Poisson intensity parameter estimation for stationary Gibbs point processes of finite interaction range

Jean-François Coeurjolly*, Nadia Morsli

Laboratory Jean Kuntzmann, Grenoble University, France

A R T I C L E   I N F O

Article history:
Received 16 October 2012
Accepted 25 February 2013
Available online 5 March 2013

Keywords:
Central limit theorem
Georgii–Nguyen–Zessin formula
Papangelou conditional intensity
Semi-parametric estimation

A B S T R A C T

We introduce a semi-parametric estimator of the Poisson intensity parameter of a spatial stationary Gibbs point process. Under very mild assumptions satisfied by a large class of Gibbs models, we establish its strong consistency and asymptotic normality. We also consider its finite-sample properties in a simulation study.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

Spatial Gibbs point processes form a major class of stochastic processes allowing the modeling of dependence of spatial point patterns. They have applications in many scientific fields such as biology, epidemiology, geography, astrophysics, physics, and economics. General references covering theoretical as well as practical aspects of these processes are e.g. Stoyan et al. (1995), Møller and Waagepetersen (2004) or Illian et al. (2008).

Gibbs point processes in \( \mathbb{R}^d \) can be defined and characterized through the Papangelou conditional intensity (see Møller and Waagepetersen, 2004) which is a function \( \lambda : \mathbb{R}^d \times \Omega \rightarrow \mathbb{R}_+ \) where \( \Omega \) is the space of locally finite configurations of points in \( \mathbb{R}^d \). The Papangelou conditional intensity can be interpreted as follows: for any \( u \in \mathbb{R}^d \) and \( x \in \Omega \), \( \lambda(u, x) \, du \) corresponds to the conditional probability of observing a point in a ball of volume \( du \) around \( u \) given the rest of the point process is \( x \). In the present paper, we assume that the Papangelou conditional intensity is decomposed as

\[
\lambda(u, x) = \beta \, \tilde{\lambda}(u, x),
\]

© 2013 Elsevier B.V. All rights reserved.

2211-6753/$ – see front matter © 2013 Elsevier B.V. All rights reserved.

http://dx.doi.org/10.1016/j.spasta.2013.02.003
where $\beta$ is a positive real parameter and $\hat{\lambda} : \mathbb{R}^d \times \Omega \rightarrow \mathbb{R}_+$. If the point process corresponds to a homogeneous Poisson point process $\hat{\lambda}(u, x) = 1$. We further assume that $\hat{\lambda}(u, \emptyset) = 1$ which in case means that $\beta = \lambda(u, \emptyset)$ and that $\lambda$ represents the higher order interaction term. Therefore, we suggest the name Poisson intensity parameter for the parameter $\beta$. The goal of this paper is to develop an estimator of $\beta$ without specifying anything on the function $\lambda$. Our semi-parametric estimator is based on a smart application of the Georgii–Nguyen–Zessin formula (see Georgii (1976) for a general presentation and Section 2.1) which is also a way to characterize a Gibbs point process. We propose a very simple ratio-estimator of $\beta$ based on a single observation of a stationary spatial Gibbs point process observed in a bounded window and we prove its asymptotic properties (strong consistency and asymptotic normality) as the window expands to $\mathbb{R}^d$. The problem considered in this paper is related to two different works that we detail below.

First, due to its easy interpretability the most popular Gibbs model is without any doubt the (isotropic) pairwise interaction point process. Its Papangelou conditional intensity is given by

$$\lambda(u, x) = \beta \tilde{\lambda}(u, x) \quad \text{with} \quad \tilde{\lambda}(u, x) = \prod_{v \in X \setminus u} g(||v - u||). \quad (1.2)$$

The function $g$ is called the (isotropic) pairwise interaction function. Estimating the model (1.2) consists in estimating $\beta$ and the function $g$. To the best of our knowledge, this awkward problem has been considered only in Diggle et al. (1987). The authors propose several methods to estimate non parametrically the function $g$. Our objective is to make some advances in that direction which justifies first to propose an estimator of $\beta$ independently of the function $g$ and to understand its properties.

Second, many popular methods exist to estimate a parametric Gibbs model. This includes methods based on the likelihood or the pseudo-likelihood and the Takacs–Fiksel method. We refer for instance to Möller and Waagepetersen (2004), Illian et al. (2008) or Möller (2010) and the references therein for an overview of these methods. A recent contribution (Baddeley and Dereudre, in press) investigates a new approach. The following exponential model is considered

$$\lambda(u, x) = \beta \tilde{\lambda}(u, x) \quad \text{with} \quad \log \tilde{\lambda}(u, x) = \theta^\top v(u, x) \quad (1.3)$$

where $\theta$ is a real $p$-dimensional parameter vector, $v(u, x) = (v_1(u, x), \ldots, v_p(u, x))^\top$ for measurable functions $v_i : \mathbb{R}^d \times \Omega \rightarrow \mathbb{R}, i = 1, \ldots, p$. Baddeley and Dereudre (in press) develops a new estimator of the parameter $\theta$ where the parameter $\beta$ is treated as a nuisance parameter. The main advantage of the estimator is that it is obtained via the resolution of a linear system of equations whereas all previous methods cited above require an optimization procedure. The ratio-estimator of $\beta$ we propose is also very computationally cheap. Hence, it could serve as a complement of the estimate of $\theta$ proposed in Baddeley and Dereudre (in press) to completely identify the model (1.3).

The rest of the paper is organized as follows. Section 2 gives the background and notation, presents examples of Gibbs point processes and our main assumptions. Section 3 deals with the core of the paper. We present our estimator and derive its asymptotic properties as the window of observation expands to $\mathbb{R}^d$. A simulation study is conducted in Section 4 where we illustrate the efficiency of our estimator. Proofs of the results are postponed to Appendix.

2. Background and assumptions

2.1. Basic definitions and tools

A point process $X$ in $\mathbb{R}^d$ is a locally finite random subset of $\mathbb{R}^d$, i.e. the number of points $N(\Lambda) = n(X, \Lambda)$ of the restriction of $X$ to $\Lambda$ is a finite random variable whenever $\Lambda$ is a bounded Borel set of $\mathbb{R}^d$ (see Daley and Vere-Jones, 1988). We let $\Omega$ be the space of locally finite point configurations. If the distribution of $X$ is translation invariant, we say that $X$ is stationary.

