8 May 2002

SHOT NOISE COX PROCESSES

JESPER MØLLER,* *Aalborg University*

Abstract

Shot noise Cox processes constitute a large class of Cox and Poisson cluster processes in $\mathbb{R}^d$, including Neyman-Scott, Poisson-Gamma, and shot noise G Cox processes. It is demonstrated that due to the structure of such models, a number of useful and general results can easily be established. The focus is on the probabilistic aspects with a view to statistical applications, particularly results for summary statistics, reduced Palm distributions, simulation with or without edge effects, conditional simulation of the intensity function, and local and global Markov properties.

Keywords: Conditional simulation; Cox processes; edge effects; Markov point processes; nearest-neighbour Markov point processes; Neyman-Scott processes; Palm distributions; Poisson cluster processes; Poisson-Gamma processes; shot noise G Cox processes; simulation; spatial point processes; summary statistics

AMS 2000 Subject Classification: Primary 60D05;60G55;62M30

Secondary 60K35;68U20

1. Introduction

Cox process models constitute one of the most important and applicable classes of point process models for aggregated or clustered point patterns caused by e.g. an (usually unobserved) environmental random heterogeneity, and they play a major role in stochastic geometry and spatial statistics, see e.g. [14, 15, 47] the references therein. Briefly, a Cox process is the natural extension of a Poisson process, considering the intensity measure of the Poisson process as a realisation of a random measure. When a Cox process X is defined on the d-dimensional Eucledian space $\mathbb{R}^d$ it is usually specified by a random field $Z(\xi) \geq 0$, $\xi \in \mathbb{R}^d$, so that the conditional distribution of X given Z is a Poisson point process on $\mathbb{R}^d$ with intensity function Z. Many recent papers

* Postal address: Department of Mathematical Sciences, Aalborg University Fredrik Bajers Vej 7G, DK-9220 Aalborg, Denmark. Email address: jm@math.auc.dk

[6, 7, 8, 9, 10, 12, 19, 26, 35, 50] deal with simulation-based inference for new flexible model classes of such Cox processes, cf. the surveys [34, 38].

This paper is concerned with *shot noise Cox processes (SNCP)*, i.e. when Z is of the form

$$Z(\xi) = \sum_j \gamma_j k(c_j, \xi) \tag{1}$$

where $k(c_j, \cdot)$ is a kernel (i.e. a density function for a continuous d-dimensional random variable) and the $(c_j, \gamma_j) \in \mathbb{R}^d \times (0, \infty)$ are the points of a Poisson point process Φ on $\mathbb{R}^d \times (0, \infty)$; further details and conditions are given in Section 2. This is a rich class of Cox process models which includes Neyman-Scott processes [39], Poisson-gamma processes [50], and shot noise G Cox processes [8] as special cases, cf. Section 2. As discussed in more detail in Sections 2 and 4, SNCPs can also be viewed as a large class of Poisson cluster processes.

The focus in the paper is on the probabilistic aspects of SNCPs with a view to statistical applications, and the aim is both to give a unified and self-contained exposition and to present a number of new results. Particularly, we demonstrate that the structure of SNCPs allow us easily to establish various useful and general results.

The paper is organised as follows. Section 2 provides some background material. Section 3 concerns results for the summary statistics and the reduced Palm distributions of a SNCP. The latter turn out to be of a particular simple form. Section 4 studies different simulation algorithms for simulation of a SNCP within a bounded window W, and for conditional simulation of Φ given the restriction of a SNCP within W. In particular the role of edge effects is discussed. Finally, Section 5 deals with local and global Markov properties of SNCPs when the kernel has a bounded support, and the similarities and differences to usual Gibbs or Markov point processes are clarified.

2. Conditions and examples

This section provides some background material, conditions, and examples used throughout the text.

We require that the Poisson point process Φ is specified by a locally finite diffuse intensity measure ζ, i.e. $\zeta(D)$ is defined for Borel sets $D \subseteq \mathbb{R}^d \times (0, \infty)$, $\zeta(D) < \infty$ whenever D is bounded, and $\zeta(\{(c, \gamma)\}) = 0$ for all $(c, \gamma) \in \mathbb{R}^d \times (0, \infty)$. Moreover, Z is assumed to be almost surely locally integrable, i.e. with probability one, $\int_B Z(\xi)\mathrm{d}\xi < \infty$ for bounded Borel sets $B \subset \mathbb{R}^d$. These restrictions become convenient for establishing the results in Sections 3–4. As verified later in Proposition 1, the intensity function $\rho(\xi) = \mathbb{E}Z(\xi)$ of X is given by

$$\rho(\xi) = \int \gamma k(c,\xi)\mathrm{d}\zeta(c,\gamma) \tag{2}$$

provided the integral is finite for all $\xi \in \mathbb{R}^d$ — in turn this condition implies that Z is almost surely locally integrable.

It is often convenient to view a SNCP as a *cluster process*, meaning that $X|\Phi$ is distributed as the superposition $\bigcup_j X_j$ of independent Poisson processes X_j, $(c_j,\gamma_j) \in \Phi$, where X_j has intensity function $\gamma_j k(c_j,\cdot)$. The point process $\mathcal{C}$ of c_j's is countable but not necessarily locally finite (this is exemplified in Example 3 below). However, for any bounded $B \subset \mathbb{R}^d$, with probability one $X_B \equiv X \cap B$ is finite, and so only finite many X_j has points in B. We refer to $\mathcal{C}$ as the *centre process* and to X_j as the *cluster* with *centre* c_j, *intensity* γ_j, and *dispersion density* $k(c_j,\cdot)$. In the literature the centre points are also called *parent* or *mother* points and the clusters for *offspring* or *daughter* points.

Often in applications ζ is of the form

$$\zeta(D) = \int\int_{(c,\gamma)\in D} \mathrm{d}c\mathrm{d}\chi(\gamma) \tag{3}$$

where χ is a locally finite measure on $(0,\infty)$. This is equivalent to assume that the distribution of Φ is invariant under translations of $\mathcal{C}$ in $\mathbb{R}^d$, and it implies that $\mathcal{C}$ is independent of the point process of cluster-intensities. By (2),

$$\rho(\xi) = \int k(c,\xi)\mathrm{d}c \int \gamma \mathrm{d}\chi(\gamma) \tag{4}$$

provided the integrals are finite. If furthermore $k(c,\xi) = k(\xi - c)$ is invariant under translations in $\mathbb{R}^d$ (where we abuse the notation and let $k(\cdot)$ denote a density function for a continuous d-dimensional random variable), we have that Z and hence X is stationary, i.e. their distributions are invariant under translations in $\mathbb{R}^d$. For short we refer to this as *the stationary case of a SNCP*. Note that the intensity then reduces to

$$\rho = \int \gamma \mathrm{d}\chi(\gamma)$$

provided the integral is finite. Finally, if also $k(\xi) = k(\|\xi\|)$ is isotropic (where we again abuse the notation), the distribution of X is invariant under motions in $\mathbb{R}^d$.

Example 1. In the stationary case of a SNCP where χ is concentrated at the parameter α (i.e. $\chi((0,\infty)) = \chi(\{\alpha\}) > 0$), we have a Neyman-Scott process [39] with intensity $\rho = \alpha\kappa$, where $\kappa = \chi(\{\alpha\})$ is the intensity of the centre process, see e.g. [47]. Two mathematically tractable models are

(I) a *Matérn cluster process* [27, 28] where

$$k(\xi) = \mathbf{1}[\|\xi\| \le r]/(\omega_d r^d)$$

is the uniform density on the ball $b(0,r)$ in $\mathbb{R}^d$ with centre 0 and radius $r > 0$, and

$$\omega_d = \pi^{d/2}/\Gamma(1 + d/2)$$

is the volume of $b(0,1)$;

(II) a *Thomas process* [49] where

$$k(\xi) = \exp(-\|\xi\|^2/(2\omega^2))/(2\pi\omega^2)^{d/2}$$

is the density for d independent normally distributed variables with mean 0 and variance $\omega^2 > 0$.

In both cases k is isotropic. We refer to the kernels in (I) and (II) several times in the sequel.

Example 2. Suppose that ζ is of the form (3) where $\kappa = \chi((0,\infty)) < \infty$. Then $\mathcal{C}$ is a stationary Poisson process with intensity κ and the cluster-intensities are independent of $\mathcal{C}$ and i.i.d. with distribution $Q = \chi/\kappa$. A Neyman-Scott process is clearly a special case of this model.

Example 3. A *shot noise G Cox process* [8] is a SNCP with ζ of the form (3) where χ is absolutely continuous with respect to Lebesgue measure with density

$$f_{\kappa,\alpha,\tau}(\gamma) = \kappa\gamma^{-\alpha-1}\exp(-\tau\gamma)/\Gamma(1-\alpha), \quad \gamma > 0. \tag{5}$$

Here "G" refers to that

$$\mu_G(B) = \sum_j \gamma_j \mathbf{1}[c_j \in B]$$

is a so-called G-measure on the Borel σ-algebra in $\mathbb{R}^d$, see [8] or [38]. Further $\kappa > 0$, $\alpha < 1$, and $\tau \geq 0$ are parameters with $\tau > 0$ if $\alpha \leq 0$. These restrictions are equivalent to local finiteness of ζ.

The intensity function exists only for $\tau \neq 0$. Then

$$\rho(\xi) = \kappa \tau^\alpha \int k(c, \xi) \mathrm{d}c$$

which reduces to $\rho = \kappa \tau^\alpha$ in the stationary case.

The distribution of Φ depends much on α as described below.

The case $\alpha < 0$: Then Φ is a special case of the marked Poisson process in Example 2, i.e. $\mathcal{C}$ is a stationary Poisson process with intensity $-\kappa \tau^\alpha / \alpha$, the γ_j are i.i.d. and independent of $\mathcal{C}$, and each γ_j is gamma distributed with shape parameter $-\alpha$ and inverse scale parameter τ.

The case $0 \leq \alpha < 1$: The situation is now less simple as $\mathcal{C}$ is not locally finite, since $\int_0^\infty f_{\kappa,\alpha,\tau}(\gamma)\mathrm{d}\gamma = \infty$. As $\{(c, \gamma) \in \Phi : c \in A\}$ and $\{(c, \gamma) \in \Phi : c \in B\}$ are independent for disjoint Borel sets $A, B \subseteq \mathbb{R}^d$, we can for simplicity assume that $\mathcal{C}$ is concentrated on a bounded Borel set B with Lebesgue measure $|B|$. Then the γ_j define an inhomogeneous Poisson process on $(0, \infty)$ with intensity function $|B| f_{\kappa,\alpha,\tau}$, the c_j are i.i.d. and independent of the γ_j, and each c_j is uniformly distributed on B.

For $\alpha = 0$, we have a *Poisson-gamma process* as μ_G is a so-called gamma-measure [14, 50].

3. Summary statistics

In this section we exploit the form (1) of Z and the so-called Slivnyak-Mecke theorem to establish different useful results for the summary statistics of a SNCP X. As is custom in point process theory, we abuse notation and write e.g. $X \setminus \eta \equiv X \setminus \{\eta\}$ when we delete a point η from X, and $X \cup \xi \equiv X \cup \{\xi\}$ when we add a point $\xi \in \mathbb{R}^d$ to X.

The following lemma presents the Slivnyak-Mecke theorem in terms of Φ and $X|\Phi$, respectively. The lemma follows from Theorem 3.1 in [29] which also states that the Poisson processes Φ and $X|\Phi$ are uniquely characterised by the equations (6) and (7). For the point process X, we let $\mathbb{R}^d$ be equipped with the usual Borel σ-algebra, N_{lf} denotes the set of locally finite subsets of $\mathbb{R}^d$, N_{lf} is equipped with the usual σ-algebra

$\mathcal{N}_{\text{lf}}$ generated by the sets $F_{B,n} = \{x \in N_{\text{lf}} : \text{card}(x_B) = n\}$ for $n = 0, 1, \ldots$ and bounded Borel sets $B \subset \mathbb{R}^d$, and $\mathcal{N}_{\text{lf}} \times \mathbb{R}^d$ is the corresponding product σ-algebra. For the point process Φ on $\mathbb{R}^d \times (0, \infty)$, we define the corresponding σ-algebras in a similar way.

