Adv. Appl. Prob. (*SGSA*) **28**, 346–355 (1996)
Printed in N. Ireland

MARKOV PROPERTIES OF CLUSTER PROCESSES

A. J. BADDELEY,* *University of Western Australia and University of Leiden*
M. N. M. VAN LIESHOUT,** *University of Warwick*
J. MØLLER,*** *University of Aarhus*

Abstract

We show that a Poisson cluster point process is a nearest-neighbour Markov point process [2] if the clusters have uniformly bounded diameter. It is typically not a finite-range Markov point process in the sense of Ripley and Kelly [12]. Furthermore, when the parent Poisson process is replaced by a Markov or nearest-neighbour Markov point process, the resulting cluster process is also nearest-neighbour Markov, provided all clusters are non-empty. In particular, the nearest-neighbour Markov property is preserved when points of the process are independently randomly translated, but not when they are randomly thinned.

MARKOV POINT PROCESS: NEAREST-NEIGHBOUR MARKOV PROCESS: CLUSTER PROCESS: CONNECTED COMPONENT RELATION

AMS 1991 SUBJECT CLASSIFICATION: PRIMARY 60D05; 60G55; 62M30

1. Introduction

Markov or Gibbs point processes [2, 9, 12, 13] form a large, flexible and understandable class of point process models with many practical advantages (see e.g. [4, 10, 11] for surveys). In this paper we consider the relationship of these models to the basic point process operation of clustering. We ask whether cluster processes are Markov, and whether the Markov property is preserved under clustering.

In a Poisson cluster process, intuitively the only 'spatial dependence' present is that between offspring of the same parent. If the offspring of a given parent all lie within distance R of the parent, then two offspring of the same parent lie at most $2R$ apart, and it is plausible to conjecture that the process is Markov with finite interaction range $2R$ in the sense of Ripley and Kelly [12].

However, this turns out to be false in general, because certain spatial configurations of the offspring points imply information about the unobserved parent points, and this information can 'propagate' over arbitrarily large distances.

In this paper we show that cluster processes have the *nearest-neighbour* Markov property in the sense of Baddeley and Møller [2] with respect to the connected

Received 19 May 1994; revision received 9 June 1995.

* Postal address: Department of Mathematics, University of Western Australia, Nedlands WA 6009, Australia.

** Postal address: Department of Statistics, University of Warwick, Coventry CV4 7AL, U.K.

*** Postal address: Department of Mathematics and Computer Science, University of Aalborg, F. Bajers Vej 7E, DK-9220 Aalborg Ø, Denmark.

component relation [2, p. 106]. We prove that (a) any Poisson cluster process with uniformly bounded clusters is a nearest-neighbour Markov point process; and (b) if a Markov or nearest-neighbour Markov point process is used as the parent process for a cluster process, and the clusters are uniformly bounded *and a.s. non-empty*, then the cluster process is again nearest-neighbour Markov. In particular, the nearest-neighbour Markov property is preserved under random displacement of points, but not under random thinning.

These results support the claim [7, 8] that nearest-neighbour Markov processes (as opposed to Ripley–Kelly Markov processes) provide a rich class of models for clustering, and further suggest that they may include good models for multiple-generation cluster processes, cf. [6]. Result (a) may also explain why statistical theory for Poisson cluster processes so closely parallels that for Markov point processes [1].

The next section recalls standard definitions; the main results are stated in Section 3 and the proofs follow in Section 4.

2. Setup

2.1. *Point processes.* We consider finite point processes X on a metric space S (typically $\mathbb{R}^d$ or a compact subset). Each realization of such a process 'is' a finite set $\boldsymbol{x} = \{x_1, \ldots, x_n\}$ of points $x_i \in S$ with $n \geqq 0$. Strictly speaking, the points may be multiply occupied, and n is the total multiplicity, but this will have probability zero in the applications considered. Realizations will also be called 'configurations' and the class of all configurations will be denoted by $\mathscr{C}$. This is the exponential space of S, see [3] or [2] for details.

Let ν be some given Borel measure on S (typically Lebesgue measure); we will consider processes whose distributions are absolutely continuous with respect to the measure μ on $\mathscr{C}$ defined by

$$\mu(F) = \sum_{k=0}^{\infty} \frac{1}{k!} \int \cdots \int \mathbf{1}[\{y_1, \cdots, y_k\} \in F]\, d\nu(y_1) \cdots d\nu(y_k). \tag{2.1}$$

If ν is totally finite (e.g. if S is compact and ν is Lebesgue measure), then μ is $e^{\nu(S)}$ times the distribution of the Poisson process on S with intensity measure ν.

