Probab. Theory Relat. Fields 92, 51–67 (1992)

Probability Theory and Related Fields

Uniform LAN condition of planar Gibbsian point processes and optimality of maximum likelihood estimators of soft-core potential functions

Shigeru Mase

Faculty of Integrated Arts and Sciences, Hiroshima University, Hiroshima, 730 Japan

Received October 23, 1989; in revised form July 29, 1991

Summary. A parametric model of planar point patterns in a bounded region is constructed using grand canonical Gibbsian point processes with soft-core potential functions. A simple and explicit condition that this model becomes a uniform locally asymptotic normal (ULAN) family will be given. From this result we can conclude that the maximum likelihood estimator of the potential function is asymptotically efficient for a wide class of loss functions.

1 Introduction

The statistical analysis of spatial point patterns has made remarkable advances as summarized in the book of Stoyan et al. (1987) or Ripley (1988). This is mainly due to the progress of the theory of spatial point process as a mathematical framework. Through these researches the importance of the Gibbsian point process as a model construction principle has been widely recognized. Besides the fact that it has a long history as a genuine physical model of point systems in equilibrium, an attractive feature of the Gibbsian process is the fact that it can offer a variety of complex point patterns starting from a simple object, potential functions, which is directly interpretable as quantitative measure of attractive and/or repulsive forces acting among points. There have been many attempts to estimate the potential function from point pattern data. The first is Ripley (1977) who used a trial-and-error method based on graphical summary statistics. Ogata and Tanemura have developed the maximum likelihood method based on a numerical approximation of likelihoods in a series of papers, see, e.g., Ogata and Tanemura (1981, 84, 85). A similar idea was in Penttinen (1984). Moyeed and Baddeley (1989) proposed an iterative procedure of estimating the maximum likelihood estimator. Moment-method type estimators were proposed by several authors, see, e.g., Glöetzl and Rauschenschwandtner (1981); Hanisch and Stoyan (1983); Fiksel (1984), and Takacs (1986). A non-parametric estimator based on Percus and Yevick approximation was given by Diggle et al. (1987). Finally an estimation procedure using the idea of pseudo-likelihood functions was proposed by Besag et al. (1982). However justifications of these estimators are almost all intuitive and a simulation study is the last resource. Except the lattice case, see, e.g., Pickard (1976, 1977, 1979, 1982), Mase (1984); Ji (1989); Younes (1989), and Gidas (1991), there are few studies of theoretical properties of estimators. This is mainly because Gibbsian point processes have a discouragingly complex structure, combinatorial in nature, which makes direct theoretical manipulations quite difficult. Although this may seem a great handicap for Gibbsian processes to be good statistical models, note that all statistical models of point patterns must inevitably have complex structures if they should represent all the interactions acting among points.

In this research we will give a rigorous basis of analyzing statistical properties of planar Gibbsian point processes and will prove the efficiency of maximum likelihood estimators of potential functions. (As papers of Ogata and Tanemura show, the calculation of maximum likelihood estimator is rather difficult and some numerical approximations are indispensable.) The class of potential functions discussed in this research is so-called soft-core potentials, that is, points can be arbitrarily close to one another. We also assume that potential functions are translation-invariant, isotropic, and of Lennard-Jones type. With each such potential function we associate a two- or three-parameters model of grand canonical (i.e., with random number of points) Gibbsian point processes on a bounded planar region G. The parameter $\theta = (z, \alpha, \beta)$, or $\theta = (z, \alpha)$ for two-parameters case, has natural meanings (at least in physical context), that is, z stands for the chemical potential, α for the inverse temperature, and β for the scale parameter. From an observation X_G of coordinates of points on G, the maximum likelihood estimator (MLE) $\hat{\theta}$ is defined.

We are interested in the asymptotic behavior of MLE's as the region G expands to $\mathbb{R}^2$ monotonically. We will give explicit and simple conditions on the potential function which guarantees the asymptotic efficiency of MLE. Our argument is based on two theories. One is the theory of asymptotic expansion of the cluster integral due to Minlos and Pogosian. Their result is used as a technical tool of detailed analysis of grand partition functions, that is, normalizing constants of grand canonical Gibbsian processes.

Another is the theory of locally asymptotic normal (LAN) families due to Le Cam. The LAN condition of Le Cam singles out a simple and essential statistical condition which guarantees a fine asymptotic statistical theory. It has been used as a basic tool of asymptotic theory of statistical models with complicated structures such as signal process with Gaussian noise, diffusion process model, and one-dimensional point process model, see, e.g., Kutoyants (1984). Also the LAN condition is examined for Gibbsian processes on two-dimensional lattice in Mase (1984) and on one-dimensional lattice in Ji (1989). In this reasearch we will use in particular the ULAN (uniformly locally asymptotic normal) condition, a variant of the LAN condition with uniformity, and relevant optimality results given in the book of Ibragimov and Has'minskii (1981).

2 Gibbsian point process

Let $(x)_n$ denote an unordered n-ple $(x_1, \ldots, x_n)$. For each Borel set $G \subset \mathbb{R}^2$, $\mathscr{C}(G)$ denotes the space of all locally finite subsets (configurations) of G. The space $\mathscr{C}(G)$ is the direct sum of $\mathscr{C}_n(G) \equiv \{(x)_n; x_i \in G\}$, $n \geqq 0$. We should set $\mathscr{C}_0(G) = \{\emptyset\}$. The set $\mathscr{C}_n(G)$ is canonically identified with the quotient space $G^n/\sim$ with respect to permutation of coordinates. Hence the Lebesgue measure $dx_1 \ldots dx_n$ induces the Lebesgue measure dc on $\mathscr{C}(G)$ canonically. It is convenient to let dc be the unit point mass at $\emptyset$ on $\mathscr{C}_0(G)$. Each function $\phi(c)$ on $\mathscr{C}(G)$ is canonically identified with symmetric functions $\phi(x)_n$ on G^n, $n \geqq 0$, and hence

$$\int_{\mathscr{C}(G)} \phi(c)\,dc = \phi(\emptyset) + \sum_{n \geqq 1} \frac{1}{n!} \int_{G^n} \phi(x)_n \, dx_1 \ldots dx_n \;,$$

where, and in the following, the notation like $f(x)_n$ means $f((x)_n)$. A pair potential function $\Phi(x)$ is an even upper semi-continuous function defined on $\mathbb{R}^2$ which may take the value $+\infty$. The interaction energy $U_\Phi(x)_n$ and the local energy $E_{z,\Phi}(x)_n$ of a configuration $(x)_n$ are defined by

$$U_\Phi(x)_n = \sum_{1 \leqq i < j \leqq n} \Phi(x_i - x_j) \;,$$

$$E_{z,\Phi}(x)_n = nz + U_\Phi(x)_n \;,$$

where z is a constant called the chemical potential. The constant e^{-z} is called the activity or fugacity. We should set $U_\Phi(c) = 0$ if $\# c \leqq 1$. A potential function is stable if there is a constant B (stability constant) such that $U_\Phi(c) \geqq -B \# c$ for all finite $c \in \mathscr{C}(\mathbb{R}^2)$. The stability is known to be equivalent to the finiteness of the integral

$$\Xi_{G,z,\Phi} = \int_{\mathscr{C}(G)} e^{-E_{z,\Phi}(c)} dc$$

for each bounded G, see Ruelle (1969, Proposition 3.2.2). The integral $\Xi_{G,z,\Phi}$ is called the grand partition function.

