AALBORG UNIVERSITY
DENMARK

Marked point processes with a view to the minicolumn hypothesis of neuroscience

Heidi Søgaard Christensen
Andreas Dyreborg Christoffersen

supervised by
professor Jesper Møller

June 10, 2016

Institute for Mathematical Sciences
Mathematics and Statistics
Fredrik Bajers Vej 7G
9220 Aalborg Ø
Tlf. 99409940

AALBORG UNIVERSITY
STUDENT REPORT

Title:
Marked point processes with a view to the minicolumn hypothesis of neuroscience

Theme:
Spatial statistics

Project Period:
Spring Semester 2016

Project Group:
G2-106c

Participants:
Heidi Søgaard Christensen
Andreas Dyreborg Christoffersen

Supervisor:
Professor Jesper Møller

Copies: 3

Page Numbers: 72

Date of Completion:
June 10, 2016

Abstract:

The organisation of neurons in the human cerebral cortex is a greatly discussed subject that has been linked to various neurological diseases. It is generally accepted by biologists that neurons are organised in minicolumns, but only few statistical analyses have been conducted in attempt to verify this. We investigate this organisation by analysing a marked point pattern consisting of pyramidal cells' three-dimensional nucleolus locations and orientation. The thesis has a hierarchical structure beginning with an introduction to the minicolumn hypothesis and the data. This is followed by an analysis of the ground point pattern in which we find, that the Strauss hardcore point process suitably models the nucleolus locations. To model the marked point pattern in relation to the hypothesised minicolumn structure we introduce a new summary statistic which accounts for the points' orientation. This and other summary statistics are used to test for independent marking, which does not seem to be a suitable model. We finish the thesis with a discussion.

The content of this report is freely available, but publication (with reference) may only be pursued due to agreement with the authors.

Contents

Preface

This master's thesis is an extension of the project Christensen and Christoffersen (2015), written by this project group during their 8th semester. Christensen and Christoffersen (2015) deal with analysing the unmarked points of a marked point pattern, which describes the nucleolus locations (points) and the orientations (marks) of pyramidal cells in the human cerebral cortex. This thesis extends the analysis of the unmarked point pattern in chapter 2 and analyses the marked point pattern in chapter 3. The treatment of unmarked point processes in chapter 2 provides a stronger theoretical background than Christensen and Christoffersen (2015) and focuses on theory that improves the analysis. Specifically, this involves Palm distributions (section 2.2), Strauss hardcore point processes (section 2.3.2), the cylindrical K-function (section 2.4.2) and global envelopes (section 2.4.3). In chapter 3 we initially introduce the most basic definitions, models and properties for marked point processes. We then generalise Palm distributions to the case of marked point processes, which in turn works as a solid theoretical foundation to introduce different functional summary statistics. To investigate the minicolumn hypothesis we introduce the modified (mark-weighted) cylindrical K-function, which is particularly useful for detecting columnar structures in the case of directional marks.

We like to thank our supervisor professor Jesper Møller for his guidance and exceptional ability to always understand and push our limits. Furthermore, we are grateful to professor Jens R. Nyengaard and CSGB for supplying the data.

Aalborg University, June 10, 2016

Heidi Søgaard Christensen
hsch11@student.aau.dk

Andreas Dyreborg Christoffersen
adch10@student.aau.dk

Danish abstract

Dette speciale er skrevet i forlængelse af projektet Christensen and Christoffersen (2015), som blev udarbejdet i forbindelse med vores 8. semester. Christensen and Christoffersen (2015) analyserer de umærkede punkter fra et mærket punktmønster, som beskriver nukleoleplaceringen samt orienteringen af pyramideceller i et menneskes hjernebark. I begge projekter undersøges organiseringen af neuroner i menneskets hjerne, som biologer generelt har accepteret som værende arrangeret i blandt andet minisøjler med en diameter på 35-60μm bestående af 80-100 neuroner. Der er dog kun lavet få statistiske analyser vedrørende denne organisering. I dette speciale udvider vi analysen for det umærkede punktmønster præsenteret i Christensen and Christoffersen (2015) og analyserer yderligere det mærkede punktmønster. Under uafhængighed mellem punkter og mærker er der en naturlig hierarkisk struktur i analysen, hvor først punkterne og dernæst mærkerne modelleres. I kapitel 2 præsenterer vi kort de mest nødvendige definitioner og egenskaber fra Christensen and Christoffersen (2015). Herudover introducerer vi Palm-fordelinger, Strauss hardcore punktprocesser, den cylindriske K-funktion samt globale envelopes, som er inkluderet for at give en stærkere teoretisk baggrund samt forbedre analysen fra Christensen and Christoffersen (2015). Vi finder at Strauss hardcore punktprocessen er en passende model til at modellere nukleoleplaceringerne for pyramidecellerne. I kapitel 3 præsenterer vi basale definitioner, modeller og egenskaber for mærkede punktprocesser. Særligt interessant er independent marking, som er den mest simple model. Herefter generaliserer vi teorien om Palm-fordelinger til tilfældet med mærkede punktprocesser, hvilket giver en stærk baggrund for at præsentere mærke-korrelationsfunktionen og det mærke-vægtede $\mathcal{K}$-mål, som er summary statistics for mærkede punktprocesser. Til at undersøge minisøjlehypotesen har vi yderligere præsenteret den modificerede cylindriske K-funktion, som er en modificering af det mærke-vægtede $\mathcal{K}$-mål, hvor struktur elementet er en cylinder hvis retning varierer over mærkerne. I dataanalysen for det mærkede punktmønster tilpasser vi en independent marking model, hvor punkterne modelleres ved Strauss hardcore modellen fra afsnit 2.5.2 og tætheden for mærkerne er bestemt ud fra et kerneestimat. Vi finder at independent marking ikke er en passende model. Vi afslutter med en diskussion om hvilke fremgangsmåder, der vil være relevante for fremtidige analyser.

Chapter 1

Introduction

We begin this master's thesis with an introduction to the problem, data and previous achievements. This will set the framework and give the reader an understanding of the goal and leitmotif of the thesis.

The organisation of neurons in the mammalian brain is a widely discussed subject in neurology and is suspected to be associated with multiple psychological and neurological diseases, such as schizophrenia, Alzheimer's disease, autism and Down's syndrome (Rafati et al., 2016, and references therein). The minicolumn hypothesis states that the neurons in the human cerebral cortex are radially organised in columns perpendicular to the pial surface of the brain. This hypothesis is generally accepted among neurologists, and each column is said to have a diameter of 35-60μm and contain 80-100 neurons. However, only few statistical analyses have been conducted and most of which are based on two-dimensional unmarked point patterns. Rafati et al. (2016) analyse a three-dimensional unmarked point pattern describing pyramidal cells' nucleolus locations and find the existence of minicolumns plausible without a priori assuming so. They analyse the organisation by fitting a Poisson line cluster point process and applying the cylindrical K-function, both presented by Møller et al. (2016).

We shall analyse the neural organisation by considering a marked point pattern supplied by CSGB. The provided data contains three-dimensional coordinates of nucleoli and apexes for pyramidal cells found in a sample from the third layer of Brodmann area four in the cerebral cortex of a human brain. We initially treated the marks as the three-dimensional vectors pointing from the nucleoli to their corresponding apex. However, regular dialogue with professor Jens R. Nyengaard, who is a member of CSGB, revealed that the length between nucleolus and apex had no biological meaning. Hence, the marks should be treated as directional vectors (i.e. unit vectors) indicating the cells' orientation. Unfortunately, this misunderstanding was first discovered in May after multiple analyses had already been conducted. Therefore, parts of the analysis have been performed under the false presumption that the observed

apex location in itself has a biological meaning. In chapter 2 we analyse both the nucleolus and apex locations which is not sensible after the aforementioned discovery.

In Christensen and Christoffersen (2015) only the unmarked points were regarded. We found that this was a regular point pattern and, by extension, three models were proposed: The Matérn hardcore point processes of type I and II and the Gibbs hardcore point process (described as the Strauss process with the interaction parameter fixed as 0). All the models consisted of two parameters; one being the hardcore distance and the other describing the intensity. It was shown that both of the parameters in the Gibbs hardcore process had a maximum profile pseudolikelihood estimate that could be expressed in closed form. However, the expression of the intensity parameter could not be easily evaluated and thus grid approximations was used. Since the Matérn processes are not described by a density, different methods for minimising the contrast function based on the summary function, K, were suggested for parameter estimation. To compare the point pattern to the suggested models, non-parametric estimates of several functional summary statistics along with pointwise envelopes were used. Christensen and Christoffersen (2015) found that neither of the three models suitably described the data.

Chapter 2

Point processes

In this chapter we introduce basic concepts of unmarked spatial point processes and use the theory to analyse the nucleolus and apex locations. Some of the theory is already introduced in Christensen and Christoffersen (2015), but is included in this thesis in order to make it more self-contained. In section 2.1 basic definitions and properties of spatial point processes are presented, and in section 2.2 the reduced Palm distribution is described. The Poisson, Strauss and Strauss hardcore point processes, which are models for analysing spatial point patterns, are introduced in section 2.3. Some of the most popular summary statistics used for preliminary analysis and model validation are described in section 2.4 along with the cylindrical K-function and the concept of global envelopes. Out of all the theory presented in this chapter the theory on Palm distributions, Strauss hardcore point processes, the cylindrical K-function and global envelopes have not been treated in Christensen and Christoffersen (2015). We end this chapter in section 2.5 with an analysis of the nucleolus and apex locations.

2.1 Basic definitions and properties

A spatial point process is a mathematical model used to describe a random point pattern. Loosely speaking, a point process is a random countable subset of some space. In order to give a more precise definition of a (locally finite) point process we first need to introduce some notation. Let S be a complete separable metric space equipped with a Borel σ-algebra $\mathcal{B}$. Define $n(x)$ to be the cardinality of $x \subseteq S$, letting $n(x) = \infty$ if x is not finite. Let x_B denote the restriction of a point configuration x to some set $B \in \mathcal{B}$. Then x is said to be *locally finite* if $n(x_B) < \infty$ for all bounded sets $B \in \mathcal{B}$. We denote the set of all locally finite point configurations on S by $N_{p,\text{lf}}$, i.e.

$$N_{p,\text{lf}} = \{x \subseteq S : n(x_B) < \infty \text{ for all bounded } B \in \mathcal{B}\}$$

and equip it with the σ-algebra

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

for $\mathbb{N}_0 = \{0\} \cup \mathbb{N}$, where $\mathbb{N} = \{1, 2, \dots\}$. A strict definition of a point process can then be formulated as follows.

Definition 2.1.1. A *locally finite point process* X defined on S is a measurable mapping on a probability space $(\Omega, \mathcal{F}, P)$ taking values in $(N_{p,\mathrm{lf}}, \mathcal{N}_{p,\mathrm{lf}})$. The distribution P_X of X is given by $P_X(F) = P(\{\omega \in \Omega : X(\omega) \in F\}) = P(X \in F)$ for $F \in \mathcal{N}_{p,\mathrm{lf}}$.

Even though a point process can be considered on any complete separable metric space S, we will restrict ourselves to the case where $S \subseteq \mathbb{R}^d$ (i.e. the d-dimensional Euclidean space) for $d \geq 1$, which covers most applications. Specifically, for the pyramidal cell data $d = 3$. Furthermore, only *simple point processes* will be considered, i.e. point processes that with probability one has no coinciding points. For treatment of point processes in a more general set-up we refer to Daley and Vere-Jones (2002).

For ease of readability most measure theoretical wording or notation will be omitted hereinafter. Particularly, it will not be explicitly written whenever a subset is an element of a certain σ-algebra, e.g. $B \subseteq S$ will actually mean $B \in \mathcal{B}$. Similarly, whenever talking about a function, it is implicitly assumed that measurability wrt. the appropriate σ-algebra is fulfilled. We shall furthermore abuse notation and write $x \backslash \xi$ instead of $x \backslash \{\xi\}$, $x \cup \xi$ instead of $x \cup \{\xi\}$, etc.

For a point process X on S define

$$N(B) = n(X_B)$$

for Borel sets $B \subseteq S$, where X_B denotes the restriction of X to B. That is $N(B)$ is the random number of points from X falling in B. The *void probability* is defined by $\nu_p(B) = P(N(B) = 0)$ for bounded Borel sets $B \subseteq S$. It can be shown (under mild conditions) that a simple point process is uniquely determined by its void probabilities (see page 35 in Daley and Vere-Jones, 2008).

Definition 2.1.2. Define the *intensity measure* by $\mu_p(B) = \mathrm{E}[N(B)]$ for $B \subseteq S$ Borel. If there exists a function ρ such that

$$\mu_p(B) = \int_B \rho(\xi) d\xi$$

for all Borel sets $B \subseteq S$, i.e. ρ is the density of μ_p wrt. Lebesgue measure, then ρ is called the *intensity function*.

It will be assumed that μ_p is locally finite, that is $\mu_p(B) < \infty$ for all bounded Borel sets $B \subseteq S$. If the intensity function exists this implies that it is locally integrable, i.e. $\int_B \rho(\xi)d\xi < \infty$ for all bounded Borel sets $B \subseteq S$. For $\xi \in S$, $\rho(\xi)d\xi$ can be interpreted as the probability of observing a point from X in the infinitesimal ball with centre ξ and volume $d\xi$.

For $S = \mathbb{R}^d$, the point process X is said to be *stationary* (resp. *isotropic*) if the distribution of X is invariant under translation (resp. rotation about the origin). If X is a stationary point process with locally finite intensity measure μ_p, then μ_p is translation invariant and thus proportional to the Lebesgue measure. Consequently (by definition 2.1.2) the intensity function ρ must be constant.

Consider two point processes X and Y on S, where X is absolutely continuous wrt. Y, i.e. $P(X \in F) = 0$ whenever $P(Y \in F) = 0$ for all $F \subseteq N_{p,\mathrm{lf}}$. Then the *density of X wrt. Y* is defined as the Radon-Nikodym derivative f that satisfies

$$P(X \in F) = \mathrm{E}[\mathbb{I}(Y \in F)f(Y)]$$

for all $F \subseteq N_{p,\mathrm{lf}}$, where $\mathbb{I}(\cdot)$ denotes the indicator function.

2.2 Palm theory

One of the most fundamental building blocks of spatial statistics is the so-called reduced Palm distribution. To define this, we first introduce the reduced Campbell measure.

Definition 2.2.1. Let X be a point process on S. Then the *reduced Campbell measure* is defined as

$$C_p^!(D) = \mathrm{E}\left[\sum_{\xi \in X} \mathbb{I}\left((\xi, X \backslash \xi) \in D\right)\right] \tag{2.1}$$

for $D \subseteq S \times N_{p,\mathrm{lf}}$.

Clearly the Campbell measure determines the intensity measure since $\mu_p(\cdot) = C_p^!(\cdot \times N_{p,\mathrm{lf}})$. In general, $C_p^!(\cdot \times F) \leq \mu_p(\cdot)$ for all $F \subseteq N_{p,\mathrm{lf}}$. Hence, $C_p^!(\cdot \times F)$ is absolutely continuous wrt. $\mu_p(\cdot)$ for any fixed $F \subseteq N_{p,\mathrm{lf}}$ (also denoted $C_p^!(\cdot \times F) \ll \mu_p(\cdot)$), i.e. $C_p^!(B \times F) = 0$ whenever $\mu_p(B) = 0$ for any Borel set $B \subseteq S$. Therefore, there exists a family of Markov kernels $P_\xi^!(F)$ such that

$$C_p^!(B \times F) = \int_B P_\xi^!(F) d\mu_p(\xi) \tag{2.2}$$

for all Borel sets $B \subseteq S$. That is, $P_\xi^!(F)$ is a measurable mapping wrt. ξ for fixed F and a probability measure on F for fixed ξ. The measurability follows

from the fact that $P^!_\xi(F)$ is the Radon-Nikodym derivative of $C^!_p(\cdot \times F)$ wrt. $\mu_p(\cdot)$. Note that this also entails that $P^!_\xi(F)$ is uniquely determined μ_p-almost every ξ (in short μ_p-a.e. ξ). If the intensity function ρ exists we can equivalently say that $P^!_\xi(F)$ is unique Lebesgue-a.e. ξ for $\rho(\xi) > 0$. Proving that $P^!_\xi(F)$ is a probability measure on F, for all fixed $\xi \in S$, is done in two steps: (1) $P^!_\xi(F)$ is a probability measure on F, for μ_p-a.e. $\xi \in S$, since

$$\int_B P^!_\xi(\emptyset) d\mu_p(\xi) = \mathrm{E}\left[\sum_{\xi \in X} \mathbb{I}\left(\xi \in B, X \backslash \xi \in \emptyset\right)\right] = 0 = \int_B 0 d\mu_p(\xi),$$

and for disjoint sets $F_i \subseteq N_{p,\mathrm{lf}}$, $i = 1, 2, \ldots$

$$\begin{aligned}\int_B P^!_\xi\left(\overset{\infty}{\underset{i=1}{\cup}} F_i\right) d\mu_p(\xi) &= \mathrm{E}\left[\sum_{\xi \in X} \mathbb{I}\left(\xi \in B, X \backslash \xi \in \overset{\infty}{\underset{i=1}{\cup}} F_i\right)\right] \\ &= \sum_{i=1}^{\infty} \mathrm{E}\left[\sum_{\xi \in X} \mathbb{I}\left(\xi \in B, X \backslash \xi \in F_i\right)\right] = \int_B \sum_{i=1}^{\infty} P^!_\xi\left(F_i\right) d\mu_p(\xi)\end{aligned}$$

and lastly

$$\int_B P^!_\xi(N_{p,\mathrm{lf}}) d\mu_p(\xi) = C^!_p(B \times N_{p,\mathrm{lf}}) = \mu_p(B) = \int_B 1 d\mu_p(\xi)$$

for all Borel sets $B \subseteq S$. (2) for every ξ in a μ_p-null set, we can choose $P^!_\xi(F)$ to be any probability measure on F. Thus, $P^!_\xi(F)$ can indeed be chosen as a Markov kernel.

Definition 2.2.2. The probability measure $P^!_\xi(\cdot)$ is called the *reduced Palm distribution at* ξ.

The (non-reduced) Palm distribution is defined in a similar manner, but based on the (non-reduced) Campbell measure obtained by substituting $X \backslash \xi$ with X in (2.1). Results for the reduced Palm distribution can easily be modified to the (non-reduced) Palm distribution if desired. However, as results and interpretations of the two distribution only differ slightly we shall just consider the reduced version in the following.

Consider a ball $B = b(\xi, \varepsilon)$ centered at $\xi \in S$ and with a sufficiently small radius $\varepsilon > 0$ such that the event $N(B) > 1$ is unlikely. Then, for all $F \subseteq N_{p,\mathrm{lf}}$, $\mu_p(B) \approx P(N(B) > 0)$ and $C^!_p(B \times F) \approx P\left(N(B) > 0, X \backslash \xi \in F\right)$ and thus

$$P^!_\xi(F) \approx \frac{C^!_p(B \times F)}{\mu_p(B)} \approx P\left(X \backslash \xi \in F \mid N(B) > 0\right).$$

Hence, the reduced Palm distribution may be interpreted as the conditional distribution of $X \backslash \xi$ given that $\xi \in X$.

It follows from the definition of the reduced Campbell measure and the reduced Palm distribution as well as the standard proof that

$$\mathrm{E}\left[\sum_{\xi\in X} h\left(\xi, X\backslash\xi\right)\right] = \int\int h\left(\xi, x\right) dP_\xi^!(x) d\mu_p(\xi) \tag{2.3}$$

for any function $h : S \times N_{p,\mathrm{lf}} \to [0,\infty)$. This formula is referred to as the *Campbell-Mecke theorem*.

For a stationary point process X on $\mathbb{R}^d$ with intensity $0 < \rho < \infty$, the reduced Palm distribution at origo, $P_0^!$, is of special interest. We will also refer to $P_0^!$ as the reduced Palm distribution at a typical point. For $\xi \in S$, $X \sim P_0^!$ if and only if $X + \xi \sim P_\xi^!$, where $X + \xi$ denotes the point process translated by ξ. Using this notation, the Campbell-Mecke theorem can be expressed as

$$\mathrm{E}\left[\sum_{\xi\in X} h\left(\xi, X\backslash\xi\right)\right] = \rho\int\int h\left(\xi, x\right) dP_\xi^!(x) d\xi = \rho\int\int h\left(\xi, x+\xi\right) dP_0^!(x) d\xi \tag{2.4}$$

for any function $h : S\times N_{p,\mathrm{lf}} \to [0,\infty)$. Let $h\left(\xi, X\backslash\xi\right) = \mathbb{I}\left(\xi \in B, X\backslash\xi - \xi \in F\right)$ for any Borel set $B \subseteq \mathbb{R}^d$ and $F \subseteq N_{p,\mathrm{lf}}$, then

$$\mathrm{E}\left[\sum_{\xi\in X_B} \mathbb{I}\left(X\backslash\xi - \xi \in F\right)\right] = \rho\int_B\int \mathbb{I}\left(x \in F\right) dP_0^!(x) d\xi = \rho|B|P_0^!(F),$$

where $|B|$ denotes the Lebesgue measure of B. Note that the left hand side is a translation invariant measure and is consequently proportional to the Lebesgue measure. Hence,

$$P_0^!(F) = \frac{1}{\rho|B|}\mathrm{E}\left[\sum_{\xi\in X_B} \mathbb{I}\left(X\backslash\xi - \xi \in F\right)\right]$$

does not depend on the choice of B.

2.3 Models

The Poisson point process is one of the most fundamental models for unmarked point processes. Due to its simple and independent nature it is often used as a null model and to define point processes specified by a density function. The Strauss and Strauss hardcore point processes are examples of point processes specified by a density wrt. the Poisson point process.

2.3.1 Poisson point process

For a Poisson point process there is no interaction between points and therefore *complete spatial randomness* (CSR) is often used as a synonym. Let $\mathrm{Poi}\left(\lambda\right)$

denote the Poisson distribution with mean $\lambda > 0$, then the Poisson point process is defined as follows.

Definition 2.3.1. Consider a point process X on S with intensity function ρ. Then X is said to be a *Poisson point process on S with intensity ρ*, if

1. $N(B) \sim \text{Poi}(\mu_p(B))$ for any Borel set $B \subseteq S$ satisfying $\mu_p(B) < \infty$.
2. $N(B_1), \ldots, N(B_n)$ are independent for any pairwise disjoint Borel sets $B_1, \ldots, B_n \subseteq S$, $n \geq 2$, with $\mu_p(B_i) < \infty$, $i = 1, \ldots, n$.

This is denoted by $X \sim \text{Poisson}(S, \rho)$.

Alternatively, a Poisson point process on S can be defined by void probabilities as a point process X fulfilling

$$\nu_p(B) = \exp(-\mu_p(B)) \tag{2.5}$$

for all bounded Borel sets $B \subseteq S$. Yet another characterization of the Poisson point process is the so-called *Poisson process expansion* presented in the following theorem.

Theorem 2.3.2. Consider a point process X with intensity measure $\mu_p(B) = \int_B \rho(\xi) d\xi$ for Borel sets $B \subseteq S$. Then $X \sim \text{Poisson}(S, \rho)$ if and only if

$$P(X_B \in F) = \sum_{n=0}^{\infty} \frac{\exp(-\mu_p(B))}{n!} \int_B \cdots \int_B \mathbb{I}(\{\xi_1, \ldots, \xi_n\} \in F) \prod_{i=1}^{n} \rho(\xi_i) d\xi_1 \cdots d\xi_n \tag{2.6}$$

for all Borel sets $B \subseteq S$ with $\mu_p(B) < \infty$ and all $F \subseteq N_{p,\text{lf}}$.

Proof for the above theorem can e.g. be found in Christensen and Christoffersen (2015). For $X \sim \text{Poisson}(S, \rho)$, the Poisson expansion can be extended by applying the standard proof to (2.6), yielding

$$\text{E}[h(X_B)] = \sum_{n=0}^{\infty} \frac{\exp(-\mu_p(B))}{n!} \int_B \cdots \int_B h(\{\xi_1, \ldots, \xi_n\}) \prod_{i=1}^{n} \rho(\xi_i) d\xi_1 \cdots d\xi_n \tag{2.7}$$

for any function $h : N_{p,\text{lf}} \to [0, \infty)$ and any Borel set $B \subseteq S$ with $\mu_p(B) < \infty$.

Theorem 2.3.3. If $X \sim \text{Poisson}(S, \rho)$, then for any function $h : S^n \times N_{p,\text{lf}} \to [0, \infty)$ with $n \in \mathbb{N}$,

$$\text{E}\left[\sum_{\xi_1, \ldots, \xi_n \in X}^{\neq} h(\xi_1, \ldots, \xi_n, X \setminus \{\xi_1, \ldots, \xi_n\}) \right]$$

$$= \int_S \cdots \int_S \text{E}[h(\xi_1, \ldots, \xi_n, X)] \prod_{i=1}^{n} \rho(\xi_i) d\xi_1 \ldots d\xi_n,$$

where $\sum^{\neq}$ is the sum over pairwise distinct points.

For proof see e.g. Møller and Waagepetersen (2004). The above theorem is generally referred to as the *extended Slivnyak-Mecke theorem*, while the case where $n = 1$ simply is called the *Slivnyak-Mecke theorem*. One use of the Slivnyak-Mecke theorem is to formally prove the following intuitive result regarding the reduced Palm distribution for a Poisson point process.

Theorem 2.3.4. For a point process $X \sim \text{Poisson}(S, \rho)$, the reduced Palm distribution at ξ can be chosen as $P_\xi^! = \text{Poisson}(S, \rho)$ for all $\xi \in S$.

Proof. The Slivnyak-Mecke and Campbell-Mecke theorems found in theorem 2.3.3 and equation (2.3), respectively, give

$$\int \mathrm{E}\left[h(\xi, X)\right] d\mu_p(\xi) = \int \int h(\xi, x) dP_\xi^!(x) d\mu_p(\xi) = \int \mathrm{E}\left[h(\xi, X_\xi^!)\right] d\mu_p(\xi),$$

where $h : S \times N_{p,\text{lf}} \to [0, \infty)$ and $X_\xi^! \sim P_\xi^!$ for $\xi \in S$. Therefore, $\mathrm{E}\left[h(\xi, X)\right] = \mathrm{E}\left[h(\xi, X_\xi^!)\right]$ μ_p-a.e. ξ and thus $P_\xi^! = \text{Poisson}(S, \rho)$ μ_p-a.e. ξ. For any ξ in a μ_p-null set, we can choose $P_\xi^! = \text{Poisson}(S, \rho)$, finishing the proof. □

2.3.2 Strauss and Strauss hardcore point processes

In Christensen and Christoffersen (2015) general theory of finite Markov point processes and particularly pairwise interaction processes were treated. The Strauss and Strauss hardcore point processes are examples of pairwise interaction processes. In the following we outline the most central theory necessary for estimation and simulation under these two models.

