Parametric estimation of pairwise Gibbs point processes with infinite range interaction

Jean-François Coeurjolly, Frédéric Lavancier

Parametric estimation of pairwise Gibbs point processes with infinite range interaction.

Jean-François Coeurjolly[1] and Frédéric Lavancier[2]

[1]Laboratory Jean Kuntzmann, Grenoble Alpes University, France,
Jean-Francois.Coeurjolly@upmf-grenoble.fr.
[2]Laboratoire de Mathématiques Jean Leray, Nantes University, France,
frederic.lavancier@univ-nantes.fr.

Abstract

This paper is concerned with statistical inference for infinite range interaction Gibbs point processes and in particular for the large class of Ruelle superstable and lower regular pairwise interaction models. We extend classical statistical methodologies such as the pseudolikelihood and the logistic regression methods, originally defined and studied for finite range models. Then we prove that the associated estimators are strongly consistent and satisfy a central limit theorem. To this end, we introduce a new central limit theorem for almost conditionally centered triangular arrays of random fields.

Keywords: Lennard-Jones potential; pseudolikelihood; central limit theorem.

1 Introduction

Spatial Gibbs point processes are an important class of models used in spatial point pattern analysis (Lieshout, 2000; Møller and Waagepetersen, 2004; Illian et al., 2008). Gibbs point processes can be viewed as modifications of the Poisson point process in order to introduce dependencies, such as attraction or repulsion, between points. These models initially arise from statistical physics to approximate the interaction between pairs of particles (Ruelle, 1969; Preston, 1976; Georgii, 1988). The most well-known example is the Lennard-Jones model (Lennard-Jones, 1924) which yields repulsion at short scales and attraction at long scales.

Assuming that the Gibbs model has a parametric form, an important question concerns the estimation of the parameters from a realization of the point process observed on a finite subset of $\mathbb{R}^d$. Popular solutions include likelihood (e.g. Ogata and Tanemura, 1981; Huang and Ogata, 1999), pseudolikelihood (e.g. Besag, 1977; Jensen and Møller, 1991; Baddeley and Turner, 2000) and logistic regression (Baddeley et al., 2014). The two latter methods are more interesting from a practical point of view as they avoid the computation of the normalizing constant in the likelihood, which is in most cases inaccessible for Gibbs point processes and

must be approximated by simulation-based methods. We focus in this paper on the pseudolikelihood and logistic regression methods.

When the Gibbs model is assumed to have a finite range interaction, consistency and asymptotic normality of the pseudolikelihood and logistic regression estimators are established in Jensen and Møller (1991); Jensen and Künsch (1994); Billiot et al. (2008); Dereudre and Lavancier (2009); Coeurjolly and Drouilhet (2010); Baddeley et al. (2014), for large families of Gibbs models. The finite range assumption means that there exists $R > 0$ such that the particles do not interact at all if they are at a distance greater than $R > 0$ apart. For the two aforementioned inference methods, this assumption turns out to be crucial from both a practical point of view and a theoretical point of view, as explained below. However this assumption may imply an artificial discontinuity of the interactions between particles, where two particles at a distance $R - \epsilon$ apart interact while they do not at a distance $R + \epsilon$, for any small $\epsilon > 0$. This is for instance the case for the widely used Strauss model, see e.g. Møller and Waagepetersen (2004). In fact, this assumption rules out many interesting Gibbs models from statistical physics like the Lennard-Jones model. The purpose of this work is to extend the pseudolikelihood and logistic regression methods to infinite range interaction Gibbs models.

From a practical point of view, an important issue is edge effects. Assume we observe a Gibbs point process with finite range interaction $R > 0$ on a window $W \subset \mathbb{R}^d$. Then the pseudolikelihood computed on W actually depends on the point process on $W \oplus R$, where $W \oplus R$ denotes the dilation of W by a ball with radius R. Some approximation or some border correction is then needed. An obvious solution is to compute the pseudolikelihood on the eroded set $W \ominus R$, since $(W \ominus R) \oplus R \subseteq W$ (see Chiu et al. (2013)) and the observation of the point process on W is sufficient for the computation. From a theoretical point of view, this border correction preserves the unbiasedness property of the pseudolikelihood score function and standard technical tools for unbiased estimating equations are available to derive the asymptotic properties of the associated estimator. If the Gibbs point process has infinite range interaction, then the pseudolikelihood computed on W depends on the point process over the whole space $\mathbb{R}^d$. It is in general impossible to apply a border correction that preserves unbiasedness of the pseudolikelihood score function. Following the previous border correction for the finite range setting, we propose to alleviate this bias by computing the pseudolikelihood and the logistic regression on an eroded set. The details are exposed in Section 2. However these procedures still lead to biased score functions and the standard ingredients to derive consistency and asymptotic normality of the estimators do not apply.

The strong consistency of the maximum pseudolikelihood estimator was studied by Mase (1995) for pairwise interaction Gibbs point processes, including the infinite range interaction case, but under the assumption that the configuration of points outside W is known. Under the more realistic setting where the point process is observed only on W, we prove the strong consistency of the pseudolikelihood estimator in Proposition 4.1. The asymptotic normality is more challenging to establish. When the pseudolikelihood score function is unbiased, the main ingredient is a central limit theorem for conditionally centered random fields proved and generalized in Guyon and Künsch (1992); Jensen and Künsch (1994); Comets and Janžura

(1998); Dedecker (1998); Coeurjolly and Lavancier (2013). It allows in particular to avoid mixing assumptions for Gibbs point process that are only known in restrictive frameworks (see for instance Heinrich (1992); Jensen (1993)). In our infinite range setting where the score function is biased, a new ingredient is needed. We prove in Theorem 3.1 a new central limit theorem for triangular arrays of *almost* conditionally centered random fields. This allows us to derive in Theorem 4.2 the asymptotic normality of the pseudolikelihood estimator for a large family of pairwise Gibbs models, namely the class of Ruelle superstable and lower regular models. Proposition 4.3 discusses similar asymptotic results for the logistic regression estimator.

In Section 2 we recall some basic facts about Gibbs point processes and we explain how to generalize inference methods for Gibbs models with infinite range interaction. Section 3 contains our main theoretical tool, a new central limit theorem, and we derive in Section 4 the asymptotic properties of our estimators. Auxiliary lemmas are gathered in Section A.

2 Background and statistical methodology

2.1 Notation

We write $\Lambda \Subset \mathbb{R}^d$ for a bounded set Λ in $\mathbb{R}^d$. A configuration of points $\mathbf{x}$ is a locally finite subset of $\mathbb{R}^d$, which means that the set $\mathbf{x}_\Lambda := \mathbf{x} \cap \Lambda$ is finite for any set $\Lambda \Subset \mathbb{R}^d$. We use the notation $\mathbf{x}_{\Lambda^c} = \mathbf{x} \backslash \mathbf{x}_\Lambda$ and denote by Ω_0 the space of all locally finite configurations of points in $\mathbb{R}^d$. For a (p, q) matrix M with real entries, we denote by $\|M\| = \mathrm{tr}(M^\top M)^{1/2}$ its Frobenius norm where tr is the trace operator and $M^\top$ is the transpose matrix of M. For a vector $z \in \mathbb{R}^p$, $\|z\|$ reduces to its Euclidean norm. For a bounded set $E \subset \mathbb{Z}^d$, $|E|$ denotes the number of elements of E, while for $z \in \mathbb{R}^p$ or $i \in \mathbb{Z}^p$, $|z|$ and $|i|$ stand for the uniform norm.

At many places in the document, we use the notation c to denote a generic positive constant which may vary from line to line.

2.2 Pairwise interaction Gibbs point processes

We briefly recall the needed background material on point processes and we refer to Daley and Vere-Jones (2003) for more details. A point process is a probability measure on Ω_0. The reference distribution on Ω_0 is the homogeneous Poisson point process with intensity $\beta > 0$, denoted by π^β. For $\Lambda \Subset \mathbb{R}^d$, we write π^β_Λ for the restriction of π^β to Λ. For any $\Delta \Subset \mathbb{R}^d$ and $\mathbf{x} \in \Omega_0$, $N_\Delta(\mathbf{x})$ denotes the number of elements of $\mathbf{x} \cap \Delta$. Let Δ_i be the unit cube centered at $i \in \mathbb{Z}^d$. We consider the following space of tempered configurations.

$$\Omega_T = \{\mathbf{x} \in \Omega_0;\ \exists t > 0, \forall n \geqslant 1, \sum_{i \in \mathbb{Z}^d, |i| \leqslant n} N^2_{\Delta_i}(\mathbf{x}) \leqslant t(2n+1)^d\}.$$

From the ergodic theorem (see Guyon (1995)), any second order stationary measure on Ω_0 is supported on Ω_T. We denote by $\Phi : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ a pair potential function, to which we associate the pairwise energy function $H_\Lambda : \Omega_T \to \mathbb{R} \cup \{+\infty\}$,

indexed by Borel sets $\Lambda \Subset \mathbb{R}^d$ and defined by

$$H_\Lambda(\mathbf{x}) = \frac{1}{2} \sum_{\substack{u,v \in \mathbf{x},\, u \neq v,\\ \{u,v\} \cap \mathbf{x}_\Lambda \neq \varnothing}} \Phi(u - v) \tag{2.1}$$

and we let

$$\Omega = \{\mathbf{x} \in \Omega_T,\, \forall \Lambda \Subset \mathbb{R}^d \; H_\Lambda(\mathbf{x}) < \infty\}.$$

Following the Dobrushin-Lanford-Ruelle formalism, see Preston (1976), we say that P is a Gibbs measure with activity parameter $\beta > 0$ for the pair potential function Φ if $P(\Omega) = 1$ and for P-almost every configuration $\mathbf{x}$ and any $\Lambda \Subset \mathbb{R}^d$, the conditional law of P given $\mathbf{x}_{\Lambda^c}$ is absolutely continuous with respect to π_Λ^β with the density $\exp\{-H_\Lambda(\mathbf{x})\}/Z_\Lambda(\mathbf{x}_{\Lambda^c})$, where $Z_\Lambda(\mathbf{x}_{\Lambda^c}) = \int_\Omega \exp\{-H_\Lambda(\mathbf{x}_\Lambda \cup \mathbf{x}_{\Lambda^c})\}\, \pi_\Lambda^\beta(\, \mathrm{d}\mathbf{x}_{\Lambda^c})$ is the normalizing constant.

We use at many places in this paper the GNZ equation, after Georgii (1976) and Nguyen and Zessin (1979b), recalled below, which is a characterization of a Gibbs measure. It is given in terms of the Papangelou conditional intensity $\lambda : \mathbb{R}^d \times \Omega \to \mathbb{R}_+$ defined for any $\Lambda \ni u$ by

$$\lambda(u, \mathbf{x}) = \beta \, \frac{e^{-H_\Lambda(\mathbf{x} \cup u)}}{e^{-H_\Lambda(\mathbf{x})}} = \beta \, e^{-\sum_{v \in \mathbf{x}} \Phi(v-u)}. \tag{2.2}$$

This quantity does not depend on Λ, provided $u \in \Lambda$, and can be viewed as the conditional probability to have a point in a vicinity of u, given that the configuration elsewhere is $\mathbf{x}$.

Theorem 2.1 (GNZ formula)**.** *A probability measure P on Ω is a Gibbs measure with activity parameter $\beta > 0$ for the pair potential function Φ if for any measurable function $f : \Omega \times \mathbb{R}^d \to \mathbb{R}$ such that the following expectations are finite,*

$$\mathrm{E}\left\{ \sum_{u \in \mathbf{X}} f(u, \mathbf{X} \backslash u) \right\} = \mathrm{E}\left\{ \int f(u, \mathbf{X}) \lambda(u, \mathbf{X}) \, \mathrm{d}u \right\} \tag{2.3}$$

where E *denotes the expectation with respect to P.*

This result can be refined by a conditional version stated in the following lemma. Its proof is actually part of the initial proof of (2.3), see also Billiot et al. (2008, Proof of Theorem 2) for a particular case. We reproduce the demonstration below.

Lemma 2.2 (Conditional GNZ formula)**.** *Let P be a Gibbs measure with activity parameter $\beta > 0$, with pair potential Φ and Papangelou conditional intensity λ. Then for any measurable function $f : \Omega \times \mathbb{R}^d \to \mathbb{R}$ and for any $\Lambda \Subset \mathbb{R}^d$ such that the following expectations are finite*

$$\mathrm{E}\left\{ \sum_{u \in \mathbf{X}_\Lambda} f(u, \mathbf{X} \backslash u) \mid \mathbf{X}_{\Lambda^c} \right\} = \mathrm{E}\left\{ \int_\Lambda f(u, \mathbf{X}) \lambda(u, \mathbf{X}) \mid \mathbf{X}_{\Lambda^c} \right\} \tag{2.4}$$

where E *denotes the expectation with respect to P.*

Proof. By definition of the conditional law of P given $\mathbf{x}'_{\Lambda^c}$

$$\begin{aligned}\mathrm{E}\left\{\sum_{u\in\mathbf{X}_\Lambda} f(u,\mathbf{X}\backslash u)\mid \mathbf{X}_{\Lambda^c}=\mathbf{x}'_{\Lambda^c}\right\} &= \int_\Omega \sum_{u\in\mathbf{x}_\Lambda} f(u,\mathbf{x}_\Lambda\cup\mathbf{x}'_{\Lambda^c}\backslash u)\frac{e^{-H_\Lambda(\mathbf{x}_\Lambda\cup\mathbf{x}'_{\Lambda^c})}}{Z_\Lambda(\mathbf{x}'_{\Lambda^c})}\pi^\beta_\Lambda(\,\mathrm{d}\mathbf{x}_\Lambda)\\ &= \mathrm{E}_{\pi^\beta_\Lambda}\left\{\sum_{u\in\mathbf{X}_\Lambda} f(u,\mathbf{X}_\Lambda\cup\mathbf{x}'_{\Lambda^c}\backslash u)\frac{e^{-H_\Lambda(\mathbf{X}_\Lambda\cup\mathbf{x}'_{\Lambda^c})}}{Z_\Lambda(\mathbf{x}'_{\Lambda^c})}\right\}\end{aligned}$$

where $\mathrm{E}_{\pi^\beta_\Lambda}$ denotes the expectation with respect to π^β_Λ. From the Slivnyak-Mecke formula (Slivnyak (1962); Mecke (1968)), we know that for any admissible measurable function h

$$\mathrm{E}_{\pi^\beta_\Lambda}\left\{\sum_{u\in\mathbf{X}} h(u,\mathbf{X}\backslash u)\right\} = \beta\,\mathrm{E}_{\pi^\beta_\Lambda}\left\{\int_{\mathbb{R}^d} h(u,\mathbf{X})\,\mathrm{d}u\right\}.$$

