily be shown that Equation 6.12 is continuous in h for a Gaussian kernel since $(1/h^d)k(-d_g(\mathbf{x},\mathbf{y})/h)c_h(\mathbf{x},\mathbf{y})^{-1}$ is continuous and hence so is $\hat{\rho}_h$. The logarithm is a continuous function and so the composition of log and $\hat{\rho}_h$ is also continuous. The second term of Equation 6.12 with the integral can be shown to be continuous in h by using the same argument as given in Lemma 6.4.1 and hence ℓ_{cv} is continuous in h.

The following theorem extends Theorem 1 of Cronie and Van Lieshout [2018] to Riemannian manifolds in the specific case of a Gaussian kernel given by Equation 6.10.

Theorem 6.4.2. *Let X be a point process over a Riemannian manifold $(\mathcal{M}, g)$ and is observed through some bounded window $W \subseteq \mathcal{M}$. Disregard the trivial case $X \cap W$ is the empty set. Suppose the same restrictions apply to $(\mathcal{M}, g)$ as in Lemma D.1.1. Then T is a continuous function in $h \in (0, \infty)$ when $c_h(\mathbf{x}, \mathbf{y}) = 1$ or given by Equations 6.8 and 6.9. We*

also obtain the following limit cases,

$$T(\hat{\rho}_h) \to 0 \qquad \text{as} \quad h \to 0,$$

for $c_h(\mathbf{x}, \mathbf{y}) = 1$ *and* $c_h(\mathbf{x}, \mathbf{y})$ *given by Equations 6.8 and 6.9 whilst,*

$$T(\hat{\rho}_h) \to \infty \qquad \text{as} \quad h \to \infty,$$

for $c_h(\mathbf{x}, \mathbf{y}) = 1$ *and,*

$$T(\hat{\rho}_h) \to Vol(W) \quad \text{as} \quad h \to \infty,$$

for $c_h(\mathbf{x}, \mathbf{y})$ *given by Equations 6.8 and 6.9.*

Proof. See Appendix D.4. See Cronie and Van Lieshout [2018, Theorem 1] for a proof in $\mathbb{R}^d$. □

From Theorem 6.4.2 we have that when $c_h(\mathbf{x}, \mathbf{y}) = 1$ there exists a minima for F given by Equation 6.17, whilst if edge correction is used a minimum occurs when $h \to \infty$, which is akin to the Euclidean approach considered by Cronie and Van Lieshout [2018], and thus, like Cronie and Van Lieshout [2018] we suggest to optimise F with no shape correction, but to include the edge correction term after a bandwidth has been selected.

6.5 Simulation study

We explore how the cross validation critera [Baddeley et al., 2015] compares to Cronie's criteria [Cronie and Van Lieshout, 2018] for bandwidth selection on manifolds. More precisely we shall run simulations on ellipsoids of varying dimensions embedded within $\mathbb{R}^3$ endowed with the canonical metric. To calculate the geodesics on the ellipsoids we use the GeographicLib `MATLAB` package [Karney, 2017] which implements the algorithms of Karney [2012]. We will suppose that we have viewed the entire point pattern, i.e. $W = \mathcal{M}$ which is possible since $\mathcal{M}$ is compact, and generate 100 realisations of each model. It should be noted that even though we have completely observed the process over $\mathcal{M}$ in these simulations this does not implicitly mean the edge correction terms become 1 as they would in the hypothetical scenario of a completely observed point pattern in $\mathbb{R}^d$. Thus the edge corrections used on compact Riemannian manifolds serve an additional purpose, not only correcting for events that may have occurred outside the observation window but also to account for the geometrical con-

straints imposed by the space; this is similar to what is encountered for point processes on linear networks Moradi et al. [2018], McSwiggan et al. [2017], Rakshit et al. [2019]. Following Cronie and Van Lieshout [2018], we will select the bandwidth using no edge correction, i.e. $c_h(\mathbf{x}, \mathbf{y}) = 1$ where we run a gridsearch over for h on the set $\{0.01, 0.02, \ldots, 4.99, 5.00\}$. Then, to assess quality we will calculate the average ISE (see Equation 6.5) using the selected bandwidths and applying local edge correction, i.e. using Equation 6.9. For the purpose of computational speed we approximate the ISE as a Riemannian sum. More precisely, we use the inverse of the local chart $\mathbf{x} = (a\sin(\theta)\cos(\phi), b\sin(\theta)\sin(\phi), c\cos(\theta))$, denoting it ψ, to map the integral over $\mathcal{M}$ to an integral over $\mathbb{R}^2$ where $\theta \in [0, \pi], \phi \in [0, 2\pi]$ thus $\det(g_{ij}) = \sin^2(\theta)a^2b^2(1 - (1 - c^2/a^2)\sin(\theta)^2\cos(\phi)^2 - (1 - c^2/b^2)\sin(\theta)^2\sin(\phi)^2)$ and so,

$$\begin{aligned} ISE(\hat{\rho}_h) &= \int_{\mathcal{M}} \left(\hat{\rho}_h(\mathbf{x}) - \rho(\mathbf{x})\right)^2 d\text{vol}(\mathbf{x}) \\ &= \int_0^{\pi}\int_0^{2\pi} \left(\hat{\rho}_h(\psi^{-1}(\theta,\phi)) - \rho(\psi^{-1}(\theta,\phi))\right)^2 \sqrt{\det(g_{ij})}\, d\theta d\phi. \end{aligned} \tag{6.18}$$

We then equidistantly partition $[0, \pi]$ and $[0, 2\pi]$ into 100, e.g. $[0, \pi] = (\cup_{i=1}^{99}[(i-1)\pi/100, i\pi/100)) \cup [99\pi/100, \pi]$ and thus we can form a Riemannain sum over these partitions to approximate the ISE as

$$ISE(\hat{\rho}_h) \approx \sum_{i,j=1}^{100} \left(\frac{\pi}{100}\frac{2\pi}{100}\right) u(\theta_i, \phi_j),$$

where $\pi/100 \cdot 2\pi/100$ corresponds to the area of each partition, which is the same for all, u is given by the integrand of Equation 6.18 and $\theta_i = (2i-1)\pi/200$ and $\phi_j = (2i-1)2\pi/200$ for $i, j = 1, \ldots, 100$. This same idea is used to evaluate the integral in the cross validation criteria, see Equation 6.12. In order for the results to be comparable we divide the average ISE by the expected number of points over $\mathcal{M}$. In order to calculate the expected number of points we can again use the same local chart applied to $\int_{\mathcal{M}} \rho(\mathbf{x})\, d\text{vol}(\mathbf{x})$ and use `MATLAB`'s `integral2` function to calculate the double integral numerically.

In the simulation conducted by Cronie and Van Lieshout [2018] they consider homogeneous, linear and modulated intensities, we follow a similar study setup considering homogeneous, log-linear and log-modulation intensity functions in the Poisson and clustered case whilst linear and modulated intensities are used for the regular case which is based in location dependent thinning. We consider the log counterparts of those consider by Cronie and Van Lieshout [2018] to ensure positivity of the intensity function since our space admits negative elements, for example if we consider the 2-dimensional sphere then $(-1, 0, 0)^T \in \mathbb{S}^2$. We shall

consider three ellipsoids with axis lengths adjusted in order to ensure their Riemmanian volume measure, i.e. their surface area, is 1. First we consider the 2-dimensional sphere with radius $r = (1/(4\pi))^{1/2}$, the ellipsoids with axis length $a = b = 0.8 \cdot (1/(4\pi))^{1/2}$ and $a = b = 0.6 \cdot (1/(4\pi))^{1/2}$ along the x- and y-axis with the axis length along the z-axis adjusted to ensure unit surface area; the former ellipsoid shall be referred to as *Ellipsoid 1* and the latter *Ellipsoid 2* in the upcoming simulation studies.

Overall, from our simulation study we have been able to replicate the results of Cronie and Van Lieshout [2018] but in the context of a Riemannian manifold instead of the typical $\mathbb{R}^d$ structure. Broadly speaking the cross-validation selection critera performs substantially better, in terms of ISE, compared to Cronie's selection criteria Cronie and Van Lieshout [2018] when the true process is Poisson. In setting of more clustered processes, and more precisely LGCP processes, we see that Cronie's approach perform significantly better and for processes that exhibit more regular behaviour we see have more mixed results where cross-validation outperforms Cronie's approach in general with some specific scenarios indicating that Cronie's method is better. In the regular case though the difference in the performance is marginal.

6.5.1 Poisson processes

As a base case we consider Poisson processes with the following intensity functions,

$$\rho_1(\mathbf{x}) = \rho \qquad \text{homogeneous} \tag{6.19}$$

$$\rho^{(2)}(\mathbf{x}) = \exp(3 + Ax) \qquad \text{log-linear} \tag{6.20}$$

$$\rho_3(\mathbf{x}) = \exp(2 + A\cos(8y)) \qquad \text{log-modulation} \tag{6.21}$$

where $\mathbf{x} = (x, y, z)^T \in \mathcal{M}$ and $\rho, A \in \mathbb{R}^d$. For the homogeneous case we shall consider $\rho \in \{50, 150, 300\}$, the log-linear case $A \in \{10, 18, 22\}$ and the log-modulation case $A \in \{3, 4, 5\}$. These cases allow us to investigate the properties of the two bandwidth selections in the presences of a low, medium and high expected number of points over $\mathcal{M}$. Figure 6.1 displays typical simulates across the three ellipsoids for when $\rho = 150$ in the homogeneous case, $A = 18$ in the log-linear case and $A = 4$ in the log modulation case (see Equations 6.19-6.21).

The results of our Poisson simulation study are presented in Table 6.1. As expected, since the true process is Poisson the cross-validation selection criteria consistently outperforms Cronie's selection criteria due to the Poisson assumption in its construction.

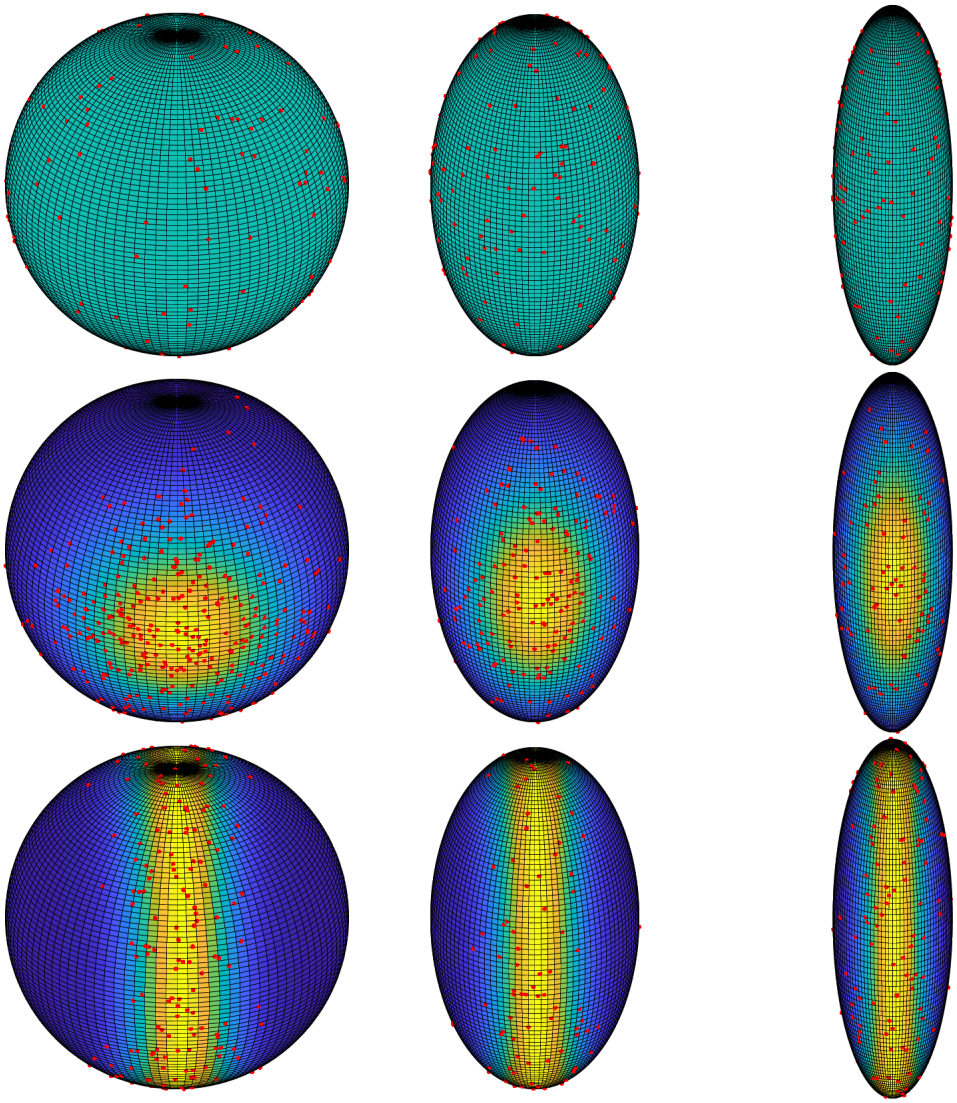


Figure 6.1: Examples of simulated Poisson processes on different ellipsoids. *Top row:* homogenoeus Poisson process with $\rho = 150$, *middle row:* log-linear Poisson process with $A = 18$, and *bottom row:* log-modulated Poisson process with $A = 4$. *Left column:* sphere with $r = (1/(4\pi))^{1/2}$, *middle column:* ellipsoid with $a = b = 0.8 \cdot (1/(4\pi))^{1/2}$, and *right column:* with ellipsoid with $a = b = 0.6 \cdot (1/(4\pi))^{1/2}$. Yellow on the surface indicates high intensity whilst the blue low intensity. Please note these images are not to scale: this refers to both the axes lengths and the intensities across the different ellipsoids and processes.

Shape	Intensity Type	Process Parameters	$\mathbb{E}[\mathbf{N}_{\mathbf{X}}(\mathcal{M})]$	Average ISE - CV	Average ISE - Cronie
Sphere	Homogeneous	$\rho = 50$	50.0000	4.0526	4.7053
Sphere	Homogeneous	$\rho = 150$	150.0000	4.9511	7.2912
Sphere	Homogeneous	$\rho = 300$	300.0000	4.9232	10.2879
Sphere	Log-linear	$A = 10$	59.5714	15.8867	60.4332
Sphere	Log-linear	$A = 18$	317.2406	51.9008	943.6023
Sphere	Log-linear	$A = 22$	802.2368	98.4829	3236.1471
Sphere	Log-modulated	$A = 3$	49.9820	21.8054	39.8503
Sphere	Log-modulated	$A = 4$	116.1563	40.0078	125.9735
Sphere	Log-modulated	$A = 5$	280.1509	70.9938	381.0162
Ellipsoid 1	Homogeneous	$\rho = 50$	50.0000	4.0749	4.5990
Ellipsoid 1	Homogeneous	$\rho = 150$	150.0000	4.7520	7.4986
Ellipsoid 1	Homogeneous	$\rho = 300$	300.0000	5.6133	10.3270
Ellipsoid 1	Log-linear	$A = 10$	43.5999	11.8202	29.1954
Ellipsoid 1	Log-linear	$A = 18$	153.7524	33.6446	309.6516
Ellipsoid 1	Log-linear	$A = 22$	313.6824	49.4227	862.5868
Ellipsoid 1	Log-modulated	$A = 3$	58.6030	20.4560	32.0236
Ellipsoid 1	Log-modulated	$A = 4$	135.9275	40.0391	105.2776
Ellipsoid 1	Log-modulated	$A = 5$	326.7017	69.5342	328.4795
Ellipsoid 2	Homogeneous	$\rho = 50$	50.0000	4.2731	4.4250
Ellipsoid 2	Homogeneous	$\rho = 150$	150.0000	5.8025	7.8692
Ellipsoid 2	Homogeneous	$\rho = 300$	300.0000	6.5221	10.5295
Ellipsoid 2	Log-linear	$A = 10$	32.3422	11.0831	15.5067
Ellipsoid 2	Log-linear	$A = 18$	75.0547	21.9043	98.0805
Ellipsoid 2	Log-linear	$A = 22$	123.1302	34.1756	225.4659
Ellipsoid 2	Log-modulated	$A = 3$	74.8816	20.5229	23.4277
Ellipsoid 2	Log-modulated	$A = 4$	175.6577	37.2837	72.6545
Ellipsoid 2	Log-modulated	$A = 5$	423.0580	62.9209	235.9694

Table 6.1: Average ISE for Poisson simulation study. The sphere has radius $(1/(4\pi))^{1/2}$. For Ellipsoid 1 and 2 the minor axis length along the x- and y-axis are the same and $0.8 \cdot (1/(4\pi))^{1/2}$ and $0.6 \cdot (1/(4\pi))^{1/2}$ respectively whilst the minor axis length along the z-axis is adjusted to ensure unit surface area of the ellipsoid. The columns for intensity type and process parameters correspond to Equations 6.19-6.21. The average ISE is divided by the expected number of points for each process to put them on a comparable scale.

6.5.2 Log Gaussian Cox processes

Next we explore how the bandwidth selection criteria perform when the underlying process exhibits aggregative behaviour. In order to do so we shall consider LGCPs. We say that X is a Cox process if, given some driving random field Z over $\mathcal{M}$, X is a Poisson process with intensity given by Z; X is then LGCP if $Z(\mathbf{x}) = \exp(U(\mathbf{x}))$ where U is a GRF over $\mathcal{M}$. The function $U : \mathcal{M} \mapsto \mathbb{R}$ is a GRF if for any $n \in \mathbb{N}$, $\mathbf{x}_1, \ldots, \mathbf{x}_n \in \mathcal{M}$ and $a_1, \ldots, a_n \in \mathbb{R}$ then $\sum_{i=1}^{n} a_i U(\mathbf{x}_i)$ is a univariate normal. Since our ellipsoids are embedded in $\mathbb{R}^3$ in order to simulate GRFs we first simulate them in $\mathbb{R}^3$ and then *extract* only those elements that lie on the ellipsoid, i.e if U is a GRF over $\mathbb{R}^3$ then on our ellipsoid our GRF is $U' = \{U(\mathbf{x}) : \mathbf{x} \in \mathcal{M}\}$: this simulation approach is similar to that considered by Cronie et al. [2020] where, in their simulation study, they simulate a LGCP over $\mathbb{R}^2$ and then only extract those elements that lie on the linear network of interest. Letting U be a GRF over $\mathcal{M}$ with mean function $\mu : \mathcal{M} \mapsto \mathbb{R}$ and covariance function $c : \mathcal{M} \times \mathcal{M} \mapsto \mathbb{R}$, then the intensity function of a LGCP X with driving random field $\exp(U)$ is

$$\rho(\mathbf{x}) = \exp\left(\mu(\mathbf{x}) + \frac{c(\mathbf{x}, \mathbf{x})}{2}\right)$$

[Møller and Waagepetersen, 1998]. In this simulation study we will set the covariance function to be exponential, that is

$$c(\mathbf{x}, \mathbf{y}) = \sigma^2 \exp\left(-\frac{\sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2}}{\gamma^2}\right),$$

where $\mathbf{x} = (x_1, x_2, x_3)^T$, $\mathbf{y} = (y_1, y_2, y_3)^T$, $\sigma^2 \in \{2\log(2), 2\log(5)\}$ and $\gamma^2 \in \{1/10, 1/50\}$. Our intensity function will be $\rho(\mathbf{x}) = \exp\left(\mu(\mathbf{x}) + \sigma^2/2\right)$. We consider three mean functions for the GRF in our simulations,

$$\mu_1(\mathbf{x}) = \log(\rho) \qquad \text{homogeneous} \tag{6.22}$$

$$\mu_2(\mathbf{x}) = 4 + 3x \qquad \text{log-linear} \tag{6.23}$$

$$\mu_3(\mathbf{x}) = 6\cos(8y) - 1 \qquad \text{log-modulation,} \tag{6.24}$$

where $\mathbf{x} = (x, y, z)^T$, with $\rho \in \{10, 50\}$. Examples of these processes are shown in Figure 6.2.

Table 6.2 displays the results of our simulation study for differing LGCPs across differently shaped ellipsoids. From this study we can see that Cronie's bandwidth selection procedure

[Cronie and Van Lieshout, 2018] is consistently better then using cross-validation. The only case when the cross-validation approach outperforms Cronie's method is for the log-modulated study with $(\sigma^2, \gamma^2) = (2\log(2), 1/10)$ for the Sphere and Ellipsoid 1. This could be due to the balance between the *fixed* and *random effect* of LGCP. More precisely the mean function, representing the fixed effect, dominates the GRF, representing the random effect and so, for this specific set of parameters, the LGCP is more like a Poisson process with intensity $\exp(\mu(\mathbf{x}))$ which explains why cross-validation outpeforms Cronie's approach in this experiment. We also note that our findings in this simulation study are comparable to the one conducted by Cronie and Van Lieshout [2018]. That is with a decreasing range of interaction (i.e. decressing γ^2) we typically see smaller ISE, whilst a larger variablity (i.e. increasing σ^2) sees larger ISE.

6.5.3 Strauss processes

Next we analyse the performance of these selection criteria for processes that exhibit more regular behaviour, more precisely how accurately these kernel intensity estimators recover the intensity of a Strauss process. Strauss, and more broadly Markov, processes have been discussed on general differentiable manifolds in Jensen and Nielsen [2001] but with the primary focus being $\mathbb{R}^d$. Letting Z be the unit rate Poisson over $\mathcal{M}$ we define a Strauss process to be the point process with density

$$f(x) \propto \beta^{n_x(\mathcal{M})} \exp\left(-\alpha \sum_{\{\mathbf{x},\mathbf{y}\}\in x} \mathbb{1}[d_g(\mathbf{x},\mathbf{y}) \leq R]\right) \tag{6.25}$$

for $x \in N_{lf}, n_x(\mathcal{M})$ is the cardinality of x in $\mathcal{M}$ and $\beta, \alpha, R > 0$, with respect to Z and where the sum is taken such that $\mathbf{x} \neq \mathbf{y}$. With $\alpha > 0$ we ensure that the density is integrable (see e.g. Møller and Waagepetersen [2003]). The parameter R controls the range of the interaction between events whilst α dictates how strong the repulsive effect between events is enforced in the Strauss process.

We define the Papangelou conditional intensity as $\lambda(\mathbf{x}, x) = f(x \cup \mathbf{x})/f(x)$ for $\mathbf{x} \in \mathcal{M}$ and $x \in N_{lf}$ [Møller and Waagepetersen, 2003] and so for the Strauss process we have,

$$\lambda(\mathbf{x}, x) = \frac{f(x \cup \mathbf{x})}{f(x)} = \beta \exp\left(-2\alpha \sum_{\mathbf{y}\in x} \mathbb{1}[d_g(\mathbf{x},\mathbf{y}) \leq R]\right),$$

notice that the Papangelou conditional intensity does not depend on the normalising con-

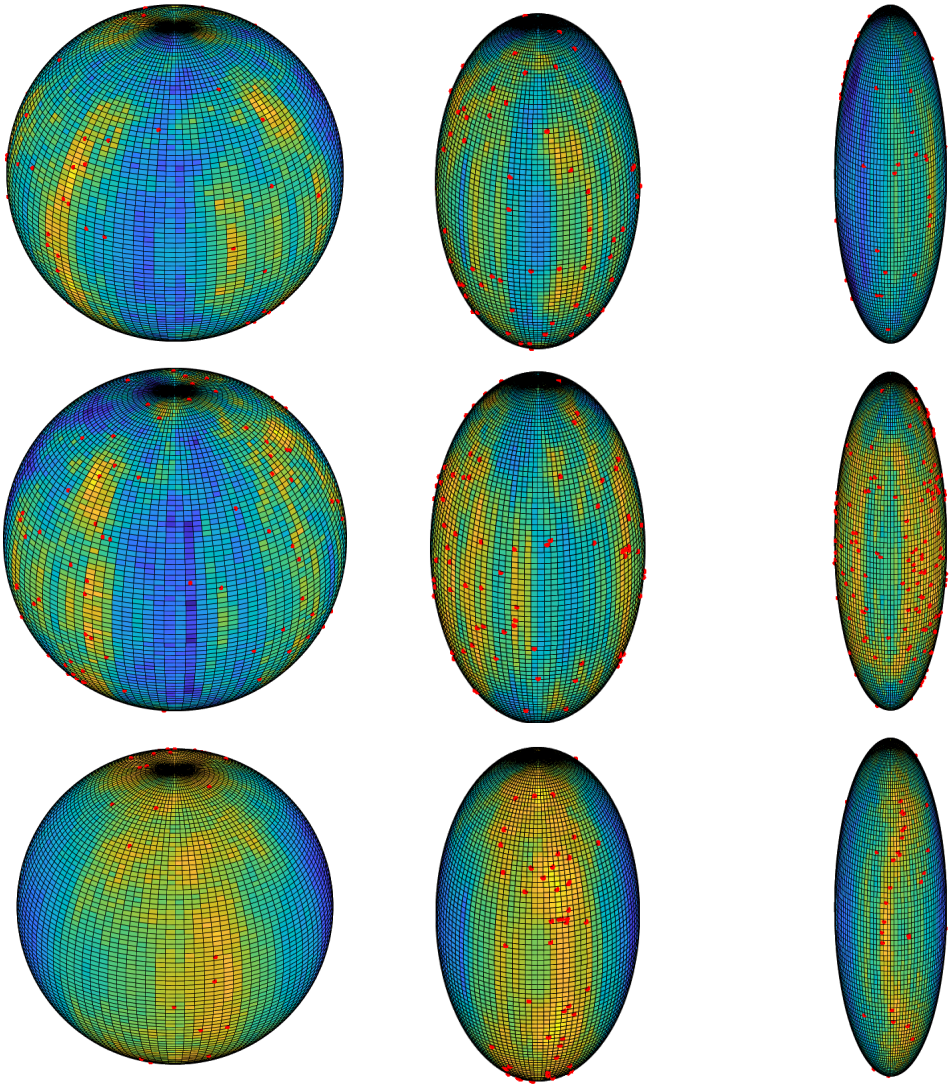


Figure 6.2: Examples of simulated LGCP on different ellipsoids, all simulated with $(\sigma^1, \gamma^2) = (2\log(2), 1/10)$ and the intensities shown are those from the realisation of the random field, U, not the true intensity function of the process and not $\exp(U)$. *Top row:* homogeneous LGCP with $\rho = 50$, *middle row:* log-linear LGCP with intensity function defined by Equation 6.23, and *bottom row:* log-modulated LGCP with intensity function defined by Equation 6.24. *Left column:* sphere with $r = (1/(4\pi))^{1/2}$, *middle column:* ellipsoid with $a = b = 0.8 \cdot (1/(4\pi))^{1/2}$, and *right column:* with ellipsoid with $a = b = 0.6 \cdot (1/(4\pi))^{1/2}$. Yellow on the surface indicates high intensity whilst the blue low intensity. Please note these images are not to scale: this refers to both the axes lengths and the intensities across the different ellipsoids and processes.

Shape	Intensity Type	Process Parameters	$\mathbb{E}[\mathbf{N}_{\mathbf{X}}(\mathcal{M})]$	Average ISE - CV	Average ISE - Cronie
Sphere	Homogeneous	$(\rho, \sigma^2, \gamma^2) = (10, 2\log(2), 1/10)$	20.0000	17.0929	8.0739
Sphere	Homogeneous	$(\rho, \sigma^2, \gamma^2) = (50, 2\log(2), 1/10)$	100.0000	97.5720	34.3151
Sphere	Homogeneous	$(\rho, \sigma^2, \gamma^2) = (10, 2\log(5), 1/10)$	50.0000	244.3839	41.9379
Sphere	Homogeneous	$(\rho, \sigma^2, \gamma^2) = (50, 2\log(5), 1/10)$	250.0000	1919.7599	204.3011
Sphere	Homogeneous	$(\rho, \sigma^2, \gamma^2) = (10, 2\log(2), 1/50)$	20.0000	4.6832	3.6429
Sphere	Homogeneous	$(\rho, \sigma^2, \gamma^2) = (50, 2\log(2), 1/50)$	100.0000	24.6142	9.6393
Sphere	Homogeneous	$(\rho, \sigma^2, \gamma^2) = (10, 2\log(5), 1/50)$	50.0000	52.5429	10.8179
Sphere	Homogeneous	$(\rho, \sigma^2, \gamma^2) = (50, 2\log(5), 1/50)$	250.0000	783.4225	54.0848
Sphere	Log-linear	$(\sigma^2, \gamma^2) = (2\log(2), 1/10)$	122.7054	155.1708	48.5685
Sphere	Log-linear	$(\sigma^2, \gamma^2) = (2\log(2), 1/50)$	122.7054	40.2732	12.4336
Sphere	Log-linear	$(\sigma^2, \gamma^2) = (2\log(5), 1/10)$	306.7636	1859.7728	213.6367
Sphere	Log-linear	$(\sigma^2, \gamma^2) = (2\log(5), 1/50)$	306.7636	1959.3217	79.6290
Sphere	Log-modulated	$(\sigma^2, \gamma^2) = (2\log(2), 1/10)$	68.8620	231.1246	144.9872
Sphere	Log-modulated	$(\sigma^2, \gamma^2) = (2\log(2), 1/50)$	68.8620	96.8251	120.0373
Sphere	Log-modulated	$(\sigma^2, \gamma^2) = (2\log(5), 1/10)$	172.1551	3045.9504	495.2031
Sphere	Log-modulated	$(\sigma^2, \gamma^2) = (2\log(5), 1/50)$	172.1551	1415.2029	326.0981
Ellipsoid 1	Homogeneous	$(\rho, \sigma^2, \gamma^2) = (10, 2\log(2), 1/10)$	20.0000	23.4645	9.5831
Ellipsoid 1	Homogeneous	$(\rho, \sigma^2, \gamma^2) = (50, 2\log(2), 1/10)$	100.0000	164.7407	45.1687
Ellipsoid 1	Homogeneous	$(\rho, \sigma^2, \gamma^2) = (10, 2\log(5), 1/10)$	50.0000	417.4968	72.0366
Ellipsoid 1	Homogeneous	$(\rho, \sigma^2, \gamma^2) = (50, 2\log(5), 1/10)$	250.0000	3018.0571	348.3341
Ellipsoid 1	Homogeneous	$(\rho, \sigma^2, \gamma^2) = (10, 2\log(2), 1/50)$	20.0000	5.7675	3.4204
Ellipsoid 1	Homogeneous	$(\rho, \sigma^2, \gamma^2) = (50, 2\log(2), 1/50)$	100.0000	28.2673	11.8004
Ellipsoid 1	Homogeneous	$(\rho, \sigma^2, \gamma^2) = (10, 2\log(5), 1/50)$	50.0000	34.3588	9.0967
Ellipsoid 1	Homogeneous	$(\rho, \sigma^2, \gamma^2) = (50, 2\log(5), 1/50)$	250.0000	471.2450	40.4640
Ellipsoid 1	Log-linear	$(\sigma^2, \gamma^2) = (2\log(2), 1/10)$	118.2890	176.2341	53.5139
Ellipsoid 1	Log-linear	$(\sigma^2, \gamma^2) = (2\log(2), 1/50)$	118.2890	40.0890	11.7100
Ellipsoid 1	Log-linear	$(\sigma^2, \gamma^2) = (2\log(5), 1/10)$	295.7225	4235.0825	468.0618
Ellipsoid 1	Log-linear	$(\sigma^2, \gamma^2) = (2\log(5), 1/50)$	295.7225	680.6555	46.9309
Ellipsoid 1	Log-modulated	$(\sigma^2, \gamma^2) = (2\log(2), 1/10)$	80.0582	1054.5760	216.5820
Ellipsoid 1	Log-modulated	$(\sigma^2, \gamma^2) = (2\log(2), 1/50)$	80.0582	103.8792	111.3561
Ellipsoid 1	Log-modulated	$(\sigma^2, \gamma^2) = (2\log(5), 1/10)$	200.1454	3024.4015	568.7395
Ellipsoid 1	Log-modulated	$(\sigma^2, \gamma^2) = (2\log(5), 1/50)$	200.1454	1169.9196	310.1984
Ellipsoid 2	Homogeneous	$(\rho, \sigma^2, \gamma^2) = (10, 2\log(2), 1/10)$	20.0000	18.3928	10.4260
Ellipsoid 2	Homogeneous	$(\rho, \sigma^2, \gamma^2) = (50, 2\log(2), 1/10)$	100.0000	93.3750	42.2984
Ellipsoid 2	Homogeneous	$(\rho, \sigma^2, \gamma^2) = (10, 2\log(5), 1/10)$	50.0000	171.5190	46.0187
Ellipsoid 2	Homogeneous	$(\rho, \sigma^2, \gamma^2) = (50, 2\log(5), 1/10)$	250.0000	1077.4310	219.4151
Ellipsoid 2	Homogeneous	$(\rho, \sigma^2, \gamma^2) = (10, 2\log(2), 1/50)$	20.0000	6.5671	4.3024
Ellipsoid 2	Homogeneous	$(\rho, \sigma^2, \gamma^2) = (50, 2\log(2), 1/50)$	100.0000	35.9551	13.2055
Ellipsoid 2	Homogeneous	$(\rho, \sigma^2, \gamma^2) = (10, 2\log(5), 1/50)$	50.0000	50.3529	11.8543
Ellipsoid 2	Homogeneous	$(\rho, \sigma^2, \gamma^2) = (50, 2\log(5), 1/50)$	250.0000	495.4170	50.0350
Ellipsoid 2	Log-linear	$(\sigma^2, \gamma^2) = (2\log(2), 1/10)$	114.4394	121.6319	52.3291
Ellipsoid 2	Log-linear	$(\sigma^2, \gamma^2) = (2\log(2), 1/50)$	114.4394	41.8724	14.2393
Ellipsoid 2	Log-linear	$(\sigma^2, \gamma^2) = (2\log(5), 1/10)$	286.0986	1681.8290	297.8267
Ellipsoid 2	Log-linear	$(\sigma^2, \gamma^2) = (2\log(5), 1/50)$	286.0986	667.2383	66.6023
Ellipsoid 2	Log-modulated	$(\sigma^2, \gamma^2) = (2\log(2), 1/10)$	103.5088	187.4522	129.5535
Ellipsoid 2	Log-modulated	$(\sigma^2, \gamma^2) = (2\log(2), 1/50)$	103.5088	100.4447	94.5360
Ellipsoid 2	Log-modulated	$(\sigma^2, \gamma^2) = (2\log(5), 1/10)$	258.7721	1320.8154	437.5811
Ellipsoid 2	Log-modulated	$(\sigma^2, \gamma^2) = (2\log(5), 1/50)$	258.7721	1373.0900	287.0485

Table 6.2: Average ISE for LGCP simulation study. The sphere has radius $(1/(4\pi))^{1/2}$. For Ellipsoid 1 and 2 the minor axis length along the x- and y-axis are the same and $0.8 \cdot (1/(4\pi))^{1/2}$ and $0.6 \cdot (1/(4\pi))^{1/2}$ respectively whilst the minor axis length along the z-axis is adjusted to ensure unit surface area of the ellipsoid. The columns for intensity type and process parameters correspond to Equations 6.22-6.24. The average ISE is divided by the expected number of points for each process to put them on a comparable scale.

stant.

This is a natural extension of the Strauss processes from $\mathbb{R}^d$ to $\mathcal{M}$ by replacing the canonical Euclidean metric with the Riemannian metric over $\mathcal{M}$. Although natural it is worth noting that the induced distance metric d_g is no longer *stationary*. For example consider $\mathbb{R}^d$ endowed with its canonical metric then for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$ we have that $\sqrt{(\mathbf{x}-\mathbf{y})^T(\mathbf{x}-\mathbf{y})} = \sqrt{(\mathbf{x}+\mathbf{z}-\mathbf{y}+\mathbf{z})^T(\mathbf{x}+\mathbf{z}-\mathbf{y}+\mathbf{z})}$ for any $\mathbf{z} \in \mathbb{R}^d$ but no such equivalent property holds more generally over all manifolds, this is a special property of Euclidean spaces. As such a Strauss process defined by Equation 6.25 is stationary in $\mathbb{R}^d$ but for a more general manifold this may not be the case, moreover it need not even be homogeneous over $\mathcal{M}$. This becomes an issue when trying to calculate the ISE as we would need to know the true intensity of the process. We instead approximate the ISE. To do so we note that the intensity of a point process, X, with Papangelou conditional intensity λ is given by $\rho(\mathbf{x}) = \mathbb{E}[\lambda(\mathbf{x}, X)]$ (see e.g. [Møller and Waagepetersen, 2003, Proposition 6.2]) and so assuming that we can simulate our process X we can approximate ρ as,

$$\rho(\mathbf{x}) \approx \frac{1}{N}\sum_{i=1}^{N} \lambda(\mathbf{x}, X_i), \tag{6.26}$$

where X_i are point processes with Papangelou conditional intensity λ. We can simulate realisations of X using a spatial birth-death-move Metropolis-Hastings (SBDMMH) algorithm and letting the chain run for a sufficiently long time that the invariant distribution has been reached [Møller and Waagepetersen, 2003]. We can thus use Equation 6.26 to approximate ρ and thus approximate the ISE. Further to this we can estimate $\mathbb{E}[N_X(\mathcal{M})]$ as $(1/N)\sum_{i=1}^{N} N_{X_i}(\mathcal{M})$ where we have N realisations X_i of our process instead of calculating the mean number of points using integration as done in the Poisson and LGCP studies. In order to consider different forms of inhomogeneity, more then that imposed by the original Strauss model by previous discussion, we shall independently thin the process using one of the three following retention probabilities,

$$p_1(\mathbf{x}) = 1 \qquad \text{pseudo-homogeneous} \tag{6.27}$$

$$p_2(\mathbf{x}) = \frac{1+3x}{1+3a} \qquad \text{linear} \tag{6.28}$$

$$p_3(\mathbf{x}) = \frac{2+4\cos(10x)}{6} \qquad \text{modulation,} \tag{6.29}$$

where $\mathbf{x} = (x, y, z)^T$ and a is the minor axis length along the x-axis of the ellipsoid. Equation 6.27 in fact refers to the original, unthinned Strauss process and we refer to it as *pseudo-*

homogeneous due to previous discussion; that on arbitrary Riemannian manifolds the process is not necessarily homogeneous. Examples of these processes are given in Figure 6.3. For our simulation study we will set $\beta = 400$, $\alpha \in \{0.01, 0.1\}$ and $R \in \{0.1, 0.2\}$, whilst in order to construct the unbiased estimate ρ as in Equation 6.26 we set $N = 10000$. Trace plots of $N_X(\mathcal{M})$ and $\sum_{\{\mathbf{x},\mathbf{y}\}\in X} \mathbb{1}[d_g(\mathbf{x}, \mathbf{y}) \leq R]$ from the SBDMMH algorithm showed that the effect of the initial distributions disappears after around 5000 iterations typically. We therefore run the chain with a burn-in of 100000 and then subsample the chain every 1000 from then on.

Results for the Strauss process are given in Table 6.3. From these results we see varying results with cross-validation outperforming Cronie's method [Cronie and Van Lieshout, 2018] in general, in particular it has improved performance in the pseudo-homogeneous case and modulated settings, whilst Cronie's method seems to give better results when the intensity function is linear. Even with these differences in performance they are extremely marginal, especially in the linear setting. Typically, it also appears that as α or R are increased there is a reduction in average ISE, but this should be examined under some suspicion as increasing the values of these parameters results in a reduction of the expected number of points and we have observed that the average ISE also drops when this occurs.

6.6 Discussion

In this chapter we have extended the theory of nonparametric kernel intensity estimation from a d-dimensional Euclidean space to a d-dimensional Riemannian manifold. We have shown that many of the desirable properties from Euclidean space extend and have adapted two bandwidth selection criteria to this scenario drawing comparable results to their Euclidean counterparts in a simulation study [Cronie and Van Lieshout, 2018]. Based on the results of the simulation we suggest using Cronie and Van Lieshout [2018] criteria over the cross validation criteria unless the point process exhibits significant regularity: conclusions similar to that of Cronie and Van Lieshout [2018]. In the final chapter we discuss how functional summary statistics can be constructed for multivariate point patterns on convex shapes. Methodology developed here is key to estimating the intensity function and subsequently used as a plug-in estimator for the summary statistics of multivariate point processes on convex shapes.

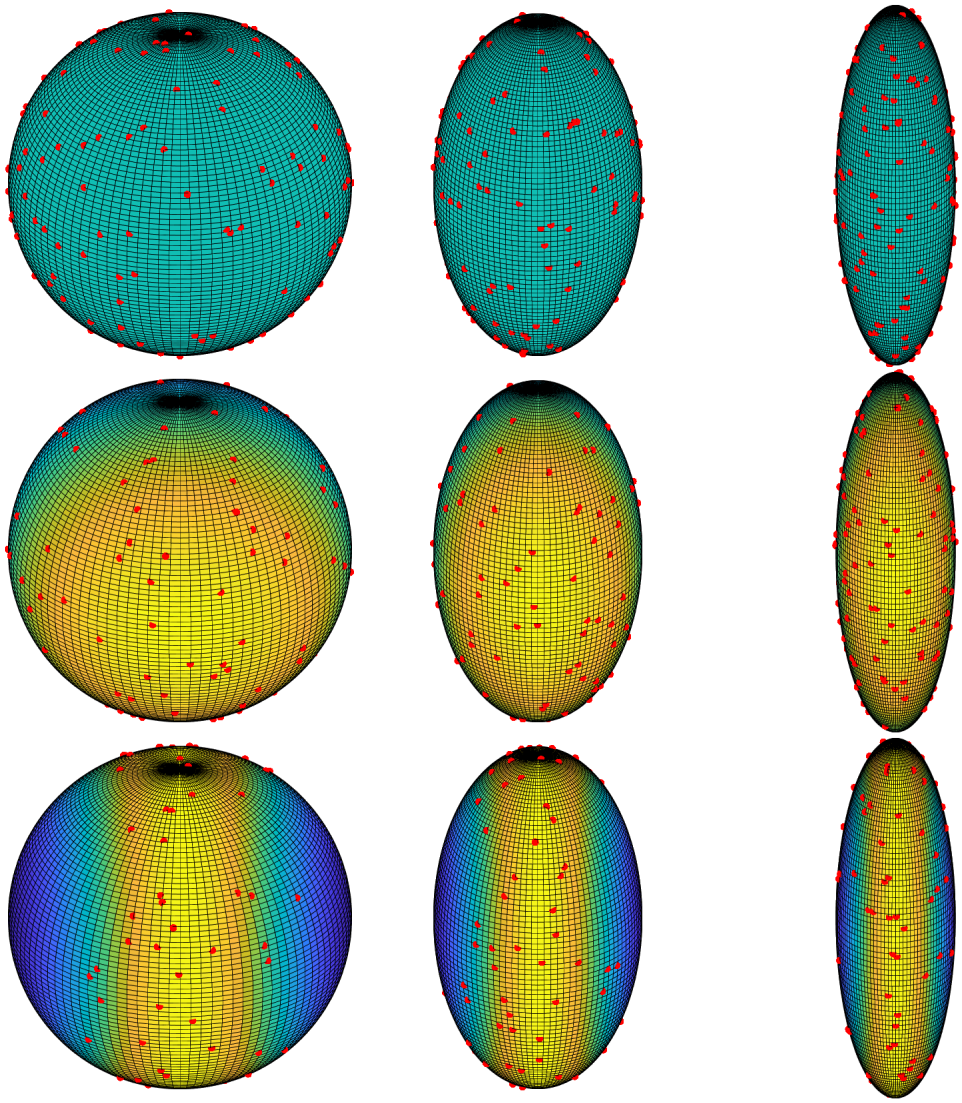


Figure 6.3: Examples of simulated Strauss processes on different ellipsoids all with parameters $(\beta, \alpha, R) = (400, 0.1, 0.1)$. *Top row:* pseudo-homogenoeus Strauss process (for the intensity plot we have simply taken the mean over the grid of points since the intensity is nearly equivalent everywhere), *middle row:* log-linear Strauss process (see Equation 6.28), and *bottom row:* log-modulated Strauss process (see Equation 6.29). *Left column:* sphere with $r = (1/(4\pi))^{1/2}$, *middle column:* ellipsoid with $a = b = 0.8 \cdot (1/(4\pi))^{1/2}$, and *right column:* with ellipsoid with $a = b = 0.6 \cdot (1/(4\pi))^{1/2}$. Yellow on the surface indicates high intensity whilst the blue low intensity. Please note these images are not to scale: this refers to both the axes lengths and the intensities across the different ellipsoids and processes.

Shape	Intensity Type	Process Parameters	$\mathbb{E}[\mathbf{N}_{\mathbf{X}}(\mathcal{M})]$	Average ISE - CV	Average ISE - Cronie
Sphere	Pseudo-homogeneous	$(\beta, \alpha, R) = (400, 0.01, 0.1)$	326.9181	4.1859	10.0066
Sphere	Pseudo-homogeneous	$(\beta, \alpha, R) = (400, 0.1, 0.1)$	158.8051	1.9340	5.5295
Sphere	Pseudo-homogeneous	$(\beta, \alpha, R) = (400, 0.01, 0.2)$	230.4917	3.0766	6.8795
Sphere	Pseudo-homogeneous	$(\beta, \alpha, R) = (400, 0.1, 0.2)$	74.0678	2.6821	4.3987
Sphere	Linear	$(\beta, \alpha, R) = (400, 0.01, 0.1)$	177.1803	6.8957	5.9619
Sphere	Linear	$(\beta, \alpha, R) = (400, 0.1, 0.1)$	85.9409	4.5972	4.4432
Sphere	Linear	$(\beta, \alpha, R) = (400, 0.01, 0.2)$	124.8093	5.1109	4.5399
Sphere	Linear	$(\beta, \alpha, R) = (400, 0.1, 0.2)$	40.1244	5.3121	4.8132
Sphere	Modulated	$(\beta, \alpha, R) = (400, 0.01, 0.1)$	147.6968	32.9222	91.9720
Sphere	Modulated	$(\beta, \alpha, R) = (400, 0.1, 0.1)$	71.8542	20.9617	50.4006
Sphere	Modulated	$(\beta, \alpha, R) = (400, 0.01, 0.2)$	104.1532	27.3495	65.5846
Sphere	Modulated	$(\beta, \alpha, R) = (400, 0.1, 0.2)$	33.4601	18.4170	30.4539
Ellipsoid 1	Pseudo-homogeneous	$(\beta, \alpha, R) = (400, 0.01, 0.1)$	327.0679	3.9288	9.1548
Ellipsoid 1	Pseudo-homogeneous	$(\beta, \alpha, R) = (400, 0.1, 0.1)$	158.9274	2.0718	5.5708
Ellipsoid 1	Pseudo-homogeneous	$(\beta, \alpha, R) = (400, 0.01, 0.2)$	230.5472	3.5758	7.4606
Ellipsoid 1	Pseudo-homogeneous	$(\beta, \alpha, R) = (400, 0.1, 0.2)$	74.0699	2.6184	4.4366
Ellipsoid 1	Linear	$(\beta, \alpha, R) = (400, 0.01, 0.1)$	194.9879	6.7941	6.2850
Ellipsoid 1	Linear	$(\beta, \alpha, R) = (400, 0.1, 0.1)$	94.6502	4.9810	4.8639
Ellipsoid 1	Linear	$(\beta, \alpha, R) = (400, 0.01, 0.2)$	137.4484	6.8035	6.0505
Ellipsoid 1	Linear	$(\beta, \alpha, R) = (400, 0.1, 0.2)$	44.2165	5.2622	4.9809
Ellipsoid 1	Modulated	$(\beta, \alpha, R) = (400, 0.01, 0.1)$	176.8703	19.3731	41.3491
Ellipsoid 1	Modulated	$(\beta, \alpha, R) = (400, 0.1, 0.1)$	86.0105	13.6570	23.8703
Ellipsoid 1	Modulated	$(\beta, \alpha, R) = (400, 0.01, 0.2)$	125.0499	17.8348	30.9726
Ellipsoid 1	Modulated	$(\beta, \alpha, R) = (400, 0.1, 0.2)$	40.1884	12.7997	16.1993
Ellipsoid 2	Pseudo-homogeneous	$(\beta, \alpha, R) = (400, 0.01, 0.1)$	327.0782	4.8086	9.7270
Ellipsoid 2	Pseudo-homogeneous	$(\beta, \alpha, R) = (400, 0.1, 0.1)$	158.9183	2.6908	5.7366
Ellipsoid 2	Pseudo-homogeneous	$(\beta, \alpha, R) = (400, 0.01, 0.2)$	230.5326	4.2225	7.6968
Ellipsoid 2	Pseudo-homogeneous	$(\beta, \alpha, R) = (400, 0.1, 0.2)$	74.1534	3.9902	5.2951
Ellipsoid 2	Linear	$(\beta, \alpha, R) = (400, 0.01, 0.1)$	217.0637	8.1579	7.8578
Ellipsoid 2	Linear	$(\beta, \alpha, R) = (400, 0.1, 0.1)$	105.4746	5.4402	5.3606
Ellipsoid 2	Linear	$(\beta, \alpha, R) = (400, 0.01, 0.2)$	152.8796	7.1587	6.7224
Ellipsoid 2	Linear	$(\beta, \alpha, R) = (400, 0.1, 0.2)$	49.1915	5.8574	5.5281
Ellipsoid 2	Modulated	$(\beta, \alpha, R) = (400, 0.01, 0.1)$	228.2123	15.8179	16.1999
Ellipsoid 2	Modulated	$(\beta, \alpha, R) = (400, 0.1, 0.1)$	111.0509	11.6062	10.7698
Ellipsoid 2	Modulated	$(\beta, \alpha, R) = (400, 0.01, 0.2)$	161.3947	13.6333	13.1285
Ellipsoid 2	Modulated	$(\beta, \alpha, R) = (400, 0.1, 0.2)$	52.0897	9.9838	9.1551

Table 6.3: Average ISE for Strauss simulation study. The sphere has radius $(1/(4\pi))^{1/2}$. For Ellipsoid 1 and 2 the minor axis length along the x- and y-axis are the same and $0.8 \cdot (1/(4\pi))^{1/2}$ and $0.6 \cdot (1/(4\pi))^{1/2}$ respectively whilst the minor axis length along the z-axis is adjusted to ensure unit surface area of the ellipsoid. The columns for intensity type and process parameters correspond to Equations 6.27-6.29. The average ISE is divided by the expected number of points for each process to put them on a comparable scale.

7

Summary statistics for multitype point patterns on convex surfaces

In this chapter we discuss our work Ward et al. [2021c]. We extend the analysis of multivariate point patterns from Euclidean space to the sphere and then towards convex shapes, with the primary focus being to construct functional summary statistics capable of providing evidence of dependence between components of an observed multitype point pattern. By using the Mapping Theorem we can shift analysis of a point process from their original metric space to the sphere where we can benefit from rotational symmetries in order to construct functional summary statistics, such as the cross $K-$function, in order to detect statistical interactions between components of the process. We explore both homogeneous and inhomogeneous processes and their ability to determine attraction and repulsion in a number of examples on sphere and ellipsoids.

7.1 Introduction

Analysis of multivariate point patterns has focused predominantly on those that arise on a Euclidean space with a rich literature available for exploratory data analysis and modelling fitting [Møller and Waagepetersen, 2003, Baddeley et al., 2015, Chiu et al., 1995]. In many applications multivariate point patterns frequently arise on surfaces that are not adequately modelled by the geometry of a Euclidean space. For example, in microbiology, researchers are interested in the spatial relationship between different cellular membrane bound molecules where the cells are more appropriately modelled by ellipsoids or capsules. In such applications current methodologies that assume the process is Euclidean would be erroneous as it

does not respect the underlying geometry with which the point process resides.

Furthermore, one may be tempted to simply exchange the Euclidean metric for a metric that respects the geometry of the original space, for example the distance of the shortest path between any two points lying on the surface commonly: the geodesic distance. Although intuitive this could lead to an improper theoretical treatment of the underlying process. This idea is explored by Ward et al. [2021b] in the setting of univariate point patterns and constructing functional summary statistics for Poisson processes. In their work they argue that simply substituting the Euclidean metric for a more appropriate one is not a sufficient solution to testing for CSR. This arises from the fact that the original space may not present an infinite number of isometries which are necessary to have analogous notions of distributional invariance such as stationarity and isotropy and hence well-defined summary statistics cannot be easily constructed. This challenge is overcome by Ward et al. [2021b] by using the Mapping Theorem Kingman [1993] to construct summary statistics for Poisson processes, such as Ripley's K-function Ripley [1977]. This issue is also encountered for point patterns on linear networks where the irregularity of the geometry imposed by the space requires special consideration when handling point processes observed on such structures Ang et al. [2012], Rakshit et al. [2017].

These issues are also apparent in the multivariate setting where lacking notions of stationarity and isotropy can lead to improper theoretical treatment. A more practical issue would be how to sample replicates from a null hypothesis of independence when the process is homogeneous on the convex space. Consider, initially, the typical example of testing for statistical correlation between components of a bivariate pattern in $\mathbb{R}^2$. Under stationarity a popular approach is to use the toroidal shift method of Lotwick and Silverman [1982] to draw simulates under the null hypothesis of independence. These are then used to construct functional summary statistics which are compared against the functional summary statistic constructed from the observed pattern. Now suppose one observes a bivariate point pattern on some surface such that it is homogeneous. Notice that to correctly implement the method of Lotwick and Silverman [1982] it is implicitly assumed that there exists a transformation from the original space to a torus such that the mapping induces a process that has constant intensity function; this need not be the case for arbitrary closed surface. More precisely, in general it may not be possible to map a point process from an arbitrary convex shape to a torus (or some other shape that is equipped with a set of isomteries) such that the mapped process has a constant intensity function as is the case for point patterns observed in a rectangular window of $\mathbb{R}^2$. Therefore methodology akin to the toroidal shift of Lotwick and Silverman [1982] is not trivially extendable to processes on asymmetric surfaces.

Limited research has been conducted in understanding multivariate point processes outside of $\mathbb{R}^d$. In the univariate setting there is extensive work for analysing point patterns on the sphere including Robeson et al. [2014], Lawrence et al. [2016], Møller and Rubak [2016], Cuevas-Pacheco and Møller [2018] and linear networks Ang et al. [2012], Rakshit et al. [2017], McSwiggan et al. [2017], Moradi et al. [2018], Rakshit et al. [2019], whilst univariate analysis outside of these spaces has been considered by Ward et al. [2021b]. The work of Jun et al. [2019] presents an approach to modelling multivariate point processes on a sphere using multivariate log Gaussian Cox processes (LGC processes): their work revolves around the analysis of rainfall on a global scale.

In this chapter we outline how exploratory data analysis can be conducted for multitype point processes on the surfaces of bounded shapes in $\mathbb{R}^3$ with particular focus on detecting statistical interactions between components of the process. Section 7.2 outlines the preliminary material and notation used for the remainder of this chapter. Section 7.3 describes briefly how statistical relationships between components of a multivariate pattern can be detected on the sphere whilst Section 7.5 uses the foundations built in Section 7.3 to analyse multivariate processes on more general surfaces in $\mathbb{R}^3$. In both Sections 7.3 and 7.5 we demonstrates our approach on spheres and ellipsoids respectively for processes which exhibit independence, attractiveness and regularity between components.

7.2 Preliminaries

In this section we outline the necessary spatial theory and notation used throughout this chapter and follows a similar framework to Cronie and van Lieshout [2016], Iftimi et al. [2019]. We refer the reader to Section 4.1 for the framework of the geometry considered in this chapter. Furthermore, we can consider multitype point patterns as a special case of marked point processes and therefore develop the theory under this more general setting.

7.2.1 Marked Point Processes

Following the notation of Møller and Waagepetersen [2003], we define $\lambda_{\mathbb{D}}(\mathbf{x})$ as the surface measure of the surface of the convex shape $\mathbb{D}$, which is endowed with its geodesic distance and Borel σ-algebra, $\mathcal{B}(\mathbb{D})$. Let $\mathcal{M}$ be a Polish space equipped with its Borel σ-algebra $\mathcal{B}(\mathcal{M})$. Define the *ground process* X_g to be a univariate locally finite simple (that is points do not coincide) point process on $\mathbb{D}$ and then, informally, if for each point $\mathbf{x}_i \in X_g$ we assign a mark m_i the resulting process X is said to be a marked point process on $\mathbb{D}$ with mark space

$\mathcal{M}$. More precisely, let $\mathcal{B}(\mathbb{D} \times \mathcal{M})$ be the Borel σ-algebra on the product space of $\mathbb{D}$ and $\mathcal{M}$ and define $N_{lf} = \{x \subset \mathbb{D} \times \mathcal{M} : |x \cap A| < \infty \text{ for all bounded } A \in \mathcal{B}(\mathbb{D} \times \mathcal{M})\}$. Define $\mathcal{N}$ to be the σ-algebra generated by the sets $\{x \in N_{lf} : |x \cap A| = m\}$ for all bounded $A \in \mathcal{B}(\mathbb{D} \times \mathcal{M})$ and $m \in \mathbb{N}$. A marked point process on $\mathbb{D}$ with mark space $\mathcal{M}$ is a measurable mapping from some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ into the measurable space $(N_{lf}, \mathcal{N})$. The special case of multitype point processes arises when $\mathcal{M}$ is finite, i.e. $\mathcal{M} = \{1, 2, \ldots, k\}$ for $k > 0$. Let $M \subseteq \mathcal{M}$ then we denote $X_M = \{(\mathbf{x}, m) \in X : m \in M\}$.

In order to integrate over $\mathbb{D} \times \mathcal{M}$ we require some reference measure over $(\mathbb{D} \times \mathcal{M}, \mathcal{B}(\mathbb{D} \times \mathcal{M}))$. For this purpose we shall take the product measure $\lambda_{\mathbb{D}} \otimes \nu$ to be our reference measure where ν is some suitable chosen reference measure on the mark space. Examples of ν can be probability measures over compact subspaces of $\mathbb{R}$ or a counting measure when $\mathcal{M} = \{1, 2, \ldots, k\}$, k a positive integer, further details can be found in e.g. Chiu et al. [1995], Daley and Vere-Jones [2003]. When well defined we can write $(\lambda_{\mathbb{D}} \otimes \nu)(D, M) = \lambda_{\mathbb{D}}(D)\nu(M)$ where $D \in \mathcal{B}(\mathbb{D})$ and $M \in \mathcal{B}(\mathcal{M})$.

Let $A = D \times M \in \mathcal{B}(\mathbb{D} \times \mathcal{M})$, then we define $N_X(A) = N_X(D \times M)$ to be the counting measure of X, i.e.

$$N_X(D \times M) = \sum_{(\mathbf{x}, m) \in X} \mathbb{1}[(\mathbf{x}, m) \in D \times M].$$

That is, N_X counts the number of points in X with spatial location lying in B and mark M. The expectation of N_X is denoted the intensity measure: $\mu(A) = \mu(D \times C) = \mathbb{E}[N_X(D \times M)]$. Throughout this paper we shall make the further assumption that μ is absolutely continuous with respect to $\lambda_{\mathbb{D}} \otimes \nu$, then by the Radon-Nikodyn theorem there exists a unique density $\rho : \mathbb{D} \times \mathcal{M} \mapsto \mathbb{R}^+$ such that

$$\mu(D \times M) = \int_D \int_M \rho(\mathbf{x}, m) \lambda_{\mathbb{D}}(d\mathbf{x}) \nu(dm),$$

with ρ denoting the intensity function of X. We note that by construction the ground process of X, X_g, has intensity measure $\mu_g(D) = \mu(D \times \mathcal{M})$ for $D \subset \mathbb{D}$ and so for any fixed $M \subset \mathcal{M}$ we have that $\mu(D \times M) \leq \mu(D \times \mathcal{M}) = \mu_g(D)$. Hence μ is absolutely continuous with respect to μ_g and therefore for any fixed $M \subset \mathcal{M}$ we have

$$\mu(D \times M) = \int_B K^{\mathbf{x}}(M) \mu_g(d\mathbf{x}).$$

Here $K^{\mathbf{x}}$ is a probability measure and can be interpreted as the probability of a point $\mathbf{x}$ having a mark in M given that $\mathbf{x}$ is an event of X_g. Further, by assumption that μ is

absolutely continuous with respect to $\lambda_{\mathbb{D}} \otimes \nu$ we have that μ_g is absolutely continuous with respect to $\lambda_{\mathbb{D}}$ and thus there exists intensity function $\rho_g : \mathbb{D} \mapsto \mathbb{R}^+$. If we also assume that $K^{\mathbf{x}}$ is absolutely continuous with respect to ν then we have

$$\mu(D \times M) = \int_B \int_M f^{\mathbf{x}}(m)\rho_g(\mathbf{x})\lambda_{\mathbb{D}}(d\mathbf{x})\nu(dm),$$

where $f^{\mathbf{x}}$ is the density of $K^{\mathbf{x}}$ with respect to ν. Hence we have the identity $\rho(\mathbf{x}, m) = f^{\mathbf{x}}(m)\rho_g(\mathbf{x})$.

We can define higher order intensity functions as densities of *factorial moment measures* with respect to the n-fold reference measure $(\lambda_{\mathbb{D}} \otimes \nu)^n$. The n^{th}-order factorial moment measure, $\alpha^{(n)}$, of X is defined as

$$\alpha^{(n)}(B_1, \ldots, B_n) = \mathbb{E} \sum_{(\mathbf{x}_1,m_1),\ldots,(\mathbf{x}_n,m_n)\in X}^{\neq} \mathbb{1}[(\mathbf{x}_1, m_1) \in B_1, \ldots, (\mathbf{x}_n, m_n) \in B_n],$$

where $B_i \in \mathcal{B}(\mathbb{D} \times \mathcal{M})$ for $i = 1, \ldots, n$ and $\sum^{\neq}$ is the sum over pairwise distinct elements. Notice that for $n = 1$, $\alpha^{(1)} = \alpha = \mu$. If $\alpha^{(n)}$ is the absolutely continuous with respect to $(\lambda_{\mathbb{D}} \otimes \nu)^n$ then there exists densities $\rho^{(n)} : (\mathbb{D} \times \nu)^n \mapsto \mathbb{R}^+$ such that for any measurable function $f : (\mathbb{D} \times \nu)^n \mapsto \mathbb{R}^+$ they satisfy the following

$$\mathbb{E} \sum_{(\mathbf{x}_1,m_1),\ldots,(\mathbf{x}_n,m_n)\in X}^{\neq} f((\mathbf{x}_1, m_1), \ldots, (\mathbf{x}_n, m_n)) =$$
$$\int_{(\mathbb{D}\times\mathcal{M})^n} \cdots \int f((\mathbf{x}_1, m_1), \ldots, (\mathbf{x}_n, m_n))\rho^{(n)}((\mathbf{x}_1, m_1), \ldots, (\mathbf{x}_n, m_n)) \prod_{i=1}^{n} \lambda_{\mathbb{D}}(d\mathbf{x}_i)\nu(dm_i),$$

where $\rho^{(n)}$ denotes the n^{th}-order intensity function and $\rho^{(n)}((\mathbf{x}_1, m_1), \ldots, (\mathbf{x}_n, m_n)) \prod_{i=1}^{n} \lambda_{\mathbb{D}}(d\mathbf{x}_i)\nu(dm_i)$ can be heuristically considered as the probability of finding points of X in infinitesimal volumes $(d\mathbf{x}_i, dm_i)$ for $i = 1, \ldots, n$. This formula is known as *Campbell's formula for marked processes* and for the remainder of this chapter we shall suppose the existence of the n^{th}-order intensity functions for all n.

Similar to $n = 1$ we can also define $K^{\mathbf{x}_1,\ldots,\mathbf{x}_n}$ to be the n^{th}-order mark distributions. By assumption of the existence of $\rho^{(n)}$ we also obtain the n^{th}-order intensity functions of the ground process X_g, $\rho_g^{(n)}$ and if we assume that $K^{\mathbf{x}_1,\ldots,\mathbf{x}_n}$ is absolutely continuous with respect to the n-fold measure ν^n with density $f^{\mathbf{x}_1,\ldots,\mathbf{x}_n} : \mathcal{M}^n \mapsto \mathbb{R}^+$ then we have the identity,

$$\rho^{(n)}((\mathbf{x}_1, m_1), \ldots, (\mathbf{x}_n, m_n)) = f^{\mathbf{x}_1,\ldots,\mathbf{x}_n}(m_1, \ldots, m_n)\rho_g^{(n)}((\mathbf{x}_1, m_1), \ldots, (\mathbf{x}_n, m_n)).$$

Further to this we say that X is independently marked if $f^{\mathbf{x}_1,\ldots,\mathbf{x}_n}(m_1,\ldots,m_n) = \prod_{i=1}^{n} f^{\mathbf{x}_i}(m_i)$ for all $n \in \mathbb{N}$.

Having defined the n^{th}-order intensity functions we can now define a central property of point processes: *the point correlation function of a marked process* (PCF), h (c.f. Equation 2.7 for PCF of unmarked processes). The PCF is defined as

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

Under a Poisson assumption with independent marking it can easily be shown that $h((\mathbf{x}, m_{\mathbf{x}}), (\mathbf{y}, m_{\mathbf{y}})) = 1$, Such a Poisson model typically serves as a benchmark of no interaction between events where if $h((\mathbf{x}, m_{\mathbf{x}}), (\mathbf{y}, m_{\mathbf{y}})) > 1$ it indicates clustering between events with marks m_1 and m_2 whilst inhibition holds if $h((\mathbf{x}, m_{\mathbf{x}}), (\mathbf{y}, m_{\mathbf{y}})) < 1$.

The n^{th}*-order correlation functions* ξ_n for $n \in \mathbb{N}$ are defined recursively (for example see White [1979], van Lieshout [2006]) and based on the n^{th}-order intensity functions. Set $\xi_1 = 1$, then for $n \geq 2$,

$$\sum_{k=1}^{l} \sum_{E_1,\ldots,E_l} \prod_{j=1}^{k} \xi_{|E_j|}\left(\{(\mathbf{x}_i, m_i) : i \in E_j\}\right) = \frac{\rho^{(n)}((\mathbf{x}_1, m_1), \ldots, (\mathbf{x}_n, m_n))}{\rho(\mathbf{x}_1, m_1) \cdots \rho(\mathbf{x}_n, m_n)},$$

where $\sum_{E_1,\ldots,E_l}$ is the sum over all possible l-sized partitions of the set $\{1,\ldots,n\}$ such that $E_j \neq \emptyset$.

The summary statistics developed in this paper are based on the *reduced Palm processes* of a marked point process X. We can define the reduced Palm process as the point process which follows the probability measure defined by the Radon-Nikodyn derivative of the *reduced Campbell measure*, $C^!(A \times N) = \mathbb{E}\sum_{(\mathbf{x},m)\in X} \mathbb{1}[((\mathbf{x}, m),\ X \setminus (\mathbf{x}, m)) \in A \times N]$ where $A \in \mathcal{B}(\mathbb{D} \times \mathcal{M})$, and $N \in \mathcal{N}$, with respect to μ (see for example Møller and Waagepetersen [2003, Appendix C.2]). Since $C^! << \mu$ we have by the Radon-Nikodyn theorem

$$C^!(A \times N) = \int_{B\times M} P^!_{(\mathbf{x},m)}(N)\mu(d\mathbf{x}, dm),$$

where $(\mathbf{x}, m) \in \mathbb{D} \times \mathcal{M}$. Here, $P^!_{(\mathbf{x},m)}$ is the Radon-Nikodyn derivative $dC^!/d\mu$ and defines a probability measure called the reduced Palm measure, see e.g. Møller and Waagepetersen [2003, Appendix C.2]. We frequently use $P^!_{(\mathbf{x},m)}(N)$ as a short hand for $P(X^!_{(\mathbf{x},m)} \in N)$ for $N \in \mathcal{N}$. The process $X^!_{(\mathbf{x},m)}$ that follows $P^!_{(\mathbf{x},m)}$ is referred to as the reduced Palm process of X. Heuristically, $P^!_{(\mathbf{x},m)}$ can be considered the conditional distribution of X given that

$(\mathbf{x}, m) \in X$. Based on this we also have the *Campbell-Mecke formula*

$$\mathbb{E} \sum_{(\mathbf{x},m)\in X}^{\neq} f((\mathbf{x}, m), X \setminus (\mathbf{x}, m)) = \int_{\mathbb{D}\times\mathcal{M}} \mathbb{E}^!_{(\mathbf{x},m)}[f((\mathbf{x}, m), X)]\mu(d\mathbf{x}, dm),$$

where $\mathbb{E}^!_{(\mathbf{x},m)}$ is the expectation under the measure $P^!_{(\mathbf{x},m)}$. Following Cronie and van Lieshout [2016] we can also define the *ν-averaged Palm measure* with respect to $M \in \mathcal{B}(\mathcal{M})$ where $\nu(M) > 0$ as,

$$P^!_{\mathbf{z},M}(N) = \frac{1}{\nu(M)} \int_M P^!_{(\mathbf{x},m)}(N)\nu(dm), \quad N \in \mathcal{N}. \tag{7.1}$$

$P^!_{\mathbf{z},M}$ defines a probability measure since $0 \leq P^!_{(\mathbf{x},m)}(\cdot) \leq 1$ and it may be interpreted as the probability of X given that X has a point at $\mathbf{x}$ with mark in M. Moreover, if we substitute the reference measure ν with a probability measure $\nu_{\mathcal{M}}$ over $\mathcal{M}$ then we have that,

$$P^!_{\mathbf{z},M}(N) = \frac{1}{\nu_{\mathcal{M}}(M)} \int_M P^!_{(\mathbf{x},m)}(N)\nu_{\mathcal{M}}(dm)$$

which may be interpreted as the conditional distribution of $Y \setminus (\mathbf{x} \times \mathcal{M})$ given that $Y \cap (\mathbf{x} \times M) \neq \emptyset$ [Cronie and van Lieshout, 2016].

As an example consider X as a multitype point process with $\mathcal{M} = \{1, \ldots, k\}$ and reference measure ν being the counting measure, then taking $M = \{i\}$ for some $i \in \{1, \ldots, k\}$ we have that,

$$P^!_{\mathbf{x},M}(N) = \frac{1}{\nu(i)}\nu(i)P^!_{(\mathbf{x},i)}(N) = P^!_{(\mathbf{x},i)}(N).$$

Further, based on Equation 7.1 we can define the expectation with respect to the probability measure $P^!_{(\mathbf{x},m)}$ as

$$\mathbb{E}^!_{\mathbf{x},M}[f(X)] = \frac{1}{\nu(M)} \int_M \mathbb{E}^!_{(\mathbf{x},m)}[f(X)]\nu(dm), \tag{7.2}$$

for measurable, non-negative functions f.

In order to define upcoming summary functional statistics we require the *generating functional of a marked point process*, G, of a point process which uniquely characterises X [Møller and Waagepetersen, 2003]. Let $u : \mathbb{D} \times \mathcal{M} \mapsto [0, 1]$ be a measurable function with bounded support then the generating functional is defined as

$$G(u) = \mathbb{E} \prod_{(\mathbf{x},m)\in X} (1 - u(\mathbf{x}, m)).$$

Further to this, given that the n^{th}-order intensity functions exists for all n and that the

following series is convergent the generating functional can be represented as an infinite series [Cronie and van Lieshout, 2016],

$$G(u) = 1 + \sum_{n=1}^{\infty} \int_{\mathbb{D}\times\mathcal{M}} \cdots \int_{\mathbb{D}\times\mathcal{M}} \rho^{(n)}((\mathbf{x}_1, m_1), \ldots, (\mathbf{x}_n, m_n)) \prod_{i=1}^{n} u(\mathbf{x}_i, m_i) \lambda_{\mathbb{D}}(d\mathbf{x}_i) \nu(m_i),$$

where we take the convention that an empty product is 1. Further to this, using Equation 7.2, we can define the generating functional $G^!_{\mathbf{x},M}$ with respect to $P^!_{\mathbf{x},M}$ as

$$G^!_{\mathbf{x},M}(u) = \frac{1}{\nu(M)} \int_M \mathbb{E}^!_{(\mathbf{x},m)} \left[\prod_{(\mathbf{y},n)\in X} (1 - u(\mathbf{y}, n)) \right] \nu(dm).$$

7.3 Summary statistics for isotropic processes on $\mathbb{S}^2$

In this section we shall discuss how to construct summary statistics for isotropic marked point processes on the sphere, extending the theory developed by van Lieshout [2006] for Euclidean point processes to spherical ones. We define the second order reduced moment measure for a spheroidal marked point process and consider how functional summary statistics can be constructed based on the work of van Lieshout [2006]. Concentrating specifically on multivariate point processes, we demonstrate, based on estimators of the summary statistics, that independence, attraction and replusion between components can be detected between components.