Throughout this section we consider a point process X on $S \subset \mathbb{R}^d$ with $|S| < \infty$, where X has a density f wrt. Poisson$(S, 1)$. That is,

$$P(X \in F) = \mathrm{E}\left[\mathbb{I}(Y \in F) f(Y)\right]$$

for $Y \sim \text{Poisson}(S, 1)$ and $F \subseteq N_{p,\text{f}}$, where $N_{p,\text{f}} = \{x \subseteq S : n(x) < \infty\}$. Often the density is only specified up to proportionality, i.e. $f = \frac{1}{c}h$ for a known function $h : N_{p,\text{f}} \to [0, \infty)$ and a normalising constant c given by

$$c = \mathrm{E}\left[h(Y)\right] = \sum_{n=0}^{\infty} \frac{\exp(-|S|)}{n!} \int_S \dots \int_S h(\{\xi_1, \dots, \xi_n\}) d\xi_1 \dots \xi_n.$$

The above expression for c is found by considering equation (2.7) for $1 = P(X \in N_{p,\text{f}}) = \frac{1}{c}\mathrm{E}\left[h(Y)\right]$. Typically, c is unknown except in the special case of a Poisson process (see e.g. proposition 3.8 in Møller and Waagepetersen, 2004).

Definition 2.3.5. Consider a function $h : N_{p,\text{f}} \to [0, \infty)$. Then h is said to be *hereditary* if

$$h(x) > 0 \Rightarrow h(y) > 0$$

for $y \subset x$. If there exists a function $\phi^* : S \to [0, \infty)$ with $\int_S \phi^*(\xi)d\xi < \infty$ and a constant $0 < \alpha < \infty$ such that

$$h(x) \leq \alpha \prod_{\xi \in x} \phi^*(\xi)$$

for all $x \in N_{p,\mathrm{f}}$, then h is said to be *Ruelle stable*.

The *Strauss point process* is defined by the density

$$f(x) = \frac{1}{c(\beta, \gamma, R)} \beta^{n(x)} \gamma^{s_R(x)}$$

for $\beta > 0$, $R > 0$, $0 \leq \gamma \leq 1$ and a normalising constant $c(\beta, \gamma, R)$. Here $s_R(x)$ denotes the number of R-close neighbours in x, i.e. $s_R(x) = \sum_{\{\xi,\eta\} \subseteq x} \mathbb{I}(\|\eta - \xi\| \leq R)$. Furthermore, we let $0^0 = 1$. Note that the density is hereditary. For $\gamma = 0$ the Strauss process is a hardcore point process with hardcore R, while $\gamma = 1$ results in a homogeneous Poisson process with intensity β as shown in Christensen and Christoffersen (2015). In Christensen and Christoffersen (2015) we furthermore show that for $\gamma > 1$ the density is not integrable, and that the restriction $\gamma \leq 1$ implies Ruelle stability and thus integrability of the density (that Ruelle stability implies integrability is also proved in Christensen and Christoffersen, 2015). Introducing a hardcore condition to the Strauss density removes the need of requiring $\gamma \leq 1$ in order to get an integrable density. That is, we define the *Strauss hardcore point process* by the density

$$f_{(\beta,\gamma,R,h)}(x) = \frac{1}{c(\beta, \gamma, R, h)} \beta^{n(x)} \gamma^{s_R(x)} \mathbb{I}(\|\eta - \xi\| > h, \forall \eta, \xi \in x) \tag{2.8}$$

for $\beta, \gamma, h > 0$, $R > h$ and a normalising constant $c(\beta, \gamma, R, h)$. Again, the density is easily seen to be hereditary. The following argument for Ruelle stability of this density is inspired by Baddeley et al. (2013). Let $h(x) = h_1(x)h_2(x)$ denote the unnormalised density for the Strauss hardcore point process, where $h_1(x) = \beta^{n(x)}\gamma^{s_R(x)}$ and $h_2(x) = \mathbb{I}(\|\eta - \xi\| > h, \forall \eta, \xi \in x)$ for $\beta, \gamma, h > 0$, $R > h$ and $x \in N_{p,\mathrm{f}}$. The assumption $|S| < \infty$ ensures that there is a finite number M of non-overlapping spheres with centres inside S and diameter h which can fit into S. Consequently, any realisation of the process has at most M points, i.e. $h_2(x) > 0$ implies that $n(x) \leq M$. Furthermore, the number of R-close neighbours has a finite upper bound:

$$0 \leq s_R(x) \leq n(x)(n(x) - 1).$$

Let $K(n) = \max(1, \beta)^n \max(1, \gamma)^{n(n-1)}$. Then $h_1(x) \leq K(n(x))$, and since $h(x) = 0$, whenever $n(x) > M$, this implies that $h(x) \leq K(M) < \infty$. Then per definition 2.3.5 with $\alpha = K(M)$ and $\phi^* = 1$, the density is Ruelle stable and hence integrable wrt. Poisson$(S, 1)$.

Estimation

Let x denote a point pattern on an observation window $W \subseteq S$. According to the following theorem, the MLE of the hardcore in the Strauss hardcore model can be found on closed form and has a very simple expression.

Theorem 2.3.6. The MLE of h in the Strauss hardcore model is the minimum distance between all pair of points in x.

Proof. Clearly, we need $h \leq \min_{\{\xi,\eta\}\subseteq x}(\|\xi-\eta\|)$ in order to get a non-zero density. The inverse normalising constant

$$c(\beta,\gamma,R,h)^{-1} = \left(\sum_{n=0}^{\infty}\frac{\exp(-|S|)}{n!}\int_S\dots\int_S \beta^{n(\{\xi_1,\dots,\xi_n\})}\gamma^{S_R(\{\xi_1,\dots,\xi_n\})} \cdot \mathbb{I}(\|\eta-\xi\|> h, \forall\eta,\xi\in\{\xi_1,\dots,\xi_n\})d\xi_1\dots d\xi_n\right)^{-1}$$

increases as h decreases. Hence the MLE of h is $\min_{\{\xi,\eta\}\subseteq x}(\|\xi-\eta\|)$. □

In order to estimate the remaining parameters in the Strauss hardcore model we numerically optimise an approximation of the profile pseudolikelihood. Consider a general point process on S specified by a hereditary and Ruelle stable density in the parametric family $\{f_\theta : \theta\in\Theta\}$, where each f_θ is a density wrt. Poisson$(S,1)$, $\Theta\subseteq\mathbb{R}^p$ and $p\geq 1$. Then the pseudolikelihood function can be expressed as

$$PL_A(\theta;x) = \exp\left(-\int_A \lambda_\theta^*(x,\xi)\,d\xi\right)\prod_{\xi\in x}\lambda_\theta^*(x\backslash\xi,\xi) \tag{2.9}$$

for any $A\subseteq S$ due to Jensen and Møller (1991). Here λ_θ^* denotes the *Papangelou conditional intensity* defined by

$$\lambda_\theta^*(x,\xi) = \frac{f_\theta(x\cup\xi)}{f_\theta(x)}$$

for $x\in N_{p,\mathrm{f}}$ and $\xi\in S\backslash x$. If $S=W$ we usually choose $A=W$ in (2.9). On the other hand, if $W\subset S$ different methods for handling edge effects are available (see e.g. Baddeley and Turner, 2000). We will in the following ignore edge effects and consider $A=W$. Note that the pseudolikelihood function does not depend on the normalising constant, which makes optimisation easier. For the Strauss hardcore model, we consider the profile pseudolikelihood with hardcore $\hat{h}=\min_{\{\xi,\eta\}\subseteq x}(\|\xi-\eta\|)$ and $\theta=(\beta,\gamma,R)$. Then the Papangelou conditional intensities in (2.9) are

$$\lambda^*_{(\beta,\gamma,R)}(x,\xi) = \beta\gamma^{n(x\cap b(\xi,R))}\mathbb{I}(||\xi-\eta||>\hat{h},\forall\eta\in x)$$

$$\lambda^*_{(\beta,\gamma,R)}(x\backslash\xi,\xi) = \beta\gamma^{n((x\backslash\xi)\cap b(\xi,R))}.$$

These give the profile pseudolikelihood

$$PL_W(\beta,\gamma,R;x) = \exp\left(-\beta\int_W \gamma^{n(x\cap b(\xi,R))}\mathbb{I}(||\xi-\eta||>\hat{h},\forall\eta\in x)d\xi\right)\prod_{\xi\in x}\beta\gamma^{n((x\backslash\xi)\cap b(\xi,R))}.$$

The integral on the right hand side can be estimated in several ways. One approach is to write the integral as an expected value under a uniform distribution and then estimate this expected value using simulations. That is

$$\begin{aligned}
&\int_W \gamma^{n(x\cap b(\xi,R))}\mathbb{I}(||\xi-\eta||>\hat{h},\forall\eta\in x)d\xi\\
&\quad= |W|\int_W \frac{1}{|W|}\gamma^{n(x\cap b(\xi,R))}\mathbb{I}(||\xi-\eta||>\hat{h},\forall\eta\in x)d\xi\\
&\quad= |W|\mathrm{E}\left[\gamma^{n(x\cap b(Y,R))}\mathbb{I}(||Y-\eta||>\hat{h},\forall\eta\in x)\right]\\
&\quad\approx \frac{|W|}{N}\sum_{i=1}^{N}\gamma^{n(x\cap b(y_i,R))}\mathbb{I}(||y_i-\eta||>\hat{h},\forall\eta\in x),
\end{aligned}$$

where Y is a stochastic variable uniformly distributed on W and y_i is a realisation of Y for $i=1,\ldots,N$ for large $N\in\mathbb{N}$. All in all this gives an approximation of the profile pseudolikelihood,

$$PL_W(\beta,\gamma,R;x) \approx \beta^{n(x)}\exp\left(-\beta\frac{|W|}{N}\sum_{i=1}^{N}\gamma^{n(x\cap b(y_i,R))}\mathbb{I}(||y_i-\eta||>\hat{h},\eta\in x)\right)\prod_{\xi\in x}\gamma^{n((x\backslash\xi)\cap b(\xi,R))}.$$

In our application the exponent is strongly negative and we rather maximize the log-profile pseudolikelihood,

$$\begin{aligned}
\log PL_W(\beta,\gamma,R;x) \approx &-\beta\frac{|W|}{N}\sum_{i=1}^{N}\gamma^{n(x\cap b(y_i,R))}\mathbb{I}\left(||y_i-\eta||>\hat{h},\eta\in x\right)\\
&+n(x)\log(\beta)+\log(\gamma)\sum_{\xi\in x}n((x\backslash\xi)\cap b(\xi,R))
\end{aligned}$$

for $\beta,\gamma>0$ and $R>\hat{h}$. For the Strauss model a similar approximation of the pseudolikelihood can be obtained:

$$PL_W(\beta,\gamma,R;x) \approx \beta^{n(x)}\exp\left(-\beta\frac{|W|}{N}\sum_{i=1}^{N}\gamma^{n(x\cap b(y_i,R))}\right)\prod_{\xi\in x}\gamma^{n((x\backslash\xi)\cap b(\xi,R))}$$

for $\beta,R>0$ and $0\leq\gamma\leq 1$.

Simulation

The birth-death Metropolis-Hastings algorithm, which is an example of a Markov chain Monte Carlo algorithm, will be used for simulating under the Strauss and Strauss hardcore model. It is an iterative algorithm that for each step attempts to remove or add a point to the current point pattern. Here the algorithm is presented in a practical manner and we refer to Baddeley et al. (2000) and the references therein for details of the underlying theory.

Consider a point process specified by an unnormalised density h with respect to Poisson$(B, 1)$, where $B \subseteq S$ Borel. For $x \in N_{p,\mathrm{f}}$, let $p(x)$ be a probability that determines whether to try to add (birth proposal) or remove (death proposal) a point to/from x. Furthermore, let $q_b(x, \cdot)$ be a density function on B used for generating a point location in a birth proposal. Similarly, let $q_d(x, \cdot)$ be a discrete density on x used for selecting a point from x in a death proposal. Then for $\xi \in B$ and $\eta \in x$, the *Hastings ratios* are defined by

$$r_b(x, \xi) = \frac{h(x \cup \xi)(1 - p(x \cup \xi))q_d(x \cup \xi, \xi)}{h(x)p(x)q_b(x, \xi)}$$

$$r_d(x, \eta) = \frac{h(x \backslash \eta)p(x \backslash \eta)q_b(x \backslash \eta)}{h(x)(1 - p(x))q_d(x, \eta)},$$

taking $a/0 = 1$ for $a \geq 0$. The acceptance probability for a birth proposal that updates the point pattern from x to $x \cup \xi$ is then taken as $\min\{1, r_b(x, \xi)\}$, while the acceptance probability of a death proposal updating x to $x \backslash \eta$ is $\min\{1, r_d(x, \eta)\}$. The algorithm requires an initial choice of point pattern Y_0, e.g. the empty point pattern or a realisation of a Poisson point process. Then the birth-death Metropolis-Hastings algorithm proceeds as follows.

Birth-death Metropolis-Hastings algorithm:
Given $Y_m = x \in N_{p,\mathrm{f}}$ for $m = 0, 1, \ldots$ we generate Y_{m+1} by:

1) Simulate $R'_m \sim \mathrm{unif}\,(0, 1)$ and $R''_m \sim \mathrm{unif}\,(0, 1)$, where unif$(0, 1)$ denotes the standard uniform distribution.

2) If $R'_m \leq p(x)$: Simulate $\xi_m \sim q_b\,(x, \cdot)$ and set

$$Y_{m+1} = \begin{cases} x \cup \xi_m & \text{if } R''_m \leq r_b\,(x, \xi_m) \\ x & \text{otherwise} \end{cases}.$$

3) If $R'_m > p(x)$:

i. If $x = \emptyset$: Set $Y_{m+1} = x$.
ii. If $x \neq \emptyset$: Draw $\eta_m \sim q_d(x, \cdot)$ and set

$$Y_{m+1} = \begin{cases} x \backslash \eta_m & \text{if } R''_m \leq r_d\,(x, \eta_m) \\ x & \text{otherwise} \end{cases}.$$

For our application $B = W$ and we choose $p(x) = 1/2$, $q_b(x, \xi) = \frac{1}{|W|}$ and $q_d(x, \eta) = \frac{1}{n(x)}$. That is, we propose birth and death with equal probability, choose a new point uniformly on W in case of a birth proposal and select an existing point uniformly from x in case of a death proposal. For the Strauss hardcore model specified by the unnormalised density

$$h_{(\beta,\gamma,R,h)}(x) = \beta^{n(x)}\gamma^{S_R(x)}\mathbb{I}(\|\eta - \xi\| > h, \forall \eta, \xi \in x),$$

the Hastings ratios are

$$\begin{aligned} r_b(x, \xi) &= \frac{\beta\gamma^{n(x\cap b(\xi,R))}|W|}{n(x \cup \xi)} \frac{\mathbb{I}(\|\eta_1 - \eta_2\| > h, \forall \eta_1, \eta_2 \in x \cup \xi)}{\mathbb{I}(\|\eta_1 - \eta_2\| > h, \forall \eta_1, \eta_2 \in x)} \\ r_d(x, \eta) &= \frac{n(x)}{\beta\gamma^{n((x\setminus\eta)\cap b(\eta,R))}|W|} \frac{\mathbb{I}(\|\xi_1 - \xi_2\| > h, \forall \xi_1, \xi_2 \in x\setminus\eta)}{\mathbb{I}(\|\xi_1 - \xi_2\| > h, \forall \xi_1, \xi_2 \in x)} \\ &= \frac{n(x)}{\beta\gamma^{n((x\setminus\eta)\cap b(\eta,R))}|W|}. \end{aligned}$$

In order to get the last equality consider the ratio of the two indicator functions in $r_d(x, \eta)$. Note that if the hardcore is fulfilled for x, then it will also be fulfilled for $x\setminus\eta$ and thus $\frac{0}{1}$ will never occur. This implies that the ratio will always equal 1, giving the simpler expression for $r_d(x, \eta)$.

For the Strauss model the Hastings ratios are

$$r_b(x, \xi) = \frac{\beta\gamma^{n(x\cap b(\xi,R))}|W|}{n(x \cup \xi)}, \quad r_d(x, \eta) = \frac{n(x)}{\beta\gamma^{n((x\setminus\eta)\cap b(\eta,R))}|W|}.$$

2.4 Summary statistics

In this section we give an introduction to the usual moment measures for point processes. Furthermore, some of the more popular functional summary statistics that typically are used for model checking will be presented and, when possible, related to the reduced Palm distribution. After a short note on estimation of summary statistics, the concept of global rank envelopes will be presented.

2.4.1 Moment measures

The nth-order moments of the count function $N(B)$ for Borel sets $B \subseteq S$, where especially the first- and second-order moments are of interest, is defined as follows.

Definition 2.4.1. For a point process X on S and a Borel set $D \subseteq S^n$ with $n \in \mathbb{N}$, the *nth order moment measure* and the *nth order factorial moment*

measure are defined by

$$\mu_p^{(n)}(D) = \mathrm{E}\left[\sum_{\xi_1,\ldots,\xi_n \in X} \mathbb{I}\left((\xi_1,\ldots,\xi_n) \in D\right)\right]$$

respectively

$$\alpha_p^{(n)}(D) = \mathrm{E}\left[\sum_{\xi_1,\ldots,\xi_n \in X}^{\neq} \mathbb{I}\left((\xi_1,\ldots,\xi_n) \in D\right)\right].$$

Applying the standard proof on the above definitions immediately yield the following useful results

$$\mathrm{E}\left[\sum_{\xi_1,\ldots,\xi_n \in X} h(\xi_1,\ldots,\xi_n)\right] = \int h(\xi_1,\ldots,\xi_n) d\mu_p^{(n)}(\xi_1,\ldots,\xi_n) \tag{2.10}$$

and

$$\mathrm{E}\left[\sum_{\xi_1,\ldots,\xi_n \in X}^{\neq} h(\xi_1,\ldots,\xi_n)\right] = \int h(\xi_1,\ldots,\xi_n) d\alpha_p^{(n)}(\xi_1,\ldots,\xi_n). \tag{2.11}$$

Note that $\mu_p = \mu_p^{(1)} = \alpha_p^{(1)}$. As stated in the following theorem the intensity measure for a Poisson point process is closely related to all of its nth order factorial moment measures.

Theorem 2.4.2. If $X \sim \text{Poisson}\,(S, \rho)$, then $\alpha_p^{(n)} = \mu_p^n$ for any $n \in \mathbb{N}$.

Proof. This result follows from Slivnyak-Meckes theorem stated in theorem 2.3.3, as

$$\begin{aligned}\int_{B_1\times\cdots\times B_n} d\alpha_p^{(n)}(\xi_1,\ldots,\xi_n) &= \int_{B_1}\cdots\int_{B_n}\prod_{i=1}^{n}\rho(\xi_i)d\xi_1\ldots d\xi_n \\ &= \int_{B_1}\cdots\int_{B_n} d\mu_p(\xi_1)\cdots d\mu_p(\xi_n)\end{aligned}$$

for all Borel sets $B_1 \times \cdots \times B_n \subseteq S^n$. Therefore, $\alpha_p^{(n)} = \mu_p^n$. □

For this thesis the intensity measure μ_p and the second-order factorial moment measure $\alpha_p^{(2)}$ will be the most relevant moments of the count function, and thus we shall only consider these onwards.

The relation between the Campbell measure and the intensity measure was discussed in section 2.2, but as it turns out the second order factorial moment measure can also be expressed in terms of the Campbell measure. Using the

standard proof on the definition of the Campbell measure (definition 2.2.1) gives

$$\mathrm{E}\left[\sum_{\xi\in X} h\left(\xi, X\backslash\xi\right)\right] = \int h\left(\xi, x\right) dC_p^!(\xi, x),$$

for any function $h : S \times N_{p,\mathrm{lf}} \to [0, \infty)$. Therefore,

$$\alpha_p^{(2)}(A \times B) = \mathrm{E}\left[\sum_{\xi\in X} \mathbb{I}\left(\xi \in A\right) \sum_{\eta\in X\backslash\xi} \mathbb{I}\left(\eta \in B\right)\right] = \int_A n\left(x_B\right) dC_p^!(\xi, x)$$

for all Borel sets $A \times B \subseteq \mathbb{R}^{2d}$.

Definition 2.4.3. If there exists a non-negative function $\rho^{(2)}$ such that the second order factorial moment measure can be written as

$$\alpha_p^{(2)}(A \times B) = \int_A \int_B \rho^{(2)}(\xi, \eta) d\xi d\eta$$

for all Borel sets $A, B \subseteq S$, then $\rho^{(2)}$ is called *the second-order intensity function*.

For $\xi, \eta \in S$, we can interpret $\rho^2(\xi, \eta)d\xi d\eta$ as the joint probability of observing two points from X in two infinitesimal balls with centres ξ and η and with volumes $d\xi$ and $d\eta$. For a Poisson point process, $\rho^{(2)}(\xi, \eta) = \rho(\xi)\rho(\eta)$ Lebesgue a.e. ξ, η due to theorem 2.4.2. In fact, the relation between the intensity and the second-order intensity is widely used to measure repulsion/aggregation of a point pattern, and since the Poisson point process is equivalent with CSR it is natural to define the following function.

Definition 2.4.4. If ρ and $\rho^{(2)}$ exist, then the *pair correlation function* is defined by

$$g(\xi, \eta) = \frac{\rho^{(2)}(\xi, \eta)}{\rho(\xi)\rho(\eta)}$$

with $\xi, \eta \in S$ and $a/0 = 0$ for $a \geq 0$.

Clearly, $g(\xi, \eta) \equiv 1$ for a Poisson process. Due to the interpretation of $\rho^{(2)}(\xi, \eta)$ as a joint probability, $g(\xi, \eta) > 1$ indicates that two points are more likely to occur simultaneously in the locations ξ and η under the point process of interest than for a Poisson process with the same intensity.

2.4.2 Functional summary statistics

To investigate how well a certain model describes an observed point pattern, different characteristics of the data should be assessed and compared to those of the model. For these types of assessments functional summary statistics are very useful tools. One of the more commonly used summary statistics is the so-called K-function, which is based on the second-order reduced moment measure.

Definition 2.4.5. Let ρ be the intensity function of X and define the measure

$$\mathcal{K}(B) = \frac{1}{|A|}\mathrm{E}\left[\sum_{\xi,\eta\in X}^{\neq} \frac{\mathbb{I}(\xi \in A, \eta - \xi \in B)}{\rho(\xi)\rho(\eta)}\right]$$

for Borel sets $A, B \subseteq S$ with $0 < |A| < \infty$ and taking $a/0 = 0$ for $a \geq 0$. If $\mathcal{K}$ does not depend on the choice of A, then X is said to be *second-order intensity reweighted stationary* and $\mathcal{K}$ is called the *second-order reduced moment measure*.

Suppose that the pair correlation function exists and is invariant under translation. With some abuse of notation we write $g(\xi - \eta) = g(\xi, \eta)$. It can then be shown (see e.g. details in Christensen and Christoffersen, 2015) that the point process is second-order intensity reweighted stationary and that

$$\mathcal{K}(B) = \int_B g(\xi)d\xi \tag{2.12}$$

for any $B \subseteq S$ Borel.

Considering $\mathcal{K}$ with a d-dimensional ball as the structuring element we obtain the K-function. This extends Ripley's K-function for stationary point processes to the case of second-order intensity reweighted stationarity and is due to Baddeley et al. (2000).

Definition 2.4.6. For a second-order intensity reweighted stationary point process, the K- and L- functions are defined by

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

for $r > 0$ and where $\omega_d = \pi^{d/2}/\Gamma(1 + d/2)$ denotes the volume of the d-dimensional unit ball.

Møller et al. (2016) propose to use a cylinder as the structuring element in $\mathcal{K}$ as a tool to detect anisotropy arising from columnar structures, i.e. point patterns where the points tend to lie around straight lines. Let $\mathbb{S}^{d-1} = \{u \in$

$\mathbb{R}^d : \|u\| \leq 1\}$ denote the unit sphere in $\mathbb{R}^d$ and let $e_d = (0, \ldots, 0, 1) \in \mathbb{R}^d$. Then

$$C(r, h) = \{(x_1, \ldots, x_d) \in \mathbb{R}^d : x_1^2 + \cdots + x_{d-1}^2 \leq r, |x_d| \leq h\}$$

is the d-dimension cylinder with midpoint at the origin of $\mathbb{R}^d$, radius $r > 0$, height $2h > 0$ and direction e_d. To obtain a cylinder with direction determined by $u \in \mathbb{S}^{d-1}$, consider a $d \times d$ rotation matrix $\mathcal{O}_u$ satisfying $u = \mathcal{O}_u e_d$. Then the d-dimensional cylinder with direction u, midpoint at the origin, radius r and height $2h$ is determined by $C_u(r, h) = \mathcal{O}_u C(r, h)$.

Definition 2.4.7. For a second-order intensity reweighted stationary point process *the cylindrical K-function in the direction* u is defined as

$$K_u(r, h) = \mathcal{K}\left(C_u(r, h)\right)$$

for $u \in \mathbb{S}^{d-1}$, $r > 0$ and $h > 0$.

