



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



Andreas Dyreborg Christoffersen, Jesper Møller
and Heidi Søgaard Christensen

Modelling columnarity of pyramidal cells in the human cerebral cortex

No. 08, August 2019

Modelling columnarity of pyramidal cells in the human cerebral cortex

Andreas Dyreborg Christoffersen, Jesper Møller
and Heidi Søgaaard Christensen

Department of Mathematical Sciences, Aalborg University

Abstract

For modelling the location of pyramidal cells in the human cerebral cortex we suggest a hierarchical point process in \mathbb{R}^3 . The model consists first of a generalised shot noise Cox process in the xy -plane, providing cylindrical clusters, and next of a Markov random field model on the z -axis, providing either repulsion, aggregation, or both within specified areas of interaction. Several cases of these hierarchical point processes are fitted to two pyramidal cell datasets, and of these a model allowing for both repulsion and attraction between the points seem adequate.

Keywords: Cylindrical K -function; Determinantal point process; Hierarchical point process model; Line cluster point process; Minicolumn hypothesis

1 Introduction and conclusions

The structuring of neurons in the human brain is a subject of great interest since abnormal structures may be linked to certain neurological diseases (see Casanova, 2007; Esiri and Chance, 2006; Casanova et al., 2006; Buxhoeveden and Casanova, 2002). A specific structure that has been extensively studied in the biological literature is the so called ‘minicolumn’ structure of the cells in the cerebral cortex (see Buxhoeveden and Casanova, 2002; Rafati et al., 2016, and references therein). Rafati et al. (2016) characterised these minicolumns as ‘linear aggregates of neurons organised vertically in units that traverse the cortical layer II–VI, and have in humans a diameter of 35 μm to 60 μm and consist typically of 80–100 neurons’.

1.1 Data

In this paper we analyse the structuring of pyramidal cells, which make up approximately 75 % to 80 % of all neurons (Buxhoeveden and Casanova, 2002) and are pyramid shaped cells, where the apical dendrite extends from the top of the pyramid. Specifically, the paper is concerned with two datasets consisting of the locations and orientations of pyramidal cells in a section of the third, respectively, fifth layer

of Brodmann’s fourth area of the human cerebral cortex. Here, each location is a three-dimensional coordinate representing the centre of a pyramidal cell’s nucleolus and each orientation is a unit vector representing the apical dendrite’s position relative to the corresponding nucleolus.

Figure 1 shows the two point pattern datasets of 634 and 548 nucleolus locations which will be referred to as L3 and L5, respectively (for plot of the orientations for L3, see Møller et al., 2019). Note that the observation window W for the cell locations is a rectangular region with side lengths $492.70\text{ }\mu\text{m}$, $132.03\text{ }\mu\text{m}$, and $407.70\text{ }\mu\text{m}$ for L3 and $488.40\text{ }\mu\text{m}$, $138.33\text{ }\mu\text{m}$, and $495.40\text{ }\mu\text{m}$ for L5. Notice also that the nucleolus locations are recorded such that the z -axis is perpendicular to the pial surface of the brain. In accordance to the minicolumn hypothesis, this implies that the minicolumns extend parallel to the z -axis.

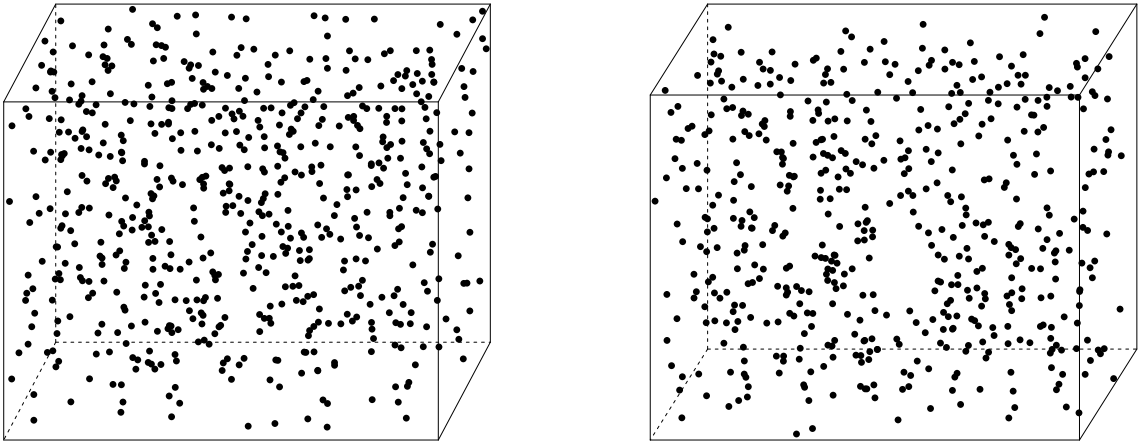


Figure 1: Visualisations of the nucleolus locations for datasets L3 (left) and L5 (right).

1.2 Background and purpose

Møller et al. (2019) found independence between locations and orientations for L3 meaning that the two components may be modelled separately; the same conclusion has afterwards been drawn for L5. As they also found a suitable inhomogeneous Poisson process model for the orientations, and since it is hard by eye to see much structure in the point patterns shown in Figure 1, the focus of this paper is on modelling the nucleolus locations. In particular, we aim at modelling the nucleolus locations for L3 respective L5 by a spatial point process with a columnar structure and discuss to what extent this relates to the minicolumn hypothesis. Note that for the two datasets we use the same notation X for the spatial point process, and we view X as a random finite subset of W .

To the best of our knowledge the so-called Poisson line cluster point process (see Møller et al., 2016) is the only existing point process model for modelling columnar structures. This model was considered by Rafati et al. (2016) in connection to the pyramidal nucleolus locations, but was not fitted to data. For each point pattern considered in the present paper, we notice later that a more advanced model than the Poisson line cluster point process is needed; below we describe such a model for X .

1.3 Hierarchical point process models

Briefly, we consider a hierarchical model for X (further details are given in Sections 3–5), noting that the observation window is a product space, $W = W_{xy} \times W_z$, where W_{xy} is a rectangular region in the xy -plane and W_z is an interval on the z -axis. First, we model the point process X_{xy} given by the projection of X onto the xy -plane; second, conditioned on X_{xy} , we model the vector X_z consisting of the z -coordinates of the points in X . Note that the dimension of X_z agrees with the number of points in X_{xy} and is denoted by n .

1.3.1 The model for X_{xy}

For X_{xy} we consider the restriction of a cluster point process to W_{xy} defined briefly as follows (further details are given in Sections 4–5). Let $\Phi \subset \mathbb{R}^2$ be a stationary point process with intensity $\kappa > 0$, and associate to each point $(\xi, \eta) \in \Phi$ a point process $X_{(\xi, \eta)} \subset \mathbb{R}^3$ that is concentrated around the line in \mathbb{R}^3 which is perpendicular to the xy -plane, with intersection point $(\xi, \eta, 0)$. We refer to $X_{(\xi, \eta)}$ as the cylindrical cluster associated to (ξ, η) . Let $P_{xy}(X_{(\xi, \eta)} \cap W)$ denote the projection onto the xy -plane of the observed part of the cylindrical cluster. For short we refer to the non-empty $P_{xy}(X_{(\xi, \eta)} \cap W)$ as the projected cluster with centre point (ξ, η) . Then we let

$$X_{xy} = \bigcup_{(\xi, \eta) \in \Phi} P_{xy}(X_{(\xi, \eta)} \cap W).$$

Further, conditional on Φ , we assume that the projected clusters are independent and each non-empty $P_{xy}(X_{(\xi, \eta)} \cap W)$ is distributed as the intersection of W_{xy} with a finite planar Poisson process translated by the centre point (ξ, η) ; this Poisson process has intensity function $a\alpha f$, where a is the length of the interval W_z , $\alpha > 0$ is a parameter, and f is the probability density function of a bivariate zero-mean isotropic normal distribution with standard deviation $\sigma > 0$. Thus, ignoring boundary effects, αa is the expected size (or number of points) of a projected cluster and σ controls the spread of points in a projected cluster. Specifically, we let first Φ be a planar stationary Poisson process and second a stationary determinantal point process (Lavancier et al., 2015), since we observe in the first case a very low expected number of points in a projected cluster and because in the second case we want a repulsive model in order to obtain less overlap between the projected clusters.

The special case with Φ a planar stationary Poisson process and X_z a homogeneous binomial point process (that is, the n points in X_z are independent and uniformly distributed on W_z) which is independent of X_{xy} corresponds to a degenerate case of a Poisson line cluster point process as considered in Møller et al. (2016). This becomes clear in Section 4.

