On thinning a spatial point process into a Poisson process using the Papangelou intensity

Frederic Paik Schoenberg
Department of Statistics, University of California, Los Angeles, CA 90095-1554, USA
and
Jiancang Zhuang
Institute of Statistical Mathematics, Tokyo, Japan.

November 30, 2010

Proposed running head: Thinning spatial point processes.

Abstract

Random thinning has been shown to produce useful diagnostics for assessing the goodness-of-fit of a temporal or space-time point process model. The technique involves keeping or deleting each point individually, with each point kept with a probability inversely proportional to the conditional intensity at that point. This method does not extend immediately to the case of a purely spatial point process defined by its Papangelou intensity, however. Here, a method for thinning a spatial point process into a Poisson process is introduced. The proposed technique involves considering each possible subset of points, and keeping or deleting the subset with the appropriate probability. A demonstration on a simulated clustered spatial point process is considered, and practical implications and shortcomings are discussed.

Keywords: K-function, Papangelou intensity, point process, residual analysis, thinning.

1 Introduction

A space-time point process N on $\mathbb{R} \times \mathbb{R}^d$ is typically characterized by its conditional intensity process λ, where the conditioning is based on all observations at all previous times. Given such a point process N, one way to evaluate a model for λ is by random thinning, keeping each point independently with a probability proportional to the reciprocal of the estimated conditional intensity at that point [Schoenberg, 2003]. In this procedure, each point (t_i, x_i) of N is retained with probability $b/\lambda(t_i, x_i \,|\, \mathcal{H}_{t_i})$, where by assumption b is a positive real number such that with probability one, $\lambda \geq b$ almost everywhere. The resulting process is a homogeneous Poisson process of rate b [Schoenberg, 2003]. The residual points, obtained after thinning using an estimate of λ in place of λ, may readily be inspected for homogeneity using standard methods.

A purely spatial point process is often characterized instead by its Papangelou intensity λ_p, where the conditioning is based on the observations at all other locations. It was conjectured in [Schoenberg, 2005] that a thinning method similar to that used for space-time point processes should be valid for purely spatial point processes as well; that is, if one were to keep each point x_i of N with probability $b/\lambda_p(x_i)$, where b is a positive number that is almost surely a lower bound for λ_p almost everywhere, then the resulting process should, under general conditions, be Poisson with rate b. Baddeley et al. (2005, p 664) noted that this conjecture is false, since for instance if the original process is a Gibbs process with inhibition then the thinning described above will generate a point process still with clustering. A somewhat different thinning method for spatial Markov and Cox point processes was proposed by Møller and Schoenberg, based on clans of ancestors and dependent thinning relative to a birth and death process.

Here, we introduce an alternative method for thinning a finite spatial point process based on its Papangelou intensity, in order to form a homogeneous Poisson process. The main idea is to consider each

possible *subset* of points, and to keep only the points in a subset chosen according to the appropriate probability. While the probability that each subset is to be kept is in general intractable for a typical point process model, we show how it may readily be estimated using only the Papangelou intensity. Following some introductory background and notation in Section 2, we provide the details of our proposed thinning procedure in Section 3. Section 4 includes a discussion of practical limitations and computational issues associated with the suggested thinning procedure. In Section 5, an illustration using a clustered point process model is described, and a discussion including suggestions for future work is presented in Section 6.

2 Spatial point processes and Papangelou intensity

Excellent treatments of spatial point processes have been given by Cressie (1993), Stoyan et al. (1995), Møller and Waagepetersen (2004), Baddeley et al. (2005), and Baddeley and Turner (2006), and for a treatment of point processes in general, see Daley and Vere-Jones (2003). We review some preliminary definitions here. Given a Euclidean space $S \subseteq \mathbb{R}^d$ equipped with Borel σ-algebra $\mathcal{B}(S)$ and Borel measure μ, let $\mathbb{N}$ be the collection of all σ- finite integer-valued measures on S, equipped with the induced Borel σ-algebra $\mathcal{B}(\mathbb{N})$, which is generated by the sets of the form $\{X \in \mathbb{N} : X(B) \neq 0\}$ where B is any Borel set on S. An element N of $\mathbb{N}$ can be written as a countable sum of δ-measures such that

$$N(A) = \sum_i \delta_{x_i}(A),$$

where $\delta_y(A) = 1$ if $y \in A$, and is 0 if $y \notin A$. A point process is a measurable mapping $N(\omega)$ from the probability space $(\Omega, \mathcal{F}, P)$ into $(\mathbb{N}, \mathcal{B}(\mathbb{N}))$. N is a *simple* point process if $P\{\omega : \exists x \in S \text{ such that } N(\omega)\{x\} > 1\} = 0$. For brevity, in what follows we will write $N(\omega)$ as N, and use P to denote the image measure on $B(\mathbb{N})$ introduced by N. A point process is *finite* if N contains only finitely many points in S almost surely.

