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

Jesper Møller and Andreas D. Christoffersen

Pair correlation functions and limiting distributions of iterated cluster point processes

No. 11, November 2017

Pair correlation functions and limiting distributions of iterated cluster point processes

Jesper Møller and Andreas D. Christoffersen

Department of Mathematical Sciences, Aalborg University

Abstract

We consider a Markov chain of point processes such that each state is a super position of an independent cluster process with the previous state as its centre process together with some independent noise process. The model extends earlier work by Felsenstein and Shimatani describing a reproducing population. We discuss when closed term expressions of the first and second order moments are available for a given state. In a special case it is known that the pair correlation function for these type of point processes converges as the Markov chain progresses, but it has not been shown whether the Markov chain has an equilibrium distribution with this, particular, pair correlation function and how it may be constructed. Assuming the same reproducing system, we construct an equilibrium distribution by a coupling argument.

Keywords: Coupling; equilibrium; independent clustering; Markov chain; pair correlation function; reproducing population; weighted determinantal and permanental point processes.

1 Introduction

This paper deals with a discrete time Markov chain of point processes $G_0, G_1, \dots$ in the d-dimensional Euclidean space $\mathbb{R}^d$, where the chain describes a reproducing population and we refer to G_n as the nth generation (of points). We make the following assumptions. Any point process considered in this paper will be viewed as a random subsets of $\mathbb{R}^d$ which is almost surely locally finite, that is, the point process has almost surely a finite number of points within any bounded subset of $\mathbb{R}^d$ (for measure theoretical details, see e.g. Daley and Vere-Jones (2003) or Møller and Waagepetersen (2004)). Recall that a point process $X \subset \mathbb{R}^d$ is stationary if its distribution is invariant under translations in $\mathbb{R}^d$, and then its intensity $\rho_X \in [0, \infty]$ is the mean number of points in X falling in any Borel subset of $\mathbb{R}^d$ of unit volume. Now, for generation 0, G_0 is stationary with intensity $\rho_{G_0} \in (0, \infty)$. Further, for generation $n = 1, 2, \dots$, conditional on the previous generations $G_0, \dots, G_{n-1}$, we obtain G_n by three basic operations for point processes:

(a) Independent clustering: To each point $x \in G_{n-1}$ is associated a (non-centred) cluster $Y_{n,x} \subset \mathbb{R}^d$. These clusters are independent identically distributed (IID) finite point processes and they are independent of $G_0, \ldots, G_{n-1}$. The cardinality of $Y_{n,x}$ has finite mean β_n and finite variance ν_n and is independent of the points in $Y_{n,x}$ which are IID, with each point following a probability density function (PDF) f_n. We refer to $x + Y_{n,x}$ (the translation of $Y_{n,x}$ by x) as the offspring/children process generated by the ancestor/parent x, and we let

$$Y_n = \bigcup_{x \in G_{n-1}} (x + Y_{n,x}) \tag{1.1}$$

be the independent cluster process given by the superposition of all offspring processes generated by the points in the previous generation G_{n-1}.

(b) Independent thinning: For all $y \in \mathbb{R}^d$, let $B_{n,y}$ be IID Bernoulli variables which are independent of Y_n, $G_0, \ldots, G_{n-1}$, and all previously generated Bernoulli variables. Let $p_n = \mathrm{P}(B_{n,y} = 1)$. For all $x \in G_{n-1}$, let

$$W_{n,x} = \{y \in x + Y_{n,x} : B_{n,y} = 1\}$$

be the independent p_n-thinned point process of $x + Y_{n,x}$, and let

$$W_n = \bigcup_{x \in G_{n-1}} W_{n,x} \tag{1.2}$$

be the independent p_n-thinned point process of Y_n.

(c) Independent noise: Let $Z_n \subset \mathbb{R}^d$ be a stationary point process with finite intensity ρ_{Z_n} and which is independent of $W_n, G_0, \ldots, G_{n-1}$. Finally, let

$$G_n = W_n \cup Z_n \tag{1.3}$$

where we interpret Z_n as noise.

Our model is an extension of the model in Shimatani (2010), which in turn is an extension of Malécot's model studied in Felsenstein (1975) (we return to this in Section 2, item (vii) and (viii)). In particular, our extension allows us to model cluster centres exhibiting clustering or regularity, and similarly the noise processes can be clustered or regular. For statistical applications, we have in mind that G_n may be observable (at least for some values of $n \geq 1$) whilst G_0 and the cluster, thinning, and superpositioning procedures in item (a)–(c) are unobservable. Our model may be of relevance for applications in population genetics and community ecology (see Shimatani (2010) and the references therein), for analyzing tropical rain forest point pattern data with multiple scales of clustering (see Wiegand et al. (2007)), and for modelling proteins with multiple noisy appearances in PhotoActivated Localization Microscopy (PALM) (see Andersen et al. (2017)). However, we leave it for other work to study the statistical applications of our model and results.

The paper is organized as follows. A discussion of the assumptions in items (a)–(c) and the related literature is given in Section 2. Section 3 focuses on the first and second order moment properties of G_n, that is, its intensity and pair correlation function (PCF); we extend results in Shimatani (2010) and show that tractable model cases for the PCF of G_0 (extending cases considered in Shimatani (2010)) are meaningful in terms of Poisson, weighted permanental, and weighted determinantal point processes (which was not observed in Shimatani (2010)). Section 4 discusses limiting cases of the PCF of G_n as $n \to \infty$ when we have the same reproduction system and under weaker conditions than in Shimatani (2010). In particular, when natural conditions are satisfied, we establish ergodicity of the Markov chain by using a coupling construction and by giving a constructive description of the chain's unique invariant distribution when extending the Markov chain backwards in time. Finally, Appendix A provides background knowledge on weighted permanental and determinantal point processes, Appendix B verifies some technical details, and Appendiks C specifies an algorithm for approximate simulation of the Markov chins invariant distribution.

2 Assumptions and related work

Items (i)–(iv) below comment on the model assumptions in items (a)–(c).

(i) The process Y_n is a stationary independent cluster process (Daley and Vere-Jones (2003)) and we have the following special cases: If G_{n-1} is a stationary Poisson process, Y_n is a Neyman-Scott process (Neyman and Scott (1958)). If in addition $\#Y_{n,x}$ follows a Poisson distribution, then $\beta_n = \nu_n$ and Y_n is a shot-noise Cox process (SNCP; Møller (2003)) driven by

$$\Lambda_n(x) = \beta_n \sum_{y \in G_{n-1}} f_n(x-y), \qquad x \in \mathbb{R}^d. \tag{2.1}$$

This is a (modified) Thomas process (Thomas (1949)) if f_n is the density of d IID zero-mean normally distributed variates with variance σ_n^2 – we denote this distribution by $N_d(\sigma_n^2)$ – and it is a Matérn cluster process (Matérn (1960, 1986)) if instead f_n is a uniform density of a d-dimensional ball with centre at the origin. However, in many applications a Poisson centre process is not appropriate. For instance, Van Lieshout & Baddeley (2002) considered a repulsive Markov point process model for the centre process, whereby it is easier to identify the clusters than under a Poisson centre process.

(ii) When $\beta_n \le \nu_n$, we may restrict attention to the general case of a stationary generalised shot-noise Cox process (GSNCP) as studied in Møller and Torrisi (2005): In this model (2.1) is extended to the case where G_{n-1} is a general stationary point process and Y_n is a Cox process driven by

$$\Lambda_n(x) = \sum_{y \in G_{n-1}} \gamma_y k_n(\{(x-y)/b_y\})/b_y^d, \qquad x \in \mathbb{R}^d, \tag{2.2}$$

where k_n is a PDF on $\mathbb{R}^d$, the γ_y and the b_y for all $y \in G_{n-1}$ are independent positive random variables which are independent of G_{n-1}, and the γ_y are identically distributed with mean β_n and variance $\nu_n - \beta_n$ (as $\#Y_{n,x}$ has mean β_n and variance $\nu_n = \mathrm{E}\{\mathrm{var}(\#Y_{n,x}|\gamma_y)\} + \mathrm{var}\{\mathrm{E}(\#Y_{n,x}|\gamma_y)\} = \beta_n + \mathrm{var}(\gamma_n)$). Further, b_y has an interpretation as a random band-width and $f_n(x) = \mathrm{E}\{k_n(x/b_y)/b_y^d\}$.

The general results for the intensity and PCF of G_n in Section 3 will be unchanged whether we consider this stationary GSNCP or the more general case in item (a)

(iii) Clearly, there is no noise (Z_n is empty with probability one) if $\rho_{Z_n} = 0$. The case $\rho_{Z_n} > 0$ may be relevant when not all points in a generation can be described as resulting from independent clustering and thinning. Note that in item (c) we could without loss of generality assume $Z_1, Z_2, \ldots$ are independent, however, it will first be in Section 4 that we assume they are IID. Further, we introduce the thinning of Y_n in item (b) only for modelling purposes; from a mathematical point of view the thinning could be omitted if in item (a) we replace each cluster $Y_{n,x}$ by what happens after the independent thinning: Namely that independent thinned clusters $Y_{n,x}^{\mathrm{th}}$ appear so that $\#Y_{n,x}^{\mathrm{th}}$ has mean $\beta_n^{\mathrm{th}} = \beta_n p_n$ and variance $\nu_n^{\mathrm{th}} = \beta_n p_n - \beta_n p_n^2 + \nu_n p_n^2$ and is independent of the points in $Y_{n,x}^{\mathrm{th}}$ which are IID with PDF f_n, whereby W_n and $Y_n^{\mathrm{th}} := \cup_{x \in G_{n-1}}(x + Y_{n,x}^{\mathrm{th}})$ are identically distributed.

(iv) Assuming no thinning ($p_n = 1$), an equivalent description of items (a) and (c) is given in terms of the Voronoi tessellation generated by G_{n-1}: For $x \in G_{n-1}$, let $C(x|G_{n-1})$ be the Voronoi cell associated to x and consisting of all points in $\mathbb{R}^d$ which are at least as close to x than to another other point in G_{n-1} (with respect to usual distance in $\mathbb{R}^d$). With probability one, since G_{n-1} is stationary and non-empty, each Voronoi cell is bounded and hence is volume is finite (see e.g. Møller (1989, 1994)). Thus we can set

$$G_n = \bigcup_{x \in G_{n-1}} (x + G_{n,x})$$

where conditional on G_{n-1}, for all $x \in G_{n-1}$, the $G_{n,x}$ are IID finite point processes with a distribution as follows: $\#G_{n,x}$ has mean $\beta_n + |C(x|G_{n-1})|\rho_{Z_n}$, variance $\nu_n + |C(x|G_{n-1})|\rho_{Z_n}$, and is independent of the points in $G_{n,x}$; and the points are IID, each following a mixture distribution so that with probability $\beta_n/(\beta_n + |C(x|G_{n-1})|\rho_{Z_n})$ the PDF is f_n and else it is a uniform distribution on $C(x|G_{n-1})$.

