A Two-Step Markov Point Process

MAJEED M. HAYAT and JOHN A. GUBNER
University of Wisconsin–Madison

ABSTRACT. Existence and uniqueness are established for a translation-invariant Gibbs measure corresponding to a spatial point process that has, in addition to inhibition and clustering, the new feature of penalizing isolated points. It is shown that this point process has the so-called two-step Markov property, and a theorem is proved that characterizes the more general m-step Markov density functions. The asymptotic normality of certain statistics of the point process is proved when the size of the observation window tends to $\mathbb{R}^2$.

Key words: isolated points, m-step Markov, Gibbs measure, asymptotic normality, mixing condition, many-body interaction.

Running head: A Two-Step Markov Point Process

Postal address: ECE Dept.
1415 Johnson Drive
Madison, WI 53706–1691
U.S.A.

E-mail address: hayat@large.ece.wisc.edu
gubner@entropy.ece.wisc.edu

This work has been submitted in part to the 1995 IEEE International Symposium on Information Theory, Whistler, British Columbia, Canada, 17–22 September 1995. This research was supported by the Office of Naval Research under Grant N00014–94–1–0366.

1. Introduction

The Markov point processes introduced by Ripley & Kelly (1977) provide a rich source for point processes featuring interaction between points. Pairwise interaction models (Strauss, 1975; Kelly & Ripley, 1976; and Strauss, 1986) are particularly useful for imposing inhibition and clustering. The range of interaction in Markov point processes is generally governed by a fixed relation. In particular, a point interacts only with surrounding points that are within a fixed neighborhood of it. Later, the more general nearest-neighbor Markov point processes were introduced by Baddeley & Møller (1989) in which the relation governing the range of interaction was allowed to be realization dependent.

In this paper, we consider a many-body interaction model for point processes. This model permits, in addition to inhibition and clustering, a new feature of penalizing isolated points. When this feature is embedded in a density function (with respect to a Poisson measure) defined on the set of all finite point configurations in bounded subsets of the plane, the resulting density function will possess a two-step Markov property. Namely, not only does the density of the one-point augmentation of a point configuration x depend on the new point and those points of x that are neighbors of the new point, but also on the neighbors' neighbors.

In Section 2, we define a class of point processes on a bounded region with density functions that have a more general m-step Markov property. A characterization theorem for such densities is proved in terms of the so-called m-step interaction functions. An m-step interaction function assumes the value one whenever its argument is not an m-clique, i.e., every two points are related in some extended sense. When $m = 1$, our definitions of an m-step Markov density function, m-step interaction function, and m-clique coincide with the definitions of the Markov density function, interaction function, and clique associated with Markov point processes. This characterization theorem is therefore a generalization of the characterization theorem of Ripley & Kelly (1977) for Markov point processes to m-step Markov point processes.

In Section 3, we define a point process on the whole plane having the two-step Markov property. To do so we use the two-step Markov density and exploit our characterization theorem to define a potential function in terms of many-body interaction functions. Using this potential function, a collection of conditional measures is defined on point configurations in $\mathbb{R}^2$. We prove a key result (Lemma 3.2) that gives conditions on the parameters of the model to guarantee a certain mixing condition. This result is used to show existence and uniqueness of a translation-invariant Gibbs measure on all of $\mathbb{R}^2$ corresponding to the conditional measures.

In Section 4, we provide conditions on the parameters of interaction to guarantee the asymptotic normality, as the observation region tends to $\mathbb{R}^2$, of some important statistical quantities of the point process. Examples of such statistics include the total number of points, the number of isolated points, and the number of pairs of points that are within a certain range of each other.

2. An m-step Markov point process on bounded sets

Let $\mathbf{X}$ be a bounded Borel subset of $\mathbb{R}^2$, and let $\mathcal{B}_{\mathbf{X}}$ be the Borel σ-algebra on $\mathbf{X}$. Let $\mathbf{X}_0 := \emptyset$, and for $n \geq 1$, let $\mathbf{X}_n$ be the set of collections of n points (not necessarily distinct) in $\mathbf{X}$. Let $\mathbf{X}_e := \bigcup_n \mathbf{X}_n$, and let $\mathcal{B}_e^{\mathbf{X}}$ be the smallest σ-algebra on $\mathbf{X}_e$ making the number of points of x that are in B a random variable for each $x \in \mathbf{X}_e$ and $B \in \mathcal{B}_{\mathbf{X}}$. The measurable space $(\mathbf{X}_e, \mathcal{B}_e^{\mathbf{X}})$ is called an exponential space (Carter & Prenter, 1972). A point process in $\mathbf{X}$ is a measurable map from a probability space into $(\mathbf{X}_e, \mathcal{B}_e^{\mathbf{X}})$. This map induces a distribution on $\mathcal{B}_e^{\mathbf{X}}$. Under the assumption that this distribution is absolutely continuous with respect to the measure $\nu_{\mathbf{X}}$ induced on $\mathcal{B}_e^{\mathbf{X}}$ by a Poisson process with constant intensity λ, this distribution can be specified by its density function f, which is the Radon–Nikodym derivative of the distribution with respect to $\nu_{\mathbf{X}}$. For each n, the density function $f(x_1, \ldots, x_n)$ is measurable on $\mathbf{X}_e$, and can be defined to be independent of the order of the arguments for each n so that if

$$\begin{aligned} p_n &:= \int_{\mathbf{X}_n} f(x)\, d\nu_{\mathbf{X}}(x) \\ &= \frac{e^{-\lambda \ell(\mathbf{X})}}{n!} \int_{\mathbf{X}^n} f(x_1, \ldots, x_n)\, dx_1 \cdots dx_n, \end{aligned}$$

then $\sum_n p_n = 1$, where ℓ is two-dimensional Lebesgue measure, and p_n is the probability of having n points. Furthermore, the density of the locations of the points is proportional to f. If $f \equiv 1$, we recover the Poisson process with constant intensity λ.