A point process N may be identified by its Janossy density f_N, which is the probability density function associated with the unordered points in a particular realization of points $\{x_1, x_2, ..., x_n\}$, relative to Lebesgue measure on the space $\mathcal{S}$. For a finite point process on a bounded Borel set $B \subseteq R^d$, the Janossy density is equivalent to the likelihood of the process; see e.g. p.213 of Daley and Vere-Jones (2003). In what remains we consider only finite point processes whose Janossy densities exist.

It is important to note that the Janossy density is the probability density associated with observing *exclusively* the given collection of points, in contrast to the *marginal* density g_N, which we define as follows. Let

$$\begin{aligned} g_N(x_1, x_2, \cdots, x_m) &= f_N(x_1, x_2, \cdots, x_m) + \sum_{n=1}^{\infty} \frac{m!}{(m+n)!} \int_S \cdots \int_S f_N(x_1, \cdots, x_m, y_1, \cdots, y_n) dy_1 \cdots dy_n, \\ &= \int_{\mathbb{N}} \mathbf{1}_{\{N \supseteq X\}} f_N(N) dP \end{aligned} \tag{1}$$

where f_N is the Janossy density. The marginal density $g_N(x_1, x_2, \cdots, x_m)$ may be thought of as the probability density associated with observing *at least* the collection of points $(x_1, ..., x_m)$.

The exvisible σ-algebra $\Phi^{\mathcal{F}}$ of $\mathcal{B}(S) \bigotimes \mathcal{F}$ is generated by the sets of the form $A \times U$ for $U \in \mathcal{J}_A$, where $A \in \mathcal{B}(S)$ and $\mathcal{J}_A$ is the σ-field generated by sets of the form $\{N(B) > k\}$ for $k \in \mathbf{R}$ and $B \subseteq A^c$ (Delay and Vere-Jones, 2006). That is, $\mathcal{J}_A$ contains all information on N outside A. An exvisible process is a measurable function on $\Phi^{\mathcal{F}}$. A typical exvisible process takes the following form: if $h(x, X)$ is a measurable function on $(S \times \mathbb{N}, \mathcal{B}(S) \bigotimes \mathcal{B}(\mathbb{N}))$ and $N(\omega)$ is a point process on S, then $h(x, N(\omega) \setminus \{x\})$ is an exvisible process.

The Campbell measure defined on $\mathcal{B}(S) \bigotimes \mathcal{F}$ is the unique measure satisfying

$$C_N(A \times F) = \int_F \int_A \mathrm{d}N(x)\, \mathrm{d}P = \int_F \int_A \delta_N(x)\, \mathrm{d}\mu\, \mathrm{d}P, \ A \in \mathcal{B}(S), \ F \in \mathcal{F}. \tag{2}$$

An $\mathcal{F}$-Papangelou intensity $\lambda_p(x, \omega)$, or simply $\lambda_p(x)$, is defined as the Radon-Nikodym derivative of

the restriction of C_N with respect to $\mu \times P$ on $\Phi^{\mathcal{F}}$. An $\mathcal{F}$-Papangelou intensity exists only when C_N is absolutely continuous with respect to $\mu \times P$ on $\Phi^{\mathcal{F}}$.