The Papangelou conditional intensity completely characterizes the Gibbs point process in terms of the Georgii–Nguyen–Zessin (GNZ) Formula (see Papangelou (2009) and Zessin (2009) for historical comments and Georgii (1976) or Nguyen and Zessin (1979) for a general presentation). The GNZ
formula states that for any measurable function \( h : \mathbb{R}^d \times \Omega \to \mathbb{R} \) such that the left or right hand side exists
\[
E \sum_{u \in X} h(u, X \setminus u) = E \int_{\mathbb{R}^d} h(u, X) \lambda(u, X) \, du.
\] (2.1)

We will not discuss how to consistently specify the Papangelou conditional intensity to ensure the existence of a Gibbs point process on \( \mathbb{R}^d \), but rather we simply assume we are given a well-defined Gibbs point process. The reader interested in a deeper presentation of Gibbs point processes and the existence of a Gibbs point process on \( \mathbb{R}^d \) is referred to Ruelle (1969), Preston (1976) or Dereudre et al. (2012), see also Section 2.2 for a few examples.

The concept of innovation for spatial point processes is proposed in Baddeley et al. (2005) and is a key-ingredient in this paper. Inspired by the GNZ formula, it is defined as follows: for some measurable function \( h : \mathbb{R}^d \times \Omega \to \mathbb{R} \), the \( h \)-weighted innovation computed in a bounded domain \( \Lambda \) is the centered random variable defined by
\[
I_\Lambda(X, h) = \sum_{u \in X_\Lambda} h(u, X \setminus u) - \int_{\Lambda} h(u, X) \lambda(u, X) \, du.
\] (2.2)

We end this paragraph with a few notation. The volume of a bounded Borel set \( \Lambda \) of \( \mathbb{R}^d \) is denoted by \( |\Lambda| \). The norm \( \| \cdot \| \) stands for the standard Euclidean norm. For two subsets \( \Lambda_1 \) and \( \Lambda_2 \) of \( \mathbb{R}^d \), \( d(\Lambda_1, \Lambda_2) \) corresponds to the minimal distance between these two subsets. Finally, \( B(u, R) \) denotes the closed ball centered at \( u \in \mathbb{R}^d \) with radius \( R \).

### 2.2. Examples of Gibbs point processes

We present in this paragraph several classical examples through their Papangelou conditional intensity. Further details on these models can be found in Møller and Waagepetersen (2004) for examples (i)–(iv) and (vi)–(vii) and to Geyer (1999) for the example (v). Let \( u \in \mathbb{R}^d \), \( x \in \Omega \) and \( R > 0 \).

(i) Strauss point process.
\[
\lambda(u, x) = \beta \gamma^{n_{[0,R]}(u,x,u)}
\]
where \( \beta > 0 \), \( \gamma \in [0, 1] \) and \( n_{[0,R]}(u,x) = \sum_{v \in \Omega} 1(\|u - v\| \leq R) \) represents the number of \( R \)-closed neighbors of \( u \) in \( x \).

(ii) Strauss point process with hard-core. Let \( 0 < \delta < R \) the Papangelou conditional intensity is
\[
\lambda(u, x) = \begin{cases} 0 & \text{if } d(u, x \setminus u) \leq \delta \\ \beta \gamma^{n_{[0,R]}(u,x,u)} & \text{otherwise} \end{cases}
\]
where \( \beta > 0 \), \( \gamma \in [0, 1] \).

(iii) Piecewise Strauss point process.
\[
\lambda(u, x) = \beta \prod_{j=1}^{p} \gamma_j^{n_{[R_{j-1}, R_j]}(u,x,u)}
\]
where \( \beta > 0 \), \( \gamma_j \in [0, 1] \), \( n_{[R_{j-1}, R_j]}(u,x) = \sum_{v \in \Omega} 1(\|v - u\| \in [R_{j-1}, R_j]) \) and \( R_0 = 0 < R_1 < \cdots < R_p = R < \infty \).

(iv) Triplets point process.
\[
\lambda(u, x) = \beta \gamma^{s_{[0,R]}(x,u) - s_{[0,R]}(x,u)}
\]
where \( \beta > 0 \), \( \gamma \in [0, 1] \) and \( s_{[0,R]}(x) \) is the number of unordered triplets that are closer than \( R \).

(v) Geyer saturation point process with saturation threshold \( s \)
\[
\lambda(u, x) = \beta \gamma^{t(x,u) - t(x,u)}
\]
where \( \beta > 0 \), \( \gamma \in [0, 1] \), \( s \geq 1 \) and \( t(x) = \sum_{v \in x} \min(s, n_{[0,R]/2}(v, x \setminus v)) \).
(vi) Lennard-Jones model with finite range. This corresponds to the pairwise interaction point process defined by

\[ \lambda(u, x) = \beta \prod_{v \in x \setminus u} g(\|v - u\|) \]

with for \( r > 0 \), \( \log g(r) = (\theta^6 r^{-6} - \theta^{12} r^{-12}) \mathbb{1}_{[0, R]}(r) \), \( \theta > 0 \) and \( \beta > 0 \).

(vii) Area-interaction point process.

\[ \lambda(u, x) = \beta \gamma A(x \cup u) - A(x \setminus u) \]

for \( \beta > 0 \), \( \gamma > 0 \) and where \( A(x) = | \bigcup_{v \in x} B(v, R/2) | \).

Examples (i)–(iii) and (v)–(vi) are pairwise interaction point processes whereas example (iv) (resp. (vii)) is based on interactions on cliques of order 3 (resp. of any order). Example (vi) is the only example which is not locally stable but only Ruelle superstable, see Ruelle (1969). We recall that the local stability property states that (for stationary Gibbs models) the Papangelou conditional intensity is uniformly bounded by a constant (see Möller and Waagepetersen, 2004).

