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

Jean-François Coeurjolly, Jesper Møller and Rasmus Waagepetersen

Palm distributions for log Gaussian Cox processes

No. 09, June 2015

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 Bajersvej 7E, DK-9220 Aalborg, email: jm@math.aau.dk,rw@math.aau.dk,

June 15, 2015

Abstract

This paper reviews useful results related to Palm distributions of spatial point processes and provides a new result regarding the characterization of Palm distributions for the class of log Gaussian Cox processes. This result is used to study functional summary statistics for a log Gaussian Cox process.

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

1 Introduction

Palm distributions [see e.g. Møller and Waagepetersen, 2003, Daley and Vere-Jones, 2008] are, at least among many applied statisticians and among most students, considered one of the more difficult topics in the field of spatial point processes. This is partly due to the general definition of Palm distributions which relies on measure theoretical results. In Section 2, in the context of finite point processes, we initially give an alternative very explicit definition

of Palm distributions in terms of their density functions. Section 2 further reviews Palm distributions in the general case as well as other concepts needed in the remaining part of the paper which deals with Palm distributions for Cox point processes and in particular log Gaussian Cox processes [Møller et al., 1998].

Section 3 establishes a characterization of Cox processes in terms of 'locally integrable joint intensities', and under mild regularity conditions Section 4 verifies a surprisingly simple characterization of the reduced Palm distribution for a log Gaussian Cox process: the reduced Palm distribution is itself a log Gaussian Cox process that only differs from the original log Gaussian Cox process in its intensity function. Section 5 applies this to study certain functional summary statistics (so-called F, G, and J-functions) for stationary log Gaussian Cox processes: We establish some new theoretical results, consider how to calculate F, G, and J using Laplace approximations, and discuss an application. Finally, Section 6 concludes the paper.

2 Palm distributions and prerequisites

Our general setting is as follows. 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 [2003] 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 $\mathbf{X}$.

Section 2.1 considers the case where S is bounded and $\mathcal{N}$ hence is all finite subsets of S, while Section 2.2 deals with the general case where S is arbitrary (i.e. including the case $S = \mathbb{R}^d$).

2.1 The finite case

Throughout this section we assume that S is bounded or just that S has finite Lebesgue measure $|S|$; let $\mathbf{Z}$ be a unit rate Poisson process on S with

distribution Π; and assume the distribution of $\mathbf{X}$ is absolutely continuous with respect to Π with density f. Thus,

$$\mathrm{E}h(\mathbf{X}) = \mathrm{E}\{f(\mathbf{Z})h(\mathbf{Z})\} \tag{1}$$

for any non-negative measurable function h on $\mathcal{N}$. Notice that for any event $F \subseteq \mathcal{N}$, denoting $1(\cdot)$ the indicator function and $\emptyset$ the empty point configuration,

$$\begin{aligned}&\mathrm{P}(\mathbf{X} \in F)\\ =&\sum_{n=0}^{\infty} \frac{\exp(-|S|)}{n!} \int_S \cdots \int_S 1(\{x_1, \ldots, x_n\} \in F) f(\{x_1, \ldots, x_n\})\, \mathrm{d}x_1 \cdots \mathrm{d}x_n\end{aligned} \tag{2}$$

where the term for $n = 0$ is $\exp(-|S|)1(\emptyset \in F)$. Therefore we shall consider probability statements in terms of $\exp(-|S|)f(\cdot)$. For example, the probability that $\mathbf{X} = \emptyset$ is $\exp(-|S|)f(\emptyset)$ while for $n \geq 1$,

$$\exp(-|S|)f(\{x_1, \ldots, x_n\})\, \mathrm{d}x_1 \cdots \mathrm{d}x_n$$

is the probability that $\mathbf{X}$ consists of precisely n points with one point in each of n infinitesimally small sets $B_1, \ldots, B_n$ around $x_1, \ldots, x_n$ with volumes $\mathrm{d}x_1, \ldots \mathrm{d}x_n$, respectively. Loosely speaking this event is '$\mathbf{X} = \{x_1, \ldots, x_n\}$'. Suppose f is hereditary, i.e. $f(\{x_1, \ldots, x_n\}) > 0$ whenever $f(\{x_0, x_1, \ldots, x_n\}) > 0$ and the points $x_0, x_1, \ldots, x_n \in S$ are pairwise distinct. We can then define the so-called nth order Papangelou conditional intensity by

$$\lambda^{(n)}(x_1, \ldots, x_n, \mathbf{x}) = f(\mathbf{x} \cup \{x_1, \ldots, x_n\})/f(\mathbf{x}) \tag{3}$$

for pairwise distinct $x_1, \ldots, x_n \in S$ and $\mathbf{x} \in \mathcal{N} \setminus \{x_1, \ldots, x_n\}$, setting $0/0 = 0$. By the previous interpretation of f, $\lambda^{(n)}(x_1, \ldots, x_n, \mathbf{x})\, \mathrm{d}x_1 \cdots \mathrm{d}x_n$ can be considered as the conditional probability of observing one point in each of the abovementioned infinitesimally small sets B_i conditional on that $\mathbf{X}$ outside $\cup_{i=1}^n B_i$ agrees with $\mathbf{x}$.

For $n = 1, 2, \ldots$, we define the nth order *joint intensity function* $\rho^{(n)}$ for $\mathbf{X}$ by

$$\rho^{(n)}(x_1, \ldots, x_n) = \mathrm{E}f(\mathbf{Z} \cup \{x_1, \ldots, x_n\}) \tag{4}$$

if $x_1, \ldots, x_n \in S$ are pairwise distinct, and set $\rho^{(n)}(x_1, \ldots, x_n) = 0$ otherwise. Assuming f is hereditary, $\rho^{(n)}(x_1, \ldots, x_n) = \mathrm{E}\lambda^{(n)}(x_1, \ldots, x_n, \mathbf{X})$ and

by the interpretation of $\lambda^{(n)}$ it follows that $\rho^{(n)}(x_1,\dots,x_n)\,\mathrm{d}x_1\cdots\mathrm{d}x_n$ can be viewed as the probability that $\mathbf{X}$ has a point in each of n infinitesimally small sets around $x_1,\dots,x_n$ with volumes $\mathrm{d}x_1,\dots\,\mathrm{d}x_n$, respectively. Loosely speaking, this event is '$x_1,\dots,x_n\in\mathbf{X}$'. Particularly, $\rho=\rho^{(1)}$ is the usual *intensity* function.

