COMMUN. STATIST.—STOCHASTIC MODELS, 14(3), 715–734 (1998)

SHOT NOISE WEIGHTED PROCESSES: A NEW FAMILY OF SPATIAL POINT PROCESSES

M. N. M. van LIESHOUT
Centrum voor Wiskunde en Informatica
P.O. Box 94079, 1090 GB Amsterdam, The Netherlands

I. S. MOLCHANOV
Department of Statistics, University of Glasgow
Glasgow G12 8QW, United Kingdom

Abstract: This paper proposes a new family of spatial point processes defined by their density with respect to a Poisson process on a bounded window. The density will be specified in terms of functionals of shot noise processes with various influence functions, for example the coverage function. Stability, Markov properties, stationary extensions and limit behaviour are studied. Examples and simulated realisations are given to indicate the applicability of shot noise weighted processes.

Keywords: coverage function, Markov point process, spatial interaction, shot noise.

1. INTRODUCTION

The simplest point pattern model is the Poisson process. Its lack of interaction between the points makes it a useful benchmark process. Indeed, rejecting a Poisson null hypothesis is often the first step in fitting a more realistic model to a given set of data. However, many point patterns do exhibit interactions between the points which, for instance, can result in clustered or regular behaviour.

A class of models designed to incorporate interaction is that of Markov point processes, introduced in the statistical literature by Ripley and Kelly [25]. Similar models are known in statistical physics under the name of Gibbs distributions [23, 27]. These models are defined by their density $p(\cdot)$ with respect to a Poisson process, typically of the form

$$p(\mathbf{x}) = \alpha\beta^{n(\mathbf{x})}\gamma^{-\psi(\mathbf{x})}. \tag{1.1}$$

Here $\mathbf{x}$ is a finite set of unordered points (called a configuration) in a bounded set $A \subset \mathbb{R}^d$, $n(\mathbf{x})$ is the number of points in $\mathbf{x}$, and $\psi(\cdot)$ is a numerical functional governing the interaction structure.

One of the first examples of (1.1) is the Strauss model [31] for which $\psi(\mathbf{x})$ denotes the number of pairs of points not more than a given distance apart. Note that interactions occur between pairs of points only. Such models are suitable for regular point patterns, but probably not for clustered ones. See [10] and [19]. Indeed, the Strauss model is ill-defined for $\gamma > 1$.

The lack of Markov models for clustered behaviour led to the development of area-interaction processes [2]. A special case had already been considered as a model of liquid vapour equilibrium in chemical physics [34]. For area-interaction process, $\psi(\mathbf{x})$ in (1.1) denotes the volume of the 'molecules' associated to $\mathbf{x}$, for instance the union of unit balls centred at the points of $\mathbf{x}$. Several generalisations have been proposed recently, for instance those using functionals from convex geometry other than the volume (cf. [1] or [20] in the discrete case).

Another generalisation replaces the volume by integrals of a potential function depending on $\mathbf{x}$. For instance, [2] mentioned models of the type $\psi(\mathbf{x}) = \int_A f(d(\mathbf{x}, u))du$, where $d(\mathbf{x}, u) = \min_i ||x_i - u||$ and $f : [0, \infty] \mapsto (-\infty, \infty]$. A similar idea leads to the concept of a shot noise weighted process, as proposed here. This class of models has $\psi(\mathbf{x}) = \int_A f(\xi_{\mathbf{x}}(u))du$ for additive functionals $\xi_{\mathbf{x}}$ reminiscent of shot noise. The area-interaction model is a special case with $\xi_{\mathbf{x}}(u)$ counting the number of 'molecules' covering a point u, but distance based functionals are considered as well.

The plan of this paper is as follows. In Section 2 we introduce shot noise weighted point processes. Section 3 considers existence, Ruelle stability and local Markov properties. Section 4 is devoted to examples and simulations. Some basic properties are derived in Section 5.

2. SET-UP AND DEFINITIONS

In this paper, we are concerned with point processes X on a locally compact complete separable metric space A. We will distinguish between finite and locally finite point processes, concentrating mostly on the former.

A finite point process X is a random element in the space $\mathfrak{N}^f(A)$ of finite point configurations (or point measures)

$$\mathbf{x} = \{x_1, \ldots, x_n\}$$

where x_i, $i = 1, \ldots, n$ are elements of A and $n = n(\mathbf{x}) \geq 0$ denotes the number of points in $\mathbf{x}$. Note that the empty set and configurations with multiple points are allowed. Writing $\mathbf{x}_B$ for $\mathbf{x}$ restricted to $B \subseteq A$, the σ-algebra $\mathcal{N}^f(A)$ on $\mathfrak{N}^f(A)$ is the smallest σ-algebra with respect to which the evaluation $\mathbf{x} \mapsto n(\mathbf{x}_B)$ is measurable for every (bounded) Borel set $B \subseteq A$. For further details, consult [6] or [18].

Definition 2.1. An *influence function* is a non-negative Borel function

$$\kappa : A \times A \to [0, \infty).$$

The *influence zone* of κ at a point $b \in A$ is

$$Z_\kappa(b) = \{a \in A : \kappa(a, b) > 0\}.$$

With each finite point configuration $\mathbf{x}$, associate a function $\xi_{\mathbf{x}} : A \to [0, \infty)$ by

$$\xi_{\mathbf{x}}(a) = \sum_{i=1}^{n(\mathbf{x})} \kappa(a, x_i).$$

If X is a finite point process, then $\xi_X(a)$, $a \in A$, is said to be the *shot noise process* generated by X with influence function κ [5, 14].

We list some obvious properties of influence functions that will be used later.

Lemma 2.2. *Let $\kappa : A \times A \to [0, \infty)$ be an influence function. Then the following hold.*

1. *Seen as a function on $\mathfrak{N}^f(A) \times A$, $\xi_{\mathbf{x}}(a)$ is measurable with respect to the σ-algebra $\mathcal{N}^f(A) \otimes \mathfrak{B}(A)$, where $\mathfrak{B}(A)$ denotes the Borel σ-algebra on A.*
2. *$0 \leq \xi_{\mathbf{x}}(a) \leq n(\mathbf{x})\kappa^*$ for all $\mathbf{x} \in \mathfrak{N}^f(A)$ and $a \in A$, where $\kappa^* = \sup\{\kappa(a, b) : a, b \in A\}$.*
3. *$\xi_{\mathbf{x} \cup \mathbf{y}}(a) = \xi_{\mathbf{x}}(a) + \xi_{\mathbf{y}}(a)$ for all $\mathbf{x}, \mathbf{y} \in \mathfrak{N}^f(A)$, $a \in A$.*
4. *If $\kappa(b, a) = 0$ for some $a, b \in A$, then $\xi_{\mathbf{x} \cup \{a\}}(b) = \xi_{\mathbf{x}}(b)$.*

Proof. The first statement can be proved as follows. For each $t > 0$,

$$\{(\mathbf{x}, a) : \xi_{\mathbf{x}}(a) < t\} = \bigcup_{n=0}^{\infty} \{(\mathbf{x}, a) : n(\mathbf{x}) = n, \xi_{\mathbf{x}}(a) < t\} .$$

Each set $\{(\mathbf{x}, a) : n(\mathbf{x}) = n, \xi_{\mathbf{x}}(a) < t\}$ in the union above is a Borel set in $A^i \times A$ due to the fact that $\xi_{\mathbf{x}}(a)$ is measurable as a sum of measurable functions $\kappa(a, x_i)$. Hence the result follows.