2.3. Main assumptions and discussion

In this section, we consider the model (1.1) and we present the assumptions we require to propose in Section 3.1 an estimator of the Poisson intensity parameter. To stress on the fact that the Papangelou conditional intensity satisfies (1.1) and that we aim at estimating \( \beta \) we denote, from now on, the corresponding Papangelou conditional intensity \( \lambda_\beta \).

The two following assumptions will be considered and discussed throughout the paper.

(a) The Papangelou conditional intensity has a finite range \( R \), i.e.

\[ \lambda_\beta(u, x) = \lambda_\beta(u, x_{B(u, R)}) , \]

for any \( u \in \mathbb{R}^d \), \( x \in \Omega \), \( \beta > 0 \).

(b) The function \( \lambda \) satisfies for any \( u \in \mathbb{R}^d \)

\[ \widetilde{\lambda}(u, \emptyset) = 1. \]

As this paper deals with stationary Gibbs point processes, we stress on the fact that the Papangelou conditional intensity is invariant by translation for any \( \beta > 0 \). We emphasize that no more assumption will be required. In particular we do not require that the Gibbs point process is locally stable. This is noticeable since local stability is often an assumption made when dealing with such models, see e.g. Billiot et al. (2008). We also underline in that the following, we do not assume that \( R \) is known. In place of this, we assume we are given an upper-bound of \( R \).

Let us discuss these assumptions regarding the examples presented in Section 2.2. Examples (i)–(vii) exhibit Papangelou conditional intensities which satisfy (1.1) and (2.3) with finite range \( R \). Example (i)–(vi) satisfy the assumption (2.4), whereas example (vii) is a counter-example since for the latter one \( \lambda(u, \emptyset) = \gamma \pi (R/2)^2 \neq 1 \) for \( \gamma > 0 \). Condition (b) can thus be viewed as an identifiability condition.

3. Semi-parametric estimator of \( \beta^* \)

3.1. Definition

In the following, we let \( \beta^* > 0 \) denote the true parameter and we denote by \( P_{\beta^*} \) the probability measure of \( X \), i.e. \( X \sim P_{\beta^*} \).

Our estimator is based on an application of the GNZ formula (2.1) for the following function \( h \) given for all \( u \in \mathbb{R}^d \), \( x \in \Omega \) and \( R \geq R \) by

\[ h(u, x) = \mathbb{1}(d(u, x) > \widetilde{R}) = \mathbb{1}(x \cap B(u, \widetilde{R}) = \emptyset) \]
The idea is formulated by the following result. Let \( \Lambda \) be a bounded region of \( \mathbb{R}^d \) and let \( N_A(\mathbf{X}; \tilde{R}) \) and \( V_A(\mathbf{X}; \tilde{R}) \) denote the following random variables
\[
N_A(\mathbf{X}; \tilde{R}) = \sum_{u \in X_A} 1 \left( (\mathbf{X} \setminus u) \cap B(u, \tilde{R}) = \emptyset \right)
\]
\[
V_A(\mathbf{X}; \tilde{R}) = \int_A 1 \left( \mathbf{X} \cap B(u, \tilde{R}) = \emptyset \right) du.
\]

The variable \( N_A \) corresponds to the number of points in \( X_A \) that are separated from other points in \( X \) by a distance greater than \( \tilde{R} \) while the variable \( V_A \) is the volume of \( \Lambda \) after digging a disk hole of radius \( \tilde{R} \) centered at each point of \( X \).

**Proposition 3.1.** Assume (1.1), (2.3) and (2.4), then we have for any \( \tilde{R} \geq R \)
\[
EN_A(\mathbf{X}; \tilde{R}) = \beta^* EV_A(\mathbf{X}; \tilde{R})
\]
where \( E \) denotes the expectation with respect to \( P_{\beta^*} \).

**Proof.** Using the GNZ formula (2.1) applied to the function \( h \) defined by (3.1), we obtain
\[
EN_A(\mathbf{X}; \tilde{R}) = E \sum_{u \in X_A} 1 \left( (\mathbf{X} \setminus u) \cap B(u, \tilde{R}) = \emptyset \right)
\]
\[
= \beta^* E \int_A 1 \left( \mathbf{X} \cap B(u, \tilde{R}) = \emptyset \right) \tilde{\lambda}(u, \mathbf{X}) du.
\]

Now, using the finite range property (2.3), the assumption (2.4) and the fact that \( \tilde{R} \geq R \) we continue with
\[
EN_A(\mathbf{X}; \tilde{R}) = \beta^* E \int_A 1 \left( \mathbf{X} \cap B(u, \tilde{R}) = \emptyset \right) \tilde{\lambda}(u, \mathbf{X}) \cap B(u, R) du
\]
\[
= \beta^* E \int_A 1 \left( \mathbf{X} \cap B(u, \tilde{R}) = \emptyset \right) \tilde{\lambda}(u, \emptyset) du
\]
\[
= \beta^* EV_A(\mathbf{X}; \tilde{R}). \quad \Box
\]

Using a classical ergodic theorem for spatial point processes obtained in Nguyen and Zessin (1979), we can expect that when the observation domain is large \( N_A(\mathbf{X}; \tilde{R}) \sim EN_A(\mathbf{X}; \tilde{R}) \) and \( V_A(\mathbf{X}; \tilde{R}) \sim EV_A(\mathbf{X}; \tilde{R}) \). Using this and Proposition 3.1, we are suggested a natural and computationally cheap ratio-estimator of the parameter \( \beta^* \). We define this estimator and present its asymptotic properties in the next section.

### 3.2. Asymptotic Properties

To derive asymptotic properties of our procedure when the window expands to \( \mathbb{R}^d \), we assume that \( \mathbf{X} \) is observed in \( \Lambda_n \) where \( (\Lambda_n)_{n \geq 1} \) is a sequence of cubes growing up to \( \mathbb{R}^d \).

