THINNING SPATIAL POINT PROCESSES INTO POISSON PROCESSES

JESPER MØLLER,* *Aalborg University*

FREDERIC PAIK SCHOENBERG,** *University of California*

Abstract

In this paper we describe methods for randomly thinning certain classes of spatial point processes. In the case of a Markov point process, the proposed method involves a dependent thinning of a spatial birth-and-death process, where clans of ancestors associated with the original points are identified, and where we simulate backwards and forwards in order to obtain the thinned process. In the case of a Cox process, a simple independent thinning technique is proposed. In both cases, the thinning results in a Poisson process if and only if the true Papangelou conditional intensity is used, and, thus, can be used as a graphical exploratory tool for inspecting the goodness-of-fit of a spatial point process model. Several examples, including clustered and inhibitive point processes, are considered.

Keywords: Area-interaction point process; clans of ancestors; coupling; Cox process; dependent and independent thinning; Markov point process; Papangelou conditional intensity; Poisson process; spatial birth-and-death process; Thomas process

2010 Mathematics Subject Classification: Primary 60G55; 62M30

1. Introduction

A useful graphical method for inspecting a fitted model $\hat{\lambda}$ for the conditional intensity λ of a temporal or space–time point process, where the conditioning is based on all observations at all previous times, is via random thinning. This technique, described in Schoenberg (2003), involves keeping each point τ_i independently with a probability proportional to $\hat{\lambda}(\tau_i)$. If the true conditional intensity λ is used in place of $\hat{\lambda}$, then the resulting process is a homogeneous Poisson process (see Schoenberg (2003)). The residual points, obtained after thinning using the estimate $\hat{\lambda}$ in place of λ, may readily be inspected for homogeneity using standard methods.

The question of how to extend this technique to the case of a purely spatial point process remains open. For example, Markov (or Gibbs) point processes are characterized by their Papangelou conditional intensities, where the conditioning is based on the observations at all other locations; see, e.g. Ripley and Kelly (1977) and Baddeley and Møller (1989). While Schoenberg (2005) conjectured that a thinning method similar to that used for space–time point processes should be valid for purely spatial point processes as well, this conjecture was shown in Baddeley *et al.* (2005, p. 664) to be false. Schoenberg and Zhuang (2008) introduced a method for thinning spatial point processes based on considering all possible subsets of points and selecting among these subsets with the appropriate probability, but this method relies

347

on certain rather restrictive assumptions and also, since it requires $O(2^n)$ computations for thinning a point pattern consisting of n points, is excessively computationally intensive and impractical for all but very small values of n. Furthermore, the method of Schoenberg and Zhuang (2008) only applies to spatial point processes with known Papangelou intensities, but not to many important classes of point processes, including Cox processes (see Cox (1955)), whose Papangelou intensities may be intractable.

Here, we introduce alternative methods for thinning a spatial point process X. For simplicity and specificity, we assume that X is finite, simple (i.e. has no multiple points), and defined on a bounded Borel set $S \subset \mathbb{R}^k$ ($k \in \{1, 2, \ldots\}$) so that X can be considered as a random subset of S. This setting covers most cases of practical interest, but our methods can easily be extended both to nonsimple Markov point processes defined on a general state space and using an exponential state space setting (see Carter and Prenter (1972), Preston (1977), and Ripley and Kelly (1977)), and to nonsimple Cox processes which may be defined on $\mathbb{R}^k$ or more complicated state spaces. For background material on spatial point processes, particularly Markov point processes and Cox processes, see Møller and Waagepetersen (2004), (2007) and the references therein.

In Section 2 we consider the case where X is a Markov point process and $\lambda(x, \xi)$ denotes its Papangelou conditional intensity, which is assumed to be bounded from below by a nonnegative deterministic function $\rho(\xi)$ for all finite point configurations $x \subset S$ and points $\xi \in S \setminus x$. We consider $X = X_0$ as an equilibrium state of a spatial birth-and-death process X_t at time 0, and we consider coupling this with another spatial birth-and-death process W_t with equilibrium distribution given by a Poisson process with intensity function ρ such that $W_t \subseteq X_t$ for all times $t \in \mathbb{R}$. The proposed method involves first simulating X_t backwards for $t \leq 0$. At the ith jump backwards in time, with respect to the neighbour relation used for defining the Markov point process, we identify the ith generation ancestors associated with the points in X, and the simulation is stopped the first time T before time 0 such a generation is empty. Second, $W_T = \varnothing$ is the empty point configuration, and we use a dependent thinning technique to obtain W_t for $T \leq t \leq 0$. This procedure has some similarity to perfect simulation algorithms for spatial point processes (see Kendall (1998), Kendall and Møller (2000), and Fernández et al. (2002)), but in contrast these algorithms assume an upper bound $V(\xi) \geq \lambda(x, \xi)$ (the so-called local stability condition) and use this to couple X_t with a birth-and-death process D_t such that $X_t \subseteq D_t$, $t \in \mathbb{R}$, and each D_t is a Poisson process with intensity function V. In fact, our procedure needs only to first simulate the jump chain of births and deaths for X_t with t running from time 0 to time T, then examine what happens at these jump times in W_t with t running from T to 0, and finally return W_0. We illustrate how this works for an area-interaction point process (see Baddeley and Van Lieshout (1995)).

