A central limit theorem for conditionally centred functional of a Markov random field

Carlo Gaetan
Dipartimento di Statistica, Università Ca' Foscari - Venezia, Italy
and
Xavier Guyon
SAMOS, Université Paris 1, France

May 5, 2004

Abstract

We prove a central limit theorem for empirical sums of a conditionally centred functional of a Markov random field on a non necessarily regular set of sites S. A studentized version of this theorem is also given with a random normalisation. Since positive definiteness of the variance of the sums is crucial for these results, we introduce the notion of conditionally separating partition and we give tools to verify such a positive definiteness. Examples of Ising an Gaussian Markov random field are studied and central limit theorems are shown regardless of phase transition.

Keyword: Central limit theorem, Markov random fields, Ising model, Random environment, Gaussian model, Conditionally centred functional, Conditionally separating partition, Irregular set.

1 Introduction

In recent years there has been interest in establishing central limit theorems (CLTs) for random fields. Bolthausen [2] obtained a CLT for a stationary field on the regular lattice $S = \mathbb{Z}^d$ under weak dependency mixing conditions (see also Dedecker, [8]); a non stationary version of this result is given in Guyon [11]. Guyon and Künsch [13] have shown that a CLT for a stationary and ergodic field on $\mathbb{Z}^d$ can be obtained without mixing conditions exploiting a conditional centring property that reminds one of martingale difference sequences on $\mathbb{Z}^1$. This idea has been applied to stationary and ergodic point random fields on $\mathbb{R}^d$ by Jensen and Künsch [15]. More recently Comets and Janzura [6] have proved a CLT for a sum of conditionally centred random fields under a moment condition and the assumption that the empirical variance does not vanish in the order of the volume. The authors applied the result to Markov random fields (MRF) on $\mathbb{Z}^d$ with shift invariant potentials. A nice consequence is the asymptotic normality of the maximum pseudo-likelihood estimator (MPLE, Besag, [1]) for MRF on $\mathbb{Z}^d$, whether phase transition occurs or not.

In this work, we establish a CLT for sums of a conditionally centred functional of a MRF defined on S, despite the regularity of S. This is motivated by many applications where S is not a regular lattice (see for example Cliff and Ord [4], Cressie [7], Haining [14] and Tiefelsdorf [17]) and shift invariance for the potentials is no longer valid. Moreover we obtain a studentized form of CLT as in Comets and Janzura [6]. A basic ingredient is the positive definiteness of the variance of the empirical means. We give tools that allows us to verify such property. These tools, based on the notion of conditionally separating partition of S, are free of regularity assumption for the lattice S, and/or shift invariance for the potentials of the MRF.

The paper is organised as follows. Section 2 gives some definitions and background materials. Section 3 contains our main results and Section 4 presents the tools to verify positive definiteness of the variance of the sum. Finally, in Section 5, we give some examples of applications.

2 Preliminaries

Let $X = (X_i, i \in S)$ be a random field on an infinite countable set S, with states in a measurable space $(E, \mathcal{E})$. For $\Lambda \subset S$, we denote $X_\Lambda = (X_i, i \in \Lambda)$

and $X^\Lambda = (X_i, i \in S \setminus \Lambda)$. By $\mathcal{F}$, $\mathcal{F}^\Lambda$ and $\mathcal{F}^i$ we also denote the σ-field generated by X, X^Λ and $X^{\{i\}}$, respectively. A configuration of X_Λ is noted by x_Λ. Let $\mathcal{G}$ be a symmetric graph on S without loops: i and j are said neighbours if $\{i, j\} \in \mathcal{G}$. The boundary (respectively the neighbourhood) of Λ is

$$\partial\Lambda = \{i \in S \setminus \Lambda : \exists j \in S \text{ with } \{i, j\} \in \mathcal{G}\} \text{ (respectively } \Lambda^* = \Lambda \cup \partial\Lambda).$$

For simplicity, we write $\partial i = \partial\{i\}$. The first and the second order neighbourhoods of i are $V_i = \{i\} \cup \partial V_i$ and $(V_i)^* = \bigcup_{k \in V_i} V_k = \{j \in S : \exists k \in S$ such that i and $j \in V_k\}$.

We suppose that X is a $\mathcal{G}$-MRF, i.e. the law of X_Λ given x^Λ depends only on $x_{\partial\Lambda}$, and we focus our attention on a derived field $Y = (Y_i)$, which is a local and multidimensional functional of X defined by

$$Y_i = f_i(X_{V_i}), \qquad \text{for all } i \in S \tag{1}$$

where $f_i : \mathbb{E}^{V_i} \longrightarrow \mathbb{R}^d$ is a family of measurable and integrable functions. The Y_i are also assumed conditionally centred, namely

$$E(Y_i / \mathcal{F}^i) = 0. \tag{2}$$

The Markov property of X entails that if $i \neq j$ are not neighbourhoods, then Y_i and Y_j are conditionally independent with respect to $\mathcal{F}^{\{i,j\}}$.

Let (Λ_n) be an increasing sequence of finite subset of S such that card(Λ_n)= $|\Lambda_n| \longrightarrow \infty$ if $n \to \infty$. In the next section we prove a CLT for the sums $S_n = \sum_{i \in \Lambda_n} Y_i$.

3 Main results

We consider first the univariate case $Y_i \in \mathbb{R}$. We denote

$$A_n = \sum_{i \in \Lambda_n} \sum_{j \in \Lambda_n \cap V_i} Y_i Y_j = \sum_{i \in \Lambda_n} Y_i S_{i,n}$$

where $S_{i,n} = \sum_{j \in \Lambda_n \cap V_i} Y_j$, and $\mu_q(Y) = \sup_{i \in S} E(|Y_i|^q)$. A_n is integrable provided $\mu_2(Y) < \infty$. In this case, due to (2) and the conditional independence, we have

$$E(A_n) = \sum_{i,j \in \Lambda_n} E(Y_i Y_j) = \mathrm{Var}(S_n) = \sigma_n^2. \tag{3}$$

