



HAL
open science

L'automate Epidémie et le modèle d'Eden face à l'irrégularité

Lucas Gerin

► **To cite this version:**

| Lucas Gerin. L'automate Epidémie et le modèle d'Eden face à l'irrégularité. 2014. hal-00981609v1

HAL Id: hal-00981609

<https://hal.science/hal-00981609v1>

Preprint submitted on 22 Apr 2014 (v1), last revised 7 Nov 2014 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

L'automate Épidémie et le modèle d'Eden face à l'irrégularité

Lucas Gerin

22 avril 2014

Résumé

Nous présentons quelques propriétés, plus ou moins bien connues, de l'automate cellulaire probabiliste Épidémie sur une grille finie, et de son analogue sur la grille infinie : le modèle d'Eden en temps discret.

Mots-clés : Automates Cellulaires Probabilistes, Modèle d'Eden, Épidémie, modèle de croissance

1 Introduction

Pour totalement comprendre un système dynamique, il faut être capable de décrire son comportement lorsqu'il est soumis à différents types de perturbations : on peut penser à des modifications des conditions initiales ou des erreurs (de différentes natures) lors de la mise à jour du système. Du point de vue de l'informatique, que ces perturbations soient destructrices ou tout à fait positives, il faut savoir les analyser pour comprendre ce que le système "calcule".

Nous discutons ici de différents types de perturbations aléatoires d'un modèle particulier d'automate cellulaire : l'Épidémie. Il s'agit d'un "modèle jouet" très simple, décrivant la propagation d'une épidémie sans guérison dans un graphe. L'objectif de cet article est de présenter le comportement d'Épidémie vis-à-vis de plusieurs sources d'aléatoire et d'irrégularité :

- aléatoire dans le schéma de mise à jour,
- irrégularité de la condition initiale,
- irrégularité, ou défauts, du réseau.

Nous allons étudier ces aspects sur deux variantes d'Épidémie : une version sur une grille carrée finie (dans la section suivante), et à la fin de l'article l'analogue sur tout le réseau \mathbb{Z}^2 . Cette dernière version est souvent appelée *modèle d'Eden* par les probabilistes.

L'automate Épidémie nous paraît un bon candidat pour l'étude de l'irrégularité des automates cellulaires. Il est à la fois suffisamment riche pour faire apparaître des phénomènes intéressants et suffisamment simples pour qu'on

soit capable de démontrer des choses rigoureusement. Les résultats présentés ici sont tous plus ou moins connus, parfois écrits d'une autre manière. Ils ne sont toutefois pas si faciles à identifier dans la littérature. L'objectif ici est de présenter certains résultats de façon autonome, quitte à ne pas toujours donner les énoncés optimaux. Les résultats sur l'Épidémie sur une grille finie sont tous démontrés, la section sur le modèle d'Eden est plutôt un guide de lecture dans la littérature, et parfois une "traduction" des énoncés pour non-probabilistes.

2 L'Épidémie sur une grille finie

2.1 Définition du modèle

Soit L un entier supérieur ou égal à 3, on note Λ la grille carrée de taille L , avec des conditions au bord torique (on identifie donc Λ avec $\mathbb{Z}/L\mathbb{Z} \times \mathbb{Z}/L\mathbb{Z}$). On pose $n = L^2$ le nombre total de cellules dans Λ , les asymptotiques seront écrites en fonction de n .

Nous munissons Λ de la distance de graphe usuelle, pour cette distance, la boule $B(c, r)$ de centre c et de rayons r est l'ensemble des cellules c' telles qu'il existe un *chemin* de $0 \leq \ell \leq r$ cellules voisines

$$c = c_0 \rightarrow c_1 \rightarrow c_2 \cdots \rightarrow c_\ell = c'.$$

Une *configuration* σ est l'un des 2^n éléments de $\{0, 1\}^\Lambda$. Pour $c \in \Lambda$, $\sigma_c \in \{0, 1\}$ est l'*état* de la cellule c dans la configuration σ .

Pour c dans Λ , $\mathcal{N}(c)$ est le voisinage de *Von Neumann* de c :

$$\mathcal{N}(c) = \{c, c + (1, 0), c + (-1, 0), c + (0, 1), c + (0, -1)\}$$

où $+$ est l'addition modulo L . Autrement dit, $\mathcal{N}(c) = B(c, 1)$.

Nous pouvons maintenant définir notre automate cellulaire comme un système dynamique aléatoire. Chaque cellule dans l'état 0 (sain), lorsqu'elle est mise à jour, se retrouve dans l'état 1 (infecté) si l'un (au moins) de ses voisins est infecté. Il n'y a pas de guérison : un 1 reste 1 pour toujours. Par ailleurs, les mises à jour se font de façon aléatoire et indépendantes les unes des autres.