With the appropriate choice of f, interaction requirements can be imposed on the resulting point process. Pairwise interaction schemes can be used to model processes featuring inhibition and clustering (Strauss, 1986; Kelly & Ripley, 1976). However, a point process may be required to have an additional property that it may be unlikely to have realizations with many "isolated" points, and it may therefore be desirable to penalize such realizations. Let $0 < d < D$ be given. Then for $x \in \mathbf{X}_e$, the number of isolated points in $x = (x_1, \ldots, x_n)$ is given by

$$I(x) := \Big|\{i : \|x_i - x_j\| > D, \forall j \neq i\}\Big|,$$

where $\|\cdot\|$ is the Euclidean norm on $\mathbb{R}^2$, and $|A|$ denotes the cardinality of the set A. Define the density f on $\mathbf{X}_e$ by

$$f(x) = a\gamma^{I(x)} \prod_{i<j} \psi(\|x_i - x_j\|), \tag{2.1}$$

where $\psi : [0, \infty) \to [0, \infty), \psi(r) = 1$ whenever $r \geq D$, $\psi(r) = 0$ if $r < d$, and ψ is upper bounded by some $\overline{\psi} < \infty$. The function ψ is responsible for pairwise interaction and will give rise to clustering and inhibition (Baddeley & Møller, 1989). What is new here is the constant $0 < \gamma \leq 1$, which is responsible for penalizing realizations with isolated points; a is a normalizing constant making f a probability density. Note that for any $y \in \mathbf{X}_e$, $f(y) = 0$ implies $f(x) = 0$ for all x with $x \supset y$. It is easy to see that for $x \in \mathbf{X}_e$ such that $f(x) > 0$, and $\xi \in \mathbf{X}$,

$$\frac{f(x \cup \{\xi\})}{f(x)} = \gamma^{I(x\cup\{\xi\})-I(x)} \prod_{i=1}^{n} \psi(\|\xi - x_i\|). \tag{2.2}$$

Observe that for each point x_i in x for which $\|\xi - x_i\| \leq D$, if x_i is an isolated point prior to the augmentation by ξ, then the augmentation results in a reduction in the number of isolated points by 1. On the other hand, if x_i is not an isolated point, then the augmentation does not contribute any change, as far as x_i is concerned, in the number of isolated points. The other possibility is that ξ is itself isolated from x, in which case, $I(x \cup \{\xi\}) - I(x) = 1$. This can be expressed symbolically as follows:

$$I(x \cup \{\xi\}) - I(x) = \begin{cases} 1, & |\{B(\{\xi\}) \cap x\}| = 0, \\ -\sum_{i=1}^{k} \mathbf{1}\{|\{B(\{\tilde{x}_i\}) \cap x\}| = 1\}, & |\{B(\{\xi\}) \cap x\}| = k > 0, \end{cases}$$

where for $A \in \mathcal{B}_{\mathbf{X}}$, $B(A) = \cup_{a\in A}\{\eta \in \mathbf{X} : \|\eta - a\| \leq D\}$, and here $(\tilde{x}_1, \ldots, \tilde{x}_k) = B(\{\xi\}) \cap x$. Hence, $I(x \cup \{\xi\}) - I(x)$ depends only on ξ and on the points $B(B(\{\xi\}) \cap x) \cap x$. The latter consists of the neighbors of ξ and their neighbors. The product term in (2.2) depends only on ξ and $B(\{\xi\}) \cap x$. Therefore, for some function g,

$$f(x \cup \{\xi\}) = f(x) g\Big(\xi, B(B(\{\xi\}) \cap x) \cap x\Big), \qquad \forall x \in \mathbf{X}_e.$$

We will prove a characterization theorem for a more general class of densities possessing what we call an m-step Markov property. Namely, we consider the case for which the dependence carries over to m steps and not just two. More precisely, fix an integer $m \geq 1$, and let $\sim$ be a reflexive and symmetric relation defined on $\mathbf{X}$. (The only relation considered in this paper is the fixed distance relation, i.e., $\xi \sim \eta$ if and only if $\|\xi - \eta\| \leq D$.) For each $x \in \mathbf{X}_e$ and $\xi \in \mathbf{X}$, define the neighborhood $E^m(\xi|x)$ of ξ with respect to x consisting of those points η in x satisfying the following condition:

$$\eta \sim \xi, \text{ or } \exists \zeta_1, \ldots, \zeta_j \in x, j < m, \text{ such that } \eta \sim \zeta_1, \zeta_1 \sim \zeta_2, \ldots, \zeta_{j-1} \sim \zeta_j, \zeta_j \sim \xi.$$

Note that $E^m(\xi|x) \subseteq E^{m+1}(\xi|x)$.

Definition. Given an integer $m \geq 1$, a function $f : \mathbf{X}_e \to [0, \infty)$ is called an *m-Markov function* if there exists a function g on $\mathbf{X} \times \mathbf{X}_e$ such that $f(x \cup \{\xi\}) = f(x)\, g(\xi, E^m(\xi|x))$ for every $x \in \mathbf{X}_e$, $\xi \in \mathbf{X}$. A point process whose density function is m-Markov is called an *m-Markov point process.*

Thus, the function f in (2.1) is a 2-Markov function. Note that the special case $f_0(x) = \gamma^{I(x)}$ is also a 2-Markov function.

Definition. A point configuration $x \in \mathbf{X}_e$ is said to be an *m-clique* if for every $\xi, \eta \in x, \xi \sim \eta$, or there exist $\zeta_1, \ldots, \zeta_j \in x, j < m$, such that $\xi \sim \zeta_1, \zeta_1 \sim \zeta_2, \ldots, \zeta_{j-1} \sim \zeta_j, \zeta_j \sim \eta$. Note that if x is an m-clique, then it is an $(m+1)$-clique.

Remark. $x \in \mathbf{X}_e$ is an m-clique if and only if $E^m(\xi|x) = x$ for all $\xi \in x$.

Definition. A function $\varphi : \mathbf{X}_e \to [0, \infty)$ is called an *m-step interaction function* if $\varphi(x) \neq 1$ implies that x is an m-clique. Note that if φ is an m-step interaction function then it is also an $(m+1)$-step interaction function.

