Centrum voor Wiskunde en Informatica

REPORT*RAPPORT*

Shot-noise-weighted processes: a new family of spatial point processes

M.N.M. van Lieshout and I.S. Molchanov

Department of Operations Reasearch, Statistics, and System Theory

BS-R9527 1995

Report BS-R9527
ISSN 0924-0659

CWI
P.O. Box 94079
1090 GB Amsterdam
The Netherlands

Shot-Noise-Weighted Processes: A New Family of Spatial Point Processes

M.N.M. van Lieshout

Department of Statistics, University of Warwick
Coventry CV4 7AL, United Kingdom
stsab@csv.warwick.ac.uk

I.S. Molchanov

CWI
P.O. Box 94079, 1090 GB Amsterdam, The Netherlands
ilia@cwi.nl

Abstract

The paper suggests a new family of of spatial point processes distributions. They are defined by means of densities with respect to the Poisson point process within a bounded set. These densities are given in terms of a functional of the shot-noise process with a given influence function built for the Poisson point process. Stationary extension, properties and examples of such processes are discussed.

AMS Subject Classification (1991): 60D05, 60G55, 62M30.

Keywords & Phrases: clique, Gibbs process, interaction, Markov property, Poisson process, shot noise, spatial point process, Strauss process.

Notes: The second author was supported by the Netherlands Organization for Scientific Research (NWO). This paper has been submitted for publication.

1. Introduction

The Poisson point process is the simplest spatial point process which is characterized by its strong (even ultimate) independence property — points in disjoint regions are independent. Although the Poison model is very simple and tractable, it can often serve only as a first approximation when studying natural phenomena represented by points in the space. In practice, point patterns exhibit interactions between points, which, for instance, can yield clustered or repulsive behaviour.

Markov point processes were introduced in [25], whereas a similar concept of Gibbs distributions [21, 27] has been used in statistical physics for a longer time. Such processes are typically defined by their densities with respect to a Poisson process. The latter serves as a reference distribution when deriving new models. Typically the density is written in the form

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

where $\mathbf{x}$ is a finite set of unordered points (called configuration) in a bounded set $A \subset \mathbb{R}^d$ (in fact, in most cases $\mathbb{R}^d$ can be replaced by a locally compact second countable metric space $\mathbb{E}$), $n(\mathbf{x})$ is the number of points in $\mathbf{x}$, and $\psi(\mathbf{x})$ is a numerical functional of $\mathbf{x}$, which actually determines statistical properties of the point process.

One of the first examples of this kind is the Strauss model [30] which can be obtained by letting $\psi(\mathbf{x})$ to be equal to the number of pairs of points that have distance no more than a given number. This is an example of a so-called pairwise interaction process, since each configuration of points in this processes interacts only via pairs of points from this configuration. These models are suitable for inhibition, but probably not so for clustering, see [11] and [18]. This led to the definition of area-interaction processes in [4], a special (attractive) case had been considered previously as a model of liquid vapour equilibrium in chemical physics [33].

In case of area-interaction process, $\psi(\mathbf{x})$ is determined by measures (or areas in the simplest case) of some sets determined by $\mathbf{x}$, for instance given by the union of unit balls centred at the points which comprise $\mathbf{x}$. This idea can be generalized by using other functionals known from convex geometry [1] or by replacing measures of sets with integrals of some functions depending on $\mathbf{x}$. The latter idea leads to the concept of a shot-noise-weighted process developed in this paper. Already [4] 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 $f : [0, \infty] \mapsto (-\infty, \infty]$.

Apart from being interesting in their own right, Markov processes are useful as prior distributions in image interpretation tasks, such as object recognition, edge detection and feature extraction [3, 17, 16]. Maximum likelihood solutions tend to suffer from multiple response and the prior distribution serves to penalize scenes with too many almost identical objects, disconnected or crossing edges etcetera. Consider for example object recognition where the task is to determine if there are any objects of a given type in the scene and to locate them. It is natural to assume that objects interact if their intersection is non-empty and the more overlap there is, the stronger the interaction, suggesting an inhibitory shot-noise model. Usually, the posterior distribution also possesses a Markov property, enabling sampling and optimization by iterative procedures that recursively update the scene by simple operations such as addition or deletion.

