HAL open science

Asymptotic properties of the maximum pseudo-likelihood estimator for stationary Gibbs point processes including the Lennard-Jones model

Jean-François Coeurjolly, Rémy Drouilhet

▶ To cite this version:

Jean-François Coeurjolly, Rémy Drouilhet. Asymptotic properties of the maximum pseudo-likelihood estimator for stationary Gibbs point processes including the Lennard-Jones model. Electronic Journal of Statistics , 2010, 4, pp.677-706. 10.1214/09-EJS494 . hal-00442750v2

HAL Id: hal-00442750

https://hal.science/hal-00442750v2

Submitted on 12 Jul 2010

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.

Asymptotic properties of the maximum pseudo-likelihood estimator for stationary Gibbs point processes including the Lennard-Jones model

Jean-François Coeurjolly[1,2] and Rémy Drouilhet[1]
[1] LJK, Grenoble University, France
[2] GIPSA-lab, Grenoble University, France

July 12, 2010

Abstract

This paper presents asymptotic properties of the maximum pseudo-likelihood estimator of a vector $\boldsymbol{\theta}$ parameterizing a stationary Gibbs point process. Sufficient conditions, expressed in terms of the local energy function defining a Gibbs point process, to establish strong consistency and asymptotic normality results of this estimator depending on a single realization, are presented. These results are general enough to no longer require the local stability and the linearity in terms of the parameters of the local energy function. We consider characteristic examples of such models, the Lennard-Jones and the finite range Lennard-Jones models. We show that the different assumptions ensuring the consistency are satisfied for both models whereas the assumptions ensuring the asymptotic normality are fulfilled only for the finite range Lennard-Jones model.

Keywords: Stationary Gibbs point processes, maximum pseudo-likelihood estimator, Lennard-Jones model.

1 Introduction

These last years, much attention has been paid to spatial point pattern data, and especially to models and methodologies for fitting them, see Møller (2008) for a recent overview of this topic and Daley and Vere-Jones (1988), Stoyan et al. (1987) Møller and Waagepetersen (2003) or Illian et al. (2008) for more general information. For spatial point pattern data, the reference model is the Poisson point process modelling a random configuration of points with no interaction between points. In particular, this leads to the independence of any two random sub-configurations lying in two non-overlapping domains. A way to introduce dependence is to consider the class of Gibbs models. In a bounded domain, a Gibbs point process is defined through its probability measure having a Radon-Nykodym derivative with respect to a Poisson point process measure proportional to $e^{-V(\varphi)}$ where $V(\varphi)$ corresponds to the energy function (i.e. a cost function expressed in terms of interactions) of the configuration of points φ. The definition of Gibbs models in $\mathbb{R}^d$ is essential when dealing with asymptotic properties of estimators

based on a point process observed in a domain aimed at converging towards $\mathbb{R}^d$. The extension of this definition is not so straightforward. The probability measure of a Gibbs point process in $\mathbb{R}^d$ has to be defined by specifying its conditional density (indirectly expressed in terms of the energy function $V(\varphi)$), see *e.g.* Preston (1976) or Section 2 for more details.

The class of Gibbs point processes is extremely rich. The energy function can penalize points, pairs or triplets of points (see *e.g.* Baddeley and Turner (2000)). More sophisticated models can also be obtained by considering interactions based on the Delaunay or the $k-$nearest neighbor graphs (Bertin et al. (1999b,c)), Voronoï tessellations (Dereudre and Lavancier (2009)) or random sets (Kendall et al. (1999), Dereudre (2009)).

Following the definition of a parametric Gibbs point process, the natural question of efficiently estimating the parameters arises. Many proposals have tried to estimate the energy function from an available point pattern data. The most well-known method 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. This parametric normalizing constant is difficult to calculate from a practical point of view. From a theoretical one, it also makes asymptotic results more complicated to obtain. An alternative approach relies on the use of the pseudo-likelihood function. The idea originated from Besag (1974) in the study of lattice processes. Besag et al. (1982) further considered this method for pairwise interaction point processes, and Jensen and Møller (1991) extended the definition of the pseudo-likelihood function to the general class of marked Gibbs point processes. The construction of the pseudo-likelihood function is based on the conditional densities which spare the computation of the scaling factor.

Our paper deals with asymptotic properties of the maximum pseudo-likelihood estimator. In order to underline our theoretical improvements, let us discuss the two main different papers discussing this topic:

- In Billiot et al. (2008), we obtain consistency and asymptotic normality for exponential family models of Gibbs point processes, that is, on models with energy functions that are linear in terms of the parameters. Moreover, we concentrate on models such that the local energy function is local and stable. The locality of the local energy expresses that the energy to insert a point x into φ, that is, $V(x|\varphi) = V(\varphi \cup x) - V(\varphi)$, depends only on the points of φ falling into some ball with a fixed radius whereas the stability of the local energy (property referred as the local stability) asserts that $V(x|\varphi)$ is bounded from below by a finite negative constant. The paper Billiot et al. (2008) 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 non necessarily stable local energy functions (actually for superstable and lower regular ones introduced by Ruelle (1970)) is obtained in Mase (1995) for specific models with only 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.

Based on this literature, the main goal of this paper is to derive asymptotic properties similar to the ones presented before (consistency and asymptotic normality) but in a more general framework. We provide asymptotic results for general Gibbs point processes with *non (necessarily) linear* and *non (necessarily) stable* local energy functions. The characteristic example we have in

mind is the *Lennard-Jones model*. This model, from statistical physics, is a stationary pairwise interaction Gibbs point process where the local energy to insert a point x into a configuration φ is parameterized as follows: for $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3) \in \mathbb{R}^3$ with $\theta_2, \theta_3 > 0$

$$V^{LJ}(x|\varphi; \boldsymbol{\theta}) := \theta_1 + 4\theta_2 \sum_{y \in \varphi} \left(\left(\frac{\theta_3}{\|y - x\|} \right)^{12} - \left(\frac{\theta_3}{\|y - x\|} \right)^{6} \right).$$

Let us notice that Mase (1995) could only propose the estimation of θ_1 and θ_2 with known θ_3. The Lennard-Jones model is of great interest from several points of view. From a physical point of view, this model arises when theoretically 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 a Pauli repulsion from the Pauli exclusion principle). In this literature, the parameters θ_2 and θ_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, this model constitutes the main example of superstable, regular and lower regular energies studied in Ruelle (1970) where the author proves the existence of ergodic measures for such models. Finally, from a statistical point of view, this model has been considered by several authors, see *e.g.* Ogata and Tanemura (1981), Goulard et al. (1996) for fitting spatial point patterns arising in forestry. In particular, let us note that, in Goulard et al. (1996), the model is fitted by using the maximum pseudo-likelihood method. As the authors do not endeavour to justify the theoretical performances of the procedure, the result proposed in Section 4 of this paper fills this gap.

The rest of the paper is organized as follows. Section 2 introduces some background and notation on Gibbs point processes (general definitions, examples). The maximum pseudo-likelihood method and asymptotic results of the derived estimator are proposed in Section 3. For general Gibbs point processes, sufficient conditions, expressed in terms of the local energy function to establish strong consistency and asymptotic normality results of this estimator are presented. While no general condition on the model is assumed to obtain the consistency, the characteristic finite range of the local energy function is required to establish the asymptotic normality. For the sake of simplicity, Section 3 (and the resulting proofs) would concentrate on non-marked Gibbs point processes. However, as we have shown in our seminal paper Billiot et al. (2008), no real mathematical difficulty occurs with the introduction of marks. Section 4 focuses on the Lennard-Jones model. We show that the general assumptions described in Section 3 are fulfilled for this model. Proofs have been postponed until Section 5.