Proposition 1 *Let X be a Markov random field on S, Y the local functional of X defined by (1). Assume that Y is conditionally centred (2) and*
(N1) : $\mu_4(Y) < \infty$;
(N2) : $M = \sup\{|V_i|, i \in S\} < \infty$;
(N3) : $\liminf_n |\Lambda_n|^{-1}\sigma_n^2 > 0$ where $\sigma_n^2 = \mathrm{Var}(S_n)$.
Then

$$\sigma_n^{-1} S_n \xrightarrow{\mathcal{D}} \mathcal{N}(0,1).$$

Proof. We adapt the proof of Theorem 3.3.1 in Guyon [11] (see also Guyon and Künsch [13]). According to Stein [16], we prove that for every $\lambda \in \mathbb{R}$

$$\lim_{n\to\infty} E((i\lambda - \overline{S}_n)e^{i\lambda\overline{S}_n}) = 0. \tag{4}$$

where $\overline{S}_n = \sigma_n^{-1}S_n$. Following Bolthausen [2] we have

$$(i\lambda - \lambda\overline{S}_n)e^{i\lambda\overline{S}_n} = A_{n,1} - A_{n,2} - A_{n,3}$$

where

$$\begin{aligned}
A_{n,1} &= i\lambda e^{i\lambda\overline{S}_n}\left(1 - \sigma_n^{-2}\sum_{j\in\Lambda_n} Y_j S_{j,n}\right), \\
A_{n,2} &= \sigma_n^{-1} e^{i\lambda\overline{S}_n}\sum_{j\in\Lambda_n} Y_j\left(1 - i\lambda\overline{S}_{j,n} - e^{-i\lambda\overline{S}_{j,n}}\right), \\
A_{n,3} &= \sigma_n^{-1}\sum_{j\in\Lambda_n} Y_j e^{i\lambda(\overline{S}_n - \overline{S}_{j,n})}
\end{aligned}$$

and $\overline{S}_{j,n} = \sigma_n^{-1}S_{j,n}$.
From (N1) we get that $E|A_{n,1}|^2 < \infty$ and

$$\begin{aligned}
E\,|A_{n,1}|^2 &= \lambda^2 E\left(1 - \sigma_n^{-2}\sum_{i\in\Lambda_n} Y_i S_{i,n}\right)^2 = \lambda^2\sigma_n^{-2}\mathrm{Var}(\sum_{i\in\Lambda_n} R_{i,n}) \qquad (5)\\
&= \lambda^2\sigma_n^{-4}\sum_{i\in\Lambda_n}\mathrm{Var}(R_{i,n}) + \sum_{i\in\Lambda_n}\sum_{j\in\Lambda_n : V_i^*\cap V_j^*\neq\emptyset}\mathrm{cov}(R_{i,n}, R_{j,n}) \\
&\leq \lambda^2\sigma_n^{-4}|\Lambda_n| \times (1 + M^4) \times \mu_4,
\end{aligned}$$

with $R_{j,n} = Y_j S_{j,n}$. The inequality follows since if $V_i^* \cap V_j^* = \emptyset$, then $R_{i,n}$ and $R_{j,n}$ are conditionally uncorrelated with respect to to $\mathcal{F}^{V_i^*\cup V_j^*}$. On the other hand we have $((V_i)^*)^* = \{j \in |\Lambda_n| : V_i^* \cap V_j^* \neq \emptyset\}$. His cardinality is bounded by $|V_i|^4$ and consequently by M^4 according to (N2).

Since $|e^{iy} - iy - 1| \leq y^2/2$ for every $y \in \mathbb{R}$, we have

$$\begin{aligned}
E|A_{n,2}| &\leq \frac{\lambda^2}{2}\sigma_n^{-3} \sum_{j\in\Lambda_n} E\left\{|Y_j| S_{j,n}^2\right\} \leq \frac{\lambda^2}{2}\sigma_n^{-3}\mu_3 \sum_{j\in\Lambda_n} |V_j|^2 \\
&\leq \frac{\lambda^2}{2} \times \sigma_n^{-3} \times \mu_3 \times M^2 \times |\Lambda_n|\,.
\end{aligned}$$

Denote $S_{j,n}^* = \overline{S}_n - \overline{S}_{j,n} = \sigma_n^{-1} \sum_{i\in(\Lambda_n \cap V_j^c)} Y_i$. Since $S_{j,n}^* \in \mathcal{F}^{V_j}$ and Y is conditionally centred, we have

$$E(A_{n,3}) = \sigma_n^{-1} \sum_{j\in\Lambda_n} E[Y_j e^{i\lambda S_{j,n}^*}] = 0.$$

The result follows since the expectation of each $A_{n,k}$, $k = 1, 2, 3$ goes to zero by (N3).

Since $\sigma_n^2 = \mathrm{Var}(S_n)$ is usually unknown, a studentized version of Proposition 1 can be useful (Comets and Janzura, [6]). According to (3) a natural estimator for σ_n^2 is A_n.

The next result requires the following definition: $C \subset S$ is a strong coding subset of S if for any $i, j \in C$, $i \neq j$, i and j are not second order neighbour sites, namely $V_i^* \cap V_j^* = \emptyset$. Now we set two additional assumptions for $\mathcal{G}$:

(M1) S is the union of K disjoint strong coding subsets C_k, $k = 1, \ldots, K$.

(M2) for every $k = 1, \ldots, K$, $\lim_n |C_k \cap \Lambda_n| = +\infty$.

Proposition 2 *Let $\xi_n = A_n^{-1/2} S_n$ if $A_n > 0$, $\xi_n = 0$ otherwise. Then, under conditions (N1-N2-N3) and (M1-M2):*

$$\xi_n \xrightarrow{\mathcal{D}} \mathcal{N}(0,1).$$

Proof. Denote $R_{i,n} = Y_i S_{i,n}$, $\tilde{R}_{i,n} = R_{i,n} - E(R_{i,n})$ and $D_{k,n} = \sum_{i\in\Lambda_n\cap C_k} \tilde{R}_{i,n}$. For large n we have

$$\frac{A_n}{\sigma_n^2} - 1 = \frac{\sum_{i\in\Lambda_n} \tilde{R}_{i,n}}{\sigma_n^2} = \frac{\sum_{k=1}^K \frac{D_{k,n}}{|C_k\cap\Lambda_n|} \frac{|C_k\cap\Lambda_n|}{|\Lambda_n|}}{\frac{\sigma_n^2}{|\Lambda_n|}}\,.$$

$\tilde{R}_{i,n}$, $i \in C_k$, have zero means and variances bounded by $\mu_4(Y)(1 + M^2)$. Moreover, they are conditionally independent with respect to $\mathcal{F}^{C_k}$. By the strong law of large numbers for L^2 centred and independent variables (Breiman, [3, Theorem 3.27]), we have for any configuration $x_{S\setminus C_k}$

$$\lim_n \frac{D_{k,n}}{|C_k \cap \Lambda_n|} = 0, \qquad P_{x_{S\setminus C_k}} - \text{a.s. .}$$

Since this limit does not depend on $x_{S\setminus C_k}$, the limit still holds almost surely for every x and we have

$$\lim_n \frac{A_n}{\sigma_n^2} = 1, \qquad \text{a.s. .}$$

On the other hand, (N3) entails that $\lim_n P(A_n \leq 0) = 0$ and we obtain the required result.