The plan of the paper is as follows. Section 2 provides definition of the shot-noise-weighted point process in a compact set. Section 3 deals with existence of the corresponding distributions and the spatial Markov property. Section 4 is devoted to examples. Basic properties are considered in Section 5.

2. Set-up and definitions

We are concerned with point processes X on a locally compact complete separable metric space $\mathbb{E}$. 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

$\mathfrak{N}^f(\mathbb{E})$ of finite point configurations. A configuration is a finite set of points denoted by $\mathbf{x} = \{x_1, \dots, x_n\}$, and $n = n(\mathbf{x})$ denotes the number of points in $\mathbf{x}$.

Writing $\mathbf{x}_B$ for $\mathbf{x}$ restricted to $B \subseteq \mathbb{E}$, the σ-algebra $\mathcal{N}^f$ on $\mathfrak{N}^f(\mathbb{E})$ is the smallest σ-algebra with respect to which the evaluation $\mathbf{x} \mapsto n(\mathbf{x}_B)$ is measurable for every (bounded) $B \in \mathfrak{B}(\mathbb{E})$, where $\mathfrak{B}(\mathbb{E})$ is the family of Borel subsets of $\mathbb{E}$. For more details, consult [7].

Definition **2.1.** Let $\kappa : \mathbb{E} \times \mathbb{E} \mapsto \mathbb{R}^+ = [0, \infty)$ be a non-negative Borel function (called the *influence function*). With each configuration $\mathbf{x}$ associate a function $\xi_{\mathbf{x}} : \mathbb{E} \mapsto \mathbb{R}^+$,

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

If X is a finite point process, then $\xi_X(a)$, $a \in \mathbb{E}$, is said to be the *shot-noise process* generated by X with influence function κ [6, 13].

Very often $\mathbb{E} = \mathbb{R}^d$, and κ depends only on the difference between its arguments, so that $\kappa(a, x) = \kappa(a - x)$. Such influence functions in $\mathbb{R}^d$ are said to be *homogeneous.* For each point $e \in \mathbb{E}$ define $Z_\kappa(e) = \{a \in \mathbb{E} : \ \kappa(a, e) > 0\}$. If κ is homogeneous, then $Z_\kappa(e) = Z_\kappa + e$, where $Z_\kappa = \{a \in \mathbb{E} : \ \kappa(a, o) > 0\}$, and o is the origin.

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

Lemma 2.2. *The following hold:*

1. *$\xi_{\mathbf{x}}(a)$ is measurable as a function on $\mathfrak{N}^f(\mathbb{E}) \times \mathbb{E}$ with respect to the σ-algebra $\mathcal{N}^f(\mathbb{E}) \otimes \mathfrak{B}(\mathbb{E})$;*
2. *$0 \le \xi_{\mathbf{x}}(a) \le n(\mathbf{x})\kappa^*$ where $\kappa^* = \sup_{a,b \in \mathbb{E}} \kappa(a, b)$;*
3. *$\xi_{\mathbf{x} \cup \mathbf{y}}(a) = \xi_{\mathbf{x}}(a) + \xi_{\mathbf{y}}(a)$;*
4. *if $\kappa(b, a) = 0$, then $\xi_{\mathbf{x} \cup \{a\}}(b) = \xi_{\mathbf{x}}(b)$.*

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

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

Now the result follows from the fact that each set in the union is a Borel set in the space $\mathbb{R}^i \times \mathbb{R}$ due to the measurability of the influence function and the fact that $\xi_{\mathbf{x}}(a)$ is measurable as a sum of measurable functions $\kappa(a, x_i)$.

Other statements can be verified by straighforward computations. □

Example **2.3.** (Coverage function) With every point $a \in \mathbb{E}$ we associate a *zone of influence* $Z(a) \in \mathcal{K}(\mathbb{E})$, the compact subsets of $\mathbb{E}$. The function $Z : \mathbb{E} \mapsto \mathcal{K}$ is supposed to be weak measurable [32], so that $\{a : \ Z(a) \cap G \neq \emptyset\}$ is a Borel set for each open $G \subset \mathbb{E}$. Our generic example will be $\mathbb{E}$ a compact subset of $\mathbb{R}^d$ and $Z(a) = B(a, r)$, the ball with radius r centred at a. Then take $\kappa(b,a) = \mathbf{1}_{b \in Z(a)}$. Note that $Z_\kappa(a) = Z(a)$. A configuration $\mathbf{x} = \{x_1, \dots, x_n\}$ gives rise to the *coverage function* $c_\mathbf{x} : \mathbb{E} \mapsto \mathbb{R}$ defined as