Lemma 1. *We have that*

$$\mathbb{E} \sum_{(c,\gamma)\in\Phi} f(\Phi \setminus (c,\gamma), (c,\gamma)) = \int \mathbb{E} f(\Phi, (c,\gamma)) \mathrm{d}\zeta(c,\gamma) \tag{6}$$

for nonnegative measurable functions f, and

$$\mathbb{E}\left[\sum_{\xi\in X} h(X \setminus \xi, \xi) \middle| \Phi \right] = \int Z(\xi) \mathbb{E}[h(X,\xi)|\Phi] \mathrm{d}\xi \tag{7}$$

for nonnegative measurable functions h.

3.1. First and second order characteristics

Expressions for the product moments $\mathbb{E}[Z(\xi_1)\cdots Z(\xi_n)]$ can easily be obtained by combining (1) and (6). Below we concentrate on the two first moments, i.e. the *intensity function* $\rho(\xi) = \mathbb{E}Z(\xi)$ and the *pair correlation function* $g(\xi,\eta) = \mathbb{E}[Z(\xi)Z(\eta)]/[\rho(\xi)\rho(\eta)]$ (provided the means exist, and taking $0/0 = 0$). These two functions are the fundamental characteristics or *summary statistics* of the first and second order properties for a spatial point process, see e.g. [47]. Moreover, Ripley's K-function [41, 42] for the stationary case and its extension [3] for the nonstationary case can be obtained by integrating the pair correlation function.

Proposition 1. *The intensity function exists and is given by*

$$\rho(\xi) = \int \gamma k(c,\xi) \mathrm{d}\zeta(c,\gamma) \tag{8}$$

provided the integral is finite for all $\xi \in \mathbb{R}^d$. Furthermore, the pair correlation function exists and is given by

$$g(\xi,\eta) = 1 + \beta(\xi,\eta)/(\rho(\xi)\rho(\eta)) \tag{9}$$

provided the integral

$$\beta(\xi,\eta) = \int \gamma^2 k(c,\xi) k(c,\eta) \mathrm{d}\zeta(c,\gamma)$$

is finite for all $\xi, \eta \in \mathbb{R}^d$.

Proof. The first result follows immediately from (1) and (6).

Suppose that $\beta(\xi,\eta) < \infty$ for all $\xi,\eta \in \mathbb{R}^d$. Then by (8) and Jensen's inequality, $\rho(\xi)^2 \le \beta(\xi,\xi) < \infty$, so $\rho(\xi) < \infty$ for all $\xi \in \mathbb{R}^d$. Using (6) twice we obtain that

$$\begin{aligned}&\mathbb{E}\sum_{j\neq k} f(\Phi \setminus \{(c_j,\gamma_j),(c_k,\gamma_k)\},(c_j,\gamma_j),(c_k,\gamma_k))\\ &= \int\int \mathbb{E}f(\Phi,(c,\gamma),(c',\gamma'))\mathrm{d}\zeta(c,\gamma)\mathrm{d}\zeta(c',\gamma')\end{aligned}$$

for nonnegative measurable functions f. Combining this with (6) and (8), we see that

$$\begin{aligned}&\mathbb{E}[Z(\xi)Z(\eta)]\\ &= \mathbb{E}\sum_{j\neq k}\gamma_j k(c_j,\xi)\gamma_k k(c_k,\eta) + \mathbb{E}\sum_j \gamma_j^2 k(c_j,\xi)k(c_j,\eta)\\ &= \int\int \gamma k(c,\xi)\gamma' k(c',\eta)\mathrm{d}\zeta(c,\gamma)\mathrm{d}\zeta(c',\gamma') + \int \gamma^2 k(c,\xi)k(c,\eta)\mathrm{d}\zeta(c,\gamma)\\ &= \rho(\xi)\rho(\eta) + \beta(\xi,\eta)\end{aligned}$$

whereby (9) is verified.

Thus $g \ge 1$, in accordance with the usual interpretation that $g \ge 1$ indicates aggregation of the points in X, see e.g. [38, 48]. Roughly speaking Proposition 1 implies that the first and second order properties of a SNCP depend only on the choice of kernel! This is exemplified below.

Example 4. Suppose that $k(c,\xi) = k(\xi - c)$ and consider first a stationary Neyman-Scott process (Example 1). By (8) and (9),

$$\rho = \alpha\kappa, \quad g(\xi,\eta) = 1 + \varphi(\xi-\eta)/\kappa, \tag{10}$$

where

$$\varphi(\xi) = \int k(\eta)k(\xi+\eta)\mathrm{d}\eta$$

is the density for the difference between two independent points which each have density k. For a Thomas process this reduces to

$$\varphi(\xi) = \left(4\pi\omega^2\right)^{-d/2}\exp\left(-\|\xi\|^2/\left(4\omega^2\right)\right),$$

while expressions for a Matérn cluster process can be found in [46, 47].

Consider next a stationary shot noise G Cox process with $\tau \neq 0$ (Example 3). Then

$$\rho = \kappa\tau^{\alpha}, \quad g(\xi,\eta) = 1 + \frac{1-\alpha}{\kappa\tau^{\alpha+2}}\varphi(\xi-\eta). \tag{11}$$

The class $\{(\rho,g) : \kappa > 0, \alpha > 0\}$ obtained from (10) and the class $\{(\rho,g) : \kappa > 0, \alpha < 1, \tau > 0\}$ obtained from (11) both agree with $\{(\rho, 1+a\varphi) : \rho > 0, a > 0\}$, so for any given kernel k we cannot distinguish between these classes. For instance, imagine that we have obtained nonparametric estimates $\hat{\rho}$ and $\hat{g}$, using only the assumption that X is a stationary point process, see e.g. [37, 48]. Then we cannot distinguish between whether $(\hat{\rho},\hat{g})$ fits a Neyman-Scott process or a shot noise G Cox process.

3.2. Reduced Palm distributions and further results for summary statistics

We now establish a very simple description of the reduced Palm distributions of a SNCP which e.g. makes it very easy to simulate from these distributions by the methods in Section 4. Related results for infinite divisible point processes can be found in [1, 22] and [29] (Theorem 6.1).

Suppose that X is a SNCP with intensity function ρ. For $\xi \in \mathbb{R}^d$, let $P^!_\xi$ denote the reduced Palm distribution at the point ξ. The reduced Palm distributions are for Lebesgue almost all $\xi \in \mathbb{R}^d$ with $\rho(\xi) > 0$ uniquely defined by the equations,

$$\mathbb{E}\sum_{\xi\in X} h(X\setminus\xi,\xi) = \int\int h(x,\xi)\mathrm{d}P^!_\xi(x)\rho(\xi)\mathrm{d}\xi \tag{12}$$

for nonnegative measurable functions h, cf. [14, 47].

Proposition 2. *For $\rho(\xi) > 0$, let*

$$Z_\xi(\eta) = \gamma_\xi k(c_\xi,\eta), \quad \eta \in \mathbb{R}^d,$$

where (c_ξ,γ_ξ) is a random variable with distribution

$$P((c_\xi,\gamma_\xi) \in D) = \int_D \gamma k(c,\xi)\mathrm{d}\zeta(c,\gamma)\,/\rho(\xi) \tag{13}$$

for Lebesgue sets $D \subseteq \mathbb{R}^d \times (0,\infty)$. Assume that $X_\xi|(\gamma_\xi,c_\xi)$ is a Poisson process with intensity function Z_ξ, and (γ_ξ,c_ξ,X_ξ) is independent of (Φ,X). Then for Lebesgue almost all $\xi \in \mathbb{R}^d$ with $\rho(\xi) > 0$,

$$P^!_\xi(F) = P(X \cup X^!_\xi \in F), \quad F \in \mathcal{N}_{\mathrm{lf}}. \tag{14}$$

Proof. The reduced Palm distributions are for Lebesgue almost all ξ with $\rho(\xi) > 0$ uniquely characterised by the equations

$$\mathbb{E}\sum_{\xi\in X} \mathbf{1}[\xi \in A, (X\setminus\xi)\cap B = \emptyset] = \int_A \mathrm{P}^!_\xi(F_{B,0})\rho(\xi)\mathrm{d}\xi \tag{15}$$

for bounded Borel sets $A, B \subset \mathbb{R}^d$, where $F_{B,0} = \{x \in N_{\mathrm{lf}} : x \cap B = \emptyset\}$. This follows from the following facts. It is well known that $\mathcal{N}_{\mathrm{lf}}$ is generated by the events $F_{B,0}$ for bounded Borel sets $B \subset \mathbb{R}^d$. Hence the class of sets $F_{B,0} \times A$ for bounded Borel sets $A, B \subset \mathbb{R}^d$, generates the product σ-algebra on $N_{\mathrm{lf}} \times \mathbb{R}^d$. This class is closed under intersection, so (15) and standard measure theoretical methods imply (12).

Now, for bounded Borel sets $A, B \subset \mathbb{R}^d$,

$$\begin{aligned}
&\mathbb{E}\sum_{\xi\in X} \mathbf{1}[\xi \in A, (X\setminus\xi)\cap B = \emptyset]\\
&= \mathbb{E}\Big(\mathbb{E}\Big(\sum_{\xi\in X} \mathbf{1}[\xi \in A, (X\setminus\xi)\cap B = \emptyset]\Big|\Phi\Big)\Big)\\
&= \mathbb{E}\int_A \sum_{(c,\gamma)\in\Phi} \gamma k(c,\xi)\exp\Big(-\int_B \sum_{(c',\gamma')\in\Phi} \gamma' k(c',\eta)\mathrm{d}\eta\Big)\mathrm{d}\xi\\
&= \int_A \mathbb{E}\sum_{(c,\gamma)\in\Phi} \gamma k(c,\xi)\exp\Big(-\int_B \sum_{(c',\gamma')\in\Phi} \gamma' k(c',\eta)\mathrm{d}\eta\Big)\mathrm{d}\xi\\
&= \int_A\int \mathbb{E}\gamma k(c,\xi)\exp\Big(-\int_B \sum_{(c',\gamma')\in\Phi\cup(c,\gamma)} \gamma' k(c',\eta)\mathrm{d}\eta\Big)\mathrm{d}\zeta(c,\gamma)\mathrm{d}\xi\\
&= \int_A\int \gamma k(c,\xi)\exp\Big(-\int_B \gamma k(c,\eta)\mathrm{d}\eta\Big)\mathrm{d}\zeta(c,\gamma)\mathrm{d}\xi\\
&\qquad\times\mathbb{E}\exp\Big(-\int_B\sum_{(c,\gamma)\in\Phi}\gamma k(c,\eta)\mathrm{d}\eta\Big)\\
&= \int_A P(X^!_\xi\cap B=\emptyset)\rho(\xi)\mathrm{d}\xi \times P(X\cap B=\emptyset)\\
&= \int_A P((X\cup X^!_\xi)\cap B=\emptyset)\rho(\xi)\mathrm{d}\xi
\end{aligned}$$

using (7) to obtain the second identity, and (6) for the fourth identity. Hence the $P^!_\xi$ given by (14) for $\rho(\xi) > 0$ satisfy (15), and so they are reduced Palm distributions.

