Spatial Jump Processes and Perfect Simulation *

Kasper K. Berthelsen and Jesper Møller

Department of Mathematical Sciences, Aalborg University, Fredrik Bajers Vej 7G
DK-9220 Aalborg, Denmark

Abstract. Spatial birth-and-death processes, spatial birth-and-catastrophe processes, and more general types of spatial jump processes are studied in detail. Particularly, various kinds of coupling constructions are considered, leading to some known and some new perfect simulation procedures for the equilibrium distributions of different types of spatial jump processes. These equilibrium distributions include many classical Gibbs point process models and a new class of models for spatial point processes introduced in the text.

1 Introduction

Many dynamic systems of interacting "objects" evolving in time can be described by a stochastic process X_t, where $t \geq 0$ is the time. Often the objects can be viewed as points living in some appropriate space S, so that $X_t \subseteq S$ for all times t. For example, S could be a planar or spatial region; or a space for some geometric objects like line segments or discs. One class of models for such stochastic processes is spatial jump processes.

In this contribution we consider the case of finite spatial jump processes, i.e. when X_t is a finite subset of S. Recall that in general a jump process is a continuous-time Markov process $X = \{X_t : t \geq 0\}$ with piecewise constant sample path (Feller 1971). More precisely, let x and y be finite subsets of a given arbitrary space S. Conditionally on the current state $X_t = x$ and the previous history $\{X_s : s < t\}$, the waiting time τ_x to the next jump (i.e. X_s stays in x for $t \leq s < t + \tau_x$, and $X_{t+\tau_x} \neq x$) is exponentially distributed and depends only on x. Furthermore, conditionally on both $X_t = x$, $\{X_s : s < t\}$, and τ_x, the next state $X_{t+\tau_x} = y$ follows some probability measure

* Submitted as a chapter in the proceedings for the second conference on *Spatial Statistics and Statistical Physics*, eds. Klaus R. Mecke and Dietrich Stoyan, Lecture Notes in Physics, Springer-Verlag

P_x which only depends on x so that either $x \subset y$ or $x \supset y$. In the special case where each jump consists in adding/deleting exactly one point, we have a spatial birth-and-death process (Preston 1977).

Spatial jump processes are interesting for several reasons. They provide a large class of models for spatio-temporal processes, which e.g. may describe many interacting particle systems studied in physics. As we shall demonstrate, they also provide mathematical tractable models for spatio-temporal processes. Specificly, we consider different coupling techniques which allow us to study the ergodic behaviour and characterise the equilibrium distribution for a large class of models in the reversible case as well as in the irreversible case. Furthermore, due to certain thinning techniques and since the sample paths are piecewise constant, spatial jump processes can often easily be simulated on a computer. Particularly, we extend the ideas of Propp & Wilson (1996) and use the thinning techniques for various perfect (or exact) simulation algorithms based on dominated coupling from the past (Kendall 1998, Kendall & Møller 2000). As in Kendall & Møller (2000), we prefer the term "perfect simulation" instead of "exact simulation", since random number generators always have defects and an algorithm may fail to deliver a simulation within practical constraints of time.

Currently perfect simulation is a hot research topic in statistics and applied probability, and as such the perfect simulation part (Sections 3.2 and 4.2) is of its own interest. Perfect simulation techniques have proven to be particular useful in statistical physics, spatial statistics, and stochastic geometry. A recent review on perfect simulation in stochastic geometry is given in Møller (2001). Häggström, van Lieshout & Møller (1999) and Georgii (2000) deal with perfect simulation and some aspects of phase transition in the Widom & Rowlinson (1970) model and related models. For a survey on the historic development of perfect simulation and a comprehensive list of references, see `http://www.dbwilson.com/exact`.

In this contribution we dicuss various constructions of spatial jump processes, study the ergodicity properties of these processes, characterise their equilibrium distributions, and show how to start spatial jump processes in equilibrium by the use of perfect simulation techniques. Readers interested in a further discussion of the statistical and computational aspects (using perfect simulations) are

referred to Berthelsen & Møller (2001). Furthermore, many of the ideas in the perfect simulation part of the text apply to Gibbs sampling (also known in statistical physics as the heat-bath algorithm) and other Metropolis-Hastings algorithms for spatial point processes (Häggström et al. 1999, Thönnes 1999, Møller & Schladitz 1999, Møller 2001, Kendall & Møller 2000).

The text is organized as follows. Section 2 provides some background material on simple jump processes (i.e. when the position of points is ignored), which becomes useful when we later couple simple jump processes with spatial jump processes. Section 3 concerns spatial birth-and-death processes and a description and comparison of the perfect simulation algorithms in Kendall & Møller (2000) and Fernández, Ferrari & Garcia (1999). Section 4 concerns spatial birth-and-catastrophe processes, i.e. when a jump consists in either adding a new point or deleting all existing points. Finally, Section 5 comments on more general cases of spatial jump processes.

2 Simple jump processes

This section is a short diversion into simple jump processes as they will eventually control everything which goes on in this text. See for example Norris (1997) and Asmussen (1987) for more details.

A simple jump process $N = \{N_t : t \geq 0\}$ is a continuous time Markov process with state space $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$ and piecewise constant right-continuous paths. It can formally be constructed as follows. Let $J_1 < J_2 < J_3 \ldots$ be the times at which N makes a jump, and set $J_0 = 0$. Further, for all $n \in \mathbb{N} = \{1, 2, \ldots\}$, let

$$H_n = \begin{cases} J_n - J_{n-1} & \text{if } J_n < \infty \\ \infty & \text{otherwise} \end{cases}$$

be the holding times, and set $M_n = N_{J_n}$. The discrete time Markov chain $M = \{M_n : n \in \mathbb{N}_0\}$ is called the jump chain (or the embedded Markov chain). For any $n \in \mathbb{N}$ and any $i \in \mathbb{N}_0$, the conditional distribution of H_n given $H_1, \ldots, H_{n-1}, M_0, \ldots, M_{n-1} = i$ is $\text{Exp}(q_i)$, the exponential distribution with mean $1/q_i$, where $q_i > 0$ is a given parameter. Furthermore, for any $j \in \mathbb{N}_0 \backslash \{i\}$,

$$\mathbb{P}(M_n = j | H_1, \ldots, H_n, M_0, \ldots, M_{n-1} = i) = q_{ij}/q_i,$$

where the $q_{ij} \geq 0$ are given parameters so that $q_i = \sum_{j \neq i} q_{ij}$. The matrix $Q = \{q_{ij} : i, j \in \mathbb{N}_0\}$ with $q_{ii} = -q_i$ is called the generator of N.

In other words, conditionally on $N_t = i$, the waiting time τ_i to the next jump is independent of the previous history $\{N_s : s < t\}$, and $\tau_i \sim \text{Exp}(q_i)$. Moreover, conditionally on both $N_t = i$, $\{N_s : s < t + \tau_i\}$, and τ_i, we have that $N_{t+\tau_i}$ jumps to j with probability q_{ij}/q_i for $j \neq i$.

In the sequel we impose the following technical conditions, which are commented below.

Conditions:
(i) $\pi = \{\pi_i : i \in \mathbb{N}_0\}$ is a given probability density function where the support $I = \{i : \pi_i > 0\}$ is given by either $I = \mathbb{N}_0$ or $I = \{0, \ldots, l\}$ for some $l \in \mathbb{N}_0$.
(ii) Q is irreducible, i.e. for all distinct $i, j \in I$ exist $i_0, i_1, \ldots i_n \in I$ such that $i_0 = i$, $i_n = j$, and $q_{i_0 i_1} \ldots q_{i_{n-1} i_n} > 0$.
(iii) π is invariant for Q, i.e. $\pi Q = 0$.
(iv) Q is non-explosive, i.e. $\mathbb{P}(\sum_n H_n < \infty | N_0 = i) = 0$ for all $i \in I$; see Fig. 1.

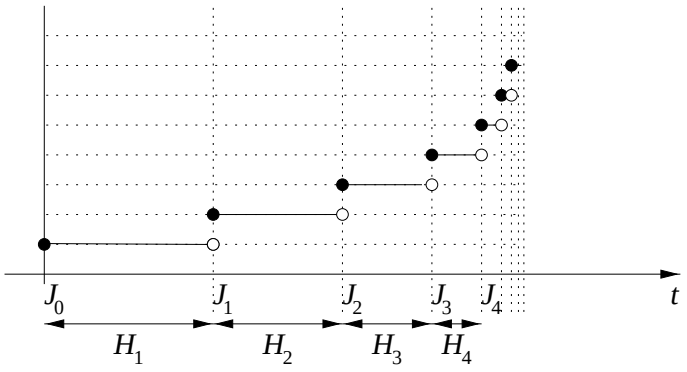

Fig. 1. Example of an explosive simple jump process.

Remarks:
(a) Condition (i) is a kind of hereditary condition on π. It is equivalent to assuming that the support is of the form $I = \{0, 1, \ldots, l\}$ or $I = \mathbb{N}_0$.

(b) Conditions (ii) and (iii) imply that π is the unique invariant distribution for Q.
(c) Sometimes we assume that Q and π are in detailed balance, i.e.

$$\pi_i q_{ij} = \pi_j q_{ji} \text{ for all } i, j \in \mathbb{N}_0. \tag{1}$$

(d) Detailed balance implies that π is invariant for Q. Combining (ii), (iv), and (1) we obtain reversibility of N.
(e) Condition (iv) holds if and only if the only bounded solution to $Qk = k$ for column vectors $k = (k_0, k_1, \ldots)^{\mathrm{T}}$ is $k = (0, 0, \ldots)^{\mathrm{T}}$ (see e.g. Asmussen (1987)). A sufficient condition for this is $\sup_{i \in I} q_i < \infty$.
(f) Conditions (ii) and (iv) imply positive recurrence, i.e. for any $i \in I$, the time between two consecutive occurrences of i in N has finite mean.
(g) Conditions (i)–(iv) imply convergence towards π, see e.g. Theorem 3.6.2 in Norris (1997) or Theorem II.4.6 in Asmussen (1987).
(h) For later purposes we notice that the jump chain has invariant dd density $\pi'_i \propto q_i \pi_i$, and it is positive recurrent (by Theorems 3.4.1 and 3.5.3 in Norris (1997), the jump chain is recurrent; and recurrence implies positive recurrence as the support I is countable). Furthermore, let L be the number of jumps in a cycle of the jump chain, i.e. the number of jumps between two successive zeros in the jump chain. Then by Kac's theorem (Meyn & Tweedie 1994),

$$\mathbb{E}L = 1/\pi'_0. \tag{2}$$

3 Spatial birth-and-death processes

3.1 General properties based on coupling constructions

This section considers finite spatial birth-and-death processes defined on some rather arbitrary space S, cf. Section 3.1.2. Section 3.1.1 concerns the simplest case where S is a singleton; this case corresponds to a simple jump process with jumps of the type $i \to i \pm 1$. The general case is described in Section 3.1.2, using a coupling to a

dominating simple birth-and-death process. Section 3.1.3 discusses a detailed balance condition and an extension of the concept of local stability (Ruelle 1969). Section 3.1.4 considers how to construct in an easy way a spatial birth-and-death process by thinning from a simpler dominating spatial birth-and-death process, extending ideas in Kendall (1998) and Kendall & Møller (2000). This coupling construction will also be used in Section 3.2 for perfect simulation purposes.

