Maximum pseudo-likelihood estimator for nearest-neighbours Gibbs point processes

Jean-Michel Billiot, Jean-François Coeurjolly, Rémy Drouilhet

Maximum pseudo-likelihood estimator for nearest-neighbours Gibbs point processes

By Jean-Michel Billiot[1], Jean-François Coeurjolly[1,2] and Rémy Drouilhet[1]

Labsad, University of Grenoble 2, France

January 3, 2006

[1]Supported by a grant from IMAG Project AlpB.

AMS 2000 subject classifications. Primary 60K35,60G55 ; secondary 62M30.

Key words and phrases : Gibbs point processes, nearest-neighbours models, Delaunay graph, pseudo-likelihood method, minimum contrast estimators.

[2]Corresponding author

Abstract

This paper is devoted to the estimation of a vector $\boldsymbol{\theta}$ parametrizing an energy function associated to some "Nearest-Neighbours" Gibbs point process, via the pseudo-likelihood method. We present some convergence results concerning this estimator, that is strong consistency and asymptotic normality, when only a single realization is observed. Sufficient conditions are expressed in terms of the local energy function and are verified on some examples.

1 Introduction

Gibbs point processes first appeared in the theory of statistical physics. Historical aspects of the mathematical theory are covered briefly in Kallenberg (1983). The importance of the Gibbs point process as a model building principle became widely recognized through these works. Indeed, the class of Gibbs point processes is interesting because it allows to introduce and study interactions between points through the modelling of an associated potential function. This resulting gain explains their use in statistical physics Ruelle (1969), Feynman (1972) (when taking interactions between molecules in models of dilute gases into account) or in ecology (when analysing competitions between plants). Within the mechanics statistics framework, Gibbs states are defined as solutions of the well known equilibrium equations refered to Dobrushin-Lanford-Ruelle (D.L.R.) equations Dobrushin (1969), Lanford and Ruelle (1969). One way to introduce Gibbs point processes consists in using a family of local specifications with respect to a weight process. The Preston's theorems (Preston (1976)) used precisely this approach in order to give sufficient conditions on local specifications for the existence of Gibbs states.

Many proposals tried to estimate the potential function from the available point pattern data generated by some Gibbs point processes. If the potential belongs to a parametric family model, the most well-known methodology is the use of the likelihood function. The main drawback of this approach is that the likelihood function contains an unknown scaling factor whose value depends on the parameters and which is difficult to calculate. The first class of models on which the estimation of the maximum likelihood has been undertaken is the class of pairwise interaction point processes. Ogata and Tanemura (1984) developed the maximum likelihood estimation method based on numerical approximations of the likelihood. Penttinen (1984) used a similar approach while applying a Monte Carlo method in a way to solve the likelihood equation by the stochastic Newton-Raphson algorithm. Moyeed and Baddeley (1991) proposed another iterative procedure for estimating the maximum likelihood estimator. For maximum likelihood by Markov chain Monte Carlo, see Geyer and Thompson (1992), Geyer (1999) and for U.L.A.N. conditions for maximum likelihood estimator, see Mase (1992). An alternative approach consists in avoiding to optimize the likelihood function (because of the scaling factor problem) and introducing a pseudo-likelihood function instead. This idea originated from Besag (1974) in the study of lattice processes. Besag et al. (1982) further considered this method for pairwise interaction point process, while Jensen and Møller (1991) generalized it to the general class of Gibbs point processes, see Mase (1995), Mase (1999), Jensen and Künsch (1994), Guyon (1991) for asymptotic properties. A third way is the Takacs-Fiksel estimation method (Takacs (1986) Fiksel (1988)), which relies on a characteristic property of Gibbs processes using Palm measure. Asymptotics properties of Takacs-Fiksel estimator are studied in Heinrich (1992), Billiot (1997). A comparison of these different procedures applied to the Strauss model is presented in Diggle et al. (1994). The non parametric setting has been undertaken by Glötzl and Rauschenschwandtner (1981) and Diggle et al. (1987) (and the references therein). Heikkinen and Penttinen (1999) proposed a semiparametric estimator based on Bayesian smoothing techniques. A general review of the problem of statistical inference on spatial point processes can be found in the recent monograph of Møller and Waggepetersen (2003).

The present study is devoted to "Nearest-Neighbour" Gibbs point models by combining stochastic geometry arguments (Stoyan et al. (1995)) and computational geometry ones (Preparata and Shamos (1988), Edelsbrunner (1988), Boissonnat and Yvinec (1995)). Such models are introduced by Baddeley and Møller (1989) where the neighbourhood relation depends on the realization of the process. Sufficient conditions (expressed in terms of the energy function) for the existence of such processes are proposed in Bertin et al. (1999b) and Bertin et al. (1999a), where some examples are also proposed. The main one is a pairwise interaction point process where the neighbourhood relation corresponds to the (slightly modified) Delaunay graph of the realization of the process.

In this paper, we study a pseudo-likelihood estimator for such processes. More precisely, our framework is restricted to stationary Gibbs point processes based on energy function related to some graph (for instance the Delaunay graph) such that the energy function is invariant by translation and such that the local energy function is stable and quasi-local (or local). The main results of this paper are convergence results (strong consistency and asymptotic normality) of maximum pseudo-likelihood estimators in this framework. These results are obtained when only a single realization is observed. Sufficient conditions are expressed in terms of the local energy function (which makes the results quite general) for some large family of parametrized energy functions. Among the different parametrizations, the exponential family is considered.

The paper is organized as follows. Section 2 is devoted to some background on Gibbs point processes and to the description of our framework. The statistical model and the pseudo-likelihood method are presented in Section 3. Consistency and asymptotic normality of the maximum pseudo-likelihood estimator are respectively proved in Section 4 and Section 5. Finally, the different sufficient conditions ensuring convergence results are verified on some examples in Section 6. A short simulation is presented to check the effectiveness of maximum pseudo-likelihood estimator.

2 Background on Gibbs point processes

2.1 Gibbs point processes

We define $\mathcal{B}$, $\mathcal{B}_b$ to be respectively the Borel σ-field and the bounded Borel boolean ring.

Let Ω denotes the class of locally finite subsets of $\mathbb{R}^d$. In particular, an element φ of Ω, also called configuration (of points), could be represented as $\varphi = \sum_{i \in \mathbb{N}} \delta_{x_i}$ which is a simple counting Radon measure in $\mathbb{R}^d$ (i.e. all the points x_i of $\mathbb{R}^d$ are distinct) where for every $\Lambda \in \mathcal{B}$, $\delta_x(\Lambda) = 1_\Lambda(x)$ is the Dirac measure and $1_A(.)$ is the indicator function of a set A. This space Ω is equipped with the vague topology, that is to say the weak topology for Radon measures with respect to the set of continuous functions vanishing outside a compact set. We also define the σ-field $\mathcal{F}$ spanned by the maps $\varphi \longrightarrow \varphi(\Lambda)\,, \Lambda \in \mathcal{B}_b$, where $\varphi(\Lambda)$ corresponds to the number of points of φ in Λ due to the Radon measure representation of φ. The set of all configurations in a measurable set $\Lambda \subset \mathbb{R}^d$ will be denoted by Ω_Λ and the corresponding σ-field $\mathcal{F}_\Lambda$ is similarly defined. Furthermore, for any $\Lambda \in \mathcal{B}_b$,

$$(\Omega, \mathcal{F}) = (\Omega_\Lambda, \mathcal{F}_\Lambda) \times (\Omega_{\Lambda^c}, \mathcal{F}_{\Lambda^c})$$

where $\Lambda^c = \mathbb{R}^d \setminus \Lambda$ denotes the complementary of Λ in $\mathbb{R}^d$. Finally, Ω_f denotes the class of all finite subsets of $\mathbb{R}^d$.

A point process on $\mathbb{R}^d$ is a Ω-valued random variable, denoted by Φ, with probability distribution P on $(\Omega, \mathcal{F})$. and the intensity measure Λ_p of P is defined as a measure on $\mathcal{B}$ such that for any $D \in \mathcal{B}$

$$\Lambda_p(D) \;=\; \int_\Omega \varphi(D)\; P(d\varphi).$$

In the stationary case, $\Lambda_p(D) \;=\; \lambda_p \nu(D)$ where the constant λ_p is called the intensity of P and ν is the Lebesgue measure on $\mathbb{R}^d$.

A Gibbs point process is usually defined using a family of local specifications with respect to a weight process (often a stationary Poisson process with distribution Q and intensity $\lambda_Q = 1$). Let Λ be a bounded region in $\mathbb{R}^d$. For such a process, given some configuration φ_{Λ^c} on Λ^c, the conditional probability on Λ is of the form, for any $Y \in \mathcal{F}$:

$$\Pi_\Lambda(\varphi, Y) = \left\{ \frac{1}{Z_\Lambda(\varphi)} \int_{\Omega_\Lambda} \exp\left(-V\left(\psi|\varphi_{\Lambda^c}\right)\right) 1_Y(\psi \cup \varphi_{\Lambda^c}) Q_\Lambda(d\psi) \right\} 1_{R_\Lambda}(\varphi),$$

where

$$Z_\Lambda(\varphi) = \int_{\Omega_\Lambda} \exp\left(-V\left(\psi|\varphi_{\Lambda^c}\right)\right) Q_\Lambda(d\psi)$$

is called the partition function and $R_\Lambda = \{\varphi \in \Omega \;\; : 0 < Z_\Lambda(\varphi) < \infty\}$.

Whereas the finite energy function $V(\varphi)$ measures the cost of any configuration, the local energy $V(\psi|\varphi)$ is defined as the energy required to add the points of ψ in φ :

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

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

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

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

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

For the general theory of Gibbs point processes, the reader may refer to Kallenberg (1983); Daley and Vere-Jones (1988); Stoyan et al. (1995) and the references therein.

2.2 Campbell and Palm measures and Glötz Theorem

The reduced Campbell measure $\mathcal{C}_p^!$ of P is a measure on $\mathcal{B}\otimes\mathcal{F}$ such that for any $D\in\mathcal{B}$ and any $Y\in\mathcal{F}$

$$\mathcal{C}_p^!(D\times Y) = \int_\Omega \int_D 1_Y(\varphi - \delta_x)\,\varphi(dx)\ P(d\varphi).$$

When some measurable function h from $\mathbb{R}^d \times \Omega$ on $\mathbb{R}^d$ is given, the following equation is often called the refined Campbell theorem

$$\int_\Omega \sum_{x\in\varphi} h(x, \varphi - \delta_x) P(d\varphi) = \int_{\mathbb{R}^d\times\Omega} h(x,\varphi)\mathcal{C}_p^!(d(x,\varphi)).$$

If the intensity measure Λ_p is $\sigma-$finite, then for Λ_p- a.a. $x\in\mathbb{R}^d$, the distribution $P_x^!$ on $(\Omega,\mathcal{F})$ exists. It is unique for Λ_p- a.a. $x\in\mathbb{R}^d$ and such that

$$\mathcal{C}_p^!(D\times Y) = \int_D P_x^!(Y)\ \Lambda_p(dx) \quad \text{for any} \quad D\in\mathcal{B},\, Y\in\mathcal{F}.$$

Then $P_x^!$ is called the reduced Palm distribution of the point process P with respect to point x. Intuitively, the Palm distribution P_x is the conditional probability of configurations of the point process given that the point x belongs to the realization φ. Therefore, we have

$$\int_\Omega \sum_{x\in\varphi} h(x, \varphi - \delta_x) P(d\varphi) = \int_{\mathbb{R}^d\times\Omega} h(x,\varphi) P_x^!(d\varphi)\Lambda_p(dx).$$

When the process is stationary, one may apply the previous equation by replacing the intensity measure Λ_p by its expression in this case $\lambda_p \nu$, and in this framework, Glötzl (1980) proved that $P \in \mathcal{G}_0(V)$ if and only if the reduced Campbell measure $\mathcal{C}_p^!$ is absolutely continuous with respect to $\nu \times P$ and :

$$\frac{d\mathcal{C}_p^!}{d(\nu \times P)}(x,\varphi) = \lambda_P \frac{dP_x^!}{dP}(\varphi) = \exp(-V(x|\varphi))$$

where $\lambda_P = \int_\Omega \exp(-V(x|\varphi)) P(d\varphi)$ is the intensity of the process P. In the particular case when $V(x|\varphi) = 0$, the point process corresponds to the stationary Poisson process Q. We know from the Slivnyak's theorem that $Q_x^! = Q$ which is one way of characterizing such process.

2.3 Description of some Gibbs models

This paper is mainly devoted to the statistical study of some nearest-neighbours Gibbs point processes first introduced in Baddeley and Møller (1989). More precisely, we are interested in models based on energy function of the form

$$V(\varphi) = \sum_{k=1}^{3} \sum_{\xi \in Del_k(\varphi)} u^{(k)}(\xi, \varphi), \tag{1}$$

where $Del_k(\varphi)$ is the set of clique of order k of the Delaunay graph defined just below. For some $\varphi \in \Omega$ in general position, one defines $Del_3(\varphi)$ by the unique decomposition into triangles ψ in which the convex hull of the circle $C(\psi)$ does not contain any point of $\varphi \setminus \psi$. The Delaunay graph is then defined by the set of edges :

$$Del_2(\varphi) = \cup_{\psi \in Del_3(\varphi)} \mathcal{P}_2(\psi).$$

In order to ensure the existence of such Gibbs state in $\mathbb{R}^d$, Bertin et al. (1999b) prove that the local stability and quasilocality properties (only expressed in terms of the energy function) are sufficient conditions of Preston's Theorem.

Without any additional modification, the previous model does not satisfy the previous assumptions. We then introduce some subgraphs. First let us denote, for some triangle ψ, by $D(\psi)$ the diameter of the circle circumscribed of ψ and by $\beta(\psi)$ the smallest angle of ψ.