For reasons which soon will be clear, we assume that the mean value in (4) is finite and

$$\int_S\cdots\int_S\rho^{(n)}(x_1,\dots,x_n)\,\mathrm{d}x_1\cdots\mathrm{d}x_n<\infty. \tag{5}$$

Combining (1) and (4) with either (2) or the extended Slivnyak-Mecke formula for the Poisson process [Møller and Waagepetersen, 2003, Theorem 3.3], it is straightforwardly seen that

$$\mathrm{E}\sum_{x_1,\dots,x_n\in\mathbf{X}}^{\neq}h(x_1,\dots,x_n)=\int_S\cdots\int_S h(x_1,\dots,x_n)\rho^{(n)}(x_1,\dots,x_n)\,\mathrm{d}x_1\dots\,\mathrm{d}x_n \tag{6}$$

for any non-negative measurable function h on S^n, where $\neq$ over the summation sign means that $x_1,\dots,x_n$ are pairwise distinct. Denoting N the number of points in $\mathbf{X}$, the left hand side in (6) with $h=1$ is seen to be the factorial moment $\mathrm{E}\{N(N-1)\cdots(N-n+1)\}$, which by (5) is assumed to be finite.

Suppose $x_1,\dots,x_n\in S$ are pairwise distinct and $\rho^{(n)}(x_1,\dots,x_n)>0$. Then we define the *reduced Palm distribution* of $\mathbf{X}$ given points at $x_1,\dots,x_n$ as the distribution $\mathrm{P}^!_{x_1,\dots,x_n}$ for the point process $\mathbf{X}^!_{x_1,\dots,x_n}$ with density

$$f_{x_1,\dots,x_n}(\mathbf{x})=\frac{f(\mathbf{x}\cup\{x_1,\dots,x_n\})}{\rho^{(n)}(x_1,\dots,x_n)},\quad \mathbf{x}\in\mathcal{N},\ \mathbf{x}\cap\{x_1,\dots,x_n\}=\emptyset, \tag{7}$$

with respect to Π. By the previous infinitesimal interpretations of f and $\rho^{(n)}$ we can view $\exp(-|S|)f_{x_1,\dots,x_n}(\mathbf{x})$ as the 'joint probability' that $\mathbf{X}$ equals the union $\mathbf{x}\cup\{x_1,\dots,x_n\}$ divided by the 'probability' that $x_1,\dots,x_n\in\mathbf{X}$. Thus $\mathrm{P}^!_{x_1,\dots,x_n}$ has an interpretation as the conditional distribution of $\mathbf{X}\setminus\{x_1,\dots,x_n\}$ given that $x_1,\dots,x_n\in\mathbf{X}$. Conversely,

$$\exp(-|S|)f(\{x_1,\dots,x_n\})=\rho^{(n)}(x_1,\dots,x_n)\mathrm{P}\left(\mathbf{X}^!_{\{x_1,\dots,x_n\}}=\emptyset\right) \tag{8}$$

provides a factorization into the 'probability' of observing $\{x_1,\dots,x_n\}$ times the conditional probability of not observing further points. Finally, if

$x_1, \ldots, x_n \in S$ are not pairwise distinct or $\rho^{(n)}(x_1, \ldots, x_n) = 0$, the choice of $\mathbf{X}^!_{x_1,\ldots,x_n}$ and its distribution $\mathrm{P}^!_{x_1,\ldots,x_n}$ is not of any importance for the results in this paper, but for simplicity and specificity we then set $\mathbf{X}^!_{x_1,\ldots,x_n} = \emptyset$.

Some remarks are in order:

(a) The (non-reduced) *Palm distribution* of $\mathbf{X}$ given points at $x_1, \ldots, x_n$ is simply the distribution of the union of $\mathbf{X}^!_{x_1,\ldots,x_n}$ and $\{x_1, \ldots, x_n\}$.

(b) We obtain immediately from (4) and (7) that for any $x_1, \ldots, x_n \in S$ and $m = 1, 2, \ldots$, $\mathbf{X}^!_{x_1,\ldots,x_n}$ has mth order joint intensity function

$$\rho^{(m)}_{x_1,\ldots,x_n}(u_1, \ldots, u_m) = \begin{cases} \frac{\rho^{(m+n)}(x_1,\ldots,x_n,u_1,\ldots,u_m)}{\rho^{(n)}(x_1,\ldots,x_n)} & \text{if } \rho^{(n)}(x_1, \ldots, x_n) > 0 \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

for pairwise distinct $u_1, \ldots, u_m \in S \setminus \{x_1, \ldots, x_n\}$.

(c) Also we easily obtain that

$$\begin{aligned} &\mathrm{E} \sum_{x_1,\ldots,x_n \in \mathbf{X}}^{\neq} h(x_1, \ldots, x_n, \mathbf{X} \setminus \{x_1, \ldots, x_n\}) \\ &= \int_S \cdots \int_S \mathrm{E}^!_{x_1,\ldots,x_n} h(x_1, \ldots, x_n, \mathbf{X}) \rho^{(n)}(x_1, \ldots, x_n) \, \mathrm{d}x_1 \cdots \mathrm{d}x_n \end{aligned} \tag{10}$$

for any non-negative measurable function h on $S^n \times \mathcal{N}$, where $\mathrm{E}^!_{x_1,\ldots,x_n}$ denotes expectation with respect to $\mathrm{P}^!_{x_1,\ldots,x_n}$. Assuming f is hereditary and rewriting the expectation in the right hand side of (10) in terms of

$$f_{x_1,\ldots,x_n}(\mathbf{x}) = f(\mathbf{x}) \lambda^{(n)}(x_1, \ldots, x_n, \mathbf{x}) / \rho^{(n)}(x_1, \ldots, x_n)$$

the finite point process case of the celebrated Georgii-Nguyen-Zessin formula is obtained [Georgii, 1976, Nguyen and Zessin, 1979].