1.3.2 The model for X_z conditioned on X_{xy}

We consider several other cases than a homogeneous binomial point process for X_z which is independent of X_{xy} . In general, conditioned on $X_{xy} = \{(x_i, y_i)\}_{i=1}^n$, we propose a Markov random field model, where the conditional probability density

function of X_z is of the form

$$f((z_i)_{i=1}^n | (x_i, y_i)_{i=1}^n) \propto \gamma_1^{s_{B_1, \theta_1}((z_i)_{i=1}^n | (x_i, y_i)_{i=1}^n)} \gamma_2^{s_{B_2, \theta_2}((z_i)_{i=1}^n | (x_i, y_i)_{i=1}^n)} \times \mathbb{I}(\|(x_i, y_i, z_i) - (x_j, y_j, z_j)\| > h \text{ for } 1 \leq i < j \leq n), \quad (1.1)$$

with notation defined as follows. We consider $\{(x_i, y_i, z_i)\}_{i=1}^n$ as a realisation of X , where (x_i, y_i) is the xy -point associated to z_i , the realisation of the i 'th point in X_z (as a technical detail, unless X_z is a binomial point process, (1.1) is not invariant under permutations of z_1, \dots, z_n since we have associated (x_i, y_i) to z_i , so we cannot view (1.1) as the density of a point process where we are conditioning on the number of points). Note that the right hand side in (1.1) is an unnormalised density and e.g. $(z_i)_{i=1}^n$ is short hand notation for (z_1, \dots, z_n) . We let $\mathbb{I}(\cdot)$ be the indicator function. Further, $\gamma_1 > 0$, $\gamma_2 > 0$, and $h \geq 0$ are unknown parameters; if $h > 0$, it is a hard core parameter ensuring a minimum distance h between all pair of points in X ; for the pyramidal cell data it seems natural to include a hard core condition since cells cannot overlap; and when $\gamma_1 = \gamma_2 = 1$ and $h = 0$, the conditional model simply reduces to the homogeneous binomial point process. Furthermore, for $k = 1, 2$,

$$s_{B_k, \theta_k}((z_i)_{i=1}^n | (x_i, y_i)_{i=1}^n) = \sum_{1 \leq i < j \leq n} \mathbb{I}((x_i, y_i, z_i) \in B_k(x_j, y_j, z_j; \theta_k)),$$

where $B_k(x, y, z; \theta_k) \subset \mathbb{R}^3$ is an interaction region, with centre of mass (x, y, z) and a 'size and shape parameter' θ_k , that determines the interaction between points. It is additionally assumed that the hard core ball, given by the three-dimensional closed ball of radius h and centre (x, y, z) does not contain $B_1(x, y, z; \theta_1)$ or $B_2(x, y, z; \theta_2)$. Finally, it is assumed that the symmetry condition

$$(x_i, y_i, z_i) \in B_k(x_j, y_j, z_j; \theta_k) \quad \text{if and only if} \quad (x_j, y_j, z_j) \in B_k(x_i, y_i, z_i; \theta_k)$$

and the disjointness condition

$$B_1(x, y, z; \theta_1) \cap B_2(x, y, z; \theta_2) = \emptyset$$

are satisfied.

These conditions ensure that we can view X_z conditioned on X_{xy} as a Markov random field with second order interactions: for $1 \leq i < j \leq n$, two z -coordinates z_i and z_j interact (in Markov random field terminology, z_i and z_j are neighbours) if and only if $\|(x_i, y_i, z_i) - (x_j, y_j, z_j)\| \leq h$ (that is, the hard core condition is not satisfied, which happens with probability 0) or (x_i, y_i, z_i) lies within the region of interaction of z_j given by the union of $B_1(x_j, y_j, z_j; \theta_1)$ and $B_2(x_j, y_j, z_j; \theta_2)$ (here the symmetry condition is needed to ensure that we can interchange the roles of i and j). The interaction can either cause repulsion/inhibition or attraction/clumping of the points in X depending on whether $\gamma_k < 1$ or $\gamma_k > 1$ for $k = 1, 2$. Thus, apart from the hard core condition, the model allows for both repulsion and attraction but within different interaction regions B_1 and B_2 .

1.3.3 The final hierarchical model and results

At the end of the paper (Section 5) we obtain a satisfactory fit of the following hierarchical model, with the following interpretation of the estimated parameters.

First, the model for X_{xy} is given as in Section 1.3.1 where the centre process Φ is a most repulsive determinantal point process (as detailed in Section 5.1). The parameter estimates are given in Table 1, where the estimated expected cluster size $\widehat{\alpha a}$ is much smaller than expected for a minicolumn when restricting it to the observation window – provided the minicolumn hypothesis is true; cf. personal communication with Jens R. Nyengaard. So we neither claim that we have a fitted model for minicolumns nor that the minicolumn hypothesis is true. Instead we have fitted a model with cylindrical clusters: from Table 1 we see, if $|W_{xy}|$ denotes the area of W_{xy} , the estimated number of projected clusters is $|W_{xy}|\hat{\kappa}$, which is approximately 260 for L3 and 142 for L5; the estimated expected size of a projected cluster is only 2.42 for L3 and 3.87 for L5.

Table 1: Minimum contrast estimates for our final model of X_{xy} (the DLCPP model in Section 5.1) for the datasets L3 and L5.

	$\hat{\kappa}$	$\hat{\sigma}$	$\widehat{\alpha a}$
L3	0.0040	5.45	2.42
L5	0.0021	6.53	3.87

Second, the model of X_z conditioned on X_{xy} has cylindrical interaction regions as illustrated in Figure 2, and (1.1) is the pairwise interaction Markov random field density

$$\begin{aligned}
f((z_i)_{i=1}^n \mid (x_i, y_i)_{i=1}^n) &\propto \prod_{1 \leq i < j \leq n} \mathbb{I}(\|(x_i, y_i, z_i) - (x_j, y_j, z_j)\| > h) \\
&\quad \times \gamma_1^{\mathbb{I}(\|(x_i, y_i) - (x_j, y_j)\| \leq r_1, |z_i - z_j| \leq t_1)} \\
&\quad \times \gamma_2^{\mathbb{I}(\|(x_i, y_i) - (x_j, y_j)\| \leq r_2, t_1 < |z_i - z_j| \leq t_2)},
\end{aligned}$$

where $\gamma_1 > 0$ and $\gamma_2 > 0$ are interaction parameters and $0 < r_2 \leq r_1$ and $0 < t_1 < t_2$ are parameters which determine the ‘range of interaction’ such that $h < \sqrt{t_k^2 + r_k^2}$ for $k = 1, 2$. The restrictions on r_1 , r_2 , t_1 , and t_2 are empirically motivated by use of functional summaries as detailed in Section 5.2. The final fitted model have parameter estimates as displayed in Table 2 where most notably $\hat{\gamma}_1 < 1$ and $\hat{\gamma}_2 > 1$. In particular the final fitted model is in accordance to the empirical findings as noted later when the so-called cylindrical K -function of Figure 3 is discussed: we have modelled repulsion within stunted cylinders (corresponding to B_1) and aggregation within elongated cylinders (corresponding to B_2), see again Figure 2. Moreover, the estimated hard core \hat{h} is greater than 6 μm , which is in accordance with ‘distance between the nucleolus and the membrane of a pyramidal cell’ (personal communication with Jens R. Nyengaard). Note that the hard core ball is much smaller than the interaction region B_1 : $2\hat{h}$ (the diameter of the hard core ball) is about half as small as $2\hat{t}_1$ (the height of B_1). Finally, comparing Tables 1–2, we note that the two ‘clustering parameters’ $2\hat{\sigma}$ and \hat{r}_2 are of the same order.

In conclusion, for each dataset we have fitted a rather complex hierarchical point process model describing columnar structures of the nucleolus locations. This model included repulsion between nucleolus locations given by a hard core condition on a

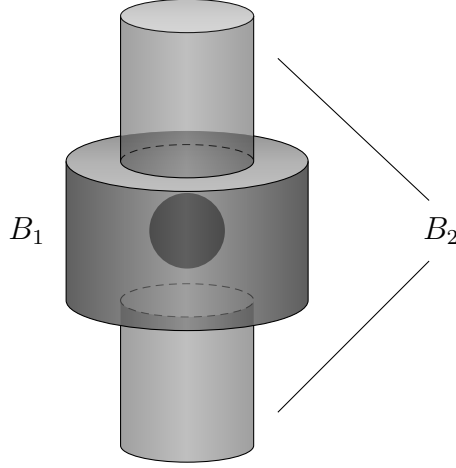


Figure 2: Visualisation of the hard core region ball (in dark) and the cylindrical interaction regions B_1 (the cylinder) and B_2 (the union of the two elongated cylinders) used in our final model for L3.

Table 2: Pseudo likelihood estimates of our final model (model 5 from Table 4 in Section 5.2) for the datasets L3 and L5.