Theorem 2.1 *A function $f : \mathbf{X}_e \to [0, \infty)$ is an m-step Markov function if and only if there exists an m-step interaction function φ such that*

$$f(x) = \prod_{y \subset x} \varphi(y), \quad \forall x \in \mathbf{X}_e. \tag{2.3}$$

Proof. Suppose that f satisfies (2.3), then

$$\begin{aligned} f(x \cup \{\xi\}) &= \prod_{y \subset x} \varphi(y) \prod_{y \subset x} \varphi(y \cup \{\xi\}) \\ &= f(x) \prod_{y \subset x \,:\, y \cup \{\xi\} \text{ is an } m\text{-clique}} \varphi(y \cup \{\xi\}) \\ &= f(x) \prod_{y \subset x\text{: } y \text{ is an } m\text{-clique, } y = E^m(\xi|y)} \varphi(y \cup \{\xi\}) \\ &=: f(x) g(\xi, E^m(\xi|x)). \end{aligned}$$

Hence, f is m-step Markov. Conversely, suppose that f is m-step Markov. Following Ripley & Kelly (1977) we recursively define the function φ as follows. Put $\varphi(\emptyset) = f(\emptyset)$ and define

$$\varphi(x) = \begin{cases} f(x)/\prod_{y \subset x, y \neq x} \varphi(y), & \text{if } x \text{ is a } m\text{-clique}, \\ 1, & \text{otherwise}, \end{cases} \tag{2.4}$$

with the convention $0/0 = 1$. It is clear that φ is well defined and that it is an m-step interaction function. It also follows from (2.4) that if x is an m-clique, then f has the form of (2.3). To avoid trivialities, assume that $f(x) > 0$ and necessarily $f(\emptyset) > 0$. If $|x| = 1$, then $\varphi(x) = f(x)/\varphi(\emptyset)$, and f will have the form of (2.3). We prove that f has the form (2.3) by induction on the cardinality of x. Suppose that f has the right form for all $x \in \mathbf{X}_e$ with $|x| = k$. Assume that $w \in \mathbf{X}_e$ and that $|w| = k + 1$. Without loss of generality, assume that w is not an m-clique so that there exist $\xi, \eta \in w$ such that $\xi \not\sim \eta$ and there are no points $\zeta_1, \ldots, \zeta_j \in x, j < m$, such that $\xi \sim \zeta_1, \zeta_1 \sim \zeta_2, \ldots, \zeta_{j-1} \sim \zeta_j, \zeta_j \sim \eta$. Hence, $w = z \cup \{\xi\} \cup \{\eta\}$ for some $z \subset w$ with $|z| = k - 1$, and

$$\begin{aligned} f(w) &= f(z \cup \{\xi\}) g(\eta, E^m(\eta | z \cup \{\xi\})) \\ &= f(z \cup \{\xi\}) g(\eta, E^m(\eta|z)) \qquad \text{(by the definition of } \xi \text{ and } \eta) \\ &= f(z \cup \{\xi\}) \frac{f(z \cup \{\eta\})}{f(z)} \qquad \text{(since } f(w) > 0 \text{ implies } f(z) > 0) \\ &= \prod_{y \subset z \cup \{\xi\}} \varphi(y) \prod_{y \subset z} \varphi(y \cup \{\eta\}) \end{aligned}$$

$$
\begin{aligned}
&= \prod_{y \subset z \cup \{\xi\}} \varphi(y) \prod_{y \subset z} \varphi(y \cup \{\eta\}) \varphi(y \cup \{\eta\} \cup \{\xi\}) \quad (\text{since } \varphi(y \cup \{\xi\} \cup \{\eta\}) = 1) \\
&= \prod_{y \subset w} \varphi(y). \qquad \square
\end{aligned}
$$

Our main application of Theorem 2.1 is that the two-step Markov density (2.1) can be expressed in the form of (2.7) below. More precisely, since the factor $\gamma^{I(x)}$ in (2.1) is proportional to a 2-Markov function, by (2.3), there exists a two-step interaction function φ_1 such that

$$
\gamma^{I(x)} = \prod_{y \subset x} \varphi_1(y). \tag{2.5}
$$

Define the one-step interaction function

$$
\varphi_2(x) := \begin{cases} \psi(\|x_1 - x_2\|), & |x| = 2, \\ 1, & \text{otherwise.} \end{cases} \tag{2.6}
$$

(Of course, φ_2 is automatically a two-step interaction function.) Equation (2.1) can now be recast as

$$
\begin{aligned}
f(x) &= a \prod_{y \subset x} \varphi_1(y) \prod_{y \subset x} \varphi_2(y) \\
&= a \prod_{y \subset x} \Phi(y),
\end{aligned} \tag{2.7}
$$

where $\Phi := \varphi_1 \varphi_2$ is a two-step interaction function. Moreover, since the relation $\sim$ is a fixed-distance relation governed by D, any $y \in \mathbf{X}_e$ having the property $\max_{i,j} \|y_i - y_j\| > 2D$, cannot be a 2-clique. Hence, $\Phi(y) = 1$ whenever $\max_{i,j} \|y_i - y_j\| > 2D$. This fact is exploited in Section 3.

Remark. The two-step Markov interaction scheme can be cast alternatively using the realization-dependent relation $\sim_x^2$ suggested by Baddeley & Møller (1989, Section 5.6). However, as the authors of the referenced paper point out, this relation does not satisfy a condition (Condition (C2) in their paper) required by their characterization theorem for realization-dependent relations. Their characterization theorem, therefore, cannot be applied in this situation.

Remark. Theorem 2.1 can be used (in analogy to the situation of Markov point processes (Ripley & Kelly, 1977, p. 190) to demonstrate the m-Markov structure of point process with

m-Markov densities. Let $\mu_{\mathbf{X}}$ be the distribution corresponding to an m-step Markov density (2.3). Then for any Borel subset Λ of $\mathbf{X}$, $A \in \mathcal{B}_e^\Lambda$, and $y \in (\Lambda^c)_e$, the conditional probability

$$\frac{\mu_{\mathbf{X}}(\{w \in \mathbf{X}_e : w \cap \Lambda \in A, w \cap \Lambda^c = y\})}{\mu_{\mathbf{X}}(\{w \in \mathbf{X}_e : w \cap \Lambda^c = y\})} = \frac{\displaystyle\int_A \prod_{z \subset x \cup y} \varphi(z)\, d\nu_\Lambda(x)}{\displaystyle\int_{\Lambda_e} \prod_{z \subset x \cup y} \varphi(z)\, d\nu_\Lambda(x)}.$$

However, members of the products above can be decomposed into two groups. Those involving $z \subset y$ (which are contained entirely in Λ^c), and all the rest. Note that the former group does not depend on x (since any z from this group does not involve any points from x), and hence the terms in this group can be brought out of the integrals. The latter group depends on y only through those points that are in $\cup_{\xi \in \Lambda} E^m(\xi|y)$. The conditional probability therefore depends solely on those points of y that are m-neighbors of Λ. This fact, when applied to the case $m = 2$, is used in the proof of Lemma 3.2.

