HAL open science

Palm distributions for log Gaussian Cox processes

Jean-François Coeurjolly, Jesper Møller, Rasmus Waagepetersen

To cite this version:

Jean-François Coeurjolly, Jesper Møller, Rasmus Waagepetersen. Palm distributions for log Gaussian Cox processes. 2015. hal-01163672v3

HAL Id: hal-01163672

https://hal.science/hal-01163672v3

Preprint submitted on 21 Dec 2015 (v3), last revised 6 Jun 2016 (v4)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Palm distributions for log Gaussian Cox processes

Jean-François Coeurjolly[1], Jesper Møller[2] and Rasmus Waagepetersen[2]

[1]Laboratory Jean Kuntzmann, Grenoble Alpes University, France, email: Jean-Francois.Coeurjolly@upmf-grenoble.fr.
[2]Department of Mathematical Sciences, Aalborg University, Fredrik Bajers Vej 7E, DK-9220 Aalborg, email: jm@math.aau.dk, rw@math.aau.dk

December 21, 2015

Abstract

This paper establishes a remarkable result regarding Palm distributions for a log Gaussian Cox process: the reduced Palm distribution for a log Gaussian Cox process is itself a log Gaussian Cox process which only differs from the original log Gaussian Cox process in the intensity function. This new result is used to study functional summaries for log Gaussian Cox processes.

Keywords: J-function; joint intensities; Laplace approximation; nearest-neighbour distribution function; spatial point process.

1 Introduction

Palm distributions (see e.g. Møller and Waagepetersen, 2004; Daley and Vere-Jones, 2008) are important in the theory and application of spatial point processes. Intuitively speaking, for a prespecified location in space, the Palm distribution of a point process, with respect to this location, plays the role of the conditional distribution of the point process given that the aforementioned location is occupied by a point in the point process.

The present paper focuses on log Gaussian Cox processes (Møller et al., 1998) which provide very flexible, useful, and popular models for modeling spatial patterns in e.g. biology and spatial epidemiology. The paper establishes a surprisingly simple characterization of Palm distributions for such a process: The reduced n-point Palm distribution is for any $n \geq 1$ itself a log Gaussian Cox process that only differs from the original log Gaussian Cox process in its intensity function. This result can be exploited for functional summaries as discussed later. The simplicity and completeness of this result is remarkable when compared with Palm distributions for other common classes of spatial point processes. Reduced Palm distributions for Gibbs point processes are also themselves Gibbs point processes but with densities only known up to a normalizing constant. For shot-noise Cox processes (Møller, 2003) one-point reduced Palm distributions have a simple characterization as cluster processes similar to shot-noise Cox processes but this is not the case for n-point Palm distributions when $n > 1$.

The paper is organized as follows. Section 2 reviews the general definition of reduced Palm distributions of any order and relates this to Cox processes. Section 3 establishes our characterization result for log Gaussian Cox processes. Section 4 applies this result to functional summaries for stationary log Gaussian Cox processes, in particular the so-called F, G, and J-functions, where we establish some new theoretical results, consider how to calculate F, G, and J using Laplace approximations, and discuss an application. Section 5 concludes the paper.

2 Palm distributions

Our general setting is as follows. For ease of exposition we view a point process as a random locally finite subset $\mathbf{X}$ of a Borel set $S \subseteq \mathbb{R}^d$; for measure theoretical details, see e.g. Møller and Waagepetersen (2004) or Daley and Vere-Jones (2003). Denoting $\mathbf{X}_B = \mathbf{X} \cap B$ the restriction of $\mathbf{X}$ to a set $B \subseteq S$, local finiteness of $\mathbf{X}$ means that $\mathbf{X}_B$ is finite almost surely (a.s.) whenever B is bounded. We denote $\mathcal{N}$ the state space consisting of the locally finite subsets (or point configurations) of S. We use the generic notation h for an arbitrary non-negative measurable function defined on $\mathcal{N}$, S^n, or $S^n \times \mathcal{N}$ for $n = 1, 2, \ldots$. Furthermore, $\mathcal{B}_0$ is the family of all bounded Borel subsets of S. Finally, recall that the void probabilities $\mathrm{P}(\mathbf{X}_K = \emptyset)$, $K \subseteq S$ compact, uniquely determine the distribution of $\mathbf{X}$.

2.1 Factorial moment measures and Palm distributions

This section provides the general definition of reduced Palm distributions of any order. For finite point processes specified by a density, a simpler and more explicit definition is available as reviewed in Coeurjolly et al. (2015).

For $n = 1, 2, \ldots$ and $B_i \in \mathcal{B}_0$, the nth order factorial moment measure $\alpha^{(n)}$ is defined by

$$\alpha^{(n)}(B_1 \times B_2 \times \cdots \times B_n) = \mathrm{E} \sum_{x_1,\ldots,x_n \in \mathbf{X}}^{\neq} 1(x_1 \in B_1, \ldots, x_n \in B_n),$$

where $1(\cdot)$ denotes the indicator function and $\neq$ over the summation sign means that $x_1, \ldots, x_n$ are pairwise distinct. If $\alpha^{(n)}$ has a density $\rho^{(n)}$ with respect to Lebesgue measure, $\rho^{(n)}$ is called the nth order joint intensity function and is determined up to a Lebesgue nullset. Therefore, we can assume that $\rho^{(n)}(x_1, \ldots, x_n)$ is invariant under permutations of $x_1, \ldots, x_n$, and we need only to consider the case where $x_1, \ldots, x_n \in S$ are pairwise distinct. Then $\rho^{(n)}(x_1, \ldots, x_n)\, \mathrm{d}x_1 \cdots \mathrm{d}x_n$ can be interpreted as the approximate probability for $\mathbf{X}$ having a point in each of infinitesimally small regions around $x_1, \ldots, x_n$ of volumes $\mathrm{d}x_1, \ldots \mathrm{d}x_n$, respectively.