By (14) we can view $P^!_\xi$ as the distribution of a Cox process with random intensity function $Z^!_\xi(\eta) = Z(\eta) + Z_\xi(\eta)$. If ζ is of the form (3), then (4) and (13) imply that γ_ξ and c_ξ are independent with $P(\gamma_\xi \in A) \propto \int_A \gamma\mathrm{d}\chi(\gamma)$ for Borel sets $A \subseteq (0,\infty)$ and

$P(c_\xi \in B) \propto \int_B k(c,\xi)\mathrm{d}c$ for Borel sets $B \subseteq \mathbb{R}^d$. In the stationary case with intensity $\rho > 0$, the reduced Palm distributions exist for all $\xi \in \mathbb{R}^d$, since we can take $\gamma_\xi = \gamma_0$, $c_\xi = c_0 + \xi$, and $X_\xi = X_0 + \xi$, so that $P^!_\xi(F) = P(X \cup (X_0 + \xi) \in F)$.

For the rest of this section we consider the stationary case with intensity $\rho > 0$. Apart from the summary statistics mentioned in Section 3.1, one usually consider the *empty space function* $F(r) = P(\mathrm{dist}(0, X) \leq r)$ and the *nearest-neighbour function* $G(r) = P^!_0(\mathrm{dist}(0, X \cup X_0) \leq r)$, where $r \geq 0$ and

$$\mathrm{dist}(A, B) = \inf\{\|\xi - \eta\| : \xi \in A, \eta \in B\}$$

denotes the shortest distance between two sets $A, B \subseteq R^d$ (we set $\mathrm{dist}(A, B) = \infty$ if A or B is empty). In general no simple expressions seem available for these two summary statistics, but they can at least by approximated by simulations using Proposition 2 and the methods in Section 4. However, the so-called J-function introduced in [25] and defined by $J(r) = (1 - G(r))/(1 - F(r))$ for $F(r) < 1$, may be calculated by numerical methods due to the following result.

Corollary 1. *For a stationary SNCP with intensity $\rho > 0$,*

$$J(r) = \frac{1}{\rho} \int \int \gamma k(c) \exp\left(- \int_{b(c,r)} \gamma k(\eta)\mathrm{d}\eta \right) \mathrm{d}c\mathrm{d}\chi(\gamma), \quad \textit{for all } r \geq 0.$$

Proof. By definition of J and Proposition 2,

$$\begin{aligned} J(r) &= P((X \cup X_0) \cap b(0,r) = \emptyset)/P(X \cap b(0,r) = \emptyset) = P(X_0 \cap b(0,r) = \emptyset) \\ &= \frac{1}{\rho} \int \int \gamma k(-c) \exp\left(- \int_{\|\eta\| \leq r} \gamma k(\eta - c)\mathrm{d}\eta \right) \mathrm{d}c\mathrm{d}\chi(\gamma) \end{aligned}$$

whereby the result follows.

Corollary 1 extends the results in [25] for Neyman-Scott processes. It follows that $J(0) = 1$ and $J(r) < 1$ is nonincreasing for $r > 0$ with

$$\lim_{r \to \infty} J(r) = \int \gamma \exp(-\gamma)\mathrm{d}\chi(\gamma)/\rho.$$

If the kernel has finite range R, i.e. $k(\xi) = 0$ for $\|\xi\| \geq R$, then $J(r)$ is constant for $r \geq 2R$.

4. Simulation and easy constructions of shot noise Cox processes

Although SNCPs are tractable for mathematical analysis, at least when compared to many other types of point process models, simulation is needed for performing statistical inference, cf. [7, 8, 9, 10, 11, 38, 50]. In this section we describe various simulation algorithms for both the restriction X_W of X to a bounded Borel set $W \subset \mathbb{R}^d$ with volume $|W| > 0$, and for Φ given X_W.

We shall several times exploit the following lemma.

Lemma 2. *Consider* $X = \bigcup_j X_j$ *as a cluster process, cf. Section 2, but suppose only that for all* $c \in \mathbb{R}^d$, $k(c, \cdot)$ *is a nonnegative Lebesgue integrable function. That is* Φ *is a Poisson process with intensity measure* ζ *as usual, and conditional on* Φ*, the clusters* X_j *are independent Poisson processes, and* X_j *associated to* $(c_j, \gamma_j) \in \Phi$ *has intensity function* $\gamma_j k(c_j, \cdot)$*. Let* $B \subseteq \mathbb{R}^d$ *be a given Borel set, and define*

$$\Phi_B = \{(c_j, \gamma_j) \in \Phi : X_j \cap B \neq \emptyset\},$$

$$p_B(c, \gamma) = 1 - \exp\left(- \gamma \int_B k(c, \xi) \mathrm{d}\xi \right),$$

and

$$\zeta_B(D) = \int_D p_B(c, \gamma) \mathrm{d}\zeta(c, \gamma)$$

for Borel sets $D \subseteq \mathbb{R}^d \times (0, \infty)$*. Then*

- **(i)** $p_B(c, \gamma)$ *is the probability that a cluster associated to* $(c, \gamma) \in \Phi$ *has at least one point falling in* B*;*

- **(ii)** Φ_B *and* $\Phi \setminus \Phi_B$ *are Poisson processes with intensity measures* ζ_B *and* $\zeta - \zeta_B$*, respectively, and* $(\Phi_B, \{X_j : (c_j, \gamma_j) \in \Phi_B\})$ *and* $(\Phi \setminus \Phi_B, \{X_j : (c_j, \gamma_j) \in \Phi \setminus \Phi_B\})$ *are independent;*

- **(iii)** *conditional on* Φ_B*, the clusters* X_j *with* $(c_j, \gamma_j) \in \Phi_B$ *are independent, and* X_j *is distributed as a Poisson process with intensity function* $\xi \to \gamma_j k(c_j, \xi)$ *when we have conditioned on that this Poisson process has at least one point in* B*;*

- **(iv)** *conditional on* Φ_B*, the point processes* $X_j \cap B$ *with* $(c_j, \gamma_j) \in \Phi_B$ *are independent, and* $X_j \cap B$ *is distributed as a Poisson process with intensity function* $\xi \to$

$\gamma_j k(c_j,\xi)\mathbf{1}[\xi \in B]$ when we have conditioned on that this Poisson process is nonempty;

(v) *conditional on $\Phi \setminus \Phi_B$, the clusters X_j with $(c_j,\gamma_j) \in \Phi \setminus \Phi_B$ are independent, and X_j is a Poisson process with intensity function $\xi \to \gamma_j k(c_j,\xi)\mathbf{1}[\xi \notin B]$.*

Proof. (i), (iii), and (v) follow immediately from the description of X as a cluster process. Since Φ_B is obtained by independent thinning of Φ with retention probability p_B we obtain immediately (ii). Similarly, (iii) implies immediately (iv).

4.1. Simulation of SNCPs with edge effects and truncation

Simulation of X_W is most easily done by applying Lemma 2. *Edge effects* play a role, as X_W is a Cox process with random intensity measure

$$Z_W(\xi) = \mathbf{1}[\xi \in W]Z(\xi) = \sum_j \mathbf{1}[\xi \in W]\gamma_j k(c_j,\xi) \tag{16}$$

which depends on all those centres $c_j \in \mathcal{C}$ with $k(c_j,\xi) > 0$ for some $\xi \in W$. We consider an extended bounded region $W_{\text{ext}} \supseteq W$ so that points in a cluster X_j with centre $c_j \notin W_{\text{ext}}$ fall in W with a negligible probability, cf. (i) in Lemma 2. Moreover, as exemplified in Example 5 below, only finite many $(c_j,\gamma_j) \in \Phi$ with $c_j \in W_{\text{ext}}$ are used in the simulations.

Example 5. Consider again Example 3 in the case $0 \leq \alpha < 1$. Then $\mathcal{C} \cap W_{\text{ext}}$ is infinite. Defining $q(t) = \zeta(W_{\text{ext}} \times [t,\infty))$ for $t > 0$, then q is a strictly decreasing function which maps $(0,\infty)$ onto $(0,\infty)$, and we can write

$$\{(c,\gamma) \in \Phi : c \in \mathcal{C} \cap W_{\text{ext}}\} = \{(c_1,\gamma_1),(c_2,\gamma_2),\ldots\}$$

where $q(\gamma_1) < q(\gamma_2) < \ldots$ are the points of a unit rate Poisson process on $(0,\infty)$. For simulation and inference, one approximates Z_W by

$$Z_W(\xi) \approx \sum_{j=1}^{J} \gamma_j k(c_j,\xi)$$

where $J < \infty$ is a "cut off". For a discussion on how q^{-1} and the tail sum $\sum_{j>J}\gamma_j$ can be evaluated, see [8, 50].

Returning to our general setting of a SNCP, we let the cut off be determined by ignoring those $(c_j, \gamma_j) \in \Phi$ with $\gamma_j \le \epsilon$, where $\epsilon \ge 0$ is a user-specified parameter (we include here the case $\epsilon = 0$ which means no truncation at all). Then X_W is approximated by $\bigcup\{X_j \cap W : (c_j, \gamma_j) \in \Phi \cap (W_{\text{ext}} \times (\epsilon, \infty))\}$ in the simulations, where we first simulate the Poisson process $\Phi \cap (W_{\text{ext}} \times (\epsilon, \infty))$, and next the associated independent processes $X_j \cap W$, where by (iv) in Lemma 2, $X_j \cap W$ is a Poisson process with intensity function $\xi \to \gamma_j k(c_j, \xi)\mathbf{1}[\xi \in W]$. In order to evaluate the error of the approximation it is convenient to consider a function $k_W^{\text{dom}} : \mathbb{R}^d \times \mathbb{R}^d \to (0, \infty)$ so that

(C1) $k_W^{\text{dom}}(c, \xi) \ge k(c, \xi)$ if $\xi \in W$, and $k_W^{\text{dom}}(c, \xi) = 0$ if $\xi \notin W$;

(C2) we can easily calculate the integral

$$a_W^{\text{dom}}(c) = \int_W k_W^{\text{dom}}(c, \xi)\mathrm{d}\xi, \quad c \in \mathbb{R}^d.$$

We illustrate this in Example 6 below.

Proposition 3. *Let*

$$M = \sum_j \mathbf{1}[c_j \notin W_{\text{ext}} \textit{ or } \gamma_j \le \epsilon]\text{card}(X_j \cap W)$$

be the number of missing points when we have made a simulation of X_W *by ignoring those clusters* X_j *with* $c_j \notin W_{\text{ext}}$ *or* $\gamma_j \le \epsilon$. *Then*

$$\mathbb{E}M \le \int \mathbf{1}[c \notin W_{\text{ext}} \textit{ or } \gamma \le \epsilon]\gamma a_W^{\text{dom}}(c)\mathrm{d}\zeta(c, \gamma).$$

Proof. By (6),

$$\begin{aligned} \mathbb{E}M &= \mathbb{E}\mathbb{E}[M|\Phi] = \mathbb{E}\sum_j \mathbf{1}[c_j \notin W_{\text{ext}} \text{ or } \gamma_j \le \epsilon] \int_W \gamma_j k(c_j, \xi)\mathrm{d}\xi \\ &= \int \mathbf{1}[c \notin W_{\text{ext}} \text{ or } \gamma \le \epsilon] \int_W \gamma k(c, \xi)\mathrm{d}\xi\mathrm{d}\zeta(c, \gamma) \end{aligned}$$

whereby (C1) implies the result.