3. Existence and uniqueness of a Gibbs measure on $\mathbb{R}^2$

The goal of this section is to define a point process on the whole plane $\mathbb{R}^2$ featuring the characteristics given by (2.1). Following the general procedure described by Preston (1976, Chapter 6) we start with a translation-invariant potential function and define on every bounded set a distribution conditional on what is outside the set. Then we show the existence and uniqueness of a translation-invariant Gibbs measure, i.e., a measure defined on point configurations in $\mathbb{R}^2$ with the correct conditional distributions.

For any Borel $\Lambda \subset \mathbb{R}^2$, let $\Lambda_{e,b}$ denote the collection of all point configurations in Λ that are finite on bounded subsets of Λ. Let $\mathcal{B}_{e,b}^\Lambda$ be the smallest σ-algebra making the number of points in point configurations in bounded Borel subsets of Λ a random variable. Note that if Λ is bounded, then $\Lambda_{e,b} = \Lambda_e$ and $\mathcal{B}_{e,b}^\Lambda = \mathcal{B}_e^\Lambda$ as in Section 2. Let $(\Omega, \mathcal{F}) = (\mathbb{R}^2_{e,b}, \mathcal{B}_{e,b}^{\mathbb{R}^2})$, and let Ω_F denote the set of all finite point configurations in Ω. We use (2.1) to define a translation-invariant potential function V as a mapping $V : \Omega_F \to (-\infty, \infty]$ as follows:

$$V(x) := -\log(f(x)/a), \tag{3.1}$$

where $\log 0 = -\infty$. Using (2.7), we can recast V as

$$V(x) = -\sum_{y \subset x} \Big(\log \varphi_1(y) + \log \varphi_2(y)\Big). \tag{3.2}$$

Since $d > 0$ (see the definition of φ_2 in Section 2), the potential (3.1) is called a hard-core potential. Furthermore, it follows from the discussion following (2.7) that $\log \Phi(y) = 0$ whenever $\max\{\|y_i - y_j\|\} > 2D$. A potential with this property is called a finite-range potential. The hard-core requirement of V has both technical and physical significance. Physically, it prohibits point configurations that contain points that are too close, a feature that is desirable for many applications (e.g., minefield modeling, biology). Technically, the hard-core requirement makes V a stable potential; namely, there exists a constant $C(d, D) \geq 0$ such that $V(x) \geq -C(d, D)|x|$, for all $x \in \Omega_F$. The stability of V guarantees that f is a valid density function. An example of such a constant is $C(d, D) = k_{\max} \log \overline{\psi}$, where $k_{\max}$ is the maximum number of points in a disk of radius D such that no two points are closer than d. The hard-core assumption also plays an important role in the existence of a translation-invariant Gibbs measure for the potential (3.1) as we shall see later in this section.

Let ν denote the measure on $(\Omega, \mathcal{F})$ corresponding to a Poisson process in $\mathbb{R}^2$ with mean measure λ times Lebesgue measure. For any Borel subset Λ in $\mathbb{R}^2$ and $s \in \Omega$, let s_Λ and ν_Λ denote the restrictions of s and ν to Λ and $\mathcal{B}_{e,b}^{\Lambda}$, respectively. Suppose that $\beta > 0$ is given. Now for any bounded $\Lambda \subset \mathbb{R}^2$, and for a given $y \in (\Lambda^c)_{e,b}$, define a conditional measure on $\mathcal{B}_{e,b}^{\Lambda}$ by

$$\mu_\Lambda(dx|y) = \frac{\exp(-\beta V_\Lambda(x|y))}{a_{\Lambda,y}} \nu_\Lambda(dx), \tag{3.3}$$

where $a_{\Lambda,y}$ is a normalizing constant making $\mu_\Lambda(\cdot|y)$ a probability, and for $y \in (\Lambda^c)_{e,b}$ and $x \in \Lambda_{e,b}$, $V_\Lambda(x|y)$ is a conditional potential (Preston, 1976, p. 98) defined by

$$V_\Lambda(x|y) := -\lim_{A \uparrow \mathbb{R}^2} \Big\{ \sum_{z \subset x \cup y_A, z \cap x \neq \emptyset} \log \varphi_1(z) + \log \varphi_2(z) \Big\},$$

where A in the above expression is bounded and Borel. The above limit exists since its argument becomes fixed, for each x and y, as soon as $d_E(A^c, \Lambda) > 2D$, where $d_E(\cdot, \cdot)$ denotes the Euclidean distance between sets in $\mathbb{R}^2$. Hence, for any A satisfying $d_E(A^c, \Lambda) > 2D$,

$$V_\Lambda(x|y) = -\sum_{z \subset x \cup y_A, z \cap x \neq \emptyset} \log \varphi_1(z) + \log \varphi_2(z). \tag{3.4}$$

In a similar way as in the proof of the claim in the Appendix, we can show that there exist nonnegative constants q_1 and q_2 (independent of y) so that $V_\Lambda(x|y) \geq -q_1 - q_2|x|$, for all

$x \in \Lambda_{e,b}$. This fact guarantees that the conditional measures (3.3) are probabilities. We claim that $V_\Lambda(\cdot|y)$ partially inherits the property of penalizing isolated points. To see this, observe that the first sum above can be written as $\sum_{z \subset x \cup y_A} \log \varphi_1(z) - \sum_{z \subset y_A} \log \varphi_1(z)$, which is equal to $[I(y_A \cup x) - I(y_A)] \log \gamma$ (see (2.5)), and is not necessarily equal to $\log \gamma$ times the number of isolated points in $y_A \cup x$ which are in x, denoted by $I(x|y_A)$. Nonetheless, it is easy to show that $I(y_A \cup x) - I(y_A) > 0$ implies $I(x|y_A) > 0$. (The proof of this fact relies on the identity $I(y_A \cup x) = I(x|y_A) + I(y_A|x)$.) Moreover, $I(x|y_A) = I(y_A \cup x) - I(y_A)$ if and only if none of the isolated points of y is a neighbor of any point of x. Hence, the resulting conditional measure partially inherits the property of penalizing isolated points.

The existence of a Gibbs measure for the above conditional distributions relies on a result due to Klein (1984). We first introduce some of his notation.

