

## GENERALISED SHOT NOISE COX PROCESSES

JESPER MØLLER,\* *Aalborg University*

GIOVANNI LUCA TORRISI,\*\* *CNR-Istituto per le Applicazioni del Calcolo ‘M. Picone’*

### Abstract

We introduce a class of Cox cluster processes called generalised shot noise Cox processes (GSNCPs), which extends the definition of shot noise Cox processes (SNCPs) in two directions: the point process that drives the shot noise is not necessarily Poisson, and the kernel of the shot noise can be random. Thereby, a very large class of models for aggregated or clustered point patterns is obtained. Due to the structure of GSNCPs, a number of useful results can be established. We focus first on deriving summary statistics for GSNCPs and, second, on how to simulate such processes. In particular, results on first- and second-order moment measures, reduced Palm distributions, the  $J$ -function, simulation with or without edge effects, and conditional simulation of the intensity function driving a GSNCP are given. Our results are exemplified in important special cases of GSNCPs, and we discuss their relation to the corresponding results for SNCPs.

*Keywords:* Cluster process; conditional simulation; Cox process; edge effect; generalised shot noise G Cox process; geometric ergodicity; Markov point process; Metropolis–Hastings algorithm; Palm distribution; perfect simulation; shot noise Cox process; spatial point process; summary statistic

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

Secondary 60K35; 68U20

### 1. Introduction

One of the most important and versatile classes of point process models for clustered point patterns is the class of Cox process models – see, e.g. [9], [10], [12], [19], and [44]. Recently, simulation-based inference for certain families of Cox processes has been studied in great detail [5], [6], [8], [49], [35], [32], [33], [34], [51]. As explained in Section 2, this research has shown the necessity of extending existing classes of Cox process models, and the present paper is therefore concerned with a rich new class of such models that, to some extent, is analytically tractable and, in particular, is open to simulation-based inference.

Recall that a point process  $X$  on  $\mathbb{R}^d$  is a Cox process driven by a random field  $Z(\xi) \geq 0$ ,  $\xi \in \mathbb{R}^d$ , if  $X \mid Z$  is a Poisson process on  $\mathbb{R}^d$  with intensity function  $Z$ . Throughout this paper,  $X$  denotes a Cox process driven by a random field of the form

$$Z(\xi) = \sum_j \gamma_j k_{b_j}(\mathbf{c}_j, \xi), \quad (1)$$

Received 15 March 2004; revision received 4 January 2005.

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

\*\* Postal address: CNR-Istituto per le Applicazioni del Calcolo ‘M. Picone’, Viale del Policlinico 137, I-00161 Rome, Italy. Email address: torrisci@iac.rm.cnr.it

where the  $(\mathbf{c}_j, b_j, \gamma_j)$  are the points of a point process  $\Phi$  on  $\mathbb{R}^d \times (0, \infty) \times (0, \infty)$ ; we identify  $\sum_j$  and  $\sum_{(\mathbf{c}_j, b_j, \gamma_j) \in \Phi}$ ; and  $b_j > 0$  is a bandwidth for the kernel  $k_{b_j}(\mathbf{c}_j, \cdot)$ , i.e.

$$k_{b_j}(\mathbf{c}_j, \boldsymbol{\xi}) = \frac{k_1(\mathbf{c}_j/b_j, \boldsymbol{\xi}/b_j)}{b_j^d}, \tag{2}$$

where  $k_1(\mathbf{c}_j, \cdot)$  is a density with respect to the Lebesgue measure on  $\mathbb{R}^d$ .

We call  $X$  a *generalised shot noise Cox process (GSNCP)*, since a *shot noise Cox process (SNCP)* is the special case in which  $\{(\mathbf{c}_j, \gamma_j)\}$  is a Poisson process on  $\mathbb{R}^d \times (0, \infty)$  with all the  $b_j$  equal and fixed; cf. Møller [31] (extending previous work by Brix [4] and Wolpert and Ickstadt [51]). We can also view a GSNCP as a Cox cluster process, since  $X \mid \Phi$  is distributed as the superposition  $\bigcup_j X_j$  of independent Poisson processes  $X_j$  with intensity functions  $\gamma_j k_{b_j}(\mathbf{c}_j, \cdot)$ . In applications of (2), the kernel  $k_1(\mathbf{c}_j, \cdot)$  is usually concentrated around  $\mathbf{c}_j$ , so we call  $\Phi_{\text{cent}} = \{\mathbf{c}_j\}$  the centre (or mother) process and  $X_j$  the cluster (or daughter) process with centre (or mother)  $\mathbf{c}_j$ , intensity  $\gamma_j$ , and dispersion density  $k_{b_j}(\mathbf{c}_j, \cdot)$ . Many of the results in this paper can easily be modified to hold for a general stochastic kernel  $k_{b_j}(\mathbf{c}_j, \cdot)$ , where  $b_j$  is a random variable with state space given by an arbitrary metric space  $E$  but, for specificity and ease of presentation, we let  $E = (0, \infty)$  and interpret  $b_j$  as a bandwidth.

The focus in this paper is on the probabilistic aspects of GSNCPs, with a view to statistical applications. Although GSNCPs are more complicated than SNCPs, we show that despite this, due to the structure of GSNCPs, a number of useful and general results can be established. Our results are exemplified for important special cases of GSNCPs, and we discuss their relation to the corresponding results for SNCPs.

The paper is organised as follows. In Section 2, we motivate the introduction of GSNCPs and discuss some other preliminaries, including examples of GSNCPs. In Section 3, we present results for summary statistics and reduced Palm distributions of GSNCPs. Particularly, results for first- and second-order properties and the  $J$ -function are obtained. In Section 4, we deal with algorithms for the simulation of a GSNCP  $X$  within a bounded window  $W$ , and for conditional simulation of  $\Phi$ , given the restriction of  $X$  to  $W$ . We quantify the effect of ignoring edge effects in a straightforward simulation algorithm for  $X$  when, e.g.  $\Phi_{\text{cent}}$  is restricted to a bounded, extended window  $W_{\text{ext}} \supset W$ . We also describe a perfect simulation algorithm inspired by the work of Brix and Kendall [7]. Further, for a certain Metropolis–Hastings algorithm for conditional simulation of  $\Phi$  given  $X \cap W$ , we discuss convergence properties; in particular, we establish geometric ergodicity. Finally, in Section 5, we briefly discuss future research and the importance of our results for statistical inference.

## 2. Background

### 2.1. Motivation

Our extension of SNCPs to GSNCPs is motivated by different statistical applications, as follows.

(a) For Neyman–Scott processes (see Example 2 in Section 2.3) and many other Cox cluster processes used for statistical analysis (see, e.g. [12], [34], and [44]),  $\Phi_{\text{cent}}$  is assumed to be a Poisson process and the  $(b_j, \gamma_j)$  are assumed to be equal to an unknown parameter. As illustrated by van Lieshout and Baddeley [49], a repulsive Markov point process model for  $\Phi_{\text{cent}}$  may be more relevant in many situations. Such models may be described by GSNCPs, but not by SNCPs.

(b) Wolpert and Ickstadt [51] considered a fully Bayesian analysis for a Poisson–gamma model (a particular case of an SNCP – see Example 3 in Section 2.3): a prior for the parameter in a parametric model of the intensity function of the Poisson process  $\Phi$  is imposed, whereby  $X$  becomes a GSNCP.

(c) Møller and Waagepetersen [34] considered likelihood inference for parametric models of SNCPs, where the likelihood is, in general, not available in closed form and the underlying random field  $Z$  is unobserved. For different fixed values of the bandwidths, they combined Markov chain Monte Carlo (MCMC) methods for calculating maximum partial likelihood estimates and likelihood ratios, whereby an approximate maximum likelihood estimate was obtained (see Section 10.3 of [34]). This procedure is rather time consuming, and it would be easier to use a Bayesian MCMC approach, imposing a prior on the bandwidth. Then a GSNCP model (with all  $b_j$  equal but random) is obtained for the likelihood term in the posterior density.

(d) SNCPs are often claimed to allow a certain degree of flexibility [30], [32], [40], [51]. In an SNCP, for the random intensity function  $Z$  in (1), the random cluster intensity  $\gamma_j$  scales the kernel  $k_b(c_j, \cdot)$ , where  $b > 0$  is the fixed bandwidth. However, the same degree of flexibility for modelling  $Z$  may be obtained by letting the  $\gamma_j$  be equal, but using different random bandwidths  $b_j$ . Of course, even more flexibility is obtained when both the  $\gamma_j$  and the  $b_j$  are random, as in a GSNCP.

(e) Kingman [22] (also see Section 5.5 of [12]) considered a model for reproducing individuals, where the  $(n + 1)$ th generation  $G_{n+1}$ , given the  $n$ th generation  $G_n$  (and previous generations  $G_0, \dots, G_{n-1}$ ), is a Poisson process with an intensity measure of the form (1), where the cluster centres are given by  $G_n$ , each  $\gamma_j$  is a function of  $(c_j, G_n)$ , and the bandwidths are equal and fixed. Thus,  $G_{n+1}$  becomes a GSNCP, but not necessarily an SNCP. Furthermore, assuming that  $G_0$  is a Poisson process, the superposition of GSNCPs  $G_0, G_1, \dots$  can be interpreted as a spatial Hawkes process (extending the definition of [17] and [18] for  $d = 1$  to  $d \geq 1$ , and allowing a more general structure for the conditional intensity function).

We consider, in particular, the following models for  $\Phi$ . Cox process models for  $\Phi$  appear naturally when we consider a hierarchical model in which  $\Phi \mid \theta$  is a Poisson process with an intensity measure  $\zeta_\theta$  depending on a random variable  $\theta$ ; see points (b) and (d) above, and Example 3 in Section 2.3. The case in which  $\Phi_{\text{cent}}$  and the  $(b_j, \gamma_j)$  are independent is particularly tractable, and, in the authors' opinion, the most important models for  $\Phi_{\text{cent}}$  are then Markov point process models and Poisson models; see point (a) above and Example 4 in Section 2.3. Markov point process models provide flexible models of inhibition in the cluster centres, while Poisson models provide a convenient mathematical framework – see [4], [31], [32], [34], [49], and [51].

## 2.2. Assumptions and other preliminaries

In this section, we specify certain conditions that are assumed to hold throughout the text, and introduce some terminology and notation.

We assume that  $\Phi$  is a random, locally finite subset of  $\Omega = \mathbb{R}^d \times (0, \infty) \times (0, \infty)$ , i.e.  $\Phi_D \equiv \Phi \cap D$  is finite whenever  $D \subset \Omega$  is bounded. Moreover,  $Z$  is assumed to be almost surely locally integrable, i.e. with probability 1,  $\int_B Z(\xi) d\xi < \infty$  for bounded Borel sets  $B \subset \mathbb{R}^d$ . Hence, for any bounded  $B \subset \mathbb{R}^d$ ,  $X_B \equiv X \cap B$  is almost surely finite and, therefore, only finitely many clusters  $X_j$  have points in  $B$ . Note that the centre process  $\Phi_{\text{cent}}$  is countable, but not necessarily locally finite (see Example 3 in Section 2.3).

If  $\Phi_{\text{cent}}$  is stationary (i.e. its distribution is invariant under translations in  $\mathbb{R}^d$ ) and  $k_1(\mathbf{c}, \boldsymbol{\xi}) \equiv k_1(\boldsymbol{\xi} - \mathbf{c})$  is invariant under translations in  $\mathbb{R}^d$ , then  $Z$  and, hence,  $X$  are stationary. If, furthermore, the distribution of  $\Phi_{\text{cent}}$  is invariant under motions in  $\mathbb{R}^d$  and  $k_1$  is isotropic, i.e.  $k_1(\mathbf{c}, \boldsymbol{\xi})$  depends only on the distance  $\|\mathbf{c} - \boldsymbol{\xi}\|$ , then the distributions of  $Z$  and  $X$  are invariant under motions in  $\mathbb{R}^d$ . Two frequently used choices of isotropic kernels are the Gaussian kernel

$$k_1(\boldsymbol{\xi}) = (2\pi)^{-d/2} \exp(-\frac{1}{2}\|\boldsymbol{\xi}\|^2) \tag{3}$$

and the uniform kernel

$$k_1(\boldsymbol{\xi}) = \mathbf{1}(\|\boldsymbol{\xi}\| \leq 1)/\omega_d, \tag{4}$$

where  $\mathbf{1}(\cdot)$  denotes the indicator function and  $\omega_d = \pi^{d/2}/\Gamma(1 + \frac{1}{2}d)$ .

We shall often refer to the intensity measure and the second-order reduced moment measure of  $\Phi$ , which we denote by  $\zeta$  and  $\zeta^{(2)}$ , respectively. Recall that, for measurable functions  $h_1: \Omega \rightarrow [0, \infty)$  and  $h_2: \Omega \times \Omega \rightarrow [0, \infty)$ ,

$$\int h_1(\mathbf{c}, b, \gamma) d\zeta(\mathbf{c}, b, \gamma) = \mathbb{E} \sum_j h_1(\mathbf{c}_j, b_j, \gamma_j), \tag{5}$$

$$\begin{aligned} \int h_2((\mathbf{c}, b, \gamma), (\mathbf{c}', b', \gamma')) d\zeta^{(2)}((\mathbf{c}, b, \gamma), (\mathbf{c}', b', \gamma')) \\ = \mathbb{E} \sum_{j \neq j'} h_2((\mathbf{c}_j, b_j, \gamma_j), (\mathbf{c}_{j'}, b_{j'}, \gamma_{j'})) \end{aligned} \tag{6}$$

– see, e.g. [44]. In applications,  $\zeta$  is often of the form

$$\zeta(D) = \iint_D d\mathbf{c} d\chi(b, \gamma) \tag{7}$$

for Borel sets  $D \subseteq \Omega$ , where  $\chi$  is a measure on  $(0, \infty)^2$ ; cf. Examples 2 to 4 in Section 2.3.

### 2.3. Examples

The following examples describe important model classes of GSNCPs. As the measures  $\zeta$  and  $\zeta^{(2)}$  play an important role in Section 3, we specify them in the examples.

**Example 1.** A particular tractable model class is obtained when  $\Phi$  is a Poisson process with locally finite intensity measure  $\zeta$ . This class contains the SNCPPs (the special cases in which all bandwidths are equal and fixed). Note that  $\zeta^{(2)} = \zeta \times \zeta$  is just a product measure, and (7) is equivalent to stationarity of  $\Phi_{\text{cent}}$ .