Other statements can be verified by straightforward computation. □

In many applications, A will be $\mathbb{R}^d$ and the influence function $\kappa(\cdot, \cdot)$ will depend on the difference between its arguments only, i.e. $\kappa(a, x) = \kappa(a - x)$. In this case κ is said to be *homogeneous* and the associated influence zone at a point x reduces to

$$Z_\kappa(x) = Z_\kappa(0) + x,$$

the translation over x of the influence zone $Z_\kappa(0)$ of the origin.

Example 2.3. (COVERAGE FUNCTION) Let A be a compact subset of $\mathbb{R}^d$ and set

$$\kappa(a, x) = \mathbf{1}\{||a - x|| \leq r\}$$

Then $\kappa(\cdot, \cdot)$ is an influence function with $Z_\kappa(x) = B(x, r) \cap A$, the ball in A of radius r centred at x.

A configuration $\mathbf{x} = \{x_1, \ldots, x_n\}$ gives rise to

$$\xi_{\mathbf{x}}(a) = \sum_{i=1}^{n} \mathbf{1}\{a \in Z(x_i)\} . \tag{2.1}$$

Thus, $\xi_{\mathbf{x}}(a)$ counts the number of balls $Z(x_i)$ covering a. Later we denote $\xi_{\mathbf{x}}$ by $c_{\mathbf{x}}$, the coverage function of $\mathbf{x}$ by balls of radius r.

More generally, given a weakly measurable [33] function Z that maps A into the family $\mathcal{K}(A)$ of compact subsets of A, $\kappa(a, x) = \mathbf{1}\{a \in Z(x)\}$ is an influence function. Note that $Z_\kappa(x) = Z(x)$.

Influence functions can be used to define spatial point process distributions. As a motivating example, consider the image analysis problem of identifying objects parameterised by a function $Z : A \to \mathcal{K}(A)$ in a noisy image. It is natural to assume that objects may interact when their intersection is non-empty and that the more overlap, the stronger the interaction. Hence we are led to consider the coverage function (measuring the amount of overlap) combined with an inhibitory potential.

Writing π_μ for the distribution of a Poisson process on A with intensity measure $\mu(\cdot)$ satisfying $0 < \mu(A) < \infty$, we define new processes by their density (or Radon-Nikodym derivative) with respect to π_μ.

Definition 2.4. A shot noise weighted process with *potential function* $f(\cdot)$ is a point process on A with density

$$p(\mathbf{x}) = \alpha \beta^{n(\mathbf{x})} \gamma^{-\int_A f(\xi_{\mathbf{x}}(a)) d\nu(a)} \tag{2.2}$$

with respect to π_μ. Here $\beta, \gamma > 0$ are model parameters, $\nu(\cdot)$ a finite Borel measure and $f : \mathbb{R} \mapsto \mathbb{R}$ a Borel function with $f(0) = 0$. The normalising constant, α, ensures that the integral of $p(x)$ is equal to 1.

Clearly, (2.2) is a particular case of (1.1). Conditions must be imposed to ensure that the model (2.2) is integrable, see Section 3. Examples and simulated realisations will be given in Section 4.

Note that (2.2) is overparameterised: taking $f_1(\cdot) = cf(\cdot)$ for some constant c is equivalent to changing γ to $\gamma_1 = \gamma^c$. However, if $f(\cdot)$ is absolutely integrable, this ambiguity can be overcome by requiring that the integral of $f(\cdot)$ is 1. The general set-up described above allows for multiple points. If this is undesirable, the reference measure $\mu(\cdot)$ must be diffuse. A generic example is taking for A a compact subset of $\mathbb{R}^d$ equipped with Lebesgue measure $\mu(\cdot)$.

3. EXISTENCE AND MARKOV PROPERTY

The existence of a point process given in terms of its density $p(\cdot)$ with respect to π_μ is ensured by Ruelle's stability condition [10, 27]. This condition requires that the energy $E(\mathbf{x}) = -\log(p(\mathbf{x})/p(\emptyset))$ has a lower bound that is linear in the number of points in $\mathbf{x}$, i.e.

$$E(\mathbf{x}) \geq -Cn(\mathbf{x}) \tag{3.1}$$

for some $C > 0$ and all $\mathbf{x}$ with $p(\mathbf{x}) > 0$. When (3.1) holds, $p(\cdot)$ (or the corresponding energy) is called *stable*. In our case (Definition 2.4), Ruelle's condition requires a linear bound on $|\int_A f(\xi_{\mathbf{x}}(a)) d\nu(a)|$ in terms of $n(\mathbf{x})$. Thus, it is sufficient to require

$$|f(\xi_{\mathbf{x}}(a))| \leq Cn(\mathbf{x}), \quad a \in A, \quad \mathbf{x} \in \mathfrak{N}^f(A), \tag{3.2}$$

for some $C > 0$.

For instance, consider the coverage function $c_{\mathbf{x}}(a)$ from Example 2.3. Since $c_{\mathbf{x}}(a) \leq n(\mathbf{x})$ for all $a \in A, x \in \mathfrak{N}^f(A)$, (3.2) is satisfied whenever

$$|f(t)| \leq Ct, \quad t \in \mathbb{R}, \tag{3.3}$$

for some $C > 0$. Note that in the situation described in Example 2.3, $f(\cdot)$ can be represented as a sequence $\{f(n), n \geq 1\}$ so that (3.3) needs to be checked for positive integers t only.

We can ensure (3.2) in a more general setting by assuming (3.3) and $\sup_{a,x} \kappa(a, x) < \infty$. For arbitrary $\kappa(\cdot, \cdot)$, if the potential function $f(\cdot)$ is bounded, (3.2) holds as well. For bounded potentials moreover, the distribution of the shot noise weighted process is uniformly absolutely continuous with respect to the distribution of a Poisson process $\pi_{\beta\mu}$ with intensity measure $\beta\mu(\cdot)$, i.e. its Radon–Nikodym derivative is uniformly bounded in $\mathbf{x}$. Examples of uniformly absolutely continuous processes include the standard area-interaction model but also 'take it or leave it' type potentials with binary values. We come back to this later in Section 4.

Note that by Lemma 2.2 and Fubini's theorem $\int_A f(\xi_{\mathbf{x}}(a))d\nu(a)$ is $\mathcal{N}^f(A)$-measurable. Hence, the function $p : \mathfrak{N}^f(A) \to \mathbb{R}$ is measurable too. Summarising we obtain the following result.

Lemma 3.1. *Under condition (3.2), density (2.2) is measurable and integrable for all values of $\beta, \gamma > 0$.*

From now on we will assume that the stability condition (3.2) holds.

In order to describe the interaction behaviour of shot noise models, we need to define a neighbourhood relation $\sim$ on A. As in [2], let $a \sim b$ if and only if their influence zones overlap:

$$Z_\kappa(a) \cap Z_\kappa(b) \neq \emptyset . \tag{3.4}$$

Following Ripley and Kelly [25], a point process given by its density $p(\cdot)$ is said to be *Markov* with respect to $\sim$ if, for all configurations $\mathbf{x} \in \mathfrak{N}^f(A)$,

(a) $p(\mathbf{x}) > 0$ implies $p(\mathbf{y}) > 0$ for all $\mathbf{y} \subseteq \mathbf{x}$;

(b) if $p(\mathbf{x}) > 0$, then $p(\mathbf{x} \cup \{u\})/p(\mathbf{x})$ depends only on u and $\{x_i \in \mathbf{x} : u \sim x_i\}$, the set of all points in $\mathbf{x}$ which are neighbours of u.

For generalisations see [3]. Usually the influence zones will be small compared to the observation window; otherwise from a computational point of view far too many points will be neighbours (cf. Section 4).