Now we consider briefly the the multivariate case, i.e. $Y_i \in \mathbb{R}^d$. Let $\|\cdot\|$ be the euclidean norm of $\mathbb{R}^d$ and, for a symmetric definite positive matrix A, denote $A^{r/s} = \Gamma\Lambda^{r/s}\Gamma^T$, for r and $s > 0$ integer numbers. Here $\Lambda^{r/s} = \mathrm{diag}(\lambda_i^{r/s})$, where (λ_i) are the eigenvalues of A and Γ is the matrix of columns eigenvectors with unit norm. We have $A_n = \sum_{i\in\Lambda_n}\sum_{j\in\Lambda_n\cap V_i} Y_i Y_j^T$, $\Sigma_n = \mathrm{Var}(S_n) = E(A_n)$, and we replace conditions (N1-N3) by :

(N1') : $\mu_4(Y) = \sup_{j\in S} E(\|Y_j\|^4) < \infty$;

(N3') : $\liminf_n |\Lambda_n|^{-1}\Sigma_n \geq \Delta$, where Δ is a positive definite matrix.

Proposition 3 *Under the conditions (N1'-N2-N3') we have*

$$\Sigma_n^{-1/2} S_n \xrightarrow{\mathcal{D}} \mathcal{N}(0, I).$$

Moreover, let $\xi_n = A_n^{-1/2} S_n$ if A_n is a positive definite matrix, $\xi_n = 0$ otherwise. Under the additional conditions (M1-M2),

$$\xi_n \xrightarrow{\mathcal{D}} \mathcal{N}(0, I).$$

Proof. The proof follows easily if we consider two linear combination $a^T\Sigma^{-1/2}S_n$ and $a^T A_n^{-1/2} S_n$ for $a \neq 0$ and we apply Propositions 1 and 2, respectively.

4 Minorization of the variance of the empirical mean

We provide some tools to verify that $\mathrm{Var}(S_n)$ is positive definite. We use the following conditional independence minorization as suggested by J.L. Jensen (see Guyon and Künsch [13] and Jensen and Künsch [15])

$$\mathrm{Var}(T) = E_{\mathcal{H}}(\mathrm{Var}(T/\mathcal{H})) + Var_{\mathcal{H}}(E(T/\mathcal{H})) \geq E_{\mathcal{H}}(\mathrm{Var}(T/\mathcal{H})). \quad (6)$$

Here, $\mathcal{H}$ is a sub σ-field of $\mathcal{F}$, and T is a $\mathcal{F}$-measurable variable with finite variance. Such as argument has been applied to ergodic Ising model on $\mathbb{Z}^d$ (Guyon and Künsch, [13]), pairwise interaction point processes (Jensen and Künsch, [15]) and Markov field dynamics (Guyon and Hardouin, [12]). Here we take $T = S_n$ and $\mathcal{H}$ will be specify below.

4.1 Conditionally separating partition

We define a specific partition of S, called conditionally separating partition (CSP), in the following way: we plunge S into S^+, a over-set of S; then we consider a subset $C \subset S^+$, and for each $i \in C$, we set $W_i \subset S$ such that $\mathcal{P} = \{W_i, i \in C\}$ is a partition of S. Note that $\mathcal{P}$ is indexed by C.

Definition 1 *A partition $\mathcal{P} = \{W_i, i \in C\}$ of S is a CSP if for every $i \in C$ we have $W_i^* \backslash \{i\} \subset S \backslash C$*

To clarify this definition, we give three examples.

$S = \mathbb{Z}^2$ and 4-nearest neighbours graph

Take $S^+ \equiv S = \mathbb{Z}^2$, $\mathcal{G}$ the 4-nearest neighbours graph, $C = \{3i, i \in \mathbb{Z}^2\}$, $V_i = \{j \in \mathbb{Z}^2 : \|i - j\|_1 \leq 1\}$. In this case $\mathcal{P} = \{W_i, i \in C\}$ where $W_i = \{j \in \mathbb{Z}^2 : \|i - j\|_\infty \leq 1\}$ is a CSP and $W_i^* = \{j \in \mathbb{Z}^2 : \|i - j\|_1 \leq 3 \text{ and } \|i - j\|_\infty \leq 2\}$ (see Figure 1). For $\Lambda_n = [-n, +n]^2$ and $C_n = C \cap \Lambda_n$, the asymptotic rate, $\liminf_{n\to\infty} \frac{|C_n|}{|\Lambda_n|}$, is positive and equal to $\frac{1}{9}$.

Regular lattice $\mathbb{Z}^2$ with holes and nearest neighbours graph

Let $\mathbb{T} = (1,1) + 2 \times \mathbb{Z}^2$ the holes set, $S = \mathbb{Z}^2 \backslash \mathbb{T}$, $S^+ = \mathbb{Z}^2$, $C = 6 \times \mathbb{Z}^2$ (see Figure 2). S^+ is an overset of S and C is not contained in S. $\mathcal{G}$ is still the

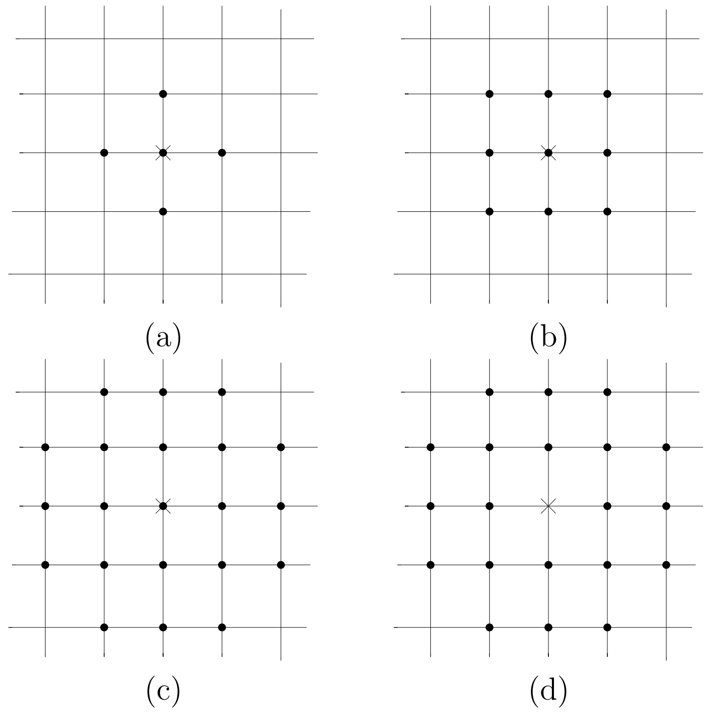


Figure 1: (a) Neighbourhood V_i of site i ($\times$) for the 4-nearest neighbours, (b) second order neighbourhood (Moore neighbourhood) W_i, (c) neighbourhood W_i^* and (d) neighbourhood U_i.

