Adv. Appl. Prob. (SGSA) **35**, 319–336 (2003)
Printed in Northern Ireland

INHOMOGENEOUS SPATIAL POINT PROCESSES BY LOCATION-DEPENDENT SCALING

UTE HAHN * AND
EVA B. VEDEL JENSEN,** *University of Aarhus*
MARIE-COLETTE VAN LIESHOUT,*** *CWI, Amsterdam*
LINDA STOUGAARD NIELSEN,** *University of Aarhus*

Abstract

A new class of models for inhomogeneous spatial point processes is introduced. These locally scaled point processes are modifications of homogeneous template point processes, having the property that regions with different intensities differ only by a scale factor. This is achieved by replacing volume measures used in the density with locally scaled analogues defined by a location-dependent scaling function. The new approach is particularly appealing for modelling inhomogeneous Markov point processes. Distance-interaction and shot noise weighted Markov point processes are discussed in detail. It is shown that the locally scaled versions are again Markov and that locally the Papangelou conditional intensity of the new process behaves like that of a global scaling of the homogeneous process. Approximations are suggested that simplify calculation of the density, for example, in simulation. For sequential point processes, an alternative and simpler definition of local scaling is proposed.

Keywords: Point process; inhomogeneity; local scaling; sequential point process

AMS 2000 Subject Classification: Primary 60G55
Secondary 62M30; 60D05

1. Introduction

Point patterns with nonhomogeneous intensity are observed quite frequently in nature and technology. For example, the number of plants per unit area in a natural environment depends on environmental conditions and topology and therefore maps showing plant locations on larger regions with changing conditions usually look inhomogeneous. In plant and animal tissue, cell size and, correspondingly, cell number often depend on the distance to the boundary of an organ. Many modern materials are designed with structural inhomogeneity, imitating natural structures in order to improve functional properties. An example is the bronze sinter filter shown in Figure 1. The data were analysed in [6]. The filter consists of almost spherical bronze particles with diameters that decrease along an axis which marks the filtering direction. Since the particles are densely packed, the number of particles per unit volume increases as the diameters

Received 29 January 2002; revision received 30 September 2002.
* Current address: Department of Mathematics, Augsburg University, D-86135 Augsburg, Germany.
Email address: hahn@math.uni-augsburg.de
** Postal address: Laboratory for Computational Stochastics, University of Aarhus, Ny Munkegade, DK-8000 Aarhus C, Denmark.
*** Postal address: CWI, PO Box 94079, 1090 GB Amsterdam, The Netherlands.

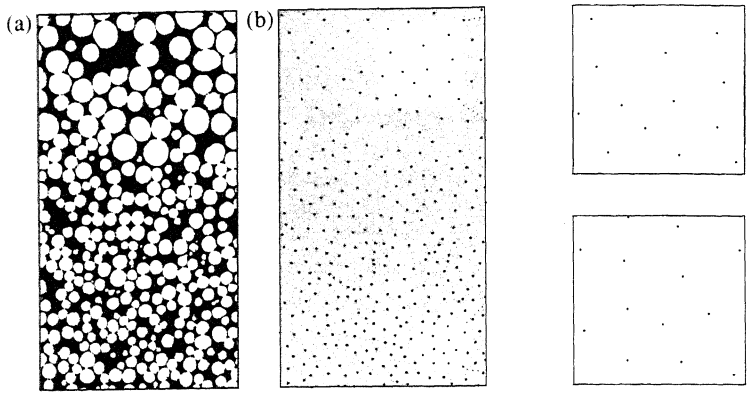

FIGURE 1: (a) Section of a bronze sinter filter with a gradient in particle size and number. (b) Centres of the particle profiles. Two enlargements from the top and bottom, containing about the same number of points, show similar geometry.

decrease. This is also observable on sections parallel to the directions of inhomogeneity: the centres of the particle section profiles form an inhomogeneous point pattern.

While it is easy to model inhomogeneous point patterns with independently positioned points by inhomogeneous Poisson point processes, situations as shown in Figure 1 require more sophisticated approaches. This pattern is characterized by repulsive interaction between the points due to the fact that it results from a packing of spheres. The packing is of similar volume fraction and similar geometry in regions with larger and with smaller sphere diameters. Therefore, regions with large sphere diameters look like scaled versions of regions with small diameters and vice versa. A similar effect can often be observed in nature, e.g. in plant communities where number density is governed by environmental conditions. For example, desert plants tend to form regular patterns with varying scale, such that distances between plants are smaller in densely covered regions. Such point patterns also appear homogeneous up to a local scale factor.

In recent years, various models have been suggested for inhomogeneous point processes with interaction. Since Markov point processes are very useful for modelling interaction in homogeneous point patterns, it is natural that they are used as starting points for inhomogeneous models. The survey by Jensen and Nielsen [9] discusses three ways of introducing inhomogeneity into a Markov model. As will be explained in more detail in Section 2, homogeneous Markov point processes are defined by a density with respect to the unit rate Poisson point process. A straightforward idea is therefore to define an inhomogeneous process by the same density (up to a constant factor) but with respect to an inhomogeneous Poisson point process [14], [12]. Inhomogeneity can also be obtained by location-dependent thinning [2], or by transformation of a homogeneous Markov point process [8].

In these three cases, the local geometry of the point pattern changes with intensity. This is illustrated in Figure 2, which shows realizations of inhomogeneous hard-core point processes obtained by the three methods. In order to obtain patterns that appear homogeneous up to a scale factor, range and strength of interaction have to be adapted to intensity. However, this is not accomplished by the first approach where the interaction between points does not depend on their locations; see Figure 2(a). Thinning, on the other hand, in general destroys the interaction structure. This leads to a Poisson-like appearance of sparse regions; see Figure 2(b). Finally,

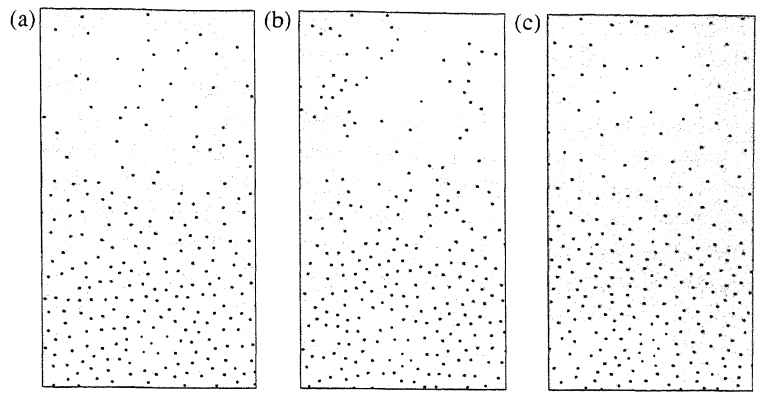

FIGURE 2: Inhomogeneous hard-core point patterns obtained by (a) defining the density with respect to an inhomogeneous Poisson point process, (b) inhomogeneous independent thinning, (c) transformation of coordinates. Note that dense and sparse regions differ in geometry. The parameters were chosen such that the processes have similar intensity to the example of Figure 1.

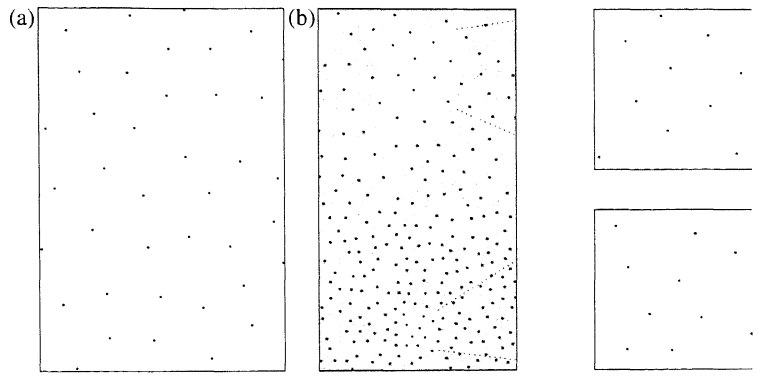

FIGURE 3: (a) Homogeneous template hard-core process and (b) inhomogeneous process obtained by local scaling. Enlargements from dense and sparse regions of the inhomogeneous pattern look similar to the template pattern.

transformation of coordinates not only introduces inhomogeneity but also local anisotropy, as shown in Figure 2(c). Therefore, these three approaches are not suitable for modelling situations as given in Figure 1.