7.3.1 Definitions

As in Møller and Rubak [2016] denote $\mathcal{O}(3)$ to be the 3×3 rotation matrices and define, for $O \in \mathcal{O}(3)$, $OX = \{(O\mathbf{x}, m) : (\mathbf{x}, m) \in X\}$ to be the rotation of the point process X by O. We then say that X is isotropic if OX and X are identically distributed for any $O \in \mathcal{O}(3)$. Notice that by isotropy, for fixed $M \in \mathcal{M}$ we have that $\mu(D \times M) = \mu(OD \times M)$ for any $D \subseteq \mathbb{S}^2$ and $O \in \mathcal{O}(3)$ is a rotationally invariant measure on the sphere. Therefore $\mu(D \times M) = \eta_M \lambda_{\mathbb{S}^2}(D)$ where η_M is a positive constant potentially depending on M, i.e. is proportional to the surface measure on the sphere [Møller and Waagepetersen, 2003]. By isotropy of X the ground process is also isotropic and thus the ground process has constant

intensity ρ_g, hence $\mu(D \times M) = \int_D K^{\mathbf{x}}(M)\rho_g \lambda_{\mathbb{S}^2}(d\mathbf{x})$. Therefore,

$$\begin{aligned} 0 &= \eta_M \lambda_{\mathbb{S}^2}(D) - \int_D K^{\mathbf{x}}(M)\rho_g \lambda_{\mathbb{S}^2}(d\mathbf{x}) \\ &= \int_D \eta_M \lambda_{\mathbb{S}^2}(d\mathbf{x}) - \int_D K^{\mathbf{x}}(M)\rho_g \lambda_{\mathbb{S}^2}(d\mathbf{x}) \\ &= \int_D \eta_M - K^{\mathbf{x}}(M)\rho_g \lambda_{\mathbb{S}^2}(d\mathbf{x}). \end{aligned}$$

Since this must hold for any $D \subseteq \mathbb{S}^2$, the integrand must be 0 meaning $\eta_M = K^{\mathbf{x}}(M)\rho_g$. Therefore the mark distribution is independent of $\mathbf{x}$ and so the superscript can be dropped. For an isotropic process we thus have,

$$\mu(D \times M) = \rho_g K(M) \lambda_{\mathbb{S}^2}(D), \tag{7.3}$$

where K is a probability measure over $\mathcal{M}$. When discussing isotropic processes later we note that it is convenient to take the reference measure ν over $\mathcal{M}$ to be K.

We say that X is *homogeneous* if the intensity function of the ground process is constant, i.e. $\rho_g(\mathbf{x}) = \rho \in \mathbb{R}_+$. In this event the intensity function of X is given by $\rho(\mathbf{x}, m) = f^{\mathbf{x}}(m)\rho$, where $f^{\mathbf{x}}$ is the density of first order mark distribution with respect to the reference measure ν over $\mathcal{M}$. We say that X has *common mark distribution* if $K^{\mathbf{x}} \equiv K$ is independent of $\mathbf{x}$, and so $f^{\mathbf{x}} \equiv f$, which is the density of K with respect to ν. Further to this, if K coincides with ν then we have that $f \equiv 1$ and so $\rho(\mathbf{x}, m) = \rho_g(\mathbf{x})$. We also define the *unit rate (marked) Poisson process* to be a marked process which has a ground Poisson process with intensity one almost everywhere, and marks that are independent of location and IID with probability measure $P_{\mathcal{M}}$ over a mark space $\mathcal{M}$. Additionally if $P_{\mathcal{M}}$ admits a density $p_{\mathcal{M}}$ with respect to the reference measure ν then the unit rate Poisson process has intensity $\rho(\mathbf{x}, m) = p_{\mathcal{M}}(m)$ or if we take $P_{\mathcal{M}}$ to be the reference measure then $\rho(\mathbf{x}, m) = 1$.

Let X be isotropic, then it can be shown that $OX^!_{(\mathbf{x},m)}$ and $X^!_{(O\mathbf{x},m)}$ are identically distributed if $O \in \mathcal{O}(3)$. We define the origin on $\mathbb{S}^2$ to be the North pole, i.e. $\mathbf{o} = (0, 0, 1)^T$ and define $O_{\mathbf{x}}$ to be the unique rotation such that $O_{\mathbf{x}}\mathbf{o} = \mathbf{x}$ for which the axis of rotation is orthogonal to the geodesic between $\mathbf{x}$ and $\mathbf{o}$ then $O_{\mathbf{x}}X^!_{(\mathbf{o},m)}$ and $X^!_{(\mathbf{x},m)}$ are identically distributed. This is the marked version of Møller and Rubak [2016, Proposition 1] and can easily be proven under the mild condition of X being absolutely continuous to the unit rate Poisson over mark space $\mathcal{M}$. Therefore from Equation 7.1 we have that,

$$P(X^!_{\mathbf{o},M} \in \cdot) = P(O^T_{\mathbf{x}}(X^!_{\mathbf{x},M}) \in \cdot),$$

for almost all $\mathbf{x} \in \mathbb{S}^2$. Similar to [Chiu et al., 1995, pg. 125] we refer to $P(X^!_{\mathbf{x},M} \in \cdot)$ as the *reduced Palm distribution with respect to the mark set M* (see also van Lieshout [2006]).

In order to be able to construct functional summary statistics we need to impose certain assumptions on our point process such that they are well defined. In $\mathbb{R}^d$, these assumptions describe how events of X_C located in $A \subset \mathbb{R}^d$ interact with other events of X_E located in $B \subset \mathbb{R}^d$ where $C, E \subset \mathcal{M}$. For Euclidean processes a common assumption imposed is for X to be *second order intensity reweighted stationary for marked processes* (SOIRWS), see for example Cronie and van Lieshout [2016], which extends an analogous definition for unmarked point processes discussed by Baddeley et al. [2000] (cf. Definition 2.3.3 for the unmarked version). We define a similar property for spheroidal processes, that is *second order intensity reweighted isotropic for marked processes* (SOIRWI), (cf. Definition 3.4.1 for the unmarked version). Define $\mathcal{K}^{CE}(B)$ for $A, B \subset \mathbb{S}^2$ and $C, E \subseteq \mathcal{M}$ by

$$\lambda_{\mathbb{S}^2}(A)\nu(C)\nu(E)\mathcal{K}^{CE}(B) = \mathbb{E}\sum_{(\mathbf{x},m_\mathbf{x}),(\mathbf{y},m_\mathbf{y})\in X}^{\neq} \frac{\mathbb{1}[(\mathbf{x},m_\mathbf{x}) \in A \times C, (O_\mathbf{x}^T\mathbf{y}, m_\mathbf{y}) \in B \times E]}{\rho(\mathbf{x},m_\mathbf{x})\rho(\mathbf{y},m_\mathbf{y})}. \tag{7.4}$$

We say that X is SOIRWI if $\mathcal{K}^{CE}$ does not depend on A. If we suppose that the PCF is isotropic, i.e. $h((\mathbf{x},m_\mathbf{x}),(\mathbf{y},m_\mathbf{y})) = h((O\mathbf{x},m_\mathbf{x}),(O\mathbf{y},m_\mathbf{y}))$ for any $O \in \mathcal{O}(3)$ then by the Campbell-Mecke theorem it is easily shown that,

$$\begin{aligned}\nu(C)\nu(E)\mathcal{K}^{CE}(B) &= \frac{1}{\lambda_{\mathbb{S}^2}(A)}\int_{A\times C}\int_{B\times E} h((\mathbf{x},m_\mathbf{x}),(\mathbf{y},m_\mathbf{y}))\lambda_{\mathbb{S}^2}(d\mathbf{x})\nu(dm_\mathbf{x})\lambda_{\mathbb{S}^2}(d\mathbf{y})\nu(dm_\mathbf{y}) \\ &= \frac{1}{\lambda_{\mathbb{S}^2}(A)}\int_{A\times C}\int_{B\times E} h((O_\mathbf{x}^T\mathbf{x},m_\mathbf{x}),(O_\mathbf{x}^T\mathbf{y},m_\mathbf{y}))\lambda_{\mathbb{S}^2}(d\mathbf{x})\nu(dm_\mathbf{x})\lambda_{\mathbb{S}^2}(d\mathbf{y})\nu(dm_\mathbf{y}).\end{aligned}$$

Letting $\mathbf{z} = O_\mathbf{x}^T\mathbf{y}$ and relabelling $m_\mathbf{x} = m_1$ and $m_\mathbf{y} = m_2$, we have

$$\begin{aligned}\nu(C)\nu(E)\mathcal{K}^{CE}(B) &= \frac{1}{\lambda_{\mathbb{S}^2}(A)}\int_{A\times C}\int_{B\times E} h((\mathbf{o},m_1),(\mathbf{z},m_2))\lambda_{\mathbb{S}^2}(d\mathbf{x})\nu(dm_1)\lambda_{\mathbb{S}^2}(d\mathbf{z})\nu(dm_2) \\ &= \int_C\int_{B\times E} h((\mathbf{o},m_1),(\mathbf{z},m_2))\lambda_{\mathbb{S}^2}(d\mathbf{x})\nu(dm_1)\lambda_{\mathbb{S}^2}(d\mathbf{z})\nu(dm_2).\end{aligned}$$

Therefore a marked point processes that has an isotropic PCF is SOIRWI. Isotropic processes are also SOIRWI. For the purpose of this chapter whenever we refer to a point process as being SOIRWI we shall suppose that their PCF is isotropic as opposed to the more general definition prescribed by Equation 7.4.

We can now define our functional summary statistics for isotropic marked point processes,

extending the work of van Lieshout [2006]. Let $C, E \subseteq \mathcal{M}$ then

$$F^E(r) = P(X \cap (B_{\mathbb{S}^2}(\mathbf{o}, r) \times E) \neq \emptyset) \tag{7.5}$$

$$D^{CE}(r) = P(X^!_{\mathbf{o},C} \cap (B_{\mathbb{S}^2}(\mathbf{o}, r) \times E) \neq \emptyset) \tag{7.6}$$

$$J^{CE}(r) = \frac{1 - D^{CE}(r)}{1 - F^E(r)} \quad \text{for } F^E(r) < 1 \tag{7.7}$$

$$K^{CE}(r) = \frac{1}{\lambda_{\mathbb{S}^2}(A)\nu(C)\nu(E)} \mathbb{E} \sum_{(\mathbf{x},m_\mathbf{x}),(\mathbf{y},m_\mathbf{y}) \in X}^{\neq} \frac{\mathbb{1}[(\mathbf{x}, m_\mathbf{x}) \in A \times C, (O_\mathbf{x}^T \mathbf{y}, m_\mathbf{y}) \in B_{\mathbb{S}^2}(\mathbf{o}, r) \times E]}{\rho(\mathbf{x}, m_\mathbf{x})\rho(\mathbf{y}, m_\mathbf{y})}, \tag{7.8}$$

for all $r \in [0, \pi]$, and where K^{CE} does not depend on $A \subseteq \mathbb{S}^2$ since the process is assumed isotropic. Note that $K^{CE}(r)$ is the marked version of the *multitype cross K function*, see for example Møller and Waagepetersen [2003] and that $K^{CE}(r) = \mathcal{K}^{CE}(B_{\mathbb{S}^2}(\mathbf{o}, r))$. Further note that under an isotropic assumption we have that $\rho(\mathbf{x}, m) = f(m)\rho_g$, where ρ_g is the constant intensity of the ground process and f is the density of K with respect to ν from Equation 7.3. Therefore $\rho(\mathbf{x}, m)$ can be subsituted in the definition of K^{CE}. Additionally, by isotropy, none of these functional summary statistics depend on the typical point $\mathbf{o}$. For example, consider $F^E(r; \mathbf{x}) = P(X \cap (B_{\mathbb{S}^2}(\mathbf{x}, r) \times E) \neq \emptyset)$ then

$$\begin{aligned} F^E(r; \mathbf{x}) &= P(X \cap (B_{\mathbb{S}^2}(\mathbf{x}, r) \times E) \neq \emptyset) \\ &= P(O_\mathbf{x}^T X \cap (B_{\mathbb{S}^2}(O_\mathbf{x}^T \mathbf{x}, r) \times E) \neq \emptyset) \\ &= P(X \cap (B_{\mathbb{S}^2}(\mathbf{o}, r) \times E) \neq \emptyset) \\ &= F^E(r). \end{aligned}$$

A similar argument can be used for K^{CE}, D^{CE} and hence J^{CE} does not depend on the typical point either. In general $J^{CE} \neq J^{EC}$ whilst $K^{CE} = K^{EC}$.

Let $C, E \subset \mathcal{M}$ be such that $C \cap E = \emptyset$. The following proposition describes how the functional summary statistics behave under independence and can be considered marked spheroidal analogues to the multitype Euclidean results given by van Lieshout and Baddeley [1999] and Møller and Waagepetersen [2003].

Proposition 7.3.1. *Let X be a marked isotropic spheroidal point process. Let $C, E \subset \mathcal{M}$ such that $C \cap E = \emptyset$ and suppose that X_C and X_E are independent. Then,*

$$\begin{aligned} D^{CE}(r) &= F^E(r) \\ J^{CE}(r) &= 1 \end{aligned}$$

$$K^{CE}(r) = 2\pi(1 - \cos(r)),$$

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

Proof. See Appendix E.1. □

7.3.2 Estimating functional summary statistics

In order to draw any conclusions concerning the relationship between X_C and X_E we need to be able to estimate these functional summary statistics. Following van Lieshout [2006], suppose that we observe the process through some window $W \subset \mathbb{S}^2$ such that $\lambda_{\mathbb{S}^2}(W) > 0$ and that the reference measure and probability measure over $\mathcal{M}$ coincide (thus $\rho(\mathbf{x}, m) = \rho_g$ where ρ_g is the constant intensity of the ground process), we can construct estimates for them as

$$\begin{aligned}
1 - \hat{D}^{CE}(r) &= \frac{1}{\rho_g \lambda_{\mathbb{S}^2}(W_{\ominus r})\nu(C)} \\
&\qquad \sum_{(\mathbf{x}, m_{\mathbf{x}}) \in X} \mathbb{1}[(\mathbf{x}, m_{\mathbf{x}}) \in W_{\ominus r} \times C] \prod_{(\mathbf{y}, m_{\mathbf{y}}) \in X} \Big(1 - \mathbb{1}[d_{\mathbb{S}^2}(\mathbf{x}, \mathbf{y}) < r, m_{\mathbf{y}} \in E]\Big), \\
1 - \hat{F}^{E}(r) &= \frac{1}{|I_{W_{\ominus r}}|} \sum_{\mathbf{p} \in I_{W_{\ominus r}}} \prod_{(\mathbf{x}, m) \in X} \Big(1 - \mathbb{1}[d_{\mathbb{S}^2}(\mathbf{p}, \mathbf{x}) \leq r, m \in E]\Big), \\
\hat{J}^{CE}(r) &= \frac{1 - \hat{D}^{CE}(r)}{1 - \hat{F}^{E}(r)} \quad \text{for } \hat{F}^{E}(r) < 1, \\
\hat{K}^{CE}(r) &= \frac{1}{\rho_g^2 \lambda_{\mathbb{S}^2}(W_{\ominus r})\nu(C)\nu(E)} \\
&\qquad \sum_{(\mathbf{x}, m_{\mathbf{x}}) \in X} \sum_{(\mathbf{y}, m_{\mathbf{y}}) \in X \setminus (\mathbf{x}, m_{\mathbf{x}})} \mathbb{1}[(\mathbf{x}, m_{\mathbf{x}}) \in W_{\ominus r} \times C, d_{\mathbb{S}^2}(\mathbf{x}, \mathbf{y}) \leq r, m_{\mathbf{y}} \in E],
\end{aligned}$$

where $A_{\ominus r}$ is the erosion of set $A \subset \mathbb{S}^2$ by distance $r \in \mathbb{R}_+$ and I is a finite grid of points taken over $\mathbb{S}^2$. A common transformation used for planar multitype point patterns is $L^{CE}(r) = \sqrt{(K^{CE}(r)/\pi)}$ which benefits from the fact that under independence $L^{CE}(r) \equiv r$ since $K^{CE}(r) = \pi r^2$ and under a Poisson hypothesis is approximately variance stabilised. Since $K^{CE}(r)$ is not of a linear form in r (see Proposition 7.3.1) we instead consider the transformation $P^{CE}(r) = \sqrt{K^{CE}(r)} - \sqrt{2\pi(1 - \cos(r))}$, and similarly to estimate P^{CE} we have $\hat{P}^{CE}(r) = \sqrt{\hat{K}^{CE}(r)} - \sqrt{2\pi(1 - \cos(r))}$. Furthermore, it is easy to show either unbiasedness or ratio-unbiasedness of these estimators.

Proposition 7.3.2. *Let X be a marked isotropic point process on $\mathbb{S}^2$ such that the same assumptions hold as in Proposition 7.3.2. Then the minus sampling estimators $\hat{F}^E$, $\hat{D}^{CE}$ and $\hat{K}^{CE}$ are unbiased whilst $\hat{J}^{CE}$ is ratio-unbiased when ρ_g, the intensity of the ground process, is known.*

Proof. See Appendix E.2. □

In the more likely event that ρ_g and potentially ν, the mark measure are unknown then $\rho_g\lambda_{\mathbb{S}^2}(W_{\ominus r})\nu(C)$ can be replaced by the unbiased estimator $N_X(W_{\ominus r}\times C)$ for $\hat{D}^{CE}$ [van Lieshout, 2006]. Using this plugin estimator leads to $\hat{D}^{CE}$ being ratio-unbiased and hence so is $\hat{J}^{CE}$. For $\hat{K}^{CE}$ it can be shown that

$$\rho_g^2\lambda_{\mathbb{S}^2}(W_{\ominus r})\nu(C)\nu(E)=\frac{\mathbb{E}[N_X(W_{\ominus r}\times C)]\mathbb{E}[N_X(W_{\ominus r}\times E)]}{\lambda_{\mathbb{S}^2}(W_{\ominus r})},$$

and so we suggest to use $N_X(W_{\ominus r}\times C)N_X(W_{\ominus r}\times E)/\lambda_{\mathbb{S}^2}(W_{\ominus r})$ as a plugin estimator for $\rho_g^2\lambda_{\mathbb{S}^2}(W_{\ominus r})\nu(C)\nu(E)$. In general, $N_X(W_{\ominus r}\times C)N_X(W_{\ominus r}\times E)/\lambda_{\mathbb{S}^2}(W_{\ominus r})$ is biased but in the event X_C and X_E are independent then unbiasedness is achieved.

In the event that the process is completely observed over $\mathbb{S}^2$ then minus sampling is not needed and $W_{\ominus r}$ can be replaced by $\mathbb{S}^2$. For the special case of a multitype point pattern where $\mathcal{M}=\{1,2,\ldots,k\}$ then $\rho(\mathbf{x},i)\nu(i)=\rho_i(\mathbf{x})$ where $i\in\mathcal{M}$ and ν is the referential counting measure over $\mathcal{M}$. We can then simplify the estimators; taking $C=\{i\}$ and $E=\{j\}$ with $i\neq j$ then $\rho_g\nu(C)=\rho_g\nu(i)=\rho_i$ and $\rho_g\nu(E)=\rho_g\nu(j)=\rho_j$, giving

$$1-\hat{D}^{ij}(r)=\frac{1}{\rho_i\lambda_{\mathbb{S}^2}(W_{\ominus r})}\sum_{\mathbf{x}\in X_i}\mathbb{1}[\mathbf{x}\in W_{\ominus r}]\prod_{\mathbf{y}\in X_j}\Big(1-\mathbb{1}[d_{\mathbb{S}^2}(\mathbf{x},\mathbf{y})<r]\Big) \tag{7.9}$$

$$1-\hat{F}^{j}(r)=\frac{1}{|I_{W_{\ominus r}}|}\sum_{\mathbf{p}\in I_{W_{\ominus r}}}\prod_{\mathbf{x}\in X_j}\Big(1-\mathbb{1}[d_{\mathbb{S}^2}(\mathbf{p},\mathbf{x})\leq r]\Big) \tag{7.10}$$

$$\hat{J}^{ij}(r)=\frac{1-\hat{D}^{ij}(r)}{1-\hat{F}^{j}(r)}\quad\text{for }\hat{F}^j(r)<1 \tag{7.11}$$

$$\hat{K}^{ij}(r)=\frac{1}{\rho_i\rho_j\lambda_{\mathbb{S}^2}(W_{\ominus r})}\sum_{\mathbf{x}\in X_i}\sum_{\mathbf{y}\in X_j}\mathbb{1}[\mathbf{x}\in W_{\ominus r},d_{\mathbb{S}^2}(\mathbf{x},\mathbf{y})\leq r]. \tag{7.12}$$

These are analagous to the summary statistics given in Møller and Waagepetersen [2003] for stationary Euclidean multitype processes. Using these functional summary statistics we can detect whether there is evidence of aggregation or repulsion between components of the

multitype processes. More precisely in the simple setting of a bivariate process $X = (X_1, X_2)$ and setting $C = \{1\}$ and $E = \{2\}$ then if the estimated $\hat{K}^{12}(r)$ is greater than $2\pi(1-\cos(r))$ this suggests that the components are not independent and are actually attracted to each other, whilst if $\hat{K}^{12}(r)$ is less than $2\pi(1-\cos(r))$ this suggests that the components repel each other. Similar conclusions can be obtained from $\hat{J}^{12}(r)$ and $\hat{J}^{21}(r)$ where values greater than 1 suggests repulsion between X_1 and X_2 whilst attraction is suggested if the values are less than 1. The intensity, ρ_i, is often unknown for $i \in \{1, \ldots, k\}$ and so plug-in estimators are necessary. In the isotropic multitype setting we can use the unbiased estimator,

$$\hat{\rho}_i = N_{X_i}(W)/\lambda_{\mathbb{S}^2}(W), \tag{7.13}$$

where W is our window of observation with $\lambda_{\mathbb{S}^2}(W) > 0$. Since $\hat{\rho}_i$ for $i \in \mathcal{M}$ is unbiased when used as plugins for Equations 7.9-7.12 $\hat{F}^{ij}$ is unbiased whilst $\hat{D}^{ij}, \hat{J}^{ij}$ and $\hat{K}^{ij}$ are ratio-unbiased.

7.3.3 Determining independence

The principal goal of this chapter is to detect presence of dependency between components in a multitype point pattern on a convex shape. Typically for Euclidean processes, estimates of functional summary statistics from the observed pattern are compared against those constructed under simulations drawn from a null hypothesis [Møller and Waagepetersen, 2003, Myllymäki et al., 2017]. Therefore determining independence ultimately comes down to how simulates from a null distribution can be sampled. Two approaches are commonly used for stationary multitype processes in $\mathbb{R}^d$,

1. The random toroidal shift approach developed by Lotwick and Silverman [1982], and
2. Simulating a specified null model, e.g. Poisson process (see for example Rajala et al. [2018]).

The approach of Lotwick and Silverman [1982] allows for the potentially unknown marginal distributions to be maintained when simulating the null. This approach can be adapted to multitype spheroidal patterns where random toroidal shifts are instead replaced by random rotations. There are though two major pitfalls to this method:

- Application of rotations requires a point pattern that has been observed over all of the sphere.

- For values of r close to π the estimates of our functional summary statistics for both the observed and null simulates are very close.

To see the second point consider $\hat{K}_X^{12}$ for an observed bivariate point pattern $X = (X_1, X_2)$. Suppose now that we randomly rotate X_1 to give a null simulate $\tilde{X} = (OX_1, X_2)$ where $O \in \mathcal{O}(3)$ is a random rotation. It is then easily shown that $\hat{K}_{\tilde{X}}^{12}(\pi) = \hat{K}_X^{12}(\pi) = 4\pi$, where Equation 7.13 is used as a plugin for ρ_i. Hence all null simulates will converge to the same value at $r = \pi$ and so the true variability of the functional summary statistics for the process is not accurately captured and for values of r close to π. As such envelope plots should be treated cautiously in this region.

The second approach imposes additional distribution assumptions on the null hypothesis, for example an independent Poisson assumption. Although stricter distributional assumptions have been made, previous work including Wiegand et al. [2012], Møller and Waagepetersen [2003], Rajala et al. [2018] have used this approach to provide evidence of dependency between components in a multitype process. This issue also highlights the difficulty of testing a hypotheses like independence in the context of spatial statistics with progress only possible with the aid of additional assumptions. Furthermore, assuming a completely specified null model allows for simulation from the null distribution even in the event of a partially observed spheroidal process. This approach also avoids falling into the potential issue of misinterpretation of the envelope plot for values of r close to π, unlike the random rotation approach [Lotwick and Silverman, 1982]. In the coming examples we will explore both these approaches for determining whether components of an observed multitype point pattern exhibit dependence or not.

7.3.4 Examples

In these examples we shall assume that the process is observed on the entirety of $\mathbb{S}^2$, and so minus sampling is not needed. Furthermore, all the processes will be bivariate and isotropic with extensions to additional components easily achieved. We shall consider four examples, first a Poisson process whose components are independent, then LGCPs which have either independent, attractive, or repulsive components. To assess the hypothesis of independence we shall construct simulation envelopes using either random rotations of the original observed process or simulate a Poisson process with the corresponding fitted intensity given by Equation 7.13. In each case we generate 199 simulates and construct 2.5% lower and 97.5% upper envelope. When generating null simulates using random rotations we randomly rotate only the second component keeping the first component fixed, whilst for the Poisson null we

simulate both the first and second components. We apply these similar simulation schemes for all examples considered in this chapter.

Poisson process

We simulate an isotropic bivariate Poisson process with independent components. The intensities are $\rho_1 = 10$ and $\rho_2 = 12$ for the first and second components respectively. Examples of the functional summary statistics are given in Figure 7.1. We can see that the observed functional summary statistics remains close the the theoretical value and within the simulation envelope.

LGCP

A LGCP can be used to model aggregation and repulsion between components of a multivariate process [Brix and Moller, 2001, Møller and Waagepetersen, 2003]. We first discuss some basic properties of multivariate LGCPs. Let $(Z_1, Z_2) : \mathbb{S}^2 \times \mathbb{S}^2 \mapsto \mathbb{R}_+ \times \mathbb{R}_+$ be a non-negative bivariate random field over the sphere such that each $Z_i, i = 1, 2$ is almost surely locally integrable. Then we say that $X = (X_1, X_2)$ is a Cox process on $\mathbb{S}^2$ if, conditioned on $Z = (Z_1, Z_2)$, X is a bivariate Poisson process with intensity function $\rho(\mathbf{x}, m) = Z_m(\mathbf{x})$ where $m \in \{1, 2\}$. We frequently refer to Z as the driving (random) field of X. We then say that X is a bivariate LGCP process on $\mathbb{S}^2$ if, conditioned on the intensity function $Z = (\exp(Y_1), \exp(Y_2))$, is a bivariate Poisson process, where $Y = (Y_1, Y_2)$ is a bivariate GRF on $\mathbb{S}^2$. The distribution of GRF Y is determined by its mean and covariance/cross-covariance functions,

$$\mu_i(\mathbf{x}) = \mathbb{E}[Y_i(\mathbf{x})], \quad c_{ij}(\mathbf{x}, \mathbf{y}) = \mathrm{Cov}(Y_i(\mathbf{x}), Y_j(\mathbf{y})),$$

for $i, j \in \{1, 2\}$. This in turn completely defines the distribution of X and the intensity function and pair correlation function are given as (see for example Møller and Waagepetersen [2003]),

$$\rho_i(\mathbf{x}) = \exp\left(\mu_i(\mathbf{x}) + \frac{c_{ii}(\mathbf{x}, \mathbf{y})}{2}\right), \quad g_{ij} = \exp(c_{ij}(\mathbf{x}, \mathbf{y})),$$

for $i, j \in \{1, 2\}$. On spheres, univariate LGCP are discussed in depth by Cuevas-Pacheco and Møller [2018] whilst multivariate LGCP are discussed by Jun et al. [2019]. Typically simplifying assumptions are made, for example the covariance/cross-covariance functions, c_{ij}, are isotropic, that is they only depends on the distance $d_{\mathbb{S}^2}(\mathbf{x}, \mathbf{y})$ between $\mathbf{x}$ and $\mathbf{y}$. In

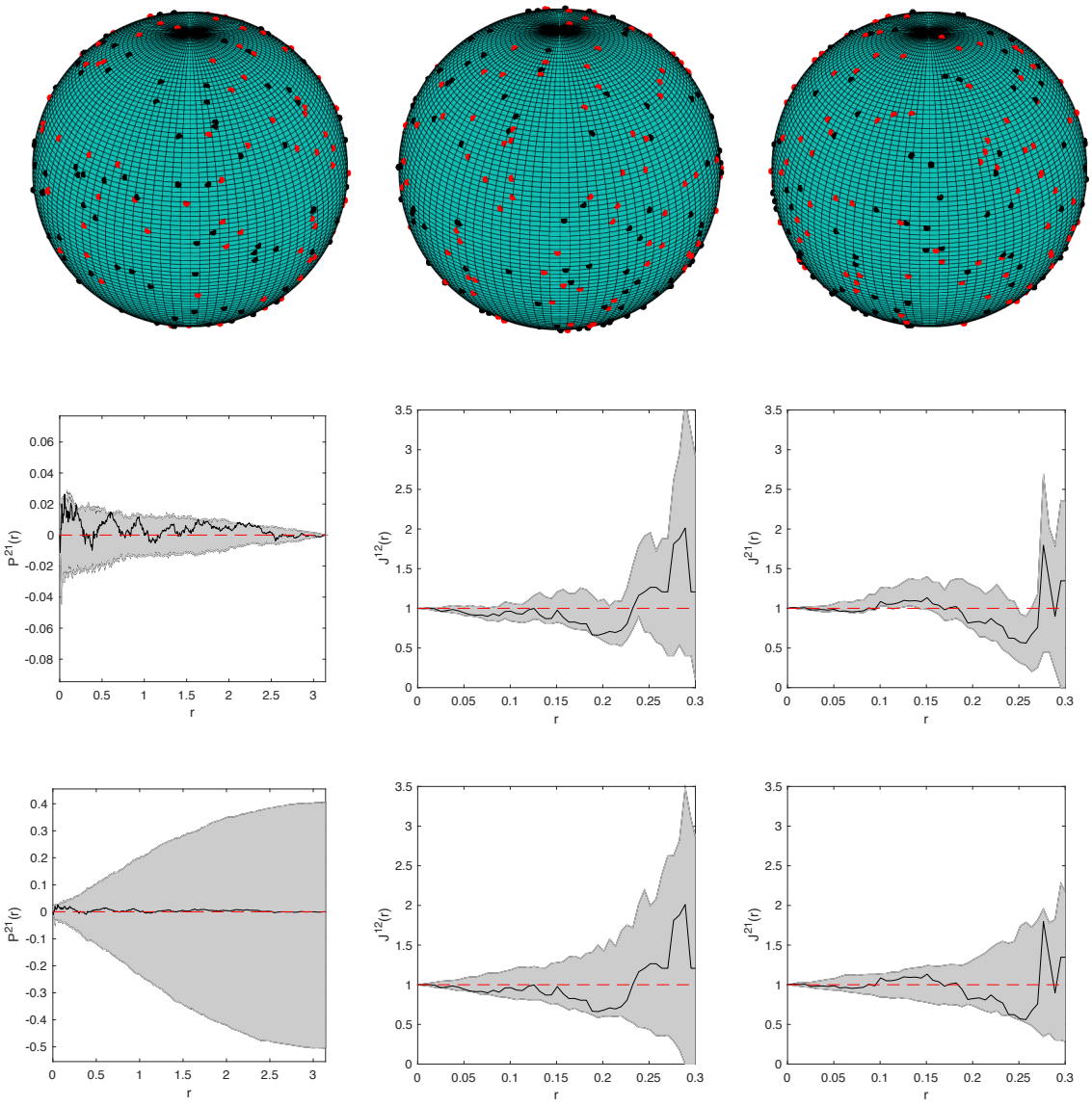


Figure 7.1: Example of an isotropic bivariate spheroidal Poisson process with constant intensity functions $\rho_1 = 10$ and $\rho_2 = 12$ for the first and second component respectively. *Top row:* realisation of the process with X_1 and X_2 being represented by the red and black points respectively. All three figures are the same realisation observed at different angles on the sphere. *Middle row:* plots of the functional summary statistics where random rotations are used to construct the envelope and *bottom row:* plots of the functional summary statistics where simulates from a homogeneous Poisson process are used to construct the envelope. Solid black line is the observed functional summary statistic, dashed red line is the theoretical value under independence and the grey area indicates the 2.5% lower and 97.5% upper simulation envelope.

Appendix E.3 we provide additional theoretical details of multivariate spheroidal LGCP.

We can simulate a bivariate LGCP with independent components by imposing the driving random field to have independent component, e.g. Y_1, Y_2 are independent GRFs then the resulting LGCP has independent components. Since the GRF has independent components we have $c_{12} = c_{21} = 0$ and $g_{12} = g_{21} = 1$, hence recovering the results of Proposition 7.3.1

In order to construct attractive and repulsive bivariate LGCP we shall follow the prescription given by Brix and Moller [2001]. Let Z be a univariate GRF with zero mean and covariance function $c(\mathbf{x}, \mathbf{y}) = \sigma^2 s(\sqrt{(\mathbf{x}-\mathbf{y})^T(\mathbf{x}-\mathbf{y})}) = \sigma^2 s(r)$, where s is the correlation function and $\sigma^2, r > 0$. In Appendix E.3 we show that a univariate GRF on the sphere with correlation function dependent on Euclidean distance instead of the geodesic distance is still isotropic. Then we define,

$$Y_1(\mathbf{x}) = \mu_1(\mathbf{x}) + a_1 Z(\mathbf{x})$$
$$Y_2(\mathbf{x}) = \mu_2(\mathbf{x}) + a_2 Z(\mathbf{x}),$$

where μ_i are continuous mean functions and $a_i \in \mathbb{R}$. The random field $Y = (Y_1, Y_2)$ is a bivariate GRF which is almost surely integrable, see Appendix E.3. The covariance and cross covariance functions are,

$$c_{ii}(r) = a_i^2 \sigma^2 s(r)$$
$$c_{12}(r) = a_1 a_2 \sigma^2 s(r),$$

respectively for $i = 1, 2$, $r > 0$, and which semi-positive definiteness of the covariance function matrix, $\{c_{ij}\}$, is satisfied in order for the GRF to be well defined [Brix and Moller, 2001]. Then the bivariate Cox process $X = (X_1, X_2)$ with driving random field $\exp(Y) = (\exp(Y_1), \exp(Y_2))$ is our LGCP. In our examples we shall assume that the correlation function s takes the *exponential* form,

$$s(r) = \exp\left(\frac{-r}{\gamma^2}\right), \tag{7.14}$$

for $\gamma^2, r > 0$. Thus, by controlling the signs of a_1 and a_2 we can enforce aggregation or repulsion. For example if we set $a_1, a_2 > 0$ then $c_{12} > 0$ which implies that $g_{12} > 1$ indicating attraction between components of X the resulting bivariate LGCP. Similarly if the signs of a_1 and a_2 differ then repulsion between the components is enforced. The magnitude of a_i $i = 1, 2$ controls the extent to which the aggregation or repulsion is imposed.

We now consider three examples:

1. Independent bivariate LGCP where we set $\sigma^2 = 1, \gamma^2 = 1$ and the mean functions of the first and second components are $\mu_1 = \log(10)$ and $\mu_2 = \log(5)$ respectively. An example is given in Figure 7.2 where the results show that the observed functional summary statistics lie within the simulation envelopes as expected.

2. Attractive bivariate LGCP where we set $\sigma^2 = 1, \gamma^2 = 1, a_1 = a_2 = 1$ and the mean functions of the first and second components are $\mu_1 = \log(15)$ and $\mu_2 = \log(10)$ respectively. An example is given in Figure 7.3 where the results show that the observed functional summary statistics lies above the simulation envelopes for P^{21} and below for both J^{12} and J^{21} highlighting the attractive nature of the process.

3. Repulsive bivariate LGCP where we set $\sigma^2 = 1, \gamma^2 = 1, a_1 = 1, a_2 = -1$ and the mean functions of the first and second components are $\mu_1 = \log(15)$ and $\mu_2 = \log(10)$ respectively. An example is given in Figure 7.4 where the results show that the observed functional summary statistics lies below the simulation envelopes for P^{21} and above for both J^{12} and J^{21} highlighting the repulsive nature of the process.

These examples demonstrate that the isotropic functional summary statistics are able to capture independence of components. Attraction and repulsion are also correctly identified based on the deviations of the summary statistics compared to the null simulates. The middle left plot P^{21} of Figures 7.1-7.4 graphically highlight the issue of simulating the null hypothesis using random rotations as the estimators of the summary statistics for the null simulates all converge to 0.

7.4 Summary statistics for inhomogeneous processes on $\mathbb{S}^2$

In this section we consider how to construct functional summary statistics for inhomogeneous marked processes. We first discuss how Equations 7.5-7.7 can be adapted to this scenario following the work of Cronie and van Lieshout [2016] who discussed the inhomogeneous J-function for point processes in $\mathbb{R}^d$. We then progress onto how the inhomogeneous cross K-function can be constructed based on Equation 7.4. We then examine how these functional summary statistics perform in determining independence, aggregation or repulsion exhibited between components of inhomogeneous spheroidal bivariate processes using simulation envelopes.

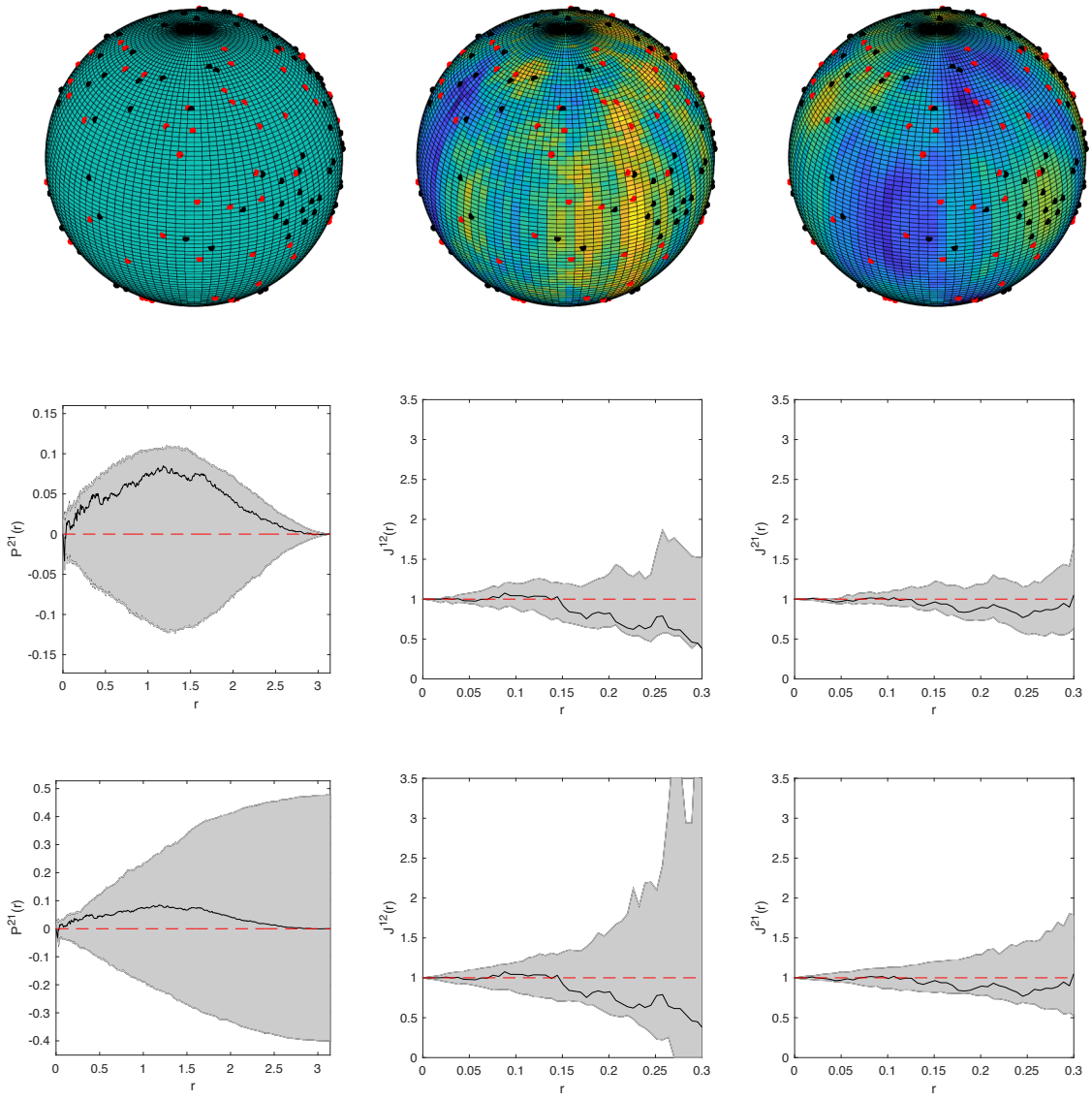


Figure 7.2: Example of an isotropic independent bivariate LGCP with parameters $\mu_1 = \log(10), \mu_2 = \log(5), \sigma^2 = 1, \gamma^2 = 1$. *Top row:* realisation of the process with X_1 and X_2 being represented by the red and black points respectively. The figure on the left does not have a surface intensity, the middle figure has the surface intensity given by the observed random field Y_1 and the right figure has the surface intensity given by the observed random field Y_2: all these figures are displayed from the same perspective, where high intensity is indicated by yellow and low by blue. *Middle row:* plots of the functional summary statistics where random rotations are used to construct the envelope and *bottom row:* plots of the functional summary statistics where simulates from a homogeneous Poisson process are used to construct the envelope. Solid black line is the observed functional summary statistic, dashed red line is the theoretical value under independence and the grey area indicates the 2.5% lower and 97.5% upper simulation envelope.

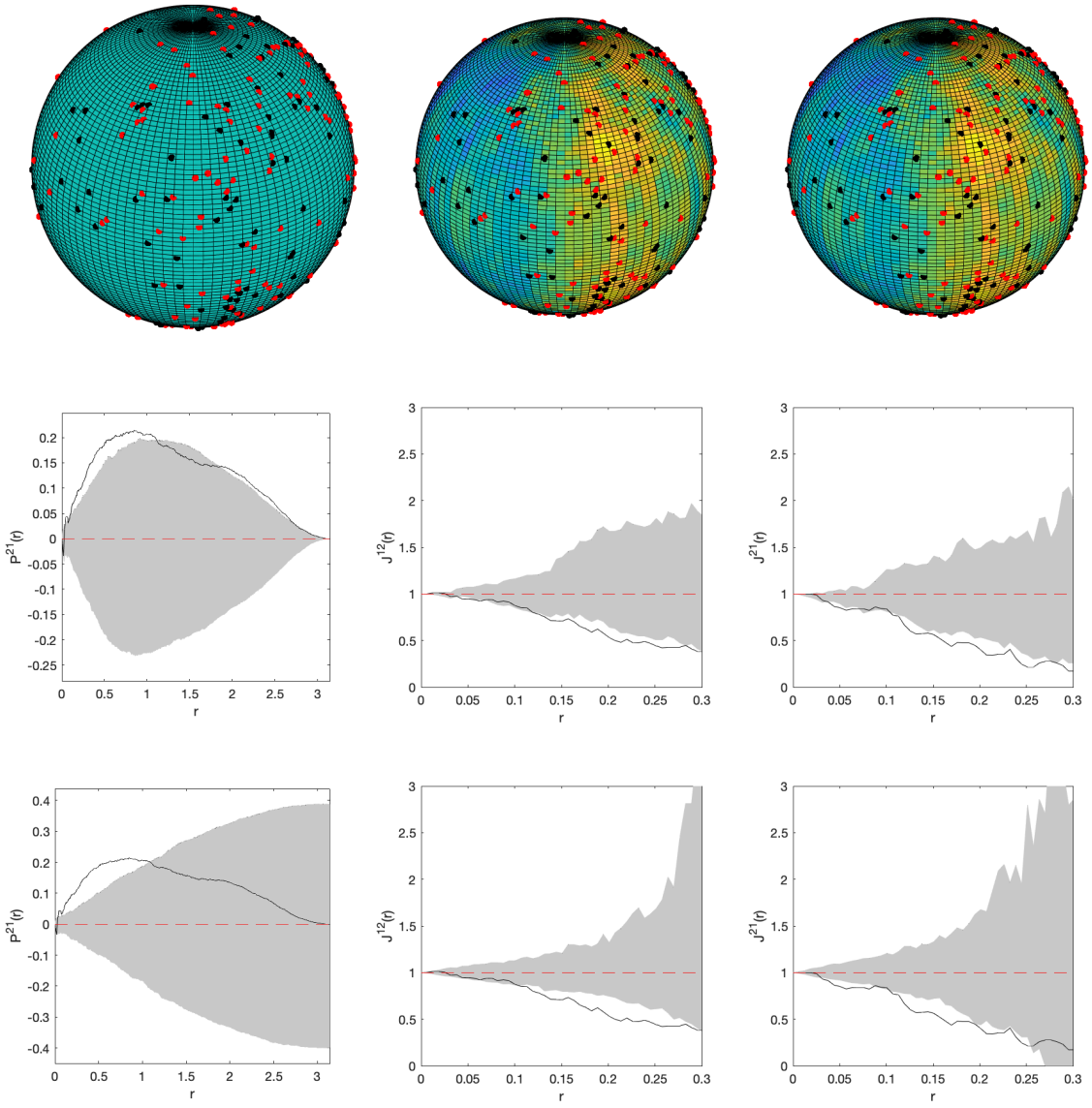


Figure 7.3: Example of an isotropic attractive bivariate LGCP with parameters $\mu_1 = \log(15), \mu_2 = \log(10), a_1 = a_2 = 1, \sigma^2 = 1, \gamma^2 = 1$. *Top row:* realisation of the process with X_1 and X_2 being represented by the red and black points respectively. The figure on the left does not have a surface intensity, the middle figure has the surface intensity given by the observed random field Y_1 and the right figure has the surface intensity given by the observed random field Y_2: all these figures are displayed from the same perspective, where high intensity is indicated by yellow and low by blue. *Middle row:* plots of the functional summary statistics where random rotations are used to construct the envelope and *bottom row:* plots of the functional summary statistics where simulates from a homogeneous Poisson process are used to construct the envelope. Solid black line is the observed functional summary statistic, dashed red line is the theoretical value under independence and the grey area indicates the 2.5% lower and 97.5% upper simulation envelope.

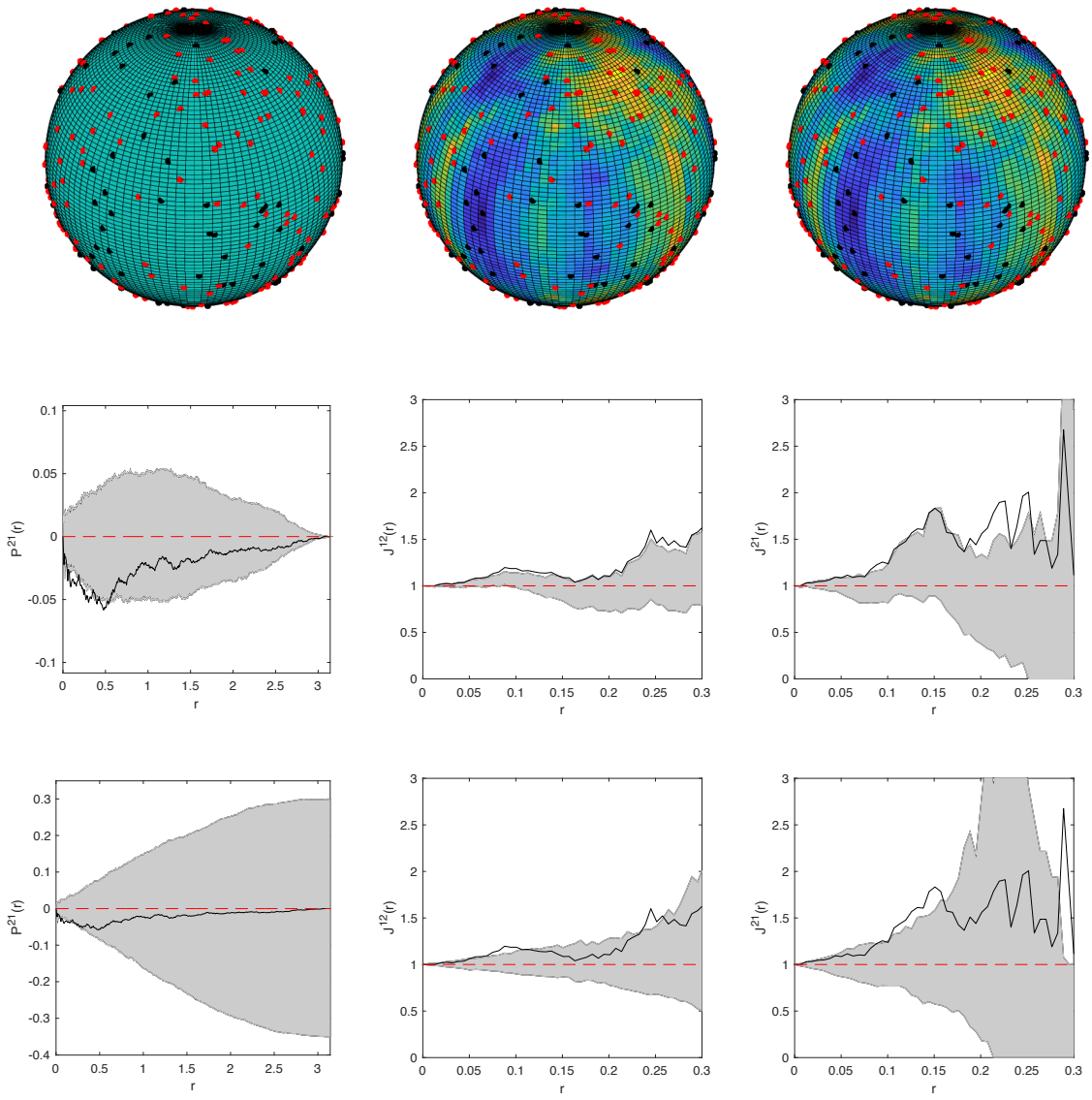


Figure 7.4: Example of an isotropic repulsive bivariate LGCP with parameters $\mu_1 = \log(15), \mu_2 = \log(10), a_1 = 1, a_2 = -1, \sigma^2 = 1, \gamma^2 = 1$. *Top row:* realisation of the process with X_1 and X_2 being represented by the red and black points respectively. The figure on the left does not have a surface intensity, the middle figure has the surface intensity given by the observed random field Y_1 and the right figure has the surface intensity given by the observed random field Y_2: all these figures are displayed from the same perspective, where high intensity is indicated by yellow and low by blue. *Middle row:* plots of the functional summary statistics where random rotations are used to construct the envelope and *bottom row:* plots of the functional summary statistics where simulates from a homogeneous Poisson process are used to construct the envelope. Solid black line is the observed functional summary statistic, dashed red line is the theoretical value under independence and the grey area indicates the 2.5% lower and 97.5% upper simulation envelope.

7.4.1 Inhomogeneous cross J function

To introduce the inhomogeneous cross J function we will follow arguments similar to those found in Cronie and van Lieshout [2016] where we use notions of rotational invariance of spheroidal processes instead of translational invariance for Euclidean processes as given originally. Let us first write $\bar{\rho}_E = \inf_{\mathbf{x}\in\mathbb{S}^2, m\in E} \rho(\mathbf{x}, m)$, we assume that our marked process X is simple and that its intensity functions of all orders exists. It is then supposed that the n^{th}-order correlation functions are isotropic in the sense,

$$\xi_n((\mathbf{x}_1, m_1), \ldots, (\mathbf{x}_n, m_n)) = \xi_n((O\mathbf{x}_1, m_1), \ldots, (O\mathbf{x}_n, m_n)),$$

for any $O \in \mathcal{O}(3)$ and $(\mathbf{x}_i, m_i) \in \mathbb{S}^2 \times \mathcal{M}$. If we also make the assumption that $\bar{\rho} \equiv \bar{\rho}_{\mathcal{M}} > 0$ then we say that X is *intensity-reweighted momented isotropic* (IRWMI), this definition is due to van Lieshout [2011] and extends a similar one found in Ward et al. [2021b] for univariate spheroidal processes. It should be noted that any isotropic process is also immediately IRWMI. Other IRWMI include Poisson processes, multitype processes with independent IRWMI components or a ground process that is IRWMI (in the sense of Ward et al. [2021b]) which is subsequently independently marked.

We now introduce the *inhomogeneous cross nearest neighbour distance distribution function* for marked spheroidal processes. Let X be a marked spheroidal processes that is IRWMI and let $C, E \subset \mathcal{M}$ with $\nu(C), \nu(E) > 0$. Denoting

$$u^r_{\mathbf{y},E}(\mathbf{x}, m) = \frac{\bar{\rho}_E \mathbb{1}[(\mathbf{x}, m) \in B_{\mathbb{S}^2}(\mathbf{y}, r) \times E]}{\rho(\mathbf{x}, m)}, \quad \mathbf{y} \in \mathbb{S}^2, \tag{7.15}$$

for $r \geq 0$ we then define the inhomogeneous cross nearest neighbour distance distribution function as,

$$D^{CE}_{\text{inhom}}(r) = 1 - G^!_{\mathbf{y},C}(1 - u^r_{\mathbf{y},E}). \tag{7.16}$$

By Theorem 7.4.1 the choice of $\mathbf{y}$ is arbitrary, this is identical to the reasoning found in Cronie and van Lieshout [2016] for the inhomogeneous D-function of Euclidean processes.

To see why this extends the isotropic cross D-function given by Equation 7.6 suppose that X is now isotropic and that the reference measure ν takes the mark measure, then Equation 7.16 becomes,

$$D^{CE}_{\text{inhom}}(r) = 1 - G^!_{\mathbf{y},C}(1 - u^r_{\mathbf{y},E})$$

$$= 1 - \frac{1}{\nu(C)} \int_C \mathbb{E}^!_{(\mathbf{y},m)} \left[\prod_{(\mathbf{x},n)\in X} \left(1 - \frac{\bar{\rho}_E \mathbb{1}[(\mathbf{x},n) \in B_{\mathbb{S}^2}(\mathbf{y},r) \times E]}{\rho(\mathbf{x},n)} \right) \right] \nu(dm)$$

since the process is isotopic with reference measure taking the mark measure then $\rho(\mathbf{x},m) \equiv \rho \in \mathbb{R}_+$ is constant and $\bar{\rho}_E = \rho$

$$\begin{aligned}
&= 1 - \frac{1}{\nu(C)} \int_C \mathbb{E}^!_{(\mathbf{y},m)} \left[\prod_{(\mathbf{x},n)\in X} (1 - \mathbb{1}[(\mathbf{x},n) \in B_{\mathbb{S}^2}(\mathbf{y},r) \times E]) \right] \nu(dm) \\
&= 1 - \frac{1}{\nu(C)} \int_C \mathbb{E}^!_{(\mathbf{y},m)} \left[\prod_{(\mathbf{x},n)\in X} \mathbb{1}[(\mathbf{x},n) \notin B_{\mathbb{S}^2}(\mathbf{y},r) \times E] \right] \nu(dm) \\
&= 1 - \frac{1}{\nu(C)} \int_C \mathbb{E}^!_{(\mathbf{y},m)} \left[\mathbb{1}[X \cap (B_{\mathbb{S}^2}(\mathbf{y},r) \times E) = \emptyset] \right] \nu(dm) \\
&= 1 - \frac{1}{\nu(C)} \int_C \mathbb{P}^!_{(\mathbf{y},m)} \left(X \cap (B_{\mathbb{S}^2}(\mathbf{y},r) \times E) = \emptyset \right]) \nu(dm) \\
&= 1 - \mathbb{P}^!_{\mathbf{y},C} \left(X \cap (B_{\mathbb{S}^2}(\mathbf{y},r) \times E) = \emptyset \right]) \\
&= \mathbb{P}^!_{\mathbf{y},C} \left(X \cap (B_{\mathbb{S}^2}(\mathbf{y},r) \times E) \neq \emptyset \right])
\end{aligned}$$

since X is isotropic then this probability does not depend on $\mathbf{y}$ and we have

$$= \mathbb{P}^!_{\mathbf{o},C} \left(X \cap (B_{\mathbb{S}^2}(\mathbf{o},r) \times E) \neq \emptyset \right])$$

which is precisely Equation 7.6.

Next we introduce the *inhomogeneous empty space function* of X_E as,

$$F^E_{\text{inhom}}(r) = 1 - G(1 - u^r_{\mathbf{y},E}),$$

where $u^r_{\mathbf{y},E}$ is given by Equation 7.15. Similarly for D^{CE}_{inhom}, Theorem 7.4.1 will show, like Cronie and van Lieshout [2016], that F^E_{inhom} does not depend on $\mathbf{y}$. Again, to see why this is an extension of the isotropic F-function given by Equation 7.5 consider X is an istropic spheroidal mark process then,

$$\begin{aligned}
F^E_{\text{inhom}}(r) &= 1 - G(1 - u^r_{\mathbf{y},E}) \\
&= 1 - \mathbb{E} \left[\prod_{(\mathbf{x},m)\in X} \left(1 - \frac{\bar{\rho}_E \mathbb{1}[(\mathbf{x},m) \in B_{\mathbb{S}^2}(\mathbf{y},r) \times E]}{\rho(\mathbf{x},m)} \right) \right]
\end{aligned}$$

since the process is isotopic $\rho(\mathbf{x}, m) \equiv \rho \in \mathbb{R}_+$ is a constant since the reference measure is also our mark measure and $\bar{\rho}_E = \rho$

$$
\begin{aligned}
&= 1 - \mathbb{E}\left[\prod_{(\mathbf{x},m)\in X} (1 - \mathbb{1}[(\mathbf{x}, m) \in B_{\mathbb{S}^2}(\mathbf{y}, r) \times E])\right] \\
&= 1 - \mathbb{E}\left[\prod_{(\mathbf{x},m)\in X} \mathbb{1}[(\mathbf{x}, m) \notin B_{\mathbb{S}^2}(\mathbf{y}, r) \times E]\right] \\
&= 1 - \mathbb{E}\left[\mathbb{1}[X \cap (B_{\mathbb{S}^2}(\mathbf{y}, r) \times E) = \emptyset]\right] \\
&= 1 - \mathbb{P}(X \cap (B_{\mathbb{S}^2}(\mathbf{y}, r) \times E) = \emptyset) \\
&= \mathbb{P}(X \cap (B_{\mathbb{S}^2}(\mathbf{y}, r) \times E) \neq \emptyset)
\end{aligned}
$$

since X is isotropic then this probability does not depend on $\mathbf{y}$ and we have

$$
= \mathbb{P}(X \cap (B_{\mathbb{S}^2}(\mathbf{o}, r) \times E) \neq \emptyset),
$$

which is precisely Equation 7.5.

Before introducing the *inhomogeneous cross J function* for spheroidal marked point processes we first need to return to the isotropic setting and provide an infinite series representation of J^{CE}, this exposition follows nearly identically to van Lieshout [2006] except where we replace translation invariance by rotational invariance and are working on $\mathbb{S}^2$ instead of $\mathbb{R}^d$. By White [1979] we have that F^E can be written as,

$$
\begin{aligned}
F^E(r) = -\sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \\
\int_{(B_{\mathbb{S}^2}(\mathbf{o},r)\times E)^n} \cdots \int \rho^{(n)}((\mathbf{x}_1, m_1), \ldots, (\mathbf{x}_n, m_n)) \lambda_{\mathbb{S}^2}(d\mathbf{x}_1)\nu(m_1) \cdots \lambda_{\mathbb{S}^2}(d\mathbf{x}_n)\nu(m_n),
\end{aligned}
$$

where our reference measure ν is taken as the mark measure over $\mathcal{M}$. It can be shown, using nearly identical arguments made by van Lieshout [2006], that if X is an isotropic spheroidal marked point process for which all intensity functions of any order exists and the first is given by the constant $\rho \in \mathbb{R}_+$, then J^{CE}, for any $C, E \subseteq \mathcal{M}$ with $\nu(C), \nu(E) > 0$, can be written as,

$$
J^{CE}(r) = \frac{1}{\nu(C)} \left(\nu(C) + \sum_{n=1}^{\infty} \frac{(-\rho)^n}{n!} J_n^{CE}(r)\right), \tag{7.17}
$$

for $r \in [0, \pi]$ and where

$$J_n^{CE}(r) = \int_C \int_{(B_{\mathbb{S}^2}(\mathbf{o},r)\times E)^n} \cdots \int \xi_{n+1}((\mathbf{y}, m), (O_{\mathbf{y}}\mathbf{x}_1, m_1), \ldots, (O_{\mathbf{y}}\mathbf{x}_n, m_n)) \\ \nu(m)\lambda_{\mathbb{S}^2}(d\mathbf{x}_1)\nu(m_1)\cdots\lambda_{\mathbb{S}^2}(d\mathbf{x}_n)\nu(m_n), \tag{7.18}$$

for $\mathbf{y} \in \mathbb{S}^2$. Through Cauchy's root test a sufficient condition for the infinite series to converge is

$$\limsup_{n\to\infty}(\rho^n J^{CE}(r)/n!)^{1/n} < 1.$$

From van Lieshout [2006] it is easily shown that Equation 7.18 does not depend on $\mathbf{y}$ if X is isotropic, hence neither does Equation 7.17. A proof of Equation 7.17 can easily be achieved based on the proof of Proposition 4.2 of van Lieshout [2006] where appropriate adaptations are made.

Based on Equation 7.17 we can now define the inhomogeneous cross J function. Let X be an IRWMI spheroidal marked point process for which its intensity functions exist of all orders and $C, E \subseteq \mathcal{M}$ such that $\nu(C), \nu(E) > 0$. We define the inhomogeneous cross J function between C and E as,

$$J_{\text{inhom}}^{CE}(r) = \frac{1}{\nu(C)}\left(\nu(C) + \sum_{n=1}^{\infty}\frac{(-\bar{\rho}_E)^n}{n!}J_n^{CE}(r)\right),$$

for $r \in [0, \pi]$ and where,

$$J_n^{CE}(r) = \int_C \int_{(B_{\mathbb{S}^2}(\mathbf{o},r)\times E)^n} \cdots \int \xi_{n+1}((\mathbf{y}, m), (O_{\mathbf{y}}\mathbf{x}_1, m_1), \ldots, (O_{\mathbf{y}}\mathbf{x}_n, m_n)) \\ \nu(m)\lambda_{\mathbb{S}^2}(d\mathbf{x}_1)\nu(m_1)\cdots\lambda_{\mathbb{S}^2}(d\mathbf{x}_n)\nu(m_n),$$

for $\mathbf{y} \in \mathbb{S}^2$. Again, as in the isotropic case, there is an implicit dependence on $\mathbf{y}$ but by IRWMI, J_n^{CE} is independent of $\mathbf{y}$ and hence so is J_{inhom}^{CE}. This definition is comparable to that given by Cronie and van Lieshout [2016] who define the inhomogeneous cross J functions for Euclidean processes. It can be seen immediately that if X is isotropic then we recover Equation 7.17. It is also important to note that if X is Poisson then $J_{\text{inhom}}^{CE}(r) \equiv 1$ and heuristically values $J_{\text{inhom}}^{CE}(r) < 1$ is indicative of points with marks in E tending to aggregate around points with marks in C whilst the converse is suggested if $J_{\text{inhom}}^{CE}(r) > 1$ at a distance $r \geq 0$.

Similar to Cronie and van Lieshout [2016], the inhomogeneous cross J function can be written as the ratio of the inhomogeneous cross nearest neighbour distribution and empty space function. The following theorem gives this relationship and a similar result is given for Euclidean processes in Theorem 1 of Cronie and van Lieshout [2016].

Theorem 7.4.1. *Let X be a spheroidal marked point process that is IRWMI. Under the further assumption that the intensity functions of all orders of X exist and that*

$$\limsup_{n\to\infty}\left(\frac{\bar{\rho}_E}{n!}\int_{(B_{\mathbb{S}^2}(\mathbf{o},r)\times E)^n}\cdots\int\frac{\rho^{(n)}(\mathbf{x}_1,m_1),\ldots,(\mathbf{x}_n,m_n)}{\rho(\mathbf{x}_1,m_1)\cdots\rho(\mathbf{x}_n,m_n)}\prod_{i=1}^{n}\lambda_{\mathbb{S}^2}(d\mathbf{x}_i)\nu(m_i)\right)^{1/n}<1,$$

then both D^{CE} and F^E are independent of $\mathbf{y}$ and

$$J^{CE}(r)=\frac{1-D^{CE}(r)}{1-F^E(r)},$$

for $r\in[0,\pi]$ and $F^E(r)\neq 1$.

Proof. The proof of Theorem 1 in Cronie and van Lieshout [2016] can be adapted to the spherical scenario. The proof also follows similar arguments to the proof of Theorem 3.5.7. □

The following proposition, similar to Proposition 2 of Cronie and van Lieshout [2016], shows that if X_C and X_E are independent then the inhomogeneous cross J function is identically 1.

Proposition 7.4.2. *Let X be a spheroidal IRWMI marked point process and consider two disjoint sets $C,E\subset\mathcal{M}$ such that they have positive measure ν. Further assume that X_C and X_E are independent and that the assumptions of Theorem 7.4.1 hold. Then,*

$$D^{CE}_{inhom}(r)=F^E_{inhom}(r)$$

and so $J^{CE}_{inhom}(r)\equiv 1$.

Proof. The proof of Proposition 2 in Cronie and van Lieshout [2016] can be adapted to the spherical scenario. □

7.4.2 Inhomogeneous cross K function

By relaxing the IRWMI assumption down to SOIRWI, which is equivalent to the first and second order correlation functions, ξ_1 and ξ_2, being rotationally invariant, we can construct the *inhomogeneous cross K function*. Moreover, the inhomogeneous cross K function takes an identical form to Equation 7.8 except we drop the assumption of isotropy in favour of SOIRWI, i.e. the PCF is isotropic. Even under the SOIRWI assumption it can easily be shown that if $C, E \subset \mathcal{M}$ are disjoint sets with positive measure ν such that X_C and X_E are independent then

$$K^{CE}_{\text{inhom}}(r) = 2\pi(1 - \cos(r)),$$

which identical to the isotropic setting [Møller and Waagepetersen, 2003]. We again maintain the heuristic understanding that if $K^{CE}_{\text{inhom}}(r) > 2\pi(1 - \cos(r))$ then this indicates that points with marks in C aggregate around points with marks in E whilst the converse holds if $K^{CE}_{\text{inhom}}(r) < 2\pi(1 - \cos(r))$. In the multitype setting if two components of the process are independent then the Cartesian product of those components is SOIRWI. In other words, let X be a multitype process with k components, suppose that X_i and X_j are independent for $i, j \in \{1, \ldots, k\}, i \neq j$ then (X_i, X_j) is SOIRWI [Møller and Waagepetersen, 2003, Proposition 4.4].

7.4.3 Estimating functional summary statistics

In order to estimate the inhomogeneous functional summary statistics we follow Cronie and van Lieshout [2016] to obtain estimates for J^{CE}_{inhom} whilst estimates for K^{CE}_{inhom} follow similarly to the isotropic setting. Let us suppose that we have observed our spheroidal point process X through some window $W \subseteq \mathbb{S}^2$. Then we can estimate our functional summary statistics as,

$$1 - \hat{D}^{CE}_{\text{inhom}}(r) = \frac{1}{\lambda_{\mathbb{S}^2}(W_{\ominus r})\nu(C)} \sum_{(\mathbf{x}, m_{\mathbf{x}}) \in X} \frac{\mathbb{1}[(\mathbf{x}, m_{\mathbf{x}}) \in W_{\ominus r} \times C]}{\rho(\mathbf{x}, m_{\mathbf{x}})} \prod_{(\mathbf{y}, m_{\mathbf{y}}) \in X} \left(1 - \frac{\bar{\rho}_E \mathbb{1}[d_{\mathbb{S}^2}(\mathbf{x}, \mathbf{y}) < r, m_{\mathbf{y}} \in E]}{\rho(\mathbf{y}, m_{\mathbf{y}})}\right) \tag{7.19}$$

$$1 - \hat{F}^{E}_{\text{inhom}}(r) = \frac{1}{|I_{W_{\ominus r}}|} \sum_{\mathbf{p} \in I_{W_{\ominus r}}} \prod_{(\mathbf{x}, m) \in X} \left(1 - \frac{\bar{\rho}_E \mathbb{1}[d_{\mathbb{S}^2}(\mathbf{p}, \mathbf{x}) \leq r, m \in E]}{\rho(\mathbf{x}, m_{\mathbf{x}})}\right) \tag{7.20}$$

$$\hat{J}^{CE}_{\text{inhom}}(r) = \frac{1-\hat{D}^{CE}(r)}{1-\hat{F}^{E}(r)} \quad \text{for } \hat{F}^{E}(r) < 1 \tag{7.21}$$

$$\hat{K}^{CE}_{\text{inhom}}(r) = \frac{1}{\lambda_{\mathbb{S}^2}(W_{\ominus r})\nu(C)\nu(E)} \sum_{(\mathbf{x},m_{\mathbf{x}})\in X} \sum_{(\mathbf{y},m_{\mathbf{y}})\in X\setminus(\mathbf{x},m_{\mathbf{x}})} \frac{\mathbb{1}[(\mathbf{x},m_{\mathbf{x}})\in W_{\ominus r}\times C, d_{\mathbb{S}^2}(\mathbf{x},\mathbf{y})\leq r, m_{\mathbf{y}}\in E]}{\rho(\mathbf{x},m_{\mathbf{x}})\rho(\mathbf{y},m_{\mathbf{y}})}. \tag{7.22}$$

The following proposition presents the first order properties of these proposed estimators.

Proposition 7.4.3. *Let X be a spheroidal marked process such that the assumptions of Theorem 7.4.1 hold. Then $\hat{D}^{CE}_{inhom}, \hat{F}^{E}_{inhom}$, and $\hat{K}^{CE}_{inhom}$ are unbiased whilst $\hat{J}^{CE}_{inhom}$ is ratio unbiased for known ρ.*

Proof. Proof of Lemma 1 Cronie and van Lieshout [2016] can be adapted to the current setting. □

For most cases the intensity function ρ is unknown in which case a plug-in estimator is typically used. Moreover, it may be the case that ν is unknown or an irregular window makes calculating $\lambda_{\mathbb{S}^2}(W_{\ominus r})$ challenging, in which case Stoyan and Stoyan [2000] advocate using the Hamiltonian principle where $\lambda_{\mathbb{S}^2}(W_{\ominus r})\nu(C)$ is replaced by

$$\sum_{(\mathbf{x},m)\in X} \frac{\mathbb{1}[(\mathbf{x},m_{\mathbf{x}})\in W_{\ominus r}\times C]}{\rho(\mathbf{x},m)}. \tag{7.23}$$

By the Campbell formula this is unbiased for $\lambda_{\mathbb{S}^2}(B)\nu(C)$ for known ρ. Even if ν is known and $\lambda_{\mathbb{S}^2}(W_{\ominus r})$ is easy to calculate, for example if a process is completely observed such that $W = \mathbb{S}^2$ we still advocate the use of Equation 7.23 instead of $\lambda_{\mathbb{S}^2}(W_{\ominus r})\nu(C)$. This is because guarantees on the range of $\hat{D}^{CE}_{\text{inhom}}$ are not achievable when ρ is replaced by a plugin estimator and the true value of $\lambda_{\mathbb{S}^2}(B)\nu(C)$ is used. To see this first note that $D^{CE}_{\text{inhom}}(0) = 0$ for any $X \in N_{lf}$ and so we would want any estimator to replicate this behaviour. If our plugin is $\hat{\rho}$ for ρ then when $r = 0$ we have

$$\hat{D}^{CE}_{\text{inhom}}(0) = 1 - \frac{1}{\lambda_{\mathbb{S}^2}(W_{\ominus r})\nu(C)} \sum_{(\mathbf{x},m_{\mathbf{x}})\in X} \frac{\mathbb{1}[(\mathbf{x},m_{\mathbf{x}})\in W_{\ominus r}\times C]}{\hat{\rho}(\mathbf{x},m_{\mathbf{x}})}. \tag{7.24}$$

There is no guarantee that this is 0 and can even be negative, this will also be the case when r is small since the point process is assumed simple. If we instead use Equation 7.23 with

the same plugin estimator for ρ as in Equation 7.24 then we gaurantee for $r = 0$ that $\hat{D}^{CE}$ is 0 and positive for small r. In all coming examples we shall use Equation 7.23 instead of $\lambda_{\mathbb{S}^2}(B)\nu(C)$.

Further to this, Iftimi et al. [2019] discusses the Hamiltonian principle when considering $\hat{K}^{CE}_{\text{inhom}}$ specifically for spatio-temporal point processes. In Appendix B of Iftimi et al. [2016], a preprint of Iftimi et al. [2019], they discuss a number of scenarios to estimating $\lambda_{\mathbb{S}^2}(B)\nu(C)\nu(E)$ by applying the Hamiltonian principle which could be analogously applied here.

In the coming example we use the kernel estimates developed in Chapter 6 where the bandwidth selected by the criteria developed by Cronie and Van Lieshout [2018] and was adapted to the setting of Riemmanian manifolds.

