ISSN 0908-1216

Characterisation results and Markov chain Monte Carlo algorithms including exact simulation for some spatial point processes

O Häggström,[1] M N M van Lieshout[2] and J Møller[3]

[1] *Department of Mathematics, Chalmers University of Technology, S-412 96 Göteborg, Sweden*
[2] *Department of Statistics, University of Warwick, Coventry CV4 7AL, United Kingdom*
[3] *Department of Mathematics, Aalborg University, Fredrik Bajers Vej 7E, DK-9220 Aalborg Ø, Denmark*

Abstract

The area-interaction process and the continuum random-cluster model are characterised in terms of certain functional forms of their respective conditional intensities. In certain cases, these two point process models can be derived from a bivariate point process model which in many respects is simpler to analyse and simulate. Using this correspondence we devise a two-component Gibbs sampler, which can be used for fast and exact simulation by extending the recent ideas of Propp and Wilson. We further introduce a Swendsen–Wang type algorithm. The importance of the results within spatial statistics as well as statistical physics is outlined.

Key words & Phrases: area-interaction process; continuum random-cluster model; exact simulation; Gibbs sampling; Markov chain Monte Carlo; nearest-neighbour Markov point process; Papangelou conditional intensity; penetrable sphere model; phase transition; spatial point processes; Swendsen–Wang algorithm; Widom–Rowlinson mixture model.

AMS classification numbers: 60G55, 62M30 (primary); 60D05, 60K35 (secondary).

1 Introduction

This paper is concerned with three related models for spatial point processes, namely the area-interaction process [1], the continuum random-cluster model [6, 24] and the penetrable spheres mixture model [35, 16, 29, 30]. These are of interest in spatial statistics in situations where the independence property of the Poisson process needs to be replaced either by attraction or by repulsion between points. They are also highly relevant in statistical physics, where the first and the third model provide the most well-known example of a phase transition in a continuous setting. After providing some background material on spatial point processes in Section 2, Section 3 gives characterisation results for the models mentioned above, providing additional motivation for their use.

In the remaining sections we devise algorithms for Markov chain Monte Carlo (MCMC) simulation of the models. The two-component Gibbs sampler for the mixture model introduced in Section 4 is used to explore how phase transition occurs; this is a problem which to a large extent lacks rigorous solution in the statistical physics literature. For accounts of MCMC techniques related to spatial statistics and statistical physics, see e.g. [4, 13, 14, 24, 32] and the references therein.

A common problem in MCMC simulation is that the available rigorous bounds for rates of convergence are not good enough to be useful in practice, so that one is forced to use the (not particularly satisfactory) method of running the chain for a reasonably long time and then just hope that it is close to the stationary distribution. Therefore, it is highly remarkable that Propp and Wilson [26] recently found a simulation technique which gives a sample from the Ising or Potts model that has *exactly* the right distribution, and which works in practice even for fairly large systems. In Section 5 we demonstrate how the two-component Gibbs sampler can be combined with Propp and Wilson's ideas in order to obtain exact samples from the mixture model (and hence also from the attractive area-interaction model and some particular cases of the continuum random-cluster model). Recently, and independently of our work, Kendall [19] has also demonstrated a way to apply the Propp–Wilson techniques in a more general point process setting. While Kendall

uses a coupling construction of spatial birth-and-death processes, our approach seems much simpler to present and implement. The possibility of doing exact simulation for general point process models by replacing spatial birth-and death processes with Markov chains generated by the Metropolis-Hastings algorithm will be studied in [20].

Section 5 also contains some empirical findings. We show that exact samples from the mixture model can be obtained in only a few steps as long as the rate of the underlying Poisson processes is small or moderate. Further we investigate empirically the phase transition behaviour. In particular, the number of steps needed per sample increases when we approach the critical parameter, and realisations tend to become dominated by one component.

The two-component Gibbs sampler in Section 4 is somewhat similar in spirit to the celebrated Swendsen–Wang algorithm [34] for simulating Potts models, in that it changes the entire configuration in one swap (as opposed to the traditional single-point updating algorithms). In Section 6, we present another algorithm which we believe is even closer in spirit to the one of Swendsen and Wang.

2 Background and notation

We consider spatial point processes X on a bounded Borel set $A \subset \mathbb{R}^d$, defined by their density $f(\cdot)$ with respect to a unit rate Poisson process. Let Ω_A denote the space of all finite point configurations in A without multiple points, that is

$$\Omega_A = \{\mathbf{x} \subset A \mid n(\mathbf{x}) < \infty\}\,.$$

Here $n(\mathbf{x})$ denotes the number of points n in the configuration $\mathbf{x} = \{x_1, \ldots, x_n\}$.

The Papangelou conditional intensity [8, 17] of X is given by

$$\lambda^*(\mathbf{x}, u) = \begin{cases} \frac{f(\mathbf{x} \cup \{u\})}{f(\mathbf{x})} & \text{if } f(\mathbf{x}) > 0 \\ 0 & \text{else} \end{cases} \tag{1}$$

for $\mathbf{x} \in \Omega_A$ and $u \in A \setminus \mathbf{x}$. In fact, there is a one-to-one correspondence between $f(\cdot)$ and $\lambda^*(\cdot, \cdot)$ if the probability density is hereditary, i.e. whenever $f(\mathbf{y}) > 0$ implies $f(\mathbf{x}) > 0$ for all $\mathbf{x} \subseteq \mathbf{y}$.

The conditional intensity can be used to define a Markov property for point processes. Given a symmetric neighbourhood relation $\sim$ on A, X is *Ripley–Kelly Markov* [27] with respect to $\sim$ if its density is hereditary and $\lambda^*(\mathbf{x}, u)$ depends only on u and its neighbours $\{x_i \in \mathbf{x} : u \sim x_i\}$ in $\mathbf{x}$. Heuristically, $\lambda^*(\mathbf{x}, u)du$ can readily be interpreted as the conditional probability of having a point in the infinitesimal region du centred at u given the rest of the pattern is $\mathbf{x}$.

For modelling purposes it may thus be useful to establish characterisation results in terms of $\lambda^*(\cdot, \cdot)$. Strauss [33] and Kelly and Ripley [18] considered the Strauss process

$$f(\mathbf{x}) = \alpha \beta^{n(\mathbf{x})} \gamma^{s(\mathbf{x})} \tag{2}$$

where α is the normalising constant, $\beta > 0$ and $\gamma \in [0, 1]$ are model parameters, and $s(\mathbf{x})$ denotes the number of pairs of points $\xi, \eta \in \mathbf{x}$ such that $\xi \sim \eta$. Any symmetric relation $\sim$ (defined on an arbitrary space) may be considered here, but usually $\xi \sim \eta$ if and only if $||\xi - \eta|| \leq R$ (for some prefixed $R > 0$). Then, if $\gamma = 0$, (2) defines a hard core process where no points are allowed to be within distance R of each other. Assuming that A contains three points ξ, η, ζ satisfying $\xi \sim \eta$, $\xi \sim \zeta$ and $\eta \not\sim \zeta$, Kelly and Ripley verified that if the conditional intensity is of the form

$$\lambda^*(\mathbf{x}, u) = g(n(\mathbf{x} \cap B_u)) \tag{3}$$

for all $\mathbf{x} \in \Omega_A$ and $u \in A \setminus \mathbf{x}$, where $g : \mathbf{N}_0 \to [0, \infty)$ and B_u denotes the closed ball centered at u with radius R, then the density on Ω_A induced by (3) is of the form (2). Conversely, (2) clearly implies (3), so the Strauss process is uniquely characterised by (3).