Let X_G denote the random element taking values in $\mathscr{C}(G)$. The Gibbsian point process $\mathbf{P}_{G,z,\Phi}$ is defined by the formula

$$\int \phi(X_G)\,d\mathbf{P}_{G,z,\Phi} = \frac{1}{\Xi_{G,z,\Phi}} \int_{\mathscr{C}(G)} \phi(c) e^{-E_{z,\Phi}(c)} dc \;.$$

If $z = 0$ and $\Phi \equiv 0$ the Gibbsian point process is the Poisson point process on G with unit intensity and is denoted by $\mathbf{Q}_G$. It is easy to see that $\mathbf{P}_{G,z,\Phi}$ is absolutely continuous with respect to $\mathbf{Q}_G$ and

$$\frac{d\mathbf{P}_{G,z,\Phi}}{d\mathbf{Q}_G}(X_G) = \exp\{-E_{z,\Phi}(X_G)\} \;.$$

For details of Gibbsian point processes, see, e.g., Ruelle (1969); Preston (1976); or Stoyan et al. (1987).

A potential function is said to be hard-core if there is a constant r_0 such that $\Phi(x) = +\infty$ whenever $|x| < r_0$. The supremum of such r_0 is called the hard-core distance. On the other hand, a potential function which has finite values except at the origin is said to be soft-core. A Gibbsian point process having a hard-core potential is supported by those configurations with points at least the hard-core distance apart one another. Therefore two hard-core Gibbsian processes having different hard-core distances are not mutually absolutely continuous. This causes some difficulties in the following study and is one of the reasons that we consider only soft-core potentials in this research. As to the stability of a given potential, there is a fine criterion.

Theorem 1 (Ruelle 1969, Proposition 3.2.8) *If there are two non-negative decreasing functions* $f_1(r)$, $f_2(r)$ *defined respectively on* $(0, a_1)$ *and* (a_2, ∞), $0 < a_1 < a_2$, *and satisfying conditions*:

$$\Phi(x) \geqq \begin{cases} f_1(|x|) & \text{if } |x| < a_1, \\ -f_2(|x|) & \text{if } |x| > a_2, \end{cases}$$

and

$$\int_0^{a_1} r f_1(r)\,\mathrm{d}r = +\infty,$$

$$\int_{a_2}^{\infty} r f_2(r)\,\mathrm{d}r < +\infty,$$

then Φ *is stable.*

A potential is said to be of Lennard-Jones type if it satisfies the preceding criterion with $f_1(r) = f_2(r) = cr^{-\lambda}$, $\exists\lambda > 2$, $\exists c > 0$. All the potentials considered in this paper is isotropic, that is, $\Phi(x)$ depends only on $|x|$. Therefore we can rewrite $\Phi(x)$ as $\Phi(|x|)$. Also we assume Φ is always continuous except at the origin. We will consider a family of local energies parametrized by three parameters $\theta = (z, \alpha, \beta)$, $-\infty < z < \infty$, and α, $\beta > 0$, as:

$$U_\theta(x)_n = \alpha \sum_{1 \leqq i < j \leqq n} \Phi(\beta r_{ij}),$$

$$E_\theta(x)_n = nz + U_\theta(x)_n,$$

where $r_{ij} = |x_i - x_j|$. By setting $\beta = 1$ we also consider the two-parameters model. We will specify a possible range Θ of the parameter θ later. The restriction to a three-parameters model is not essential, but most practical potentials are of this form, see Grandy (1987, Chap. 7). Following are examples of soft-core potentials. They are of Lennard-Jones type.

(1) Soft-sphere potential (this leads actually to a two-parameters model);

$$\Phi(r) = \frac{1}{r^n}, \quad n > 2.$$

(2) Repulsive exponential potential;

$$\Phi(r) = \frac{1}{r^n} \mathrm{e}^{-r}, \quad n > 2.$$

(3) Lennard-Jones potential;

$$\Phi(r) = \frac{1}{r^m} - \frac{c}{r^n}, \quad c > 0, \quad m > n > 2.$$

We take the asymptotic viewpoint and let regions $G = G_n$ expand monotonically to $\mathbb{R}^2$. In the rest of the paper we will fix $\{G_n\}$ and Φ. Because of notational simplicity, we use following notations; $\mathbf{P}_{n,\theta}$ for the Gibbsian distribution corresponding to G_n and θ, $\mathbf{Q}_n$ for $\mathbf{Q}_{G_n}$, X_n for X_{G_n}, and U_θ (resp. E_θ) for the interaction (resp. local) energy corresponding to Φ and θ.

3 Uniform LAN condition

The main result of this paper is a consequence of a general optimality theory of maximum likelihood estimators under the ULAN condition. In the following we cite this condition (slightly modified to our situation) and a relevant theorem for convenience of reference.

A statistical experiment $\{\mathbf{P}_{n,\theta}\}$, $\theta \in \Theta$, for observations X_n is said to satisfy the ULAN condition if there are non-degenerate matrices $\phi_n(\theta)$ and if the following conditions N1–N6 are satisfied;

(N1) For $\forall$compact $K \subset \Theta$, $\forall$sequence $\theta_n \in K$, and $\forall$sequence u_n with $u_n \to u$ and $\theta_n + \phi_n(\theta_n)u_n \in K$, likelihood ratios have the expression

$$Z_{n,\theta_n}(u_n) \equiv \frac{d\mathbf{P}_{n,\theta_n+\phi_n(\theta_n)u_n}}{d\mathbf{P}_{n,\theta_n}}(X_n) = \exp\left\{(\Delta_{n,\theta_n}, u) - \frac{|u|^2}{2} + \psi_n(u_n, \theta_n)\right\},$$

where $\mathscr{L}(\Delta_{n,\theta_n}|\mathbf{P}_{n,\theta_n}) \to \mathscr{N}(0, J)$ as $n \to +\infty$, J being the identity matrix, and the sequence $\psi_n(u_n, \theta_n) \to 0$ in $\{\mathbf{P}_{n,\theta_n}\}$-probability.

(N2) For $\forall$compact $K \subset \Theta$,

$$\lim_{n\to\infty} \sup_{\theta\in K} \operatorname{trace}(\phi_n(\theta)\phi_n(\theta)^{\mathrm{T}}) = 0.$$

(N3) For $\forall$compact $K \subset \Theta$, $\exists\beta > 0$, $\exists m > 0$, $\exists D = D(K)$, and $\exists a = a(K)$,

$$\sup_{\theta\in K} \sup_{\substack{u,v\in\Theta_{n,\theta}\\ |u|,|v|<R}} |u-v|^{-\beta}\mathbf{E}_{n,\theta}|Z_{n,\theta}^{1/m}(u) - Z_{n,\theta}^{1/m}(v)|^m < D(1+R^a),$$

where $\Theta_{n,\theta} = \{u; \theta + \phi_n(\theta)u \in \Theta\}$. The MLE $\hat{\theta}_n$ of the parameter θ based on observation X_n is one of the values which maximize the likelihood function $p_n(y, X_n)$ with respect to $y \in \Theta$. The class $\mathbf{W}_p$ consists of those functions $w(t)$, $t \in \mathbb{R}^3$, satisfying following conditions;

1) $w(t) \geqq 0$, $w(0) = 0$, $w(0) \not\equiv 0$, and $w(-t) = w(t)$,
2) $w(t)$ is continuous at $t = 0$,
3) $\{t; w(t) < c\}$ is convex for all $c > 0$, and is bounded for all sufficiently small $c > 0$.

With each $w \in \mathbf{W}_p$ we associate loss functions $W_n(t, \theta) = w(\phi_n^{-1}(\theta)(t-\theta))$. The MLE is said to be w-asymptotically efficient in K if for $\forall$non-empty open $U \subset K$

$$\lim_{n\to\infty}\left(\inf_{\{T_n\}}\sup_{\theta\in U}\mathbf{E}_{n,\theta}W_n(T_n,\theta) - \sup_{\theta\in U}\mathbf{E}_{n,\theta}W_n(\hat{\theta}_n,\theta)\right) = 0,$$

where the infimum is taken over all estimators $\{T_n\}$ of the parameter θ.