Moreover, for any measurable $F \subseteq \mathcal{N}$, define the nth order reduced Campbell measure $C^{(n)!}$ as the measure on $S^n \times \mathcal{N}$ given by

$$C^{(n)!}(B_1 \times B_2 \times \cdots \times B_n \times F) = \\ \mathrm{E} \sum_{x_1,\ldots,x_n \in \mathbf{X}}^{\neq} 1(x_1 \in B_1, \ldots, x_n \in B_n, \mathbf{X} \setminus \{x_1, \ldots, x_n\} \in F).$$

Note that $C^{(n)!}(\cdot \times F)$, as a measure on S^n, is absolutely continuous with respect to $\alpha^{(n)}$, with a density $P^!_{x_1,\ldots,x_n}(F)$ which is determined up to an $\alpha^{(n)}$ nullset, and so we can assume that $P^!_{x_1,\ldots,x_n}(\cdot)$ is a point process distribution on $\mathcal{N}$, called the nth order reduced Palm distribution given $x_1, \ldots, x_n$. We denote by $\mathbf{X}^!_{x_1,\ldots,x_n}$ a point process distributed according to $P^!_{x_1,\ldots,x_n}$. Again we need only to consider the case where $x_1, \ldots, x_n$ are pairwise distinct. Then $P^!_{x_1,\ldots,x_n}$ can be interpreted as the conditional distribution of $\mathbf{X} \setminus \{x_1, \ldots, x_n\}$ given that $x_1, \ldots, x_n \in \mathbf{X}$.

If $\rho^{(n)}$ exists, then by standard measure theoretical arguments we obtain

$$\begin{aligned}&\mathrm{E}\sum_{x_1,\ldots,x_n\in\mathbf{X}}^{\neq} h(x_1,\ldots,x_n,\mathbf{X}\setminus\{x_1,\ldots,x_n\})\\&=\int_S\cdots\int_S \mathrm{E}h(x_1,\ldots,x_n,\mathbf{X}^!_{x_1,\ldots,x_n})\rho^{(n)}(x_1,\ldots,x_n)\,\mathrm{d}x_1\cdots\mathrm{d}x_n. \qquad (1)\end{aligned}$$

Suppose $\rho^{(m+n)}$ exists for an $m\geq 1$ and $n\geq 1$. Then, for pairwise distinct $u_1,\ldots,u_m,x_1,\ldots,x_n\in S$, it follows easily by expressing $\alpha^{(m+n)}$ as an expectation of the form (1) that $\mathbf{X}^!_{x_1,\ldots,x_n}$ has mth order joint intensity function

$$\rho^{(m)}_{x_1,\ldots,x_n}(u_1,\ldots,u_m)=\begin{cases}\frac{\rho^{(m+n)}(u_1,\ldots,u_m,x_1,\ldots,x_n)}{\rho^{(n)}(x_1,\ldots,x_n)} & \text{if } \rho^{(n)}(x_1,\ldots,x_n)>0,\\ 0 & \text{otherwise.}\end{cases} \qquad (2)$$

2.2 Cox processes

Let $\mathbf{\Lambda}=\{\Lambda(x)\}_{x\in S}$ be a nonnegative random field such that $\mathbf{\Lambda}$ is locally integrable a.s., that is, for any $B\in\mathcal{B}_0$, the integral $\int_B\Lambda(x)\,\mathrm{d}x$ exists and is finite a.s. In the sequel, $\mathbf{X}$ conditional on $\mathbf{\Lambda}$ is assumed to be a Poisson process with intensity function $\mathbf{\Lambda}$; we say that $\mathbf{X}$ is a Cox process driven by $\mathbf{\Lambda}$. We also assume that $\mathbf{\Lambda}$ has moments of any order $n=1,2,\ldots$. Then the joint intensities of $\mathbf{X}$ exist: For any $n=1,2,\ldots$ and pairwise distinct $x_1,\ldots,x_n\in S$,

$$\rho^{(n)}(x_1,\ldots,x_n)=\mathrm{E}\left\{\prod_{i=1}^n\Lambda(x_i)\right\}. \qquad (3)$$

The following lemma, which is verified in Appendix A, gives a characterization of the reduced Palm distributions and their void probabilities.