	$\hat{\gamma}_1$	$\hat{\gamma}_2$	\hat{h}	\hat{r}_1	\hat{t}_1	\hat{r}_2	\hat{t}_2
L3	0.41	1.78	6.25	20	11.5	11	35.5
L5	0.51	1.68	6.77	24.25	15.5	14.75	37.25

small scale and a stunted cylindrical interaction region on a larger scale, as well as clustering between nucleolus locations given by an elongated cylindrical interaction region.

1.4 Model fitting

In Møller et al. (2016) parameter estimation for the degenerate PLCPP model was simply done by a moment based procedure which included minimisation of a certain contrast between a theoretical second order moment functional summary and its empirical estimate. In the present paper we use a similar minimum contrast procedure for estimating the parameters of models for X_{xy} . For the models of X_z conditioned on X_{xy} we find it convenient to use a maximum pseudo likelihood procedure as detailed in Section 5.2. Moreover, each fitted model is evaluated by considering informative global extreme rank length (GERL) envelope procedures (Mrkvička et al., 2018; Myllymäki et al., 2017) for various functional summaries.

1.5 Outline

The remainder of this paper explains how we arrive at the final model given in Section 1.3.3 after fitting several other models. In Section 2 we introduce some basic concepts and definitions needed for the models in the subsequent sections. In Section 3 we investigate how the nucleolus locations deviate from complete spatial

randomness (that is, when X is a homogeneous Poisson process), and in Section 4 we also notice a deviation from a fitted degenerate PLCPP model. Finally, in Section 5 we introduce and fit various generalisations of the degenerate PLCPP model as briefly described in Sections 1.3–1.4.

2 Preliminaries

The point processes X , X_{xy} , and X_z introduced above are viewed as the restriction to the bounded sets W , W_{xy} , and W_z of a locally finite point process $Y \subset \mathbb{R}^d$ with $d = 3, 2, 1$, respectively. Briefly speaking, this means that Y is a random subset of \mathbb{R}^d satisfying that $Y_B = Y \cap B$ is finite for any bounded set $B \subset \mathbb{R}^3$; for a more rigorous definition of point processes, see e.g. Daley and Vere-Jones (2003) or Møller and Waagepetersen (2004). Below we recall a few basic statistical tools needed in this paper, using the generic notation Y for a locally finite point process defined on \mathbb{R}^d (apart from the cases above, we have in mind that Y could also be the centre process Φ from Section 1.3).

2.1 Moments

For each integer $n \geq 1$, to describe the n 'th order moment properties of Y , we consider the so-called n 'th order intensity function $\lambda^{(n)} : (\mathbb{R}^d)^n \rightarrow [0, \infty)$ given that it exists. This means that for any pairwise distinct and bounded Borel sets $B_1, \dots, B_n \subset \mathbb{R}^d$,

$$\mathbb{E}[n(Y_{B_1}) \cdots n(Y_{B_n})] = \int_{B_1} \cdots \int_{B_n} \lambda^{(n)}(x_1, \dots, x_n) dx_1 \cdots dx_n$$

is finite, where $n(Y_B)$ denotes the cardinality of Y_B .

The first order intensity function $\lambda^{(1)} = \lambda$ is of particular interest and is simply referred to as the intensity function. Heuristically, $\lambda(u) du$ can be interpreted as the probability of observing a point from Y in the infinitesimal ball of volume du centred at u . If the intensity function $\lambda(\cdot) \equiv \lambda$ is constant, then $\lambda|B| = \mathbb{E}[n(Y_B)]$ for any bounded Borel set $B \subset \mathbb{R}^d$, where $|\cdot|$ is the Lebesgue measure. In this case Y is said to be homogeneous and otherwise inhomogeneous. Clearly, stationarity of Y (meaning that its distribution is invariant under translations in \mathbb{R}^d) implies homogeneity.

2.2 Functional summaries

In order to determine an appropriate model for an observed point pattern, we consider functional summaries, which reflect/summarise different properties of the point pattern and are useful for model fitting and control. The main examples are considered below.

To summarise the second order moment properties, it is custom to consider the pair correlation function (PCF), g , which is defined as the ratio of the second and

first order intensity function, that is,

$$g(x_1, x_2) = \frac{\lambda^{(2)}(x_1, x_2)}{\lambda(x_1)\lambda(x_2)}, \quad x_1, x_2 \in \mathbb{R}^d.$$

Heuristically, $g(x_1, x_2)$ can be interpreted as the probability of simultaneously observing a point from X in each of the two infinitesimal balls of volume dx_1 and dx_2 centred at respectively x_1 and x_2 relative to the probability of independently observing a point in the two infinitesimal balls. The PCF is said to be stationary when (with abuse of notation) $g(x_1, x_2) = g(x_1 - x_2)$ and isotropic when $g(x_1, x_2) = g(\|x_1 - x_2\|)$.

If the PCF is stationary, it is closely related to the so-called second order reduced moment measure, \mathcal{K} , given by

$$\mathcal{K}(B) = \int_B g(x) dx,$$

where $B \subset \mathbb{R}^d$ is a Borel set (see Møller and Waagepetersen, 2004). If Y is stationary and B has centre of mass at the origin of \mathbb{R}^d , then $\lambda\mathcal{K}(B)$ can be interpreted as the expected number of further points falling within B given that Y has a point at the origin; and when considering scalings of B , we refer to B as a structuring element. The simplest example occurs when B is a ball centred at the origin and with radius $r > 0$; then $K(r) = \mathcal{K}(B)$ becomes the K -function introduced by Ripley (1976); and often we instead consider a transformation of the K -function, which is called the L -function and defined by $L(r) = (K(r)/\omega_d)^{1/d}$, where ω_d is the volume of the d -dimensional unit ball. In particular, if Y is a stationary Poisson process, then $L(r) = r$.

For detecting cylindrical structures, Møller et al. (2016) introduced the cylindrical K -function which corresponds to $\mathcal{K}(B)$ when B is a cylinder of height $2t$, base-radius r , and centre of mass at the origin. Note that Ripley's K -function depends only on one argument, r , while the cylindrical K -function depends both on r , t , and the direction of the cylinder. However, when $d = 3$ and since the minicolumns are expected to extend along the z -axis, we only consider cylinders extending in this direction, effectively reducing the number of arguments to two.

We will also consider the commonly used F -, G -, and J -functions when performing model control; see van Lieshout and Baddeley (1996) for definitions. Briefly, if Y is stationary, $F(r)$ is the probability that Y has a point within distance $r > 0$ from a fixed location in \mathbb{R}^d ; $G(r)$ is the probability that Y has another point within distance $r > 0$ from an arbitrary fixed point in Y ; and $J(r) = (1 - G(r))/(1 - F(r))$ when $F(r) < 1$.

The functional summaries will in the following be used both for model fitting as described in Section 2.3 and for model checking using GERL envelope procedures as mentioned in Section 1.4. In the GERL envelope procedure, the distribution of the empirical functional summary under the hypothesis of interest is estimated by simulations. The procedure is a refinement of the global rank envelope procedure (Myllymäki et al., 2017), where it is recommended to use 2499 simulations for a single one-dimensional functional summary and at least 9999 simulations for a single two-dimensional functional summary (Mrkvička et al., 2016). However, we consider

a concatenation of the L -, G -, F -, and J -functions, as well as the cylindrical K -function in which case Mrkvíčka et al. (2017) recommend using more simulations. Particularly for a concatenation of k one-dimensional summary functions they recommend using $k \times 2499$ simulations. We do however have a different setup since we are concatenating both one- and two-dimensional summary functions. For the GERL envelope procedure, Mrkvíčka et al. (2018) suggest that a lower number of simulations may be enough. Therefore, we use 9999 simulations. Since we consider a concatenation of one- and two-dimensional functional summaries, we ensure that each of the functional summaries are weighted equally in the GERL envelope test by evaluating them at the same number of arguments (Mrkvíčka et al., 2017). Specifically we consider 64^2 r -values for each of the L -, G -, F -, and J -functions and a square grid over 64 r -values and 64 t -values for the cylindrical K -function.

2.3 Minimum contrast estimation

For parametric point process models, minimum contrast estimation is a computationally simple fitting procedure introduced by Diggle and Gratton (1984) that is applicable when a closed form expression of a functional summary, T , exists. The idea is to minimise the distance from the theoretical function T to its empirical estimate \hat{T} for the data. Specifically, if T depends on the parameter vector θ and is a function of ‘distance’ $r > 0$ (as for example in case of Ripley’s K -function), the minimum contrast estimate of θ is given by

$$\hat{\theta} = \operatorname{argmin}_{\theta} \int_{r_{\min}}^{r_{\max}} |T(\theta, r)^q - \hat{T}(r)^q|^p dr, \quad (2.1)$$

where $r_{\min} < r_{\max}$, q , and p are positive tuning parameters. General recommendations on q are given in Guan (2009) and Diggle (2014), when $T(r) = g(r)$ or $T(r) = K(r)$. Unless otherwise stated, we let $p = 2$, $q = 1/4$, $r_{\min} = 0$, and r_{\max} be one fourth of the shortest side length of the relevant observation window (the rectangular window W_{xy} in our case).