Definition. Let C_n be the square of side length $2n$ centered at the origin, and let

$$\mathsf{U}_{nm} := \{s \in \Omega : |s \cap (\mathsf{C}_n \setminus \mathsf{C}_{n-1})| \leq m\ell(\mathsf{C}_n \setminus \mathsf{C}_{n-1})\}.$$

Let $\mathsf{D} := \{s \in \Omega : V(y) < \infty \text{ for all } y \in \Omega_F, y \subset s\}$. Define $\mathsf{U}_m := \mathsf{D} \cap (\cap_{n \geq 1} \mathsf{U}_{nm})$, and $\mathsf{U}_\infty := \cup_{m \geq 1} \mathsf{U}_m$.

Remark. For the potential given in (3.1), the set D consists of those elements of Ω for which no two points are within a distance d. Furthermore, $\mathsf{D} \subseteq \cap_{n \geq 1} \mathsf{U}_{nm}$ for large enough m, and hence $\mathsf{U}_\infty = \mathsf{D}$.

Theorem 3.1 (Klein 1984, Theorem 2.2 and Remark 2.3) *Let $\hat{V}(x) = \sum_{z \subset x} \hat{\varphi}(z)$ be any translation-invariant potential satisfying the following conditions:*

(i) $\hat{V}$ is stable, i.e., for some constant $\hat{C} \geq 0$, $\hat{V}(x) \geq -\hat{C}|x|$, for all $x \in \Omega_F$.

(ii) For any bounded Borel Λ, any $x \in \Lambda_{e,b} \cap \mathsf{U}_\infty$, and any $m \geq 1$, $|\hat{V}_\Lambda(x|s) - \hat{V}_\Lambda(x|s_{\mathsf{C}_k})| \leq \varepsilon_s(k)$, where $\varepsilon_s(k)$ converges uniformly to zero for all $s \in \mathsf{U}_m \cap (\Lambda^c)_{e,b}$ as $k \to \infty$.

(iii) For some $c > 0$, $\sum_{z \subset x \cup s, z \cap x \cap s \neq \emptyset} \hat{\varphi}(z) \geq -c|x||s|$ for all $x, s \in \Omega_F$ with $|x|, |s| \geq 1$.

(iv) There exist $r > 0$ and $l \in \mathbb{N}$ such that $\hat{\varphi}(x) = \infty$, for all x with $|x| = l$ and $\max_{i,j} \|x_i - x_j\| < r$.

Then for any nonnegative β and λ, there exists a translation-invariant Gibbs measure μ, *corresponding to the conditional measures obtained from the potential* $\hat{V}$*. Furthermore,* $\mu(\mathsf{U}_\infty) = 1$*.*

We now show that the potential V given in (3.1) satisfies conditions $(i)-(iv)$ of Theorem 3.1. The stability condition (i) has been demonstrated earlier in this section as a consequence of the hard-core property of V. By the finite-range property of V, $|V_\Lambda(x|s) - V_\Lambda(x|s_{\mathsf{C}_k})| = 0$ for all $s \in (\Lambda^c)_{e,b}$ if k is large enough so that $d_E(\mathsf{C}_k^c, \Lambda) > 2D$, and hence (ii) is satisfied. To see that (iii) is satisfied, note that from (2.5), (2.6), and (3.2) we have

$$\begin{aligned}
-\sum_{z\subset x\cup s, z\cap x\cap s\neq\emptyset} \log\Phi(z) &= -\Big\{\sum_{z\subset x\cup s}\log\varphi_1(z) - \sum_{z\subset x}\log\varphi_1(z) - \sum_{z\subset s}\log\varphi_1(z)\Big\} \\
&\quad - \sum_{\xi\in x,\eta\in s}\log\psi(\|\xi-\eta\|) \\
&= -(I(x\cup s) - I(x) - I(s))\log\gamma - \sum_{\xi\in x,\eta\in s}\log\psi(\|\xi-\eta\|) \\
&\geq -(|x|+|s|)\log\gamma^{-1} - |x||s|\log\overline{\psi} \\
&\geq -(2\log\gamma^{-1} + \log\overline{\psi})|x||s|.
\end{aligned}$$

Clearly, (iv) holds by the hard-core assumption for $l = 2$. This establishes the existence of a translation-invariant Gibbs measure corresponding to the potential given in (3.2). □

We now establish uniqueness of the Gibbs measure. Let $\{U_j\}_{j\in\mathbb{Z}^2}$ be the partition of the plane given by $U_j = \{u = (u_1, u_2) \in \mathbb{R}^2 : Dj_i \leq u_i \leq D(j_i + 1), i = 1, 2\}$. Let $\mathcal{D}_j$ be the σ-algebra generated by the point process (distributed according to μ) restricted to U_j. Following Föllmer (1982), define the interaction coefficients

$$C_{ij} := \sup\Big\{\frac{R(\mu_{U_j}(\cdot|x), \mu_{U_j}(\cdot|y))}{r(x_{U_i}, y_{U_i})} : x = y \text{ off } U_i\Big\}, \tag{3.5}$$

where $R(\cdot,\cdot)$ is the Vasserstein metric for probability measures on $((U_0)_e, \mathcal{B}_{e,b}^{U_0})$, and $r(\cdot,\cdot)$ is a metric on $(U_0)_e$ used by Jensen (1993b). (See the Appendix for a more formal definition of these metrics.) Define the matrices $\mathbf{C} := ((C_{ij}))$, and $\mathbf{D} := \sum_{k=0}^\infty \mathbf{C}^n$. Note that by the translation invariance of μ and the symmetry of $\mathbf{C}$, $C_{ij} = C_{0,j-i} = C_{0,i-j} =: \tilde{C}_{i-j}$, and $D_{ij} = D_{0,j-i} = D_{0,i-j} =: \tilde{D}_{i-j}$. For each $i \neq (0,0)$, $\tilde{C}_i$ measures the dependence on the points in U_i, of the conditional distribution of the point process in U_0 given $\{\mathcal{D}_j, j \neq (0,0)\}$. Föllmer (1982, eq. (3.10)) has shown that if $\sum_i \tilde{C}_i < 1$, then any two measures on $(\Omega, \mathcal{F})$

having the same conditional measures are identical. Jensen (1993b, Lemma 3.1) provided an estimate for $\tilde{C}_i$ for conditional measures obtained from conditional potentials of the form

$$\hat{V}_\Lambda(x|y) = \sum_{\xi,\eta \in x \cup y_\Lambda, \{\xi,\eta\} \cap x \neq \emptyset} \hat{\varphi}(\|\xi - \eta\|), \tag{3.6}$$

which is a pairwise interaction conditional potential. In our case however, the conditional potential $V_\Lambda(x|y)$ (see (3.4)) has two ingredients: φ_2 which is a pairwise interaction function; and φ_1 which is a many-body interaction function. In Lemma 3.2 below, we provide a modification to Jensen's result and obtain a sufficient condition for the inequality $\sum_i \tilde{C}_i < 1$ in terms of the parameters of the potential V.