**Example 2.** A Neyman–Scott process [37] is obtained when  $\Phi_{\text{cent}}$  is a stationary Poisson process on  $\mathbb{R}^d$  and the cluster intensities  $\gamma_j = \gamma$  and the bandwidths  $b_j = b$  are equal and fixed. For the Gaussian kernel (3) we have a (modified) Thomas process [46], and for the uniform kernel (4) we have a Matérn cluster process [25], [26].

Natural extensions of this model include the GSNCPs obtained if  $\Phi_{\text{cent}}$  is a stationary point process on  $\mathbb{R}^d$  with intensity  $\rho_{\text{cent}} < \infty$ ,  $\Phi_{\text{cent}}$  is independent of the  $(b_j, \gamma_j)$ , and

- (i) the  $(b_j, \gamma_j) = (b, \gamma)$  are identical, with distribution  $Q$ ;
- (ii) the  $b_j = b$  are identical, with distribution  $Q_1$ , the  $\gamma_j$  are independent and identically distributed (i.i.d.) with distribution  $Q_2$ , and  $b$  and  $\{\gamma_j\}$  are independent; or
- (iii) the  $(b_j, \gamma_j)$  are i.i.d. with distribution  $Q$ .

We call such extensions the *generalised Neyman–Scott processes*. Then (7) holds, with

$$\chi = \rho_{\text{cent}} Q \tag{8}$$

and  $Q = Q_1 \times Q_2$  in case (ii). This follows by conditioning on  $\Phi_{\text{cent}}$  on the right-hand side of (5) and using standard arguments, where we first let  $h_1(\mathbf{c}, b, \gamma)$  in (5) be an indicator function  $\mathbf{1}(\mathbf{c} \in A, b \in B, \gamma \in C)$  (where  $A \subseteq \mathbb{R}^d, B \subseteq (0, \infty)$ , and  $C \subseteq (0, \infty)$  are Borel sets). Furthermore, using similar arguments and (6), we obtain

$$\int h_2((\mathbf{c}, b, \gamma), (\mathbf{c}', b', \gamma')) d\zeta^{(2)}((\mathbf{c}, b, \gamma), (\mathbf{c}', b', \gamma')) \tag{9}$$

$$= \begin{cases} \iint h_2((\mathbf{c}, b, \gamma), (\mathbf{c}', b, \gamma)) dQ(b, \gamma) d\zeta_{\text{cent}}^{(2)}(\mathbf{c}, \mathbf{c}') & \text{in case (i),} \\ \iiint h_2((\mathbf{c}, b, \gamma), (\mathbf{c}', b, \gamma')) dQ_1(b) dQ_2(\gamma) dQ_2(\gamma') d\zeta_{\text{cent}}^{(2)}(\mathbf{c}, \mathbf{c}') & \text{in case (ii),} \\ \iiint h_2((\mathbf{c}, b, \gamma), (\mathbf{c}', b', \gamma')) dQ(b, \gamma) dQ(b', \gamma') d\zeta_{\text{cent}}^{(2)}(\mathbf{c}, \mathbf{c}') & \text{in case (iii),} \end{cases}$$

where  $\zeta_{\text{cent}}^{(2)}$  denotes the second-order reduced moment measure for  $\Phi_{\text{cent}}$ . In particular, if  $\Phi_{\text{cent}}$  is a stationary Poisson process, then  $d\zeta_{\text{cent}}^{(2)}(\mathbf{c}, \mathbf{c}') = \rho_{\text{cent}}^2 d\mathbf{c} d\mathbf{c}'$  (see, e.g. [44]).

**Example 3.** Suppose that  $\{(\mathbf{c}_j, \gamma_j)\}$  is a Poisson process on  $\mathbb{R}^d \times (0, \infty)$  with intensity function

$$\lambda_\theta(\mathbf{c}, \gamma) \equiv \lambda_\theta(\gamma) = \kappa \gamma^{-\alpha-1} \exp(-\tau \gamma) / \Gamma(1 - \alpha), \tag{10}$$

where  $\theta = (\alpha, \kappa, \tau)$  is a parameter with  $\alpha < 1, \kappa > 0$ , and  $\tau > 0$  (these restrictions are equivalent to local integrability of  $\lambda_\theta$ ; cf. [4] and [31]). When all the  $b_j$  in (1) are equal and fixed,  $X$  is called a shot noise G Cox process (SNGCP) [4], [31]; a Poisson–gamma process [51] is the special case of  $\alpha = 0$ . In the stationary case, a SNGCP has intensity equal to  $\kappa \tau^{\alpha-1}$ .

We shall later refer to the following properties of the Poisson process  $\{(\mathbf{c}_j, \gamma_j)\}$ . The point processes  $\{\mathbf{c}_j\}$  and  $\{\gamma_j\}$  are independent, and their distributions depend on  $\alpha$  as follows.

(i)  $\alpha < 0$ . In this case,  $\Phi_{\text{cent}}$  is a stationary Poisson process with intensity  $-\kappa \tau^\alpha / \alpha$ , and the  $\gamma_j$  are independent and gamma distributed with shape parameter  $-\alpha$  and inverse scale parameter  $\tau$ .

(ii)  $0 \leq \alpha < 1$ . In this case,  $\Phi_{\text{cent}}$  is not locally finite. However,  $\{(\mathbf{c}_j, \gamma_j) : \mathbf{c}_j \in A\}$  and  $\{(\mathbf{c}_j, \gamma_j) : \mathbf{c}_j \in B\}$  are independent for disjoint Borel sets  $A, B \subset \mathbb{R}^d$ . If  $A$  has finite Lebesgue measure  $|A|$ , then the points in  $\Phi_{\text{cent}} \cap A$  are independent and uniformly distributed on  $A$ , and the corresponding  $\gamma_j$  form an inhomogeneous Poisson process on  $(0, \infty)$  with intensity function  $|A| \lambda_\theta(\gamma)$ .

*Generalised shot noise G Cox processes* can naturally be obtained in different ways. For specificity, let  $\pi$  be a distribution imposed on the parameter  $\theta = (\alpha, \kappa, \tau)$ , and let  $\{(\mathbf{c}_j, \gamma_j)\} \mid \theta$  be a Poisson process with intensity function  $\lambda_\theta$ ; cf. [51]. Furthermore, assume that  $\{b_j\}$  and  $(\theta, \{(\mathbf{c}_j, \gamma_j)\})$  are independent, and that the  $b_j$  are either identical or i.i.d. with distribution  $\nu$ . Then we obtain a GSNCP for which (7) holds, with

$$\chi(A) = \iint_A \mathbb{E} \lambda_\theta(\gamma) d\nu(b) d\gamma \tag{11}$$

for Borel sets  $A \subseteq (0, \infty) \times (0, \infty)$ , where the expectation is with respect to  $\pi$  (this follows by arguments similar to those in Example 2).

Furthermore,

$$\int h_2((\mathbf{c}, b, \gamma), (\mathbf{c}', b', \gamma')) d\zeta^{(2)}((\mathbf{c}, b, \gamma), (\mathbf{c}', b', \gamma')) = \iiint \iiint h_2((\mathbf{c}, b, \gamma), (\mathbf{c}', b, \gamma')) E(\lambda_\theta(\gamma)\lambda_\theta(\gamma')) d\mathbf{c} d\mathbf{c}' d\nu(b) d\gamma d\gamma' \tag{12}$$

when the  $b_j$  are identical and, when the  $b_j$  are i.i.d.,

$$\int h_2((\mathbf{c}, b, \gamma), (\mathbf{c}', b', \gamma')) d\zeta^{(2)}((\mathbf{c}, b, \gamma), (\mathbf{c}', b', \gamma')) = \iiint \iiint \iiint h_2((\mathbf{c}, b, \gamma), (\mathbf{c}', b', \gamma')) E(\lambda_\theta(\gamma)\lambda_\theta(\gamma')) d\mathbf{c} d\mathbf{c}' d\nu(b) d\nu(b') d\gamma d\gamma'. \tag{13}$$

**Example 4.** As in van Lieshout and Baddeley [49], suppose that  $\Phi_{\text{cent}}$  is a finite Markov (or Gibbs) point process [2], [34], [41], [42], [47] defined on a bounded Borel set  $B \subset \mathbb{R}^d$ . This means that  $\Phi_{\text{cent}}$  has a density  $p$  with respect to the unit-rate Poisson process on  $B$ , so that

$$p(\mathbf{c}) = \prod_{\eta \subseteq \mathbf{c}} \varphi(\eta) \tag{14}$$

for finite subsets  $\mathbf{c} \subset B$ , where  $\varphi(\eta) \geq 0$  is a so-called interaction function such that  $\varphi(\eta) = 1$  whenever  $\eta$  contains two points further than  $R$  units apart. Here  $R < \infty$  is a parameter specifying the range of interaction. Note that if  $p(\mathbf{c}) > 0$  and we define the so-called Papangelou conditional intensity by  $\lambda^*(\mathbf{c}, \xi) := p(\mathbf{c} \cup \{\xi\})/p(\mathbf{c})$ , then, for  $\xi \notin \mathbf{c}$ ,

$$\lambda^*(\mathbf{c}, \xi) = \prod_{\eta \subseteq \mathbf{c} \cup b(\xi, R)} \varphi(\eta \cup \{\xi\}), \tag{15}$$

where  $b(\xi, R)$  denotes the ball in  $\mathbb{R}^d$  with centre  $\xi$  and radius  $R$ .

In [49] it was also assumed that  $Z(\xi) = \varepsilon + \sum_j h(\xi | c_j)$ , where  $h(\xi | c_j)$  is a nonnegative integrable function and  $\varepsilon > 0$  is a parameter. If instead  $\varepsilon = 0$ , we have a GSNCP with nonrandom  $(\gamma_j, b_j)$  but, clearly, (7) does not hold, since  $\Phi_{\text{cent}}$  is contained in  $B$ . If, moreover,  $\Phi_{\text{cent}}$  is independent of the  $(b_j, \gamma_j)$ , which are either i.i.d. or identical with distribution  $Q$ , then

$$\int h_1(\mathbf{c}, b, \gamma) d\zeta(\mathbf{c}, b, \gamma) = \iint h_1(\xi, b, \gamma) E \lambda^*(\Phi_{\text{cent}}, \xi) d\xi dQ(b, \gamma). \tag{16}$$

This follows from the fact that  $\Phi_{\text{cent}}$  has intensity function  $\rho_{\text{cent}}(\xi) = E \lambda^*(\Phi_{\text{cent}}, \xi)$  (see Proposition 6.2 of [34]), but a closed form expression for  $\rho_{\text{cent}}$  is in general unknown (except in the Poisson case, where  $\lambda^*(\Phi_{\text{cent}}, \xi)$  does not actually depend on  $\Phi_{\text{cent}}$ ). Furthermore,

$$\begin{aligned} &\int h_2((\mathbf{c}, b, \gamma), (\mathbf{c}', b', \gamma')) d\zeta^{(2)}((\mathbf{c}, b, \gamma), (\mathbf{c}', b', \gamma')) \\ &= \iiint \iiint h_2((\mathbf{c}, b, \gamma), (\mathbf{c}', b, \gamma')) \\ &\quad \times E \left( \prod_{\eta \subseteq \Phi_{\text{cent}} \cap (b(\mathbf{c}, R) \cup b(\mathbf{c}', R))} \varphi(\eta \cup \{\mathbf{c}\}) \varphi(\eta \cup \{\mathbf{c}'\}) \varphi(\eta \cup \{\mathbf{c}, \mathbf{c}'\}) \right) \\ &\quad \times d\mathbf{c} d\mathbf{c}' dQ(b, \gamma) dQ(b', \gamma') \end{aligned} \tag{17}$$

– again see, e.g. Proposition 6.2 of [34]. Also,  $\zeta^{(2)}$  is not known in closed form.

Suppose we instead consider a stationary Markov (or Gibbs) point process  $\Phi_{\text{cent}}$  defined on  $\mathbb{R}^d$ , with an interaction function  $\varphi$  of finite range of interaction  $R$ , where  $\varphi$  is invariant under translations in  $\mathbb{R}^d$  [13], [34], [38], [39], [43]. Briefly, such a point process can be specified by a Papangelou conditional intensity  $\lambda^*(\mathbf{c}, \boldsymbol{\xi})$  of the form (15), but now defined for locally finite subsets  $\mathbf{c} \subset \mathbb{R}^d$  and points  $\boldsymbol{\xi} \in \mathbb{R}^d \setminus \mathbf{c}$  (for details, see the references mentioned above). If  $\Phi_{\text{cent}}$  is independent of the  $(b_j, \gamma_j)$ , which are either i.i.d. or identical with distribution  $Q$ , then we have a GSNCP. Here (7) holds with  $\chi$  of the form (8), provided that

$$\rho_{\text{cent}} = E \lambda^*(\Phi_{\text{cent}}, \mathbf{0}) = E \prod_{\eta \subseteq \Phi_{\text{cent}} \cap b(\mathbf{0}, R)} \varphi(\eta \cup \{\mathbf{0}\}) \tag{18}$$

is finite. However, a closed form expression for  $\rho_{\text{cent}}$  is in general not known, and likewise for  $\zeta^{(2)}$ , which is still given by (17).

### 3. Summary statistics

This section deals with summary statistics such as the intensity function, the pair correlation function, and the  $J$ -function for GSNCPs.

#### 3.1. First- and second-order characteristics

Expressions for the product moments  $E(Z(\boldsymbol{\xi}_1) \cdots Z(\boldsymbol{\xi}_n))$  in terms of the moment measures for  $\Phi$  can be obtained, as in [31], by using the Slivnyak–Mecke theorem for the Poisson process  $X \mid \Phi$ . In this paper we concentrate on the two most fundamental summary statistics, namely the *intensity function*  $\rho(\boldsymbol{\xi}) = E Z(\boldsymbol{\xi})$  and the *pair correlation function*  $g(\boldsymbol{\xi}, \boldsymbol{\eta}) = E(Z(\boldsymbol{\xi})Z(\boldsymbol{\eta}))/[\rho(\boldsymbol{\xi})\rho(\boldsymbol{\eta})]$  (provided that the means exist, and taking  $0/0 = 0$ ). In the examples below we discuss to what extent closed form expressions for  $\rho(\boldsymbol{\xi})$  and  $g(\boldsymbol{\xi}, \boldsymbol{\eta})$  can be derived for GSNCPs as introduced in Examples 1 to 4 in Section 2.3.

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

$$\rho(\boldsymbol{\xi}) = \int \gamma k_b(\mathbf{c}, \boldsymbol{\xi}) d\zeta(\mathbf{c}, b, \gamma), \tag{19}$$

provided that the integral is finite for all  $\boldsymbol{\xi} \in \mathbb{R}^d$ .