Definition 1 *Given any $\beta_0 \in]0, \pi/3]$, we introduce the following particular subset of $Del_3(\varphi)$:*

$$Del_{3,\beta}^{\beta_0}(\varphi) = \{\psi \in Del_3(\varphi) : \beta(\psi) > \beta_0\}.$$

The β-Delaunay graph of order β_0 of any configuration φ is the Delaunay subgraph defined by :

$$Del_{2,\beta}^{\beta_0}(\varphi) = \bigcup_{\psi \in Del_{3,\beta}^{\beta_0}(\varphi)} \mathcal{P}_2(\psi).$$

The model obtained by replacing the original Delaunay graph by the β-Delaunay subgraph of order β_0 in (1) satisfies the previous sufficient conditions of Preston's Theorem. From now on, this model is called the β-Delaunay model. In this spirit, some other models may be defined (see *e.g.* Bertin et al. (1999b), Bertin et al. (1999a)) but we advice the reader to keep in mind the β-Delaunay model as the main example in order to illustrate the statistical results developped in this work.

The framework of this paper is restricted to stationary Gibbs point processes based on energy function related to some graph, denoted $\mathcal{G}_2(\varphi)$ for some finite configuration φ ($\mathcal{G}_k(\varphi)$ representing the set of cliques of order k), of the form

$$V(\varphi) = \sum_{k=1}^{K_{max}} \left\{ \sum_{\xi \in G_k(\varphi)} u^{(k)}(\xi;\varphi) \right\} \tag{2}$$

$$= \theta^{(1)}|\varphi| + \sum_{k=2}^{K_{max}} \left\{ \sum_{\xi \in G_k(\varphi)} u^{(k)}(\xi;\varphi) \right\}, \quad \text{when } u^{(1)} \equiv \theta^{(1)}$$

and satisfying Assumptions $\mathbf{E_1}$, $\mathbf{E_2^{loc}}$ or more generally $\mathbf{E_2^{qloc}}$, $\mathbf{E_3}$ defined by :

$\mathbf{E_1}$ $V(\cdot)$ is invariant by translation.

$\mathbf{E_2^{loc}}$ Locality of the local energy : there exists some fixed range denoted by D such that for any $\varphi \in \Omega$ one has

$$V(0|\varphi) = V(0|\varphi \cap \mathcal{B}(0,D)).$$

$\mathbf{E_2^{qloc}}$ Quasi-locality of the local energy : there exists a nonnegative function ε vanishing asymptotically such that for any $\varphi \in \Omega$ one has

$$|V(0|\varphi) - V(0|\varphi \cap \mathcal{B}(0,D))| < \varepsilon(D).$$

$\mathbf{E_3}$ Stability of the local energy : there exists $K \geq 0$ such that for any $\varphi \in \Omega$,

$$V(0|\varphi) \geq -K.$$

This framework includes some classical point processes such as :

- models based on the usual complete graph $\mathcal{G}_2(\varphi) = \mathcal{P}_2(\varphi)$) with pairwise interaction function satisfying a hard-core or inhibition condition and with finite range.
- k-nearest neighbours models with pairwise interaction function bounded and with finite range (see Bertin et al. (1999c)).
- Widom-Rowlinson or area interaction model.
- ...

3 Statistical model and inference method

3.1 Statistical model

We consider Gibbs point processes with energy function $V(\cdot;\boldsymbol{\theta})$ parametrized as follows

As a statistical model we consider a parametrized version of (2) where the different $u^{(k)}(\xi,\varphi)$ depend on a vector parameters $\boldsymbol{\theta}^{(k)}$ and then denoted from now by $u^{(k)}(\xi;\varphi,\boldsymbol{\theta}^{(k)})$. It is assumed that the vector of parameters $\boldsymbol{\theta}=(\theta_1,\ldots,\theta_{p+1})=\left(\theta^{(1)},\boldsymbol{\theta}^{(2)},\ldots,\boldsymbol{\theta}^{(K_{max})}\right)\in\boldsymbol{\Theta}$ where $\boldsymbol{\Theta}$ is an open bounded set of $\mathbb{R}^{p+1}$.

Our data consist in the realization of a point process with energy function $V(\cdot;\boldsymbol{\theta}^\star)$ in a domain $\Lambda\subset\mathbb{R}^d$ satisfying Assumptions $\mathbf{E_1}$ to $\mathbf{E_3}$. Thus, $\boldsymbol{\theta}^\star$ is the true parameter to be estimated. The Gibbs measure will be denoted by $P_{\boldsymbol{\theta}^\star}$. From (2), we obtain easily the energy to insert a point x in a configuration φ.

$$V(x|\varphi;\boldsymbol{\theta})=\sum_{k=1}^{K_{max}}\left\{\sum_{\xi\in G_k(\varphi\cup\{x\})\setminus G_k(\varphi)}u^{(k)}(\xi;\boldsymbol{\theta}^{(k)},\varphi\cup\{x\})-\sum_{\xi\in G_k(\varphi)\setminus G_k(\varphi\cup\{x\})}u^{(k)}(\xi;\boldsymbol{\theta}^{(k)},\varphi)\right\}\tag{3}$$

Theoretical results presented in the next sections are valid for a general energy function $V(\cdot;\boldsymbol{\theta})$. But among this class of models, we will focus on energy functions described by (2) and such that

$$u^{(k)}(\xi;\boldsymbol{\theta}^{(k)},\varphi)=\boldsymbol{\theta}^{(k)^T}\boldsymbol{u}^{(k)}(\xi;\varphi).$$

The energy can be rewritten

$$V(\varphi;\boldsymbol{\theta})=\sum_{k=1}^{K_{max}}\sum_{\xi\in G_k(\varphi)}\boldsymbol{\theta}^{(k)^T}\boldsymbol{u}^{(k)}(\xi;\varphi)=\sum_{k=1}^{K_{max}}\boldsymbol{\theta}^{(k)^T}\boldsymbol{u}^{(k)}(\varphi)\tag{4}$$

where for any finite configuration φ

$$\boldsymbol{u}^{(k)}(\varphi)=\sum_{\xi\in G_k(\varphi)}\boldsymbol{u}^{(k)}(\xi;\varphi)\quad\text{and}\quad\boldsymbol{u}(\varphi)=(u_1(\varphi),\ldots,u_{p+1}(\varphi))=\left(\boldsymbol{u}^{(1)}(\varphi),\ldots,\boldsymbol{u}^{(K_{max})}(\varphi)\right)$$

For two finite configurations φ and ψ, by denoting

$$\boldsymbol{u}^{(k)}(\psi|\varphi)=\boldsymbol{u}^{(k)}(\psi\cup\varphi)-\boldsymbol{u}^{(k)}(\varphi)\quad\text{and}\quad\boldsymbol{u}(\psi|\varphi)=\left(\boldsymbol{u}^{(1)}(\psi|\varphi),\ldots,\boldsymbol{u}^{(K_{max})}(\psi|\varphi)\right)\tag{5}$$

we have for any point x

$$V(\varphi;\boldsymbol{\theta})=\sum_{k=1}^{K_{max}}\boldsymbol{\theta}^{(k)^T}\boldsymbol{u}^{(k)}(\varphi)=\boldsymbol{\theta}^T\boldsymbol{u}(\varphi)\quad\text{and}\quad V(x|\varphi;\boldsymbol{\theta})=\sum_{k=1}^{K_{max}}\boldsymbol{\theta}^{(k)^T}\boldsymbol{u}^{(k)}(x|\varphi)=\boldsymbol{\theta}^T\boldsymbol{u}(x|\varphi),\tag{6}$$

where $\boldsymbol{u}(x|\varphi)=(u_1(x|\varphi,\ldots,u_{p+1}(x|\varphi)=(\boldsymbol{u}^{(1)}(x|\varphi),\ldots,\boldsymbol{u}^{K_{max}}(x|\varphi))$. The local specification of the Gibbs point process associated to an energy function defined by (6) belongs to an exponential family.

3.2 Pseudo-likelihood

As precised in the introduction, the idea of maximum pseudo-likelihood is due to Besag (1975) 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. With 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

$$PL_\Lambda(\varphi;\boldsymbol{\theta}) = \exp\left(-\int_\Lambda \exp\left(-V(x|\varphi;\boldsymbol{\theta})\right)dx\right)\prod_{x\in\varphi_\Lambda}\exp\left(-V(x|\varphi\setminus x;\boldsymbol{\theta})\right). \tag{7}$$

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

$$LPL_\Lambda(\varphi;\boldsymbol{\theta}) = -\int_\Lambda \exp\left(-V(x|\varphi;\boldsymbol{\theta})\right)dx - \sum_{x\in\varphi_\Lambda} V(x|\varphi\setminus x;\boldsymbol{\theta}) \tag{8}$$

3.3 Main statistical tools

Let us start by presenting a particular case of Campbell Theorem combined with Glötz Theorem that is widely used in our future proofs. For some finite configuration φ (resp. for some set G) and for all x, we denote by φ_x (resp. G_x) the configuration φ (resp. the set G) translated of x.

Corollary 1 *If the probability measure P is stationary and if the function $h(\cdot,\cdot)$ (used in Campbell Theorem) can be decomposed into $h(x,\varphi) = \mathbf{1}(x\in\Lambda)g(x,\varphi)$ for $\Lambda\subset\mathbb{R}^d$ where $g(\cdot,\cdot)$ is such that $g(x,\varphi_x) = g(0,\varphi)$ for all x, then the refined Campbell theorem combined with Glötz Theorem allow us to obtain*

$$\boldsymbol{E}_P\Big(\sum_{x\in\Phi_\Lambda\setminus x} g(x,\Phi\setminus x)\Big) = |\Lambda|\ \boldsymbol{E}_P\Big(\ g(0,\Phi)\exp(-V(0|\Phi))\ \Big) \tag{9}$$

Let us now present a version of an ergodic theorem obtained by Nguyen and Zessin (1979) and widely used in this paper. Let $\widetilde{D}>0$ and denote by Λ_0 the following fixed domain

$$\Lambda_0 = \left\{ z\in\mathbb{R}^2, -\frac{\widetilde{D}}{2} \le |z| \le \frac{\widetilde{D}}{2} \right\},$$

where for all $z\in\mathbb{R}^2$, $|z| = \max(z_1,z_2)$.

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

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

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*$-\sigma-$*algebra of* $\mathcal{F}$ *consisting of translation invariant (with probability 1) sets. Assume there exists a nonnegative and integrable random variable* Y *such that* $|H_G| \leq Y$ *a.s. for every convex* $G \subset \Lambda_0$*. Then,*

$$\lim_{n \to +\infty} \frac{1}{|G_n|} H_{G_n} = \frac{1}{|\Lambda_0|} E(H_{\Lambda_0} | \mathcal{I}), \quad a.s.$$

for each regular sequence $G_n \to I\!R^d$.

4 Consistency of the maximum pseudo-likelihood estimator

Maximizing the pseudo-likelihood is equivalent to minimize $U_n(\boldsymbol{\theta})$ defined by

$$U_n(\boldsymbol{\theta}) = -\frac{1}{|\Lambda_n|} LPL_{\Lambda_n}(\varphi; \boldsymbol{\theta}).$$

We denote by $\widehat{\boldsymbol{\theta}}_n = \widehat{\boldsymbol{\theta}}_n(\varphi)$ the maximum pseudo-likelihood estimator based on the configuration φ, alternatively defined as

$$\widehat{\boldsymbol{\theta}}_n(\varphi) = \mathrm{argmin}_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} U_n(\boldsymbol{\theta})$$

In this section, the existence of an ergodic measure is ensured, relatively to our framework, by Assumptions $\mathbf{E_1}$, $\mathbf{E_2^{qloc}}$ and $\mathbf{E_3}$. The following Assumptions are needed to derive the almost sure convergence of this estimator.

$\mathbf{C_1}$ $(\Lambda_n)_{n \geq 1}$ is a regular sequence of domains such that $\Lambda_n \to \mathbb{R}^2$ as $n \to +\infty$.

$\mathbf{C_2}$ For all $\boldsymbol{\theta} \in \boldsymbol{\Theta}$,

$$V(0|\cdot; \boldsymbol{\theta}) \in L^1(P_{\boldsymbol{\theta}^\star}).$$

$\mathbf{C_3}$ For all $\boldsymbol{\theta} \in \boldsymbol{\Theta} \setminus \boldsymbol{\theta}^\star$

$$P_{\boldsymbol{\theta}^\star}\Big(\{\varphi, \; V(0|\varphi; \boldsymbol{\theta}) \neq V(0|\varphi; \boldsymbol{\theta}^\star)\} \Big) > 0$$

$\mathbf{C_4}$ For all $\boldsymbol{\theta}, \boldsymbol{\theta}' \in \boldsymbol{\Theta}$, there exists $c > 0$ such that $P_{\boldsymbol{\theta}^\star}-$almost surely, we have

$$|V(0|\Phi; \boldsymbol{\theta}) - V\left(0|\Phi; \boldsymbol{\theta}'\right)| \leq ||\boldsymbol{\theta} - \boldsymbol{\theta}'||^c g(0, \Phi) \tag{10}$$

where $g(\cdot, \cdot)$ is a function such that for all x, $g(0, \Phi) = g(x, \Phi_x)$ and such that $g(0, \cdot) \in L^1(P_{\boldsymbol{\theta}^\star})$.