In the present paper, we propose alternative inhomogeneous point process models that aim to preserve local geometry. As in the three approaches discussed above, the inhomogeneous model is obtained by modifying a homogeneous 'template' process that yields the interaction. The idea is that inhomogeneity is obtained by scaling the template process with a location-dependent scaling factor. A large scaling factor thereby results in low intensity and large interaction distances, whereas a small scaling factor yields high intensity and small interaction range. In regions with constant scaling factor, the point process should locally behave like a scaled version of the template; see Figure 3.

The method and results presented in this paper are applicable to homogeneous template processes that are given by a density with respect to a homogeneous Poisson point process;

however, the main emphasis will be on Markov point processes. The definition of Markov point processes and other prerequisites are recalled in Section 2.

Calculating the density function of a point process for a given point pattern usually implies evaluating distances, areas, etc. The local scaling model proposed in Section 3 changes the way such quantities are measured according to a location-dependent scaling function.

Sections 4 and 5 give a closer look at the important classes of distance-interaction and shot noise weighted processes. In particular, we show that locally scaled Markov point processes are again Markov, now with respect to a location-dependent relation. Useful approximations of local scaling simplifying calculations, for example, in simulation are presented in Section 6.

For the class of sequential point processes, another approach to obtaining local scaling by means of conditional intensities is suggested in Section 7. The paper concludes with a critical discussion.

2. Prerequisites

Let $\mathbb{B}^k$ denote the set of all full-dimensional bounded subsets of $\mathbb{R}^k$ and write $\mathcal{B}_k$ for the Borel σ-algebra on $\mathbb{R}^k$. We consider finite point processes X on sets $\mathcal{X} \in \mathbb{B}^k$. A point process X on $\mathcal{X}$ is a random variable taking values in $\Omega_{\mathcal{X}}$, the set of all finite subsets $\boldsymbol{x} = \{x_1, \ldots, x_n\}$ of $\mathcal{X}$, equipped with the smallest σ-algebra for which the number of points placed in a Borel set $B \subseteq \mathcal{X}$ is a random variable.

We will concentrate on point processes X that have a density f_X with respect to the restriction of the unit rate Poisson point process Π to $\mathcal{X}$. A point process X on $\mathcal{X}$ is called *homogeneous* if f_X is the restriction to $\mathcal{X}$ of a translation-invariant function defined on all finite subsets of $\mathbb{R}^k$; see [11].

Markov point processes in the sense of Ripley and Kelly [13] are particularly useful for modelling point patterns with interaction. They are defined with respect to a symmetric and reflexive relation $\sim$ on $\mathcal{X}$. Two points $x_1, x_2 \in \mathcal{X}$ are said to be neighbours if $x_1 \sim x_2$, and a finite subset $\boldsymbol{x} \subset \mathcal{X}$ is called a clique if all points in $\boldsymbol{x}$ are neighbours. Note that in graph theory a clique is defined in a different way, cf. for example [7].

For a Markov point process, the hereditary condition holds, i.e. $f_X(\boldsymbol{x}) > 0$ implies that $f_X(\boldsymbol{y}) > 0$ for all $\boldsymbol{y} \subseteq \boldsymbol{x}$. Furthermore, the Papangelou conditional intensity

$$\lambda_X(x \mid \boldsymbol{x}) = \begin{cases} \dfrac{f_X(\boldsymbol{x} \cup \{x\})}{f_X(\boldsymbol{x})}, & f_X(\boldsymbol{x}) > 0, \\ 0, & \text{otherwise}, \end{cases}$$

for $x \notin \boldsymbol{x}$, depends only on those points in $\boldsymbol{x}$ which are neighbours of x. If we let $\mathrm{d}x$ be an infinitesimal region around x and $\nu^k(\mathrm{d}x)$ the k-dimensional volume (Lebesgue measure) of $\mathrm{d}x$, then $\lambda(x \mid \boldsymbol{x})\nu^k(\mathrm{d}x)$ can be interpreted as the conditional probability of finding a point from the process in $\mathrm{d}x$ given that the configuration elsewhere is $\boldsymbol{x}$; see e.g. [16].

By the Hammersley–Clifford theorem (see [13]), a Ripley–Kelly Markov point process X has density with respect to the unit rate Poisson point process of the form

$$f_X(\boldsymbol{x}) = \prod_{\boldsymbol{y} \subseteq \boldsymbol{x}} \varphi(\boldsymbol{y}), \qquad \boldsymbol{x} \in \Omega_{\mathcal{X}},$$

where φ is an interaction function, i.e. $\varphi(\boldsymbol{y}) = 1$ when the set $\boldsymbol{y}$ is not a clique. We will always assume that the interaction function φ is defined on *all* finite subsets of $\mathbb{R}^k$. A Markov point process X is thereby homogeneous if φ is translation invariant (for a proof see [10, p. 29]).

Before defining local scaling of point processes, let us consider global scaling with a constant factor $c > 0$. As a transformation of coordinates, global scaling maps a point process X on $\mathcal{X}$ to a process $X_c = cX$ on the set $c\mathcal{X} = \{x : x/c \in \mathcal{X}\}$; see also [8].

The unit rate Poisson point process Π on $\mathcal{X}$ with intensity measure ν^k is transformed into a Poisson point process Π_c on $c\mathcal{X}$ with intensity measure $c^{-k}\nu^k$.

Let f_X be the density of the original process with respect to Π. Then the scaled process X_c has density $f_{X_c}^{(c)}$ with respect to Π_c,

$$f_{X_c}^{(c)}(\boldsymbol{x}) = f_X\left(\frac{\boldsymbol{x}}{c}\right).$$

(The superscript (c) in $f_{X_c}^{(c)}$ is used to indicate that the density is with respect to Π_c instead of Π.) The conditional intensity associated with $f_{X_c}^{(c)}$ is

$$\lambda_{X_c}^{(c)}(x \mid \boldsymbol{x}) = \lambda_X\left(\frac{x}{c} \mid \frac{\boldsymbol{x}}{c}\right). \tag{1}$$

The density of X_c with respect to Π is

$$f_{X_c}(\boldsymbol{x}) = \mathrm{e}^{(1-c^{-k})\nu^k(c\mathcal{X})} c^{-kn(\boldsymbol{x})} f_{X_c}^{(c)}(\boldsymbol{x}), \qquad \boldsymbol{x} \in \Omega_{c\mathcal{X}},$$

where $n(\boldsymbol{x})$ is the number of points in $\boldsymbol{x}$.

3. Local scaling of homogeneous point processes

In this section, we give a general definition of a locally scaled version of a homogeneous template process.

The concept of scale invariance plays a crucial role in the definition. This concept relates to global scaling with a constant factor $c > 0$. Note that, under scaling with a factor c, a measure μ on $(\mathbb{R}^k, \mathcal{B}_k)$ is transformed into μ_c where $\mu_c(A) = \mu(c^{-1}A)$ for $A \in \mathcal{B}_k$.

Definition 1. Let $g(\boldsymbol{x}; \mu^*)$ be a real-valued measurable function defined on $\Omega_{\mathbb{R}^k}$, depending on a set $\mu^* = (\mu^1, \ldots, \mu^m)$ of measures on $(\mathbb{R}^k, \mathcal{B}_k)$. The function g is called *scale invariant* if, for all $\boldsymbol{x} \in \Omega_{\mathbb{R}^k}$ and all $c > 0$,

$$g(c\boldsymbol{x}; \mu_c^*) = g(\boldsymbol{x}; \mu^*),$$

where $\mu_c^* = (\mu_c^1, \ldots, \mu_c^m)$.

The classical homogeneous point process models that appear in the spatial statistics literature have a density which is the restriction to $\Omega_{\mathcal{X}}$ of a scale-invariant function $g(\cdot\,; \mu^*)$, where $\mu^* = \nu^* = (\nu^0, \ldots, \nu^k)$ is the set of d-dimensional volume (Hausdorff) measures ν^d in $\mathbb{R}^k$, $d = 0, 1, \ldots, k$. A comprehensive set of examples will be given in the sections to follow. Recall that ν^0 is the counting measure, thus $\nu^0(\boldsymbol{x}) = n(\boldsymbol{x})$, and ν^1 is the length measure in $\mathbb{R}^k$. Note also that $\nu_c^d(A) = \nu^d(c^{-1}A) = c^{-d}\nu^d(A)$ for $A \in \mathcal{B}_k$.

