

# Estimation spatio-temporelle d'un système de particules

Xavier Guyon<sup>(\*)</sup> et Besnik Pumo<sup>(+)</sup>

<sup>(\*)</sup>SAMOS, Université Paris 1 et <sup>(+)</sup>INH, Angers

January 22, 2004

## Abstract

Soit  $X$  un processus de contact à temps discret sur  $\mathbb{Z}^2$  tel qu'il est défini par Durrett et Levin. On s'intéresse à l'estimation des paramètres de ce modèle lorsqu'on dispose d'une observation spatio-temporelle du processus, à savoir  $T$  observations successives de  $X$  sur un sous-ensemble fini et fixé de sites  $S$ . L'asymptotique de cette estimation est étudiée dans le temps et en situation de non-extinction de  $X$ . La vraisemblance étant numériquement incalculable, on propose d'utiliser une pseudo-vraisemblance marginale (PVM) facile à calculer et qui capte bien l'information que l'on a sur le modèle. On montre alors que l'estimation du maximum de PVM est convergente et asymptotiquement gaussienne. Des études expérimentales confirment ces résultats.

*Mots clés* : processus de contact, pseudo-vraisemblance marginale, processus super-critique, sous-ergodicité, estimateur convergent, normalité asymptotique, estimation spatio-temporelle.

*Code AMS* : 62M30, 62M05, 62F12

## 1 Introduction et description du modèle

Considérons, à titre d'illustration, le problème de la modélisation et de l'estimation de la dynamique d'une espèce végétale évoluant sur le lattice régulier  $\mathbb{Z}^2$ . Dépendant de paramètres biologiques, cette évolution est déterminée en spécifiant, pour chaque site  $s \in \mathbb{Z}^2$ , la probabilité conditionnelle que le site  $s$  soit dans l'état  $X_{t+1}(s) = y \in \{0, 1\}$  au temps  $t + 1$  étant donné l'état  $X_t = x \in \Omega = \{0, 1\}^{\mathbb{Z}^2}$  au temps  $t$ . L'état 0 (resp. 1) code l'absence (resp. la présence) d'une plante au site  $s$ . L'objectif est d'estimer les paramètres du modèle lorsqu'on dispose d'une *observation spatio-temporelle* du processus aux instants  $t = 0, \dots, T$  sur un sous-ensemble fini et fixé  $S$  de sites de  $\mathbb{Z}^2$ . Nous nous intéresserons à l'*asymptotique temporelle* de cette estimation ( $T \rightarrow \infty$ ) lorsqu'il n'y a pas extinction de l'espèce. Fiocco (1997) et Fiocco et Zwet (2002) étudient un autre contexte statistique, celui de l'observation complète d'une seule réalisation  $x_t$  pour  $t$  grand.

### 1.1 Le modèle de processus de contact

La règle d'évolution est celle d'un système de particules particulier, un *Processus de Contact* (PC), processus à temps discret markovien. On supposera que la transition en un site  $s$  est stationnaire dans l'espace et dans le temps, dépendant localement de la configuration passée  $x$  par l'intermédiaire de  $x(\mathcal{N}_1(s))$  où  $\mathcal{N}_1(s)$  est le voisinage d'ordre 1 de  $s$ ,

le voisinage d'ordre  $d \geq 1$  de  $s$  étant défini par :

$$\mathcal{N}_d(s) = \{t \in \mathbb{Z}^2 : \|s - t\|_1 \leq d\}$$

La règle d'évolution est la suivante (Durrett and Levin, 1994) :

- a. Chaque plante en vie à l'instant  $t$  meurt au temps  $t + 1$  avec la probabilité  $\gamma$ ,
- b. Si une plante en  $s$  survit, alors elle donne naissance à une plante en  $u \in \partial s$  avec une probabilité  $\lambda$ , ces reproductions étant indépendantes pour les 4 sites voisins  $u$  de  $s$ ,
- c. Si une ou plusieurs plantes sont dispersées en  $s$ , ou si la plante en  $s$  survit en  $t + 1$ , alors  $X_{t+1}(s) = 1$ ; sinon,  $X_{t+1}(s) = 0$ .

De plus, on suppose qu'il y a indépendance temporelle entre tous les événements définis en (a) et (b).

Ce modèle dépend du paramètre  $\theta = (\gamma, \lambda)$  et nous supposons que  $\theta \in ]0, 1[^2$ . D'autres règles d'évolution peuvent être envisagées sur le lattice  $\mathbb{Z}^d$  (cf. Durrett et Levin, 1994 ; Durrett, 1995 ; Griffeath, 1981 ; Harris, 1974 ; Pumo et Le Corff, 2001). Signalons que les techniques utilisées dans cet article peuvent s'étendre au cas où il n'y a pas stationnarité dans l'espace et/ou dans le temps.

## 1.2 Processus supercritique

L'état  $\mathbf{0}$  partout sur  $\mathbb{Z}^2$  est un état absorbant du processus. Aussi, afin de donner un sens à l'étude asymptotique, on imposera une condition **(I)** de non extinction sur  $S$ , à savoir qu'il y a une infinité d'instants tels que au moins une plante soit présente sur  $S$ .

Cette condition est à rapprochée de la notion de processus supercritique. Le processus de contact est dit *supercritique* (Durrett, 1995) si, lorsque seule l'origine est occupée à l'instant initial et si on définit le temps d'extinction de l'espèce par

$$\tau = \inf\{t : X_t(s) = 0 \text{ pour tout } s \in \mathbb{Z}^2\}$$

alors  $P(\tau = +\infty) > 0$  : il y a non extinction avec une probabilité  $> 0$ .

On sait alors (Durrett, 1995) que pour chaque taux de reproduction  $\lambda > 0$ , il existe une valeur critique  $\gamma_c(\lambda)$  du taux de décès telle que  $P(\tau = \infty) = 0$  si  $\gamma > \gamma_c(\lambda)$  et  $P(\tau = \infty) > 0$  si  $\gamma < \gamma_c(\lambda)$ . Ainsi, si  $\gamma > \gamma_c(\lambda)$  l'espèce disparaîtra *p.s.* alors que si  $\gamma < \gamma_c(\lambda)$ , l'espèce survivra indéfiniment avec une probabilité  $> 0$ .

## 1.3 Simulation du PC

Soient  $x, y \in \Omega = \{0, 1\}^{\mathbb{Z}^2}$  deux configurations sur  $\mathbb{Z}^2$ ,  $S$  un sous-ensemble fini de  $\mathbb{Z}^2$ . On supposera que  $X$  est observé aux instants  $t = 0, 1, \dots, T$  sur  $S_1$ , le 1-voisinage de  $S$ ,

$$S_1 = \{u \in \mathbb{Z}^2 : \exists v \in S \text{ tel que } \|u - v\|_1 \leq 1\}.$$

La loi de  $X_{t+1}(S)$  conditionnellement à  $X_t = x$  dépend uniquement de la configuration  $x(S_1)$  de  $X_t$  sur  $S_1$ , de transition  $P_S(x, y) = P(X_{t+1}(S) = y(S) \mid X_t(S_1) = x(S_1))$ .

A l'inverse des modèles de Gibbs (Geman et Geman, 1984; Guyon 1995), la simulation d'un système de particules ne nécessite pas de recourir à un algorithme itératif du type MCMC. Pour le modèle de PC, un seul pas d'algorithme suivant (a), (b) et (c) simule exactement la règle d'évolution. Décrivons cet algorithme.

Soient  $\{V_s, s \in S\}$  des variables i.i.d. de Bernoulli de paramètre  $(1 - \gamma)$  modélisant les survies sur  $S$ ,  $\{N_{(s',s)}, s \in S \text{ et } s' \in \partial s\}$  des variables i.i.d. de Bernoulli de paramètre  $\lambda$  modélisant les naissances en  $s \in S$  issues de  $s'$ , les variables  $V$  étant indépendantes des variables  $N$ . Notons  $z(s) \in \{0, 1, 2, 3, 4, 5\}$  le nombre de plantes présentes au site  $s \in S$  après les étapes (a) et (b). Pour une configuration initiale  $x$ , l'état  $z(s)$  au site  $s \in S$  après une itération est réalisé par :

$$z(s) = x(s) \times V_s + \sum_{u \in \partial s} x(u) \times V_u \times N_{(u,s)} \quad (1)$$

La simulation  $x \mapsto y$  sur  $S$  est alors réalisée par :

$$\forall s \in S : y(s) = \mathbf{1}(z(s) \geq 1)$$

## 2 Pseudo-vraisemblance marginale

### 2.1 La transition $P_S(x, y)$ est incalculable

Si la simulation d'un PC ne pose pas de difficulté, le calcul numérique de la transition  $P_S(x, y)$  est impossible dès que  $S$  n'est pas réduit à quelques points. Il y existe deux raisons à cela :

1. d'un part, il faut examiner chaque configuration particulière  $x(S_1)$  pour expliciter la loi de  $z(S)$ , ce qui est de complexité  $2^{\#(S)}$ ;
2. d'autre part, le fait que seule  $z(s) = 0$  ou  $z(s) \geq 1$  est observable entraîne une nouvelle complexité de l'ordre de  $2^{\#(S)}$ . En effet, pour  $y$  une configuration sur  $S$ , si  $S_0(y) = \{s \in S : y(s) = 0\}$ ,  $S_1(y) = \{s \in S : y(s) = 1\}$ , et  $E = \{1, 2, 3, 4, 5\}$ , alors

$$P_S(x, y) = \sum_{w \in E^{S_1(y)}} P_S(x; Z(S_0(y)) = 0, Z(S_1(y)) = w)$$

### 2.2 Pseudo-transition marginale $M_S(x, y)$

Une façon de contourner cette difficulté consiste à remplacer la transition jointe sur  $S$ ,  $P_S(x, y)$ , par une *pseudo-transition marginale*  $M_S(x, y)$  facile à calculer et qui capte bien l'information que l'on a sur le modèle. Le produit de ces pseudo-transitions aux instants successifs d'observation définira la *pseudo-vraisemblance marginale* (PVM) qui servira à estimer  $\theta = (\gamma, \lambda)$ . Cette idée a été proposée par Besag afin d'estimer les paramètres d'un champ de Markov à partir d'une pseudo-vraisemblance conditionnelle, produit des densités conditionnelles aux sites d'observation (Besag, 1974). Définissons la PVM du PC.