2.2 The general case

The concepts and results in the previous section extend to the general case where S is any Borel subset of $\mathbb{R}^d$. However, if $|S| = \infty$, the unit rate Poisson process on S will be infinite and we can not in general assume that $\mathbf{X}$ is absolutely continuous with respect to the distribution of this process.

Thus we do not longer have the direct definitions (4) and (7) of $\rho^{(n)}$ and $\mathbf{X}^!_{x_1,\ldots,x_n}$ in terms of density functions.

In fact (6) is usually taken as the definition of the nth order joint intensity $\rho^{(n)}$ for $\mathbf{X}$, provided there exists such a non-negative measurable function. Technically speaking, viewing the left hand side in (6) as an integral $\int h \, \mathrm{d}\alpha^{(n)}$, where $\alpha^{(n)}$ is called the nth order factorial moment measure on S^n, $\rho^{(n)}$ is assumed to be a density for $\alpha^{(n)}$ with respect to Lebesgue measure on S^n. Here, $\alpha^{(n)}$ is required to be a locally finite measure, i.e.

$$\int_B \cdots \int_B \rho^{(n)}(x_1, \ldots, x_n) \, \mathrm{d}x_1 \cdots \mathrm{d}x_n < \infty \quad \text{for all } B \in \mathcal{B}_0. \tag{11}$$

Thereby (10) can be used as the definition of the reduced Palm distributions. Their existence follows by viewing the left hand side in (10) as defining a measure $C^!$ (the so-called nth order reduced Campbell measure) on $S^n \times \mathcal{N}$ and noticing that for events $F \subseteq \mathcal{N}$, $C^!(\cdot \times F)$ is absolutely continuous with respect to $\alpha^{(n)}$ [see e.g. Møller and Waagepetersen, 2003].

We observe the following:

(e) Clearly $\rho^{(n)}(x_1, \ldots, x_n)$ and $\mathrm{P}^!_{x_1,\ldots,x_n}$ are then only uniquely determined up to a Lebesgue nullset of S^n. For ease of exposition, we shall usually ignore such nullsets.

(f) We have that $\rho^{(n)}(x_1, \ldots, x_n)$ and $\mathrm{P}^!_{x_1,\ldots,x_n}$ are invariant under permutations of the points $x_1, \ldots, x_n$. Moreover, $\left(\mathbf{X}^!_{x_1,\ldots,x_m}\right)^!_{x_{m+1},\ldots,x_n} = \mathbf{X}^!_{x_1,\ldots,x_n}$ if $0 < m < n$.

(g) In order to specify the reduced Palm distribution by (10) it is in fact enough to consider functions h of the form

$$h(x_1, \ldots, x_n, \mathbf{X}) = 1[x_i \in B_i, i = 1, \ldots, n, \mathbf{X} \cap K = \emptyset]$$

for sets $B_1, \ldots, B_n \in \mathcal{B}_0$ and compact sets $K \subseteq S$. From this it follows that $\mathbf{X}^!_{x_1,\ldots,x_n} \cap B$ and $(\mathbf{X} \cap B)^!_{x_1,\ldots,x_n}$ have the same distribution if $B \in \mathcal{B}_0$.

(h) Suppose that X is stationary, i.e. its distribution is invariant under translations in $\mathbb{R}^d$ and so $S = \mathbb{R}^d$ (unless $\mathbf{X} = \emptyset$ which is not a case of our interest). Then, for any $x \in \mathbb{R}^d$ and if o denotes the origin in $\mathbb{R}^d$, $X^!_x$ and $\{x + y : y \in X^!_o\}$ are identically distributed. The reduced Palm distribution $P^!_o$ is often considered as the 'conditional distribution for the further points in $\mathbf{X}$ given a typical point of $\mathbf{X}$'.

3 Cox processes

Let $\mathbf{\Lambda}$ 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, $\mathbf{X}$ is assumed to be a Cox process driven by a random intensity function $\mathbf{\Lambda} = \{\Lambda(x) : x \in S\}$, that is, conditional on $\mathbf{\Lambda}$, $\mathbf{X}$ is a Poisson process with intensity function $\mathbf{\Lambda}$. This ensures that for any bounded Borel set $B \subseteq S$, $\mathbf{X}_B$ is absolutely continuous with respect to the unit rate Poisson process on B, and we denote its density f_B. By Fubini's theorem, for any $n = 1, 2, \ldots$ and pairwise distinct points $x_1, \ldots, x_n \in B$,

$$f_B(\{x_1, \ldots, x_n\}) = \exp(|B|)\mathrm{E}\left[\exp\left\{-\int_B \Lambda(x)\,\mathrm{d}x\right\}\prod_{i=1}^{n}\Lambda(x_i)\right]. \tag{12}$$

Also we assume that $\mathbf{\Lambda}$ has moments of any order $n = 1, 2, \ldots$. Then the joint intensities are finite and given by

$$\rho^{(n)}(x_1, \ldots, x_n) = \mathrm{E}\left\{\prod_{i=1}^{n}\Lambda(x_i)\right\} \tag{13}$$

for any $n = 1, 2, \ldots$ and pairwise distinct $x_1, \ldots, x_n \in S$. In addition (11) is assumed. The following lemma is used in Section 4 and is verified in Appendix A.

