

## CONSISTENCY OF LIKELIHOOD ESTIMATION FOR GIBBS POINT PROCESSES<sup>1</sup>

BY DAVID DEREUDRE\* AND FRÉDÉRIC LAVANCIER<sup>†,‡</sup>

*University of Lille I\**, *University of Nantes<sup>†</sup>* and *Inria<sup>‡</sup>*

Strong consistency of the maximum likelihood estimator (MLE) for parametric Gibbs point process models is established. The setting is very general. It includes pairwise pair potentials, finite and infinite multibody interactions and geometrical interactions, where the range can be finite or infinite. The Gibbs interaction may depend linearly or nonlinearly on the parameters, a particular case being hardcore parameters and interaction range parameters. As important examples, we deduce the consistency of the MLE for all parameters of the Strauss model, the hardcore Strauss model, the Lennard–Jones model and the area-interaction model.

**1. Introduction.** Gibbs point processes are popular and widely used models in spatial statistics to describe the repartition of points or geometrical structures in space. They initially arose from statistical physics where they are models for interacting continuum particles; see, for instance, [23]. They are now used in as different domains as astronomy, biology, computer science, ecology, forestry, image analysis and materials science. The main reason is that Gibbs point processes provide a clear interpretation of the interactions between the points, such as attraction or repulsion depending on their relative position. We refer to [2, 13, 20] and [26] for classical text books on spatial statistics and stochastic geometry, including examples and applications of Gibbs point processes.

Assuming a parametric form of the Gibbs interaction, the natural method to estimate the parameters is likelihood inference. A practical issue, however, is that the likelihood depends on an intractable normalizing constant, called the partition function in the statistical physics literature, that has to be approximated. Some sparse data approximations were first proposed in the 1980s, for example, in [21], before simulation-based methods have been developed [11]. A comparative simulation study carried out in [19] demonstrates that Monte-Carlo approximation of the partition function provides the best results in practice. To avoid the latter approximation, other estimation methods have been introduced, including pseudo-likelihood and moments based methods; see the books cited above. With modern computers, the Monte-Carlo approximation of the partition function is no longer

---

Received April 2015; revised March 2016.

<sup>1</sup>Supported in part by the Labex CEMPI (ANR-11-LABX-0007-01).

*MSC2010 subject classifications.* 62F12, 62M30, 60G55, 60D05.

*Key words and phrases.* Parametric estimation, variational principle, Strauss model, Lennard–Jones model, area-interaction model.

an important issue and maximum likelihood estimation for spatial data is feasible and widely used in practice.

From a theoretical point of view, very few is known about the asymptotic properties of the maximum likelihood estimator (MLE) for Gibbs point processes. The asymptotic here means that the window containing the point pattern increases to the whole space  $\mathbb{R}^d$ . It is commonly believed that the MLE is consistent and more efficient, at least asymptotically, than the other estimation methods. This conviction has been supported by several simulation studies, ever since [21]; see also [6]. The present work is concerned with the basic question of consistency. The latter is conjectured to hold in a very general setting but no proof was so far available in the continuous case of Gibbs point processes on  $\mathbb{R}^d$ . This is in contrast with the discrete case of Gibbs interactions on a lattice, where consistency is established for most standard parametric models, regardless of the occurrence of phase transition (when the Gibbs measure is not unique); see [16] and [12]. In fact, the continuous case is more challenging in that most parametric Gibbs models involve irregular parameters as hardcore parameters (controlling the support of the measure) or interaction range parameters. These specificities result in a discontinuous likelihood contrast function, even asymptotically, and some further technical difficulties (for instance the true unknown Gibbs measure is not absolutely continuous with respect to the Gibbs measure associated to an estimation of the hardcore parameter). In an unpublished manuscript [18], Mase addressed the consistency of the MLE for superstable and regular pairwise interactions (a formal definition will be given later). His main tool was the variational principle for Gibbs processes, following the initial idea developed in the discrete case in [16]. He restricted his study to pairwise interactions that are linear in their parameters, which yields a convex contrast function. His result does not imply any restriction on the parameter space, thus including the possibility of phase transition, but the parametric interactions considered in [18] remain nonetheless rather restrictive and do not include hardcore or interaction range parameters.

We prove in this paper that the MLE is strongly consistent for a wide class of stationary Gibbs interactions, without any major restriction on the parameter space. Our assumptions include finite and infinite-body interactions with finite or infinite range, and we do not assume any continuity with respect to the parameters. Our result covers in particular the consistency of the MLE of all parameters of the Strauss model (including the range of interaction), the hardcore Strauss model (including the hardcore parameter), the Lennard–Jones model (including the exponent parameters) and the area-interaction process (including the radius of balls). An important ingredient of the proof is the variational principle, as in [16] and [18], which guarantees the identifiability of the parameters. Our original contribution is the formulation of a minimum contrast result in the presence of discontinuities and, in order to apply it, new controls of physical quantities as the pressure and the mean energy of the Gibbs measure with respect to the parameters.

Beyond consistency, the next natural question concerns the asymptotic distribution of the MLE. This problem is even more arduous and we do not address it in the present paper. Nonetheless, let us briefly mention the state of the art on this question. In 1992, Mase [17] proved that the MLE of regular parameters in a certain class of Gibbs interactions is uniform locally asymptotic normal. However, his result, relying on strong cluster estimates, is established for Gibbs measure generating very sparse point patterns, which implies restrictive conditions on the parameter space. In the same period, Jensen [14] proved the asymptotic normality of the MLE under the Dobrushin uniqueness region, where the Gibbs process satisfies mixing properties. Here again, as noticed in [11], this assumption implies strong restrictions on the parameter space. Without these conditions, phase transition may occur and some long-range dependence phenomena can appear. The MLE might then exhibit a nonstandard asymptotic behaviour, in the sense that the rate of convergence might differ from the standard square root of the size of the window and the limiting law might be non-Gaussian. To our knowledge, the only well understood example is the estimation of the inverse temperature in the Ising model on a square lattice studied in [22]. For this example, the asymptotic law of the MLE is always Gaussian and the rate of convergence is standard except at the critical temperature where it is faster. Generalizing this result to other Gibbs models, especially in the continuous case, is hazardous. The main reason is that the occurrence of phase transition is in general not well understood, and this phenomenon can be of very different nature depending on the model. Characterizing the asymptotic distribution of the MLE in a general setting of Gibbs models still remains a challenging open question.

The remainder of the paper is organized as follows. Parametric Gibbs point processes and the MLE procedure are described in Section 2. Section 3 contains our main result, namely minimal conditions on the Gibbs interaction to ensure strong consistency of the MLE. While the latter result is established in a very general setting, we present in Section 4 standard families of models where our main result applies. Specifically, we deal with finite range pair potentials with or without hardcore (including the Strauss model and the hardcore Strauss model), infinite range pair potentials (including the Lennard–Jones model) and infinite-body interactions (including the area-interaction model). Section 5 contains the proof of our main result and the Supplementary Material [5] contains the other proofs.

## 2. Gibbs point processes and the MLE.

*2.1. State space, reference measure and notation.* We consider the continuous space  $\mathbb{R}^d$  of arbitrary dimension  $d \geq 1$ . The Lebesgue measure on  $\mathbb{R}^d$  is denoted by  $\lambda^d$  and the symbol  $\Lambda$  will always refer to a bounded Borel subset of  $\mathbb{R}^d$ . For  $x \in \mathbb{R}^d$ ,  $|x|$  denotes the Euclidean norm of  $x$  while for  $\Lambda \subset \mathbb{R}^d$ ,  $|\Lambda| := \lambda^d(\Lambda)$ .

A *configuration* is a subset  $\omega$  of  $\mathbb{R}^d$  which is locally finite, meaning that  $\omega \cap \Lambda$  has finite cardinality  $N_\Lambda(\omega) := \#(\omega \cap \Lambda)$  for every bounded Borel set  $\Lambda$ . The space  $\Omega$  of all configurations is equipped with the  $\sigma$ -algebra  $\mathcal{F}$  generated by the counting variables  $N_\Lambda$ .

For convenience, we will often write  $\omega_\Lambda$  in place of  $\omega \cap \Lambda$  and  $\omega_{\Lambda^c}$  for  $\omega \cap \Lambda^c = \omega \setminus \omega_\Lambda$ . Similarly, for every  $\omega$  and every  $x$  in  $\omega$  we abbreviate  $\omega \cup \{x\}$  to  $\omega \cup x$  and  $\omega \setminus \{x\}$  to  $\omega \setminus x$ .

As usual, we take the reference measure on  $(\Omega, \mathcal{F})$  to be the distribution  $\pi$  of the Poisson point process with intensity measure  $\lambda^d$  on  $\mathbb{R}^d$ . Recall that  $\pi$  is the unique probability measure on  $(\Omega, \mathcal{F})$  such that the following hold for all subsets  $\Lambda$ : (i)  $N_\Lambda$  is Poisson distributed with parameter  $\lambda^d(\Lambda)$ , and (ii) given  $N_\Lambda = n$ , the  $n$  points in  $\Lambda$  are independent with uniform distribution on  $\Lambda$ . The Poisson point process restricted to  $\Lambda$  will be denoted by  $\pi_\Lambda$ .

Translation by a vector  $u \in \mathbb{R}^d$  is denoted by  $\tau_u$ , either acting on  $\mathbb{R}^d$  or on  $\Omega$ . A probability  $P$  on  $\Omega$  is said stationary if  $P = P \circ \tau_u^{-1}$  for any  $u$  in  $\mathbb{R}^d$ . In this paper, we consider only stationary probability measures  $P$  with finite intensity  $E_P(N_{[0,1]^d})$ , where  $E_P$  stands for the expectation with respect to  $P$ . We denote by  $\mathcal{P}$  the space of such probability measures.

We denote by  $\Delta_0, \mathcal{I}_n$  and  $\Lambda_n$  the following sets:

$$\begin{aligned} \Delta_0 &= [0, 1)^d, & \mathcal{I}_n &= \{-n, -n + 1, \dots, n - 1\}^d, \\ \Lambda_n &= \bigcup_{k \in \mathcal{I}_n} \tau_k(\Delta_0) = [-n, n)^d. \end{aligned}$$

In the following, some infinite range interaction processes will be considered. To ensure their existence, we must restrict the set of configurations to the so-called set of tempered configurations as in [25]. The definition of the latter may depend on the type of interactions at hand. In the case of superstable pairwise interactions, it is simply defined by

$$\Omega_T = \left\{ \omega \in \Omega; \exists t > 0, \forall n \geq 1, \sum_{i \in \mathcal{I}_n} N_{\tau_i(\Delta_0)}^2(\omega) \leq t(2n)^d \right\}.$$

Unless specified otherwise, this is the definition we consider in the sequel. From the ergodic theorem, any second-order stationary measure on  $\Omega$  is supported on  $\Omega_T$ , so the restriction to  $\Omega_T$  is a mild assumption in practice.

Several other notation are introduced throughout the next sections. For convenience to the reader, we summarize the most important of them below:

- $\Omega, \Omega_T$  Space of configurations and tempered configurations.
- $\omega_\Lambda, \omega_{\Lambda^c}$  Configuration  $\omega$  inside  $\Lambda$ , outside  $\Lambda$  respectively.
- $\delta, \theta$  Hardcore ( $\delta$ ) and other ( $\theta$ ) parameters of the interaction.
- $I, \Theta, \mathcal{K}$   $\delta \in I = [\delta_{\min}, \delta_{\max}]$ ,  $\theta \in \Theta$  and  $\mathcal{K} \subset \Theta$  is compact.
- $\Omega_\infty^\delta$  Space of configurations with locally finite energy.
- $H_\Lambda^\theta(\omega)$  Energy (or Hamiltonian) of  $\omega$  in  $\Lambda$ .

- $f_{\Lambda}^{\delta, \theta}(\omega)$  Conditional density of  $\omega$  inside  $\Lambda$ , given  $\omega_{\Lambda^c}$ , see (7).
- $Z_{\Lambda}^{\delta, \theta}$  Partition function, see (11).
- $\mathcal{G}^{\delta, \theta}$  Set of stationary Gibbs measures associated to  $f_{\Lambda}^{\delta, \theta}(\cdot)$ .
- $\mathcal{G}$   $\mathcal{G} = \bigcup_{\delta \in I, \theta \in \Theta} \mathcal{G}^{\delta, \theta}$
- $\omega^*$  Configuration for the true unknown parameters  $\delta^*$  and  $\theta^*$ .
- $K_n^{\delta, \theta}(\omega_{\Lambda_n}^*)$  Likelihood contrast function on  $\Lambda_n$ , see (9).
- $H^{\theta}(P)$  Mean energy of  $P \in \mathcal{G}$ , see (15).
- $\mathcal{I}(P)$  Specific entropy of  $P \in \mathcal{G}$ , see (18).
- $p(\delta, \theta)$  Pressure associated to  $P^{\delta, \theta} \in \mathcal{G}^{\delta, \theta}$ , see (19).

