HAL
open science

Takacs-Fiksel method for stationary marked Gibbs point processes

Jean-François Coeurjolly, David Dereudre, Rémy Drouilhet, Frédéric Lavancier

To cite this version:

Jean-François Coeurjolly, David Dereudre, Rémy Drouilhet, Frédéric Lavancier. Takacs-Fiksel method for stationary marked Gibbs point processes. Scandinavian Journal of Statistics, 2012, 39 (3), pp.416-443. 10.1111/j.1467-9469.2011.00738.x . hal-00502004v2

HAL Id: hal-00502004

https://hal.science/hal-00502004v2

Submitted on 6 Jan 2011

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.

Takacs-Fiksel method for stationary marked Gibbs point processes

J.-F. Coeurjolly[1,2], D. Dereudre[3], R. Drouilhet[2] and F. Lavancier[4]

[1] GIPSA-lab, Grenoble University, France,
[2] Laboratoire Jean Kuntzmann, Grenoble University, France
[3] LAMAV UVHC FR 2956, Lille Nord de France University, France,
[4] Laboratoire de Mathématiques Jean Leray, Nantes University, France.

January 6, 2011

Abstract

This paper studies a method to estimate the parameters governing the distribution of a stationary marked Gibbs point process. This procedure, known as the Takacs-Fiksel method, is based on the estimation of the left and right hand sides of the Georgii-Nguyen-Zessin formula and leads to a family of estimators due to the possible choices of test functions. We propose several examples illustrating the interest and flexibility of this procedure. We also provide sufficient conditions based on the model and the test functions to derive asymptotic properties (consistency and asymptotic normality) of the resulting estimator. The different assumptions are discussed for exponential family models and for a large class of test functions. A short simulation study is proposed to assess the correctness of the methodology and the asymptotic results.

Keywords: stationary marked Gibbs point processes, parametric estimation, Takacs-Fiksel method, asymptotic properties, ergodic theorem, central limit theorem.

1 Introduction

Spatial point pattern data arise in a wide range of applications and since the seventies [10], [31] various statistical methods have been developed to study these kinds of data (see [36], [37] or [27] for recent reviews). In particular, a spatial point process is often modelled as the realization of a Gibbs distribution, defined through an interaction function, also called Hamiltonian. Gibbs models are extensions of the well-known Poisson process since they constitute a way to introduce dependence between points. Inference for parametric models in this setting has known a large development during the last decade. The most popular method to estimate the parameters is certainly the maximum likelihood estimator (MLE). It involves an intractable normalizing constant, but recent developments in computational statistics, in particular perfect simulations, have made inference feasible for many Gibbs models (see [4]). Although the MLE suffers from a lack of theoretical justifications (only some results for sparse patterns are proposed in [33]), some comparison studies, as in [18], have shown that it outperforms the other estimation methods. Nevertheless, the computation of the MLE remains very time-consuming and even extremely difficult to perform for some models. It is thus necessary to have quick alternative estimators at one's disposal, at least to propose relevant initial values for the MLE computation. The maximum pseudo-likelihood estimator (MPLE for short) constitutes one of them. Proposed by Besag in [7] and popularized by J.L. Jensen and J. Møller in [29] and A. Baddeley and R. Turner in [1] for spatial point processes, this method has the advantage of being theoretically well-understood (see [9], [16]) and of being much

faster to compute than the MLE. Another estimation procedure is the Takacs-Fiksel estimator, which arose from [19], [45], [20]. It can be viewed, in some sense, as a generalization of the MPLE. As a matter of fact, the Takacs-Fiksel method is not very popular, nor really used in practice. The main reason is certainly its relatively poor performances, in terms of mean square error, observed for some particular cases as in [18]. However, we think that this procedure deserves some consideration for several reasons that we expose below.

The Takacs-Fiksel procedure is based on the Georgii-Nguyen-Zessin formula (GNZ formula for short). Empirical counterparts of the left and the right hand sides of this equation are considered, and the induced estimator is such that the difference of these two terms is close to zero. Since the GNZ formula is valid for any test functions, the Takacs-Fiksel procedure does not only lead to one particular estimator but to a family of estimators, depending on the choice of the test functions. This flexibility is the main advantage of the procedure. We present several examples in the present paper. Let us summarize them in order to underline the interest of this method (see Section 3.2 for more details). First, this procedure can allow us to achieve estimations that likelihood-type methods cannot. As an example, we focus on the quermass model, which gathers the area interaction point process as a particular case (see [30], [14]). This model is sometimes used for geometric random objects. From a data set, one typically does not observe the point pattern but only some geometric sets arising from these points. The non-observability of the points makes the likelihood-type inference unfeasible. As we will show, this problem may be solved thanks to the Takacs-Fiksel procedure, provided that the test functions are chosen properly.

Another motivation is the possibility to choose test functions depending on the Hamiltonian in order to construct quicker estimators which do not require the computation of an integral for each value of the parameter. This improvement appears crucial for rigid models, such as those involved in stochastic geometry (see [17]), for which the MLE is prohibitively time-consuming and the MPLE still remains difficult to implement. Moreover, for some models, it is even possible to obtain explicit estimators which do not require any simulation nor optimization. This is illustrated for the Strauss model.

Therefore, it appears important to us to understand the theoretical properties of this procedure. This problem is the main objective of the present paper. We prove the consistency and the asymptotic normality of the induced estimator in a very general setting. In particular, we obtain a central limit theorem for the Takacs-Fiksel estimator with the classical rate of convergence, i.e. square root of the volume of the observation domain. This asymptotic result leads to the following comment: as a quick consistent estimator, the Takacs-Fiksel estimator appears to be a very good starting point for refined algorithms. Among them, let us mention the first step Newton-Raphson algorithm (as used in [26]) which allows an accurate approximation of the MLE, starting from a consistent estimator, in only one step. Although the theoretical justifications are missing in the Gibbs framework, it is well-known that this procedure leads to an efficient estimator in the classical iid case (see [32]). Another possibility could be to exploit the local asymptotic normality of the model in order to construct an adaptive estimator from the Takacs-Fiksel estimator. The Local Asymptotic Normality property (LAN) has only been proved for restrictive models in [33], but one can hope that it remains true for most Gibbs models. This procedure also leads to an efficient estimator. All these possibilities are interesting prospects for future investigations.

Some asymptotic properties of the Takacs-Fiksel procedure have already been investigated in two previous studies: one by L. Heinrich in [25] and the one by J.-M. Billiot in [8]. These papers have different frameworks and are based on different tools, but they both involve regularity and integrability type assumptions on the Hamiltonian and a theoretical condition which ensures that the contrast function (associated to the Takacs-Fiksel procedure) has a unique minimum. In [25], the consistency and the asymptotic normality are obtained for a quite large class of test functions. These results are, however, proved under the Dobrushin condition (see Theorems 2 and 3 in [25]) which implies the uniqueness of the underlying Gibbs measure and some mixing properties. This condition imposes a dramatic reduction of the space of possible values for the parameters of the model. In [8], the author focuses only on pairwise interaction point processes (which excludes the quermass model for example). The author mainly obtained the consistency for a specific class of test functions. In the case of a multi-Strauss pairwise interaction point process, the author also proved that the identifiability condition holds for the class of test functions he considered.

In contrast, our asymptotic results are proved in a very general setting, i.e. for a large class of stationary marked Gibbs models and test functions. The method employed to prove asymptotic normality is based on a conditional centering assumption, first appeared in [24] for the Ising model and generalized to certain spatial point processes in [28]. The main restriction that this method induces is only the finite range of the Hamiltonian. There are no limitations on the space of parameters and, in particular, the possible presence of phase transition does not affect the asymptotic behavior of the estimator. Moreover, the test functions may depend on the parameters. This extension seems important to us because, as emphasized in Section 3.2.2, such test functions can lead to quick and/or explicit estimators. All the general hypotheses assumed for the asymptotic results are discussed. For this, we focus on exponential family models, that is, on models whose interaction function is linear in the parameters. We show that our integrability and regularity assumptions are not restrictive since they are valid for a large class of models such as the Multi-Strauss marked point process, the Strauss-disc type point process, the Geyer's triplet point process, the quermass model and for all test functions used as a motivation for this work. In the setting of the exponential family models, we also discuss the classical identifiability condition which is required for the Takacs-Fiksel procedure. To the best of our knowledge, this is the first attempt to discuss it. We will specially dwell on questions like: what choices of test functions (and how many test functions) lead to a unique minimum of the contrast function? We propose general criteria and provide examples. It seems commonly admitted that to achieve the identification of the Takacs-Fiksel procedure, one should at least choose as many test functions as the number of parameters. As a consequence of our study, it appears that one should generally strictly choose more test functions than the number of parameters to achieve identification.

The rest of the paper is organized as follows. Section 2 introduces notation and a short background on marked Gibbs point processes. The Takacs-Fiksel method is presented in Section 3. It is based on the GNZ formula which is recalled in Section 3 also. Several examples of test functions are given. They aim at illustrating our interest in considering the Takacs-Fiksel procedure. The asymptotic results of the induced estimator are proposed in Section 4. Our results are obtained from a single realization observed in a domain whose volume is supposed to increase to infinity. Some integrability and regularity assumptions made for the Hamiltonian and for the test functions are discussed in this section while in Section 5 the identifiability condition is specifically dealt with. In Section 6, the very special situation where the energy function is not hereditary is considered. The GNZ formula is no longer valid in this setting, but it has been recently extended in [16] thanks to a slight modification. This leads to a natural generalization of the Takacs-Fiksel procedure. In Section 7, we propose a brief simulation study in order to assess the correctness of the asymptotic results we obtained in Section 4. Finally, Section 8 contains the proofs of the asymptotic results.

2 Background and notation

2.1 General notation, configuration space

Subregions of $\mathbb{R}^d$ will typically be denoted by Λ or Δ and will always be assumed to be Borel with positive Lebesgue measure. We write $\Lambda \Subset \mathbb{R}^d$ if Λ is bounded. Λ^c denotes the complementary set of Λ inside $\mathbb{R}^d$. 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 $\Lambda \Subset \mathbb{R}^d$, $|\Lambda|$ is the volume of Λ; For $x \in \mathbb{R}^d$, $|x|$ corresponds to its uniform norm while $\|x\|$ is its Euclidean norm. For all $x \in \mathbb{R}^d, \rho > 0$, let $\mathcal{B}(x,\rho) := \{y \in \mathbb{R}^d, \|y-x\| < \rho\}$. For a matrix $\underline{\mathbf{M}}$, let $\|\underline{\mathbf{M}}\|$ be the Frobenius norm of $\underline{\mathbf{M}}$ defined by $\|\underline{\mathbf{M}}\|^2 = Tr(\underline{\mathbf{M}}^T\underline{\mathbf{M}})$, where Tr is the trace operator.

The space $\mathbb{R}^d$ is endowed with the Borel σ-algebra $\mathcal{B}(\mathbb{R}^d)$ and the Lebesgue measure λ. Let $\mathbb{M}$ be a measurable space, which aims at being the mark space, endowed with the σ-algebra $\mathcal{M}$ and the probability measure λ^{m}. The state space of the point processes will be $\mathbb{S} := \mathbb{R}^d \times \mathbb{M}$ measured by $\mu := \lambda \otimes \lambda^{\mathrm{m}}$. We shall denote for short $x^m = (x,m)$ an element of $\mathbb{S}$.

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

the application $\varphi \mapsto \cup_{(y,m)\in\varphi}\{(y-x,m)\}$). For the sake of simplicity, we set $\varphi \cup x^m := \varphi \cup \{x^m\}$ and $\varphi \setminus x^m := \varphi \setminus \{x^m\}$.

2.2 Marked 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 $\mathbb{R}^d$. Let us briefly recall their definition.

A marked point process Φ is an Ω-valued random variable, with probability distribution P on $(\Omega, \mathcal{F})$. The most prominent marked point process is the marked Poisson process π^z with intensity measure $z\lambda \otimes \lambda^{\mathrm{m}}$ on $\mathbb{S}$, with $z > 0$ (see for example [34] for definition and properties). For $\Lambda \Subset \mathbb{R}^d$, let us denote by π^z_Λ the marginal probability measure in Λ of the Poisson process with intensity z. Without loss of generality, the intensity z is fixed to 1 and we simply write π and π_Λ in place of π^1 and π^1_Λ.

Let $\theta \in \mathbb{R}^p$ (for some $p \geq 1$). For any $\Lambda \Subset \mathbb{R}^d$, let us consider the parametric function $V_\Lambda(.;\theta)$ from Ω into $\mathbb{R} \cup \{+\infty\}$. From a physical point of view, $V_\Lambda(\varphi;\theta)$ is the energy of φ_Λ in Λ given the outside configuration φ_{Λ^c}. In this paper, we focus on stationary point processes on $\mathbb{R}^d$, i.e. with stationary (i.e. T-invariant) probability measure. For any $\Lambda \Subset \mathbb{R}^d$, we therefore consider $V_\Lambda(.;\theta)$ to be T-invariant, i.e. $V_\Lambda(\tau_x\varphi;\theta) = V_\Lambda(\varphi;\theta)$ for any $x \in \mathbb{R}^d$. Furthermore, $(V_\Lambda(.;\theta))_{\Lambda\Subset\mathbb{R}^d}$ is a compatible family of energies, i.e. for every $\Lambda \subset \Lambda' \Subset \mathbb{R}^d$, there exists a measurable function $\psi_{\Lambda,\Lambda'}$ from Ω into $\mathbb{R} \cup \{+\infty\}$ such that

$$\forall \varphi \in \Omega \quad V_{\Lambda'}(\varphi;\theta) = V_\Lambda(\varphi;\theta) + \psi_{\Lambda,\Lambda'}(\varphi_{\Lambda^c};\theta). \tag{1}$$

A stationary marked Gibbs point process is characterized by a stationary marked Gibbs measure usually defined as follows (see [41]).

Definition 1 *A probability measure P_θ on Ω is a stationary marked Gibbs measure for the compatible family of T-invariant energies $(V_\Lambda(.;\theta))_{\Lambda\Subset\mathbb{R}^d}$ if for every $\Lambda \Subset \mathbb{R}^d$, for P_θ-almost every outside configuration φ_{Λ^c}, the law of P_θ given φ_{Λ^c} admits the following conditional density with respect to π_Λ:*

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

where $Z_\Lambda(\varphi_{\Lambda^c};\theta)$ is a normalization called the partition function.

The existence of a Gibbs measure on Ω which satisfies these conditional specifications is a difficult issue. We refer the interested reader to [42], [41], [5], [13], [15] for the technical and mathematical development of the existence problem.

In a first step, we assume that the family of energies is hereditary (the non-hereditary case will be considered in Section 6), which means that for any $\Lambda \Subset \mathbb{R}^d$, for any $\varphi \in \Omega$, and for all $x^m \in \Lambda \times \mathbb{M}$,

$$V_\Lambda(\varphi;\theta) = +\infty \Rightarrow V_\Lambda(\varphi \cup x^m;\theta) = +\infty, \tag{2}$$

or equivalently, for all $x^m \in \varphi_\Lambda$, $f_\Lambda(\varphi_\Lambda|\varphi_{\Lambda^c};\theta) > 0 \Rightarrow f_\Lambda(\varphi_\Lambda \setminus x^m|\varphi_{\Lambda^c};\theta) > 0$.

The minimal assumption of our paper is then:

> **[Mod]**: For any $\theta \in \mathbf{\Theta}$, where $\mathbf{\Theta}$ is a compact subset of $\mathbb{R}^p$, there exists a stationary marked Gibbs measure P_θ for the compatible T-invariant hereditary family $(V_\Lambda(.;\theta))_{\Lambda\Subset\mathbb{R}^d}$. Our data consist in the realization of a marked point process Φ with stationary marked Gibbs measure $P_{\theta^\star}$, where $\theta^\star \in \mathring{\mathbf{\Theta}}$ is the unknown parameter vector to estimate.