Under local scaling, the constant scaling factor c is replaced by a nonconstant location-dependent scaling function $c : \mathbb{R}^k \to \mathbb{R}_+$. The globally scaled measures ν_c^d can easily be extended to this case.

Definition 2. Let c be a positive Borel measurable function on $\mathbb{R}^k$. Then the *locally scaled d-dimensional volume measure* ν_c^d is defined by

$$\nu_c^d(A) = \int_A c(u)^{-d} \nu^d(\mathrm{d}u)$$

for all $A \in \mathcal{B}_k$.

In the following, we will assume that c is bounded from below and above, i.e. there exist $\underline{c}, \bar{c}$ such that $0 < \underline{c} \leq c(u) \leq \bar{c} < \infty$ for $u \in \mathbb{R}^k$. This assumption implies in particular that $\nu_c^d(A) < \infty$ whenever $\nu^d(A) < \infty$.

We can now present the definition of locally scaled point processes.

Definition 3. Let X be a homogeneous point process on $\mathcal{X}$, with density f_X with respect to Π of the form

$$f_X(\boldsymbol{x}) \propto g(\boldsymbol{x}; \nu^*), \qquad \boldsymbol{x} \in \Omega_{\mathcal{X}},$$

where g is scale invariant. Let c be a positive, Borel measurable function in $\mathbb{R}^k$ and let Π_c be the Poisson point process with the locally scaled volume measure ν_c^k as intensity measure. Let $\mathcal{X}' \in \mathbb{B}^k$ be arbitrary and suppose that $g(\cdot\,; \nu_c^*)$ is integrable on $\Omega_{\mathcal{X}'}$ with respect to Π_c. A *locally scaled point process* X_c on $\mathcal{X}'$ with template X is defined by the following density with respect to Π_c,

$$f_{X_c}^{(c)}(\boldsymbol{x}) \propto g(\boldsymbol{x}; \nu_c^*), \qquad \boldsymbol{x} \in \Omega_{\mathcal{X}'},$$

where ν_c^* is the set of locally scaled volume measures.

If $c : \mathbb{R}^k \to \mathbb{R}_+$ is constant, $c(u) \equiv c$, say, then the density with respect to Π_c of the scaled process on $\mathcal{X}' = c\mathcal{X}$ becomes

$$f_{X_c}^{(c)}(\boldsymbol{x}) \propto g(\boldsymbol{x}; \nu_c^*) = g(c^{-1}\boldsymbol{x}; \nu^*), \qquad \boldsymbol{x} \in \Omega_{\mathcal{X}'}.$$

Local scaling with a constant scaling function is thereby equivalent to global scaling. In the general case where c is not constant, local scaling does not necessarily correspond to a mapping. Therefore, there is no natural choice of $\mathcal{X}'$ which is related to $\mathcal{X}$, and the set $\mathcal{X}'$ can be arbitrary. In particular, we may choose $\mathcal{X}' = \mathcal{X}$. Note that the density of the locally scaled process X_c with respect to the unit rate Poisson point process Π is

$$f_{X_c}(\boldsymbol{x}) \propto f_{X_c}^{(c)}(\boldsymbol{x}) \prod_{x \in \boldsymbol{x}} c(x)^{-k}.$$

Locally scaled Markov point processes are again Markov, but now with respect to a relation $\sim_c$ which in general is different from the template relation $\sim$. Local scaling of two general Markov model classes, distance-interaction processes and shot noise processes, is discussed in detail in Sections 4 and 5. For these classes, conditions on the scaling function which ensure integrability of $g(\cdot, \nu_c^*)$ will be given, and it will be shown that the Papangelou conditional intensity of the locally scaled process,

$$\lambda_{X_c}^{(c)}(x \mid \boldsymbol{x}) = \begin{cases} \dfrac{f_{X_c}^{(c)}(\boldsymbol{x} \cup \{x\})}{f_{X_c}^{(c)}(\boldsymbol{x})}, & f_{X_c}^{(c)}(\boldsymbol{x}) > 0, \\ 0, & \text{otherwise}, \end{cases}$$

satisfies a local analogue of (1),

$$\lambda_{X_c}^{(c)}(x \mid \boldsymbol{x}) = \lambda_X\left(\frac{x}{c(x)} \;\middle|\; \frac{\boldsymbol{x}}{c(x)}\right). \tag{2}$$

if c is constant in a $\sim_c$-neighbourhood of x. The locally scaled processes thereby behave locally like a scaled version of the template process and the local geometry is preserved. In particular, if the template is locally isotropic in the sense that $\lambda_X(x \mid \cdot)$ is invariant under rotations around x, then so is the locally scaled process. Indeed, let R_x be a rotation around x. Then

$$\begin{aligned}\lambda_{X_c}^{(c)}(x \mid R_x\boldsymbol{x}) &= \lambda_X\left(\frac{x}{c(x)} \;\middle|\; \frac{R_x\boldsymbol{x}}{c(x)}\right) = \lambda_X\left(\frac{x}{c(x)} \;\middle|\; R_{x/c(x)}\left(\frac{\boldsymbol{x}}{c(x)}\right)\right)\\ &= \lambda_X\left(\frac{x}{c(x)} \;\middle|\; \frac{\boldsymbol{x}}{c(x)}\right) = \lambda_{X_c}^{(c)}(x \mid \boldsymbol{x}).\end{aligned} \tag{3}$$

Remark 1. The representation $g(\boldsymbol{x}; v^*)$ of a given function $f : \Omega_{\mathcal{X}} \to \mathbb{R}$ is, in general, not unique. Therefore, a given homogeneous template process does not correspond to a unique locally scaled process. For instance, the simple neighbour relation $u \sim v \iff v^1([u, v]) < r$ indicating that two points u and v are less than a distance r apart could also be written as $u \sim v \iff v^2(b(u, r/2) \cap b(v, r/2)) > 0$, which means that the balls of radius $r/2$ around the points have nonempty intersection. In order to find a natural extension of a homogeneous template to a locally scaled process, we recommend that the simplest possible representation of the density function be used.

4. Distance-interaction processes

In many Markov point process models, higher-order interactions are functions of pairwise distances. The densities are of the form

$$f_X(\boldsymbol{x}) = \prod_{y \subseteq \boldsymbol{x}} \varphi(D(y)), \tag{4}$$

where $D(y) = y$ if $n(y) < 2$ and, for $n(y) \geq 2$,

$$D(y) = \{v^1([u, v]) : \{u, v\} \subseteq y,\ u \neq v\}$$

denotes the set of all pairwise distances of points in y, with $[u, v]$ being the line segment connecting the points u and v. Such processes shall be called *distance-interaction processes*. This class includes pairwise interaction processes, such as the hard-core process and the Strauss process, as well as higher-order processes, e.g. the triplets process [5].

Assume that $\varphi(\{x\}) = \beta$ and that $\varphi(D(y)) = 1$ for $n(y) \geq 2$ unless $v^1([u, v]) \leq r$ for all $\{u, v\} \subseteq y$. The process X is thereby homogeneous and Markov with respect to the relation $u \sim v \iff v^1([u, v]) \leq r$.

According to (4), $f_X(\boldsymbol{x}) \propto g(\boldsymbol{x}; v^*)$ for $\boldsymbol{x} \in \Omega_{\mathcal{X}}$, where

$$g(\boldsymbol{x}; v^*) = \beta^{v^0(\boldsymbol{x})} \prod_{y \subseteq_2 \boldsymbol{x}} \varphi(\{v^1([u, v]) : \{u, v\} \subseteq y,\ u \neq v\})$$

and $y \subseteq_2 \boldsymbol{x}$ is short for $\{y \subseteq \boldsymbol{x} : v^0(y) \geq 2\}$. The function g is clearly scale invariant.

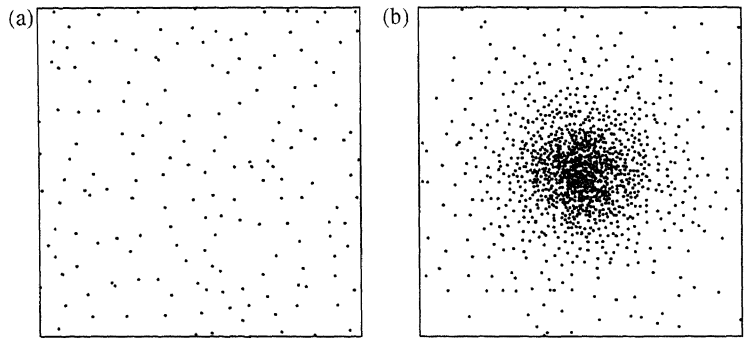

FIGURE 4: (a) Homogeneous template Strauss process X on $\mathcal{X} = [-1, 1]^2$ with parameters $\beta = 200, \gamma = 0.1, r = 0.1$. (b) Inhomogeneous Strauss process X_c on $\mathcal{X}' = [-1, 1]^2$ obtained by local scaling of X with $c(u) = 0.1 + \|u\|^2$.

