



**HAL**  
open science

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

Jean-François Coeurjolly, Frédéric Lavancier

► **To cite this version:**

Jean-François Coeurjolly, Frédéric Lavancier. Parametric estimation of pairwise Gibbs point processes with infinite range interaction. *Bernoulli*, 2017, 23 (2), pp.1299-1334. 10.3150/15-BEJ779 . hal-01092225v3

**HAL Id: hal-01092225**

**<https://hal.science/hal-01092225v3>**

Submitted on 1 Oct 2015

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

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

Jean-François Coeurjolly<sup>1</sup> and Frédéric Lavancier<sup>2,3</sup>

<sup>1</sup>Univ. Grenoble Alpes, France, [Jean-Francois.Coeurjolly@upmf-grenoble.fr](mailto:Jean-Francois.Coeurjolly@upmf-grenoble.fr).

<sup>2</sup>Laboratoire de Mathématiques Jean Leray, Nantes University, France,

[Frederic.Lavancier@univ-nantes.fr](mailto:Frederic.Lavancier@univ-nantes.fr).

<sup>3</sup>Inria, Centre Rennes Bretagne Atlantique, France

## 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, provided the pairwise interaction function tends sufficiently fast to zero. 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, 2000b) 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 border correction is often used to make the pseudolikelihood score unbiased. An obvious solution is to compute the pseudolikelihood on the eroded set  $W \ominus R$ , and in view of  $(W \ominus R) \oplus R \subseteq W$  (see Chiu et al. (2013)) the observation of the point process on  $W$  is sufficient for the computation. From a theoretical point of view, 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. We propose in Section 2 a family of contrast functions that involve an eroded set, following the previous border correction, and a truncated range of interaction. The details are exposed in Section 2. However these contrast functions 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 our pseudolikelihood estimator in Proposition 3.1. Our result is valid for a large family of pairwise Gibbs models, namely the class of Ruelle superstable and lower regular models. 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 establish in Appendix A a new central limit theorem for triangular arrays of *almost* conditionally centered random fields. This allows us to derive in Theorem 3.3 the asymptotic normality of our pseudolikelihood estimator. Assume the potential decreases with a hyperbolic rate with exponent  $\gamma_2$ . While  $\gamma_2 > d$  is (in general) required to ensure the existence of a Gibbs measure and the consistency of the pseudolikelihood estimator, we require the condition  $\gamma_2 > 2d$  to ensure a central limit theorem. The asymptotic normality when  $d < \gamma_2 < 2d$  remains an open question. Nonetheless our result includes the important example of the Lennard-Jones model in dimension  $d = 2$ , for which  $\gamma_2 = 6$ . Proposition 3.4 discusses similar asymptotic results for the logistic regression estimator.

The remainder of this paper is organized as follows. 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. We derive in Section 3 the asymptotic properties of our estimators. Section 4 contains a simulation study concerning the estimation of the Lennard-Jones potential, where some recommendations are derived for the practical choice of tuning parameters in the pseudolikelihood contrast function. Appendix A contains our main technical tool, namely a central limit theorem for almost conditionally centered random fields, and Appendix B gathers auxiliary lemmas.

## 2 Background and statistical methodology

### 2.1 Notation

We write  $\Lambda \Subset \mathbb{R}^d$  for a bounded set  $\Lambda$  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  $\mathbf{x}_{\Lambda^c} = \mathbf{x} \setminus \mathbf{x}_\Lambda$  and denote by  $\Omega_0$  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  $\|M\| = \text{tr}(M^\top M)^{1/2}$  its Frobenius norm where  $\text{tr}$  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  $\Omega_0$ . The reference distribution on  $\Omega_0$  is the homogeneous Poisson point process with intensity  $\beta > 0$ , denoted by  $\pi^\beta$ . For  $\Lambda \Subset \mathbb{R}^d$ , we write  $\pi_\Lambda^\beta$  for the restriction of  $\pi^\beta$  to  $\Lambda$ . 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  $\Delta_i$  be the unit cube centered at  $i \in \mathbb{Z}^d$ . We consider the following space of tempered configurations.

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

From the ergodic theorem (see Guyon (1995)), any second order stationary measure on  $\Omega_0$  is supported on  $\Omega_T$ . We denote by  $\Phi : \mathbb{R}^d \rightarrow \mathbb{R} \cup \{+\infty\}$  a pair potential function, to which we associate the pairwise energy function  $H_\Lambda : \Omega_T \rightarrow \mathbb{R} \cup \{+\infty\}$ , indexed by Borel sets  $\Lambda \Subset \mathbb{R}^d$  and defined by

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

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  $\Phi$  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  $\pi_\Lambda^\beta$  with the density  $\exp\{-H_\Lambda(\mathbf{x})\}/Z_\Lambda(\mathbf{x}_{\Lambda^c})$ , where  $Z_\Lambda(\mathbf{x}_{\Lambda^c}) = \int_\Omega \exp\{-H_\Lambda(\mathbf{x}_\Lambda \cup \mathbf{x}_{\Lambda^c})\} \pi_\Lambda^\beta(d\mathbf{x}_{\Lambda^c})$  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  $\lambda : \mathbb{R}^d \times \Omega \rightarrow \mathbb{R}_+$  defined for any  $\Lambda \ni u$  by

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

This quantity does not depend on  $\Lambda$ , 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  $\Omega$  is a Gibbs measure with activity parameter  $\beta > 0$  for the pair potential function  $\Phi$  if for any measurable function  $f : \Omega \times \mathbb{R}^d \rightarrow \mathbb{R}$  such that the following expectations are finite,*

$$\mathbb{E} \left\{ \sum_{u \in \mathbf{X}} f(u, \mathbf{X} \setminus u) \right\} = \mathbb{E} \left\{ \int f(u, \mathbf{X}) \lambda(u, \mathbf{X}) du \right\} \quad (2.3)$$

where  $\mathbb{E}$  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  $\Phi$  and Papangelou conditional intensity  $\lambda$ . Then for any measurable function  $f : \Omega \times \mathbb{R}^d \rightarrow \mathbb{R}$  and for any  $\Lambda \in \mathbb{R}^d$  such that the following expectations are finite*

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

where  $\mathbb{E}$  denotes the expectation with respect to  $P$ .

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

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

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

$$\mathbb{E}_{\pi_\Lambda^\beta} \left\{ \sum_{u \in \mathbf{X}} h(u, \mathbf{X} \setminus u) \right\} = \beta \mathbb{E}_{\pi_\Lambda^\beta} \left\{ \int_{\mathbb{R}^d} h(u, \mathbf{X}) du \right\}.$$

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

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

□

The existence of a Gibbs measure  $P$  satisfying the above definition and characterization is a difficult question. Sufficient conditions on the pair potential  $\Phi$  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  $\Phi$ , 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).

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

Examples of potentials satisfying **[ $\Phi$ ]** 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  $\Theta$  is an open bounded set of  $\mathbb{R}^p$ . We stress this assumption by writing  $\lambda_\theta$  instead of  $\lambda$  and  $\Phi_\theta$  instead of  $\Phi$ . For brevity, assumption **[ $\Phi$ ]** now means that  $\Phi_\theta$  fulfills this assumption for any  $\theta \in \Theta$ .

Assume that we observe the point process  $\mathbf{X}$  in  $W_n$  where  $(W_n)_{n \geq 1}$  is a sequence of bounded domains which converges to  $\mathbb{R}^d$  as  $n \rightarrow \infty$ . 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

$$\text{LPL}_{W_n}(\mathbf{X}; \theta) = \sum_{u \in \mathbf{X}_{W_n}} \log \lambda_\theta(u, \mathbf{X} \setminus u) - \int_{W_n} \lambda_\theta(u, \mathbf{X}) \, du \quad (2.5)$$

$$\text{LRL}_{W_n}(\mathbf{X}; \theta) = \sum_{u \in \mathbf{X}_{W_n}} \log \frac{\lambda_\theta(u, \mathbf{X} \setminus u)}{\lambda_\theta(u, \mathbf{X} \setminus u) + \rho} - \int_{W_n} \rho \log \frac{\lambda_\theta(u, \mathbf{X}) + \rho}{\rho} \, du \quad (2.6)$$

where  $\rho$  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  $\Phi_\theta$  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  $\Lambda \subseteq \mathbb{R}^d$  and some  $\kappa \geq 0$  the notation  $\Lambda \ominus \kappa$  stands for the domain  $\Lambda$  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  $\text{LPL}_{W_n \ominus R}(\mathbf{X}; \theta)$  (or  $\text{LRL}_{W_n \ominus R}(\mathbf{X}; \theta)$ ) with respect to  $\theta$ . The key-ingredient is that both scores constitute unbiased estimating functions, since by application of the GNZ formula (2.3) their expectation vanishes when  $\theta$  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) that depend on two sequences of positive real numbers  $(\alpha_n)$  and  $(R_n)$

$$\widetilde{\text{LPL}}_{W_n \ominus \alpha_n, R_n}(\mathbf{X}; \theta) = \sum_{u \in \mathbf{X}_{W_n \ominus \alpha_n}} \log \lambda_\theta(u, \mathbf{X}_{u, R_n} \setminus u) - \int_{W_n \ominus \alpha_n} \lambda_\theta(u, \mathbf{X}_{u, R_n}) du \quad (2.7)$$

$$\widetilde{\text{LRL}}_{W_n \ominus \alpha_n, R_n}(\mathbf{X}; \theta) = \sum_{u \in \mathbf{X}_{W_n \ominus \alpha_n}} \log \frac{\lambda_\theta(u, \mathbf{X}_{u, R_n} \setminus u)}{\lambda_\theta(u, \mathbf{X}_{u, R_n} \setminus u) + \rho} - \int_{W_n \ominus \alpha_n} \rho \log \frac{\lambda_\theta(u, \mathbf{X}_{u, R_n}) + \rho}{\rho} du \quad (2.8)$$

where we denote  $\mathbf{X}_{u, R_n} = \mathbf{X}_{B(u, R_n) \cap W_n}$ . These expressions are computable from the single observation of  $\mathbf{X}$  on  $W_n$ , provided the integrals are approximated as usual by numerical scheme or by Monte-Carlo. Since they depend on two sequences  $\alpha_n$  and  $R_n$ , (2.7) and (2.8) actually form a general family of contrast functions, important particular cases being the choices

- $\alpha_n = R_n$ , which agrees with the classical border correction for finite range interaction models with range  $R$  taking  $R_n = R$ ;
- $R_n = \infty$ , accounting for the maximal possible range of interaction;
- $R_n = \infty$  and  $\alpha_n = 0$ , which is a particular case of the previous choice where in addition no erosion is considered.

We study in the next section the asymptotic properties of estimators derived from (2.7) and (2.8) for a wide class of sequences  $\alpha_n$  and  $R_n$ , and based on a simulation study in Section 4, we give some recommendations for the choice of these sequences in practice. From a theoretical point view, these contrast functions introduce new challenges since the gradients of  $\widetilde{\text{LPL}}_{W_n \ominus \alpha_n, R_n}(\mathbf{X}; \theta)$  and  $\widetilde{\text{LRL}}_{W_n \ominus \alpha_n, R_n}(\mathbf{X}; \theta)$  are no longer unbiased estimating equations in the infinite range case. To overcome this difficulty we prove a new central limit theorem in Appendix A for almost conditionally centered random fields.

### 3 Asymptotic properties

We present asymptotic properties of the maximum pseudolikelihood estimate, derived from (2.7), for infinite range Gibbs point process. Similar results for the max-



imum logistic regression derived from (2.8) are presented at the end of this section without proof. We assume the window of observation expands to  $\mathbb{R}^d$  as follows.