*Proof.* The proof follows immediately from (5).

**Example 5.** By (19), if (7) holds and  $k_1(\mathbf{c}, \boldsymbol{\xi}) \equiv k_1(\boldsymbol{\xi} - \mathbf{c})$  is invariant under translations,  $\rho \equiv \rho(\boldsymbol{\xi})$  is given by

$$\rho = \int \gamma d\chi(b, \gamma). \tag{20}$$

This reduces as follows for the GSNCPs in Examples 1 to 3: if  $\Phi$  is a Poisson process and the bandwidths are equal and fixed, the results for  $\rho$  obtained in [31] for SNCPs apply. For a generalised Neyman–Scott process,  $\rho = \rho_{\text{cent}} E \gamma$ , where the mean is with respect to  $(b, \gamma) \sim Q$ ; cf. (8). For a generalised shot noise G Cox process,  $\rho = E(\kappa \tau^{\alpha-1})$ ; cf. (11).

However, for the Markov point process considered in Example 4, a closed form expression for  $\rho(\boldsymbol{\xi})$  is in general unknown, since  $E \lambda^*(\Phi_{\text{cent}}, \boldsymbol{\xi})$  is not known in closed form in either (16) or (18).

**Proposition 2.** *The pair correlation function exists and is given by*

$$g(\xi, \eta) = \frac{\beta_1(\xi, \eta) + \beta_2(\xi, \eta)}{\rho(\xi)\rho(\eta)},$$

provided that  $\rho(\xi) < \infty$  for all  $\xi \in \mathbb{R}^d$  and the integrals

$$\beta_1(\xi, \eta) = \int \gamma \gamma' k_b(\mathbf{c}, \xi) k_{b'}(\mathbf{c}', \eta) d\zeta^{(2)}((\mathbf{c}, b, \gamma), (\mathbf{c}', b', \gamma'))$$

and

$$\beta_2(\xi, \eta) = \int \gamma^2 k_b(\mathbf{c}, \xi) k_b(\mathbf{c}, \eta) d\zeta(\mathbf{c}, b, \gamma)$$

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

*Proof.* The proof follows along the lines of the proof of Proposition 1 of [31], but using (6).

**Example 6.** Suppose that  $k_1(\mathbf{c}, \xi) \equiv k_1(\xi - \mathbf{c})$  is invariant under translations.

If  $\Phi$  is a Poisson process so that  $\Phi_{\text{cent}}$  is stationary and  $\rho$ , given by (20), is finite, then, by Example 1 and Proposition 2,

$$g(\xi, \eta) = 1 + \beta_2(\xi, \eta)/\rho^2, \tag{21}$$

where

$$\beta_2(\xi, \eta) = \int \gamma^2 \phi_b(\xi - \eta) d\chi(b, \gamma)$$

and

$$\phi_b(\xi) = \int k_b(\mathbf{c}) k_b(\xi + \mathbf{c}) d\mathbf{c}.$$

For the Gaussian kernel (3), this reduces to

$$\phi_b(\xi) = \frac{1}{(4\pi b^2)^{d/2}} \exp\left(-\frac{\|\xi\|^2}{4b^2}\right).$$

Consider a generalised Neyman–Scott process in which  $\Phi_{\text{cent}}$  is a stationary Poisson process with intensity  $\rho_{\text{cent}} < \infty$ . Then

$$\beta_2(\xi, \eta) = \rho_{\text{cent}} E(\gamma^2 \phi_b(\xi - \eta)),$$

where the mean is with respect to  $(b, \gamma) \sim Q$ . For each of the cases (i) to (iii) in Example 2, we obtain the following from (9) and Proposition 2. In case (i),  $\beta_1(\xi, \eta) = \rho_{\text{cent}}^2 E(\gamma^2)$ , so

$$g(\xi, \eta) = \frac{E(\gamma^2) + E(\gamma^2 \phi_b(\xi - \eta))/\rho_{\text{cent}}}{(E \gamma)^2}.$$

In case (ii),  $\beta_1(\xi, \eta) = \rho^2$ , so

$$g(\xi, \eta) = 1 + \frac{E(\gamma^2 \phi_b(\xi - \eta))}{\rho_{\text{cent}} (E \gamma)^2}. \tag{22}$$

Case (iii) is just the Poisson case above with  $d\chi(b, \gamma) = \rho_{\text{cent}} dQ(b, \gamma)$ , and  $g$  is again of the form (22).

Next, consider the specific example of a generalised shot noise G Cox process in Example 3. By (12) and (13), both when the  $b_j$  are identical and when the  $b_j$  are i.i.d., it is straightforward to derive that  $g$  is of the form (21), with

$$\beta_2(\xi, \eta) = E((1 - \alpha)\kappa\tau^{\alpha-2}) E\phi_b(\xi - \eta),$$

where the means now are with respect to the independent random variables  $b$  and  $\theta = (\alpha, \kappa, \tau)$ .

Note that  $g \geq 1$  for the GSNCPs in Example 6. This is in accordance with the usual interpretation that  $g \geq 1$  indicates aggregation of the points in  $X$  [34], [44]. It seems to be an open problem to what extent  $g \geq 1$  for the Markov point process considered in Example 4. Most Markov point process models are repulsive, that is,  $\varphi(\eta) \leq 1$  whenever  $\text{card}(\eta) \geq 2$  (see [3], [14], [28], and [43] for examples of both repulsive and attractive Markov point processes). In the special case of no interaction, i.e.  $\varphi(\eta) = 1$  whenever  $\text{card}(\eta) \geq 2$ , we clearly have  $g = 1$ . So, in the repulsive case, since aggregation in  $X$  is expected to be more pronounced than is the case if there is no interaction, we may conjecture that  $g(\xi, \eta) \geq 1$ , at least when  $\xi$  and  $\eta$  are sufficiently close. However, the authors' attempt to verify this, by combining (16) to (18) and Proposition 2, failed.

**3.2. Reduced Palm distributions and  $J$ -functions**

A simple description of the reduced Palm distribution of an SNCP was established in [31]. This section extends this to GSNCPs and discusses how to use this to derive certain properties of van Lieshout and Baddeley's  $J$ -function [48].

We first need some additional notation and assumptions. Denote the state space of  $X$  by  $N_{\text{lf}}$ , the set of locally finite subsets of  $\mathbb{R}^d$ . Let  $N_{\text{lf}}$  be equipped with the  $\sigma$ -field  $\mathcal{N}_{\text{lf}}$  generated by the sets  $F_{B,n} = \{x \in N_{\text{lf}} : \text{card}(x \cap B) = n\}$  for  $n = 0, 1, \dots$  and bounded Borel sets  $B \subset \mathbb{R}^d$ . Let the product space  $N_{\text{lf}} \times \mathbb{R}^d$  be equipped with the product  $\sigma$ -field  $\mathcal{N}_{\text{lf}} \otimes \mathcal{B}^d$ , where  $\mathcal{B}^d$  is the Borel  $\sigma$ -field on  $\mathbb{R}^d$ . Similarly, for the point process  $\Phi$  we define  $\sigma$ -fields by replacing  $\mathbb{R}^d$  above with  $\Omega = \mathbb{R}^d \times (0, \infty) \times (0, \infty)$ .

Assume that the intensity function  $\rho(\xi)$  for  $X$  exists for Lebesgue almost all  $\xi \in \mathbb{R}^d$ ; cf. Proposition 1. Recall that the reduced Palm distribution  $P_\xi^!$  of  $X$  at  $\xi \in \mathbb{R}^d$  and the reduced Palm distribution  $P_{(c,b,\gamma)}^!$  of  $\Phi$  at  $(c, b, \gamma) \in \mathbb{R}^d \times (0, \infty) \times (0, \infty)$  are given by the Campbell–Mecke formula: for Lebesgue almost all  $\xi \in \mathbb{R}^d$ , with  $\rho(\xi) > 0$ , and for  $\zeta$ -almost all  $(c, b, \gamma) \in \mathbb{R}^d \times (0, \infty) \times (0, \infty)$ ,  $P_\xi^!$  and  $P_{(c,b,\gamma)}^!$  are uniquely given by

$$E \sum_{\xi \in X} f(X \setminus \{\xi\}, \xi) = \iint f(x, \xi) \rho(\xi) dP_\xi^!(x) d\xi \tag{23}$$

for nonnegative, measurable functions  $f$ , and

$$E \sum_{(c,b,\gamma) \in \Phi} h(\Phi \setminus \{(c, b, \gamma)\}, (c, b, \gamma)) = \iint h(\phi, (c, b, \gamma)) dP_{(c,b,\gamma)}^!(\phi) d\zeta(c, b, \gamma) \tag{24}$$

for nonnegative, measurable functions  $h$ ; see [10] and [44].

When  $\rho(\xi) > 0$ , we define the following. Let

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

where  $(c_\xi, b_\xi, \gamma_\xi)$  is a random variable with distribution

$$P((c_\xi, b_\xi, \gamma_\xi) \in D) = \int_D [\gamma k_b(c, \xi) / \rho(\xi)] d\zeta(c, b, \gamma)$$

for Borel sets  $D \subseteq \Omega$ . Conditional on  $(\mathbf{c}_\xi, b_\xi, \gamma_\xi)$ , let  $X_\xi$  and  $X^\xi$  denote independent point processes, where  $X_\xi$  is a Poisson process on  $\mathbb{R}^d$  with intensity function  $Z_\xi$  and  $X^\xi$  is a GSNCP driven by

$$Z^\xi(\eta) = \sum_{(\mathbf{c}, b, \gamma) \in \Phi(\mathbf{c}_\xi, b_\xi, \gamma_\xi)} \gamma k_b(\mathbf{c}, \eta), \quad \eta \in \mathbb{R}^d.$$

Here,  $\Phi^{(\mathbf{c}, b, \gamma)}$  denotes a point process with law  $P_{(\mathbf{c}, b, \gamma)}^1$ .

**Proposition 3.** For Lebesgue almost all  $\xi \in \mathbb{R}^d$  with  $\rho(\xi) > 0$ ,

$$P_\xi^1(F) = P(X^\xi \cup X_\xi \in F), \quad F \in \mathcal{M}_f.$$

*Proof.* By (23), and arguing as in the proof of Proposition 2 of [31], it suffices to show that

$$E \sum_{\xi \in X} \mathbf{1}(\xi \in A, (X \setminus \{\xi\}) \cap B = \emptyset) = \int_A P((X^\xi \cup X_\xi) \cap B = \emptyset) \rho(\xi) \, d\xi \tag{25}$$

for bounded  $A, B \in \mathcal{B}^d$ . It also follows from Proposition 2 of [31] that the left-hand side of (25) is given by

$$\int_A E \left( \sum_{(\mathbf{c}, b, \gamma) \in \Phi} \gamma k_b(\mathbf{c}, \xi) \exp \left( - \int_B \sum_{(\mathbf{c}', b', \gamma') \in \Phi} \gamma' k_{b'}(\mathbf{c}', \eta) \, d\eta \right) \right) d\xi.$$

By (24),

$$\begin{aligned} & E \left( \sum_{(\mathbf{c}, b, \gamma) \in \Phi} \gamma k_b(\mathbf{c}, \xi) \exp \left( - \int_B \sum_{(\mathbf{c}', b', \gamma') \in \Phi} \gamma' k_{b'}(\mathbf{c}', \eta) \, d\eta \right) \right) \\ &= \iint \gamma k_b(\mathbf{c}, \xi) \exp \left( - \int_B \sum_{(\mathbf{c}', b', \gamma') \in \Phi \cup \{(\mathbf{c}, b, \gamma)\}} \gamma' k_{b'}(\mathbf{c}', \eta) \, d\eta \right) dP_{(\mathbf{c}, b, \gamma)}^1(\phi) \, d\zeta(\mathbf{c}, b, \gamma) \\ &= \int \gamma k_b(\mathbf{c}, \xi) \exp \left( - \int_B \gamma k_b(\mathbf{c}, \eta) \, d\eta \right) \\ &\quad \times E \exp \left( - \int_B \sum_{(\mathbf{c}', b', \gamma') \in \Phi(\mathbf{c}, b, \gamma)} \gamma' k_{b'}(\mathbf{c}', \eta) \, d\eta \right) d\zeta(\mathbf{c}, b, \gamma) \\ &= \rho(\xi) P((X^\xi \cup X_\xi) \cap B = \emptyset), \end{aligned}$$

where we have used the conditional independence of  $X^\xi$  and  $X_\xi$ , given  $(\mathbf{c}_\xi, b_\xi, \gamma_\xi)$ , to obtain the last equality. Thereby, (25) is obtained.

**Remark 1.** The complication in using Proposition 3 lies in the necessity of determining  $P_{(\mathbf{c}, b, \gamma)}^1$ . This reduced Palm distribution is particularly simple for SNCs – see Proposition 2 of [31]. Suppose first that  $\Phi$  is a Cox process such that  $\Phi \mid \theta$  is a Poisson process with intensity function  $\lambda_\theta$ , where  $\theta$  is a random variable. Then

$$\int h(\phi) dP_{(\mathbf{c}, b, \gamma)}^1(\phi) = E(h(\Phi)\lambda_\theta(\mathbf{c}, b, \gamma))/\lambda(\mathbf{c}, b, \gamma), \tag{26}$$

where the mean is with respect to  $(\Phi, \theta)$ ,  $h \geq 0$  is any measurable function, and  $\lambda(\mathbf{c}, b, \gamma) \equiv E \lambda_\theta(\mathbf{c}, b, \gamma) \in (0, \infty)$ . This follows from first conditioning on  $\theta$  on the left-hand side of (24)

and using the Slivnyak–Mecke theorem and, next, by taking expectation and then conditioning on  $\Phi$ . Now suppose that  $\Phi_{\text{cent}}$  is a stationary Markov point process (see Example 4) with intensity  $\rho_{\text{cent}} > 0$ . If we assume, for simplicity, that  $(b_j, \gamma_j) = (b, \gamma)$  i.e. they are all equal and fixed, then  $P_c^! \equiv P_{(c,b,\gamma)}^!$  is concentrated on  $N_{\text{lf}}$  and satisfies

$$\int h(\phi) dP_c^!(\phi) = E(h(\Phi_{\text{cent}})\lambda^*(\Phi_{\text{cent}}, \mathbf{c}))/\rho_{\text{cent}}. \tag{27}$$

This follows from [38] or Equation (28) of [31] (which, incidentally, should be corrected by replacing  $\tilde{\lambda}(x, \xi)$  with  $\tilde{\lambda}(x, \xi)/\tilde{\rho}(\xi)$  (in the notation of [31])).