Soient  $x$  et  $y$  deux configurations du PC. Pour  $s$  un site de  $S$  et  $A$  un sous-ensemble fini de  $\mathbb{Z}^2$ , on définit :

$$m(x, A) = \sum_{s \in A} x(s), \text{ et } v(x, s) = m(x, \partial s) = \sum_{u \in \partial s} x(u).$$

$m(x, A)$  est le nombre de sites de  $A$  occupés par  $x$  ;  $v(x, s)$  est le nombre de sites voisins de  $s$  occupés par  $x$ . Le modèle étant isotropique, l'évolution de  $x$  vers  $y(s)$  ne dépend de  $x$  que par l'intermédiaire de  $x(s)$  et de  $v(x, s)$ , ou encore du résumé  $c(x, s)$  de la configuration de 1-voisinage de  $s$  :

$$c(x, s) = (x(s), v(x, s)) \in \mathcal{C}_1 = \{0, 1\} \times \{0, 1, 2, 3, 4\}. \quad (2)$$

Notons que pour caractériser cette évolution, il est équivalent de conditionner par  $c(x, s)$  ou par  $x(\mathcal{N}_1(s))$ . Plus précisément, pour  $s \in S$ , la loi de  $X_{t+1}(s)$  conditionnelle à la configuration  $x$  à l'instant  $t$  suit une loi de Bernoulli de paramètre  $1 - p(x, s; \theta)$ , où :

$$p(x, s; \theta) = P_\theta(X_{t+1}(s) = 0 \mid x) = \gamma^{x(s)} \{\gamma + (1 - \gamma)(1 - \lambda)\}^{v(x, s)}. \quad (3)$$

Puisque  $X_{t+1}(s) = 0$  si  $c(x, s) = (0, 0)$ , les seuls sites pour lesquels  $x$  est informatif dans la transition  $t \mapsto t + 1$  sont ceux de  $I(x, S)$

$$\begin{aligned} I(x, S) &= \{s \in S, \text{ t.q. } c(x, s) \neq (0, 0)\} \\ &= \{s \in S, \text{ t.q. } x(\mathcal{N}_1(s)) \neq 0\} \end{aligned}$$

On définit alors la *pseudo-transition marginale* comme le produit sur  $I(x, S)$  des transitions marginales,

$$M_S(x, y; \theta) = \prod_{s \in I(x, S)} P_\theta(X_{t+1}(s) = y(s) \mid x) = \prod_{s \in I(x, S)} p(x, s; \theta)^{1-y(s)} (1 - p(x, s; \theta))^{y(s)} \quad (4)$$

avec la convention  $M(0, 0; \theta) = 1$  si  $I(x, S) = \emptyset$ . Notons :

$$\begin{aligned} \delta &= \gamma + (1 - \gamma)(1 - \lambda) = 1 - \lambda + \gamma\lambda \\ \eta &= \gamma + (1 - \gamma)(1 - \lambda)^2 \end{aligned}$$

$\gamma$  contrôle la non-survie au site  $s$  (liée à  $x(s)$ ),  $\delta$  la non-prolifération issue d'un site voisin de  $s$  (liée à  $v(x, s)$ ). La propriété suivante précise l'indépendance ou non des variables conditionnelles en deux sites distincts.

**Theorem 1** *Soient  $s$  et  $s'$  deux sites de  $S$  et  $x$  une configuration de  $X_t$ . Trois cas se présentent :*

(1) *Si  $s' \in \partial s$  alors :*

$$P(X_{t+1}(\{s, s'\}) = (0, 0) \mid x) = \gamma^{m(x, \{s, s'\})} \delta^{m(x, \partial\{s, s'\})}$$

(2) *Si  $s' \in \partial(\partial s) \setminus \{s\}$ , on a , notant  $A \Delta B = (A \setminus B) \cup (B \setminus A)$  :*

$$P(X_{t+1}(\{s, s'\}) = (0, 0) \mid x) = \gamma^{m(x, \{s, s'\})} \delta^{m(x, \partial s \Delta \partial s')} \eta^{m(x, \partial s \cap \partial s')}$$

(3) *Si  $s$  et  $s'$  sont non-voisins et si  $\partial s \cap \partial s' = \emptyset$ , les variables conditionnelles  $(X_{t+1}(s) \mid x)$  et  $(X_{t+1}(s') \mid x)$  sont indépendantes.*

*Preuve :*

(1) Dans ce cas,  $\partial s \cap \partial s' = \phi$  ; le résultat découle directement de la définition de  $\gamma$  et de  $\delta$ .

(2) Soit maintenant  $u \in \partial s \cap \partial s'$  un site tel que  $x(u) = 1$ . La probabilité qu'il n'y ait pas de naissance en  $s$  et en  $s'$  issue du site  $u$  est égale à  $\eta$ . On en déduit (2).

(3) Enfin, (1) montre que sous les hypothèses faites sur  $s$  et  $s'$ , les variables qui entrent dans la génération de  $X_{t+1}(s)$  et de  $X_{t+1}(s')$  sont, à  $x$  donné, indépendantes.

D'où le résultat annoncé. ..... ■

En particulier, si  $s$  ou  $s'$  sont deux sites voisins, ou si  $s \neq s'$  et  $\partial s \cap \partial s' \neq \emptyset$ ,  $X_{t+1}(s)$  et  $X_{t+1}(s')$  sont, conditionnellement à  $x_t$ , dépendantes, de covariance :

$$\begin{aligned} \text{Cov}(X_{t+1}(s), X_{t+1}(s') \mid x_t) &= \text{Cov}(1 - X_{t+1}(s), 1 - X_{t+1}(s') \mid x_t) \\ &= P(X_{t+1}(\{s, s'\}) = (0, 0) \mid x_t) - P(X_{t+1}(s) = 0 \mid x_t)P(X_{t+1}(s') = 0 \mid x_t) \\ &= p(x_t, s; \theta) p(x_t, s'; \theta) [b(x_t, s, s'; \theta) - 1] \end{aligned}$$

où

$$b(x_t, s, s'; \theta) = \begin{cases} \delta^{-x_t(s) - x_t(s')} & \text{si } s' \in \mathcal{N}_1(s) \setminus s \\ \delta^{-2m(x_t, \{s, s'\})} \eta^{m(x_t, \{s, s'\})} & \text{si } s' \in \mathcal{N}_2(s) \setminus \mathcal{N}_1(s) \\ 1 & \text{si } s' \in S \setminus \mathcal{N}_2(s). \end{cases} \quad (5)$$

### 2.3 Pseudo-vraisemblance marginale

Notons  $\mathbf{x}(T) = (x_0, x_1, \dots, x_T)$  la suite des  $(T+1)$  configurations initiales de  $X$ . L'observation de  $\mathbf{x}(T)$  sur  $S_1$  permet de calculer chaque pseudo-transition (4) pour  $t = 0, 1, \dots, T-1$ . Définissons  $n(x) = \sharp(I(x, S))$  le nombre de sites de  $S$  où  $x$  est informatif,  $n(\mathbf{x}(T)) = \sum_{t=0}^{T-1} n(x_t)$  le nombre total de ces sites pour la suite d'observations  $\mathbf{x}(T)$ .

Pour  $c \in \mathcal{C}_1$ , notons  $n(x, c)$  le nombre de sites  $s \in S$  de configuration  $c(x, s)$  égale à  $c$ ,

$$n(x, c) := \sharp\{s \in S : c(x, s) = c\}, \text{ et } n(\mathbf{x}(T), c) = \sum_{t=0}^{T-1} n(x_t, c) \quad (6)$$

On a :

$$n(\mathbf{x}(T)) = \sum_{c \neq (0,0)} n(\mathbf{x}(T), c). \quad (7)$$

La log-pseudo-vraisemblance marginale renormalisée de  $\mathbf{x}(T)$  observée sur  $S_1$  est définie par :

$$l_{T,S}(\mathbf{x}(T); \theta) = \frac{1}{n(\mathbf{x}(T))} \sum_{t=0}^{T-1} \log M_S(x_t, x_{t+1}; \theta).$$

$S$  étant fixé, on oubliera le plus souvent l'indexation par  $S$ , notant  $l_T(\mathbf{x}(T); \theta)$ ,  $I(x)$  ou  $M(x, y; \theta)$  pour  $l_{T,S}(\mathbf{x}(T); \theta)$ ,  $I(x, S)$  ou  $M_S(x, y; \theta)$ . De même, quand on travaillera conditionnellement à une suite  $\mathbf{x}(T)$  fixée, on notera  $n(T)$  pour  $n(\mathbf{x}(T))$ ,  $n(T, c)$  pour  $n(\mathbf{x}(T), c)$ ,  $l_T(\theta)$  pour  $l_T(\mathbf{x}(T); \theta)$ .

La pseudo-vraisemblance marginale (PVM) s'écrit alors, avec  $p(x, s; \theta) = \gamma^{x(s)} \delta^{v(x,s)}$

$$l_T(\theta) = \frac{1}{n(T)} \sum_{t=0}^{T-1} \sum_{s \in I(x_t)} \{[1 - x_{t+1}(s)] \log p(x_t, s; \theta) + x_{t+1}(s) \log [1 - p(x_t, s; \theta)]\}. \quad (8)$$

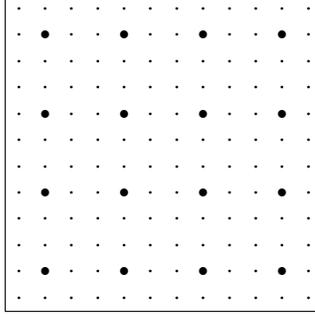


Figure 1: Ensemble  $\mathbf{K}$  de codage fort

L'estimateur du maximum de PVM de  $\theta$  (EMPV) pour l'observation  $\mathbf{x}(T)$  sur  $S_1$  est une valeur qui maximise cette PVM,

$$\hat{\theta}_T = \arg \max_{\theta} l_T(\theta). \quad (9)$$

## 2.4 Pseudo-vraisemblance marginale de codage

Un sous-ensemble  $\mathbf{K} \subset \mathbb{Z}^2$  est un *sous-ensemble de codage fort* si deux sites  $s \neq s'$  de  $\mathbf{K}$  ne sont pas voisins et si de plus leurs voisinages ne se coupent pas :  $\partial s \cap \partial s' = \emptyset$ . Un exemple est  $\mathbf{K} = 3 \times \mathbb{Z}^2$  (cf. Figure 1) : il est facile de voir que l'on peut partitionner  $\mathbb{Z}^2$  avec 9 tels ensembles de codage fort.

La propriété importante qui nous sera utile pour un ensemble de codage fort est que, conditionnellement à toute configuration  $x$ , les variables aléatoires  $\{(X_{t+1}(s) \mid X_t = x), s \in \mathbf{K}\}$  sont indépendantes.