2.2. *Gibbs point processes.* From a general point of view, a family of interaction energies is a collection  $\mathcal{H} = (H_{\Lambda})$ , indexed by bounded Borel sets  $\Lambda$ , of measurable functions from  $\Omega_T$  to  $\mathbb{R} \cup \{+\infty\}$  such that for every  $\Lambda \subset \Lambda'$ , there exists a measurable function  $\varphi_{\Lambda, \Lambda'}$  from  $\Omega_T$  to  $\mathbb{R} \cup \{+\infty\}$  and for every  $\omega \in \Omega_T$

$$(1) \quad H_{\Lambda'}(\omega) = H_{\Lambda}(\omega) + \varphi_{\Lambda, \Lambda'}(\omega_{\Lambda^c}).$$

This decomposition is equivalent to (6.11) and (6.12) in [23], page 92. In physical terms,  $H_{\Lambda}(\omega) = H_{\Lambda}(\omega_{\Lambda} \cup \omega_{\Lambda^c})$  represents the potential energy of the configuration  $\omega_{\Lambda}$  inside  $\Lambda$  given the configuration  $\omega_{\Lambda^c}$  outside  $\Lambda$ . In words, (1) is just a compatibility property stating that the difference between the energy on  $\Lambda'$  and the energy on a subset  $\Lambda$  only depends on the exterior configuration  $\omega_{\Lambda^c}$ .

As explained in the Introduction, we aim at including a large class of interactions in our study, to cover for instance infinite-body interactions like the area-interaction process considered in Section 4.4. Nevertheless, most of standard parametric Gibbs models are pairwise interaction point processes. They are introduced below and we will come back to this important class of Gibbs models in Section 4. *Pairwise potential interactions* take the particular form

$$(2) \quad H_{\Lambda}(\omega) = zN_{\Lambda}(\omega) + \sum_{\{x, y\} \in \omega, \{x, y\} \cap \omega_{\Lambda} \neq \emptyset} \phi(x - y),$$

where  $z > 0$  is the intensity parameter and  $\phi$  is the pair potential, a function from  $\mathbb{R}^d$  to  $\mathbb{R} \cup \{+\infty\}$  which is symmetric, that is,  $\phi(-x) = \phi(x)$  for all  $x$  in  $\mathbb{R}^d$ . In connection with (1), if  $H_{\Lambda}$  satisfies (2), then  $\varphi_{\Lambda, \Lambda'}(\omega_{\Lambda^c}) = zN_{\Lambda'}(\omega_{\Lambda^c}) + \sum_{\{x, y\} \in \omega_{\Lambda^c}, \{x, y\} \cap \Lambda' \neq \emptyset} \phi(x - y)$ .

Let us present two well-known examples. We will use them through the paper to illustrate our notation and assumptions.

EXAMPLE 1. The Strauss pair potential, defined for some possible hardcore parameter  $\delta \geq 0$ , some interacting parameter  $\beta \geq 0$ , and some range of interaction  $R \geq 0$  corresponds in (2) to

$$(3) \quad \phi(x) = \begin{cases} \infty, & \text{if } |x| < \delta, \\ \beta, & \text{if } \delta \leq |x| < R, \\ 0, & \text{if } |x| \geq R, \end{cases}$$

if  $R > \delta$ , while  $\phi(x) = \infty \mathbb{1}_{|x| < \delta}$  if  $R \leq \delta$ . As we will see later, the theoretical properties of this model strongly differ whether  $\delta = 0$ , which corresponds to the standard Strauss model, or  $\delta > 0$ , which is the hardcore Strauss model.

EXAMPLE 2. The general Lennard–Jones  $(n, m)$  pair potential is defined for some  $d < m < n$ , and some  $A > 0, B \in \mathbb{R}$  by (2) with

$$(4) \quad \phi(x) = A|x|^{-n} - B|x|^{-m}, \quad x \in \mathbb{R}^d.$$

The standard Lennard–Jones model in dimension  $d = 2$  and  $d = 3$  corresponds to  $n = 12$  and  $m = 6$ .

For a family of energies  $\mathcal{H} = (H_\Lambda)$ , we denote by  $\Omega_\infty$  the space of configurations which have a locally finite energy, that is,  $\omega \in \Omega_\infty$  if and only if, for any bounded Borel set  $\Lambda$ ,  $H_\Lambda(\omega)$  is finite.

The Gibbs measures  $P$  associated to  $\mathcal{H}$  are defined through their local conditional specification, as described below. For every  $\Lambda$  and every  $\omega$  in  $\Omega_\infty \cap \Omega_T$ , the conditional density  $f_\Lambda$  of  $P$  with respect to  $\pi_\Lambda$  is defined by

$$(5) \quad f_\Lambda(\omega) = \frac{1}{Z_\Lambda(\omega_{\Lambda^c})} e^{-H_\Lambda(\omega)},$$

where  $Z_\Lambda(\omega_{\Lambda^c})$  is the normalization constant, or partition function, given by

$$Z_\Lambda(\omega_{\Lambda^c}) = \int e^{-H_\Lambda(\omega'_\Lambda \cup \omega_{\Lambda^c})} \pi_\Lambda(d\omega'_\Lambda).$$

Some regularity assumptions on  $\mathcal{H}$  are required to ensure that  $0 < Z_\Lambda(\omega_{\Lambda^c}) < +\infty$ , implying that the local density is well-defined. They will be part of our general hypothesis in Section 3 and are fulfilled for all our examples in Section 4.

We are now in position to define the Gibbs measures associated to  $\mathcal{H}$ ; see, for instance, [8].

DEFINITION 1. A probability measure  $P$  on  $\Omega$  is a *Gibbs measure* for the family of energies  $\mathcal{H}$  if  $P(\Omega_\infty \cap \Omega_T) = 1$  and, for every bounded borel set  $\Lambda$ , for any measurable and integrable function  $g$  from  $\Omega$  to  $\mathbb{R}$ ,

$$(6) \quad \int g(\omega) P(d\omega) = \int \int g(\omega'_\Lambda \cup \omega_{\Lambda^c}) f_\Lambda(\omega'_\Lambda \cup \omega_{\Lambda^c}) \pi_\Lambda(d\omega'_\Lambda) P(d\omega).$$

Equivalently, for  $P$ -almost every  $\omega$  the conditional law of  $P$  given  $\omega_{\Lambda^c}$  is absolutely continuous with respect to  $\pi_\Lambda$  with the density  $f_\Lambda$  defined in (5).

The equations (6) are called the Dobrushin–Lanford–Ruelle (DLR) equations. Conditions on  $\mathcal{H}$  are mandatory to ensure the existence of a measure  $P$  satisfying (6). For general interactions, we will later assume that both (5) is well-defined and  $P$  exists. On the other hand, the unicity of  $P$  does not necessarily hold, leading to phase transition. As explained in the [Introduction](#), our consistency results are not affected by this phenomenon.

2.3. *Parametric Gibbs models and the MLE procedure.* We consider a parametric Gibbs model that depends on a hardcore parameter  $\delta \in I$ , where  $I = [\delta_{\min}, \delta_{\max}]$  with  $0 \leq \delta_{\min} \leq \delta_{\max} \leq \infty$ , and on a parameter  $\theta \in \Theta$  with  $\Theta \subset \mathbb{R}^p$  and  $\Theta \neq \emptyset$ . We stress the dependence on these parameters by adding some superscripts to our notation, for example, the partition function  $Z_\Lambda(\omega_{\Lambda^c})$  becomes  $Z_\Lambda^{\delta, \theta}(\omega_{\Lambda^c})$  and the interaction energy  $H_\Lambda(\omega)$  becomes  $H_\Lambda^\theta(\omega)$ . In particular, the reason why we write  $H_\Lambda^\theta(\omega)$  and not  $H_\Lambda^{\delta, \theta}(\omega)$  (and similarly for other quantities) is due to the fact that under our assumptions, the interaction energy will not depend on  $\delta$ , as explained below.

Specifically, we assume that the conditional density writes for any  $\delta \in I$  and any  $\theta \in \Theta$

$$(7) \quad f_\Lambda^{\delta, \theta}(\omega) = \frac{1}{Z_\Lambda^{\delta, \theta}(\omega_{\Lambda^c})} e^{-H_\Lambda^\theta(\omega)} \mathbb{1}_{\Omega_\infty^\delta}(\omega), \quad \omega \in \Omega_T,$$

where

$$\omega \in \Omega_\infty^\delta \Leftrightarrow \inf_{\{x, y\} \in \omega} |x - y| \geq \delta,$$

while for any  $\theta \in \Theta$  and any  $\omega \in \Omega_T$ ,  $H_\Lambda^\theta(\omega) < \infty$ . This specific parametric form clearly indicates that the hardcore parameter  $\delta$  only rules the support  $\Omega_\infty^\delta$  of the measure, and has no effect on the interaction energy  $H_\Lambda^\theta$ . This is what happens for most models in spatial statistics; see Section 4. A counter-example is the Diggle–Gratton model [7], Section 6, where the interaction energy depends also on  $\delta$ . This situation is not covered by our study.

Let  $\mathcal{G}^{\delta, \theta}$  be the set of stationary Gibbs measures defined by the conditional density (7) and with finite intensity. A minimal condition is to assume that for any value of the parameters, (7) is well-defined and that the latter set is not empty.

[EXISTENCE]. For any  $\delta \in I$ , any  $\theta \in \Theta$ , any bounded set  $\Lambda$  in  $\mathbb{R}^d$  and any  $\omega \in \Omega_\infty^\delta \cap \Omega_T$ ,  $Z_\Lambda^{\delta, \theta}(\omega_{\Lambda^c}) < \infty$  and the set  $\mathcal{G}^{\delta, \theta}$  is not empty.

EXAMPLE 1 (Continued). For the hardcore Strauss model defined in (3) with  $\delta > 0$ , the hardcore parameter is  $\delta$  and the parameter  $\theta$  corresponds to  $\theta = (z, R, \beta)$ . The assumption [EXISTENCE] for this model holds if we take  $I = \mathbb{R}_+$  and  $\Theta = \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}$ . On the other hand, for the standard Strauss model corresponding to the case  $\delta = 0$  in (3), there is no hardcore parameter, meaning that  $I = \{0\}$  and the only parameter of the model is  $\theta = (z, R, \beta)$ . The existence of this model, that is, [EXISTENCE], holds iff  $\theta$  belongs to  $\Theta = \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}_+$ . Note that  $\beta$  needs to be nonnegative when there is no hardcore, contrary to the hardcore case. We refer to [25] for the existence results involving stable and superstable pairwise potentials, a class of interactions which includes the Strauss model.

EXAMPLE 2 (Continued). For the Lennard–Jones model defined in (4), there is no hardcore parameter ( $\delta = 0$ ) and the parameter  $\theta$  corresponds to  $(z, A, B, n, m) \in \Theta \subset \mathbb{R}^5$  where  $\Theta$  is just defined via the constraints  $z \geq 0, A > 0$  and  $d < m < n$ . The existence is also proved in [25].

Let  $\delta^* \in I$  and  $\theta^*$  in  $\Theta$ . We denote by  $\omega^*$  a realization of a Gibbs measure belonging to  $\mathcal{G}^{\delta^*, \theta^*}$ . The parameters  $\delta^*$  and  $\theta^*$  represent the true unknown parameters that we want to estimate. The MLE of  $(\delta^*, \theta^*)$  from the observation of  $\omega^*$  on  $\Lambda_n$  is defined as follows.

DEFINITION 2. Let  $\mathcal{K}$  be a compact subset of  $\Theta$  such that  $\theta^* \in \mathcal{K}$ . The MLE of  $\delta^*$  and  $\theta^*$  from the observation of  $\omega^*$  in  $\Lambda_n$  is defined by

$$(8) \quad (\hat{\delta}_n, \hat{\theta}_n) = \operatorname{argmax}_{(\delta, \theta) \in I \times \mathcal{K}} f_{\Lambda_n}^{\delta, \theta}(\omega_{\Lambda_n}^*).$$

In Lemma 2 below, we give sufficient conditions which ensure the existence of the argmax in (8). Let us note that we consider here the MLE with free boundary condition, meaning that the configuration  $\omega_{\Lambda_n^c}^*$  outside  $\Lambda_n$  is not involved. This is the most natural setting given that we observe  $\omega^*$  only on  $\Lambda_n$ . A MLE procedure that depends on the outside configuration could have been considered as well and similar theoretical results would have been proved. To implement this alternative procedure in practice yet, the interaction has to be finite range so that the outside configuration reduces to boundary effects that can be handled by minus sampling. For infinite range potentials, as considered in Section 4.3, this method is not feasible and the MLE with free boundary condition makes more sense.