We now consider the  $J$ -function for a stationary GSNCP with intensity  $\rho \in (0, \infty)$ , assuming that  $k_1(\mathbf{c}, \xi) = k_1(\xi - \mathbf{c})$  and that (7) holds.

**Corollary 1.** *Under the assumptions above, for all  $r \geq 0$ ,*

$$J(r) = \frac{1}{\rho I_1(r)} \iint I_2(\mathbf{c}, b, \gamma, r) I_3(\mathbf{c}, b, \gamma, r) d\mathbf{c} d\chi(b, \gamma) \tag{28}$$

where

$$I_1(r) = E \exp\left(-\sum_j \int_{b(\mathbf{c}_j, r)} \gamma_j k_{b_j}(\xi) d\xi\right),$$

$$I_2(\mathbf{c}, b, \gamma, r) = \gamma k_b(-\mathbf{c}) \exp\left(-\int_{b(\mathbf{c}, r)} \gamma k_b(\eta) d\eta\right),$$

and

$$I_3(\mathbf{c}, b, \gamma, r) = E \exp\left(-\sum_{(\mathbf{c}', b', \gamma') \in \Phi(\mathbf{c}, b, \gamma)} \int_{b(\mathbf{c}', r)} \gamma' k_{b'}(\xi) d\xi\right).$$

*Proof.* By definition,

$$J(r) = P_0^!({x \in N_{\text{lf}} : x \cap b(\mathbf{0}, r) = \emptyset}) / P(X \cap b(\mathbf{0}, r) = \emptyset).$$

Hence, by Proposition 3,

$$J(r) = P((X^0 \cup X_0) \cap b(\mathbf{0}, r) = \emptyset) / P(X \cap b(\mathbf{0}, r) = \emptyset)$$

$$= \frac{1}{\rho P(X \cap b(\mathbf{0}, r) = \emptyset)} \iint P(X^0 \cap b(\mathbf{0}, r) = \emptyset \mid (\mathbf{c}_0, b_0, \gamma_0) = (\mathbf{c}, b, \gamma))$$

$$\times P(X_0 \cap b(\mathbf{0}, r) = \emptyset \mid (\mathbf{c}_0, b_0, \gamma_0) = (\mathbf{c}, b, \gamma)) \gamma k_b(-\mathbf{c}) d\mathbf{c} d\chi(b, \gamma).$$

Since the latter three probabilities are equal to  $I_1(r)$ ,  $I_2(\mathbf{c}, b, \gamma, r)/\gamma k_b(-\mathbf{c})$ , and  $I_3(\mathbf{c}, b, \gamma, r)$ , respectively, we obtain (28).

**Example 7.** It is well known that, for stationary Poisson cluster processes and SNCs,  $J(r) \leq 1$  and  $J$  is nonincreasing [48], [31]. Below, we show that these properties hold for certain GSNCPs. We let the situation be as in Corollary 1, and recall that  $d\zeta(\mathbf{c}, b, \gamma) = d\mathbf{c} d\chi(b, \gamma)$ . Notice that if, for  $\zeta$ -almost all  $(\mathbf{c}, b, \gamma)$  and  $r > 0$ ,

$$I_1(r) \geq I_3(\mathbf{c}, b, \gamma, r), \tag{29}$$

then (20) and (28) imply that  $J(r) < 1$  for  $r > 0$ .

If  $\Phi$  is a Poisson process, so that  $\Phi_{\text{cent}}$  is stationary, then  $I_3(\mathbf{c}, b, \gamma, r) = I_1(r)$  and, so, by (28),

$$J(r) = \frac{1}{\rho} \iint \gamma k_b(-\mathbf{c}) \exp\left(-\int_{b(\mathbf{c},r)} \gamma k_b(\boldsymbol{\eta}) \, d\boldsymbol{\eta}\right) \, d\mathbf{c} \, d\chi(b, \gamma).$$

This is a slight extension of Corollary 1 of [31] (where, incidentally, there is a minor typo:  $k(\mathbf{c})$  should read  $k(-\mathbf{c})$  (in the notation of [31]). It follows that  $J$  is nonincreasing and  $J(r) < 1$  for  $r > 0$ .

Now consider a generalised Neyman–Scott process. Suppose that  $\Phi_{\text{cent}}$  is a mixed Poisson process driven by a positive random variable  $\theta$  (i.e.  $\Phi_{\text{cent}} \mid \theta$  is a stationary Poisson process with intensity  $\theta$ ), where  $(\theta, \Phi_{\text{cent}})$  is independent of the  $(b_j, \gamma_j)$ , and these are i.i.d. with distribution  $Q$  (case (iii) of Example 2). In other words,  $\Phi$  is a Cox process driven by a random measure  $\Lambda$  given by  $d\Lambda(\mathbf{c}, b, \gamma) = \theta \, d\mathbf{c} \, dQ(b, \gamma)$ . Let  $\pi$  denote the distribution of  $\theta$  and define another positive random variable  $\theta'$  with distribution

$$\pi'(A) = \frac{1}{E\theta} \int_A \theta \, d\pi(\theta)$$

for Borel sets  $A \subseteq (0, \infty)$ , and suppose that

$$\pi \leq_{\text{st}} \pi', \tag{30}$$

i.e.  $\pi((t, \infty)) \leq \pi'((t, \infty))$  for all  $t > 0$  (the usual stochastic order). Equation (30) is satisfied if  $\pi$  is a gamma distribution, for instance. We claim that (30) implies (29): since  $\Phi$  is a Cox process, a slight modification of the proof of (26) shows that

$$\begin{aligned} I_3(\mathbf{c}, b, \gamma, r) &= E\left(\exp\left(-\sum_j \int_{b(\mathbf{c}_j,r)} \gamma_j k_{b_j}(\boldsymbol{\xi}) \, d\boldsymbol{\xi}\right)\theta\right) (E\theta)^{-1} \\ &= E \exp\left(-\sum_j \int_{b(\mathbf{c}'_j,r)} \gamma_j k_{b_j}(\boldsymbol{\xi}) \, d\boldsymbol{\xi}\right), \end{aligned}$$

where  $\{(\mathbf{c}'_j, b_j, \gamma_j)\} = \Phi'$  is a Cox process driven by  $\Lambda'$ , itself given by

$$d\Lambda'(\mathbf{c}, b, \gamma) = \theta' \, d\mathbf{c} \, dQ(b, \gamma).$$

For locally finite measures  $\mu_1$  and  $\mu_2$  on  $\Omega$ , define a partial order  $\leq$  by

$$\mu_1 \leq \mu_2 \quad \text{whenever } \mu_1(B) \leq \mu_2(B) \text{ for all bounded Borel sets } B \subset \Omega.$$

Furthermore, let  $\leq_{\text{st}}$  denote the usual stochastic order for locally finite random measures  $\Gamma_1$  and  $\Gamma_2$  on  $\Omega$ , i.e.

$$\Gamma_1 \leq_{\text{st}} \Gamma_2 \Leftrightarrow E f(\Gamma_1) \leq E f(\Gamma_2) \quad \text{whenever } f \text{ is increasing with respect to } \leq.$$

By (30),  $\Lambda \leq_{\text{st}} \Lambda'$  and, so, by Theorem 7.4.7 of [36],  $\Phi \leq_{\text{st}} \Phi'$ . Now,  $I_3(\mathbf{c}, b, \gamma, r) = I_3(r)$  depends only on  $r$ , and it is of the same form as  $I_1(r)$  except that  $\Phi_{\text{cent}}$  is replaced by  $\Phi'_{\text{cent}}$ . Consequently, (29) holds since the function

$$f(\mu) = \exp\left(-\sum_{(\mathbf{c},b,\gamma) \in \mu} \int_{b(\mathbf{c},r)} \gamma k_b(\boldsymbol{\xi}) \, d\boldsymbol{\xi}\right),$$

where  $\mu$  is a locally finite measure on  $\Omega$ , is nonincreasing with respect to  $\leq$ ,  $I_1(r) = E f(\Phi)$  and  $I_3(r) = E f(\Phi')$ .

Next, consider a generalised shot noise G Cox process as in Example 3, and assume that  $\theta = (\alpha, \kappa, \tau)$  is a random variable with distribution  $\pi$ , and that the  $b_j$  are i.i.d. with distribution  $\nu$ , and independent of  $(\theta, \{(c_j, \gamma_j)\})$ . Then, the Cox process  $\Phi$  is driven by a  $\Lambda$  such that  $d\Lambda(c, b, \gamma) = \lambda_\theta(\gamma) dc d\nu(b) d\gamma$ , where  $\lambda_\theta(\gamma)$  is given by (10), and we have

$$I_3(c, b, \gamma, r) = E \left( \exp \left( - \sum_j \int_{b(c_j, r)} \gamma_j k_{b_j}(\xi) d\xi \right) \lambda_\theta(\gamma) \right) (E \lambda_\theta(\gamma))^{-1}.$$

Arguing, as above, for a generalised Neyman–Scott process, (29) holds if  $\pi_\gamma \leq_{st} \pi$  for any  $\gamma > 0$ , where

$$\pi_\gamma(A) = \frac{\int_A \lambda_\theta(\gamma) d\pi(\theta)}{\int \lambda_\theta(\gamma) d\pi(\theta)}$$

for Borel sets  $A \subseteq (0, \infty)$ . For instance, if  $(\alpha, \kappa)$  is fixed and  $\tau$  follows a gamma distribution with inverse scale parameter  $\nu$ , then  $\pi_\gamma$  is also a gamma distribution with the same shape parameter but inverse scale parameter  $\nu + \gamma$  and, so,  $\pi_\gamma \leq_{st} \pi$  for any  $\gamma > 0$ .

Finally, consider a GSNCP in which  $\Phi_{cent}$  is a stationary Markov point process on  $\mathbb{R}^d$ , with intensity  $\rho_{cent} \in (0, \infty)$ , and the  $(b_j, \gamma_j) = (b, \gamma)$  are all equal and fixed. Then, (28) becomes

$$J(r) = \frac{1}{I_1(r)} \int I_2(c, b, \gamma, r) I_3(c, b, \gamma, r) dc,$$

where

$$I_3(c, b, \gamma, r) = E \left( \exp \left( - \sum_j \int_{b(c_j, r)} \gamma_j k_b(\xi) d\xi \right) \frac{\lambda^*(\Phi_{cent}, c)}{\rho_{cent}} \right);$$

cf. (27). However, the authors do not know how to get any further with this expression.

#### 4. Simulation of generalised shot noise Cox processes

In this section, we consider various simulation algorithms for the restriction  $X_W = X \cap W$  of  $X$  to a Borel set  $W \subset \mathbb{R}^d$  with volume  $|W| \in (0, \infty)$ , and for conditional simulation of  $\Phi$ , given  $X_W$ .

##### 4.1. Simulation with edge effects and truncation

Clearly,  $X_W$  is a Cox process driven by the random field  $Z_W(\xi) = \mathbf{1}(\xi \in W)Z(\xi)$ . For the simulation of  $X_W$ , a truncation may be required due to the possibly infinite point process  $\Phi$  entering  $Z_W$ . Assume that  $W_{ext} \subseteq \mathbb{R}^d$ ,  $B \leq \infty$ , and  $\varepsilon \geq 0$  such that  $\zeta(D)$ , where  $D = W_{ext} \times (0, B) \times (\varepsilon, \infty)$ , is finite. For instance, for the specific example of a generalised shot noise G process in Example 3, the condition on  $\varepsilon$  is needed if  $P(\alpha \geq 0) > 0$ .

An approximate simulation of  $X_W$  is obtained by first simulating the point process  $\Phi \cap D$  and, next, the corresponding independent Poisson processes  $X_j \cap W$  with intensity functions  $\xi \mapsto \gamma_j k_{b_j}(c_j, \xi) \mathbf{1}(\xi \in W)$ ,  $(c_j, b_j, \gamma_j) \in \Phi \cap D$ . As was noticed in [31], edge effects may enter in applications since  $W_{ext}$  is typically a bounded window such that  $W \subset W_{ext}$ . Below, we quantify the error of such approximate simulations by extending Proposition 3 of [31].

Let

$$M_W = \sum_j \mathbf{1}(\{c_j \notin W_{ext}\} \cup \{b_j \geq B\} \cup \{\gamma_j \leq \varepsilon\}) \text{card}(X_j \cap W)$$

denote the number of missing points when we make an approximate simulation of  $X_W$  by ignoring clusters  $X_j$  with  $\mathbf{c}_j \notin W_{\text{ext}}, b_j \geq B$ , or  $\gamma_j \leq \varepsilon$ .

Furthermore, let  $q_W$  be the probability that some cluster  $X_j$  with  $\mathbf{c}_j \notin W_{\text{ext}}, b_j \geq B$ , or  $\gamma_j \leq \varepsilon$  has a point in  $W$ . Finally, assume that there exists a function  $k_W^{\text{dom}}: (0, \infty) \times \mathbb{R}^d \times \mathbb{R}^d \rightarrow [0, \infty)$  satisfying the following conditions.

**Condition 1.** If  $\xi \in W$  then  $k_W^{\text{dom}}(b, \mathbf{c}, \xi) \geq k_b(\mathbf{c}, \xi)$ ; otherwise,  $k_W^{\text{dom}}(b, \mathbf{c}, \xi) = 0$ .

**Condition 2.** The integral

$$a_W^{\text{dom}}(\mathbf{c}, b) = \int_W k_W^{\text{dom}}(b, \mathbf{c}, \xi) \, d\xi, \quad (\mathbf{c}, b) \in \mathbb{R}^d \times (0, \infty),$$

can be easily computed.