By definition of the Papangelou conditional intensity (2.2), we also have for any $u\in\Lambda$, $\beta\, e^{-H_\Lambda(\mathbf{x}\cup u)} = e^{-H_\Lambda(\mathbf{x})}\lambda(u,\mathbf{x})$. Using these two facts, we conclude by

$$\begin{aligned}&\mathrm{E}\left\{\sum_{u\in\mathbf{X}_\Lambda} f(u,\mathbf{X}\backslash u)\mid \mathbf{X}_{\Lambda^c}=\mathbf{x}'_{\Lambda^c}\right\}\\ &\qquad= \beta\,\mathrm{E}_{\pi^\beta_\Lambda}\left\{\int_\Lambda f(u,\mathbf{X}_\Lambda\cup\mathbf{x}'_{\Lambda^c})\frac{e^{-H_\Lambda(\mathbf{X}_\Lambda\cup\mathbf{x}'_{\Lambda^c}\cup u)}}{Z_\Lambda(\mathbf{x}'_{\Lambda^c})}\,\mathrm{d}u\right\}\\ &\qquad= \mathrm{E}_{\pi^\beta_\Lambda}\left\{\int_\Lambda f(u,\mathbf{X}_\Lambda\cup\mathbf{x}'_{\Lambda^c})\lambda(u,\mathbf{X}_\Lambda\cup\mathbf{x}'_{\Lambda^c})\frac{e^{-H_\Lambda(\mathbf{X}_\Lambda\cup\mathbf{x}'_{\Lambda^c})}}{Z_\Lambda(\mathbf{x}'_{\Lambda^c})}\,\mathrm{d}u\right\}\\ &\qquad= \int_\Omega\int_\Lambda f(u,\mathbf{x}_\Lambda\cup\mathbf{x}'_{\Lambda^c})\lambda(u,\mathbf{x}_\Lambda\cup\mathbf{x}'_{\Lambda^c})\frac{e^{-H_\Lambda(\mathbf{x}_\Lambda\cup\mathbf{x}'_{\Lambda^c})}}{Z_\Lambda(\mathbf{x}'_{\Lambda^c})}\,\mathrm{d}u\,\pi^\beta_\Lambda(\,\mathrm{d}\mathbf{x}_\Lambda)\\ &\qquad= \mathrm{E}\left\{\int_\Lambda f(u,\mathbf{X})\lambda(u,\mathbf{X})\mid \mathbf{X}_{\Lambda^c}=\mathbf{x}'_{\Lambda^c}\right\}.\end{aligned}$$

□

The existence of a Gibbs measure P satisfying the above definition and characterization is a difficult question. Sufficient conditions on the pair potential Φ can be found in Ruelle (1969) and are also discussed in Preston (1976). The special case of finite range potentials, i.e. compactly supported functions Φ, is treated in Bertin et al. (1999). As we are mainly interested in this paper by infinite range potentials, we introduce the following assumption, that leads to the existence of at least one stationary Gibbs measure, as proved in Ruelle (1969).

[Φ] The potential Φ is bounded from below and there exist $0<r_1<r_2<\infty$, $c>0$ and $\gamma_1,\gamma_2>d$ such that $\Phi(u)\geqslant c\|u\|^{-\gamma_1}$ for $\|u\|\leqslant r_1$ and $|\Phi(u)|\leqslant c\|u\|^{-\gamma_2}$ for $\|u\|\geqslant r_2$.

Examples of potentials satisfying **[Φ]** are $\Phi(u)=\|u\|^{-\gamma}$ with $\gamma>d$ and $\Phi(u)=e^{-\|u\|}\|u\|^{-\gamma}$ with $\gamma>d$, in which cases the assumption is satisfied with $\gamma_1=\gamma_2=\gamma$. Another important example is the general Lennard-Jones pair potential defined for some $d<\gamma_2<\gamma_1$ and some $A,B>0$ by $\Phi(u)=A\|u\|^{-\gamma_1}-B\|u\|^{-\gamma_2}$. The standard

Lennard-Jones model corresponds to $d = 2$, $\gamma_1 = 12$ and $\gamma_2 = 6$. The main interest of this model is that it can model repulsion at small scales and attraction at large scales.

2.3 Inference for infinite range Gibbs point processes

In this section, we extend the usual statistical methodologies available for finite range Gibbs point processes to the infinite range case. We assume that the Gibbs measure is parametric, in that the explicit expression of the associated Papangelou conditional intensity (2.2) is entirely determined by the knowledge of some parameter $\theta \in \Theta$, including the activity parameter $\beta > 0$, where Θ is an open bounded set of $\mathbb{R}^p$. We stress this assumption by writing λ_θ instead of λ and Φ_θ instead of Φ. For brevity, assumption **[Φ]** now means that Φ_θ fulfills this assumption for any $\theta \in \Theta$.

Assume that we observe the point process $\mathbf{X}$ in W_n where $(W_n)_{n\geqslant 1}$ is a sequence of bounded domains which converges to $\mathbb{R}^d$ as $n \to \infty$. As outlined in the introduction, the pseudolikelihood and the logistic regression methods are popular alternatives to the maximum likelihood as they do not involve the normalizing constant. The associated estimators are respectively defined as the maximum of

$$\mathsf{LPL}_{W_n}(\mathbf{X};\theta) = \sum_{u\in\mathbf{X}_{W_n}} \log \lambda_\theta(u, \mathbf{X}\backslash u) - \int_{W_n} \lambda_\theta(u, \mathbf{X})\,\mathrm{d}u \tag{2.5}$$

$$\mathsf{LRL}_{W_n}(\mathbf{X};\theta) = \sum_{u\in\mathbf{X}_{W_n}} \log \frac{\lambda_\theta(u, \mathbf{X}\backslash u)}{\lambda_\theta(u, \mathbf{X}\backslash u) + \rho} - \int_{W_n} \rho \log \frac{\lambda_\theta(u, \mathbf{X}) + \rho}{\rho}\,\mathrm{d}u \tag{2.6}$$

where ρ is some fixed positive real number.

A problem however occurs. The integrals in (2.5) and (2.6) are not computable in practice because for values of u close to the boundary of W_n, $\lambda_\theta(u, \mathbf{X})$ depends on $\mathbf{X}_{W_n^c}$ which is not observed. When $\mathbf{X}$ has a finite range $0 < R < \infty$, meaning that Φ_θ is compactly supported on the euclidean ball $B(0, R)$ or equivalently that for any $u \in \mathbb{R}^d$ and any $\mathbf{x} \in \Omega$, $\lambda_\theta(u, \mathbf{x}) = \lambda_\theta(u, \mathbf{x}_{B(u,R)})$, we can simply substitute W_n by $W_n \ominus R$ in (2.5) and (2.6), where for a bounded domain $\Lambda \subset \mathbb{R}^d$ and some $\kappa \geqslant 0$ the notation $\Lambda \ominus \kappa$ stands for the domain Λ eroded by the ball $B(0, \kappa)$. Using this border correction $\lambda_\theta(u, \mathbf{X})$ can be indeed computed for any $u \in W_n \ominus R$. As a remaining practical issue, the integrals have to be approximated by some numerical scheme or by Monte-Carlo, see Baddeley et al. (2014) for an efficient solution.

The asymptotic properties of the pseudolikelihood and the logistic regression estimators are well understood in this finite range setting, see the references in introduction. Maximizing the log-pseudolikelihood (or the logistic regression likelihood) on $W_n \ominus R$ is equivalent to cancel the score, i.e. the gradient of $\mathsf{LPL}_{W_n\ominus R}(\mathbf{X};\theta)$ (or $\mathsf{LRL}_{W_n\ominus R}(\mathbf{X};\theta)$) with respect to θ. The key-ingredient is that both scores constitute unbiased estimating functions, since by application of the GNZ formula (2.3) their expectation vanishes when θ corresponds to the true parameter of the underlying Gibbs measure. Standard theoretical tools for unbiased estimating equations (see e.g. Guyon (1995)) can therefore be used to study the consistency and asymptotic normality of the associated estimators.

In the infinite range setting, the situation becomes more delicate since for any u, $\lambda_\theta(u, \mathbf{X})$ depends on $\mathbf{X}_\Lambda$ for any $\Lambda \subset \mathbb{R}^d$. In this case, we introduce the following modifications of (2.5) and (2.6)

$$\widetilde{\mathsf{LPL}}_{W_n \ominus \alpha_n}(\mathbf{X}; \theta) = \sum_{u \in \mathbf{X}_{W_n \ominus \alpha_n}} \log \lambda_\theta(u, \mathbf{X}_{B_{u,n}} \backslash u) - \int_{W_n \ominus \alpha_n} \lambda_\theta(u, \mathbf{X}_{B_{u,n}}) \, \mathrm{d}u \tag{2.7}$$

$$\widetilde{\mathsf{LRL}}_{W_n \ominus \alpha_n}(\mathbf{X}; \theta) = \sum_{u \in \mathbf{X}_{W_n \ominus \alpha_n}} \log \frac{\lambda_\theta(u, \mathbf{X}_{B_{u,n}} \backslash u)}{\lambda_\theta(u, \mathbf{X}_{B_{u,n}} \backslash u) + \rho} - \int_{W_n \ominus \alpha_n} \rho \log \frac{\lambda_\theta(u, \mathbf{X}_{B_{u,n}}) + \rho}{\rho} \, \mathrm{d}u \tag{2.8}$$

where $B_{u,n} = B(u, \alpha_n)$ and $(\alpha_n)_{n \geqslant 1}$ is a sequence of positive numbers, which agree with the classical border correction for finite range interaction models when $\alpha_n = R$. Whatever the range of interaction is, each term above can be computed in practice from the single observation of $\mathbf{X}$ in W_n, provided the integrals are approximated as usual by numerical scheme or by Monte-Carlo. From a theoretical point view, these modifications introduce new challenges since the gradients of $\widetilde{\mathsf{LPL}}_{W_n \ominus \alpha_n}(\mathbf{X}; \theta)$ and $\widetilde{\mathsf{LRL}}_{W_n \ominus \alpha_n}(\mathbf{X}; \theta)$ are no longer unbiased estimating equations in the infinite range case. To overcome this difficulty we prove a new central limit theorem in the next section that allows us to deduce in Section 4 the consistency and asymptotic normality of the estimators.

3 A new central limit theorem

When the Gibbs point process has a finite range, the asymptotic normality of the pseudolikelihood or the logistic regression estimators are essentially derived from a central limit theorem for conditionally centered random fields, see the references in introduction. This connection comes from the fact that in the finite range case, the score function of the pseudolikelihood (or the logistic regression) is not only centered, as noticed in the previous section, but also conditionally centered, by application of the conditional GNZ formula (2.4). As already mentioned, this property allows to prove the asymptotic normality without mixing assumptions, which is crucial for Gibbs point processes.

In the infinite range case, the score functions of the log-pseudolikelihood and the logistic regression are neither centered, nor conditionally centered, and a new central limit theorem is needed. In our Theorem 3.1 below, the conditional centering condition is replaced by condition (d) and we avoid mixing assumptions. The other conditions are mainly due to our non-stationary setting induced by the border correction with the sequence α_n. They allow in particular to control the asymptotic behavior of the empirical covariance matrix in (3.1).

For two square matrices A, B we write $A \geqslant B$ when $A - B$ is a positive semi-definite matrix.

Theorem 3.1. *For $n \in \mathbb{N}$ and $j \in \mathbb{Z}^d$, let $X_{n,j}$ be a triangular array field in a measurable space S. For $n \in \mathbb{N}$, let $\mathcal{I}_n \subset \mathbb{Z}^d$ and $\alpha_n \in \mathbb{R}_+$ such that $|\mathcal{I}_n| \to \infty$ and $\alpha_n \to \infty$ as $n \to \infty$. Define $S_n = \sum_{j \in \mathcal{I}_n} Z_{n,j}$ where $Z_{n,j} = f_{n,j}(X_{n,k}, k \in \mathcal{K}_{n,j})$ with*

$\mathcal{K}_{n,j} = \{k \in \mathbb{Z}^d, |k-j| \leqslant \alpha_n\}$ and where $f_{n,j} : S^{\mathcal{K}_{n,j}} \to \mathbb{R}^p$ is a measurable function. If $\alpha_n^{3d} = o(|\mathcal{I}_n|)$ as $n \to \infty$ and

(a) $\mathrm{E}\, Z_{n,j} = 0$ *and* $\sup_{n \geqslant 1} \sup_{j \in \mathcal{I}_n} \mathrm{E}\, \|Z_{n,j}\|^4 < \infty$,

(b) for any sequences $(\mathcal{I}_n)$ and (α_n) as above, as $n \to \infty$,

$$|\mathcal{I}_n|^{-1} \sum_{j,k \in \mathcal{I}_n} \big\| \mathrm{E}(Z_{n,j} Z_{n,k}^\top) \big\| = \mathcal{O}(1) \quad \text{and} \quad |\mathcal{I}_n|^{-1} \sum_{\substack{j,k \in \mathcal{I}_n \\ |k-j| > \alpha_n}} \big\| \mathrm{E}(Z_{n,j} Z_{n,k}^\top) \big\| = o(1),$$

then denoting $\Sigma_n = \mathrm{Var}(S_n)$ and

$$\widehat{\Sigma}_n = \sum_{j \in \mathcal{I}_n} \sum_{\substack{k \in \mathcal{I}_n \\ |k-j| \leqslant \alpha_n}} Z_{n,j} Z_{n,k}^\top,$$

we have the following convergence

$$|\mathcal{I}_n|^{-1}\, \mathrm{E}\, \|\widehat{\Sigma}_n - \Sigma_n\| \to 0. \tag{3.1}$$

If in addition

(c) there exists a positive definite matrix Q such that $|\mathcal{I}_n|^{-1}\Sigma_n \geqslant Q$ for n sufficiently large,

(d) as $n \to \infty$

$$|\mathcal{I}_n|^{-1/2} \sum_{j \in \mathcal{I}_n} \mathrm{E}\, \|\mathrm{E}\,(Z_{n,j} |\, X_{n,k}, k \neq j)\| \to 0,$$

then

$$\Sigma_n^{-1/2} S_n \xrightarrow{d} \mathcal{N}(0, I_p) \tag{3.2}$$

where $\xrightarrow{d}$ stands for the convergence in distribution.