Définition 1 (L'Épidémie dans la dynamique α -synchrone). *Soit $\alpha \in]0, 1[$, l'automate Épidémie α -synchrone est le processus aléatoire à temps discret $(\sigma^t)_{t \geq 0}$ à valeurs dans $\{0, 1\}^\Lambda$ défini par $\sigma^0 \in \{0, 1\}^\Lambda$ et l'évolution suivante :*

Soit $t \geq 0$, connaissant σ^t au temps t , la configuration σ^{t+1} est construite ainsi : chaque cellule de Λ est mise à jour avec probabilité α (indépendamment du passé et des $n - 1$ autres cellules) ;

- Si c est mise à jour, $\sigma_c^{t+1} = 1$ si et seulement si au moins un élément de $\mathcal{N}(c)$ à l'instant t est dans l'état 1.*
- Sinon, $\sigma_c^{t+1} = \sigma_c^t$.*

La suite (σ^t) est alors une chaîne de Markov à temps discret à valeurs dans $\{0, 1\}^\Lambda$. Clairement, elle finit par atteindre l'une des deux seules configurations fixes 0^Λ ou 1^Λ . La configuration 0^Λ est isolée :

$$\sigma^t \text{ atteint } 0^\Lambda \Leftrightarrow \sigma^0 = 0^\Lambda.$$

Nous excluons le cas trivial $\sigma^0 = 0^\Lambda$ dans la suite.

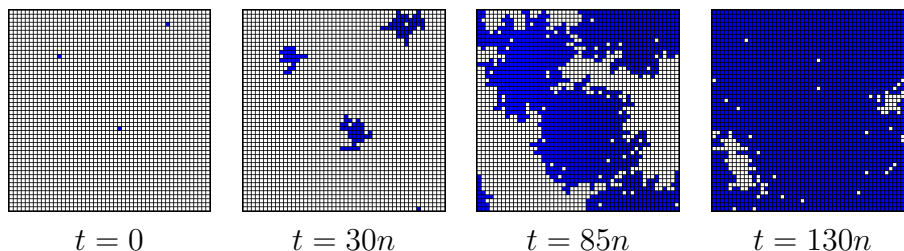


FIGURE 1 – Une simulation d'Épidémie α -synchrone pour $\alpha = 0.1$ et $L = 50$, avec σ^0 qui n'a que 3 cellules dans l'état 1. Les simulations ont été obtenues avec l'application FiatLux : <http://fiatlux.loria.fr/>.

Définition 2 (Temps de convergence). *Pour toute configuration initiale $\sigma^0 \neq 0^\Lambda$, la variable aléatoire $T_n(\sigma^0)$ est le Temps de convergence : le premier instant t où toutes les cellules sont infectées :*

$$T_n(\sigma^0) = \min \{t \geq 0 \text{ tel que } \sigma^t = 1^\Lambda\}.$$

Dans cette section nous allons nous concentrer sur le comportement asymptotique de l'espérance de cette variable aléatoire : $\mathbb{E}[T_n(\sigma^0)]$, dans les cas le *pire* et *typique*.

2.2 Le pire temps de convergence moyen

On considère le *pire* temps de convergence moyen (WECT en anglais) pour Épidémie, c'est-à-dire le temps moyen de convergence lorsque $\sigma^0 = \sigma^w$ où σ^w est tel que

$$\mathbb{E}[T_n(\sigma^w)] = \max_{\substack{\sigma^0 \in \{0,1\}^\Lambda \\ \sigma^0 \neq 0^\Lambda}} \mathbb{E}[T_n(\sigma^0)].$$

Il est évident que les tels σ^w sont exactement les n configurations avec un unique 1.

Avant d'énoncer le résultat, voici quelques motivations pour l'étude du WECT :

1. En dimension un, [FMST06] ont analysé le WECT des Automates Cellulaires Élémentaires avec deux états quiescents (voir leur article pour la définition). Leur travail a révélé que le comportement asymptotique du

WECT donne une classification pertinente à l'intérieur de cette famille naturelle d'Automates Cellulaires. Plus précisément, ils ont démontré que ces règles peuvent être classées en 5 familles, selon que le WECT croît en $n \log n$, n^2 , n^3 , $n2^n$, ou est infini. Cette approche a été étendue dans [FG09] pour une famille d'Automates Cellulaires de dimension deux (en particulier pour l'Épidémie).

2. Une autre motivation vient de la théorie de la complexité, si l'on voit les Automates Cellulaires comme des modèles de calcul. Avec ce point de vue, il est naturel de se demander ce qui se passe dans le pire cas.
3. Enfin, si l'on considère les Automates Cellulaires comme des modèles (simplifiés!) de systèmes physiques ou biologiques, ce problème devient un problème de *relaxation* : que se passe-t-il lorsque l'on force le système à s'éloigner le plus possible de son équilibre.

Théorème 1 (Pire temps de convergence moyen sur une grille). *Pour Épidémie sur une grille à n cellules, pour tout $\alpha \in (0, 1)$, si n est assez grand,*

$$\frac{\sqrt{n}}{8\alpha} \leq \mathbb{E}[T_n(\sigma^w)] \leq 3\frac{\sqrt{n}}{\alpha}.$$

Remarque. – Dans [FG09], un résultat tout à fait analogue (mais un peu moins précis car il y avait des termes en $\log(n)$ de chaque côté) avait été obtenu pour la dynamique asynchrone.

- Il est très vraisemblable que $\mathbb{E}[T_n(\sigma^w)]/\sqrt{n}$ tende vers une constante, et même que l'on ait convergence en probabilités de la suite de variables aléatoires $T_n(\sigma^w)/\sqrt{n}$. Nous n'avons pas cherché à le démontrer, mais l'outil habituel pour faire cela (la sous-additivité, voir [BS10] par exemple) ne semble pas fonctionner aussi facilement que cela, principalement en raison des conditions toriques au bord.

Démonstration.