3.1.1 The simple case

A simple jump process with generator Q is a simple birth-and-death process if $q_{ij} = 0$ whenever $|i-j| > 1$. For convenience set $b_i = q_{i,i+1}$ and $d_{i+1} = q_{i+1,i}$ for $i \in \mathbb{N}_0$. Throughout this section we assume that

$$\pi_i b_i = \pi_{i+1} d_{i+1} > 0 \quad \text{whenever } \pi_{i+1} > 0. \tag{3}$$

This assumption is equivalent to the conditions (i)–(iii) and the detailed balance condition (1) in Section 2. Regarding the final condition (iv) in Section 2, by Reuter & Ledermann (1953), a simple birth-and-death process is non-explosive if

$$\sum_{i \in I} \frac{1}{b_i} = \infty. \tag{4}$$

In the sequel we often refer to the following Examples 1–3.

Example 1: Consider a standard immigration birth-and-death process, i.e. $b_i \equiv \beta > 0$ and $d_{i+1} = i+1$ for $i \in \mathbb{N}_0$. Clearly, (3) and (4) are satisfied, and

$$\pi_i = e^{-\beta} \beta^i / i! \tag{5}$$

is Poisson. By (h) in Section 2, $\pi'_{i+1} \propto \pi_{i+1} + \pi_i$. Hence by (2), $\mathbb{E}L = 2e^{\beta}$.

Example 2: For $i \in \mathbb{N}_0$, let $b_i = p$ and $d_{i+1} = 1-p$, where $0 < p < \frac{1}{2}$. Then (3) and (4) are satisfied, and

$$\pi_i = (1-q)q^i \tag{6}$$

is geometric with parameter $q = p/(1-p)$. The corresponding jump chain is a random walk with reflecting barrier at 0. Using (h) and (2), we obtain that $\mathbb{E}L = 2q/(1-q)$.

Example 3: In this example we start by specifying π and then determine the birth and death rates such that (3) and (4) hold. Assume that

$$\pi_i = \frac{\Gamma(i+\alpha)}{\Gamma(\alpha)i!}(1-q)^\alpha q^i \tag{7}$$

is negative binomial, where $\alpha > 0$ and $0 < q < 1$. Letting $q = p/(1-p)$, one solution to (3) is an extension of Example 2 such that

$$b_i = p(i+\alpha)/(i+1) \quad \text{and} \quad d_{i+1} = 1-p.$$

Then (4) is seen to hold.

3.1.2 The spatial case

In the spatial case we consider stochastic processes $X = \{X_t : t \geq 0\}$ defined on a metric space S with metric $d(\cdot,\cdot)$. We equip S with the Borel σ-algebra $\mathcal{B}$ and let λ denote an arbitrary diffuse probability measure on $\mathcal{B}$ (here "diffuse" means that $\lambda(\{\xi\}) = 0$ for all $\xi \in S$). For simplicity we consider X_t to be a finite subset of S, though everything in the sequel easily extend to the case where X_t is allowed to have multiple points (then λ is not necessarily diffuse). For any finite $x \subseteq S$, let $n(x)$ denote the number of points in x. The state space of X_t is $\Omega = \bigcup_{i=0}^{\infty} \Omega_i$, where $\Omega_i = \{x \subseteq S : n(x) = i\}$ is the set of finite point configurations of cardinality i; note that $\Omega_0 = \{\emptyset\}$ where $\emptyset$ denotes the empty point configuration. We equip Ω with the smallest σ-algebra $\mathcal{F}$ making the mappings $n_B(x) = n(x \cap B)$ measurable for all bounded $B \in \mathcal{B}$. For technical reasons S is assumed to be a Polish space (i.e. every Cauchy sequence is convergent, and S admits a countable dense set). Then Ω is a Polish space with respect to the Hausdorff (or Prohorov) metric, and its Borel σ-algebra is equal to $\mathcal{F}$; for details, see Daley & Vere-Jones (1988). Finally, for $x \in \Omega$ and $\xi \in S$, we write $x\backslash\xi$ for $x\backslash\{\xi\}$, and $x \cup \xi$ for $x \cup \{\xi\}$.

Now, consider a spatial birth-and-death process X. We specify this by two measurable functions $\beta, \delta : \Omega \times S \to [0,\infty)$. For any $x \in \Omega$, $\beta(x,\xi)$ is the birth rate at which a point $\xi \in S$ is added to x, and if $x \neq \emptyset$, $\delta(x\backslash\eta,\eta)$ is the death rate at which a point $\eta \in x$ is deleted from x. Roughly speaking, for an infinitesimal time interval

$[t, t+dt]$ and an infinitesimal ball $d\xi$ with centre ξ,

$$\begin{aligned}
&\mathbb{P}(X_{t+dt} \setminus x \text{ is a point in } d\xi \,|X_t = x) \approx \beta(x,\xi)\lambda(d\xi)dt,\\
&\mathbb{P}(X_{t+dt} = x\backslash\eta \,|X_t = x) \approx \delta(x\backslash\eta, \eta)dt,\\
&\mathbb{P}(\text{more than one jump in } [t, t+dt]|X_t = x) \approx 0.
\end{aligned}$$

More precisely, define in a similar way as in Section 2, the jump times $J_1 < J_2 < \ldots$, the holding times $H_1, H_2, \ldots$, and the jump chain $Y = \{Y_n : n \in \mathbb{N}_0\}$ for X, where $Y_0 = X_0$. Further, for $x \in \Omega$, let $B(x) = \int_S \beta(x,\xi)\lambda(d\xi)$ and $D(x) = \sum_{\eta\in x} \delta(x\backslash\eta, \eta)$ be the total birth and death rates, assuming $B(x) < \infty$ and setting $D(\emptyset) = 0$. Then the conditional distribution of H_n given $H_1, \ldots, H_{n-1}, Y_0, \ldots, Y_{n-1} = x$ is $\mathrm{Exp}\,(B(x) + D(x))$. If we also condition on H_n, the probability for a birth at time J_n is $B(x)/(B(x) + D(x))$. If we further condition on that a birth happens at time J_n, then $Y_n = x \cup \xi$, where ξ has density $\beta(x,\cdot)/B(x)$ with respect to λ. If we instead condition on that a death happens at time J_n, then $Y_n = x\backslash\eta$, where $\eta \in x$ is selected with probability $\delta(x\backslash\eta, \eta)/D(x)$.

It remains to clarify if such a spatial birth-and-death process is non-explosive. Given a generator Q for a non-eplosive simple birth-and-death process N as in Section 3.1, we assume henceforth that

$$\beta(x,\xi) = b_{n(x)}\bar{b}(x,\xi) \quad \text{and} \quad \delta(x,\xi) = d_{n(x)+1}\bar{d}(x,\xi)/(n(x)+1) \tag{8}$$

where

$$0 \le \bar{b} \le 1 \quad \text{and} \quad \bar{d} \ge 1 \tag{9}$$

are measurable functions. By (8) and (9),

$$B(x) \le b_{n(x)} \quad \text{and} \quad D(x) \ge d_{n(x)}. \tag{10}$$

It is then possible to couple N and X so that

$$N_t \ge n(X_t) \quad \text{for all } t \ge 0 \tag{11}$$

if $N_0 \ge n(X_0)$. Briefly, (X, N) is a jump process where X and N evolve independently of each other as long as $N_t > n(X_t)$, while if $N_t = n(x) = n$ and $X_t = x$, (X_t, N_t) jumps to (x', n') with a rate specified as follows:

$$\beta(x,\xi) \;\text{ for } (x', n') = (x \cup \xi, n+1),$$

$$b_n - B(x) \quad \text{for } (x', n') = (x, n+1),$$
$$d_n \delta(x \backslash \eta, \eta)/D(x) \quad \text{for } (x', n') = (x \backslash \eta, n-1),$$
$$(D(x) - d_n)\delta(x \backslash \eta, \eta)/D(x) \quad \text{for } (x', n') = (x \backslash \eta, n).$$

By (10) and (11), since Q is non-explosive, it is intuitively clear that X is non-explosive. A formal proof can be found in Preston (1977).

Combining (f) in Section 2, (11), and the renewal theorem (see e.g. Theorem V.1.2 in Asmussen (1987)), we obtain the following results. Write X_{t-} for the state of X just before time t, let $\mathcal{T} = \inf\{t > 0 : X_{t-} \neq \emptyset, X_t = \emptyset\}$ be the return time to $\emptyset$ when $X_0 = \emptyset$, and set $\mu = \mathbb{E}(\mathcal{T} | X_0 = \emptyset)$. Then $\emptyset$ is an ergodic atom, i.e. $\mu < \infty$. Hence X jumps to $\emptyset$ infinite often, and at each such jump time X regenerates (i.e. X at such a time is independent of its past history). As $t \to \infty$, X_t converges in distribution towards an equilibrium distribution κ, say, independent of the initial state of X_0, i.e. for all bounded functions $f : \Omega \to \mathbb{R}$ which are continuous almost everywhere with respect to κ ,

$$\mathbb{E}[f(X_t)|X_0] \to \int f(x)\kappa(dx) \quad \text{as } t \to \infty. \tag{12}$$

Further,

$$\mu \int f(x)\kappa(dx) = \mathbb{E}\left[\int_0^{\mathcal{T}} f(X_t)dt \middle| X_0 = \emptyset\right], \tag{13}$$

and the support of κ is given by

$$\text{supp}(\kappa) = \Omega_0 \cup \big\{\{x_1, \ldots, x_n\} \in \Omega : n \in \mathbb{N} \text{ and } \beta(\{x_1, \ldots, x_{i-1}\}, x_i) > 0 \text{ for } i = 1, \ldots, n\big\}. \tag{14}$$

Finally, we often have geometrically fast convergence towards κ, cf. Møller (1989). For instance, this is the case when b_i and d_i are specified as in Examples 1–3.