Under the ULAN condition a class of estimators including MLE can be shown to be w-asymptotically efficient for a wide class of loss functions w.

Theorem 2 (Ibragimov and Has'minskii 1981, Chap. 3) *Let conditions* N1–N4 *and* N6 *be satisfied with* $\beta > 3$ *in* N3. *Let* K *be an arbitrary compact set in* Θ. *Then uniformly in* $\theta \in K$:
(1) *the estimator* $\hat{\theta}_n$ *is consistent,*
(2) *the estimator* $\hat{\theta}_n$ *is asymptotically normal with the parameters* $(\theta, \phi_n^2(\theta))$,
(3) *all the moments of the random variables* $\phi_n^{-1}(\theta)(\hat{\theta}_n - \theta)$ *converges as* $n \to \infty$ *to the corresponding moments of the normal distribution* $\mathcal{N}(0, J)$.
For each $\varepsilon > 0$

$$\lim_{n \to \infty} \sup_{\theta \in K} \mathbf{P}_{n,\theta}\{|\phi_n^{-1}(\theta)(\hat{\theta}_n - \Delta_{n,\theta})| > 0\} = 0.$$

If the condition N5 *is also satisfied,* $\hat{\theta}_n$ *is w-asymptotically efficient for every* $w \in \mathbf{W}_p$ *and every compact* $K \subset \Theta$.

Note that, since $w(t) = |t|^2 \in \mathbf{W}_p$, $\hat{\theta}_n$ is also asymptotically efficient under the quadratic loss $W = |\phi_n^{-1}(\theta)(t - \theta)|^2$. As to the definitions of the uniform consistency and the uniform convergence in law, see the book of Ibragimov and Has'minskii.

4 Strong cluster estimate

A key tool in this research is the strong cluster estimate of the Ursell function associated with a given local energy. From a technical reason, we need to consider a complex-valued pair potential function Φ and a complex chemical potential z. A complex-valued potential Φ is stable if its real part is stable. Assume that Φ satisfies the following estimate with $p > 4$;

$$|e^{-\Phi(x)} - 1| \leqq K(1 + |x|)^{-p} \quad \text{for} \quad \forall x \in \mathbb{R}^2 .$$

Define the constant $q_0 = e^{\mathrm{Re}\{z\} + B + 1}K$, an absolute constant $A_0 = \max_{n \geqq 2} 2n^{n-2}e^{-n}/n!$ (< 0.147), and the function

$$q(r) = q_0(1 + r)^{-p/2} .$$

Theorem 3 (Minlos and Pogosian (1977, Theorem 1 and Lemma 4)) *If*

$$\int q(|x|)\,dx < 1 , \qquad (1)$$

then the logarithm of the grand partition function $\Xi_{G,z,\Phi} = \int_{\mathscr{C}(G)} e^{-E_{z,\Phi}(c)} dc$ *can be defined and has the representation*

$$\log \Xi_{G,z,\Phi} = \int_{\mathscr{C}(G)} \Psi_{z,\Phi}(c)\,dc .$$

The Ursell function $\Psi_{z,\Phi}(c)$ *satisfies the strong cluster estimate*

$$|\Psi_{z,\Phi}(c)| \leqq \frac{A_0}{K} q_0 \sum_{\gamma \in \mathscr{L}_c} \prod_{(x,y) \in \gamma} q(|x - y|) ,$$

where the sum is taken over the set $\mathscr{L}_c$ *of all chain graphs on* c, *and the product is taken over all vertices* (x, y) *of a chain* γ. *The inequality* (1) *is satisfied in particular if* $p > 4$ *and* $q_0 < \frac{1}{2}$.

The proof of this theorem is given in Minlos and Pogosian (1977) for real-valued Φ and z. Its proof can be modified easily to the complex-valued case. This fact is used implicitly in Pogosian (1984). The explicit form of $\Psi_{z,\Phi}$ is

$$\Psi_{z,\Phi}(c) = e^{-\#cz} \sum_{\gamma} \prod_{(x,y)\in\gamma} [e^{-\Phi(x-y)} - 1] ,$$

where the sum is taken over all connected graphs on c and the product is taken over all edges (x, y) of γ, see Ruelle (1969, Chap. 4).

Asymptotic arguments in this paper are based on the following asymptotic expansion of integrals of strong cluster functions. This result goes back to Yang and Lee (1952) and is elaborated extensively in Pogosian (1984).

Theorem 4 (Pogosian (1984, Theorem 7)) *Let a translation invariant function $\Psi(c)$ satisfy the strong cluster estimate*

$$|\Psi(c)| \leqq \sum_{\gamma\in\mathscr{L}_c} \prod_{(x,y)\in\gamma} q(x-y) ,$$

where the function q is of the form $q(x) = q_+(|x|)(1+|x|)^{-p}$ with $q_+(r)$ being non-negative, bounded and integrable on $[0, \infty)$. Assume that $p \geqq 5$, $\int q(x)\,dx < 1$, and $\sup q(x) < 1/2$. *For each bounded convex set G define the cluster integral*

$$Q(G) = \int_{\mathscr{C}(G)} \Psi(c)\,dc .$$

Then $Q(G)$ has the following asymptotic expansion as $G \uparrow \mathbb{R}^2$:

$$Q(G) = \left[\int_{\mathscr{C}(\mathbb{R}^2)} \frac{\Psi(c\cup\{0\})}{1+\#c}\,dc \right] |G| + R(G)|V(G)| , \tag{2}$$

where $|G|$ means the area of G and $V(G) = \{x\in G;\, \mathrm{dist}(x, \partial G) \leqq 1\}$. The residual term $R(G)$ can be bounded as $|R(V)| \leqq C(q)$. The constant $C(q)$ depends only on q and, if $q_1 \leqq q_2$, then $C(q_1) \leqq C(q_2)$.

Remark. The proportional constant to $|G|$ of the first term of the right hand side of (2) can be interpreted as the Gibbs specific free energy in the context of statistical physics if Ψ is the Ursell function of a Gibbsian point process. An explicit form of constants $C(q)$ is not necessary but, for example, can be chosen to be

$$C(q) = \left(1 - \int q(x)\,dx\right)^{-2} \left(\lambda(0) + 8\int_0^\infty r\lambda(r)\,dr\right) ,$$

where $\lambda(r) = \int_{|x|\geqq r} (q + q^{*2} + q^{*3} + \cdots)\,dx$.

5 Auxiliary lemmas

We collect auxiliary lemmas in this section which will be used later. These lemmas are used in particular to show the existence of moment generating functions and the local uniformity of several estimates with respect to $\theta\in\Theta$.

Lemma 1 *Let a continuous potential Φ satisfy two conditions*

$$\Phi(r) \geqq c_1/r^q \quad as \quad r \to +0\,,$$
$$|\Phi(r)| \leqq c_2/r^p \quad as \quad r \to \infty\,,$$

with some positive constants c_1, c_2, p, and q. Then there is a constant K such that

$$|e^{-\Phi(r)} - 1| \leqq K(1+r)^{-p} \quad for \quad \forall r > 0\,.$$

Proof. Easy. □

Lemma 2 *Consider two convergent sequences $\alpha_n \to \alpha > 0$, $\beta_n \to \beta > 0$ and numbers $p > 1$, $q > 1$, $K > 0$, $c_1 > 0$, and c_2. c_2 may be zero. Also consider a potential Φ and a family of potentials $\{\Phi_\rho(r)\}$. Let $\Phi(r)$ be continuous for $r > 0$ and $\sup_\rho|\Phi_\rho(r)|$ be bounded on any interval $[r_1, r_2]$, $0 < r_1 < r_2 < \infty$. Assume the following conditions;*

$$|e^{-\alpha\Phi(\beta r)} - 1| \leqq K(1+r)^{-p} \quad for \quad \forall r > 0\,,$$

$$\Phi(r) \approx c_1/r^q \quad and \quad \sup_\rho|\Phi_\rho(r)| = O(\Phi(r)) \quad as \quad r \to +0,$$

$$\Phi(r) \approx c_2/r^p \quad and \quad \sup_\rho|\Phi_\rho(r)| = O(|\Phi(r)|) \quad as \quad r \to \infty\,.$$