Proof. We have

$$|\mathcal{I}_n|^{-1} \|\widehat{\Sigma}_n - \Sigma_n\| \leqslant |\mathcal{I}_n|^{-1} \|\widehat{\Sigma}_n - \mathrm{E}(\widehat{\Sigma}_n)\| + |\mathcal{I}_n|^{-1} \|\mathrm{E}(\widehat{\Sigma}_n) - \Sigma_n\|.$$

The last term satisfies

$$|\mathcal{I}_n|^{-1} \|\mathrm{E}(\widehat{\Sigma}_n) - \Sigma_n\| \leqslant |\mathcal{I}_n|^{-1} \sum_{\substack{j,k \in \mathcal{I}_n \\ |k-j| > \alpha_n}} \big\| \mathrm{E}(Z_{n,j} Z_{n,k}^\top) \big\| = o(1)$$

by assumption (b). For $m, m' = 1, \ldots, p$, let $\Delta_{mm'} = \Big(\widehat{\Sigma}_n - \mathrm{E}(\widehat{\Sigma}_n)\Big)_{mm'}$. We have $\Delta_{mm'} = \sum_{j \in \mathcal{I}_n} U_{n,j}$ where

$$U_{n,j} = \sum_{k \in \mathcal{I}_{n,j}} \big\{ Z_{n,j} Z_{n,k}^\top - \mathrm{E}(Z_{n,j} Z_{n,k}^\top) \big\}_{mm'}$$

and $\mathcal{I}_{n,j} = \{k \in \mathcal{I}_n : |k-j| \leqslant \alpha_n\}$. The assertion (3.1) is proved if we show that for any m, m', $\mathrm{Var}(|\mathcal{I}_n|^{-1}\Delta_{mm'}) \to 0$ as $n \to \infty$. It is clear that $U_{n,j}$ depends only on $X_{n,k}$ for $|k-j| \leqslant 2\alpha_n$. So if we let $j, j' \in \mathcal{I}_n$ such that $|j-j'| > 4\alpha_n$ then

$$\begin{aligned}\mathrm{E}(U_{n,j}U_{n,j'}) &= \mathrm{E}\left\{\mathrm{E}\left(U_{n,j}U_{n,j'} \mid \ X_{n,k}, |k-j| > 2\alpha_n\right)\right\} \\ &= \mathrm{E}\left\{U_{n,j'}\,\mathrm{E}\left(U_{n,j} \mid \ X_{n,k}, |k-j| > 2\alpha_n\right)\right\} \\ &= \mathrm{E}\left\{U_{n,j'}\,\mathrm{E}\left(U_{n,j}\right)\right\} = 0\end{aligned}$$

whereby we deduce that

$$\mathrm{Var}(\Delta_{mm'}) = \sum_{\substack{j,j' \in \mathcal{I}_n \\ |j-j'| \leqslant 4\alpha_n}} \mathrm{E}(U_{n,j}U_{n,j'}).$$

Now, by condition (a) and Hölder's inequality

$$\begin{aligned}\mathrm{E}\,U_{n,j}^2 &= \sum_{k,k' \in \mathcal{I}_{n,j}} \mathrm{E}\left[\left\{Z_{n,j}Z_{n,k}^\top - \mathrm{E}(Z_{n,j}Z_{n,k}^\top)\right\}_{mm'}\left\{Z_{n,j}Z_{n,k'}^\top - \mathrm{E}(Z_{n,j}Z_{n,k'}^\top)\right\}_{mm'}\right] \\ &\leqslant c\sum_{k,k' \in \mathcal{I}_{n,j}} \mu_4 \\ &\leqslant c\left(\sup_{j \in \mathcal{I}_n} \#\left\{k \in \mathcal{I}_n, |k-j| \leqslant \alpha_n\right\}\right)^2 = \mathcal{O}(\alpha_n^{2d}).\end{aligned}$$

From Cauchy-Schwarz's inequality, we continue with

$$\mathrm{Var}(\Delta_{mm'}) \leqslant \sum_{\substack{j,j' \in \mathcal{I}_n \\ |j-j'| \leqslant 4\alpha_n}} \mathrm{E}(U_{n,j}^2)^{1/2}\,\mathrm{E}(U_{n,j'}^2)^{1/2} = \mathcal{O}\left(\alpha_n^{3d}|\mathcal{I}_n|\right)$$

leading to

$$\mathrm{Var}(|\mathcal{I}_n|^{-1}\Delta_{mm'}) = \mathcal{O}\left(\frac{\alpha_n^{3d}}{|\mathcal{I}_n|}\right) = o(1)$$

which completes the proof of (3.1).

We now focus on (3.2) and we let

$$\overline{S}_n = \Sigma_n^{-1/2}S_n, \quad S_{n,j} = \sum_{k \in \mathcal{I}_{n,j}} Z_{n,k} \quad \text{and} \quad \overline{S}_{n,j} = \Sigma_n^{-1/2}S_{n,j}$$

where we recall the notation $\mathcal{I}_{n,j} = \mathcal{K}_{n,j} \cap \mathcal{I}_n$. According to Stein's method (see Bolthausen, 1982), in order to show (3.2) it suffices to prove that for all $u \in \mathbb{R}^p$ such that $\|u\| = 1$ and for all $\omega \in \mathbb{R}$

$$\mathrm{E}\left\{(\mathbf{i}\omega - u^\top\overline{S}_n)e^{\mathbf{i}\omega u^\top\overline{S}_n}\right\} \to 0$$

as $n \to \infty$ where $\mathbf{i} = \sqrt{-1}$. Letting $v = \omega u$, this is equivalent to show that for all $v \in \mathbb{R}^p$, $\|\,\mathrm{E}(A_n)\| \to 0$ where $A_n = (\mathbf{i}v - \overline{S}_n)e^{iv^\top\overline{S}_n}$. We decompose the term A_n in

the same spirit as Bolthausen (1982) : $A_n = A_{n,1} - A_{n,2} - A_{n,3}$ where

$$\begin{aligned}
A_{n,1} &= \mathbf{i} e^{\mathbf{i} v^\top \overline{S}_n} (I_p - \Sigma_n^{-1/2} \sum_{j \in \mathcal{I}_n} Z_{n,j} S_{n,j}^\top \Sigma_n^{-1/2}) v \\
&= \mathbf{i} e^{\mathbf{i} v^\top \overline{S}_n} (I_p - \Sigma_n^{-1/2} \widehat{\Sigma}_n \Sigma_n^{-1/2}) v \\
A_{n,2} &= e^{\mathbf{i} v^\top \overline{S}_n} \Sigma_n^{-1/2} \sum_{j \in \mathcal{I}_n} Z_{n,j} (1 - \mathbf{i} v^\top \overline{S}_{n,j} - e^{-\mathbf{i} v^\top \overline{S}_{n,j}}) \\
A_{n,3} &= \Sigma_n^{-1/2} \sum_{j \in \mathcal{I}_n} Z_{n,j} e^{\mathbf{i} v^\top (\overline{S}_n - \overline{S}_{n,j})}
\end{aligned}$$

and prove in the following that $\| \operatorname{E} A_{n,r} \| \to 0$ for $r = 1, 2, 3$ as $n \to \infty$.

First, assumption (c) implies that $|\mathcal{I}_n|^{-1} \Sigma_n$ is a positive definite matrix for n sufficiently large, which is now assumed in the following. By ℓ we denote the constant $p / \lambda_{\min}(Q)$ where $\lambda_{\min}(M)$ stands for the smallest eigenvalue of a positive definite squared matrix M. For n sufficiently large, $\lambda_{\min}(|\mathcal{I}_n|^{-1} \Sigma_n) \geqslant \lambda_{\min}(Q) > 0$ whereby we deduce

$$\|\Sigma_n^{-1/2}\| = |\mathcal{I}_n|^{-1/2} \operatorname{tr}(|\mathcal{I}_n| \Sigma_n^{-1})^{1/2} \leqslant \ell^{1/2} |\mathcal{I}_n|^{-1/2}. \tag{3.3}$$

Using this result and the sub-multiplicative property of the Frobenius norm, we get

$$\begin{aligned}
\|\operatorname{E} A_{n,1}\| &\leqslant \|v\| \operatorname{E} \|I_p - \Sigma_n^{-1/2} \widehat{\Sigma}_n \Sigma_n^{-1/2}\| \\
&\leqslant \|v\| \, \|\Sigma_n^{-1/2}\|^2 \operatorname{E} \|\widehat{\Sigma}_n - \Sigma_n\| \\
&\leqslant \ell \|v\| \, |\mathcal{I}_n|^{-1} \operatorname{E} \|\widehat{\Sigma}_n - \Sigma_n\|
\end{aligned}$$

whereby we deduce that $\| \operatorname{E} A_{n,1} \| \to 0$ from (3.1).

Second, since $|1 - e^{-\mathbf{i}z} - \mathbf{i}z| \leqslant z^2/2$ for any $z \in \mathbb{R}$, we have

$$\begin{aligned}
\|A_{n,2}\| &\leqslant \frac{1}{2} \|\Sigma_n^{-1/2}\| \sum_{j \in \mathcal{I}_n} \|Z_{n,j}\| (v^\top \overline{S}_{n,j})^2 \\
&\leqslant \frac{\|v\|^2}{2} \|\Sigma_n^{-1/2}\|^3 \sum_{j \in \mathcal{I}_n} B_{n,j} \leqslant \frac{\ell^{3/2} \|v\|^2}{2} |\mathcal{I}_n|^{-3/2} \sum_{j \in \mathcal{I}_n} B_{n,j}
\end{aligned}$$

where

$$B_{n,j} \;=\; \|Z_{n,j}\| \, \|\sum_{k \in \mathcal{I}_{n,j}} Z_{n,k}\|^2 \;=\; \|Z_{n,j}\| \sum_{k,k' \in \mathcal{I}_{n,j}} Z_{n,k}^\top Z_{n,k'} \;\geqslant 0.$$

Let us decompose $B_{n,j} = B_{n,j}^{(1)} + B_{n,j}^{(2)}$ where $B_{n,j}^{(1)} = B_{n,j} \mathbf{1}(\|Z_{n,j}\| \leqslant |\mathcal{I}_n|^{1/6})$ and $B_{n,j}^{(2)} = B_{n,j} \mathbf{1}(\|Z_{n,j}\| > |\mathcal{I}_n|^{1/6})$. By assumption (b), we have

$$\| \operatorname{E} B_{n,j}^{(1)} \| \leqslant |\mathcal{I}_n|^{1/6} \sum_{k,k' \in \mathcal{I}_{n,j}} |\operatorname{E}(Z_{n,k}^\top Z_{n,k'})| = \mathcal{O}(|\mathcal{I}_n|^{1/6} |\mathcal{I}_{n,j}|) = \mathcal{O}(|\mathcal{I}_n|^{1/6} \alpha_n^d). \tag{3.4}$$

By assumption (a), using Hölder and Bienaymé-Chebyshev inequalities, we continue

with

$$\begin{aligned}
\mathrm{E}\, B_{n,j}^{(2)} &\leqslant \sum_{k,k' \in \mathcal{I}_{n,j}} \mathrm{E}\left\{ \|Z_{n,j}\| \, \|Z_{n,k}\| \, \|Z_{n,k'}\| \, \mathbf{1}(\|Z_{n,j} > |\mathcal{I}_n|^{1/6}\|) \right\} \\
&\leqslant \sum_{k,k' \in \mathcal{I}_{n,j}} \mu_4^{3/4} \, \mathrm{P}(\|Z_{n,j}\| > |\mathcal{I}_n|^{1/6})^{1/4} \\
&\leqslant \mu_4^{3/4} \sum_{k,k' \in \mathcal{I}_{n,j}} \mu_4^{1/4} |\mathcal{I}_n|^{-1/6} \\
&\leqslant c\mu_4 |\mathcal{I}_n|^{-1/6} \alpha_n^{2d}. \qquad (3.5)
\end{aligned}$$

Combining (3.4)-(3.5), we deduce that as $n \to \infty$

$$\|EA_{n,2}\| \leqslant \mathrm{E}\, \|A_{n,2}\| = \mathcal{O}\left(\frac{\alpha_n^d}{|\mathcal{I}_n|^{1/2}} |\mathcal{I}_n|^{1/6} \right) \, + \, \mathcal{O}\left(\frac{\alpha_n^{2d}}{|\mathcal{I}_n|^{1/2}} \, |\mathcal{I}_n|^{-1/6} \right) = o(1)$$

since $\alpha_n^{3d} = o(|\mathcal{I}_n|)$.

Third, for any $j \in \mathcal{I}_n$, $\overline{S}_n - \overline{S}_{n,j}$ does not depend on $X_{n,j}$. This yields

$$\mathrm{E}\, A_{n,3} = \Sigma_n^{-1/2} \sum_{j \in \mathcal{I}_n} \mathrm{E}\left\{ e^{\mathbf{i} v^\top (\overline{S}_n - \overline{S}_{n,j})} \, \mathrm{E}(Z_{n,j} \mid X_{n,k}, k \neq j) \right\}$$

whereby we deduce, in view of (3.3), that

$$\|\, \mathrm{E}\, A_{n,3}\| \leqslant c|\mathcal{I}_n|^{-1/2} \sum_{j \in \mathcal{I}_n} \mathrm{E}\, \|\mathrm{E}\, (Z_{n,j} \mid X_{n,k}, k \neq j)\|$$

which tends to 0 by assumption (d). □

4 Applications to infinite range pairwise interaction Gibbs point processes

In this section, we present asymptotic properties of the maximum pseudolikelihood estimate, derived from (2.7), for infinite range Gibbs point process. Similar results for the maximum logistic regression derived from (2.8) are presented at the end of this section without proof. We focus on exponential family models of pairwise interaction Gibbs point processes and rewrite the model (2.2) for any $u \in \mathbb{R}^d$ and $\mathbf{x} \in \Omega$ as

$$\lambda_\theta(u, \mathbf{x}) = \beta e^{-\sum_{v \in \mathbf{x}} \Phi_\theta(v-u)} = e^{-\theta^\top t(u,\mathbf{x})} \qquad (4.1)$$

with $\theta_1 = -\log \beta$ and $t = (t_1, \ldots, t_p)^\top$ where $t_1(u, \mathbf{x}) = 1$ and

$$t_m(u, \mathbf{x}) = \sum_{v \in \mathbf{x}} g_m(v - u), \quad m = 2, \ldots, p. \qquad (4.2)$$

In that connection, our framework amounts to assume that $\Phi = \sum_{m=2}^p \theta_m g_m$. For convenience we let $g_1 = 0$ and we denote by g the p-dimensional vector $g = (0, g_2, \ldots, g_p)^\top$. We make the following assumption on g.