Providing λ_p exists, from the properties of Radon-Nikodym derivatives follows the Georgii-Nguyen-Zessin formula

$$\mathbf{E}\left[\int_S h(x)\,\mathrm{d}N\right] = \mathbf{E}\left[\int_S h(x)\,\delta_N(x)\,\mathrm{d}\mu\right] = \mathbf{E}\left[\int_S h(x)\,\lambda_p(x)\,\mathrm{d}\mu\right], \tag{3}$$

for any non-negative exvisible process $h(x)$. This formula is often written as

$$\mathbf{E}\left[\int_S h(x, N \setminus \{x\})\,\mathrm{d}N\right] = \mathbf{E}\left[\int_S h(x, N)\,\lambda_p(x)\,\mathrm{d}\mu\right], \tag{4}$$

for any non-negative $h(u,\, X)$ that is a measurable function on $(S \times (N),\, \mathcal{B}(S) \bigotimes \mathcal{B}(\mathbb{N}))$, and conversely, any exvisible process λ_p satisfying (3) is a Papangelou intensity of N (Georgii, 1976; Nguyen and Zessin, 1979; Baddeley et al., 2005; Baddeley and Turner, 2006). A simple point process is a homogeneous Poisson process if its Papangelou intensity is constant.

The Papangelou intensity may also be characterized in terms of the likelihood of a particular realization $N = \{x_1, ..., x_n\}$ of the point process:

$$\lambda_p(x) = \frac{L(N \cup x)}{L(N)}, \tag{5}$$

for $x \notin N$, and for $x \in N$,

$$\lambda_p(x) = \frac{L(N)}{L(N \setminus x)}. \tag{6}$$

In equations (5-6), $\lambda_p(x)$ refers to the Papangelou intensity at x given the configuration of other points in N, and is thus sometimes written $\lambda_p(x|N)$. It is important to bear in mind that λ_p is random: if the points of N are changed, then $\lambda_p(x)$ may change as well.

For a Poisson process with integrable non-negative rate function $\lambda(x)$, the Papangelou intensity is deterministic and equal to $\lambda(x)$ almost everywhere. For such a process, the likelihood of any realization

$\{x_1, ..., x_n\}$ is simply

$$\exp\{-\int_S \lambda(x)dx\} \prod_{i=1}^{n} \lambda(x_i).$$

3 Thinning spatial point processes

Consider the following thinning procedure. Given a realization $\{x_1, x_2, ..., x_n\}$ of points of a finite spatial point process N, where each x_i is a location in the bounded Borel subspace $\mathcal{S} \subseteq R^d$, let $\{M_1, M_2, ..., M_{2^n-1}\}$ denote the $2^n - 1$ distinct non-empty subsets of these n points, including the realization N itself. To emphasize that M_k depends on N, one might alternatively use the notation $M_{\{k;N(\omega)\}}$ in place of M_k, but for simplicity we will generally suppress this notation in what follows.

Let $p_0 = 0$, and for $k = 1, 2, ..., 2^n - 1$, let $p_k = exp\{-b\mu(\mathcal{S})\}b^{|M_k|}/g_N(M_k)$, where $|M_k|$ denotes the cardinality of M_k, b is a scalar, and g_N is the marginal density of N. Intuitively, $g_N(M_k)$ may be thought of as the likelihood that N contains (but is not necessarily limited to) all the points in M_k. Suppose that N is thinned so that with probability p_k, the subset M_k of points is retained, and all other points are discarded. Thus, if none of the subsets $M_1, ..., M_{2^n-1}$ is kept, then M contains no points at all. The following result shows that this thinning results in a homogeneous Poisson process of rate b.

Theorem 3.1. Suppose that N is a finite spatial point process on the bounded Borel subspace $\mathcal{S} \subseteq R^d$, with Janossy density f_N and Papangelou intensity $\lambda_p(x)$. Let $A = \mu(\mathcal{S})$, and suppose that with probability one $g_N(M_k) > 0$ for all M_k, and that there exists a scalar $b > 0$ such that with probability one, $\sum_{k=0}^{2^n-1} \exp\{-bA\}b^{|M_k|}/g_N(M_k) < 1$. Let u be a uniform random variable on $[0, 1]$ that is independent of N. Then $M(\omega) := \sum_k M_{\{k;N(\omega)\}} \mathbf{1}_{\{\sum_{j=0}^{k-1} p_j \leq u < \sum_{j=0}^{k} p_j\}}$ is a homogeneous Poisson process on $\mathcal{S}$ with rate b.

Proof.