**[W<sub>n</sub>]** ( $W_n$ ) is an increasing sequence of convex compact sets, such that  $W_n \rightarrow \mathbb{R}^d$  as  $n \rightarrow \infty$ .

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

$$\lambda_\theta(u, \mathbf{x}) = \beta e^{-\sum_{v \in \mathbf{x}} \Phi_\theta(v-u)} = e^{-\theta^\top t(u, \mathbf{x})} \quad (3.1)$$

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

$$t_m(u, \mathbf{x}) = \sum_{v \in \mathbf{x}} g_m(v - u), \quad m = 2, \dots, p. \quad (3.2)$$

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  $g = (0, g_2, \dots, g_p)^\top$ . We make the following assumption on  $g$ .

**[g]** For all  $m \geq 2$ ,  $g_m$  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  $\forall \theta \in \Theta$ ,  $\theta_2 g_2(x) \geq c_g \|x\|^{-\gamma_1}$
- (ii)  $\forall m \geq 3$ ,  $g_m(x) = o(\|x\|^{-\gamma_1})$  as  $\|x\| \rightarrow 0$
- (iii)  $\forall m \geq 2$  and  $\forall \|x\| \geq r_0$ ,  $|g_m(x)| \leq c \|x\|^{-\gamma_2}$ .

Since  $\Theta$  is bounded, **[g]** implies **[Φ]** which yields that for any  $\theta \in \Theta$  there exists a Gibbs measure  $P_\theta$ . Assumption **[g]** allows us to specify which function  $g_m$  is responsible for the behavior at the origin of  $\Phi_\theta$ , 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,  $\theta^*$  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  $P_{\theta^*}$  on  $W_n$ .

For exponential family models (3.1) the score function of the log-pseudolikelihood defined by (2.7) writes  $s_{W_n \ominus \alpha_n, R_n}(\mathbf{X}; \theta)$  where for any  $\Delta \subset \mathbb{R}^d$

$$s_{\Delta, R_n}(\mathbf{X}; \theta) = \int_{\Delta} t(u, \mathbf{X}_{u, R_n}) \lambda_\theta(u, \mathbf{X}_{u, R_n}) du - \sum_{u \in \mathbf{X}_\Delta} t(u, \mathbf{X}_{u, R_n} \setminus u). \quad (3.3)$$

Our first result establishes the strong consistency of the maximum pseudolikelihood based on (2.7) for infinite range Gibbs point processes and for a wide class of sequences  $(\alpha_n, R_n)$ . 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 3.1.** *Assume that  $[\mathbf{W}_n]$  and  $[\mathbf{g}]$  hold. Then for any  $\mathbf{x} \in \Omega$  the function  $\theta \rightarrow -\widetilde{\text{LPL}}_{W_n \ominus \alpha_n, R_n}(\mathbf{x}; \theta)$  is a convex function with Hessian matrix given by*

$$\begin{aligned} -\frac{d}{d\theta d\theta^\top} \widetilde{\text{LPL}}_{W_n \ominus \alpha_n, R_n}(\mathbf{x}; \theta) &= -\frac{d}{d\theta^\top} s_{W_n \ominus \alpha_n, R_n}(\mathbf{x}; \theta) \\ &= \int_{W_n \ominus \alpha_n} t(u, \mathbf{x}_{u, R_n}) t(u, \mathbf{x}_{u, R_n})^\top \lambda_\theta(u, \mathbf{x}_{u, R_n}) du. \end{aligned} \quad (3.4)$$

Moreover if  $\alpha_n |W_n|^{-1/d} \rightarrow 0$  and  $R_n \rightarrow \infty$  as  $n \rightarrow \infty$ , and if for any  $y \in \mathbb{R}^p \setminus \{0\}$

$$P \{y^\top t(0, \mathbf{X}) \neq 0\} > 0, \quad (3.5)$$

then the maximum pseudolikelihood estimator

$$\hat{\theta}_{\widetilde{\text{LPL}}} = \operatorname{argmax}_{\theta \in \Theta} \widetilde{\text{LPL}}_{W_n \ominus \alpha_n, R_n}(\mathbf{X}; \theta)$$

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

*Proof.* By  $[\mathbf{W}_n]$  and the basic assumption on  $\alpha_n$ , we are ensured that  $W_n \ominus \alpha_n$  is a sequence of regular bounded domains of  $\mathbb{R}^d$  and that  $|W_n \ominus \alpha_n| \rightarrow \infty$  as  $n \rightarrow \infty$ . 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  $P_{\theta^*}$  is ergodic. Since  $\Theta$  is an open bounded set, and by convexity of  $\theta \rightarrow -\widetilde{\text{LPL}}_{W_n \ominus \alpha_n, R_n}(\mathbf{x}; \theta)$ , then from Guyon (1995, Theorem 3.4.4) we only need to prove that  $K_n(\theta, \theta^*) = |W_n \ominus \alpha_n|^{-1} \left\{ \widetilde{\text{LPL}}_{W_n \ominus \alpha_n, R_n}(\mathbf{X}; \theta^*) - \widetilde{\text{LPL}}_{W_n \ominus \alpha_n, R_n}(\mathbf{X}; \theta) \right\} \rightarrow K(\theta, \theta^*)$  almost surely as  $n \rightarrow \infty$ , where  $\theta \rightarrow K(\theta, \theta^*)$  is a nonnegative function which vanishes at  $\theta = \theta^*$  only. We decompose  $K_n(\theta, \theta^*)$  as the sum of the three terms  $T_1 + T_2(\theta^*) - T_2(\theta)$  where for any  $\theta \in \Theta$

$$\begin{aligned} T_1 &= |W_n \ominus \alpha_n|^{-1} \left\{ \text{LPL}_{W_n \ominus \alpha_n}(\mathbf{X}; \theta^*) - \text{LPL}_{W_n \ominus \alpha_n}(\mathbf{X}; \theta) \right\} \\ T_2(\theta) &= |W_n \ominus \alpha_n|^{-1} \left\{ \widetilde{\text{LPL}}_{W_n \ominus \alpha_n, R_n}(\mathbf{X}; \theta) - \text{LPL}_{W_n \ominus \alpha_n}(\mathbf{X}; \theta) \right\}. \end{aligned}$$

Lemma B.2 shows in particular that  $\lambda_\theta(0, \mathbf{X})$  and  $|\theta^\top t(0, \mathbf{X})| \lambda_{\theta^*}(0, \mathbf{X})$  have finite expectation under  $P_{\theta^*}$ . 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  $T_1 \rightarrow K(\theta, \theta^*)$  almost surely as  $n \rightarrow \infty$  where

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

which is a nonnegative function that vanishes at  $\theta = \theta^*$  only, under the identifiability condition (3.5). So the rest of the proof consists in proving that  $T_2(\theta) \rightarrow 0$  almost surely for any  $\theta \in \Theta$ . We have  $T_2(\theta) = T'_1 + T'_2$  where

$$\begin{aligned} T'_1 &= |W_n \ominus \alpha_n|^{-1} \sum_{u \in \mathbf{X}_{W_n \ominus \alpha_n}} \theta^\top \{t(u, \mathbf{X}_{u, R_n} \setminus u) - t(u, \mathbf{X} \setminus u)\} \\ T'_2 &= |W_n \ominus \alpha_n|^{-1} \int_{W_n \ominus \alpha_n} \{\lambda_\theta(u, \mathbf{X}) - \lambda_\theta(u, \mathbf{X}_{u, R_n})\} du. \end{aligned}$$

Let us first look at  $T'_1$ . By boundedness of  $\Theta$ , it is sufficient to prove that  $|W_n \ominus \alpha_n|^{-1} \sum_{u \in \mathbf{X}_{W_n \ominus \alpha_n}} \{t_m(u, \mathbf{X}_{u, R_n} \setminus u) - t_m(u, \mathbf{X} \setminus u)\}$  tends to 0 almost surely. Let  $\delta_n$  be a sequence of real numbers such that  $\delta_n \rightarrow \infty$  and  $\delta_n |W_n|^{-1/d} \rightarrow 0$  as  $n \rightarrow \infty$ . For brevity, let  $\mathbf{X}_{(u, R_n)^c} := \mathbf{X}_{\mathbb{R}^d \setminus \{W_n \cap B(u, R_n)\}}$ .

$$\begin{aligned}
& \sum_{u \in \mathbf{X}_{W_n \ominus \alpha_n}} \{t_m(u, \mathbf{X}_{u, R_n} \setminus u) - t_m(u, \mathbf{X} \setminus u)\} \\
&= \sum_{u \in \mathbf{X}_{W_n \ominus \alpha_n}} \sum_{v \in \mathbf{X}_{(u, R_n)^c}} g_m(v - u) \\
&= \sum_{u \in \mathbf{X}_{W_n \ominus (\alpha_n + \delta_n)}} \sum_{v \in \mathbf{X}_{(u, R_n)^c}} g_m(v - u) + \sum_{u \in \mathbf{X}_{W_n \ominus \alpha_n \setminus W_n \ominus (\alpha_n + \delta_n)}} \sum_{v \in \mathbf{X}_{(u, R_n)^c}} g_m(v - u) \quad (3.6)
\end{aligned}$$

In the first sum above,  $\|u - v\| \geq R_n \wedge (\alpha_n + \delta_n)$  and using the same arguments and the same notation as in (ii) of Lemma B.1 we get that the absolute value of this sum is lower than  $c(R_n \wedge (\alpha_n + \delta_n))^{-\gamma'} \sum_{u \in \mathbf{X}_{W_n \ominus (\alpha_n + \delta_n)}} H(u, \mathbf{X} \setminus u)$  for some  $\gamma' > 0$ . Hence

$$\begin{aligned}
& |W_n \ominus \alpha_n|^{-1} \left| \sum_{u \in \mathbf{X}_{W_n \ominus (\alpha_n + \delta_n)}} \sum_{v \in \mathbf{X}_{(u, R_n)^c}} g_m(v - u) \right| \\
&\leq c(R_n \wedge (\alpha_n + \delta_n))^{-\gamma'} \frac{|W_n \ominus (\alpha_n + \delta_n)|}{|W_n \ominus \alpha_n|} |W_n \ominus (\alpha_n + \delta_n)|^{-1} \sum_{u \in \mathbf{X}_{W_n \ominus (\alpha_n + \delta_n)}} H(u, \mathbf{X} \setminus u).
\end{aligned}$$

By Lemma B.2, the random variable  $|H(0, \mathbf{X})| \lambda_{\theta^*}(0, \mathbf{X})$  has finite expectation under  $P_{\theta^*}$ . Moreover our assumptions ensure that  $W_n \ominus (\alpha_n + \delta_n)$  is a sequence of regular bounded domains of  $\mathbb{R}^d$  with  $|W_n \ominus (\alpha_n + \delta_n)| \rightarrow \infty$  as  $n \rightarrow \infty$ . So by the ergodic theorem  $|W_n \ominus (\alpha_n + \delta_n)|^{-1} \sum_{u \in \mathbf{X}_{W_n \ominus (\alpha_n + \delta_n)}} H(u, \mathbf{X} \setminus u) \rightarrow \mathbb{E} \{H(0, \mathbf{X}) \lambda_{\theta^*}(0, \mathbf{X})\}$  almost surely whereby

$$|W_n \ominus \alpha_n|^{-1} \left| \sum_{u \in \mathbf{X}_{W_n \ominus (\alpha_n + \delta_n)}} \sum_{v \in \mathbf{X}_{(u, R_n)^c}} g_m(v - u) \right| \rightarrow 0 \quad (3.7)$$

almost surely. For the second sum in (3.6), using the notation  $|t_m|(u, \mathbf{X}) = \sum_{v \in \mathbf{X}} |g_m(v - u)|$  we have

$$\begin{aligned}
& |W_n \ominus \alpha_n|^{-1} \left| \sum_{u \in \mathbf{X}_{W_n \ominus \alpha_n \setminus W_n \ominus (\alpha_n + \delta_n)}} \sum_{v \in \mathbf{X}_{(u, R_n)^c}} g_m(v - u) \right| \\
&\leq |W_n \ominus \alpha_n|^{-1} \sum_{u \in \mathbf{X}_{W_n \ominus \alpha_n \setminus W_n \ominus (\alpha_n + \delta_n)}} |t_m|(u, \mathbf{X}) \\
&= |W_n \ominus \alpha_n|^{-1} \sum_{u \in \mathbf{X}_{W_n \ominus \alpha_n}} |t_m|(u, \mathbf{X}) - |W_n \ominus \alpha_n|^{-1} \sum_{u \in \mathbf{X}_{W_n \ominus (\alpha_n + \delta_n)}} |t_m|(u, \mathbf{X})
\end{aligned}$$

which tends to 0 almost surely by application of the ergodic theorem, Lemma B.2 and since  $|W_n \ominus (\alpha_n + \delta_n)| / |W_n \ominus \alpha_n| \rightarrow 1$  as  $n \rightarrow \infty$ . Combining this result with (3.7) in (3.6) shows that  $T'_1 \rightarrow 0$  almost surely.