For $0 < D_1 < D_2$, let $A(D_1, D_2)$ denote the annulus consisting of the set difference of two concentric squares (centered at the origin), the outer square having side-length D_2 and the inner one having side-length D_1. Let κ be the maximum number of points in $A(D, D+2d)$ such that every pair of points are more than d units apart. Let $\kappa_{i,5}$ and $\kappa_{i,3}$ be the maximum number of isolated points in $A(D, 5D)$, and $A(D, 3D)$, respectively. Finally, let $\hat{\kappa} = \max\{2D, D^3\}$.

Lemma 3.2 *Let* $\omega = \lambda D^2 \overline{\psi}^{\beta\kappa}$*, and let*

$$\delta = (1/48)\gamma^{-\kappa_{i,5}\beta} e^{\omega} \min\Big\{\omega + 1 - e^{-\omega}, \beta\{\hat{\kappa}(\omega^2 + 2\omega)\log \overline{\psi} + \kappa_{i,3}(\omega + 1 - e^{-\omega})\log \gamma^{-1}\}\Big\}.$$

Then $\sum_i \tilde{C}_i < \delta$*. In particular,* $\sum_i \tilde{C}_i < 1$ *if* β *is chosen small enough.*

Proof. See the Appendix.

Lemma 3.2 also plays a key role in the next section.

4. Asymptotic normality

In this section, we assume that the function ψ in (2.1), which is responsible for pairwise interaction, is piecewise constant, i.e., $\psi(r) = \sum_{i=1}^{M+1} \psi_i \mathbf{1}_{[r_{i-1}, r_i)}(r)$, where $0 = r_0 < d = r_1 < \cdots < r_M = D, r_{M+1} = \infty, \psi_1 = 0, \psi_{M+1} = 1$, and $M \geq 1$. Observe that the density (2.1) can now be recast as

$$f(x) = a\gamma^{I(x)} \prod_{i=1}^{M} \psi_i^{S_i(x)},$$

where $S_i(x)$ is the number of pairs of points that are r_{i-1} to r_i units apart, $i = 1, \ldots, M$. To motivate the analysis, suppose that we are interested in the following hypothesis testing problem: The null hypothesis is assigned to a Poisson (totally random) process while the alternative hypothesis is assigned to a process with the Gibbs distribution μ discussed in Section 3. One can form a likelihood ratio test by evaluating the likelihood function and comparing it to a threshold. Any performance analysis of such a detection scheme requires knowledge of the distribution of the likelihood function. The likelihood function depends on random variables such as $I(x)$, $|x|$, and $S_i(x)$ ($i = 1, \ldots, M$). Due to computational complexity, the distribution of these quantities is not available except in special cases. Hence, to do any kind of performance analysis, it is useful to develop asymptotic approximations for the joint distribution of the random variables $S_i(x)$, $I(x)$, and $|x|$.

Let Λ_n, $n = 1, 2, \ldots$, be a sequence of subsets of $\mathbb{R}^2$ such that $n/\ell(\Lambda_n)$ converges to a finite constant as $n \to \infty$. Consider the point process on $\mathbb{R}^2$ with Gibbs measure μ as in the previous section. Let n be fixed, and define $X := (N, S_1, \ldots, S_M, I)$, where N is the total number of points, I is the total number of isolated points, and S_i is the number of pairs of points that are within a distance r_{i-1} to r_i, ($i = 1, \ldots, M$), all in Λ_n. The asymptotic normality of X relies on Theorem 4.1 and on Lemma 3.2. Let $\mathcal{J}_n$ be the set of $j \in \mathbb{Z}^2$ for which $\Lambda_n \cap U_j$ is not empty. Following Jensen (1993b), we assume that as $n \to \infty$, $|\partial\mathcal{J}_n|/|\mathcal{J}_n| \to 0$, where $\partial\mathcal{J}_n := \{z = (z_1, z_2) \in \mathcal{J}_n : \exists t = (t_1, t_2) \notin \mathcal{J}_n \text{ with } \max_i |z_i - t_i| = 1, i = 1, 2\}$ is the boundary of $\mathcal{J}_n$. For each $k = (k_1, k_2) \in \mathbb{Z}^2$, let T_k denote a shift operator on Ω defined by $(T_k x)_j = x_j + (k_1 D, k_2 D)$.

Theorem 4.1 (Jensen 1993b, Theorem 2.2) *Let μ be a translation-invariant Gibbs measure with $\sum_k \tilde{C}_k < 1$, $\int_\Omega r(x_{U_0}, s)^2 \, d\mu(x) < \infty$ for some $s \in (U_0)_e$, and let $h : \Omega \to \mathbb{R}$ be a measurable function dependent only on the restriction of Ω to finitely many of the U_i. If for some $\alpha > 0, c > 0,$ and $b > 4/\alpha$,*

(i) $\int_\Omega |h - \int_\Omega h \, d\mu|^{2+\alpha} \, d\mu < \infty$, and

(ii) $\tilde{D}_k \leq c|k|^{4-b}$,

then, $\sum_{i \in \mathcal{J}_n} \Big(h \circ T_i - \int_\Omega h \, d\mu\Big)/|\mathcal{J}_n|^{\frac{1}{2}}$ is asymptotically normally distributed with zero mean and variance $\sum_i \mathrm{cov}_\mu(h, h \circ T_i) < \infty$.

We now apply Theorem 4.1 to the Gibbs measure corresponding to the potential V of Section 3. Let N_j and I_j be the number of points and the number of isolated points, respectively, in x which are in $U_j \cap \Lambda_n$, where U_j is defined in Section 3. Following Jensen (1993a), for $i = 1, \ldots, M$, and $x \in \Omega$, let

$$S_i^j(x) = \frac{1}{2} \sum_{\xi,\eta \in x_{(\Lambda_n \cap U_j)}, \xi \neq \eta} \mathbf{1}_{[r_{i-1}, r_i)}(\|\xi - \eta\|) + \frac{1}{2} \sum_{\xi \in x_{(\Lambda_n \cap U_j)}, \eta \in x_{(\Lambda_n \cap U_j^c)}} \mathbf{1}_{[r_{i-1}, r_i)}(\|\xi - \eta\|). \quad (4.1)$$

Let $X_j := (N_j, S_1^j, \ldots, S_M^j, I_j)$, here the dependence on x is omitted for brevity. Clearly, $X = \sum_{\mathcal{J}_n} X_j$.