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

that is, $c_\mathbf{x}$ counts the number of influence zones $Z(x_i)$ covering a. Then $c_\mathbf{x}(a) = 0$ for all $a \in \mathbb{E} \setminus U(\mathbf{x})$, where

$$U(\mathbf{x}) = \bigcup_{i=1}^{n(\mathbf{x})} Z(x_i).$$

The fourth statement of Lemma 2.2 now becomes $\xi_{\mathbf{x}\cup\{a\}}(b) = \xi_\mathbf{x}(b)$, if $b \notin Z(a)$. Moreover, if $Z(a) \cap U(\mathbf{x}) = \emptyset$, then $\xi_{\mathbf{x}\cup\{a\}}(b) = \xi_\mathbf{x}(b)$ for all $b \in U(\mathbf{x})$.

The simplest random point configuration is defined in the following way. Let μ be a finite measure on $\mathbb{E}$, and let N be a Poisson random variable with mean $\mu(\mathbb{E})$. Then N mutually independent (and independent of N) points with distribution

$$\mathbf{P}\{x_i \in B\} = \frac{\mu(B)}{\mu(\mathbb{E})}, \quad B \in \mathfrak{B}(\mathbb{E}),$$

give the random configuration $\mathbf{x} = \{x_1, \dots, x_N\}$ which is said to be the Poisson point process on $\mathbb{E}$ with intensity measure μ (cf. [7, 8, 29]). Its distribution on $\mathfrak{N}^f(\mathbb{E})$ is denoted by π_μ.

We shall restrict attention to processes that are absolutely continuous with respect to π_μ and define them by their density (or Radon-Nikodym derivative) with respect to π_μ.

Definition **2.4.** A shot-noise-weighted process with *potential function* f is a point process on $\mathbb{E}$ with density

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

with respect to π_μ. Here $\beta, \gamma > 0$ are model parameters, ν a finite Borel measure and $f : \mathbb{R} \mapsto \mathbb{R}$ a Borel function with $f(0) = 0$.

The model (2.2) is overparameterised as taking $f_1 = cf$ for some constant c is equivalent to changing γ to γ^c. If f is absolutely integrable, this ambiguity can be overcome by requiring that the integral of f is 1. The framework above does allow for multiple points. If this is undesirable, the reference measure μ must be diffuse.

3. Existence and Markov property

The existence of any point process given in terms of the density p with respect to the Poisson point process is ensured by Ruelle's stability condition [11, 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 $C > 0$. Then the density (or the corresponding energy) is called *stable*. In our case, Ruelle's condition requires a linear bound on $|\int_{\mathbb{E}} f(\xi_{\mathbf{x}}(a)) d\nu(a)|$ in terms of $n(\mathbf{x})$. For this, it suffices to require

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

for some $C > 0$.

For instance, the coverage function $c_{\mathbf{x}}(a)$ from Example 2.3 is bounded by $c_{\mathbf{x}}(a) \leq n(\mathbf{x})$, whence (3.2) is satisfied if

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

for some $C > 0$. Note that in the framework of Example 2.3, f can be represented as a sequence $f(n)_{n\geq 1}$ so that (3.3) must be checked for positive integers t only. We can ensure (3.2) in a more general setting by assuming (3.3) together with $\kappa^* = \sup_{a,x} \kappa(a,x) < \infty$. If $\kappa^* = \infty$ and f is bounded, i.e. $\sup_t f(t) < \infty$, then (3.2) is also valid.

Note that by Lemma 2.2 and Fubini's theorem, the function $p : \mathfrak{N}^f \mapsto \mathbb{R}$ is measurable. Summarizing we obtain the following result.

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

If f is bounded, then the distribution of the shot-noise-weighted process (Definition 2.4) 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 but also 'take it or leave it' type functions f with binary values. We come back to this later in Section 4.