Theorem 3.2. *The shot noise weighted process (Definition 2.4) is Ripley–Kelly Markov with respect to the relation (3.4).*

Proof. Note that there are no zero-likelihood configurations, hence the density is hereditary (condition **(a)**). To check **(b)**, write

$$\frac{p(\mathbf{x} \cup \{u\})}{p(\mathbf{x})} = \beta \exp\left[-(\log\gamma)\left\{\int_A f(\xi_{\mathbf{x}\cup\{u\}}(a))d\nu(a) - \int_A f(\xi_{\mathbf{x}}(a))d\nu(a)\right\}\right] .$$

Split $\mathbf{x}$ in $\mathbf{y} \cup \mathbf{z}$ where

$$\mathbf{y} = \{x_i \in \mathbf{x} : Z_\kappa(x_i) \cap Z_\kappa(u) = \emptyset\}$$

and $\mathbf{z} = \mathbf{x} \backslash \mathbf{y}$. By Lemma 2.2, for all $a \notin Z_\kappa(u)$ we have that $\xi_{\mathbf{x}\cup\{u\}}(a) = \xi_{\mathbf{x}}(a)$. On the other hand, for $a \in Z_\kappa(u)$,

$$\xi_{\mathbf{x}\cup\{u\}}(a) = \sum_{y_i \in \mathbf{y}} \kappa(a, y_i) + \sum_{z_i \in \mathbf{z}} \kappa(a, z_i) + \kappa(a, u) = 0 + \xi_{\mathbf{z}}(a) + \kappa(a, u) .$$

Similarly, $\xi_{\mathbf{x}}(a) = \xi_{\mathbf{z}}(a)$ for $a \in Z_\kappa(u)$. Hence

$$\begin{aligned} &\frac{p(\mathbf{x} \cup \{u\})}{p(\mathbf{x})} \\ &\quad = \beta \exp\left[-(\log\gamma)\left\{\int_{Z_\kappa(u)} [f(\xi_{\mathbf{z}}(a) + \kappa(a,u)) - f(\xi_{\mathbf{z}}(a))]\, d\nu(a)\right\}\right] \end{aligned} \tag{3.5}$$

completing the proof. □

One of the most important results for Markov point processes is the Hammersley–Clifford theorem [25] stating that $p(\mathbf{x})$, $\mathbf{x} \in \mathfrak{N}^f(A)$, can be factorised as a product of *clique interaction functions*

$$p(\mathbf{x}) = \prod_{\substack{\mathbf{y} \subseteq \mathbf{x} \\ \mathbf{y} \text{ is a clique}}} \phi(\mathbf{y}).$$

Here a clique is any configuration $\mathbf{x}$ for which all its members are neighbours ($s \sim t$ for all $s, t \in \mathbf{x}$). The following result gives the interaction functions $\phi(\cdot)$ for the shot noise model.

Theorem 3.3. *The interaction functions of a shot noise weighted process (2.2) are*

$$\phi(\emptyset) = \alpha, \tag{3.6}$$

$$\phi(\{u\}) = \beta\gamma^{-\int_A f(\xi_{\{u\}}(a))d\nu(a)}, \tag{3.7}$$

$$\phi(\mathbf{x}) = \exp\left\{-(\log\gamma)\int_A \sum_{\mathbf{y}\subseteq\mathbf{x}} (-1)^{n(\mathbf{x}\setminus\mathbf{y})} f(\xi_{\mathbf{y}}(a))d\nu(a)\right\}, n(\mathbf{x}) \geq 2. \tag{3.8}$$

Proof. The proof is based on induction with respect to the number of points. The case $n(\mathbf{x}) = 0$ is straightforward from the Hammersley–Clifford formula. For $n(\mathbf{x}) = 1$, note that $\phi(\{u\}) = p(\{u\})/\phi(\emptyset)$.

If $n(\mathbf{x}) = 2$, then for $u \neq v \in A$,

$$\phi(\{u,v\}) = \frac{p(\{u,v\})}{\phi(\emptyset)\phi(\{u\})\phi(\{v\})} = \gamma^{-\int_A \left(f(\xi_{\{u,v\}}(a)) - f(\xi_{\{u\}}(a)) - f(\xi_{\{v\}}(a))\right)d\nu(a)}$$

in accordance with (3.8).

Next assume that formula (3.8) holds for configurations with up to $n \geq 2$ points and let $\mathbf{x}$ be such that $n(\mathbf{x}) = n$. Then, writing '$\subset$' for the proper subset ordering,

$$\begin{aligned} \phi(\mathbf{x} \cup \{u\}) &= \frac{p(\mathbf{x} \cup \{u\})}{\prod_{\mathbf{y}\subset\mathbf{x}\cup\{u\}} \phi(\mathbf{y})} \\ &= \gamma^{-\int_A \left[f(\xi_{\mathbf{x}\cup\{u\}}(t)) - \sum_{\mathbf{y}\subset\mathbf{x}\cup\{u\}} \sum_{\mathbf{z}\subseteq\mathbf{y}} (-1)^{n(\mathbf{y}\setminus\mathbf{z})} f(\xi_{\mathbf{z}}(t))\right] d\nu(t)}. \end{aligned}$$

Now

$$\sum_{\mathbf{y}\subset\mathbf{x}\cup\{a\}}\sum_{\mathbf{z}\subseteq\mathbf{y}}(-1)^{n(\mathbf{y}\setminus\mathbf{z})}f(\xi_{\mathbf{z}}(t)) = \sum_{\mathbf{z}\subset\mathbf{x}\cup\{a\}} f(\xi_{\mathbf{z}}(t)) \sum_{\mathbf{z}\subseteq\mathbf{y}\subset\mathbf{x}\cup\{a\}} (-1)^{n(\mathbf{y}\setminus\mathbf{z})}$$
$$= -\sum_{\mathbf{z}\subset\mathbf{x}\cup\{a\}} (-1)^{n(\mathbf{x}\cup\{a\}\setminus\mathbf{z})} f(\xi_{\mathbf{z}}(t)),$$

where we have used that by Newton's binomium the inner sum in the penultimate formula above equals $-(-1)^{n(\mathbf{x}\cup\{a\}\setminus\mathbf{z})}$. Hence

$$f(\xi_{\mathbf{x}\cup\{a\}}(t)) - \sum_{\mathbf{y}\subset\mathbf{x}\cup\{a\}}\sum_{\mathbf{z}\subseteq\mathbf{y}}(-1)^{n(\mathbf{y}\setminus\mathbf{z})}f(\xi_{\mathbf{z}}(t)) = \sum_{\mathbf{z}\subseteq\mathbf{x}\cup\{a\}} (-1)^{n(\mathbf{x}\cup\{a\}\setminus\mathbf{z})} f(\xi_{\mathbf{z}}(t))$$

and the proof is complete. □

The highest $n(\mathbf{x})$ with $\phi(\mathbf{x}) \neq 1$ is said to be the order of interaction. In most cases shot noise weighted processes exhibit infinite order of interactions. Note that two such processes generated by the functions $f(t)$ and $f(t) + ct$ for some $c \in \mathbb{R}$ have the same order of interactions, so that the linear part of f is not important to determine interaction order.