Concerning the remaining term  $T'_2$ , we have

$$T'_2 = |W_n \ominus \alpha_n|^{-1} \int_{W_n \ominus \alpha_n} \lambda_\theta(u, \mathbf{X}) \left\{ 1 - e^{-\theta^\top \sum_{v \in \mathbf{X}_{(u, R_n)^c}} g(v-u)} \right\} du.$$

We can use exactly the same decomposition as in (3.6) by introducing  $\delta_n$ , then use Lemmas B.1 and B.2 to apply the ergodic theorem, leading to  $T'_2 \rightarrow 0$  almost surely. These details are omitted. Hence  $T_2(\theta) \rightarrow 0$  almost surely for any  $\theta \in \Theta$  and the proof of Proposition 3.1 is completed.  $\square$

As a preliminary result towards the asymptotic normality of  $\widehat{\theta}_{\text{LPL}}$ , we state in the next lemma general conditions on the sequences  $\alpha_n$  and  $R_n$  leading to the equivalence in probability of the score functions, up to  $|W_n|^{1/2}$ .

**Lemma 3.2.** *Assume that  $[\mathbf{W}_n]$  and  $[\mathbf{g}]$  hold. Let  $\mathcal{D}$  be the set of all sequences  $(\alpha_n, R_n)$  such that  $\alpha_n |W_n|^{-1/d} \rightarrow 0$  and such that there exists  $0 < \gamma' < \gamma_2 - d$  such that  $\alpha_n^{-\gamma'} |W_n|^{1/2} \rightarrow 0$  and  $R_n^{-\gamma'} |W_n|^{1/2} \rightarrow 0$ . Then, for any  $(\alpha_n, R_n) \in \mathcal{D}$  and  $(\alpha'_n, R'_n) \in \mathcal{D}$  we have*

$$s_{W_n \ominus \alpha_n, R_n}(\mathbf{X}; \theta^*) - s_{W_n \ominus \alpha'_n, R'_n}(\mathbf{X}; \theta^*) = o_P(|W_n|^{1/2}). \quad (3.8)$$

*Proof.* For  $\Delta \in \mathbb{R}^d$ , we denote for short  $s_{\Delta, R_n} := s_{\Delta, R_n}(\mathbf{X}; \theta^*)$  and set

$$s'_{\Delta, R_n} = \int_{\Delta} t(u, \mathbf{X}_{u, R_n}) \lambda_{\theta^*}(u, \mathbf{X}) du - \sum_{u \in \mathbf{X}_{\Delta}} t(u, \mathbf{X}_{u, R_n} \setminus u). \quad (3.9)$$

We prove below that for any  $(\alpha_n, R_n) \in \mathcal{D}$ ,  $s_{W_n \ominus \alpha_n, R_n} - s'_{W_n, \infty} = o_P(|W_n|^{1/2})$ , whereby (3.8) is an immediate consequence. We have

$$s_{W_n \ominus \alpha_n, R_n} - s'_{W_n, \infty} = A + B + C \quad (3.10)$$

with  $A = s_{W_n \ominus \alpha_n, R_n} - s_{W_n \ominus \alpha_n, \infty}$ ,  $B = s_{W_n \ominus \alpha_n, \infty} - s'_{W_n \ominus \alpha_n, \infty}$  and  $C = s'_{W_n \ominus \alpha_n, \infty} - s'_{W_n, \infty}$ . Let us prove that each of these three terms is  $o_P(|W_n|^{1/2})$ .

First, we have  $A = A_1 + A_2$  with

$$\begin{aligned} A_1 &= \int_{W_n \ominus \alpha_n} \{t(u, \mathbf{X}_{u, R_n}) \lambda_{\theta^*}(u, \mathbf{X}_{u, R_n}) - t(u, \mathbf{X}_{W_n}) \lambda_{\theta^*}(u, \mathbf{X}_{W_n})\} du, \\ A_2 &= \sum_{u \in \mathbf{X}_{W_n \ominus \alpha_n}} \{t(u, \mathbf{X}_{W_n} \setminus u) - t(u, \mathbf{X}_{u, R_n} \setminus u)\}. \end{aligned}$$

Both terms above are handled similarly and we give the details for  $A_2$  only. Denoting  $A_{2,m}$  the  $m$ -th coordinate of  $A_2$ , we obtain using the same arguments and the same notation as in Lemma B.1 (ii) that for any  $m$

$$|A_{2,m}| \leq \sum_{u \in \mathbf{X}_{W_n \ominus \alpha_n}} \sum_{v \in \mathbf{X}_{W_n \setminus B(u, R_n)}} |g_m(v-u)| \leq R_n^{-\gamma'} \sum_{u \in \mathbf{X}_{W_n}} H(u, \mathbf{X}).$$

Applying the GNZ formula and Lemma B.2, we deduce that  $E|A_{2,m}| = \mathcal{O}(R_n^{-\gamma'} |W_n|)$  showing that  $|W_n|^{-1/2} A_2 = o_P(1)$ . The same result holds for  $A_1$  by the arguments developed in Lemma B.1 (ii)-(iii).

Second,

$$B = \int_{W_n \ominus \alpha_n} t(u, \mathbf{X}_{W_n}) \{ \lambda_{\theta^*}(u, \mathbf{X}_{W_n}) - \lambda_{\theta^*}(u, \mathbf{X}) \} du.$$

Note that

$$| \lambda_{\theta}(u, \mathbf{x}) - \lambda_{\theta}(u, \mathbf{x}_{W_n}) | = \lambda_{\theta}(u, \mathbf{x}) \left| 1 - e^{\sum_{v \in \mathbf{x}_{W_n^c}} \Phi_{\theta}(v-u)} \right|$$

where in the last sum  $\|v - u\| \geq \alpha_n$  since  $u \in W_n \ominus \alpha_n$ . The same arguments as in the proof of Lemma B.1 (iii) thus apply and we get for all  $u \in W_n \ominus \alpha_n$ ,  $| \lambda_{\theta}(u, \mathbf{x}) - \lambda_{\theta}(u, \mathbf{x}_{W_n}) | \leq c \alpha_n^{-\gamma'} e^{cG(u, \mathbf{x})} \lambda_{\theta}(u, \mathbf{x}) H(u, \mathbf{x})$ . From Lemma B.1 (i) and Lemma B.2, we obtain

$$\mathbb{E} |B| = \mathcal{O}(\alpha_n^{-\gamma'} |W_n \ominus \alpha_n|) = \mathcal{O}(\alpha_n^{-\gamma'} |W_n|)$$

and thus  $|W_n|^{-1/2} B = o_P(1)$ .

Third,

$$\begin{aligned} C &= - \int_{W_n \setminus (W_n \ominus \alpha_n)} t(u, \mathbf{X}_{W_n}) \lambda_{\theta^*}(u, \mathbf{X}) du + \sum_{u \in \mathbf{X}_{W_n \setminus (W_n \ominus \alpha_n)}} t(u, \mathbf{X}_{W_n \setminus u}) \\ &= - \sum_{j \in \mathcal{J}_n} s'_{\Delta_{n,j}, \infty} \end{aligned}$$

where  $\Delta_j$  is the unit cube centered at  $j \in \mathbb{Z}^d$ ,  $\Delta_{n,j} = \Delta_j \cap (W_n \setminus (W_n \ominus \alpha_n))$  and  $\mathcal{J}_n \subset \mathbb{Z}^d$  is the set such that  $W_n \setminus (W_n \ominus \alpha_n) = \cup_{j \in \mathcal{J}_n} \Delta_{n,j}$ . We have

$$\text{Var}(C) = \sum_{j, k \in \mathcal{J}_n} \text{Cov}(s'_{\Delta_{n,j}, \infty}, s'_{\Delta_{n,k}, \infty}).$$

It is not difficult to check that all results of Lemma B.4 (i)-(iii) remain true if the ball  $B_{u,n} = B(u, \alpha_n)$  therein is replaced by  $W_n$ , or in other words  $s'_{\Delta}$  is replaced by  $s'_{\Delta, \infty}$ . Therefore, from (iii) of Lemma B.4

$$\| \text{Var}(C) \| = \mathcal{O}(|\mathcal{J}_n|) = \mathcal{O}(|W_n \setminus (W_n \ominus \alpha_n)|) = \mathcal{O}(|W_n|^{1-1/d} \alpha_n)$$

showing that  $|W_n|^{-1} \| \text{Var}(C) \| \rightarrow 0$ . Hereby  $|W_n|^{-1/2} C = o_P(1)$  and the proof is completed.  $\square$

The next result establishes the asymptotic normality of the score function associated to the modified pseudolikelihood  $\widetilde{\text{LPL}}_{W_n \ominus \alpha_n, R_n}(\mathbf{X}; \theta)$  at the true value of the parameter  $\theta = \theta^*$ , whenever  $(\alpha_n, R_n)$  belongs to the set  $\mathcal{D}$  introduced in Lemma 3.2. The proof relies on a new central limit theorem stated in Appendix A. As a consequence we deduce the asymptotic normality of  $\hat{\theta}_{\widetilde{\text{LPL}}}$ .

These results require the following notation: let  $\Sigma_{\infty}$  and  $U_{\infty}$  the  $(p, p)$  matrices

$$\begin{aligned} \Sigma_{\infty} &= \mathbb{E} \{ t(0, \mathbf{X}) t(0, \mathbf{X})^{\top} \lambda_{\theta^*}(0, \mathbf{X}) \} \\ &\quad + \int_{\mathbb{R}^d} \mathbb{E} \{ t(0, \mathbf{X}) t(v, \mathbf{X})^{\top} \lambda_{\theta^*}(0, \mathbf{X}) \lambda_{\theta^*}(v, \mathbf{X}) \} \{ 1 - e^{-\Phi_{\theta^*}(v)} \} dv \\ &\quad + \int_{\mathbb{R}^d} \mathbb{E} \{ \lambda_{\theta^*}(0, \mathbf{X}) \lambda_{\theta^*}(v, \mathbf{X}) \} g(v) g(v)^{\top} e^{-\Phi_{\theta^*}(v)} dv \end{aligned} \quad (3.11)$$

$$U_{\infty} = \mathbb{E} \{ t(0, \mathbf{X}) t(0, \mathbf{X})^{\top} \lambda_{\theta^*}(0, \mathbf{X}) \}. \quad (3.12)$$

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

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

We denote by  $\xrightarrow{d}$  the convergence in distribution.