Consider a second-order intensity reweighted stationary point process X. Then, per definition of $\mathcal{K}$ and equation (2.11), it follows that

$$\begin{aligned}\int\int \mathbb{I}(\xi \in A, \eta \in B) d\mathcal{K}(\eta) d\xi &= \mathrm{E}\left[\sum_{\xi,\eta \in X}^{\neq} \frac{\mathbb{I}(\xi \in A, \eta - \xi \in B)}{\rho(\xi)\rho(\eta)}\right] \\ &= \int \frac{\mathbb{I}(\xi \in A, \eta - \xi \in B)}{\rho(\xi)\rho(\eta)} d\alpha_p^{(2)}(\xi, \eta)\end{aligned}$$

for all Borel sets $A, B \subseteq S$. Using the standard proof, we can then relate $\mathcal{K}$, $\alpha_p^{(2)}$ and ρ by

$$\int\int h(\xi, \eta) d\mathcal{K}(\eta) d\xi = \int \frac{h(\xi, \eta - \xi)}{\rho(\xi)\rho(\eta)} d\alpha_p^{(2)}(\xi, \eta)$$

for functions $h : S \times S \to [0, \infty)$. The second-order reduced moment measure also has an interpretation as an expectation under the reduced Palm distribution in form of

$$\mathcal{K}(B) = \mathrm{E}\left[\sum_{\eta \in X_\xi^!} \frac{\mathbb{I}(\eta - \xi \in B)}{\rho(\eta)}\right]$$

for Lebesgue-a.e. $\xi \in S$. This follows by Campbell-Meckes theorem stated in equation (2.3),

$$\begin{aligned}|A|\mathcal{K}(B) &= \mathrm{E}\left[\sum_{\xi \in X} \sum_{\eta \in X \setminus \xi} \frac{\mathbb{I}(\xi \in A, \eta - \xi \in B)}{\rho(\xi)\rho(\eta)}\right] \\ &= \int \mathbb{I}(\xi \in A) \int \sum_{\eta \in x} \frac{\mathbb{I}(\eta - \xi \in B)}{\rho(\eta)} dP_\xi^!(x) d\xi\end{aligned} \tag{2.13}$$

for all Borel sets $A, B \subseteq S$ with $0 < |A| < \infty$.

Consider now a stationary point process X on $\mathbb{R}^d$. Imitating the steps in (2.13) for a constant intensity $0 < \rho < \infty$ and switching to the reduced Palm distribution at a typical point, we get

$$\mathcal{K}(B) = \frac{1}{\rho|A|} \int \mathbb{I}(\xi \in A) \int \sum_{\eta \in x} \mathbb{I}(\eta \in B) dP_0^!(x) d\xi = \mathrm{E}[n(X_0^! \cap B)]/\rho,$$

where $X_0^! \sim P_0^!$ and $A, B \subseteq \mathbb{R}^d$ are Borel sets. Specifically, for $B = b(0, r)$ for $r > 0$, we have $\rho K(r) = \mathrm{E}[n(X_0^! \cap b(0, r))]$, i.e. $\rho K(r)$ can be interpreted as the expected number of further points within distance r of the typical point. For the cylindrical K-function, we interpret $\rho K_u(r, h)$ as the mean number of further points within a cylinder $\xi_0 + C_u(r, h)$ with midpoint in a typical point, ξ_0.

Consider a Poisson point process X on S. Since the pair correlation function is always 1 for the Poisson process, it follows from equation (2.12) that $\mathcal{K}(B) = |B|$ for all Borel sets $B \subseteq S$. Thus, $K(r) = \omega_d r^d$ and $L(r) = r$ for all $r > 0$. Therefore, for a general point process, $L(r) - r > 0$ (resp. $L(r) - r < 0$) indicates aggregation (resp. repulsion) within distance $r > 0$. Furthermore, for the Poisson process, $K_u(r, h) = 2\omega_{d-1} r^{d-1} h$ for all $u \in \mathbb{S}^{d-1}$, $r > 0$ and $h > 0$. Hence, for a general point process $K_u(r, h) > 2\omega_{d-1} r^{d-1} h$ indicates columnar clustering.

Definition 2.4.8. Suppose that X is a stationary point process on $\mathbb{R}^d$ with intensity $0 < \rho < \infty$. For $r > 0$ the *empty space function*, or simply the *F-function*, is defined by

$$F(r) = P(X \cap b(0, r) \neq \emptyset),$$

and the *nearest neighbour function*, or simply the *G-function*, is defined by

$$G(r) = \frac{1}{\rho|A|} \mathrm{E} \left[\sum_{\xi \in X_A} \mathbb{I}((X \backslash \xi) \cap b(\xi, r) \neq \emptyset) \right]$$

for any Borel set $A \subset \mathbb{R}^d$ with $0 < |A| < \infty$. On this basis, the *J-function* is defined by

$$J(r) = \frac{1 - G(r)}{1 - F(r)}$$

for $F(r) < 1$.

If X is a stationary Poisson point proces with intensity ρ, it follows from equation (2.5) that

$$F(r) = 1 - P(N(b(0, r)) = 0) = 1 - \exp(-\rho \omega_d r^d)$$

for $r > 0$. Furthermore, due to the extended Slivnyak-Mecke theorem (found in theorem 2.3.3) and equation (2.5),

$$\begin{aligned} G(r) &= \frac{1}{|A|} \int \mathrm{E}[\mathbb{I}(\xi \in A, X \cap b(\xi, r) \neq \emptyset)] d\xi \\ &= \frac{1}{|A|} \int_A (1 - P(X \cap b(\xi, r) = \emptyset))\, d\xi \\ &= 1 - \exp(-\rho \omega_d r^d) \end{aligned} \tag{2.14}$$

for $r > 0$. Thus, $J(r) = 1$ for all $r > 0$.

For any stationary point process X, the G-function has an interpretation as a probability under the reduced Palm distribution at a typical point. This follows from equation (2.4), since

$$G(r) = \frac{1}{|A|} \int \int \mathbb{I}(\xi \in A, (x + \xi) \cap b(\xi, r) \neq \emptyset) dP_0^!(x) d\xi = \int \mathbb{I}(x \cap b(0, r) \neq \emptyset) dP_0^!(x)$$

$$= P\left(X_0^! \cap b(0, r) \neq \emptyset\right)$$

for $r > 0$. The G-function can thus be interpreted as the probability that the point process has any further points within a certain distance of the typical point. Note that this result gives an alternative way of deriving expression (2.14) by using theorem (2.3.4) and void probabilities.

For $r > 0$, $F(r) < G(r)$ (equivalently $J(r) < 1$) indicates aggregation, while $F(r) > G(r)$ (equivalently $J(r) > 1$) indicates regularity.

Estimation

Estimation for the presented summary statistics (except for the cylindrical K-function) were discussed in Christensen and Christoffersen (2015) and thus will only be described very briefly. Several nonparametric estimates for the various summary statistics have been given in the literature, see e.g. Møller and Waagepetersen (2004) for an overview of estimates and relevant references. Most of the estimates take into account the edge effect that arise when observing the point patterns on a window $W \subset S$. Different choices of edge corrections are available. A popular choice of edge correction factor, which will also be used throughout this thesis, is $1/|W \cap W_{\eta - \xi}|$, where $W_\xi = \{\eta + \xi : \eta \in W\}$ for $\xi \in S$. This factor appears for example in the unbiased estimate of $\mathcal{K}(B)$ (see e.g. Møller and Waagepetersen, 2004) given by

$$\hat{\mathcal{K}}(B) = \sum_{\xi, \eta \in x}^{\neq} \frac{\mathbb{I}(\eta - \xi \in B)}{\rho(\xi)\rho(\eta)|W \cap W_{\eta - \xi}|},$$

where $B \subseteq S$ is Borel and x is an observed point pattern on a bounded window $W \subseteq S$. Furthermore, it is assumed that $|W \cap W_\xi| > 0$ for all $\xi \in B$. The

estimate of the cylindrical K-function is found in Møller et al. (2016) and simply corresponds to the above estimate of $\mathcal{K}(B)$ with B as the directional cylinder.

2.4.3 Global envelopes

In Christensen and Christoffersen (2015) pointwise envelopes were used to make inference based on the summary statistics introduced in section 2.4.2 that depends on some distance r. The pointwise envelopes can be used to make inference based on a fixed distance, but most often we are interested in making inference based on an interval of distances. Using the pointwise envelopes for this kind of multiple testing leads to a greatly underestimated type I error, as pointed out and discussed by several authors (e.g. Ripley (1977) and later Loosmore and Ford (2006) and Baddeley et al. (2014)). This pitfall has led to the use of *global envelopes* that allow simultaneous inference to be made in a way that controls the type I error rate.

The pointwise envelopes are briefly described as follows. Let X_1 denote an observed point pattern and let $X_2, \ldots, X_{s+1}$ denote s point patterns simulated under a null hypothesis. Furthermore, for $r > 0$ and $i = 1, \ldots, s+1$, let $\hat{T}_i(r)$ denote the estimator of some summary statistic (e.g. the K- or J- function) for X_i. Let also $\hat{T}^{(k)}_{low}(r)$ and $\hat{T}^{(k)}_{upp}(r)$ denote the kth smallest and largest values among $\hat{T}_1(r), \ldots, \hat{T}_{s+1}(r)$ for some choice of $k \in \{1, 2, \ldots, \lfloor (s+1)/2 \rfloor\}$ and r. With this set-up $\hat{T}^{(k)}_{low}(r)$ and $\hat{T}^{(k)}_{upp}(r)$ are the boundaries of the *pointwise envelopes*. If there are no ties in the ordering of $\hat{T}_i(r)$, we have that $\hat{T}_1(r)$ lies outside the pointwise envelopes with probability $\alpha = 2k/(s+1)$ for a fixed r.

As alternatives to pointwise envelopes, Myllymäki et al. (2015) suggest approaches to construct a global envelope test on an interval I. Of these we will only consider the *rank envelope test*, where the envelope boundaries are directly based on $\hat{T}_i$, $i = 1, \ldots, s+1$. To each $\hat{T}_i$ we assign an *extreme rank* R_i, defined by

$$R_i = \max\{k : \hat{T}^{(k)}_{low}(r) \leq \hat{T}_i(r) \leq \hat{T}^{(k)}_{upp}(r) \text{ for all } r \in I\}$$

for $i = 1, \ldots, s+1$. The extreme ranks only give a weak ordering of $\hat{T}_i$ (as $\max R_i < s+1$), and therefore do not lead directly to an exact test. Even though Myllymäki et al. (2015) introduces a way of constructing an exact test, we will only present the more simple approach, where a conservative and liberal p-value is constructed directly from the position of $\hat{T}_1$ among the estimated summary statistics. More precisely, define

$$p_- = \frac{1}{s+1} \sum_{i=1}^{s+1} \mathbb{I}(R_i < R_1) \quad \text{and} \quad p_+ = \frac{1}{s+1} \sum_{i=1}^{s+1} \mathbb{I}(R_i \leq R_1).$$

Then p_- and p_+ forms the lower and upper bounds for the p-value of the global test concerning the null hypothesis. For a chosen significance level α, the test obtained by rejecting the null hypothesis when $p_- \leq \alpha$ is liberal, while rejecting when $p_+ \leq \alpha$ leads to a conservative test, cf. Myllymäki et al. (2015). That is, compared to an exact p-value the liberal p_- tends to overstate the evidence against the null hypothesis, while the conservative p_+ tends to understate it. We can however at significance level α safely reject the null hypothesis if $p_+ \leq \alpha$ or infer that there is no evidence against the null hypothesis if $p_- > \alpha$. To get a graphical interpretation, Myllymäki et al. (2015) define the $100(1-\alpha)\%$ *global rank envelope* by the bounds $\hat{T}_{low}^{(k_\alpha)}(r)$ and $\hat{T}_{upp}^{(k_\alpha)}(r)$ for $r \in I$, where

$$k_\alpha = \max\left\{ k : \sum_{i=1}^{s+1} \mathbb{I}(R_i < k) \leq \alpha(s+1) \right\}.$$

They then show that $\hat{T}_1(r)$ lies strictly outside the $100(1-\alpha)\%$ global rank envelope for at least one $r \in I$ if and only if $p_+ \leq \alpha$, which leads to a rejection of the null hypothesis. Furthermore, if there are no pointwise ties in $\hat{T}_i(r)$ for $i = 1, \ldots, s+1$ with probability 1, then $\hat{T}_1(r)$ lies strictly inside the global rank envelopes for all $r \in I$ if and only if $p_- > \alpha$. Lastly, $\hat{T}_1(r)$ does not fall outside the envelopes, but coincides with $\hat{T}_{low}^{(k_\alpha)}(r)$ or $\hat{T}_{upp}^{(k_\alpha)}(r)$ for at least one $r \in I$ if and only if $p_- \leq \alpha < p_+$. According to Myllymäki et al. (2015) the width of the p-interval, (p_-, p_+), depends on the number of simulations s, the smoothness of the functional summary statistic and the value of R_1. In this thesis we follow their general recommendation of using $s = 2499$ for $\alpha = 0.05$.

2.5 Analysis of pyramidal cells' nucleolus and apex locations

In this section the locations of the nucleoli and apexes are analysed as two separate point patterns. In Christensen and Christoffersen (2015) we found that the nucleolus point pattern posses a higher degree of regularity than the proposed models. Even though this conclusion could change when considering global rather than pointwise envelopes, we will here propose a new model: the Strauss hardcore model presented in section 2.3.2, that besides the hardcore also allows for further interaction between pairs of points.

In this thesis and in Christensen and Christoffersen (2015) the pair correlation function as well as the K- and J-functions are used for preliminary analysis and model check. However, these summary statistics or their presented estimates require stationarity, which does not apply to the Strauss hardcore (and Strauss) point process (or generally a finite Markov point process) as it is defined on a bounded space. Strictly, we could extend the definition of

Markov point processes to cover the case on $\mathbb{R}^d$, for example by doing a *local specification* (for details see e.g. Møller and Waagepetersen (2004) and the references therein). We will however not do this as the empirical estimates of the summary statistics still make sense, even though they not necessarily have the same interpretation.

The analysis was performed using `R` version 3.2.1. The `R`-package `spatstat` was used for estimating the pair correlation function and the K-function as well as for simulating under Poisson. The approximate (profile) pseudo-likelihood was optimised using the `optim`-function from the `stats`-package. Estimation of the J-function and the cylindrical K-function for three-dimensional point patterns and simulation under the Strauss and Strauss hardcore models using the Metropolis-Hastings algorithm were implemented by the project group, partially using `Rcpp`. Global rank envelopes were computed using the `R`-packages `spptest` and `spatstat`.

2.5.1 Preliminary plots

The nucleolus and apex point patterns are both located on a three-dimensional box with dimension $[0, 1382] \times [0, 145] \times [0, 505]\,\mu m^3$, each containing 2085 nucleoli/apexes. Figure 2.1 and 2.3 show projections of the nucleolus and apex locations onto the xy-, xz- and yz-plane. Considering the projections onto the xz- and yz-plane there is an absence of nucleoli at the top and some in the bottom, while the absence of apexes is mostly seen at the bottom. The non-parametric kernel estimate of the inhomogeneous intensity function plotted in figure 2.2 and 2.4 also illustrate these bare areas. This may indicate that the apex tend to lie closer to the pile surface than its related nucleolus, i.e. the pyramidal cells tend to have a common direction. We shall discuss this further in section 3.7. In order to eliminate this source of inhomogeneity each window was eroded. The eroded window for the nucleolus point pattern is chosen as $[0, 1382] \times [0, 145] \times [10, 485]\,\mu m^3$, resulting in a loss of 25 points. For the apex point pattern the eroded window is chosen as $[0, 1382] \times [0, 145] \times [20, 505]\,\mu m^3$, resulting in a loss of 7 points. Figure 2.5 - 2.8 display the projected apex and nucleolus locations and their corresponding kernel estimates of the inhomogeneous intensity function for the eroded window. Even though these plots still may indicate slight inhomogeneity, it will in the following be assumed that the underlying point processes are homogeneous. The intensity can therefore be estimated by dividing the observed number of points with the size of the observation window, resulting in the estimates

$$\hat{\rho}_{nucl} = 2.16 * 10^{-5} \quad \text{and} \quad \hat{\rho}_{apex} = 2.14 * 10^{-5}.$$

In figure 2.1 and more clearly in figure 2.3 thick, empty, parallel lines traverse the window in the xz-plane. This tendency is more evident in figure 2.9, where

the z-axis has been stretched. Consultation with Jens R. Nyengaard revealed that this most likely is an artefact originating from the data collection. Data was collected using a counting frame, quadratic in the xz-plane. Presumably, the step length in the z-direction was longer than the counting frame leading to areas, where the cells was not counted. The data should most likely be recollected in order to perform a meaningful analysis. We will however continue to analyse these data in order to demonstrate some of the presented theory.

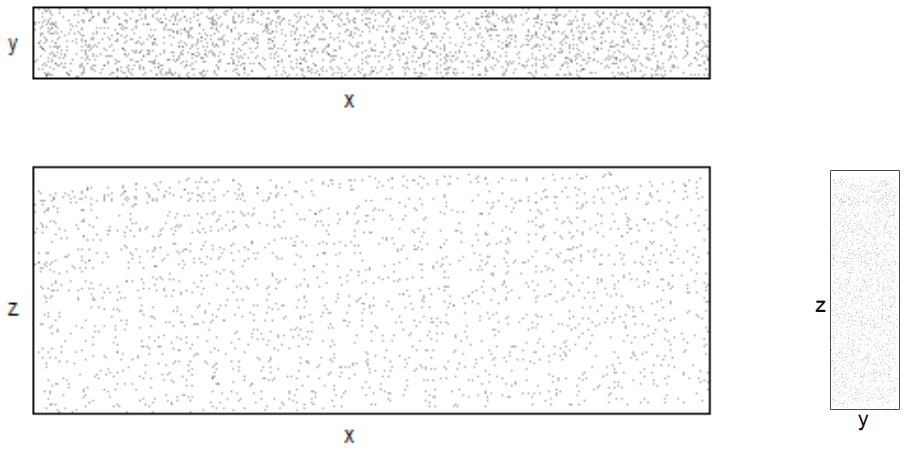

Figure 2.1: Projection of nucleolus locations onto the xy-plane (top, window: $[0, 1382] \times [0, 145]$ μm^2), xz-plane (bottom left, window: $[0, 1382] \times [0, 505]$ μm^2) and yz-plane (bottom right, window: $[0, 145] \times [0, 505]$ μm^2).

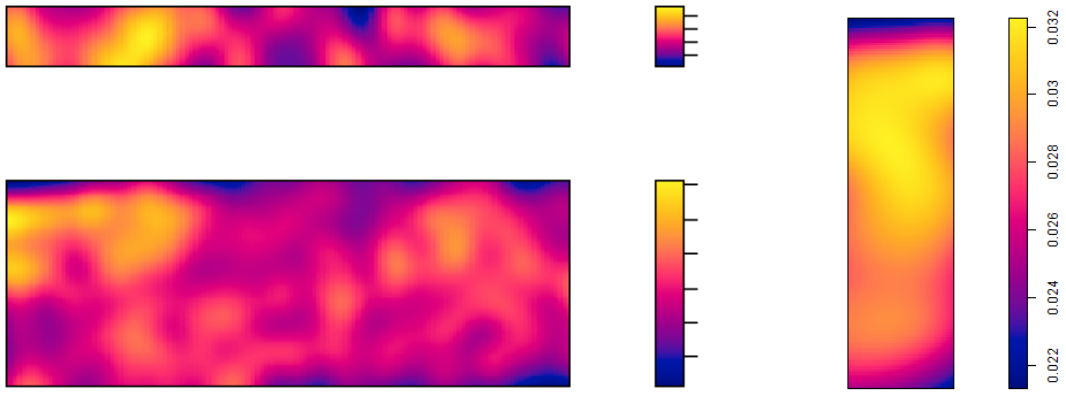

Figure 2.2: Non-parametric kernel estimate of the inhomogeneous intensity function for the projected nucleolus locations from figure 2.1. The plot for the yz-plane has been scaled, so proportions between plots are physically incorrect.

Estimates of the pair correlation function as well as the J- and K-function for both point patterns are shown in figure 2.10 and 2.11 along with 95%-global

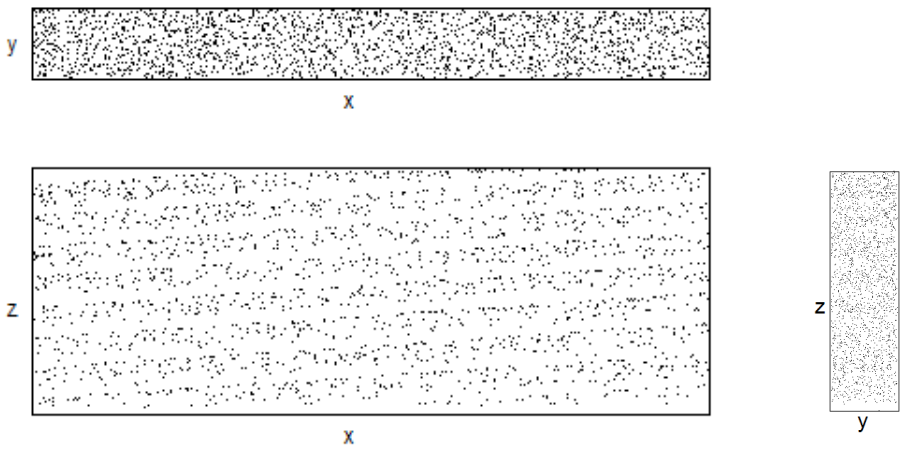

Figure 2.3: Projection of apex locations onto the xy-plane (top, window: $[0, 1382] \times [0, 145]$ μm^2), xz-plane (bottom left, window: $[0, 1382] \times [0, 505]$ μm^2) and yz-plane (bottom right, window: $[0, 145] \times [0, 505]$ μm^2).

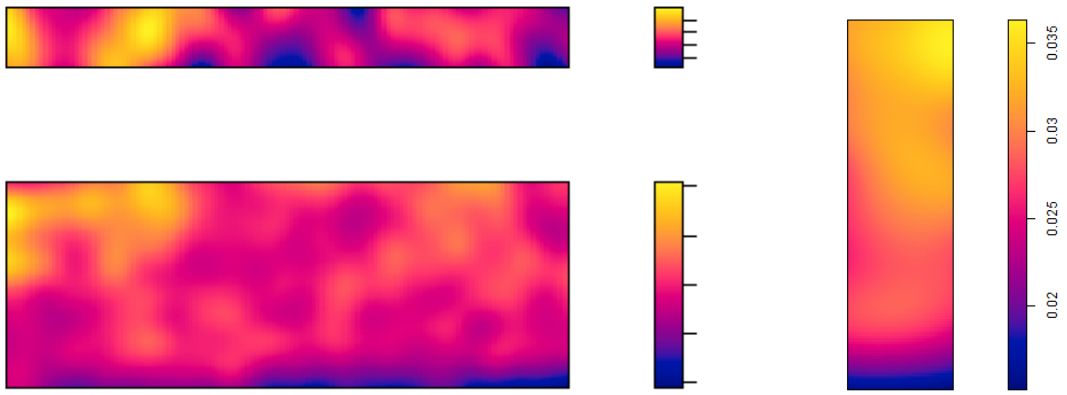

Figure 2.4: Non-parametric kernel estimate of the inhomogeneous intensity function for the projected apex locations from figure 2.3. The plot for the yz-plane has been scaled, so proportions between plots are physically incorrect.

rank envelopes under the homogeneous Poisson point processes with intensity $\hat{\rho}_{nucl}$ and $\hat{\rho}_{apex}$, respectively. In this type of plots, involving envelopes, red marks indicate values of the estimated test function that falls strictly outside the envelopes. The p-interval for the rank envelope test is also displayed in these plots. As $p_+ < 0.05$ for all of the considered summary statistics, we can safely reject the Poisson models at significance level 0.05. The envelopes reveal that both point patterns inherit regularity (mostly for distances less than 25 μm). In figure 2.12 and 2.13 $\hat{K}_u(r, t) - 2\pi hr^2$ is plotted for u equal to each of the three main axes and $h = 40, 60$ and 80 μm for respectively the nucleolus

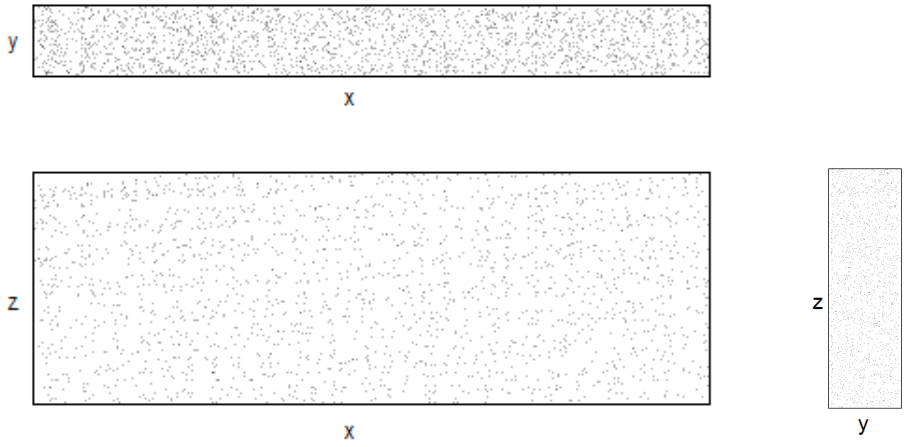

Figure 2.5: After erosion of the window. Projection of nucleolus locations onto the xy-plane (top, window: $[0, 1382] \times [0, 145]$ μm^2), xz-plane (bottom left, window: $[0, 1382] \times [10, 485]$ μm^2) and yz-plane (bottom right, window: $[0, 145] \times [10, 485]$ μm^2).