Définissons enfin la pseudo-vraisemblance de codage. Si  $L$  est un sous-ensemble de  $\mathbb{Z}^2$ , on note  $I^L(x, S) = I(x, S) \cap L$ ,  $n^L(x) = \sharp(I^L(x, S))$  et  $n^L(T) = \sum_{t=0}^{T-1} n^L(x_t)$  : la *log-pseudo-vraisemblance du  $\mathbf{K}$ -codage* sur  $S$  est alors définie par :

$$l_T^{\mathbf{K}}(\theta) = \frac{1}{n^{\mathbf{K}}(T)} \sum_{t=0}^{T-1} \sum_{s \in I^{\mathbf{K}}(x_t)} \{[1 - x_{t+1}(s)] \log p(x_t, s; \theta) + x_{t+1}(s) \log[1 - p(x_t, s; \theta)]\}. \quad (10)$$

## 3 La PVM permet d'identifier le paramètre $\theta$

Dans toute la suite de cet article, *on supposera réalisée la condition suivante de non-extinction de  $X$  sur  $S$* , à savoir, l'observation  $\mathbf{x} = (x_t, t \geq 0)$  sur  $S$  apporte une information infinie :

$$(\mathbf{I}) : I_{\infty} = \{\mathbf{x} = (x_t, t \geq 0) \text{ tel que } n(\mathbf{x}(T)) \longrightarrow \infty \text{ si } T \longrightarrow \infty\} \quad (11)$$

Nous allons définir deux quantités, *lim inf* de fréquences empiriques : la première,  $\pi_c$ , sera utile pour démontrer que la PVM rend  $\theta$  identifiable ; la seconde,  $\pi_{c,0}^{\mathcal{L}}$ , permettra de montrer qu'une matrice de pseudo-information est définie positive. Pour chacune de ces quantités, on va établir une condition de sous-ergodicité.

Pour  $c \in \mathcal{C}_1$ ,  $\pi_c$  est définie par :

$$\pi_c = \underline{\lim}_{T \rightarrow \infty} \frac{n(T, c)}{n(T)}. \quad (12)$$

Définissons  $\pi_{c,0}^{\mathcal{L}}$ . Soit  $\mathbf{L} = \{L_i, i \in \mathcal{I}\}$  une partition de  $S$  (un ensemble de parties non-vides et disjointes qui recouvrent  $S$ ),  $\mathcal{L} = \{L_{i(t)}, t \geq 0\}$ , une suite infinie de parties de  $\mathbf{L}$ , et  $(c, 0)$  la configuration de  $\mathcal{N}_2(0)$  telle que  $x(\mathcal{N}_1(0)) = c$  et  $x(\mathcal{N}_2(0) \setminus \mathcal{N}_1(0)) \equiv 0$ . Soit  $n^L(x_t, c, 0)$  le nombre de sites  $s \in S \cap L$  tels que  $x_t(\mathcal{N}_2(s))$  soit de configuration  $(c, 0)$ .  $\pi_{c,0}^{\mathcal{L}}$  est défini par :

$$\pi_{c,0}^{\mathcal{L}} = \underline{\lim}_{T \rightarrow \infty} \frac{\sum_{t=0}^{T-1} n^{L_{i(t)}}(x_t, c, 0)}{n(T)}$$

### 3.1 Un résultat de sous-ergodicité

On a le résultat suivant :

**Theorem 2** Soit  $\mathcal{C}_1^*$  l'ensemble des configurations sur  $\mathcal{N}_1(0)$  telles que  $x(0) = 1$  et  $\mathbf{L} = \{L_i, i \in \mathcal{I}\}$  une partition de  $S$ . Alors il existe  $\alpha > 0$  et une suite de parties  $\mathcal{L}$  tels que, pour tout  $c \in \mathcal{C}_1^*$ , et tout  $x \in I_\infty$ , on a :

$$\pi_{c,0}^{\mathcal{L}} \geq \alpha.$$

*Preuve* : On utilisera le résultat suivant : soit  $(Y_n, n \geq 1)$  une suite infinie de variables indépendantes de Bernoulli de paramètres  $(q(k_n), n \geq 1)$ , où  $k_n \in \mathcal{K} = \{1, \dots, K\}$ , un ensemble fini. Supposons d'autre part que  $\beta = \min\{q(k), k = 1, \dots, K\} > 0$ . Alors

$$\underline{\lim}_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1, N} Y_n \geq \frac{\beta}{2}. \quad (13)$$

En effet, soit  $\{n_{k,1}, \dots, n_{k,N_k}, \dots\}$  le sous-ensemble des valeurs croissantes de  $\{1, 2, \dots\}$  tel que  $\{Y_{n_{k,l}} \sim \mathcal{B}(q(k)), l = 1, 2, \dots\}$ . L'ensemble  $\mathcal{K}_\infty = \{k \text{ tel que } N_k \rightarrow \infty\}$  étant non vide, la loi des grands nombres donne en effet :

$$\underline{\lim}_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1, N} Y_n \geq \underline{\lim}_{N \rightarrow \infty} \frac{1}{N} \sum_{k \in \mathcal{K}_\infty} N_k \cdot \frac{1}{N_k} \sum_{l=1, N_k} Y_{n_{k,l}} \geq \frac{\beta}{2}.$$

Soit  $\mathcal{C}_3^*$  l'ensemble des configurations sur  $\mathcal{N}_3(0)$  telles que  $x(0) = 1$ . On va définir deux suites d'instant. La première,

$$0 \leq T_1 < T_2 < \dots < T_{m_T} \leq T - 2$$

est la suite croissante des instants  $t$  entre 0 et  $T - 2$  telle que, pour un site  $s \in S$  au moins,  $x_t(\mathcal{N}_1(s)) \neq (0, 0)$ . On notera  $s_{T_k}$  un tel site au temps  $T_k$  :  $x_{T_k}(\mathcal{N}_1(s_{T_k})) \neq (0, 0)$ . La deuxième,

$$0 \leq T_1^* < T_2^* < \dots < T_{m_T}^* \leq T - 2$$

est la suite croissante des instants  $t$  entre 0 et  $T - 2$  qui sont informatifs, c'est à dire tels que pour un site au moins de  $S$ ,  $x_t(s) = 1$  ; on notera  $s_{T_k^*}$  un tel site au temps  $T_k^*$ . De plus, on choisira  $s_{T_k} = s_{T_k^*}$  chaque fois que  $T_k = T_k^*$ .

On définit alors  $\mathcal{L}$  de la façon suivante : (i) pour un instant  $t = T_k$  informatif,  $L_{i(T_k+1)}$  est la composante de la partition qui contient  $s_{T_k}$ ; (ii) aux autres instants,  $L_{i(t)}$  est un ensemble arbitraire de la partition.

Soit  $c$  une configuration de  $\mathcal{C}_1^*$ . Définissons :

$$Z_{k^*} = \mathbf{1}\{X_{T_{k^*}^*+1}(\mathcal{N}_2(s_{T_{k^*}^*}^*)) = (c, 0)\}, k^* = 1, m_T^*$$

$P(X_{t+1}(\mathcal{N}_2(s)) \mid x_t)$  dépendant uniquement de  $x_t(\mathcal{N}_3(s))$ , les variables  $(Z_{k^*}, k^* = 1, \dots, m_T^*)$  sont des variables indépendantes de Bernoulli de paramètre :

$$q(x_{T_{k^*}^*}(\mathcal{N}_3(s_{T_{k^*}^*}^*)), c) = P(x_{T_{k^*}^*+1}(\mathcal{N}_2(s_{T_{k^*}^*}^*)) = (c, 0) \mid x_{T_{k^*}^*}(\mathcal{N}_3(s_{T_{k^*}^*}^*))).$$

Le fait que  $x_{T_{k^*}^*}(s_{T_{k^*}^*}^*) = 1$  garantit que toute configuration  $(c, 0)$  de  $x_{T_{k^*}^*+1}(\mathcal{N}_2(s_{T_{k^*}^*}^*))$  pour  $c \in \mathcal{C}_1^*$  est atteinte avec une probabilité  $> 0$  à partir de toute configuration de  $\mathcal{C}_3^*$  :  $q(x_{T_{k^*}^*}(\mathcal{N}_3(s_{T_{k^*}^*}^*)), c) > 0$ . On a donc :

$$\beta^* = \inf\{q(x_{T_{k^*}^*}(\mathcal{N}_3(s_{T_{k^*}^*}^*)), c) : x_{T_{k^*}^*}(\mathcal{N}_3(s_{T_{k^*}^*}^*)) \in \mathcal{C}_3^*, c \in \mathcal{C}_1^*, T_{k^*}^*, \text{ pour } k \geq 1\} > 0.$$

Définissons :

$$W_k = X_{T_k+1}(s_{T_k}), k = 1, m_T.$$

$P(X_{t+1}(s) \mid x_t)$  dépendant uniquement de  $x_t(\mathcal{N}_1(s))$ , les variables  $(W_k, k = 1, m_T)$  sont des variables indépendantes de Bernoulli de paramètres,

$$q(x_{T_k}(\mathcal{N}_1(s_{T_k}))) = P(X_{T_k+1}(s_{T_k}) = 1 \mid x_{T_k}(\mathcal{N}_1(s_{T_k}))).$$

Le fait que  $x_t(\mathcal{N}_1(s_{T_k})) \neq (0, 0)$  garantit que  $X_{T_k+1}(s_{T_k}) = 1$  est atteinte avec une probabilité  $> 0$ , On a donc :

$$\beta = \inf\{q(x_{T_k}(\mathcal{N}_1(s_{T_k}))), x_{T_k}(\mathcal{N}_1(s_{T_k})) \neq (0, 0), T_k, \text{ pour } k \geq 1\} > 0.$$

D'une part,

$$\sum_{t=0}^{T-2} n^{L_{i(t+1)}}(x_{t+1}, c, 0) \geq \sum_{k=1}^{m_T^*} n^{L_{i(T_{k^*}^*+1)}}(x_{T_{k^*}^*+1}, c, 0) \geq \sum_{k=1}^{m_T^*} Z_{k^*}. \quad (14)$$

En effet, chaque  $Z_{k^*}$  égale à 1 contribue pour 1 à  $n^{L_{i(T_{k^*}^*+1)}}(x_{T_{k^*}^*+1}, c, 0)$ .

D'autre part,

$$\sum_{t=0}^{T-2} n^{L_{i(t+1)}}(x_{t+1}) \leq \sum_{k=1}^{m_T} n(x_{t+1}) \leq \sharp(S) \cdot \frac{m_T}{\sum_{k=1, m_T} W_k} \cdot m_T^*. \quad (15)$$

En effet, à chaque instant  $t$  il y a au plus  $\sharp(L_{i(t)}) \leq \sharp(S)$  sites informatifs dans  $S$  et ces instants informatifs sont en nombre au plus  $m_T$ . De plus  $m_T^* \geq \sum_{k=1, m_T} W_k$ .

