Consistency of likelihood estimation for Gibbs point processes

David Dereudre,
Laboratoire Paul Painlevé, University Lille 1
(Joint work with **Frédéric Lavancier**, Laboratoire Jean Leray, Université Nantes)

Conference of the GDR GeoSto, Nantes, 6th April 2016

General problem

Let P be an infinite volume Gibbs Point Process in $\mathbb{R}^d$ for an energy function which depends on a parameter θ.

General problem

Let P be an infinite volume Gibbs Point Process in $\mathbb{R}^d$ for an energy function which depends on a parameter θ.
Let γ be a realization of P.

General problem

Let P be an infinite volume Gibbs Point Process in $\mathbb{R}^d$ for an energy function which depends on a parameter θ.
Let γ be a realization of P.
Let $\hat{\theta}_\Lambda$ the MLE of θ based on the observation ω_Λ.

General problem

Let P be an infinite volume Gibbs Point Process in $\mathbb{R}^d$ for an energy function which depends on a parameter θ.
Let γ be a realization of P.
Let $\hat{\theta}_\Lambda$ the MLE of θ based on the observation ω_Λ.

Question

What are the asymptotic properties of $(\hat{\theta}_\Lambda)$ *when* Λ *goes to* $\mathbb{R}^d$ *?*

General problem

Let P be an infinite volume Gibbs Point Process in $\mathbb{R}^d$ for an energy function which depends on a parameter θ.
Let γ be a realization of P.
Let $\hat{\theta}_\Lambda$ the MLE of θ based on the observation ω_Λ.

Question

What are the asymptotic properties of $(\hat{\theta}_\Lambda)$ *when* Λ *goes to* $\mathbb{R}^d$ *?*

- Mase(1992) The MLE is consistent, asymptotically normal and efficient for exponential models with small interaction (cluster expansion conditions).

General problem

Let P be an infinite volume Gibbs Point Process in $\mathbb{R}^d$ for an energy function which depends on a parameter θ.
Let γ be a realization of P.
Let $\hat{\theta}_\Lambda$ the MLE of θ based on the observation ω_Λ.

Question

What are the asymptotic properties of $(\hat{\theta}_\Lambda)$ when Λ goes to $\mathbb{R}^d$?

- Mase(1992) The MLE is consistent, asymptotically normal and efficient for exponential models with small interaction (cluster expansion conditions).
- Jensen(1993) The MLE is asymptotically normal under the Dobrushin uniqueness condition.

General problem

Let P be an infinite volume Gibbs Point Process in $\mathbb{R}^d$ for an energy function which depends on a parameter θ.
Let γ be a realization of P.
Let $\hat{\theta}_\Lambda$ the MLE of θ based on the observation ω_Λ.

Question

What are the asymptotic properties of $(\hat{\theta}_\Lambda)$ *when* Λ *goes to* $\mathbb{R}^d$ *?*

- Mase(1992) The MLE is consistent, asymptotically normal and efficient for exponential models with small interaction (cluster expansion conditions).
- Jensen(1993) The MLE is asymptotically normal under the Dobrushin uniqueness condition.
- Mase(2000) The MLE is consistent for exponential models.

General problem

Let P be an infinite volume Gibbs Point Process in $\mathbb{R}^d$ for an energy function which depends on a parameter θ.
Let γ be a realization of P.
Let $\hat{\theta}_\Lambda$ the MLE of θ based on the observation ω_Λ.

Question

What are the asymptotic properties of $(\hat{\theta}_\Lambda)$ when Λ goes to $\mathbb{R}^d$?

- Mase(1992) The MLE is consistent, asymptotically normal and efficient for exponential models with small interaction (cluster expansion conditions).
- Jensen(1993) The MLE is asymptotically normal under the Dobrushin uniqueness condition.
- Mase(2000) The MLE is consistent for exponential models.
- Other estimators are consistent and asymptotically normal : MPLE, Takacs-Fiksel estimators, variational estimators.

Finite volume Gibbs point process

Let Λ be a bounded window in $\mathbb{R}^d$, $\mathcal{C}_\Lambda$ the space of finite configurations in Λ and π_Λ the law of the Poisson Point Process in Λ with intensity 1.

