BULLETIN
BRAZILIAN
MATHEMATICAL
SOCIETY

Bull Braz Math Soc, New Series 33(3), 351-367

A primer on perfect simulation for spatial point processes

Kasper K. Berthelsen and Jesper Møller

Abstract. This primer provides a self-contained exposition of the case where spatial birth-and-death processes are used for perfect simulation of locally stable point processes. Particularly, a simple dominating coupling from the past (CFTP) algorithm and the CFTP algorithms introduced in [13], [14], and [5] are studied. Some empirical results for the algorithms are discussed.

Keywords: Coupling from the past (CFTP), clans of ancestors, dominated CFTP, exact simulation, local stability, spatial birth-and-death process, Strauss process.

Mathematical subject classification: Primary 60G55, 62M30, 65C05; Secondary: 60D05; 60K35, 68U20.

1 Introduction

One of the most exciting and important recent developments in Markov chain Monte Carlo (MCMC) is perfect or exact simulation. Following the seminal work by [19] many new perfect simulation ideas have appeared, particularly for spatial point processes, cf. the survey in [17]; see also Wilson's web site (http://dimacs.rutgers.edu/∼dwilson/exact.html). The aims of this paper are to review and compare the performance of some perfect simulation algorithms which apply on a rather general class of point processes, viz. locally stable point processes. For simplicity, apart from Section 8, we consider only finite point processes.

We focus on algorithms based on dominated (or horizontal) coupling from the past (CFTP) using spatial birth-and-death processes; alternative and efficient perfect samplers have been developed for some special models, see [17] and the references therein. In [14] dominated CFTP is treated in a general context

Received 30 June 2002.

and applied on locally stable point processes using either spatial birth-and-death processes or a Metropolis-Hastings algorithm. In this paper we give an alternative and self-contained exposition of the case where spatial birth-and-death processes are used. A spatial birth-and-death process is a continuous time Markov process where each transition consists in either adding a new point to the process (a birth) or deleting an existing point from the process (a death). Background material on spatial birth-and-death processes can be found in [18] and [16], but it is not needed in the present paper. Extensions of the algorithms considered in this paper are given in [4] using spatial jump processes. Another extension which is not treated in this paper, is Wilson's 2000a read-once version of CFTP. This algorithm applies also on locally stable point processes, and it drastically reduces the storage requirements.

The paper is organized as follows. Section 2 describes the setting for spatial point processes used in this paper, and it is explained what is meant by local stability. Section 3 specifies a coupling construction which is underlying the perfect samplers considered later. Section 4 discusses a very simple perfect simulation algorithm, and we show that it is too slow for practical purposes. Section 5 describes a more efficient algorithm based on so-called upper and lower processes [13, 14]. Section 6 describes an alternative algorithm using so-called clans of ancestors [5]. Section 7 discusses some empirical findings for the various perfect simulation algorithms. Section 8 concludes with some comments on extensions to infinite point processes.

2 Background

Throughout this paper we consider a fairly general setting for a spatial point process χ defined on a space S, equipped with a σ-algebra $\mathcal{B}$ which contains all singleton sets, and a diffuse probability measure λ, i.e. $\{\xi\} \in \mathcal{B}$ and $\lambda(\{\xi\}) = 0$ for all $\xi \in S$. For simplicity we assume χ to be a finite subset of S, though everything in the sequel easily extend to the case where χ is allowed to have multiple points and λ is not necessarily diffuse.

The state space of χ is the set of all finite point configurations $\Omega = \bigcup_{i=0}^{\infty}\{x \subseteq S : n(x) = i\}$, where $n(x)$ denotes the number of points in x; for $i = 0$ we have the empty point configuration $x = \emptyset$. We equip Ω with the smallest σ-algebra making the mappings $n_B(x) = n(x \cap B)$ measurable for all $B \in \mathcal{B}$. Further, ν denotes a Poisson point process on S with intensity measure $\beta\lambda$, where $\beta > 0$ is a parameter. In other words, if χ follows ν, then $n(\chi)$ is Poisson distributed with mean β, and conditionally on $n(\chi) = i$, the i points in χ are independent and each point has distribution λ. Specifically one may

think of $S = [0, 1]^2$ as the unit square, $\mathcal{B}$ as the Borel sets, and λ as the uniform distribution, in which case ν is a standard Poisson process. However, our general setting covers many other cases, including situations where χ can be interpret as a multitype or marked point process, see e.g. [1] and [15].

We assume that the distribution of χ is specified by an unnormalized density ϕ with respect to ν, so that ϕ is non-increasing in the following sense:

$$\phi(x \cup \xi) \leq \phi(x) \quad \text{for all } x \in S \text{ and } \xi \in S \setminus x \tag{1}$$

(we abuse the notation and write $x \cup \xi$ for $x \cup \{\xi\}$, $x \setminus \eta$ for $x \setminus \{\eta\}$, etc., when $x \in \Omega$, $\xi \in S \setminus x$, $\eta \in x$). This condition implies integrability of ϕ with respect to ν. Particularly, (1) is needed for the perfect simulation algorithms considered in this paper.