When \( u \) is a point close to the boundary of the window \( \Lambda_n \), we can only evaluate \( h(u, \mathbf{X} \cap \Lambda_n \cap B(u, \tilde{R})) \). One simple strategy for eliminating the edge effect bias is the border method, i.e. by considering the erosion of \( \Lambda_n \) by \( R \) defined by
\[
\Lambda_{n,\tilde{R}} = \Lambda_n \ominus \tilde{R} = \{ u \in \Lambda_n : B(u, \tilde{R}) \subseteq \Lambda_n \}.
\]

Then we define our estimator by
\[
\hat{\beta}_n = \frac{N_{\Lambda_{n,\tilde{R}}}(\mathbf{X}; \tilde{R})}{V_{\Lambda_{n,\tilde{R}}}(\mathbf{X}; \tilde{R})}.
\]

(3.3)
We let $F$ denote the empty space function (see Møller and Waagepetersen, 2004) given, in the stationary case, by

$$F(r) = P_{\beta^*} (X \cap B(0, r) \neq \emptyset) = 1 - P_{\beta^*} (X \cap B(0, r) = \emptyset)$$

for $r > 0$. Following this definition, by $F_{u, v} : \mathbb{R}_+ \to [0, 1]$ for any $u, v \in \mathbb{R}^d$ we denote the function given by

$$F_{u, v}(r) = 1 - P_{\beta^*} (X \cap B(u, r) = \emptyset, X \cap B(v, r) = \emptyset)$$

for $r > 0$. We now state our main result.

**Proposition 3.2.** Assume (1.1), (2.3) and (2.4), and let $\tilde{R} \geq R$. Then we have the following statements as $n \to \infty$.

(i) $\hat{\beta}_n$ is a strongly consistent estimator of $\beta^*$.

(ii) If $P_{\beta^*}$ is ergodic, the following convergence in distribution holds

$$|\Lambda_n, \tilde{R}|^{1/2} (\hat{\beta}_n - \beta^*) \overset{d}{\to} \mathcal{N}(0, \sigma^2) \quad (3.4)$$

where

$$\sigma^2 = \sigma^2(\beta^*, \tilde{R}) = \frac{\beta^*}{1 - F(\tilde{R})} + \frac{\beta^{*2}}{(1 - F(\tilde{R}))^2} \int_{B(0, \tilde{R})} (1 - F_{0, v}(\tilde{R})) \, dv. \quad (3.5)$$

(iii) The following convergence in distribution holds

$$|\Lambda_n, \tilde{R}|^{1/2} \left( \frac{\hat{\beta}_n - \beta^*}{\hat{\sigma}_n} \right) \overset{d}{\to} \mathcal{N}(0, 1) \quad (3.6)$$

where

$$\hat{\sigma}_n^2 = \hat{\sigma}_n^2(X; \tilde{R}) = |\Lambda_n, \tilde{R}| \left( \frac{\hat{\beta}_n}{V_{\Lambda_n, \tilde{R}}(X; \tilde{R})} + \frac{\hat{\beta}_n^2 W_{\Lambda_n, \tilde{R}}(X; \tilde{R})}{V_{\Lambda_n, \tilde{R}}^2(X; \tilde{R})} \right) \quad (3.7)$$

is a strongly consistent estimator of $\sigma^2$ where

$$W_{\Lambda_n, \tilde{R}}(X; \tilde{R}) = \int_{\Lambda_n, \tilde{R}} \int_{B(u, \tilde{R}) \cap \Lambda_n, \tilde{R}} 1 (X \cap B(u, \tilde{R}) = \emptyset) 1 (X \cap B(v, \tilde{R}) = \emptyset) \, du \, dv.$$

If the measure $P_{\beta^*}$ is not ergodic, then it can be represented as a mixture of ergodic measures (see Georgii (1988, Theorem 14.10)). In this case, the asymptotic distribution in (3.4) becomes a mixture of Gaussian distributions. Since a mixture of standard Gaussian distribution is a standard Gaussian distribution, this explains why $P_{\beta^*}$ is assumed to be ergodic for the result (ii) and omitted for the result (iii).

We underline the interesting and unexpected fact that the asymptotic variance $\sigma^2(\beta^*, \tilde{R})$ (and its estimate) are independent on $\lambda$, which is the unspecified part of the Papangelou conditional intensity.

### 4. Simulation study

In this section, we investigate the efficiency of the Poisson intensity estimate in a simulation study. We consider the models (i)–(v) described in Section 2.2. We set $R = 0.05$ and $\beta^* = 200$ for all the examples except for the model g2 where we set $\beta^* = 50$. We consider

- Strauss point process where models s1 and s2 have respectively $\gamma = 0.2$ and $\gamma = 0.8$.
- Strauss point process with hard-core with $\delta = R/2$ and where models shc1 and shc2 have respectively $\gamma = 0.2$ and $\gamma = 0.8$. 

• Piecewise Strauss point process with \( p = 3 \) and \((R_1, R_2, R_3) = (R/3, 2R/3, R)\) and where models ps1 and ps2 have respectively \((\gamma_1, \gamma_2, \gamma_3) = (0.8, 0.5, 0.2)\) and \((\gamma_1, \gamma_2, \gamma_3) = (0.2, 0.8, 0.2)\).

• Triplet point process where models t1 and t2 have respectively \( \gamma = 0.2 \) and \( \gamma = 0.8 \).

• Geyer saturation point process with saturation threshold \( s = 1 \) where models g1 and g2 have respectively \( \gamma = 0.5 \) and \( \gamma = 1.5 \).

The estimate of the Poisson intensity we propose depends on a parameter \( \tilde{R} \) and ought to work for any value larger than \( R \). To illustrate this, \( \beta^* \) is estimated for different values of \( R \). We consider \( \tilde{R} = R \times p \), where \( p \) is a parameter varying from 0.9 to 1.2. Each replication has been generated on the domain \([0, L]^2\) and the estimates are computed on \([\tilde{R}, L - \tilde{R}]^2\). To illustrate the convergence as the window grows up, the simulations are done with \( L = 1 \) and \( L = 2 \). The simulations are done using the R package \texttt{spatstat} (Baddeley and Turner, 2005). The empirical results based on 500 replications are presented in Table 1.

The results are very satisfactory. We observe that for values of \( \tilde{R} \geq R \) (i.e. for values equal or larger than the finite range), the estimates are almost unbiased whereas a non negligible bias is observed when \( R < \tilde{R} \). The bias is negative when the point process is repulsive (i.e. when \( \lambda(u, x) \geq 1 \)) and positive when the point process is attractive (i.e. when \( \lambda(u, x) \leq 1 \)) in particular for the model g2.

