J. Appl. Prob. **55,** 789–809 (2018)
*doi:*10.1017/jpr.2018.50
© *Applied Probability Trust* 2018

PAIR CORRELATION FUNCTIONS AND LIMITING DISTRIBUTIONS OF ITERATED CLUSTER POINT PROCESSES

JESPER MØLLER * ** AND
ANDREAS D. CHRISTOFFERSEN,* *Aalborg University*

Abstract

We consider a Markov chain of point processes such that each state is a superposition of an independent cluster process with the previous state as its centre process together with some independent noise process and a thinned version of the previous state. The model extends earlier work by Felsenstein (1975) and Shimatani (2010) describing a reproducing population. We discuss when closed-form 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

2010 Mathematics Subject Classification: Primary 60G55; 60J20
Secondary 60D05; 62M30

1. Introduction

In this paper we address a discrete-time Markov chain of point processes $G_0, G_1, \ldots$ 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 subset 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 $\mathbb{R}^d$; see, e.g. [3] or [17] for details. 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 $n = 1, 2, \ldots$, conditional on the previous generations $G_0, \ldots, G_{n-1}$, we obtain G_n by four basic operations for point processes.

(a) *Independent clustering.* To each point $x \in G_{n-1}$ is associated a (noncentred) cluster $Y_{n,x} \subset \mathbb{R}^d$. These clusters are independent and identically distributed (i.i.d.) 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, finite variance ν_n, and is independent of the points in $Y_{n,x}$, which are i.i.d. with each point

Received 20 November 2017; revision received 2 July 2018.
* Postal address: Department of Mathematical Sciences, Aalborg University, Skjernvej 4A, 9220 Aalborg Øst, Denmark.
** Email address: jm@math.aau.dk

https://doi.org/10.1017/jpr.2018.50 Published online by Cambridge University Press

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 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 i.i.d. Bernoulli variables, which are independent of Y_n, $G_0, \ldots, G_{n-1}$, and all previously generated Bernoulli variables. Let $p_n = \mathbb{P}(B_{n,y} = 1)$. For all $x \in G_{n-1}$, let

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

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

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

be the independent p_n-thinned point process of Y_n. Note that with probability 1, $W_n \cap G_{n-1} = \varnothing$, since, by assumption on the cluster points, the origin is not contained in $Y_{n,x}$.

(c) *Independent retention.* For all $x \in \mathbb{R}^d$, let $Q_{n,x}$ be i.i.d. Bernoulli variables which are independent of Y_n, $G_0, \ldots, G_{n-1}$, and all previously generated Bernoulli variables. Let $q_n = \mathbb{P}(Q_{n,x} = 1)$, and

$$G_{n-1}^{\text{thin}} = \{x \in G_{n-1} \colon Q_{n,x} = 1\}$$

be the independent q_n-thinned point process of G_{n-1}.

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

$$G_n = W_n \cup G_{n-1}^{\text{thin}} \cup Z_n, \tag{1.3}$$

where we interpret Z_n as noise. For ease of presentation, we assume with probability 1 that $W_n \cup G_{n-1}^{\text{thin}}$ and Z_n are disjoint. Thus, W_n, G_{n-1}^{thin}, and Z_n are pairwise disjoint almost surely.

When we later interpret our results, for any point $x \in G_{n-1}^{\text{thin}}$, since $x \in G_{n-1} \cap G_n$, we consider x both as its own ancestor and its own child.

Our model is an extension of the model of Shimatani [20], which, in turn, is an extension of Malécot's model [4] (we return to this in Section 2, items (vii) and (viii)). In particular, our extension allows us to model cluster centres exhibiting clustering or regularity, points from previous generations can be retaining, and the noise processes can also exhibit clustering or regularity (that is, they are not assumed to be Poisson processes). 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 items (a), (b), and (d) are unobservable. Our model may be of relevance for applications in population genetics and community ecology (see [20] and the references therein), for analyzing tropical rain forest point pattern data with multiple scales of clustering (see [24]), and for modelling proteins with multiple noisy appearances in photoactivated localization microscopy (see [1]). However, the study of the statistical applications of our model and results is beyond the scope of this paper.

The paper is organized as follows. A discussion of the assumptions in items (a)–(d) and the related literature is presented in Section 2. In Section 3 we focus on the first- and second-order

https://doi.org/10.1017/jpr.2018.50 Published online by Cambridge University Press

moment properties of G_n, that is, its intensity and pair correlation function (PCF); we extend the model cases and results of Shimatani [20] and show that tractable model cases for the PCF of G_0 are meaningful in terms of Poisson and other point processes, including weighted permanental and weighted determinantal point processes (which was not observed in [20]). In Section 4 we discuss 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 [20]. 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, in Appendix A we provide background material on weighted permanental and determinantal point processes, and in Appendix B we specify an algorithm for approximate simulation of the Markov chain's invariant distribution.

2. Assumptions and related work

In items (i)–(iv) below, we comment on the model assumptions in items (a)–(d).

(i) The process Y_n is a stationary independent cluster process [3] and we have the following special cases: if G_{n-1} is a stationary Poisson process, Y_n is a Neyman–Scott process [19]; 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); (see [13]) 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 [22] if f_n is the density of d i.i.d. 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 (see [8] and [9]) if instead f_n is a uniform density of a d-dimensional ball with its centre at the origin. However, in many applications a Poisson centre process is not appropriate. For instance, van Lieshout and Baddeley [23] 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 $\beta_n \leq \nu_n$, we may consider Y_n as a stationary generalised shot-noise Cox process (GSNCP; see [15]). 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

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

where k_n is a PDF on $\mathbb{R}^d$, the γ_y and b_y for all $y \in G_{n-1}$ are independent positive random variables 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 $\nu_n = \mathbb{E}\{\mathrm{var}(\#Y_{n,x} \mid \gamma_y)\} + \mathrm{var}(\mathbb{E}\{\#Y_{n,x} \mid \gamma_y\}) = \beta_n + \mathrm{var}(\gamma_n)$). Further, b_y has an interpretation as a random bandwidth and

$$f_n(x) = \mathbb{E}\left\{\frac{k_n(x/b_y)}{b_y^d}\right\}.$$

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 1) if $\rho_{Z_n} = 0$. The $\rho_{Z_n} > 0$ case may be relevant when not all points in a generation can be described as resulting from

https://doi.org/10.1017/jpr.2018.50 Published online by Cambridge University Press

independent clustering and thinning as in (a)–(c). Note that in item (d) we could, without loss of generality, assume that $Z_1, Z_2, \ldots$ are independent. Further, we introduce the thinning of Y_n in item (b) only for modelling purposes and for comparison with [20]; from a mathematical point of view this 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 $Y_{n,x}^{\text{thin}}$ appear so that $\#Y_{n,x}^{\text{thin}}$ has mean $\beta_n^{\text{thin}} = \beta_n p_n$ and variance $\nu_n^{\text{thin}} = \beta_n p_n - \beta_n p_n^2 + \nu_n p_n^2$, and is independent of the points in $Y_{n,x}^{\text{thin}}$ which are i.i.d. with PDF f_n, whereby W_n and $Y_n^{\text{thin}} := \bigcup_{x \in G_{n-1}} (x + Y_{n,x}^{\text{thin}})$ are identically distributed.