Lemma 1. *Let $\mathbf{X}$ be a Cox process satisfying the conditions above. Then, for any $n=1,2,\ldots$, pairwise distinct $x_1,\ldots,x_n\in S$, and compact $K\subseteq S$,*

$$\rho^{(n)}(x_1,\ldots,x_n)\mathrm{E}\left\{h\left(x_1,\ldots,x_n,\mathbf{X}^!_{x_1,\ldots,x_n}\right)\right\}=\mathrm{E}\left\{h(x_1,\ldots,x_n,\mathbf{X})\prod_{i=1}^n\Lambda(x_i)\right\} \qquad (4)$$

and

$$\rho^{(n)}(x_1,\ldots,x_n)\mathrm{P}(\mathbf{X}^!_{x_1,\ldots,x_n}\cap K=\emptyset)=\mathrm{E}\left[\exp\left\{-\int_K\Lambda(u)\,\mathrm{d}u\right\}\prod_{i=1}^n\Lambda(x_i)\right]. \qquad (5)$$

3 Reduced Palm distributions for log Gaussian Cox processes

For the remainder of this paper, let $\mathbf{X}$ be a Cox process driven by $\mathbf{\Lambda} = \{\Lambda(x)\}_{x\in S}$, where $\Lambda(x) = \exp\{Y(x)\}$ and $\mathbf{Y} = \{Y(x)\}_{x\in S}$ is a Gaussian process with mean function μ and covariance function c so that $\mathbf{\Lambda}$ is locally integrable a.s. (simple conditions ensuring this are given in Møller et al., 1998). Then $\mathbf{X}$ is a log Gaussian Cox process (LGCP) as introduced by Coles and Jones (1991) in astronomy and independently by Møller et al. (1998) in statistics.

For $x, y \in S$, define the so-called pair correlation function by

$$g(x,y) = \begin{cases} \rho^{(2)}(x,y)/\{\rho(x)\rho(y)\} & \text{if } x \neq y \\ \exp\{c(x,x)\} & \text{if } x = y \end{cases}$$

taking $0/0 = 0$ in the case $x \neq y$. By Møller et al. (1998, Theorem 1),

$$\rho(x) = \exp\{\mu(x) + c(x,x)/2\}, \qquad g(x,y) = \exp\{c(x,y)\}, \tag{6}$$

and for pairwise distinct $x_1, \ldots, x_n \in S$,

$$\rho^{(n)}(x_1, \ldots, x_n) = \left\{\prod_{i=1}^{n} \rho(x_i)\right\}\left\{\prod_{1\le i<j\le n} g(x_i, x_j)\right\} \tag{7}$$

is strictly positive.

For $u, x_1, \ldots, x_n \in S$, define

$$\mu_{x_1,\ldots,x_n}(u) = \mu(u) + \sum_{i=1}^{n} c(u, x_i).$$

Combining (2) and (6)-(7), we obtain for any pairwise distinct $u_1, \ldots, u_m, x_1, \ldots, x_n \in S$ with $m > 0$ and $n > 0$,

$$\rho^{(m)}_{x_1,\ldots,x_n}(u_1, \ldots, u_m) = \left\{\prod_{i=1}^{m} \rho_{x_1,\ldots,x_n}(u_i)\right\}\left\{\prod_{1\le i<j\le m} g(u_i, u_j)\right\}, \tag{8}$$

where

$$\rho^{(m)}_{x_1,\ldots,x_n}(u_1, \ldots, u_m) = \exp\left\{\mu_{x_1,\ldots,x_n}(u) + c(u,u)/2\right\}.$$

Thereby the following proposition follows.

Proposition 1. *For the LGCP* $\mathbf{X}$ *and any pairwise distinct* $x_1, \ldots, x_n \in S$, $\mathbf{X}^!_{x_1,\ldots,x_n}$ *has m'th order joint intensity* (8) *which agrees with the m'th order joint intensity function for an LGCP with mean function* $\mu_{x_1,\ldots,x_n}$ *and covariance function c for the underlying Gaussian process.*

Proposition 1 indicates that also $\mathbf{X}^!_{x_1,\ldots,x_n}$ could be an LGCP. A sufficient condition, considered by Macchi (1975), is the existence of a number $a = a(B) > 1$ for each set $B \in \mathcal{B}_0$ such that

$$\mathrm{E}\left[\exp\left\{a \int_B \Lambda(u)\,\mathrm{d}u\right\}\right] < \infty. \tag{9}$$

However, we have not been successful in verifying this condition which seems too strong to hold for any of the covariance function models we have considered, including when c is constant (then $\mathbf{X}$ is a mixed Poisson process) or weaker cases of correlation, e.g. if c is a stationary exponential covariance function. The case where c is constant is closely related to the log normal distribution which is not uniquely determined by its moments (Heyde, 1963).

Accordingly we use instead Lemma 1 when establishing the following theorem, which implies that the LGCPs $\mathbf{X}$ and $\mathbf{X}^!_{x_1,\ldots,x_n}$ share the same pair correlation function and differ only in their intensity functions.

Theorem 1. *For pairwise distinct* $x_1, \ldots, x_n \in S$, $\mathbf{X}^!_{x_1,\ldots,x_n}$ *is an LGCP with underlying Gaussian process* $\mathbf{Y}_{x_1,\ldots,x_n}$ *where* $\mathbf{Y}_{x_1,\ldots,x_n}$ *has mean function* $\mu_{x_1,\ldots,x_n}$ *and covariance function c.*

Let $\tilde{\mathbf{Y}} = \mathbf{Y} - \mu$ be the centered Gaussian process with covariance function c. Theorem 1 is a consequence of the fact that $\mathbf{Y}_{x_1,\ldots,x_n}$ is absolutely continuous with respect to $\tilde{\mathbf{Y}}$, with density $\exp\left\{\sum_{i=1}^n \tilde{y}(x_i) - \sum_{i,j=1}^n c(x_i, x_j)/2\right\}$ when $\tilde{\mathbf{y}}$ is a realization of $\tilde{\mathbf{Y}}$. This result is related to the Cameron-Martin-Girsanov formula for one-dimensional Gaussian processes. A short selfcontained proof covering our spatial setting is given in Appendix A.