We will use in the following the equivalent definition of the MLE:

$$(\hat{\delta}_n, \hat{\theta}_n) = \operatorname{argmin}_{(\delta, \theta) \in I \times \mathcal{K}} K_n^{\delta, \theta}(\omega_{\Lambda_n}^*),$$

where  $K_n^{\delta, \theta}$  is the contrast function

$$(9) \quad K_n^{\delta, \theta}(\omega_{\Lambda_n}^*) = \frac{\ln(Z_{\Lambda_n}^{\delta, \theta})}{|\Lambda_n|} + \frac{H_{\Lambda_n}^{\theta}(\omega_{\Lambda_n}^*)}{|\Lambda_n|} + \infty \mathbb{1}_{\tilde{\delta}_n(\omega_{\Lambda_n}^*) < \delta}$$

with

$$(10) \quad \tilde{\delta}_n(\omega_{\Lambda_n}) = \min_{\{x, y\} \in \omega_{\Lambda_n}} |x - y|$$

and

$$(11) \quad Z_{\Lambda}^{\delta, \theta} = Z_{\Lambda}^{\delta, \theta}(\emptyset) = \int e^{-H_{\Lambda}^{\theta}(\omega_{\Lambda})} \mathbb{1}_{\Omega_{\infty}^{\delta}}(\omega_{\Lambda}) \pi_{\Lambda}(d\omega_{\Lambda}).$$

In statistical mechanics, the first term in (9) is called the finite volume pressure while the second term corresponds to the specific energy.

In the following lemma, we give an explicit expression for  $\hat{\delta}_n$ .



LEMMA 1. Under [EXISTENCE], the MLE of the hardcore parameter  $\delta$  is  $\hat{\delta}_n = \tilde{\delta}_n$  if  $\tilde{\delta}_n \in I$  and  $\hat{\delta}_n = \delta_{\max}$  otherwise, where  $\tilde{\delta}_n$  is given by (10), consequently

$$(12) \quad \hat{\theta}_n = \operatorname{argmin}_{\theta \in \mathcal{K}} K_n^{\hat{\delta}_n, \theta}(\omega_{\Lambda_n}^*).$$

PROOF. If  $\delta > \tilde{\delta}_n$ , then  $K_n^{\delta, \theta}(\omega_{\Lambda_n}^*) = \infty$  and so  $\hat{\delta}_n \leq \tilde{\delta}_n$ . For any  $\delta \leq \tilde{\delta}_n$  and any  $\theta \in \Theta$ ,  $K_n^{\delta, \theta}(\omega_{\Lambda_n}^*) = |\Lambda_n|^{-1}(\ln(Z_{\Lambda_n}^{\delta, \theta}) + H_{\Lambda_n}^\theta(\omega_{\Lambda_n}^*))$ . Note that if  $\delta < \delta'$ , then  $\Omega_\infty^\delta \subset \Omega_\infty^{\delta'}$  and from (11) the function  $\delta \mapsto Z_{\Lambda_n}^{\delta, \theta}$  is decreasing. Therefore, for any  $\theta \in \Theta$  and for any  $\delta \leq \tilde{\delta}_n$ ,  $\delta \mapsto K_n^{\delta, \theta}(\omega_{\Lambda_n}^*)$  is decreasing, proving that  $\hat{\delta}_n \geq \tilde{\delta}_n$ . Hence,  $\hat{\delta}_n = \tilde{\delta}_n$  if  $\tilde{\delta}_n \in I$ . The other statements of Lemma 1 are straightforward.  $\square$

This lemma provides a useful result for practical purposes, since it states that the MLE for the hardcore parameter is just the minimal distance between pairs of points observed in the pattern, as it is usually implemented. Let us remark that a crucial assumption is that the space of parameters in [EXISTENCE] is the Cartesian product  $I \times \Theta$ , that is, the parametric model is well-defined for any value of  $\theta \in \Theta$ , for whatever  $\delta$  is. This is the case for the examples considered in Section 4.

As explained in the Introduction, the computation of  $\hat{\theta}_n$  given by (12) is more difficult because there is in general no closed form expression for  $Z_{\Lambda_n}^{\delta, \theta}$ . In practice, the optimization (12) is conducted from an approximated contrast function, based on Markov chain Monte-Carlo methods where the error of approximation can be controlled; see, for instance, [11]. Even to ensure the existence of  $\hat{\theta}_n$ , we need some extra assumptions, as detailed in the following lemma.

[ARGMAX]. For any  $\omega \in \Omega_T$  and any  $n \geq 1$ , the function  $\theta \mapsto H_{\Lambda_n}^\theta(\omega_{\Lambda_n})$  is lower semicontinuous over  $\mathcal{K}$ , that is, for any  $\theta \in \mathcal{K}$   $\liminf_{\theta' \rightarrow \theta} H_{\Lambda_n}^{\theta'}(\omega_{\Lambda_n}) \geq H_{\Lambda_n}^\theta(\omega_{\Lambda_n})$ , and there exists an upper semicontinuous version  $\theta \mapsto \tilde{H}_{\Lambda_n}^\theta(\omega_{\Lambda_n})$  such that for every  $\theta \in \mathcal{K}$ ,  $\tilde{H}_{\Lambda_n}^\theta = H_{\Lambda_n}^\theta$   $\pi_{\Lambda_n}$ -almost surely.

EXAMPLE 1 (Continued). For all  $x \neq 0$ , the function  $\theta \mapsto \phi(x)$  in (3) is lower semicontinuous which ensures that the associated energy  $\theta \mapsto H_{\Lambda_n}^\theta$  for the Strauss model is lower semicontinuous as well. Changing the value of  $\phi$  at the discontinuity points, it is easy to obtain an upper semicontinuous version of the energy function which is  $\pi_{\Lambda_n}$ -almost surely equal to  $H_{\Lambda_n}^\theta$ . Therefore, [ARGMAX] holds for the Strauss model.

EXAMPLE 2 (Continued). [ARGMAX] also holds for the Lennard–Jones model since the function  $\theta \mapsto \phi(x)$  in (4) is continuous for any  $x \neq 0$ .

LEMMA 2. Under [EXISTENCE] and [ARGMAX], the MLE of  $\theta$  exists, that is, there exists at least one  $\hat{\theta}_n \in \mathcal{K}$  such that  $K_n^{\hat{\theta}_n, \hat{\theta}_n}(\omega_{\Lambda_n}^*) = \min_{\theta \in \mathcal{K}} K_n^{\hat{\theta}_n, \theta}(\omega_{\Lambda_n}^*)$ .

PROOF. It is sufficient to show that  $\theta \mapsto K_n^{\hat{\theta}_n, \theta}(\omega_{\Lambda_n}^*)$  is lower semicontinuous. As a direct consequence of [ARGMAX],  $\theta \mapsto H_{\Lambda_n}^\theta(\omega_{\Lambda_n}^*)$  is lower semicontinuous. For the finite volume pressure, this comes from [ARGMAX] and the Fatou’s lemma since

$$\begin{aligned} \liminf_{\theta' \mapsto \theta} Z_{\Lambda_n}^{\hat{\theta}_n, \theta'} &= \liminf_{\theta' \mapsto \theta} \int e^{-\tilde{H}_{\Lambda_n}^{\theta'}(\omega_{\Lambda_n})} \mathbb{1}_{\Omega_{\infty}^{\hat{\theta}_n}}(\omega_{\Lambda}) \pi_{\Lambda_n}(d\omega_{\Lambda_n}) \\ &\geq \int \liminf_{\theta' \mapsto \theta} e^{-\tilde{H}_{\Lambda_n}^{\theta'}(\omega_{\Lambda_n})} \mathbb{1}_{\Omega_{\infty}^{\hat{\theta}_n}}(\omega_{\Lambda}) \pi_{\Lambda_n}(d\omega_{\Lambda_n}) \\ &\geq \int e^{-\tilde{H}_{\Lambda_n}^{\theta}(\omega_{\Lambda_n})} \mathbb{1}_{\Omega_{\infty}^{\hat{\theta}_n}}(\omega_{\Lambda}) \pi_{\Lambda_n}(d\omega_{\Lambda_n}) = Z_{\Lambda_n}^{\hat{\theta}_n, \theta}. \quad \square \end{aligned}$$

**3. Consistency of the MLE for general Gibbs interactions.** We detail in this section the minimal assumptions on the parametric family of Gibbs measure that imply the strong consistency of the MLE. All of them are fulfilled by the examples of Section 4.

The first ones, [EXISTENCE] and [ARGMAX] introduced in the previous section, imply the existence of a Gibbs measure for any  $\delta \in I$  and  $\theta \in \Theta$ , and the existence of the MLE.

We assume the following mild assumptions on the family of energies. They gather the stability of the energy function, a standard stationary decomposition in terms of mean energy per unit volume and they deal with boundary effects. We put  $\mathcal{G} = \bigcup_{\delta \in I, \theta \in \Theta} \mathcal{G}^{\delta, \theta}$ .

[STABILITY]. For any compact set  $\mathcal{K} \subset \Theta$ , there exists a constant  $\kappa \geq 0$  such that for any  $\Lambda$ , any  $\theta \in \mathcal{K}$  and any  $\omega \in \Omega_T$

$$(13) \quad H_{\Lambda}^{\theta}(\omega_{\Lambda}) \geq -\kappa N_{\Lambda}(\omega).$$

[MEANENERGY]. There exist measurable functions  $H_0^{\theta}$  and  $\partial H_n^{\theta}$  from  $\Omega_T$  to  $\mathbb{R}$  such that for all  $n \geq 1$  and all  $\theta \in \Theta$

$$(14) \quad H_{\Lambda_n}^{\theta} = \sum_{k \in \mathcal{I}_n} H_0^{\theta} \circ \tau_{-k} + \partial H_{\Lambda_n}^{\theta},$$

where for all  $P \in \mathcal{G}$  and all  $\theta \in \Theta$  the mean energy of  $P$  defined by

$$(15) \quad H^{\theta}(P) = E_P(H_0^{\theta})$$

is finite and for any compact set  $\mathcal{K} \subset \Theta$ , for  $P$ -almost every  $\omega \in \Omega_T$ ,

$$\lim_{n \rightarrow \infty} \frac{1}{|\Lambda_n|} \sup_{\theta \in \mathcal{K}} |\partial H_{\Lambda_n}^{\theta}(\omega)| = 0.$$

[BOUNDARY]. For all  $P \in \mathcal{G}$ , for any compact set  $\mathcal{K} \subset \Theta$  and for  $P$ -almost every  $\omega \in \Omega_T$

$$\lim_{n \rightarrow \infty} \frac{1}{|\Lambda_n|} \sup_{\theta \in \mathcal{K}} |H_{\Lambda_n}^\theta(\omega_{\Lambda_n}) - H_{\Lambda_n}^\theta(\omega)| = 0.$$

To prove the consistency of the MLE, we need to control the asymptotic behaviour of the finite volume pressure and the specific energy in (9). The following assumption deals with the latter.

[REGULARITY]. For any  $P$  in  $\mathcal{G}$ , for any compact set  $\mathcal{K} \subset \Theta$ , there exists a function  $g$  from  $\mathbb{R}_+ \rightarrow \mathbb{R}_+$  with  $\lim_{r \rightarrow 0} g(r) = 0$  such that for any  $\theta \in \mathcal{K}$  and any  $r > 0$

$$(16) \quad E_P \left( \sup_{\substack{\theta' \in \mathcal{K} \\ |\theta - \theta'| \leq r}} |H_0^\theta - H_0^{\theta'}| \right) \leq g(r).$$

Moreover, there exists a bounded subset  $\Lambda_0$  such that for any  $\eta > 0$  and any  $\theta_0 \in \mathcal{K}$ , there exists a finite subset  $\mathcal{N}(\theta_0) \subset B(\theta_0, \eta) \cap \Theta$  and  $r(\theta_0) > 0$  such that  $r(\theta_0) < \eta$  and for any  $\Lambda \supset \Lambda_0$ ,

$$(17) \quad \max_{\theta \in \mathcal{N}(\theta_0)} \inf_{\theta' \in B(\theta_0, r(\theta_0)) \cap \mathcal{K}} \inf_{\omega_\Lambda \in \Omega_\infty^{\delta_{\min}}} \left( \frac{H_\Lambda^\theta(\omega_\Lambda) - H_\Lambda^{\theta'}(\omega_\Lambda)}{N_\Lambda(\omega_\Lambda)} \right) \geq -g(r(\theta_0)).$$

Assumption (16) implies in particular that the mean energy  $H^\theta(P)$  is continuous with respect to  $\theta$ . Let us note that this does not imply that the specific energy or  $\theta \mapsto H_0^\theta$  is continuous with respect to  $\theta$ . As a counterexample, the Strauss model in Example 1, or more generally any model from Sections 4.1 and 4.2, has a discontinuous specific energy but the mean energy is continuous, which is verified in the proof of Theorem 2. Assumption (17) is related to the regularity of the coefficient  $\kappa$  in [STABILITY] with respect to the parameter  $\theta$ .