As underlined by one of the reviewers the sign of the bias when \( \tilde{R} \geq R \) is violated can be explained by the fact that \( V \) is underestimated for attractive point processes and overestimated for repulsive point processes. The closer \( R \) is to the finite range \( \tilde{R} \), the lower the standard deviations are. The number of points with no \( \tilde{R} \)-close neighbors is lower and lower as \( \tilde{R} \) grows up, which explains this variance increase. Finally, we can also highlight that the asymptotic results are partly checked. When the window varies from \([0, 1]^2\) to \([0, 2]^2\), the standard deviations of the estimates are divided by 2 which turns out to be the square root of the ratios of the domains areas.

Table 1 points out the dangers to choose a wrong value for the finite range parameter \( R \). From a practical point of view, if this parameter is not set by the user, it has to be estimated. When the Papangelou conditional intensity is entirely modeled the parameters can be estimated using for instance the maximum likelihood or the pseudo-likelihood (e.g. Möller and Waagepetersen, 2004) and the irregular parameters (such as the finite range \( R \)) can be estimated using the profile pseudo-likelihood (Baddeley and Turner, 2000). Such an approach cannot be used here since we do not want to model the higher order interaction terms. We noticed that for a given (repulsive) point pattern our estimate is increasing with the value of \( \tilde{R} \) until a certain value where the regime is changing. Therefore we propose to estimate the finite range parameter \( R \) as follows: (a) Select a reasonable grid of values for \( \tilde{R} \) (we chose \([0.02, 0.08]\) in our example). (b) Compute \( \hat{\beta}(\tilde{R}) \) for each value of the grid. (c) We consider a piecewise linear regression of \( \hat{\beta}(\tilde{R}) \) in terms of \( \tilde{R} \) and we define \( R \) as the breakpoint on the slope (for this step we used the R package \texttt{segmented} (Muggeo, 2008)). Empirical results for the models s1 and s2 are presented in Table 2. Using the strategy described above, we manage to estimate \( R \) quite well. The resulting estimates of the parameter \( \beta^* \) seem to be still unbiased. We note that the standard deviation is greater than in the case where \( R \) is exactly known (i.e. column \( p = 1 \) in Table 1). However the loss of efficiency is quite reasonable.

Finally, we investigate the finite sample properties of the estimator of the asymptotic variance of \( \hat{\beta} \) given by (3.7) and the asymptotic normality result (3.6). Table 3 focuses on the models s1 and s2 and displays the 95% empirical coverage rates (i.e. the fraction of asymptotic confidence intervals constructed from the asymptotic normality result (3.6) covering the true parameter value, here \( \beta^* = 200 \)). As expected, the theoretical coverage rate is respected whatever the value of \( \tilde{R} \geq R \) and far from the result when \( R = 90\% R \). Despite Proposition 3.2 does not cover this case, it is interesting to note that the empirical results are still quite satisfactory if we replace \( R \) by its estimated value.

Appendix A. Proof of Proposition 3.2

We use the following additional notation in the proofs. The notation \( \sum_{x \neq y} \) stands for summation over distinct pairs of points, \( \Lambda^c \) denotes the complementary set of \( \Lambda \) in \( \mathbb{R}^d \). Let \( \tau = (\tau_x)_{x \in \mathbb{R}^d} \) be the shift group, where \( \tau_x \) is the translation by the vector \(-x \in \mathbb{R}^d\). For brevity, when there is no ambiguity, we skip the dependence on \( X \) and on \( R \) of the variables \( N^\Lambda_{n,R}, V^\Lambda_{n,R}, W^\Lambda_{n,R}, \sigma^2_{n,R} \).
Table 1
Mean and standard deviations of estimates of the Poisson intensity parameter $\beta^*$ for different parameters $\bar{R}$ based on 500 replications of different models generated on the window $[0, L]^2$ and estimated on the window $[\bar{R}, L - \bar{R}]^2$ for $L = 1, 2$. The second column, $\pi$, indicates the average number of points (over the 500 replications) falling into $[0, L]^2$, i.e., the Monte-Carlo estimation of $E_n(X \cap [0, L]^2)$.

<table>
<thead>
<tr>
<th>Model</th>
<th>$\pi$</th>
<th>Mean (std dev.) for $\bar{R} = R \times p$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$p = 0.9$</td>
</tr>
<tr>
<td>$s_1$</td>
<td>$L = 1$</td>
<td>99</td>
</tr>
<tr>
<td></td>
<td>$L = 2$</td>
<td>393</td>
</tr>
<tr>
<td>$s_2$</td>
<td>$L = 1$</td>
<td>156</td>
</tr>
<tr>
<td></td>
<td>$L = 2$</td>
<td>622</td>
</tr>
<tr>
<td>$shc_1$</td>
<td>$L = 1$</td>
<td>94</td>
</tr>
<tr>
<td></td>
<td>$L = 2$</td>
<td>379</td>
</tr>
<tr>
<td>$shc_2$</td>
<td>$L = 1$</td>
<td>130</td>
</tr>
<tr>
<td></td>
<td>$L = 2$</td>
<td>514</td>
</tr>
<tr>
<td>$ps_1$</td>
<td>$L = 1$</td>
<td>111</td>
</tr>
<tr>
<td></td>
<td>$L = 2$</td>
<td>445</td>
</tr>
<tr>
<td>$ps_2$</td>
<td>$L = 1$</td>
<td>134</td>
</tr>
<tr>
<td></td>
<td>$L = 2$</td>
<td>535</td>
</tr>
<tr>
<td>$t_1$</td>
<td>$L = 1$</td>
<td>159</td>
</tr>
<tr>
<td></td>
<td>$L = 2$</td>
<td>634</td>
</tr>
<tr>
<td>$t_2$</td>
<td>$L = 1$</td>
<td>178</td>
</tr>
<tr>
<td></td>
<td>$L = 2$</td>
<td>727</td>
</tr>
<tr>
<td>$g_1$</td>
<td>$L = 1$</td>
<td>110</td>
</tr>
<tr>
<td></td>
<td>$L = 2$</td>
<td>440</td>
</tr>
<tr>
<td>$g_2$</td>
<td>$L = 1$</td>
<td>70</td>
</tr>
<tr>
<td></td>
<td>$L = 2$</td>
<td>276</td>
</tr>
</tbody>
</table>

Table 2
Mean and standard deviations of estimates of the finite range parameter and related Poisson intensity parameter estimates (i.e., $\hat{R}(\hat{R})$) based on 500 replications of Strauss models on the window $[0, L]^2$.