Then, for $\forall\varepsilon > 0$, there are $\eta > 0$ and N such that, if $|x+yi| \leqq \eta$ and $n \geqq N$,

$$|e^{(x+yi)\Phi_p(r) - \alpha_n\Phi(\beta_n r)} - 1| \leqq (1+\varepsilon)K(1+r)^{-p} \quad for \quad \forall r > 0.$$

Proof. The proof is divided into three parts.
1) As $r \to +0$, we have, from the conditions,

$$\sup_\rho|\Phi_\rho(r)| = O(\alpha\Phi(\beta r))$$

$$\alpha_n\Phi(\beta_n r) - \alpha\Phi(\beta r) = o(\alpha\Phi(\beta r))\,.$$

Hence there are $\exists\delta > 0$, $\exists N$, $\exists r_1 > 0$ such that, if $|x+yi| \leqq \delta$, $n \geqq N$, and $r \leqq r_1$,

$$|(x+yi)\Phi_\rho(r) - \alpha_n\Phi(\beta_n r) + \alpha\Phi(\beta r)| \leqq \tfrac{1}{2}\alpha\Phi(\beta r) \quad \text{for } \forall\rho\,.$$

From the elementary inequality $|e^z - 1| \leqq |z|e^{|z|}$ for complex z and the inequality

$$|e^{(x+yi)\Phi_\rho(r) - \alpha_n\Phi(\beta_n r)} - 1| \leqq e^{-\alpha\Phi(\beta r)}|e^{(x+yi)\Phi_\rho(r) - \alpha_n\Phi(\beta_n r) + \alpha\Phi(\beta r)} - 1| + |e^{-\alpha\Phi(\beta r)} - 1|\,, \tag{3}$$

it follows

$$|e^{(x+yi)\Phi_\rho(r) - \alpha_n\Phi(\beta_n r)} - 1| \leqq \tfrac{1}{2}\alpha\Phi(\beta r)e^{-\alpha\Phi(\beta r)/2} + K(1+r)^{-p}$$
$$= K(1+r)^{-p}[1 + o(1)]\,.$$

Therefore, there is $\exists r_1 > 0$ for $\forall\varepsilon > 0$ such that the assertion is valid for $r \leqq r_1$.
2) As $r \to +\infty$, we have, from the conditions,

$$\alpha_n\Phi(\beta_n r) - \alpha\Phi(\beta r) = o(1/r^p)\,.$$

Hence, for $\forall\varepsilon$, there exist $\exists\delta > 0$, $\exists N$, $\exists r_2 > 0$ such that, if $|x + yi| \leqq \delta$, $n \geqq N$, and $r \geqq r_2$,

$$|(x + yi)\Phi_\rho(r) - \alpha_n\Phi(\beta_n r) + \alpha\Phi(\beta r)| \leqq \frac{\varepsilon}{r^p} \quad \text{for } \forall\rho \,.$$

From the preceding inequality (3)

$$|e^{(x+yi)\Phi_\rho(r) - \alpha_n\Phi(\beta_n r)} - 1| \leqq e^{-\alpha\Phi(\beta r)}\frac{\varepsilon}{r^p}e^{\varepsilon/r^p} + K(1 + r)^{-p}$$

$$= K(1 + r)^{-p}[1 + o(1)] \,.$$

Therefore, there is $\exists r_2 > 0$ for $\forall\varepsilon > 0$ such that the assertion is valid for $r \geqq r_2$.
3) Fix an arbitrary interval $[r_1, r_2]$, $0 < r_1 < r_2 < \infty$. From the uniform continuity on compact intervals of continuous functions, there exist $\exists\delta > 0$, $\exists N$, $\exists r_2 > 0$ for $\forall\varepsilon$, such that, if $|x + yi| \leqq \delta$, $n \geqq N$, and $r \geqq r_2$,

$$|(x + yi)\Phi_\rho(r) - \alpha_n\Phi(\beta_n r) + \alpha\Phi(\beta r)| \leqq \varepsilon \quad \text{for } \forall\rho \,.$$

Using the inequality (3) again,

$$|e^{(x+yi)\Phi_\rho(r) - \alpha_n\Phi(\beta_n r)} - 1| \leqq e^{-\alpha\Phi(\beta r)}\varepsilon e^{\varepsilon} + K(1 + r)^{-p} = K(1 + r)^{-p}[1 + o(1)] \,.$$

Therefore the assertion holds for $r \in [r_1, r_2]$ and, combining three cases, we can finish the proof. □

Lemma 3 *Under the same situation as in Lemma 2, there is a constant $\delta > 0$ for each $\varepsilon > 0$ such that potentials $\Phi + (x + iy)\Phi_\rho$, $\forall\rho$, is stable with the stability constant $(1 + \varepsilon)B$, B being a stability constant of Φ, if $|x + iy| \leqq \delta$.*

Proof. From Theorem 1, there is a positive η such that the potential

$$\Phi_0 = \Phi - \eta \sup_\rho |\Phi_\rho|$$

is stable. Let B_0 be its stability constant. We can assume $B_0 > B$ without loss of generality. If $|x| \leqq \delta$, then

$$\sum_{i<j} \{\Phi(r_{ij}) + x\Phi_\rho(r_{ij})\} \geqq \left(1 - \frac{\delta}{\eta}\right)\sum_{i<j} \Phi(r_{ij}) + \frac{\delta}{\eta}\sum_{i<j} \Phi_0(r_{ij})$$

$$\geqq \left(1 - \frac{\delta}{\eta}\right)Bn - \frac{\delta}{\eta}B_0 n \,.$$

Therefore, if we let $\delta = \varepsilon\eta B/(B_0 - B)$, the assertion holds. □

Lemma 4 *Let p and q be positive. Assume that $\Phi(r) \to +\infty$ as $r \to 0$, and that $r^p\Phi(r) \to c_2$ as $r \to +\infty$. Let $K_\alpha = \sup_{r>0}(1 + r)^p|e^{-\alpha\Phi(r)} - 1|$. Then K_α is monotone non-decreasing and is larger than both 1 and $\alpha|c_2|$. $\lim_{\alpha\to\infty} K_\alpha < \infty$ iff Φ is both non-negative and of bounded range. If so, $\lim_{\alpha\to\infty} K_\alpha = (1 + r_0)^p$, where $r_0 = \sup\{r;\, \Phi(r) > 0\}$.*

Proof. Letting $r \to 0$ (resp. $\to \infty$) in $(1+r)^p|e^{-\alpha\Phi(r)} - 1|$, we see $K_\alpha \geqq 1$ (resp. $\geqq \alpha|c_2|$). If $K_\alpha = (1+r)^p|e^{-\alpha\Phi(r)} - 1|$ for some r, then $\Phi(r) \neq 0$ and, if $\alpha < \alpha'$,

$$K_\alpha = (1+r)^p|e^{-\alpha\Phi(r)} - 1| \leqq (1+r)^p|e^{-\alpha'\Phi(r)} - 1| \leqq K_{\alpha'} .$$

If $K_\alpha = \lim_{r\to 0}(1+r)^p|e^{-\alpha\Phi(r)} - 1|$, $K_\alpha = 1$. Hence $K_{\alpha'} \geqq 1 = K_\alpha$. Also if $K_\alpha = \lim_{r\to\infty}(1+r)^p|e^{-\alpha\Phi(r)} - 1|$, $K_\alpha = \alpha|c_2|$. Hence $K_{\alpha'} \geqq \alpha'|c_2| \geqq \alpha|c_2| = K_\alpha$. Therefore K_α is monotone non-decreasing.
If $\Phi(r) > 0$, then

$$\sup_\alpha K_\alpha \geqq \sup_\alpha (1+r)^p|e^{-\alpha\Phi(r)} - 1| = (1+r)^p .$$