For a moment consider any unnormalized density ϕ with respect to ν. Local stability of ϕ means that for some constant $K > 0$ and all $x \in \Omega$ and all $\xi \in S \setminus x$,

$$\phi(x \cup \xi) \leq K\phi(x) \tag{2}$$

[20]. This is a basic assumption in many papers: for example, [7] establishes geometric ergodicity of a birth-death type Metropolis-Hastings algorithm for locally stable point processes [8]; and [14] show that it is a sufficient condition for applying dominated CFTP based on spatial birth-and-death processes and Metropolis-Hastings algorithms. Local stability is in fact a rather weak condition satisfied by most models considered in the statistical literature on spatial point processes, cf. the discussion in [14]. The concept of local stability is extended in [4] to cases where the dominating measure ν is not necessary a Poisson process.

As K can be absorbed into the parameter β we may without loss of generality set $K = 1$ in (2), whereby (1) is obtained. Below we consider just two examples where (1) is satisfied.

Example 1. Suppose that λ is the uniform distribution on $S = [0, 1]^2$ and

$$\phi(x) = \gamma^{s_R(x)} \tag{3}$$

taking $0^0 = 1$, where $s_R(x) = \sum_{\{\xi,\eta\} \subseteq x} \mathbb{1}[||\xi - \eta|| \leq R]$ is the number of R-close pairs of points in x, and where $0 \leq \gamma \leq 1$ and $R > 0$ are parameters. This specifies a Strauss process on the unit square [21, 11]. Clearly, ϕ is locally stable.

Example 2. Let S and λ be specified as in Example 1, but let now

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

where $U_x = \cup_{\xi \in x} \text{ball}(\xi, R)$ is the union of closed balls with centers $\xi \in x$ and of radius R, where $R > 0$ and $\gamma > 0$ are parameters. This is an area interaction point process [22, 2]. The process is said to be attractive for $\gamma > 1$, and repulsive for $\gamma < 1$, since

$$\phi(x \cup \xi)/\phi(x) = \gamma^{-\lambda(U_{x\cup\xi} \setminus U_x)} \tag{4}$$

is increasing ($\gamma > 1$) or decreasing ($\gamma < 1$) in x. It follows from (4) that (1) holds in the attractive case, but not in the repulsive case. If $\gamma < 1$ we therefore redefine ν as a Poisson process with intensity measure $(\beta/\gamma^{\pi R^2})\lambda$, and redefine ϕ by

$$\phi(x) = \gamma^{n(x)\pi R^2 - \lambda(U_x)}.$$

Then (1) is satisfied.

3 Coupling construction

Below we construct two time-stationary and reversible spatial birth-and-death processes $X = \{X_t : t \in \mathbb{R}\}$ and $D = \{D_t : t \in \mathbb{R}\}$ with equilibrium distributions given by $X_t \sim \phi$ (with respect to ν) and $D_t \sim \nu$. The two processes are coupled so that D dominates X in the sense that

$$X_t \subseteq D_t \quad \text{for all } t \in \mathbb{R}. \tag{5}$$

This is obtained by letting (D, X) be a continuous time Markov processes with the following types of transitions: either a new point is added to both D and X, or a birth happens in D but not in X, or a point in X is deleted from both D and X, or a point in D but not in X is deleted. The coupling construction is underlying the perfect samplers in Sections 4–6.

We first specify how D_t can be generated forwards in time $t \geq 0$. For any $x \in \Omega$ and $t \geq 0$, if we condition on that $D_t = x$, and τ is the waiting time for the next transition in D after time t, then

- τ is exponentially distributed with mean $1/(\beta + n(x))$;
- with probability $\beta/(\beta + n(x))$ a birth happens in D at time $t + \tau$: draw a point $\xi_{t+\tau} \sim \lambda$ and set $D_{t+\tau} = x \cup \xi_{t+\tau}$ — for later use in the coupling construction, generate also a "mark" $R_{t+\tau} \sim \text{Uniform}[0, 1]$;

- else a death happens in D at time $t + \tau$:
  draw randomly uniformly a point $\eta_{t+\tau}$ from x and set $D_{t+\tau} = x \setminus \eta_{t+\tau}$.

Furthermore, the conditional distributions of τ, the event of a birth or death, and the generation of either $(\xi, R_{t+\tau})$ or η are assumed to be mutually independent and independent of the previous history given $D_t = x$. In other words, a birth of a new point in D happens with rate β and follows the distribution λ, each point in D dies with rate 1, and births and deaths in D are independent events.

It is easily verified that $\{D_t : t \geq 0\}$ is reversible with invariant distribution ν, and all the marks associated to the birth times are mutually independent and independent of $\{D_t : t \geq 0\}$. Hence we can easily start in equilibrium $D_0 \sim \nu$, and by reversibility, D_t is easily generated backwards in time $t < 0$ together with the associated marks for (forwards) births $D_t = D_{t-} \cup \xi_t$, where $t-$ refers to the situation just before time t. Moreover, it is not hard to verify that D is non-explosive and $\emptyset$ is an ergodic atom at which D regenerates, see Fig. 1.

