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

J.-M. Billiot, J.-F. Coeurjolly, R. Drouilhet
University of Grenoble 2, France

June 2, 2006
Objective

number of points=266

number of points=472
Outline

1. Gibbs point processes
   - Basic definition
   - Existence conditions based on the energy function
   - Description of some Gibbs models

2. Statistical model and inference method

3. Asymptotic results
   - Consistency of the mple estimator
   - Asymptotic normality of the mple estimator

4. Description of some examples and short simulation
Outline

1. Gibbs point processes
   - Basic definition
   - Existence conditions based on the energy function
   - Description of some Gibbs models

2. Statistical model and inference method

3. Asymptotic results
   - Consistency of the mple estimator
   - Asymptotic normality of the mple estimator

4. Description of some examples and short simulation
Point processes: definition and notation

**Notation**

- $\mathcal{B}_b$: set of bounded borelian of $\mathbb{R}^d$.
- $\Omega_f$, $\Omega$, $\Omega_\Lambda$: set of finite configurations of $\mathbb{R}^d$, set of configurations in $\mathbb{R}^d$, set of configurations in $\Lambda \subset \mathbb{R}^d$:
- Let $\Lambda \subset \mathbb{R}^d$ and $\varphi \in \Omega$, $\varphi_\Lambda := \varphi \cap \Lambda \in \Omega_\Lambda$

**Point process in some bounded $\Lambda \subset \mathbb{R}^d$**

A point process in $\Lambda$ is a random variable $\Phi_\Lambda$ with values in $\Omega_\Lambda$ equipped with the smallest $\sigma$-field which make measurable all the maps $i_\Delta : \varphi \in \Omega_\Lambda \rightarrow |\varphi_\Delta|$ with $\Delta \subset \Lambda \in \mathcal{B}_b$. 
Useful notation

\[ \oint_{\Lambda} d\varphi g(\varphi) := \sum_{n=0}^{+\infty} \frac{1}{n!} \int_{\Lambda} \cdots \int_{\Lambda} dx_1 \cdots dx_n g(\{x_1, \cdots, x_n\}) \]

i.e. \( \oint_{\Lambda} d\varphi \) means the summation over all configuration \( \varphi \) in \( \Lambda \).

### Poisson point process and Gibbs point process

We define:

\[ Q_{\Lambda}(F) = \frac{1}{\exp(|\Lambda|)} \sum_{n=0}^{+\infty} \frac{1}{n!} \int_{\Lambda} \cdots \int_{\Lambda} dx_1 \cdots dx_n 1_F(\{x_1, \cdots, x_n\}) \]
Useful notation

\[ \oint_{\Lambda} d\varphi g(\varphi) := \sum_{n=0}^{+\infty} \frac{1}{n!} \int_{\Lambda} \cdots \int_{\Lambda} dx_1 \cdots dx_n g(\{x_1, \cdots, x_n\}) \]

i.e. \( \oint_{\Lambda} d\varphi \) means the summation over all configuration \( \varphi \) in \( \Lambda \).

Poisson point process and Gibbs point process

We define:

→ a poisson point process with intensity 1 in \( \Lambda \) with probability measure \( Q_{\Lambda} \)

\[ Q_{\Lambda}(F) = \frac{1}{\exp(|\Lambda|)} \oint_{\Lambda} d\varphi 1_{F}(\varphi) \]
Useful notation

\[ \oint_{\Lambda} d\varphi \ g(\varphi) := \sum_{n=0}^{+\infty} \frac{1}{n!} \int_{\Lambda} \cdots \int_{\Lambda} dx_1 \cdots dx_n \ g(\{x_1, \cdots, x_n\}) \]

i.e. \( \oint_{\Lambda} d\varphi \) means the summation over all configuration \( \varphi \) in \( \Lambda \).

Poisson point process and Gibbs point process

We define:

\( \rightarrow \) a Gibbs point process in \( \Lambda \) with probability measure \( P_\Lambda \)

\[ P_\Lambda(F) = \frac{1}{Z_\Lambda} \oint_{\Lambda} d\varphi \ 1_F(\varphi) \exp(-V(\varphi)) \]

where \( Z_\Lambda < +\infty \) as soon as \( V(\varphi) > -K|\varphi| \) (i.e. \( V(\cdot) \) is stable).
**Useful notation**

\[
\oint_{\Lambda} d\varphi g(\varphi) := \sum_{n=0}^{+\infty} \frac{1}{n!} \int_{\Lambda} \cdots \int_{\Lambda} dx_1 \cdots dx_n g(\{x_1, \cdots, x_n\})
\]

i.e. \(\oint_{\Lambda} d\varphi\) means the summation over all configuration \(\varphi\) in \(\Lambda\).