**Theorem 3.3.** *Under the assumptions of Lemma 3.2 with  $\gamma_2 > 2d$  and the assumption that  $\Sigma_\infty$  is a positive definite matrix, then we have the two following convergences in distribution for any  $(\alpha_n, R_n) \in \mathcal{D}$ , as  $n \rightarrow \infty$ ,*

(i)

$$|W_n|^{-1/2} s_{W_n \ominus \alpha_n, R_n}(\mathbf{X}; \theta^*) \xrightarrow{d} \mathcal{N}(0, \Sigma_\infty),$$

(ii)

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

Some remarks on this theorem are in order. The condition  $\gamma_2 > 2d$  is clearly the most restrictive one. Nonetheless it includes the standard Lennard-Jones model in dimension  $d = 2$  for which  $\gamma_2 = 6$ . Under **[g]**, existence of the model is ensured if  $\gamma_2 > d$  but it remains an open problem to prove the asymptotic normality of the pseudolikelihood estimator when  $d < \gamma_2 < 2d$ . Concerning the set  $\mathcal{D}$  of possible sequences  $(\alpha_n, R_n)$ , it includes the natural choices  $(\alpha_n, R_n) = (\alpha_n, \alpha_n)$  and  $(\alpha_n, R_n) = (\alpha_n, \infty)$  discussed in Section 2.3, provided  $\alpha_n$  tends to infinity at a good rate. However,  $\mathcal{D}$  does not include the particular case  $(\alpha_n, R_n) = (0, \infty)$ , whereas this choice leads to a consistent estimator as proved in Proposition 3.1. In fact, when the erosion parameter  $\alpha_n$  does not tend to infinity, some edge effects occur due to the infinite range of the process. These edge effects are negligible with respect to  $|W_n|$  but not with respect to  $|W_n|^{1/2}$ . Finally, following Coeurjolly and Rubak (2013), it is possible to construct a fast estimator of the asymptotic covariance matrices  $\Sigma_\infty$  and  $U_\infty$ , but its asymptotic properties are out of the scope of the present paper.

*Proof.* (i) Since  $\gamma_2 > 2d$ , there exists  $\varepsilon > 0$  and  $a > 0$  such that

$$\frac{d}{2(\gamma_2 - d - \varepsilon)} < a < \frac{1}{2 + \varepsilon}. \quad (3.13)$$

For such  $\varepsilon, a$ , we let  $\gamma' = \gamma_2 - d - \varepsilon$  and consider the particular case  $R_n = \alpha_n = |W_n|^{a/d}$ . Then,  $\alpha_n |W_n|^{-1/d} \rightarrow 0$  and  $\alpha_n^{-\gamma'} |W_n|^{1/2} = R_n^{-\gamma'} |W_n|^{1/2} = |W_n|^{\frac{d/2 - a(\gamma_2 - d - \varepsilon)}{d}} \rightarrow 0$ . This particular choice of  $(\alpha_n, R_n)$  thus belongs to  $\mathcal{D}$ . From Lemma 3.2, if we prove (i) for this choice, then the same convergence holds true for all sequences  $(\alpha_n, R_n)$  in  $\mathcal{D}$ , completing the proof of (i). Henceforth we let  $R_n = \alpha_n = |W_n|^{a/d}$  where  $a$  is such that (3.13) holds.

Denoting by  $\Delta_j$  the unit cube centered at  $j \in \mathbb{Z}^d$ , we let  $\Delta_{n,j} = \Delta_j \cap (W_n \ominus \alpha_n)$  and  $\mathcal{I}_n \subset \mathbb{Z}^d$  the set such that  $W_n \ominus \alpha_n = \cup_{j \in \mathcal{I}_n} \Delta_{n,j}$ . At several places in the proof the sequence  $\rho_n = |\mathcal{I}_n|^{1/2} / \alpha_n^{\gamma'}$  is involved. Then,

$$\rho_n = \frac{|\mathcal{I}_n|^{1/2}}{\alpha_n^{\gamma'}} = \mathcal{O} \left\{ |W_n|^{\frac{d/2 - a(\gamma_2 - d - \varepsilon)}{d}} \right\} \quad (3.14)$$

tends to 0.

We write for short  $s_{\Delta_{n,j}} = s_{\Delta_{n,j},R_n}(\mathbf{X};\theta^*) = s_{\Delta_{n,j},\alpha_n}(\mathbf{X};\theta^*)$  and  $B_{u,n} = B(u, \alpha_n)$ . Note that for any  $u \in W_n \ominus \alpha_n$ ,  $t(u, \mathbf{X}_{u,R_n}) = t(u, \mathbf{X}_{u,\alpha_n}) = t(u, \mathbf{X}_{B_{u,n}})$  and similarly  $\lambda_{\theta^*}(u, \mathbf{X}_{u,R_n}) = \lambda_{\theta^*}(u, \mathbf{X}_{u,\alpha_n}) = \lambda_{\theta^*}(u, \mathbf{X}_{B_{u,n}})$ . Therefore for any  $j \in \mathcal{I}_n$ ,

$$s_{\Delta_{n,j}} = \int_{\Delta_{n,j}} t(u, \mathbf{X}_{B_{u,n}}) \lambda_{\theta^*}(u, \mathbf{X}_{B_{u,n}}) du - \sum_{u \in \mathbf{X}_{\Delta_{n,j}}} t(u, \mathbf{X}_{B_{u,n}} \setminus u).$$

Letting  $Z_{n,j} = s_{\Delta_{n,j}} - \mathbb{E}(s_{\Delta_{n,j}})$ , we have

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

where  $S_n = \sum_{j \in \mathcal{I}_n} Z_{n,j}$ . Define

$$\hat{\Sigma}_n = \sum_{j \in \mathcal{I}_n} \sum_{\substack{k \in \mathcal{I}_n \\ |k-j| \leq \alpha_n}} Z_{n,j} Z_{n,k}^\top \quad \text{and} \quad \Sigma_n = \mathbb{E} \hat{\Sigma}_n.$$

The proof of (i) is completed if we show that  $\Sigma_n^{-1/2} S_n \xrightarrow{d} \mathcal{N}(0, I_p)$ ,  $\Sigma_n^{-1/2} \mathbb{E}(s_{W_n \ominus \alpha_n}) \rightarrow 0$  and  $|W_n|^{-1} \Sigma_n \rightarrow \Sigma_\infty$ . Let us prove the first convergence by application of Theorem A.1.

By  $[\mathbf{W}_n]$  and the definition of  $\mathcal{I}_n$ , we have  $|\mathcal{I}_n| = \mathcal{O}(|W_n|)$ , see e.g. Coeurjolly and Møller (2014, Lemma A.1). From (3.13),  $\alpha_n^{(2+\varepsilon)d} = o(|\mathcal{I}_n|)$  which, following the remark after Theorem A.1, satisfies the assumption of Theorem A.1 if Assumption (a) of this theorem is satisfied for any  $q \geq 1$ . And the latter holds by definition of  $Z_{n,j}$  and Lemma B.3.

Concerning assumption (b), we use for short the notation  $s'_\Delta = s'_{\Delta, \alpha_n}$  for any  $\Delta \in \mathbb{R}^d$ , see (3.9), namely

$$s'_\Delta = \int_{\Delta} t(u, \mathbf{X}_{B_{u,n}}) \lambda_{\theta^*}(u, \mathbf{X}) du - \sum_{u \in \mathbf{X}_\Delta} t(u, \mathbf{X}_{B_{u,n}} \setminus u).$$

Note that from the GNZ formula  $\mathbb{E} s'_\Delta = 0$ . We have from Lemma B.4, for any sequence  $\mathcal{J}_n \subset \mathcal{I}_n$  such that  $|\mathcal{J}_n| \rightarrow \infty$ ,

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

which is  $\mathcal{O}(|\mathcal{J}_n|)$  by Lemma B.3 and from (3.14).

Since  $\Sigma_\infty$  is assumed to be a positive definite matrix, assumption (c) holds if we prove that  $|\mathcal{I}_n|^{-1} \Sigma_n \rightarrow \Sigma_\infty$  as  $n \rightarrow \infty$ . For this, let

$$\Sigma'_n = \text{Var}(s'_{W_n \ominus \alpha_n}) = \sum_{j,k \in \mathcal{I}_n} \mathbb{E} \left\{ s'_{\Delta_{n,j}} (s'_{\Delta_{n,k}})^\top \right\} \quad \text{and} \quad \hat{\Sigma}'_n = \sum_{\substack{j,k \in \mathcal{I}_n \\ |k-j| \leq \alpha_n}} s'_{\Delta_{n,j}} (s'_{\Delta_{n,k}})^\top.$$

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

$$T_1 = |\mathcal{I}_n|^{-1} \| \mathbb{E}(\widehat{\Sigma}'_n) - \mathbb{E}(\widehat{\Sigma}'_n) \|, \quad T_2 = |\mathcal{I}_n|^{-1} \| \mathbb{E}(\widehat{\Sigma}'_n) - \Sigma'_n \|, \quad T_3 = \| |\mathcal{I}_n|^{-1} \Sigma'_n - \Sigma_\infty \|.$$

First applying Lemma B.4

$$\begin{aligned} T_1 &\leq |\mathcal{I}_n|^{-1} \sum_{j \in \mathcal{I}_n} \sum_{\substack{k \in \mathcal{I}_n \\ |k-j| \leq \alpha_n}} \| \mathbb{E}(Z_{n,j} Z_{n,k}^\top) - \mathbb{E}\{s'_{\Delta_{n,j}}(s'_{\Delta_{n,k}})^\top\} \| \\ &= |\mathcal{I}_n|^{-1} \sum_{j \in \mathcal{I}_n} \sum_{\substack{k \in \mathcal{I}_n \\ |k-j| \leq \alpha_n}} \| \text{Cov}(s_{\Delta_{n,j}}, s_{\Delta_{n,k}}) - \text{Cov}(s'_{\Delta_{n,j}}, s'_{\Delta_{n,k}}) \| \\ &= \mathcal{O}(\alpha_n^{-\gamma'}) + \mathcal{O}(\alpha_n^{d-2\gamma'}) \end{aligned}$$

and  $T_1 \rightarrow 0$  since  $\gamma_2 > 2d$  implies  $2\gamma' - d > 0$  as soon as  $\varepsilon < d/2$ , which can be assumed without loss of generality. Second, from (iii) in Lemma B.4

$$\begin{aligned} T_2 &\leq |\mathcal{I}_n|^{-1} \sum_{j \in \mathcal{I}_n} \sum_{\substack{k \in \mathcal{I}_n \\ |k-j| > \alpha_n}} \| \mathbb{E}\{s'_{\Delta_{n,j}}(s'_{\Delta_{n,k}})^\top\} \| \\ &\leq c |\mathcal{I}_n|^{-1} \sum_{j \in \mathcal{I}_n} \sum_{\substack{k \in \mathcal{I}_n \\ |k-j| > \alpha_n}} \frac{1}{|k-j|^{\gamma_2}} \leq c \sum_{|i| > \alpha_n} \frac{1}{|i|^{\gamma_2}} = o(1). \end{aligned}$$

Finally  $T_3 \rightarrow 0$  from (ii) in Lemma B.4, which concludes the proof of condition (c) of Theorem A.1.