2 Background and notation

For the sake of simplicity, we consider Gibbs point processes in dimension $d = 2$.

2.1 General notation, configuration space

Subregions of $\mathbb{R}^2$ will typically be denoted by Λ or Δ and will always be assumed to be Borel with positive Lebesgue measure. We write $\Lambda \Subset \mathbb{R}^2$ if Λ is bounded. Λ^c denotes the complementary set of Λ inside $\mathbb{R}^2$. The notation $|.|$ will be used without ambiguity for different kind of objects. For a countable set $\mathcal{J}$, $|\mathcal{J}|$ represents the number of elements belonging to $\mathcal{J}$; For $\Lambda \Subset \mathbb{R}^2$, $|\Lambda|$ is the volume of Λ; For a vector $x \in \mathbb{R}^2$, $|x|$ corresponds to its uniform norm while $\|x\|$ is simply its euclidean norm. For all $x \in \mathbb{R}^2, \rho > 0$ and $i \in \mathbb{Z}^2$, let $\mathcal{B}(x, \rho) := \{y \in \mathbb{R}^2,\ |y - x| < \rho\}$ and $\mathbb{B}(i, \rho) := \mathcal{B}(i, \rho) \cap \mathbb{Z}^2$.

A configuration is a subset φ of $\mathbb{R}^2$ which is locally finite in that $\varphi_\Lambda := \varphi \cap \Lambda$ has finite cardinality $N_\Lambda(\varphi) := |\varphi_\Lambda|$ for all $\Lambda \Subset \mathbb{R}^2$. The space Ω of all configurations is equipped with the σ-algebra $\mathcal{F}$ that is generated by the counting variables $N_\Lambda(\varphi)$ with $\Lambda \Subset \mathbb{R}^2$. Finally, let $T = (\tau_x)_{x\in\mathbb{R}^2}$ be the shift group, where $\tau_x : \Omega \to \Omega$ is the translation by the vector $-x \in \mathbb{R}^2$.

2.2 Gibbs point processes

Our results will be expressed for general stationary Gibbs point processes. Since we are interested in asymptotic properties, we have to consider these point processes acting on the infinite volume $\mathbb{R}^2$. Let us briefly recall their definition.

A point process Φ is a Ω-valued random variable, with probability distribution P on $(\Omega, \mathcal{F})$. The most prominent point process is the (homogeneous) Poisson process with intensity $z > 0$. Recall that its probability measure π^z is the unique probability measure on $(\Omega, \mathcal{F})$ such that the following holds for $\Lambda \Subset \mathbb{R}^2$: (i) N_Λ is Poisson distributed with parameter $z|\Lambda|$, and (ii) conditionally to $N_\Lambda = n$, the n points in Λ are independent with uniform distribution on Λ, for each interger $n \geq 1$. For $\Lambda \Subset \mathbb{R}^2$, let us denote by π^z_Λ the marginal probability measure in Λ of the Poisson process with intensity z.

Let $\boldsymbol{\theta} \in \mathbb{R}^p$ (for some $p \geq 1$). For any $\Lambda \Subset \mathbb{R}^2$, let us consider the parametric function $V_\Lambda(.;\boldsymbol{\theta})$ from Ω into $\mathbb{R} \cup \{+\infty\}$. From a physical point of view, $V_\Lambda(\varphi;\boldsymbol{\theta})$ is the energy of φ_Λ in Λ given the outside configuration φ_{Λ^c}.

In this article, we focus on stationary point processes on $\mathbb{R}^2$, i.e. with T-invariant probability measure. For any $\Lambda \Subset \mathbb{R}^2$, we therefore consider $V_\Lambda(.;\boldsymbol{\theta}))$ to be T-invariant, i.e. $V_\Lambda(\tau_x\varphi;\boldsymbol{\theta}) = V_\Lambda(\varphi;\boldsymbol{\theta})$ for any $x \in \mathbb{R}^2$. Furthermore, we assume that the family of energies is hereditary, which means that for any $\Lambda \Subset \mathbb{R}^2$, $\varphi \in \Omega$, and $x \in \Lambda$: $V_\Lambda(\varphi;\boldsymbol{\theta})) = +\infty \Rightarrow V_\Lambda(\varphi \cup \{x\};\boldsymbol{\theta})) = +\infty$.

In such a context, a Gibbs measure is usually defined as follows (see Preston (1976)).

Definition 1 *A probability measure $P_{\boldsymbol{\theta}}$ on Ω is a Gibbs measure for the family of energies $(V_\Lambda(.;\boldsymbol{\theta}))_{\Lambda\Subset\mathbb{R}^2}$ if for every $\Lambda \Subset \mathbb{R}^2$, for $P_{\boldsymbol{\theta}}$-almost every outside configuration φ_{Λ^c}, the law of $P_{\boldsymbol{\theta}}$ given φ_{Λ^c} admits the following density with respect to π^z_Λ:*

$$f_\Lambda(\varphi_\Lambda|\varphi_{\Lambda^c};\boldsymbol{\theta}) = \frac{1}{Z_\Lambda(\varphi_{\Lambda^c};\boldsymbol{\theta})} e^{-V_\Lambda(\varphi;\boldsymbol{\theta})},$$

where $Z_\Lambda(\varphi_{\Lambda^c};\boldsymbol{\theta}) := \int_{\Omega_\Lambda} e^{-V_\Lambda(\varphi_\Lambda\cup\varphi_{\Lambda^c};\boldsymbol{\theta})}\pi^z_\Lambda(d\varphi_\Lambda)$ is called the partition function.

Without loss of generality, the intensity of the Poisson process, z is fixed to 1 and we simply write π and π_Λ in place of π^1 and π^1_Λ. In the previous definition, we implicitly assume the consistency of the family $(f_\Lambda(.|.;\boldsymbol{\theta}))_{\Lambda\Subset\mathbb{R}^2}$: for any $\Delta \subset \Lambda \Subset \mathbb{R}^2$

$$f_\Delta(\varphi_\Delta|\varphi_{\Delta^c};\boldsymbol{\theta}) = \frac{f_\Lambda(\varphi_\Delta \cup \varphi_{\Lambda\setminus\Delta}|\varphi_{\Lambda^c};\boldsymbol{\theta})}{f_\Lambda(\varphi_{\Lambda\setminus\Delta}|\varphi_{\Lambda^c};\boldsymbol{\theta})} = \frac{f_\Lambda(\varphi_\Delta \cup \varphi_{\Lambda\setminus\Delta}|\varphi_{\Lambda^c};\boldsymbol{\theta})}{\int_{\Omega_\Delta} f_\Lambda(\psi_\Delta \cup \varphi_{\Lambda\setminus\Delta}|\varphi_{\Lambda^c};\boldsymbol{\theta})\pi_\Delta(d\psi_\Delta)}.$$

A sufficient condition to directly fulfill this basic ingredient is to assume the compatibility of the family $(V_\Lambda(.))_{\Lambda\Subset\mathbb{R}^2}$: for every $\Delta \subset \Lambda \Subset \mathbb{R}^2$, the function $\varphi \to V_\Lambda(\varphi;\boldsymbol{\theta}) - V_\Delta(\varphi;\boldsymbol{\theta})$ from Ω into $\mathbb{R} \cup \{+\infty\}$ is measurable and only depends on φ_{Λ^c}.