When considering the special case of multitype patterns, ν can be taken as the standard counting measure in which case the estimates simplify as the reference measure is known. For this case $\rho(\mathbf{x}, i) = \nu(i)\rho_i(\mathbf{x}) = \rho_i(\mathbf{x})$, let $i, j \in \mathcal{M} = \{1, \dots, k\}$ such that $i \neq j$ then the estimates are,

$$1 - \hat{D}^{ij}_{\text{inhom}}(r) = \frac{1}{\lambda_{\mathbb{S}^2}(W_{\ominus r})} \sum_{\mathbf{x} \in X_i} \frac{\mathbb{1}[\mathbf{x} \in W_{\ominus r}]}{\rho_i(\mathbf{x})} \prod_{\mathbf{y} \in X_j} \left(1 - \frac{\bar{\rho}_j \mathbb{1}[d_{\mathbb{S}^2}(\mathbf{x}, \mathbf{y}) < r]}{\rho_j(\mathbf{y})}\right) \tag{7.25}$$

$$1 - \hat{F}^{j}_{\text{inhom}}(r) = \frac{1}{|I_{W_{\ominus r}}|} \sum_{\mathbf{p} \in I_{W_{\ominus r}}} \prod_{\mathbf{x} \in X_j} \left(1 - \frac{\bar{\rho}_j \mathbb{1}[d_{\mathbb{S}^2}(\mathbf{p}, \mathbf{x}) \leq r]}{\rho_j(\mathbf{x})}\right) \tag{7.26}$$

$$\hat{J}^{ij}_{\text{inhom}}(r) = \frac{1 - \hat{D}^{ij}(r)}{1 - \hat{F}^{j}(r)} \quad \text{for } \hat{F}^{j}(r) < 1 \tag{7.27}$$

$$\hat{K}^{CE}_{\text{inhom}}(r) = \frac{1}{\lambda_{\mathbb{S}^2}(W_{\ominus r})} \sum_{\mathbf{x} \in X_i} \sum_{\mathbf{y} \in X_j} \frac{\mathbb{1}[\mathbf{x} \in W_{\ominus r}, d_{\mathbb{S}^2}(\mathbf{x}, \mathbf{y}) \leq r]}{\rho_i(\mathbf{x})\rho_j(\mathbf{y})}, \tag{7.28}$$

Plugin estimators can be used for $\rho_i, i \in \{1, \dots, k\}$. For $\hat{D}^{CE}_{\text{inhom}}$ instead of using $\lambda_{\mathbb{S}^2}(W_{\ominus r})$ we use the estimator $\sum_{\mathbf{x} \in X_i \cap W_{\ominus r}} 1/\hat{\rho}_i(\mathbf{x})$ to ensure that $\hat{D}^{CE}_{\text{inhom}}$ is 0 when $r = 0$ and is positive for small r, where $\hat{\rho}_i$ is our plugin estimator.

7.4.4 Determining independence

In the inhomogeneous setting, determining whether dependency exists between marks is more complex. A toroidal shift approach is considered by Cronie and van Lieshout [2016] where not only is the pattern translated but so is the underlying intensity function for Euclidean

processes. This approach can be adapted to the spherical setting where translations are replaced with rotations. Following Cronie and van Lieshout [2016], we first consider the random measure which places mass $1/\rho(\mathbf{x}, m)$ at points $(\mathbf{x}, m) \in X$

$$\Xi^X = \sum_{(\mathbf{x},m)\in X} \frac{\delta_{(\mathbf{x},m)}}{\rho(\mathbf{x}, m)},$$

then for $B \times C \subset \mathcal{B}(\mathbb{S}^2 \times \mathcal{M})$ we have that,

$$\Xi^X(B \times C) = \sum_{(\mathbf{x},m)\in X} \frac{\mathbb{1}[(\mathbf{x}, m) \in B \times C]}{\rho(\mathbf{x}, m)}.$$

We define the rotation of Ξ^X as,

$$\begin{aligned}
\Xi^X_O(B \times C) &= \Xi^X(OB \times C) \\
&= \sum_{(\mathbf{x},m)\in X} \frac{\mathbb{1}[(O^T\mathbf{x}, m) \in B \times C]}{\rho(\mathbf{x}, m)} \\
&= \sum_{(\mathbf{x},m)\in O^T X} \frac{\mathbb{1}[(\mathbf{x}, m) \in B \times C]}{\rho(O\mathbf{x}, m)},
\end{aligned}$$

where $O \in \mathcal{O}(3)$ and OX is taken to mean the spatial coordinates of X being rotated by the rotation O. We say that Ξ^X is isotropic if $\Xi^X \overset{d}{=} \Xi^X_O$ for any $O \in \mathcal{O}(3)$. Notice that if X is IRWMI then Ξ^X is moment isotropic, that is the factorial moment measures of Ξ^X and Ξ^X_O are identical for any $O \in \mathcal{O}(3)$. Using this and an identical argument to Proposition 4 Cronie and van Lieshout [2016] we can show the following proposition.

Proposition 7.4.4. *Suppose that X is a spheroidal marked point process for which the assumptions of Theorem 7.4.1 hold and let C, E be disjoint subsets of $\mathcal{M}$ with strictly positive ν measure. If X_C and X_E are independent and the measure Ξ^X is isotropic then the estimates $D^{CE}_{inhom}, F^{E}_{inhom}, J^{CE}_{inhom}$, and K^{CE}_{inhom} can be written as a function of (Ξ^{X_C}, Ξ^{X_E}) and $(\Xi^{X_C}, \Xi^{X_E}_O) \overset{d}{=} (\Xi^{X_C}, \Xi^{X_E})$.*

Proof. Proposition 4 Cronie and van Lieshout [2016] can be adapted to the current setting. □

This proposition implies that we can generate samples from our null hypothesis by randomly rotating one or both of the components X_C and/or X_E whilst also rotating the intensity function to account for this rotation.

Although this approach is quite general extending Lotwick and Silverman [1982] toroidal shift method to inhomogeneous processes, this approach does fall into the same issues as its isotropic counterpart. That is we require the point process to be completely observed on $\mathbb{S}^2$ and that for values of r close to π the estimates of our functional summary statistics are very close to each other. Thus in scenarios where the point process is not completely observed or distant interactions are important we may consider imposing further assumptions on the null hypothesis in order to generate null samples, e.g. Poisson processes with independent components, and thus construct the simulation envelopes akin to the second approach discussed for isotropic processes.

7.4.5 Examples

Estimates of the intensity are constructed using the techniques discussed in Chapter 6. We select the bandwidth using the Cronie's criteria [Cronie and Van Lieshout, 2018] without any edge correction and then construct the estimator applying local edge correction, see Equation 6.9. When used as a plugin estimator into Equations 7.25-7.28 we use the leave-one-out estimator for ρ_i, i.e. if $\mathbf{x} \in X_i$ then we have the estimate,

$$\hat{\rho}_i(\mathbf{x}) = \sum_{\mathbf{y} \in X \cap W \setminus \mathbf{x}} \frac{c_h(\mathbf{x}, \mathbf{y})^{-1}}{h^d} k\left(\frac{d_g(\mathbf{x}, \mathbf{y})}{h}\right), \tag{7.29}$$

and $c_h(\mathbf{x}, \mathbf{y})$ is given by Equation 6.9. We use the leave-one-out estimator since Baddeley et al. [2000] found that Equation 7.29 leads to less bias, compared to Equation 6.7, for both Poisson and clustered processes. The situation is less clear for regular processes.

We consider four examples,

1. Independent bivariate Poisson process with intensity $\rho_1(\mathbf{x}) = \exp(\log(6) + z)$ and $\rho_2(\mathbf{x}) = \exp(\log(5) + 2x)$ for the first and second components respectively where $\mathbf{x} = (x, y, z)^T \in \mathbb{S}^2$. An example is given in Figure 7.5 where the results show that the observed functional summary statistics lie within the simulation envelopes as expected.

2. Independent bivariate LGCP where we set $\sigma^2 = 1, \gamma^2 = 1$. The mean functions of the first and second components are $\mu_1(\mathbf{x}) = \mu_2(\mathbf{x}) = \log(5) + x$ respectively, $\mathbf{x} = (x, y, z)^T \in \mathbb{S}^2$. An example is given in Figure 7.6 where the results show that the observed functional summary statistics lie within the simulation envelopes as expected.

3. Attractive bivariate LGCP where we set $\sigma^2 = 1, \gamma^2 = 1, a_1 = a_2 = 1$. The mean functions of the first and second components are $\mu_1(\mathbf{x}) = \mu_2(\mathbf{x}) = \log(5) + x^2$ respectively,

$\mathbf{x} = (x, y, z)^T \in \mathbb{S}^2$. An example is given in Figure 7.7 where the results show that the observed functional summary statistics lies above the simulation envelopes for P^{21} and below for both J^{12} and J^{21} highlighting the attractive nature of the process.

4. Repulsive bivariate LGCP where we set $\sigma^2 = 1, \gamma^2 = 1, a_1 = 1, a_2 = -1$. The mean functions of the first and second components are $\mu_1(\mathbf{x}) = \log(5) + y^2$ and $\mu_2(\mathbf{x}) = \log(5) + x^2$ respectively, $\mathbf{x} = (x, y, z)^T \in \mathbb{S}^2$. An example is given in Figure 7.8 where the results show that the observed functional summary statistics lies below the simulation envelopes for P^{21} and above for both J^{12} and J^{21}, highlighting the repulsive nature of the process.

These examples demonstrate that the functional summary statistics developed in this section for inhomogeneous multitype point processes are capable of providing evidence supporting whether components are independent, attractive or repulsive based on comparison with simulates from a null model. The middle left plots of P^{21} given in Figures 7.5-7.8 again highlight the issue of using random rotations to simulate the null hypothesis, as was the case for the isotropic examples considered earlier; all the functional summary statistics converge to a single point at $r = \pi$, although not 0 in the inhomogeneous case.

7.5 Summary statistics on convex shapes

In this section we demonstrate how functional summary statistics can be constructed for point patterns observed on convex shapes and used to determine independence, attraction or repulsion. Following the discussion given by Ward et al. [2021b] there are a number of impracticalities to defining functional summary statistics directly on a convex shape due to the potential asymmetric nature of the convex shape. This asymmetry means analogous notions of translation and rotational invariance cannot be well-defined, which in turn means functional summary statistics cannot be constructed directly. Instead we can use the concepts introduced by Ward et al. [2021b] and map point processes from their original space onto the sphere, where rotational symmetries can now be utilised to construct well-defined functional summary statistics.

7.5.1 Mapping marked point processes to $\mathbb{S}^2$

The following theorem shows that marked point processes on convex shapes can be mapped to a marked point process on the sphere under an appropriately transformed intensity function.

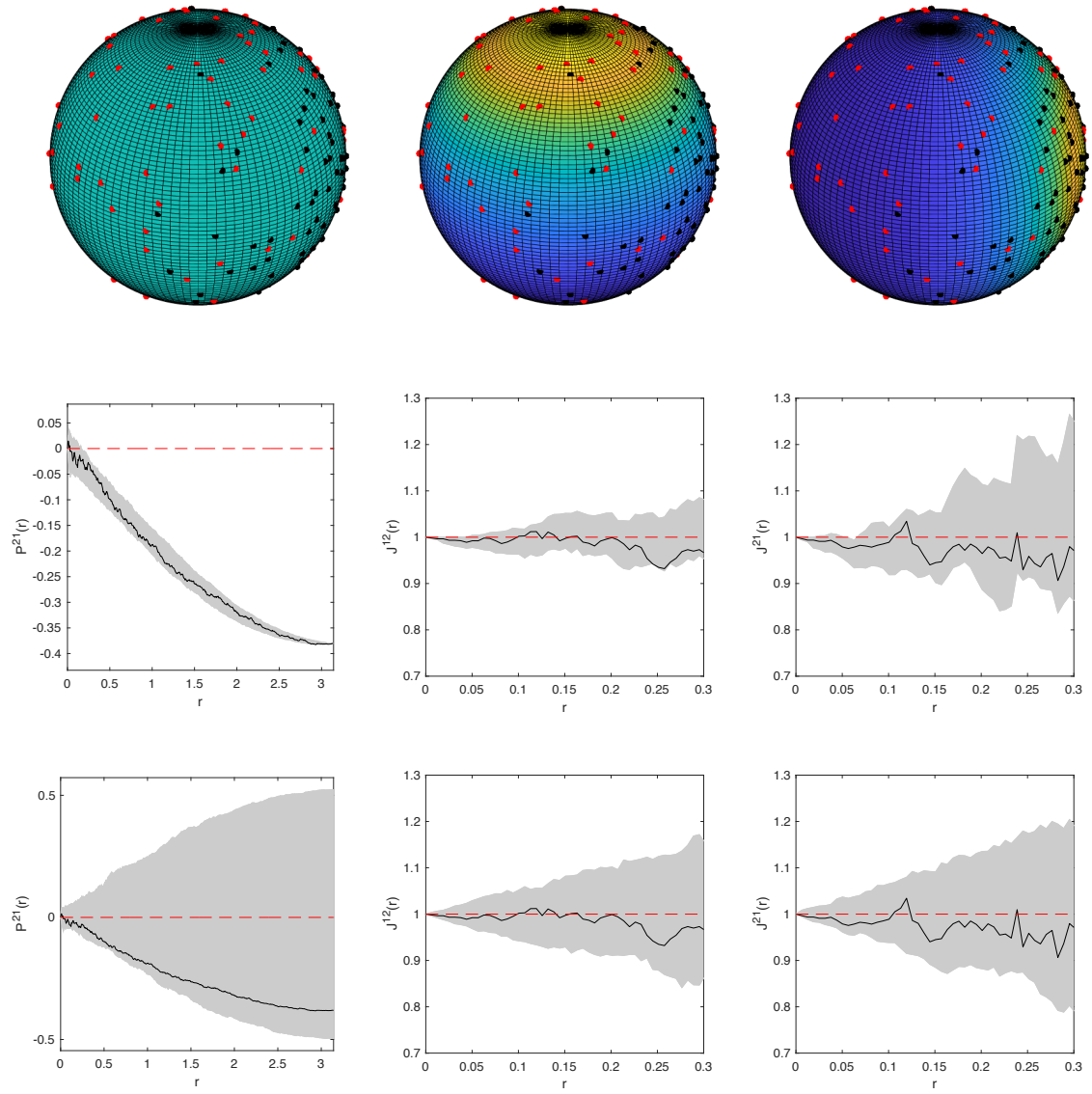


Figure 7.5: Example of an inhomogeneous independent bivariate Poison process with intensity functions $\rho_1(\mathbf{x}) = \exp(\log(6) + z), \rho_2(\mathbf{x}) = \exp(\log(5) + 2x), \mathbf{x} = (x, y, z)^T \in \mathbb{S}^2$ for the first and second components respectively. *Top row:* realisation of the process with X_1 and X_2 being represented by the red and black points respectively. The figure on the left does not have a surface intensity, the middle figure has the surface intensity given by ρ_1 and the right figure has the surface intensity given by ρ_2: all these figures are displayed from the same perspective, where high intensity is indicated by yellow and low by blue. *Middle row:* plots of the functional summary statistics where random rotations are used to construct the envelope and *bottom row:* plots of the functional summary statistics where simulates from an inhomogeneous Poisson process with the fitted intensity are used to construct the envelope. Solid black line is the observed functional summary statistic, dashed red line is the theoretical value under independence and the grey area indicates the 2.5% lower and 97.5% upper simulation envelope.

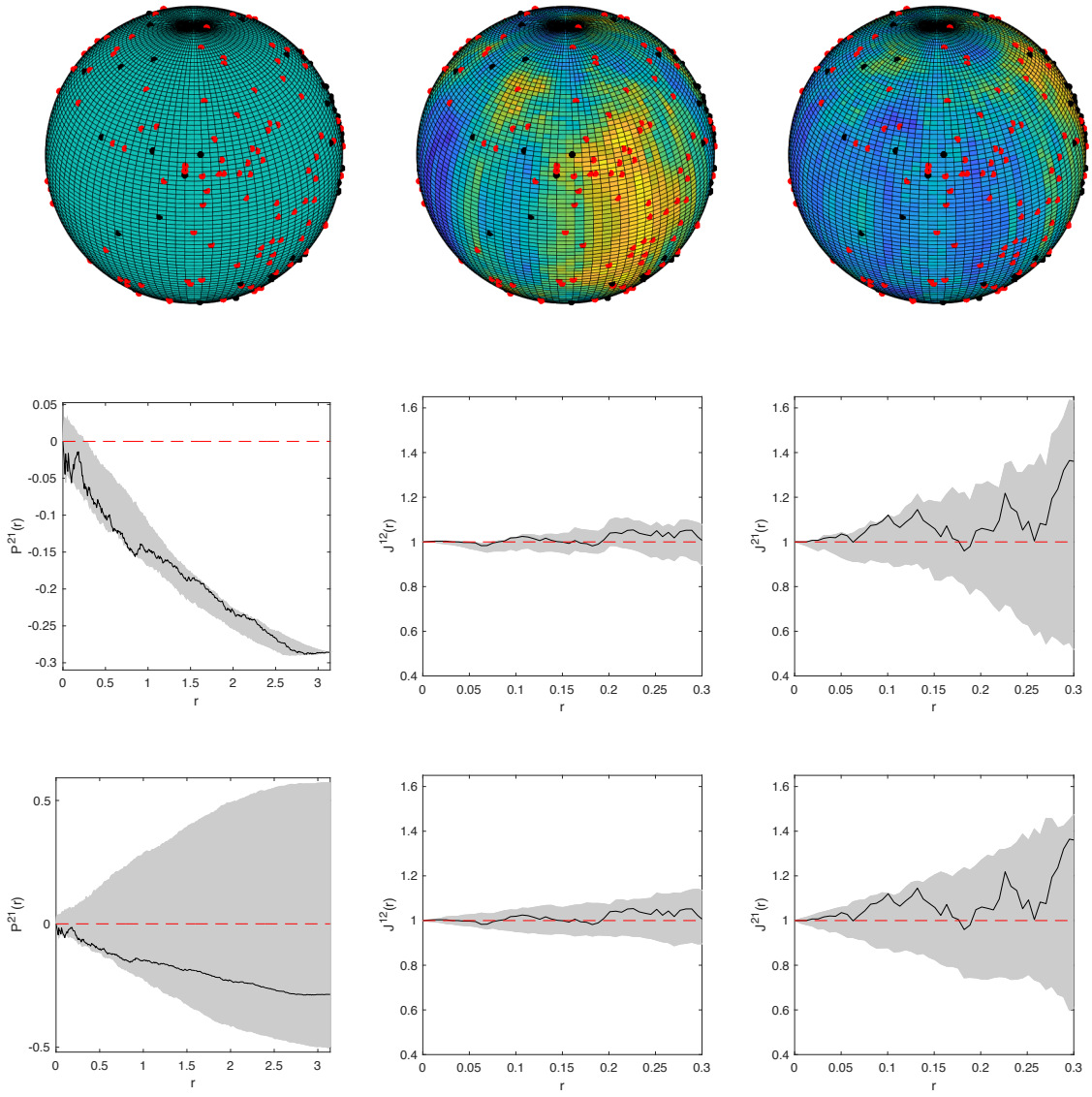


Figure 7.6: Example of an inhomogeneous independent bivariate LGCP with parameters $\mu_1(\mathbf{x}) = \mu_2(\mathbf{x}) = \log(5) + x, \sigma^2 = 1, \gamma^2 = 1, \mathbf{x} = (x, y, z)^T \in \mathbb{S}^2$. *Top row:* realisation of the process with X_1 and X_2 being represented by the red and black points respectively. The figure on the left does not have a surface intensity, the middle figure has the surface intensity given by the observed random field Y_1 and the right figure has the surface intensity given by the observed random field Y_2: all these figures are displayed from the same perspective, where high intensity is indicated by yellow and low by blue. *Middle row:* plots of the functional summary statistics where random rotations are used to construct the envelope and *bottom row:* plots of the functional summary statistics where simulates from an inhomogeneous Poisson process with the fitted intensity are used to construct the envelope. Solid black line is the observed functional summary statistic, dashed red line is the theoretical value under independence and the grey area indicates the 2.5% lower and 97.5% upper simulation envelope.

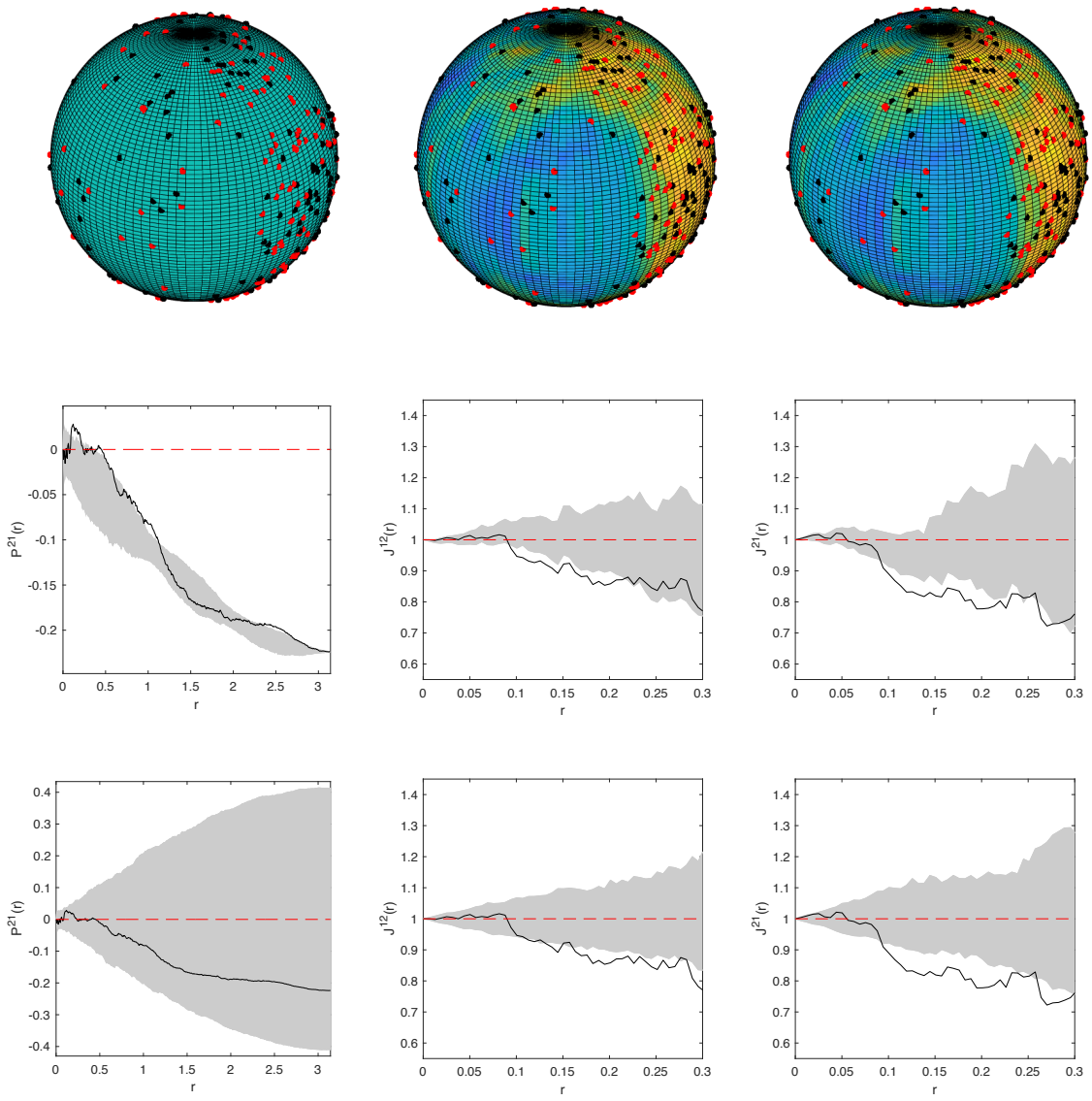


Figure 7.7: Example of an inhomogeneous attractive bivariate LGCP with parameters $\mu_1(\mathbf{x}) = \mu_2(\mathbf{x}) = \log(5) + x^2, a_1 = a_2 = 1, \sigma^2 = 1, \gamma^2 = 1, \mathbf{x} = (x, y, z)^T \in \mathbb{S}^2$. *Top row:* realisation of the process with X_1 and X_2 being represented by the red and black points respectively. The figure on the left does not have a surface intensity, the middle figure has the surface intensity given by the observed random field Y_1 and the right figure has the surface intensity given by the observed random field Y_2: all these figures are displayed from the same perspective, where high intensity is indicated by yellow and low by blue. *Middle row:* plots of the functional summary statistics where random rotations are used to construct the envelope and *bottom row:* plots of the functional summary statistics where simulates from an inhomogeneous Poisson process with the fitted intensity are used to construct the envelope. Solid black line is the observed functional summary statistic, dashed red line is the theoretical value under independence and the grey area indicates the 2.5% lower and 97.5% upper simulation envelope.

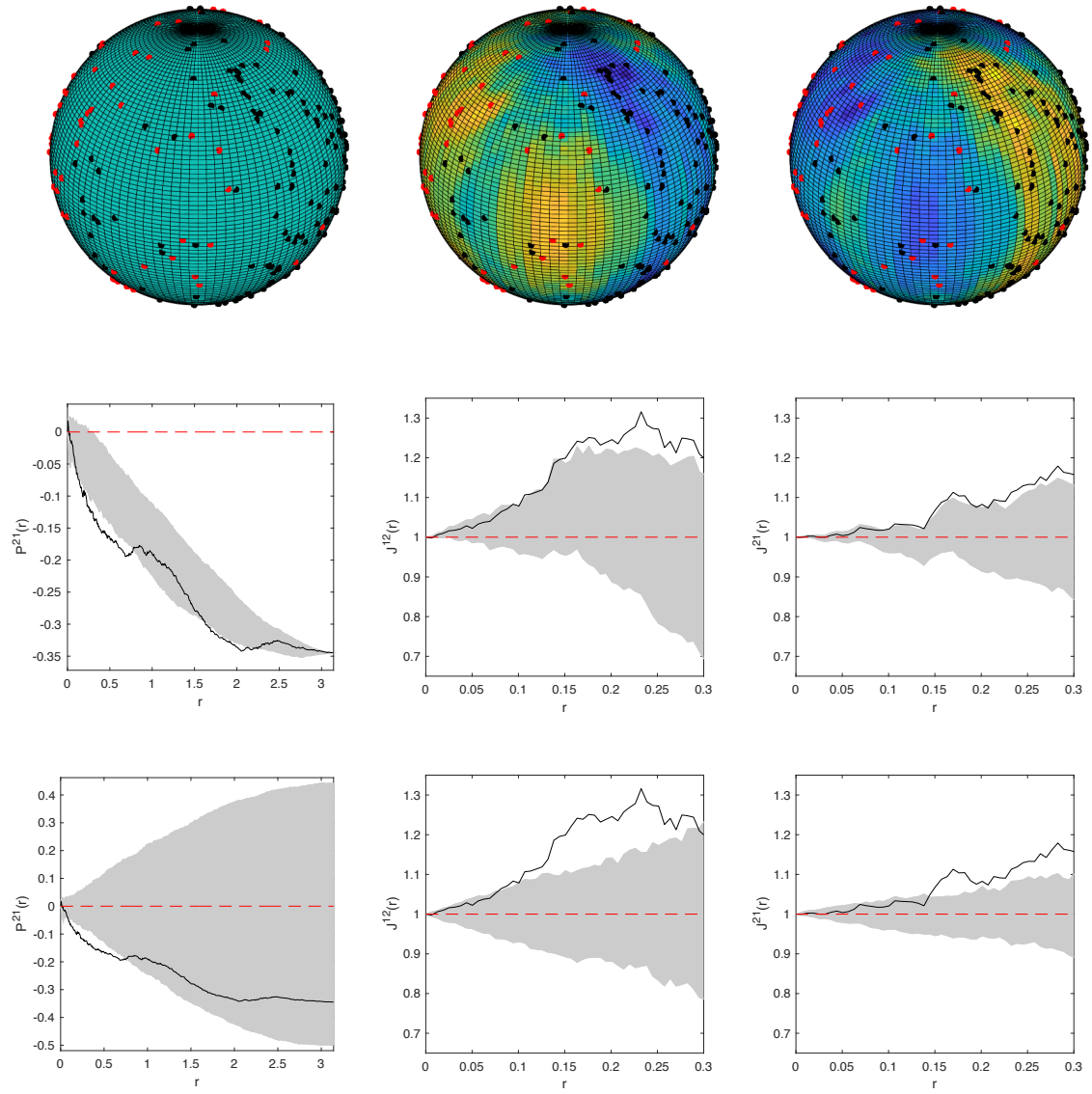


Figure 7.8: Example of an inhomogeneous repulsive bivariate LGCP with parameters $\mu_1(\mathbf{x}) = \log(5) + y^2, \mu_2(\mathbf{x}) = \log(5) + x^2, a_1 = 1, a_2 = -1, \sigma^2 = 1, \gamma^2 = 1, \mathbf{x} = (x, y, z)^T \in \mathbb{S}^2$. *Top row:* realisation of the process with X_1 and X_2 being represented by the red and black points respectively. The figure on the left does not have a surface intensity, the middle figure has the surface intensity given by the observed random field Y_1 and the right figure has the surface intensity given by the observed random field Y_2: all these figures are displayed from the same perspective, where high intensity is indicated by yellow and low by blue. *Middle row:* plots of the functional summary statistics where random rotations are used to construct the envelope and *bottom row:* plots of the functional summary statistics where simulates from an inhomogeneous Poisson process with the fitted intensity are used to construct the envelope. Solid black line is the observed functional summary statistic, dashed red line is the theoretical value under independence and the grey area indicates the 2.5% lower and 97.5% upper simulation envelope.

Theorem 7.5.1. *Let $\mathbb{D}$ be a convex surface such that the same assumptions for Theorem 4.3.2 hold. Let X be a marked point process on $\mathbb{D}$ with intensity function ρ and $f : \mathbb{D} \mapsto \mathbb{S}^2$ be a bijective measurable mapping. Then the point process $Y = f(X) \equiv \{(f(\mathbf{x}), m) : (\mathbf{x}, m) \in X\}$, is a marked point process on $\mathbb{S}^2$ with intensity function,*

$$\rho^*(\mathbf{x}, m) = \begin{cases} \rho(f^{-1}(\mathbf{x}), m) l_1(f^{-1}(\mathbf{x})) J_{(1,f^*)}(\mathbf{x}) \sqrt{1 - x_{m_1}^2 - x_{n_1}^2}, & \mathbf{x} = (x_{r_1}, x_{s_1}, x_{t_1})^T \in f(\mathbb{D}_1) \\ \vdots & \\ \rho(f^{-1}(\mathbf{x}), m) l_p(f^{-1}(\mathbf{x})) J_{(p,f^*)}(\mathbf{x}) \sqrt{1 - x_{m_p}^2 - x_{n_p}^2}, & \mathbf{x} = (x_{r_p}, x_{s_p}, x_{t_p})^T \in f(\mathbb{D}_p) \end{cases}$$

where $x_{o_i} = \tilde{g}_i(x_{m_i}, x_{n_i})$ and $\tilde{g}_i, l_i, J_{(i,f^*)}, m_i, n_i, o_i, r_i, s_i, t_i$ for $i = 1, \ldots, p$ and f^* are defined as in Theorem 4.3.2.

Proof. Proof of Last and Penrose [2018, Theorem 5.1] can be adapted from the unmarked to the marked setting and when the surface is convex. □

By using Theorem 7.5.1 we can convert a point process over $\mathbb{D}$ to a point process over $\mathbb{S}^2$ and thus utilise the estimators developed previously for analysis of spheroidal marked processes. More precisely, we can use Equations 7.19-7.22 to estimate our functional summary statistics where instead of ρ in the original statement of the estimators we can use ρ^* as given by Theorem 7.5.1.

As in the spherical case, we would need to construct an estimator of the unknown intensity function. Suppose we have some estimator $\hat{\rho}$ of ρ then we can estimate ρ^* from Theorem 7.5.1 as,

$$\hat{\rho}^*(\mathbf{x}, m) = \begin{cases} \hat{\rho}(f^{-1}(\mathbf{x}), m) l_1(f^{-1}(\mathbf{x})) J_{(1,f^*)}(\mathbf{x}) \sqrt{1 - x_{m_1}^2 - x_{n_1}^2}, & \mathbf{x} = (x_{r_1}, x_{s_1}, x_{t_1})^T \in f(\mathbb{D}_1) \\ \vdots & \\ \hat{\rho}(f^{-1}(\mathbf{x}), m) l_p(f^{-1}(\mathbf{x})) J_{(p,f^*)}(\mathbf{x}) \sqrt{1 - x_{m_p}^2 - x_{n_p}^2}, & \mathbf{x} = (x_{r_p}, x_{s_p}, x_{t_p})^T \in f(\mathbb{D}_p) \end{cases} \tag{7.30}$$

Iftimi et al. [2019] proposed a Voronoi intensity estimator for marked spatio-temporal point processes which could be adapted to this setting to estimate ρ. Additionally Iftimi et al. [2016, Section 6.1.1], a preprint of Iftimi et al. [2019], outlines a number of simplfying assumptions that can make calculating $\hat{\rho}$ easier, e.g. a common mark distribution.

As the focus of this chapter is on multitype patterns we can take our reference measure to be the standard counting measure and so $\rho(\mathbf{x}, i) = \rho_i(\mathbf{x})$ for $i \in \{1, \ldots, k\}$. Thus we need

only consider two scenarios:

1. $\rho_i(\mathbf{x})$ is constant over $\mathbb{D}$ and so can be estimated by $\hat{\rho}_i(\mathbf{x}) = N_{X_i}(W)/\lambda_{\mathbb{D}}(W)$,
2. $\rho_i(\mathbf{x})$ is not constant over $\mathbb{D}$ and so can be estimated using a kernel intensity estimator for each X_i, $i \in \{1, \ldots, k\}$.

Both of these can then be used in conjunction with Theorem 7.5.1 and Equation 7.30 to construct an estimate of the intensity function on the sphere as either,

$$\hat{\rho}_i^*(\mathbf{x}) = \begin{cases} \frac{N_{X_i}(W)}{\lambda_{\mathbb{D}}(W)} l_1(f^{-1}(\mathbf{x})) J_{(1,f^*)}(\mathbf{x}) \sqrt{1 - x_{m_1}^2 - x_{n_1}^2}, & \mathbf{x} = (x_{r_1}, x_{s_1}, x_{t_1})^T \in f(\mathbb{D}_1) \\ \vdots & \\ \frac{N_{X_i}(W)}{\lambda_{\mathbb{D}}(W)} l_p(f^{-1}(\mathbf{x})) J_{(p,f^*)}(\mathbf{x}) \sqrt{1 - x_{m_p}^2 - x_{n_p}^2}, & \mathbf{x} = (x_{r_p}, x_{s_p}, x_{t_p})^T \in f(\mathbb{D}_p), \end{cases} \tag{7.31}$$

$$\hat{\rho}_i^*(\mathbf{x}) = \begin{cases} \hat{\rho}_i(f^{-1}(\mathbf{x})) l_1(f^{-1}(\mathbf{x})) J_{(1,f^*)}(\mathbf{x}) \sqrt{1 - x_{m_1}^2 - x_{n_1}^2}, & \mathbf{x} = (x_{r_1}, x_{s_1}, x_{t_1})^T \in f(\mathbb{D}_1) \\ \vdots & \\ \hat{\rho}_i(f^{-1}(\mathbf{x})) l_p(f^{-1}(\mathbf{x})) J_{(p,f^*)}(\mathbf{x}) \sqrt{1 - x_{m_p}^2 - x_{n_p}^2}, & \mathbf{x} = (x_{r_p}, x_{s_p}, x_{t_p})^T \in f(\mathbb{D}_p), \end{cases} \tag{7.32}$$

for the homogeneous and inhomogeneous case respectively and $\mathbf{x} \in \mathbb{S}^2$ and $\hat{\rho}_i$ is an estimator of the intensity for X_i. In the coming examples we will use the nonparametric kernel approach discussed in Chapter 6 when ρ_i is inhomogeneous.

7.5.2 Determining independence

On $\mathbb{S}^2$ we considered two approaches to examining whether a process has independence between subsets to the mark space: (1) random rotations and (2) full distributional specification of the null hypothesis. For the first approach, by Proposition 7.4.4, we know that if the process is IRWMI on $\mathbb{S}^2$ then we can apply random rotations to simulate the null, assuming the intensity function is also suitably adjusted to test independence. Alternatively, for the second approach, so long as the specified null process is IRWMI and/or SOIRWI then from Monte Carlo simulates the functional summary statistics can be constructed and independence tested for.

When mapping point processes from our convex shape to the sphere the assumptions of IRWMI and/or SOIRWI of the resulting process may not hold. Even though this may be the case in the coming examples we show that the functional summary statistics are still able to capture attractive and repulsive behaviour of a process. Although, in general, the resulting spheroidal process may not satisfy the assumptions of IRWMI and/or SOIRWI there are a few cases where these assumptions still hold for processes mapped from $\mathbb{D}$ to $\mathbb{S}^2$: (1) Poisson process with an common mark distribution is IRWMI, (2) a multitype Poisson process with independent components is IRWMI, and (3) A multitype process with independent components is SOIRWI, see e.g [Møller and Waagepetersen, 2003, Proposition 4.4].

We highlight the special case when the process is Poisson. If we considered the second approach and we specify a Poisson null then we know the null model is IRWMI and hence, if the process is completely observed, we could in fact use random rotations rather then simulating Poisson processes over $\mathbb{D}$. This allows for a more general hypothesis with less distributional assumptions to be tested.

7.5.3 Examples: Homogeneous multitype processes

We now exemplify these functional summary statistics across various bivariate processes. We shall consider point processes that have been completely observed over an ellipsoid with semi-major axis lengths $a = b = 0.8$ and $c = 1.4398$ along the x, y, and $z-$axes, the length along the $z-$axis is set such that the surface area of the ellipsoid is 4π: identical to the unit sphere. To simulate a LGCP on an ellipsoid we follow similarly to the spherical case where a zero mean Gaussian field is first simulated in $\mathbb{R}^3$ with correlation function given by Equation 7.14. Then focusing only on points $\mathbf{x} \in \mathbb{D}$, given the random field, we can simulate a Poisson process. The result is a LGCP. We use a construction similar to Brix and Moller [2001] to model attractive and repulsive behaviour of components. We consider four homogeneous examples where we use Equation 7.31 and have used the function $f(\mathbf{x}) = (x/a, y/b, z/c)^T$ for $\mathbf{x} = (x, y, z)^T \in \mathbb{D}$ to map from the ellipsoid to the sphere (cf. Equation 4.9),

1. Independent bivariate Poisson process with constant intensity functions $\rho_1 = \rho_2 = 10$ for the first and second components. An example is given in Figure 7.9 where the results show that the observed functional summary statistics lie within the simulation envelopes as expected.

2. Independent bivariate LGCP where we set $\sigma^2 = 1, \gamma^2 = 0.2$ and the mean functions of

the first and second components are $\mu_1 = \mu_2 = \log(5)$. An example is given in Figure 7.10 where the results show that the observed functional summary statistics lie well within the simulation envelopes as expected.

3. Attractive bivariate LGCP where we set $\sigma^2 = 1, \gamma^2 = 0.2, a_1 = a_2 = 1$ and the mean functions of the first and second components are $\mu_1 = \mu_2 = \log(5)$. An example is given in Figure 7.11 where the results show that the observed functional summary statistics lies above the simulation envelopes for P^{21} and below for both J^{12} and J^{21} highlighting the attractive nature of the process.

4. Repulsive bivariate LGCP where we set $\sigma^2 = 1, \gamma^2 = 0.2, a_1 = 1, a_2 = -1$ and the mean functions of the first and second components are $\mu_1 = \mu_2 = \log(5)$. An example is given in Figure 7.12 where the results show that the observed functional summary statistics lies below the simulation envelopes for P^{21} and above for both J^{12} and J^{21} highlighting the repulsive nature of the process.

Although the assumptions necessary to construct well defined functional summary statistics may not hold when the process is mapped from the ellipsoid to the sphere, these examples demonstrate that we can still determine whether components of a multitype point process exhibit independence, attraction or repulsion. These functional summary statistics therefore serve as useful exploratory tools when first investigating point processes on convex shapes.

7.5.4 Examples: Inhomogeneous multitype processes

Next we consider four inhomogeneous point processes. We use Equation 7.32 with our kernel estimator for the intensity function developed in Chapter 6 and bandwidth selected using Cronie's criteria [Cronie and Van Lieshout, 2018] adapted to the Riemannian manifold setting. We shall use the function $f(\mathbf{x}) = (x/a, y/b, z/c)^T$ for $\mathbf{x} = (x, y, z)^T \in \mathbb{D}$ to map from the ellipsoid to the sphere.

1. Independent bivariate Poisson process with intensity functions $\rho_1(\mathbf{x}) = \exp(\log(6) + z)$ and $\rho_2(\mathbf{x}) = \exp(\log(6) + 2x)$ for the first and second components respectively where $\mathbf{x} = (x, y, z)^T \in \mathbb{D}$. An example is given in Figure 7.13 where the results show that the observed functional summary statistics lie well within the simulation envelopes as expected.

2. Independent bivariate LGCP where we set $\sigma^2 = 1, \gamma^2 = 0.2$ and the mean functions of the first and second components are $\mu_1(\mathbf{x}) = \mu_2(\mathbf{x}) = \log(6) + x$ with $\mathbf{x} = (x, y, z)^T \in \mathbb{D}$.

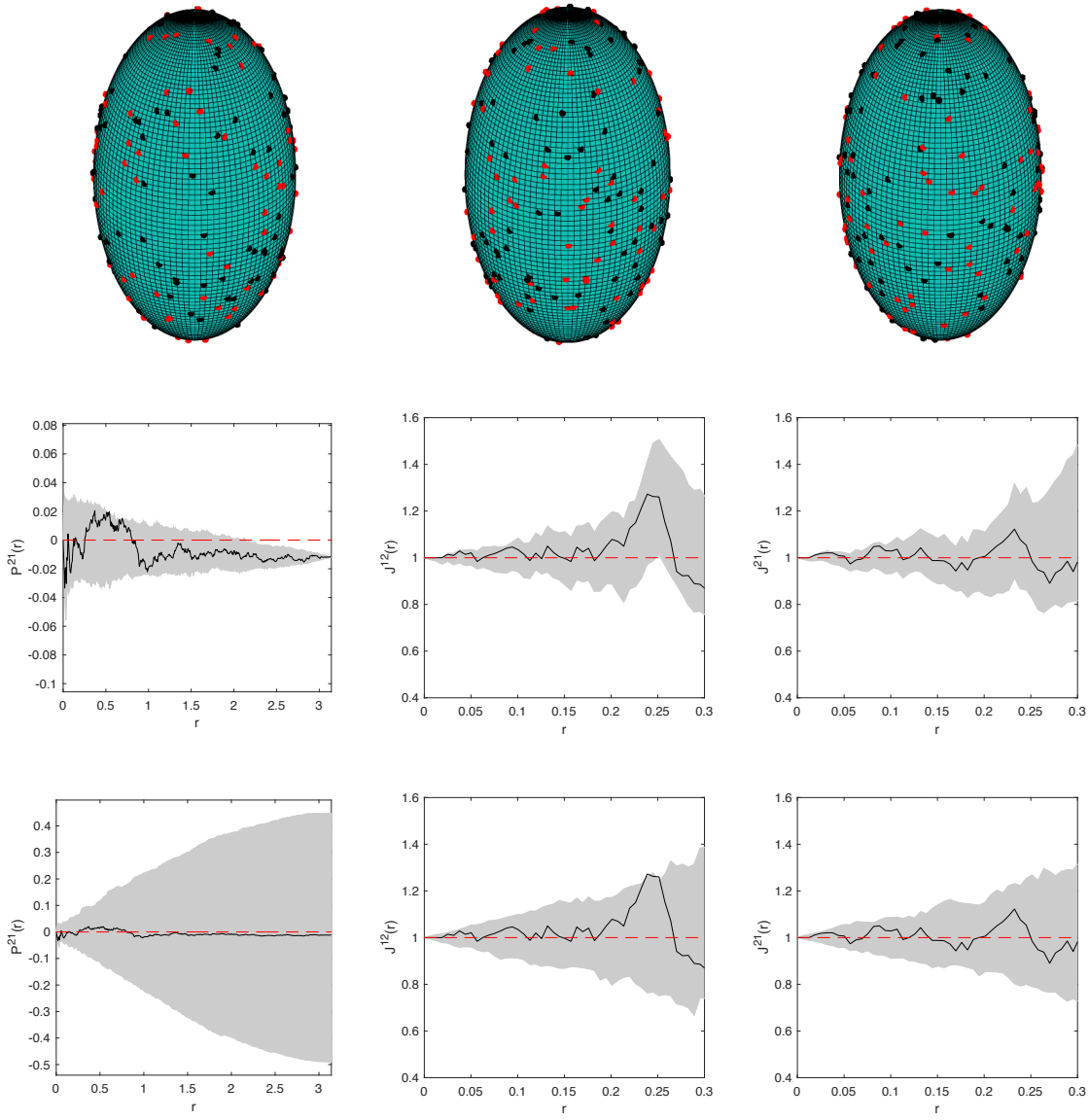


Figure 7.9: Example of a homogeneous independent bivariate Poisson process on an ellipsoid with dimensions $(a, b, c) = (0.8, 0.8, 1.4398)$ and intensity function $\rho_1 = \rho_2 = \log(10)$. *Top row:* realisation of the process with X_1 and X_2 being represented by the red and black points respectively. The three figures show the same realisation from different perspectives of the ellipsoid. *Middle row:* plots of the functional summary statistics where random rotations are used to construct the envelope and *bottom row:* plots of the functional summary statistics where simulates from a homogeneous Poisson process are used to construct the envelope. Solid black line is the observed functional summary statistic, dashed red line is the theoretical value under independence and the grey area indicates the 2.5% lower and 97.5% upper simulation envelope.

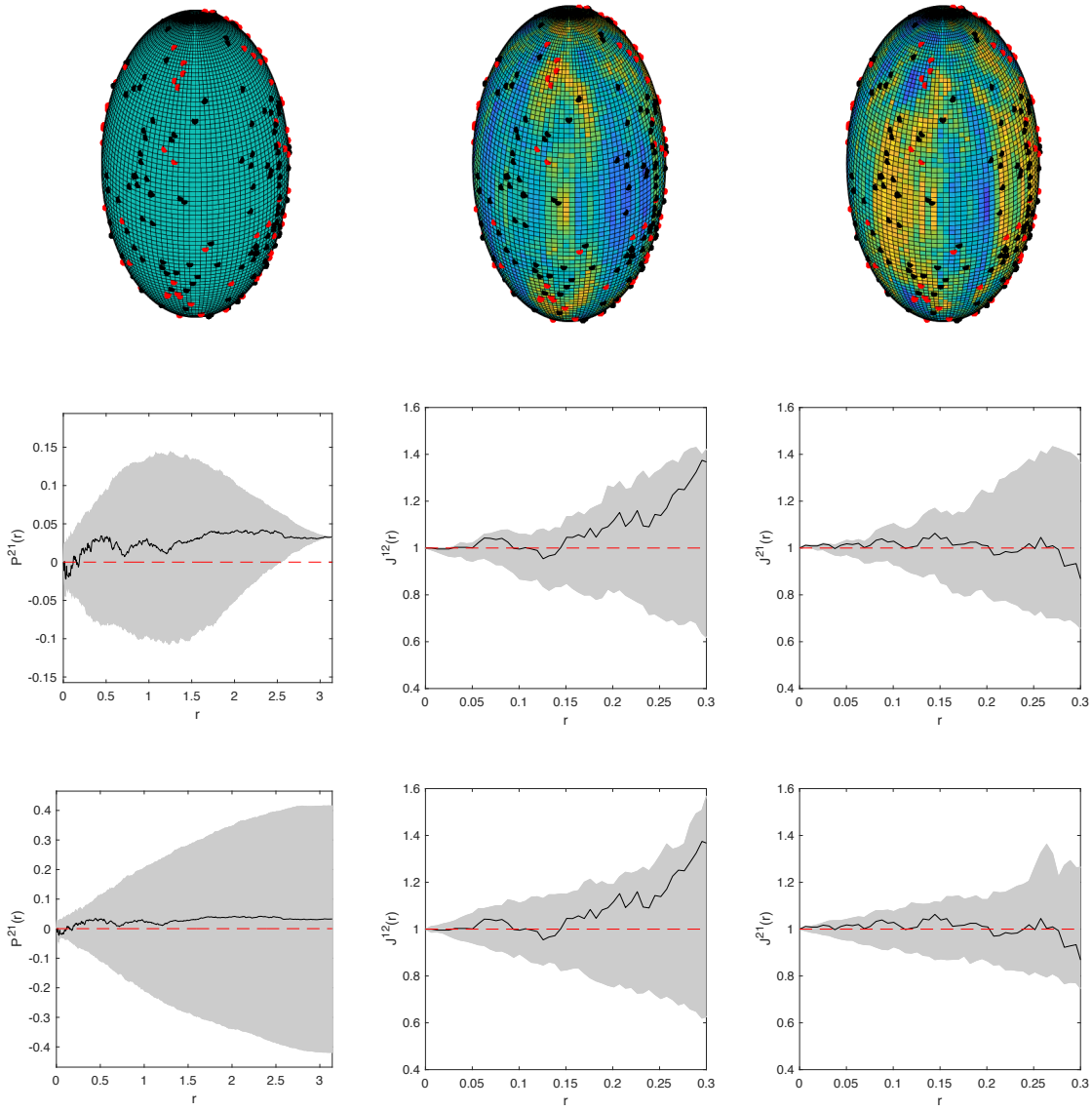


Figure 7.10: Example of a homogeneous independent bivariate LGCP on an ellipsoid with dimensions $(a, b, c) = (0.8, 0.8, 1.4398)$ and parameters $\mu_1 = \mu_2 = \log(5), \sigma^2 = 1, \gamma^2 = 0.2$. *Top row:* realisation of the process with X_1 and X_2 being represented by the red and black points respectively. The three figures show the same realisation from the same perspectives of the ellipsoid, the left figure displays the data without any surface intensity, the middle figure displays the realisation of Y_1 as the surface intensity and the right figure displays the realisation of Y_2 as the surface intensity: high intensity is indicated by yellow whilst blue indicates low intensity. *Middle row:* plots of the functional summary statistics where random rotations are used to construct the envelope and *bottom row:* plots of the functional summary statistics where simulates from a homogeneous Poisson process are used to construct the envelope. Solid black line is the observed functional summary statistic, dashed red line is the theoretical value under independence and the grey area indicates the 2.5% lower and 97.5% upper simulation envelope.

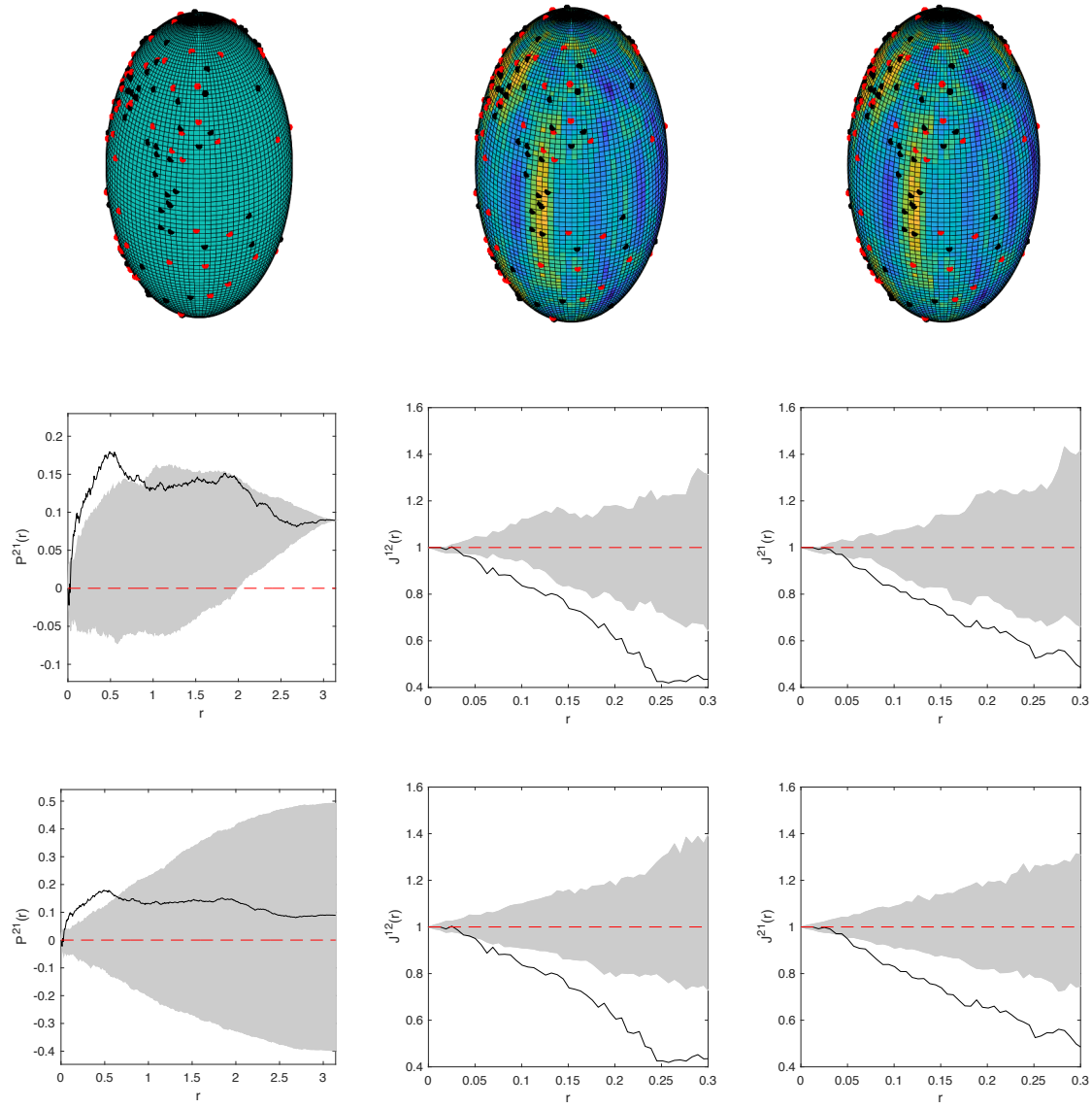


Figure 7.11: Example of a homogeneous attractive bivariate LGCP on an ellipsoid with dimensions $(a, b, c) = (0.8, 0.8, 1.4398)$ and parameters $\mu_1 = \mu_2 = \log(5), a_1 = a_2 = 1, \sigma^2 = 1, \gamma^2 = 0.2$. *Top row:* realisation of the process with X_1 and X_2 being represented by the red and black points respectively. The three figures show the same realisation from the same perspectives of the ellipsoid, the left figure displays the data without any surface intensity, the middle figure displays the realisation of Y_1 as the surface intensity and the right figure displays the realisation of Y_2 as the surface intensity: high intensity is indicated by yellow whilst blue indicates low intensity. *Middle row:* plots of the functional summary statistics where random rotations are used to construct the envelope and *bottom row:* plots of the functional summary statistics where simulates from a homogeneous Poisson process are used to construct the envelope. Solid black line is the observed functional summary statistic, dashed red line is the theoretical value under independence and the grey area indicates the 2.5% lower and 97.5% upper simulation envelope.

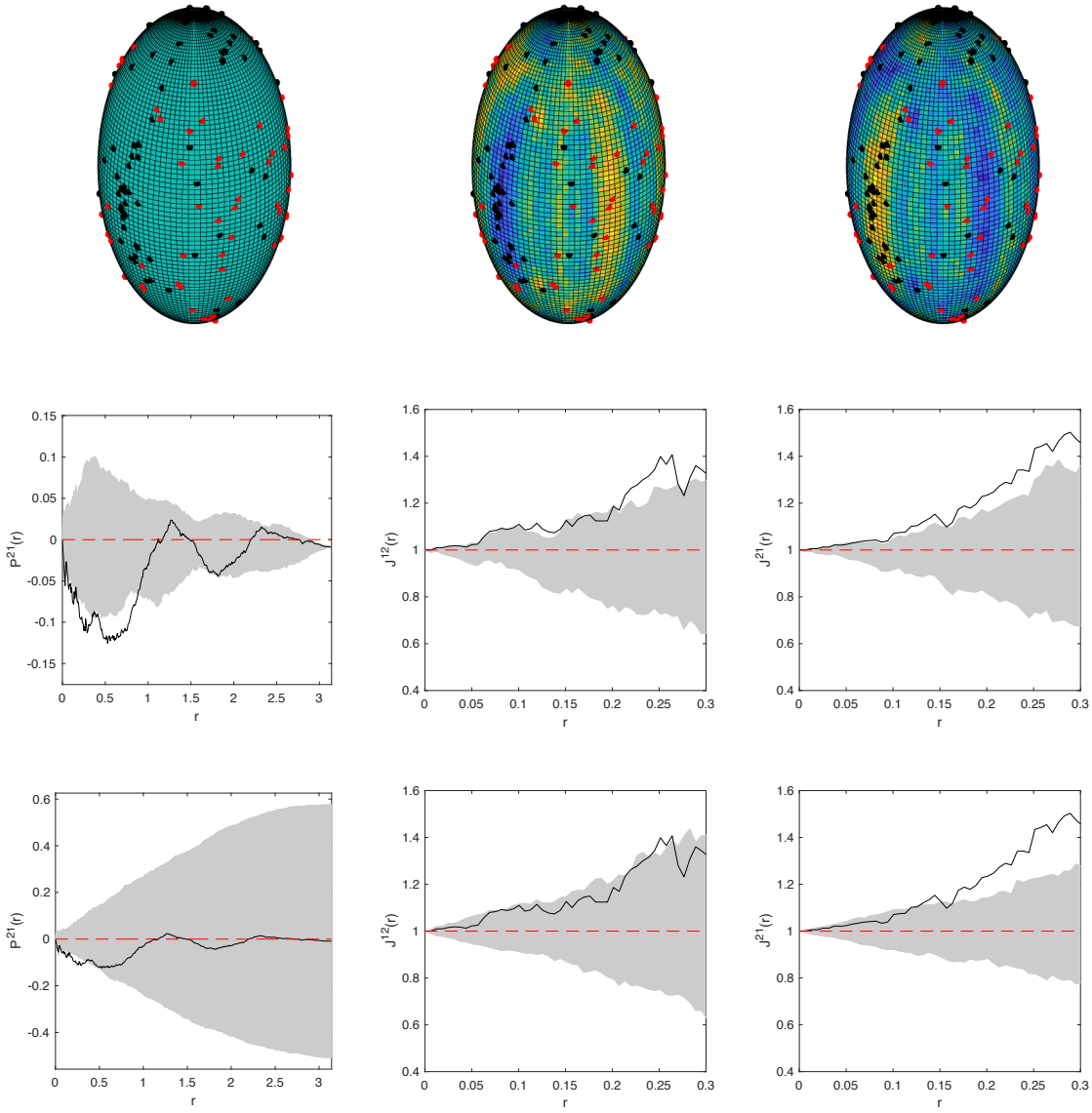


Figure 7.12: Example of a homogeneous repulsive bivariate LGCP on an ellipsoid with dimensions $(a, b, c) = (0.8, 0.8, 1.4398)$ and parameters $\mu_1 = \mu_2 = \log(5), a_1 = 1, a_2 = -1, \sigma^2 = 1, \gamma^2 = 0.2$. *Top row:* realisation of the process with X_1 and X_2 being represented by the red and black points respectively. The three figures show the same realisation from the same perspectives of the ellipsoid, the left figure displays the data without any surface intensity, the middle figure displays the realisation of Y_1 as the surface intensity and the right figure displays the realisation of Y_2 as the surface intensity: high intensity is indicated by yellow whilst blue indicates low intensity. *Middle row:* plots of the functional summary statistics where random rotations are used to construct the envelope and *bottom row:* plots of the functional summary statistics where simulates from a homogeneous Poisson process are used to construct the envelope. Solid black line is the observed functional summary statistic, dashed red line is the theoretical value under independence and the grey area indicates the 2.5% lower and 97.5% upper simulation envelope.

An example is given in Figure 7.14 where the results show that the observed functional summary statistics lie well within the simulation envelopes as expected.

3. Attractive bivariate LGCP where we set $\sigma^2 = 1, \gamma^2 = 0.2, a_1 = a_2 = 1$ and the mean functions of the first and second components are $\mu_1(\mathbf{x}) = \mu_2(\mathbf{x}) = \log(5) + x^2$ with $\mathbf{x} = (x, y, z)^T \in \mathbb{D}$. An example is given in Figure 7.15 where the results show that the observed functional summary statistics typically lie above the simulation envelopes for P^{21} and below for both J^{12} and J^{21} highlighting the attractive nature of the process.

4. Repulsive bivariate LGCP where we set $\sigma^2 = 1, \gamma^2 = 0.2, a_1 = 1, a_2 = -1$ and the mean functions of the first and second components are $\mu_1(\mathbf{x}) = \log(5) + y^2$ and $\mu_2(\mathbf{x}) = \log(5) + x^2$ respectively and $\mathbf{x} = (x, y, z)^T \in \mathbb{D}$. An example is given in Figure 7.16 where the results show that the observed functional summary statistics typically lie below the simulation envelopes for P^{21} and above for both J^{12} and J^{21} highlighting the repulsive nature of the process.

For the last two examples, under the Poisson null the plot of P^{21} (bottom left plot in Figures 7.11 and 7.16), only lies marginally outside the simulation envelope. Based solely on these plots we would suppose there is little evidence to suggest deviations away from independence, but when also considering the plots of J^{12} and J^{21} there is clear indication of attraction or repulsion in the process. This example highlights the importance of examining many different second order statistics [Baddeley and Silverman, 1984].

7.6 Discussion

In this chapter we discussed our final contribution to the spatial statistics literature. We detailed how functional summary statistics can be constructed for marked spheroidal processes by extending the theory from $\mathbb{R}^d$ [Cronie and van Lieshout, 2016]. Under the framework of marked processes we specifically focused on multivariate point patterns highlighting that these functional summary statistics were capable of capturing attractive or repulsive behaviour between its components in an observed point process. Using the Mapping Theorem [Last and Penrose, 2018] we were able to extend this methodology further for point processes lying on the surface of convex shapes. Numerical examples showed that the functional summary statistics constructed for marked point processes on convex shapes are still capable of capturing attractive and repulsive behaviour even when the assumptions of IRWMI or SOIRWI do not hold for the mapped point process.

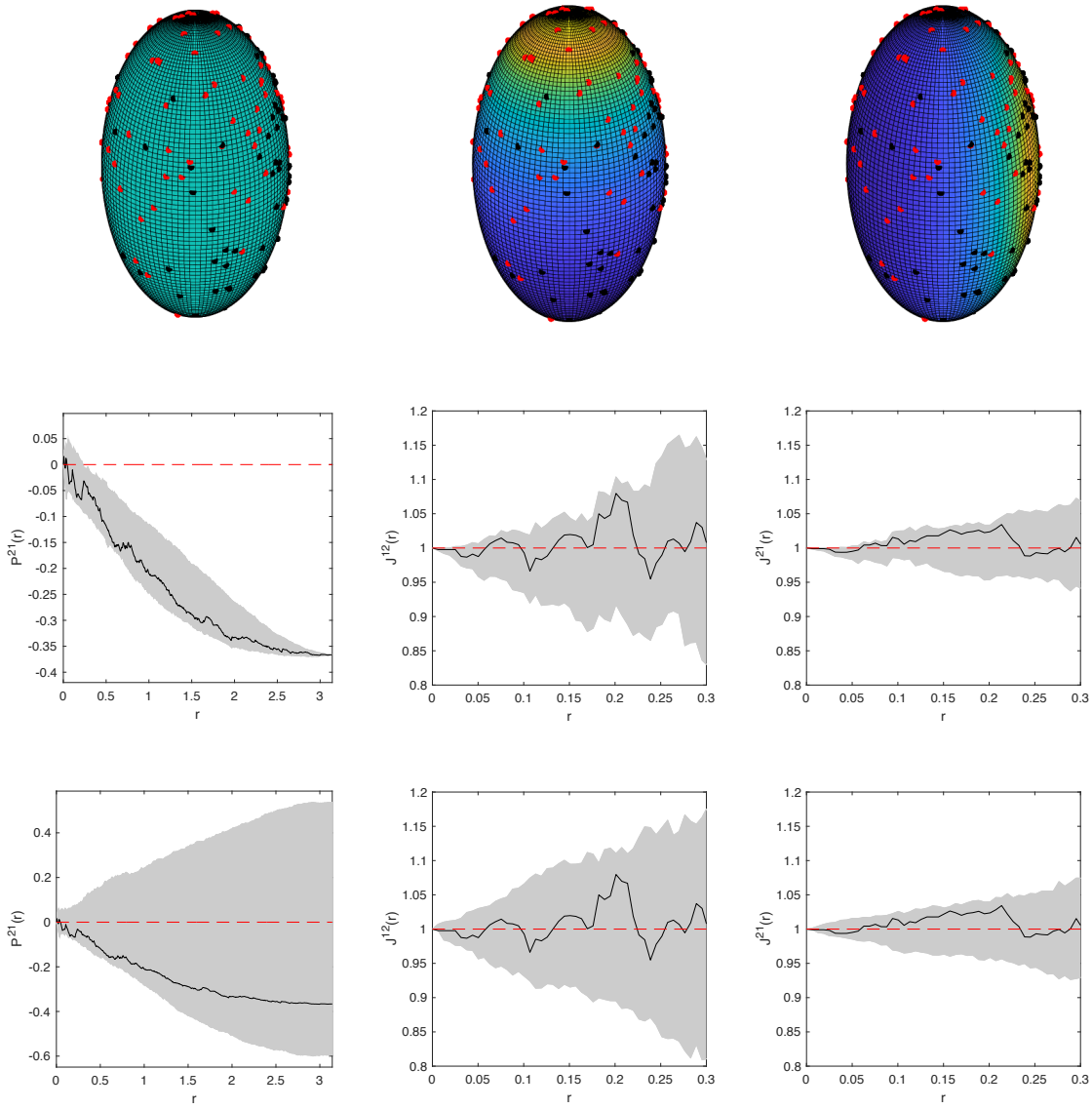


Figure 7.13: Example of an inhomogeneous independent bivariate Poisson process on an ellipsoid dimensions $(a, b, c) = (0.8, 0.8, 1.4398)$ and intensity functions $\rho_1 = \exp(\log(6) + z)$ and $\rho_2 = \exp(\log(6) + 2x)$ for the first and second component respectively. *Top row:* realisation of the process with X_1 and X_2 being represented by the red and black points respectively. The three figures show the same realisation from the same perspectives of the ellipsoid. The left figure has no surface intensity, the middle figure displays ρ_1 as its surface intensity, whilst the right figure displays ρ_2 as its surface intensity: high intensity is indicated by yellow whilst low intensity is indicated by blue. *Middle row:* plots of the functional summary statistics where random rotations are used to construct the envelope and *bottom row:* plots of the functional summary statistics where simulates from an inhomogeneous Poisson process with the fitted intensity function are used to construct the envelope. Solid black line is the observed functional summary statistic, dashed red line is the theoretical value under independence and the grey area indicates the 2.5% lower and 97.5% upper simulation envelope.

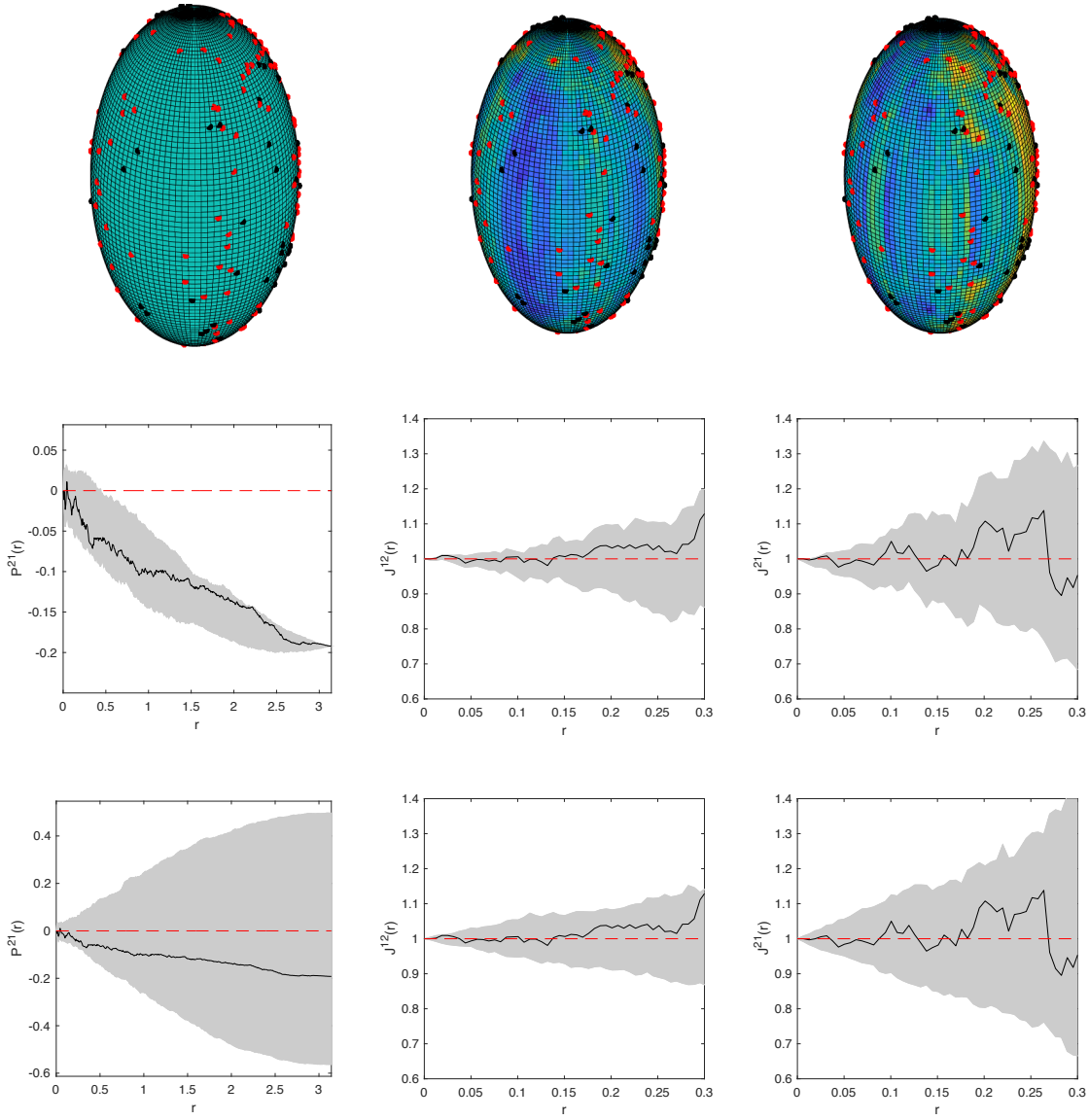


Figure 7.14: Example of an inhomogeneous independent bivariate LGCP on an ellipsoid with dimensions $(a, b, c) = (0.8, 0.8, 1.4398)$ and parameters $\mu_1(\mathbf{x}) = \mu_2(\mathbf{x}) = \log(6) + x, \sigma^2 = 1, \gamma^2 = 0.2$. *Top row:* realisation of the process with X_1 and X_2 being represented by the red and black points respectively. The three figures show the same realisation from the same perspectives of the ellipsoid, the left figure displays the data without any surface intensity, the middle figure displays the realisation of Y_1 as the surface intensity and the right figure displays the realisation of Y_2 as the surface intensity: high intensity is indicated by yellow whilst blue indicates low intensity. *Middle row:* plots of the functional summary statistics where random rotations are used to construct the envelope and *bottom row:* plots of the functional summary statistics where simulates from an inhomogeneous Poisson process with the fitted intensity function are used to construct the envelope. Solid black line is the observed functional summary statistic, dashed red line is the theoretical value under independence and the grey area indicates the 2.5% lower and 97.5% upper simulation envelope.