Finite volume Gibbs point process

Let Λ be a bounded window in $\mathbb{R}^d$, $\mathcal{C}_\Lambda$ the space of finite configurations in Λ and π_Λ the law of the Poisson Point Process in Λ with intensity 1.

The energy H is a functional from Ω_Λ to $\mathbb{R} \cup \{+\infty\}$.

Definition

The Finite volume Gibbs point process in Λ is the probability measure on Ω_Λ which is absolutely continuous with respect to π_Λ with density

$$f_\Lambda = \frac{1}{Z_\Lambda} e^{-H}.$$

Examples : pairwise potential interactions

$$H_\Lambda(\gamma) = z\ N_\Lambda(\gamma) + \sum_{\{x,y\}\in\gamma} \phi(x-y).$$

Examples : pairwise potential interactions

$$H_\Lambda(\gamma) = z\ N_\Lambda(\gamma) + \sum_{\{x,y\}\in\gamma} \phi(x-y).$$

- **The Strauss pair potential :**

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

Examples : pairwise potential interactions

$$H_\Lambda(\gamma) = z\ N_\Lambda(\gamma) + \sum_{\{x,y\}\in\gamma} \phi(x-y).$$

- **The Strauss pair potential :**

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

- **The Lennard-Jones pair potential :**

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

(The standard Lennard-Jones model, $n = 12$ and $m = 6$).

Examples :

The area interaction :

$$H_\Lambda(\gamma) = z\ N_\Lambda(\gamma) + \beta \text{Volume}\left(\bigcup_{x \in \gamma} B(x, R) \right).$$

Examples :

The area interaction :

$$H_\Lambda(\gamma) = z\ N_\Lambda(\gamma) + \beta \text{Volume}\left(\bigcup_{x\in\gamma} B(x,R)\right).$$

The Quermass interaction :

$$H_\Lambda(\gamma) = z\ N_\Lambda(\gamma) + \sum_{i=1}^{d+1} \beta_i M_i \left(\bigcup_{x\in\gamma} B(x,R)\right).$$

($(M_i)_{1\le i\le d+1}$ are the Minkowski's functionals)

Infinite volume Gibbs point process

$\mathcal{C}$ is the space of locally finite configurations of points in $\mathbb{R}^d$.

Infinite volume Gibbs point process

$\mathcal{C}$ is the space of locally finite configurations of points in $\mathbb{R}^d$.
A *family of interaction energies* is a collection $\mathcal{H} = (H_\Lambda)$ of measurable functions from Ω to $\mathbb{R} \cup \{+\infty\}$ such that for $\Lambda \subset \Lambda'$

$$H_{\Lambda'}(\gamma) = H_\Lambda(\gamma) + \varphi_{\Lambda,\Lambda'}(\gamma_{\Lambda^c}).$$

Infinite volume Gibbs point process

$\mathcal{C}$ is the space of locally finite configurations of points in $\mathbb{R}^d$.
A *family of interaction energies* is a collection $\mathcal{H} = (H_\Lambda)$ of measurable functions from Ω to $\mathbb{R} \cup \{+\infty\}$ such that for $\Lambda \subset \Lambda'$

$$H_{\Lambda'}(\gamma) = H_\Lambda(\gamma) + \varphi_{\Lambda,\Lambda'}(\gamma_{\Lambda^c}).$$

The local conditional densities :

$$f_\Lambda(\gamma) = \frac{1}{Z_\Lambda(\gamma_{\Lambda^c})} e^{-H_\Lambda(\gamma)},$$

Definition

A probability measure P on is a *Gibbs measure* if for every Λ,

$$P(d\gamma_\Lambda | \gamma_{\Lambda^c}) = f_\Lambda(\gamma)\pi_\Lambda(d\gamma_\Lambda).$$

The MLE procedure

We consider a familly of parametric energy functionals (H_Λ^θ) where $\theta \in \mathring{\Theta} \subset \mathbb{R}^p$.
P^* is a Gibbs point process for $(H_\Lambda^{\theta^*})$ with unknown θ^*.
γ^* is a realization of P^*.
$\Lambda_n = [-n, n]^d$ are the observation windows.

The MLE procedure

We consider a familly of parametric energy functionals (H_Λ^θ) where $\theta \in \mathring{\Theta} \subset \mathbb{R}^p$.
P^* is a Gibbs point process for $(H_\Lambda^{\theta^*})$ with unknown θ^*.
γ^* is a realization of P^*.
$\Lambda_n = [-n, n]^d$ are the observation windows.