In Section 3 we consider the case where $\lambda(\xi)$ is a nonnegative random function, X is a Cox process driven by λ, and λ is bounded from below by a nonnegative deterministic function ρ (note that λ is hence not the Papangeleou conditional intensity of X). This case is much simpler, since X can be viewed as a superposition of a Cox process driven by $\lambda - \rho$ and an independent Poisson process with intensity function ρ. We show how an independent thinning technique applies to obtain a Poisson process with intensity function ρ. This is illustrated in the case of a modified Thomas process (see Thomas (1949)).

Our main result is the following.

Theorem 1. *For both a Markov point process and a Cox process model, the thinning results in a Poisson process with intensity function ρ if and only if the true λ is used.*

A consequence of Theorem 1 is that thinning can be used as a graphical exploratory tool for visualizing the fit of a spatial point process model. In Sections 2.4 and 3.3 we show illustrations, using the L-function to check if the thinned process is Poisson. The L-function is a stabilized variance version of Ripley's K-function (see Ripley (1976), (1977)) and both can be extended to the inhomogeneous case (see Baddeley *et al.* (2000)), though many other techniques for checking whether a point process is Poisson have been developed; see, e.g. Cressie (1993) and Stoyan *et al.* (1995). In practical applications, in both Bayesian and frequentist settings, λ is typically replaced by an estimate $\hat{\lambda}$ (for a discussion on how such estimates can be obtained, see Møller and Waagpetersen (2004), (2007) and the references therein), but, for simplicity and specificity, in all examples we consider simulated data where the true λ is known and hence not estimated.

2. Thinning Markov point processes

Suppose that X has a density f with respect to the homogeneous Poisson process on S with intensity equal to unity. We assume that f is hereditary, that is, $f(x \cup \{\xi\}) > 0$ whenever $f(x) > 0$ for a finite point configuration $x \subset S$ and a point $\xi \in S \setminus x$. The Papangelou conditional intensity (see Kallenberg (1984)) is defined by

$$\lambda(x, \xi) = \frac{f(x \cup \{\xi\})}{f(x)},$$

where the hereditary condition ensures that f and λ are in a one-to-one correspondence. (Here and throughout this paper, we use the convention that $0/0 = 0$.) Heuristically, $\lambda(x, \xi)\,\mathrm{d}\xi$ can be interpreted as the conditional probability of X having a point in an infinitesimally small region containing ξ and of size $\mathrm{d}\xi$ given that the rest of X is x.

Let '$\sim$' denote a symmetric relation on S, and let $N_\xi = \{\eta \in S : \xi \sim \eta\}$ denote the neighbours to $\xi \in S$. If, for any finite $x \subset S$ and $\xi \in S \setminus x$, $\lambda(x, \xi)$ depends only on ξ and N_ξ, then X is said to be a Markov point process (with respect to '$\sim$'); cf. Ripley and Kelly (1977) and Van Lieshout (2000). Obviously, any hereditary density f defines a Markov point process if we let all pairs of points be neighbours ($\xi \sim \eta$ for all $\xi, \eta \in S$), but, as we note in Sections 2.3–2.4, we have much more restricted relations in mind.

2.1. Lower bound on the Papangelou conditional intensity

Assume that, for any finite $x \subset S$ and $\xi \in S \setminus x$,

$$\lambda(x, \xi) \geq \rho(\xi) \geq 0, \tag{1}$$

where ρ is a (deterministic) Borel function. Since the Poisson process with intensity function ρ is considered below, in order to avoid the trivial case where this process is almost surely empty, we also assume that the Lebesgue integral $\int_S \rho(\xi)\,\mathrm{d}\xi$ is positive. Thus (1) is a rather strong condition on λ. For example, pairwise interaction processes are in general excluded; cf. the Strauss process, the hard-core Gibbs point process, and many other examples in Baddeley and Møller (1989), Van Lieshout (2000), and Møller and Waagepetersen (2004). However, condition (1) is satisfied for a saturation process (see Geyer (1999)) and for the area-interaction point process studied in Section 2.4.

2.2. Coupling

Consider a spatial birth-and-death process $(X_t; t \in \mathbb{R})$ with birth rate λ and death rate equal to unity, whereby $(X_t; t \in \mathbb{R})$ is reversible with respect to the equilibrium distribution

specified by f (see Preston (1977) and Ripley (1977)). Let X_{t-} denote the state just before time t. The spatial birth-and-death process is a jump process, i.e. it is a continuous-time Markov process where a jump at time t is either a birth $X_t = X_{t-} \cup \{\xi\}$ of a new point ξ or a death $X_t = X_{t-} \setminus \{\eta\}$ of an old point $\eta \in X_{t-}$. If the process is in a state x (a finite subset of S) after a jump, the waiting time to the next jump is exponentially distributed with mean $1/A(x)$, where $A(x) = B(x) + n(x)$, $B(x) = \int_S \lambda(x, \xi) \, d\xi$, and $n(x)$ is the number of points in x. Conditional on t being a jump time and $X_{t-} = x$, a birth $X_t = x \cup \{\xi\}$ happens with probability $B(x)/A(x)$, in which case the newborn point ξ has density $\lambda(x, \xi)/B(x)$, otherwise a death $X_t = x \setminus \{\eta\}$ happens, where η is a uniformly picked point from x. For the present paper, as we describe in Section 2.3, we need to generate only a finite number of jumps of X_t for $t \leq 0$.