If $\varphi(D(\boldsymbol{y})) \leq 1$ for $n(\boldsymbol{y}) \geq 2$, then X is repulsive since each clique $\boldsymbol{y} \subseteq \boldsymbol{x}$ contributes a penalty $\varphi(D(\boldsymbol{y}))$ to the density. In this case, $g(\boldsymbol{x}; \nu_c^*) \leq \beta^{n(\boldsymbol{x})}$ for any scaling function $c : \mathbb{R}^k \to \mathbb{R}_+$. Hence, $g(\cdot\,; \nu_c^*)$ is integrable on $\Omega_{\mathcal{X}'}$ for any $\mathcal{X}' \in \mathbb{B}^k$ and, therefore, locally scaled versions of such processes do exist. Otherwise, integrability has to be proved case by case and may require certain restrictions on the scaling function c. The locally scaled process has a density of the form

$$f_{X_c}^{(c)}(\boldsymbol{x}) \propto \beta^{n(\boldsymbol{x})} \prod_{\boldsymbol{y} \subseteq_2 \boldsymbol{x}} \varphi(D_c(\boldsymbol{y})), \qquad \boldsymbol{x} \in \Omega_{\mathcal{X}'},$$

where $D_c(\boldsymbol{y}) = \{\nu_c^1([u, v]) : \{u, v\} \subseteq \boldsymbol{y}, u \neq v\}$.

Example 1. (*Strauss process.*) A Strauss process X on $\mathcal{X} \subseteq \mathbb{R}^k$ with intensity parameter $\beta > 0$, interaction parameter $\gamma \in [0, 1]$ and interaction distance r has density

$$f_X(\boldsymbol{x}) \propto \beta^{n(\boldsymbol{x})} \gamma^{s(\boldsymbol{x})}, \qquad s(\boldsymbol{x}) = \sum_{\{u,v\} \subseteq \boldsymbol{x}}^{\neq} \mathbf{1}(\nu^1([u, v]) \leq r), \qquad \boldsymbol{x} \in \Omega_{\mathcal{X}},$$

where $s(\boldsymbol{x})$ is the number of r-close pairs in $\boldsymbol{x}$ [15]. (The superscript $\neq$ in the summation indicates that u and v are different.) For $\gamma = 0$ we obtain the hard-core process, for $\gamma = 1$ the Poisson point process with intensity β. The locally scaled Strauss process has density

$$f_{X_c}^{(c)}(\boldsymbol{x}) \propto \beta^{n(\boldsymbol{x})} \gamma^{s_c(\boldsymbol{x})}, \qquad s_c(\boldsymbol{x}) = \sum_{\{u,v\} \subseteq \boldsymbol{x}}^{\neq} \mathbf{1}(\nu_c^1([u, v]) \leq r), \qquad \boldsymbol{x} \in \Omega_{\mathcal{X}'}.$$

Figure 4 shows a realization.

The locally scaled process is Markov with respect to the relation $u \sim v \iff \nu_c^1([u, v]) \leq r$. Thus, the shape of the neighbourhood

$$\partial_c(x) = \{y : \nu_c^1([x, y]) \leq r\} = b_c(x, r)$$

in the locally scaled process depends on the scaling function c. It is not necessarily convex but always star shaped; see Figure 5. A neighbourhood $\partial_c(x)$ is called star shaped if it is star shaped with respect to x, which means that, if $u \in \partial_c(x)$, then $[x, u] \subseteq \partial_c(x)$.

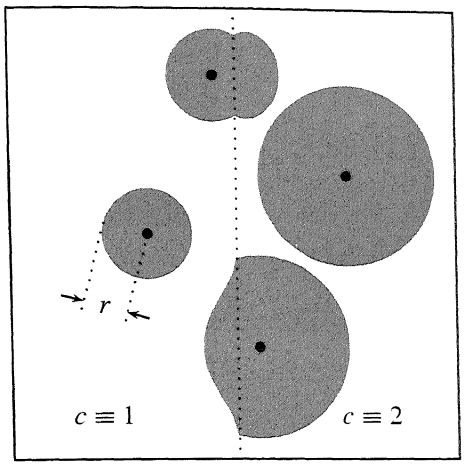

FIGURE 5: Neighbourhoods of four points in a locally scaled distance-interaction process, with scale factor $c \equiv 1$ in the left half and $c \equiv 2$ in the right half of the domain.

In regions where c is constant, the neighbourhood $\partial_c(x) = b_c(x, r)$ is ball shaped and thus similar to the neighbourhood $\partial(x) = b(x, r)$ in the homogeneous and isotropic template; see Figure 5. More precisely, we have the following result.

Proposition 1. *If $c(u) = \tilde{c}$ for all $u \in b(x, \tilde{c}r)$, then $b_c(x, r) = b(x, \tilde{c}r)$.*

Proof. With straightforward calculations it can be shown that $v \in b(x, \tilde{c}r)$ implies that $v \in b_c(x, r)$ and $v \notin b(x, \tilde{c}r)$ implies that $v \notin b_c(x, r)$.

Locally scaled distance-interaction processes have the desired property that in regions where c is constant the process behaves like a scaled version of the template process.

Proposition 2. *Let X be a distance-interaction point process with conditional intensity λ_X. Suppose that $c(u) = \tilde{c}$ for all $u \in b(x, \tilde{c}r)$. Then the conditional intensity of the locally scaled process X_c is given by*

$$\lambda_{X_c}^{(c)}(x \mid \boldsymbol{x}) = \lambda_X\left(\frac{x}{\tilde{c}} \,\middle|\, \frac{\boldsymbol{x}}{\tilde{c}}\right).$$

Note that $c(x) = \tilde{c}$ under the assumption.

Proof. First notice that the assumption implies that $b(x, \tilde{c}r) = b_c(x, r)$; see Proposition 1. The conditional intensity of the locally scaled process is of the form

$$\lambda_{X_c}^{(c)}(x \mid \boldsymbol{x}) = \beta \prod_{\boldsymbol{y} \subseteq_1 \boldsymbol{x}} \varphi(D_c(\boldsymbol{y} \cup \{x\})),$$

where $\boldsymbol{y} \subseteq_1 \boldsymbol{x}$ is short for $\{\boldsymbol{y} \subseteq \boldsymbol{x} : n(\boldsymbol{y}) \geq 1\}$. Suppose that $\boldsymbol{y} \subseteq b(x, \tilde{c}r)$. Since

$$v_c^1([u, v]) = \tilde{c}^{-1} v^1([u, v]) = v^1\left(\left[\frac{u}{\tilde{c}}, \frac{v}{\tilde{c}}\right]\right) \quad \text{for any } u, v \in b(x, \tilde{c}r),$$

we get $D_c(\boldsymbol{y} \cup \{x\}) = D(\boldsymbol{y}/\tilde{c} \cup \{x/\tilde{c}\})$.

On the other hand, suppose that there exists a $u \in y$ such that $u \notin b(x, \tilde{c}r)$. Thus, $v_c^1([u, x]) > r$ and therefore $\varphi(D_c(\boldsymbol{y} \cup \{x\})) = 1 = \varphi(D(\boldsymbol{y}/\tilde{c} \cup \{x/\tilde{c}\}))$. It follows that

$$\lambda_{X_c}^{(c)}(x \mid \boldsymbol{x}) = \beta \prod_{\boldsymbol{y} \subseteq_1 \boldsymbol{x}} \varphi\left(D\left(\frac{\boldsymbol{y}}{\tilde{c}} \cup \left\{\frac{x}{\tilde{c}}\right\}\right)\right) = \lambda_X\left(\frac{x}{\tilde{c}} \mid \frac{\boldsymbol{x}}{\tilde{c}}\right).$$

5. Shot noise weighted processes

Shot noise weighted processes are based on geometric quantities other than pairwise distances. Write $C_{\boldsymbol{x}}(u) = \sum_{x \in \boldsymbol{x}} \mathbf{1}(u \in b(x, r))$ for the template coverage function. Then a shot noise weighted point process with potential function $p(\cdot)$ is defined by

$$f_X(\boldsymbol{x}) \propto \beta^{n(\boldsymbol{x})} \gamma^{-\int p(C_{\boldsymbol{x}}(u)) v^k(\mathrm{d}u)}, \qquad \boldsymbol{x} \in \Omega_{\mathcal{X}}, \tag{5}$$

where $\gamma > 0$ and p is a function on the nonnegative integers $\mathbb{N}_0$ with $p(0) = 0$. The integral $\int p(C_{\boldsymbol{x}}(u)) v^k(\mathrm{d}u)$ is taken on all $\mathbb{R}^k$. The special case $p(n) = \mathbf{1}(n \geq 1)$ is known as an area-interaction point process [1].