Let $f: \mathscr{C} \to [0, \infty)$ be the density of a point process X with respect to μ. We say f is *hereditary* if $f(\boldsymbol{x}) > 0$ implies $f(\boldsymbol{z}) > 0$ for all $\boldsymbol{z} \subseteq \boldsymbol{x}$ and *hereditary excluding* $\varnothing$ if this holds except when $\boldsymbol{z} = \varnothing$.

2.2. *Markov point processes.* This subsection collects necessary definitions from [2], [10], [12].

Define $u, v \in S$ to be *r-close*, written $u \sim v$, if $d(u, v) < r$ where d is the metric of S. This defines a relation $\sim$ on S which is clearly symmetric and reflexive. (The results of this paper extend to the case where $\sim$ is any symmetric reflexive relation

on S which is measurable in the product space. Theorem 2 in Section 3 requires two such relations.)

Definition 1. (*Ripley–Kelly*) A point process X is Markov with respect to the static relation $\sim$ if its density f satisfies

(i) f is hereditary;

(ii) for any $\boldsymbol{x} \in \mathcal{C}$ such that $f(\boldsymbol{x}) > 0$ and $u \in S$, the ratio

$$\frac{f(\boldsymbol{x} \cup \{u\})}{f(\boldsymbol{x})} \tag{2.2}$$

depends only on u and on $\{x_i \in \boldsymbol{x} : u \sim x_i\}$.

Now define for each $\boldsymbol{x} \in \mathcal{C}$ the 'connected component relation' [2, Appendix III] between points of $\boldsymbol{x}$ by $x_i \underset{\boldsymbol{x}}{\sim} x_j$ iff $x_i \sim z_1 \sim \cdots \sim z_n \sim x_j$ for some $z_1, \cdots, z_n \in \boldsymbol{x}$. In other words, two points of $\boldsymbol{x}$ are related under $\underset{\boldsymbol{x}}{\sim}$ if they are in the same connected component of the finite graph whose edges connect every pair of r-close points in $\boldsymbol{x}$.

Definition 2. (*Baddeley–Møller*) A point process X is nearest-neighbour Markov with respect to the dynamic relation $\underset{\boldsymbol{x}}{\sim}$ if its density f is hereditary and the ratio (2.2) depends only on u, on

$$\text{Nbd}\,(u \mid \boldsymbol{x} \cup \{u\}) = \{x_i \in \boldsymbol{x} : u \underset{\boldsymbol{x} \cup \{u\}}{\sim} x_i\},$$

and on the relations $\underset{\boldsymbol{x}}{\sim}$, $\underset{\boldsymbol{x} \cup \{u\}}{\sim}$ restricted to Nbd $(u \mid \boldsymbol{x} \cup \{u\})$.

Clearly if X is Markov with respect to $\sim$, it is also nearest-neighbour Markov with respect to both $\sim$ and $\underset{\boldsymbol{x}}{\sim}$.

Other dynamic relations $\underset{\boldsymbol{x}}{\sim}$, for instance based on the Voronoi tessellation of pattern $\boldsymbol{x}$, can be used to define a nearest-neighbour Markov density as well. For details, see [2]. A spatial Markov property in terms of splitting sets was proved in [5].

Analogues of the Hammersley–Clifford theorem proved in [12] and in [2] give explicit expressions for the density f when X is Markov and nearest-neighbour Markov, respectively. Define a configuration $\boldsymbol{z}$ to be a *clique* with respect to $\sim$ (or an *r-clique*) if all pairs of points in $\boldsymbol{z}$ are r-close, $z_i \sim z_j$ for all $z_i, z_j \in \boldsymbol{z}$. Then [12] X is Markov iff

$$f(\boldsymbol{x}) = \prod_{\boldsymbol{z} \subseteq \boldsymbol{x}} \varphi(\boldsymbol{z}) \tag{2.3}$$

where $\varphi(\boldsymbol{z}) \geqq 0$ with $\varphi(\boldsymbol{z}) \neq 1$ only if $\boldsymbol{z}$ is a clique.

In the nearest-neighbour case, a subconfiguration $\boldsymbol{z} \subseteq \boldsymbol{x}$ is termed a *clique with*

respect to $\underset{x}{\sim}$ if all pairs of points in z are $\underset{x}{\sim}$-neighbours, $z_i \underset{x}{\sim} z_j$ for all $z_i, z_j \in z$. The maximal r-cliques are also called *connected components.*

An analogue of (2.3) for dynamic relations $\underset{x}{\sim}$ is given in [2, Theorem 4.13]. For the 'connected component relation' this specialises to the following result.