When the PCF has a closed form expression, alternative estimation procedures can be used, including the second order composite likelihood (see Guan, 2006; Waagepetersen, 2007), adapted second order composite likelihood (see Lavancier et al., 2018), and Palm likelihood (see Ogata and Katsura, 1991; Prokešová et al., 2016; Baddeley et al., 2016).

3 Complete spatial randomness

The most natural place to begin our point pattern analysis is by testing whether a homogeneous Poisson process X with intensity $\lambda > 0$ (we then view Y as a stationary Poisson process with the same intensity), also called complete spatial randomness (CSR), adequately describe each nucleolus point pattern dataset. Recall that this means that $n(X)$ is Poisson distributed with parameter $\lambda|W|$ and conditional on $n(X)$ the points in X are independent and uniformly distributed within W . Even when CSR is not an appropriate model, deviations from the model can be useful for

determining whether the points of a homogeneous point pattern tend to e.g. attract or repel each other.

The CSR model is fully specified by its intensity, which naturally is estimated by $n(X)/|W|$, which is equal to 2.37×10^{-5} for L3 and 1.63×10^{-5} for L5. For this fitted model Figure 3 summarises the results of the GERL envelope procedure based on the concatenation of the L -, G -, F -, J -, and cylindrical K -functions as discussed in Section 2.2. Particularly, the left column depicts the part concerning the empirical functional summaries $\hat{L}(r) - r$, $\hat{G}(r)$, $\hat{F}(r)$, and $\hat{J}(r)$ along with the corresponding 95% envelope. The right column depicts the empirical cylindrical K -function along with the areas at which it falls outside the 95% envelope. It is observed that the empirical functional summaries \hat{L} , \hat{F} , and \hat{J} fall strictly outside the envelope for midrange values of r in a manner that indicates repulsion between points at this range. For small and large r -values the observed point patterns resemble the Poisson process. This behaviour could suggest a kind of clustering, where clusters of points from a Poisson process are somewhat separated. The separation of these clusters seems to be more pronounced for L3 than for L5. Further, in the right column of Figure 3, the empirical cylindrical K -function falls above the upper global rank envelopes for cylinders that have a height larger than approximately $35 \mu\text{m}$ for both datasets and a base radius of approximately $5 \mu\text{m}$ to $15 \mu\text{m}$ for L3 and $5 \mu\text{m}$ to $20 \mu\text{m}$ for L5. Furthermore, the observed cylindrical K -functions falls below the lower 95% GERL envelope for cylinders with a height of approximately $10 \mu\text{m}$ to $30 \mu\text{m}$ and a base radius larger than $5 \mu\text{m}$. Hence, for elongated cylinders extending in the z -direction, we tend to see more points in the data than we expect under CSR, while for stunted cylinders we tend to see fewer points. This seems to be in correspondence with columnar structures where the columns extend in the z -direction.

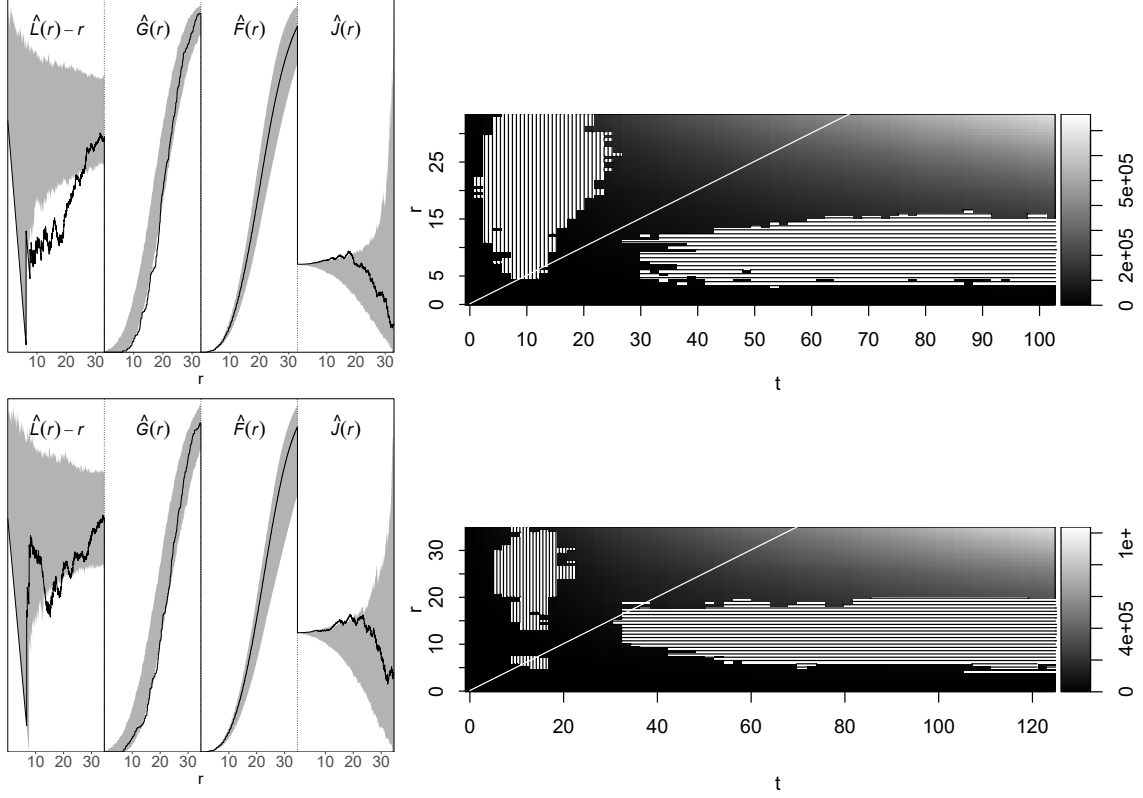


Figure 3: Results of the GERL envelope procedure under CSR based on a concatenation of the empirical L -, G -, F -, J -, and cylindrical K -functions. Left: concatenation of the one-dimensional empirical functional summaries for the data (solid line) together with 95% envelopes (grey region); for ease of visualisation, the functions have been scaled. Right: empirical cylindrical K -function (grey scale) where shaded vertical/horizontal lines indicate that the function falls above/below the 95% envelope. The white line indicates the values for which the cylinder height is equal to the base diameter. Top: results for the dataset L3. Bottom: results for the dataset L5.

4 The degenerate Poisson line cluster point process

Møller et al. (2016) presented the so-called Poisson line cluster point process (PLCPP) which is useful for modelling columnar structures. Specifically, we consider a degenerate PLCPP $Y \subset \mathbb{R}^3$ constructed as follows.

1. Generate a stationary Poisson process $\Phi = \{(\xi_i, \eta_i)\}_{i=1}^\infty \subset \mathbb{R}^2$ with finite intensity $\kappa > 0$. Each point $(\xi_i, \eta_i) \in \Phi$ corresponds to an infinite line l_i in \mathbb{R}^3 which is perpendicular to the xy -plane, that is, $l_i = \{(\xi_i, \eta_i, z) \mid z \in \mathbb{R}\}$.
2. Conditional on Φ , generate independent stationary Poisson processes $L_1 \subset l_1, L_2 \subset l_2, \dots$ with identical and finite intensity $\alpha > 0$.
3. Generate point processes $X_1, X_2, \dots \subset \mathbb{R}^3$ by independently displacing the points of L_1, L_2, \dots by the zero-mean isotropic normal distribution with standard deviation $\sigma > 0$.
4. Finally, set $Y = \bigcup_{i=1}^\infty X_i$ and $X = Y_W$.

Some comments to the construction in items 1–4 are in order.

In the general definition of the PLCPP in Møller et al. (2016), the lines l_1, l_2, \dots are modelled as a stationary Poisson line process. That is, the lines are not required to be perpendicular to the xy -plane nor does the Poisson line process need to be degenerate (meaning that the lines are not required to be mutually parallel). Further, the dispersion density (used in item 3) can be arbitrary. However, the construction is still such that Y becomes stationary. Furthermore, it turns out that it does not matter whether we consider a three-dimensional normal distribution for displacements in item 3 or a bivariate normal distribution with displacements of the xy -coordinates for the points of L_1, L_2, \dots .