Assume that $\varnothing$ (the empty point configuration) is an ergodic state of $(X_t; t \in \mathbb{R})$. This condition implies that f specifies the unique equilibrium distribution; see Preston (1977), where examples satisfying this condition are discussed. For instance, ergodicity of $\varnothing$ is implied by the local stability condition,

$$\lambda(x, \xi) \leq V(\xi), \tag{2}$$

where V is an integrable function. In turn (2) is satisfied for most models used in practice; see Møller and Waagepetersen (2004, Section 6.1.1).

Suppose that $X = X_0$ follows f, and imagine that we have generated X_t forwards for $t \geq 0$ and backwards for $t \leq 0$ (by reversibility, the same generation can be used forwards and backwards). We can then couple $(X_t; t \in \mathbb{R})$ with a lower spatial birth-and-death process $(W_t; t \in \mathbb{R})$ obtained as follows. Each time t where $X_t = \varnothing$ is a regeneration time, and then we set $W_t = \varnothing$. These regeneration times split $((X_t, W_t); t \in \mathbb{R})$ into independent and identically distributed cycles, and so it suffices to consider the generation of $(W_t; t \in \mathbb{R})$ within each cycle. If a birth $X_t = X_{t-} \cup \{\xi\}$ happens then, with probability $\rho(\xi)/\lambda(x, \xi)$ (independently of what else has happened at previous jump times), we let $W_t = W_{t-} \cup \{\xi\}$, otherwise $W_t = W_{t-}$ is unchanged. If a death $X_t = X_{t-} \setminus \{\eta\}$ happens then we let $W_t = W_{t-} \setminus \{\eta\}$ (meaning that $W_t = W_{t-}$ is unchanged if $\eta \notin W_{t-}$). Thereby $(W_t; t \in \mathbb{R})$ is a spatial birth-and death process with birth rate ρ and death rate equal to unity, and W_0 is then a Poisson process on S with intensity function ρ (see Appendix G of Møller and Waagepetersen (2004)). Thus, Theorem 1 follows in the case of a Markov point process satisfying condition (1) (strictly speaking, the Markov property of the point process is not of importance; all that is required is that the density is hereditary and that (1) holds). Clearly, $W_t \subseteq X_t$ for all $t \in \mathbb{R}$.

2.3. Dependent thinning procedure

As above, suppose that (1) is satisfied and $\varnothing$ is an ergodic state of the spatial birth-and-death processes $(X_t; t \in \mathbb{R})$, where $X = X_0$ follows f. Assume also that X is a Markov point process (see Section 2.1). We can then generate the Poisson process W_0 with intensity function ρ within a random but finite number of steps, as described below.

Denote by $Z_0, Z_1, \ldots$ the jump chain of $(X_t; t \leq 0)$ considered backwards in time, meaning that $Z_0 = X_0$, Z_1 is the state just before the first jump time before time 0 occurs, and so on. Similarly, define $Y_0, Y_1, \ldots$, again in reverse chronological order, as the jump chain of $(W_t; t \leq 0)$. Let $G_0 = Z_0$, and, for $i = 1, 2, \ldots$, define recursively the ith generation ancestors of X by

$$G_i = \{\xi \in Z_i : N_\xi \cap G_{i-1} \neq \varnothing\}.$$

Since $\varnothing$ is an ergodic state, we can define a discrete nonnegative random variable I by $I = 0$ if $X = \varnothing$, and $G_0 \neq \varnothing, \ldots, G_{I-1} \neq \varnothing, G_I = \varnothing$ otherwise. We also define a discrete

nonnegative random variable J by $J = 0$ if $X = \varnothing$, and $Z_0 \neq \varnothing, \ldots, Z_{J-1} \neq \varnothing, Z_J = \varnothing$ otherwise.

Now, the dependent thinning procedure works as follows. If $X_0 = \varnothing$ then simply $W_0 = \varnothing$. In most practical applications we expect $X_0 = \varnothing$ to be a very unlikely event. Below we assume that $X_0 \neq \varnothing$.

First, for each $i = 1, \ldots, I$, simulate Z_i and determine G_i. Here we use the fact that, conditional on $Z_{i-1} = x$, a (backwards) birth $Z_i = x \cup \{\xi\}$ happens with probability $B(x)/A(x)$, in which case the newborn point ξ has density $\lambda(x, \xi)/B(x)$, otherwise a (backwards) death $Z_i = x \setminus \{\eta\}$ happens, where η is a uniformly picked point from x. This requires the evaluation of the function $B(x)$, which may necessitate numerical integration.

Second, considering the jump chain for the lower birth-and-death process forwards in time, set $Y_I = \varnothing$, and, for $i = I - 1, \ldots, 0$, generate Y_i in the same way as in the coupling construction in Section 2.2, but where of course we only need to consider the ancestors of X, since all other points will be irrelevant for the output W_0. Specifically, for $i = I - 1, \ldots, 1$, let U_i be a uniform random variable on $[0, 1]$ which is independent of what has so far been generated forwards from step I to step i; that is, (Z_I, Z_{I-1}, Y_I) if $i = I - 1$, and $(Z_I, \ldots, Z_i, Y_I, \ldots, Y_{i+1}, U_{I-1}, \ldots, U_{i+1})$ if $i < I - 1$. Then

- if $Z_i = Z_{i+1} \cup \{\xi_i\}$ is a (forwards) birth and $U_i \leq \rho(\xi_i)/\lambda(Z_{i+1}, \xi_i)$, then $Y_i = Y_{i+1} \cup \{\xi_i\}$, and $Y_i = Y_{i+1}$ otherwise;

- if $Z_i = Z_{i+1} \setminus \{\eta_i\}$ is a (forwards) death then $Y_i = Y_{i+1} \setminus \{\eta_i\}$.