The existence of a Gibbs measure on Ω which satisfies these conditional specifications is a difficult issue. We refer the interested reader to Ruelle (1969); Preston (1976); Bertin et al. (1999a); Dereudre (2005); Dereudre et al. (2010) for the technical and mathematical development of the existence problem. The minimal assumption of our paper is then:

[**Mod-E**]: Our data consist in the realization of a point process Φ with Gibbs measure $P_{\boldsymbol{\theta}^\star}$, where $\boldsymbol{\theta}^\star \in \mathring{\boldsymbol{\Theta}}$, $\boldsymbol{\Theta}$ is a compact subset of $\mathbb{R}^p$ and, for any $\boldsymbol{\theta} \in \boldsymbol{\Theta}$, there exists a stationary Gibbs measure $P_{\boldsymbol{\theta}}$ for the family $(V_\Lambda(.;\boldsymbol{\theta}))_{\Lambda \Subset \mathbb{R}^2}$.

In the rest of this paper, the reader has mainly to keep in mind the concept of local energy defined as the energy required to insert a point x into the configuration φ and expressed for any $\Lambda \ni x$ by

$$V\left(x|\varphi;\boldsymbol{\theta}\right) := V_\Lambda(\varphi \cup \{x\}) - V_\Lambda(\varphi).$$

From the compatibility of the family of energies, the local energy does not depend on Λ.

Our asymptotic normality result will require the following locality property assumption.

[**Mod-L**]: There exists $D \geq 0$ such that for all $\varphi \in \Omega$

$$V\left(0|\varphi;\boldsymbol{\theta}\right) = V\left(0|\varphi_{\mathcal{B}(0,D)};\boldsymbol{\theta}\right).$$

2.3 Example : Lennard-Jones models

Let us present the main example studied in this paper. We call LJ-type model the stationary pairwise interaction point process defined for some $D \in]0,+\infty]$ by

$$V_\Lambda^{LJ}\left(\varphi;\boldsymbol{\theta}\right) := \theta_1|\varphi_\Lambda| + H_\Lambda^{LJ}\left(\varphi;\boldsymbol{\theta}\right) \text{ with } H_\Lambda^{LJ}\left(\varphi;\boldsymbol{\theta}\right) := \sum_{\substack{x_1\in\varphi_\Lambda \\ x_2\in\varphi_{\Lambda^c}}} g^{LJ}(||x_1-x_2||;\boldsymbol{\theta})$$

and

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

As a direct consequence, the local energy function is expressed as

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

where $\boldsymbol{\theta} = (\theta_1,\theta_2,\theta_3) \in \mathbb{R}\times(\mathbb{R}^+)^2$. The cases $D=+\infty$ and $D<+\infty$ respectively correpond to the Lennard-Jones model (briefly presented in the introduction) and the Lennard-Jones model with finite range.

Ruelle (1970) has proved the existence of an ergodic measure for superstable, regular and lower regular potentials. The Lennard-Jones model (including the finite range one) is known to be the characteristic example of such a family of models for which Ruelle managed to prove the existence of ergodic measures for any $\boldsymbol{\theta} \in \mathbb{R}\times(\mathbb{R}^+)^2$. In order to ensure [**Mod-E**], it is required to assume that $\theta_2^\star, \theta_3^\star > 0$. Finally, [**Mod-L**] is satisfied for the LJ-type model with $D<+\infty$ since the parameter D corresponds for pairwise interaction point processes to the range of the Gibbs point process.

3 Asymptotic results of the Maximum pseudo-likelihood estimator

3.1 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 Gibbs point processes. Using our notation and up to a scalar factor, the pseudo-likelihood defined for a configuration φ and a domain of observation Λ is denoted by $PL_\Lambda(\varphi;\boldsymbol{\theta})$ and given by

$$PL_\Lambda(\varphi;\boldsymbol{\theta}) = \exp\left(-\int_\Lambda e^{-V\left(x|\varphi;\boldsymbol{\theta}\right)}dx\right)\prod_{x\in\varphi_\Lambda} e^{-V\left(x|\varphi\setminus x;\boldsymbol{\theta}\right)}. \tag{1}$$

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

$$LPL_\Lambda(\varphi;\boldsymbol{\theta}) = -\int_\Lambda e^{-V\left(x|\varphi;\boldsymbol{\theta}\right)}dx - \sum_{x\in\varphi_\Lambda} V(x|\varphi\setminus x;\boldsymbol{\theta}). \tag{2}$$

The point process is assumed to be observed in a domain $\Lambda_n \oplus \widetilde{D} = \cup_{x\in\Lambda_n}\mathcal{B}(x,\widetilde{D})$ for some $\widetilde{D} < +\infty$. For the asymptotic normality result, it is also assumed that $\widetilde{D} \geq D$ and 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 $i\in\mathbb{Z}^2$, $\Delta_i = \Delta_i(\widetilde{D})$ is the square centered at i with side-length $\widetilde{D}$. As a consequence, as $n\to+\infty$, $\Lambda_n \to \mathbb{R}^2$ such that $|\Lambda_n|\to+\infty$ and $\dfrac{|\partial\Lambda_n|}{|\Lambda_n|}\to 0$.

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

$$\widehat{\boldsymbol{\theta}}_n(\varphi) = \underset{\boldsymbol{\theta}\in\boldsymbol{\Theta}}{\arg\max}\ LPL_{\Lambda_n}(\varphi;\boldsymbol{\theta}) = \underset{\boldsymbol{\theta}\in\boldsymbol{\Theta}}{\arg\min}\, U_n(\varphi;\boldsymbol{\theta}).$$

The following basic notation are introduced: for $j,k=1,\ldots,p$ and $\Lambda \Subset \mathbb{R}^2$

- Gradient vector of U_n: $\boldsymbol{U}_n^{(1)}(\varphi;\boldsymbol{\theta}) := -|\Lambda_n|^{-1}\boldsymbol{LPL}_{\Lambda_n}^{(1)}(\varphi;\boldsymbol{\theta})$ where

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

- Hessian matrix of U_n: $\underline{\boldsymbol{U}}_n^{(2)}(\varphi;\boldsymbol{\theta}) := -|\Lambda_n|^{-1}\underline{\boldsymbol{LPL}}_{\Lambda_n}^{(2)}(\varphi;\boldsymbol{\theta})$

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

Finally, note that from the decomposition of the observation domain Λ_n, one has

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

3.2 Consistency of the MPLE

The assumption **[C]** gathers the following four assumptions:

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

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

[C2] Identifiability condition : there exists $A_1, \ldots, A_\ell$, $\ell \geq p$ events in Ω such that:

- the ℓ events A_i are disjoint and satisfy $P_{\boldsymbol{\theta}^\star}(B_i) > 0$
- for all $(\varphi_1, \ldots, \varphi_\ell) \in A_1 \times \cdots \times A_\ell$

$$\left\{ \begin{array}{l} D(0|\varphi_i; \boldsymbol{\theta}) = 0 \\ i = 1 \ldots, \ell \end{array} \right. \quad \Rightarrow \quad \boldsymbol{\theta} = \boldsymbol{\theta}^\star$$

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

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

[C4] For all $\varphi \in \Omega$, $V(0|\varphi; \boldsymbol{\theta})$ is continuously differentiable in $\boldsymbol{\theta}$ and for all $j = 1, \ldots, p$

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

Theorem 1 *Under the assumptions **[Mod-E]** and **[C]**, for $P_{\boldsymbol{\theta}^\star}-$almost every φ, the maximum pseudo-likelihood estimate $\widehat{\boldsymbol{\theta}}_n(\varphi)$ converges towards $\boldsymbol{\theta}^\star$ as n tends to infinity.*