Borne inférieure. Soit c le seul 1 dans σ^w , et soit c' l'une des cellules de Λ qui est à distance $\lfloor L/2 \rfloor$ de c . Il suffit de montrer que $\mathbb{E}[\tau_{c'}] \geq \frac{L}{8\alpha}$, avec

$$\tau_{c'} = \min \{ t \geq 0, \sigma_{c'}^t = 1 \},$$

car bien sûr $T_n(\sigma^w) \geq \tau_{c'}$.

Posons $k = \lfloor L/4\alpha \rfloor$, nous allons montrer que si L est assez grand

$$\mathbb{P}(\tau_{c'} < k) \leq 1/2. \tag{1}$$

Ceci implique la borne inférieure car alors

$$\mathbb{E}[T_n(\sigma^w)] \geq \mathbb{E}[\tau_{c'}] \geq \sum_{r=1}^k \mathbb{P}(\tau_{c'} \geq r) \geq \sum_{r=1}^k 1/2 = \frac{1}{2} (\lfloor L/4\alpha \rfloor + 1) \geq \frac{\sqrt{n}}{8\alpha}.$$

On se concentre donc sur (1). La preuve n'est pas difficile mais un peu technique, voici l'idée générale. Comme souvent dans ce genre de problèmes, il y a un compromis à comprendre entre les deux phénomènes suivants :

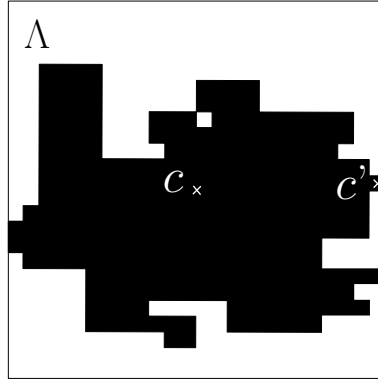


FIGURE 2 – La configuration à l’instant $\tau_{c'}$.

- Il y a potentiellement beaucoup de chemins (ce sera le terme 3^j dans la preuve) allant de c à c' le long desquels des mises à jours successives finiraient par contaminer la cellule c' ;
- D’un autre côté, si l’on fixe un tel chemin il est très improbable (ce sera l’objet du Lemme 1) que les cellules de ce chemin particulier soient mises à jour dans le bon ordre avant k , si k est bien choisi.

On détaille maintenant la preuve. Supposons que $\tau_{c'} \leq k$, cela veut dire qu’il existe un j avec $\lfloor L/2 \rfloor \leq j \leq k$ et un chemin \mathcal{P} de j cellules toutes disjointes, voisines deux à deux, allant de c à c' :

$$\mathcal{P} = \{c = c_0 \rightarrow c_1 \rightarrow c_2 \rightarrow \dots \rightarrow c_j = c'\}$$

tel que, durant les k premières unités de temps, les cellules c_1, \dots, c_j sont mises à jour *dans cet ordre*. Autrement dit,

$$\begin{aligned} \mathbb{P}(\tau_{c'} \leq k) &= \mathbb{P} \left(\bigcup_{j=\lfloor L/2 \rfloor}^k \bigcup_{\mathcal{P}, |\mathcal{P}|=j} \{c_1, c_2, \dots, c_j \text{ sont mises à jour dans cet ordre et avant } k\} \right) \\ &\leq \sum_{j=\lfloor L/2 \rfloor}^k \sum_{\mathcal{P}, |\mathcal{P}|=j} \mathbb{P}(c_1, c_2, \dots, c_j \text{ sont mises à jour dans cet ordre et avant } k), \end{aligned}$$

(ici nous avons utilisé le fait que $\mathbb{P}(\cup) \leq \sum \mathbb{P}$, souvent appelée *union bound*). La seconde somme porte sur les chemins de longueur j allant de c à c' . Fixons un tel \mathcal{P} and majorons cette dernière probabilité. Parmi les temps $\{1, 2, \dots, k\}$, il y a des instants $t_1 < t_2 < \dots < t_j$ tels que, au temps t_j , la cellule c_j est mise à jour. Chaque mise à jour ayant lieu indépendamment et avec probabilité α , nous avons

$$\mathbb{P}(c_1, c_2, \dots, c_j \text{ sont mises à jour dans cet ordre et avant } k) = \mathbb{P}(\text{Binom}(k, \alpha) \geq j),$$

où $\text{Binom}(k, \alpha)$ suit la loi binomiale de paramètres k, α . Le lemme suivant est alors utile, c’est par exemple (2.5) dans [JLR00].

Lemme 1 (Déviations à droite pour la binomiale). *Pour tous $j \geq k\alpha$,*

$$\mathbb{P}(\text{Binom}(k, \alpha) \geq j) \leq \exp\left(-3\frac{(j - k\alpha)^2}{2k\alpha + j}\right).$$

Puisqu'il y a moins que 3^j chemins de longueur j allant de c à c' (c'est une majoration grossière mais suffisante ici), nous avons

$$\mathbb{P}(\tau_{c'} \leq k) = \sum_{j=\lfloor L/2 \rfloor}^k 3^j \exp\left(-3\frac{(j - k\alpha)^2}{2k\alpha + j}\right).$$