**Poisson point process and Gibbs point process**

We define:

→ a Gibbs point process in \(\mathbb{R}^d\) with conditional probability measure \(P_\Lambda(\cdot | \varphi^o)\) for all \(\varphi^o\)

\[
P_\Lambda(F | \varphi^o) = \frac{1}{Z_\Lambda(\varphi^o)} \oint_{\Lambda} d\varphi \mathbf{1}_F(\varphi) \exp \left( -V(\varphi | \varphi^o_{\Lambda^c}) \right)
\]

where \(V(\varphi | \varphi^o_{\Lambda^c}) := V(\varphi \cup \varphi^o_{\Lambda^c}) - V(\varphi^o_{\Lambda^c})\) is the energy required to insert the points of \(\varphi\) in \(\varphi^o_{\Lambda^c}\).
Framework (of the presentation)

Restricted to stationary Gibbs point processes based on energy function related to some graph $G_2(\varphi)$:

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

with $G_k(\varphi)$: set of cliques of order $k$ of $\varphi$

satisfying the following Assumptions

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

$E_2$ Locality of the local energy: $\exists D > 0$ such that

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

(can be replaced by a quasi-locality assumption).

$E_3$ Stability of the local energy: $\exists K \geq 0$ such that

$$V(0|\varphi) \geq -K.$$
Framework (of the presentation)

Restricted to stationary Gibbs point processes based on energy function related to some graph $G_2(\varphi)$:

$$V(\varphi) = \theta |\varphi| + \sum_{\xi \in G_2(\varphi)} u(\xi; \varphi), \quad \theta \in \mathbb{R}$$

$\implies$ pairwise interaction point processes.

satisfying the following Assumptions

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

$E_2$ Locality of the local energy: $\exists D > 0$ such that

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

(can be replaced by a quasi-locality assumption).

$E_3$ Stability of the local energy: $\exists K \geq 0$ such that

$$V(0|\varphi) \geq -K.$$
Framework (of the presentation)

Restricted to stationary Gibbs point processes based on energy function related to some graph $G_2(\varphi)$:

$$V(\varphi) = \theta |\varphi| + \sum_{\xi \in G_2(\varphi)} u(||\xi||), \quad \theta \in \mathbb{R}$$

$\implies$ pairwise interaction point processes.

satisfying the following Assumptions

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

$E_2$ Locality of the local energy: $\exists D > 0$ such that

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

(can be replaced by a quasi-locality assumption).

$E_3$ Stability of the local energy: $\exists K \geq 0$ such that

$$V(0|\varphi) \geq -K.$$
Framework (of the presentation)

Restricted to stationary Gibbs point processes based on energy function related to some graph $G_2(\varphi)$:

$$V(\varphi) = \theta |\varphi| + \sum_{\xi \in G_2(\varphi)} u(||\xi||), \quad \theta \in \mathbb{R}$$

$\implies$ pairwise interaction point processes.

satisfying the following Assumptions

**E$_1$** $V(\cdot)$ is invariant by translation.

**E$_2$** Locality of the local energy: $\exists D > 0$ such that

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

(can be replaced by a quasi-locality assumption).

**E$_3$** Stability of the local energy: $\exists K \geq 0$ such that

$$V(0|\varphi) \geq -K.$$
This framework includes:

- models based on the usual complete graph $G(\varphi) = \mathcal{P}_2(\varphi)$ with pairwise interaction function satisfying a hard-core or inhibition condition and with finite range.

- models based on the (slightly modified) Delaunay graph $G_2(\varphi) = \text{Del}_{2,\beta_0}(\varphi)$ with pairwise interaction function bounded and with finite range.

Definition of $\text{Del}_{2,\beta}^\beta_0(\varphi)$

Let $\text{Del}_3(\varphi)$ denote the “Delaunay triangles”, let $\beta_0 \in [0, \pi/3]$ and let $\beta(\psi)$ denote the smallest angle of some triangle $\psi$. Then,

$$\text{Del}_{3,\beta}^\beta_0(\varphi) = \{\psi \in \text{Del}_3(\psi), \beta(\psi) \geq \beta_0\} \text{ and } \text{Del}_{2,\beta}^\beta_0 = \bigcup_{\psi \in \text{Del}_{3,\beta}^\beta_0} \mathcal{P}_2(\psi)$$

Slight abuse: for $\beta_0$ small enough, $\text{Del}_{2,\beta}^\beta_0(\varphi) \simeq \text{Del}_2(\varphi)$
This framework includes

- models based on the usual complete graph $G(\varphi) = \mathcal{P}_2(\varphi)$ with pairwise interaction function satisfying a hard-core or inhibition condition and with finite range.