Remark 1 *If one only assumes the existence of an ergodic measure, in particular without Assumption* $\mathbf{E_3}$ *(taken into account to express Assumptions* $\mathbf{C_2}$ *and* $\mathbf{C_4}$*) then*

- *the condition* $\mathbf{C_2}$ *becomes : for all* $\boldsymbol{\theta} \in \boldsymbol{\Theta}$*, the variables* $V(0|\cdot;\boldsymbol{\theta})\exp\big(-V(0|\cdot;\boldsymbol{\theta}^\star)\big)$ *and* $\exp\big(-V(0|\cdot;\boldsymbol{\theta})\big)$ *are* $P_{\boldsymbol{\theta}^\star}$*-integrable.*
- *the function* $g(\cdot,\cdot)$ *occuring in Assumption* $\mathbf{C_4}$ *is now such that for all* $\boldsymbol{\theta} \in \boldsymbol{\Theta}$*,* $g(0,\cdot)\exp\big(-V(0|\cdot;\boldsymbol{\theta})\big) \in L^1(P_{\boldsymbol{\theta}^\star})$*.*

These Assumptions have been verified in Mase (1995) for the Ruelle class of pairwise interaction function with $\boldsymbol{\theta} = (\beta, z)$ *where* β *represents the inverse temperature and* z *the chemical potential.*

Proposition 3 *Assume* $P_{\boldsymbol{\theta}^\star}$ *stationary, then under Assumptions* $\mathbf{C_1}$ *to* $\mathbf{C_4}$*, we have* $P_{\boldsymbol{\theta}^\star}-$*almost surely, as* $n \to +\infty$

$$\widehat{\boldsymbol{\theta}}_n(\Phi) \to \boldsymbol{\theta}^\star \tag{11}$$

Due to the decomposition of stationary measures as a mixture of ergodic measures (see Preston (1976)), one only needs to prove Proposition 3 by assuming that $P_{\boldsymbol{\theta}^\star}$ is ergodic. Therefore, in Lemmas 4 to 6, $P_{\boldsymbol{\theta}^\star}$ 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). Define

$$K_n(\boldsymbol{\theta},\boldsymbol{\theta}^\star) = U_n(\boldsymbol{\theta}) - U_n(\boldsymbol{\theta}^\star)$$

Lemma 4 *For all* $\boldsymbol{\theta} \in \boldsymbol{\Theta}$*, under Assumptions* $\mathbf{C_1}$ *and* $\mathbf{C_2}$*, we have* $P_{\boldsymbol{\theta}^\star}-$*almost surely, as* $n \to +\infty$

$$U_n(\boldsymbol{\theta}) \to U(\boldsymbol{\theta}) = \boldsymbol{E}_{P_{\boldsymbol{\theta}^\star}}\Big(\exp\left(-V(0|\Phi;\boldsymbol{\theta})\right) \;+\; V(0|\Phi;\boldsymbol{\theta})\exp\big(-V(0|\Phi;\boldsymbol{\theta}^\star)\big)\Big) \tag{12}$$

Proof. Under Assumptions $\mathbf{C_1}$, $\mathbf{C_2}$ and $\mathbf{E_1}$, one can apply Theorem 2 (Nguyen and Zessin (1979)) to the process

$$H_{1,\Lambda_n} = \int_{\Lambda_n} \exp\left(-V(x|\varphi;\boldsymbol{\theta})\right) dx.$$

And from Corollary 1, we obtain $P_{\boldsymbol{\theta}^\star}-$almost surely as $n \to +\infty$

$$\frac{1}{|\Lambda_n|} H_{1,\Lambda_n} \to \boldsymbol{E}_{P_{\boldsymbol{\theta}^\star}} \exp\left(-V(0|\Phi;\boldsymbol{\theta})\right). \tag{13}$$

Now, define

$$H_{2,\Lambda_n} = \sum_{x \in \Phi_{\Lambda_n}} V(x|\Phi\setminus x;\boldsymbol{\theta})$$

Let $G \subset \Lambda_0$, we clearly have

$$|H_{2,G}| \le \sum_{x\in\Phi_G} |V(x|\Phi\setminus x;\boldsymbol{\theta})| \le \sum_{x\in\Phi_{\Lambda_0}} |V(x|\Phi\setminus x;\boldsymbol{\theta})|$$

Under Assumption $\mathbf{C_2}$ and from Corollary 1, we have

$$\boldsymbol{E}_{P_{\boldsymbol{\theta}^\star}}\left(\sum_{x\in\Phi_{\Lambda_0}} |V(x|\Phi\setminus x;\boldsymbol{\theta})|\right) = |\Lambda_0|\boldsymbol{E}_{P_{\boldsymbol{\theta}^\star}}\Big(|V(0|\Phi;\boldsymbol{\theta})|\exp\Big(-V(0|\Phi;\boldsymbol{\theta}^\star)\Big)\Big) < +\infty$$

This means that for all $G \subset \Lambda_0$, there exists a random variable $Y \in L^1(P_{\boldsymbol{\theta}^\star})$ such that $|H_{2,G}| \leq Y$. Thus, under Assumption $\mathbf{C_1}$ and from Theorem 2 (Nguyen and Zessin (1979)) and from Corollary 1, we have $P_{\boldsymbol{\theta}^\star}-$almost surely

$$\frac{1}{|\Lambda_n|} H_{1,\Lambda_n} \to \frac{1}{|\Lambda_0|} \boldsymbol{E}_{P_{\boldsymbol{\theta}^\star}} \Big(\sum_{x \in \Phi_{\Lambda_0}} V\left(x|\Phi \setminus x; \boldsymbol{\theta}\right) \Big) = \boldsymbol{E}_{P_{\boldsymbol{\theta}^\star}} \Big(V\left(0|\Phi; \boldsymbol{\theta}\right) \exp\Big(- V\left(0|\Phi; \boldsymbol{\theta}^\star\right) \Big) \Big). \tag{14}$$

We have the result by combining (13) and (14). ■

Lemma 5 *Under the conditions of Lemma 4, the function $U_n(\cdot)$ defines a constrast function, that is there exists a function $K(\cdot, \boldsymbol{\theta}^\star)$ such that $P_{\boldsymbol{\theta}^\star}-$almost surely the following holds for all $\boldsymbol{\theta} \in \boldsymbol{\Theta}$:*

$$K_n(\boldsymbol{\theta}, \boldsymbol{\theta}^\star) \to K(\boldsymbol{\theta}, \boldsymbol{\theta}^\star)$$

where $K(\cdot, \boldsymbol{\theta}^\star)$ is a positive which, under Assumption $\mathbf{C_3}$ is zero if and only if $\boldsymbol{\theta} = \boldsymbol{\theta}^\star$.

Proof. From Lemma 4, the function $K(\boldsymbol{\theta}, \boldsymbol{\theta}^\star) \geq 0$ can be written

$$K(\boldsymbol{\theta}, \boldsymbol{\theta}^\star) = \boldsymbol{E}_{P_{\boldsymbol{\theta}^\star}} \Big(\exp(-V\left(0|\Phi; \boldsymbol{\theta}^\star\right)) \Big(\exp(V\left(0|\Phi; \boldsymbol{\theta}\right) - V\left(0|\Phi; \boldsymbol{\theta}^\star\right)) - (1 + V\left(0|\Phi; \boldsymbol{\theta}\right) - V\left(0|\Phi; \boldsymbol{\theta}^\star\right)) \Big) \Big) \tag{15}$$

The result is obtained using Assumption $\mathbf{C_3}$ and by noting that the function $t \mapsto \exp(t) - (1+t)$ is positive and is zero if and only if $t = 0$. ■

Lemma 6 *Under Assumption $\mathbf{C_4}$, the functions $\boldsymbol{\theta} \mapsto U_n(\boldsymbol{\theta})$ and $\boldsymbol{\theta} \mapsto K(\boldsymbol{\theta}, \boldsymbol{\theta}^\star)$ are continuous in $\boldsymbol{\theta}$. Moreover, the modulus of continuity of $U_n(\boldsymbol{\theta})$ defined by*

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

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

$$P_{\boldsymbol{\theta}^\star} \left(\limsup_{n \to +\infty} \left(W_n \left(\frac{1}{k} \right) \geq \varepsilon_k \right) \right) = 0. \tag{16}$$

Proof. Under Assumption $\mathbf{C_4}$, it is sufficient to prove (16). Denote by

$$W_{1,n}\Big(\frac{1}{k}\Big) = \sup \Big\{ \Big| \frac{1}{|\Lambda_n|} \int_{\Lambda_n} \Big(\exp\Big(- V\left(x|\Phi; \boldsymbol{\theta}\right) \Big) - \exp\Big(- V\left(x|\Phi; \boldsymbol{\theta}'\right) \Big) \Big) dx \Big|, \boldsymbol{\theta}, \boldsymbol{\theta}' \in \boldsymbol{\Theta}, ||\boldsymbol{\theta} - \boldsymbol{\theta}'|| \leq \frac{1}{k} \Big\}$$

and

$$W_{2,n}\Big(\frac{1}{k}\Big) = \sup \left\{ \Big| \sum_{x \in \Phi_{\Lambda_n}} V\left(x|\Phi \setminus x; \boldsymbol{\theta}\right) - V\left(x|\Phi \setminus x; \boldsymbol{\theta}'\right) \Big|, \boldsymbol{\theta}, \boldsymbol{\theta}' \in \boldsymbol{\Theta}, ||\boldsymbol{\theta} - \boldsymbol{\theta}'|| \leq \frac{1}{k} \right\}$$

Under Assumptions $\mathbf{E_3}$ and $\mathbf{C_4}$, one can prove that $P_{\boldsymbol{\theta}^\star}-$almost surely

$$W_{1,n}\left(\frac{1}{k}\right) \leq \frac{\exp(K)}{k^c} \frac{1}{|\Lambda_n|} \int_{\Lambda_n} g(x, \Phi) dx$$

$$W_{2,n}\left(\frac{1}{k}\right) \leq \frac{1}{k^c} \frac{1}{|\Lambda_n|} \sum_{x \in \Phi_{\Lambda_n}} g(x, \Phi \setminus x)$$

Since $g(0,\cdot) \in L^1(P_{\boldsymbol{\theta}^\star})$, from Theorem 1 (Nguyen and Zessin (1979)) there exists $N_0 \in \mathbb{N}$ such that for all $n \geq N_0$ we have

$$W_{1,n}\left(\frac{1}{k}\right) \leq \frac{2\exp(K)}{k^c}\boldsymbol{E}_{P_{\boldsymbol{\theta}^\star}}(g(0,\Phi)) \quad \text{and} \quad W_{2,n}\left(\frac{1}{k}\right) \leq \frac{2\exp(K)}{k^c}\boldsymbol{E}_{P_{\boldsymbol{\theta}^\star}}(g(0,\Phi)).$$

And so for all $n \geq N_0$,

$$W_n\left(\frac{1}{k}\right) \leq \frac{\delta}{k^c} \qquad \text{with } \delta = 4\exp(K)\boldsymbol{E}_{P_{\boldsymbol{\theta}^\star}}(g(0,\Phi)).$$

Since

$$\limsup_{n\to+\infty}\left\{W_n\left(\frac{1}{k}\right) \geq \varepsilon_k\right\} \subset \left\{\frac{\delta}{k^c} \geq \varepsilon_k\right\}.$$

Thus, it is sufficient to choose $\varepsilon_k = \delta' k^{-c}$ with $\delta' > \delta$ to obtain the result. ■

Proof of Proposition 3

Lemma 5 and Lemma 6 ensure the fact that we can apply Property 3.6 of Guyon (1992) which asserts almost sure convergence for minimum contrast estimators. ■

The following proposition describes conditions $\mathbf{C_2}, \mathbf{C_3}$ and $\mathbf{C_4}$ in the case of an exponential family. New conditions are denoted $\mathbf{C_2^{exp}}$, $\mathbf{C_3^{exp}}$ and $\mathbf{C_4^{exp}}$. For this result, let us consider energy functions described by (6).

Proposition 7 *Conditions* $\mathbf{C_2}$ *and* $\mathbf{C_4}$ *(resp.* $\mathbf{C_3}$*) can be replaced by* $\mathbf{C_{2,4}^{exp}}$ *(resp.* $\mathbf{C_3^{exp}}$*) where*

$\mathbf{C_{2,4}^{exp}}$ *There exists* $\varepsilon > 0$ *such that for all* $i = 1, \ldots, p+1$

$$u_i(0|\varphi) \in L^{1+\varepsilon}(P_{\boldsymbol{\theta}^\star}).$$

$\mathbf{C_3^{exp}}$ *Identifiability condition : There exists* $A_1, \ldots, A_{p+1}$*,* $p+1$ *disjoint events of* Ω *such that* $P_{\boldsymbol{\theta}^\star}(A_i) > 0$ *and such that for all* $\varphi_1, \ldots, \varphi_{p+1} \in A_1 \times \cdots \times A_{p+1}$ *the* $(p+1)\times(p+1)$ *matrix with entries* $u_j(0|\varphi_i)$ *is constant and invertible.*

Proof.

- Denote by $||\cdot||_q$ the norm defined for $z \in \mathbb{R}^p$ by $||z||_q = \left(\sum_{i=1}^p |z_i|^q\right)^{1/q}$ with the obvious notation $||\cdot|| = ||\cdot||_2$. We have from Hölder's inequality

  $$|V(0|\Phi;\boldsymbol{\theta}) - V(0|\Phi;\boldsymbol{\theta}')| = \left|(\boldsymbol{\theta}-\boldsymbol{\theta}')\,\boldsymbol{u}(0|\Phi)\right| \leq ||\boldsymbol{\theta}-\boldsymbol{\theta}'||_{\frac{1+\varepsilon}{\varepsilon}}||\boldsymbol{u}(0|\Phi)||_{1+\varepsilon}.$$

  Since, $\boldsymbol{\Theta}$ is a bounded set there exists a constant $\kappa = \kappa(\varepsilon, \boldsymbol{\Theta})$ such that we have

  $$||\boldsymbol{\theta}-\boldsymbol{\theta}'||_{\frac{1+\varepsilon}{\varepsilon}} \leq ||\boldsymbol{\theta}-\boldsymbol{\theta}'||_2^{\frac{\varepsilon}{1+\varepsilon}} \; ||\boldsymbol{\theta}-\boldsymbol{\theta}'||_{\frac{2}{\varepsilon}}^{\frac{1}{1+\varepsilon}} \leq \kappa\, ||\boldsymbol{\theta}-\boldsymbol{\theta}'||_2^{\frac{\varepsilon}{1+\varepsilon}}.$$

  Thus, we have (10), with $c = \frac{\varepsilon}{1+\varepsilon}$ and $g(0,\cdot) = ||\boldsymbol{u}(0|\Phi)||_{1+\varepsilon}$. And so, $\mathbf{C_{2,4}^{exp}}$ implies $\mathbf{C_4}$ and obviously $\mathbf{C_2}$.

- Assumption $\mathbf{C_3^{exp}}$ means that for all $\boldsymbol{y} \in \mathrm{IR}^{p+1} \setminus \{\mathbf{0}\}$ and for all $\varphi_1, \ldots, \varphi_{p+1} \in A_1 \times \cdots \times A_{p+1}$, the matrix $(p+1) \times (p+1)$ with entries $(\underline{\boldsymbol{U}})_{i,j} = u_j\left(0|\varphi_i\right)$ (that does not depend on $\varphi_1, \ldots, \varphi_{p+1}$) is such that $\underline{\boldsymbol{U}}\boldsymbol{y} \neq \mathbf{0}$. So there exists $i_0(\boldsymbol{y}) \in \{1, \ldots, p+1\}$ such that $\boldsymbol{y}^T \underline{\boldsymbol{U}}_{i_0(\boldsymbol{y}),\cdot} = \boldsymbol{y}^T \boldsymbol{u}\left(0|\varphi_{i_0(\boldsymbol{y})}\right) \neq 0$. Therefore, for all $\boldsymbol{y} \in \mathrm{IR}^{p+1} \setminus \{\mathbf{0}\}$

$$P_{\boldsymbol{\theta}^\star}\left(\left\{\varphi,\ \ \boldsymbol{y}^T \boldsymbol{u}(0|\varphi) \neq 0\right\}\right) > P_{\boldsymbol{\theta}^\star}(A_{i_0(\boldsymbol{y})}) > 0,$$

which ends the proof.