Theorem 4.2 *Let δ be defined as in Lemma 3.2 and suppose that $\delta < 1$, then as $n \to \infty$, $(X - \mathsf{E}[X])/|\mathcal{J}_n|^{\frac{1}{2}}$ converges in distribution to a zero-mean normally distributed $\mathbb{R}^{M+2}$-valued random variable with covariance matrix $\sum_{j=1}^{\infty} \mathsf{E}[(X_j - \mathsf{E}[X_j])(X_j - \mathsf{E}[X_j])^T]$.*

Proof. By Lemma 3.2, $\sum_k \tilde{C}_k < \delta < 1$. Since $r(x_{U_0}, \emptyset) \leq |x_{U_0}| + 1$ (see the Appendix), and since $|x_{U_0}|$ is upper bounded almost surely by a constant as a consequence of the hard-core property, the integrability condition in the statement of Theorem 4.1 is fulfilled with $s = \emptyset$. For $x \in \Omega$, let $h(x)$ be any linear combination of the components of the vector X, and let $h_0(x)$ be the corresponding linear combination of the components of X_0. Observe that $h(x) = \sum_{i \in \mathcal{J}_n} (h_0 \circ T_i)(x)$. Note that h_0 depends only on points of x in U_0 and those U_j that are direct neighbors of U_0. By the hard-core property of V, $|h_0|$ is upper bounded almost surely by a constant. The hypothesis (i) is therefore satisfied for any α. To show that condition (ii) is met, note that $\sum_i \tilde{D}_i = \sum_{n=0}^{\infty} \sum_i (\mathbf{C}^n)_{0i} \leq \sum_{n=0}^{\infty} \delta^n = (1-\delta)^{-1}$. By taking $\alpha = 1$, $b = 5$, and using the above summability result above we establish (ii). Hence, by Theorem 4.1, $h(x)$ is asymptotically normally distributed with the the covariance matrix $\sum_k \mathrm{cov}_\mu(h_0, h_0 \circ T_k) < \infty$. Since h is an arbitrary linear combination of the components of X, the theorem follows. □

Appendix. Proof of Lemma 3.2

We first follow Jensen (1993b) and give a precise definition of the the metric r and the Vasserstein metric R introduced in Section 3. Let ρ_1 be the discrete metric on Ω, i.e,

$\rho_1(t,s) = 1$ if $t \neq s$ and zero otherwise. Define the symmetric metric ρ_2 on Ω by (assuming $|s| \leq |t|$)

$$\rho_2(s,t) := \min_{\pi \in \Sigma_{|s|,|t|}} \sum_{i=1}^{|s|} \|s_i - t_{\pi(i)}\| + |t| - |s|,$$

where $\Sigma_{n,m}, n \leq m$, is the set of all one-to-one mappings from $\{1,\ldots,n\}$ into $\{1,\ldots,m\}$. Put $r = \rho_1 + \rho_2$ and note that $r(s,\emptyset) = |s| + 1, s \neq \emptyset$, and in particular,

$$r(s,\emptyset) \leq |s| + 1, \qquad s \in \Omega. \tag{A.1}$$

Also,

$$r(s,t) \geq 1, \quad \text{for all } s,t \in \Omega. \tag{A.2}$$

For a bounded Borel subset Λ in $\mathbb{R}^2$, and any two measures θ_1 and θ_2 on $(\Lambda_{e,b}, \mathcal{B}^{\Lambda}_{e,b})$, define the Vasserstein metric

$$R(\theta_1,\theta_2) := \sup_{g:\delta(g)<\infty} \Big| \int g\, d\theta_1 - \int g\, d\theta_2 \Big| / \delta(g), \tag{A.3}$$

where for $g : \Lambda_{e,b} \to \mathbb{R}$, $\delta(g) = \sup_{s\neq t} |g(s) - g(t)|/r(s,t)$. We now proceed with the proof of Lemma 3.2 and claim that for $x \in (U_0)_e$, and $s \in (U_0^c)_{e,b}$,

$$V_{U_0}(x|s) \geq -\kappa|x| \log \overline{\psi} - \kappa_{i,5} \log \gamma^{-1}, \tag{A.4}$$

where κ and $\kappa_{i,5}$ are defined in Section 3 immediately preceding Lemma 3.2. To verify the claim we take $A = \cup_{i=(i_1,i_2)\in\mathbb{Z}, i\neq(0,0), |i_1|,|i_2|\leq 2} U_i$ which clearly satisfies $d_E(A^c, U_0) > 2D$. Now we identify $\kappa_{i,5}$ as the maximum possible number of isolated points in A (Since there can be at most three isolated points in each square U_i, and since there are 24 squares in A, $\kappa_{i,5} \leq 72$.). From (3.4) and (2.5) we have

$$\begin{aligned}
V_{U_0}(x|s) &= -\sum_{z\subset x\cup s_A, z\cap x\neq\emptyset} \log \varphi_1(z) - \sum_{\xi,\eta\in x\cup y_A, \{\xi,\eta\}\cap x\neq\emptyset} \log \psi(\|\xi - \eta\|) \\
&= \Big\{ I(x \cup y_A) - I(y_A) \Big\} \log \gamma^{-1} - \sum_{\xi,\eta\in x\cup y_A, \{\xi,\eta\}\cap x\neq\emptyset} \log \psi(\|\xi - \eta\|) \\
&\geq -I(y_A) \log \gamma^{-1} - \kappa|x| \log \overline{\psi} \\
&\geq -\kappa_{i,5} \log \gamma^{-1} - \kappa|x| \log \overline{\psi}.
\end{aligned}$$

We now establish that

$$\tilde{C}_i \leq 2\gamma^{-\kappa_{i,5}\beta} e^{\omega} \{\omega + 1 - e^{-\omega}\}. \tag{A.5}$$

