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

I.T. Andersen, U. Hahn, E.C. Arnspang, L.N. Nejsum and E.B.V. Jensen

Double Cox cluster processes – with applications to photoactivated localization microscopy

No. 09, October 2017

Double Cox cluster processes

– with applications to photoactivated localization microscopy

I.T. Andersen[1], U. Hahn[1], E.C. Arnspang[2,3], L.N. Nejsum[2,3] and E.B.V. Jensen[1]

[1]Department of Mathematics, Aarhus University, Denmark
[2]Department of Clinical Medicine, Aarhus University, Denmark
[3]Associated with the Interdisciplinary Nanoscience Center (iNANO), Aarhus University, Denmark

Abstract

Photoactivated localization microscopy (PALM) is an ingenious super-resolution imaging technique that produces 2D point patterns of proteins. Individual proteins may appear as small artificial clusters of points, due to multiple blinking of individual fluorophores. The proteins may also cluster together, and in such cases a pertinent model for a PALM point pattern describes clustering at two different scales. Despite the importance of the imaging technique, statistical methods for analyzing PALM data have remained relatively under-studied. In the present paper, we develop a model-based framework for analysis of PALM data. We focus on a subclass of independent cluster processes, denoted *double Cox cluster processes* (DCCPs), for which both the parent process (of proteins) and the observed process are Cox cluster processes. Parametric models for DCCPs with a Neyman-Scott process as parent process are developed together with statistical inference procedures, based on moment methods. To illustrate the proposed methodology, we analyze a data set from a PALM acquisition. In contrast to earlier model-free methods, the analysis provides information, directly relating to the performance of the proteins. The paper also represents an independent contribution to point process theory.

Keywords: double Cox cluster process, independent clustering, moment-based inference, photoactivated localization microscopy, point process

1 Introduction

PALM is an ingenious super-resolution imaging technique by means of which proteins may be localized with a precision down to 20 nm, even though the resolution of the microscope is limited to 250 nm, due to the diffraction limit. During the recording of the image, PALM uses that fluorescent proteins will emit light at separate

Corresponding author: I.T. Andersen, ita@math.au.dk

times and their positions thereby become resolvable also in cases where the proteins are closer than the conventional resolution limit of the microscope.

PALM was introduced in the seminal paper Betzig et al. (2006), published in *Science*, and two years later the technique was the "Method of the Year" in *Nature Methods*. In 2014, the Nobel Prize in Chemistry was awarded to Eric Betzig, William E. Moerner and Stefan Hell for the development of super-resolved fluorescence microscopy, which brought "optical microscopy into the nanodimension".

The raw data in PALM is a list of 2D coordinates of the observed fluorescent proteins. Individual proteins may fluoresce multiple times at slightly varying positions so each protein gives rise to a small cluster. The challenge is to make inference on the interactions between the proteins as well as on the multiple appearances of each protein.

Despite of the importance of the PALM imaging technique, methods for analyzing the resulting point patterns for properties such as clustering have remained relatively under-studied (Rubin-Delanchy et al., 2015). Methods based on pair correlation analysis have been developed by Sengupta et al. (2011, 2013); Veatch et al. (2012). However, these are model-free methods in the sense that they are not based on a point process model for the proteins and the estimated parameters do not directly relate to the performance of the proteins. In Rubin-Delanchy et al. (2015), a model-based Bayesian cluster algorithm is developed, but the model and algorithm do not account for multiple appearances of proteins.

In the present paper, we develop a model-based framework for analysis of PALM data. Although this motivating application is in 2D, we consider point processes in $\mathbb{R}^d$. The parent process (of proteins) is modelled as a stationary point process X in $\mathbb{R}^d$. To each $x \in X$, we associate a point process Y_x, representing the multiple appearances of a protein. Conditional on X, the processes $\{Y_x : x \in X\}$ are independent and identically distributed (i.i.d.). If Y is a point process, having the common distribution of the Y_xs, we suppose that the intensity function of Y has the form $\rho_Y(y) = \tau k(y)$ where $\tau > 0$ is the mean number of multiple appearances and k is a probability density function on $\mathbb{R}^d$. In the application from PALM, k is the so-called point spread function and typically of Gaussian form

$$k(y) = \exp(-\|y\|^2/(2\sigma^2))/(2\pi\sigma^2)^{d/2}, \qquad y \in \mathbb{R}^d.$$

The observed process

$$Z = \bigcup_{x \in X} (x + Y_x)$$

is thus modelled as an *independent cluster process* (Lieshout and Baddeley, 2002; Illian et al., 2008, p. 370). These processes can be classified into various subclasses, including *Poisson cluster processes*, for which X is assumed to be a Poisson process, and *Cox cluster processes* (CCPs), for which the Y_xs are assumed to be Poisson processes. Bayesian inference for CCPs with X modelled by a repulsive Markov point process has been studied in Lieshout and Baddeley (2002). Considering the PALM applications where proteins typically are clustered, we are instead interested in models for Z for which X shows clustering.

We will primarily focus on parametric models for Z for which X is a CCP and Y is a Poisson process. In such cases, we call Z a *double Cox cluster process*

(DCCP), since both the parent process X and Z itself are CCPs. DCCPs with Neyman-Scott processes as parent processes is a very flexible model class. We show that moment-based inference is particularly simple for this type of processes. An important example is the *double Thomas process* (TTP), obtained by assuming that X is a Thomas process and the probability density k is Gaussian. TTPs have also been used in ecology for modelling two scales of interaction in rain forest data (Wiegand et al., 2007).

The present paper is organized as follows. In Section 2, we introduce the notation and basic properties of point processes used in this paper. In Section 3, we give a summary of first- and second-order moment properties of independent cluster processes while DCCPs together with the associated moment-based inference are treated in Section 4. Parametric models for DCCPs with a Neyman-Scott parent process are discussed in Section 5. In Section 6, we apply the developed moment-based inference to a data set from a PALM acquisition. The model used in the analysis also includes background noise modelled by a Poisson process. A discussion regarding further generalizations of the model, Bayesian inference and future work, may be found in Section 7. Some derivations are deferred to Appendix A while additional information about model fitting in Section 6 may be found in Appendix B.

2 Preliminaries

This section introduces the notation and basic properties of the point process models considered in this paper. For further details, see Møller and Waagepetersen (2004); Illian et al. (2008); Chiu et al. (2013).

A spatial point process Z in $\mathbb{R}^d$ is a random locally finite subset of $\mathbb{R}^d$. We assume that Z has a well-defined intensity function ρ_Z and second-order product density $\rho_Z^{(2)}$, such that the intensity measure α_Z and the second-order factorial moment measure $\alpha_Z^{(2)}$ are given by

$$\alpha_Z(C) = \mathbb{E}\Big[\sum_{z\in Z} \mathbb{1}(z \in C)\Big] = \int_C \rho_Z(z)\,\mathrm{d}z,$$
$$\alpha_Z^{(2)}(C_1 \times C_2) = \mathbb{E}\Big[\sum_{z_1,z_2\in Z}^{\neq} \mathbb{1}(z_1 \in C_1, z_2 \in C_2)\Big] = \int_{C_1}\int_{C_2} \rho_Z^{(2)}(z_1,z_2)\,\mathrm{d}z_1\,\mathrm{d}z_2,$$

for C, C_1 and C_2 in the Borel σ-algebra on $\mathbb{R}^d$. Here $\sum^{\neq}$ denotes summation over distinct pairs. The interaction between pairs of points can be measured by the *pair correlation function*

$$g_Z(z_1,z_2) = \rho_Z^{(2)}(z_1,z_2)/\big(\rho_Z(z_1)\rho_Z(z_2)\big).$$

A *stationary* process, i.e. a process with translation invariant distribution, has a constant intensity function and a translation invariant second-order product density, i.e. $\rho_Z^{(2)}(z_1,z_2) = \rho_Z^{(2)}(z_1 - z_2)$. A process is said to be *isotropic* if its distribution is rotation invariant. For stationary and isotropic processes, we have $\rho_Z^{(2)}(z_1,z_2) = \rho_Z^{(2)}(\|z_1 - z_2\|)$ and the pair correlation function is effectively a function on $\mathbb{R}$,

$$g_Z(r) = \rho_Z^{(2)}(r)/\rho_Z^2, \qquad r \in \mathbb{R}.$$

The function can be interpreted as the mean number of points at distance r from a "typical" point in Z, relative to the mean number for a Poisson process with same intensity.