(iv) Assuming there is no thinning of Y_n ($p_n = 1$) for $n = 1, 2, \ldots$, an equivalent description of items (a), (c), and (d) can be given in terms of the Voronoi tessellation generated by G_{n-1}, that is, for $x \in G_{n-1}$, let $C(x \mid G_{n-1})$ 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 any other point in G_{n-1} (with respect to usual distance in $\mathbb{R}^d$). With probability 1, since G_{n-1} is stationary and nonempty, each Voronoi cell is bounded and, hence, its volume is finite; see, e.g. [11] and [12]. Thus, we can set

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

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

In items (v) and (vi) below, we discuss earlier work on the model for $G_0, G_1, \ldots$, where G_0 is a stationary Poisson process, all $G_n = Y_n$ for $n \geq 1$ (that is, no thinning, no retention, and no noise), and $f_n = f$ 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) This replicated SNCP was originally studied by Malécot; see the discussion and references in Felsenstein [4], where three conditions were 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 the author also considered more general offspring densities f and the $d = 2, 3$ cases. Felsenstein [4] also noted that (I) is incompatible with (II) and (III) since $G_1, G_2, \ldots$ are not stationary Poisson processes and the use of (II) and (III) in a model will give rise to the formation of increasingly larger clumps of individuals separated by increasingly greater distances, concluding that it is irrelevant from a biological modelling viewpoint.

(vi) Kingman [5] considered the case where β is replaced by a nonnegative function b which is allowed to depend on the cluster centre x and the previous generation, so a cluster with its centre at 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.

https://doi.org/10.1017/jpr.2018.50 Published online by Cambridge University Press

Then G_n is a Cox process: G_n conditional on G_{n-1} is a Poisson process with intensity function

$$\Lambda_n(x) = \sum_{y \in G_{n-1}} b(y, G_{n-1}) f(x - y), \qquad x \in \mathbb{R}^d. \tag{2.2}$$

In this setting, Kingman verified that it is impossible for G_n to be a stationary Poisson process, however, replacing $f(x-y)$ in (2.2) by a more general density, which may depend on $G_{n-1}-x$, the author noted 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, Shimantani [20] considered first the case of items (a) and (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, Shimantani [20] considered the $f \sim N_2(\sigma^2)$ case and when $\beta p = 1$ or, equivalently, when the intensities $\rho_{G_0} = \rho_{G_1} = \cdots$ are invariant over generations. He then showed that as $n \to \infty$, the PCF for G_n diverges. 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, the result in [20] agrees with the results in [4] and [5]. Note that Shimantani 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 Sections 3.2.2 and 4; see also [5, Section 3].

Next, Shimantani [20] extended the model by including noise as in item (d) by making the following assumptions. The noise processes Z_n are stationary Poisson processes, satisfying $0 < \rho_{Z_1} = \rho_{Z_2} = \cdots$ and $\rho_{G_0} = \rho_{G_1} = \cdots$, 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, Shimantani [20] showed that the PCF of G_n converges uniformly as $n \to \infty$ and argued that this limiting case may be 'biologically valid'. However, we address some points arising from [20].

- The author did not show that there exists an underlying point process having this limiting case as its PCF, although it was claimed that the modified replicated Neyman–Scott process was able to reach 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)), the author did not argue as to the existence of an underlying point process nor the form it might take.

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.* [24] studied this in the special case of a double Thomas cluster process G_2 when $d = 2$, that is, 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 [1] for more general functions f_n. Moreover, the authors extended the double Thomas process to the case where $\rho_{Z_1} = 0$ and $\rho_{Z_2} > 0$; this type of model was also considered in [1]. In any case, our general results for intensities and PCFs in Section 3 will cover all these cases.

https://doi.org/10.1017/jpr.2018.50 Published online by Cambridge University Press

(x) If for each generation we assume no thinning ($p_1 = p_2 = \cdots = 1$), no noise ($\rho_{Z_1} = \rho_{Z_2} = \cdots = 0$), no retention ($q_1 = q_2 = \cdots = 0$), as well as $\beta_1 = \beta_2 = \cdots$ and $f_1 = f_2 = \cdots$, then the superposition $\bigcup_{n=0}^{\infty} G_n$ is known as a spatial Hawkes process; see [16] 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 [20]. Specifically, points from one generation can be retained in the next generation, the noise is an arbitrary stationary point process (not necessarily a stationary Poisson process as in [20]), and we do not assume the same reproduction system.

3.1. Intensities

By induction, G_n is seen to be stationary for $n = 0, 1, \ldots$. We determine its intensity 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 $n = 1, 2, \ldots$, it holds that G_n is stationary with a positive and finite intensity*

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

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

3.2. PCFs

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 $(u, v) \to g_X(u - v)$ with $(u, v) \in \mathbb{R}^d \times \mathbb{R}^d$ if for any bounded Borel function $h : \mathbb{R}^d \times \mathbb{R}^d \to [0, \infty)$ with compact support,

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

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

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

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 the reduced PCF $g_X - 1$ rather than g_X, and $g_X = 1$ if X is a Poisson process. It is convenient when g_X is isotropic, meaning that there is a function $g_{X,o}$ so that for all $x \in \mathbb{R}^d$, $g_X(x) = g_{X,o}(\|x\|)$ depends only on x through $\|x\|$. 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

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

https://doi.org/10.1017/jpr.2018.50 Published online by Cambridge University Press

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

$$g_X - 1 = ah * \tilde{h} \quad \text{for real constants } a, \tag{3.4}$$

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 [20], for the initial generation process G_0, Shimatani briefly discussed the special case of (3.4) when $h \sim N_2(\frac{1}{2}\tau^2)$ (so $h * \tilde{h} \sim N_2(\tau^2)$) whilst the noise processses are stationary Poisson processes. However, if $a \neq 0$ the author 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 need 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 [2]. Also, generalized shot-noise Cox processes [16] have PCFs of the form (3.4) with $a > 0$. Moreover, (3.4) holds for many other cases of point process models for X: if the Fourier transform $\mathcal{F}(g_X - 1)$ is well defined and nonnegative, if $h = \tilde{h}$, and if $a := \int (g_X - 1) \in (0, \infty)$, then (3.4) holds with

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

provided this inverse transform is well defined. Extensions of (3.4) are discussed in Section 3.2.4.

We will need the following lemma in Section 3.2.3.

Lemma 3.1. *Suppose that g_X is of the form (3.4). Then*

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

for any integrable real function f defined on $\mathbb{R}^d$ and for any $u, v \in \mathbb{R}^d$.