3.1.3 Detailed balance and local stability conditions

Let the situation be as in Section 3.1.2. In this section we discuss a detailed balance condition which characterises κ in (13). Furthermore, we motivate an extension of the important concept of local stability (Ruelle 1969, Geyer 1999, Kendall & Møller 2000).

In order to specify the invariant distribution κ, we assume that ϕ is an unnormalised density with respect to the probability measure

$$\nu = \sum_{i=0}^{\infty} \pi_i \lambda_i \tag{15}$$

where π is given by (3), λ_0 is the probability measure concentrated at the empty point configuration, and λ_i $(i \geq 1)$ denotes the distribution for a binomial point process given by i independent points in S with common distribution λ. If π is a Poisson distribution with mean β, then ν is simply a Poisson point process with intensity measure $\beta\lambda$. Note that for a point process $Z \sim \phi$, conditionally on $n(Z) = i$, Z has unnormalised density ϕ with respect to λ_i, i.e. no matter the choice of π.

Now, as in Ripley (1977), we obtain that the detailed balance condition,

$$\bar{b}(x, \xi)\phi(x) = \bar{d}(x, \xi)\phi(x \cup \xi) > 0 \quad \text{whenever } \phi(x \cup \xi) > 0 \tag{16}$$

ensures that κ has a density proportional to ϕ, and X is time reversible with respect to κ (Ripley (1977) assumes π to be Poisson, but we do not need to make this restriction). Note that (9) and (16) always imply that ϕ is non-increasing,

$$\phi(x \cup \xi) \leq \phi(x) \tag{17}$$

which in turn implies integrability of ϕ with respect to ν no matter the choice of π. We refer to (17) as local stability of κ with respect to ν for the following reasons.

Geyer (1999) and Kendall & Møller (2000) consider the common case (5) where π is Poisson and

$$\phi(x \cup \xi) \leq K\phi(x), \quad \bar{d} \equiv 1, \quad \bar{b}(x, \xi) = \begin{cases} \phi(x \cup \xi)/\phi(x) & \text{if } \phi(x) \neq 0 \\ 0 & \text{otherwise} \end{cases} \tag{18}$$

for some constant K (independent of x and ξ). Then points are dying with uniform rate 1, and $\bar{b}$ is the Papangelou conditional intensity (Kallenberg 1984, Daley & Vere-Jones 1988). Note that (18) implies (16), and the existence of the uniform upper bound K on $\bar{b}$

is local stability in the sense of Ruelle (1969). As we can replace β in (5) by $\beta' = \beta K$, and ϕ in (18) by $\phi'(x) = \phi(x)/K^{n(x)}$, we can without loss of generality assume $K = 1$. Then (9) and (17) are satisfied.

On the other hand, if π is not Poisson, then in general we do not have Ruelle local stability. For instance, let π be given by (6), and let $\tilde{\nu} = \sum_i \tilde{\pi}_i \lambda_i$ where $\tilde{\pi}_i = e^{-1}/i!$ is Poisson. Then $\tilde{\phi}(x) = \phi(x) q^{n(x)} n(x)!$ is the corresponding unnormalised density with respect to $\tilde{\nu}$, but (17) does not ensure that

$$\tilde{\phi}(x \cup \xi)/\tilde{\phi}(x) = (\phi(x \cup \xi)/\phi(x)) q(n(x) + 1)$$

is bounded by a constant K.

We now consider two examples where (17) is satisfied. The first example is frequently used later in the text, and it shows the effect of the specification of π. The other demonstrates that the specification of π and ϕ sometimes requires a little thought.

Example 4: Suppose that

$$\phi(x) = \gamma^{s_R(x)} \tag{19}$$

where $s_R(x) = \sum_{\{\xi,\eta\} \subseteq x} \mathbb{1}[d(\xi, \eta) \leq R]$, and $0 \leq \gamma \leq 1$ and $R \geq 0$ are given parameters. In the case where π is Poisson, we obtain a usual Strauss process (Strauss 1975, Kelly & Ripley 1976); if further $\gamma = 0$, we have a Gibbs hard core process (taking $0^0 = 1$ in (19)). Clearly, ϕ is non-increasing.

Simulated realisations of different Strauss processes are shown in Fig. 2. For this the perfect simulation algorithm in Section 3.2.2 has been used. Due to the thinning procedure in that algorithm (as illustrated later in Fig. 4), the first point pattern in Fig. 2 contains the two others.

Fig. 3 shows the mean number of points for different values of γ when π is either Poisson or geometric; the means are estimated by Monte Carlo methods using the perfect simulation algorithms in Sections 3.2.2 and 4.2. In the geometric case, for γ close to 1, the mean is very sensitive to changes in γ, while for γ not close to 1, the number of points is very small.

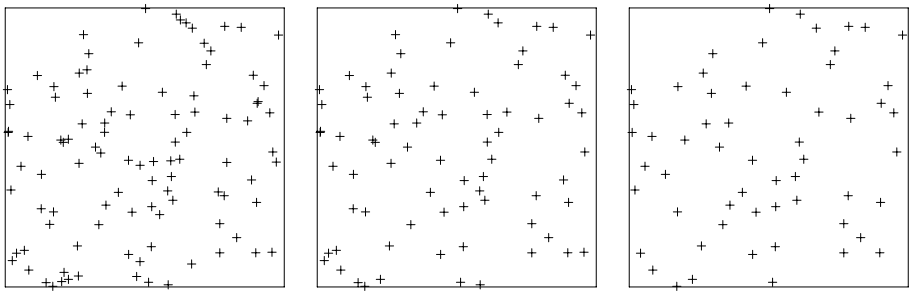

Fig. 2. Simulation of a usual Strauss process ($\pi_i \propto \beta^i/i!$) on $S = [0,1]^2$, when $\beta = 100$, $R = 0.05$, and $\gamma = 1, 0.5, 0$ (from left to right).

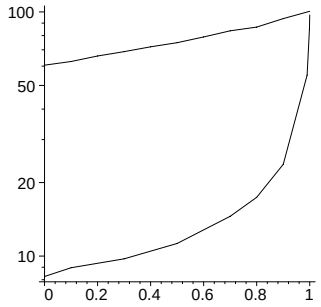

Fig. 3. Mean number of points versus γ in a usual Strauss process ($\pi_i \propto \beta^i/i!$; upper curve) and in a "geometric-Strauss" process ($\pi_i \propto q^i$; lower curve). In both models, $R = 0.05$ and π has mean 100 (i.e. $\beta = 100$ and $q = 100/101$). Each mean number of points is estimated from 500 i.i.d. samples.

Example 5: Assume that $S \subset \mathbb{R}^k$ has Lebesgue measure 1, λ is the uniform distribution on S, and $\pi_i \propto \beta^i/i!$ is Poisson. Consider an area interaction point process with unnormalised density

$$\phi(x) = \gamma^{-\lambda(U_x)}$$

with respect to ν (Widom & Rowlinson 1970, Baddeley & van Lieshout 1995, Häggström et al. 1999, Møller 2001). Here $U_x = \cup_{\xi \in x} \mathrm{ball}(\xi, R)$ where $\mathrm{ball}(\xi, R)$ denotes the closed ball centred in ξ with radius R, and $R > 0$ and $\gamma \geq 0$ are given parameters. The process is said to be attractive for $\gamma > 1$ and repulsive for $\gamma < 1$, as

$$\phi(x \cup \xi)/\phi(x) = \gamma^{-\lambda(U_{x \cup \xi} \setminus U_x)} \tag{20}$$

is non-decreasing ($\gamma > 1$) or non-increasing ($\gamma < 1$) in x. It follows from (20) that (17) holds in the attractive case but not in the

repulsive case. In the latter case we therefore redefine ν by taking $\pi_i \propto (\beta\gamma^{-B})^i/i!$ where B is the volume of a ball in $\mathbb{R}^k$ with radius R, and redefine ϕ by

$$\phi(x) = \gamma^{n(x)B-\lambda(U_x)}.$$

Then (17) is satisfied.

3.1.4 Coupling construction

Consider again the situation in Section 3.1.2. This section presents a coupling constructing where X is obtained by thinning from a dominating spatial birth-and-death process D, which is easily constructed. The coupling construction extends that of Kendall (1998) and Kendall & Møller (2000), and it is advantageous for simulation of X as the computation of the integral $B(x) = \int_S \beta(x,\xi)\lambda(d\xi)$ is not needed. Furthermore, the coupling construction becomes later useful for perfect simulation.

As will be clarified below, we need to assume that

$$b_{i-1} \le b_i \quad \text{and} \quad d_i/i \ge d_{i+1}/(i+1) \quad \text{for } i \in I\backslash\{0\}. \tag{21}$$

For instance, (21) is satisfied in Examples 1–3.

The dominating process D has birth and death rates

$$\beta^D(x,\xi) = b_{n(x)} \quad \text{and} \quad \delta^D(x,\xi) = d_{n(x)+1}/(n(x)+1).$$

This choice corresponds to a maximal birth rate and a minimal death rate in (8), cf. (9). By (16), D has invariant distribution ν. As the total birth and death rates satisfy

$$B^D(x) = b_{n(x)} \quad \text{and} \quad D^D(x) = d_{n(x)},$$

we can take $N = n(D)$. The process $D_t,\ t \ge 0$, can easily be generated: The simple birth-and-death process N is straightforwardly generated, and $D_t,\ t \ge 0$, is constructed forwards in time by

$$N_t = N_{t-} + 1 \Rightarrow D_t = D_{t-} \cup \xi_t,$$
$$N_t = N_{t-} - 1 \Rightarrow D_t = D_{t-}\backslash\eta_t,$$

where $\xi_t \sim \lambda$ and η_t is a uniformly selected point from D_{t-} (and where ξ_t and η_t are independent of the history of D before time t).