Additional summary functions are distance functions such as the F-function and G-function, defined for a stationary process Z. Here, F is the distribution function of the distance from the origin to the nearest point in Z while G is the distribution function of the distance from a typical point of Z to the nearest neighbour in Z. Finally, the J-function is defined by

$$J(r) = (1 - G(r))/(1 - F(r)) \qquad \text{for } F(r) < 1,$$

see Lieshout and Baddeley (1996).

For Poisson processes, $g(r) \equiv J(r) \equiv 1$ and $G(r) \equiv F(r)$, while for cluster processes $g(r) > 1 > J(r)$ and $G(r) > F(r)$ for small r values.

The CCPs studied in the present paper form a subclass of the class of *Cox processes* (Cox, 1955; Møller and Waagepetersen, 2004). Let $\{\Lambda(z)\}$ be a non-negative random field. A point process is then a Cox process with driving field Λ, if conditionally on Λ, the process is a Poisson process with intensity function Λ. Note that a CCP (as defined in Section 1) is actually a Cox process with driving field $\Lambda(z) = \sum_{x \in X} \rho_Y(z - x)$, where ρ_Y is the common intensity function of the Y_x processes.

3 Moment properties of independent cluster processes

Let $Z = \cup_{x \in X}(x + Y_x)$ be an independent cluster process, as defined in Section 1. In particular, X is stationary. We suppose that X has a finite and positive intensity ρ_X and pair correlation function g_X. The pair correlation function of Y is denoted g_Y.

Using conditioning on X, it is easy to show that the intensity of Z is $\rho_Z = \rho_X \tau$. Since Z is stationary, the pair correlation function of Z is a function of one argument only and takes the following form

$$g_Z(z) = \rho_X^{-1} \int_{\mathbb{R}^d} k(z + x)k(x)g_Y(z + x, x) \, \mathrm{d}x + \int_{\mathbb{R}^d} h_k(z - x)g_X(x) \, \mathrm{d}x, \qquad (3.1)$$

$z \in \mathbb{R}^d$, where

$$h_k(z) = \int_{\mathbb{R}^d} k(z + x)k(x) \, \mathrm{d}x, \qquad z \in \mathbb{R}^d. \qquad (3.2)$$

The equation (3.1) holds under the assumption that the processes $x + Y_x$, $x \in X$, do not overlap. We will make this assumption in the following. Note that h_k is the probability density of $U_1 - U_2$ where U_1 and U_2 are independent, both with probability density k.

The first term in (3.1) corresponds to the contribution from pairs of points from the same cluster and the second term is the contribution from pairs of points from different clusters. The formula (3.1) for g_Z follows from Illian et al. (2008, (6.2.10)), see also Felsenstein (1975); Shimatani (2001). In Appendix A, a short proof of (3.1) in the notation used in the present paper may be found.

Let N be the number of points in Y. If, conditional on N, the points in Y are i.i.d., then (3.1) simplifies to, cf. Appendix A,

$$g_Z(z) = \frac{\mathbb{E}(N^2) - \mathbb{E}(N)}{\rho_X \tau^2} h_k(z) + \int_{\mathbb{R}^d} h_k(z-x) g_X(x) \, \mathrm{d}x, \qquad z \in \mathbb{R}^d. \tag{3.3}$$

Note that $N \equiv 1$ corresponds to the situation of noisy observations of a spatial point process in which case (3.3) reduces to a convolution

$$g_Z(z) = \int_{\mathbb{R}^d} h_k(z-x) g_X(x) \, \mathrm{d}x, \qquad z \in \mathbb{R}^d.$$

In the case where Y is a Poisson process, Z is a CCP and (3.3) simplifies to

$$g_Z(z) = \rho_X^{-1} h_k(z) + \int_{\mathbb{R}^d} h_k(z-x) g_X(x) \, \mathrm{d}x, \qquad z \in \mathbb{R}^d. \tag{3.4}$$

Let us now consider the case where the probability density k is Gaussian

$$k(y) = \exp(-\|y\|^2/2\sigma^2)/(2\pi\sigma^2)^{d/2}, \qquad y \in \mathbb{R}^d,$$

which is often used as a model for point spread functions in PALM. If X is isotropic, we can use polar decomposition in $\mathbb{R}^d$ and express the pair correlation function (3.4) of Z more explicitly as

$$\begin{aligned} g_Z(z) = {} & (4\pi\sigma^2)^{-d/2} \exp(-\|z\|^2/(4\sigma^2)) \\ & \times \Big[\rho_X^{-1} + \int_0^\infty \exp\big(-r^2/(4\sigma^2)\big) F\big(\|z\| r/(2\sigma^2)\big) g_X(r) r^{d-1} \, \mathrm{d}r \Big], \end{aligned} \tag{3.5}$$

where

$$F(r) = \int_{\mathbb{S}^{d-1}} \exp(r u_1) \, \mathrm{d}u^{d-1}, \qquad r > 0, \tag{3.6}$$

and u_1 is the first coordinate of $u \in \mathbb{S}^{d-1}$. For $d = 2$,

$$F(r) = 2\pi I_0(r), \qquad r > 0,$$

where I_0 is the modified Bessel function of the first kind with index 0.

Relation (3.5) may be used for numerical evaluation of g_Z for a specific choice of model for g_X.

Equation (3.3) implies that the relationship between the Fourier transforms of g_X and g_Z is of the form

$$\mathcal{F}(g_Z) = \mathcal{F}(h_k) \left(\frac{\mathbb{E}(N^2) - \mathbb{E}(N)}{\rho_X \tau^2} + \mathcal{F}(g_X) \right)$$

where the d-dimensional Fourier transform is given by

$$\mathcal{F}(f)(x) = \int_{\mathbb{R}^d} \exp(-ix \cdot y) f(y) \, \mathrm{d}y.$$

4 Double Cox cluster processes (DCCPs)

For the PALM application, it is interesting to consider the case where the parent process X is a cluster process. An independent cluster process Z is said to be a *double Cox cluster process* (DCCP) if X is a CCP and Y is Poisson. Note that when Z is a DCCP, both X and Z are CCPs.

The intensity and pair correlation function of a DCCP Z are easily obtained, using the results of Section 3. We have

$$Z = \bigcup_{x \in X} (x + Y_x),$$

as previously, but now

$$X = \bigcup_{\phi \in \Phi} (\phi + U_\phi),$$

where Φ is a stationary process with intensity κ, and conditional on Φ, the processes $\{U_\phi : \phi \in \Phi\}$ are i.i.d. Poisson processes with intensity function $\rho_U(u) = \mu k_X(u)$, say. Here, $\mu > 0$ and k_X is a probability density on $\mathbb{R}^d$. Note that $\rho_X = \mu\kappa$.

The intensity of Z is $\rho_Z = \tau\mu\kappa$. Using that X is a CCP, we find, cf. (3.4),

$$g_X(x) = \kappa^{-1} h_{k_X}(z) + \int_{\mathbb{R}^d} h_{k_X}(x - y) g_\Phi(y) \, \mathrm{d}y, \qquad x \in \mathbb{R}^d, \tag{4.1}$$

and, using that (3.4) also holds for g_Z, we find

$$\begin{aligned} g_Z(z) = (\mu\kappa)^{-1} h_k(z) + \kappa^{-1} \int_{\mathbb{R}^d} h_k(z - x) h_{k_X}(x) \, \mathrm{d}x \\ + \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} h_k(z - x) h_{k_X}(x - y) g_\Phi(y) \, \mathrm{d}y \, \mathrm{d}x, \qquad z \in \mathbb{R}^d. \end{aligned} \tag{4.2}$$

In the particular case where Φ is Poisson, X is a Neyman-Scott process (Neyman and Scott, 1958; Illian et al., 2008, p. 374) and (4.1) and (4.2) reduce to

$$g_X(x) = 1 + \kappa^{-1} h_{k_X}(x), \qquad x \in \mathbb{R}^d, \tag{4.3}$$

and

$$g_Z(z) = 1 + (\mu\kappa)^{-1} h_k(z) + \kappa^{-1} \int_{\mathbb{R}^d} h_k(z - x) h_{k_X}(x) \, \mathrm{d}x, \qquad z \in \mathbb{R}^d. \tag{4.4}$$

Moment-based inference for DCCPs may be performed, using a discrepancy measure such as

$$D(\theta) = \int_{r_1}^{r_2} (\hat{g}_Z(r)^q - g_Z(r; \theta)^q)^p \, \mathrm{d}r, \tag{4.5}$$

where $\hat{g}_Z$ is a kernel estimator of g_Z and $g_Z(\cdot; \theta)$ is of the form (4.2). Note that it may be difficult to estimate all components of the parameter θ with high precision. As an example, suppose h_{k_X} is much more concentrated at the origin than h_k. Then, (4.4) is approximately

$$g_Z(z) \approx 1 + ((\mu\kappa)^{-1} + \kappa^{-1}) h_k(z).$$

In such an extreme case, there is no information about the parent process in g_Z.