<table>
<thead>
<tr>
<th></th>
<th>$\hat{R}$ Mean (std dev.)</th>
<th>$\hat{\beta}$ Mean (std dev.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s_1$ $L = 1$</td>
<td>0.052 (0.004)</td>
<td>197.7 (38.8)</td>
</tr>
<tr>
<td>$L = 2$</td>
<td>0.051 (0.003)</td>
<td>198.9 (22.1)</td>
</tr>
<tr>
<td>$s_2$ $L = 1$</td>
<td>0.051 (0.004)</td>
<td>201.7 (38.9)</td>
</tr>
<tr>
<td>$L = 2$</td>
<td>0.051 (0.003)</td>
<td>198.4 (21.9)</td>
</tr>
</tbody>
</table>

Table 3
Empirical coverage rates (i.e., the fraction of 95% confidence intervals covering the true parameter value $\beta^* = 200$) based on 500 replications of Strauss models on the window $[0, L]^2$.

<table>
<thead>
<tr>
<th>$\bar{R} = p \times R$</th>
<th>$\hat{R}$ Mean (std dev.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p = 0.9$</td>
<td></td>
</tr>
<tr>
<td>$p = 1$</td>
<td></td>
</tr>
<tr>
<td>$p = 1.1$</td>
<td></td>
</tr>
<tr>
<td>$p = 1.2$</td>
<td></td>
</tr>
<tr>
<td>$s_1$ $L = 1$</td>
<td>77.4%</td>
</tr>
<tr>
<td>$L = 2$</td>
<td>40.0%</td>
</tr>
<tr>
<td>$s_2$ $L = 1$</td>
<td>82.6%</td>
</tr>
<tr>
<td>$L = 2$</td>
<td>69.8%</td>
</tr>
</tbody>
</table>

Proof. (i) There exists at least one stationary Gibbs measure. If this measure is unique, it is ergodic. Otherwise, it can be represented as a mixture of ergodic measures (see Georgii (1988, Theorem 14.10)). Therefore, we can assume, for this proof, that $P_{\beta^*}$ is ergodic. We apply a general ergodic theorem for spatial point processes obtained by Nguyen and Zessin (1979). For this, we have to check that $N_\Lambda$ and $V_\Lambda$ are additive, invariant by translation and satisfy $EN_{\Lambda_0} < \infty$ and $EV_{\Lambda_0} < \infty$, where for instance
\( \Omega_0 \) is the unit cube \([0, 1]^d \). The additivity property is straightforward. Regarding the invariance by translation, let \( x \in \Omega_0, \Lambda \subset \mathbb{R}^d \) and \( y \in \mathbb{R}^d \), we have

\[
N_{\tau^y \Lambda}(\tau_y x; \tilde{R}) = \sum_{u \in \tau^y \Lambda} 1_{\tau^y \Lambda}(u) \left( d(u, \tau_y x \setminus u) > \tilde{R} \right)
\]

\[
= \sum_{u \in \tau^y \Lambda} 1_{\Lambda}(u + y) \left( d(u, \tau_y x \setminus u) > \tilde{R} \right)
\]

\[
= \sum_{v \in \tau^y \Lambda} 1_{\Lambda}(v) \left( d(v - y, \tau_y x \setminus (v - y)) > \tilde{R} \right)
\]

\[
= \sum_{v \in \Lambda} 1_{\Lambda}(v) \left( d(v, x \setminus v) > \tilde{R} \right) = N_{\Lambda}(x; \tilde{R}).
\]

Following similar arguments, we also have \( V_{\tau^y \Lambda}(\tau_y x; \tilde{R}) = V_{\Lambda}(x; \tilde{R}) \). Let us now focus on the integrability conditions. First, we have

\[
E |V_{\Lambda_0}| = EV_{\Lambda_0} = E \int_{\Lambda_0} 1_{(X \cap B(u, \tilde{R}) = \emptyset)} \, du = |\Lambda_0|(1 - F(\tilde{R})) = 1 - F(\tilde{R}) < \infty.
\]

Second, from Proposition 3.1

\[
E |N_{\Lambda_0}| = EN_{\Lambda_0} = \beta^* EV_{\Lambda_0} = \beta^*(1 - F(\tilde{R})).
\]

Therefore as \( n \to \infty \), we derive the following almost sure convergences

\[
|A_{n, \tilde{R}}|^{-1} N_{A_{n, \tilde{R}}} \longrightarrow EN_{\Lambda_0} = \beta^*(1 - F(\tilde{R}))
\]

\[
|A_{n, \tilde{R}}|^{-1} V_{A_{n, \tilde{R}}} \longrightarrow EV_{\Lambda_0} = 1 - F(\tilde{R}). \tag{A.1}
\]

Since the function \((y, z) \mapsto \frac{y}{z}\) is continuous for any \((y, z)\) in the open set \(\{(y, z) \in \mathbb{R}^2 : z \neq 0\}\), we deduce the expected result for any fixed \(\tilde{R} \geq R\).

