



**HAL**  
open science

# Maximum pseudo-likelihood estimator for general marked Gibbs point processes and applications to the Lennard-Jones model

Jean-François Coeurjolly, Rémy Drouilhet

► **To cite this version:**

Jean-François Coeurjolly, Rémy Drouilhet. Maximum pseudo-likelihood estimator for general marked Gibbs point processes and applications to the Lennard-Jones model. 2009. hal-00442750v1

**HAL Id: hal-00442750**

**<https://hal.science/hal-00442750v1>**

Submitted on 22 Dec 2009 (v1), last revised 12 Jul 2010 (v2)

**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.

# Maximum pseudo-likelihood estimator for general marked Gibbs point processes and applications to the Lennard-Jones model

Jean-François Coeurjolly<sup>1,2</sup> and Rémy Drouilhet<sup>1</sup>

<sup>1</sup> Sagag Team, Dept Statistics, LJK, Grenoble University, France

<sup>2</sup> SigmaPhy Team, GIPSA-lab, Grenoble University, France

December 22, 2009

## Abstract

This paper is devoted to the maximum pseudo-likelihood estimator of a vector  $\boldsymbol{\theta}$  parametrizing a stationary marked Gibbs point process which is not necessarily a locally stable exponential family model. Sufficient conditions, expressed in terms of the local energy function, to establish strong consistency and asymptotic normality results of this estimator depending on a single realization are presented. These results constitute an extension of the ones obtained in Billiot et al. (2008) where the local energy function was assumed to be parametrically linear and stable. By applying these tools, we finally obtain the main results: consistency for both the Lennard-Jones model and the finite range Lennard-Jones model and asymptotic normality for the finite range Lennard-Jones model.

**Keywords:** stationary marked Gibbs point processes, maximum pseudo-likelihood, Lennard-Jones model

## 1 Introduction

The class of Gibbs point processes is interesting because it allows us to introduce and study interactions between points through the modelling of an associated energy function. Historical aspects of the mathematical theory are covered briefly in Kallenberg (1983). Among many models, a very well-known model coming from statistical physics is the Lennard-Jones model Lennard-Jones (1931). This model is a stationary pairwise interaction Gibbs point process where the local energy to insert a point  $x$  (of  $\mathbb{R}^2$ ) into a configuration of points  $\varphi$  is parametrized as follows

$$V^{LJ}(x|\varphi; \boldsymbol{\theta}) := \theta_1 + H^{LJ}(x|\varphi; \boldsymbol{\theta}) \quad \text{with } H^{LJ}(x|\varphi; \boldsymbol{\theta}) := \sum_{y \in \varphi} g^{LJ}(\|x - y\|; \boldsymbol{\theta}) \quad (1)$$

where  $g^{LJ}(\cdot; \boldsymbol{\theta})$  is defined for  $r > 0$  as

$$g^{LJ}(r; \boldsymbol{\theta}) := 4\theta_2 \left( \left( \frac{\theta_3}{r} \right)^{12} - \left( \frac{\theta_3}{r} \right)^6 \right),$$

where  $\theta_2, \theta_3 > 0$ . This simple mathematical potential is a tool for modelling a pair of neutral atoms or molecules subject to two distinct forces in the limit of large separation and small separation: an attractive force at long ranges (van der Waals force, or dispersion force) and a repulsive force at short ranges (the result of overlapping electron orbitals, referred to as Pauli repulsion from Pauli exclusion principle). The parameters  $\theta_2$  and  $\theta_3$  are often referred to as the depth potential and the (finite) distance at which the interparticle potential is zero. From a probabilistic point of view, the Lennard-Jones model is of major interest since it constitutes the main example of superstable, regular<sup>1</sup> and lower regular potentials studied by Ruelle (1970) who has proved the existence of ergodic measures for such models. In this paper, we also consider the finite range Lennard-Jones model which is defined by replacing the pairwise interaction function  $g^{LJ}(\cdot, \boldsymbol{\theta})$  in (1) by the following one

$$g^{LJ}(r; \boldsymbol{\theta}) := 4\theta_2 \left( \left( \frac{\theta_3}{r} \right)^{12} - \left( \frac{\theta_3}{r} \right)^6 \right) \mathbf{1}_{[0, D]}(r), \quad (2)$$

where  $D$  is a positive real representing the range of the model. Once you know that the model exists, you may be interested in estimating its parameters based on a single observation in a bounded window. One among many other methods of estimation is the maximization of the pseudo-likelihood. This paper aims at proving the following result:

**Theorem 1** *The maximum pseudo-likelihood estimate based on a single observation in a bounded window (assumed to grow to  $\mathbb{R}^2$ ) is consistent for both the Lennard-Jones model and the finite range Lennard-Jones model. The asymptotic normality holds only for the finite range Lennard-Jones model.*

Many proposals tried to estimate the energy function from the available point pattern data generated by some marked Gibbs point processes. If the energy belongs to a parametric family model, the most well-known methodology is the use of the likelihood function, see *e.g.* Møller and Waagepetersen (2003) and the references therein. The main drawback of this approach is that the likelihood function contains an unknown scaling factor whose value depends on the parameters and which is difficult to calculate. An alternative approach relies on the use of the pseudo-likelihood. This idea originated from Besag (1974) in the study of lattice processes. Besag et al. (1982) further considered this method for pairwise interaction point processes, while Jensen and Møller (1991) generalized it to the general class of marked Gibbs point processes.

In order to underline our theoretical contributions, let us present the different papers discussing the asymptotic properties of the maximum pseudo-likelihood estimator. Let us discuss the main two works: Billiot et al. (2008) obtained consistency and asymptotic normality for exponential family models of marked Gibbs point processes. In particular, they concentrate on models such that the local energy function has a finite range and is locally stable. This paper extends several papers (Jensen and Møller (1991), Jensen and Künsch (1994)) and includes a large class of examples of practical interest: area-interaction point process, Multi-Strauss marked point process based on the complete graph or the  $k$ -nearest-neighbors graph, or the Geyer's triplet point process to name a few. Another work has been undertaken by Mase. The consistency for superstable and lower regular potentials (introduced by Ruelle (1970)) is obtained in Mase (1995) for specific models with two parameters -the chemical potential and the inverse temperature- which can be viewed as particular exponential family models. Mase (2000) extended his work to the context of marked point processes and provided asymptotic normality by adding the assumption of finite range.

---

<sup>1</sup>A pairwise potential  $\Psi$  is said to be regular if  $\int_{\mathbb{R}^2} |1 - e^{\beta\Psi(x)}| dx < +\infty$  for some  $\beta > 0$ .

For convenience, a LJ-type model designates either the Lennard-Jones model or the finite range Lennard-Jones model. Let us underline that the previous studies do not allow us to derive asymptotic results for the maximum pseudo-likelihood estimate of the parameters of a LJ-type model since

- the model is superstable, regular and lower regular but not locally stable.
- the local energy function has a parametrically nonlinear form, not considered by Mase (1995, 2000). Actually, his results can be applied to obtain the same results as Theorem 1 in the very restrictive case where  $\theta_3$  is known.

Note that even if one linearizes the model by setting  $\lambda_2 := 4\theta_2\theta_3^{12}$  and  $\lambda_3 := 4\theta_2\theta_3^6$ , the results of Billiot et al. (2008) are not well-suited to the estimation of the parameter vector  $(\theta_1, \lambda_2, \lambda_3)$  due to the local stability requirement.

Based on this literature, the aim of this paper is to propose asymptotic results for non locally stable exponential family models. We keep the generality of Billiot et al. (2008) in the sense that sufficient conditions ensuring consistency and asymptotic normality are expressed in terms of the local energy. Finally, the main result of this paper is to fulfill the different assumptions for the LJ-type model.

The rest of the paper is decomposed as follows. Section 2 introduces some background and notation on marked Gibbs point processes and on the maximum pseudo-likelihood method. Asymptotic results are proposed in Section 3. The main part of this paper is Section 4 devoted to the proof of Theorem 1 which consists in verifying general assumptions described in Section 3 for the LJ-type model. Proofs have been postponed until Section 5.

## 2 Background and notation

### 2.1 Background on marked Gibbs point processes

For the sake of simplicity, the framework of this paper is restricted to two-dimensional marked Gibbs point processes. However, all the results must remain valid in the general  $d$ -dimensional ( $d \geq 1$ ) case. Define  $\mathcal{B}^2$  the Borel  $\sigma$ -algebra on  $\mathbb{R}^2$ ,  $\mathcal{B}_b^2$  the set of bounded Borel subsets of  $\mathbb{R}^2$  and  $\lambda^2$  the Lebesgue measure on  $\mathbb{R}^2$ . Denote also by  $\mathbb{M}$ ,  $\mathcal{M}$  and  $\lambda^m$  the mark space and its corresponding  $\sigma$ -algebra and probability measure. Let  $\mathbb{S} := \mathbb{R}^2 \times \mathbb{M}$ ,  $\mathcal{B} := \mathcal{B}^2 \otimes \mathcal{M}$  and  $\mu := \lambda^2 \otimes \lambda^m$  denote respectively the state space and its corresponding  $\sigma$ -algebra and measure.

For short, let us denote  $x^m = (x, m)$  for any  $x \in \mathbb{R}^2$  and any mark  $m \in \mathbb{M}$  and  $|\Lambda| := \lambda^2(\Lambda)$  for any  $\Lambda \in \mathcal{B}^2$ . In addition,  $|I|$  designates the number of elements of some countable set  $I$ ,  $\Lambda^c$  is the complementary of some set  $\Lambda$  in  $\mathbb{R}^2$  and  $\|\cdot\|$  is the  $\ell^2$ -norm. Let us define for all  $i = (i_1, i_2) \in \mathbb{Z}^2$ ,  $d > 0$  and  $\rho \geq 0$   $\Delta_i(d) := \{z \in \mathbb{R}^2, d(i_j - \frac{1}{2}) \leq z_j \leq d(i_j + \frac{1}{2}), j = 1, 2\}$  and  $\mathbb{B}(i, \rho) := \{k \in \mathbb{Z}^2 : |k - i| \leq \rho\}$  with  $|i| := \max(|i_1|, |i_2|)$ .

Let  $\tilde{\Omega}$  denote the set of so-called configurations -of marked points-  $\varphi := \{x_i^{m_i}\}_{i \in I}$  where  $I$  is a subset of  $\mathbb{N}$  and  $((x_i, m_i))_{i \in I}$  is a sequence of elements of  $\mathbb{S}$ . In particular, any element  $\varphi \in \tilde{\Omega}$  has the following representation  $\varphi = \sum_{i \in I} \delta_{x_i^{m_i}}$  as an integer-valued measure on  $\mathbb{S}$  such that for every  $F \in \mathcal{B}_b^2$ ,  $\varphi(F) \in \mathbb{N}$ , where  $\delta_{x^m}$  is the Dirac measure at some element  $x^m \in \mathbb{S}$ . The subset of  $\tilde{\Omega}$  with elements  $\varphi$  satisfying  $|\varphi| := \varphi(\mathbb{S}) < +\infty$  is denoted by  $\tilde{\Omega}_f$ . The space  $\tilde{\Omega}$  is equipped with the  $\sigma$ -algebra  $\mathcal{F}$  generated by the family of sets  $\{\varphi \in \tilde{\Omega} : \varphi(F) = n\}$  with  $n \in \mathbb{N}$  and  $F \in \mathcal{B}_b^2$ . For every  $F \in \mathcal{B}^2$  and  $\varphi \in \tilde{\Omega}$  represented as  $\varphi = \sum_{i \in I} \delta_{x_i^{m_i}}$ , one introduces  $\varphi_F := \sum_{i \in I, x_i^{m_i} \in F} \delta_{x_i^{m_i}}$  which can be viewed as the configuration of marked points of  $\varphi$  restricted to  $F$ . Furthermore, for every  $\Lambda \in \mathcal{B}_b^2$ ,  $\varphi_\Lambda$  conveniently denotes  $\varphi_{\Lambda \times \mathbb{M}}$ .

A marked point process is a  $\tilde{\Omega}$ -valued random variable, denoted by  $\Phi$ , with probability distribution  $P$  on  $(\tilde{\Omega}, \mathcal{F})$ . The intensity measure  $N_P$  of  $P$  is defined as a measure on  $\mathcal{B}^2$  such that for any  $F \in \mathcal{B}_b^2$ :

$$N_P(F) = \int_{\tilde{\Omega}} \varphi(F) P(d\varphi) := \mathbf{E}(\Phi(F)).$$

In the stationary case,  $N_P(F) = \nu_P \lambda^2(F)$  where  $\nu_P$  is called the intensity of  $P$ . A marked Gibbs point process is usually defined using a family of local specifications with respect to a weight process (often a stationary marked Poisson process with distribution  $Q$  and intensity  $\lambda_Q = 1$ ). Let  $\Lambda$  be a bounded region in  $\mathbb{R}^2$ . For such a process, given some configuration  $\varphi_{\Lambda^c}$  on  $\Lambda^c$ , the conditional probability on  $\Lambda$  is of the form, for any  $F \in \mathcal{F}$

$$\Pi_{\Lambda}(\varphi, F) = \left\{ \frac{1}{Z_{\Lambda}(\varphi)} \int_{\tilde{\Omega}_{\Lambda}} e^{-V(\psi|\varphi_{\Lambda^c})} \mathbf{1}_F(\psi \cup \varphi_{\Lambda^c}) Q_{\Lambda}(d\psi) \right\} \mathbf{1}_{R_{\Lambda}}(\varphi),$$

with the partition function

$$Z_{\Lambda}(\varphi) = \int_{\tilde{\Omega}_{\Lambda}} e^{-V(\psi|\varphi_{\Lambda^c})} Q_{\Lambda}(d\psi)$$

and  $R_{\Lambda} = \{\varphi \in \tilde{\Omega} : 0 < Z_{\Lambda}(\varphi) < +\infty\}$  where

$$\int f(\psi) Q_{\Lambda}(\psi) := e^{-\mu(\Lambda \times \mathbb{M})} \sum_{n=0}^{+\infty} \frac{1}{n!} \int f(\underbrace{\{x_1^{m_1}, \dots, x_n^{m_n}\}}_{\psi}) d\mu^{\otimes n}(x_1^{m_1}, \dots, x_n^{m_n}).$$

Let us define the subset of all admissible configurations  $\Omega := \{\varphi \in \tilde{\Omega} : \varphi \in \cap_{\Lambda \in \mathcal{B}_b^2} R_{\Lambda}\}$  and denote by  $\Omega_f := \tilde{\Omega}_f \cap \Omega$ . Whereas the finite energy function  $V(\varphi)$  (for any  $\varphi \in \Omega_f$ ) measures the cost of any configuration, the local energy  $V(\psi|\varphi)$  (for any  $\varphi, \psi \in \Omega_f$ ) represents the energy required to add the points of  $\psi$  in  $\varphi$ :

$$V(\psi|\varphi) = V(\psi \cup \varphi) - V(\varphi).$$

Let us notice that when  $\psi$  is a singleton  $\{x^m\}$ , we denote by a slight abuse  $V(x^m|\varphi)$  instead of  $V(\{x^m\}|\varphi)$ . It is well-known that the collection of probability kernels  $(\Pi_{\Lambda})_{\Lambda \in \mathcal{B}_b^2}$  satisfies the set of compatibility and measurability conditions which define a local specification in the Preston's sense (Preston (1976)). The main condition is the consistency :

$$\Pi_{\Lambda} \Pi_{\Lambda'} = \Pi_{\Lambda} \quad \text{for } \Lambda' \subset \Lambda.$$

Notice that some conditions are needed to ensure the existence of a probability measure  $P$  related to any local energy  $V$  and some weight process that satisfies the so-called Dobrushin-Lanford-Ruelle (D.L.R.) equations :

$$P(F|F_{\Lambda^c})(\varphi) = \Pi_{\Lambda}(\varphi, F) \quad \text{for } P \text{ a.e. } \varphi \in \Omega \quad \text{for any } \Lambda \in \mathcal{B}_b^2 \text{ and } F \in \mathcal{F}.$$

For the general theory of Gibbs point processes, the reader may refer to Kallenberg (1983); Stoyan et al. (1995) and the references therein.

For some finite configuration  $\varphi$  (resp. some set  $G$ ) and for all  $x \in \mathbb{R}^2$ ,  $\varphi_x$  (resp.  $G_x$ ) denotes the configuration  $\varphi$  (resp. the set  $G$ ) translated of  $x$ . Finally, in this work a non-marked point process can be viewed as a particular case of marked point processes with  $\mathbb{M} = \{0\}$ .

The framework of this paper is restricted to stationary marked Gibbs point processes based on an energy function invariant by translation,  $V(\varphi; \theta)$ , parametrized by some  $\theta \in \Theta$ , where  $\Theta$  is some compact set of  $\mathbb{R}^p$ . The following assumption [**Mod**] describes the general assumption on the models considered

**[Mod-E]** Our data consist in the realization of a point process with energy function  $V(\cdot; \boldsymbol{\theta}^*)$  such that the associated Gibbs measure  $P_{\boldsymbol{\theta}^*}$  exists. The vector  $\boldsymbol{\theta}^*$  is thus the true parameter to be estimated, assumed to be in  $\hat{\Theta}$ .