Proof. This 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 3.1, which we verify in Section 3.2.3. We use the following notation. Define

$$c_n = \frac{\mathbb{E}\{\#Y_{n,x}(\#Y_{n,x} - 1)\}}{\beta_n^2} = \frac{\nu_n + \beta_n^2 - \beta_n}{\beta_n^2} \quad \text{if } \beta_n > 0, \tag{3.5}$$

with $c_n = 0$ if $\beta_n = 0$. If $\beta_n = \nu_n > 0$, as in the case when $\#Y_{n,x}$ follows a (nondegenerated) Poisson distribution, then $c_n = 1$. The case of overdispersion (underdispersion), that is, $\nu_n > \beta_n$ ($\nu_n < \beta_n$) corresponds to $c_n > 1$ ($c_n < 1$). Denote by δ_0 the Dirac delta function defined on $\mathbb{R}^d$. Recall that for any integrable real function h defined on $\mathbb{R}^d$, $h * \delta_0 = \delta_0 * h = h$, and for any $a \in \mathbb{R}$, $a\delta_0 * a\delta_0 = a^2\delta_0$, where we understand $0\delta_0$ as 0. Finally, let

$$\ast_{i=1}^{n} h_i = h_1 * \cdots * h_n,$$

where each h_i is either of the form $a_i\delta_0$ with a_i a real constant, or it is an integrable real function defined on $\mathbb{R}^d$.

https://doi.org/10.1017/jpr.2018.50 Published online by Cambridge University Press

Theorem 3.1. *Suppose that g_{G_0} and 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 $n = 1, 2, \ldots$,*

$$g_{G_n}(u) - 1 = \left(\frac{\rho_{G_0}}{\rho_{G_n}}\right)^2 af_0 * \tilde{f}_0 * \overset{n}{\underset{i=1}{\ast}} \{(\beta_i p_i f_i + q_i\delta_0) * (\beta_i p_i \tilde{f}_i + q_i\delta_0)\}(u) \tag{3.6}$$

$$+ \sum_{i=1}^{n} \frac{\rho_{G_{i-1}}}{\rho_{G_n}^2} \{c_i (b_i p_i)^2 f_i * \tilde{f}_i + \beta_i p_i q_i (f_i + \tilde{f}_i)\}$$

$$* \overset{n}{\underset{j=i+1}{\ast}} \{(\beta_j p_j f_j + q_j\delta_0) * (\beta_j p_j \tilde{f}_j + q_j\delta_0)\}(u) \tag{3.7}$$

$$+ \sum_{i=1}^{n-1} \left(\frac{\rho_{Z_i}}{\rho_{G_n}}\right)^2 b_i f_{Z_i} * \tilde{f}_{Z_i}$$

$$* \overset{n}{\underset{j=i+1}{\ast}} \{(\beta_j p_j f_j + q_j\delta_0) * (\beta_j p_j \tilde{f}_j + q_j\delta_0)\}(u) \tag{3.8}$$

$$+ \left(\frac{\rho_{Z_n}}{\rho_{G_n}}\right)^2 b_n f_{Z_n} * \tilde{f}_{Z_n}(u), \tag{3.9}$$

where the $(n-i)$th fold convolution in (3.7) is interpreted as 1 if $i = n$ and the sum in (3.8) is interpreted as 0 if $n = 1$.

The terms in (3.6)–(3.9) have the following interpretation when $u \neq 0$: the right-hand side of (3.6) corresponds to pairs of nth generation points with different 0th generation ancestors; the ith term in (3.7) corresponds to pairs of nth generation points when they have a common $(i-1)$th generation ancestor initiated by Z_{i-1} if $i > 1$ or by G_0 if $i = 1$; the ith term in (3.8) corresponds to pairs of nth generation points with different ith generation ancestors initiated by Z_i and (3.9) 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 noting the flexibility of our model for G_1 and the effect of the choice of its centre process G_0: for simplicity, suppose there is no noise and no retention of points from one generation to the next (that is, $\rho_{Z_n} = q_n = 0$ for $n = 1, 2, \ldots$), and G_0 is stationary and either a Poisson, a weighted determinantal, or permanental point process with a Gaussian kernel. Specifically, suppose that $d = 2$, G_0 has intensity $\rho_{G_0} = 100$, and using the 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 the largest possible to ensure existence of the process [6]; for the α-weighted permanental point process, $\alpha = \frac{1}{2}$ (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 $f_0 \sim N_2(\frac{1}{8}\tau^2)$, which by (A.1) and (A.2) means 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 3.1,

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

https://doi.org/10.1017/jpr.2018.50 Published online by Cambridge University Press

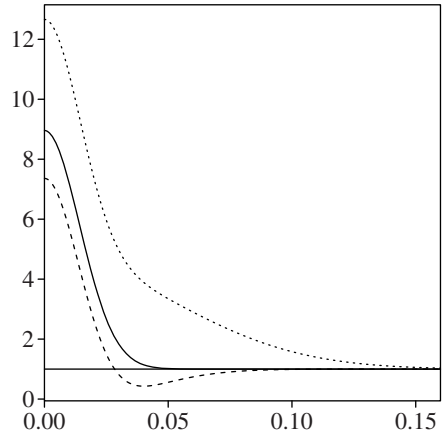


FIGURE 1: The PCFs of G_1 when G_0 is a determinantal (*dashed line*), Poisson (*solid line*), or weighted permanental (*dotted line*) point process, with parameters and Gaussian offspring PDF as specified in the text. The solid horizontal line is the PCF for a Poisson process.

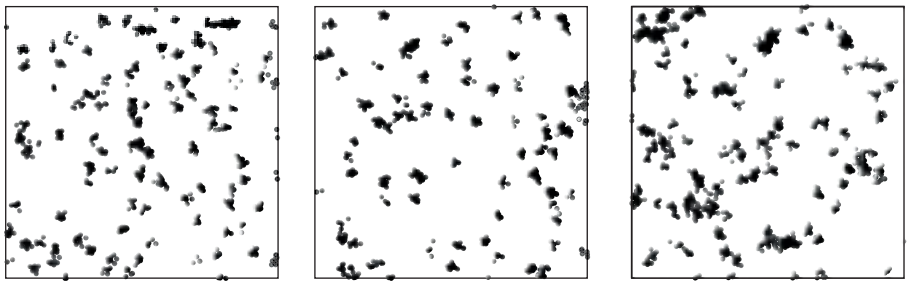


FIGURE 2: Simulations of G_1 restricted to a unit square when G_0 is a determinantal (*left panel*), Poisson (*middle panel*), or weighted permanental (*right panel*) point process; see Figure 1 and the text.

In Figure 1 we present the isotropic PCF $g_{G_1,o}(r) = g_{G_1}(u)$ as a function of the interpoint distance $r = \|u\|$ in the 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. In Figure 2 we present 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 (left panel) and a higher degree of clustering when G_0 is a weighted permanental point process (right panel). 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 since the spread of clusters then decreases.

3.2.3. *Proof of Theorem 3.1.* Shimatani [20] verified Theorem 3.1 in the special case where $q_1 = q_2 = \cdots = 0$, $b_1 = b_2 = \cdots = 0$ (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 (3.8) and (3.9) are 0. If $c_1 = c_2 = \cdots = 0$ then (3.7) is 0 and by (3.5), with probability 1, $\#Y_{n,x} \in \{0, 1\}$ for all $x \in G_{n-1}$ and $n = 1, 2, \ldots$. Consequently, the proof of Theorem 3.1 is trivial if $c_1 = c_2 = \cdots = 0$ and

https://doi.org/10.1017/jpr.2018.50 Published online by Cambridge University Press

both G_0 and $Z_1, Z_2, \ldots$ are stationary Poisson processes, since then $a = 0, b_1 = b_2 = \cdots = 0$, $G_1, G_2, \ldots$ are stationary Poisson processes, and the class of stationary Poisson processes is closed under i.i.d. random shifts of the points, thinning, and superposition. The general proof of Theorem 3.1 follows by induction using the next lemma together with Lemma 3.1.