nearest neighbours graph: two thirds of the sites have 2-nearest neighbours whereas one third have 4-nearest neighbours. $\mathcal{P} = \{W_i, i \in C\}$ is a CSP if we take $W_i = \{j \in S : \|i - j\|_\infty \leq 1\}$. For $U_i = W_i^* \backslash \{i\}$, there are three types of (W_i, W_i^*, U_i), namely

1. If $i \in (0,0) + 6 \times \mathbb{Z}^2$, W_i, W_i^* and U_i contain 5, 9 and 8 sites respectively (see Figure 3);

2. for $i \in (3,0) + 6 \times \mathbb{Z}^2$, or $i \in (0,3) + 6 \times \mathbb{Z}^2$, W_i, W_i^* and U_i contain 7, 13 and 12 sites (see Figure 4);

3. for $i \in (3,3) + 6 \times \mathbb{Z}^2$, W_i, $W_i^* \equiv U_i$ contain 8 and 16 sites (see Figure 5).

If $\Lambda_n = [-n, +n]^2$, the asymptotic rate of C is positive and equal to $\frac{4}{27}$.

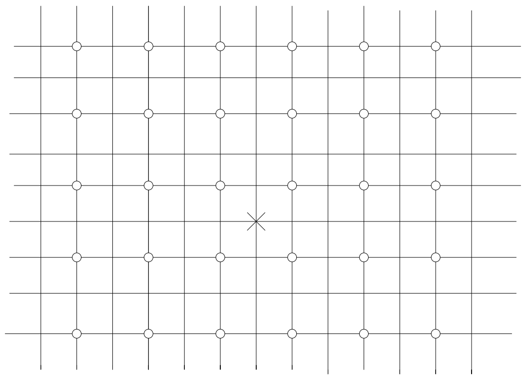
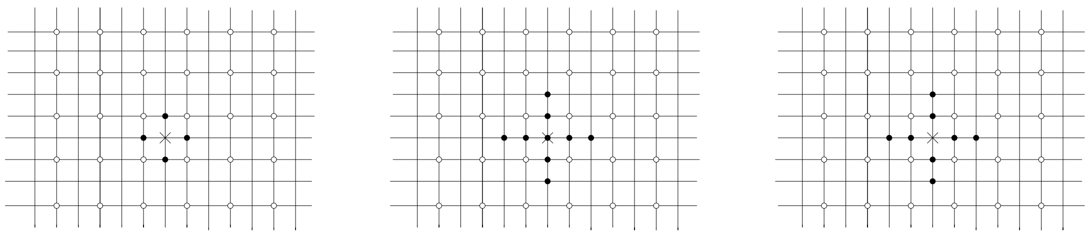


Figure 2: Regular lattice with holes ($\circ$).



(a) (b) (c)

Figure 3: (a) Neighbourhood W_i, (b) neighbourhood W_i^*, (c) neighbourhood U_i for $i \in (0,0) + 6 \times \mathbb{Z}^2$.

A finite and irregular lattice S

This example deals with a finite and irregular lattice S with 41 sites ($\circ$ and $\bullet$) and a graph defined as in Figure 6. We take $S^+ = S \cup \{\times\}$ and we consider C with 10 points ($\bullet$) and the point $\times$. Note that $C \nsubseteq S$. Then the partition $\mathcal{P} = \{W_i,\ i \in C\}$, where W_i is delimited by the dotted lines, is a CSP. The partition rate is equal to $\frac{9}{42}$.

Given a CSP $\mathcal{P}$ we can rearrange the terms of S_n as

$$S_n = \sum_{i \in \Lambda_n} Y_i = \sum_{i \in C_n} \sum_{j \in W_{i,n}} Y_j = \sum_{i \in C_n} G_{i,n}$$

where $C_n = \{i \in C : W_{i,n} \neq \emptyset\}$, $W_{i,n} = W_i \cap \Lambda_n$ and $G_{i,n} = \sum_{j \in W_{i,n}} Y_j$. Moving from this remark, Lemma 1 gives a simple and useful property for

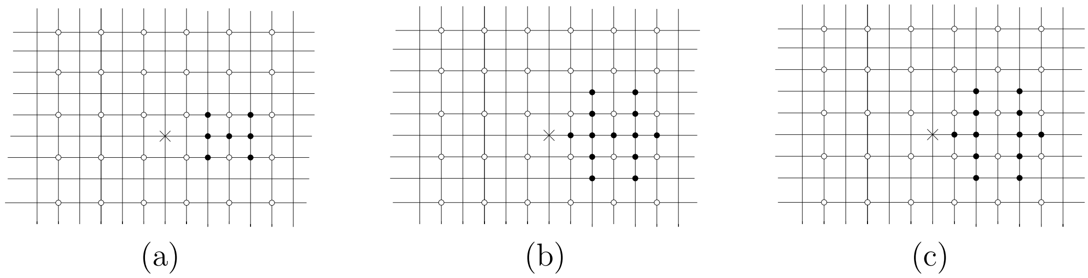


Figure 4: (a) Neighbourhood W_i, (b) neighbourhood W_i^*, (c) neighbourhood U_i for $i \in (3,0) + 6 \times \mathbb{Z}^2$.

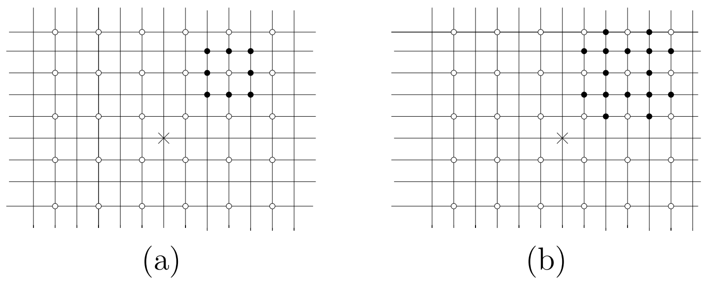


Figure 5: (a) Neighbourhood W_i, (b) neighbourhood $W_i^* \equiv U_i$ for $i \in (3,3) + 6 \times \mathbb{Z}^2$.

the partial sums $G_{i,n}$.