**Proposition 4.** Given these definitions,

$$E M_W \leq \int \mathbf{1}(\{\mathbf{c} \notin W_{\text{ext}}\} \cup \{b \geq B\} \cup \{\gamma \leq \varepsilon\}) \gamma a_W^{\text{dom}}(\mathbf{c}, b) \, d\zeta(\mathbf{c}, b, \gamma) \tag{31}$$

and

$$q_W \leq 1 - \exp\left(- \int \mathbf{1}(\{\mathbf{c} \notin W_{\text{ext}}\} \cup \{b \geq B\} \cup \{\gamma \leq \varepsilon\}) \gamma a_W^{\text{dom}}(\mathbf{c}, b) \, d\zeta(\mathbf{c}, b, \gamma)\right). \tag{32}$$

*Proof.* The proof of (31) is similar to that of Proposition 3 of [31]. Conditional on  $\Phi$ , the clusters  $X_j$  are independent Poisson processes with intensity functions  $\gamma_j k_{b_j}(\mathbf{c}_j, \cdot)$ , so

$$q_W = 1 - E \prod_j \left[ \exp\left(-\gamma_j \int_W k_{b_j}(\mathbf{c}_j, \xi) \, d\xi\right) \right]^{\mathbf{1}(\{\mathbf{c}_j \notin W_{\text{ext}}\} \cup \{b_j \geq B\} \cup \{\gamma_j \leq \varepsilon\})}. \tag{33}$$

Hence, by Jensen’s inequality and (5),

$$\begin{aligned} q_W &\leq 1 - \exp\left(- E \sum_j \mathbf{1}(\{\mathbf{c}_j \notin W_{\text{ext}}\} \cup \{b_j \geq B\} \cup \{\gamma_j \leq \varepsilon\}) \gamma_j \int_W k_{b_j}(\mathbf{c}_j, \xi) \, d\xi\right) \\ &= 1 - \exp\left(- \int \mathbf{1}(\{\mathbf{c} \notin W_{\text{ext}}\} \cup \{b \geq B\} \cup \{\gamma \leq \varepsilon\}) \gamma \int_W k_b(\mathbf{c}, \xi) \, d\xi \, d\zeta(\mathbf{c}, b, \gamma)\right). \end{aligned}$$

Thereby, using Conditions 1 and 2, (32) is obtained.

**Remark 2.** When  $\Phi$  is a Poisson process, (32) can be improved – see [34]. Indeed, using (33), Conditions 1 and 2, and arguments similar to the derivation of the generating functional of Poisson processes (see [10] and [34]), we find that

$$\begin{aligned} 1 - q_W &= \exp\left[- \int \mathbf{1}(\{\mathbf{c} \notin W_{\text{ext}}\} \cup \{b \geq B\} \cup \{\gamma \leq \varepsilon\}) \right. \\ &\quad \times \left. \left( 1 - \exp\left(-\gamma \int k_b(\mathbf{c}, \xi) \, d\xi\right) \right) \right] d\zeta(\mathbf{c}, b, \gamma) \\ &\geq \exp\left[- \int \mathbf{1}(\{\mathbf{c} \notin W_{\text{ext}}\} \cup \{b \geq B\} \cup \{\gamma \leq \varepsilon\}) \right. \\ &\quad \times \left. \left( 1 - \exp(-\gamma a_W^{\text{dom}}(\mathbf{c}, b)) \right) \right] d\zeta(\mathbf{c}, b, \gamma). \end{aligned}$$

Note also that, by the coupling inequality (see [23]), the upper bound on  $q_W$  in (32) is also an upper bound on the total variation distance between the law of  $X_W$  and the law of the truncated process.

**Example 8.** Suppose that (7) holds. For specific models of  $\chi$  and  $k_1$ , the upper bounds in Proposition 4 can be calculated using methods similar to those in [7] and [31]. For example, let  $W = b(\mathbf{0}, R)$  and  $W_{\text{ext}} = b(\mathbf{0}, R + r)$ , where  $R$  and  $r$  are two positive constants, and let  $k_1$  be the Gaussian kernel (3). Define  $k_W^{\text{dom}}(b, \xi - c) = \sup_{\eta \in W} k_b(\eta - c)$  for  $\xi \in W$ , and set  $\sigma_d = 2\pi^{d/2} / \Gamma(\frac{1}{2}d)$ . Proceeding as in Example 5 of [31],

$$a(\|c\|, b) := a_W^{\text{dom}}(c, b) = \left[ \frac{\omega_d R^d}{(2\pi b^2)^{d/2}} \right] \exp \left[ -\mathbf{1}(\|c\| > R) \frac{(\|c\| - R)^2}{2b^2} \right]$$

and the integral in (31) and (32) is given by

$$\sigma_d \int_0^\infty \int_0^\infty \gamma \, d\chi(b, \gamma) \int_0^\infty s^{d-1} a(s, b) \, ds - \sigma_d \int_0^B \int_\varepsilon^\infty \gamma \, d\chi(b, \gamma) \int_0^{R+r} s^{d-1} a(s, b) \, ds.$$

This may be determined by numerical methods for specific models of generalised Neyman–Scott and generalised shot noise G Cox processes; see [7] and [31].

**4.2. Simulation without edge effects and truncation**

Perfect simulation of  $X_W$  can be obtained by independent thinning of the nonempty clusters in a Cox process  $X_W^{\text{dom}}$  driven by

$$Z_W^{\text{dom}}(\xi) = \sum_j \gamma_j k_W^{\text{dom}}(b_j, c_j, \xi),$$

where  $k_W^{\text{dom}}$  satisfies Conditions 1 and 2. The details for SNCPPs are carefully discussed in [7], [31], and [34], so in this section we give only a description of the algorithm and discuss some applications for GSNCPs.

Let  $X_j^{\text{dom}}, (c_j, b_j, \gamma_j) \in \Phi$ , denote the clusters of  $X_W^{\text{dom}}$ , and let  $\Phi_W^{\text{dom}} = \{(c_j, b_j, \gamma_j) \in \Phi : X_j^{\text{dom}} \neq \emptyset\}$ . In addition to Conditions 1 and 2, assume that the following condition holds.

**Condition 3.** *It is feasible to simulate  $\Phi_W^{\text{dom}}$  (and, hence,  $\Phi_W^{\text{dom}}$  is almost surely finite).*

Example 9, below, demonstrates that having the ability to check Condition 3 will depend very much on the choice of model for  $\Phi$ , the kernel  $k_1$ , etc.

**Perfect simulation algorithm for GSNCPs.**

- (a) Generate the point process  $\Phi_W^{\text{dom}} = \{(c_1, b_1, \gamma_1), \dots, (c_N, b_N, \gamma_N)\}$ .
- (b) For each  $j = 1, \dots, N$ , generate
  - (i)  $X_j^{\text{dom}}$ , which is distributed as a conditional Poisson process with intensity function  $\gamma_j k_W^{\text{dom}}(b_j, c_j, \cdot)$ , given that it is nonempty; and
  - (ii)  $X'_j$ , which is an independent thinning of  $X_j^{\text{dom}}$  with retention probabilities

$$k_{b_j}(c_j, \xi) / k_W^{\text{dom}}(b_j, c_j, \xi) \quad \text{for } \xi \in X_j^{\text{dom}}.$$

- (c) Return  $\bigcup_j X'_j$ .

For the loop in step (b), it is implicit that the generation of the processes in parts (i) and (ii) is independent of previous generations. The output in step (c) follows the same distribution as  $X_W$ ; see, e.g. Proposition 4 of [31].

**Example 9.** For the arguments below, it is useful to notice that if  $\Phi$  is a Cox process, we can obtain  $\Phi_W^{\text{dom}}$  by independent thinning of  $\Phi$  with retention probabilities

$$p_W^{\text{dom}}(\mathbf{c}, b, \gamma) = 1 - \exp(-\gamma a_W^{\text{dom}}(\mathbf{c}, b)), \quad (\mathbf{c}, b, \gamma) \in \Omega.$$

Suppose that  $\Phi$  is a Poisson process. Then, clearly,  $\Phi_W^{\text{dom}}$  is a Poisson process on  $\Omega$ , with intensity measure

$$\zeta_W^{\text{dom}}(D) = \int_D p_W^{\text{dom}}(\mathbf{c}, b, \gamma) \, d\zeta(\mathbf{c}, b, \gamma).$$

Condition 3 means that  $\zeta_W^{\text{dom}}(\Omega) < \infty$  and we are able to generate  $\Phi_W^{\text{dom}}$ . Example 6 of [31] gives a detailed discussion of this condition for SNCs, and this discussion easily extends to the present case.

Now suppose that  $\Phi$  is a generalised Neyman–Scott process, where the  $(b_j, \gamma_j)$  are i.i.d. with distribution  $Q$ , and  $\Phi_{\text{cent}}$  is a Cox process driven by a random field  $Y(\cdot)$ , which is bounded by a constant  $M$  and independent of the  $(b_j, \gamma_j)$ . Furthermore, let  $k_1, W, W_{\text{ext}}, k_W^{\text{dom}}$ , and  $a$  be defined as in Example 8. Then  $\Phi_W^{\text{dom}}$  is a Cox process on  $\Omega$  driven by the random measure  $dm(\mathbf{c}, b, \gamma) = p_W^{\text{dom}}(\mathbf{c}, b, \gamma)Y(\mathbf{c}) \, d\mathbf{c} \, dQ(b, \gamma)$ , and

$$E \Phi_W^{\text{dom}} = E m(\Omega) \leq M \sigma_d \int_0^\infty \int_{(0, \infty)^2} s^{d-1} [1 - \exp(-\gamma a(s, b))] \, ds \, dQ(b, \gamma) =: \beta_W^{\text{dom}},$$

where we assume that  $\beta_W^{\text{dom}}$  is finite and can be determined by numerical integration (see, e.g. page 628 of [31]). At least in principle, simulation of  $\Phi_W^{\text{dom}}$  is straightforward by thinning. First, simulate a Poisson variate  $N$  with mean  $\beta_W^{\text{dom}}$ . Second, generate independent points  $(s_j \mathbf{u}_j, b_j, \gamma_j)$  for  $j = 1, \dots, N$ , where  $\mathbf{u}_j$  is a uniformly distributed unit vector in  $\mathbb{R}^d$ ,  $(s_j, b_j, \gamma_j)$  has distribution

$$P(A) \propto \iint \mathbf{1}((s, b, \gamma) \in A) s^{d-1} [1 - \exp(-\gamma a(s, b))] \, ds \, dQ(b, \gamma), \quad A \subseteq (0, \infty)^3,$$

and  $\mathbf{u}_j$  is independent of  $(s_j, b_j, \gamma_j)$  (here rejection sampling may be useful). Third, simulate  $Y(s_j \mathbf{u}_j)$  for  $j = 1, \dots, N$ , and assign each  $(s_j \mathbf{u}_j, b_j, \gamma_j)$  to  $\Phi_W^{\text{dom}}$  with probability  $Y(s_j \mathbf{u}_j)/M$  (where these assignments are independent).

Similar ideas apply for a generalised shot noise G Cox process. Consider the case of Example 3, in which the  $b_j$  are i.i.d. with distribution  $\nu$ , and again let  $k_1, W$ , etc. be as in Example 8. Then  $\Phi_W^{\text{dom}}$  is a Cox process driven by the random measure

$$dm_\theta(\mathbf{c}, b, \gamma) = p_W^{\text{dom}}(\mathbf{c}, b, \gamma) \lambda_\theta(\gamma) \, d\mathbf{c} \, d\nu(b) \, d\gamma,$$

where  $\lambda_\theta$  is given by (10), and, so,  $\Phi_W^{\text{dom}}$  is almost surely finite if

$$E \Phi_W^{\text{dom}} = \sigma_d \int_0^\infty \int_0^\infty \int_0^\infty s^{d-1} [1 - \exp(-\gamma a(s, b))] E \lambda_\theta(\gamma) \, ds \, d\nu(b) \, d\gamma$$

is finite. We can then, at least in principle, simulate  $\Phi_W^{\text{dom}}$  by generating first  $\theta$  and then  $\Phi_W^{\text{dom}}$  as a Poisson process with intensity measure  $m_\theta$ .

Finally, consider the Markov point process setting of Example 4, where, for simplicity, we assume that  $\Phi_{\text{cent}}$  is a finite Markov point process contained in the bounded window  $W_{\text{ext}}$ . We cannot exploit the same ideas as above, since  $\Phi_{\text{cent}}$  is not easily viewed as a Cox process. However, dominated coupling from the past can be used for making a perfect simulation of  $\Phi_{\text{cent}}$  if this process is locally stable [21]. Then, since  $\Phi_{\text{cent}}$  is independent of the  $(b_j, \gamma_j)$ , which are either i.i.d. or identical (as in Example 4), it may be straightforward to simulate  $\Phi_W^{\text{dom}}$  using acceptance–rejection sampling for the clusters. That is, we simulate  $X_j^{\text{dom}}$  without conditioning and accept (i.e. return)  $X_j^{\text{dom}}$  if it is nonempty, repeating the procedure until acceptance is obtained.

**4.3. Conditional simulation**

Assume that  $X_W = x = \{x_1, \dots, x_m\} \neq \emptyset$  is observed. Simulation of the conditional distribution of  $\Phi$ , given that  $X_W = x$ , is needed to predict  $\Phi$  and also to perform likelihood and Bayesian inference based on MCMC methods; see Section 4.3 of [31] and the references therein. In this section, following [15] (see also [1], [14], [24], [29], [34], and [49]), we first describe the birth–death–move Metropolis–Hastings algorithm for conditional simulation of the process  $\Phi \mid (X_W = x)$  (this is a particular case of Green’s reversible jump MCMC algorithm [16]). Second, we give sufficient conditions that guarantee geometric ergodicity of the algorithm (or, in fact,  $V$ -uniform ergodicity; the reader is referred to [27] or Section 7.2 of [34] for background material on Markov chains).

Throughout this section, we assume the following: the process  $\Phi$  is almost surely finite and contained in  $D = W_{\text{ext}} \times B \times (\varepsilon, \infty)$ , with  $W \subseteq W_{\text{ext}} \subseteq \mathbb{R}^d$ ,  $B \subseteq (0, \infty)$ , and  $\varepsilon \geq 0$ . Depending on the context, it may be natural to consider cases in which  $B$  is a bounded or unbounded interval and, for technical reasons, when establishing geometric ergodicity, we will first take  $\varepsilon$  to be positive. Furthermore,  $\nu$  denotes a Poisson process on  $D$  with a diffuse intensity measure  $\mu$  such that  $0 < \mu(D) < \infty$ . Finally,  $\Phi$  has a density  $p$  with respect to  $\nu$ .

Given these assumptions,  $X_W \mid \Phi$  is a Poisson process and it has a density, with respect to a unit-rate Poisson process on  $W$ , given by

$$f(x \mid \Phi) = \exp\left(|W| - \int_W Z(\xi) \, d\xi\right) \prod_{\xi \in x} Z(\xi).$$

Hence, an unnormalised density for  $\Phi \mid (X_W = x)$  with respect to  $\nu$  is given by

$$\pi(\phi \mid x) = f(x \mid \phi)p(\phi).$$

The corresponding normalizing constant is unknown, but depends only on the data  $x$ .

We turn now to the description of the birth–death–move Metropolis–Hastings algorithm. It generates a Markov chain defined on

$$\Omega_x = \{\phi \subset D : \text{card}(\phi) < \infty, \pi(\phi \mid x) > 0\}.$$

Let  $0 < q_1 < 1$  and  $0 < q_2 \leq 1$  be given numbers and, for  $\phi \in \Omega_x$  and  $(c, b, \gamma) \in D$ , define

$$r[\phi, (c, b, \gamma)] = \frac{(1 - q_1)\pi(\phi \cup \{(c, b, \gamma)\} \mid x)\mu(D)}{q_1\pi(\phi \mid x)(\text{card}(\phi) + 1)}. \tag{34}$$

If  $\phi = \{(c_1, b_1, \gamma_1), \dots, (c_n, b_n, \gamma_n)\} \in \Omega_x$  is the current state of the chain, the next state is generated as follows.