■

5 Asymptotic normality for maximum pseudo-likelihood estimates

In this section, the existence of an ergodic measure is ensured, relatively to our framework, by Assumptions $\mathbf{E_1}$, $\mathbf{E_2^{loc}}$ and $\mathbf{E_3}$. The main tool used hereafter is a central limit theorem proposed by Jensen and Künsch (1994) which justifies the need of $\mathbf{E_2^{loc}}$ instead of $\mathbf{E_2^{qloc}}$.

To ensure the asymptotic normality for the maximum pseudo-likelihood estimator, the following assumptions are needed. Denote, for some real z, by $[z]$ the integer part of z.

$\mathbf{N_1}$ The point process is observed in a domain $\Lambda_n \oplus D = \cup_{x \in \Lambda_n} \mathcal{B}(x, D)$, where $\Lambda_n \subset \mathrm{IR}^2$ can de decomposed into $\cup_{i \in I_n} \Lambda_{(i)}$ where for $i = (i_1, i_2)$

$$\Lambda_{(i)} = \left\{ z \in \mathrm{IR}^2, \widetilde{D}\left(i_j - \frac{1}{2}\right) \leq z_j \leq \widetilde{D}\left(i_j - \frac{1}{2}\right), j = 1, 2 \right\}$$

for some $\widetilde{D} > 0$. As $n \to +\infty$, we also assume that $\Lambda_n \to \mathrm{IR}^2$ such that $|\Lambda_n| \to +\infty$ and $\dfrac{|\partial \Lambda_n|}{|\Lambda_n|} \to 0$

$\mathbf{N_2}$ $V\left(0|\cdot;\boldsymbol{\theta}\right)$ is twice times differentiable in $\boldsymbol{\theta} = \boldsymbol{\theta}^\star$ and for all $j, k = 1, \ldots, p+1$, there exists $\varepsilon > 0$ such that the variables

$$\frac{\partial V}{\partial \theta_j}\left(0|\cdot;\boldsymbol{\theta}^\star\right)^{3+\varepsilon} \quad \text{and} \quad \frac{\partial^2 V}{\partial \theta_j \partial \theta_k}\left(0|\cdot;\boldsymbol{\theta}^\star\right) \ \in L^1(P_{\boldsymbol{\theta}^\star})$$

$\mathbf{N_3}$ The matrix

$$\underline{\boldsymbol{\Sigma}}(\widetilde{D}, \boldsymbol{\theta}^\star) = \widetilde{D}^{-2} \sum_{|i| \leq \left[\frac{D}{\widetilde{D}}\right]+1} \boldsymbol{E}_{\boldsymbol{\theta}^\star}\left(\mathbf{LPL}^{(1)}_{\Lambda_0}\left(\Phi;\boldsymbol{\theta}^\star\right) \mathbf{LPL}^{(1)}_{\Lambda_i}\left(\Phi;\boldsymbol{\theta}^\star\right)^T\right) \tag{17}$$

is symmetric and definite positive. The vector $\mathbf{LPL}^{(1)}_{\Lambda_i}\left(\varphi;\boldsymbol{\theta}\right)$ is defined for any finite configuration φ and for all $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ and $j = 1, \ldots, p+1$ by

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

$\mathbf{N_4}$ $\forall \boldsymbol{y} \in \mathbb{R}^{p+1} \setminus \{\mathbf{0}\}$

$$P_{\boldsymbol{\theta}^\star}\left(\left\{\varphi,\ \boldsymbol{y}^T \boldsymbol{V}^{(1)}(0|\varphi;\boldsymbol{\theta}^\star) \neq 0\right\}\right) > 0,$$

where for $i = 1, \ldots, p+1$, $(\boldsymbol{V}^{(1)}(0|\varphi;\boldsymbol{\theta}^\star))_i = \frac{\partial V}{\partial \theta_i}(0|\varphi;\boldsymbol{\theta}^\star)$.

$\mathbf{N_5}$ There exists a neighborhood $\mathcal{V}$ of $\boldsymbol{\theta}^\star$ such that $V(\cdot;\boldsymbol{\theta})$ is twice times continuously differentiable for all $j,k = 1, \ldots, p+1$, we have

$$\left|\frac{\partial V}{\partial \theta_j}(0|\Phi;\boldsymbol{\theta}) - \frac{\partial V}{\partial \theta_j}(0|\Phi;\boldsymbol{\theta}^\star)\right| \leq ||\boldsymbol{\theta} - \boldsymbol{\theta}^\star||^{c_1}\ h_1(0,\Phi),$$

and

$$\left|\frac{\partial^2 V}{\partial \theta_j \partial \theta_k}(0|\Phi;\boldsymbol{\theta}) - \frac{\partial^2 V}{\partial \theta_j \partial \theta_k}(0|\Phi;\boldsymbol{\theta}^\star)\right| \leq ||\boldsymbol{\theta} - \boldsymbol{\theta}^\star||^{c_2}\ h_2(0,\Phi),$$

with $c_1, c_2 > 0$ and $h_1(\cdot,\cdot), h_2(\cdot,\cdot)$ two functions such that, for all x, $h_i(0,\Phi) = h_i(x,\Phi_x)$ and such that $h_1(0,\cdot)^2$ and $h_2(0,\cdot) \in L^1(P_{\boldsymbol{\theta}^\star})$.

Remark 2 *Assumption $\mathbf{N_1}$ is similar to the one of Jensen (1993), Jensen and Künsch (1994) and Heinrich (1992). Among other things, $\mathbf{N_1}$ ensures that Λ_n is a regular sequence of domains such that $\Lambda_n \to \mathbb{R}^2$.*

Remark 3 *Similarly to Remark 1,*

- *the integrability condition occuring in Assumption $\mathbf{N_2}$ becomes : $\left(\frac{\partial V}{\partial \theta_j}(0|\cdot;\boldsymbol{\theta}^\star)\exp\left(-V(0|\cdot;\boldsymbol{\theta}^\star)\right)\right)^{3+\varepsilon}$ and $\frac{\partial^2 V}{\partial \theta_j \partial \theta_k}(0|\cdot;\boldsymbol{\theta}^\star)\exp\left(-V(0|\cdot;\boldsymbol{\theta}^\star)\right)$ are $P_{\boldsymbol{\theta}^\star}$-integrable.*
- *the functions $h_1(\cdot,\cdot)$ and $h_2(\cdot,\cdot)$ occuring in Assumption $\mathbf{N_5}$ are now such that for all $\boldsymbol{\theta} \in \mathcal{V}$, the variables $h_1(0,\cdot)^2 \exp\left(-V(0|\cdot;\boldsymbol{\theta})\right)$ and $h_2(0,\cdot)\exp\left(-V(0|\cdot;\boldsymbol{\theta})\right)$ are $P_{\boldsymbol{\theta}^\star}$-integrable. Moreover, it is also assumed, for all $j,k = 1,\ldots,p+1$, the $P_{\boldsymbol{\theta}^\star}$-integrability of the variables $\left(\frac{\partial V}{\partial \theta_j}(0|\cdot;\boldsymbol{\theta}^\star)\right)^2 \exp\left(-V(0|\cdot;\boldsymbol{\theta})\right)$ and $\frac{\partial^2 V}{\partial \theta_j \partial \theta_k}(0|\cdot;\boldsymbol{\theta}^\star)\exp\left(-V(0|\cdot;\boldsymbol{\theta})\right)$.*

These Assumptions have been verified in Mase (1999) for the Ruelle class of pairwise interaction function with $\boldsymbol{\theta} = (\beta, z)$ where β represents the inverse temperature and z the chemical potential.

For $\boldsymbol{\theta}$ in a neighborhood of $\boldsymbol{\theta}^\star$, we can define, under Assumption $\mathbf{N_5}$, $\boldsymbol{U}_n^{(1)}(\boldsymbol{\theta})$ as the vector derivative of U_n. More precisely under Assumption $\mathbf{N_1}$, we can write

$$\boldsymbol{U}_n^{(1)}(\boldsymbol{\theta}) = |\Lambda_n|^{-1} \sum_{i \in I_n} \mathbf{LPL}_{\Lambda_i}^{(1)}(\varphi;\boldsymbol{\theta}). \tag{18}$$

For $\boldsymbol{\theta}$ in a neighborhood of $\boldsymbol{\theta}^\star$, we can also define, under Assumption $\mathbf{N_5}$, the Hessian matrix $\underline{\boldsymbol{U_n}}^{(2)}(\boldsymbol{\theta})$ given for $j,k = 1,\ldots,p+1$ by

$$\begin{aligned}\underline{\boldsymbol{U_n}}^{(2)}(\boldsymbol{\theta}) = & \frac{1}{|\Lambda_n|}\int_{\Lambda_n}\left(\frac{\partial V}{\partial \theta_j}(x|\varphi;\boldsymbol{\theta})\frac{\partial V}{\partial \theta_k}(x|\varphi;\boldsymbol{\theta}) - \frac{\partial^2 V}{\partial \theta_j \partial \theta_k}(x|\varphi;\boldsymbol{\theta})\right)\exp\left(-V(x|\varphi;\boldsymbol{\theta})\right)dx \\ & + \frac{1}{|\Lambda_n|}\sum_{x \in \varphi_{\Lambda_n}} \frac{\partial^2 V}{\partial \theta_j \partial \theta_k}(x|\varphi \setminus x;\boldsymbol{\theta}).\end{aligned} \tag{19}$$

Proposition 8 *Assume $P_{\boldsymbol{\theta}^\star}$ stationary, then under Assumptions $\mathbf{N_1}$ to $\mathbf{N_5}$, we have, for any $\widetilde{D}$ fixed, the following convergence in distribution as $n \to +\infty$*

$$|\Lambda_n|^{1/2}\, \underline{\widehat{\boldsymbol{\Sigma}}}_n(\widetilde{D},\widehat{\boldsymbol{\theta}}_n)^{-1/2}\, \underline{\boldsymbol{U_n}}^{(2)}(\widehat{\boldsymbol{\theta}}_n)\, \left(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}^\star\right) \to \mathcal{N}\left(0, \underline{\boldsymbol{I}}_{p+1}\right), \tag{20}$$

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

$$\underline{\widehat{\boldsymbol{\Sigma}}}_n(\widetilde{D},\boldsymbol{\theta}) = |\Lambda_n|^{-1}\widetilde{D}^{-2}\sum_{i\in I_n}\sum_{|j-i|\leq\left[\frac{D}{\widetilde{D}}\right]+1, j\in I_n} \mathbf{LPL}^{(1)}_{\Lambda_i}(\varphi;\boldsymbol{\theta})\,\mathbf{LPL}^{(1)}_{\Lambda_j}(\varphi;\boldsymbol{\theta})^T \tag{21}$$

By similar arguments of Jensen and Künsch (1994), due to the decomposition of stationary measures as a mixture of ergodic measures (see Preston (1976)), one only needs to prove Proposition 8 by assuming that $P_{\boldsymbol{\theta}^\star}$ is ergodic. Therefore, in Lemmas 9 to 11, $P_{\boldsymbol{\theta}^\star}$ is assumed to be ergodic.

The proof of this result is based on a general result obtained by Guyon (1992) (Proposition 3.7), giving conditions for which a miminum contrast estimator is asymptotically normal. The following Lemmas are needed to ensure these conditions. The first one ensures a central limit theorem for $\boldsymbol{U}_n^{(1)}(\boldsymbol{\theta}^\star)$.

Lemma 9 *Under Assumptions $\mathbf{N_1}$, $\mathbf{N_2}$ and $\mathbf{N_3}$,*
(a) we have, for any fixed $\widetilde{D}$, the following convergence in distribution as $n \to +\infty$

$$|\Lambda_n|^{1/2}\underline{\boldsymbol{\Sigma}}(\widetilde{D},\boldsymbol{\theta}^\star)^{-1/2}\, \boldsymbol{U_n}^{(1)}(\boldsymbol{\theta}^\star) \to \mathcal{N}\left(0, \underline{\boldsymbol{I}}_{p+1}\right) \tag{22}$$

where the matrix $\underline{\boldsymbol{\Sigma}}(\widetilde{D},\boldsymbol{\theta}^\star)$ is defined by (17).
(b) Moreover, we have $P_{\boldsymbol{\theta}^\star}-$almost surely as $n \to +\infty$

$$\underline{\widehat{\boldsymbol{\Sigma}}}_n(\widetilde{D},\boldsymbol{\theta}^\star) \to \underline{\boldsymbol{\Sigma}}(\widetilde{D},\boldsymbol{\theta}^\star). \tag{23}$$

Proof. (a) The idea is to apply to $\boldsymbol{U}_n^{(1)}(\boldsymbol{\theta}^\star)$ a central limit theorem obtained by Jensen and Künsch (1994), Theorem 2.1. The following conditions have to be fullfilled to apply this result.

- (i) For all $i \in I_n$ and for all $j = 1,\ldots,p+1$ $\boldsymbol{E}_{P_{\boldsymbol{\theta}^\star}}\left(\left|(\mathbf{LPL}^{(1)}_{\Lambda_i}(\Phi;\boldsymbol{\theta}^\star))_j\right|^3\right) < +\infty.$
- (ii) For all $i \in I_n$, $\boldsymbol{E}_{P_{\boldsymbol{\theta}^\star}}\left((\mathbf{LPL}^{(1)}_{\Lambda_i}(\Phi;\boldsymbol{\theta}^\star))_j|\Phi_{\Lambda^c_{(i)}}\right) = 0.$
- (iii) The set I_n is such that $\dfrac{|\partial I_n|}{|I_n|} \to 0$, as $n \to +\infty$.
- (iv) The matrix $\mathbb{V}\mathrm{ar}_{P_{\boldsymbol{\theta}^\star}}\left(|\Lambda_n|^{1/2}\boldsymbol{U}_n^{(1)}(\boldsymbol{\theta}^\star)\right)$ converges to the matrix $\underline{\boldsymbol{\Sigma}}(\widetilde{D},\boldsymbol{\theta})$, which is definite positive under Assumption $\mathbf{N_3}$.