On déduit alors de (14) et (15) que :

$$\frac{\sum_{t=0}^{T-2} n^{L_{i(t+1)}}(x_{t+1}, c, 0)}{\sum_{t=0}^{T-2} n^{L_{i(t+1)}}(x_{t+1})} \geq \frac{1}{\sharp(S)} \cdot \frac{1}{m_T^*} \sum_{k=1}^{m_T^*} Z_{k^*} \cdot \frac{1}{m_T} \sum_{k=1}^{m_T} W_k.$$

Le choix de  $\mathcal{L}$  entraîne que :

$$\frac{\sum_{t=0}^{T-2} n^{L_i(t+1)}(x_{t+1})}{n(T) - n(1)} = \frac{\sum_{t=0}^{T-2} n^{L_i(t+1)}(x_{t+1})}{\sum_{t=0}^{T-2} n(x_{t+1})} \geq \frac{1}{\#(S)}$$

En effet, chaque fois qu'un instant  $t$  est informatif,  $n^{L_i(t+1)}(x_{t+1}) \geq \frac{1}{\#(S)}n(x_{t+1})$ , et pour les autres instants,  $n^{L_i(t+1)}(x_{t+1}) = n(x_{t+1}) = 0$ . On a d'autre part :

$$\frac{\sum_{t=0}^{T-2} n^{L_i(t+1)}(x_{t+1}, c, 0)}{n(T)} \geq \frac{\sum_{t=0}^{T-2} n^{L_i(t+1)}(x_{t+1}, c, 0)}{\sum_{t=0}^{T-2} n^{L_i(t+1)}(x_{t+1})} \cdot \frac{\sum_{t=0}^{T-2} n^{L_i(t+1)}(x_{t+1})}{n(T) - n(1)} \cdot \frac{n(T) - n(1)}{n(T)}$$

Appliquant le résultat (13), on obtient :

$$\liminf_{T \rightarrow \infty} \frac{\sum_{t=0}^{T-2} n^{L_i(t+1)}(x_{t+1}, c, 0)}{n(T)} \geq \frac{\beta \cdot \beta^*}{4 \cdot \{\#(S)\}^2}.$$

Le théorème 3 est ainsi démontré avec  $\alpha = \frac{\beta \cdot \beta^*}{4 \cdot \{\#(S)\}^2}$ . ..... ■

*Remarques :*

(1)  $\pi_c \geq \alpha > 0$  puisque  $\pi_c \geq \pi_{c,0}^{\mathcal{L}}$  en prenant  $\mathbf{L} = \{S\}$  et  $\mathcal{L} = \{S, S, \dots\}$ .

(2) La technique utilisée permet de s'affranchir de l'hypothèse d'invariance dans le temps et/ou dans l'espace à la condition d'imposer des conditions de minoration sur les probabilités locales de survie et/ou de naissance. Mais une condition de non-extinction de type **(I)** devra toujours être exigée pour établir un résultat de sous-ergodicité.

### 3.2 Identifiabilité de $\theta$

Cette propriété de sous-ergodicité montre que, sous **(I)** et pour  $T$  assez grand, la PVM rend  $\theta$  identifiable. En effet :

- si l'observation  $\mathbf{x}(T)$  réalise deux configurations  $c_a$  et  $c_b$  non-colinéaires, alors  $\theta \mapsto l_T(\theta)$  est injective;
- sous **(I)**, la sous-ergodicité assure que  $c_a$  et  $c_b$  sont réalisées avec une probabilité tendant vers 1.

## 4 Convergence et loi asymptotique de l'estimateur du MPV

### 4.1 Enoncé du résultat

Pour  $f : U \rightarrow R$  une fonction réelle deux fois dérivable sur un ouvert  $U$  de  $\mathbb{R}^d$ , on note  $f^{(1)}(\theta)$  le vecteur des dérivées premières et  $f^{(2)}(\theta)$  la matrice des dérivées secondes de  $f(\theta)$ . Le résultat suivant établit la convergence et normalité asymptotique de la suite des estimateurs  $(\hat{\theta}_T)$  du maximum de PVM donnés par (8).

*Quelques notations :* notons

$$q(c; \theta) = P(X_{t+1}(s) = 0 \mid c(x_t, s) = c).$$

Rappelons que  $n(T, c)$  (resp.  $n(T)$ ) défini par (6) (resp. par (7)) est le nombre cumulé de sites de  $S$  de configuration locale  $c$  entre les instants 0 et  $T - 1$  (resp. le nombre total de sites informatifs). On définit alors les deux matrices  $2 \times 2$  :

$$\begin{aligned} A_T(\theta_o) &= \frac{1}{n(T)} \sum_{t=0}^{T-1} \sum_{s \in I(x_t)} \frac{p^{(1)}(x_t, s; \theta_o)}{p(x_t, s; \theta_o)} \frac{t[p^{(1)}(x_t, s; \theta_o)]}{1 - p(x_t, s; \theta_o)} \\ &= \sum_{c \neq (0,0)} \frac{n(T, c)}{n(T)} \times \frac{q^{(1)}(c; \theta_o)^t [q^{(1)}(c; \theta_o)]}{q(c; \theta_o) [1 - q(c; \theta_o)]} \end{aligned} \quad (16)$$

et, avec  $b(x_t, s, s'; \theta_o)$  défini par (5),

$$B_T(\theta_o) = \frac{1}{n(T)} \sum_{t=0}^{T-1} \sum_{s \in I(x_t)} \sum_{s' \in I(x_t)} [b(x_t, s, s'; \theta_o) - 1] \frac{p^{(1)}(x_t, s; \theta_o)^t [p^{(1)}(x_t, s'; \theta_o)]}{[1 - p(x_t, s; \theta_o)] [1 - p(x_t, s'; \theta_o)]} \quad (17)$$

**Theorem 3** *Supposons que la vraie valeur  $\theta_o = (\gamma_o, \lambda_o)$  appartienne à l'intérieur d'un compact  $\Theta \subset ]0, 1]^2$ . Alors, conditionnellement à la survie du processus dans  $S$  :*

(1) *L'estimateur du maximum de PVM est convergent :*

$$\lim_{T \rightarrow \infty} \hat{\theta}_T \stackrel{p.s.}{=} \theta_o.$$

(2) *Cet estimateur est asymptotiquement gaussien,*

$$\sqrt{n(T)} [A_T(\theta_o) + B_T(\theta_o)]^{-1/2} A_T(\theta_o) (\hat{\theta}_T - \theta_o) \sim \mathcal{N}_2(0, I_2). \quad (18)$$

où  $I_2$  est la matrice identité d'ordre 2.

## 4.2 Preuve de la convergence de l'estimateur

Observant  $x$  à l'instant  $t$ , on définit  $n_1(x, c)$  comme le nombre de sites  $s \in S$  occupé à l'instant  $t + 1$  et de configuration  $c(x, s) = c$  à l'instant  $t$ , et  $n_1(T, c)$  par :

$$n_1(T, c) = \sum_{t=0}^{T-1} n_1(x_t, c). \quad (19)$$

**Lemma 4** *Soient  $A \subset S$ ,  $I^A(x) = I(x) \cap A$ , et  $n^A(x) = \sharp(I^A(x))$ . Alors :*

$$Var \left[ \sum_{s \in I^A(x_t)} X_{t+1}(s) \mid x_t \right] \leq \frac{13}{4} n^A(x_t). \quad (20)$$

*Preuve :* Utilisant le théorème 1, on obtient :

$$Var \left[ \sum_{s \in I^A(x_t)} X_{t+1}(s) \mid x_t \right] = \sum_{s \in I^A(x_t)} \sum_{s' \in I^A \cap \mathcal{N}_2(s)(x_t)} Cov[X_{t+1}(s), X_{t+1}(s') \mid x_t].$$

Conditionnellement à  $x_t$ , les  $X_{t+1}(s)$  sont des variables de Bernoulli :  $Var(X_{t+1}(s) \mid x_t)$  et  $Cov(X_{t+1}(s), X_{t+1}(s') \mid x_t)$  sont donc majorées par 1/4. Le résultat découle du fait que  $\sharp(\mathcal{N}_2(s)) = 13$ . ..... ■

**Lemma 5** Fixons  $c \in \mathcal{C}_1$ . Sous la condition  $n(T, c) \rightarrow \infty$ , et pour l'effectif  $n_1(T, c)$  défini par (19), on a :

$$\frac{n_1(T, c)}{n(T, c)} \xrightarrow{ps} 1 - q(c; \theta_o) \text{ si } T \longrightarrow \infty \quad (21)$$

*Preuve* : On utilise le résultat suivant (Shirayev, p. 523; Arnaudière et Fraysse, p. 466) : soit  $(a_n)$  une suite à termes positifs,  $b_n = \sum_{k=1}^n a_k$ ; alors,

$$\sum_{n=1}^{\infty} \frac{a_n}{b_n^2} < \infty. \quad (22)$$

Pour  $I(x, c) = \{s : s \in I(x), c(x, s) = c\}$ , posons :

$$\begin{aligned} \zeta_t(c; \theta_o) &= \sum_{s \in I(x_{t-1}, c)} \{X_t(s) - [1 - q(c; \theta_o)]\} \text{ pour } 1 \leq t \leq T \\ S_T(c; \theta_o) &= \sum_{t=1}^T \zeta_t(c; \theta_o), \text{ et } \mathcal{F}_T = \sigma(X_t, t \leq T) \end{aligned} \quad (23)$$

$(S_T(c; \theta_o))$  est une  $(\mathcal{F}_T)$ -martingale. Le résultat annoncé est alors une conséquence du Théorème 2.18 de Hall et Heyde appliquée à  $X_T = \zeta_T(c; \theta_o)$ ,  $U_T = n(T, c)$ . Il suffit pour cela de montrer que

$$\sum_{T=1}^{\infty} \frac{E_{\theta_o}([\zeta_T(c; \theta_o)]^2 | x_{T-1})}{[n(T, c)]^2} < \infty,$$

Ceci est une conséquence de (22) et du lemme 4 puisque :

$$E_{\theta_o}([\zeta_T(c; \theta_o)]^2 | x_{T-1}) \leq (13/4) \cdot n(x_{T-1}, c).$$

..... ■

*Convergence de l'estimateur* : on utilise le théorème 3.4.3 (Guyon, 1995). Définissons

:

$$\begin{aligned} k_c(\theta_o, \theta) &= [1 - q(c; \theta_o)] \left\{ \log \frac{1 - q(c; \theta_o)}{1 - q(c; \theta)} - \log \frac{q(c; \theta_o)}{q(c; \theta)} \right\} + \log \frac{q(c; \theta_o)}{q(c; \theta)} \text{ et} \\ K(\theta_o, \theta) &= \sum_{c \neq (0,0)} \pi_c \times k_c(\theta_o, \theta). \end{aligned}$$

où  $\pi_c$  est défini par (12).