Definition

Let $\mathcal{K}$ be a compact subset of $\mathring{\Theta}$ such that $\theta^* \in \mathcal{K}$. The MLE of θ^* from the observation $\gamma^*_{\Lambda_n}$ is defined by

$$(\hat{\theta}_n) = \underset{\theta \in \mathcal{K}}{\operatorname{argmax}} f^\theta_{\Lambda_n}(\gamma^*_{\Lambda_n}).$$

Corollaries of our main Theorems

- The MLE of (z^*, β^*, R^*) in the Strauss model is consistent

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

- The MLE of $(z^*, A^*, B^*, n^*, m^*)$ in the Lennard-Jones model is consistent

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

- The MLE of (z^*, β^*, R^*) in the Area Process is consistent

$$H_\Lambda(\gamma) = z \; N_\Lambda(\gamma) + \beta \text{Volume}\left(\bigcup_{x\in\gamma} B(x, R)\right).$$

Assumptions of our main Theorem

[Stability] : For any compact set $\mathcal{K} \subset \Theta$, there exists a constant $\kappa \geq 0$ such that for any Λ, any $\theta \in \mathcal{K}$ and any $\gamma \in \Omega$

$$H_\Lambda^\theta(\gamma_\Lambda) \geq -\kappa N_\Lambda(\gamma)$$

Assumptions of our main Theorem

[Stability] : For any compact set $\mathcal{K} \subset \Theta$, there exists a constant $\kappa \geq 0$ such that for any Λ, any $\theta \in \mathcal{K}$ and any $\gamma \in \Omega$

$$H_\Lambda^\theta(\gamma_\Lambda) \geq -\kappa N_\Lambda(\gamma)$$

[MeanEnergy] : The following decompositions holds

$$H_{\Lambda_n}^\theta = \sum_{k \in \{-n, n-1\}^d} H_0^\theta \circ \tau_{-k} + \partial H_{\Lambda_n}^\theta$$

with for any $P \in \mathcal{G}$

$$E_P(H_0^\theta) < +\infty$$

and

$$\lim_{n \to \infty} \frac{1}{|\Lambda_n|} \sup_{\theta \in \mathcal{K}} \left| \partial H_{\Lambda_n}^\theta \right| \overset{P-as}{=} 0.$$

Assumptions of our main Theorem

[Boundary] : For all $P \in \mathcal{G}$, for any compact set $\mathcal{K} \subset \Theta$ and for P-almost every $\gamma \in \Omega$

$$\lim_{n\to\infty} \frac{1}{|\Lambda_n|} \sup_{\theta\in\mathcal{K}} \left| H^{\theta}_{\Lambda_n}(\gamma_{\Lambda_n}) - H^{\theta}_{\Lambda_n}(\gamma) \right| = 0.$$

Assumptions of our main Theorem

[Boundary] : For all $P \in \mathcal{G}$, for any compact set $\mathcal{K} \subset \Theta$ and for P-almost every $\gamma \in \Omega$

$$\lim_{n \to \infty} \frac{1}{|\Lambda_n|} \sup_{\theta \in \mathcal{K}} \left| H^{\theta}_{\Lambda_n}(\gamma_{\Lambda_n}) - H^{\theta}_{\Lambda_n}(\gamma) \right| = 0.$$

[Regularity] : For all $P \in \mathcal{G}$, for any compact set $\mathcal{K} \subset \Theta$

$$E_P \left(\sup_{\substack{\theta' \in \mathcal{K} \\ |\theta - \theta'| \leq r}} \left| H^{\theta}_0 - H^{\theta'}_0 \right| \right) \underset{r \mapsto 0}{\longmapsto} 0.$$

Assumptions of our main Theorem

[Regularity] : (second part)
For any $\eta > 0$ and any $\theta_0 \in \Theta$, there exists $\theta \subset B(\theta_0, \eta)$ and $r_0 > 0$ such that for any Λ and any γ_Λ

$$\inf_{\theta' \in B(\theta_0, r_0)} \left(\frac{H_\Lambda^\theta(\gamma_\Lambda) - H_\Lambda^{\theta'}(\gamma_\Lambda)}{N_\Lambda(\gamma_\Lambda)} \right) \geq g(r_0) \underset{r_0 \mapsto 0}{\longmapsto} 0$$

where g is a funtion such that $g(r) \mapsto 0$ when $r \mapsto 0$.