5 DCCPs with a Neyman-Scott process as parent process

In this section, we discuss different parametric models for the probability density k_X and the resulting expressions for g_X and g_Z. We focus on the case where X is a Neyman-Scott process. Having PALM applications in mind, k will be Gaussian.

If both k and k_X are of Gaussian form (with not necessarily the same parameters), then X is a Thomas process and Z is called a *double Thomas process* with abbrevation TTP. This process has also been considered in ecology, see Wiegand et al. (2007).

The pair correlation function g_Z has a closed form when Z is a TTP. If the Gaussian probability density k_X is parametrized by ω, then the pair correlation function of X is, cf. (4.3),

$$g_X(x) = 1 + \kappa^{-1}(4\pi\omega^2)^{-d/2}\exp(-\|x\|^2/(4\omega^2)), \qquad x \in \mathbb{R}^d, \tag{5.1}$$

and, using (4.4), we thus find for $z \in \mathbb{R}^d$

$$\begin{aligned} g_Z(z) = 1 &+ (\mu\kappa)^{-1}(4\pi\sigma^2)^{-d/2}\exp\bigl(-\|z\|^2/(4\sigma^2)\bigr) \\ &+ \kappa^{-1}\bigl(4\pi(\omega^2+\sigma^2)\bigr)^{-d/2}\exp\bigl(-\|z\|^2/[4(\omega^2+\sigma^2)]\bigr). \end{aligned} \tag{5.2}$$

More generally, if

$$g_X(x) = 1 + \sum_{i=1}^{n} \alpha_i \exp(-\|x\|^2/\beta_i), \qquad x \in \mathbb{R}^d, \tag{5.3}$$

with $\alpha_i, \beta_i > 0$, then the last term in (5.2) is replaced by, cf. (3.4),

$$\sum_{i=1}^{n} \alpha_i\bigl(\beta_i/(\beta_i+4\sigma^2)\bigr)^{d/2}\exp\bigl(-\|z\|^2/(\beta_i+4\sigma^2)\bigr).$$

As shown in Møller and Christoffersen (*in preparation*), (5.3) is the pair correlation function of a point process that can be obtained by an iterative scheme.

Furthermore, the class of Neyman-Scott processes also includes processes for which the pair correlation function g_X follows the flexible Matérn type covariance model (Jalilian et al., 2013; Jónsdóttir et al., 2013). As an example, suppose that g_X is of exponential form

$$g_X(x) = 1 + \kappa^{-1}\alpha\exp(-\|x\|/\eta), \qquad x \in \mathbb{R}^d, \tag{5.4}$$

with $\alpha, \eta > 0$. In order to find a probability density k_X such that (4.3) is satisfied, we recall that h_{k_X} is a probability density and therefore,

$$\alpha = \bigl(2^d\pi^{(d-1/2)}\Gamma([d+1]/2]\eta^d\bigr)^{-1}. \tag{5.5}$$

Using Jónsdóttir et al. (2013, p. 515), we find

$$k_X(x) = \bigl(2^{3(d-1)/4}\pi^{d/2}\Gamma([d+1]/4)\eta^d\bigr)^{-1}(\|x\|/\eta)^{-(d-1)/4}K_{-(d-1)/4}(\|x\|/\eta), \tag{5.6}$$

where K_ν is a modified Bessel function of the second kind with index ν.

In Sengupta et al. (2011), the correlation between proteins are exactly assumed to give rise to an exponential form of g_X for $d = 2$, however, no parametric point process model is stated for the proteins. More generally, we can find the probability density k_X corresponding to a pair correlation function of the form

$$g_X(x) = 1 + \alpha\|x\|^\nu K_\nu(\|x\|/\eta), \qquad x \in \mathbb{R}^d,\ \alpha, \eta, \nu > 0.$$

When X is a Neyman-Scott process with k_X of the form (5.6) and k is Gaussian, then Z is said to be a *Thomas exponential process* (TEP). Its pair correlation function is of the form, cf. (4.4),

$$\begin{aligned} g_Z(z) = 1 &+ (\mu\kappa)^{-1}(4\pi\sigma^2)^{-d/2}\exp\bigl(-\|z\|^2/(4\sigma^2)\bigr) \\ &+ \kappa^{-1}(4\pi\sigma^2)^{-d/2}\alpha\int_{\mathbb{R}^d}\exp\bigl(-\|z-x\|^2/(4\sigma^2)\bigr)\exp(-\|x\|/\eta)\,\mathrm{d}x, \end{aligned} \tag{5.7}$$

where α is given in (5.5). Using polar coordinates in $\mathbb{R}^d$, the last integral in (5.7) can be rewritten as

$$\exp\bigl(-\|z\|^2/(4\sigma^2)\bigr)\int_0^\infty \exp\bigl(-r^2/(4\sigma^2)\bigr)F\bigl(\|z\|r/(2\sigma^2)\bigr)\exp(-r/\eta)r^{d-1}\,\mathrm{d}r,$$

where the function F is defined in (3.6).

In Figure 1, g_X and g_Z are shown for $d = 2$ for some examples of TTPs, using (5.1) and (5.2), and some examples of TEPs, using (5.4) and (5.7). Note that the pair correlation function g_X is quite different in the case of TTP and TEP, but g_Z has a similar appearances for the two types of processes.

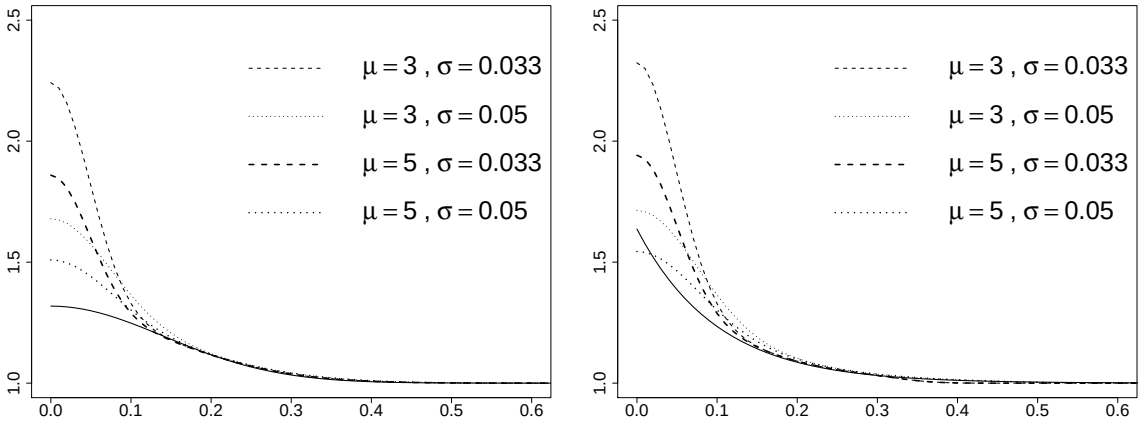

Figure 1: Plots of g_X (full drawn lines) and g_Z (stippled and dotted lines) for some examples of the TTP (left) and the TEP (right). The parameters $\kappa = 25$ and $\omega = \eta = 0.1$ are fixed while the remaining parameters (which do not have an impact on g_X) are $\mu = 3$ or 5 and $\sigma = 0.033$ or 0.050.