Let us note that **[Mod]** ensures the existence of at least one stationary Gibbs measure. When this Gibbs measure is not unique, we say that the phase transition occurs. In this situation the set of Gibbs measures is a Choquet simplex and any Gibbs measure is a mixture of extremal ergodic Gibbs measures. If the Gibbs measure is unique, it is necessary ergodic (see [22] for more details about these properties).

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^m into the configuration φ and expressed for any $\Lambda \ni x$ by

$$V\left(x^m|\varphi;\theta\right) := V_\Lambda(\varphi \cup x^m;\theta) - V_\Lambda(\varphi;\theta). \tag{3}$$

From the compatibility of the family of energies, i.e. (1), this definition does not depend on Λ.

3 The Takacs-Fiksel estimation procedure

3.1 Presentation

The basic ingredient for the definition of the Takacs-Fiksel method is the so-called GNZ formula. This equation was proved by Georgii, Nguyen and Zessin in the seventies but other authors such as Papangelou and Takahashi also contributed to its establishment. See [40] and [46] for historical comments and [21] or [38] for a general presentation.

Lemma 1 (Georgii-Nguyen-Zessin Formula) *Under **[Mod]**, for any measurable function $h(\cdot,\cdot;\theta) : \mathbb{S}\times\Omega\to\mathbb{R}$ such that the following quantities are defined and finite, then*

$$\mathbf{E}\left(\int_{\mathbb{R}^d\times\mathbb{M}} h\left(x^m,\Phi;\theta\right)e^{-V(x^m|\Phi;\theta^\star)}\mu(dx^m)\right)=\mathbf{E}\left(\sum_{x^m\in\Phi} h\left(x^m,\Phi\setminus x^m;\theta\right)\right), \tag{4}$$

where $\mathbf{E}$ denotes the expectation with respect to $P_{\theta^\star}$.

For stationary marked Gibbs point processes, (4) reduces to

$$\mathbf{E}\left(h\left(0^M,\Phi;\theta\right)e^{-V\left(0^M|\Phi;\theta^\star\right)}\right)=\mathbf{E}\left(h\left(0^M,\Phi\setminus 0^M;\theta\right)\right), \tag{5}$$

where M denotes a random variable with probability distribution λ^{m}.

A second tool used throughout the paper is the ergodic Theorem established in [39]. Let us give a simpler form, here, which is sufficient in this paper.

Lemma 2 (Ergodic result) *Under **[Mod]**, we assume that $P_{\theta^\star}$ is ergodic. Then for any family of measurable functions F_Λ, indexed by the bounded sets Λ, from Ω to $\mathbb{R}$ which are additive (i.e. $F_{\Lambda\cup\Lambda'}=F_\Lambda+F_{\Lambda'}-F_{\Lambda\cap\Lambda'}$), shift invariant (i.e. $F_\Lambda(\varphi)=F_{\tau(\Lambda)}(\tau(\varphi))$ for any translation τ) and integrable (i.e. $\mathbf{E}(|F_{[0,1]^d}|)<+\infty$), we have that for $P_{\theta^\star}$-almost every φ*

$$\lim_{n\to+\infty}|\Lambda_n|^{-1}F_{\Lambda_n}(\varphi)=\mathbf{E}(F_{[0,1]^d}),$$

where $\Lambda_n=[-n,n]^d$ (other regular domains $(\Lambda_n)_{n\geq 1}$ converging towards $\mathbb{R}^d$ could be also considered).

Let $h(\cdot,\cdot;\theta):\mathbb{S}\times\Omega\to\mathbb{R}$ and let us define for any $\varphi\in\Omega$, $\theta\in\mathbf{\Theta}$ and $\Lambda\Subset\mathbb{R}^d$

$$C_\Lambda(\varphi;h,\theta):=\int_{\Lambda\times\mathbb{M}} h(x^m,\varphi;\theta)e^{-V(x^m|\varphi;\theta)}\mu(dx^m)-\sum_{x^m\in\varphi_\Lambda} h(x^m,\varphi\setminus x^m;\theta). \tag{6}$$

Assume that we observe the realization of a marked point process Φ satisfying **[Mod]** in a domain Λ_n. For appropriate choices of the functional h and a sequence of domain Λ_n, then it is possible to apply the ergodic result in Lemma 2 to prove that the first and second terms of $|\Lambda_n|^{-1}C_{\Lambda_n}(\Phi;h,\theta)$ respectively converge $P_{\theta^\star}$-almost surely to the left and right terms of (5).

The latter observation is the basic argument to define the Takacs-Fiksel method. Let us give K functions $h_k(\cdot,\cdot;\theta):\mathbb{S}\times\Omega\to\mathbb{R}$ (for $k=1,\ldots,K$), then the Takacs-Fiksel estimator is simply defined by

$$\widehat{\theta}_n(\varphi):=\widehat{\theta}_n^{TF}(\varphi)=\arg\min_{\theta\in\mathbf{\Theta}}\sum_{k=1}^K C_{\Lambda_n}(\varphi;h_k,\theta)^2, \tag{7}$$

where $\arg\min_{\theta\in\mathbf{\Theta}}F(\theta)$ means the parameter θ which minimizes the function F. Under identification assumption (16) given later, this minimum is obtained for a unique θ provided that n is large enough.

3.2 Some examples

In this section, some examples of models and test functions h, involved in (6), are provided. The choices made in previous studies are presented in 3.2.1. The two examples presented in 3.2.2 and 3.2.3 show the relevance of the Takacs-Fiksel procedure to provide quick estimates. The last example, in 3.2.4, is concerned with the possible identification of a special marked point process, the quermass model, and shows that appropriate choices of test functions can solve identification problems when points are not observed.

There is no asymptotic consideration in this section. Therefore, the different estimates are defined over a window, say Λ, and are denoted by $\widehat{\theta}$.

3.2.1 Classical examples

Let us first quote the particular case when the Takacs-Fiksel estimator reduces to the maximum pseudo-likelihood estimator introduced by [29] for spatial point processes. The MPLE is obtained by maximizing the log-pseudo-likelihood contrast function, given by

$$LPL_\Lambda(\varphi;\theta) = -\int_{\Lambda\times\mathbb{M}} e^{-V(x^m|\varphi;\theta)}\mu(dx^m) \; - \sum_{x^m\in\varphi_\Lambda} V(x^m|\varphi\setminus x^m;\theta). \tag{8}$$

Therefore, with the choice $h_k(x^m,\varphi;\theta) = \frac{\partial}{\partial\theta_k}V(x^m|\varphi;\theta)$, $k=1,\dots,p$, the estimator $\widehat{\theta}$, defined by (7), solves the system $C_\Lambda(\varphi;h_k,\theta)=0$, $k=1,\dots,p$, which means that $\widehat{\theta}$ is the root of the gradient vector of LPL_Λ, i.e. $\widehat{\theta}$ is the MPLE.

The first empirical study of the Takacs-Fiksel estimator can be found in [18], where this estimate is compared to other estimators for some unmarked pairwise interaction point processes with two parameters. In this context different test functions $h_{r_1},\cdots,h_{r_K}$ were used, where for $r>0$

$$h_r(x,\varphi;\theta) = \left|\varphi_{\mathcal{B}(x,r)}\right| = \sum_{y\in\varphi}\mathbf{1}_{[0,r]}(\|y-x\|).$$

The integral term involved in (6) is approximated by discretization and the induced estimation (7) is then assessed. Note that with this choice of test functions, the sum term in (6), when normalized by $|\Lambda|^{-1}$, is an estimation of $\rho^2\mathcal{K}(r)$, where $\mathcal{K}(\cdot)$ is the reduced second order function and ρ denotes the intensity of the stationary point process Φ, i.e. for all $B\in\mathcal{B}(\mathbb{R}^d)$, $\mathbf{E}(\Phi(B))=\rho|B|$.

The latter choice requires the computation of the integral in (6) for all θ. A more convenient choice could be the one first proposed by Fiksel:

$$h_r(x,\varphi;\theta) = \left|\varphi_{\mathcal{B}(x,r)}\right| e^{V(x|\varphi;\theta)} = e^{V(x|\varphi;\theta)}\sum_{y\in\varphi}\mathbf{1}_{[0,r]}(\|y-x\|). \tag{9}$$

In the stationary case, this leads to the following approximation thanks to the ergodic theorem

$$\frac{1}{|\Lambda|}\int_\Lambda h_r(x,\varphi;\theta)e^{-V(x|\varphi;\theta)}dx \approx \mathbf{E}\left|\Phi_{\mathcal{B}(0,r)}\right| = \rho\pi r^2. \tag{10}$$

The integral term in (6) is thus easily approximated by $|\varphi_\Lambda|\pi r^2$ for all θ, while the sum can be explicitly computed. These historical choices of test functions have a natural extension in the marked case.

3.2.2 Some choices leading to quick estimations

The main advantage of the Takacs-Fiksel procedure is to provide quick consistent estimators, that might supply initial values for a more evolved procedure. A simple way to achieve this goal is to generalize (9) and consider test functions of the form

$$h(x^m,\varphi;\theta) = \tilde{h}(x^m,\varphi)e^{V(x^m|\varphi;\theta)},$$

where $\tilde{h}(x^m,\varphi)$ does not depend on θ. So, the integral term in (6) has to be computed only once and not for all θ, while the sum term in (6) does not require any approximation. Hence, the optimisation problem (7) may be resolved very quickly.

In some particular examples, explicit formulas may even be obtained for the integral term, as in (10). In the same spirit, an explicit estimator for the Strauss interaction is provided below.

3.2.3 An example of explicit estimator for the Strauss process

The (non-marked) Strauss process with range of interaction $R > 0$ is given for any $\Lambda \Subset \mathbb{R}^d$ by

$$V_\Lambda(\varphi;\theta) = \theta_1|\varphi_\Lambda| + \theta_2 \sum_{\substack{x,y\in\varphi \\ \{x,y\}\cap\Lambda\neq\emptyset}} \mathbf{1}_{[0,R]}(\|x-y\|), \tag{11}$$

where $\theta_1 \in \mathbb{R}$ and $\theta_2 > 0$ are the two parameters of the model. Alternatively,

$$V(x|\varphi;\theta) = \theta_1 + \theta_2 \sum_{y\in\varphi} \mathbf{1}_{[0,R]}(\|y-x\|).$$

Let us consider the following family of test functions, for $k \in \mathbb{N}\setminus\{0\}$,

$$h_k(x,\varphi;\theta) = \begin{cases} e^{(k-1)\theta_2} & \text{if } \left|\varphi_{\mathcal{B}(x,R)}\right| = k-1, \\ 0 & \text{otherwise.} \end{cases} \tag{12}$$

This choice gives in (6)

$$C_\Lambda(\varphi;h_k,\theta) = e^{-\theta_1}V_{k-1,\Lambda}(\varphi) - e^{(k-1)\theta_2}N_{k-1,\Lambda}(\varphi),$$

where $N_{k,\Lambda}(\varphi)$ denotes the number of points $x \in \varphi_\Lambda$ such that $|\mathcal{B}(x,R)\cap(\varphi\setminus x)| = k$ and $V_{k,\Lambda}(\varphi)$ denotes the volume of the set $\{y \in \Lambda, |\mathcal{B}(y,R)\cap\varphi| = k\}$.

Several explicit estimators may be obtained following (7) from (at least) two test functions as above. Let us quote the simplest one, corresponding to the choice h_1 and h_2 in (7). This leads to the contrast function $C_\Lambda(\varphi;h_1,\theta)^2 + C_\Lambda(\varphi;h_2,\theta)^2$ which vanishes at the unique point $(\widehat{\theta}_1(\varphi),\widehat{\theta}_2(\varphi))$ with

$$\widehat{\theta}_1(\varphi) = \ln\left(\frac{V_{0,\Lambda}(\varphi)}{N_{0,\Lambda}(\varphi)}\right), \quad \widehat{\theta}_2(\varphi) = \ln\left(\frac{V_{1,\Lambda}(\varphi)}{N_{1,\Lambda}(\varphi)}\right) - \ln\left(\frac{V_{0,\Lambda}(\varphi)}{N_{0,\Lambda}(\varphi)}\right). \tag{13}$$

This estimator of (θ_1,θ_2) is completely explicit, provided the quantities $N_{k,\Lambda}(\varphi)$ and $V_{k,\Lambda}(\varphi)$ are available. They can be easily approximated by computational geometry tools.

3.2.4 A solution for unobservability issues

The quermass model introduced in [30] is a marked point process which aims at modelling random sets in $\mathbb{R}^2$. This is a generalization of the well-known Boolean model to interacting random balls. Let us denote by x^R a marked point where x and $R > 0$ (i.e. the mark) respectively represent the center and the radius of the associated ball $\mathcal{B}(x,R)$. For a finite configuration φ, i.e. with a finite support instead of $\mathbb{R}^2$, the quermass energy is defined for $(\theta_1,\theta_2,\theta_3,\theta_4) \in \mathbb{R}^4$ by:

$$V(\varphi;\theta) = \theta_1|\varphi| + \theta_2\,\mathcal{P}(\Gamma) + \theta_3\,\mathcal{A}(\Gamma) + \theta_4\,\mathcal{E}(\Gamma) \quad \text{where } \Gamma = \bigcup_{(x,R)\in\varphi} \mathcal{B}(x,R)$$

and $\mathcal{P}(\Gamma)$, $\mathcal{A}(\Gamma)$ and $\mathcal{E}(\Gamma)$ denote respectively the perimeter, the area and the Euler-Poincaré characteristic (i.e. number of components minus number of holes) of the set Γ. To extend this definition to the infinite support $\mathbb{R}^2$, it is convenient to suppose that the radii of the balls are almost surely uniformly bounded (i.e. $\lambda^{\mathrm{m}}([0,R_0]) = 1$ for some $R_0 > 0$). In this case the family of energies (V_Λ) is defined by

$$V_\Lambda(\varphi;\theta) = V\left(\varphi_{\Lambda\oplus B(0,2R_0)};\theta\right) - V\left(\varphi_{\Lambda\oplus B(0,2R_0)\setminus\Lambda};\theta\right).$$

This definition may be extended to unbounded radius, though a restriction to the so-called tempered configurations is needed to ensure the existence of the associated Gibbs measure. We refer to [14] for more details.

When $\theta_2 = \theta_3 = \theta_4 = 0$, this model reduces to the Boolean model (see [44] for a survey). The area process (see [3]) is also a particular case, taking $\theta_2 = \theta_4 = 0$.

In practice, one only observes the random set Γ, so the marked points x^R in φ are unknown. A challenging task is then to estimate the parameters $(\theta_1, \theta_2, \theta_3, \theta_4)$ in the presence of this unobservability issue. In particular, a direct application of the maximum likelihood or pseudo-likelihood method is impossible to estimate all the parameters, and especially θ_1 which requires the observation of the number of points in φ. For the other parameters, which are related to the observable functionals $\mathcal{P}$, $\mathcal{A}$ and $\mathcal{E}$, the MLE has been investigated in [35].

Let us show that the Takacs-Fiksel procedure may be used to estimate θ_1 in spite of this unobservability issue. Indeed, it is possible to choose some test function h such that both the integral and the sum in (6) are computable. The unobservability issue occurs mainly for the sum term. Let us consider the following example of test function:

$$h_{per}(x^R, \varphi; \theta) = \mathcal{P}\left(\mathcal{C}(x, R) \cap \Gamma^c\right), \tag{14}$$

where $\mathcal{C}(x, R)$ is the sphere $\{y, ||y - x|| = R\}$. For any finite configuration φ, we then have

$$\sum_{x^R \in \varphi} h_{per}(x^R, \varphi \setminus x^R; \theta) = \mathcal{P}(\Gamma),$$

so that this sum is computable even if each term $h_{per}(x^R, \varphi \setminus x^R; \theta)$ is not. If the configuration φ is infinite then for any bounded set Λ, $\sum_{x^R \in \varphi_\Lambda} h_{per}(x^R, \varphi \setminus x^R; \theta)$ is equal to the perimeter of Γ restricted to Λ plus a boundary term which is asymptotically negligible with respect to the volume of Λ.