Finally, the following assumption relates the pressure and the specific entropy defined below, through the variational principle. The variational principle is conjectured to hold for all models of statistical mechanics even if it is not proved in such a general setting. It is established for stable and finite-range interactions in [4] and for infinite-range pair potentials that are regular and nonintegrably divergent at the origin in [9, 10].

For any  $\delta \in I$  and  $\theta \in \Theta$ , the *specific entropy* of  $P^{\delta, \theta} \in \mathcal{G}^{\delta, \theta}$  is the limit

$$(18) \quad \mathcal{I}(P^{\delta, \theta}) = \lim_{n \rightarrow \infty} \frac{1}{|\Lambda_n|} \mathcal{I}_{\Lambda_n}(P^{\delta, \theta}),$$

where  $\mathcal{I}_{\Lambda_n}(P^{\delta, \theta}) = \int \ln f_{\Lambda_n}^{\delta, \theta}(\omega_{\Lambda_n}) P_{\Lambda_n}^{\delta, \theta}(d\omega)$  is the finite volume entropy of  $P^{\delta, \theta}$ . Note that this limit always exists; see [8].

[VARPRIN]. For any  $\delta \in I$  and any  $\theta \in \Theta$ , the *pressure* defined as the following limit exists and is finite:

$$(19) \quad p(\delta, \theta) = \lim_{n \rightarrow \infty} \frac{1}{|\Lambda_n|} \ln(Z_{\Lambda_n}^{\delta, \theta}).$$

Moreover, the map  $\delta \mapsto p(\delta, \theta)$  is right continuous on  $I$ , that is,  $\lim_{\delta' \rightarrow \delta, \delta' \geq \delta} p(\delta', \theta) = p(\delta, \theta)$ . In addition, for any  $\delta, \delta'$  in  $I$  with  $\delta' \geq \delta$ , for any  $\theta, \theta'$  in  $\Theta$ , for all  $P^{\delta', \theta'} \in \mathcal{G}^{\delta', \theta'}$  the inequality

$$(20) \quad p(\delta, \theta) \geq -\mathcal{I}(P^{\delta', \theta'}) - H^\theta(P^{\delta', \theta'})$$

holds. If  $(\delta', \theta') = (\delta^*, \theta^*)$ , then equality holds in (20) if and only if  $(\delta, \theta) = (\delta^*, \theta^*)$ .

The formulation of the variational principle presented here is slightly weaker than the classical one in statistical mechanics where the inequality (20) is required for any probability measure  $P$  on  $\Omega_T$ , and not only for  $P^{\delta', \theta'} \in \mathcal{G}^{\delta', \theta'}$ . Note that the right continuity of the pressure holds trivially when  $\delta^*$  is known, that is,  $I = \{\delta^*\}$ . Let us also remark that the last part in [VARPRIN] concerning  $(\delta^*, \theta^*)$  is just an identifiability assumption.

We are now in position to state the consistency of the MLE.

**THEOREM 1.** *Under the assumptions [EXISTENCE], [ARGMAX], [STABILITY], [MEANENERGY], [BOUNDARY], [REGULARITY] and [VARPRIN], for any  $(\delta^*, \theta^*) \in I \times \mathcal{K}$  and any  $P \in \mathcal{G}^{\delta^*, \theta^*}$ , the MLE  $(\hat{\delta}_n, \hat{\theta}_n)$  is well-defined and converges  $P$ -almost surely to  $(\delta^*, \theta^*)$  when  $n$  goes to infinity.*

**EXAMPLE 1 (Continued).** The Strauss and the hardcore Strauss models satisfy all the assumptions of Theorem 1 which ensures the consistency of the MLE of  $\delta^*$ ,  $\beta^*$  and  $R^*$ . The assumptions [EXISTENCE] and [ARGMAX] have already been discussed earlier and the stability assumption is obvious here since  $\phi \geq 0$ . The assumptions [MEANENERGY] and [VARPRIN] involve a variational characterisation of infinite volume Gibbs measures which is standard in statistical mechanics [4]. On the contrary, [REGULARITY] is a technical assumption which is specific to the study of consistency of the MLE. The Strauss model belongs to the large class of finite range pairwise interactions developed in detail in Sections 4.1 and 4.2, and [REGULARITY] is verified for these models in the Supplementary Material [5] (see the proof of Theorem 2 when the hardcore parameter  $\delta^*$  is positive and the proof of Theorem 3 when  $\delta^* = 0$ ).

**EXAMPLE 2 (Continued).** The Lennard–Jones model also satisfies the assumptions of Theorem 1. The assumptions [EXISTENCE] and [ARGMAX] have been checked in Section 2.3. The stability assumption is already known (see, for

instance, Section 3.2 in [24]) and is proved using the theory of positive type functions and Bochner's theorem. The assumptions [MEANENERGY] and [VARPRIN] are proved in [9, 10] which deal with the variational principle in the general setting of infinite range pairwise interactions. Again the [REGULARITY] assumption is more specific and requires fine computations given in the Supplementary Material [5]; see the proof of Theorem 4 which includes as a particular case the Lennard–Jones model.

As illustrated in the two examples above, all the assumptions of Theorem 1, except [REGULARITY], are natural and standard in statistical mechanics. They generally do not require much efforts to be checked since they are already proved for many models in the literature. In fact, only assumption [REGULARITY] is really specific to the study of consistency of the MLE and needs a particular attention. Its proof strongly depends on the underlying model. Nonetheless, the techniques we develop in the Supplementary Material [5] in the cases of the Strauss, the hard-core Strauss and the Lennard–Jones models (and their generalizations) provide tools and strategies to verify [REGULARITY]. We think that they could be easily extended to the study of other models.

As a noticeable example stated in the following corollary, if the energy  $H_\Lambda^\theta$  depends linearly on the parameter  $\theta$ , which is called the exponential model, then the assumptions [ARGMAX] and [REGULARITY] are automatically satisfied. Specifically, for any  $1 \leq k \leq p$ , let  $(H_\Lambda^k)$  be a family of interaction energies satisfying (1). For any  $\theta \in \Theta$  and any bounded set  $\Lambda$ , an exponential model takes the form

$$(21) \quad H_\Lambda^\theta = \sum_{k=1}^p \theta_k H_\Lambda^k.$$

Let us introduce the following modification of [MEANENERGY].

[MEANENERGY']. Assumption [MEANENERGY] holds with  $H_0^\theta = \sum_{k=1}^p \theta_k H_0^k$  where  $H_0^k$  comes from the decomposition (14) applied to  $H_{\Lambda_n}^k$ .

**COROLLARY 1.** *Assume that the family of interaction energies satisfy (21) and that  $\Theta$  is an open subset of  $\mathbb{R}^p$ . Then under the assumptions [EXISTENCE], [STABILITY], [MEANENERGY'], [BOUNDARY] and [VARPRIN], for any  $(\delta^*, \theta^*) \in I \times \mathcal{K}$  and any  $P \in \mathcal{G}^{\delta^*, \theta^*}$ , the MLE  $(\hat{\delta}_n, \hat{\theta}_n)$  is well-defined and converges  $P$ -almost surely to  $(\delta^*, \theta^*)$  when  $n$  goes to infinity.*

**REMARK 1.** In Theorem 1, there is no topological assumption on the set  $\Theta$  except that  $\Theta \neq \emptyset$ . However, for identifiability reasons or in order to check [REGULARITY], some assumptions on  $\Theta$  will often come out, depending on the model. This is illustrated in Corollary 1 where we assume  $\Theta$  to be an open set. This assumption is usual and will still be used in the specific examples of the next

section. Nevertheless, this hypothesis is in general not necessary for consistency. For some models, it is possible to check all assumptions of Theorem 1 when  $\Theta$  is not open and  $\theta^*$  is on the boundary of  $\Theta$ . As an example, it is not difficult to prove that the estimation of  $\beta^* = 0$  in the Strauss model (where  $\delta^* = 0$  and  $R^*$  is known) is consistent, in which case  $\theta = \beta$ ,  $\Theta = [0, \infty)$  and (21) is satisfied. This can be done as in the proof of Corollary 1, where the assumption (17) in [REGULARITY] is verified by taking, for any  $\eta > 0$  and  $\beta_0 \geq 0$ ,  $\mathcal{N}(\beta_0) = \{\beta_0 + \eta/2\}$  and  $r < \eta/2$ . Note that the same conclusion is not true if in this model we try to estimate both  $\beta^*$  and  $R^*$  when  $\beta^* = 0$ , because then  $R^*$  is not identifiable. Finally, the study of consistency for parameters on the boundary of  $\Theta$  is specific to the model at hand and it is difficult to draw general results in this case.

**4. Consistency for pairwise interactions and other examples.**

4.1. *Finite range piecewise continuous pair potentials with hardcore.* In this section, we illustrate how the general result of Theorem 1 applies to the class of finite-range pairwise interactions with hardcore. The following framework gathers all potentials of this type described in [2, 13, 20] and [26], including the hardcore Strauss model of Example 1 (detailed at the end of this section in Corollary 2), the hardcore piecewise constant pairwise interaction model and the Fiksel model.

Specifically, we consider conditional densities given by (7) with a nonvanishing hardcore parameter, that is,  $\delta_{\min} > 0$ , and a Hamiltonian  $H_\Lambda^\theta$  defined as in (2) by

$$(22) \quad H_\Lambda^\theta(\omega) = zN_\Lambda(\omega) + \sum_{\{x,y\} \in \omega, \{x,y\} \cap \omega_\Lambda \neq \emptyset} \phi^{\beta,R}(x-y),$$

where  $z > 0$ ,  $\beta \in \mathbb{R}^p$  and  $R = (R_1, \dots, R_q)$  is a multidimensional interaction parameter in  $\mathbb{R}^q$  belonging to the subset  $\mathcal{R} = \{R \in \mathbb{R}^q, 0 < R_1 < \dots < R_q\}$ . The pair potential interaction  $\phi^{\beta,R}$  is defined piecewise as follows. For any  $x \in \mathbb{R}^d$ , such that  $|x| \neq R_1, R_2, \dots, R_q$

$$(23) \quad \phi^{\beta,R}(x) = \mathbb{1}_{[0, R_1)}(|x|)\varphi_1^{\beta,R}(|x|) + \dots + \mathbb{1}_{(R_{q-1}, R_q)}(|x|)\varphi_q^{\beta,R}(|x|),$$

while

$$(24) \quad \phi^{\beta,R}(x) = \min(\varphi_k^{\beta,R}(|x|), \varphi_{k+1}^{\beta,R}(|x|)) \quad \text{if } |x| = R_k \text{ for some } k$$

with the convention  $\varphi_{q+1}^{\beta,R} = 0$ . For any  $k = 1, \dots, q$ , we assume that there exists an open subset  $\mathcal{B}$  of  $\mathbb{R}^p$  such that

$$(25) \quad (\beta, R, x) \mapsto \varphi_k^{\beta,R}(|x|) \text{ is continuous on } \mathcal{B} \times \mathcal{R} \times \mathbb{R}^d.$$

Note that from (23)–(25), the function  $(\beta, R) \mapsto \phi^{\beta,R}(x)$  is lower semicontinuous which is crucial for the assumption [ARGMAX].

In this setting, the parameters to estimate are  $\delta^*$  and  $\theta = (z^*, \beta^*, R^*)$  that we assume belong to  $I = [\delta_{\min}, \infty)$  and  $\Theta = \mathbb{R}_+ \times \mathcal{B} \times \mathcal{R}$ , respectively. For identifiability reasons, we further assume that

$$(26) \quad R_1^* > \delta^* \quad \text{and} \quad \lambda^d(x \in \mathbb{R}^d, \mathbb{1}_{[\delta^*, +\infty)}(|x|)(\phi^{\beta^*, R^*}(x) - \phi^{\beta, R}(x)) \neq 0) > 0$$

for any  $(\beta, R) \neq (\beta^*, R^*)$ . This implies that if  $(\delta, \theta) \neq (\delta^*, \theta^*)$  then  $\mathcal{G}^{\delta, \theta} \cap \mathcal{G}^{\delta^*, \theta^*} = \emptyset$ . Notice that the assumption  $R_1^* > \delta^*$  is necessary to ensure identifiability of  $R^*$  and  $\delta^*$ , but we do not need to add the constraint  $R_1 > \delta$  in the MLE optimisation (8), as confirmed by the choice of  $\mathcal{K}$  in the theorem below. This is in particular important to agree with the setting of Lemma 1, where it is necessary that the optimisation (8) is carried out on a Cartesian product  $I \times \mathcal{K}$ .

The proof of the following theorem is given in the Supplementary Material [5].