[g] For all $m \geqslant 2$, g_m is bounded from below and there exist $\gamma_1, \gamma_2 > d$ and $c_g, r_0 > 0$ such that

(i) $\forall \|x\| < r_0$ and $\forall \theta \in \Theta$, $\theta_2\, g_2(x) \geqslant c_g \|x\|^{-\gamma_1}$

(ii) $\forall m \geqslant 3$, $g_m(x) = o(\|x\|^{-\gamma_1})$ as $\|x\| \to 0$

(iii) $\forall m \geqslant 2$ and $\forall \|x\| \geqslant r_0$, $|g_m(x)| \leqslant c\|x\|^{-\gamma_2}$.

Since Θ is bounded, **[g]** implies **[Φ]** which yields that for any $\theta \in \Theta$ there exists a Gibbs measure P_θ. Assumption **[g]** allows us to specify which function g_m is responsible for the behavior at the origin of Φ_θ, namely g_2. Note that the Lennard-Jones model defined in Section 2.2 (and the other examples presented in this section) fits this setting with $\theta_2 = A$, $\theta_3 = -B$, $g_2(u) = \|u\|^{-\gamma_1}$ and $g_3(u) = \|u\|^{-\gamma_2}$. In the sequel, $\theta^\star$ stands for the true parameter vector to estimate. In other words, we assume observing a realization of a spatial point process $\mathbf{X}$ with Gibbs measure $P_{\theta^\star}$ on W_n.

For exponential family models (4.1) the score function of the log-pseudolikelihood defined by (2.7) writes $s_{W_n \ominus \alpha_n}(\mathbf{X}; \theta)$ where for any $\Delta \subset \mathbb{R}^d$

$$s_\Delta(\mathbf{X}; \theta) = \int_\Delta t(u, \mathbf{X}_{B_{u,n}}) \lambda_\theta(u, \mathbf{X}_{B_{u,n}})\, \mathrm{d}u - \sum_{u \in \mathbf{X}_\Delta} t(u, \mathbf{X}_{B_{u,n}} \backslash u). \tag{4.3}$$

Our first result establishes the strong consistency of the maximum pseudolikelihood based on (2.7) for infinite range Gibbs point processes. In close relation, Mase (1995) proved the strong consistency of estimators derived from (2.5). As pointed out in Section 2.3, the form (2.5) of log-pseudolikelihood is however unusable as it can only be computed if $\mathbf{X}$ is observed on $\mathbb{R}^d$. We obtain the same result but for estimators derived from the computable pseudolikelihood given by (2.7).

Proposition 4.1. *Assume that for any $n \in \mathbb{N}$, W_n is a convex, compact set and α_n a sequence of real numbers satisfying $W_n \to \mathbb{R}^d$, $\alpha_n \to \infty$ and $\alpha_n = o(|W_n|)$ as $n \to \infty$. Then, the function $\theta \to -\widetilde{\mathsf{LPL}}_{W_n \ominus \alpha_n}(\mathbf{x}; \theta)$ is a convex function for any $\mathbf{x} \in \Omega$ with Hessian matrix given by*

$$\begin{aligned} -\frac{\mathrm{d}}{\mathrm{d}\theta\, \mathrm{d}\theta^\top} \widetilde{\mathsf{LPL}}_{W_n \ominus \alpha_n}(\mathbf{X}; \theta) &= -\frac{\mathrm{d}}{\mathrm{d}\theta^\top} s_{W_n \ominus \alpha_n}(\mathbf{X}; \theta) \\ &= \int_{W_n \ominus \alpha_n} t(u, \mathbf{X}_{B_{u,n}}) t(u, \mathbf{X}_{B_{u,n}})^\top \lambda_\theta(u, \mathbf{X}_{B_{u,n}})\, \mathrm{d}u. \end{aligned} \tag{4.4}$$

In addition, assume that **[g]** *holds and that for any $y \in \mathbb{R}^p \backslash \{0\}$*

$$P\left\{ y^\top t(0, \mathbf{X}) \neq 0 \right\} > 0 \tag{4.5}$$

then the maximum pseudolikelihood estimator

$$\widehat{\theta}_{\widetilde{\mathsf{LPL}}} = \mathrm{argmax}_{\theta \in \Theta} \widetilde{\mathsf{LPL}}_{W_n \ominus \alpha_n}(\mathbf{X}; \theta)$$

converges almost surely to $\theta^\star$ as $n \to \infty$.

Proof. The basic assumption on W_n and α_n ensures that $W_n \ominus \alpha_n$ is a sequence of regular bounded domains of $\mathbb{R}^d$ and that $|W_n \ominus \alpha_n| \to \infty$ as $n \to \infty$. Since any stationary Gibbs measure can be represented as a mixture of ergodic measures (Preston, 1976), it is sufficient to prove consistency for ergodic measures. So, we

assume here that $P_{\theta^\star}$ is ergodic. Since Θ is an open bounded set, and by convexity of $\theta \to -\widetilde{\mathsf{LPL}}_{W_n\ominus\alpha_n}(\mathbf{x};\theta)$, then from Guyon (1995, Theorem 3.4.4) we only need to prove that $K_n(\theta,\theta^\star) = |W_n \ominus \alpha_n|^{-1}\left\{\widetilde{\mathsf{LPL}}_{W_n\ominus\alpha_n}(\mathbf{X};\theta^\star) - \widetilde{\mathsf{LPL}}_{W_n\ominus\alpha_n}(\mathbf{X};\theta)\right\} \to K(\theta,\theta^\star)$ almost surely as $n\to\infty$, where $\theta \to K(\theta,\theta^\star)$ is a nonnegative function which vanishes at $\theta=\theta^\star$ only. We decompose $K_n(\theta,\theta^\star)$ as the sum of the three terms $T_1 + T_2(\theta^\star) - T_2(\theta)$ where for any $\theta\in\Theta$

$$T_1 = |W_n \ominus \alpha_n|^{-1}\left\{\mathsf{LPL}_{W_n\ominus\alpha_n}(\mathbf{X};\theta^\star) - \mathsf{LPL}_{W_n\ominus\alpha_n}(\mathbf{X};\theta)\right\}$$
$$T_2(\theta) = |W_n \ominus \alpha_n|^{-1}\left\{\widetilde{\mathsf{LPL}}_{W_n\ominus\alpha_n}(\mathbf{X};\theta) - \mathsf{LPL}_{W_n\ominus\alpha_n}(\mathbf{X};\theta)\right\}.$$

Lemma A.2 shows in particular that $\lambda_\theta(0,\mathbf{X})$ and $|\theta^\top t(0,\mathbf{X})|\lambda_{\theta^\star}(0,\mathbf{X})$ have finite expectation under $P_{\theta^\star}$. Hence, using the ergodic theorem for spatial processes of Nguyen and Zessin (1979a), we can follow the proof of Mase (1995) or the proof of Billiot et al. (2008, Theorem 1) to prove that $T_1 \to K(\theta,\theta^\star)$ almost surely as $n\to\infty$ where

$$K(\theta,\theta^\star) = \mathrm{E}\left(\lambda_{\theta^\star}(0,\mathbf{X})\left[e^{(\theta^\star-\theta)^\top t(0,\mathbf{X})} - \{1+(\theta^\star-\theta)^\top t(0,\mathbf{X})\}\right]\right)$$

which is a nonnegative function that vanishes at $\theta=\theta^\star$ only, under the identifiability condition (4.5). So the rest of the proof consists in proving that $T_2(\theta)\to 0$ almost surely for any $\theta\in\Theta$. We have $T_2(\theta) = T_1' + T_2'$ where

$$T_1' = |W_n \ominus \alpha_n|^{-1}\sum_{u\in\mathbf{X}_{W_n\ominus\alpha_n}} \theta^\top\left\{t(u,\mathbf{X}_{B_{u,n}}\backslash u) - t(u,\mathbf{X}\backslash u)\right\}$$
$$T_2' = |W_n \ominus \alpha_n|^{-1}\int_{W_n\ominus\alpha_n}\left\{\lambda_\theta(u,\mathbf{X}) - \lambda_\theta(u,\mathbf{X}_{B_{u,n}})\right\}\mathrm{d}u.$$

By Lemma A.1, the boundedness of Θ and **[g]**, there exists $\gamma'>0$ such that

$$|T_1'| \leqslant \frac{c}{\alpha_n^{\gamma'}}|W_n \ominus \alpha_n|^{-1}\sum_{u\in\mathbf{X}_{W_n\ominus\alpha_n}} H(u,\mathbf{X}\backslash u)$$
$$|T_2'| \leqslant \frac{c}{\alpha_n^{\gamma'}}|W_n \ominus \alpha_n|^{-1}\int_{W_n\ominus\alpha_n} e^{cG(u,\mathbf{X})}H(u,\mathbf{X})\lambda_\theta(u,\mathbf{X})\,\mathrm{d}u.$$

By Lemma A.2, the random variables $|H(0,\mathbf{X})|\lambda_{\theta^\star}(0,\mathbf{X})$ and $e^{cG(0,\mathbf{X})}H(0,\mathbf{X})\lambda_\theta(0,\mathbf{X})$ have finite expectations under $P_{\theta^\star}$. Hence, using again the ergodic theorem

$$|W_n \ominus \alpha_n|^{-1}\sum_{u\in\mathbf{X}_{W_n\ominus\alpha_n}} H(u,\mathbf{X}\backslash u) \to \mathrm{E}\left\{H(0,\mathbf{X})\lambda_{\theta^\star}(0,\mathbf{X})\right\}$$
$$|W_n \ominus \alpha_n|^{-1}\int_{W_n\ominus\alpha_n} e^{cG(u,\mathbf{X})}H(u,\mathbf{X})\lambda_\theta(u,\mathbf{X})\,\mathrm{d}u \to \mathrm{E}\left\{e^{cG(0,\mathbf{X})}H(0,\mathbf{X})\lambda_\theta(0,\mathbf{X})\right\}$$

almost surely as $n\to\infty$, whereby we deduce that $T_2(\theta)\to 0$ almost surely. □

The next result establishes the asymptotic normality of the score function associated to the modified pseudolikelihood $\widetilde{\mathsf{LPL}}_{W_n\ominus\alpha_n}(\mathbf{X};\theta)$ at the true value of the

parameter $\theta = \theta^\star$. The proof relies on the central limit theorem of Theorem 3.1. As a consequence we deduce the asymptotic normality of the associated estimator.

These results require the following notation: let Σ_n, Σ_∞ and U_∞ the (p,p) matrices

$$\begin{aligned}
\Sigma_n &= \mathrm{Var}\{s_{W_n \ominus \alpha_n}(\mathbf{X};\theta^\star)\} \\
\Sigma_\infty &= \mathrm{E}\left\{t(0,\mathbf{X})t(0,\mathbf{X})^\top \lambda_{\theta^\star}(0,\mathbf{X})\right\} \\
&\quad + \int_{\mathbb{R}^d} \mathrm{E}\left\{t(0,\mathbf{X})t(v,\mathbf{X})^\top \lambda_{\theta^\star}(0,\mathbf{X})\lambda_{\theta^\star}(v,\mathbf{X})\right\}\{1-e^{-\Phi_{\theta^\star}(v)}\}\,\mathrm{d}v \\
&\quad + \int_{\mathbb{R}^d} \mathrm{E}\left\{\lambda_{\theta^\star}(0,\mathbf{X})\lambda_{\theta^\star}(v,\mathbf{X})\right\} g(v)g(v)^\top e^{-\Phi_{\theta^\star}(v)}\,\mathrm{d}v \qquad (4.6) \\
U_\infty &= \mathrm{E}\left\{t(0,\mathbf{X})t(0,\mathbf{X})^\top \lambda_{\theta^\star}(0,\mathbf{X})\right\}. \qquad (4.7)
\end{aligned}$$

The matrices Σ_∞ and U_∞ are indeed correctly defined, as **[g]** implies on the one hand that all the expectations involved are uniformly bounded in v by Lemmas A.1-A.2, and on the other hand that

$$\int_{\mathbb{R}^d} |1-e^{-\Phi_{\theta^\star}(v)}|\,\mathrm{d}v < \infty \quad \text{and} \quad \int_{\mathbb{R}^d} \|g(v)g(v)^\top\| e^{-\Phi_{\theta^\star}(v)}\,\mathrm{d}v < \infty.$$

Theorem 4.2. *Under the assumptions of Proposition 4.1 with $\gamma_2 > 5d/2$, the assumption that Σ_∞ is a positive definite matrix and if $\alpha_n = c_a|W_n|^{a/d}$ where $c_a > 0$ and a is such that $\frac{d}{2(\gamma_2-d)} < a < \frac{1}{3}$, then we have the two following convergences in distribution as $n \to \infty$*

(i)

$$\Sigma_n^{-1/2} s_{W_n \ominus \alpha_n}(\mathbf{X};\theta^\star) \xrightarrow{d} \mathcal{N}(0, I_p),$$

(ii)

$$|W_n|^{1/2}\left(\widehat{\theta}_{\widetilde{\mathsf{LPL}}} - \theta^\star\right) \xrightarrow{d} \mathcal{N}\left(0, U_\infty^{-1}\Sigma_\infty U_\infty^{-1}\right).$$

The restriction $\gamma_2 > 5d/2$ includes the standard Lennard-Jones model in dimension $d = 2$ for which $\gamma_2 = 6$. Note that the choice of the sequence α_n, or equivalently of a, is always possible since $\gamma_2 > 5d/2$ implies $d/\{2(\gamma_2 - d)\} < 1/3$.

As shown in the proof of (i), $\Sigma_n^{-1/2}$ can be replaced by $|W_n \ominus \alpha_n|^{-1/2}\Sigma_\infty^{-1/2}$. On the basis of Coeurjolly and Rubak (2013), it should be possible to construct a fast estimator of the asymptotic covariance matrices Σ_∞ and U_∞. This is not investigated here.

Proof. At several places in the proof the sequence $\rho_n = |\mathcal{I}_n|^{1/2}/\alpha_n^{\gamma'}$ for some $\gamma' = \gamma_2 - d - \varepsilon$ and $0 < \varepsilon < \gamma_2 - d$ is involved. Then

$$\rho_n = \frac{|\mathcal{I}_n|^{1/2}}{\alpha_n^{\gamma'}} = \mathcal{O}\left\{|W_n|^{\frac{d/2 - a(\gamma_2-d-\varepsilon)}{d}}\right\} \qquad (4.8)$$

tends to 0 since $a(\gamma_2 - d) > d/2$ and there is no restriction to choose ε sufficiently small to satisfy $a\varepsilon < a(\gamma_2 - d) - d/2$.