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

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

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

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

where  $Y(u, \mathbf{x}) = \| \{ |t_m|(u, \mathbf{x}) \}_{m \geq 1} \| H(u, \mathbf{x}) e^{cG(u, \mathbf{x})} \lambda_{\theta^*}(u, \mathbf{x})$  using the notation of the lemma. Since  $|\Delta_{n,j}| \leq 1$ , we deduce from the stationarity of  $\mathbf{X}$  and Lemma B.2 that

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

Hence

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



tends to 0 from (3.14). All conditions of Theorem A.1 are therefore satisfied, which yields that  $\Sigma_n^{-1/2} S_n \xrightarrow{d} \mathcal{N}(0, I_p)$ . The convergence  $|W_n|^{-1} \Sigma_n \rightarrow \Sigma_\infty$  is an immediate consequence of assumption (c) checked above. It remains to prove that  $\Sigma_n^{-1/2} \mathbb{E}(s_{W_n \ominus \alpha_n}) \rightarrow 0$ . This is a consequence of the GNZ formula, (A.3), Lemma B.1 and the condition (3.13) since

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

(ii) It is worth repeating that  $\theta \rightarrow -\widetilde{\text{LPL}}_{W_n \ominus \alpha_n, R_n}(\mathbf{x}; \theta)$  is a convex function with Hessian matrix given by (3.4). Following Lemmas B.1-B.2 and arguments developed in the proof of Proposition 3.1, we leave the reader to check that almost surely

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

and

$$-|W_n \ominus \alpha_n|^{-1} \frac{d}{d\theta d\theta^\top} \text{LPL}_{W_n \ominus \alpha_n}(\mathbf{X}; \theta) \rightarrow \mathbb{E} \{t(0, \mathbf{X}) t(0, \mathbf{X})^\top \lambda_\theta(0, \mathbf{X})\}$$

as  $n \rightarrow \infty$ , which equals to  $U_\infty$  when  $\theta = \theta^*$ . We also note that (3.5) implies that  $U_\infty$  is a positive definite matrix. These facts and (i) allow us to apply Guyon (1995, Theorem 3.4.5) to deduce the result.  $\square$

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 3.1 and Theorem 3.3.

**Proposition 3.4.** *Under the assumptions of Proposition 3.1, the maximum logistic regression estimator defined by*

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

*converges almost surely to  $\theta^*$  as  $n \rightarrow \infty$  and under the assumptions of Theorem 3.3 it satisfies the following convergence in distribution*

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

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

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

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  $\mathbf{x} \in \Omega$ .

## 4 Simulation study

In this section, we present results of simulation experiments assessing the performance estimation of maximum pseudolikelihood estimators in the Lennard-Jones model, given by

$$\log \lambda_\theta(u, \mathbf{x}) = \log(\beta) - \sum_{v \in \mathbf{x}} \Phi(v - u) \quad \text{with} \quad \Phi(u) = 4\varepsilon \left\{ \left( \frac{\sigma}{\|u\|} \right)^6 - \left( \frac{\sigma}{\|u\|} \right)^{12} \right\}. \quad (4.1)$$

We chose  $\beta = 100$  and  $\sigma = 0.1$  and considered three cases where  $\varepsilon$  takes the values 0.1, 0.5 and 1 respectively, which, following Baddeley and Dereudre (2013) we call low, moderate and high rigidity models. The realizations are generated using the Metropolis-Hastings algorithm, implemented in the R package `spatstat` (Baddeley and Turner, 2005; Baddeley et al., 2015), on  $W_n = [-n, n]^2$  and for  $n = 1/2, 1, 2$ . To take into account the infinite range characteristic of the Lennard-Jones model, the processes are simulated on  $[-n - 2, n + 2]^2$  and then clipped to  $W_n$ . Figure 1 depicts some typical realizations on  $[-1, 1]^2$ .

For each model and each observation window, we considered three versions of maximum pseudolikelihood estimators given by (2.7) of the parameter vector  $\{\log(\beta), \sigma, \varepsilon\}^\top$ : (i)  $\alpha_n = R_n \in [0.05, 0.3]$ , (ii)  $\alpha_n \in [0.05, 0.3]$  and  $R_n = \infty$ , (iii)  $\alpha_n = 0$ ,  $R_n = \infty$ . We remind that the values  $\alpha_n = 0$  and  $R_n = \infty$  respectively mean that no border erosion is considered (i.e.  $W_n \ominus \alpha_n = W_n$ ) and the maximal possible range of interaction in  $W_n$  is taken into account (i.e.  $\lambda(u, \mathbf{x}_{u, R_n}) = \lambda(u, \mathbf{x}_{W_n})$ ). Writing  $\alpha_n$  and/or  $R_n \in [0.05, 0.3]$  means that we evaluated the estimates for 30 values regularly sampled in  $[0.05, 0.3]$ .

We computed the pseudolikelihood estimator by using a  $100 \times 100$  grid of quadrature points to discretize the integral involved in (2.7). We did not use the Berman-Turner approximation, implemented in `spatstat` for a large class of models excluding (4.1) (see Baddeley and Turner (2000a)), because the latter may artificially lead to biased estimates for very repulsive patterns. As suggested by Baddeley and Dereudre (2013), to minimise numerical problems (overflow, instability, slow convergence) we rescaled the interpoint distances to a unit equal to the true value of  $\sigma$ .

We define the weighted mean squared error WMSE by

$$\text{WMSE} = \frac{\hat{E} \left\{ \left( \widehat{\log \beta} - \log \beta \right)^2 \right\}}{(\log \beta)^2} + \frac{\hat{E} \{ (\hat{\sigma} - \sigma)^2 \}}{\sigma^2} + \frac{\hat{E} \{ (\hat{\varepsilon} - \varepsilon)^2 \}}{\varepsilon^2}$$

and we consider in the following its root  $\text{RWMSE} = \sqrt{\text{WMSE}}$ . Similarly, we define the root-weighted squared bias and the root-weighted variance respectively denoted by  $\text{RWSQ}$  and  $\text{RWV}$ .

Tables 1 and 2 summarize the simulation study based on 100 replications, where we report the values of  $\text{RWMSE}$ ,  $\text{RWSB}$  and  $\text{RWV}$ . When  $\alpha_n$  and/or  $R_n$  vary, we report in Table 1 the smallest value of  $\text{RWMSE}$  and the associated value  $\alpha_{\text{opt}}$  of  $\alpha_n$  between brackets. To be consistent, we report in Table 2 the values of  $\text{RWSB}$  and  $\text{RWV}$  associated to  $\alpha_{\text{opt}}$ . We observe that the three versions of the estimates have a  $\text{RWMSE}$  decreasing with  $n$  for the three Lennard-Jones models. In case (i)

where  $\alpha_n = R_n$ , the optimal value seems to be around  $\alpha_n = 0.15$ . A closer look at the estimates showed us that their average behavior (sample mean and standard deviation) fluctuate quite a lot with  $\alpha_n$ . In case (ii) where  $R_n = \infty$ , we observed that the biases of the estimates do not fluctuate that much with  $\alpha_n$ . Since the estimates had smaller standard deviation when the amount of information is maximal, i.e. when  $\alpha_n$  is low, this explains why the smallest value of  $\alpha_n$  led in almost all cases to the smallest RWMSE. Surprisingly, the third situation corresponding to  $\alpha_n = 0$  and  $R_n = \infty$  produced very interesting results which are optimal or close to the optimal ones in all cases considered. This estimator may be very time consuming to evaluate for very large datasets since all the points are involved in the evaluations of the Papangelou conditional intensity. Nonetheless, for the setting considered in this simulation study the computational time differences were negligible. The situation  $\alpha_n = 0$  and  $R_n = \infty$  is supported by Proposition 3.1 (consistency) but not by Theorem 3.3 (asymptotic normality). However, the normal QQ-plots in Figures 2-4 seem to show a convergence to a Gaussian behavior for all our choices of  $\alpha_n$  and  $R_n$ , with approximatively the same rate of convergence, i.e.  $|W_n|^{-1/2}$ , if we refer to the decreasing rate of the slopes in each QQ-plot. Note that the Gaussian behavior is less clear in the low rigidity Lennard-Jones model than in the moderate and high rigidity cases, but this seems specific to the model rather than to the estimators. In conclusion, to estimate the parameters of a Lennard-Jones model using the pseudolikelihood method, we recommend to use no erosion and no finite range correction.

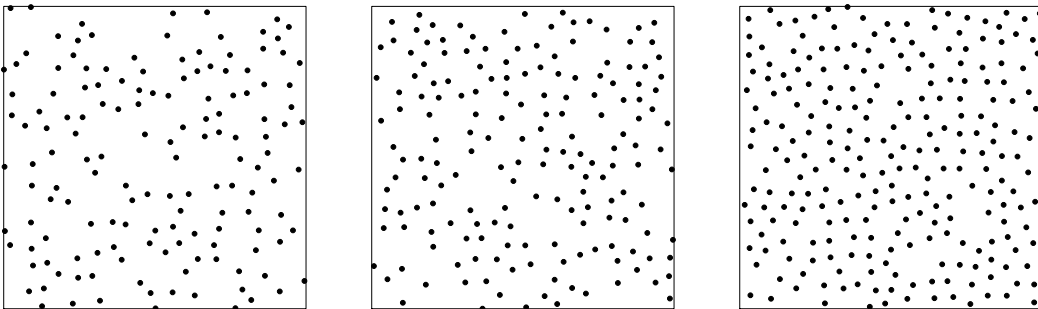


Figure 1: Typical realizations on  $[-1, 1]^2$  of a Lennard-Jones model with parameters  $\beta = 100$ ,  $\sigma = 0.1$  and  $\varepsilon = 0.1$  (left),  $0.5$  (middle) and  $1$  (right).

## A 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 conditionally centered, by application of the conditional GNZ formula (2.4). In the infinite range case, the score functions of the log-pseudolikelihood and the logistic regression are neither centered, nor conditionally centered. In the following theorem, the

	RWMSE		
	$[-1/2, 1/2]^2$	$[-1, 1]^2$	$[-2, 2]^2$
<b>Low (<math>\varepsilon = 0.1</math>)</b>			
$\alpha_n = R_n \in [0.05, 0.3]$	3.26 (0.13)	1.25 (0.13)	0.62 (0.12)
$\alpha_n \in [0.05, 0.3], R_n = \infty$	3.72 (0.05)	1.79 (0.05)	0.63 (0.06)
$\alpha_n = 0, R_n = \infty$	3.5	1.66	0.69
<b>Moderate (<math>\varepsilon = 0.5</math>)</b>			
$\alpha_n = R_n \in [0.05, 0.3]$	0.65 (0.12)	0.34 (0.14)	0.2 (0.15)
$\alpha_n \in [0.05, 0.3], R_n = \infty$	0.68 (0.05)	0.38 (0.05)	0.19 (0.05)
$\alpha_n = 0, R_n = \infty$	0.59	0.33	0.18
<b>High (<math>\varepsilon = 1</math>)</b>			
$\alpha_n = R_n \in [0.05, 0.3]$	1.04 (0.08)	0.42 (0.16)	0.13 (0.16)
$\alpha_n \in [0.05, 0.3], R_n = \infty$	1.34 (0.05)	0.36 (0.05)	0.16 (0.05)
$\alpha_n = 0, R_n = \infty$	1.23	0.27	0.17

Table 1: Root-weighted mean squared errors (RWMSE) of parameters estimates for different Lennard-Jones models. The results are based on 100 replications. The realizations are generated on  $[-n-2, n+2]^2$  for  $n = 1/2, 1, 2$  and the window of observation corresponds to  $[-n, n]^2$ . When it makes sense, we indicate between brackets the value  $\alpha_{\text{opt}}$  of  $\alpha_n$  leading to the minimal value of RWMSE.