Lemma 1. *A point process X is nearest-neighbour Markov with respect to the connected component relation* $\underset{x}{\sim}$ *iff*

$$f(x) = \prod_{\text{cliques } z \subseteq x} \varphi(z) \tag{2.4}$$

where $\varphi(\cdot) \geqq 0$ *is such that whenever* z *is a* $\underset{z}{\sim}$*-clique with* $\varphi(z) > 0$ *then* $\varphi(w) > 0$ *for all* $w \subseteq z$.

Equivalently, X is nearest-neighbour Markov w.r.t. $\underset{x}{\sim}$ iff

$$f(x) = f(\varnothing) \prod_{k=1}^{K} \Phi(x_{D_k}) \tag{2.5}$$

where $x_D, \cdots x_{D_K}$ are the connected components of x and $\Phi(\cdot) \geqq 0$ is such that if x is a $\underset{x}{\sim}$-clique and $z \subseteq x$ is a $\underset{z}{\sim}$-clique then $\Phi(x) > 0$ implies $\Phi(z) > 0$.

2.3. *Cluster processes.* For $\xi \in S$ let Q_ξ be the distribution of a finite point process Z_ξ. This will be the process of offspring of a parent at point ξ.

We shall assume that

(A) Q_ξ is absolutely continuous with respect to μ, with density q_ξ;

(B) $(\xi, z) \mapsto q_\xi(z)$ is Borel measurable $S \times \mathscr{C} \to \mathbb{R}_+$;

(C) (uniform boundedness) $Z_\xi \subseteq b(\xi, R)$ a.s. for some fixed $R > 0$;

(D) q_ξ is hereditary excluding $\varnothing$ (see Section 2.1).

Here $b(\xi, R)$ denotes the closed ball in the metric d with centre ξ and radius R.

Given a finite point process x, construct a marked point process $\{(x_i, Z_i)\}$ where x_i are the points of x and the marks Z_i are conditionally independent, finite point processes such that Z_i has distribution Q_{x_i}. Form the superposition $y = \bigcup_i Z_i$; then y is the *cluster process with parents* x *and cluster distributions* Q_ξ.

Lemma 2. *The distribution of the cluster process y described above is absolutely continuous with respect to μ, with density*

$$f(y) = \mathbb{E}\Big(\sum_{C_1, \cdots, C_n} \prod_{i=1}^{n} q_{x_i}(y_{C_i}) \Big) \tag{2.6}$$

where $\mathbb{E}$ denotes expectation with respect to the distribution of x, and the sum is over all ordered partitions of $y = \{y_1, \cdots, y_m\}$ into n subconfigurations $y_{C_1}, \cdots, y_{C_n}$ (allowing empty sets) where $n = n(x)$ is the random number of points in x.

Strictly speaking, by an 'ordered partition' we mean a mapping ε from the points of $\boldsymbol{y}$ into those of $\boldsymbol{x}$, and $\boldsymbol{y}_{C_i}$ is a convenient notation for $\varepsilon^{-1}\{x_i\}$. An 'unordered partition' is simply a partition of $\boldsymbol{y}$ into disjoint subconfigurations.

3. Statement of results

Theorem 1. *Let* $\boldsymbol{x}$ *be a Poisson point process on* S *with finite intensity measure* λ, *and* $\boldsymbol{y}$ *a cluster process with parent process* $\boldsymbol{x}$ *and clusters satisfying the assumptions* (*A*)–(*D*) *of Section* 2.3. *Then* $\boldsymbol{y}$ *is a nearest-neighbour Markov point process with respect to the connected component relation at distance* $2R$.

For example, consider a Neyman–Scott process in a bounded set $S = B \subset \mathbb{R}^d$. Let λ be Lebesgue measure restricted to B. Thus $\boldsymbol{x}$ is a unit rate Poisson process on B. Assume that a parent at ξ has a Poisson (ω) number of offspring, positioned i.i.d. with probability density $h(\cdot - \xi)$ where h is supported on the ball $b(0, R)$ of radius R. Let ν be Lebesgue measure restricted to the dilated set $B_{\oplus R} = \{a \in \mathbb{R}^d : d(a, B) < R\}$. Then the density of $\boldsymbol{y}$ is (for $\boldsymbol{y} \neq \varnothing$)