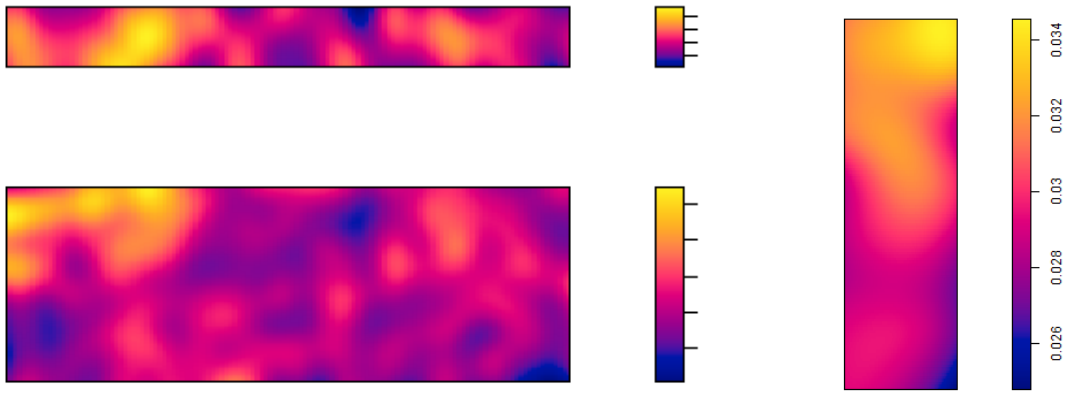

Figure 2.6: Non-parametric kernel estimate of the inhomogeneous intensity function for the projected nucleolus locations on the eroded window, from figure 2.5. The plot for the yz-plane has been scaled, so proportions between plots are physically incorrect.

and apex point pattern along with envelopes under Poisson. Note that $2\pi hr^2$ is the theoretical value of $K_u(r, t)$ for any Poisson process. For the nucleolus point pattern, the null hypothesis, that the point pattern is Poisson, is rejected when $h = 40$ μm or $h = 60$ μm and u is either in the x- or y-direction. The null hypothesis is however accepted for the remaining combination of cylinder heights and directions. When the estimated cylindrical K-function lies above the Poisson envelopes, it indicates columnar structures, but here the estimate

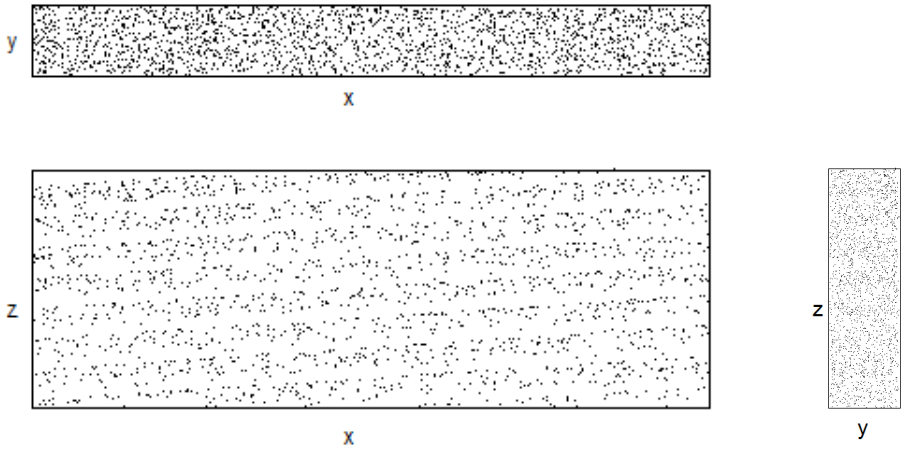

Figure 2.7: After erosion of the window. Projection of apex locations onto the xy-plane (top, window: $[0, 1382] \times [0, 145]$ μm^2), xz-plane (bottom left, window: $[0, 1382] \times [20, 505]$ μm^2) and yz-plane (bottom right, window: $[0, 145] \times [20, 505]$ μm^2).

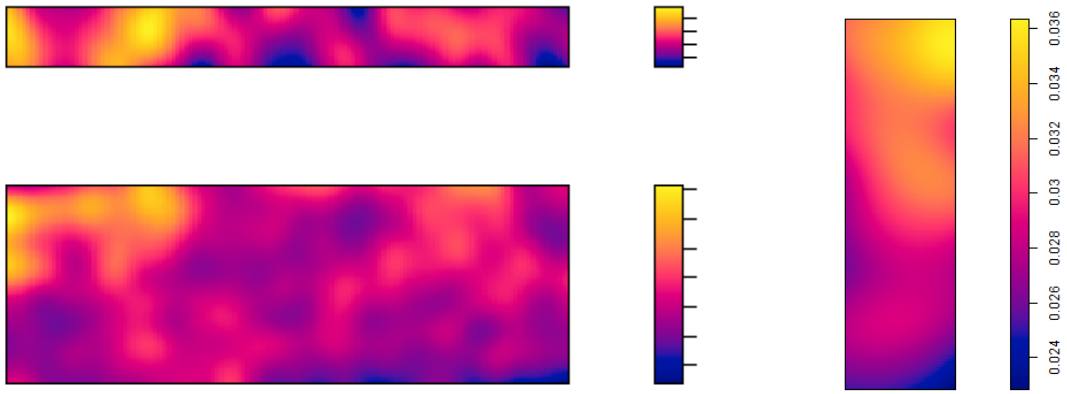

Figure 2.8: Non-parametric kernel estimate of the inhomogeneous intensity function for the projected apex locations on the eroded window, from figure 2.7. The plot for the yz-plane has been scaled, so proportions between plots are physically incorrect.

lies either within or below the envelopes. This could possibly be due to the empty belts seen in figure 2.9. Furthermore, there is no evidence supporting the minicolumn hypothesis, as the estimated cylindrical K-function in the z-direction lies inside the Poisson envelopes for all three values of h. Again, this conclusion may be affected by the flaws in the data collection. For the apex point pattern the hypothesis of a Poisson process is rejected for all the considered combinations of h and u, except for $h = 40$ and $u = (1, 0, 0)$. The

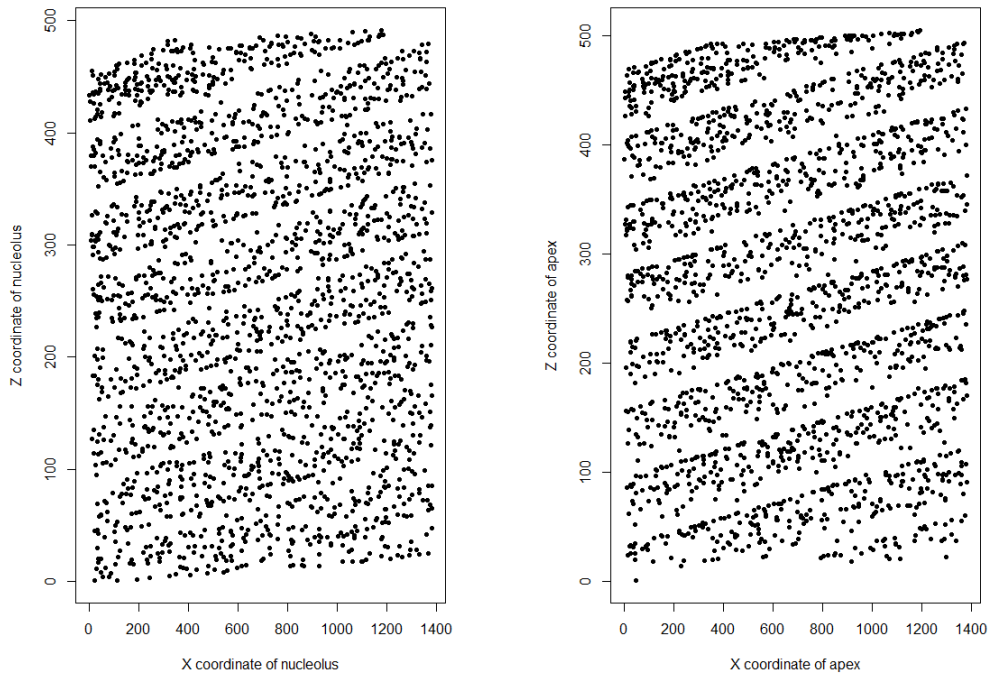

Figure 2.9: Nucleolus (left) and apex (right) locations projected onto the xz-plane with the z-axis is stretched in order to make the empty belts more visible.

results are somewhat diffuse and no clear indication of minicolumns are seen. Besides flaws in the data collection the results may also reflect that the apex locations not are meaningful. In figure 2.14 estimates of the pair correlation function as well as the K- and J-functions for the nucleolus and apex point patterns are plotted in order to compare the two. Each summary statistic leads to the same conclusion, that the apex point pattern is not as regular as the nucleolus point pattern.

2.5.2 Fitting a Strauss hardcore model

In section 2.3.2 we found that the MLE of the hardcore, h, in the Strauss hardcore model is the minimum distance between a pair of point in the observed point pattern, giving the estimates

$$\hat{h}_{nucl} = 2.36 \quad \text{and} \quad \hat{h}_{apex} = 1.62$$

for the nucleolus and apex point patterns, respectively. The remaining parameters, β, γ and R, are estimated by numerical optimisation of the approximate log profile pseudolikelihood derived in section 2.3.2. A choice of four different starting values for the optimisation give somewhat different estimates (see table 2.1). The start value of β is chosen as the homogeneous estimate of the

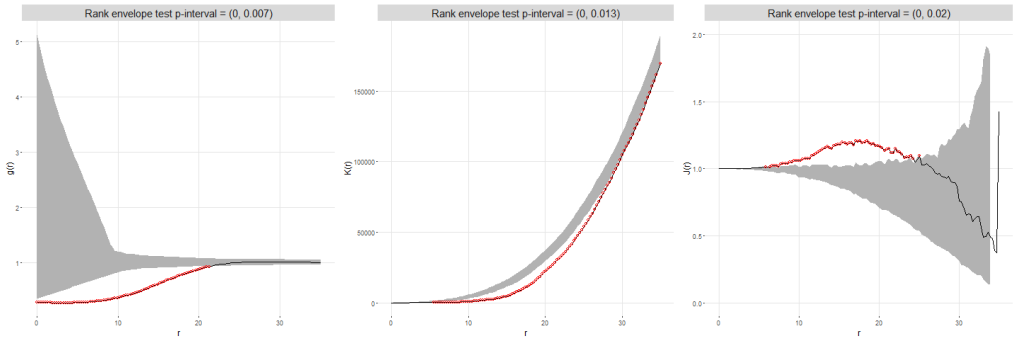

Figure 2.10: 95%-global rank envelopes under Poisson together with the estimate for the nucleolus point pattern of the pair correlation function (left), K-function (middle) and J-function (right) for $0 < r \leq 35$.

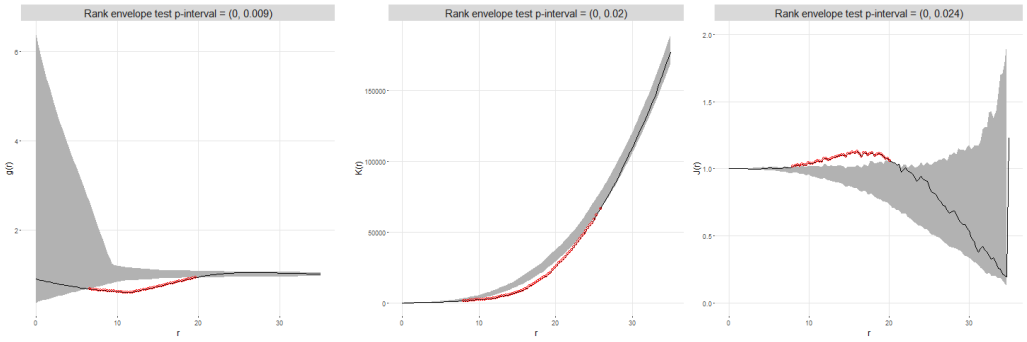

Figure 2.11: 95%-global rank envelopes under Poisson together with the estimate for the apex point pattern of the pair correlation function (left), K-function (middle) and J-function (right) for $0 < r \leq 35$.

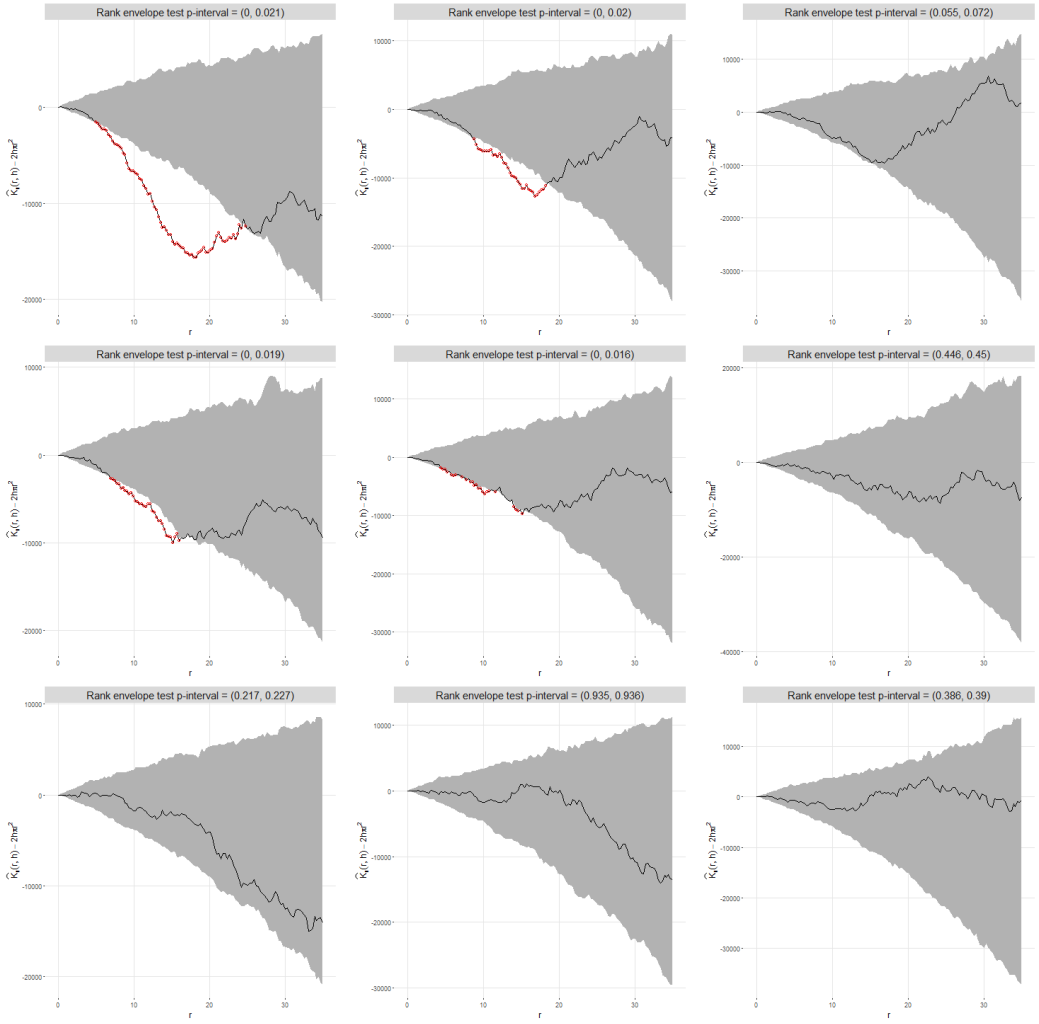

Figure 2.12: 95%-global rank envelopes under Poisson for the cylindrical K-function minus the theoretical value for Poisson concerning the nucleolus point pattern. Computed for the three main directions (x, y and z from top to bottom), $h = 40, 60$ and 80 μm (from left to right) and $0 < r \leq 35$.

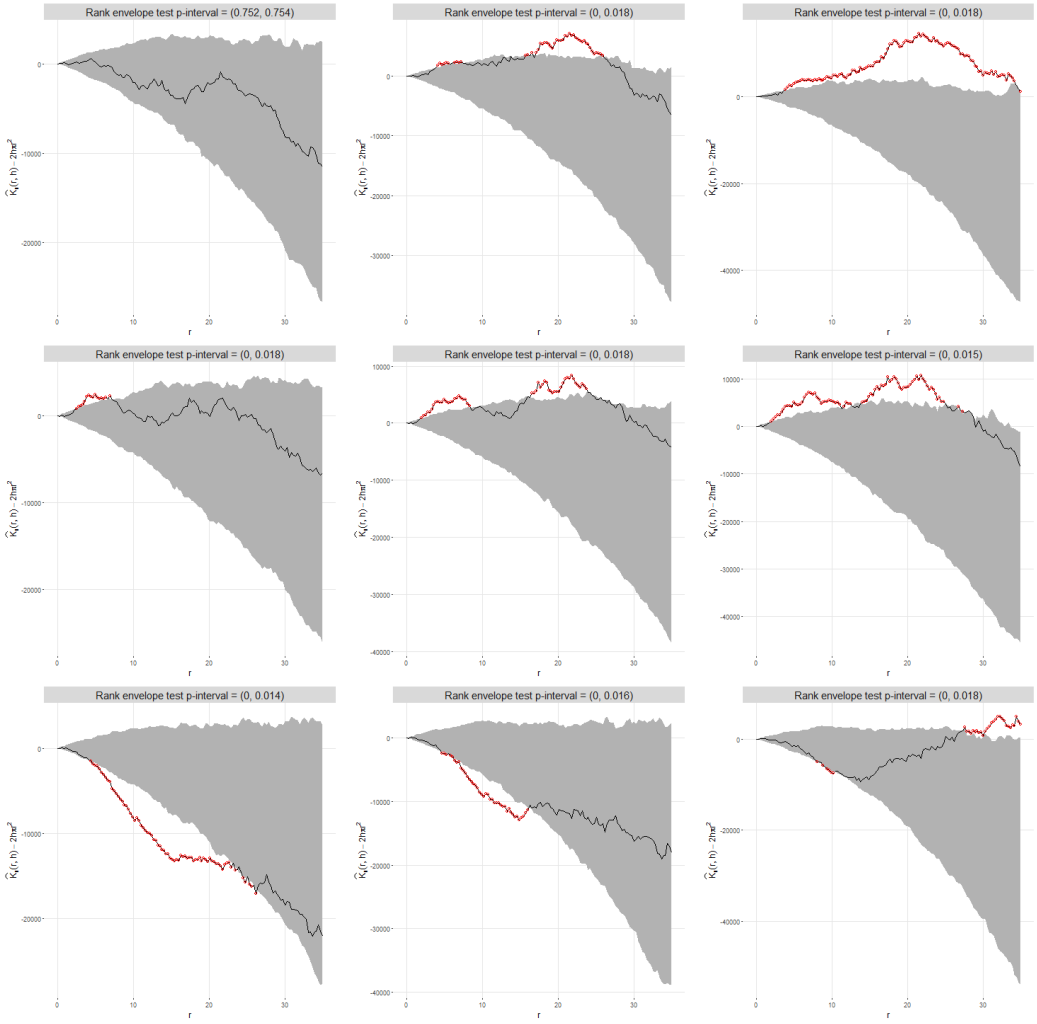

Figure 2.13: 95%-global rank envelopes under Poisson for the cylindrical K-function minus the theoretical value for Poisson concerning the apex point pattern. Computed for the three main directions (x, y and z from top to bottom), $h = 40, 60$ and 80 μm (from left to right) and $0 < r \leq 35$.

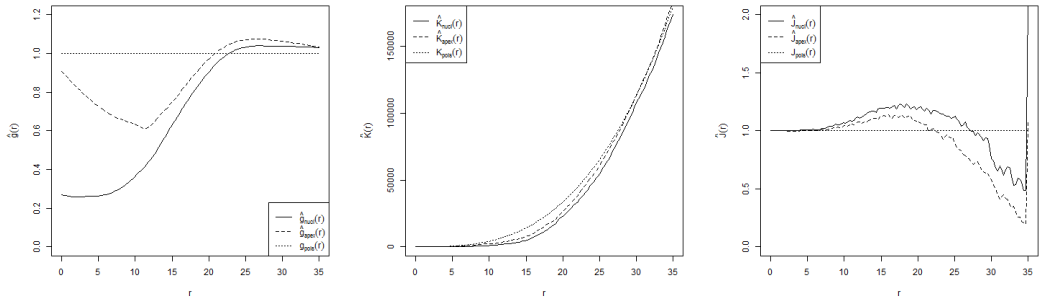

Figure 2.14: Estimated summary statistics for the nucleolus (solid) and apex (dashed) point pattern along with the theoretical value for the Poisson process (dotted), with $0 < r \leq 35$. Pair correlation function (left), K-function (middle) and J-function (right).

Data	$\boldsymbol{\beta_0}$	$\boldsymbol{\gamma_0}$	$\boldsymbol{R_0}$	$\hat{\boldsymbol{\beta}}$	$\hat{\boldsymbol{\gamma}}$	$\hat{\boldsymbol{R}}$	**LPPL**
nucleolus	$2.16 * 10^{-5}$	0.5	10	$2.97 * 10^{-5}$	0.29	15.37	-23976
nucleolus	$2.16 * 10^{-5}$	0.5	20	$2.52 * 10^{-5}$	0.71	20.00	-24089
nucleolus	$2.16 * 10^{-5}$	1	10	$2.42 * 10^{-5}$	0.22	10.92	-24060
nucleolus	$2.16 * 10^{-5}$	0.5	30	$2.60 * 10^{-5}$	0.92	30.74	-24163
apex	$2.14 * 10^{-5}$	0.5	10	$2.42 * 10^{-5}$	0.57	15.64	-24312
apex	$2.14 * 10^{-5}$	0.5	20	$2.58 * 10^{-5}$	0.76	20.75	-24362
apex	$2.14 * 10^{-5}$	1	10	$2.32 * 10^{-5}$	0.47	10.86	-24367
apex	$2.14 * 10^{-5}$	0.5	30	$2.34 * 10^{-5}$	0.96	30.60	-24413

Table 2.1: Results for each optimisation with start values (β_0, γ_0, R_0). LPPL is the value of the maximised log profile pseudolikelihood for the Strauss hardcore model.

intensity, while the start values of γ and R are chosen as 0.5 or 1 respectively 10, 20 or 30. Not all combinations are pursued. For the nucleolus point pattern the estimates

$$\hat{\beta}_{nucl} = 2.97 * 10^{-5}, \quad \hat{\gamma}_{nucl} = 0.29, \quad \hat{R}_{nucl} = 15.37$$

yield the largest value of the pseudolikelihood. For the apex point pattern the pseudolikelihood is largest for the estimates

$$\hat{\beta}_{apex} = 2.42 * 10^{-5}, \quad \hat{\gamma}_{apex} = 0.57 \quad \hat{R}_{apex} = 15.64.$$

Model check

For simulation of 2499 point patterns under the two fitted Strauss hardcore models, the Metropolis-Hastings algorithm was applied in accordance with the

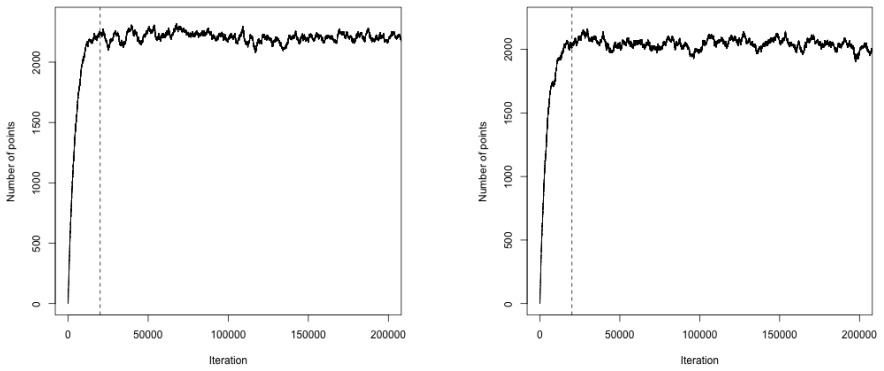

Figure 2.15: Number of points in the first 200,000 simulated point patterns from the fitted Strauss hardcore models for the nucleolus (left) and apex (right) point patterns.

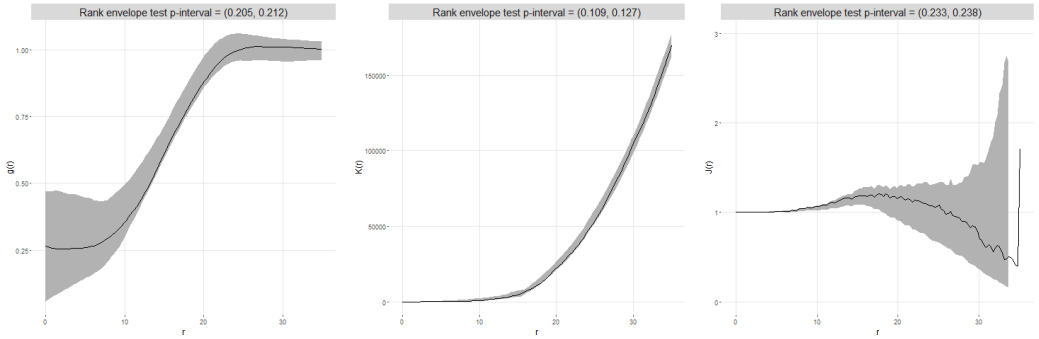

Figure 2.16: 95%-global rank envelopes under fitted Strauss hardcore model together with the estimate for the nucleolus point pattern of the pair correlation function (left), K-function (middle) and J-function (right) for $0 < r \leq 35$.

description in section 2.3.2. In order to ensure the hardcore condition $Y_0 = \emptyset$ was chosen. A total of 2.6 million iterations for each model were performed. Burn-in was chosen to be 20,000 iterations as the number of points in the simulated point patterns seem to have stabilised at this stage (see figure 2.15). Every thousand of the remaining simulations were extracted, and of these 2499 point patterns were sampled and used for creating global rank envelopes. The resulting envelopes and the associated p-intervals for the pair correlation function and for the K- and J-functions are shown in figure 2.16 and 2.17. For the nucleolus point pattern, $p_- > 0.05$ for the three p-intervals, i.e. there is no evidence against the fitted Strauss hardcore model at significance level 0.05. For the apex point pattern, the K-function provide slight evidence against the Strauss hardcore model with $p_- = 0.046$. However, this is not a strong evidence against the model, especially as the two other summary statistics gives p-intervals with $p_- > 0.05$.

In order to investigate whether the apex and nucleolus point patterns can be

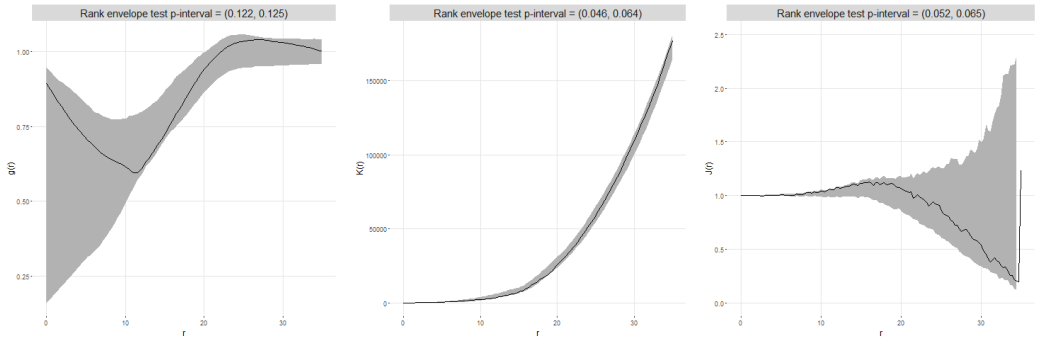

Figure 2.17: 95%-global rank envelopes under fitted Strauss hardcore model together with the estimate for the apex point pattern of the pair correlation function (left), K-function (middle) and J-function (right) for $0 < r \leq 35$.