Often we consider a non-negative covariance function c or equivalently $g \geq 1$, which is interpreted as 'attractiveness of the LGCP at all ranges', but even more can be said: A coupling between $\mathbf{X}$ and $\mathbf{X}^!_{x_1,\ldots,x_n}$ is obtained by taking $Y_{x_1,\ldots,x_n}(x) = Y(x) + \sum_{i=1}^n c(x, x_i)$. Thus, if $c \geq 0$ and we are given pairwise distinct points $x_1, \ldots, x_n \in S$, we can consider $\mathbf{X}$ as being included in $\mathbf{X}^!_{x_1,\ldots,x_n}$, since $\mathbf{X}$ can be obtained by an independent thinning of $\mathbf{X}^!_{x_1,\ldots,x_n}$, with inclusion probabilities $\exp\{-\sum_{i=1}^n c(x, x_i)\}$, $x \in X \setminus \{x_1, \ldots, x_n\}$. This

property clearly shows the attractiveness of the LGCP if $c \geq 0$ (equivalently $g \geq 1$).

4 Functional summary statistics for stationary log Gaussian Cox processes

Throughout this section, let $S = \mathbb{R}^d$ and assume that the LGCP $\mathbf{X}$ is stationary, i.e., its distribution is invariant under translations in $\mathbb{R}^d$. By (6)-(7), this is equivalent to stationarity of the underlying Gaussian process $\mathbf{Y}$, that is, the intensity ρ is constant and the pair correlation function $g(x, y) = \tilde{g}(x-y)$ is translation invariant, where $x, y \in \mathbb{R}^d$, $\tilde{g}(x) = \exp\{\tilde{c}(x)\}$, $\tilde{c}(x) = c(o, x)$, and where o denotes the origin in $\mathbb{R}^d$. It is custom to call $\mathrm{P}_o^!$ the reduced Palm distribution at a typical point, noticing that for any $x \in \mathbb{R}^d$, $\mathbf{X}_o^!$ and $\mathbf{X}_x^! - x = \{y - x : y \in \mathbf{X}_x^!\}$ are identically distributed.

Denote $B(o, r)$ the ball in $\mathbb{R}^d$ of radius $r > 0$ and centered at o. Popular tools for exploratory purposes as well as model fitting and model checking are based on the following functional summaries: The pair correlation function $\tilde{g}(x)$, usually provided it depends only on the distance $\|x\|$, and the related Ripley's K-function given by

$$K(r) = \frac{1}{\rho}\mathrm{E}\,\#\left\{\mathbf{X}_o^! \cap B(o, r)\right\} = \int_{B(o,r)} \tilde{g}(x)\,\mathrm{d}x,$$

so $\rho K(r)$ is the expected number of r-close points to a typical point in $\mathbf{X}$; the empty space function

$$F(r) = \mathrm{P}\left\{\mathbf{X} \cap B(o, r) \neq \emptyset\right\},$$

i.e., the probability of $\mathbf{X}$ having a point in a ball placed at an arbitrary fixed location; the nearest-neighbour distribution function

$$G(r) = \mathrm{P}\left\{\mathbf{X}_o^! \cap B(o, r) \neq \emptyset\right\},$$

which is the probability of having an r-close point to a typical point in $\mathbf{X}$; and the J-function

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

with the convention $a/0 = 0$ for any $a \geq 0$; see e.g. Møller and Waagepetersen (2004). Section 4.1 establishes some new results for these theoretical functions and Section 4.2 discusses how they can be calculated using a Laplace approximation. Section 4.3 illustrates this calculation and Section 4.4 discusses an application for a real dataset.

4.1 New formulae

By conditioning on $\mathbf{Y}$, we see that

$$1 - F(r) = \mathrm{E}\left(\exp\left[-\int_{B(o,r)} \exp\{Y(x)\}\,\mathrm{d}x\right]\right). \tag{10}$$

Using the Slivnyak-Mecke formula, Møller et al. (1998) showed that

$$1 - G(r) = \frac{1}{\rho}\,\mathrm{E}\left(\exp\left[Y(o) - \int_{B(o,r)} \exp\{Y(x)\}\,\mathrm{d}x\right]\right). \tag{11}$$

Since the nearest-neighbour distribution function for $\mathbf{X}$ is the same as the empty space function for $\mathbf{X}_o^!$, which is an LGCP with underlying Gaussian process $Y_o(x) = Y(x) + \tilde{c}(x)$, and since $\tilde{g}(x) = \exp\{\tilde{c}(x)\}$, we obtain an alternative expression

$$1 - G(r) = \mathrm{E}\left(\exp\left[-\int_{B(o,r)} \tilde{g}(x)\exp\{Y(x)\}\,\mathrm{d}x\right]\right). \tag{12}$$

Therefore, we also obtain a new expression for the J-function,

$$J(r) = \frac{\mathrm{E}\left(\exp\left[-\int_{B(o,r)} \tilde{g}(x)\exp\{Y(x)\}\,\mathrm{d}x\right]\right)}{\mathrm{E}\left(\exp\left[-\int_{B(o,r)} \exp\{Y(x)\}\,\mathrm{d}x\right]\right)}. \tag{13}$$