We denote by Δ_j the unit cube centered at $j \in \mathbb{Z}^d$, by $\Delta_{n,j} = \Delta_j \cap (W_n \ominus \alpha_n)$ and by $\mathcal{I}_n \subset \mathbb{Z}^d$ the set such that $W_n \ominus \alpha_n = \cup_{j\in\mathcal{I}_n}\Delta_{n,j}$. We write for short $s_{\Delta_{n,j}} = s_{\Delta_{n,j}}(\mathbf{X};\theta^\star)$ and we let $Z_{n,j} = s_{\Delta_{n,j}} - \mathrm{E}(s_{\Delta_{n,j}})$. Then we have

$$s_{W_n \ominus \alpha_n} = S_n + \mathrm{E}(s_{W_n \ominus \alpha_n})$$

where $S_n = \sum_{j\in\mathcal{I}_n} Z_{n,j}$. Clearly $\mathrm{Var}(S_n) = \mathrm{Var}(s_{W_n\ominus\alpha_n}) = \Sigma_n$. The proof of (i) is completed if we show that $\Sigma_n^{-1/2} S_n \xrightarrow{d} \mathcal{N}(0, I_p)$ and $\Sigma_n^{-1/2}\,\mathrm{E}(s_{W_n\ominus\alpha_n}) \to 0$. Let us prove the first convergence by application of Theorem 3.1.

From the assumptions on the set W_n and by definition of $\mathcal{I}_n$, we have $|\mathcal{I}_n| = \mathcal{O}(|W_n|)$, see e.g. Coeurjolly and Møller (2014, Lemma A.1). Therefore the choice of α_n clearly satisfies the hypothesis of Theorem 3.1, namely $\alpha_n^{3d} = o(|\mathcal{I}_n|)$.

Assumption (a) of this theorem holds by definition of $Z_{n,j}$ and Lemma A.3.

Concerning assumption (b), let us introduce the notation, for any $\Delta \subset \mathbb{R}^d$,

$$s'_\Delta = \int_\Delta t(u, \mathbf{X}_{B_{u,n}})\lambda_{\theta^\star}(u, \mathbf{X})\,\mathrm{d}u - \sum_{u\in\mathbf{X}_\Delta} t(u, \mathbf{X}_{B_{u,n}}\backslash u).$$

Note that from the GNZ formula $\mathrm{E}\, s'_\Delta = 0$. We have from Lemma A.4

$$\begin{aligned}
&\sum_{j,k\in\mathcal{I}_n} \big\|\,\mathrm{E}(Z_{n,j}Z_{n,k}^\top)\big\| = \sum_{j,k\in\mathcal{I}_n} \big\|\,\mathrm{Cov}(s_{\Delta_{n,j}}, s_{\Delta_{n,k}})\big\| \\
&\leqslant \sum_{j,k\in\mathcal{I}_n} \big\|\,\mathrm{Cov}(s_{\Delta_{n,j}}, s_{\Delta_{n,k}}) - \mathrm{Cov}(s'_{\Delta_{n,j}}, s'_{\Delta_{n,k}})\big\| + \sum_{j,k\in\mathcal{I}_n} \big\|\,\mathrm{Cov}(s'_{\Delta_{n,j}}, s'_{\Delta_{n,k}})\big\| \\
&\leqslant \sum_{j,k\in\mathcal{I}_n} \left(\frac{c}{\alpha_n^{\gamma'}(1+|k-j|^{\gamma_2})} + \frac{c}{\alpha_n^{2\gamma'}}\right) + \sum_{\substack{j,k\in\mathcal{I}_n\\ |j-k|\leqslant 2r_0}} \big\|\,\mathrm{Cov}(s'_{\Delta_{n,j}}, s'_{\Delta_{n,k}})\big\| + \sum_{\substack{j,k\in\mathcal{I}_n\\ |j-k|> 2r_0}} c|k-j|^{-\gamma_2} \\
&\leqslant c\,\alpha_n^{-\gamma'}|\mathcal{I}_n| + c|\mathcal{I}_n|\rho_n^2 + c|\mathcal{I}_n|\big\|\,\mathrm{Var}(s'_{\Delta_{n,j}})\big\| + c|\mathcal{I}_n|
\end{aligned}$$

which is $\mathcal{O}(|\mathcal{I}_n|)$ by Lemma A.3 and from (4.8). We derive with the same ingredients that

$$\sum_{\substack{j,k\in\mathcal{I}_n\\ |k-j|>\alpha_n}} \big\|\,\mathrm{E}(Z_{n,j}Z_{n,k}^\top)\big\| = o(|\mathcal{I}_n|) + \mathcal{O}(|\mathcal{I}_n|\rho_n^2) = o(|\mathcal{I}_n|)$$

which proves assumption (b).

Since Σ_∞ is assumed to be a positive definite matrix, assumption (c) holds if we prove that $|\mathcal{I}_n|^{-1}\,\mathrm{Var}(S_n) \to \Sigma_\infty$ as $n \to \infty$. For this, let $\Sigma'_n = \mathrm{Var}(s'_{W_n\ominus\alpha_n})$ and

$$\widehat{\Sigma}_n = \sum_{j\in\mathcal{I}_n}\sum_{\substack{k\in\mathcal{I}_n\\ |k-j|\leqslant\alpha_n}} Z_{n,j}Z_{n,k}^\top, \qquad \widehat{\Sigma}'_n = \sum_{j\in\mathcal{I}_n}\sum_{\substack{k\in\mathcal{I}_n\\ |k-j|\leqslant\alpha_n}} s'_{\Delta_{n,j}}(s'_{\Delta_{n,k}})^\top.$$

We have $\|\,|\mathcal{I}_n|^{-1}\Sigma_n - \Sigma_\infty\| \leqslant T_1 + T_2 + T_3 + T_4$ where

$$\begin{aligned}
T_1 &= |\mathcal{I}_n|^{-1}\|\Sigma_n - \mathrm{E}(\widehat{\Sigma}_n)\|, \qquad T_2 = |\mathcal{I}_n|^{-1}\|\,\mathrm{E}(\widehat{\Sigma}_n) - \mathrm{E}(\widehat{\Sigma}'_n)\|, \\
T_3 &= |\mathcal{I}_n|^{-1}\|\,\mathrm{E}(\widehat{\Sigma}'_n) - \Sigma'_n\|, \qquad T_4 = \||\mathcal{I}_n|^{-1}\Sigma'_n - \Sigma_\infty\|.
\end{aligned}$$

First $T_1 \leqslant |\mathcal{I}_n|^{-1}\,\mathrm{E}\,\|\widehat{\Sigma}_n - \Sigma_n\|$ which tends to 0 from (3.1) in Theorem 3.1, as we already proved assumptions (a)-(b). Second applying Lemma A.4

$$\begin{aligned}
T_2 &\leqslant |\mathcal{I}_n|^{-1}\sum_{j\in\mathcal{I}_n}\sum_{\substack{k\in\mathcal{I}_n\\ |k-j|\leqslant\alpha_n}} \big\|\,\mathrm{E}(Z_{n,j}Z_{n,k}^\top) - \mathrm{E}\{s'_{\Delta_{n,j}}(s'_{\Delta_{n,k}})^\top\}\big\| \\
&= |\mathcal{I}_n|^{-1}\sum_{j\in\mathcal{I}_n}\sum_{\substack{k\in\mathcal{I}_n\\ |k-j|\leqslant\alpha_n}} \big\|\,\mathrm{Cov}(s_{\Delta_{n,j}}, s_{\Delta_{n,k}}) - \mathrm{Cov}(s'_{\Delta_{n,j}}, s'_{\Delta_{n,k}})\big\| \\
&= \mathcal{O}(\alpha_n^{-\gamma'}) + \mathcal{O}(\alpha_n^{1-2\gamma'})
\end{aligned}$$

and $T_2 \to 0$ since the assumption $\gamma_2 > 5d/2$ implies $2\gamma' - 1 > 0$ for ε sufficiently small. Third, note that $\Sigma_n' = \sum_{j,k\in\mathcal{I}_n} \mathrm{E}\{s'_{\Delta_{n,j}}(s'_{\Delta_{n,k}})^\top\}$, so that from (iii) in Lemma A.4

$$\begin{aligned} T_3 &\leqslant |\mathcal{I}_n|^{-1} \sum_{j\in\mathcal{I}_n} \sum_{\substack{k\in\mathcal{I}_n \\ |k-j|>\alpha_n}} \| \mathrm{E}\{s'_{\Delta_{n,j}}(s'_{\Delta_{n,k}})^\top\}\| \\ &\leqslant c\,|\mathcal{I}_n|^{-1} \sum_{j\in\mathcal{I}_n} \sum_{\substack{k\in\mathcal{I}_n \\ |k-j|>\alpha_n}} \frac{1}{|k-j|^{\gamma_2}} \leqslant c \sum_{|i|>\alpha_n} \frac{1}{|i|^{\gamma_2}} = o(1). \end{aligned}$$

Finally $T_4 \to 0$ from (ii) in Lemma A.4, which concludes the proof of condition (c) of Theorem 3.1.

To prove assumption (d), we apply the conditional GNZ formula (2.4) to write, for any $j \in \mathcal{I}_n$,

$$\begin{aligned} &\mathrm{E}\left(Z_{n,j} \,|\, X_{\Delta_{n,k}}, k \neq j\right) \\ &\quad = \mathrm{E}\left[\int_{\Delta_{n,j}} t(u, \mathbf{X}_{B_{u,n}}) \left\{\lambda_{\theta^\star}(u,\mathbf{X}) - \lambda_{\theta^\star}(u,\mathbf{X}_{B_{u,n}})\right\} \mathrm{d}u \mid X_{\Delta_{n,k}}, k \neq j\right] \\ &\quad\quad - \mathrm{E}\int_{\Delta_{n,j}} t(u, \mathbf{X}_{B_{u,n}}) \left\{\lambda_{\theta^\star}(u,\mathbf{X}) - \lambda_{\theta^\star}(u,\mathbf{X}_{B_{u,n}})\right\} \mathrm{d}u. \end{aligned}$$

From Lemma A.1, we have for any $u \in \mathbb{R}^d$ and $\mathbf{x} \in \Omega$

$$\|t(u, \mathbf{x}_{B_{u,n}})\| \, |\lambda_{\theta^\star}(u,\mathbf{x}) - \lambda_{\theta^\star}(u,\mathbf{x}_{B_{u,n}})| \leqslant \frac{c}{\alpha_n^{\gamma'}} Y(u,\mathbf{x})$$

where $Y(u,\mathbf{x}) = \|\{|t_m|(u,\mathbf{x})\}_{m\geqslant 1}\| H(u,\mathbf{x}) e^{c\,G(u,\mathbf{x})} \lambda_{\theta^\star}(u,\mathbf{x})$ using the notation of the lemma. Since $|\Delta_{n,j}| \leqslant 1$, we deduce from the stationarity of $\mathbf{X}$ and Lemma A.2 that

$$\mathrm{E}\left\|\mathrm{E}\left(Z_{n,j} \mid \mathbf{X}_{\Delta_{n,k}}, k \neq j\right)\right\| \leqslant \frac{c}{\alpha_n^{\gamma'}} \mathrm{E}\{Y(0,\mathbf{X})\} = \mathcal{O}(\alpha_n^{-\gamma'}).$$

Hence

$$|\mathcal{I}_n|^{-1/2} \sum_{j\in\mathcal{I}_n} \mathrm{E}\,\|\,\mathrm{E}\left(Z_{n,j} \mid \mathbf{X}_{\Delta_{n,k}}, k \neq j\right)\| = \mathcal{O}(\rho_n)$$

tends to 0 from (4.8). All conditions of Theorem 3.1 are therefore satisfied, which yields that $\Sigma_n^{-1/2} S_n \xrightarrow{d} \mathcal{N}(0, I_p)$. It remains to prove that $\Sigma_n^{-1/2}\,\mathrm{E}(s_{W_n\ominus\alpha_n}) \to 0$. This is a consequence of the GNZ formula, (3.3), Lemma A.1 and the condition $a(\gamma_2 - d) > d/2$ since

$$\begin{aligned} \|\Sigma_n^{-1/2}\,\mathrm{E}(s_{W_n\ominus\alpha_n})\| &\leqslant \|\Sigma_n^{-1/2}\|\,\|\,\mathrm{E}(s_{W_n\ominus\alpha_n})\| \\ &\leqslant c\,|\mathcal{I}_n|^{-1/2}\alpha_n^{-\gamma'}\,\mathrm{E}\int_{W_n\ominus\alpha_n} Y(u,\mathbf{X})\,\mathrm{d}u = \mathcal{O}\left(\rho_n\right) = o(1). \end{aligned}$$

(ii) It is worth repeating that $\theta \to -\widetilde{\mathsf{LPL}}_{W_n\ominus\alpha_n}(\mathbf{x};\theta)$ is a convex function with Hessian matrix given by (4.4). Following Lemmas A.1-A.2 and arguments developed in the proof of Proposition 4.1, we leave the reader to check that almost surely

$$|W_n \ominus \alpha_n|^{-1} \left\{ \frac{\mathrm{d}}{\mathrm{d}\theta\,\mathrm{d}\theta^\top} \widetilde{\mathsf{LPL}}_{W_n\ominus\alpha_n}(\mathbf{X};\theta) - \frac{\mathrm{d}}{\mathrm{d}\theta\,\mathrm{d}\theta^\top} \mathsf{LPL}_{W_n\ominus\alpha_n}(\mathbf{X};\theta) \right\} \to 0$$

and

$$-|W_n \ominus \alpha_n|^{-1} \frac{\mathrm{d}}{\mathrm{d}\theta\, \mathrm{d}\theta^\top} \mathsf{LPL}_{W_n \ominus \alpha_n}(\mathbf{X}; \theta) \to \mathrm{E}\left\{ t(0, \mathbf{X}) t(0, \mathbf{X})^\top \lambda_\theta(0, \mathbf{X}) \right\}$$

as $n \to \infty$, which equals to U_∞ when $\theta = \theta^\star$. We also note that (4.5) implies that U_∞ is a positive definite matrix. These facts and (i) allow us to apply Guyon (1995, Theorem 3.4.5) to deduce the result. □

The following proposition focuses on the maximum logistic regression and states its strong consistency and asymptotic normality. The result is given without proof, but we claim that it follows by the same arguments as those involved in the proofs of Proposition 4.1 and Theorem 4.2.