Lemma 3.2. *If $\rho_{G_{n-1}} > 0$, $\rho_{G_n} > 0$, and $g_{G_{n-1}}$ and g_{Z_n} exist, then g_{G_n} also exists and is*

$$\begin{aligned} g_{G_n}(u-v) - 1 = {} & \left(\frac{\rho_{G_{n-1}}\beta_n p_n}{\rho_{G_n}}\right)^2 \Bigg[\iint \{g_{G_{n-1}}(x_1 - x_2) - 1\} f_n(u - x_1) f_n(v - x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2 \\ & + \frac{c_n}{\rho_{G_{n-1}}} f_n * \tilde{f}_n(u - v)\Bigg] \\ & + \left(\frac{\rho_{G_{n-1}} q_n}{\rho_{G_n}}\right)^2 \{g_{G_{n-1}}(u - v) - 1\} + \left(\frac{\rho_{Z_n}}{\rho_{G_n}}\right)^2 \{g_{Z_n}(u - v) - 1\} \\ & + \frac{\rho_{G_{n-1}}^2 \beta_n p_n q_n}{\rho_{G_n}^2}\Bigg[\int \{g_{G_{n-1}}(v - x) - 1\}\{f_n(u - x) + \tilde{f}_n(u - x)\}\,\mathrm{d}x \\ & + \frac{1}{\rho_{G_{n-1}}}\{f_n(u - v) + \tilde{f}_n(u - v)\}\Bigg] \quad \textit{for any } u, v \in \mathbb{R}^d. \end{aligned}$$

Proof. Note that Y_n is stationary with intensity

$$\rho_{Y_n} = \rho_{G_{n-1}}\beta_n. \tag{3.10}$$

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

$$\begin{aligned} \rho_{Y_n}^2 g_{Y_n}(u - v) = {} & \rho_{G_{n-1}}^2 \beta_n^2 \iint g_{G_{n-1}}(x_1 - x_2) f_n(u - x_1) f_n(v - x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2 \\ & + \rho_{G_{n-1}} c_n \beta_n^2 f_n * \tilde{f}_n(u - v) \quad \text{for any } u, v \in \mathbb{R}^d, \end{aligned} \tag{3.11}$$

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 (3.10), W_n is stationary with intensity

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

and PCF

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

where the first identify follows from the fact that PCFs are invariant under independent thinning, and where (3.11) is used to obtain the second identity. Also, for disjoint Borel sets $A_1, A_2 \subseteq \mathbb{R}^d$, it follows from items (a)–(c) that

$$\begin{aligned} &\mathbb{E}\{\#(W_n \cap A_1)\#(G_{n-1}^{\text{thin}} \cap A_2)\} \\ &\quad = \rho_{W_n}\rho_{G_{n-1}^{\text{thin}}} \int_{A_1}\int_{A_2} \left\{\frac{1}{\rho_{G_{n-1}}} f_n(x_1 - x_2) + g_{G_{n-1}} * \tilde{f}_n(x_1 - x_2)\right\} \mathrm{d}x_1\,\mathrm{d}x_2. \end{aligned}$$

https://doi.org/10.1017/jpr.2018.50 Published online by Cambridge University Press

Furthermore, by (1.3), (3.2), and Fubini's theorem, we readily see that G_n has the PCF

$$\rho_{G_n}^2 g_{G_n}(x) = \rho_{W_n}^2 g_{W_n}(x) + (\rho_{G_{n-1}}^{\text{thin}})^2 g_{G_{n-1}^{\text{thin}}}(x) + \rho_{Z_n}^2 g_{Z_n}(x) + 2\rho_{W_n}\rho_{Z_n} + 2\rho_{Z_n}\rho_{G_{n-1}}^{\text{thin}} + 2\rho_{W_n}\rho_{G_{n-1}^{\text{thin}}}\left\{ g_{G_{n-1}} * f_n(x) + \frac{1}{\rho_{G_{n-1}}} f_n(x) \right\},$$

where the six terms on the right-hand side correspond to pairs of points from W_n, Z_n, G_{n-1}^{thin}, W_n and Z_n, Z_n and G_{n-1}^{thin}, and W_n and G_{n-1}^{thin}, respectively, where the latter three cases can be ordered in two ways. Combining all this with the first identity in (3.1) together with (3.12), we easily obtain

$$\begin{aligned} g_{G_n}(u-v) - 1 = {} & \left(\frac{\rho_{G_{n-1}} \beta_n p_n}{\rho_{G_n}} \right)^2 \{ g_{W_n}(u-v) - 1 \} + \left(\frac{\rho_{Z_n}}{\rho_{G_n}} \right)^2 \{ g_{Z_n}(u-v) - 1 \} \\ & + \left(\frac{\rho_{G_{n-1}} q_n}{\rho_{G_n}} \right)^2 \{ g_{G_{n-1}}(u-v) - 1 \} \\ & + \rho_{W_n} \rho_{G_{n-1}^{\text{thin}}} \left[\int (g_{G_{n-1}}(v-x) - 1) \{ f_n(u-x) + \tilde{f}_n(u-x) \} \, \mathrm{d}x \right. \\ & \left. + \frac{1}{\rho_{G_{n-1}}} \{ f_n(u-v) + \tilde{f}_n(u-v) \} \right]. \end{aligned}$$

Combined with (3.13) this implies the result in Lemma 3.2. □

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, that is,

$$g_{G_0} - 1 = a_0 + a_1 f_{0,1} * \tilde{f}_{0,1} + \cdots + a_k f_{0,k} * \tilde{f}_{0,k}, \tag{3.14}$$

$$g_{Z_n} - 1 = b_{n,0} + b_{n,1} f_{Z_n,1} * \tilde{f}_{Z_n,1} + \cdots + b_{n,l} f_{Z_n,l} * \tilde{f}_{Z_n,l}, \qquad n = 1, 2, \ldots, \tag{3.15}$$

for real constants $a_0, \ldots, a_k, b_{n,1}, \ldots, b_{n,l}$ and PDFs $f_{0,1}, \ldots, f_{0,k}, f_{Z_n,1}, \ldots, f_{Z_n,l}$. For instance, the superposition of k independent Poisson, weighted permanental, or weighted determinantal point processes has a PCF of the form of (3.14) or (3.15). Also, Theorem 3.1 provides examples of PCFs of the form of (3.14) or (3.15). Assuming (3.14) and (3.15), Theorem 3.1 is immediately generalised by replacing $a f_0 * \tilde{f}_0$ in (3.6) by (3.14), $b_n f_{Z_n} * \tilde{f}_{Z_n}$ in (3.9) by (3.15), and similarly for $b_i f_{Z_i} * \tilde{f}_{Z_i}$ in (3.8).