3.3 Asymptotic normality of the MPLE

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

[N1] For all $\varphi \in \Omega$, $V(0|\varphi; \boldsymbol{\theta})$ is differentiable in $\boldsymbol{\theta} = \boldsymbol{\theta}^\star$. For all $k = 1, \ldots, 3$ and for all $\lambda_1, \ldots, \lambda_k$, k positive integers such that $\sum_{i=1}^k \lambda_i = 3$ and for $\Delta \Subset \mathbb{R}^2$

$$\boldsymbol{E}\left(\int_{\Delta^k} \prod_{i=1}^k \left| \frac{\partial V}{\partial \theta_j}(0^M|\Phi; \boldsymbol{\theta}^\star) \right|^{\lambda_i} e^{-V(\{x_1, \ldots, x_k\}|\Phi; \boldsymbol{\theta}^\star)} dx_1 \ldots dx_k \right) < +\infty.$$

[N2] There exists a neighbourhood $\mathcal{V}(\boldsymbol{\theta}^\star)$ of $\boldsymbol{\theta}^\star$ such that for all $\varphi \in \Omega$, $V(0|\varphi; \boldsymbol{\theta})$ is twice continuously differentiable in $\boldsymbol{\theta} \in \mathcal{V}$ and, for all $j, k = 1, \ldots, p$ and $\boldsymbol{\theta} \in \mathcal{V}(\boldsymbol{\theta}^\star)$,

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

and

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

[N3] There exists $A_1, \ldots, A_\ell$, $\ell \geq p$ events in Ω such that:

- the ℓ events A_i are disjoint and satisfy $P_{\boldsymbol{\theta}^\star}(A_i) > 0$
- for all $(\varphi_1, \ldots, \varphi_\ell) \in A_1 \times \cdots \times A_\ell$ the (ℓ, p) matrix with entries $\frac{\partial V}{\partial \theta_j}(0|\varphi_i; \boldsymbol{\theta}^\star)$ is injective.

[N4] There exists $A_0, \ldots, A_\ell$, $\ell \geq p$ disjoint sub-events of $\overline{\Omega} := \{\varphi \in \Omega : \varphi_{\Delta_i} = \emptyset, 1 \leq |i| \leq 2\}$ such that

– for $j = 0, \ldots, \ell$, $P_{\boldsymbol{\theta}^\star}(A_j) > 0$.

– for all $(\varphi_0, \ldots, \varphi_\ell) \in A_0 \times \cdots \times A_\ell$ the (ℓ, p) matrix with entries $\left(\boldsymbol{LPL}^{(1)}_{\overline{\Lambda}}(\varphi_i; \boldsymbol{\theta}^\star)\right)_j - \left(\boldsymbol{LPL}^{(1)}_{\overline{\Lambda}}(\varphi_0; \boldsymbol{\theta}^\star)\right)_j$ is injective, with $\overline{\Lambda} := \cup_{i \in \mathbb{B}(0,1)}$.

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

$$\left(\underline{\boldsymbol{U}}^{(2)}(\boldsymbol{\theta}^\star)\right)_{j,k} := \boldsymbol{E}\left(\frac{\partial V}{\partial \theta_j}(0|\Phi; \boldsymbol{\theta}^\star)\frac{\partial V}{\partial \theta_k}(0|\Phi; \boldsymbol{\theta}^\star)\, e^{-V\left(0|\Phi; \boldsymbol{\theta}^\star\right)}\right) \tag{3}$$

and

$$\underline{\boldsymbol{\Sigma}}(\widetilde{D}, \boldsymbol{\theta}^\star) = \widetilde{D}^{-2} \sum_{i \in \mathbb{B}(0,1)} \boldsymbol{E}\left(\boldsymbol{LPL}^{(1)}_{\Delta_0}(\Phi; \boldsymbol{\theta}^\star)\, \boldsymbol{LPL}^{(1)}_{\Delta_i}(\Phi; \boldsymbol{\theta}^\star)^T\right), \tag{4}$$

are definite positive.

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

Theorem 2 *Under the assumptions **[Mod]**, **[C]**, **[N1]**, **[N2]** and **[N3]**, we have the following convergence in distribution as $n \to +\infty$*

$$|\Lambda_n|^{1/2}\, \underline{\boldsymbol{U}}^{(2)}(\boldsymbol{\theta}^\star)\, \left(\widehat{\boldsymbol{\theta}}_n(\Phi) - \boldsymbol{\theta}^\star\right) \to \mathcal{N}\left(0, \underline{\boldsymbol{\Sigma}}(\widetilde{D}, \boldsymbol{\theta}^\star)\right), \tag{5}$$

*where $\underline{\boldsymbol{\Sigma}}(\widetilde{D}, \boldsymbol{\theta}^\star)$ is defined by (4). In addition under the assumption **[N4]***

$$|\Lambda_n|^{1/2}\, \widehat{\underline{\boldsymbol{\Sigma}}}_n(\Phi; \widehat{\boldsymbol{\theta}}_n(\Phi))^{-1/2}\, \underline{\boldsymbol{U}}^{(2)}_n(\Phi; \widehat{\boldsymbol{\theta}}_n(\Phi))\, \left(\widehat{\boldsymbol{\theta}}_n(\Phi) - \boldsymbol{\theta}^\star\right) \to \mathcal{N}\left(0, \underline{\boldsymbol{I}}_p\right), \tag{6}$$

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

$$\widehat{\underline{\boldsymbol{\Sigma}}}_n(\varphi; \boldsymbol{\theta}) = |\Lambda_n|^{-1} \sum_{i \in I_n} \sum_{j \in \mathbb{B}(i,1) \cap I_n} \boldsymbol{LPL}^{(1)}_{\Delta_i}(\varphi; \boldsymbol{\theta})\, \boldsymbol{LPL}^{(1)}_{\Delta_j}(\varphi; \boldsymbol{\theta})^T. \tag{7}$$

In the following the assumption **[N]** will stand for the assumptions **[N1]**, **[N2]**, **[N3]** and **[N4]**.

4 Applications to the LJ-type model

This section focuses on the LJ-type model presented in Section 2.3 and aims at proving the following result.

Proposition 3
(i) Theorem 1 holds for the LJ-type model (with $D \in]0, +\infty]$), that is for the Lennard-Jones and the finite-range Lennard-Jones model.
(ii) Theorem 2 holds only for the finite-range Lennard-Jones model.

The proof of Proposition 3 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:

- Integrabilility type assumptions, i.e. Assumptions **[C1]**, **[C4]**, **[N1]** 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 Φ be a stationary pairwise interaction Gibbs point process assumed to be super-stable, regular and lower regular. For $i = 1, 2$, define $H_i(x|\varphi) = \sum_{y\in\varphi} g_i(||x-y||)$ with g_i a continuous function. 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 $k \geq 0$,*

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

Proof. For all finite configuration φ

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

where

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

Now, the assumptions ensure that g_ε is lower regular in the Ruelle sense. We may now apply the same argument as in Lemma 3 of Mase (1995) to prove the integrability of the random variable $e^{-H_\varepsilon(0|\Phi)}$. ■

Before verifying the different assumptions, let us denote by

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

Since $\boldsymbol{\Theta}$ 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 $\boldsymbol{\theta} \in \boldsymbol{\Theta}$

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

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

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

which satisfies the assumptions of Lemma 4 as soon as $\varepsilon < \left(\frac{\theta_3^\star}{\theta_3}\right)^{12}\frac{\theta_2^\star}{\theta_2}$, 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

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

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

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

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

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

Proof.