- models based on the (slightly modified) Delaunay graph $G_2(\varphi) = \text{Del}_{2,\beta}(\varphi)$ with pairwise interaction function bounded and with finite range.

**Definition of $\text{Del}_{2,\beta}^{\beta_0}(\varphi)$**

Let $\text{Del}_3(\varphi)$ denote the “Delaunay triangles”, let $\beta_0 \in [0, \pi/3]$ and let $\beta(\psi)$ denote the smallest angle of some triangle $\psi$. Then,

$$\text{Del}_{3,\beta}^{\beta_0}(\varphi) = \{\psi \in \text{Del}_3(\psi), \beta(\psi) \geq \beta_0\} \quad \text{and} \quad \text{Del}_{2,\beta}^{\beta_0} = \bigcup_{\psi \in \text{Del}_{3,\beta}^{\beta_0}} \mathcal{P}_2(\psi)$$

Slight abuse: for $\beta_0$ small enough, $\text{Del}_{2,\beta}^{\beta_0}(\varphi) \simeq \text{Del}_2(\varphi)$
Outline

1. Gibbs point processes
   - Basic definition
   - Existence conditions based on the energy function
   - Description of some Gibbs models

2. Statistical model and inference method

3. Asymptotic results
   - Consistency of the maximum likelihood estimator
   - Asymptotic normality of the maximum likelihood estimator

4. Description of some examples and short simulation
Parametrization of the function $u(\cdot)$

- Let $\theta \in \Theta$ where $\Theta$ is a compact of $\mathbb{R}^{p+1}$.
- Energy function described by:

$$V(\varphi; \theta) = \theta_1 |\varphi| + \sum_{\xi \in G_2(\varphi)} u(||\xi||; \theta)$$

Local energy: energy to insert $x$ in some configuration $\varphi$
Model parametrization (1)

**Parametrization of the function \( u(\cdot) \)**

- Let \( \theta \in \Theta \) where \( \Theta \) is a compact of \( \mathbb{R}^{p+1} \).
- Energy function described by:

\[
V(\varphi; \theta) = \theta_1|\varphi| + \sum_{\xi \in G_2(\varphi)} u(||\xi||; \theta)
\]

**Local energy: energy to insert \( x \) in some configuration \( \varphi \)**

\[
V(x|\varphi; \theta) = \theta_1 + \sum_{\xi \in G_2(\varphi \cup \{x\})} u(||\xi||; \theta) - \sum_{\xi \in G_2(\varphi)} u(||\xi||; \theta).
\]
Model parametrization (1)

Parametrization of the function \( u(\cdot) \)

- Let \( \theta \in \Theta \) where \( \Theta \) is a compact of \( \mathbb{R}^{p+1} \).
- Energy function described by:

\[
V(\varphi; \theta) = \theta_1 |\varphi| + \sum_{\xi \in G_2(\varphi)} u(||\xi||; \theta)
\]

Local energy: energy to insert \( x \) in some configuration \( \varphi \)

\[
V(x|\varphi; \theta) = \theta_1 + \sum_{y \in \varphi} u(||y - x||; \theta) \quad \text{when } G_2(\varphi) = \mathcal{P}_2(\varphi).
\]
Model parametrization (1)

Parametrization of the function $u(\cdot)$

- Let $\theta \in \Theta$ where $\Theta$ is a compact of $\mathbb{R}^{p+1}$.
- Energy function described by:

$$V(\phi; \theta) = \theta_1|\phi| + \sum_{\xi \in G_2(\phi)} u(||\xi||; \theta)$$

Local energy: energy to insert $x$ in some configuration $\phi$

$$V(x|\phi; \theta) = \theta_1 + \sum_{\xi \in G_2(\phi \cup x) \setminus G_2(\phi)} u(||\xi||; \theta) - \sum_{\xi \in G_2(\phi) \setminus G_2(\phi \cup \{x\})} u(||\xi||; \theta),$$

for a general graph such as $G_2(\phi) = Del_{2,\beta}(\phi)$.

J.-M. Billiot, J.-F. Coeurjolly, R. Drouilhet  University of Grenobl  MPLE for nearest-neighbours Gibbs point processes
Inhibition Delaunay Gibbs Process

Local energy:
\[ V(0|\varphi) := V(0 \cup \varphi) - V(\varphi) \]

\[ V_2^+(0|\varphi) = \sum_{\xi^+ \in G(0 \cup \varphi), \xi^+ \notin G(\varphi)} u(||\xi^+||; \theta) \]

\[ V_2^-(0|\varphi) = \sum_{\xi^- \in G(\varphi), \xi^- \notin G(0 \cup \varphi)} u(||\xi^-||; \theta) \]
Inhibition Delaunay Gibbs Process