In items (v)–(vi) below we discuss earlier work on the model for $G_0, G_1, \ldots$, where G_0 is a stationary Poisson process, all $G_n = Y_n$ (no thinning and no noise) for $n \geq 1$, $f_n = f$ and $\beta_n = \beta$ do not depend on $n \geq 1$. We may refer to this as a replicated SNCP. Frequently in the literature, a so-called replicated Thomas process is considered, that is, $f \sim N_d(\sigma^2)$.

(v) Apperently this replicated SNCP was originally studied by Malécot, see the discussion and references in Felsenstein (1975) where the following three conditions are stated: "(I) individuals are distributed randomly on the line with equal expected density everywhere; (II) each individual reproduces independently, the number of offspring being drawn from a Poisson distribution with a mean of one; and (III) each offspring migrates independently, the displacements being drawn from some distribution m(x), which we will take to be a normal distribution." (In our notation, $d = 1$, $\beta = 1$, and $f \sim N_1(\sigma^2)$, but

Felsenstein (1975) considered also more general offspring densities f and the cases $d = 2, 3$.) Felsenstein (1975) showed that "(I) is incompatible with (II)–(III)" because $G_1, G_2, \ldots$ are not stationary Poisson processes and "a model embodying (II) and (III) will lead to the formation of larger and larger clumps of individuals separated by greater and greater distances", and then he concluded "This model is therefore biologically irrelevant".

(vi) Kingman (1977) considered the case where β is replaced by a non-negative function b which is allowed to depend on the cluster centre x and the previous generation, so a cluster with centre x is a Poisson process with intensity function $b(x, G_{n-1})f(\cdot - x)$; e.g., as in the Voronoi case discussed in item (iv), $b(x, G_{n-1})$ may depend on G_{n-1} in a neighbourhood of x. Then G_n is a Cox process: G_n conditional on G_{n-1} is a Poisson process with intensity function

$$\Lambda_n(x) = \sum_{y \in G_{n-1}} b(y, G_{n-1}) f(x - y), \qquad x \in \mathbb{R}^d. \tag{2.3}$$

In this setting Kingman (1977) verified that it is impossible for G_n to be a stationary Poisson process, however, replacing $f(x - y)$ in (2.3) by a more general density which may depend on $C_{n-1} - x$, Kingman (1977) noticed that it is possible for G_n to be a stationary Poisson process. A trivial example is the Voronoi case in item (iv) when $G_n = Z_n$ for $n \geq 1$.

Recently, Shimatani (2010) considered first the case of items (a)–(b) and no noise, when $d = 2$ and there is the same reproduction system so that $f_n = f$, $\beta_n = \beta > 0$, $\nu_n = \nu$, and $p_n = p \in (0, 1]$ do not depend on $n \geq 1$.

(vii) In particular, Shimatani (2010) considered the case $f \sim N_2(\sigma^2)$ and when $\beta p = 1$ or equivalently when the intensities $\rho_{G_0} = \rho_{G_1} = \ldots$ are invariant over generations, and then he showed that as $n \to \infty$, the PCF for G_n diverges and "all offspring will eventually have the same ancestor". It follows from item (iii) that the model is equivalent to a replicated Neyman-Scott process; this becomes a replicated Thomas process when each cluster size is Poisson distributed, and hence Shimatani (2010) result agrees with the results in Felsenstein (1975) and Kingman (1977). Note that Shimatani (2010) implicitly assumed that a cluster can have more than one point. Otherwise the PCF of G_n becomes equal to 1; we discuss this rather trivial case again in Section 3.2.2 and 4; see also Section 3 in Kingman (1977).

Next Shimatani (2010) extended the model by including noise as in item (c) and by making the following assumptions: The noise processes Z_n are stationary Poisson processes, satisfying $0 < \rho_{Z_1} = \rho_{Z_2} = \ldots$ and $\rho_{G_0} = \rho_{G_1} = \ldots$, meaning that $\beta p \leq 1$. As there is no noise if $\beta p = 1$ it is also assumed that $\beta p < 1$.

(viii) Then Shimatani (2010) showed that the PCF of G_n converges uniformly as $n \to \infty$ and he argued that this limiting case may be "biologically valid" (Shimatani, 2010, Section 2.4). However, there are some flaws in the paper by Shimatani (2010) which we address:

- He was not showing that there exists an underlying point process having this limiting case as its PCF, although he claimed that "this modified replicated Neyman-Scott process reaches an equilibrium state". In Section 4, for our more general model, we prove the existence of such an underlying point process.
- When G_0 is not a stationary Poisson process but its PCF is of a particular form (which we specify later in connection to (3.4)), he did not argue that there exists an underlying point process and what it could be. In Section 3, we verify this existence under our more general model.

Finally, we remark on a few related cases.

(ix) Whilst we study the processes G_n for all $n = 1, 2, \ldots$, often in the spatial point process literature the focus is on either G_1 or G_2, assuming $p_n = 1$ and $\rho_{Z_n} = 0$ for $n = 1$ or $n = 1, 2$, respectively. Wiegand et al. (2007) studied this in the special case of a double Thomas cluster process G_2 when $d = 2$, i.e., when G_0 is a stationary Poisson process, (2.1) holds for both $G_1 = Y_1$ and $G_2 = Y_2$, and $f_n \sim N_2(\sigma_n^2)$ for $n = 1, 2$; see also Andersen et al. (2017) for more general functions f_n. Moreover, Wiegand et al. (2007) extended the double Thomas process to the case where $\rho_{Z_1} = 0$ and $\rho_{Z_2} > 0$; this type of model is also considered in Andersen et al. (2017). In any case, our general results for intensities and PCFs in Section 3 will cover all these cases.

(x) Incidentally, when $p_1 = p_2 = \ldots = 1$, $\rho_{Z_1} = \rho_{Z_2} = \ldots = 0$, $\beta_1 = \beta_2 = \ldots$, and $f_1 = f_2 = \ldots$, the superposition $\bigcup_{n=0}^{\infty} G_n$ is known as a spatial Hawkes process, see Møller and Torrisi (2007) and the references therein.

3 First and second order moment properties

In this section we determine the intensity and the PCF of G_n for $n = 1, 2, \ldots$, under more general assumptions than in Shimatani (2010). Specifically, the noise is an arbitrary stationary point process (not necessarily a stationary Poisson process as in Shimatani (2010)) and we do not assume the same reproduction system.

3.1 Intensities

By induction G_n is seen to be stationary for $n = 0, 1, \ldots$ Its intensity is determined in the following proposition where for notational convenience we define $Z_0 = G_0$ so that $\rho_{Z_0} = \rho_{G_0}$.

Proposition 3.1. *For $n = 1, 2, \ldots$, we have that G_n is stationary with a positive and finite intensity given by*

$$\rho_{G_n} = \rho_{G_{n-1}} \beta_n p_n + \rho_{Z_n} = \rho_{Z_n} + \sum_{i=0}^{n-1} \rho_{Z_i} \prod_{j=i+1}^{n} \beta_j p_j. \tag{3.1}$$

Proof. Using induction for $n = 1, 2, \ldots$, the proposition follows immediately from items (a)–(c), where the term $\rho_{Z_i} \prod_{j=i+1}^{n} \beta_j p_j$ is the contribution to the intensity caused by the clusters with centres Z_i and after independent thinning. $\square$

3.2 Pair correlation functions

3.2.1 Preliminaries.

Recall that a stationary point process $X \subset \mathbb{R}^d$ with intensity $\rho_X \in (0,\infty)$ has a translation invariant PCF (pair correlation function) $(u,v) \mapsto g_X(u-v)$ with $(u,v) \in \mathbb{R}^d \times \mathbb{R}^d$ if for any bounded Borel function $h : \mathbb{R}^d \times \mathbb{R}^d \mapsto [0,\infty)$ with compact support,

$$\mathrm{E} \sum_{x_1,x_2 \in X:\, x_1 \neq x_2} h(x_1,x_2) = \rho_X^2 \iint h(x_1,x_2) g_X(x_1-x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2 < \infty. \tag{3.2}$$

Equivalently, for any bounded and disjoint Borel sets $A, B \subset \mathbb{R}^d$, denoting $N(A)$ the cardinality of $X \cap A$, the covariance between $N(A)$ and $N(B)$ exists and is given by

$$\mathrm{cov}\{N(A), N(B)\} = \rho_X^2 \int_A \int_B \{g_X(x_1-x_2) - 1\}\,\mathrm{d}x_1\,\mathrm{d}x_2.$$

Some remarks are in order. Note that g_X is uniquely determined except for nullsets with respect to Lebesgue measure on $\mathbb{R}^d$, but we ignore such nullsets in the following. Thus the translation invariance of the PCF is implied by the stationarity of X. Our results below are presented in terms of $g_X - 1$ rather than g_X, and $g_X = 1$ if X is a Poisson process. It is convenient when g_X is isotropic, not at least when considering plots of this function: This means that $g_X(x) = g_{X,o}(\|x\|)$ for all $x \in \mathbb{R}^d$. With a slight abuse of terminology, we also refer to g_X and $g_{X,o}$ as PCFs.

For a PDF h on $\mathbb{R}^d$, let $\tilde{h}(x) := h(-x)$ and let

$$h * \tilde{h}(x_1 - x_2) = \int h(x_1 - y) h(x_2 - y)\,\mathrm{d}y \tag{3.3}$$

be the convolution of h and $\tilde{h}$. Note that if U and V are IID random variables with PDF h, then $U - V$ has PDF $h * \tilde{h}$. In the following section we consider the case

$$g_X - 1 = a\ h * \tilde{h} \tag{3.4}$$

for real constants a, where X in particular, may refer to the initial generation process, G_0, or the noise process, Z_n. This corresponds to X being a Poisson process if $a = 0$, a point process with positive association between its points (attractiveness, clustering, or clumping) if $a > 0$, and a point process with negative association between its points (repulsiveness or regularity) if $a < 0$. In Shimatani (2010), for the initial generation process G_0, he briefly discussed the special case of (3.4) when $h \sim N_2(\tau^2/2)$ (so $h * \tilde{h} \sim N_2(\tau^2)$) whilst the noise processses are stationary Poisson processes. However, if $a \neq 0$ he did not argue if an underlying point process with PCF g_X exists. Indeed, as detailed in Appendix A, there exist α-weighted determinantal point processes satisfying (3.4) if $\alpha = -1/a$ is a positive integer, and there exist Cox processes given by α-weighted permanental point processes satisfying (3.4) if $\alpha = 1/a$ is a positive half-integer. Additionally, h needs not to be Gaussian when dealing with weighted determinantal and permanental point processes; e.g. h may be the density of a normal-variance mixture distribution (Barndorff-Nielsen et al., 1982). Notice also that (3.4) holds for many other cases of point process models for X: If the Fourier transform $\mathcal{F}\left(g_{G_0}-1\right)$ is well-defined and non-negative, if $h=\tilde{h}$, and if $a := \int (g_X - 1) \in (0,\infty)$, then (3.4) holds with

$$h = \mathcal{F}^{-1}\left\{\sqrt{\mathcal{F}(g_X-1)}\right\}/\sqrt{a}$$

provided this inverse transform is well-defined.