Thereby we can return $W_0 = Y_0$ within the $2I$ steps given above. In many applications, the mean value of J might be extremely large (see Berthelsen and Møller (2002)). In the extreme case where all points are neighbours, $I = J$. However, in most application examples, $I \ll J$. See Fernández *et al.* (2002) and Berthelsen and Møller (2002).

2.4. Example: area-interaction point process

For S a bounded planar region, the area-interaction point process X has Papangelou conditional intensity

$$\lambda(x, \xi) = \beta \gamma^{-|b(\xi, r) \setminus \cup_{\eta \in x} b(\eta, r)|}, \tag{3}$$

where β, γ, and r are positive parameters, $|\cdot|$ denotes the area, and $b(\xi, r)$ is the disc with radius r centred at ξ (see Baddeley and Van Lieshout (1995)). For $\gamma > 1$, $\lambda(x, \xi)$ is increasing in x (the attractive case, originally studied by Widom and Rowlinson (1970)). For $\gamma < 1$, $\lambda(x, \xi)$ is decreasing in x (the repulsive case). For $\gamma = 1$, X is simply a homogeneous Poisson process on S with intensity β. Consequently, both (1) and (2) are satisfied, with

$$\rho(\xi) = \beta \gamma^{-\pi r^2}, \qquad V(\xi) = \beta, \qquad \text{if } \gamma \geq 1$$

and

$$\rho(\xi) = \beta, \qquad V(\xi) = \beta \gamma^{-\pi r^2}, \qquad \text{if } 0 < \gamma \leq 1.$$

An inhomogeneous version of the area-interaction point process, satisfying (1) and (2), is obtained by replacing all instances of β above with a nonnegative Borel function $\beta(\xi)$, assuming that $\int_S \beta(\xi) \, d\xi < \infty$, which ensures the existence of the process. Clearly, no matter which version is used, (3) implies that X is Markov with respect to the relation given by $\xi \sim \eta$ if and only if $b(\xi, r) \cap b(\eta, r) \neq \varnothing$, i.e. when the distance between ξ and η is less than or equal to $2r$.

To illustrate the use of the proposed method as a graphical technique for inspecting the fit of a spatial point process model, in Figures 1 and 2 we show examples of random thinnings of
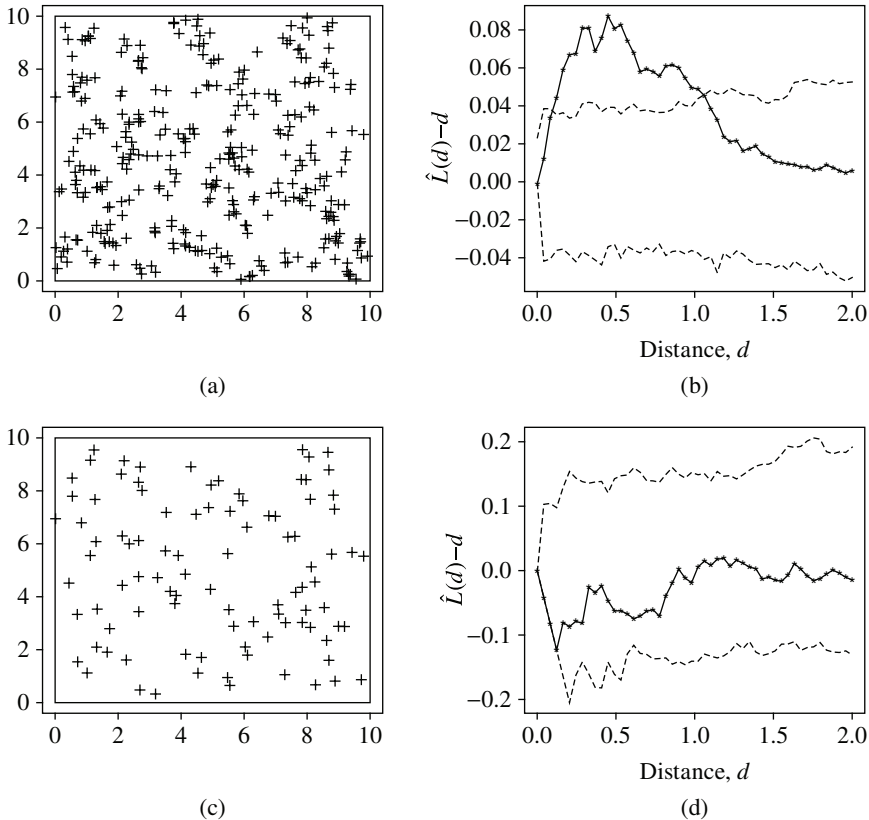
FIGURE 1: Thinning of a simulated attractive area-interaction process on $S = [0, 10] \times [0, 10]$, with $\gamma = 10^3$, $\beta = 7$, and $r = 0.3$. (a) The simulated area-interaction process, X. (b) The estimated centred L-function corresponding to X, and the empirical pointwise 95% upper and lower envelopes obtained by simulating 1000 homogeneous Poisson processes on $[0, 10] \times [0, 10]$, each with an expected number of points equal to the number observed in X. (c) The thinned process W corresponding to X. (d) The estimated centred L-function corresponding to W, and the empirical pointwise 95% upper and lower envelopes from 1000 simulated homogeneous Poisson processes, each with an expected number of points equal to the number observed in W.