**[Mod-L]** There exists  $D \geq 0$  such that for all  $(m, \varphi) \in \mathbb{M} \times \Omega_f$

$$V(0^m | \varphi; \boldsymbol{\theta}) = V(0^m | \varphi_{B(0,D)}; \boldsymbol{\theta}),$$

where  $\mathcal{B}(x, r)$  denotes the ball centered at  $x \in \mathbb{R}^2$  with radius  $r > 0$ .

We do not want to discuss here general assumptions ensuring the existence of an ergodic measure, i.e. **[Mod-E]**.

## 2.2 Maximum pseudo-likelihood method

The idea of maximum pseudo-likelihood is due to Besag (1974) who first introduced the concept for Markov random fields in order to avoid the normalizing constant. This work was then widely extended and Jensen and Møller (1991) (Theorem 2.2) obtained a general expression for marked Gibbs point processes. Using our notation and up to a scalar factor, the pseudo-likelihood defined for a configuration  $\varphi$  and a domain of observation  $\Lambda$  is denoted by  $PL_{\Lambda}(\varphi; \boldsymbol{\theta})$  and given by

$$PL_{\Lambda}(\varphi; \boldsymbol{\theta}) = \exp\left(-\int_{\Lambda \times \mathbb{M}} e^{-V(x^m | \varphi; \boldsymbol{\theta})} \mu(dx^m)\right) \prod_{x^m \in \varphi_{\Lambda}} e^{-V(x^m | \varphi \setminus x^m; \boldsymbol{\theta})}. \quad (3)$$

It is more convenient to define (and work with) the log-pseudo-likelihood, denoted by  $LPL_{\Lambda}(\varphi; \boldsymbol{\theta})$ .

$$LPL_{\Lambda}(\varphi; \boldsymbol{\theta}) = -\int_{\Lambda \times \mathbb{M}} e^{-V(x^m | \varphi; \boldsymbol{\theta})} \mu(dx^m) - \sum_{x^m \in \varphi_{\Lambda}} V(x^m | \varphi \setminus x^m; \boldsymbol{\theta}). \quad (4)$$

For a practical point of view, the point process is assumed to be observed in a domain  $\Lambda_n \oplus D^{\vee} = \cup_{x \in \Lambda_n} \mathcal{B}(x, D^{\vee})$  for some  $D^{\vee} < +\infty$ . For the asymptotic normality result, it is also assumed that  $\Lambda_n \subset \mathbb{R}^2$  can be decomposed into  $\cup_{i \in I_n} \Delta_i$  where  $I_n = \mathbb{B}(0, n)$  and for  $i = (i_1, i_2) \in \mathbb{Z}^2$ ,  $\Delta_i = \Delta_i(\tilde{D})$  for some  $\tilde{D} > 0$  fixed from now on. As a consequence, as  $n \rightarrow +\infty$ ,  $\Lambda_n \rightarrow \mathbb{R}^2$  such that  $|\Lambda_n| \rightarrow +\infty$  and  $\frac{|\partial \Lambda_n|}{|\Lambda_n|} \rightarrow 0$ .

Define for any configuration  $\varphi$ ,  $U_n(\varphi; \boldsymbol{\theta}) = -\frac{1}{|\Lambda_n|} LPL_{\Lambda_n}(\varphi; \boldsymbol{\theta})$ . The maximum pseudo-likelihood estimate (MPLE) denoted by  $\hat{\boldsymbol{\theta}}_n(\varphi)$  is then defined by

$$\hat{\boldsymbol{\theta}}_n(\varphi) = \arg \max_{\boldsymbol{\theta} \in \Theta} LPL_{\Lambda_n}(\varphi; \boldsymbol{\theta}) = \arg \min_{\boldsymbol{\theta} \in \Theta} U_n(\varphi; \boldsymbol{\theta}).$$

We will also need the following basic notation:

- Gradient vector of  $U_n$ :  $\mathbf{U}_n^{(1)}(\varphi; \boldsymbol{\theta}) := -|\Lambda_n|^{-1} \mathbf{LPL}_{\Lambda_n}^{(1)}(\varphi; \boldsymbol{\theta})$  where for any bounded Borel set  $\Lambda$ ,  $\left(\mathbf{LPL}_{\Lambda}^{(1)}(\varphi; \boldsymbol{\theta})\right)_j$  is defined for  $j = 1, \dots, p$  by

$$\left(\mathbf{LPL}_{\Lambda}^{(1)}(\varphi; \boldsymbol{\theta})\right)_j = \int_{\Lambda \times \mathbb{M}} \frac{\partial V}{\partial \theta_j}(x^m | \varphi; \boldsymbol{\theta}) e^{-V(x^m | \varphi; \boldsymbol{\theta})} \mu(dx^m) - \sum_{x^m \in \varphi_{\Lambda}} \frac{\partial V}{\partial \theta_j}(x^m | \varphi \setminus x^m; \boldsymbol{\theta})$$

- Hessian matrix of  $U_n$ :  $\underline{U}_n^{(2)}(\varphi; \boldsymbol{\theta}) := -|\Lambda_n|^{-1} \underline{\mathbf{LPL}}_{\Lambda_n}^{(2)}(\varphi; \boldsymbol{\theta})$  where for any bounded Borel set  $\Lambda$ ,  $(\underline{\mathbf{LPL}}_{\Lambda}^{(2)}(\varphi; \boldsymbol{\theta}))_{j,k}$  is defined for  $j, k = 1, \dots, p$  by

$$\begin{aligned} (\underline{\mathbf{LPL}}_{\Lambda}^{(2)}(\varphi; \boldsymbol{\theta}))_{j,k} &= \int_{\Lambda \times \mathbb{M}} \left( \frac{\partial^2 V}{\partial \theta_j \partial \theta_k}(x^m | \varphi; \boldsymbol{\theta}) - \frac{\partial V}{\partial \theta_j}(x^m | \varphi; \boldsymbol{\theta}) \frac{\partial V}{\partial \theta_k}(x^m | \varphi; \boldsymbol{\theta}) \right) e^{-V(x^m | \varphi; \boldsymbol{\theta})} \mu(dx^m) \\ &\quad + \sum_{x^m \in \varphi_{\Lambda}} \frac{\partial V}{\partial \theta_j}(x^m | \varphi \setminus x^m; \boldsymbol{\theta}) \frac{\partial V}{\partial \theta_k}(x^m | \varphi \setminus x^m; \boldsymbol{\theta}) \end{aligned}$$

Finally, note that from the decomposition of the observation domain  $\Lambda_n$ , one has

$$\underline{U}_n^{(1)}(\varphi; \boldsymbol{\theta}) = |\Lambda_n|^{-1} \sum_{i \in I_n} \mathbf{LPL}_{\Delta_i}^{(1)}(\varphi; \boldsymbol{\theta}) \quad \text{and} \quad \underline{U}_n^{(2)}(\varphi; \boldsymbol{\theta}) = |\Lambda_n|^{-1} \sum_{i \in I_n} \underline{\mathbf{LPL}}_{\Delta_i}^{(2)}(\varphi; \boldsymbol{\theta}).$$

### 3 Asymptotic results of the MPLE

In the rest of the paper, let  $M$  be a random variable with distribution  $\lambda^m$ .

#### 3.1 Consistency of the MPLE

The assumption [C] gathers the following four assumptions:

[C1] For all  $\boldsymbol{\theta} \in \Theta$ ,

$$\mathbf{E} \left( e^{-V(0^M | \Phi; \boldsymbol{\theta})} \right) < +\infty \quad \text{and} \quad \mathbf{E} \left( |V(0^M | \Phi; \boldsymbol{\theta})| e^{-V(0^M | \Phi; \boldsymbol{\theta}^*)} \right) < +\infty.$$

[C2] Identifiability condition : there exists  $A_1, \dots, A_{\ell}$ ,  $\ell \geq p$  events in  $\mathcal{B}^2$  and  $A_1^m, \dots, A_{\ell}^m$  events in  $\mathcal{M}$  such that:

- the  $\ell$  events  $B_i := A_i^m \times A_i$  ( $i = 1, \dots, \ell$ ) are disjoint and satisfy  $\lambda^m \otimes P_{\boldsymbol{\theta}^*}(B_i) > 0$
- for all  $((m_1, \varphi_1), \dots, (m_{\ell}, \varphi_{\ell})) \in B_1 \times \dots \times B_{\ell}$

$$\begin{cases} D(0^{m_i} | \varphi_i; \boldsymbol{\theta}) = 0 \\ i = 1, \dots, \ell \end{cases} \Rightarrow \boldsymbol{\theta} = \boldsymbol{\theta}^*$$

where  $D(0^{m_i} | \varphi_i; \boldsymbol{\theta}) := V(0^{m_i} | \varphi_i; \boldsymbol{\theta}) - V(0^{m_i} | \varphi_i; \boldsymbol{\theta}^*)$

[C3] The function  $U_n(\varphi; \cdot)$  is continuous for  $P_{\boldsymbol{\theta}^*}$ -a.e.  $\varphi$ .

[C4] For all  $(m, \varphi) \in \mathbb{M} \times \Omega_f$   $V(0^m | \varphi; \boldsymbol{\theta})$  is continuously differentiable in  $\boldsymbol{\theta}$  and for all  $j = 1, \dots, p$

$$\mathbf{E} \left( \max_{\boldsymbol{\theta} \in \Theta} \left( \left| \frac{\partial V}{\partial \theta_j}(0^M | \Phi; \boldsymbol{\theta}) \right| e^{-V(0^M | \Phi; \boldsymbol{\theta})} \right)^2 \right) < +\infty.$$

**Theorem 2** Under the assumptions [Mod-E] and [C], for  $P_{\boldsymbol{\theta}^*}$ -almost every  $\varphi$ , the maximum pseudo-likelihood estimate  $\hat{\boldsymbol{\theta}}_n(\varphi)$  converges towards  $\boldsymbol{\theta}^*$  as  $n$  tends to infinity.

### 3.2 Asymptotic normality of the MPLE

For establishing the asymptotic normality of the MPLE we need to assume the four additional following assumptions:

[N1(k)] For all  $(m, \varphi) \in \mathbb{M} \times \Omega_f$ ,  $V(0^m | \varphi; \boldsymbol{\theta})$  is differentiable in  $\boldsymbol{\theta} = \boldsymbol{\theta}^*$  and for all  $k' \leq k$  and for all  $\lambda_1, \dots, \lambda_{k'}$   $k'$  positive integers such that  $\sum_{i=1}^{k'} \lambda_i = k$

$$\mathbf{E} \left( \int_{(\Delta \times \mathbb{M})^{k'}} \prod_{i=1}^{k'} \left| \frac{\partial V}{\partial \theta_j} (0^M | \Phi; \boldsymbol{\theta}^*) \right|^{\lambda_i} e^{-V(\{x_1, \dots, x_{k'}\} | \Phi; \boldsymbol{\theta}^*)} \mu(dx_i^{m_i}) \right) < +\infty.$$

[N2] There exists a neighbourhood  $\mathcal{V}(\boldsymbol{\theta}^*)$  of  $\boldsymbol{\theta}^*$  such that for all  $(m, \varphi) \in \mathbb{M} \times \Omega_f$   $V(0^m; \boldsymbol{\theta})$  is twice continuously differentiable in  $\boldsymbol{\theta} \in \mathcal{V}$  and, for all  $j, k = 1, \dots, p$  and  $\boldsymbol{\theta} \in \mathcal{V}(\boldsymbol{\theta}^*)$ ,

$$\mathbf{E} \left( \left| \frac{\partial^2 V}{\partial \theta_j \partial \theta_k} (0^M | \Phi; \boldsymbol{\theta}) \right| e^{-V(0^M | \Phi; \boldsymbol{\theta})} \right) < +\infty, \quad \mathbf{E} \left( \left| \frac{\partial^2 V}{\partial \theta_j \partial \theta_k} (0^M | \Phi; \boldsymbol{\theta}) \right| e^{-V(0^M | \Phi; \boldsymbol{\theta}^*)} \right) < +\infty,$$

and

$$\mathbf{E} \left( \left( \left| \frac{\partial V}{\partial \theta_j} (0^M | \Phi; \boldsymbol{\theta}) \right| e^{-V(0^M | \Phi; \boldsymbol{\theta})} \right)^2 \right) < +\infty.$$

[N3] There exists  $A_1, \dots, A_\ell$ ,  $\ell \geq p$  events in  $\mathcal{B}^2$  and  $A_1^m, \dots, A_\ell^m$  events in  $\mathcal{M}$  such that:

- the  $\ell$  events  $B_j := A_j^m \times A_j$  are disjoint and satisfy  $\lambda^m \otimes P_{\boldsymbol{\theta}^*}(B_j) > 0$
- for all  $((m_1, \varphi_1), \dots, (m_\ell, \varphi_\ell)) \in B_1 \times \dots \times B_\ell$  the  $(\ell, p)$  matrix with entries  $\frac{\partial V}{\partial \theta_j} (0^{m_i} | \varphi_i; \boldsymbol{\theta}^*)$  is injective.

[N4] For some  $\bar{D} > 0$  and some  $\bar{\Lambda} := \cup_{i \in \mathbb{B}(0, \lceil \frac{\bar{D}}{2} \rceil)}$ , there exists  $A_0, \dots, A_\ell$ ,  $\ell \geq p$  disjoint subevents of  $\bar{\Omega} := \left\{ \varphi \in \Omega : \varphi_{\Delta_i(\bar{D})} = \emptyset, 1 \leq |i| \leq 2 \left\lceil \frac{\bar{D}}{2} \right\rceil \right\}$  such that

- for  $j = 0, \dots, \ell$ ,  $P_{\boldsymbol{\theta}^*}(A_j) > 0$ .
- for all  $(\varphi_0, \dots, \varphi_\ell) \in A_0 \times \dots \times A_\ell$  the  $(\ell, p)$  matrix with entries  $\left( \mathbf{LPL}_{\bar{\Lambda}}^{(1)}(\varphi_i; \boldsymbol{\theta}^*) \right)_j - \left( \mathbf{LPL}_{\bar{\Lambda}}^{(1)}(\varphi_0; \boldsymbol{\theta}^*) \right)_j$  is injective.

The assumptions [N3] and [N4] will ensure (see Section 5 for more details) that the matrices  $\underline{\mathbf{U}}^{(2)}(\boldsymbol{\theta}^*)$  and  $\underline{\boldsymbol{\Sigma}}(\tilde{D}, \boldsymbol{\theta}^*)$  respectively defined by

$$\left( \underline{\mathbf{U}}^{(2)}(\boldsymbol{\theta}^*) \right)_{j,k} := \mathbf{E} \left( \frac{\partial V}{\partial \theta_j} (0^M | \Phi; \boldsymbol{\theta}^*) \frac{\partial V}{\partial \theta_k} (0^M | \Phi; \boldsymbol{\theta}^*) e^{-V(0^M | \Phi; \boldsymbol{\theta}^*)} \right) \quad (5)$$

and

$$\underline{\boldsymbol{\Sigma}}(\tilde{D}, \boldsymbol{\theta}^*) = \tilde{D}^{-2} \sum_{i \in \mathbb{B}(0, \lceil \frac{\tilde{D}}{2} \rceil)} \mathbf{E} \left( \mathbf{LPL}_{\Delta_0}^{(1)}(\Phi; \boldsymbol{\theta}^*) \mathbf{LPL}_{\Delta_i}^{(1)}(\Phi; \boldsymbol{\theta}^*)^T \right), \quad (6)$$

are definite positive. Let us note that the matrix  $\underline{\boldsymbol{\Sigma}}(\tilde{D}, \boldsymbol{\theta}^*)$  is actually independent of  $\tilde{D}$ . Therefore, one may note in the following  $\underline{\boldsymbol{\Sigma}}(\boldsymbol{\theta}^*) = \underline{\boldsymbol{\Sigma}}(\tilde{D}, \boldsymbol{\theta}^*) = \underline{\boldsymbol{\Sigma}}(1, \boldsymbol{\theta}^*)$ .



**Remark 1** *Let us make a little summary on the different notation. The real  $D$  denotes the range of the local energy function. The real  $\tilde{D}$  is associated to the geometry of the observation domain  $\Lambda_n$  and is chosen by the user independently of the value of  $D$ . Finally, the parameter  $\bar{D}$  occurs only in the verification of the assumption [N4]. It can be chosen independently of  $\tilde{D}$ . For the sake of simplicity, it is often chosen such that  $\tilde{D} > D$ .*

Observe also that, when the energy function is linear, the expressions of the assumptions [N1(k)] and [N2] are clearly simpler (see Billiot et al. (2008)) and that [C2] and [N3] are similar.