6 Data example

Data originates from a PALM study (Arnspang et al., *in preparation*) and consists of $N = 1324$ observed positions of fluorescent proteins in a sub-section of a super-resolution PALM image, as shown in Figure 2. A single protein may fluoresce a

number of times at slightly varying positions and may thus appear as a small cluster. Furthermore, the proteins are expected to cluster, hence DCCPs seem to be a natural choice of model class for the data.

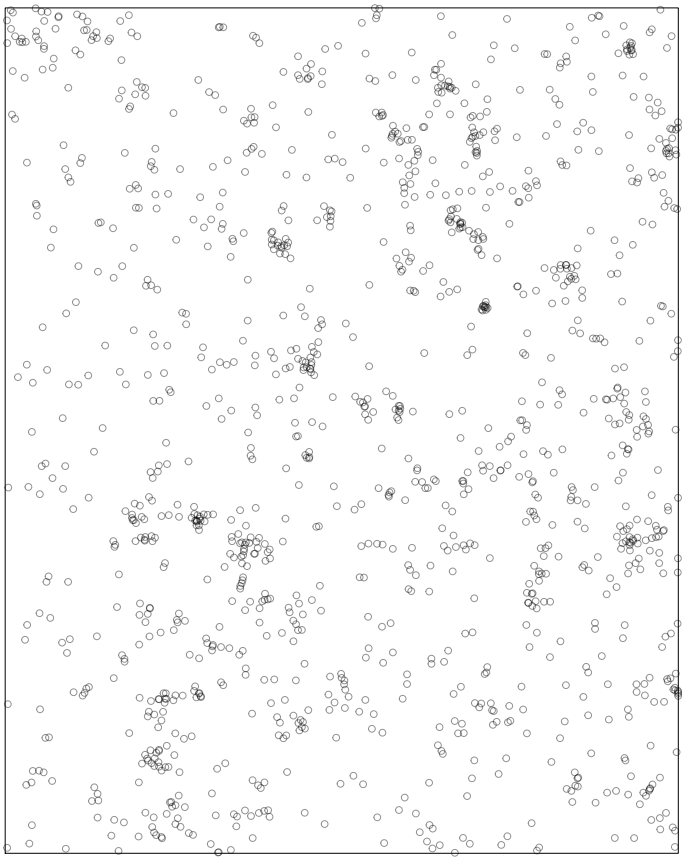

Figure 2: Observed point pattern consisting of $N = 1324$ positions of fluorescent proteins in a sub-section of size $3355\,\mathrm{nm} \times 4188\,\mathrm{nm}$ of a PALM image.

We use moment-based inference and start by fitting a TTP and a TEP, using the discrepancy measure (4.5) with $d = 2$, $g_Z(\,\cdot\,;\theta)$ of the form (5.2) or (5.7) and $\theta = (\kappa, \omega, \mu, \sigma, \tau)$ or $\theta = (\kappa, \eta, \mu, \sigma, \tau)$, respectively. Note that $g_Z(\,\cdot\,;\theta)$ does not depend on τ so the estimation procedure is supplemented by the estimating equation $\hat{\rho}_Z = \mu\kappa\tau$.

In Figure 3, $\hat{g}_Z$ (full drawn line) is plotted together with $g_Z(r;\hat{\theta})$ for both models (stippled lines). See Appendix B for numerical details.

Both TTP and TEP provide a very good fit to the observed pair correlation function. However, these models do not explain the quite high fraction of single points that is visible in the plotted data. The J-functions shown in Figure 4 reveal that the pure DCCP models are not sufficient to describe the data. We therefore extend the model by superimposing a Poisson point process, which may be interpret as background noise from the microscope. This model has an additional parameter, namely the fraction a of points that originate from the DCCP model. Below we propose a two step method to fit the model, by minimizing the discrepancy of both the pair correlation function and the J-function. For sake of readability, the procedure is described only for the TTP model; fitting a TEP model is completely analogous.

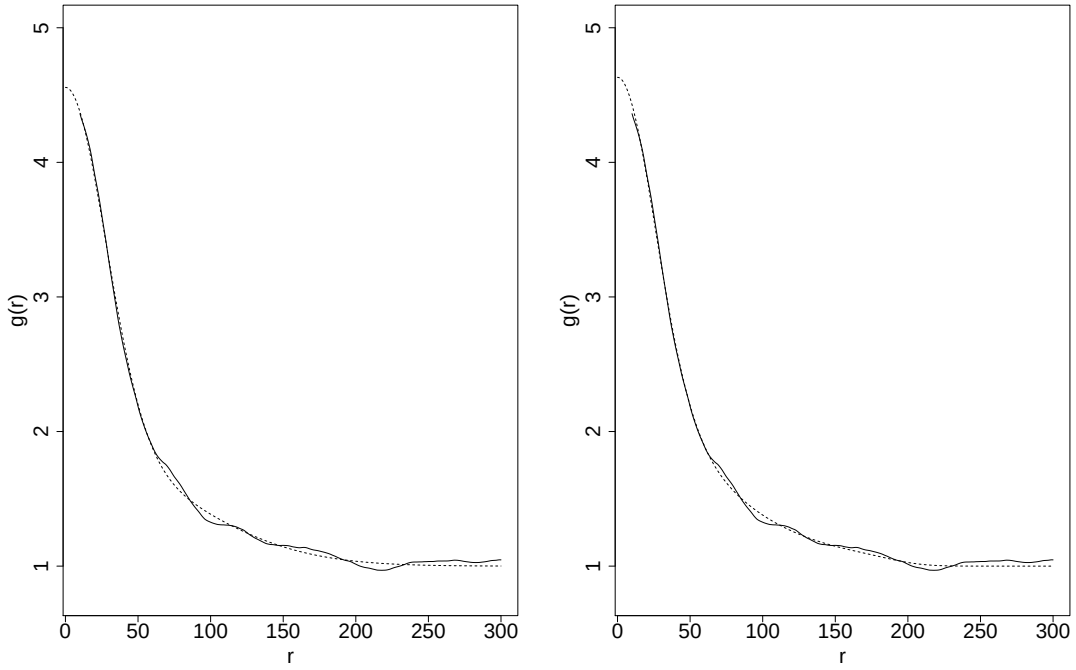

Figure 3: Estimated pair correlation function (full drawn line) of the observed point pattern and the theoretical pair correlation function of the fitted processes (stippled lines) for the TTP (left) and the TEP (right).

Let $\tilde{Z}$ denote a stationary and isotropic point process that results from independent superposition of a TTP process Z and a stationary Poisson point process. The process $\tilde{Z}$ has pair correlation function

$$g_{\tilde{Z}}(r) = 1 + a^2(g_Z(r) - 1), \tag{6.1}$$

and J-function

$$J_{\tilde{Z}}(r) = 1 + aJ_Z(r) + (1 - a), \tag{6.2}$$

where $a = \rho_Z/\rho_{\tilde{Z}}$, see Illian et al. (2008, p.371) and Lieshout and Baddeley (1996).

From (5.2) and (6.1), we see that $g_{\tilde{Z}}$ can be written in the form

$$g_{\tilde{Z}}(r) = 1 + \tfrac{a^2}{\kappa} f(r; \omega, \sigma, \mu).$$

This is the pair correlation function of a TTP Z_0 with parameters $\kappa_0 = \kappa/a^2$, $\omega_0 = \omega$, $\sigma_0 = \sigma$, and $\mu_0 = \mu$. Estimates of these parameters and of τ_0 are obtained by fitting Z_0 to the data as previously described. In the extended model, $\hat{\kappa}_0$ is an estimate of κ/a^2 and with $a\rho_{\tilde{Z}} = \mu\kappa\tau$, it follows that $\hat{\tau}_0$ is an estimate of τa. The remaining parameter estimates remain valid in the extended model.

The extra parameter a of the extended model is subsequently estimated by minimizing the distance between the observed J-function and $J_{\tilde{Z}}$. Since J_Z is not known in closed form, this was done by simulation of superposition processes for a sequence of different values of a, followed by minimizing the discrepancy measure similar to (4.5) based on the J-function with respect to a. We used medians of the estimated J-functions based on the simulations as estimates of the theoretical J-functions and obtained estimates for a approximately equal to 0.6 for both processes.