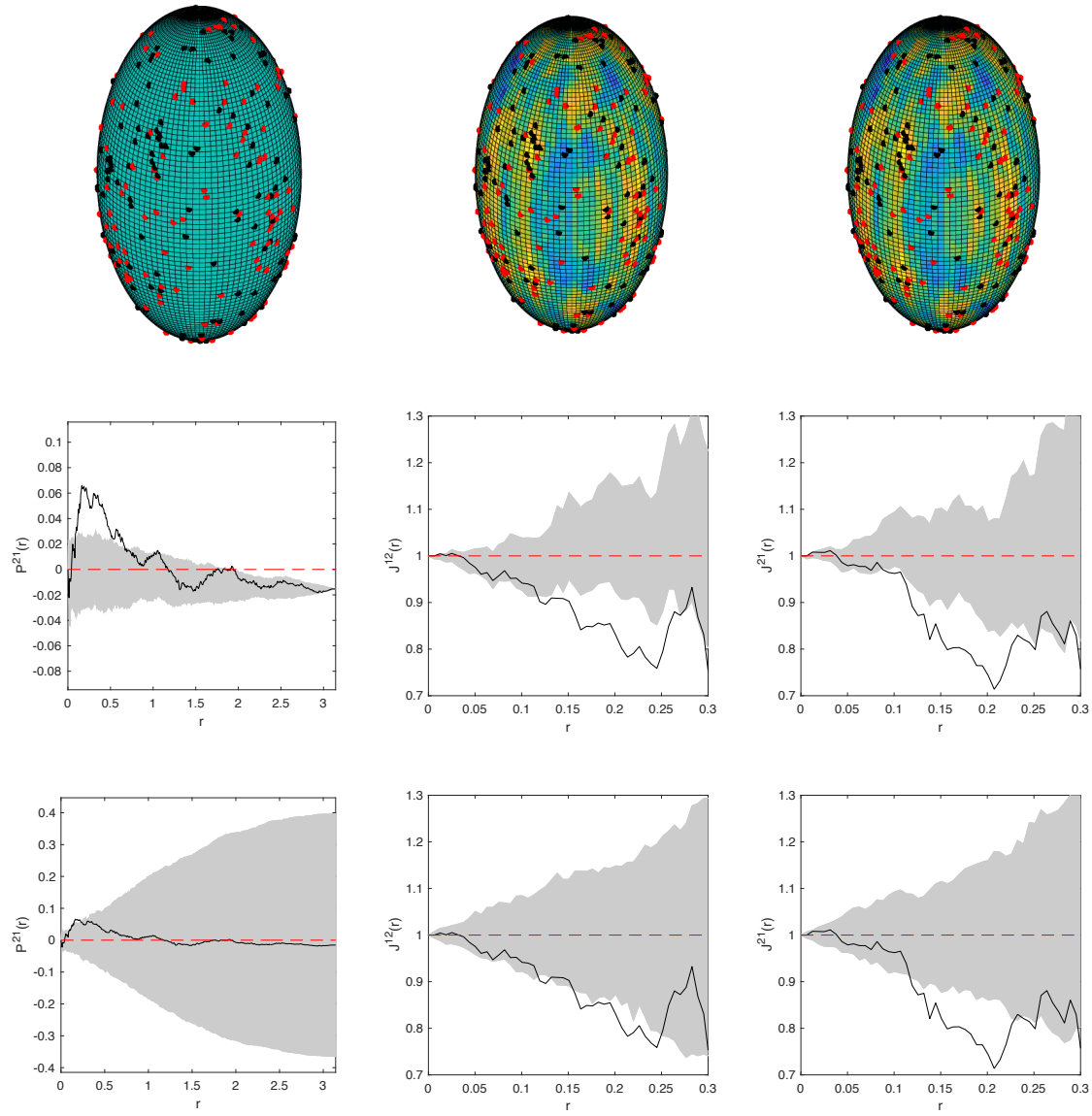


Figure 7.15: Example of an inhomogeneous attractive bivariate LGCP on an ellipsoid with dimensions $(a, b, c) = (0.8, 0.8, 1.4398)$ and parameters $\mu_1(\mathbf{x}) = \mu_2(\mathbf{x}) = \log(6) + x^2, a_1 = a_2 = 1, \sigma^2 = 1, \gamma^2 = 0.2$. *Top row:* realisation of the process with X_1 and X_2 being represented by the red and black points respectively. The three figures show the same realisation from the same perspectives of the ellipsoid, the left figure displays the data without any surface intensity, the middle figure displays the realisation of Y_1 as the surface intensity and the right figure displays the realisation of Y_2 as the surface intensity: high intensity is indicated by yellow whilst blue indicates low intensity. *Middle row:* plots of the functional summary statistics where random rotations are used to construct the envelope and *bottom row:* plots of the functional summary statistics where simulates from an inhomogeneous Poisson process with the fitted intensity function are used to construct the envelope. Solid black line is the observed functional summary statistic, dashed red line is the theoretical value under independence and the grey area indicates the 2.5% lower and 97.5% upper simulation envelope.

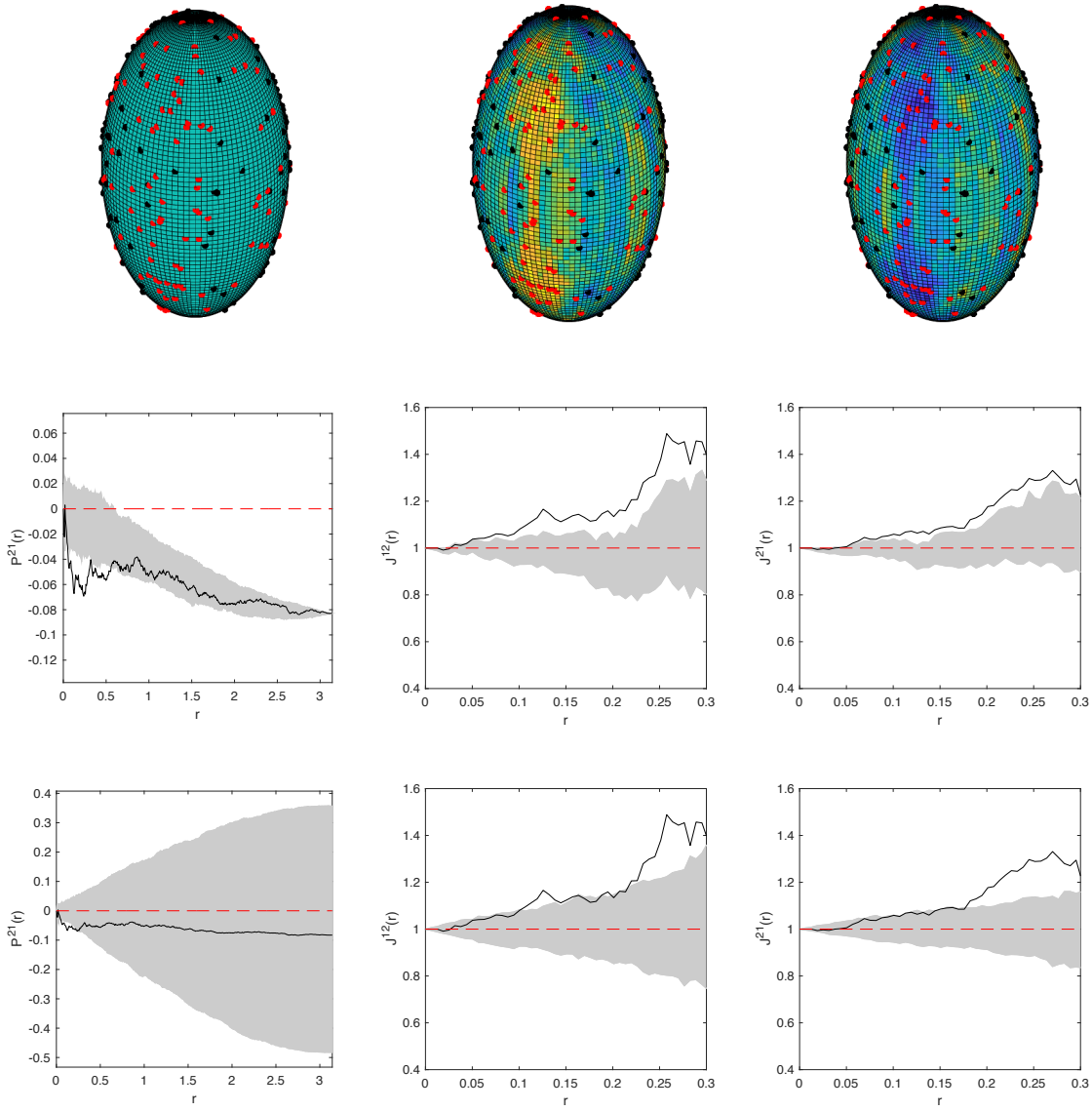


Figure 7.16: Example of an inhomogeneous repulsive bivariate LGCP on an ellipsoid with dimensions $(a, b, c) = (0.8, 0.8, 1.4398)$ and parameters $\mu_1(\mathbf{x}) = \log(6) + y^2, \mu_2(\mathbf{x}) = \log(6) + x^2, a_1 = 1, a_2 = -1, \sigma^2 = 1, \gamma^2 = 0.2$. *Top row:* realisation of the process with X_1 and X_2 being represented by the red and black points respectively. The three figures show the same realisation from the same perspectives of the ellipsoid, the left figure displays the data without any surface intensity, the middle figure displays the realisation of Y_1 as the surface intensity and the right figure displays the realisation of Y_2 as the surface intensity: high intensity is indicated by yellow whilst blue indicates low intensity. *Middle row:* plots of the functional summary statistics where random rotations are used to construct the envelope and *bottom row:* plots of the functional summary statistics where simulates from an inhomogeneous Poisson process with the fitted intensity function are used to construct the envelope. Solid black line is the observed functional summary statistic, dashed red line is the theoretical value under independence and the grey area indicates the 2.5% lower and 97.5% upper simulation envelope.

8

Conclusions

The focus of this thesis has been to develop theory and methodology for the analysis of point processes outside of a Euclidean space, a typical assumption made in the current literature. We summarise the contributions made and discuss how this work can be extended and future avenues that could be fruitfully pursued.

In Chapter 3 we reviewed the current literature on spheroidal point processes making particular reference to the discussion of the inhomogeneous K-function given by Møller and Rubak [2016], Lawrence et al. [2016] which is of importance in the subsequent chapters. Building on the work of van Lieshout [2011] we have defined IRWMI processes and provide extensions of the inhomogeneous F, H, and J-functions from $\mathbb{R}^{2,3}$ to $\mathbb{S}^2$ demonstrating them on three types of nonstationary processes over the sphere: Poisson, LGCP, and location dependent thinning. The development of these functional summary statistics serve as the foundation for the analysis of point patterns in the two subsequent chapters which discuss how to construct functional summary statistics for Poisson processes over convex shapes and formal hypothesis testing for CSR over convex shapes.

Chapter 4 presents our first major contribution. We demonstrated that construction of functional summary statistics outside of symmetric spaces such as $\mathbb{R}^d$ and $\mathbb{S}^{d-1}$ is non-trivial. We show that for Poisson point processes existing on convex shapes it is in fact possible to construct functional summary statistics through the Mapping Theorem [Kingman, 1993]. This allows for analysis to be conducted on a sphere where rotational symmetries can be exploited and thus functional summary statistics can be well-defined. We have demonstrated

the methodology through simulation envelopes demonstrating that the described summary statistics are able to identify when a process is truly CSR on $\mathbb{D}$. The summary statistics are also capable of highlighting when a process may be clustered or regular based on their deviations from an assumed CSR process.

Building on Chapter 4, Chapter 5 discusses a formal Monte Carlo approach to hypothesis testing of CSR on a convex surface. We derived two test statistics: one based on the work of Lawrence [2018] and the other based on a standardisation considered in Lagache et al. [2013]. Properties of the latter test statistic were derived. Both statistics were investigated for their efficacy in determining CSR through an extensive simulation study across ellipsoids of varying dimensions and processes that exhibited differing levels of clustering or regularity. We showed that T_1 outperforms T_2 (see Equation 5.4) for the majority of the experiments and that calculation of T_1 benefits from quicker computational calculation. Based on these results we suggested that T_1 would be a more favourable test statistic to use when testing CSR but this should not be at the expense of a critical examination of envelope plots [Diggle, 2003].

If a point pattern is suspected to not be CSR we would then be interested in recovering the intensity function of the point pattern. Chapter 6 discusses precisely this when the space is a Riemannian manifold. In this chapter we extended the theory of nonparametric kernel intensity estimation for point processes from a Euclidean space endowed with the standard Euclidean metric to a Riemannian manifold endowed with a Riemmanian metric. We have provided formulas for the bias and variance for the kernel intensity estimators under a Poisson assumption whilst also providing asymptotic results when the expected number of points across the manifold increases. Furthermore, we have also discussed two bandwidth selection protocols Baddeley et al. [2015], Cronie and Van Lieshout [2018]. We have investigated these two selection criteria through an extensive simulation study across ellipsoids of varying dimensions. Based on the simulation studies we can provide similar tentative conclusions as those drawn by Cronie and Van Lieshout [2018]: in the Poisson setting one should use the cross validation criteria, whilst in the clustered setting Cronie's selection criteria is more appropriate. In the regular setting cross validation seems to outperform Cronie's criteria in general, although this difference is small in most cases. In addition to this, Cronie's approach is computationally more efficient and therefore in practice would suggest using Cronie's selection criteria unless the pattern shows significant regularity, in which case cross validation would be more appropriate.

The final contribution made discusses how summary statistics can be constructed for multitype processes on convex shapes. In Chapter 7 we discussed how functional summary

statistics can be constructed for marked spheroidal processes by extending the theory from $\mathbb{R}^d$ [Cronie and van Lieshout, 2016] to the current setting. We focused on multitype point processes with the goal of determining whether a pattern exhibits attractiveness or repulsiveness between its component. Examples are explored on the sphere demonstrating that these functional summary statistics are able to capture these behaviours of point patterns. Using the Mapping Theorem [Last and Penrose, 2018] we were able to further extend this methodology for point processes lying on the surface of convex shapes. Again, examples show that the functional summary statistics constructed for multitype point processes on convex shapes are still capable of capturing attractive and repuslive behaviour even when the assumptions of IRWMI or SOIRWI may be violated.

8.1 Future work

There are a number of natural extensions following our work. The major focus here has been on the development of tools for exploratory data analysis and the next step would be to consider modelling techniques for point processes observed on general spaces. Due the difficulty of defining summary statistics for point process that are not Poisson we would expect techniques such as minimum contrast [Møller and Waagepetersen, 2003] to be ineffective. We anticipate that Markov processes on convex shapes can be fit using maximum-likelihood, psuedo-likelihood, and logistic regression [Baddeley and Turner, 1998, Møller and Waagepetersen, 2003, Baddeley et al., 2014], whilst LGCP can be fit using maximum-likelihood and pseudo-likelihood techniques also [Møller and Waagepetersen, 2003]. Typically following model fitting determining whether the model is adequate is suitable follows: for spatial point patterns this can be assessed through functional summary statistics where the observed is compared to realisations generated from the fitted model. Although construction of summary statistics for general point processes is not possible on convex shapes, it can be shown that both Markov and Cox processes remain Markov and Cox respectively when mapped to the sphere, this fact coupled with the thinning technique developed by Møller and Schoenberg [2010] may allow for functional summary statistics to be constructed that can be used for model validation. The thinning technique of Møller and Schoenberg [2010] was successfully applied to spheroidal LGCP in Cuevas-Pacheco and Møller [2018] to determine whether the model fit well.

An extension of the work in Chapter 6 would be to relax the Riemannian assumption. It may not always be plausible to suppose a smooth Riemannian structure on the underlying surface, instead a more practical approach would be to use a piecewise linear approximation

of the surface which may be more easily obtained in real world applications. In such a setting we expect the theory can be easily adapted and the results still hold. If in addition to this we also drop the compactness assumption and instead suppose we have a manifold with a boundary this allows for new types of data to be analysed such as point patterns that exist on uneven surfaces, e.g. tree locations on mountainous or hilly regions. Under the assumption of a Riemannian manifold it may be possible to construct geometrically corrected functional summary statistics using the volume density function introduced in Appendix D.1. The arguments follow similarly to those made by Ang et al. [2012], Rakshit et al. [2017] for point processes on linear networks and these ideas will be pursued in upcoming work.

In Chapter 7 we focused on testing independence between components of a multitype point pattern. Another hypothesis that is commonly considered is *random labelling*: that the marks are all IID and do not depend upon the ground process. These ideas have been explored in a number of works including van Lieshout and Baddeley [1999], Cronie and van Lieshout [2016], Iftimi et al. [2019] for spatial and spatio-temporal processes in $\mathbb{R}^d$ and could be extended to processes lying on general surfaces.

References

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

P. Alonso-Ruiz and E. Spodarev. Estimation of entropy for Poisson marked point processes. *Advances in Applied Probability*, 49(1):258–278, 2017.

Q. W. Ang, A. J. Baddeley, and G. Nair. Geometrically corrected second order analysis of events on a linear network, with applications to ecology and criminology. *Scandinavian Journal of Statistics*, 39(4):591–617, 2012.

A. J. Baddeley and R. D. Gill. Kaplan-Meier estimators of distance distributions for spatial point processes. *Annals of Statistics*, 25(1):263–292, 1997.

A. J. Baddeley and B. W. Silverman. A cautionary example on the use of second-order methods for analyzing point patterns. *Biometrics*, 40(4):1089–1093, 1984.

A. J. Baddeley and R. Turner. Practical maximum pseudolikelihood for spatial point patterns. *Advances in Applied Probability*, 30(2):273, 1998.

A. J. Baddeley, J. Møller, and R. Waagepetersen. Non- and semi-parametric estimation of interaction in inhomogeneous point patterns. *Statistica Neerlandica*, 54(3):329–350, 2000.

A. J. Baddeley, J. F. Coeurjolly, E. Rubak, and R. Waagepetersen. Logistic regression for spatial gibbs point processes. *Biometrika*, 101(2):377–392, 2014.

A. J. Baddeley, E. Rubak, and R. Turner. *Spatial Point Patterns: Methodology and Applications with R*. CRC press, Florida, 2015.

M. Berman and P. J. Diggle. Estimating weighted integrals of the second-order intensity of a spatial point process. *Journal of the Royal Statistical Society: Series B*, 51(1):81–92, 1989.

J. E. Besag. Discussion of the paper by Ripley. *Journal of the Royal Statistical Society: Series B*, 39(2):193–195, 1977.

A. L. Besse. *Manifolds all of whose Geodesics are Closed*. Springer, Berlin, Heidelberg, 1978.

P. Billingsley. *Probability and Measure*. John Wiley & Sons, New York, second edition, 1986.

A. Brix and J. Moller. Space-time multi type log Gaussian Cox processes with a view to modelling weeds. *Scandinavian Journal of Statistics*, 28(3):471–488, 2001.

C. Cabriel, N. Bourg, P. Jouchet, G. Dupuis, C. Leterrier, A. Baron, M.-A. Badet-Denisot, B. Vauzeilles, E. Fort, and S. Lévêque-Fort. Combining 3D single molecule localization strategies for reproducible bioimaging. *Nature Communications*, 10(1):1980, 2019.

M. P. d. Carmo. *Riemannian Geometry*. Birkhäuser, Boston, 1992.

M. P. d. Carmo. *Differential Geometry of Curves and Surfaces*. Dover Publications, New York, second edition, 2016.

S. N. Chiu, D. Stoyan, W. S. Kendall, and J. Mecke. *Stochastic Geometry and its Applications*. John Wiley & Sons, Chichester, second edition, 1995.

J. F. Coeurjolly, J. Møller, and R. Waagepetersen. A tutorial on Palm distributions for spatial point processes. *International Statistical Review*, 85(3):404–420, 2017.

E. A. Cohen, A. V. Abraham, S. Ramakrishnan, and R. J. Ober. Resolution limit of image analysis algorithms. *Nature Communications*, 10(1):793, 2019.

N. Cressie. *Statistics for Spatial Data*. John Wiley & Sons, New York, revised edition, 1993.

O. Cronie and M. N. van Lieshout. Summary statistics for inhomogeneous marked point processes. *Annals of the Institute of Statistical Mathematics*, 68(4):905–928, 2016.

O. Cronie and M. N. Van Lieshout. A non-model-based approach to bandwidth selection for kernel estimators of spatial intensity functions. *Biometrika*, 105(2):455–462, 2018.

O. Cronie, M. Moradi, and J. Mateu. Inhomogeneous higher-order summary statistics for point processes on linear networks. *Statistics and Computing*, 30(5):1221–1239, 2020.

L. Cucala. *Espacements bidimensionnels et données entachées d'erreurs dans l'analyse des processus ponctuels spatiaux*. PhD thesis, Université des Sciences Sociales - Toulouse I, France, 2006.

L. Cucala. Intensity estimation for spatial point processes observed with noise. *Scandinavian Journal of Statistics*, 35(2):322–334, 2008.

F. Cuevas, D. Allard, and E. Porcu. Fast and exact simulation of Gaussian random fields defined on the sphere cross time. *Statistics and Computing*, 30(1):187–194, 2020.

F. Cuevas-Pacheco and J. Møller. Log Gaussian Cox processes on the sphere. *Spatial Statistics*, 26:69–82, 2018.

D. Daley and D. Vere-Jones. *An Introduction to the Theory of Point Processes. Volume 1: Elementary Theory and Methods*. Springer, New York, second edition, 2003.

M. Deserno. How to generate equidistributed points on the surface of a sphere. 2004. URL https://www.cmu.edu/biolphys/deserno/pdf/sphere_equi.pdf.

P. J. Diggle. A kernel method for smoothing point process data. *Journal of the Royal Statistical Society: Series C*, 34(2):138–147, 1985.

P. J. Diggle. Displaced amacrine cells in the retina of a rabbit: analysis of a bivariate spatial point pattern. *Journal of Neuroscience Methods*, 18(1-2):115–125, 1986.

P. J. Diggle. *Statistical Analysis of Spatial Point Patterns*. Arnold, London, second edition, 2003.

P. J. Diggle. *Statistical Analysis of Spatial and Spatio-temporal Point Patterns*. CRC press, Florida, third edition, 2014.

A. K. Gustavsson, P. N. Petrov, M. Y. Lee, Y. Shechtman, and W. E. Moerner. 3D single-molecule super-resolution microscopy with a tilted light sheet. *Nature Communications*, 9(1):123, 2018.

K. H. Hanisch. Some remarks on estimators of the distribution function of nearest neighbour distance in stationary spatial point processes. *Series Statistics*, 15(3):409–412, 1984.

G. Henry and D. Rodriguez. Kernel density estimation on Riemannian manifolds: Asymptotic results. *Journal of Mathematical Imaging and Vision*, 34(3):235–239, 2009.

W. Hoeffding. *Breakthroughs in Statistics: Foundations and Basic Theory*, chapter A Class of Statistics with Asymptotically Normal Distribution, pages 308–334. Springer, New York, 1992.

A. Iftimi, O. Cronie, and F. Montes. The second-order analysis of marked spatio-temporal point processes, with an application to earthquake data. *arXiv:1611.04808*, 2016.

A. Iftimi, O. Cronie, and F. Montes. Second-order analysis of marked inhomogeneous spatiotemporal point processes: Applications to earthquake data. *Scandinavian Journal of Statistics*, 46(3):661–685, 2019.

P. Jagers. On Palm probabilities. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 26(1):17–32, 1973.

E. B. V. Jensen and L. S. Nielsen. A review on inhomogeneous Markov point processes. *Lecture Notes-Monograph Series*, 37:297–318, 2001.

M. C. Jones. Simple boundary correction for kernel density estimation. *Statistics and Computing*, 3(3):135–146, 1993.

M. Jun, C. Schumacher, and R. Saravanan. Global multivariate point pattern models for rain type occurrence. *Spatial Statistics*, 31:100355, 2019.

C. F. F. Karney. Algorithms for geodesics. *Journal of Geodesy*, 87(1):43–55, 2012.

C. F. F. Karney. *GeographicLib*. MATLAB Central File Exchange, 2017. URL https://www.mathworks.com/matlabcentral/fileexchange/50605-geographiclib. Retrieved October 5, 2017.

J. F. C. Kingman. *Poisson Processes*. Oxford University Press, Oxford, 1993.

N. P. Kopytov and E. A. Mityushov. The method for uniform distribution of points on surfaces in multi-dimensional Euclidean space. *Intellectual Archive*, 2013. URL http://www.intellectualarchive.com/?link=item&id=1170.

T. Lagache, G. Lang, N. Sauvonnet, and J. C. Olivo-Marin. Analysis of the spatial organization of molecules with robust statistics. *PLoS ONE*, 8(12):1–7, 2013.

A. Lang, J. Potthoff, M. Schlather, and D. Schwab. Continuity of random fields on Riemannian manifolds. *Communications on Stochastic Analysis*, 10(2):185–193, 2016.

G. Lang and E. Marcon. Testing randomness of spatial point patterns with the Ripley statistic. *ESAIM: Probability and Statistics*, 17:767–788, 2012.

G. Last and M. Penrose. *Lectures on the Poisson Process*. Cambridge University Press, Cambridge, 2018.

T. J. Lawrence. Point pattern analysis on a sphere. Master's thesis, The University of Western Australia, 2018.

T. J. Lawrence, A. J. Baddeley, R. K. Milne, and G. Nair. Point pattern analysis on a region of a sphere. *Stat*, 5(1):144–157, 2016.

A. Le Brigant and S. Puechmorel. Approximation of densities on Riemannian manifolds. *Entropy*, 21(1):43, 2019.

J. M. Lee. *Introduction to Riemannian Manifolds*. Springer, Cham, 2018.

H. Lotwick and B. Silverman. Methods for analysing spatial processes of several types of points. *Journal of the Royal Statistical Society: Series B*, 44(3):406–413, 1982.

B. Matérn. *Spatial Variation*. Lecture Notes in Statistics. Springer, Berlin, second edition, 2013.

G. McSwiggan, A. J. Baddeley, and G. Nair. Kernel density estimation on a linear network. *Scandinavian Journal of Statistics*, 44(2):324–345, 2017.

J. Møller and E. Rubak. Functional summary statistics for point processes on the sphere with an application to determinantal point processes. *Spatial Statistics*, 18(A):4–23, 2016.

J. Møller and F. P. Schoenberg. Thinning spatial point processes into poisson processes. *Advances in Applied Probability*, 42(2):347–358, 2010.

J. Møller and R. Waagepetersen. Some recent developments in statistics for spatial point patterns. *Annual Review of Statistics and its Application*, 4(1):317–342, 2017.

J. Møller and R. P. Waagepetersen. Log Gaussian Cox Processes. *Scandinavian Journal of Statistics*, 25(3):451–482, 1998.

J. Møller and R. P. Waagepetersen. *Statistical Inference and Simulation for Spatial Point Processes*. CRC Press, Florida, 2003.

J. Møller, M. Nielsen, E. Porcu, and E. Rubak. Determinantal point process models on the sphere. *Bernoulli*, 24(2):1171–1201, 2018.

J. Møller, H. S. Christensen, F. Cuevas-Pacheco, and A. D. Christoffersen. Structured space-sphere point processes and K-functions. *Methodology and Computing in Applied Probability*, 23(2):569–591, 2021.

M. M. Moradi, F. J. Rodríguez-Cortés, and J. Mateu. On kernel-based intensity estimation of spatial point patterns on linear networks. *Journal of Computational and Graphical Statistics*, 27(2):302–311, 2018.

M. Myllymäki, T. Mrkvicka, P. Grabarnik, H. Seijo, and U. Hahn. Global envelope tests for spatial processes. *Journal of the Royal Statistical Society: Series B*, 79(2):381–404, 2017.

J. Nash. The imbedding problem for Riemannian manifolds. *Annals of Mathematics*, 63(1): 20–63, 1956.

C. Palm. Intensitatsschwankungen im fernsprechverker. *Ericsson Technics*, 44:1–189, 1943.

P. J. E. Peebles. *The large-scale structure of the universe*. Princeton University Press, Princeton, 1980.

B. Pelletier. Kernel density estimation on Riemannian manifolds. *Statistics & probability letters*, 73(3):297–304, 2005.

X. Pennec. Intrinsic statistics on Riemannian manifolds: Basic tools for geometric measurements. *Journal of Mathematical Imaging and Vision*, 25(1):127–154, 2006.

L. D. Pitt. Positively correlated normal variables are associated. *The Annals of Probability*, 10(2):496–499, 1982.

T. Rajala, D. J. Murrell, and S. C. Olhede. Detecting multivariate interactions in spatial point patterns with gibbs models and variable selection. *Journal of the Royal Statistical Society: Series C*, 67(5):1237–1273, 2018.

S. Rakshit, G. Nair, and A. J. Baddeley. Second-order analysis of point patterns on a network using any distance metric. *Spatial Statistics*, 22(1):129–154, 2017.

S. Rakshit, T. Davies, M. M. Moradi, G. McSwiggan, G. Nair, J. Mateu, and A. J. Baddeley. Fast kernel smoothing of point patterns on a large network using two-dimensional convolution. *International Statistical Review*, 87(3):531–556, 2019.

B. D. Ripley. Modelling Spatial Patterns. *Journal of the Royal Statistical Society: Series B*, 39(2):172–212, 1977.

B. D. Ripley. *Statistical Inference for Spatial Processes*. Cambridge University Press, Cambridge, 1991.

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

S. Said, L. Bombrun, Y. Berthoumieu, and J. H. Manton. Riemannian Gaussian distributions on the space of symmetric positive definite matrices. *IEEE Transactions on Information Theory*, 63(4):2153–2170, 2017.

M. Schlather, A. Malinowski, M. Oesting, D. Boecker, K. Strokorb, S. Engelke, J. Martini, F. Ballani, O. Moreva, J. Auel, P. J. Menck, S. Gross, U. Ober, P. Ribeiro, B. D. Ripley, R. Singleton, B. Pfaff, and R Core Team. *RandomFields: Simulation and Analysis of Random Fields*, 2020. URL https://cran.r-project.org/package=RandomFields. R package version 3.3.8.

B. W. Silverman. *Density Estimation for Statistics and Data Analysis*. CRC press, London, 1986.

D. Stoyan. On estimators of the nearest neighbour distance distribution function for stationary point processes. *Metrika*, 64(2):139–150, 2006.

D. Stoyan and P. Grabarnik. Second-order characteristics for stochastic structures connected with Gibbs point processes. *Mathematische Nachrichten*, 151(1):95–100, 1991.

D. Stoyan and H. Stoyan. On one of Matérn's hard-core point process models. *Mathematische Nachrichten*, 122(1):205–214, 1985.

D. Stoyan and H. Stoyan. Improving ratio estimators of second order point process characteristics. *Scandinavian Journal of Statistics*, 27(4):641–656, 2000.

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

L. W. Tu. *An Introduction to Manifolds*. Springer, New York, second edition, 2011.

M. N. M. van Lieshout. A J -function for marked point patterns. *Annals of the Institute of Statistical Mathematics*, 58(2):235–259, 2006.

M. N. M. van Lieshout. A J-function for inhomogeneous point processes. *Statistica Neerlandica*, 65(2):183–201, 2011.

M. N. M. van Lieshout. On estimation of the intensity function of a point process. *Methodology and Computing in Applied Probability*, 14(3):567–578, 2012.

M. N. M. van Lieshout and A. J. Baddeley. A nonparametric measure of spatial interaction in point patterns. *Statistica Neerlandica*, 50(3):344–361, 1996.

M. N. M. van Lieshout and A. J. Baddeley. Indices of dependence between types in multivariate point patterns. *Scandinavian Journal of Statistics*, 26(4):511–532, 1999.

M. P. Wand and M. C. Jones. *Kernel smoothing*. CRC press, New York, 1994.

S. Ward, H. Battey, and E. A. K. Cohen. Estimation of the intensity function of a spatial point process on a Riemannian manifold. (manuscript in preparation), 2021a.

S. Ward, E. A. K. Cohen, and N. Adams. Testing for complete spatial randomness on three dimensional bounded convex shapes. *Spatial Statistics*, 41:100489, 2021b.

S. Ward, E. A. K. Cohen, and N. Adams. Functional summary statistics for multitype point patterns on three dimensional convex surfaces. (manuscript in preparation), 2021c.

S. D. White. The hierarchy of correlation functions and its relation to other measures of galaxy clustering. *Monthly Notices of the Royal Astronomical Society*, 186(2):145–154, 1979.

T. Wiegand, A. Huth, S. Getzin, X. Wang, Z. Hao, C. V. S. Gunatilleke, and I. A. U. N. Gunatilleke. Testing the independent species' arrangement assertion made by theories of stochastic geometry of biodiversity. *Proceedings of the Royal Society B: Biological Sciences*, 279(1741):3312–3320, 2012.

K. Wolter. *Introduction to Variance Estimation*. Springer, New York, first edition, 2007.

A

Appendix to Chapter 3

A.1 Proof of Theorem 3.5.1

Proof. Let $\{B_1, \ldots, B_n\}$ be a partition of $B_{\mathbb{S}^2}(\mathbf{o}, r)$ into n sets, such that as n increases the area of each B_i decreases. Then,

$$
\begin{aligned}
F(r) &= 1 - \mathbb{P}_{X \cap B_{\mathbb{S}^2}(\mathbf{o},r)}(\emptyset) \\
&= 1 - \mathbb{P}((X_{B_1} = \emptyset) \cap \cdots \cap (X_{B_n} = \emptyset)) \\
&= 1 - \left(1 - \sum_{i=1}^{n} \mathbb{P}(X_{B_i} \neq \emptyset) + \sum_{\substack{i,j=1 \\ i<j}}^{n} \mathbb{P}((X_{B_i} \neq \emptyset) \cap (X_{B_j} \neq \emptyset)) - \cdots \right) \qquad \text{(A.1)} \\
&= \sum_{i=1}^{n} \mathbb{P}(X_{B_i} \neq \emptyset) - \sum_{\substack{i,j=1 \\ i<j}}^{n} \mathbb{P}((X_{B_i} \neq \emptyset) \cap (X_{B_j} \neq \emptyset)) + \cdots , \qquad \text{(A.2)}
\end{aligned}
$$

where (A.1) follows from the inclusion-exclusion principle. Next we shall consider the first term. Define $a_n(x) = \sum_{i=1}^{n} \mathbb{1}[x_{B_i} \neq \emptyset]$ and $a(x) = \sum_{\mathbf{x} \in x} \mathbb{1}[\mathbf{x} \in B_{\mathbb{S}^2}(\mathbf{o}, r)]$ where $x \in N_{lf}$. Then it can easily be seen that a_n is a monotonically increasing sequence since as the number of partitions increases the number of partitions containing more than one point of x decreases. Further the maximimum of a_n is a. To see this consider small neighbourhoods of each $\mathbf{x} \in x$ such that these neighbourhoods are all disjoint. This is possible since we only consider locally

finite point configurations of N_{lf} which have unique points, i.e. no two elements of $x \in N_{lf}$ can be identical. For the space $N_{lf,\text{multi}} = \{x \in N_{lf} | \text{ there exists } \mathbf{x}, \mathbf{y} \in x \text{ such that } \mathbf{x} = \mathbf{y}\} \subset N_{lf}$ by assuming simplicity of X we define $\mathbb{P}(X \in N_{lf,\text{multi}}) = 0$. Label these $B_{\mathbf{x}}$, then $\mathbb{1}[x_{B_{\mathbf{x}}} \neq \emptyset] = 1$, $\forall \mathbf{x} \in x$ and so for this partition $a_n = a$ and cannot increase as this would require at least one point of x to be in two different disjoint $B_{\mathbf{x}}$: a contradiction. Hence as $n \to \infty$, $a_n \to a$. Therefore,

$$\begin{aligned}\sum_{i=1}^{n} \mathbb{P}(X_{B_i} \neq \emptyset) &= \mathbb{E} \sum_{i=1}^{n} \mathbb{1}[X_{B_i} \neq \emptyset] \\ &= \int_{N_{lf}} \sum_{i=1}^{n} \mathbb{1}[x_{B_i} \neq \emptyset] d\,\mathbb{P}(x).\end{aligned}$$

Since $a_n(x) \leq a(x)$ and by assumption $\alpha^{(1)} \equiv \alpha$ exists, and noticing that $\mathbb{E}[a(X)] = \mathbb{E} \sum_{\mathbf{x} \in X} \mathbb{1}[\mathbf{x} \in B_{\mathbb{S}^2}(\mathbf{o}, r)] = \alpha(B_{\mathbb{S}^2}(\mathbf{o}, r)) < \infty$, we can therefore apply the dominated convergence theorem when taking the limit as n increases and the volumes of B_i decrease, i.e.

$$\begin{aligned}\lim_{\substack{n \to \infty \\ \lambda_{\mathbb{S}^2}(B_i) \to 0}} \sum_{i=1}^{n} \mathbb{P}(X_{B_i} \neq \emptyset) &= \lim_{\substack{n \to \infty \\ \lambda_{\mathbb{S}^2}(B_i) \to 0}} \int_{N_{lf}} \sum_{i=1}^{n} \mathbb{1}[x_{B_i} \neq \emptyset] d\,\mathbb{P}(x) \\ &= \int_{N_{lf}} \lim_{\substack{n \to \infty \\ \lambda_{\mathbb{S}^2}(B_i) \to 0}} \sum_{i=1}^{n} \mathbb{1}[x_{B_i} \neq \emptyset] d\,\mathbb{P}(x) \\ &= \int_{N_{lf}} \sum_{\mathbf{y} \in x} \mathbb{1}[y \in B_{\mathbb{S}^2}(\mathbf{o}, r)] d\,\mathbb{P}(x) \\ &= \mathbb{E} \sum_{\mathbf{x} \in X} \mathbb{1}[x_{B_i} \in B_{\mathbb{S}^2}(\mathbf{o}, r)] \\ &= \alpha(B_{\mathbb{S}^2}(\mathbf{o}, r)),\end{aligned}$$

noting by our previous discussion that $\mathbb{P}(X \in N_{lf,\text{multi}}) = 0$ and so we can disregard elements of $N_{lf,\text{multi}}$. Hence this identity holds almost surely.

An identical approach can be used for the remaining terms of A.2 in the following. By

considering the k^{th} term of A.2 we have,

$$(-1)^{k+1} \sum_{\substack{i_1,\ldots,i_k=1 \\ i_1<\cdots<i_k}}^{n} \mathbb{P}((X_{B_{i_1}} \neq \emptyset) \cap \cdots \cap (X_{B_{i_k}} \neq \emptyset))$$
$$= \frac{(-1)^{k+1}}{k!} \sum_{i_1=1}^{n} \cdots \sum_{\substack{i_k=1 \\ i_k \notin \{i_1,\ldots,i_{k-1}\}}}^{n} \mathbb{P}((X_{B_{i_1}} \neq \emptyset) \cap \cdots \cap (X_{B_{i_k}} \neq \emptyset)).$$

Define,

$$a_n(x) = \sum_{i_1,\ldots,i_k \in \{1,\ldots,n\}}^{\neq} \mathbb{1}[x_{B_{i_1}} \neq \emptyset, \ldots, x_{B_{i_k}} \neq \emptyset]$$
$$a(x) = \sum_{\mathbf{x}_1,\ldots,\mathbf{x}_k \in x}^{\neq} \mathbb{1}[\mathbf{x}_1 \in B_{\mathbb{S}^2}(\mathbf{o}, r), \ldots, \mathbf{x}_k \in B_{\mathbb{S}^2}(\mathbf{o}, r)],$$

for $x \in N_{lf}$. By identical arguments as in the case for $k = 1$ we have that $a_n(x)$ is increasing for a shrinking partition of $B_{\mathbb{S}^2}(\mathbf{o}, r)$ and attains its maximum of $a(x)$. Thus by using the dominated convergence theorem again we have,

$$\lim_{\substack{n\to\infty \\ \lambda_{\mathbb{S}^2}(B_i)\to 0}} (-1)^{k+1} \sum_{\substack{i_1,\ldots,i_k=1 \\ i_1<\cdots<i_k}}^{n} \mathbb{P}((X_{B_{i_1}} \neq \emptyset) \cap \cdots \cap (X_{B_{i_k}} \neq \emptyset))$$
$$= \frac{-(-1)^k}{k!} \alpha^{(k)}(B_{\mathbb{S}^2}(\mathbf{o}, r), \ldots, B_{\mathbb{S}^2}(\mathbf{o}, r)),$$

and so gives the infinite series representation of the F - function.

The series representation for the H - function follows an identical argument to that of the F - function, instead using the factorial moment measure for the reduced Palm point process, $X^!_{\mathbf{x}}$. □

A.2 Proof of Corollary 3.5.2

Proof. The infinite series for F follows immediately from Theorem 3.5.1 and Equation 2.1. Further, by assumption we also know that,

$$\alpha^{!(n)}_{\mathbf{o}}(B_1 \times \cdots \times B_n) = \int_{B_1} \cdots \int_{B_n} \rho^{!(n)}_{\mathbf{o}}(\mathbf{x}_1, \ldots, \mathbf{x}_n) \lambda_{\mathbb{S}^2}(d\mathbf{x}_1) \cdots \lambda_{\mathbb{S}^2}(d\mathbf{x}_n).$$

Then we have the following relationship between the n^{th}-order product intensities of X and $X^!_{\mathbf{o}}$ (see for example Coeurjolly et al. ((2017))),

$$\rho^{!(n)}_{\mathbf{x}}(\mathbf{x}_1,\ldots,\mathbf{x}_m) = \frac{\rho^{(n+1)}(\mathbf{x},\mathbf{x}_1,\ldots,\mathbf{x}_n)}{\rho(\mathbf{x})}, \tag{A.3}$$

and so for our point process X,

$$\alpha^{!(n)}_{\mathbf{o}}(B_1\times\cdots\times B_n) = \int_{B_1}\cdots\int_{B_n}\frac{\rho^{(n+1)}_{\mathbf{o}}(\mathbf{o},\mathbf{x}_1,\ldots,\mathbf{x}_n)}{\rho}\lambda_{\mathbb{S}^2}(d\mathbf{x}_1)\cdots\lambda_{\mathbb{S}^2}(d\mathbf{x}_n),$$

and then by Theorem 3.5.1 we have the infinite series for the H-function in terms of the n^{th}-order product intensities. □

A.3 Proof of Theorem 3.5.4

Our proof follows identically to the proof of Proposition 4.2 in van Lieshout ((2006)), although we restrict our attention to $\mathbb{S}^2$ and only focus on unmarked point processes.

Proof. From Corollary 3.5.2 and assuming that all factorial moment intensities exist,

$$\begin{aligned}\rho(1-H(r)) =&\rho+\sum_{n=1}^{\infty}\frac{(-1)^n}{n!}\\ &\int_{B_{\mathbb{S}^2}(\mathbf{o},r)}\cdots\int_{B_{\mathbb{S}^2}(\mathbf{o},r)}\rho^{(n+1)}(0,\mathbf{x}_1,\ldots,\mathbf{x}_n)\lambda_{\mathbb{S}^2}(d\mathbf{x}_1)\cdots\lambda_{\mathbb{S}^2}(d\mathbf{x}_n),\end{aligned}$$

then using the Definition 3.5.3 of the n^{th}-order correlation function,

$$\begin{aligned}\rho(1-H(r)) =&\rho+\rho\sum_{n=1}^{\infty}\frac{(-\rho)^n}{n!}\int_{B_{\mathbb{S}^2}(\mathbf{o},r)}\cdots\int_{B_{\mathbb{S}^2}(\mathbf{o},r)}\\ &\sum_{k=0}^{n}\sum_{D_1,\ldots,D_k}\xi_{|D_1|}(\mathbf{x}_{D_1})\cdots\xi_{|D_k|}(\mathbf{x}_{D_k})\lambda_{\mathbb{S}^2}(d\mathbf{x}_1)\cdots\lambda_{\mathbb{S}^2}(d\mathbf{x}_n),\end{aligned}$$

where the set $\{D_1,\ldots,D_k\}$ now partitions $\{0,1,\ldots,n\}$. Then,

$$\begin{aligned}\rho(1-H(r)) = \rho+\rho\sum_{n=1}^{\infty}\frac{(-\rho)^n}{n!}\sum_{k=0}^{n}\sum_{D_1,\ldots,D_k}&\\ \int_{B_{\mathbb{S}^2}(\mathbf{o},r)^{|D_1|}}\xi_{|D_1|}(\mathbf{x}_{D_1})\lambda^{(|D_1|)}_{\mathbb{S}^2}(d\mathbf{x}_{D_1})\cdots\int_{B_{\mathbb{S}^2}(\mathbf{o},r)^{|D_k|}}&\xi_{|D_k|}(\mathbf{x}_{D_k})\lambda^{(|D_k|)}_{\mathbb{S}^2}(d\mathbf{x}_{D_k})\end{aligned}$$

where $B_{\mathbb{S}^2}(\mathbf{o},r)^{|D_i|}$ is the Cartesian product of $B_{\mathbb{S}^2}(\mathbf{o},r)$, $|D_i|$ times, $d\mathbf{x}_{D_i} = d\mathbf{x}_{a_1} \times \cdots \times d\mathbf{x}_{a_{|D_i|}}$ such that $a_1, \ldots, a_{|D_i|} \in \mathbf{x}_{D_i}$ and $\lambda_{\mathbb{S}^2}^{(n)}$ is the n - fold Lebesgue measure on $\mathbb{S}^2$. Then we pull out the terms $\xi_{|D_i|}(\mathbf{x}_{D_i})$ such that $0 \in D_i$ giving,

$$\rho(1-H(r)) = \rho + \rho\sum_{n=1}^{\infty}\frac{(-\rho)^n}{n!}\sum_{D\subseteq\{1,\ldots,n\}} J_{|D|}(r)\sum_{k=1}^{n-|D|}\sum_{\substack{D_1,\ldots,D_k\neq\emptyset\\ \cup D_j=\{1,\ldots,n\}\setminus D}} I_{|D_1|}\cdots I_{|D_k|}, \quad \text{(A.4)}$$

where $I_{|D_i|} = \int_{B_{\mathbb{S}^2}(\mathbf{o},r)^{|D_i|}} \xi_{|D_i|}(\mathbf{x}_{|D_i|})\lambda_{\mathbb{S}^2}^{(|D_i|)}(d\mathbf{x}_{|D_i|})$ and we have taken the convention that $\sum_{k=1}^{0} = 1$. It can be shown that this can then be factorised as,

$$\left(\rho J_0(r) + \rho\sum_{n=1}^{\infty}\frac{(-\rho)^n}{n!}J_n(t)\right)\left(1+\sum_{m=1}^{\infty}\frac{(-\rho)^m}{m!}\sum_{k=1}^{m}\sum_{\substack{D_1,\ldots,D_k\\ \cup D_j=\{1,\ldots,m\}}} I_{|D_1|}\cdots I_{|D_k|}\right). \quad \text{(A.5)}$$

To see why this factorisation holds we expand Equation A.5 and then work term by term. By expanding A.5, we have the following terms,

$$\rho J_0(r) \quad \text{(A.6)}$$

$$\rho\sum_{n=1}^{\infty}\frac{(-\rho)^n}{n!}J_n(t) \quad \text{(A.7)}$$

$$\rho J_0(r)\left(\sum_{m=1}^{\infty}\frac{(-\rho)^m}{m!}\sum_{k=1}^{m}\sum_{\substack{D_1,\ldots,D_k\\ \cup D_j=\{1,\ldots,m\}}} I_{|D_1|}\cdots I_{|D_k|}\right) \quad \text{(A.8)}$$

$$\left(\rho\sum_{n=1}^{\infty}\frac{(-\rho)^n}{n!}J_n(t)\right)\left(\sum_{m=1}^{\infty}\frac{(-\rho)^m}{m!}\sum_{k=1}^{m}\sum_{\substack{D_1,\ldots,D_k\\ \cup D_j=\{1,\ldots,m\}}} I_{|D_1|}\cdots I_{|D_k|}\right) \quad \text{(A.9)}$$

For term A.6, this is identical to ρ by the definition of $\xi_1 \equiv 1$ and so $J_0(r) = 1$, which gives the first term of A.4. In A.4 consider the second term but in the second summand only consider the case when $D = \{1,\ldots,n\}$. Thus $\sum_{k=1}^{n-|D|} = \sum_{k=1}^{0} = 1$ and therefore we have A.7. Next consider when $D = \emptyset$, then pulling $J_0(r)$ out we have A.8. Finally we consider all

other terms over the sum of D,

$$\rho\sum_{n=1}^{\infty}\frac{(-\rho)^n}{n!}\sum_{\substack{D\subseteq\{1,\ldots,n\}\\ D\neq\emptyset, D\neq\{1,\ldots,n\}}} J_{|D|}(r)\sum_{k=1}^{n-|D|}\sum_{\substack{D_1,\ldots,D_k\neq\emptyset\\ \cup D_j=\{1,\ldots,n\}\setminus D}} I_{|D_1|}\cdots I_{|D_k|}. \tag{A.10}$$

Notice that by integrating out $\mathbf{x}_i$ in $J_{|D|}(r), I_{|D_i|},\ i=1,2,\ldots$ the specific elements of the sets D and $D_j,\ j=1,2,\ldots$ are not relevant and only the size of the sets are. Acknowledging this fact and that there are $\binom{n}{x}$ ways of selecting a subset of $\{1,\ldots,n\}$ of size d, A.10 can be rewritten as,

$$\rho\sum_{n=1}^{\infty}\frac{(-\rho)^n}{n!}\sum_{d=1}^{n-1}\binom{n}{d}J_d(r)\sum_{k=1}^{n-d}\sum_{\substack{D_1,\ldots,D_k\neq\emptyset\\ \cup D_j=\{1,\ldots,n-d\}}} I_{|D_1|}\cdots I_{|D_k|}$$

By expanding $\binom{n}{d}$ and switching the order of the first two sums we have,

$$\left(\rho\sum_{n=1}^{\infty}\frac{(-\rho)^n}{n!}J_n(t)\right)\left(\sum_{m=1}^{\infty}\frac{(-\rho)^m}{m!}\sum_{k=1}^{m}\sum_{\substack{D_1,\ldots,D_k\\ \cup D_j=\{1,\ldots,m\}}} I_{|D_1|}\cdots I_{|D_k|}\right),$$

which is identically A.9. By the definition of the n^{th}-order correlation function and Corollary 3.5.2, the second term of A.5 is $1-F(r)$ and so we have,

$$\rho(1-H(r))=\left(\rho+\rho\sum_{n=1}^{\infty}\frac{(-\rho)^n}{n!}J_n(t)\right)(1-F(r))$$

Since $J(r)=(1-H(r))/(1-F(r))$, we have the final result. □

A.4 Proof of Theorem 3.5.7

Before proceeding with the proof of Theorem 3.5.7 we first discuss the generating functional of a point process. The generating functional, $G_X(u)$ of a point process X, very much like a generating function for a random variable, can be used to derive the n^{th} order factorial moment measures based on derivatives of $G_X(u)$ (e.g. see Chiu et al. ((1995))). Conversely, assuming that the product intensity functions of all orders exists and letting u be a measurable function taking values in $[0,1]$ with bounded support then as shown by e.g. van

Lieshout ((2011)) we have an infinite series representation of the generating functional,

$$
\begin{aligned}
G_X(1-u) &= \mathbb{E}\prod_{\mathbf{x}\in X}(1-u(\mathbf{x})) \\
&= \mathbb{E}\left[1-\sum_{\mathbf{x}\in X}u(\mathbf{x})+\frac{1}{2!}\sum_{\mathbf{x}_1,\mathbf{x_2}\in X}^{\neq}u(\mathbf{x}_1)u(\mathbf{x}_2)-\cdots\right] \\
&= \mathbb{E}\left[1+\sum_{n=1}^{\infty}\frac{(-1)^n}{n!}\sum_{\mathbf{x}_1,\cdots,\mathbf{x}_n}^{\neq}u(\mathbf{x}_1)\cdots u(\mathbf{x}_n)\right] \\
&= 1+\sum_{n=1}^{\infty}\frac{(-1)^n}{n!}\mathbb{E}\sum_{\mathbf{x}_1,\cdots,\mathbf{x}_n}^{\neq}u(\mathbf{x}_1)\cdots u(\mathbf{x}_n) \\
&= 1+\sum_{n=1}^{\infty}\frac{(-1)^n}{n!} \\
&\quad \int_{\mathbb{S}^2}\cdots\int_{\mathbb{S}^2}u(\mathbf{x}_1)\cdots u(\mathbf{x}_n)\rho^{(n)}(\mathbf{x}_1,\ldots,\mathbf{x}_n)\lambda_{\mathbb{S}^2}(d\mathbf{x}_1)\cdots\lambda_{\mathbb{S}^2}(d\mathbf{x}_n), \qquad \text{(A.11)}
\end{aligned}
$$

assuming the right hand side is absolutely convergent. In the event the n^{th} order product intensities do not exists but the n^{th} order factorial moment measures do and are locally finite then can be replaced with $\alpha^{(n)}(d\mathbf{x}_1,\ldots,d\mathbf{x}_n)$. Equipped with this identity we can now progress with the proof of Theorem 3.5.7 which follows analogously to the proof of Theorem 1 in van Lieshout ((2011)).

Proof. To begin the proof we first show the following identity,

$$
\begin{aligned}
&\mathbb{E}\sum_{\mathbf{x}_1,\ldots,\mathbf{x}_n\in X^!_{\mathbf{x}}}^{\neq}\prod_{i=1}^{n}\frac{\mathbb{1}[\mathbf{x}_i\in B_{\mathbb{S}^2}(\mathbf{x},r)]}{\rho(\mathbf{x}_i)} \\
&\qquad = \int_{B_{\mathbb{S}^2}(\mathbf{o},r)}\cdots\int_{B_{\mathbb{S}^2}(\mathbf{o},r)}\frac{\rho^{(n+1)}(\mathbf{0},\mathbf{x}_1,\ldots,\mathbf{x}_n)}{\rho(\mathbf{0})\rho(\mathbf{x}_1)\cdots\rho(\mathbf{x}_n)}\lambda_{\mathbb{S}^2}(d\mathbf{x}_1)\cdots\lambda_{\mathbb{S}^2}(d\mathbf{x}_n).
\end{aligned}
$$

Starting with the left-hand side,

$$
\begin{aligned}
&\mathbb{E}\sum_{\mathbf{x}_1,\ldots,\mathbf{x}_n\in X^!_{\mathbf{x}}}^{\neq}\prod_{i=1}^{n}\frac{\mathbb{1}[\mathbf{x}_i\in B_{\mathbb{S}^2}(\mathbf{x},r)]}{\rho(\mathbf{x}_i)} \\
&\qquad = \int_{B_{\mathbb{S}^2}(\mathbf{x},r)}\cdots\int_{B_{\mathbb{S}^2}(\mathbf{x},r)}\frac{\rho^{!(n)}(\mathbf{x}_1,\ldots,\mathbf{x}_n)}{\rho(\mathbf{x}_1)\cdots\rho(\mathbf{x}_n)}\lambda_{\mathbb{S}^2}(d\mathbf{x}_1)\cdots\lambda_{\mathbb{S}^2}(d\mathbf{x}_n)
\end{aligned}
$$

$$= \int_{B_{\mathbb{S}^2}(\mathbf{x},r)} \cdots \int_{B_{\mathbb{S}^2}(\mathbf{x},r)} \frac{\rho^{(n+1)}(\mathbf{x},\mathbf{x}_1,\ldots,\mathbf{x}_n)}{\rho(\mathbf{x})\rho(\mathbf{x}_1)\cdots\rho(\mathbf{x}_n)} \lambda_{\mathbb{S}^2}(d\mathbf{x}_1)\cdots\lambda_{\mathbb{S}^2}(d\mathbf{x}_n)$$
$$= \int_{B_{\mathbb{S}^2}(\mathbf{o},r)} \cdots \int_{B_{\mathbb{S}^2}(\mathbf{o},r)} \frac{\rho^{(n+1)}(\mathbf{0},\mathbf{x}_1,\ldots,\mathbf{x}_n)}{\rho(\mathbf{0})\rho(\mathbf{x}_1)\cdots\rho(\mathbf{x}_n)} \lambda_{\mathbb{S}^2}(d\mathbf{x}_1)\cdots\lambda_{\mathbb{S}^2}(d\mathbf{x}_n)$$

where the first line holds by the Campbell-Mecke Theorem applied to the reduced Palm process and the penultimate line holds by Equation A.3, and the final one holds by rotational invariance of the n^{th} order correlation functions, and this holds for almost all $\mathbf{x} \in \mathbb{S}^2$. In the original proof given by van Lieshout ((2011)) they use Fubini's theorem to equate integrands to show the above identity. We have the following identity,

$$\prod_{\mathbf{x}\in X} \left(1 - \frac{\bar{\rho}\mathbb{1}[\mathbf{x} \in B_{\mathbb{S}^2}(\mathbf{y},r)]}{\rho(\mathbf{x})}\right) = 1 + \sum_{n=1}^{\infty} \frac{(-\bar{\rho})^n}{n!} \sum_{\mathbf{x}_1,\ldots,\mathbf{x}_n \in X}^{\neq} \prod_{i=1}^{n} \frac{\mathbb{1}[\mathbf{x}_i \in B_{\mathbb{S}^2}(\mathbf{y},r)]}{\rho(\mathbf{x}_i)}. \tag{A.12}$$

This identity holds since the product on the left-hand side is almost surely a finite product since $X \cap B_{\mathbb{S}^2}(\mathbf{x},r)$ is almost surely finite. To see why this identity follows, an identical argument can be used as in the series representation of the generating functional given previously. Not only does the identity hold for a point process X but it also holds for its reduced Palm process, $X^!_{\mathbf{x}}$. Thus taking expectations of both sides the following identity holds for almost all $\mathbf{y} \in \mathbb{S}^2$,

$$\begin{aligned} G_{X^!_{\mathbf{y}}}(1-u_r^{\mathbf{y}}) =& 1 + \sum_{n=1}^{\infty} \frac{(-\bar{\rho})^n}{n!} \\ & \int_{B_{\mathbb{S}^2}(\mathbf{o},r)} \cdots \int_{B_{\mathbb{S}^2}(\mathbf{o},r)} \frac{\rho^{(n+1)}(\mathbf{0},\mathbf{x}_1,\ldots,\mathbf{x}_n)}{\rho(\mathbf{0})\rho(\mathbf{x}_1)\cdots\rho(\mathbf{x}_n)} \lambda_{\mathbb{S}^2}(d\mathbf{x}_1)\cdots\lambda_{\mathbb{S}^2}(d\mathbf{x}_n), \end{aligned} \tag{A.13}$$

which holds provided that the power series on the right-hand side is absolutely convergent. Further for all $\mathbf{y} \in \mathbb{S}^2$, by the series representation given by Equation A.11, and rotational invariance of the n^{th} order correlation function we also have,

$$\begin{aligned} G_X(1-u_r^{\mathbf{y}}) =& 1 + \sum_{n=1}^{\infty} \frac{(-\bar{\rho})^n}{n!} \\ & \int_{B_{\mathbb{S}^2}(\mathbf{o},r)} \cdots \int_{B_{\mathbb{S}^2}(\mathbf{o},r)} \frac{\rho^{(n)}(\mathbf{x}_1,\ldots,\mathbf{x}_n)}{\rho(\mathbf{x}_1)\cdots\rho(\mathbf{x}_n)} \lambda_{\mathbb{S}^2}(d\mathbf{x}_1)\cdots\lambda_{\mathbb{S}^2}(d\mathbf{x}_n), \end{aligned} \tag{A.14}$$

where the right-hand side is assumed absolutely convergent. By taking Equation A.13, it

can be shown that this equates to,

$$G_{X^!_{\mathbf{y}}}(1-u_r^{\mathbf{y}}) = 1+\sum_{n=1}^{\infty}\frac{(-\bar{\rho})^n}{n!}\sum_{D\subseteq\{1,\ldots,n\}} J_{|D|}(r)\sum_{k=1}^{n-|D|}\sum_{\substack{D_1,\ldots,D_k\neq\emptyset\\ \cup D_j=\{1,\ldots,m\}}} I_{|D_1|}\cdots I_{|D_k|}, \tag{A.15}$$

where $I_n = \int_{B_{\mathbb{S}^2}(\mathbf{o},r)}\cdots\int_{B_{\mathbb{S}^2}(\mathbf{o},r)}\xi_n(\mathbf{x}_1,\ldots,\mathbf{x}_n)\lambda_{\mathbb{S}^2}(d\mathbf{x}_1)\cdots\lambda_{\mathbb{S}^2}(d\mathbf{x}_n)$. This can then be factored as,

$$\left(1+\sum_{n=1}^{\infty}\frac{(-\bar{\rho})^n}{n!}J_n(r)\right)\left(1+\sum_{m=1}^{\infty}\frac{(-\bar{\rho})^m}{m!}\sum_{k=1}^{m}\sum_{\substack{D_1,\ldots,D_k\neq\emptyset\\ \cup D_j=\{1,\ldots,m\}}} I_{|D_1|}\cdots I_{|D_k|}\right). \tag{A.16}$$

Equation A.15 and A.16 follow by identical arguments used in the derivation of the series representation of the isotropic J-function, in particular Equations A.4 and A.5. Noting the the right-hand side of the product in Equation A.16 is equivalently $G(1-u_r^{\mathbf{y}})$, the representation of J_{inhom} follows by Definition 3.5.6. □

A.5 Proof of Proposition 3.5.8

We demonstrate the unbiasedness of $\hat{\mathrm{F}}_{\text{inhom}}$ and ratio - unbiasedness of $\hat{\mathrm{H}}_{\text{inhom}}$ and $\hat{\mathrm{J}}_{\text{inhom}}$ for a point process that has been observed over all of $\mathbb{S}^2$. The proof for windowed observations follows identically.

Proof. Consider taking expectations of $1-\hat{F}_{\text{inhom}}(r) = \frac{\sum_{\mathbf{p}\in P}\prod_{\mathbf{x}\in X\cap B(\mathbf{p},r)}\left(1-\frac{\bar{\rho}}{\rho(\mathbf{x})}\right)}{|P|}$ then,

$$\begin{aligned}
1-\mathbb{E}[\hat{F}_{\text{inhom}}(r)] &= \mathbb{E}\frac{1}{|P|}\sum_{\mathbf{p}\in P}\prod_{\mathbf{x}\in X\cap B(\mathbf{p},r)}\left(1-\frac{\bar{\rho}}{\rho(\mathbf{x})}\right)\\
&= \frac{1}{|P|}\sum_{\mathbf{p}\in P}\mathbb{E}\prod_{\mathbf{x}\in X\cap B(\mathbf{p},r)}\left(1-\frac{\bar{\rho}}{\rho(\mathbf{x})}\right)\\
&= \frac{1}{|P|}\sum_{\mathbf{p}\in P}\mathbb{E}\prod_{\mathbf{x}\in X\cap B(\mathbf{y},r)}\left(1-\frac{\bar{\rho}}{\rho(\mathbf{x})}\right)\\
&= \mathbb{E}\prod_{\mathbf{x}\in X\cap B(\mathbf{y},r)}\left(1-\frac{\bar{\rho}}{\rho(\mathbf{x})}\right)\\
&= G(1-u_r^{\mathbf{y}}),
\end{aligned}$$

where the third line follows from the fact that the point process is assumed IRWMI and hence unbiasedness of $\hat{F}_{\text{inhom}}(r)$ follows.

To show ratio-unbiasedness of $\hat{H}_{\text{inhom}}(r)$ define the random variables

$$Y = \sum_{\mathbf{x}\in X} \prod_{\mathbf{y}\in(X\setminus\{\mathbf{x}\})\cap B(\mathbf{x},r)} \left(1 - \frac{\bar{\rho}}{\rho(\mathbf{y})}\right),$$

and $Z = N_X(\mathbb{S}^2)$, then taking expectations of Y, leads to,

$$\begin{aligned}
\mathbb{E}[Y] &= \mathbb{E} \sum_{\mathbf{x}\in X} \prod_{\mathbf{y}\in X\setminus\{\mathbf{x}\}\cap B(\mathbf{x},r)} \left(1 - \frac{\bar{\rho}}{\rho(\mathbf{y})}\right) \\
&= \int_{\mathbb{S}^2} \mathbb{E} \prod_{\mathbf{y}\in X^!_{\mathbf{x}}} \left(1 - \frac{\bar{\rho}}{\rho(\mathbf{y})}\right) \alpha(d\mathbf{x}) \\
&= \int_{\mathbb{S}^2} G^!_{\mathbf{x}}(1 - u_r^{\mathbf{x}})\alpha(d\mathbf{x}) \\
&= \int_{\mathbb{S}^2} G^!_{\mathbf{y}}(1 - u_r^{\mathbf{y}})\alpha(d\mathbf{x}) \\
&= G^!_{\mathbf{y}}(1 - u_r^{\mathbf{y}})\mathbb{E}[N_X(\mathbb{S}^2)],
\end{aligned}$$

where the second line follows by the Campbell-Mecke Theorem, the penultimate line follows due to the assumption of IRWMI and the final line follows from $\int_{\mathbb{S}^2} \alpha(d\mathbf{x}) = \alpha(\mathbb{S}^2) = \mathbb{E}[N_X(\mathbb{S}^2)]$. Clearly, $\mathbb{E}[Z] = \mathbb{E}[N_X(\mathbb{S}^2)]$ and so ratio-unbiasedness of $\hat{H}_{\text{inhom}}$ follows. □

A.6 Proof of Proposition 3.6.2

Proof. Let us define the zero mean Gaussian random field $Z(\mathbf{x}) = U(\mathbf{x}) - \mu(\mathbf{x})$ with the same correlation function as U. Then $J_{\text{inhom}}(r) \leq 1$ if

$$\text{Cov}\left(\mathrm{e}^{Z(\mathbf{o})}, \exp\left(-\bar{\mu}\int_{B_{\mathbb{S}^2}(\mathbf{o},r)} \mathrm{e}^{Z(\mathbf{x})}\lambda_{\mathbb{S}^2}(d\,\mathbf{x})\right)\right) \leq 0. \tag{A.17}$$

Consider a Riemann partition of $B_{\mathbb{S}^2}(\mathbf{o}, r)$, that is define the sequence $B(n)$ for $n > 1$ such that $B(n) = \cup_{i=1}^n B_{i,n}$ where $\cup_{i=1}^n B_{i,n} = B_{\mathbb{S}^2}(\mathbf{o}, r)$, $B_{i,n} \cap B_{j,n} = \emptyset$ for $i \neq j$ and $\lambda_{\mathbb{S}^2}(B_{i,n}) \to 0$ as $n \to \infty$. Then for each $B_{i,n}$ take any $\mathbf{x}_{i,n} \in B_{i,n}$ and by almost sure path continuity of Z

(see Proposition 3.6.1) we have that

$$\exp\left(-\bar{\mu}\int_{B_{\mathbb{S}^2}(\mathbf{o},r)} \mathrm{e}^{Z(\mathbf{x})}\lambda_{\mathbb{S}^2}(d\,\mathbf{x})\right) = \lim_{n\to\infty}\exp\left(-\bar{\mu}\sum_{i=1}^{n}\mathrm{e}^{Z(\mathbf{x}_{i,n})}\lambda_{\mathbb{S}^2}(B_{i,n})\right).$$

For each n define $\mathbf{Z} = (Z(\mathbf{o}), Z(\mathbf{x}_{1,n}), \ldots, Z(\mathbf{x}_{n,n}))^T$, then since the correlation function is nonnegative by construction we have that the elements of $\mathbf{Z}$ are *associated* by Pitt's Theorem ((Pitt, 1982)) . That is random variables $Z_1, \ldots, Z_n$ are said to be associated if

$$\mathrm{Cov}(f(\mathbf{Z}), g(\mathbf{Z})) \geq 0,$$

for any pair of *increasing* functions f and g; a function $f : \mathbb{R}^n \mapsto \mathbb{R}$ is said to be increasing if it is a non - decreasing function of any single component. Let us define $f(\mathbf{Z}) = \mathrm{e}^{Z_1}$ and $g(\mathbf{Z}) = -\exp\left(-\bar{\mu}\sum_{i=1}^{n}\mathrm{e}^{Z_i}\lambda_{\mathbb{S}^2}(B_{i,n})\right)$, then since both f and g are increasing we have that for any $n > 0$,

$$\begin{aligned}\mathrm{Cov}\left(\mathrm{e}^{Z(\mathbf{o})}, \exp\left(-\bar{\mu}\sum_{i=1}^{n}\mathrm{e}^{Z(\mathbf{x}_{i,n})}\lambda_{\mathbb{S}^2}(B_{i,n})\right)\right) &= \mathrm{Cov}(f(\mathbf{Z}), -g(\mathbf{Z})) \\ &= -\mathrm{Cov}(f(\mathbf{Z}), g(\mathbf{Z})) \\ &\leq 0.\end{aligned}$$

Thus by taking the limit as $n \to \infty$ and applying the dominated convergence theorem we obtain A.17. □

B

Appendix to Chapter 4

B.1 Proof of Lemma 4.3.1

Proof. To show that f is bijective we need to show that it is both injective and surjective. For surjectivity we need to show that for any $\mathbf{x}' \in \mathbb{S}^2$, there exists $\mathbf{x} \in \mathbb{D}$ such that $f(\mathbf{x}) = \mathbf{x}'$. To do this first fix any $\mathbf{x}' \in \mathbb{S}^2$ and define the half line, $L_{\mathbf{x}'} = \{\mathbf{y} \in \mathbb{R}^3 : \mathbf{y} = \lambda \mathbf{x}',\ \lambda \in \mathbb{R}^+\}$, where $\mathbb{R}^+$ is the positive real line including 0. Then since $\mathbb{D}$ is compact (i.e. closed and bounded) and $\mathbf{o}$ is in the interior of $\mathbb{D}$ then the half line must intersect the $\mathbb{D}$ and so there exists $\mathbf{x} \in \mathbb{D}$ such that $\mathbf{x} \in L_{\mathbf{x}'}$. Therefore there must exists $\lambda \in \mathbb{R}^+, \mathbf{x} = \lambda \mathbf{x}'$. Taking norms of both sides and noting that since $\mathbf{x}' \in \mathbb{S}^2$ meaning $||\mathbf{x}'|| = 1$ then $\lambda = ||\mathbf{x}||$ and so, $\mathbf{x}/||\mathbf{x}|| = \mathbf{x}'$ and so f is surjective.

For injectivity we need to show that for any $\mathbf{x}, \mathbf{y} \in \mathbb{D}$ if $f(\mathbf{x}) = f(\mathbf{y})$ then $\mathbf{x} = \mathbf{y}$. Fix $\mathbf{x}, \mathbf{y} \in \mathbb{D}$ such that $f(\mathbf{x}) = f(\mathbf{y})$ and define $\mathbf{x}' = f(\mathbf{x}) = f(\mathbf{y})$. Again define the line $L_{\mathbf{x}'}$ as previous and by convexity of $\mathbb{D}$, the fact that $\mathbf{o} \in \mathbb{D}$ and since $L_{\mathbf{x}'}$ is a half line then there is precisely only one intersection of $L_{\mathbf{x}'}$ with $\mathbb{D}$. Therefore both $\mathbf{x}$ and $\mathbf{y}$ must be this point of intersection meaning $\mathbf{x} = \mathbf{y}$. Hence f is bijective. □

B.2 Proof of Theorem 4.3.2

In this section we show that a Poisson process lying on an arbitrary bounded convex space $\mathbb{D} \subset \mathbb{R}^3$ can be mapped to another Poisson process on a sphere, known as the Mapping

Theorem Kingman ((1993)). We also show that no two different Poisson processes on $\mathbb{D}$ map to the same Poisson process on the sphere under the same mapping.

Before starting the main theorem of this section we first introduce a lemma (see for example Møller and Waagepetersen ((2003, Proposition 3.1, pp 15-16))) which is an expansion of the probability measure for a Poisson process on an arbitrary metric space. We shall state it in the context of an arbitrary convex shape in $\mathbb{R}^3$, represented by $\mathbb{D}$.

Lemma S1. *(Møller and Waagepetersen ((2003))) X is a Poisson process with intensity function $\rho : \mathbb{D} \mapsto \mathbb{R}$ on $\mathbb{D}$ if and only if for all $B \subseteq \mathbb{D}$ with $\mu(B) = \int_B \rho(\mathbf{x})d\mathbf{x} < \infty$ and all $F \subseteq N_{lf}$,*

$$P(X_B \in F) = \sum_{n=0}^{\infty} \frac{\exp(-\mu(B))}{n!} \int_B \cdots \int_B \mathbb{1}[\{\mathbf{x}_1, \ldots, \mathbf{x}_n\} \in F] \prod_{i=1}^{n} \rho(\mathbf{x}_i)\lambda_{\mathbb{D}}(d\mathbf{x}_1) \cdots \lambda_{\mathbb{D}}(d\mathbf{x}_n), \tag{B.1}$$

where the integral for $n = 0$ is read as $\mathbb{1}[\emptyset \in F]$.

Proof. See Proposition 3.1 of Møller and Waagepetersen ((2003)). □

Now we give the main theorem of our work which shows that a Poisson process on $\mathbb{D}$ is mapped to a Poisson process on $\mathbb{S}^2$. This is known as the Mapping Theorem Kingman ((1993)). Here we use Lemma S1.

Before beginning this theorem we lay down a little notation in order to avoid confusion. $\mathbf{x} = (x, y, z)$ will be an element of $\mathbb{R}^3$ where we may subscript with an $n \in \mathbb{N}$ when we are referring to a single vector within a set. x will be an element of N_{lf} and may also be subscripted with $n \in \mathbb{N}$ when we are referring to a single element in a set of finite point configurations. Notice that we are using x to be both the first element of $\mathbf{x}$ and an element in N_{lf}, based on context it will be clear to which we are referring too.