- (a) With probability  $q_1 q_2$ , make a birth step:
  - generate  $(c, b, \gamma)$  with law  $\mu/\mu(D)$ ;
  - with probability  $\min\{1, r[\phi, (c, b, \gamma)]\}$ , return  $\phi \cup \{(c, b, \gamma)\}$  as the next state.
- (b) With probability  $(1 - q_1)q_2$ , make a death step (if  $n > 0$ ):
  - generate  $i$  according to the uniform law on  $\{1, \dots, n\}$ ;
  - with probability  $\min\{1, r[\phi \setminus \{(c_i, b_i, \gamma_i)\}, (c_i, b_i, \gamma_i)]^{-1}\}$ , return  $\phi \setminus \{(c_i, b_i, \gamma_i)\}$  as the next state.
- (c) With probability  $1 - q_2$ , make a move step (if  $n > 0$ ):
  - generate  $i$  according to the uniform law on  $\{1, \dots, n\}$ ;
  - generate  $(c, b, \gamma)$  with law  $\mu/\mu(D)$ ;
  - with probability  $\min\{1, \pi((\phi \setminus \{(c_i, b_i, \gamma_i)\}) \cup \{(c, b, \gamma)\} \mid x) / \pi(\phi \mid x)\}$ , return  $(\phi \setminus \{(c_i, b_i, \gamma_i)\}) \cup \{(c, b, \gamma)\}$  as the next state.
- (d) Otherwise, return  $\phi$  as the next state.

Assuming that the initial state is in  $\Omega_x$ , the chain stays in  $\Omega_x$ . Note that the empty point configuration is not contained in  $\Omega_x$ , since  $x \neq \emptyset$ . Furthermore, the chain is reversible with invariant (unnormalised) density  $\pi(\cdot \mid x)$ . This follows along similar lines as in the proofs of Propositions 7.11, 7.12, and 7.15 of [34].

We consider two situations in which irreducibility of the chain is satisfied. If, for any  $\phi \in \Omega_x$ ,  $\phi' \in \Omega_x$ ,  $\psi \subset \phi$ , and  $\psi' \subset \phi'$ , we have  $\phi \cup \psi' \in \Omega_x$  and  $(\phi \setminus \psi) \cup \phi' \in \Omega_x$ , then irreducibility holds. Briefly, this follows because the chain can then move up from  $\phi$  to  $\phi \cup \psi'$ , and then down to  $\phi'$ . Note that  $f(x \mid \phi) > 0$  if  $f(x \mid \phi') > 0$  and  $\phi' \subset \phi$ , and that  $\pi(\phi \mid x) > 0$  if and only if  $f(x \mid \phi) > 0$  and  $p(\phi) > 0$ . Thus, irreducibility holds if, for all  $\phi \in \Omega_x$  and  $(c, b, \gamma) \in D$ ,

$$p(\phi) > 0 \Rightarrow p(\phi \cup \{(c, b, \gamma)\}) > 0. \tag{35}$$

This condition is satisfied for SNCs (since  $p = 1$  in [31]). However, (35) is not always satisfied for other models of interest. For example, if a hard-core condition is imposed, so that  $p(\phi \cup \{(c, b, \gamma)\}) = 0$  if a cluster centre from  $\phi$  is sufficiently close to  $c$ , then (35) is violated. In fact, the opposite will often hold:  $p$  is said to be hereditary on  $\Omega_x$  if, for all  $\phi \in \Omega_x$  and  $(c, b, \gamma) \in D$ ,

$$p(\phi \cup \{(c, b, \gamma)\}) > 0 \Rightarrow p(\phi) > 0. \tag{36}$$

In that case, we need to impose further conditions: if  $q_2 < 1$  (i.e. moves are possible),

$$k_1(c, x_i) > 0 \quad \text{for all } c \in W_{\text{ext}} \text{ and } i = 1, \dots, m, \tag{37}$$

(i.e.  $f(x \mid \phi) > 0$  for all finite, nonempty  $\phi \subset D$ ), and (36) holds, then we have irreducibility. Briefly, this follows since  $\pi(\cdot \mid x)$  is now hereditary on  $\Omega_x$ , so, if  $\phi \in \Omega_x$ ,  $\phi' \in \Omega_x$ ,  $(c, b, \gamma) \in \phi$ , and  $(c', b', \gamma') \in \phi'$ , the chain can first move down from  $\phi$  to  $(c, b, \gamma)$ , then move to  $(c', b', \gamma')$  and, finally, move up to  $\phi'$ . Note that (37) is clearly satisfied for a positive kernel such as the Gaussian kernel (3), but might easily be violated for the uniform kernel (4).

In the sequel, irreducibility is assumed. Note that the chain is clearly aperiodic, since it can stay in the same state for one or more transitions. Below, conditions stronger than those above will be assumed in order to establish  $V$ -uniform ergodicity. We consider the cases with and without moves separately (i.e. the two cases  $q_2 < 1$  and  $q_2 = 1$ ).

The following Proposition 5 concerns the case  $q_2 = 1$ , in which we refer to our Metropolis–Hastings algorithm as the birth–death algorithm. The proposition subsumes Proposition 5 of [31] (where the  $b_j$  are equal and fixed,  $\Phi$  is a Poisson process, and  $p = 1$ ). The following conditions are assumed to hold.

- (a) The density  $p$  is locally stable, i.e. there is a finite constant  $\Lambda$  such that

$$p(\phi \cup \{(c, b, \gamma)\}) \leq \Lambda p(\phi)$$

for all finite  $\phi \subset D$  and  $(c, b, \gamma) \in D$ .

- (b) For any positive integer  $K$ , there exists a positive function  $h_K$  such that

$$h_K(c, b, \gamma)p(\phi) \leq p(\phi \cup \{(c, b, \gamma)\})$$

for all  $(c, b, \gamma) \in D$  and  $\phi \in \Omega_x$  for which  $\text{card}(\phi) \leq K$ .

- (c) There exists a positive constant  $\delta > 0$  such that

$$\int_W k_b(c, \xi) \, d\xi \geq \delta$$

for any  $(c, b) \in W_{\text{ext}} \times B$ .

- (d) For  $i = 1, \dots, m$ ,

$$\int_D \mathbf{1}(k_b(c, x_i) > 0) \, d\mu(c, b, \gamma) > 0.$$

- (e) There exist positive constants  $\delta'_1, \dots, \delta'_m, \delta''_1, \dots, \delta''_m$  such that

$$\delta'_i \leq k_b(c, x_i) \leq \delta''_i \quad \text{whenever } k_b(c, x_i) > 0 \text{ and } (c, b) \in W_{\text{ext}} \times B.$$

Conditions (a) and (b) are automatically satisfied for SNCPS (as  $p = 1$ ). For GSNCPs, condition (b) implies (35) and, so, irreducibility is ensured. Conditions (a) to (e) are further discussed in Example 10, below.

**Proposition 5.** *Assume that  $q_2 = 1$ ,  $\varepsilon > 0$ , and conditions (a) to (e) are satisfied. Then, for any function  $V(\phi) = \beta^{\text{card}(\phi)}$ ,  $\phi \in \Omega_x$ , with  $\beta > 1$ , the birth–death algorithm is  $V$ -uniformly ergodic.*

*Proof.* We show first that, for any positive integer  $N$ , the set  $S_N = \{\phi \in \Omega_x : \text{card}(\phi) \leq N\}$  is a small set. This means that  $P^t(\phi, F) \geq \varepsilon' Q(F)$  for any  $\phi \in S_N$  and any measurable set  $F \subseteq \Omega_x$ , where  $\varepsilon' > 0$  is some positive constant,  $Q$  is some nonzero measure on  $\Omega_x$ , and  $P^t(\phi, \cdot)$  denotes the  $t$ -step transition probability of the Metropolis–Hastings chain when it starts in  $\phi$ .

We shall use the following bounds. Let  $\phi \in \Omega_x$  and  $(c, b, \gamma) \in D$ . Arguing as at the beginning of the proof of Proposition 5 of [31], using conditions (a), (c), and (e), it is seen that

$$\exp(-\gamma/b^d)p(\phi \cup \{(c, b, \gamma)\})/p(\phi) \leq \pi(\phi \cup \{(c, b, \gamma)\} \mid x)/\pi(\phi \mid x) \leq M, \tag{38}$$

where  $M > 0$  is a constant. Furthermore,  $0 < \mu(D) < \infty$  implies that, for any positive integer  $K$ , there exist a constant  $L_K > \varepsilon$  and a nonvoid interval  $B_K \subseteq B$  such that  $\rho_K := \inf_{B_K} \in (0, \infty)$ ,  $((1 - q_1)/q_1) \exp(-L_K/\rho_K^d)\mu(D)\Lambda \leq 1$ , and  $\mu(W_{\text{ext}} \times B \times (\varepsilon, L_K)) > 0$ . Furthermore, by condition (d), and since  $h_K$  is a positive function,

$$\int \mathbf{1}(k_b(\mathbf{c}, \mathbf{x}_i) > 0, b \in B_K, \gamma < L_K)h_K(\mathbf{c}, b, \gamma) \, d\mu(\mathbf{c}, b, \gamma) > 0, \quad i = 1, \dots, m. \tag{39}$$

Let  $m' > \max\{m, ((1 - q_1)/q_1)M\mu(D) - 1\}$  be an integer and set  $a = \min\{q_1, 1 - q_1\}$ . For any  $\phi = \{(\mathbf{c}_1, b_1, \gamma_1), (\mathbf{c}_2, b_2, \gamma_2), \dots, (\mathbf{c}_k, b_k, \gamma_k)\} \in S_N$  and measurable  $F \subseteq \Omega_x$ ,

$$P^{m'+N}(\phi, F) \geq \int_D \dots \int_D \int_D \dots \int_D \times \prod_{l=1}^{m'} a \min\{1, r[\phi \cup \{(\mathbf{v}_1, u_1, \sigma_1), \dots, (\mathbf{v}_{l-1}, u_{l-1}, \sigma_{l-1})\}, (\mathbf{v}_l, u_l, \sigma_l)]\} \tag{40}$$

$$\times \prod_{j=1}^k \frac{a}{k + m' - j + 1} \min\{1, r[(\phi \setminus \{(\mathbf{c}_1, b_1, \gamma_1), \dots, (\mathbf{c}_j, b_j, \gamma_j)\}) \cup \{(\mathbf{v}_1, u_1, \sigma_1), \dots, (\mathbf{v}_{m'}, u_{m'}, \sigma_{m'})\}, (\mathbf{c}_j, b_j, \gamma_j)]^{-1}\} \tag{41}$$

$$\times \prod_{i=1}^{N-k} a(1 - \min\{1, r[\{(\mathbf{v}_1, u_1, \sigma_1), \dots, (\mathbf{v}_{m'}, u_{m'}, \sigma_{m'})\}, (\mathbf{v}'_i, u'_i, \sigma'_i)]\}) \tag{42}$$

$$\times \mathbf{1}(\{(\mathbf{v}_1, u_1, \sigma_1), \dots, (\mathbf{v}_{m'}, u_{m'}, \sigma_{m'})\} \in F) \times \frac{d\mu(\mathbf{v}_1, u_1, \sigma_1)}{\mu(D)} \dots \frac{d\mu(\mathbf{v}_{m'}, u_{m'}, \sigma_{m'})}{\mu(D)} \frac{d\mu(\mathbf{v}'_1, u'_1, \sigma'_1)}{\mu(D)} \dots \frac{d\mu(\mathbf{v}'_{N-k}, u'_{N-k}, \sigma'_{N-k})}{\mu(D)},$$

corresponding to first adding  $(\mathbf{v}_1, u_1, \sigma_1), \dots, (\mathbf{v}_{m'}, u_{m'}, \sigma_{m'})$  to  $\phi$ ; next, deleting the points in  $\phi$ ; and, finally, making no changes when  $N - k$  births are proposed. From (34) and the last inequality in (38), we obtain, for each of the terms in line (41) of the previous expression,  $r[\dots]^{-1} \geq (q_1/(1 - q_1))(m' + 1)/M\mu(D) \geq 1$ , i.e.  $\min\{1, r[\dots]^{-1}\} = 1, j = 1, \dots, k$ , and, for each of the terms in line (42),  $r[\dots] \leq ((1 - q_1)/q_1)M\mu(D)/(m' + 1) \leq 1$ , i.e.  $1 - \min\{1, r[\dots]\} = 1 - ((1 - q_1)/q_1)M\mu(D)/(m' + 1), i = 1, \dots, N - k$ . Furthermore, combining (34), condition (b), and the first inequality in (38), we obtain, for each of the terms in line (40),

$$r[\dots] \geq \frac{(1 - q_1) \exp(-L_{m'+N}/\rho_{m'+N}^d)\mu(D)}{q_1(k + l)} h_{m'+N}(\mathbf{v}_l, u_l, \sigma_l)$$

if  $\sigma_l < L_{m'+N}$  and  $u_l \in B_{m'+N}, l = 1, \dots, m'$ . Therefore,

$$P^{m'+N}(\phi, F) \geq a^{m'+N} \left(\frac{1 - q_1}{q_1}\right)^{m'} \frac{[\exp(-L_{m'+N}/\rho_{m'+N}^d)\mu(D)]^{m'}}{[\prod_{l=1}^{m'} (k + l)][\prod_{j=1}^k (m' + j)]} \times \left[1 - \frac{(1 - q_1)M\mu(D)}{q_1(m' + 1)}\right]^{N-k} Q(F),$$

where

$$\begin{aligned}
 Q(F) &= \int_{\{(\mathbf{v}_1, u_1, \sigma_1) \in D: u_1 \in B_{m'+N}, \sigma_1 < L_{m'+N}\}} \cdots \int_{\{(\mathbf{v}_{m'}, u_{m'}, \sigma_{m'}) \in D: u_{m'} \in B_{m'+N}, \sigma_{m'} < L_{m'+N}\}} \\
 &\quad \times \mathbf{1}(\{(\mathbf{v}_1, u_1, \sigma_1), \dots, (\mathbf{v}_{m'}, u_{m'}, \sigma_{m'})\} \in F) \\
 &\quad \times h_{m'+N}(\mathbf{v}_1, u_1, \sigma_1) \cdots h_{m'+N}(\mathbf{v}_{m'}, u_{m'}, \sigma_{m'}) \\
 &\quad \times \frac{d\mu(\mathbf{v}_1, u_1, \sigma_1)}{\mu(D)} \cdots \frac{d\mu(\mathbf{v}_{m'}, u_{m'}, \sigma_{m'})}{\mu(D)}.
 \end{aligned}$$

Consequently,  $P^{m'+N}(\phi, F) \geq \varepsilon' Q(F)$ , where

$$\begin{aligned}
 \varepsilon' &= a^{m'+N} \left( \frac{1 - q_1}{q_1} \right)^{m'} \frac{[\exp(-L_{m'+N}/\rho_{m'+N}^d)\mu(D)]^{m'}}{(N + 1) \cdots (N + m')} \\
 &\quad \times \frac{m'!}{(N + m')!} \left[ 1 - \frac{(1 - q_1)M\mu(D)}{q_1(m' + 1)} \right]^N.
 \end{aligned}$$

Notice that  $Q$  is a nonzero measure because, if we set  $x_i = \mathbf{x}_m$  for  $i \geq m$ , then

$$\begin{aligned}
 Q(\Omega_x) &\geq \int_D \cdots \int_D \mathbf{1}(k_{u_i}(\mathbf{v}_i, \mathbf{x}_i) > 0, u_i \in B_{m'+N}, \sigma_i < L_{m'+N}, i = 1, \dots, m') \\
 &\quad \times h_{m'+N}(\mathbf{v}_1, u_1, \sigma_1) \cdots h_{m'+N}(\mathbf{v}_{m'}, u_{m'}, \sigma_{m'}) \frac{d\mu(\mathbf{v}_1, u_1, \sigma_1)}{\mu(D)} \cdots \frac{d\mu(\mathbf{v}_{m'}, u_{m'}, \sigma_{m'})}{\mu(D)},
 \end{aligned}$$

which is positive – see (39). Thus,  $S_N$  is a small set.