The following lemma is needed in Section 3.2.3.

Lemma 3.2. *Suppose g_X is of the form (3.4). Then for any function f*

$$\iint \{g_X(x_1-x_2)-1\} f(u-x_1) f(v-x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2 = ah * \tilde{h} * f * \tilde{f}(u-v) \tag{3.5}$$

for any $u,v \in \mathbb{R}^d$.

Proof. Follows from (3.3) and (3.4) using Fubini's theorem and the fact that the convolution operation is commutative and associative. □

3.2.2 First main result.

This section concerns our first main result, Theorem 3.3. We use the following notation. Define

$$c_n = \mathrm{E}\{\#Y_{n,x}(\#Y_{n,x}-1)\}/\beta_n^2 = (\nu_n + \beta_n^2 - \beta_n)/\beta_n^2 \qquad \text{if } \beta_n > 0, \tag{3.6}$$

with $c_n = 0$ if $\beta_n = 0$. If $\beta_n = \nu_n > 0$, as in the case when $\#Y_{n,x}$ follows a (non-degenerated) Poisson distribution, then $c_n = 1$. The case of overdispersion (underdispersion), that is, $\nu_n > \beta_n$ ($\nu_n < \beta_n$) corresponds to $c_n > 1$ ($c_n < 1$).

Theorem 3.3. *Suppose g_{G_0} and $g_{G_{Z_n}}$ are of the form (3.4), that is, $g_{G_0}-1 = af_0 * \tilde{f}_0$ and $g_{Z_n} - 1 = b_n f_{Z_n} * \tilde{f}_{Z_n}$ for $n=1,2,\ldots$. Then, for all $u \in \mathbb{R}^d$ and $n = 1,2,\ldots$,*

$$g_{G_n}(u) - 1 = \left(\frac{\rho_{G_0}}{\rho_{G_n}} \prod_{i=1}^{n} \beta_i p_i\right)^2 af_0 * \tilde{f}_0 * \cdots * f_n * \tilde{f}_n(u) \tag{3.7}$$

$$+ \sum_{i=1}^{n} \frac{c_i \rho_{G_{i-1}}}{\rho_{G_n}^2} \left(\prod_{j=i}^{n} \beta_j p_j\right)^2 f_i * \tilde{f}_i * \cdots * f_n * \tilde{f}_n(u) \tag{3.8}$$

$$+ \sum_{i=1}^{n-1} \left(\frac{\rho_{Z_i}}{\rho_{G_n}} \prod_{j=i+1}^{n} \beta_j p_j\right)^2 b_i f_{Z_i} * \tilde{f}_{Z_i} * f_{i+1} * \tilde{f}_{i+1} * \cdots * f_n * \tilde{f}_n(u) \tag{3.9}$$

$$+ \left(\frac{\rho_{Z_n}}{\rho_{G_n}}\right)^2 b_n f_{Z_n} * \tilde{f}_{Z_n}(u) \tag{3.10}$$

where the sum in (3.9) is interpreted as zero if $n = 1$.

The terms in (3.7)–(3.10) have the following interpretations: The right side of (3.7) corresponds to pairs of n-th generation points with different 0-th generation ancestors; the i-th term in (3.8) corresponds to such pairs when they have a common

$(i-1)$-th generation ancestor; the i-th term in (3.9) corresponds to pairs of n-th generation points with different i-th generation ancestors initiated by the noise process Z_i; and the term in (3.10) corresponds to point pairs in Z_n.

Later in Section 4.1, our main interest is in the behaviour of g_{G_n} as $n \to \infty$ when we have the same reproduction system, but for the moment, it is worth noticing the flexibility of our model for G_1 and the effect of the choice of its centre process G_0: Suppose there is no noise and G_0 is stationary and either a Poisson or an weighted determinantal or permanental point process with a Gaussian kernel. Specifically, $d = 2$, G_0 has intensity $\rho_{G_0} = 100$, and using a notation as in Appendix A, the Gaussian kernel has an auto-correlation function of the form $R(x) = \exp(-\|x/\tau\|^2)$, where the value of τ depends on the type of process: For the α-weighted determinantal point process, we consider the most repulsive case, that is, a determinantal point process ($\alpha = 1$) and $\tau = 1/\sqrt{\rho_{G_0}\pi}$ is largest possible to ensure existence of the process (Lavancier et al. (2015)); for the α-weighted permanental point process, $\alpha = 1/2$ (the most attractive case when it is also a Cox process, see Appendix A) and $\tau = 0.1$ is an arbitrary value (any positive number can be used). Note that $R^2 = (\sqrt{\pi}\tau)^2 f_0 * \tilde{f}_0$ where $f_0 \sim N_2(\tau^2/8)$, which by (A.1) and (A.2) mean that (3.4) is satisfied with $a = 2(\sqrt{\pi}\tau)^2$ and $a = -(\sqrt{\pi}\tau)^2$ for the weighted permanental and determinantal point processes, respectively, and $a = 0$ in case of the Poisson process. Moreover, let the number of points in a cluster be Poisson distributed with mean $\beta_1 = 10$, $p_1 = 1$, and $f_1 \sim N_2(\sigma^2)$, with $\sigma = 0.01$. Then, by Theorem 3.3,

$$g_{G_1}(u) - 1 = \frac{a}{\sqrt{2\pi(2\sigma^2 + \tau^2/4)}} \exp\left\{-\frac{\|u\|^2}{2(2\sigma^2 + \tau^2/4)}\right\} + \frac{1}{\rho_{G_0}\sqrt{4\pi\sigma^2}} \exp\left\{-\frac{\|u\|^2}{4\sigma^2}\right\}.$$

Figure 1 shows the isotropic PCF $g_{G_{1,o}}(r) = g_{G_1}(u)$ as a function of the inter-point distance $r = \|u\|$ in case of each of the three models of G_0, where using an obvious notation, $g^{\mathrm{det}}_{G_{1,o}} < g^{\mathrm{Pois}}_{G_{1,o}} < g^{\mathrm{wper}}_{G_{1,o}}$. Most notable is the fact that $g^{\mathrm{det}}_{G_{1,o}}(r)$ exhibits repulsion at midrange distances r. For $g^{\mathrm{wper}}_{G_{1,o}}$, we see a high degree of clustering, which is persistent for large values of r; this will of course be even more pronounced if we increase the value of τ; whilst decreasing σ will increase the peak at small values of r. Figure 2 shows simulations of G_1 in each of the three cases of the model of G_0. As expected, we clearly see a higher degree of repulsion when G_0 is a determinantal point process (the left most plot) and a higher degree of clustering when G_0 is a weighted permanental point process (the right most plot). In particular, the clusters are more distinguishable when G_0 is a determinantal point process, and this will be even more pronounced if decreasing σ because the spread of clusters then decrease. When G_0 is a weighted permanental point process, the clusters overlap more.

3.2.3 Proof of Theorem 3.3.

Shimatani (2010) verified Theorem 3.3 when both $b_1 = b_2 = \cdots = 0$ (as is the case if $Z_1, Z_2, \ldots$ are stationary Poisson processes) and $c_1 = c_2 = \cdots > 0$, in which case the terms in (3.9)-(3.10) are zero. If $c_1 = c_2 = \cdots = 0$, then (3.8) is zero and by (3.6), with probability one, $\#Y_{n,x} \in \{0, 1\}$ for all $x \in G_{n-1}$ and $n = 1, 2, \ldots$

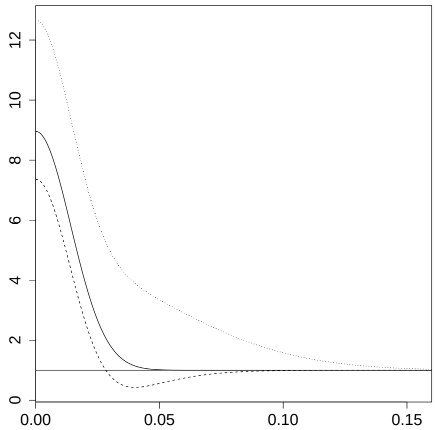

Figure 1: The PCFs of G_1 when G_0 is a determinantal, Poisson, or weighted permanental point process (dashed, solid, and dotted respectively), with parameters and Gaussian offspring PDF as specified in the text. The solid horizontal line is the theoretical PCF for a Poisson process.

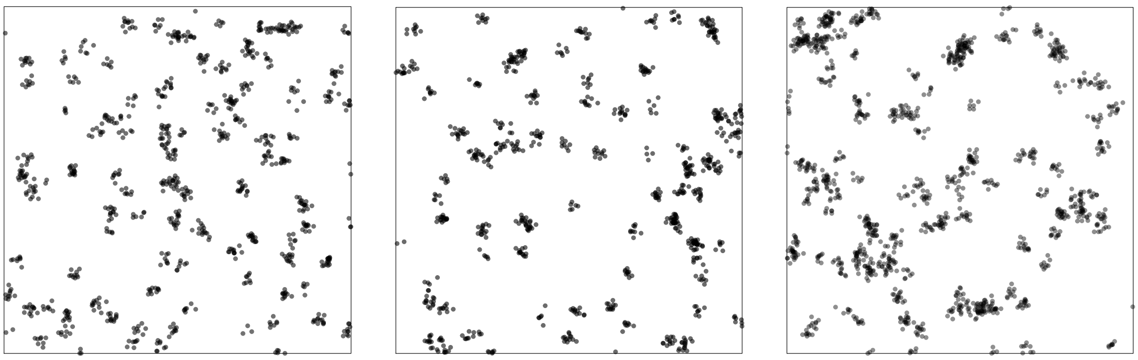

Figure 2: Simulations of G_1 restricted to a unit square when G_0 is a determinantal (left panel), Poisson (middle panel), or weighted permanental (right panel) point process, see Figure 1 and the text.