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

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

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

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

Let $\boldsymbol{\theta} \in \boldsymbol{\Theta} \setminus \boldsymbol{\theta}^\star$ and consider the following configuration sets defined for $k \geq 1$ and for η small enough by

$$A_0 = \{\varphi \in \Omega : |\varphi \cap \mathcal{B}(0,D)| = 0\} \tag{8}$$

$$A_k(\eta) = \left\{ \varphi \in \Omega : |\varphi \cap \mathcal{B}(0,D)| = |\varphi \cap \mathcal{B}((0, Dk^{-1/12}), \eta)| = 1 \right\}, \tag{9}$$

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 $k = 2, 4$

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

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,

$$D(0|\varphi_0; \boldsymbol{\theta}) = \theta_1 - \theta_1^\star + Z(\varphi_0, D; \boldsymbol{\theta}) - Z(\varphi_0, D; \boldsymbol{\theta}^\star) = 0.$$

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

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

Hence for η small enough, and for both models

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

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

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

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

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

with

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

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

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

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

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

Hence for η small enough, and for both models

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

leading to $\theta_2\theta_3^6=\theta_2^\star\theta_3^{\star 6}$. 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: $\theta_2\theta_3^{12}=\theta_2^\star\theta_3^{\star 12}$. By computing the ratio of the two last equations, one obtains $\theta_3=\theta_3^\star$ and then $\theta_2=\theta_2^\star$.

4.1.3 Assumption [C4]

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

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

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

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

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

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

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

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

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

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

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 $g_{j,\varepsilon}(r) = \widetilde{g}_j^{\inf}(r) - \varepsilon\left|g^{\inf}(r)\right|$. On the one hand, one has

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

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

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

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

4.2 Assumptions [N]

4.2.1 Assumption [N1]

Let us present two auxiliary lemmas.

Lemma 6 *Let φ be the realization of a stationary pairwise interaction point process with local energy function defined by*

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

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

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

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 $\boldsymbol{\theta}\in\boldsymbol{\Theta}$, $g(r;\boldsymbol{\theta})\geq g_{min}$, then*

$$e^{-V\left(\{x_1,\ldots,x_K\}|\varphi;\boldsymbol{\theta}\right)}\leq c_K\prod_{k=1}^{K}e^{-V\left(x_k|\varphi;\boldsymbol{\theta}\right)}\qquad \textit{with } c_K=e^{-\frac{K(K-1)}{2}g_{min}}$$

Proof. The proof is immediate since

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

■

Let $k=1,\ldots,3$ and let $\lambda_1,\ldots,\lambda_k$, k positive integers such that $\sum_{i=1}^{k}\lambda_i=3$ and define the random variable

$$A(\Phi):=\int_{\Delta^k}\prod_{i=1}^{k}\left|\frac{\partial V}{\partial\theta_j}\left(x_i|\Phi;\boldsymbol{\theta}^\star\right)\right|^{\lambda_i}e^{-V\left(\{x_1,\ldots,x_k\}|\Phi;\boldsymbol{\theta}^\star\right)}dx_i.$$

From Lemma 7, we have

$$\begin{aligned}
\boldsymbol{E}\left(A(\Phi)\right) &\leq \boldsymbol{E}\left(c_k \int_{\Delta^k} \prod_{i=1}^{k} \left|\frac{\partial V}{\partial \theta_j}\left(x_i|\Phi;\boldsymbol{\theta}^\star\right)\right|^{\lambda_i} e^{-V\left(x_i|\Phi;\boldsymbol{\theta}^\star\right)} dx_i\right) \\
&= c_k \int_{\Delta^k} \boldsymbol{E}\left(\prod_{i=1}^{k} \left|\frac{\partial V}{\partial \theta_j}\left(x_i|\Phi;\boldsymbol{\theta}^\star\right)\right|^{\lambda_i} e^{-V\left(x_i|\Phi;\boldsymbol{\theta}^\star\right)}\right) dx_1 \ldots dx_k \\
&\leq c_k \int_{\Delta^k} \prod_{i=1}^{k} \boldsymbol{E}\left(\left|\frac{\partial V}{\partial \theta_j}\left(x_i|\Phi;\boldsymbol{\theta}^\star\right)\right|^{k} e^{-\frac{k}{\lambda_i} V\left(x_i|\Phi;\boldsymbol{\theta}^\star\right)}\right)^{1/k} dx_1 \ldots dx_k \\
&= c_k \prod_{i=1}^{k} \int_{\Delta} \boldsymbol{E}\left(\left|\frac{\partial V}{\partial \theta_j}\left(x_i|\Phi;\boldsymbol{\theta}^\star\right)\right|^{k} e^{-\frac{k}{\lambda_i} V\left(x_i|\Phi;\boldsymbol{\theta}^\star\right)}\right)^{1/k} dx_i \\
&= c_k |\Delta|^k \prod_{i=1}^{k} \boldsymbol{E}\left(\left|\frac{\partial V}{\partial \theta_j}\left(0|\Phi;\boldsymbol{\theta}^\star\right)\right|^{k} e^{-\frac{k}{\lambda_i} V\left(0|\Phi;\boldsymbol{\theta}^\star\right)}\right)^{1/k}
\end{aligned}$$

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

4.2.2 Assumption [N2]

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

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

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 $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ $g_{j,k,\varepsilon}(r) := g^{LJ}(r;\boldsymbol{\theta}) - \varepsilon \left|\frac{\partial^2 g^{LJ}}{\partial \theta_j \partial \theta_k}(r;\boldsymbol{\theta})\right|$. 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 $\varepsilon < \frac{\left(\theta^{\inf}\right)^2}{12}$. 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 $\varepsilon < \frac{\left(\theta^{\inf}\right)^3}{132}$.

4.2.3 Assumption [N3]

Let $\boldsymbol{y} = (y_1, y_2, y_3) \in \mathbb{R}^3$ and $g(\boldsymbol{y}, \varphi) := \boldsymbol{y}^T \boldsymbol{V}_{LJ}^{(1)}\left(0|\varphi;\boldsymbol{\theta}^\star\right)$. Let $\varphi_0 \in A_0$ and $\varphi_k(\eta) \in A_k(\eta)$ $(k = 2, 4)$ where A_0 and $A_k(\eta)$ are defined by (8) and (9). Assume $g(\boldsymbol{y}, \varphi_k) = 0$ for $k = 0, 2, 4$.