4. Same reproduction system

Throughout this section we assume the same reproduction system over generations, that is, in items (a)–(d), $\beta_n = \beta$, $\nu_n = \nu$, $f_n = f$, $p_n = p$, and $q_n = q$ do not depend on n, $Z_1, Z_2, \ldots$ being i.i.d. stationary point processes, so $\rho_{Z_n} = \rho_Z$ for $n = 1, 2, \ldots$, 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 [20]. By (3.1), we have either

$$\text{if } \rho_Z = 0 \quad \text{then } \beta p + q = 1; \quad \text{if } \rho_Z > 0 \quad \text{then } \beta p + q < 1. \tag{4.1}$$

Thus, if $\rho_Z > 0$,

$$\rho_G = \frac{\rho_Z}{1 - \beta p - q}. \tag{4.2}$$

https://doi.org/10.1017/jpr.2018.50 Published online by Cambridge University Press

4.1. Limiting PCF

Under the assumptions above and in Theorem 3.1, setting $0^0 = 1$, we see that the PCF simplifies, after a straightforward calculation, to

$$\begin{aligned} g_{G_n}(u) - 1 = {} & af_0 * \tilde{f}_0 * \sum_{k_1=0}^{n}\sum_{k_2=0}^{n}\binom{n}{k_1}\binom{n}{k_2}q^{2n-k_1-k_2}(\beta p)^{k_1+k_2} f^{*k_1} * \tilde{f}^{*k_2}(u) \\ & + \left\{\frac{c(\beta p)^2 f * \tilde{f} + \beta pq(f+\tilde{f})}{\rho_G} + \left(\frac{\rho_Z}{\rho_G}\right)^2 bf_Z * \tilde{f}_Z\right\} \\ & * \sum_{i=0}^{n-1}\sum_{k_1=0}^{i}\sum_{k_2=0}^{i}\binom{i}{k_1}\binom{i}{k_2}q^{2i-k_1-k_2}(\beta p)^{k_1+k_2} f^{*k_1} * \tilde{f}^{*k_2}(u) \end{aligned} \tag{4.3}$$

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

$$c = \begin{cases} \dfrac{\nu + \beta^2 - \beta}{\beta^2} & \text{if } \beta > 0, \\ 0 & \text{if } \beta = 0, \end{cases}$$

f^{*n} is the nth convolution power of f if $n > 0$, and $f^{*0} * \tilde{f}^{*0} = \delta_0$. For instance, consider the case $f \sim N_d(\sigma^2)$ and $f_Z \sim N_d(\kappa^2)$, and suppose that $d \geq 3$ if $\rho_Z > 0$. Then the binomial formula combined with either choice in (4.1) implies that the first double sum in (4.3) tends to 0 as $n \to \infty$, and, hence,

$$\begin{aligned} g_G(u) - 1 := {} & \lim_{n\to\infty} g_{G_n}(u) - 1 \\ = {} & \frac{c}{\rho_G}\sum_{i=0}^{\infty}\sum_{k_1=0}^{i}\sum_{k_2=0}^{i}\binom{i}{k_1}\binom{i}{k_2}\frac{q^{2i-k_1-k_2}(\beta p)^{2+k_1+k_2}}{\{2\pi(2+k_1+k_2)\sigma^2\}^{d/2}} \\ & \qquad \times \exp\left\{-\frac{\|u\|^2}{2(2+k_1+k_2)\sigma^2}\right\} \\ & + \frac{2}{\rho_G}\sum_{i=0}^{\infty}\sum_{k_1=0}^{i}\sum_{k_2=0}^{i}\binom{i}{k_1}\binom{i}{k_2}\frac{q^{2i-k_1-k_2+1}(\beta p)^{1+k_1+k_2}}{\{2\pi(1+k_1+k_2)\sigma^2\}^{d/2}} \\ & \qquad \times \exp\left\{-\frac{\|u\|^2}{2(1+k_1+k_2)\sigma^2}\right\} \\ & + \left(\frac{\rho_Z}{\rho_G}\right)^2\sum_{i=0}^{\infty}\sum_{k_1=0}^{i}\sum_{k_2=0}^{i}\binom{i}{k_1}\binom{i}{k_2}\frac{q^{2i-k_1-k_2}(\beta p)^{k_1+k_2}}{[2\pi\{(k_1+k_2)\sigma^2+2\kappa^2\}]^{d/2}} \\ & \qquad \times \exp\left[-\frac{\|u\|^2}{2\{(k_1+k_2)\sigma^2+2\kappa^2\}}\right] \end{aligned} \tag{4.4}$$

is finite. Shimatani [20] showed only that this is finite under the assumptions $d = 2, b = q = 0$, and $c > 0$. The author also noted that $\beta p = 1$ and $\rho_Z = 0$ (which is the first choice in (4.1) for $q = 0$) imply divergence of g_{G_n} as $n \to \infty$, whilst $\beta p < 1$ and $\rho_Z > 0$ (which is the second choice in (4.1) for $q = 0$) imply convergence. Further, in the case of convergence and when $\beta p \approx 1$, the author discussed an approximation of $g_G(u)$ that depends on whether $\|u\|$ is close to 0 or not.

https://doi.org/10.1017/jpr.2018.50 Published online by Cambridge University Press

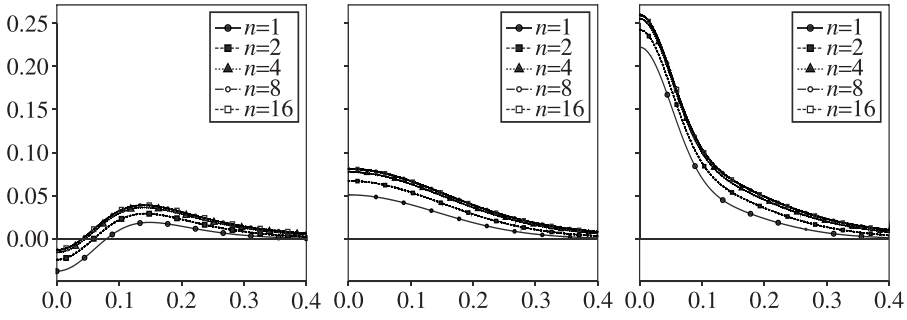


FIGURE 3: The reduced PCFs $g_{G_n} - 1$ when the noise processes are either determinantal, Poisson, or weighted permanental point processes (*left to right*), with parameters and Gaussian offspring PDF as specified in the text. The solid horizontal line is the PCF -1 for a Poisson process.

In general (that is, without making the assumption of normal distributions and so on), if we assume $g_{G_n} - 1$ has a finite limit and the second choice in (4.1) is satisfied, then

$$g_G(u) - 1 = \left\{ \frac{c(\beta p)^2 f * \tilde{f} + \beta p q (f + \tilde{f})}{\rho_G} + \left(\frac{\rho_Z}{\rho_G} \right)^2 b f_Z * \tilde{f}_Z \right\} * \sum_{i=0}^{\infty} \sum_{k_1=0}^{i} \sum_{k_2=0}^{i} \binom{i}{k_1} \binom{i}{k_2} q^{2i-k_1-k_2} (\beta p)^{k_1+k_2} f^{*k_1} * \tilde{f}^{*k_2}(u), \tag{4.5}$$

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

Considering the situation at the end of Section 3.2.2, assume that $d = 2$, $q = 0$, $f \sim N_d(\sigma^2)$, 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 (4.4), where $d = 2$, and κ^2 is replaced by $\frac{1}{8}\kappa^2$. 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 that $\kappa = 0.1$ in the case of weighted permanental noise and $\kappa = 1/\sqrt{\rho_Z \pi}$ in the case of determinantal noise (the most repulsive Gaussian determinantal point process). Shimatani [20] discussed the case where $\beta p = 0.99$—a plot (omitted here) indicates 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 since 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. In Figure 3 we further illustrate that it is only for small or moderate interpoint distances that the three limiting PCFs clearly differ.