En étudiant les variations de cette suite, on se convainc que l'expression $j \mapsto 3^j \exp\left(-3\frac{(j - k\alpha)^2}{2k\alpha + j}\right)$ est décroissante pour $L/2 \leq j \leq k$ (rappelons que $k\alpha \approx L/8$) et ainsi, en oubliant les parties entières pour alléger les équations,

$$\begin{aligned} \mathbb{P}(\tau_{c'} \leq k) &\leq k \times \max_j \left\{ 3^j \exp\left(-3\frac{(j - k\alpha)^2}{2k\alpha + j}\right) \right\} \\ &\leq k \exp\left(\log(3)L/2 - 3\frac{(L/2 - k\alpha)^2}{2k\alpha + L/2}\right) \\ &\leq \frac{L}{4\alpha} \exp\left(\log(3)L/2 - 3L\frac{(1/2 - 1/4)^2}{2/4 + 1/2}\right) \leq \frac{L}{4\alpha} \exp(-0.2 \times L), \end{aligned}$$

et donc est plus petite que $1/2$ si L est assez grand (dépendant de α) et nous avons démontré (1).

Borne supérieure. Nous allons montrer que pour L assez grand et $k \geq 2L/\alpha$ alors

$$\mathbb{P}(T_n(\sigma^w) > k) \leq L^2 \exp(-k\alpha/32). \quad (2)$$

La borne supérieure en découlera car

$$\begin{aligned} \mathbb{E}[T_n(\sigma^w)] &= \sum_{k \geq 1} \mathbb{P}(T_n(\sigma^w) \geq k) \\ &\leq 2\sqrt{n}/\alpha + \sum_{k \geq 2\sqrt{n}/\alpha} \mathbb{P}(T_n(\sigma^w) \geq k) \\ &\leq 2\sqrt{n}/\alpha + L^2 \sum_{k \geq 2\sqrt{n}/\alpha} \exp(-k\alpha/32) \leq 3\sqrt{n}/\alpha \end{aligned}$$

pour L assez grand.

Démontrons (2). Tout d'abord, pour chaque $c' \neq c$, on choisit arbitrairement (et de façon déterministe) un chemin $\mathcal{P}_{c'}$ parmi les plus courts chemins allant de c à c' : $\mathcal{P}_{c'}$ peut s'écrire

$$\mathcal{P}_{c'} = \{c = c_0 \rightarrow c_1 \rightarrow c_2 \rightarrow \dots \rightarrow c_j = c'\}$$

où $j \leq L/2$ est la distance entre c et c' .

Dire que $T_n(\sigma^w) > k$ c'est dire qu'il y a une cellule c' qui est toujours dans l'état 0 au temps k . En particulier, il y a un $j \leq L/2$ et une cellule c' à distance j tels que les cellules c_1, c_2, \dots, c_j de son chemin associé $\mathcal{P}_{c'}$ n'ont pas été mises à jour dans cet ordre avant k :

$$\begin{aligned} \mathbb{P}(T_n(\sigma^w) > k) &= \mathbb{P}\left(\bigcup_{c' \in \Lambda} \{c_1, c_2, \dots, c_j \text{ ne sont pas mises à jour dans cet ordre avant } k\}\right) \\ &\leq \text{card}(\Lambda) \max_{c' \in \Lambda} \mathbb{P}(c_1, c_2, \dots, c_j \text{ ne sont pas mises à jour dans cet ordre avant } k). \end{aligned}$$

et par le même argument que pour la borne inférieure

$$\mathbb{P}(c_1, c_2, \dots, c_j \text{ ne sont pas mises à jour dans cet ordre avant } k) = \mathbb{P}(\text{Binom}(k, \alpha) < j),$$

nous avons donc besoin de la majoration suivante, c'est (2.6) dans [JLR00].

Lemme 2 (Déviations à gauche pour la binomiale). *Pour tous $j \leq k\alpha$,*

$$\mathbb{P}(\text{Binom}(k, \alpha) < j) \leq \exp\left(-\frac{(k\alpha - j)^2}{2k\alpha}\right).$$

Pour tout $j \leq L/2$, nous avons $j \leq k\alpha$ (rappelons que $k \geq 2L/\alpha$) et on peut alors appliquer le Lemme :

$$\begin{aligned} \mathbb{P}(T_n(\sigma^w) > k) &\leq L^2 \max_{1 \leq j \leq L/2} \exp\left(-\frac{(k\alpha - j)^2}{2k\alpha}\right) \\ &\leq L^2 \exp\left(-\frac{(k\alpha - L/2)^2}{2k\alpha}\right) \\ &\leq L^2 \exp\left(-\frac{(k\alpha - 3k\alpha/4)^2}{2k\alpha}\right) \quad \text{car } (k\alpha - L/2)^2 \geq k\alpha/2 \\ &\leq L^2 \exp(-k\alpha/32), \end{aligned}$$

et nous avons démontré (2). □

2.3 Le temps de convergence moyen typique

Nous allons maintenant estimer le temps de convergence moyen typique, lorsque σ^0 est choisi uniformément au hasard dans Q^Λ .

$$\text{Typ}_n := \frac{1}{2^n} \sum_{\sigma^0 \in \{0,1\}^\Lambda} \mathbb{E}[T_n(\sigma^0)].$$

Comme attendu, Typ_n croît beaucoup moins vite que dans le pire cas.

Théorème 2 (Temps de convergence moyen typique sur une grille finie). *Pour n assez grand,*

$$\frac{1}{4\alpha} \log n \leq \text{Typ}_n \leq \frac{6}{\alpha} (\log n)^{3/2}.$$

Démonstration. Nous suivons ([Ger08], Chap.2).