<u>Condition (i)</u> : let us write

$$\boldsymbol{E}_{P_{\boldsymbol{\theta}^\star}}\left(\left|(\mathbf{LPL}^{(1)}_{\Lambda_i}(\Phi;\boldsymbol{\theta}^\star))_j\right|^3\right) \leq 2\times(T_1+T_2) \tag{24}$$

where the terms T_1 et T_2 are respectively defined by

$$T_1 = \boldsymbol{E}_{P_{\boldsymbol{\theta}^\star}}\left(\left| \int_{\Lambda_{(i)}} \frac{\partial V}{\partial \theta_j}(x|\Phi;\boldsymbol{\theta}^\star)\exp(-V(x|\Phi;\boldsymbol{\theta}^\star))dx \right|^3 \right)$$

$$T_2 = \boldsymbol{E}_{P_{\boldsymbol{\theta}^\star}}\left(\left| \sum_{x\in\Phi_{\Lambda_{(i)}}} \frac{\partial V}{\partial \theta_j}(x|\Phi\setminus x;\boldsymbol{\theta}^\star) \right|^3 \right)$$

Under Assumption $\mathbf{N_2}$, Hölder's inequality and the stationarity of $P_{\boldsymbol{\theta}^\star}$, we can prove that

$$\begin{aligned} T_1 &\leq \exp(3K)|\Lambda_0|^2 \times \boldsymbol{E}_{P_{\boldsymbol{\theta}^\star}}\left(\int_{\Lambda_{(i)}} \left| \frac{\partial V}{\partial \theta_j}(x|\Phi;\boldsymbol{\theta}^\star)\right|^3 dx \right) \\ &\leq \exp(3K)|\Lambda_0|^3 \times \boldsymbol{E}_{P_{\boldsymbol{\theta}^\star}}\left(\left| \frac{\partial V}{\partial \theta_j}(0|\Phi;\boldsymbol{\theta}^\star)\right|^3 \right) < +\infty. \end{aligned} \tag{25}$$

We have from Hölder's inequality

$$T_2 \leq \boldsymbol{E}_{P_{\boldsymbol{\theta}^\star}}\left(|\Phi_{\Lambda_{(i)}}|^2 \sum_{x\in\Phi_{\Lambda_{(i)}}} \left| \frac{\partial V}{\partial \theta_j}(x|\Phi\setminus x;\boldsymbol{\theta}^\star) \right|^3 \right).$$

And from Corollary 1, it follows

$$T_2 \leq \; |\Lambda_0| \; \boldsymbol{E}_{P_{\boldsymbol{\theta}^\star}}\left(|\Phi_{\Lambda_0}|^2 \left| \frac{\partial V}{\partial \theta_j}(0|\Phi;\boldsymbol{\theta}^\star)\right|^3 \exp\left(-V(0|\Phi;\boldsymbol{\theta}^\star)\right) \right).$$

Again from Hölder's inequality and under Assumption $\mathbf{E_3}$, we can prove that for all $\eta > 0$,

$$T_2 \leq \; |\Lambda_0|\exp(K) \; \boldsymbol{E}_{P_{\boldsymbol{\theta}^\star}}\left(|\Phi_{\Lambda_0}|^{2(1+\frac{1}{\eta})}\right)^{\frac{\eta}{1+\eta}} \; \boldsymbol{E}_{P_{\boldsymbol{\theta}^\star}}\left(\left| \frac{\partial V}{\partial \theta_j}(0|\Phi;\boldsymbol{\theta}^\star)\right|^{3(1+\eta)} \right)^{\frac{1}{1+\eta}}.$$

Under Assumption $\mathbf{E_3}$ it is well-known that, for all $z > 0$, $\boldsymbol{E}_{P_{\boldsymbol{\theta}^\star}}(|\Phi_{\Lambda_0}|^z) < +\infty$. Now, let $\varepsilon = 3\eta$, there exists $\kappa = \kappa(\varepsilon)$ such that

$$T_2 \;\leq\; \kappa\,|\Lambda_0|\; \boldsymbol{E}_{P_{\boldsymbol{\theta}^\star}}\left(\left| \frac{\partial V}{\partial \theta_j}(0|\Phi;\boldsymbol{\theta}^\star)\right|^{3+\varepsilon} \right)^{\frac{1}{1+\varepsilon/3}} < +\infty \tag{26}$$

under Assumption $\mathbf{N_2}$. Condition (i) is obtained by combining (24), (25) and (26)
Condition (ii) : From the stationarity of the process, it is sufficient to prove that

$$\boldsymbol{E}_{P_{\boldsymbol{\theta}^\star}}\left((\mathbf{LPL}^{(1)}_{\Lambda_0}(\Phi;\boldsymbol{\theta}^\star))_j | \Phi_{\Lambda_0^c} \right) = 0.$$

Let us write for any finite configuration φ

$$(\mathbf{LPL}^{(1)}_{\Lambda_0}(\varphi;\boldsymbol{\theta}^\star))_j = -\int_{\Lambda_0} \frac{\partial V}{\partial \theta_j}(x|\varphi;\boldsymbol{\theta}^\star)\exp(-V(x|\varphi;\boldsymbol{\theta}^\star))dx + \int_{\Lambda_0} \frac{\partial V}{\partial \theta_j}(x|\varphi\setminus x;\boldsymbol{\theta}^\star)\,\varphi(dx). \tag{27}$$

Denote respectively by $G_1(\varphi)$ and $G_2(\varphi)$ the first and the second right-hand term of (27) and by $E_i = \boldsymbol{E}_{P_{\boldsymbol{\theta}^\star}}\left(G_i(\Phi)|\Phi_{\Lambda_0^c} = \varphi_{\Lambda_0^c}\right)$. From the definition of Gibbs point processes,

$$E_2 = \frac{1}{Z_{\Lambda_0}(\varphi_{\Lambda_0^c})}\int_{\Omega_{\Lambda_0}} Q(d\varphi_{\Lambda_0})\int_{\mathbb{R}^2}\varphi_{\Lambda_0}(dx)\mathbf{1}_{\Lambda_0}(x)\frac{\partial V}{\partial\theta_j}(x|\varphi\setminus x;\boldsymbol{\theta}^\star)\exp\left(-V\left(\varphi_{\Lambda_0}|\varphi_{\Lambda_0^c};\boldsymbol{\theta}^\star\right)\right).$$

Denote by $\varphi' = (\varphi_{\Lambda_0}, \varphi'_{\Lambda_0^c})$. Since Q is a Poisson process we can write

$$\begin{aligned} E_2 &= \frac{1}{Z_{\Lambda_0}(\varphi_{\Lambda_0^c})}\int_{\Omega} Q(d\varphi')\int_{\mathbb{R}^2}\varphi'(dx)\mathbf{1}_{\Lambda_0}(x)\frac{\partial V}{\partial\theta_j}(x|\varphi\setminus x;\boldsymbol{\theta}^\star)\exp\left(-V\left(\varphi_{\Lambda_0}|\varphi_{\Lambda_0^c};\boldsymbol{\theta}^\star\right)\right)\\ &= \frac{1}{Z_{\Lambda_0}(\varphi_{\Lambda_0^c})}\int_{\Omega} Q(d\varphi')\int_{\mathbb{R}^2}\varphi'(dx)\mathbf{1}_{\Lambda_0}(x)\frac{\partial V}{\partial\theta_j}\left(x|\varphi'_{\Lambda_0}\cup\varphi_{\Lambda_0^c}\setminus x;\boldsymbol{\theta}^\star\right)\exp\left(-V\left(\varphi'_{\Lambda_0}|\varphi_{\Lambda_0^c};\boldsymbol{\theta}^\star\right)\right)\end{aligned}$$

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

$$E_2 = \frac{1}{Z_{\Lambda_0}(\varphi_{\Lambda_0^c})}\int_{\Lambda_0} dx\int_{\Omega} Q_x^!(d\varphi')\frac{\partial V}{\partial\theta_j}\left(x|\varphi'_{\Lambda_0}\cup\varphi_{\Lambda_0^c};\boldsymbol{\theta}^\star\right)\exp\left(-V\left(\varphi'_{\Lambda_0}\cup x|\varphi_{\Lambda_0^c};\boldsymbol{\theta}^\star\right)\right).$$

Since from Slivnyak-Mecke Theorem, $Q = Q_x^!$, one can obtain

$$\begin{aligned} E_2 &= \frac{1}{Z_{\Lambda_0}(\varphi_{\Lambda_0^c})}\int_{\Omega} Q(d\varphi')\int_{\Lambda_0} dx\,\frac{\partial V}{\partial\theta_j}\left(x|\varphi'_{\Lambda_0}\cup\varphi_{\Lambda_0^c};\boldsymbol{\theta}^\star\right)\exp\left(-V\left(\varphi'_{\Lambda_0}\cup x|\varphi_{\Lambda_0^c};\boldsymbol{\theta}^\star\right)\right)\\ &= \frac{1}{Z_{\Lambda_0}(\varphi_{\Lambda_0^c})}\int_{\Omega} Q(d\varphi_{\Lambda_0})\int_{\Lambda_0} dx\frac{\partial V}{\partial\theta_j}(x|\varphi;\boldsymbol{\theta}^\star)\exp\left(-V(x|\varphi;\boldsymbol{\theta}^\star)\right)\exp\left(-V\left(\varphi_{\Lambda_0}|\varphi_{\Lambda_0^c};\boldsymbol{\theta}^\star\right)\right)\\ &= -E_1\end{aligned}$$

<u>Condition (iii) :</u> this condition is equivalent to Assumption $\mathbf{N_1}$.
<u>Condition (iv) :</u> let us start by noting that the vector $\mathbf{LPL}_{\Lambda_i}^{(1)}(\varphi;\boldsymbol{\theta}^\star)$ depends only on $\varphi_{\Lambda_{(j)}}$ for j such that $|j-i|\leq\left[\frac{D}{\tilde{D}}\right]+1$. From (18), we can obtain

$$\begin{aligned}\mathbb{V}\mathrm{ar}_{\boldsymbol{\theta}^\star}\left(|\Lambda_n|^{1/2}\boldsymbol{U}_n^{(1)}(\boldsymbol{\theta}^\star)\right) &= |\Lambda_n|^{-1}\mathbb{V}\mathrm{ar}_{P_{\boldsymbol{\theta}^\star}}\left(\mathbf{LPL}_{\Lambda_i}^{(1)}(\Phi;\boldsymbol{\theta}^\star)\right)\\ &= |\Lambda_n|^{-1}\sum_{i,j\in I_n}\boldsymbol{E}_{P_{\boldsymbol{\theta}^\star}}\left(\mathbf{LPL}_{\Lambda_i}^{(1)}(\Phi;\boldsymbol{\theta}^\star)\,\mathbf{LPL}_{\Lambda_j}^{(1)}(\Phi;\boldsymbol{\theta}^\star)^T\right)\\ &= |\Lambda_n|^{-1}\sum_{i\in I_n}\left\{\sum_{|j-i|\leq\left[\frac{D}{\tilde{D}}\right]+1,j\in I_n}\boldsymbol{E}_{P_{\boldsymbol{\theta}^\star}}\left(\mathbf{LPL}_{\Lambda_i}^{(1)}(\Phi;\boldsymbol{\theta}^\star)\,\mathbf{LPL}_{\Lambda_j}^{(1)}(\Phi;\boldsymbol{\theta}^\star)^T\right)\right.\\ &\qquad\left.+\sum_{|j-i|>\left[\frac{D}{\tilde{D}}\right]+1,j\in I_n}\boldsymbol{E}_{P_{\boldsymbol{\theta}^\star}}\left(\mathbf{LPL}_{\Lambda_i}^{(1)}(\Phi;\boldsymbol{\theta}^\star)\,\mathbf{LPL}_{\Lambda_j}^{(1)}(\Phi;\boldsymbol{\theta}^\star)^T\right)\right\}.\end{aligned}$$

Let $j\in I_n$ such that $|j-i|>\left[\frac{D}{\tilde{D}}\right]+1$, then using condition (ii)

$$\begin{aligned}\boldsymbol{E}_{P_{\boldsymbol{\theta}^\star}}\left(\mathbf{LPL}_{\Lambda_i}^{(1)}(\Phi;\boldsymbol{\theta}^\star)\,\mathbf{LPL}_{\Lambda_j}^{(1)}(\Phi;\boldsymbol{\theta}^\star)^T\right) &= \boldsymbol{E}_{P_{\boldsymbol{\theta}^\star}}\left(\boldsymbol{E}\left(\mathbf{LPL}_{\Lambda_i}^{(1)}(\Phi;\boldsymbol{\theta}^\star)\,\mathbf{LPL}_{\Lambda_j}^{(1)}(\Phi;\boldsymbol{\theta}^\star)^T|\mathbf{LPL}_{\Lambda_j}^{(1)}(\Phi;\boldsymbol{\theta}^\star)\right)\right)\\ &= \boldsymbol{E}_{P_{\boldsymbol{\theta}^\star}}\left(\boldsymbol{E}\left(\mathbf{LPL}_{\Lambda_i}^{(1)}(\Phi;\boldsymbol{\theta}^\star)|\mathbf{LPL}_{\Lambda_j}^{(1)}(\Phi;\boldsymbol{\theta}^\star)\right)\mathbf{LPL}_{\Lambda_j}^{(1)}(\Phi;\boldsymbol{\theta}^\star)^T\right)\\ &= 0\end{aligned}$$

Now, denote by $\widetilde{I}$ the following set

$$\widetilde{I} = \left\{ k \in I_n, |k-i| \leq \left[\frac{D}{\widetilde{D}} \right] + 1, \forall i \in \partial I_n \right\}$$

and (for the sake of simplicity) by $E_{i,j}$ the following mean

$$E_{i,j} = \boldsymbol{E}_{P_{\boldsymbol{\theta}^\star}} \left(\mathbf{LPL}^{(1)}_{\Lambda_i} (\Phi; \boldsymbol{\theta}^\star) \, \mathbf{LPL}^{(1)}_{\Lambda_j} (\Phi; \boldsymbol{\theta}^\star)^T \right).$$