Local energy:
\[ V(0|\varphi) := V(0 \cup \varphi) - V(\varphi) = \theta_1 + V_2^+(0|\varphi) - V_2^-(0|\varphi) \]

where

\[ V_2^+(0|\varphi) = \sum_{\xi^+ \in G(0 \cup \varphi) \setminus G(\varphi)} u(||\xi^+||; \theta) \]

\[ V_2^-(0|\varphi) = \sum_{\xi^- \in G(\varphi) \setminus G(0 \cup \varphi)} u(||\xi^-||; \theta) \]
Inhibition Delaunay Gibbs Process

**Local energy:**

\[ V(0|\varphi) := V(0 \cup \varphi) - V(\varphi) = \theta_1 + V_2^+(0|\varphi) - V_2^-(0|\varphi) \]

where

\[ V_2^+(0|\varphi) = \sum_{\xi^+ \in G(0 \cup \varphi), \xi^+ \notin G(\varphi)} u(||\xi^+||; \theta) \]

\[ V_2^-(0|\varphi) = \sum_{\xi^- \in G(\varphi), \xi^- \notin G(0 \cup \varphi)} u(||\xi^-||; \theta) = 0 \]

when \( G(\varphi) \subset G(0 \cup \varphi) \)
Inhibition Delaunay Gibbs Process

Local energy:
\[ V(0|\varphi) := V(0 \cup \varphi) - V(\varphi) = \theta_1 + V_2^+(0|\varphi) - V_2^-(0|\varphi) \]

when \( G_2(\varphi) = \text{Del}_{2,\beta}(\varphi) \)

\[ V_2^+(0|\varphi) = \sum_{\xi^+ \in G(0 \cup \varphi) \atop \xi^+ \notin G(\varphi)} u(||\xi^+||; \theta) \]

\[ V_2^-(0|\varphi) = \sum_{\xi^- \in G(\varphi) \atop \xi^- \notin G(0 \cup \varphi)} u(||\xi^-||; \theta) \]
Inhibition Delaunay Gibbs Process

Local energy:
\[ V(0|\varphi) := V(0 \cup \varphi) - V(\varphi) = \theta_1 + V_2^+(0|\varphi) - V_2^-(0|\varphi) \]

when \( G_2(\varphi) = Del_{2,\beta}^\beta(\varphi) \)

\[ V_2^+(0|\varphi) = \sum_{\xi^+ \in G(0 \cup \varphi) \setminus \xi^+ \notin G(\varphi)} u(||\xi^+||; \theta) \]

\[ V_2^-(0|\varphi) = \sum_{\xi^- \in G(\varphi) \setminus \xi^- \notin G(0 \cup \varphi)} u(||\xi^-||; \theta) \]
Inhibition Delaunay Gibbs Process
Model parametrization (2)

Particular case: exponential family

\[ u(||\xi||; \theta) = \sum_{i=2}^{p+1} \theta_i u_i(||\xi||). \]

⇒ Exponential energy function: \( V(\varphi; \theta) = \theta^T u(\varphi) \), where \( u(\cdot) = (u_1(\cdot), \ldots, u_{p+1}(\cdot)) \) with

\[ u_1(\varphi) = |\varphi| \quad \text{and} \quad u_i(\varphi) = \sum_{\xi \in G_2(\varphi)} u_i(||\xi||). \]

⇒ Exponential local energy function: \( V(x|\varphi; \theta) = \theta^T u(x|\varphi) \), with

\[ u(x|\varphi) = u(\varphi \cup \{x\}) - u(\varphi). \]
Particular case: exponential family

\[ u(||\xi||; \theta) = \sum_{i=2}^{p+1} \theta_i u_i(||\xi||). \]

\[ \Rightarrow \text{Exponential energy function: } V(\varphi; \theta) = \theta^T u(\varphi), \text{ where} \]
\[ u(\cdot) = (u_1(\cdot), \ldots, u_{p+1}(\cdot)) \text{ with} \]
\[ u_1(\varphi) = |\varphi| \quad \text{and} \quad u_i(\varphi) = \sum_{\xi \in G_2(\varphi)} u_i(||\xi||). \]

\[ \Rightarrow \text{Exponential local energy function: } V(x|\varphi; \theta) = \theta^T u(x|\varphi), \text{ with} \]
\[ u(x|\varphi) = u(\varphi \cup \{x\}) - u(\varphi). \]
Particular case: exponential family

\[ u(||\xi||; \theta) = \sum_{i=2}^{p+1} \theta_i u_i(||\xi||). \]

\[ \Rightarrow \text{Exponential energy function: } V(\varphi; \theta) = \theta^T u(\varphi), \]
where