simulated homogeneous area-interaction processes with $\gamma > 1$ and $\gamma < 1$, respectively. The aggregation in the area-interaction process for $\gamma > 1$ can be seen in Figure 1(a) and is confirmed in Figure 1(b), which shows the estimated centred L-function, $\hat{L}(d) - d$, corresponding to the realization in Figure 1(a), along with the pointwise 95% upper and lower envelopes based on simulations of 1000 homogeneous Poisson processes whose rates are equivalent to those observed by the process in Figure 1(a). Specifically, in terms of Ripley's K-function (see Ripley (1976), (1977)), $L(d) = \sqrt{K(d)/\pi}$, where $d > 0$ denotes the distance, and we use Ripley's nonparametric estimate $\hat{K}$ to obtain the estimate $\hat{L}(d) = \sqrt{\hat{K}(d)/\pi}$. Note that in the special case of a stationary Poisson process, the centred L-function, $L(d) - d$, is equal to 0. Figure 1(c) shows W, the random thinning of the process shown in Figure 1(a), using the method described in Section 2.3. Here $I = 4192$, while we stopped the backward run after 10^7 iterations, so $J > 10^7$. The number of points has decreased from 330 in Figure 1(a) to
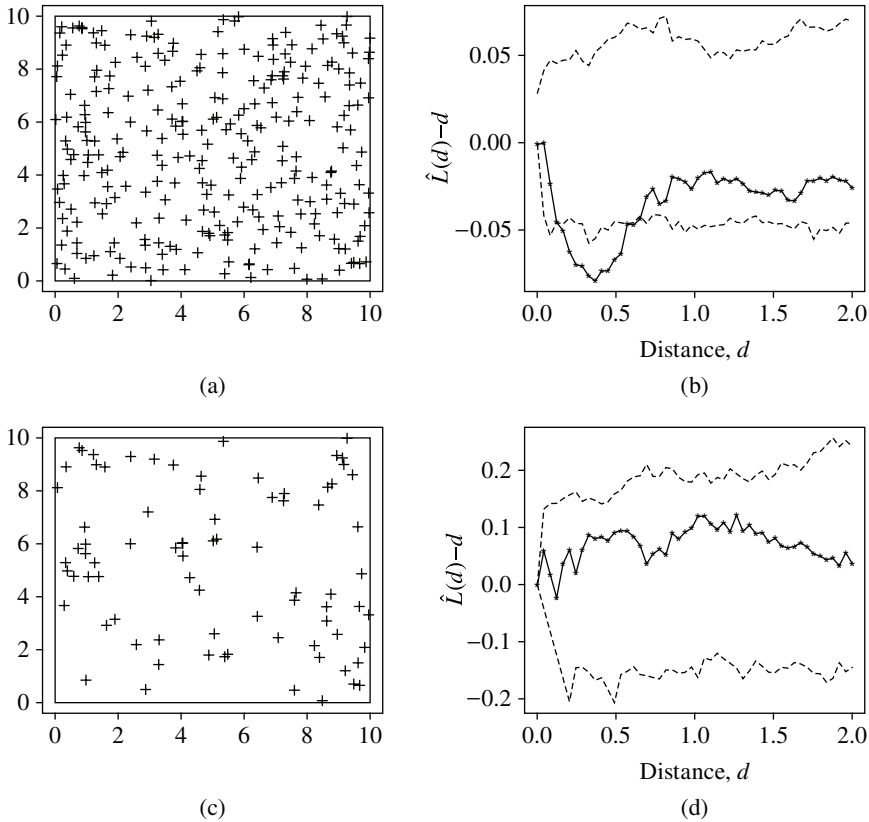
FIGURE 2: The same four plots as in Figure 1, but for the inhibitive case where $\gamma = 10^{-3}$, $\beta = 1$, and $r = 0.3$. (a) The simulated area-interaction process, X. (b) The estimated centred L-function for X, and the empirical pointwise 95% upper and lower envelopes obtained by simulating 1000 homogeneous Poisson processes on $[0, 10] \times [0, 10]$, each with an expected number of points equal to the number observed in the process X. (c) The thinned process W corresponding to X. (d) The estimated centred L-function corresponding to W, and the empirical pointwise 95% upper and lower envelopes from 1000 simulated homogeneous Poisson processes on $[0, 10] \times [0, 10]$, each with an expected number of points equal to the number observed in the process W.

96 in Figure 1(c). We see that several of the clusters of points in Figure 1(a) have been almost completely removed in Figure 1(c). Figure 1(d), which shows the estimated L-function for the thinned process in Figure 1(c), demonstrates that the clustering in the area-interaction process with $\gamma > 1$ is removed by the thinning procedure, resulting in a homogeneous Poisson process, and illustrates how this method may be used in practice in conjunction with the L-function as a visualization tool for examining the fit of the model used in the random thinning.