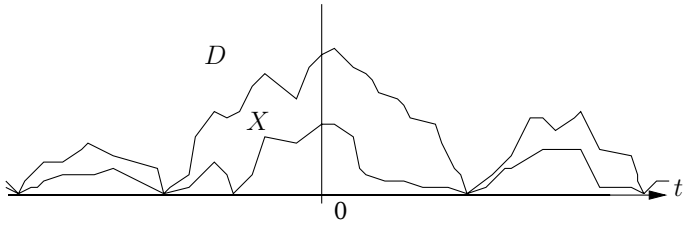

Figure 1: Upper curve: the dominating spatial birth-and-death process D; lower curve: the spatial birth-and-death process X. The horizontal axis is time and the vertical line corresponds to the state space Ω with $\emptyset$ placed at the bottom. Each time $D_t = \emptyset$, the jump process (D, X) regenerates.

We show next how X_t can be coupled to D_t forwards in time $t \in \mathbb{R}$. For $x \in \Omega$ and $\xi \in S \setminus x$, define

$$b(x, \xi) = \phi(x \cup \xi)/\phi(x) \tag{6}$$

(setting $0/0 = 0$). By (1), $b \leq 1$, and βb is a so-called Papangelou conditionally intensity [10]. Consider a cycle of D given by $\{D_t : \tau_1 \leq t < \tau_2\}$ where τ_1 and τ_2 are two successive times at which D enters $\emptyset$, i.e. $D_{\tau_1-} \neq \emptyset$, $D_{\tau_1} = \emptyset$, and $\tau_2 = \inf\{t > \tau_1 : D_{t-} \neq \emptyset, D_t = \emptyset\}$ (with probability one, D enters $\emptyset$ infinite often, and $-\infty < \tau_1 < \tau_2 < \infty$). Then set $X_{\tau_1} = \emptyset$ and construct X_t forwards

in time $t \in (\tau_1, \tau_2)$, according to the following rules:

$$\begin{aligned} D_t = D_{t-} &\Rightarrow X_t = X_{t-} \\ D_t = D_{t-} \cup \xi_t &\Rightarrow X_t = \begin{cases} X_{t-} \cup \xi_t & \text{if } R_t \le b(X_{t-}, \xi_t) \\ X_{t-} & \text{otherwise} \end{cases} \\ D_t = D_{t-} \setminus \eta_t &\Rightarrow X_t = X_{t-} \setminus \eta_t. \end{aligned}$$

Using this coupling construction for all cycles of D, (5) is obviously satisfied.

It follows immediately from the coupling construction that X is a spatial birth-and-death process with birth rate βb and death rate 1. As ϕ satisfies the detailed balance condition $\phi(x)b(x, \xi) = \phi(x \cup \xi)$, we obtain that X is reversible with invariant (unnormalized) density ϕ. Hence, since (D, X) is time-stationary, X_t follows ϕ for any fixed time $t \in \mathbb{R}$.

In the case where $\phi(x) = \alpha^{n(x)}$ with $0 \le \alpha \le 1$, we have that (D, X) is reversible, X and $\{D_t \setminus X_t : t \in \mathbb{R}\}$ are independent spatial birth-and-death processes, and for any fixed time $t \in \mathbb{R}$, X_t and $D_t \setminus X_t$ are independent Poisson processes with intensity measures $\alpha\beta\lambda$ and $(1-\alpha)\beta\lambda$, respectively. However, it is easily checked that (D, X) is in general not reversible, and apart from the Poisson case above, it seems complicated to obtain a closed form expression for the equilibrium distribution of (D, X).

4 The simple dominated CFTP algorithm

A jump in D happens when $D_t \neq D_{t-}$, in which case t is called a jump time. In order to generate a simulation of $X_0 \sim \phi$ we need only to consider the jump chain (or embedded Markov chain) of $\{D_t : t < 0\}$, its associated marks for forwards births, and the states of X when $\{D_t : t < 0\}$ jumps. This is described in detail below.

Let $\ldots, Z_{-2}, Z_{-1}, Z_0$ denote the jump chain of $\{D_t : t < 0\}$ so that $Z_0 = D_0 \sim \nu$. This can be generated backwards in time together with the associated marks for forwards births as follows. For $i = 0, -1, -2, \ldots,$

- with probability $\beta/(\beta + n(Z_i))$ make a backwards birth:
  draw $\eta_i \sim \lambda$ and set $Z_{i-1} = Z_i \cup \eta_i$;
- else make a backwards death:
  draw randomly uniformly $\xi_i \in Z_i$, set $Z_{i-1} = Z_i \setminus \xi_i$, and generate the associated mark $R_i \sim \text{Uniform}[0, 1]$ for the forwards birth $Z_i = Z_{i-1} \cup \xi_i$.