Consequently, assuming $(\theta_2, \theta_3, \theta_4)$ known, θ_1 may be estimated thanks to (7) with the above choice of test function.

The joint estimation of all four parameters might be achieved thanks to additional test functions sharing the same property as above, i.e. such that the sum in (6) is observable.

In the particular case of the area process, $\theta_2 = \theta_4 = 0$ and R is constant (i.e. $\lambda^{\mathrm{m}} = \delta_R$), it suffices to find one more test function to ensure an identifiable estimation (see Example 2 in Section 5 for more details). A possible additional test function is

$$h_{iso}(x^R, \varphi; \theta) = \begin{cases} 1 & \text{if} \quad \mathcal{P}\left(\mathcal{C}(x, R)\right) = 2\pi R \\ 0 & \text{otherwise.} \end{cases} \tag{15}$$

In this case $\sum_{x^R \in \varphi} h_{iso}(x^R, \varphi \setminus x^R; \theta)$ corresponds to the number of isolated balls in Γ.

4 Asymptotic results for the Takacs-Fiksel estimator

We present in this section asymptotic results for the Takacs-Fiksel estimator for a point process satisfying **[Mod]** and assumed to be observed in a domain Λ_n, where $(\Lambda_n)_{n \geq 1}$ is a sequence of increasing cubes whose size goes to $+\infty$ as n goes to $+\infty$.

First, for a function g depending on θ, we denote by $\mathbf{g}^{(1)}(\theta)$ (resp. $\underline{\mathbf{g}}^{(2)}(\theta)$) the gradient vector of length p (resp. the Hessian matrix of size (p, p)) evaluated at θ. Let us rewrite the Takacs-Fiksel estimator as

$$\widehat{\theta}_n(\varphi) = \arg\min_{\theta \in \mathbf{\Theta}} U_{\Lambda_n}(\varphi; \mathbf{h}, \theta),$$

with $U_{\Lambda_n}(\varphi; \mathbf{h}, \theta) = |\Lambda_n|^{-2} \sum_{k=1}^{K} C_{\Lambda_n}(\varphi; h_k, \theta)^2$, where $\mathbf{h} = (h_1, \ldots, h_K)$ and C_{Λ_n} is given by (6).

4.1 Consistency

The consistency is obtained under the following assumptions, denoted by **[C]**: for any Gibbs measure $P_{\theta^\star}$, for all $\theta \in \mathbf{\Theta}$, $k = 1, \ldots, K$ and $\varphi \in \Omega$

[C1] For all $x \in \mathbb{R}^d$, $h_k(x^m, \varphi; \theta) = h_k(0^m, \tau_x \varphi; \theta)$ and

$$\mathbf{E}\left(\left|h_k\left(0^M, \Phi; \theta\right)\right| e^{-V\left(0^M | \Phi; \theta'\right)}\right) < +\infty, \text{ for } \theta' = \theta, \theta^\star.$$

[C2] $U_{\Lambda_n}(\varphi; \mathbf{h}, \cdot)$ is a continuous function for $P_{\theta^\star}-$a.e. φ.

[C3]

$$\sum_{k=1}^{K} \mathbf{E}\left(h_k(0^M,\Phi;\theta)\left(e^{-V(0^M|\Phi;\theta)}-e^{-V(0^M|\Phi;\theta^\star)}\right)\right)^2=0 \quad \Longrightarrow \theta=\theta^\star. \tag{16}$$

[C4] h_k and f_k, defined by $f_k(x^m,\varphi;\theta):=h_k(x^m,\varphi;\theta)e^{-V(x^m|\varphi;\theta)}$, are continuously differentiable and

$$\mathbf{E}\left(\max_{\theta\in\mathbf{\Theta}}|f_k(0^M,\Phi;\theta)|\right)<+\infty \quad \text{and} \quad \mathbf{E}\left(\max_{\theta\in\mathbf{\Theta}}|h_k(0^M,\Phi;\theta)|e^{-V(0^M|\Phi;\theta^\star)}\right)<+\infty,$$

$$\mathbf{E}\left(\max_{\theta\in\mathbf{\Theta}}\|\mathbf{f_k}^{(1)}(0^M,\Phi;\theta)\|\right)<+\infty \quad \text{and} \quad \mathbf{E}\left(\max_{\theta\in\mathbf{\Theta}}\|\mathbf{h_k}^{(1)}(0^M,\Phi;\theta)\|e^{-V(0^M|\Phi;\theta^\star)}\right)<+\infty.$$

Theorem 3 *Assuming [**Mod**] and [**C**] then, as $n\to+\infty$, the Takacs-Fiksel estimator $\widehat{\theta}_n(\varphi)$ converges towards $\theta^\star$ for $P_{\theta^\star}-$a.e. φ.*

Assumptions **[C1]**, **[C2]** and **[C4]** are related to the regularity and the integrability of the different test functions and the local energy function. Some general criteria may be proposed to verify these assumptions, see Section 4.3 for a discussion. Assumption **[C3]** corresponds to an identifiability condition and requires much more attention. It is well-known that such an assumption is fulfilled when $\mathbf{h}=\mathbf{V}^{(1)}$ (leading to the MPLE) under mild assumptions (see Assumption **[Ident]** proposed by [9]). The question to know if this remains true for more general test functions is difficult (actually it is untrue in several cases). This will be discussed specifically in Section 5.

4.2 Asymptotic normality

We need the following assumptions denoted by **[N]**: For any Gibbs measure $P_{\theta^\star}$, $k=1,\dots,K$, $\Lambda\Subset\mathbb{R}^d$, $\varphi\in\Omega$ and θ in a neighborhood $\mathcal{V}(\theta^\star)$ of $\theta^\star$:

[N1] $\mathbf{E}\left(|C_\Lambda(\Phi;h_k,\theta^\star)|^3\right)<+\infty$.

[N2] For any sequence of bounded domains Γ_n such that $\Gamma_n\to 0$ as $n\to+\infty$, $\mathbf{E}\left(C_{\Gamma_n}(\Phi;h_k,\theta^\star)^2\right)\to 0$.

[N3] $C_\Lambda(\varphi;h_k,\theta^\star)$ depends only on $\varphi_{\Lambda\oplus B(0,D)}$ for some $D\geq 0$ (which is uniform in Λ, φ, $\theta^\star$).

[N4] h_k and f_k (defined in **[C4]**) are twice continuously differentiable in θ and

$$\mathbf{E}\left(\|\underline{\mathbf{h}}^{(2)}(0^M,\Phi;\theta)\|e^{-V(0^M|\Phi;\theta^\star)}\right)<+\infty \quad \text{and} \quad \mathbf{E}\left(\|\underline{\mathbf{f}}^{(2)}(0^M,\Phi;\theta)\|\right)<+\infty.$$

Let us remark that Assumption **[N3]** leads us to consider in general that V has a finite range, which means that there exists $D\geq 0$ such that for all $(m,\varphi)\in\mathbb{M}\times\Omega$ and all $\theta\in\mathbf{\Theta}$

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

The same kind of finite range property is also expected for (h_k).

Theorem 4 *Under Assumptions [**Mod**], [**C**] and [**N**], for any ergodic Gibbs measure $P_{\theta^\star}$ the following convergence in distribution holds as $n\to+\infty$*

$$|\Lambda_n|^{1/2}\underline{\mathcal{E}}(\mathbf{h},\theta^\star)\underline{\mathcal{E}}(\mathbf{h},\theta^\star)^T\left(\widehat{\theta}_n(\Phi)-\theta^\star\right)\xrightarrow{d}\mathcal{N}\left(0,\underline{\mathcal{E}}(\mathbf{h},\theta^\star)\underline{\mathbf{\Sigma}}(\mathbf{h},\theta^\star)\underline{\mathcal{E}}(\mathbf{h},\theta^\star)^T\right), \tag{17}$$

where $\underline{\mathcal{E}}(\mathbf{h},\theta^\star)$ is the (p,K) matrix defined for $i=1,\dots,p$ and $k=1,\dots,K$ by

$$\left(\underline{\mathcal{E}}(\mathbf{h},\theta^\star)\right)_{ik}=\mathbf{E}\left(h_k(0^M,\Phi;\theta^\star)\left(\mathbf{V}^{(1)}(0^M|\Phi;\theta^\star)\right)_i e^{-V(0^M|\Phi;\theta^\star)}\right)$$

and where $\underline{\mathbf{\Sigma}}(\mathbf{h},\theta^\star)$ is the (K,K) matrix defined by

$$\underline{\mathbf{\Sigma}}(\mathbf{h},\theta^\star)=D^{-d}\sum_{|\ell|\leq 1}\mathbf{E}\left(\widetilde{\mathbf{C}}_{\Delta_0(D)}(\Phi;\mathbf{h},\theta^\star)\widetilde{\mathbf{C}}_{\Delta_\ell(D)}(\Phi;\mathbf{h},\theta^\star)^T\right), \tag{18}$$

where, for all $k\in\mathbb{Z}^d$, $\Delta_k(D)$ is the cube centered at kD with side-length D and where, for any bounded domain Λ, $\widetilde{\mathbf{C}}_\Lambda(\Phi;\mathbf{h},\theta^\star):=\left(C_\Lambda(\Phi;h_k,\theta^\star)\right)_{k=1,\dots,K}$.

Remark 1 *While Assumptions **[N1-3]** will ensure that a central limit theorem holds for* $\mathbf{U}^{(1)}_{\Lambda_n}(\Phi;\mathbf{h},\theta^\star)$*, Assumption **[N4]** is required to prove that* $\underline{\mathbf{U}}^{(2)}_{\Lambda_n}(\Phi;\mathbf{h},\theta)$ *is uniformly bounded in a neighborhood of* $\theta^\star$*. These two statements allow us to apply a general central limit theorem for minimum contrast estimators (e.g. Theorem 3.4.5 of [23]).*

Remark 2 *In Theorem 4, if the Gibbs measure* $P_{\theta^\star}$ *is stationary (and then not necessarily ergodic), it is a mixture of ergodic measures and the left hand term in (17) converges in distribution to a mixture of normal distributions. We may also propose an asymptotic result valid in the presence of a phase transition. Indeed, if the matrix* $\underline{\mathcal{E}}(\mathbf{h},\theta^\star)\underline{\mathbf{\Sigma}}(\mathbf{h},\theta^\star)\underline{\mathcal{E}}(\mathbf{h},\theta^\star)^T$ *is positive definite (for any extremal ergodic measure of the Choquet simplex of stationary Gibbs measures), then we may define a consistent empirical version* $S(\Phi)$ *of* $\left[\underline{\mathcal{E}}(\mathbf{h},\theta^\star)\underline{\mathbf{\Sigma}}(\mathbf{h},\theta^\star)\underline{\mathcal{E}}(\mathbf{h},\theta^\star)^T\right]^{-1/2}\underline{\mathcal{E}}(\mathbf{h},\theta^\star)\underline{\mathcal{E}}(\mathbf{h},\theta^\star)^T$ *(see Section 4 of [12] for more details) to obtain:*

$$|\Lambda_n|^{1/2}S(\Phi)\left(\widehat{\theta}_n(\Phi)-\theta^\star\right)\xrightarrow{d}\mathcal{N}\left(0,\underline{\mathbf{I}}\right).$$

Remark 3 *Following Section 3.2.1, let us underline that (17) is coherent with the asymptotic normality of the MPLE established in [11], i.e. with the case* $K=p$ *and* $\mathbf{h}=\mathbf{V}^{(1)}$*. Indeed, with similar assumptions to the ones presented in the present paper, Equation (4.4) in Theorem 2 [11] states that*

$$|\Lambda_n|^{1/2}\underline{\mathbf{A}}(\theta^\star)\left(\widehat{\theta}_n(\Phi)-\theta^\star\right)\xrightarrow{d}\mathcal{N}\left(0,\underline{\mathbf{\Sigma}}(\mathbf{V}^{(1)},\theta^\star)\right), \tag{19}$$

where $\underline{\mathbf{A}}(\theta^\star)$ *is the symmetric* (p,p) *matrix given for* $i,k=1,\ldots,p$ *by*

$$(\underline{\mathbf{A}}(\theta^\star))_{ik}=\mathbf{E}\left(\left(\mathbf{V}^{(1)}(0^M|\Phi;\theta^\star)\right)_k\left(\mathbf{V}^{(1)}(0^M|\Phi;\theta^\star)\right)_i e^{-V(0^M|\Phi;\theta^\star)}\right).$$

Since $\mathbf{h}=\mathbf{V}^{(1)}$*, then* $\underline{\mathcal{E}}(\mathbf{V}^{(1)},\theta^\star):=\underline{\mathbf{A}}(\theta^\star)$ *and therefore (17) reduces to*

$$|\Lambda_n|^{1/2}\underline{\mathbf{A}}(\theta^\star)^2\left(\widehat{\theta}_n(\Phi)-\theta^\star\right)\xrightarrow{d}\mathcal{N}\left(0,\underline{\mathbf{A}}(\theta^\star)\underline{\mathbf{\Sigma}}(\mathbf{V}^{(1)},\theta^\star)\underline{\mathbf{A}}(\theta^\star)\right)$$

which is exactly (19) by assuming that $\underline{\mathbf{A}}(\theta^\star)$ *is invertible.*

Remark 4 *The question how to choose the test functions in order to minimize the norm of the asymptotic covariance matrix is difficult to answer, still open and is a perspective for future work.*

4.3 Discussion

The present paragraph is devoted to the discussion of Assumptions **[Mod]**, **[C]** (except **[C3]**) and **[N]**. In the previous sections, we have expressed the different assumptions in a very general way. Our aim, here, is to make these assumptions concrete for a wide range of models and a wide range of test functions, in order to illustrate that our setting is not restrictive. In particular, we will focus on exponential family models having a local energy of the form:

$$V(x^m|\varphi;\theta):=\theta^T\mathbf{V}(x^m|\varphi)=\theta_1V_1(x^m|\varphi)+\ldots+\theta_pV_p(x^m|\varphi), \tag{20}$$

with $\mathbf{V}=(V_1,\ldots,V_p)$ a vector function from $\mathbb{S}\times\Omega(\mathbb{S})$ to $\mathbb{R}^p$.

Let us consider the following assumptions: for all $(m,\varphi)\in\mathbb{M}\times\Omega$.

[Exp] For $i=1,\cdots,p$, for all $x\in\mathbb{R}^d$, $V_i(x^m|\varphi)=V_i(0^m|\tau_x\varphi)$ and there exist $\kappa_i^{(\inf)},\kappa_i^{(\sup)}\geq 0$, $k_i\in\mathbb{N}, D>0$ such that one of the two following assumptions is satisfied :

$$\theta_i\geq 0 \text{ and } -\kappa_i^{(\inf)}\leq V_i(0^m|\varphi)=V_i(0^m|\varphi_{\mathcal{B}(0,D)})\leq\kappa_i^{(\sup)}|\varphi_{\mathcal{B}(0,D)}|^{k_i}.$$

or

$$-\kappa_i^{(\inf)}\leq V_i(0^m|\varphi)=V_i(0^m|\varphi_{\mathcal{B}(0,D)})\leq\kappa_i^{(\sup)}.$$

$\widetilde{[\mathbf{Exp}]}$ Assumption **[Exp]** with $k_i = 0$ or 1 for all i (when $\theta_i \geq 0$).