From the stationarity of the process, we can write

$$\begin{aligned}
\mathbb{V}\mathrm{ar}_{P_{\boldsymbol{\theta}^\star}} \left(|\Lambda_n|^{1/2} \boldsymbol{U}_n^{(1)}(\boldsymbol{\theta}^\star) \right) &= |\Lambda_n|^{-1} \left(\sum_{i \in I_n \setminus \widetilde{I}} \; \sum_{|j-i| \leq \left[\frac{D}{\widetilde{D}}\right] + 1, j \in I_n} E_{i,j} + \sum_{i \in \widetilde{I}} \; \sum_{|j-i| \leq \left[\frac{D}{\widetilde{D}}\right] + 1, j \in I_n} E_{i,j} \right). \\
&= |I_n \setminus \widetilde{I}| \, |\Lambda_n|^{-1} \sum_{|i| \leq \left[\frac{D}{\widetilde{D}}\right] + 1} E_{0,i} \; + \; |\widetilde{I}| \, |\Lambda_n|^{-1} \sum_{|j-i_0| \leq \left[\frac{D}{\widetilde{D}}\right] + 1} E_{i_0,j},
\end{aligned}$$

for some $i_0 \in \partial I_n$. From the definition of the set I_n, we have as $n \to +\infty$

$$\mathbb{V}\mathrm{ar}_{P_{\boldsymbol{\theta}^\star}} \left(|\Lambda_n|^{1/2} \boldsymbol{U}_n^{(1)}(\boldsymbol{\theta}^\star) \right) \to \sum_{|i| \leq \left[\frac{D}{\widetilde{D}}\right] + 1} E_{0,i} \; = \; \underline{\boldsymbol{\Sigma}} \left(\widetilde{D}, \boldsymbol{\theta}^\star \right).$$

(b) According to (21), it is easy to see that $\widehat{\underline{\boldsymbol{\Sigma}}}_n(\widetilde{D}, \boldsymbol{\theta}^\star)$ is defined such that as $n \to +\infty$,

$$\boldsymbol{E}_{P_{\boldsymbol{\theta}^\star}} \left(\widehat{\underline{\boldsymbol{\Sigma}}}_n(\widetilde{D}, \boldsymbol{\theta}^\star) \right) \to \underline{\boldsymbol{\Sigma}}(\widetilde{D}, \boldsymbol{\theta}^\star).$$

We leave the reader to check that under Assumption $\mathbf{N_1}$ and from Theorem 1 (Nguyen and Zessin (1979)), we have $P_{\boldsymbol{\theta}^\star}$−almost surely as $n \to +\infty$, $\widehat{\underline{\boldsymbol{\Sigma}}}_n(\widetilde{D}, \boldsymbol{\theta}^\star) \to \underline{\boldsymbol{\Sigma}}(\widetilde{D}, \boldsymbol{\theta}^\star)$. ■

Remark 4 *From the previous proof, we can note that Assumption* $\mathbf{N_3}$ *is fullfilled as soon as one can prove that for n sufficiently large the matrix* $\mathbb{V}ar_{P_{\boldsymbol{\theta}^\star}} \left(|\Lambda_n|^{1/2} \boldsymbol{U}_n^{(1)}(\boldsymbol{\theta}^\star) \right)$ *is definite positive.*

Lemma 10 *Under Assumptions* $\mathbf{N_1}$, $\mathbf{N_2}$ *and* $\mathbf{N_5}$, *there exists a neighborhood* $\mathcal{V}$ *of* $\boldsymbol{\theta}^\star$ *on which* $U_n(\cdot)$ *is twice times continuously differentiable and a random variable* Y *such that for all* $j, k = 1, \ldots, p+1$ *and for all* $\boldsymbol{\theta} \in \mathcal{V}$ *we have,*

$$\left| \left(\underline{\boldsymbol{U_n}}^{(2)}(\boldsymbol{\theta}) \right)_{j,k} \right| \leq Y.$$

Proof. Let $j, k = 1, \ldots, p+1$. Under Assmuption $\mathbf{N_5}$, there exists a neighborhood $\mathcal{V}$ of $\boldsymbol{\theta}^\star$ such that we can write for any configuration φ

$$\begin{aligned}
\left(\underline{\boldsymbol{U_n}}^{(2)}(\boldsymbol{\theta}) \right)_{j,k} &= -\frac{1}{|\Lambda_n|} \int_{\Lambda_n} \frac{\partial^2 V}{\partial \theta_j \partial \theta_k} (x|\varphi; \boldsymbol{\theta}) \exp\left(-V\left(x|\varphi; \boldsymbol{\theta}\right)\right) dx \\
&\quad + \frac{1}{|\Lambda_n|} \int_{\Lambda_n} \frac{\partial V}{\partial \theta_j} (x|\varphi; \boldsymbol{\theta}) \frac{\partial V}{\partial \theta_k} (x|\varphi; \boldsymbol{\theta}) \exp\left(-V\left(x|\varphi; \boldsymbol{\theta}\right)\right) dx \\
&\quad + \frac{1}{|\Lambda_n|} \sum_{x \in \varphi_{\Lambda_n}} \frac{\partial^2 V}{\partial \theta_j \partial \theta_k} (x|\varphi \setminus x; \boldsymbol{\theta}). \qquad (28)
\end{aligned}$$

Denote respectively by R_1, R_2, R_3 the three right-hand terms of the previous equation. Under Assumption $\mathbf{N_5}$, one can choose the neighborhood $\mathcal{V}$ such that $||\boldsymbol{\theta} - \boldsymbol{\theta}^\star|| \leq \kappa$. Thus, one can obtain

$$
\begin{aligned}
|R_1| &\leq \exp(K)\frac{1}{|\Lambda_n|}\int_{\Lambda_n}\left(\kappa^{c_2}h_2(x,\varphi) + \left|\frac{\partial^2 V}{\partial\theta_j\partial\theta_k}(x|\varphi;\boldsymbol{\theta}^\star)\right|\right)dx \\
|R_2| &\leq \exp(K)\frac{1}{|\Lambda_n|}\int_{\Lambda_n}\left(\kappa^{2c_1}h_1(x,\varphi)^2 + \kappa^{c_1}h_1(x,\varphi)\left|\frac{\partial V}{\partial\theta_j}(x|\varphi;\boldsymbol{\theta}^\star)\right|\right. \\
&\quad \left. + \kappa^{c_1}h_1(x,\varphi)\left|\frac{\partial V}{\partial\theta_k}(x|\varphi;\boldsymbol{\theta}^\star)\right| + \left|\frac{\partial V}{\partial\theta_j}(x|\varphi;\boldsymbol{\theta}^\star)\frac{\partial V}{\partial\theta_k}(x|\varphi;\boldsymbol{\theta}^\star)\right|\right) \\
|R_3| &\leq \frac{1}{|\Lambda_n|}\sum_{x\in\varphi_{\Lambda_n}}\left(\kappa^{c_2}h_2(x,\varphi\setminus x) + \left|\frac{\partial^2 V}{\partial\theta_j\partial\theta_k}(x|\varphi\setminus x;\boldsymbol{\theta}^\star)\right|\right)
\end{aligned}
$$

Under Assumptions $\mathbf{N_1}$ and $\mathbf{N_2}$, from Theorem 2 (Nguyen and Zessin (1979)), and using the stationarity of $P_{\boldsymbol{\theta}^\star}$, there exists $N_0 \in \mathbb{N}$ such that for all $n \geq N_0$, we have $P_{\boldsymbol{\theta}^\star}-$almost surely

$$
\begin{aligned}
|R_1| &\leq 2\times\exp(K)\boldsymbol{E}_{P_{\boldsymbol{\theta}^\star}}\left(\kappa^{c_2}h_2(0,\Phi) + \left|\frac{\partial^2 V}{\partial\theta_j\partial\theta_k}(0|\Phi;\boldsymbol{\theta}^\star)\right|\right) \\
|R_2| &\leq 2\times\exp(K)\left\{\boldsymbol{E}_{P_{\boldsymbol{\theta}^\star}}\left(k^{2c_1}h_1(0,\Phi)^2 + \left|\frac{\partial V}{\partial\theta_j}(0|\Phi;\boldsymbol{\theta}^\star)\frac{\partial V}{\partial\theta_k}(0|\Phi;\boldsymbol{\theta}^\star)\right|\right)\right. \\
&\quad \left. +\boldsymbol{E}_{P_{\boldsymbol{\theta}^\star}}\left(\kappa^{c_1}h_1(0,\Phi)\left|\frac{\partial V}{\partial\theta_j}(0|\Phi;\boldsymbol{\theta}^\star)\right| + \kappa^{c_1}h_1(0,\Phi)\left|\frac{\partial V}{\partial\theta_k}(0|\Phi;\boldsymbol{\theta}^\star)\right|\right)\right\} \\
|R_3| &\leq 2\times\exp(K)\boldsymbol{E}_{P_{\boldsymbol{\theta}^\star}}\left(\kappa^{c_2}h_2(0,\Phi) + \left|\frac{\partial^2 V}{\partial\theta_j\partial\theta_k}(0|\Phi;\boldsymbol{\theta}^\star)\right|\right)
\end{aligned}
$$

Consequenlty, for n large enough, there exists a positive constant κ' such that $\left|\left(\underline{\boldsymbol{U_n}}^{(2)}(\boldsymbol{\theta})\right)_{j,k}\right| \leq \kappa'$, which implies the result. ■

Lemma 11 *Under Assumptions* $\mathbf{N_1}$ *and* $\mathbf{N_2}$*, we have almost surely, as* $n \to +\infty$

$$\underline{\boldsymbol{U_n}}^{(2)}(\boldsymbol{\theta}^\star) \to \underline{\boldsymbol{U}}^{(2)}(\boldsymbol{\theta}^\star)$$

where $\underline{\boldsymbol{U}}^{(2)}(\boldsymbol{\theta}^\star)$ *is the* $(p+1)\times(p+1)$ *matrix whose entry is*

$$\left(\underline{\boldsymbol{U}}^{(2)}(\boldsymbol{\theta}^\star)\right)_{j,k} = \boldsymbol{E}_{P_{\boldsymbol{\theta}^\star}}\left(\frac{\partial V}{\partial\theta_j}(0|\Phi;\boldsymbol{\theta}^\star)\frac{\partial V}{\partial\theta_k}(0|\Phi;\boldsymbol{\theta}^\star)\exp\left(-V\left(0|\Phi;\boldsymbol{\theta}^\star\right)\right)\right). \tag{29}$$

Furthermore, under Assumption $\mathbf{N_4}$*,* $\underline{\boldsymbol{U}}^{(2)}$ *is a symmetric definite positive matrix.*

Proof. Let $j,k = 1,\ldots,p+1$. Under Assumptions $\mathbf{N_1}$ and $\mathbf{N_2}$ and from Theorem 2 (Nguyen and Zessin (1979)), we have almost surely, as $n \to +\infty$

$$
\begin{aligned}
\left(\underline{\boldsymbol{U_n}}^{(2)}(\boldsymbol{\theta}^\star)\right)_{j,k} \to &-\frac{1}{|\Lambda_0|}\boldsymbol{E}_{P_{\boldsymbol{\theta}^\star}}\left(\int_{\Lambda_0}\frac{\partial^2 V}{\partial\theta_j\partial\theta_k}(x|\Phi;\boldsymbol{\theta}^\star)\exp\left(-V\left(x|\Phi;\boldsymbol{\theta}^\star\right)\right)dx\right) \\
&+\frac{1}{|\Lambda_0|}\boldsymbol{E}_{P_{\boldsymbol{\theta}^\star}}\left(\int_{\Lambda_0}\frac{\partial V}{\partial\theta_j}(x|\Phi;\boldsymbol{\theta}^\star)\frac{\partial V}{\partial\theta_k}(x|\Phi;\boldsymbol{\theta}^\star)\exp\left(-V\left(x|\Phi;\boldsymbol{\theta}^\star\right)\right)dx\right) \\
&+\frac{1}{|\Lambda_0|}\boldsymbol{E}_{P_{\boldsymbol{\theta}^\star}}\left(\sum_{x\in\varphi_{\Lambda_0}}\frac{\partial^2 V}{\partial\theta_j\partial\theta_k}(x|\Phi\setminus x;\boldsymbol{\theta}^\star)\right)
\end{aligned} \tag{30}
$$

Equation (29) is obtained using Corollary 1. And under Assumption $\mathbf{N_4}$, it is easy to see that $\underline{\boldsymbol{U}}^{(2)}$ is a symmetric definite positive matrix. ■

Proof of Proposition 8 Using Lemmas 9 à 11, one can apply a classical result concerning asymptotic normality for minimum contrast estimators, *e.g.* Proposition 3.7 de Guyon (1992), in order to prove as $n \to +\infty$

$$|\Lambda_n|^{1/2}\underline{\widehat{\boldsymbol{\Sigma}}}_n(\widetilde{D},\boldsymbol{\theta}^\star)^{-1/2}\,\underline{\boldsymbol{U}}_n^{(2)}(\boldsymbol{\theta}^\star)\,\left(\widehat{\boldsymbol{\theta}}_n-\boldsymbol{\theta}^\star\right)\to\mathcal{N}\left(0,\underline{\boldsymbol{I}}_{p+1}\right).$$

The result is then obtained using the fact that $\widehat{\boldsymbol{\theta}}_n$ is a consistent estimator of $\boldsymbol{\theta}^\star$. ■

Let us precise, as in Section 4, the different Assumptions for energy functions that can be written as (4).

Proposition 12 *For energy functions described by (4), Assumptions* $\mathbf{N_2}$ *and* $\mathbf{N_5}$ *(resp.* $\mathbf{N_4}$*) can be replaced by* $\mathbf{N_{2,5}^{exp}}$ *(resp.* $\mathbf{N_4^{exp}}$*)*

$\mathbf{N_{2,5}^{exp}}$ *For* $i=1,\ldots,p+1,$ *there exists* $\varepsilon>0$ *such that* $u_i(0|\cdot)\in L^{3+\varepsilon}(P_{\boldsymbol{\theta}^\star})$.

$\mathbf{N_4^{exp}}=\mathbf{C_3^{exp}}$

The proof is trivial.