Assumptions of our main Theorem

[VariationalPrinciple] : The *pressure* exits ;

$$p(\theta) := \lim_{n\to\infty} \frac{1}{|\Lambda_n|} \ln(Z^{\theta}_{\Lambda_n}).$$

In addition, for any θ, θ' in Θ and $\mu^{\theta'} \in \mathcal{G}^{\theta'}$,

$$\mathcal{I}(\mu^{\theta'}) + E_{\mu^{\theta'}}(H^{\theta}_0) \geq -p(\theta)$$

and the equality holds if and only if $\theta' = \theta$.
($\mathcal{I}(\mu^{\theta'})$ is the specific entropy of $\mu^{\theta'}$ with respect to π)

Assumptions of our main Theorem

[VariationalPrinciple] : The *pressure* exits ;

$$p(\theta) := \lim_{n\to\infty} \frac{1}{|\Lambda_n|} \ln(Z^\theta_{\Lambda_n}).$$

In addition, for any θ, θ' in Θ and $\mu^{\theta'} \in \mathcal{G}^{\theta'}$,

$$\mathcal{I}(\mu^{\theta'}) + E_{\mu^{\theta'}}(H_0^\theta) \geq -p(\theta)$$

and the equality holds if and only if $\theta' = \theta$.
($\mathcal{I}(\mu^{\theta'})$ is the specific entropy of $\mu^{\theta'}$ with respect to π)
This assumption is satisfied for any finite range interaction (Der. 2015)

Main Theorems

Theorem (Dereudre, Lavancier)

Under the assumptions **[Stability]**, **[MeanEnergy]**, **[Boundary]**, **[Regularity]** *and* **[VariationalPrinciple]**, *for any* $\theta^* \in \mathcal{K}$ *and any* $P^* \in \mathcal{G}^{\theta^*}$, *the MLE* $\hat{\theta}_n$ *converges* P^**-almost surely to* θ^* *when* n *goes to infinity.*

Main Theorems

Theorem (Dereudre, Lavancier)

Under the assumptions **[Stability]**, **[MeanEnergy]**, **[Boundary]**, **[Regularity]** *and* **[VariationalPrinciple]**, *for any* $\theta^* \in \mathcal{K}$ *and any* $P^* \in \mathcal{G}^{\theta^*}$, *the MLE* $\hat{\theta}_n$ *converges* P^**-almost surely to* θ^* *when* n *goes to infinity.*

Variants of this theorem are given in the paper :

- with an hardcore part
- in the pairwise setting
- in the linear setting

Sketch of the proof

$$\hat{\theta}_n(\gamma) = \underset{\theta \in \mathcal{K}}{\text{argmax}} f^{\theta}_{\Lambda_n}(\gamma_{\Lambda_n}).$$

- **Step 1 (A limiting contrast function)** :

$$K_n(\theta, \gamma) := \frac{1}{|\Lambda_n|} \log f^{\theta}_{\Lambda_n}(\gamma) \overset{P^* - as}{\longrightarrow} K(\theta).$$

(Classical thermodynamic arguments)

Sketch of the proof

$$\hat{\theta}_n(\gamma) = \underset{\theta \in \mathcal{K}}{\operatorname{argmax}} f^{\theta}_{\Lambda_n}(\gamma_{\Lambda_n}).$$

- **Step 1 (A limiting contrast function)** :

$$K_n(\theta, \gamma) := \frac{1}{|\Lambda_n|} \log f^{\theta}_{\Lambda_n}(\gamma) \overset{P^*-as}{\longrightarrow} K(\theta).$$

(Classical thermodynamic arguments)

- **Step 2 (Identification)** :

$$\operatorname{argmax}_{\theta} K(\theta) = \{\theta^*\}.$$

(Variational principle)

Sketch of the proof

$$\hat{\theta}_n(\gamma) = \underset{\theta \in \mathcal{K}}{\text{argmax}} f^{\theta}_{\Lambda_n}(\gamma_{\Lambda_n}).$$