\[ u(\cdot) = (u_1(\cdot), \ldots, u_{p+1}(\cdot)) \]
with

\[ u_1(\varphi) = |\varphi| \quad \text{and} \quad u_i(\varphi) = \sum_{\xi \in G_2(\varphi)} u_i(||\xi||). \]

\[ \Rightarrow \text{Exponential local energy function: } V(x|\varphi; \theta) = \theta^T u(x|\varphi), \]
with

\[ u(x|\varphi) = u(\varphi \cup \{x\}) - u(\varphi). \]
Let $d_1 < d_2 < \ldots < d_{p+1}$ some fixed real numbers. Define for $i = 2, \ldots, p + 1,$

\[ u_i(||\xi||) = 1(||\xi|| \in ]d_{i-1}, d_i]). \]

Thus, for our two Gibbs models

\[ V(\varphi; \theta) = \theta_1|\varphi| + \sum_{i=2}^{p+1} \theta_i u_i(\varphi), \]

where $u_i(\varphi)$ is interpreted as

- (when $G_2(\varphi) = \mathcal{P}_2(\varphi)$) the number of points of $\varphi$ in the class of distances $]d_{i-1}, d_i].$  
- (when $G_2(\varphi) = \text{Del}_{2,\beta}^0(\varphi)$) the number of Delaunay edges of $\varphi$ in the class of distances $]d_{i-1}, d_i].$
An example

\[ \theta = (1, 2, 4), \ d = (0, 20, 80) \]

\[ G_2(\varphi) = P_2(\varphi) \]

number of points = 266

\[ G_2(\varphi) = Del_{2,\beta}^\beta(\varphi) \]

number of points = 472
An example

\[ \theta = (1, 2, 4), \quad d = (0, 20, 80) \]

\[ G_2(\varphi) = P_2(\varphi) \]

\[ G_2(\varphi) = Del_{2,\beta}^\beta(\varphi) \]

Small 213 (0.6\%), Medium 15 (0\%), Large 35017 (99.4\%)

Small 415 (30.3\%), Medium 64 (4.7\%), Large 891 (65\%)
Inference method

Data

- Realization of a p.p. with energy function $V(\cdot; \theta^*)$ in some domain $\Lambda \subset \mathbb{R}^d$ satisfying Assumptions $E_1$ to $E_3$.
- $\theta^*$ true parameter to be estimated, $P_{\theta^*}$ associated Gibbs measure.

Usual parametric methods

- Maximum likelihood estimator: drawback = computation of the normalizing constant.
- Maximum pseudo-likelihood estimator (Besag (1968), Jensen and Møller (1991), ...)
- Takacs-Fiksel estimator (based on the refined Campbell theorem): competitive with respect to the MPLE.
Inference method

Data

- Realization of a p.p. with energy function \( V(\cdot; \theta^*) \) in some domain \( \Lambda \subset \mathbb{R}^d \) satisfying Assumptions \( E_1 \) to \( E_3 \).
- \( \theta^* \) true parameter to be estimated, \( P_{\theta^*} \) associated Gibbs measure.

Usual parametric methods

- Maximum likelihood estimator: drawback = computation of the normalizing constant.
- Maximum pseudo-likelihood estimator (Besag (1968), Jensen and Møller (1991), ...)
- Takacs-Fiksel estimator (based on the refined Campbell theorem): competitive with respect to the MPLE.
- Pseudo-likelihood function (Jensen and Møller (1991))

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

- Log-pseudo-likelihood function

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

Different contributions to asympt. results (when \( G_2(\varphi) = P_2(\varphi) \))


- Jensen and Kunsch (1994): asymptotic normality of the MPLE, exponential family \( \theta = (z, \beta) \)

Outline

1. Gibbs point processes
   - Basic definition
   - Existence conditions based on the energy function
   - Description of some Gibbs models

2. Statistical model and inference method

3. Asymptotic results
   - Consistency of the mple estimator
   - Asymptotic normality of the mple estimator

4. Description of some examples and short simulation
Asymptotic results - Introduction

Definition of the estimator

- Define
  \[ U_n(\theta) = -\frac{1}{|\Lambda_n|} LPL_{\Lambda_n}(\varphi; \theta) \]

- Maximum pseudo-likelihood estimator:
  \[ \hat{\theta}_n(\varphi) = \arg\max_{\theta \in \Theta} LPL_{\Lambda_n}(\varphi; \theta) = \arg\min_{\theta \in \Theta} U_n(\theta) \]