Proposition 4.3. *Under the assumptions of Proposition 4.1 and Theorem 4.2, the maximum logistic regression estimator defined by*

$$\widehat{\theta}_{\widetilde{\mathsf{LRL}}} = \mathrm{argmax}_{\theta \in \Theta} \widetilde{\mathsf{LRL}}_{W_n \ominus \alpha_n}(\mathbf{X}; \theta)$$

converges almost surely to $\theta^\star$ as $n \to \infty$ and satisfies the following convergence in distribution

$$|W_n|^{1/2} \left(\widehat{\theta}_{\widetilde{\mathsf{LRL}}} - \theta^\star \right) \xrightarrow{d} \mathcal{N}\left(0, V_\infty^{-1} \Gamma_\infty V_\infty^{-1}\right)$$

where denoting $h(u, \mathbf{x}) = \rho\, t(u, \mathbf{x}) / \{\lambda_{\theta^\star}(u, \mathbf{x}) + \rho\}$ for any $u \in \mathbb{R}^d, \mathbf{x} \in \Omega$,

$$\begin{aligned}
\Gamma_\infty &= \mathrm{E}\left\{ h(0, \mathbf{X}) h(0, \mathbf{X})^\top \lambda_{\theta^\star}(0, \mathbf{X}) \right\} \\
&\quad + \int_{\mathbb{R}^d} \mathrm{E}\left\{ h(0, \mathbf{X}) h(v, \mathbf{X})^\top \lambda_{\theta^\star}(0, \mathbf{X}) \lambda_{\theta^\star}(v, \mathbf{X}) \right\} \{1 - e^{-\Phi_{\theta^\star}(v)}\}\, \mathrm{d}v \\
&\quad + \int_{\mathbb{R}^d} \mathrm{E}\left\{ \lambda_{\theta^\star}(0, \mathbf{X}) \lambda_{\theta^\star}(v, \mathbf{X}) \Delta_v h(0, \mathbf{X}) \Delta_0 h(v, \mathbf{X})^\top \right\} e^{-\Phi_{\theta^\star}(v)}\, \mathrm{d}v \\
V_\infty &= \frac{1}{\rho} \mathrm{E}\left\{ h(0, \mathbf{X}) h(0, \mathbf{X})^\top \lambda_{\theta^\star}(0, \mathbf{X}) \right\}
\end{aligned}$$

with $\Delta_v h(u, \mathbf{x}) = h(u, \mathbf{x} \cup v) - h(u, \mathbf{x})$ for any $u, v \in \mathbb{R}^d$ and $\mathbf{x} \in \Omega$.

A Auxiliary results

We gather in this section several auxiliary results. They are established under the setting, assumptions and notation of Section 4. In particular, we recall that Δ_j is the cube centered at $j \in \mathbb{Z}^d$ with volume 1, $\Delta_{n,j} = \Delta_j \cap (W_n \ominus \alpha_n)$, $W_n = \cup_{j \in \mathcal{I}_n} \Delta_{n,j}$, $B_{u,n} = B(u, \alpha_n)$ and

$$s_\Delta = s_\Delta(\mathbf{X}; \theta^\star) = \int_\Delta t(u, \mathbf{X}_{B_{u,n}}) \lambda_{\theta^\star}(u, \mathbf{X}_{B_{u,n}})\, \mathrm{d}u - \sum_{u \in \mathbf{X}_\Delta} t(u, \mathbf{X}_{B_{u,n}} \backslash u) \tag{A.1}$$

$$s'_\Delta = s'_\Delta(\mathbf{X}; \theta^\star) = \int_\Delta t(u, \mathbf{X}_{B_{u,n}}) \lambda_{\theta^\star}(u, \mathbf{X})\, \mathrm{d}u - \sum_{u \in \mathbf{X}_\Delta} t(u, \mathbf{X}_{B_{u,n}} \backslash u). \tag{A.2}$$

Lemma A.1. *Let $j \in \mathcal{I}_n$ and $u \in \Delta_{n,j}$, assume* **[g]**, *set $\gamma' = \gamma_2 - d - \varepsilon$ where $0 < \varepsilon < \gamma_2 - d$ and define*

$$|t_m|(u, \mathbf{x}) = \sum_{v \in \mathbf{x}} |g_m(v-u)|$$
$$G(u, \mathbf{x}) = \sum_{v \in \mathbf{x}} \|v-u\|^{-\gamma_2} \mathbf{1}(\|v-u\| \geqslant r_0)$$
$$H(u, \mathbf{x}) = \sum_{v \in \mathbf{x}} \|v-u\|^{-d-\varepsilon} \mathbf{1}(\|v-u\| \geqslant r_0).$$

Then, if $\alpha_n \geqslant r_0$

(i) $|t_m|(u, \mathbf{x}_{B_{u,n}})| \leqslant |t_m|(u, \mathbf{x})$

(ii) $|t_m(u, \mathbf{x}) - t_m(u, \mathbf{x}_{B_{u,n}})| \leqslant c \min\{G(u, \mathbf{x}), \alpha_n^{-\gamma'} H(u, \mathbf{x})\}$

(iii) $\forall \theta \in \Theta$, $|\lambda_\theta(u, \mathbf{x}) - \lambda_\theta(u, \mathbf{x}_{B_{u,n}})| \leqslant c\, e^{c\, G(u,\mathbf{x})} \lambda_\theta(u, \mathbf{x}) \min\{G(u, \mathbf{x}), \alpha_n^{-\gamma'} H(u, \mathbf{x})\}$.

Proof. The first statement is straightforward from the definition. For the second one, from **[g]** and since $\alpha_n \geqslant r_0$,

$$\begin{aligned} |t_m(u, \mathbf{x}) - t_m(u, \mathbf{x}_{B_{u,n}})| &= |\sum_{v \in \mathbf{x}} g_m(v-u) \mathbf{1}(\|v-u\| \geqslant \alpha_n)| \\ &\leqslant c \sum_{v \in \mathbf{x}} \|v-u\|^{-\gamma_2} \mathbf{1}(\|v-u\| \geqslant \alpha_n), \end{aligned}$$

which is clearly lower than $c\, G(u, \mathbf{x})$. Pushing one step further, we get

$$|t_m(u, \mathbf{x}) - t_m(u, \mathbf{x}_{B_{u,n}})| \leqslant c\, \alpha_n^{-\gamma'} \sum_{v \in \mathbf{x}} \|v-u\|^{-d-\varepsilon} \mathbf{1}(\|v-u\| \geqslant \alpha_n) \leqslant c\, \alpha_n^{-\gamma'} H(u, \mathbf{x}),$$

which proves *(ii)*. For the third statement, since for all x, $|1 - e^x| < |x| e^{|x|}$, we have

$$\begin{aligned} |\lambda_\theta(u, \mathbf{x}) - \lambda_\theta(u, \mathbf{x}_{B_{u,n}})| &= \lambda_\theta(u, \mathbf{x}) \Big| 1 - e^{\sum_{v \in \mathbf{x}_{B_{u,n}^c}} \Phi_\theta(v-u)} \Big| \\ &\leqslant \lambda_\theta(u, \mathbf{x}) \Big| \sum_{v \in \mathbf{x}_{B_{u,n}^c}} \Phi_\theta(v-u) \Big| e^{\left| \sum_{v \in \mathbf{x}_{B_{u,n}^c}} \Phi_\theta(v-u) \right|}. \end{aligned}$$

The result follows from the same inequalities as before, noting that

$$\Big| \sum_{v \in \mathbf{x}_{B_{u,n}^c}} \Phi_\theta(v-u) \Big| = \Big| \sum_{m=2}^{p} \theta_m \sum_{v \in \mathbf{x}_{B_{u,n}^c}} g_m(v-u) \Big| \leqslant c \sum_{v \in \mathbf{x}} |g_m(v-u)| \mathbf{1}(\|v-u\| \geqslant \alpha_n)$$

where $c = (p-1) \sup_{\theta \in \Theta} \sup_m |\theta_m| < \infty$, since Θ is bounded. □

Lemma A.2. *Under the assumption* **[g]**, *then for any $\theta \in \Theta$ we have the following statements where* E *denotes the expectation with respect to $P_{\theta^\star}$.*

(i) For any $q \geqslant 0$, $\mathrm{E}\{\lambda_\theta(0, \mathbf{X})^q\} < \infty$.

(ii) *Let $f : \mathbb{R}^d \to \mathbb{R}$ be a measurable function such that $|f(u)| \leqslant c(1 + \|u\|)^{-\gamma}$ with $\gamma > d$, then for any $q \geqslant 0$*

$$\mathrm{E}\left\{e^{q\left|\sum_{u\in\mathbf{X}} f(\|u\|)\right|}\right\} < \infty.$$

(iii) *For any $q \geqslant 0$, $q' > 0$ and $\theta \in \Theta$, $\mathrm{E}\{|t_m|(0,\mathbf{X})^q \lambda_\theta(0,\mathbf{X})^{q'}\} < \infty$.*

(iv) *Let f_1 and f_2 be two functions as in (ii), then for any $q_1, q_2, q_3 \geqslant 0$ and $q' > 0$,*

$$\mathrm{E}\left\{|t_m|(0,\mathbf{X})^{q_1}\left|\sum_{u\in\mathbf{X}} f_1(\|u\|)\right|^{q_2} e^{q_3\left|\sum_{u\in\mathbf{X}} f_2(\|u\|)\right|}\lambda_\theta(0,\mathbf{X})^{q'}\right\} < \infty.$$

Proof. The first statement is a consequence of Proposition 5.2 (a) in Ruelle (1970). It relies on the following property, see also Mase (1995, Lemma 2). If $\psi : \mathbb{R}^+ \to \mathbb{R}^+$ is a decreasing function with $\int_0^\infty \psi(t)t^{d-1}\,\mathrm{d}t < \infty$, then for any $q \geqslant 0$,

$$\mathrm{E}(e^{q\sum_{u\in\mathbf{X}}\psi(\|u\|)}) < \infty.$$

The proof of (ii) is an easy consequence of this property. We deduce in particular that all moments of $\sum_{u\in\mathbf{X}} f(\|u\|)$ exist and are finite. Assuming (iii) is true, then (iv) is a straightforward consequence of the previous properties and Hölder's inequality. Let us prove (iii). For any $\varepsilon > 0$, using the fact that for any $q \geqslant 0, \kappa > 0$, $x \mapsto x^q e^{-\kappa x}$ is bounded on $[0,\infty)$, we have

$$\begin{aligned}
|t_m|(0,\mathbf{x})^q \lambda_\theta(0,\mathbf{x})^{q'} &= |t_m|(0,\mathbf{x})^q e^{-q'\sum_{k=2}^p \theta_k t_k(0,\mathbf{x})} \\
&= |t_m|(0,\mathbf{x})^q e^{-q'\varepsilon|\theta_m t_m(0,\mathbf{x})|} e^{q'\varepsilon|\theta_m t_m(0,\mathbf{x})| - q'\sum_{k=2}^p \theta_k t_k(0,\mathbf{x})} \\
&\leqslant c\; e^{-q'\sum_{u\in\mathbf{x}}\tilde{\Phi}_\theta(u)},
\end{aligned}$$

where $\tilde{\Phi}_\theta(u) = \sum_{k=2}^p \theta_k g_k(u) - \varepsilon|\theta_m g_m(u)|$. The proof of (iii) is completed in view of (i) if we show that $\tilde{\Phi}_\theta$ satisfies **[Φ]** for any θ. Write $\tilde{\Phi}_\theta(u) = \tilde{\Phi}_1(u) + \tilde{\Phi}_2(u)$ with

$$\tilde{\Phi}_1(u) = \frac{\theta_2}{2} g_2(u) + \sum_{k=3}^p \theta_k g_k(u), \quad \tilde{\Phi}_2(u) = \frac{\theta_2}{2} g_2(u) - \varepsilon|\theta_m g_m(u)|.$$

From **[g]**, we deduce that there exists $r > 0$ such that $\|u\| < r$ implies $\tilde{\Phi}_1(u) > c\|u\|^{-\gamma_1}$. Moreover if $m = 2$, $\tilde{\Phi}_2(u) > 0$ for all $\|u\| < r_0$, provided $\varepsilon < 1/2$. If $m \geqslant 3$, there exists r' such that $\|u\| < r'$ implies $|\theta_m g_m(u)| < c_g\|u\|^{-\gamma_1}/(4\varepsilon)$ where c_g is the constant in **[g]**, yielding $\tilde{\Phi}_2(u) > (c_g/4)\|u\|^{-\gamma_1}$. In all cases, we obtain that for some $r_1 > 0$, $\|u\| < r_1$ implies $\tilde{\Phi}(u) > c\|u\|^{-\gamma_1}$. On the other hand, it is clear that if $\|u\| > r_0$ then $|\tilde{\Phi}(u)| \leqslant c\|u\|^{-\gamma_2}$ and that $\tilde{\Phi}_\theta$ is bounded from below, proving that it satisfies **[Φ]**. □

Lemma A.3. *Let $j \in \mathcal{I}_n$ and $s_m = (s_{\Delta_{n,j}})_m$, respectively $s'_m = (s'_{\Delta_{n,j}})_m$, be the m-th coordinate of $s_{\Delta_{n,j}}$ given by* (A.1)*, respectively $s'_{\Delta_{n,j}}$ given by* (A.2)*. Under* **[g]***, if $\alpha_n \geqslant r_0$ then, for any $q \in \mathbb{N}$, $\mathrm{E}(|s_m|^q) < \infty$ and $\mathrm{E}(|s'_m|^q) < \infty$.*

Proof. The proof being similar for s_m and s'_m, we only give the details concerning s_m. From (A.1) and the binomial formula, the statement is a consequence of

$$\mathrm{E}\left\{\left|\sum_{u\in\mathbf{X}_{\Delta_{n,j}}} t_m(u,\mathbf{X}_{B_{u,n}}\backslash u)\right|^{p_1}\left|\int_{\Delta_{n,j}} t_m(u,\mathbf{X}_{B_{u,n}})\lambda_{\theta^\star}(u,\mathbf{X}_{B_{u,n}})\,\mathrm{d}u\right|^{p_2}\right\}<\infty$$

for any $p_1,p_2\in\mathbb{N}$. Applying the Cauchy-Schwarz's inequality, we consider each term above separately. First, for any $p\in\mathbb{N}$, by Hölder's inequality and using Lemma A.1 we get

$$\begin{aligned}
&\mathrm{E}\left\{\left|\int_{\Delta_{n,j}} t_m(u,\mathbf{X}_{B_{u,n}})\lambda_{\theta^\star}(u,\mathbf{X}_{B_{u,n}})\,\mathrm{d}u\right|^{p}\right\}\\
&\qquad\leqslant c\int_{\Delta_{n,j}}\mathrm{E}\left\{\left|t_m(u,\mathbf{X}_{B_{u,n}})\right|^p\lambda_{\theta^\star}(u,\mathbf{X}_{B_{u,n}})^p\right\}\,\mathrm{d}u\\
&\qquad\leqslant c\int_{\Delta_{n,j}}\mathrm{E}\left[|t_m(u,\mathbf{X})|^p\left\{\lambda_{\theta^\star}(u,\mathbf{X})^p+|\lambda_{\theta^\star}(u,\mathbf{X}_{B_{u,n}})-\lambda_{\theta^\star}(u,\mathbf{X})|^p\right\}\right]\,\mathrm{d}u\\
&\qquad\leqslant c\int_{\Delta_{n,j}}\mathrm{E}\left[|t_m(u,\mathbf{X})|^p\,\lambda_{\theta^\star}(u,\mathbf{X})^p\left\{1+G^p(u,\mathbf{x})e^{c\,p\,G(u,\mathbf{x})}\right\}\right]\,\mathrm{d}u
\end{aligned}$$

which is finite by Lemma A.2 and the stationarity of $\mathbf{X}$.