6 Some examples

In this section, it is assumed that the sequence of domains satisfies $\mathbf{N_1}$ (which implies $\mathbf{C_1}$). Moreover, we only focus on examples satisfying the following convenient Assumption denoted by $\mathbf{M}$:

M There exists $K_1, K_2>0$ such that for any finite configuration φ, we have for all x

$$-K_1\le u_i(x|\varphi)\le K_2,\qquad \text{for } i=1,\ldots,p+1.$$

Quite obviously, Assumption $\mathbf{M}$ ensures $\mathbf{C_{2,4}^{exp}}$ and $\mathbf{N_{2,5}^{exp}}$. Let us now present a Corollary of Propositions 3 and 8.

Corollary 13 *Under Assumption* $\mathbf{M}$ *and* $\mathbf{C_3^{exp}}$*, the consistency of the maximum pseudo-likelihood, that is the result (11), is valid. And in addition with* $\mathbf{N_3^{exp}}$*, its asymptotic normality property, that is the result (20), is ensured.*

6.1 Pairwise β-Delaunay model

We first deal with our main example : β-Delaunay (of order some small enough fixed β_0) model with multi-Strauss pairwise interaction function. In other words,

$$V(\varphi;\boldsymbol{\theta}) = \theta^{(1)}|\varphi| + \sum_{\xi \in Del_{2,\beta}^{\beta_0}(\varphi)} u^{(2)}(\xi;\varphi,\boldsymbol{\theta}^{(2)}) = \boldsymbol{\theta}^T \boldsymbol{u}(\varphi) \qquad (31)$$

with $u_1(\varphi) = |\varphi|$ and for any $i \in \{2,\ldots,p+1\}$,

$$u_i(\varphi) = \sum_{\xi \in Del_{2,\beta}^{\beta_0}(\varphi)} \mathbf{1}_{]d_{i-1},d_i]}(\|\xi\|)$$

where $0 = d_1 \leq d_2 \leq \ldots \leq d_{p+1}$ are some fixed real numbers. Literally, $u_i(\varphi)$ $(i > 1)$ corresponds to the number of (β-Delaunay) edges of length between d_{i-1} and d_i. We may also notice that the range of the pairwise interaction function is d_{p+1}, that is $u^{(2)}(\xi;\varphi,\boldsymbol{\theta}^{(2)}) = 0$ when $\|\xi\| > d_{p+1}$.

In Bertin et al. (1999a), it is proved that this model satisfies Assumption **M**. Let us now verify the technical conditions $\mathbf{C_3^{exp}}$ and $\mathbf{N_3^{exp}}$.

Proposition 14 *Assumption* $\mathbf{C_3^{exp}}$ *is satisfied for the* β*-Delaunay model with multi-Strauss pairwise interaction function.*

Proof. Denote by Δ the following domain

$$\Delta = \left\{ z \in \mathrm{IR}^2 : -D \leq z_i \leq D, i = 1,2 \right\}$$

and by A_1 the event $A_1 = \{\varphi, \varphi_\Delta = \emptyset\}$. We clearly have for all $\varphi_1 \in A_1$, $\boldsymbol{u}(0|\varphi_1) = (1,0,\ldots,0)^T$. Now, let us give for $j = 2,\ldots,p+1$, the points $c_{1,j}$ and $c_{2,j}$ such that the distances $d(0,c_{1,j}) = d(0,c_{2,j}) = d(c_{1,j},c_{2,j}) = \frac{d_{j-1}+d_j}{2}$. Denote for $j = 2,\ldots,p+1$ the following events for some $\eta > 0$

$$A_j(\eta) = \Big\{ \varphi \in \Omega : \varphi_\Delta = \{z_1,z_2\}, z_1 \in \mathcal{B}(c_{1,j},\eta), z_{2,j} \in \mathcal{B}(c_{2,j},\eta) \Big\}.$$

One can choose η such that for all $\varphi \in A_j(\eta)$, the distances $d(0,z_1), d(0,z_2$ and $d(z_1,z_2)$ are comprised between d_{j-1} and d_j. One can also choose η such that the smallest angle of the triangle with vertices $\{0,c_{1,j},c_{2,j}\}$ is strictly greater thant β_0, which means that $\{0,c_{1,j},c_{2,j}\} \in Del_3^{\beta_0}(\varphi_j)$. Now, it is easy to see that the matrix $\underline{\boldsymbol{U}}$ defined in Proposition 7 is given by

$$\underline{\boldsymbol{U}} = (u_j(0|\varphi_i))_{1 \leq i,j \leq p+1} = \begin{pmatrix} 1 & 0 & \cdots & \cdots & 0 \\ 1 & 3 & \ddots & \vdots & \vdots \\ 1 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & 3 & 0 \\ 1 & 0 & \ldots & 0 & 3 \end{pmatrix}$$

and is clearly invertible, which ends the proof. ■

Proposition 15 *Assumption* $\mathbf{N_3}$ *is satisfied for the* β*-Delaunay model with multi-Strauss pairwise interaction function.*

Proof. From Remark 4, it is sufficient to prove that the matrix $\mathbb{V}\mathrm{ar}_{P_{\boldsymbol{\theta}^\star}}(|\Lambda_n|^{1/2}\boldsymbol{U}_n^{(1)}(\Phi;\boldsymbol{\theta}^\star))$ is definite positive for n sufficiently large. Let $\widetilde{D} > D$, $\boldsymbol{y} \in \mathbb{R}^{p+1}$ and let $\widetilde{\Lambda} = \cup_{|i|\leq 1}\Lambda_{(i)}$, by the same argument of Jensen and Künsch (1994) (Equation (3.2)), we can write

$$\boldsymbol{y}^T\mathbb{V}\mathrm{ar}_{P_{\boldsymbol{\theta}^\star}}\left(|\Lambda_n|^{1/2}\boldsymbol{U}_n^{(1)}\right)\boldsymbol{y} \geq |\Lambda_n|\,\boldsymbol{E}_{P_{\boldsymbol{\theta}^\star}}\left(\mathbb{V}\mathrm{ar}_{P_{\boldsymbol{\theta}^\star}}\left(\boldsymbol{y}^T\boldsymbol{U}_n^{(1)}|\Phi_{\Lambda_\ell}, \ell \notin 3\mathbb{Z}^2\right)\right).$$

Now, following the proof of Lemma 9 ((a) condition (iv)), one can prove that there exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$,

$$\boldsymbol{y}^T\mathbb{V}\mathrm{ar}_{P_{\boldsymbol{\theta}^\star}}\left(|\Lambda_n|^{1/2}\right)\boldsymbol{y} \geq \frac{1}{2}\boldsymbol{E}_{P_{\boldsymbol{\theta}^\star}}\left(\mathbb{V}\mathrm{ar}_{P_{\boldsymbol{\theta}^\star}}\left(\boldsymbol{y}^T\mathbf{LPL}_{\widetilde{\Lambda}}^{(1)}(\Phi;\boldsymbol{\theta}^\star)\,|\Phi_{\Lambda_\ell}, 1 \leq |\ell| \leq 2\right)\right).$$

The aim is to prove that the function $h(\Phi) = \boldsymbol{y}^T\mathbf{LPL}_{\widetilde{\Lambda}}^{(1)}(\Phi;\boldsymbol{\theta}^\star)$ is not almost surely a constant, when the variables $\Phi_{\Lambda_\ell}, 1 \leq |\ell| \leq 2$ are (for example) fixed to $\emptyset$. Assume that the function $h(\cdot)$ explicitly given for any finite configuration φ by

$$h(\varphi) = \sum_{k=1}^{p+1} y_k \left\{ \int_{\widetilde{\Lambda}} u_k(x|\varphi_{\Lambda_0})\exp\left(-\boldsymbol{\theta}^{\star T}\boldsymbol{u}(x|\varphi_{\Lambda_0})\right)dx \;\;-\;\; \sum_{x\in\varphi_{\Lambda_0}} u_k(x|\varphi_{\Lambda_0}\setminus x)\right\}$$

is constant for all $\varphi \in \Omega' = \{\varphi \in \Omega : \varphi_{\Lambda_\ell} = \emptyset, 1 \leq |\ell| \leq 2\}$.

Denote by $A_0 = \left\{\varphi \in \Omega' : \varphi_{\Lambda_0} = \emptyset\right\}$ and by $A_1 = \left\{\varphi \in \Omega' : |\varphi_{\Lambda_0}| = 1\right\}$. It is clear that, $P_{\boldsymbol{\theta}^\star}(A_0) > 0$ and $P_{\boldsymbol{\theta}^\star}(A_1) > 0$. We have for all $\varphi_0 \in A_0$ and for all $\varphi_1 \in A_1$

$$h(\varphi_0) = y_1|\widetilde{\Lambda}|\exp(-\theta_1^\star) \quad \text{and} \quad h(\varphi_1) = y_1|\widetilde{\Lambda}|\exp(-\theta_1^\star) - y_1.$$

Assuming $h(\cdot)$ constant implies that $y_1 = 0$ and then $h(\cdot)$ vanishes. We now consider particular configurations of two points in Λ_0 and empty in $\Lambda_\ell \setminus \Lambda_0, 1 \leq |\ell| \leq 2$. Let us first introduce the following sets for any $j \in \{1, \cdots, p-1\}$ and any $\eta > 0$

$$D_j(\eta) = \left\{(z_1, z_2) \in \Lambda_0^2 : z_1 \in \mathcal{B}\left((0,0), \frac{\eta}{4}\right) \text{ and } z_2 \in \mathcal{B}\left((d_j, 0), \frac{3\eta}{4}\right)\right\}$$

$$D_j^-(\eta) = \left\{(z_1, z_2) \in \Lambda_0^2 : z_1 \in \mathcal{B}\left((0,0), \frac{\eta}{4}\right) \text{ and } z_2 \in \mathcal{B}\left((d_j - \frac{\eta}{2}, 0), \frac{\eta}{4}\right)\right\} \subset D_j(\eta)$$

$$D_j^+(\eta) = \left\{(z_1, z_2) \in \Lambda_0^2 : z_1 \in \mathcal{B}\left((0,0), \frac{\eta}{4}\right) \text{ and } z_2 \in \mathcal{B}\left((d_j + \frac{\eta}{2}, 0), \frac{\eta}{4}\right)\right\} \subset D_j(\eta)$$

When η is small enough, the couple of points $(z_1, z_2) \in D_j(\eta)$ (resp. $D_j^-(\eta)$ and $D_j^+(\eta)$) are such that $d_{j-1} < d_j - \eta < d(z_1, z_2) < d_j + \eta < d_{j+1}$ (resp. $d_{j-1} < d_j - \eta < d(z_1, z_2) < d_j$ and $d_j < d(z_1, z_2) < d_j + \eta < d_{j+1}$).

We now derive the corresponding events for any $j \in \{1, \cdots, p-1\}$ and any $\eta > 0$

$$A_j(\eta) = \left\{\varphi \in \Omega' : \varphi_{\Lambda_0} = \{z_1, z_2\} \text{ with } (z_1, z_2) \in D_j(\eta)\right\}$$

$$A_j^-(\eta) = \left\{\varphi \in \Omega' : \varphi_{\Lambda_0} = \{z_1, z_2\} \text{ with } (z_1, z_2) \in D_j^-(\eta)\right\} \subset A_j(\eta)$$

$$A_j^+(\eta) = \left\{\varphi \in \Omega' : \varphi_{\Lambda_0} = \{z_1, z_2\} \text{ with } (z_1, z_2) \in D_j^+(\eta)\right\} \subset A_j(\eta)$$

satisfying $P_{\boldsymbol{\theta}^\star}(A_j(\eta)) > 0$, $P_{\boldsymbol{\theta}^\star}(A_j^-(\eta)) > 0$ and $P_{\boldsymbol{\theta}^\star}(A_j^+(\eta)) > 0$.

Let us fix some $\varphi \in A_j(\eta)$. There exists some unique couple of points $(z_1, z_2) \in D_j(\eta)$, for which we define the following domain

$$\widetilde{\Lambda}(z_1, z_2) = \left\{ x \in \widetilde{\Lambda} : \{x, z_1, z_2\} \in Del_3^{\beta_0}(\varphi \cup \{x\}) \right\}.$$

Since $\{z_1, z_2\} \notin Del_2^{\beta_0}(\varphi)$, we then derive that

$$0 = h(\varphi) = \sum_{k=2}^{p+1} y_k \int_{\widetilde{\Lambda}(z_1,z_2)} u_k(x|\{z_1, z_2\}) \exp\left(-\boldsymbol{\theta}^{\star T} \boldsymbol{u}(x|\{z_1, z_2\})\right) dx. \tag{32}$$

When $\{x, z_1, z_2\} \in Del_3^{\beta_0}(\varphi \cup \{x\})$, we decompose $u_k(x|\{z_1, z_2\})$ into two additive terms in order to isolate the contribution of x :

$$u_k(x|\{z_1, z_2\}) = u_k^{-x}(z_1, z_2) + u_k^x(z_1, z_2)$$

with $u_1^{-x}(z_1, z_2) = 0$ and $u_1^x(z_1, z_2) = 1$, and for $k \neq 1$, $u_k^{-x}(z_1, z_2) = \mathbf{1}_{[d_{k-1}, d_k[}(\|z_1 - z_2\|)$ and $u_k^x(z_1, z_2) = \sum_{j=1,2} \mathbf{1}_{[d_{k-1}, d_k[}(\|x - z_j\|)$.