We wish to couple X to D so that D dominates X in the sense that

$$X_t \subseteq D_t \quad \text{for all } t \geq 0. \tag{22}$$

For simplicity, as in Kendall (1998) and Kendall & Møller (2000), we consider the case where $d_i \propto i$. This is fulfilled in Example 1. Further we assume that $\bar{d}$ and $\bar{b}$ are specified by (18) where ϕ satisfies (17). Then X_t can be constructed iteratively for $t \geq 0$ as follows. Each time $N_t = N_{t-} + 1$ we associate to the birth in D at time t a mark $R_t \sim \text{Uniform}[0,1]$ which is independent of $\{D_s : 0 \leq s \leq t\}$ and previous marks $R_s,\ 0 \leq s < t$. Initially let $X_0 \subseteq D_0$ and then

$$N_t = N_{t-} \qquad \Rightarrow X_t = X_{t-} \tag{23}$$

$$N_t = N_{t-} + 1 \Rightarrow X_t = \begin{cases} X_{t-} \cup \xi_t \text{ if } R_t \leq \bar{b}(X_{t-}, \xi_t) b_{n(X_{t-})} / b_{N_{t-}} \\ X_{t-} \qquad\qquad \text{otherwise} \end{cases} \tag{24}$$

$$N_t = N_{t-} - 1 \Rightarrow X_t = X_{t-} \backslash \eta_t. \tag{25}$$

Clearly this construction satisfies (22) as illustrated in Fig. 4. It is straightforwardly verified that the birth and death rates of X are given by (8). Note that (22) and (24) imply the need of the first restriction in (21).

Example 3 (and hence also Example 2) does not satisfy $d_i \propto i$, but it is still to some extend possible to use the coupling construction above, letting now $d_i = (1-p)i$ and $b_i = (i+\alpha)p$. Then the detailed balance condition (1) is fulfilled, and Q is non-explosive as (4) is satisfied.

Now, consider the situation where we neither assume $d_i \propto i$ nor that $\bar{b}$ and $\bar{d}$ are specified by (18). In general we have that (22) together with (25) imply a lower bound on the death rate

$$\delta(x \backslash \eta, \eta) = d_{n(x)} \bar{d}(x \backslash \eta, \eta) / n(x) \geq d_{n(y)} / n(y) \quad \text{for } \eta \in x, y \in \Omega, x \subseteq y, x \neq \emptyset.$$

This explains why the second restriction in (21) is needed. By (8), for each $\eta \in X_{t-}$, we need a rate

$$d_{n(X_{t-})} \bar{d}(X_{t-} \backslash \eta, \eta) / n(X_{t-}) - d_{n(D_{t-})} / n(D_{t-}) \tag{26}$$

for deleting η from X but not from D. In the coupling construction below we show how this is possible. However, readers who prefer to

skip these technical details should move on to Section 3.2 where only the construction in (23)–(25) is used.

The process X_t is iteratively generated forwards in time $t \geq 0$ as follows. Initially let $X_0 \subseteq D_0$, and let $X_t = X_{t-}$ whenever $N_t = N_{t-}$. Consider first the case $t = 0$. Set $x = X_t$, $i = n(x)$, and $j = n(D_t)$. Let $\tau = \min\{s > t : N_{s-} \neq N_s\}$ be the first jump time in N after time t. In accordance with (26), for each $\eta \in x$, define

$$a_\eta = d_i \bar{d}(x \backslash \eta, \eta)/i - d_j/j \quad \text{and} \quad A(x) = \sum_{\eta \in x} a_\eta, \tag{27}$$

setting $A(x) = 0$ if $x = \emptyset$. Then generate $\tau' \sim \text{Exp}(A(x))$ (independently of τ and the history of D and X before and including time t), setting $\tau' = \infty$ if $A(x) = 0$. If $\tau' < \tau - t$, then increase t by τ' and set $X_t = x \backslash \eta$, where $\eta \in x$ is chosen with probability $a_\eta / A(x)$ (independently of τ, τ', and the history of D and X before time t). Else set $t = \tau$ and let

$$N_t = N_{t-} + 1 \Rightarrow X_t = \begin{cases} x \cup \xi_t & \text{if } R_t \leq \bar{b}(x, \xi_t) b_i / b_j \\ x & \text{otherwise} \end{cases} \tag{28}$$

$$N_t = N_{t-} - 1 \Rightarrow X_t = x \backslash \eta_t, \tag{29}$$

where in (28) we associate to the birth in D a mark $R_t \sim \text{Uniform}[0, 1]$ (independent of τ, τ', the history of D before and including time t, and the history of X before time t). Repeating this procedure it is straightforwardly verified that (22) is satisfied and the birth and death rates of X are given by (8).

3.2 Perfect simulation

In this section we discuss perfect simulation procedures based on coupling from the past (Propp & Wilson 1996) for a point process with a distribution κ with respect to ν in (15). We assume that κ is the equilibrium distribution of a spatial birth-and-death process satisfying the conditions in Sections 3.1.2 and 3.1.4. We use the coupling construction in Section 3.1.4 in two ways, extending the methods of Kendall & Møller (2000) in Sections 3.2.1 and 3.2.2, and of Fernández et al. (1999) in Section 3.2.3. Finally, in Section 3.2.3 we compare the methods and report on some empirical findings.

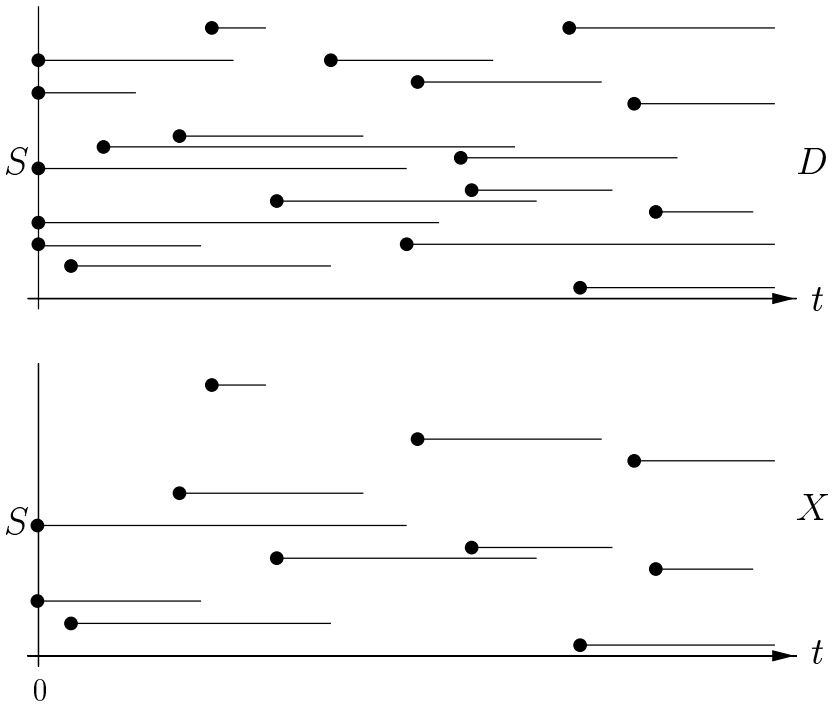

Fig. 4. Illustration of the coupling construction when $\bar{d} \equiv 1$, $X_0 \subseteq D_0$ and $S = [0, 1]$. Top: the dominating process D. Bottom: the target process X obtained by thinning from D as described in (24).

Throughout this section we assume for simplicity that $d_i \propto i$, and $\bar{d}$ and $\bar{b}$ are specified by (18) where ϕ satisfies (17). However, it is possible to extend our perfect samplers to the case of the more complicated coupling construction described at the end of Section 3.1.4.

3.2.1 Dominated coupling from the past

In this section we describe the simplest version of dominated coupling from the past (dominated CFTP), also called horizontal CFTP, cf. the survey in Møller (2001). For further details on dominated CFTP in a general setting, see Kendall & Møller (2000). Refined versions of dominated CFTP are given in Sections 3.2.2 and 3.2.3.

The basic idea in dominated CFTP is as follows. Start the dominating process D from Section 3.1.4 in equilibrium at time 0, and extend it both forwards and backwards in time. This is generally easily done, since D has equilibrium distribution ν and D is time reversible.

Further, for all forwards births $D_t \backslash D_{t-} \neq \emptyset$ with $-\infty < t < \infty$, generate independent marks $R_t \sim \text{Uniform}[0,1]$, where the R_t's are independent of $\{D_t : -\infty < t < \infty\}$. Let $\tau_i,\ i \in \mathbb{N}_0$, denote the times D enters $\emptyset$, i.e. $D_{\tau_i -} \neq \emptyset$ and $D_{\tau_i} = \emptyset$, and assume that $\tau_i < \tau_{i+1}$ for all $i \in \mathbb{N}_0$. For each $i \in \mathbb{N}_0$, construct a target process X_t on the time interval $[\tau_i, \tau_{i+1}]$ in exactly the same way as in Section 3.1.4 except that we are now starting with $X_{\tau_i} = \emptyset$. Then (D, X) regenerates each time D enters $\emptyset$, $X_t \subseteq D_t$ for all $-\infty < t < \infty$, and X is a spatial birth-and-death process with equilibrium distribution κ, where κ has unnormalised density ϕ. Because of time-stationarity of (D, X), for any two fixed times, e.g. 0 and -1, X_0 and X_{-1} follow the same distribution, which is then the invariant distribution κ. Hence $X_0 \sim \kappa$.

Let $-T_\emptyset \leq 0$ denote the first time that D enters $\emptyset$ when D is generated backwards from time 0. As $\emptyset$ is an ergodic atom, $T_\emptyset$ is finite almost surely. Note that for the generation of X_0, it suffices to consider the jump chains of D and X on the time interval $[-T_\emptyset, 0]$. Let $T'_\emptyset$ be the number of jumps D makes on $[-T_\emptyset, 0]$. As illustrated in Fig. 5, for the perfect simulation procedure described above, we need only to

(I) generate $D_0 \sim \nu$;
(II) generate the jump chain of D (together with the marks for forwards births), for $T'_\emptyset$ steps backwards in time from time 0;
(III) construct the jump chain of $\{X_t : -T_\emptyset \leq t \leq 0\}$ forwards in time;
(IV) return $X_0 \sim \kappa$.

However, for many applications $T'_\emptyset$ will be infeasibly large as shown in the following example.