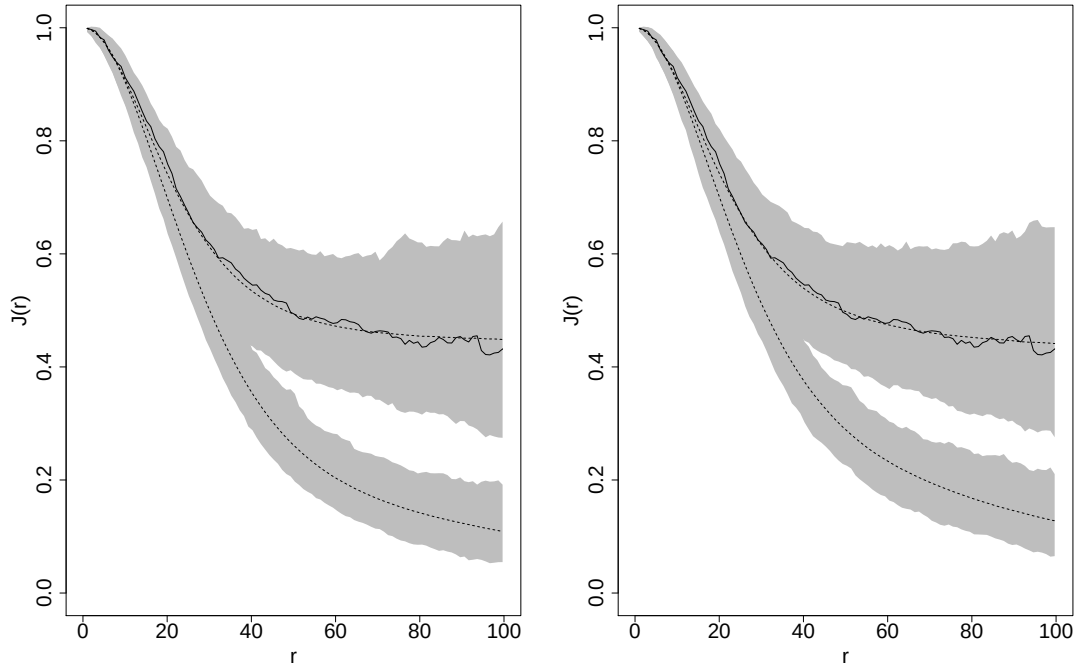

Figure 4: Estimated J-function for the observed point pattern (full drawn line), compared to the ones for the fitted (i) TTP (left) and TEP (right) and (ii) DCCPs superimposed with approximately 40% Poisson noise, (ii) being the process which coincides with the estimate from the observed point pattern. 95% global envelopes (grey areas) and median estimates (stippled lines) for $r = 0, 1, \ldots, 100$, are based on 2000 simulations.

We clearly get a much better fit when the superposition models are used, as illustrated in Figure 4, where the estimated J-function of the observed point pattern (full drawn line) is compared to estimates of J-functions for the fitted DCCPs with and without superimposed background noise. Thus, approximately $(1 - a)100\% = 40\%$ of the points are modelled as background noise.

All estimates from the fitted DCCPs and the superposition models are found in Table 1 (top), together with estimated relative mean squared error (RMSE), estimated relative bias (Rbias) and measured discrepancy of the fitted models. Both the TTP and TEP with background noise provide a very good fit to the data, which is also visualized in Figure 5, where realizations from the fitted models are shown. We tested the models further using envelope tests, based on the pair correlation function and the J-, G- and F-functions, and found no significant deviations.

	TTP				TEP			
	Est	Est_2	RMSE	Rbias	Est	Est_2	RMSE	Rbias
$\hat{a}$		57.5%				59.5%		
$\hat{\kappa}$ $(\times 10^{-5})$	2.98	**0.99**	.17	.13	2.39	**0.84**	.09	.05
$\hat{\mu}$	2.62	**2.62**	.14	$-.05$	4.99	**4.99**	.33	$-.04$
$\hat{\omega}/\hat{\eta}$	53.06	**53.06**	.06	.09	51.92	**51.92**	.29	.30
$\hat{\tau}$	1.21	**2.11**	.11	.11	0.80	**1.34**	.59	.37
$\hat{\sigma}$	19.36	**19.36**	.02	.08	17.42	**17.42**	.06	.14
$D(\hat{\theta})$	0.0134				0.0106			

	TTP				TEP			
	Est	Est_2	RMSE	Rbias	Est	Est_2	RMSE	Rbias
$\hat{a}$		57.5%				58.0%		
$\hat{\kappa}$ $(\times 10^{-5})$	3.03	**1.00**	.06	.02	2.55	**0.86**	.06	.01
$\hat{\mu}$	2.42	**2.42**	.09	.16	3.36	**3.36**	.25	.28
$\hat{\omega}/\hat{\eta}$	54.26	**54.26**	.03	$-.02$	55.89	**55.89**	.07	.00
$\hat{\tau}$	1.29	**2.25**	.03	-.09	1.10	**1.90**	.05	$-.14$
$\sigma = 20$	—	—	—	—	—	—	—	—
$D(\hat{\theta})$	0.0134				0.0169			

Table 1: Parameter estimates for the TTP (left) and the TEP (right), before (Est) and after (Est_2) correction for $(1-a)100\%$ background noise. Estimated relative mean square error (RMSE) and estimated relative bias (Rbias) based on 200 simulations of the fitted processes are also shown together with discrepancies between observed and fitted pair correlation functions. The top table shows results for the models with unknown σ and the bottom table shows the results for the models with fixed $\sigma = 20$.

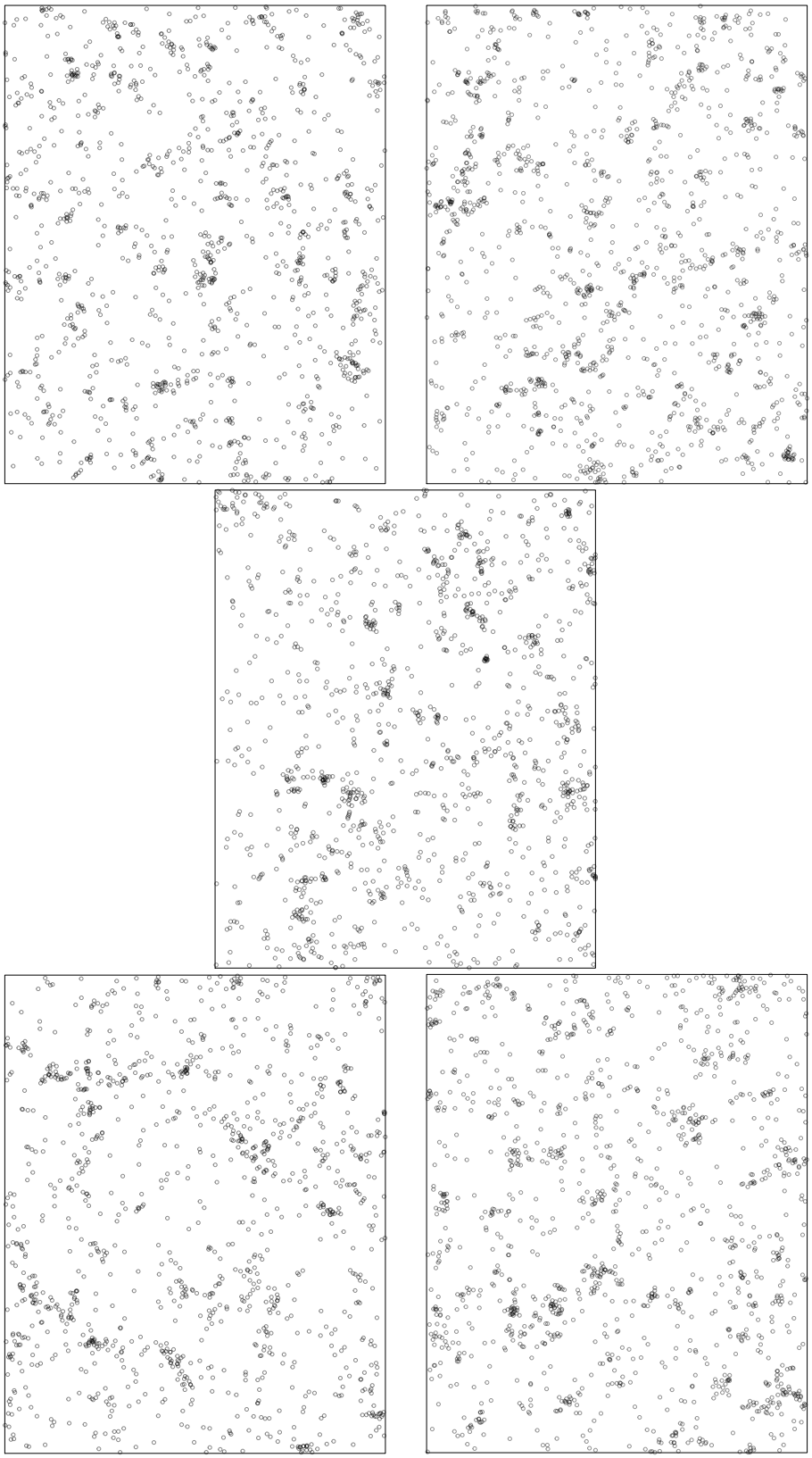

Figure 5: The observed point pattern (middle) together with two realizations from the fitted TTP (top) and fitted TEP (bottom) with superimposed Poisson noise.