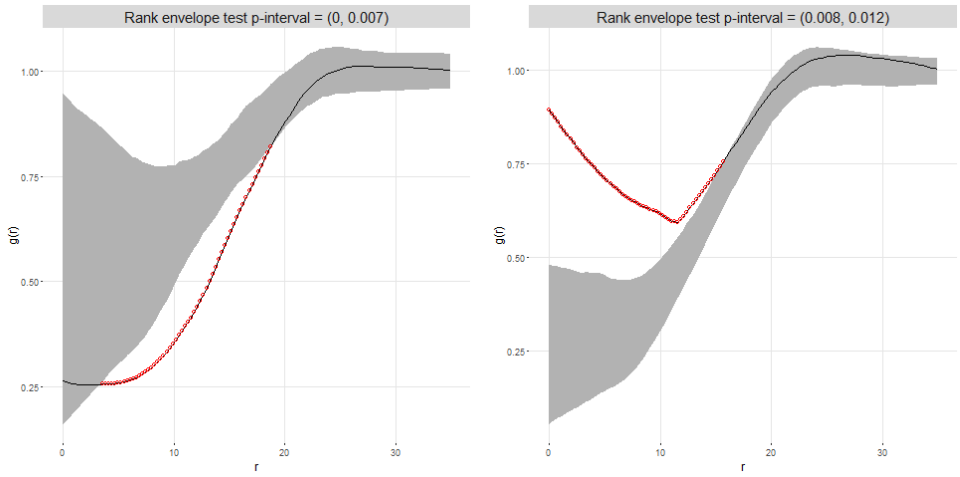

Figure 2.18: Left (respectively right): 95%-global rank envelopes based on the fitted Strauss model for the apex (respectively nucleoli) point pattern together with the estimated pair correlation for the nucleoli (respectively apex) point pattern.

described by the same model, the pair correlation function of one point pattern is compared to the envelopes of the model based on the other point pattern (see figure 2.18). We only consider the pair correlation function here, but both the K- and J-functions result in the same conclusion. As both p-intervals has $p_+ < 0.05$ there is evidence against the hypothesis that the nucleolus and apex point patterns can be described by the same model.

2.5.3 Fitting a Strauss model

The pair correlation function for the apex point pattern indicates that the underlying process may not be a hardcore process. This is also consistent with the small hardcore estimate. In this section Strauss models are fitted to both the

Data	$\boldsymbol{\beta_0}$	$\boldsymbol{\gamma_0}$	$\boldsymbol{R_0}$	$\hat{\boldsymbol{\beta}}$	$\hat{\gamma}$	$\hat{\boldsymbol{R}}$	**LPPL**
nucleolus	$2.16 * 10^{-5}$	0.5	10	$2.25 * 10^{-5}$	0.20	10.56	-24067
nucleolus	$2.16 * 10^{-5}$	0.5	20	$2.54 * 10^{-5}$	0.69	20.27	-24090
nucleolus	$2.16 * 10^{-5}$	1	10	$2.44 * 10^{-5}$	0.21	11.26	-24064
nucleolus	$2.16 * 10^{-5}$	0.5	30	$2.55 * 10^{-5}$	0.94	31.78	-24169
apex	$2.14 * 10^{-5}$	0.5	10	$2.42 * 10^{-5}$	0.51	15.62	-24305
apex	$2.14 * 10^{-5}$	0.5	20	$2.68 * 10^{-5}$	0.74	20.80	-24367
apex	$2.14 * 10^{-5}$	1	10	$2.27 * 10^{-5}$	0.50	10.89	-24369
apex	$2.14 * 10^{-5}$	0.5	30	$4.65 * 10^{-5}$	0.76	30.51	-24546

Table 2.2: Results for each optimisation with start values (β_0, γ_0, R_0). LPPL is the value of the maximised log profile pseudolikelihood for the Strauss model.

nucleolus and apex point patterns in order to investigate whether a hardcore is necessary. The parameter estimation is done in a similar fashion to what we did for the Strauss hardcore model, and in accordance with section 2.3.2. Results for the optimisation of the (approximate) log profile pseudolikelihood is shown in table 2.2. Note that we used the same for four combinations of start values as earlier. For the nucleolus point pattern the estimates

$$\hat{\beta}_{nucl} = 2.44 * 10^{-5}, \quad \hat{\gamma}_{nucl} = 0.21, \quad \hat{R}_{nucl} = 11.26$$

optimise the pseudolikelihood, and for the apex point pattern we get

$$\hat{\beta}_{apex} = 2.42 * 10^{-5}, \quad \hat{\gamma}_{apex} = 0.51, \quad \hat{R}_{apex} = 15.62.$$

Model check

Simulations used for the envelopes under the fitted Strauss models are constructed in a similar manner as for the Strauss hardcore model. Figure 2.19 and 2.20 display global rank envelopes for the pair correlation function and for the K- and J-functions along with the associated p-intervals. For the nucleolus point pattern, each rank envelope test rejects the hypothesis of a Strauss model, since $p_+ < 0.05$, indicating that the hardcore is necessary. However, for the apex point pattern, no evidence is found against the Strauss model, since $p_- > 0.05$, confirming our suspicion that the apexes do not come from a hardcore process. This is consistent with the fact that an apex location in itself does not correspond to a biological point, but only can be viewed relative to the nucleolus location as an indicator of the pyramidal cell's orientation.

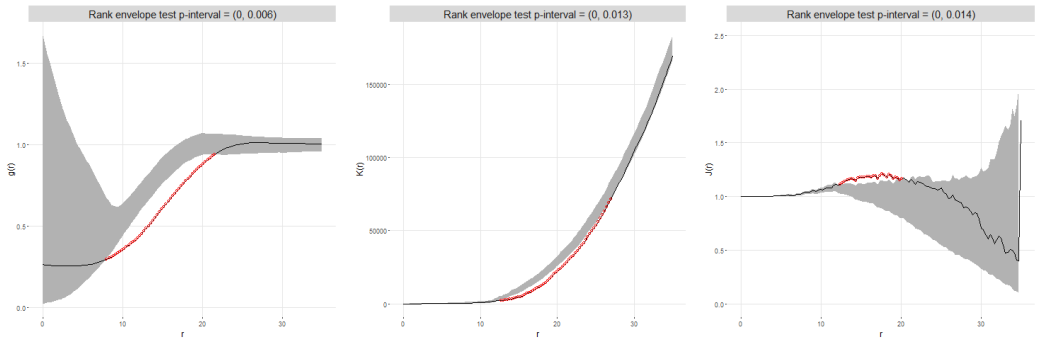

Figure 2.19: 95%-global rank envelopes under fitted Strauss model together with the estimate for the nucleolus point pattern of the pair correlation function (left), K-function (middle) and J-function (right) for $0 < r \leq 35$.

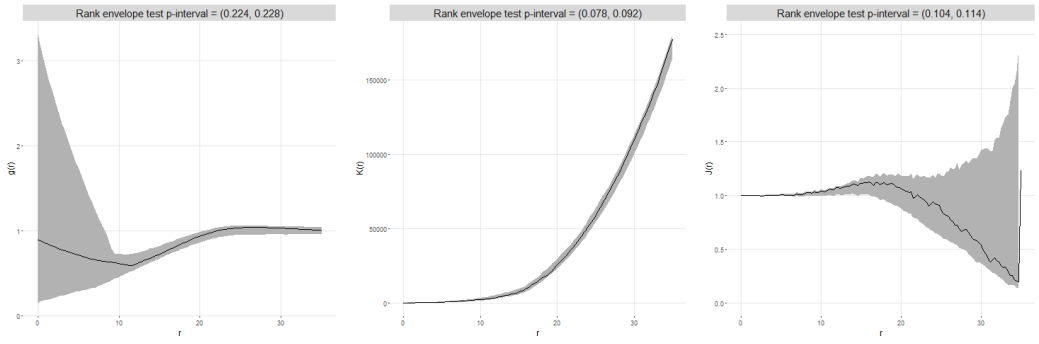

Figure 2.20: 95%-global rank envelopes under fitted Strauss model together with the estimate for the apex point pattern of the pair correlation function (left), K-function (middle) and J-function (right) for $0 < r \leq 35$.

Chapter 3

Marked point processes

In this chapter we describe theory involving marked point processes and analyse the pyramidal cell data, taking into account the cells' orientation. Most literature focus on real or discrete marks, when handling marked point processes, and often assumptions of stationarity and isotropy are made for the sake of simplicity. For our application these are unjust restrictions and more general theory is needed. We will therefore extend some of the theory, described in the literature, to cover a wider range of applications; especially applications with multidimensional marks.

The theory presented in the following includes an introduction to marked point processes in section 3.1 and a description of the most common marked point process models in section 3.2. These models are independent marking, the marked Poisson process, the random field model and constructed marks. In section 3.3 moment measures for marked point processes are defined and in turn used for defining the mark and two-point mark distributions. We then extend the concept of Palm distributions from section 2.2 to cover marked point processes. The theory of Palm distributions for marked point processes is vaguely described by Illian et al. (2008) in the stationary case, while the treatment by Daley and Vere-Jones (2008) is more general and mathematically precise but also cumbersome. We aim in between to give a simple but precise presentation.

In section 3.5 and 3.6 we present the mark correlation function and the mark-weighted $\mathcal{K}$-measure, which are summary statistics for marked point processes, and their estimates. Here we also propose the modified cylindrical K-function, that is specifically used to detect columnar structures in marked point patterns with directional marks. Lastly, we analyse the pyramidal cell data and discuss possible extensions of the analysis.

3.1 Basic definitions

Point patterns emerge in a wide variety of applications for some of which additional information associated with each point is available. This information can be handled simultaneously with the point pattern by using the theory of marked point processes. Let X be a point process on $S \subseteq \mathbb{R}^d$ with $d \geq 1$, and, for some space M, let $m_\xi \in M$ be a stochastic variable that describes the additional information of the point $\xi \in X$. Then m_ξ is said to be the *mark* attached to ξ and $\Phi = \{(\xi, m_\xi) : \xi \in X\}$, defined on $S \times M$, is called a *marked point process* with points in S and *mark space* M. We call X the *ground process* of Φ. Marked point processes can more generally be defined when S and M are complete separable metric spaces, but this is beyond the scope of this thesis (for details see e.g. Daley and Vere-Jones, 2002). The marks are often either continuous or discrete variables on $\mathbb{R}^k$, $k \geq 1$. Specifically, for our application marks and point are continuous variables on $\mathbb{S}^2$ and $\mathbb{R}^3$, respectively.

Formally, S and M are both equipped with a Borel σ-algebra, but as mentioned in section 2.1 the use of measure theoretical wording will be kept to a minimum, and thus we shall not introduce these σ-algebras.

The pyramidal cell data can be viewed as a marked point pattern, where the points describe the nucleolus locations, while the marks describe the directional (unit) vectors between nucleolus and apex. Hence, these data form a marked point pattern on $\mathbb{R}^3 \times \mathbb{S}^2$. Other examples of marked point processes are the location of different tree sorts, where the marks describe the species (e.g. maple, pine, oak, ect.), hinting that the mark space need not be Euclidean. Discrete cases like this are however generally coded with numbers and are referred to as multitype point processes.

The definitions of stationarity and isotropy for marked point processes are analogous to those of unmarked point processes. However, the marks need not belong to the same space as the points. *Translation* or *rotation* of a marked point process is therefore defined as an operation on the points only and leave the marks unchanged. A marked point process is then said to be *stationary* (resp. *isotropic*) if its distribution is invariant under translation (resp. rotation).

Throughout this thesis we consider a marked point process, $\Phi = \{(\xi, m_\xi) : \xi \in X\}$, on $S \times M$, where the ground process, X, constitutes a locally finite point process with a locally finite intensity measure, μ_p. We denote by N_{lf} the space of locally finite marked point configurations on $S \times M$, i.e. N_{lf} consists of marked point configurations in which the points constitute a locally finite point configuration on S.

3.2 Models

Depending on the nature of the marks, different models are used to analyse marked point patterns. In our application the marks are directional vectors, and thus multitype point processes will not be useful models. A general point of interest, regardless of the marks' nature, is whether the marks and points are independent. If this is the case we can separately build a model for our points and marks, otherwise interactions have to be accounted for.

3.2.1 Independent and conditional independent marking

Daley and Vere-Jones (2002) define independent marking as a model, where the marks are assumed to be pairwise independent given the ground process X such that the conditional distribution of a mark m_ξ given the ground process does not depend on $X \backslash \xi$. Hence this model describes a conditional independence and not a pure independence as implied by its name. We will refer to this model as *conditional independent marking*. Illian et al. (2008) give a different and less precise definition of independent marking, where the marks at least are assumed to be iid. (independent and identically distributed). We will define *independent marking* as a model with iid. marks independent of the ground process. Note that conditional independent marking is a broader class of models containing the independent marking. Due to the two models' independent nature, they are natural null models when analysing marked point patterns.

Simulation under conditional independent marking can be done hierarchically: First, simulate points from some appropriate model. Then, given the points, independently simulate each mark from a distribution that may depend on its corresponding point. Simulation under independent marking is done similarly, but the marks are simulated from the same distribution that is independent of the points.

For the pyramidal cell data, the conditional independent marking or independent marking model seem appropriate if the direction of a cell does not depend on the other cells and their direction. There is however physical restrictions, as the cells cannot overlap, which may affect the orientations, introducing some interaction between marks and other marks/points.

3.2.2 Marked Poisson process

The marked Poisson process with conditional independent marks is another model that can be used for marks of various types.

Definition 3.2.1. Let $X \sim \text{Poisson}(S, \rho)$ and assume that the marks $\{m_\xi : \xi \in X\}$ are mutually independent conditional on X. Then $\Phi = \{(\xi, m_\xi) : \xi \in X\}$ is called a *marked Poisson process with conditional independent marks.*

If the distribution of a mark in the above definition depends only on its corresponding point, then the resulting marked point process is a case of conditional independent marking. As stated by the following theorem a marked Poisson process and its marks are Poisson point processes under certain conditions.

Theorem 3.2.2. Let Φ be a marked Poisson process satisfying conditional independent marking with marks in $M \subseteq \mathbb{R}^k$, $k \geq 1$, and assume that $m_\xi|\xi$ has density p_ξ wrt. the Lebesgue measure on $\mathbb{R}^k$. Then

1. $\Phi \sim \text{Poisson}(S \times M, \lambda)$, where $\lambda(\xi, m) = \rho(\xi)p_\xi(m)$.

2. if the function κ on M, defined by $\kappa(m) = \int_S \lambda(\xi, m)d\xi$, is locally integrable, then the point process $\{m_\xi : \xi \in S\} \sim \text{Poisson}(M, \kappa)$.

Proof. Let $B \subseteq S \times M$ be any bounded Borel set and let $\mu_p(A) = \int_A \rho(\xi)d\xi$, $A \subseteq S$ Borel, be the intensity measure for X. For $\xi \in S$, define $B_M^\xi = \{m \in M : (\xi, m) \in B\}$ and $B_S = \{\xi \in S : \exists m \in M \text{ such that } (\xi, m) \in B\}$. Note that B_S is a bounded Borel subset of S. The void probability for Φ in B can then be written as

$$P(\Phi \cap B = \emptyset) = \text{E}\left[P(\Phi \cap B = \emptyset | X_{B_S})\right]$$

$$= \sum_{n=0}^{\infty} \frac{\exp(-\mu_p(B_S))}{n!} \int_{B_S} \cdots \int_{B_S} P(m_i \notin B_M^{\xi_i}, i = 1, \ldots, n | \{\xi_1, \ldots, \xi_n\}) \prod_{i=1}^{n} \rho(\xi_i)d\xi_1 \ldots d\xi_n$$

$$= \sum_{n=0}^{\infty} \frac{\exp(-\mu_p(B_S))}{n!} \int_{B_S} \cdots \int_{B_S} \prod_{i=1}^{n} \left(1 - \int_{B_M^{\xi_i}} p_{\xi_i}(m)dm\right) \rho(\xi_i)d\xi_1 \ldots d\xi_n$$

$$= \exp(-\mu_p(B_S)) \sum_{n=0}^{\infty} \frac{1}{n!} \left(\int_{B_S} \left(1 - \int_{B_M^\xi} p_\xi(m)dm\right) \rho(\xi)d\xi\right)^n$$

$$= \exp(-\mu_p(B_S)) \exp\left(\int_{B_S} \left(1 - \int_{B_M^\xi} p_\xi(m)dm\right) \rho(\xi)d\xi\right)$$

$$= \exp\left(-\int_{B_S} \int_{B_M^\xi} \rho(\xi)p_\xi(m)dmd\xi\right)$$

$$= \exp\left(-\int_B \lambda(\xi, m)d(\xi, m)\right),$$

where the first equality follows from the law of iterated expectation and the second equality from equation (2.7). Since this is the void probability under Poisson$(S \times M, \lambda)$, first part of the proposition is proved.

Let now $B \subseteq M$ be a bounded Borel set. Then it follows from part 1 in the proposition that

$$P\left(\{m_\xi : (\xi, m_\xi) \in \Phi\} \cap B = \emptyset\right) = P(\Phi \cap (S \times B) = \emptyset)$$
$$= \exp\left(-\int_B \int_S \lambda(x, m) dx dm\right),$$

which is the void probability under Poisson(M, κ) with $\kappa(m) = \int_S \lambda(\xi, m) d\xi$. □

The marked Poisson process does not seem relevant in our analysis of the pyramidal cell data, as we in section 2.5 found that the Poisson process was not a suitable model for the nucleolus locations.

3.2.3 Random field model

In order to define random field models we first introduce the concept of random fields.

Briefly described, a random field $\{Z(\xi) : \xi \in I\}$, also denoted $\{Z(\xi)\}$, is a generalisation of a stochastic process, where the index set I no longer is required to be real, but can e.g. be multidimensional. For the random field models, we consider $I = S$. That is, the random field is a collection of random variables indexed by elements of S. This enables us to define *random field models* that based on a point process allocate marks by a random field $\{Z(\xi)\}$ that is independent of the points. More precisely, the mark attached to a point ξ is generated by $m_\xi = Z(\xi)$. Due to the construction of this model, two points in close vicinity typically have similar marks.

In relation to the pyramidal cell data this means that nucleoli that are close to each other are prone to have similar orientations, which may seem reasonable. On the other hand the random field model does not allow for interaction between a cell's direction and nearby nucleoli An interaction that could be relevant as the cells cannot overlap.

3.2.4 Constructed marks

Constructed marks reflect some characteristic of the neighbourhood of their related point, e.g. the local intensity around the point. Examples of such constructed marks are the points' distance to their nearest neighbour or the number of further points within a certain distance of the mark's point. Constructed marks can in many instances give an indication of a suitable marked

point process model for an observed marked point pattern. Consider for example a point pattern with real marks. If the observed marks tend to stimulate (resp. weaken) each other when their points are close, the correlation between the observed marks and constructed marks, increasing with the local intensities, will be positive (resp. negative).

We will use constructed marks as a tool to investigate whether the pyramidal cells' orientation is affected by nearby nucleoli, but we will not pursue it as a model.

3.3 Moment measures

To describe the first- as well as second-order properties of a marked point process more summary characteristics, than those presented for unmarked point processes, are necessary, since the marks too need consideration.

3.3.1 First-order properties

The first-order properties of marked point processes are their intensity measure and mark distribution. The *intensity measure* for a marked point proces Φ is defined by

$$\mu(A \times B) = \mathrm{E}\left[\sum_{(\xi, m_\xi)\in\Phi} \mathbb{I}(\xi \in A, m_\xi \in B)\right] \tag{3.1}$$

for Borel sets $A \subseteq S$ and $B \subseteq M$. Note that $\mu(\cdot \times M) = \mu_p(\cdot)$ and, for any Borel set $B \subseteq M$, $\mu(\cdot \times B) \leq \mu_p(\cdot)$. Then, for a fixed Borel set $B \subseteq M$, $\mu(\cdot \times B)$ is easily seen to be a measure on S that is absolutely continuous wrt. μ_p. Therefore, there exists a Radon-Nikodym derivative, $Q_\xi(B)$, such that

$$\mu(A \times B) = \int_A Q_\xi(B) d\mu_p(\xi), \tag{3.2}$$

for all Borel sets $A \subseteq S$ and $B \subseteq M$. Note that $Q_\xi(B)$ is uniquely determined μ_p-almost every ξ. If the intensity function ρ exists we can equivalently say that $Q_\xi(B)$ is unique Lebesgue-a.e. ξ for $\rho(\xi) > 0$. Furthermore, if Q_ξ has a density, q_ξ, wrt. the Lebesgue measure then

$$\mu(A \times B) = \int_A \int_B q_\xi(m)\rho(\xi) dm d\xi$$

for all Borel sets $A \subseteq S$ and $B \subseteq M$. We can choose $Q_\xi(B)$ such that it is a Markov kernel: Per definition, $Q_\xi(B)$ is a Radon Nikodym derivative and thus a measurable mapping wrt. ξ for fixed $B \subseteq M$ Borel. Furthermore, Q_ξ is a measure μ_p-a.e. ξ, since equation (3.2) yields that, for $B = \emptyset$,

$$\int_A Q_\xi(\emptyset) d\mu_p(\xi) = \mu(A \times \emptyset) = 0 = \int_A 0 d\mu_p(\xi),$$

and, for disjoint $B_i \subseteq M$, $i = 1, 2, \ldots$,

$$\begin{aligned}\int_A Q_\xi\left(\overset{\infty}{\underset{i=1}{\cup}} B_i\right) d\mu_p(\xi) &= \mathrm{E}\left[\sum_{(\xi,m_\xi)\in\Phi} \mathbb{I}\left(\xi \in A, m_\xi \in \overset{\infty}{\underset{i=1}{\cup}} B_i\right)\right] \\ &= \mathrm{E}\left[\sum_{(\xi,m_\xi)\in\Phi} \sum_{i=1}^{\infty} \mathbb{I}\left(\xi \in A, m_\xi \in B_i\right)\right] \\ &= \int_A \sum_{i=1}^{\infty} Q_\xi(B_i) d\mu_p(\xi),\end{aligned}$$

for every Borel set $A \subseteq S$. In addition, since $\mu(A \times M) = \mu_p(A)$ for any Borel set $A \subseteq S$, it follows from (3.2) that $Q_\xi(M) = 1$ μ_p-a.e. ξ. Thus Q_ξ is a probability measure for μ_p-a.e. ξ. Lastly, since $Q_\xi(B)$ only is unique μ_p-a.e. ξ, we can choose Q_ξ to be any probability measure on M whenever ξ lies in a μ_p-null set. All in all, this proves that $Q_\xi(B)$ can be chosen as a Markov kernel.

Definition 3.3.1. The probability measure Q_ξ is called the *mark distribution at* ξ.

If $Q_\xi = Q$ does not depend on $\xi \in S$, Q is simply called the *mark distribution*. In that case equation (3.2) reduces to $\mu(A \times B) = Q(B)\mu_p(A)$ for all Borel sets $A \subseteq S$ and $B \subseteq M$, and Q may be interpreted as the distribution of a typical mark. In general, if the intensity function, ρ, exists then

$$\mu(A \times B) = \int_A \frac{1}{\rho(\xi)|A|} \mu(A \times B) d\mu_p(\xi)$$

for all Borel sets $B \subseteq M$ and $A \subseteq S$ satisfying $0 < |A| < \infty$. Hence, we can take

$$Q_\xi(B) = \frac{\mu(A \times B)}{\rho(\xi)|A|} \tag{3.3}$$

for all $\xi \in S$, entailing that $Q_\xi(B)$ can be interpreted as the conditional probability of $m_\xi \in B$ given ξ. In case of stationarity equation (3.3) implies that $Q_\xi = Q$ for all $\xi \in S$.

By combining equation (3.1) and (3.2) with the standard proof, the so-called *Campbell theorem* is immediately obtained,

$$\mathrm{E}\left[\sum_{(\xi,m_\xi)\in\Phi} h(\xi, m_\xi)\right] = \int \int h(\xi, m) dQ_\xi(m) d\mu_p(\xi),$$

for any function $h : S \times M \to [0, \infty)$.

Theorem 3.3.2. For a marked point process, where the distribution of each mark m_ξ, conditional on the ground process X, only depends on ξ, we can take $Q_\xi(B) = P(m_\xi \in B|\xi)$ for all $\xi \in S$ and any Borel set $B \subseteq M$.

Proof. From equation (3.1) and (3.2) and the law of iterated expectation we get that

$$\begin{aligned}\int_A Q_\xi(B)d\mu_p(\xi) &= \mathrm{E}\left[\sum_{\xi\in X}\mathbb{I}(\xi \in A)\mathrm{E}\left[\mathbb{I}(m_\xi \in B)|X\right]\right] \\ &= \mathrm{E}\left[\sum_{\xi\in X}\mathbb{I}(\xi \in A)P(m_\xi \in B|\xi)\right] \\ &= \int_A P(m_\xi \in B|\xi)d\mu_p(\xi)\end{aligned}$$

for all Borel sets $A \subseteq S$ and $B \subseteq M$. Note that the last equality follows from equation (2.10) and the fact that $\mathbb{I}(\xi \in A)P(m_\xi \in B|\xi)$ is a function of ξ only. Thus, we can choose $Q_\xi(B) = P(m_\xi \in B|\xi)$ for all $\xi \in S$. □

Note that the above theorem holds under the assumption of conditional independent marking or independent marking. In the particular case of independent marking, or simply the case where the marks are identically distributed and independent of the ground process, $Q_\xi(B) = P(m \in B)$. That is, $Q_\xi = Q$ is the distribution of a single mark.

3.3.2 Second-order properties

In this section we describe the second-order factorial moment measure and the two-point mark distribution. Furthermore, a weighted analogue to the second-order factorial moment measure is introduced.