Lemma 1 *For two different sites of C , l and k, $G_{l,n}$ and $G_{k,n}$ are conditionally independent with respect to $\mathcal{F}^C$.*

Thus we have

$$\mathrm{Var}(S_n/\mathcal{F}^C) = \sum_{i \in C_n} \mathrm{Var}(G_{i,n}/\mathcal{F}^C). \tag{7}$$

According to (6), a strategy for verifying $(N3)$ consists in three steps :

1. bounding from below $\mathrm{Var}(G_{i,n}/\mathcal{F}^C)$, $i \in C$;
2. bounding from below $E_{\mathcal{F}^C}(\mathrm{Var}(G_{i,n}/\mathcal{F}^C))$;
3. controlling the asymptotic rate of C, i.e. $\liminf_{n\to\infty} \frac{|C_n|}{|\Lambda_n|}$.

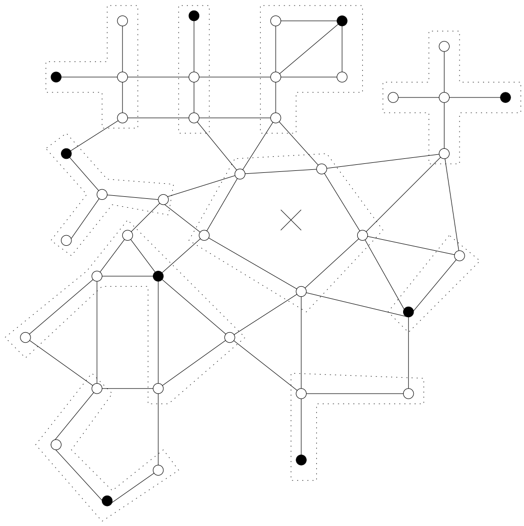


Figure 6: An example of irregular finite lattice.

Step 3 is purely combinatorial and requires a direct examination of $\mathcal{G}$ and S^+. For the other steps, it is sufficient to look at points $i \in C$ provided that $|\partial\Lambda_n|/|\Lambda_n| \longrightarrow 0$. We can also restrict our investigation to points $i \in C \cap S$ since, if $i \notin S$, then $G_{i,n}$ is $\mathcal{F}^C$-constant.

Thus we focus on $i \in C \cap S$ such that $i \notin \partial\Lambda_n$. In this case, we have $(G_{i,n}/\mathcal{F}^C) = g_i(X_i, x_{U_i})$. No universal tools are available for bounding general expressions of $\mathrm{Var}(G_{i,n}/\mathcal{F}^C)$ from below. This variance is positive provided that $(G_{i,n}/\mathcal{F}^C)$ is not constant; therefore we have to find x_{U_i} such that $(G_{i,n}/\mathcal{F}^C)$ is not constant. We need only study minorizations for sites $i \in C_1 \subset C \cap S$ provided $\liminf_{n\to\infty} |C_{1,n}| \,/\, |C_n| = \kappa > 0$, with $C_{1,n} = C_1 \cap \Lambda_n$. In some cases, we can choose C_1 such that the geometry of U_i does not depend on $i \in C_1$ and $\pi_{U_i}(\cdot/x_{U_i})$ is bounded from below by $c\times\ \pi(\cdot/x_{U_i})$, where c is a strictly positive constant and $\pi(\cdot/x_{U_i})$ only depends on x_{U_i}.

For the second step, note that

$$E_{\mathcal{F}^C}(\mathrm{Var}(G_{i,n}/\mathcal{F}^C)) = \int \mathrm{Var}(G_{i,n}/x_U)\pi_U(x_U)\lambda(dx_U).$$

A bound for this expression is obtained when π_U is bounded from zero and $\mathrm{Var}(G_{i,n}/x_U)$ is positive over a set of x_U with positive measure. If the state space is compact and $\pi_U(x_U/x_{\partial U})$ is strictly positive and continuous in

$(x_U, x_{\partial U})$, we have

$$\pi_U(x_U) = \int \pi_U(x_U/x_{\partial U})\pi_{\partial U}(x_{\partial U})\lambda(x_{\partial U}) \geq \rho \int \pi_{\partial U}(x_{\partial U})\lambda(x_{\partial U}) = \rho$$

with $\rho = \inf_{x_U, x_{\partial U}} \pi_U(x_U/x_{\partial U}) > 0$.

5 Examples

5.1 Isotropic Ising model in a random environment on $\mathbb{Z}^2$

Let $S = \mathbb{Z}^2$, $p > 0$ a positive probability and $\mathcal{G}(p)$ the percolation graph on S defined as follow : let $L = \{L_{i,j}, i, j \in S, \|i - j\|_1 = 1\}$ be a collection of independent identically distributed Bernoulli random variables with parameter p; then, for each $i \in S$, the neighbourhood of i is

$$\partial i = \{j \in S \text{ s.t. } \|i - j\|_1 = 1 \text{ and } L_{i,j} = 1\}$$

An nearest neighbourhood Ising model for this graph is a probability π on $\{-1, +1\}^{\mathbb{Z}^2}$ such that

$$\pi_i(x_i \mid x^i) = \pi_i(x_i \mid x_{\partial i}) = \frac{\exp\{x_i(\alpha + \beta v_i)\}}{2\cosh(\alpha + \beta v_i)} \tag{8}$$

where $v_i = \sum_{j \in \partial i} x_j$.

For $p = 1$, this is the 4-nearest neighbour Ising model and for some values of $\theta = (\alpha, \beta)$, there are more than one probability satisfying (8) (Georgii, [10]). This causes difficulties in studying asymptotic properties of local estimators like maximum pseudo-likelihood estimator (MPLE), coding estimator (see Comets [5] and Guyon [11]). For $0 < p < 1$, the graph is not regular and potentials are not shift invariant.