It is easy to verify that $\phi(\mathbf{x}) = 1$ whenever $\mathbf{x}$ is not a clique: take $u, v \in \mathbf{x}$ with $u \not\sim v$ and rewrite the integrand in the exponent of (3.8) as

$$\sum_{\mathbf{y}\subseteq\mathbf{x}\setminus\{u,v\}} (-1)^{n(\mathbf{x}\setminus\mathbf{y})} \left[f(\xi_{\mathbf{y}}(a)) + f(\xi_{\mathbf{y}\cup\{u,v\}}(a)) - f(\xi_{\mathbf{y}\cup\{u\}}(a)) - f(\xi_{\mathbf{y}\cup\{v\}}(a))\right].$$

Note that for $a \in Z(u)$, $f(\xi_{\mathbf{y}}(a)) = f(\xi_{\mathbf{y}\cup\{v\}}(a))$ and also $f(\xi_{\mathbf{y}\cup\{u\}}(a)) = f(\xi_{\mathbf{y}\cup\{u,v\}}(a))$; for $a \notin Z(u)$ on the other hand, $f(\xi_{\mathbf{y}}(a)) = f(\xi_{\mathbf{y}\cup\{u\}}(a))$ and $f(\xi_{\mathbf{y}\cup\{v\}}(a)) = f(\xi_{\mathbf{y}\cup\{u,v\}}(a))$. Hence $\phi(\mathbf{x}) = \gamma^0 = 1$.

4. EXAMPLES AND INTERPRETATIONS

In this Section we give some specific examples of shot noise weighted processes and suggestions on how to simulate them.

Stationary Poisson process. If $\gamma = 1$, regardless of the choice of potential and influence function, the shot noise weighted model is a Poisson process with intensity measure $\beta\mu(\cdot)$. The interaction functions of order 2 and higher are identically 1, in accordance with the Poisson process' interpretation of 'spatial randomness' [7]. The lower order interaction functions are $\phi(\emptyset) = e^{(1-\beta)\mu(A)}$ and $\phi(\{u\}) = \beta$, $u \in A$.

Inhomogeneous Poisson process. If the potential function $f(\cdot)$ is linear, i.e. $f(t) = ct$ for all $t \in \mathbb{R}$ and some $c \in \mathbb{R}$, (2.2) defines an (inhomogeneous) Poisson process with intensity measure $\beta\gamma^{-c\nu_\kappa(\cdot)}\mu(\cdot)$, where

$$\nu_\kappa(x) = \int_A \kappa(a, x) d\nu(a) \,, \quad x \in A \,.$$

Again, there are no interactions of orders higher than 1,

$$\phi(\emptyset) = \exp\left[\int_A (1 - \beta\gamma^{-c\nu_\kappa(a)}) d\mu(a)\right] ,$$

but now $\phi(\{u\}) = \beta\gamma^{-c\nu_\kappa(u)}$ may be location dependent.

Area-interaction process. An example that does exhibit interactions between the members in a configuration is the area–interaction model [2]. This model is obtained by taking the coverage function introduced in Example 2.3 and a potential function with $f(n) = 1$ for all strictly positive integers, zero otherwise:

$$p(\mathbf{x}) = \alpha\beta^{n(\mathbf{x})}\gamma^{-\nu(U(\mathbf{x}))}.$$

The model produces clustered patterns for values of $\gamma > 1$, regular ones for $\gamma < 1$. The special case $\gamma > 1$, $Z(x) = B(x, r)$, a ball of fixed radius centred at x and Lebesgue measure for $\nu(\cdot)$ is the *penetrable sphere model* introduced in [34] for liquid-vapour equilibrium. These models have interactions of arbitrary large order, except in the Poisson case $\gamma = 1$. For details, see [2, 13, 26, 34].

A typical realisation of the penetrable sphere model, obtained using the exact Gibbs sampler of [12] is given in Figure 4.1. Other simulated realisations can be found in [2, 12, 17].

Truncated at 1. Another example that can deal with clustering and inhibition is again taking the framework of Example 2.3 but now truncating the binary potential function at 1, so that $f(n) = \mathbf{1}\{n = 1\}$, $n \geq 1$. Then

$$p(\mathbf{x}) = \alpha\beta^{n(\mathbf{x})}\gamma^{-\nu(\{a : c_{\mathbf{x}}(a) = 1\})} .$$

Note that $p(\cdot)$ is integrable and uniformly absolutely continuous with respect to $\pi_{\beta\mu}$, since $|f(\cdot)|$ is bounded by 1 (cf. Section 3). For $\gamma > 1$ the model tends to have smallish 1–covered regions, indicating clustering. For $\gamma < 1$ the model assigns most mass to configurations having largish 1–covered regions, hence inhibition. Like the area-interaction process, this model has interactions of infinite order.

Direct simulation from $p(\cdot)$ is difficult due to the dimension and the normalising constant α that cannot be evaluated explicitly. On the other hand, the conditional intensities (3.5) are 'local' and easy to compute provided the influence zones are not too large. This observation can be used to construct Markov chain Monte Carlo samplers whose transition probabilities are based on the likelihood ratio (3.5). For our shot noise process, we have used the Metropolis–Hastings sampler of [11]. Briefly, given an initial configuration $\mathbf{x}_0$,

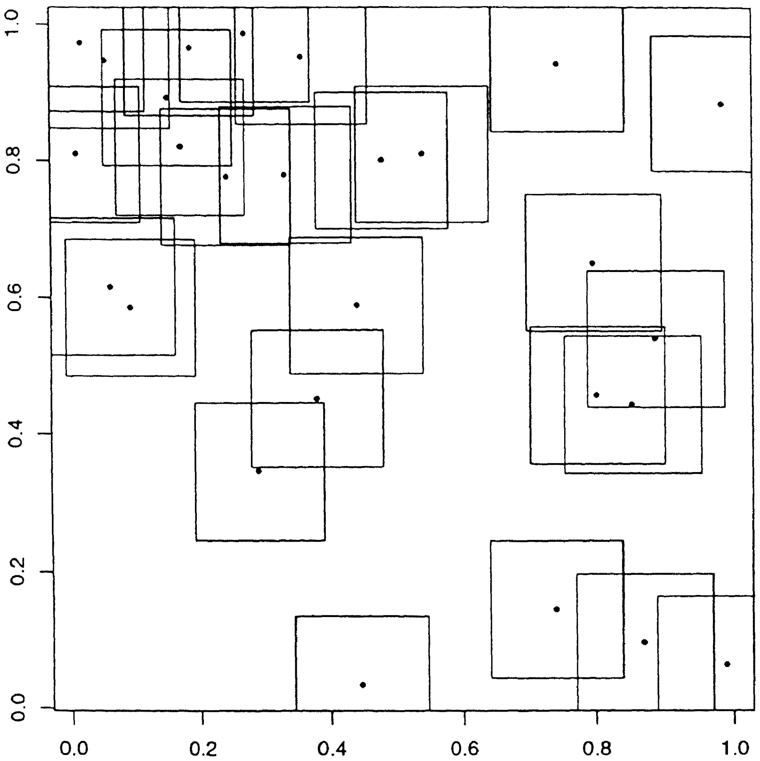

Figure 4.1: Realisation of the (symmetric) penetrable sphere model with $\beta = 50$, $\gamma = e^{50}$ and square influence zones of side .2

with probability 1/2 propose adding a point ('birth'); with probability 1/2 propose deleting one of the points in $\mathbf{x}$ if any ('death'). Births are proposed uniformly with respect to $\mu(\cdot)$ on A and accepted with probability

$$\min\left\{1, \frac{p(\mathbf{x}_0 \cup \{u\})\mu(A)}{p(\mathbf{x}_0)(n(\mathbf{x}_0)+1)}\right\}$$

If the new point u is accepted, set $\mathbf{x}_1 = \mathbf{x}_0 \cup \{u\}$; otherwise $\mathbf{x}_1 = \mathbf{x}_0$. Similarly, we delete x_i from $\mathbf{x}_0 = \{x_1, \ldots, x_n\}$ with probability $1/n$. The death of x_i is accepted with probability

$$\min\left\{1, \frac{p(\mathbf{x}_0 \setminus \{x_i\})n(\mathbf{x}_0)}{p(\mathbf{x}_0)\mu(A)}\right\}$$

and if so, $\mathbf{x}_1 = \mathbf{x}_0 \setminus \{x_i\}$. Otherwise $\mathbf{x}_1 = \mathbf{x}_0$. Continuing in this fashion, we obtain a sequence $\mathbf{x}_k, k \in \mathbb{N}_0$, which converges to $p(\cdot)$ as $k \to \infty$.