$K(\theta_o, \theta)$  admet un unique minimum en  $\theta = \theta_o$  : en effet,  $K(\theta_o, \theta_o) = 0$ , et  $K(\theta_o, \theta) > 0$  si  $\theta \neq \theta_o$  car  $k_c(\theta_o, \theta) > 0$  et  $\pi_c > 0$  pour  $c \in \mathcal{C}_1^*$  (Théorème 2).

Notons  $U_T(\theta) = -l_T(\theta)$  l'opposé de la log-pseudo-vraisemblance. On a encore :

$$\begin{aligned} U_T(\theta) &= -\frac{1}{n(T)} \sum_{t=0}^{T-1} \sum_{s \in I(x_t)} \left\{ X_{t+1}(s) \cdot \log \frac{1 - p(x_t, s; \theta)}{p(x_t, s; \theta)} + \log[p(x_t, s; \theta)] \right\} \\ &= -\sum_{c \neq (0,0)} \frac{n(T, c)}{n(T)} \left\{ \frac{n_1(T, c)}{n(T, c)} \cdot \log \frac{1 - q(c; \theta)}{q(c; \theta)} + \log[q(c; \theta)] \right\} \end{aligned} \quad (24)$$

Puisque  $\frac{n_1(T,c)}{n(T,c)} \xrightarrow{ps} 1 - q(c; \theta_o)$ , on obtient :

$$\underline{\lim}_{T \rightarrow \infty} [U_T(\theta) - U_T(\theta_o)] \geq K(\theta_o, \theta) \text{ p.s. sur } I_\infty.$$

Soit, pour  $k$  un entier positif et :

$$W_T(1/k) = \sup\{|U_T(\alpha) - U_T(\beta)| : \|\alpha - \beta\| < 1/k, \alpha, \beta \in \Theta\}.$$

$\Theta \subseteq ]0, 1]^2$  étant compact, il existe  $D < \infty$  telle que :

$$\sup_{c \in \mathcal{C}_1; \alpha, \beta \in \Theta} \{|\log[q(c; \alpha)] - \log[q(c; \beta)]|, |\log[1 - q(c; \alpha)] - \log[1 - q(c; \beta)]|\} \leq D\|\alpha - \beta\|.$$

On en déduit que  $\sup_T W_T(1/k) \leq 3D/k$ . Choissant  $\epsilon_k = 3D/k$ , on obtient  $P_o(W_T(1/k) \leq \epsilon_k) = 1$ . Cette condition assure la convergence p.s. de la suite des estimateurs du PMV.

■

### 4.3 Preuve de la convergence gaussienne

On utilise le Théorème 3.4.5 (Guyon, 1995). Remarquons que, dans les conditions (H1), (H2) et (H3) de ce théorème, la condition (H2) peut être remplacée par :

$$(H2') : \exists (a_n) \rightarrow \infty \text{ et une suite } (J_n) \text{ satisfaisant les conditions (1) et (2)}$$

On va vérifier que (H1), (H2') et (H3) sont satisfaites. Définissons  $J_T(\theta_o) = A_T(\theta_o) + B_T(\theta_o)$  où  $A_T(\theta_o)$  et  $B_T(\theta_o)$  sont définies par (16) et (17). On procédera en trois étapes :

1. Calcul de  $U_T^{(1)}(\theta_o)$  et  $U_T^{(2)}(\theta_o)$ ,
2. Calcul des matrices de pseudo-information  $I_T = A_T(\theta_o)$  et  $J_T = J_T(\theta_o)$  et existence de  $I = A(\theta_o) > 0$  et  $J = J(\theta_o) > 0$  telles que  $A_T(\theta_o) \geq A(\theta_o)$  et  $J_T(\theta_o) \geq J(\theta_o)$ ,
3. Théorème Central Limite pour  $\sqrt{n(T)}J_T^{-1/2}(\theta_o)U_T^{(1)}(\theta_o)$ .

#### (1) : Calcul de $U_T^{(1)}(\theta_o)$ et $U_T^{(2)}(\theta_o)$

Soient :

$$u_{t+1}(x_t, s; \theta) = \frac{X_{t+1}(s)}{1 - p(x_t, s; \theta)} - 1, \text{ et } \xi_{t+1}(\theta) = \sum_{s \in I(x_t)} u_{t+1}(x_t, s; \theta).$$

Utilisant (24), on obtient :

$$U_T^{(1)}(\theta_o) = \frac{1}{n(T)} \sum_{t=0}^{T-1} \xi_{t+1}(\theta_o), \tag{25}$$

$$U_T^{(2)}(\theta_o) = A_T'(\theta_o) + C_T(\theta_o), \text{ où :} \tag{26}$$

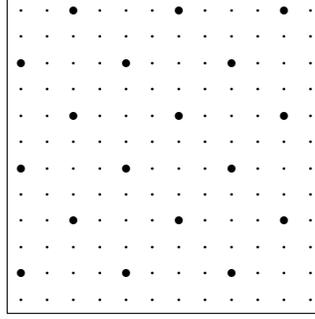


Figure 2: Ensemble  $L = \{4 \times \mathbb{Z}^2\} \cup \{4 \times \mathbb{Z}^2 + (2, 2)\}$

$$A'_T(\theta_o) = \sum_{c \neq (0,0)} \frac{n(T, c)}{n(T)} \times \frac{n_1(T, c)}{n(T, c)} \times \frac{q^{(1)}(c; \theta_o)^t q^{(1)}(c; \theta_o)}{q(c; \theta_o)[1 - q(c; \theta_o)]^2} \text{ et} \quad (27)$$

$$C_T(\theta_o) = \sum_{c \neq (0,0)} \frac{n(T, c)}{n(T)} \left\{ \frac{n_1(T, c)}{n(T, c)} \frac{1}{1 - q(c; \theta_o)} - 1 \right\} (\log q)^{(2)}(c; \theta_o). \quad (28)$$

## (2) : Calcul de $A_T(\theta_o)$ , $J_T(\theta_o)$ , et minoration

On a :

$$\text{Var} [\xi_{t+1}(\theta_o) | x_t] = \sum_{s \in I(x_t)} \sum_{s' \in I(x_t)} \text{Cov}[u_{t+1}(x_t, s; \theta_o), u_{t+1}(x_t, s'; \theta_o) | x_t].$$

Utilisant le Théorème 1, on obtient :

$$J_T(\theta_o) = \frac{1}{n(T)} \sum_{t=0}^{T-1} \text{Var}[\xi_{t+1}(\theta_o) | x_t] \quad (29)$$

Soient  $L = (\{4 \times \mathbb{Z}^2\} \cup \{4 \times \mathbb{Z}^2 + (2, 2)\}) \cap S$  (cf. Figure 2) et  $I_0^L(x_t) = \{s \in L : x_t(\mathcal{N}_2(s)) = (c, 0), c \in \mathcal{C}_1^*\}$ .  $L$  étant un sous-ensemble de codage fort, les variables conditionnelles  $\{(X_{t+1}(s) | x_{t+1}(S_1 \setminus I_0^L(x_t)), x_t), s \in I_0^L(x_t)\}$  sont indépendantes. En effet pour  $A \supseteq S_1 \setminus I_0^L(x_t)$  et  $s \notin A$  :

$$P[X_{t+1}(s) = 0 | x_{t+1}(A), x_t] = \gamma_o^{x_i(s)} \cdot (1 - \lambda_o)^{v(x_{t+1}, s)}. \quad (30)$$

On partitionne alors  $S$  avec 8 de ces ensembles, que l'on dénotera  $L_1, \dots, L_8$ . Soient  $\mathbf{L} = \{L_1, \dots, L_8\}$  et  $\mathcal{L}$  la suite associées à  $\mathbf{L}$  définie au Théorème 2. On a le résultat suivant :

**Lemma 6** Soient  $A(\theta_o)$ ,  $J(\theta_o)$  les matrices définies par :

$$A(\theta_o) = \sum_{c \neq (0,0)} \pi_c \times \frac{q^{(1)}(c; \theta_o)^t [q^{(1)}(c; \theta_o)]}{q(c; \theta_o)[1 - q(c; \theta_o)]}$$

$$J(\theta_o) = \sum_{c \neq (0,0)} \pi_{c,0}^{\mathcal{L}} \times \frac{q^{(1)}(c; \theta_o)^t [q^{(1)}(c; \theta_o)]}{[1 - q(c; \theta_o)]^2 [q(c; \theta_o)]^2}.$$

Pour  $\theta_o \in \Theta$  et sous la condition (I) :

- a)  $A(\theta_o)$  est définie positive et  $\underline{\lim}_{T \rightarrow \infty} A_T(\theta_o) = A(\theta_o)$ ,
- b)  $J(\theta_o)$  est définie positive,
- c) Il existe  $\epsilon_o > 0$  tel que  $\underline{\lim}_{T \rightarrow \infty} J_T(\theta_o) \geq \epsilon_o \times J(\theta_o)$

*Preuve :*

- a)  $A(\theta_o)$  est semi-définie positive. En effet pour tout  $\alpha = (\alpha_1, \alpha_2) \in \mathbb{R}^2$  :

$${}^t \alpha A(\theta_o) \alpha = \sum_{c \neq (0,0)} \pi_c q(c; \theta_o) [(1 - q(c; \theta_o))]^{-1} \left( {}^t \alpha (\log q)^{(1)}(c; \theta_o) \right)^2 \geq 0$$

En effet,  $\forall c \in \mathcal{C}_1^*$ ,  $\pi_c > 0$ ,  $0 < q(c; \theta_o) < 1$ . Si  ${}^t \alpha A(\theta_o) \alpha = 0$ , alors  ${}^t \alpha (\log q)^{(1)}(c_a; \theta_o) = {}^t \alpha (\log q)^{(1)}(c_b; \theta_o) = 0$  où  $c_a = (c_{a,1}, c_{a,2})$  et  $c_b = (c_{b,1}, c_{b,2})$  sont deux configurations non-colinéaires de  $\mathcal{C}_1^*$ . Puisque  $(\gamma, \lambda) \rightarrow (\gamma, \delta)$  est bijective, on obtient :

$$\begin{aligned} \alpha_1 \cdot c_{a,1}/\gamma_o + \alpha_2 \cdot c_{a,2}/\delta_o &= 0 \\ \alpha_1 \cdot c_{b,1}/\gamma_o + \alpha_2 \cdot c_{b,2}/\delta_o &= 0. \end{aligned}$$

On en déduit  $\alpha_1 = \alpha_2 = 0$ . Ainsi  $A(\theta_o)$  est définie positive.

b) On montre de façon analogue que  $J(\theta_o) > 0$  : en effet, le Théorème 2 donne  $\pi_{c,0}^{\mathcal{L}} > 0$  pour tout  $c \in \mathcal{C}_1^*$ .

c) On utilise l'idée suivante de J. Jensen (Guyon-Künsch, 1992) : pour minorer  $Var(X)$ , on conditionne  $X$  par  $\mathcal{G}$ , une  $\sigma$ -algèbre qui est à notre choix et on utilise la décomposition et minoration suivante :

$$Var(X) = E_{\mathcal{G}}(Var(X | \mathcal{G})) + Var_{\mathcal{G}}(E(X | \mathcal{G})) \geq E_{\mathcal{G}}(Var(X | \mathcal{G})). \quad (31)$$

Cette idée est utilisée dans d'autres contextes : modèle d'Ising (Guyon-Künsch, 1992), processus ponctuel (Jensen-Künsch, 1994), dynamique de processus latticiels (Guyon-Hardouin, 2001).