Definition 3.3.3. For Borel sets $A_1, A_2 \subseteq S$ and $B_1, B_2 \subseteq M$, *the second-order factorial moment measure* is defined as

$$\begin{aligned}&\alpha^{(2)}(A_1 \times B_1 \times A_2 \times B_2) \\ &\quad = \mathrm{E}\left[\sum_{(\xi,m_\xi),(\eta,m_\eta)\in\Phi}^{\neq} \mathbb{I}\left(\xi \in A_1, m_\xi \in B_1, \eta \in A_2, m_\eta \in B_2\right)\right], \qquad (3.4)\end{aligned}$$

where $\sum^{\neq}$ is the sum over all pairs of different marked points, i.e. $\xi \neq \eta$.

For fixed $B_1, B_2 \subseteq M$ Borel, $\alpha^{(2)}(\cdot \times B_1 \times \cdot \times B_2)$ is a measure on $S \times S$. Note also that $\alpha^{(2)}(\cdot \times M \times \cdot \times M) = \alpha_p^{(2)}(\cdot \times \cdot)$ is the second-order factorial moment

measure of the ground process. Moreover, $\alpha^{(2)}(\cdot \times B_1 \times \cdot \times B_2) \leq \alpha_p^{(2)}(\cdot \times \cdot)$ for all Borel sets $B_1, B_2 \subseteq M$, i.e. $\alpha^{(2)}(\cdot \times B_1 \times \cdot \times B_2) \ll \alpha_p^{(2)}(\cdot \times \cdot)$. Hence, there exists a Radon-Nikodym derivative, $Q_{\xi,\eta}(B_1 \times B_2)$, such that

$$\alpha^{(2)}(A_1 \times B_1 \times A_2 \times B_2) = \int_{A_1 \times A_2} Q_{\xi,\eta}\left(B_1 \times B_2\right) d\alpha_p^{(2)}(\xi, \eta). \tag{3.5}$$

It can be shown, analogously to the proof concerning the mark distribution Q_ξ, that $Q_{\xi,\eta}$ can be chosen as a Markov kernel.

Definition 3.3.4. The probability measure $Q_{\xi,\eta}$ is called *the two-point mark distribution at ξ and η.*

If the second-order intensity function exists, i.e. $d\alpha_p^{(2)}(\xi, \eta) = \rho^{(2)}(\xi, \eta) d\xi d\eta$, we have that

$$\alpha^{(2)}(A_1 \times B_1 \times A_2 \times B_2) = \int_{A_1 \times A_2} \frac{\alpha^{(2)}(A_1 \times B_1 \times A_2 \times B_2)}{\rho^{(2)}(\xi, \eta)|A_1||A_2|} d\alpha_p^{(2)}(\xi, \eta)$$

for all Borel sets $B_1, B_2 \subseteq M$ and $A_1, A_2 \subseteq S$ satisfying $0 < |A_1|, |A_2| < \infty$. Therefore, we can take

$$Q_{\xi,\eta}\left(B_1 \times B_2\right) = \frac{\alpha^{(2)}(A_1 \times B_1 \times A_2 \times B_2)}{\rho^{(2)}(\xi, \eta)|A_1||A_2|} \tag{3.6}$$

for all $\xi, \eta \in S$. Hence, $Q_{\xi,\eta}$ may be interpreted as the joint, conditional probability of $m_\xi \in B_1$ and $m_\eta \in B_2$ given ξ and η. For a stationary marked point process, equation (3.6) implies that (with a slight abuse of notation) $Q_{\xi,\eta} = Q_{\eta-\xi}$ for all $\xi, \eta \in S$. If the marked point process additionally is isotropic, then (again abusing notation) $Q_{\xi,\eta} = Q_{\|\xi-\eta\|}$ for all $\xi, \eta \in S$.

Using equation (3.4) and (3.5) as well as the standard proof it follows that

$$\mathrm{E}\left[\sum_{(\xi,m_\xi),(\eta,m_\eta)\in\Phi}^{\neq} h\left(\xi, m_\xi, \eta, m_\eta\right)\right] = \int\int h\left(\xi, m, \eta, l\right) dQ_{\xi,\eta}\left(m, l\right) d\alpha_p^{(2)}(\xi, \eta) \tag{3.7}$$

for any function $h : S \times M \times S \times M \to [0, \infty)$. This result is referred to as the *two-point Campbell theorem.*

Theorem 3.3.5. For a marked point process, where the joint distribution of each pair of marks m_ξ and m_η, conditional on the ground process X, only depends on ξ and η, we can take $Q_{\xi,\eta} = P(m_\xi \in B_1, m_\eta \in B_2 | \xi, \eta)$ for all $\xi, \eta \in S$ and any Borel sets $B_1, B_2 \subseteq M$.

Proof. This prove is analogous to that of theorem 3.3.2. Using equation (2.11) and the law of iterated expectation on the definition of $Q_{\xi,\eta}$ yields that

$$\begin{aligned}
&\int_{A_1\times A_2} Q_{\xi,\eta}\left(B_1\times B_2\right) d\alpha_p^{(2)}(\xi,\eta) \\
&= \mathrm{E}\left[\sum_{\xi,\eta\in X}^{\neq} \mathbb{I}\left(\xi\in A_1,\eta\in A_2\right)\mathrm{E}\left[\mathbb{I}\left(m_\xi\in B_1, m_\eta\in B_2\right)|X\right]\right] \\
&= \int_{A_1\times A_2} P(m_\xi\in B_1, m_\eta\in B_2|\xi,\eta) d\alpha_p^{(2)}(\xi,\eta)
\end{aligned}$$

for all Borel sets $A_1, A_2 \subseteq S$ and $B_1, B_2 \subseteq M$. That is, we can choose $Q_{\xi,\eta} = P(m_\xi \in B_1, m_\eta \in B_2|\xi,\eta)$ for all $\xi,\eta \in S$ and any Borel sets $B_1, B_2 \subseteq M$. □

Note that the above theorem holds in the case of conditional independent marking, under which,

$$\begin{aligned}
Q_{\xi,\eta}(B_1\times B_2) &= P(m_\xi\in B_1, m_\eta\in B_2|\xi,\eta) = P(m_\xi\in B_1|\xi)P(m_\eta\in B_2|\eta) \\
&= Q_\xi(B_1)Q_\eta(B_2)
\end{aligned}$$

for all $\xi,\eta \in S$ and Borel sets $B_1, B_2 \subseteq M$, where the last equality follows from theorem 3.3.2. Furthermore, for a marked point process with identically distributed marks independent of the ground process, theorem 3.3.5 yields that $Q_{\xi,\eta}(B_1\times B_2) = P(m\in B_1, l\in B_2)$ for all $\xi,\eta\in S$, all Borel sets $B_1, B_2 \subseteq M$ and marks m and l. Hence, $Q_{\xi,\eta}$ does not depend on ξ or η and is the joint distribution of any two marks. If the marks in addition are mutually independent, i.e. under independent marking, $Q_{\xi,\eta}(B_1\times B_2) = P(m\in B_1)P(m\in B_2) = Q(B_1)Q(B_2)$ for all $\xi,\eta\in S$ and Borel sets $B_1, B_2 \subseteq M$.

Analogously to $\alpha^{(2)}$, we define the measure $\alpha_t^{(2)}$

Definition 3.3.6. For Borel sets $A_1, A_2 \subseteq S$ and any function $t : M\times M \to [0,\infty)$, we define the measure

$$\alpha_t^{(2)}(A_1\times A_2) = \mathrm{E}\left[\sum_{(\xi,m_\xi),(\eta,m_\eta)\in\Phi}^{\neq} \mathbb{I}\left(\xi\in A_1,\eta\in A_2\right) t(m_\xi,m_\eta)\right].$$

If there exists a non-negative function $\rho_t^{(2)}$ such that, for all Borel sets $A_1, A_2 \subseteq S$,

$$\alpha_t^{(2)}(A_1\times A_2) = \int_{A_1}\int_{A_2} \rho_t^{(2)}(\xi,\eta)d\xi d\eta,$$

then $\rho_t^{(2)}$ is called the *second-order t-intensity function*.

Note that $\alpha_t^{(2)}(A_1 \times A_2) = \alpha^{(2)}(A_1 \times B_1 \times A_2 \times B_2)$ for $t(m_1, m_2) = \mathbb{I}(m_1 \in B_1, m_2 \in B_2)$, in which instance we shall denote $\rho_t^{(2)}$ by $\rho_{B_1,B_2}^{(2)}$. Per definition of $\alpha_t^{(2)}$ and $\alpha^{(2)}$ it follows from the standard proof that

$$\begin{aligned} \int h(\xi,\eta) d\alpha_t^{(2)}(\xi,\eta) &= \mathrm{E}\left[\sum_{(\xi,m_\xi),(\eta,m_\eta)\in\Phi}^{\neq} h(\xi,\eta) t(m_\xi, m_\eta) \right] \\ &= \int h(\xi,\eta) t(m_\xi, m_\eta) d\alpha^{(2)}(\xi, m_\xi, \eta, m_\eta) \end{aligned} \tag{3.8}$$

for any function $h : S \times S \to [0, \infty)$. Provided that the second-order intensity and the second-order t-intensity functions exists it follows as a consequence of equation (3.8) that

$$\int_{A_1}\int_{A_2} \rho_t^{(2)}(\xi,\eta) d\xi d\eta = \int_{A_1}\int_{A_2}\int t(m_1, m_2)\rho^{(2)}(\xi,\eta) dQ_{\xi,\eta}(m_1, m_2) d\xi d\eta$$

for all Borel sets $A_1, A_2 \subseteq S$. Therefore,

$$\frac{\rho_t^{(2)}(\xi,\eta)}{\rho^{(2)}(\xi,\eta)} = \int t(m_1, m_2) dQ_{\xi,\eta}(m_1, m_2) \tag{3.9}$$

for $\rho^{(2)}(\xi,\eta) > 0$ and Lebesgue a.e. ξ, η, which in turn implies that

$$\rho_{B_1,B_2}^{(2)}(\xi,\eta) = \rho^{(2)}(\xi,\eta) Q_{\xi,\eta}(B_1 \times B_2)$$

for all Borel sets $B_1, B_2 \subseteq M$ Lebesgue a.e. ξ, η. This implies that $\rho_{M,M}^{(2)}(\xi,\eta) = \rho^{(2)}(\xi,\eta)$ for Lebesgue-a.e. ξ, η.

3.4 Palm theory

In this section we introduce the theory of reduced Palm distributions for marked point processes. Most of the concepts are similar to those presented for unmarked point processes in section 2.2. Therefore, some of the results and their proofs immediately extend from the unmarked to the marked case and will not be discussed in details.

Definition 3.4.1. The *reduced Campbell measure* of a marked point process Φ on $S \times M$ is defined by

$$C^!(D \times F) = \mathrm{E}\left[\sum_{(\xi,m_\xi)\in\Phi} \mathbb{I}((\xi, m_\xi) \in D, \Phi\backslash(\xi, m_\xi) \in F) \right], \tag{3.10}$$

for $D \subseteq S \times M$ Borel and $F \subseteq N_{\mathrm{lf}}$.

Note that $C^!(\cdot \times N_{\text{lf}}) = \mu(\cdot)$. Performing a similar argument as in section 2.2 for the relation between α_p and $C_p^!$, we can relate $\alpha^{(2)}$ and $C^!$ by

$$\alpha^{(2)}(A_1 \times B_1 \times A_2 \times B_2) = \int \mathbb{I}(\xi \in A_1, m \in B_1) \sum_{(\eta, m_\eta) \in \phi} \mathbb{I}(\eta \in A_2, m_\eta \in B_2) dC^!(\xi, m, \phi)$$

for all Borel sets $A_1, A_2 \subseteq S$ and $B_1, B_2 \subseteq M$.

For fixed $F \subseteq N_{\text{lf}}$, $C^!(\cdot \times F)$ is clearly absolutely continuous wrt. the intensity measure μ. Hence, there exists a Radon-Nikodym derivative $P^!_{(\xi,m)}(F)$ such that

$$C^!(D \times F) = \int_D P^!_{(\xi,m)}(F) d\mu(\xi, m) \tag{3.11}$$

for all Borel sets $D \subseteq S \times M$ and all $F \subseteq N_{\text{lf}}$. Note that $P^!_{(\xi,m)}(F)$ is uniquely determined μ-a.e. (ξ, m). If q_ξ and ρ exist, we can equivalently say, due to equation (3.2), that $P^!_{(\xi,m)}(F)$ is unique Lebesgue a.e. ξ, m for $q_\xi(m)\rho(\xi) > 0$. By carrying out an argument similar to that for $P^!_\xi(F)$ in section 2.2, we can show that $P^!_{(\xi,m)}(F)$ can be chosen as a Markov kernel.

Definition 3.4.2. The probability measure $P^!_{(\xi,m)}$ is called the *reduced Palm distribution at point ξ with mark m.*

We can consider a non-reduced version of the Palm distribution by specifying the Campbell measure with $\Phi \backslash (\xi, m_\xi)$ substituted by Φ in (3.10). However, the reduced Palm distribution is the one of interest and the results can easily be modified to the non-reduced Palm distribution, and hence will not be covered.

The marked reduced Palm distribution, $P^!_{(\xi,m)}$, also has an interpretation as a conditional distribution. Assume that Φ has continuous marks, and let $A = b(\xi, \varepsilon_\xi)$ and $B = b(m, \varepsilon_m)$ denote balls with sufficiently small radii $\varepsilon_\xi, \varepsilon_m > 0$, where $\xi \in S$ and $m \in M$. Heuristically it is unlikely that $\sum_{(\xi, m_\xi) \in \Phi} \mathbb{I}(\xi \in A, m_\xi \in B) > 1$, i.e.

$$\mu(A \times B) \approx P\left(\sum_{(\eta, m_\eta) \in \Phi} \mathbb{I}(\eta \in A, m_\eta \in B) > 0 \right)$$

and

$$C^!(A \times B \times F) \approx P\left(\sum_{(\eta, m_\eta) \in \Phi} \mathbb{I}(\eta \in A, m_\eta \in B) > 0, \Phi \backslash (\xi, m) \in F \right)$$

for all $F \subseteq N_{\text{lf}}$. Furthermore, since A and B are sufficiently small balls, $C^!(A \times B \times F) \approx P^!_{(\xi,m)}(F)\mu(A \times B)$, and thus

$$P^!_{(\xi,m)}(F) \approx P\left(\Phi \backslash (\xi, m) \in F \,\middle|\, \sum_{(\eta, m_\eta) \in \Phi} \mathbb{I}(\eta \in b(\xi, \varepsilon_\xi), m_\eta \in b(m, \varepsilon_m)) > 0 \right)$$

for all $F \subseteq N_{\text{lf}}$. That is, the reduced Palm distribution $P^!_{(\xi,m)}$ can be interpreted as the conditional distribution of $\Phi\backslash(\xi, m)$ given that $(\xi, m) \in \Phi$.

It follows from equation (3.2) and (3.11) that

$$\begin{aligned} C^!(D \times F) &= \int \mathbb{I}((\xi, m) \in D) P^!_{(\xi,m)}(F) d\mu(\xi, m) \\ &= \int\int\int \mathbb{I}((\xi, m) \in D, \phi \in F) dP^!_{(\xi,m)}(\phi) dQ_\xi(m) d\mu_p(\xi) \end{aligned} \tag{3.12}$$

for all Borel sets $D \subseteq S \times M$ and all $F \subseteq N_{\text{lf}}$. Applying the standard proof we get the *Campbell-Mecke theorem* for marked point processes, stating that

$$\mathrm{E}\left[\sum_{(\xi,m_\xi)\in\Phi} h(\xi, m_\xi, \Phi\backslash(\xi, m_\xi))\right] = \int\int\int h(\xi, m, \phi) dP^!_{(\xi,m)}(\phi) dQ_\xi(m) d\mu_p(\xi) \tag{3.13}$$

for any function $h : S \times M \times N_{\text{lf}} \to [0, \infty)$.

The following theorem give a more concrete understanding of the reduced Palm distribution for marked point process under certain independence assumption. That is, under independent marking, the marked reduced Palm distribution marginalises into a ground process distributed with the unmarked reduced Palm distribution and iid. marks that are independent of the ground process. Let $\stackrel{d}{=}$ denote equality in distribution, then the result can more formally be stated as follows .

Theorem 3.4.3. Consider a marked point process $\Phi = \{(\xi, m_\xi) : \xi \in X\}$ with ground process X and consider for any $\xi \in S$ and $m \in M$ the related reduced Palm marked point process $\Phi^!_{(\xi,m)} = \{(\eta, l_\eta) : \eta \in Y_{(\xi,m)}\}$ with ground process $Y_{(\xi,m)}$. If Φ is under independent marking, then $\{l_\eta : \eta \in Y_{(\xi,m)}\} \stackrel{d}{=} \{m_\xi : \xi \in X\}$ is independent of $Y_{(\xi,m)} \stackrel{d}{=} X^!_\xi$, i.e. the distribution of $Y_{(\xi,m)}$ does not depend on m.

Proof. To prove this theorem, the left and then the right hand side (abbreviated LHS and RHS) of the Campbell-Mecke theorem for marked point processes, stated in equation (3.13), are considered for a specific choice of h. That is, $h(\xi, m, \phi) = h_1(\xi)h_2(m)h_3(\phi) = h_1(\xi)h_2(m)h_{31}(x)h_{32}(\{m_1, m_2, \dots\})$, where ϕ is a marked point configuration with points $x = \{\xi_1, \xi_2, \dots\}$ and marks $\{m_{\xi_1}, m_{\xi_2}, \dots\}$. Note that since the marks are identically distributed and independent of X, the mark distribution $Q_\xi = Q$ and $m_\xi \sim Q$ for all $\xi \in S$

as derived in section 3.3.1. Then, by iterated expectations,

$$
\begin{aligned}
LHS &= \mathrm{E}\left[\sum_{(\xi,m_\xi)\in\Phi} h_1(\xi)h_2(m_\xi)h_{31}(X\backslash\xi)h_{32}(\{m_\eta:\eta\in X\}\backslash m_\xi)\right] \\
&= \mathrm{E}\left[\mathrm{E}\left[\sum_{(\xi,m_\xi)\in\Phi} h_1(\xi)h_2(m_\xi)h_{31}(X\backslash\xi)h_{32}(\{m_\eta:\eta\in X\}\backslash m_\xi)\middle| X\right]\right] \\
&= \mathrm{E}\left[\sum_{(\xi,m_\xi)\in\Phi} h_1(\xi)h_{31}(X\backslash\xi)\right]\mathrm{E}\left[h_2(m_\xi)\right]\mathrm{E}\left[h_{32}(\{m_\eta:\eta\in X\}\backslash m_\xi)\right] \\
&= \int\int h_1(\xi)h_{31}(x)dP^!_\xi(x)d\mu_p(\xi)\int h_2(m)dQ(m)\mathrm{E}\left[h_{32}(\{m_\eta:\eta\in X\}\backslash m_\xi)\right],
\end{aligned}
$$

where the third equality follows from the assumption that $\{m_\eta:\eta\in X\}$ are iid. and independent of X and the last equality follows from the Campbell-Mecke theorem for unmarked point processes stated in equation (2.3).

Consider now the right hand side,

$$
\begin{aligned}
RHS &= \int\int\int h_1(\xi)h_2(m)h_3(\phi)dP^!_{(\xi,m)}(\phi)dQ(m)d\mu_p(\xi) \\
&= \int\int h_1(\xi)h_2(m)\mathrm{E}\left[h_3\left(\Phi^!_{(\xi,m)}\right)\right]dQ(m)d\mu_p(\xi) \\
&= \int\int h_1(\xi)h_2(m)\mathrm{E}\left[h_{31}(Y_{(\xi,m)})h_{32}(\{l_\eta:\eta\in Y_{(\xi,m)}\})\right]dQ(m)d\mu_p(\xi).
\end{aligned}
$$

By proposing $Y_{(\xi,m)}\sim P^!_\xi$ and $\{l_\eta:\eta\in Y_{(\xi,m)}\}\stackrel{d}{=}\{m_\eta:\eta\in X\}$ independent of $Y_{(\xi,m)}$, we can show that

$$
\begin{aligned}
RHS &= \int\int h_1(\xi)h_2(m)\mathrm{E}\left[h_{31}(Y_{(\xi,m)})\right]\mathrm{E}\left[h_{32}(\{l_1,l_2,\dots\})\right]dQ(m)d\mu_p(\xi) \\
&= \int h_1(\xi)\mathrm{E}\left[h_{31}(Y_{(\xi,m)})\right]d\mu_p(\xi)\int h_2(m)dQ(m)\mathrm{E}\left[h_{32}(\{l_1,l_2,\dots\})\right] \\
&= LHS.
\end{aligned}
$$

This satisfies the Campbell-Mecke theorem and thus ends the proof.

□

For a fixed Borel set $B\subseteq S$ and a fixed $F\subseteq N_{\mathrm{lf}}$, clearly $C^!(\cdot\times B\times F)\ll\mu_p(\cdot)$. Thus there exists a Radon Nikodym derivative, $P^{*!}_\xi(B\times F)$, such that

$$
C^!(A\times B\times F)=\int_A P^{*!}_\xi(B\times F)d\mu_p(\xi) \tag{3.14}
$$

for all Borel sets $A \subseteq S$ and $B \subseteq M$ and all $F \subseteq N_{\text{lf}}$. Per equation (3.12),

$$P_\xi^{*!}(B \times F) = \int_B P^!_{(\xi,m)}(F) dQ_\xi(m)$$

for μ_p-a.e. ξ, and thus $P^!_{(\xi,m)}$ is the Radon-Nikodym derivative of $P_\xi^{*!}$ with respect to Q_ξ. It then follows from proposition A1.5.II in Daley and Vere-Jones (2002) that $P_\xi^{*!}$ is a probability measure on $M \times N_{\text{lf}}$, where the condition that $P^!_{(\xi,\cdot)}(F)$ is measurable as a function of m is satisfied, since $P^!_{(\cdot,\cdot)}(F)$ is measurable on $S \times M$ for fixed F.

Definition 3.4.4. The probability measure $P_\xi^{*!}$ is called the *reduced Palm distribution at point ξ for its mark and the rest of the marked point process.*

Note that, if $(m_\xi^{*!}, \Phi_\xi^{*!}) \sim P_\xi^{*!}$, then

$$P(m_\xi^{*!} \in B) = P_\xi^{*!}(B \times N_{\text{lf}}) = Q_\xi(B)$$

for all Borel sets $B \subseteq M$ and

$$P(\Phi_\xi^{*!} \in F) = P_\xi^{*!}(M \times F) = \int_M P^!_{(\xi,m)}(F) dQ_\xi(m)$$

for all $F \subseteq N_{\text{lf}}$. That is, $m_\xi^{*!} \sim Q_\xi$ and $(\Phi_\xi^{*!} | m_\xi^{*!} = m) \sim P^!_{(\xi,m)}$. We may interpret $P_\xi^{*!}$ as the joint distribution of the mark m_ξ and $\Phi \backslash (\xi, m_\xi)$ conditional on $\xi \in X$.

Let Φ be a stationary marked point process with intensity ρ, i.e. $Q_\xi = Q$ and $d\mu_p(\xi) = \rho d\xi$ for all $\xi \in S$. It then follows from the Campbell-Mecke theorem, stated in equation (3.13), that

$$\begin{aligned} \mathrm{E}\left[\sum_{(\xi,m_\xi)\in\Phi} h(\xi, m_\xi, \Phi\backslash(\xi, m_\xi))\right] &= \rho \int\int\int h(\xi, m, \phi) dP^!_{(\xi,m)}(\phi) dQ(m) d\xi \\ &= \rho \int\int\int h(\xi, m, \phi + \xi) dP^!_{(0,m)}(\phi) dQ(m) d\xi. \end{aligned} \tag{3.15}$$

For $\xi \in S$ and $m \in M$, $\Phi^!_{(0,m)} \sim P^!_{(0,m)}$ if and only if $\Phi^!_{(0,m)} + \xi \sim P^!_{(\xi,m)}$. We say that $P^!_{(0,m)}$ is the reduced Palm distribution in a typical point with mark m for the rest of the marked point process. Furthermore, $P_0^{*!}(B \times F) = \int_B P^!_{(0,m)}(F) dQ(m)$ and equation (3.15) can be written as

$$\mathrm{E}\left[\sum_{(\xi,m_\xi)\in\Phi} h(\xi, m_\xi, \Phi\backslash(\xi, m_\xi) - \xi)\right] = \rho \int\int h(\xi, m, \phi) dP_0^{*!}(\phi, m) d\xi$$

for any function $h : S \times M \times N_{\text{lf}} \to [0, \infty)$. Specifically, for $h(\xi, m, \phi) = \mathbb{I}(\xi \in A, m \in B, \phi \in F)$, where $A \subseteq S$ Borel, $B \subseteq M$ Borel and $F \subseteq N_{\text{lf}}$,

$$\mathrm{E}\left[\sum_{(\xi,m_\xi)\in\Phi} \mathbb{I}(\xi \in A, m_\xi \in B, (\Phi\backslash(\xi, m_\xi) - \xi) \in F)\right] = \rho |A| P_0^{*!}(B \times F).$$

Thus, for any function $h : M \times N_{\mathrm{lf}} \to [0, \infty)$ and any Borel set $A \subseteq S$ with $0 < |A| < \infty$,

$$\mathrm{E}\left[h\left(m_0^{*!}, \Phi_0^{*!}\right)\right] = \frac{1}{\rho|A|}\mathrm{E}\left[\sum_{(\xi, m_\xi)\in\Phi} \mathbb{I}\left(\xi \in A\right) h\left(m_\xi, (\Phi\backslash(\xi, m_\xi)) - \xi\right)\right], \quad (3.16)$$

where $(m_0^{*!}, \Phi_0^{*!}) \sim P_0^{*!}$. We say that $P_0^{*!}$ is the reduced Palm distribution in a typical point for its mark and the rest of the marked point process.

3.5 Functional summary statistics

Several functional summary statistics have been proposed in the literature in order to investigate relations between marks and points. One example is the mark correlation function e.g. described in Illian et al. (2008), Stoyan and Stoyan (1994) and Penttinen and Stoyan (1989). These sources however limit themselves to stationary and isotropic processes with either discrete or real marks. In our application the marks are two-dimensional, and thus we extend the definitions to the case where the marks are in $\mathbb{R}^k$, $k \geq 1$. We also present a new summary statistic for detecting columnar structures.