(ii) From the definition of the innovations process (2.2), we have

\[
|A_{n, \tilde{R}}|^{-1/2} (\beta_n - \beta^*) = |A_{n, \tilde{R}}|^{1/2} \frac{N_{A_{n, \tilde{R}}} - \beta^* V_{A_{n, \tilde{R}}}}{V_{A_{n, \tilde{R}}}^{1/2}} = |A_{n, \tilde{R}}|^{-1/2} I_{A_{n, \tilde{R}}}(X, h) \frac{A_{n, \tilde{R}}^{-1/2} V_{A_{n, \tilde{R}}}}{A_{n, \tilde{R}}^{-1} V_{A_{n, \tilde{R}}}}. \tag{A.2}
\]

Henceforth from Slutsky’s Theorem and (A.1), the result will be deduced if we prove that \( |A_{n, \tilde{R}}|^{-1/2} I_{A_{n, \tilde{R}}}(X, h) \) tends to a zero mean Gaussian distribution. This will be done by combining two papers Coeurjolly et al. (2012) and Coeurjolly and Rubak (in press) dealing with asymptotic normality and variance calculation of innovations processes. For brevity, we write \( I_A \) in place of \( I_{A_{\Lambda_0}} \).

Following Coeurjolly et al. (2012, Lemma 3), we have to assume that \( P_{\beta^*} \) is ergodic and to check that for any bounded domain \( \Lambda \subset \mathbb{R}^d \), (I) \( E|I_A|^3 < \infty \), (II) \( E|I_n|^2 \to 0 \) for any sequence of bounded domains \( I_n \subset \mathbb{R}^d \) such that \( I_n \to 0 \) as \( n \to \infty \) and (III) \( I_A \) depends on \( X_{A_{\Lambda_0}} \). (III) follows from the finite range property (2.3) and the definition of \( h \) given by (3.1). We now focus on verifying (I). We have from Cauchy–Schwarz inequality

\[
E |I_A|^3 \leq 4 \left( EN_A^3 + \beta^* EV_A^3 \right). \tag{A.3}
\]

On the one hand, we have since for any \( m \geq 1, h^m(u, x) = h(u, x) \)

\[
EN_A^3 = EN_A + 3E \sum_{u, v \in X_A} h(u, x \setminus u) h(v, x \setminus v)
\]

\[
+ E \sum_{u, v, w \in X_A} h(u, x \setminus u) h(v, x \setminus v) h(w, x \setminus w). \tag{A.4}
\]
Let $T_1, T_2$ and $T_3$ denote the three terms of the right-hand side of (A.4). Following the proof of (i) in Proposition 3.2, $ET_i \leq \beta^*|\Lambda|$. Now, for any $u, v \in \mathbb{R}^d$ and $x \in \Omega$, we define the second order Papanastasiou intensity by

$$\lambda_{\beta^*}(u, v, x) = \lambda_{\beta^*}(u, x)\lambda_{\beta^*}(v, x \cup u) = \lambda_{\beta^*}(v, x)\lambda_{\beta^*}(u, x \cup v).$$

By using iterated versions of the GNZ formula (2.1), we derive the following computations

$$T_2 = 3E \int_A \int_A h(u, x \cup v)h(v, x \cup u)\lambda_{\beta^*}(u, v, x)dvdu$$

$$= 3\beta^2E \int_A \int_A (d(u, x \cup v) > \tilde{R})(d(v, x \cup u) > \tilde{R})\lambda_{\beta^*}(u, x)\lambda_{\beta^*}(v, x \cup u)dvdu$$

$$= 3\beta^2E \int_A \int_A (d(u, x \cup v) > \tilde{R})(d(v, x \cup u) > \tilde{R})\lambda_{\beta^*}^2(u, 0)dvdu$$

$$\leq 3\beta^2|\Lambda|^2.$$  

Using similar ideas, we leave the reader to check that $T_3 \leq \beta^3|\Lambda|^3$. On the other hand, we also check that

$$EV^3_1 = E \int_A \int_A \int_A (d(u, x) > \tilde{R})(d(v, x) > \tilde{R})(d(w, x) > \tilde{R})dvdwdu$$

$$\leq |\Lambda|^3.$$  

We finally obtain

$$E|\Lambda|^3 \leq 4\beta^2|\Lambda| + 12\beta^2|\Lambda|^2 + 8\beta^3|\Lambda|^3 < \infty$$

whereby (i) is checked. Following the proof of (I), we may prove that for any bounded domain $\Lambda$

$$ET_2^2 \leq 2\beta^*|\Lambda| + 4\beta^2|\Lambda|^2$$

which proves (II). Therefore we can invoke (Coeurjolly et al., 2012, Lemma 3) to obtain the asymptotic normality of $\tilde{\beta}_n$.

The derivation of the asymptotic variance of $|\Lambda_{n,R}|^{-1}Var|\Lambda_{n,R}|(X, h)$ is done using Lemma B.1 in Appendix B with the function $g = h$ given by (3.1). We omit the details and leave the reader to check that $h$ fulfills the integrability assumptions in Lemma B.1 and focus on the derivations of the quantities $A_1, A_2$ and $A_3$. First, it is clear that $A_1 = \beta^*(1 - F(\tilde{R}))$. Second, note that from (2.3) we have for any $v \in B^c(0, R)$ and $x \in \Omega$

$$\lambda_{\beta^*}(0, x)\lambda_{\beta^*}(v, x) - \lambda_{\beta^*}(0, v, x) = 0.$$  

Since $B(0, R) \subseteq B(0, \tilde{R})$

$$E \int_{B(0, \tilde{R}) \setminus B(0, R)} (\lambda_{\beta^*}(0, x)\lambda_{\beta^*}(v, x) - \lambda_{\beta^*}(0, v, x))dv = 0$$

which means that we can substitute $R$ by $\tilde{R}$ in the term $A_2$. Then, we derive

$$A_2 = E \int_{B(0, \tilde{R})} \mathbb{1}(d(0, x) > \tilde{R})\mathbb{1}(d(v, x) > \tilde{R})(\beta^2 - \lambda_{\beta^*}(0, v, x))dv.$$  

Third,

$$A_3 = E \int_{B(0, \tilde{R})} (\mathbb{1}(d(0, x \cup v) > \tilde{R}) - \mathbb{1}(d(0, x) > \tilde{R}))$$

$$\times (\mathbb{1}(d(v, x \cup v) > \tilde{R}) - \mathbb{1}(d(v, x) > \tilde{R}))\lambda_{\beta^*}(0, v, x)dv$$

$$= E \int_{B(0, \tilde{R})} \mathbb{1}(d(0, x) > \tilde{R})\mathbb{1}(d(v, x) > \tilde{R})\lambda_{\beta^*}(0, v, x)dv.$$
which leads to
\[ A_2 + A_3 = \beta^* E \int_{B(0, \tilde{R})} \mathbb{1} (d(0, \mathbf{X}) > \tilde{R}) \mathbb{1} (d(v, \mathbf{X}) > \tilde{R}) \, dv \]
\[ = \beta^* \int_{B(0, \tilde{R})} (1 - F_{0,v}(\tilde{R})) \, dv. \]

In other words
\[ |\Lambda_{n, \tilde{R}}|^{-1} \text{Var} \Lambda_{n, \tilde{R}}(\mathbf{X}, h) \to \beta^* (1 - F(\tilde{R})) + \beta^2 \int_{B(0, \tilde{R})} (1 - F_{0,v}(\tilde{R})) \, dv \]
which leads to the result from (A.1) and (A.2).