Similarly, Figure 2 illustrates the random thinning of an area-interaction process with $\gamma < 1$. The inhibition in Figure 2(a) may not be immediately obvious by eye, but Figure 2(b) and 2(d), which show the estimated centred L-function for the realization of the area-interaction process and its corresponding thinned process, demonstrate that the inhibition in the original process is statistically significant, compared to the homogeneous Poisson process, for distances of 0.3 to 0.6, and that this inhibition is removed by the random thinning procedure. The thinning

procedure allows the difficult task of examining the fit of a model for the clustering or inhibition in a spatial point process model to be simplified to that of merely inspecting the homogeneity of a thinned process. However, as these examples show, the thinned process may have substantially fewer points than the original process, which results in a loss of information and an impediment to the examination of homogeneity.

3. Thinning Cox processes

In the sequel, let $\lambda(\xi)$ be a random nonnegative function defined for all $\xi \in S$ such that, conditional on λ, X is a Poisson process with intensity function λ. In other words, X is a Cox process driven by λ. We assume that, almost surely, $\int_S \lambda(\xi) \, d\xi$ is finite, meaning that X is a finite point process.

3.1. Lower bound on the random intensity

Assume that, with probability 1,

$$\lambda(\xi) \geq \rho(\xi) \geq 0, \tag{4}$$

where, as in Section 2, ρ is a (deterministic) Borel function with $\int_S \rho(\xi) \, d\xi > 0$. Condition (4) implies that X can be viewed as the superposition $Q \cup R$ of a Cox process Q driven by $\gamma = \lambda - \rho$ and an independent Poisson process R with intensity function ρ. For example, γ may be log Gaussian and Q then a log Gaussian Cox process (see Møller *et al.* (1998)), or a shot noise process and Q then a shot noise Cox process (see Møller (2003)).

3.2. Independent thinning procedure

If λ were known then we could directly obtain a Poisson process W on S with intensity function ρ as an independent thinning of X with retention probabilities $\rho(\xi)/\lambda(\xi)$. Indeed, since conditional on λ, the independent thinning X is a Poisson process with intensity function ρ, and W is a Poisson process independent of λ and with intensity function ρ. This verifies Theorem 1 in the case of a Cox point process satisfying condition (4).

In practice we usually only observe a realization of $X = x$, and, hence, we first need to generate λ conditional on $X = x$. Hence, the independent thinning procedure works by

- firstly generating a realization of λ conditional on $X = x$;

- secondly generating W as an independent thinning of X with retention probabilities $\rho(\xi)/\lambda(\xi)$, $\xi \in x$.

How to simulate λ conditional on $X = x$ depends on the particular model. For example, if γ is log Gaussian, a Langevin–Hastings algorithm can be used (see Møller *et al.* (1998) and Møller and Waagepetersen (2004, Section 10.2.3)), and if γ is a shot-noise process, a birth–death Metropolis–Hastings algorithm applies (Møller (2003) and Møller and Waagepetersen (2004, Section 10.2.2)). We run one of these Metropolis–Hastings algorithms until it is effectively in equilibrium, and then return an (approximate) simulation of λ conditional on $X = x$.

3.3. Example: Thomas process

Let S be a bounded planar region, and let Q be a modified Thomas process (see Thomas (1949)), i.e. a Cox process driven by

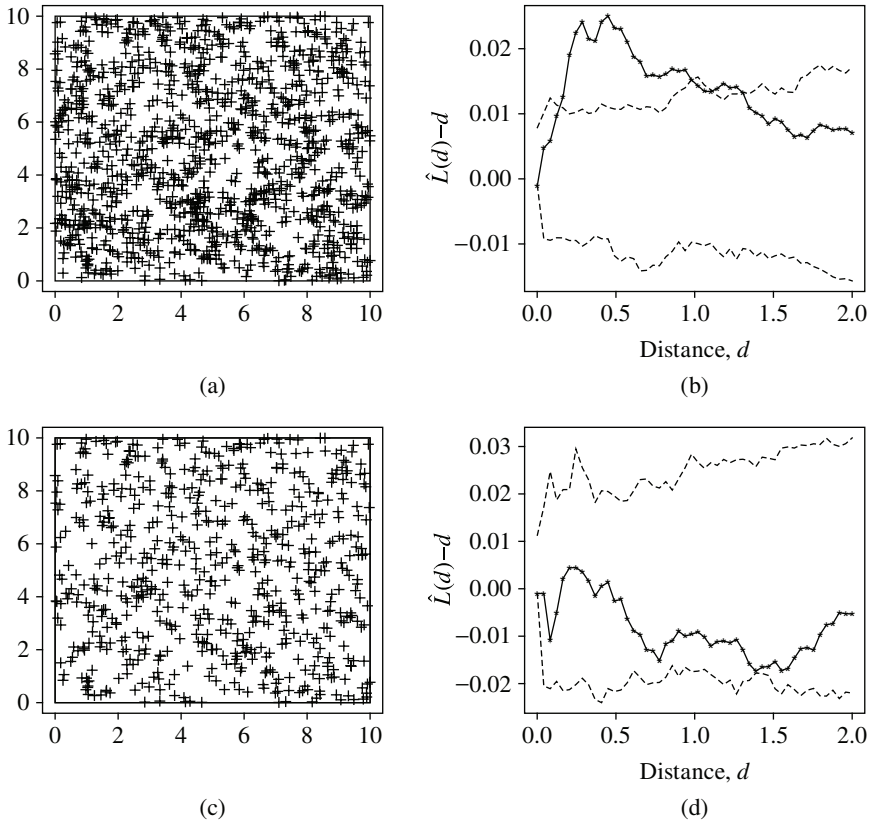$$\gamma(\xi) = \omega \sum_{\eta \in \Phi} \varphi(\xi - \eta), \qquad \xi \in S,$$