Since, $g(\boldsymbol{y},\varphi_0)=y_1$, we have $y_1=0$. Now,

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

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

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

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

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

Hence for η small enough,

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

leading to the equation $\theta_3^\star y_2+6\theta_2^\star y_3=0$. By considering the linear combination $\sqrt{2}g(\boldsymbol{y},\varphi_2)-g(\boldsymbol{y},\varphi_4)$, we may obtain the equation $\theta_3^\star y_2+12\theta_2^\star y_3=0$ 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 $k=1,\cdots,\ell$ and for all $\varphi_k\in A_k$ and $\varphi_0\in A_0$:

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

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

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

Concerning this assumption, we choose $\varphi_0\in A_0=\left\{\varphi\in\overline{\Omega}:\varphi_{\Delta_0}=\emptyset\right\}$. Let $\boldsymbol{y}\in\mathbb{R}^3$ then

$$\int_{\overline{\Lambda}}\boldsymbol{y}^T\boldsymbol{V}_{LJ}^{(1)}\left(x|\varphi_0;\boldsymbol{\theta}^\star\right)e^{-V^{LJ}\left(x|\varphi_0;\boldsymbol{\theta}^\star\right)}dx=y_1e^{-\theta_1^\star}\left|\overline{\Lambda}\right|\quad\text{and}\quad\sum_{x\in\varphi_0\cap\overline{\Lambda}}\boldsymbol{y}^T\boldsymbol{V}_{LJ}^{(1)}\left(x|\varphi_0\setminus x;\boldsymbol{\theta}^\star\right)=0.$$

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

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

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 $x \in \overline{\Lambda}$, then one may prove that for $j = 2,3$

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

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

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

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

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

where

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

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

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

Since

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

For η and ε chosen small enough, one may prove that

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

leading to

$$2y_2 4\theta_3^{\star 12} + 2y_3 4\theta_2^\star 12\theta_3^{\star 11} = 0 \Leftrightarrow \theta_3^\star y_2 + 12\theta_2^\star y_3 = 0. \tag{11}$$

This means that

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

With similar arguments, we obtain that

$$2y_2 4\theta_3^{\star 6} + 2y_3 4\theta_2^\star 6\theta_3^{\star 5} = 0 \Leftrightarrow \theta_3^\star y_2 + 6\theta_2^\star y_3 = 0. \tag{12}$$

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

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

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

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

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

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

Hence for k large enough and for η small enough, we may obtain

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

which leads to $y_1=0$.
Case 2: $\theta_3^\star\ge D$**.** First note that for all $x\in\Lambda_2$,

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

On the one hand, for k large enough and for η small enough, we may have

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

On the other hand, we have for k large enough

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

Therefore for k large enough and for η small enough, we have

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

which leads to $y_1=0$.

5 Annex: proofs of Theorems 1 and 2

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 point process Φ with probability measure P is stationary. Let $\Lambda \Subset \mathbb{R}^2$, $\varphi \in \Omega$ and let g be a function satisfying $g(x,\varphi) = g(0,\tau_x\varphi)$ for all $x \in \mathbb{R}^2$. Define $f(\varphi) = g(0,\varphi)e^{-V(0|\varphi)}$ and assume that $f \in L^1(P)$. Then,*

$$\boldsymbol{E}\left(\sum_{x\in\Phi_\Lambda} g(x,\Phi\setminus x)\right) = \boldsymbol{E}\left(\int_\Lambda g(x,\Phi)e^{-V(x|\Phi)}dx\right) = |\Lambda|\ \boldsymbol{E}\Big(\, g\,(0,\Phi)\, e^{-V(0|\Phi)}\,\Big) \tag{13}$$

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 Δ_0 be 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 $x \in \mathbb{R}^2$,*

$$H_{\tau_x G}(\tau_x\varphi) = H_G(\varphi), \quad \textit{for a.e. } \varphi$$

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

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

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

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

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

5.1 Proof of Theorem 1

Due to the decomposition of stationary measures as a mixture of ergodic measures (see Preston (1976)), one only needs to prove Theorem 1 by assuming that $P_{\boldsymbol{\theta}^\star}$ is ergodic. From now on, $P_{\boldsymbol{\theta}^\star}$ 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 $U_n(\Phi;\boldsymbol{\theta})$.

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

$$H_{1,\Lambda_n}(\varphi) = \int_{\Lambda_n} e^{-V(x|\varphi;\boldsymbol{\theta})}dx \quad \text{and} \quad H_{2,\Lambda_n}(\varphi) = \sum_{x\in\Phi_{\Lambda_n}} V\left(x|\varphi\setminus x;\boldsymbol{\theta}\right).$$

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

$$\frac{1}{|\Lambda_n|} H_{1,\Lambda_n}(\Phi) \to \boldsymbol{E}\Big(e^{-V\left(0|\Phi;\boldsymbol{\theta}\right)}\Big). \tag{14}$$

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

$$|H_{2,G}(\varphi)| \le \sum_{x\in\varphi_G} |V\left(x|\varphi\setminus x;\boldsymbol{\theta}\right)| \le \sum_{x\in\varphi_{\Delta_0}} |V\left(x|\varphi\setminus x;\boldsymbol{\theta}\right)|.$$

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

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

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

$$\frac{1}{|\Lambda_n|}H_{2,\Lambda_n}(\Phi)\to\frac{1}{|\Delta_0|}\boldsymbol{E}\Big(\sum_{x\in\Phi_{\Delta_0}}V\left(x|\Phi\setminus x;\boldsymbol{\theta}\right)\Big)=\boldsymbol{E}\left(V\left(0|\Phi;\boldsymbol{\theta}\right)e^{-V\left(0|\Phi;\boldsymbol{\theta}^\star\right)}\right). \tag{15}$$

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

$$U_n(\Phi;\boldsymbol{\theta})\to U(\boldsymbol{\theta})=\boldsymbol{E}\Big(e^{-V\left(0|\Phi;\boldsymbol{\theta}\right)}\ +\ V\left(0|\Phi;\boldsymbol{\theta}\right)e^{-V\left(0|\Phi;\boldsymbol{\theta}^\star\right)}\Big) \tag{16}$$

Step 2. $U_n(\cdot;\boldsymbol{\theta})$ is a contrast function

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

$$K(\boldsymbol{\theta},\boldsymbol{\theta}^\star)=\boldsymbol{E}\Big(e^{-V\left(0|\Phi;\boldsymbol{\theta}^\star\right)}\Big(e^{V\left(0|\Phi;\boldsymbol{\theta}\right)-V\left(0|\Phi;\boldsymbol{\theta}^\star\right)}-\Big(1+V\left(0|\Phi;\boldsymbol{\theta}\right)-V\left(0|\Phi;\boldsymbol{\theta}^\star\right)\Big)\Big)\Big). \tag{17}$$

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

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

for $P_{\boldsymbol{\theta}^\star}-$a.e. φ. Let us consider the ℓ events A_j $(j=1,\ldots,\ell)$ defined in Assumption **[C2]**. The previous equation is at least true for $\varphi_j\in A_j$, which leads under Assumption **[C2]** to $\boldsymbol{\theta}=\boldsymbol{\theta}^\star$. Therefore, $K(\boldsymbol{\theta},\boldsymbol{\theta}^\star)=0\Rightarrow\boldsymbol{\theta}=\boldsymbol{\theta}^\star$. The converse is trivial.

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

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

is such that there exists a sequence $(\varepsilon_k)_{k\geq1}$, with $\varepsilon_k\to0$ as $k\to+\infty$ such that for all $k\geq1$

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

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

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

and

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

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

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

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

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

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

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

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

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

with $\gamma_2:=\boldsymbol{E}\left(\sum_{j=1}^p\max_{\boldsymbol{\theta}\in\boldsymbol{\Theta}}\left|\frac{\partial V}{\partial\theta_j}\left(0|\Phi;\boldsymbol{\theta}\right)e^{-V\left(0|\Phi;\boldsymbol{\theta}^\star\right)}\right|^2\right)$. Hence, for all $k\geq 1$ and for all $\boldsymbol{\theta},\boldsymbol{\theta}'\in\boldsymbol{\Theta}$ 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

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

Since γ_1 and γ_2 are independent of $\boldsymbol{\theta}$ and $\boldsymbol{\theta}'$, we have for all $n\geq n_0(k)$

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

Finally, since

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

for $P_{\boldsymbol{\theta}^\star}-$a.e. φ, 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.