**Theorem 3** *Under the assumptions [Mod], [C], [N1(3)] [N2] and [N3], we have, for any fixed  $\tilde{D}$ , the following convergence in distribution as  $n \rightarrow +\infty$*

$$|\Lambda_n|^{1/2} \underline{U}^{(2)}(\theta^*) \left( \widehat{\theta}_n(\Phi) - \theta^* \right) \rightarrow \mathcal{N}(0, \underline{\Sigma}(\theta^*)), \quad (7)$$

where  $\underline{\Sigma}(\theta^*) = \underline{\Sigma}(1, \theta^*)$ . In addition under the assumptions [N1(4)] and [N4]

$$|\Lambda_n|^{1/2} \widehat{\underline{\Sigma}}_n(\Phi; D^\vee, \tilde{D}, \widehat{\theta}_n(\Phi))^{-1/2} \underline{U}_n^{(2)}(\Phi; \widehat{\theta}_n(\Phi)) \left( \widehat{\theta}_n(\Phi) - \theta^* \right) \rightarrow \mathcal{N}(0, \underline{I}_p), \quad (8)$$

where for some  $\theta$  and any configuration  $\varphi$ , the matrix  $\widehat{\underline{\Sigma}}_n(\varphi; D^\vee, \tilde{D}, \theta)$  is defined by

$$\widehat{\underline{\Sigma}}_n(\varphi; D^\vee, \tilde{D}, \theta) = |\Lambda_n|^{-1} \sum_{i \in I_n} \sum_{j \in \mathbb{B}(i, \lceil \frac{D^\vee}{\tilde{D}} \rceil) \cap I_n} \mathbf{LPL}_{\Delta_i}^{(1)}(\varphi; \theta) \mathbf{LPL}_{\Delta_j}^{(1)}(\varphi; \theta)^T. \quad (9)$$

In the following the assumption [N] will gather the assumptions [N1(4)], [N2], [N3] and [N4].

**Remark 2** *As noted, previously the matrix  $\underline{\Sigma}(\tilde{D}, \theta^*)$  does not depend on  $\tilde{D}$ , whereas its estimation depends strongly on  $\tilde{D}$ .*

## 4 Applications to the LJ-type model

This Section is devoted to the proof of Theorem 1. The LJ-type model is defined for some  $D \in ]0, +\infty]$  by

$$\begin{aligned} V^{LJ}(x|\varphi; \theta^*) &:= \theta_1^* + H^{LJ}(x|\varphi; \theta^*) \\ H^{LJ}(x|\varphi; \theta^*) &:= \sum_{y \in \varphi} g^{LJ}(\|y - x\|; \theta^*) \\ \text{and } g^{LJ}(r; \theta^*) &:= 4\theta_2^* \left( \left( \frac{\theta_3^*}{r} \right)^{12} - \left( \frac{\theta_3^*}{r} \right)^6 \right) \mathbf{1}_{[0, D]}(r), \end{aligned}$$

where  $\theta_2^*, \theta_3^* > 0$  and  $\theta_1^* \in \mathbb{R}$ . Recall that the cases  $D = +\infty$  and  $D < +\infty$  respectively correspond to the Lennard-Jones model and the Lennard-Jones model with finite range.

Let us first discuss the assumption [Mod-E]. Ruelle (1970) has proved the existence of an ergodic measure for superstable, regular and lower regular potentials. The Lennard-Jones model is known to be the characteristic example of such a family of models. Of course, [Mod-E] is also satisfied for the finite range Lennard-Jones model.

Thus, from Theorem 2 and 3, the proof of Theorem 1 consists in verifying Assumptions [C] for the LJ-type model and [N] only for the finite range Lennard-Jones model. In the following, we will deal with two types of assumptions:

- Integrability type assumptions, i.e. Assumptions [C1], [C4], [N1(4)] and [N2].
- Identifiability type assumptions, i.e. Assumptions [C2], [N3] and [N4].

Note that [C3] is obvious since  $g^{LJ}(r, \cdot)$  is continuous. For the integrability type assumptions, the following Lemma will be widely used.

**Lemma 4** *Let  $\Phi$  be a stationary pairwise interaction Gibbs point process with Hamiltonian  $H(\cdot)$  assumed to be superstable, regular and lower regular. Let  $H_i(x|\varphi) = \sum_{y \in \varphi} g_i(\|x-y\|)$  for  $i = 1, 2$ . Assume that there exists  $\varepsilon > 0$  such that there exists a positive and decreasing function  $g(\cdot)$  such that  $g_\varepsilon(r) := g_2(r) - \varepsilon|g_1(r)| \geq -g(r)$  for all  $r > 0$  and  $\int_0^{+\infty} rg(r)dr < +\infty$ . Then for all  $k \geq 0$ ,*

$$\mathbf{E} \left( |H_1(0|\Phi)|^k e^{-H_2(0|\Phi)} \right) < +\infty.$$

**Proof.** For all finite configuration  $\varphi$

$$\begin{aligned} |H_1(0|\varphi)|^k e^{-H_2(0|\varphi)} &= |H_1(0|\varphi)|^k e^{-\varepsilon|H_1(0|\varphi)|} e^{-(H_2(0|\varphi) - \varepsilon H_1(0|\varphi))} \\ &\leq c(\varepsilon, k) e^{-(H_2(0|\varphi) - \varepsilon H_1(0|\varphi))}, \quad \text{with } c(\varepsilon, k) = \left( \frac{k}{\varepsilon e} \right)^k \\ &\leq c(\varepsilon, k) e^{-H_\varepsilon(0|\varphi)}, \end{aligned}$$

where

$$H_\varepsilon(0|\varphi) := \sum_{x \in \varphi} g_\varepsilon(\|x\|).$$

Now, the assumptions made on the function  $g_\varepsilon$  ensure that the potential associated to the local hamiltonian  $H_\varepsilon(0|\varphi)$  is lower regular. We may now apply the same argument as in Lemma 3 of Mase (1995) to prove the integrability of the random variable  $e^{-H_\varepsilon(0|\Phi)}$ . ■

Let us denote by

$$\theta_i^{\inf} := \inf_{\theta \in \Theta} \theta_i \quad \theta_i^{\sup} := \sup_{\theta \in \Theta} \theta_i \quad \theta^{\inf} := \min(\theta_2^{\inf}, \theta_3^{\inf}) \quad \text{and} \quad \theta^{\sup} := \max(\theta_2^{\sup}, \theta_3^{\sup})$$

Since  $\Theta$  is a compact set of  $\mathbb{R} \times (]0, +\infty[)^2$  then  $\theta^{\inf} > 0$  and  $\theta^{\sup} < +\infty$ .

## 4.1 Assumptions [C]

### 4.1.1 Assumption [C1]

The first part is a direct application of Lemma 4. For the second part, one has to prove that for all  $\theta \in \Theta$

$$\mathbf{E} \left( H^{LJ}(0|\Phi; \theta) e^{-H^{LJ}(0|\Phi; \theta^*)} \right) < +\infty$$

Let  $g_\varepsilon(r) = g^{LJ}(r; \theta^*) - \varepsilon|g^{LJ}(r; \theta)|$ . We have

$$g_\varepsilon(r) := \begin{cases} 4\theta_2^* \left( \frac{(\theta_3^*)^{12} - \varepsilon \frac{\theta_2^*}{\theta_3^*} \theta_3^{12}}{r^{12}} - \frac{(\theta_3^*)^6 - \varepsilon \frac{\theta_2^*}{\theta_3^*} \theta_3^6}{r^6} \right) & \text{if } r \leq \theta_3 \\ 4\theta_2^* \left( \frac{(\theta_3^*)^{12} + \varepsilon \frac{\theta_2^*}{\theta_3^*} \theta_3^{12}}{r^{12}} - \frac{(\theta_3^*)^6 + \varepsilon \frac{\theta_2^*}{\theta_3^*} \theta_3^6}{r^6} \right) & \text{if } r \geq \theta_3 \end{cases}$$

which satisfies the assumptions of Lemma 4 as soon as  $\varepsilon < \left( \frac{\theta_3^*}{\theta_3} \right)^{12} \frac{\theta_2^*}{\theta_2}$ , that is, as soon as  $\varepsilon < \left( \frac{\theta^{\inf}}{\theta^{\sup}} \right)^{13}$ .

### 4.1.2 Assumption [C2]

Let us denote for  $n \geq 1$ ,  $C_n = \mathcal{B}(0, n) \setminus \mathcal{B}(0, n-1)$  and define for  $m, n \geq 1$  the following configuration sets

$$\begin{aligned} U_{m,n} &= \{\varphi \in \Omega : |\varphi_{C_n}| \leq m|C_n|\} \\ U_m &= \bigcap_{n \geq 1} U_{m,n}. \end{aligned}$$

In order to prove [C2], we need the following Lemma.

**Lemma 5** *Let  $R \in \mathbb{R}^+$ ,  $\theta \in \Theta$  and  $\varphi \in U_m$ , let us denote by*

$$Z(\varphi, R; \theta) := \sum_{x \in \varphi_{\mathcal{B}(0, R)^c}} g^{LJ}(\|x\|; \theta),$$

then for all  $\delta > 0$  there exists  $R_0$  such that for all  $R \geq R_0$ ,  $|Z(\varphi, R; \theta)| \leq \delta$ .

**Proof.**

$$\begin{aligned} |Z(\varphi, R; \theta)| &= \left| \sum_{x \in \varphi_{\mathcal{B}(0, R)^c}} g^{LJ}(\|x\|; \theta) \right| \leq \sum_{n \geq \lceil R \rceil} \sum_{x \in \varphi_{C_n}} |g^{LJ}(\|x\|; \theta)| \\ &\leq \sum_{n \geq \lceil R \rceil} |\varphi_{C_n}| \times \sup_{x \in C_n} |g^{LJ}(\|x\|; \theta^*)|. \end{aligned}$$

There exists a constant  $k = k(R)$  such that for all  $n \geq \lceil R \rceil$ ,  $\sup_{x \in C_n} |g^{LJ}(\|x\|; \theta^*)| \leq kn^{-6}$ . Therefore,

$$\left| \sum_{x \in \varphi_{\mathcal{B}(0, R)^c}} g^{LJ}(\|x\|; \theta) \right| \leq km \sum_{n \geq \lceil R \rceil} |C_n| \times n^{-6} = \mathcal{O} \left( \sum_{n \geq \lceil R \rceil} n^{-5} \right),$$

which leads to the result since the previous series is convergent. ■

Let  $\theta \in \Theta \setminus \theta^*$  and consider the following configuration sets defined for  $k \geq 1$  and for  $\eta$  small enough by

$$\begin{aligned} A_0 &= \{\varphi \in \Omega : \varphi(\mathcal{B}(0, D)) = 0\} \\ A_k(\eta) &= \left\{ \varphi \in \Omega : \varphi(\mathcal{B}(0, D)) = \varphi \left( \mathcal{B} \left( \left( 0, \frac{D}{k^{1/12}} \right), \eta \right) \right) = 1 \right\}, \end{aligned}$$

where  $D$  is any positive real for the Lennard-Jones model and corresponds to the range of the function  $g^{LJ}(\cdot)$  for the finite range Lennard-Jones model. There exists  $m \geq 1$  such that for all  $\eta > 0$  and for  $k = 2, 4$

$$P_{\theta^*}(A_0 \cap U_m) > 0 \quad \text{and} \quad P_{\theta^*}(A_k(\eta) \cap U_m) > 0.$$

Now, let  $\varphi_0 \in A_0 \cap U_m$ ,  $\varphi_2 \in A_2(\eta) \cap U_m$  and  $\varphi_4 \in A_4(\eta) \cap U_m$ . First,

$$D(0|\varphi_0; \theta) = \theta_1 - \theta_1^* + Z(\varphi_0, D; \theta) - Z(\varphi_0, D; \theta^*) = 0.$$

For the Lennard-Jones model, according to Lemma 5 one has, for  $D$  large enough,

$$|Z(\varphi_0, D; \theta) - Z(\varphi_0, D; \theta^*)| \leq \frac{1}{2} |\theta_1 - \theta_1^*|.$$

Hence for  $\eta$  small enough, and for both models

$$\begin{aligned} 0 &= |D(0|\varphi_0; \boldsymbol{\theta})| \\ &\geq |\theta_1 - \theta_1^*| - |Z(\varphi_0, D; \boldsymbol{\theta}) - Z(\varphi_0, D; \boldsymbol{\theta}^*)| \\ &\geq \frac{1}{2}|\theta_1 - \theta_1^*|, \end{aligned}$$

which leads to  $\theta_1 = \theta_1^*$ . Moreover,

$$\begin{aligned} D(0|\varphi_2; \boldsymbol{\theta}) &= 4\theta_2 \left( 2 \left( \frac{\theta_3}{D} \right)^{12} - \sqrt{2} \left( \frac{\theta_3}{D} \right)^6 \right) - 4\theta_2^* \left( 2 \left( \frac{\theta_3^*}{D} \right)^{12} - \sqrt{2} \left( \frac{\theta_3^*}{D} \right)^6 \right) \\ &\quad + f_2(\varphi_2) + Z(\varphi_2, D; \boldsymbol{\theta}) - Z(\varphi_2, D; \boldsymbol{\theta}^*) \\ D(0|\varphi_4; \boldsymbol{\theta}) &= 4\theta_2 \left( 4 \left( \frac{\theta_3}{D} \right)^{12} - 2 \left( \frac{\theta_3}{D} \right)^6 \right) - 4\theta_2^* \left( 4 \left( \frac{\theta_3^*}{D} \right)^{12} - 2 \left( \frac{\theta_3^*}{D} \right)^6 \right) \\ &\quad + f_4(\varphi_4) + Z(\varphi_4, D; \boldsymbol{\theta}) - Z(\varphi_4, D; \boldsymbol{\theta}^*), \end{aligned}$$

where for any  $\varphi_k \in A_k(\eta)$  ( $k = 2, 4$ ), there exists a positive function  $\tilde{f}_k(\eta)$  converging towards zero as  $\eta \rightarrow 0$  such that  $|f_k(\varphi_k)|$  is bounded by  $\tilde{f}_k(\eta)$ . Now, we have

$$\begin{aligned} 2D(0|\varphi_2; \boldsymbol{\theta}) - D(0|\varphi_4; \boldsymbol{\theta}) &= \frac{4(2 - 2\sqrt{2})}{D^6} (\theta_2\theta_3^6 - \theta_2^*\theta_3^{*6}) + 2f(\varphi_2) - f_4(\varphi_4) + Z'(\varphi_2, \varphi_4, D; \boldsymbol{\theta}, \boldsymbol{\theta}^*) \\ &= 0 \end{aligned}$$

with

$$Z'(\varphi_2, \varphi_4, D; \boldsymbol{\theta}, \boldsymbol{\theta}^*) := 2(Z(\varphi_2, D; \boldsymbol{\theta}) - Z(\varphi_2, D; \boldsymbol{\theta}^*)) - (Z(\varphi_4, D; \boldsymbol{\theta}) - Z(\varphi_4, D; \boldsymbol{\theta}^*)).$$

For  $\eta$  small enough, we have, for any  $\varphi_k \in A_k(\eta)$  ( $k = 2, 4$ ),

$$|2f(\varphi_2) - f_4(\varphi_4)| \leq 2\tilde{f}_2(\eta) + \tilde{f}_4(\eta) \leq \frac{1}{4} \left| \frac{4(2 - 2\sqrt{2})}{D^6} \right| |\theta_2\theta_3^6 - \theta_2^*\theta_3^{*6}|.$$

For the finite range Lennard-Jones model,  $Z'(\varphi_2, \varphi_4, D; \boldsymbol{\theta}, \boldsymbol{\theta}^*) = 0$ . For the Lennard-Jones model, according to Lemma 5, one has for  $D$  large enough

$$|Z'(\varphi_2, \varphi_4, D; \boldsymbol{\theta}, \boldsymbol{\theta}^*)| \leq \frac{1}{4} \left| \frac{4(2 - 2\sqrt{2})}{D^6} \right| |\theta_2\theta_3^6 - \theta_2^*\theta_3^{*6}|.$$

Hence for  $\eta$  small enough, and for both models

$$\begin{aligned} 0 &= \left| \frac{4(2 - 2\sqrt{2})}{D^6} (\theta_2\theta_3^6 - \theta_2^*\theta_3^{*6}) + 2f(\varphi_2) - f_4(\varphi_4) + Z'(\varphi_2, \varphi_4, D; \boldsymbol{\theta}, \boldsymbol{\theta}^*) \right| \\ &\geq \left| \frac{4(2 - 2\sqrt{2})}{D^6} \right| |\theta_2\theta_3^6 - \theta_2^*\theta_3^{*6}| - |2f(\varphi_2) - f_4(\varphi_4)| - |Z'(\varphi_2, \varphi_4, D; \boldsymbol{\theta}, \boldsymbol{\theta}^*)| \\ &\geq \frac{1}{2} \left| \frac{4(2 - 2\sqrt{2})}{D^6} \right| |\theta_2\theta_3^6 - \theta_2^*\theta_3^{*6}| \end{aligned}$$

leading to  $\theta_2\theta_3^6 = \theta_2^*\theta_3^{*6}$ . By considering the combination  $\sqrt{2}D(0|\varphi_2; \boldsymbol{\theta}) - D(0|\varphi_4; \boldsymbol{\theta})$  and using similar arguments as previously, one obtains:  $\theta_2\theta_3^{12} = \theta_2^*\theta_3^{*12}$ . By computing the ratio of the two last equations, one obtains  $\theta_3 = \theta_3^*$  and then  $\theta_2 = \theta_2^*$ .