$$f(\boldsymbol{y}) = \sum_{n=1}^{\infty} \frac{\mathrm{e}^{-\nu(B)}\omega^m \mathrm{e}^{-\omega n}}{n!} \int_B \cdots \int_B \prod_{j=1}^{m} \left(\sum_{i=1}^{n} h(y_j - x_i) \right) dx_1 \cdots dx_n$$

where $m = n(\boldsymbol{y})$. This is easily (or via the proof of Theorem 1) factorised as

$$f(\boldsymbol{y}) = \omega^m \mathrm{e}^{-m\omega} \mathrm{e}^{-\beta} \sum_{C_1, \cdots, C_k}^{\neq \varnothing} \mathrm{e}^{\omega(m-k)} J(\boldsymbol{y}_{C_1}) \cdots J(\boldsymbol{y}_{C_k}), \qquad \text{if } \boldsymbol{y} \neq \varnothing \tag{3.7}$$

where $\beta = (1 - \mathrm{e}^{-\omega})\nu(B)$, the superscript $\neq \varnothing$ indicates that the sum is over all unordered partitions of $\boldsymbol{y}$ into disjoint *non-empty* subsets, and

$$J(\boldsymbol{y}_C) = \int_B \prod_{y_j \in \boldsymbol{y}_C} h(y_j - \xi)\, d\xi. \tag{3.8}$$

Since $J(\boldsymbol{y}_C) = 0$ unless $\boldsymbol{y}_C$ is a $\underset{\boldsymbol{y}}{\sim}$-clique, the only non-zero terms in (3.7) are those for partitions which are refinements of the partition of $\boldsymbol{x}$ into connected components. Thus (3.7) factorises into terms associated with each connected component, and according to (2.5) we have a nearest-neighbour Markov process, provided the positivity condition stated below (2.5) is satisfied.

A special case of Neyman–Scott is the Matérn cluster process, in which $h \equiv$ constant on $b(0, R)$: then we have $J(\boldsymbol{y}_C) = \lambda(\bigcap_{y_j \in \boldsymbol{y}_C} b(y_j, R))$ i.e. $J(\boldsymbol{y}_C)$ is the volume occupied within B by the intersection of the balls of radius R centred at the points of $\boldsymbol{y}_C$. In this case the positivity condition is clearly satisfied, so that the Matérn cluster process is nearest-neighbour Markov.

In Theorem 1, nearest-neighbour Markov cannot be replaced by Ripley–Kelly Markov, as the following example shows.

Counterexample 1. A Neyman–Scott process with uniformly bounded clusters is *not* (in general) a Markov point process at any fixed range $s < \infty$.

For example, for a Matérn cluster process, first consider ranges $s < 2R$. If f were a Markov function at range $0 < s \leqq R$, then *a fortiori* f would be a Markov function at range $s' > R$ so assume that $R < s < 2R$ and consider a configuration of three points y_1, y_2, y_3 such that $\|y_1 - y_2\| < s$, $\|y_2 - y_3\| < s$, but $\|y_1 - y_3\| > s$. If f were a Markov function at range s then $f(\{y_1, y_2, y_3\})f(\{y_2\}) = f(\{y_1, y_2\})f(\{y_2, y_3\})$.

Substituting (3.7) gives

$$[1 + e^{\omega}J(y_1, y_2) + e^{\omega}J(y_2, y_3)] = [1 + e^{\omega}J(y_1, y_2)][1 + e^{\omega}J(y_2, y_3)].$$

This is clearly a contradiction, since the J terms are non-zero. Hence f is not a Markov density in the Ripley–Kelly sense at distance $s < 2R$. For $s \geqq 2R$ one can use similar arguments involving chains of more than three points.

Next, we consider cluster processes generated from a parent process which is Markov or nearest-neighbour Markov. In general the cluster process is not Markov.

Counterexample 2. Let $\boldsymbol{x}$ be a Ripley–Kelly Markov point process (finite range r) and $\boldsymbol{y}$ the result of thinning the points independently with retention probability q, $0 < q < 1$. Then, in general, $\boldsymbol{y}$ is *not* a nearest-neighbour Markov point process (and *a fortiori* it is not a Ripley–Kelly Markov process) for any $R < \infty$.

This can be checked from (2.6), since random thinning is the special case of clustering in which $Z_\xi = \{\xi\}$ with probability q, and $Z_\xi = \varnothing$ otherwise. For any given pair of points $y_i, y_j \in \boldsymbol{y}$ there are (potentially) non-zero summands in (2.6) of the form $q_{x_1}(y_i)q_{x_2}(\varnothing) \cdots q_{x_{N-1}}(\varnothing)q_{x_N}(y_j)$ involving both y_i, y_j. Hence $\boldsymbol{y}$ is not nearest-neighbour Markov according to (2.4).