Also if $\Phi(r) < 0$, then

$$\sup_\alpha K_\alpha \geqq \sup_\alpha (1+r)^p|e^{-\alpha\Phi(r)} - 1| = +\infty .$$

The second half of the lemma follows from the last two relations. □

Lemma 5 *Let K_α and $K_{\alpha,\beta}$, $\alpha, \beta > 0$, be the smallest constants such that inequalities*

$$|e^{-\alpha\Phi(r)} - 1| \leqq K_\alpha(1+r)^{-p} \quad \text{for} \quad \forall r > 0 ,$$
$$|e^{-\alpha\Phi(\beta r)} - 1| \leqq K_{\alpha,\beta}(1+r)^{-p} \quad \text{for} \quad \forall r > 0 ,$$

hold. Then

$$K_{\alpha,\beta} = \begin{cases} K_\alpha & \text{if} \quad \beta \geqq 1 \\ \beta^{-p}K_\alpha & \text{if} \quad \beta < 1 . \end{cases}$$

Proof. Easy. □

Lemma 6 *Let B and $B_{\alpha,\beta}$, $\alpha, \beta > 0$, be the smallest constants such that the following inequalities hold for $\forall(x)_n$,*

$$\sum_{1 \leqq i < j \leqq n} \Phi(r_{ij}) \geqq -Bn ,$$
$$\sum_{1 \leqq i < j \leqq n} \alpha\Phi(\beta r_{ij}) \geqq -B_{\alpha,\beta}n .$$

Then $B_{\alpha,\beta} = \alpha B$.

Proof. Easy. □

Lemma 7 *Let A be a square matrix. Define functions $F_1(x) = f(Ax)$ and $F_2(x) = e^{f(Ax)}$. Then gradient vectors $\partial_x F_a = \left(\frac{\partial}{\partial x_i} F_a\right)$ and Hessian matrices $\partial_x^2 F_a = \left(\frac{\partial^2}{\partial x_i \partial x_j} F_a\right)$ are given by*

$$\partial_x F_1 = A^{\mathrm{T}}(\partial_x f)(Ax) ,$$
$$\partial_x F_2 = F_2 \times A^{\mathrm{T}}(\partial_x f)(Ax) ,$$
$$\partial_x^2 F_1 = A^{\mathrm{T}}[(\partial_x^2 f)(Ax)]A ,$$
$$\partial_x^2 F_2 = F_2 \times A^{\mathrm{T}}[(\partial_x f)(Ax)(\partial_x f)(Ax)^{\mathrm{T}} + (\partial_x^2 f)(Ax)]A.$$

Proof. Straightforward. □

Lemma 8 *Let $\mathscr{M}_n$ be the space of positive-definite $n \times n$ matrices. For $X = \{x_{ij}\} \in \mathscr{M}_n$ let $P(X)$ be the vector $(x_{ij})_{i \leqq j}$. The Jacobian $\partial P(X^{-1})/\partial P(X)$ of the transform $X \to X^{-1}$ is $|X|^{-n-1}$. Also the Jacobian of the transform $X \to X^2$ is $\prod_{i \leqq j}(\lambda_i + \lambda_j)$ where $\{\lambda_i\}$ is the eigenvalues of X. Transforms $X \to X^{1/2}$ and $X \to X^{-1/2}$ are holomorphic functions in variables $P(X)$ on $\mathscr{M}_n$. In particular, for each $X_0 \in \mathscr{M}$, there is a positive constants ε and c such that, if $\|X_1 - X_0\|$, $\|X_2 - X_0\| \leqq c$, then*

$$\|X_1^{1/2} - X_2^{1/2}\|, \quad \|X_1^{-1/2} - X_2^{-1/2}\| \leqq \varepsilon .$$

Proof. As to Jacobian formulas, see Rogers (1980, Chap. 14). The rest follows from the well-known implicit function theorem, see Hörmander (1973, Chap. 2). □

Lemma 9 *Let a function $f(z_1, \ldots, z_n)$ of n-complex variables be holomorphic in a polydisk $M = \{|z_i| < r_i\}$. If $|f| \leqq C$ in M, then Cauchy's estimate*

$$|(\partial^\alpha f)(0)| \leqq C \prod_i \alpha_i!/r_i^{\alpha_i}$$

holds for every multi-index $\alpha = \{\alpha_i\}$.

Proof. See Hörmander (1973, Chap. 2). □

Lemma 10 *Let $X = (X_1, X_2, \ldots)^{\mathrm{T}}$ be a random vector and $a = (a_1, a_2, \ldots)^{\mathrm{T}}$ be its mean vector. Let $f(t) = \mathbf{E}\{e^{(t, X)}\}$ be the moment generating function of X and let $g(t) = \log f(t)$. If ∂_t denotes the vector $(\partial/\partial t_1, \partial/\partial t_2, \ldots)^{\mathrm{T}}$ of partial differential operators, then*

$$(h, \partial_t)^k g(t)|_{t=0} = \mathbf{E}\{(h, X - a)^k\} \quad k = 2, 3, \ldots .$$

Proof. Straightforward. □

6 ULAN condition of Gibbsian point process

In this section we will show that the Gibbsian model satisfies the ULAN condition under several regularity conditions. These regularity conditions seem restrictive and may be weakened in several points. First we formulate our basic assumptions:
(A1) The potential function Φ is isotropic and, except at the origin, finite and two-times continuously differentiable.
(A2) There are constants $c_1 > 0$, c_2, $p > 12$ and $q > 2$ such that, if $r \to +0$,

$$r^q \Phi(r) \to c_1, \; r\Phi'(r), \; r^2 \Phi''(r) = O(\Phi(r))$$

and, if $r \to \infty$,

$$r^p \Phi(r) \to c_2, \; r\Phi'(r), \; r^2 \Phi''(r) = O(|\Phi(r)|) .$$

Let B and K_α, $\alpha > 0$, be the smallest constants satisfying inequalities

$$\sum_{1 \leqq i < j \leqq n} \Phi(r_{ij}) \geqq -Bn \quad \text{for} \quad \forall (x)_n,$$

$$|e^{-\alpha\Phi(r)} - 1| \leqq K_\alpha (1 + r)^{-p} \quad \text{for} \quad \forall r > 0 .$$

For $\theta = (z, \alpha, \beta) \in \mathbb{R} \times (0, \infty)^2$, let $B_\theta = \alpha B$, $K_\theta = K_\alpha$ if $\beta \geqq 1$, $= \beta^{-p} K_\alpha$ if $\beta < 1$, $q_0(\theta) = e^{z + B_\theta + 1} K_\theta$, and $q_\theta(r) = q_0(\theta)(1 + r)^{-p/2}$. If $\theta = (z, \alpha)$, we should consider $\beta = 1$.

(A3) The parameter space Θ is a bounded open subset of Θ_0, where Θ_0 consists of those θ such that both $q_0(\theta) < \frac{1}{2}$ and $\int q_\theta(|x|)\,dx < 1$.

For a two-parameters model we need one more condition. Let $p(\theta) = c_\theta(0)$ be the specific free energy corresponding to the parameter $\theta = (z, \alpha)$.

(A4) The Hessian matrix of the specific free energy $p(\theta)$ is positive definite on the closure of Θ.