### 4.1.3 Assumption [C4]

For all  $\varphi \in \Omega_f$  and for any  $\boldsymbol{\theta} \in \Theta$ ,  $V^{LJ}(0|\varphi; \boldsymbol{\theta})$  is clearly differentiable in  $\boldsymbol{\theta}$ . First, note that [C4] is trivial for  $j = 1$ . For  $j = 2, 3$ , let us define:

$$X_j(\varphi; \boldsymbol{\theta}) := \left| \frac{\partial V^{LJ}}{\partial \theta_j}(0|\varphi; \boldsymbol{\theta}) \right| e^{-V^{LJ}(0|\varphi; \boldsymbol{\theta})}.$$

Our aim will be to prove that for  $j = 2, 3$  and for all  $k > 0$

$$\mathbf{E} \left( \max_{\boldsymbol{\theta} \in \Theta} X_j(\Phi; \boldsymbol{\theta})^k \right) < +\infty. \quad (10)$$

In particular, the Assumption [C4] corresponds to (10) with  $k = 2$ . Let us notice that for all  $\varphi \in \Omega$  and for all  $\boldsymbol{\theta} \in \Theta$

$$V^{LJ}(0|\varphi; \boldsymbol{\theta}) \geq V^{\text{inf}}(0|\varphi) := \theta^{\text{inf}} + \sum_{x \in \varphi} g^{\text{inf}}(\|x\|),$$

with for some  $r > 0$ ,  $g^{\text{inf}}(r) := 4\theta^{\text{inf}} \left( \frac{(\theta^{\text{inf}})^{12}}{r^{12}} - \frac{(\theta^{\text{sup}})^6}{r^6} \right)$ . Let us also underline that for  $j = 2, 3$

$$\frac{\partial g^{LJ}}{\partial \theta_j}(r; \boldsymbol{\theta}) \geq \tilde{g}_j^{\text{inf}}(r) \quad \text{with} \quad \tilde{g}_j^{\text{inf}}(r) := \begin{cases} 4 \left( \frac{(\theta^{\text{inf}})^{12}}{r^{12}} - \frac{(\theta^{\text{sup}})^6}{r^6} \right) & \text{if } j = 2, \\ 4m \left( \frac{12(\theta^{\text{inf}})^{11}}{r^{12}} - \frac{6(\theta^{\text{sup}})^5}{r^6} \right) & \text{if } j = 3. \end{cases}$$

Therefore, by defining  $\tilde{V}_j^{\text{inf}}(0|\varphi) := \sum_{x \in \varphi} \tilde{g}_j^{\text{inf}}(\|x\|)$ , the result (10) will be ensured by proving

$$\mathbf{E} \left( \tilde{V}_j^{\text{inf}}(0|\Phi) e^{-V^{\text{inf}}(0|\Phi)} \right) < +\infty.$$

According to Lemma 4, in order to prove this, let us denote by  $g_{j,\varepsilon}(\cdot)$  the function defined for  $j = 2, 3$ , for some  $\varepsilon > 0$  and for  $r > 0$  by  $g_{j,\varepsilon}(r) = \tilde{g}_j^{\text{inf}}(r) - \varepsilon |g^{\text{inf}}(r)|$ . On the one hand, one has

$$g_{2,\varepsilon}(r) = \begin{cases} 4 \left( \frac{(\theta^{\text{inf}})^{13} - \varepsilon (\theta^{\text{inf}})^{12}}{r^{12}} - \frac{\theta^{\text{inf}} (\theta^{\text{sup}})^6 - \varepsilon (\theta^{\text{sup}})^6}{r^6} \right) & \text{if } r \leq \frac{(\theta^{\text{inf}})^2}{\theta^{\text{sup}}}, \\ 4 \left( \frac{(\theta^{\text{inf}})^{13} + \varepsilon (\theta^{\text{inf}})^{12}}{r^{12}} - \frac{\theta^{\text{inf}} (\theta^{\text{sup}})^6 + \varepsilon (\theta^{\text{sup}})^6}{r^6} \right) & \text{if } r \geq \frac{(\theta^{\text{inf}})^2}{\theta^{\text{sup}}}, \end{cases}$$

which satisfies the assumptions of Lemma 4 as soon as  $\varepsilon < \theta^{\text{inf}}$ . On the other hand

$$g_{3,\varepsilon}(r) = \begin{cases} 4\theta^{\text{inf}} \left( \frac{(\theta^{\text{inf}})^{12} - 12\varepsilon (\theta^{\text{inf}})^{11}}{r^{12}} - \frac{(\theta^{\text{sup}})^6 - 6\varepsilon (\theta^{\text{sup}})^5}{r^6} \right) & \text{if } r \leq \left( 2 \frac{(\theta^{\text{inf}})^{11}}{(\theta^{\text{sup}})^5} \right)^{1/6}, \\ 4\theta^{\text{inf}} \left( \frac{(\theta^{\text{inf}})^{12} + 12\varepsilon (\theta^{\text{inf}})^{11}}{r^{12}} - \frac{(\theta^{\text{sup}})^6 + 6\varepsilon (\theta^{\text{sup}})^5}{r^6} \right) & \text{if } r \geq \left( 2 \frac{(\theta^{\text{inf}})^{11}}{(\theta^{\text{sup}})^5} \right)^{1/6}, \end{cases}$$

which satisfies the assumptions of Lemma 4 as soon as  $\varepsilon < \theta^{\text{inf}}/12$ , which ends the proof.

## 4.2 Assumptions [N]

### 4.2.1 Assumption [N1(4)]

Actually, the LJ-type model satisfies [N1(k)] for any  $k \geq 1$ . In order to prove this, let us first present two auxiliary lemmas.

**Lemma 6** *Let  $\Phi$  a stationary pairwise interaction point process with local energy function defined by*

$$V(x|\varphi; \boldsymbol{\theta}) = \theta_1 + H(x|\varphi; \boldsymbol{\theta}) \quad \text{with} \quad H(x|\varphi; \boldsymbol{\theta}) = \sum_{y \in \varphi} g(\|y - x\|; \boldsymbol{\theta}).$$

Let  $K < +\infty$  and let  $x_1, \dots, x_K \in \mathbb{R}^2 \setminus \varphi$ ,  $x_i \neq x_j$  for  $i, j = 1, \dots, K$  (where  $K < +\infty$ ), then

$$\begin{aligned} H(\{x_1, \dots, x_K\}|\varphi; \boldsymbol{\theta}) &= \sum_{k=1}^K H(x_k|\varphi; \boldsymbol{\theta}) + H(\{x_1, \dots, x_K\}; \boldsymbol{\theta}) \\ V(\{x_1, \dots, x_K\}|\varphi; \boldsymbol{\theta}) &= \sum_{k=1}^K V(x_k|\varphi; \boldsymbol{\theta}) + H(\{x_1, \dots, x_K\}; \boldsymbol{\theta}) \end{aligned}$$

This result comes from the definition of the local energy.

**Lemma 7** *Using the same notation and under the same assumptions of Lemma 6, assume that there exists  $g_{\min}$  such that for all  $r > 0$  and any  $\boldsymbol{\theta} \in \Theta$ ,  $g(r; \boldsymbol{\theta}) \geq g_{\min}$ , then*

$$e^{-V(\{x_1, \dots, x_K\}|\varphi; \boldsymbol{\theta})} \leq c_K \prod_{k=1}^K e^{-V(x_k|\varphi; \boldsymbol{\theta})} \quad \text{with} \quad c_K = e^{-\frac{K(K-1)}{2}g_{\min}}$$

**Proof.** The proof is immediate since

$$H(\{x_1, \dots, x_K\}; \boldsymbol{\theta}) = \sum_{i < j} g(\|x_i - x_j\|; \boldsymbol{\theta}) \geq \frac{K(K-1)}{2}g_{\min}.$$

■

Let  $k \geq 1$ ,  $k' \leq k$  and let  $\lambda_1, \dots, \lambda_{k'}$   $k'$  positive integers such that  $\sum_{i=1}^{k'} \lambda_i = k$  and define the random variable

$$A(\Phi) := \int_{(\Delta \times \mathbb{M})^{k'}} \prod_{i=1}^{k'} \left| \frac{\partial V}{\partial \theta_j} (x_i|\Phi; \boldsymbol{\theta}^*) \right|^{\lambda_i} e^{-V(\{x_1, \dots, x_{k'}\}|\Phi; \boldsymbol{\theta}^*)} dx_i.$$

From Lemma 7, we have

$$\begin{aligned} \mathbf{E}(A(\Phi)) &\leq \mathbf{E} \left( c_{k'} \int_{(\Delta \times \mathbb{M})^{k'}} \prod_{i=1}^{k'} \left| \frac{\partial V}{\partial \theta_j} (x_i|\Phi; \boldsymbol{\theta}^*) \right|^{\lambda_i} e^{-V(x_i|\Phi; \boldsymbol{\theta}^*)} dx_i \right) \\ &= c_{k'} \int_{(\Delta \times \mathbb{M})^{k'}} \mathbf{E} \left( \prod_{i=1}^{k'} \left| \frac{\partial V}{\partial \theta_j} (x_i|\Phi; \boldsymbol{\theta}^*) \right|^{\lambda_i} e^{-V(x_i|\Phi; \boldsymbol{\theta}^*)} \right) dx_1 \dots dx_{k'} \\ &\leq c_{k'} \prod_{i=1}^{k'} \int_{(\Delta \times \mathbb{M})^{k'}} \left| \frac{\partial V}{\partial \theta_j} (x|\Phi; \boldsymbol{\theta}^*) \right|^k e^{-\frac{\lambda_i}{k}V(x|\Phi; \boldsymbol{\theta}^*)} dx, \end{aligned}$$

by using Hölder's inequality. The result is then a simple consequence of (10) and Lemma 4.

### 4.2.2 Assumption [N2]

For all  $\varphi \in \Omega_f$ , it is clear that for all  $\boldsymbol{\theta} \in \Theta$ ,  $V(0|\varphi; \boldsymbol{\theta})$  is twice continuously differentiable in  $\boldsymbol{\theta}$ . According to Lemma 4 and the fact that [N1(4)] is satisfied, it is sufficient to prove that for all  $j, k = 1, 2, 3$

$$\mathbf{E} \left( \left| \frac{\partial^2 V^{LJ}}{\partial \theta_j \partial \theta_k} (0|\Phi; \boldsymbol{\theta}) \right| e^{-V^{LJ}(0|\Phi; \boldsymbol{\theta})} \right) < +\infty.$$

This is obvious when either  $j$  or  $k$  equals 1 and when  $j = k = 2$  (since  $\frac{\partial^2 g^{LJ}}{(\partial \theta_2)^2}(r; \cdot) = 0$ ). Now, for the other cases, define for  $\boldsymbol{\theta} \in \Theta$   $g_{j,k,\varepsilon}(r) := g^{LJ}(r; \boldsymbol{\theta}) - \varepsilon \left| \frac{\partial^2 g^{LJ}}{\partial \theta_j \partial \theta_k}(r; \boldsymbol{\theta}) \right|$ . We have

$$g_{2,3,\varepsilon}(r) = g_{3,2,\varepsilon}(r) = \begin{cases} 4 \left( \frac{\theta_2 \theta_3^{12} - 12\varepsilon \theta_3^{11}}{r^{12}} - \frac{\theta_3^6 - 6\varepsilon \theta_3^5}{r^6} \right) & \text{if } r \leq 2^{1/6} \\ 4 \left( \frac{\theta_2 \theta_3^{12} + 12\varepsilon \theta_3^{11}}{r^{12}} - \frac{\theta_3^6 + 6\varepsilon \theta_3^5}{r^6} \right) & \text{otherwise} \end{cases}$$

which satisfies the assumptions of Lemma 4 as soon as  $\varepsilon < \frac{\theta_2 \theta_3}{12}$ , that is, as soon as  $\varepsilon < \frac{(\theta^{\text{inf}})^2}{12}$ . Finally,

$$g_{3,3,\varepsilon}(r) = \begin{cases} 4 \left( \frac{\theta_2 \theta_3^{12} - 132\varepsilon \theta_3^{10}}{r^{12}} - \frac{\theta_2 \theta_3^6 - 30\varepsilon \theta_3^4}{r^6} \right) & \text{if } r \leq \left(\frac{132}{30}\right)^{1/6} \theta_3 \\ 4 \left( \frac{\theta_2 \theta_3^{12} + 132\varepsilon \theta_3^{10}}{r^{12}} - \frac{\theta_2 \theta_3^6 + 30\varepsilon \theta_3^4}{r^6} \right) & \text{otherwise} \end{cases}$$

which satisfies the assumptions of Lemma 4 as soon as  $\varepsilon < \frac{\theta_2 \theta_3^2}{132}$ , that is, as soon as  $\varepsilon < \frac{(\theta^{\text{inf}})^3}{132}$ .

### 4.2.3 Assumption [N3]

Let  $\mathbf{y} = (y_1, y_2, y_3) \in \mathbb{R}^3$  and let us denote for any finite configuration  $\varphi$  by  $g(\mathbf{y}, \varphi) := \mathbf{y}^T \mathbf{V}_{LJ}^{(1)}(0|\varphi; \boldsymbol{\theta}^*)$ . Let  $\varphi_0 \in A_0$  and  $\varphi_k(\eta) \in A_k(\eta)$  ( $k = 2, 4$ ) as in the previous section. Assume  $g(\mathbf{y}, \varphi_k) = 0$  for  $k = 0, 2, 4$ . Since,  $g(\mathbf{y}, \varphi_0) = y_1$ , we have  $y_1 = 0$ . Now,

$$\begin{aligned} g(\mathbf{y}, \varphi_2) &= 4y_2 \left( 2 \left( \frac{\theta_3^*}{D} \right)^{12} - \sqrt{2} \left( \frac{\theta_3^*}{D} \right)^6 \right) + 4y_3 \theta_2^* \left( 2 \frac{12\theta_3^{*11}}{D^{12}} - \sqrt{2} \frac{6\theta_3^{*5}}{D^6} \right) + f_2(\mathbf{y}, \varphi_2) \\ g(\mathbf{y}, \varphi_4) &= 4y_2 \left( 4 \left( \frac{\theta_3^*}{D} \right)^{12} - 2 \left( \frac{\theta_3^*}{D} \right)^6 \right) + 4y_3 \theta_2^* \left( 4 \frac{12\theta_3^{*11}}{D^{12}} - 2 \frac{6\theta_3^{*5}}{D^6} \right) + f_4(\mathbf{y}, \varphi_4), \end{aligned}$$

where for any  $\varphi_k \in A_k(\eta)$  ( $k = 2, 4$ ), there exists a positive function  $\tilde{f}_k(\mathbf{y}, \eta)$  converging towards zero as  $\eta \rightarrow 0$  such that  $|f_k(\mathbf{y}, \varphi_k)|$  is bounded  $\tilde{f}_k(\mathbf{y}, \eta)$ . Now, we have

$$2g(\mathbf{y}, \varphi_2) - g(\mathbf{y}, \varphi_4) = 4(2 - 2\sqrt{2}) \frac{\theta_3^{*5}}{D^6} (\theta_3^* y_2 + 6\theta_2^* y_3) + 2f_2(\mathbf{y}, \varphi_2) - f_4(\mathbf{y}, \varphi_4) = 0.$$

For  $\eta$  small enough, we have, for any  $\varphi_k \in A_k(\eta)$  ( $k = 2, 4$ ),

$$|2f(\mathbf{y}, \varphi_2) - f_4(\mathbf{y}, \varphi_4)| \leq 2|\tilde{f}_2(\mathbf{y}, \eta)| + |\tilde{f}_4(\mathbf{y}, \eta)| \leq \frac{1}{2} \left| 4(2 - 2\sqrt{2}) \frac{\theta_3^{*5}}{D^6} (\theta_3^* y_2 + 6\theta_2^* y_3) \right|.$$

Hence for  $\eta$  small enough,

$$0 = |2g(\mathbf{y}, \varphi_2) - g(\mathbf{y}, \varphi_4)| \geq \frac{1}{2} \left| 4(2 - 2\sqrt{2}) \frac{\theta_3^{*5}}{D^6} (\theta_3^* y_2 + 6\theta_2^* y_3) \right|,$$

leading to the equation  $\theta_3^* y_2 + 6\theta_2^* y_3 = 0$ . By considering the linear combination  $\sqrt{2}g(\mathbf{y}, \varphi_2) - g(\mathbf{y}, \varphi_4)$ , we may obtain the equation  $\theta_3^* y_2 + 12\theta_2^* y_3 = 0$  with similar arguments. Both equations lead to  $y_2 = y_3 = 0$ .