Example 6: Consider the jump chain $\{\ldots, M_{-2}, M_{-1}, M_1, M_2, \ldots\}$ from Section 2 when this is in equilibrium and extended to all times $n \in \mathbb{Z}\backslash\{0\}$. Let $L_0 = T'_\emptyset + T''_\emptyset$ where $T'_\emptyset = \inf\{n \in \mathbb{N} : M_{-n} = 0\}$ and $T''_\emptyset = \inf\{n \in \mathbb{N} : M_n = 0\}$. By reversibility, $T'_\emptyset$ and $T''_\emptyset$ are identically distributed, so $\mathbb{E}T'_\emptyset = \mathbb{E}L_0/2$. By time-stationarity, $\mathbb{P}(L_0 = l) = \mathbb{P}(L = l)l/\mathbb{E}L$, hence $\mathbb{E}L_0 \geq \mathbb{E}L$, i.e. $\mathbb{E}T'_\emptyset \geq \mathbb{E}L/2$. Combining this with (2), it is straightforwardly derived that

$$\mathbb{E}T'_\emptyset \geq \mu_\pi/\pi_1 \tag{30}$$

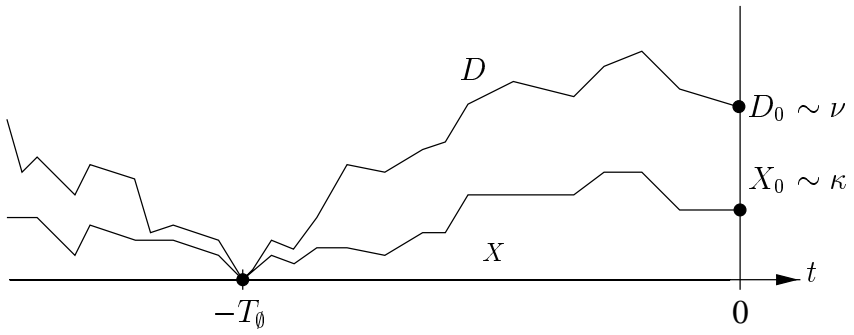

Fig. 5. Illustration of the idea behind dominated CFTP.

where μ_π denotes the mean of the distribution π.

Consider first the Poisson case $\pi_i \propto \beta^i/i!$. By (30), $\mathbb{E}T'_\emptyset \geq e^\beta$ is at least exponentially growing in β. For example, if $\beta = 100$ (corresponding to $\mu_\pi = 100$),

$$\mathbb{E}T'_\emptyset \geq e^{100} \approx 2.7 \times 10^{43}. \tag{31}$$

Consider next the case where $\pi_i \propto q^i \Gamma(i+\alpha)/i!$ is negative binomial as in Example 3, where $\alpha > 0$ and $0 < q < 1$. Let $d_i = i$ and $b_i = (i+\alpha)q$ (as noticed in Section 3.1.4, Q is then non-explosive). By (30),

$$\mathbb{E}T'_\emptyset \geq (1-q)^{-(1+\alpha)}. \tag{32}$$

If $(1-q)^{1-\alpha} \leq \alpha q(1+q)$, a better bound is obtained as follows. The probability for a transition $i \to i+1$ in M is $q(i+\alpha)/(q(i+\alpha)+i)$. As i increases, $q(i+\alpha)/(q(i+\alpha)+i)$ decreases towards $p \equiv q/(1+q) < 1/2$. Hence it suffices to consider a random walk $\{S_n : n \in \mathbb{N}_0\}$ starting in $S_0 = N_0$ and with $\mathbb{P}(S_{n+1} - S_n = 1) = 1 - \mathbb{P}(S_{n+1} - S_n = -1) = p$, since we can couple the random walk to M so that $T''_\emptyset \geq \mathcal{T}_0$ where $\mathcal{T}_0 = \inf\{n \in \mathbb{N}_0 : S_n = 0\}$. By Wald's identity, $N_0 = (1-2p)\mathbb{E}(\mathcal{T}_0|N_0)$, and so $\mathbb{E}N_0 \leq (1-2p)\mathbb{E}T''_\emptyset$, i.e.

$$\mathbb{E}T'_\emptyset \geq \alpha q(1+q)/(1-q)^2. \tag{33}$$

For instance, in the case of a geometric distribution ($\alpha = 1$), this is a better bound than that in (32) when $q > (\sqrt{5}-1)/2 \approx 0.618$, and we obtain

$$\mathbb{E}T'_\emptyset \geq 20100 \tag{34}$$

if $q = 100/101$ and $\alpha = 1$ (corresponding to $\mu_\pi = 100$).

This indicates that the running time may be much smaller in the geometric case than in the Poisson case (at least when $\mu_\pi = 100$ in both cases). In the geometric case we expect the distribution of $T'_\emptyset$ to be heavy-tailed so that very long running times may occur. However, in the negative binomial case with $\alpha \leq 1$, a smaller coalescence time is obtained using the algorithm in Section 4.2.

3.2.2 Upper and lower processes

A faster perfect simulation procedure can be obtained by constructing upper and lower processes $U^s = \{U_t^s : s \leq t \leq 0\}$ and $L^s = \{L_t^s : s \leq t \leq 0\}$ for $s \leq 0$. These ar constructed forward in time as follows. Initially set $U_s^s = D_s$ and $L_s^s = \emptyset$. For $s < t \leq 0$, if $d = D_{t-}$, $u = U_{t-}^s$, and $l = L_{t-}^s$, let

$$\begin{aligned} D_t = d \quad &\Rightarrow\ U_t^s = u \quad \text{and} \quad L_t^s = l, \\ D_t = d \backslash \eta_t \quad &\Rightarrow\ U_t^s = u \backslash \eta_t \quad \text{and} \quad L_t^s = l \backslash \eta_t, \end{aligned} \tag{35}$$

$$D_t = d \cup \xi_t \ \Rightarrow\ U_t^s = \begin{cases} u \cup \xi_t & \text{if } R_t \leq \alpha_{\max}(d, u, l, \xi_t) \\ u & \text{otherwise} \end{cases}$$

$$\text{and } L_t^s = \begin{cases} l \cup \xi_t & \text{if } R_t \leq \alpha_{\min}(d, u, l, \xi_t) \\ l & \text{otherwise,} \end{cases} \tag{36}$$

where

$$\alpha_{\max}(d, u, l, \xi) = \max\{\bar{b}(x, \xi) b_{n(x)} / b_{n(d)} : l \subseteq x \subseteq u\}, \tag{37}$$

$$\alpha_{\min}(d, u, l, \xi) = \min\{\bar{b}(x, \xi) b_{n(x)} / b_{n(d)} : l \subseteq x \subseteq u\}. \tag{38}$$

Other choices of $\alpha_{\max}$ and $\alpha_{\min}$ are possible and may be convenient for computational reasons, cf. the discussion at the end of this section.

The construction in (35)–(38) ensures the sandwiching property

$$L_t^s \subseteq X_t \subseteq U_t^s \subseteq D_t, \quad s \leq t \leq 0, \tag{39}$$

the funnelling property,

$$L_t^{s'} \subseteq L_t^s \subseteq U_t^s \subseteq U_t^{s'}, \quad s \leq s' \leq t \leq 0, \tag{40}$$

and the coalescence property,

$$L_t^s = U_t^s \quad \Rightarrow \quad L_{t'}^s = U_{t'}^s \quad \text{for } s \le t \le t' \le 0, \tag{41}$$

see Fig. 6. Hence once a pair of upper and lower processes have coalesced, they stay in coalescence, and at time 0 they are equal to $X_0 \sim \kappa$.

Now the perfect simulation algorithm works as follows. Pick a sequence of times $\dots s_2 < s_1 < 0$, where $\lim_{i\to\infty} s_i = -\infty$. Then start D in equilibrium at time 0 and generate it backwards until time s_1. For forwards birth times $t \in [s_1, 0]$ in D, generate i.i.d. marks $R_t \sim \text{Uniform}[0,1]$ independent of D. Then generate the upper and lower processes U^{s_1} and L^{s_1} as in (35)–(36). If U^{s_1} and L^{s_1} are in a common state at time 0 then stop. If not then extend D (together with the marks for forwards births) further backwards from $t = s_1$ to $t = s_2$, construct U^{s_2} and L^{s_2}, and check if $U_0^{s_2} = L_0^{s_2}$. This backwards-forwards step is repeated until the upper and lower chains have coalesced by time 0, i.e. at the coalescence time $-T_s = \sup\{s_i : U_0^{s_i} = L_0^{s_i}\}$. By (39), we have $U_0^{-T_s} = X_0$ and hence $U_0^{-T_s} \sim \kappa$. This procedure is illustrated in Fig. 6.

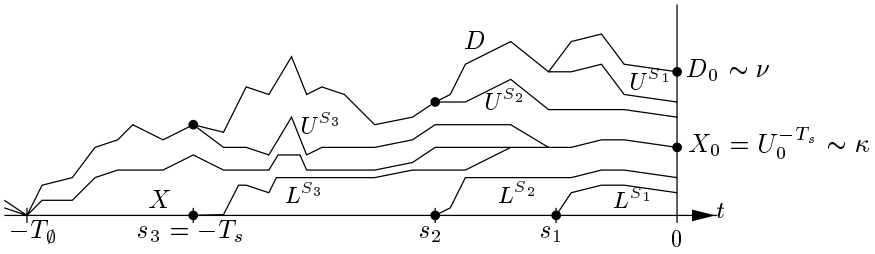

Fig. 6. Illustration of sandwiching, funnelling, and coalescence properties.

Notice that we need only to generate the jump chains of D and the upper and lower processes. Hence it is sensible to let $\{s_i\} \subseteq \{J_{-i}\}$, where $\dots J_{-2} \le J_{-1} < 0$ are the jump times of D before time 0. As $-T_s \le -T \le J_{-n(D_0)}$, where $-T = \sup\{s \le 0 : U_0^s = L_0^s\}$ is the "true coalescence time", it is natural to let $s_1 \le J_{-n(D_0)}$. For efficiency reasons a doubling scheme is usually used (Propp &

Wilson 1996), i.e.

$$s_1 = J_{-n}, s_2 = J_{-2n}, s_3 = J_{-4n}, s_4 = J_{-8n}, \ldots, \tag{42}$$

where $n \in \mathbb{N}$ is a user-specified parameter. Then we write T_n for T_s. Note that $T' \leq T'_n$ for the coalescence times in the jump chains, i.e. T' and T'_s are defined by $-T = J_{-T'}$ and $-T_s = J_{-T'_s}$. The doubling scheme can be refined by letting

$$s_1 = J_{-T'_{\min}}, s_2 = J_{-T'_{\min}-n}, s_3 = J_{-T'_{\min}-2n}, s_4 = J_{-T'_{\min}-4n}, \ldots, \tag{43}$$

writing $T_{n,\min}$ for T_s, and where $T'_{\min} = \inf\{k \in \mathbb{N}_0 : D_{J_{-k}} \cap D_0 = \emptyset\}$ specifies the number of jumps D has to go through from the first point in D_0 is born until time zero. Often in applications, $T_n \ll T_\emptyset$, $T'_n \ll T'_\emptyset$ and $T'_{n,\min} \leq T'_n$; this is illustrated later in Example 7.