Some patterns generated this way are given in Figure 4.2. The time series of the sufficient statistics $n(\mathbf{x})$ and $\nu(\{a : c_{\mathbf{x}}(a) = 1\})$ are plotted in Figure 4.3.

Truncated at k. More generally, consider the previous example but truncate at $k > 1$, so that $f(n) = \mathbf{1}\{n \leq k\}$, $n \geq 1$. The corresponding density

$$p(\mathbf{x}) = \alpha \beta^{n(\mathbf{x})} \gamma^{-\nu(\{a : c_{\mathbf{x}}(a) \leq k\})}$$

allows some more overlap than for the case $k = 1$. For $\gamma < 1$, likely configurations contain m-tuples with $m \leq k$.

Pair coverage interaction. Consider the coverage function from Example 2.3 with the potential function $f(n) = \mathbf{1}\{n = 2\}$. Then $\int_A f(\xi_{\mathbf{x}}(a))\nu(da)$ is the ν-measure of the set $U_2(\mathbf{x})$ of points in A covered by exactly two sets $Z(x_i)$, $i = 1, \ldots, n(\mathbf{x})$, and, therefore,

$$p(\mathbf{x}) = \alpha \beta^{n(\mathbf{x})} \gamma^{-\nu(U_2(\mathbf{x}))}.$$

For $\gamma > 1$ there tend to be many high order overlaps or no overlaps at all, while, for $\gamma < 1$, objects tend to come in pairs.

Note that $f(\cdot)$ takes binary values. Hence the pair coverage interaction model is well-defined and uniformly absolutely continuous with respect to $\beta\mu(\cdot)$ (Section 3). For simulation purposes, we used the Metropolis–Hastings sampler described above for the truncation model. Some examples are shown in Figures 4.4–4.5.

Odd and even. Again take the coverage function example with one of the following potential functions:

1. $f(2k-1) = 1$, $f(2k) = 0$, $k \geq 1$;
2. $f(2k-1) = 2k-1$, $f(2k) = 0$, $k \geq 1$.

In both cases X has interactions of infinite order. If $\gamma > 1$, then in both cases points tend to come in clusters with even numbers of points. For $\gamma < 1$ the models tend to have odd coverage, model 2 particularly favouring high amounts of overlap.

Now, $f(\cdot)$ is linearly bounded in absolute value. Thus, by Lemma 3.1, the odd and even model is well-defined.

Distance influence function. Let us consider an example not related to the coverage function. Set $\kappa(a, x) = ||a - x||$, the distance between a and x in the observation window $A \subset \mathbb{R}^2$. Furthermore, let $\nu(\cdot)$ be Lebesgue measure

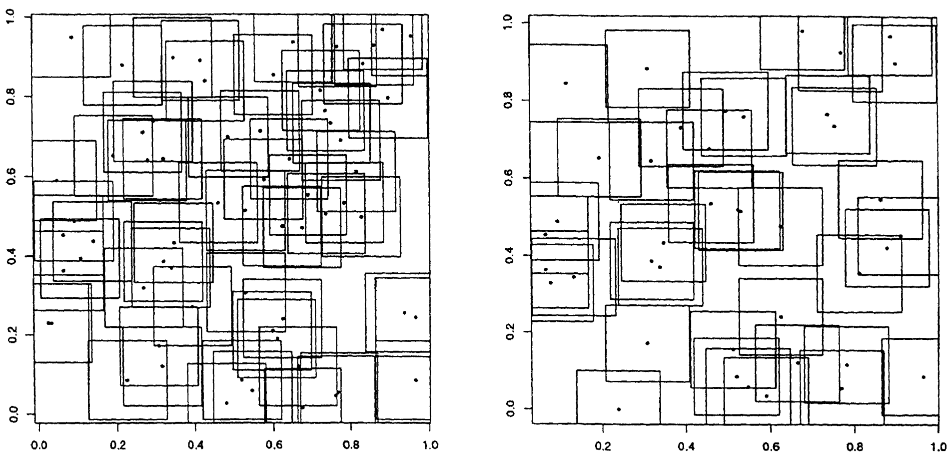

Figure 4.2: Samples of the 1-truncated model after 20000 steps of the Metropolis–Hastings algorithm for $\beta = 50$ and square influence zones of side .2; the values of γ are e^{50} (left) and e^{-50} (right).

and set $f(t) = \mathbf{1}\{t \leq 1\}$. Then the associated density is well-defined and is of the form

$$p(\mathbf{x}) = \alpha \beta^{n(\mathbf{x})} \gamma^{-\nu(M_1(\mathbf{x}))},$$

where $M_1(\mathbf{x}) = \left\{a \in A : \sum_{i=1}^{n(\mathbf{x})} ||a - x_i|| \leq 1\right\}$. If $\mathbf{x}$ contains at least two points at distance greater than 1, then $\nu(M_1(\mathbf{x})) = 0$, so that $\gamma > 1$ makes such configurations more probable. On the contrary, for $\gamma < 1$ points tend to appear in *one* cluster which is contained within a ball of radius 1. For instance, if $n(\mathbf{x}) = 2$, then $M_1(\mathbf{x})$ is an ellipse with foci x_1 and x_2.

5. PROPERTIES

In this Section we consider the behaviour of shot noise weighted point processes under some elementary operations, as well as their limit behaviour.

First note that the family of shot noise weighted point processes is closed under taking Radon-Nikodym derivatives. This means that the densities of such a process with respect to another process from this family has the same form (2.2).

If A, κ and ν are group-invariant (e.g., with respect to rotations), then X is distribution-invariant with respect to the same group.