Strauss [33] suggested (2) with $\gamma > 1$ as a model for the clustering of Californian redwood seedlings around older stumps but, as pointed out in [18], (2) is only well-defined for $0 \leq \gamma \leq 1$. Note that the model exhibits interactions between pairs of points only. Pairwise interaction models appear to be a useful and flexible class of models for regular patterns, but probably not so for clustered patterns [9, 11, 24]. A more promising way of modelling attraction between points in a spatial pattern is to allow interaction terms of higher order, or to generalise the Markov property to depend on the configuration as in the definition of *nearest-neighbour Markov point processes* (cf. [2]). Examples of the former are penetrable sphere [35] or area-interaction models [1]. For the latter, consider the

connected component relation $\underset{\mathbf{x}}{\sim}$ on $\mathbf{x} \in \Omega_A$ defined by

$$\xi \underset{\mathbf{x}}{\sim} \eta \Leftrightarrow \xi = x_1 \sim x_2 \sim \cdots \sim x_m = \eta \tag{4}$$

for some subconfiguration $\{x_1, \ldots x_m\} \subseteq \mathbf{x}$. Here for specificity we write $u \sim v$ whenever $||u - v|| \leq R$, but any symmetric relation may be considered as well. Then, as shown in Baddeley et al. [3], the general expression for the density of a nearest-neighbour point process defined with respect to (4) becomes

$$f(\mathbf{x}) = \alpha \prod_{\mathbf{y} \in C(\mathbf{x})} \Phi(\mathbf{y}) \tag{5}$$

where $C(\mathbf{x})$ is the set of (maximal) connected components defined by $\mathbf{x}$ and $\Phi(\cdot) \geq 0$ satisfies certain regularity conditions (if $f(\cdot) > 0$ then strict positivity of $\Phi(\cdot)$ is the only condition).

In fact, the Ripley and Kelly [27] Markov point processes as well as certain Poisson cluster processes [8] are special cases of (5), cf. Baddeley et al. [3]. Lattice processes with a density of a form similar to (5) have recently been studied by Møller and Waagepetersen [25].

3 Characterisation results

The Strauss process is Ripley–Kelly Markov, hence a fortiori of the form (5) with $\Phi(\mathbf{y}) = \beta^{n(\mathbf{y})}\gamma^{s(\mathbf{y})}$. In this paper, we consider two other models with respectively, $\Phi(\mathbf{y}) = \beta^{n(\mathbf{y})}/\gamma$ and $\Phi(\mathbf{y}) = \beta^{n(\mathbf{y})}\gamma^{-|U\mathbf{y}|}$, where $U_{\mathbf{y}} = \cup_i B_{y_i}$ and $|\cdot|$ denotes Lebesgue measure. In both cases, the densities are well-defined for all $\beta, \gamma > 0$ where the models exhibit regularity for $0 < \gamma < 1$ and clustering for $\gamma > 1$. The hard core process may be considered as a limiting case of both models.

First, consider the continuum random-cluster model defined by its density

$$f(\mathbf{x}) = \alpha\beta^{n(\mathbf{x})}\gamma^{-c(\mathbf{x})}, \tag{6}$$

where $c(\mathbf{x})$ denotes the number of connected components in $U_{\mathbf{x}}$. This corresponds to $\Phi(\mathbf{y}) = \beta^{n(\mathbf{y})}/\gamma$. This model seems to have been rediscovered many times, e.g. in [6] and [23]; the earliest appearance in the literature we are aware of is in [21]. We use the name "continuum random-cluster model" because of the strong analogy with the random-cluster representation of the Ising/Potts models introduced by Fortuin

and Kasteleyn [10] (see [5] and [15] for recent reviews). We have the following characterisation result.

Theorem 1 *A density $f(\cdot)$ is a continuum random-cluster process if and only if*

$$\lambda^*(\mathbf{x}, u) = g(c(\mathbf{x} \cup \{u\}) - c(\mathbf{x})) \tag{7}$$

for all $\mathbf{x} \in \Omega_A, u \in A \setminus \mathbf{x}$ and a function $g : \mathbb{Z} \to (0, \infty)$.

Note that

$$c(\mathbf{x}, u) = 1 + c(\mathbf{x}) - c(\mathbf{x} \cup \{u\})$$

is the number of 'clusters' $U_{\mathbf{y}}, \mathbf{y} \in C(\mathbf{x})$ generated by $\mathbf{x}$ which are intersected by the disc associated with u, so (7) is equivalent to $\frac{f(\mathbf{x} \cup \{u\})}{f(\mathbf{x})} = g(c(\mathbf{x}, u))$.

Note also that the result still holds if the fixed range relation $\sim$ is replaced by any other symmetric relation defined on a finite measure space, cf. [24]. For instance, one may consider configurations of path-connected sets like e.g. discs, where two discs are related if and only if they overlap each other.

Moreover, the positivity condition on $\lambda^*(\cdot, \cdot)$ is necessary: if $g(n) = \beta \mathbf{1}\{n = 1\}$ then (7) gives a hard core process (all $||x_i - x_j|| \geq 2R$).

Proof : Clearly, if $f(\cdot)$ belongs to the family of continuum random-cluster processes, its conditional intensity is of the form specified in (7). Reversely, assuming (7) holds, we proceed to prove that $f(\cdot)$ is a continuum random-cluster process by induction with respect to $n(\mathbf{x})$, the number of points. Setting

$$\beta = g(0) \text{ and } \gamma = \frac{g(0)}{g(1)} \tag{8}$$

then (6) holds for $n(\mathbf{x}) \leq 1$. Assuming the statement holds for configurations with up to $n(\mathbf{x}) = n$ points and writing $k = c(\mathbf{x} \cup \{u\}) - c(\mathbf{x})$,

$$f(\mathbf{x} \cup \{u\}) = f(\mathbf{x})\lambda^*(\mathbf{x}, u) = \alpha \beta^n \gamma^{-c(\mathbf{x})} g(k).$$

Combining this with the definitions in (8) give that $f(\cdot)$ has the desired form if $k = 0$ or $k = 1$. Suppose $k < 0$. Then the disc associated with u intersects at least two of the clusters generated by $\mathbf{x}$, and it is easily seen

that we can delete a disc from one of these clusters without changing the total number of clusters generated by $\mathbf{x}$, i.e.

$$c(\mathbf{x} \cup \{u\}) = c((\mathbf{x} \cup \{u\}) \setminus \{x_i\})$$

for some $x_i \in \mathbf{x}$. Hence, by the induction hypothesis and (8),

$$\begin{aligned} f(\mathbf{x} \cup \{u\}) &= f((\mathbf{x} \cup \{u\}) \setminus \{x_i\})\lambda^*((\mathbf{x} \cup \{u\}) \setminus \{x_i\}, x_i) \\ &= \alpha\beta^n\gamma^{-c((\mathbf{x}\cup\{u\})\setminus\{x_i\})}g(0) \\ &= \alpha\beta^{n+1}\gamma^{-c(\mathbf{x}\cup\{u\})}. \end{aligned}$$

□

Next turn to the model specified by $\Phi(\mathbf{y}) = \beta^{n(\mathbf{y})}\gamma^{-|U_\mathbf{y}|}$ or, equivalently, by density

$$f(\mathbf{x}) = \alpha\beta^{n(\mathbf{x})}\gamma^{-|U_\mathbf{x}|} \tag{9}$$

This is the area-interaction model studied by Baddeley and Van Lieshout [1], a generalisation of the (marginal) penetrable sphere model introduced by Widom and Rowlinson [35] in statistical physics. Contrary to the continuum random-cluster model, (9) is Markov in the Ripley–Kelly sense [27] but it has interactions of arbitrarily high order, cf. [1]. In the one dimensional case, it can also be presented as a *pairwise* interaction sequential neighbours Markov process [2], since

$$f(\mathbf{x}) = \alpha\gamma^{-2R}\beta^n \exp[-(\log\gamma)\sum_{i=1}^{n-1}\min(x_{i+1} - x_i, 2R)]$$

for $\mathbf{x} = \{x_1, \ldots, x_n\}$ with $x_1 < x_2 < \ldots < x_n$.