4.2. Second main result

Although Shimatani [20] showed convergence of g_{G_n} in the special case considered above, the author 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 (4.5). In order to show that $G_0, G_1, \ldots$ does indeed converge

https://doi.org/10.1017/jpr.2018.50 Published online by Cambridge University Press

to a limiting distribution under more general conditions, and to specify what this is, we construct, in accordance with items (a)–(d), a Markov chain $\ldots, G^{\mathrm{st}}_{-1}, G^{\mathrm{st}}_0, G^{\mathrm{st}}_1, \ldots$ 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 (d). 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 a point $x \in Z_n$, consider the family of all generations initiated by the ancestor x, that is, the family

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

where

$$W^{(1)}_{n,x} = \begin{cases} W_{n,x} \cup \{x\} & \text{if } Q_{1,x} = 1, \\ W_{n,x} & \text{if } Q_{1,x} = 0, \end{cases}$$

is defined by the reproduction mechanism of independent clustering, independent thinning, and independent retention given in items (a)–(c) (with $\beta_n = \beta$ and $\nu_n = \nu$), $W^{(2)}_{n,x}$ is the offspring and retained points 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 nth 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, i.i.d.). Finally, for all integers n, let

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

It will be evident from the next theorem that any G^{st}_n has intensity ρ_G given by (4.2) and PCF g_G given by (4.5) (provided $g_G(u-v)$ is a locally integrable function of $(u,v) \in \mathbb{R}^d \times \mathbb{R}^d$); a formal proof can be found in [14, Appendix B]. The proof of Theorem 4.1 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 next result.

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

$$T^{\mathrm{st}}_{0,K} = \sup\{m \in \{1,2,\ldots\} \colon W^{(m)}_{0,x} \cap K \neq \varnothing \text{ for some } x \in G^{\mathrm{st}}_0\} \tag{4.6}$$

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

$$\mathbb{E}\{T^{\mathrm{st}}_{0,K}\} \leq |K| \rho_G \frac{\beta p + q}{1 - \beta p - q}$$

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

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

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

https://doi.org/10.1017/jpr.2018.50 Published online by Cambridge University Press

By the law of total expectation, conditioning on G_0^{st} and using Campbell's theorem, we obtain

$$\begin{aligned}\mathbb{E}\{N\} &= \rho_G \int \sum_{m=1}^{\infty} \int_K (\beta p f + q\delta_0)^{*m}(y-x)\,\mathrm{d}y\,\mathrm{d}x \\ &= \rho_G \int \sum_{m=1}^{\infty} \int_K \sum_{k=0}^{m} \binom{m}{k} q^{m-k} (\beta p)^k f^{*k}(y-x)\,\mathrm{d}y\,\mathrm{d}x \\ &= |K|\rho_G \frac{\beta p + q}{1-\beta p - q}\end{aligned}$$

using Fubini's theorem in the last identity. Further, the families initiated by the points in G_0^{st} are almost surely pairwise disjoint, so N is almost surely the number of points in K belonging to some family initiated by a point $x \in G_0^{\mathrm{st}}$. Consequently, $\mathbb{P}(T_{0,K}^{\mathrm{st}} \leq N) = 1$, whereby the lemma follows. □

We are now ready to state our second main result.

Theorem 4.1. *Suppose that* $\ldots, Z_{-1}, Z_0, Z_1 \ldots$ *are i.i.d. stationary point processes and* $\beta_n = \beta$, $\nu_n = \nu$, $f_n = f$, $p_n = p$, $q_n = q$, *and* $\rho_{Z_n} = \rho_Z$ *do not depend on* $n \geq 1$, *where* $\beta p + q < 1$ *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 with items (a)–(d). Let* Π *be the distribution of any* G_n^{st} *and* $\mathcal{N}$ *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 with items (a)–(d). 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 $\ldots, G_{-1}^{\mathrm{st}}, G_0^{\mathrm{st}}, G_1^{\mathrm{st}}, \ldots$ since $G_1^{\mathrm{st}}, G_2^{\mathrm{st}}, \ldots$ 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 $x \in G_{n-1} \cap G_{n-1}^{\mathrm{st}}$, the same Bernoulli variables $B_{n,y}$ for all times $n = 1, 2, \ldots$ and all offspring $y \in Y_{n,x}$ with ancestor $x \in G_{n-1} \cap G_{n-1}^{\mathrm{st}}$, and the same Bernoulli variables $Q_{n,x}$ for all times $n = 1, 2, \ldots$ and all retained points $x \in G_{n-1} \cap G_{n-1}^{\mathrm{st}}$. Let $K \subset \mathbb{R}^d$ be compact. In accordance with (4.6), for $\omega \in \mathcal{N}$, let

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

be the last time a point in K is a member of a family initiated by some point in ω, and let $\Omega = \{\omega \in \mathcal{N} \colon T_K(\omega) < \infty\}$. 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 $\omega \in \Omega$,

$$\lim_{n\to\infty} \mathbb{P}(G_n \cap K = \varnothing \mid G_0 = \omega) = \lim_{n\to\infty} \mathbb{P}(G_n^{\mathrm{st}} \cap K = \varnothing, n > T_K(\omega))$$

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

$$\lim_{n\to\infty} \mathbb{P}(G_n \cap K = \varnothing \mid G_0 = \omega) = \lim_{n\to\infty} \mathbb{P}(G_n^{\mathrm{st}} \cap K = \varnothing) = \mathbb{P}(G_0^{\mathrm{st}} \cap K = \varnothing).$$

https://doi.org/10.1017/jpr.2018.50 Published online by Cambridge University Press

Thus, recalling that the distribution of a random closed set $X \subseteq \mathbb{R}^d$ (e.g. a locally finite point process) is uniquely characterized by the void probabilities $\mathbb{P}(X \cap K = \varnothing)$ 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 to Π. In turn, this implies uniqueness of the invariant distribution Π. □

In Theorem 4.1, under mild conditions, we can take $\Omega = \mathcal{N}$. For instance, this is easily seen to be the case if there exists $\varepsilon > 0$ so that $f(x) > 0$ whenever $\|x\| \le \varepsilon$. In the special case of $c = 0$, 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 only makes sense if the processes have equal intensities; see [6]. Under the assumptions of both Theorems 3.1 and 4.1, and using (4.5), we have

$$\begin{aligned}\gamma &= \frac{c(\beta p)^2 + 2\beta pq}{\rho_G\{1-(\beta p+q)^2\}} + \frac{b\rho_Z^2}{\rho_G^2\{1-(\beta p+q)^2\}} \\ &= \frac{1}{1+\beta p+q}\left\{\frac{c(\beta p)^2+2\beta pq}{\rho_Z} + b(1-\beta p-q)\right\},\end{aligned}$$

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 values of the parameters. 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 [6] and, consequently, $\gamma \ge b_{\min}$. The case of $\gamma = b_{\min}$ occurs exactly when $\beta p + q = 0$ (that is, when offspring are never produced or no points are retained after the thinning procedures in items (b) and (c)) 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 B. Simulation was initially carried out 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 the definition of L- and J-functions, see, e.g. [17], and for the envelopes, see [18]). 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 $\sigma = 0.1$, $f_Z \sim N_d(\frac{1}{8}\kappa^2)$, $\rho_G = 100$, $p = 1$, $\beta = 0.3$, $q = 0$, and, consequently, $\rho_Z = 70$.