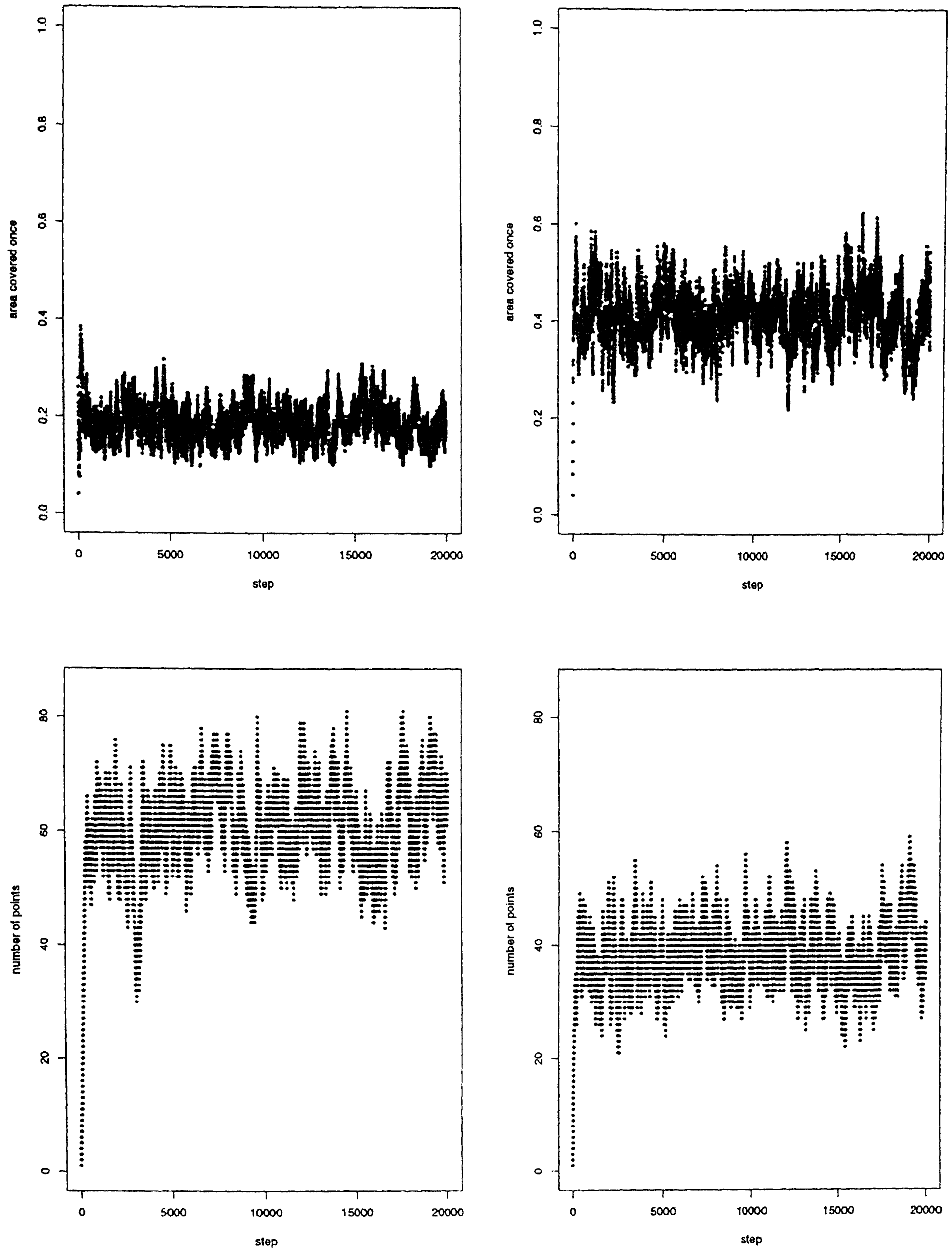

Figure 4.3: Time series of the sufficient statistics of the 1-truncated model over 20000 steps of the Metropolis-Hastings algorithm for $\beta = 50$ and square influence zones of side .2; the values of γ are e^{50} (left) and e^{-50} (right).

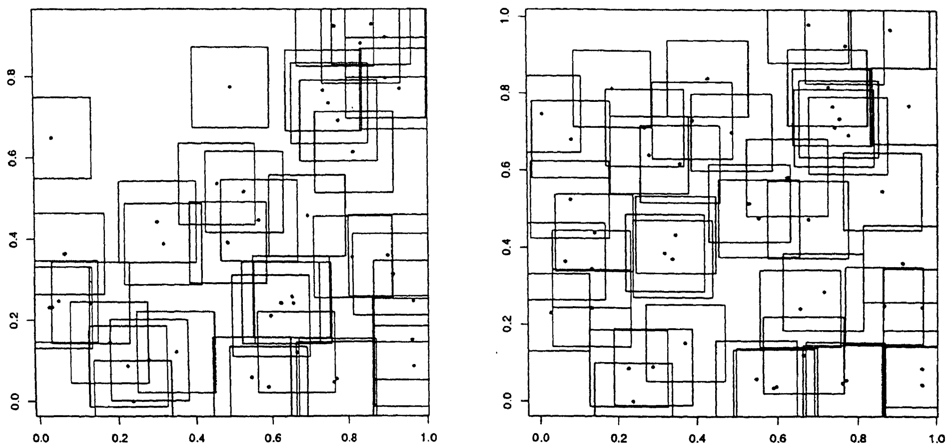

Figure 4.4: Samples of the pair coverage model after 20000 steps of the Metropolis–Hastings algorithm for $\beta = 50$ and square influence zones of side .2; the values of γ are e^{50} (left) and e^{-50} (right).

If two independent point processes X_1, X_2 are absolutely continuous with respect to π_μ with densities $p_1(\cdot)$ and $p_2(\cdot)$ respectively, then the *superposition* $X_1 \cup X_2$ is also absolutely continuous with respect to π_μ and has density

$$p(\mathbf{x}) = e^{-\mu(A)} \sum_{\varphi:\mathbf{x}\to\{1,2\}} p_1(\varphi^{-1}(1))p_2(\varphi^{-1}(2)) \tag{5.1}$$

where the sum is over all ordered partitions of $\mathbf{x} \in \mathfrak{N}^f(A)$. In the case of two independent shot noise weighted processes, (5.1) reads

$$p(\mathbf{x}) = e^{-\mu(A)}\alpha^2\beta^{n(\mathbf{x})} \times \sum_{\varphi:\,\mathbf{x}\to\{1,2\}} \exp\left[-\log\gamma \int_A \left(f(\xi_{\varphi^{-1}(1)}(a)) + f(\xi_{\varphi^{-1}(2)}(a))\right) d\nu(a)\right].$$

Hence, in general, $X_1 \cup X_2$ is not a shot noise weighted process; it is for linear potential function $f(\cdot)$, a well-known property of Poisson processes [7, 30].

Next consider independent thinning: each point in a realisation of a point process is retained independently of every other point with probability p. Then the Janossy densities of the thinned shot noise weighted process (cf. [6, p. 122]) are given by

$$j_n^{th}(x_1,\ldots,x_n) = \sum_{m=0}^{\infty} \frac{1}{m!} \int_{A^m} p^n j_{n+m}(x_1,\ldots,x_n,y_1,\ldots,y_m)(1-p)^m d\mu(y_1)\ldots d\mu(y_m)$$