Consequently, the proof of Theorem 3.3 is trivial if $c_1 = c_2 = \cdots = 0$ and both G_0 and $Z_1, Z_2, \ldots$ are stationary Poisson processes, because then $a = 0$, $b_1 = b_2 = \cdots = 0$, $G_1, G_2, \ldots$ are stationary Poisson processes, and IID random shifts of the points in a stationary Poisson process generate a stationary Poisson process. The general proof of Theorem 3.3 follows by induction from the following Lemma 3.4 when applying Lemma 3.2.

Lemma 3.4. *If $\rho_{G_{n-1}} > 0$, $\rho_{G_n} > 0$, and $g_{G_{n-1}}$ and g_{Z_n} exist, then g_{G_n} exists and is given by*

$$
\begin{aligned}
g_{G_n}(u-v) - 1 = {} & \left(\frac{\rho_{G_{n-1}}\beta_n p_n}{\rho_{G_n}}\right)^2 \\
& \cdot \left[\iint \{g_{G_{n-1}}(x_1 - x_2) - 1\} f_n(u - x_1)\tilde{f}_n(v - x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2 \right. \\
& \left. + \frac{c_n}{\rho_{G_{n-1}}} f_n * \tilde{f}_n(u - v)\right] + \left(\frac{\rho_{Z_n}}{\rho_{G_n}}\right)^2 \{g_{Z_n}(u - v) - 1\}
\end{aligned}
\tag{3.11}
$$

for any $u, v \in \mathbb{R}^d$.

Proof. Note that Y_n is stationary with intensity

$$\rho_{Y_n} = \rho_{G_{n-1}}\beta_n. \tag{3.12}$$

It follows straightforwardly from (1.1), (3.2), and Fubini's theorem that its PCF is given by

$$\begin{aligned}\rho_{Y_n}^2 g_{Y_n}(u-v) = \rho_{G_{n-1}}^2\beta_n^2 \iint g_{G_{n-1}}(x_1-x_2)f_n(u-x_1)f_n(v-x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2 \\ + \rho_{G_{n-1}}c_n\beta_n^2 f_n * \tilde{f}_n(u-v)\end{aligned} \tag{3.13}$$

for any $u, v \in \mathbb{R}^d$, where the two terms on the right hand side correspond to pairs of points from Y_n belonging to different clusters and the same cluster, respectively. Hence by (1.2) and (3.12), W_n is stationary with intensity

$$\rho_{W_n} = p_n\rho_{Y_n} = \rho_{G_{n-1}}\beta_n p_n \tag{3.14}$$

and PCF

$$\begin{aligned} g_{W_n}(u-v) &= g_{Y_n}(u-v) \\ &= \iint g_{G_{n-1}}(x_1-x_2)f_n(u-x_1)f_n(v-x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2 \\ &\quad + \frac{c_n}{\rho_{G_{n-1}}} f_n * \tilde{f}_n(u-v)\end{aligned} \tag{3.15}$$

where the first identify follows from the fact that PCFs are invariant under independent thinning, and where (3.13) is used to obtain the second identity. Furthermore, it follows straightforwardly from (1.3), (3.2), and Fubini's theorem that G_n has PCF given by

$$\rho_{G_n}^2 g_{G_n}(x) = \rho_{W_n}^2 g_{W_n}(x) + 2\rho_{W_n}\rho_{Z_n} + \rho_{Z_n}^2 g_{Z_n}(x)$$

where the three terms on the right hand side correspond to pairs of points from W_n, from W_n and Z_n (which can be ordered in two ways), and from Z_n, respectively. Combining this with the first identity in (3.1) and (3.14), we easily obtain

$$g_{G_n}(x) - 1 = \left(\frac{\rho_{G_{n-1}}\beta_n p_n}{\rho_{G_n}}\right)^2 \{g_{W_n}(x) - 1\} + \left(\frac{\rho_{Z_n}}{\rho_{G_n}}\right)^2 \{g_{Z_n}(x) - 1\}$$

which combined with (3.15) imply (3.11). □

3.2.4 Extension.

More generally than in Section 3.2.2 we may consider the case where the PCF of the initial generation G_0 and the noise Z_n are affine expressions:

$$g_{G_0} - 1 = a_0 + a_1 f_{0,1} * \tilde{f}_{0,1} + \cdots + a_k f_{0,k} * \tilde{f}_{0,k} \tag{3.16}$$

and

$$g_{Z_n} - 1 = b_{n,0} + b_{n,1} f_{Z_n,1} * \tilde{f}_{Z_n,1} + \cdots + b_{n,l} f_{Z_n,l} * \tilde{f}_{Z_n,l}, \qquad n = 1, 2, \ldots, \tag{3.17}$$

for real constants $a_0, \dots, a_k, b_{n,1}, \dots, b_{n,l}$ and PDFs $f_{0,1}, \dots, f_{0,k}, f_{Z_n,1}, \dots, f_{Z_n,l}$. For instance, the superposition of k independent Poisson, weighted permanental, or weigthed determinantal point processes has a PCF of the form (3.16). Then we have

$$\begin{aligned} g_{G_n}(u) - 1 = &\left(\frac{\rho_{G_0}}{\rho_{G_n}} \prod_{i=1}^{n} \beta_i p_i\right)^2 \left(a_0 + \sum_{i=1}^{k} a_i f_{0,i} * \tilde{f}_{0,i} * f_1 * \tilde{f}_1 \ldots * f_n * \tilde{f}_n(u)\right) \\ &+ \sum_{i=1}^{n} \frac{c_i \rho_{G_{i-1}}}{\rho_{G_n}^2} \left(\prod_{j=i}^{n} \beta_j p_j\right)^2 f_i * \tilde{f}_i * \ldots * f_n * \tilde{f}_n(u) \\ &+ \sum_{i=1}^{n-1} \left(\frac{\rho_{Z_i}}{\rho_{G_n}} \prod_{j=i+1}^{n} \beta_j p_j\right)^2 \\ &\quad \cdot \left(b_{i,0} + \sum_{j=1}^{l} b_{i,j} f_{Z_i,j} * \tilde{f}_{Z_i,j} * f_{i+1} * \tilde{f}_{i+1} * \ldots * f_n * \tilde{f}_n(u)\right) \\ &+ \left(\frac{\rho_{Z_n}}{\rho_{G_n}}\right)^2 \left(b_{n,0} + \sum_{j=1}^{l} b_{n,j} f_{Z_n,j} * \tilde{f}_{Z_n,j}(u)\right). \end{aligned}$$

Essentially, this follows from Theorem 3.3 by replacing $af_0 * \tilde{f}_0$ in (3.7) by 3.16, and $b_i f_{Z_i} * \tilde{f}_{Z_i}$ in (3.9) and (3.10) by 3.17.

4 Same reproduction system

Throughout this section we assume the same reproduction system over generations, that is, in items (a)–(c), $\beta_n = \beta$, $\nu_n = \nu$, $f_n = f$, $p_n = p$ do not depend on n, $Z_1, Z_2, \dots$ are IID stationary point processes, so $\rho_{Z_n} = \rho_Z$ for $n = 1, 2, \dots,$ and $\rho_{G_0} = \rho_{G_1} = \cdots = \rho_G > 0$. Note that the noise process Z_n and the initial generation process G_0 need not be Poisson processes and the offspring densities need not be Gaussian as in Shimatani (2010). By (3.1), this implies either

$$\beta p = 1 \quad \text{and} \quad \rho_Z = 0, \tag{4.1}$$

or

$$\beta p < 1 \quad \text{and} \quad \rho_Z > 0. \tag{4.2}$$

In case of (4.2),

$$\rho_G = \rho_Z / (1 - \beta p). \tag{4.3}$$

4.1 Limiting pair correlation function

Under the assumptions above and in Theorem 3.3, the PCF simplifies such that

$$\begin{aligned}
g_{G_n}(u) - 1 &= (\beta p)^{2n} a f_0 * \tilde{f}_0 * f^{*n} * \tilde{f}^{*n}(u) \\
&\quad + \frac{c}{\rho_G} \sum_{i=1}^{n} (\beta p)^{2(n-i+1)} f^{*(n-i+1)} * \tilde{f}^{*(n-i+1)}(u) \\
&\quad + \left(\frac{\rho_Z}{\rho_G}\right)^2 \sum_{i=1}^{n} (\beta p)^{2(n-i)} b f_Z * \tilde{f}_Z * f^{*(n-i)} * \tilde{f}^{*(n-i)}(u) \\
&= (\beta p)^{2n} a f_0 * \tilde{f}_0 * f^{*n} * \tilde{f}^{*n}(u) + \frac{c}{\rho_G} \sum_{i=1}^{n} (\beta p)^{2i} f^{*i} * \tilde{f}^{*i}(u) \\
&\quad + \left(\frac{\rho_Z}{\rho_G}\right)^2 b f_Z * \tilde{f}_Z * \sum_{i=0}^{n-1} (\beta p)^{2i} f^{*i} * \tilde{f}^{*i}(u),
\end{aligned} \tag{4.4}$$

for $n = 1, 2, \ldots$, where

$$c = (\nu + \beta^2 - \beta)/\beta^2 \quad \text{if } \beta > 0, \qquad c = 0 \quad \text{if } \beta = 0,$$

and where f^{*n} is the n-th convolution power of f. For instance, consider the case $f_0 \sim N_d(\tau^2)$, $f \sim N_d(\sigma^2)$, and $f_Z \sim N_d(\kappa^2)$. Then

$$\begin{aligned}
g_G(u) - 1 &:= \lim_{n\to\infty} g_{G_n}(u) - 1 \\
&= \frac{c}{\rho_G} \sum_{i=1}^{\infty} \frac{(\beta p)^{2i}}{(4\pi i\sigma^2)^{d/2}} \exp\left(-\frac{\|u\|^2}{4i\sigma^2}\right) \\
&\quad + b\left(\frac{\rho_Z}{\rho_G}\right)^2 \sum_{i=0}^{\infty} \frac{(\beta p)^{2i}}{\{4\pi(i\sigma^2+\kappa^2)\}^{d/2}} \exp\left\{-\frac{\|u\|^2}{4(i\sigma^2+\kappa^2)}\right\}
\end{aligned} \tag{4.5}$$

is finite if and only if $\beta p < 1$ or both $\beta p = 1$ and $d \geq 3$. Shimatani (2010) considered this special case for $d = 2$, $b = 0$, and $c > 0$; he noticed that (4.1) implies divergence of g_{G_n} as $n \to \infty$ whilst (4.2) implies convergence, where in the latter case, when $\beta p \approx 1$, he discussed an approximation of $g_G(u)$ that depends on whether $\|u\|$ is close to 0 or not.