We can now verify the following geometric drift condition, which implies  $V$ -uniform ergodicity: for each  $\beta > 1$ , there exist constants  $b < \infty$  and  $b' < 1$  such that, for any  $\phi \in \Omega_x$ ,

$$E(\beta^{\text{card}(X_1)} \mid X_0 = \phi) \leq b' \beta^{\text{card}(\phi)} + b \mathbf{1}(\phi \in S_N), \tag{43}$$

where  $\{X_n\}_{n \geq 0}$  denotes the Metropolis–Hastings chain and  $N > \beta M \mu(D)$ . The proof follows along lines similar to those of the proofs of Proposition 3.3 of [14], Proposition 7.14 of [34], and Proposition 6, below. Finally, (43) implies  $V$ -uniform ergodicity; cf. Proposition 7.9 of [34].

**Example 10.** In applications,  $W_{\text{ext}}$  is bounded and it does not make sense to consider arbitrary, large bandwidths, so let us suppose that both  $W_{\text{ext}}$  and the interval  $B$  are bounded, and let  $\varepsilon > 0$ . For technical reasons, we also assume that  $W_{\text{ext}}$  and  $B$  are both closed. Furthermore, suppose that  $\{c_j\}$  has density  $\tilde{p}$  with respect to the unit-rate Poisson process on  $W_{\text{ext}}$ , and is also independent of the  $(b_j, \gamma_j)$ , which are i.i.d. with density  $q$  (this setting covers, e.g. the finite Markov point process considered in Example 4). Let the Poisson process  $\nu$  be specified by the intensity measure

$$d\mu(\mathbf{c}, b, \gamma) = q(b, \gamma) \, d\mathbf{c} \, db \, d\gamma.$$

Then,  $\Phi$  has the following density with respect to  $\nu$ :

$$p(\{(\mathbf{c}_1, b_1, \gamma_1), \dots, (\mathbf{c}_n, b_n, \gamma_n)\}) = \tilde{p}(\{\mathbf{c}_1, \dots, \mathbf{c}_n\}).$$

Conditions (c) to (e) are then satisfied if  $k_1$  is positive and continuous on  $W_{\text{ext}} \times W_{\text{ext}}$ , e.g. if  $k_1$  is the Gaussian kernel (3). If, on the other hand,  $k_1$  is the uniform kernel (4), then condition (c)

is satisfied if the diameter of  $W_{\text{ext}}$  is less than or equal to the left endpoint of the interval  $B$ ; condition (e) still holds; and condition (d) is satisfied provided that

$$\int_{W_{\text{ext}}} \int_B \int_{\varepsilon}^{\infty} \mathbf{1}(\|c - x_i\| < b) q(b, \gamma) \, dc \, db \, d\gamma > 0, \quad i = 1, \dots, m.$$

The latter condition may be expected to hold for almost any natural choice of  $W_{\text{ext}}$ ,  $B$ ,  $\varepsilon$ , and  $q$ . Suppose, also, that  $\tilde{p}$  is a Markov density, i.e.

$$\tilde{p}(c) = \prod_{\eta \subseteq c} \varphi(\eta)$$

for all finite  $c \subset W_{\text{ext}}$ , where  $\varphi$  is an interaction function; cf. (14). Then, condition (a) is satisfied for most choices of  $\varphi$  used in practice (including repulsive models and the examples considered below) – see Chapter 6 of [34]. However, condition (b) is usually not satisfied: for instance, if we impose a hard-core condition, so that  $\tilde{p}(c) = 0$  whenever  $c$  contains two points separated by a distance less than the diameter of the hard core, then condition (b) is violated. Two exceptions are the Strauss process [20], [45] and the area-interaction process [3], [50]. For a Strauss process,

$$\tilde{p}(\{c_1, \dots, c_n\}) \propto \beta^n \theta^{\sum_{1 \leq i < j \leq n} \mathbf{1}(\|c_i - c_j\| \leq R)},$$

where  $\beta > 0$ ,  $0 < \theta \leq 1$ , and  $R > 0$  are parameters. Then, condition (a) is satisfied with  $\Lambda = \beta$ , and condition (b) holds with  $h_K(c, b, \gamma) = \beta \theta^{K(K+1)/2}$ . For an area-interaction process,

$$\tilde{p}(\{c_1, \dots, c_n\}) \propto \beta^n \theta^{-|\cup_{i=1}^n b(c_i, R)|},$$

where  $\beta > 0$ ,  $\theta > 0$ , and  $R > 0$  are parameters and  $b(c_i, R)$  is the ball in  $\mathbb{R}^d$  with centre  $c_i$  and radius  $R$ . Then, conditions (a) and (b) are satisfied with

$$\Lambda = \beta \quad \text{and} \quad h_K(c, b, \gamma) = \beta \theta^{-|b(\mathbf{0}, R)|} \quad \text{if } \theta \geq 1,$$

and with

$$\Lambda = \beta \theta^{-|b(\mathbf{0}, R)|} \quad \text{and} \quad h_K(c, b, \gamma) = \beta \quad \text{if } \theta \leq 1.$$

The following proposition concerns the case  $q_2 < 1$  when certain conditions, including the following, are satisfied.

- (f) The density  $p(\phi)$  is constant and positive for all  $\phi \in \Omega_x$  with  $\text{card}(\phi) = 1$ .
- (g)  $D = W_{\text{ext}} \times B \times (\varepsilon, L)$  with  $W_{\text{ext}}$ ,  $B$ , and  $\varepsilon$  defined as before, and where  $L > \varepsilon$  is finite.

Note that the conditions in Proposition 6 imply irreducibility, and we do not need to assume that condition (b) is satisfied. (As noted in Example 10, condition (b) is violated in most Markov models.) Moreover, condition (d) is implied by (37), since  $\mu(D) > 0$ , but we will not use it in the proof.

**Proposition 6.** *Assume that  $q_2 < 1$ , that  $\varepsilon > 0$ , and that conditions (37), (a), (c), (e), (f), and (g) are satisfied. Then, for any function  $V(\phi) = \beta^{\text{card}(\phi)}$  with  $\phi \in \Omega_x$  and  $\beta > 1$ , the birth–death–move algorithm is  $V$ -uniformly ergodic.*

*Proof.* As in the proof of Proposition 5, the first step is to show that  $S_N$  is a small set for any positive integer  $N$ . Let  $m' > N$  be an integer,  $a = \min\{q_1q_2, (1 - q_1)q_2, 1 - q_2\}$ ,

$$\phi = \{(c_1, b_1, \gamma_1), (c_2, b_2, \gamma_2), \dots, (c_k, b_k, \gamma_k)\} \in S_N,$$

$(c'_0, b'_0, \gamma'_0) = (c_k, b_k, \gamma_k)$ , and  $F \subseteq \Omega_x$ . Then,

$$\begin{aligned} P^{m'}(\phi, F) &\geq a^{m'} \frac{1}{k!} \prod_{j=1}^{k-1} \min\{1, r[\phi \setminus \{(c_1, b_1, \gamma_1), \dots, (c_j, b_j, \gamma_j)\}, (c_j, b_j, \gamma_j)]^{-1}\} \\ &\quad \times \int_{D^{m'-k+1}} \prod_{j=1}^{m'-k+1} \min\left\{1, \frac{\pi(\{(c'_j, b'_j, \gamma'_j)\} | x)}{\pi(\{(c'_{j-1}, b'_{j-1}, \gamma'_{j-1})\} | x)}\right\} \\ &\quad \times \mathbf{1}(\{(c'_{m'-k+1}, b'_{m'-k+1}, \gamma'_{m'-k+1})\} \in F) \\ &\quad \times \frac{d\mu(c'_1, b'_1, \gamma'_1)}{\mu(D)} \dots \frac{d\mu(c'_{m'-k+1}, b'_{m'-k+1}, \gamma'_{m'-k+1})}{\mu(D)}, \end{aligned}$$

corresponding to first deleting all except one point of  $\phi$  and then applying the move step  $m' - k + 1$  times. (Note that condition (b) was used in the proof of Proposition 5, when we considered the  $m'$  additions to  $\phi$ ; in the present proof, condition (b) is not needed because we do not consider any births.) A straightforward computation shows that, for any  $j = 1, \dots, k - 1$ ,

$$\begin{aligned} &\min\{1, r[\phi \setminus \{(c_1, b_1, \gamma_1), \dots, (c_j, b_j, \gamma_j)\}, (c_j, b_j, \gamma_j)]^{-1}\} \\ &\geq \min\left\{1, \frac{q_1}{(1 - q_1)M\mu(D)}\right\} > 0, \end{aligned}$$

where  $M > 0$  is the upper bound in (38) (which was also obtained without using condition (b)). By conditions (37), (e), (f), and (g), for all  $(c, b, \gamma), (c', b', \gamma') \in D$ ,

$$\begin{aligned} \frac{\pi(\{(c, b, \gamma)\} | x)}{\pi(\{(c', b', \gamma')\} | x)} &= \frac{\exp(-\int_W \gamma k_b(c, \xi) d\xi)}{\exp(-\int_W \gamma' k_{b'}(c', \xi) d\xi)} \left(\frac{\gamma}{\gamma'}\right)^m \prod_{j=1}^m \left[\frac{k_b(c, x_j)}{k_{b'}(c', x_j)}\right] \\ &\geq \exp\left(-\frac{\gamma}{b^d}\right) \left(\frac{\gamma \delta'}{\gamma' \delta''}\right)^m \geq \exp\left(-\frac{\gamma}{b^d}\right) \left(\frac{\varepsilon \delta'}{L \delta''}\right)^m, \end{aligned}$$

where  $\delta' = \min\{\delta'_1, \dots, \delta'_m\}$  and  $\delta'' = \max\{\delta''_1, \dots, \delta''_m\}$ . Therefore,

$$\begin{aligned} P^{m'}(\phi, F) &\geq \frac{a^{m'}}{k!} \left[\min\left\{1, \frac{q_1}{(1 - q_1)M\mu(D)}\right\}\right]^{k-1} \\ &\quad \times \int_D \dots \int_D \prod_{j=1}^{m'-k+1} \min\left\{1, \exp\left(-\frac{\gamma'_j}{b_j{}^d}\right) \left(\frac{\varepsilon \delta'}{L \delta''}\right)^m\right\} \\ &\quad \times \mathbf{1}(\{(c'_{m'-k+1}, b'_{m'-k+1}, \gamma'_{m'-k+1})\} \in F) \\ &\quad \times \frac{d\mu(c'_1, b'_1, \gamma'_1)}{\mu(D)} \dots \frac{d\mu(c'_{m'-k+1}, b'_{m'-k+1}, \gamma'_{m'-k+1})}{\mu(D)}. \end{aligned}$$

Since  $\mu(D) > 0$ , there exists  $B' \subseteq B$  such that  $\rho := \inf B' \in (0, \infty)$  and  $\mu(W_{\text{ext}} \times B' \times (\varepsilon, L)) > 0$ . Thus,

$$\begin{aligned}
 P^{m'}(\phi, F) &\geq \frac{a^{m'}}{k!} \left[ \min \left\{ 1, \frac{q_1}{(1 - q_1)M\mu(D)} \right\} \right]^{k-1} \\
 &\quad \times \int_D \cdots \int_D \prod_{j=1}^{m'-k+1} \min \left\{ 1, \exp \left( -\frac{\gamma'_j}{b_j{}^d} \right) \left( \frac{\varepsilon \delta'}{L\delta''} \right)^m \right\} \\
 &\quad \times \mathbf{1}(b'_1 \in B', \dots, b'_{m'-k+1} \in B', \{(c'_{m'-k+1}, b'_{m'-k+1}, \gamma'_{m'-k+1})\} \in F) \\
 &\quad \times \frac{d\mu(c'_1, b'_1, \gamma'_1)}{\mu(D)} \cdots \frac{d\mu(c'_{m'-k+1}, b'_{m'-k+1}, \gamma'_{m'-k+1})}{\mu(D)} \\
 &\geq \frac{a^{m'}}{k!} \left[ \min \left\{ 1, \frac{q_1}{(1 - q_1)M\mu(D)} \right\} \right]^{k-1} \\
 &\quad \times \left[ \min \left\{ 1, \exp \left( -\frac{L}{\rho_j^d} \right) \left( \frac{\varepsilon \delta'}{L\delta''} \right)^m \right\} \right]^{m'-k+1} \left( \frac{\mu(D')}{\mu(D)} \right)^{m'-k} Q(F),
 \end{aligned}$$

where  $D' = W_{\text{ext}} \times B' \times (\varepsilon, L)$  and

$$Q(F) = \int_{D'} \mathbf{1}(\{(v, u, \sigma)\} \in F) \frac{d\mu(v, u, \sigma)}{\mu(D)}.$$

Note that  $Q$  is a nonzero measure since, by (37) and condition (f),  $Q(\Omega_x) = \mu(D')/\mu(D) > 0$ . Furthermore,  $P^{m'}(\phi, F) \geq \varepsilon' Q(F)$ , where

$$\begin{aligned}
 \varepsilon' &= \frac{a^{m'}}{N!} \left[ \min \left\{ 1, \frac{q_1}{(1 - q_1)M\mu(D)} \right\} \right]^{N-1} \\
 &\quad \times \left[ \min \left\{ 1, \exp \left( -\frac{L}{\rho^d} \right) \left( \frac{\varepsilon \delta'}{L\delta''} \right)^m \right\} \right]^{m'+1} \left( \frac{\mu(D')}{\mu(D)} \right)^{m'} > 0.
 \end{aligned}$$

Thus,  $S_N$  is a small set.