Assume that X is observed on the neighbourhood Λ_n^* of $\Lambda_n = [-n, +n]^2$ and we concentrate on MPLE $\hat{\theta}$, a maximiser of the logarithm of the pseudo-likelihood

$$U_n(\theta) = \sum_{i \in \Lambda_n} \log \pi_i(x_i \mid x^i; \theta).$$

Derivation of asymptotic normality for $\hat{\theta}$ involves examination of asymptotic properties of the derivative of U_n

$$U_n^{(1)}(\theta) = \sum_{i\in\Lambda_n} (\log \pi_i)^{(1)}(x_i \mid x^i;\theta) = \sum_{i\in\Lambda_n} \binom{1}{v_i} (X_i - \tanh(\alpha+\beta v_i)).$$

If we consider $Y_i = (a+bv_i)(X_i - \tanh(\alpha+\beta v_i))$, $(a,b)\neq 0$, then Y_i satisfies (2). More generally, for some non zero function b, we prove the CLT for the sum of conditionally centred functional

$$Y_i = b(v_i)(X_i - \tanh(\alpha+\beta v_i))$$

To verify $(N3)$, choose $C = 3\times\mathbb{Z}^2$, $W_i = \{j : \|j-i\|_\infty \leq 1\} : \mathcal{P} = \{W_i, i\in C\}$ is a CSP of S. Define, for $(V_i)^*$ the second order neighbour of i,

$$C_1 = \{i \in C \text{ s.t. } L_{k,l} = 1 \text{ if } k,l \in (V_i)^* \text{ and } \|k-l\|_1 = 1\}$$

C_1 is the subset of C of site i such the 16 pairs of nearest neighbour sites of $(V_i)^*$ are all connected. It is easy to see that

$$\lim_n \frac{|C_{1,n}|}{|\Lambda_n|} = \frac{p^{16}}{9} > 0$$

We have

$$\mathrm{Var}(\sum_{i\in\Lambda_n} Y_i) \geq \sum_{i\in C_n} E\{\mathrm{Var}(Y_i + \sum_{j\in W_{i,n}\setminus\{i\}} Y_j \mid \mathcal{F}^C)\}.$$

Look at sites $i\in C_1\cap[-(n-2),(n-2)]^2$. As $(V_i)^*\subseteq\Lambda_n$, we have

$$Y_i + \sum_{j\in W_i\setminus\{i\}} Y_j = G_i(X_i, x_{A_i}) + g_i,$$

where g_i is $\mathcal{F}^C$-constant and $A_i = i + A_0$ where $A_0 = \{j\in S : \|j\|_1 = 1 \text{ or } 2\}$.

For $i\in C_1$, $b:\mathbb{V}\to\mathbb{R}$ with $\mathbb{V} = \{-4,-2,0,2,4\}$. Suppose that b is not zero. The crucial functional step in Guyon and Künsch [13] (see Proof of Theorem 3, page 190-193) is still valid here, without any hypothesis of shift invariance, stationarity or ergodicity for the Ising model. This result is the following there exists a configuration $x^0_{A_0}$ on A_0 such that, uniformly in $i\in C_{1,n}$, $X_i\mapsto G_i(X_i, x^0_{A_i})$ is not constant, $x^0_{A_i}$ being the configuration $x^0_{A_0}$

shifted from i. Then, according to (8), $(X_i \mid x^0_{A_i})$ is not constant and there exists $\delta > 0$ such that, for every $i \in C_{1,n}$, $\mathrm{Var}(Y_i + \sum_{W_i \setminus \{i\}} Y_j \mid \mathcal{F}^C) \geq \delta > 0$. Therefore, $E(\mathrm{Var}(Y_i + \sum_{W_i \setminus \{i\}} Y_j \mid \mathcal{F}^C)) \geq \delta \times \pi(x^0_{A_i})$. On the other hand, a bound from below for $\pi(x^0_{A_i})$ is a consequence of Bayes formula:

$$\pi(x_{A_i}) = \sum_{x_{\partial A_i}} \pi(x_{A_i} \mid x_{\partial A_i}) \pi(x_{\partial A_i}) \geq \varepsilon$$

where $\varepsilon = \inf_{y_{A_i}, y_{\partial A_i}} \pi(y_{A_i} \mid y_{\partial A_i}) > 0$. Thus we have, for large n,

$$\mathrm{Var}(S_n) \geq p^{16} \times \frac{\delta \times \varepsilon}{10} \times |\Lambda_n| \, .$$

For $p = 0$, CLT is trivial because the random variables Y_i are independent and identically distributed. For $p = 1$, we are in presence of the 4-nearest neighbour. Ising model (Guyon and Künsch, [13]): CLT for Y is valid regardless of phase transition, or non stationarity, or non ergodicity of the model.

The result obtain here for the percolation graph can be generalized to more general random environment : the main property of the random graph that is need is the sub-ergodicity of the subset C_1 : $\liminf_n \frac{|C_{1n}|}{|C_n|} > 0$.

5.2 Isotropic Ising model on $S = \mathbb{Z}^2 \setminus \mathbb{T}$

Consider the non-stationary nearest neighbours isotropic Ising model on $S = \mathbb{Z}^2 \setminus \mathbb{T}$, $\mathbb{T} = (1,1) + 2 \times \mathbb{Z}^2$, $\Lambda_n = [-n, +n]^2 \cap S$, and the centred functional

$$Y_i = b(v_i)(X_i - \tanh(\alpha + \beta v_i))$$

with $v_i = \sum_{j \in \partial i} x_j$. Take the CPS as defined in example 2. If we focus on $C_1 = (3,0) + 6 \times \mathbb{Z}^2$, (N3) is fulfilled and $\lim_{n \to \infty} \frac{|C_{1,n}|}{|\Lambda_n|} = \frac{1}{27}$. For $i = (3,0)$, $Y_i + \sum_{W_i \setminus \{i\}} Y_j = G_i(X_i, x_{A_i}) + g_i$, where g_i is $\mathcal{F}^C$-constant, and $A_i = \{j \in S : \|j - i\|_1 = 1 \text{ or } 2,\ j \neq (3, \pm 1) \text{ and } j \neq (3, \pm 2)\}$. We can show that there exists x_{A_i} such that $x \mapsto G_i(x, x_{A_i})$ is not constant and we can apply same argument as before to prove CLT for Y.

5.3 Ising model on an irregular lattice

Consider S is an infinite countable set equipped with a graph $\mathcal{G}$ satisfying $(N2)$. Suppose also that there exists a CPS, $\mathcal{P} = \{W_i, i \in C\}$, with basis

(S, C) such that for every i, $V_i = \{i\} \cup \partial i \subset W_i$. Suppose that (Λ_n) is a strictly increasing sequence such that $\liminf_n \frac{|C_n|}{|\Lambda_n|} > 0$.