Let $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$ and define

$$T_0 = \inf\{i \in \mathbb{N}_0 : Z_{-i} = \emptyset\}.$$

Furthermore, define recursively $Y_{-T_0}, \ldots, Y_0$, setting $Y_{-T_0} = \emptyset$ and using the rules

$$Z_i = Z_{i-1} \cup \xi_i \quad \Rightarrow \quad Y_i = \begin{cases} Y_{i-1} \cup \xi_i & \text{if } R_i \le b(Y_{i-1}, \xi_i) \\ Y_{i-1} & \text{otherwise} \end{cases} \tag{7}$$

$$Z_i = Z_{i-1} \backslash \eta_i \quad \Rightarrow \quad Y_i = Y_{i-1} \backslash \eta_i \tag{8}$$

for $i = -T_0 + 1, \ldots, 0$. Let $\cdots < \tau_{-2} < \tau_{-1} < \tau_0$ denote the jump times of D before time 0. Then $(X_{\tau_{-T_0}}, \ldots, X_{\tau_0})$ and $(Y_{-T_0}, \ldots, Y_0)$ follow the same distribution. Especially, $Y_0 \sim \phi$, since $X_{\tau_0} = X_0$ almost surely. This suggests the following perfect sampler.

The simple dominated CFTP algorithm.

1. Generate backwards $Z_0, \ldots, Z_{-T_0}$, starting with $Z_0 \sim \nu$, and generate the associated marks R_i for forwards births $Z_i = Z_{i-1} \cup \xi_i$;
2. set $Y_{-T_0} = \emptyset$ and construct $Y_{-T_0+1}, \ldots, Y_0$ as in (7)–(8);
3. return $Y_0 \sim \phi$; see Fig. 2.

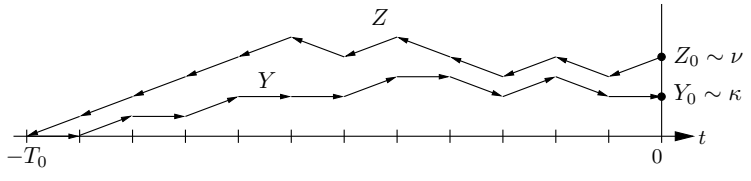

Figure 2: Illustration of the simple dominated CFTP algorithm.

Proposition. The mean number of steps involved in the backwards construction of the simple dominated CFTP algorithm is bounded from below by

$$\mathbb{E}T_0 \ge \exp(\beta) - 1/2. \tag{9}$$

Proof. Let $Z_1, Z_2, \ldots$ denote the jump chain of $\{D_t : t > 0\}$. Set $M_i = n(Z_i)$ for $i \in \mathbb{Z}$, $T_0^- = T_0$, $T_0^+ = \inf\{i \in \mathbb{N}_0 : M_i = 0\}$, $L = \inf\{i \in \mathbb{N} : M_{i+T_0^+} = 0\}$, and $L_0 = T_0^- + T_0^+$ if $M_0 \neq 0$ and $L_0 = L$ otherwise. By time-stationarity,

$$\mathbb{E}L_0 \ge \mathbb{E}L = 1/\pi_0, \tag{10}$$

where π denotes the invariant probability density function of M. By reversibility, T_0^- and T_0^+ are identically distributed, so

$$2\mathbb{E}T_0 = \mathbb{E}\big(T_0^- + T_0^+\big) = \mathbb{E}\big(L_0\mathbf{1}[M_0 \neq 0]\big). \tag{11}$$

Further,

$$\mathbb{E}\big(L_0\mathbf{1}[M_0 = 0]\big) = \pi_0\mathbb{E}(L_0|M_0 = 0) = \pi_0\mathbb{E}L = 1. \tag{12}$$

Combining (10)–(12) we obtain that

$$\mathbb{E}T_0 \geq (1/\pi_0 - 1)/2.$$

Finally, by detailed balance of M,

$$\pi_i\beta/(\beta + i) = \pi_{i+1}(i + 1)/(\beta + i + 1)$$

so by induction

$$\pi_{i+1} = \pi_0\beta^i(\beta + i + 1)/(i + 1)!, \quad i \in \mathbb{N}_0,$$

whereby $1/\pi_0 = 2\exp(\beta)$, and so (9) follows.

Remark. Since $\mathbb{E}T_0$ is at least exponentially growing in β, the simple dominated CFTP algorithm is infeasible for real applications of interest. For instance, if $\beta = 100$, then $\mathbb{E}T_0 \geq e^{100} - 1/2 \approx 2.7 \times 10^{43}$.

5 Upper and lower processes

A much faster perfect simulation algorithm is given in [14], using upper and lower processes $U^j = \{U_j^j, \ldots, U_0^j\}$ and $L^j = \{L_j^j, \ldots, L_0^j\}$ which are started at times $j = 0, -1, -2, \ldots$. For each j, the upper and lower processes are constructed as follows. Initially set $U_j^j = Z_j$ and $L_j^j = \emptyset$. For $i = j+1, \ldots, 0$, if $z = Z_{i-1}$, $u = U_{i-1}^j$, and $l = L_{i-1}^j$, use the rules

$$Z_i = z\backslash\eta_i \quad \Rightarrow \quad U_i^j = u\backslash\eta_i \quad \text{and} \quad L_i^j = l\backslash\eta_i, \tag{13}$$

$$Z_i = z \cup \xi_i \quad \Rightarrow \quad U_i^j = \begin{cases} u \cup \xi_i & \text{if } R_i \leq \alpha_{\max}(u, l, \xi_i) \\ u & \text{otherwise} \end{cases}$$

$$\text{and} \quad L_i^j = \begin{cases} l \cup \xi_i & \text{if } R_i \leq \alpha_{\min}(u, l, \xi_i) \\ l & \text{otherwise,} \end{cases} \tag{14}$$

where

$$\alpha_{\max}(u, l, \xi) = \max\{b(x, \xi) : l \subseteq x \subseteq u\} \tag{15}$$

and

$$\alpha_{\min}(u, l, \xi) = \min\{b(x, \xi) : l \subseteq x \subseteq u\}. \tag{16}$$

Notice that U^j, L^j, U^{j-1}, L^{j-1}, ... are coupled by the same R_i, ξ_i, η_i for $i > j$.