Van Lieshout (2011) established for a general stationary point process the approximation $J(r) - 1 \approx -\rho\{K(r) - \omega_d r^d\}$, where $\omega_d = |B(o,1)|$ is Ripley's K-function for a stationary Poisson process, and it is therefore not so surprising that often empirical J and K-functions lead to the same practical interpretations. In particular, if for our LGCP $\tilde{c} \geq 0$, i.e. $\tilde{g} \geq 1$, then we have $K(r) - \omega_d r^d \geq 0$, and so we expect that $J(r) \leq 1$. Indeed Van Lieshout (2011) verified this in the case of a stationary LGCP with $\tilde{g} \geq 1$. This result immediately follows by the new expression (13).

4.2 Laplace approximation

This section discusses a Laplace approximation of $1 - G(r)$; the Laplace approximation of $1 - F(r)$ is similar.

For $\Delta > 0$, define $\mathcal{G}(\Delta) = \{(\Delta i_1, \ldots, \Delta i_d) \,|\, i_1, \ldots, i_d \in \mathbb{Z}\}$ as the grid of quadrature points. Further, for $v \in \mathcal{G}(\Delta)$, let $A_v^\Delta = [v_1 - \Delta/2, v_1 + \Delta/2[\times \cdots \times [v_d - \Delta/2, v_d + \Delta/2[$ be the grid cell associated with v. For any non-negative Borel function $\ell : \mathbb{R}^d \to \mathbb{R}$, we use the approximation

$$\int_{B(o,r)} \exp\{Y(x)\}\ell(x)\,\mathrm{d}x \approx \sum_{v \in \mathcal{G}(\Delta) \cap B(o,r)} w_v \ell(v) \exp\{Y(v)\}, \tag{14}$$

where the quadrature weight $w_v = |A_v^\Delta \cap B(o,r)|$.

Denote by M and Σ the mean vector and the covariance matrix of the normally distributed vector $\{Y(v)\}_{v \in \mathcal{G}(\Delta) \cap B(o,r)}$. Then (12) and (14) give

$$1 - G(r) \approx \int \exp\{h(y)\}\,\mathrm{d}y \tag{15}$$

where y is the vector $(y_v)_{v \in \mathcal{G}(\Delta) \cap B(o,r)}$ of dimension $m = \#\{\mathcal{G}(\Delta) \cap B(o,r)\}$ and

$$h(y) = -\sum_v w_v \tilde{g}(v) \exp(y_v) - \frac{1}{2}(y - M)^\top \Sigma^{-1}(y - M) - \frac{1}{2}\log\{(2\pi)^m |\Sigma|\}.$$

The gradient vector for h is

$$\nabla h(y) = -d(y) - \Sigma^{-1}(y - M), \tag{16}$$

where $d(y) = \{w_v \tilde{g}(v) \exp(y_v)\}_{v \in \mathcal{G}(\Delta) \cap B(o,r)}$, and the negated Hessian matrix for h is

$$H(y) = D(y) + \Sigma^{-1},$$

where $D(y) = \mathrm{diag}\{d(y)_v, v \in \mathcal{G}(\Delta) \cap B(o,r)\}$. Since $H(y)$ is a positive definite matrix, h has a unique maximum $\hat{y}$ which can be found using Newton-Raphson iterations

$$y^{(l+1)} = y^{(l)} + H^{-1}\{y^{(l)}\}\nabla h\{y^{(l)}\}. \tag{17}$$

Therefore, the logarithm of the Laplace approximation of the right hand side in (15) (see e.g. Stigler (1986)) gives

$$\log\{1-G(r)\} \approx -\sum_{v\in\mathcal{G}(\Delta)\cap B(o,r)} w_v \tilde{g}(v)\exp(\hat{y}_v) + \frac{1}{2}(\hat{y}-M)^\top d(\hat{y}) - \frac{1}{2}\log|D(\hat{y})\Sigma + I|. \tag{18}$$

where I is the $m \times m$ identity matrix. For the computation of $\Sigma^{-1}(y-M)$ in (16) we solve $LL^\top z = y - M$ where L is the Cholesky factor of Σ. In the same way, considering the QR decomposition of the matrix $D(y)\Sigma + I$, the computation of $H^{-(1)}\{y^{(l)}\}\nabla h\{y^{(l)}\}$ in (17) is done by first solving $Q\{y^{(l)}\}R\{y^{(l)}\}\tilde{z} = \nabla h\{y^{(l)}\}$ and second by evaluating $\Sigma\tilde{z}$. Finally, in (18), $|D(\hat{y})\Sigma + I| = |R(\hat{y})|$.