Lemma

Under certain Assumptions, \( U_n(\cdot) \) defines a contrast function: there exists a function \( K(\cdot, \theta^*) \) such that \( P_{\theta^*}\)-a.s.
\[ U_n(\theta) - U_n(\theta^*) \to K(\theta, \theta^*) , \text{ where } K(\cdot, \theta^*) \text{ is a positive function and is zero if and only if } \theta = \theta^*. \]

\[ \Rightarrow \text{ results on minimum contrast estimators } (\text{Guyon (1992))} \]
Asymptotic results - Introduction

**Definition of the estimator**

- Define

\[
U_n(\theta) = -\frac{1}{|\Lambda_n|} L_{PL}\Lambda_n (\varphi; \theta)
\]

- Maximum pseudo-likelihood estimator:

\[
\hat{\theta}_n(\varphi) = \arg\max_{\theta \in \Theta} L_{PL}\Lambda_n (\varphi; \theta) = \arg\min_{\theta \in \Theta} U_n(\theta)
\]

**Lemma**

Under certain Assumptions, \( U_n(\cdot) \) defines a contrast function: there exists a function \( K(\cdot, \theta^*) \) such that \( P_{\theta^*}\text{-a.s.} \)

\[
U_n(\theta) - U_n(\theta^*) \rightarrow K(\theta, \theta^*), \text{ where } K(\cdot, \theta^*) \text{ is a positive function and is zero if and only if } \theta = \theta^*.
\]

\[\rightarrow\] results on minimum contrast estimators (Guyon (1992))
Asymptotic results - Introduction

Definition of the estimator

- Define

\[ U_n(\theta) = -\frac{1}{|\Lambda_n|} LPL_{\Lambda_n} (\varphi; \theta) \]

- Maximum pseudo-likelihood estimator:

\[ \widehat{\theta}_n(\varphi) = \arg\max_{\theta \in \Theta} LPL_{\Lambda_n} (\varphi; \theta) = \arg\min_{\theta \in \Theta} U_n(\theta) \]

Lemma

Under certain Assumptions, \( U_n(\cdot) \) defines a contrast function: there exists a function \( K(\cdot, \theta^*) \) such that \( P_{\theta^*}\text{-a.s.} \)

\[ U_n(\theta) - U_n(\theta^*) \rightarrow K(\theta, \theta^*), \] where \( K(\cdot, \theta^*) \) is a positive function and is zero if and only if \( \theta = \theta^* \).

\[ \implies \text{results on minimum contrast estimators (Guyon (1992))} \]
Consistency of the MPLE: assumptions - general case

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

\( C_2 \) For all \( \theta \in \Theta \),
\[
V(0|\cdot; \theta) \in L^1(P_{\theta^*}).
\]

\( C_3 \) For all \( \theta \in \Theta \setminus \theta^* \),
\[
P_{\theta^*}\left( \{ \varphi, \ V(0|\varphi; \theta) \neq V(0|\varphi; \theta^*) \} \right) > 0
\]

\( C_4 \) For all \( \theta, \theta' \in \Theta \), there exists \( c > 0 \) such that \( P_{\theta^*} \)-a.s.
\[
|V(0|\Phi; \theta) - V(0|\Phi; \theta')| \leq ||\theta - \theta'||^c g(0, \Phi)
\]

where \( g(\cdot, \cdot) \) is a function such that for all \( x \),
\[
g(0, \Phi) = g(x, \Phi_x)
\]
and such that \( g(0, \cdot) \in L^1(P_{\theta^*}) \).
Consistency of the MPLE: assumptions - exponential case

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

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

$$u_i(0|\cdot) \in L^{1+\varepsilon}(P_{\theta^*}).$$

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

**Proposition (consistency)**

Assume $P_{\theta^*}$ stationary, then under Assumptions $C_1$ to $C_4$ in the general case or under Assumptions $C_1$, $C_{2,4}^{\text{exp}}$ and $C_3^{\text{exp}}$ in the exponential case, we have $P_{\theta^*} -$ almost surely, as $n \to +\infty$,

$$\hat{\theta}_n(\Phi) \to \theta^*$$

**Tools**

- Glötz Theorem, refined Campbell theorem.
- General ergodic theorems obtained by Nguyen and Zessin (1979).
- General result concerning the consistency of minimum contrast estimators obtained by Guyon (1992).
Consistency: statement of the result

**Proposition (consistency)**

Assume $P_{\theta^\star}$ stationary, then under Assumptions $C_1$ to $C_4$ in the general case or under Assumptions $C_1$, $C_{2,4}^{\text{exp}}$ and $C_3^{\text{exp}}$ in the exponential case, we have $P_{\theta^\star}$—almost surely, as $n \to +\infty$,

$$\hat{\theta}_n(\Phi) \to \theta^\star$$

**Tools**

- Glötz Theorem, refined Campbell theorem.
- General ergodic theorems obtained by Nguyen and Zessin (1979).
- General result concerning the consistency of minimum contrast estimators obtained by Guyon (1992).
Asymptotic normality: assumptions - general case (1)