FIGURE 3: Thinning of a simulated superposition of a homogeneous Poisson process with rate $\rho = 6$ and a Thomas process on $S = [0, 10] \times [0, 10]$ with $\kappa = 2$, $\omega = 3$, and $\sigma = 0.2$. (a) The simulated superposition, X. (b) The estimated centred L-function for X, and the empirical pointwise 95% upper and lower envelopes obtained via 1000 simulated homogeneous Poisson processes, each with an expected number of points equal to the number observed in X. (c) The thinned process W corresponding to X. (d) The estimated centred L-function for W, and the empirical pointwise 95% upper and lower envelopes obtained via 1000 simulated homogeneous Poisson processes, each with an expected number of points equal to the number observed in W.

where φ is the bivariate normal density with mean 0 and covariance matrix $\sigma^2 I$, Φ is a homogeneous Poisson process with intensity κ defined on a bounded Borel set $S_{\text{ext}} \supseteq S$, and ω, σ, and κ are positive parameters. Here S_{ext} is chosen sufficiently large that edge effects can effectively be ignored; see Møller (2003) and Møller and Waagepetersen (2004, Section 5.5). In the following examples,

$$S = [0, 10] \times [0, 10] \quad \text{and} \quad S_{\text{ext}} = [-10, 20] \times [-10, 20].$$

As an illustration, Figure 3 shows the thinning of the superposition $X = Q \cup R$ when the Poisson process R has constant rate $\rho = 6$ and the Thomas process has parameters $(\kappa, \omega, \sigma) = (2, 3, 0.2)$. We see from Figure 3(b) that the process is highly clustered, which is a result of the fact that points tend to be clustered around the cluster centres Φ. Figure 3(c) shows the corresponding thinned process using the method described in Section 3.2, and the centred
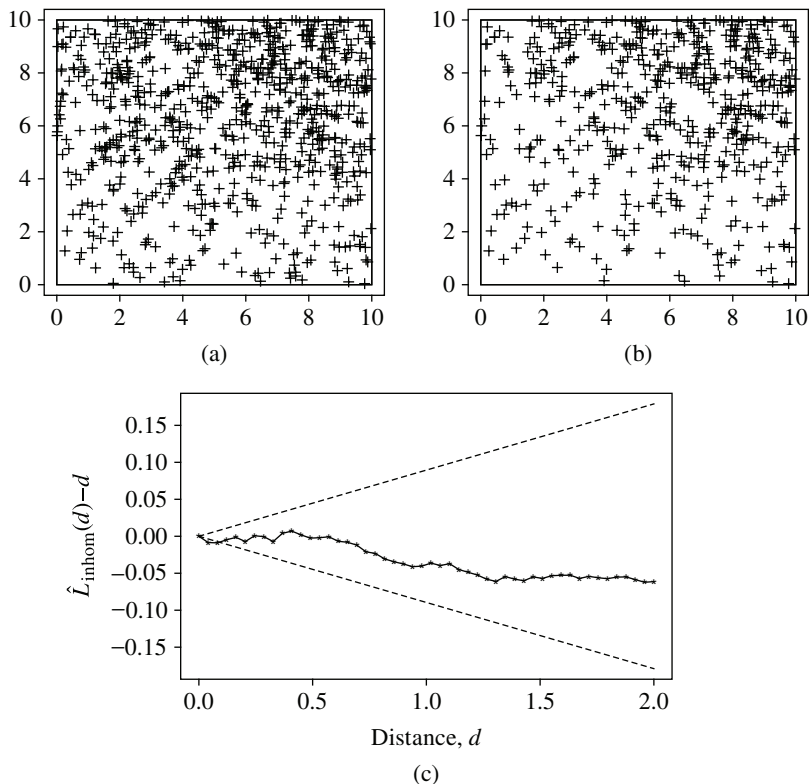
FIGURE 4: Thinning of a simulated superposition of an inhomogeneous Poisson process and Thomas process on $S = [0, 10] \times [0, 10]$, where the Poisson process has intensity function $\rho(x, y) = \exp(ax+by)$ with $(a, b) = (0.1, 0.2)$, and the Thomas process has parameters $(\kappa, \omega, \sigma) = (1, 5, 0.5)$. (a) The simulated superposition, X. (b) The thinned process W corresponding to X. (c) The centred estimate of the inhomogeneous $L(d)$ corresponding to W, using intensity ρ, and the pointwise 95% upper and lower envelopes based on the normal approximation in Veen and Schoenberg (2005), i.e. $\pm 1.96\sqrt{2\pi 10^2}d / \mathrm{E}(\#R)$, where $\mathrm{E}(\#R) = \int_S \rho(x, y)\, dx\, dy = [\exp(10a) - 1][\exp(10b) - 1]/(ab)$ is the expected number of points in the Poisson process R.

L-function estimate in Figure 3(d) verifies the homogeneity of the resulting thinned process and, thus, confirms that the clustering in the initial point process was accurately characterized.