Proof. In order to show that $Y \equiv f(X)$ is a Poisson process we show that its distribution function can be expanded as given by Equation B.1. Then $\forall B \subseteq \mathbb{D}$ and $\forall F \subseteq N_{lf}$ and noting that f is a measurable map (since the map is bijective and hence an inverse exists) we have that,

$$\begin{aligned} P(Y_B \in F) &= P(f^{-1}(Y_B) \in f^{-1}(F)) \\ &= P(X_{f^{-1}(B)} \in f^{-1}(F)) \end{aligned}$$

Now we define $f_i^{-1}(F)$, for $i = 1, \ldots, n$ where $F \subseteq N_{lf}$,

$$\begin{aligned} F = &\quad \{x \in N_{lf} : x = \{\mathbf{x}_1, \ldots, \mathbf{x}_m\}, m \in \mathbb{N}, x \in F\} \\ f^{-1}(F) = &\quad \{f^{-1}(x) : f^{-1}(x) = \{f^{-1}(\mathbf{x}_1), \ldots, f^{-1}(\mathbf{x}_m)\}, m \in \mathbb{N}, x \in F\} \end{aligned}$$

we want to partition $f^{-1}(F)$ over each $f^{-1}(\mathbb{D}_i), i = 1, \ldots, n$

$$\begin{aligned} f^{-1}(F) = &\quad \left\{\cup_{i=1}^n f^{-1}(x_i) : f^{-1}(x_i) = \{f^{-1}(\mathbf{x}_{(i,1)}), \ldots, f^{-1}(\mathbf{x}_{(i,m_i)})\},\right. \\ &\quad \left. f^{-1}(\mathbf{x}_{(i,j)}) \in f^{-1}(\mathbb{D}_i), j = 1, \ldots, m_i,\ m_i \in \mathbb{N},\ i = 1, \ldots, n, \cup_{i=1}^n x_i \in F\right\}. \end{aligned}$$

To understand the notation $\mathbf{x}_{(i,j)}$ consider first a single element $x \in F$. Then since F is a subset of N_{lf} this means that $|x| \in \mathbb{N}$. Then define $m_i = |x \cap \mathbb{D}_i|$, hence $\sum_{i=1}^n m_i = |x|$. Then $\mathbf{x}_{(i,j)}$ is j^{th} element of $x \cap \mathbb{D}_i$ such that $j = 1, \ldots, m_i$. We define $f_i^{-1}(F) \equiv \{f^{-1}(x) : f^{-1}(x) \equiv \{f^{-1}(\mathbf{x}_1), \ldots, f^{-1}(\mathbf{x}_n)\},\ f^{-1}(\mathbf{x}_i) \in f^{-1}(\mathbb{D}_i), n \in \mathbb{N}, \exists y \in F \text{ such that } x \subseteq y\}$. Then,

$$\begin{aligned} P(X_{f^{-1}(B)} \in f^{-1}(F)) &= P(\{X_{f^{-1}(B)\cap\mathbb{D}_1}, \ldots, X_{f^{-1}(B)\cap\mathbb{D}_n}\} \in f^{-1}(F)) \\ &= P(X_{f^{-1}(B)\cap\mathbb{D}_1} \in f_1^{-1}(F), \ldots, X_{f^{-1}(B)\cap\mathbb{D}_n} \in f_n^{-1}(F)) \\ &= P(X_{f_1^{-1}(B)} \in f_1^{-1}(F), \ldots, X_{f_n^{-1}(B)} \in f_n^{-1}(F)) \\ &= \prod_{i=1}^n P(X_{f_i^{-1}(B)} \in f_i^{-1}(F)) \end{aligned}$$

where $f_i^{-1}(A) = \{(x, y, z)^T \in A : (x, y, z)^T \in \mathbb{D}_i\}$ if $A \subseteq \mathbb{D}$. We emphasize the dual meaning of $f_i^{-1}(A)$ where the defintion depends on the nature of A, i.e. if $A \subseteq \mathbb{D}$ or if $A \subseteq N_{lf}$. Without loss of generality lets suppose that all projections of each $\mathbb{D}_i$ onto $\mathbb{R}^2$ are invertible, if not then we can divide $\mathbb{D}_i$ into further subsets $\cup_{j=1}^m \mathbb{D}_{i,j}$ such that the projection of each $\mathbb{D}_{i,j}$ is then invertible. For example an ellipsoid is defined by the zero-set equation $x^2/a^2 + y^2/b^2 + z^2/c^2 = 1$, but if the entire space were projected down to $\mathbb{R}^2$ its inverse does not exists, instead we divide the ellipsoid into the upper and lower hemiellipsoids and then the projections restricted to these segments of the ellipsoids are then invertible. Let us also define the projection of $\mathbb{D}_i$ to $\mathbb{R}^2$ as $P_{\mathbb{D}_i}$. Further, for $\mathbf{x} = (x, y, z)^T$, then $\mathbf{x}$ lies on $\mathbb{D}$ if $g(\mathbf{x}) = 0$, we define $\tilde{g}$ to be the rearrangement of g such that z is a function of x and y, i.e. $z = \tilde{g}(x, y)$. Then by Lemma S1,

$$P(X_{f_i^{-1}(B)} \in f_i^{-1}(F))$$

$$= \sum_{n=0}^{\infty} \frac{\exp(-\mu(f_i^{-1}(B)))}{n!}$$
$$\int_{f_i^{-1}(B)} \cdots \int_{f_i^{-1}(B)} \mathbb{1}[\{\mathbf{x}_1, \ldots, \mathbf{x}_n\} \in f_i^{-1}(F)] \prod_{i=1}^{n} \rho(\mathbf{x}_i) \lambda_{\mathbb{D}_i}(d\mathbf{x}_1) \cdots \lambda_{\mathbb{D}_i}(d\mathbf{x}_n)$$
$$= \sum_{n=0}^{\infty} \frac{\exp(-\mu(f_i^{-1}(B))}{n!}$$
$$\iint_{P_{\mathbb{D}_i}[f_i^{-1}(B)]} \cdots \iint_{P_{\mathbb{D}_i}[f_i^{-1}(B)]} \mathbb{1}[\{(x_1, y_1, \tilde{g}_1(x_1, y_1))^T, \ldots, (x_n, y_n, \tilde{g}_1(x_n, y_n))^T\} \in f_i^{-1}(F)]$$
$$\prod_{i=1}^{n} \rho\left(x_i, y_i, \tilde{g}_i(x_i, y_i)\right) \sqrt{1 + \left(\frac{\partial \tilde{g}_i}{\partial x_i}\right)^2 + \left(\frac{\partial \tilde{g}_i}{\partial y_i}\right)^2} dx_i dy_i$$
$$= \sum_{n=0}^{\infty} \frac{\exp(-\mu(f_i^{-1}(B))}{n!}$$
$$\iint_{P_{\mathbb{D}_i}[f_i^{-1}(B)]} \cdots \iint_{P_{\mathbb{D}_i}[f_i^{-1}(B)]} \mathbb{1}[\{(x_1, y_1, \tilde{g}_1(x_1, y_1))^T, \ldots, (x_n, y_n, \tilde{g}_1(x_n, y_n))^T\} \in f_i^{-1}(F)]$$
$$\prod_{i=1}^{n} \rho\left(x_i, y_i, \tilde{g}_i(x_i, y_i)\right) l_i(x_i, y_i) dx_i dy_i,$$

where $l_i(x_i, y_i) = \sqrt{1 + (\partial \tilde{g}_i/\partial x_i)^2 + (\partial \tilde{g}_i/\partial y_i)^2}$. Now consider the indicator term, we need to show that when $x_i \mapsto x_i/||\mathbf{x}||$ and $y_i \mapsto y_i/||\mathbf{x}||$ then $\mathbb{1}[\{\mathbf{x}_1, \ldots, \mathbf{x}_n\} \in f_i^{-1}(F)] \mapsto \mathbb{1}[\{\mathbf{y}_1, \ldots, \mathbf{y}_n\} \in F_i]$, where $F_i = \{x : x = \{\mathbf{x}_1, \ldots, \mathbf{x}_m\}, \mathbf{x}_j \in f(\mathbb{D}_i), j = 1, \ldots, m, m \in \mathbb{N}, \exists y \in F \text{ such that } x \subseteq y\}$. Let us consider first an individual point $\mathbf{x} \in \mathbb{D}$. Further let us define $r = |\mathbf{x}| = \sqrt{x^2 + y^2 + z^2}$. Then since $z = \tilde{g}_i(x, y)$, r is thus a function of x and y, let us write $r(x, y) = ||\mathbf{x}||$. Thus we can rewrite $z = \tilde{g}_i(x, y) = \sqrt{r^2(x, y) - x^2 - y^2}$. Then suppose we apply the transformations $x' = x/r(x, y)$ and $y' = y/r(x, y)$, we have $z = \sqrt{r^2(x, y) - r^2(x, y)x'^2 - r^2(x, y)y'^2} \Rightarrow z = r(x, y)\sqrt{1 - x'^2 - y'^2}$. Therefore,

$$\mathbb{1}[\{\mathbf{x}_1, \ldots, \mathbf{x}_n\} \in f_i^{-1}(F)] = \mathbb{1}[\{(x_1, y_1, \tilde{g}_1(x_1, y_1))^T, \ldots, (x_n, y_n, \tilde{g}_1(x_n, y_n))^T\} \in f_i^{-1}(F)]$$

apply transformations $x' = x/r(x, y)$ and $y' = y/r(x, y)$,

$$= \mathbb{1}\left[\left\{\left(r(x_1, y_1)x'_1, r(x_1, y_1)y'_1, r(x_1, y_1)\sqrt{1 - x'^2_1 - y'^2_1}\right)^T, \ldots,\right.\right.$$
$$\left.\left.\left(r(x_n, y_n)x'_n, r(x_n, y_n)y'_n, r(x_n, y_n)\sqrt{1 - x'^2_n - y'^2_n}\right)^T\right\} \in f_i^{-1}(F)\right]$$

$$= \mathbb{1}\left[\left\{\left(x_1', y_1', \sqrt{1-x_1'^2-y_1'^2}\right)^T, \ldots, \left(x_n', y_n', \sqrt{1-x_n'^2-y_n'^2}\right)^T\right\} \in F_i\right]$$
$$= \mathbb{1}[\{\mathbf{y}_1, \ldots, \mathbf{y}_n\} \in F_i]$$

Let us return to $P(X_{f_i^{-1}(B)} \in f_i^{-1}(F))$. We apply the transformation $f^*(x,y) = (x/r(x,y), y/r(x,y))^T$ and define the inverse as $f^{*-1}(x,y)$. We have,

$$\begin{aligned}
&P(X_{f_i^{-1}(B)} \in f_i^{-1}(F)) \\
&= \sum_{n=0}^{\infty} \frac{\exp(-\mu(f_i^{-1}(B))}{n!} \iint\limits_{f^*(P_{\mathbb{D}_i}[f_i^{-1}(B)])} \cdots \iint\limits_{f^*(P_{\mathbb{D}_i}[f_i^{-1}(B)])} \mathbb{1}[\{\mathbf{y}, \ldots, \mathbf{y}_n\} \in F_i] \\
&\qquad \prod_{i=1}^{n} \left[\rho\left(f_1^{*-1}(x_i', y_i'), f_2^{*-1}(x_i', y_i'), \tilde{g}_i(f_1^{*-1}(x_i', y_i'), f_2^{*-1}(x_i', y_i'))\right)\right. \\
&\qquad\qquad \left. l_i(f_1^{*-1}(x_i', y_i'), f_2^{*-1}(x_i', y_i')) J_{(i,f^*)}(x_i', y_i') dx_i' dy_i'\right],
\end{aligned}$$

where $J_{(i,f^*)}(x_i', y_i')$ is the Jacobian of the transformation f^*. The Jacobian of the transformation is defined as,

$$J_{(i,f^*)}(\mathbf{x}) = \frac{1}{J_{(i,f)}(f^{*-1}(\mathbf{x}))}$$

where we can use the inverse property to obtain $J_{(i,f^{*-1})}$,

$$J_{(i,f^{*-1})}(\mathbf{x}) = \det\left[\begin{pmatrix} \frac{\partial x'}{\partial x} & \frac{\partial x'}{\partial y} \\ \frac{\partial y'}{\partial x} & \frac{\partial y'}{\partial y} \end{pmatrix}\right]$$

The entries of $J_{(i,f^{*-1})}(\mathbf{x})$ are given as follows,

$$\begin{aligned}
\frac{\partial x'}{\partial x} &= \frac{y^2 + \tilde{g}_i^2(x,y) - x\tilde{g}(x,y)\frac{\partial \tilde{g}}{\partial x}}{r^3(x,y)} \\
\frac{\partial x'}{\partial y} &= \frac{-x\left(y + \tilde{g}(x,y)\frac{\partial \tilde{g}}{\partial y}\right)}{r^3(x,y)} \\
\frac{\partial y'}{\partial x} &= \frac{-y\left(x + \tilde{g}(x,y)\frac{\partial \tilde{g}}{\partial x}\right)}{r^3(x,y)} \\
\frac{\partial y'}{\partial y} &= \frac{x^2 + \tilde{g}_i^2(x,y) - y\tilde{g}(x,y)\frac{\partial \tilde{g}}{\partial y}}{r^3(x,y)},
\end{aligned}$$

where $r(x,y) = ||\mathbf{x}||$, and so

$$\begin{aligned} J_{(i,f^{*-1})}(\mathbf{x}) &= J_{(i,f^{*-1})}(x,y) \\ &= \frac{1}{r^6(x,y)} \det\left[\begin{pmatrix} y^2 + \tilde{g}_i^2(x,y) - x\tilde{g}(x,y)\frac{\partial \tilde{g}}{\partial x} & -x\left(y + \tilde{g}(x,y)\frac{\partial \tilde{g}}{\partial y}\right) \\ -y\left(x + \tilde{g}(x,y)\frac{\partial \tilde{g}}{\partial x}\right) & x^2 + \tilde{g}_i^2(x,y) - y\tilde{g}(x,y)\frac{\partial \tilde{g}}{\partial y} \end{pmatrix}\right]. \end{aligned}$$

Therefore,

$$J_{(i,f^*)}(\mathbf{x'}) = J_{(i,f^*)}(x',y') = \frac{1}{J_{(i,f)}(f_1^{*-1}(x',y'), f_2^{*-1}(x',y'))}$$

Projecting onto the sphere,

$$\begin{aligned} P(X_{f_i^{-1}(B)} \in f_i^{-1}(F)) = \sum_{n=0}^{\infty} \frac{\exp(-\mu(f_i^{-1}(B))}{n!} \\ \int_{P_{\mathbb{S}^2}^{-1}[f^*(P_{\mathbb{D}_i}[f_i^{-1}(B)])]} \cdots \int_{P_{\mathbb{S}^2}^{-1}[f^*(P_{\mathbb{D}_i}[f_i^{-1}(B)])]} \mathbb{1}[\{\mathbf{y}, \ldots, \mathbf{y}_n\} \in F_i] \prod_{i=1}^{n} \rho_i^*(\mathbf{y}_i) \lambda_{\mathbb{S}^2}(d\mathbf{y}_i), \quad \text{(B.2)} \end{aligned}$$

where $\rho_i^*(\mathbf{y}) = \rho(f_1^{*-1}(x)\ , f_2^{*-1}(y),\ \tilde{g}_i(f_1^{*-1}(x),\ f_2^{*-1}(y)))\ l_i(f_1^{*-1}(x), f_2^{*-1}(y))\ J_{i,f^{-1}}(x,y)$ $\sqrt{1-x^2-y^2}$, $\mathbf{y} = (x,y,z)^T \in \mathbb{S}^2$.

We now need to show that $P_{\mathbb{S}^2}^{-1}[f^*(P_{\mathbb{D}_i}[f_i^{-1}(B)])] = B \cap f(\mathbb{D}_i)$. Equivalently, we can show that $f^*(P_{\mathbb{D}_i}[f_i^{-1}(B)]) = P_{\mathbb{S}^2}[B \cap f(\mathbb{D}_i)]$. It is easy to see that,

$$\begin{aligned} f^*(P_{\mathbb{D}_i}[f_i^{-1}(B)]) &= \{(x/||\mathbf{x}||, y/||\mathbf{x}||)^2 \in \mathbb{R}^2 : f(\mathbf{x}) \in B \cap f(\mathbb{D}_i), \mathbf{x} = (x,y,z)^T \in \mathbb{D}_i\} \\ P_{\mathbb{S}^2}[B \cap f(\mathbb{D}_i)] &= \{(x,y) \in \mathbb{R}^2 : (x,y,z)^T \in B \cap f(\mathbb{D}_i)\}. \end{aligned}$$

Then since f is bijective (see Lemma 4.3.1) this means that for all $\mathbf{x} \in \mathbb{D}$, there exists $\mathbf{y} \in \mathbb{S}^2$ such that $\mathbf{y} = f(\mathbf{x})$. Further since $\mathbb{D}_i,\ i = 1, \ldots, n$ partition $\mathbb{D}$ this means that $f(\mathbb{D}_i),\ i = 1, \ldots, n$ partitions $\mathbb{S}^2$ and so $\mathbf{x} \in f^{-1}(B) \cap \mathbb{D}_i \Rightarrow \mathbf{y} = f(\mathbf{x}) \in B \cap f(\mathbb{D}_i)$. Hence taking the set $P_{\mathbb{S}^2}[B \cap f(\mathbb{D}_i)]$,

$$\begin{aligned} P_{\mathbb{S}^2}[B \cap f(\mathbb{D}_i)] &= \{(x,y) \in \mathbb{R}^2 : (x,y,z)^T \in B \cap f(\mathbb{D}_i)\} \\ &= \{(x,y) \in \mathbb{R}^2 : (x,y,z)^T = f(\mathbf{x}') \in B \cap f(\mathbb{D}_i), \mathbf{x}' \in \mathbb{D}_i\} \\ &= \{(x,y) \in \mathbb{R}^2 : (x,y,z)^T = (x'/||\mathbf{x}'||, y'/||\mathbf{x}'||, z'/||\mathbf{x}'||)^T \in B \cap f(\mathbb{D}_i), \\ &\qquad \mathbf{x}' = (x',y',z')^T \in \mathbb{D}_i\} \\ &= \{(x'/||\mathbf{x}'||, y'/||\mathbf{x}'||)^T \in \mathbb{R}^2 : (x'/||\mathbf{x}'||, y'/||\mathbf{x}'||, z'/||\mathbf{x}'||)^T \in B \cap f(\mathbb{D}_i), \\ &\qquad \mathbf{x}' = (x',y',z')^T \in \mathbb{D}_i\} \end{aligned}$$

$$= f^*(P_{\mathbb{D}_i}[f_i^{-1}(B)]).$$

Therefore,

$$P(X_{f_i^{-1}(B)} \in f_i^{-1}(F)) = \sum_{n=0}^{\infty} \frac{\exp(-\mu(f_i^{-1}(B))}{n!} \int_{B\cap f(\mathbb{D}_i)} \cdots \int_{B\cap f(\mathbb{D}_i)} \mathbb{1}[\{\mathbf{y}, \ldots, \mathbf{y}_n\} \in F_i] \prod_{i=1}^{n} \rho_i^*(\mathbf{y}_i) \lambda_{\mathbb{S}^2}(d\mathbf{y}_i).$$

With this identity we thus have,

$$\begin{aligned} P(Y_B \in F) &= \prod_{i=1}^{n} P(X_{f_i^{-1}(B)} \in f_i^{-1}(F)) \\ &= \prod_{i=1}^{n} \left(\sum_{n=0}^{\infty} \frac{\exp(-\mu(f_i^{-1}(B))}{n!} \int_{B\cap f(\mathbb{D}_i)} \cdots \int_{B\cap f(\mathbb{D}_i)} \mathbb{1}[\{\mathbf{y}, \ldots, \mathbf{y}_n\} \in F_i] \prod_{i=1}^{n} \rho_i^*(\mathbf{y}_i) \lambda_{\mathbb{S}^2}(d\mathbf{y}_i) \right). \end{aligned}$$

Consider the multiplication of two of the multiplicands $i \neq j$. Define $\rho_{i,j}^*(\mathbf{x}) = \rho_i^*(\mathbf{x})$ if $\mathbf{x} \in \mathbb{S}^2 \cap f(\mathbb{D}_i)$ and $\rho_{i,j}^*(\mathbf{x}) = \rho_j^*(\mathbf{x})$ if $\mathbf{x} \in \mathbb{S}^2 \cap f(\mathbb{D}_j)$ and also let $k_{p,n}(\{\mathbf{y}_i\}_{i=1}^n) = \mathbb{1}[\{\mathbf{y}_1, \ldots, \mathbf{y}_n\} \in F_p] \prod_{i=1}^n \rho_i^*(\mathbf{y}_i)$, for $p = i, j$. Then we have,

$$\begin{aligned} &P(X_{f_i^{-1}(B)} \in f_i^{-1}(F)) P(X_{f_j^{-1}(B)} \in f_j^{-1}(F)) \\ &= \left(\sum_{n=0}^{\infty} \frac{\exp(-\mu(f_i^{-1}(B)))}{n!} \int_{B\cap f(\mathbb{D}_i)} \cdots \int_{B\cap f(\mathbb{D}_i)} k_{i,n}(\{\mathbf{y}_i\}_{i=1}^n) \lambda_{\mathbb{S}^2}(d\mathbf{y}_1) \cdots \lambda_{\mathbb{S}^2}(d\mathbf{y}_n) \right) \\ &\quad \times \left(\sum_{m=0}^{\infty} \frac{\exp(-\mu(f_j^{-1}(B)))}{m!} \int_{B\cap f(\mathbb{D}_j)} \cdots \int_{B\cap f(\mathbb{D}_j)} k_{j,m}(\{\mathbf{y}_i\}_{i=1}^m) \lambda_{\mathbb{S}^2}(d\mathbf{y}_1) \cdots \lambda_{\mathbb{S}^2}(d\mathbf{y}_m) \right) \\ &= \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \exp(-\mu(f_i^{-1}(B))) \exp(-\mu(f_j^{-1}(B))) \\ &\quad \times \left(\frac{1}{n!} \int_{B\cap f(\mathbb{D}_i)} \cdots \int_{B\cap f(\mathbb{D}_i)} k_{i,n}(\{\mathbf{y}_i\}_{i=1}^n) \lambda_{\mathbb{S}^2}(d\mathbf{y}_1) \cdots \lambda_{\mathbb{S}^2}(d\mathbf{y}_n) \right) \\ &\quad\quad \times \left(\frac{1}{m!} \int_{B\cap f(\mathbb{D}_j)} \cdots \int_{B\cap f(\mathbb{D}_j)} k_{j,m}(\{\mathbf{y}_i\}_{i=1}^m) \lambda_{\mathbb{S}^2}(d\mathbf{y}_1) \cdots \lambda_{\mathbb{S}^2}(d\mathbf{y}_m) \right) \end{aligned}$$

we then use the substitution $m' = m + n$,

$$
\begin{aligned}
&= \sum_{n=0}^{\infty}\sum_{m=n}^{\infty} \exp(-\mu(f_i^{-1}(B)))\exp(-\mu(f_j^{-1}(B))) \\
&\qquad \times \left(\frac{1}{n!}\int_{B\cap f(\mathbb{D}_i)}\cdots\int_{B\cap f(\mathbb{D}_i)} k_{i,n}(\{\mathbf{y}_i\}_{i=1}^{n})\lambda_{\mathbb{S}^2}(d\mathbf{y}_1)\cdots\lambda_{\mathbb{S}^2}(d\mathbf{y}_n)\right) \\
&\qquad\qquad \times \left(\frac{1}{(m-n)!}\int_{B\cap f(\mathbb{D}_j)}\cdots\int_{B\cap f(\mathbb{D}_j)} k_{j,m-n}(\{\mathbf{y}_i\}_{i=1}^{m-n})\lambda_{\mathbb{S}^2}(d\mathbf{y}_1)\cdots\lambda_{\mathbb{S}^2}(d\mathbf{y}_{m-n})\right) \\
&= \sum_{m=0}^{\infty}\sum_{n=0}^{m} \exp(-\mu(f_i^{-1}(B)))\exp(-\mu(f_j^{-1}(B)))\frac{1}{m!}\frac{m!}{n!(m-n)!} \\
&\qquad \times \left(\int_{B\cap f(\mathbb{D}_i)}\cdots\int_{B\cap f(\mathbb{D}_i)} k_{i,n}(\{\mathbf{y}_i\}_{i=1}^{n})\lambda_{\mathbb{S}^2}(d\mathbf{y}_1)\cdots\lambda_{\mathbb{S}^2}(d\mathbf{y}_n)\right) \\
&\qquad\qquad \times \left(\int_{B\cap f(\mathbb{D}_j)}\cdots\int_{B\cap f(\mathbb{D}_j)} k_{j,m-n}(\{\mathbf{y}_i\}_{i=1}^{m-n})\lambda_{\mathbb{S}^2}(d\mathbf{y}_1)\cdots\lambda_{\mathbb{S}^2}(d\mathbf{y}_{m-n})\right) \\
&= \sum_{m=0}^{\infty} \frac{\exp(-\mu(f_i^{-1}(B)))\exp(-\mu(f_j^{-1}(B)))}{m!} \\
&\qquad \sum_{n=0}^{m}\binom{m}{n}\left(\int_{B\cap f(\mathbb{D}_i)}\cdots\int_{B\cap f(\mathbb{D}_i)} k_{i,n}(\{\mathbf{y}_i\}_{i=1}^{n})\lambda_{\mathbb{S}^2}(d\mathbf{y}_1)\cdots\lambda_{\mathbb{S}^2}(d\mathbf{y}_n)\right) \\
&\qquad\qquad \times \left(\int_{B\cap f(\mathbb{D}_j)}\cdots\int_{B\cap f(\mathbb{D}_j)} k_{j,m-n}(\{\mathbf{y}_i\}_{i=1}^{m-n})\lambda_{\mathbb{S}^2}(d\mathbf{y}_1)\cdots\lambda_{\mathbb{S}^2}(d\mathbf{y}_{m-n})\right) \\
&= \sum_{m=0}^{\infty} \frac{\exp(-\mu(f_i^{-1}(B)))\exp(-\mu(f_j^{-1}(B)))}{m!} \\
&\qquad \sum_{n=0}^{m}\binom{m}{n}\left(\int_{B\cap f(\mathbb{D}_i)}\cdots\int_{B\cap f(\mathbb{D}_i)} k_{i,n}(\{\mathbf{y}_i\}_{i=1}^{n})\lambda_{\mathbb{S}^2}(d\mathbf{y}_1)\cdots\lambda_{\mathbb{S}^2}(d\mathbf{y}_n)\right) \\
&\qquad\qquad \times \left(\int_{B\cap f(\mathbb{D}_j)}\cdots\int_{B\cap f(\mathbb{D}_j)} k_{j,m-n}(\{\mathbf{y}_i\}_{i=n+1}^{m})\lambda_{\mathbb{S}^2}(d\mathbf{y}_{n+1})\cdots\lambda_{\mathbb{S}^2}(d\mathbf{y}_m)\right) \\
&= \sum_{m=0}^{\infty} \frac{\exp(-\mu(f_i^{-1}(B)))\exp(-\mu(f_j^{-1}(B)))}{m!} \\
&\qquad \times \left(\int_{B\cap f(\mathbb{D}_i)} + \int_{B\cap f(\mathbb{D}_j)}\right)\cdots \\
&\qquad\qquad \left(\int_{B\cap f(\mathbb{D}_i)} + \int_{B\cap f(\mathbb{D}_j)}\right) k_{i,n}(\{\mathbf{y}\}_{i=1}^{n})k_{j,m-n}(\{\mathbf{y}\}_{i=n+1}^{m})\lambda_{\mathbb{S}^2}(d\mathbf{y}_1)\cdots\lambda_{\mathbb{S}^2}(d\mathbf{y}_m)
\end{aligned}
$$

Define $k_{i,j,m}(\{\mathbf{y}_i\}_{i=1}^m) = \mathbb{1}[\{\mathbf{y}_1,\ldots,\mathbf{y}_n\} \in F_{i,j}] \prod_{i=1}^n \rho_{i,j}^*(\mathbf{y}_i)$, where $F_{i,j} = \{x : x = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}, n \in \mathbb{N}, \mathbf{x}_i \in \mathbb{D}_i \cup \mathbb{D}_j, \exists y \in F \text{ such that } x \subseteq y\}$. Then it can be seen that $k_{i,n}(\{\mathbf{y}\}_{i=1}^n)$ $k_{j,m-n}(\{\mathbf{y}\}_{i=n+1}^m) = k_{i,j,m}(\{\mathbf{y}_i\}_{i=1}^m)$ and so,

$$
\begin{aligned}
&= \sum_{m=0}^{\infty} \frac{\exp(-\mu(f_i^{-1}(B)))\exp(-\mu(f_j^{-1}(B)))}{m!} \\
&\qquad \times \left(\int_{B\cap f(\mathbb{D}_i)} + \int_{B\cap f(\mathbb{D}_j)}\right)\cdots \\
&\qquad\qquad \left(\int_{B\cap f(\mathbb{D}_i)} + \int_{B\cap f(\mathbb{D}_j)}\right) k_{i,j,m}(\{\mathbf{y}_i\}_{i=1}^m)\lambda_{\mathbb{S}^2}(d\mathbf{y}_1)\cdots\lambda_{\mathbb{S}^2}(d\mathbf{y}_m) \\
&= \sum_{m=0}^{\infty} \frac{\exp(-\mu(f_i^{-1}(B)))\exp(-\mu(f_j^{-1}(B)))}{m!} \\
&\qquad \times \int_{B\cap(f(\mathbb{D}_i)\cup f(\mathbb{D}_j))}\cdots\int_{B\cap(f(\mathbb{D}_i)\cup f(\mathbb{D}_j))} k_{i,j,m}(\{\mathbf{y}_i\}_{i=1}^m)\lambda_{\mathbb{S}^2}(d\mathbf{y}_1)\cdots\lambda_{\mathbb{S}^2}(d\mathbf{y}_m)
\end{aligned}
$$

Define $B_{i,j} = B\cap(f(\mathbb{D}_i)\cup f(\mathbb{D}_j))$ and $f_{i,j}^{-1}(B) = f_i^{-1}(B)\cup f_j^{-1}(B)$ and noting that $f_i^{-1}(B)\cap f_j^{-1}(B) = \emptyset$,

$$
\begin{aligned}
&= \sum_{m=0}^{\infty} \frac{\exp(-\mu(f_{i,j}^{-1}(B)))}{m!}\int_{B_{i,j}}\cdots\int_{B_{i,j}} k_{i,j,m}(\{\mathbf{y}_i\}_{i=1}^m)\lambda_{\mathbb{S}^2}(d\mathbf{y}_1)\cdots\lambda_{\mathbb{S}^2}(d\mathbf{y}_m) \\
&= \sum_{m=0}^{\infty} \frac{\exp(-\mu(f_{i,j}^{-1}(B)))}{m!}\int_{B_{i,j}}\cdots\int_{B_{i,j}} \mathbb{1}[\{\mathbf{y}_1,\ldots,\mathbf{y}_m\}\in F_{i,j}]\prod_{i=1}^m \rho_{i,j}^*(\mathbf{y}_i)\lambda_{\mathbb{S}^2}(d\mathbf{y}_1)\cdots\lambda_{\mathbb{S}^2}(d\mathbf{y}_m)
\end{aligned}
$$

We can then repeat this argument for each multiplicand in Equation B.2 and reduce the probability measure $P(Y_B \in F)$ to,

$$
P(Y_B\in F) = \sum_{n=0}\frac{\exp(-\mu(f^{-1}(B)))}{n!}\int_B\cdots\int_B \mathbb{1}[\{\mathbf{y},\ldots,\mathbf{y}_n\}\in F]\prod_{i=1}^n \rho^*(\mathbf{y}_i)d\mathbf{y}_i,
$$

where,

$$
\rho^*(\mathbf{x}) = \begin{cases} \rho_1^*(\mathbf{x}), & \mathbf{x}\in f(\mathbb{D}_1) \\ \quad\vdots & \\ \rho_n^*(\mathbf{x}), & \mathbf{x}\in f(\mathbb{D}_n). \end{cases} \tag{B.3}
$$

Then defining $\mu^*(B) = \mu(f^{-1}(B))$ we need to show that $\mu^*(B) = \int_B \rho^*(\mathbf{x})d\mathbf{x}$. This follows

since,

$$\begin{aligned}
\mu^*(B) &= \mu(f^{-1}(B)) \\
&= \int_{f^{-1}(B)} \rho(\mathbf{x}) d\mathbf{x} \\
&= \sum_{i=1}^{n} \int_{f_i^{-1}(B)} \rho(\mathbf{x}) d\mathbf{x} \\
&= \sum_{i=1}^{n} \int_{B \cap f(\mathbb{D}_j)} \rho_i^*(\mathbf{x}) d\mathbf{x} \\
&= \int_B \rho^*(\mathbf{x}) d\mathbf{x},
\end{aligned}$$

where the penultimate line follows by a the same argument used before, first projecting from $\mathbb{D}_i$ to $\mathbb{R}^2$, then applying the transformation $x' \mapsto x/||\mathbf{x}||$ and $y' \mapsto y/||\mathbf{x}||$, then doing the inverse projection back to the sphere. This finishes the proof by again applying Lemma S1. □

B.3 Proof of Lemma 4.3.4

Proof. By assumption we suppose there is at least one point in $\mathbb{D}$ such that $\rho_1(\mathbf{x}) \neq \rho_2(\mathbf{x})$. Denote this point $\mathbf{x}^*$ and define $\mathbf{y}^* \equiv f(\mathbf{x}^*) \in f(\mathbb{D}_j)$ for some $j \in \{1, \ldots, n\}$. Then at the point $\mathbf{y}^*$ the intensities of Y_1 and Y_2 are,

$$\begin{aligned}
\rho_1^*(\mathbf{y}^*) &= \rho_1(f^{-1}(\mathbf{y}^*)) l_1(f^{-1}(\mathbf{y}^*)) J_{(1,f^{-1})}(\mathbf{y}^*) \sqrt{1 - y_1^{*2} - y_2^{*2}} \\
\rho_2^*(\mathbf{y}^*) &= \rho_1(f^{-1}(\mathbf{y}^*)) l_1(f^{-1}(\mathbf{y}^*)) J_{(1,f^{-1})}(\mathbf{y}^*) \sqrt{1 - y_1^{*2} - y_2^{*2}},
\end{aligned}$$

respectively. Then $\rho_1^*(\mathbf{y}^*) \neq \rho_2^*(\mathbf{y}^*)$, since $\rho_1(f^{-1}(\mathbf{y}^*)) = \rho_1(\mathbf{x}^*) \neq \rho_2(\mathbf{x}^*) = \rho_2(f^{-1}(\mathbf{y}^*))$. □

B.4 Proof of Theorem 4.4.1

In this section we derive the means for the estimators of F_{inhom}-, H_{inhom}-, and K_{inhom}-functions. To do this we will make use of the Campbell-Mecke Theorem ((Møller and Waagepetersen, 2003)) on $\mathbb{S}^2$,

Theorem S2. *Let X be a point process on $\mathbb{S}^2$ and h be any non-negative, measurable function*

such that $h : \mathbb{S}^2 \times N_{lf} \mapsto \mathbb{R}$. *Then,*

$$\mathbb{E}\sum_{\mathbf{x}\in X} h(\mathbf{x}, X \setminus \{\mathbf{x}\}) = \int_{\mathbb{S}^2} \mathbb{E}[h(\mathbf{x}, X^!_{\mathbf{x}})]\mu(d\mathbf{x}),$$

and if the intensity function exists then,

$$\mathbb{E}\sum_{\mathbf{x}\in X} h(\mathbf{x}, X \setminus \{\mathbf{x}\}) = \int_{\mathbb{S}^2} \mathbb{E}[h(\mathbf{x}, X^!_{\mathbf{x}})]\rho(\mathbf{x})\lambda_{\mathbb{S}^2}(d\mathbf{x}).$$

Proof. See Møller and Waagepetersen ((2003, Appendix C, pp. 248-249)). □

Noting that the reduced Palm process, $X^!_{\mathbf{x}}$, for a Poisson process, X, is again the same Poisson process Coeurjolly et al. ((2017)) we get the Slivnyak-Mecke Theorem ((Møller and Waagepetersen, 2003, Theorem 3.2, p. 21)),

$$\mathbb{E}\sum_{\mathbf{x}\in X} h(\mathbf{x}, X \setminus \{\mathbf{x}\}) = \int_{\mathbb{S}^2} \mathbb{E}[h(\mathbf{x}, X)]\rho(\mathbf{x})\lambda_{\mathbb{S}^2}(d\mathbf{x}).$$

We now give the proof the of Theorem 4.4.1.

Proof. Proofs for the expectation of $\hat{F}_{\text{inhom}}$-, and $\hat{K}_{\text{inhom}}$-functions are found in van Lieshout ((2011)) and Lawrence et al. ((2016)), Møller and Rubak ((2016)) respectively. Ratio-unbiasedness of the $\hat{H}_{\text{inhom}}$-function is also found in van Lieshout ((2011)). The proofs found in van Lieshout ((2011)) are in $\mathbb{R}^n$ but can be extended easily to $\mathbb{S}^2$. Then for the expectation of $\hat{H}_{\text{inhom}}(r)$,

$$\begin{aligned}
&\mathbb{E}\left[\frac{1}{N_X(\mathbb{S}^2)}\sum_{\mathbf{x}\in X}\prod_{\mathbf{y}\in X\setminus\{\mathbf{x}\}}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{y}\in B_{\mathbb{S}^2}(\mathbf{x},r)]}{\rho(\mathbf{y})}\right)\right] \\
&\quad= \mathbb{E}\left[\sum_{\mathbf{x}\in X}\frac{1}{N_X(\mathbb{S}^2)}\prod_{\mathbf{y}\in X\setminus\{\mathbf{x}\}}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{y}\in B_{\mathbb{S}^2}(\mathbf{x},r)]}{\rho(\mathbf{y})}\right)\right] \\
&\quad= \mathbb{E}\left[\sum_{\mathbf{x}\in X}\frac{1}{|X|}\prod_{\mathbf{y}\in X\setminus\{\mathbf{x}\}}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{y}\in B_{\mathbb{S}^2}(\mathbf{x},r)]}{\rho(\mathbf{y})}\right)\right] \\
&\quad= \mathbb{E}\left[\sum_{\mathbf{x}\in X}\frac{1}{|X\setminus\{\mathbf{x}\}\cup\{\mathbf{x}\}|}\prod_{\mathbf{y}\in X\setminus\{\mathbf{x}\}}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{y}\in B_{\mathbb{S}^2}(\mathbf{x},r)]}{\rho(\mathbf{y})}\right)\right],
\end{aligned}$$

applying the Slivnyak-Mecke Theorem,

$$= \int_{\mathbb{S}^2} \mathbb{E}\left[\frac{1}{|X \cup \{\mathbf{x}\}|} \prod_{\mathbf{y}\in X}\left(1 - \frac{\bar{\rho}\mathbb{1}[\mathbf{y} \in B_{\mathbb{S}^2}(\mathbf{x}, r)]}{\rho(\mathbf{y})}\right)\right] \rho(\mathbf{x})\lambda_{\mathbb{S}^2}(d\mathbf{x})$$

$$= \int_{\mathbb{S}^2} \mathbb{E}\left[\frac{1}{N_X(\mathbb{S}^2)+1} \prod_{\mathbf{y}\in X}\left(1 - \frac{\bar{\rho}\mathbb{1}[\mathbf{y} \in B_{\mathbb{S}^2}(\mathbf{x}, r)]}{\rho(\mathbf{y})}\right)\right] \rho(\mathbf{x})\lambda_{\mathbb{S}^2}(d\mathbf{x}). \tag{B.4}$$

We then take the expectation in the integrand and handle it separately, using the definition of a Poisson process being the independent distribution of a Poisson number of points.

$$\mathbb{E}\left[\frac{1}{N_X(\mathbb{S}^2)+1} \prod_{\mathbf{y}\in X}\left(1 - \frac{\bar{\rho}\mathbb{1}[\mathbf{y} \in B_{\mathbb{S}^2}(\mathbf{x}, r)]}{\rho(\mathbf{y})}\right)\right]$$

$$= \mathbb{E}\left[\frac{1}{N_X(\mathbb{S}^2)+1}\mathbb{E}\left[\prod_{\mathbf{y}\in X}\left(1 - \frac{\bar{\rho}\mathbb{1}[\mathbf{y} \in B_{\mathbb{S}^2}(\mathbf{x}, r)]}{\rho(\mathbf{y})}\right) \middle| N_X(\mathbb{S}^2) = n\right]\right]$$

$$= \mathbb{E}\left[\frac{1}{N_X(\mathbb{S}^2)+1}\mathbb{E}\left[\prod_{i=1}^{n}\left(1 - \frac{\bar{\rho}\mathbb{1}[\mathbf{X}_i \in B_{\mathbb{S}^2}(\mathbf{x}, r)]}{\rho(\mathbf{X}_i)}\right)\right]\right],$$

where $\mathbf{X}_i$ are independently distributed across $\mathbb{S}^2$ with density $\frac{\rho(\mathbf{x})}{\mu(\mathbb{S}^2)}$. Then taking the first expectation,

$$\mathbb{E}\left[\prod_{i=1}^{n}\left(1 - \frac{\bar{\rho}\mathbb{1}[\mathbf{X}_i \in B_{\mathbb{S}^2}(\mathbf{x}, r)]}{\rho(\mathbf{X}_i)}\right)\right] = \overbrace{\int_{\mathbb{S}^2}\cdots\int_{\mathbb{S}^2}}^{n}\prod_{i=1}^{n}\left(1 - \frac{\bar{\rho}\mathbb{1}[\mathbf{x}_i \in B_{\mathbb{S}^2}(\mathbf{x}, r)]}{\rho(\mathbf{x}_i)}\right)\frac{\rho(\mathbf{x}_i)}{\mu(\mathbb{S}^2)}\lambda_{\mathbb{S}^2}(d\mathbf{x})_i$$

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

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

Returning to the expectation over $N_X(\mathbb{S}^2)$ we have,

$$\mathbb{E}\left[\frac{1}{N_X(\mathbb{S}^2)+1} \prod_{\mathbf{y}\in X}\left(1 - \frac{\bar{\rho}\mathbb{1}[\mathbf{y} \in B_{\mathbb{S}^2}(\mathbf{x}, r)]}{\rho(\mathbf{y})}\right)\right]$$

$$= \mathbb{E}\left[\frac{1}{N_X(\mathbb{S}^2)+1}\left(1 - \frac{\bar{\rho}}{\mu(\mathbb{S}^2)}2\pi(1-\cos r)\right)^{N_X(\mathbb{S}^2)}\right],$$

define $A \equiv 1 - \frac{\bar{\rho}}{\mu(\mathbb{S}^2)} 2\pi(1-\cos r)$, then

$$
\begin{aligned}
&= \mathbb{E}\left[\frac{1}{N_X(\mathbb{S}^2)+1} A^{N_X(\mathbb{S}^2)}\right], \\
&= \sum_{n=0}^{\infty} \frac{A^n}{n+1} \frac{\lambda^n \mathrm{e}^{-\lambda}}{n!},
\end{aligned}
$$

where $\lambda \equiv \mu(\mathbb{S}^2)$,

$$
\begin{aligned}
&= \frac{\mathrm{e}^{-\lambda}}{A\lambda} \sum_{n=0}^{\infty} \frac{(A\lambda)^{n+1}}{(n+1)!} \\
&= \frac{\mathrm{e}^{-\lambda}}{A\lambda} \sum_{n=1}^{\infty} \frac{(A\lambda)^{n}}{n!} \\
&= \frac{\mathrm{e}^{-\lambda}}{A\lambda} \left(\sum_{n=0}^{\infty} \frac{(A\lambda)^{n}}{n!} - 1 \right) \\
&= \frac{\mathrm{e}^{-\lambda}}{A\lambda} \left(\mathrm{e}^{A\lambda} - 1 \right) \\
&= \frac{\mathrm{e}^{-\bar{\rho}2\pi(1-\cos r)} - \mathrm{e}^{-\mu(\mathbb{S}^2)}}{\mu(\mathbb{S}^2) - \bar{\rho}2\pi(1-\cos r)},
\end{aligned}
$$

plugging this into Equation B.4,

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

and so,

$$
\mathbb{E}[\hat{H}_{\mathrm{inhom}}(r)] = 1 - \frac{\mathrm{e}^{-\bar{\rho}2\pi(1-\cos r)} - \mathrm{e}^{-\mu(\mathbb{S}^2)}}{1 - \frac{\bar{\rho}2\pi(1-\cos r)}{\mu(\mathbb{S}^2)}}.
$$

□

B.5 Proof of Corollary 4.4.2

Proof.

$$
\mathrm{Bias}(\hat{H}_{\mathrm{inhom}}(r)) = \left(1 - \frac{\mathrm{e}^{-\bar{\rho}2\pi(1-\cos r)} - \mathrm{e}^{-\mu(\mathbb{S}^2)}}{1 - \frac{\bar{\rho}2\pi(1-\cos r)}{\mu(\mathbb{S}^2)}} \right) - \left(1 - \mathrm{e}^{-\bar{\rho}2\pi(1-\cos r)} \right)
$$

$$= \frac{\exp(\mu(\mathbb{S}^2)) - \frac{\bar{\rho}2\pi(1-\cos r)}{\mu(\mathbb{S}^2)} \exp(-2\pi(1-\cos r)\bar{\rho})}{1 - \frac{\bar{\rho}2\pi(1-\cos r)}{\mu(\mathbb{S}^2)}}$$
$$= \frac{\mu(\mathbb{S}^2)\exp(\mu(\mathbb{S}^2)) - \bar{\rho}2\pi(1-\cos r)\exp(-2\pi(1-\cos r)\bar{\rho})}{\mu(\mathbb{S}^2) - \bar{\rho}2\pi(1-\cos r)}$$

Taking the absolute value of the bias and the numerator is bounded above by,

$$\begin{aligned}
&|\mu(\mathbb{S}^2)\exp(\mu(\mathbb{S}^2)) - \bar{\rho}2\pi(1-\cos r)\exp(-2\pi(1-\cos r)\bar{\rho})| \\
&\qquad \leq \mu(\mathbb{S}^2)\exp(\mu(\mathbb{S}^2)) + \bar{\rho}2\pi(1-\cos r)\exp(-2\pi(1-\cos r)\bar{\rho}) \\
&\qquad \leq \mu(\mathbb{S}^2)\exp(\mu(\mathbb{S}^2)),
\end{aligned}$$

where the fist line follows from the triangle inequality. The denominator is bounded below,

$$\begin{aligned}
|\mu(\mathbb{S}^2) - \bar{\rho}2\pi(1-\cos r)| &\geq \mu(\mathbb{S}^2) + \bar{\rho}2\pi(1-\cos r) \\
&\geq \mu(\mathbb{S}^2),
\end{aligned}$$

where again the first line follows from the triangle inequality, since $|a-b+b| \leq |a-b|+|b| \Rightarrow |a|-|b| \leq |a-b|$, for $a, b \in \mathbb{R}$. Hence the absolute of the bias is bounded by,

$$|\text{Bias}(\hat{H}_{\text{inhom}}(r))| \leq \exp(\mu(\mathbb{S}^2)).$$

The final inequality follows by noting that $\mu(\mathbb{S}^2) = \int_{\mathbb{S}^2} \rho(\mathbf{x})\lambda_{\mathbb{S}^2}(d\mathbf{x}) \geq \int_{\mathbb{S}^2} \bar{\rho}\lambda_{\mathbb{S}^2}(d\mathbf{x}) = 4\pi\bar{\rho}$. □

B.6 Proof of Theorem 4.4.3

In this section we derive the variance of the functional summary statistics. Throughout this section we will frequently refer to the area of a spherical cap, at any point $\mathbf{o} \in \mathbb{S}^2$ with geodesic distance r by $B_{\mathbb{S}^2}(\mathbf{o}, r)$. We will also make use of the extended Campbell-Mecke Theorem Møller and Waagepetersen ((2003)) throughout,

Theorem S3. *Let X be a point process on $\mathbb{S}^2$ and h be any non-negative, measurable function such that $h : \left(\times_{i=1}^{n} \mathbb{S}^2\right) \times N_{lf} \mapsto \mathbb{R}$. Then,*

$$\mathbb{E} \sum_{\mathbf{x}_1,\ldots,\mathbf{x}_n \in X}^{\neq} h(\mathbf{x}_1, \ldots, \mathbf{x}_n, X \setminus \{\mathbf{x}_1, \ldots, \mathbf{x}_n\})$$

$$= \int_{\mathbb{S}^2} \cdots \int_{\mathbb{S}^2} \mathbb{E}[h(\mathbf{x}_1, \ldots, \mathbf{x}_n, X^!_{\mathbf{x}_1,\ldots,\mathbf{x}_n})]\alpha(\lambda_{\mathbb{S}^2}(d\mathbf{x})_1, \ldots, \lambda_{\mathbb{S}^2}(d\mathbf{x})_n)$$

and if the intensity function of order n exists then,

$$\mathbb{E} \sum_{\mathbf{x}_1,\ldots,\mathbf{x}_n \in X}^{\neq} h(\mathbf{x}_1, \ldots, \mathbf{x}_n, X \setminus \{\mathbf{x}_1, \ldots, \mathbf{x}_n\})$$
$$= \int_{\mathbb{S}^2} \cdots \int_{\mathbb{S}^2} \mathbb{E}[h(\mathbf{x}_1, \ldots, \mathbf{x}_n, X^!_{\mathbf{x}_1,\ldots,\mathbf{x}_n})]\rho^{(n)}(\mathbf{x}_1, \ldots, \mathbf{x}_n)\lambda_{\mathbb{S}^2}(\lambda_{\mathbb{S}^2}(d\mathbf{x})_1) \cdots \lambda_{\mathbb{S}^2}(\lambda_{\mathbb{S}^2}(d\mathbf{x})_n),$$

where $X^!_{\mathbf{x}_1,\ldots,\mathbf{x}_n}$ is the n^{th}-order reduced Palm process of X.

Proof. See Møller and Waagepetersen ((2003, Appendix C, pp. 248-249)). □

Again, as in the case for when $n = 1$, for any order n the reduced Palm process of order n, $X^!_{\mathbf{x}_1,\ldots,\mathbf{x}_n}$, for a Poisson process, X, is again the same Poisson process. Hence we get the extended Slivnyak-Mecke Theorem ((Møller and Waagepetersen, 2003, Theorem 3.3, p. 22)),

$$\begin{aligned} &\mathbb{E} \sum_{\mathbf{x}_1,\ldots,\mathbf{x}_n \in X}^{\neq} h(\mathbf{x}_1, \ldots, \mathbf{x}_n, X \setminus \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}) \\ &\quad = \int_{\mathbb{S}^2} \cdots \int_{\mathbb{S}^2} \mathbb{E}[h(\mathbf{x}_1, \ldots, \mathbf{x}_n, X)] \prod_{i=1}^{n} \rho(\mathbf{x}_i)\lambda_{\mathbb{S}^2}(\lambda_{\mathbb{S}^2}(d\mathbf{x})_1) \cdots \lambda_{\mathbb{S}^2}(\lambda_{\mathbb{S}^2}(d\mathbf{x})_n), \end{aligned} \tag{B.5}$$

To derive the variance of $\hat{K}_{\text{inhom}}(r)$, before which we require the following lemma,

Lemma S4. *Let X be a finite set and define X^n as the Cartesian product of X n times, i.e. $X^n = X \times \cdots \times X$, and the following sets,*

$$Y = \{(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4)^T \in X^4 : \mathbf{x}_1 \in X, \mathbf{x}_2 \in X \setminus \{\mathbf{x}_1\}, \mathbf{x}_3 \in X, \mathbf{x}_4 \in X \setminus \{\mathbf{x}_3\}\}$$
$$Y_1 = \{(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4)^T \in X^4 : \mathbf{x}_1 \in X, \mathbf{x}_2 \in X \setminus \{\mathbf{x}_1\}, \mathbf{x}_3 \in X \setminus \{\mathbf{x}_1, \mathbf{x}_2\}, \mathbf{x}_4 \in X \setminus \{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}\}$$
$$Y_2 = \{(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4)^T \in X^4 : \mathbf{x}_1 \in X, \mathbf{x}_2 \in X \setminus \{\mathbf{x}_1\}, \mathbf{x}_3 \in \{\mathbf{x}_1, \mathbf{x}_2\}, \mathbf{x}_4 \in \{\mathbf{x}_1, \mathbf{x}_2\} \setminus \{\mathbf{x}_3\}\}$$
$$Y_3 = \{(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4)^T \in X^4 : \mathbf{x}_1 \in X, \mathbf{x}_2 \in X \setminus \{\mathbf{x}_1\}, \mathbf{x}_3 \in X \setminus \{\mathbf{x}_1, \mathbf{x}_2\}, \mathbf{x}_4 \in \{\mathbf{x}_1, \mathbf{x}_2\} \setminus \{\mathbf{x}_3\}\}$$
$$Y_4 = \{(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4)^T \in X^4 : \mathbf{x}_1 \in X, \mathbf{x}_2 \in X \setminus \{\mathbf{x}_1\}, \mathbf{x}_3 \in \{\mathbf{x}_1, \mathbf{x}_2\}, \mathbf{x}_4 \in X \setminus \{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}\},$$

then $Y = \cup_{i=1,\ldots,4} Y_i$ and $Y_i, i = 1, \ldots, 4$ are pairwise disjoint.

Proof. Pairwise disjointness of the sets $Y_i, i = 1 \ldots 4$ follows from the definitions of the sets.

To prove equality we will show the following,

$$\mathbf{r} \in Y \Rightarrow \mathbf{r} \in \cup_{i=1,\ldots,4} Y_i, \quad \text{and} \tag{B.6}$$

$$\mathbf{r} \in \cup_{i=1,\ldots,4} Y_i \Rightarrow \mathbf{r} \in Y \tag{B.7}$$

Statement B.7, holds by considering each set Y_i in turn. In particular it is clear by the definitions of the sets that for each $i = 1, \ldots, 4, Y_i \subseteq Y$ and so Statement B.7 holds.

To show Statement B.6 let $\mathbf{x} \in Y$ and fix $\mathbf{x}_1 \in X$ and $\mathbf{x}_2 \in X \setminus \{\mathbf{x}_1\}$. Then there are two possibilities for $\mathbf{x}_3$, either $\mathbf{x}_3 \in X \setminus \{\mathbf{x}_1, \mathbf{x}_2\}$ or $\mathbf{x}_3 \in \{\mathbf{x}_1, \mathbf{x}_2\}$. If the former holds then $\mathbf{x}_4$ can either be in $X \setminus \{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ or $\{\mathbf{x}_1, \mathbf{x}_2\} \setminus \{\mathbf{x}_3\}$. If the first holds then $\mathbf{x} \in Y_1$ and if the second holds then $\mathbf{x} \in Y_3$. Considering all possible combinations proves Statement B.6. Hence it follows that $Y = \cup_{i=1,\ldots,4} Y_i$. □

We now proceed with the proof for the variance of estimates of the inhomogeneous functional summary statistics. The proof is spread over three parts, one for each functional summary statistic.

Proof. VARIANCE OF $\hat{K}_{\text{INHOM}}(r)$

We expand the variance as $\text{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}^2[X]$,

$$\begin{aligned}
\text{Var}(\tilde{K}_{\text{inhom}}(r)) &= \text{Var}\left(\frac{1}{4\pi} \sum_{\mathbf{x} \in X} \sum_{\mathbf{y} \in X \setminus \{\mathbf{x}\}} \frac{\mathbb{1}[d(\mathbf{x}, \mathbf{y}) \leq r]}{\rho(\mathbf{x})\rho(\mathbf{y})}\right) \\
&= \frac{1}{16\pi^2}\Bigg[\underbrace{\mathbb{E}\left(\sum_{\mathbf{x} \in X} \sum_{\mathbf{y} \in X \setminus \{\mathbf{x}\}} \frac{\mathbb{1}[d(\mathbf{x}, \mathbf{y}) \leq r]}{\rho(\mathbf{x})\rho(\mathbf{y})}\right)^2}_{(1)} - \underbrace{\mathbb{E}^2 \sum_{\mathbf{x} \in X} \sum_{\mathbf{y} \in X \setminus \{\mathbf{x}\}} \frac{\mathbb{1}(d(\mathbf{x}, \mathbf{y}) \leq r)}{\rho(\mathbf{x})\rho(\mathbf{y})}}_{(2)}\Bigg],
\end{aligned} \tag{B.8}$$

We deal with each of the terms individually. First consider term (2) of the previous equation, this is simply the inhomogeneous K-function for a Poisson process,

$$\begin{aligned}
\frac{1}{16\pi^2}\mathbb{E}^2 \sum_{\mathbf{x} \in X} \sum_{\mathbf{y} \in X \setminus \{\mathbf{x}\}} \frac{\mathbb{1}(d(\mathbf{x}, \mathbf{y}) \leq r)}{\rho(\mathbf{x})\rho(\mathbf{y})} &= \mathbb{E}^2 \frac{1}{4\pi} \sum_{\mathbf{x} \in X} \sum_{\mathbf{y} \in X \setminus \{\mathbf{x}\}} \frac{\mathbb{1}(d(\mathbf{x}, \mathbf{y}) \leq r)}{\rho(\mathbf{x})\rho(\mathbf{y})} \\
&= K^2_{\text{inhom}}(r) \\
&= 4\pi^2(1 - \cos r)^2,
\end{aligned}$$

where the penultimate equality follows from the definition of K_{inhom} (taking the arbitrary area B to be $\mathbb{S}^2$) and the final equality follows from Theorem 4.4.1. To handle term (1) we first expand the square,

$$\begin{aligned}
&\frac{1}{16\pi^2}\mathbb{E}\left(\sum_{\mathbf{x}\in X}\sum_{\mathbf{y}\in X\setminus\{\mathbf{x}\}}\frac{\mathbb{1}[d(\mathbf{x},\mathbf{y})\le r]}{\rho(\mathbf{x})\rho(\mathbf{y})}\right)^2 \\
&\quad=\frac{1}{16\pi^2}\mathbb{E}\sum_{\mathbf{x}\in X}\sum_{\mathbf{y}\in X\setminus\{\mathbf{x}\}}\frac{\mathbb{1}[d(\mathbf{x},\mathbf{y})\le r]}{\rho(\mathbf{x})\rho(\mathbf{y})}\sum_{\mathbf{x}'\in X}\sum_{\mathbf{y}'\in X\setminus\{\mathbf{x}'\}}\frac{\mathbb{1}[d(\mathbf{x}',\mathbf{y}')\le r]}{\rho(\mathbf{x}')\rho(\mathbf{y}')} \\
&\quad=\frac{1}{16\pi^2}\mathbb{E}\sum_{\mathbf{x}\in X}\sum_{\mathbf{y}\in X\setminus\{\mathbf{x}\}}\sum_{\mathbf{x}'\in X}\sum_{\mathbf{y}'\in X\setminus\{\mathbf{x}'\}}\frac{\mathbb{1}[d(\mathbf{x},\mathbf{y})\le r]\mathbb{1}[d(\mathbf{x}',\mathbf{y}')\le r]}{\rho(\mathbf{x})\rho(\mathbf{y})\rho(\mathbf{x}')\rho(\mathbf{y}')}
\end{aligned}$$

Let us define the summand as,

$$f_r(\mathbf{x},\mathbf{y},\mathbf{x}',\mathbf{y}')=\frac{\mathbb{1}[d(\mathbf{x},\mathbf{y})\le r]\mathbb{1}[d(\mathbf{x}',\mathbf{y}')\le r]}{\rho(\mathbf{x})\rho(\mathbf{y})\rho(\mathbf{x}')\rho(\mathbf{y}')},$$

and then by Lemma S4 we can divide the sum in the expectation into 4 terms,

$$\begin{aligned}
&\underbrace{\frac{1}{16\pi^2}\mathbb{E}\sum_{\mathbf{x}\in X}\sum_{\mathbf{y}\in X\setminus\{\mathbf{x}\}}\sum_{\mathbf{x}'\in X\setminus\{\mathbf{x},\mathbf{y}\}}\sum_{\mathbf{y}'\in X\setminus\{\mathbf{x}',\mathbf{x},\mathbf{y}\}}f_r(\mathbf{x},\mathbf{y},\mathbf{x}',\mathbf{y}')}_{(a)} \\
&\quad+\underbrace{\frac{1}{16\pi^2}\mathbb{E}\sum_{\mathbf{x}\in X}\sum_{\mathbf{y}\in X\setminus\{\mathbf{x}\}}\sum_{\mathbf{x}'\in\{\mathbf{x},\mathbf{y}\}}\sum_{\mathbf{y}'\in\{\mathbf{x},\mathbf{y}\}\setminus\{\mathbf{x}'\}}f_r(\mathbf{x},\mathbf{y},\mathbf{x}',\mathbf{y}')}_{(b)} \\
&\quad+\underbrace{\frac{1}{16\pi^2}\mathbb{E}\sum_{\mathbf{x}\in X}\sum_{\mathbf{y}\in X\setminus\{\mathbf{x}\}}\sum_{\mathbf{x}'\in X\setminus\{\mathbf{x},\mathbf{y}\}}\sum_{\mathbf{y}'\in\{\mathbf{x},\mathbf{y}\}\setminus\{\mathbf{x}'\}}f_r(\mathbf{x},\mathbf{y},\mathbf{x}',\mathbf{y}')}_{(c)} \\
&\quad+\underbrace{\frac{1}{16\pi^2}\mathbb{E}\sum_{\mathbf{x}\in X}\sum_{\mathbf{y}\in X\setminus\{\mathbf{x}\}}\sum_{\mathbf{x}'\in\{\mathbf{x},\mathbf{y}\}}\sum_{\mathbf{y}'\in X\setminus\{\mathbf{x},\mathbf{y},\mathbf{x}'\}}f_r(\mathbf{x},\mathbf{y},\mathbf{x}',\mathbf{y}')}_{(d)}.
\end{aligned}$$

We handle these terms independently. For term (a) we can directly apply the extended Slivnyak-Mecke Theorem given by Equation B.5,

$$\frac{1}{16\pi^2}\mathbb{E}\sum_{\mathbf{x}\in X}\sum_{\mathbf{y}\in X\setminus\{\mathbf{x}\}}\sum_{\mathbf{x}'\in X\setminus\{\mathbf{x},\mathbf{y}\}}\sum_{\mathbf{y}'\in X\setminus\{\mathbf{x}',\mathbf{x},\mathbf{y}\}}f_r(\mathbf{x},\mathbf{y},\mathbf{x}',\mathbf{y}')$$

$$
\begin{aligned}
&= \frac{1}{16\pi^2} \int_{\mathbb{S}^2} \cdots \int_{\mathbb{S}^2} f_r(\mathbf{x}, \mathbf{y}, \mathbf{x}', \mathbf{y}') \rho(\mathbf{x}) \rho(\mathbf{y}) \rho(\mathbf{x}') \rho(\mathbf{y}') \lambda_{\mathbb{S}^2}(d\mathbf{x}) \lambda_{\mathbb{S}^2}(d\mathbf{y}) \lambda_{\mathbb{S}^2}(d\mathbf{x})' \lambda_{\mathbb{S}^2}(d\mathbf{y})' \\
&= \frac{1}{16\pi^2} \int_{\mathbb{S}^2} \cdots \int_{\mathbb{S}^2} \mathbb{1}[d(\mathbf{x}, \mathbf{y}) \leq r] \mathbb{1}[d(\mathbf{x}', \mathbf{y}') \leq r] \lambda_{\mathbb{S}^2}(d\mathbf{x}) \lambda_{\mathbb{S}^2}(d\mathbf{y}) \lambda_{\mathbb{S}^2}(d\mathbf{x})' \lambda_{\mathbb{S}^2}(d\mathbf{y})' \\
&= \frac{1}{16\pi^2} \left(\int_{\mathbb{S}^2} \int_{\mathbb{S}^2} \mathbb{1}[d(\mathbf{x}, \mathbf{y}) \leq r] \lambda_{\mathbb{S}^2}(d\mathbf{x}) \lambda_{\mathbb{S}^2}(d\mathbf{y}) \right)^2 \\
&= \frac{1}{16\pi^2} \left(\int_{\mathbb{S}^2} \lambda_{\mathbb{S}^2}(B_{\mathbb{S}^2}(\mathbf{o}, r)) \lambda_{\mathbb{S}^2}(d\mathbf{y}) \right)^2 \\
&= \frac{1}{16\pi^2} \left(4\pi \lambda_{\mathbb{S}^2}(B_{\mathbb{S}^2}(\mathbf{o}, r)) \right)^2 \\
&= \lambda_{\mathbb{S}^2}(B_{\mathbb{S}^2}(\mathbf{o}, r))^2,
\end{aligned}
$$

where the penultimate equality follows since the area of the spherical cap is constant for a fixed geodesic radius for any centre, $\mathbf{0}$ indicates any arbitrary point in $\mathbb{S}^2$. Term (b) can be handled in a similar manner as term (a),

$$
\begin{aligned}
&\frac{1}{16\pi^2} \mathbb{E} \sum_{\mathbf{x} \in X} \sum_{\mathbf{y} \in X \setminus \{\mathbf{x}\}} \sum_{\mathbf{x}' \in \{\mathbf{x}, \mathbf{y}\}} \sum_{\mathbf{y}' \in \{\mathbf{x}, \mathbf{y}\} \setminus \{\mathbf{x}'\}} f_r(\mathbf{x}, \mathbf{y}, \mathbf{x}', \mathbf{y}') \\
&\qquad = \frac{1}{16\pi^2} \mathbb{E} \sum_{\mathbf{x} \in X} \sum_{\mathbf{y} \in X \setminus \{\mathbf{x}\}} f_r(\mathbf{x}, \mathbf{y}, \mathbf{x}, \mathbf{y}) + f_r(\mathbf{x}, \mathbf{y}, \mathbf{y}, \mathbf{x}) \\
&\qquad = \frac{1}{16\pi^2} \mathbb{E} \sum_{\mathbf{x} \in X} \sum_{\mathbf{y} \in X \setminus \{\mathbf{x}\}} \frac{\mathbb{1}[d(\mathbf{x}, \mathbf{y}) \leq r] \mathbb{1}[d(\mathbf{x}, \mathbf{y}) \leq r]}{\rho(\mathbf{x}) \rho(\mathbf{y}) \rho(\mathbf{x}) \rho(\mathbf{y})} \\
&\qquad\qquad + \frac{\mathbb{1}[d(\mathbf{x}, \mathbf{y}) \leq r] \mathbb{1}[d(\mathbf{y}, \mathbf{x}) \leq r]}{\rho(\mathbf{x}) \rho(\mathbf{y}) \rho(\mathbf{y}) \rho(\mathbf{x})} \\
&\qquad = \frac{1}{8\pi^2} \mathbb{E} \sum_{\mathbf{x} \in X} \sum_{\mathbf{y} \in X \setminus \{\mathbf{x}\}} \frac{\mathbb{1}[d(\mathbf{x}, \mathbf{y}) \leq r]}{\rho^2(\mathbf{x}) \rho^2(\mathbf{y})}
\end{aligned}
$$

Hence by the extended Slivnyak-Mecke Theorem again we have,

$$
\begin{aligned}
\frac{1}{8\pi^2} \mathbb{E} \sum_{\mathbf{x} \in X} \sum_{\mathbf{y} \in X \setminus \{\mathbf{x}\}} \frac{\mathbb{1}[d(\mathbf{x}, \mathbf{y}) \leq r]}{\rho^2(\mathbf{x}) \rho^2(\mathbf{y})} &= \frac{1}{8\pi^2} \int_{\mathbb{S}^2} \int_{\mathbb{S}^2} \frac{\mathbb{1}[d(\mathbf{x}, \mathbf{y}) \leq r]}{\rho^2(\mathbf{x}) \rho^2(\mathbf{y})} \rho(\mathbf{x}) \rho(\mathbf{y}) \lambda_{\mathbb{S}^2}(d\mathbf{x}) \lambda_{\mathbb{S}^2}(d\mathbf{y}) \\
&= \frac{1}{8\pi^2} \int_{\mathbb{S}^2} \int_{\mathbb{S}^2} \frac{\mathbb{1}[d(\mathbf{x}, \mathbf{y}) \leq r]}{\rho(\mathbf{x}) \rho(\mathbf{y})} \lambda_{\mathbb{S}^2}(d\mathbf{x}) \lambda_{\mathbb{S}^2}(d\mathbf{y})
\end{aligned}
$$

We now consider term (c),

$$
\frac{1}{16\pi^2} \mathbb{E} \sum_{\mathbf{x} \in X} \sum_{\mathbf{y} \in X \setminus \{\mathbf{x}\}} \sum_{\mathbf{x}' \in X \setminus \{\mathbf{x}, \mathbf{y}\}} \sum_{\mathbf{y}' \in \{\mathbf{x}, \mathbf{y}\} \setminus \{\mathbf{x}'\}} f(\mathbf{x}, \mathbf{y}, \mathbf{x}', \mathbf{y}')
$$

$$= \frac{1}{16\pi^2}\mathbb{E}\sum_{\mathbf{x}\in X}\sum_{\mathbf{y}\in X\setminus\{\mathbf{x}\}}\sum_{\mathbf{x}'\in X\setminus\{\mathbf{x},\mathbf{y}\}}\underbrace{f_r(\mathbf{x},\mathbf{y},\mathbf{x}',\mathbf{x})}_{(I)}+\underbrace{f_r(\mathbf{x},\mathbf{y},\mathbf{x}',\mathbf{y})}_{(II)}.$$

We can handle these terms independently.

$$\begin{aligned}
&\frac{1}{16\pi^2}\mathbb{E}\sum_{\mathbf{x}\in X}\sum_{\mathbf{y}\in X\setminus\{\mathbf{x}\}}\sum_{\mathbf{x}'\in X\setminus\{\mathbf{x},\mathbf{y}\}} f_r(\mathbf{x},\mathbf{y},\mathbf{x}',\mathbf{x})\\
&\quad= \frac{1}{16\pi^2}\mathbb{E}\sum_{\mathbf{x}\in X}\sum_{\mathbf{y}\in X\setminus\{\mathbf{x}\}}\sum_{\mathbf{x}'\in X\setminus\{\mathbf{x},\mathbf{y}\}}\frac{\mathbb{1}[d(\mathbf{x},\mathbf{y})\le r]\mathbb{1}[d(\mathbf{x}',\mathbf{x})\le r]}{\rho(\mathbf{x})\rho(\mathbf{y})\rho(\mathbf{x}')\rho(\mathbf{x})}\\
&\quad= \frac{1}{16\pi^2}\int_{\mathbb{S}^2}\int_{\mathbb{S}^2}\int_{\mathbb{S}^2}\frac{\mathbb{1}[d(\mathbf{x},\mathbf{y})\le r]\mathbb{1}[d(\mathbf{x}',\mathbf{x})\le r]}{\rho^2(\mathbf{x})\rho(\mathbf{y})\rho(\mathbf{x}')}\rho(\mathbf{x})\rho(\mathbf{y})\rho(\mathbf{x}')\lambda_{\mathbb{S}^2}(d\mathbf{x})\lambda_{\mathbb{S}^2}(d\mathbf{y})\lambda_{\mathbb{S}^2}(d\mathbf{x})'\\
&\quad= \frac{1}{16\pi^2}\int_{\mathbb{S}^2}\int_{\mathbb{S}^2}\int_{\mathbb{S}^2}\frac{\mathbb{1}[d(\mathbf{x},\mathbf{y})\le r]\mathbb{1}[d(\mathbf{x}',\mathbf{x})\le r]}{\rho(\mathbf{x})}\lambda_{\mathbb{S}^2}(d\mathbf{x})\lambda_{\mathbb{S}^2}(d\mathbf{y})\lambda_{\mathbb{S}^2}(d\mathbf{x})'\\
&\quad= \frac{1}{16\pi^2}\int_{\mathbb{S}^2}\int_{\mathbb{S}^2}\frac{\mathbb{1}[d(\mathbf{x},\mathbf{y})\le r]}{\rho(\mathbf{x})}\left(\int_{\mathbb{S}^2}\mathbb{1}[d(\mathbf{x}',\mathbf{x})\le r]\lambda_{\mathbb{S}^2}(d\mathbf{x})'\right)\lambda_{\mathbb{S}^2}(d\mathbf{x})\lambda_{\mathbb{S}^2}(d\mathbf{y})\\
&\quad= \frac{\lambda_{\mathbb{S}^2}(B_{\mathbb{S}^2}(\mathbf{o},r))}{16\pi^2}\int_{\mathbb{S}^2}\frac{1}{\rho(\mathbf{x})}\left(\int_{\mathbb{S}^2}\mathbb{1}[d(\mathbf{x},\mathbf{y})\le r]\lambda_{\mathbb{S}^2}(d\mathbf{y})\right)\lambda_{\mathbb{S}^2}(d\mathbf{x})\\
&\quad= \frac{\lambda_{\mathbb{S}^2}(B_{\mathbb{S}^2}(\mathbf{o},r))^2}{16\pi^2}\int_{\mathbb{S}^2}\frac{1}{\rho(\mathbf{x})}\lambda_{\mathbb{S}^2}(d\mathbf{x}),
\end{aligned}$$

where the second equality follows by the Slivnyak-Mecke Theorem. By an identical argument term II is given by,

$$\frac{1}{16\pi^2}\mathbb{E}\sum_{\mathbf{x}\in X}\sum_{\mathbf{y}\in X\setminus\{\mathbf{x}\}}\sum_{\mathbf{x}'\in X\setminus\{\mathbf{x},\mathbf{y}\}} f_r(\mathbf{x},\mathbf{y},\mathbf{x}',\mathbf{y}) = \frac{\lambda_{\mathbb{S}^2}(B_{\mathbb{S}^2}(\mathbf{o},r))^2}{16\pi^2}\int_{\mathbb{S}^2}\frac{1}{\rho(\mathbf{y})}\lambda_{\mathbb{S}^2}(d\mathbf{y})$$

Term (d) can be handled in an identical manner as term (c). To see this consider the summation over $\mathbf{y}'$ in term (d). Since $\mathbf{x}' \in \{\mathbf{x},\mathbf{y}\}$ this means that the set $X\setminus\{\mathbf{x},\mathbf{y},\mathbf{x}'\}$ is identical to $X\setminus\{\mathbf{x},\mathbf{y}\}$ and so the summations over $\mathbf{x}'$ and $\mathbf{y}'$ can be interchanged. Therefore,

$$\frac{1}{16\pi^2}\mathbb{E}\sum_{\mathbf{x}\in X}\sum_{\mathbf{y}\in X\setminus\{\mathbf{x}\}}\sum_{\mathbf{x}'\in\{\mathbf{x},\mathbf{y}\}}\sum_{\mathbf{y}'\in X\setminus\{\mathbf{x},\mathbf{y},\mathbf{x}'\}} f_r(\mathbf{x},\mathbf{y},\mathbf{x}',\mathbf{y}') = \frac{2\lambda_{\mathbb{S}^2}(B_{\mathbb{S}^2}(\mathbf{o},r))^2}{16\pi^2}\int_{\mathbb{S}^2}\frac{1}{\rho(\mathbf{x})}\lambda_{\mathbb{S}^2}(d\mathbf{x}).$$