On applique (31) pour chaque  $t = 0, \dots, T-1$  avec  $X = (\xi_{t+1}(\theta_o) | x_t)$  et  $\mathcal{G} = \mathcal{G}_{t+1} = \sigma\{X_{t+1}(s), s \in S_1 \setminus I_0^L(x_t)\}$  où  $L = L_{i(t)}$ . Utilisant (30), on a, pour  $s \in I_0^L(x_t)$  et si  $v(X_{t+1}(s)) \geq 1$  :

$$E_{\mathcal{G}_{t+1}}[Var(X_{t+1}(s) | \mathcal{G}_{t+1}, x_t)] \geq \epsilon_o \quad (32)$$

où  $\epsilon_o = \gamma_o(1 - \gamma_o)(1 - \lambda_o)^5[1 - (1 - \lambda_o)^4]$ . En effet  $Var(X_{t+1}(s) | \mathcal{G}_{t+1}, x_t) \geq \gamma_o(1 - \lambda_o)^5$  car  $\gamma_o \cdot (1 - \lambda_o)^4 \leq P(X_{t+1}(s) = 0 | \mathcal{G}_{t+1}, x_t) \leq 1 - \lambda_o$ . (32) est alors une conséquence de la majoration  $P[v(X_{t+1}, s) = 0 | x_t] \leq \gamma_o + (1 - \gamma_o)(1 - \lambda_o)^4$  et de

$$E_{\mathcal{G}_{t+1}}[Var(X_{t+1}(s) | \mathcal{G}_{t+1}, x_t)] \geq \gamma_o(1 - \lambda_o)^5 \cdot \sum_{\nu \geq 1} P[v(X_{t+1}, s) = \nu | x_t].$$

A partir de (32), on obtient, si  $s \in I_0^L(x_t)$  :

$$E_{\mathcal{G}_{t+1}}[Var(u_{t+1}(x_t, s; \theta_o) | \mathcal{G}_{t+1}, x_t)] \geq \epsilon_o \frac{p^{(1)}(x_t, s; \theta_o) {}^t p^{(1)}(x_t, s; \theta_o)}{[1 - p(x_t, s; \theta_o)]^2 [p(x_t, s; \theta_o)]^2}.$$

Puisque les variables conditionnelles  $\{(X_{t+1}(s) \mid \mathcal{G}_{t+1}, x_t), s \in I_0^L(x_t)\}$  sont indépendantes :

$$\begin{aligned}
E_{\mathcal{G}_{t+1}} [Var (\xi_{t+1}(\theta_o) \mid \mathcal{G}_{t+1}, x_t)] &= E_{\mathcal{G}_{t+1}} [Var (\sum_{s \in S_1} u_{t+1}(x_t, s; \theta_o) \mid \mathcal{G}_{t+1}, x_t)] \\
&= E_{\mathcal{G}_{t+1}} [Var (\sum_{s \in I_0^L(x_t)} u_{t+1}(x_t, s; \theta_o) \mid \mathcal{G}_{t+1}, x_t)] \\
&= E_{\mathcal{G}_{t+1}} (\sum_{s \in I_0^L(x_t)} Var (u_{t+1}(x_t, s; \theta_o) \mid \mathcal{G}_{t+1}, x_t)) \\
&\geq \epsilon_o \sum_{s \in I_0^L(x_t)} \frac{p^{(1)}(x_t, s; \theta_o) {}^t [p^{(1)}(x_t, s; \theta_o)]}{[1 - p(x_t, s; \theta_o)]^2 [p(x_t, s; \theta_o)]^2}.
\end{aligned}$$

c) découle alors de cette minoration (29) et de l'inégalité :

$$\frac{1}{n(T)} \sum_{t=0}^{T-1} Var[\xi_{t+1}(\theta_o) \mid x_t] \geq \frac{1}{n(T)} \sum_{k=1}^{T-1} E_{\mathcal{G}_{t+1}} [Var (\xi_{t+1}(\theta_o) \mid \mathcal{G}_{t+1}, x_t)].$$

..... ■

### (3) : Le TCL pour $\sqrt{n(T)} J_T^{-1/2}(\theta_o) U_T^{(1)}(\theta_o)$

**Lemma 7** Lorsque  $T \rightarrow \infty$  :

$$\sqrt{n(T)} J_T^{-1/2}(\theta_o) U_T^{(1)}(\theta_o) \xrightarrow{Loi} \mathcal{N}_2(0, I_2)$$

*Preuve* : On va montrer que pour tout  $\alpha \in \mathbb{R}^2$  non-nul, on a :

$$\frac{1}{\sqrt{n(T)}} {}^t \alpha J_T^{-1/2}(\theta_o) \sum_{t=0}^{T-1} \xi_{t+1}(\theta_o) \sim \mathcal{N}(0, {}^t \alpha \alpha).$$

On utilise un résultat sur les suites triangulaires (Hall et Heyde, Théorème 3.2 ; Dacunha-Castelle et Duflo, Théorème 2.8.43) appliqué à  $(S_{T,t}, 1 \leq t \leq T)$ ,  $(\mathcal{F}_{T,t}, 1 \leq t \leq T)$  où :

$$S_{T,t} = \frac{1}{\sqrt{n(T)}} {}^t \alpha J_T^{-1/2}(\theta_o) \sum_{t'=0}^{t-1} \xi_{t'+1}(\theta_o), \text{ et } \mathcal{F}_{T,t} = \sigma(X_{t'}, 1 \leq t' \leq t) \equiv \mathcal{F}_t.$$

Puisque  $E[\xi_{t+1}(\theta_o) \mid \mathcal{F}_{T,t}] = E[\xi_{t+1}(\theta_o) \mid x_t] = 0$  pour  $t \leq T$ ,  $\{S_{T,t}, \mathcal{F}_{T,t}, 1 \leq t \leq T\}$  est une martingale centrée de carré intégrable. Utilisant (29), on obtient :

$$V_{T,T} = \frac{1}{n(T)} {}^t \alpha J_T^{-1/2}(\theta_o) \sum_{t=0}^{T-1} Var[\xi_{t+1}(\theta_o) \mid x_t] J_T^{-1/2}(\theta_o) \alpha = {}^t \alpha \alpha > 0.$$

Reste à vérifier la condition de Lindeberg. Soient  ${}^t \mathbf{b} = (b_1, b_2)$  où pour  $i = 1, 2$  :

$$b_i = \sup_{\theta_o \in \Theta, c \in \mathcal{C}_1} |\log q_i^{(1)}(c; \theta_o)| \cdot \frac{2 - q(c; \theta_o)}{1 - q(c; \theta_o)}$$

et  $B^2$  le maximum des modules des termes de la matrice  $\mathbf{b}^t \mathbf{b}$ . On a :

$$\begin{aligned}
& \sum_{t=0}^{T-1} E \left( \frac{1}{n(T)} [{}^t \boldsymbol{\alpha} J_T^{-1/2}(\theta_o) \xi_{t+1}(\theta_o)]^2 \cdot \mathbf{1}_{\left| \frac{1}{\sqrt{n(T)}} {}^t \boldsymbol{\alpha} J_T^{-1/2}(\theta_o) \xi_{t+1}(\theta_o) \right| > \epsilon} \mid x_t \right) \\
& \leq \frac{B^2}{n(T)} {}^t \boldsymbol{\alpha} J_T^{-1}(\theta_o) \boldsymbol{\alpha} \cdot \sum_{t=0}^{T-1} P \left( \left| {}^t \boldsymbol{\alpha} J_T^{-1/2}(\theta_o) \xi_{t+1}(\theta_o) \right| > \epsilon \sqrt{n(T)} \mid x_t \right) \\
& \leq \frac{B^2}{n(T)} {}^t \boldsymbol{\alpha} J_T^{-1}(\theta_o) \boldsymbol{\alpha} \times B^2 {}^t \boldsymbol{\alpha} J_T^{-1}(\theta_o) \boldsymbol{\alpha} \cdot \sum_{t=0}^{T-1} \frac{\text{Var}[\sum_{s \in I(x_t)} X_{t+1}(s) \mid x_t]}{\epsilon^2 n(T)} \\
& \leq \frac{B^4}{n(T)} [{}^t \boldsymbol{\alpha} J_T^{-1}(\theta_o) \boldsymbol{\alpha}]^2 \cdot \frac{13}{4\epsilon^2}.
\end{aligned}$$

Pour  $\boldsymbol{\alpha}$  fixé,  ${}^t \boldsymbol{\alpha} J_T^{-1}(\theta_o) \boldsymbol{\alpha}$  est borné en  $T$ . En effet, d'après le lemme (6),  $\underline{\lim}_T J_T(\theta_o) \geq \epsilon_o J(\theta_o) > 0$ . Ainsi si  $\mu_{J_T}(\theta_o), \mu_J(\theta_o)$  dénotent respectivement les plus petites valeurs propres de  $J_T(\theta_o)$  et  $J(\theta_o)$ ,  $\mu_{J_T}(\theta_o) \geq \epsilon_o \mu_J(\theta_o) > 0$  et il s'en suit que :

$$|{}^t \boldsymbol{\alpha} J_T^{-1}(\theta_o) \boldsymbol{\alpha}| \leq \text{sup}_\alpha |{}^t \boldsymbol{\alpha} J_T^{-1}(\theta_o) \boldsymbol{\alpha}| = [\mu_{J_T}(\theta_o)]^{-1} \leq [\epsilon_o \mu_J(\theta_o)]^{-1}.$$

En conséquence, pour tout  $\epsilon > 0$ , ce majorant tend vers 0 quand  $T \rightarrow \infty$ , et la condition de Lindeberg est bien satisfaite. ■

Vérifions les conditions (H1), (H2') et (H3). Soit  $a_T = \sqrt{n(T)}$ .