The construction in (13)–(16) ensures the sandwiching property

$$L_i^j \subseteq Y_i \subseteq U_i^j \subseteq Z_i, \quad j \le i \le 0, \tag{17}$$

the funneling property

$$L_i^{j'} \subseteq L_i^j \subseteq U_i^j \subseteq U_i^{j'}, \quad j \le j' \le i \le 0, \tag{18}$$

and the coalescence property

$$L_i^j = U_i^j \quad \Rightarrow \quad L_{i'}^j = U_{i'}^j \quad \text{for } j \le i \le i' \le 0, \tag{19}$$

see Fig. 3. The sandwiching property explains why the U^j and L^j are called upper and lower processes: they bound the "target process" Y. The definitions (15)–(16) seem natural as they provide the minimal upper and maximal lower processes so that (17) is satisfied for all possible realizations of the marks R_i. By (17) and (19), once a pair of upper and lower processes have coalesced, they stay in coalescence, and at time 0 they are equal to $Y_0 \sim \phi$.

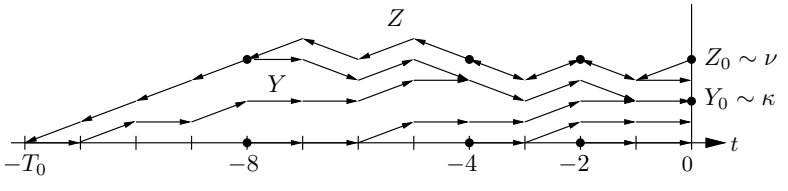

Figure 3: Illustration of sandwiching, funnelling, and coalescence properties for $T_0 = 12$ and $j = -2, -4, -8$ in (17)–(19).

The time

$$T = \inf\{j \in \mathbb{N}_0 : U_0^{-j} = L_0^{-j}\}$$

is called the true coalescence time for upper and lower processes. The funneling property (18) suggests that instead of searching for T it may be advantageous

to search for a larger coalescence time. Therefore, consider any sequence of (possibly random) integers $\dots j_2 < j_1 < 0$ such that $\lim_{k\to\infty} j_k = -\infty$, and let

$$T_{\{j_k\}} = \inf\{-j_k : U_0^{j_k} = L_0^{j_k}\}$$

be the first time that $U_0^j = L_0^j = Y_0$ when pairs of upper and lower processes are started at times $j = j_1, j_2, \dots$. Further, let

$$T_{\min} = \inf\{i \in \mathbb{N}_0 : Z_{-i} \cap Z_0 = \emptyset\}$$

be the time just before the first point in Z_0 is born. For $-T_{\min} < j \le 0$, we have that $U_0^j \supseteq Z_j \cap Z_0 \neq \emptyset$ and $L_0^j \cap Z_j = \emptyset$, so clearly

$$T_{\min} \le T \le T_{\{j_k\}} \le T_0. \tag{20}$$

For efficiency reasons a doubling scheme is usually used [19, 24], i.e. $j_k = -2^{k-1}n$, where $n \in \mathbb{N}$ is chosen by the user; then we write T_n for $T_{\{j_k\}}$. Typically in applications $T_n \ll T_0$, cf. Section 7. Taking (20) into account, we propose to replace n by $T_{\min}$ in the doubling scheme; then we write T_* for $T_{\{j_k\}}$. See also the empirical results in Section 7.

Given a sequence of (possibly random) integers $\dots j_2 < j_1 < 0$ such that $\lim_{k\to\infty} j_k = -\infty$, we have the following perfect sampler, where we set $j_0 = 0$.

The dominated CFTP algorithm based on upper and lower processes

1. Generate $Z_0 \sim \nu$;
2. repeat the following steps 3.–4. for $k = 1, 2, \dots$ until $U_0^{j_k} = L_0^{j_k}$;
3. generate backwards $Z_{j_{k-1}-1}, \dots, Z_{j_k}$ and generate the associated marks $R_i \sim$ Uniform[0, 1] each time $Z_i \setminus Z_{i-1} \neq \emptyset,\ j_k < i \le j_{k-1}$;
4. generate forwards $(U_{j_k}^{j_k}, L_{j_k}^{j_k}), \dots, (U_0^{j_k}, L_0^{j_k})$ as in (13)–(14);
5. return $U_0^{-T_{\{j_k\}}} \sim \phi$.