Clearly, to prove a characterisation result for (9) we need a positivity condition, since the conditional intensity of a hard core process

$$f(\mathbf{x}) = \alpha\beta^{n(\mathbf{x})}\mathbf{1}\{||x_i - x_j|| \geq 2R \text{ for all } i \neq j\}$$

can be written as

$$\lambda^*(\mathbf{x}, u) = \beta\mathbf{1}\{\rho(u, \mathbf{x}) \geq 2R\} = \beta\mathbf{1}\{|B_u \cap U_\mathbf{x}| = 0\}.$$

The hard core process does not fall within the class of area-interaction processes, although it can be seen as a limiting case.

Theorem 2 *Given a bounded Borel set $A \subseteq \mathbb{R}^2$ containing an open ball of radius $4R$, a density $f : \Omega_A \to (0,\infty)$ is an area-interaction density if and only if*

$$\lambda^*(\mathbf{x}, u) = g(|B_u \cap U_{\mathbf{x}}|)$$

for all $\mathbf{x} \in \Omega_A, u \in A \setminus \mathbf{x}$ and a left-continuous function $g : [0, \pi R^2] \to (0,\infty)$.

Note that although the characterisation Theorem 2 is stated for the planar case only, generalisations to higher dimensions are straightforward. In order to prove Theorem 2 we need the following Lemma.

Lemma 1 *If $g : [0, \pi R^2] \to (0,\infty)$ is left-continuous and $\frac{g(s+t)g(0)}{g(s)g(t)} = 1$ for all $s,t \in [0, \pi R^2]$ such that $s + t \in [0, \pi R^2]$, then $g(s) = g(0)\gamma^s$ for some $\gamma > 0$ and all $s \in [0, \pi R^2]$.*

Proof : Extend the function $g(\cdot)$ onto the whole of $(0,\infty)$ as follows.

$$g(k\pi R^2 + s) = \left(\frac{g(\pi R^2)}{g(0)}\right)^k g(s), \quad s \in (0, \pi R^2], k \in \mathbb{N}.$$

Then $g(\cdot)$ is left-continuous on $(0,\infty)$. Moreover, for $s,t \in [0, \pi R^2]$ with $s+t > \pi R^2$,

$$\frac{g(0)g(s+t)}{g(s)g(t)} = \frac{g(\pi R^2)g(s+t-\pi R^2)}{g(s)g(t)}.$$

Now choose $c_1, c_2 > 0$ such that $s - c_1\pi R^2 \geq 0$, $t - c_2\pi R^2 \geq 0$, and $c_1 + c_2 = 1$. Then

$$\begin{aligned}
\frac{g(0)g(s+t)}{g(s)g(t)} &= \frac{g(\pi R^2)g(s - c_1\pi R^2)g(t - c_2\pi R^2)}{g(s)g(t)g(0)} \\
&= \frac{g(\pi R^2)g(s - c_1\pi R^2)g(t - c_2\pi R^2)g(0)g(0)}{g(0)g(c_1\pi R^2)g(s - c_1\pi R^2)g(c_2\pi R^2)g(t - c_2\pi R^2)} \\
&= \frac{g(\pi R^2)g(0)}{g(c_1\pi R^2)g(c_2\pi R^2)}
\end{aligned}$$

The latter expression equals 1 by assumption, since $c_1 + c_2 = 1$. By Hamel's theorem $g(t) = g(0)e^{-\lambda t}$ where $\lambda = \frac{-1}{\pi R^2} \log \frac{g(\pi R^2)}{g(0)}$. In particular

$\equiv [0, \pi R^2]$,

$$g(s) = g(0)\gamma^s, \quad \gamma = \left(\frac{g(\pi R^2)}{g(0)}\right)^{1/(\pi R^2)} > 0.$$

□

Theorem 2: If $f(\cdot)$ is an area-interaction process,

$$\frac{f(\mathbf{x} \cup \{u\})}{f(\mathbf{x})} = \beta\gamma^{-|B_u|+|B_u \cap U_\mathbf{x}|} = g(|B_u \cap U_\mathbf{x}|)$$

ontinuous function $g(t) = \beta\gamma^{-\pi R^2}\gamma^t$.

ove the reverse statement, take $s, t \in [0, \pi R^2]$ such that $s + t \in$ We will show that $g(s+t)g(0) = g(s)g(t)$, so assume without nerality that $t > 0$ and $0 < s < \pi R^2$.

se $\epsilon > 0$. By the Heine–Borel theorem and since A contains an of radius $4R$, there exist $u, v \in A$, $\mathbf{y} \in \Omega_A$, such that $|B_u \cap B_v| =$ $| = t' \in [t-\epsilon, t]$, where $D = B_v \cap U_\mathbf{y}$ and $||y_i - u|| = 2R$ for all ow

$$\begin{aligned} (s)g(t') &= \frac{f(\{u,v\} \cup \mathbf{y})}{f(\{v\} \cup \mathbf{y})} \frac{f(\{v\} \cup \mathbf{y})}{f(\mathbf{y})} \\ &= \frac{f(\{u,v\} \cup \mathbf{y})}{f(\{u\} \cup \mathbf{y})} \frac{f(\{u\} \cup \mathbf{y})}{f(\mathbf{y})} = g(s+t')g(0). \end{aligned}$$

$t' \to t$. Since $t' \leq t$, by left continuity $g(s+t)g(0) = g(s)g(t)$.

equently, by the Lemma, $g(s) = g(0)\gamma^s$, for some $\gamma > 0$. Now, $= g(0)\gamma^{\pi R^2} = g(\pi R^2)$, we obtain $g(s) = \beta\gamma^{s-\pi R^2}$ and hence

$$f(\{u\}) = f(\emptyset)g(0) = \alpha\beta\gamma^{-\pi R^2} = \alpha\beta\gamma^{-|U_{\{u\}}|}.$$

tion with respect to the number of points $n(\mathbf{x}) = n$,

$$\begin{aligned} f(\mathbf{x} \cup \{u\}) &= g(|B_u \cap U_\mathbf{x}|)f(\mathbf{x}) \\ &= g(0)\gamma^{|B_u \cap U_\mathbf{x}|}\alpha\beta^n\gamma^{-|U_\mathbf{x}|} \\ &= \alpha\beta^{n+1}\gamma^{-\pi R^2+|B_u \cap U_\mathbf{x}|-|U_\mathbf{x}|} \\ &= \alpha\beta^{n+1}\gamma^{-|U_{\mathbf{x}\cup\{u\}}|}. \end{aligned}$$

□

4 Relationship to a mixture model and Gibbs sampling

The penetrable sphere model and a particular case of the continuum random-cluster model are both related to Widom and Rowlinson's [35] mixture model as described below. So far this correspondence has mainly been used for studying phase transition behaviour; we return to this in Section 5. In the present section we shall investigate the relationships for the purpose of simulation.