As in Lemma 3.1 of Jensen (1993b), we use (A.4), (A.1), and the estimate $|g(x) - g(\emptyset)| \leq \delta(g) r(x, \emptyset)$ to obtain

$$\begin{aligned} &\left| \int g(x)\, d\mu_{U_0}(x|y) - \int g(x)\, d\mu_{U_0}(x|z) \right| \\ &\quad \leq \; 2\delta(g) \int r(x,\emptyset) \exp(\beta\{\kappa |x| \log \overline{\psi} + \kappa_{i,5} \log \gamma^{-1}\} + \lambda D^2)\, d\nu_{U_0}(x) \\ &\quad \leq \; 2\delta(g) \int (1 + |x|) \exp(\beta\{\kappa |x| \log \overline{\psi} + \kappa_{i,5} \log \gamma^{-1}\} + \lambda D^2)\, d\nu_{U_0}(x) \\ &\quad = \; 2\delta(g) \gamma^{-\kappa_{i,5}\beta} e^{\omega} \{\omega + 1 - e^{-\omega}\}. \qquad \text{(A.6)} \end{aligned}$$

Here, $y_{U_j} = z_{U_j}$ for all $j \geq 1, j \neq i$. Now using the estimate (A.6) in (3.5) and along with (A.3) establishes (A.5). To establish the estimate

$$\tilde{C}_i \leq 2\beta\gamma^{-\kappa_{i,5}\beta} e^{\omega} \left\{ \hat{\kappa}(\omega^2 + 2\omega) \log \overline{\psi} + 2\kappa_{i,3}(\omega + 1 - e^{\omega}) \log \gamma^{-1} \right\} \qquad \text{(A.7)}$$

we follow again the steps in the proof of Lemma 3.1(ii) of Jensen (1993b) and use (A.4) and (A.1) to obtain the estimate

$$\begin{aligned} &\left| \int g(x)\, d\mu_{U_0}(x|y) - \int g(x)\, d\mu_{U_0}(x|z) \right| \\ &\quad \leq \; 2\delta(g) \int r(x,\emptyset) \beta \{V_{U_0}(x|y) - V_{U_0}(x|z)\} \\ &\qquad \cdot \exp(\beta\{\kappa |x| \log \overline{\psi} + \kappa_{i,5} \log \gamma^{-1}\} + \lambda D^2)\, d\nu_{U_0}(x). \qquad \text{(A.8)} \end{aligned}$$

Using (3.4), and (2.5) we write

$$\begin{aligned} &|\{V_{U_0}(x|y) - V_{U_0}(x|z)\}| \\ &\quad = \; \Bigg| - \sum_{w \subset x \cup y_A, w \cap x \neq \emptyset} \log \varphi_1(w) + \sum_{w \subset x \cup z_A, w \cap x \neq \emptyset} \log \varphi_1(w) \\ &\qquad - \sum_{\xi,\eta \in x \cup y_A, \{\xi,\eta\} \cap x \neq \emptyset} \log \psi(\|\xi - \eta\|) + \sum_{\xi,\eta \in x \cup z_A, \{\xi,\eta\} \cap x \neq \emptyset} \log \psi(\|\xi - \eta\|) \Bigg| \\ &\quad \leq \; \rho_2(y_{U_i}, x_{U_i}) + \Big(|I(x \cup z) - I(x \cup y)| + |I(y) - I(z)| \Big) \log \gamma^{-1} \\ &\quad \leq \; \rho_2(y_{U_i}, z_{U_i}) + 2\kappa_{i,3} \log \gamma^{-1}, \qquad \text{(A.9)} \end{aligned}$$

where $\kappa_{i,3}$ is defined in Section 3 preceding Lemma 3.2. The first inequality above follows from the discussion following equation (3.3) of Jensen (1993b), and the second inequality follows from the fact that y and z agree everywhere except on U_i, and therefore $|\{I(x \cup z) - I(x \cup y)|$ is at most equal to the maximum number ($\kappa_{i,3}$) of isolated points in U_i and in those U_j

that are its immediate neighbors. A similar argument applies to the difference $|I(y) - I(z)|$. Substituting (A.9) in (A.8) yields

$$\begin{aligned} &\left| \int g(x)\, d\mu_{U_0}(x|y) - \int g(x)\, d\mu_{U_0}(x|z) \right| \\ &\quad \leq \; 2\delta(g)\beta\gamma^{-\kappa_{i,5}\beta} e^{\omega} \left\{ \hat{\kappa}(\omega^2 + 2\omega)\rho_2(y_{U_i}, z_{U_i}) \log \overline{\psi} + \kappa_{i,3}(\omega + 1 - e^{\omega}) \log \gamma^{-1} \right\}. \end{aligned} \quad \text{(A.10)}$$

Substituting (A.10) in (3.5) and using (A.3) yields (A.7). Note that the conditional distribution of points in U_0 given its surroundings depends only on U_j for those $j = (j_1, j_2)$ for which $|j_1|, |j_2| \leq 2$ (see the Remark at the end of Section 2). We conclude therefore that $\tilde{C}_i$ is nonzero at most at 24 sites. Combining this with the estimates (A.5) and (A.7) completes the proof. □

References

Baddeley, A. & Møller, J. (1989). Nearest-neighbour Markov point processes and random sets. *Int. Statist. Rev.*, vol. 57, pp. 89–121.

Carter, D. S. & and Prenter, P. M. (1972). Exponential spaces and counting processes. *Z. Wahrscheinlichkeitstheorie*, vol. 21, pp. 1–19.

Föllmer, H. (1982). A covariance estimate for Gibbs measures. *J. Functl. Anal.*, no. 46, pp. 387–395.

Jensen, J. L. (1993a). A note on asymptotic expansions for sums of over a weakly dependent random field with application to the Poisson and Strauss Processes. *Ann. Statist. Math.*, vol. 45, no. 2, pp. 353–360.

Jensen, J. L. (1993b). Asymptotic normality of estimates in spatial point processes. *Scand. J. Statist.*, vol. 20, pp. 97–109.

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

Klein, D. (1984). Convergence of grand canonical Gibbs measures. *Commun. Math. Phys.*, vol. 92, pp. 295–308.

C. J. Preston, C. J. (1976). *Random Fields.* Springer Lecture Notes in Mathematics 534. Springer, Heidelberg.

Ripley, B. D. & Kelly, F. P. (1977). Markov point processes. *J. London Math. Soc.*, vol. 2, pp. 188–192.

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

Strauss, D. J. (1986). On a general class of models for interaction. *SIAM Rev.*, vol. 28, no. 4, pp. 513–527.

Majeed M. Hayat and John A. Gubner, Department of Electrical and Computer Engineering, University of Wisconsin, Madison, WI 53706, U.S.A. (E-mail: hayat@large.ece.wisc.edu, gubner@entropy.ece.wisc.edu)