It must first be verified that M is a simple point process. M clearly inherits its simplicity and σ-finiteness from N, and by construction M must be non-negative integer-valued since the same is true for

N. M is measurable since for any non-negative integer k and any Borel subset $B \subseteq \mathcal{S}$, the set

$$\begin{aligned}
\{\omega : M(B) > k\} &= \{\omega : N(B) > k, u \in [p_{k-1}, p_k), M_k \cap B > k\} \\
&= \bigcup_{i=k+1}^{\infty} \bigcup_{j=k+1}^{i} \left[\{N(B) = i\} \cap \{u \in [\sum_{j=0}^{k-1} p_j, \sum_{j=0}^{k} p_j)\} \cap \{M_k \cap B = j\} \right] \\
&= \bigcup_{i=k+1}^{\infty} \bigcup_{\{M_k(N) : |M_k \cap B| > i\}} \left[\{N(B) = i\} \cap \{u \in [\sum_{j=0}^{k-1} p_j, \sum_{j=0}^{k} p_j)\} \right].
\end{aligned}$$

Since u and N are jointly measurable (as is tacitly implied in the assumption that u and N are independent), the set $\{M(B) > k\}$ is the countable union of measurable sets and is therefore measurable.

If $X \neq \emptyset$, the Janossy density function for M is given by the sum, over all possible configurations of N, of the probability that N contains X multiplied by the probability that the thinning X is selected from N, given N. This yields

$$\begin{aligned}
f_M(X) &= \int_{\mathbb{N}} \mathbf{1}_{\{N \supseteq X\}} f_N(N) \exp\{-b\mu(\mathcal{S})\} b^{|X|} / g_N(X) dP \\
&= \exp\{-b\mu(\mathcal{S})\} b^{|X|} \int_{\mathbb{N}} \mathbf{1}_{\{N \supseteq X\}} f_N(N) dP / g_N(X) \\
&= \exp\{-b\mu(\mathcal{S})\} b^{|X|},
\end{aligned}$$

which implies that M is a homogeneous Poisson process with rate b.

4 Practical considerations.

There are several practical limitations to the implementation of the thinning procedure described in the preceding Section. Foremost among these are the computation and approximation of thinning probabilities, the verification of the condition on b in Theorem 3.1, and the enumeration of subsets of N. Some of these problems may readily be remedied while others are extremely problematic. This Section describes some issues and ideas for overcoming some of these obstacles.

We consider here the typical case where a finite spatial point process N on a bounded region $B \subseteq \mathbf{R}^d$ is specified by its Papangelou intensity $\lambda_p(x)$, which is readily calculable for any location x and any realization of points, $x_1, ..., x_n$. We assume below that the probability $q = P\{N(B) = \emptyset\}$ is known as well.

In order to implement the thinning technique proposed in Theorem 3.1, the function $g(M_k)$ is required. Although $g(M_k)$ cannot typically be written expressly in terms of λ_p, the function can readily be estimated using only λ_p. Let $\{m_1, m_2, ..., m_c\}$ denote the points in the subset M_k. Suppose that a particular realization X of points includes M_k as well as some other points $Y = \{y_1, ..., y_d\}$, where $Y \cap M_k = \emptyset$. Repeatedly using equation (5), the likelihood for X is given by

$$L(X) = L(Y)\lambda_p(m_1|Y)\lambda_p(m_2|Y, m_1)...\lambda_p(m_c|Y, m_1, ..., m_{c-1}). \tag{7}$$

Recall that, for a finite process N such as that considered here, the Janossy density f is equivalent to the likelihood function L. Therefore, using (1) and (7), and provided f has no atoms in M_k (which is tacitly implied by assuming the existence of λ_p everywhere), $g(M_k)$ may be written as

$$\begin{aligned} g(M_k) &= f(M_k) + \sum_{n=1}^{\infty} \frac{1}{(c+n)!} \int_S \cdots \int_S L(M_k \cup \{y_1, \cdots, y_n\}) dy_1 \cdots dy_n \\ &= f(M_k) + \sum_{n=1}^{\infty} \frac{1}{(c+n)!} \int_S \cdots \int_S L(y_1, \cdots, y_n) \prod_{i=1}^{c} \lambda_p(m_i|y_1, \cdots, y_n, m_1, ..., m_{i-1}) dy_1 \cdots dy_n, \end{aligned} \tag{8}$$

where for $i = 1$, $\{m_1, ..., m_{i-1}\} = \emptyset$.