Formally, let X and Y refer to the configurations of the two types. Then the joint density of (X, Y) at $(\mathbf{x}, \mathbf{y})$ with respect to the product measure of two independent unit rate Poisson processes is

$$f(\mathbf{x}, \mathbf{y}) = \alpha \beta_1^{n(\mathbf{x})} \beta_2^{n(\mathbf{y})} 1\{d(\mathbf{x}, \mathbf{y}) > R\} \tag{10}$$

where, as before, α is the normalising constant, $\beta_1, \beta_2 > 0$ are model parameters, and $d(\mathbf{x}, \mathbf{y})$ is the shortest distance between a point in $\mathbf{x}$ and a point in $\mathbf{y}$. In other words, only points of different types interact and they are not allowed to be within a distance R of each other. Hence the conditional distribution of X given Y is a homogeneous Poisson process on $A \setminus U_Y$ with intensity β_1 (see [35]). The marginal distributions of X and Y are area-interaction models (9) with interaction parameter $\gamma \geq 1$, as shown in [35].

It is interesting to note that there exists a simple, explicit relation between the normalising constants α_{mix} of the mixture model and α_1 of the marginal distribution of X as follows. Writing π for the distribution of a unit rate Poisson process on A and integrating out the second component,

$$\begin{aligned} f(\mathbf{x}) &= \int \alpha_{\text{mix}} \beta_1^{n(\mathbf{x})} \beta_2^{n(\mathbf{y})} 1\{d(\mathbf{x}, \mathbf{y}) > R\} d\pi(\mathbf{y}) \\ &= \int \alpha_{\text{mix}} \beta_1^{n(\mathbf{x})} \beta_2^{n(\mathbf{y})} 1\{\mathbf{y} \cap U_{\mathbf{x}} = \emptyset\} d\pi(\mathbf{y}) \\ &= \alpha_{\text{mix}} \beta_1^{n(\mathbf{x})} \exp[\beta_2 |A \setminus U_{\mathbf{x}}| - |A|]. \end{aligned}$$

Provided periodic boundary conditions are imposed on A (making A into a torus) and balls B_u are defined with respect to geodesic distance $||\cdot||$ or alternatively, $|U_{\mathbf{x}}|$ is defined as the Lebesgue measure of $U_{\mathbf{x}} \cap A$,

$$f(\mathbf{x}) = \alpha_1 \beta_1^{n(\mathbf{x})} (e^{\beta_2})^{-|U_{\mathbf{x}}|},$$

where $\alpha_1 = \alpha_{\rm mix} e^{-(1-\beta_2)|A|}$. Hence X is an area-interaction model with parameters $(\beta, \gamma) = (\beta_1, e^{\beta_2})$. Similarly, Y is an area-interaction model with parameters $(\beta, \gamma) = (\beta_2, e^{\beta_1})$ and normalising constant $\alpha_2 = \alpha_{\rm mix} e^{-(1-\beta_1)|A|}$.

For simulating the area-interaction model, it is simplest to apply MCMC methods on the mixture model (10) and marginalise, since this avoids calculation of the areas of overlapping balls in U_X. In particular, if $\beta = \log \gamma$ we can sample from both X and Y as they are identically distributed. Moreover, using the fact that the conditional distribution of Y given X is a Poisson process, moments of the marginal area-interaction model are easily obtained in terms of the mixture model as indicated in the following Lemma.

Lemma 2 *The moments of (10) and its marginals satisfy*

$$\mathbb{E}_1 |U_X| = |A| - \mathbb{E}_{\rm mix} \frac{n(Y)}{\beta_2} \tag{11}$$

$$\mathrm{Var}_1 |U_X| = (\mathrm{Var}_{\rm mix} n(Y) - \mathbb{E}_{\rm mix}\, n(Y))/\beta_2^2 \tag{12}$$

$$\mathrm{Cov}_1(n(X), |U_X|) = -\mathrm{Cov}_{\rm mix}(n(X), n(Y))/\beta_2 \tag{13}$$

In particular, $n(Y)$ is overdispersed. Similar expressions hold with the roles of X and Y exchanged. In the symmetric case $\beta_1 = \beta_2 = \beta$,

$$\mathbb{E}_1 |U_X| = \mathbb{E}_1 |U_Y| = |A| - \mathbb{E}_{\rm mix}[n(X) + n(Y)]/(2\beta)$$

and

$$\mathrm{Var}|U_X| = \mathrm{Var}|U_Y| = (\mathrm{Var}_{\rm mix} n(X) + \mathrm{Var}_{\rm mix} n(Y) - \mathbb{E}_{\rm mix}(n(X) + n(Y)))/(2\beta^2).$$

In the inhibitory case $\gamma < 1$ in (9) a mixture can be defined by its density

$$f(\mathbf{x}, \mathbf{y}) = \alpha \beta_1^{n(\mathbf{x})} \beta_2^{n(\mathbf{y})} \mathbf{1}\{\mathbf{y} \subset U_{\mathbf{x}}\}. \tag{14}$$

Then arguments similar to those for $\gamma > 1$ give the marginal density for X as

$$f(\mathbf{x}) = \alpha e^{-|A|} \beta_1^{n(\mathbf{x})} (e^{-\beta_2})^{-|U_{\mathbf{x}}|}, \tag{15}$$

again writing $|U_{\mathbf{x}}|$ for the Lebesgue measure of $U_{\mathbf{x}}$ either restricted to A or in the geodesic sense. Hence X (but not Y!) is an area-interaction model.

Turning attention to the superposition $Z = X \cup Y$, its density (with respect to a unit rate Poisson process on A) can be written as

$$\begin{aligned} f(\mathbf{z}) &= e^{-|A|} \sum f_{\text{mix}}(\mathbf{z}_1, \mathbf{z}_2) \\ &= e^{-|A|} \alpha_{\text{mix}} \sum \beta_1^{n(\mathbf{z}_1)} \beta_2^{n(\mathbf{z}_2)} \mathbf{1}\{d(\mathbf{z}_1, \mathbf{z}_2) > R\} \end{aligned}$$

where the sum is over all ordered partitions of $\mathbf{z}$ into two groups $\mathbf{z}_1$ and $\mathbf{z}_2$. In the symmetric case $\beta_1 = \beta_2 = \beta$ this reduces to

$$f(\mathbf{z}) = \alpha_{\text{mix}} e^{-|A|} \beta^{n(\mathbf{z})} N(\mathbf{z})$$

where $N(\mathbf{z})$ denotes the number of partitions such that

$$(\cup_{\eta \in \mathbf{z}_1} B(\eta, R/2)) \cap (\cup_{\zeta \in \mathbf{z}_2} B(\zeta, R/2)) = \emptyset.$$

Since $N(\mathbf{z}) = 2^{c(\mathbf{z})}$, the superposition density is a continuum random-cluster model with radii $R/2$ and parameters β and $\gamma = 1/2$. This relation is exploited in [12] in order to understand the phase transition of the mixture model in percolation terms.