Further we note that the area of spherical cap is given by $\lambda_{\mathbb{S}^2}(B_{\mathbb{S}^2}(\mathbf{o},r)) = 2\pi(1-\cos r)$. Thus collecting all the terms gives the form of the variance,

$$\mathrm{Var}(\hat{K}_{\mathrm{inhom}}(r)) = \frac{1}{8\pi^2}\int_{\mathbb{S}^2}\int_{\mathbb{S}^2}\frac{\mathbb{1}[d(\mathbf{x},\mathbf{y})\le r]}{\rho(\mathbf{x})\rho(\mathbf{y})}\lambda_{\mathbb{S}^2}(d\mathbf{x})\lambda_{\mathbb{S}^2}(d\mathbf{y}) + (1-\cos r)^2\int_{\mathbb{S}^2}\frac{1}{\rho(\mathbf{x})}\lambda_{\mathbb{S}^2}(d\mathbf{x}). \tag{B.9}$$

□

Proof. VARIANCE OF $\hat{F}_{\text{INHOM}}(r)$

Restating the estimator for $F_{\text{inhom}(r)}$,

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

Taking the variance using $\text{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}^2[X]$,

$$\begin{aligned}\text{Var}(\hat{F}_{\text{inhom}}(r)) &= \frac{1}{|P|^2}\text{Var}\sum_{\mathbf{p}\in P}\prod_{\mathbf{x}\in X}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{x}\in B_{\mathbb{S}^2}(\mathbf{p},r)]}{\rho(\mathbf{x})}\right)\\ &= \frac{1}{|P|^2}\mathbb{E}\left(\sum_{\mathbf{p}\in P}\prod_{\mathbf{x}\in X}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{x}\in B_{\mathbb{S}^2}(\mathbf{p},r)]}{\rho(\mathbf{x})}\right)\right)^2 - \frac{1}{|P|^2}\mathbb{E}^2\sum_{\mathbf{p}\in P}\prod_{\mathbf{x}\in X}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{x}\in B_{\mathbb{S}^2}(\mathbf{p},r)]}{\rho(\mathbf{x})}\right)\end{aligned} \tag{B.10}$$

Dealing with each term independently, we have

$$\begin{aligned}&\mathbb{E}\left(\sum_{\mathbf{p}\in P}\prod_{\mathbf{x}\in X}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{x}\in B_{\mathbb{S}^2}(\mathbf{p},r)]}{\rho(\mathbf{x})}\right)\right)^2\\ &\quad= \mathbb{E}\sum_{\mathbf{p}\in P}\prod_{\mathbf{x}\in X}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{x}\in B_{\mathbb{S}^2}(\mathbf{p},r)]}{\rho(\mathbf{x})}\right)\sum_{\mathbf{p}'\in P}\prod_{\mathbf{y}\in X}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{y}\in B_{\mathbb{S}^2}(\mathbf{p}',r)]}{\rho(\mathbf{y})}\right)\\ &\quad= \sum_{\mathbf{p}\in P}\sum_{\mathbf{p}'\in P}\mathbb{E}\prod_{\mathbf{x}\in X}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{x}\in B_{\mathbb{S}^2}(\mathbf{p},r)]}{\rho(\mathbf{x})}\right)\prod_{\mathbf{y}\in X}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{y}\in B_{\mathbb{S}^2}(\mathbf{p}',r)]}{\rho(\mathbf{y})}\right)\end{aligned} \tag{B.11}$$

From the proof of Theorem 1 given by van Lieshout ((2011)), we have the following identity,

$$\prod_{\mathbf{y}\in X}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{y}\in B_{\mathbb{S}^2}(\mathbf{x},r)]}{\rho(\mathbf{y})}\right) = 1+\sum_{n=1}^{\infty}\frac{(-\bar{\rho})^n}{n!}\sum_{\mathbf{x}_1,\ldots,\mathbf{x}_n\in X}^{\neq}\prod_{i=1}^{n}\frac{\mathbb{1}[\mathbf{x}_i\in B_{\mathbb{S}^2}(\mathbf{x},r)]}{\rho(\mathbf{x}_i)},$$

and using the convention that a sum over an emptyset is 0, i.e. $\sum_{k=1}^{0} = \sum_{x\subseteq\emptyset} = \sum_{\emptyset\in x} = 1$,

$$\prod_{\mathbf{y}\in X}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{y}\in B_{\mathbb{S}^2}(\mathbf{x},r)]}{\rho(\mathbf{y})}\right) = \sum_{n=0}^{\infty}\frac{(-\bar{\rho})^n}{n!}\sum_{\mathbf{x}_1,\ldots,\mathbf{x}_n\in X}^{\neq}\prod_{i=1}^{n}\frac{\mathbb{1}[\mathbf{x}_i\in B_{\mathbb{S}^2}(\mathbf{x},r)]}{\rho(\mathbf{x}_i)}. \tag{B.12}$$

Taking just the expectation from Equation B.11 we expand the first product over $\mathbf{x}$ using the previous identity to give,

$$
\begin{aligned}
&\mathbb{E}\prod_{\mathbf{x}\in X}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{x}\in B_{\mathbb{S}^2}(\mathbf{p},r)]}{\rho(\mathbf{x})}\right)\prod_{\mathbf{y}\in X}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{y}\in B_{\mathbb{S}^2}(\mathbf{p}',r)]}{\rho(\mathbf{y})}\right)\\
&=\mathbb{E}\left(\sum_{n=0}^{\infty}\frac{(-\bar{\rho})^n}{n!}\sum_{\mathbf{x}_1,\ldots,\mathbf{x}_n\in X}^{\neq}\prod_{i=1}^{n}\frac{\mathbb{1}[\mathbf{x}_i\in B_{\mathbb{S}^2}(\mathbf{p},r)]}{\rho(\mathbf{x}_i)}\right)\prod_{\mathbf{y}\in X}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{y}\in B_{\mathbb{S}^2}(\mathbf{p}',r)]}{\rho(\mathbf{y})}\right)\\
&=\mathbb{E}\sum_{n=0}^{\infty}\frac{(-\bar{\rho})^n}{n!}\sum_{\mathbf{x}_1,\ldots,\mathbf{x}_n\in X}^{\neq}\left(\prod_{i=1}^{n}\frac{\mathbb{1}[\mathbf{x}_i\in B_{\mathbb{S}^2}(\mathbf{p},r)]}{\rho(\mathbf{x}_i)}\right.\\
&\qquad\left.\prod_{\mathbf{y}\in\{X\setminus\{\mathbf{x}_1,\ldots,\mathbf{x}_n\},\{\mathbf{x}_1,\ldots,\mathbf{x}_n\}\}}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{y}\in B_{\mathbb{S}^2}(\mathbf{p}',r)]}{\rho(\mathbf{y})}\right)\right)\\
&=\mathbb{E}\sum_{n=0}^{\infty}\frac{(-\bar{\rho})^n}{n!}\sum_{\mathbf{x}_1,\ldots,\mathbf{x}_n\in X}^{\neq}\left(\prod_{i=1}^{n}\frac{\mathbb{1}[\mathbf{x}_i\in B_{\mathbb{S}^2}(\mathbf{p},r)]}{\rho(\mathbf{x}_i)}\prod_{\mathbf{y}\in\{\mathbf{x}_1,\ldots,\mathbf{x}_n\}}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{y}\in B_{\mathbb{S}^2}(\mathbf{p}',r)]}{\rho(\mathbf{y})}\right)\right.\\
&\qquad\left.\prod_{\mathbf{y}\in X\setminus\{\mathbf{x}_1,\ldots,\mathbf{x}_n\}}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{y}\in B_{\mathbb{S}^2}(\mathbf{p}',r)]}{\rho(\mathbf{y})}\right)\right)\\
&=\mathbb{E}\sum_{n=0}^{\infty}\frac{(-\bar{\rho})^n}{n!}\sum_{\mathbf{x}_1,\ldots,\mathbf{x}_n\in X}^{\neq}\left(\prod_{i=1}^{n}\frac{\mathbb{1}[\mathbf{x}_i\in B_{\mathbb{S}^2}(\mathbf{p},r)]}{\rho(\mathbf{x}_i)}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{x}_i\in B_{\mathbb{S}^2}(\mathbf{p}',r)]}{\rho(\mathbf{x}_i)}\right)\right.\\
&\qquad\left.\prod_{\mathbf{y}\in X\setminus\{\mathbf{x}_1,\ldots,\mathbf{x}_n\}}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{y}\in B_{\mathbb{S}^2}(\mathbf{p}',r)]}{\rho(\mathbf{y})}\right)\right)\\
&=\sum_{n=0}^{\infty}\frac{(-\bar{\rho})^n}{n!}\mathbb{E}\sum_{\mathbf{x}_1,\ldots,\mathbf{x}_n\in X}^{\neq}\left(\prod_{i=1}^{n}\frac{\mathbb{1}[\mathbf{x}_i\in B_{\mathbb{S}^2}(\mathbf{p},r)]}{\rho(\mathbf{x}_i)}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{x}_i\in B_{\mathbb{S}^2}(\mathbf{p}',r)]}{\rho(\mathbf{x}_i)}\right)\right.\\
&\qquad\left.\prod_{\mathbf{y}\in X\setminus\{\mathbf{x}_1,\ldots,\mathbf{x}_n\}}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{y}\in B_{\mathbb{S}^2}(\mathbf{p}',r)]}{\rho(\mathbf{y})}\right)\right)
\end{aligned}
$$

Using the extended Slivnyak-Mecke Theorem,

$$
\begin{aligned}
&=\sum_{n=0}^{\infty}\frac{(-\bar{\rho})^n}{n!}\overbrace{\int_{\mathbb{S}^2}\cdots\int_{\mathbb{S}^2}}^{n}\\
&\qquad\mathbb{E}\left(\prod_{i=1}^{n}\frac{\mathbb{1}[\mathbf{x}_i\in B_{\mathbb{S}^2}(\mathbf{p},r)]}{\rho(\mathbf{x}_i)}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{x}_i\in B_{\mathbb{S}^2}(\mathbf{p}',r)]}{\rho(\mathbf{x}_i)}\right)\right.
\end{aligned}
$$

$$\prod_{\mathbf{y}\in X}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{y}\in B_{\mathbb{S}^2}(\mathbf{p}',r)]}{\rho(\mathbf{y})}\right)\Bigg)\prod_{i=1}^{n}\rho(\mathbf{x}_i)\lambda_{\mathbb{S}^2}(d\mathbf{x}_i)$$

$$=\sum_{n=0}^{\infty}\frac{(-\bar{\rho})^n}{n!}\overbrace{\int_{\mathbb{S}^2}\cdots\int_{\mathbb{S}^2}}^{n}$$

$$\prod_{i=1}^{n}\frac{\mathbb{1}[\mathbf{x}_i\in B_{\mathbb{S}^2}(\mathbf{p},r)]}{\rho(\mathbf{x}_i)}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{x}_i\in B_{\mathbb{S}^2}(\mathbf{p}',r)]}{\rho(\mathbf{x}_i)}\right)$$

$$\mathbb{E}\prod_{\mathbf{y}\in X}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{y}\in B_{\mathbb{S}^2}(\mathbf{p}',r)]}{\rho(\mathbf{y})}\right)\prod_{i=1}^{n}\rho(\mathbf{x}_i)\lambda_{\mathbb{S}^2}(d\mathbf{x}_i)$$

$$=\sum_{n=0}^{\infty}\frac{(-\bar{\rho})^n}{n!}\mathbb{E}\prod_{\mathbf{y}\in X}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{y}\in B_{\mathbb{S}^2}(\mathbf{p}',r)]}{\rho(\mathbf{y})}\right)$$

$$\overbrace{\int_{\mathbb{S}^2}\cdots\int_{\mathbb{S}^2}}^{n}\prod_{i=1}^{n}\frac{\mathbb{1}[\mathbf{x}_i\in B_{\mathbb{S}^2}(\mathbf{p},r)]}{\rho(\mathbf{x}_i)}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{x}_i\in B_{\mathbb{S}^2}(\mathbf{p}',r)]}{\rho(\mathbf{x}_i)}\right)\prod_{i=1}^{n}\rho(\mathbf{x}_i)\lambda_{\mathbb{S}^2}(d\mathbf{x}_i)$$

Note the expectation is just the definition of the generating functional of X which does not depend on the point $\mathbf{p}'$ (see proof of Theorem 1 by van Lieshout ((2011))),

$$=\sum_{n=0}^{\infty}\frac{(-\bar{\rho})^n}{n!}G(1-u_r^{\mathbf{y}})\overbrace{\int_{\mathbb{S}^2}\cdots\int_{\mathbb{S}^2}}^{n}\prod_{i=1}^{n}\frac{\mathbb{1}[\mathbf{x}_i\in B_{\mathbb{S}^2}(\mathbf{p},r)]}{\rho(\mathbf{x}_i)}$$

$$\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{x}_i\in B_{\mathbb{S}^2}(\mathbf{p}',r)]}{\rho(\mathbf{x}_i)}\right)\prod_{i=1}^{n}\rho(\mathbf{x}_i)\lambda_{\mathbb{S}^2}(d\mathbf{x}_i)$$

$$=\sum_{n=0}^{\infty}\frac{(-\bar{\rho})^n}{n!}G(1-u_r^{\mathbf{y}})$$

$$\overbrace{\int_{\mathbb{S}^2\cap B_{\mathbb{S}^2}(\mathbf{p},r)}\cdots\int_{\mathbb{S}^2\cap B_{\mathbb{S}^2}(\mathbf{p},r)}}^{n}\prod_{i=1}^{n}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{x}_i\in B_{\mathbb{S}^2}(\mathbf{p}',r)]}{\rho(\mathbf{x}_i)}\right)\lambda_{\mathbb{S}^2}(d\mathbf{x})_1\cdots\lambda_{\mathbb{S}^2}(d\mathbf{x})_n$$

$$=\sum_{n=0}^{\infty}\frac{(-\bar{\rho})^n}{n!}G(1-u_r^{\mathbf{y}})\left(\int_{\mathbb{S}^2\cap B_{\mathbb{S}^2}(\mathbf{p},r)}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{x}\in B_{\mathbb{S}^2}(\mathbf{p}',r)]}{\rho(\mathbf{x})}\right)\lambda_{\mathbb{S}^2}(d\mathbf{x})\right)^n$$

$$=G(1-u_r^{\mathbf{y}})\sum_{n=0}^{\infty}\frac{(-\bar{\rho})^n}{n!}\left(\lambda_{\mathbb{S}^2}(\mathbb{S}^2_{B_{\mathbb{S}^2}(\mathbf{p},r)})-\int_{\mathbb{S}^2_{B_{\mathbb{S}^2}(\mathbf{p},r)\cap B_{\mathbb{S}^2}(\mathbf{p}',r)}}\frac{\bar{\rho}}{\rho(\mathbf{x})}\lambda_{\mathbb{S}^2}(d\mathbf{x})\right)^n$$

Using the series definition for the exponential function,

$$
\begin{aligned}
&= G(1-u_r^{\mathbf{y}}) \exp\left(-\bar{\rho}\left(\lambda_{\mathbb{S}^2}(\mathbb{S}^2_{B_{\mathbb{S}^2}(\mathbf{p},r)}) - \int_{\mathbb{S}^2_{B_{\mathbb{S}^2}(\mathbf{o},r)\cap B_{\mathbb{S}^2}(\mathbf{p}',r)}} \frac{\bar{\rho}}{\rho(\mathbf{x})}\lambda_{\mathbb{S}^2}(d\mathbf{x})\right)\right) \\
&= G(1-u_r^{\mathbf{y}}) \exp\left(\bar{\rho}\int_{\mathbb{S}^2_{B_{\mathbb{S}^2}(\mathbf{p},r)\cap B_{\mathbb{S}^2}(\mathbf{p}',r)}} \frac{\bar{\rho}}{\rho(\mathbf{x})}\lambda_{\mathbb{S}^2}(d\mathbf{x}) - \bar{\rho}\lambda_{\mathbb{S}^2}(B_{\mathbb{S}^2}(\mathbf{o},r))\right) \\
&= G(1-u_r^{\mathbf{y}}) \exp\left(-\bar{\rho}\lambda_{\mathbb{S}^2}(B_{\mathbb{S}^2}(\mathbf{o},r))\right) \exp\left(\int_{\mathbb{S}^2_{B_{\mathbb{S}^2}(\mathbf{p},r)\cap B_{\mathbb{S}^2}(\mathbf{p}',r)}} \frac{\bar{\rho}^2}{\rho(\mathbf{x})}\lambda_{\mathbb{S}^2}(d\mathbf{x})\right)
\end{aligned}
$$

Substituting this into the first term of Equation B.10 gives,

$$
\begin{aligned}
&\frac{1}{|P|^2}\mathbb{E}\left(\sum_{\mathbf{p}\in P}\prod_{\mathbf{x}\in X}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{x}\in B_{\mathbb{S}^2}(\mathbf{p},r)]}{\rho(\mathbf{x})}\right)\right)^2 \\
&= G(1-u_r^{\mathbf{y}}) \exp\left(-\bar{\rho}\lambda_{\mathbb{S}^2}(B_{\mathbb{S}^2}(\mathbf{o},r))\right) \frac{1}{|P|^2}\sum_{\mathbf{p}\in P}\sum_{\mathbf{p}'\in P} \exp\left(\int_{\mathbb{S}^2_{B_{\mathbb{S}^2}(\mathbf{p},r)\cap B_{\mathbb{S}^2}(\mathbf{p}',r)}} \frac{\bar{\rho}^2}{\rho(\mathbf{x})}\lambda_{\mathbb{S}^2}(d\mathbf{x})\right)
\end{aligned}
$$

The second term of Equation B.10, by unbiasedness of $\hat{F}_{\text{inhom}}$ shown in Theorem 4.4.1, gives,

$$
\begin{aligned}
\mathbb{E}^2\frac{1}{|P|}\sum_{\mathbf{p}\in P}\prod_{\mathbf{x}\in X}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{x}\in B_{\mathbb{S}^2}(\mathbf{p},r)]}{\rho(\mathbf{x})}\right) &= \mathbb{E}^2\left[1-\hat{F}_{\text{inhom}}(r)\right] \\
&= \left(1-\mathbb{E}[\hat{F}_{\text{inhom}}(r)]\right)^2 \\
&= G^2(1-u_r^{\mathbf{y}}),
\end{aligned}
$$

where the generating functional does not depend on $\mathbf{y}$. Thus the variance is,

$$
\begin{aligned}
&G(1-u_r^{\mathbf{y}}) \exp\left(-\bar{\rho}\lambda_{\mathbb{S}^2}(B_{\mathbb{S}^2}(\mathbf{o},r))\right) \\
&\qquad \times \frac{1}{|P|^2}\sum_{\mathbf{p}\in P}\sum_{\mathbf{p}'\in P} \exp\left(\int_{\mathbb{S}^2_{B_{\mathbb{S}^2}(\mathbf{p},r)\cap B_{\mathbb{S}^2}(\mathbf{p}',r)}} \frac{\bar{\rho}^2}{\rho(\mathbf{x})}\lambda_{\mathbb{S}^2}(d\mathbf{x})\right) - G^2(1-u_r^{\mathbf{y}}),
\end{aligned}
$$

For a Poisson process, $G(1-u_r^{\mathbf{y}}) = \exp\left(-\bar{\rho}\lambda_{\mathbb{S}^2}(\mathcal{S}^2_{B_{\mathbb{S}^2}(\mathbf{o},r)})\right)$, thus the variance is,

$$\exp\left(-2\bar{\rho}\lambda_{\mathbb{S}^2}(B_{\mathbb{S}^2}(\mathbf{o},r))\right)\frac{1}{|P|^2}\sum_{\mathbf{p}\in P}\sum_{\mathbf{p}'\in P}\exp\left(\int_{\mathbb{S}^2_{B_{\mathbb{S}^2}(\mathbf{p},r)\cap B_{\mathbb{S}^2}(\mathbf{p}',r)}}\frac{\bar{\rho}^2}{\rho(\mathbf{x})}\lambda_{\mathbb{S}^2}(d\mathbf{x})\right) - \exp\left(-2\bar{\rho}\lambda_{\mathbb{S}^2}(\mathcal{S}^2_{B_{\mathbb{S}^2}(\mathbf{o},r)})\right).$$

□

Before the proof for the variance of $\hat{H}_{\text{inhom}}(r)$ we introduce the exponential integral. The exponential integral, denoted $\text{Ei}(x)$, is defined as the following integral,

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

It can then be shown that the exponential integral has the following infinite series representation,

$$\text{Ei}(x) = \gamma + \log(x) + \sum_{k=1}^{\infty}\frac{x^k}{k\cdot k!}, \tag{B.13}$$

where γ is known as the Euler-Mascheroni constant and defined as,

$$\gamma = \lim_{n\to\infty}\left(-\log(n) + \sum_{k=1}^{n}\frac{1}{k}\right).$$

The variance for $\hat{H}_{\text{inhom}}(r)$ will be given in terms of $\text{Ei}(x)$.

Proof. VARIANCE OF $\hat{H}_{\text{INHOM}}(r)$

Restating the estimator for $\hat{H}_{\text{inhom}}(r)$,

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

we note that this estimator is only well defined when $\mathbb{1}[N_X(\mathbb{S}^2) > 0]$. In the event that $N_X(\mathbb{S}^2) = 0$ we shall define $\hat{H}_{\text{inhom}}(r) = 0$, in which case we can write our estimator as,

$$\hat{H}_{\text{inhom}}(r) = \mathbb{1}[N_X(\mathbb{S}^2) > 0]\left(1 - \frac{\sum_{\mathbf{x}\in X}\prod_{\mathbf{y}\in X\setminus\{\mathbf{x}\}}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{y}\in B_{\mathbb{S}^2}(\mathbf{x},r)]}{\rho(\mathbf{y})}\right)}{N_X(\mathbb{S}^2)}\right).$$

By using the law of total variance, conditioning on $N_X(\mathbb{S}^2) = n$,

$$\mathrm{Var}(\hat{H}_{\mathrm{inhom}}(r)) = \underbrace{\mathbb{E}[\mathrm{Var}(\hat{H}_{\mathrm{inhom}}(r)|N_X(\mathbb{S}^2 = n)]}_{(1)} + \underbrace{\mathrm{Var}(\mathbb{E}[\hat{H}_{\mathrm{inhom}}(r)|N_X(\mathbb{S}^2) = n])}_{(2)}. \quad \text{(B.14)}$$

The variance in term (1) is,

$$\begin{aligned}
\mathrm{Var}(\hat{H}_{\mathrm{inhom}}(r)|N_X(\mathbb{S}^2 = n) &= \mathbb{1}[n > 0]\mathrm{Var}\left(1 + \frac{1}{n}\sum_{i=1}^{n}\prod_{\substack{j=1\\ j\neq i}}^{n}\left(1 - \frac{\bar{\rho}\mathbb{1}[\mathbf{X}_j \in B(\mathbf{X}_i, r)]}{\rho(\mathbf{X}_j)}\right)\right), \\
&= \frac{\mathbb{1}[n > 0]}{n^2}\mathrm{Var}\left(\sum_{i=1}^{n}\prod_{\substack{j=1\\ j\neq i}}^{n}\left(1 - \frac{\bar{\rho}\mathbb{1}[\mathbf{X}_j \in B(\mathbf{X}_i, r)]}{\rho(\mathbf{X}_j)}\right)\right),
\end{aligned}$$

where $\mathbf{X}_k,\ k = 1, \ldots, n$ are independently distributed points with density proportional to $\rho(\mathbf{x}_k)$, by definition of a Poisson process. We use the identity $\mathrm{Var}(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} \sum_{j=1}^{n} \mathrm{Cov}(X_i, X_j)$ and define $L_i = \prod_{\substack{j=1\\ j\neq i}}^{n}\left(1 - \frac{\bar{\rho}\mathbb{1}[\mathbf{X}_j \in B(\mathbf{X}_i, r)]}{\rho(\mathbf{X}_j)}\right)$,

$$\begin{aligned}
&= \frac{\mathbb{1}[n > 0]}{n^2}\sum_{p=1}^{n}\sum_{q=1}^{n}\mathrm{Cov}\left(L_p, L_q\right) \\
&= \frac{\mathbb{1}[n > 0]}{n^2}\sum_{p=1}^{n}\sum_{q=1}^{n}\Big(\mathbb{E}[L_p L_q] - \mathbb{E}[L_p]\mathbb{E}[L_q]\Big) \\
&= \frac{\mathbb{1}[n > 0]}{n^2}\sum_{p,q\in\{1,\ldots,n\}}^{\neq}\left(\underbrace{\mathbb{E}[L_p L_q]}_{(a)} - \underbrace{\mathbb{E}[L_p]\mathbb{E}[L_q]}_{(b)}\right) \\
&\qquad + \frac{\mathbb{1}[n > 0]}{n^2}\sum_{p=1}^{n}\left(\underbrace{\mathbb{E}[L_p^2]}_{(c)} - \underbrace{\mathbb{E}^2[L_p]}_{(d)}\right) \qquad \text{(B.15)}
\end{aligned}$$

Taking term (a),

$$\mathbb{E}[L_p L_q] = \mathbb{E}\left[\prod_{\substack{j=1\\ j\neq p}}^{n}\left(1 - \frac{\bar{\rho}\mathbb{1}[\mathbf{X}_j \in B_{\mathbb{S}^2}(\mathbf{X}_p, r)]}{\rho(\mathbf{X}_j)}\right)\prod_{\substack{j=1\\ j\neq q}}^{n}\left(1 - \frac{\bar{\rho}\mathbb{1}[\mathbf{X}_j \in B_{\mathbb{S}^2}(\mathbf{X}_q, r)]}{\rho(\mathbf{X}_j)}\right)\right]$$

using iterated expectation, conditioning on $\mathbf{X_p} = \mathbf{x}_p, \mathbf{X_q} = \mathbf{x}_q$,

$$
\begin{aligned}
&= \mathbb{E}\left[\mathbb{E}\left[\prod_{\substack{j=1\\ j\neq p}}^{n}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{X}_j\in B_{\mathbb{S}^2}(\mathbf{X}_p,r)]}{\rho(\mathbf{X}_j)}\right)\right.\right.\\
&\qquad\left.\left.\prod_{\substack{j=1\\ j\neq q}}^{n}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{X}_j\in B_{\mathbb{S}^2}(\mathbf{X}_q,r)]}{\rho(\mathbf{X}_j)}\right)\,\middle|\,\mathbf{X_p}=\mathbf{x}_p,\mathbf{X_q}=\mathbf{x}_q\right]\right]\\
&= \mathbb{E}\left[\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{X}_p\in B_{\mathbb{S}^2}(\mathbf{X}_q,r)]}{\rho(\mathbf{X}_p)}\right)\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{X}_q\in B_{\mathbb{S}^2}(\mathbf{X}_p,r)]}{\rho(\mathbf{X}_q)}\right)\right.\\
&\times\left.\mathbb{E}\left[\prod_{\substack{j=1\\ j\neq p,q}}^{n}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{X}_j\in B_{\mathbb{S}^2}(\mathbf{X}_p,r)]}{\rho(\mathbf{X}_j)}\right)\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{X}_j\in B_{\mathbb{S}^2}(\mathbf{X}_q,r)]}{\rho(\mathbf{X}_j)}\right)\,\middle|\,\mathbf{X_p}=\mathbf{x}_p,\mathbf{X_q}=\mathbf{x}_q\right]\right]
\end{aligned}
$$

The expectation conditioned on $(\mathbf{X}_q, \mathbf{X}_p)$ is,

$$
\begin{aligned}
&\mathbb{E}\left[\prod_{\substack{j=1\\ j\neq p,q}}^{n}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{X}_j\in B_{\mathbb{S}^2}(\mathbf{X}_p,r)]}{\rho(\mathbf{X}_j)}\right)\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{X}_j\in B_{\mathbb{S}^2}(\mathbf{X}_q,r)]}{\rho(\mathbf{X}_j)}\right)\,\middle|\,\mathbf{X_p}=\mathbf{x}_p,\mathbf{X_q}=\mathbf{x}_q\right]\\
&\quad=\overbrace{\int_{\mathbb{S}^2}\cdots\int_{\mathbb{S}^2}}^{n-2}\prod_{\substack{j=1\\ j\neq p,q}}^{n}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{x}_j\in B_{\mathbb{S}^2}(\mathbf{x}_p,r)]}{\rho(\mathbf{x}_j)}\right)\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{x}_j\in B_{\mathbb{S}^2}(\mathbf{x}_q,r)]}{\rho(\mathbf{x}_j)}\right)\frac{\rho(\mathbf{x}_j)}{\mu(\mathbb{S}^2)}\lambda_{\mathbb{S}^2}(d\mathbf{x}_j)\\
&\quad=\left(\int_{\mathbb{S}^2}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{x}\in B_{\mathbb{S}^2}(\mathbf{x}_p,r)]}{\rho(\mathbf{x})}\right)\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{x}\in B_{\mathbb{S}^2}(\mathbf{x}_q,r)]}{\rho(\mathbf{x})}\right)\frac{\rho(\mathbf{x})}{\mu(\mathbb{S}^2)}\lambda_{\mathbb{S}^2}(d\mathbf{x})\right)^{n-2}\\
&\quad=\left(\int_{\mathbb{S}^2}\frac{\rho(\mathbf{x})}{\mu(\mathbb{S}^2)}-\frac{\bar{\rho}}{\mu(\mathbb{S}^2)}\mathbb{1}[\mathbf{x}\in B_{\mathbb{S}^2}(\mathbf{x}_p,r)]-\frac{\bar{\rho}}{\mu(\mathbb{S}^2)}\mathbb{1}[\mathbf{x}\in B_{\mathbb{S}^2}(\mathbf{x}_q,r)]\right.\\
&\qquad\left.+\frac{\bar{\rho}^2}{\mu(\mathbb{S}^2)}\frac{\mathbb{1}[\mathbf{x}\in B_{\mathbb{S}^2}(\mathbf{x}_p,r),\mathbf{x}\in B_{\mathbb{S}^2}(\mathbf{x}_q,r)]}{\rho(\mathbf{x})}\lambda_{\mathbb{S}^2}(d\mathbf{x})\right)^{n-2}\\
&\quad=\left(1-\frac{2\bar{\rho}}{\mu(\mathbb{S}^2)}2\pi(1-\cos r)+\frac{\bar{\rho}^2}{\mu(\mathbb{S}^2)}\int_{B_{\mathbb{S}^2}(\mathbf{x}_p,r)\cap B_{\mathbb{S}^2}(\mathbf{x}_q,r)}\frac{1}{\rho(\mathbf{x})}\lambda_{\mathbb{S}^2}(d\mathbf{x})\right)^{n-2}
\end{aligned}
$$

We then define $A_1(\mathbf{x}_p, \mathbf{x}_q) \equiv 1 - \frac{2\bar{\rho}}{\mu(\mathbb{S}^2)} 2\pi(1-\cos r) + \frac{\bar{\rho}^2}{\mu(\mathbb{S}^2)} \int_{B_{\mathbb{S}^2}(\mathbf{x}_p,r)\cap B_{\mathbb{S}^2}(\mathbf{x}_q,r)} \frac{1}{\rho(\mathbf{x})} \lambda_{\mathbb{S}^2}(d\mathbf{x})$ and returning to $\mathbb{E}[L_p L_q]$,

$$\begin{aligned}
\mathbb{E}[L_p L_q] &= \mathbb{E}\left[\left(1 - \frac{\bar{\rho}\mathbb{1}[\mathbf{X}_p \in B_{\mathbb{S}^2}(\mathbf{X}_q, r)]}{\rho(\mathbf{X}_p)}\right)\left(1 - \frac{\bar{\rho}\mathbb{1}[\mathbf{X}_q \in B_{\mathbb{S}^2}(\mathbf{X}_p, r)]}{\rho(\mathbf{X}_q)}\right) A_1^{n-2}(\mathbf{X}_p, \mathbf{X}_q)\right] \\
&= \int_{\mathbb{S}^2}\int_{\mathbb{S}^2}\left(1 - \frac{\bar{\rho}\mathbb{1}[\mathbf{x}_p \in B_{\mathbb{S}^2}(\mathbf{x}_q, r)]}{\rho(\mathbf{x}_p)}\right)\left(1 - \frac{\bar{\rho}\mathbb{1}[\mathbf{x}_q \in B_{\mathbb{S}^2}(\mathbf{x}_p, r)]}{\rho(\mathbf{x}_q)}\right) \\
&\qquad A_1^{n-2}(\mathbf{x}_p, \mathbf{x}_q)\frac{\rho(\mathbf{x}_p)\rho(\mathbf{x}_p)}{\mu^2(\mathbb{S}^2)}\lambda_{\mathbb{S}^2}(d\mathbf{x}_p)\lambda_{\mathbb{S}^2}(d\mathbf{x}_q) \\
&= \frac{1}{\mu^2(\mathbb{S}^2)}\int_{\mathbb{S}^2}\int_{\mathbb{S}^2}\left(\rho(\mathbf{x}_p) - \bar{\rho}\mathbb{1}[\mathbf{x}_p \in B_{\mathbb{S}^2}(\mathbf{x}_q, r)]\right)\left(\rho(\mathbf{x}_q) - \bar{\rho}\mathbb{1}[\mathbf{x}_q \in B_{\mathbb{S}^2}(\mathbf{x}_p, r)]\right) \\
&\qquad A_1^{n-2}(\mathbf{x}_p, \mathbf{x}_q)\lambda_{\mathbb{S}^2}(d\mathbf{x}_p)\lambda_{\mathbb{S}^2}(d\mathbf{x}_q) \\
&= \frac{1}{\mu^2(\mathbb{S}^2)}\int_{\mathbb{S}^2}\int_{\mathbb{S}^2}\left(\rho(\mathbf{x}) - \bar{\rho}\mathbb{1}[\mathbf{x} \in B_{\mathbb{S}^2}(\mathbf{y}, r)]\right)\left(\rho(\mathbf{y}) - \bar{\rho}\mathbb{1}[\mathbf{y} \in B_{\mathbb{S}^2}(\mathbf{x}, r)]\right) \\
&\qquad A_1^{n-2}(\mathbf{x}, \mathbf{y})\lambda_{\mathbb{S}^2}(d\mathbf{x})\lambda_{\mathbb{S}^2}(d\mathbf{y})
\end{aligned}$$

Then taking term (b) of Equation B.15 and examining one of the expectations,

$$\mathbb{E}[L_p] = \mathbb{E}\left[\prod_{\substack{j=1\\ j\neq p}}^{n}\left(1 - \frac{\bar{\rho}\mathbb{1}[\mathbf{X}_j \in B_{\mathbb{S}^2}(\mathbf{X}_p, r)]}{\rho(\mathbf{X}_j)}\right)\right]$$

using iterated expectation conditioning on $\mathbf{X}_p = \mathbf{x}_p$,

$$\begin{aligned}
&= \mathbb{E}\left[\mathbb{E}\left[\prod_{\substack{j=1\\ j\neq p}}^{n}\left(1 - \frac{\bar{\rho}\mathbb{1}[\mathbf{X}_j \in B_{\mathbb{S}^2}(\mathbf{X}_p, r)]}{\rho(\mathbf{X}_j)}\right)\middle|\mathbf{X}_p = \mathbf{x}_p\right]\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[\prod_{\substack{j=1\\ j\neq p}}^{n}\left(1 - \frac{\bar{\rho}\mathbb{1}[\mathbf{X}_j \in B_{\mathbb{S}^2}(\mathbf{x}_p, r)]}{\rho(\mathbf{X}_j)}\right)\right]\right].
\end{aligned}$$

Taking the condition expectation,

$$\mathbb{E}\left[\prod_{\substack{j=1\\ j\neq p}}^{n}\left(1 - \frac{\bar{\rho}\mathbb{1}[\mathbf{X}_j \in B_{\mathbb{S}^2}(\mathbf{x}_p, r)]}{\rho(\mathbf{X}_j)}\right)\right] = \overbrace{\int_{\mathbb{S}^2}\cdots\int_{\mathbb{S}^2}}^{n-1}\prod_{\substack{j=1\\ j\neq p}}^{n}\left(1 - \frac{\bar{\rho}\mathbb{1}[\mathbf{x}_j \in B_{\mathbb{S}^2}(\mathbf{x}_p, r)]}{\rho(\mathbf{x}_j)}\right)\frac{\rho(\mathbf{x}_j)}{\mu(\mathbb{S}^2)}\lambda_{\mathbb{S}^2}(d\mathbf{x}_j)$$

$$= \left(\int_{\mathbb{S}^2} \left(1 - \frac{\bar{\rho} \mathbb{1}[\mathbf{x} \in B_{\mathbb{S}^2}(\mathbf{x}_p, r)]}{\rho(\mathbf{x})} \right) \frac{\rho(\mathbf{x})}{\mu(\mathbb{S}^2)} \lambda_{\mathbb{S}^2}(d\mathbf{x}) \right)^{n-1}$$
$$= \left(1 - \frac{\bar{\rho}}{\mu(\mathbb{S}^2)} 2\pi(1 - \cos r) \right)^{n-1}.$$

Hence,

$$\mathbb{E}[L_p] = \left(1 - \frac{\bar{\rho}}{\mu(\mathbb{S}^2)} 2\pi(1 - \cos r) \right)^{n-1}$$

Next we will deal with term (c).

$$\mathbb{E}[L_p^2] = \mathbb{E}\left[\left(\prod_{\substack{j=1 \\ j \neq p}}^{n} \left(1 - \frac{\bar{\rho} \mathbb{1}[\mathbf{X}_j \in B_{\mathbb{S}^2}(\mathbf{X}_p, r)]}{\rho(\mathbf{X}_j)} \right) \right)^2 \right]$$
$$= \mathbb{E}\left[\prod_{\substack{j=1 \\ j \neq p}}^{n} \left(1 - \frac{\bar{\rho} \mathbb{1}[\mathbf{X}_j \in B_{\mathbb{S}^2}(\mathbf{X}_p, r)]}{\rho(\mathbf{X}_j)} \right)^2 \right]$$
$$= \mathbb{E}\left[\mathbb{E}\left[\prod_{\substack{j=1 \\ j \neq p}}^{n} \left(1 - \frac{\bar{\rho} \mathbb{1}[\mathbf{X}_j \in B_{\mathbb{S}^2}(\mathbf{x}_p, r)]}{\rho(\mathbf{X}_j)} \right)^2 \middle| \mathbf{X}_p = \mathbf{x}_p \right] \right]$$
$$= \mathbb{E}\left[\prod_{\substack{j=1 \\ j \neq p}}^{n} \mathbb{E}\left[\left(1 - \frac{\bar{\rho} \mathbb{1}[\mathbf{X}_j \in B_{\mathbb{S}^2}(\mathbf{x}_p, r)]}{\rho(\mathbf{X}_j)} \right)^2 \middle| \mathbf{X}_p = \mathbf{x}_p \right] \right]$$
$$= \mathbb{E}\left[\prod_{\substack{j=1 \\ j \neq p}}^{n} \int_{\mathbb{S}^2} \left(1 - \frac{\bar{\rho} \mathbb{1}[\mathbf{x} \in B_{\mathbb{S}^2}(\mathbf{x}_p, r)]}{\rho(\mathbf{x})} \right)^2 \frac{\rho(\mathbf{x})}{\mu(\mathbb{S}^2)} \lambda_{\mathbb{S}^2}(d\mathbf{x}) \right]$$
$$= \mathbb{E}\left[\prod_{\substack{j=1 \\ j \neq p}}^{n} \int_{\mathbb{S}^2} \frac{\rho(\mathbf{x})}{\mu(\mathbb{S}^2)} - \frac{2\bar{\rho}}{\mu(\mathbb{S}^2)} \mathbb{1}[\mathbf{x} \in B_{\mathbb{S}^2}(\mathbf{x}_q, r)] + \frac{\bar{\rho}^2}{\mu(\mathbb{S}^2)} \frac{\mathbb{1}[\mathbf{x} \in B_{\mathbb{S}^2}(\mathbf{x}_p, r)]}{\rho(\mathbf{x})} \lambda_{\mathbb{S}^2}(d\mathbf{x}) \right]$$
$$= \mathbb{E}\left[\prod_{\substack{j=1 \\ j \neq p}}^{n} \left(1 - \frac{2\bar{\rho}}{\mu(\mathbb{S}^2)} 2\pi(1 - \cos r) + \frac{\bar{\rho}^2}{\mu(\mathbb{S}^2)} \int_{\mathbb{S}^2 \cap B_{\mathbb{S}^2}(\mathbf{x}_p, r)} \frac{1}{\rho(\mathbf{x})} \lambda_{\mathbb{S}^2}(d\mathbf{x}) \right) \right]$$
$$= \mathbb{E}\left[\left(1 - \frac{2\bar{\rho}}{\mu(\mathbb{S}^2)} 2\pi(1 - \cos r) + \frac{\bar{\rho}^2}{\mu(\mathbb{S}^2)} \int_{\mathbb{S}^2 \cap B_{\mathbb{S}^2}(\mathbf{X}_p, r)} \frac{1}{\rho(\mathbf{x})} \lambda_{\mathbb{S}^2}(d\mathbf{x}) \right)^{n-1} \right]$$

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

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

$$= \frac{1}{\mu(\mathbb{S}^2)} \int_{\mathbb{S}^2} A_2^{n-1}(\mathbf{y}) \rho(\mathbf{y}) \lambda_{\mathbb{S}^2}(d\mathbf{y})$$

The final term (d) is identical to that of (b). So plugging into Equation B.15 gives,

$$\begin{aligned}
&\mathrm{Var}(\hat{H}_{\mathrm{inhom}}(r) | N_X(\mathbb{S}^2 = n) = \frac{\mathbb{1}[n>0]}{n^2} \sum_{p,q \in \{1,\ldots,n\}}^{\neq} \\
&\left(\frac{1}{\mu^2(\mathbb{S}^2)} \int_{\mathbb{S}^2} \int_{\mathbb{S}^2} \left(\rho(\mathbf{x}) - \bar{\rho}\mathbb{1}[\mathbf{x} \in B_{\mathbb{S}^2}(\mathbf{y},r)]\right) \left(\rho(\mathbf{y}) - \bar{\rho}\mathbb{1}[\mathbf{y} \in B_{\mathbb{S}^2}(\mathbf{x},r)]\right) A_1^{n-2}(\mathbf{x},\mathbf{y}) \lambda_{\mathbb{S}^2}(d\mathbf{x}) \lambda_{\mathbb{S}^2}(d\mathbf{y}) \right. \\
&\qquad - \left(1 - \frac{\bar{\rho}}{\mu(\mathbb{S}^2)} 2\pi(1-\cos r)\right)^{n-1} \left(1 - \frac{\bar{\rho}}{\mu(\mathbb{S}^2)} 2\pi(1-\cos r)\right)^{n-1} \Bigg) \\
&\qquad + \frac{\mathbb{1}[n>0]}{n^2} \sum_{p=1}^{n} \left(\frac{1}{\mu(\mathbb{S}^2)} \int_{\mathbb{S}^2} A_2^{n-1}(\mathbf{y}) \rho(\mathbf{y}) \lambda_{\mathbb{S}^2}(d\mathbf{y}) - \left(1 - \frac{\bar{\rho}}{\mu(\mathbb{S}^2)} 2\pi(1-\cos r)\right)^{2n-2} \right) \\
&= \frac{1}{\mu^2(\mathbb{S}^2)} \\
&\quad \int_{\mathbb{S}^2} \int_{\mathbb{S}^2} \left(\rho(\mathbf{x}) - \bar{\rho}\mathbb{1}[\mathbf{x} \in B_{\mathbb{S}^2}(\mathbf{y},r)]\right) \left(\rho(\mathbf{y}) - \bar{\rho}\mathbb{1}[\mathbf{y} \in B_{\mathbb{S}^2}(\mathbf{x},r)]\right) \\
&\qquad \left(\frac{\mathbb{1}[n>0](n-1)}{n} A_1^{n-2}(\mathbf{x},\mathbf{y}) \right) \lambda_{\mathbb{S}^2}(d\mathbf{x}) \lambda_{\mathbb{S}^2}(d\mathbf{y}) \\
&\qquad + \frac{1}{\mu(\mathbb{S}^2)} \int_{\mathbb{S}^2} \left(\frac{\mathbb{1}[n>0]}{n} A_2^{n-1}(\mathbf{y}) \right) \rho(\mathbf{y}) \lambda_{\mathbb{S}^2}(d\mathbf{y}) - \mathbb{1}[n>0] \left(1 - \frac{\bar{\rho}}{\mu(\mathbb{S}^2)} 2\pi(1-\cos r)\right)^{2n-2}
\end{aligned}$$

We need to then take the expectation of this variance over $N_X(\mathbb{S}^2)$. By application of Tonelli's Theorem we can interchange the expectation over $N_X(\mathbb{S}^2)$ with the integrals over $\mathbf{x}$ and $\mathbf{y}$. This comes by showing that both $A_1(\mathbf{x},\mathbf{y})$ and $A_2(\mathbf{x})$ are non-negative for all $\mathbf{x},\mathbf{y} \in \mathbb{S}^2$. Obviously $\rho(\mathbf{x}) - \bar{\rho}\mathbb{1}[\mathbf{x} \in B_{\mathbb{S}^2}(\mathbf{y},r)]$ and $\rho(\mathbf{y}) - \bar{\rho}\mathbb{1}[\mathbf{y} \in B_{\mathbb{S}^2}(\mathbf{x},r)]$ are greater than or equal to 0 since $\bar{\rho} = \inf_{\mathbf{x} \in \mathbb{S}^2} \rho(\mathbf{x})$. It then follows since the integrand of $A_1(\mathbf{x}_p,\mathbf{x}_q)$, $(1 - \bar{\rho}\mathbb{1}[\mathbf{x} \in B_{\mathbb{S}^2}(\mathbf{x}_p,r)]/\rho(\mathbf{x}))\,(1 - \bar{\rho}\mathbb{1}[\mathbf{x} \in B_{\mathbb{S}^2}(\mathbf{x}_q,r)]/\rho(\mathbf{x}))\,\rho(\mathbf{x})/\mu(\mathbb{S}^2)$, is then non-negative for all $\mathbf{x}$ and so the integral over $\mathbf{f}$ is non-negative and thus Tonelli's Theorem can be applied. A near identitical argument can be applied to $A_2(\mathbf{x})$ to show that it is always non-negative and hence Tonelli's Theorem can then be applied. We then calculate the following expectations,

a) $\mathbb{E}\left[\frac{\mathbb{1}[N_X(\mathbb{S}^2)>0](N_X(\mathbb{S}^2)-1)}{N_X(\mathbb{S}^2)}A_1^{N_X(\mathbb{S}^2)-2}(\mathbf{x},\mathbf{y})\right]$

b) $\mathbb{E}\left[\frac{\mathbb{1}[N_X(\mathbb{S}^2)>0]}{N_X(\mathbb{S}^2)}A_2^{N_X(\mathbb{S}^2)-1}(\mathbf{x})\right]$

c) $\mathbb{E}\left[\mathbb{1}[N_X(\mathbb{S}^2)>0]\left(1-\frac{\bar{\rho}}{\mu(\mathbb{S}^2)}2\pi(1-\cos r)\right)^{2N_X(\mathbb{S}^2)-2}\right]$

Setting $\Lambda=\mu(\mathbb{S}^2)$, expectation (a) is,

$$\begin{aligned}
&\mathbb{E}\left[\frac{\mathbb{1}[N_X(\mathbb{S}^2)>0](N_X(\mathbb{S}^2)-1)}{N_X(\mathbb{S}^2)}A_1^{N_X(\mathbb{S}^2)-2}(\mathbf{x},\mathbf{y})\right]\\
&\quad=\mathbb{E}\left[\mathbb{1}[N_X(\mathbb{S}^2)>0]A_1^{N_X(\mathbb{S}^2)-2}(\mathbf{x},\mathbf{y})\right]-\mathbb{E}\left[\frac{\mathbb{1}[N_X(\mathbb{S}^2)>0]}{N_X(\mathbb{S}^2)}A_1^{N_X(\mathbb{S}^2)-2}(\mathbf{x},\mathbf{y})\right]\\
&\quad=\sum_{n=0}^{\infty}\mathbb{1}[n>0]A_1^{n-2}(\mathbf{x},\mathbf{y})\frac{\Lambda^n\mathrm{e}^{-\Lambda}}{n!}-\sum_{n=0}^{\infty}\frac{\mathbb{1}[n>0]}{n}A_1^{n-2}(\mathbf{x},\mathbf{y})\frac{\Lambda^n\mathrm{e}^{-\Lambda}}{n!}\\
&\quad=\frac{\mathrm{e}^{-\Lambda}}{A_1^2(\mathbf{x},\mathbf{y})}\sum_{n=1}^{\infty}\frac{(\Lambda A_1(\mathbf{x},\mathbf{y}))^n}{n!}-\frac{\mathrm{e}^{-\Lambda}}{A_1^2(\mathbf{x},\mathbf{y})}\sum_{n=1}^{\infty}\frac{1}{n}\frac{(\Lambda A_1(\mathbf{x},\mathbf{y}))^n}{n!}\\
&\quad=\frac{\mathrm{e}^{-\Lambda}}{A_1^2(\mathbf{x},\mathbf{y})}\left(\sum_{n=0}^{\infty}\frac{(\Lambda A_1(\mathbf{x},\mathbf{y}))^n}{n!}-1\right)-\frac{\mathrm{e}^{-\Lambda}}{A_1^2(\mathbf{x},\mathbf{y})}(\mathrm{Ei}(\Lambda A_1(\mathbf{x},\mathbf{y}))-\gamma-\log(\Lambda A_1(\mathbf{x},\mathbf{y})))\\
&\quad=\frac{\mathrm{e}^{-\Lambda}}{A_1^2(\mathbf{x},\mathbf{y})}\left(\mathrm{e}^{\Lambda A_1(\mathbf{x},\mathbf{y})}-1-\mathrm{Ei}(\Lambda A_1(\mathbf{x},\mathbf{y}))+\gamma+\log(\Lambda A_1(\mathbf{x},\mathbf{y}))\right)
\end{aligned}$$

where the penultimate line follows from Equation B.13. Similarly for (b) and (c),

$$\begin{aligned}
\mathbb{E}\left[\frac{\mathbb{1}[N_X(\mathbb{S}^2)>0]}{N_X(\mathbb{S}^2)}A_2^{N_X(\mathbb{S}^2)-1}(\mathbf{x})\right]&=\frac{\mathrm{e}^{-\Lambda}}{A_2(\mathbf{x})}\left(\gamma+\log(\Lambda A_2(\mathbf{x}))-\mathrm{Ei}(\Lambda A_2(\mathbf{x}))\right)\\
\mathbb{E}\left[\mathbb{1}[N_X(\mathbb{S}^2)>0]C^{2N_X(\mathbb{S}^2)-2}\right]&=\frac{\mathrm{e}^{-\Lambda}}{C^2}\left(\mathrm{e}^{\Lambda C^2}-1\right),
\end{aligned}$$

where $C=\left(1-\frac{\bar{\rho}}{\mu(\mathbb{S}^2)}2\pi(1-\cos r)\right)$. Then taking expectations of $\mathrm{Var}(\hat{H}_{\mathrm{inhom}}(r)|N_X(\mathbb{S}^2)=n)$ over $N_X(\mathbb{S}^2)$ gives,

$$\begin{aligned}
&\mathbb{E}[\mathrm{Var}(\hat{H}_{\mathrm{inhom}}(r)|N_X(\mathbb{S}^2)=n)]\\
&=\frac{1}{\mu^2(\mathbb{S}^2)}\\
&\quad\int_{\mathbb{S}^2}\int_{\mathbb{S}^2}\left(\rho(\mathbf{x})-\bar{\rho}\mathbb{1}[\mathbf{x}\in B_{\mathbb{S}^2}(\mathbf{y},r)]\right)\left(\rho(\mathbf{y})-\bar{\rho}\mathbb{1}[\mathbf{y}\in B_{\mathbb{S}^2}(\mathbf{x},r)]\right)\\
&\qquad\qquad\frac{\mathrm{e}^{-\Lambda}}{A_1^2(\mathbf{x},\mathbf{y})}\left(\mathrm{e}^{\Lambda A_1(\mathbf{x},\mathbf{y})}-1-\mathrm{Ei}(\Lambda A_1(\mathbf{x},\mathbf{y}))+\gamma+\log(\Lambda A_1(\mathbf{x},\mathbf{y}))\right)\lambda_{\mathbb{S}^2}(d\mathbf{x})\lambda_{\mathbb{S}^2}(d\mathbf{y})
\end{aligned}$$

$$+\frac{1}{\mu(\mathbb{S}^2)}\int_{\mathbb{S}^2}\frac{\mathrm{e}^{-\Lambda}}{A_2(\mathbf{x})}\left(\gamma+\log(\Lambda A_2(\mathbf{x}))-\mathrm{Ei}(\Lambda A_2(\mathbf{x}))\right)\rho(\mathbf{y})\lambda_{\mathbb{S}^2}(d\mathbf{y})$$
$$-\frac{\mathrm{e}^{-\Lambda}}{\left(1-\frac{\bar{\rho}}{\mu(\mathbb{S}^2)}2\pi(1-\cos r)\right)^2}\left(\mathrm{e}^{\Lambda\left(1-\frac{\bar{\rho}}{\mu(\mathbb{S}^2)}2\pi(1-\cos r)\right)^2}-1\right)$$

Term (2) of Equation B.14, the expectation conditioned on $N_X(\mathbb{S}^2)=n$,

$$\begin{aligned}
&\mathbb{E}[\hat{H}_{\mathrm{inhom}}(r)|N_X(\mathbb{S}^2)=n]=\\
&\qquad\mathbb{E}\left[1-\frac{\mathbb{1}[N_X(\mathbb{S}^2)>0]}{N_X(\mathbb{S}^2)}\sum_{\mathbf{x}\in X}\prod_{\mathbf{y}\in X\setminus\{\mathbf{x}\}}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{y}\in B_{\mathbb{S}^2}(\mathbf{x},r)]}{\rho(\mathbf{y})}\right)\middle|N_X(\mathbb{S}^2)=n\right]\\
&\qquad=1-\frac{\mathbb{1}[n>0]}{n}\sum_{i=1}^{n}\mathbb{E}\left[\prod_{\substack{j=1\\j\neq i}}^{n}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{x}_j\in B(\mathbf{x}_i,r)]}{\rho(\mathbf{x}_j)}\right)\right]\\
&\qquad=1-\frac{\mathbb{1}[n>0]}{n}\sum_{i=1}^{n}\mathbb{E}\left[L_i\right],
\end{aligned}$$

where L_i is as defined in Equation B.15,

$$\begin{aligned}
&=1-\frac{\mathbb{1}[n>0]}{n}\sum_{i=1}^{n}\left(1-\frac{\bar{\rho}}{\mu(\mathbb{S}^2)}2\pi(1-\cos r)\right)^{n-1}\\
&=1-\mathbb{1}[n>0]\left(1-\frac{\bar{\rho}}{\mu(\mathbb{S}^2)}2\pi(1-\cos r)\right)^{n-1}\\
&=1-\mathbb{1}[n>0]C^{n-1},
\end{aligned}$$

where $C=1-\frac{\bar{\rho}}{\mu(\mathbb{S}^2)}2\pi(1-\cos r)$. Then taking the variance over $N_X(\mathbb{S}^2)$ gives,

$$\begin{aligned}
&\mathrm{Var}(1-\mathbb{1}[N_X(\mathbb{S}^2)>0]C^{N_X(\mathbb{S}^2)-1})\\
&\qquad=\mathrm{Var}(\mathbb{1}[N_X(\mathbb{S}^2)>0]C^{N_X(\mathbb{S}^2)-1})\\
&\qquad=\mathbb{E}\left[\left(\mathbb{1}[N_X(\mathbb{S}^2)>0]C^{N_X(\mathbb{S}^2)-1}\right)^2\right]-\mathbb{E}^2\left[\mathbb{1}[N_X(\mathbb{S}^2)>0]C^{N_X(\mathbb{S}^2)-1}\right]\\
&\qquad=\mathbb{E}\left[\mathbb{1}[N_X(\mathbb{S}^2)>0]C^{2N_X(\mathbb{S}^2)-2}\right]-\mathbb{E}^2\left[\mathbb{1}[N_X(\mathbb{S}^2)>0]C^{N_X(\mathbb{S}^2)-1}\right]\\
&\qquad=\frac{\mathrm{e}^{-\Lambda}}{C^2}\left(\mathrm{e}^{\Lambda C^2}-1\right)-\frac{\mathrm{e}^{-2\Lambda}}{C^2}\left(\mathrm{e}^{\Lambda C}-1\right)^2,
\end{aligned}$$

where $\Lambda = \mu(\mathbb{S}^2)$. Therefore the variance of $\hat{H}_{\text{inhom}}(r)$ is,

$$
\begin{aligned}
&\operatorname{Var}(\hat{H}_{\text{inhom}}(r)) \\
&= \frac{1}{\mu^2(\mathbb{S}^2)} \int_{\mathbb{S}^2} \int_{\mathbb{S}^2} \left(\rho(\mathbf{x}) - \bar{\rho}\mathbb{1}[\mathbf{x} \in B_{\mathbb{S}^2}(\mathbf{y}, r)]\right) \left(\rho(\mathbf{y}) - \bar{\rho}\mathbb{1}[\mathbf{y} \in B_{\mathbb{S}^2}(\mathbf{x}, r)]\right) \\
&\frac{\mathrm{e}^{-\mu(\mathbb{S}^2)}}{A_1^2(\mathbf{x}, \mathbf{y})} \left(\mathrm{e}^{\mu(\mathbb{S}^2)A_1(\mathbf{x},\mathbf{y})} - 1 - \operatorname{Ei}(\mu(\mathbb{S}^2)A_1(\mathbf{x}, \mathbf{y})) + \gamma + \log(\mu(\mathbb{S}^2)A_1(\mathbf{x}, \mathbf{y}))\right) \lambda_{\mathbb{S}^2}(d\mathbf{x})\lambda_{\mathbb{S}^2}(d\mathbf{y}) \\
&\quad + \frac{1}{\mu(\mathbb{S}^2)} \int_{\mathbb{S}^2} \frac{\mathrm{e}^{-\mu(\mathbb{S}^2)}}{A_2(\mathbf{x})} \left(\gamma + \log(\mu(\mathbb{S}^2)A_2(\mathbf{x})) - \operatorname{Ei}(\mu(\mathbb{S}^2)A_2(\mathbf{x}))\right) \rho(\mathbf{y})\lambda_{\mathbb{S}^2}(d\mathbf{y}) \\
&\quad - \frac{\mathrm{e}^{-2\mu(\mathbb{S}^2)}}{\left(1 - \frac{\bar{\rho}}{\mu(\mathbb{S}^2)} 2\pi(1 - \cos r)\right)^2} \left(\mathrm{e}^{\mu(\mathbb{S}^2)\left(1 - \frac{\bar{\rho}}{\mu(\mathbb{S}^2)} 2\pi(1-\cos r)\right)} - 1\right)^2
\end{aligned}
\tag{B.16}
$$

The final part of the proof is to ensure that the integrands are truly Lebesgue integrable, that is the integrals given in the previous equation are finite. We shall work with the second, $\int_{\mathbb{S}^2} (\mathrm{e}^{-\mu(\mathbb{S}^2)}/A_2(\mathbf{x})) \left(\gamma + \log(\mu(\mathbb{S}^2)A_2(\mathbf{x})) - \operatorname{Ei}(\mu(\mathbb{S}^2)A_2(\mathbf{x}))\right) \rho(\mathbf{y})\lambda_{\mathbb{S}^2}(d\mathbf{y})$, the first then follows by a similar argument. By using the series expansion we know that $\operatorname{Ei}(\mu(\mathbb{S}^2)A_2(\mathbf{x})) - \gamma - \log(\mu(\mathbb{S}^2)A_2(\mathbf{x})) = \sum_{n=1}^{\infty} (\left(\mu(\mathbb{S}^2)A_2(\mathbf{x})\right)^n)/(n \cdot n!)$ then it can be bound from above as,

$$
\begin{aligned}
\operatorname{Ei}(\mu(\mathbb{S}^2)A_2(\mathbf{x})) - \gamma - \log(\mu(\mathbb{S}^2)A_2(\mathbf{x})) &\leq \sum_{n=0}^{\infty} \frac{\left(\mu(\mathbb{S}^2)A_2(\mathbf{x})\right)^n}{n \cdot n!} \\
&\sum_{n=0}^{\infty} \frac{\left(\mu(\mathbb{S}^2)A_2(\mathbf{x})\right)^n}{n!} \\
&= \mathrm{e}^{\mu(\mathbb{S}^2)A_2(\mathbf{x})},
\end{aligned}
$$

and bound from below as,

$$
\begin{aligned}
\operatorname{Ei}(\mu(\mathbb{S}^2)A_2(\mathbf{x})) - \gamma - \log(\mu(\mathbb{S}^2)A_2(\mathbf{x})) &\geq \sum_{n=1}^{\infty} \frac{\left(\mu(\mathbb{S}^2)A_2(\mathbf{x})\right)^n}{(n+1)!} \\
&= \frac{1}{\mu(\mathbb{S}^2)A_2(\mathbf{x})} \left(\sum_{n=0}^{\infty} \frac{\left(\mu(\mathbb{S}^2)A_2(\mathbf{x})\right)^{n+1}}{(n+1)!} - 1\right) \\
&= \frac{1}{\mu(\mathbb{S}^2)A_2(\mathbf{x})} \left(\mathrm{e}^{\mu(\mathbb{S}^2)A_2(\mathbf{x})} - 1\right)
\end{aligned}
$$

Hence the integrand is bounded,

$$\begin{aligned}\frac{\mathrm{e}^{-\mu(\mathbb{S}^2)}}{\mu(\mathbb{S}^2)A_2^2(\mathbf{x})}\left(\mathrm{e}^{\mu(\mathbb{S}^2)A_2(\mathbf{x})}-1\right)\leq \\ \frac{\mathrm{e}^{-\mu(\mathbb{S}^2)}}{A_2(\mathbf{x})}\left(\gamma+\log(\mu(\mathbb{S}^2)A_2(\mathbf{x}))-\mathrm{Ei}(\mu(\mathbb{S}^2)A_2(\mathbf{x}))\right)\leq\frac{\mathrm{e}^{-\mu(\mathbb{S}^2)}}{A_2(\mathbf{x})}\mathrm{e}^{\mu(\mathbb{S}^2)A_2(\mathbf{x})}.\end{aligned} \tag{B.17}$$

We show that the lower and upper bounds can be bounded further such that they do not depend on r or $\mathbf{x}$. First we bound $\mathrm{e}^{-\mu(\mathbb{S}^2)}\mathrm{e}^{\mu(\mathbb{S}^2)A_2(\mathbf{x})}$ below and above,

$$\begin{aligned}\mathrm{e}^{-\mu(\mathbb{S}^2)}\mathrm{e}^{\mu(\mathbb{S}^2)A_2(\mathbf{x})}&=\exp\left(-4\pi\bar{\rho}(1-\cos r)+\frac{\bar{\rho}^2}{\mu(\mathbb{S}^2)}\int_{\mathbb{S}^2\cap B_{\mathbb{S}^2}(\mathbf{x},r)}\frac{1}{\rho(\mathbf{y})}\lambda_{\mathbb{S}^2}(d\mathbf{y})\right)\\&\leq\exp\left(\frac{\bar{\rho}^2}{\mu(\mathbb{S}^2)}\int_{\mathbb{S}^2\cap B_{\mathbb{S}^2}(\mathbf{x},r)}\frac{1}{\rho(\mathbf{y})}\lambda_{\mathbb{S}^2}(d\mathbf{y})\right)\\&\leq\exp\left(\frac{\bar{\rho}^2}{\mu(\mathbb{S}^2)}\int_{\mathbb{S}^2\cap B_{\mathbb{S}^2}(\mathbf{x},r)}\frac{1}{\bar{\rho}}\lambda_{\mathbb{S}^2}(d\mathbf{y})\right)\\&=\exp\left(\frac{4\pi\bar{\rho}}{\mu(\mathbb{S}^2)}\right)\\&\leq\mathrm{e},\end{aligned}$$

where the final inequality follows from $\mu(\mathbb{S}^2)\geq 4\pi\bar{\rho}$. The lower bound is,

$$\begin{aligned}\mathrm{e}^{-\mu(\mathbb{S}^2)}\mathrm{e}^{\mu(\mathbb{S}^2)A_2(\mathbf{x})}&=\exp\left(-4\pi\bar{\rho}(1-\cos r)+\frac{\bar{\rho}^2}{\mu(\mathbb{S}^2)}\int_{\mathbb{S}^2\cap B_{\mathbb{S}^2}(\mathbf{x},r)}\frac{1}{\rho(\mathbf{y})}\lambda_{\mathbb{S}^2}(d\mathbf{y})\right)\\&\geq\exp\left(-4\pi\bar{\rho}(1-\cos r)\right)\\&\geq\exp\left(-8\pi\bar{\rho}\right).\end{aligned}$$

Next we need to show that $A_2(\mathbf{x})$ is strictly greater than 0,

$$\begin{aligned}A_2(\mathbf{x})&=1-\frac{4\pi\bar{\rho}}{\mu(\mathbb{S}^2)}(1-\cos r)+\frac{\bar{\rho}^2}{\mu(\mathbb{S}^2)}\int_{\mathbb{S}^2\cap B_{\mathbb{S}^2}(\mathbf{x},r)}\frac{1}{\rho(\mathbf{y})}\lambda_{\mathbb{S}^2}(d\mathbf{y})\\&=\int_{\mathbb{S}^2}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{y}\in B_{\mathbb{S}^2}(\mathbf{x},r)]}{\rho(\mathbf{y})}\right)^2\frac{\rho(\mathbf{y})}{\mu(\mathbb{S}^2)}\lambda_{\mathbb{S}^2}(d\mathbf{y})\\&=\int_{\mathbb{S}^2\cap B_{\mathbb{S}^2}(\mathbf{x},r)}\left(1-\frac{\bar{\rho}}{\rho(\mathbf{y})}\right)^2\frac{\rho(\mathbf{y})}{\mu(\mathbb{S}^2)}\lambda_{\mathbb{S}^2}(d\mathbf{y})+\int_{\mathbb{S}^2\setminus B_{\mathbb{S}^2}(\mathbf{x},r)}\frac{\rho(\mathbf{y})}{\mu(\mathbb{S}^2)}\lambda_{\mathbb{S}^2}(d\mathbf{y}).\end{aligned}$$

Then for $r\in[0,\pi)$ $\mu_L(\mathbb{S}^2\setminus B_{\mathbb{S}^2}(\mathbf{x},r))>0$ and since $\rho(\mathbf{y})\geq\bar{\rho}>0$ this means that the second term is strictly greater than 0. Further the first term is always non-negative since

$\rho(\mathbf{y}) \geq \bar{\rho} > 0$. Then if $r = \pi$ we have that,

$$\begin{aligned} A_2(\mathbf{x}) &= \int_{\mathbb{S}^2} \left(1 - \frac{\bar{\rho}}{\rho(\mathbf{y})}\right)^2 \frac{\rho(\mathbf{y})}{\mu(\mathbb{S}^2)} \lambda_{\mathbb{S}^2}(d\mathbf{y}) \\ &= \int_{\mathbb{S}^2 \cap E} \left(1 - \frac{\bar{\rho}}{\rho(\mathbf{y})}\right)^2 \frac{\rho(\mathbf{y})}{\mu(\mathbb{S}^2)} \lambda_{\mathbb{S}^2}(d\mathbf{y}) + \int_{\mathbb{S}^2 \setminus E} \left(1 - \frac{\bar{\rho}}{\rho(\mathbf{y})}\right)^2 \frac{\rho(\mathbf{y})}{\mu(\mathbb{S}^2)} \lambda_{\mathbb{S}^2}(d\mathbf{y}), \end{aligned}$$

then by assumption $\rho(\mathbf{y}) > \bar{\rho}$ for all $\mathbf{y} \in E \subset \mathbb{S}^2$, hence the first term is strictly greater than 0 whilst the second term is non-negative and so we have shown that for any $r \in [0, \pi]$ $A_2(\mathbf{x}) > 0$, therefore when taking the reciprocal of $A_2(\mathbf{x})$ it is well defined for all $\mathbf{x} \in \mathbb{S}^2$. Then we can bound the absolute value of the reciprocal of $A_2(\mathbb{S}^2)$. Further, in the case when $r \in [0, \pi)$, $\int_{\mathbb{S}^2 \setminus B_{\mathbb{S}^2}(\mathbf{x}, r)} \frac{\rho(\mathbf{y})}{\mu(\mathbb{S}^2)} \lambda_{\mathbb{S}^2}(d\mathbf{y}) \geq \int_{\mathbb{S}^2 \setminus B_{\mathbb{S}^2}(\mathbf{x}, r)} \frac{\bar{\rho}}{\mu(\mathbb{S}^2)} \lambda_{\mathbb{S}^2}(d\mathbf{y}) = (1 - \cos r) \frac{2\pi\bar{\rho}}{\mu(\mathbb{S}^2)}$ then define,

$$\bar{A}_2 = \begin{cases} (1 - \cos r) \frac{2\pi\bar{\rho}}{\mu(\mathbb{S}^2)}, & r \in [0, \pi) \\ \int_{\mathbb{S}^2 \setminus E} \left(1 - \frac{\bar{\rho}}{\rho(\mathbf{y})}\right)^2 \frac{\rho(\mathbf{y})}{\mu(\mathbb{S}^2)} \lambda_{\mathbb{S}^2}(d\mathbf{y}), & r = \pi, \end{cases}$$

which does not depend on $\mathbf{x}$. Then $A_2(\mathbf{x}) \geq \bar{A}_2 > 0$ and so $0 \leq \frac{1}{A_2(\mathbf{x})} \leq \frac{1}{\bar{A}_2}$, which means we can bound the the absolute value of the reciprocal of $A_2(\mathbb{S}^2)$. Further we can obtain a non-zero lower bound for $\frac{1}{A_2(\mathbf{x})}$ since $A_2(\mathbf{x}) = 1 - \frac{4\pi\bar{\rho}}{\mu(\mathbb{S}^2)}(1 - \cos r) + \frac{\bar{\rho}^2}{\mu(\mathbb{S}^2)} \int_{\mathbb{S}^2 \cap B_{\mathbb{S}^2}(\mathbf{x}, r)} \frac{1}{\rho(\mathbf{y})} \lambda_{\mathbb{S}^2}(d\mathbf{y}) \leq 1 + \frac{\bar{\rho}^2}{\mu(\mathbb{S}^2)} \int_{\mathbb{S}^2} \frac{1}{\rho(\mathbf{y})} \lambda_{\mathbb{S}^2}(d\mathbf{y})$. Let us define $\tilde{A}_2 = 1 + \frac{\bar{\rho}^2}{\mu(\mathbb{S}^2)} \int_{\mathbb{S}^2} \frac{1}{\rho(\mathbf{y})} \lambda_{\mathbb{S}^2}(d\mathbf{y})$ then returning to Equation B.17 we have,

$$\frac{1}{\mu(\mathbb{S}^2)\tilde{A}_2^2} \left(\mathrm{e}^{-8\pi\bar{\rho}} - \mathrm{e}^{-\mu(\mathbb{S}^2)}\right) \leq \frac{\mathrm{e}^{-\mu(\mathbb{S}^2)}}{A_2(\mathbf{x})} \left(\gamma + \log(\mu(\mathbb{S}^2) A_2(\mathbf{x})) - \mathrm{Ei}(\mu(\mathbb{S}^2) A_2(\mathbf{x}))\right) \leq \frac{\mathrm{e}}{\bar{A}_2},$$

and therefore,

$$\left| \frac{\mathrm{e}^{-\mu(\mathbb{S}^2)}}{A_2(\mathbf{x})} \left(\gamma + \log(\mu(\mathbb{S}^2) A_2(\mathbf{x})) - \mathrm{Ei}(\mu(\mathbb{S}^2) A_2(\mathbf{x}))\right) \right| \leq \max \left\{ \left| \frac{1}{\mu(\mathbb{S}^2)\tilde{A}_2^2} \left(\mathrm{e}^{-8\pi\bar{\rho}} - \mathrm{e}^{-\mu(\mathbb{S}^2)}\right) \right|, \frac{\mathrm{e}}{\bar{A}_2} \right\}$$

Since the right hand side of this inequality is a constant this means that it is Lebesgue integrable over $\mathbb{S}^2$ and so by the dominated convergence theorem so is the left hand side, thus showing that the integrands are truly Lebesgue integrable. An identical approach can be used to show that the first term of Equation B.16 is also Lebesgue integrable. □

B.7 Proof of Theorem 4.4.4

Proof. Starting with the expectation,

$$\mathbb{E}[\hat{J}_{\text{inhom}}(r)] = \int_{N_{lf}} \hat{J}_{\text{inhom}}(r) P_X(dx),$$

where $P_X(X \in F), F \subseteq N_{lf}$ is the probability measure of X. Define $N_{lf,0} = \{x \in N_{lf} | n_x(\mathbb{S}^2) = 0\}$ and $N_{lf,1} = \{x \in N_{lf} | n_x(\mathbb{S}^2) > 0\}$, then $N_{lf} = N_{lf,0} \cup N_{lf,1}$ and $N_{lf,0} \cap N_{lf,1} = \emptyset$ and so,

$$\mathbb{E}[\hat{J}_{\text{inhom}}(r)] = \int_{N_{lf,0}} \hat{J}_{\text{inhom}}(r) P_X(dx) + \int_{N_{lf,1}} \hat{J}_{\text{inhom}}(r) P_X(dx),$$

taking the convention that $\frac{0}{0} = 1$, the first term is finite,

$$\begin{aligned} &= \int_{N_{lf,0}} P_X(dx) + \int_{N_{lf,1}} \hat{J}_{\text{inhom}}(r) P_X(dx), \\ &= P_X(X \in N_{lf,0}) + \int_{N_{lf,1}} \hat{J}_{\text{inhom}}(r) P_X(dx), \end{aligned}$$

it can be shown that $P_X(X \in N_{lf,0}) = P_{N_X(\mathbb{S}^2)}(N_X(\mathbb{S}^2) = 0)$ then,

$$= P(N_X(\mathbb{S}^2) = 0) + \int_{N_{lf,1}} \hat{J}_{\text{inhom}}(r) P_X(dx).$$

We now show that the second term can be bounded and hence the expectation is well defined. Taking the integrand and noting that random variables are now deterministic,

$$\hat{J}_{\text{inhom}}(r) = \frac{|P| \sum_{\mathbf{x} \in x} \prod_{\mathbf{y} \in x \setminus \mathbf{x}} \left(1 - \frac{\bar{\rho} \mathbb{1}[\mathbf{y} \in B_{\mathbb{S}^2}(\mathbf{x}, r)]}{\rho(\mathbf{y})}\right)}{n_x(\mathbb{S}^2) \sum_{\mathbf{p} \in P} \prod_{\mathbf{z} \in x} \left(1 - \frac{\bar{\rho} \mathbb{1}[\mathbf{z} \in B_{\mathbb{S}^2}(\mathbf{p}, r)]}{\rho(\mathbf{z})}\right)}.$$

First we note that $\hat{J}_{\text{inhom}}(r) \geq 0$ and is so bounded below. Working with the numerator we have that,

$$\begin{aligned} \sum_{\mathbf{x} \in x} \prod_{\mathbf{y} \in x \setminus \mathbf{x}} \left(1 - \frac{\bar{\rho} \mathbb{1}[\mathbf{y} \in B_{\mathbb{S}^2}(\mathbf{x}, r)]}{\rho(\mathbf{y})}\right) &\leq \sum_{\mathbf{x} \in x} \prod_{y \in x \setminus \mathbf{x}} 1 \\ &= n_x(\mathbb{S}^2). \end{aligned}$$

Now working with the denominator showing that it is bounded below and hence its reciprocal bounded above. By the assumption we have that there exists $\tilde{\mathbf{p}} \in P$ such that for any $\mathbf{x} \in B_{\mathbb{S}^2}(\tilde{\mathbf{p}}, r), \rho(\mathbf{x}) \neq \bar{\rho}$, then

$$\sum_{\mathbf{p}\in P}\prod_{\mathbf{z}\in x}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{z}\in B_{\mathbb{S}^2}(\mathbf{p},r)]}{\rho(\mathbf{z})}\right) = \prod_{\mathbf{z}\in x}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{z}\in B_{\mathbb{S}^2}(\tilde{\mathbf{p}},r)]}{\rho(\mathbf{z})}\right) + \sum_{\mathbf{p}\in P\setminus\tilde{p}}\prod_{\mathbf{z}\in x}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{z}\in B_{\mathbb{S}^2}(\mathbf{p},r)]}{\rho(\mathbf{z})}\right).$$

Then the first term is strictly greater than 0 by our assumption. Thus,

$$\sum_{\mathbf{p}\in P}\prod_{\mathbf{z}\in x}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{z}\in B_{\mathbb{S}^2}(\mathbf{p},r)]}{\rho(\mathbf{z})}\right) \geq \prod_{\mathbf{z}\in x}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{z}\in B_{\mathbb{S}^2}(\tilde{\mathbf{p}},r)]}{\rho(\mathbf{z})}\right) > 0.$$

Further by the assumption we can define $\bar{\rho}_{\tilde{\mathbf{p}}} = \inf_{\mathbf{x}\in B_{\mathbb{S}^2}(\tilde{\mathbf{p}},r)}\rho(\mathbf{x})$ and then,

$$\sum_{\mathbf{p}\in P}\prod_{\mathbf{z}\in x}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{z}\in B_{\mathbb{S}^2}(\mathbf{p},r)]}{\rho(\mathbf{z})}\right) \geq \prod_{\mathbf{z}\in x}\left(1-\frac{\bar{\rho}}{\bar{\rho}_{\tilde{\mathbf{p}}}}\right) = \left(1-\frac{\bar{\rho}}{\bar{\rho}_{\tilde{\mathbf{p}}}}\right)^{n_x(\mathbf{S}^2)},$$

and so,

$$\left|\hat{J}_{\text{inhom}}(r)\right| \leq \frac{|P|n_x(\mathbb{S}^2)}{n_x(\mathbb{S}^2)\left(1-\frac{\bar{\rho}}{\bar{\rho}_{\tilde{\mathbf{p}}}}\right)^{n_x(\mathbf{S}^2)}} = |P|\left(1-\frac{\bar{\rho}}{\bar{\rho}_{\tilde{\mathbf{p}}}}\right)^{-n_x(\mathbf{S}^2)}.$$

We now show that the right hand side of the above inequality is integrable. Define the sets $N_{lf,1,(i)} = \{x \in N_{lf,1} : n(x) = i\}$, then $N_{lf,1} = \bigcup_{i=1}^{\infty} N_{lf,1,(i)}$ and $N_{lf,1,(i)} \cap N_{lf,1,(j)} = \emptyset$ for $i \neq j$ and hence we have partitioned the space $N_{lf,1}$. Thus,

$$\int_{N_{lf,1}} |P|\left(1-\frac{\bar{\rho}}{\bar{\rho}_{\tilde{\mathbf{p}}}}\right)^{-n_x(\mathbf{S}^2)} P_X(dx) = \sum_{i=1}^{\infty}\int_{N_{lf,1,(i)}} |P|\left(1-\frac{\bar{\rho}}{\bar{\rho}_{\tilde{\mathbf{p}}}}\right)^{-n_x(\mathbb{S}^2)} P_X(dx)$$

$n_x(\mathbb{S}^2) = i$ for all $x \in N_{lf,1,(i)}$ and so is constant over each partition of the space,

$$\begin{aligned}
&= \sum_{i=1}^{\infty} |P|\left(1-\frac{\bar{\rho}}{\bar{\rho}_{\tilde{\mathbf{p}}}}\right)^{-i}\int_{N_{lf,1,(i)}} P_X(dx)\\
&= \sum_{i=1}^{\infty} |P|\left(1-\frac{\bar{\rho}}{\bar{\rho}_{\tilde{\mathbf{p}}}}\right)^{-i} P_X(X \in N_{lf,1,(i)})
\end{aligned}$$

but it is easy to see that $P_X(X \in N_{lf,1,(i)}) = P_{N_X(\mathbb{S}^2)}(N_X(\mathbb{S}^2) = i)$, and since X is Poisson and defining $a = (1 - \frac{\bar{\rho}}{\bar{\rho}_{\tilde{\mathbf{P}}}})$ and $\lambda = \mu(\mathbb{S}^2)$,

$$\begin{aligned}
&= |P| \sum_{i=1}^{\infty} a^{-i} \frac{\lambda^n \mathrm{e}^{-\lambda}}{i!} \\
&= |P| \mathrm{e}^{-\lambda} \sum_{i=1}^{\infty} \frac{(\lambda/a)^i}{i!} \\
&= |P| \mathrm{e}^{-\lambda} \left(\sum_{i=0}^{\infty} \frac{(\lambda/a)^i}{i!} - 1 \right) \\
&= |P| \mathrm{e}^{-\lambda} \left(\mathrm{e}^{-\lambda/a} - 1 \right).
\end{aligned}$$

Hence we have shown that,

$$\int_{N_{lf,1}} \left| \hat{J}_{\mathrm{inhom}}(r) \right| P_X(dx) \leq |P| \mathrm{e}^{-\lambda} \left(\mathrm{e}^{-\lambda/a} - 1 \right)$$

and so by the dominated convergence theorem the expectation of $\hat{J}_{\mathrm{inhom}}(r)$ exists.