Clearly this problem may arise whenever clusters are permitted to be empty, i.e. when a parent point may have no offspring. When this is excluded, we do obtain a Markov property.

Theorem 2. *Let* $\boldsymbol{x}$ *be a Markov or nearest-neighbour Markov point process at range* r *and* $\boldsymbol{y}$ *the associated cluster process satisfying* (*A*)–(*D*) *of Section* 2.3 *and moreover*

(E) *the clusters are* non-empty a.s.

Then $\boldsymbol{y}$ *is a nearest-neighbour Markov point process for the connected component relation at range* $2R + r$.

Corollary 1. *Let* $\boldsymbol{x}$ *be as above, and let* $\boldsymbol{y}$ *be the process obtained from* $\boldsymbol{x}$ *by independently translating each point:* $y_i = x_i + v_i$, *where the vectors* v_i *are i.i.d., have a probability density, and satisfy* $\|v_i\| < R$ *a.s. Then* $\boldsymbol{y}$ *is a nearest-neighbour Markov point process for the connected component relation at range* $2R + r$.

4. Proofs

Proof of Lemma 1. Suppose that (2.4) holds. Let $\boldsymbol{x} \in \mathscr{C}$, $\xi \in S$ and let $\boldsymbol{x}_{D_1}, \cdots, \boldsymbol{x}_{D_K}$ and $\boldsymbol{w} \cup \{\xi\}$ denote the connected components of $\boldsymbol{x} \cup \{\xi\}$. Then, if $\boldsymbol{x}_{D_{K+1}}, \cdots, \boldsymbol{x}_{D_L}$ are the connected components of $\boldsymbol{w}$, we have that $\boldsymbol{x}_{D_1}, \cdots, \boldsymbol{x}_{D_L}$ are the connected components of $\boldsymbol{x}$, and

$$f(\boldsymbol{x} \cup \{\xi\}) = \varphi(\varnothing)\Big[\prod_{i=1}^{K} \prod_{\varnothing \neq \boldsymbol{z} \subseteq \boldsymbol{x}_{D_i}} \varphi(\boldsymbol{z})\Big] \prod_{\varnothing \neq \boldsymbol{z} \subseteq \boldsymbol{w} \cup \{\xi\}} \varphi(\boldsymbol{z})$$

while

$$f(\boldsymbol{x}) = \varphi(\varnothing)\Big[\prod_{i=1}^{K} \prod_{\varnothing \neq \boldsymbol{z} \subseteq \boldsymbol{x}_{D_i}} \varphi(\boldsymbol{z})\Big] \prod_{j=K+1}^{L} \prod_{\varnothing \neq \boldsymbol{z} \subseteq \boldsymbol{x}_{D_j}} \varphi(\boldsymbol{z}).$$

Hence $f(\boldsymbol{x} \cup \{\xi\}) > 0$ implies $f(\boldsymbol{x}) > 0$ (as $\boldsymbol{z} \subseteq \boldsymbol{x}_{D_j}$ for $j > K$ implies that $\boldsymbol{z} \subseteq \boldsymbol{w}$) and $f(\boldsymbol{x} \cup \{\xi\})/f(\boldsymbol{x})$ satisfies the conditions of Definition 2. Thus X is nearest-neighbour Markov.

Conversely, suppose X is nearest-neighbour Markov. By the analogue of the Hammersley–Clifford theorem [2, Theorem 4.13],

$$f(\boldsymbol{x}) = \prod_{\boldsymbol{y} \subseteq \boldsymbol{x}} \varphi(\boldsymbol{y})^{\chi(\boldsymbol{y} \mid \boldsymbol{x})} \qquad \text{(taking } 0^0 = 0\text{)} \tag{4.9}$$

where $\chi(\boldsymbol{y} \mid \boldsymbol{x}) = 1$ if $\boldsymbol{y}$ is a $\underset{\boldsymbol{x}}{\sim}$-clique and 0 otherwise; and $\varphi: \mathscr{C} \to \mathbb{R}_+$ satisfies

(I1) $\varphi(\boldsymbol{x}) > 0$ implies $\varphi(\boldsymbol{y}) > 0$ for all $\boldsymbol{y} \subseteq \boldsymbol{x}$,

(I2) $\varphi(\boldsymbol{x}) > 0$ and $\varphi(\{\xi\} \cup \text{Nbd}\,(\xi \mid \boldsymbol{x} \cup \{\xi\})) > 0$ imply $\varphi(\boldsymbol{x} \cup \{\xi\}) > 0$.