This construction can be straightforwardly extended to the case of a multitype point process $(X^{(1)}, \ldots, X^{(k)})$ with density proportional to

$$\beta^{n(\mathbf{z})} \mathbf{1}\{d(\mathbf{x}^{(i)}, \mathbf{x}^{(j)}) > R, \forall i \neq j\} \tag{16}$$

for $\mathbf{z} = \mathbf{x}^{(1)} \cup \cdots \cup \mathbf{x}^{(k)}$. The superposition is then a continuum random-cluster model but with $\gamma = 1/k$. However, for $k \geq 3$, $X^{(i)}$ is not an area-interaction process.

The conditional distribution of one type of points given the others is very easy to sample from, so we opted for a Gibbs sampler approach. In the bivariate case this is given by altering between the conditional distribution of $X|Y$ and $Y|X$: initialise with any $\mathbf{y}_0$, e.g. from a Poisson process (rate β_2). Then the sweeps of the two-component Gibbs sampler are given by the following steps for each $k = 0, 1, \ldots$

1. $\mathbf{x}_{k+1} \sim \text{Poisson}(\beta_1)$ on $A \setminus U_{\mathbf{y}_k}$;
2. $\mathbf{y}_{k+1} \sim \text{Poisson}(\beta_2)$ on $A \setminus U_{\mathbf{x}_{k+1}}$.

The Poisson processes in steps 1 and 2 above are easily implemented by thinning a Poisson process in the whole window A (that is, just delete those points lying in $U_{\mathbf{y}_k}$ or $U_{\mathbf{x}_{k+1}}$). It is possible to specify a similar two-component Gibbs sampler for the inhibitory case of the area-interaction

process using (14)–(15) but the first step, where $X|Y = \mathbf{y}$ is Poisson(β_1) restricted to the event $H_{\mathbf{x},\mathbf{y}} = \{U_{\mathbf{x}} \supseteq \mathbf{y}\}$ may be too slow in practice, so it would be preferable to replace this step with a single point updating procedure. The other step, where $Y|X = \mathbf{x}$ is a Poisson(β_2) process on $U_{\mathbf{x}}$ is still easy to perform by thinning.

5 Exact simulation

We now combine Propp and Wilson's [26] ideas of exact simulation based on coupling Markov chains from the past with the two-component Gibbs sampler introduced in Section 4.

Our setup differs from that in [26] mainly in two respects. First, the state space in [26] is finite while ours is infinite. Second, our state space does not have any maximal (or minimal) element with respect to the partial order introduced below. The first difference turns out to be inconsequential, but the second requires special treatment.

Introduce a partial order $\preceq$ on the space $\Omega_A^2 = \Omega_A \times \Omega_A$ of mixed configurations by $(\mathbf{x},\mathbf{y}) \preceq (\mathbf{x}',\mathbf{y}')$ (or $(\mathbf{x}',\mathbf{y}') \succeq (\mathbf{x},\mathbf{y})$) if $\mathbf{x} \subseteq \mathbf{x}'$ and $\mathbf{y} \supseteq \mathbf{y}'$. In order to be able to adapt the Propp–Wilson ideas we verify first that the two-component Gibbs sampler respects the partial order $\preceq$. Let $\preceq_d$ denote stochastic domination with respect to $\preceq$, i.e. two Ω_A^2-valued random elements (X,Y) and (X',Y') satisfy $(X,Y) \preceq_d (X',Y')$ if there exists a coupling of (X,Y) and (X',Y') such that $(X,Y) \preceq (X',Y')$ a.s. (see [22] for a general discussion of coupling and stochastic domination).

Lemma 3 *Fix $(\mathbf{x},\mathbf{y}),(\mathbf{x}',\mathbf{y}') \in \Omega_A^2$ such that $(\mathbf{x},\mathbf{y}) \preceq (\mathbf{x}',\mathbf{y}')$. Let, for $i = 0,1,\ldots$, (X_i,Y_i) be the Ω_A^2-valued random element obtained by starting with $(X_0,Y_0) = (\mathbf{x},\mathbf{y})$ and running i iterations of the two-component Gibbs sampler in Section 4, and define (X_i',Y_i') similarly. Then $(X_i,Y_i) \preceq_d (X_i',Y_i')$ for all i.*

Proof: The case $i = 0$ is trivial. To prove the case $i = 1$ we consider the following coupling. Let $Z_{1,x}$ and $Z_{1,y}$ be two independent Poisson processes on A with rates β_1 and β_2, respectively. Let $X_1 = Z_{1,x} \setminus U_{Y_0}$ and $X_1' = Z_{1,x} \setminus U_{Y_0'}$, and then let $Y_1 = Z_{1,y} \setminus U_{X_1}$ and $Y_1' = Z_{1,y} \setminus U_{X_1'}$. Clearly, this gives the right marginal distributions of (X_1,Y_1) and (X_1',Y_1'). Since $Y_0 \supseteq Y_0'$ we get $U_{Y_0} \supseteq U_{Y_0'}$ whence $X_1 \subseteq X_1'$. This in turn implies $U_{X_1} \subseteq U_{X_1'}$ so that $Y_1 \supseteq Y_1'$, and we have $(X_1,Y_1) \preceq (X_1',Y_1')$. The cases

$i = 2, 3, \ldots$ follow similarly by induction. □

We call an element $(\mathbf{x}, \mathbf{y}) \in \Omega_A^2$ *quasimaximal* if $\mathbf{y} = \emptyset$ and $U_{\mathbf{x}} \supseteq A$. Similarly, $(\mathbf{x}, \mathbf{y})$ is called *quasiminimal* if $\mathbf{x} = \emptyset$ and $U_{\mathbf{y}} \supseteq A$.

Lemma 4 *Fix $(\mathbf{x}, \mathbf{y}) \in \Omega_A^2$, and for $i = 0, 1, \ldots$, let (X_i, Y_i) be the Ω_A^2-valued random element obtained by taking $(X_0, Y_0) = (\mathbf{x}, \mathbf{y})$ and running i iterations of the two-component Gibbs sampler. Define (X_i', Y_i') similarly, with an arbitrary distribution of (X_0', Y_0'). If $(\mathbf{x}, \mathbf{y})$ is quasimaximal, then*

$$(X_i, Y_i) \succeq_d (X_i', Y_i')$$

for all $i \geq 1$, while if $(\mathbf{x}, \mathbf{y})$ is quasiminimal, then

$$(X_i, Y_i) \preceq_d (X_i', Y_i')$$

for all $i \geq 1$.