Remark. The finiteness of constants B_θ and K_θ under conditions A_1 and A_2 is guaranteed by Theorem 1 and Lemma 1. Since $\int q_\theta(|x|)\,dx = 8\pi q_0(\theta)/(p-2)(p-4)$ and $p > 12$, Θ_0 is defined actually by the sole condition $q_0(\theta) < \frac{1}{2}$. Three potentials cited previously satisfy conditions A1 and A2 if $n > 12$. From lemmas proved in the last section, it can be shown that, if $(z_0, \alpha_0, \beta_0) \in \Theta_0$, then $\{(z, \alpha, \beta)\};\ z \leqq z_0,\ \alpha \leqq \alpha_0,\ \beta \geqq \beta_0)\} \subset \Theta_0$. The specific free energy $p(z, \alpha)$ is analytic on Θ_0. From Hölder inequality applied to grand partition functions, $p(\theta)$ is seen to be a convex function of θ and, hence, has positive semi-definite Hessian matrix.

Symbols such as U_θ, E_θ, X_n, $\mathbf{P}_{n,\theta}$, $\mathbf{E}_{n,\theta}$, and $Z_{n,\theta}$ have the similar meanings as before. The grand partition function corresponding to G_n and θ is denoted by $\Xi_{n,\theta}$. The Ursell function corresponding to θ is denoted by Ψ_θ. Let $\partial E_\theta(c)$ be the gradient vector

$$\left(\# c,\ \sum_{x_i, x_j \in c} \Phi(\beta r_{ij}),\ \alpha \sum_{x_i, x_j \in c} r_{ij} \Phi'(\beta r_{ij})\right) \quad \text{if} \quad \theta = (z, \alpha, \beta)\,, \tag{4}$$

$$\left(\# c,\ \sum_{x_i, x_j \in c} \Phi(r_{ij})\right) \quad \text{if} \quad \theta = (z, \alpha)$$

The following is the main result and the rest of the paper is devoted to its proof.

Theorem 5 *Let $M_{n,\theta}$ and $V_{n,\theta}$ be the mean vector and the covariance matrix of $\partial E_\theta(X_n)$ with respect to $\mathbf{P}_{n,\theta}$. Define the matrix $\phi_n(\theta) = V_{n,\theta}^{-1/2}$ and the random vector*

$$\Delta_{n,\theta} = -\,\phi_n(\theta)\left[\partial E_\theta(X_n) - M_{n,\theta}\right].$$

Under conditions A1–A3, the ULAN *condition* N1–N5 *holds. For a two-parameters model, the* ULAN *condition* N1–N6 *holds if conditions A1–A4 are satisfied.*

Proof. First consider the moment generating function of $\partial E_t(X_n)$

$$\Xi_{n,t}(\xi) = \int_{\mathscr{C}(G)} e^{\xi^T \partial E_t(c) - E_t(c)}\, dc, \quad \xi \in \mathbb{C}^3\,,$$

which exists at least in a neighborhood of the origin from conditions A1–A3 and Lemma 2. Also the existence of partial derivatives up to degree 2 of $\Xi_{n,t}$ with respect to t can be shown. Then, from Theorem 4,

$$\log \Xi_{n,t}(\xi) = c_t(\xi)|G_n| + R_{n,t}(\xi)|V(G_n)|\,.$$

From preceding lemmas, Theorem 3, and 4, it can be shown that, for each fixed $t \in \Theta$, there is a polydisk $D_\rho = \{\xi;\ \max_i |\xi_i| < \rho\}$, a constant T, and a neighborhood

W_t of t such that $c_s(\xi)$ and $R_{n,s}(\xi)$ are holomorphic (as a function of $\xi \in \mathbb{C}^3$) in D_ρ and have estimates

$$|c_s(\xi)| \leqq T,\ |R_{n,s}(\xi)| \leqq T \quad \text{for} \quad \forall \xi \in D_\rho,\ \forall s \in W_t, \quad \text{and } \forall n\,.$$

From these estimates and Lemma 8, it follows that there is a constant T' such that absolute values of all the partial derivatives of $c_s(\xi)$ and $R_{n,s}(\xi)$ up to degree 4 are bounded by T' for $\forall \xi \in D_{\rho/2}$, $\forall s \in W_t$, and $\forall n$.

Since

$$\partial_\xi \log \Xi_{n,s}(\xi)|_{\xi=0} = M_{n,s}\,,$$

$$\partial_\xi^2 \log \Xi_{n,s}(\xi)|_{\xi=0} = V_{n,s}\,,$$

$M_{n,s}/|G_n|$ and $V_{n,s}/|G_n|$ converge to $(\partial_\xi c_s)(0)$ and $(\partial_\xi^2 c_s)(0)$ respectively uniformly in $s \in W_t$. Note that these convergences are locally uniform in t. Hence

$$|G_n|^{-1/2} \phi_n(t)^{-1} \to (\partial_\xi^2 c_t)(0)^{1/2}\,,$$

$$|G_n|^{1/2} \phi_n(t) \to (\partial_\xi^2 c_t)(0)^{-1/2}\,,$$

locally uniformly in t. From these relations, we can conclude that conditions N2 and N5 hold.

Next we will prove the asymptotic normality of Δ_{n,t_n}. The sequence $\{t_n\}$ is assumed to be in a compact set K and we can assume without loss of generality that $t_n \to \exists t_0$. Consider the function $f_{n,t}(\xi) = \Xi_{n,t}(\xi)/\Xi_{n,t}(0)$ and expand $\log f_{n,t}(\xi)$ as

$$\log f_{n,t}(\xi) = \xi^{\mathrm{T}} M_{n,t} + \tfrac{1}{2} \xi^{\mathrm{T}} V_{n,t} \xi + \tfrac{1}{6} [(\xi^{\mathrm{T}} \partial_\xi)^3 \log \Xi_{n,t}](\xi')\,,$$

with $\xi' = \rho \xi$, $0 < \exists \rho < 1$. Then the moment generating function $f_n(\xi) = \mathbf{E}_{n,t_n}\{\mathrm{e}^{-\xi^{\mathrm{T}} \Delta_{n,t_n}}\}$ of $-\Delta_{n,t_n}$ can be expanded as

$$\begin{aligned} \log f_n(\xi) &= \log f_{n,t_n}(\phi_n(t_n)\xi) - (\phi_n(t_n)\xi)^{\mathrm{T}} M_{n,t_n} \\ &= \tfrac{1}{2}|\xi|^2 + \text{residual term}\,, \end{aligned}$$

where the residual term is a combination of partial derivatives of degree three of c_{t_n} and R_{n,t_n}. From the preceding estimates of these derivatives and the order relation $\phi_n = O(|G_n|^{-1/2})$, the residual term converges to 0 as $n \to \infty$. This proves $\mathscr{L}(\Delta_{n,t_n} | \mathbf{P}_{n,t_n}) \to \mathscr{N}(0, J)$, a part of Condition N1.

The rest of Condition N1 is proved as follows. The log-likelihood ratio $\log Z_{n,t_n}(u_n)$ is expanded as

$$\text{(5)} \qquad \log Z_{n,t_n}(u_n) = u_n^{\mathrm{T}} \Delta_{n,t_n} - \tfrac{1}{2}|u_n|^2 - \tfrac{1}{2}[\phi_n(t_n)u_n]^{\mathrm{T}} \Delta^*_{n,t_n+\phi_n(t_n)u_n'}[\phi_n(t_n)u_n]$$

where $u_n' = \rho_n u_n$ with $0 < \exists \rho_n < 1$ and

$$\Delta^*_{n,t} = (\partial^2 E_t)(X_n) - \mathbf{E}\{(\partial^2 E_t)(X_n)\}\,.$$

From the same reasoning used to show that $V_{n,t} = O(|G_n|)$, we can show that

$$\mathbf{E}_{n,t_n}\{|u_n^{\mathrm{T}} \Delta^*_{n,t_n} u_n|^2\} = O(|G_n|)$$

and, hence, $\mathbf{E}_{n,t_n}\{|u_n^{\mathrm{T}} \Delta^*_{n,t_n} u_n|\}$ is of order $O(|G_n|^{1/2})$. Therefore, the third term of the right-hand-side of (5), that is, $\psi_n(u_n, t_n)$ in Condition N1, converges to 0 in $\{\mathbf{P}_{n,t_n}\}$-probability and we can complete the proof of Condition N1.