In general, if we assume (4.2) and that $g_{G_n} - 1$ has a finite limit, we have

$$g_G(u) - 1 = \frac{c}{\rho_G} \sum_{i=1}^{\infty} (\beta p)^{2i} f^{*i} * \tilde{f}^{*i}(u) + \left(\frac{\rho_Z}{\rho_G}\right)^2 b f_Z * \tilde{f}_Z * \sum_{i=0}^{\infty} (\beta p)^{2i} f^{*i} * \tilde{f}^{*i}(u) \tag{4.6}$$

which does not depend on a or f_0. Here, as $\beta p \uparrow 1$, the second term goes to zero, meaning that the less noise we consider, the less it matters which type of noise process we choose. On the other hand, as $\beta p \downarrow 0$, $g_G - 1$ tends to $b f_Z * \tilde{f}_Z$, which simply is the PCF of the noise process Z_n.

Considering the situation at the end of Section 3.2.2, assume that $d = 2$, $f \sim N_d(\sigma^2)$, and $g_{Z_n} - 1 = b f_Z * \tilde{f}_Z$ (corresponding to (3.4)) with $f_Z \sim N_d(\kappa^2/8)$ and $b = 0$, $b = -(\sqrt{\pi}\kappa)^2$, and $b = 2(\sqrt{\pi}\kappa)^2$ for the Poisson, determinantal, and weighted permanental point process, respectively. Then $g_G(u)$ is given by (4.5), where $d = 2$ and κ^2 is replaced by $\kappa^2/8$. Also assume that $p = 1$, $\sigma = 0.1$, $\rho_G = 100$, and the number of points in a cluster is Poisson distributed (implying $c = 1$) with mean $\beta = 0.8$, so $\rho_Z = 20$. Finally, assume $\kappa = 0.1$ in case of weighted permanental noise and $\kappa = 1/\sqrt{\rho_Z \pi}$ in case of determinantal noise (the most repulsive Gaussian determinantal point process). Shimantani Shimatani (2010) discussed the case where $\beta p = 0.99$ – a plot (omitted here) shows that the limiting PCFs corresponding to the three models of noise processes are then effectively equal. By lowering βp, the reproduction system is diminished, and hence depending on the model type, a higher degree of regularity or clustering is obtained. This will also increase the rate of convergence because the number of generations initialized by a single point will be fewer. Note that in Figure 3 the convergence is already rapid as g_{G_8} and $g_{G_{16}}$ are practically indistinguishable. Figure 3 further shows that it is only for small inter-point distances that the three limiting PCFs differ – and only slightly.

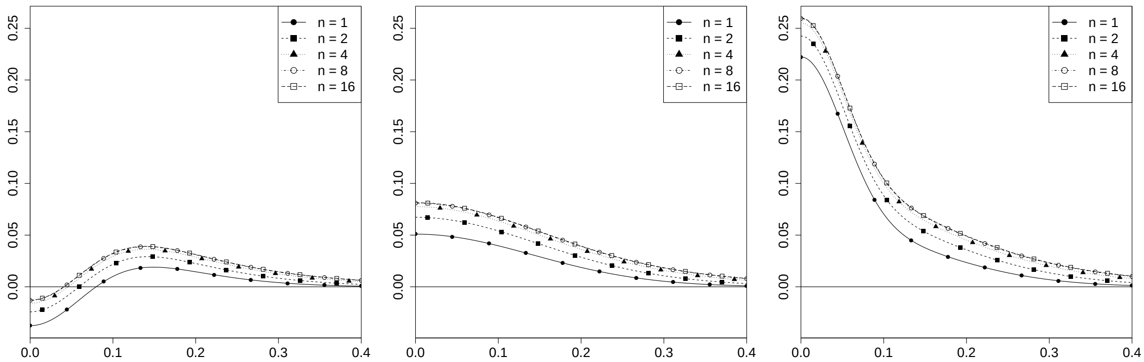

Figure 3: The PCFs of G_n when the noise processes are either determinantal, Poisson or weighted permanental point processes (left to right), with parameters and Gaussian offspring PDF as specified in the text. The solid horizontal line is the theoretical PCF for a Poisson process.

4.2 Second main result

Although Shimatani (2010) showed convergence of g_{G_n} in the special case considered above, he did not clarify whether the Markov chain $G_0, G_1, \ldots$ converges in distribution to a limit so that this limiting distribution (also called the equilibrium, invariant, or stationary distribution) has a PCF given by (4.6). In order to show that $G_0, G_1, \ldots$ is indeed converging to a limiting distribution under more general conditions, and to specify what this is, we construct in accordance with items (a)–(c) a Markov chain $\ldots, G_{-1}^{\mathrm{st}}, G_0^{\mathrm{st}}, G_1^{\mathrm{st}}, \ldots$ with times given by all integers n and so that this chain is time-stationary (its distribution is invariant under discrete time shifts), as follows. First, we generate noise processes as in item (c): Let $\ldots, Z_{-1}, Z_0, Z_1, \ldots$ be independent stationary Poisson processes on $\mathbb{R}^d$ with intensity ρ_Z. Second, for any integer n and point $x \in Z_n$, we consider the family of all generations initiated by the ancestor x, that is, the family

$$F_{n,x} = \bigcup_{m=1}^{\infty} W_{n,x}^{(m)}$$

where $W^{(1)}_{n,x} = W_{n,x}$ is defined by the reproduction mechanism of independent clustering and independent thinning given in items (a)–(b) (with $\beta_n = \beta$ and $\nu_n = \nu$), $W^{(2)}_{n,x}$ is the retained offspring generated by the points in $W^{(1)}_{n,x}$ (using the same reproduction mechanism as before), and so on. In other words, $W^{(m)}_{n,x}$ is the set of $(m+n)$-th generation points with common ancestor n-th generation ancestor $x \in Z_n$. Moreover, we assume that conditional on $\ldots, Z_{-1}, Z_0, Z_1, \ldots$, the families $F_{n,x}$ for all integers n and $x \in Z_n$ are independent (and hence IID). Finally, for all integers n, we let

$$G^{\mathrm{st}}_n = W^{\mathrm{st}}_n \cup Z_n \qquad \text{with } W^{\mathrm{st}}_n = \bigcup_{m=1}^{\infty} \bigcup_{x \in Z_{n-m}} W^{(m)}_{n-m,x}. \tag{4.7}$$

For completeness, we show in Appendix B that any G^{st}_n has intensity ρ_G given by (4.3) and PCF g_G given by (4.6), although this should be evident from Theorem 4.2 below. The proof of Theorem 4.2 is based on a coupling construction between $G_1, G_2, \ldots$ and $G^{\mathrm{st}}_1, G^{\mathrm{st}}_2, \ldots$ together with the following result.

Lemma 4.1. *Suppose $\beta_n = \beta$, $\nu_n = \nu$, $f_n = f$, $p_n = p$, and $\rho_{Z_n} = \rho_Z$ do not depend on $n \geq 1$, where $\beta p < 1$ and $\rho_Z > 0$. Let $K \subset \mathbb{R}^d$ be a compact set and let*

$$T^{\mathrm{st}}_{0,K} = \sup\big\{m \in \{1, 2, \ldots\} : W^{(m)}_{0,x} \cap K \neq \emptyset \text{ for some } x \in G^{\mathrm{st}}_0\big\} \tag{4.8}$$

be the last time a point in K is a member of a family initiated by some point in the 0-th generation G^{st}_0. Then

$$\mathrm{E}\left(T^{\mathrm{st}}_{0,K}\right) \leq |K| \rho_G \frac{\beta p}{1 - \beta p}$$

is finite, and so $T^{\mathrm{st}}_{0,K} < \infty$ almost surely.

Proof. Let $K \subset \mathbb{R}^d$ be compact and define

$$N = \sum_{x \in G^{\mathrm{st}}_0} \#(F_{0,x} \cap K).$$

We have

$$\mathrm{E}(N) = \rho_G \int \left\{ \sum_{m=1}^{\infty} \int_K (\beta p)^m f^{*m}(y - x)\, \mathrm{d}y \right\} \mathrm{d}x = |K| \rho_G \frac{\beta p}{1 - \beta p}$$

using Fubini's theorem in the last identity. Further, the families initiated by the points in G^{st}_0 are almost surely pairwise disjoint, so N is almost surely the number of points in K belonging to some family initiated by a point $x \in G^{\mathrm{st}}_0$. Consequently, $\mathrm{P}(T^{\mathrm{st}}_{0,K} \leq N) = 1$, whereby the lemma follows. □

We are now ready to state our second main result.

Theorem 4.2. *Suppose* $\ldots, Z_{-1}, Z_0, Z_1 \ldots$ *are IID stationary point processes and* $\beta_n = \beta$, $\nu_n = \nu$, $f_n = f$, $p_n = p$, *and* $\rho_{Z_n} = \rho_Z$ *do not depend on* $n \geq 1$, *where* $\beta p < 1$ *and* $\rho_Z > 0$. *Then* $\ldots, G_{-1}^{\mathrm{st}}, G_0^{\mathrm{st}}, G_1^{\mathrm{st}}, \ldots$ *is a time-stationary Markov chain constructed in accordance to items (a)–(c). Let* Π *be the distribution of any* G_n^{st} *and let* $\mathcal{N}$ *be the space of all locally finite subsets of* $\mathbb{R}^d$. *Then there exists a (measurable) subset* $\Omega \subseteq \mathcal{N}$ *so that* $\Pi(\Omega) = 1$ *and for any compact set* $K \subset \mathbb{R}^d$ *and all* $\omega \in \Omega$, *conditional on* $G_0 = \omega$, *there is a coupling between* $G_1, G_2, \ldots$ *and* $\ldots, G_{-1}^{\mathrm{st}}, G_0^{\mathrm{st}}, G_1^{\mathrm{st}}, \ldots$, *and there exists a random time* $T_K(\omega) \in \{0, 1, \ldots\}$ *so that* $G_n \cap K = G_n^{\mathrm{st}} \cap K$ *for all integers* $n > T_K(\omega)$. *In particular, for any* $\omega \in \Omega$ *and conditional on* $G_0 = \omega$, G_n *converges in distribution to* Π *as* $n \to \infty$, *and so* Π *is the unique invariant distribution of the chain* $G_0, G_1, \ldots$.