5.2 Proof of Theorem 2

Step 1. Asymptotic normality of $\boldsymbol{U}_n^{(1)}(\Phi;\boldsymbol{\theta}^\star)$

The aim is to prove the following convergence in distribution as $n \to +\infty$

$$|\Lambda_n|^{1/2}\, \boldsymbol{U}_n^{(1)}(\Phi;\boldsymbol{\theta}^\star) \to \mathcal{N}\left(0, \underline{\boldsymbol{\Sigma}}(\widetilde{D},\boldsymbol{\theta}^\star)\right) \tag{19}$$

where the matrix $\underline{\boldsymbol{\Sigma}}(\widetilde{D},\boldsymbol{\theta}^\star)$ is defined by (4).

The idea is to apply to $\boldsymbol{U}_n^{(1)}(\Phi;\boldsymbol{\theta}^\star)$ 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 $j = 1,\ldots,p$

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

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

(iii) The matrix $\mathbb{V}\mathrm{ar}\left(|\Lambda_n|^{1/2}\boldsymbol{U}_n^{(1)}(\Phi;\boldsymbol{\theta}^\star)\right)$ converges to the matrix $\underline{\boldsymbol{\Sigma}}(\widetilde{D},\boldsymbol{\theta}^\star)$.

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

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

Recall that for any configuration φ

$$\left(\boldsymbol{LPL}_{\Delta_0}^{(1)}(\varphi;\boldsymbol{\theta}^\star)\right)_j = -\int_{\Delta_0} \frac{\partial V}{\partial \theta_j}(x|\varphi;\boldsymbol{\theta}^\star)\, e^{-V(x|\varphi;\boldsymbol{\theta}^\star)} dx + \int_{\Delta_0} \frac{\partial V}{\partial \theta_j}(x|\varphi\setminus x;\boldsymbol{\theta}^\star)\,\varphi(dx). \tag{20}$$

Denote respectively by $G_1(\varphi)$ and $G_2(\varphi)$ the first and the second right-hand term of (20) and by $E_i = \boldsymbol{E}\left(G_i(\Phi)|\Phi_{\Delta_0^c} = \varphi_{\Delta_0^c}\right)$. Let us define for any φ, the measure $\mu_\varphi := \sum_{x\in\varphi}\delta_x$. From the definition of Gibbs point processes,

$$E_2 = \frac{1}{Z_{\Delta_0}(\varphi_{\Delta_0^c})}\int_{\Omega_{\Delta_0}} \pi_{\Delta_0}(d\varphi_{\Delta_0})\int_{\mathbb{R}^2}\mu_{\varphi_{\Delta_0}}(dx)\mathbf{1}_{\Delta_0}(x)\frac{\partial V}{\partial\theta_j}(x|\varphi\setminus x;\boldsymbol{\theta}^\star)\, e^{-V_{\Delta_0}(\varphi;\boldsymbol{\theta}^\star)}.$$

Since π is a Poisson process,

$$\int_{\Omega_{\Delta_0}} \pi_{\Delta_0}(d\varphi_{\Delta_0}) f(\varphi) = \int_{\Omega_{\Delta_0}} \pi_{\Delta_0}(d\varphi_{\Delta_0})\int_{\Omega_{\Delta_0^c}} \pi_{\Delta_0^c}(d\varphi'_{\Delta_0^c}) f(\varphi)$$

and therefore, by introducing $\psi := \varphi_{\Delta_0}\cup\varphi'_{\Delta_0^c}$

$$E_2 = \frac{1}{Z_{\Delta_0}(\varphi_{\Delta_0^c})}\int_{\Omega}\pi(d\psi)\int_{\mathbb{R}^2}\mu_\psi(dx)\mathbf{1}_{\Delta_0}(x)\frac{\partial V}{\partial\theta_j}\left(x|\psi_{\Delta_0}\cup\varphi_{\Delta_0^c}\setminus x;\boldsymbol{\theta}^\star\right) e^{-V_{\Delta_0}\left(\psi_{\Delta_0}\cup\varphi_{\Delta_0^c};\boldsymbol{\theta}^\star\right)}.$$

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

$$E_2 = \frac{1}{Z_{\Delta_0}(\varphi_{\Delta_0^c})} \int_{\Delta_0} dx \int_{\Omega} \pi_x^!(d\psi) \frac{\partial V}{\partial \theta_j} \left(x|\psi_{\Delta_0} \cup \varphi_{\Delta_0^c}; \boldsymbol{\theta}^\star\right) e^{-V_{\Delta_0}\left(x \cup \psi_{\Delta_0} \cup \varphi_{\Delta_0^c}; \boldsymbol{\theta}^\star\right)},$$

where $\pi_x^!$ stands for the reduced Palm distribution of the Poisson point process. Since from Slivnyak-Mecke Theorem (see *e.g.* Møller and Waagepetersen (2003)), $\pi = \pi_x^!$, one can obtain

$$\begin{aligned} E_2 &= \frac{1}{Z_{\Delta_0}(\varphi_{\Delta_0^c})} \int_{\Omega} \pi(d\psi) \int_{\Delta_0} dx \, \frac{\partial V}{\partial \theta_j} \left(x|\psi_{\Delta_0} \cup \varphi_{\Delta_0^c}; \boldsymbol{\theta}^\star\right) e^{-V_{\Delta_0}\left(x \cup \psi_{\Delta_0} \cup \varphi_{\Delta_0^c}; \boldsymbol{\theta}^\star\right)} \\ &= \frac{1}{Z_{\Delta_0}(\varphi_{\Delta_0^c})} \int_{\Omega_{\Delta_0}} \pi_{\Delta_0}(d\varphi_{\Delta_0}) \int_{\Delta_0} dx \frac{\partial V}{\partial \theta_j} \left(x|\varphi; \boldsymbol{\theta}^\star\right) e^{-V\left(x|\varphi; \boldsymbol{\theta}^\star\right)} e^{-V_{\Delta_0}\left(\varphi; \boldsymbol{\theta}^\star\right)} \\ &= -E_1 \end{aligned}$$

<u>Condition (ii)</u> : For any bounded domain Δ one may write for $j = 1, \ldots, p$

$$\left| \left(\boldsymbol{LPL}_{\Delta}^{(1)} \left(\Phi; \boldsymbol{\theta}^\star\right) \right)_j \right|^3 \leq 4 \left| \int_{\Delta} \frac{\partial V}{\partial \theta_j} \left(x|\Phi; \boldsymbol{\theta}^\star\right) e^{-V\left(x|\Phi; \boldsymbol{\theta}^\star\right)} dx \right|^3 + 4 \left| \sum_{x \in \varphi_\Delta} \frac{\partial V}{\partial \theta_j} \left(x|\Phi \setminus x; \boldsymbol{\theta}^\star\right) \right|^3.$$

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

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

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

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

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

Let $j \in I_n \cap \mathbb{B}(i,1)^c$, since $\boldsymbol{LPL}^{(1)}_{\Delta_i}(\varphi;\boldsymbol{\theta}^\star)$ is a measurable function of $\varphi_{\Delta_i^c}$, we have by using condition (i):

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

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

$$\widetilde{I}_n = I_n \cap \left(\cup_{i\in\partial I_n}\mathbb{B}(i,1)\right).$$