Note that, in the case of the connected component relation, $\xi \underset{\boldsymbol{y}}{\sim} \eta$ implies $\xi \underset{\boldsymbol{x}}{\sim} \eta$ for $\boldsymbol{x} \supseteq \boldsymbol{y}$, so that $\chi(\boldsymbol{y} \mid \boldsymbol{y}) = 1$ implies $\chi(\boldsymbol{y} \mid \boldsymbol{x})$ for any $\boldsymbol{x} \supseteq \boldsymbol{y}$.

To prove that (4.9) reduces to (2.4) we need to show that, if $\varphi(\boldsymbol{y}) > 0$ for all $\boldsymbol{y} \subseteq \boldsymbol{x}$ with $\chi(\boldsymbol{y} \mid \boldsymbol{x}) = 1$, then $\varphi(\boldsymbol{y}) > 0$ for all $\boldsymbol{y} \subseteq \boldsymbol{x}$.

To prove this, suppose $\boldsymbol{v}, \boldsymbol{w} \subseteq \boldsymbol{x}$ are disjoint connected components of $\boldsymbol{x}$ (i.e. with respect to $\underset{\boldsymbol{x}}{\sim}$). If $\xi \in \boldsymbol{v}$ then Nbd $(\xi \mid \boldsymbol{w} \cup \{\xi\}) = \{\xi\}$, and by assumption $\varphi(\{\xi\}) > 0$, so (I2) gives $\varphi(\boldsymbol{w} \cup \{\xi\}) > 0$. Similarly, if $\{\xi, \eta\} \subseteq \boldsymbol{v}$ with $\eta \sim \xi$ then Nbd $(\eta \mid \boldsymbol{w} \cup \{\xi, \eta\}) = \{\xi, \eta\}$, and by assumption $\varphi(\{\xi, \eta\}) > 0$, so (I2) gives $\varphi(\boldsymbol{w} \cup \{\xi, \eta\}) > 0$. Continuing in this way we obtain that $\varphi(\boldsymbol{y}) > 0$ for all $\boldsymbol{y} \subseteq \boldsymbol{x}$.

Hence if X is nearest-neighbour Markov then its density is of the form (2.4) where φ satisfies (I1) and hence the condition stated in the lemma.

Proof of Lemma 2. The clusters Z_i being conditionally independent given $\boldsymbol{x}$, we have for any measurable event F

$$\begin{aligned}\mathbb{P}\{\boldsymbol{y} \in F \mid \boldsymbol{x}\} &= \int \mathbf{1}\Big[\bigcup_i Z_i \in F\Big] q_{x_1}(Z_1) \cdots q_{x_n}(Z_n)\, d\mu(Z_1) \cdots d\mu(Z_n) \\ &= \sum_{C_1, \cdots, C_n} \int \mathbf{1}[\boldsymbol{y} \in F] q_{x_1}(\boldsymbol{y}_{C_1}) \cdots q_{x_n}(\boldsymbol{y}_{C_n})\, d\mu(\boldsymbol{y}).\end{aligned}$$

The last line was obtained using (2.1) by rewriting each integral over Z_i as a sum of multiple integrals with respect to ν and regrouping. The result follows by taking expectations with respect to $\boldsymbol{x}$.

Proof of Theorem 1. By (2.6), the density of $\boldsymbol{y}$ with respect to μ is (for $\boldsymbol{y} \neq \varnothing$)

$$f(\boldsymbol{y}) = \sum_{n=1}^{\infty} \frac{e^{-\lambda(S)}}{n!} \int_S \cdots \int_S \sum_{C_1,\cdots,C_n} \prod_{i=1}^{n} q_{x_i}(\boldsymbol{y}_{C_i})\, d\lambda(x_1) \cdots d\lambda(x_n) \tag{4.10}$$

$$= \sum_{n=1}^{\infty} \frac{e^{-\lambda(S)}}{n!} \sum_{C_1,\cdots,C_n} \prod_{i=1}^{n} \int_S q_\xi(\boldsymbol{y}_{C_i})\, d\lambda(\xi); \tag{4.11}$$

here the inner sum is over all *ordered* partitions of $\boldsymbol{y}$ into n disjoint, possibly empty, sets. Since the parent process is Poisson, the number of non-empty clusters is Poisson distributed with mean $\beta = \int_S (1 - q_\xi(\varnothing))\, d\lambda(\xi)$, so that for $\boldsymbol{y} = \varnothing$ we have $f(\varnothing) = e^{-\beta}$.