**N₁** The point process is observed in a domain 
\( \Lambda_n \oplus D = \bigcup_{x \in \Lambda_n} B(x, D) \), where \( \Lambda_n \subset \mathbb{R}^2 \) can be decomposed into \( \bigcup_{i \in I_n} \Lambda(i) \) where for \( i = (i_1, i_2) \)

\[
\Lambda(i) = \left\{ z \in \mathbb{R}^2, \tilde{D} \left( i_j - \frac{1}{2} \right) \leq z_j \leq \tilde{D} \left( i_j - \frac{1}{2} \right), j = 1, 2 \right\}
\]

for some \( \tilde{D} > 0 \). As \( n \to +\infty \), we also assume that \( \Lambda_n \to \mathbb{R}^2 \) such that \( |\Lambda_n| \to +\infty \) and \( \left| \frac{\partial \Lambda_n}{\partial \Lambda_n} \right| \to 0 \)

**N₂** \( V(0|\cdot; \theta) \) is twice times differentiable in \( \theta = \theta^* \) and for all \( j, k = 1, \ldots, p + 1 \), there exists \( \varepsilon > 0 \) such that the variables

\[
\frac{\partial V}{\partial \theta_j} (0|\cdot; \theta^*) \in L^{1+\varepsilon} \quad \text{and} \quad \frac{\partial^2 V}{\partial \theta_j \partial \theta_k} (0|\cdot; \theta^*) \in L^1(P_{\theta^*})
\]
Asymptotic normality: assumptions - general case (2)

The matrix

$$\Sigma(\tilde{D}, \theta^*) = \tilde{D}^{-2} \sum_{|i| \leq \left\lfloor \frac{D}{D} \right\rfloor + 1} E_{\theta^*} \left( LPL_{\Lambda_0}^{(1)}(\Phi; \theta^*) LPL_{\Lambda_i}^{(1)}(\Phi; \theta^*)^T \right)$$

is symmetric and definite positive.

The vector $LPL_{\Lambda_i}^{(1)}(\varphi; \theta)$ is defined for $j = 1, \ldots, p + 1$ by

$$\left( LPL_{\Lambda_i}^{(1)}(\varphi; \theta) \right)_j = \int_{\Lambda(i)} \frac{\partial V}{\partial \theta_j}(x|\varphi; \theta) \exp \left( -V(x|\varphi; \theta) \right) dx - \sum_{x \in \varphi \setminus x} \frac{\partial V}{\partial \theta_j}(x|\varphi \setminus x; \theta).$$

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

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

where for $i = 1, \ldots, p + 1$, $(V^{(1)}(0|\varphi; \theta^*))_i = \frac{\partial V}{\partial \theta_i}(0|\varphi; \theta^*).$
Asymptotic normality: assumptions - general case (3)

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

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

and

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

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

Assumptions $N_2$ and $N_5$ (resp. $N_4$) can be replaced by $N_{2,5}^{\text{exp}}$ (resp. $N_4^{\text{exp}}$) where

$N_{2,5}^{\text{exp}}$ For $i = 1, \ldots, p + 1$, there exists $\varepsilon > 0$ such that

$$u_i(0|\cdot) \in L^{3+\varepsilon}(P_{\theta^*}).$$

$N_4^{\text{exp}} = C_3^{\text{exp}}$
Proposition (asymptotic normality)

Assume $P_{\theta^*}$ stationary, then under Assumptions $N_1$ to $N_5$ in the general case or under Assumptions $N_1$ $N_{2,5}$ $C^\text{exp}_3$ and $N_3$ in the exponential case, we have, for any fixed $\tilde{D}$ fixed

$$|\Lambda_n|^{1/2} \sum_n(\tilde{D}, \hat{\theta}_n)^{-1/2} U_n^{(2)}(\hat{\theta}_n) \left( \hat{\theta}_n - \theta^* \right) \to \mathcal{N}(0, I_{p+1}),$$

where for some $\theta$ and some finite configuration $\varphi$

$$\hat{\Sigma}_n(\tilde{D}, \theta) = |\Lambda_n|^{-1} \tilde{D}^{-2} \sum_{i \in I_n} \sum_{|j-i| \leq \left\lfloor \frac{D}{\tilde{D}} \right\rfloor + 1, j \in I_n} \mathbf{LPL}_{\Lambda_i}^{(1)}(\varphi; \theta) \mathbf{LPL}_{\Lambda_j}^{(1)}(\varphi; \theta)^T$$

Tools

- Asympt. normality for minimum contrast estimators (Guyon (1992)).
Asymptotic normality: statement of the result