Next, we establish  $V$ -uniform ergodicity by checking the geometric drift condition (43) with  $N > \max\{\beta, (1 - q_1)/q_1\}M\mu(D)$ . Assume that the current state  $\phi$  of the Metropolis–Hastings chain is such that  $\phi \notin S_N$ , and set  $k = \text{card}(\phi)$ . As for the birth–death Metropolis–Hastings algorithm, here the probability of accepting a death is again 1. Now, letting

$$B(\phi) = \int_D \min\{1, r[\phi, (c, b, \gamma)]\} \frac{d\mu(c, b, \gamma)}{\mu(D)}$$

denote the probability of accepting a birth, we have

$$B(\phi) \leq \frac{(1 - q_1)M\mu(D)}{q_1(k + 1)} \leq \frac{(1 - q_1)M\mu(D)}{q_1(N + 1)},$$

since  $\phi \notin S_N$ . Therefore,

$$\begin{aligned}
 &E(\beta^{\text{card}(X_1)} \mid X_0 = \phi) \\
 &= q_1 q_2 \beta^{k+1} B(\phi) + q_1 q_2 \beta^k (1 - B(\phi)) + (1 - q_1) q_2 \beta^{k-1} + (1 - q_2) \beta^k \\
 &\leq \beta^k \left[ (1 - q_1) q_2 (\beta - 1) \frac{M\mu(D)}{N + 1} + (1 - q_1) q_2 \left( \frac{1}{\beta} - 1 \right) + 1 \right], \tag{44}
 \end{aligned}$$

where we have used the fact that  $\beta > 1$ . Therefore, when  $\phi \notin S_N$ , (43) follows from setting  $b'$  equal to the contents of the square brackets in (44). Indeed, it is clear that  $b' < 1$ , since  $N > \max\{\beta, (1 - q_1)/q_1\}M\mu(D)$ . Finally, when  $\phi \in S_N$ ,

$$E(\beta^{\text{card}(X_1)} \mid X_0 = \phi) \leq \beta^{\text{card}(\phi)+1} \leq \beta^{N+1},$$

whereby (43) is verified.

**Example 11.** Condition (37) is satisfied if  $k_1$  is positive, but it is easily violated by the uniform kernel. If the situation is as in Example 10, then condition (c) is satisfied if  $k_1$  is continuous and positive. Furthermore, condition (f) means that  $\varphi(\eta)$  is a positive constant  $\beta$  whenever  $\text{card}(\eta) = 1$ . If this is not the case, condition (f) can still be obtained by redefining the intensity measure of  $\nu$  as  $d\mu(\mathbf{c}, b, \gamma) = \varphi(\{\mathbf{c}\})q(b, \gamma) d\mathbf{c} db d\gamma$ . Finally, condition (g) is needed for technical reasons (cf. the proof above), and the existence of the upper bound  $L$  on the  $\gamma_j$  may, in many applications, be a weaker assumption.

## 5. Concluding remarks

Although we have demonstrated that, to some extent, GSNCPs possess many appealing properties, it remains to investigate their statistical aspects and practical uses. The results in Section 3.1 for the intensity and pair correlation function may be useful for model checking and parameter estimation, particularly minimum contrast estimation; see [34]. The results in Section 3.2 on reduced Palm distributions and  $J$ -functions may, to some extent, be useful for statistical applications [34], [47], [48]. The most important part of this paper from a statistical viewpoint is probably Section 4 on simulation algorithms, since they provide ways of performing simulation-based inference, as mentioned in Section 2.1; see [34]. The algorithms in Sections 4.1 and 4.2 for simulation of GSNCPs may be useful in connection to model checking, while the conditional simulation algorithm in Section 4.3 has applications in both Bayesian and likelihood inference [31].

In this paper, we have concentrated on the probabilistic aspects of GSNCPs. Future research should address the following: (i) Markov properties of GSNCPs; cf. the discussion in [31] for SNCPs; (ii) implementation and experimentation with the simulation algorithms discussed in Section 4; (iii) convergence properties of MCMC hybrid algorithms (or Gibbs within Metropolis algorithms – see, e.g. [11]) for conditional simulation, when we extend the setting of Section 4.3 by imposing a prior distribution on the hyperparameters of the distribution of  $\Phi$  (e.g. imposing a prior on  $\theta$  in Example 3); and (iv) application of our current understanding of GSNCPs to related models, e.g. spatial Hawkes processes (see point (e) in Section 2.1).

## Acknowledgements

This research was supported by the Danish Natural Science Research Council, the Network in Mathematical Physics and Stochastics (MaPhySto; funded by a grant from the Danish National Research Foundation), and the Murst Project ‘Processi Stocastici e Applicazioni a Filtraggio, Controllo, Simulazione e Finanza Matematica’.

## References

- [1] ALLEN, M. P. AND TILDESLEY, D. J. (1987). *Computer Simulation of Liquids*. Oxford University Press.
- [2] BADDELEY, A. AND MØLLER, J. (1989). Nearest-neighbour Markov point processes and random sets. *Internat. Statist. Rev.* **2**, 89–121.

- [3] BADDELEY, A. J. AND VAN LIESHOUT, M. N. M. (1995). Area-interaction point processes. *Ann. Inst. Statist. Math.* **46**, 601–619.
- [4] BRIX, A. (1999). Generalized gamma measures and shot-noise Cox processes. *Adv. Appl. Prob.* **31**, 929–953.
- [5] BRIX, A. AND CHADŒUF, J. (2002). Spatio-temporal modeling of weeds by shot-noise G Cox processes. *Biometrical J.* **44**, 83–99.
- [6] BRIX, A. AND DIGGLE, P. J. (2001). Spatio-temporal prediction for log-Gaussian Cox processes. *J. R. Statist. Soc. B* **63**, 823–841.
- [7] BRIX, A. AND KENDALL, W. S. (2002). Simulation of cluster point processes without edge effects. *Adv. Appl. Prob.* **34**, 267–280.
- [8] BRIX, A. AND MØLLER, J. (2001). Space-time multi type log Gaussian Cox processes with a view to modelling weeds. *Scand. J. Statist.* **28**, 471–488.
- [9] COX, D. R. AND ISHAM, V. (1980). *Point Processes*. Chapman and Hall, London.
- [10] DALEY, D. J. AND VERE-JONES, D. (2003). *An Introduction to the Theory of Point Processes, Elementary Theory and Methods*, Vol. 1, 2nd edn. Springer, New York.
- [11] DELLAPORTAS, P. AND ROBERTS, G. O. (2003). An introduction to MCMC. In *Spatial Statistics and Computational Methods* (Lecture Notes Statist. **173**), ed. J. Møller. Springer, New York, pp. 1–41.
- [12] DIGGLE, P. J. (2003). *Statistical Analysis of Spatial Point Patterns*, 2nd edn. Arnold, London.
- [13] GEORGI, H.-O. (1976). Canonical and grand canonical Gibbs states for continuum systems. *Commun. Math. Phys.* **48**, 31–51.
- [14] GEYER, C. J. (1999). Likelihood inference for spatial point processes. In *Stochastic Geometry: Likelihood and Computation*, eds O. E. Barndorff-Nielsen, W. S. Kendall and M. N. M. van Lieshout, Chapman and Hall/CRC, Boca Raton, FL, pp. 79–140.
- [15] GEYER, C. J. AND MØLLER, J. (1994). Simulation procedures and likelihood inference for spatial point processes. *Scand. J. Statist.* **21**, 359–373.
- [16] GREEN, P. J. (1995). Reversible jump MCMC computation and Bayesian model determination. *Biometrika* **82**, 711–732.
- [17] HAWKES, A. G. (1971). Spectra of some self-exciting and mutually exciting point processes. *Biometrika* **58**, 83–90.
- [18] HAWKES, A. G. AND OAKES, D. (1974). A cluster representation of a self-exciting process. *J. Appl. Prob.* **11**, 493–503.
- [19] KARR, A. F. (1991). *Point Processes and Their Statistical Inference*. Marcel Dekker, New York.
- [20] KELLY, F. P. AND RIPLEY, B. D. (1976). A note on Strauss's model for clustering. *Biometrika* **63**, 357–360.
- [21] KENDALL, W. S. AND MØLLER, J. (2000). Perfect simulation using dominating processes on ordered spaces, with application to locally stable point processes. *Adv. Appl. Prob.* **32**, 844–865.
- [22] KINGMAN, J. F. C. (1977). Remarks on the spatial distribution of a reproducing population. *J. Appl. Prob.* **14**, 577–583.
- [23] LINDVALL, T. (1992). *Lectures on the Coupling Method*. John Wiley, New York.
- [24] MCKEAGUE, I. W. AND LOIZEAUX, M. A. (2002). Perfect sampling for point process cluster modelling. In *Spatial Cluster Modelling*, eds A. B. Lawson and D. Denison, Chapman and Hall/CRC, Boca Raton, FL, pp. 87–107.
- [25] MATÉRN, B. (1960). *Spatial Variation*, Meddelanden från Statens Skogsforskningsinstitut, Stockholm.
- [26] MATÉRN, B. (1986). *Spatial Variation* (Lecture Notes Statist. **36**), 2nd edn. Springer, Berlin.
- [27] MEYN, S. P. AND TWEEDIE, R. L. (1993). *Markov Chains and Stochastic Stability*. Springer, London.
- [28] MØLLER, J. (1994). Contribution to the discussion of N. L. Hjort and H. Omre (1994): Topics in spatial statistics. *Scand. J. Statist.* **21**, 346–349.
- [29] MØLLER, J. (1999). Markov chain Monte Carlo and spatial point processes. In *Stochastic Geometry: Likelihood and Computation*, eds O. E. Barndorff-Nielsen, W. S. Kendall and M. N. M. van Lieshout (Monogr. Statist. Appl. Prob. **80**), Chapman and Hall/CRC, Boca Raton, FL, pp. 141–172.
- [30] MØLLER, J. (2003). A comparison of spatial point process models in epidemiological applications. In *Highly Structured Stochastic Systems*, eds P. J. Green, N. L. Hjort and S. Richardson, Oxford University Press, pp. 264–268.
- [31] MØLLER, J. (2003). Shot noise Cox processes. *Adv. Appl. Prob.* **35**, 614–640.
- [32] MØLLER, J. AND WAAGEPETERSEN, R. P. (2002). Statistical inference for Cox processes. In *Spatial Cluster Modelling*, eds A. B. Lawson and D. Denison, Chapman and Hall/CRC, Boca Raton, FL, pp. 37–60.
- [33] MØLLER, J. AND WAAGEPETERSEN, R. P. (2003). An introduction to simulation-based inference for spatial point processes. In *Spatial Statistics and Computational Methods* (Lecture Notes Statist. **173**), ed. J. Møller, Springer, New York, pp. 143–198.
- [34] MØLLER, J. AND WAAGEPETERSEN, R. P. (2003). *Statistical Inference and Simulation for Spatial Point Processes*. Chapman and Hall/CRC, Boca Raton, FL.
- [35] MØLLER, J., SYVERVEEN, A. R. AND WAAGEPETERSEN, R. P. (1998). Log Gaussian Cox processes. *Scand. J. Statist.* **25**, 451–482.

- [36] MÜLLER, A. AND STOYAN, D. (2002). *Comparison Methods for Stochastic Models and Risks*. John Wiley, Chichester.
- [37] NEYMAN, J. AND SCOTT, E. L. (1958). Statistical approach to problems of cosmology. *J. R. Statist. Soc. B* **20**, 1–43.
- [38] NGUYEN, X. X. AND ZESSIN, H. (1979). Integral and differential characterizations of Gibbs processes. *Math. Nachr.* **88**, 105–115.
- [39] PRESTON, C. (1976). *Random Fields* (Lecture Notes Math. **534**). Springer, Berlin.
- [40] RICHARDSON, S. (2003). Spatial models in epidemiological applications. In *Highly Structured Stochastic Systems*, eds P. J. Green, N. L. Hjort and S. Richardson, Oxford University Press, pp. 237–259.
- [41] RIPLEY, B. D. (1977). Modelling spatial patterns (with discussion). *J. R. Statist. Soc. B* **39**, 172–212.
- [42] RIPLEY, B. D. AND KELLY, F. P. (1977). Markov point processes. *J. Lond. Math. Soc.* **15**, 188–192.
- [43] RUELLE, D. (1969). *Statistical Mechanics: Rigorous Results*. W. A. Benjamin, Reading, MA.
- [44] STOYAN, D., KENDALL, W. S. AND MECKE, J. (1995). *Stochastic Geometry and Its Applications*, 2nd edn. John Wiley, Chichester.
- [45] STRAUSS, D. J. (1975). A model for clustering. *Biometrika* **63**, 467–475.
- [46] THOMAS, M. (1949). A generalization of Poisson's binomial limit for use in ecology. *Biometrika* **36**, 18–25.
- [47] VAN LIESHOUT, M. N. M. (2000). *Markov Point Processes and Their Applications*. Imperial College Press, London.
- [48] VAN LIESHOUT, M. N. M. AND BADDELEY, A. J. (1996). A nonparametric measure of spatial interaction in point patterns. *Statistica Neerlandica* **50**, 344–361.
- [49] VAN LIESHOUT, M. N. M. AND BADDELEY, A. J. (2002). Extrapolating and interpolating spatial patterns. In *Spatial Cluster Modelling*, eds A. B. Lawson and D. Denison, Chapman and Hall/CRC, Boca Raton, FL, pp. 61–86.
- [50] WIDOM, B. AND ROWLINSON, J. S. (1970). A new model for the study of liquid–vapor phase transitions. *J. Chem. Phys.* **52**, 1670–1684.
- [51] WOLPERT, R. L. AND ICKSTADT, K. (1998). Poisson/gamma random field models for spatial statistics. *Biometrika* **85**, 251–267.