**THEOREM 2.** *Consider the Gibbs model defined by (7) with the Hamiltonian (22) and the pair potential (23)–(24), under the assumptions  $\delta_{\min} > 0$ , (25) and (26). Let  $\mathcal{K}$  be a compact subset of  $\mathbb{R}_+ \times \mathcal{B} \times \mathcal{R}$  such that  $(z^*, \beta^*, R^*) \in \mathcal{K}$ . Then the MLE of  $(\delta^*, z^*, \beta^*, R^*)$  given by (8) is strongly consistent.*

**COROLLARY 2 (Hardcore Strauss model).** *Consider the hardcore Strauss model of Example 1 with parameters  $\delta^* \geq \delta_{\min}$  where  $\delta_{\min} > 0$ ,  $z^* > 0$ ,  $\beta^* \in \mathbb{R}$  and  $R^* > \delta^*$ . Let  $\mathcal{K}$  be a compact subset of  $\mathbb{R}_+ \times \mathbb{R} \times \mathbb{R}_+$  such that  $(z^*, \beta^*, R^*) \in \mathcal{K}$ . Then the MLE of  $(\delta^*, z^*, \beta^*, R^*)$  given by (8) is strongly consistent.*

**PROOF.** The hardcore Strauss model corresponds to the Hamiltonian (22) and the pair potential (23) where  $q = p = 1$  and  $\varphi_1^{\beta, R} \equiv \beta$ . In this setting,  $\mathcal{R} = \mathbb{R}_+$ . The assumptions (25) and (26) thus hold trivially true for  $\mathcal{B} = \mathbb{R}$  and Theorem 2 applies.  $\square$

**4.2. Finite range piecewise continuous pair potentials without hardcore.** We consider the same setting as in the previous section except that there is no hardcore, that is  $\delta^* = 0$  and  $\delta^*$  is known, meaning that it has not to be fitted, or equivalently  $\delta_{\min} = \delta_{\max} = 0$  implying  $I = \{0\}$  in (8). The Hamiltonian  $H_\Lambda^\theta$  is defined as in (22) and (23)–(24) with the same continuity assumptions (25). The standard example we have in mind for this section is the Strauss model of Example 1 or any of the examples mentioned in the previous section without assuming the presence of a hardcore part.

Whereas this setting may appear to be simpler than the previous one from a statistical point of view (there is one parameter less to fit), it requires additional assumptions to ensure the existence of the model and to control its regularity. This has been already pointed out for the Strauss model of Example 1 after the statement of [EXISTENCE]. Consequently, we further assume that the Hamiltonian  $H_\Lambda^\theta$  satisfies [STABILITY] and that for any  $(\beta, R)$  the pair potential  $\phi^{\beta, R}$  is superstable,

which means that it is the sum of a stable pair potential plus a nonnegative potential which is positive around zero (see [25]). The identifiability assumptions remain similar as in the previous section, namely (26) with  $\delta^* = 0$ .

To handle the second part of [REGULARITY], we need a last hypothesis on the potential function. We assume that for any  $(\beta_0, R_0) \in \mathcal{B} \times \mathcal{R}$  and any  $\eta > 0$  there exists  $\beta_1 \in B(\beta_0, \eta) \cap \mathcal{B}$  such that for any  $k = 1, \dots, q$

$$(27) \quad \varphi_k^{\beta_1, R_0} > \varphi_k^{\beta_0, R_0}.$$

**THEOREM 3.** *Consider the Gibbs model defined by (7) with the Hamiltonian (22) and the pair potential (23)–(24), under the assumptions  $\delta_{\min} = \delta_{\max} = 0$ , (25) and (26) where  $\delta^* = 0$ . Assume further that the pair potential  $\phi^{\beta, R}$  is superstable and (27). Let  $\mathcal{K}$  be a compact subset of  $\mathbb{R}_+ \times \mathcal{B} \times \mathcal{R}$  such that  $(z^*, \beta^*, R^*) \in \mathcal{K}$ . Then the MLE of  $(z^*, \beta^*, R^*)$  given by (8) is strongly consistent.*

This theorem is proved in the Supplementary Material [5]. It can be easily applied to all standard finite range pairwise potentials (without hardcore) described in [2, 13, 20] and [26]. In particular, condition (27) turns out to be nonrestrictive. Because of its central role in spatial statistics, we focus in the following corollary on the Strauss model of Example 1.

**COROLLARY 3 (Strauss model).** *Consider the Strauss model of Example 1 with parameters  $z^* > 0, \beta^* > 0, R^* > 0$  and there is no hardcore, that is,  $\delta^* = 0$  is known. Let  $\mathcal{K}$  be a compact subset of  $\mathbb{R}_+^3$  such that  $(z^*, \beta^*, R^*) \in \mathcal{K}$ . Then the MLE of  $(z^*, \beta^*, R^*)$  given by (8) is strongly consistent.*

**PROOF.** Recall that the Strauss model corresponds in (23) to  $q = p = 1$  and  $\varphi_1^{\beta, R} \equiv \beta$ . As in the proof of Corollary 2, the assumptions (25) and (26) (where  $\delta^* = 0$ ) hold true with the choice  $\mathcal{B} = \mathbb{R}_+$ . Moreover, the associated pair potential  $\phi^{\beta, R}$  is positive for any  $\beta > 0$  and  $R > 0$ , which shows that it is superstable and that [STABILITY] holds true. Finally, the assumption (27) is verified if we choose  $\beta_1 > \beta_0$ . Therefore, Theorem 3 applies.  $\square$

**4.3. Infinite range pair potentials with a smooth parametrization.** In this section, we consider pairwise models in the spirit of the Lennard–Jones model given in Example 2. Our general setting concerns infinite range pair potentials without hardcore, that is,  $\delta_{\min} = \delta_{\max} = 0$ , that are uniformly regular and nonintegrably divergent at the origin in the sense of Ruelle [25]. Specifically, we assume that there exist two positive decreasing functions  $\psi$  and  $\chi$  from  $\mathbb{R}_+$  to  $\mathbb{R}$  and  $r_0 > 0$  with

$$\int_{r_0}^{+\infty} \psi(t)t^{d-1} dt < \infty, \quad \int_0^{r_0} \chi(t)t^{d-1} dt = +\infty$$



such that for any parameter  $\beta \in \mathcal{B}$ , where  $\mathcal{B}$  is an open set in  $\mathbb{R}^p$ , the function  $\phi^\beta : \mathbb{R}^d \rightarrow \mathbb{R}$  satisfies for any  $x \in \mathbb{R}^d$

$$(28) \quad \phi^\beta(x) \geq \chi(|x|) \quad \text{if } |x| \leq r_0 \quad \text{and} \quad |\phi^\beta(x)| \leq \psi(|x|) \quad \text{if } |x| > r_0.$$

The Hamiltonian  $H_\Lambda^\theta$  is defined as in (2) by

$$(29) \quad H_\Lambda^\theta(\omega) = zN_\Lambda(\omega) + \sum_{\{x,y\} \in \omega, \{x,y\} \cap \omega_\Lambda \neq \emptyset} \phi^\beta(x - y),$$

where  $\theta = (z, \beta)$  is in  $\Theta = \mathbb{R}_+ \times \mathcal{B}$ .

We assume that  $\phi^\beta$  is a symmetric function on  $\mathbb{R}^d \setminus \{0\}$  and that for any  $x \in \mathbb{R}^d \setminus \{0\}$ , the map  $\beta \mapsto \phi^\beta(x)$  is differentiable on  $\mathcal{B}$ . Denoting by  $\nabla\phi^\beta(x)$  its gradient, we also assume that for any compact set  $\mathcal{K} \subset \mathcal{B}$ , for any  $\beta \in \mathcal{K}$  and  $x \in \mathbb{R}^d$  with  $|x| > r_0$

$$(30) \quad |\nabla\phi^\beta(x)| \leq \psi(|x|)$$

and

$$(31) \quad \sup_{\beta \in \mathcal{K}} \sup_{x \in \mathbb{R}^d} \mathbb{1}_{|x| \leq r_0} \sup_{\beta' \in \mathcal{K}} e^{-\phi^{\beta'}(x)} \max(\phi^\beta(x), |\nabla\phi^\beta(x)|) < \infty.$$

In addition, we suppose that for any compact set  $\mathcal{K} \subset \mathcal{B}$  there exists an open set  $U$  in  $\mathbb{R}^p$  and a stable pair potential  $\tilde{\phi}$  from  $\mathbb{R}^d$  to  $\mathbb{R}$  such that for any  $u \in U$ , any  $x \in \mathbb{R}^d$  and any  $\beta \in \mathcal{K}$

$$(32) \quad \nabla\phi^\beta(x) \cdot u \geq \tilde{\phi}(x),$$

where  $v \cdot u$  denotes the scalar product of vectors  $u, v$  in  $\mathbb{R}^p$ .

Finally, for identifiability reasons, we assume that for any  $\beta \neq \beta'$  in  $\mathcal{B}$ ,

$$(33) \quad \lambda^d(x \in \mathbb{R}^d, \phi^\beta(x) \neq \phi^{\beta'}(x)) > 0.$$

**THEOREM 4.** *Let  $(\phi^\beta)_{\beta \in \mathcal{B}}$  be a family of pair potentials which are uniformly regular, nonintegrably divergent at the origin and satisfy assumptions (30)–(33). Let  $\mathcal{K}$  be a compact subset of  $\Theta$  such that  $\theta^* = (z^*, \beta^*)$  belongs to  $\mathcal{K}$ . Then the MLE of  $(z^*, \beta^*)$  given by (8) is strongly consistent.*

The proof of this theorem can be found in the Supplementary Material [5]. As a fundamental example, the following corollary focuses on the Lennard–Jones model of Example 2. Let the triangle domain  $\mathcal{T} = \{(n, m) \in \mathbb{R}^2, d < m < n\}$ .

**COROLLARY 4 (Lennard–Jones model).** *Consider the Lennard–Jones model of Example 2 given by (29) and (4) with parameters  $z > 0$  and  $\beta = (A, B, n, m) \in \mathcal{B}$  where  $\mathcal{B} = \mathbb{R}_+ \times \mathbb{R} \times \mathcal{T}$ . Let  $\mathcal{K}$  be a compact subset of  $\mathbb{R}_+ \times \mathcal{B}$  such that  $(z^*, \beta^*) \in \mathcal{K}$ . Then the MLE of  $(z^*, \beta^*)$  given by (8) is strongly consistent.*

PROOF. Denote by  $\mathcal{K}_{\mathcal{T}}$  the projection of  $\mathcal{K}$  onto  $\mathcal{T}$ , that is,  $(n, m) \in \mathcal{K}_{\mathcal{T}}$  if and only if there exists  $z > 0$ ,  $A > 0$  and  $B$  such that  $(z, A, B, n, m) \in \mathcal{K}$ . It is easy to find  $r_0 > 0$  and two positive constants  $c_0, c_1$  such that  $\phi^\beta$  given by (4) is uniformly regular and nonintegrably divergent at the origin for  $r_0$ ,  $\psi(t) = c_0 t^{-m_0}$  and  $\chi(t) = t^{-n_0}$  with  $d < m_0 < m$  and  $d < n_0 < n$  for all  $(n, m) \in \mathcal{K}_{\mathcal{T}}$ .

The map  $\beta \mapsto \phi^\beta$  is clearly differentiable on  $\mathcal{B}$  with

$$\nabla \phi^\beta(x) = (|x|^{-n}, |x|^{-m}, -A \ln(|x|)|x|^{-n}, B \ln(|x|)|x|^{-m}).$$

In adjusting the constant  $c_0$  in the definition of  $\psi$ , we show that assumption (30) holds. Assumptions (31) and (33) are obvious. It remains to show (32). Let  $u_0$  be the vector  $(1, 1, 1, 1)$  in  $\mathbb{R}^4$  and  $u$  any vector in the open ball  $B(u_0, 1/2)$  in  $\mathbb{R}^4$ . Denoting  $\mathcal{K}_{\mathcal{B}}$  the projection of  $\mathcal{K}$  onto  $\mathcal{B}$ , we find that for any  $\beta \in \mathcal{K}_{\mathcal{B}}$  and  $x \in \mathbb{R}^d$

$$\begin{aligned} \nabla \phi^\beta(x) \cdot u &\geq \frac{1}{2} \mathbb{1}_{[0,1]}(|x|)(|x|^{-n} - |b_0 \ln(|x|)|x|^{-m}) \\ &\quad - \frac{3}{2} \mathbb{1}_{(1,+\infty)}(|x|)(a_1 + |b_1|) \ln(|x|)|x|^{-m}, \end{aligned}$$

where  $a_1, b_0$  and  $b_1$  are chosen so that for any  $A$  and  $B$ ,  $A < a_1$  and  $b_0 < B < b_1$ . There exists  $\delta > 0$  such that for all  $(n, m) \in \mathcal{K}_{\mathcal{T}}$ ,  $m < n - \delta$ , whereby

$$\nabla \phi^\beta(x) \cdot u \geq c \mathbb{1}_{[0,1]}(|x|)|x|^{-n_0} - c \mathbb{1}_{(1,+\infty)}(|x|) \ln(|x|)|x|^{-m_0}$$

for some  $c > 0$ . The right-hand term of this inequality is regular nonintegrably divergent at the origin. Therefore, it is stable and (32) is proved.  $\square$

We can also deduce from Theorem 4 the same kind of result as in [18] concerning exponential models. Note, however, that the proof in [18] crucially relies on the convexity of  $\beta \mapsto \phi^\beta(x)$ , whereas the following corollary is obtained by different techniques.