Example 6. Let $W = b(0, R)$ and $W_{\text{ext}} = b(0, R+r)$, and consider the stationary case of a SNCP where the kernel $k(c, \xi) = k(\xi - c)$ is given by either the uniform kernel (I) or the Gaussian kernel (II) in Example 1. In case (I) $X_j \cap W = \emptyset$ whenever $c_j \notin W_{\text{ext}}$, so it is only the truncation of the γ_j which play a role for the error done when making simulations with $\epsilon > 0$. Setting $k_W^{\mathrm{dom}}(c,\xi) = \sup_{\eta \in W} k(c,\eta)$ for $\xi \in W$, (C1)–(C2) are satisfied: for $s = \|c\| \geq 0$, in case (I)

$$a_I(s) \equiv a_W^{\mathrm{dom}}(c) = (R/r)^d \mathbf{1}[s \leq R + r],$$

while in case (II)

$$a_{II}(s) \equiv a_W^{\mathrm{dom}}(c) = \left[\omega_d R^d / \left(2\pi\omega^2\right)^{d/2}\right] \exp\left[-\mathbf{1}[s > R]\left(s - R\right)^2/(2\omega^2)\right].$$

Thus in case (I)

$$\mathbb{E}M \leq \omega_d (R(R+r)/r)^d \int_0^\epsilon \gamma \mathrm{d}\chi(\gamma),$$

while in case (II) by making a shift to polar coordinates

$$\mathbb{E}M \leq \sigma_d \int\int \mathbf{1}[s > R + r \text{ or } \gamma \leq \epsilon] \gamma a_{II}(s) \mathrm{d}s \mathrm{d}\chi(\gamma)$$

where

$$\sigma_d = 2\pi^{d/2}/\Gamma(d/2)$$

is the surface area of $b(0,1)$. These integrals may easily be determined by numerical methods when χ has a density like e.g. (5).

4.2. Simulation of SNCPs without edge effects and truncation

Recently Brix and Kendall [11] showed how edge effects and truncation can be avoided when making simulation of a SNCP (which they call a Cox Poisson cluster process). This section provides a short and easy description of their method using Lemma 2.

The idea is to obtain a simulation of X_W by independent thinning of the nonempty clusters in a Cox process X_W^{dom} with random intensity function

$$Z_W^{\mathrm{dom}}(\xi) = \sum_j \gamma_j k_W^{\mathrm{dom}}(c_j, \xi) \tag{18}$$

where k_W^{dom} satisfies (C1)–(C2). By (C1), (16), and (18), X_W^{dom} dominates X_W in the sense that $Z_W^{\mathrm{dom}} \geq Z_W$. Let X_j^{dom}, $(c_j,\gamma_j) \in \Phi$, denote the clusters of X^{dom}, and let

$$\Phi_W^{\mathrm{dom}} = \{(c_j,\gamma_j) \in \Phi : X_j^{\mathrm{dom}} \neq \emptyset\}.$$

By (i)-(ii) in Lemma 2, Φ_W^{dom} is a Poisson process on $\mathbb{R}^d \times (0,\infty)$ with intensity measure

$$\zeta_W^{\mathrm{dom}}(D) = \int_D p_W^{\mathrm{dom}}(c,\gamma) \mathrm{d}\zeta(c,\gamma)$$

where

$$p_W^{\rm dom}(c,\gamma) = 1 - \exp\left(-\gamma \int_W k_W^{\rm dom}(c,\xi){\rm d}\xi \right) = 1 - \exp(-\gamma a_W^{\rm dom}(c)).$$

Finally, assume that

(C3)

$$\beta_W^{\rm dom} \equiv \int p_W^{\rm dom}(c,\gamma){\rm d}\zeta(c,\gamma) < \infty$$

and we can easily calculate $\beta_W^{\rm dom}$ (at least by numerical methods).

This means that $\Phi_W^{\rm dom}$ is almost surely finite, and we can make perfect simulations of X_W as follows, cf. Proposition 4 below.

Perfect simulation algorithm for SNCPs:

(a) generate the Poisson process $\Phi_W^{\rm dom} = \{(c_1,\gamma_1),\ldots,(c_N,\gamma_N)\}$;

(b) for each $j = 1,\ldots,N$ generate

- **(i)** $X_j^{\rm dom}$ which is distributed as a conditional Poisson process with intensity function $\gamma_j k_W^{\rm dom}(c_j,\cdot)$ given that it is nonempty;
- **(ii)** X_j' which is an independent thinning of $X_j^{\rm dom}$ with retention probabilities $k(c_j,\xi)/k_W^{\rm dom}(c_j,\xi)$ for $\xi \in X_j^{\rm dom}$;

(c) return $\bigcup_j X_j'$.

The generation of the Poisson process $\Phi_W^{\rm dom}$ in (a) is rather straightforward for the specific examples of SNCPs considered in this paper, cf. [11] and Example 7 below. For the loop in (b) it is implicit that the generation of processes in (i)–(ii) is independent of previous generations. The generation of $X_j^{\rm dom}$ in (i) will be straightforward in our examples where $k_W^{\rm dom}(c,\xi) = \sup_{\eta\in W} k(c,\eta)\mathbf{1}[\xi \in W]$ is constant for $\xi \in W$ when $c = c_j$ is fixed.

Proposition 4. *The output in (c) above follows the same distribution as X_W.*

Proof. We consider $X^{\rm dom} = \bigcup_j X_j^{\rm dom}$ as a cluster process (including the empty clusters!) where conditional on Φ the clusters are independent Poisson processes so that $X_j^{\rm dom}$ associated to $(c_j,\gamma_j) \in \Phi$ has intensity function $\gamma_j k_W^{\rm dom}(c_j,\cdot)$. By (iv) in

Lemma 2 we have correctly generated the nonempty clusters in (i) of the algorithm. If we are making an independent thinning of each of the empty or nonempty clusters X_j^{dom} in the same way as in (ii) of the algorithm (thinning an empty X_j^{dom} simply means that $X_j' = \emptyset$), we obtain that each X_j' given $(c_j, \gamma_j) \in \Phi$ is a Poisson process with intensity function $\xi \to \mathbf{1}[\xi \in W]\gamma_j k(c_j, \xi)$, and all these clusters X_j' are conditionally independent given Φ. Hence $\bigcup_j X_j'$ and X_W are identically distributed. Of course we need here only to consider the union of nonempty X_j', so the output in (c) and X_W are identically distributed.

Example 7. Let the situation be as in Example 6.

In case (I) the centres of Φ_W^{dom} are contained in $b(0, R+r)$, $p_W^{\mathrm{dom}}(c,\gamma) = 1 - \exp(-\gamma(R/r)^d)$ for $\|c\| \le R + r$, and

$$\beta_W^{\mathrm{dom}} = \omega_d (R+r)^d \int \big[1 - \exp\big(-\gamma(R/r)^d\big)\big] \mathrm{d}\chi(\gamma).$$

In case (II) $p_W^{\mathrm{dom}}(c,\gamma) = 1 - \exp(-\gamma a_{II}(\|c\|))$ is also a radically symmetric function of c which decays fast to zero, and

$$\beta_W^{\mathrm{dom}} = \sigma_d \int \int_0^\infty s^{d-1} \left[1 - \exp\left(-\frac{\gamma\omega_d R^d}{(2\pi\omega^2)^{d/2}} \exp\left(-\frac{\mathbf{1}[s>R]}{2\omega^2}\big(s-R\big)^2\right)\right)\right] \mathrm{d}s \mathrm{d}\chi(\gamma).$$

If χ is concentrated at α and $\kappa = \chi(\{\alpha\})$ as in Example 1, we have a Matérn cluster process in case (I) and a Thomas process in case (II). For the Matérn cluster process, the centres in Φ_W^{dom} form a homogeneous Poisson process on $b(0, R+r)$ with rate

$$\beta_W^{\mathrm{dom}} = \kappa\omega_d (R+r)^d \big[1 - \exp\big(-\alpha(R/r)^d\big)\big].$$

For the Thomas process,

$$\beta_W^{\mathrm{dom}} = \kappa\sigma_d \int_0^\infty s^{d-1} \left[1 - \exp\left(-\frac{\alpha\omega_d R^d}{(2\pi\omega^2)^{d/2}} \exp\left(-\frac{\mathbf{1}[s>R]}{2\omega^2}\big(s-R\big)^2\right)\right)\right] \mathrm{d}s$$

is finite and easily determined by numerical integration. We can first generate $N \sim \mathrm{po}(\beta_W^{\mathrm{dom}})$ and next the N i.i.d. centres c_j, where the direction of c_j is uniformly distributed and independent of $s_j = \|c_j\|$, and s_j has a density proportional to

$$s^{d-1} \left[1 - \exp\left(-\frac{\alpha\omega_d R^d}{(2\pi\omega^2)^{d/2}} \exp\left(-\frac{\mathbf{1}[s>R]}{2\omega^2}\big(s-R\big)^2\right)\right)\right], \quad s > 0$$

(each s_j can be generated by e.g. rejection sampling [43]).

Consider a shot noise G Cox process with parameter (α, τ, κ) as in Example 3. Then in case (I),

$$\beta_W^{\text{dom}} = \kappa\omega_d(R+r)^d[(\tau + (R/r)^d)^\alpha - \tau^\alpha]/\alpha \quad \text{if } \alpha < 0$$

and

$$\beta_W^{\text{dom}} = \frac{\kappa\omega_d(R+r)^d}{\Gamma(1-\alpha)} \int_0^\infty \left[1 - \exp\left(-(R/r)^d\gamma\right)\right]\gamma^{-\alpha-1}\exp(-\tau\gamma)\mathrm{d}\gamma \quad \text{if } 0 \leq \alpha < 1.$$

And in case (II),

$$\beta_W^{\text{dom}} = \frac{\kappa\sigma_d}{\Gamma(1-\alpha)} \int_0^\infty \int_0^\infty \left[1 - \exp\left(-\frac{\gamma\omega_d R^d}{(2\pi\omega^2)^{d/2}} \exp\left(-\frac{\mathbf{1}[s > R]}{2\omega^2}(s-R)^2\right)\right)\right] s^{d-1}\gamma^{-\alpha-1}\exp(-\tau\gamma)\mathrm{d}s\mathrm{d}\gamma.$$

In all cases β_W^{dom} is finite, and it is known or can be determined by numerical integration. Finally, Φ_W^{dom} is a finite inhomogeneous Poisson process with intensity function $(c, \gamma) \to (1 - \exp(-\gamma a_W^{\text{dom}}(c)))f_{\kappa,\alpha,\tau}(\gamma)$, and we can generate this along similar lines as above for the Thomas process.

4.3. Conditional simulation

Suppose we have observed a finite point configuration $X_W = x$, and let

$$f(x|\Phi) = \exp\left(|W| - \int_W Z(\xi)\mathrm{d}\xi\right) \prod_{\xi \in x} Z(\xi) \tag{20}$$

denote the density of X_W given Φ with respect to the unit rate Poisson on W. For simplicity we assume in this section that ζ has support $D = W_{\text{ext}} \times (\epsilon, \infty)$ where $W_{\text{ext}} \supseteq W$, $\epsilon > 0$, and $0 < \zeta(D) < \infty$. Then Φ given $X_W = x$ is a finite point process on D with unnormalised density $\pi(\phi|x) = f(x|\phi)$ with respect to the marginal distribution of Φ, where the normalising constant is the "likelihood" $L(x) = \mathbb{E}f(x|\Phi)$. A closed form expression of $L(x)$ is in general unknown. Thus simulation from $\pi(\cdot|x)$ is needed for making prediction of Φ as well as for performing likelihood and Bayesian inference based on MCMC methods, see e.g. [4, 7, 23, 24, 26, 38, 50].