4.3 Numerical illustration

To illustrate the Laplace approximations of the G and J-functions (Section 4.2) we consider three planar stationary LGCPs with intensity $\rho = 50$ and spherical covariance function

$$\tilde{c}(x) = \begin{cases} \sigma^2\left[1 - \frac{2}{\pi}\left\{\frac{\|x\|}{\alpha}\sqrt{1-\left(\frac{\|x\|}{\alpha}\right)^2} + \sin^{-1}\frac{\|x\|}{\alpha}\right\}\right] & \text{if } \|x\| \le \alpha, \\ 0 & \text{otherwise,} \end{cases}$$

with variance $\sigma^2 = 4$ and scale parameters $\alpha = 0.1, 0.2, 0.3$, respectively. We evaluate the approximations of $G(r)$ and $J(r)$ at $r \in \mathcal{R}$, where $\mathcal{R}$ is a regular grid of 50 values between 0.01 and 0.25. For $r \in \mathcal{R}$, we define the grid $\mathcal{G}(\Delta_r)$ with $\Delta_r = 2r/q$, where q is a fixed integer. Such a choice implies that $\#\{\mathcal{G}(\Delta_r) \cap [-r,r]^2\} = q^2$, and so we have at least $q^2\pi/4$ quadrature points in $B(o,r)$. For a given q, denote by G_q, F_q, and J_q the corresponding Laplace approximations of G, F, and J, respectively. Figure 1 shows the resulting curves with $q = 16$. To see how far these Cox processes deviate from the Poisson case (which would correspond to $\sigma^2 = 0$), we also plot the G-function in the Poisson case, namely $1-G(r) = \exp(-\rho\pi r^2)$. To study the role of q, we report in Table 1 the maximal differences $\max_{r\in\mathcal{R}}|G_{16}(r)-G_q(r)|$ and $\max_{r\in\mathcal{R}}|J_{16}(r)-J_q(r)|$ for $q = 4, 8, 12$. As expected, each difference is decreasing with q and is already very small for $q = 12$ (less than 4×10^{-3}

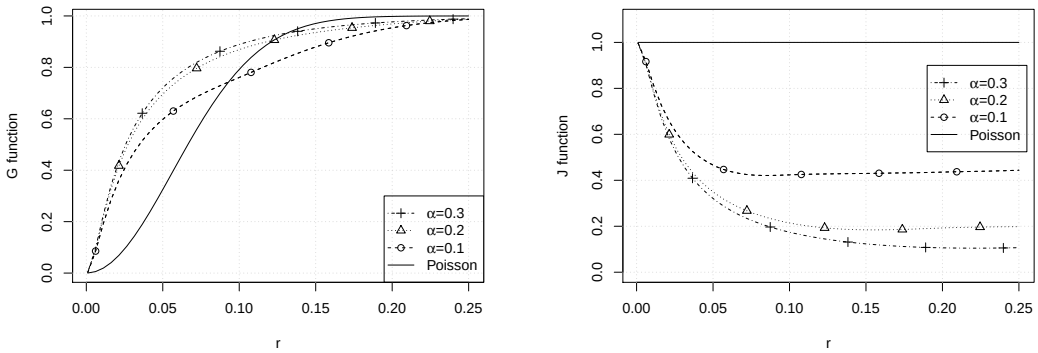

Figure 1: For three planar stationary LGCPs with intensity $\rho = 50$ and $\tilde{c}$ given by a spherical covariance function, with variance $\sigma^2 = 4$ and scale parameters $\alpha = 0.1, 0.2, 0.3$, respectively, Laplace approximations of the G-function (left) and the J-function (right).

except for the J-function and $\alpha = 1$). This justifies our choice $q = 16$ in Figure 1.

	$\max_{r\in\mathcal{R}} \lvert H_{16}(r) - H_q(r)\rvert$, $H = G, J$		
	$q = 4$	$q = 8$	$q = 12$
$\alpha = 0.1$, G	59.9	8.4	2.1
J	505.9	96.1	20.5
$\alpha = 0.2$, G	14.3	1.6	0.5
J	109.0	13.8	3.5
$\alpha = 0.3$, G	4.2	0.5	0.1
J	22.1	3.1	0.3

Table 1: For the same three LGCPs as in Figure 1, maximal differences between the Laplace approximations G_q and G_{16}, and between the Laplace approximations J_q and J_{16}, with $q = 4, 8, 12$. Results are multiplied by 10^3.

The Laplace approximation of the G-function could also be derived using the classical formula (11). To check the agreement of the numerical approximations based on equations (11) and (12), Table 2 shows the maximal difference between the two approximations of the G-function respective the J-function. In agreement with the theoretical developments, in both cases,

the difference does not exceed 4×10^{-4} when $q = 16$.

	$q = 4$	$q = 8$	$q = 12$	$q = 16$
$\alpha = 0.1$, G	3.8	8.4	4.4	3.2
J	15	8.5	6.1	3.9
$\alpha = 0.2$, G	4.7	3.1	3.9	1.5
J	4.7	3.1	2.1	1.7
$\alpha = 0.3$, G	0.1	1.9	1.4	1.1
J	0.2	2.0	1.5	1.3

Table 2: For the same three LGCPs as in Figure 1, maximal differences between the Laplace approximations of the G-function based on (11) respective (12), with $q = 4, 8, 12, 16$, and similarly for the J-function. Results are multiplied by 10^4.

4.4 Scots pine saplings dataset

The left panel in Figure 2 shows the locations of 126 Scots pine saplings in a 10 by 10 metre square. The dataset is included in the `R` package `spatstat` as `finpines`, and it has previously been analyzed by Penttinen et al. (1992), Stoyan and Stoyan (1994), and Møller et al. (1998). The first two papers fitted a Matérn cluster process, using the K-function (or its equivalent L-function) and its nonparametric estimate both for parameter estimation and model checking, while the third paper considered an LGCP with exponential covariance function and used the pair correlation function for parameter estimation and the F and G-functions for model checking. Møller et al. (1998) concluded that both models provide a reasonable fit although when also including a third-order functional summary (i.e. one based on $\mathbf{X}^{!}_{o,x}$) the LGCP model showed a better fit. Below we supplement this analysis by using the J-function and the approximation established in Section 4.2.