The calculation of $\alpha_{\max}$ and $\alpha_{\min}$ is particular simple in the following cases. A point processes is attractive respectively repulsive if

$$\bar{b}(x,\xi) \leq \bar{b}(y,\xi) \quad \text{whenever} \quad x \subseteq y, \xi \notin y, \tag{44}$$
$$\bar{b}(x,\xi) \geq \bar{b}(y,\xi) \quad \text{whenever} \quad x \subseteq y, \xi \notin y. \tag{45}$$

For example, for the point processes considered in Examples 4 and 5, either (44) or (45) is satisfied. In the attractive case (44) we obtain by (21) and (37)–(38),

$$\alpha_{\max}(d,u,l,\xi) = \bar{b}(u,\xi) b_{n(u)}/b_{n(d)},$$
$$\alpha_{\min}(d,u,l,\xi) = \bar{b}(l,\xi) b_{n(l)}/b_{n(d)}.$$

In the repulsive case (45), it is computational convenient to redefine $\alpha_{\max}$ and $\alpha_{\min}$ by

$$\alpha_{\max}(d,u,l,\xi) = \bar{b}(l,\xi) b_{n(u)}/b_{n(d)},$$
$$\alpha_{\min}(d,u,l,\xi) = \bar{b}(u,\xi) b_{n(l)}/b_{n(d)}.$$

Then the perfect simulation algorithm still works as (39)–(41) hold. Note that it is only in the attractive case that U^s and L^s are individual Markov chains. Observe also that in Fig. 2, the point pattern for $\gamma = 0$ is not contained in that for $\gamma = 0.5$. This is possible because the Strauss model is non-attractive.

3.2.3 Clan of ancestors

An alternative perfect simulation algorithm is given by Fernández et al. (1999). We assume that ϕ has finite range of interaction, i.e. there exists an $R < \infty$ such that for any $x \in \Omega$ and $\xi \in S \backslash x$, $\bar{b}(x, \xi) = \bar{b}(x \cap \mathrm{ball}(\xi, R), \xi)$. This is fulfilled in Examples 4 and 5.

The algorithm is based on two central concepts defined as follows. Consider again the coupling construction in Section 3.1.4. When we have a birth $D_t = D_{t-} \cup \xi_t$, the *parents* of $\xi = \xi_t$ is the set $\mathrm{pa}(\xi) = D_{t-} \cap \mathrm{ball}(\xi, R)$ (we suppress in the notation that $\mathrm{pa}(\xi)$ depends on D_{t-}). Whether $X_t = X_{t-} \cup \xi$ or $X_t = X_{t-}$ in (24) depends on D_{t-} only through $\mathrm{pa}(\xi)$, or in fact only through those points in $\mathrm{pa}(\xi)$ which have not been accepted earlier in the thinning of D. In this sense whether ξ should be added to X depends not only on its parents, but also on the parents of its parents, and so on. Therefore we define recursively the ith generation ancestors of ξ by $\mathrm{pa}^i(\xi) = \cup_{\eta \in \mathrm{pa}^{i-1}(\xi)} \mathrm{pa}(\eta)$, $i = 2, 3, \ldots$, where $\mathrm{pa}^1(\xi) = \mathrm{pa}(\xi)$ (again we suppress the dependence of D when defining $\mathrm{pa}^i(\xi)$). Furthermore, let $\mathrm{an}(\xi) = \cup_{i \in \mathbb{N}} \mathrm{pa}^i(\xi)$ be the union of all generations of ancestors of ξ. If $\xi \in D_0$, then the points in $\mathrm{an}(\xi)$ are the only points which can have any influence on whether ξ is in X_0 or not. So $C(D_0) = \cup_{\xi \in D_0} \mathrm{an}(\xi) \cup D_0$ is the set of *all* points in D which can have any influence on the configuration of X_0. In other words, if $[-T_C, 0]$ is the time interval in which the points in $C(D_0)$ are living, then $T_C \leq T_\emptyset$, and X_0 is unaffected if we set $X_{-T_C} = \emptyset$ and generate X_t forwards in time $t \geq -T_C$ as usual. We refer to $C(D_0)$ as the *clan of ancestors*. An example of a clan is shown in Fig. 7.

Now, briefly the perfect simulation algorithm works as follows (noticing that we only need to generate the relevant jump chains). Let $D_0 \sim \nu$. As noticed above we only need to generate the dominating process D back until time $-T_C$, which may be determined as follows. Initially all points in D_0 are set to be in the clan. Then whenever a member of the clan dies in D (when D is considered backwards in time), e.g. $D_t = D_{t-} \cup \{\xi_t\}$, all points in D_{t-} within distance R of ξ_t are the parents of ξ_t and are therefore added to the clan. When all members of the clan have died, we have found the time T_C and the generation of D stops. Finally, we set $X_{-T_C} = \emptyset$, construct X_t, $t \geq -T_C$, as in Section 3.2.1, and return $X_0 \sim \kappa$.

As noticed, $T_C \leq T_\emptyset$, and it is not hard to see that $T \leq T_C$. Similarly, $T' \leq T'_C \leq T'_\emptyset$, where $T_C = -J_{-T'_C}$. Note that the running time of the algorithm depends only on $\bar{b}$ through R, and no monotonicity properties such as (44) and (45) are required.

Example 7: Consider a Gibbs hard core process as given by (19) with $\gamma = 0$ and hard core parameter $R \geq 0$. Let S be a unit square, and ν a homogeneous Poisson point process of rate $\beta = 100$. Fig. 8 shows the mean coalescence times of T'_C , T'_n, and $T'_{n,\min}$ versus R, when $n = 1$ in the doubling schemes. For all values of R in Fig. 8, the mean coalescence times are much smaller than $\mathbb{E}T'_\emptyset \geq e^{100}$, cf. (31). For small values of R, the clan method and the refined doubling scheme are almost equal (in the sense that $\mathbb{E}(T'_C) \approx \mathbb{E}(T'_{n,\min})$) and slightly better than the (ordinary) doubling scheme (where the mean coalesce times are about 2/3 smaller). As R increases, T'_n and $T'_{n,\min}$ do not increase as rapidly as T'_C, and the clan method is clearly slower than both doubling schemes for large values of R. This is not so surprising, as an increase in R implies an increase in the number of parents for each point, and hence a larger clan reaching further back in time. The specification of β is of course also crucial: T'_C tends to be smaller than T'_n for small values of β, while the opposite is the case for large values of β.

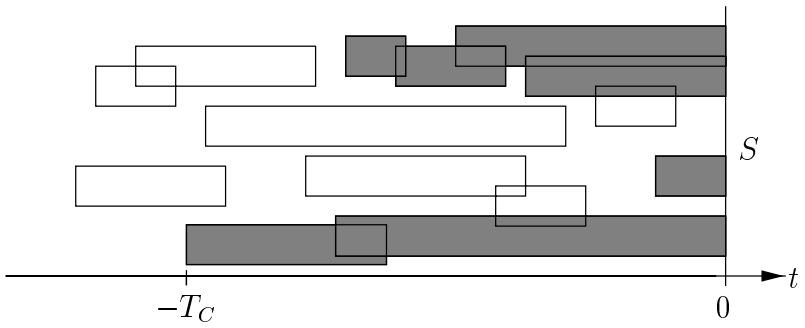

Fig. 7. Example of a clan of ancestors where shaded blocks represents members of the clan.

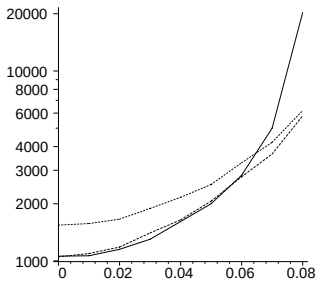

Fig. 8. Mean coalescence times versus the hard core parameter R for a hard core process. Full line: $\mathbb{E}(T'_C)$; upper dotted line: $\mathbb{E}(T'_n)$; lower dotted line: $\mathbb{E}(T'_{n,\min})$. Each mean value is estimated from 500 independent perfect simulations.

4 Spatial birth-and-catastrophe process

4.1 Description and construction

4.1.1 The simple case

We start by specifying a simple birth-and-catastrophe process $N = \{N_t : t \geq 0\}$ with generator Q given by $q_{i+1,0} = d_{i+1}$ and $q_{i,i+1} = b_i$, while all other off-diagonal elements are zero. Hence $q_i = b_i + d_i$ for all $i \in \mathbb{N}_0$, setting $d_0 = 0$. Note that d_i now has another interpretation than in Section 3.

As in Section 2, (i)–(iv) are assumed to hold. For simplicity we assume that $I = \mathbb{N}_0$ and $d_{i+1} > 0$ and $b_i > 0$ for all $i \in \mathbb{N}_0$. Then (ii) is satisfied. Further (iii) is equivalent to that both

$$\pi_{i+1} = \pi_0 \frac{b_0}{q_1} \frac{b_1}{q_2} \cdots \frac{b_i}{q_{i+1}}, \quad i \in \mathbb{N}_0, \tag{46}$$

and

$$\lim_{n \to \infty} \frac{b_0 \cdots b_i}{q_1 \cdots q_i} = 0 \tag{47}$$

are satisfied. To see this, rewrite $\pi Q = 0$ as

$$\pi_0 b_0 = \sum_{i=1}^{\infty} \pi_i d_i \quad \text{and} \quad \pi_i b_i = \pi_{i+1} q_{i+1} \quad \text{for all } i \in \mathbb{N}_0. \tag{48}$$

¿From the latter equation, using induction, we obtain (46). Combining the equations in (48),

$$\pi_0 b_0 = \sum_{i=1}^{\infty} \pi_{i-1} b_{i-1} - \pi_i b_i = \pi_0 b_0 - \lim_{i \to \infty} \pi_i b_i,$$

whereby (47) is obtained.