We can use the Metropolis-Hastings algorithm in [18] for conditional simulation. Briefly, suppose $\phi = \{(c_1, \gamma_1), \ldots, (c_n, \gamma_n)\}$ (with $\pi(\phi|x) > 0$) is the current state of the Metropolis-Hastings chain. Then we make either a birth or a death proposal, each with probability 1/2. If a birth is proposed, we generate a point (c, γ) with distribution

$\zeta/\zeta(D)$, and with probability $\min\{1, r(\phi, (c,\gamma))\}$, where

$$r(\phi, (c,\gamma)) = \frac{\pi(\phi \cup (c,\gamma)|x)\zeta(D)}{\pi(\phi|x)(n+1)},$$

we let $\phi \cup (c,\gamma)$ be the next state of the chain, and we retain ϕ otherwise. If a death is proposed, we let $\Phi \setminus (c_i,\gamma_i)$ be the next state of the chain with probability $(1/n)\min\{1, 1/r(\phi \setminus (c_i,\gamma_i), (c_i,\gamma_i))\}$, $i = 1,\ldots,n$ (we set $1/r(\phi \setminus (c_i,\gamma_i), (c_i,\gamma_i)) = 0$ if $\pi(\phi \setminus (c_i,\gamma_i)) = 0$), and we retain ϕ otherwise. By (1) and (20),

$$\frac{\pi(\phi \cup (c,\gamma)|x)}{\pi(\phi|x)} = \exp\left(-\gamma \int_W k(c,\xi)\mathrm{d}\xi \right) \prod_{\xi \in x} \left(1 + \frac{\gamma k(c,\xi)}{\sum_{(c_j,\gamma_j)\in\phi} \gamma_j k(c_j,\xi)} \right) \qquad (21)$$

where the integral can easily be calculated for the uniform kernel (I) and the Gaussian kernel (II) (Example 1) if e.g. W is rectangular.

Alternative and more complicated algorithms using auxiliary techniques or by running a spatial birth-death process have been proposed [26, 50]. The Metropolis-Hastings algorithm described above is proposed because of its simplicity and since its theoretical properties are well-understood (see [31] for definitions of the following concepts): It is reversible with invariant (unnormalised) density $\pi(\cdot|x)$; this follows along similar lines as in [18], noticing that $\pi(\psi|x) > 0$ implies that $\pi(\psi \cup (c,\gamma)|x) > 0$ (however, $\pi(\psi \cup (c,\gamma)|x) > 0$ does not imply that $\pi(\psi|x) > 0$, since e.g. $\pi(\emptyset|x) = 0$ when $x \neq \emptyset$). Further, it is irreducible and aperiodic on the support

$$\begin{aligned}\Omega_x &= \{\phi \subset D : \mathrm{card}(\phi) < \infty, \pi(\phi|x) > 0\} \\ &= \{\phi \subset D : \mathrm{card}(\phi) < \infty, \text{ for each } \xi \in x \text{ exists } (c,\gamma) \in \phi \text{ with } k(c,\xi) > 0\}\end{aligned}$$

(basically because if $\pi(\phi|x) > 0$ and $\pi(\phi'|x) > 0$ then $\pi(\phi \cup \phi'|x) > 0$ and the Metropolis-Hastings chain can move from ϕ to $\phi \cup \phi'$ to ϕ', and it can stay at ϕ). Furthermore, Proposition 5 below states that under weak conditions on the kernel, the chain is geometrical ergodic. For example, the conditions are satisfied for the uniform and Gaussian kernels (I)–(II) considered in Example 6. Furthermore, we have that with probability one, $\int_{(c,\gamma)\in D: k(c,\xi)>0} \mathrm{d}\zeta(c,\gamma) > 0$ for all $\xi \in X_W$.

Proposition 5. *Let $x = \{x_1,\ldots,x_m\} \neq \emptyset$ so that $\int_{(c,\gamma)\in D: k(c,x_i)>0} \mathrm{d}\zeta(c,\gamma) > 0$, $i = 1,\ldots,m$. Assume there exist strictly positive constants $\delta, \delta_i', \delta_i''$ so that for all $c \in W_{\mathrm{ext}}$ and $i = 1,\ldots,m$ we have that $\int_W k(c,\xi)\mathrm{d}\xi \geq \delta$ and if $k(c,x_i) > 0$ then $\delta_i' \leq k(c,x_i) \leq \delta_i''$. Then the Metropolis-Hastings chain is geometrical ergodic.*

Proof. The corresponding result and proof in [17] do not immediately apply since they rely much on the assumption that $\pi(\cdot|x)$ is locally stable, that is

$$\pi(\phi \cup (c,\gamma)|x) \leq M\pi(\phi|x) \tag{22}$$

for some constant M and all $\phi \cup (c,\gamma) \in \Omega_x$ with $(c,\gamma) \notin \phi$. This is not satisfied in our case, since $\pi(\emptyset|x) = 0$. However, we start by verifying that (22) holds whenever $\phi \in \Omega_x$. Next we show that for any integer N, $\{\phi \in \Omega_x : \mathrm{card}(\phi) \leq N\}$ is a so-called small set for the chain. Thereby geometric ergodicity follows along similar lines as in the proof of Proposition 3.3. in [17].

Let $\phi \in \Omega_x$, and set $\delta' = \min\{\delta'_1, \ldots, \delta'_m\}$ and $\delta'' = \max\{\delta''_1, \ldots, \delta''_m\}$. By (21),

$$\frac{\pi(\phi \cup (c,\gamma)|x)}{\pi(\phi|x)} \leq \exp(-\gamma\delta) \prod_{i=1}^{m} \frac{\gamma\delta''_i + \epsilon\delta'}{\epsilon\delta'} \tag{23}$$

because of the conditions on the kernel and since $\gamma_j > \epsilon > 0$. The upper bound in (23) is a bounded function of $\gamma > \epsilon$ which attains its maximum at either the solution to the equation

$$\sum_{i=1}^{m} \frac{\delta''_i}{\gamma\delta''_i + \epsilon\delta'} = \delta$$

or, if there is no such solution, at the limit $\gamma = \epsilon$. Thus $\pi(\phi \cup (c,\gamma)|x)/\pi(\phi|x)$ is bounded from above by a constant M. Furthermore, it is bounded from below by $\exp(-\gamma\delta''|W|)$.

Since $\zeta(D) > 0$ and $\int_{(c,\gamma)\in D: k(c,x_i)>0} \mathrm{d}\zeta(c,\gamma) > 0$, $i = 1, \ldots, m$, there exists a constant $L > \epsilon$ so that $\zeta(W_{\mathrm{ext}} \times (\epsilon, L)) > 0$, $\int_{(c,\gamma)\in D: k(c,x_i)>0, \gamma_i < L} \mathrm{d}\zeta(c,\gamma) > 0$, $i = 1, \ldots, m$, and $\exp(-L\delta''|W|)\zeta(D) \leq 1$. Let $m' \geq \max\{m, M\zeta(D) - 1\}$ be an integer. Then, if $N \geq \mathrm{card}(\phi)$ and $P^t(\phi, F)$ denotes the transition probability for the Metropolis-Hastings chain when it starts in $\phi \in \Omega_x$ and after t-steps belongs to an event $F \subseteq \Omega_x$,

$$P^{m'+N}(\phi, F) \geq \frac{[\frac{1}{2}\exp(-L\delta''|W|)\zeta(D)]^{m'}}{(\mathrm{card}(\phi)+1)\cdots(\mathrm{card}(\phi)+m')} \times \frac{1}{2[\mathrm{card}(\phi)+m']\cdots 2[1+m']} \times \left[\frac{1}{2} - \frac{M\zeta(D)}{2(m'+1)}\right]^{N-\mathrm{card}(\phi)} \times \int_{(c_1,\gamma_1)\in D:\gamma_1<L} \cdots \int_{(c_{m'},\gamma_{m'})\in D:\gamma_{m'}<L} \mathbf{1}[\{(c_1,\gamma_1),\ldots,(c_{m'},\gamma_{m'})\} \in F] \frac{\mathrm{d}\zeta(c_1,\gamma_1)}{\zeta(D)} \cdots \frac{\mathrm{d}\zeta(c_{m'},\gamma_{m'})}{\zeta(D)} \tag{24}$$

corresponding to first adding $(c_1,\gamma_1),\ldots,(c_{m'},\gamma_{m'})$ to ϕ, next deleting the points in ϕ, and finally making no changes when $N - \mathrm{card}(\phi)$ births are proposed. Hence $P^{m'+N}(\phi, F) \geq \epsilon' Q(F)$ where

$$\epsilon' = \left(\frac{1}{2}\right)^{m'+N} \frac{[\exp(-L\delta''|W|)\zeta(D)]^{m'}}{(N+1)\cdots(N+m')} \frac{m'!}{(N+m')!} \left[1 - \frac{M\zeta(D)}{m'+1}\right]^{N}$$

is a strictly positive constant and $Q(F)$ denotes the integral (24). Note that Q is a finite nonzero measure, since if we set $x_j = x_m$ for $j \geq m$,

$$Q(\Omega_x) \geq \int_D \cdots \int_D \mathbf{1}[k(c_j, x_j) > 0, \gamma_j < L, j = 1,\ldots,m'] \frac{\mathrm{d}\zeta(c_1,\gamma_1)}{\zeta(D)} \cdots \frac{\mathrm{d}\zeta(c_{m'},\gamma_{m'})}{\zeta(D)}$$

is strictly positive. Thus $\{\phi \in \Omega_x : \mathrm{card}(\phi) \leq N\}$ is a small set.

5. Markov properties of shot noise Cox processes

The investigation of Markov properties of SNCPs provides deeper understanding and is relevant to the considerations of edge effects.

Throughout Sections 5.1– 5.3 we assume that X is a SNCP with uniformly bounded clusters, i.e. the kernel $k(c, \cdot)$ has support contained in $b(c, r)$ where $r > 0$ is a parameter which does not depend on $c \in \mathbb{R}^d$. For example, the uniform kernel (I) (but not the Gaussian kernel (II)) in Example 1 clearly satisfies this condition. For $B \subseteq \mathbb{R}^d$, we set

$$B_{\oplus r} = \bigcup_{\eta \in B} b(\eta, r) = \{\eta \in \mathbb{R}^d : \mathrm{dist}(\eta, B) \leq r\},$$

and call $A \subseteq B$ a maximal connected component of B if $A_{\oplus r}$ is a connected set, but $(A \cup \eta)_{\oplus r}$ is not for any $\eta \in B \setminus A$. Moreover, we let $C(B)$ denote the set of maximal connected components of B.

It is intuitively clear that the connected components in $C(X)$ are conditional independent in some sense, since offspring in different components have different parents, and the parents generating nonempty clusters is a Poisson process. This intuition is to some extend made precise in Sections 5.1 and 5.3. The connection to usual Gibbs or Markov point processes [16, 40, 44, 45, 47] and to nearest-neighbour Markov point processes [2] is discussed in Sections 5.1 and 5.2.

5.1. Local Markov properties

This section extends the results in [5, 8] concerning the nearest-neighbour Markov property (in the sense of [2]) for finite Neyman-Scott and shot noise G Cox processes to the general case of the restriction X_W of the SNCP X to a bounded Borel set $W \subset \mathbb{R}^d$.

We need first to introduce some notation. By (20) X_W is almost surely finite with density

$$f_W(x) = \mathbb{E}\Big[\exp\left(|W| - \int_W Z(\xi)\mathrm{d}\xi\right) \prod_{\xi \in x} Z(\xi)\Big] \quad \text{for finite } x \subset W$$

with respect to the unit rate Poisson process Π_W on W (indeed this is true for any Cox process X with random intensity function Z). Clearly f_W is hereditary, i.e. for finite point configurations $x \subset W$ and points $\xi \in W \setminus x$, we have that $f_W(x) > 0$ if $f_W(x \cup \xi) > 0$. The Papangelou conditional intensity [20] is defined by

$$\lambda_W(x, \xi) = f_W(x \cup \xi)/f_W(x)$$

taking $0/0 = 0$.

In general X_W is not a Markov process in the usual Ripley-Kelly [44] sense, since $\lambda_W(x, \xi)$ will depend on points $\eta \in x \setminus b(\xi, R)$ for any R less than half the diameter of W, cf. [5]. However, as shown after Proposition 6 below, $\lambda_W(x, \xi)$ depends only on x through y_ξ, i.e. the subset $y \subseteq x$ with $y \cup \xi \in C(x \cup \xi)$. Thus X_W is a nearest-neighbour Markov process in the sense of Baddeley and Møller [2].