- *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 $\gamma \approx -5.38 \times 10^{-3}$.
- *Poisson noise.* Let the number of points in a cluster be Poisson distributed with intensity β, implying $c = 1$. Then $\gamma \approx 9.89 \times 10^{-4}$.
- *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 $\gamma \approx 3.39$.

https://doi.org/10.1017/jpr.2018.50 Published online by Cambridge University Press

Case 2. In this case the clusters are more separated. Let $d = 2$, $f \sim N_d(\sigma^2)$ with $\sigma = 0.01$, $f_Z \sim N_d(\kappa^2/8)$, $\rho_G = 100$, $p = 1$, $\beta = 0.95$, $q = 0$, 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 $c = 5$.

- *Determinantal noise.* Let $\kappa = 1/\sqrt{\rho_Z \pi}$. Then $\gamma \approx 0.463$.
- *Poisson noise.* We have $\gamma \approx 0.463$.
- *Weighted permanental noise.* Let $\kappa = 1$. Then $\gamma \approx 0.624$.

In Figure 4 we present the results of simulations of G_0^{st} under each of the three models of the noise processes (left to right) in cases 1 and 2 (upper and lower). Based on these simulations, in Figure 5 we present 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 [18]) of a stationary Poisson process with the same intensity as used in Figure 4. The first simulated point pattern of case 1 appears slightly less clustered than the second, whilst the last appears 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 appear to be 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 a few highly separated clusters. All three models of case 2 are, as expected, significantly different from the stationary Poisson process.

Appendix A. Weighted determinantal and permanental point processes

When defining stationary weighted determinantal/permanental point processes, the main ingredients are a symmetric function $C\colon \mathbb{R}^d \to \mathbb{R}$ and a real number α. Before stating 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

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

where σ denotes a permutation of $\{1, \ldots, n\}$ 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

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

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 $x_1, \ldots, x_n \in \mathbb{R}^d$, the $n \times n$ matrix with (i, j)th entry $C(x_i - x_j)$ is denoted by $[C](x_1, \ldots, x_n)$. Thus,

$$\operatorname*{per}_{\alpha}[C](x_1, \ldots, x_n) = \sum_{\sigma} \alpha^{\#\sigma} C(x_1 - x_{\sigma_1}) \cdots C(x_n - x_{\sigma_n}).$$

https://doi.org/10.1017/jpr.2018.50 Published online by Cambridge University Press

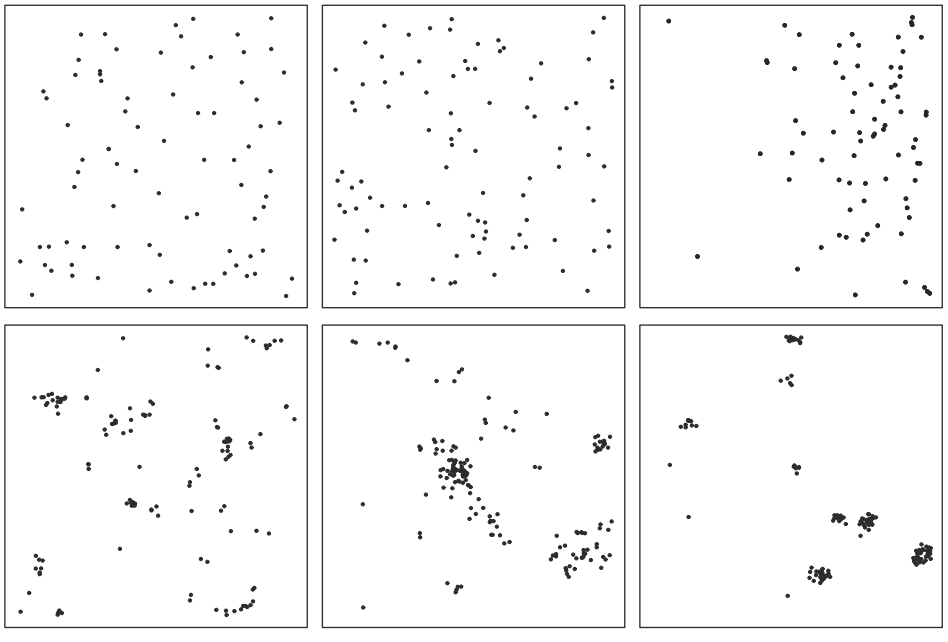


FIGURE 4: Simulations of G_0^{st} restricted to a unit square when the noise processes are either determinantal (*left*), Poisson (*middle*), or weighted permanental (*right*) point processes, with parameters as specified in the text. The rows correspond to cases 1 and 2, respectively.

Note that the weighted permanent/determinant can be negative if the mapping $\mathbb{R}^d \times \mathbb{R}^d \ni (u, v) \to C(u - v)$ 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 nth order joint intensity $\rho_X^{(n)}$ for $n = 1, 2, \ldots$ if for any bounded and pairwise disjoint Borel sets $A_1, \ldots, A_n \subset \mathbb{R}^d$,

$$\mathbb{E}\{N(A_1) \cdots N(A_n)\} = \int_{A_1} \cdots \int_{A_n} \rho_X^{(n)}(x_1, \ldots, x_n)\, \mathrm{d}x_1 \cdots \mathrm{d}x_n < \infty.$$

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 $n = 1, 2, \ldots$, the nth order joint intensity exists and is

$$\rho_X^{(n)}(x_1, \ldots, x_n) = \operatorname*{per}_{\alpha}[C](x_1, \ldots, x_n),$$

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 needed to ensure the existence of $\mathrm{PPP}_\alpha(C)$; see [10] and [21]. To exclude the trivial case where X is empty, we assume that $\alpha C(0) > 0$. Note that C must be an auto-covariance function, $\alpha > 0$ since $\rho_X = \alpha C(0)$, and

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

https://doi.org/10.1017/jpr.2018.50 Published online by Cambridge University Press

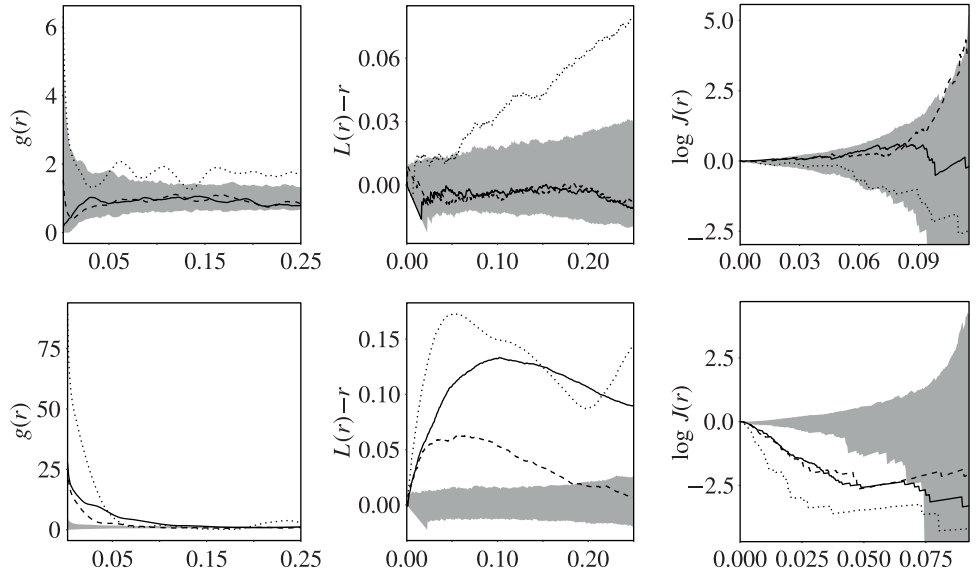


FIGURE 5: Empirical PCFs, L-functions, and J-functions (*left to right*) based on the simulations of G_0^{st} from Figure 4 when the noise processes are either determinantal (*dashed*), Poisson (*solid*), or weighted permanental (*dotted*). The rows correspond to cases 1 and 2, respectively. The shaded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