The calculation of $\alpha_{\max}$ and $\alpha_{\min}$ is particularly simple in the following cases. A point process is attractive if

$$b(x, \xi) \le b(y, \xi) \qquad \text{whenever} \qquad x \subseteq y, \xi \notin y, \tag{21}$$

and repulsive if

$$b(x, \xi) \ge b(y, \xi) \qquad \text{whenever} \qquad x \subseteq y, \xi \notin y. \tag{22}$$

For the Strauss and area interaction point processes (Examples 1 and 2), either (21) or (22) is satisfied. In the attractive case (21), $\alpha_{\max}(u, l, \xi) = b(u, \xi)$ and $\alpha_{\min}(u, l, \xi) = b(l, \xi)$, while in the repulsive case (22), $\alpha_{\max}(u, l, \xi) = b(l, \xi)$ and $\alpha_{\min}(u, l, \xi) = b(u, \xi)$. Note that it is only in the attractive case that U^j and L^j are individual Markov chains.

It other situations it may be quite time consuming to calculate $\alpha_{\max}$ and $\alpha_{\min}$ by (15) and (16). For instance, if $b(x, \xi) = b_a(x, \xi)b_r(x, \xi)$ factorizes into two terms, where $b_a(x, \xi) \leq K_a$ is increasing in x, $b_r(x, \xi) \leq K_r$ is decreasing in x, and $K_a K_r \leq 1$, it may be convenient to redefine $\alpha_{\max}$ and $\alpha_{\min}$ by

$$\alpha_{\max}(u, l, \xi) = b_a(u, \xi)b_r(l, \xi) \quad \text{and} \quad \alpha_{\min}(u, l, \xi) = b_a(l, \xi)b_r(u, \xi).$$

Since (17)–(19) are satisfied with this choice of $\alpha_{\max}$ and $\alpha_{\min}$, the algorithm still works.

Example 1 (continued): Perfect simulations of different Strauss processes with $\gamma = 1$ (the Poisson case), $\gamma = 0.5$, and $\gamma = 0$ are shown in Fig. 4, using the same dominating process (and associated marks) in all three cases. Due to the thinning procedure in the algorithm, the point pattern with $\gamma = 1$ contains the two others. The point pattern with $\gamma = 0$ does not contain the point pattern with $\gamma = 0.5$, because the Strauss process is repulsive.

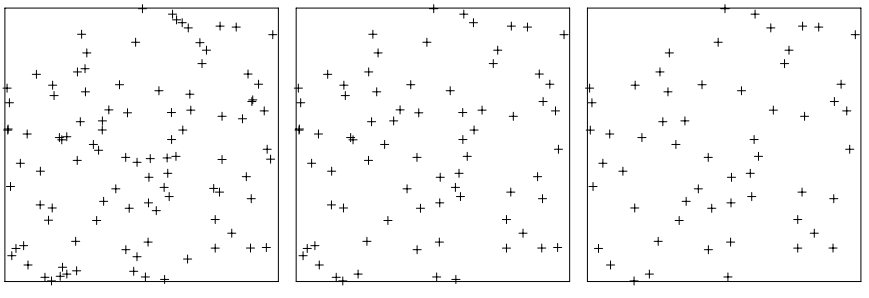

Figure 4: Simulation of a Strauss process on $S = [0, 1]^2$, when $\beta = 100$, $R = 0.05$, and $\gamma = 1, 0.5, 0$ (from left to right).

6 Clan of ancestors

In this section we consider an alternative algorithm due to [5]. For simplicity we assume that S is a metric space and ϕ has finite range of interaction, i.e. there exists an $R < \infty$ such that for any $x \in \Omega$ and $\xi \in S\backslash x$, $b(x, \xi) =$

$b(x \cap \text{ball}(\xi, R), \xi)$, where $\text{ball}(\xi, R)$ denotes the ball with center ξ and radius R. This is fulfilled in Examples 1 and 2.

In order to understand the following definitions it may be useful to consider Fig. 5 and to keep in mind how the simple dominated CFTP algorithm (Section 4) works. For $\xi \in \cup_{i \le 0} Z_i$, let $I(\xi)$ be the time at which ξ was born, i.e. $I(\xi) = i$ if $\xi = \xi_i$ in (7). We call

$$\text{an}^1(\xi) = Z_{I(\xi)-1} \cap \text{ball}(\xi, R)$$

the first generation of ancestors of ξ, define recursively the jth generation of ancestors of ξ by

$$\text{an}^j(\xi) = \cup_{\eta \in \text{an}^{j-1}(\xi)} \text{an}^1(\eta), \quad j = 2, 3, \ldots,$$

and call $\text{an}(\xi) = \cup_{j \in \mathbb{N}} \text{an}^j(\xi)$ the ancestors of ξ. If $I(\xi) = i$, then $Y_{i-1} \cap \text{ball}(\xi, R) \subseteq \text{an}(\xi)$, so the ancestors of $\xi = \xi_i$ are the only points in Z which are needed in (7) in order to determine whether or not $\xi_i \in Y_i$. Hence Y_0 depends only on Z through $Z_C = C(Z_0) \cup Z_0$, where

$$C(Z_0) = \cup_{\xi \in Z_0} \text{an}(\xi)$$

is called the *clan of ancestors* of Z_0. Finally, let

$$T_C = \inf\{i \in \mathbb{N}_0 : Z_{-i} \cap Z_C = \emptyset\}$$

specify the time interval in which the points in Z_C are living. Then $T_C \le T_0$, and Y_0 is unaffected if we set $Y_{-T_C} = \emptyset$ and generate Y_i forwards in time $i \ge -T_C$ as usual, but considering only the transitions in $Z_{-T_C} \cap Z_C, \ldots, Z_0 \cap Z_C$.