Second, we can prove by induction and successive application of the GNZ formula, see Corollary 3.1 in Decreusefond and Flint (2014), that

$$\begin{aligned}
&\mathrm{E}\left[\left\{\sum_{u\in\mathbf{X}_{\Delta_{n,j}}} t_m(u,\mathbf{X}_{B_{u,n}}\backslash u)\right\}^p\right]\\
&\qquad=\sum_{k=1}^{p}\sum_{(\mathcal{P}_1,\ldots,\mathcal{P}_k)\in\mathcal{T}_p^k}\mathrm{E}\int_{\Delta_{n,j}^k}\lambda_{\theta^\star}(\{u_1,\ldots,u_k\},\mathbf{X})\prod_{\ell=1}^{k}t_m^{|\mathcal{P}_\ell|}(u_\ell,\mathbf{X}_{B_{u,n}}\cup\{\mathbf{u}\backslash u_\ell\})\,\mathrm{d}\mathbf{u}
\end{aligned}$$

where $\mathcal{T}_p^k$ is the set of all partitions of $\{1,\ldots,p\}$ into k subsets, $|\mathcal{P}|$ is the cardinality of $\mathcal{P}$, $\mathbf{u}=(u_1,\ldots,u_k)$ and $\mathbf{u}\backslash u_\ell=(u_1,\ldots,u_{\ell-1},u_{\ell+1},\ldots,u_k)$. Since

$$\lambda_{\theta^\star}(\{u_1,\ldots,u_k\},\mathbf{X})=\prod_{\ell=1}^{k}\lambda_{\theta^\star}(u_\ell,\mathbf{X})\prod_{i=1,i\neq\ell}^{k}e^{-\Phi_{\theta^\star}(u_i-u_\ell)},$$

we obtain by application of Hölder's inequality,

$$\begin{aligned}
&\mathrm{E}\left\{\left|\sum_{u\in\mathbf{X}_{\Delta_{n,j}}} t_m(u,\mathbf{X}_{B_{u,n}}\backslash u)\right|^{p}\right\}\leqslant\\
&\sum_{k=1}^{p}\sum_{\mathcal{P}\in\mathcal{T}_p^k}\prod_{\ell=1}^{k}\mathrm{E}^{1/k}\int_{\Delta_{n,j}^k}\lambda_{\theta^\star}(u_\ell,\mathbf{X})^k|t_m(u_\ell,\mathbf{X}_{B_{u,n}}\cup\{\mathbf{u}\backslash u_\ell\})|^{k|\mathcal{P}|}\prod_{i=1,i\neq\ell}^{k}e^{-k\Phi_{\theta^\star}(u_i-u_\ell)}\,\mathrm{d}\mathbf{u}.
\end{aligned}$$

The proof is completed if we show that all expectations above are finite. To that end, note that

$$t_m(u_\ell, \mathbf{X}_{B_{u,n}} \cup \{\mathbf{u} \backslash u_\ell\}) = t_m(u_\ell, \mathbf{X}_{B_{u,n}}) + \sum_{h=1, h\neq\ell}^{k} g_m(u_h - u_\ell)$$

whereby, denoting $q = k|\mathcal{P}|$

$$\begin{aligned}
&\mathrm{E} \int_{\Delta^k_{n,j}} \lambda_{\theta^\star}(u_\ell, \mathbf{X})^k |t_m(u_\ell, \mathbf{X}_{B_{u,n}} \cup \{\mathbf{u} \backslash u_\ell\})|^q \prod_{i=1, i\neq\ell}^{k} e^{-k\Phi_{\theta^\star}(u_i - u_\ell)} \, \mathrm{d}\mathbf{u} \\
&\leqslant \sum_{r=0}^{q} \binom{q}{r} \int_{\Delta^k_{n,j}} \left| \sum_{h=1, h\neq\ell}^{k} g_m(u_h - u_\ell) \right|^r \times \\
&\qquad\qquad \prod_{i=1, i\neq\ell}^{k} e^{-k\Phi_{\theta^\star}(u_i - u_\ell)} \, \mathrm{E}\left\{ |t_m(u_\ell, \mathbf{X}_{B_{u,n}})|^{q-r} \lambda_{\theta^\star}(u_\ell, \mathbf{X})^k \right\} \, \mathrm{d}\mathbf{u} \\
&\leqslant c \sum_{r=0}^{q} \sum_{h=1, h\neq\ell}^{k} \int_{\Delta^k_{n,j}} |g_m(u_h - u_\ell)|^r \times \\
&\qquad\qquad \prod_{i=1, i\neq\ell}^{k} e^{-k\Phi_{\theta^\star}(u_i - u_\ell)} \, \mathrm{E}\left\{ |t_m(u_\ell, \mathbf{X}_{B_{u,n}})|^{q-r} \lambda_{\theta^\star}(u_\ell, \mathbf{X})^k \right\} \, \mathrm{d}\mathbf{u}.
\end{aligned}$$

The last expectation is finite in view of Lemma A.2, so the above expression is lower than

$$\begin{aligned}
&c \sum_{r=0}^{q} \sum_{h=1, h\neq\ell}^{k} \int_{\Delta^k_{n,j}} |g_m(u_h - u_\ell)|^r \prod_{i=1, i\neq\ell}^{k} e^{-k\Phi_{\theta^\star}(u_i - u_\ell)} \, \mathrm{d}\mathbf{u} \\
&\qquad\qquad \leqslant c \left\{ \int_{\mathbb{R}^d} e^{-k\Phi_{\theta^\star}(v)} \, \mathrm{d}v \right\}^{k-2} \sum_{r=0}^{q} \int_{\mathbb{R}^d} |g_m(v)|^r e^{-k\Phi_{\theta^\star}(v)} \, \mathrm{d}v,
\end{aligned}$$

which is finite from **[g]**. □

Lemma A.4. *The following properties hold under the assumption* **[g]**.

(i) For Λ_1, Λ_2 two bounded Borel sets of $\mathbb{R}^d$

$$\begin{aligned}
&\mathrm{Cov}(s'_{\Lambda_1}, s'_{\Lambda_2}) = \mathrm{E} \int_{\Lambda_1 \cap \Lambda_2} t(u, \mathbf{X}_{B_{u,n}}) t(u, \mathbf{X}_{B_{u,n}})^\top \lambda_{\theta^\star}(u, \mathbf{X}) \, \mathrm{d}u \\
&+ \mathrm{E} \int_{\Lambda_1} \int_{\Lambda_2} t(u, \mathbf{X}_{B_{u,n}}) t(v, \mathbf{X}_{B_{v,n}})^\top \{ \lambda_{\theta^\star}(u, \mathbf{X}) \lambda_{\theta^\star}(v, \mathbf{X}) - \lambda_{\theta^\star}(\{u, v\}, \mathbf{X}) \} \, \mathrm{d}u \, \mathrm{d}v \\
&+ \mathrm{E} \int_{\Lambda_1} \int_{\Lambda_2} \Delta_v t(u, \mathbf{X}_{B_{u,n}}) \{ \Delta_u t(v, \mathbf{X}_{B_{v,n}}) \}^\top \lambda_{\theta^\star}(\{u, v\}, \mathbf{X}) \, \mathrm{d}u \, \mathrm{d}v
\end{aligned}$$

where for any $u, v \in \mathbb{R}^d$, $\mathbf{x} \in \Omega$ and any measurable function $f : \mathbb{R}^d \times \Omega \to \mathbb{R}^p$, the difference operator Δ_v is defined by $\Delta_v f(u, \mathbf{x}) = f(u, \mathbf{x} \cup v) - f(u, \mathbf{x})$.

(ii) *Let $(\Delta_n)_{n\geqslant 1}$ be a sequence of increasing domains such that $\Delta_n \to \mathbb{R}^d$ as $n \to \infty$, then*

$$|\Delta_n|^{-1} \operatorname{Var}(s'_{\Delta_n}) \to \Sigma_\infty$$

where Σ_∞ is defined by (4.6).

(iii) *Let $j, k \in \mathcal{I}_n$. Then if $\alpha_n \geqslant r_0$ and $|k-j| > 2r_0$,*

$$\left\| \operatorname{Cov}(s'_{\Delta_{n,j}}, s'_{\Delta_{n,k}}) \right\| \leqslant c|k-j|^{-\gamma_2}.$$

(iv) *For $j, k \in \mathcal{I}_n$ denote*

$$\mathcal{C}_{n,jk} = \operatorname{Cov}(s_{\Delta_{n,j}}, s_{\Delta_{n,k}}) - \operatorname{Cov}(s'_{\Delta_{n,j}}, s'_{\Delta_{n,k}}).$$

Then if $\alpha_n \geqslant r_0$, for any $j, k \in \mathcal{I}_n$,

$$\begin{aligned}
\|\mathcal{C}_{n,jk}\| &\leqslant \frac{c}{\alpha_n^{\gamma'}(1+|k-j|^{\gamma_2})} + \frac{c}{\alpha_n^{2\gamma'}} \\
\sum_{j,k\in\mathcal{I}_n} \|\mathcal{C}_{n,jk}\| &= \mathcal{O}(\alpha_n^{-\gamma'}|\mathcal{I}_n|) + \mathcal{O}(\alpha_n^{-2\gamma'}|\mathcal{I}_n|^2) \\
\sum_{\substack{j,k\in\mathcal{I}_n \\ |k-j|\leqslant\alpha_n}} \|\mathcal{C}_{n,jk}\| &= \mathcal{O}(\alpha_n^{-\gamma'}|\mathcal{I}_n|) + \mathcal{O}(\alpha_n^{1-2\gamma'}|\mathcal{I}_n|) \\
\sum_{\substack{j,k\in\mathcal{I}_n \\ |k-j|>\alpha_n}} \|\mathcal{C}_{n,jk}\| &= o(|\mathcal{I}_n|) + \mathcal{O}(\alpha_n^{-2\gamma'}|\mathcal{I}_n|^2),
\end{aligned}$$

as $n \to \infty$, where we recall that $\gamma' = \gamma_2 - d - \varepsilon$ with $0 < \varepsilon < \gamma_2 - d$.

Proof. (i) is a slight extension of Coeurjolly and Rubak (2013, Lemma 3.1) where the case $\Lambda_1 = \Lambda_2$ was considered. The proof is omitted.

For (ii), we note that for any $u, v \in \mathbb{R}^d$, $m \geqslant 1$ and $\mathbf{x} \in \Omega$

$$\lambda_{\theta^\star}(u,\mathbf{x})\lambda_{\theta^\star}(v,\mathbf{x}) - \lambda_{\theta^\star}(\{u,v\},\mathbf{x}) = \lambda_{\theta^\star}(u,\mathbf{x})\lambda_{\theta^\star}(v,\mathbf{x})\{1 - e^{-\Phi_{\theta^\star}(v-u)}\} \tag{A.3}$$

and

$$\Delta_v t_m(u,\mathbf{x}) = t_m(u,\mathbf{x}\cup v) - t_m(u,\mathbf{x}) = \begin{cases} 0 & \text{if } m = 1 \\ g_m(v) & \text{if } m \geqslant 2 \end{cases} \tag{A.4}$$

which leads to $\Delta_u t(v,\mathbf{x}) = g(v)$. Letting $|t|(u,\mathbf{x}) = \{|t_m|(u,\mathbf{x})\}_{m\geqslant 1}$ for any $u \in \mathbb{R}^d$ and $\mathbf{x} \in \Omega$, we have for any $u, v \in \mathbb{R}^d$

$$\|t(u,\mathbf{X}_{B_{u,n}})t(v,\mathbf{X}_{B_{v,n}})^\top\| \leqslant \| \, |t|(u,\mathbf{X})|t|(v,\mathbf{X})^\top\|.$$

The result is derived using the dominated convergence theorem, the stationarity of $\mathbf{X}$ and since from Lemma A.2 the random variables $\| \, |t|(0,\mathbf{X})|t|(0,\mathbf{X})^\top\|\lambda_{\theta^\star}(0,\mathbf{X})$ and $\| \, |t|(0,\mathbf{X})|t|(v,\mathbf{X})^\top\|\lambda_{\theta^\star}(0,\mathbf{X})\lambda_{\theta^\star}(v,\mathbf{X})$ have expectation uniformly bounded in v while by **[g]**

$$\int_{\mathbb{R}^d} |1 - e^{-\Phi_{\theta^\star}(v)}| \, \mathrm{d}v < \infty \quad \text{and} \quad \int_{\mathbb{R}^d} \|g(v)g(v)^\top\| e^{-\Phi_{\theta^\star}(v)} \, \mathrm{d}v < \infty.$$

To prove (iii), we apply (i) to the disjoint sets $\Delta_{n,j}$, $\Delta_{n,k}$ and relations (A.3)-(A.4) to get

$$\begin{aligned}
&\left\|\mathrm{Cov}(s'_{\Delta_{n,j}}, s'_{\Delta_{n,k}})\right\| \\
&\leqslant \mathrm{E}\int_{\Delta_{n,j}}\int_{\Delta_{n,k}} \| \, |t|(u,\mathbf{X})|t|(v,\mathbf{X})^\top \| \lambda_{\theta^\star}(u,\mathbf{X})\lambda_{\theta^\star}(v,\mathbf{X})|1-e^{-\Phi_{\theta^\star}(v-u)}| \,\mathrm{d}u\,\mathrm{d}v \\
&\quad + \mathrm{E}\int_{\Delta_{n,j}}\int_{\Delta_{n,k}} \|g(v-u)g(v-u)^\top\| \, \lambda_{\theta^\star}(u,\mathbf{X})\lambda_{\theta^\star}(v,\mathbf{X})e^{-\Phi_{\theta^\star}(v-u)} \,\mathrm{d}u\,\mathrm{d}v. \quad \text{(A.5)}
\end{aligned}$$

Since $|k-j| > 2r_0$, we deduce from **[g]** that for any $(u,v) \in \Delta_{n,j} \times \Delta_{n,k}$ and any $m \geqslant 2$, $|g_m(v-u)| \leqslant c|k-j|^{-\gamma_2}$. This leads to

$$\|g(v-u)g(v-u)^\top\| \leqslant c|k-j|^{-\gamma_2}.$$

Similarly since $\Phi_{\theta^\star} = \sum_{m=2}^p \theta^\star_m g_m$, for any $(u,v) \in \Delta_{n,j}\times\Delta_{n,k}$, $e^{-\Phi_{\theta^\star}(v-u)} \leqslant e^{|\Phi_{\theta^\star}(v-u)|} \leqslant c$ and

$$|1-e^{-\Phi_{\theta^\star}(v-u)}| \leqslant |\Phi_{\theta^\star}(v-u)|e^{|\Phi_{\theta^\star}(v-u)|} \leqslant c|k-j|^{-\gamma_2}. \quad \text{(A.6)}$$

Plugging these inequalities in (A.5) shows (ii), as the remaining terms have finite expectations from Lemma A.2.