Borne inférieure. Le nombre de cellules dans l'état 0 dans σ^0 , qui est une Binomiale de paramètres $(n, 1/2)$, est plus grand que $n/2$ avec plus de 50% de chance. Pour de tels σ^0 , la convergence est bien sûr plus longue que le temps nécessaire à mettre à jour ces cellules. Ainsi

$$\mathbb{E} [T_n(\sigma^0)] \geq \mathbb{E} [\max \{\mathcal{G}_1, \mathcal{G}_2, \dots, \mathcal{G}_{n/2}\}],$$

où les \mathcal{G}_i sont des variables aléatoires indépendantes suivant la loi géométrique de paramètre α . Pour un grand k nous avons (c'est plus ou moins dans le folklore, voir [SR90] par exemple)

$$\frac{2 \log(k)}{3\alpha} \leq \mathbb{E} [\max \{\mathcal{G}_1, \mathcal{G}_2, \dots, \mathcal{G}_k\}] \leq \frac{2 \log(k)}{\alpha} \quad (3)$$

et alors, pour n assez grand, $\mathbb{E} [T_n(\sigma^0)] \geq \frac{2}{3\alpha} \log(n/2) \geq \frac{1}{2\alpha} \log(n)$ lorsque σ^0 a plus de $n/2$ cellules dans l'état 0. Maintenant,

$$\begin{aligned} \text{Typ}_n &= \frac{1}{2^n} \sum_{\sigma^0 \in \{0,1\}^\Lambda} \mathbb{E} [T_n(\sigma^0)] \geq \frac{1}{2^n} \sum_{\substack{\sigma^0 \text{ avec} \\ \text{plus de } n/2 \text{ 0}}} \mathbb{E} [T_n(\sigma^0)] \\ &\geq \frac{1}{2^n} \text{card} \{ \sigma^0 \text{ avec plus de } n/2 \text{ 0} \} \frac{1}{2\alpha} \log(n) \\ &\geq \frac{1}{2^n} \frac{2^n}{2} \frac{1}{2\alpha} \log(n), \end{aligned}$$

d'où la borne inférieure.

Borne supérieure. La première étape est de montrer que, avec grande probabilité, il n'y a pas dans σ^0 de boule de rayon $3\sqrt{\log(L)}$ remplie uniquement de 0.

Précisément, posons $a(L) = \lfloor 3\sqrt{\log L} \rfloor$, et définissons l'événement

$$\begin{aligned} A &= \bigcup_{c \in \Lambda} \{ \text{la boule de centre } c \text{ et de rayon } a(L) \text{ est vide à l'instant 0.} \} \\ &= \bigcup_{c \in \Lambda} \{ \forall c' \in B(c, a(L)), \sigma_{c'}^0 = 0 \}. \end{aligned}$$

Par la *union bound*,

$$\mathbb{P}(A) \leq \sum_{c \in \Lambda} \mathbb{P}(\forall c' \in B(c, a(L)), \sigma_{c'}^0 = 0).$$

Chaque boule $B(c, a(L))$ contenant plus de $2a(L)^2$ cellules, elle est remplie de 0 avec une probabilité inférieure à $(1/2)^{2a(L)^2}$, on obtient alors (on oublie à nouveau les parties entières pour alléger les notations)

$$\mathbb{P}(A) \leq \text{card}(\Lambda)(1/2)^{2a(L)^2} \leq L^2(1/2)^{2a(L)^2} \leq \exp(2 \log L - 2 \times 3^2 \log(L) \log 2) \leq 1/L = 1/\sqrt{n}$$

(pour un grand n). Écrivons maintenant

$$\begin{aligned}
\frac{1}{2^n} \sum_{\sigma^0} \mathbb{E}[T(\sigma^0)] &= \frac{1}{2^n} \sum_{\substack{\sigma^0 \text{ tels que} \\ A \text{ est vrai}}} \mathbb{E}[T(\sigma^0)] + \frac{1}{2^n} \sum_{\substack{\sigma^0 \text{ tels que} \\ A \text{ est faux}}} \mathbb{E}[T(\sigma^0)], \\
&\leq \mathbb{P}(A) \max_{\substack{\sigma^0 \text{ tels que} \\ A \text{ est vrai}}} \mathbb{E}[T(\sigma^0)] + \mathbb{P}(\text{non } A) \max_{\substack{\sigma^0 \text{ tels que} \\ A \text{ est faux}}} \mathbb{E}[T(\sigma^0)], \\
&\leq 1/\sqrt{n} \max_{\substack{\sigma^0 \text{ tels que} \\ A \text{ est vrai}}} \mathbb{E}[T(\sigma^0)] + 1 \times \max_{\substack{\sigma^0 \text{ tels que} \\ A \text{ est faux}}} \mathbb{E}[T(\sigma^0)]. \quad (4)
\end{aligned}$$

Majorons chacun des deux termes :