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 non-negative measurable function h on $S^n \times \mathcal{N}$,*

$$\rho^{(n)}(x_1, \ldots, x_n)\mathrm{E}\left\{h\left(x_1, \ldots, x_n, \mathbf{X}^!_{x_1,\ldots,x_n}\right)\right\} = \mathrm{E}\left\{h(x_1, \ldots, x_n, \mathbf{X})\prod_{i=1}^{n}\Lambda(x_i)\right\}. \tag{14}$$

In particular, if $\rho^{(n)}(x_1, \ldots, x_n) > 0$, the void probabilities of $\mathbf{X}^!_{x_1,\ldots,x_n}$ are given by

$$\mathrm{P}(\mathbf{X}^!_{x_1,\ldots,x_n} \cap K = \emptyset) = \mathrm{E}\left[\prod_{i=1}^{n}\Lambda(x_i)\exp\left\{-\int_K \Lambda(u)\,\mathrm{d}u\right\}/\rho^{(n)}(x_1, \ldots, x_n)\right] \tag{15}$$

for compact $K \subseteq S$.

A natural question is when the distribution of the Cox process $\mathbf{X}$ and hence also those of the associated reduced Palm processes $\mathbf{X}^!_{x_1,\ldots,x_n}$ are uniquely determined by the joint intensities. A sufficient condition is the existence of a number $a = a(B) > 1$ for each set $B \in \mathcal{B}_0$ such that

$$\mathrm{E}\left[\exp\left\{a\int_B \Lambda(u)\,\mathrm{d}u\right\}\right] < \infty. \tag{16}$$

This together with (17) below were established in Macchi [1975, Theorem 8 (ii)] for one-dimensional Cox processes; for the proof there are no changes induced by higher dimension, and so we have the following proposition where (18) follows by combining (8) with (17).

Proposition 1. *Suppose the Cox process* $\mathbf{X}$ *satisfies* (16)*. Then, for any* $B \in \mathcal{B}_0$ *and pairwise distinct* $x_1, \ldots, x_n \in B$*,*

$$\begin{aligned}&\exp(-|B|)f_B(\{x_1,\ldots,x_n\})\\ &= \sum_{m=0}^{\infty}\frac{(-1)^m}{m!}\int_B\cdots\int_B \rho^{(n+m)}(x_1,\ldots,x_n,u_1,\ldots,u_m)\,\mathrm{d}u_1\cdots\mathrm{d}u_m.\end{aligned} \tag{17}$$

In particular, the void probability of $\mathbf{X}^!_{x_1,\ldots,x_n}$ *is*

$$\begin{aligned}&\mathrm{P}\left(\mathbf{X}^!_{x_1,\ldots,x_n}\cap K=\emptyset\right)\\ &=\sum_{m=0}^{\infty}\frac{(-1)^m}{m!}\int_K\cdots\int_K \rho^{(m)}_{x_1,\ldots,x_n}(u_1,\ldots,u_m)\,\mathrm{d}u_1\cdots\mathrm{d}u_m\end{aligned} \tag{18}$$

for compact $K \subseteq S$*.*

Proposition 1 has been applied for Cox permanental point processes, which can be defined in terms of their joint intensities being certain permanents of a covariance function, see Macchi [1975] and McCullagh and Møller [2006].

4 Reduced Palm distributions for log Gaussian Cox processes

Henceforth, let $\Lambda(x) = \exp\{Y(x)\}$ where $\mathbf{Y} = \{Y(x) : x \in S\}$ 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 discussed later in connection to our main result (Theorem 1 below). In the sequel, we let $\mathbf{X}$ be the Cox process driven by the random intensity function $\Lambda(x) = \exp\{Y(x)\}$. Then $\mathbf{X}$ is a *log Gaussian Cox process (LGCP)* (introduced by Coles and Jones [1991] in astronomy and independently by Møller et al. [1998] in statistics).

For $x, y \in S$, we 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],

$$\rho(x) = \exp\{\mu(x) + c(x,x)/2\} \quad \text{and} \quad g(x,y) = \exp\{c(x,y)\} \tag{19}$$

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

$$\rho^{(n)}(x_1, \ldots, x_n) = \left\{ \prod_{i=1}^{n} \rho(x_i) \right\} \left\{ \prod_{1 \leq i < j \leq n} g(x_i, x_j) \right\} \tag{20}$$

is strictly positive. Finally, we assume that (11) is satisfied; this is e.g. the case if c is continuous and μ restricted to any $B \in \mathcal{B}_0$ is bounded from above, cf. (19)-(20).

For the LGCP restricted to $B \in \mathcal{B}_0$, the density f_B as given by (12) is not expressible on closed form. Combining (8) and (20) shows that f_B has some similarity to the density for a pairwise interaction point process with first and second order interaction functions given by ρ and g. However, while a pairwise interaction point process has a normalizing constant given by its probability for having no points in B, according to (8), f_B for a LGCP involves the void probability for $\mathbf{X}^!_{x_1,\ldots,x_n}$ which in general depends on $\{x_1, \ldots, x_n\}$.

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

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

Combining (9) and (19)-(20), for pairwise distinct $x_1, \ldots, x_{m+n} \in S$ with

$m > 0$ and $n > 0$,

$$\begin{aligned}
&\rho^{(m)}_{x_1,\ldots,x_n}(x_{1+n},\ldots,x_{m+n}) \\
&= \left[\prod_{i=n+1}^{m+n}\left\{\rho(x_i)\prod_{j=1}^{n} g(x_i,x_j)\right\}\right]\left\{\prod_{1+n\le i<j\le m+n} g(x_i,x_j)\right\} \\
&= \prod_{i=n+1}^{m+n}\exp\left\{\mu_{x_1,\ldots,x_n}(x_i)+c(x_i,x_i)\right\} \qquad (21)
\end{aligned}$$

and hence we obtain the following result.