Consider the case where the point process N may readily be simulated repeatedly. Then a simple empirical estimator is suggested immediately by equation (8). That is, one may simulate the process N repeatedly, obtaining the point patterns $Y^{(1)}, Y^{(2)}, ...$ For each simulation $Y^{(j)}$, one obtains the value $\tilde{g}(M_k)^{(j)} := \prod_{i=1}^{c} \lambda_p(m_i|Y^{(j)}, m_1, ..., m_{i-1})$. One may then use the mean of these estimates $\hat{g}(M_k) := avg\{\tilde{g}(M_k)^{(j)}\}$ as an estimate of $g(M_k)$. As the number of simulations of N increases, this estimate will

converge to $g(M_k)$ almost surely, since the frequency of any realization in the simulations will converge almost surely to the likelihood of that realization.

A variety of methods exist for simulating various types of spatial point processes; see e.g. Møller and Waagepetersen (2004) for a review. In the general case where only the Papangelou intensity λ_p and the empty set probability q are available, and where there exists a constant a such that $\lambda_p(x) < a$ for all x with probability one, a very simple rejection sampling procedure may be used. Consider beginning by keeping the empty set as a realization of N with probability q. Then with probability $1 - q$, consider simulating a homogeneous Poisson process on B with rate a, obtaining the k points $x_1, x_2, ..., x_k$. [Simulating a Poisson process is a trivial task; see e.g. p.24 of Møller and Waagepetersen (2004).] One may then keep the simulation with probability $a^{-k} \prod_{i=1}^{k} \lambda_p(x_i|x_1, ..., x_{i-1})$ if $k > 0$, and with probability 0 if $k = 0$, repeating until a simulation is retained. Through repeated use of equation (5), one sees that this retention probability is proportional to the ratio of the Janossy density of $\{x_1, ..., x_k\}$ for the point process N to the Janossy density of this realization for the Poisson process of rate a. Since the Janossy density must integrate to unity for both N and the Poisson process, it follows that the resulting retained process will have a Janossy density identical to that of N.

Regarding the conditions of Theorem 3.1, the hypothesis that $g(M_K) > 0$ for any M_K is satisfied for many point processes. Recall that $g(M_k)$ represents the likelihood of observing at least the collection of points M_k. If, for instance, $q > 0$ and $\lambda_p(x) > 0$ for all x, then the likelihood for any realization is positive; hence so is $g(M_k)$. Regarding the determination of b satisfying the condition that $\sum_{k=0}^{2^n-1} \exp\{-bA\} b^{|M_k|}/g(M_k) < 1$ with probability one, in general no such b can be found. However, one may typically find a small positive number b satisfying this inequality with very high probability in order to obtain a thinning that closely approximates a homogeneous Poisson process, as in the example given in the next Section. It is important to note that the Papangelou intensity λ_p depends on the spatial

units of the region B considered. Hence, provided a positive lower bound for λ_p exists, one can always rescale the space so that this lower bound is unity. Suppose b is this lower bound, and that $b \approx 1$. Then $\sum \exp\{-bA\} b^{|M_k|}/g(M_k) = \exp\{-A\} \sum 1/g(M_k)$. Note as above that, by iterative use of equation (5), $f_N(X) = L(X) = q\lambda_p(x_1|\emptyset)\lambda_p(x_2|x_1)...\lambda_p(x_n|x_1, ..., x_{n-1})$, where $X = \{x_1, ..., x_n\}$. Thus

$$g(M_k) = \int_{\mathbb{N}} \mathbf{1}_{\{X \supseteq M_k\}} f_N(X) dP = q \int_{\mathbb{N}} \mathbf{1}_{\{X \supseteq M_k\}} \prod \lambda_p(x_i|x_1, ..., x_{i-1})\, dP > q \int_{\mathbb{N}} \mathbf{1}_{\{X \supseteq M_k\}}\, dP,$$

so that the condition in Theorem 3.1 holds provided that $\exp\{-A\} \sum 1/[q \int_{\mathbb{N}} \mathbf{1}_{\{X \supseteq M_k\}}\, dP] < 1$.