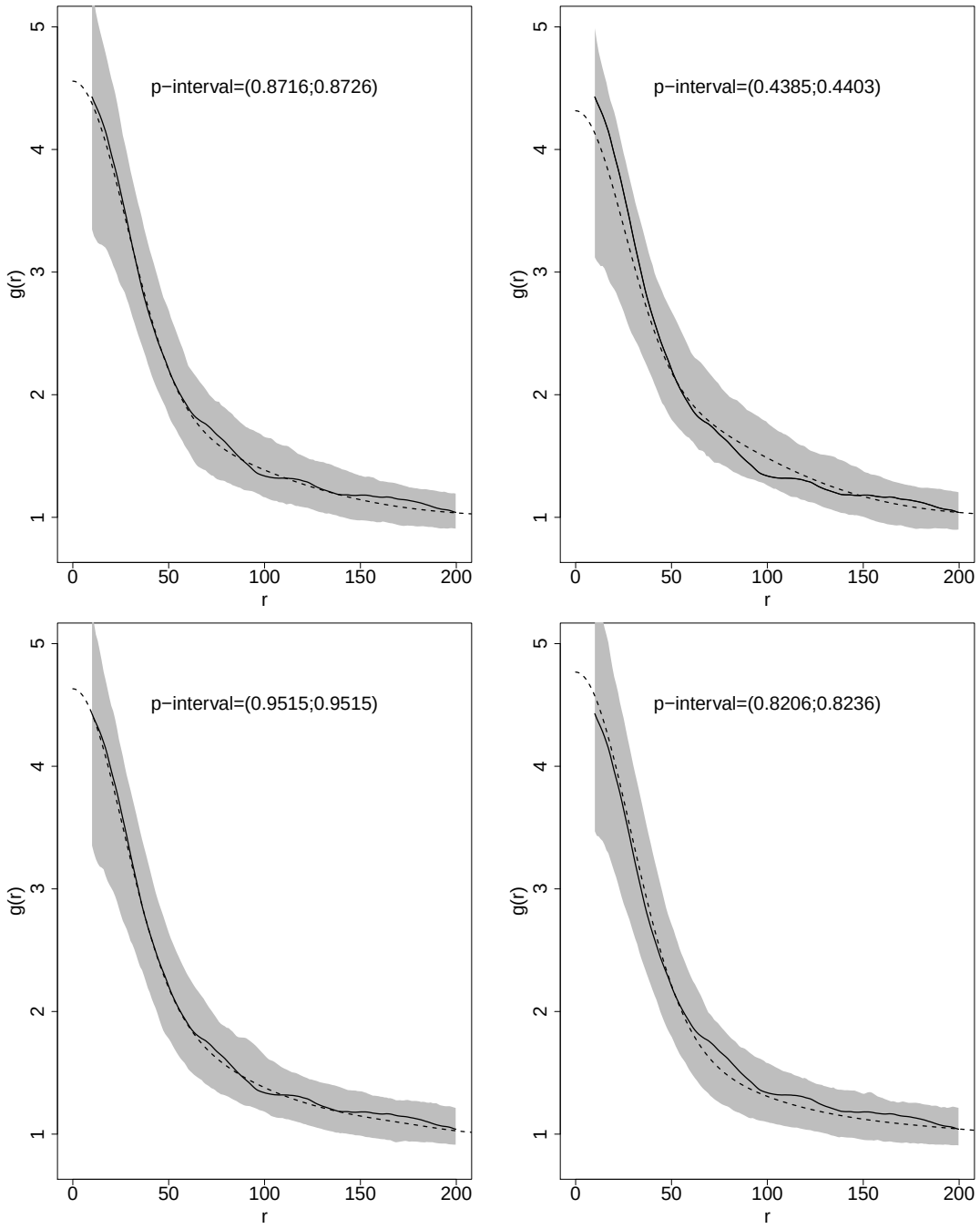

Figure 6: 95% global envelopes (grey areas) for $r = 10, 11, \ldots, 200$, based on 2000 simulations and the pair correlation function for the fitted TTP (top left), the TTP with alternative parameters (top right), the fitted TEP (bottom left) and the TEP with alternative parameters (bottom right). The estimated pair correlation function of the observed point pattern (full drawn line) is shown together with the theoretical pair correlation functions (stippled lines).

As shown in Table 1 (top), the RMSEs of the estimators are quite large. Based on simulations of the fitted processes, we also found that the estimators were correlated, one obvious reason is that in the noise free models we use the estimating equation $\hat{\rho}_Z = \mu\kappa\tau$. In fact, the parameter estimates obtained in the TTP model also gave an acceptable fit in the TEP model and vice versa, see the global envelope tests based on the pair correlation functions in Figure 6. The same conclusion was obtained when using envelope tests based on F-, G- and J-functions. As a consequence, the observed data can be generated, using parent processes with somewhat different degree of clustering, see Figure 7 showing simulations from the two fitted parent processes.

However, if prior information about σ is available, it is possible to estimate the parent process with higher precision. For instance, if we let $\sigma = 20$, we obtain the results reported in Table 1 (bottom). The parameter estimates in the TTP and TEP models are now much more alike, and the MSE and bias of the estimators are reduced in this simpler model.

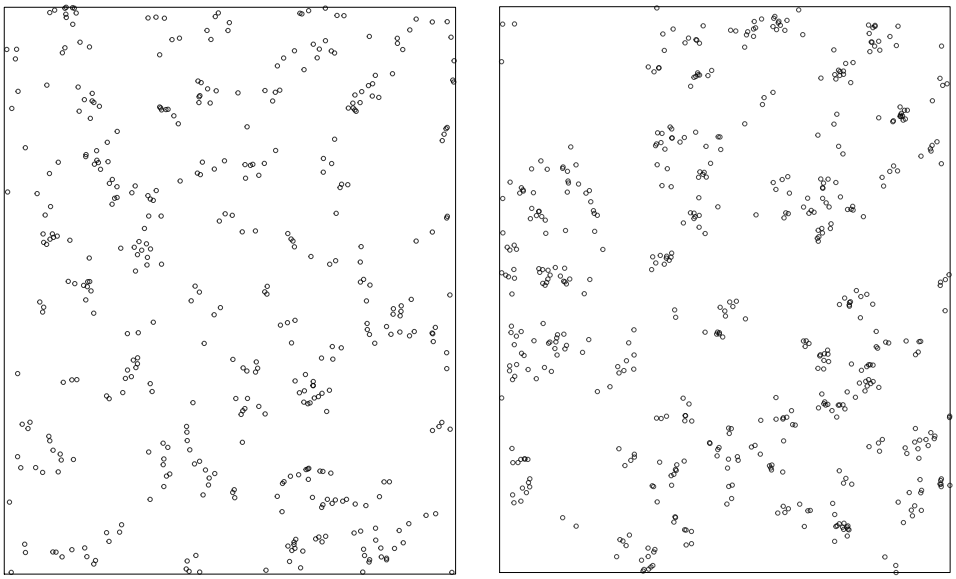

Figure 7: Realizations from the fitted parent processes, i.e. a Thomas process (left) and a Neyman-Scott process with an exponential pair correlation function (right).