- $\max_{\substack{\sigma^0 \text{ tels que} \\ A \text{ est vrai}}} \mathbb{E}[T(\sigma^0)]$ est un $\mathcal{O}(\sqrt{n})$ par la borne supérieure du WECT ;
- Si A est faux, chaque $\mathbf{0}$ est à moins de $a(L)$ cellules d'un $\mathbf{1}$. L'automate a donc convergé au pire après le succès de $a(L)$ collectionneurs de coupons successifs avec n coupons (on appelle *collectionneur de coupons* le processus qui consiste à attendre que les n coupons aient été tirés au sort). D'après (3), un collectionneur de coupons à n coupons (les n cellules ici) prend un temps moyen inférieur à $2 \log(n)/\alpha$ et donc nous avons

$$\max_{\sigma^0; A \text{ est faux}} \mathbb{E}[T(\sigma^0)] \leq a(L) \frac{2}{\alpha} \log(n).$$

Et (4) donne enfin

$$\text{Typ}_n \leq c^{\text{st}} \sqrt{n} \frac{1}{\sqrt{n}} + 3 \sqrt{\log(\sqrt{n})} \frac{2}{\alpha} \log(n) \leq \frac{6}{\alpha} \log(n)^{3/2},$$

pour n assez grand. □

Remarque. *En fait, en modifiant légèrement (2), il est possible d'améliorer la borne supérieure pour passer de $\mathcal{O}(\log(n)^{3/2})$ à $\mathcal{O}(\log(n))$. L'idée est qu'une boule de $\mathbf{0}$ de rayon $\log(L)$ est entièrement remplie de $\mathbf{1}$ en moins de $\mathcal{O}(\log(L))$ étapes (avec grande probabilité).*

3 Épidémie sur \mathbb{Z}^2 : modèle d'Eden

On considère maintenant l'analogie d'Épidémie sur la grille infinie \mathbb{Z}^2 , on l'appelle généralement *Modèle d'Eden* [Ede61] ou *Modèle de Richardson* [Ric73]. Soit $\alpha > 0$, on considère le processus aléatoire $(\sigma^t)_{t \geq 0}$ à valeurs dans $\{0, 1\}^{\mathbb{Z}^2}$ défini de la façon suivante :

- $\sigma_{\mathbf{0}}^0 = 1$ et $\sigma_c^0 = 0$ pour $c \neq \mathbf{0}$, où $\mathbf{0}$ désigne l'origine de \mathbb{Z}^2 .
- Au temps $t + 1$, chaque $\mathbf{0}$ dans \mathbb{Z}^2 qui a un voisin $\mathbf{1}$ dans σ^t devient $\mathbf{1}$ avec probabilité α , indépendamment du passé et des autres cellules.

On s'intéresse au comportement asymptotique de $(\sigma^t)_{t \geq 0}$. Bien sûr en temps fini l'automate ne converge pas vers une configuration fixe, mais on observe

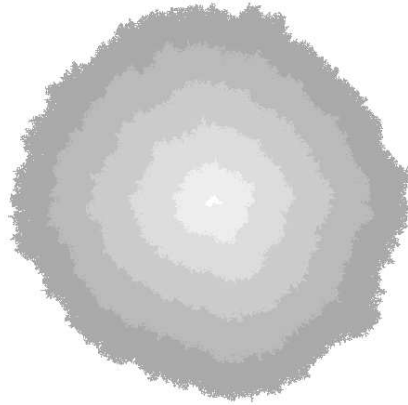


FIGURE 3 – Le modèle d’Eden avec $\alpha = 0.02$, à différents instants jusqu’à 10^6 .

sur les simulations (voir Figure 3) que la composante de 1 semble croître selon une forme particulière, peut-être un disque.

Richardson a en effet démontré que l’ensemble des cellules dans l’état 1 a une *forme limite*, dans le sens suivant.

Théorème 3 (Théorème de forme pour le modèle d’Eden). *Soit B_t l’ensemble des cellules dans l’état 1 au temps t . Il existe un ensemble déterministe $B_\star \subset \mathbb{R}^2$ compact, convexe et non vide tel que pour tout $\varepsilon > 0$,*

$$\mathbb{P} \left(B_\star(1 - \varepsilon) \subset \frac{B_t}{t} \subset B_\star(1 + \varepsilon) \right) \xrightarrow{t \rightarrow +\infty} 1.$$

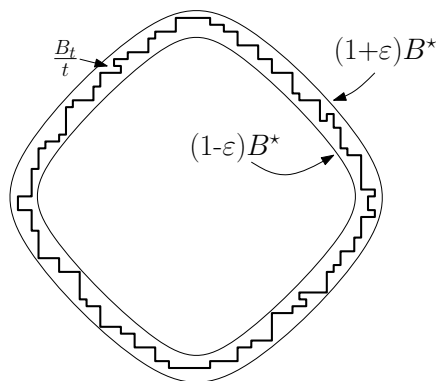


FIGURE 4 – Ce que dit le Théorème de forme, pour t grand.

Le résultat est donc dû à D.Richardson, [CD81] l’ont amélioré en démontrant la convergence presque-sûre.