Proposition 2. *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 intensity* (21) *which agrees with the m'th order intensity function for a LGCP with mean function* $\mu_{x_1,\ldots,x_n}$ *and covariance function* c *for the underlying Gaussian process.*

Proposition 2 indicates that also $\mathbf{X}^!_{x_1,\ldots,x_n}$ could be a LGCP. However, we have not been successful in verifying Macchi's condition (16) in Proposition 1, 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 our main result, namely that $\mathbf{X}^!_{x_1,\ldots,x_n}$ is a LGCP, at least under mild regularity conditions: Consider a stationary zero-mean unit-variance Gaussian process $\mathbf{Z}$ defined on $\mathbb{R}^d$ with a correlation function $\mathrm{Cov}\{Z(x),Z(y)\} = r(x-y)$ satisfying certain regularity conditions in Møller et al. [1998] so that $\mathbf{Z}$ is almost surely continuous. Let $\sigma(\cdot) \ge 0$ and $\mu(\cdot)$ be continuous functions on $S \setminus C$ where $|C| = 0$. For $x \in S$, define $Y_0(x) = \sigma(x)Z(x)$ and $Y(x) = Y_0(x) + \mu(x)$. Then $\mathbf{Y}$ has mean function μ and a covariance function of the form

$$c(x,y) = \sigma(x)\sigma(y)r(x-y), \quad x,y \in S.$$

Thereby, $\mathbf{Y}$ is continuous on $S \setminus C$ a.s., and so $\mathbf{\Lambda}$ is locally integrable a.s. Moreover, assume that r is strictly positive definite. If r is continuous and integrable on $\mathbb{R}^d$, strictly positive definiteness is equivalent to that the Fourier transform of r is non-negative and not identically zero.

Theorem 1. *Under the regularity conditions above for the LGCP* $\mathbf{X}$ *driven by* $\mathbf{\Lambda} = \exp(\mathbf{Y})$*, there exists a Lebesgue nullset* $B \subset \mathbb{R}^d$ *such that for any pairwise distinct* $x_1, \ldots, x_n \in S \setminus B$*,* $\mathbf{X}^!_{x_1,\ldots,x_n}$ *is a LGCP where the underlying Gaussian process has mean function* $\mu_{x_1,\ldots,x_n}$ *and the same covariance function* $c(x, y) = \sigma(x)\sigma(y)r(x - y)$ *as* $\mathbf{Y}$*.*

Theorem 1 is verified in Appendix A, where $B \supseteq C$. That such nullsets may appear is no surprise, cf. (e) in Section 2.2. The regularity conditions used in Theorem 1 are commonly imposed in the literature on LGCPs, where indeed further restrictions are often used: frequently $\mathbf{Y}$ is often assumed to be continuous a.s., but our nullset C allows 'stepwise discontinuities'; and $\mathbf{X}$ is second order intensity reweighted stationary [Baddeley et al., 2000] if and only if $\sigma(\cdot)$ is constant, but we allow $\sigma(\cdot)$ to be any non-negative continuous function.

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'. Even more can be said when Theorem 1 applies: If $\mathbf{Y}_{x_1,\ldots,x_n}$ denotes the Gaussian field associated with $\mathbf{X}^!_{x_1,\ldots,x_n}$, then a coupling between $\mathbf{X}$ and $\mathbf{X}^!_{x_1,\ldots,x_n}$ is obtained by taking $Y_{x_1,\ldots,x_n}(x) = Y(x) + \sum_{i=1}^n c(x, x_i)$. In particular, if $c \geq 0$ and we are given pairwise distinct points $x_1, \ldots, x_n$ from $\mathbf{X}$, we can consider $\mathbf{X} \setminus \{x_1, \ldots, x_n\}$ as being included in $\mathbf{X}^!_{\{x_1,\ldots,x_n\}}$. This property clearly shows the attractiveness of the LGCP if $c \geq 0$ (equivalently $g \geq 1$).

5 Functional summary statistics for stationary log Gaussian Cox processes

In the remainder, we let $S = \mathbb{R}^d$ and assume that $\mathbf{X}$ is a stationary LGCP satisfying the regularity conditions for Theorem 1. By (19)-(20), stationarity of the LGCP is equivalent to stationarity of the underlying Gaussian process. It implies that the intensity ρ is constant and the pair correlation function $g(x, y) = g_0(x-y)$ is translation invariant, where $g_0(x) = \exp\{c_0(x)\}$, $c_0(x) = c(o, x)$, and $x, y \in \mathbb{R}^d$.

Denote $B(o, r)$ the ball in $\mathbb{R}^d$ of radius $r > 0$ and centered at the origin o. Then apart from g_0 and the related Ripley's K-function given by

$$K(r) = \int_{B(o,r)} g_0(x)\,\mathrm{d}x,$$

popular tools for exploratory purposes as well as model fitting and model checking are the following functional summary statistics: The *empty space function*

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

the *nearest-neighbour distribution function*

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

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 [2003].

Section 5.1 establishes some new results for these theoretical functions and Section 5.2 discusses how they can be calculated using a Laplace approximation. Section 5.3 illustrates this calculation and Section 5.4 discusses an application for a real dataset.

5.1 New formulae

By conditioning on $\mathbf{Y}$, we see that the empty space function for $\mathbf{X}$ is given by

$$1 - F(r) = \mathrm{E}\left(\exp\left[-\int_{B(o,r)} \exp\{Y(x)\}\,\mathrm{d}x\right]\right). \tag{22}$$

Using the Slivnyak-Mecke formula, Møller et al. [1998] showed that

$$1 - G(r) = \frac{1}{\rho}\,\mathrm{E}\left(\exp\left[Y(o) - \int_{B(o,r)} \exp\{Y(x)\}\,\mathrm{d}x\right]\right). \tag{23}$$

Since the nearest-neighbour distribution function for $\mathbf{X}$ is the same as the empty space function for $\mathbf{X}_o^!$, which is a LGCP with underlying Gaussian process $Y_o(x) = Y(x) + c_0(x)$, and since $g_0(x) = \exp\{c_0(x)\}$, we obtain an alternative expression

$$1 - G(r) = \mathrm{E}\left(\exp\left[-\int_{B(o,r)} g_0(x)\exp\{Y(x)\}\,\mathrm{d}x\right]\right). \tag{24}$$