Returning to the degenerate PLCPP of items 1–4, we imagine that each X_i is a cylindrical cluster of points around the line l_i , where these cylindrical clusters are parallel to the z -axis. Furthermore, the interpretation of the parameters κ , α , and σ in terms of a Poisson cluster point process is similar to that in Section 1.3.1 except that we now also consider lines not intersecting W : if Y as defined by items 1–4 is restricted to a subset $S \subset \mathbb{R}^3$ bounded by two planes parallel to the xy -plane, for specificity $S = \{(x, y, z) \in \mathbb{R}^3 \mid z \in W_z\}$, this restricted point process can be seen as a (modified) Thomas process (see Thomas, 1949; Møller and Waagepetersen, 2004) on \mathbb{R}^2 along with independent z -coordinates following a uniform distribution on W_z .

To see this, first note that conditional on $\Phi = \{(\xi_i, \eta_i)\}_{i=1}^\infty$ and for all $i = 1, 2, \dots$, X_i is a Poisson process in \mathbb{R}^3 with intensity function $\lambda_i((x, y, z)) = \alpha f(x - \xi_i, y - \eta_i)$, where f is the probability density function of the bivariate isotropic normal distribution given in item 3. In turn, this implies that Y conditioned on Φ is a Poisson process in \mathbb{R}^3 with intensity function $\sum_{i=1}^\infty \lambda_i((x, y, z))$. Further, since $\lambda_i(x, y, z) = \lambda_i(x, y)$ does not depend on z for all $i = 1, 2, \dots$, the projection of Y_S onto the xy -plane, $P_{xy}(Y_S)$, conditioned on Φ is a Poisson process with intensity $a \sum_{i=1}^\infty \lambda_i(x, y)$, where a is the length of the interval W_z . Since Φ is a stationary Poisson process, $P_{xy}(Y_S)$ is a Thomas process with centre process intensity κ and expected cluster

size αa (that is, the expected number of points in $X_i \cap S$). Finally, from items 2–4 it follows that the z -coordinates of X_z are independent and uniformly distributed on W_z , and they are independent of X_{xy} .

Consequently, simulating $X = Y_W$ is straightforwardly done by simulating a Thomas point process (on a larger set than W_{xy} in order to avoid boundary effects) along with independent uniform z -coordinates on W_z . For simulating the Thomas point process we apply standard software from the R-package `spatstat` (Baddeley et al., 2016). Similarly, fitting a degenerate PLCPP to a realisation of X is simply a matter of fitting a Thomas process to the point pattern consisting of the xy -coordinates of the points in that realisation. Since the K -function of the Thomas process has a closed form expression, the model can be fitted using minimum contrast estimation with $T(r) = K(r)$ in (2.1). Table 3 summarises the parameter estimates, where most notably the expected cluster size $\widehat{\alpha a}$ is < 1 for both L3 and L5. Understanding each cylindrical cluster within W as (a part of) a minicolumn, ‘these parameter estimates result in very unnatural models for the datasets, since each minicolumn within W is expected to consist of less than one point’ (personal communication with Jens R. Nyengaard).

Table 3: Minimum contrast estimates of the degenerate PLCPP.

	$\hat{\kappa}$	$\hat{\sigma}$	$\widehat{\alpha a}$
L3	0.027	2.86	0.36
L5	0.0085	4.58	0.95

Despite the fact that the fitted degenerate PLCPP models are somewhat unnatural and hardly can be interpreted as a model with (hypothesised) minicolumnar structures, GERL envelope procedure based on a concatenation of the F -, G -, and J -functions show that the Thomas process suitably fit the projected locations with a p -value of 0.76 for L3 and 0.87 for L5. However, results from the concatenated GERL envelope procedure described in Section 2.2 indicated that the model did not suitably describe the three-dimensional nucleolus locations with a p -value of 10^{-4} for both L3 and L5. Specifically, Figure 4 shows the empirical cylindrical K -function and indicates where it deviates from the 95% envelope. Clearly, the model does account for some of the columnarity of the data as opposed to CSR, but the empirical cylindrical K -function for L3 still falls above the 95% envelope. Furthermore, the empirical cylindrical K -function for both datasets falls below the 95% envelope similar to what was seen under CSR, indicating a lack of regularity, which in fact is supported by the one-dimensional functional summaries (not shown). This could suggest that the cylindrical clusters should be more distinct; motivating us to generalise the degenerate PLCPP model as in the following section.

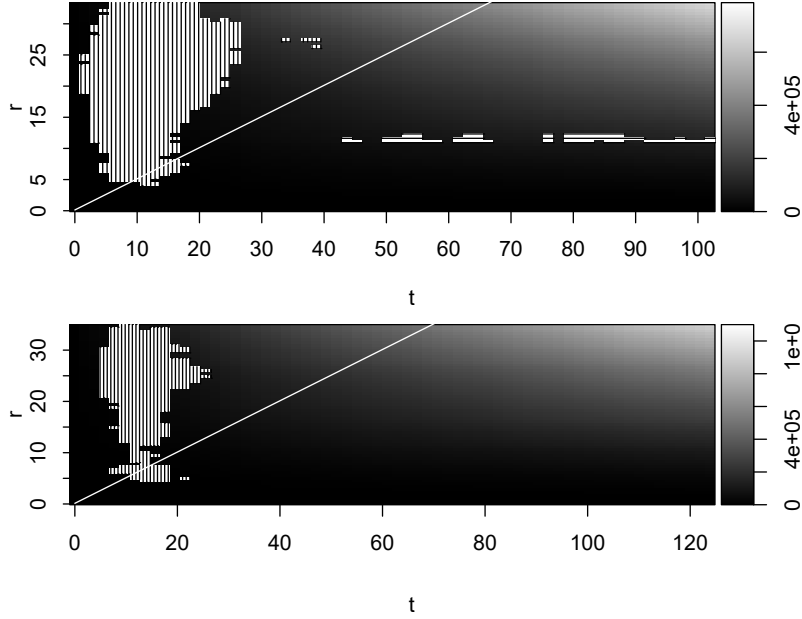


Figure 4: Empirical estimates of the cylindrical K -function (grey scale) where shaded vertical/horizontal lines indicate that the function falls above/below the 95% GERL envelope under the fitted degenerate PLCPP and based on the concatenation described in Section 2.2. The white line indicates the values for which the cylinder height is equal to the base diameter. Top: results for the dataset L3. Bottom: results for the dataset L5.

5 A generalisation of the degenerate PLCPP

As some but not all features of the data were explained by the degenerate PLCPP fitted in Section 4, we propose in this section two generalisations as follows.

1. The centre process Φ is a planar stationary point process.
2. X_z conditioned on X_{xy} follows a Markov random field model.

The first modification is straightforward and for this specific application we chose a repulsive centre process to obtain more distinguishable cylindrical clusters; this is detailed in Section 5.1. Further, the assumption of stationarity of Φ is made in order to apply a similar minimum contrast estimation procedure as in Section 4, so implicitly we make the assumption that the PCF or the K -function is expressible on closed form. For the second modification we suggest a conditional model inspired by the multiscale point process and particularly the Strauss hard core point process (see e.g. Møller and Waagepetersen, 2004) which will allow for further repulsion or even aggregation between the points; this is detailed in Section 5.2.

5.1 A determinantal point process model for the centre points

Consider a point process $Y \subset \mathbb{R}^3$ specified by items 1–4 in Section 4 with the exception that the centre process Φ now is an arbitrary stationary planar point

process. Then, recalling the notation from Section 4, $P_{xy}(Y_S)$ is a planar Cox process (see Møller and Waagepetersen, 2004) and even a planar generalised shot-noise Cox process (see Møller and Torrisi, 2005) driven by the random intensity function $\Lambda(x, y) = a \sum_{i=1}^{\infty} \lambda_i(x, y)$ for $(x, y) \in \mathbb{R}^2$. Moreover, $P_{xy}(Y_S)$ corresponds to the Thomas process, but with a different centre point process (unless of course Φ is a stationary Poisson process).

In this section we focus on the case where Φ is a stationary determinantal point process (DPP; see Lavancier et al., 2015), in which case we will refer to Y as the determinantal line cluster point process (DLCPP). A DPP is defined in terms of its n 'th order intensity function for $n = 1, 2, \dots$: let $C : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{C}$ be a function and $\lambda^{(n)}$ the n 'th order intensity function of Φ , then Φ is called a DPP with kernel C if

$$\lambda^{(n)}(x_1, \dots, x_n) = \det[C](x_1, \dots, x_n) \quad \text{for } n = 1, 2, \dots, x_1, \dots, x_n \in \mathbb{R}^2,$$

where $\det[C](x_1, \dots, x_n)$ is the determinant of the $n \times n$ matrix with (i, j) 'th entry $C(x_i, x_j)$. For further details on DPPs, we refer to Lavancier et al. (2015) and the references therein. When Φ is a DPP, we call $P_{xy}(Y_S)$ a determinantal Thomas point process (DTPP). The DTPP is discussed to some extent in Møller and Christoffersen (2018), where a closed form expression of its PCF is found. Thus, the DLCPP can be fitted by fitting a DTPP to the projected data using a minimum contrast procedure (see Section 2.3).