To describe interaction behaviour, we need to define when points are "neighbours". As in [4], let $a \sim b$ if and only if

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

A process given by its density $p(\cdot)$ is a *Markov point process* [25] with respect to $\sim$ if, for all configurations $\mathbf{x}$,

(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 $N(\{u\}) \cap \mathbf{x} = \{x_i \in \mathbf{x} : u \sim x_i\}$ (the set of all points in $\mathbf{x}$ which are neighbours to u).

For generalizations see [2]. We will usually require that the influence zones are bounded. Otherwise far too many points are becoming neighbours.

Lemma 3.2. *The shot-noise-weighted process is Ripley-Kelly Markov with respect to the relation (3.4).*

Proof. Note that there are no zero-likelihood configurations (or forbidden states in statistical physics parlance), so the density is hereditary. To check **(b)**, write

$$\frac{p(\mathbf{x}\cup\{a\})}{p(\mathbf{x})} = \beta \exp\left\{-(\log\gamma)\left\{\int_{\mathbb{E}} f(\xi_{\mathbf{x}\cup\{a\}}(t))d\nu(t) - \int_{\mathbb{E}} f(\xi_{\mathbf{x}}(t))d\nu(t)\right\}\right\}.$$

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

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

and $\mathbf{z} = \mathbf{x}\setminus\mathbf{y}$. By Lemma 2.2, for all $t \notin Z_\kappa(a)$ we have that $\xi_{\mathbf{x}\cup\{a\}}(t) = \xi_{\mathbf{x}}(t)$. For all $t \in Z_\kappa(a)$,

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

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

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

completing the proof. □

One of the most important results for Markov point processes is the Hammersley–Clifford theorem [25] stating that $p(\cdot)$ can be factorized in a product of *clique interaction functions*:

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

A clique is any configuration $\mathbf{x}$ where all its members are neighbours ($\forall s,t\in\mathbf{x} : s\sim t$).

Theorem 3.3. *The interaction functions of a shot-noise-weighted process are*

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

$$\phi(\{a\}) = \beta\gamma^{-\int_{\mathbb{E}} f(\xi_{\{a\}}(t))d\nu(t)}, \tag{3.7}$$

$$\phi(\mathbf{x}) = \exp\left\{-(\log\gamma)\int_{\mathbb{E}}\sum_{\mathbf{y}\subseteq\mathbf{x}}(-1)^{|\mathbf{x}\setminus\mathbf{y}|}f(\xi_{\mathbf{y}}(t))d\nu(t)\right\} \quad \text{if } n(\mathbf{x})\geq 2. \tag{3.8}$$

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 $\phi(\{a\}) = p(\{a\})/\phi(\emptyset)$.

If $n(\mathbf{x}) = 2$, then for $a,b\in\mathbb{E}$ distinct,

$$\phi(\{a,b\}) = \frac{p(\{a,b\})}{\phi(\emptyset)\phi(\{a\})\phi(\{b\})} = \gamma^{-\int_{\mathbb{E}}\left(f(\xi_{\{a,b\}}(t)) - f(\xi_{\{a\}}(t)) - f(\xi_{\{b\}}(t))\right)d\nu(t)}.$$

Supposing the formula holds for configurations of $n(\mathbf{x}) = n \geq 2$ points, let $\mathbf{x}$ be such that $n(\mathbf{x}) = n$. Then

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

Note that "$\subset$" means proper subset relationship. Now

$$\begin{aligned}\sum_{\mathbf{y}\subset\mathbf{x}\cup\{a\}}\sum_{\mathbf{z}\subseteq\mathbf{y}}(-1)^{|\mathbf{y}\setminus\mathbf{z}|}f(\xi_{\mathbf{z}}(t)) &= \sum_{\mathbf{z}\subseteq\mathbf{y}\subset\mathbf{x}\cup\{a\}}(-1)^{|\mathbf{y}\setminus\mathbf{z}|}f(\xi_{\mathbf{z}}(t)) \\ &= \sum_{\mathbf{z}\subset\mathbf{x}\cup\{a\}}f(\xi_{\mathbf{z}}(t))\sum_{\mathbf{z}\subseteq\mathbf{y}\subset\mathbf{x}\cup\{a\}}(-1)^{|\mathbf{y}\setminus\mathbf{z}|} \\ &= -\sum_{\mathbf{z}\subset\mathbf{x}\cup\{a\}}(-1)^{|\mathbf{x}\cup\{a\}\setminus\mathbf{z}|}f(\xi_{\mathbf{z}}(t))\end{aligned}$$