#### 4.2.4 Assumption [N4]

The assumption [N4] may be rewritten for all  $k = 1, \dots, \ell$  and for all  $\varphi_k \in A_k$  and  $\varphi_0 \in A_0$ :

$$\left( \forall \mathbf{y} \in \mathbb{R}^3, \mathbf{y}^T \left( \mathbf{LPL}_{\bar{\Lambda}}^{(1)}(\varphi_k; \boldsymbol{\theta}^*) - \mathbf{LPL}_{\bar{\Lambda}}^{(1)}(\varphi_0; \boldsymbol{\theta}^*) \right) = \mathbf{y}^T (\mathbf{L}(\varphi_k; \boldsymbol{\theta}^*) - \mathbf{R}(\varphi_k; \boldsymbol{\theta}^*)) = 0 \right) \implies \mathbf{y} = 0.$$

where for any configuration  $\varphi \in \bar{\Omega}$  and  $\varphi_0 \in A_0$

$$\begin{aligned} \mathbf{L}(\varphi; \boldsymbol{\theta}^*) &:= \int_{\bar{\Lambda}} \mathbf{V}_{LJ}^{(1)}(x|\varphi; \boldsymbol{\theta}^*) e^{-V^{LJ}(x|\varphi; \boldsymbol{\theta}^*)} dx - \int_{\bar{\Lambda}} \mathbf{V}_{LJ}^{(1)}(x|\varphi_0; \boldsymbol{\theta}^*) e^{-V^{LJ}(x|\varphi_0; \boldsymbol{\theta}^*)} dx \\ \mathbf{R}(\varphi; \boldsymbol{\theta}^*) &:= \sum_{x \in \varphi \cap \bar{\Lambda}} \mathbf{V}_{LJ}^{(1)}(x|\varphi \setminus x; \boldsymbol{\theta}^*) - \sum_{x \in \varphi_0 \cap \bar{\Lambda}} \mathbf{V}_{LJ}^{(1)}(x|\varphi_0 \setminus x; \boldsymbol{\theta}^*). \end{aligned}$$

Concerning this assumption, we choose  $\bar{D} > D$  and  $\varphi_0 \in A_0 = \left\{ \varphi \in \bar{\Omega} : \varphi_{\Delta_0(\bar{D})} = \emptyset \right\}$ . Let  $\mathbf{y} \in \mathbb{R}^3$  then

$$\int_{\bar{\Lambda}} \mathbf{y}^T \mathbf{V}_{LJ}^{(1)}(x|\varphi_0; \boldsymbol{\theta}^*) e^{-V^{LJ}(x|\varphi_0; \boldsymbol{\theta}^*)} dx = y_1 e^{-\theta_1^*} |\bar{\Lambda}| \quad \text{and} \quad \sum_{x \in \varphi_0 \cap \bar{\Lambda}} \mathbf{y}^T \mathbf{V}_{LJ}^{(1)}(x|\varphi_0 \setminus x; \boldsymbol{\theta}^*) = 0.$$

Consider the following configuration set defined for  $\eta, \varepsilon > 0$

$$A_2(\eta, \varepsilon) = \left\{ \varphi \in \bar{\Omega} : \varphi_{\Delta_0(\bar{D})} = \{z_1, z_2\} \text{ where } z_1 \in \mathcal{B}(0, \eta), z_2 \in \mathcal{B}((0, 2\eta + \varepsilon), \eta) \right\}.$$

Note that for  $z_1 \in \mathcal{B}(0, \eta), z_2 \in \mathcal{B}((0, 2\eta + \varepsilon), \eta), \varepsilon \leq \|z_2 - z_1\| \leq \varepsilon + 4\eta$ . Let  $\varphi_2 \in A_2(\eta, \varepsilon)$  and  $x \in \bar{\Lambda}$ , then one may prove that for  $j = 2, 3$

$$\begin{aligned} V^{LJ}(x|\varphi_2; \boldsymbol{\theta}^*) &= \theta_1^* + 2g^{LJ}(\|x\|; \boldsymbol{\theta}^*) + f(x, \eta, \varepsilon) \\ \frac{\partial V^{LJ}}{\partial \theta_j}(x|\varphi_2; \boldsymbol{\theta}^*) &= 2 \frac{\partial g^{LJ}}{\partial \theta_j}(\|x\|; \boldsymbol{\theta}^*) + f_j(x, \eta, \varepsilon) \end{aligned}$$

where  $f(x, \eta, \varepsilon)$  and  $f_j(x, \eta, \varepsilon)$  are such that

$$\lim_{(\eta, \varepsilon) \rightarrow (0, 0)} f(x, \eta, \varepsilon) = \lim_{(\eta, \varepsilon) \rightarrow (0, 0)} f_j(x, \eta, \varepsilon) = 0.$$

On the one hand, one may prove that there exists a function  $f_L(\mathbf{y}, \eta, \varepsilon)$  such that  $\lim_{(\eta, \varepsilon) \rightarrow (0, 0)} f_L(x, \eta, \varepsilon) = 0$  and such that

$$\mathbf{y}^T \mathbf{L}(\varphi_2; \boldsymbol{\theta}^*) = \mathbf{y}^T \mathbf{I} - y_1 e^{-\theta_1^*} |\bar{\Lambda}| + f_L(\mathbf{y}, \eta, \varepsilon)$$

where

$$\mathbf{I} := \int_{\bar{\Lambda}} \mathbf{h}(\|x\|; \boldsymbol{\theta}^*) e^{-\theta_1^* - 2g^{LJ}(\|x\|; \boldsymbol{\theta}^*)} dx \quad \text{and} \quad \mathbf{h}(r; \boldsymbol{\theta}^*) := \left( 1, 2 \frac{\partial g^{LJ}}{\partial \theta_2}(r; \boldsymbol{\theta}^*), 2 \frac{\partial g^{LJ}}{\partial \theta_3}(r; \boldsymbol{\theta}^*) \right)^T.$$

On the other hand, there exists a function  $f_R(\mathbf{y}, \eta, \varepsilon)$  such that  $\lim_{\eta \rightarrow 0} f_R(x, \eta, \varepsilon) = 0$

$$\mathbf{y}^T \mathbf{R}(\varphi_2; \boldsymbol{\theta}^*) = 2y_1 + 2y_2 4 \left( \left( \frac{\theta_3^*}{\varepsilon} \right)^{12} - \left( \frac{\theta_3^*}{\varepsilon} \right)^6 \right) + 2y_3 4 \theta_2^* \left( \frac{12\theta_3^{*11}}{\varepsilon^{12}} - \frac{6\theta_3^{*5}}{\varepsilon^6} \right) + f_R(\mathbf{y}, \eta, \varepsilon).$$

Since

$$\begin{aligned} \varepsilon^{12} \mathbf{y}^T (\mathbf{L}(\varphi_2; \boldsymbol{\theta}^*) - \mathbf{R}(\varphi_2; \boldsymbol{\theta}^*)) &= \varepsilon^{12} \left( \mathbf{y}^T \mathbf{I} - y_1 e^{-\theta_1^*} |\bar{\Lambda}| + f_L(\mathbf{y}, \eta, \varepsilon) - f_R(\mathbf{y}, \eta, \varepsilon) \right) \\ &\quad - \varepsilon^6 \left( 2y_2 4 \theta_3^{*6} + 2y_3 4 \theta_2^* 6 \theta_3^{*5} \right) + 2y_2 4 \theta_3^{*12} + 2y_3 4 \theta_2^* 12 \theta_3^{*11}. \end{aligned}$$



For  $\eta$  and  $\varepsilon$  chosen small enough, one may prove that

$$0 = |\varepsilon^{12} \mathbf{y}^T (\mathbf{L}(\varphi_2; \boldsymbol{\theta}^*) - \mathbf{R}(\varphi_2; \boldsymbol{\theta}^*))| \geq \frac{1}{2} \left| 2y_2 4\theta_3^{*12} + 2y_3 4\theta_2^* 12\theta_3^{*11} \right|$$

leading to

$$2y_2 4\theta_3^{*12} + 2y_3 4\theta_2^* 12\theta_3^{*11} = 0 \Leftrightarrow \theta_3^* y_2 + 12\theta_2^* y_3 = 0. \quad (11)$$

This means that

$$\mathbf{y}^T \mathbf{R}(\varphi_2; \boldsymbol{\theta}^*) = 2y_1 - \frac{1}{\varepsilon^6} \left( 2y_2 4\theta_3^{*6} + 2y_3 4\theta_2^* 6\theta_3^{*5} \right) + f_R(\mathbf{y}, \eta, \varepsilon).$$

With similar arguments, we obtain that

$$2y_2 4\theta_3^{*6} + 2y_3 4\theta_2^* 6\theta_3^{*5} = 0 \Leftrightarrow \theta_3^* y_2 + 6\theta_2^* y_3 = 0. \quad (12)$$

Equations (11) and (12) lead to  $y_2 = y_3 = 0$ . Now consider the following configuration set defined for some  $k \geq 1$  and  $\eta > 0$

$$A_k(\eta) = \left\{ \varphi \in \overline{\Omega} : \varphi_{\Delta_0(\overline{D})} = \varphi(\mathcal{B}(0, \eta)) = k \right\}$$

and let  $\varphi_k \in A_k(\eta)$ . Then, one may prove that there exists a function  $\tilde{f}_L(\mathbf{y}, \eta)$  such that  $\lim_{\eta \rightarrow 0} \tilde{f}_L(\mathbf{y}, \eta) = 0$  and such that

$$\mathbf{y}^T (\mathbf{L}(\varphi_k; \boldsymbol{\theta}^*) - \mathbf{R}(\varphi_k; \boldsymbol{\theta}^*)) = y_1 \int_{\overline{\Lambda}} e^{-\theta_1^*} \left( e^{-kg^{LJ}(\|x\|; \boldsymbol{\theta}^*)} - 1 \right) dx - ky_1 + \tilde{f}_L(\mathbf{y}, \eta) = 0.$$

Let us denote by  $\Lambda_1 := \mathcal{B}(0, \min(\theta_3^*, D))$  and  $\Lambda_2 := \mathcal{B}(0, D) \setminus \Lambda_1$ . Now let us consider two cases. **Case 1:**  $\theta_3^* \leq D$ . First note that for all  $x \in \overline{\Lambda}$ ,  $g(\|x\|; \boldsymbol{\theta}^*) \geq 0$ . Then, for  $k$  large enough and for  $\eta$  small enough, we have

$$\left| \frac{1}{k} \int_{\Lambda_1} e^{-\theta_1^*} \left( e^{-kg(\|x\|; \boldsymbol{\theta}^*)} - 1 \right) dx \right| \leq \frac{|\Lambda_1|}{k} e^{-\theta_1^*} \leq \frac{1}{4} \quad \text{and} \quad \left| \frac{1}{k} \tilde{f}_L(\mathbf{y}, \eta) \right| \leq \frac{|y_1|}{4}.$$

Hence for  $k$  large enough and for  $\eta$  small enough, we may obtain

$$\begin{aligned} 0 &= \frac{1}{k} \left| \mathbf{y}^T (\mathbf{L}(\varphi_k; \boldsymbol{\theta}^*) - \mathbf{R}(\varphi_k; \boldsymbol{\theta}^*)) \right| \\ &\geq |y_1| - \left| y_1 \frac{1}{k} \int_{\Lambda_1} e^{-\theta_1^*} \left( e^{-kg(\|x\|; \boldsymbol{\theta}^*)} - 1 \right) dx + \frac{1}{k} \tilde{f}_L(\mathbf{y}, \eta) \right| \\ &\geq |y_1| - \frac{|y_1|}{4} - \frac{|y_1|}{4} = \frac{|y_1|}{2}, \end{aligned}$$

which leads to  $y_1 = 0$ .

**Case 2:**  $\theta_3^* \geq D$ . First note that for all  $x \in \Lambda_2$ ,

$$g(\|x\|; \boldsymbol{\theta}^*) \leq g_m := g(D; \boldsymbol{\theta}^*) = 4\theta_2^* \left( \left( \frac{\theta_3^*}{D} \right)^{12} - \left( \frac{\theta_3^*}{D} \right)^6 \right) < 0.$$

On the one hand, for  $k$  large enough and for  $\eta$  small enough, we may have

$$\left| \frac{1}{k} y_1 \int_{\Lambda_1} e^{-\theta_1^*} \left( e^{-kg(\|x\|; \boldsymbol{\theta}^*)} - 1 \right) dx + \frac{1}{k} \tilde{f}_L(\mathbf{y}, \eta) - y_1 \right| \leq \frac{|y_1|}{2} + |y_1| \leq \frac{3}{2} |y_1|.$$

On the other hand, we have for  $k$  large enough

$$\begin{aligned} \frac{1}{k} \left| y_1 \int_{\Lambda_2} e^{-\theta_1^*} \left( e^{-kg(\|x\|; \boldsymbol{\theta}^*)} - 1 \right) dx \right| &= \frac{|y_1|}{k} \int_{\Lambda_2} e^{-\theta_1^*} \left( e^{-kg(\|x\|; \boldsymbol{\theta}^*)} - 1 \right) dx \\ &\geq \frac{|y_1|}{k} e^{-\theta_1^*} |\Lambda_2| (e^{-kg_m} - 1) = |y_1| e^{-\theta_1^*} \frac{e^{k|g_m|} - 1}{k} \\ &\geq 2|y_1|. \end{aligned}$$

Therefore for  $k$  large enough and for  $\eta$  small enough, we have

$$0 = \frac{1}{k} \left| \mathbf{y}^T (\mathbf{L}(\varphi_k; \boldsymbol{\theta}^*) - \mathbf{R}(\varphi_k; \boldsymbol{\theta}^*)) \right| \geq 2|y_1| - \frac{3}{2}|y_1| = \frac{|y_1|}{2},$$

which leads to  $y_1 = 0$ .

## 5 Annex: proofs

### 5.1 Tools

Let us start by presenting a particular case of the Campbell Theorem combined with the Glötz Theorem that is widely used in our future proofs.

**Corollary 8** *Assume that the (marked) point process  $\Phi$  with probability measure  $P$  is stationary. Let  $\Lambda$  be a bounded Borel set, let  $\varphi \in \Omega$  and let  $g$  be a function satisfying  $g(x^m, \varphi_x) = g(0^m, \varphi)$  for all  $x^m \in \mathbb{S}$ . Define  $M$  a random variable with its distribution  $\lambda^m$  and  $f(m, \varphi) = g(0^m, \varphi)e^{-V(0^m|\varphi)}$  and assume that  $f \in L^1(\lambda^m \otimes P)$ . Then,*

$$\mathbf{E} \left( \sum_{x^m \in \Phi_\Lambda} g(x^m, \Phi \setminus x^m) \right) = |\Lambda| \mathbf{E} \left( g(0^M, \Phi) e^{-V(0^M|\Phi)} \right) \quad (13)$$

**Proof.** see Corollary 3 of Billiot et al. (2008) ■

Let us now present a version of an ergodic theorem obtained by Nguyen and Zessin (1979) and widely used in this paper. Let  $\Delta_0$  a fixed bounded domain

**Theorem 9 (Nguyen and Zessin (1979))** *Let  $\{H_G, G \in \mathcal{B}_b\}$  be a family of random variables, which is covariant, that for all  $x \in \mathbb{R}^2$ ,*

$$H_{G_x}(\varphi_x) = H_G(\varphi), \quad \text{for a.e. } \varphi$$

*and additive, that is for every disjoint  $G_1, G_2 \in \mathcal{B}_b$ ,*

$$H_{G_1 \cup G_2} = H_{G_1} + H_{G_2}, \quad \text{a.s.}$$

*Let  $\mathcal{I}$  be the sub- $\sigma$ -algebra of  $F$  consisting of translation invariant (with probability 1) sets. Assume there exists a nonnegative and integrable random variable  $Y$  such that  $|H_G| \leq Y$  a.s. for every convex  $G \subset \Delta_0$ . Then,*

$$\lim_{n \rightarrow +\infty} \frac{1}{|G_n|} H_{G_n} = \frac{1}{|\Delta_0|} E(H_{\Delta_0} | \mathcal{I}), \quad \text{a.s.}$$

*for each regular sequence  $G_n \rightarrow \mathbb{R}^2$ .*

## 5.2 Proof of Theorem 2

Due to the decomposition of stationary measures as a mixture of ergodic measures (see Preston (1976)), one only needs to prove Theorem 2 by assuming that  $P_{\theta^*}$  is ergodic. From now on,  $P_{\theta^*}$  is assumed to be ergodic. The tool used to obtain the almost sure convergence is a convergence theorem for minimum contrast estimators established by Guyon (1992).

We proceed in three stages.