7 Discussion

In the present paper, we have developed a model-based framework for analysis of PALM data. The focus has been on moment-based inference for double Cox cluster processes, in particular the double Thomas process and the Thomas exponential process.

The model considered in the present paper may be extended in various ways. Apart from superposition of independent Poisson noise which was used in the data example, we may in addition include independent thinning. For the PALM application, this may be a relevant extension of the model since it is expected that not all proteins are actually observed.

In Møller and Christoffersen (*in preparation*), the pair correlation function for such a model is studied in an iterative scheme, described as a discrete time Markov chain of point processes. Their basic Lemma 3.4 simplifies under no noise and thinning to the relation (3.3) between g_X and g_Z in the present paper. See also Møller and Torrisi (2005, Proposition 2) for the special case of generalized shot noise Cox processes. The focus of Møller and Christoffersen (*in preparation*) is, however, different from the focus of the present paper. The iterative scheme is the primary object of study in their paper, including limiting distributions, while the aim of the present paper is parametric modelling and statistical inference for double Cox cluster point processes, motivated by the applications in PALM.

Our set-up includes as a special case modelling of noisy observations of spatial point processes. Earlier references on inference for spatial point processes subject to noise are Lund and Rudemo (2000); Cucala (2008); Bar-Hen et al. (2013). The latter reference considers the influence of measurement errors on descriptive statistics for testing complete spatial randomness.

In the analysis of the data example, we have found that the observed data could be generated by parent processes with somewhat different degree of clustering. If prior information about some of the model parameters are available, we encourage to use this information because then it is possible to estimate the parent process with much higher precision.

In Sengupta et al. (2011, 2013); Veatch et al. (2012), the convolution in the last term in (10) seems to be replaced by $h_{k_X}(z)$ as a simple approximation, see e.g. Veatch et al. (2012, p. 5). In the models considered for the data example in the present paper, we found that parameter estimates, obtained with such approximation, deviate up to 43% from the estimates, obtained by numerical integration based on (3.5). Hence, when a Gaussian point spread function is used in PALM, better estimates can be obtained, using (3.5).

An alternative to the proposed method is to construct a non-parametric estimate $\hat{g}_X$ of g_X, using the relation between the Fourier transforms of g_X and g_Z, derived in Section 3, and then use moment-based inference directly on the estimate $\hat{g}_X$ and the parametric model $g_X(\cdot;\theta)$.

An alternative method to moment-based inference is Bayesian inference, which has the advantage, that it also provides a posterior estimate of the intensity surface of the parent process X. Bayesian inference can often improve the quality of the parameter estimates, see e.g. a comparison of the methods for Neyman-Scott processes in Kopeckỳ and Mrkvička (2016).

Unfortunately, the Bayesian approach is time-consuming as the posterior density is intractable, hence it must be found using MCMC methods. Additional background noise complicates Bayesian inference for double Cox cluster processes further, as an additional classification algorithm, similar to the one in Redenbach et al. (2015), is required to classify points as either cluster points or background points.

Preliminary studies indicate that a Bayesian approach can be used for double Cox cluster point processes without background noise, for instance, in the case where the parameters for the multiple appearances are known. Future work includes further investigation as to what extent, Bayesian inference can be applied to double Cox cluster processes with background noise.

8 Acknowledgement

This research has been supported by Centre for Stochastic Geometry and Advanced Bioimaging, funded by grant 8721 from the Villum Foundation. The microscope that produced the point pattern in Figure 2, was funded by a Junior Group Leader Fellowship from the Lundbeck Foundation, the Carlsberg Foundation and MEMBRANES (AU) to LNN. The research relating to the PALM acquisition analyzed in Section 6 has also been supported by grant CF14-0786 from the Carlsberg Foundation and grant 0602-02518B from FNU Sapere aude: to EAC.

References

Arnspang, E. C., H. H. Jensen, P. Sengupta, U. Hahn, I. T. Andersen, E. B. V. Jensen, K. Mortensen, J. Lippincott-Schwartz, and L. N. Nejsum. Pair correlation analysis of fixed photoactivatable localization microscopy (PALM) and powerspectral analysis of live PALM applied on the water channel aquaporin-3. *In preparation*.

Baddeley, A., E. Rubak, and R. Turner (2015). *Spatial Point Patterns: Methodology and Applications with R*. London: Chapman and Hall/CRC Press.

Bar-Hen, A., J. Chadoeuf, H. Dessard, and P. Monestiez (2013). Estimating second order characteristics of point processes with known independent noise. *Statistics and Computing 23*(3), 297–309.

Betzig, E., G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino, M. W. Davidson, J. Lippincott-Schwartz, and H. F. Hess (2006). Imaging intracellular fluorescent proteins at nanometer resolution. *Science 313*(5793), 1642–1645.

Chiu, S. N., D. Stoyan, W. S. Kendall, and J. Mecke (2013). *Stochastic Geometry and its Applications*. John Wiley & Sons.

Cox, D. R. (1955). Some statistical methods connected with series of events. *Journal of the Royal Statistical Society. Series B (Methodological)*, 129–164.

Cucala, L. (2008). Intensity estimation for spatial point processes observed with noise. *Scandinavian Journal of Statistics 35*(2), 322–334.

Felsenstein, J. (1975). A pain in the torus: some difficulties with models of isolation by distance. *The American Naturalist 109*(967), 359–368.

Illian, J., A. Penttinen, H. Stoyan, and D. Stoyan (2008). *Statistical Analysis and Modelling of Spatial Point Patterns*, Volume 70. John Wiley & Sons.

Jalilian, A., Y. Guan, and R. Waagepetersen (2013). Decomposition of variance for spatial Cox processes. *Scandinavian Journal of Statistics 40*(1), 119–137.

Jónsdóttir, K. Ý., A. Rønn-Nielsen, K. Mouridsen, and E. B. Vedel Jensen (2013). Lévy-based modelling in brain imaging. *Scandinavian Journal of Statistics 40*(3), 511–529.

Kopeckỳ, J. and T. Mrkvička (2016). On the Bayesian estimation for the stationary Neyman-Scott point processes. *Applications of Mathematics 61*(4), 503–514.

Lieshout, M. N. M. van and A. J. Baddeley (1996). A nonparametric measure of spatial interaction in point patterns. *Statistica Neerlandica 50*(3), 344–361.

Lieshout, M. N. M. van and A. J. Baddeley (2002). *Extrapolating and interpolating spatial patterns*. In: A. B. Lawson and D. G. T. Denison (Eds.), Spatial Cluster Modelling, Chapman & Hall/CRC, Boca Raton, 61-86.

Lund, J. and M. Rudemo (2000). Models for point processes observed with noise. *Biometrika*, 235–249.

Møller, J. and A. D. Christoffersen. Pair correlation functions and limiting distributions of iterated cluster point processes. *In preparation*.

Møller, J. and G. L. Torrisi (2005). Generalised shot noise Cox processes. *Advances in Applied Probability 37*(01), 48–74.

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

Neyman, J. and E. L. Scott (1958). Statistical approach to problems of cosmology. *Journal of the Royal Statistical Society. Series B (Methodological)*, 1–43.

Redenbach, C., A. Särkkä, and M. Sormani (2015). Classification of points in superpositions of Strauss and Poisson processes. *Spatial Statistics 12*, 81–95.

Rubin-Delanchy, P., G. L. Burn, J. Griffié, D. J. Williamson, N. A. Heard, A. P. Cope, and D. M. Owen (2015). Bayesian cluster identification in single-molecule localization microscopy data. *Nature Methods 12*(11), 1072–1078.

Sengupta, P., T. Jovanovic-Talisman, and J. Lippincott-Schwartz (2013). Quantifying spatial organization in point–localization superresolution images using pair correlation analysis. *Nature Protocols 8*(2), 345–354.

Sengupta, P., T. Jovanovic-Talisman, D. Skoko, M. Renz, S. L. Veatch, and J. Lippincott-Schwartz (2011). Probing protein heterogeneity in the plasma membrane using PALM and pair correlation analysis. *Nature Methods 8*(11), 969–975.