Existence of the variance of $\hat{J}_{\mathrm{inhom}}(r)$ follows simply based on the existence of the expectation,

$$\mathrm{Var}(\hat{J}_{\mathrm{inhom}}(r)) = \int_{N_{lf}} \left(\hat{J}_{\mathrm{inhom}}(r) - \mathbb{E}[\hat{J}_{\mathrm{inhom}}(r)] \right)^2 P_X(dx).$$

Partitioning the space N_{lf} again into $N_{lf,0} \equiv \{x \in N_{lf} | N(x) = 0\}$ and $N_{lf,1} \equiv \{x \in N_{lf} | N(x) > 0\}$ we have that,

$$\mathrm{Var}(\hat{J}_{\mathrm{inhom}}(r)) \tag{B.18}$$

$$= \int_{N_{lf,0}} \left(\hat{J}_{\mathrm{inhom}}(r) - \mathbb{E}[\hat{J}_{\mathrm{inhom}}(r)] \right)^2 P_X(dx) + \int_{N_{lf,1}} \left(\hat{J}_{\mathrm{inhom}}(r) - \mathbb{E}[\hat{J}_{\mathrm{inhom}}(r)] \right)^2 P_X(dx). \tag{B.19}$$

Taking the convention that $\frac{0}{0} = 1$ then the first term is simply,

$$\begin{aligned}
\int_{N_{lf,0}} \left(\hat{J}_{\mathrm{inhom}}(r) - \mathbb{E}[\hat{J}_{\mathrm{inhom}}(r)] \right)^2 P_X(dx) &= \mathbb{E}^2[\hat{J}_{\mathrm{inhom}}(r)] \int_{N_{lf,0}} P_X(dx) \\
&= \mathbb{E}^2[\hat{J}_{\mathrm{inhom}}(r)] P_X(X \in N_{lf,0}),
\end{aligned}$$

then we note that $P_X(X \in N_{lf,0}) = P(N_X(\mathbb{S}^2) = 0)$ and so,

$$= \mathbb{E}^2[\hat{J}_{\text{inhom}}(r)]P(N_X(\mathbb{S}^2) = 0),$$

which is finite since the expectation exists. The second term of Equation B.19 over the space $N_{lf,1}$ can also be shown to be finite. Since $\mathbb{E}[\hat{J}_{\text{inhom}}(r)]$ is finite we just need to show that $\hat{J}_{\text{inhom}}(r)$ is bounded in order show that the integrand is bounded and hence integrable. But from proving the expectation exists we have that,

$$0 \leq \hat{J}_{\text{inhom}}(r) \leq |P| \left(1 - \frac{\bar{\rho}}{\bar{\rho}_{\tilde{\mathbf{p}}}}\right)^{-n_x(\mathbf{S}^2)},$$

and hence the square of $\hat{J}_{\text{inhom}}(r)$ is also bounded and so the integrand of second term in Equation B.19 is bounded above and thus the variance exists. □

Remark S5. *From the proof of this theorem the requirement of the process being Poisson was only needed for a closed form of the distribution for its corresponding counting process $N_X(\mathbb{S}^2)$. We can drop the requirement of X being Poisson but instead require that the probability generating function of $N_X(\mathbb{S}^2)$, $G_{N_X(\mathbb{S}^2)}(s)$, has radius of convergence $|s| \leq \left(1 - \frac{\bar{\rho}}{\bar{\rho}_{\tilde{\mathbf{p}}}}\right)^{-2}$, where $\bar{\rho}_{\tilde{\mathbf{p}}} = \inf_{\mathbf{x} \in B_{\mathbb{S}^2}(\tilde{\mathbf{p}}, r)} \rho(\mathbf{x})$ and there exists $\tilde{\mathbf{p}} \in P$ such that for any $\mathbf{x} \in B_{\mathbb{S}^2}(\tilde{\mathbf{p}}, r), \rho(\mathbf{x}) \neq \bar{\rho}$. This condition would be sufficient for the theorem to hold true also.*

B.8 Proof of Propositon 4.4.5

Proof. Define,

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

$$Y \equiv 1 - \hat{F}_{\text{inhom}}(r) = \frac{1}{|P|} \sum_{\mathbf{p} \in P} \prod_{\mathbf{y} \in X} \left(1 - \frac{\bar{\rho} \mathbb{1}[\mathbf{y} \in B_{\mathbb{S}^2}(\mathbf{p}, r)]}{\rho(\mathbf{y})}\right)$$

Then,

$$\text{Cov}(X, Y) = \underbrace{\mathbb{E}[XY]}_{(a)} - \underbrace{\mathbb{E}[X]\mathbb{E}[Y]}_{(b)}$$

Term (a) is given by,

$\mathbb{E}[XY]$

$$= \mathbb{E}\left[\left(\frac{1}{N_X(\mathbb{S}^2)}\sum_{\mathbf{x}\in X}\prod_{\mathbf{y}\in X\setminus\{\mathbf{x}\}}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{y}\in B_{\mathbb{S}^2}(\mathbf{x},r)]}{\rho(\mathbf{y})}\right)\right)\left(\frac{1}{|P|}\sum_{\mathbf{p}\in P}\prod_{\mathbf{y}\in X}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{y}\in B_{\mathbb{S}^2}(\mathbf{p},r)]}{\rho(\mathbf{y})}\right)\right)\right]$$
$$= \frac{1}{|P|}\sum_{\mathbf{p}\in P}\mathbb{E}\left[\frac{1}{N_X(\mathbb{S}^2)}\sum_{\mathbf{x}\in X}\left(\prod_{\mathbf{y}\in X\setminus\{\mathbf{x}\}}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{y}\in B_{\mathbb{S}^2}(\mathbf{x},r)]}{\rho(\mathbf{y})}\right)\right)\left(\prod_{\mathbf{y}\in X}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{y}\in B_{\mathbb{S}^2}(\mathbf{p},r)]}{\rho(\mathbf{y})}\right)\right)\right]$$
$$= \frac{1}{|P|}\sum_{\mathbf{p}\in P}\mathbb{E}\left(\frac{1}{N_X(\mathbb{S}^2)}\sum_{i=1}^{N_X(\mathbb{S}^2)}\underbrace{\mathbb{E}\left[\left(\prod_{\substack{j=1\\j\neq i}}^{n}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{X}_j\in B_{\mathbb{S}^2}(\mathbf{X}_i,r)]}{\rho(\mathbf{X}_j)}\right)\right)\left(\prod_{j=1}^{n}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{X}_j\in B_{\mathbb{S}^2}(\mathbf{p},r)]}{\rho(\mathbf{X}_j)}\right)\right)\middle|N_X(\mathbb{S}^2)=n\right]}_{(\star)}\right),$$

where $X_k,\ k=1,\ldots,n$ are independently distributed on $\mathbb{S}^2$ with density proportional to $\rho(\mathbf{x}_k)$, when conditioned on $N_X(\mathbb{S}^2)$. Taking $(\star)$,

$$\mathbb{E}\left[\left(\prod_{\substack{j=1\\j\neq i}}^{n}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{X}_j\in B_{\mathbb{S}^2}(\mathbf{X}_i,r)]}{\rho(\mathbf{X}_j)}\right)\right)\left(\prod_{j=1}^{n}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{X}_j\in B_{\mathbb{S}^2}(\mathbf{p},r)]}{\rho(\mathbf{X}_j)}\right)\right)\middle|N_X(\mathbb{S}^2)=n\right]$$
$$= \mathbb{E}\left[\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{X}_i\in B_{\mathbb{S}^2}(\mathbf{p},r)]}{\rho(\mathbf{X}_i)}\right)\left(\prod_{\substack{j=1\\j\neq i}}^{n}\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{X}_j\in B_{\mathbb{S}^2}(\mathbf{X}_i,r)]}{\rho(\mathbf{X}_j)}\right)\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{X}_j\in B_{\mathbb{S}^2}(\mathbf{p},r)]}{\rho(\mathbf{X}_j)}\right)\right)\right],$$

Using iterated expectations and conditioning on $\mathbf{X}_i$,

$$= \mathbb{E}\left[\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{X}_i\in B_{\mathbb{S}^2}(\mathbf{p},r)]}{\rho(\mathbf{X}_i)}\right)\prod_{\substack{j=1\\j\neq i}}^{n}\mathbb{E}\left[\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{X}_j\in B_{\mathbb{S}^2}(\mathbf{x},r)]}{\rho(\mathbf{X}_j)}\right)\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{X}_j\in B_{\mathbb{S}^2}(\mathbf{p},r)]}{\rho(\mathbf{X}_j)}\right)\right]\right]$$
$$= \mathbb{E}\left[\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{X}_i\in B_{\mathbb{S}^2}(\mathbf{p},r)]}{\rho(\mathbf{X}_i)}\right)\mathbb{E}^{n-1}\left[\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{Y}\in B_{\mathbb{S}^2}(\mathbf{x},r)]}{\rho(\mathbf{Y})}\right)\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{Y}\in B_{\mathbb{S}^2}(\mathbf{p},r)]}{\rho(\mathbf{Y})}\right)\right]\right] \tag{B.20}$$

where $\mathbf{Y}$ is distributed with density proportional to $\rho(\mathbf{y})$ on $\mathbb{S}^2$. It is then easy to show that,

$$\mathbb{E}\left[\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{Y}\in B_{\mathbb{S}^2}(\mathbf{x},r)]}{\rho(\mathbf{Y})}\right)\left(1-\frac{\bar{\rho}\mathbb{1}[\mathbf{Y}\in B_{\mathbb{S}^2}(\mathbf{p},r)]}{\rho(\mathbf{Y})}\right)\right]$$

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

Let us define $A(\mathbf{x},\mathbf{p}) = \mathbb{E}\left[\left(1 - \frac{\bar{\rho}\mathbb{1}[\mathbf{Y}\in B_{\mathbb{S}^2}(\mathbf{x},r)]}{\rho(\mathbf{Y})}\right)\left(1 - \frac{\bar{\rho}\mathbb{1}[\mathbf{Y}\in B_{\mathbb{S}^2}(\mathbf{p},r)]}{\rho(\mathbf{Y})}\right)\right]$, and so returning to B.20,

$$\mathbb{E}\left[\left(1 - \frac{\bar{\rho}\mathbb{1}[\mathbf{X}_i \in B_{\mathbb{S}^2}(\mathbf{p},r)]}{\rho(\mathbf{X}_i)}\right) A^{n-1}(\mathbf{X}_i,\mathbf{p})\right] = \mathbb{E}\left[\left(1 - \frac{\bar{\rho}\mathbb{1}[\mathbf{X} \in B_{\mathbb{S}^2}(\mathbf{p},r)]}{\rho(\mathbf{X})}\right) A^{n-1}(\mathbf{X},\mathbf{p})\right],$$

where $\mathbf{X}$ has density proportional to $\rho(\mathbf{x})$ on $\mathbb{S}^2$,

$$= \int_{\mathbb{S}^2} \left(1 - \frac{\bar{\rho}\mathbb{1}[\mathbf{x} \in B_{\mathbb{S}^2}(\mathbf{p},r)]}{\rho(\mathbf{x})}\right) A^{n-1}(\mathbf{x},\mathbf{p}) \frac{\rho(\mathbf{x})}{\mu(\mathbb{S}^2)} \lambda_{\mathbb{S}^2}(d\mathbf{x}).$$

And so,

$$\begin{aligned}
\mathbb{E}[XY] &= \frac{1}{|P|}\sum_{\mathbf{p}\in P} \mathbb{E}\left[\frac{1}{N_X(\mathbb{S}^2)} \sum_{i=1}^{N_X(\mathbb{S}^2)} \int_{\mathbb{S}^2} \left(1 - \frac{\bar{\rho}\mathbb{1}[\mathbf{x} \in B_{\mathbb{S}^2}(\mathbf{p},r)]}{\rho(\mathbf{x})}\right) A^{N_X(\mathbb{S}^2)-1}(\mathbf{x},\mathbf{p}) \frac{\rho(\mathbf{x})}{\mu(\mathbb{S}^2)} \lambda_{\mathbb{S}^2}(d\mathbf{x})\right] \\
&= \frac{1}{|P|}\sum_{\mathbf{p}\in P} \mathbb{E}\left[\int_{\mathbb{S}^2} \left(1 - \frac{\bar{\rho}\mathbb{1}[\mathbf{x} \in B_{\mathbb{S}^2}(\mathbf{p},r)]}{\rho(\mathbf{x})}\right) A^{N_X(\mathbb{S}^2)-1}(\mathbf{x},\mathbf{p}) \frac{\rho(\mathbf{x})}{\mu(\mathbb{S}^2)} \lambda_{\mathbb{S}^2}(d\mathbf{x})\right] \\
&= \frac{1}{|P|}\sum_{\mathbf{p}\in P} \int_{\mathbb{S}^2} \left(1 - \frac{\bar{\rho}\mathbb{1}[\mathbf{x} \in B_{\mathbb{S}^2}(\mathbf{p},r)]}{\rho(\mathbf{x})}\right) \mathbb{E}\left[A^{N_X(\mathbb{S}^2)-1}(\mathbf{x},\mathbf{p})\right] \frac{\rho(\mathbf{x})}{\mu(\mathbb{S}^2)} \lambda_{\mathbb{S}^2}(d\mathbf{x})
\end{aligned}$$

Then it can easily be shown that,

$$\mathbb{E}\left[A^{N_X(\mathbb{S}^2)-1}(\mathbf{x},\mathbf{p})\right] = \frac{\exp\left\{-2\bar{\rho}2\pi(1-\cos r) - \int_{B_{\mathbb{S}^2}(\mathbf{x},r)\cap B_{\mathbb{S}^2}(\mathbf{p},r)} \frac{\bar{\rho}^2}{\rho(\mathbf{y})} \lambda_{\mathbb{S}^2}(d\mathbf{y})\right\}}{A(\mathbf{x},\mathbf{p})}.$$

Then from Theorem 4.4.1 we have that,

$$\begin{aligned}
\mathbb{E}[X] &= \exp(2\pi(1-\cos r)\bar{\rho}) \\
\mathbb{E}[Y] &= \mu^2(\mathbb{S}) \frac{\exp(2\pi(1-\cos r)\bar{\rho}) - \exp(\mu(\mathbb{S}^2))}{\mu(\mathbb{S}^2) - 2\pi(1-\cos r)\bar{\rho}}.
\end{aligned}$$

And so we have the covariance between $\hat{H}_{\text{inhom}}(r)$ and $\hat{F}_{\text{inhom}}(r)$. □

B.9 Taylor series expansion for $\hat{J}_{\text{INHOM}}$

We discuss Taylor series expansions in general and their use to approximate moments of random variables Wolter ((2007)), after which we apply this to the $\hat{J}_{\text{inhom}}$-function. For any function $f : \mathbb{R}^2 \mapsto \mathbb{R}$, we have its Taylor expansion up to second order around $\boldsymbol{\theta} = (\theta_x, \theta_y)$ as,

$$f(x,y) = f(\theta_x,\theta_y) + \frac{\partial f}{\partial x}\bigg|_{\theta_X,\theta_Y}(x-\theta_x) + \frac{\partial f}{\partial y}\bigg|_{\theta_X,\theta_Y}(y-\theta_y) + \\ \frac{1}{2}\left[\frac{\partial^2 f}{\partial x^2}\bigg|_{\theta_X,\theta_Y}(x-\theta_x)^2 + 2\frac{\partial^2 f}{\partial x \partial y}\bigg|_{\theta_X,\theta_Y}(x-\theta_x)(y-\theta_y) + \frac{\partial^2 f}{\partial y^2}\bigg|_{\theta_X,\theta_Y}(y-\theta_y)^2\right] + R(x,y),$$

where $R(x,y)$ is a remainder term. Thus using random variables X and Y, with $\boldsymbol{\theta} = (\mathbb{E}[X], \mathbb{E}[Y]) \equiv (\mu_X, \mu_Y)$, whilst also assuming that $\mathbb{E}[R(X,Y)]$ is close to 0 then,

$$\mathbb{E}[f(X,Y)] \approx f(\mu_X,\mu_Y) + \frac{1}{2}\left[\frac{\partial^2 f}{\partial x^2}\text{Var}(X) + 2\frac{\partial^2 f}{\partial x \partial y}\text{Cov}(X,Y) + \frac{\partial^2 f}{\partial y^2}\text{Var}(Y)\right].$$

Then to approximate the variance we first note that using the first order Taylor series expansion $\mathbb{E}[f(X,Y)] \approx f(\mu_X,\mu_Y)$. Then,

$$\begin{aligned}
\text{Var}(f(X,Y)) &= \mathbb{E}[(f(X,Y) - \mathbb{E}[f(X,Y)])^2] \\
&\approx \mathbb{E}[(f(X,Y) - f(\mu_X,\mu_Y))^2] \\
&\approx \mathbb{E}\left[\left(f(\mu_X,\mu_Y) + \frac{\partial f}{\partial x}\bigg|_{\mu_X,\mu_Y}(X-\mu_X) + \frac{\partial f}{\partial y}\bigg|_{\mu_X,\mu_Y}(Y-\mu_Y) - f(\mu_X,\mu_Y)\right)^2\right] \\
&= \mathbb{E}\left[\left(\frac{\partial f}{\partial x}\bigg|_{\mu_X,\mu_Y}(X-\mu_X) + \frac{\partial f}{\partial y}\bigg|_{\mu_X,\mu_Y}(Y-\mu_Y)\right)^2\right] \\
&= \frac{\partial f}{\partial x}\bigg|^2_{\mu_X,\mu_Y}\text{Var}(X) + 2\frac{\partial f}{\partial x}\bigg|_{\mu_X,\mu_Y}\frac{\partial f}{\partial y}\bigg|_{\mu_X,\mu_Y}\text{Cov}(X,Y) + \frac{\partial f}{\partial y}\bigg|^2_{\mu_X,\mu_Y}\text{Var}(Y)
\end{aligned}$$

Hence defining $f(X,Y) = \frac{X}{Y}$, we have the following approximations to the first and second order moments of $f(X,Y)$,

$$\mathbb{E}\left[\frac{X}{Y}\right] \approx \frac{\mu_X}{\mu_Y} - \frac{\text{Cov}(X,Y)}{\mu_Y^2} + \frac{\text{Var}(Y)\mu_X}{\mu_Y^3} \tag{B.21}$$

$$\text{Var}\left(\frac{X}{Y}\right) \approx \frac{\mu_X}{\mu_Y}\left[\frac{\text{Var}(X)}{\mu_X^2} - 2\frac{\text{Cov}(X,Y)}{\mu_X\mu_Y} + \frac{\text{Var}(Y)}{\mu_Y^2}\right]. \tag{B.22}$$

Then since $\hat{J}_{\text{inhom}}(r)$ is defined as the ratio of two random variables, in particular $1 - \hat{H}_{\text{inhom}}(r)$ and $1 - \hat{F}_{\text{inhom}}(r)$. Thus combining Proposition 4.4.5 with the Taylor series expansions given in Equations B.21 and B.22 provides an estimate for the expectation and variance of $\hat{J}_{\text{inhom}}(r)$.

B.10 Proof of Proposition 4.6.4

Proof. Let us first start with X_1. Define $Y_1 \sim PPP(\rho, \mathbb{D})$ to be the homogeneous Poisson process which is thinned to give X_1. Then $\forall B \subseteq \mathbb{D}$ we can rewrite the counting measure for X_1 as,

$$N_{X_1}(B) = \sum_{\mathbf{x} \in Y_1 \cap B} \mathbb{1}[N_{Y_1 \setminus \{\mathbf{x}\}}(B_{\mathbb{D}}(\mathbf{x}, R)) = 0],$$

where $N_{Y_1 \setminus \{\mathbf{x}\}}$, is the random counting measure for the process Y_1 without the point $\mathbf{x}$. Then taking expectations and using the Slivnyak-Mecke Theorem,

$$\begin{aligned}
\mathbb{E}[N_{X_1}(B)] &= \mathbb{E}\left[\sum_{\mathbf{x} \in Y_1 \cap B} \mathbb{1}[N_{Y_1 \setminus \{\mathbf{x}\}}(B_{\mathbb{D}}(\mathbf{x}, R)) = 0]\right] \\
&= \int_B \mathbb{E}\left[\mathbb{1}[N_{Y_1}(B_{\mathbb{D}}(\mathbf{x}, R)) = 0]\right] \rho \lambda_{\mathbb{D}}(d\mathbf{x}) \\
&= \rho \int_B \mathbb{P}(N_{Y_1}(B_{\mathbb{D}}(\mathbf{x}, R)) = 0) \lambda_{\mathbb{D}}(d\mathbf{x}) \\
&= \rho \int_B \mathrm{e}^{-\rho \lambda_{\mathbb{D}}(B_{\mathbb{D}}(\mathbf{x}, R))} \lambda_{\mathbb{D}}(d\mathbf{x})
\end{aligned}$$

For $N_{X_2}(B)$ a few more steps are required in order to take into account the mark associated with each point. Similarly to the counting measure for X_1, we can rewrite the counting measure for X_2 as,

$$N_{X_2}(B) = \sum_{\mathbf{x} \in Y_2 \cap B} \mathbb{1}[M_{\mathbf{x}} \leq M_{\mathbf{y}}, \forall \mathbf{y} \in (Y_2 \setminus \{\mathbf{x}\}) \cap B_{\mathbb{D}}(\mathbf{x}, R)].$$

By again taking expectations and using the Slivnyak-Mecke Theorem,

$$\mathbb{E}[N_{X_2}(B)] = \mathbb{E}\left[\sum_{\mathbf{x} \in Y_2 \cap B} \mathbb{1}[M_{\mathbf{x}} \leq M_{\mathbf{y}}, \forall \mathbf{y} \in (Y_2 \setminus \{\mathbf{x}\}) \cap B_{\mathbb{D}}(\mathbf{x}, R)]\right]$$

$$= \int_B \mathbb{E}[\mathbb{1}[M_{\mathbf{x}} \leq M_{\mathbf{y}}, \forall \mathbf{y} \in Y_2 \cap B_{\mathbb{D}}(\mathbf{x}, R)]]\rho\lambda_{\mathbb{D}}(d\mathbf{x})$$
$$= \int_B \mathbb{P}(M_{\mathbf{x}} \leq M_{\mathbf{y}}, \forall \mathbf{y} \in Y_2 \cap B_{\mathbb{D}}(\mathbf{x}, R))\rho\lambda_{\mathbb{D}}(d\mathbf{x})$$

Define $\lambda_{\mathbf{x}} = \rho\lambda_{\mathbb{D}}(B_{\mathbb{D}}(\mathbf{x}, R))$, then the probability is calculated as follows,

$$\begin{aligned}
&\mathbb{P}(M_{\mathbf{x}} \leq M_{\mathbf{y}}, \forall \mathbf{y} \in Y_2 \cap B_{\mathbb{D}}(\mathbf{x}, R)) \\
&\quad = \sum_{n=0}^{\infty} \mathbb{P}\left(M_{\mathbf{x}} \leq M_{\mathbf{y}}, \forall \mathbf{y} \in Y_2 \cap B_{\mathbb{D}}(\mathbf{x}, R) | N_{Y_2}(B_{\mathbb{D}}(\mathbf{x}, R)) = n\right) \mathbb{P}(N_{Y_2}(B_{\mathbb{D}}(\mathbf{x}, R)) = n) \\
&\quad = \sum_{n=0}^{\infty} \frac{\lambda_{\mathbf{x}}^n \mathrm{e}^{-\lambda_{\mathbf{x}}}}{n!} \mathbb{P}\left(M_{\mathbf{x}} \leq M_{\mathbf{y}}, \forall \mathbf{y} \in Y_2 \cap B_{\mathbb{D}}(\mathbf{x}, R) | N_{Y_2}(B_{\mathbb{D}}(\mathbf{x}, R)) = n\right)
\end{aligned}$$

Let us label the points $\mathbf{y}_i \in Y_2$, for $i = 1, \ldots, n$ to be the n points in $B_{\mathbb{D}}(\mathbf{x}, R)$ coming from the process Y_2. Then the event $M_{\mathbf{x}} \leq M_{\mathbf{y}}, \forall \mathbf{y} \in Y_2 \cap B_{\mathbb{D}}(\mathbf{x}, R)$ given that $N_{Y_2}(B_{\mathbb{D}}(\mathbf{x}, R)) = n$ is identical to the event that the mark associated to the point $\mathbf{x}$ is the smallest of all marks. In other words we are concerned with the event $M_{\mathbf{x}} \leq \min\{M_{\mathbf{y}_1}, \ldots, M_{\mathbf{y}_n}\}$. Let us define $M_{\min} = \min\{M_{\mathbf{y}_1}, \ldots, M_{\mathbf{y}_n}\}$ and $f_{M^{\min}}(m)$ to be the density function of $M_{\min}$, then by using order statistics we know that $f_{M^{\min}}(m) = n f_{M_{\mathbf{x}}}(m)(1 - F_{M_{\mathbf{x}}}(m))^{n-1}$, where $f_{M_{\mathbf{x}}}(m)$ and $F_{M_{\mathbf{x}}}(m)$ are the density and cumulative density functions of an individual markis respectively. Then we have,

$$\begin{aligned}
&\mathbb{P}\left(M_{\mathbf{x}} \leq M_{\mathbf{y}}, \forall \mathbf{y} \in Y_2 \cap B_{\mathbb{D}}(\mathbf{x}, R) | N_{Y_2}(B_{\mathbb{D}}(\mathbf{x}, R)) = n\right) \\
&\quad = \mathbb{P}(M_{\mathbf{x}} \leq M^{\min}) \\
&\quad = \int_{\mathbb{M}^{\min}} \mathbb{P}(M_{\mathbf{x}} \leq m) f_{M^{\min}}(m) dm \\
&\quad = \int_{\mathbb{M}^{\min}} F_{M_{\mathbf{x}}}(m) n f_{M_{\mathbf{x}}}(m)(1 - F_{M_{\mathbf{x}}}(m))^{n-1} dm \\
&\quad = [F_{M_{\mathbf{x}}}(m)(1 - F_{M_{\mathbf{x}}}(m))^n]^{\mathbb{M}^{\min}} + \int_{\mathbb{M}^{\min}} f_{M_{\mathbf{x}}}(m)(1 - F_{M_{\mathbf{x}}}(m))^n dm \\
&\quad = [F_{M_{\mathbf{x}}}(m)(1 - F_{M_{\mathbf{x}}}(m))^n]^{\mathbb{M}^{\min}} + \left[\frac{1}{n+1}(1 - F_{M_{\mathbf{x}}}(m))^{n+1}\right]^{\mathbb{M}^{\min}} \\
&\quad = \frac{1}{n+1},
\end{aligned}$$

where the last line follows since the range of $\mathbb{M}^{\min}$ is identical to the range of $M_{\mathbf{x}}$. Returning to $\mathbb{P}(\mathbf{x} \in X_2)$,

$$\begin{aligned}
\mathbb{P}(M_{\mathbf{x}} \leq M_{\mathbf{y}}, \forall \mathbf{y} \in Y_2 \cap B_{\mathbb{D}}(\mathbf{x}, R)) &= \sum_{n=0}^{\infty} \frac{\lambda_{\mathbf{x}}^n \mathrm{e}^{-\lambda_{\mathbf{x}}}}{n!} \frac{1}{n+1} \\
&= \frac{\mathrm{e}^{-\lambda_{\mathbf{x}}}}{\lambda_{\mathbf{x}}} \sum_{n=0}^{\infty} \frac{\lambda_{\mathbf{x}}^{n+1}}{(n+1)!} \\
&= \frac{\mathrm{e}^{-\lambda_{\mathbf{x}}}}{\lambda_{\mathbf{x}}} \left[\sum_{n=0}^{\infty} \frac{\lambda_{\mathbf{x}}^n}{n!} - 1 \right] \\
&= \frac{\mathrm{e}^{-\lambda_{\mathbf{x}}}}{\lambda_{\mathbf{x}}} (\mathrm{e}^{\lambda_{\mathbf{x}}} - 1) \\
&= \frac{1 - \mathrm{e}^{-\lambda_{\mathbf{x}}}}{\lambda_{\mathbf{x}}} \\
&= \frac{1 - \mathrm{e}^{-\rho \lambda_{\mathbb{D}}(B_{\mathbb{D}}(\mathbf{x}, R))}}{\rho \lambda_{\mathbb{D}}(B_{\mathbb{D}}(\mathbf{x}, R))}.
\end{aligned}$$

Thus returning to the expectation of $N_{X_2}(B)$,

$$\mathbb{E}[N_{X_2}(B)] = \int_B \frac{1 - \mathrm{e}^{-\rho \lambda_{\mathbb{D}}(B_{\mathbb{D}}(\mathbf{x}, R))}}{\lambda_{\mathbb{D}}(B_{\mathbb{D}}(\mathbf{x}, R))} \lambda_{\mathbb{D}}(d\mathbf{x}).$$

□

B.11 Proof of Corollary 4.6.5

Proof. Notice that since $\rho \lambda_{\mathbb{D}}(B_{\mathbb{D}}(\mathbf{x}, R))$ is always positive this means that $\frac{1-\mathrm{e}^{-\rho \lambda_{\mathbb{D}}(B_{\mathbb{D}}(\mathbf{x}, R))}}{\lambda_{\mathbb{D}}(B_{\mathbb{D}}(\mathbf{x}, R))} \geq 0$, $\forall \rho \in \mathbb{R}^+$ and $\mathbf{x} \in \mathbb{D}$. Then $\frac{1-\mathrm{e}^{-\rho_1 \lambda_{\mathbb{D}}(B_{\mathbb{D}}(\mathbf{x}, R))}}{\lambda_{\mathbb{D}}(B_{\mathbb{D}}(\mathbf{x}, R))} < \frac{1-\mathrm{e}^{-\rho_2 \lambda_{\mathbb{D}}(B_{\mathbb{D}}(\mathbf{x}, R))}}{\lambda_{\mathbb{D}}(B_{\mathbb{D}}(\mathbf{x}, R))}$, $\forall \rho_1 < \rho_2$ and so the supremum is when ρ is taken to infinity giving the final result. □

B.12 Proof of Proposition 4.6.6

Proof. Let $Y = f(X)$, where $f(\mathbf{x}) = \mathbf{x}/||\mathbf{x}||$ then,

$$\mathbb{E}[N_X(\mathbb{D})] = \mathbb{E}\left[\sum_{\mathbf{c} \in X_p} N_{X_{\mathbf{c}}}(\mathbb{S}^2) \right]$$

$$
\begin{aligned}
&= \int_{\mathbb{D}} \mathbb{E}[N_{X_{\mathbf{c}}}(\mathbb{S}^2)]\rho d\mathbf{c} \\
&= \lambda\rho \int_{\mathbb{D}} d\mathbf{c} \\
&= \lambda_{\mathbb{D}}(\mathbb{D})\rho\lambda.
\end{aligned}
$$

□

C

Appendix to Chapter 5

C.1 Proof of Theorem 5.1.1

In this section we shall rederive the expectation and variance of our estimators for $K_{\text{inhom}}(r)$ in the scenario when ρ is unknown. We restate our estimator as,

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

where $Y = f(X)$ and

$$\tilde{\rho}(\mathbf{x}) = \begin{cases} l_1(f^{-1}(\mathbf{x}))J_{(1,f^*)}(\mathbf{x})\sqrt{1-\mathbf{x}_1^2-\mathbf{x}_2^2}, & \mathbf{x}\in f(\mathbb{D}_1) \\ \vdots & \\ l_n(f^{-1}(\mathbf{x}))J_{(n,f^*)}(\mathbf{x})\sqrt{1-\mathbf{x}_1^2-\mathbf{x}_2^2}, & \mathbf{x}\in f(\mathbb{D}_n) \end{cases} \tag{C.2}$$

The proof is then,

Proof. We can rewrite our estimator as,

$$\tilde{K}_{\text{inhom}}(r) = \frac{\mathbb{1}[N_Y(\mathbb{S}^2) > 1]\lambda_{\mathbb{D}}^2(\mathbb{D})}{4\pi N_Y(\mathbb{S}^2)(N_Y(\mathbb{S}^2)-1)} \sum_{\mathbf{x}\in Y} \sum_{\mathbf{y}\in Y\setminus\{x\}} \frac{\mathbb{1}[d(\mathbf{x},\mathbf{y})\leq r]}{\tilde{\rho}(\mathbf{x})\tilde{\rho}(\mathbf{y})} \tag{C.3}$$

Further, note that $N_Y(\mathbb{S}^2) = N_X(\mathbb{D}^2)$, where $f(\mathbf{x}) = \frac{\mathbf{x}}{|\mathbf{x}|}$. We then take expectations (using iterated expectations conditioning on $N_Y(\mathbb{S}^2)$) of $\tilde{K}_{\text{inhom}}(r)$ to get the bias,

$$\begin{aligned}
\mathbb{E}(\tilde{K}_{\text{inhom}}(r)) &= \mathbb{E}\left[\frac{\mathbb{1}[N_Y(\mathbb{S}^2) > 1]\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\{x\}} \frac{\mathbb{1}[d(\mathbf{x},\mathbf{y}) \le r]}{\tilde{\rho}(\mathbf{x})\tilde{\rho}(\mathbf{y})}\right] \\
&= \frac{\lambda_{\mathbb{D}}^2(\mathbb{D})}{4\pi}\mathbb{E}\left(\frac{\mathbb{1}[N_Y(\mathbb{S}^2) > 1]}{N_Y(\mathbb{S}^2)(N_Y(\mathbb{S}^2)-1)}\mathbb{E}\left[\sum_{\mathbf{x},\mathbf{y}\in Y}^{\neq} \frac{\mathbb{1}[d(\mathbf{x},\mathbf{y}) \le r]}{\tilde{\rho}(\mathbf{x})\tilde{\rho}(\mathbf{y})}\middle| N_Y(\mathbb{S}^2) = n\right]\right) \\
&= \frac{\lambda_{\mathbb{D}}^2(\mathbb{D})}{4\pi}\mathbb{E}\left(\frac{\mathbb{1}[N_Y(\mathbb{S}^2) > 1]}{N_Y(\mathbb{S}^2)(N_Y(\mathbb{S}^2)-1)}\mathbb{E}\left[\sum_{\mathbf{i},\mathbf{j}\in\{1,\ldots,n\}}^{\neq} \frac{\mathbb{1}[d(\mathbf{Y}_i,\mathbf{Y}_j) \le r]}{\tilde{\rho}(\mathbf{Y}_i)\tilde{\rho}(\mathbf{Y}_j)}\right]\right) \\
&= \frac{\lambda_{\mathbb{D}}^2(\mathbb{D})}{4\pi}\mathbb{E}\left(\frac{\mathbb{1}[N_Y(\mathbb{S}^2) > 1]N_Y(\mathbb{S}^2)(N_Y(\mathbb{S}^2)-1)}{N_Y(\mathbb{S}^2)(N_Y(\mathbb{S}^2)-1)}\mathbb{E}\left[\frac{\mathbb{1}[d(\mathbf{Y},\mathbf{Y}) \le r]}{\tilde{\rho}(\mathbf{Y})\tilde{\rho}(\mathbf{Y})}\right]\right) \\
&= \frac{\lambda_{\mathbb{D}}^2(\mathbb{D})}{4\pi}P(N_Y(\mathbb{S}^2) > 1)\mathbb{E}\left[\frac{\mathbb{1}[d(\mathbf{Y},\mathbf{Y}) \le r]}{\tilde{\rho}(\mathbf{Y})\tilde{\rho}(\mathbf{Y})}\right] \\
&= \frac{\lambda_{\mathbb{D}}^2(\mathbb{D})}{4\pi}P(N_Y(\mathbb{S}^2) > 1)\int_{\mathbb{S}^2}\int_{\mathbb{S}^2} \frac{\mathbb{1}[d(\mathbf{x},\mathbf{y}) \le r]}{\tilde{\rho}(\mathbf{x})\tilde{\rho}(\mathbf{y})}\frac{\rho^*(\mathbf{x})\rho^*(\mathbf{y})}{\rho^2\lambda_{\mathbb{D}}^2(\mathbb{D})}\lambda_{\mathbb{S}^2}(d\mathbf{x})\lambda_{\mathbb{S}^2}(d\mathbf{y}) \\
&= P(N_Y(\mathbb{S}^2) > 1)2\pi(1-\cos r).
\end{aligned}$$

The bias follows by noting that $P(N_Y(\mathbb{S}^2) > 1) = 1 - P(N_Y(\mathbb{S}^2) \le 1)$. The variance can be calculated using the law of total variance we have,

$$\text{Var}(\tilde{K}_{\text{inhom}}(r)) = \underbrace{\mathbb{E}[\text{Var}(\tilde{K}_{\text{inhom}}(r)|N_Y(\mathbb{S}^2))]}_{(1)} + \underbrace{\text{Var}[\mathbb{E}[\tilde{K}_{\text{inhom}}(r)|N_Y(\mathbb{S}^2)]]}_{(2)}. \tag{C.4}$$

Considering term (2) first,

$$\mathbb{E}[\tilde{K}_{\text{inhom}}(r)|N_Y(\mathbb{S}^2) = n] = \frac{\mathbb{1}[n > 1]\lambda_{\mathbb{D}}^2(\mathbb{D})}{4\pi n(n-1)}\mathbb{E}\left[\sum_{\mathbf{x},\mathbf{y}\in Y}^{\neq} \frac{\mathbb{1}[d(\mathbf{x},\mathbf{y}) \le r]}{\tilde{\rho}(\mathbf{x})\tilde{\rho}(\mathbf{y})}\middle| N_Y(\mathbb{S}^2) = n\right]$$

Using the fact that given $N_Y(\mathbb{S}^2) = n$, each $\mathbf{x} \in Y$ is independently and identically distributed with density $\rho^*(\mathbf{x})/\mu(\mathbb{D})$, where $\mu(\mathbb{D}) = \rho\lambda_{\mathbb{D}}(\mathbb{D})$, the expectation becomes,

$$\begin{aligned}
&= \frac{\mathbb{1}[n > 1]\lambda_{\mathbb{D}}^2(\mathbb{D})}{4\pi n(n-1)}\mathbb{E}\left[\sum_{i,j\in\{1,\ldots,n\}}^{\neq} \frac{\mathbb{1}[d(\mathbf{Y}_i,\mathbf{Y}_j) \le r]}{\tilde{\rho}(\mathbf{Y}_i)\tilde{\rho}(\mathbf{Y}_j)}\middle| N_Y(\mathbb{S}^2) = n\right] \\
&= \frac{\mathbb{1}[n > 1]\lambda_{\mathbb{D}}^2(\mathbb{D})}{4\pi n(n-1)}\sum_{i,j\in\{1,\ldots,n\}}^{\neq} \mathbb{E}\left[\frac{\mathbb{1}[d(\mathbf{Y}_i,\mathbf{Y}_j) \le r]}{\tilde{\rho}(\mathbf{Y}_i)\tilde{\rho}(\mathbf{Y}_j)}\middle| N_Y(\mathbb{S}^2) = n\right]
\end{aligned}$$

$$= \frac{\mathbb{1}[n>1]\lambda_{\mathbb{D}}^2(\mathbb{D})}{4\pi n(n-1)} n(n-1)\mathbb{E}\left[\frac{\mathbb{1}[d(\mathbf{X},\mathbf{Y})\le r]}{\tilde{\rho}(\mathbf{X})\tilde{\rho}(\mathbf{Y})}\right]$$
$$= \frac{\mathbb{1}[n>1]\lambda_{\mathbb{D}}^2(\mathbb{D})}{4\pi}\mathbb{E}\left[\frac{\mathbb{1}[d(\mathbf{X},\mathbf{Y})\le r]}{\tilde{\rho}(\mathbf{X})\tilde{\rho}(\mathbf{Y})}\right],$$

where $\mathbf{X}$ and $\mathbf{Y}$ are independent random vectors distributed in $\mathbb{S}^2$ with density $\rho^*(\mathbf{x})/\mu(\mathbb{D})$. Noting that the joint density of $\mathbf{X}$ and $\mathbf{Y}$ is $\rho^*(\mathbf{x})\rho^*(\mathbf{y})/\mu^2(\mathbb{D})$, the expectation becomes,

$$\begin{aligned}
\mathbb{E}[\tilde{K}_{\text{inhom}}(r)|N_Y(\mathbb{S}^2)=n] &= \frac{\mathbb{1}[n>1]\lambda_{\mathbb{D}}^2(\mathbb{D})}{4\pi}\int_{\mathbb{S}^2}\int_{\mathbb{S}^2}\frac{\mathbb{1}[d(\mathbf{x},\mathbf{y})\le r]}{\tilde{\rho}(\mathbf{y})\tilde{\rho}(\mathbf{y})}\frac{\rho^*(\mathbf{x})\rho^*(\mathbf{y})}{\mu^2(\mathbb{D})}\lambda_{\mathbb{S}^2}(d\mathbf{x})\lambda_{\mathbb{S}^2}(d\mathbf{y}) \\
&= \frac{\mathbb{1}[n>1]\lambda_{\mathbb{D}}^2(\mathbb{D})}{4\pi}\frac{1}{\lambda_{\mathbb{D}}^2(\mathbb{D})}\int_{\mathbb{S}^2}\int_{\mathbb{S}^2}\mathbb{1}[d(\mathbf{x},\mathbf{y})\le r]\lambda_{\mathbb{S}^2}(d\mathbf{x})\lambda_{\mathbb{S}^2}(d\mathbf{y}) \\
&= \frac{\mathbb{1}[n>1]}{4\pi}\int_{\mathbb{S}^2}\int_{\mathbb{S}^2}\mathbb{1}[d(\mathbf{x},\mathbf{y})\le r]\lambda_{\mathbb{S}^2}(d\mathbf{x})\lambda_{\mathbb{S}^2}(d\mathbf{y}) \\
&= \mathbb{1}[n>1]2\pi(1-\cos r)
\end{aligned}$$

Hence term (2) is,

$$\begin{aligned}
\text{Var}[\mathbb{E}[\tilde{K}_{\text{inhom}}(r)|N_Y(\mathbb{S}^2)]] &= \text{Var}(\mathbb{1}[N_Y(\mathbb{S}^2)>1]2\pi(1-\cos r)) \\
&= 4\pi^2(1-\cos r)^2\left[\mathbb{E}(\mathbb{1}^2[N_Y(\mathbb{S}^2)>1]) - \mathbb{E}^2(\mathbb{1}[N_Y(\mathbb{S}^2)>1])\right] \\
&= 4\pi^2(1-\cos r)^2\left[P(N_Y(\mathbb{S}^2)>1) - P^2(N_Y(\mathbb{S}^2)>1)\right] \\
&= 4\pi^2(1-\cos r)^2 P(N_Y(\mathbb{S}^2)>1)\left[1-P(N_Y(\mathbb{S}^2)>1)\right] \\
&= 4\pi^2(1-\cos r)^2(1-P(N_Y(\mathbb{S}^2)\le 1))P(N_Y(\mathbb{S}^2)\le 1),
\end{aligned}$$

where, since $N_Y(\mathbb{S}^2)\sim\text{Poisson}(\rho\lambda_{\mathbb{D}}(\mathbb{D}))$,

$$P(N_Y(\mathbb{S}^2)\le 1) = 1 - e^{-\rho\lambda_{\mathbb{D}}(\mathbb{D})} - \rho\lambda_{\mathbb{D}}(\mathbb{D})e^{-\rho\lambda_{\mathbb{D}}(\mathbb{D})}.$$

Now consider term (1), we calculate $\text{Var}(\tilde{K}_{\text{inhom}}(r)|N_Y(\mathbb{S}^2))$,

$$\begin{aligned}
\text{Var}(\tilde{K}_{\text{inhom}}(r)|N_Y(\mathbb{S}^2)=n) &= \text{Var}\left(\frac{\mathbb{1}[N_Y(\mathbb{S}^2)>1]\lambda_{\mathbb{D}}^2(\mathbb{D})}{4\pi N_Y(\mathbb{S}^2)(N_Y(\mathbb{S}^2)-1)}\sum_{\mathbf{x},\mathbf{y}\in Y}^{\neq}\frac{\mathbb{1}[d(\mathbf{x},\mathbf{y})\le r]}{\tilde{\rho}(\mathbf{x})\tilde{\rho}(\mathbf{y})}\middle| N_Y(\mathbb{S}^2)=n\right) \\
&= \frac{\mathbb{1}^2[n>1]\lambda_{\mathbb{D}}^4(\mathbb{D})}{16\pi^2}\text{Var}\left(\frac{1}{n(n-1)}\sum_{\mathbf{x},\mathbf{y}\in Y}^{\neq}\frac{\mathbb{1}[d(\mathbf{x},\mathbf{y})\le r]}{\tilde{\rho}(\mathbf{x})\tilde{\rho}(\mathbf{y})}\middle| N_Y(\mathbb{S}^2)=n\right) \\
&= \frac{\mathbb{1}[n>1]\lambda_{\mathbb{D}}^4(\mathbb{D})}{16\pi^2}\text{Var}\left(\frac{1}{n(n-1)}\sum_{\mathbf{x},\mathbf{y}\in Y}^{\neq}\frac{\mathbb{1}[d(\mathbf{x},\mathbf{y})\le r]}{\tilde{\rho}(\mathbf{x})\tilde{\rho}(\mathbf{y})}\middle| N_Y(\mathbb{S}^2)=n\right)
\end{aligned}$$

$$= \frac{\mathbb{1}[n > 1]\lambda_{\mathbb{D}}^4(\mathbb{D})}{16\pi^2}\text{Var}\left(\frac{1}{n(n-1)}\sum_{i,j\in\{1,\ldots,n\}}^{\neq}\frac{\mathbb{1}[d(\mathbf{Y}_i,\mathbf{Y}_j)\leq r]}{\tilde{\rho}(\mathbf{Y}_i)\tilde{\rho}(\mathbf{Y}_j)}\right)$$

Here we follow a similar argument to that of Lang and Marcon ((2012)) through U-statistics. Noting that $\frac{\mathbb{1}[d(\mathbf{Y}_i,\mathbf{Y}_j)\leq r]}{\tilde{\rho}(\mathbf{Y}_i)\tilde{\rho}(\mathbf{Y}_j)} = \frac{\mathbb{1}[d(\mathbf{Y}_j,\mathbf{Y}_i)\leq r]}{\tilde{\rho}(\mathbf{Y}_j)\tilde{\rho}(\mathbf{Y}_i)}$, i.e. it is symmetric in its arguments, we rewrite the summation,

$$\begin{aligned}\frac{1}{n(n-1)}\sum_{i,j\in\{1,\ldots,n\}}^{\neq}\frac{\mathbb{1}[d(\mathbf{Y}_i,\mathbf{Y}_j)\leq r]}{\tilde{\rho}(\mathbf{Y}_i)\tilde{\rho}(\mathbf{Y}_j)} &= \frac{1}{n(n-1)}\sum_{1\leq i<j\leq n}\frac{\mathbb{1}[d(\mathbf{Y}_i,\mathbf{Y}_j)\leq r]}{\tilde{\rho}(\mathbf{Y}_i)\tilde{\rho}(\mathbf{Y}_j)} + \frac{\mathbb{1}[d(\mathbf{Y}_j,\mathbf{Y}_i)\leq r]}{\tilde{\rho}(\mathbf{Y}_j)\tilde{\rho}(\mathbf{Y}_i)} \\ &= \frac{2}{n(n-1)}\sum_{1\leq i<j\leq n}\frac{\mathbb{1}[d(\mathbf{Y}_i,\mathbf{Y}_j)\leq r]}{\tilde{\rho}(\mathbf{Y}_i)\tilde{\rho}(\mathbf{Y}_j)} \\ &= \binom{n}{2}^{-1}\sum_{1\leq i<j\leq n}\frac{\mathbb{1}[d(\mathbf{Y}_i,\mathbf{Y}_j)\leq r]}{\tilde{\rho}(\mathbf{Y}_i)\tilde{\rho}(\mathbf{Y}_j)}\end{aligned}$$

This is form of a U-statistic and variances of this class of statistics can be decomposed using the work of Hoeffding ((1992)). Using the same notation as Hoeffding ((1992)), we define some quantities and derive a number of expectations,

$$\begin{aligned}U_n &= \binom{n}{2}^{-1}\sum_{1\leq i<j\leq n}\frac{\mathbb{1}[d(\mathbf{Y}_i,\mathbf{Y}_j)\leq r]}{\tilde{\rho}(\mathbf{Y}_i)\tilde{\rho}(\mathbf{Y}_j)} \\ \Phi(\mathbf{y}_1,\mathbf{y}_2) &= \frac{\mathbb{1}[d(\mathbf{y}_1,\mathbf{y}_2)\leq r]}{\tilde{\rho}(\mathbf{y}_1)\tilde{\rho}(\mathbf{y}_2)} \\ \Phi_1(\mathbf{y}_1) &\equiv \Phi_1(\mathbf{y}_1,\mathbf{Y}_2) = \mathbb{E}[\Phi(\mathbf{y}_1,\mathbf{Y}_2)] = \mathbb{E}[\Phi(\mathbf{Y}_1,\mathbf{Y}_2)|\mathbf{Y}_1=\mathbf{y}_1] \\ &= \mathbb{E}\left[\frac{\mathbb{1}[d(\mathbf{y}_1,\mathbf{Y}_2)\leq r]}{\tilde{\rho}(\mathbf{y}_1)\tilde{\rho}(\mathbf{Y}_2)}\right] \\ &= \int_{\mathbb{S}^2}\frac{\mathbb{1}[d(\mathbf{y}_1,\mathbf{y}_2)\leq r]}{\tilde{\rho}(\mathbf{y}_1)\tilde{\rho}(\mathbf{y}_2)}\frac{\rho^*(\mathbf{y}_2)}{\rho\lambda_{\mathbb{D}}(\mathbb{D})}\lambda_{\mathbb{S}^2}(d\mathbf{y}_2) \\ &= \frac{1}{\tilde{\rho}(\mathbf{y}_1)\lambda_{\mathbb{D}}(\mathbb{D})}\int_{\mathbb{S}^2}\mathbb{1}[d(\mathbf{y}_1,\mathbf{y}_2)\leq r]\lambda_{\mathbb{S}^2}(d\mathbf{y}_2) \\ &= \frac{2\pi(1-\cos r)}{\tilde{\rho}(\mathbf{y}_1)\lambda_{\mathbb{D}}(\mathbb{D})} \\ \Phi_2(\mathbf{y}_1,\mathbf{y}_2) &= \mathbb{E}[\Phi(\mathbf{y}_1,\mathbf{y}_2)] = \mathbb{E}[\phi(\mathbf{Y}_1,\mathbf{Y}_2)|\mathbf{Y}_1=\mathbf{y}_1,\mathbf{Y}_2=\mathbf{y}_2] = \Phi(\mathbf{y}_1,\mathbf{y}_2) \\ \mathbb{E}[\Phi_1(\mathbf{Y}_1)] &= \int_{\mathbb{S}^2}\frac{2\pi(1-\cos r)}{\tilde{\rho}(\mathbf{y}_1)\lambda_{\mathbb{D}}(\mathbb{D})}\frac{\rho^*(\mathbf{y}_1)}{\rho\lambda_{\mathbb{D}}(\mathbb{D})}\lambda_{\mathbb{S}^2}(d\mathbf{y}_1) \\ &= \frac{4\pi\cdot 2\pi(1-\cos r)}{\lambda_{\mathbb{D}}^2(\mathbb{D})}\end{aligned}$$

$$
\begin{aligned}
\mathbb{E}[\Phi_2(\mathbf{Y}_1,\mathbf{Y}_2)] &= \int_{\mathbb{S}^2}\int_{\mathbb{S}^2} \frac{\mathbb{1}[d(\mathbf{y}_1,\mathbf{y}_2)\le r]}{\tilde{\rho}(\mathbf{y}_1)\tilde{\rho}(\mathbf{y}_2)} \frac{\rho^*(\mathbf{y}_1)\rho^*(\mathbf{y}_2)}{\rho^2\lambda_{\mathbb{D}}^2(\mathbb{D})} \lambda_{\mathbb{S}^2}(d\mathbf{y}_1)\lambda_{\mathbb{S}^2}(d\mathbf{y}_2) \\
&= \frac{4\pi\cdot 2\pi(1-\cos r)}{\lambda_{\mathbb{D}}^2(\mathbb{D})} \\
\mathbb{E}[\Phi_1^2(\mathbf{Y}_1)] &= \int_{\mathbb{S}^2} \frac{4\pi^2(1-\cos r)^2}{\tilde{\rho}^2(\mathbf{y}_1)\lambda_{\mathbb{D}}^2(\mathbb{D})} \frac{\rho^*(\mathbf{y}_1)}{\rho\lambda_{\mathbb{D}}(\mathbb{D})} \lambda_{\mathbb{S}^2}(d\mathbf{y}_1) \\
&= \frac{4\pi^2(1-\cos r)^2}{\lambda_{\mathbb{D}}^3(\mathbb{D})} \int_{\mathbb{S}^2} \frac{1}{\tilde{\rho}(\mathbf{y}_1)} \lambda_{\mathbb{S}^2}(d\mathbf{y}_1) \\
\mathbb{E}[\Phi_2^2(\mathbf{Y}_1,\mathbf{Y}_2)] &= \int_{\mathbb{S}^2}\int_{\mathbb{S}^2} \frac{\mathbb{1}[d(\mathbf{y}_1,\mathbf{y}_2)\le r]}{\tilde{\rho}^2(\mathbf{y}_1)\tilde{\rho}^2(\mathbf{y}_2)} \frac{\rho^*(\mathbf{y}_1)\rho^*(\mathbf{y}_2)}{\rho^2\lambda_{\mathbb{D}}^2(\mathbb{D})} \lambda_{\mathbb{S}^2}(d\mathbf{y}_1)\lambda_{\mathbb{S}^2}(d\mathbf{y}_2) \\
&= \frac{1}{\lambda_{\mathbb{D}}^2(\mathbb{D})} \int_{\mathbb{S}^2}\int_{\mathbb{S}^2} \frac{\mathbb{1}[d(\mathbf{y}_1,\mathbf{y}_2)\le r]}{\tilde{\rho}(\mathbf{y}_1)\tilde{\rho}(\mathbf{y}_2)} \lambda_{\mathbb{S}^2}(d\mathbf{y}_1)\lambda_{\mathbb{S}^2}(d\mathbf{y}_2) \\
\zeta_1 &= \mathrm{Var}(\Phi_1(\mathbf{Y}_1)) \\
&= \mathbb{E}[\Phi_1^2(\mathbf{Y}_1)] - \mathbb{E}^2[\Phi_1(\mathbf{Y}_1)] \\
&= \frac{4\pi^2(1-\cos r)^2}{\lambda_{\mathbb{D}}^3(\mathbb{D})} \int_{\mathbb{S}^2} \frac{1}{\tilde{\rho}(\mathbf{y}_1)} \lambda_{\mathbb{S}^2}(d\mathbf{y}_1) - \frac{16\pi^2\cdot 4\pi^2(1-\cos r)^2}{\lambda_{\mathbb{D}}^4(\mathbb{D})} \\
&= \frac{4\pi^2(1-\cos r)^2}{\lambda_{\mathbb{D}}^3(\mathbb{D})} \left(\int_{\mathbb{S}^2} \frac{1}{\tilde{\rho}(\mathbf{y}_1)} \lambda_{\mathbb{S}^2}(d\mathbf{y}_1) - \frac{16\pi^2}{\lambda_{\mathbb{D}}(\mathbb{D})} \right) \\
\zeta_2 &= \mathrm{Var}(\Phi_2(\mathbf{Y}_1,\mathbf{Y}_2)) \\
&= \mathbb{E}[\Phi_2^2(\mathbf{Y}_1,\mathbf{Y}_2)] - \mathbb{E}^2[\Phi_2(\mathbf{Y}_1,\mathbf{Y}_2)] \\
&= \frac{1}{\lambda_{\mathbb{D}}^2(\mathbb{D})} \int_{\mathbb{S}^2}\int_{\mathbb{S}^2} \frac{\mathbb{1}[d(\mathbf{y}_1,\mathbf{y}_2)\le r]}{\tilde{\rho}(\mathbf{y}_1)\tilde{\rho}(\mathbf{y}_2)} \lambda_{\mathbb{S}^2}(d\mathbf{y}_1)\lambda_{\mathbb{S}^2}(d\mathbf{y}_2) - \frac{16\pi^2\cdot 4\pi^2(1-\cos r)^2}{\lambda_{\mathbb{D}}^4(\mathbb{D})} \\
&= \frac{1}{\lambda_{\mathbb{D}}^2(\mathbb{D})} \left(\int_{\mathbb{S}^2}\int_{\mathbb{S}^2} \frac{\mathbb{1}[d(\mathbf{y}_1,\mathbf{y}_2)\le r]}{\tilde{\rho}(\mathbf{y}_1)\tilde{\rho}(\mathbf{y}_2)} \lambda_{\mathbb{S}^2}(d\mathbf{y}_1)\lambda_{\mathbb{S}^2}(d\mathbf{y}_2) - \frac{64\pi^4(1-\cos r)^2}{\lambda_{\mathbb{D}}^2(\mathbb{D})} \right)
\end{aligned}
$$

Then using the variance derived by Hoeffding ((1992)) for U-statistics, the variance of our U_n can be decomposed as,

$$
\begin{aligned}
\mathrm{Var}\left(U_n\right) &= \binom{n}{2}^{-1} \sum_{k=1}^{2} \binom{2}{k}\binom{n-2}{2-k}\zeta_k \\
&= \frac{4(n-2)}{n(n-1)}\zeta_1 + \frac{2}{n(n-1)}\zeta_2
\end{aligned}
$$

Then,

$$\begin{aligned}
&\mathrm{Var}(\tilde{K}_{\mathrm{inhom}}(r)|N_Y(\mathbb{S}^2)=n) = \frac{\mathbb{1}[n>1]\lambda_{\mathbb{D}}^4(\mathbb{D})}{16\pi^2}\left(\frac{4(n-2)}{n(n-1)}\zeta_1 + \frac{2}{n(n-1)}\zeta_2\right)\\
&= \lambda_{\mathbb{D}}(\mathbb{D})(1-\cos r)^2\left(\int_{\mathbb{S}^2}\frac{1}{\tilde{\rho}(\mathbf{y}_1)}\lambda_{\mathbb{S}^2}(d\mathbf{y}_1) - \frac{16\pi^2}{\lambda_{\mathbb{D}}(\mathbb{D})}\right)\cdot\frac{\mathbb{1}[n>1](n-2)}{n(n-1)}\\
&+ \frac{\lambda_{\mathbb{D}}^2(\mathbb{D})}{8\pi^2}\left(\int_{\mathbb{S}^2}\int_{\mathbb{S}^2}\frac{\mathbb{1}[d(\mathbf{y}_1,\mathbf{y}_2)\le r]}{\tilde{\rho}(\mathbf{y}_1)\tilde{\rho}(\mathbf{y}_2)}\lambda_{\mathbb{S}^2}(d\mathbf{y}_1)\lambda_{\mathbb{S}^2}(d\mathbf{y}_2) - \frac{64\pi^4(1-\cos r)^2}{\lambda_{\mathbb{D}}^2(\mathbb{D})}\right)\frac{\mathbb{1}[n>1]}{n(n-1)}
\end{aligned}$$

Taking expectations gives,

$$\begin{aligned}
&\mathbb{E}[\mathrm{Var}(\tilde{K}_{\mathrm{inhom}}(r)|N_Y(\mathbb{S}^2))]\\
&= \lambda_{\mathbb{D}}(\mathbb{D})(1-\cos r)^2\left(\int_{\mathbb{S}^2}\frac{1}{\tilde{\rho}(\mathbf{y}_1)}\lambda_{\mathbb{S}^2}(d\mathbf{y}_1) - \frac{16\pi^2}{\lambda_{\mathbb{D}}(\mathbb{D})}\right)\mathbb{E}\left[\frac{\mathbb{1}[N_Y(\mathbb{S}^2)>1](N_Y(\mathbb{S}^2)-2)}{N_Y(\mathbb{S}^2)(N_Y(\mathbb{S}^2)-1)}\right]\\
&+ \frac{\lambda_{\mathbb{D}}^2(\mathbb{D})}{8\pi^2}\left(\int_{\mathbb{S}^2}\int_{\mathbb{S}^2}\frac{\mathbb{1}[d(\mathbf{y}_1,\mathbf{y}_2)\le r]}{\tilde{\rho}(\mathbf{y}_1)\tilde{\rho}(\mathbf{y}_2)}\lambda_{\mathbb{S}^2}(d\mathbf{y}_1)\lambda_{\mathbb{S}^2}(d\mathbf{y}_2) - \frac{64\pi^4(1-\cos r)^2}{\lambda_{\mathbb{D}}^2(\mathbb{D})}\right)\mathbb{E}\left[\frac{\mathbb{1}[N_Y(\mathbb{S}^2)>1]}{N_Y(\mathbb{S}^2)(N_Y(\mathbb{S}^2)-1)}\right]
\end{aligned}$$

The expectations can be simplified as follows and defining $\lambda = \rho\lambda_{\mathbb{D}}(\mathbb{D})$,

$$\begin{aligned}
\mathbb{E}\left[\frac{\mathbb{1}[N_Y(\mathbb{S}^2)>1](N_Y(\mathbb{S}^2)-2)}{N_Y(\mathbb{S}^2)(N_Y(\mathbb{S}^2)-1)}\right] &= \sum_{n=0}^{\infty}\frac{\mathbb{1}[n>1](n-2)}{n(n-1)}\frac{\lambda^n e^{-\lambda}}{n!}\\
&= \sum_{n=3}^{\infty}\frac{(n-2)}{n(n-1)}\frac{\lambda^n e^{-\lambda}}{n!}\\
&= \sum_{n=3}^{\infty}\frac{1}{n^2(n-1)^2}\frac{\lambda^n e^{-\lambda}}{(n-3)!}\\
&= \sum_{n=0}^{\infty}\frac{1}{(n+3)^2(n+2)^2}\frac{\lambda^{n+3} e^{-\lambda}}{n!}\\
&= \lambda^3\sum_{n=0}^{\infty}\frac{1}{(n+3)^2(n+2)^2}\frac{\lambda^n e^{-\lambda}}{n!}\\
&= \lambda^3\mathbb{E}\left[\frac{1}{(N_Y(\mathbb{S}^2)+3)^2(N_Y(\mathbb{S}^2)+2)^2}\right]
\end{aligned}$$

Similarly the other expectation is,

$$\mathbb{E}\left[\frac{\mathbb{1}[N_Y(\mathbb{S}^2)>1]}{N_Y(\mathbb{S}^2)(N_Y(\mathbb{S}^2)-1)}\right] = \lambda^2\mathbb{E}\left[\frac{1}{(N_Y(\mathbb{S}^2)+2)^2(N_Y(\mathbb{S}^2)+1)^2}\right],$$

and so,

$$\begin{aligned}
&\mathbb{E}[\mathrm{Var}(\tilde{K}_{\mathrm{inhom}}(r)|N_Y(\mathbb{S}^2))] \\
&\quad = \rho^3\lambda_{\mathbb{D}}^4(\mathbb{D})(1-\cos r)^2\left(\int_{\mathbb{S}^2}\frac{1}{\tilde{\rho}(\mathbf{y}_1)}\lambda_{\mathbb{S}^2}(d\mathbf{y}_1) - \frac{16\pi^2}{\lambda_{\mathbb{D}}(\mathbb{D})}\right)\mathbb{E}\left[\frac{1}{(N_Y(\mathbb{S}^2)+3)^2(N_Y(\mathbb{S}^2)+2)^2}\right] \\
&\quad + \frac{\rho^2\lambda_{\mathbb{D}}^4(\mathbb{D})}{8\pi^2}\left(\int_{\mathbb{S}^2}\int_{\mathbb{S}^2}\frac{\mathbb{1}[d(\mathbf{y}_1,\mathbf{y}_2)\le r]}{\tilde{\rho}(\mathbf{y}_1)\tilde{\rho}(\mathbf{y}_2)}\lambda_{\mathbb{S}^2}(d\mathbf{y}_1)\lambda_{\mathbb{S}^2}(d\mathbf{y}_2) - \frac{64\pi^4(1-\cos r)^2}{\lambda_{\mathbb{D}}^2(\mathbb{D})}\right) \\
&\qquad \times \mathbb{E}\left[\frac{1}{(N_Y(\mathbb{S}^2)+2)^2(N_Y(\mathbb{S}^2)+1)^2}\right].
\end{aligned}$$

Combining everything gives the variance of $\tilde{K}_{\mathrm{inhom}}(r)$,

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

□

C.2 Proof of Lemma 5.1.2

Proof. Define $S = N!e^{N-k}/(N-k)!$ and $T = (e+p)^N$, then $R = S/T$. Then,

$$\begin{aligned}
\mathbb{E}[S] &= \sum_{n=0}^{\infty}\frac{n!e^{n-k}}{(n-k)!}\frac{\lambda^n e^{-\lambda}}{n!} \\
&= e^{-k}e^{-\lambda}\sum_{n=k}^{\infty}\frac{(e\lambda)^n}{(n-k)!} \\
&= e^{-k}e^{-\lambda}\sum_{l=0}^{\infty}\frac{(e\lambda)^{l+k}}{l!} \\
&= e^{-k}e^{-\lambda}(e\lambda)^k\sum_{l=0}^{\infty}\frac{(e\lambda)^l}{l!} \\
&= e^{-k}e^{-\lambda}(e\lambda)^k e^{e\lambda} \\
&= \lambda^k e^{\lambda(e-1)}.
\end{aligned}$$

Further,

$$\begin{aligned}\mathbb{E}[T] &= \sum_{n=0}^{\infty}(e+p)^n \frac{\lambda^n e^{-\lambda}}{n!} \\ &= e^{-\lambda}\sum_{n=0}^{\infty}\frac{(\lambda(e+p))^n}{n!} \\ &= e^{-\lambda}e^{\lambda(e+p)} \\ &= e^{\lambda(e+p-1)}.\end{aligned}$$

Therefore,

$$\frac{\mathbb{E}[S]}{\mathbb{E}[T]} = \frac{\lambda^k e^{\lambda(e-1)}}{e^{\lambda(e+p-1)}} = \lambda^k e^{-p\lambda}$$

□

C.3 Decreasing power

We see, in Tables 6.1-6.3, that as a becomes smaller, and therefore c becomes larger a reducing empirical power of our test for both regular and cluster processes, for the same R and κ respectively. This effect could be due to a combination of mapping from the ellipsoid to the sphere and an artefact of the test statistic being proposed over a finite grid of points rather than being consider over the entire range of $[0, \pi]$. To see this consider just the Matérn II process with a fixed hardcore distance R. Further, let us only consider a cross section of the ellipsoid, more specifically the ellipse such that it's major and minor axis lengths are c and a respectively, see Figure C.1. Let us consider the point $(0, a)$ lying on the ellipse and take the point to the right of it which is precisely the hardcore distance R away from it, let us label this point $\mathbf{x}$. To find $\mathbf{x}$ use the parametrisation $(x, y) = (c\cos t, a\sin t)$ for $t \in [0, 2\pi)$. Then we solve the following equation for t to find $\mathbf{x}$,

$$R = \int_t^{\pi/2} (c^2\sin^2 s + a^2\cos^2 s)^{1/2} ds.$$

Let us label $\tilde{t}$ as the solution to this equation and we then find $\mathbf{x}$ as $(c\cos\tilde{t}, a\sin\tilde{t})$. Then apply the map onto the unit circle which gives us $(0, a) \mapsto (0, 1)$ and $(c\cos\tilde{t}, a\sin\tilde{t}) \mapsto (\cos\tilde{t}, \sin\tilde{t})$ and so our new hardcore distance on the circle is $R' = \cos^{-1}(\sin\tilde{t})$. It should be noted that this calculation is dependent upon where on an ellipsoid the event of interest, $\mathbf{x}$ is as $B_{\mathbb{D}}(\mathbf{x}, R)$ does not map to $B_{\mathbb{D}}(f(\mathbf{x}), r)$ for some $r > 0$ and where f is our mapping from

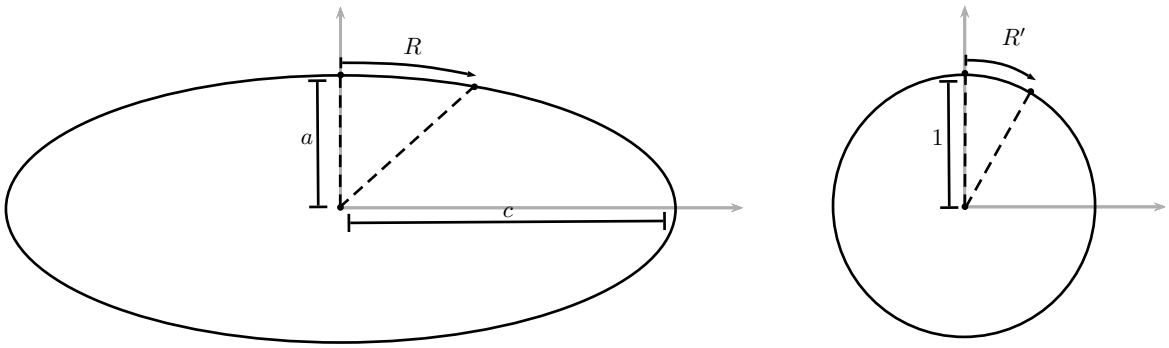


Figure C.1: Example of hardcore distance reduction due to mapping to a sphere.

the ellipsoid to the sphere. Using the example when $a = 0.4000$ we have that $c = 3.1602$ and thus for a hardcore distance of $R = 0.2$, on the cross sectional ellipse we have an effective hardcore distance of $R' = 0.0633$, which is an extremely small hardcore distance and since our finite grid of points is too coarse (points are only separated a distance of 0.02 apart) it results in a loss in power of our test. Furthermore, examining the standardised inhomogeneous K-function in Figure C.2 we can further see the effect of our mapping and taking only a finite grid of points along $[0, \pi]$ as the negative deviation reduces as ellipsoid *deforms* further away from the unit sphere. It should be noted though that even though the power of our test reduces as we move further away from the sphere Figure C.2 still indicates evidence that of regularity as for small r for all ellipsoids the observed inhomogeneous K-function falls below the simulation envelope. This highlights the importance of a proper examination of graphical representations of functional summary statistics as opposed to the use of formal hypothesis testing Diggle ((2003)). Another consideration would be to potential use a two sided hypothesis testing procedure which may provide greater power when the true underlying process is not CSR.