Proof. Obviously, $\ldots, G_{-1}^{\mathrm{st}}, G_0^{\mathrm{st}}, G_1^{\mathrm{st}}, \ldots$ is a time-stationary Markov chain constructed in accordance to items (a)–(c). To verify the remaining part of the theorem, we may assume that G_0 and G_0^{st} are independent. Then, conditional on G_0, we have a coupling between $G_1, G_2, \ldots$ and $\ldots, G_{-1}^{\mathrm{st}}, G_0^{\mathrm{st}}, G_1^{\mathrm{st}}, \ldots$ because $G_1^{\mathrm{st}}, G_2^{\mathrm{st}}, \ldots$ and $G_1, G_2, \ldots$ are generated by the same noise processes $Z_1, Z_2, \ldots$, the same offspring processes $Y_{n,x}$ for all times $n = 1, 2, \ldots$ and all ancestors $x \in G_{n-1} \cap G_{n-1}^{\mathrm{st}}$, and the same Bernoulli variables $B_{n,y}$ for all times $n = 1, 2, \ldots$ and all offspring $y \in Y_{n,x}$ with ancestor $x \in G_{n-1} \cap G_{n-1}^{\mathrm{st}}$. Let $K \subset \mathbb{R}^d$ be compact. In accordance with (4.8), let

$$T_K(\omega) = \sup\big\{m \in \{1, 2, \ldots\} : W_{0,x}^{(m)} \cap K \neq \emptyset \text{ for some } x \in \omega\big\}$$

be the last time a point in K is a member of a family initiated by some point in ω, and let $\Omega = \{\omega \in \mathcal{N} : T_K(\omega) < \infty\}$. By Lemma 4.1 and the coupling construction, $\Pi(\Omega) = 1$ and $G_n \cap K = G_n^{\mathrm{st}} \cap K$ whenever $n > T_K(\omega)$, so for any $\omega \in \Omega$,

$$\lim_{n \to \infty} \mathrm{P}\left(G_n \cap K = \emptyset | G_0 = \omega\right) = \lim_{n \to \infty} \mathrm{P}\left(G_n^{\mathrm{st}} \cap K = \emptyset, n > T_K(\omega)\right)$$

because G_0 is independent of $(G_0^{\mathrm{st}}, T_K(\omega))$. Since the sequence of events $\{\omega : 1 > T_K(\omega)\} \subseteq \{\omega : 2 > T_K(\omega)\} \subseteq \ldots$ increases to Ω, we obtain

$$\lim_{n \to \infty} \mathrm{P}\left(G_n \cap K = \emptyset | G_0 = \omega\right) = \lim_{n \to \infty} \mathrm{P}\left(G_n^{\mathrm{st}} \cap K = \emptyset\right) = \mathrm{P}\left(G_0^{\mathrm{st}} \cap K = \emptyset\right).$$

Thus, recalling that the distribution of a random closed set $X \subseteq \mathbb{R}^d$ (including a locally finite point process) is uniquely characterized by the void probabilities $\mathrm{P}(X \cap K = \emptyset)$ for all compact sets $K \subset \mathbb{R}^d$, we have verified that conditional on $G_0 = \omega$, the chain $G_1, G_2 \ldots$ converges in distribution towards Π. In turn, this implies uniqueness of the invariant distribution Π. □

In Theorem 4.2, under mild conditions, we can take $\Omega = \mathcal{N}$. For instance, this is easily seen to be the case if there exists $\epsilon > 0$ so that $f(x) > 0$ whenever $\|x\| \leq \epsilon$. In the special case $c = 0$, Π is just a stationary Poisson process, and so $\Omega = \mathcal{N}$. Moreover, the integral

$$\gamma := \int (g_G - 1)$$

is a rough measure of the amount of positive/negative association between the points in G_n^{st}. Note that comparing γ with the corresponding measure for another stationary point process makes only sense if the processes have equal intensities, see Lavancier et al. (2015). Under the assumptions in both Theorem 3.3 and 4.2, by (4.6),

$$\gamma = \frac{c(\beta p)^2}{\rho_G\left\{1-(\beta p)^2\right\}} + \frac{b\rho_Z^2}{\rho_G^2\left\{1-(\beta p)^2\right\}} = \frac{1}{1+\beta p}\left\{\frac{c(\beta p)^2}{\rho_Z} + b(1-\beta p)\right\}$$

which does not depend on f or f_Z. Furthermore, γ may take any positive value and some negative values depending on how we choose the parameter $(\beta, \nu, p, b, \rho_Z)$. This means we may have an equilibrium distribution exhibiting any degree of clustering or some degree of regularity. In fact, γ can only be negative when b is negative, e.g when Z_n is a determinantal point process. In this case b has a lower bound, $b_{\min}$, that ensures the existence of the determinantal point process (Lavancier et al., 2015) and consequently, $\gamma \geq b_{\min}$. The case $\gamma = b_{\min}$ happens exactly when $\beta p = 0$ (i.e. when no points are retained from the previous generation process) and thus $G_n = Z_n$ is a determinantal point process.

For approximate simulation of G_0^{st} under each of the three models of the noise processes, we use the algorithm described in Appendix C. Simulation was initially done with parameters and set-up corresponding to that of Figure 3. However, the resulting point patterns were not distinguishable from a stationary Poisson process when comparing empirical estimates of the PCF, L-function, or J-function of the simulations to 95% global rank envelopes under each model (for definition of L- and J-functions, see e.g. Møller and Waagepetersen (2004), and for the envelopes, see Myllymäki et al. (2016)). Therefore, in order to better distinguish the three models, we consider two cases as follows.

Case 1:

This case is based on minimizing γ under determinantal noise and on maximizing γ under weighted permanental noise. Let $d = 2$, $f \sim N_d(\sigma^2)$, with $\sigma = 0.1$, $f_Z \sim N_d(\kappa^2/8)$, $\rho_G = 100$, $p = 1$, $\beta = 0.3$, and consequently $\rho_Z = 70$.

- In case of determinantal noise: Let $\kappa = 1/\sqrt{\rho_Z \pi}$ (the most repulsive Gaussian determinantal point process) and the number of points in a cluster be Bernoulli distributed with parameter β, implying $c = 0$ (each point has at most one offspring). Then $\gamma \approx -5.38 \times 10^{-3}$.
- In case of Poisson noise: Let the number of points in a cluster be Poisson distributed with intensity β, implying $c = 1$. Then $\gamma \approx 9.89 \times 10^{-4}$.
- In case of weighted permanental noise: Let $\kappa = 1$ and the number of points in a cluster be negative binomially distributed with probability of success equal to 0.12 and dispersion parameter equal to 0.11, implying $c = 10$. Then $\gamma \approx 3.39$.

Case 2:

This case is such that the clusters are more separated. Let $d = 2$, $f \sim N_d(\sigma^2)$, with $\sigma = 0.01$, $f_Z \sim N_d(\kappa^2/8)$, $\rho_G = 100$, $p = 1$, $\beta = 0.95$, and consequently $\rho_Z = 5$. Also, let the number of points in a cluster be negative binomially distributed with probability of success equal to 0.208 and dispersion parameter equal to 0.25, implying $c = 5$.

- In case of determinantal noise: Let $\kappa = 1/\sqrt{\rho_Z \pi}$. Then $\gamma \approx 0.463$.
- In case of Poisson noise: $\gamma \approx 0.463$.
- In case of weighted permanental noise: Let $\kappa = 1$. Then $\gamma \approx 0.624$.

Figure 4 shows simulations of G_0^{st} under each of the three models of the noise processes (left to right) in Case 1 and 2 (top and bottom). Based on these simulations, Figure 5 shows empirical estimates of functional summary statistics based on the simulated point patterns from Figure 4 along with 95% global rank envelopes based on 2499 simulations (as recommended in Myllymäki et al. (2016)) of a stationary Poisson process with the same intensity as used in Figure 4. The first simulated point pattern of Case 1 looks slightly less clustered than the second, whilst the last looks more clustered. This is in accordance with the values of γ and the corresponding functional summary statistics in Figure 5. Additionally, Figure 5 reveals that the case of Poisson noise is not distinguishable from the stationary Poisson process, while the case of weighted permanental noise is more clustered. The case of determinantal noise is not distinguishable from the stationary Poisson process by the PCF or L-function, but is shown to be more regular by the J-function. In Case 2, the clusters of the point pattern simulated under determinantal noise looks more separated than the clusters of the point pattern simulated under Poisson noise. The clusters of the point pattern simulated under weighted permanental noise are clustered to such a degree that it gives the illusion of few highly separated clusters. All three models of Case 2 are as expected significantly different from the stationary Poisson process.

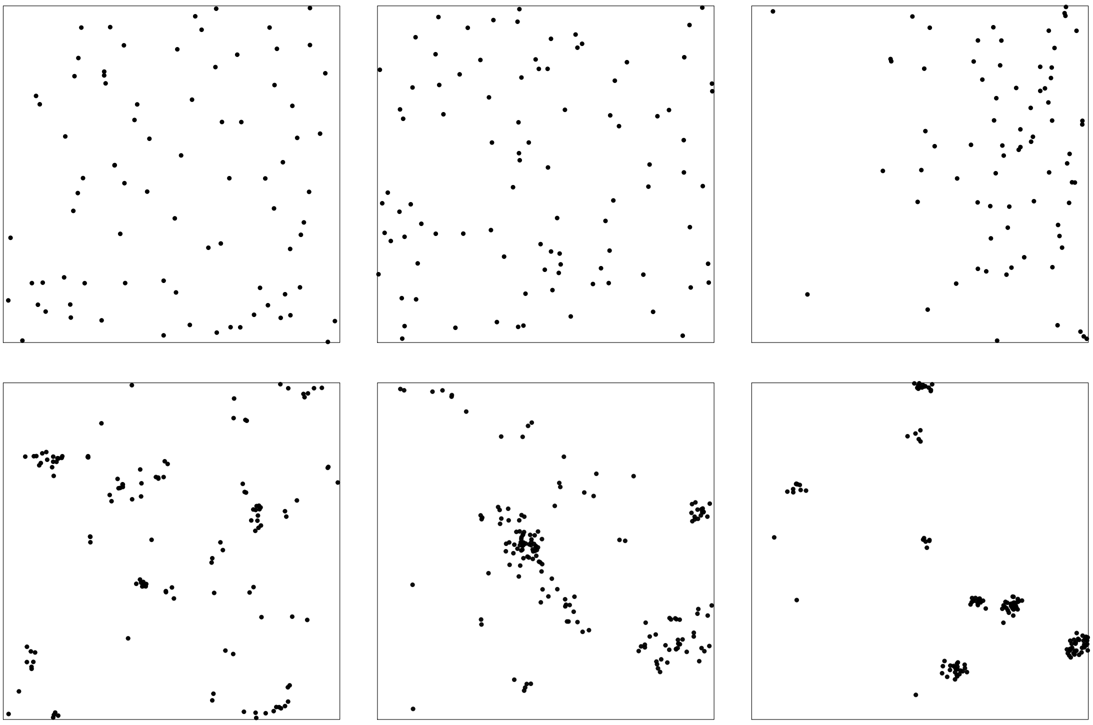

Figure 4: Simulations of G_0^{st} restricted to a unit square when the noise processes are either determinantal (left panel), Poisson (middle panel), or weighted permanental (right panel) point processes, with parameters as specified in the text. The rows corresponds to Case 1 and 2, respectively.