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

$$J(r) = \frac{\mathrm{E}\left(\exp\left[-\int_{B(o,r)} g_0(x)\exp\{Y(x)\}\,\mathrm{d}x\right]\right)}{\mathrm{E}\left(\exp\left[-\int_{B(o,r)}\exp\{Y(x)\}\,\mathrm{d}x\right]\right)}. \tag{25}$$

Van Lieshout [2011] established for a general stationary point process the approximation $J(r)-1 \approx -\rho\{K(r)-\pi r^2\}$, where πr^2 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 $c_0 \geq 0$, i.e. $g_0 \geq 1$, then we have $K(r)-\pi r^2 \geq 0$, and so we expect that $J(r) \leq 1$. Indeed Van Lieshout [2011] verified this in the case of a stationary LGCP with $g_0 \geq 1$. This result immediately follows by the new expression (25).

5.2 Laplace approximation

In this section we discuss Laplace approximation of $1-G(r)$ (the approximation of $1-F(r)$ is similar). The notation introduced is also used in the proof of Theorem 1 in the Appendix. For $\Delta > 0$ define the grid of quadrature points $\mathcal{G}(\Delta) = \{(\Delta i_1, \ldots, \Delta i_d) \,|\, i_1, \ldots, i_d \in \mathbb{Z}\}$. Further let for $v \in \mathcal{G}(\Delta)$, $A_v^\Delta = [v_1 - \Delta/2, v_1 + \Delta/2[\times \cdots \times [v_d - \Delta/2, v_d + \Delta/2[$ be the grid cell associated with v. For any non-negative Borel function $\ell : \mathbb{R}^d \to \mathbb{R}$, we use the approximation

$$\int_{B(o,r)} \exp\{Y(x)\}\ell(x)\,\mathrm{d}x \approx \sum_{v\in\mathcal{G}(\Delta)\cap B(o,r)} w_v \ell(v)\exp\{Y(v)\} \tag{26}$$

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 $\{Y(v)\}_{v\in\mathcal{G}(\Delta)}$. Then the approximation (26) gives

$$1 - G(r) \approx \int \exp\{h(y)\}\,\mathrm{d}y \tag{27}$$

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

$$h(y) = -\sum_v w_v g_0(v)\exp(y_v) - \frac{1}{2}(y-M)^\top\Sigma^{-1}(y-M) - \frac{1}{2}\log\{(2\pi)^m|\Sigma|\}.$$

The gradient vector for h is

$$\nabla h(y) = -d(y) - \Sigma^{-1}(y - M) \tag{28}$$

where $d(y) = \{w_v g_0(v) \exp(y_v)\}_{v \in \mathcal{G}(\Delta) \cap B(o,r)}$, and the negated Hessian matrix for h is

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

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

$$y^{(l+1)} = y^{(l)} + H^{-1}\{y^{(l)}\} \nabla h\{y^{(l)}\}. \tag{29}$$

Now, the logarithm of the Laplace approximation of the right hand side in (27) (see e.g. Stigler [1986]) equals

$$\log\{1-G(r)\} \approx -\sum_{v \in \mathcal{G}(\Delta) \cap B(o,r)} w_v g_0(v) \exp(\hat{y}_v) + \frac{1}{2}(\hat{y}-M)^\top d(\hat{y}) - \frac{1}{2} \log |D(\hat{y})\Sigma + I|. \tag{30}$$

For the computation of $\Sigma^{-1}(y-M)$ in (28) 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 (29) 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, $|D(\hat{y})\Sigma + I| = |R(\hat{y})|$.

5.3 Numerical illustration

To illustrate the Laplace approximations of the G- and J-functions (Section 5.2) we consider three planar stationary LGCPs with intensity $\rho = 50$ and spherical covariance function 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 we denote by G_q, F_q, and J_q the corresponding Laplace approximations of G, F, and J. Figure 1 shows the resulting curves with $q = 16$. To appreciate 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} except for the J-function and $\alpha = 1$). This justifies our choice $q = 16$ in Figure 1.

The Laplace approximation of the G-function could also be derived using the classical formula (23). To check the agreement of the numerical approximations of equations (23) and (24), we computed in Table 2 the maximal difference between the two approximations of the G- and J-functions. In agreement with the theoretical developments, the difference does not exceed 4×10^{-4} when $q = 16$.

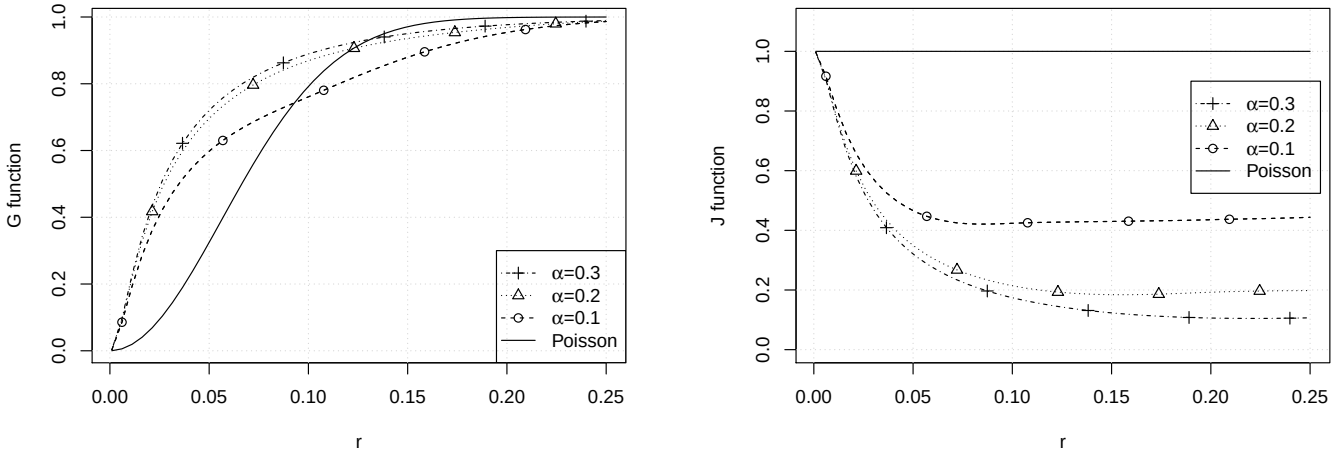

Figure 1: Laplace approximations of the G- (left) and J-functions (right) for three planar stationary LGCPs with intensity $\rho = 50$ and a spherical covariance function with variance $\sigma^2 = 4$ and scale parameters $\alpha = 0.1, 0.2, 0.3$, respectively.