**Proposition (asymptotic normality)**

Assume $P_{\theta^*}$ stationary, then under Assumptions $N_1$ to $N_5$ in the general case or under Assumptions $N_1 N_{2,5}^\text{exp} C_3^\text{exp}$ and $N_3$ in the exponential case, we have, for any fixed $\tilde{D}$ fixed

$$\left|\Lambda_n\right|^{1/2} \hat{\Sigma}_n(\tilde{D}, \hat{\theta}_n)^{-1/2} \underline{U}_n^{(2)}(\hat{\theta}_n) \left(\hat{\theta}_n - \theta^*\right) \to \mathcal{N}(0, I_{p+1}),$$

where for some $\theta$ and some finite configuration $\varphi$

$$\hat{\Sigma}_n(\tilde{D}, \theta) = \left|\Lambda_n\right|^{-1} \tilde{D}^{-2} \sum_{i \in I_n} \sum_{|j-i| \leq \left\lceil\frac{D}{\tilde{D}}\right\rceil + 1, j \in I_n} \text{LPL}^{(1)}_{\Lambda_i}(\varphi; \theta) \text{LPL}^{(1)}_{\Lambda_j}(\varphi; \theta)^T.$$

**Tools**

- Asympt. normality for minimum contrast estimators (Guyon (1992)).
Outline

1. Gibbs point processes
   - Basic definition
   - Existence conditions based on the energy function
   - Description of some Gibbs models

2. Statistical model and inference method

3. Asymptotic results
   - Consistency of the mle estimator
   - Asymptotic normality of the mle estimator

4. Description of some examples and short simulation
A useful corollary

A particular class of exponential family

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

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

Assumption $M \implies C_{2,4}^{\exp}$ and $N_{2,5}^{\exp}$.

Corollary

Assume $P_{\theta^*}$ stationary, then under Assumption $M$ and $C_3^{\exp}$, the consistency is valid. And in addition under Assumption $N_3$ the asymptotic normality is ensured.
Back to the multi-Strauss pairwise interaction p.p.

\[ V(\varphi; \theta) = \theta_1 \|\varphi\| + \sum_{i=2}^{p+1} \theta_i \sum_{\xi \in \text{Del}_{2,\beta}^{\beta_0}(\varphi)} 1(\|\xi\| \in ]d_{i-1}, d_i]). \]

Assumption \( M, C_3^{\exp} \) and \( N_3 \)

- **Assumption \( M \):** proved in Bertin, Billiot and Drouilhet (1999).
- **Assumption \( C_3^{\exp} \):** verified by considering particular sets of configurations of two points in a domain \( \Delta = \{ z \in \mathbb{R}^2, -D \leq z_i \leq D, i = 1, 2 \} \).
- **Assumption \( N_3 \):** verified for this model by using an inequality obtained by Jensen and Künsch and then by considering particular sets of configurations of three points in \( \bigcup_{|i| \leq 1} \Lambda(i) \).
Short simulation study

Parameters

\[ \theta^* = (0, 2, 4), \quad d = (0, 20, 80) \]

\[ m = 5000 \text{ replications generated in the domain } [-600, 600]^2. \]
Short simulation study

Parameters

- $\theta^* = (0, 2, 4), \ d = (0, 20, 80)$
- $m = 5000$ replications generated in the domain $[-600, 600]^2$.

<table>
<thead>
<tr>
<th>Domain $\Lambda_n$</th>
<th>Estimations of $\theta_2^*$</th>
<th>Estimations of $\theta_3^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (Std Dev.)</td>
<td>Mean of Estim. (Std Dev.)</td>
</tr>
<tr>
<td>$[-250, 250]^2$</td>
<td>2.068 (0.104)</td>
<td>4.382 (0.786)</td>
</tr>
<tr>
<td>$[-350, 350]^2$</td>
<td>2.049 (0.071)</td>
<td>4.223 (0.551)</td>
</tr>
<tr>
<td>$[-450, 450]^2$</td>
<td>2.041 (0.056)</td>
<td>4.144 (0.436)</td>
</tr>
</tbody>
</table>
Perspectives

- concerning the multi-Strauss pairwise interaction point process based on the Delaunay graph: automatic estimation of the different $d_i$, $i = 1, \ldots, p + 1$.
- A larger simulation study is needed:
  1. to compare models based on the Delaunay graph and the complete graph.
  2. to investigate other nearest-neighbour models, models based on cliques of order larger than 2, marked nearest-neighbour Gibbs point processes,\ldots
- Nonparametric estimation of the pairwise interaction function for nearest-neighbour Gibbs point processes.
Delaunay graph, delaunay triangulation
Delaunay graph, delaunay triangulation