Then Equation (32) becomes for any $\boldsymbol{y} \in \mathrm{I\!R}^{p+1}$ with $y_1 = 0$

$$\begin{aligned} h(\varphi) = 0 = \exp\left(-\boldsymbol{\theta}^{\star T} \boldsymbol{u}^{-x}(z_1, z_2)\right) \left(\boldsymbol{y}^T \boldsymbol{u}^{-x}(z_1, z_2) f_1(z_1, z_2) + \boldsymbol{y}^T \boldsymbol{f}(z_1, z_2)\right) \\ \Longleftrightarrow \boldsymbol{y}^T \boldsymbol{u}^{-x}(z_1, z_2) f_1(z_1, z_2) + \boldsymbol{y}^T \boldsymbol{f}(z_1, z_2) = 0 \end{aligned} \tag{33}$$

where $\boldsymbol{u}^{-x} = (u_1^{-x}, \cdots, u_{p+1}^{-x})$, $\boldsymbol{u}^x = (u_1^x, \cdots, u_{p+1}^x)$ and $\boldsymbol{f} = (f_1, f_2, \ldots, f_{p+1})$ with

$$f_k(z_1, z_2) = \int_{\widetilde{\Lambda}(z_1,z_2)} u_k^x(z_1, z_2) \exp\left(-\boldsymbol{\theta}^{\star T} \boldsymbol{u}^x(z_1, z_2)\right) dx$$

Since each f_k is continuous, one could assert that for any $\varepsilon > 0$, there exists $\eta > 0$ such that for any $(z_1, z_2) \in D_j(\eta)$, $|f_k(z_1, z_2) - \widetilde{f}_k| < \varepsilon$ where $\widetilde{f}_k = f_k((0,0), (d_j, 0))$ is positive. We then set $\delta_k(z_1, z_2) = f_k(z_1, z_2) - \widetilde{f}_k$ and $\boldsymbol{\delta} = (\delta_1, \cdots, \delta_{p+1})$. We now apply the equation (33) for some fixed $\varphi_j^- \in A_j^-(\eta)$ and $\varphi_j^+ \in A_j^+(\eta)$. By denoting $(z_1^-, z_2^-) \in D_j^-(\eta)$ and $(z_1^+, z_2^+) \in D_j^+(\eta)$ such that $\varphi_j^- \cap \Lambda_0 = \{z_1^-, z_2^-\}$ and $\varphi_j^+ \cap \Lambda_0 = \{z_1^+, z_2^+\}$, we have

$$\begin{aligned} y_j f_1(z_1^-, z_2^-) + \boldsymbol{y}^T \boldsymbol{f}(z_1^-, z_2^-) = 0 \\ y_{j+1} f_1(z_1^+, z_2^+) + \boldsymbol{y}^T \boldsymbol{f}(z_1^+, z_2^+) = 0 \end{aligned}$$

By substracting these two terms, we can obtain

$$(y_{j+1} - y_j)\widetilde{f}_1 = y_j \delta_1(z_1^-, z_2^-) - y_{j+1}\delta_1(z_1^+, z_2^+) + \boldsymbol{y}^T \left(\boldsymbol{\delta}(z_1^-, z_2^-) - \boldsymbol{\delta}(z_1^+, z_2^+)\right) \tag{34}$$

By the previous continuity argument on the f_k, on can choose $\eta > 0$ (depending on $\boldsymbol{y}$) small enough such that the absolute value of the right-hand term of (34) could be lower than any

$\varepsilon > 0$. Thus, by assuming that $y_j \neq y_{j+1}$ and choosing $\varepsilon = \frac{1}{2}|y_{j+1} - y_j|\widetilde{f_1}$, there exists η such that $|y_{j+1} - y_j|\widetilde{f_1} \leq \frac{1}{2}|y_{j+1} - y_j|\widetilde{f_1}$ which leads to an obvious contradiction. Thus, (34) holds only if $y_{j+1} = y_j$. By iterating this argument, we obtain that $y_2 = y_3 = \ldots = y_{p+1}$ and by applying this result on the equation (33), one may assert that for any $y \in \mathbb{R}$

$$y \sum_{k=1}^{p+1} f_k(z_1, z_2) = 0,$$

which implies that $y = 0$ since $f_k(z_1, z_2)$ is positive for any $(z_1, z_2) \in D_j(\eta)$ (j arbitrarily chosen in $\{1, \cdots, p+1\}$) ■

We propose a simulation study to verify the consistency of maximum pseudo-likelihood estimator. We consider the model (31) with the vector of parameters $\boldsymbol{\theta} = (0, 2, 4)$. The vector of bounds $\boldsymbol{d}$ is assumed to be known and fixed to $\boldsymbol{d} = (0, 20, 80)$. The simulation procedure used here is a direct adaptation to the Delaunay energies of the Geyer and Møller proposal (Geyer and Møller (1994), Geyer (1999)). We refer the reader to Bertin et al. (1999d) for a detail of the used algorithm. One simulation of such a point process is proposed in Figure 1. Table 1 summarizes the different results obtained via $m = 5000$ replications each one is generated after one million of iterations of the algorithm. One may verify that both the bias and the standard deviation become smaller and smaller as the domain Λ_n grows.

6.2 Other examples of pairwise interaction models

In order to satisfy $\mathbf{C_3^{exp}}$ and $\mathbf{N_3^{exp}}$ for models on the complete graph or on the k nearest-neighbours graph with multi-Strauss pairwise interaction function, we can chosse as in Jensen and Künsch (1994) a configuration with one point or two points. On the delaunay graph, it may be interesting to study multi-Strauss interaction function on the circumradius or on the smallest angle of each Delaunay triangle. As discussed previously the identifiability assumption $\mathbf{C_3^{exp}}$ holds easily but $\mathbf{N_3^{exp}}$ needs more attention. Otherwise, for pairwise Delaunay model, we can replace the assumption on the smallest angle by a hard-core assumption and then Ω, by the set of admissible configurations $\Omega_\delta = \{\varphi \in \Omega : \forall x, y \in \varphi \times \varphi, x \neq y \quad ||x - y|| \geq \delta\}$.

Références

A.J. Baddeley and J. Møller. Nearest-Neighbour Markov Point Processes and Random Sets. *Int. Statist. Rev.*, 57(2) :89–121, 1989.

E. Bertin, J.-M. Billiot, and R. Drouilhet. Existence of Delaunay Pairwise Gibbs Point Processes with Superstable Component. *J. of Statist. Physics*, 95 :719–744, 1999a.

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

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

J. Besag. Statistical analysis of non-lattice data. *The statistician*, 24 :192–236, 1975.

J. Besag, R. Milne, and S. Zachary. Point process limits of lattice processes. *Ann. Appl. Prob.*, 19 :210–216, 1982.

J.E. Besag. Spatial interaction and the statistical analysis of lattice system. *J. R. Statist. Soc. Ser. B*, 26 :192–236, 1974.

J.-M. Billiot. Asymptotic properties of takacs-fiksel estimation method for gibbs point processes. *Statistics*, 33 :69–89, 1997.

J.-D. Boissonnat and M. Yvinec. *Géométrie Algorithmique.* Ediscience international, Paris, 1995.

D.J. Daley and D. Vere-Jones. *An introduction to the theory of point processes.* Springer Series in Statistics, Springer-Verlag, New York, Berlin, Heidelberg, London, Paris, Tokyo, 1988.

P.J. Diggle, T. Fiksel, P. Grabarnik, Y. Ogata, D. Stoyan, and M. Tanemura. On Paramater Estimation for Pairwise Interaction Point Process. *Int. Statist. Rev.*, 62(1) :99–117, 1994.

P.J. Diggle, D.J. Gates, and A. Stibbard. A non parametric estimator for pairwise-interaction point process. *Biometrika*, 74 :763–770, 1987.

R.L. Dobrushin. Gibbsian random field. The general case. *Func. Anal. Appl.*, 3 :22–28, 1969.

H. Edelsbrunner. *Algorithms in combinatorial geometry.* Texts and Monographs in Computer Science, Springer Verlag, New York, 1988.

R.P. Feynman. *Statistical mechanics. A set of lectures.* W.A. Benjamin, Reading, Mass, 1972.

T. Fiksel. Estimation of interaction potentials of gibbsian point processes. *Statistics*, 19(1) : 77–86, 1988.

C.J. Geyer. Likelihood inference for spatial point processes. In W.S. Kendall & M.N.M. van Lieshout O.E. Barndorff-Nielsen, editor, *Stochastic Geometry : likelihood and Computation*, pages 79–140, Bocca Raton, Florida, 1999.

C.J. Geyer and J. Møller. Simulation Procedures and Likelihood Inference for Spatial Point Processes. *Scand. J. of Statist.*, 21 :359–373, 1994.

C.J. Geyer and E. Thompson. Constrained monte carlo maximum likelihood for dependent data (with discussion). *J. R. Statist. Soc. Ser. B*, 54 :657–699, 1992.

E. Glötzl. Locale Energien und Potentiale für Punkprozesse. *Math. Nachr.*, 96 :195–206, 1980.

E. Glötzl and B. Rauschenschwandtner. On statistics of gibbsian point processes. In V.M. Zolotarev P. Revesz, L. Schmetterer, editor, *Lectures Notes in Statistics*, pages 83–93, 1981.

X. Guyon. Méthodes de pseudo-vraisemblance et de codage pour les processus ponctuels de gibbs. Technical Report 6, SAMOS, Université Paris 1, 1991.

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

J. Heikkinen and A. Penttinen. Bayesian smoothing in the estimation of the pair potential function of gibbs point processes. *Bernoulli*, 5 :1119–1136, 1999.

L. Heinrich. Mixing properties of gibbsian point normality of takacs-fiksel estimates. Technical Report 92-051, Universität Bielefeld, 1992.

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

J.L. Jensen and H.R. Künsch. On asymptotic normality of pseudo likelihood estimates of pairwise interaction processes. *Ann. Inst. Statist. Math.*, 46 :475–486, 1994.

J.L. Jensen and J. Møller. Pseudolikelihood for exponential family models of spatial point processes. *Ann. Appl. Probab.*, 1 :445–461, 1991.

O. Kallenberg. *Random Measures.* Academie-Verlag and Academic Press, Berlin, London, 1983.

O.E. Lanford and D. Ruelle. Observables at infinity and states with support range correlations in statistical mechanics. *Commun. Math. Phys.*, 13 :194–215, 1969.

S. Mase. Uniform l.a.n. condition of planar gibbsian point processes and optimality of maximum likelihood estimators of soft-core potential functions. *Prob. Theor. Relat. Fields*, 92 :51–67, 1992.

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, 1999.

J. Møller and R. Waggepetersen. *Statistical Inference and Simulation for Spatial Point Processes.* Chapman and Hall/CRC, Boca Raton, 2003.

R.A. Moyeed and A.J. Baddeley. Stochastic approximation of the m.l.e. for a spatial point pattern. *Scand. J. Statist.*, 18 :39–50, 1991.

X.X. Nguyen and H. Zessin. Ergodic theorems for Spatial Process. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 48 :133–158, 1979.

Y. Ogata and M. Tanemura. Likelihood analysis of spatial point patterns. *J. R. Statist. Soc. Ser. B*, 46 :496–518, 1984.

A. Penttinen. *Modelling interactions in spatial point patterns : parameter estimation by the maximum likehood method.* PhD thesis, University of Jyväskylä, 1984.

F.P. Preparata and M.I.S. Shamos. *Computational Geometry, an Introduction.* Springer Verlag, New York, 1988.

C.J. Preston. *Random Fields*, volume 534. Springer-Verlag, Berlin, Heidelberg, New York, 1976.

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

D. Stoyan, W.S. Kendall, and J. Mecke. *Stochastic Geometry and its Applications.* Wiley, Chichester, second edition, 1995.

R. Takacs. Estimator for the pair-potential of a gibbsian point process. *Statistics*, 17 :429–433, 1986.

Authors : Jean-Michel Billiot, Jean-François Coeurjolly and Rémy Drouilhet.

Address : LABSAD, BSHM, 1251 avenue centrale BP 47 - 38040 GRENOBLE Cedex 09.

E-mail addresses :

`Jean-Michel.Billiot@upmf-grenoble.fr`

`Jean-Francois.Coeurjolly@upmf-grenoble.fr`

`Remy.Drouilhet@upmf-grenoble.fr`

Corresponding author : Jean-François Coeujolly.

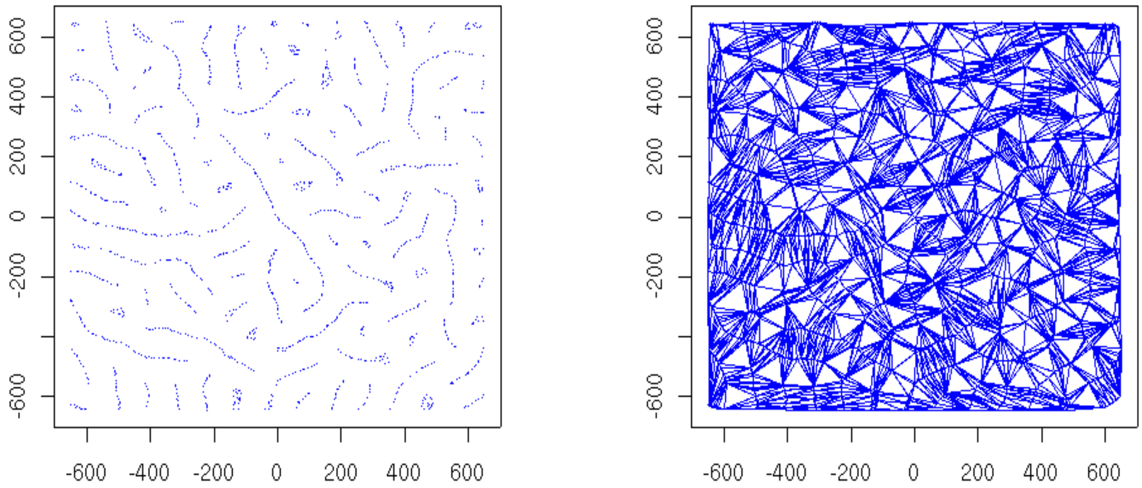

Fig. 1 – Example of the points, with (left) and without the edges (right), of the realization of the β-Delaunay model with multi-Strauss pairwise interaction function where parameters $\boldsymbol{\theta}$ and $\boldsymbol{d}$ are respectively fixed to $(0, 2, 4)$ and $(0, 20, 80)$.

Domain Λ_n	Mean of Estim. of θ_2	(Std Dev.)	Mean of Estim. of θ_3	(Std Dev.)
$[-250, 250]^2$	2.068	0.104	4.382	0.786
$[-350, 350]^2$	2.049	0.071	4.223	0.551
$[-450, 450]^2$	2.041	0.056	4.144	0.436

Tab. 1 – Empirical mean and standard deviation of maximum pseudo-likelihood estimates of parameters of $\theta_2 = 2$ and $\theta_3 = 4$ representing the levels of a multi-Strauss pairwise interaction function where the vector of bounds is assumed to be known and fixed to $\boldsymbol{d} = (0, 20, 80)$. These results are obtained from $m = 5000$ replications of the point process described by (31) generated in the domain $[-600, 600]^2$. Three sizes of domains Λ_n have been considered.