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

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

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. It is easy to see that two such processes generated by the functions f 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. For this, take $s,t\in\mathbf{x}$: $s\not\sim t$ and write the integrand in the exponent in (3.8) as

$$\sum_{\mathbf{y}\subseteq\mathbf{x}\setminus\{s,t\}}(-1)^{|\mathbf{x}\setminus\mathbf{y}|}\left\{f(\xi_{\mathbf{y}}(a))+f(\xi_{\mathbf{y}\cup\{\mathbf{s},\mathbf{t}\}}(a))-f(\xi_{\mathbf{y}\cup\{\mathbf{s}\}}(a))-f(\xi_{\mathbf{y}\cup\{\mathbf{t}\}}(a))\right\}.$$

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

Direct simulation of the shot-noise-weighted process is difficult due to the dimension and the normalizing constant α that cannot be evaluated explicitly. On the other hand however, the conditional intensities (3.5) are "local" and very easy to compute. Therefore, it is possible to use a general techniques for spatial Markov point processes [2, 22, 23] that construct a Markov chain whose values are configurations of points and transitions are simple operations such as adding/deleting a point or replacing a point by another one. These transition probabilities are given by the likelihood ratio or conditional intensity (3.5). If this is done carefully, the Markov chain has the shot-noise-weighted process as its stationary distribution and converges to it from any initial configuration.

4. Examples and interpretations

Stationary Poisson process. If $\gamma=1$, regardless of the choice of potential function and influence function, the shot-noise-weighted model is a Poisson process with intensity measure $\beta\mu(\cdot)$. This is a model of spatial randomness and in particular, the interaction functions of order higher than 2 are identically 1. The lower order interaction functions are $\phi(\emptyset)=\alpha=e^{(1-\beta)\mu(\mathbb{E})}$, $\phi(\{a\})=\beta$.

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

$$\nu_\kappa(x) = \int_{\mathbb{E}} f(\kappa(a,x))d\nu(a)\,, \quad x \in \mathbb{E}.$$

Again, there are no interactions of orders higher than 1, $\phi(\emptyset) = \exp\left\{\int_{\mathbb{E}}(1 - \beta\gamma^{-c\nu_\kappa(a)})d\mu(a)\right\}$, but now $\phi(\{a\}) = \beta\gamma^{-c\nu_\kappa(a)}$ may be location dependent.

Area-interaction process. An example that does exhibit interactions between the members in a configuration is the area–interaction model [4]. This model is obtained by taking the coverage function introduced in Example 2.3 and potential function $f(n) = 1$ for all strictly positive integers, zero otherwise:

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

The model yields clustering for $\gamma > 1$, regularity for $\gamma < 1$. The special case $\gamma > 1$, $Z(a) = B(a, r)$, a ball of fixed radius centred at a and Lebesgue measure for ν is the penetrable sphere model introduced by [33] for liquid-vapour equilibrium. These models have interactions of arbitrary large order, except in the Poisson case $\gamma = 1$. For details, see [4, 12, 26, 33].

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: $f(n) = \mathbf{1}_{n=1}$, $n \geq 1$. Then

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

For $\gamma > 1$ the model tends to have smallish 1–covered regions, hence clustering. For $\gamma < 1$ the model tends to have largish 1–covered regions, hence inhibition. As the area-interaction process, this model has interactions of infinite order.

Truncated at k. Take the previous example but truncate at $k > 1$, so that $f(n) = \mathbf{1}_{n \leq k}$, $n \geq 1$. This allows some more overlap, namely of type up to k, and the density is given by

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

For $\gamma > 1$, the points more probably appear in 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 $\int f(\xi_{\mathbf{x}}(a))\nu(da)$ is the ν–measure of the set $U_2(\mathbf{x})$ of points in $\mathbb{E}$ 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.

Odd and even. Take again the coverage function example with one of 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. If $\gamma<1$, then the first example leads to the model with repulsion and largish 1–covered regions, while the second case gives model with mostly odd-numbered clusters.