Now $q_\xi(\boldsymbol{z}) = 0$ whenever $\boldsymbol{z} \not\subseteq b(\xi, R)$; hence if $q_\xi(\boldsymbol{z}) \neq 0$ then all pairs of points in $\boldsymbol{z}$ are $2R$-close, i.e. $\boldsymbol{z}$ is a clique with respect to the finite range relation with distance $2R$. Hence the integral in (4.11) is non-zero only when the partition consists of $2R$-cliques.

For $\boldsymbol{y} \neq \varnothing$, let $\boldsymbol{y}_{D_1}, \cdots, \boldsymbol{y}_{D_K}$ be the connected components of $\boldsymbol{y}$ for the relation $\underset{\boldsymbol{y}}{\sim}$ with range $2R$. Then the integral in (4.11) is non-zero only when the partition is a refinement of $D_1, \cdots, D_K$. Let $C_1, \cdots, C_k$ be an (unordered) partition refining $D_1, \cdots, D_K$ and consisting of *non-empty* sets. This contributes a term $\alpha \prod_{i=1}^{k} \int_S q_\xi(\boldsymbol{y}_{C_i})\, d\lambda(\xi)$ to the density. Since $\int_S q_\xi(\varnothing)\, d\lambda(\xi) = \lambda(S) - \beta$, the coefficient α is

$$\sum_{n=k}^{\infty} \frac{e^{-\lambda(S)}}{n!} (\lambda(S) - \beta)^{n-k} n(n-1) \cdots (n-k+1) = e^{-\beta}.$$

The class of all partitions that are refinements of $D_1, \cdots, D_K$ is the cartesian product of the set of partitions of each D_i. Hence, for $\boldsymbol{y} \neq \varnothing$,

$$f(\boldsymbol{y}) = e^{-\beta} \prod_{i=1}^{K} \Phi(\boldsymbol{y}_{D_i}) \tag{4.12}$$

where

$$\Phi(\boldsymbol{z}) = \sum_{k \geq 1} \sum_{\boldsymbol{z}_{C_1},\cdots,\boldsymbol{z}_{C_k}} \prod_{j=1}^{k} \int_S q_\xi(\boldsymbol{z}_{C_j})\, d\lambda(\xi) \tag{4.13}$$

where $\boldsymbol{z}_{C_1}, \cdots, \boldsymbol{z}_{C_k}$ range over all (unordered) partitions of $\boldsymbol{z}$ into non-empty subconfigurations.

Since the offspring densities q_ξ are hereditary excluding $\varnothing$, clearly Φ is hereditary excluding $\varnothing$, and hence f is hereditary. According to (2.4) the density (4.12) is nearest-neighbour Markov with respect to the connected component relation at range $2R$.

Proof of Theorem 2. The density $p(\boldsymbol{x})$ of $\boldsymbol{x}$ can be factorised as in (2.4). By (2.6),

the density of y with respect to μ is

$$f(y) = \int_{\mathscr{C}} \sum_{C_1,\cdots,C_n(x)} \prod_{i=1}^{n(x)} q_{x_i}(y_{C_i}) p(x)\, d\mu(x) \tag{4.14}$$

where the sum ranges over all ordered partitions of y into disjoint, possibly empty, subsets. Since $q_\xi(\varnothing) = 0$, the integrand of (4.14) can be rephrased as

$$\sum_{\varepsilon} \prod_{i=1}^{n} q_{x_i}(y_{\varepsilon^{-1}(i)}) p(x) \tag{4.15}$$

where ε ranges over all surjective mappings of the points of y onto the points of x, identified with mappings from $\{1, \cdots, m\}$ onto $\{1, \cdots, n\}$.

We can restrict attention to those ε such that

$$d(y_i, x_{\varepsilon(i)}) \leqq R \qquad \text{for all } i \tag{4.16}$$

since all other terms are zero. For such ε, if $z \subseteq x$ is an r-clique and $\varepsilon^{-1}(z) = w \subseteq y$, then w *must be a* $(2R + r)$*-clique.* To see this, take $y_i, y_j \in w$ and apply the triangle inequality: $d(y_i, y_j) \leqq d(y_i, x_{\varepsilon(i)}) + d(x_{\varepsilon(i)}, x_{\varepsilon(j)}) + d(x_{\varepsilon(j)}, y_j) \leqq R + r + R$. By a similar argument, if $z \subseteq x$ is a clique with respect to the relation $\underset{x}{\sim}$ at distance r, then w is a clique with respect to the relation $\underset{y}{\sim}$ at distance $2R + r$.