For our data we want to obtain a DLCPP with as much repulsion as possible between the centre lines of the cylindrical clusters. Therefore, we let Φ be the ‘most repulsive DPP’ (in the sense of Lavancier et al., 2015), which is the jinc-like DPP given by the kernel $C(x_1, x_2) = \sqrt{\rho/\pi} J_1(2\sqrt{\pi\rho}\|x_1 - x_2\|) / \|x_1 - x_2\|$, where J_1 is the first order Bessel function of the first kind and $\|\cdot\|$ denotes the usual planar distance (for more information on this particular DPP, see Lavancier et al., 2018; Biscio and Lavancier, 2016).

Simulation of the DTPP is done by first simulating a DPP with intensity κ (on a larger region than W_{xy} in order to avoid boundary effects), for which we use the functionality of **spatstat**, then secondly generating for each cluster a Poisson distributed number of points with intensity αa , and finally displacing these points by a bivariate zero-mean isotropic normal distribution.

The parameter estimates of the jinc-like DTPP model were obtained by minimum contrast with $T(r) = g(r)$; see Table 1 for the results and the accompanying discussion in Section 1.3.3. Despite the expectation under the minicolumn hypothesis of having much higher values of $\widehat{\alpha a}$ than in Table 1 (see again Section 1.3.3), simulations of the fitted jinc-like DPP in the xy -plane seem in reasonable correspondence to the projected data; see Figure 5. Furthermore, results from the GERL envelope procedure based on a concatenation of the F -, G -, and J -functions do not provide any evidence against the jinc-like DPP model for the projected points with p -values of 0.67 for L3 and 0.83 for L5.

Since the jinc-like DTPP model fits the projected data well, we proceeded and added independent uniform z -coordinates on W_z to the simulations, thereby obtaining simulations of the jinc-like DLCPP. Figure 6 summarises the result of the 95% GERL test based on the concatenation of functional summaries as described in Section 2.2. The left column depicts the part of the one-dimensional functional

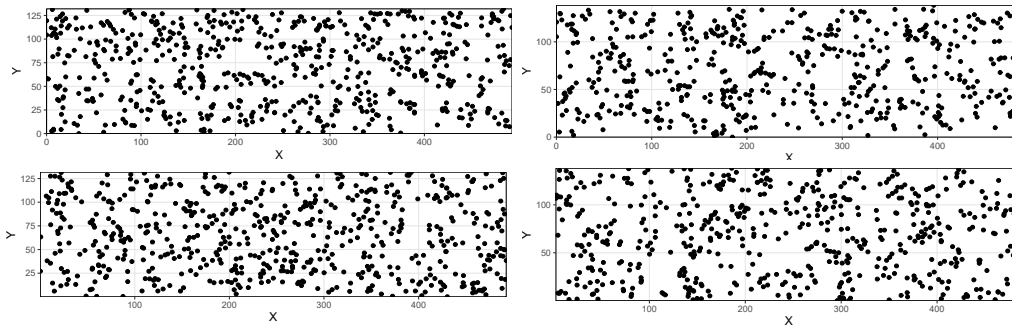


Figure 5: Projection of observed nucleolus locations onto the xy -plane (left) and simulations from the fitted jinc-like DTPP (right) for the datasets L3 (top) and L5 (bottom).

summaries along with 95% envelopes, while the right column shows the empirical cylindrical K -function along with shaded regions that indicate where it deviates from the corresponding envelope. These plots show that the models do not account for the regularity of the data. This leads us to our next generalisation in Section 5.2.

5.2 A Markov random field model for the z -coordinates

Motivated by the observations at the end of the previous section, in this section we propose to model the z -coordinates conditioned on the xy -coordinates by a pairwise interaction point process as given in (1.1). Thereby, our hierarchical model construction yields a more flexible model for X but we ignore edge effects in the sense that we have only specified a model for first $P_{xy}(Y_S)$ and second X_z conditioned on $X_{xy} = P_{xy}(Y_S) \cap W_{xy}$, thereby ignoring a possible influence of points in $Y \setminus W$ when (1.1) is used in the latter step (unless it specifies a binomial point process). This simplification is just made for mathematical convenience; indeed it would be interesting to construct a model taking edge effects into account so that Y becomes stationary, but we leave this challenging issue for future research. Below we first specify the ingredients of the conditional probability density function given in (1.1) for various models and discuss the overall conclusions, next describe how to find parameter estimates, and finally discuss how well the estimated models fit the data. Note that although we have not specified a stationary model for Y , it may still make sense to interpret plots of empirical cylindrical K -functions and \hat{F} , \hat{G} , \hat{J} , and \hat{L} -functions, since we have stationarity in the xy -plane and approximately stationarity in the z -direction (as the density (1.1) is invariant under ‘translations of (z_1, \dots, z_n) within W_z ’).

In our search for a suitable model for the nucleolus locations, we considered many special cases of (1.1). Table 4 summarises five selected models, where $b((x, y, z); r)$ is the ball with centre (x, y, z) and radius r , and where $c((x, y, z); r, t)$ and $d((x, y, z); r, t)$ denote the cylinder and double cone, respectively, with centre of mass at (x, y, z) , height $2t$, base radius r , and extending in the z -direction. First, we considered model 1 which is a hard core model if $h > 0$ and one of the simplest ways of modelling regularity; note that model 1 with $h = 0$ is the binomial point process with a uniform density as considered in Section 4. Though accounting for small distance repulsion, when fitted to the data, model 1 turned out not to account for

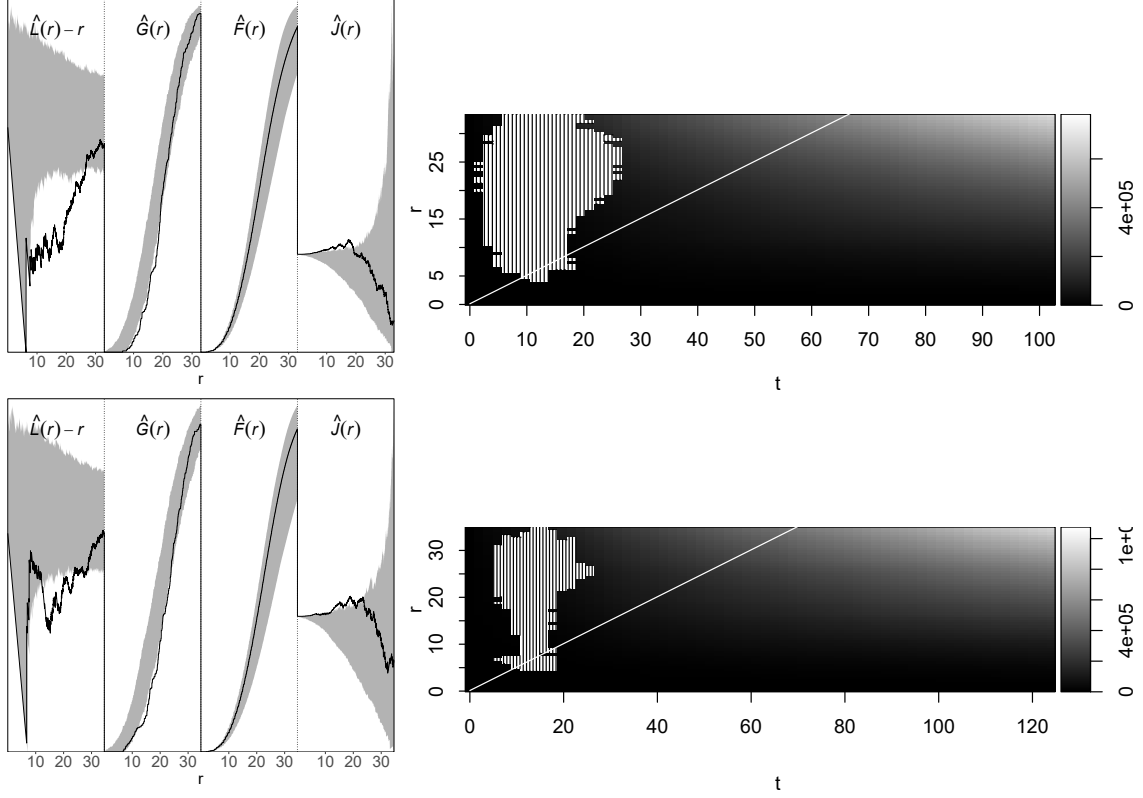


Figure 6: Results of the GERL envelope procedure under the fitted DLCPP based on a concatenation of the empirical L -, G -, F -, J -, and cylindrical K -functions. Left: concatenation of the one-dimensional empirical functional summaries for the data (solid line) together with 95% envelopes (grey region); for ease of visualisation, the functions have been scaled. Right: empirical cylindrical K -function (grey scale) where shaded vertical/horizontal lines indicate that the function falls above/below the 95% envelope. The white line indicates the values for which the cylinder height is equal to the base diameter. Top: results for the dataset L3. Bottom: results for the dataset L5.