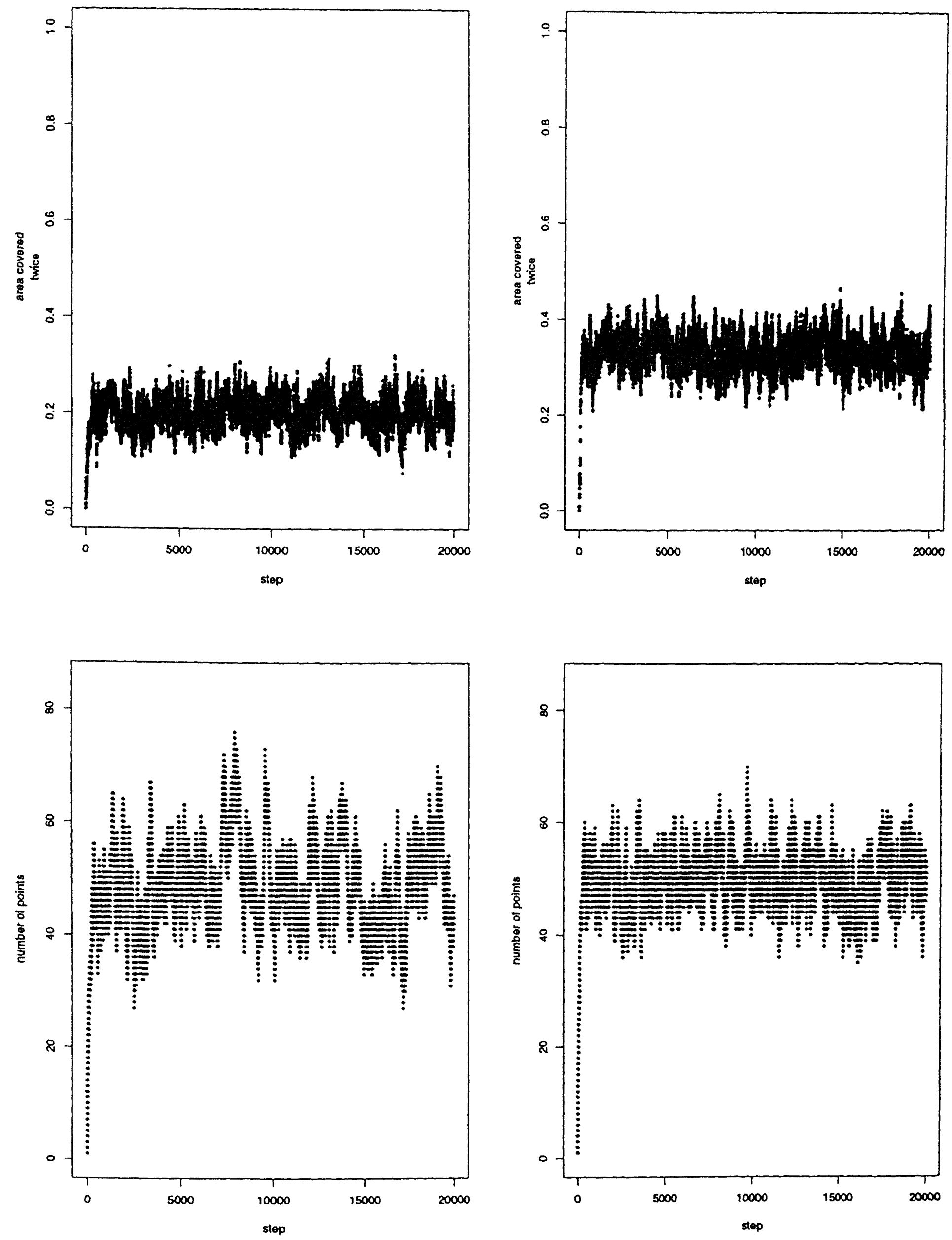

Figure 4.5: Time series of the sufficient statistics of pair coverage model over 20000 steps of the Metropolis–Hastings algorithm for $\beta = 50$ and square influence zones of side .2; the values of γ are e^{50} (left) and e^{-50} (right).

$$= \alpha e^{-\mu(A)}(p\beta)^n \sum_{m=0}^{\infty} \frac{1}{m!} \int_{A^m} (\beta(1-p))^m \gamma^{-\int_A f(\xi_{x\cup y}(a))d\nu(a)} d\mu(y_1)\dots d\mu(y_m).$$

Hence, the density of the thinned process with respect to π_μ is

$$p^{th}(\mathbf{x}) = \alpha(p\beta)^{n(\mathbf{x})} e^{\mu(A)} \int_{\mathfrak{N}^f(A)} (\beta(1-p))^{n(\mathbf{y})} \gamma^{-\int_A f(\xi_{\mathbf{x}\cup\mathbf{y}}(a))d\nu(a)} d\pi_\mu(\mathbf{y}).$$

Equivalently, the thinned process is absolutely continuous with respect to the original process with density

$$\frac{p^{th}(\mathbf{x})}{p(\mathbf{x})} = p^{n(\mathbf{x})} e^{\mu(A)} \mathbf{E}_{\pi_\mu}\left[(\beta(1-p))^{n(Y)} \gamma^{-\int_A [f(\xi_{\mathbf{x}\cup Y}(a)) - f(\xi_{\mathbf{x}}(a))]d\nu(a)}\right]. \quad (5.2)$$

Again, if $f(\cdot)$ is a linear function (Poisson case), then the thinned process is Poisson with intensity measure $p\beta\gamma^{-\nu_\kappa(\cdot)}\mu(\cdot)$ [30].

Let us study the convergence of the shot noise weighted point process as $\gamma \to 0$ or ∞.

Theorem 5.1. *Let $P_{\beta,\gamma}$ be the distribution of the shot noise weighted process with density (2.2) for given influence and potential functions.*

1. *Assume that $f(\cdot)$ is bounded and let*
$$\mathsf{H} = \left\{\mathbf{x} : \int_A f(\xi_{\mathbf{x}}(a))d\nu(a) = \max_{\mathbf{x}} \int_A f(\xi_{\mathbf{x}}(a))d\nu(a)\right\}.$$
Then, as $\gamma \to 0$ for fixed β, $P_{\beta,\gamma}$ converges in distribution to a uniform process on H, i.e. a random configuration corresponding to the distribution $\pi_{\beta\mu}$ on H.
2. *Suppose that $\beta \to 0$ and $\gamma \to 0$ in such a way that $\beta\gamma^{-\nu_\kappa(a)} \to \zeta(a) \in (0,\infty)$, $a \in A$. If f is positive and sublinear, i.e. $f(t+s) \le f(t) + f(s)$ for all s and t, then $P_{\beta,\gamma}$ converges in distribution to a Poisson process with intensity $\zeta(a)\mu(da)$ restricted to the set of configurations*
$$\mathsf{HC} = \left\{\mathbf{x} : \int_A f(\sum_{i=1}^{n(\mathbf{x})} \kappa(a,x_i))d\nu(a) = \sum_{i=1}^{n(\mathbf{x})} \int_A f(\kappa(a,x_i))d\nu(a)\right\}.$$
3. *If f is strictly positive except in 0, then as $\gamma \to \infty$ with $\beta < \infty$ fixed, $P_{\beta,\gamma}$ converges to a process that is empty with probability 1.*

Proof.

1. Write $f^* = \max_{\mathbf{x}} \int f(\xi_{\mathbf{x}}(a))d\nu(a)$. Then
$$\int \gamma^{f^* - \int f(\xi_{\mathbf{x}}(a))d\nu(a)} d\pi_{\beta\mu}(\mathbf{x}) \to \pi_{\beta\mu}(\mathsf{H}) \quad \text{as} \quad \gamma \to 0.$$
Hence
$$p_{\beta,\gamma}(\mathbf{x}) = \frac{\gamma^{f^* - \int f(\xi_{\mathbf{x}}(a)d\nu(a)}}{\int \gamma^{f^* - \int f(\xi_{\mathbf{x}}(a)d\nu(a)} d\pi_{\beta\mu}(\mathbf{x})} \to \frac{\mathbf{1}_{\mathbf{x}\in\mathsf{H}}}{\pi_{\beta\mu}(\mathsf{H})} \quad \text{as} \quad \gamma \to 0,$$
which is equivalent to the first assertion.

2. By sublinearity,

$$\gamma^{\int(\sum_i f(\kappa(a,x_i))-f(\sum_i \kappa(a,x_i)))d\nu(a)} \to 0 \quad \text{as} \quad \gamma \to 0\,.$$

if $\mathbf{x} \notin \mathsf{HC}$. Hence

$$p_{\beta,\gamma}(\mathbf{x}) \to \frac{\prod_{i=1}^{n(\mathbf{x})} \zeta(x_i)\, \mathbf{1}_{\mathbf{x}\in\mathsf{HC}}}{\int_{\mathsf{HC}} \prod_{i=1}^{n(\mathbf{x})} \zeta(x_i)\, d\pi_\mu(\mathbf{x})} \quad \text{as} \quad \gamma \to 0\,.$$

3. Note that the density converges pointwise to zero unless the pattern is empty. □

Note that the set HC in Theorem 5.1 contains at least all singletons and also point configurations which are similar to realisations of hard-core point processes [30].

Limit behaviour of a different kind occurs when the space A expands. Suppose on each compact subset A of $\mathbb{R}^d$ we have defined a shot noise weighted process by a fixed homogeneous influence function $\kappa(\cdot,\cdot)$ with bounded zone Z_κ and with potential $f(\cdot)$. Provided that a stability condition similar to (3.2) is satisfied for each A, using methods of Preston [23] in the same way as it has been done in [2] (with evident changes) one can prove that there exists a stationary extension to the whole of $\mathbb{R}^d$. However, the extension need not be unique, i.e. there may be phase transition [23, p. 46]. Indeed, the area-interaction process exhibits a phase transition [28]. Existence of a stationary extension is important with respect to edge effects, indicates that shot noise weighted models can be considered as the restriction to a bounded sampling window of a stationary point process and justifies estimation techniques such as the Takacs–Fiksel method (see below).