Before stating Proposition 6 we need to introduce some further notation and a lemma. For nonempty finite point configurations $x \subset W$, define

$$\Psi_W(x) = \sum_i \sum_{\{x^1,\ldots,x^i\}} \prod_{j=1}^{i} \int \exp\left(-\gamma \int_W k(c,\xi)\mathrm{d}\xi \right) \prod_{\xi \in x^j} \gamma k(c,\xi)\mathrm{d}\zeta(c,\gamma) \qquad (25)$$

where the first sum is over $i = 1, \ldots, \mathrm{card}(x)$ and the second sum is over all unordered partitions of x into nonempty point configurations $x^1, \ldots, x^i$ (it will be apparent from the proof of Proposition 6 that $x^1, \ldots, x^i$ correspond to the nonempty $X_j \cap W$).

Lemma 3. *We have that*

$$e^{(i-1)|W|} \int \cdots \int h(x^1, \ldots, x^i) \mathrm{d}\Pi_W(x^1) \ldots \mathrm{d}\Pi_W(x^i) = \int \sum_{(x^1, \ldots, x^i)} h(x^1, \ldots, x^i) \mathrm{d}\Pi_W(x)$$

for any integer $i \geq 1$ and any measurable function $h(x^1, \ldots, x^i) \geq 0$, where the sum is over all ordered partitions of x into i nonempty point configurations.

Proof. This follows immediately from the expansion

$$\Pi_W(F) = \sum_{n=0}^{\infty} \frac{e^{-|W|}}{n!} \int_W \cdots \int_W \mathbf{1}[\{x_1, \ldots, x_n\} \in F] \mathrm{d}x_1 \cdots \mathrm{d}x_n$$

where F is any event for finite point configurations contained in W, and where the term for $n = 0$ is read as $\mathbf{1}[\emptyset \in F]$.

Proposition 6. *The density of X_W is given by*

$$f_W(x) = c_W \prod_{y \in C(x)} \Psi_W(y) \tag{26}$$

with normalising constant

$$c_W = \exp\left(|W| - \int p_W(c, \gamma) \mathrm{d}\zeta(c, \gamma)\right)$$

and where the product in (26) is set equal to one if $x = \emptyset$.

Proof. Since neither the proof of the related result in Theorem 1 in [5] nor the modified proof given in Section 4.2.2 in [8] immediately extend to the present more general situation, we give here another (and simpler) proof.

By (ii) in Lemma 2,

$$P(X_W = \emptyset) = P(\Phi_W = \emptyset) = \exp\left(-\int p_W(c, \gamma) \mathrm{d}\zeta(c, \gamma)\right) = e^{-|W|} c_W \tag{27}$$

whereby (26) is verified for $x = \emptyset$. Hence we need only to check that

$$P(X_W \in F)/P(X_W = \emptyset) = e^{|W|} \int_F \prod_{y \in C(x)} \Psi_W(y) \mathrm{d}\Pi_W(x) \tag{28}$$

for any event F of finite point configurations contained in W and with $\emptyset \notin F$. By (iv) in Lemma 2, conditional on Φ_W, $X_j \cap W$ associated to $(c_j, \gamma_j) \in \Phi_W$ has density

$$q_{(c_j,\gamma_j)}(x^j) = \exp\left(|W| - \gamma_j \int_W k(c_j,\xi)\mathrm{d}\xi\right)\left[\prod_{\xi\in x^j}\gamma_j k(c_j,\xi)\right]\Big/ p_W(c_j,\gamma_j)$$

with respect to Π_W. Hence, combining (27) with (ii) and (iv) in Lemma 2, the left hand side in (28) is equal to

$$\begin{aligned}
&\sum_{i=1}^{\infty}\frac{1}{i!}\int\int\cdots\int\int \mathbf{1}[x^1 \neq \emptyset,\ldots,x^i \neq \emptyset, x^1\cup\ldots\cup x^i \in F]\prod_{j=1}^{i} q_{(c_j,\gamma_j)}(x^j)\\
&\qquad\qquad \mathrm{d}\Pi_W(x^1)\mathrm{d}\zeta_W(c_1,\gamma_1)\cdots\mathrm{d}\Pi_W(x^i)\mathrm{d}\zeta_W(c_i,\gamma_i)\\
=&\sum_{i=1}^{\infty}\frac{1}{i!}\int\int\cdots\int\int \mathbf{1}[x^1 \neq \emptyset,\ldots,x^i \neq \emptyset, x^1\cup\ldots\cup x^i \in F]\\
&\qquad\qquad \prod_{j=1}^{i}\exp\left(|W| - \gamma_j\int_W k(c_j,\xi)\mathrm{d}\xi\right)\left[\prod_{\xi\in x^j}\gamma_j k(c_j,\xi)\right]\\
&\qquad\qquad \mathrm{d}\Pi_W(x^1)\mathrm{d}\zeta(c_1,\gamma_1)\cdots\mathrm{d}\Pi_W(x^i)\mathrm{d}\zeta(c_i,\gamma_i).
\end{aligned}$$

It is straightforwardly verified that this is equal to the right hand side in (28), using first Lemma 3 and next the fact that $\prod_{\xi\in x^j}\gamma_j k(c_j,\xi) = 0$ whenever $(x_j)_{\oplus r}$ is not a connected set.

It follows from (26) that

$$\lambda_W(x,\xi) = \lambda_W(y_\xi,\xi) = \Psi_W(y_\xi \cup \xi)/\prod_{z\in C(y_\xi)}\Psi_W(z). \tag{29}$$

This verifies the claim above that X_W is a nearest-neighbour Markov point process. Densities of a similar form as in (26) but for various kind of lattice and point process models have been considered in several papers [5, 13, 32, 33, 36]. Using a terminology similar to that in [36], we may call a finite point process with a density of the product form (26) a *Markov connected component point process.*

5.2. Integral and differential characterisations

There are many equivalent ways of defining or characterising a Gibbs or Markov point process $\tilde{X}$ on $\mathbb{R}^d$, cf. [16, 40, 44, 45]. Below we first briefly present the integral and differential equations in [16, 40], which are most convenient for our purpose. We next show how these equations can be extended to SNCPs provided the random set $X_{\oplus r}$ has bounded connected components.

Suppose that E is a so-called energy function, i.e. a real measurable function defined for all finite $x \subset \mathbb{R}^d$ (we have for simplicity excluded the case where $E(x) = \infty$; this case is of interest when dealing with so-called hard core processes, but it is not of relevance for the present paper). Furthermore, assume E has finite range of interaction $R \geq 0$, i.e. for all finite $x \subset \mathbb{R}^d$ and points $\xi \in \mathbb{R}^d \setminus x$, $E(x) - E(x \cup \xi)$ depends only on x through $x \cap b(\xi, R)$. Then for any locally finite point configuration $x \subset \mathbb{R}^d$ and any point $\xi \in \mathbb{R}^d \setminus x$, define

$$\tilde{\lambda}(x, \xi) = \exp[E(x \cap b(\xi, R)) - E((x \cap b(\xi, R)) \cup \xi)].$$

Finally, assume that E is stable in the sense of Ruelle [45], i.e. there exists a constant $K \geq 0$ so that $E(x) \geq -K\mathrm{card}(x)$ for all finite $x \subset \mathbb{R}^d$. Then $\tilde{X}$ is a Gibbs point process with energy function E if and only if

$$\mathbb{E} \sum_{\xi \in \tilde{X}} h(\tilde{X} \setminus \xi, \xi) = \int \mathbb{E}[\tilde{\lambda}(\tilde{X}, \xi) h(\tilde{X}, \xi)] \mathrm{d}\xi \tag{30}$$

for nonnegative measurable functions h. From the integral equation (30) follows that $\tilde{\rho}(\xi) = \mathbb{E}\tilde{\lambda}(\tilde{X}, \xi)$ is an intensity function for $\tilde{X}$, and for Lebesgue almost all ξ with $\tilde{\rho}(\xi) > 0$, the reduced Palm distribution $\tilde{P}_\xi^!$ for $\tilde{X}$ is absolutely continuous with respect to the distribution $\tilde{P}$ for $\tilde{X}$, and we have the differential equation

$$\frac{\mathrm{d}\tilde{P}_\xi^!}{\mathrm{d}\tilde{P}}(x) = \tilde{\lambda}(x, \xi), \tag{31}$$

cf. (12).

Remarkably, although the SNCP X is in general not a Gibbs process in the sense above, it satisfies both (30) and (31) with $(\tilde{X}, \tilde{\lambda})$ replaced by (X, λ) where

$$\lambda(X, \xi) = \mathbb{E}[Z(\xi)|X].$$

Indeed this is true for any Cox process specified by a random intensity function Z, since by (7) and basic properties for conditional mean values,

$$\mathbb{E} \sum_{\xi \in X} h(X \setminus \xi, \xi) = \mathbb{E}\mathbb{E}\left[\sum_{\xi \in X} h(X \setminus \xi, \xi) \middle| Z \right] = \mathbb{E} \int \mathbb{E}(h(X, \xi)|Z) Z(\xi) \mathrm{d}\xi$$
$$= \int \mathbb{E}[h(X, \xi) Z(\xi)] \mathrm{d}\xi = \int \mathbb{E}\{\mathbb{E}[Z(\xi)|X] h(X, \xi)\} \mathrm{d}\xi.$$

In addition, for a SNCP X with bounded connected components, we have the following extension of (29) for bounded regions W to the "infinite volume" $\mathbb{R}^d$.

Proposition 7. *Suppose that with probability one any $y \in C(X)$ is finite, and let Y_ξ denote the $y \subseteq X$ with $y \cup \xi \in C(X \cup \xi)$. Then for Lebesgue almost all ξ with $\rho(\xi) > 0$, with probability one,*

$$\lambda(X,\xi) = \lim_{s\to\infty} \lambda_{b(\xi,s)}(X \cap b(\xi,s),\xi) = \Psi(Y_\xi \cup \xi) / \prod_{z \in C(Y_\xi)} \Psi(z) \tag{33}$$

where $\Psi = \Psi_{\mathbb{R}^d}$ is given by (25) with W replaced by $\mathbb{R}^d$.

Proof. With probability one, the Poisson process $X|\Phi$ has no atoms, so $\xi \notin X$. Since X is locally finite, only finite many $y \in C(X)$ has $\mathrm{dist}(y,\xi) \leq 2r$, and Y_ξ is the union of such y, so Y_ξ is finite almost surely. Consequently, with probability one, $Y_\xi \subset b(\xi, s-r)$ for all sufficiently large s. As $\Psi_{b(\xi,s)}(x) = \Psi(x)$ for finite $x \subset b(\xi, s-r)$, we obtain that the second identity in (33) is satisfied almost surely, cf. (29).