The interaction function of a shot noise weighted process is

$$\varphi(\boldsymbol{y}) = \begin{cases} \beta \gamma^{-m(\boldsymbol{y})}, & n(\boldsymbol{y}) = 1, \\ \gamma^{-m(\boldsymbol{y})}, & n(\boldsymbol{y}) > 1, \end{cases}$$

where

$$m(\boldsymbol{y}) = v^k\left(\bigcap_{y \in \boldsymbol{y}} b(y, r)\right) \sum_{l=1}^{n(\boldsymbol{y})} \binom{n(\boldsymbol{y})}{l} (-1)^{n(\boldsymbol{y})-l} p(l);$$

compare with Theorem 3.3 in [17]. As usual, $\varphi(\varnothing)$ is the normalizing constant. Homogeneous shot noise weighted processes are Markov with respect to the overlapping objects relation

$$u \sim v \iff b(u, r) \cap b(v, r) \neq \varnothing \iff \|u - v\| \leq 2r,$$

which means that the neighbourhood of a point is a ball with radius $R = 2r$.

It is easy to show that

$$g(\boldsymbol{x}; v^*) = \beta^{v^0(\boldsymbol{x})} \gamma^{-\int p(\sum_{x \in \boldsymbol{x}} \mathbf{1}(v^1([x,u]) \leq r)) v^k(\mathrm{d}u)}$$

scale invariant. The locally scaled shot noise weighted process has density

$$f_{X_c}^{(c)}(\boldsymbol{x}) \propto g(\boldsymbol{x}; v_c^*) = \beta^{n(\boldsymbol{x})} \gamma^{-\int p(C_{c,\boldsymbol{x}}(u)) v_c^k(\mathrm{d}u)}, \qquad \boldsymbol{x} \in \Omega_{\mathcal{X}'}, \tag{6}$$

a scaled coverage function $C_{c,\boldsymbol{x}}(u) = \sum_{x \in \boldsymbol{x}} \mathbf{1}(u \in b_c(x, r))$.

Van Lieshout and Molchanov [17] show that (5) is integrable if there exists some constant with $0 < C < \infty$ and

$$|p(n)| \leq Cn \quad \text{for all } n \in \mathbb{N}_0. \tag{7}$$

A similar result holds for the scaled process.

Proposition 3. *Under the condition (7), $g(\cdot\,; v_c^*)$ is integrable on $\Omega_{\mathcal{X}'}$ for any $\mathcal{X}' \in \mathbb{B}^k$ and hence the locally scaled process defined by (6) does exist.*

Proof. We show that there exists an $M > 0$ such that $g(\boldsymbol{x}; \nu_c^*) \leq M^{n(\boldsymbol{x})}$ for all $\boldsymbol{x} \in \Omega_{\mathcal{X}'}$. This is fulfilled if

$$\left| \int_{\mathbb{R}^k} p(C_{c,\boldsymbol{x}}(u)) \nu_c^k(\mathrm{d}u) \right| \leq M' n(\boldsymbol{x}) \tag{8}$$

for some $0 < M' < \infty$ and all $\boldsymbol{x} \in \Omega_{\mathcal{X}'}$.

Let $S_C(\boldsymbol{x})$ denote the support of $C_{c,\boldsymbol{x}}$. As $C_{c,\boldsymbol{x}}(u) \leq n(\boldsymbol{x})$, we have

$$\left| \int_{\mathbb{R}^k} p(C_{c,\boldsymbol{x}}(u)) \nu_c^k(\mathrm{d}u) \right| \leq C n(\boldsymbol{x}) \nu_c^k(S_C(\boldsymbol{x}))$$

with C as in (7). Since $S_C(\boldsymbol{x}) = \bigcup_{x \in \boldsymbol{x}} b_c(x, r) \subseteq \mathcal{X}' \oplus b(0, \bar{c}r)$, where $\oplus$ denotes Minkowski addition, and $\nu_c^k(B) \leq \underline{c}^{-k} \nu^k(B)$ for all $B \in \mathcal{B}_k$, (8) holds with

$$M' = C \underline{c}^{-k} \nu^k(\mathcal{X}' \oplus b(0, \bar{c}r)).$$

The locally scaled shot noise weighted process has the interaction function

$$\varphi_c(\boldsymbol{y}) = \begin{cases} \beta \gamma^{-m_c(\boldsymbol{y})}, & n(\boldsymbol{y}) = 1, \\ \gamma^{-m_c(\boldsymbol{y})}, & n(\boldsymbol{y}) > 1, \end{cases}$$

where

$$m_c(\boldsymbol{y}) = \nu_c^k\left(\bigcap_{y \in \boldsymbol{y}} b_c(y, r) \right) \sum_{l=1}^{n(\boldsymbol{y})} \binom{n(\boldsymbol{y})}{l} (-1)^{n(\boldsymbol{y})-l} p(l).$$

It follows that X_c is Markov with respect to the overlapping objects relation

$$u \sim_c v \iff b_c(u, r) \cap b_c(v, r) \neq \varnothing \iff \exists w : \nu_c^1([u, w]) \leq r \wedge \nu_c^1([w, v]) \leq r.$$

The neighbourhood of a point x is

$$\partial_c(x) = \bigcup_{w \in b_c(x,r)} b_c(w, r),$$

which in general is not ball shaped, but contains all points that are $2r$-close to x with respect to ν_c^1. Additionally, it is possible that two points are neighbours in X_c although their scaled distance is larger than $2r$, since the triangular inequality does not necessarily hold for scaled distances defined by ν_c^1. However, in analogy with the results obtained for the distance-interaction processes, the following proposition holds.

Proposition 4. *For a shot noise weighted process, if*

$$c(u) = \tilde{c} \quad \textit{for all } u \in b(x, 2\tilde{c}r),$$

then

$$\partial_c(x) = b(x, 2\tilde{c}r).$$

Proof. Consider $w \in b_c(x, r)$. Then $b(w, \tilde{c}r) \subset b(x, 2\tilde{c}r)$ and thus $c(u) = \tilde{c}$ for all $u \in b(w, \tilde{c}r)$. Therefore, Proposition 1 yields that $b_c(x, r) = b(x, \tilde{c}r)$ and $b_c(w, r) = b(w, \tilde{c}r)$, and hence

$$\partial_c(x) = \bigcup_{w \in b_c(x,r)} b_c(w, r) = \bigcup_{w \in b(x,\tilde{c}r)} b(w, \tilde{c}r) = b(x, 2\tilde{c}r).$$

If the scaling function is constant in a neighbourhood of a point x, then the conditional intensity of a locally scaled shot noise weighted process again behaves as though it is under global scaling.