Proof: We only need to prove the lemma for $i = 1$, as the general case follows using Lemma 3. However, the case $i = 1$ follows directly if we use the same coupling as in the proof of Lemma 3. □

We are now ready to describe the algorithm for exact simulation. For $i = 0, -1, -2, \ldots$, let $Z_{i,x}$ and $Z_{i,y}$ be independent Poisson processes on A with respective intensities β_1 and β_2. Let $k_1, k_2, \ldots$ be a strictly increasing sequence of positive integers, and let $(\mathbf{x}, \mathbf{y})$ and $(\mathbf{x}', \mathbf{y}')$ be fixed elements of Ω_A^2 such that $(\mathbf{x}, \mathbf{y})$ is quasiminimal and $(\mathbf{x}', \mathbf{y}')$ is quasimaximal. Then for $i = 1, 2, \ldots$ we generate two coupled Markov chains in accordance to the two-component Gibbs sampler by setting

$$({}^iX_{-k_i}, {}^iY_{-k_i}) = (\mathbf{x}, \mathbf{y}), \quad ({}^iX'_{-k_i}, {}^iY'_{-k_i}) = (\mathbf{x}', \mathbf{y}')$$

and for $j = 1, \ldots, k_i$,

$${}^iX_{j-k_i} = Z_{j-k_i,x} \setminus U_{{}^iY_{j-1-k_i}}, \quad {}^iY_{j-k_i} = Z_{j-k_i,y} \setminus U_{{}^iX_{j-k_i}}$$

$${}^iX'_{j-k_i} = Z_{j-k_i,x} \setminus U_{{}^iY'_{j-1-k_i}}, \quad {}^iY'_{j-k_i} = Z_{j-k_i,y} \setminus U_{{}^iX'_{j-k_i}}.$$

Clearly $({}^iX_{j-k_i}, {}^iY_{j-k_i}) \preceq ({}^iX'_{j-k_i}, {}^iY'_{j-k_i})$ for all $j = 0, \ldots, k_i$. Letting $I = \inf\{i \in \mathbb{N} : ({}^iX_0, {}^iY_0) = ({}^iX_0', {}^iY_0')\}$, (taking $\inf \emptyset = \infty$), we show below that $I < \infty$ a.s. We stop when $i = I$ and set $(X_0, Y_0) = ({}^IX_0, {}^IY_0)$.

Theorem 3 *The above algorithm terminates a.s., and the distribution of the obtained sample* (X_0, Y_0) *is given by (10).*

Proof : Note that for any $m \geq 0$ we have that if

$$Z_{-m-1,y} = \emptyset \tag{17}$$

then for any i such that $k_i \geq m$ we get $({}^iX_{-m}, {}^iY_{-m}) = ({}^iX'_{-m}, {}^iY'_{-m})$ and hence also $({}^iX_0, {}^iY_0) = ({}^iX'_0, {}^iY'_0)$. That $I < \infty$ a.s. now follows from the observation that with probability 1, (17) occurs for some m. Let M denote the smallest such m. Moreover, define Markov chains $({}^iX''_{j-k_i}, {}^iY''_{j-k_i}), j = 0, \ldots, k_i, i = 1, 2, \ldots$, in exactly the same way as the chains $({}^iX, {}^iY)$ and $({}^iX', {}^iY')$ except that $({}^iX''_{-k_i}, {}^iY''_{-k_i}) = (X, Y)$, where (X, Y) follows the stationary distribution (10). Then for $k_i \geq M$ we have that ${}^iX_{-M} = {}^iX'_{-M} = {}^iX''_{-M} = Z_{-M,x}$. It follows then from Lemmas 3 and 4 and the coupling construction that $(X_0, Y_0) = ({}^iX_0, {}^iY_0) = ({}^iX'_0, {}^iY'_0) = ({}^iX''_0, {}^iY''_0)$ when $k_i \geq M$, so $(X_0, Y_0) = \lim_{i\to\infty}({}^iX''_0, {}^iY''_0)$ a.s. Hence the distribution of (X_0, Y_0) is given by the stationary distribution (10). □

Propp and Wilson give an argument for preferring the sequence $k_i = 2^i$, and we have used the same in our simulation studies.

Fixing the dimension $d \geq 2$, let us say $d = 2$, it is known that phase transition behaviour occurs in the penetrable sphere model in the symmetric case $\beta = \log\gamma$, whenever β is very large, and does not occur when β is very small. This is a consequence of the phase transition behaviour in the mixture model (10), which was demonstrated by Ruelle [31] and later in [6] and [12] using percolation arguments analogous to the random-cluster derivation (see e.g. [5]) of the phase transition occurring in Ising and Potts models. Similarly as in the Ising/Potts models, phase transition means that infinite-volume limits fail to be unique. In particular, realisations tend, even for large systems, to be dominated by a single type of point (despite the symmetry of the model).

One would believe that the occurrence of phase transition is monotone in β, so that there exists a β_c such that

$$\beta \begin{cases} > \beta_c & \Rightarrow \text{ phase transition} \\ < \beta_c & \Rightarrow \text{ no phase transition} \end{cases}$$

but this is not known. Of course, this is a statement about infinite volume limits, but the effect should be visible already for moderately large A. A

measure of the amount of consensus is the fraction ϕ of pairs of points that are of the same type, i.e.

$$\phi = \frac{\binom{n_1}{2} + \binom{n_2}{2}}{\binom{n}{2}}$$

where n_i is the number of points of type i, and $n = n_1 + n_2$. Note that ϕ is undefined when $n = 0, 1$. In these cases, we will arbitrarily set $\phi = 0$. For large systems, one expects the following behaviour. When $\beta < \beta_c$, ϕ should be close to $\frac{1}{2}$ with high probability, it should start increasing rapidly as β approaches β_c, and tend to 1 as $\beta \to \infty$.

We have simulated independent samples from the symmetric mixture model using our exact simulation procedure for various values of β and R. We took $A = [0, 1]^2$ with the restricted Lebesgue measure.

Figure 1 depict simulated mean and quantiles of ϕ over 100 simulation against the canonical parameter $\theta = \log \beta$ of the mixture distribution (considered as an exponential family). It can be seen that for small values of θ, indeed the Monte Carlo approximation of $E_\theta \phi$ is close to $1/2$, increasing to 1 as θ increases. The rapid increase becomes more aparent when the interaction range is smaller - or equivalently, the window size is bigger.

The phase transition behaviour can also be observed from plots of the fraction $p = \max(n_1, n_2)/(n_1 + n_2)$ of points of the most frequent type. Plots of the Monte Carlo mean of p (Figure 1) look very similar to the plots for ϕ, and with increasing θ, the histograms of p (which we omit) become bimodal, due to the fact that realisations tend to consist predominantly of one type.

Examples of simulated realisations of point patterns for $\theta = \log(30)$, $\log(80)$ at radius $R = .2$ and for $\theta = \log(100), \log(200)$ at radius $R = .1$ can be seen in Figure 2. This figure also shows that for larger θ, one of the components dominates.

Finally, in Figure 3 we plotted Monte Carlo estimates of $E_\theta I = E_\theta \log k_I$, the expected number of steps until coalescence. The plot demonstrates the feasibility of exact simulation: for small to moderate θ, convergence is reached in only a few steps. However, as θ approaches phase transition, the number of steps needed to obtain coalescence increases rapidly.