The dominated CFTP algorithm based on the clan of ancestors

1. Generate backwards $Z_0, \ldots, Z_{-T_C}$, i.e. starting with $Z_0 \sim \nu$;
2. set $Y_{-T_C} = \emptyset$ and generate forwards $Y_{-T_C+1}, \ldots, Y_0$ as in (7)– (8), but so that $Y_{i+1} = Y_i$ is unchanged whenever $Z_{i+1} \cap Z_C = Z_i \cap Z_C$ is unchanged;
3. return $Y_0 \sim \phi$.

It is not hard to see that $T \le T_C$, so

$$T \le T_C \le T_0. \tag{23}$$

Note that T_C depends only on b through R, and no monotonicity properties such as (21) and (22) are required. The algorithm can easily be modified to perfect Metropolis-Hastings simulation of locally stable point processes [14], and to perfect Gibbs sampling of the [22] model [9] and related models [6]. The case of the Widom-Rowlinson model turns out to be particularly simple.

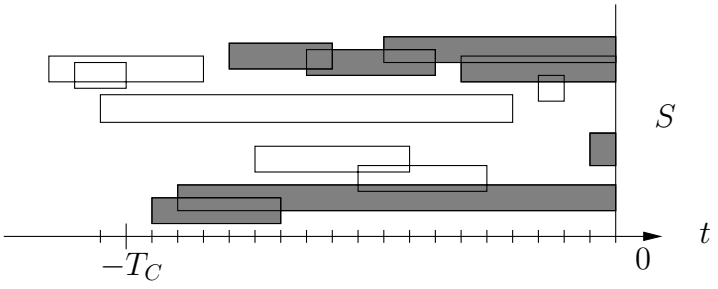

Figure 5: Example of a clan of ancestors when S is a line segment. The points in D agree with the midpoints of the vertical edges of the rectangles. Each horizontal edge of a rectangle shows the life time of the corresponding point in D. The vertical edges are all of length R. Shaded rectangles represent members of the clan.

7 Empirical findings

In this section we present some empirical findings for the dominated CFTP algorithm based on upper and lower processes (Section 5) and the clan algorithm (Section 6), respectively. The algorithms are applied on a Strauss process defined on the unit square with $\beta = 100$ and $R > 0$ (Example 1). Note that as the interaction parameter γ increases, the interaction/repulsion between the points in the Strauss process decreases. Below we consider three values of γ: $\gamma = 0$ (a so-called hard core process), $\gamma = 0.5$, and $\gamma = 1$ (a Poisson process on the unit square with rate $\beta = 100$).

First we consider the algorithm based on upper and lower processes, using the doubling scheme with either $n = 1$ or n replaced by $T_{\min}$. Recall that T_1 and T_* denote the corresponding coalescence times, cf. Section 5. The number of steps involved in the backward-forwards construction in the two cases are given by

$$N_1 \equiv T_1 + (1 + 2 + 4 + \ldots + T_1) = 3T_1 - 1$$

and

$$N_{\min} \equiv T_* + (T_{\min} + 2T_{\min} + 4T_{\min} + \ldots + T_*) = 3T_* - T_{\min},$$

respectively. It makes sense to compare N_1 and $N_{\min}$ because the "basic algorithm" is the same in the two cases.

The left plot in Fig. 6 shows how the means $\mathbb{E}N_1$ and $\mathbb{E}N_{\min}$ depend on $R > 0$ when $\gamma = 0$. Each mean is estimated by the empirical average based on 500

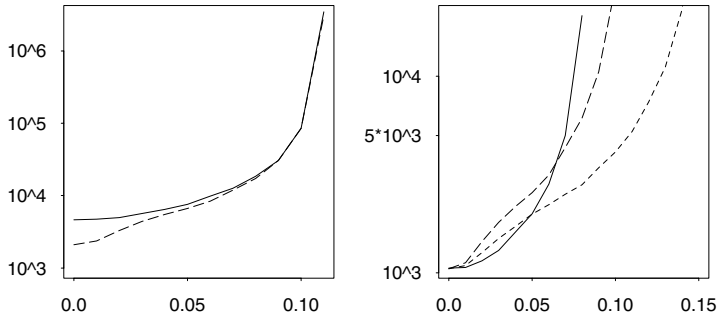

Figure 6: Various mean values related to the CFTP algorithms, where each mean is estimated from 500 independent runs. Left plot: $\mathbb{E}N_1$ (full line) and $\mathbb{E}N_{\min}$ (dotted line) versus R. Right plot: $\mathbb{E}T_C$ (full line) and $\mathbb{E}T_*$ when $\gamma = 0$ (upper dotted line) and $\gamma = 0.5$ (lower dotted line) versus R.

independent runs of the algorithm. For all values of R in the plot, $\mathbb{E}N_{\min} < \mathbb{E}N_1$, but the difference decreases as R increases. Based on this and other results (not shown here) we prefer to replace n by $T_{\min}$ in the doubling scheme.