Proposition 5. *Let X be a shot noise weighted point process with conditional intensity λ_X. Suppose that $c(u) = \tilde{c}$ for all $u \in b(x, 2\tilde{c}r)$. Then the conditional intensity of the locally scaled process X_c is given by*

$$\lambda_{X_c}^{(c)}(x \mid \boldsymbol{x}) = \lambda_X\left(\frac{x}{\tilde{c}} \;\middle|\; \frac{\boldsymbol{x}}{\tilde{c}}\right).$$

Proof. Since the conditional intensity can be written as a product of interactions, $\lambda_X(x \mid \boldsymbol{x}) = \prod_{\boldsymbol{y} \subseteq \boldsymbol{x} \cap \partial(x)} \varphi(\boldsymbol{y} \cup \{x\})$, we only need to show that

$$\prod_{\boldsymbol{y} \subseteq \boldsymbol{x} \cap \partial_c(x)} \varphi_c(\boldsymbol{y} \cup \{x\}) = \prod_{\boldsymbol{y}/\tilde{c} \subseteq \boldsymbol{x}/\tilde{c} \cap \partial(x/\tilde{c})} \varphi\left(\frac{\boldsymbol{y} \cup \{x\}}{\tilde{c}}\right).$$

This is fulfilled if $m_c(\boldsymbol{y} \cup \{x\}) = m((\boldsymbol{y} \cup \{x\}/\tilde{c}))$ for finite $\boldsymbol{y} \subset \partial_c(x)$, $\boldsymbol{y} \neq \varnothing$, i.e. if

$$\nu_c^k\left(\bigcap_{z \in \boldsymbol{y} \cup \{x\}} b_c(z, r)\right) = \nu^k\left(\bigcap_{z \in (\boldsymbol{y} \cup \{x\})/\tilde{c}} b(z, r)\right). \tag{9}$$

By the assumption, $\nu_c([z, w]) = \tilde{c}^{-1}\nu([z, w])$ for any $w, z \in b(x, 2\tilde{c}r)$. Thus, for all $z \in \partial_c(x) = b(x, 2\tilde{c}r)$ (Proposition 4), we have

$$w \in b_c(z, r) \cap b_c(x, r) \iff w \in b(z, \tilde{c}r) \cap b(x, \tilde{c}r).$$

Therefore, $b_c(z, r) \cap b_c(x, r) = b(z, \tilde{c}r) \cap b(x, \tilde{c}r)$ and

$$\bigcap_{z \in \boldsymbol{y} \cup \{x\}} b_c(z, r) = \bigcap_{z \in \boldsymbol{y} \cup \{x\}} b(z, \tilde{c}r) = \bigcap_{z \in (\boldsymbol{y} \cup \{x\})/\tilde{c}} \tilde{c}b(z, r),$$

which immediately leads to (9).

Remark 2. In the present section, the focus has been on local scaling of homogeneous shot noise weighted processes defined using balls $b(x, r)$. It is possible to define a more general type of homogeneous shot noise weighted process with $b(x, r)$ replaced by $x + Z$, where Z is an arbitrary bounded subset of $\mathbb{R}^k$. Expressing Z with generalized radius-vector functions, the homogeneous template process can easily be generalized to local scaling.

6. Approximation of local scaling

For simulation of locally scaled Markov point processes using, for example, the Metropolis–Hastings algorithm (see [5]), expressions of the form $g(\boldsymbol{x}; \nu_c^*)$ have to be evaluated. This usually involves integration with respect to scaled d-dimensional volume measures ν_c^d. In the locally scaled distance-interaction processes introduced in Section 4, for example, we deal with scaled distances

$$\nu_c^1([u, v]) = \int_{[u,v]} c(w)^{-1}\, \mathrm{d}w = \|u - v\|\overline{c^{-1}}(u, v),$$

where $\overline{c^{-1}}(u, v)$ is the integral mean of the inverse scaling function $w \to 1/c(w)$ on the segment $[u, v]$.

For certain scaling functions c, such integrals can be expressed explicitly. However, if we strive to design programs that handle arbitrary scaling functions, we would have to resort to numeric algorithms. Time-consuming calculations can be avoided by defining *approximately scaled* processes that require only pointwise evaluation of the scaling function.

Markov point processes comprise a large variety of models that are essentially different to each other, for example, with respect to the order of interaction, or to the dimensionality of volume measures involved in the definition of their density. There is no best recipe for approximate local scaling of all possible models. A general method for pointwise local scaling of Markov point processes, based on averaging the interaction functions, is presented in Subsection 6.1. Tailor-made, more intuitive approaches for distance-interaction processes and shot noise weighted processes are suggested in the following subsections.

6.1. Local scaling by φ-averaging for Markov point processes

In order to restrict the evaluation of c to the points in the pattern $\boldsymbol{x} = \{x_1, \ldots, x_n\}$, we interpret local scaling as an average of global scalings with scaling factors $c(x_1), \ldots, c(x_n)$. In the context of finite-order interaction Markov point processes where $\varphi(\boldsymbol{y}) \equiv 1$ if $n(\boldsymbol{y}) > m$ for some $m < \infty$, we propose to construct the density of the locally scaled process from averaged interaction functions. Doing so, cliques $\boldsymbol{y}$ are only influenced by their own scaling factors, and not by scale factors from points outside $\boldsymbol{y}$. Thus, the paradigm of local interaction is preserved. Locally scaled interaction functions are defined as the geometric mean of the corresponding interactions in globally scaled patterns,

$$\hat{\varphi}_c(\boldsymbol{y}) = \left(\prod_{y \in \boldsymbol{y}} \varphi\left(\frac{y}{c(y)} \right) \right)^{1/n(\boldsymbol{y})}, \qquad n(\boldsymbol{y}) \geq 1. \tag{10}$$

Thus we obtain the density $\hat{f}^{(c)}_{X_c}$ *by local φ-averaging* as

$$\hat{f}^{(c)}_{X_c}(\boldsymbol{x}) \propto \prod_{\boldsymbol{y} \subseteq_1 \boldsymbol{x}} \hat{\varphi}_c(\boldsymbol{y}),$$

where $\boldsymbol{y} \subseteq_1 \boldsymbol{x}$ is again short for $\{\boldsymbol{y} \subseteq \boldsymbol{x} : n(\boldsymbol{y}) \geq 1\}$. The use of the geometric mean in (10) is motivated by the fact that interaction functions are usually of the form $\varphi(\boldsymbol{y}) = \exp(-p(\boldsymbol{y}))$, where $p(\cdot)$ is the so-called potential function. This notion stems from statistical physics, where Markov point processes were first described as Gibbs processes. Taking the geometric mean of φ means taking the arithmetic mean of the potential function,

$$\hat{\varphi}_c(\boldsymbol{y}) = \exp\left(-\frac{1}{n(\boldsymbol{y})} \sum_{y \in \boldsymbol{y}} p\left(\frac{y}{c(y)} \right) \right), \quad n(\boldsymbol{y}) \geq 1.$$

The following example of a Strauss process shows how this concept is applied to a distance-interaction process.

Example 2. (*Strauss process, continued.*) An approximately locally scaled Strauss process X_c obtained by φ-averaging has second-order interaction

$$\hat{\varphi}_c(\{u, v\}) = \gamma^{(1/2)(\mathbf{1}(\|u-v\| \leq c(u)r) + \mathbf{1}(\|u-v\| \leq c(v)r))}.$$

Therefore, the density takes the form $\hat{f}^{(c)}_{X_c}(\boldsymbol{x}) \propto \beta^{n(\boldsymbol{x})} \gamma^{\hat{s}_c(\boldsymbol{x})}$, where $\hat{s}_c(\boldsymbol{x})$ is calculated from the number of neighbours given by the directed relation

$$u \rightsquigarrow v \iff \|u - v\| \leq c(u)r.$$

The number of directed neighbours divided by two,

$$\hat{s}_c(\boldsymbol{x}) = \sum_{\{u,v\}\subseteq\boldsymbol{x}}^{\neq} \frac{1}{2}(\mathbf{1}(u \rightsquigarrow v) + \mathbf{1}(v \rightsquigarrow u)),$$

can be considered as an approximation of the true number $s_c(\boldsymbol{x})$ of neighbours under local scaling; compare this with Section 4.

6.2. Local scaling by c-averaging for distance-interaction processes

Locally scaled distance-interaction processes as introduced in Section 4 require only the calculation of scaled pairwise distances $v_c^1([u,v]) = \overline{c^{-1}}(u,v)\|u-v\|$. A natural idea is to replace the integral mean $\overline{c^{-1}}(u,v)$ by a simpler mean $\widehat{c^{-1}}(u,v)$ of the inverse scaling factors $c(u)^{-1}$ and $c(v)^{-1}$. We propose to use the harmonic mean $\widehat{c^{-1}}(u,v) = 2/(c(u)+c(v))$. The original neighbour relation $u \sim_c v \iff v_c^1([u,v]) \le r$ is thus approximated by

$$u \sim_{\hat{c}} v \iff \|u-v\| \le \tfrac{1}{2}(c(u)+c(v))r. \tag{11}$$

This relation allows for a nice geometric interpretation. Two points u, v are neighbours if and only if the balls $b(u, \frac{1}{2}c(u)r)$ and $b(v, \frac{1}{2}c(v)r)$ overlap. Note that (11) actually means that the scaling function c itself is locally replaced by the arithmetic mean $\frac{1}{2}(c(u)+c(v))$. Therefore, we call this approach *local scaling by c-averaging*.