	RWSB and RWV					
	$[-1/2, 1/2]^2$		$[-1, 1]^2$		$[-2, 2]^2$	
<b>Low (<math>\varepsilon = 0.1</math>)</b>						
$\alpha_n = R_n \in [0.05, 0.3]$	1.82	2.70	0.57	1.11	0.07	0.62
$\alpha_n \in [0.05, 0.3], R_n = \infty$	2.49	2.76	0.82	1.59	0.03	0.63
$\alpha_n = 0, R_n = \infty$	2.36	2.59	0.78	1.46	0.20	0.66
<b>Moderate (<math>\varepsilon = 0.5</math>)</b>						
$\alpha_n = R_n \in [0.05, 0.3]$	0.23	0.60	0.16	0.30	0.07	0.19
$\alpha_n \in [0.05, 0.3], R_n = \infty$	0.04	0.66	0.10	0.37	0.02	0.19
$\alpha_n = 0, R_n = \infty$	0.07	0.58	0.02	0.33	0.02	0.18
<b>High (<math>\varepsilon = 1</math>)</b>						
$\alpha_n = R_n \in [0.05, 0.3]$	0.43	0.71	0.16	0.39	0.07	0.12
$\alpha_n \in [0.05, 0.3], R_n = \infty$	0.11	1.27	0.13	0.34	0.12	0.10
$\alpha_n = 0, R_n = \infty$	0.06	1.23	0.05	0.26	0.12	0.11

Table 2: Root-weighted squared biases (RWSB) and variances (RWS) of parameters estimates for different Lennard-Jones models. The setting is the same as in Table 1. When  $\alpha_n$  and  $R_n$  vary, we report the values leading to the minimal RWMSE, i.e. the values associated to  $\alpha_n = \alpha_{\text{opt}}$  as given in Table 1.

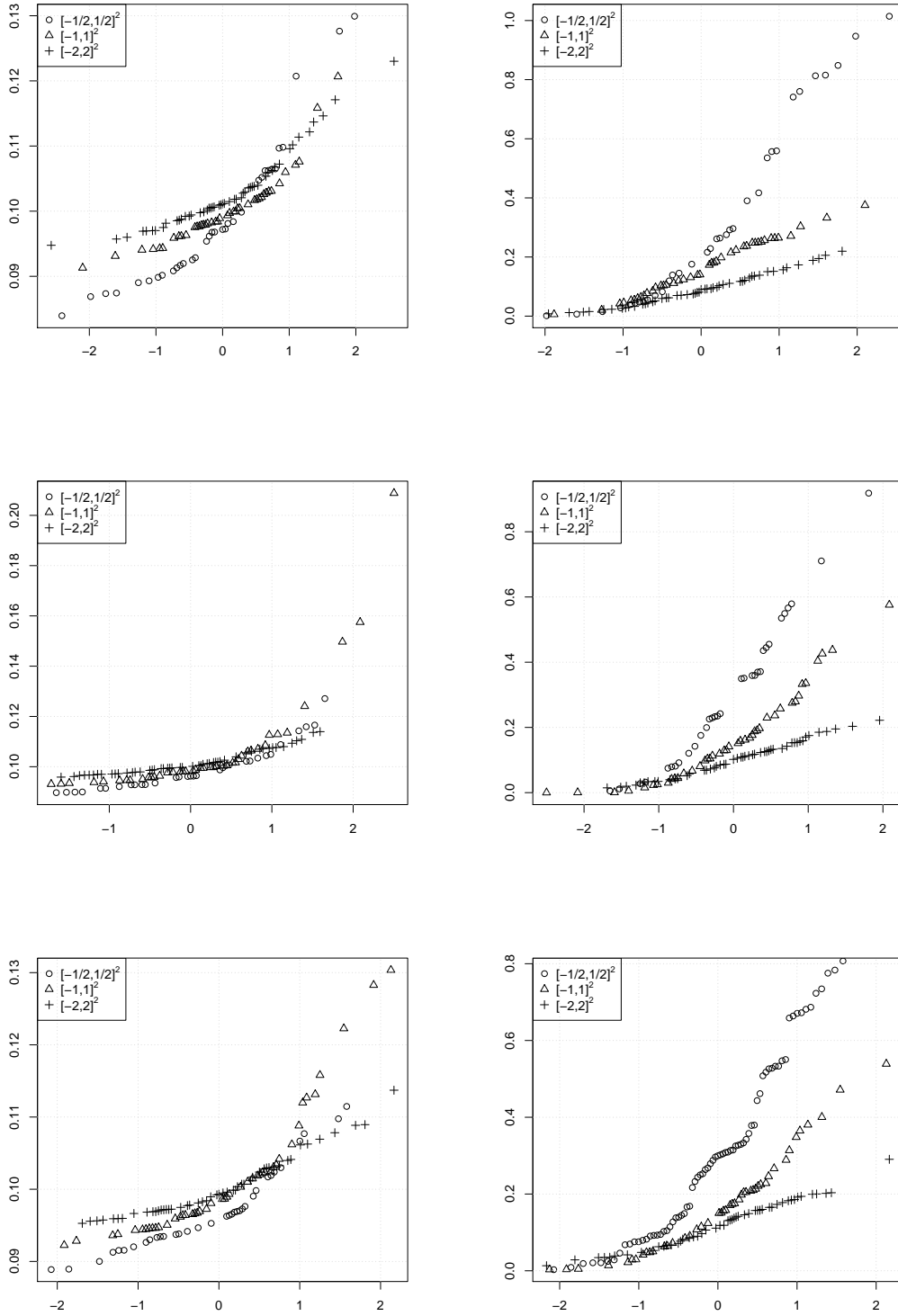


Figure 2: Normal QQ-plots for estimates of the parameter  $\varepsilon$  (left) and  $\sigma$  (right) for the low interaction Lennard-Jones model, i.e.  $(\log(\beta), \varepsilon, \sigma) = (\log(100), 0.1, 0.1)$ . The first row (resp. the second and third) corresponds to estimates obtained with  $\alpha_n = R_n = \alpha_{\text{opt}}$  (resp. with  $\alpha_n = \alpha_{\text{opt}}, R_n = \infty$  and with  $\alpha_n = 0, R_n = \infty$ ). The optimal values  $\alpha_{\text{opt}}$  are given in Table 1.

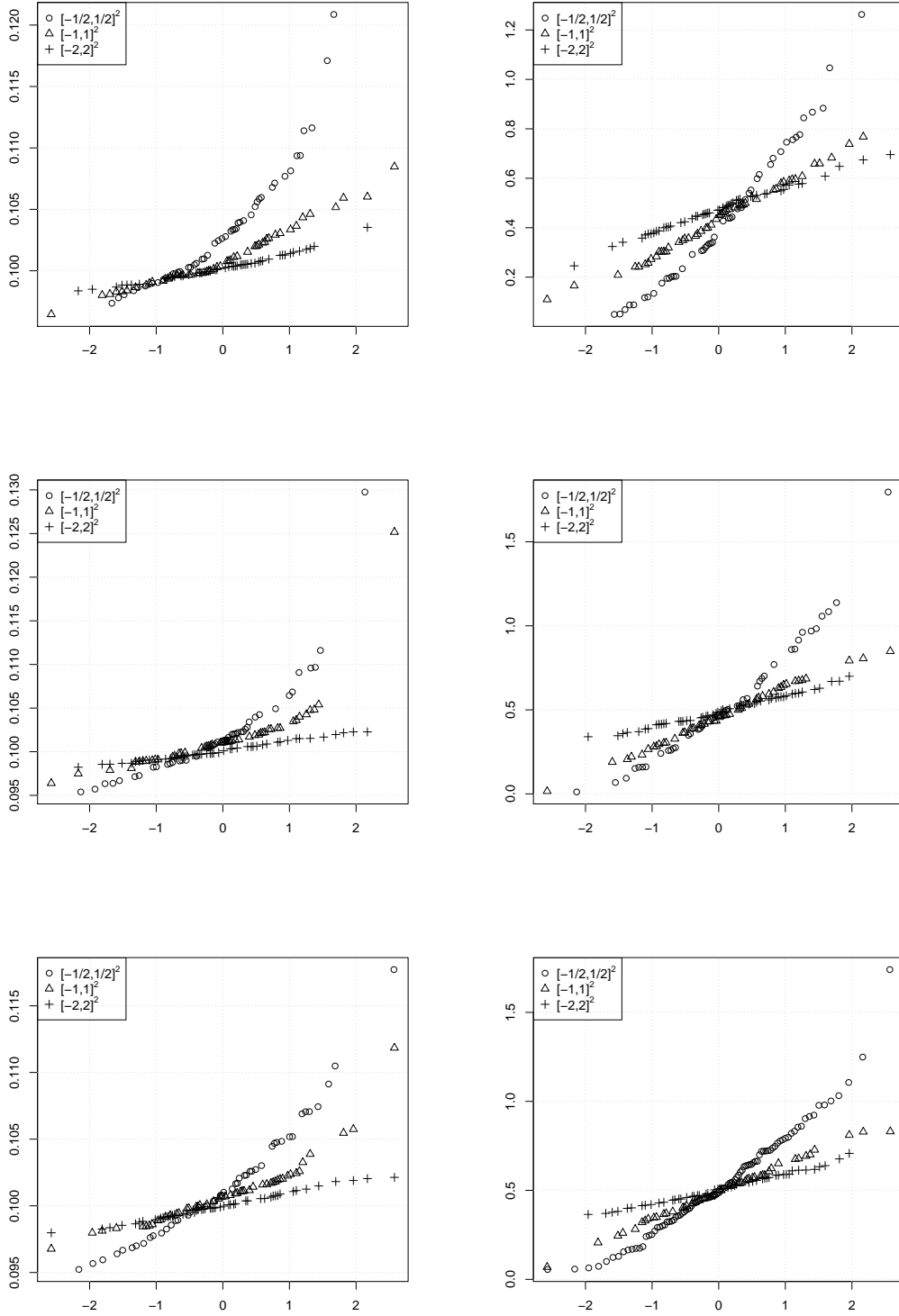


Figure 3: Normal QQ-plots for estimates of the parameter  $\varepsilon$  (left) and  $\sigma$  (right) for the moderate interaction Lennard-Jones model, i.e.  $(\log(\beta), \varepsilon, \sigma) = (\log(100), 0.5, 0.1)$ . The first row (resp. the second and third) corresponds to estimates obtained with  $\alpha_n = R_n = \alpha_{\text{opt}}$  (resp. with  $\alpha_n = \alpha_{\text{opt}}, R_n = \infty$  and with  $\alpha_n = 0, R_n = \infty$ ). The optimal values  $\alpha_{\text{opt}}$  are given in Table 1.