3.1 Lien avec la percolation de premier passage

Le modèle d'Eden est un modèle de croissance dynamique, mais on peut aussi le voir comme un objet statique. Pour cela posons, comme précédemment, pour chaque $c' \in \mathbb{Z}^2$

$$\tau_{c'} = \min \{t \geq 0, \sigma_{c'}^t = 1\}.$$

Comme pour l'Épidémie, $\tau_{c'} \leq k$ si et seulement s'il y a un chemin \mathcal{P} de $j \leq k$ cellules voisines disjointes allant de $\mathbf{0}$ à c'

$$\mathcal{P} = \{\mathbf{0} \rightarrow c_1 \rightarrow c_2 \rightarrow \dots \rightarrow c_j = c'\}$$

tel que des mises à jour successives le long de \mathcal{P} ont changé c' en 1. On peut alors se convaincre que, pour c' fixé,

$$\tau_{c'} \stackrel{\text{(loi)}}{=} \min_{\mathcal{P}: \mathbf{0} \rightarrow c'} \sum_{i=1}^{|\mathcal{P}|} g_{c_i}, \quad (5)$$

où $\{g_c, c \in \mathbb{Z}^2\}$ est une famille de v.a. i.i.d. géométriques de paramètre α . La variable aléatoire g_{c_i} est ici le temps qu'il a fallu pour mettre à jour c_i , à partir du moment où l'un de ses voisins était 1.

Autrement dit, un moyen de construire $\tau_{c'}$ est de tirer au sort sur chaque cellule de \mathbb{Z}^2 des temps aléatoires g_{c_i} indépendants, et alors $\tau_{c'}$ est la somme de ces temps sur le chemin $\mathbf{0} \rightarrow c'$ tel que la somme est la minimale. Ce modèle de chemin minimal aléatoire est appelé Percolation de Premier Passage (PPP, *first-passage percolation* en Anglais) et a été étudié mathématiquement pour la première fois par Hammersley et Welsh [HW65]. L'article [BS10] est une introduction moderne à la PPP et à ses liens avec les modèles de croissance.

Pour résumer, le lien entre $\tau_{c'}$ dans le modèle d'Eden et la percolation de premier passage est le suivant :

Proposition 1 (Le modèle d'Eden est équivalent à la PPP). *Soit $\{g_c, c \in \mathbb{Z}^2\}$ une famille de v.a. i.i.d. géométrique de paramètre α . Alors*

$$\left\{ \tau_c, c \in \mathbb{Z}^2 \right\} \stackrel{\text{(loi)}}{=} \left\{ \min_{\mathcal{P}: \mathbf{0} \rightarrow c} \sum_{i=1}^{|\mathcal{P}|} g_{c_i}, c \in \mathbb{Z}^2 \right\}.$$

où le min est pris sur tous les chemins \mathcal{P} de $|\mathcal{P}|$ cellules voisines

$$\mathcal{P} = (\mathbf{0} = c_0 \rightarrow c_1 \rightarrow \dots \rightarrow c_{|\mathcal{P}|} = c).$$

Le lien entre modèle d'Eden et PPP est habituellement attribué à Richardson, même s'il n'est pas explicitement question de PPP dans [Ric73]. Étrangement, il semble qu'il n'existe pas de preuve de la Proposition 1 dans la littérature, c'est souvent considéré comme "bien connu", mais il n'est pas si facile d'écrire une preuve complète et rigoureuse (une difficulté est de démontrer l'égalité en loi pour toute la famille des τ_c , et pas seulement pour c fixée).

3.2 Modèle d'Eden et structure du réseau

Depuis [Ric73], on attribue parfois à Eden la conjecture suivante :

Conjecture (Conjecture d'Eden). *Pour le modèle d'Eden en temps continu, l'ensemble des cellules dans l'état 1 a asymptotiquement une forme circulaire : B_* est une boule euclidienne.*

Le modèle d'Eden en temps continu est défini de façon tout à fait analogue au temps discret, sauf que chaque cellule voisine de B_t attend un temps exponentiel de paramètre 1 avant de devenir 1, au lieu d'une géométrie de paramètre α . Il y a effectivement des simulations dans [Ede61], mais en réalité les conclusions qui en sont tirées sont beaucoup plus prudentes :

As yet the samples of configurations computed in this way appear to be too few to justify anything more than a few qualitative statements. It is to be seen that the colony is essentially circular in outline.

En 1984, H.Kesten a eu l'intuition que cette conjecture devait être fausse au moins en grande dimension, pour des raisons essentiellement géométriques. Il l'a démontré effectivement en dimension supérieure à 600000 (voir [Kes86]), la dimension a été abaissée depuis à 35 dans [Dha88] puis [CEG11].

Ceci peut sembler anecdotique (et très éloigné d'un quelconque modèle physique ou biologique !) mais ces résultats nous disent tout de même quelque chose d'intéressant : le modèle d'Eden, même "vu de loin", *se souvient* du réseau discret sur lequel il est défini. On sait que cela n'est pas vrai pour la position d'une marche aléatoire simple dans \mathbb{Z}^d dont la loi asymptotique est la loi normale, qui est invariante par rotation. Le modèle d'Eden est ainsi sensible à la structure microscopique du réseau. En réalité, la seule propriété que l'on connaît de B_* est celle qu'il a héritée de \mathbb{Z}^2 : il est symétrique par rapport aux deux axes.

3.3 Modèle d'Eden et défauts aléatoires

Dans la mesure où il semble impossible pour l'instant de déterminer la forme B_* , on se demande si cet objet limite est assez robuste. Par exemple, que se passe-t-il dans le Théorème 3 si certaines cellules sont *immunisées* contre la maladie ? Est-ce que la croissance de B_t , si elle a toujours lieu, est toujours linéaire ?

Supposons qu'à l'instant initial chaque cellule est immunisée indépendamment avec probabilité p : elle reste dans l'état 0 indéfiniment, et ne peut contaminer ses voisins. Clairement la croissance infinie de B_t n'est pas possible si l'événement suivant E est réalisé :

soit $\mathbf{0}$ est immunisée,

soit il y a un chemin de cellules voisines, toutes immunisées, qui fait le contour de $\mathbf{0}$.