A major practical limitation to the implementation of the thinning procedure described in Theorem 3.1 is the computational burden required. The simulation procedure described above may be inefficient, and faster simulation methods, such as the Gibbs sampler strategies outlined in Møller and Waagepetersen (2004) may be of assistance. More importantly, unless the number of points in the realization of N is very small, merely enumerating all $2^n - 1$ non-empty subsets of N may be computationally infeasible. Thus the thinning method proposed here may presently only be used for very small datasets, such as the example given in the next Section. One way to address this problem is to divide the observation region B into small (non-random) subregions, and to thin the points in each subregion individually, as in the example in the next Section. Provided the subregions are sufficiently small so that N does not have many points in any single subregion of B, thinning each subregion in this manner may be computationally feasible. In addition, future improvements in computational efficiency in general, as well as the development of clever computational algorithms, may render the method proposed in Theorem 3.1 applicable to larger datasets. In particular, in some cases one may perhaps be able to begin the thinning procedure by first considering those subsets which are most likely to be kept and stopping once one such subset is retained, thus eliminating the need to consider all possible subsets every time this thinning is implemented. We leave the development of such a scheme as a subject for future work.

5 Example.

As an illustration of the thinning described in Theorem 3.1, consider a spatial point process N on the square $B = [0, z] \times [0, z]$ such that $\lambda_p(x) = \lambda_1$ if there is another point within distance δ of x, and $\lambda_p(x) = \lambda_2$ otherwise. If $\lambda_1 > \lambda_2$ then the process will be clustered; $\lambda_1 < \lambda_2$ results in inhibition. Let $q = \exp(-\lambda_2 z^2)$. Here, we choose $(\lambda_1, \lambda_2, \delta, z) = (0.5, 0.2, 0.1, 50)$, so that the process is clustered, as seen in Figure 1a, which was simulated using the simple rejection method described in Section 3.

In thinning the realization depicted in Figure 1a, the square B is subdivided into 2500 square cells of unit area. Within each cell, each subset M_k of the points in N is considered, and the function $g(M_k)$ is estimated as described in the previous Section, using 2000 simulations of N. That is, for each simulation, a value of $\prod_{i=1}^{c} \lambda_p(m_i|Y^{(j)}, m_1, ..., m_{i-1})$ is obtained, and the mean of these values is used as an estimate of $g(M_k)$.

Letting $b = 0.05$, the sum $\sum_{k=0}^{2^n-1} \exp\{-bA\} b^{|M_k|}/g(M_k)$ will not be less than unity with probability one, but among all 2500 cells in the realization in Figure 1a, the sum $\sum_{k=0}^{2^n-1} \exp\{-bA\} b^{|M_k|}/g(M_k)$ never exceeded 0.90, so the thinned process in Figure 1b should be very well-approximated by a Poisson process of rate $b = 0.05$. The thinned process in Figure 1b has 114 points, while the expected number is $b \times 2500 = 125$.

A commonly used diagnostic used to investigate clustering or inhibition in spatial point patterns is the function $K(r)$, which represents the number, normalized relative to the homogeneous Poisson process, of additional points within radius r of any given point in the point pattern (see Ripley, 1981). Figure 2 shows the K-function for both the realization of N in Figure 1a and the realization of the thinned process M in Figure 1b, in order to verify that while the process N is clustered, this clustering is removed by the thinning procedure described in Theorem 3.1.

6 Conclusion.

Results involving thinning point processes to form Poisson processes should be extended in future research to the more practical case where b is random. A comparison of the resulting Cox processes one might obtain in the situation where b is random would be most welcome, especially for the case where both λ and b are random. In addition, while the conditional intensity of the process resulting from thinning a Poisson process on the line with unknown intensity using the maximum likelihood estimate of the intensity was investigated in Schoenberg (2002), the extension of such results to the spatial case and to more complex point processes is another topic for future work.

Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant No. 0306526. We thank Yosihiko Ogata for helpful remarks and encouragement.

References

Baddeley, A., and Turner, R. (2006). Modelling spatial point patterns in R. In *Case Studies in Spatial Point Pattern Modelling*, A. Baddeley, P. Gregori, J. Mateu, R. Stoica, and D. Stoyan, editors, Lecture Notes in Statistics, 185, pp. 23-74. Springer-Verlag, New York, 2006.

Baddeley, A., Turner, R., Møller, J. and Hazelton, M. (2005). Residual analysis for spatial point processes (with discussion). *J. R. Statist. Soc. B*, 67, 617–666.

Cressie, N. (1993). *Statistics for Spatial Data, revised edition.* Wiley, New York.