4.1.2 The spatial case

We next construct a spatial birth-and-catastrophe process $X = \{X_t : t \geq 0\}$, proceeding along similar lines as in Sections 3.1.2–3.1.4. The process is specified by two measurable functions $\beta : \Omega \times S \to [0, \infty)$ and $\delta : \Omega \to [0, \infty)$, where for any $x \in \Omega$, $\beta(x, \xi)$ is the birth rate at which a point $\xi \in S$ is added to x, and if $x \neq \emptyset$, $\delta(x)$ is the catastrophe rate a which all points are deleted from x. As in Sections 2 and 3.1.2, let $J_1 < J_2 < \ldots$ be the jump times, $H_1, H_2, \ldots$ the holding times, and $Y = \{Y_n : n \in \mathbb{N}_0\}$ the jump chain of X with $Y_0 = X_0$. The conditional distribution of H_n given $H_1, \ldots H_{n-1}, Y_0, \ldots, Y_{n-1} = x$ is $\mathrm{Exp}(B(x)+\delta(x))$, where $B(x) = \int_S \beta(x, \xi)\lambda(d\xi)$ is the total birth rate (assuming $B(x) < \infty$) and we set $\delta(\emptyset) = 0$. If we also condition on H_n, the probability for a birth at time J_n is $B(x)/(B(x) + \delta(x))$. If we further condition on that a birth happens at time J_n, then $Y_n = x \cup \xi$ where ξ has density $\beta(x, \cdot)/B(x)$. If instead a catastrophe happens at time J_n, then $Y_n = \emptyset$.

We assume that β and δ are related to the rates for N by

$$\beta(x, \xi) = b_{n(x)}\bar{b}(x, \xi) \quad \text{and} \quad \delta(x) = d_{n(x)}\bar{d}(x), \tag{49}$$

where $0 \leq \bar{b} \leq 1$ and $\bar{d} \geq 1$ are measurable functions. Then, by (49), $B(x) \leq b_{n(x)}$ and $\delta(x) \geq d_{n(x)}$. Furthermore, we assume that $d_1 \geq d_2 \geq \ldots$. Then X and N can be coupled so that N dominates $n(X)$ as in (11). Briefly, (X, N) is a jump process where a transition from (x, n) to (x', n') happens with a rate given as follows, where $m = n(x) \leq n$:

$$\begin{aligned}
&b_n \ \text{ for } (x', n') = (x, n + 1) \text{ and } m < n,\\
&\beta(x, \xi) \ \text{ for } (x', n') = (x \cup \xi, n) \text{ and } m < n,\\
&\beta(x, \xi) \ \text{ for } (x', n') = (x \cup \xi, n + 1) \text{ and } m = n,\\
&b_n - B(x) \ \text{ for } (x', n') = (x, n + 1) \text{ and } m = n,\\
&d_n \ \text{ for } (x', n') = (\emptyset, 0) \text{ and } m \leq n,\\
&\delta(x) - d_n \ \text{ for } (x', n') = (\emptyset, n) \text{ and } m \leq n.
\end{aligned}$$

As in Section 3.1.2 we obtain the following: X is non-explosive; $\emptyset$ is an ergodic atom at which X regenerates; as in (12), X_t converges in

distribution towards an equilibrium distribution κ, say, independent of the initial state of X_0 as $t \to \infty$; and κ has a support given by (14).

The equilibrium distribution is uniquely characterised by the equations

$$B(\emptyset)\kappa(\emptyset) = \int_{x \neq \emptyset} \delta(x)\kappa(dx) \tag{50}$$

and

$$\int_{F_i} [B(x) + \delta(x)]\kappa(dx) = \iint_{x \cup \xi \in F_i} \beta(x, \xi)\lambda(d\xi)\kappa(dx) \tag{51}$$

for all $i \in \mathbb{N}$ and $F_i \in \mathcal{F}$ with $F_i \subseteq \Omega_i$ (this follows from (56) in Section 5). If κ has an unnormalised density ϕ with respect to ν given by (15), then (50) and (51) become

$$b_0 \pi_0 \phi(\emptyset) \int \bar{b}(\emptyset, \xi)\lambda(d\xi) = \sum_{i=1}^{\infty} d_i \pi_i \int \bar{d}(x)\phi(x)\lambda_i(dx) \tag{52}$$

and

$$\int_{F_i} \pi_i \phi(x) \left(b_i \int \bar{b}(x, \xi)\lambda(d\xi) + d_i \bar{d}(x) \right) \lambda_i(dx) = \iint_{x \cup \xi \in F_i} \pi_{i-1} b_{i-1} \bar{b}(x, \xi)\phi(x)\lambda_{i-1}(dx). \tag{53}$$

Using (48), then (52) and (53) are easily seen to be satisfied if, for example, ϕ is non-increasing as in (17), $\bar{b}$ is the Papangelou intensity as in (18), and

$$\bar{d}(x) = 1 + \left(b_{n(x)} - B(x) \right) / d_{n(x)} \quad \text{for } x \neq \emptyset.$$

In the sequel we assume that $\bar{b}$ and $\bar{d}$ are defined in this way for a given unnormalised density ϕ which is non-increasing. These assumptions imply that the total rates in N and X agree in the sense that for any $x \in \Omega$, $B(x) + \delta(x) = q_{n(x)}$. In other words, if $X_t = x$, the waiting time τ to the next jump is $\mathrm{Exp}(q_{n(x)})$-distributed, and we may generate $\xi \sim \lambda$ and return $X_{t+\tau} = x \cup \xi$ with probability $\beta(x, \xi)/q_{n(x)}$, and return $X_{t+\tau} = \emptyset$ otherwise.

4.2 Perfect simulation

In this section we show that the simple perfect simulation procedure in Section 3.2.1 is feasible when the dominating simple jump process is of the irreversible type presented in Section 4.1. In addition to the assumptions in Section 4.1, we assume for simplicity that $q_i \equiv 1$ is constant for all $i \in \mathbb{N}$ and $b_0 \leq b_1$. Equivalently, by (48), $b_i = \pi_{i+1}/\pi_i \leq 1$ is non-decreasing for $i \in \mathbb{N}_0$, and $d_i = 1 - b_i$ for $i \in \mathbb{N}$. Hence the Poisson case $\pi_i \propto \beta^i/i!$ is excluded, since $\pi_{i+1}/\pi_i = \beta/(i+1)$ is decreasing in i. If $\pi_i \propto \Gamma(i+\alpha)q^i/i!$ is negative binomial (Example 3), $\pi_{i+1}/\pi_i = q(i+\alpha)/(i+1) \leq 1$ is non-decreasing if and only if $\alpha \leq 1$.

Now, a coupling construction is easily specified: Suppose N_t is generated for all times $t \geq 0$, and for all jump times t of N_t, there are generated mutually independent points $\xi_t \sim \lambda$ and numbers $R_t \sim$ Uniform$[0,1]$ (independent of N). Let $n(X_0) \leq N_0$ and generate X_t forwards in time $t > 0$ as follows. Suppose that $X_{t-} = x$ and $N_{t-} = j$ with $n(x) = i \leq j$. If $N_t = j$ is unchanged, then $X_t = x$ is unchanged, while

$$N_t \neq j \Rightarrow X_t = \begin{cases} x \cup \xi_t & \text{if } N_t = j+1 \text{ and } R_t \leq \bar{b}(x, \xi_t) b_i / b_j \\ \emptyset & \text{otherwise.} \end{cases} \tag{54}$$

Since $\{b_j\}$ is non-decreasing, we obtain a spatial birth-and-catastrophe process with rates as required, and by induction, $N_t \geq n(X_t)$ for all $t \geq 0$.

For perfect simulation, we let $N_0 \sim \pi$ and imagine that N_t is generated backwards in time $t \leq 0$ until it reaches zero, i.e. at time $-\tilde{T}_\emptyset \equiv \sup\{t \leq 0 : N_t = 0\}$. The number of jumps in $[-\tilde{T}_\emptyset, 0]$ is simply given by N_0, since N_t can only move one step up when considered forwards in time between $-\tilde{T}_\emptyset$ and 0. Hence, in accordance with (54), we let $Y_0 = \emptyset$ and generate Y_i, $i = 1, \ldots, N_0$, by

$$Y_i = \begin{cases} Y_{i-1} \cup \xi_i & \text{if } R_i \leq \bar{b}(Y_{i-1}, \xi_i) b_{n(Y_{i-1})} / b_{i-1} \\ \emptyset & \text{otherwise,} \end{cases} \tag{55}$$

where $\xi_i \sim \lambda$ and $R_i \sim$ Uniform$[0,1]$. Since $Y_0, \ldots, Y_{N_0}$ can be considered as the part of the jump chain of X_t ("started in the infinite

past") where $t \in [-\tilde{T}_\emptyset, 0]$, we conclude that $Y_{N_0} \sim \kappa$ (a formal argument follows exactly the same lines as in Section 3.2.1).

So perfect simulation is very simple:

(I) generate $N_0 \sim \pi$;
(II) generate the jump chain $Y_0, \ldots, Y_{N_0}$;
(III) return $Y_{N_0} \sim \kappa$.

The number of jumps is given by $N_0 \sim \pi$. This should be compared to the case in Example 6: If $\pi_i \propto q^i$ is geometric, $\mu_\pi = q/(1-q)$ is $(1+q)/(1-q)$ times smaller than the lower bound on $\mathbb{E}T'_\emptyset$ in (33). Hence it is much faster; for example, $(1+q)/(1-q) = 201$ if $\mu_\pi = 100$, cf. (34).

Unless $\bar{b}$ is very close to 1 (corresponding to a weak interaction), we might by (55) expect $n(Y_{N_0})$ to be small. This is in accordance with the observations in Example 4, and it explains why the perfect simulation algorithm is so fact. Furthermore, in contrast to the algorithms in Section 3.2, there is no need for storing any information: our perfect sampler is a simple example of a so called read-once algorithm (Wilson 2000).