	$\max_{r\in\mathcal{R}} \lvert H_{16} - H_q \rvert$, $H = G, J$		
	$q = 4$	$q = 8$	$q = 12$
$\alpha = 0.1$, 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: Maximal differences between the Laplace approximations G_q and G_{16} and F_q and F_{16} with $q = 4, 8, 12$ for the same three LGCPs as in Figure 1. Results are multiplied by 10^3.

	$q = 4$	$q = 8$	$q = 12$	$q = 16$
$\alpha = 0.1$, 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: Maximal differences between the Laplace approximations of the G- and J-functions for the same three LGCPs as in Figure 1 and derived from (23) or (24) using $q = 4, 8, 12, 16$. Results are multiplied by 10^4.

5.4 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 two first papers fitted a Matérn cluster process using the K-function (or its equivalent L-function) both for parameter estimation and model checking, while the third paper considered a 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

statistic (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 the previous section.

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 the value $q = 12$ (no improvements were noticed with higher values of q). For the exact expression of the J-function for the Matérn cluster process we refer to Møller and Waagepetersen [2003]. 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. 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.

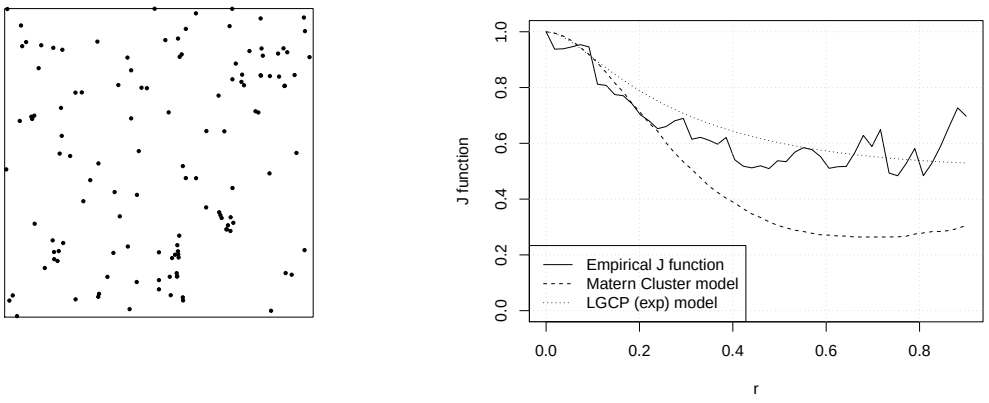

Figure 2: Left panel: Locations of 126 Scots pine saplings in a 10 by 10 metre square. Right panel: Non-parametric estimate of the J-function (solid curve) and fitted J-functions for the Matérn cluster process (dashed curve) and the LGCP with exponential covariance function (dotted curve).

6 Concluding remarks

This paper has provided a description of Palm distribution intended for a mathematically less inclined audience. Reduced Palm distributions are known for some models of spatial point processes: A Poisson process is characterized by the fact that its distribution agrees with its reduced Palm distribution; the reduced Palm distribution of a Gibbs point process is absolutely continuous with respect to its distribution, with density equal to its Papangelou conditional intensity (this is equivalent to the Georgii-Nguyen-Zessin formula discussed in (d), Section 2.1); and the Palm distributions for infinitely divisible point processes such as Poisson cluster processes are also known [see e.g. Daley and Vere-Jones, 2008, Chiu et al., 2013, and the references therein]. We established a further result: Under mild regularity conditions, the reduced Palm distributions for a LGCP are again LGCPs with the same pair correlation function. This was used to derive new results for the G and J-functions. Furthermore, we considered a Laplace approximation for the calculation of F, G, and J-functions, which was used to supplement previous analyses of the Scots pine saplings dataset modelled by a LGCP or a Matérn cluster process.

We expect that our results for the reduced Palm distributions for a LGCP can be exploited further regarding third-order and higher order functional summary statistics (one such characteristic was briefly mentioned in Section 5.4), 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? A fundamental question is if the reduced Palm distributions for *any* LGCP are LGCPs; e.g. it is essential in our proof of Theorem 1 that we work with a strictly positive definite covariance function, but is this condition really needed? We leave these and other open questions for future work.

A Proofs

Proof of Lemma 1: By conditioning on $\mathbf{\Lambda}$, the right hand side of (10) becomes

$$\begin{aligned}
&\mathrm{E}\mathrm{E}\left\{\sum_{x_1,\ldots,x_n\in\mathbf{X}} h(x_1,\ldots,x_n,\mathbf{X}\setminus\{x_1,\ldots,x_n\})|\mathbf{\Lambda}\right\}\\
=&\,\mathrm{E}\mathrm{E}\left\{\int_S\cdots\int_S h(x_1,\ldots,x_n,\mathbf{X})\prod_{i=1}^n\Lambda(x_i)\,\mathrm{d}x_1\cdots\mathrm{d}x_n|\mathbf{\Lambda}\right\} \qquad (31)\\
=&\int_S\cdots\int_S \mathrm{E}\left\{h(x_1,\ldots,x_n,\mathbf{X})\prod_{i=1}^n\Lambda(x_i)\right\}\mathrm{d}x_1\cdots\mathrm{d}x_n. \qquad (32)
\end{aligned}$$

Here, in (31) we use that $\mathbf{X}$ given $\mathbf{\Lambda}$ is a Poisson process and apply the extended Slivnyak-Mecke theorem [Møller and Waagepetersen, 2003], and in (32) we use Fubini's theorem. Combining (10) and (32), we deduce (14). Finally (15) follows from (14) with $h(x_1,\ldots,x_n,\mathbf{x}) = 1[\mathbf{x}\cap K=\emptyset]$.