Shimatani, K. (2001). Multivariate point processes and spatial variation of species diversity. *Forest Ecology and Management 142*(1), 215–229.

Veatch, S. L., B. B. Machta, S. A. Shelby, E. N. Chiang, D. A. Holowka, and B. A. Baird (2012). Correlation functions quantify super-resolution images and estimate apparent clustering due to over-counting. *PloS ONE 7*(2), e31457.

Wiegand, T., S. Gunatilleke, N. Gunatilleke, and T. Okuda (2007). Analyzing the spatial structure of a Sri Lankan tree species with multiple scales of clustering. *Ecology 88*(12), 3088–3102.

Appendix A

In order to derive the second-order factorial moment measure $\alpha_Z^{(2)}$ of Z, let C_1, C_2 be two Borel sets in $\mathbb{R}^d$. Then,

$$\begin{aligned}
&\alpha_Z^{(2)}(C_1 \times C_2) \\
&\quad = \mathbb{E} \sum_{z_1,z_2 \in Z}^{\neq} \mathbb{1}(z_1 \in C_1, z_2 \in C_2) \\
&\quad = \mathbb{E} \sum_{x \in X} \sum_{y_1,y_2 \in Y_x}^{\neq} \mathbb{1}(x + y_1 \in C_1, x + y_2 \in C_2) \\
&\qquad + \mathbb{E} \sum_{x_1,x_2 \in X}^{\neq} \Big(\sum_{y_1 \in Y_{x_1}} \mathbb{1}(x_1 + y_1 \in C_1) \Big) \Big(\sum_{y_2 \in Y_{x_2}} \mathbb{1}(x_2 + y_2 \in C_2) \Big),
\end{aligned}$$

where the last equality sign holds since the processes $x + Y_x$, $x \in X$, do not overlap, as assumed in Section 3. Using the conditional independence of the processes $\{Y_x : x \in X\}$, given X, we get

$$\begin{aligned}
&\alpha_Z^{(2)}(C_1 \times C_2) \\
&\quad = \mathbb{E} \sum_{x \in X} \int_{\mathbb{R}^d \times \mathbb{R}^d} \mathbb{1}(y_1 \in C_1 - x, y_2 \in C_2 - x) \rho_Y^{(2)}(y_1, y_2) \, \mathrm{d}y_2 \, \mathrm{d}y_1 \\
&\qquad + \tau^2 \mathbb{E} \sum_{x_1,x_2 \in X}^{\neq} \left(\int_{C_1} k(y_1 - x_1) \, \mathrm{d}y_1 \right) \left(\int_{C_2} k(y_2 - x_2) \, \mathrm{d}y_2 \right).
\end{aligned}$$

It follows that

$$\begin{aligned}
&\alpha_Z^{(2)}(C_1 \times C_2) \\
&\quad = \rho_X \int_{\mathbb{R}^d} \int_{\mathbb{R}^d \times \mathbb{R}^d} \mathbb{1}(y_1 \in C_1 - x, y_2 \in C_2 - x) \rho_Y^{(2)}(y_1, y_2) \, \mathrm{d}y_2 \, \mathrm{d}y_1 \, \mathrm{d}x \\
&\qquad + \tau^2 \int_{C_1} \int_{C_2} \int_{\mathbb{R}^d \times \mathbb{R}^d} k(y_1 - x_1) k(y_2 - x_2) \rho_X^{(2)}(x_1, x_2) \, \mathrm{d}x_2 \, \mathrm{d}x_1 \, \mathrm{d}y_2 \, \mathrm{d}y_1.
\end{aligned}$$

Since the first term can be rewritten as

$$\rho_X \int_{C_1} \int_{C_2} \int_{\mathbb{R}^d} \rho_Y^{(2)}(z_1 - x, z_2 - x) \, \mathrm{d}x \, \mathrm{d}z_2 \, \mathrm{d}z_1,$$

the second-order product density of Z takes the form

$$\begin{aligned}
\rho_Z^{(2)}(z_1, z_2) = \rho_X &\int_{\mathbb{R}^d} \rho_Y^{(2)}(z_1 - x, z_2 - x) \, \mathrm{d}x \\
&+ \tau^2 \int_{\mathbb{R}^d \times \mathbb{R}^d} k(z_1 - x_1) k(z_2 - x_2) \rho_X^{(2)}(x_1, x_2) \, \mathrm{d}x_2 \, \mathrm{d}x_1,
\end{aligned}$$

or, equivalently, the pair correlation function of Z becomes

$$\begin{aligned}
g_Z(z_1, z_2) = \frac{1}{\rho_X} &\int_{\mathbb{R}^d} \frac{\rho_Y^{(2)}(z_1 - x, z_2 - x)}{\tau^2} \, \mathrm{d}x \\
&+ \int_{\mathbb{R}^d \times \mathbb{R}^d} k(z_1 - x_1) k(z_2 - x_2) g_X(x_1, x_2) \, \mathrm{d}x_2 \, \mathrm{d}x_1.
\end{aligned}$$

If $g_X(x_1, x_2) = g_X(x_1 - x_2)$, say, the second term takes the form

$$\int_{\mathbb{R}^d} h_k(z_1 - z_2 - x) g_X(x) \, \mathrm{d}x,$$

where

$$h_k(z) = \int_{\mathbb{R}^d} k(z + x) k(x) \, \mathrm{d}x.$$

Putting these results together, we obtain (3.1).

Let N be the number of points in Y. Then, $\rho_Y^{(2)}$ takes a simple form if, conditional on N, the points in Y are i.i.d. with probability density k. Note that $\mathbb{E}(N) = \tau$. The second factorial moment measure $\alpha_Y^{(2)}$ of Y becomes

$$\begin{aligned}
\alpha_Y^{(2)}(C_1 \times C_2) &= \mathbb{E} \sum_{y_1, y_2 \in Y}^{\neq} \mathbb{1}(y_1 \in C_1, y_2 \in C_2) \\
&= \sum_{n=2}^{\infty} \mathbb{P}(N = n) \mathbb{E}\Big(\sum_{y_1, y_2 \in Y}^{\neq} \mathbb{1}(y_1 \in C_1, y_2 \in C_2) \;\Big|\; N = n \Big) \\
&= \sum_{n=2}^{\infty} \mathbb{P}(N = n) n(n-1) \int_{C_1} k(y_1) \, \mathrm{d}y_1 \int_{C_2} k(y_2) \, \mathrm{d}y_2 \\
&= \big(\mathbb{E}(N^2) - \mathbb{E}(N) \big) \int_{C_1} k(y_1) \, \mathrm{d}y_1 \int_{C_2} k(y_2) \, \mathrm{d}y_2.
\end{aligned}$$

We find

$$g_Y(y_1, y_2) = \big(\mathbb{E}(N^2) - \mathbb{E}(N) \big) / \tau^2,$$

and (3.3) follows. If Y is a Poisson process, $g_Y(y_1, y_2) \equiv 1$ and the pair correlation function of Z becomes

$$g_Z(z_1, z_2) = \rho_X^{-1} h_k(z_1 - z_2) + \int_{\mathbb{R}^d} h_k(z_1 - z_2 - x) g_X(x) \, \mathrm{d}x.$$

Appendix B

All estimates of summary functions were obtained using default settings of functions in the `spatstat` library of `R` (Baddeley et al., 2015). The parameter estimates were obtained with the `mincontrast` algorithm using tuning parameters $r_1 = 10$, $r_2 = 300$, $q = 1/4$, $p = 2$ and `startpar` with `kappa = 0.00001`, `mu = 5`, `omega = 75` and `sigma = 20` (or `eta = 75`). The algorithm was tested for several choices of `startpar`. The chosen `startpar` gave the best fit for both models, and the obtained estimates were robust in the sense that several combinations of `startpar` led to approximately the same results. Furthermore, the results were robust to alternative choices of r_1 and kernel bandwidth. This was tested because it is well-known, that the algorithm may be affected by a bias near $r = 0$ due to kernel smoothing. The parameter estimates were supplemented by estimated relative mean squared error (RMSE) and estimated relative bias (Rbias). These estimates were obtained by re-estimation of the parameters based on 200 simulations from the fitted models, using same tuning parameters. Envelope tests were performed with the `devtools` and `spptest` libraries of `R`, using `rank_envelopes` based on 2000 simulations.