   We may also consider an inhomogeneous version where $\rho(\xi)$ depends on ξ. For instance, Figure 4(a) shows a simulation of such a process with $\rho(x, y) = \exp(ax+by)$, where (x, y) are the Cartesian coordinates of ξ. Unlike the previous examples, in this case the thinned process is an inhomogeneous Poisson process with intensity function ρ, rather than a homogeneous Poisson process. Figure 4(b) shows the resulting thinned process, and Figure 4(c) displays a centred version of the estimated inhomogeneous L-function (see Baddeley *et al.* (2000) and Veen and Schoenberg (2005)), confirming that the process in Figure 4(c) is an inhomogeneous Poisson process. Figures 3 and 4 illustrate how the random thinning method may be used in practice, in conjunction with the L-function or other standard graphical tools for inspecting clustering and inhibition, as a graphical tool for inspecting the fit of a spatial point process model. An important area for future work is the comparison of the methods proposed here to

alternative graphical tools for point process residual analysis, such as the rescaled residuals used in Schoenberg (2003) and the quadrat residuals described in Baddeley *et al.* (2005).

Acknowledgements

References

BADDELEY, A. J. AND MØLLER, J. (1989). Nearest-neighbour Markov point processes and random sets. *Internat. Statist. Rev.* **2,** 89–121.

BADDELEY, A. AND TURNER, R. (2005). Spatstat: an R package for analyzing spatial point patterns. *J. Statist. Software* **12,** 1–42.

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

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

BADDELEY, A., TURNER, R., MØLLER, J. AND HAZELTON, M. (2005). Residual analysis for spatial point processes. With discussion and a reply by the authors. *J. R. Statist. Soc. B* **67,** 617–666.

BERTHELSEN, K. K. AND MØLLER, J. (2002). A primer on perfect simulation for spatial point processes. *Bull. Brazilian Math. Soc.* **33,** 351–367.

CARTER, D. S. AND PRENTER, P. M. (1972). Exponential spaces and counting processes. *Z. Wahrscheinlichkeitsth.* **21,** 1–19.

COX, D. R. (1955). Some statistical models connected with series of events. *J. R. Statist. Soc. B* **17,** 129–157.

CRESSIE, N. A. C. (1993). *Statistics for Spatial Data*. (Reprint of the 1991 edn.) John Wiley, New York.

FERRARI, P. A., FERNÁNDEZ, R. AND GARCIA, N. L. (2002). Perfect simulation for interacting point processes, loss networks and Ising models. *Stoch. Process. Appl.* **102,** 63–88.

GEYER, C. J. (1999). Likelihood inference for spatial point processes. In *Stochastic Geometry* (Toulouse, 1996; Monogr. Statist. Appl. Prob. **80**), eds O. E. Barndorff-Nielsen, W. S. Kendall and M. N. M. van Lieshout, Chapman and Hall/CRC, Boca Raton, FL, pp. 79–140.

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

KENDALL, W. S. (1998). Perfect simulation for the area-interaction point process. In *Probability Towards 2000* (New York, 1995; Lecture Notes Statist. **128**), eds L. Accardi and C. C. Heyde, Springer, New York, pp. 218–234.

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

MØLLER, J. (2003). Shot noise Cox processes. *Adv. Appl. Prob.* **35,** 614–640.

MØLLER, J. AND WAAGEPETERSEN, R. P. (2004). *Statistical Inference and Simulation for Spatial Point Processes*. Chapman and Hall/CRC, Boca Raton, FL.

MØLLER, J. AND WAAGEPETERSEN, R. P. (2007). Modern statistics for spatial point processes. With discussion. *Scand. J. Statist.* **34,** 643–711.

MØLLER, J., SYVERSVEEN, A. R. AND WAAGEPETERSEN, R. P. (1998). Log Gaussian Cox processes. *Scand. J. Statist.* **25,** 451–482.

PRESTON, C. J. (1977). Spatial birth-and-death processes. *Bull. Internat. Statist. Inst.* **46,** 371–391.

RIPLEY, B. D. (1976). The second-order analysis of stationary point processes. *J. Appl. Prob.* **13,** 255–266.

RIPLEY, B. D. (1977). Modelling spatial patterns. With discussion. *J. R. Statist. Soc. B* **39,** 172–212.

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

SCHOENBERG, F. P. (2003). Multidimensional residual analysis of point process models for earthquake occurrences. *J. Amer. Statist. Assoc.* **98,** 789–795.

SCHOENBERG, F. P. (2005). Comment on 'Residual analysis for spatial point processes' by Baddeley, Turner, Møller and Hazelton. *J. R. Statist. Soc. B* **67,** 661.

SCHOENBERG, F. P. AND ZHUANG, J. (2008). Residual analysis via thinning for spatial point processes. Submitted to *Ann. Instit. Statist. Math.*

STOYAN, D., KENDALL, W. S. AND MECKE, J. (1995). *Stochastic Geometry and Its Applications*, 2nd edn. John Wiley, Chichester.

THOMAS, M. (1949). A generalization of Poisson's binomial limit for use in ecology. *Biometrika* **36,** 18–25.

VAN LIESHOUT, M. N. M. (2000). *Markov Point Processes and Their Applications*. Imperial College Press, London.

VEEN, A. AND SCHOENBERG, F. P. (2005). Assessing spatial point process models for California earthquakes using weighted K-functions: analysis of California earthquakes. In *Case Studies in Spatial Point Process Models*, eds A. Baddeley *et al.* Springer, New York, pp. 293–306.

WIDOM, B. AND ROWLINSON, J. S. (1970). A new model for the study of liquid-vapor phase transitions. *J. Chemical Physics* **52,** 1670–1684.