*Step 1. Convergence of  $U_n(\Phi; \theta)$ .*

Decompose  $U_n(\varphi; \theta) = \frac{1}{|\Lambda_n|} (H_{1,\Lambda_n}(\varphi) + H_{2,\Lambda_n}(\varphi))$  with

$$H_{1,\Lambda_n}(\varphi) = \int_{\Lambda_n \times \mathbb{M}} e^{-V(x^m|\varphi; \theta)} \mu(dx^m) \quad \text{and} \quad H_{2,\Lambda_n}(\varphi) = \sum_{x^m \in \Phi_{\Lambda_n}} V(x^m|\varphi \setminus x^m; \theta).$$

Under the assumption [C1], one can apply Theorem 9 (Nguyen and Zessin (1979)) to the process  $H_{1,\Lambda_n}$ . And from Corollary 8, we obtain  $P_{\theta^*}$ -almost surely as  $n \rightarrow +\infty$

$$\frac{1}{|\Lambda_n|} H_{1,\Lambda_n} \rightarrow \mathbf{E} \left( \int_{\mathbb{M}} e^{-V(0^m|\Phi; \theta)} \lambda^{\mathbb{m}}(dm) \right) = \mathbf{E} \left( e^{-V(0^M|\Phi; \theta)} \right). \quad (14)$$

Now, let  $G \subset \Delta_0$ , we clearly have

$$|H_{2,G}(\varphi)| \leq \sum_{x^m \in \varphi_G} |V(x^m|\varphi \setminus x^m; \theta)| \leq \sum_{x^m \in \varphi_{\Delta_0}} |V(x^m|\varphi \setminus x^m; \theta)|.$$

Under the assumption [Mod] and from Corollary 8, we have

$$\mathbf{E} \left( \sum_{x^m \in \Phi_{\Delta_0}} |V(x^m|\Phi \setminus x^m; \theta)| \right) = |\Delta_0| \mathbf{E} \left( |V(0^M|\Phi; \theta)| e^{-V(0^M|\Phi; \theta^*)} \right) < +\infty$$

This means that for all  $G \subset \Delta_0$ , there exists a random variable  $Y \in L^1(P_{\theta^*})$  such that  $|H_{2,G}| \leq Y$ . Thus, under the assumption [C1] and from Theorem 9 (Nguyen and Zessin (1979)) and from Corollary 8, we have  $P_{\theta^*}$ -almost surely

$$\frac{1}{|\Lambda_n|} H_{2,\Lambda_n} \rightarrow \frac{1}{|\Delta_0|} \mathbf{E} \left( \sum_{x^m \in \Phi_{\Delta_0}} V(x^m|\Phi \setminus x^m; \theta) \right) = \mathbf{E} \left( V(0^M|\Phi; \theta) e^{-V(0^M|\Phi; \theta^*)} \right). \quad (15)$$

We have the result by combining (14) and (15):  $P_{\theta^*}$ -almost surely

$$U_n(\cdot; \theta) \rightarrow U(\theta) = \mathbf{E} \left( e^{-V(0^M|\Phi; \theta)} + V(0^M|\Phi; \theta) e^{-V(0^M|\Phi; \theta^*)} \right) \quad (16)$$

*Step 2.  $U_n(\cdot; \theta)$  a contrast function*

Recall that  $U_n(\cdot; \theta)$  is a contrast function if there exists a function  $K(\cdot, \theta^*)$  (i.e. nonnegative function equal to zero if and only if  $\theta = \theta^*$ ) such that  $P_{\theta^*}$ -almost surely  $U_n(\varphi; \theta) - U_n(\varphi; \theta^*) \rightarrow K(\theta, \theta^*)$ . From Step 1, we have

$$K(\theta, \theta^*) = \mathbf{E} \left( e^{-V(0^M|\Phi; \theta^*)} \left( e^{V(0^M|\Phi; \theta) - V(0^M|\Phi; \theta^*)} - \left( 1 + V(0^M|\Phi; \theta) - V(0^M|\Phi; \theta^*) \right) \right) \right). \quad (17)$$

Since the function  $t \mapsto e^t - (1+t)$  is nonnegative and is equal to zero if and only if  $t = 0$ ,  $K(\theta, \theta^*) \geq 0$  and

$$\begin{aligned} K(\theta, \theta^*) = 0 &\Leftrightarrow e^{V(0^m|\varphi; \theta) - V(0^m|\varphi; \theta^*)} - \left( 1 + V(0^m|\varphi; \theta) - V(0^m|\varphi; \theta^*) \right) = 0 \\ &\Leftrightarrow D(0^m|\varphi; \theta) := V(0^m|\varphi; \theta) - V(0^m|\varphi; \theta^*) = 0 \end{aligned}$$

for  $\lambda^m \times P_{\boldsymbol{\theta}^*}$  - a.e.  $(m, \varphi)$ . Let us consider the  $\ell$  events  $B_j$  ( $j = 1, \dots, \ell$ ) defined in Assumption [C2]. The previous Equation is at least true for  $(m_j, \varphi_j) \in B_j$ , which leads under Assumption [C2] to  $\boldsymbol{\theta} = \boldsymbol{\theta}^*$ . Therefore,  $K(\boldsymbol{\theta}, \boldsymbol{\theta}^*) = 0 \Rightarrow \boldsymbol{\theta} = \boldsymbol{\theta}^*$ . The converse is trivial.

Before ending this step, note that the assumption [C3] asserts that for any  $\varphi$ ,  $U_n(\varphi; \cdot)$  and  $K(\cdot, \boldsymbol{\theta}^*)$  are continuous functions.

*Step 3. Modulus of continuity.*

The modulus of continuity of the contrast process defined for all  $\varphi \in \Omega$  and all  $\eta > 0$  by

$$W_n(\varphi, \eta) = \sup \left\{ \left| U_n(\varphi; \boldsymbol{\theta}) - U_n(\varphi; \boldsymbol{\theta}') \right| : \boldsymbol{\theta}, \boldsymbol{\theta}' \in \Theta, \|\boldsymbol{\theta} - \boldsymbol{\theta}'\| \leq \eta \right\}$$

is such that there exists a sequence  $(\varepsilon_k)_{k \geq 1}$ , with  $\varepsilon_k \rightarrow 0$  as  $k \rightarrow +\infty$  such that for all  $k \geq 1$

$$P \left( \limsup_{n \rightarrow +\infty} \left( W_n \left( \Phi, \frac{1}{k} \right) \geq \varepsilon_k \right) \right) = 0. \quad (18)$$

Let us start to write  $W_n(\varphi, \frac{1}{k}) \leq W_{1,n}(\varphi, \frac{1}{k}) + W_{2,n}(\varphi, \frac{1}{k})$  with

$$\begin{aligned} W_{1,n} \left( \varphi, \frac{1}{k} \right) &:= \sup \left\{ W'_{1,\Lambda_n}(\varphi; \boldsymbol{\theta}, \boldsymbol{\theta}') : \boldsymbol{\theta}, \boldsymbol{\theta}' \in \Theta, \|\boldsymbol{\theta} - \boldsymbol{\theta}'\| \leq \frac{1}{k} \right\} \\ W_{2,n} \left( \varphi, \frac{1}{k} \right) &:= \sup \left\{ W'_{2,\Lambda_n}(\varphi; \boldsymbol{\theta}, \boldsymbol{\theta}') : \boldsymbol{\theta}, \boldsymbol{\theta}' \in \Theta, \|\boldsymbol{\theta} - \boldsymbol{\theta}'\| \leq \frac{1}{k} \right\}. \end{aligned}$$

and

$$\begin{aligned} W'_{1,\Lambda_n}(\varphi; \boldsymbol{\theta}, \boldsymbol{\theta}') &:= \frac{1}{|\Lambda_n|} \int_{\Lambda_n \times \mathbb{M}} \left| e^{-V(x^m|\varphi; \boldsymbol{\theta})} - e^{-V(x^m|\varphi; \boldsymbol{\theta}')} \right| \mu(dx^m) \\ W'_{2,\Lambda_n}(\varphi; \boldsymbol{\theta}, \boldsymbol{\theta}') &:= \frac{1}{|\Lambda_n|} \sum_{x^m \in \varphi_{\Lambda_n}} \left| V(x^m|\varphi \setminus x^m; \boldsymbol{\theta}) - V(x^m|\varphi \setminus x^m; \boldsymbol{\theta}') \right|. \end{aligned}$$

Let  $k \geq 1$  and let  $\boldsymbol{\theta}, \boldsymbol{\theta}' \in \Theta$  such that  $\|\boldsymbol{\theta} - \boldsymbol{\theta}'\| \leq \frac{1}{k}$ , then under the assumption [C1] and from Theorem 9 and Corollary 8, we have  $P_{\boldsymbol{\theta}^*}$ -almost surely as  $n \rightarrow +\infty$

$$\begin{aligned} W'_{1,\Lambda_n}(\Phi; \boldsymbol{\theta}, \boldsymbol{\theta}') &\longrightarrow \mathbf{E} \left( \left| e^{-V(0^M|\Phi; \boldsymbol{\theta})} - e^{-V(0^M|\Phi; \boldsymbol{\theta}')} \right| \right) \\ W'_{2,\Lambda_n}(\Phi; \boldsymbol{\theta}, \boldsymbol{\theta}') &\longrightarrow \mathbf{E} \left( \left| V(0^M|\Phi; \boldsymbol{\theta}) - V(0^M|\Phi; \boldsymbol{\theta}') \right| e^{-V(0^M|\Phi; \boldsymbol{\theta}^*)} \right) \end{aligned}$$

Under Assumption [C4], one may apply the mean value theorem in  $\mathbb{R}^p$  as follows: there exist  $\boldsymbol{\xi}^{(1)}, \dots, \boldsymbol{\xi}^{(p)} \in \prod_{j=1}^p [\min(\theta_j, \theta'_j), \max(\theta_j, \theta'_j)]$  such that for all  $(m, \varphi) \in \mathbb{M} \times \Omega_f$

$$e^{-V(0^m|\varphi; \boldsymbol{\theta})} - e^{-V(0^m|\varphi; \boldsymbol{\theta}')} = \sum_{j=1}^p (\theta_j - \theta'_j) \frac{\partial V}{\partial \theta_j} \left( 0^m|\varphi; \boldsymbol{\xi}^{(j)} \right) e^{-V(0^m|\varphi; \boldsymbol{\xi}^{(j)})}.$$

This leads, under Assumption [C4], to the following inequality

$$\begin{aligned} \mathbf{E} \left( \left| e^{-V(0^M|\Phi; \boldsymbol{\theta})} - e^{-V(0^M|\Phi; \boldsymbol{\theta}')} \right| \right)^2 &\leq \mathbf{E} \left( \left| e^{-V(0^M|\Phi; \boldsymbol{\theta})} - e^{-V(0^M|\Phi; \boldsymbol{\theta}')} \right|^2 \right) \\ &\leq \mathbf{E} \left( \|\boldsymbol{\theta} - \boldsymbol{\theta}'\|^2 \sum_{j=1}^p \left| \frac{\partial V}{\partial \theta_j} \left( 0^M|\Phi; \boldsymbol{\xi}^{(j)} \right) e^{-V(0^M|\Phi; \boldsymbol{\xi}^{(j)})} \right|^2 \right) \\ &\leq \left( \frac{1}{k} \right)^2 \gamma_1^2, \end{aligned}$$

with  $\gamma_1 := \mathbf{E} \left( \sum_{j=1}^p \max_{\boldsymbol{\theta} \in \Theta} \left| \frac{\partial V}{\partial \theta_j} (0^M | \Phi; \boldsymbol{\theta}) e^{-V(0^M | \Phi; \boldsymbol{\theta})} \right|^2 \right) < +\infty$ . In such a way, one may also prove that

$$\mathbf{E} \left( \left| V(0^M | \Phi; \boldsymbol{\theta}) - V(0^M | \Phi; \boldsymbol{\theta}') \right| e^{-V(0^M | \Phi; \boldsymbol{\theta}^*)} \right)^2 \leq \left( \frac{1}{k} \right)^2 \gamma_2,$$

with  $\gamma_2 := \mathbf{E} \left( \sum_{j=1}^p \max_{\boldsymbol{\theta} \in \Theta} \left| \frac{\partial V}{\partial \theta_j} (0^M | \Phi; \boldsymbol{\theta}) e^{-V(0^M | \Phi; \boldsymbol{\theta}^*)} \right|^2 \right)$ . Hence, for all  $k \geq 1$  and for all  $\boldsymbol{\theta}, \boldsymbol{\theta}' \in \Theta$  such that  $\|\boldsymbol{\theta} - \boldsymbol{\theta}'\| \leq \frac{1}{k}$  there exists  $n_0(k) \geq 1$  such that for all  $n \geq n_0(k)$ , we have

$$W'_{1, \Lambda_n}(\varphi; \boldsymbol{\theta}, \boldsymbol{\theta}') \leq \frac{2}{k} \gamma_1 \quad \text{and} \quad W'_{2, \Lambda_n}(\varphi; \boldsymbol{\theta}, \boldsymbol{\theta}') \leq \frac{2}{k} \gamma_2, \text{ for } P_{\boldsymbol{\theta}^*} - \text{a.e. } \varphi.$$

Since  $\gamma_1$  and  $\gamma_2$  are independent of  $\boldsymbol{\theta}$  and  $\boldsymbol{\theta}'$ , we have for all  $n \geq n_0(k)$

$$W_n \left( \varphi, \frac{1}{k} \right) \leq W_{1, n} \left( \varphi, \frac{1}{k} \right) + W_{2, n} \left( \varphi, \frac{1}{k} \right) \leq \frac{2}{k} (\gamma_1 + \gamma_2) := \frac{c}{k}, \text{ for } P_{\boldsymbol{\theta}^*} - \text{a.e. } \varphi.$$

Finally, since

$$\limsup_{n \rightarrow +\infty} \left\{ W_n \left( \varphi, \frac{1}{k} \right) \geq \frac{c}{k} \right\} = \bigcap_{m \in \mathbb{N}} \bigcup_{n \geq m} \left\{ W_n \left( \varphi, \frac{1}{k} \right) \geq \frac{c}{k} \right\} \subset \bigcup_{n \geq n_0(k)} \left\{ W_n \left( \varphi, \frac{1}{k} \right) \geq \frac{c}{k} \right\}$$

for  $P_{\boldsymbol{\theta}^*}$ -a.e.  $\varphi$ , the expected result (18) is proved.

*Conclusion step.* The Steps 1, 2 and 3 ensure the fact that we can apply Property 3.6 of Guyon (1992) which asserts the almost sure convergence for minimum contrast estimators.

### 5.3 Proof of Theorem 3

*Step 1. Asymptotic normality of  $U_n^{(1)}(\Phi; \boldsymbol{\theta}^*)$*

The aim is to prove that for any fixed  $\tilde{D}$ , the following convergence in distribution as  $n \rightarrow +\infty$

$$|\Lambda_n|^{1/2} U_n^{(1)}(\Phi; \boldsymbol{\theta}^*) \rightarrow \mathcal{N} \left( 0, \underline{\Sigma}(\tilde{D}, \boldsymbol{\theta}^*) \right) \quad (19)$$

where the matrix  $\underline{\Sigma}(\tilde{D}, \boldsymbol{\theta}^*)$  is defined by (6).

The idea is to apply to  $U_n^{(1)}(\Phi; \boldsymbol{\theta}^*)$  a central limit theorem obtained by Jensen and Künsch (1994), Theorem 2.1. The following conditions have to be fulfilled to apply this result. For all  $j = 1, \dots, p$

(i) For all  $i \in \mathbb{Z}^2$ ,  $\mathbf{E} \left( \left( \mathbf{LPL}_{\Delta_i}^{(1)}(\Phi; \boldsymbol{\theta}^*) \right)_j \middle| \Phi_{\Delta_i^c} \right) = 0$ .

(ii) For all  $i \in \mathbb{Z}^2$ ,  $\mathbf{E} \left( \left| \left( \mathbf{LPL}_{\Delta_i}^{(1)}(\Phi; \boldsymbol{\theta}^*) \right)_j \right|^3 \right) < +\infty$ .

(iii) The matrix  $\text{Var} \left( |\Lambda_n|^{1/2} U_n^{(1)}(\Phi; \boldsymbol{\theta}^*) \right)$  converges to the matrix  $\underline{\Sigma}(\tilde{D}, \boldsymbol{\theta}^*)$ .