Distance influence function. Let us consider an example not related to the coverage function. Set $\kappa(a,x)=\rho(a,x)$, i.e. the influence function $\kappa(a,x)$ is equal to the distance between a and x in $\mathbb{E}=\mathbb{R}^2$. Furthermore, take ν to be the Lebesgue measure and the potential function $f(t)=\mathbf{1}_{t\leq 1}$. Then

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

where $M_1(\mathbf{x})=\left\{a\in\mathbb{R}^2:\ \sum_{i=1}^{n(\mathbf{x})}\rho(a,x_i)\leq 1\right\}$. 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 a ellipse with foci x_1 and x_2.

5. Properties

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 $\mathbb{E}$, κ and ν are group-invariant (e.g., w.r.t. rotations), then X is distribution-invariant with respect to the same group.

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

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

where the sum is over all ordered partitions of $\mathbf{y}$. In the case of two independent shot-noise-weighted processes, (5.1) reads

$$p(\mathbf{y})=e^{-\mu(\mathbb{E})}\alpha^2\beta^{n(\mathbf{y})}\sum_{\varphi:\mathbf{y}\to\{1,2\}}\exp\left[-\log\gamma\int\Big(f(\xi_{\mathbf{y}_{\varphi^{-1}(1)}}(t))+f(\xi_{\mathbf{y}_{\varphi^{-1}(2)}}(t))\Big)\,d\nu(t)\right].$$

In general, $X_1 \cup X_2$ is not a shot-noise-weighted process. If we assume moreover that f is a linear function this reduces to the familiar superposition property of Poisson point processes [8, 29].

Now consider independent thinning. Under this operation each point in a realization of a point process is retained independent of every other point with probability p. Then the Janossy densities of the thinned shot-noise-weighted process (cf. [7, p. 122]) are given by

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

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

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

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

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

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

If f is a linear function (Poisson case), then the thinned process is a Poisson with intensity measure $p\beta\gamma^{-\nu_\kappa(\cdot)}\mu(\cdot)$, which is a familiar property of the Poisson point process, see [29].

Let us study the convergence of the shot-noise-weighted point process as $\gamma \to 0, \infty$.

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

1. *Assume that f is bounded and let*

$$\mathsf{H} = \left\{ \mathbf{x} : \int_{\mathbb{E}} f(\xi_{\mathbf{x}}(a)) d\nu(a) = \max_{\mathbf{x}} \int_{\mathbb{E}} f(\xi_{\mathbf{x}}(a)) d\nu(a) \right\} .$$

   If $\gamma \to 0$ for fixed β, then $P_{\beta,\gamma}$ 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 $\beta\gamma^{-\nu_\kappa(a)} \to \zeta(a) \in (0, \infty)$, $a \in \mathbb{E}$. 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 $\zeta(\cdot)\mu(da)$ restricted to the set of configurations*

$$\mathsf{HC} = \left\{ \mathbf{x} : \int_{\mathbb{E}} f(\sum_{i=1}^{n(\mathbf{x})} \kappa(a, x_i)) d\nu(a) = \sum_{i=1}^{n(\mathbf{x})} \int_{\mathbb{E}} f(\kappa(a, x_i)) d\nu(a) \right\} .$$

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 realizations of hard-core point processes [29].

The intensity and moment measures of shot-noise-weighted process X can be found using the expression for the density (2.2). Suppose that μ is diffuse, i.e. X has no multiple points. The intensity of X is given by $\lambda(dx) = \alpha\beta\gamma^{-\nu_\kappa(x)}\mu(dx)$. This is the probability to have exactly one point in the neighbourhood of x. Furthermore, the second-order cumulant factorial moment measure of X is given by
$$\alpha^{-1}\gamma^{\int(f(\kappa(a,u)+\kappa(a,v)) - f(\kappa(a,u)) - f(\kappa(a,v)))d\nu(a)}\mu(du)\mu(dv)\,.$$

Below we assume that $\mathbb{E} = \mathbb{R}^d$ and the influence function is homogeneous with bounded Z_κ. Existence of a stationary extension is an important problem when defining distributions given in terms of the density with respect to the Poisson process. This means that the model can be considered as the restriction to a bounded sampling window of a stationary point process on the whole of $\mathbb{R}^d$. By using methods of Preston [22] in the same way as it has been done in [4] (with evident changes) one can prove that this is true as soon as (3.2) is satisfied. Note that this result does not exclude the possibility that the Gibbs state is not unique, i.e. there may be “phase transition” [21, p. 46].