Let us proceed to the proof of Condition N3. First note the inequality

$$\mathbf{E}_{n,t}|Z_{n,t}(u)^{1/4} - Z_{n,t}(v)^{1/4}|^4 \leqq 8\mathbf{E}_{n,t}\{Z_{n,t}(u) + Z_{n,t}(v)\} = 16 .$$

Hence, if $|u - v| \geqq 1$,

$$|u - v|^{-4}\mathbf{E}_{n,t}|Z_{n,t}(u)^{1/4} - Z_{n,t}(v)^{1/4}|^4 \leqq 16 .$$

Using Lemma 7, we have the relation

$$\text{(6)}\quad \mathbf{E}_{n,t}|Z_{n,t}(u)^{1/4} - Z_{n,t}(v)^{1/4}|^4 = \frac{1}{4^4}|u - v|^4\mathbf{E}_{n,t+\phi_n(t)u'}|(\phi_n(t)h)^{\mathrm{T}}\Delta_{n,t+\phi_n(t)u'}|^4 ,$$

where $h = 1/|u - v|\cdot(u - v)$ and $u' = v + \rho(u - v)$, $0 < \rho < 1$. The expectation of the right-hand-side of this equation is shown, from Lemma 6, to be equal to

$$\text{(7)}\quad [(\phi_n(t)h)^{\mathrm{T}}\partial_\xi]^4 \log \Xi_{n,t+\phi_n(t)u'}(\xi)|_{\xi=0} = [(\phi_n(t)h)^{\mathrm{T}}\partial_\xi]^4 c_{t+\phi_n(t)u'}(\xi)|_{\xi=0} \times |G_n|$$
$$+ [(\phi_n(t)h)^{\mathrm{T}}\partial_\xi]^4 R_{n,t+\phi_n(t)u'}(\xi)|_{\xi=0} \times |V(G_n)| .$$

Recall that $\phi_n(t)$ is of order $|G_n|^{1/2}$ locally uniformly in t, and that partial derivatives $[\partial^4/\partial\xi_i\ldots]c_t(\xi)|_{\xi=0}$ and $[\partial^4/\partial\xi_i\ldots]R_{n,t}(\xi)|_{\xi=0}$ are bounded in n and locally uniformly bounded in t. For ε small enough, the compact set $K' = \{\theta + u; \theta \in K, |u| \leqq \varepsilon\}$ become a subset of Θ. Then, from some n on, $t + \phi_n(t)u' \in K'$ for $\forall t \in K$ and $|\forall u'| < 1$. Consequently we can choose a constant $D = D(K)$ so that Condition N3 is valid for $\beta = 4$, $m = 4$ and $a = 0$. This finishes the proof of Condition N3. The proof of Condition N4 can be done as follows. First note the equality

$$\text{(8)}\qquad \mathbf{E}_{n,t}\{Z_{n,t}^{1/2}(u)\} = \int_{\mathscr{C}(G_n)} \left[\frac{\mathrm{e}^{-U_t(c)}}{\Xi_{n,t}}\right]^{1/2} \left[\frac{\mathrm{e}^{-U_{t+v}(c)}}{\Xi_{n,t+v}}\right]^{1/2} \mathrm{d}c ,$$

where $v = \phi_n(t)u$. From the expansion (2) the logarithm of the right hand side of (8) can be expressed as the sum of two terms

$$I_{n,t,v}^{(1)}|G_n| + I_{n,t,v}^{(2)}|V(G_n)| , \quad \text{say,}$$

where the term $I_{n,t,v}^{(1)}$ is equal to

$$p(t + \tfrac{1}{2}v) - \tfrac{1}{2}(p(t) + p(t + v)) .$$

If $\partial^2 p$ denotes the Hessian matrix of p, the last expression can be rewritten as

$$-\tfrac{1}{8}v^{\mathrm{T}}(\partial^2 p)(t + \rho v)v ,$$

with $0 < \exists\rho < 1$. From the assumption A4, the smallest eigenvalue of $(\partial p)(t)$ is positive and is bounded by some $p_0 > 0$ from below uniformly for $t \in \Theta$. Hence

$$I_{n,t,v}^{(1)} \leqq -\tfrac{1}{8}p_0|v|^2 .$$

On the other hand, the term $I_{n,t,v}^{(2)}$, that is,

$$R_{n,t+\frac{1}{2}v} - \tfrac{1}{2}(R_{n,t} + R_{n,t+v}) ,$$

can be bounded in absolute value by $r|v|^2$ with some constant r which can be chosen uniformly for $t \in \Theta$ and $v \in \Theta - t$.

Therefore, recalling that $\phi_n(t) = O(|G_n|^{-1/2})$ and $|V(G_n)| = O(|G_n|^{1/2})$, we can conclude that there are positive constants w_1 and w_2 which are independent of $t \in \Theta$, such that

$$\sup_{u \in \Theta_{n,t}} |u|^N \mathbf{E}_{n,t} Z_{n,t}^{1/2}(u) = \sup_{v \in \Theta - t} |\phi_n^{-1}(t) u|^N \exp\{I_{n,t,v}^{(1)}|G_n| + I_{n,t,v}^{(2)}|V(G_n)|\}$$

$$\leqq w_1 |G_n|^N \exp\{-w_2|G_n|\} ,$$

for n large enough. This completes the proof of the condition N4. □

7 Concluding remarks

Our study is an attempt to the rigorous statistical analysis of planar Gibbsian point process. Assumptions and results of this paper are by no means complete and will be improved in many points in future studies. So following comments may be useful.

The assumption that Θ is bounded and convex in A3 is added in order to ease the proof of Condition N3 and N4 and may not be essential. We restricted ourselves to soft-core potentials. However, as is easily seen, our analysis is also valid for hard-core potentials if the hard-core distance is the same for all parameters (necessarily the case of the two-parameters model). If the hard-core distance varies with the parameter, several problems may occur. The first is how to parametrize this moving hard-core distance. The next is the fact that two Gibbsian point processes with different hard-core distances are not mutually absolutely continuous as noted previously. The LAN condition and the relevant theory have still meanings if we define the likelihood $d\mathbf{P}_{n,u}/d\mathbf{P}_{n,v}$ to be the Radon-Nikodym derivative of the absolutely continuous part of $\mathbf{P}_{n,u}$ with respect to $\mathbf{P}_{n,v}$, see, Ibragimov and Has'minskii (1981). Nevertheless, as is usual in statistics, moving supports of probability measures may cause a special irregularity and need a separate consideration.

We do not know how restrictive the assumption A4 is. It can be shown that $p(z, \alpha)$ is strictly convex, that is, the surface $y = p(z, \alpha)$ contain no segment. Also we are not aware of a variant of this assumption which is suitable for a three-parameters model. We should note that the process with a parameter (z, α, β) on a domain G is formally equivalent to the process with the parameter $(z - 2\log\beta, \alpha)$ on the domain βG.

If we restrict ourselves to the two-parameters model, the potential function Φ need not to be differentiable. A possible natural generalization of our results to a multi-parameter case may be, as in Gidas (1991), to introduce the parametrized local energies of the form

$$E_\theta(x)_n = \theta_0 n + \sum_{k=1}^{m} \theta_i \left[\sum_{1 \leqq i < j \leqq n} \Phi_k(r_{ij}) \right],$$

where $\theta = (\theta_0, \theta_1, \ldots, \theta_m)$ and $\{\Phi_k\}$ are fixed potential functions. It is likely that our arguments are also valid with appropriate modifications.