Jensen [15] proved a central limit theorem for stationary Gibbs point processes. Assuming again a homogeneous influence function with bounded zone and stability for each compact $A \subset \mathbb{R}^d$, it can be shown that the associated shot noise weighted process satisfies the conditions of Theorem 2.2 in [15]. In particular, this yields a central limit theorem for *additive* functionals such as the number of points. However, the integrated shot noise [14]

$$\int_A f(\xi_{\mathbf{x}}(a))\, d\nu(a)\,,$$

in general is a non-additive functional of the observation window.

Statistical estimation of the parameters β and γ of shot noise weighted processes can be performed by means of one of the following techniques.

- maximum likelihood techniques using Markov chain Monte Carlo simulations or stochastic approximation to evaluate the unknown normalisation parameter α [8, 21, 22];

- the Takacs–Fiksel estimation method, see [9, 32] and [24, p. 54–55];
- solving the pseudo-likelihood equations [4, 16], which are of the same form as the pseudo-likelihood equations for the Strauss model [24, p. 53] and are a special case of the Takacs–Fiksel method, see [2, 8, 29].

The estimation of $f(\cdot)$ and $\kappa(\cdot,\cdot)$ is a difficult non-parametric statistical problem.

ACKNOWLEDGEMENTS

This research was carried out while Molchanov was at the Centre for Mathematics and Computer Science (CWI), Van Lieshout at the Department of Statistics, University of Warwick, and was facilitated by a visiting grant of the European Science Foundation's initiative on highly structured stochastic systems. Van Lieshout's research has been supported by grant SCI/180/94/103 "Applications of stochastic geometry in the analysis of spatial data" of the Nuffield foundation; Molchanov was supported by the Netherlands Organisation for Scientific Research (NWO). The authors are also grateful to CWI and Warwick University for hospitality and to the referees for helpful comments.

REFERENCES

[1] Baddeley, A.J., Kendall, W.S. and Lieshout, M.N.M. van, Quermass-interaction processes, Research Report 293, Department of Statistics, University of Warwick (1996).

[2] Baddeley, A.J. and Lieshout, M.N.M. van, Area-interaction point processes, Ann. Inst. Statist. Math. **47** (1995), 601–619.

[3] Baddeley, A.J. and Møller, J., Nearest-neighbour Markov point processes and random sets, Internat. Statist. Rev. **57** (1989), 89–121.

[4] Besag, J., Some methods of statistical analysis for spatial data, Bull. Inst. Intern. Statist. **47**(2) (1977), 77–92.

[5] Daley, D.J., The definition of a multidimensional generalization of shot noise, J. Appl. Probab. **8** (1971), 128–135.

[6] Daley, D.J. and Vere-Jones, D., *An Introduction to the Theory of Point Processes*, Springer (1988), New York.

[7] Diggle, P.J., *Statistical Analysis of Spatial Point Patterns*, Academic-Press (1983), London.

[8] Diggle, P.J., Fiksel, T., Grabarnik, P., Ogata, Y., Stoyan, D. and Tanemura, M., On parameter estimation for pairwise interaction point processes, Internat. Statist. Rev. **62** (1994), 99–117.

[9] Fiksel, T., Estimation of interaction partials of Gibbsian point processes, Statistics **19** (1988), 77–86.

[10] Gates, D.J. and Westcott, M., Clustering estimates for spatial point distributions with unstable potentials, Ann. Inst. Statist. Math. **38** (1986), 123–135.

[11] Geyer, C.J. and Møller, J., Simulation procedures and likelihood inference for spatial point processes, Scand. J. Statist. **21** (1994), 359-373.

[12] O Häggström, O., Lieshout, M.N.M. and Møller, J., Characterisation results and Markov chain Monte Carlo algorithms including exact simulation for some spatial point processes, Technical Report 2040, Department of Mathematics, University of Aalborg (1996).

[13] Hammersley, J.M., Lewis, J.W.E. and Rowlinson, J.S., Relationships between the multinomial and Poisson models of stochastic processes, and between the canonical and grand canonical ensembles in statistical mechanics, with illustrations and Monte Carlo methods for the penetrable sphere model of liquid-vapour equilibrium, Sankhya: The Indian Journal of Statistics, series A **37** (1975), 457–491.

[14] Heinrich, L. and Schmidt, V., Normal convergence of multidimensional shot noise and rates of this convergence, Adv. in Appl. Probab. **17** (1985), 709–730.

[15] Jensen, J.L., Asymptotic normality of estimates in spatial point processes, Scand. J. Statist. **20** (1993), 97–109.

[16] Jensen, J.L. and Møller, J., Pseudolikelihood for exponential family models of spatial processes, Ann. Appl. Probab. **1** (1991), 445–461.

[17] Kendall, W.S., Perfect simulation for the area-interaction point process. To appear in *Proceedings of the Symposium on Probability towards the year 2000*. L. Accardi and C. Heyde (Eds), Springer-Verlag (1997), Berlin.

[18] Matheron, G., Random sets and integral geometry, Wiley (1975), New York.

[19] Møller, J., Markov chain Monte Carlo and spatial point processes, To appear in *Proceedings Seminaire Européen de Statistique, "Stochastic geometry, likelihood, and computation"*. O. Barndorff-Nielsen, W.S. Kendall and M.N.M. van Lieshout (Eds), Chapman and Hall (1997), London.

[20] Møller, J. and Waagepetersen, R., Markov connected component fields, Research Report 96-2009, Department of Mathematics and Computer Science, Aalborg University, 1996.

[21] Ogata, Y. and Tanemura, M., Likelihood estimation of soft-core interaction potentials for Gibbsian point patterns, Ann. Inst. Statist. Math. **41** (1989), 583–600.

[22] Pentinnen, A.K., Modelling interaction in spatial point patterns: Parameter estimation by the maximum likelihood method, Jyväskylä Studies in Computer Science, Economics and Statistics **7** (1984).

[23] Preston, C.J., *Random Fields*, Springer (1976), Berlin.

[24] Ripley, B.D., *Statistical Inference for Spatial Processes*, Cambridge Univ. Press (1988), Cambridge.

[25] Ripley, B.D. and Kelly, F.P., Markov point processes, J. London Math. Soc. **15** (1977), 188–192.

[26] Rowlinson, J.S., Penetrable sphere models of liquid-vapor equilibrium, Adv. Chem. Physics **41** (1980), 1–57.

[27] Ruelle, D., *Statistical Mechanics*, Wiley (1969), New York.

[28] Ruelle, D., Existence of a phase transition in a continuous classical system, Phys. Review Letters **27** (1971), 1040–1041.

[29] Särkkä, A., Pseudo-likelihood approach for pair potential estimation of Gibbs processes, Jyväskylä Studies in Computer Science, Economics and Statistics **22** (1993).

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

[31] Strauss, D.J., A model for clustering, Biometrika **63** (1975), 467–475.

[32] Takacs, R., Estimator for the pair-potential of a Gibbsian point process, Math. Operationsforsch. Statist. Ser. Statistik **17** (1986), 429–433.

[33] Wagner, D., Survey of measurable selection theorem, SIAM J. Control Optim. **15** (1979), 859–903.

[34] Widom, B. and Rowlinson, J.S., New model for the study of liquid-vapor phase transitions, J. Chemical Physics **52** (1970), 1670–1684.

Received: 10/9/1995
Revised: 12/6/1996
Accepted: 2/26/1997