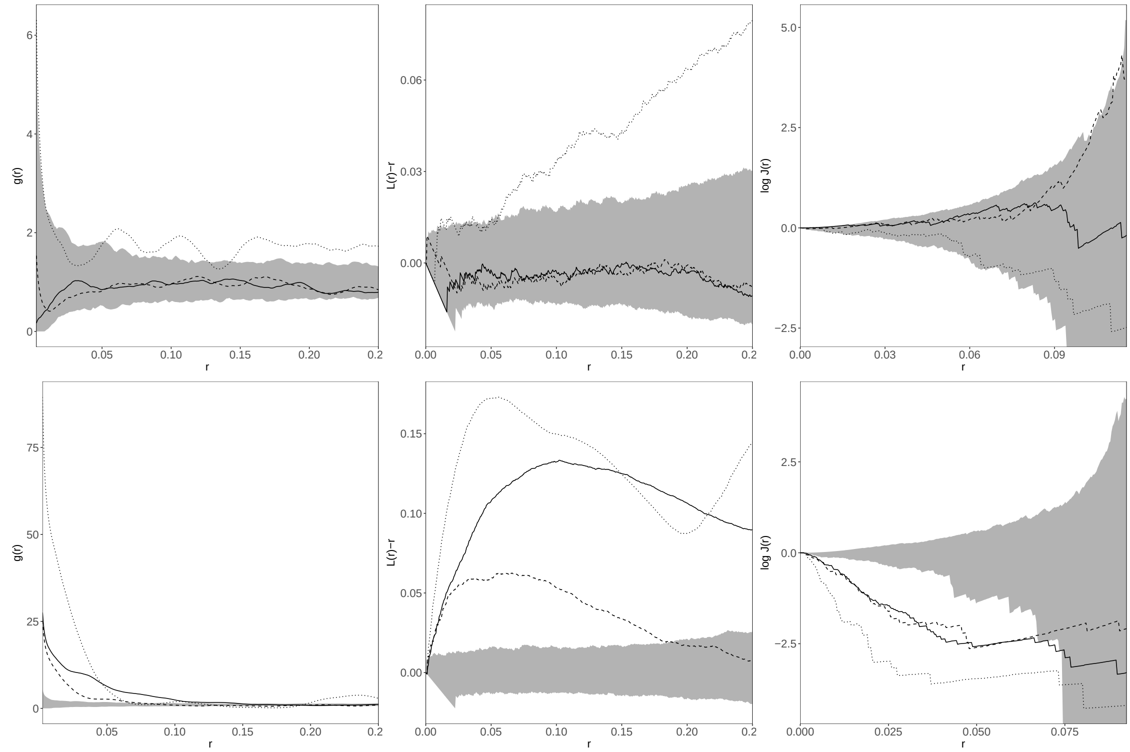

Figure 5: Empirical PCFs, L-functions, and J-functions (left to right) based on the simulations of G_0^{st} from Figure 4 when the noise processes are either determinantal (dashed), Poisson (solid), or weighted permanental (dotted). The rows corresponds to Case 1 and 2, respectively. The grey regions are 95% global rank envelopes based on 2499 simulations of a stationary Poisson process with the same intensity as G_0^{st}.

A Weighted determinantal and permanental point processes

When defining stationary weighted determinantal/permanental point processes, the main ingredients are a symmetric function $C : \mathbb{R}^d \mapsto \mathbb{R}$ and a real number α. Before giving the definitions of these point processes we recall the following.

For a real $n \times n$ matrix A with (i,j)-th entry $a_{i,j}$, the α-weighted permanent of A is defined by

$$\operatorname{per}_\alpha(A) = \sum_\sigma \alpha^{\#\sigma} a_{1,\sigma_1} \cdots a_{n,\sigma_n}$$

where σ denotes a permutation of $\{1, \dots, n\}$ and $\#\sigma$ is the number of its cycles. This is the usual permanent of A if $\alpha = 1$. Moreover, the α-weighted determinant of A is given by

$$\det_\alpha(A) = \operatorname{per}_{-\alpha}(-A).$$

This is the usual determinant of A if $\alpha = -1$. Often we just write $\operatorname{per}_\alpha A$ for $\operatorname{per}_\alpha(A)$, and $\det_\alpha A$ for $\det_\alpha(A)$.

For any $G_1, \dots, G_n \in \mathbb{R}^d$, the $n \times n$ matrix with (i,j)-th entry $C(G_i - G_j)$ is denoted by $[C](G_1, \dots, G_n)$. Thus

$$\operatorname{per}_\alpha[C](G_1, \dots, G_n) = \sum_\sigma \alpha^{\#\sigma} C(G_1 - G_{\sigma_1}) \cdots C(G_n - G_{\sigma_n}).$$

Note that the weighted permanent/determinant can be negative if the mapping $\mathbb{R}^d \times \mathbb{R}^d \ni (u,v) \mapsto C(u-v)$ is not positive semi-definite. When this mapping is positive semi-definite, C is an auto-covariance function, with corresponding auto-correlation function $R(x) = C(x)/C(0)$ provided $C(0) > 0$.

A locally finite point process $X \subset \mathbb{R}^d$ has n-th order joint intensity $\rho_X^{(n)}$ for $n = 1, 2, \dots$ if for any bounded and pairwise disjoint Borel sets $A_1, \dots, A_n \subset \mathbb{R}^d$,

$$\mathrm{E}\left[N(A_1) \cdots N(A_n)\right] = \int_{A_1} \int_{A_n} \rho_X^{(n)}\left(G_1, \dots, G_n\right) \mathrm{d}G_1 \cdots \mathrm{d}G_n < \infty.$$

Note that $\rho_X^{(n)}$ is unique except for a Lebesgue nullset in $\mathbb{R}^{dn}$ (we ignore nullsets in the following). Thus, if X is stationary, $\rho_X^{(1)}$ is constant and agrees with the intensity ρ_X, and $\rho_X > 0$ implies that $g_X(u-v) = \rho_X^{(2)}(u,v)/\rho_X^2$ is the PCF.

If for all $n = 1, 2, \dots$, the n-th order joint intensity exists and is given by

$$\rho_X^{(n)}(G_1, \dots, G_n) = \operatorname{per}_\alpha[C](G_1, \dots, G_n)$$

we say that X is a stationary α-weighted permanental point process with kernel C and write $X \sim \mathrm{PPP}_\alpha(C)$. Conditions are need to ensure the existence of $\mathrm{PPP}_\alpha(C)$, see Shirai and Takahashi (2003) and McCullagh and Møller (2006). To exclude the trivial case where X is empty we assume $\alpha C(0) > 0$. Note that C must be an auto-covariance function and $\alpha > 0$ because $\rho_X = \alpha C(0)$ and

$$g_X(x) - 1 = R(x)^2/\alpha. \tag{A.1}$$

This reflects that the process exhibits a positive association between its points. In fact, if C is an auto-covariance function and $k = 2\alpha$ is a positive integer, then $X \sim \mathrm{PPP}_\alpha(C)$ exists and it is a Cox process: Conditional on IID zero-mean stationary Gaussian processes $\Phi_1, \dots, \Phi_k$ on $\mathbb{R}^d$ with auto-covariance function $C/2$, we can let X be a Poisson process with intensity function $\Lambda(x) = \Phi_1(x)^2 + \dots + \Phi_k(x)^2$, $x \in \mathbb{R}^d$. In particular, if $\alpha = 1$, then X is the boson process introduced by Macchi (1975).

If for all $n = 1, 2, \dots$, the n-th order joint intensity exists and is given by

$$\rho_X^{(n)}(G_1, \dots, G_n) = \det_\alpha[C](G_1, \dots, G_n)$$

we say that X is a stationary α-weighted determinantal point process with kernel C and write $X \sim \mathrm{DPP}_\alpha(C)$. To exclude the trivial case where X is empty we assume $\alpha C(0) > 0$. Again C needs to be an auto-covariance function and $\alpha > 0$ because $\rho_X = \alpha C(0)$ and

$$g_X(x) - 1 = -R(x)^2/\alpha. \tag{A.2}$$

If $\alpha = 1$, then X is the fermion process introduced by Macchi (1975) (it is usually called the determinantal point process). We have the following existence result: If C is continuous and square integrable, existence of $X \sim \mathrm{DPP}_1(C)$ is equivalent to that the Fourier transform of C is bounded by 0 and 1 (Lavancier et al. (2015)). When α is a positive integer, $X \sim \mathrm{DPP}_\alpha(C)$ can be identified with the superposition $G_1 \cup \dots \cup G_\alpha$ of independent processes $G_i \sim \mathrm{DPP}_\alpha(C/\alpha)$, $i = 1, \dots, \alpha$. In general, the process is not well-defined if $0 < \alpha < 1$, cf. McCullagh and Møller (2006).

B The intensity and PCF of the invariant distribution

Let the situation be as in Theorem 4.2. Below we verify (4.2) and (4.6).

Note that the G_n are identically distributed and $G_0^{\mathrm{st}} = W_0^{\mathrm{st}} \cup Z_0$ where $W_0^{\mathrm{st}} = \bigcup_{m=1}^\infty \bigcup_{x \in Z_{-m}} W_{0,x}$, cf. (4.7). Hence, for Borel sets $A \subseteq \mathbb{R}^d$ with $|A| < \infty$,

$$\begin{aligned} \mathrm{E}\{\#(W_0^{\mathrm{st}} \cap A)\} &= \int \rho_Z \left\{ \sum_{m=1}^\infty \int_A (\beta p)^m f^{*m}(y - x) \, \mathrm{d}y \right\} \mathrm{d}x \\ &= |A| \sum_{m=1}^\infty \rho_Z (\beta p)^m = |A| \rho_Z \frac{\beta p}{1 - \beta p} \end{aligned} \tag{B.1}$$

using Fubini's theorem in the second identity, so W_0^{st} has intensity

$$\rho_W = \rho_Z \frac{\beta p}{1 - \beta p} \tag{B.2}$$

whereby it follows that G_0^{st} has intensity ρ_G as given by (4.2).