the repulsion at larger scales. Second, we considered model 2 which is a conditional Strauss model with a hard core condition (see Møller and Waagepetersen, 2004, and the references therein). For this model the scale of repulsion for the z -coordinates seemed too great for points with similar xy -coordinates, and therefore we found it natural to replace the spherical interaction region with a cylinder, yielding model 3. However, model 3 did not correct the problem, and continuing with a single region of interaction we next suggested model 4 with a region given by a cylinder minus a double cone. Model 4 does to a smaller degree penalise the occurrence of points with similar xy -coordinates. However, this model was not suitable either. Models 1–4 were discarded by GERL tests with extremely small p -values. Finally, we considered model 5 which is a more flexible model that allows for both repulsion and aggregation within cylinder shaped interaction regions, cf. the discussion in Section 1.3.3. For simplicity all the models were also considered without a hard core condition, that is $h = 0$, but was in every case found inadequate.

Table 4: Specific choices of the parameters $\gamma_1, \gamma_2, \theta_1, \theta_2$ and the interaction regions $B_1(\cdot; \theta_1), B_2(\cdot; \theta_2)$ for five models given by the density (1.1). For each model, a hard core parameter $h \geq 0$ is included. Apart from the specified restrictions, it is required for models 2–5 that $B_1(\cdot; \theta_1) \not\subseteq b(\cdot; h)$ (for model 2 this means that $r > h$ as already indicated) and in addition for model 5 that $B_2(\cdot; \theta_2) \not\subseteq b(\cdot; h)$ where $\theta_2 = (r_2, t_2)$ with $r_1 \geq r_2 > 0$ and $t_2 > t_1$.

Model	γ_1	γ_2	$B_1(\cdot; \theta_1)$	$B_2(\cdot; \theta_2)$	θ_1
1	1	1	\emptyset	\emptyset	–
2	>0	1	$b(\cdot; r)$	\emptyset	$r > h$
3	>0	1	$c(\cdot; r, t)$	\emptyset	$r, t > 0$
4	>0	1	$c(\cdot; r, t) \setminus d(\cdot; r, t)$	\emptyset	$r, t > 0$
5	>0	>0	$c(\cdot; r_1, t_1)$	$c(\cdot; r_2, t_2) \setminus c(\cdot; r_1, t_1)$	$r_1, t_1 > 0$

The likelihood function corresponding to (1.1) involves a normalising constant which needs to be approximated by Markov chain Monte Carlo methods. We propose an easier alternative based on the pseudo likelihood function (Besag, 1975) defined as follows when the data is given by $\{(x_i, y_i, z_i)\}_{i=1}^n \subset W$. For $i = 1, \dots, n$, the i 'th full conditional density associated to (1.1) is

$$\begin{aligned}
f(z_i \mid (z_1, \dots, z_{i-1}, z_{i+1}, \dots, z_n), (x_j, y_j)_{j=1}^n) \\
= \mathbb{I}(\|(x_i, y_i, z_i) - (x_j, y_j, z_j)\| > h \text{ for } j \neq i) \gamma_1^{s_{1,i}} \gamma_2^{s_{2,i}} / c_i
\end{aligned} \tag{5.1}$$

where we define

$$s_{k,i} = \sum_{j: j \neq i} \mathbb{I}((x_j, y_j, z_j) \in B_k((x_i, y_i, z_i); \theta_k)), \quad k = 1, 2,$$

and where the normalising constant is given by

$$\begin{aligned}
c_i = & \sum_{k=0}^{n-1} \sum_{l=0}^{n-1} \gamma_1^k \gamma_2^l \int_{W_z} \mathbb{I}(\|(x_i, y_i, z) - (x_j, y_j, z_j)\| > h \text{ for } j \neq i) \\
& \times \mathbb{I}\left(\sum_{j:j \neq i} \mathbb{I}((x_j, y_j, z_j) \in B_1((x_i, y_i, z); \theta_1)) = k\right) \\
& \times \mathbb{I}\left(\sum_{j:j \neq i} \mathbb{I}((x_j, y_j, z_j) \in B_2((x_i, y_i, z); \theta_2)) = l\right) dz.
\end{aligned}$$

To estimate the model parameters we maximise the log pseudo likelihood given by

$$\begin{aligned}
LP(\gamma_1, \gamma_2, h, \theta_1, \theta_2) \\
= \sum_{i=1}^n \log f(z_i | (z_1, \dots, z_{i-1}, z_{i+1}, \dots, z_n), (x_j, y_j)_{j=1}^n). \tag{5.2}
\end{aligned}$$

Clearly, by (5.1) the maximum pseudo likelihood estimate (MPLE) \hat{h} of h is the minimum distance between any distinct pair of points (x_i, y_i, z_i) and (x_j, y_j, z_j) in the data. This in fact also corresponds to the maximum likelihood estimate. For $h = \hat{h}$ and for fixed θ_1 and θ_2 , we easily obtain the following. For each of models 2–4, the MPLE of γ_1 exists if and only if $s_{1,i} \neq 0$ for some i , and then the log pseudo likelihood function is strictly concave with respect to $\log \gamma_1$. For model 5, the MPLE of (γ_1, γ_2) exists if and only if $s_{1,i} \neq 0$ for some i and $s_{2,j} \neq 0$ for some j , and then the log pseudo likelihood function is strictly concave with respect to $(\log \gamma_1, \log \gamma_2)$. Therefore, the (profile) log pseudo likelihood can be maximised by a combination of a grid search over θ_1 and θ_2 and numerical optimisation with respect to γ_1 and γ_2 . Table 2 shows the maximum pseudo likelihood estimates of model 5 for the two datasets, where for the numerical optimisation we used `optim` (a general-purpose optimisation function from the R-package `stats`).

Each of the five models in Table 4 were fitted to L3 and L5 by finding the maximum pseudo likelihood estimate, and model checking was performed using GERL envelope procedures based on the concatenation of functional summaries as discussed in Section 2.2. For the fitted models, model 5 was the most appropriate with p -values of 0.34 for L3 and 0.03 for L5 when using the GERL envelope procedure; the 95% GERL envelope is visualised in Figure 7. Thus no evidence is seen against the fitted models summarised in Table 2 for L3 while only slight evidence is present for L5. We note that for both datasets the fitted models are such that B_1 is a stunted cylinder and models repulsion since $\hat{\gamma}_1 < 1$, while $c(\cdot, r_2, t_2)$ is elongated and B_2 models aggregation, since $\hat{\gamma}_2 > 1$. Hence, when standing in some point $(x_1, y_1, z_1) \in X$ it is less likely to observe a z -coordinate if the corresponding xy -coordinates are similar to (x_1, y_1) . Specifically, if (x_1, y_1) and (x_2, y_2) lies within distance 20 μm for L3 and 24.25 μm for L5, it is less likely to observe a z -coordinate z_2 (associated to (x_2, y_2)) with $|z_1 - z_2|$ less than 11.5 μm for L3 and 15.5 μm for L5. Analogously, given that (x_1, y_1) and (x_2, y_2) lies within distance 11 μm for L3 and 14.75 μm L5, it is more likely to observe z_2 if $|z_1 - z_2|$ is in the interval from 11.5 μm to 35.5 μm for L3 or from 15.5 μm to 37.25 μm for L5.