Let X be an Ising model on $(S, \mathcal{G})$ with conditional laws

$$\text{for } i \in S: \pi_i(x_i \mid x_{\partial i}) = \frac{\exp\{x_i(\alpha a_i + \beta v_i)\}}{2\cosh(\alpha a_i + \beta v_i)}, \tag{9}$$

where $v_i = \sum_{j\in\partial i} b_{ij} x_j$, $\theta = (\alpha, \beta)$ is a parameter and (a_i), $(b_{ij} = b_{ji})$ are known weights. The conditionally centred functional Y is

$$Y_i = b_i(x_{\partial i})(X_i - \tanh(\alpha a_i + \beta v_i)$$

and $S_n = \sum_{i\in\Lambda_n} Y_i = \sum_{i\in C_n} G_{i,n}$.

If $i \in C$ and $(V_i)^* \subset \Lambda_n$, $G_{i,n} = G_i = G_i(X_i, x_{\partial i\cup\partial^2 i}) + g_i$ where g_i is $\mathcal{F}^C$-constant, $\partial^2 i = \{k, k \neq i : \exists j \in \partial i \text{ s.t. } k \in \partial j\}$, and

$$G_i(X_i, x_{A_i}) = b_i(x_{\partial i})X_i + \sum_{j\in\partial i} b_j(X_i, w_{j,i})\{x_j - m_j(X_i, w_{j,i})\}$$

where, for $j \in \partial i$, $w_{j,i} = (x_k, k \in \partial j\backslash\{i\})$. We can verify (N3) by the following steps:

1. find $C_1 \subset C$ such that:
   (a) $\liminf_n \frac{|C_{1,n}|}{|C_n|} > 0$;
   (b) for each $i \in C_1$, $\exists x_{\partial i\cup\partial^2 i}$ s.t. $\Delta_i = G_i(+1, x_{\partial i\cup\partial^2 i}) - G_i(-1, x_{\partial i\cup\partial^2 i}) \neq 0$;
   (c) get a uniform lower bound for $|\Delta_i|$ over C_1;
2. get a uniform lower bound for $\pi_i(x_i \mid x_{\partial i})$ and $\pi_{\partial i\cup\partial^2 i}(x_{\partial i\cup\partial^2 i})$ in $i \in C_1$, x_i, $x_{\partial i}$, $x_{\partial i\cup\partial^2 i}$.

Step 1 requires *ad hoc* strategies. Step 2 follows easily provided (a_i) and (b_{ij}) are bounded since conditional probabilities are positive and continuous in (a_i, b_{ij}), x_i, $x_{\partial i}$ and $x_{\partial i\cup\partial^2 i}$.

Now we consider $Y_i = \binom{a_i}{v_i}(X_i - \tanh(\alpha a_i + \beta v_i))$. To prove CLT for (Y_i), we need only to consider linear real functions $b_i(v_i) = a a_i + b v_i$, $(a, b) \neq 0$. Setting, for $j \in \partial i$, $v_{j,i} = v_j - b_{i,j} x_i = \sum_{k\in\partial j\backslash\{i\}} b_{k,j} x_k$, we can write

$$\begin{aligned} \Delta_i &= g_i(x_{\partial i}; a, b) + h_i(x_{\partial^2 i}, a, b; \theta), \text{ with } g_i(x_{\partial i}) = 2b_i(v_i) \text{ and} \\ h_i(x_{\partial^2 i}) &= -\sum_{\xi\in\{-1,+1\}} \sum_{j\in\partial i} \xi\{\tanh(\alpha a_j + \xi\beta b_{i,j} + \beta v_{j,i})\}\{a a_i + b(\xi b_{i,j} + v_{j,i})\} \end{aligned}$$

To simplify, look at the particular case where $\partial i \cap \partial^2 i = \emptyset$ for any $i \in C_1$.

1. If $b \neq 0$, for any $x_{\partial^2 i}$, Δ_i takes two values $\Delta_i' \neq \Delta_i''$ provided there exists two configurations $x'_{\partial i}$ and $x''_{\partial i}$ such that $v_i' \neq v_i''$. As $\max\{|\Delta_i'|, |\Delta_i''|\} \geq \frac{1}{2}|b|\,|v_i' - v_i''|$, we have a lower bound for $|\Delta_i|$ if we can obtain a lower bound for $|v_i' - v_i''|$.

2. If $b = 0$, then $\Delta_i = a a_i D_i(x_{\partial^2 i})$ where $D_i(x_{\partial^2 i}) = (2 - \sum_{j \in \partial i} \tanh(\alpha a_j - \beta b_{i,j} + \beta v_{j,i}) - \tanh(\alpha a_j + \beta b_{i,j} + \beta v_{j,i})$. Then (N3) holds under the conditions :

$$\inf_{i \in C_1^*} |a_i| > 0 \text{ and } \inf_{i \in C_1^*} \sup_{x_{\partial^2 i}} |D_i(x_{\partial^2 i})| > 0.$$

An irregular lattice which comes from a Poisson process on $\mathbb{R}$

Now we shall consider a particular example of irregular lattice. Suppose that S is a realisation of an homogeneous Poisson process on $\mathbb{R}$, and write $S = \{i_k, k \in \mathbb{Z}\}$ with $i_k < i_l$ if $k < l$. Consider the 2-nearest neighbour Ising model on S, with weight $a_i \equiv 0$ and $b_{i,j} = f(|i-j|)$ for $f : (\mathbb{R}^+)^* \longrightarrow (\mathbb{R}^+)^*$ decreasing.

For $C = \{i_{3k}, k \in \mathbb{Z}\}$ and $W_{i_k} = \{i_{k-1}, i_k, i_{k+1}\}$, the partition $\mathcal{P} = \{W_i, i \in C\}$ is a CSP. Define

$$C_1 = \{i_l \in C : \inf\{|i_l - i_{l-1}|, |i_l - i_{l+1}|\} \leq 1\}$$

and $\Lambda_n = [-n, +n] \cap S$. It is easy to verify that $\lim_n \frac{|C_{1,n}|}{|\Lambda_n|} > 0$.

On the other hand, because $v_{i_l} = b_{i_l,i_{l-1}} x_{i_{l-1}} + b_{i_l,i_{l+1}} x_{i_{l+1}}$, the range of variation of v_{i_l} is $e_{i_l} = 2(b_{i_l,i_{l-1}} + b_{i_l,i_{l+1}})$ and $\inf_{i \in C_1} e_i > 0$. Thus (N3) is satisfied.