(H1) est satisfaite : en effet, les dérivées secondes de  $p(x_t, s; \theta)$  en  $\theta$  sont continues, et  $\theta_o$  est un point intérieur à  $\Theta$ , un compact de  $]0, 1[$ .

Soient :  $U_T^{(2)}(\theta_o) = A'_T(\theta_o) + C_T(\theta_o)$  où  $A'_T(\theta_o)$  et  $C_T(\theta_o)$  sont donnés par (27) et (28),  $k_{i,j}$ ,  $i, j = 1, 2$  définis par :

$$k_{i,j} = \max \left( \sup_{c \in \mathcal{C}_1, \theta_o \in \Theta} \left| \frac{\{q^{(1)} {}^t q^{(1)}(c; \theta_o)\}_{i,j}}{q(c; \theta_o) [1 - q(c; \theta_o)]^2} \right|, \sup_{c \in \mathcal{C}_1, \theta_o \in \Theta} \left| \frac{(\log q)^{(2)}_{i,j}(c; \theta_o)}{[1 - q(c; \theta_o)]} \right| \right),$$

et  $K^2 = \max_{i,j} (k_{i,j})$ . Alors, pour tout  $i, j = 1, 2$  :

$$|[A'_T(\theta_o) + C_T(\theta_o)]_{i,j}| \leq \left[ \sum_{c \neq (0,0)} n_1(T, c) / n(T) \cdot 2 + 1 \right] \cdot K^2.$$

(H2') :  $a_T = \sqrt{n(T)} \rightarrow \infty$  puisqu'il n'y a pas extinction du processus sur  $S$ . Le lemme 6 établit l'existence de  $J(\theta_o) > 0$  et le lemme 7 la normalité asymptotique  $\sqrt{a_T} J_T^{-1/2}(\theta_o) U_T^{(1)}(\theta_o) \rightarrow \mathcal{N}_2(0, I_2)$  quand  $T \rightarrow \infty$ .

(H3) : Le lemme 6 entraîne que  $\underline{\lim}_{T \rightarrow \infty} A_T(\theta_o) = A(\theta_o)$ . Enfin quand  $T \rightarrow \infty$ ,  $C_T(\theta_o) \xrightarrow{p.s.} 0$  et  $A'_T(\theta_o) - A_T(\theta_o) \xrightarrow{p.s.} 0$  puisque  $n_1(T, c) / n(T, c) \xrightarrow{p.s.} 1 - q(c; \theta_o)$ .

Le théorème 3 est démontré. ■

## 5 Etudes expérimentales

Nous présentons dans ce paragraphe quelques résultats empiriques obtenus à partir de simulations réalisées à l'aide du logiciel Splus de différents PC sur le lattice  $S = (1, \dots, 64) \times$

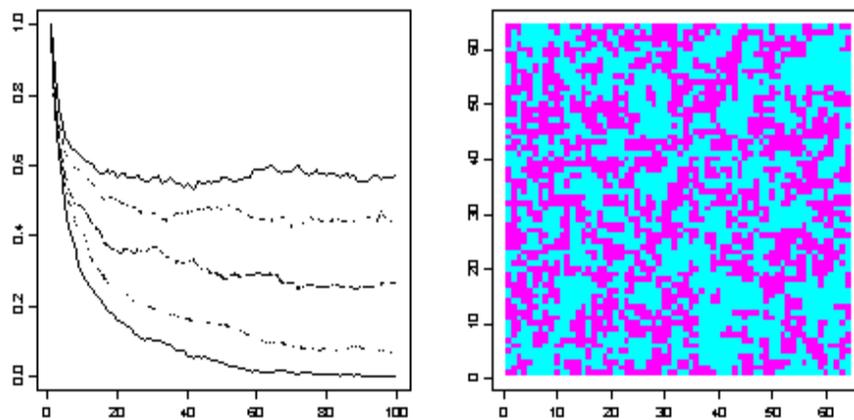


Figure 3: A gauche : Evolution de la proportion des sites occupés pour  $t = 0, 100$  pour les 5 PC de paramètres  $\gamma_o \in \{0.32, 0.35, 0.38, 0.41, 0.44\}$ ,  $\lambda_o = 0.25$ . A droite : Configuration à l'instant  $t = 100$  du PC de paramètres  $\gamma_o = 0.35$ ,  $\lambda_o = 0.25$  (les sites occupés sont en noir).

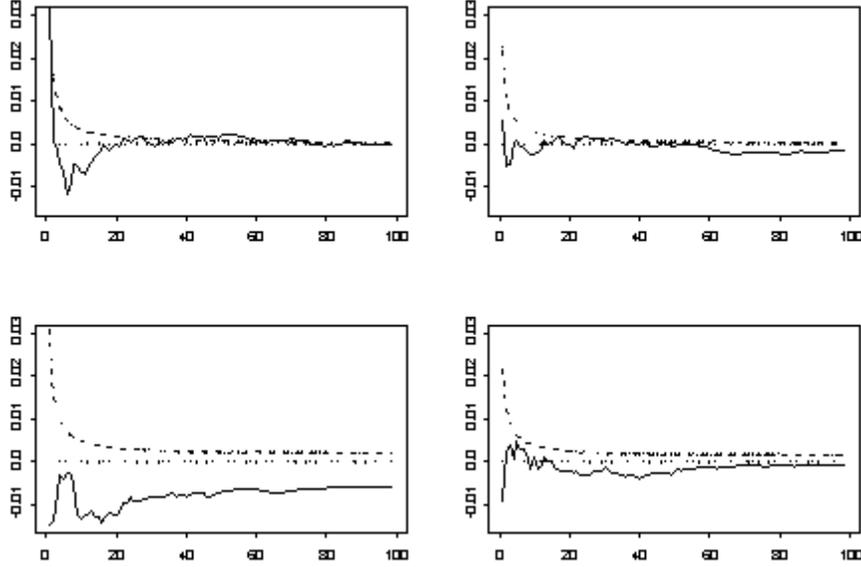


Figure 4: A gauche, cas supercritique ( $\gamma_o = 0.35, \lambda_o = 0.25$ ) ; à droite, le cas souscritique ( $\gamma_o = 0.50, \lambda_o = 0.25$ ) ; en haut évolution (ligne pleine) du biais  $\hat{\gamma}_T - \gamma_o$  et  $\hat{\sigma}_{\hat{\gamma}_T} * 100$  ; en bas évolution (ligne pleine) du biais  $\hat{\lambda}_T - \lambda_o$  et  $\hat{\sigma}_{\hat{\lambda}_T} * 100$ .

(1, ..., 64). Pour éliminer les effets de bords on suppose que  $S$  est un tore, i.e. que des sites  $s, w \in S$  sont voisins s'il existe  $z \in \partial 0$  tel que  $s_i - w_i = z_i \text{ modulo}(64)$  pour  $i = 1, 2$  (Buttell & al. 1993).

Les simulations sont présentées Figure 3 pour 5 PC de paramètres  $\theta_o = (\gamma_o, \lambda_o)$  avec  $\gamma_o = 0.32, 0.35, 0.38, 0.41, 0.44$  et  $\lambda_o = 0.25$ , ceci pour une séquence temporelle  $t = 0, \dots, 100$ . Initialement, tous les sites sont occupés. Le graphique de gauche donne l'évolution de la proportion des sites occupés pour les 5 processus ; l'image de droite donne une configuration spatiale à l'instant  $T = 100$  du PC de paramètres  $\gamma_o = 0.35$  et  $\lambda_o = 0.25$ .

La figure 4 figure l'évolution de  $(T, \hat{\theta}_T)$  pour  $T = 1, \dots, 99$  dans deux situations : à gauche, une situation supercritique ( $\gamma_o = 0.35, \lambda_o = 0.25$ ) ; à droite, une situation souscritique où p.s. il y a extinction du processus ( $\gamma_o = 0.50, \lambda_o = 0.25$ ). Le cas supercritique permet de constater la convergence de l'estimateur alors qu'un biais asymptotique apparaît dans le cas sous-critique.

L'étude empirique de la loi limite de l'estimateur du maximum de PVM a été réalisée sur la base de 100 simulations pour  $t = 1, 99$  dans les deux cas supercritique (à gauche,  $\gamma_o = 0.35, \lambda_o = 0.25$ ) et souscritique (à droite,  $\gamma_o = 0.50, \lambda_o = 0.25$ ). Pour chacun des paramètres, on a testé la normalité ( $H_0$ ) des estimateurs à partir du test du Chi2 au niveau 5% et en définissant 9 classes d'états équiprobables. La statistique  $\chi^2$  utilisée suit donc un loi du Chi2 à 6 degré de liberté sous l'hypothèse de normalité. Les résultats de ces tests sont résumés dans le tableau suivant :

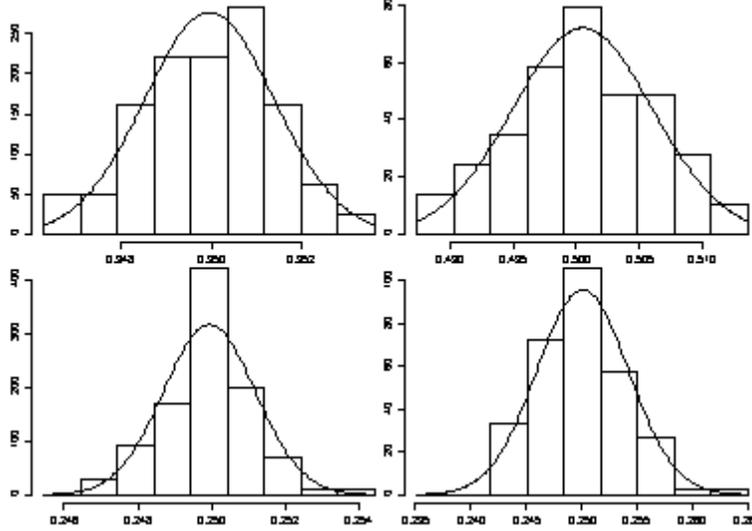


Figure 5: Histogrammes des 100 estimations de  $\gamma_o$  et  $\lambda_o$  : à gauche, le cas supercritique ( $\gamma_o = 0.35, \lambda_o = 0.25$ ) ; à droite, le cas souscritique ( $\gamma_o = 0.45, \lambda_o = 0.25$ ).