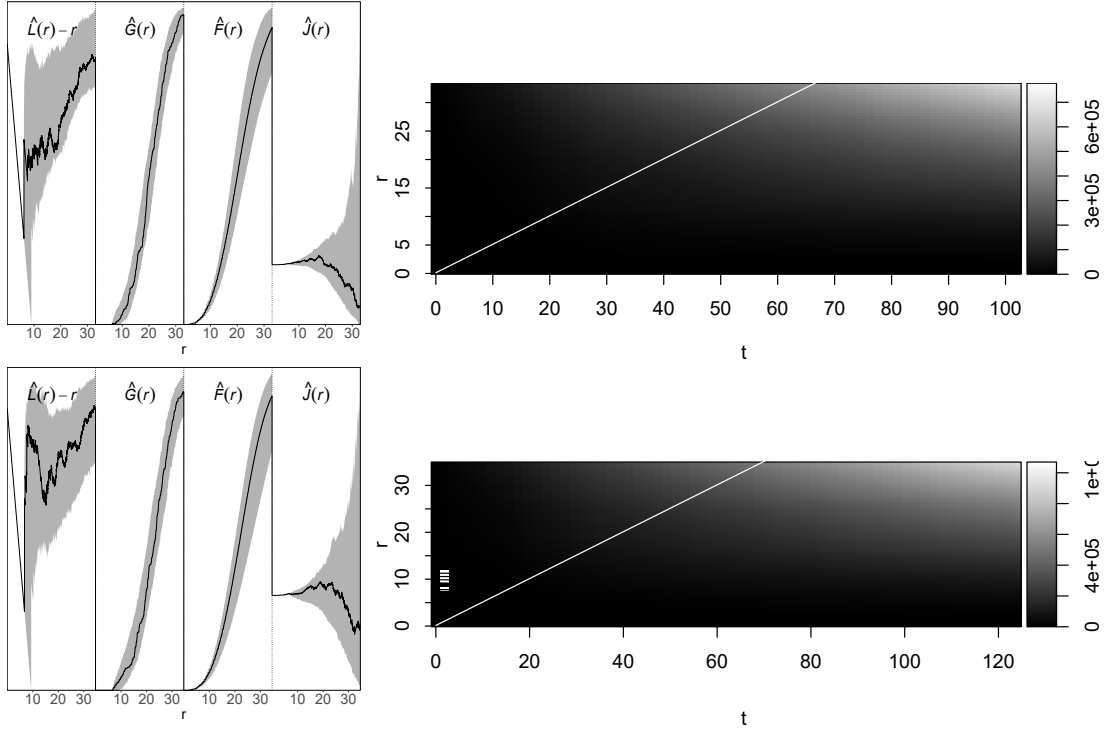


Figure 7: Results of the GERL envelope procedure under the fitted model 5 based on a concatenation of the empirical L -, G -, F -, J -, and cylindrical K -functions. Left: concatenation of the one-dimensional empirical functional summaries for the data (solid line) together with 95% envelopes (grey region); for ease of visualisation, the functions have been scaled. Right: empirical cylindrical K -function (grey scale) where shaded vertical/horizontal lines indicate that the function falls above/below the 95% envelope. The white line indicates the values for which the cylinder height is equal to the base diameter. Top: results for the dataset L3. Bottom: results for the dataset L5.

Finally, note that simulations from each of models 1–5 can straightforwardly be obtained using a Metropolis-Hastings algorithm for a fixed number of points and given a realisation of the xy -coordinates. Specifically, we used (Algorithm 7.1 in Møller and Waagepetersen, 2004) but with a systematic updating scheme cycling over the point indexes 1 to n , using a uniform proposal for a new point in W_z and a Hastings ratio calculated from the full conditional (5.1). We successively updated each point 100 times under the systematic updating scheme, corresponding to 63400 and 54800 point updates for L3 and L5, respectively.

Acknowledgements

This work was supported by The Danish Council for Independent Research | Natural Sciences, grant DFF – 7014-00074 ‘Statistics for point processes in space and beyond’, and by the ‘Centre for Stochastic Geometry and Advanced Bioimaging’, funded by grant 8721 from the Villum Foundation. We are thankful to Ali H. Rafati for collecting the data analysed in this paper and to Jens R. Nyengaard and Ninna Vihrs for helpful comments.

References

- Baddeley, A., Rubak, E., and Turner, R. (2016). *Spatial Point Patterns: Methodology and Applications with R*. Chapman & Hall/CRC, New York.
- Besag, J. (1975). Statistical analysis of non-lattice data. *The Statistician*, 24:179–195.
- Biscio, C. A. N. and Lavancier, F. (2016). Quantifying repulsiveness of determinantal point processes. *Bernoulli*, 22:2001–2028.
- Buxhoeveden, D. P. and Casanova, M. F. (2002). The minicolumn hypothesis in neuroscience. *Brain*, 125:935–951.
- Casanova, M. F. (2007). Schizophrenia seen as a deficit in the modulation of cortical minicolumns by monoaminergic systems. *International Review of Psychiatry*, 19:361–372.
- Casanova, M. F., van Kooten, I. A. J., Switala, A. E., van Engeland, H., Heinsen, H., Steinbusch, H. W. M., Hof, P. R., Trippe, J., Stone, J., and Schmitz, C. (2006). Minicolumnar abnormalities in autism. *Acta Neuropathologica (Berl)*, 112:287–303.
- Daley, D. J. and Vere-Jones, D. (2003). *An Introduction to the Theory of Point Processes. Volume I: Elementary Theory and Methods*. Springer-Verlag, New York, second edition.
- Diggle, P. J. (2014). *Statistical Analysis of Spatial and Spatio-temporal Point Patterns*. Chapman & Hall/CRC, Boca Raton, Florida.
- Diggle, P. J. and Gratton, R. J. (1984). Monte Carlo methods of inference for implicit statistical models (with discussion). *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 46:193–227.

- Esiri, M. M. and Chance, S. A. (2006). Vulnerability to Alzheimer’s pathology in neocortex: the roles of plasticity and columnar organization. *Journal of Alzheimer’s Disease*, 9:79–89.
- Guan, Y. (2006). A composite likelihood approach in fitting spatial point process models. *Journal of the American Statistical Association*, 101:1502–1512.
- Guan, Y. (2009). A minimum contrast estimation procedure for estimating the second-order parameters of inhomogeneous spatial point processes. *Statistics and Its Interface*, 2:91–99.
- Lavancier, F., Møller, J., and Rubak, E. (2015). Determinantal point process models and statistical inference. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 77:853–877.
- Lavancier, F., Poinas, A., and Waagepetersen, R. P. (2018). Adaptive estimating function inference for non-stationary determinantal point processes. Available on arXiv:1806.06231.
- van Lieshout, M. N. M. and Baddeley, A. J. (1996). A nonparametric measure of spatial interaction in point patterns. *Statistica Neerlandica*, 50:344–361.
- Møller, J., Christensen, H. S., Cuevas-Pacheco, F., and Christoffersen, A. D. (2019). Structured space-sphere point processes and K -functions. *Methodology and Computing in Applied Probability*. Available at <https://doi.org/10.1007/s11009-019-09712-w>.
- Møller, J. and Christoffersen, A. D. (2018). Pair correlation functions and limiting distributions of iterated cluster point processes. *Journal of Applied Probability*, 55:789–809.
- Møller, J., Safavimanesh, F., and Rasmussen, J. G. (2016). The cylindrical K -function and Poisson line cluster point process. *Biometrika*, 103:937–954.
- Møller, J. and Torrisi, G. L. (2005). Generalised shot noise Cox processes. *Advances in Applied Probability*, 37:48–74.
- Møller, J. and Waagepetersen, R. P. (2004). *Statistical Inference and Simulation for Spatial Point Processes*. Chapman & Hall/CRC, Boca Raton, Florida.
- Mrkvička, T., Myllymäki, M., and Hahn, U. (2017). Multiple Monte Carlo testing, with applications in spatial point processes. *Statistics and Computing*, 27:1239–1255.
- Mrkvička, T., Myllymäki, M., Jílek, M., and Hahn, U. (2018). A one-way ANOVA test for functional data with graphical interpretation. Available on arXiv:1612.03608.
- Mrkvička, T., Soubeyrand, S., Myllymäki, M., Grabarnik, P., and Hahn, U. (2016). Monte Carlo testing in spatial statistics, with applications to spatial residuals. *Spatial Statistics*, 18:40–53.
- Myllymäki, M., Mrkvička, T., Grabarnik, P., Seijo, H., and Hahn, U. (2017). Global envelope tests for spatial processes. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 79:381–404.
- Ogata, Y. and Katsura, K. (1991). Maximum likelihood estimates of the fractal dimension for random spatial patterns. *Biometrika*, 78:463–474.

- Prokešová, M., Dvořák, J., and Jensen, E. B. V. (2016). Two-step estimation procedures for inhomogeneous shot-noise Cox processes. *Annals of the Institute of Statistical Mathematics*, 69:1–30.
- Rafati, A. H., Safavimanesh, F., Dorph-Petersen, K.-A., Rasmussen, J. G., Møller, J., and Nyengaard, J. R. (2016). Detection and spatial characterization of minicolumnarity in the human cerebral cortex. *Journal of Microscopy*, 261:115–126.
- Ripley, B. D. (1976). The second-order analysis of stationary point processes. *Journal of Applied Probability*, 13:255–266.
- Thomas, M. (1949). A generalization of Poisson’s binomial limit for use in ecology. *Biometrika*, 36:18–25.
- Waagepetersen, R. P. (2007). An estimating function approach to inference for inhomogeneous Neyman-Scott processes. *Biometrics*, 63:252–258.