5 The general case

In a general setting, a spatial jump process $X = \{X_t : t \geq 0\}$ may be specified by measurable functions $\beta, \delta : \Omega \times \Omega \to [0, \infty)$, where for any $x \in \Omega$, $\beta(x, y)$ is the rate at which a point configuration $y \in \Omega$ is added to x, and if $z = x \cup y \in \Omega$, $\delta(x, y)$ is the rate at which y is deleted from z; in order to ensure that the process really jumps, we set $\beta(x, \emptyset) = 0$ and $\delta(x, \emptyset) = 0$. Let further $B(x) = \int \beta(x, y)\nu(dy)$ and $D(z) = \sum_{x \subseteq z} \delta(x, z \backslash x)$, where ν is defined as in (15) with respect to a given probability density function $\pi = \{\pi_i\}$ on $\mathbb{N}_0$. Then X can be defined in terms of its jump times, holding times, and jump chain in the same way as in Section 3.1.2. Especially, for the jump chain $\{Y_n\}$, in case of a transition $Y_n = x \cup y$ given that $Y_{n-1} = x$, y has density $\beta(x, \cdot)/B(x)$ with respect to ν (assuming $B(x) < \infty$), and in case of a transition $Y_n = x$ given that $Y_{n-1} = z = x \cup y$, y has been selected with probability $\delta(x, y)/D(z)$.

Let us, as before, restrict attention to the non-explosive case. For instance, X is non-explosive if $B(x) + D(x)$ is uniformly bounded.

By Proposition 8.1 in Preston (1977), a probability measure κ on $(\Omega, \mathcal{F})$ is invariant if and only if

$$\int_{F_i} [B(x) + D(x)]\, \kappa(dx) = \iint_{x \cup y \in F_i} \beta(x,y)\nu(dy)\kappa(dx) + \int \sum_{x \subseteq z : x \in F_i} \delta(x, z \backslash x)\kappa(dz) \quad (56)$$

for all $i \in \mathbb{N}_0$ and $F_i \in \mathcal{F}$ with $F_i \subseteq \Omega_i$. Here invariance means that $X_t \sim \kappa$ for all $t \geq 0$ if $X_0 \sim \kappa$. In particular, if X converges in distribution towards κ, then κ is the unique invariant distribution. For example, assume that Q and π are in balance, cf. (1), and

$$\beta(x,y) = q_{n(x),n(x)+n(y)}\bar{b}(x,y)$$

and

$$\delta(x,y) = q_{n(x)+n(y),n(x)}\bar{d}(x,y) / \binom{n(x)+n(y)}{n(x)}.$$

Then, if ϕ is an unnormalised density for κ with respect to ν, there is a detailed balance condition:

$$\phi(x)\bar{b}(x,y) = \phi(x \cup y)\bar{d}(x,y) \quad \text{for all } x, y \in \Omega. \quad (57)$$

It is straightforwardly verified that this implies both (56) and reversibility of X.

In Sections 3.1 and 4.1 we demonstrated that a coupling of X to a simple jump process N is very useful. However, at the present level of generality it seems difficult to specify a coupling so that N dominates $n(X)$. Probably we need to consider this problem case by case; indeed the coupling constructions in Sections 3.1 and 4.1 depend much on the choice of β and δ. For making perfect simulations from a given invariant distribution κ, we may simply try to use one of the algorithms in Sections 3.2 and 4.2.

Example 8: Let $\pi_i \propto \beta^i/i!$ be Poisson, where $\beta > 0$,

$$q_{i,i+j} = p^i(\beta(1-p))^j/j!, \quad q_{i+j,i} = \binom{i+j}{i} p^i(1-p)^j,$$

where $0 < p < 1$, and

$$\bar{d} \equiv 1, \quad \bar{b}(x,y) = \phi(x \cup y)/\phi(y),$$

with $\phi > 0$ non-increasing. Then the detailed balance conditions (1) and (57) are satisfied, and X develops as follows. Suppose that $X_t = x$. Then

(I) with rate $p^{n(x)}e^{\beta(1-p)}$, first propose a birth $x \to x \cup y$, drawn from a Poisson process with intensity measure $\beta(1-p)\lambda$, next with probability $\phi(x \cup y)/\phi(x)$ accept $x \cup y$, and retain x otherwise;

(II) with rate 1 make an independent thinning of the points in x, where each point is retained with probability p.

Note that nothing may happen with rate $2p^{n(x)}$ (namely if $y = \emptyset$ in (I) or all points are retained in (II)). Observe also that y is independent of x in (I), and the retention probability p does not depend on x in (II).

Both N and X are non-explosive as $q_i = p^i e^{\beta(1-p)} + 1 - 2p^i$ and $B(x)+D(x) \leq q_{n(x)}$ are bounded. It is easily seen that $q_{i,i+j} \geq q_{i-k,i+j}$ for all integers $0 < k \leq i$ and $j \geq 1$ if and only if $p \geq \beta/(2+\beta)$. Hence, if $p \geq \beta/(2+\beta)$, we obtain that N dominates X by specifying a jump process (X, N) as follows. The rate for a jump $(x, n) \to (x', n')$ is as follows:

$$\begin{aligned}
&\beta(x,y) \quad \text{for } (x',n') = (x \cup y, n(x \cup y)) \text{ and } n < n', \\
&q_{n,n'} - \textstyle\int_{n(x \cup y)=n'} \beta(x,y)\nu(dy) \quad \text{for } x' = x \text{ and } n < n', \\
&\beta(x,y) \quad \text{for } (x',n') = (x \cup y, n) \text{ and } n(x \cup y) \leq n, \\
&q_{n,n'} \quad \text{for } x' = x \backslash y \text{ and } n' \leq n - n(y).
\end{aligned}$$

Because of this domination, we have the usual ergodicity properties: $\emptyset$ is an ergodic atom at which X regenerates, and X converges in distribution towards κ (independent of the initial state of X_0).

It is tempting to try to extend this coupling construction to a dominating spatial jump process $D \supseteq X$, but this idea fails because of the accept-rejection step (I). However, we can still simulate X_0 from its equilibrium distribution as any of the perfect simulation algorithms in Section 3.2 apply (for any $\beta > 0$ and $0 < p < 1$).

Acknowledgment

This research will be a part of KKB's PhD dissertation. JM was supported by the European Union's research network "Statistical and Computational Methods for the Analysis of Spatial Data, ERB-FMRX-CT96-0096", by the Centre for Mathematical Physics and Stochastics (MaPhySto), funded by a grant from the Danish National Research Foundation, and by the Danish Natural Science Research Council. We thank Bjarne Højgaard for helpful discussions.

References

Asmussen, S. (1987). *Applied Probability and Queues*, Wiley, Chichester.

Baddeley, A. & van Lieshout, M. N. M. (1995). Area-interaction Point Processes, *Ann. Inst. Statist. Math.* **46**: 601–619.

Berthelsen, K. K. & Møller, J. (2001). Perfect simulation and inference for spatial point processes. In preparation.

Daley, D. J. & Vere-Jones, D. (1988). *An Introduction to Point Processes*, Springer, Berlin.

Feller, W. (1971). *An Introduction to Probability Theory and Its Applications*, Vol. 2, 2nd edn, John Wiley and Sons Inc., New York.

Fernández, R., Ferrari, P. A. & Garcia, N. L. (1999). Perfect simulation for interacting point processes, loss networks and Ising models. Unpublished.

Georgii, H.-O. (2000). Phase transitions and percolation in Gibbsian particle models, *in* K. R. Mecke & D. Stoyan (eds), *Statistical Physics and Spatial Statistics*, Vol. 554 of *Lecture Notes in Physics*, Springer, Berlin, pp. 267–294.

Geyer, C. J. (1999). Likelihood inference for spatial processes, *in* O. E. Barndorff-Nielsen, W. Kendall & M. van Lieshout (eds), *Stochastic Geometry, Likelihood and Computation*, Chapman & Hall, pp. 79–140.

Häggström, O., van Lieshout, M. N. M. & Møller, J. (1999). Characterization results and Markov chain Monte Carlo algorithms including exact simulation for some spatial point processes, *Bernoulli* **5**: 641–658.

Kallenberg, O. (1984). An informal guide to the theory of conditioning in point processes, *Int. Statist. Rev.* **52**: 151–164.

Kelly, F. P. & Ripley, B. D. (1976). A note on Strauss's model for clustering, *Biometrika* **63**: 357–360.

Kendall, W. S. (1998). Perfect simulation for the area-interaction point process, *in* L. Accardi & C. Heyde (eds), *Probability Towards 2000*, Springer, New York, pp. 218–234.

Kendall, W. S. & Møller, J. (2000). Perfect simulation using dominating processes on ordered spaces, with application to locally stable point processes, *Adv. Appl. Prob.* **32**: 844–865.

Meyn, S. P. & Tweedie, R. L. (1994). *Markov Chains and Stochastic Stability*, Springer-Verlag, London.

Møller, J. (1989). On the rate of convergence of spatial birth-and-death processes, *Ann. Inst. Statist. Math.* **41**: 565–581.

Møller, J. (2001). A review of perfect simulation in stochastic geometry, *in* I. V. Basawa, C. C. Heyde & R. L. Taylor (eds), *Selected Proceedings of the Symposium on Inference for Stochastic Processes*, IMS Lecture Notes & Monographs Series. To appear.

Møller, J. & Schladitz, K. (1999). Extensions of Fill's algorithm for perfect simulation, *J. Roy. Statist. Soc. Ser B* **61**: 1–15.

Norris, J. R. (1997). *Markov Chains*, Cambridge University Press, New York.

Preston, C. (1977). Spatial birth-and-death processes, *Bull. Int. Statist. Inst.* **46**: 371–391.

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

Reuter, G. E. H. & Ledermann, W. (1953). On the differential equations for the transition probabilities of Markov processes with enumerably many states, *Proc. Camb. Phil. Soc.* **49**: 247–262.

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

Ruelle, D. (1969). *Statistical Mechanis: Rigorous Results*, W. A. Benjamin, Reading, Massachusetts.

Strauss, D. J. (1975). A model for clustering, *Biometrika* **62**: 467–475.

Thönnes, E. (1999). Perfect simulation of some point processes for the impatient user, *Adv. Appl. Prob.* **31**: 69–87.

Widom, B. & Rowlinson, J. S. (1970). A new model for the study of liquid-vapor phase transitions, *J. Chem. Phys.* **52**: 1670–1684.

Wilson, D. B. (2000). How to Couple from the Past Using a Read-Once Source of Randomness, *Random Structures and Algorithms* **16**: 85–113.