Type de PC	PC supercritique		PC souscritique	
$(H_0)$	$\hat{\gamma}_T \sim \mathcal{N}_1$	$\hat{\lambda}_T \sim \mathcal{N}_1$	$\hat{\gamma}_T \sim \mathcal{N}_1$	$\hat{\lambda}_T \sim \mathcal{N}_1$
Valeur du $\chi^2$	0.98	8.54	4.58	3.68

On retient l'hypothèse de normalité dans tous les cas puisque le seuil de rejet est  $\chi_{0.95}^2(6) = 12.59$ . Les histogrammes empiriques des estimations présentés figure 5 confirme cette conclusion.

Une validation complémentaire est de comparer les écarts-types théoriques estimés  $\hat{\sigma}_{\hat{\gamma}}, \hat{\sigma}_{\hat{\lambda}}, \hat{\sigma}_{\hat{\gamma}\mathbf{K}}, \hat{\sigma}_{\hat{\lambda}\mathbf{K}}$  avec les écarts-types empiriques  $s_{\hat{\gamma}}, s_{\hat{\lambda}}, s_{\hat{\gamma}\mathbf{K}}, s_{\hat{\lambda}\mathbf{K}}$  pour les estimateurs du maximum de PVM et ceux de  $\mathbf{K}$ -codage. Les estimations  $\hat{\sigma}_{\hat{\gamma}}, \hat{\sigma}_{\hat{\lambda}}$  ont été obtenues à partir de 18 , où  $A_T = A_T(\hat{\theta}_T)$ ,  $B_T = B_T(\hat{\theta}_T)$ , et  $\hat{\sigma}_{\hat{\gamma}\mathbf{K}}, \hat{\sigma}_{\hat{\lambda}\mathbf{K}}$  à partir de cette même équation en se basant sur les observations sur  $\mathbf{K}$  et avec  $B_T(\hat{\theta}_o) = \mathbf{0}$ . Le modèle retenu pour cette étude est le PC supercritique avec  $\gamma_o = 0.35$  et  $\lambda_o = 0.25$ . Les résultats sont présentés au tableau 1. On remarque qu'il y a peu de différences entre les écarts-types théoriques estimés (à partir d'une simulation) et les écarts-types empiriques (obtenus à partir des 100 estimations). D'autre part, on constate que les écarts-types des estimateurs maximum de PVM sont approximativement dans un rapport de 1 à 3 par rapport à ceux de l'estimation par codage. Ceci est à rapprocher du fait qu'il faut 9 codages disjoints pour recouvrir  $S$ .

Le tableau 2 présente les estimations de  $\gamma_o$  et  $\lambda_o$  pour 12 PC de paramètres  $(\gamma_o, \lambda_o) \in (0.2, 0.4, 0.6) \times (0.1, 0.2, 0.3, 0.4)$  sur le même réseau  $S$ , avec  $T = 4$  et une configuration initiale où 40% de sites choisis au hasard sont occupés.

On termine avec l'étude du test de sous-hypothèse :  $H_0 : \lambda = a_o$  où  $a_o$  est connu. Ce test est étudié dans (Guyon, 1995, Théorème 3.4.6; Bayomog et al. 1996). Pour le PC de paramètres  $\gamma_o = 0.3, \lambda_o = 0.3$ , on obtient  $\tilde{\gamma}_T = 0.307$  pour l'estimateur contraint.

	$\gamma_o = 0.35$			$\lambda_o = 0.25$		
MPV	$\hat{\gamma}$	$\hat{\sigma}_{\hat{\gamma}}$	$s_{\hat{\gamma}_i}$	$\hat{\lambda}$	$\hat{\sigma}_{\hat{\lambda}}$	$s_{\hat{\lambda}}$
MPV du K-codage	$\hat{\gamma}^K$	$\hat{\sigma}_{\hat{\gamma}_i^K}$	$s_{\hat{\gamma}_i^K}$	$\hat{\lambda}^K$	$\hat{\sigma}_{\hat{\lambda}^K}$	$s_{\hat{\lambda}^K}$
	0.3572	0.0076	0.0079	0.2434	0.0059	0.0064
	0.3456	0.0219	0.0218	0.2361	0.0183	0.0181

Table 1: Estimations du MPVM et de **K**-codage des paramètres du PC de paramètres  $\gamma_o = 0.35, \lambda_o = 0.25$  :  $\hat{\gamma}$  (resp.  $\hat{\lambda}$ ), estimation de  $\gamma_o$  (resp.  $\lambda_o$ ) ;  $\hat{\sigma}_{\hat{\gamma}}$  (resp.  $\hat{\sigma}_{\hat{\lambda}}$ ), l'écart-type estimé à partir d'une simulation en appliquant le Théorème 3 ;  $s_{\hat{\gamma}}$  (resp.  $s_{\hat{\lambda}}$ ) l'écart-type empirique obtenu à partir des 100 simulations. Les estimations du **K**-codage sont obtenues de façon analogue en utilisant le codage **K** de la figure 1.

$\gamma$	$\lambda = 0.1$					$\lambda = 0.2$				
	$\hat{\gamma}_T$	$\hat{\sigma}_{\hat{\gamma}_T}$	$\hat{\lambda}_T$	$\hat{\sigma}_{\hat{\lambda}_T}$	$n(T)$	$\hat{\gamma}_T$	$\hat{\sigma}_{\hat{\gamma}_T}$	$\hat{\lambda}_T$	$\hat{\sigma}_{\hat{\lambda}_T}$	$n(T)$
0.2	0.210	0.006	0.104	0.003	14598	0.189	0.006	0.193	0.004	15286
0.4	0.391	0.008	0.106	0.004	12402	0.399	0.008	0.188	0.005	13565
0.6	0.597	0.009	0.100	0.005	9224	0.607	0.009	0.206	0.008	10452
$\gamma$	$\lambda = 0.3$					$\lambda = 0.4$				
	$\hat{\gamma}_T$	$\hat{\sigma}_{\hat{\gamma}_T}$	$\hat{\lambda}_T$	$\hat{\sigma}_{\hat{\lambda}_T}$	$n(T)$	$\hat{\gamma}_T$	$\hat{\sigma}_{\hat{\gamma}_T}$	$\hat{\lambda}_T$	$\hat{\sigma}_{\hat{\lambda}_T}$	$n(T)$
0.2	0.202	0.006	0.294	0.006	15761	0.198	0.008	0.400	0.008	15843
0.4	0.394	0.008	0.305	0.007	14729	0.393	0.009	0.400	0.009	15106
0.6	0.596	0.009	0.292	0.010	11468	0.604	0.009	0.386	0.012	12402

Table 2: Estimation des paramètres et de leurs écarts-types à partir de simulations de PC sur un lattice  $64 \times 64$  et avec  $T = 4$  : par exemple, pour le PC de paramètres  $\gamma_o = 0.2, \lambda_o = 0.4, \hat{\gamma}_4 = 0.198, \hat{\lambda}_4 = 0.400, \hat{\sigma}_{\hat{\gamma}_4} = 0.008, \hat{\sigma}_{\hat{\lambda}_4} = 0.008$ .

La statistique de test  $\Delta = 2 \cdot a_T [U_T(\tilde{\theta}) - U_T(\hat{\theta})]$  vaut 0.001. La seule valeur propre non-nulle de  $J_T^{1/2}(\hat{\theta}_T) [I_T^{-1}(\hat{\theta}_T) - R_0 \tilde{I}_T^{-1}(\tilde{\gamma}_T) {}^t R_0] J_T^{1/2}(\hat{\theta}_T)$  valant  $w = 0.782$ , l'hypothèse nulle ( $H_0$ ) :  $\lambda_o = 0.3$  est acceptée au niveau  $5^\circ / \circ$  puisque  $\Delta \stackrel{\mathcal{L}}{\sim} w \cdot \chi_{0.95}^2(1)$  et  $\chi_{0.95}^2(1) = 3.81$ .

## References

- Arnaudiès, J.M., Fraysse, H. (1994) *Cours de mathématiques - Analyse*, Dunod.
- Bayomog, S., Guyon, X., Hardouin, C., Yao, J. (1996) Test de différence de contrastes et somme pondérée de khi-deux, *The Canadian Journal of Statistics*. **24**, 115-130.
- Besag, J. (1974) Spatial Interaction and the Statistical Analysis of Lattice Systems. *JRSS, B*, **36**, 192-225.
- Buttell, L., Cox, J.T., Durrett, R. (1993) Estimating the critical values of stochastic growth models. *J. Appl. Prob.*, **30**, 455-461.

- Dacunha-Castelle, D., Duflo, M. (1986) *Probability & Statistics - volume 2*, Springer-Verlag.
- Durrett, R., Levin, S.A. (1994) Stochastic spatial models : a user's guide to ecological applications. *Phil. Trans. R. Soc. Lond. B*, **343**, 329-350.
- Durrett, R. (1995) *Ten lectures on particle systems*. Cours de Saint. Flour (1993). L.N.M. n° **1608**, 97-201, Springer.
- Fiocco, M. (1997) Statistical estimation for the supercritical contact process, *PHD Thesis*, Leiden University, Mathematical Institute, The Netherlands.
- Fiocco, M., Zwet, W.R. (2002) Statistics for the contact process, *Statistica Neerlandica*, **56**, No. 2, 243-251.
- Geman, D., Geman, S. (1984) Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images, *IEEE-PAMI* **6**, 721-741.
- Griffeath, D. (1981) The basic contact processes. *Stochastic Processes and their Applications*, **11**, 151-185.
- Guyon, X. (1995) *Random fields on a network : modelling, statistics and applications*, Springer-Verlag.
- Guyon, X., Künsch, H.R., (1992) Asymptotic comparison of estimators in the Ising model, L.N.S. n° 74, Springer, 177-198.
- Guyon, X., Hardouin, C. (2001) Markov chain Markov field dynamics : Models and Statistics, *Statistics*, **35**, 593-627.
- Hall, P., Heyde, C.C.(1980) *Martingale limit theory and its application*, Academic Press.
- Harris, T.E. (1974) Contact interactions on a lattice. *Ann. Prob.*, **2**, 969-988.
- Jensen, J.L., Künsch, H.R. (1994) On asymptotic normality of pseudo-likelihood estimate for pairwise interaction processes. *Ann. Inst. Statist. Math.*, **46**, 475-486.
- Pumo, B., Le Corff, J. (2001) Parameter estimation of the contact process on a lattice. *Proceeding of 10-th Applied Statistic Models and Data Analysis, Compiègne, France*, 866-871.
- Shiryaev, A.N. (1996) *Probability*, 2nd edition, Springer-Verlag.
- Splus 2000 *Guide to Statistics*, Data Analysis Products Division. Mathsoft, Seattle.