[H] For all $x \in \mathbb{R}^d$, $h(x^m, \varphi; \theta) = h(0^m, \tau_x\varphi; \theta)$ and there exist $\kappa > 0$, $k \in \mathbb{N}, D > 0$ such that $h(0^m, \varphi; \theta) = h(0^m, \varphi_{\mathcal{B}(0,D)}; \theta)$, such that $h(0^m, \varphi; \cdot)$ is twice continuously differentiable in θ and such that $|Y(\varphi, m)| \leq \kappa |\varphi_{\mathcal{B}(0,D)}|^k$, where

$$Y(\varphi, m) := \max\left(|h(0^m, \varphi; \theta)|, \|\mathbf{h}^{(1)}(0^m, \varphi; \theta)\|, \|\underline{\mathbf{h}}^{(2)}(0^m, \varphi; \theta)\|\right). \tag{21}$$

$\widetilde{[\mathbf{H}]}$ $h(0^m, \varphi; \theta) = \widetilde{h}(0^m, \varphi; \theta) e^{\theta^T \mathbf{V}(0^m|\varphi)}$ with $\widetilde{h}$ satisfying **[H]**.

Let us underline that the different constants involved in these assumptions are assumed to be independent of m, φ, θ. Note also that if the test function h is independent of θ, $Y(\varphi, m)$ obviously reduces to $|h(0^m, \varphi)|$.

These assumptions are common and very simple to check. Assumption **[Exp]** has already been investigated in [9]. It includes a wide variety of models such as the overlap area point process, the multi-Strauss marked point process, the $k-$nearest-neighbor multi-Strauss marked point process, the Strauss type disc process, the Geyer's triplet point process, the area process, some special cases of quermass process (for instance when $\lambda^{\mathbb{m}}$ has a compact support not containing 0 and $\theta_4 = 0$), etc. Among these models, the only one that does not satisfy $\widetilde{[\mathbf{Exp}]}$ is the Geyer's triplet point process (see [9], p.242).

On the other hand, the test functions $h(x^m, \varphi; \theta) = 1$, $h(x^m, \varphi; \theta) = \mathbf{V}_k^{(1)}(x^m|\varphi; \theta) = V_k(x^m|\varphi)$, $h(x^m, \varphi; \theta) = |\varphi_{\mathcal{B}(x,r)}|$ satisfy **[H]** (for the second one, it is implied by **[Exp]**). Note that the functional described by (12) for the Strauss model depending on θ, the functionals h_{per} and h_{iso} in (14), (15) also satisfy **[H]**. In a similar way, test functions like $e^{\theta^T \mathbf{V}(x^m|\varphi)}$, $e^{\theta^T \mathbf{V}(x^m|\varphi)/2}$, $|\varphi_{\mathcal{B}(x,r)}| e^{\theta^T \mathbf{V}(x^m|\varphi)}$, $\mathbf{1}_{[0,r]}(d(x^m, \varphi)) e^{\theta^T \mathbf{V}(x^m|\varphi)}$ satisfy $\widetilde{[\mathbf{H}]}$.

We show that most of all the assumptions required in Theorems 3 and 4 are not too restrictive.

Proposition 5 *(i) Under Assumption **[Exp]**, Assumption **[Mod]** is fulfilled.*
*(ii) For a test function satisfying **[H]** (resp. $\widetilde{[\mathbf{H}]}$) and a model satisfying **[Exp]** (resp. $\widetilde{[\mathbf{Exp}]}$), then Assumptions **[C]** (excepted **[C3]**) and **[N]** are fulfilled.*

Proof. *(i)* From **[Exp]**, it is easy to check that for any fixed θ, there exists a positive constant $K(\theta)$ such that $V(0^m|\varphi; \theta) \geq -K(\theta)$. Therefore the local energy is stable for any θ. On the other hand it is finite range. These two properties imply the existence of ergodic measures (see *e.g.* [5], Proposition 1). Among them, there exists at least one stationary measure due to the invariance by translation of the family of energy functions.

(ii) **[C2]** and **[N3]** are quite obvious to check. Now, from the stability of the local energy, we have that for all $(m, \varphi) \in \mathbb{M} \times \Omega$, $\theta^T \mathbf{V}(0^m|\varphi) \geq -\rho$ for $\rho < +\infty$, independent of m, φ, θ (this is possible because $\mathbf{\Theta}$ is compact). Let us also underline that this property ensures that for every $\Lambda \Subset \mathbb{R}^d$, every $c \in \mathbb{R}$, $\mathbf{E}(e^{c|\Phi_\Lambda|}) < +\infty$ (see *e.g.* Proposition 11 of [6]), which obviously implies that $\mathbf{E}(|\Phi_\Lambda|^\alpha) < +\infty$ and $\mathbf{E}(|\Phi_\Lambda|^\alpha e^{c|\Phi_\Lambda|}) < +\infty$ for every $\alpha > 0$. Now, under **[H]** and **[Exp]** (or $\widetilde{[\mathbf{H}]}$ and $\widetilde{[\mathbf{Exp}]}$), the expectations in **[C1]**, **[C4]**, **[N1]** and **[N4]** are clearly finite. Let us focus on **[C1]** for example (the justification for the other assumptions is similar). We have for any $\theta, \theta' \in \mathbf{\Theta}$

$$\mathbf{E}\left(|h(0^M, \Phi; \theta)| e^{-\theta'^T \mathbf{V}(0^M|\Phi)}\right)$$

$$\leq \begin{cases} e^\rho \mathbf{E}\left(|h(0^M, \Phi; \theta)|\right) \leq c \times e^\rho \mathbf{E}\left(|\Phi_{\mathcal{B}(0,D)}|^\alpha\right) & \text{under } [\mathbf{H}] \text{ and } [\mathbf{Exp}] \\ e^\rho \mathbf{E}\left(|\widetilde{h}(0^M, \Phi; \theta)| e^{\theta^T \mathbf{V}(0^M|\Phi)}\right) \leq c \times e^\rho \mathbf{E}\left(|\Phi_{\mathcal{B}(0,D)}|^\alpha e^{c|\Phi_{\mathcal{B}(0,D)}|}\right) & \text{under } \widetilde{[\mathbf{H}]} \text{ and } \widetilde{[\mathbf{Exp}]}, \end{cases}$$

for some constants α and c. Note that, if we had not assumed $\widetilde{[\mathbf{Exp}]}$ for test functions of the form $\widetilde{[\mathbf{H}]}$, then one would have had expectations of the form $\mathbf{E}\left(e^{|\Phi_\Lambda|^k}\right)$ for some $k > 1$ which is not necessarily finite under the local stability property. Assumption **[N2]** is proved similarly and by using the dominated convergence theorem. ■

Remark 5 *By following ideas in [11], it is possible to fulfill the integrability type assumptions for more complicated models such as the Lennard-Jones model (which is not locally stable and nonlinear in terms of the parameters). The using of Ruelle's estimates [43] plays a crucial role in this case of superstable interaction. For the sake of conciseness and simplicity, we do not investigate this in the present paper.*

5 Identifiability : Assumption [C3]

The Assumption **[C3]** is related to the identifiability of the estimation procedure. It is more complicated to verify than the other assumptions and an investigation to obtain a criterion or a characterization seems necessary. We address this question in this section.

In the following, we consider that the interaction has an exponential form as in (20). Then **[C3]** is equivalent to: $\theta = \theta^\star$ is the unique solution of the nonlinear system of equations in θ defined by

$$\mathbf{E}\left(h_k(0^M, \Phi; \theta)\left(e^{-\theta^T \mathbf{V}(0^M|\Phi)} - e^{-\theta^{\star T}\mathbf{V}(0^M|\Phi)}\right)\right) = 0, \qquad 1 \leq k \leq K. \tag{22}$$

If h_k and $\mathbf{V}$ are sufficiently regular, each equation in (22) gives a $(p-1)$-dimensional manifold of solutions in $\mathbf{\Theta}$ containing $\theta^\star$. So it is clear that the choice $K \geq p$ is in general necessary to prove that the system (22) admits the unique solution $\theta^\star$.

In Section 5.1, we investigate the delicate case $K = p$ in detail. In opposition to the linear case where p hyperplanes in $\mathbb{R}^p$ have in general a unique common point, the intersection of p $(p-1)$-dimensional manifolds does not generally reduce to a single point. So, when $K = p$, there is no guarantee that (22) has a unique solution $\theta^\star$. This is illustrated by a simple example at the beginning of Section 5.1. In Proposition 6, we provide a criterion to ensure that the system in (22) admits the only one solution $\theta^\star$. Some examples, for which the criterion is available, are presented and the rigidity of the criterion, when $p \geq 3$, is also evoked. In the case where $p = 2$, we show that our criterion is not far from being necessary.

The case $K > p$ is studied in Section 5.2. The identification problem should be simpler since, in general, $p+1$ $(p-1)$-dimensional manifolds in $\mathbb{R}^p$ have no common point. We give a sufficient criterion to prove the identification but we think that it is far from being necessary.

Before presenting these two sections, let us give further notation. We denote by P_V the law of $\mathbf{V}(0^M, \Phi)$ in $\mathbb{R}^p$. We also define the function Ψ_θ, for each $\theta \in \mathbf{\Theta}$, by

$$\begin{array}{rccc} \Psi_\theta : & \mathbb{R}^p & \longrightarrow & \mathbb{R}^K \\ & \mathbf{v} & \longmapsto & \left(\begin{array}{c} \mathbf{E}\Big(h_1(0^M, \Phi; \theta)\Big|\mathbf{V}(0^M|\Phi) = \mathbf{v}\Big) \\ \vdots \\ \mathbf{E}\Big(h_K(0^M, \Phi; \theta)\Big|\mathbf{V}(0^M|\Phi) = \mathbf{v}\Big) \end{array}\right) \end{array}. \tag{23}$$

We will see that this function plays a crucial role in the identification problem.

5.1 The case $K = p$

First of all, let us give a simple example to show that the identification problem is delicate in the situation where $K = p$. Let us consider that $K = p = 2$, $V_1 = 1$, and let us choose the simple test functions $h_1 = 1$ and $h_2 = e^{\theta^T \mathbf{V}(x^m|\varphi)}$. Then $\tilde{\theta} = (\tilde{\theta}_1, \tilde{\theta}_2)$ with $\tilde{\theta}_1 = \theta_1^\star - \ln(\mathbf{E}(e^{-\theta_2^\star V_2(0^M|\Phi)}))$ and $\tilde{\theta}_2 = 0$ is always a solution of the system in (22). Therefore if $\theta_2^\star \neq 0$ and if $\tilde{\theta}$ defined before is in $\mathbf{\Theta}$, then the system in (22) admits at least two solutions.

In the following, we first give a sufficient criterion to prove the identifiability and propose some examples. Next, we show the rigidity of our criterion which seems constraining when $p \geq 3$.

5.1.1 Criterion for identifiability

Assumption **[Det]** gathers the two following assumptions:

[Det(≠)] For every θ in $\mathbf{\Theta}$, $det(\mathbf{v}_1, \ldots, \mathbf{v}_p) det\big(\Psi_\theta(\mathbf{v}_1), \ldots, \Psi_\theta(\mathbf{v}_p)\big)$ is not $(P_V)^{\otimes p}$-a.s. identically null

[Det(≥)] For every θ in $\mathbf{\Theta}$, there exists $\epsilon = \pm 1$ such that for $(P_V)^{\otimes p}$-a.s. every $(\mathbf{v}_1, \ldots, \mathbf{v}_p)$ in $(\mathbb{R}^p)^p$

$$\epsilon\, det(\mathbf{v}_1, \ldots, \mathbf{v}_p) det\big(\Psi_\theta(\mathbf{v}_1), \ldots, \Psi_\theta(\mathbf{v}_p)\big) \geq 0.$$

When $\epsilon = 1$ (respectively $\epsilon = -1$), **[Det(≥)]** means that Ψ_θ preserves the sign (respectively the opposite sign) of the determinant.

The criterion is the following.

Proposition 6 *If $K = p$ then Assumption* **[Det]** *ensures that Assumption **[C3]** holds.*

Proof. Denoting by ζ the real function $x \mapsto \ln\left(\frac{e^x - 1}{x}\right)$ with the convention $\zeta(0) = 0$, the equations (22) become

$$(\theta^\star - \theta)^T \mathbf{X}_k(\theta, \theta^\star) = 0, \qquad 1 \leq k \leq p, \tag{24}$$

where the vector $\mathbf{X}_k(\theta, \theta^\star)$ is defined by

$$\mathbf{X}_k(\theta, \theta^\star) = \mathbf{E}\left(h_k(0^M, \Phi; \theta) e^{-\theta^{\star T}\mathbf{V}(0^M|\Phi)} e^{\zeta\left((\theta^\star - \theta)^T \mathbf{V}(0^M|\Phi)\right)} \mathbf{V}(0^M|\Phi)\right). \tag{25}$$

Therefore the system (22) admits the unique solution $\theta^\star$ if the family of vectors $(\mathbf{X}_k(\theta, \theta^\star))_{1 \leq k \leq p}$ is independent in $\mathbb{R}^p$. Let us give a formula of the determinant of these vectors which shows that it is not null.

Conditioning by the law of $\mathbf{V}(0^M|\Phi)$ and using the multi-linearity of the determinant, we obtain

$$\begin{aligned}
& det(\mathbf{X}_1(\theta, \theta^\star), \ldots, \mathbf{X}_p(\theta, \theta^\star)) \\
= & \int \cdots \int det\left(\mathbf{E}\left[h_1(0^M, \Phi; \theta) e^{-\theta^{\star T}\mathbf{v}_1} e^{\zeta\left((\theta^\star - \theta)^T \mathbf{v}_1\right)} \mathbf{v}_1 \middle| \mathbf{V}(0^M|\Phi) = \mathbf{v}_1\right], \ldots, \right. \\
& \left. \mathbf{E}\left[h_p(0^M, \Phi; \theta) e^{-\theta^{\star T}\mathbf{v}_p} e^{\zeta\left((\theta^\star - \theta)^T \mathbf{v}_p\right)} \mathbf{v}_p \middle| \mathbf{V}(0^M|\Phi) = \mathbf{v}_p\right]\right) P_V(d\mathbf{v}_1) \cdots P_V(d\mathbf{v}_p) \\
= & \int \cdots \int e^{\sum_{k=1}^p -\theta^{\star T}\mathbf{v}_k + \zeta\left((\theta^\star - \theta)^T \mathbf{v}_k\right)} det\left(\mathbf{v}_1, \ldots, \mathbf{v}_p\right) \prod_{k=1}^p \mathbf{E}\left[h_k(0^M, \Phi; \theta)|\mathbf{v}_k\right] P_V(d\mathbf{v}_1) \cdots P_V(d\mathbf{v}_p) \\
= & \frac{1}{p!} \sum_{\sigma \in S_p} \int \cdots \int e^{\sum_{k=1}^p -\theta^{\star T}\mathbf{v}_{\sigma(k)} + \zeta\left((\theta^\star - \theta)^T \mathbf{v}_{\sigma(k)}\right)} det\left(\mathbf{v}_{\sigma(1)}, \ldots, \mathbf{v}_{\sigma(p)}\right) \\
& \prod_{k=1}^p \mathbf{E}\left[h_k(0^M, \Phi; \theta)|\mathbf{v}_{\sigma(k)}\right] P_V(d\mathbf{v}_1) \cdots P_V(d\mathbf{v}_p),
\end{aligned}$$

where S_p is the set of all permutations in $\{1, \ldots, p\}$. Denoting by $\epsilon(\sigma)$ the signature of σ, we obtain

$$\begin{aligned}
& det(\mathbf{X}_1(\theta, \theta^\star), \ldots, \mathbf{X}_p(\theta, \theta^\star)) \qquad (26) \\
= & \frac{1}{p!} \int \cdots \int e^{\sum_{k=1}^p -\theta^{\star T}\mathbf{v}_k + \zeta\left((\theta^\star - \theta)^T \mathbf{v}_k\right)} det\left(\mathbf{v}_1, \ldots, \mathbf{v}_p\right) \sum_{\sigma \in S_p} \epsilon(\sigma) \prod_{k=1}^p \mathbf{E}\left[h_k(0^M, \Phi; \theta)|\mathbf{v}_{\sigma(k)}\right] \\
& P_V(d\mathbf{v}_1) \cdots P_V(d\mathbf{v}_p) \\
= & \frac{1}{p!} \int \cdots \int e^{\sum_{k=1}^p -\theta^{\star T}\mathbf{v}_k + \zeta\left((\theta^\star - \theta)^T \mathbf{v}_k\right)} det(\mathbf{v}_1, \ldots, \mathbf{v}_p) det\big(\Psi_\theta(\mathbf{v}_1), \ldots, \Psi_\theta(\mathbf{v}_p)\big) P_V(d\mathbf{v}_1) \cdots P_V(d\mathbf{v}_p).
\end{aligned}$$

From Assumption **[Det]** , this determinant is not null. The proposition is proved. ■

Now let us give some examples for which the criterion is valid.