3.5.1 Mark correlation functions

We define the *non-normalised mark correlation function* by

$$c_t(\xi, \eta) = \int t(m, l) dQ_{\xi,\eta}(m, l),$$

for some test function $t : M \times M \to [0, \infty)$. Due to the interpretation of $Q_{\xi,\eta}$, we can interpret $c_t(\xi, \eta)$ as the mean of $t(m_\xi, m_\eta)$ conditional on $\xi, \eta \in X$. Note that

$$c_t(\xi, \eta) = \frac{\rho_t^{(2)}(\xi, \eta)}{\rho^{(2)}(\xi, \eta)} \quad (3.17)$$

for Lebesgue-a.e. ξ, η as per equation (3.9).

It can be an interpretational advantage to normalise c_t by c_t^*, which is defined similar to c_t but under the assumption of independent marking, i.e.

$$c_t^* = \int\int t(m, l) dQ(m) dQ(l). \quad (3.18)$$

Due to theorem 3.3.2, $c_t^* = \mathrm{E}\left[t(m, l)\right]$ for $m, l \sim Q$ independent. This normalisation yields the mark correlation function.

Definition 3.5.1. The *mark correlation function* is defined as

$$k_t(\xi, \eta) = \frac{c_t(\xi, \eta)}{c_t^*}, \quad (3.19)$$

where we take $a/0 = 1$ for $a \geq 0$.

For a stationary marked point process, $k_t(\xi, \eta) = k_t(\xi - \eta)$ (abusing notation), since $Q_{\xi,\eta} = Q_{\xi-\eta}$ for all $\xi, \eta \in \mathbb{R}^d$ as discussed in section 3.3.2. If in addition the marked point process is isotropic, then (again abusing notation) $k_t(\xi, \eta) = k_t(\|\xi - \eta\|)$, since $Q_{\xi,\eta} = Q_{\|\xi-\eta\|}$ for all $\xi, \eta \in \mathbb{R}^d$, also noted in section 3.3.2.

The mark correlation function is obviously defined in such a way that under independent marking,

$$k_t(\xi, \eta) = 1$$

for any $\xi, \eta \in S$. If $k_t(\xi, \eta) > 1$ (resp. $k_t(\xi, \eta) < 1$) it indicates that the test function evaluated in m_ξ and m_η is greater (resp. smaller) than under independent marking. The interpretation naturally depends on the choice of test function. Consider a test function that decreases as the marks become more alike, e.g. the angle between two vectors. Then $k_t(\xi, \eta) < 1$ (resp. $k_t(\xi, \eta) > 1$) indicates that m_ξ and m_η tend to be more alike (resp. different) for the specified point process than under independent marking. On the other hand, if the test function increases as the marks become more alike, the interpretation is the other way around.

3.5.2 Mark-weighted $\mathcal{K}$-measure

The so-called mark-weighted K-function is discussed in Illian et al. (2008), Stoyan and Stoyan (1994) and Penttinen and Stoyan (1989) for real marks under the assumption of stationarity and isotropy. In this section we generalise these results by considering a mark-weighted version of the $\mathcal{K}$-measure, introduced in definition 2.4.5, that allows for multidimensional (Euclidean) marks.

Definition 3.5.2. Assume that the intensity function, ρ, exists and that

$$\mathcal{K}_t(B) = \frac{1}{|A|} \mathrm{E}\left[\sum_{(\xi,m_\xi),(\eta,m_\eta)\in\Phi}^{\neq} \frac{\mathbb{I}(\xi \in A)\mathbb{I}(\eta - \xi \in B)t(m_\xi, m_\eta)}{\rho(\xi)\rho(\eta)c_t^*} \right] \tag{3.20}$$

does not depend on the choice of $A \subseteq S$ Borel with $0 < |A| < \infty$, where $\frac{a}{0} = 0$ for $a \geq 0$, $B \subseteq S$ Borel and $t : M \times M \to [0, \infty)$ is some test function. Then $\mathcal{K}_t$ is called the *mark-weighted $\mathcal{K}$-measure*.

Note that, when $t = 1$, $\mathcal{K}_t(B) = \mathcal{K}(B)$ for all Borel sets $B \subseteq S$.

Assume that $\mathcal{K}_t$ exists for some choice of test function t. By applying the Campbell-Mecke theorem for marked point processes, stated in equation

(3.13), we see that $\mathcal{K}_t$ has a Palm interpretation. That is,

$$\begin{aligned}
&\mathcal{K}_t(B)|A| \\
&= \int\int\int \sum_{(\eta,m_\eta)\in\phi} \frac{\mathbb{I}(\xi\in A)\mathbb{I}(\eta-\xi\in B)t(m,m_\eta)}{\rho(\xi)\rho(\eta)c_t^*} dP^!_{(\xi,m)}(\phi)dQ_\xi(m)d\mu_p(\xi) \\
&= \int \mathbb{I}(\xi\in A)\int \sum_{(\eta,m_\eta)\in\phi} \frac{\mathbb{I}(\eta-\xi\in B)t(m,m_\eta)}{\rho(\eta)c_t^*} dP^{*!}_\xi(m,\phi)d\xi.
\end{aligned}$$

Therefore,

$$\mathcal{K}_t(B) = \mathrm{E}\left[\sum_{(\eta,m_\eta)\in\Phi^{*!}_\xi} \frac{\mathbb{I}(\eta-\xi\in B)t(m^{*!}_\xi,m_\eta)}{\rho(\eta)c_t^*}\right] \tag{3.21}$$

for Lebesgue a.e ξ, where $(m^{*!}_\xi,\Phi^{*!}_\xi)\sim P^{*!}_\xi$. In the stationary case

$$\mathcal{K}_t(B) = \frac{1}{\rho c_t^*}\mathrm{E}\left[\sum_{(\eta,m_\eta)\in\Phi^{*!}_0} \mathbb{I}(\eta\in B)t(m^{*!}_0,m_\eta)\right] \tag{3.22}$$

for all Borel sets $B\subseteq S$.

Consider a marked point process with translation invariant pair and mark correlation functions. Then by the two-point Campbell theorem, found in (3.7), the right hand side of equation (3.20) can be written as

$$\begin{aligned}
&\frac{1}{|A|}\int\int \frac{\mathbb{I}(\xi\in A)\mathbb{I}(\eta-\xi\in B)t(m_\xi,m_\eta)}{\rho(\xi)\rho(\eta)c_t^*} dQ_{\xi,\eta}(m_\xi,m_\eta)d\alpha_p^{(2)}(\xi,\eta) \\
&= \frac{1}{|A|}\int\int \frac{\mathbb{I}(\xi\in A)\mathbb{I}(\eta-\xi\in B)\rho^{(2)}(\xi,\eta)c_t(\xi,\eta)}{\rho(\xi)\rho(\eta)c_t^*} d\xi d\eta \\
&= \frac{1}{|A|}\int\int \mathbb{I}(\xi\in A)\mathbb{I}(\eta-\xi\in B)g(\eta-\xi)k_t(\eta-\xi)d\xi d\eta \\
&= \frac{1}{|A|}\int\int \mathbb{I}(\xi\in A)\mathbb{I}(\omega\in B)g(\omega)k_t(\omega)d\xi d\omega = \int_B g(\omega)k_t(\omega)d\omega
\end{aligned}$$

with $\omega=\eta-\xi$. This expression does not depend on A, and hence translation invariance of g and k_t entail the existence of $\mathcal{K}_t$ for any test function $t: M\times M\to[0,\infty)$. Due to equation (2.12) this furthermore yields that

$$\mathcal{K}_t(B) = \int_B k_t(\xi)d\mathcal{K}(\xi)$$

for all Borel sets $B\subseteq S$. I Φ is stationary, both g and k_t are translation invariant and thus $\mathcal{K}_t$ exists for any test function $t: M\times M\to[0,\infty)$.

It follows from definition 3.5.2 and equation (3.8) that

$$\int\int \mathbb{I}(\xi \in A, \eta \in B) d\mathcal{K}_t(\eta) d\xi = \int \frac{\mathbb{I}(\xi \in A, \eta - \xi \in B)}{\rho(\xi)\rho(\eta)c_t^*} d\alpha_t^{(2)}(\xi, \eta), \quad (3.23)$$

which under stationarity entails

$$\alpha_t^{(2)}(A \times B) = \rho^2 c_t^* \int_A \mathcal{K}_t(B - \xi) d\xi$$

for any Borel sets $A, B \subseteq S$.

Theorem 3.5.3. Under independent marking $\mathcal{K}_t(B) = \mathcal{K}(B)$.

Proof. Consider a marked point process Φ with ground process X. Then, by iterated expectation and the definition of c_t^* in equation (3.18),

$$\mathcal{K}_t(B) = \frac{1}{|A|} \mathrm{E}\left[\mathrm{E}\left[\sum_{(\xi, m_\xi), (\eta, m_\eta) \in \Phi}^{\neq} \frac{\mathbb{I}(\xi \in A)\mathbb{I}(\eta - \xi \in B) t(m_\xi, m_\eta)}{\rho(\xi)\rho(\eta)c_t^*} \,\middle|\, X \right] \right]$$

$$= \frac{1}{|A|} \mathrm{E}\left[\sum_{\xi, \eta \in X}^{\neq} \frac{\mathbb{I}(\xi \in A)\mathbb{I}(\eta - \xi \in B)}{\rho(\xi)\rho(\eta)c_t^*} \mathrm{E}\left[t(m, l) \right] \right] = \mathcal{K}(B),$$

where $m, l \sim Q$ are independent. □

Based on the mark-weighted $\mathcal{K}$-measure we define the following generalisations of the K-function and the associated L-function.

Definition 3.5.4. If $\mathcal{K}_t$ exists, we define the *mark-weighted K-function* and the *mark-weighted L-function* by

$$K_t(r) = \mathcal{K}_t(b(0, r)) \quad \text{and} \quad L_t(r) = (K_t(r)/\omega_d)^{1/d}$$

for $r > 0$.

Theorem 3.5.3 can then be used to check for independent marking by plotting $K_t(r)$ against $K(r)$, or equivalently $K_t(r) - K(r)$ against r for $r > 0$.

3.5.3 Modified cylindrical K-function

In this section we introduce a new summary statistic for marked point patterns with marks describing the direction of the points; as for the pyramidal cell data. This new summary statistic is strongly inspired by the cylindrical K-function and the mark-weighted $\mathcal{K}$-measure. Specifically, the new summary statistic is similar to the mark-weighted $\mathcal{K}$-measure with the cylinder as structuring element, but where the cylinder's direction varies across points.

In the following, for $u \in \mathbb{S}^{d-1}$, $d \geq 2$, let $C_u^2(r,h) = C_u(r,h)$ and $C_u^1(r,h) = \mathcal{O}_u\tilde{C}(r,h)$ where

$$\tilde{C}(r,h) = \{(x_1,\ldots,x_d) \in \mathbb{R}^d : x_1^2 + \cdots + x_{d-1}^2 \leq r, 0 \leq x_d \leq h\}$$

and $\mathcal{O}_u$ is the rotation matrix described in section 2.4.2. Note that $C_u^1(r,h)$ and $C_u^2(r,h)$ both are cylinders with direction u and radius r, but $C_u^1(r,h)$ has height h, while $C_u^2(r,h)$ has height $2h$.

Definition 3.5.5. Consider a marked point process with mark space $M \subseteq \mathbb{S}^{d-1}$, $d \geq 2$. Assume that the intensity function, ρ, exists and that, for $i = 1$ or $i = 2$,

$$K_t^{\circ i}(r,h) = \frac{1}{|A|}\mathrm{E}\left[\sum_{(\xi,m_\xi),(\eta,m_\eta)\in\Phi}^{\neq} \frac{\mathbb{I}\big(\xi \in A, \eta - \xi \in C_{m_\xi}^i(r,h)\big)t(m_\xi,m_\eta)}{\rho(\xi)\rho(\eta)c_t^*}\right]$$

does not depend on the choice of $A \subseteq S$ with $0 < |A| < \infty$, where $r > 0$, $h > 0$, $t : M \times M \to [0,\infty)$ is some test function and $\frac{a}{0} = 0$ for $a \geq 0$. Then $K_t^{\circ i}$ is called the *modified (mark-weighted) cylindrical K-function*.

In a similar manner to the derivation of equation (3.21), we can show that $K_t^{\circ i}$ has a Palm interpretation. That is,

$$K_t^{\circ i}(r,h) = \mathrm{E}\left[\sum_{(\eta,m_\eta)\in\Phi_\xi^{*!}} \frac{\mathbb{I}\big(\eta - \xi \in C_{m_\xi^{*!}}^i(r,h)\big)t(m_\xi^{*!},m_\eta)}{\rho(\eta)c_t^*}\right]$$

for Lebesgue-a.e. ξ and $(m_\xi^{*!}, \Phi_\xi^{*!}) \sim P_\xi^{*!}$. This implies in the stationary case that

$$\rho c_t^* K_t^{\circ i}(r,h) = \mathrm{E}\left[\sum_{(\eta,m_\eta)\in\Phi_0^{*!}} \mathbb{I}\big(\eta \in C_{m_0^{*!}}^i(r,h)\big)t(m_0^{*!},m_\eta)\right].$$

Hence, for $t = 1$ we can interpret $\rho K_t^{\circ i}(r,h)$ as the expected number of further points in the cylinder with direction of the typical mark.

3.6 Estimation

In this section we describe how to estimate the summary statistics presented in section 3.5. Throughout the section we consider a realisation ϕ of a marked point process Φ, where the points of ϕ lie in a bounded window $W \subseteq S$.

3.6.1 Estimating mark correlation functions

In order to estimate the mark correlation function it is convenient to assume that $\rho^{(2)}(\xi,\eta) = \rho^{(2)}(\|\xi-\eta\|)$ and $\rho_t^{(2)}(\xi,\eta) = \rho_t^{(2)}(\|\xi-\eta\|)$. Note that stationary and isotropic marked point processes naturally meet these assumptions, since by equation (3.9), $\rho^{(2)}(\xi,\eta) = \rho^{(2)}(\|\xi-\eta\|)$ implies $\rho_t^{(2)}(\xi,\eta) = \rho_t^{(2)}(\|\xi-\eta\|)$.

Due to equation (3.17) and (3.19),

$$k_t(r) = \frac{1}{c_t^*}\frac{\rho_t^{(2)}(r)}{\rho^{(2)}(r)}$$

for $r > 0$, and $k_t(r)$ can thus be estimated by separately estimating $\rho^{(2)}(r)$, $\rho_t^{(2)}(r)$ and c_t^*. A general kernel estimate of the second-order intensity function (see e.g. Stoyan and Stoyan, 1994) is given by

$$\hat{\rho}^{(2)}(r) = \frac{1}{\sigma_d r^{d-1}} \overset{\neq}{\sum_{(\xi,m_\xi),(\eta,m_\eta)\in\phi}} \frac{k^{(h)}(r-\|\eta-\xi\|)}{|W\cap W_{\eta-\xi}|}$$

for $r > 0$, where σ_d is the surface area of the d-dimensional unit ball and it is assumed that $|W\cap W_{\eta-\xi}| > 0$. Furthermore, for $u\in\mathbb{R}$, $k^{(h)}(u) = k(u/h)/h$ for a kernel $k(\cdot)$ and bandwidth $h > 0$. Similarly, the second-order t-intensity function can be estimated by

$$\hat{\rho}_t^{(2)}(r) = \frac{1}{\sigma_d r^{d-1}} \overset{\neq}{\sum_{(\xi,m_\xi),(\eta,m_\eta)\in\phi}} \frac{k^{(h)}(r-\|\eta-\xi\|)t(m_\xi,m_\eta)}{|W\cap W_{\eta-\xi}|}.$$

For the kernel estimates, Illian et al. (2008) recommend to use the uniform kernel given by $k(u) = \frac{1}{2}\mathbb{I}(|u|\le 1)$ and as general rule of thumb to choose a bandwidth of order $h \approx 0.1/\sqrt{\rho}$ for $d = 2$ and $h \approx 0.05/\sqrt[3]{\rho}$ for $d = 3$. These recommendations have been followed in our application.

Finally, c_t^* may be estimated by

$$\hat{c}_t^* = \frac{1}{n(\phi)^2} \sum_{(\xi,m_\xi),(\eta,m_\eta)\in\phi} t(m_\xi,m_\eta), \tag{3.24}$$

where $n(\phi)$ is the number of points in ϕ, which is seen to be an unbiased estimate under independent marking, since

$$\begin{aligned}
\mathrm{E}\left[\hat{c}_t^*\right] &= \mathrm{E}\left[\frac{1}{n(\Phi_W)^2} \sum_{(\xi,m_\xi),(\eta,m_\eta)\in\Phi_W} t(m_\xi,m_\eta)\right] \\
&= \mathrm{E}\left[\frac{1}{n(\Phi_W)^2} \sum_{(\xi,m_\xi),(\eta,m_\eta)\in X_W} \mathrm{E}\left[t(m_\xi,m_\eta)|X\right]\right] \\
&= \mathrm{E}\left[\frac{n(\Phi_W)^2}{n(\Phi_W)^2} c_t^*\right] = c_t^*.
\end{aligned}$$

3.6.2 Estimating the mark-weighted $\mathcal{K}$-measure

The mark-weighted $\mathcal{K}$-measure introduced in section 3.5.2 can be estimated in a similar fashion as the second-order reduced moment measure, $\mathcal{K}$ (see e.g. Møller and Waagepetersen, 2004).

Theorem 3.6.1. Assume that $\mathcal{K}_t$ exists for the function $t : M \times M \to [0, \infty)$ and $|W \cap W_\eta| > 0$ for all $\xi \in B$, where $B \subseteq S$ Borel. Then

$$\hat{\mathcal{K}}_t(B) = \sum_{(\xi,m_\xi),(\eta,m_\eta)\in\phi}^{\neq} \frac{\mathbb{I}(\eta - \xi \in B)t(m_\xi, m_\eta)}{\rho(\xi)\rho(\eta)c_t^*|W \cap W_{\eta-\xi}|}$$

is an unbiased estimate of $\mathcal{K}_t(B)$.

Proof. Let $\Phi_W = \{(\xi, m_\xi) \in \Phi : \xi \in X_W\}$. Applying equation (3.8) and a more general version of (3.23) (obtained from using the standard proof) gives

$$\begin{aligned}
\mathrm{E}\left[\hat{\mathcal{K}}_t(B)\right] &= \mathrm{E}\left[\sum_{(\xi,m_\xi),(\eta,m_\eta)\in\Phi_W}^{\neq} \frac{\mathbb{I}(\eta - \xi \in B)t(m_\xi, m_\eta)}{\rho(\xi)\rho(\eta)c_t^*|W \cap W_{\eta-\xi}|}\right] \\
&= \int \frac{\mathbb{I}(\xi \in W, \eta \in W, \eta - \xi \in B)}{\rho(\xi)\rho(\eta)c_t^*|W \cap W_{\eta-\xi}|} d\alpha_t(\xi, \eta) \\
&= \int \frac{\mathbb{I}(\xi \in W \cap W_{\xi-\eta}, \eta - \xi \in B)}{\rho(\xi)\rho(\eta)c_t^*|W \cap W_{\eta-\xi}|} d\alpha_t(\xi, \eta) \\
&= \int_B \int \frac{\mathbb{I}(\xi \in W \cap W_{-\eta})}{|W \cap W_\eta|} d\xi d\mathcal{K}_t(\eta) = \mathcal{K}_t(B),
\end{aligned}$$

where the last equality follows from the fact that $|W \cap W_\eta| = |W \cap W_{-\eta}|$ (derived in Christensen and Christoffersen, 2015). □

Hence, in case of stationarity an unbiased estimate of the mark weighted K-function is

$$\hat{K}_t(r) = \frac{1}{\rho^2 c_t^*} \sum_{(\xi,m_\xi),(\eta,m_\eta)\in\phi}^{\neq} \frac{\mathbb{I}(\|\eta - \xi\| \leq r)t(m_\xi, m_\eta)}{|W \cap W_{\eta-\xi}|}$$

for $r > 0$. In practice, ρ^2 and c_t^* need to be estimated, which introduces some bias. We will estimate c_t^* by (3.24) and ρ^2 by $n(\phi)(n(\phi) - 1)/|W|^2$, the latter being an unbiased estimate for a Poisson process.

3.6.3 Estimating the modified cylindrical K-function

The definition of the modified cylindrical K-function was strongly inspired by the cylindrical K-function and the mark-weighted $\mathcal{K}$-measure, and so is its

estimate. That is,

$$\hat{K}_t^{\circ i}(r,h) = \sum_{(\xi,m_\xi),(\eta,m_\xi)\in\phi}^{\neq} \frac{\mathbb{I}(\eta-\xi \in C^i_{m_\xi}(r,h))t(m_\xi,m_\eta)}{\rho(\xi)\rho(\eta)c_t^*|W\cap W_{\eta-\xi}|}$$

where i is either 1 or 2, $r>0$ and $h>0$. Again, c_t^* will be estimated by (3.24) and, assuming stationarity, ρ^2 is estimated by $n(\phi)(n(\phi)-1)/|W|^2$.

3.7 Analysis of pyramidal cells' location and orientation

In this section we will apply the theory presented in this chapter to the pyramidal cell data. As mentioned earlier, the points describe are the three-dimensional coordinates of the pyramidal cells' nucleoli, and the marks are unit vectors describing their orientation First, a number of plots will be presented and used in a discussion of the data. This will include considerations of the marks' distribution and their relation to the nucleolus locations. We then test for independent marking and finish with a discussion of possible extensions of the analysis.

3.7.1 Investigate data

In section 2.5.1 we chose an eroded observation window for the nucleolus locations. In this marked point pattern analysis we shall use the same observation window, that is $[0,1382]\times[0,145]\times[10,485]$ μm^3, which contains a total of 2060 pyramidal cells.

Figure 3.1 displays the marks as three-dimensional unit vectors and gives an indication of the distribution of the marks. A great part of the marks point in the (approximate) positive z-direction, i.e. directed towards the pile surface of the brain. In fact, only 87 of the 2060 marks have a negative z-coordinate. Figure 3.2 confirms that the marks tend to point upwards and also shows that most have a negative x-coordinate. Furthermore, not much variation is found in the y-direction, i.e. the mark's y-coordinate tend to be closer to zero than the x- and z-coordinate. To explore the cell directions further we consider the mark as spherical angles, denoted ϕ and θ. Here $\phi\in[0,\pi]$ is the polar angle between the positive z-axis and the mark, while $\theta\in(-\pi,\pi]$ is the azimuth angle between the positive x-axis and the projection of the mark onto the xy-plane. Figure 3.3 displays a kernel density estimate of the joint distribution of the spherical angles along with marginal histograms. Here we used the three-dimensional normal kernel and a small bandwidth. This provides evidence that there is an absence of marks in an area roughly defined by $\theta\in(-2.5,-0.5)\cup(0.5,2.5)$ and $\phi\in(0.75,2.75)$. There is furthermore a

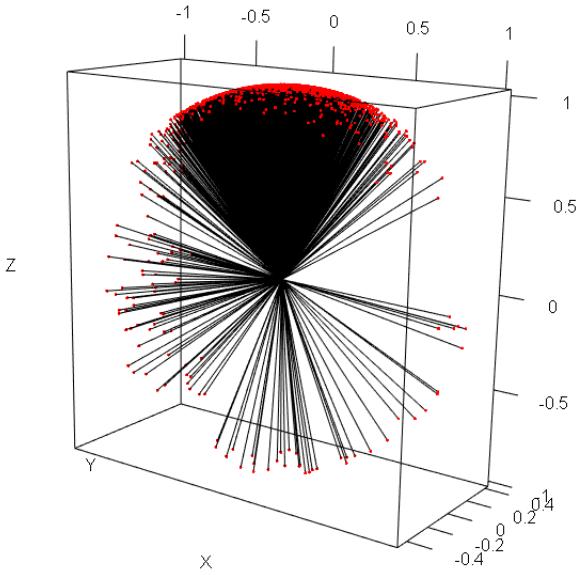

Figure 3.1: Plot of the marks as unit vectors.

high concentration of marks with polar angle $\phi \in (0, 0.75)$, especially for an azimuth angle $\theta \in (1, \pi)$. This is consistent with what we saw in figure 3.2. Figure 3.4 and 3.5, that display plots of the polar (resp. azimuth) angle versus the x-, y- and z-coordinate of the nucleolus, give no apparent indication of dependency between marks and nucleolus locations. As discussed in section 3.2.4, correlation between constructed marks describing the local intensity and the observed marks can help to identify a suitable model for the marks. In figure 3.6 the nearest neighbour distances are therefore plotted against the spherical angles, but no correlation is revealed.

3.7.2 Test for independent marking

Under independent marking the nucleolus locations and the marks are simulated separately. For this analysis we simulate the nucleolus locations from the Strauss hardcore model fitted in section 2.5.2 since we found this an appropriate model.