- **Step 1 (A limiting contrast function)** :

$$K_n(\theta,\gamma) := \frac{1}{|\Lambda_n|} \log f^{\theta}_{\Lambda_n}(\gamma) \overset{P^*-as}{\longrightarrow} K(\theta).$$

(Classical thermodynamic arguments)

- **Step 2 (Identification)** :

$$\text{argmax}_{\theta} K(\theta) = \{\theta^*\}.$$

(Variational principle)

- **Step 3 (Convergence of argmax)** :

$$\text{argmax}_{\theta} K_n(\theta,\gamma) \overset{P^*-as}{\longrightarrow} \text{argmax}_{\theta} K(\theta).$$

(main contribution in the present work)

Step 1 : A limiting contrast function

$$
\begin{aligned}
K_n(\theta,\gamma) &= \frac{1}{|\Lambda_n|}\log f^{\theta}_{\Lambda_n}(\gamma) \\
&= \frac{1}{|\Lambda_n|}\Big(-\log(Z^{\theta}_{\Lambda_n}) + H^{\theta}_{\Lambda_n}(\gamma_{\Lambda_n})\Big)
\end{aligned}
$$

We assume that P^* is ergodic. So, For P^*-almost all γ

$$
K_n(\theta,\gamma) \longmapsto -p(\theta) - E_{P^*}(H^{\theta}_0) := K(\theta).
$$

Step 2 : Identification

Recall the variational principle,
for any θ, θ' in Θ and $\mu^{\theta'} \in \mathcal{G}^{\theta'}$

$$\mathcal{I}(\mu^{\theta'}) + E_{\mu^{\theta'}}(H_0^{\theta}) \geq -p(\theta) \qquad (1)$$

with equality if and only if $\theta = \theta'$.
So

$$\begin{aligned} K(\theta) &= -p(\theta) - E_{P^*}(H_0^{\theta}) \\ &\leq I(P^*), \end{aligned}$$

with equality if and only if $\theta = \theta^*$.

Step 3 : Convergence of argmax

- Mase 2000 (exponential models) : The contrast functions are are strictly concave. The convergence of argmax is direct.

Step 3 : Convergence of argmax

- Mase 2000 (exponential models) : The contrast functions are are strictly concave. The convergence of argmax is direct.
- Guyon 1995 : Control of the modulus of continuity.

Step 3 : Convergence of argmax

- Mase 2000 (exponential models) : The contrast functions are are strictly concave. The convergence of argmax is direct.
- Guyon 1995 : Control of the modulus of continuity.

Lemma (Dereudre-Lavancier)

If a family of random contrast functions (K_n^θ) *satisfies*

- $K_n^\theta \overset{P^*-as}{\mapsto} K^\theta$
- $argmax_\theta K^\theta = \{\theta^*\}$
- $\theta \mapsto \mathcal{K}^\theta$ *is upper semicontinuous*
- *There exists a function* $g : \mathbb{R}^+ \to R^+$ *with* $g(x) \mapsto 0$ *when* $x \mapsto 0$ *such that* $\forall \varepsilon > 0$, $\forall \theta$, $\exists \theta' \in B(\theta, g(\varepsilon))$, $\exists r > 0$

$$P^* \left(\limsup_{n\to\infty} \left\{ \sup_{B(\theta,r)} K_n^{\cdot} - K_n^{\theta'} \geq \varepsilon \right\} \right) = 0.$$

Then the $argmax_\theta K_n^\theta$ *converges* P^**-almost surely to* $argmax_\theta K^\theta$.

Bibliography

- D. Dereudre, F. Lavancier, *Consistency of likelihood estimation for Gibbs point processes*, To appear in Annals of Stat.
- D. Dereudre, *Variational principle for Gibbs Point Processes with finite range interaction*, Elect. Comm. Prob. (2015)
- X. Guyon. *Random Fields on a Network*. Springer-Verlag (1995).
- J. L. Jensen. *Asymptotic normality of estimates in spatial point processes*. Scandinavian Journal of Statistics (1993).
- S. Mase. *Uniform LAN condition of planar Gibbsian point processes and optimality of maximum likelihood estimators of soft-core potential functions*. Probability Theory and Related Fields, (1992).
- S. Mase. *Asymptotic properties of MLE's of Gibbs models on* $\mathbb{R}^d$. Unpublished manuscript, (2002).