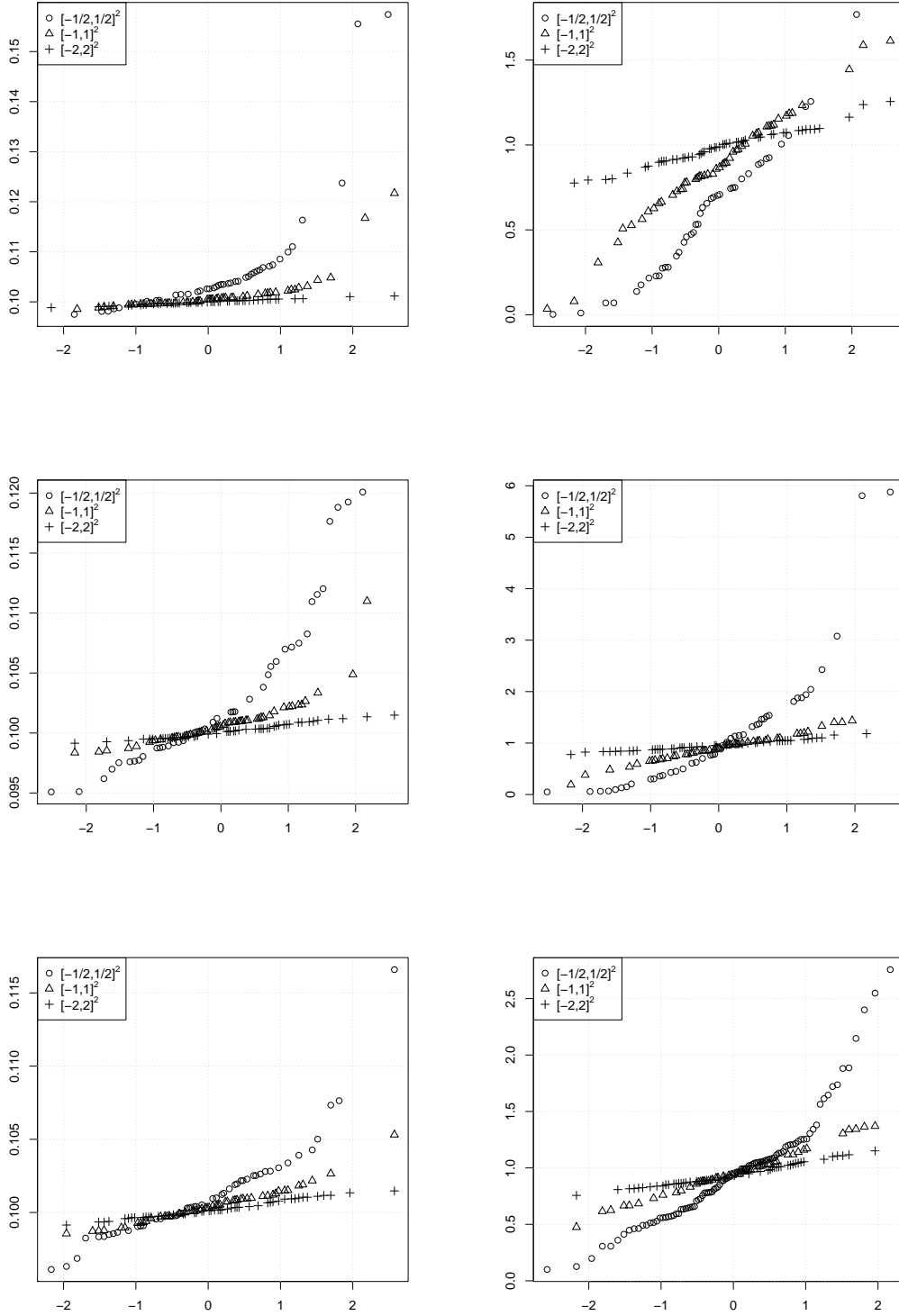


Figure 4: Normal QQ-plots for estimates of the parameter  $\varepsilon$  (left) and  $\sigma$  (right) for the high interaction Lennard-Jones model, i.e.  $(\log(\beta), \varepsilon, \sigma) = (\log(100), 1, 0.1)$ . The first row (resp. the second and third) corresponds to estimates obtained with  $\alpha_n = R_n = \alpha_{\text{opt}}$  (resp. with  $\alpha_n = \alpha_{\text{opt}}, R_n = \infty$  and with  $\alpha_n = 0, R_n = \infty$ ). The optimal values  $\alpha_{\text{opt}}$  are given in Table 1.

conditional centering condition is replaced by condition (d), which turns out to be sufficient for our application to  $s_{W_n \ominus \alpha_n, R_n}(\mathbf{X}; \theta^*)$  in Theorem 3.3. The other conditions are mainly due to the non-stationary setting induced by the presence of  $\alpha_n$  and  $R_n$ . They allow in particular to control the asymptotic behavior of the empirical covariance matrix in (A.1). For two square matrices  $A, B$  we write  $A \geq B$  when  $A - B$  is a positive semi-definite matrix.

**Theorem A.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  $|\mathcal{I}_n| \rightarrow \infty$  and  $\alpha_n \rightarrow \infty$  as  $n \rightarrow \infty$ . 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  $\mathcal{K}_{n,j} = \{k \in \mathbb{Z}^d, |k - j| \leq \alpha_n\}$  and where  $f_{n,j} : S^{\mathcal{K}_{n,j}} \rightarrow \mathbb{R}^p$  is a measurable function. We define  $\hat{\Sigma}_n$  and  $\Sigma_n$  by*

$$\hat{\Sigma}_n = \sum_{j \in \mathcal{I}_n} \sum_{\substack{k \in \mathcal{I}_n \\ |k-j| \leq \alpha_n}} Z_{n,j} Z_{n,k}^\top \quad \text{and} \quad \Sigma_n = \mathbb{E} \hat{\Sigma}_n.$$

We assume that

- (a)  $\mathbb{E} Z_{n,j} = 0$  and there exists  $q \geq 1$  such that  $\sup_{n \geq 1} \sup_{j \in \mathcal{I}_n} \mathbb{E} \|Z_{n,j}\|^{4q} < \infty$ ,
- (b) for any sequence  $\mathcal{J}_n \subset \mathcal{I}_n$  such that  $|\mathcal{J}_n| \rightarrow \infty$  as  $n \rightarrow \infty$ ,

$$|\mathcal{J}_n|^{-1} \sum_{j,k \in \mathcal{J}_n} \|\mathbb{E}(Z_{n,j} Z_{n,k}^\top)\| = \mathcal{O}(1).$$

Then if  $\alpha_n^{\frac{4q-1}{2q-1}d} = o(|\mathcal{I}_n|)$  as  $n \rightarrow \infty$ ,

$$|\mathcal{I}_n|^{-1} (\hat{\Sigma}_n - \Sigma_n) \rightarrow 0 \quad \text{in } L^{2q}. \quad (\text{A.1})$$

If in addition

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

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

then

$$\Sigma_n^{-1/2} S_n \xrightarrow{d} \mathcal{N}(0, I_p). \quad (\text{A.2})$$

Before detailing the proof, let us remark that if assumption (a) is valid for any  $q \geq 1$  then the result remains true if  $\alpha_n^{(2+\varepsilon)d} = o(|\mathcal{I}_n|)$  for any  $\varepsilon > 0$ .

*Proof.* For  $m, m' = 1, \dots, p$ , let  $\Delta_{mm'} = (\hat{\Sigma}_n - \Sigma_n)_{mm'}$ . Let  $q \geq 1$  be as in assumption (a), the assertion (A.1) will be proved if we prove that  $\mathbb{E} (|\mathcal{I}_n|^{-2q} \Delta_{mm'}^{2q}) = o(1)$ . We have  $\Delta_{mm'} = \sum_{j \in \mathcal{I}_n} U_{n,j}$  where

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



and  $\mathcal{I}_{n,j} = \{k \in \mathcal{I}_n : |k - j| \leq \alpha_n\}$ . Let  $j_1, \dots, j_{2q} \in \mathcal{I}_n$  such that  $|j_k - j_1| > 4\alpha_n$  for  $k = 2, \dots, 2q$ . It is clear that for any  $j \in \mathcal{I}_n$ ,  $U_{n,j}$  depends only on  $X_{n,k}$  for  $|k - j| \leq 2\alpha_n$ . So,

$$\begin{aligned} \mathbb{E}(U_{n,j_1} \dots U_{n,j_{2q}}) &= \mathbb{E} \left\{ \mathbb{E}(U_{n,j_1} \dots U_{n,j_{2q}} \mid X_{n,k}, |k - j_1| > 2\alpha_n) \right\} \\ &= \mathbb{E} \left\{ U_{n,j_2} \dots U_{n,j_{2q}} \mathbb{E}(U_{n,j_1} \mid X_{n,k}, |k - j_1| > 2\alpha_n) \right\} \\ &= \mathbb{E} \left\{ U_{n,j_2} \dots U_{n,j_{2q}} \mathbb{E}(U_{n,j_1}) \right\} = 0 \end{aligned}$$

whereby we deduce that

$$\mathbb{E}(\Delta_{mm'}^{2q}) = \sum_{\substack{j_1, \dots, j_{2q} \in \mathcal{I}_n \\ |j_k - j_1| \leq 4\alpha_n, k=2, \dots, 2q}} \mathbb{E}(U_{n,j_1} \dots U_{n,j_{2q}}).$$

Now, by condition (a) and Hölder's inequality, we have for any  $j \in \mathcal{I}_n$

$$\begin{aligned} \mathbb{E} U_{n,j}^{2q} &= \sum_{k_1, \dots, k_{2q} \in \mathcal{I}_{n,j}} \mathbb{E} \left[ \left\{ Z_{n,j} Z_{n,k_1}^\top - \mathbb{E}(Z_{n,j} Z_{n,k_1}^\top) \right\}_{mm'} \dots \left\{ Z_{n,j} Z_{n,k_{2q}}^\top - \mathbb{E}(Z_{n,j} Z_{n,k_{2q}}^\top) \right\}_{mm'} \right] \\ &\leq c \sum_{k_1, \dots, k_{2q} \in \mathcal{I}_{n,j}} \mu_{4q} \\ &\leq c \left( \sup_{j \in \mathcal{I}_n} \#\{k \in \mathcal{I}_n, |k - j| \leq \alpha_n\} \right)^{2q} = \mathcal{O}(\alpha_n^{2qd}). \end{aligned}$$

From Hölder's inequality, we continue with

$$\mathbb{E}(\Delta_{mm'}^{2q}) \leq \sum_{\substack{j_1, \dots, j_{2q} \in \mathcal{I}_n \\ |j_k - j_1| \leq 4\alpha_n, k=2, \dots, 2q}} \mathbb{E}(U_{n,j_1}^{2q})^{1/2q} \mathbb{E}(U_{n,j_{2q}}^{2q})^{1/2q} = \mathcal{O} \left\{ \alpha_n^{(4q-1)d} |\mathcal{I}_n| \right\}$$

leading to

$$\mathbb{E}(|\mathcal{I}_n|^{-2q} \Delta_{mm'}^{2q}) = \mathcal{O} \left( \frac{\alpha_n^{(4q-1)d}}{|\mathcal{I}_n|^{2q-1}} \right) = o(1)$$

by assumption on  $\alpha_n$ , which completes the proof of (A.1).

We now focus on (A.2) and we let

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

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 (A.2) it suffices to prove that for all  $u \in \mathbb{R}^p$  such that  $\|u\| = 1$  and for all  $\omega \in \mathbb{R}$

$$\mathbb{E} \left\{ (\mathbf{i}\omega - u^\top \bar{S}_n) e^{\mathbf{i}\omega u^\top \bar{S}_n} \right\} \rightarrow 0$$

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

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

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

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  $\ell$  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,  $\lambda_{\min}(|\mathcal{I}_n|^{-1} \Sigma_n) \geq \lambda_{\min}(Q) > 0$  whereby we deduce

$$\| \Sigma_n^{-1/2} \| = |\mathcal{I}_n|^{-1/2} \text{tr}(|\mathcal{I}_n| \Sigma_n^{-1})^{1/2} \leq \ell^{1/2} |\mathcal{I}_n|^{-1/2}. \quad (\text{A.3})$$

Using this result, Jensen's inequality and the sub-multiplicative property of the Frobenius norm, we get for  $q \geq 1$  satisfying (a) and the assumption on  $\alpha_n$

$$\begin{aligned} \| \mathbb{E} A_{n,1} \|^{2q} &\leq \|v\|^{2q} \mathbb{E} \|I_p - \Sigma_n^{-1/2} \widehat{\Sigma}_n \Sigma_n^{-1/2}\|^{2q} \\ &\leq \|v\|^{2q} \| \Sigma_n^{-1/2} \|^{4q} \mathbb{E} \| \widehat{\Sigma}_n - \Sigma_n \|^{2q} \\ &\leq \ell^{2q} \|v\|^{2q} |\mathcal{I}_n|^{-2q} \mathbb{E} \| \widehat{\Sigma}_n - \Sigma_n \|^{2q} \end{aligned}$$

whereby we deduce that  $\| \mathbb{E} A_{n,1} \| \rightarrow 0$  from (A.1).