We now focus on (iv). Let us write $s_{\Delta_{n,j}} = s'_{\Delta_{n,j}} + I_j$ where $I_j = \int_{\Delta_{n,j}} \pi_n(u,\mathbf{x})\,\mathrm{d}u$ and

$$\pi_n(u,\mathbf{x}) = t(u,\mathbf{x}_{B_{u,n}})\{\lambda_{\theta^\star}(u,\mathbf{x}_{B_{u,n}}) - \lambda_{\theta^\star}(u,\mathbf{x})\}.$$

We have

$$\begin{aligned}
&\mathrm{Cov}(s_{\Delta_{n,j}},s_{\Delta_{n,k}}) - \mathrm{Cov}(s'_{\Delta_{n,j}}, s'_{\Delta_{n,k}}) \\
&\qquad = \mathrm{E}(s'_{\Delta_{n,j}} I_k^\top) + \mathrm{E}(s'_{\Delta_{n,k}} I_j^\top) + \mathrm{E}(I_j I_k^\top) - \mathrm{E}(s_{\Delta_{n,j}})\,\mathrm{E}\,(s_{\Delta_{n,k}})^\top. \quad \text{(A.7)}
\end{aligned}$$

Let us control each term in (A.7). From the GNZ formula

$$\mathrm{E}(s'_{\Delta_{n,j}} I_k^\top) = \mathrm{E}\int_{\Delta_{n,j}}\int_{\Delta_{n,k}} t(u,\mathbf{X})\lambda_{\theta^\star}(u,\mathbf{X})\{\pi_n(v,\mathbf{X}) - \pi_n(v,\mathbf{X}\cup u)\}^\top \,\mathrm{d}u\,\mathrm{d}v.$$

By definition of $\lambda_{\theta^\star}$ and t (see (2.2) and (4.2)), we have for any $u,v \in \mathbb{R}^d$ and $\mathbf{x} \in \Omega$

$$\pi_n(v,\mathbf{x}\cup u) = e^{-\Phi_{\theta^\star}(v-u)}[\pi_n(v,\mathbf{x}) + g(v-u)\{\lambda_{\theta^\star}(v,\mathbf{x}_{B_{v,n}}) - \lambda_{\theta^\star}(v,\mathbf{x})\}],$$

whereby

$$\begin{aligned}
\mathrm{E}(s'_{\Delta_{n,j}} I_k^\top) &= \mathrm{E}\int_{\Delta_{n,j}}\int_{\Delta_{n,k}} t(u,\mathbf{X})\lambda_{\theta^\star}(u,\mathbf{X})\big\{1-e^{-\Phi_{\theta^\star}(v-u)}\big\}\pi_n(v,\mathbf{X})^\top \,\mathrm{d}u\,\mathrm{d}v \\
&- \mathrm{E}\int_{\Delta_{n,j}}\int_{\Delta_{n,k}} t(u,\mathbf{X})\lambda_{\theta^\star}(u,\mathbf{X})e^{-\Phi_{\theta^\star}(v-u)}g(v-u)^\top\{\lambda_{\theta^\star}(v,\mathbf{x}_{B_{v,n}}) - \lambda_{\theta^\star}(v,\mathbf{x})\}\,\mathrm{d}u\,\mathrm{d}v.
\end{aligned} \quad \text{(A.8)}$$

[g] implies **[Φ]** which in turn yields $|1-e^{-\Phi_{\theta^\star}(v-u)}| \leqslant 1 + e^{-\Phi_{\theta^\star}(v-u)} \leqslant c$ since $\Phi_{\theta^\star}$ is bounded from below. On the other hand, for any $m \geqslant 2$, denoting $\tilde{\Phi}_{\theta^\star} =$

$\Phi_{\theta^\star} - \varepsilon|\theta^\star_m g_m|$ for some $\varepsilon > 0$, we have $|g_m|e^{-\Phi_{\theta^\star}} = |g_m|e^{-\varepsilon|\theta^\star_m g_m|}e^{-\tilde{\Phi}_{\theta^\star}} \leqslant c$ since $x \mapsto xe^{-\kappa x}$ is bounded on $[0,\infty)$ for any $\kappa > 0$ and $\tilde{\Phi}_{\theta^\star}$ satisfies **[Φ]** as seen in the proof of Lemma A.2. This proves that for any u, v, $\|e^{-\Phi_{\theta^\star}(v-u)}g(v-u)^\top\|$ is bounded. Moreover, from (A.6), we know that if $|k-j| > 2r_0$, then $|1-e^{-\Phi_{\theta^\star}(v-u)}| \leqslant c|k-j|^{-\gamma_2}$ and similarly $\|e^{-\Phi_{\theta^\star}(v-u)}g(v-u)^\top\| \leqslant c|k-j|^{-\gamma_2}$. We deduce that for any $u \in \Delta_{n,j}$, any $v \in \Delta_{n,k}$ and any j, k, $|1-e^{-\Phi_{\theta^\star}(v-u)}| \leqslant c(1+|k-j|)^{-\gamma_2}$ and $\|e^{-\Phi_{\theta^\star}(v-u)}g(v-u)^\top\| \leqslant c(1+|k-j|)^{-\gamma_2}$. Plugging these inequalities in (A.8) and applying Lemmas A.1-A.2 to the remaining terms shows that for any j, k

$$\|\operatorname{E}(s'_{\Delta_{n,j}} I_k^\top)\| \leqslant \frac{c}{\alpha_n^{\gamma'}(1+|k-j|^{\gamma_2})}. \tag{A.9}$$

The same inequality obviously holds for $\|\operatorname{E}(s'_{\Delta_{n,k}} I_j^\top)\|$. For the two last terms in the right hand side of (A.7), namely

$$\operatorname{E}(I_j I_k^\top) = \operatorname{E}\int_{\Delta_{n,j}}\int_{\Delta_{n,k}} \pi_n(u,\mathbf{X})\pi_n(v,\mathbf{X})^\top \,\mathrm{d}u\,\mathrm{d}v$$

and, after application of the GNZ formula,

$$\operatorname{E}(s_{\Delta_{n,j}})\operatorname{E}(s_{\Delta_{n,k}})^\top = \int_{\Delta_{n,j}}\int_{\Delta_{n,k}} \operatorname{E}\pi_n(u,\mathbf{X})\operatorname{E}\pi_n(v,\mathbf{X})^\top \,\mathrm{d}u\,\mathrm{d}v,$$

we deduce from Lemmas A.1-A.2 that their norm is bounded by $\alpha_n^{-2\gamma'}$ for any j, k, up to a positive constant. The latter and (A.9) prove the first inequality in (iv) from which the three remaining statements are easily deduced. □

References

A. Baddeley and R. Turner. Practical maximum pseudolikelihood for spatial point patterns (with discussion). *Australian and New Zealand Journal of Statistics*, 42 (3):283–322, 2000.

A. Baddeley, J.-F. Coeurjolly, E. Rubak, and R. Waagepetersen. Logistic regression for spatial gibbs point processes. *Biometrika*, 101(2):377–392, 2014.

E. Bertin, J.-M. Billiot, and R. Drouilhet. Existence of "nearest-neighbour" spatial Gibbs models. *Advances in Applied Probability*, 31:895–909, 1999.

J. Besag. Some methods of statistical analysis for spatial data. *Bulletin of the International Statistical Institute*, 47:77–92, 1977.

J.-M. Billiot, J.-F. Coeurjolly, and R. Drouilhet. Maximum pseudolikelihood estimator for exponential family models of marked Gibbs point processes. *Electronic Journal of Statistics*, 2:234–264, 2008.

E. Bolthausen. On the central limit theorem for stationary mixing random fields. *The Annals of Probability*, 10:1047–1050, 1982.

S. N. Chiu, D. Stoyan, W. S. Kendall, and J. Mecke. *Stochastic geometry and its applications*. John Wiley & Sons, third edition, 2013.

J.-F. Coeurjolly and R. Drouilhet. Asymptotic properties of the maximum pseudo-likelihood estimator for stationary Gibbs point processes including the Lennard-Jones model. *Electronic Journal of Statistics*, 4:677–706, 2010.

J.-F. Coeurjolly and F. Lavancier. Residuals for stationary marked Gibbs point processes. *Journal of the Royal Statistical Society, Series B*, 75(2):247–276, 2013.

J.-F. Coeurjolly and J. Møller. Variational approach for spatial point process intensity estimation. *Bernoulli*, 20(3):1097–1125, 2014.

J.-F. Coeurjolly and E. Rubak. Fast covariance estimation for innovations computed from a spatial Gibbs point process. *Scandinavian Journal of Statistics*, 40(4):669–684, 2013.

F. Comets and M. Janžura. A central limit theorem for conditionally centred random fields with an application to Markov fields. *Journal of applied probability*, 35(3): 608–621, 1998.

D. J. Daley and D. Vere-Jones. *An Introduction to the Theory of Point Processes. Volume I: Elementary Theory and Methods*. Springer-Verlag, New York, second edition, 2003.

L. Decreusefond and I. Flint. Moment formulae for general point processes. *Journal of Functional Analysis*, 267(2):452–476, 2014.

J. Dedecker. A central limit theorem for stationary random fields. *Probability Theory and Related Fields*, 110(3):397–426, 1998.

D. Dereudre and F. Lavancier. Campbell equilibrium equation and pseudo-likelihood estimation for non-hereditary Gibbs point processes. *Bernoulli*, 15(4):1368–1396, 2009.

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

Hans-Otto Georgii. *Gibbs measures and phase transitions*. Walter de Gruyter, Berlin, 1988.

X. Guyon. *Random fields on a network. Modeling, Statistics and Applications*. Springer Verlag, New York, 1995.

X. Guyon and H. R. Künsch. Asymptotic comparison of estimators in the ising model. In *Stochastic Models, Statistical Methods, and Algorithms in Image Analysis*, volume 74 of *Lecture Notes in Statistics*, pages 177–198. Springer, New York, 1992.

L. Heinrich. Mixing properties of Gibbsian point processes and asymptotic normality of Takacs-Fiksel estimates. *preprint*, 1992.

F. Huang and Y. Ogata. Improvements of the maximum pseudo-likelihood estimators in various spatial statistical models. *Journal of Computational and Graphical Statistics*, 8(3):510–530, 1999.

J. Illian, A. Penttinen, H. Stoyan, and D. Stoyan. *Statistical analysis and modelling of spatial point patterns*. Statistics in Practice. Wiley, 2008.

J. L. Jensen. Asymptotic normality of estimates in spatial point processes. *Scandinavian Journal of Statistics*, 20:97–109, 1993.

J. L. Jensen and H. R. Künsch. On asymptotic normality of pseudolikelihood estimates of pairwise interaction processes. *Annals of the Institute of Statistical Mathematics*, 46:475–486, 1994.

J. L. Jensen and J. Møller. Pseudolikelihood for exponential family models of spatial point processes. *Annals of Applied Probability*, 1:445–461, 1991.

J. E. Lennard-Jones. On the determination of molecular fields. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, pages 463–477, 1924.

M.N.M. Van Lieshout. *Markov point processes and their applications*. World Scientific, 2000.

S. Mase. Consistency of the maximum pseudo-likelihood estimator of continuous state space Gibbs processes. *Annals of Applied Probability*, 5:603–612, 1995.

J. Mecke. Eine charakteristische eigenschaft der doppelt stochastischen poissonschen prozesse. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 11(1): 74–81, 1968.

J. Møller and R. P. Waagepetersen. *Statistical inference and simulation for spatial point processes*. Chapman and Hall/CRC, Boca Raton, 2004.

X. Nguyen and H. Zessin. Ergodic theorems for spatial processes. *Zeitschrift für Wahrscheinlichkeitstheorie verwiete Gebiete*, 48:133–158, 1979a.

X. Nguyen and H. Zessin. Integral and differential characterizations Gibbs processes. *Mathematische Nachrichten*, 88(1):105–115, 1979b.

Y. Ogata and M. Tanemura. Estimation of interaction potentials of spatial point patterns through the maximum likelihood procedure. *Annals of the Institute of Statistical Mathematics*, 33:315–338, 1981.

C. Preston. *Random fields*. Number 534 in Lecture Notes in Mathematics. Springer-Verlag, Berlin, 1976.

D. Ruelle. *Statistical Mechanics*. Benjamin, New York-Amsterdam, 1969.

D. Ruelle. Superstable interactions in classical statistical mechanics. *Communications in Mathematical Physics*, 1970.

I.M. Slivnyak. Some properties of stationary flows of homogeneous random events. *Theory of Probability & Its Applications*, 7(3):336–341, 1962.