Condition (i) : From the stationarity of the process, it is sufficient to prove that

$$\mathbf{E} \left( \left( \mathbf{LPL}_{\Delta_0}^{(1)}(\Phi; \boldsymbol{\theta}^*) \right)_j \middle| \Phi_{\Delta_0^c} \right) = 0.$$

Recall that for any configuration  $\varphi$

$$\left( \mathbf{LPL}_{\Delta_0}^{(1)}(\varphi; \boldsymbol{\theta}^*) \right)_j = - \int_{\Delta_0 \times \mathbb{M}} \frac{\partial V}{\partial \theta_j}(x^m | \varphi; \boldsymbol{\theta}^*) e^{-V(x^m | \varphi; \boldsymbol{\theta}^*)} \mu(dx^m) + \int_{\Delta_0} \frac{\partial V}{\partial \theta_j}(x^m | \varphi \setminus x^m; \boldsymbol{\theta}^*) \varphi(dx^m). \quad (20)$$

Denote respectively by  $G_1(\varphi)$  and  $G_2(\varphi)$  the first and the second right-hand term of (20) and by  $E_i = \mathbf{E}(G_i(\Phi) | \Phi_{\Delta_0^c} = \varphi_{\Delta_0^c})$ . From the definition of Gibbs point processes,

$$E_2 = \frac{1}{Z_{\Delta_0}(\varphi_{\Delta_0^c})} \int_{\Omega_{\Delta_0}} Q(d\varphi_{\Delta_0}) \int_{\mathbb{S}} \varphi_{\Delta_0}(dx^m) \mathbf{1}_{\Delta_0}(x) \frac{\partial V}{\partial \theta_j}(x^m | \varphi \setminus x^m; \boldsymbol{\theta}^*) e^{-V(\varphi_{\Delta_0} | \varphi_{\Delta_0^c}; \boldsymbol{\theta}^*)}.$$

Denote by  $\varphi' = (\varphi_{\Delta_0}, \varphi'_{\Delta_0^c})$ . Since  $Q$  is a Poisson process we can write

$$\begin{aligned} E_2 &= \frac{1}{Z_{\Delta_0}(\varphi_{\Delta_0^c})} \int_{\Omega} Q(d\varphi') \int_{\mathbb{S}} \varphi'(dx^m) \mathbf{1}_{\Delta_0}(x) \frac{\partial V}{\partial \theta_j}(x^m | \varphi \setminus x^m; \boldsymbol{\theta}^*) e^{-V(\varphi_{\Delta_0} | \varphi_{\Delta_0^c}; \boldsymbol{\theta}^*)} \\ &= \frac{1}{Z_{\Delta_0}(\varphi_{\Delta_0^c})} \int_{\Omega} Q(d\varphi') \int_{\mathbb{S}} \varphi'(dx^m) \mathbf{1}_{\Delta_0}(x) \frac{\partial V}{\partial \theta_j}(x^m | \varphi'_{\Delta_0} \cup \varphi_{\Delta_0^c} \setminus x^m; \boldsymbol{\theta}^*) e^{-V(\varphi'_{\Delta_0} | \varphi_{\Delta_0^c}; \boldsymbol{\theta}^*)} \end{aligned}$$

Now, from Campbell Theorem (applied to the Poisson measure  $Q$ )

$$E_2 = \frac{1}{Z_{\Delta_0}(\varphi_{\Delta_0^c})} \int_{\Delta_0 \times \mathbb{M}} \mu(dx^m) \int_{\Omega} Q_{x^m}^1(d\varphi') \frac{\partial V}{\partial \theta_j}(x^m | \varphi'_{\Delta_0} \cup \varphi_{\Delta_0^c}; \boldsymbol{\theta}^*) e^{-V(\varphi'_{\Delta_0} \cup x^m | \varphi_{\Delta_0^c}; \boldsymbol{\theta}^*)}.$$

Since from Slivnyak-Mecke Theorem (see *e.g.* Møller and Waagepetersen (2003)),  $Q = Q_x^1$ , one can obtain

$$\begin{aligned} E_2 &= \frac{1}{Z_{\Delta_0}(\varphi_{\Delta_0^c})} \int_{\Omega} Q(d\varphi') \int_{\Delta_0 \times \mathbb{M}} \mu(dx^m) \frac{\partial V}{\partial \theta_j}(x^m | \varphi'_{\Delta_0} \cup \varphi_{\Delta_0^c}; \boldsymbol{\theta}^*) e^{-V(\varphi'_{\Delta_0} \cup x^m | \varphi_{\Delta_0^c}; \boldsymbol{\theta}^*)} \\ &= \frac{1}{Z_{\Delta_0}(\varphi_{\Delta_0^c})} \int_{\Omega} Q(d\varphi_{\Delta_0}) \int_{\Delta_0 \times \mathbb{M}} \mu(dx^m) \frac{\partial V}{\partial \theta_j}(x^m | \varphi; \boldsymbol{\theta}^*) e^{-V(x^m | \varphi; \boldsymbol{\theta}^*)} e^{-V(\varphi_{\Delta_0} | \varphi_{\Delta_0^c}; \boldsymbol{\theta}^*)} \\ &= -E_1 \end{aligned}$$

Condition (ii) : For any bounded domain  $\Delta$  and any finite configuration  $\varphi$ , one may write for  $j = 1, \dots, p$

$$\left| \left( \mathbf{LPL}_{\Delta}^{(1)}(\varphi; \boldsymbol{\theta}^*) \right)_j \right|^3 \leq 4 \left| \int_{\Delta \times \mathbb{M}} \frac{\partial V}{\partial \theta_j}(x^m | \varphi; \boldsymbol{\theta}^*) e^{-V(x^m | \varphi; \boldsymbol{\theta}^*)} \mu(dx^m) \right|^3 + 4 \left| \sum_{x^m \in \varphi_{\Delta}} \frac{\partial V}{\partial \theta_j}(x^m | \varphi \setminus x^m; \boldsymbol{\theta}^*) \right|^3.$$

The assumption **[N1(3)]** ensures the integrability of the first right-hand term. For the second one, note that

$$\begin{aligned} T_2 &:= \left| \sum_{x^m \in \varphi_{\Delta}} \frac{\partial V}{\partial \theta_j}(x^m | \varphi \setminus x^m; \boldsymbol{\theta}^*) \right|^3 \\ &\leq \sum_{\substack{x_1^{m_1}, x_2^{m_2}, x_3^{m_3} \in \varphi_{\Delta} \\ x_1^{m_1} \neq x_1^{m_2}, x_2^{m_2} \neq x_3^{m_3}, x_2^{m_2} \neq x_3^{m_3}}} \left| \frac{\partial V}{\partial \theta_j}(x_1^{m_1} | \varphi \setminus x_1^{m_1}; \boldsymbol{\theta}^*) \right| \left| \frac{\partial V}{\partial \theta_j}(x_2^{m_2} | \varphi \setminus x_2^{m_2}; \boldsymbol{\theta}^*) \right| \left| \frac{\partial V}{\partial \theta_j}(x_3^{m_3} | \varphi \setminus x_3^{m_3}; \boldsymbol{\theta}^*) \right| \\ &\quad + 3 \sum_{x_1^{m_1}, x_2^{m_2} \in \varphi_{\Delta}, x_1^{m_1} \neq x_2^{m_2}} \left| \frac{\partial V}{\partial \theta_j}(x_1^{m_1} | \varphi \setminus x_1^{m_1}; \boldsymbol{\theta}^*) \right|^2 \left| \frac{\partial V}{\partial \theta_j}(x_2^{m_2} | \varphi \setminus x_2^{m_2}; \boldsymbol{\theta}^*) \right| \\ &\quad + \sum_{x_1^{m_1} \in \varphi_{\Delta}} \left| \frac{\partial V}{\partial \theta_j}(x_2^{m_2} | \varphi \setminus x_1^{m_1}; \boldsymbol{\theta}^*) \right|^3. \end{aligned}$$

The result is obtained by using the assumption **[N1(3)]** and iterated versions of Corollary 8.

Condition (iii): let us start by noting that from the assumption **[Mod-L]**, the vector  $\mathbf{LPL}_{\Delta_i}^{(1)}(\varphi; \boldsymbol{\theta}^*)$  depends only on  $\varphi_{\Delta_j}$  for  $j \in \mathbb{B}\left(i, \left[\frac{D}{2}\right]\right)$ . Let  $E_{i,j} := \mathbf{E}\left(\mathbf{LPL}_{\Delta_i}^{(1)}(\Phi; \boldsymbol{\theta}^*) \mathbf{LPL}_{\Delta_j}^{(1)}(\Phi; \boldsymbol{\theta}^*)^T\right)$ . Based on our definitions, we have

$$\begin{aligned} \mathbb{V}\text{ar}\left(|\Lambda_n|^{1/2} \mathbf{U}_n^{(1)}(\Phi; \boldsymbol{\theta}^*)\right) &= |\Lambda_n|^{-1} \mathbb{V}\text{ar}\left(\sum_{i \in I_n} \mathbf{LPL}_{\Delta_i}^{(1)}(\Phi; \boldsymbol{\theta}^*)\right) \\ &= |\Lambda_n|^{-1} \sum_{i,j \in I_n} E_{i,j} \\ &= |\Lambda_n|^{-1} \sum_{i \in I_n} \left( \sum_{j \in I_n \cap \mathbb{B}\left(i, \left[\frac{D}{2}\right]\right)} E_{i,j} + \sum_{j \in I_n \cap \mathbb{B}\left(i, \left[\frac{D}{2}\right]\right)^c} E_{i,j} \right). \end{aligned}$$

Let  $j \in I_n \cap \mathbb{B}\left(i, \left[\frac{D}{2}\right]\right)^c$ , since  $\mathbf{LPL}_{\Delta_i}^{(1)}(\varphi; \boldsymbol{\theta}^*)$  is a measurable function of  $\varphi_{\Delta_i^c}$  we have by using condition (i):

$$\begin{aligned} \mathbf{E}\left(\mathbf{LPL}_{\Delta_i}^{(1)}(\Phi; \boldsymbol{\theta}^*) \mathbf{LPL}_{\Delta_j}^{(1)}(\Phi; \boldsymbol{\theta}^*)^T\right) &= \mathbf{E}\left(\mathbf{E}\left(\mathbf{LPL}_{\Delta_i}^{(1)}(\Phi; \boldsymbol{\theta}^*) \mathbf{LPL}_{\Delta_j}^{(1)}(\Phi; \boldsymbol{\theta}^*)^T \mid \Phi_{\Delta_i^c}\right)\right) \\ &= \mathbf{E}\left(\mathbf{E}\left(\mathbf{LPL}_{\Delta_i}^{(1)}(\Phi; \boldsymbol{\theta}^*) \mid \Phi_{\Delta_i^c}\right) \mathbf{LPL}_{\Delta_j}^{(1)}(\Phi; \boldsymbol{\theta}^*)^T\right) \\ &= 0 \end{aligned}$$

Denote by  $\tilde{I}_n$  the following set

$$\tilde{I}_n = I_n \cap \left( \cup_{i \in \partial I_n} \mathbb{B}\left(i, \left[\frac{D}{2}\right]\right) \right).$$

We now obtain

$$\begin{aligned} \mathbb{V}\text{ar}\left(|\Lambda_n|^{1/2} \mathbf{U}_n^{(1)}(\Phi; \boldsymbol{\theta}^*)\right) &= |\Lambda_n|^{-1} \sum_{i \in I_n} \sum_{j \in I_n \cap \mathbb{B}\left(i, \left[\frac{D}{2}\right]\right)} E_{i,j} \\ &= |\Lambda_n|^{-1} \left( \sum_{i \in I_n \setminus \tilde{I}_n} \sum_{j \in I_n \cap \mathbb{B}\left(i, \left[\frac{D}{2}\right]\right)} E_{i,j} + \sum_{i \in \tilde{I}_n} \sum_{j \in I_n \cap \mathbb{B}\left(i, \left[\frac{D}{2}\right]\right)} E_{i,j} \right) \end{aligned}$$

Using the stationarity and the definition of the domain  $\Lambda_n$ , one obtains

$$|\Lambda_n|^{-1} \sum_{i \in I_n \setminus \tilde{I}_n} \sum_{j \in I_n \cap \mathbb{B}\left(i, \left[\frac{D}{2}\right]\right)} E_{i,j} = |\Lambda_n|^{-1} |I_n \setminus \tilde{I}_n| \sum_{j \in \mathbb{B}\left(0, \left[\frac{D}{2}\right]\right)} E_{0,j} \rightarrow \boldsymbol{\Sigma}(\tilde{D}, \boldsymbol{\theta}^*) \quad \text{as } n \rightarrow +\infty$$

and

$$|\Lambda_n|^{-1} \left| \sum_{i \in \tilde{I}_n} \sum_{j \in I_n \cap \mathbb{B}\left(i, \left[\frac{D}{2}\right]\right)} E_{i,j} \right| \leq |\Lambda_n|^{-1} |\tilde{I}_n| \sum_{j \in \mathbb{B}\left(0, \left[\frac{D}{2}\right]\right)} |E_{0,j}| \rightarrow 0 \quad \text{as } n \rightarrow +\infty.$$

Hence as  $n \rightarrow +\infty$

$$\begin{aligned} \text{Var} \left( |\Lambda_n|^{1/2} \mathbf{U}_n^{(1)}(\Phi; \boldsymbol{\theta}^*) \right) &= |\Lambda_n|^{-1} \sum_{i \in I_n} \sum_{j \in I_n \cap \mathbb{B}(i, \lceil \frac{D}{B} \rceil)} E_{i,j} \\ &\xrightarrow{n \rightarrow +\infty} \underbrace{|\Lambda_n|^{-1}}_{\bar{D}-2} \sum_{k \in \mathbb{B}(0, \lceil \frac{D}{B} \rceil)} E_{0,k} = \underline{\Sigma}(\bar{D}, \boldsymbol{\theta}^*). \end{aligned} \quad (21)$$

*Step 2. Domination of  $\underline{\mathbf{U}}_n^{(2)}(\Phi; \boldsymbol{\theta})$  in a neighborhood of  $\boldsymbol{\theta}^*$  and convergence of  $\underline{\mathbf{U}}_n^{(2)}(\Phi; \boldsymbol{\theta}^*)$*  Let  $j, k = 1, \dots, p$ , recall that  $\left( \underline{\mathbf{U}}_n^{(2)}(\varphi; \boldsymbol{\theta}) \right)_{j,k}$  is defined in a neighborhood  $\mathcal{V}(\boldsymbol{\theta}^*)$  of  $\boldsymbol{\theta}^*$  for any configuration  $\varphi$  by

$$\begin{aligned} \left( \underline{\mathbf{U}}_n^{(2)}(\varphi; \boldsymbol{\theta}) \right)_{j,k} &= -\frac{1}{|\Lambda_n|} \int_{\Lambda_n \times \mathbb{M}} \frac{\partial^2 V}{\partial \theta_j \partial \theta_k} (x^m | \varphi; \boldsymbol{\theta}) \exp(-V(x^m | \varphi; \boldsymbol{\theta})) \mu(dx^m) \\ &\quad + \frac{1}{|\Lambda_n|} \int_{\Lambda_n \times \mathbb{M}} \frac{\partial V}{\partial \theta_j} (x^m | \varphi; \boldsymbol{\theta}) \frac{\partial V}{\partial \theta_k} (x^m | \varphi; \boldsymbol{\theta}) \exp(-V(x^m | \varphi; \boldsymbol{\theta})) \mu(dx^m) \\ &\quad + \frac{1}{|\Lambda_n|} \sum_{x^m \in \varphi_{\Lambda_n}} \frac{\partial^2 V}{\partial \theta_j \partial \theta_k} (x^m | \varphi \setminus x^m; \boldsymbol{\theta}). \end{aligned} \quad (22)$$

Under the assumption **[N1(3)]** and **[N2]**, from Theorem 9 (Nguyen and Zessin (1979)) and from Corollary 8, there exists  $n_0 \in \mathbb{N}$  such that for all  $n \geq n_0$

$$\begin{aligned} \left| \left( \underline{\mathbf{U}}_n^{(2)}(\varphi; \boldsymbol{\theta}) \right)_{j,k} \right| &\leq 2\mathbf{E} \left( \left( \left| \frac{\partial^2 V}{\partial \theta_j \partial \theta_k} (0^M | \Phi; \boldsymbol{\theta}) \right| + \left| \frac{\partial V}{\partial \theta_j} (0^M | \Phi; \boldsymbol{\theta}) \frac{\partial V}{\partial \theta_k} (0^M | \Phi; \boldsymbol{\theta}) \right| \right) e^{-V(0^M | \Phi; \boldsymbol{\theta})} \right) \\ &\quad + 2 \times \mathbf{E} \left( \left| \frac{\partial^2 V}{\partial \theta_j \partial \theta_k} (0^M | \Phi; \boldsymbol{\theta}) \right| e^{-V(0^M | \Phi; \boldsymbol{\theta}^*)} \right) \end{aligned}$$

Note that from Theorem 9 (Nguyen and Zessin (1979)),  $\underline{\mathbf{U}}_n^{(2)}(\cdot; \boldsymbol{\theta}^*)$  converges almost surely as  $n \rightarrow +\infty$  towards  $\underline{\mathbf{U}}^{(2)}(\boldsymbol{\theta}^*)$  defined by (5). Note that  $\underline{\mathbf{U}}^{(2)}(\boldsymbol{\theta}^*)$  is a symmetric positive matrix since for all  $\mathbf{y} \in \mathbb{R}^p$

$$\mathbf{y}^T \underline{\mathbf{U}}^{(2)}(\boldsymbol{\theta}^*) \mathbf{y} = \mathbf{E} \left( \left( \mathbf{y}^T \mathbf{V}^{(1)}(0^M | \Phi; \boldsymbol{\theta}^*) \right)^2 e^{-V(0^M | \Phi; \boldsymbol{\theta}^*)} \right) \geq 0,$$

where for  $j = 1, \dots, p$ ,  $(m, \varphi) \in \mathbb{M} \times \Omega_f$  and for  $\boldsymbol{\theta} \in \mathcal{V}(\boldsymbol{\theta}^*)$   $\left( \mathbf{V}^{(1)}(x^m | \varphi; \boldsymbol{\theta}^*) \right)_j := \frac{\partial V}{\partial \theta_j} (x^m | \varphi; \boldsymbol{\theta})$  and it is a definite matrix under the assumption **[N3]**.

*Conclusion Step* Under the assumptions **[Mod]** and **[Ident]**, and using Steps 1 and 2, one can apply a classical result concerning asymptotic normality for minimum contrast estimators *e.g.* Proposition 3.7 de Guyon (1992) in order to obtain (7).