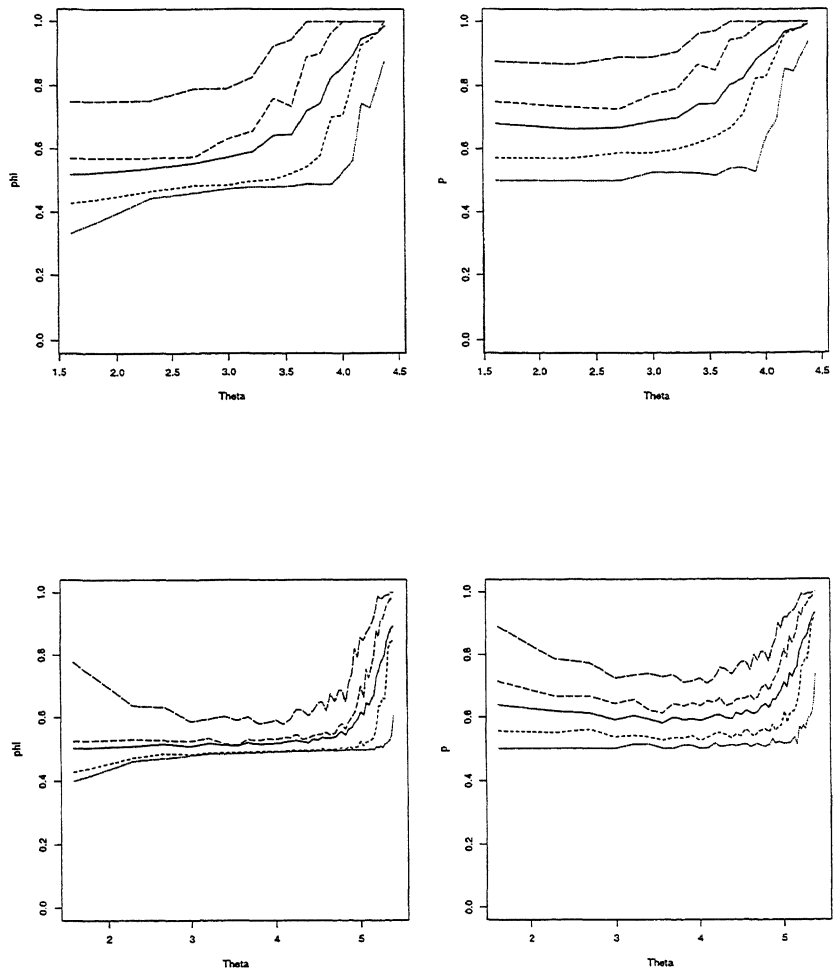


Figure 1: Monte Carlo estimates of the means of ϕ (left) and p (right) as a function of θ (solid line) at interaction radius $R = .2$ (top) and $R = .1$ (bottom). The dotted lines denote the .5, .25, .75 and .95 quantiles.

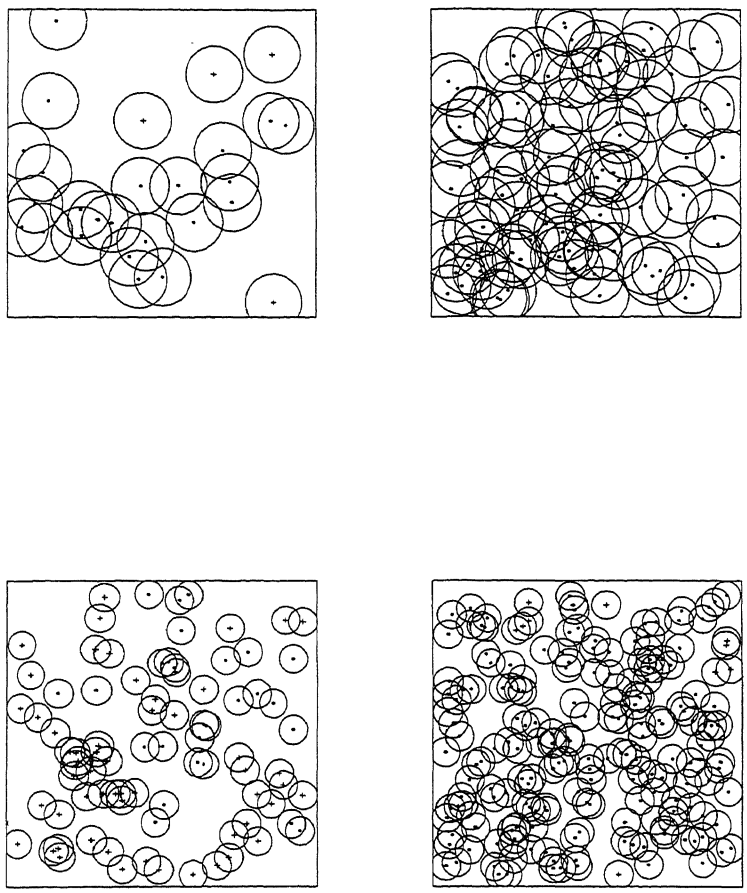
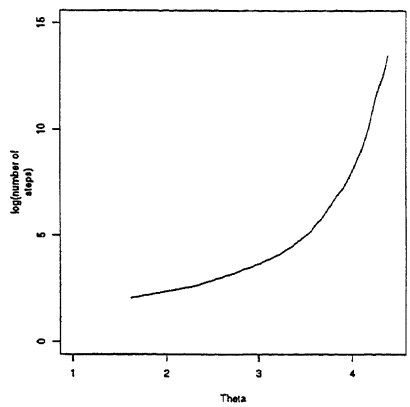
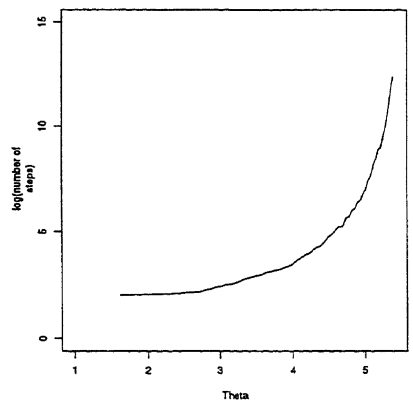


Figure 2: Realisation of the mixture model with $\beta = 30, R = .2$ (top left), $\beta = 80, R = .2$ (top right), $\beta = 100, R = .1$ (bottom left) and $\beta = 200, R = .1$ (bottom right).





3: Monte Carlo estimates of $\log k_I$ as a function of θ at interaction $R = .2$ (top) and $R = .1$ (bottom).

6 A Swendsen–Wang type algorithm

In this Section we present an algorithm which is similar to one of Swendsen and Wang [34]. This algorithm has independently been discovered by Chayes and Machta [7].

The algorithm works for the important symmetric case $\beta_1 = \beta_2 = \beta$ of the mixture model (and hence also for the area-interaction model with $\gamma = e^\beta$ and for the continuum random-cluster process with $\gamma = \frac{1}{2}$).

Initialising with any $(\mathbf{x}, \mathbf{y})$, an iteration of the algorithm consists of

1. choosing a new value of $(\mathbf{x}, \mathbf{y})$ according to its conditional distribution given $\mathbf{x} \cup \mathbf{y}$, i.e. flipping a fair coin independently for each connected component of $U_{\mathbf{x}\cup\mathbf{y}}$ to determine whether the points should be of the first or the second type;
2. replacing $\mathbf{y}$ by a Poisson (β) process on $A \setminus U_{\mathbf{x}}$.

It is immediate that the mixture measure given by (10) is invariant under step 1 of this algorithm, and we have already seen in Section 4 that it is invariant under step 2. The algorithm can also be extended to the case of a multitype point process (16), and also to the $\gamma < 1$ case of the continuum random-cluster model, even when γ^{-1} is not an integer. In the latter case, the algorithm goes as follows. First paint the points of each connected component red with probability γ (independently for different connected components), and then replace all the red connected components by a Poisson (β) process on the part of A not occupied by the remaining connected components.