COROLLARY 5. Assume that in (29) the family of pair potentials  $(\phi^\beta)_{\beta \in \mathcal{B}}$  is uniformly regular, nonintegrably divergent at the origin, satisfies the identifiability condition (33) and that

$$\phi^\beta(x) = \sum_{i=1}^p \beta_i \phi_i(x),$$

where  $\phi_1(x) \geq \chi(|x|)$  if  $|x| < r_0$  while for  $i \geq 2$ ,  $\phi_i(x) = o(\phi_1(x))$  when  $x \rightarrow 0$ . Let  $\mathcal{K}$  be a compact subset of  $\mathbb{R}_+ \times \mathcal{B}$  such that  $\theta^* = (z^*, \beta^*)$  belongs to  $\mathcal{K}$ . Then the MLE of  $(z^*, \beta^*)$  given by (8) is strongly consistent.

PROOF. In view of Theorem 4, we just need to check (30)–(32). Let  $\beta = (\beta_1, \dots, \beta_p) \in \mathcal{K}$  and  $\beta' = (\beta_1, \dots, \beta_{p-1}, \beta'_p) \in \mathcal{K}$  with  $\beta'_p \neq \beta_p$ . Since  $\phi^\beta$  is uniformly regular, we have  $|\phi^\beta(x) - \phi^{\beta'}(x)| \leq 2\psi(|x|)$ , meaning that  $|\phi_p(x)| \leq$

$c\psi(|x|)$  for some  $c > 0$ . We obtain likewise  $|\phi_i(x)| \leq c\psi(|x|)$  for any  $i = 1, \dots, p$ . Consequently,  $|\nabla\phi^\beta(x)| \leq \sum_{i=1}^p |\phi_i(x)| \leq pc\Psi(|x|)$  and (30) is proved.

The relation (31) holds trivially true by choosing  $r_0$  sufficiently small. Finally, letting  $U$  be any open set included in  $\mathcal{K}$ , we have  $\nabla\phi^\beta(x) \cdot u = \phi^u(x)$  which is stable by assumption.  $\square$

4.4. *Examples of infinite-body interactions.* Our general result in Theorem 1 is also adapted to nonpairwise potential models. For instance, it is not difficult to generalize the results of the three previous sections to the case of a finite-body interaction of order greater than two, as for instance a triplet or quadruplet interaction. For brevity reasons, we do not include this generalization in the present paper. Instead, we give in this section examples arising from stochastic geometry which involve infinite-body interactions. Specifically, we focus on the area-interaction and the Quermass-interaction processes, but other models could have been considered as well.

The area-interaction process [1], also called the Widom–Rowlinson model in the statistical physics community [28], is probably the most popular model of infinite-body interaction. For  $R \geq 0$ ,  $\omega \in \Omega$  and  $\Lambda$  a bounded set in  $\mathbb{R}^d$ , we introduce the notation

$$\mathcal{A}^R(\omega_\Lambda) = \lambda^d \left( \bigcup_{x \in \omega_\Lambda} B(x, R) \right).$$

The Hamiltonian  $H_\Lambda^\theta$  of the area-interaction process is defined for any  $\omega \in \Omega$  and any bounded set  $\Lambda$  by

$$(34) \quad H_\Lambda^\theta(\omega) = zN_\Lambda(\omega) + \beta(\mathcal{A}^R(\omega_{\Lambda \oplus B(0,2R)}) - \mathcal{A}^R(\omega_{\Lambda \oplus B(0,2R) \setminus \Lambda})),$$

where the parameter  $\theta = (z, R, \beta)$  belongs to  $\Theta = \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}$  and the operator  $\oplus$  stands for the Minkowski sum acting on the sets in  $\mathbb{R}^d$ . Note that the unknown radius  $R$  is part of the parameters and is estimated consistently as stated in the following proposition proved in [5].

PROPOSITION 1. *Let  $\mathcal{H} = (H^\theta)_\Lambda$  be the family of area-interaction energies defined in (34) and  $\mathcal{K}$  a compact subset of  $\Theta$  such that  $\theta^* = (z^*, R^*, \beta^*)$  belongs to  $\mathcal{K}$ . Then the MLE of  $\theta^*$  given by (8) is strongly consistent.*

The Quermass-interaction model is a generalization of the area-interaction process, where not only the volume of the union of balls is involved in the Hamiltonian but also the other Minkowski functionals. We denote by  $M_k, k = 1, \dots, d + 1$ , the  $d + 1$  Minkowski functionals in  $\mathbb{R}^d$  and for short

$$\mathcal{M}_k^R(\omega_\Lambda) = M_k \left( \bigcup_{x \in \omega_\Lambda} B(x, R) \right),$$

where  $R \geq 0$ ,  $\omega \in \Omega$  and  $\Lambda$  is a bounded subset of  $\mathbb{R}^d$ . Recall that for  $d = 2$ ,  $M_1$  corresponds to the Euler–Poincaré characteristic,  $M_2$  is the perimeter and  $M_3$  is the area. We refer to [2] for more details about Minkowski functionals.

The Hamiltonian  $H_\Lambda^\theta$  of the Quermass-interaction process is defined for any  $\omega \in \Omega$  and any bounded set  $\Lambda$  by

$$(35) \quad H_\Lambda^\theta(\omega) = zN_\Lambda(\omega) + \sum_{k=1}^{d+1} \beta_k(\mathcal{M}_k^R(\omega_{\Lambda \oplus B(0,2R)}) - \mathcal{M}_k^R(\omega_{\Lambda \oplus B(0,2R) \setminus \Lambda})),$$

where  $R \geq 0$ ,  $z > 0$  and  $\beta = (\beta_1, \dots, \beta_{d+1}) \in \mathbb{R}^{d+1}$ .

This model has been introduced in [15]. Its existence on  $\mathbb{R}^d$  has been solved so far only when  $d \leq 2$  in [3]. Therefore, we restrict the following study to the case  $d \leq 2$ . Moreover, we assume that  $R$  is known and we only consider the MLE estimation of  $\theta := (z, \beta)$  in (35). The reason is that we did not succeed to prove that (17) in [REGULARITY] holds in presence of the Euler–Poincaré characteristic  $M_1$  when  $R$  is part of the unknown parameter. With the assumption that  $R$  is known, the Quermass-interaction process becomes an infinite-body interactions exponential model and Corollary 1 applies. In this framework, the set of parameters is  $\Theta = \mathbb{R}_+ \times \mathbb{R}^3$ .

**PROPOSITION 2.** *Let  $\mathcal{H} = (H^\theta)_\Lambda$  be the family of Quermass-interaction energies defined in (35) for  $d = 2$  and let  $\mathcal{K}$  be a compact subset of  $\Theta$  such that  $\theta^* = (z^*, \beta^*)$  belongs to  $\mathcal{K}$ . Then the MLE of  $\theta^*$  given by (8) is strongly consistent.*

**5. Proofs.**

5.1. *Proof of Theorem 1.* The proof is organised as follows. In a first step (Lemma 3), we show that the hardcore parameter  $\hat{\delta}_n$  converges to  $\delta^*$ . So, according to Lemma 2, it remains to prove that the minimizer of  $\theta \rightarrow K_n^{\hat{\delta}_n, \theta}(\omega_{\Lambda_n}^*)$  converges to  $\theta^*$ . This result is guaranteed via a general lemma on the convergence of minimizers of possibly nonregular contrast functions (Lemma 4). The final step of the proof therefore consists in checking the assumptions of this lemma, namely:

- the almost sure convergence of the contrast functions to a lower semicontinuous function admitting  $\theta^*$  as a minimizer [assumptions (i), (ii), (iv) of Lemma 4],
- the control, when  $n$  goes to infinity, of the infimum of contrast functions evaluated on small balls with respect to the minimum of contrast functions evaluated only on a finite number of points [assumption (v)].

Note that Lemma 4 and the verification of its assumptions in the setting of Theorem 1 and for all models considered in Section 4 are the main contributions of the present paper.

Let us now enter into the details of the proof. Let  $P$  be a probability measure in  $\mathcal{G}^{\delta^*, \theta^*}$  and let  $\omega^*$  be a realisation of  $P$ . If  $P$  is not ergodic, it can be represented as the mixture of ergodic stationary Gibbs measures; see [23]. Therefore, the proof of the consistency of the MLE reduces to the case when  $P$  is ergodic, which is assumed henceforth.

Let us start by proving the convergence of the hardcore parameter.

LEMMA 3. *The MLE  $\hat{\delta}_n$  converges  $P$ -almost surely to  $\delta^*$  when  $n$  goes to infinity.*

PROOF. From Lemma 1,  $\hat{\delta}_n = \min(\tilde{\delta}_n, \delta_{\max})$  where  $\tilde{\delta}_n$  is given by (10). Clearly,  $\tilde{\delta}_n$  is a decreasing sequence and  $\tilde{\delta}_n \geq \delta^*$ . So it remains to prove that, for any  $\delta' > \delta^*$ ,  $\tilde{\delta}_n$  is smaller than  $\delta'$  for  $n$  large enough. By the DLR equations (6) and the definition of local densities (7), there exists a bounded set  $\Lambda$  such that

$$P(\text{there exist } x, y \in \omega_\Lambda \text{ such that } \delta^* < |x - y| < \delta') > 0.$$

Thanks to the ergodic theorem, this implies that

$$P(\text{there exist } n \geq 1 \text{ and } x, y \in \omega_{\Lambda_n} \text{ such that } \delta^* < |x - y| < \delta') = 1$$

which proves the expected result.  $\square$

According to Lemmas 2 and 3, we have now to prove the consistency of  $\hat{\theta}_n$  where

$$\hat{\theta}_n = \operatorname{argmin}_{\theta \in \mathcal{K}} K_n^{\hat{\delta}_n, \theta}(\omega_{\Lambda_n}^*).$$

The strong consistency relies on a minimum contrast function result stated in the next lemma, where neither the contrast function nor its limit need to be continuous with respect to  $\theta$ . Its proof is given in the Supplementary Material [5].

LEMMA 4. *For any  $\theta \in \Theta$ , let  $(h_n^\theta)_{n \geq 1}$  be a family of parametric measurable functions from  $\Omega_T$  to  $\mathbb{R}$ . Let  $\mathcal{K}$  be a compact subset of  $\Theta$  and let  $P$  be a probability measure on  $\Omega_T$ . We assume that:*

- (i) *For any  $\theta \in \Theta$ ,  $h_n^\theta$  converges  $P$ -almost surely to a finite real number (denoted by  $h^\theta$ ) when  $n$  goes to infinity.*
- (ii) *The function  $\theta \mapsto h^\theta$  admits a unique minimum over  $\Theta$ :*

$$\theta^* = \operatorname{argmin}_{\theta \in \Theta} h^\theta,$$

and  $\theta^* \in \mathcal{K}$ .

- (iii) *For  $P$ -almost every  $\omega \in \Omega_T$  and for  $n$  sufficiently large,  $\theta \mapsto h_n^\theta$  admits an infimum over  $\mathcal{K}$  attained by at least one element  $\hat{\theta}_n(\omega)$ .*

(iv) The function  $\theta \mapsto h^\theta$  is lower semicontinuous on  $\Theta$  (i.e., for any  $\theta \in \Theta$ ,  $\liminf_{\theta' \rightarrow \theta} h^{\theta'} \geq h^\theta$ ).

(v) There exists a function  $g_0$  from  $\mathbb{R}_+$  to  $\mathbb{R}_+$  satisfying

$$\lim_{x \rightarrow 0} g_0(x) = 0$$

and such that for any  $\varepsilon > 0$  and for any  $\theta \in K$ , there exists a finite subset  $\mathcal{N}(\theta)$  included in  $B(\theta, g_0(\varepsilon)) \cap \Theta$  and  $r(\theta) > 0$  such that

$$(36) \quad P\left(\limsup_{n \rightarrow \infty} \left( \min_{\theta' \in \mathcal{N}(\theta)} h_n^{\theta'} - \inf_{\theta' \in B(\theta, r(\theta)) \cap K} h_n^{\theta'} \right) \geq \varepsilon\right) = 0.$$

Then the sequence  $\hat{\theta}_n = \operatorname{argmin}_{\theta \in K} h_n(\theta)$  converges  $P$ -almost surely to  $\theta^*$  when  $n$  goes to infinity.