By (30) with $(\tilde{X}, \tilde{\lambda})$ replaced by (X, λ), repeating the arguments at the beginning of the proof of Proposition 2, the first identity in (33) follows if for any bounded Borel sets $A, B \subset \mathbb{R}^d$,

$$\mathbb{E} \sum_{\xi \in X} \mathbf{1}[\xi \in A, (X \setminus \xi) \cap B = \emptyset] = \int_A \mathbb{E}\left\{ \mathbf{1}[X \cap B] \Psi(Y_\xi \cup \xi) / \prod_{z \in C(Y_\xi)} \Psi(z) \right\} \mathrm{d}\xi. \tag{34}$$

By (7)

$$\mathbb{E} \sum_{\xi \in X} \mathbf{1}[\xi \in A, (X \setminus \xi) \cap B = \emptyset] = \mathbb{E}\mathbb{E}\left[\sum_{\xi \in X} \mathbf{1}[\xi \in A, (X \setminus \xi) \cap B = \emptyset] \middle| \Phi \right]$$
$$= \mathbb{E} \int_A Z(\xi) \mathbf{1}[X \cap B = \emptyset] \mathrm{d}\xi = \int_A \mathbb{E}(Z(\xi) \mathbf{1}[X \cap B = \emptyset]) \mathrm{d}\xi$$

so let us consider the latter mean value for an arbitrary fixed $\xi \in A$ and write $X_s = X_{b(\xi,s)}$, $f_s = f_{b(\xi,s)}$, and $\Pi_s = \Pi_{b(\xi,s)}$. Then

$$\begin{aligned} &\mathbb{E}(Z(\xi) \mathbf{1}[X \cap B = \emptyset]) \\ &= \lim_{s\to\infty} \mathbb{E}(Z(\xi) \mathbf{1}[X_s \cap B = \emptyset, Y_\xi \subset b(\xi, s-r)]) \\ &= \lim_{s\to\infty} \mathbb{E} \int Z(\xi) \mathbf{1}[x \cap B = \emptyset, y_\xi \subset b(\xi, s-r)] \\ &\qquad \exp\left(|b(\xi,s)| - \int_{b(\xi,s)} Z(\eta) \mathrm{d}\eta \right) \prod_{\eta \in x} Z(\eta) \mathrm{d}\Pi_s(x) \end{aligned} \tag{35}$$

using the monotone convergence theorem and the facts that B is bounded and Y_ξ is finite almost surely to obtain the first equality, and the definition of X as a Cox process for the next equality, where y_ξ denotes the $y \subseteq x$ with $y \cup \xi \in C(x \cup \xi)$. By Fubini's theorem we can interchange the order of the expectation and integration in (35), and

$$\mathbb{E}\left[Z(\xi)\exp\left(|b(\xi,s)| - \int_{b(\xi,s)} Z(\eta)\mathrm{d}\eta\right)\prod_{\eta\in x} Z(\eta)\right] = f_s(x\cup\xi) = \lambda_s(x,\xi)f_s(x)$$

where $y_\xi \subset b(\xi, s-r)$ implies that

$$\lambda_s(x,\xi) = \Psi(y_\xi \cup \xi)/\prod_{z\in C(y_\xi)}\Psi(z).$$

If $(Y_s)_\xi$ denotes the $y \subseteq X_s$ with $y\cup\xi \in C(X_s\cup\xi)$, then $(Y_s)_\xi = Y_\xi$ for $(Y_s)_\xi \subset b(\xi, s-r)$, and so

$$\begin{aligned}
&\mathbb{E}(Z(\xi)\mathbf{1}[X\cap B=\emptyset])\\
&= \lim_{s\to\infty}\int 1[x\cap B=\emptyset, y_\xi\subset b(\xi,s-r)]\lambda_s(x,\xi)f_s(x)\mathrm{d}\Pi_s(x)\\
&= \lim_{s\to\infty}\mathbb{E}\left\{\mathbf{1}[X_s\cap B=\emptyset,(Y_s)_\xi\subset b(\xi,s-r)]\Psi(Y_\xi\cup\xi)/\prod_{z\in C(Y_\xi)}\Psi(z)\right\}\\
&=\mathbb{E}\left\{\mathbf{1}[X\cap B]\Psi(Y_\xi\cup\xi)/\prod_{z\in C(Y_\xi)}\Psi(z)\right\}
\end{aligned}$$

using again the monotone convergence theorem. Thereby (34) is verified.

The condition that the connected components in $C(X)$ are almost surely finite can be rephrased as the condition that the balls $b(\xi, r)$, $\xi \in X$, do not percolate. This is clearly the case if the balls $b(c_j, 2r)$, $c_j \in \mathcal{C}$, do not percolate. Some results are known when $\mathcal{C}$ is a stationary Poisson process: for $d = 1$ there is no percolation (Theorem 3.1 in [30]), but for $d \geq 2$ only some rather wide bounds are known for the "critical density" of $\mathcal{C}$ (see e.g. Theorem 3.10 in [30]). We refrain from a further discussion of percolation in the present paper.

5.3. Global Markov properties

A global Markov property for nearest-neighbour Markov point processes can be established for X_W when $W \subset \mathbb{R}^d$ is a bounded Borel set, cf. [21, 33]. Corollary 2 below establishes a related result for the "infinite volume", i.e. a global Markov property for both X_W and X_{W^c}, where $W^c = \mathbb{R}^d \setminus W$.

We start by establishing a slightly more general result, where $\mathcal{D}(\cdot)$ read as "distribution of $\cdot$" and $\mathcal{D}(\cdot|\cdots)$ as "conditional distribution of $\cdot$ given $\cdots$". Since we deal with the "infinite volume", the proof is different from that in [21].

Proposition 8. *Let $B \subset \mathbb{R}^d$ be a given Borel set, let $C = C_B(X)$ denote the union of those $y \in C(X)$ with $\mathrm{dist}(y, B) \leq r$, and assume that C is finite almost surely. Then*

$$\mathcal{D}(X \setminus C|C) = \mathcal{D}(Y \setminus (B_{\oplus r} \cup C_{\oplus 2r})) \tag{36}$$

where Y is independent of X and $\mathcal{D}(Y) = \mathcal{D}(X)$.

Proof. The proof is rather trivial if $C = \emptyset$, so let us assume that $C \neq \emptyset$. Below we recursively define $X^{(1)}$ as the union of clusters in X with offspring in $U^{(0)} = B_{\oplus r}$, $X^{(2)}$ as the union of clusters in X with offspring in $U^{(1)} = X^{(1)}_{\oplus 2r}$ but no offspring in $U^{(0)}$, $X^{(3)}$ as the union of clusters in X with offspring in $U^{(2)} = X^{(2)}_{\oplus 2r}$ but no offspring in $U^{(0)} \cup U^{(1)}$, etc. Clearly, with probability one, since C is finite and nonempty, $C = \bigcup_{i=1}^{I} X^{(i)}$ where $1 \leq I < \infty$, $X^{(i)} \neq \emptyset$ for $i = 1, \ldots, I$, and $X^{(I+1)} = \emptyset$. Furthermore, we let $X^{(0)} = X \setminus (B_{\oplus r} \cup C_{\oplus 2r})$, and let $\Phi_i \subseteq \Phi$ be the process of centres and intensities corresponding to the nonempty clusters for $X^{(i)}$, $i = 0, 1, 2, \ldots, I$. Thereby $X = \bigcup_{i=0}^{I} X^{(i)}$ and $\Phi = \bigcup_{i=0}^{I} \Phi_i$ where in each union the $I + 1$ sets are disjoint point processes.

Case $i = 1$: Using a notation as in Lemma 2, if $p_1 = p_{U^{(0)}}$, $\zeta_1 = \zeta_{U^{(0)}} = p_1\zeta$ (meaning that ζ_1 is the measure defined by $\zeta_1(D) = \int_D p_1(c,\gamma)\mathrm{d}\zeta(c,\gamma)$ for Borel sets $D \subseteq \mathbb{R}^d \times (0, \infty)$), and $\Phi_1 = \Phi_{U^{(0)}}$, then $X^{(1)}|\Phi_1$ is the union of independent clusters with a distribution as in (iii) in Lemma 2 (with $B = U^{(0)}$). Furthermore, set $U^{(1)} = X^{(1)}_{\oplus 2r}$, $\zeta_1' = (1 - p_1)\zeta$, $\Phi_1' = \Phi \setminus \Phi_1$, and $X_1' = X \setminus X^{(1)}$. Then by (ii) in Lemma 2, $\Phi_1'|(\Phi_1, X^{(1)})$ is a Poisson process with intensity measure ζ_1', and by (v) in Lemma 2, $X_1'|(\Phi_1, X^{(1)}, \Phi_1')$ is the union of independent clusters where the cluster associated to $(c_j, \gamma_j) \in \Phi_1'$ is a Poisson process with intensity function $\xi \to \gamma_j k(c_j, \xi)\mathbf{1}[\xi \notin U^{(0)}]$.

Case $i \geq 2$: We consider only this case as long as $X^{(i-1)} \neq \emptyset$. Conditional on $T_i \equiv (\Phi_1, X^{(1)}, \ldots, \Phi_{i-1}, X^{(i-1)})$, set

$$V_i = U^{(i-1)} \setminus \bigcup_{j=0}^{i-2} U^{(j)},$$

and use (i)–(iii) in Lemma 2 with $p_i = p_{V_i}$, $\zeta_i = p_i\zeta_{i-1}'$, and $\Phi_i = \Phi_{V_i}$ to conclude that $X^{(i)}|(T_i,\Phi_i)$ is the union of independent clusters with a distribution as in (iii) in Lemma 2 (with $B=V_i$). Set $U^{(i)}=X^{(i)}_{\oplus 2r}$, $\zeta_i'=(1-p_i)\zeta_{i-1}$, $\Phi_i'=\Phi\setminus\bigcup_{j=1}^i\Phi_j$, and $X_i'=X\setminus\bigcup_{j=1}^i X^{(j)}$. By (ii) and (v) in Lemma 2, $\Phi_i'|(T_i,\Phi_i,X^{(i)})$ is a Poisson process with intensity measure ζ_i', and $X_i'|(T_i,\Phi_i,X^{(i)},\Phi_i')$ is the union of independent clusters where the cluster associated to $(c_j,\gamma_j)\in\Phi_i'$ is a Poisson process with intensity function $\xi\to\gamma_j k(c_j,\xi)\mathbf{1}[\xi\notin V_i]$. Note that $\zeta_i'=\big[\prod_{j=1}^i(1-p_j)\big]\zeta$, where

$$\prod_{j=1}^{i}(1-p_j(c,\gamma))=\exp\bigg(-\gamma\int_{\bigcup_{j=0}^{i-1}U^{(j)}}k(c,\xi)\mathrm{d}\xi\bigg)$$

as the sets $U^{(0)},V_2,\ldots,V_i$ are disjoint and $U^{(0)}\cup V_2\cup\ldots\cup V_i=\bigcup_{j=0}^{i-1}U^{(j)}$. Note also that the first time we obtain that $X^{(i)}=\emptyset$, then $I=i-1$ and

$$V_0\equiv\bigcup_{j=0}^{I}U^{(j)}=B_{\oplus r}\cup X^{(1)}_{\oplus 2r}\cup\ldots\cup X^{(I)}_{\oplus 2r}=B_{\oplus r}\cup C_{\oplus 2r}.$$

Case $i=0$: It follows now that $\Phi_0|(\Phi_1,X^{(1)},\ldots,\Phi_I,X^{(I)})$ is a Poisson process with intensity measure $\zeta_0=p_0\zeta$, where

$$p_0(c,\gamma)\equiv\prod_{j=1}^{I}(1-p_j(c,\gamma))=\exp\bigg(-\gamma\int_{V_0}k(c,\xi)\mathrm{d}\xi\bigg),$$

and $X^{(0)}|(\Phi_0,\Phi_1,X^{(1)},\ldots,\Phi_I,X^{(I)})$ is the union of independent clusters where the cluster associated to $(c_j,\gamma_j)\in\Phi_0$ is a Poisson process with intensity function $\xi\to\gamma_j k(c_j,\xi)\mathbf{1}[\xi\notin V_0]$. By definition of C, $X\setminus C=X\setminus V_0=X^{(0)}$, so

$$\mathcal{D}(X\setminus C|\Phi_1,X^{(1)},\ldots,\Phi_I,X^{(I)})=\mathcal{D}(X^{(0)}|\Phi_0,V_0)=\mathcal{D}(Y\setminus V_0),\qquad(37)$$

whereby (36) follows.

The condition in Proposition 8 that C is almost surely finite is satisfied if e.g. B is bounded and X has almost surely no infinite connected components. The proof above depend on having a finite C or equivalently a finite I (this was used in (37) to obtain a well-defined conditional distribution). Intuitively, due to the strong independence properties in a cluster process, one may propose that Proposition 8 remains true without this restriction. Possibly this can be established by a limit argument, but we do not attempt to verify this here.