Let $A_1, A_2 \subseteq \mathbb{R}^d$ be disjoint Borel sets with $|A_i| < \infty$, $i = 1, 2$. Using similar arguments as in (B.1) and exploiting the fact that $Z_0, Z_{-1}, \dots$ are IID point processes with a PCF of the form $g_Z = 1 + bf_Z * \tilde{f}_Z$ as well as the independence between Z_0 and W_0^{st}, we obtain

$$
\begin{aligned}
&\mathrm{E}\{\#(G_0^{\mathrm{st}} \cap A_1)\#(G_0^{\mathrm{st}} \cap A_2)\} \\
&\quad = \rho_Z^2|A_1||A_2| + \rho_Z^2 \int_{A_1}\int_{A_2} bf_Z * \tilde{f}_Z(x_1 - x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2 + 2\rho_Z\rho_W|A_1||A_2| && \text{(B.3)} \\
&\qquad + \sum_{m_1=1}^{\infty}\sum_{m_2=1:\,m_1\neq m_2}^{\infty} \rho_Z^2(\beta p)^{m_1+m_2}|A_1||A_2| && \text{(B.4)} \\
&\qquad + \sum_{m=1}^{\infty}\rho_Z^2(\beta p)^{2m}|A_1||A_2| \\
&\qquad + \sum_{m=1}^{\infty}\rho_Z^2(\beta p)^{2m}\int_{A_1}\int_{A_2} f_Z * \tilde{f}_Z * f^{*m} * \tilde{f}^{*m}(y_1 - y_2)\,\mathrm{d}y_1\,\mathrm{d}y_2 && \text{(B.5)} \\
&\qquad + \sum_{m=1}^{\infty}\mathrm{E}\Big\{\sum_{x\in Z_{-m}} \#(W_{0,x} \cap A_1)\#(W_{0,x} \cap A_2)\Big\} && \text{(B.6)}
\end{aligned}
$$

Here, the first two term of (B.3) corresponds to pairs of points from Z_0 with one point falling in A_1 and the other in A_2, the second term corresponds to pairs of points either from $Z_0 \cap A_1$ and $W_0^{\mathrm{st}} \cap A_2$ or from $Z_0 \cap A_2$ and $W_0^{\mathrm{st}} \cap A_1$. Moreover, the term in (B.4) corresponds to pairs of points, with one point falling in A_1 and the other in A_2 of two families initiated by ancestors from different generations, while the term in (B.5) corresponds to such pairs of points in two different families initiated by ancestors from the same generation, and finally the term in (B.6) corresponds to pairs of points from the same family, falling in A_1 and A_2, respectively. Using (4.3) and (B.2), we observe that (B.3)–(B.5) simplify to

$$
\rho_G^2 + \sum_{m=0}^{\infty}\rho_Z^2(\beta p)^{2m}\int_{A_1}\int_{A_2} f_Z * \tilde{f}_Z * f^{*m} * \tilde{f}^{*m}(y_1 - y_2)\,\mathrm{d}y_1\,\mathrm{d}y_2,
$$

whilst the term in (B.6) is equal to

$$
\begin{aligned}
&\sum_{m=1}^{\infty}\int \rho_Z \sum_{i=0}^{m-1}\int (\beta p)^i f^{*i}(y - x)c\beta^2p^2 \\
&\qquad \cdot \int_{A_1}\int_{A_2}(\beta p)^{2(m-1-i)} f^{*(m-i)}(y_1 - y) f^{*(m-i)}(y_2 - y)\,\mathrm{d}y_1\,\mathrm{d}y_2\,\mathrm{d}y\,\mathrm{d}x
\end{aligned} \tag{B.7}
$$

where y corresponds to a i-th generation point in the family initiated by $x \in Z_{-m}$, and where $c\beta^2p^2$ is the expected number of pairs of points y_1 and y_2 which are $(m-1-i)$-th generation points of that ancestor. By Fubini's theorem, (B.7) reduces to

$$
\begin{aligned}
&\rho_Z c \sum_{m=1}^{\infty}\sum_{i=0}^{m-1}(\beta p)^{2m-i}\int_{A_1}\int_{A_2} f^{*(m-i)} * \tilde{f}^{*(m-i)}(y_1 - y_2)\,\mathrm{d}y_1\,\mathrm{d}y_2 \\
&\qquad = \rho_G c \sum_{k=1}^{\infty}(\beta p)^{2k}\int_{A_1}\int_{A_2} f^{*k} * \tilde{f}^{*k}(y_1 - y_2)\,\mathrm{d}y_1\,\mathrm{d}y_2
\end{aligned}
$$

where (4.3) has been used. Combining these results we finally see that G_0^{st} has PCF g_G as given by (4.6).

C Simulating the limiting process

This appendix presents an approximate simulation procedure for simulating G_0^{st} on a bounded region $R \subset \mathbb{R}^d$. It is available in `R` through the package `icpp`, which can be obtained at `https://github.com/adchSTATS/icpp`. The implementation utilizes existing functions from the packages `spatstat` and `RandomFields` to simulate the noise process.

We make the following assumptions. Let the situation be as in Theorem 4.2 and let $f \sim N_d(\sigma^2)$ with $\sigma > 0$. Also, without loss of generality, assume no thinning (i.e. $p = 1$). Let $R_{\oplus r} = \{\xi \in \mathbb{R}^d : b(\xi, r) \cap R \neq \emptyset\}$ where $b(\xi, r)$ is a closed ball with centre ξ and radius $r \geq 0$. Denote n the number of iterations in our approximate simulation algorithm, that is, $-n$ is the starting time when ignoring what happens previously. Note that $\sqrt{n}\sigma$ is the standard deviation of the nth convolution power of f. To account for edge effects, let $r = 4\sqrt{n}\sigma$ where 4 is an arbitrary non-negative value ensuring that a point of $G_{-n}^{\mathrm{st}} \setminus R_{\oplus r}$ would generate a nth generation offspring in R with very low probability, at most 1/15787. In the approximate simulation procedure, we ignore those points of $G_0^{\mathrm{st}} \cap R$ which are generated by an ith generation ancestor x when $i < -n$ or both $-n \leq i < 0$ and $x \notin R_{\oplus 4\sqrt{-i}\sigma}$. This is our algorithm in pseudocode where "parallel-for" means a parallel for loop:

---

parallel-for $i = -n$ to 0 **do**
    simulate $Z_i' := Z_i \cap R_{\oplus 4\sqrt{-i}\sigma}$
end parallel-for
set $O := Z_{-n}'$
if $n \neq 0$ **then**
    for $i = -(n-1)$ to 0 **do**
        parallel-for $x \in O$ **do**
            simulate the 1st generation offspring process, O_x, with parent x
        end parallel-for
        set $O := Z_i' \bigcup \left(\bigcup_{x \in O} O_x \cap R_{\oplus 4\sqrt{-i}\sigma}\right)$
    end for
end if
return O

Note that $\rho_Z \sum_{i=0}^{n} (\beta p)^i$ is the intensity of the stationary point process obtained by ignoring those points of G_0^{st} which are generated by an ith generation ancestor with $i < -n$. We base the choice of n on this fact by considering a precision parameter $\varepsilon > 0$ and letting

$$n = \sup\left\{m \in \{1, 2, \ldots\} : \left\| \rho_Z \sum_{i=0}^{m} (\beta p)^i - \rho_G \right\| \leq \varepsilon \right\}.$$

To exemplify, let $\rho_G = 100$ and $\beta p = 0.8$ implying that $\rho_Z = 20$, and let $\varepsilon = 2.22 \times 10^{-16}$, then $n = 159$. If instead $\beta p = 0.99$, then $n = 3609$.

Acknowledgements

Supported by The Danish Council for Independent Research | Natural Sciences, grant DFF – 7014-00074 "Statistics for point processes in space and beyond", and by the "Centre for Stochastic Geometry and Advanced Bioimaging", funded by grant 8721 from the Villum Foundation. We thank Ina Trolle Andersen, Yongtao Guan, Ute Hahn, Eva B. Vedel Jensen, and Morten Nielsen for helpful comments.

References

I. T. Andersen, U. Hahn, E. A. Christensen, L. N. Nejsum, and E. B. V. Jensen. Double Cox cluster processes – with applications to photoactivated localization microscopy. 2017. In preparation.

O.E Barndorff-Nielsen, J. Kent, and M. Sørensen. Normal variance-mean mixtures and z-distributions. *International Statistical Review*, 50:145–159, 1982.

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

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

J. F. C. Kingman. Remarks on the spatial distribution of a reproducing population. *Journal of Applied Probability*, 14:577–583, 1977.

Frédéric Lavancier, Jesper Møller, and Ege Rubak. Determinantal point process models and statistical inference. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 77:853–877, 2015.

Lieshout, M. N. M. van and A. J. Baddeley. Extrapolating and interpolating spatial patterns. In A. B. Lawson and D. Denison, editors, *Spatial Cluster Modelling*, pages 61–86. Chapman & Hall/CRC, Boca Raton, Florida, 2002.

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

B. Matérn. *Spatial Variation*, Meddelanden från Statens Skogforskningsinstitut, **49** (5). 1960.

B. Matérn. *Spatial Variation.* Lecture Notes in Statistics 36, Springer-Verlag, Berlin, 1986.

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

J. Møller. Random tessellations in $\mathbf{R}^d$. *Advances in Applied Probability*, 21:37–73, 1989.

J. Møller. *Lectures on Random Voronoi Tessellations.* Lecture Notes in Statistics 87. Springer-Verlag, New York, 1994.

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

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

J. Møller and G. L. Torrisi. The pair correlation function of spatial hawkes processes. *Statistics and Probability Letters*, 77:995–1003, 2007.

Jesper Møller and Rasmus Waagepetersen. *Statistical Inference and Simulation for Spatial Point Processes.* Chapman & Hall/CRC, Boca Raton, Florida, 2004.

Mari Myllymäki, Tomáš Mrkvička, Pavel Grabarnik, Henri Seijo, and Ute Hahn. Global envelope tests for spatial processes. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 79:381 – 404, 2016.

J. Neyman and E. L. Scott. Statistical approach to problems of cosmology. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 20:1–43, 1958.

I. K. Shimatani. Spatially explicit neutral models for population genetics and community ecology: Extensions of the Neyman-Scott clustering process. *Theoretical Population Biology*, 77:32–41, 2010.

T. Shirai and Y. Takahashi. Random point fields associated with certain Fredholm determinants I: fermion, Poisson and boson point processes. *Journal of Functional Analysis*, 205:414–463, 2003.

M. Thomas. A generalization of Poisson's binomial limit for use in ecology. *Biometrika*, 36:18–25, 1949.

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