We now obtain

$$\begin{aligned}
\mathbb{V}\mathrm{ar}\left(|\Lambda_n|^{1/2}\boldsymbol{U}^{(1)}_n(\Phi;\boldsymbol{\theta}^\star)\right) &= |\Lambda_n|^{-1}\sum_{i\in I_n}\sum_{j\in I_n\cap\mathbb{B}(i,1)}E_{i,j}\\
&= |\Lambda_n|^{-1}\left(\sum_{i\in I_n\setminus\widetilde{I}_n}\sum_{j\in I_n\cap\mathbb{B}(i,1)}E_{i,j}+\sum_{i\in\widetilde{I}_n}\sum_{j\in I_n\cap\mathbb{B}(i,1)}E_{i,j}\right)
\end{aligned}$$

Using the stationarity and the definition of the domain Λ_n, one obtains

$$|\Lambda_n|^{-1}\sum_{i\in I_n\setminus\widetilde{I}_n}\sum_{j\in I_n\cap\mathbb{B}(i,1)}E_{i,j} = |\Lambda_n|^{-1}|I_n\setminus\widetilde{I}_n|\sum_{j\in\mathbb{B}(0,1)}E_{0,j}\to\underline{\boldsymbol{\Sigma}}(\widetilde{D},\boldsymbol{\theta}^\star)\quad\text{as } n\to+\infty$$

and

$$|\Lambda_n|^{-1}\left|\sum_{i\in\widetilde{I}_n}\sum_{j\in I_n\cap\mathbb{B}\left(i,\lceil\frac{D}{\widetilde{D}}\rceil\right)}E_{i,j}\right|\le|\Lambda_n|^{-1}|\widetilde{I}_n|\sum_{j\in\mathbb{B}(0,1)}|E_{0,j}|\to 0\quad\text{as } n\to+\infty.$$

Hence as $n\to+\infty$

$$\begin{aligned}
\mathbb{V}\mathrm{ar}\left(|\Lambda_n|^{1/2}\boldsymbol{U}^{(1)}_n(\Phi;\boldsymbol{\theta}^\star)\right) &= |\Lambda_n|^{-1}\sum_{i\in I_n}\sum_{j\in I_n\cap\mathbb{B}(i,1)}E_{i,j}\\
&\overset{n\to+\infty}{\longrightarrow} \underbrace{|I_n||\Lambda_n|^{-1}}_{\widetilde{D}^{-2}}\sum_{k\in\mathbb{B}(0,1)}E_{0,k} = \underline{\boldsymbol{\Sigma}}(\widetilde{D},\boldsymbol{\theta}^\star).
\end{aligned}\tag{21}$$

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

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

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

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

Note that from Theorem 9 (Nguyen and Zessin (1979)), $\underline{\boldsymbol{U}}_n^{(2)}(\cdot;\boldsymbol{\theta}^\star)$ converges almost surely as $n \to +\infty$ towards $\underline{\boldsymbol{U}}^{(2)}(\boldsymbol{\theta}^\star)$ defined by (3). Note that $\underline{\boldsymbol{U}}^{(2)}(\boldsymbol{\theta}^\star)$ is a symmetric positive matrix since for all $\boldsymbol{y} \in \mathbb{R}^p$

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

where for $j = 1,\ldots,p$, $\varphi \in \Omega$ and for $\boldsymbol{\theta} \in \mathcal{V}(\boldsymbol{\theta}^\star)$ $\left(\boldsymbol{V}^{(1)}(x|\varphi;\boldsymbol{\theta}^\star)\right)_j := \frac{\partial V}{\partial\theta_j}(x|\varphi;\boldsymbol{\theta})$ 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 of Guyon (1992) in order to obtain (5).

It remains to prove (6). This may de done in two different steps. The first one consists in verifying the positive definiteness of the matrix $\underline{\boldsymbol{\Sigma}}(\widetilde{D},\boldsymbol{\theta}^\star)$. The proof is strictly similar to the one of Billiot et al. (2008) (p. 261) except that the assumption **[SDP]** is now simply replaced by the more general one assumption **[N4]**. Now, the convergence in probability of $\widehat{\underline{\boldsymbol{\Sigma}}}_n(\Phi;\widehat{\boldsymbol{\theta}}_n(\Phi))$ towards $\underline{\boldsymbol{\Sigma}}(\widetilde{D},\boldsymbol{\theta}^\star)$ is obtained by applying Proposition 9 of Coeurjolly and Lavancier (2010).

Acknowledgements

We are grateful to the referee and the associate editor for their comments which helped us in improving a previous version.

References

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

E. Bertin, J.-M. Billiot, and R. Drouilhet. Existence of "Nearest-Neighbour" Gibbs Point Models. *Ann. Appl. Probab.*, 31:895–909, 1999a.

E. Bertin, J.-M. Billiot, and R. Drouilhet. Spatial Delaunay Gibbs Point Processes. *Stochastic Models*, 15(2):181–199, 1999b.

E. Bertin, J.-M. Billiot, and R. Drouilhet. k-Nearest-Neighbour Gibbs Point Processes. *Markov Processes and Related Fields*, 5(2):219–234, 1999c.

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. Prob.*, 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.

J.-F. Coeurjolly and F. Lavancier. Residuals for stationary marked Gibbs point processes. *submitted for publication*, 2010. `http://hal.archives-ouvertes.fr/hal-00453102/fr/`.

D. Daley and D. Vere-Jones. *An introduction to the Theory of Point Processes.* Springer Verlag, New York, 1988.

D. Dereudre. Gibbs Delaunay tessellations with geometric hardcore condition. *J. Stat. Phys.*, 121(3-4):511–515, 2005.

D. Dereudre. The existence of quermass-interaction processes for nonlocally stable interaction and nonbounded convex grains. *Adv. in Appl. Probab*, 41(3):664–681, 2009.

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

D. Dereudre, R. Drouilhet, and H.O. Georgii. Existence of Gibbsian point processes with geometry-dependent interactions. *submitted*, 2010. http://arxiv.org/abs/1003.2875.

M. Goulard, A. Särkkä, and P. Grabarnik. Parameter estimation for marked Gibbs point processes through the maximum pseudo-likelihood method. *Scandinavian Journal of Statistics*, 23(3):365–379, 1996.

X. Guyon. *Champs aléatoires sur un réseau.* Masson, Paris, 1992.

J. Illian, A. Penttinen H., and Stoyan. *Statistical analysis and modelling of spatial point patterns.* Wiley-Interscience, 2008.

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.

W. S. Kendall, M. N. M. Van Lieshout, and A. J. Baddeley. Quermass-interaction processes conditions for stability. *Advances on applied probability*, 31:315–342, 1999.

S. Mase. Consistency of maximum pseudo-likelihood estimator of continuous state space gibbsian process. *Ann. Appl. Probab.*, 5:603–612, 1995.

S. Mase. Marked gibbs processes and asymptotic normality of maximum pseudo-likelihhod estimators. *Math. Nachr.*, 209:151–169, 2000.

J. Møller. Parametric methods for spatial point processes. Technical Report Research Report R-2008-04, Department of Mathematical Sciences, Aalborg University, 2008.

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.

Y. Ogata and M. Tanemura. Estimation of interaction potentials of spatial point patterns through the maximum likelihood procedure. *Annals of the Institute of Statistical Mathematics*, 33(1): 315–338, 1981.

C.J. Preston. *Random fields*. Springer Verlag, 1976.

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

D. Ruelle. Superstable interactions in classical statistical mechanics. *Commun. Math. Phys.*, 18: 127–159, 1970.

D. Stoyan, W.S. Kendall, J. Mecke, and L. Ruschendorf. *Stochastic geometry and its applications*. John Wiley and Sons, Chichester, 1987.