Example 1 (linear case) *If the function Ψ_θ is linear and invertible then Assumption* **[Det(≥)]** *is clearly satisfied and* **[Det(≠)]** *holds as soon as the support of P_V is not included in a hyperplane. In particular, if $h_k = V_k$ for every $1 \leq k \leq p$ then Ψ_θ is equal to the identity function. This situation corresponds to the pseudo-likelihood procedure for which we regain the identifiability via our criterion.*

Example 2 (area process) *For the area process defined in Section 3.2.4 with $\lambda^{\mathrm{m}} = \delta_R$ (i.e. the radii of balls are constant), it is easy to check that the functions h_{per} and h_{iso} respectively defined by* (14) *and* (15) *give a function Ψ which satisfies Assumption* **[Det]** *. Indeed the support of P_V is the segment $\{1\} \times [0, \pi R^2]$ in $\mathbb{R}^2$ and for a vector $\mathbf{v} = (1, v_2)$ the image $\Psi(\mathbf{v})$ is $(\psi_1(v_2), 0)$ if $v_2 \neq \pi R^2$ and $(2\pi R, 1)$ if $v_2 = \pi R^2$. Therefore, it follows that* $[\mathbf{Det}(\geq)]$ *is satisfied and noting that $0 < P_V((1, \pi R^2)) < 1$ we deduce that* $[\mathbf{Det}(\neq)]$ *holds too.*

Example 3 (a general example with $p = 2$) *Example 2 is included in a more general setting when $p = 2$. Indeed let us suppose that the function Ψ_θ has the form $\Psi_\theta(v_1, v_2) = \big(g_\theta(v_1, v_2), g_\theta(v_1, v_2) f_\theta(v_2/v_1)\big)$ where g_θ is a nonnegative scalar function and f_θ is a monotone scalar function. Then Ψ_θ satisfies* $[\mathbf{Det}(\geq)]$ *and* $[\mathbf{Det}(\neq)]$ *holds if $g_\theta(v_1, v_2) f_\theta(v_2/v_1)$ is not $P_V^{\otimes 2}$-a.s. constant when $g_\theta(v_1, v_2)$ is not null.*

Example 4 (functions of the type $h_k e^{\theta^T \mathbf{V}}$) *Let us suppose that the functions (h_k) ensure that Ψ_θ satisfies* $[\mathbf{Det}(\geq)]$ *then for any nonnegative function g_θ from $\mathbb{R}^p$ to $\mathbb{R}$ the functions $(\tilde{h}_k) = \big(g_\theta(\mathbf{V}) h_k\big)$ also provide a function $\tilde{\Psi}_\theta$ satisfying* $[\mathbf{Det}(\geq)]$ *. This remark is related to Section 3.2.2 where it is suggested to choose functions $(\tilde{h}_k)$ of the form $(e^{\theta^T \mathbf{V}} h_k)$ to simplify the integral in (6).*
As an immediate consequence, the test functions $\big(V_k e^{\theta^T \mathbf{V}}\big)$, considered in [8] for the particular multi-Strauss point process, satisfy $[\mathbf{Det}(\geq)]$ *.*

5.1.2 Rigidity of the criterion

In this section, we give some comments about the rigidity of the criterion. In Proposition 7 below, we show that a function Ψ_θ, satisfying $[\mathbf{Det}(\geq)]$, has a strong linear structure since, under very reasonable assumptions, the image of any hyperplane is included in a hyperplane. For example, in the classical setting where $V_1 = 1$, the function Ψ_θ is defined from the affine space $H = \{1\} \times \mathbb{R}^{p-1}$ and if Ψ_θ is assumed to be continuous then the image of any $p-2$ dimensional affine space in H is included in a hyperplane. This property clearly shows that Ψ_θ is very rigid when $p \geq 3$.

However, when $p = 2$, we show in Proposition 8 that our criterion is not far from being necessary. Indeed, we present a large class of examples which do not satisfy our criterion and for which the identifiability fails.

Proposition 7 *Let Ψ be a continuous function from $\mathcal{D}$ to $\mathbb{R}^p$ satisfying* $[\mathbf{Det}(\geq)]$ *, where the domain $\mathcal{D}$ is a subset of $\mathbb{R}^p$ with the following property: for any $(x_i)_{1 \leq i \leq p} \in \mathcal{D}^p$ such that $det(x_1, \ldots, x_p) = 0$, then for any neighborhood $\mathcal{V}$ of (x_i), there exist (x_i^+) and (x_i^-) in $\mathcal{V} \cap \mathcal{D}^p$ such that $det(x_1^+, \ldots, x_p^+) > 0$ and $det(x_1^-, \ldots, x_p^-) < 0$. Then for any hyperplane H in $\mathbb{R}^p$ the image $\Psi(H \cap \mathcal{D})$ is included in a hyperplane of $\mathbb{R}^p$.*

Proof. Let H be a hyperplane in $\mathbb{R}^p$. To prove that $\Psi(H \cap \mathcal{D})$ is included in a hyperplane, it is sufficient to prove that the dimension of the vectorial space generated by the vectors in $\Psi(H \cap \mathcal{D})$ is not equal to p. Let us suppose that it is equal to p, then there exists $(x_i)_{1 \leq i \leq p}$ in $(H \cap \mathcal{D})^p$ such that $det(\Psi(x_1), \ldots, \Psi(x_p)) \neq 0$. Since $dim(H) = p - 1$, we have $det(x_1, \ldots, x_p) = 0$. By continuity of Ψ and by the local properties of $\mathcal{D}$ assumed in Proposition 7, we find $(x_i^+)_{1 \leq i \leq p}$ and $(x_i^-)_{1 \leq i \leq p}$ in $\mathcal{D}^p$ such that $det(x_1^+, \ldots, x_p^+) det(\Psi(x_1^+), \ldots, \Psi(x_p^+)) > 0$ and $det(x_1^-, \ldots, x_p^-) det(\Psi(x_1^-), \ldots, \Psi(x_p^-)) < 0$, which contradicts Assumption $[\mathbf{Det}(\geq)]$. ■

In the case where $\mathcal{D} = \mathbb{R}^p$, if we assume that $\Psi(\mathbb{R}^p)$ is not reduced to a hyperplane and that Ψ is differentiable at the origin, then we can show that Ψ satisfies $[\mathbf{Det}(\geq)]$ if and only if $\Psi(x) = g(x) A x$, where A is an invertible matrix and g a nonnegative scalar function. It means that Ψ is quasi linear and so the rigidity of Ψ is very strong.

Now let us focus on the case where $p = 2$ and let us show that, while our criterion seems very constraining, it is not far from being necessary in this case. We suppose that $p = 2$, $V_1 = 1$ and that the support of V_2 is included in an interval $[a, b]$. Let us remark that this case occurs for the area process with $[a, b] = [0, \pi R^2]$. First of all, it is easy to check visually, depending on the geometry of γ_θ defined by the curve $\Psi_\theta(\{1\} \times [a, b])$, whether Ψ_θ satisfies $[\mathbf{Det}]$ (see figure 1 for examples).

Moreover let us show that the criterion is not far from being necessary. Suppose that the functions (h_i) do not depend on θ and that $\Psi := \Psi_\theta$ satisfies for $\epsilon = \pm 1$ Assumption $[\widetilde{\mathbf{Det}}]$ decomposed into the three following assumptions:

$[\widetilde{\mathbf{Det}}(\neq)]$ $det\,(\mathbf{v}_1, \mathbf{v}_2)\, det\,(\Psi(\mathbf{v}_1), \Psi(\mathbf{v}_2))$ is not $P_V^{\otimes 2}$-a.s. identically null

$[\widetilde{\mathbf{Det}}(\geq)]$ there exists $\delta > 0$ such that for $P_V^{\otimes 2}$-a.s. every $(\mathbf{v}_1, \mathbf{v}_2)$ in $(\{1\} \times [a, a+\delta])^2$

$$\epsilon\, det\,(\mathbf{v}_1, \mathbf{v}_2)\, det\,(\Psi(\mathbf{v}_1), \Psi(\mathbf{v}_2)) \geq 0$$

$[\widetilde{\mathbf{Det}}(\leq)]$ there exists $\delta > 0$ such that for $P_V^{\otimes 2}$-a.s. every $(\mathbf{v}_1, \mathbf{v}_2)$ in $(\{1\} \times [b-\delta, b])^2$

$$\epsilon\, det\,(\mathbf{v}_1, \mathbf{v}_2)\, det\,(\Psi(\mathbf{v}_1), \Psi(\mathbf{v}_2)) \leq 0.$$

See Figure 1 for an example of such Ψ. Obviously, this situation is not exactly the opposite of Assumption $[\mathbf{Det}]$, but it is strongly related to it. Then, we have the following proposition which proves that the identifiability fails for this large class of examples.

Proposition 8 *If the functions (h_i) are nonnegative, if Ψ satisfies $[\widetilde{\mathbf{Det}}]$ and if $det\Big(E_{P_V}(\Psi(\mathbf{v})\mathbf{v}^T)\Big) \neq 0$ then **[C3]** fails.*

Let us note that even if the assumption $det\Big(E_{P_V}(\Psi(\mathbf{v})\mathbf{v}^T)\Big) \neq 0$ seems unnatural, it is in general satisfied.

Proof. Let us show that (22) admits another solution than $\theta^\star$. We only give here the main lines of the proof.

We denote by $\mathcal{O}$ the set in $\mathbb{R}^2$ containing the vectors $\mathbf{u}$ which are orthogonal to at least one vector $\mathbf{v}$ in $\{1\} \times [a, b]$, i.e. $\mathcal{O} = \{\mathbf{u} \in \mathbb{R}^2, \exists \mathbf{v} \in \{1\} \times [a, b], \mathbf{u}^T\mathbf{v} = 0\}$. In fact, $\mathcal{O}$ is the union of a cone $\mathcal{O}^+$ in the upper half plane and a cone $\mathcal{O}^-$ in the lower half plane. For any $\delta > 0$, the expression of the determinant in (26) can be split in two parts

$$\begin{aligned} & det(\mathbf{X}_1(\theta, \theta^\star), \mathbf{X}_2(\theta, \theta^\star)) \qquad (27) \\ = \; & \frac{1}{2} \int\int_{[a,a+\delta]^2} e^{\sum_{k=1}^2 -\theta^{\star T}\mathbf{v}_k + \zeta\left((\theta^\star - \theta)^T \mathbf{v}_k\right)} det(\mathbf{v}_1, \mathbf{v}_2) det\big(\Psi(\mathbf{v}_1), \Psi(\mathbf{v}_2)\big) P_V(d\mathbf{v}_1) P_V(d\mathbf{v}_2) \\ & + \frac{1}{2} \int\int_{[a,b]^2 \setminus [a,a+\delta]^2} e^{\sum_{k=1}^2 -\theta^{\star T}\mathbf{v}_k + \zeta\left((\theta^\star - \theta)^T \mathbf{v}_k\right)} det(\mathbf{v}_1, \mathbf{v}_2) det\big(\Psi(\mathbf{v}_1), \Psi(\mathbf{v}_2)\big) P_V(d\mathbf{v}_1) P_V(d\mathbf{v}_2). \end{aligned}$$

Let $\mathbf{u} \neq 0$ in $\mathcal{O}^+$ and $\theta = \theta^\star + \alpha\mathbf{u}$ with $\alpha > 0$. From $[\widetilde{\mathbf{Det}}(\neq)]$ and $[\widetilde{\mathbf{Det}}(\geq)]$ since ζ is increasing and $\zeta(x) \sim x$ as $x \to +\infty$, we deduce that the first integral in (27) dominates the second one when α goes to infinity. Therefore $det(\mathbf{X}_1(\theta, \theta^\star), \mathbf{X}_2(\theta, \theta^\star))$ has the sign ϵ (defined before $[\widetilde{\mathbf{Det}}]$) when α is large enough. Similarly, if $\mathbf{u}$ is in $\mathcal{O}^-$ then from $[\widetilde{\mathbf{Det}}(\neq)]$ and $[\widetilde{\mathbf{Det}}(\leq)]$ $det(\mathbf{X}_1(\theta, \theta^\star), \mathbf{X}_2(\theta, \theta^\star))$ has the sign $-\epsilon$ when α is large enough and it implies that the sign of $det(\mathbf{X}_1(\theta, \theta^\star), \mathbf{X}_2(\theta, \theta^\star))$, with $\theta = \theta^\star + \mathbf{u}$, is different for $\mathbf{u}$ in $\mathcal{O}^+$ or in $\mathcal{O}^-$ as soon as $|\mathbf{u}|$ is large enough. We deduce that there exists a continuous curve $t \mapsto \mathbf{u}(t)$ which crosses $\mathcal{O}$ such that $det(\mathbf{X}_1(\theta(t), \theta^\star), \mathbf{X}_2(\theta(t), \theta^\star)) = 0$ for every $\theta(t) = \theta^\star + \mathbf{u}(t)$. Let us note that the assumption $det\Big(E_{P_V}(\Psi(\mathbf{v})\mathbf{v}^T)\Big) \neq 0$ ensures that $det(\mathbf{X}_1(\theta^\star, \theta^\star), \mathbf{X}_2(\theta^\star, \theta^\star)) \neq 0$, and so $\mathbf{u}(t)$ is never null. Let us show that there exists t_0 such that $\theta(t_0)$ is a solution of the system in (22).

Since the functions (h_i) are nonnegative and from Definition (25), we obtain that for every t, $\mathbf{X}_1(\theta(t), \theta^\star)$ is collinear to a vector in $\{1\} \times [a, b]$. By continuity of the function $t \mapsto \mathbf{X}_1(\theta(t), \theta^\star)^T \mathbf{u}(t)$ and by the mean value theorem, there exists t_0 such that $\mathbf{u}(t_0)$ is orthogonal to $\mathbf{X}_1(\theta(t_0), \theta^\star)$. Since the determinant $det(\mathbf{X}_1(\theta(t_0), \theta^\star), \mathbf{X}_2(\theta(t_0), \theta^\star)) = 0$, $\mathbf{u}(t_0)$ is also orthogonal to $\mathbf{X}_2(\theta(t_0), \theta^\star)$ and it follows that the system in (24) provides at least two solutions $\theta^\star$ and $\theta(t_0)$. Identification assumption (22) or $[\mathbf{C3}]$ fail. ■

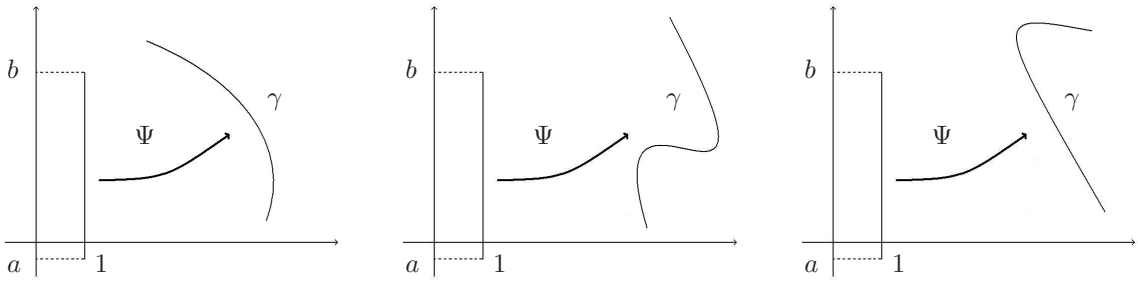

Figure 1: on the left (respectively in the middle), an example of function Ψ satisfying (respectively not satisfying) $[\mathbf{Det}(\geq)]$. On the right, an example of Ψ satisfying $[\widetilde{\mathbf{Det}}]$.