Fisher et al. (1987) describes several distributions for modelling spherical data. By considering figure 3.1, 3.2 and 3.3 it becomes apparent that the marks has a girdle form distribution, which excludes many spherical distributions. The girdle seems to be concentrated around the great circle in the plane which is approximately normal to the vector $(0, 1, 0)$. The intensity of the marks is however not constant on this girdle, as the marks are highly concentrated around the (approximate) positive z-axis. Except for the high concentration of polar angles between 0 and 1 radians the marks' intensity appears rather

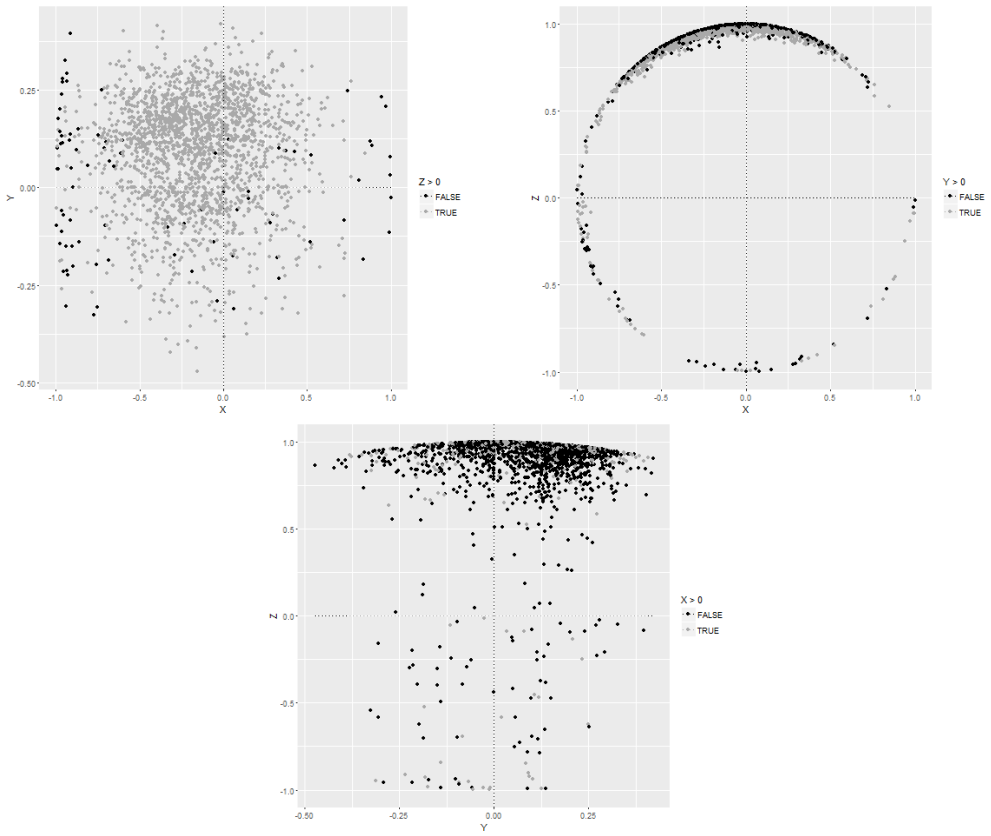

Figure 3.2: Projection of the marks onto the xy-plane (top left), the xz-plane (top right) and the yz-plane (bottom). Colours indicate the sign of the excluded coordinate (z, y and x, respectively).

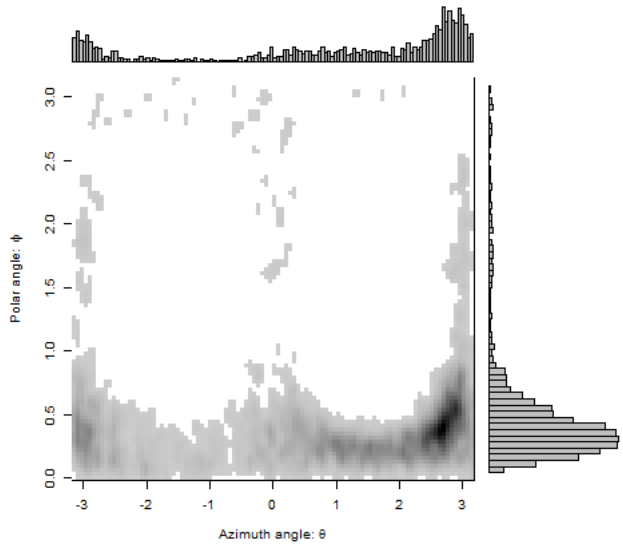

Figure 3.3: Non-parametric kernel density estimate of the marks' spherical angles together with marginal histograms.

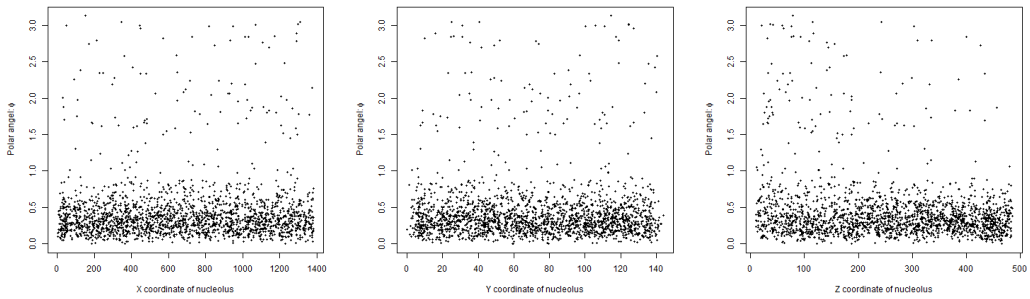

Figure 3.4: The marks' polar angles plotted against the x- (left), y- (right) and z-coordinate (left) of the nucleoli.

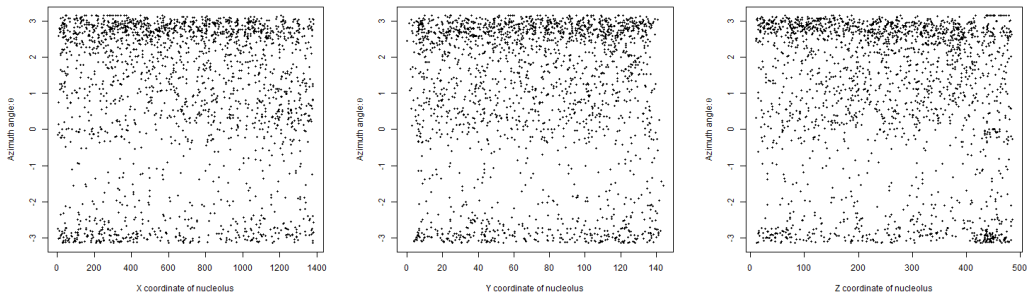

Figure 3.5: The marks' azimuth angles plotted against the x- (left), y- (right) and z-coordinate (left) of the nucleoli.

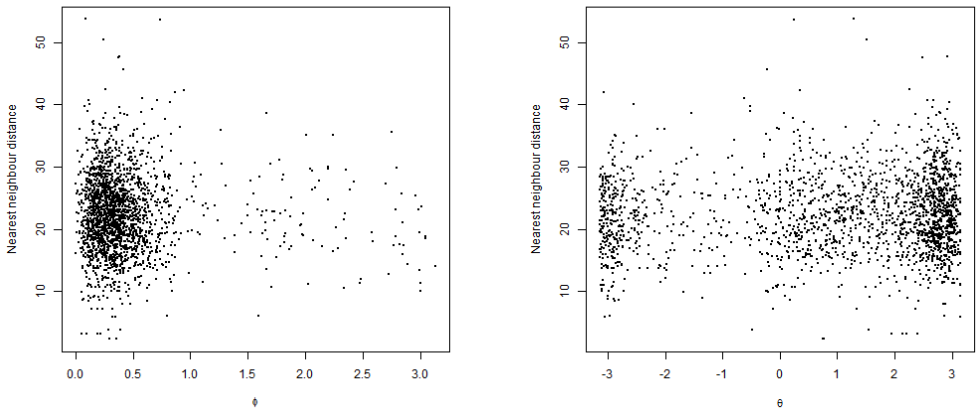

Figure 3.6: The nearest neighbour distance of the nucleolus against the mark's polar (left) and azimuth (right) angle.

constant on the girdle. These characteristics are not mimicked by the spherical distributions described by Fisher et al. (1987). The marks will therefore be simulated from a kernel density estimate for spherical data proposed by Bai and Zhao (1988). For observed marks $m_1, \ldots, m_n \in \mathbb{S}^{k-1} \subset \mathbb{R}^k$, $k \geq 2$ The kernel estimate is given by

$$\hat{f}_h(m) = \frac{c_{h,k}(L)}{n} \sum_{i=1}^{n} L\left(\frac{1 - m^T m_i}{h^2}\right), \tag{3.25}$$

where $h > 0$ is the bandwidth, $c_{h,k}(L)$ is a normalising constant and $L(\cdot)$ is the directional kernel, which is a non-negative function defined on $[0, \infty)$. For a more detailed treatment see García-Portugués et al. (2014), references therein and Bai and Zhao (1988). For the von Mises-Fisher kernel, given by $L(r) = \exp(-r)$, the kernel density estimate (3.25) can be interpreted as a mixture of von Mises-Fisher distributions with directional means m_i, $i = 1, \ldots, n$ and concentration parameter $1/h^2$ (for proof see García-Portugués et al., 2014, and references therein). That is,

$$\hat{f}_h(m) = \frac{1}{n} \sum_{i=1}^{n} f_{vMF}(m; m_i, 1/h^2) = \frac{1}{n} \sum_{i=1}^{n} C_k(1/h^2) \exp\left(\frac{m^T m_i}{h^2}\right)$$

A description of the von Mises-Fisher distribution is given by Fisher et al. (1987). Specifically, for our case with $k = 3$,

$$C_3(1/h^2) = \frac{1/h^2}{2\pi \left(\exp(1/h^2) - \exp(-1/h^2)\right)}.$$

Each mark can then be simulated from a von Mises-Fisher distribution with directional mean chosen uniformly from the observed marks and concentration $1/h^2$. Based on figure 3.7 we have chosen $h = 0.1$, which corresponds to a concentration parameter of 100.

Simulation under the von Mises-Fisher distribution was done using the `rmovMF` function from the package `movMF` in `R`. The nucleolus locations was simulated using the Metropolis-Hastings algorithm as described in section 2.3.2. In total 2499 marked point patterns were simulated in the above described manner and used for constructing global rank envelopes (as described in section 2.4.3) for $\hat{k}_{t_1}$, $\hat{K}_{t_1}$, $\hat{K}_{t_1} - \hat{K}$ (figure 3.8), $\hat{K}^{\circ 1}_{t_1}$, $\hat{K}^{\circ 2}_{t_1}$ (figure 3.9), $\hat{K}^{\circ 1}_{t_2}$ and $\hat{K}^{\circ 2}_{t_2}$ (figure 3.10) with

$$\begin{aligned} t_1(m_1, m_2) &= \arccos(m_1 \cdot m_2) \\ t_2(m_1, m_2) &= 1. \end{aligned}$$

For $\hat{K}^{\circ 1}_{t_1}$, $\hat{K}^{\circ 2}_{t_1}$, $\hat{K}^{\circ 1}_{t_2}$ and $\hat{K}^{\circ 2}_{t_2}$ we considered three different values of h, specifically $h = 40, 60$ and 80 μm, and let r vary from 0 to 35 μm. This yields a total of 15 summary statistics, of which 10 (including $\hat{k}_{t_1}$, $\hat{K}_{t_1}$ and $\hat{K}_{t_1} - \hat{K}$) have a

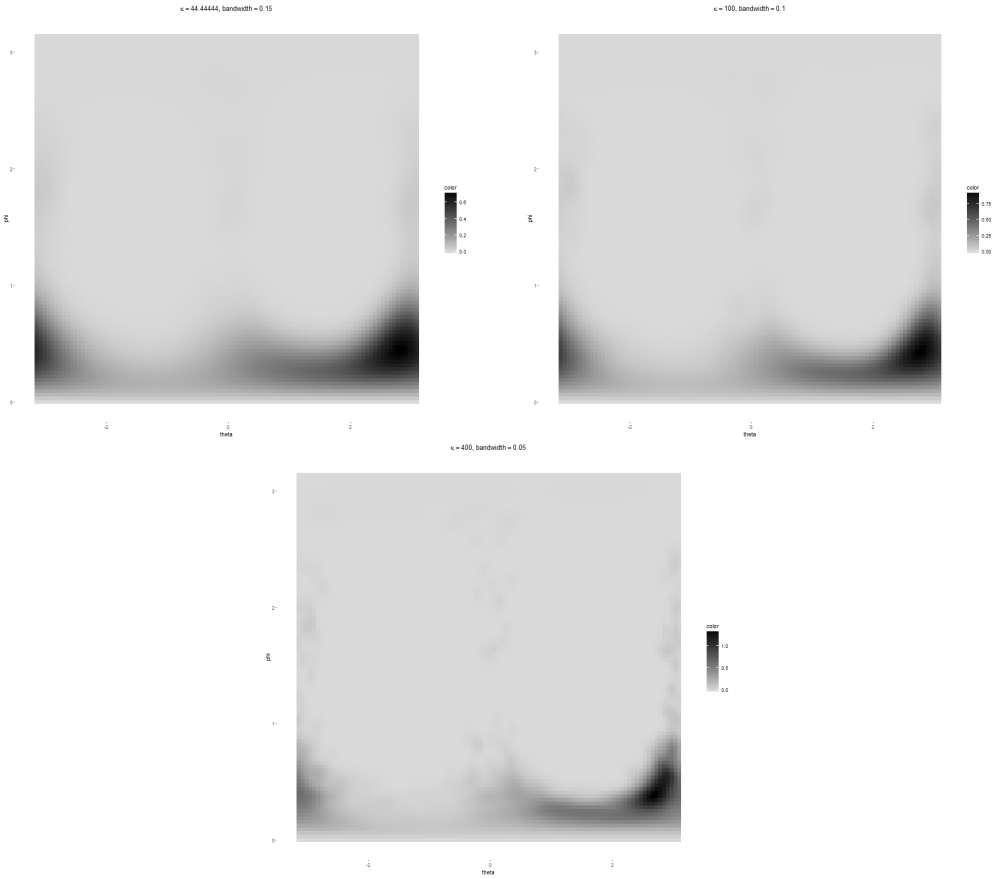

Figure 3.7: Directional kernel density estimates of the marks' spherical angles with bandwidth 0.15 (top left), 0.1 (top right) and 0.05 (bottom).

conservative p-value, p_+, smaller than 0.05. Furthermore, two of the tests has a p-interval containing 0.05, and thus we reject the hypothesis of independent marking.

For the mark-weighted K-function with test function t_1, that describes the angle between pair of marks, the estimate falls below the envelopes for several values of $r \leq 10$ (see figure 3.8). This may indicate that marks with points lying closer than 10 μm tend to have more similar directions than under independent marking, indicating some interaction between points and marks. For the mark correlation function, in figure 3.8, the estimate only falls below and not above the envelopes, indicating that the marks are more alike than expected under independent marking. However, the estimate mostly falls outside the envelopes for $r \geq 20$ μm, which is not in agreement with K_{t_1}. $K_{t_1} - K$ also rejects independent marking for $r \geq 25$ μm.

The main purpose of the modified cylindrical K-function is not to test for independent marking, but rather columnar structures. When $t = t_2$, we simply test for columnar structures with direction as the typical mark. For $t = t_1$, we have a weight that is small when the points in the cylinder have similar direction as the typical mark. We do not get a clear indication of minicolumns

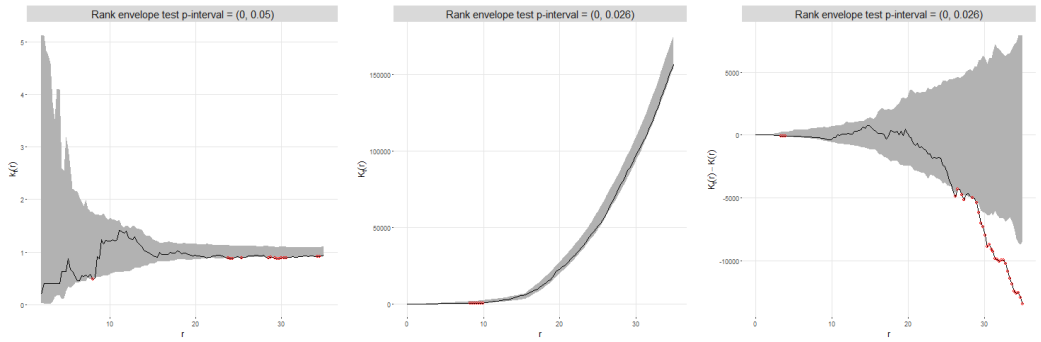

Figure 3.8: Estimates of k_{t_1} (left), K_{t_1} (middle) and $K_{t_1} - K$ (right) along with 95% global rank envelopes under independent marking. Test function is $t_1 =$ the angle between two marks.

from either figure 3.10 or 3.9. However, the estimate of the modified cylindrical K-function tends to fall below the envelopes more for $t = t_1$ than for $t = t_2$. This may be an indication that marks, whose points lie in the same cylinder, tend to be more similar than under independent marking.

In conclusion, independent marking (with a Strauss hardcore ground process) does not seem to be an adequate model for the pyramidal cell data.

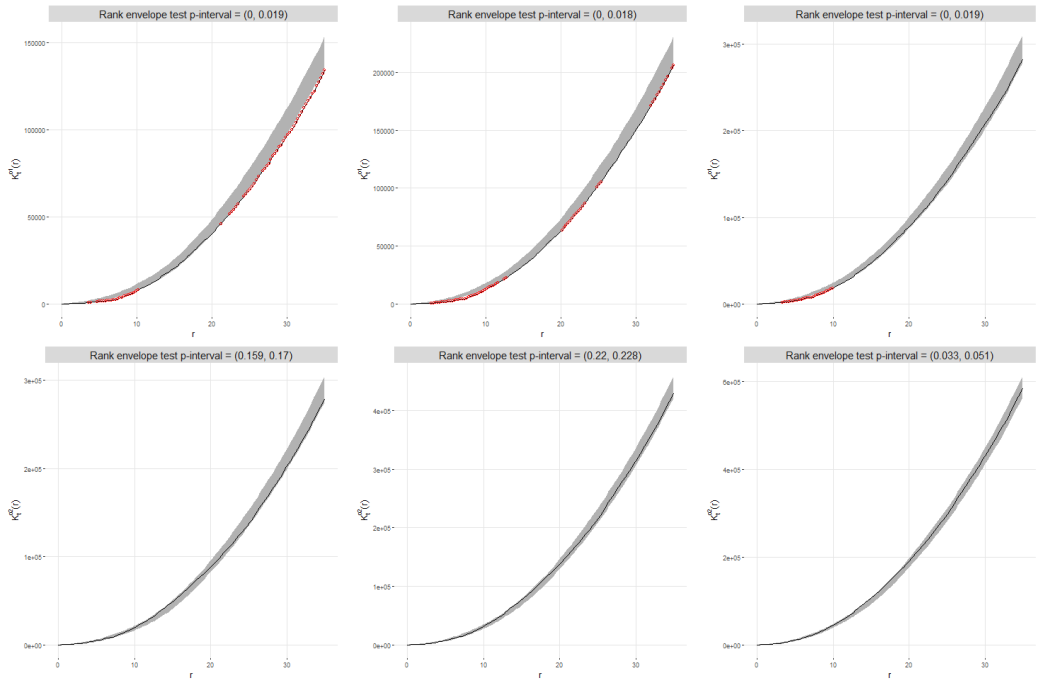

Figure 3.9: Estimates of $K_{t_1}^{\circ 1}$ (top) and $K_{t_1}^{\circ 2}$ (bottom) with $h = 40,\ 60$ and 80 μm from left to right along with 95% global rank envelopes under independent marking. Test function is $t_1 =$ the angle between two marks.

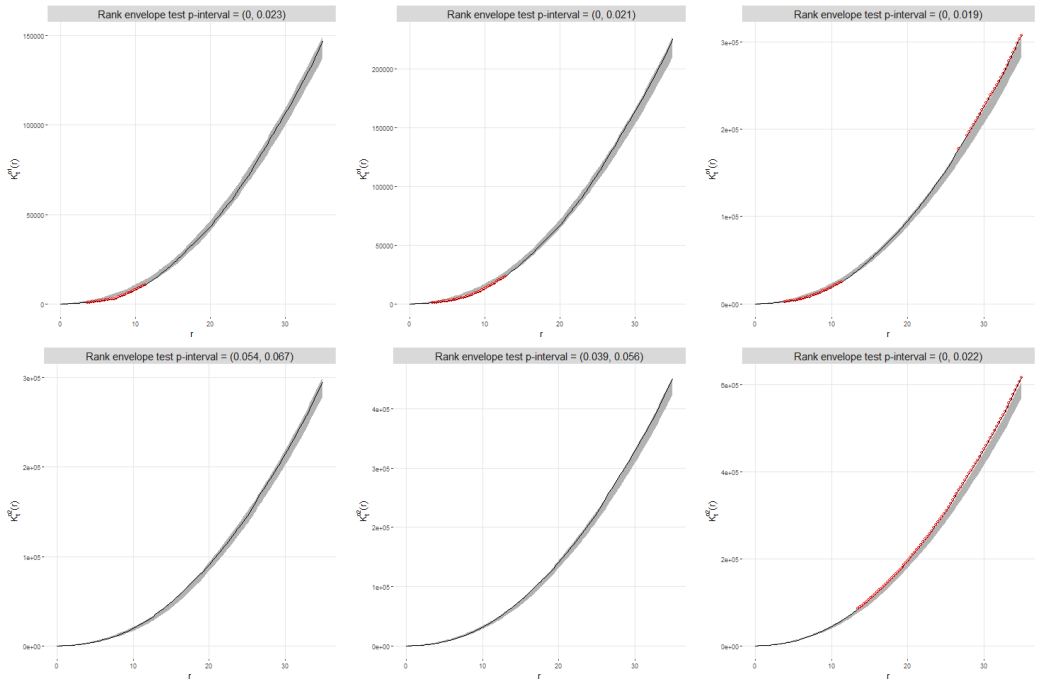

Figure 3.10: Estimates of $K_{t_2}^{\circ 1}$ (top) and $K_{t_2}^{\circ 2}$ (bottom) with $h = 40,\ 60$ and 80 μm from left to right along with 95% global rank envelopes under independent marking. Test function is $t_2 = 1$.

3.7.3 Discussion

The conclusion from the previous section may depend on several factors, e.g. the bandwidth for the kernel density estimate of the mark distribution. It would be preferable to investigate how a different (sensible) choice of bandwidth value would affect the outcome of the analysis, specifically whether a new bandwidth would lead to rejection of independent marking. The performed tests of independent marking not only depends on the estimated mark density, but also on the chosen model for the ground process. Even though the Strauss hardcore point process was accepted as a suitable model for the points it may influence the results regarding independent marking. A more thorough examination of how well the Strauss hardcore model describes different aspects of the point pattern would possibly be desired. Another possible direction to pursue in a further analysis is to model the data with marks, that are no longer required to be mutually independent, but still independent of the points. However, we saw some indication that the marks actually do depend on nearby points. Another option would thus be to build a hierarchical model, where the marks are described by some distribution conditional on the points $x = \{\xi_1, \xi_2, \ldots, \xi_n\}$. One possibility is to consider a Markov random field model with the joint mark density for unit vectors $m_1, \ldots, m_n$ (conditional on

X) given by

$$f(m_1, \ldots, m_n | X = x) \propto \exp\left(\beta \sum_{i,j:\|\xi_i - \xi_j\| \leq r} m_i \cdot m_j\right),$$

where $\beta > 0$ is a concentration parameter and $r > 0$ determines the radius of the neighbourhood that affects the mark of a point. Since $m_1, \ldots, m_n$ are unit vectors,

$$\exp(\beta m_i \cdot m_j) \leq \exp(\beta)$$

for $i, j = 1, \ldots n$. That is, the density is bounded by

$$f(m_1, \ldots, m_n | X = x) \leq \exp(\beta n) < \infty$$

and thus Ruelle stable, which implies integrability. Hence, the density is indeed well-defined. Note that the density function increases as more marks in the neighbourhoods have similar directions. That is, a mark tend to be similar to the marks in its point's neighbourhood.

References

A. Baddeley, P. J. Diggle, A. Hardegen, T. Lawrence, R. K. Milne, and G. Nair. On tests of spatial pattern based on simulation envelopes. *Ecological Monographs*, 84:477–489, 2014.

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

Adrian Baddeley and Rolf Turner. Practical maximum pseudolikelihood for spatial point patterns. *Australian and New Zealand Journal of Statistics*, 27:24–39, 2000.

Adrian Baddeley, Rolf Turner, Jorge Mateau, and Andrew Bewan. Hybrids of gibbs point process models and their implementation. *Journal of Statistical Software*, 55, 2013.

Z. D. Bai and C. Radhakrishna Rao AMD L. C. Zhao. Kernel estimators of density function of directional data. *Journal of Multivariate Analysis*, 27: 24–39, 1988.

Heidi Søgaard Christensen and Andreas Dyreborg Christoffersen. Rumlige punktprocesser, 2015.

D. J. Daley and D. Vere-Jones. *An Introduction to the Theory of Point Processes. Volume I: Elementary Theory and Methods*. Springer, New York, 2002.

D. J. Daley and D. Vere-Jones. *An Introduction to the Theory of Point Processes. Volume II: General Theory and Structure*. Springer, New York, 2008.

N. I. Fisher, T. Lewis, and B. J. J. Embleton. *Statistical analysi of spherical data*. Cambridge University Press, Cambridge, 1987.

Eduardo García-Portugués, Rosa M. Crujeiras1, and Wenceslao González-Manteiga. Kernel density estimation for directional-linear data. 2014. Available at arXiv:1210.3214v5.

Janine Illian, Antti Penttinen, Helga Stoyan, and Dietrich Stoyan. *Statistical Analysis and Modelling of Spatial Point Patterns*. John Wiley and Sons, Chichester, 2008.

Jens Ledet Jensen and Jesper Møller. Pseudolikelihood for exponential family models of spatial point processes. *The Annals of Applied Probability*, 1, 1991.

N. B. Loosmore and E. D. Ford. Statistical inference using the G or K point pattern spatial statistics. *Ecology*, 87:1925–1931, 2006.

Jesper Møller and Rasmus Plenge Waagepetersen. *Statistical Inference and Simulation for Spatial Point Processes*. Chapman & Hall/CRC, Boca Raton, Fla., 2004.

Jesper Møller, Farzaneh Safavimanesh, and Jakob G. Rasmussen. The cylindrical K-function and poisson line cluster point processes. 2016. Available at arXiv: 1503.07423.

M. Myllymäki, T. Mrkvicka, P. Grabarnik, H. Seijo, and U. Hahn. Global envelope tests for spatial processes. *On arxiv: 1307.0239v4*, 2015.

Antti Penttinen and Dietrich Stoyan. Statistical analysis for a class of line segment processes. *Scandinavian Journal of Statistics*, 16:153–168, 1989.

Ali Hoseinpoor Rafati, Farzaneh Safavimanesh, Karl-Anton Dorph-Petersen, Jakob G. Rasmussen, Jesper Møller, and Jens R. Nyengaard. Detection and spatial characterization of minicolumnarity in the human cerebral cortex. *Journal of Microscopy*, 261, 2016.

B. D. Ripley. Modelling spatial patterns. *Journal of the Royal Statistical Society*, 39:172–212, 1977.

Dietrich Stoyan and Helga Stoyan. *Fractals, Random Shapes and Point Fields: Methods of Geometrical Statistics*. John Wiley and Sons, Chichester, 1994.

Subject index

Notation index