Let us show that the assumptions (i), (ii), (iii), (iv) and (v) of Lemma 4 hold for the family of contrast functions

$$h_n^\theta(\omega_{\Lambda_n}^*) = K_n^{\hat{\delta}_n, \theta}(\omega_{\Lambda_n}^*) = \frac{\ln(Z_{\Lambda_n}^{\hat{\delta}_n, \theta})}{|\Lambda_n|} + \frac{H_{\Lambda_n}^\theta(\omega_{\Lambda_n}^*)}{|\Lambda_n|}.$$

To prove (i), first note, from [MEANENERGY], [BOUNDARY] and the ergodic theorem, that for any  $\theta \in \Theta$  and  $P$ -almost all  $\omega$

$$\lim_{n \rightarrow \infty} \frac{1}{|\Lambda_n|} H_{\Lambda_n}^\theta(\omega_{\Lambda_n}) = E_P(H_0^\theta) = H^\theta(P).$$

Second, recall from assumption [VARPRIN] that for any  $\delta \in I$  and  $\theta \in \Theta$ ,  $\ln(Z_{\Lambda_n}^{\delta, \theta})/|\Lambda_n|$  converges to  $p(\delta, \theta)$  and  $\delta \mapsto p(\delta, \theta)$  is right continuous. On the other hand, from Lemma 3,  $\hat{\delta}_n \rightarrow \delta^*$  almost surely with  $\hat{\delta}_n > \delta^*$ . Let  $\theta \in \Theta$ ,  $\epsilon > 0$  and  $\delta^+ > \delta^*$  such that  $|p(\delta^*, \theta) - p(\delta^+, \theta)| < \epsilon$ . Let  $n$  be sufficiently large so that  $\hat{\delta}_n < \delta^+$ ,  $|\ln(Z_{\Lambda_n}^{\delta^+, \theta})/|\Lambda_n| - p(\delta^+, \theta)| < \epsilon$  and  $|\ln(Z_{\Lambda_n}^{\delta^*, \theta})/|\Lambda_n| - p(\delta^*, \theta)| < \epsilon$ . Then, since  $\delta \mapsto Z_{\Lambda_n}^{\delta, \theta}$  is decreasing, we have

$$\ln(Z_{\Lambda_n}^{\hat{\delta}_n, \theta})/|\Lambda_n| - p(\delta^*, \theta) \leq \ln(Z_{\Lambda_n}^{\delta^*, \theta})/|\Lambda_n| - p(\delta^*, \theta) < \epsilon$$

and

$$\begin{aligned} p(\delta^*, \theta) - \ln(Z_{\Lambda_n}^{\hat{\delta}_n, \theta})/|\Lambda_n| &\leq p(\delta^*, \theta) - p(\delta^+, \theta) + p(\delta^+, \theta) - \ln(Z_{\Lambda_n}^{\delta^+, \theta})/|\Lambda_n| \\ &< 2\epsilon, \end{aligned}$$

proving that  $\ln(Z_{\Lambda_n}^{\hat{\delta}_n, \theta})/|\Lambda_n|$  converges to  $p(\delta^*, \theta)$  almost surely.

Hence, for any  $\theta \in \Theta$  and for  $P$ -almost every  $\omega^*$

$$\lim_{n \rightarrow \infty} h_n^\theta(\omega_{\Lambda_n}^*) = h^\theta,$$

where

$$h^\theta = p(\delta^*, \theta) + H^\theta(P)$$

which proves assumption (i).

To prove assumption (ii), note from [VARPRIN] that for any  $\theta \in \Theta$ , any  $P^{\delta^*,\theta} \in \mathcal{G}^{\delta^*,\theta}$  and any  $P^{\delta^*,\theta^*} \in \mathcal{G}^{\delta^*,\theta^*}$

$$p(\delta^*, \theta) \geq -\mathcal{I}(P^{\delta^*,\theta^*}) - H^\theta(P^{\delta^*,\theta^*})$$

while  $-\mathcal{I}(P^{\delta^*,\theta^*}) = p(\delta^*, \theta^*) + H^\theta(P^{\delta^*,\theta^*})$  so that

$$h^\theta \geq h^{\theta^*}.$$

From [VARPRIN], the equality holds if and only if  $\theta = \theta^*$  which proves assumption (ii).

Assumption (iii) is given by Lemma 2.

The function  $\theta \mapsto H^\theta(P)$  is continuous thanks to (16) in assumption [REGULARITY]. By assumption [VARPRIN], we have for any  $\theta$  and  $\theta'$  in  $\Theta$

$$p(\delta^*, \theta') \geq p(\delta^*, \theta) + H^\theta(P^{\delta^*,\theta}) - H^{\theta'}(P^{\delta^*,\theta}),$$

where  $P^{\delta^*,\theta} \in \mathcal{G}^{\delta^*,\theta}$ . By continuity of  $\theta \mapsto H^\theta(P)$ , it follows that  $\theta \mapsto p(\delta^*, \theta)$  is lower semicontinuous. Therefore,  $\theta \mapsto h^\theta$  is lower semicontinuous and assumption (iv) holds.

It remains to prove assumption (v). Let us start with some preliminary results.

Thanks to [MEANENERGY], [BOUNDARY], (16) in [REGULARITY] and the ergodic theorem, we have that, for any  $\theta \in \Theta$ , any  $r > 0$  and  $P$ -almost every  $\omega$ ,

$$(37) \quad \limsup_{n \rightarrow \infty} \frac{1}{|\Lambda_n|} \sup_{\substack{\theta' \in \mathcal{K} \\ |\theta - \theta'| < r}} |H_{\Lambda_n}^\theta(\omega_{\Lambda_n}) - H_{\Lambda_n}^{\theta'}(\omega_{\Lambda_n})| \leq g(r).$$

This inequality allows us to control the variation of the infimum of the specific energy in the contrast functions. The following lemma deals with the variation of the finite volume pressure.

**LEMMA 5.** *For any  $\eta > 0$  and any  $\theta_0 \in \mathcal{K}$ , there exists a finite subset  $\mathcal{N}(\theta_0) \subset B(\theta_0, \eta) \cap \Theta$  and  $0 < r(\theta_0) < \eta$  such that for  $n \geq 1$  and any  $\delta \in I$*

$$(38) \quad \min_{\theta \in \mathcal{N}(\theta_0)} \frac{1}{|\Lambda_n|} \ln(Z_{\Lambda_n}^{\delta,\theta}) - \inf_{\theta \in B(\theta_0, r(\theta_0)) \cap \mathcal{K}} \frac{1}{|\Lambda_n|} \ln(Z_{\Lambda_n}^{\delta,\theta}) \leq e^{\kappa+1} g(r(\theta_0)),$$

where  $g$  comes from assumption [REGULARITY] and  $\kappa$  from assumption [STABILITY].

**Proof.** For any  $\delta \in I$  and any  $\theta, \theta'$  in  $\Theta$

$$\begin{aligned} \ln(Z_{\Lambda_n}^{\delta,\theta'}) - \ln(Z_{\Lambda_n}^{\delta,\theta}) &= \ln\left(\int \frac{1}{Z_{\Lambda_n}^{\delta,\theta}} e^{-H_{\Lambda_n}^{\theta'}(\omega_{\Lambda_n})} \mathbb{1}_{\Omega_\infty^\delta}(\omega_{\Lambda_n}) \pi_{\Lambda_n}(d\omega_{\Lambda_n})\right) \\ (39) \quad &= \ln\left(\int e^{-H_{\Lambda_n}^{\theta'}(\omega_{\Lambda_n}) + H_{\Lambda_n}^\theta(\omega_{\Lambda_n})} \mathbb{1}_{\Omega_\infty^\delta}(\omega_{\Lambda_n}) P_{\Lambda_n}^{\delta,\theta}(d\omega_{\Lambda_n})\right) \\ &= \ln E_{P_{\Lambda_n}^{\delta,\theta}}(e^{-H_{\Lambda_n}^{\theta'} + H_{\Lambda_n}^\theta}), \end{aligned}$$

where  $P_{\Lambda_n}^{\delta,\theta}$  is the probability measure  $(1/Z_{\Lambda_n}^{\delta,\theta})e^{-H_{\Lambda_n}^\theta} \mathbb{1}_{\Omega_\infty^\delta} \pi_{\Lambda_n}$ . By Jensen’s inequality, we have that

$$\ln(Z_{\Lambda_n}^{\delta,\theta'}) - \ln(Z_{\Lambda_n}^{\delta,\theta}) \geq E_{P_{\Lambda_n}^{\delta,\theta}}(H_{\Lambda_n}^\theta - H_{\Lambda_n}^{\theta'}).$$

For any  $\eta > 0$  and  $\theta_0$  in  $\mathcal{K}$ , we choose  $\mathcal{N}(\theta_0)$  and  $r(\theta_0) > 0$  as in (17). Denoting  $[x]_+ = \max(x, 0)$  and writing for short  $\mathcal{N}$  for  $\mathcal{N}(\theta_0)$  and  $B$  for  $B(\theta_0, r(\theta_0)) \cap \mathcal{K}$ , we obtain, for any  $\delta \geq \delta_{\min}$ ,

$$\begin{aligned} & \min_{\theta \in \mathcal{N}} \frac{1}{|\Lambda_n|} \ln(Z_{\Lambda_n}^{\delta,\theta}) - \inf_{\theta \in B} \frac{1}{|\Lambda_n|} \ln(Z_{\Lambda_n}^{\delta,\theta}) \\ &= \min_{\theta \in \mathcal{N}} \sup_{\theta' \in B} \frac{1}{|\Lambda_n|} (\ln(Z_{\Lambda_n}^{\delta,\theta}) - \ln(Z_{\Lambda_n}^{\delta,\theta'})) \\ &\leq \min_{\theta \in \mathcal{N}} \sup_{\theta' \in B} \frac{1}{|\Lambda_n|} (E_{P_{\Lambda_n}^{\delta,\theta}}(H_{\Lambda_n}^{\theta'} - H_{\Lambda_n}^\theta)) \\ &= \frac{1}{|\Lambda_n|} \min_{\theta \in \mathcal{N}} \sup_{\theta' \in B} E_{P_{\Lambda_n}^{\delta,\theta}} \left( \frac{H_{\Lambda_n}^{\theta'} - H_{\Lambda_n}^\theta}{N_{\Lambda_n}} \right) \\ &\leq \frac{1}{|\Lambda_n|} \min_{\theta \in \mathcal{N}} \sup_{\theta' \in B} \sup_{\omega_{\Lambda_n} \in \Omega_\infty^\delta} \left[ \frac{H_{\Lambda_n}^{\theta'}(\omega_{\Lambda_n}) - H_{\Lambda_n}^\theta(\omega_{\Lambda_n})}{N_{\Lambda_n}(\omega_{\Lambda_n})} \right]_+ E_{P_{\Lambda_n}^{\delta,\theta}}(N_{\Lambda_n}) \\ &\leq \frac{1}{|\Lambda_n|} \max_{\theta \in \mathcal{N}} E_{P_{\Lambda_n}^{\delta,\theta}}(N_{\Lambda_n}) \left[ \min_{\theta \in \mathcal{N}} \sup_{\theta' \in B} \sup_{\omega_{\Lambda_n} \in \Omega_\infty^\delta} \left( \frac{H_{\Lambda_n}^{\theta'}(\omega_{\Lambda_n}) - H_{\Lambda_n}^\theta(\omega_{\Lambda_n})}{N_{\Lambda_n}(\omega_{\Lambda_n})} \right) \right]_+ \\ &= \frac{1}{|\Lambda_n|} \max_{\theta \in \mathcal{N}} E_{P_{\Lambda_n}^{\delta,\theta}}(N_{\Lambda_n}) \left[ -\max_{\theta \in \mathcal{N}} \inf_{\theta' \in B} \inf_{\omega_{\Lambda_n} \in \Omega_\infty^\delta} \left( \frac{H_{\Lambda_n}^\theta(\omega_{\Lambda_n}) - H_{\Lambda_n}^{\theta'}(\omega_{\Lambda_n})}{N_{\Lambda_n}(\omega_{\Lambda_n})} \right) \right]_+. \end{aligned}$$

Using [REGULARITY], we therefore obtain

$$(40) \quad \min_{\theta \in \mathcal{N}} \frac{1}{|\Lambda_n|} \ln(Z_{\Lambda_n}^{\delta,\theta}) - \inf_{\theta \in B} \frac{1}{|\Lambda_n|} \ln(Z_{\Lambda_n}^{\delta,\theta}) \leq \frac{1}{|\Lambda_n|} g(r(\theta_0)) \max_{\theta \in \mathcal{N}} E_{P_{\Lambda_n}^{\delta,\theta}}(N_{\Lambda_n}).$$