(iii) We check that the proposed estimate \( \hat{\sigma}_n^2 \) tends almost surely towards \( \sigma^2 \) as \( n \to \infty \) (a convergence in probability would be sufficient). From (3.7), (A.1) and (3.5), we need to prove that
\[ |\Lambda_{n, \tilde{R}}|^{-1} W_{\Lambda_{n, \tilde{R}}} \to \int_{B(0, \tilde{R})} (1 - F_{0,v}(\tilde{R})) \, dv. \]

We decompose \( W_{\Lambda_{n, \tilde{R}}} \) as \( W_{\Lambda_{n, \tilde{R}}}^{(1)} + W_{\Lambda_{n, \tilde{R}}}^{(2)} \) where
\[ W_{\Lambda_{n, \tilde{R}}}^{(1)} = \int_{A_{n, \tilde{R}}} \int_{B(u, \tilde{R})} \mathbb{1} (d(u, \mathbf{X}) > \tilde{R}) \mathbb{1} (d(v, \mathbf{X}) > \tilde{R}) \, dv \, du \]
\[ W_{\Lambda_{n, \tilde{R}}}^{(2)} = \int_{A_{n, \tilde{R}} \setminus A_{n, \tilde{R}}} \int_{B(u, \tilde{R}) \cap A_{n, \tilde{R}}} \mathbb{1} (d(u, \mathbf{X}) > \tilde{R}) \mathbb{1} (d(v, \mathbf{X}) > \tilde{R}) \, dv \, du \]
where \( A_{n, \tilde{R}} \) is the set \( \{ u \in A_{n, \tilde{R}} : B(u, \tilde{R}) \subseteq A_{n, \tilde{R}} \} \). Regarding the first term, we can now apply the ergodic theorem to prove that almost surely
\[ |\Lambda_{n, \tilde{R}}|^{-1} W_{\Lambda_{n, \tilde{R}}}^{(1)} \to E \int_{B(0, \tilde{R})} \mathbb{1} (d(0, \mathbf{X}) > \tilde{R}) \mathbb{1} (d(v, \mathbf{X}) > \tilde{R}) \, dv \]
where the latter comes from the finite range assumption (2.3) and the fact that \( \tilde{R} \geq R \). Since \( |\Lambda_{n, \tilde{R}}| \sim |\Lambda_{n, \tilde{R}}| \) as \( n \to \infty \), the proof is done by verifying that \( |\Lambda_{n, \tilde{R}}|^{-1} W_{\Lambda_{n, \tilde{R}}}^{(2)} \to 0 \) which is straightforward since
\[ |\Lambda_{n, \tilde{R}}|^{-1} W_{\Lambda_{n, \tilde{R}}}^{(2)} \leq \frac{|\Lambda_{n, \tilde{R}} \setminus A_{n, \tilde{R}}|}{|\Lambda_{n, \tilde{R}}|} |B(u, \tilde{R})| = o(1). \]

Appendix B. Auxiliary lemma

The following result is a particular case of Coeurjolly and Rubak (in press, Proposition 3.2).

**Lemma B.1.** Assume the Papangelou conditional intensity of the point process \( \mathbf{X} \) satisfies (1.1) and (2.3) with finite range \( R \). Let \( g : \mathbb{R}^d \times \Omega \to \mathbb{R} \) be a nonnegative measurable function satisfying for any \( u \in \mathbb{R}^d \) and \( \mathbf{X} \in \Omega \)
\[ g(u, \mathbf{X}) = g(0, \tau_u \mathbf{X}) \quad \text{and} \quad g(u, \mathbf{X}) = g(u, \mathbf{X} \ominus B(u, \tilde{R})) \]
with \( \tilde{R} \geq R \). We also assume that the following integrability conditions are fulfilled
\[ E[g^2(0, \mathbf{X}) \lambda_{\beta^*}(0, \mathbf{X})] < \infty \]
\[ E \int_{B(0, \tilde{R})} g(0, \mathbf{X}) g(v, \mathbf{X}) |\lambda_{\beta^*}(0, \mathbf{X})\lambda_{\beta^*}(v, \mathbf{X}) - \lambda_{\beta^*}(0, v, \mathbf{X})| \, dv < \infty \]
\[ E \int_{B(0, \tilde{R})} \left| (g(0, \mathbf{X} \cup v) - g(0, \mathbf{X})) (g(v, \mathbf{X} \cup 0) - g(v, \mathbf{X})) \right| \lambda_{\beta^*}(0, v, \mathbf{X}) \, dv < \infty. \]
Then, as \( n \to \infty \)

\[
|A_n|^{-1} \text{Var} \Lambda_n (X, g) \longrightarrow A_1 + A_2 + A_3
\]

where

\[
A_1 = E[g^2(0, X)\lambda_{\beta^*}(0, X)]
\]

\[
A_2 = E \int_{B(0, R)} g(0, X)g(v, X)(\lambda_{\beta^*}(0, X)\lambda_{\beta^*}(v, X) - \lambda_{\beta^*}(0, v, X))dv
\]

\[
A_3 = E \int_{B(0, R)} (g(0, X \cup v) - g(0, X))(g(v, X \cup 0) - g(v, X))\lambda_{\beta^*}(0, v, X)dv.
\]

References