A similar effect also occurs when examining cluster processes. By mapping the pattern from the ellipsoid to the sphere we will distort the isotropic nature of our offspring density relative to its parent. In particular this will cause the cluster size to contract and so if our finite grid of points is too coarse we will struggle to detect deviations away from CSR.

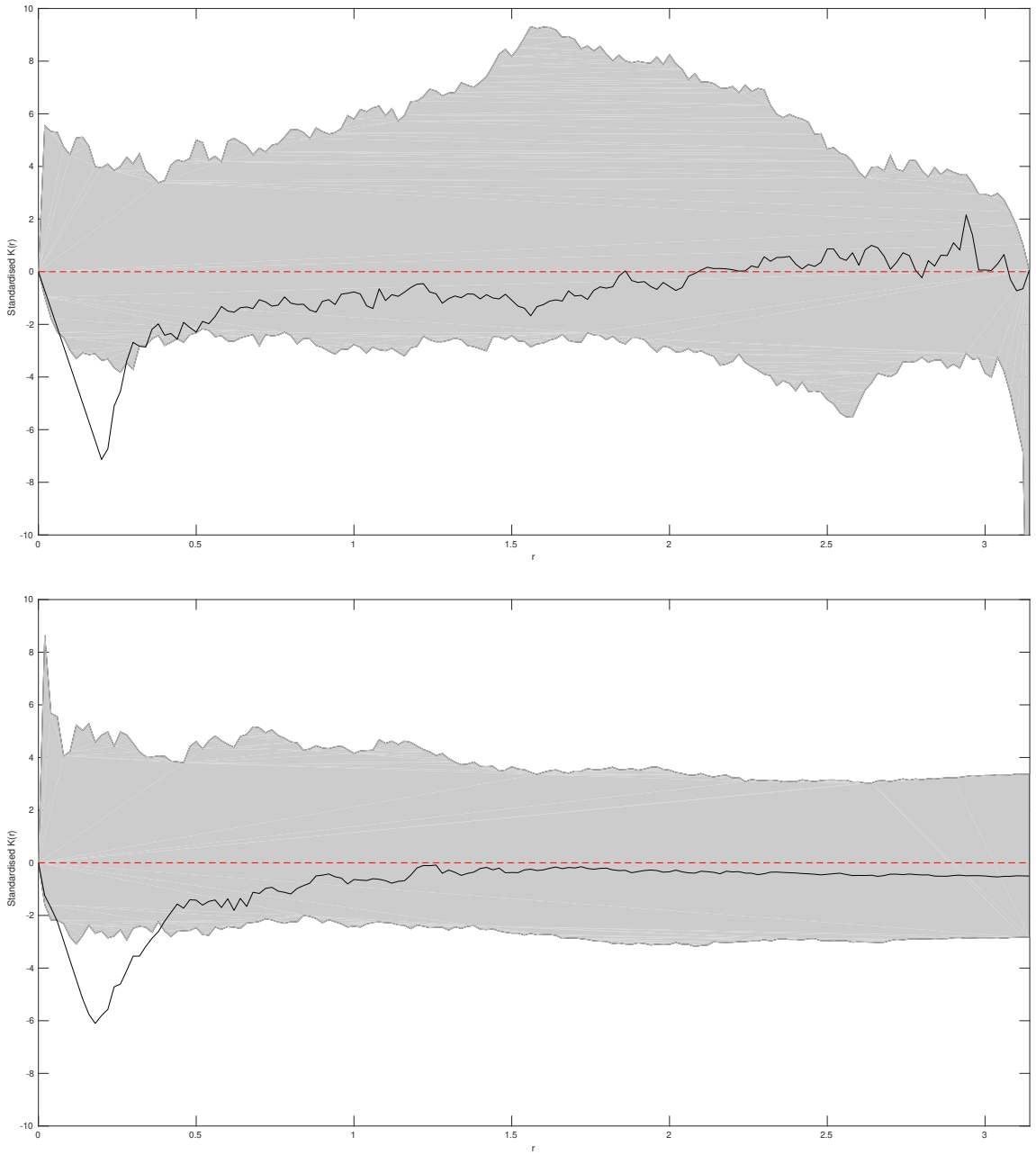


Figure continued on following page.

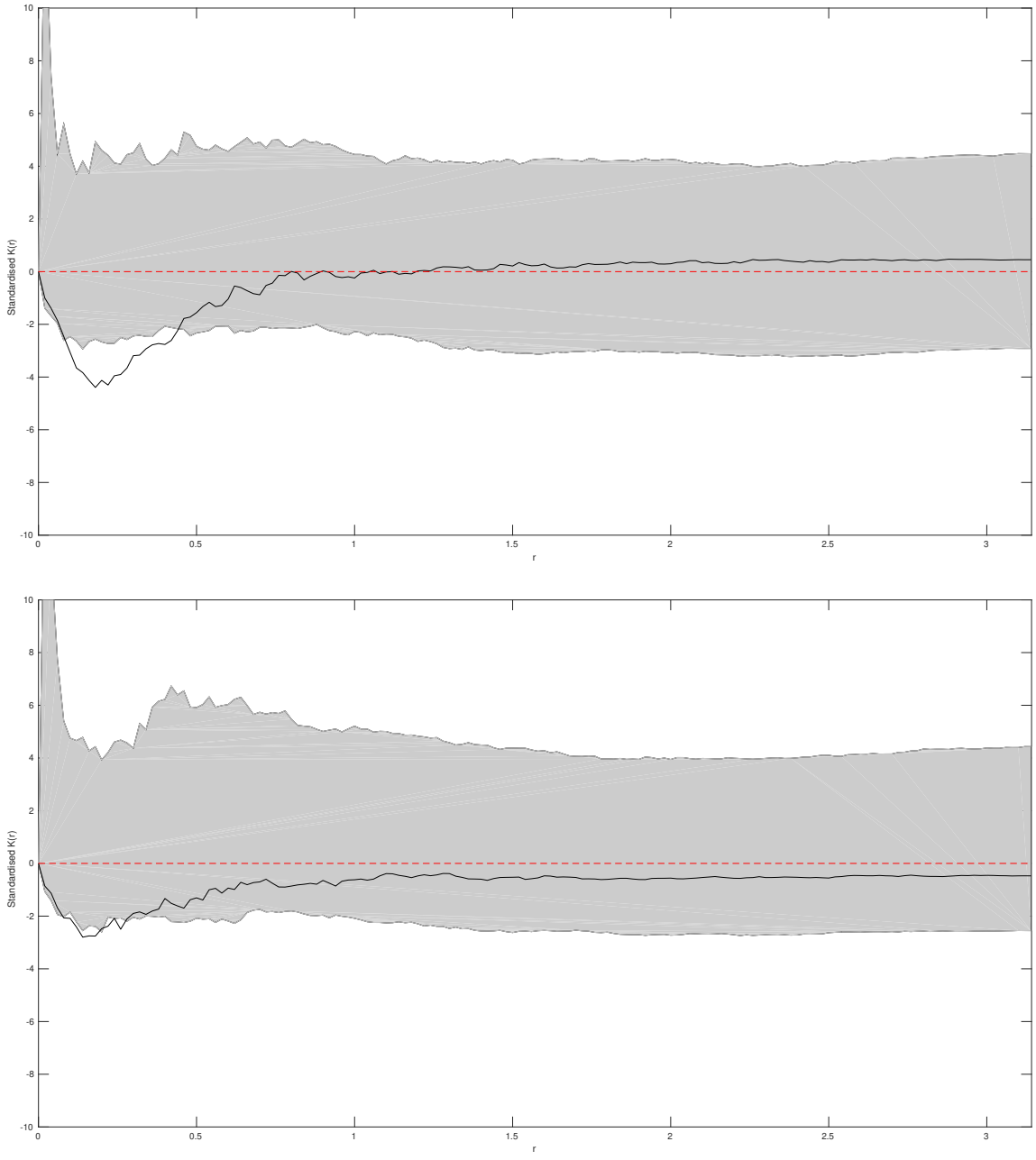


Figure C.2: Plots of the standardised inhomogeneous K-function for a Matérn II process with hardcore distance $R = 0.2$. From top left to bottom right $a = 1$ (sphere), $a = 0.8, a = 0.6$ and $a = 0.4$. Black line is the estimated standardised $\hat{K}_{\text{inhom}}(r)$ for our observed data, dashed red line is the theoretical functional summary statistic for a Poisson process, and the grey shaded area represents the simulation envelope from 999 Monte Carlo simulations of Poisson processes fitted to the observed data.

D

Appendix to Chapter 6

D.1 Limit properties of edge correction factor

The following lemma is discusses the asymptotics of the edge correction factors as h goes to either 0 or infinity. It is used in the proofs of Proposition 6.3.1 and Theorem 6.4.2.

Lemma D.1.1. *Suppose that $(\mathcal{M}, g)$ is a Riemannian manifold then the edge correction factors given by Equation 6.8 and 6.9 with a Gaussian kernel (see Equation 6.10) we have the following limit properties when considered as a function of h for fixed $\mathbf{x}, \mathbf{y} \in \mathcal{M}$,*

$$c_h(\mathbf{x}, \mathbf{y}) \to \frac{K}{(2\pi)^{n/2}} \quad as \quad h \to 0,$$
$$h^n c_h(\mathbf{x}, \mathbf{y}) \to \frac{Vol(W)}{(2\pi)^{n/2}} \quad as \quad h \to \infty,$$

where $K = \int_{\mathbb{R}^d} \exp(-||\mathbf{x}||^2/2) d\mathbf{x} < \infty$.

Before providing the proof we make the following remark. If we supposed that our Riemannian manifold is embedded within some Euclidean space and we then take g to be the canonical Riemannian metric, for example the $d-1$ dimensional sphere in $\mathbb{R}^d$ inherits the Euclidean metric in $\mathbb{R}^d$, then $||\mathbf{x}||^2$ is identically the Euclidean norm squared and so $K = (2\pi)^{n/2}$ and so,

$$c_h(\mathbf{x}, \mathbf{y}) \to 1 \quad \text{as} \quad h \to 0$$

which is equivalent to the limiting behaviour as $h \to 0$ of the edge correction terms in Euclidean space given by Cronie and Van Lieshout ((2018)). Moreover, we can also retrieve the limiting behaviour as $h \to \infty$ if we supposed that $\mathcal{M}$ is $\mathbb{R}^d$ endowed with its Euclidean metric then trivially $\text{Vol}(W) = \lambda_{\mathbb{R}^d}(W)$ and so we recover the results of Cronie and Van Lieshout ((2018)). The proof of Lemma D.1.1 is now given.

Proof. Since the edge correction factors are effectively symmetric in $\mathbf{x}$ and $\mathbf{y}$ we need only consider one of Equations 6.8 and 6.9: we will use Equation 6.8. We first recall that by Gauss's Lemma Carmo ((1992)), the exponential map is radially isometric. We also require the so-called *volume density function* Besse ((1978)), Pelletier ((2005)), Henry and Rodriguez ((2009)). The volume density function, $\theta_{\mathbf{p}}(\mathbf{q}) : \mathcal{M} \mapsto \mathbb{R}$, can be defined as the Radon-Nikodyn derivative of the pullback of the volume form by $\exp_{\mathbf{p}}$ with respect to the Lebesgue measure, $d\mathbf{x} = dx^1 \wedge \cdots dx^d$, over the tangent space defined by $g_{\mathbf{p}}$ on $T_{\mathbf{p}}\mathcal{M}$,

$$\theta_{\mathbf{p}}(\mathbf{q}) = \frac{d \exp_{\mathbf{p}}^* \text{vol}(\mathbf{x})}{d\mathbf{x}}(\log_{\mathbf{p}} \mathbf{q}),$$

see Besse ((1978)) for more details on the volume density function. Moreover, the volume density function can be represented in locally in normal coordinates, that is the local chart $(B_{\mathcal{M}}(\mathbf{p}, r), \log_{\mathbf{p}})$, where $r < r^*$ and r^* is the global injectivity radius. Then in this coordinate system we have that,

$$\theta_{\mathbf{p}}(\mathbf{q}) = \det(g_{ij}(\log_{\mathbf{p}}(\mathbf{q})))^{1/2},$$

where g_{ij} is the metric tensor g expressed in the local coordinate system given by $log_{\mathbf{p}}$ Pelletier ((2005)), Henry and Rodriguez ((2009)). Since g is smooth over $\mathcal{M}$ this implies that $\theta_{\mathbf{p}}$ is also smooth and $g_{ij}(\mathbf{p}) = \delta_{ij}$ and so $\theta_{\mathbf{p}}(\mathbf{p}) = 1$. By defining the volume density function we can integrate over local regions of $\mathcal{M}$ by mapping to the tangent space at $\mathbf{p}$, i.e. suppose f is an integrable function over $B_{\mathcal{M}}(\mathbf{p}, r)$ where $r < r^*$ and let $(B_{\mathcal{M}}(\mathbf{p}, r), \log_{\mathbf{p}})$ be the normal coordinates at $\mathbf{p}$ then,

$$\int_{B_{\mathcal{M}}(\mathbf{p},r)} f \, d\text{vol} = \int_{\log_{\mathbf{p}}(B_{\mathcal{M}}(\mathbf{p},r))} ((\exp_{\mathbf{p}})^* f) \, d \exp_{\mathbf{p}}^* \text{vol} = \int_{B_{T_{\mathbf{p}}\mathcal{M}}(\mathbf{0},r)} (f \circ \exp_{\mathbf{p}}) \cdot (\theta_{\mathbf{p}} \circ \exp_{\mathbf{p}}) \, d\mathbf{x}$$

where, by Gauss's lemma, we have that $\log_{\mathbf{p}}(B_{\mathcal{M}}(\mathbf{p}, r)) = B_{T_{\mathbf{p}}\mathcal{M}}(\mathbf{0}, r)$. We now show the result of the lemma, first dividing the integral of $c_h(\mathbf{x}, \mathbf{y})$ into two parts where we take $r \leq r^*$,

$$c_h(\mathbf{x}, \mathbf{y}) = \frac{1}{(2\pi)^{d/2}} \int_{\mathcal{M}} \frac{1}{h^d} \exp\left(\frac{-d_g^2(\mathbf{x}, \mathbf{z})}{2h^2}\right) d\text{vol}(\mathbf{z})$$

$$= \frac{1}{(2\pi)^{d/2}} \int_{\mathcal{M} \setminus B_{\mathcal{M}}(\mathbf{x},r)} \frac{1}{h^d} \exp\left(\frac{-d_g^2(\mathbf{x},\mathbf{z})}{2h^2}\right) d\text{vol}(\mathbf{z}) + \frac{1}{(2\pi)^{d/2}} \int_{B_{\mathcal{M}}(\mathbf{x},r)} \frac{1}{h^d} \exp\left(\frac{-d_g^2(\mathbf{x},\mathbf{z})}{2h^2}\right) d\text{vol}(\mathbf{z}). \quad \text{(D.1)}$$

The first term can easily be shown to go to 0 as h goes to 0 as follows,

$$\frac{1}{(2\pi)^{d/2}} \int_{\mathcal{M} \setminus B_{\mathcal{M}}(\mathbf{x},r)} \frac{1}{h^d} \exp\left(\frac{-d_g^2(\mathbf{x},\mathbf{z})}{2h^2}\right) d\text{vol}(\mathbf{z}) \leq \frac{1}{h^d} \int_{\mathcal{M} \setminus B_{\mathcal{M}}(\mathbf{x},r)} \exp\left(\frac{-r^2}{2h^2}\right) d\text{vol}(\mathbf{z}) = \frac{\text{Vol}(\mathcal{M} \setminus B_{\mathcal{M}}(\mathbf{x},r))}{h^d} \exp\left(\frac{-r^2}{2h^2}\right),$$

where the second line follows since $\mathcal{M}$ is compact and thus $\text{Vol}(\mathcal{M})$ is finite. The right hand side tends to 0 as h goes to 0 and hence the limit of this term is also 0 since it must be greater than or equal to 0.

Turning to the second term of Equation D.1 we have that,

$$\frac{1}{(2\pi)^{d/2}} \int_{B_{\mathcal{M}}(\mathbf{x},r)} \frac{1}{h^d} \exp\left(\frac{-d_g^2(\mathbf{x},\mathbf{z})}{2h^2}\right) d\text{vol}(\mathbf{z}) = \frac{1}{(2\pi)^{d/2}} \int_{B_{T_{\mathbf{x}}\mathcal{M}}(\mathbf{0},r)} \frac{1}{h^d} \exp\left(\frac{-d_g^2(\mathbf{x},\exp_{\mathbf{x}}(\mathbf{z}))}{2h^2}\right) \theta_{\mathbf{x}}(\exp_{\mathbf{x}}(\mathbf{z})) d\mathbf{z},$$

by radial isometry we have that $d_g^2(\mathbf{x},\exp_{\mathbf{x}}(\mathbf{z})) = ||\mathbf{z}||^2$ and so,

$$= \frac{1}{(2\pi)^{d/2}} \int_{B_{T_{\mathbf{x}}\mathcal{M}}(\mathbf{0},r)} \frac{1}{h^d} \exp\left(\frac{-||\mathbf{z}||^2}{2h^2}\right) \theta_{\mathbf{x}}(\exp_{\mathbf{x}}(\mathbf{z})) d\mathbf{z}$$

define $K = \int_{\mathbb{R}^d} \exp(-||\mathbf{x}||^2/2) d\mathbf{x}$ and for the time being assume $K < \infty$

$$= \frac{K}{(2\pi)^{d/2}} \int_{B_{T_{\mathbf{x}}\mathcal{M}}(\mathbf{0},r)} \frac{1}{Kh^d} \exp\left(\frac{-||\mathbf{z}||^2}{2h^2}\right) \theta_{\mathbf{x}}(\exp_{\mathbf{x}}(\mathbf{z})) d\mathbf{z}$$

and hence we notice that $\lim_{h\to 0}(1/Kh^d) \exp\left(-||\mathbf{z}||^2/2h^2\right)$ takes the form of a Dirac - delta. Since $\theta_{\mathbf{x}}(\exp_{\mathbf{x}}(\mathbf{z}))$ depends smoothly on $\mathbf{z}$, that $\theta_{\mathbf{x}}(\exp_{\mathbf{x}}(\mathbf{0})) = \theta_{\mathbf{x}}(\mathbf{x}) = 1$ and is defined over the compact space of $B_{T_{\mathbf{x}}\mathcal{M}}(\mathbf{0},r)$ we can use the standard Dirac-delta property $\int_{\mathbb{R}^d} f(\mathbf{x})\delta(d\mathbf{x}) = f(\mathbf{0})$ where δ is the Dirac-delta function to give,

$$= \frac{K}{(2\pi)^{d/2}} \theta_{\mathbf{x}}(\mathbf{0})$$

$$= \frac{K}{(2\pi)^{d/2}},$$

and hence our result for Riemannian manifolds assuming that $K < \infty$, which is indeed true by the Nash embedding theorem Nash ((1956)). To properly apply Nash's theorem we first need to define what an isometric embedding is. Let $F : \mathcal{M} \mapsto \mathcal{N}$ be a smooth function between Riemannian manifolds $(\mathcal{M}, g_1)$ and $(\mathcal{N}, g_2)$, F is then an isometric embedding of $\mathcal{M}$ in $\mathcal{N}$ if it preserves the metric, that is for all $\mathbf{p} \in \mathcal{M}$ and $X_\mathbf{p}, Y_\mathbf{p} \in T_\mathbf{p}\mathcal{M}$ then,

$$(g_1)_\mathbf{p}(X_\mathbf{p}, Y_\mathbf{p}) = (g_2)_{F(\mathbf{p})}(F_{*,\mathbf{p}}X_\mathbf{p}, F_{*,\mathbf{p}}Y_\mathbf{p}).$$

As such Nash's embedding theorem states that any compact d-dimensional Riemannain manifold $(\mathcal{M}, g)$ without boundary can be isometrically embedded into $\mathbb{R}^m$ for some $m \in \mathbb{N}$ under its canonical metric, i.e. the Euclidean metric. Let $F : \mathcal{M} \mapsto \mathbb{R}^m$ be the isometric embedding as defined by Nash's theorem and let $T_{F(\mathbf{p})}\mathbb{R}^m$ be the tangent space at $F(p)$. Let $(U, x^1, \ldots, x^d)$ be a local coordinate system with $\mathbf{p} \in U \subset \mathcal{M}$, we can then write $X_\mathbf{p} = \sum_{i=1}^d a^i \partial/\partial x^i|_\mathbf{p}$ to be a vector in the tangent space at $\mathbf{p}$ written using the basis imposed by the local coordinates. Let $(V, y^1, \ldots, y^m)$ be the standard coordinate system of $\mathbb{R}^m$, set $F^i = y^i \circ F$ be the i^{th} component of F and $F(\mathbf{p}) \in V$ then we can write $F_{*,\mathbf{p}}X_\mathbf{p}$ as ((Tu, 2011, Proposition 8.11)),

$$F_{*,\mathbf{p}}X_\mathbf{p} = \sum_{j=1}^m \left(\sum_{i=1}^d a^i \left. \frac{\partial F^j}{\partial x^i} \right|_\mathbf{p} \right) \left. \frac{\partial}{\partial y^j} \right|_{F(\mathbf{p})}.$$

Letting e be the canonical Euclidean metric on $\mathbb{R}^m$ and since $y^1, \ldots, y^m$ is our standard coordinate system we have that,

$$\begin{aligned}
&(e)_{F(\mathbf{p})}(F_{*,\mathbf{p}}X_\mathbf{p}, F_{*,\mathbf{p}}Y_\mathbf{p}) \\
&\qquad = \begin{pmatrix} a^1 & \cdots & a^d \end{pmatrix} \begin{pmatrix} \left.\frac{\partial F^1}{\partial x^1}\right|_\mathbf{p} & \cdots & \left.\frac{\partial F^m}{\partial x^1}\right|_\mathbf{p} \\ \vdots & \ddots & \vdots \\ \left.\frac{\partial F^1}{\partial x^d}\right|_\mathbf{p} & \cdots & \left.\frac{\partial F^m}{\partial x^d}\right|_\mathbf{p} \end{pmatrix} \begin{pmatrix} \left.\frac{\partial F^1}{\partial x^1}\right|_\mathbf{p} & \cdots & \left.\frac{\partial F^1}{\partial x^d}\right|_\mathbf{p} \\ \vdots & \ddots & \vdots \\ \left.\frac{\partial F^m}{\partial x^1}\right|_\mathbf{p} & \cdots & \left.\frac{\partial F^m}{\partial x^d}\right|_\mathbf{p} \end{pmatrix} \begin{pmatrix} b^1 \\ \vdots \\ b^d \end{pmatrix} \\
&\qquad = \mathbf{v}_{X_\mathbf{p}} J_F^T J_F \mathbf{v}_{Y_\mathbf{p}},
\end{aligned}$$

where the matrix $J_F = [\partial F^i/\partial x^j]$ corresponds to the Jacobian of the transformation F, $Y_\mathbf{p} = \sum_{i=1}^d b^i \partial/\partial x^i|_\mathbf{p}$ and $\mathbf{v}_{X_\mathbf{p}}$ and $\mathbf{v}_{Y_\mathbf{p}}$ are the representations of $X_\mathbf{p}$ and $Y_\mathbf{p}$ in the local

coordinate system imposed by $x^1, \ldots, x^d$. Therefore by Nash's theorem we have that,

$$g_{\mathbf{p}}(X_{\mathbf{p}}, Y_{\mathbf{p}}) = (e)_{F(\mathbf{p})}(F_{*,\mathbf{p}}X_{\mathbf{p}}, F_{*,\mathbf{p}}Y_{\mathbf{p}}) = \mathbf{v}_{X_{\mathbf{p}}} J_F^T J_F \mathbf{v}_{Y_{\mathbf{p}}}.$$

Therefore

$$\begin{aligned} K &= \int_{\mathbb{R}^d} \exp\left(-\frac{||\mathbf{x}||^2}{2}\right) d\mathbf{x} \\ &= \int_{\mathbb{R}^d} \exp\left(-\frac{\mathbf{x}^T J_F^T J_F \mathbf{x}}{2}\right) d\mathbf{x} \\ &= (2\pi)^{d/2} \sqrt{\det\left((J_F^T J_F)^{-1}\right)} \\ &= \left(\frac{(2\pi)^d}{\det\left(J_F^T J_F\right)}\right)^{1/2} \end{aligned}$$

where the last line follows if $J_F^T J_F$ is positive definite.

To show the limiting behaviour as $h \to \infty$ we follow a similar arugment to Cronie and Van Lieshout ((2018)) noting that $\lim_{h\to\infty} k(d_g(\mathbf{x}, \mathbf{y})/h) = (2\pi)^{-d/2}$ by Equation 6.10. Therefore,

$$\begin{aligned} \lim_{h\to\infty} h^d c_h(\mathbf{x}, \mathbf{y}) &= \lim_{h\to\infty} \int_W \frac{1}{(2\pi)^{d/2}} \exp\left(-\frac{d_g(\mathbf{x}, \mathbf{y})}{2h^2}\right) d\text{vol}(\mathbf{z}) \\ &= \int_W \frac{1}{(2\pi)^{d/2}} \left(\lim_{h\to\infty} \exp\left(-\frac{d_g(\mathbf{x}, \mathbf{y})}{2h^2}\right)\right) d\text{vol}(\mathbf{z}) \\ &= \int_W \frac{1}{(2\pi)^{d/2}} d\text{vol}(\mathbf{z}) \\ &= \frac{\text{Vol}(W)}{(2\pi)^{d/2}}, \end{aligned}$$

where the second line follows by the dominated convergence theorem.

□

D.2 Proof of Proposition 6.3.1

In order to show pointwise unbiasedness and consistency we require the following identities given by Cucala ((2006)) translated to the setting of Riemannian manifolds,

$$\mathbb{E}[Z] = \left(1 - \mathrm{e}^{-\mu(W)}\right) \int_W f(\mathbf{x}) \rho_1(\mathbf{x}) \, d\text{vol}(\mathbf{x}) \tag{D.2}$$

$$\mathrm{Var}(Z) = A(\mu(W)) \int_W f^2(\mathbf{x})\rho_1(\mathbf{x})\, d\mathrm{vol}(\mathbf{x}) \\ - \left(\int_W f(\mathbf{x})\rho_1(\mathbf{x})\, d\mathrm{vol}(\mathbf{x})\right)\left(A(\mu(W)) - \mathrm{e}^{-\mu(W)} - \mathrm{e}^{-2\mu(W)}\right), \tag{D.3}$$

where $Z = (\mathbb{1}[N_X(W) \neq 0]/N_X(W)) \sum_{\mathbf{x}\in X\cap W} f(\mathbf{x})$, X is a Poisson process with intensity function ρ, $A(m) = \mathbb{E}[(1/N_X(W))\mathbb{1}[N_X(W) \neq 0], \rho_1(\mathbf{x}) = \rho(\mathbf{x})/\mu(\mathcal{M})$ and $f : \mathcal{M} \mapsto \mathbb{R}$ is measurable and nonnegative. We now present the proof of Proposition 6.3.1.

Proof. We first show asymptotic unbiasedness. For $i = 1, 2$ and from Equation D.2 we have that,

$$\begin{aligned}
\mathbb{E}\left[\hat{\rho}^{(i)}_{h,1}(\mathbf{x})\right] &= \left(1 - \mathrm{e}^{-\mu(W)}\right) \int_{\mathcal{M}} \mathbb{1}[\mathbf{y} \in W] \frac{c_h^{-1}(\mathbf{x}, \mathbf{y})}{h^d} k\left(\frac{d_g(\mathbf{x}, \mathbf{y})}{h}\right) \rho_1(\mathbf{y})\, d\mathrm{vol}(\mathbf{y}) \\
&= \left(1 - \mathrm{e}^{-\mu(W)}\right) \left[\int_{\mathbb{B}_{\mathcal{M}}(\mathbf{x},r)} \mathbb{1}[\mathbf{y} \in W] \frac{c_h^{-1}(\mathbf{x}, \mathbf{y})}{h^d} k\left(\frac{d_g(\mathbf{x}, \mathbf{y})}{h}\right) \rho_1(\mathbf{x})\, d\mathrm{vol}(\mathbf{y}) \right. \\
&\qquad \left. + \int_{\mathcal{M}\setminus\mathbb{B}_{\mathcal{M}}(\mathbf{x},r)} \mathbb{1}[\mathbf{y} \in W] \frac{c_h^{-1}(\mathbf{x}, \mathbf{y})}{h^d} k\left(\frac{d_g(\mathbf{x}, \mathbf{y})}{h}\right) \rho_1(\mathbf{y})\, d\mathrm{vol}(\mathbf{y})\right]
\end{aligned}$$

where $0 < r < r^*$, $r*$ being the global injectivity radius of $\mathcal{M}$. Applying an identical argument to that used in the proof of Lemma D.1.1 to show that the first term of Equation D.1 goes to 0, we can show that the second term here also goes to 0 as h goes to 0. Thus we shall focus on the integral in the first term,

$$\begin{aligned}
&\int_{\mathbb{B}_{\mathcal{M}}(\mathbf{x},r)} \mathbb{1}[\mathbf{y} \in W] \frac{c_h^{-1}(\mathbf{x}, \mathbf{y})}{h^d} k\left(\frac{d_g(\mathbf{x}, \mathbf{y})}{h}\right) \rho_1(\mathbf{y})\, d\mathrm{vol}(\mathbf{y}) \\
&= \int_{B_{T_{\mathbf{x}}\mathcal{M}}(\mathbf{0},r)} \mathbb{1}[\exp_{\mathbf{x}}(\mathbf{y}) \in W] \frac{c_h^{-1}(\mathbf{x}, \exp_{\mathbf{x}}(\mathbf{y}))}{h^d} k\left(\frac{d_g(\mathbf{x}), \exp_{\mathbf{x}}(\mathbf{y})}{h}\right) \rho_1(\exp_{\mathbf{x}}(\mathbf{y}))\theta_{\mathbf{x}}(\exp_{\mathbf{x}}(\mathbf{y}))d\mathbf{y} \\
&= \int_{B_{T_{\mathbf{x}}\mathcal{M}}(\mathbf{0},r)} \mathbb{1}[\exp_{\mathbf{x}}(\mathbf{y}) \in W] \frac{c_h^{-1}(\mathbf{x}, \exp_{\mathbf{x}}(\mathbf{y}))}{(2\pi)^{d/2}h^d} \exp\left(-\frac{d_g^2(\mathbf{x}, \exp_{\mathbf{x}}(\mathbf{y}))}{2h^2}\right) \rho_1(\exp_{\mathbf{x}}(\mathbf{y}))\theta_{\mathbf{x}}(\exp_{\mathbf{x}}(\mathbf{y}))\, d\mathbf{y} \\
&= \int_{B_{T_{\mathbf{x}}\mathcal{M}}(\mathbf{0},r)} \frac{c_h^{-1}(\mathbf{x}, \exp_{\mathbf{x}}(\mathbf{y}))}{(2\pi)^{d/2}h^d} \exp\left(-\frac{||\mathbf{y}||^2}{2h^2}\right) \mathbb{1}[\exp_{\mathbf{x}}(\mathbf{y}) \in W]\rho_1(\exp_{\mathbf{x}}(\mathbf{y}))\theta_{\mathbf{x}}(\exp_{\mathbf{x}}(\mathbf{y}))\, d\mathbf{y} \\
&= \int_{B_{T_{\mathbf{x}}\mathcal{M}}(\mathbf{0},r)} \frac{c_h^{-1}(\mathbf{x}, \exp_{\mathbf{x}}(\mathbf{y}))}{(2\pi)^{d/2}h^d} \exp\left(-\frac{||\mathbf{y}||^2}{2h^2}\right) g(\mathbf{y})\, d\mathbf{y},
\end{aligned}$$

where $g(\mathbf{y}) = \mathbb{1}[\exp_{\mathbf{x}}(\mathbf{y}) \in W]\rho_1(\exp_{\mathbf{x}}(\mathbf{y}))\theta_{\mathbf{x}}(\exp_{\mathbf{x}}(\mathbf{y}))$. By Lemma D.1.1 for any $0 < \epsilon < K/((2\pi)^{d/2})$ we can find a small enough h' such that for $h < h'$, $K/((2\pi)^{d/2}) - \epsilon < c_h(\mathbf{x}, \mathbf{y}) < K/((2\pi)^{d/2}) + \epsilon$. Setting $K' = K/((2\pi)^{d/2})$ we can bound the previous integral, denoting it

$I(h)$, as

$$\int_{B_{T_{\mathbf{x}}\mathcal{M}}(\mathbf{0},r)} \frac{1}{(2\pi)^{d/2}h^d(K'+\epsilon)} \exp\left(-\frac{||\mathbf{y}||^2}{2h^2}\right) g(\mathbf{y})\, d\mathbf{y} < I(h) \\ < \int_{B_{T_{\mathbf{x}}\mathcal{M}}(\mathbf{0},r)} \frac{1}{(2\pi)^{d/2}h^d(K'-\epsilon)} \exp\left(-\frac{||\mathbf{y}||^2}{2h^2}\right) g(\mathbf{y})\, d\mathbf{y}. \tag{D.4}$$

Taking the left hand side of the previous inequality we have that,

$$\begin{aligned}\int_{B_{T_{\mathbf{x}}\mathcal{M}}(\mathbf{0},r)} \frac{1}{(2\pi)^{d/2}h^d(K'+\epsilon)} \exp\left(-\frac{||\mathbf{y}||^2}{2h^2}\right) g(\mathbf{y})\, d\mathbf{y} \\ &= \frac{K}{(2\pi)^{d/2}(K'+\epsilon)} \int_{B_{T_{\mathbf{x}}\mathcal{M}}(\mathbf{0},r)} \frac{1}{Kh^d} \exp\left(-\frac{||\mathbf{y}||^2}{2h^2}\right) g(\mathbf{y})\, d\mathbf{y} \\ &= \frac{K'}{(K'+\epsilon)} \int_{B_{T_{\mathbf{x}}\mathcal{M}}(\mathbf{0},r)} \frac{1}{Kh^d} \exp\left(-\frac{||\mathbf{y}||^2}{2h^2}\right) g(\mathbf{y})\, d\mathbf{y}\end{aligned}$$

Then as h goes to 0, $\frac{1}{Kh^d}\exp\left(-||\mathbf{y}||^2/2h^2\right)$ acts as a Dirac-delta and so we have as $h \to 0$ and noting that $g(\mathbf{0}) = \mathbb{1}[\exp_{\mathbf{x}}(\mathbf{0}) \in W]\rho_1(\exp_{\mathbf{x}}(\mathbf{0}))\theta_{\mathbf{x}}(\exp_{\mathbf{x}}(\mathbf{0})) = \rho_1(\mathbf{x})$,

$$\frac{K'}{(K'+\epsilon)} g(\mathbf{0}) = \frac{K'}{(K'+\epsilon)} g(\mathbf{0}) = \frac{K'}{(K'+\epsilon)} \rho_1(\mathbf{x}).$$

We can use a similar argument for the right hand side of Inequality D.4 and therefore have,

$$\frac{K'}{(K'+\epsilon)} \rho_1(\mathbf{x}) < \lim_{h\to 0} I(h) < \frac{K'}{(K'-\epsilon)} \rho_1(\mathbf{x}),$$

for $h < h'$. Then since we can make ϵ arbitrarily small we must have that,

$$\lim_{h\to 0} I(h) = \rho_1(\mathbf{y}).$$

Then since $1 - e^{-\mu(W)}$ goes to 1 as $\mu(W)$ goes to infinity we have asymptotic unbiasedness.

To show that the variance goes to 0 take Equation D.3 giving,

$$\operatorname{Var}\left(\hat{\rho}^{(i)}_{h,1}(\mathbf{x})\right) = A(\mu(W)) \int_{\mathcal{M}} \mathbb{1}[\mathbf{y} \in W] \frac{c_h^{-2}(\mathbf{x},\mathbf{y})}{h^{2d}} k^2\left(\frac{d_W(\mathbf{x},\mathbf{y})}{h}\right) \rho_1(\mathbf{y})\, d\mathrm{vol}(\mathbf{y})$$
$$- \left(\int_{\mathcal{M}} \mathbb{1}[\mathbf{y} \in W] \frac{c_h^{-1}(\mathbf{x},\mathbf{y})}{h^{n}} k\left(\frac{d_g(\mathbf{x},\mathbf{y})}{h}\right) \rho_1(\mathbf{y})\, d\mathrm{vol}(\mathbf{y})\right)\left(A(\mu(W)) - \mathrm{e}^{-\mu(W)} - \mathrm{e}^{-2\mu(W)}\right). \tag{D.5}$$

Taking the first bracket in the second term notice that,

$$\int_{\mathcal{M}} \mathbb{1}[\mathbf{y} \in W] \frac{c_h^{-1}(\mathbf{x},\mathbf{y})}{h^{n}} k\left(\frac{d_g(\mathbf{x},\mathbf{y})}{h}\right) \rho_1(\mathbf{y})\, d\mathrm{vol}(\mathbf{y}) = \frac{\mathbb{E}\left[\hat{\rho}^{(i)}_{h,1}(\mathbf{x})\right]}{1-\mathrm{e}^{-\mu(W)}},$$

and hence as $h \to 0$ and $\mu(W) \to \infty$ this integral is of $O(1)$, i.e. it is bounded in this limit. Examining $A(\mu(W))$ we have,

$$\begin{aligned}
A(\mu(W)) &= \mathbb{E}\left[\frac{\mathbb{1}[N_X(W) \neq 0]}{N_X(W)}\right] \\
&= \sum_{n=1}^{\infty} \frac{\mathrm{e}^{-\mu(W)}(\mu(W))^n}{n \cdot n!} \\
&\leq \sum_{n=1}^{\infty} \frac{2\mathrm{e}^{-\mu(W)}(\mu(W))^n}{(n+1)!} \\
&= \frac{2\mathrm{e}^{-\mu(W)}}{\mu(W)} \sum_{n=1}^{\infty} \frac{(\mu(W))^{n+1}}{(n+1)!} \\
&= \frac{2\mathrm{e}^{-\mu(W)}}{\mu(W)} \sum_{n=2}^{\infty} \frac{(\mu(W))^{n}}{n!} \\
&\leq \frac{2\mathrm{e}^{-\mu(W)}}{\mu(W)} \sum_{n=0}^{\infty} \frac{(\mu(W))^{n}}{n!} = \frac{2}{\mu(W)},
\end{aligned}$$

where the first inequality follows since $k \geq 1 \Rightarrow 1/k \leq 2/(k+1)$. Then $A(\mu(W))$ goes to zero as $\mu(W)$ goes to infinity and hence the second term of Equation D.5 goes to 0 as h goes to 0 and $\mu(W)$ goes to infinity.

Then taking the integral in the first term of Equation D.5 we can split the integral. By Lemma D.1.1, for any $0 < \epsilon < \min(r^*, K/((2\pi)^{d/2}))$, we have that there exists a h' such that for $h < h'$ we have $0 < K/((2\pi)^{d/2}) - \epsilon < c_h(\mathbf{x},\mathbf{y})$ and so taking $h < h'$ we can split into to over $B_{\mathcal{M}}(\mathbf{x},h)$ and $\mathcal{M} \setminus B_{\mathcal{M}}(\mathbf{x},h)$, and like previous arguments the integral over $\mathcal{M} \setminus B_{\mathcal{M}}(\mathbf{x},r)$ goes to zero. Thus focusing on the integral over $B_{\mathcal{M}}(\mathbf{x},h)$ and defining

$K' = K/((2\pi)^{d/2})$ we have,

$$\begin{aligned}
&\int_{B_{\mathcal{M}}(\mathbf{x},h)} \mathbb{1}[\mathbf{y} \in W] \frac{c_h^{-2}(\mathbf{x},\mathbf{y})}{h^{2d}} k^2\left(\frac{d_g(\mathbf{x},\mathbf{y})}{h}\right) \rho_1(\mathbf{y})\, d\text{vol}(\mathbf{y}) \\
&\quad = \int_{B_{T_{\mathbf{x}}\mathcal{M}}(\mathbf{0},h)} \mathbb{1}[\exp_{\mathbf{x}}(\mathbf{y}) \in W] \frac{c_h^{-2}(\mathbf{x},\exp_{\mathbf{x}}(\mathbf{y}))}{(2\pi)^{d/2}h^{2d}} \exp\left(-\frac{||\mathbf{y}||^2}{h^2}\right) \rho_1(\exp_{\mathbf{x}}(\mathbf{y}))\theta_{\mathbf{x}}(\exp_{\mathbf{x}}(\mathbf{y}))\, d\mathbf{y} \\
&\quad \leq \frac{1}{(K'-\epsilon)^2} \int_{B_{T_{\mathbf{x}}\mathcal{M}}(\mathbf{0},h)} \frac{\mathbb{1}[\exp_{\mathbf{x}}(\mathbf{y}) \in W]}{(2\pi)^{d/2}h^{2d}} \exp\left(-\frac{||\mathbf{y}||^2}{h^2}\right) \rho_1(\exp_{\mathbf{x}}(\mathbf{y}))\theta_{\mathbf{x}}(\exp_{\mathbf{x}}(\mathbf{y}))\, d\mathbf{y} \\
&\quad \leq \frac{1}{(K'-\epsilon)^2} \int_{B_{T_{\mathbf{x}}\mathcal{M}}(\mathbf{0},h)} \frac{1}{(2\pi)^{d/2}h^{2d}} \exp\left(-\frac{||\mathbf{y}||^2}{h^2}\right) \rho_1(\exp_{\mathbf{x}}(\mathbf{y}))\theta_{\mathbf{x}}(\exp_{\mathbf{x}}(\mathbf{y}))\, d\mathbf{y}
\end{aligned}$$

using the change of variables $\mathbf{z} = \mathbf{y}/h$

$$\begin{aligned}
&= \frac{1}{(K'-\epsilon)^2} \int_{B_{T_{\mathbf{x}}\mathcal{M}}(\mathbf{0},1)} \frac{h^d}{(2\pi)^{d/2}h^{2d}} \exp\left(-||\mathbf{z}||^2\right) \rho_1(\exp_{\mathbf{x}}(h\mathbf{z}))\theta_{\mathbf{x}}(\exp_{\mathbf{x}}(h\mathbf{z}))\, d\mathbf{z} \\
&= \frac{1}{h^d(K'-\epsilon)^2} \int_{B_{T_{\mathbf{x}}\mathcal{M}}(\mathbf{0},1)} \frac{1}{(2\pi)^{d/2}} \exp\left(-||\mathbf{z}||^2\right) \rho_1(\exp_{\mathbf{x}}(h\mathbf{z}))\theta_{\mathbf{x}}(\exp_{\mathbf{x}}(h\mathbf{z}))\, d\mathbf{z}.
\end{aligned}$$

Notice that both $\rho_1(\exp_{\mathbf{x}}(h\mathbf{z}))$ and $\theta_{\mathbf{x}}(\exp_{\mathbf{x}}(h\mathbf{z}))$ are bounded for small h since $\lim_{h\to 0} \rho_1(\exp_{\mathbf{x}}(h\mathbf{z})) = \rho_1(\mathbf{x})$ and $\lim_{h\to 0} \theta_{\mathbf{x}}(\exp_{\mathbf{x}}(h\mathbf{z})) = 1$ and $\exp(-||\mathbf{z}||^2) \leq \exp(-||\mathbf{0}||) = 1$ means that the integral is of $O(1)$ as $h \to 0$ and thus the first term of Equation D.5 is,

$$A(\mu(W)) \int_{\mathcal{M}} \mathbb{1}[\mathbf{y} \in W] \frac{c_h^{-2}(\mathbf{x},\mathbf{y})}{h^{2d}} k^2\left(\frac{d_W(\mathbf{x},\mathbf{y})}{h}\right) \rho_1(\mathbf{y})\, d\text{vol}(\mathbf{y}) \leq P\frac{A(\mu(W))}{h^d},$$

where P is some real number independent of h and $\mu(W)$. Thus the variance goes to 0 if $A(\mu(W))/h^d \to 0$ as $h \to 0$ and $\mu(W) \to \infty$.

□

D.3 Proof of Lemma 6.4.1

Proof. It is trivial that if the $c_h(\mathbf{x},\mathbf{y}) = 1$ then it is continuous. Now the edge correction factors given by Equations 6.8 and 6.9 are symmetric so we only need to consider one of them: we shall focus on the global one edge correction, i.e. Equation 6.8. Fix $\epsilon, h_0 > 0$ and define $f_{\mathbf{x}}(\mathbf{z},h) = (1/(2\pi)^{d/2}) \exp(-d_g^2(\mathbf{x},\mathbf{z})/(2h^2))$. Then for fixed $\mathbf{x},\mathbf{z} \in \mathcal{M}$, $f_{\mathbf{x}}$ is continuous in h, by continuity of $1/h^2$, exp and that the composition of continuous functions are also continuous. Thus by definition of continuity, for $\epsilon' = \epsilon/\text{Vol}(\mathcal{M})$ there exists $\delta > 0$

such that,

$$|h - h_0| < \delta \Rightarrow |g_{\mathbf{x}}(\mathbf{z}, h) - g_{\mathbf{x}}(\mathbf{z}, h_0)| < \epsilon' = \frac{\epsilon}{\text{Vol}(W)}.$$

Then notice that for fixed $\mathbf{x} \in \mathcal{M}$,

$$\begin{aligned}
\left| \int_W g_{\mathbf{x}}(\mathbf{z}, h)\, d\text{vol}(\mathbf{z}) - \int_W g_{\mathbf{x}}(\mathbf{z}, h_0)\, d\text{vol}(\mathbf{z}) \right| &= \left| \int_W g_{\mathbf{x}}(\mathbf{z}, h) - g_{\mathbf{x}}(\mathbf{z}, h_0)\, d\text{vol}(\mathbf{z}) \right| \\
&\leq \int_W |g_{\mathbf{x}}(\mathbf{z}, h) - g_{\mathbf{x}}(\mathbf{z}, h_0)|\ d\text{vol}(\mathbf{z}) \\
&< \int_W \frac{\epsilon}{\text{Vol}(W)}\, d\text{vol}(\mathbf{z}) \\
&= \epsilon,
\end{aligned}$$

and so by definition $\int_W g_{\mathbf{x}}(\mathbf{z}, h)\, d\text{vol}(\mathbf{z})$ is continuous in h. Therefore, since $h \mapsto 1/h^d$ is continuous and the product of continuous functions is also continuous this implies that $c_h(\mathbf{x}, \mathbf{y})$ is continuous over $h \in (0, \infty)$. □

D.4 Proof of Theorem 6.4.2

Proof. Parts of our proof follow similarly to that of Cronie and Van Lieshout ((2018)). To show continuity of T it is clear that $\hat{\rho}_h$ is continuous by Lemma 6.4.1. Further, since we ignore X being the empty set we have that $\hat{\rho}_h(\mathbf{x}) \geq h^{-n} k(d_g(\mathbf{x}, \mathbf{x})/h) c_h^{-1}(\mathbf{x}, \mathbf{y}) > 0$ and hence taken the reciprical of $\hat{\rho}_h(\mathbf{x})$ is well-defined and more importantly continuous, hence T is continuous.

We now let h go to 0. Then we have for $\mathbf{x} \in X$ and $i = 1, 2$,

$$\begin{aligned}
\hat{\rho}_h^{(i)}(\mathbf{x}) &= \sum_{\mathbf{y} \in X \cap W} \frac{1}{((2\pi)^{d/2}) h^d} \exp\left(\frac{-d_g^2(\mathbf{x}, \mathbf{y})}{2h^2} \right) c_h^{-1}(\mathbf{x}, \mathbf{y}) \\
&\geq \frac{c_h^{-1}(\mathbf{x}, \mathbf{x})}{h^d} \exp\left(\frac{-d_g^2(\mathbf{x}, \mathbf{x})}{2h^2} \right) = \frac{1}{h^d c_h(\mathbf{x}, \mathbf{x})},
\end{aligned}$$

and so,

$$\Rightarrow \frac{1}{\hat{\rho}_h(\mathbf{x})} \leq h^d c_h(\mathbf{x}, \mathbf{x}),$$

where the second line follows since it is a sum of positive real numbers and so we remove all

terms except for $\mathbf{y} = \mathbf{x}$ since $\mathbf{x} \in X$. Thus we have that,

$$T(\hat{\rho}_h) \leq h^d \sum_{\mathbf{x} \in X \cap W} c_h(\mathbf{x}, \mathbf{x}),$$

Thus given an observed X by Lemma D.1.1 all our edge correction terms go to $K/()(2\pi)^{d/2})$ for $K = \int_{\mathbb{R}^d} \exp(-||\mathbf{x}||^2) d\mathbf{x}$ which is bounded and thus we have that $T(\hat{\rho}_h)$ goes to 0 as h goes to 0.

Now suppose that h goes to infinity then,

$$\frac{c_h^{-1}(\mathbf{x}, \mathbf{y})}{(2\pi)^{d/2} h^d} \exp\left(\frac{-d_g^2(\mathbf{x}, \mathbf{y})}{2h^2}\right) \to \begin{cases} \frac{1}{\text{Vol}(W)}, & \text{if } c_h(\mathbf{x}, \mathbf{y}) \text{ is given by Equation 6.8 or 6.9} \\ 0, & \text{if } c_h(\mathbf{x}, \mathbf{y}) = 1, \end{cases}$$

by Lemma D.1.1 and so,

$$T(\hat{\rho}_h) \to \begin{cases} \sum_{\mathbf{x} \in X} \frac{1}{\sum_{\mathbf{y} \in X} 1/\text{Vol}(W)} = \text{Vol}(W), & \text{if } c_h(\mathbf{x}, \mathbf{y}) \text{ is given by Equation 6.8 or 6.9} \\ \infty, & \text{if } c_h(\mathbf{x}, \mathbf{y}) = 1. \end{cases}$$

□

E

Appendix to Chapter 7

E.1 Proof of Proposition 7.3.1

Proof. This proof follows nearly identical to ((Møller and Waagepetersen, 2003, Proposition 4.5)). Clearly $J^{CE}(r) = 1$ if $D^{CE}(r) = F^{E}(r)$. We can rewrite D^{CE} as an expectation over the point process X rather then its reduced Palm process $X^{!}_{\mathbf{o},C}$. Let our finite reference measure ν over $\mathcal{M}$ coincide with the mark distribution and so under isotropy we have that $\rho(\mathbf{x}, m) = \rho$ and A be an arbitrary subset of $\mathbb{S}^2$ such that $\lambda_{\mathbb{S}^2}(A) > 0$ then,

$$
\begin{aligned}
D^{CE}(r) &= P(X^{!}_{\mathbf{o},C} \cap (B_{\mathbb{S}^2}(\mathbf{o}, r) \times E) \neq \emptyset) \\
&= P(X^{!}_{\mathbf{x},C} \cap (B_{\mathbb{S}^2}(\mathbf{x}, r) \times E) \neq \emptyset) && \text{(E.1)} \\
&= \frac{1}{\lambda_{\mathbb{S}^2}(A)} \int_A P(X^{!}_{\mathbf{x},C} \cap (B_{\mathbb{S}^2}(\mathbf{x}, r) \times E) \neq \emptyset) \lambda_{\mathbb{S}^2}(d\mathbf{x}) \\
&= \frac{1}{\lambda_{\mathbb{S}^2}(A)\nu(C)} \int_A \int_C P(X^{!}_{(\mathbf{x},m)} \cap (B_{\mathbb{S}^2}(\mathbf{x}, r) \times E) \neq \emptyset) \lambda_{\mathbb{S}^2}(d\mathbf{x}) \nu(dm) && \text{(E.2)} \\
&= \frac{1}{\rho\lambda_{\mathbb{S}^2}(A)\nu(C)} \int_A \int_C P(X^{!}_{(\mathbf{x},m)} \cap (B_{\mathbb{S}^2}(\mathbf{x}, r) \times E) \neq \emptyset) \rho\lambda_{\mathbb{S}^2}(d\mathbf{x}) \nu(dm) \\
&= \frac{1}{\rho\lambda_{\mathbb{S}^2}(A)\nu(C)} \mathbb{E} \sum_{(\mathbf{x},m) \in X} \mathbb{1}[(\mathbf{x}, m) \in A \times C, (X \setminus (\mathbf{x}, m)) \cap (B_{\mathbb{S}^2}(\mathbf{x}, r) \times E) \neq \emptyset], \\
& && \text{(E.3)}
\end{aligned}
$$

where E.1 follows by isotropy of X, E.2 follows by definition of $P^!_{\mathbf{x},C}$ and E.3 by the Campbell-Mecke theorem. Let us define $X_{A\times C} = X \cap (A \times C)$ for $A \subset \mathbb{S}^2$ and $C \subset \mathcal{M}$, then this expectation can be rewritten as,

$$
\begin{aligned}
&= \frac{1}{\rho\lambda_{\mathbb{S}^2}(A)\nu(C)}\mathbb{E}\sum_{(\mathbf{x},m)\in X_{A\times C}} \mathbb{1}[X_{B_{\mathbb{S}^2}(\mathbf{x},r)\times E} \neq \emptyset] \\
&= \frac{1}{\rho\lambda_{\mathbb{S}^2}(A)\nu(C)}\mathbb{E}\sum_{(\mathbf{x},m)\in X_{A\times C}} \mathbb{E}\left[\mathbb{1}[X_{B_{\mathbb{S}^2}(\mathbf{x},r)\times E} \neq \emptyset]|X_c\right] \\
&= \frac{1}{\rho\lambda_{\mathbb{S}^2}(A)\nu(C)}\mathbb{E}\sum_{(\mathbf{x},m)\in X_{A\times C}} P(X_{B_{\mathbb{S}^2}(\mathbf{x},r)\times E} \neq \emptyset) \\
&= \frac{1}{\rho\lambda_{\mathbb{S}^2}(A)\nu(C)}\mathbb{E}\sum_{(\mathbf{x},m)\in X_{A\times C}} F^E(r) \\
&= \frac{\rho\lambda_{\mathbb{S}^2}(A)\nu(C)}{\rho\lambda_{\mathbb{S}^2}(A)\nu(C)}F^E(r) \\
&= F^E(r).
\end{aligned}
$$

To show that $K^{CE}(r) = 2\pi(1-\cos(r))$ we note that under independence $\rho^{(2)}((\mathbf{x}, m_{\mathbf{x}}), (\mathbf{y}, m_{\mathbf{y}})) = \rho(\mathbf{x}, m_{\mathbf{x}})\rho(\mathbf{y}, m_{\mathbf{y}})$ when $m_1 \in C$ and $m_2 \in E$ and so the result can be shown by application of the Campbell theorem and noting that $\lambda_{\mathbb{S}^2}(B_{\mathbb{S}^2}(\mathbf{o}, r)) = 2\pi(1 - \cos(r))$. □

E.2 Proof of Proposition E.2

Proof. Starting with $\hat{F}^E$ we have that,

$$
\begin{aligned}
\mathbb{E}[1 - \hat{F}^E(r)] &= \frac{1}{|I_{W_{\ominus r}}|}\sum_{\mathbf{p}\in I_{W_{\ominus r}}} \mathbb{E}\prod_{(\mathbf{x},m)\in X}\left(1 - \mathbb{1}[d_{\mathbb{S}^2}(\mathbf{p}, \mathbf{x}) \leq r, m \in E]\right) \\
&= \frac{1}{|I_{W_{\ominus r}}|}\sum_{\mathbf{p}\in I_{W_{\ominus r}}} P(X \cap (B_{\mathbb{S}^2}(\mathbf{p}, r) \times E) = \emptyset) \\
&= \frac{1}{|I_{W_{\ominus r}}|}\sum_{\mathbf{p}\in I_{W_{\ominus r}}} P(X \cap (B_{\mathbb{S}^2}(\mathbf{o}, r) \times E) = \emptyset) \\
&= P(X \cap (B_{\mathbb{S}^2}(\mathbf{o}, r) \times E) = \emptyset) \\
&= 1 - F^E(r).
\end{aligned}
$$

For $\hat{D}^{CE}$ and by the Campbell-Mecke Theorem,

$$\begin{aligned}
\mathbb{E}[1-\hat{D}^{CE}(r)] &= \mathbb{E}\frac{1}{\rho_g \lambda_{\mathbb{S}^2}(W_{\ominus r})\nu(C)} \\
&\qquad \sum_{(\mathbf{x},m_{\mathbf{x}})\in X_{W\ominus r}} \mathbb{1}[(\mathbf{x},m_{\mathbf{x}})\in W_{\ominus r}\times C] \prod_{(\mathbf{y},m_{\mathbf{y}})\in X}\Big(1-\mathbb{1}[d_{\mathbb{S}^2}(\mathbf{x},\mathbf{y})<r, m_{\mathbf{y}}\in E]\Big) \\
&= \frac{1}{\rho_g \lambda_{\mathbb{S}^2}(W_{\ominus r})\nu(C)} \int_{W_{\ominus r}}\int_{C} \\
&\qquad \mathbb{E}^{!}_{(\mathbf{x},m)} \prod_{(\mathbf{y},m_{\mathbf{y}})\in X}\Big(1-\mathbb{1}[d_{\mathbb{S}^2}(\mathbf{x},\mathbf{y})<r, m_{\mathbf{y}}\in E]\Big)\rho(\mathbf{x},m)\lambda_{\mathbb{S}^2}(d\mathbf{x})\nu(m)
\end{aligned}$$

since X is isotropic and we have taken the reference measure ν to be the probability measure over $\mathcal{M}$ we have that $\rho(\mathbf{x},m)=\rho_g$ and so

$$\begin{aligned}
&= \frac{1}{\lambda_{\mathbb{S}^2}(W_{\ominus r})\nu(C)} \int_{W_{\ominus r}}\int_{C} \mathbb{E}^{!}_{(\mathbf{x},m)} \prod_{(\mathbf{y},m_{\mathbf{y}})\in X}\Big(1-\mathbb{1}[d_{\mathbb{S}^2}(\mathbf{x},\mathbf{y})<r, m_{\mathbf{y}}\in E]\Big)\lambda_{\mathbb{S}^2}(d\mathbf{x})\nu(m) \\
&= \frac{1}{\lambda_{\mathbb{S}^2}(W_{\ominus r})} \int_{W_{\ominus r}} \mathbb{E}^{!}_{(\mathbf{x},C)} \prod_{(\mathbf{y},m_{\mathbf{y}})\in X}\Big(1-\mathbb{1}[d_{\mathbb{S}^2}(\mathbf{x},\mathbf{y})<r, m_{\mathbf{y}}\in E]\Big)\lambda_{\mathbb{S}^2}(d\mathbf{x}) \\
&= \frac{1}{\lambda_{\mathbb{S}^2}(W_{\ominus r})} \int_{W_{\ominus r}} P(X^{!}_{\mathbf{x},C}\cap(B_{\mathbb{S}^2}(\mathbf{x},r)\times E)=\emptyset)\lambda_{\mathbb{S}^2}(d\mathbf{x}) \\
&= \frac{1}{\lambda_{\mathbb{S}^2}(W_{\ominus r})} \int_{W_{\ominus r}} 1-P(X^{!}_{\mathbf{x},C}\cap(B_{\mathbb{S}^2}(\mathbf{x},r)\times E)\neq\emptyset)\lambda_{\mathbb{S}^2}(d\mathbf{x}) \\
&= \frac{1}{\lambda_{\mathbb{S}^2}(W_{\ominus r})} \int_{W_{\ominus r}} 1-P(X^{!}_{\mathbf{o},C}\cap(B_{\mathbb{S}^2}(\mathbf{o},r)\times E)\neq\emptyset)\lambda_{\mathbb{S}^2}(d\mathbf{x}) \\
&= 1-P(X^{!}_{\mathbf{o},C}\cap(B_{\mathbb{S}^2}(\mathbf{o},r)\times E)\neq\emptyset) \\
&= 1-D^{CE}(r)
\end{aligned}$$

For $\hat{K}^{CE}$ unbiasedness follows by application of the Campbell Theorem and the fact the independence between X_C and X_E implies $\rho^{(2)}((\mathbf{x},m_{\mathbf{x}}),(\mathbf{y},m_{\mathbf{y}}))=\rho(\mathbf{x},m_{\mathbf{x}})\rho(\mathbf{y},m_{\mathbf{y}})$ when $m_1\in C$ and $m_2\in E$. Ratio-unbiasedness of $\hat{J}^{CE}$ follows by unbiasedness of $\hat{F}^{E}$ and $\hat{D}^{CE}$. □

E.3 Bivariate LGCP on $\mathbb{S}^2$

In this section we discuss bivariate LGCP on $\mathbb{S}^2$ in more detail. Define $c(\mathbf{x},\mathbf{y})=\{c_{ij}(\mathbf{x},\mathbf{y})\}$ to be the covariance function matrix and that the covariance function matrix is isotropic,

i.e. $c_{ij}(\mathbf{x},\mathbf{y}) = c_{ij}(d_{\mathbb{S}^2}(\mathbf{x},\mathbf{y})) = c_{ij}(r)$, where $r > 0$. We assume that the mean functions μ_i of the GRF are continuous and that the covariance functions, c_{ii}, take the following form,

$$c_{ii}(d) = \sigma^2 s_{ii}(r), \tag{E.4}$$

where $s_{ii} : \mathbb{R} \mapsto [-1, 1]$ is a known correlation function. In order for almost sure integrability of the GRF conditions need to be imposed on μ_i and c_{ij}. These conditions are outlined across Cuevas-Pacheco and Møller ((2018)) and Brix and Moller ((2001)). More precisely, Proposition 1 of Cuevas-Pacheco and Møller ((2018)) (see also Proposition 3.6.1) provides conditions on the *variogram* of a zero mean univariate GRF to have continuous sample paths: suppose there exist numbers $s \in (0, 1], l \in (0, 1)$, and $m > 0$ such that,

$$\beta(\mathbf{x},\mathbf{y}) \leq m d_{\mathbb{S}^2}(\mathbf{x},\mathbf{y})^{l/2},$$

for $d_{\mathbb{S}^2}(\mathbf{x},\mathbf{y}) < s$, where $\beta(\mathbf{x},\mathbf{y}) = \mathbb{E}[(X_0(\mathbf{x}) - X_0(\mathbf{y}))^2]/2$ is the variogram and X_0 is a zero mean univariate GRF. Thus in the bivariate setting if β_i is the variogram of X_i the i^{th} component of X then we have the sufficient condition,

$$\beta_i(\mathbf{x},\mathbf{y}) \leq m d_{\mathbb{S}^2}(\mathbf{x},\mathbf{y})^{l/2},$$

for almost sure integrability. Under Equation E.4 we can simplify this condition to,

$$1 - s_{ii}(r) \leq \frac{m}{\sigma^2} d^{l/2},$$

for $d < s$, which is identically the condition derived by Møller and Waagepetersen ((1998)) for Euclidean processes. In addition to this condition we also reqiuire that the covariance function matrix is semi-positive definite and Brix and Moller ((2001)) provide conditions on the covariance and cross covariance functions for this to hold.

In this work we shall suppose that our covariance functions are of the form,

$$c_{ii}(\mathbf{x},\mathbf{y}) = \sigma^2 s_{ii}(||\mathbf{x} - \mathbf{y}||), \tag{E.5}$$

where $\sigma^2, \gamma > 0$, $||\cdot||$ is the Euclidean norm, where $s_{ii} : \mathbb{R} \mapsto [-1, 1]$, or $i = 1, 2$ is known. This form is distinctly different to that seen in Cuevas-Pacheco and Møller ((2018)) for univariate LGCP where they supposed that the correlation function depended on the geodesic distance rather then on the Euclidean distance in our case.

Although we are using a Euclidean distance it is still possible to prove isotropy of the

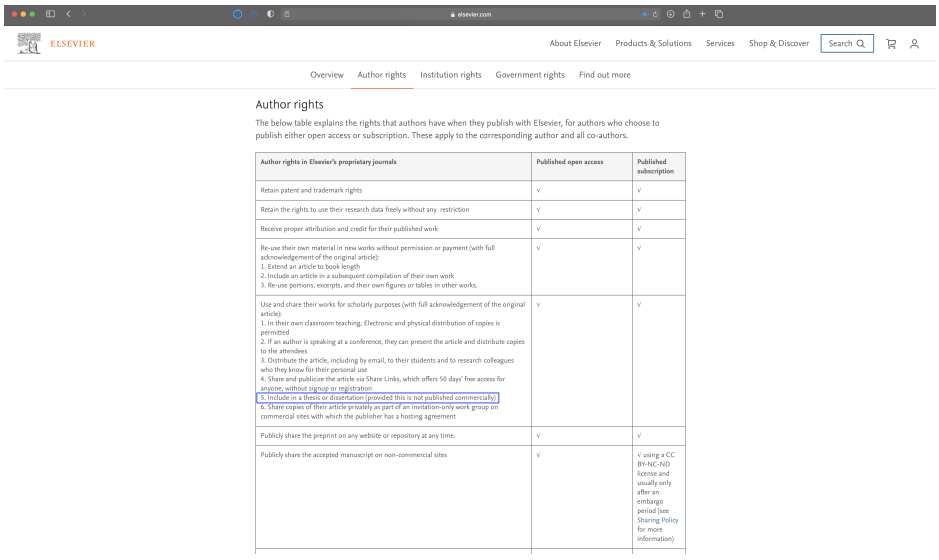
resulting LGCP on the sphere when the mean functions are constant. To see this consider first the bivariate GRF $\tilde{Y}$ in $\mathbb{R}^3$ with constant mean functions and correlation function given by Equation E.5 such that the conditions of Theorem 3.4.1 of Adler ((2010)) are satisfied then the field is isotropic in $\mathbb{R}^3$. If we then define $Y(\mathbf{x}) = \tilde{Y}(\mathbf{x})$ for $\mathbf{x} \in \mathbb{S}^2$ then isotropy of $\tilde{Y}$ implies isotropy of Y and Y is precisely the Gaussian field defined by Equation E.5 with constant mean function. Therefore, since Y is isotropic we also have that X the LGCP with driving field $\exp(Y)$ is also isotropic. An alternative argument is to see that,

$$||\mathbf{x} - \mathbf{y}|| = 2\sin\left(\frac{d_{\mathbb{S}^2}(\mathbf{x}, \mathbf{y})}{2}\right),$$

for $\mathbf{x}, \mathbf{y} \in \mathbb{S}^2$, obtained using simple geometric arguments and hence Equation E.5 can be written as a function of $d_{\mathbb{S}^2}(\mathbf{x}, \mathbf{y})$ and therefore the resulting GRF is isotropic on $\mathbb{S}^2$ and hence so is X. This is similar to a result in Møller and Waagepetersen ((1998)) who show that if X is a LGCP and Y is a GRF in $\mathbb{R}^d$ where if Y is translationally and rotationally invariant then under any rigid motions ϕ in $\mathbb{R}^d$ $(X, Y) \stackrel{d}{=} (\phi(X), Y(\phi(\dot{c})))$. Furthermore, assuming that the conditions of Theorem 3.4.1 of Adler ((2010)) are satisfied then we have almost sure continuity over all of $\mathbb{R}^3$ which implies continuity specifically on $\mathbb{S}^2$ and hence almost sure integrability is satisfied also. The simplicity of using Equation E.5 is that we can simulate GRFs in $\mathbb{R}^3$ and then restrict our view only to those points that lie on the sphere and this subset inherits the properties of the GRF it is embedded in. In order to simulate the GRFs we use the `RandomFields` package available in `R` Schlather et al. ((2020)).

Copyright statement

Chapters 3 - 5 and Appendices A - C reproduce material from our paper 'Testing for complete spatial randomness on three dimensional bounded convex shapes' ((Ward et al., 2021b)). A screenshot of the Authors rights is given below. More details can be found at `https://www.elsevier.com/about/policies/copyright` where a complete list of permissions and limitations can be found.



ELSEVIER

About Elsevier | Products & Solutions | Services | Shop & Discover | Search

Overview | Author rights | Institution rights | Government rights | Find out more

Author rights

The below table explains the rights that authors have when they publish with Elsevier, for authors who choose to publish either open access or subscription. These apply to the corresponding author and all co-authors.

Author rights in Elsevier's proprietary journals	Published open access	Published subscription
Retain patent and trademark rights	√	√
Retain the rights to use their research data freely without any restriction	√	√
Receive proper attribution and credit for their published work	√	√
Re-use their own material in new works without permission or payment (with full acknowledgement of the original article): 1. Extend an article to book length 2. Include an article in a subsequent compilation of their own work 3. Re-use portions, excerpts, and their own figures or tables in other works.	√	√
Use and share their works for scholarly purposes (with full acknowledgement of the original article): 1. In their own classroom teaching. Electronic and physical distribution of copies is permitted 2. If an author is speaking at a conference, they can present the article and distribute copies to the attendees 3. Distribute the article, including by email, to their students and to research colleagues who they know for their personal use 4. Share and publicize the article via Share Links, which offers 50 days' free access for anyone, without signup or registration 5. Include in a thesis or dissertation (provided this is not published commercially) 6. Share copies of their article privately as part of an invitation-only work group on commercial sites with which the publisher has a hosting agreement	√	√
Publicly share the preprint on any website or repository at any time.	√	√
Publicly share the accepted manuscript on non-commercial sites	√	√ using a CC BY-NC-ND license and usually only after an embargo period (see Sharing Policy for more information)

Figure F.1: Copyright agreement associated with Ward et al. ((2021b)). The blue box highlights that the material can be reproduced as part of a thesis or dissertation. Additional details on permissions, resuse and limitations can be found at `https://www.elsevier.com/about/policies/copyright`.