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

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

where

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

Let us decompose  $B_{n,j} = B_{n,j}^{(1)} + B_{n,j}^{(2)}$  where  $B_{n,j}^{(1)} = B_{n,j} \mathbf{1}(\|Z_{n,j}\| \leq |\mathcal{I}_n|^\tau)$  and  $B_{n,j}^{(2)} = B_{n,j} \mathbf{1}(\|Z_{n,j}\| > |\mathcal{I}_n|^\tau)$  with  $1/\tau = 2(4p-1)$ . By assumption (b), we have

$$\| \mathbb{E} B_{n,j}^{(1)} \| \leq |\mathcal{I}_n|^\tau \sum_{k, k' \in \mathcal{I}_{n,j}} | \mathbb{E}(Z_{n,k}^\top Z_{n,k'}) | = \mathcal{O}(|\mathcal{I}_n|^\tau |\mathcal{I}_{n,j}|) = \mathcal{O}(|\mathcal{I}_n|^\tau \alpha_n^d). \quad (\text{A.4})$$

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

with

$$\begin{aligned}
\mathbb{E} B_{n,j}^{(2)} &\leq \sum_{k,k' \in \mathcal{I}_{n,j}} \mathbb{E} \{ \|Z_{n,j}\| \|Z_{n,k}\| \|Z_{n,k'}\| \mathbf{1}(\|Z_{n,j}\| > |\mathcal{I}_n|^\tau) \} \\
&\leq \alpha_n^{2d} \mu_{4q}^{\frac{3}{4q}} \mathbb{P}(\|Z_{n,j}\| > |\mathcal{I}_n|^\tau)^{\frac{4q-3}{4q}} \\
&\leq \alpha_n^{2d} \mu_{4q}^{\frac{3}{4q}} (|\mathcal{I}_n|^{-4q\tau} \mathbb{E} \|Z_{n,j}\|^{4q})^{\frac{4q-3}{4q}} \\
&\leq \mu_{4q} \alpha_n^{2d} |\mathcal{I}_n|^{-(4q-3)\tau}.
\end{aligned} \tag{A.5}$$

Combining (A.4)-(A.5), we deduce that as  $n \rightarrow \infty$

$$\| \mathbb{E} A_{n,2} \| \leq \mathbb{E} \| A_{n,2} \| = \mathcal{O} \left( \frac{\alpha_n^d}{|\mathcal{I}_n|^{1/2}} |\mathcal{I}_n|^\tau \right) + \mathcal{O} \left( \frac{\alpha_n^{2d}}{|\mathcal{I}_n|^{1/2}} |\mathcal{I}_n|^{-(4q-3)\tau} \right) = o(1)$$

by definition of  $\tau$  and  $\alpha_n$ .

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

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

whereby we deduce, in view of (A.3), that

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

which tends to 0 by assumption (d).  $\square$

## B Auxiliary results

We gather in this section several auxiliary results. They are established under the setting, assumptions and notation of Section 3. In particular, we recall that  $\Delta_j$  is the cube centered at  $j \in \mathbb{Z}^d$  with volume 1,  $\Delta_{n,j} = \Delta_j \cap (W_n \ominus \alpha_n)$ ,  $\mathcal{I}_n \subset \mathbb{Z}^d$  is the set such that  $W_n \ominus \alpha_n = \cup_{j \in \mathcal{I}_n} \Delta_{n,j}$ ,  $B_{u,n} = B(u, \alpha_n)$  and for any  $\Delta \Subset \mathbb{R}^d$

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

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

**Lemma B.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*

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

Then, if  $\alpha_n \geq r_0$

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

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

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

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

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

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

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

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

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

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

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

- (i) For any  $q \geq 0$ ,  $\mathbb{E}\{\lambda_\theta(0, \mathbf{X})^q\} < \infty$ .
- (ii) Let  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  be a measurable function such that  $|f(u)| \leq c(1 + \|u\|)^{-\gamma}$  with  $\gamma > d$ , then for any  $q \geq 0$

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

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

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

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

*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  $\psi : \mathbb{R}^+ \rightarrow \mathbb{R}^+$  is a decreasing function with  $\int_0^\infty \psi(t)t^{d-1} dt < \infty$ , then for any  $q \geq 0$ ,

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

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  $q \geq 0, \kappa > 0, x \mapsto x^q e^{-\kappa x}$  is bounded on  $[0, \infty)$ , we have

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

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  $[\Phi]$  for any  $\theta$ . 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  $[\mathbf{g}]$ , 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  $m \geq 3$ , 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  $[\mathbf{g}]$ , 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  $|\tilde{\Phi}(u)| \leq c\|u\|^{-\gamma_2}$  and that  $\tilde{\Phi}_\theta$  is bounded from below, proving that it satisfies  $[\Phi]$ .  $\square$

**Lemma B.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 (B.1), respectively  $s'_{\Delta_{n,j}}$  given by (B.2). Under  $[\mathbf{g}]$ , if  $\alpha_n \geq r_0$  then, for any  $q \in \mathbb{N}$ ,  $\mathbb{E}(|s_m|^q) < \infty$  and  $\mathbb{E}(|s'_m|^q) < \infty$ .*

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

$$\mathbb{E} \left\{ \left| \sum_{u \in \mathbf{X}_{\Delta_{n,j}}} t_m(u, \mathbf{X}_{B_{u,n}} \setminus u) \right|^{p_1} \left| \int_{\Delta_{n,j}} t_m(u, \mathbf{X}_{B_{u,n}}) \lambda_{\theta^*}(u, \mathbf{X}_{B_{u,n}}) du \right|^{p_2} \right\} < \infty$$

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 B.1

we get

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

which is finite by Lemma B.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

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

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

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

we obtain by application of Hölder's inequality,

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

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

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

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

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

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

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

which is finite from **[g]**. □

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

(i) For  $\Lambda_1, \Lambda_2$  two bounded Borel sets of  $\mathbb{R}^d$

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

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

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

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

where  $\Sigma_\infty$  is defined by (3.11).

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

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

(iv) For any  $j, k \in \mathcal{I}_n$ , if  $\alpha_n \geq r_0$ , then

$$\left\| \text{Cov}(s_{\Delta_{n,j}}, s_{\Delta_{n,k}}) - \text{Cov}(s'_{\Delta_{n,j}}, s'_{\Delta_{n,k}}) \right\| \leq \frac{c}{\alpha_n^{\gamma'}(1 + |k - j|^{\gamma_2})} + \frac{c}{\alpha_n^{2\gamma'}}$$

as  $n \rightarrow \infty$ , 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  $u, v \in \mathbb{R}^d$ ,  $m \geq 1$  and  $\mathbf{x} \in \Omega$

$$\lambda_{\theta^*}(u, \mathbf{x})\lambda_{\theta^*}(v, \mathbf{x}) - \lambda_{\theta^*}(\{u, v\}, \mathbf{x}) = \lambda_{\theta^*}(u, \mathbf{x})\lambda_{\theta^*}(v, \mathbf{x})\{1 - e^{-\Phi_{\theta^*}(v-u)}\} \quad (\text{B.3})$$

and

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

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

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

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

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

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

$$\begin{aligned} & \left\| \text{Cov}(s'_{\Delta_{n,j}}, s'_{\Delta_{n,k}}) \right\| \\ & \leq \mathbb{E} \int_{\Delta_{n,j}} \int_{\Delta_{n,k}} \| |t|(u, \mathbf{X})|t|(v, \mathbf{X})^\top \| \lambda_{\theta^*}(u, \mathbf{X})\lambda_{\theta^*}(v, \mathbf{X}) |1 - e^{-\Phi_{\theta^*}(v-u)}| du dv \\ & \quad + \mathbb{E} \int_{\Delta_{n,j}} \int_{\Delta_{n,k}} \|g(v-u)g(v-u)^\top\| \lambda_{\theta^*}(u, \mathbf{X})\lambda_{\theta^*}(v, \mathbf{X}) e^{-\Phi_{\theta^*}(v-u)} du dv. \end{aligned} \quad (\text{B.5})$$

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

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

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

$$|1 - e^{-\Phi_{\theta^*}(v-u)}| \leq |\Phi_{\theta^*}(v-u)| e^{|\Phi_{\theta^*}(v-u)|} \leq c|k - j|^{-\gamma_2}. \quad (\text{B.6})$$



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

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

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

We have

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

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

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

By definition of  $\lambda_{\theta^*}$  and  $t$  (see (2.2) and (3.2)), we have for any  $u, v \in \mathbb{R}^d$  and  $\mathbf{x} \in \Omega$

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

whereby

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

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

$$\| \text{E}(s'_{\Delta_{n,j}} I_k^\top) \| \leq \frac{c}{\alpha_n^{\gamma'} (1 + |k-j|^{\gamma_2})}. \quad (\text{B.9})$$

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

$$\text{E}(I_j I_k^\top) = \text{E} \int_{\Delta_{n,j}} \int_{\Delta_{n,k}} \pi_n(u, \mathbf{X}) \pi_n(v, \mathbf{X})^\top du dv$$

and, after application of the GNZ formula,

$$\text{E}(s_{\Delta_{n,j}}) \text{E}(s_{\Delta_{n,k}})^\top = \int_{\Delta_{n,j}} \int_{\Delta_{n,k}} \text{E} \pi_n(u, \mathbf{X}) \text{E} \pi_n(v, \mathbf{X})^\top du dv,$$

we deduce from Lemmas B.1-B.2 that their norm is bounded by  $\alpha_n^{-2\gamma'}$  for any  $j, k$ , up to a positive constant. The latter and (B.9) prove (iv).  $\square$

## Acknowledgements

The authors would like to thank sincerely the referee for providing many interesting remarks and relevant suggestions. Her/his work allowed us to have a new look on a previous version of this paper and allowed us to improve our results in several directions. The research of J.-F. Coeurjolly is partially funded by Persyval-lab EA Oculo-Nimbus.

## References

- A. Baddeley and D. Dereudre. Variational estimators for the parameters of Gibbs point process models. *Bernoulli*, 19(3):905–930, 2013.
- A. Baddeley and R. Turner. Practical maximum pseudolikelihood for spatial point patterns. *Australian and New Zealand Journal of Statistics*, 42:283–322, 2000a.
- 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, 2000b.
- A. Baddeley and R. Turner. Spatstat: an R package for analyzing spatial point patterns. *Journal of Statistical Software*, 12:1–42, 2005.
- A. Baddeley, J.-F. Coeurjolly, E. Rubak, and R. Waagepetersen. Logistic regression for spatial Gibbs point processes. *Biometrika*, 101(2):377–392, 2014.
- A. Baddeley, E. Rubak, and R. Turner. *Spatial Point Patterns: Methodology and Applications with R*. Chapman and Hall/CRC Press, London, 2015. In press.
- 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, Chichester, third edition, 2013.
- J.-F. Coeurjolly and R. Drouilhet. Asymptotic properties of the maximum pseudolikelihood 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.
- H. O. 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, Chichester, 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, London, 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*, 18(2):127–159, 1970.
- I.M. Slivnyak. Some properties of stationary flows of homogeneous random events. *Theory of Probability & Its Applications*, 7(3):336–341, 1962.