Next we compare the dominated CFTP algorithm based on upper and lower processes and the clan algorithm. As these algorithms are not immediately comparable, there is little sense in comparing the number of steps involved in the backwards-forwards construction in the two algorithms. Instead we just consider the means $\mathbb{E}T_*$ and $\mathbb{E}T_C$ for $R > 0$ and either $\gamma = 0$ or $\gamma = 0.5$, though this is of course not telling the whole story about which algorithm is the fastest.

The right plot in Fig. 6 shows $\mathbb{E}T_*$ and $\mathbb{E}T_C$ versus R when $\gamma = 0$ and $\gamma = 0.5$, respectively. Note that T_C does not depend on γ, and each mean in the plot is estimated by the empirical average based on 500 independent runs of the algorithm. All means in the plot are much smaller than $\mathbb{E}T_0 \geq e^{100} - 1/2$, cf. (9). As expected the means agree as R tends to 0, and $\mathbb{E}T_*$ decreases as γ increases. For both $\gamma = 0$ and $\gamma = 0.5$, it is only for rather small values of γ that $\mathbb{E}T_*$ is larger than $\mathbb{E}T_C$. The picture changes as γ tends to 1, since in the limit T_* agrees with $T_{\min}$ which is smaller than T_C, cf. (20) and (23). Furthermore, as R increases, $\mathbb{E}T_*$ becomes much smaller than $\mathbb{E}T_C$.

We have also investigated empirically how $\mathbb{E}T_C$ and $\mathbb{E}T_*$ depend on β, and obtained similar conclusions as above. Further empirical results for the Strauss process and other locally stable point processes can be found in [3].

Finally, all things considered our conclusion is that the dominated CFTP algorithm based on upper and lower processes using the doubling scheme with n replaced by $T_{\min}$ seems to be the best choice.

8 Perfect simulation of infinite point processes

Often one considers point processes with infinitely many points contained in an "infinite volume" such as $\mathbb{R}^d$. In order to avoid edge-effects, a perfect sample within a bounded region may be achieved by extending simulations both backwards in time and space [12, 5, 6]. This is sometimes possible, for example if b is sufficiently close to 1 and the interaction radius R is sufficiently small. The constructions in the abovementioned papers are rather straightforward, but particularly the algorithm in [5] allows a detailed mathematical analysis. Such coupling constructions may be of great theoretical interest, but in our opinion they remain so far unpractical for applications of real interest.

Acknowledgment. This paper 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. JM also acknowledges the invitation and support for attending the 5th Brazilian School of Probability.

References

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

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

[3] Berthelsen, K. K. & Møller, J. 2001a. Perfect simulation and inference for spatial point processes. In preparation.

[4] Berthelsen, K. K. & Møller, J. 2001b. Spatial jump processes and perfect simulation, *Technical report*, R-01-2008, Department of Mathematical Sciences, Aalborg University. Submitted.

[5] Fernández, R., Ferrari, P. A. & Garcia, N. L. 2000. Perfect simulation for interacting point processes, loss networks and Ising models. Submitted to *Stoch. Process. Appl*.

[6] 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.

[7] 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.

[8] Geyer, C. J. & Møller, J. 1994. Simulation procedures and likelihood inference for spatial point processes, *Scand. J. Statist.* **21**: 359–373.

[9] 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.

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

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

[12] Kendall, W. S. 1997. Perfect simulation for spatial point processes, *Proc. ISI 51st session, Istanbul* **3**: 163–166.

[13] 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.

[14] 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.

[15] Lieshout, M. N. M. van 2000. *Markov point processes and their applications*, Imperial College Press, London.

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

[17] Møller, J. 2001. A review of perfect simulation in stochastic geometry, *in* C. C. H. I. V. Basawa & R. L. Taylor (eds), *Selected Proceedings of the Symposium on Inference for Stochastic Processes*, Vol. 37, IMS Lecture Notes & Monographs Series, pp. 333–355.

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

[19] 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.

[20] Ruelle, D. 1969. *Statistical Mechanis: Rigorous Results*, W. A. Benjamin, Reading, Massachusetts.

[21] Strauss, D. J. 1975. A model for clustering, *Biometrika* **62**: 467–475.

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

[23] Wilson, D. B. 2000a. How to couple from the past using a read-once source of randomness, *Random Structures and Algorithms* **16**: 85–113.

[24] Wilson, D. B. 2000b. Layered multishift coupling for use in perfect sampling algorithms (with a primer on CFTP), *in* N. Madras (ed.), *Monte Carlo Methods*, Vol. 26 of *Fields Institute Communications*, Amer. Stat. Soc., pp. 141–176.

Kasper K. Berthelsen and Jesper Møller
Department of Mathematical Sciences
Aalborg University
Fredrik Bajers Vej 7G, DK-9220 Aalborg
DENMARK