Note that we can weaken the hypothesis about the Poisson process, including dependence and/or inhomogeneity for other point processes.

5.4 Gaussian MRF on a irregular lattice

Let X be a Gaussian MRF on $(S, \mathcal{G})$. The conditional law in each site is

$$X_i | x_{\partial i} \sim \mathcal{N}(\alpha v_i, w_i) \tag{10}$$

where $w_i > 0$, $v_i = \sum_{j \in \partial i} b_{ij} x_j$ (a_i, b_{ij}) are known weights, α is an unknown parameter.

We set $a_i = w_i^{-1}$. The conditional specification (10) is coherent if for any $i, j \in S$, $i \neq j$, $a_i b_{ij} = a_j b_{ji}$ and for any finite subset $\Lambda \subset S$, the symmetric matrix $J_\Lambda = (J_\Lambda(i,j))_{i,j\in\Lambda}$, where $J_\Lambda(i,i) = a_i$, $J_\Lambda(i,j) = -\alpha a_i b_{ij}$, $i, j \in \Lambda$, $i \neq j$, is positive definite.

If X is observed on a increasing sequence (Λ_n) of S, asymptotic normality of MPLE $\widehat{\alpha}_n$ can be proved under the contraction condition (Guyon, [11, Section 4.3]) :

$$|\alpha| \, \{\sup_i \sum_{j\in\partial i} |b_{ij}|\} < 1 \tag{11}$$

This condition entails that there is not phase transition and X is α-mixing (Doukhan, [9]). The last property allow us to establish the asymptotic normality.

We prove asymptotic normality of MPLE $\widehat{\alpha}_n$ by means of proposition 1 without condition (11). The verification of condition (N3) still requires careful examination. The conditionally centred functional Y is

$$Y_i = (\log \pi(x_i \mid x_{\partial i})^{(1)}_\alpha = a_i v_i (X_i - \alpha v_i).$$

Let $\mathcal{P} = \{W_i, i \in C\}$ be a PCS with $C \subseteq S$ such that for any i, $V_i = \{i\} \cup \partial i \subseteq W_i$. For all $(V_i)^* \subseteq \Lambda_n$ we have

$$G_{i,n} = G_i = a_i v_i (X_i - \alpha v_i) + \sum_{j\in\partial i} a_j v_j (x_j - \alpha v_j) + g_i$$

where g_i is $\mathcal{F}^C(X)$-constant. For $j \in \partial i$, we denote $v_j = b_{ji} X_i + v_{ji}$ with $v_{ji} = \sum_{k\in\partial j, k\neq i} b_{jk} x_k$. We can also write G_i as:

$$G_i = c_i X_i + d_{ij} X_i^2 + e_{ij}$$

where $c_i = \sum_{j\in\partial i} a_j b_{ij} (x_j - 2\alpha v_{ji})$ and $d_{ij} = -\alpha\{\sum_{j\in\partial i} a_j b_{ij}^2\}$. Note that c_i, d_{ij} and e_{ij} are $\mathcal{F}^C(X)$-constant.

A lower bound for $Var(G_i \mid \mathcal{F}^C(X))$ is determined by noting that if $Z \sim \mathcal{N}(\mu, \sigma^2)$, and $G = cZ + dZ^2$ then

$$Var(G) = (c + 2d\mu)^2 \sigma^2 + 2d^2\sigma^4 \geq 2d^2\sigma^4.$$

We obtain

$$Var(G_i \mid \mathcal{F}^C(X)) \geq 2\alpha^2 \{\sum_{j\in\partial i} a_j b_{ij}^2\}^2 \times a_i^{-2} = 2\alpha^2 \{\sum_{j\in\partial i} b_{ij} \times b_{ji}\}^2.$$

The lower bound does not depend on $x_{\partial i}$, thus a sufficient condition for (N3) is

$$\alpha \neq 0 \text{ and } \liminf_n \frac{\sum_{i \in C_n} \{\sum_{j \in \partial i} b_{ij} \times b_{ji}\}^2}{|\Lambda_n|} > 0.$$

References

[1] Besag, J., 1974. Spatial interaction and the statistical analysis of lattice systems. Journal of the Royal Statistical Society, series B 36, 192–236.

[2] Bolthausen, E., 1982. On the central limit theorem for stationary mixing random fields. Annals of Probability 10, 1047–1050.

[3] Breiman, L., 1992. Probability. SIAM.

[4] Cliff, A. D., Ord, J. K., 1981. Spatial Processes: Models and Applications. Pion Ltd.

[5] Comets, F., 1992. On consistency of a class of estimators for exponential families of Markov random fields on the lattice. The Annals of Statistics 20, 455–468.

[6] Comets, F., Janzura, M., 1998. A central limit theorem for conditionally centered random fields with an application to markov fields. Journal of Applied Probability 35, 608–621.

[7] Cressie, N., 1991. Statistics for Spatial Data. Wiley, New York.

[8] Dedecker, J., 1998. A central limit theorem for stationary random fields. Proba. Theory Relat. Fields 110, 397–437.

[9] Doukhan, P., 1994. Mixing: Properties and Examples. Lecture Notes in Statistics 85, Springer, New York.

[10] Georgii, H.-O., 1988. Gibbs Measures and Phase Transitions. Walter de Gruyter and Company.

[11] Guyon, X., 1995. Random Fields on a Network. Springer, New York.

[12] Guyon, X., Hardouin, C., 2001. Markov field dynamics : models and statistics. Statistics 35, 593–627.

[13] Guyon, X., Künsch, H. R., 1992. Asymptotic comparison of estimators in the Ising model. In: Lecture Notes in Statistics. Vol. 74. Springer, Berlin, pp. 177–198.

[14] Haining, R., 1990. Spatial Data Analysis in the Social and Environmental Sciences. Cambridge University Press, Cambridge.

[15] Jensen, J., Künsch, H. R, 1994. On asymptotic normality of pseudo-likelihood estimation for pairwise interaction processes. Annals of the Institute of Mathematical Statistics 46, 475–486.

[16] Stein, C., 1973. A bound for the error of normal approximation of a sum of dependent random variables. In: Proc. Sixth Berkeley Symp. Math. Statist. Prob. Vol. 2. pp. 583–602.

[17] Tiefelsdorf, M., 2000. Modelling Spatial Processes. Lecture Notes in Earth Sciences Vol. 87. Springer, Berlin.