Daley, D. and Vere-Jones, D. (2003). *An Introduction to Theory of Point Processes* (2nd edition). New York: Springer.

Georgii, H.O. (1976). Canonical and grand canonical Gibbs states for continuum systems. *Communications in Mathematical Physics* 48(1), 31–51.

Møller, J., and Schoenberg, F.P. (2010). Thinning spatial point processes into Poisson processes. *Adv. Appl. Prob.* 42, 347–358.

Møller, J. and Waagepetersen, R.P. (2004). Statistical Inference and Simulation for Spatial Point Processes. London: Chapman and Hall.

Nguyen, X.X. and Zessin, H. (1979). Integral and differential characterization of the Gibbs process. *Math. Nachr.* 88:105–115.

Ripley, B. (1981). *Spatial Statistics.* Wiley, NY.

Schoenberg, F.P. (2002). On rescaled Poisson processes and the Brownian bridge. *Ann. Int. Stat. Math.* , 54(2), 445–457.

Schoenberg, F. P. (2003). Multidimensional residual analysis of point process models for earthquake occurrences. *Journal of the American Statistical Association*, 98, 789-795.

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

Stoyan, D., Kendall, W.S., and Mecke, J. (1995). *Stochastic Geometry and its Applications, 2nd ed.* J. Wiley & Sons, Chichester.

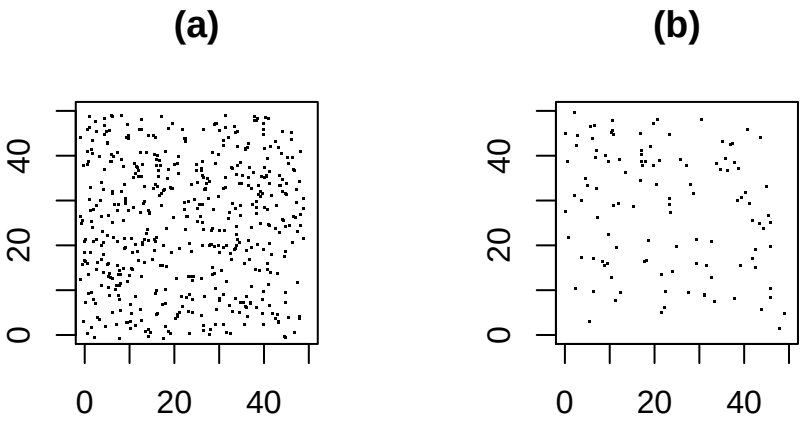

Figure 1: (a) Realization of the clustered process N described in Section 4. (b) Thinned process M based on thinning N using the method in Theorem 3.1.

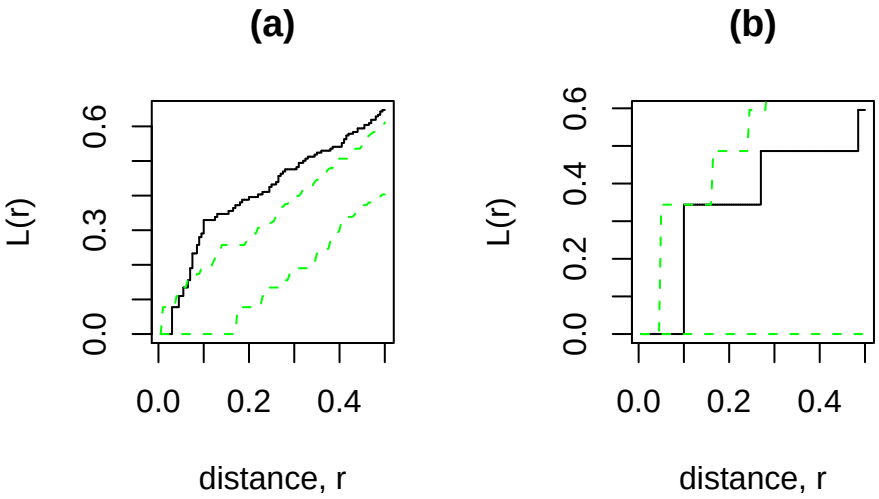

Figure 2: (a) K-function for the realization of the process N from Figure 1a. (b) K-function for the thinned process M of Figure 1b. In both cases, 95% confidence bounds based on simulated homogeneous Poisson process are shown.