Corollary 2. *Let ∂W denote the boundary of a given bounded Borel set $W \subset \mathbb{R}^d$, and assume that $C = C_{\partial W}(X)$ is finite almost surely. Then X_W and X_{W^c} are conditionally independent given C, and*

$$\mathcal{D}(X_W \setminus C|C) = \mathcal{D}(Y_W \setminus (C_{\oplus 2r} \cup \partial W_{\oplus r})) \tag{38}$$

and

$$\mathcal{D}(X_{W^c} \setminus C|C) = \mathcal{D}(Y_{W^c} \setminus (C_{\oplus 2r} \cup \partial W_{\oplus r})). \tag{39}$$

Proof. We have that X_A and X_B are independent for Borel sets $A, B \subset \mathbb{R}^d$ with $\mathrm{dist}(A, B) > 2r$, as $\mathcal{D}(X_A|Z) = \mathcal{D}(X_A|Z_{A_{\oplus r}})$ and $\mathcal{D}(X_B|Z) = \mathcal{D}(X_B|Z_{B_{\oplus r}})$ where $Z_{A_{\oplus r}} = \{Z(\xi) : \xi \in A_{\oplus r}\}$ and $Z_{B_{\oplus r}} = \{Z(\xi) : \xi \in B_{\oplus r}\}$ are independent. Thus $Y_W \setminus (\partial W)_{\oplus r}$ and $Y_{W^c} \setminus (\partial W)_{\oplus r}$ are independent since $\mathrm{dist}(W \setminus (\partial W)_{\oplus r}, W^c \setminus (\partial W)_{\oplus r}) > 2r$. Hence, since C and Y are independent, conditional on C, we have that $Y_W \setminus (C \cup \partial W)_{\oplus r}$ and $Y_{W^c} \setminus (C \cup \partial W)_{\oplus r}$ are independent. The results follow then immediately from Proposition 8 with $B = \partial W$.

By (38)–(39) the conditional distribution of $X_W \setminus C$ given C depends only on C through $C \cap W$, and the conditional distribution of $X_{W^c} \setminus C$ given C depends only on C through $C \setminus W$. We call C the *splitting set*, $C \cap W$ the *inner splitting set*, and $C \setminus W$ the *outer splitting set* for W. It seems plausible to claim that these splitting sets are minimal, i.e. that for example $C \cap W$ is the smallest random closed set $A \subseteq W$ with the property that $\mathcal{D}(X_W|A, X_A, X_{W^c}) = \mathcal{D}(X_W|A, X_A)$ (note that $C = X_C$). We shall not attempt to verify this here.

Acknowledgements

This research was supported by the Centre for Mathematical Physics and Stochastics (MaPhySto), funded by a grant from the Danish National Research Foundation, and by the Danish Natural Science Research Council. Helpful discussions with Hans-Otto Georgii, Bjarne Højgaard, Joseph Mecke, Rasmus P. Waagepetersen, and Sergei Zuyev are acknowledged.

References

[1] Ambartzumian, R. V. (1966). On an equation for stationary point processes. *Dokl. Akad. Nauk Armjanskoi SSR* **42,** 141–147.

[2] Baddeley, A. and Møller, J. (1989). Nearest-neighbour Markov point processes and random sets. *International Statistical Review* **2,** 89–121.

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

[4] Baddeley, A. J. and van Lieshout, M. N. M. (1993). Stochastic geometry models in high-level vision. In *Statistics and Images, Advances in Applied Statistics, a supplement to the Journal of Applied Statistics*. ed. K. V. Mardia and G. K. Kanji. vol. 20. Carfax Publishing, Abingdon ch. 11, pp. 235–256.

[5] Baddeley, A. J., van Lieshout, M. N. M. and Møller, J. (1996). Markov properties of cluster processes. *Advances in Applied Probability (SGSA)* **28,** 346–355.

[6] Benes, V., Bodlak, K., Møller, J. and Waagepetersen, R. P. (2002). Bayesian analysis of log Gaussian Cox process models for disease mapping. *Technical report* R-02-2001. Department of Mathematical Sciences, Aalborg University.

[7] Best, N. G., Ickstadt, K. and Wolpert, R. (2000). Spatial Poisson regression for health and exposure data measured at disparate resolutions. *Journal of the American Statistical Association* **95,** 1076–1088.

[8] Brix, A. (1999). Generalized gamma measures and shot-noise Cox processes. *Advances in Applied Probability* **31,** 929–953.

[9] Brix, A. and Chadoeuf, J. Spatio-temporal modeling of weeds and shot-noise G Cox processes 2000. Submitted.

[10] Brix, A. and Diggle, P. J. (2001). Spatio-temporal prediction for log-Gaussian Cox processes. *Journal of the Royal Statistical Society Series B* **63,** 823–841.

[11] BRIX, A. AND KENDALL, W. S. (2001). Simulation of cluster point processes without edge effects. Research Report 382, Department of Statistics, University of Warwick.

[12] BRIX, A. AND MØLLER, J. (2001). Space-time multitype log Gaussian Cox processes with a view to modelling weed data. *Scandinavian Journal of Statistics* **28,** 471–488.

[13] CHIN, Y. C. AND BADDELEY, A. (2000). Markov interacting component processes. *Advances in Applied Probability* **32,** 597–619.

[14] DALEY, D. J. AND VERE-JONES, D. (1988). *An Introduction to the Theory of Point Processes.* Springer-Verlag, New York.

[15] DIGGLE, P. J. (1983). *Statistical Analysis of Spatial Point Patterns.* Academic Press, London.

[16] GEORGII, H.-O. (1976). Canonical and grand canonical Gibbs states for continuum systems. *Communications of Mathematical Physics* **48,** 31–51.

[17] GEYER, C. J. (1999). Likelihood inference for spatial point processes. In *Stochastic Geometry: Likelihood and Computation.* ed. O. E. Barndorff-Nielsen, W. S. Kendall, and M. N. M. van Lieshout. Chapman and Hall/CRC, London, Boca Raton. pp. 79–140.

[18] GEYER, C. J. AND MØLLER, J. (1994). Simulation procedures and likelihood inference for spatial point processes. *Scandinavian Journal of Statistics* **21,** 359–373.

[19] HEIKKINEN, J. AND ARJAS, E. (1998). Non-parametric Bayesian estimation of a spatial Poisson intensity. *Scandinavian Journal of Statistics* **25,** 435–450.

[20] KALLENBERG, O. (1984). An informal guide to the theory of conditioning in point processes. *International Statistical Review* **52,** 151–164.

[21] KENDALL, W. S. (1990). A spatial Markov property for nearest-neighbour Markov point processes. *Journal of Applied Probability* **28,** 767–778.

[22] Kerstan, J. and Matthes, K. (1964). Verallgemeinerung eines Satzes von Sliwnjak. *Rev. Roumaine Math. Pures Appl.* **IX,** 811–829.

[23] Lawson, A. B. (1993). Discussion contribution. *Journal of the Royal Statistical Society* **B 55,** 61–62.

[24] Lieshout, M. N. M. van and Baddeley, A. J. (1995). Markov chain Monte Carlo methods for clustering of image features. In *Proceedings of the 5th IEE International Conference on Image Processing and its Applications.* vol. 410 of *IEE Conference Publication.* IEE Press, London. pp. 241–245.

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

[26] Lieshout, M. N. M. van and Baddeley, A. J. (2001). Extrapolating and interpolating spatial patterns. *Technical report* PNA-RO117. Centrum voor Wiskunde en Informatica. To appear in Spatial Cluster Modelling, eds. Andrew B. Lawson and David Denison, Chapman and Hall.

[27] Matérn, B. (1960). Spatial Variation. Meddelanden fran Statens Skogforskningsinstitut, Band 49, No. 5.

[28] Matérn, B. (1986). *Spatial Variation.* Lecture Notes in Statistics. Springer-Verlag, Berlin.

[29] Mecke, J. (1967). Stationäre zufällige Maße auf lokalkompakten Abelschen Gruppen. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* **9,** 36–58.

[30] Meester, R. and Roy, R. (1996). *Continuum Percolation.* Cambridge University Press, New York.

[31] Meyn, S. P. and Tweedie, R. L. (1993). *Markov Chains and Stochastic Stability.* Springer-Verlag, London.

[32] MØLLER, J. (1994). Contribution to the discussion of N.L. Hjort and H. Omre (1994): Topics in spatial statistics. *Scandinavian Journal of Statistics* **21,** 346–349.

[33] MØLLER, J. (1999). Markov chain Monte Carlo and spatial point processes. In *Stochastic Geometry: Likelihood and Computations.* ed. O. E. Barndorff-Nielsen, W. S. Kendall, and M. N. M. van Lieshout. No. 80 in Monographs on Statistics and Applied Probability. Chapman and Hall/CRC, Boca Raton. pp. 141–172.

[34] MØLLER, J. (2001). A comparison of spatial point process models in epidemiological applications. In *Highly Structured Stochastic Systems.* ed. P. J. Green, N. L. Hjort, and S. Richardson. Oxford University Press, Oxford. To appear.

[35] MØLLER, J., SYVERSVEEN, A. R. AND WAAGEPETERSEN, R. P. (1998). Log Gaussian Cox processes. *Scandinavian Journal of Statistics* **25,** 451–482.

[36] MØLLER, J. AND WAAGEPETERSEN, R. P. (1998). Markov connected component fields. *Advances in Applied Probability (SGSA)* **30,** 1–35.

[37] MØLLER, J. AND WAAGEPETERSEN, R. P. (2001). Simulation based inference for spatial point processes. In *Spatial Statistics and Computational Methods.* ed. M. B. Hansen and J. Møller. Lecture Notes in Statistics. Springer-Verlag. To appear.

[38] MØLLER, J. AND WAAGEPETERSEN, R. P. (2001). Statistical inference for Cox processes. *Technical report* R-01-2024. Department of Mathematical Sciences, Aalborg University. To appear in Spatial Cluster Modelling, eds. Andrew B. Lawson and David Denison, Chapman and Hall.

[39] NEYMAN, J. AND SCOTT, E. L. (1958). Statistical approach to problems of cosmology. *Journal of the Royal Statistical Society Series B* **20,** 1–43.

[40] NGUYEN, X. X. AND ZESSIN, H. (1979). Integral and differential characterizations of Gibbs processes. *Mathematische Nachrichten* **88,** 105–115.

[41] RIPLEY, B. D. (1976). The second-order analysis of stationary point processes. *Journal of Applied Probability* **13,** 255–266.

[42] Ripley, B. D. (1977). Modelling spatial patterns (with discussion). *Journal of the Royal Statistical Society Series B* **39,** 172–212.

[43] Ripley, B. D. (1987). *Stochastic Simulation.* Wiley, New York.

[44] Ripley, B. D. and Kelly, F. P. (1977). Markov point processes. *Journal of the London Mathematical Society* **15,** 188–192.

[45] Ruelle, D. (1969). *Statistical Mechanics: Rigorous Results.* W.A. Benjamin, Reading, Massachusetts.

[46] Santaló, L. (1976). *Integral Geometry and Geometric Probability.* Addison–Wesley, Reading, MA.

[47] Stoyan, D., Kendall, W. S. and Mecke, J. (1995). *Stochastic Geometry and Its Applications* second ed. Wiley, Chichester.

[48] Stoyan, D. and Stoyan, H. (1994). *Fractals, Random Shapes and Point Fields.* Wiley, Chichester.

[49] Thomas, M. (1949). A generalization of Poisson's binomial limit for use in ecology. *Biometrika* **36,** 18–25.

[50] Wolpert, R. L. and Ickstadt, K. (1998). Poisson/gamma random field models for spatial statistics. *Biometrika* **85,** 251–267.