We fitted both models by minimum contrast estimation (method `kppm` in `spatstat`) which compares a non-parametric estimate of the K-function with its theoretical value. When approximating the J-function for the LGCP, we used $q = 12$; no improvements were noticed with higher values of q. The right panel in Figure 2 shows the theoretical J-functions for the two fitted models together with a non-parametric estimate of the J-function, considering 50 distances (r-values) on a regular grid between 0 and 0.9 metre (for

the exact expression of the J-function for the Matérn cluster process, see e.g. Møller and Waagepetersen, 2004). Clearly, the fitted LGCP provides a better fit than the fitted Matérn cluster process. Indeed, the maximal difference between the non-parametric estimate and the theoretical J-function equals 0.43 for the Matérn cluster model and 0.20 for the LGCP model.

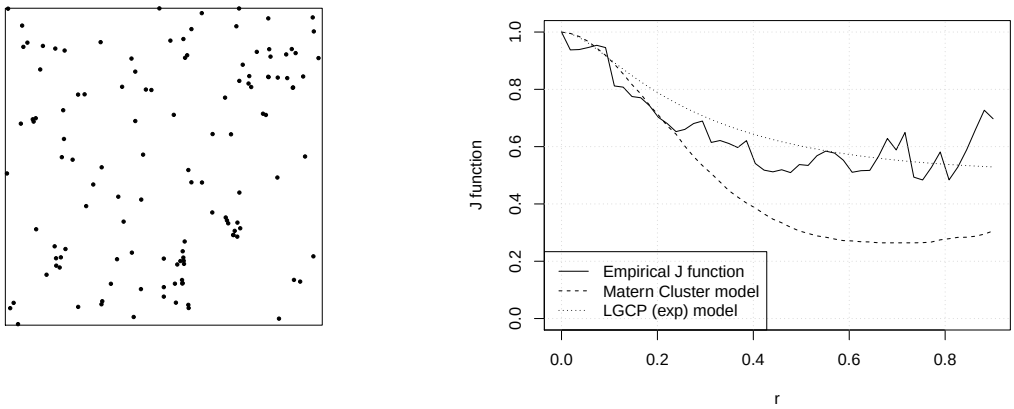

Figure 2: Left panel: Locations of 126 Scots pine saplings in a 10 by 10 metre square. Right panel: Non-parametric estimate of the J-function (solid curve) and fitted J-functions for the Matérn cluster process (dashed curve) and the LGCP with exponential covariance function (dotted curve).

5 Concluding remarks

We expect that our results for the reduced Palm distributions for an LGCP can be exploited further regarding third-order and higher order functional summaries (one such characteristic was briefly mentioned in Section 4.4), parameter estimation procedures, model checking, etc. For example, for any point process, the pair correlation function (when it exists) is invariant under independent thinning; could this property be exploited in connection to LGCPs where we know how the pair correlation function is related to those of the reduced Palm distributions?

Acknowledgments

J. Møller and R. Waagepetersen were supported by the Danish Council for Independent Research — Natural Sciences, grant 12-124675, "Mathematical and Statistical Analysis of Spatial Data", and by the "Centre for Stochastic Geometry and Advanced Bioimaging", funded by grant 8721 from the Villum Foundation. J.-F. Coeurjolly was supported by ANR-11-LABX-0025 PERSYVAL-Lab (2011, project OculoNimbus).

A Proofs

Proof of Lemma 1: By conditioning on $\mathbf{\Lambda}$, (1) becomes

$$\begin{aligned}
&\mathrm{EE}\left\{\sum_{x_1,\ldots,x_n\in\mathbf{X}}^{\neq} h(x_1,\ldots,x_n,\mathbf{X}\setminus\{x_1,\ldots,x_n\})\,\middle|\,\mathbf{\Lambda}\right\}\\
&=\mathrm{EE}\left\{\int_S\cdots\int_S h(x_1,\ldots,x_n,\mathbf{X})\prod_{i=1}^n\Lambda(x_i)\,\mathrm{d}x_1\cdots\mathrm{d}x_n\,\middle|\,\mathbf{\Lambda}\right\} \qquad (19)\\
&=\int_S\cdots\int_S \mathrm{E}\left\{h(x_1,\ldots,x_n,\mathbf{X})\prod_{i=1}^n\Lambda(x_i)\right\}\mathrm{d}x_1\cdots\mathrm{d}x_n. \qquad (20)
\end{aligned}$$

Here, in (19) we use that $\mathbf{X}$ given $\mathbf{\Lambda}$ is a Poisson process and apply the extended Slivnyak-Mecke theorem (Møller and Waagepetersen, 2004), and in (20) we use Fubini's theorem. Combining (1) and (20), we deduce (4). Finally, (5) follows from (4) with $h(x_1,\ldots,x_n,\mathbf{x})=1[\mathbf{x}\cap K=\emptyset]$.