5.2 The case K>p

In the case where $K > p$, we noticed, in the introduction, that the identification problem should be simpler. Nevertheless, we did not find a satisfactory criterion to prove it. The following Proposition 9 gives a sufficient criterion which is probably far from being necessary. It is based on a slight modification of Assumption **[Det]** which does not seem to be the appropriate tool in this setting. However, in the case where $p = 2$ and $K = 3$, this condition reduces to a nice geometrical property which can be checked easily.

First, let us present the criterion. We denote by $\mathcal{A}$ the set of all subsets with p elements in $\{1, \ldots, K\}$,

$$\mathcal{A} = \Big\{ I \subset \{1, \ldots, K\}, \text{ such that } \#(I) = p \Big\}.$$

We say that Assumption **[Det']** is satisfied if, for every θ in $\mathbf{\Theta}$, there exists a family of real coefficients $(c_I)_{I\in\mathcal{A}}$ such that the two following assumptions hold:

[Det'(≠)] $\displaystyle\sum_{I\in\mathcal{A}} c_I det(\mathbf{v}_1, \ldots, \mathbf{v}_p) det\big(\Psi_\theta^I(\mathbf{v}_1), \ldots, \Psi_\theta^I(\mathbf{v}_p)\big)$ is not $(P_V)^{\otimes p}$-a.s. identically null

[Det'(≥)] for $(P_V)^{\otimes p}$-a.s. every $(\mathbf{v}_1, \ldots, \mathbf{v}_p)$ in $(\mathbb{R}^p)^p$

$$\sum_{I\in\mathcal{A}} c_I det(\mathbf{v}_1, \ldots, \mathbf{v}_p) det\big(\Psi_\theta^I(\mathbf{v}_1), \ldots, \Psi_\theta^I(\mathbf{v}_p)\big) \geq 0,$$

where $\Psi_\theta^I(\mathbf{v})$ denotes the p-dimensional vector extracted from $\Psi_\theta(\mathbf{v})$ for the coordinates given by I. In the particular case where $K = p$, Assumption **[Det']** becomes Assumption **[Det]** exactly.

Our criterion is the following

Proposition 9 *Assumption* **[Det']** *ensures that Assumption [C3] holds.*

Proof. As in the proof of Proposition 6, to show that (22) admits the unique solution $\theta^\star$, it is sufficient to prove that there exists I in $\mathcal{A}$ such that, for all $\theta \neq \theta^\star$, $det\big((\mathbf{X}_i(\theta, \theta^\star))_{i\in I}\big) \neq 0$. It is equivalent to: for every $\theta \neq \theta^\star$ in $\mathbf{\Theta}$ there exists a family of real coefficients $(c_I)_{I\in\mathcal{A}}$ such that

$$\sum_{I\in\mathcal{A}} c_I det\big((\mathbf{X}_i(\theta, \theta^\star))_{i\in I}\big) > 0. \tag{28}$$

With calculations as in (26) we obtain

$$\begin{aligned} & p! \sum_{I\in\mathcal{A}} c_I det\big((\mathbf{X}_i(\theta, \theta^\star))_{i\in I}\big) \\ = & \int e^{\sum_{k=1}^p -\theta^{\star T}\mathbf{v}_k + \zeta\left((\theta^\star-\theta)^T\mathbf{v}_k\right)} \left(\sum_{I\in\mathcal{A}} c_I det(\mathbf{v}_1, \ldots, \mathbf{v}_p) det\big(\Psi_\theta^I(\mathbf{v}_1), \ldots, \Psi_\theta^I(\mathbf{v}_p)\big) \right) P_V(d\mathbf{v}_1) \cdots P_V(d\mathbf{v}_p). \end{aligned}$$

Thanks to **[Det']** , this quantity is positive. ■

In the case where $p = 2$ and $K = 3$, **[Det'(≥)]** is satisfied if and only if there exist a, b, c in $\mathbb{R}^3$ such that for every $\mathbf{v}_1$ and $\mathbf{v}_2$ with $det(\mathbf{v}_1, \mathbf{v}_2) > 0$

$$a\, det\left(\Psi_\theta^{\{1,2\}}(\mathbf{v}_1), \Psi_\theta^{\{1,2\}}(\mathbf{v}_2)\right) + b\, det\left(\Psi_\theta^{\{1,3\}}(\mathbf{v}_1), \Psi_\theta^{\{1,3\}}(\mathbf{v}_2)\right) + c\, det\left(\Psi_\theta^{\{2,3\}}(\mathbf{v}_1), \Psi_\theta^{\{2,3\}}(\mathbf{v}_2)\right) \geq 0. \quad (29)$$

If we denote by $\wedge$ the vectorial product in $\mathbb{R}^3$, the inequality in (29) means that the following set

$$\left\{\Psi_\theta(\mathbf{v}_1) \wedge \Psi_\theta(\mathbf{v}_2),\ \text{for all } \mathbf{v}_1, \mathbf{v}_2 \text{ such that } det(\mathbf{v}_1, \mathbf{v}_2) > 0\right\} \quad (30)$$

is included in the half space in $\mathbb{R}^3$ with equation $cx - by + az \geq 0$. In the setting where $V_1 = 1$ and V_2 is included in an interval $[a, b]$, as for the area process, this condition becomes a geometrical characteristic of the curve $\gamma_\theta = \Psi_\theta(\{1\} \times [a, b])$ in $\mathbb{R}^3$, which is easy to check visually.

6 Extension in the presence of non-hereditary interaction

In several recent papers, Gibbs processes with non hereditary interactions are considered, in particular in the domain of stochastic geometry (see [13], [15]). The parametric estimation of such models has also been investigated. The first results in this direction have been given in [16] via a pseudo-likelihood procedure based on a generalization of the Georgii-Nguyen-Zessin formula (4). The same kind of generalization is possible for the Takacs-Fiksel procedure. We address this improvement in this section.

In the following, we do not assume that the energy $V_\Lambda(\varphi, \theta)$ satisfies the heredity assumption (2). The first consequences are that the local energy $V(x^m|\varphi; \theta)$ is not defined in general and that the Georgii-Nguyen-Zessin formula is not available. Let us begin by presenting the generalization of this formula, as stated in [16], Proposition 2, which is valid in the hereditary and non-hereditary settings.

We first need to recall the concept of removable points which has been introduced in [16], Definition 3.

Definition 2 *A point x^m in a configuration φ is called removable if there exists a bounded set Λ containing x such that $V_\Lambda(\varphi \backslash x^m, \theta) < +\infty$. We denote by $\mathcal{R}_\theta(\varphi)$ the set of removable points in φ.*

Let us remark that the removable set is only related to the support of the underlying Gibbs measure. The local energy $V(x^m|\varphi \backslash x^m; \theta)$ of any removable point $x^m \in \mathcal{R}_\theta(\varphi)$ can then be defined by the classical expression (3) where Λ comes from Definition 2. In the hereditary case, all the points of φ are removable and we regain the classical definition of the local energy.

The generalization of the Georgii-Nguyen-Zessin formula is the following equation

$$\mathbf{E}\left(\int_{\mathbb{R}^d \times \mathbb{M}} h\left(x^m, \Phi; \theta\right) e^{-V(x^m|\Phi;\theta^\star)} \mu(dx^m)\right) = \mathbf{E}\left(\sum_{x^m \in \mathcal{R}_{\theta^\star}(\Phi)} h\left(x^m, \Phi \setminus x^m; \theta\right)\right). \quad (31)$$

Let us notice that the only difference with the classical formula is that the sum is restricted to the removable points. Now, let us present the consequences of this formula on the Takacs-Fiksel procedure. We have to consider the two following cases:

- When the support of the Gibbs measure does not depend on θ: the set of removable points $\mathcal{R}_\theta(\varphi)$ does not depend on θ either. In this case, the Takacs-Fiksel estimor is defined by (7), and C_Λ is as in (6) to the exception of the sum which is restricted to the removable points:

$$C_\Lambda(\varphi; h, \theta) := \int_{\Lambda \times \mathbb{M}} h(x^m, \varphi; \theta) e^{-V(x^m|\varphi;\theta)} \mu(dx^m) - \sum_{x^m \in \mathcal{R}_{\theta^\star}(\varphi) \cap \Lambda} h(x^m, \varphi \setminus x^m; \theta). \quad (32)$$

  The sum is computable because by assumption, the set $\mathcal{R}_{\theta^\star}(\varphi)$ does not depend on $\theta^\star$. In this situation, with the same assumptions **[C]** and **[N]**, the consistency and the asymptotic normality of the estimator may be proved as in Section 7.

- When the support of the Gibbs measure depends on some parameters $\theta_{hc} = (\theta_1, \ldots, \theta_q)$, $q \leq p$ (called the hardcore parameters): the remaining parameters $\theta_{sm} = (\theta_{q+1}, \ldots, \theta_p)$ are supposed to parameterize the classical (or smooth) interaction between points. The set of removable points $\mathcal{R}_\theta(\varphi)$ therefore depends on θ_{hc} only. The estimation issue is more complicated in this case. Indeed, Assumption **[C2]** requires some regularities of the interaction with respect to the parameter θ, such as continuity, which clearly fail to be true for the support parameter θ_{hc}. The Takacs-Fiksel procedure is therefore unable to estimate θ_{hc}. Note that this problem is not specific to the presence of non-hereditary interactions, but arises as soon as some hardcore parameters have to be estimated. In [16], the authors solve this problem in both the hereditary and non-hereditary setting, by introducing a two-step estimation procedure. We can follow the same strategy here. In a first step, the estimator $\hat{\theta}_{hc}$ of the hardcore parameter is defined in a natural way according to the observed support of the point process (see Section 4.2.1 in [16]). Then, in a second step, the Takacs-Fiksel estimator $\hat{\theta}_{sm}$ is defined by (7) with

$$\begin{aligned} C_\Lambda(\varphi; h, \theta_{sm}) \quad := \quad & \int_{\Lambda \times \mathbb{M}} h(x^m, \varphi; \theta_{sm}, \hat{\theta}_{hc}) e^{-V(x^m|\varphi;\theta_{sm},\hat{\theta}_{hc})} \mu(dx^m) \\ & - \sum_{x^m \in \mathcal{R}_{\hat{\theta}_{hc}}(\varphi) \cap \Lambda} h(x^m, \varphi \setminus x^m; \theta_{sm}, \hat{\theta}_{hc}). \end{aligned} \tag{33}$$

  Let us remark that the estimator $\hat{\theta}_{hc}$ is plugged in the computation of C_Λ. In particular, the removable points are determined with respect to $\hat{\theta}_{hc}$. As in [16], the regularity and integrability assumptions of type **[C]** for $\hat{\theta}_{sm}$ and conditions on the support of the Gibbs measure are required in order to obtain the consistency of $(\hat{\theta}_{hc}, \hat{\theta}_{sm})$. The asymptotic normality is more difficult to obtain and no general results are available. In fact, it seems that there is no hope to expect asymptotic normality without managing the rate of convergence of $\hat{\theta}_{hc}$, which should strongly depend on the model.

7 A short simulation study

The first aim of the present paper is to explore the asymptotic properties of the Takacs-Fiksel estimate. Given a model, the important question of correctly choosing the test functions is an open question and will not be treated here. Also, a complete comparison between all the existing parametric estimation methods has not been attempted. In this section, we just present a brief simulation study to illustrate the methodology and the asymptotic results.

The model considered in this section is the (non-marked) Strauss process described by (11) with two parameters. Table 1 and Figure 2 give an empirical comparison of the approximated maximum likelihood estimate (MLE) described in [26], the maximum pseudo-likelihood estimate (MPLE), which is a particular case of the Takacs-Fiksel estimate, and the explicit Takacs-Fiksel estimate (TFE) described in Section 3.2.3 and in particular in (13). To generate Strauss point processes and to compute the MLE, the `R` package `spatstat` is used. The MPLE and the TFE are computed respectively from (8) and (6) where the integrals were approximated by Monte-Carlo.[1] Strauss point processes are generated on the window $[0,3]^2 \oplus R$ where R is set to 0.05. Recall that this parameter corresponds to the finite range of the Strauss process. Then the estimates are computed on the window $[0,\tau]^2 \oplus R$ for $\tau = 1, 2, 3$, using a R-erosion. We use the parametrization of `spatstat`, i.e. the intensity parameter is $\beta^\star := e^{-\theta_1^\star}$ and the interaction parameter is $\gamma^\star := e^{-\theta_2^\star}$. We set $\beta^\star = 100$ and $\gamma^\star = 0.5$. The estimation of $(\beta^\star, \gamma^\star)$ by the TFE is thus $(\widehat{\beta}, \widehat{\gamma}) = (e^{-\widehat{\theta}_1}, e^{-\widehat{\theta}_2})$ where $(\widehat{\theta}_1, \widehat{\theta}_2)$ are defined in (13).

As expected, the three estimates have decreasing variance when the domain of observation grows. Since the explicit TFE is obtained very quickly, its behavior appears very satisfactory in comparison with the MLE and MPLE.

[1] The MPLE is also implemented in `spatstat` but we did not use it. In fact, it appears that the implementation therein leads to biased estimations.

Estimates of the parameter $\beta^\star = 100$			
W_τ	MLE	MPLE	TFE
$[0,1]^2$	103.50 (18.24)	101.19 (16.92)	101.44 (18.95)
$[0,2]^2$	101.03 (8.04)	100.33 (8.66)	99.99 (8.71)
$[0,3]^2$	100.38 (6.58)	99.55 (5.47)	99.96 (6.27)

Estimates of the parameter $\gamma^\star = 0.5$			
W_τ	MLE	MPLE	TFE
$[0,1]^2$	0.50 (0.17)	0.50 (0.16)	0.52 (0.22)
$[0,2]^2$	0.50 (0.08)	0.50 (0.08)	0.51(0.11)
$[0,3]^2$	0.51 (0.05)	0.49 (0.05)	0.51 (0.06)

Table 1: Empirical means and standard deviations between brackets for the MLE, MPLE and the (explicit) TFE based on $m = 500$ replications of a Strauss point process with parameters $\beta^\star = 100, \gamma^\star = 0.5$, and (known) interaction range $R = 0.05$, generated in the window $[0,3]^2 \oplus R$ and estimated on the window $W_\tau = [0,\tau]^2 \oplus R$ with R-erosion for $\tau = 1, 2, 3$.