Let us control  $E_{P_{\Lambda_n}^{\delta,\theta}}(N_{\Lambda_n})$  by entropy inequalities. By definition of the entropy, the assumption [STABILITY] and the standard inequality  $Z_{\Lambda_n}^{\delta,\theta} \geq \pi_{\Lambda_n}(\{\emptyset\}) = e^{-|\Lambda_n|}$ , we find that

$$\begin{aligned} \mathcal{I}_{\Lambda_n}(P_{\Lambda_n}^{\delta,\theta}) &= \int \ln \left( \frac{e^{-H_{\Lambda_n}^\theta} \mathbb{1}_{\Omega_\infty^\delta}(\omega_{\Lambda_n})}{Z_{\Lambda_n}^{\delta,\theta}} \right) P_{\Lambda_n}^{\delta,\theta} \\ (41) \quad &\leq -\ln(Z_{\Lambda_n}^{\delta,\theta}) - \int H_{\Lambda_n}^\theta P_{\Lambda_n}^{\delta,\theta} \\ &\leq |\Lambda_n| + \kappa E_{P_{\Lambda_n}^{\delta,\theta}}(N_{\Lambda_n}). \end{aligned}$$



The entropy is also characterised by

$$\mathcal{I}_{\Lambda_n}(P_{\Lambda_n}^{\delta,\theta}) = \sup_g \left( \int g P_{\Lambda_n}^{\delta,\theta} - \ln \left( \int e^g \pi_{\Lambda_n} \right) \right),$$

where the supremum is over all bounded measurable functions  $g$ ; see, for example, [27]. Choosing  $g = (\kappa + 1)N_{\Lambda_n} \mathbb{1}_{N_{\Lambda_n} \leq C}$  for any constant  $C > 0$ , we get

$$\begin{aligned} \mathcal{I}_{\Lambda_n}(P_{\Lambda_n}^{\delta,\theta}) &\geq (\kappa + 1)E_{P_{\Lambda_n}^{\delta,\theta}}(N_{\Lambda_n} \mathbb{1}_{N_{\Lambda_n} \leq C}) - \ln \left( \int e^{(\kappa+1)N_{\Lambda_n} \mathbb{1}_{N_{\Lambda_n} \leq C}} \pi_{\Lambda_n} \right) \\ (42) \quad &\geq (\kappa + 1)E_{P_{\Lambda_n}^{\delta,\theta}}(N_{\Lambda_n} \mathbb{1}_{N_{\Lambda_n} \leq C}) - (e^{\kappa+1} - 1)|\Lambda_n|. \end{aligned}$$

Let  $C$  goes to infinity and combine (41) and (42), we obtain that for any  $(\delta, \theta)$

$$(43) \quad E_{P_{\Lambda_n}^{\delta,\theta}}(N_{\Lambda_n}) \leq e^{\kappa+1}|\Lambda_n|.$$

The inequalities (40) and (43) imply (38).

We are now in position to prove (v). The function  $g_0$  in (v) depends on  $g$  in (37) and Lemma 5 as follows. For any  $x > 0$ , we choose  $g_0(x) > 0$  such that

$$g(2g_0(x)) + e^{\kappa+1}g(g_0(x)) < x.$$

This choice is always possible since  $g(u)$  tends to 0 when  $u$  goes to 0.

Let  $\varepsilon > 0$  and  $\theta_0 \in \mathcal{K}$  and consider the finite subset  $\mathcal{N}(\theta_0) \subset B(\theta_0, \eta) \cap \Theta$  and the positive number  $r(\theta_0)$ , with  $r(\theta_0) < \eta$ , given by Lemma 5 where  $\eta = g_0(\varepsilon)$ . In the following, we write for short  $\mathcal{N}$  for  $\mathcal{N}(\theta_0)$  and  $B$  for  $B(\theta_0, r(\theta_0)) \cap \mathcal{K}$ . Using (37), Lemma 5 and the definition of  $g_0$ , we have that for  $P$ -almost every  $\omega$ ,

$$\begin{aligned} &\limsup_{n \rightarrow \infty} \left( \min_{\theta \in \mathcal{N}} h_n^\theta(\omega_{\Lambda_n}) - \inf_{\theta \in B} h_n^\theta(\omega_{\Lambda_n}) \right) \\ &= \limsup_{n \rightarrow \infty} \frac{1}{|\Lambda_n|} \left( \min_{\theta \in \mathcal{N}} (H_{\Lambda_n}^\theta(\omega_{\Lambda_n}) + \ln(Z_{\Lambda_n}^{\hat{\delta}_n, \theta})) \right. \\ &\quad \left. - \inf_{\theta \in B} (H_{\Lambda_n}^\theta(\omega_{\Lambda_n}) + \ln(Z_{\Lambda_n}^{\hat{\delta}_n, \theta})) \right) \\ &\leq \limsup_{n \rightarrow \infty} \frac{1}{|\Lambda_n|} \left( \max_{\theta \in \mathcal{N}} H_{\Lambda_n}^\theta(\omega_{\Lambda_n}) - \inf_{\theta \in B} H_{\Lambda_n}^\theta(\omega_{\Lambda_n}) \right) \\ &\quad + \min_{\theta \in \mathcal{N}} \ln(Z_{\Lambda_n}^{\hat{\delta}_n, \theta}) - \inf_{\theta \in B} \ln(Z_{\Lambda_n}^{\hat{\delta}_n, \theta}) \\ &\leq \max_{\theta \in \mathcal{N}} \limsup_{n \rightarrow \infty} \frac{1}{|\Lambda_n|} \sup_{\theta' \in B} |H_{\Lambda_n}^\theta(\omega_{\Lambda_n}) - H_{\Lambda_n}^{\theta'}(\omega_{\Lambda_n})| \\ &\quad + \limsup_{n \rightarrow \infty} \frac{1}{|\Lambda_n|} \left( \min_{\theta \in \mathcal{N}} \ln(Z_{\Lambda_n}^{\hat{\delta}_n, \theta}) - \inf_{\theta \in B} \ln(Z_{\Lambda_n}^{\hat{\delta}_n, \theta}) \right) \\ &\leq g(2g_0(\varepsilon)) + e^{\kappa+1}g(g_0(\varepsilon)) < \varepsilon, \end{aligned}$$

which proves (v).

**Acknowledgements.** The authors are grateful to Shigeru Mase for supplying his unpublished manuscript [18]. They also thank Jean-Francois Coeurjolly for fruitful discussions and the anonymous referees for their suggestions which helped improve this paper.

## SUPPLEMENTARY MATERIAL

**Supplement to “Consistency of likelihood estimation for Gibbs point processes”** (DOI: [10.1214/16-AOS1466SUPP](https://doi.org/10.1214/16-AOS1466SUPP); .pdf). This supplementary material provides the proofs of Lemma 4, Corollary 1, Theorems 2–4 and Propositions 1 and 2.

## REFERENCES

- [1] BADDELEY, A. J. and VAN LIESHOUT, M. N. M. (1995). Area-interaction point processes. *Ann. Inst. Statist. Math.* **47** 601–619. [MR1370279](#)
- [2] CHIU, S. N., STOYAN, D., KENDALL, W. S. and MECKE, J. (2013). *Stochastic Geometry and Its Applications*, 3rd ed. Wiley, Chichester. [MR3236788](#)
- [3] DEREUDRE, D. (2009). The existence of Quermass-interaction processes for nonlocally stable interaction and nonbounded convex grains. *Adv. in Appl. Probab.* **41** 664–681. [MR2571312](#)
- [4] DEREUDRE, D. (2016). Variational principle for Gibbs point processes with finite range interaction. *Electron. Commun. Probab.* **21** Paper No. 10, 11. [MR3485379](#)
- [5] DEREUDRE, D. and LAVANCIER, F. (2016). Supplement to “Consistency of likelihood estimation for Gibbs point processes.” DOI:[10.1214/16-AOS1466SUPP](https://doi.org/10.1214/16-AOS1466SUPP).
- [6] DIGGLE, P. J., FIKSEL, T., GRABARNIK, P., OGATA, Y., STOYAN, D. and TANEMURA, M. (1994). On parameter estimation for pairwise interaction point processes. In *International Statistical Review/Revue Internationale de Statistique* 99–117.
- [7] DIGGLE, P. J. and GRATTON, R. J. (1984). Monte Carlo methods of inference for implicit statistical models. *J. Roy. Statist. Soc. Ser. B* **46** 193–227. [MR0781880](#)
- [8] GEORGII, H.-O. (1988). *Gibbs Measures and Phase Transitions. de Gruyter Studies in Mathematics* **9**. de Gruyter, Berlin. [MR0956646](#)
- [9] GEORGII, H.-O. (1994). Large deviations and the equivalence of ensembles for Gibbsian particle systems with superstable interaction. *Probab. Theory Related Fields* **99** 171–195. [MR1278881](#)
- [10] GEORGII, H.-O. (1995). The equivalence of ensembles for classical systems of particles. *J. Stat. Phys.* **80** 1341–1378. [MR1349785](#)
- [11] GEYER, C. J. and MØLLER, J. (1994). Simulation procedures and likelihood inference for spatial point processes. *Scand. J. Statist.* **21** 359–373. [MR1310082](#)
- [12] GUYON, X. (1995). *Random Fields on a Network: Modeling, Statistics, and Applications*. Springer, New York. [MR1344683](#)
- [13] ILLIAN, J., PENTTINEN, A., STOYAN, H. and STOYAN, D. (2008). *Statistical Analysis and Modelling of Spatial Point Patterns*. Wiley, Chichester. [MR2384630](#)
- [14] JENSEN, J. L. (1993). Asymptotic normality of estimates in spatial point processes. *Scand. J. Statist.* **20** 97–109. [MR1229287](#)
- [15] KENDALL, W. S., VAN LIESHOUT, M. N. M. and BADDELEY, A. J. (1999). Quermass-interaction processes: Conditions for stability. *Adv. in Appl. Probab.* **31** 315–342. [MR1724554](#)

- [16] KÜNSCH, H. (1981). Thermodynamics and statistical analysis of Gaussian random fields. *Z. Wahrsch. Verw. Gebiete* **58** 407–421. [MR0639149](#)
- [17] MASE, S. (1992). Uniform LAN condition of planar Gibbsian point processes and optimality of maximum likelihood estimators of soft-core potential functions. *Probab. Theory Related Fields* **92** 51–67. [MR1156450](#)
- [18] MASE, S. (2002). Asymptotic properties of MLE's of Gibbs models on  $\mathbb{R}^d$ . Unpublished manuscript.
- [19] MATEU, J. and MONTES, F. (2001). Likelihood inference for Gibbs processes in the analysis of spatial point patterns. *Int. Stat. Rev.* **69** 81–104.
- [20] MØLLER, J. and WAAGEPETERSEN, R. (2003). *Statistical Inference and Simulation for Spatial Point Processes*. Chapman & Hall, Boca Raton.
- [21] OGATA, Y. and TANEMURA, M. (1984). Likelihood analysis of spatial point patterns. *J. Roy. Statist. Soc. Ser. B* **46** 496–518. [MR0790635](#)
- [22] PICKARD, D. K. (1987). Inference for discrete Markov fields: The simplest nontrivial case. *J. Amer. Statist. Assoc.* **82** 90–96. [MR0883337](#)
- [23] PRESTON, C. (1976). *Random Fields*. Springer, Berlin. [MR0448630](#)
- [24] RUELLE, D. (1969). *Statistical Mechanics: Rigorous Results*. Benjamin, New York. [MR0289084](#)
- [25] RUELLE, D. (1970). Superstable interactions in classical statistical mechanics. *Comm. Math. Phys.* **18** 127–159. [MR0266565](#)
- [26] VAN LIESHOUT, M. N. M. (2000). *Markov Point Processes and Their Applications*. World Scientific, Singapore.
- [27] VARADHAN, S. R. S. (1988). Large deviations and applications. In *École D'Été de Probabilités de Saint-Flour XV–XVII, 1985–87. Lecture Notes in Math.* **1362** 1–49. Springer, Berlin. [MR0983371](#)
- [28] WIDOM, B. and ROWLINSON, J. S. (1970). New model for the study of liquid-vapor phase transitions. *J. Chem. Phys.* **52** 1670–1684.

LABORATOIRE PAUL PAINLEVÉ  
UNIVERSITÉ LILLE 1  
CITÉ SCIENTIFIQUE  
59655 VILLENEUVE D'ASCQ CEDEX  
FRANCE  
E-MAIL: david.dereudre@math.univ-lille1.fr

LABORATOIRE DE MATHÉMATIQUES JEAN LERAY  
UNIVERSITÉ DE NANTES  
BP 92208  
2, RUE DE LA HOUSSINIÈRE  
F-44322 NANTES CEDEX 03  
FRANCE  
E-MAIL: frederic.lavancier@univ-nantes.fr