Example 3. (*Strauss process, continued.*) As for φ-averaging, the density of an approximately locally scaled Strauss process obtained by c-averaging is of the form

$$\hat{f}_{X_c}^{(c)}(\boldsymbol{x}) \propto \beta^{n(\boldsymbol{x})}\gamma^{\hat{s}_c(\boldsymbol{x})}, \qquad \hat{s}_c(\boldsymbol{x}) = \sum_{\{u,v\}\subseteq\boldsymbol{x}}^{\neq} \mathbf{1}(u \sim_{\hat{c}} v), \qquad \boldsymbol{x} \in \Omega_{\mathcal{X}'}.$$

Now, $\hat{s}_c(\boldsymbol{x})$ is the number of neighbour pairs with respect to the approximate neighbour relation $\sim_{\hat{c}}$ given by (11).

6.3. Local scaling by influence zones for shot noise weighted processes

Shot noise weighted processes as defined in Section 5 require the evaluation of the coverage function $C_{c,\boldsymbol{x}}(u) = \sum_{x\in\boldsymbol{x}} \mathbf{1}(u \in b_c(x,r))$, which gives the number of 'influence zones' $b_c(x,r)$ covering a point u.

In Proposition 1 we saw that $b_c(x,r) = b(x,\tilde{c}r)$ if $c(u) = \tilde{c} = c(x)$ for all $u \in b(x,\tilde{c}r) = b(x,c(x)r)$. Assuming that c does not vary very much in $b(x,c(x)r)$, we can use this result to approximate the influence zones by $b_c(x,r) \approx b(x,c(x)r)$ and thus obtain the coverage function $\hat{C}_{c,\boldsymbol{x}}(u) = \sum_{x\in\boldsymbol{x}} \mathbf{1}(u \in b(x,c(x)r))$. Calculating the density function

$$\hat{f}_{X_c}^{(c)}(\boldsymbol{x}) \propto \beta^{n(\boldsymbol{x})}\gamma^{-\int p(\hat{C}_{c,\boldsymbol{x}}(u))v_c^k(\mathrm{d}u)}$$

still requires integration with respect to the locally scaled measure v_c^k. But even when dealing with homogeneous processes, the integral $\int p(C_{c,\boldsymbol{x}}(u))\,\mathrm{d}u$ is usually approximated by grid methods. Once the coverage function $\hat{C}_{c,\boldsymbol{x}}(\cdot)$ is known, evaluating

$$\int_{\mathbb{R}^k} p(\hat{C}_{c,\boldsymbol{x}}(u))v_c^k(\mathrm{d}u) = \int_{\mathbb{R}^k} p(\hat{C}_{c,\boldsymbol{x}}(u))c(u)^{-k}v^k(\mathrm{d}u)$$

is therefore no bigger a problem than evaluating the corresponding integral in a homogeneous (template) setting.

7. Sequential local scaling

A sequential point process $\vec{X}$ on $\mathcal{X} \in \mathbb{B}^k$ is a random variable taking values in $\Omega_{s,\mathcal{X}}$, the set of all finite sequences $\vec{x} = (x_1, \ldots, x_n)$ of points in $\mathcal{X}$. The set $\Omega_{s,\mathcal{X}}$ is equipped with the σ-algebra $\mathcal{F}_{s,\mathcal{X}}$ generated by the Borel σ-algebras $\mathcal{B}_n$ on $\mathcal{X}^n$. In the following, we will consider sequential point processes that have a density $f_{\vec{X}}$ with respect to the probability measure μ on $\Omega_{s,\mathcal{X}}$, where, for all $F \in \mathcal{F}_{S,\mathcal{X}}$,

$$\mu(F) = \sum_{n=0}^{\infty} \frac{\mathrm{e}^{-\nu^k(\mathcal{X})}}{n!} \int_{\mathcal{X}} \cdots \int_{\mathcal{X}} \mathbf{1}((x_1, \ldots, x_n) \in F)\, \mathrm{d}\nu^k(x_1) \ldots \mathrm{d}\nu^k(x_n).$$

This measure on the space $(\Omega_{s,\mathcal{X}}, \mathcal{F}_{s,\mathcal{X}})$ of sequential point sets corresponds to the unit rate Poisson point process for unordered (nonsequential) point processes on $(\Omega_{\mathcal{X}}, \mathcal{F}_{s,\mathcal{X}})$.

For sequential point processes, the Papangelou conditional intensity,

$$\lambda_{\vec{X}}(x \mid (x_1, \ldots, x_n)) = \begin{cases} \dfrac{f_{\vec{X}}(x_1, \ldots, x_n, x)}{f_{\vec{X}}(x_1, \ldots, x_n)}, & f_{\vec{X}}(x_1, \ldots, x_n) > 0, \\ 0, & \text{otherwise}, \end{cases} \tag{12}$$

gains a particularly intuitive meaning, since $\lambda_{\vec{X}}(x \mid \vec{x})\nu^k(\mathrm{d}x)$ relates to the conditional probability of adding a new point x in a region $\mathrm{d}x$ to an existing sequence $\vec{x} = (x_1, \ldots, x_n)$.

In analogy to hereditary (unordered) point processes, we define hereditary sequential point processes as follows.

Definition 4. A sequential point process $\vec{X}$ with density $f_{\vec{X}}$ is called *hereditary* if $f_{\vec{X}}(x_1, \ldots, x_n) > 0$ implies that $f_{\vec{X}}(\vec{y}) > 0$ for all subsequences $\vec{y}$ of $(x_1, \ldots, x_n)$.

If a sequential point process is hereditary, then, as a consequence of (12), the density can be written as

$$f_{\vec{X}}(x_1, \ldots, x_n) \propto \prod_{j=1}^{n} \lambda_{\vec{X}}(x_j \mid \vec{x}_{<j}), \qquad \vec{x}_{<j} = (x_1, \ldots, x_{j-1}).$$

This gives rise to a straightforward idea of defining local scaling in the sequential context by means of locally scaled conditional intensities. We start again with a homogeneous, now sequential, template process $\vec{X}$, which means that $f_{\vec{X}}$ and, hence, $\lambda_{\vec{X}}$ are translation invariant and in principle defined for any finite point sequence in $\mathbb{R}^k$. Motivated by the effect of global scaling on the conditional intensity, cf. (1), we require that the Papangelou conditional intensity of the scaled sequential process $\vec{X}_c$ fulfils

$$\lambda^{(c)}_{\vec{X}_c}(x \mid \vec{x}) = \lambda_{\vec{X}}\left(\frac{x}{c(x)} \,\middle|\, \frac{\vec{x}}{c(x)}\right), \tag{13}$$

where c as before is a measurable, bounded scaling function that is bounded away from 0. Note that, for locally scaled unordered point processes, the corresponding property (2) is only fulfilled in regions where the scaling function c is constant.

Definition 5. Suppose that $\vec{X}$ is a homogeneous hereditary sequential point process on $\mathcal{X}$ with Papangelou conditional intensity $\lambda_{\vec{X}}$ given by (12). Let μ_c be the distribution of a sequential

Poisson point process with the locally scaled volume measure ν_c^k as intensity measure, and let $\mathcal{X}' \in \mathbb{B}^k$ be arbitrary.

Then the *locally scaled sequential point process* $\vec{X}_c$ on $\mathcal{X}'$ with template $\vec{X}$ is defined by the density

$$f_{\vec{X}_c}^{(c)}(x_1, \ldots, x_n) \propto \prod_{j=1}^{n} \lambda_{\vec{X}}\left(\frac{x_j}{c(x_j)} \,\middle|\, \frac{\vec{x}_{<j}}{c(x_j)}\right) \tag{14}$$

with respect to μ_c, provided that $f_{\vec{X}_c}$ is integrable on $\Omega_{s,\mathcal{X}'}$.

Sequential templates can easily be obtained from homogeneous unordered point processes, since any unordered point process X on $\mathcal{X}$ with a density f_X with respect to the unit rate Poisson point process can be converted into a corresponding ordered point process $\vec{X}$ by defining the density $f_{\vec{X}}$ with respect to μ as

$$f_{\vec{X}}(x_1, \ldots, x_n) = f_X(\{x_1, \ldots, x_n\}). \tag{15}$$

This means that every permutation of $\vec{x}$ has the same density; see also [4, p. 122]. We shall refer to the process $\vec{X}$ with density (15) as the *sequentialized version* of the process X.