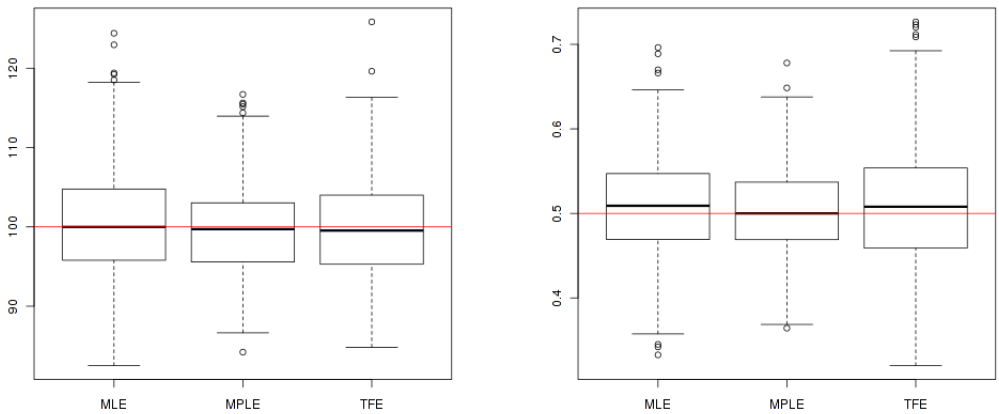

Figure 2: Boxplots of the 3 estimates of $\beta^\star$ (left) and $\gamma^\star$ (right) when $\tau = 3$, based on $m = 500$ replications.

Empirical covariance of $\tau \times (\widehat{\theta_1}, \widehat{\theta_2})^T$			MC approximation of
$\tau = 1$	$\tau = 2$	$\tau = 3$	$(\mathcal{E}^T)^{-1}\underline{\mathbf{\Sigma}}\mathcal{E}^{-1}$
$\begin{pmatrix} 0.037 & -0.051 \\ -0.051 & 0.185 \end{pmatrix}$	$\begin{pmatrix} 0.031 & -0.048 \\ -0.048 & 0.185 \end{pmatrix}$	$\begin{pmatrix} 0.035 & -0.051 \\ -0.051 & 0.179 \end{pmatrix}$	$\begin{pmatrix} 0.034 & -0.045 \\ -0.045 & 0.154 \end{pmatrix}$

Table 2: Comparison between the renormalized empirical covariance matrix of the estimates $(\widehat{\theta_1}, \widehat{\theta_2})$ (based on $m = 500$ replications of the Strauss process on $[0, \tau]^2 \oplus R$ as in Table 1) and the asymptotic covariance matrix (approximated by Monte-Carlo simulations).

Now, we focus more particularly on the explicit TFE given by (13). We wish to assess whether the empirical covariance matrix of $(\widehat{\theta_1}, \widehat{\theta_2})^T$ agrees with the asymptotic covariance matrix deduced from (17) in Theorem 4. For this reason, we come back to the parameterization of the Strauss model of Section 3.2.3. From (17), we deduce that the asymptotic covariance matrix of $\tau \times (\widehat{\theta_1}, \widehat{\theta_2})^T$, as τ tends to $+\infty$, is $\mathcal{E}^{-1}(\mathbf{h}, \theta^\star)^T \underline{\mathbf{\Sigma}}(\mathbf{h}, \theta^\star)\mathcal{E}^{-1}(\mathbf{h}, \theta^\star)$ (a simplification occurs here since $K = p = 2$). Moreover, we can prove that the matrix $\mathcal{E}(\mathbf{h}, \theta^\star)$ reduces to

$$\mathcal{E}(\mathbf{h}, \theta^\star) = |\Delta|^{-1} \begin{pmatrix} \mathbf{E}\left(N_{0,\Delta}\right) & e^{\theta_2^\star}\mathbf{E}\left(N_{1,\Delta}\right) \\ 0 & e^{\theta_2^\star}\mathbf{E}\left(N_{1,\Delta}\right) \end{pmatrix}$$

where Δ is any bounded domain. Since the matrices $\mathcal{E}(\mathbf{h}, \theta^\star)$ and $\underline{\mathbf{\Sigma}}(\mathbf{h}, \theta^\star)$ depend on expectations with respect to $P_\theta^\star$, they are approximated by Monte-Carlo simulations. Table 2 contains the empirical covariance of $\tau \times (\widehat{\theta_1}, \widehat{\theta_2})^T$, for $\tau = 1, 2, 3$ as well as the asymptotic covariance matrix approximated by Monte-Carlo simulations. Finally, in order to appreciate the Gaussian limit, Figure 3 compares the empirical distribution of the two estimates $(\widehat{\theta_1}, \widehat{\theta_2})$ with the expected limit when $\tau = 3$. Note that the same kind of results as in Table 2 and Figure 3 could have been easily obtained for $(\widehat{\beta}, \widehat{\gamma})$ by use of the delta method.

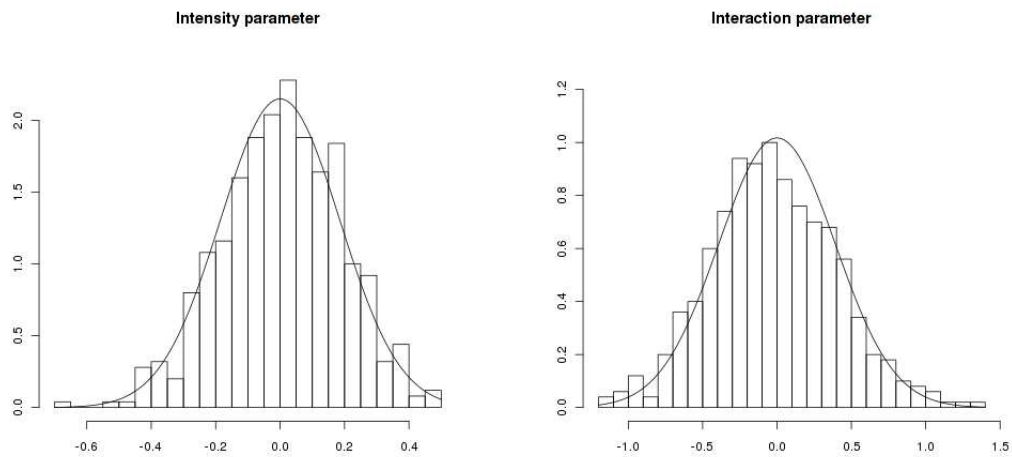

Figure 3: Histogram of $\tau(\widehat{\theta}_j - \theta_j^\star)$ for $j = 1$ (intensity parameter, left) and $j = 2$ (interaction parameter, right), for the large window, i.e. $\tau = 3$. The curves correspond to the densities of the asymptotic Gaussian laws, which are centered with variance .034 (left) and .154 (right) according to Table 2.

8 Proofs of asymptotic results

8.1 Proof of Theorem 3

If there is more than one stationary Gibbs measure, then some non ergodic Gibbs measures automatically exist because, in the convex set of all Gibbs measures, only the extremal measures are ergodic. But any stationary Gibbs measure can be represented as a mixture of ergodic measures ([22], Theorem 14.10). Due to this decomposition, we can assume that $P_{\theta^\star}$ is ergodic to prove the consistency of the Takacs-Fiksel estimate. The proof is split into two steps.

Step 1. U_{Λ_n} is a contrast function. Under **[C1]**, the ergodic Lemma 2 and the GNZ formula given in (5) may be applied to prove that $P_{\theta^\star}$-a.s.

$$\begin{aligned} |\Lambda_n|^{-1} C_{\Lambda_n}(\Phi, h_k; \theta) &\to \mathbf{E}\left(h_k(0^M, \Phi; \theta) e^{-V(0^M|\Phi;\theta)}\right) - \mathbf{E}\left(h_k(0^M, \Phi \setminus 0^M; \theta)\right) \\ &= \mathbf{E}\left(h_k(0^M, \Phi; \theta)\left(e^{-V(0^M|\Phi;\theta)} - e^{-V(0^M|\Phi;\theta^\star)}\right)\right). \end{aligned}$$

Therefore, as $n \to +\infty$, one obtains $P_{\theta^\star}-$a.s.

$$U_{\Lambda_n}(\Phi; \mathbf{h}, \theta) \to U(\theta) := \sum_{k=1}^{K} \mathbf{E}\left(h_k(0^M, \Phi; \theta)\left(e^{-V(0^M|\Phi;\theta)} - e^{-V(0^M|\Phi;\theta^\star)}\right)\right)^2.$$

Note that $U(\theta^\star) = 0$. In addition with Assumptions **[C2]** and **[C3]**, this proves that U_{Λ_n} is a continuous contrast function vanishing only at $\theta^\star$.

Step 2. Modulus of continuity

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

$$W_n(\varphi, \eta) = \sup\left\{\left|U_{\Lambda_n}(\varphi; \mathbf{h}, \theta) - U_{\Lambda_n}(\varphi; \mathbf{h}, \theta')\right| : \theta, \theta' \in \mathbf{\Theta}, ||\theta - \theta'|| \le \eta\right\}.$$

This step aims at proving that there exists a sequence $(\varepsilon_\ell)_{\ell \ge 1}$, with $\varepsilon_\ell \to 0$ as $\ell \to +\infty$ such that for all $\ell \ge 1$

$$P\left(\limsup_{n \to +\infty}\left(W_n\left(\Phi, \frac{1}{\ell}\right) \ge \varepsilon_\ell\right)\right) = 0. \tag{34}$$

Let $\theta, \theta' \in \mathbf{\Theta}$, then

$$\begin{aligned} |U_{\Lambda_n}(\varphi; \mathbf{h}, \theta) - U_{\Lambda_n}(\varphi; \mathbf{h}, \theta')| \le \sum_{k=1}^{K} \Big\{ &|\Lambda_n|^{-1} |C_{\Lambda_n}(\varphi; h_k, \theta) - C_{\Lambda_n}(\varphi; h_k, \theta')| \times \\ &|\Lambda_n|^{-1}\big(|C_{\Lambda_n}(\varphi; h_k, \theta)| + |C_{\Lambda_n}(\varphi; h_k, \theta')|\big)\Big\}. \end{aligned} \tag{35}$$

Under Assumption **[C4]**, there exists $n_0 \ge 1$ such that for all $n \ge n_0$ we have for $P_{\theta^\star}-$a.e. φ

$$\begin{aligned} |\Lambda_n|^{-1}\left(|C_{\Lambda_n}(\varphi; h_k, \theta)| + |C_{\Lambda_n}(\varphi; h_k, \theta')|\right) &\le 2|\Lambda_n|^{-1} \int_{\Lambda_n \times \mathbb{M}} \max_{\theta \in \mathbf{\Theta}} |h_k(x^m, \varphi; \theta)| e^{-V(x^m|\varphi;\theta)} \mu(dx^m) \\ &\quad + 2|\Lambda_n|^{-1} \sum_{x^m \in \varphi_{\Lambda_n}} \max_{\theta \in \mathbf{\Theta}} |h_k(x^m, \varphi \setminus x^m; \theta)| \\ &\le \gamma_1, \end{aligned}$$

where

$$\gamma_1 := 4 \times \max_{k=1,\dots,K}\left(\mathbf{E}\left(\max_{\theta \in \mathbf{\Theta}} |f_k(0^M, \Phi; \theta)|\right) + \mathbf{E}\left(\max_{\theta \in \mathbf{\Theta}} |h_k(0^M, \Phi; \theta)| e^{-V(0^M|\Phi;\theta^\star)}\right)\right).$$

Therefore for all $n \ge n_0$

$$|U_{\Lambda_n}(\varphi; \mathbf{h}, \theta) - U_{\Lambda_n}(\varphi; \mathbf{h}, \theta')| \le \gamma_1 \times \frac{1}{|\Lambda_n|} \sum_{k=1}^{K} \left(A_{\Lambda_n}(\varphi; h_k, \theta, \theta') + B_{\Lambda_n}(\varphi; h_k, \theta, \theta')\right),$$

where

$$\begin{aligned} A_{\Lambda_n}(\varphi; h_k, \theta, \theta') &:= \int_{\Lambda_n \times \mathbb{M}} |f_k(x^m, \varphi; \theta) - f_k(x^m, \varphi; \theta')| \mu(dx^m) \\ B_{\Lambda_n}(\varphi; h_k, \theta, \theta') &= \sum_{x^m \in \varphi_{\Lambda_n}} |h_k(x^m, \varphi \setminus x^m; \theta) - h_k(x^m, \varphi \setminus x^m; \theta')|. \end{aligned}$$

Under Assumption **[C4]**, one may apply the mean value theorem in $\mathbb{R}^p$ as follows: there exist $\xi^{(1)}, \ldots, \xi^{(p)} \in [\min(\theta_1, \theta_1'), \max(\theta_1, \theta_1')] \times \ldots \times \left[\min(\theta_p, \theta_p'), \max(\theta_p, \theta_p')\right]$ such that for all $\varphi \in \Omega$

$$\begin{aligned} A_{\Lambda_n}(\varphi; h_k, \theta, \theta') &= \int_{\Lambda_n \times \mathbb{M}} \sum_{j=1}^{p} (\theta_j - \theta_j')(\mathbf{f_k}^{(1)}(x^m, \varphi; \xi_j^{(1)}))_j \; \mu(dx^m) \\ &\leq p \times \|\theta - \theta'\| \int_{\Lambda_n \times \mathbb{M}} \max_{\theta \in \mathbf{\Theta}} \|\mathbf{f_k}^{(1)}(x^m, \varphi; \theta)\| \; \mu(dx^m) \end{aligned}$$

In a similar way, one may prove that for $P_{\theta^\star}-$a.e. φ

$$B_{\Lambda_n}(\varphi; h_k, \theta, \theta') \leq p \times \|\theta - \theta'\| \sum_{x^m \in \varphi_{\Lambda_n}} \max_{\theta \in \mathbf{\Theta}} \|\mathbf{h_k}^{(1)}(x^m, \varphi \setminus x^m; \theta)\|.$$

Under Assumption **[C4]**, there exists $n_1(k) \geq 1$ such that for all $n \geq n_1(k)$, we have for $P_\theta^\star-$a.e. φ

$$\frac{1}{|\Lambda_n|}\big(A_{\Lambda_n}(\varphi; h_k, \theta, \theta') + B_{\Lambda_n}(\varphi; h_k, \theta, \theta')\big) \leq \gamma_2 \|\theta - \theta'\|$$

where

$$\gamma_2 := 2p \times \max_{k=1,\ldots,K} \left(\mathbf{E}\left(\max_{\theta \in \mathbf{\Theta}} \|\mathbf{f_k}^{(1)}(0^M, \Phi; \theta)\| \right) + \mathbf{E}\left(\max_{\theta \in \mathbf{\Theta}} \|\mathbf{h_k}^{(1)}(0^M, \Phi; \theta)\| e^{-V(0^M|\Phi;\theta^\star)} \right) \right).$$

We finally obtain the following upper-bound for $P_{\theta^\star}-$a.e. φ, for all θ, θ' such that $\|\theta - \theta'\| \leq 1/\ell$ and for all $n \geq N = \max(n_0, \max_k n_1(k))$

$$|U_{\Lambda_n}(\varphi; \mathbf{h}, \theta) - U_{\Lambda_n}(\varphi; \mathbf{h}, \theta')| \leq \gamma \times \frac{1}{\ell},$$

with $\gamma = K \times \gamma_1 \times \gamma_2$ and therefore $W_n(\varphi, 1/\ell) \leq \gamma \times \frac{1}{\ell}$. Finally, since

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

for $P_{\theta^\star}-$a.e. φ, the expected result (34) is proved.
Conclusion step. Steps 1 and 2 ensure the fact that we can apply Theorem 3.4.3 of [23] which asserts the almost sure convergence for minimum contrast estimators.

8.2 Proof of Theorem 4

The proof is based on a classical result concerning asymptotic normality for minimum contrast estimators *e.g.* Theorem 3.4.5 of [23]. We split it into two different steps.

Step 1. Asymptotic normality of $\mathbf{U}_{\Lambda_n}^{(1)}(\Phi; \mathbf{h}, \theta^\star)$.