We should note that the theory of Minlos and Pogosian also covers Gibbsian processes on lattice points and our results may have a discrete version. Recently a Bayesian approach to the image restoration using Gibbsian processes on lattice points has been successfully applied, see Geman and Geman (1984). On this account, although it is unknown that a Bayesian estimation has a practical meaning in the problem of estimation of potential functions, it may be interesting to note that the Bayesian estimator as well as the maximum likelihood estimator is asymptotically efficient under the ULAN condition, see Ibragimov and Has'minskii (1981, Theorem 2.1 and 2.2).

Finally it is appropriate to state practical implications of our result. The initial motivation of the present study was to understand the MLE method of Ogata and Tanemura. To be precise, the MLE used by Ogata and Tanemura is not our MLE but the conditional MLE with the condition on number of points, that is, they modeled point patterns by canonical Gibbsian processes. Also the iterative procedure of Moyeed and Baddeley (1989) calculates the conditional MLE. However a direct study of the efficiency of conditional MLE seems to be impossible because we have less knowledge of canonical Gibbsian processes. Nevertheless we can show that under several assumptions efficiencies of two MLE's are asymptotically equivalent (in the sense that they are both asymptotically normal and have the same asymptotic variance), see Mase (1991). Therefore the conditional MLE is also asymptotically efficient and the conditional inference of Ogata and Tanemura can be justified.

Also our result, together with theorems of Jensen (1990), implies that the pseudo-likelihood estimator is inefficient. This fact is known in the lattice case, see e.g., Gidas (1991), but there seem no proofs in the continuous state space case. Jensen proved that the pseudo-likelihood estimator is asymptotically normal and showed its asymptotic variance explicitly. (By the way, he used a completely different framework from ours, that is, a strong-mixing type condition, and is very important methodologically.) This variance is different from that of MLE. On the other hand, it is known that under the LAN condition MLE is BAN, that is, it has the smallest asymptotic variance in every asymptotically normal regular estimator, see Ibragimov and Has'minskii (1981, Chap. 2, Theorem 9.1). Therefore the pseudo-likelihood estimator cannot be asymptotically efficient.

References

1. Besag, J., Milne, R., Zachary, S.: Point process limits of lattice processes. J. Appl. Probab. **19**, 210–216 (1982)
2. Diggle, P.J., Gates, D.J., Stibbard, A.: A nonparametric estimator for pairwise-interaction point processes. Biometrika **74**, 763–770 (1987)
3. Fiksel, T.: Estimation of interaction potentials of Gibbsian point processess. Operationsforsch. Stat. Ser. Stat. **20**, 270–278 (1984)
4. Geman, S., Geman, D.: Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. IEEE Trans. PAMI **6**, 721–741 (1984)
5. Gidas, B.: Parameter estimation for Gibbs distributions, I: Fully observed data. In: Chellapa, R., Jain, A. (eds.) Markov random fields; theory and applications. New York: Academic Press 1991
6. Glötzl, E. and Rauschenschwandtner, B.: On the statistics of Gibbsian processes. In: Re've'sz, P. et al. (eds.) The first Pannonian symposium on mathematical statistics, Bad Tatzmannsdorf, Austria, 1979. (Lect. Notes Stat., vol. 8, pp. 83–93) Berlin Heidelberg New York: Springer 1981
7. Grandy, Jr., W.T.: Foundations of Statistical Mechanics. Volume I: Equilibrium theory. Dordrecht: Reidel 1987

8. Hanisch, K.-H., Stoyan, D.: Remarks on statistical inference and prediction for a hard-core clustering model. Math. Operationsforsch. Stat. Ser. Stat. **14**, 559–567 (1983)
9. Hörmander, L.: An introduction to complex analysis in several variables. Amsterdam: North-Holland 1973
10. Ibragimov, I.A., Has'minskii, R.Z.: Statistical estimation, asymptotic theory. Berlin Heidelberg New York: Springer 1981
11. Jensen, J.L.: Asymptotic normality of estimates in spatial point processes. Research Reports No. 210, Department of Theoretical Statistics, University of Aarhus (1990)
12. Ji, C.: Estimating Functionals of one-dimensional Gibbs states. Probab. Theory Relat. Fields **82**, 155–175 (1989)
13. Kutoyants, Yu. A.: Parameter estimation for stochastic processes. Berlin: Heddermann 1984
14. Mase, S.: Locally asymptotic normality of Gibbs models on a lattice. J. Appl. Probab. **16**, 585–602 (1984)
15. Mase, S.: Asymptotic equivalence of grand canonical MLE and canonical MLE of pair potential functions of Gibbsian point process model. (Submitted, 1991)
16. Minlos, R.A., Pogosian, S.K.: Estimates of ursell functions, group functions, and their derivatives. Theor. Math. Phys. **31**, 408–418 (1977)
17. Moyeed, R.A., Baddeley, A.J.: Stochastic approximation of the MLE for a spatial point pattern. Report BS-R8926, Center for Mathematics and Computer Sciences, Stichting Mathematisch Centrum. (1989)
18. Ogata, Y., Tanemura, M.: Approximation of likelihood functions in estimating interaction potentials from spatial point patterns. J.R. Stat. Soc. Ser. B**46**, 496–518 (1984)
19. Ogata, Y., Tanemura, M.: Estimation of interaction potentials of marked spatial point processes through the maximum-likelihood method. Biometrics **41**, 421–433 (1985)
20. Ogata, Y., Tanemura, M.: Estimation of interaction potentials of spatial point patterns through the maximum likelihood procedure. Ann. Inst. Stat. Math. **B33**, 315–338 (1981)
21. Penttinen, A.: Modelling interactions in spatial point patterns: parameter estimation by the maximum-likelihood method. Jyväskyla Stud. Comp. Sci., Econ. Stat. **7**, 1–107 (1984)
22. Pickard, D.K.: Asymptotic inference for an Ising lattice. J. Appl. Probab. **13**, 486–497 (1976)
23. Pickard, D.K.: Asymptotic inference for an Ising lattice, II. Adv. Appl. Probab. **9**, 476–501 (1977)
24. Pickard, D.K.: Asymptotic inference for an Ising lattice, III, Non-zero field and ferromagnetic states. J. Appl. Probab. **16**, 12–24 (1979)
25. Pickard, D.K.: Inference for general Ising models. Adv. Appl. Probab. **14**, 345–357 (1982)
26. Pogosian, S.K.: Asymptotic expansion in the local limit theorem for the particle number in the grand canonical ensemble. In: Fritz, J. et al. (eds.) Random Fields. vol. 2, pp. 889–913. Colloquia Mathematica Societatis János Bolyai. Amsterdam: North-Holland 1981
27. Pogosian, S.K.: Asymptotic expansion of the logarithm of the partition function. Commun. Math. Phys. **95**, 227–245 (1984)
28. Preston, C.: Random fields. Berlin Heidelberg New York: Springer 1976
29. Ripley, B.D.: Modelling spatial patterns (with discussion). J. R. Stat. Soc. Ser. B **39**, 172–212 (1977)
30. Ripley, B.D.: Statistical inference for spatial processes. Cambridge: Cambridge University Press 1988
31. Rodgers, G.S.: Matrix derivatives. New York: Marcel Dekker 1980
32. Ruelle, D.: Statistical mechanics rigorous results. New York: Benjamin 1969
33. Stoyan, D., Kendall, W.S., Mecke, J.: Stochastic geometry and its applications. New York: Wiley 1987
34. Takacs, R.: Estimator for the pair-potential of a Gibbsian point process. Statistics **17**, 429–433 (1986)
35. Yang, C.N., Lee, T.D.: Statistical theory of equations of state and phase transition, 1. theory of condensation. Phys. Rev. **87**, 404–409 (1952)
36. Younes, L.: Parametric inference for imperfectly observed Gibbsian fields. Probab. Theory Relat. Fields **82**, 625–645 (1989)