Clearly, the sequentialized process $\vec{X}$ is hereditary if the unordered process X is hereditary. Combining (12) and (15), the Papangelou conditional intensity of a sequentialized process becomes $\lambda_{\vec{X}}(x \mid x_1, \ldots, x_n) = \lambda_X(x \mid \{x_1, \ldots, x_n\})$. Thereby, the Papangelou conditional intensity of the corresponding locally scaled point process $\vec{X}_c$ is

$$\lambda_{\vec{X}_c}^{(c)}(x \mid \vec{x}) = \lambda_X\left(\frac{x}{c(x)} \,\middle|\, \frac{\boldsymbol{x}}{c(x)}\right),$$

cf. (13), and the density (14) of $\vec{X}_c$ becomes

$$f_{\vec{X}_c}^{(c)}(x_1, \ldots, x_n) \propto \prod_{j=1}^{n} \lambda_X\left(\frac{x_j}{c(x_j)} \,\middle|\, \frac{\boldsymbol{x}_{<j}}{c(x_j)}\right). \tag{16}$$

The locally scaled process $\vec{X}_c$ is locally stable, i.e. $\lambda_{\vec{X}_c}^{(c)}$ is bounded above, if the homogeneous unordered template X is locally stable, i.e. if λ_X is bounded above. Then (16) is integrable. As before, a constant scaling function simply yields a globally scaled sequential point process. Note in particular that a locally scaled sequential point process is isotropic in the sense of (3) if the template is isotropic.

Example 4. (*Strauss process, continued.*) For a Strauss process X, $\lambda_X(x \mid \boldsymbol{x}) = \beta\gamma^{s(x;\boldsymbol{x})}$ where $s(x; \boldsymbol{x}) = s(\boldsymbol{x} \cup \{x\}) - s(\boldsymbol{x})$ denotes the number of points in $\boldsymbol{x}$ that are closer than r to the new point $x \notin \boldsymbol{x}$. Here

$$f_{\vec{X}_c}^{(c)}(x_1, \ldots, x_n) \propto \beta^n \gamma^{s_c(x_1, \ldots, x_n)},$$

with

$$\begin{aligned} s_c(x_1, \ldots, x_n) &= \sum_{i<j} \mathbf{1}\left(\left\| \frac{x_j}{c(x_j)} - \frac{x_i}{c(x_j)} \right\| \leq r\right) \\ &= \sum_{i<j} \mathbf{1}(x_j \rightsquigarrow_c x_i), \end{aligned}$$

where

$$x_j \rightsquigarrow_c x_i \iff \|x_j - x_i\| \leq c(x_j)r.$$

This is the same directed neighbour relation as in φ-averaging, see Subsection 6.1.

Specializing to the hard-core model ($\gamma = 0$), we obtain a sequential inhibition model in which each point upon arrival keeps a distance $c(x_j)r$ away from previously arrived points; see also [3].

8. Discussion

Inhomogeneity in natural structures may be caused by very different mechanisms. Correspondingly, there is a myriad of ways to define inhomogeneous models. Therefore, some restrictions have to be introduced that replace the usual homogeneity condition. The three models for inhomogeneous point processes described in the introduction stand for three different situations. In the first model, the interaction between points is independent of location. In the second model, inhomogeneity results from a (physical) location-dependent thinning, and in the third from (physical) deformation of the matrix. In this paper, we introduced yet another approach which yields models for patterns that are homogeneous up to a local scale factor. Such point processes may describe a spatial arrangement of spheres with diameters that vary with location, see Figure 1, or situations where both intensity and interaction are governed by the same external factor, such as desert plant communities that are ruled by water supply.

When it comes to choosing an appropriate model for a given situation, there will sometimes be prior information about the physical genesis of the patterns that strongly suggests one of the approaches. In general, however, it will be necessary to define criteria for the best choice which can then be used to develop model tests. These criteria will strongly depend on the modelling purpose. Often it is desirable to pick the model that reflects best the local arrangement of neighbouring points. Therefore, a test on the local scaling assumption could exploit scale-invariant local geometric properties, as, for example, shape factors of corresponding Voronoi cells. Statistical inference will be studied in detail in a forthcoming paper.

The intensity $\lambda_c(x)$ of a globally scaled point process is proportional to the intensity λ of the template, $\lambda_c(x) = c^{-k}\lambda(x/c) \approx c^{-k}\lambda$, since the intensity of a homogeneous template is approximately constant. Analogously, the intensity of a locally scaled process is (approximately)

$$\lambda_c(x) \approx c(x)^{-k}\lambda. \tag{17}$$

This allows us firstly to model practically any inhomogeneous intensity and, secondly, to retrieve the scaling function (up to a proportionality constant) from a given or estimated density. In this aspect the scaling function plays a similar role to the survival probability of the thinning model.

Once the scaling function has been estimated, it can be used to subsequently fit the parameters of the template process and thus to complete the model specification. A similar approach has been followed by Nielsen and Jensen [11] for fitting the transformation model. Furthermore, possible empirical relations between an estimated scaling function and explanatory variables such as water supply in the desert vegetation case can be used for prediction purposes.

Future work will concentrate on validating the approximation (17) as well as on development of model tests and other statistical methods. Moreover, the idea of local scaling will be extended to other random sets.

Acknowledgements

The authors thank Markus Kiderlen for fruitful discussions and an anonymous referee for valuable suggestions to improve the paper. M. N. M. van Lieshout acknowledges the hospitality of the Laboratory of Computational Stochastics, Aarhus University, during her visit in December 2000. This research was supported by the Centre for Mathematical Physics and Stochastics (MaPhySto), funded by a grant from the Danish National Research Foundation.

References

[1] BADDELEY, A. J. AND VAN LIESHOUT, M. N. M. (1995). Area-interaction point processes. *Ann. Inst. Statist. Math.* **47**, 601–619.

[2] BADDELEY, A. J., MØLLER, J. AND WAAGEPETERSEN, R. (2000). Non- and semiparametric estimation of interaction in inhomogeneous point patterns. *Statist. Neerlandica* **54**, 329–350.

[3] CLAUSEN, W. H. O., RASMUSSEN, H. H. AND RASMUSSEN, M. H. (2000). Inhomogeneous point processes. Doctoral Thesis, Aalborg University.

[4] DALEY, D. J. AND VERE-JONES, D. (1988). *An Introduction to the Theory of Point Processes*. Springer, New York.

[5] GEYER, C. (1999). Likelihood inference for spatial point patterns. In *Stochastic Geometry—Likelihood and Computation*, eds O. E. Barndorff-Nielsen, W. S. Kendall and M. N. M. van Lieshout, Chapman and Hall, London, pp. 79–140.

[6] HAHN, U. *et al.* (1999). Stereological analysis and modelling of gradient structures. *J. Microscopy* **195**, 113–124.

[7] HARARY, F. (1969). *Graph Theory*. Addison-Wesley, Reading, MA.

[8] JENSEN, E. B. V. AND NIELSEN, L. S. (2000). Inhomogeneous Markov point processes by transformation. *Bernoulli* **6**, 761–782.

[9] JENSEN, E. B. V. AND NIELSEN, L. S. (2001). A review on inhomogeneous spatial point processes. In *Selected Proc. Symp. Inference for Stoch. Process.* (IMS Lecture Notes Monogr. Ser. **37**), eds I. V. Basawa, C. C. Heyde and R. L. Taylor, Institute of Mathematical Statistics, Beachwood, OH, pp. 297–318.

[10] NIELSEN, L. S. (2001). Point process models allowing for interaction and inhomogeneity. Doctoral Thesis, University of Aarhus.

[11] NIELSEN, L. S. AND JENSEN, E. B. V. (2003). Statistical inference for transformation inhomogeneous point processes. To appear in *Scand. J. Statist.*

[12] OGATA, Y. AND TANEMURA, M. (1986). Likelihood estimation of interaction potentials and external fields of inhomogeneous spatial point patterns. In *Proc. Pacific Statist. Congress 1985*, eds I. S. Francis, B. F. J. Manly and F. C. Lam, Elsevier, Amsterdam, pp. 150–154.

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

[14] STOYAN, D. AND STOYAN, H. (1998). Non homogeneous Gibbs process models for forestry—a case study. *Biometrical J.* **40**, 521–531.

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

[16] VAN LIESHOUT, M. N. M. (2000). *Markov Point Processes and Their Applications*. World Scientific, Singapore.

[17] VAN LIESHOUT, M. N. M. AND MOLCHANOV, I. S. (1998). Shot noise weighted processes: a new family of spatial point processes. *Commun. Statist. Stoch. Models* **14**, 715–734.