Let $y_{D_1}, \cdots, y_{D_K}$ be the connected components of y with respect to the relation at distance $2R + r$. Then we can rewrite (4.15) as

$$\begin{aligned} \prod_{\text{cliques } z \subseteq x} \varphi(z) \sum_{\varepsilon} \prod_{i=1}^{n} q_{x_i}(y_{\varepsilon^{-1}(i)}) &= \sum_{\varepsilon} \left[\prod_{i=1}^{n} q_{x_i}(y_{\varepsilon^{-1}(i)}) \prod_{k=1}^{K} \prod_{\text{cliques } z \subseteq x : \varepsilon^{-1}(z) \subseteq D_k} \varphi(z) \right] \\ &= \sum_{\varepsilon} \prod_{k=1}^{K} \left[\prod_{i : \varepsilon^{-1}(i) \subseteq D_k} q_{x_i}(y_{\varepsilon^{-1}(i)}) \prod_{\text{cliques } z \subseteq x : \varepsilon^{-1}(z) \subseteq D_k} \varphi(z) \right]. \end{aligned} \tag{4.17}$$

Any ε of the type described above can be represented as an ordered set of K surjective mappings $\varepsilon_k : D_k \to D_k' = \{i \mid d(x_i, y_j) \leqq R \text{ for some } j \in D_i\}$ automatically satisfying the norm condition (4.16). Note that $x_{D_k'}$, $k = 1, \cdots, K$ form a disjoint partition of x. Thus (4.17) is

$$\prod_{k=1}^{K} \sum_{\varepsilon_k} \left[\prod_{i \in D_k'} q_{x_i}(y_{\varepsilon_k^{-1}(i)}) \prod_{\text{cliques } z \subseteq x_{D_k'}} \varphi(z) \right].$$

Integrating over x and exploiting the form (2.1) of μ yields

$$f(y) = \prod_{k=1}^{K} \int_{\mathscr{C}} \sum_{\varepsilon_k} \prod_{i=1}^{n(v)} q_{v_i}(y_{\varepsilon_k^{-1}(i)}) \prod_{\text{cliques } z \subseteq v} \varphi(z)\, d\mu(v).$$

Thus, f factorises as required by (2.4). The hereditary property follows as in the previous proof.

Acknowledgements

This research was carried out at CWI, Amsterdam (Baddeley and Van Lieshout), the Free University of Amsterdam (Van Lieshout) and the University of Aarhus (Møller).

References

[1] BADDELEY, A. J. AND VAN LIESHOUT, M. N. M. (1996) A nonparametric measure of spatial interaction in point patterns. *Statist. Neerland.* To appear.

[2] BADDELEY, A. J. AND MØLLER, J. (1989) Nearest-neighbour Markov point processes and random sets. *Int. Statist. Rev.* **57,** 89–121.

[3] CARTER, D. S. AND PRENTER, P. M. (1972) Exponential spaces and counting processes. *Z. Wahrscheinlichkeitsth.* **21,** 1–19.

[4] DIGGLE, P. J., FIKSEL, T., OGATA, Y., STOYAN, D. AND TANEMURA, M. (1994) On parameter estimation for pairwise interaction processes. *Int. Statist. Rev.* **62,** 99–117.

[5] KENDALL, W. S. (1990) A spatial Markov property for nearest-neighbour Markov point processes. *J. Appl. Prob.* **28,** 767–778.

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

[7] MØLLER, J. (1994) Discussion contribution. *Scand. J. Statist.* **21,** 346–349.

[8] MØLLER, J. (1994) Markov chain Monte Carlo and spatial point processes. Research Report 293. Department of Theoretical Statistics, University of Aarhus. To appear in *Proc. Seminaire Européen de Statistique Toulouse 1996 'Stochastic Geometry: Theory and Applications'*. ed. O. Barndorff-Nielsen, W. S. Kendall, G. Letac and M. N. M. van Lieshout. Chapman and Hall, London.

[9] PRESTON, C. J. (1976) *Random Fields.* Springer, Berlin.

[10] RIPLEY, B. D. (1988) *Statistical Inference for Spatial Processes.* Cambridge University Press, Cambridge.

[11] RIPLEY, B. D. (1989). Gibbsian interaction models. In *Spatial Statistics: Past, Present and Future.* pp. 1–19. ed. D. A. Griffiths. Image, New York.

[12] RIPLEY, B. D. AND KELLY, F. P. (1977) Markov point processes. *J. London Math. Soc.* **15,** 188–192.

[13] RUELLE, D. (1969) *Statistical Mechanics.* Wiley, New York.