Proof of Theorem 1: By (15) and (19)-(20) we just have to show the existence of a Lebesgue nullset $B\subseteq S$ such that for any compact $K\subseteq S$ and pairwise distinct points $x_1,\ldots,x_n\in S\setminus B$,

$$\begin{aligned}
&\mathrm{E}\exp\left[\sum_{i=1}^n Y_o(x_i)-\sum_{i,j=1}^n c(x_i,x_j)/2-\int_K\exp\{\mu(u)+Y_o(u)\}\,\mathrm{d}u\right]\\
=&\mathrm{E}\exp\left[-\int_K\exp\{Y_o(u)+\mu_{x_1,\ldots,x_n}(u)\}\,\mathrm{d}u\right]. \qquad (33)
\end{aligned}$$

For $\Delta>0$ define the grid $\mathcal{G}(\Delta)$ with associated grid cells A_v^Δ, $v\in\mathcal{G}(\Delta)$ as in Section 5.2. In the following we let $\Delta=k^{-1}$ for integer $k\geq 1$. For $u\in S$, let $g^k(u)$ be the grid point in $\mathcal{G}(k^{-1})$ closest to u, discarding the Lebesgue nullset C_k of those $u\in S$ where there is not a unique such grid point. Further, let $C_0=C$ and $B=\cup_{k=0}^\infty C_k$. Then B has Lebesgue measure zero, and we let $v_i^k=g^k(x_i)$, $i=1,\ldots,n$. Furthermore, for $v\in\mathcal{G}(k^{-1})$ we let $w_v^k=|K\cap A_v^{k^{-1}}|$. Then by Waagepetersen [2004, Lemma 1] and dominated convergence, the left and right hand sides of (33) are equal to the limits as $k\to\infty$ of respectively

$$\mathrm{E}\exp\left[\sum_{i=1}^n Y_o(v_i^k)-\sum_{i,j=1}^n c(v_i^k,v_j^k)/2-\sum_{v\in\mathcal{G}(k^{-1})} w_v^k\exp\{Y_o(v)+\mu(v)\}\right] \qquad (34)$$

and

$$\mathrm{E}\exp\left[-\sum_{v\in\mathcal{G}(k^{-1})} w_v^k \exp\left\{Y_o(v)+\mu(v)+\sum_{i=1}^{n} c(v,v_i^k)\right\}\right]. \qquad (35)$$

Let $H^k=\{v_1^k,\ldots,v_n^k\}$ (a set of cardinality n for all sufficiently large k) and order the elements in the set $I^k=(\mathcal{G}(k^{-1})\cap K)\cup H^k$ when considering the Gaussian vector $\mathbf{Y}_o^k=\{Y_o(v)\}_{v\in I^k}$ with mean vector $M_k=\{\mu(v)\}_{v\in I^k}$ and covariance matrix $\Sigma_k=\{c(u,v)\}_{u,v\in I^k}$. Then (34) and (35) can be rewritten as multivariate integrals involving the Gaussian density of $\mathbf{Y}_o^k$ at each realization $y=(y_v)_{v\in I^k}$ (see also Section 5.2). Considering (35), making a shift of coordinates from y to $z=\{y_v+\sum_{i=1}^n c(v,v_i^k)\}_{v\in I^k}=y+\sum_{i=1}^n \Sigma_{k,v_i^k}$ where Σ_{k,v_i^k} denotes the column in Σ_k corresponding to index v_i^k, we notice that the exponent for (35) becomes

$$\begin{aligned}
&-\sum_{v\in\mathcal{G}(k^{-1})} w_v^k \exp\left\{y_v+\mu(v)+\sum_{i=1}^{n} c(v,v_i^k)\right\}-\frac{1}{2}y^{\mathsf{T}}\Sigma_k^{-1}y\\
=&\sum_{i=1}^{n} z_{v_i^k}-\frac{1}{2}\sum_{i,j=1}^{n} c(v_i^k,v_j^k)-\sum_{v\in\mathcal{G}(k^{-1})} w_v^k\exp\{z_v+\mu(v)\}-\frac{1}{2}z^{\mathsf{T}}\Sigma_k^{-1}z
\end{aligned}$$

which agrees with the exponent in (34). Thereby the proof is completed.

References

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

S.N. Chiu, D. Stoyan, W.S. Kendall, and J. Mecke. *Stochastic Geometry and Its Applications*. Wiley, Chichester, third edition, 2013.

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.

H. O. Georgii. Canonical and grand canonical Gibbs states for continuum systems. *Communications in Mathematical Physics*, 48:31–51, 1976.

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

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

P. McCullagh and J. Møller. The permanental process. *Advances in Applied Probability*, 38:873–888, 2006.

J. Møller and R.P. Waagepetersen. *Statistical Inference and Simulation for Spatial Point Processes*. Chapman and Hall/CRC, Boca Raton, 2003.

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

X. Nguyen and H. Zessin. Integral and differential characterizations of Gibbs processes. *Mathematische Nachrichten*, 88:105–115, 1979.

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.

R. Waagepetersen. Convergence of posteriors for discretized log Gaussian Cox processes. *Statistics and Probability Letters*, 66:229–235, 2004.

Acknowledgements

This research was supported by Centre for Stochastic Geometry and Advanced Bioimaging, funded by a grant from the Villum Foundation.