It remains to prove (8). The proof is strictly similar to the one of Billiot et al. (2008) (p. 261) except that the assumption **[SDP]** is now replaced by the more general one assumption **[N4]**. We keep it for a better understanding. The final result is proved in three substeps :

(i) We first prove that the matrix  $\underline{\Sigma}(\boldsymbol{\theta}^*) = \underline{\Sigma}(\bar{D}, \boldsymbol{\theta}^*)$  is a symmetric definite positive matrix. From Equation (21), it is sufficient to prove that the matrix  $\text{Var}(|\Lambda_n(\bar{D})|^{-1/2} \mathbf{LPL}_{\Lambda_n(\bar{D})}^{(1)}(\Phi; \boldsymbol{\theta}^*))$  is definite positive for  $n$  large enough. Let  $\mathbf{y} \in \mathbb{R}^p \setminus \{0\}$ , the aim is to prove that

$$\mathbf{V} := \mathbf{y}^T \text{Var} \left( |\Lambda_n(\bar{D})|^{-1/2} \mathbf{LPL}_{\Lambda_n(\bar{D})}^{(1)}(\Phi; \boldsymbol{\theta}^*) \right) \mathbf{y} > 0.$$



Let  $\bar{\Lambda} = \cup_{|i| \leq \mathbb{B}(i, \lceil \frac{D}{2} \rceil)} \Delta_i(\bar{D})$ , using the same argument of Jensen and Künsch (1994) (Equation (3.2)), one can write

$$V \geq |\Lambda_n(\bar{D})|^{-1} \mathbf{E} \left( \text{Var} \left( \mathbf{y}^T \mathbf{LPL}_{\Lambda_n(\bar{D})}^{(1)}(\Phi; \boldsymbol{\theta}^*) | \Phi_{\Delta_k(\bar{D})}, k \notin (2 \lceil \frac{D}{2} \rceil + 1)\mathbb{Z}^2 \right) \right).$$

Note that for  $i \neq j \in I_n$ ,  $\text{Cov}(\mathbf{y}^T \mathbf{LPL}_{\Delta_i(\bar{D})}^{(1)}(\varphi; \boldsymbol{\theta}^*), \mathbf{y}^T \mathbf{LPL}_{\Delta_j(\bar{D})}^{(1)}(\varphi; \boldsymbol{\theta}^*) | \Phi_{\Delta_k(\bar{D})}, k \notin (2 \lceil \frac{D}{2} \rceil + 1)\mathbb{Z}^2) = 0$  due to the independence of  $\mathbf{LPL}_{\Delta_i(\bar{D})}^{(1)}(\varphi; \boldsymbol{\theta}^*)$  and  $\mathbf{LPL}_{\Delta_j(\bar{D})}^{(1)}(\varphi; \boldsymbol{\theta}^*)$  conditionally on  $\Phi_{\Delta_k(\bar{D})}, k \notin (2 \lceil \frac{D}{2} \rceil + 1)\mathbb{Z}^2$  when  $i, j \in I_n \cap (2 \lceil \frac{D}{2} \rceil + 1)\mathbb{Z}^2$  and  $\mathbf{LPL}_{\Delta_i(\bar{D})}^{(1)}(\varphi; \boldsymbol{\theta}^*)$  or  $\mathbf{LPL}_{\Delta_j(\bar{D})}^{(1)}(\varphi; \boldsymbol{\theta}^*)$  is constant when either  $i$  or  $j \notin I_n \cap (2 \lceil \frac{D}{2} \rceil + 1)\mathbb{Z}^2$ . As a direct consequence,

$$\begin{aligned} V &\geq |\Lambda_n(\bar{D})|^{-1} \mathbf{E} \left( \text{Var} \left( \mathbf{y}^T \sum_{i \in I_n} \mathbf{LPL}_{\Delta_i(\bar{D})}^{(1)}(\Phi; \boldsymbol{\theta}^*) | \Phi_{\Delta_k(\bar{D})}, k \notin (2 \lceil \frac{D}{2} \rceil + 1)\mathbb{Z}^2 \right) \right) \\ &= |\Lambda_n(\bar{D})|^{-1} \sum_{i \in I_n} \mathbf{E} \left( \text{Var} \left( \mathbf{y}^T \mathbf{LPL}_{\Delta_i(\bar{D})}^{(1)}(\Phi; \boldsymbol{\theta}^*) | \Phi_{\Delta_k(\bar{D})}, k \notin (2 \lceil \frac{D}{2} \rceil + 1)\mathbb{Z}^2 \right) \right) \\ &= |\Lambda_n(\bar{D})|^{-1} \sum_{\ell \in I_n \cap (2 \lceil \frac{D}{2} \rceil + 1)\mathbb{Z}^2 \setminus \bar{I}_n} \mathbf{E} \left( \text{Var} \left( \mathbf{y}^T \sum_{i \in I_n \cap \mathbb{B}(\ell, \lceil \frac{D}{2} \rceil)} \mathbf{LPL}_{\Delta_i(\bar{D})}^{(1)}(\Phi; \boldsymbol{\theta}^*) | \Phi_{\Delta_k(\bar{D})}, k \notin (2 \lceil \frac{D}{2} \rceil + 1)\mathbb{Z}^2 \right) \right) \\ &\quad + |\Lambda_n(\bar{D})|^{-1} \sum_{\ell \in (2 \lceil \frac{D}{2} \rceil + 1)\mathbb{Z}^2 \cap \bar{I}_n} \mathbf{E} \left( \text{Var} \left( \mathbf{y}^T \sum_{i \in I_n \cap \mathbb{B}(\ell, \lceil \frac{D}{2} \rceil)} \mathbf{LPL}_{\Delta_i(\bar{D})}^{(1)}(\Phi; \boldsymbol{\theta}^*) | \Phi_{\Delta_k(\bar{D})}, k \notin (2 \lceil \frac{D}{2} \rceil + 1)\mathbb{Z}^2 \right) \right) \end{aligned}$$

Following the proof of Step 1, condition (iii) one may prove that the second right-hand term tends to 0 as  $n \rightarrow +\infty$ . Therefore by using the stationarity, we have for  $n$  large enough

$$\begin{aligned} V &\geq \frac{1}{2} |\Lambda_n(\bar{D})|^{-1} \left| I_n \cap (2 \lceil \frac{D}{2} \rceil + 1)\mathbb{Z}^2 \right| \times \mathbf{E} \left( \text{Var} \left( \mathbf{y}^T \mathbf{LPL}_{\bar{\Lambda}}^{(1)}(\Phi; \boldsymbol{\theta}^*) | \Phi_{\Delta_k(\bar{D})}, 1 \leq |k| \leq 2 \lceil \frac{D}{2} \rceil \right) \right) \\ &= \frac{\bar{D}^{-2} |I_n \cap (2 \lceil \frac{D}{2} \rceil + 1)\mathbb{Z}^2|}{2 |I_n|} \times \mathbf{E} \left( \text{Var} \left( \mathbf{y}^T \mathbf{LPL}_{\bar{\Lambda}}^{(1)}(\Phi; \boldsymbol{\theta}^*) | \Phi_{\Delta_k(\bar{D})}, 1 \leq |k| \leq 2 \lceil \frac{D}{2} \rceil \right) \right) \\ &\geq \frac{\bar{D}^{-2}}{2} \left( \frac{3}{4 \lceil \frac{D}{2} \rceil + 1} \right)^2 \times \mathbf{E} \left( \text{Var} \left( \mathbf{y}^T \mathbf{LPL}_{\bar{\Lambda}}^{(1)}(\Phi; \boldsymbol{\theta}^*) | \Phi_{\Delta_k(\bar{D})}, 1 \leq |k| \leq 2 \lceil \frac{D}{2} \rceil \right) \right) \end{aligned}$$

Assume there exists some positive constant  $c$  such that  $P_{\boldsymbol{\theta}^*}$ -a.s.  $\mathbf{y}^T \mathbf{LPL}_{\bar{\Lambda}}^{(1)}(\Phi; \boldsymbol{\theta}^*) = c$  when the variables  $\Phi_{\Delta_k(\bar{D})}, 1 \leq |k| \leq 2 \lceil \frac{D}{2} \rceil$  are (for example) fixed to  $\emptyset$ . By assuming [SDP] it follows that for any  $\varphi_i \in A_i$  for  $i = 0, \dots, \ell$  (with  $\ell \geq p$ ),  $\mathbf{y}^T \left( \mathbf{LPL}_{\bar{\Lambda}}^{(1)}(\varphi_i; \boldsymbol{\theta}^*) - \mathbf{LPL}_{\bar{\Lambda}}^{(1)}(\varphi_0; \boldsymbol{\theta}^*) \right) = 0$ . Since for all  $(\varphi_0, \dots, \varphi_\ell) \in A_0 \times \dots \times A_\ell$ , the matrix with entries  $\left( \mathbf{LPL}_{\bar{\Lambda}}^{(1)}(\varphi_i; \boldsymbol{\theta}^*) \right)_j - \left( \mathbf{LPL}_{\bar{\Lambda}}^{(1)}(\varphi_0; \boldsymbol{\theta}^*) \right)_j$  is assumed to be injective, this leads to  $\mathbf{y} = 0$  and hence to some contradiction. Therefore, when the variables  $\Phi_{\Delta_k(\bar{D})}, 1 \leq |k| \leq 2 \lceil \frac{D}{2} \rceil$  are fixed to  $\emptyset$ , the variable  $\mathbf{y}^T \mathbf{LPL}_{\bar{\Lambda}}^{(1)}(\Phi; \boldsymbol{\theta}^*)$  is almost surely not a constant. Hence,  $\Sigma(\boldsymbol{\theta}^*)$  is a symmetric definite positive matrix.

(ii) Convergence of  $\widehat{\Sigma}_n(\varphi; D^\vee, \bar{D}, \boldsymbol{\theta})$ .

Let us recall that for any  $\varphi \in \Omega$ ,  $D^\vee \geq D$  and  $\theta \in \Theta$  we define

$$\widehat{\underline{\Sigma}}_n(\varphi; D^\vee, \tilde{D}, \theta) = \frac{\tilde{D}^{-2}}{|I_n|} \sum_{i \in I_n} \sum_{j \in I_n \cap \mathbb{B}(i, \lceil \frac{D^\vee}{D} \rceil)} \mathbf{LPL}_{\Delta_i}^{(1)}(\varphi; \theta)^T \mathbf{LPL}_{\Delta_j}^{(1)}(\varphi; \theta)$$

We also define

$$X_i(\varphi) := X_i(\varphi)^{k,\ell} = \sum_{j \in I_n \cap \mathbb{B}(i, \lceil \frac{D^\vee}{D} \rceil)} \left( \mathbf{LPL}_{\Delta_i}^{(1)}(\varphi; \theta) \right)_k \left( \mathbf{LPL}_{\Delta_j}^{(1)}(\varphi; \theta) \right)_\ell,$$

$Y_i(\varphi) := X_i(\varphi) - \mathbf{E}(X_i(\Phi))$  and  $\bar{Y}_n(\varphi) = |I_n|^{-1} \sum_{i \in I_n} Y_i(\varphi)$ . Since one may notice that  $\mathbf{E}(X_i(\Phi)) = \tilde{D}^2 \left( \underline{\Sigma}(\tilde{D}, \theta) \right)_{k,\ell}$ , we have

$$\bar{Y}_n(\varphi) = \tilde{D}^2 \left( \widehat{\underline{\Sigma}}(\varphi; D^\vee, \tilde{D}, \theta) - \underline{\Sigma}(\tilde{D}, \theta) \right)_{k,\ell}.$$

Thus, the aim is to prove  $\bar{Y}_n(\varphi) \rightarrow 0$  for  $P_{\theta^*}$ -a.e.  $\varphi$ . Since the process  $\{Y_i, i \in \mathbb{Z}^2\}$  is stationary, from Property 3.1 p.96 of Guyon (1992), it is sufficient to prove

- (a)  $\mathbf{E}(Y_0(\Phi)^2) < +\infty$
- (b)  $\mathbf{E}(|I_n| \bar{Y}_n(\Phi)^2) < +\infty$ .

(a) We leave the reader to verify that **[N1(4)]** ensures this integrability condition.

(b) Note that  $Y_i(\varphi)$  depends only on  $\varphi_{\Delta_j}$  for  $j \in \mathbb{B}\left(i, \lceil \frac{D^\vee}{D} \rceil + \lceil \frac{D}{D} \rceil\right)$ . Hence, by choosing  $j \in I_n \cap \mathbb{B}(i, 2\alpha_{D^\vee, D})^c$  with  $\alpha_{D^\vee, D} = \alpha_{D^\vee, D}(\tilde{D}) := \lceil \frac{D^\vee}{D} \rceil + \lceil \frac{D}{D} \rceil$ , then the covariance between  $Y_i(\Phi)$  and  $Y_j(\Phi)$  is zero. Indeed, let  $A = \cup_{i \in \mathbb{B}(i, \alpha_{D^\vee, D})} \Delta_i$

$$\begin{aligned} \mathbf{E}(Y_i(\Phi)Y_j(\Phi)) &= \mathbf{E}(\mathbf{E}(Y_i(\Phi)Y_j(\Phi)|\Phi_A)) \\ &= \mathbf{E}(Y_i(\Phi)\mathbf{E}(Y_j(\Phi)|\Phi_A)) \\ &= 0. \end{aligned}$$

Then, we obtain

$$\begin{aligned} \mathbf{E}(|I_n| \bar{Y}_n(\Phi)^2) &= \frac{1}{|I_n|} \sum_{i,j \in I_n} \mathbf{E}(Y_i(\Phi)Y_j(\Phi)) \\ &= \frac{1}{|I_n|} \sum_{i \in I_n} \left( \sum_{j \in I_n \cap \mathbb{B}(i, 2\alpha_{D^\vee, D})} \mathbf{E}(Y_i(\Phi)Y_j(\Phi)) + \sum_{j \in I_n \cap \mathbb{B}(i, 2\alpha_{D^\vee, D})^c} \mathbf{E}(Y_i(\Phi)Y_j(\Phi)) \right) \\ &= \frac{1}{|I_n|} \sum_{i \in I_n} \sum_{j \in I_n \cap \mathbb{B}(i, 2\alpha_{D^\vee, D})} \mathbf{E}(Y_i(\Phi)Y_j(\Phi)) \\ &\stackrel{n \rightarrow +\infty}{\rightarrow} \sum_{k \in \mathbb{B}(0, 2\alpha_{D^\vee, D})} \mathbf{E}(Y_0(\Phi)Y_k(\Phi)) \leq c \mathbf{E}(Y_0(\Phi)^2), \end{aligned}$$

where  $c$  is a constant depending only on  $D, \tilde{D}, D^\vee$ . Therefore, for all  $D^\vee \geq D$  and for all  $\theta \in \Theta$ , we have for  $P_{\theta^*}$ -a.e.  $\varphi$  as  $n \rightarrow +\infty$

$$\widehat{\underline{\Sigma}}_n(\varphi; D^\vee, \tilde{D}, \theta) \rightarrow \underline{\Sigma}(\tilde{D}, \theta) = \underline{\Sigma}(\theta). \quad (23)$$

(iii) Since for any  $\varphi$ , the functions  $\underline{U}_n^{(2)}(\varphi; \cdot)$  and  $\widehat{\underline{\Sigma}}_n(\varphi; D^\vee, \tilde{D}, \cdot)$  are continuous, it follows from Step 2 and (23) that one obtains for  $P_{\theta^*}$ -a.e.  $\varphi$ , as  $n \rightarrow +\infty$

$$\underline{U}_n^{(2)}(\varphi; \widehat{\theta}) \rightarrow \underline{U}^{(2)}(\theta^*) \quad \text{and} \quad \widehat{\underline{\Sigma}}_n(\varphi; D^\vee, \tilde{D}, \widehat{\theta}) \rightarrow \underline{\Sigma}(\theta^*).$$

Finally, note that the previous convergence also implies that for  $n$  large enough  $\widehat{\underline{\Sigma}}_n(\Phi; D^\vee, \tilde{D}, \widehat{\theta})$  is almost surely a symmetric definite positive matrix.

## References

- J. Besag. Spatial interaction and the statistical analysis of lattice system. *J. R. Statist. Soc. Ser. B*, 26:192–236, 1974.
- J. Besag, R. Milne, and S. Zachary. Point process limits of lattice processes. *Ann. Appl. Probab.*, 19:210–216, 1982.
- 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.
- X. Guyon. *Champs aléatoires sur un réseau*. Masson, Paris, 1992.
- J.L. Jensen and H.R. Künsch. On asymptotic normality of pseudo likelihood estimates of pairwise interaction processes. *Ann. Inst. Statist. Math.*, 46:475–486, 1994.
- J.L. Jensen and J. Møller. Pseudolikelihood for exponential family models of spatial point processes. *Ann. Appl. Probab.*, 1:445–461, 1991.
- O. Kallenberg. *Random Measures*. Akademie-Verlag and Academic Press, Berlin, London, 1983.
- J.E. Lennard-Jones. Cohesion. *Proceedings of the Physical Society*, 43:461–482, 1931.
- S. Mase. Marked gibbs processes and asymptotic normality of maximum pseudo-likelihood estimators. *Math. Nachr.*, 209:151–169, 2000.
- S. Mase. Consistency of maximum pseudo-likelihood estimator of continuous state space gibbsian process. *Ann. Appl. Probab.*, 5:603–612, 1995.
- J. Møller and R. Waagepetersen. *Statistical Inference and Simulation for Spatial Point Processes*. Chapman and Hall/CRC, Boca Raton, 2003.
- X.X. Nguyen and H. Zessin. Ergodic theorems for Spatial Process. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 48:133–158, 1979.
- C.J. Preston. *Random Fields*, volume 534. Springer-Verlag, Berlin, Heidelberg, New York, 1976.
- D. Ruelle. Superstable interactions in classical statistical mechanics. *Commun. Math. Phys.*, 18: 127–159, 1970.
- D. Stoyan, W.S. Kendall, and J. Mecke. *Stochastic Geometry and its Applications*. Wiley, Chichester, second edition, 1995.