We believe (although we have no rigorous justification) that this algorithm approaches stationarity much faster and mixes better than the two-component Gibbs sampler when β is large (i.e. in the phase transition region of the parameter space). The reason should be that when the Gibbs sampler starts with a configuration with mostly points of the first (or second) type, then it will tend to stay in this state for an astronomical amount of time provided that β is large and A is large compared to a disc with radius R, while the Swendsen–Wang type algorithm will jump back and forth between the two states. One might suggest that the slowness of the Gibbs sampler could be solved by allowing $\mathbf{x}$ and $\mathbf{y}$ to change places occasionally, but this is presumably not the case, the reason being the following. Suppose β is large and $A = [-M, M]^2$, where $M >> R$, and we start with a "dense" crowd of points of type 1 in $[-M, M] \times [-M, 0]$ and a

similar crowd of points of type 2 in $[-M, M]\times[0, M]$. This highly improbable type of configuration will remain for a long time using the modified Gibbs sampler (although the two types will sometimes interchange regions), while on the other hand the Swendsen–Wang type algorithm does not seem to exhibit such a phenomenon.

It would be very nice if the Swendsen–Wang type algorithm could be combined with the ideas of the previous section in order to obtain exact samples, but unfortunately it seems very difficult to find any useful monotonicity property of the algorithm. Propp and Wilson [26] make a similar remark about the original Swendsen–Wang algorithm. They are still able to obtain exact samples in reasonable time also in the phase transition regime by simulating the Fortuin–Kasteleyn random-cluster model rather than the Ising/Potts models directly. The corresponding thing to do here would be first to simulate the continuum random-cluster model and then to update the components as in step 1 in our Swendsen–Wang type algorithm but, in the absence of any simple monotonicity relation in the continuum random-cluster model, we cannot see any suitable way of doing it.

Acknowledgements

This work was initiated when the second author visited the third author at the Department of Theoretical Statistics in Aarhus, and a similar visit from the first author at the Department of Mathematics and Computer Science in Aalborg; the hospitality and support from these departments is gratefully acknowledged. The research has been funded by the Danish Informatics Network in Agricultural Sciences, the Danish Natural Science Research Council, the Dutch Mathematical Research Institute, the Swedish Natural Science Research Council and grant SCI/180/94/103 (Applications of stochastic geometry in the analysis of spatial data) of the Nuffield foundation.

References

[1] A.J. Baddeley and M.N.M. van Lieshout. Area-interaction point processes. *Annals of the Institute of Statistical Mathematics* 46:601–619, 1995.

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

[3] A.J. Baddeley, M.N.M. van Lieshout and J. Møller. Markov properties of cluster processes. *Advances in Applied Probability (SGSA)* 28:346–355, 1996.

[4] J.E. Besag and P.J. Green. Spatial statistics and Bayesian computation (with discussion). *Journal of the Royal Statistical Society, Series B* 55:25–37, 1993.

[5] C. Borgs and J.T. Chayes. The covariance matrix of the Potts model: a random cluster analysis. *Journal of Statistical Physics* 82:1235–1297, 1996.

[6] J.T. Chayes, L. Chayes and R. Kotecky. The analysis of the Widom–Rowlinson model by stochastic geometric methods, *Communications in Mathematical Physics* 172:551–569, 1995.

[7] L. Chayes and J. Machta. Graphical representations and cluster algorithms part II. In preparation.

[8] D.J. Daley and D. Vere-Jones. *An introduction to the theory of point processes*. Springer, New York, 1988.

[9] P.J. Diggle, T. Fiksel, P. Grabarnik, Y. Ogata, D. Stoyan and M. Tanemura. On parameter estimation for pairwise interaction processes. *International Statistical Review* 62:99–117, 1994.

[10] C.M. Fortuin and P.W. Kasteleyn. On the random-cluster model. I. Introduction and relation to other models. *Physica* 57:535–559, 1972.

[11] D.J. Gates and M. Westcott. Clustering estimates for spatial point distributions with unstable potentials. *Annals of the Institute of Statistical Mathematics* 38:123–135, 1986.

[12] H.-O. Georgii and O. Häggström. Phase transition in continuum Potts models. *Communications in Mathematical Physics*, 181:507-528, 1996.

[13] C.J. Geyer. Likelihood inference for spatial point processes. In: *Proceedings Seminaire Européen de Statistique, "Stochastic geometry, likelihood, and computation"*. O. Barndorff–Nielsen, W.S. Kendall and M.N.M. van Lieshout (Eds), Chapman and Hall, 1996. To appear.

[14] C.J. Geyer and J. Møller. Simulation procedures and likelihood inference for spatial point processes. *Scandinavian Journal of Statistics* 21:359–373, 1994.

[15] G. Grimmett. The stochastic random-cluster process, and the uniqueness of random-cluster measures. *Annals of Probability* 23:1461–1510, 1995.

[16] 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.

[17] O. Kallenberg. An informal guide to the theory of conditioning in point processes. *International Statistical Review* 52:151–164, 1984.

[18] F.P. Kelly and B.D. Ripley. On Strauss's model for clustering. *Biometrika* 63:357–360, 1976.

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

[20] W.S. Kendall and J. Møller. Perfect Metropolis-Hastings simulation of locally stable spatial point processes. (In preparation)

[21] W. Klein. Potts-model formulation of continuum percolation. *Physical Review* B26:2677–2678, 1982.

[22] T. Lindvall. *Lectures on the Coupling Method*. Wiley, New York, 1992.

[23] J. Møller. Discussion contribution. *Scandinavian Journal of Statistics* 21:346–349, 1994.

[24] J. Møller. Markov chain Monte Carlo and spatial point processes. In: *Proceedings Seminaire Européen de Statistique, "Stochastic geometry, likelihood, and computation"*. O. Barndorff–Nielsen, W.S. Kendall and M.N.M. van Lieshout (Eds), Chapman and Hall, 1996. To appear.

[25] J. Møller and R. Waagepetersen. Markov connected component fields. Research Report 96-2009, Department of Mathematics and

Computer Science, Aalborg University, 1996. To appear in *Advances in Applied Probability (SGSA)*.

[26] J.G. Propp and D.B. Wilson. Exact sampling with coupled Markov chains and applications to statistical mechanics. *Random Structures and Algorithms*, 9:223–252, 1996.

[27] B.D. Ripley and F.P. Kelly. Markov point processes. *Journal of the London Mathematical Society* 15:188–192, 1977.

[28] S. Ross. *A first course in probability (second edition)*. Collier MacMillan Publishers, London, 1984.

[29] J.S. Rowlinson. Penetrable sphere models of liquid-vapor equilibrium. *Advances in Chemical Physics*, 41:1–57, 1980.

[30] J.S. Rowlinson. Probability densities for some one-dimensional problems in statistical mechanics. In G.R. Grimmett and D.J.A. Welsh, editors, *Disorder in physical systems*, pages 261–276, Clarendon Press, Oxford, 1990.

[31] D. Ruelle. Existence of a phase transition in a continuous classical system. *Physical Review Letters* 27:1040–1041, 1971.

[32] A.D. Sokal. Monte Carlo methods in statistical mechanics: foundations and new algorithms. In *Cours de Troisième Cycle de la Physique en Suisse Romande*, Lausanne, 1989.

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

[34] R.H. Swendsen and J.-S. Wang. Nonuniversal critical dynamics in Monte Carlo simulations. *Physical Review Letters* 58:86–88, 1987.

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