Proof of Theorem 1: By (5) and (6)-(7), we just have to show that for any compact $K\subseteq S$ and pairwise distinct points $x_1,\ldots,x_n\in S$,

$$\begin{aligned}
&\mathrm{E}\exp\left[-\int_K\exp\left\{\tilde{Y}(u)+\mu_{x_1,\ldots,x_n}(u)\right\}\mathrm{d}u\right]\\
&=\mathrm{E}\exp\left[\sum_{i=1}^n\tilde{Y}(x_i)-\sum_{i,j=1}^n c(x_i,x_j)/2-\int_K\exp\left\{\mu(u)+\tilde{Y}(u)\right\}\mathrm{d}u\right].
\end{aligned}$$

This follows by showing that the distribution of $\{\tilde{Y}(u)+\sum_{i=1}^n c(u,x_i)\}_{u\in S}$ is absolutely continuous with respect to the distribution of $\tilde{\mathbf{Y}}=\{\tilde{Y}(u)\}_{u\in S}$,

with density $\exp\left\{\sum_{i=1}^n \tilde{y}(x_i) - \sum_{i,j=1}^n c(x_i,x_j)/2\right\}$ when $\tilde{\mathbf{y}}$ is a realization of $\tilde{\mathbf{Y}}$. Since the distribution of a random field is determined by its finite dimensional distributions, we just need to verify the agreement of the characteristic functions of the probability measures Q_1 and Q_2 given by

$$Q_1(B) = P\left(\{\tilde{Y}(u) + \sum_{i=1}^n c(u,x_i))\}_{u\in U} \in B\right)$$

and

$$Q_2(B) = \mathrm{E}\left(1\left[\{\tilde{Y}(u)\}_{u\in U} \in B\right] \exp\left\{\sum_{i=1}^n \tilde{Y}(x_i) - \sum_{i,j=1}^n c(x_i,x_j)/2\right\}\right),$$

for any Borel set $B \subseteq \mathbb{R}^{n+m}$, any pairwise distinct locations $u_1,\ldots,u_m \in \mathbb{R}^d \setminus \{x_1,\ldots,x_n\}$, with $m > 0$ and $U = \{x_1,\ldots,x_n,u_1,\ldots,u_m\}$. Let $\Sigma = \{c(u,v)\}_{u,v\in U}$ denote the covariance matrix of $\{\tilde{Y}(u)\}_{u\in U}$, and let $c_{x_1,\ldots,x_n} = \{\sum_{i=1}^n c(u,x_i)\}_{u\in U} = \Sigma e$, where e consists of n 1's followed by m 0's. For $t \in \mathbb{R}^{n+m}$, the characteristic function of Q_2 is (with $i^2 = -1$)

$$\begin{aligned}
&\mathrm{E}\exp\left[i\{\tilde{Y}(u)\}_{u\in I}^{\mathsf{T}} t + \{\tilde{Y}(u)\}_{u\in I}^{\mathsf{T}} e + e^{\mathsf{T}}\Sigma e/2\right] \\
&= \exp\left(e^{\mathsf{T}}\Sigma e/2\right) \mathrm{E}\exp\left[i\{\tilde{Y}(u)\}_{u\in I}^{\mathsf{T}}(t - ie)\right] \\
&= \exp\left(e^{\mathsf{T}}\Sigma e/2 + ie^{\mathsf{T}}\Sigma t - t^{\mathsf{T}}\Sigma t/2 - e^{\mathsf{T}}\Sigma e/2\right) \\
&= \exp\left(ie^{\mathsf{T}}\Sigma t - t^{\mathsf{T}}\Sigma t/2\right).
\end{aligned}$$

The last expression is the characteristic function of Q_1 which concludes the proof. For the second last equality in the above derivation we considered $\{\tilde{Y}(u)\}_{u\in U}$ as a complex Gaussian vector and used the expression for its characteristic function.

References

J.-F. Coeurjolly, J. Møller, and R. Waagepetersen. Conditioning in spatial point processes. Submitted for publication. Available at arXiv:1512.05871, 2015.

P. Coles and B. Jones. A lognormal model for the cosmological mass distribution. *Monthly Notices of the Royal Astronomical Society*, 248:1–13, 1991.

D.J. Daley and D. Vere-Jones. *An Introduction to the Theory of Point Processes, Volume I: Elementary Theory and Methods*. Springer, New York, second edition, 2003.

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

C. Heyde. On a property of the lognormal distribution. *Journal of Royal Statistical Society: Series B (Statistical Methodology)*, 25:392–393, 1963.

O. Macchi. The coincidence approach to stochastic point processes. *Advances in Applied Probability*, 7:83–122, 1975.

J. Møller. Shot noise Cox processes. *Advances in Applied Probability*, 35: 614–640, 2003.

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

J. Møller, A.R. Syversveen, and R.P. Waagepetersen. Log Gaussian Cox processes. *Scandinavian Journal of Statistics*, 25:451–482, 1998.

A. Penttinen, D. Stoyan, and H.M. Henttonen. Marked point processes in forest statistics. *Forest Science*, 38:806–824, 1992.

S.M. Stigler. Memoir on the probability of the causes of events, Laplace, Pierre Simon. (english translation). *Statistical Science*, 1:364–378, 1986.

D. Stoyan and H. Stoyan. *Fractals, Random Shapes, and Point Fields: Methods of Geometrical Statistics*. Wiley, Chichester, 1994.

M.N.M. Van Lieshout. A J-function for inhomogeneous point processes. *Statistica Neerlandica*, 65:183–201, 2011.