We start with a Lemma which states a central limit theorem. Its proof uses a conditional centering assumption as in [28], which holds due to the particular form of the contrast function (6). This scheme of proof is now well-known and we mainly refer to previous works for the technical details.

Lemma 10 *Under the Assumptions **[N1]**, **[N2]** and **[N3]**, the following convergence holds in distribution, as $n \to +\infty$*

$$|\Lambda_n|^{-1/2} \; \widetilde{\mathbf{C}}_{\Lambda_n}(\Phi; \mathbf{h}, \theta^\star) \xrightarrow{d} \mathcal{N}(0, \underline{\mathbf{\Sigma}}(\mathbf{h}, \theta^\star)), \tag{36}$$

where $\underline{\mathbf{\Sigma}}(\mathbf{h}, \theta^\star)$ and $\widetilde{\mathbf{C}}_{\Lambda_n}(\varphi; \mathbf{h}, \theta^\star)$ are defined in Theorem 4.

Proof. The vector $\widetilde{\mathbf{C}}_{\Lambda_n}(\varphi;\mathbf{h},\theta^\star)$ (of length K) corresponds to the vector of the h_k-residuals for $k=1,\ldots,K$ computed on the same domain Λ_n with $\widehat{\theta}=\theta^\star$, see [2] for a definition and practical study of this concept of residuals. The asymptotic behavior of the residuals process has been investigated in [12]. In particular, with the notation of the present paper, the asymptotic normality of the vector $\widetilde{\mathbf{C}}_{\Lambda_n}(\Phi;\mathbf{h},\widehat{\theta})$ for general $\widehat{\theta}$ is obtained (see Proposition 4 in [12]). When $\widehat{\theta}=\theta^\star$, the assumptions and the asymptotic covariance matrix of Proposition 4 in [12] respectively reduce to **[N1-3]** and (18). ■

Now, according to the definition of $U_{\Lambda_n}(\varphi;\mathbf{h},\theta^\star)$, we have

$$\mathbf{U}^{(1)}_{\Lambda_n}(\varphi;\mathbf{h},\theta^\star)=2|\Lambda_n|^{-2}\sum_{k=1}^{K}\mathbf{C}^{(1)}_{\Lambda_n}(\varphi;h_k,\theta^\star)C_{\Lambda_n}(\varphi;h_k,\theta^\star).$$

Under Assumption **[C4]**, one may apply the ergodic Lemma 2, in order to derive $P^\star_\theta-$a.s., as $n\to+\infty$

$$|\Lambda_n|^{-1}\mathbf{C}^{(1)}_{\Lambda_n}(\Phi;h_k,\theta^\star)\to\mathbf{E}\left(\mathbf{f_k}^{(1)}(0^M,\Phi;\theta^\star)-\mathbf{h_k}^{(1)}(0^M,\Phi;\theta^\star)e^{-V(0^M|\Phi;\theta^\star)}\right). \tag{37}$$

It is easily checked that this expectation reduces to $-\mathcal{E}(h_k,\theta^\star)$, this vector of length p being defined by $\mathcal{E}(h_k,\theta^\star):=\mathbf{E}\left(h_k(0^M,\Phi;\theta^\star)\mathbf{V}^{(1)}(0^M|\Phi;\theta^\star)e^{-V(0^M|\Phi;\theta^\star)}\right)$. Let us denote by $\underline{\mathcal{E}}(\mathbf{h},\theta^\star)$ the (p,K) matrix $(\mathcal{E}(h_1,\theta^\star),\ldots,\mathcal{E}(h_K,\theta^\star))$, then we get the following decomposition

$$\begin{aligned}|\Lambda_n|^{1/2}\mathbf{U}^{(1)}_{\Lambda_n}(\Phi;\mathbf{h},\theta^\star) &= 2|\Lambda_n|^{1/2}\times|\Lambda_n|^{-2}\sum_{k=1}^{K}\mathbf{C}^{(1)}_{\Lambda_n}(\Phi;h_k,\theta^\star)C_{\Lambda_n}(\Phi;h_k,\theta^\star)\\ &= -2|\Lambda_n|^{-1/2}\underline{\mathcal{E}}(\mathbf{h},\theta^\star)\widetilde{\mathbf{C}}_{\Lambda_n}(\Phi;\mathbf{h},\theta^\star)\\ &\quad +2\sum_{k=1}^{K}\left[|\Lambda_n|^{-1}\mathbf{C}^{(1)}_{\Lambda_n}(\Phi;h_k,\theta^\star)-(-\mathcal{E}(h_k,\theta^\star))\right]|\Lambda_n|^{-1/2}C_{\Lambda_n}(\Phi;h_k,\theta^\star).\end{aligned}$$

According to (37) and Lemma 10, Slutsky's Theorem implies that for any $k\in\{1,\ldots,K\}$,

$$\left(|\Lambda_n|^{-1}\mathbf{C}^{(1)}_{\Lambda_n}(\Phi;h_k,\theta^\star)-\mathcal{E}(h_k,\theta^\star)\right)|\Lambda_n|^{-1/2}C_{\Lambda_n}(\Phi;h_k,\theta^\star)\overset{P}{\to}0,$$

as $n\to+\infty$, the zero here being a vector of length p. Using again Lemma 10, we finally reach the following convergence in distribution as $n\to+\infty$

$$|\Lambda_n|^{1/2}\mathbf{U}^{(1)}_{\Lambda_n}(\Phi;\mathbf{h},\theta^\star)\overset{d}{\to}\mathcal{N}\left(0,4\underline{\mathcal{E}}(\mathbf{h},\theta^\star)\,\underline{\mathbf{\Sigma}}(\mathbf{h},\theta^\star)\,\underline{\mathcal{E}}(\mathbf{h},\theta^\star)^T\right),$$

where $\underline{\mathbf{\Sigma}}(\mathbf{h},\theta^\star)$ is defined by (18).

Step 2. Convergence of $\underline{\mathbf{U}}^{(2)}_{\Lambda_n}(\Phi;\mathbf{h},\theta)$ for $\theta\in\mathcal{V}(\theta^\star)$

According to our definition and Assumption **[N4]**, the (p,p) matrix $\underline{\mathbf{U}}^{(2)}_{\Lambda_n}(\varphi;\mathbf{h},\theta)$ is defined for $i,j=1,\ldots,p$ by

$$\left(\underline{\mathbf{U}}^{(2)}_{\Lambda_n}(\varphi;\mathbf{h},\theta^\star)\right)_{ij}=2|\Lambda_n|^{-2}\sum_{k=1}^{K}\left\{\left(\underline{\mathbf{C}}^{(2)}_{\Lambda_n}(\varphi;h_k;\theta)\right)_{ij}C_{\Lambda_n}(\varphi;h_k,\theta)+\left(\mathbf{C}^{(1)}_{\Lambda_n}(\varphi;h_k,\theta)\right)_i\left(\mathbf{C}^{(1)}_{\Lambda_n}(\varphi;h_k,\theta)\right)_j\right\}.$$

Note also that $\mathbf{C}^{(1)}_{\Lambda_n}(\varphi;h_k,\theta)$ and $\underline{\mathbf{C}}^{(2)}_{\Lambda_n}(\varphi;h_k,\theta)$ are defined by

$$\begin{aligned}\mathbf{C}^{(1)}_{\Lambda_n}(\varphi;h_k,\theta) &= \int_{\Lambda_n\times\mathbb{M}}\mathbf{f}^{(1)}_k(x^m,\varphi;\theta)\mu(dx^m)-\sum_{x^m\in\varphi_{\Lambda_n}}\mathbf{h}^{(1)}_k(x^m,\varphi\setminus x^m;\theta)\\ \underline{\mathbf{C}}^{(2)}_{\Lambda_n}(\varphi;h_k,\theta) &= \int_{\Lambda_n\times\mathbb{M}}\underline{\mathbf{f}}^{(2)}_k(x^m,\varphi;\theta)\mu(dx^m)-\sum_{x^m\in\varphi_{\Lambda_n}}\underline{\mathbf{h}}^{(2)}_k(x^m,\varphi\setminus x^m;\theta).\end{aligned}$$

Under Assumption **[N4]**, then for all $i, j = 1, \ldots, p$ and for any $k = 1, \ldots, K$, each normalized term $|\Lambda_n|^{-1} C_{\Lambda_n}(\Phi; h_k, \theta)$, $|\Lambda_n|^{-1} \mathbf{C}^{(1)}_{\Lambda_n}(\Phi; h_k, \theta)$ and $|\Lambda_n|^{-1} \underline{\mathbf{C}}^{(2)}_{\Lambda_n}(\Phi; h_k, \theta)$ satisfies the assumptions of the ergodic Lemma 2. Therefore, for any $\theta \in \mathcal{V}(\theta^\star)$, there exists a matrix $\underline{\mathbf{U}}^{(2)}(\mathbf{h}, \theta)$ such that $P^\star_\theta-$a.s.

$$\underline{\mathbf{U}}^{(2)}_{\Lambda_n}(\Phi; \mathbf{h}, \theta) \to \underline{\mathbf{U}}^{(2)}(\mathbf{h}, \theta).$$

This justifies that, for n large enough, in a neighborhood of $\theta^\star$, $\left(\underline{\mathbf{U}}^{(2)}_{\Lambda_n}(\varphi; \mathbf{h}, \theta)\right)_{ij}$ is uniformly bounded by $2 \times \max_{\theta \in \mathcal{V}(\theta^\star)} \left|\left(\underline{\mathbf{U}}^{(2)}(\mathbf{h}, \theta)\right)_{ij}\right|$ for $P^\star_\theta-$a.e. φ. When $\theta = \theta^\star$, recall, from (5), that $|\Lambda_n|^{-1} C_{\Lambda_n}(\Phi; h_k, \theta^\star)$ converges almost surely to zero and that (37) holds. Hence, $\underline{\mathbf{U}}^{(2)}(\mathbf{h}, \theta^\star)$ reduces to $2\underline{\mathcal{E}}(\mathbf{h}, \theta^\star)\underline{\mathcal{E}}(\mathbf{h}, \theta^\star)^T$.

Conclusion Step. From Theorem 3.4.5 of [23], Steps 1 and 2 ensure that the normalized difference $|\Lambda_n|^{1/2} \left(\underline{\mathbf{U}}^{(2)}(\mathbf{h}, \theta^\star) \left(\widehat{\theta}_n(\Phi) - \theta^\star\right) - \mathbf{U}^{(1)}_{\Lambda_n}(\Phi; \mathbf{h}, \theta^\star)\right)$ converges in probability to 0, which is the expected result.

Acknowledgements

We are grateful to the associate editor, the anonymous referee and to Hans Zessin for their valuable comments and advices.

References

[1] 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.

[2] A. Baddeley, R. Turner, J. Møller, and M. Hazelton. Residual analysis for spatial point processes. *Journal of the Royal Statistical Society (series B)*, 67:1–35, 2005.

[3] A. J. Baddeley and M. N. M. Van Lieshout. Area-interaction point processes. *Ann. Inst. Statist. Math.*, 47(4):601–619, 1995.

[4] K.K. Berthelsen and J. Møller. Likelihood and non-parametric Bayesian MCMC inference for spatial point processes based on perfect simulation and path sampling. *Scandinavian Journal of Statistics*, 30(3):549–564, 2003.

[5] E. Bertin, J.-M. Billiot, and R. Drouilhet. Existence of "nearest-neighbour" spatial Gibbs models. *Adv. Appli. Prob.*, 31:895–909, 1999.

[6] E. Bertin, J.-M. Billiot, and R. Drouilhet. R-local Delaunay inhibition model. *Journal of Statistical Physics*, 132(4):649–667, 2008.

[7] J. Besag. Spatial interaction and the statistical analysis of lattice system. *Journal of the Royal Statistical Society (series B)*, 26:192–236, 1974.

[8] J.-M. Billiot. Asymptotic properties of Takacs-Fiksel estimation method for Gibbs point processes. *Statistics*, 30(1):69–89, 1997.

[9] 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.

[10] D. Brillinger. Statistical inference for stationary point processes. *Stochastic processes and related topics, vol 1. 55-99 Ed. Madan Lal Puri. Academic Press, New-York-London*, 1975.

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

[12] J.-F. Coeurjolly and F. Lavancier. Residuals for stationary marked Gibbs point processes. *submitted*, 2010. `http://arxiv.org/abs/1002.0857`.

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

[14] D. Dereudre. The existence of quermass-interaction processes for nonlocally stable interaction and nonbounded convex grains. *Adv. Appli. Prob.*, 41(3):664–681, 2009.

[15] D. Dereudre, R. Drouilhet, and H.O. Georgii. Existence of Gibbsian point processes with geometry-dependent interactions. *to appear in Prob. Theory and Related Fields*, 2010.

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

[17] D. Dereudre and F. Lavancier. Practical simulation and estimation for Gibbs Delaunay-Voronoï tessellations with geometric hardcore interaction. *Computational Stat. and Data Analysis*, 55:498–519, 2011.

[18] P. J. Diggle, T. Fiksel, P. Grabarnik, Y. Ogata, D. Stoyan, and M. Tanemura. On parameter estimation for pairwise interaction point processes. *International Statistical Review*, 62(1):99–117, 1994.

[19] T. Fiksel. Estimation of parameterized pair potentials of marked and non-marked Gibbsian point processes. *Elektron. Inform. Kybernet.*, 20:270–278, 1984.

[20] T. Fiksel. Estimation of interaction potentials of Gibbsian point processes. *Math. Operationsf. Statist. Ser. Statist.*, 19:77–86, 1988.

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

[22] H.O. Georgii. *Gibbs measures and Phase Transitions*. de Gruyter, Berlin, 1988.

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

[24] X. Guyon and H.R. Künsch. *Asymptotic comparison of estimators in the Ising model*. Lecture notes in Statistics 74. Springer, Berlin, 1992.

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

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

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

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

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

[30] W. S. Kendall, M. N. M. Van Lieshout, and A. J. Baddeley. Quermass-interaction processes conditions for stability. *Adv. Appli. Prob.*, 31:315–342, 1999.

[31] K. Krickeberg. Statistical problems of point processes. *Rosz. Pol. Tow. Mat., Ser. III, Mat. Stosow. 13, 29-57*, 1978.

[32] E.L. Lehman. *Elements of large-sample theorey.* Springer, 1999.

[33] S. Mase. Uniform LAN condition of planar Gibbsian point processes and optimality of maximum likelihood estimators of soft-core potential function. *Probab. Theory and Related Fields*, 92:51–67, 1992.

[34] K Matthes, J. Kerstan, and J. Mecke. *Infinitely Divisible Point Processes.* J. Wiley, 1978.

[35] J. Møller and K. Helisová. Power diagrams and interaction processes for union of discs. *Adv. Appli. Prob.*, 40(2):321–347, 2008.

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

[37] J. Møller and R. Waagepetersen. Modern statistics for spatial point processes. *Scan. J. Stat.*, 34:643–684, 2007.

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

[39] X.X. Nguyen and H. Zessin. Ergodic theorems for spatial processes. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 48:133–158, 1979.

[40] F. Papangelou. The Armenian Connection: Reminiscences from a time of interactions. *Journal of contemporary mathematical analysis*, 44(1):14–19, 2009.

[41] C.J. Preston. *Random fields.* Springer Verlag, 1976.

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

[43] D. Ruelle. Superstable interactions in classical statistical mechanics. *Comm. Math. Phys.*, 1970.

[44] D. Stoyan, W.S. Kendall, and J. Mecke. *Stochastic geometry and its applications.* John Wiley and Sons, Chichester, 1995.

[45] R. Takacs. Estimator for the pair-potential of a Gibbsian point process. *Math. Operationsf. Statist. Ser. Statist.*, 17:429–433, 1986.

[46] H. Zessin. Der papangelou prozess. *Journal of contemporary mathematical analysis*, 44(1):36–44, 2009.