Jensen [14] proved a central limit theorem for Gibbs point processes in view of applications to the maximum pseudo likelihood estimates. It is possible to show that, if

f is bounded and the influence function κ is homogeneous with compact Z_κ, then the corresponding shot-noise-weighted process satisfies conditions of Theorem 2.2 [14]. In particular, this yields the central limit theorem for *additive* functionals like the number of points. For shot-noise-weighted process the sufficient statistics are $n(\mathbf{x})$ as usual, and

$$\int_{\mathbb{E}} f(\xi_{\mathbf{x}}(a))\, d\nu(a)\,,$$

which is the integrated shot-noise [13]. Unfortunately, the latter (for non-linear f which is only interesting) becomes a non-additive functional of the observation window.

Statistical estimation for shot-noise-weighted processes can be performed using the usual techniques as for area-interaction point processes. It is possible to estimate β and γ using three basic methods:

- maximum likelihood techniques using Markov chain Monte Carlo simulations or stochastic approximation techniques to find out the unknown normalization parameter α [9, 19, 20];
- the Takacs-Fiksel estimation method, see [10, 31] and [24, p. 54–55];
- the pseudo likelihood equations [5, 15], which have exactly 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 [4, 9, 28].

The estimation of f and κ is a difficult non-parametric statistical problem.

Acknowledgements

This research 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. The authors are also grateful to CWI and Warwick University for hospitality.

References

[1] A.J. Baddeley, W.S. Kendall, and M.N.M. van Lieshout. Quermass-interaction processes. Unpublished manuscript, 1995.

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

[3] A.J. Baddeley and M.N.M. van Lieshout. Stochastic geometry in high-level vision. In K. V. Mardia and G. K. Kanji, editors, *Statistics and images*, volume 1 of *Advances in Applied Statistics*, pages 231–256, Carfax, 1993. Abingdon.

[4] A.J. Baddeley and M.N.M. van Lieshout. Area-interaction point processes. *Ann. Inst. Statist. Math.*, 1995. To appear.

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

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

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

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

[9] P.J. Diggle, T. Fiksel, P. Grabarnik, Y. Ogata, D. Stoyan, and M. Tanemura. On parameter estimation for pairwise interaction point processes. *Internat. Statist. Rev.*, 62:99–117, 1993.

[10] T. Fiksel. Estimation of interaction partials of Gibbsian point processes. *statistics*, 19:77–86, 1988.

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

[12] J.M. Hammersley, J.W.E. Lewis, and J.S. Rowlinson. 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*, 37:457–491, 1975.

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

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

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

[16] M.N.M. van Lieshout and A.J. Baddeley. Markov chain Monte Carlo methods for clustering of image features. In *Proceedings of the fifth international conference on image processing and its applications*, volume 410 of *IEE Conference Publication*, pages 241–245, London, 1995. IEE.

[17] R. Molina and B.D. Ripley. Using spatial models as priors in astronomical image analysis. *J. Appl. Statist.*, 16:193–206, 1989.

[18] J. Møller. Markov chain Monte Carlo and spatial point processes. Technical Report 293, 1994.

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

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

[21] C.J. Preston. *Random fields*. Springer, Berlin, 1976.

[22] C.J. Preston. Spatial birth-and-death processes. *Bull. Inst. Intern. Statist.*, 46:371–391, 1977.

[23] B.D. Ripley. Modelling spatial patterns (with discussion). *J. Roy. Statist. Soc. Ser. B*, 39:172–212, 1977.

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

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

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

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

[28] A. Särkkä. On parameter estimation of Gibbs point processes through the pseudolikelihood method. Technical Report 4, Department of Statistics, University of Jyväskylä, 1989.

[29] D. Stoyan, W.S. Kendall, and J. Mecke. *Stochastic Geometry and Its Applications*. Wiley, Chichester, 1987.

[30] D.J. Strauss. A model for clustering. *Biometrika*, 63:467–475, 1975.

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

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

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