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 i.i.d. zero-mean stationary Gaussian processes $\Phi_1, \ldots, \Phi_k$ on $\mathbb{R}^d$ with auto-covariance function $C/2$, we can let X be a Poisson process with intensity function $\Lambda(x) = \Phi_1(x)^2 + \cdots + \Phi_k(x)^2$, $x \in \mathbb{R}^d$. In particular, if $\alpha = 1$ then X is the boson process introduced in [7].

If, for all $n = 1, 2, \ldots$, the nth order joint intensity exists and is

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

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 that $\alpha C(0) > 0$. Again, C needs to be an auto-covariance function, $\alpha > 0$ since $\rho_X = \alpha C(0)$, and

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

If $\alpha = 1$ then X is the fermion process introduced in [7] (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 the Fourier transform of C being bounded by 0 and 1; see [6]. When α is a positive integer, $X \sim \mathrm{DPP}_\alpha(C)$ can be identified with the superposition $G_1 \cup \cdots \cup G_\alpha$ of independent processes $G_i \sim \mathrm{DPP}_\alpha(C/\alpha)$, $i = 1, \ldots, \alpha$. In general, the process is not well defined if $0 < \alpha < 1$; see [10].

https://doi.org/10.1017/jpr.2018.50 Published online by Cambridge University Press

Appendix B. Simulating the limiting process

In this appendix we present an approximate simulation procedure for simulating a special case of G_0^{st} on a bounded region $R \subset \mathbb{R}^d$. It is available in R through the package icpp; available at https://github.com/adchSTATS/icpp. The implementation utilizes existing functions from the packages spatstat and RandomFields to simulate the noise process.

For simplicity and specificity, we make the following assumptions. Let the situation be as in Theorem 4.1, but with $q = 0$ and let $f \sim N_d(\sigma^2)$ with $\sigma > 0$. Also, without loss of generality, assume no thinning (that is, $p = 1$). Let $R_{\oplus r} = \{\xi \in \mathbb{R}^d : b(\xi, r) \cap R \neq \varnothing\}$, where $b(\xi, r)$ is a closed ball with its centre at ξ and radius $r \geq 0$. Denote by n the number of iterations in our approximate simulation algorithm, that is, $-n$ is the starting time when ignoring what occurred 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 nonnegative 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 $G_0^{\mathrm{st}} \cap R$ that are generated by an ith generation ancestor x when $i < -n$ or both $-n \leq i < 0$ and $x \notin R_{\oplus 4\sqrt{-i}\sigma}$.

Algorithm 1. The algorithm is 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 $O := Z_i' \bigcup (\bigcup_{x \in O} O_x \cap R_{\oplus 4\sqrt{-i}\sigma})$
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

This work was 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, Henrike Häbel, Eva B. Vedel Jensen, and Morten Nielsen for helpful comments.

https://doi.org/10.1017/jpr.2018.50 Published online by Cambridge University Press

References

[1] Andersen, I. T. *et al.* (2018). Double Cox cluster processes - with applications to photoactivated localization microscopy. *Spatial Statist.* **27,** 58–73.

[2] Barndorff-Nielsen, O., Kent, J. and Sørensen, M. (1982). Normal variance-mean mixtures and z distributions. *Internat. Statist. Rev.* **50,** 145–159.

[3] Daley, D. J. and Vere-Jones, D. (2003). *An Introduction to the Theory of Point Processes,* Vol. I, *Elementary Theory and Methods*, 2nd edn. Springer, New York.

[4] Felsenstein, J. (1975). A pain in the torus: some difficulties with models of isolation by distance. *Amer. Naturalist* **109,** 359–368.

[5] Kingman, J. F. C. (1977). Remarks on the spatial distribution of a reproducing population. *J. Appl. Prob.* **14,** 577–583.

[6] Lavancier, F., Møller, J. and Rubak, E. (2015). Determinantal point process models and statistical inference. *J. R. Statist. Soc. B* **77,** 853–877.

[7] Macchi, O. (1975). The coincidence approach to stochastic point processes. *Adv. Appl. Prob.* **7,** 83–122.

[8] Matérn, B. (1960). *Spatial Variation*. Meddelanden från Statens Skogforskningsinstitut, Stockholm.

[9] Matérn, B. (1986). *Spatial Variation* (Lecture Notes Statist. **36**), 2nd edn. Springer, Berlin.

[10] McCullagh, P. and Møller, J. (2006). The permanental process. *Adv. Appl. Prob.* **38,** 873–888.

[11] Møller, J. (1989). Random tessellations in $\mathbf{R}^d$. *Adv. Appl. Prob.* **21,** 37–73.

[12] Møller, J. (1994). *Lectures on Random Voronoi Tessellations* (Lecture Notes Statist. **87**). Springer, New York.

[13] Møller, J. (2003). Shot noise Cox processes. *Adv. Appl. Prob.* **35,** 614–640.

[14] Møller, J. and Christoffersen, A. D. (2018). Pair correlation functions and limiting distributions of iterated cluster point processes. Preprint. Available at https://arxiv.org/abs/1711.08984.

[15] Møller, J. and Torrisi, G. L. (2005). Generalised shot noise Cox processes. *Adv. Appl. Prob.* **37,** 48–74.

[16] Møller, J. and Torrisi, G. L. (2007). The pair correlation function of spatial Hawkes processes. *Statist. Prob. Lett.* **77,** 995–1003.

[17] Møller, J. and Waagepetersen, R. P. (2004). *Statistical Inference and Simulation for Spatial Point Processes*. Chapman & Hall/CRC, Boca Raton, FL.

[18] Myllymäki, M. *et al.* (2017). Global envelope tests for spatial processes. *J. R. Statist. Soc. B* **79,** 381–404.

[19] Neyman, J. and Scott, E. L. (1958). Statistical approach to problems of cosmology. *J. R. Statist. Soc. B* **20,** 1–43.

[20] Shimatani, I. K. (2010). Spatially explicit neutral models for population genetics and community ecology: extensions of the Neyman-Scott clustering process. *Theoret. Pop. Biol.* **77,** 32–41.

[21] Shirai, T. and Takahashi, Y. (2003). Random point fields associated with certain Fredholm determinants. I. Fermion, Poisson and boson point processes. *J. Functional Anal.* **205,** 414–463.

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

[23] Van Lieshout, M. N. M. and Baddeley, A. J. (2002). Extrapolating and interpolating spatial patterns. In *Spatial Cluster Modelling*, Chapman & Hall/CRC, Boca Raton, FL, pp. 61–86.

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

https://doi.org/10.1017/jpr.2018.50 Published online by Cambridge University Press