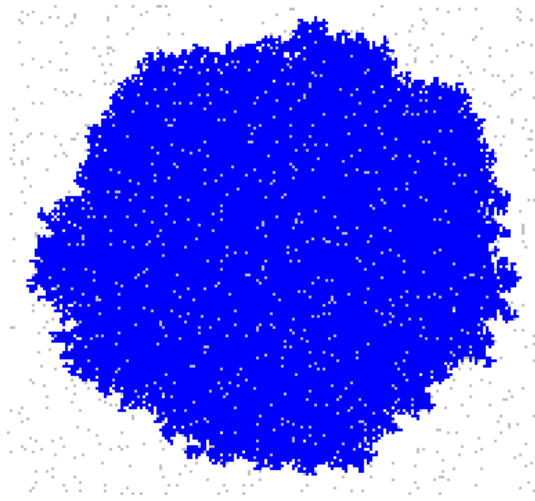


FIGURE 5 – Le modèle d’Eden avec des cellules immunisées (en gris clair).

Bien sûr $\mathbb{P}(E) \geq p$ mais on peut démontrer (voir [Gri99] par exemple) que si p est suffisamment petit alors $\mathbb{P}(E) < 1$. Dans ce cas où E n’est pas réalisé, alors il est possible que l’épidémie se propage indéfiniment, et alors B_t a une croissance infinie et linéaire, comme dans le cas simple. Ceci a été démontré rigoureusement par [GM04], nous avons besoin de quelques notations pour énoncer le résultat.

Soit \vec{n} la cellule de coordonnées $(n, 0)$. Soit \mathcal{A} l’ensemble (aléatoire) des entiers n tels qu’il existe un chemin de cellules non-immunisées joignant l’origine à \vec{n} .

Si E n’est pas réalisé, alors une infinité de cellules dans la grille peuvent être contaminés et si $n \in \mathcal{A}$, le premier instant $\tau_{\vec{n}}$ où \vec{n} est dans B_t est fini, presque-sûrement.

Théorème 4 (Croissance linéaire pour le modèle d’Eden avec défauts). *Il existe une constante $\mu > 0$ telle que, si E n’est pas réalisé, presque-sûrement,*

$$\lim_{\substack{n \rightarrow +\infty, \\ n \in \mathcal{A}}} \frac{\tau_{\vec{n}}}{n} = \mu,$$

où la limite est prise le long de la sous-suite aléatoire $n \in \mathcal{A}$.

Références

- [BS10] N.D. Blair-Stahn. First passage percolation and competition models. [arXiv:1005.0649](https://arxiv.org/abs/1005.0649), 2010.
- [CD81] J.T. Cox and R. Durrett. Some limit theorems for percolation processes with necessary and sufficient conditions. *Annals of Probability*, 9(4) :583–603, 1981.

- [CEG11] O. Couronné, N. Enriquez, and L. Gerin. Construction of a short path in high dimensional first-passage percolation. *Electronic Communications in Probability*, 16 :22–28, 2011.
- [Dha88] D. Dhar. First passage percolation in many dimensions. *Physics Letter A*, 130 :308–310, 1988.
- [Ede61] M. Eden. A two-dimensional growth process. In *Proc. 4th Berkeley Sympos. Math. Statist. and Prob., Vol. IV*, pages 223–239. Univ. California Press, Berkeley, Calif., 1961.
- [FG09] N. Fatès and L. Gerin. Examples of fast and slow convergence of 2d asynchronous cellular systems. *Journal of Cellular Automata*, 4 :323–337, 2009.
- [FMST06] N. Fatès, M. Morvan, N. Schabanel, and É. Thierry. Fully asynchronous behavior of double-quiescent elementary cellular automata. *Theoretical Computer Science*, 362 :1–16, 2006.
- [Ger08] L. Gerin. Aspects probabilistes des automates cellulaires, et d’autres problèmes en informatique théorique. Thèse de l’Université Nancy 1, 2008.
- [GM04] O. Garet and R. Marchand. Asymptotic shape for the chemical distance and first-passage percolation on the infinite Bernoulli cluster. *ESAIM Probab. Stat.*, 8 :169–199 (electronic), 2004.
- [Gri99] G. Grimmett. *Percolation*, volume 321. Springer-Verlag, second edition, 1999.
- [HW65] J. M. Hammersley and D. J. A. Welsh. First-passage percolation, subadditive processes, stochastic networks, and generalized renewal theory. In *Proc. Internat. Res. Semin., Statist. Lab., Univ. California, Berkeley, Calif*, pages 61–110. Springer-Verlag, New York, 1965.
- [JLR00] S. Janson, T. Łuczak, and A. J. Ruciński. *Random graphs*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2000.
- [Kes86] K. Kesten. Aspects of first passage percolation. In *École d’été de probabilités de Saint-Flour, XIV—1984*, volume 1180 of *Lecture Notes in Math.*, pages 125–264. Springer, Berlin, 1986.
- [Ric73] D. Richardson. Random growth in a tessellation. *Proc. Cambridge Philos. Soc.*, 74 :515–528, 1973.
- [SR90] W. Szpankowski and V. Rego. Yet another application of a binomial recurrence order statistics. *Computing*, 43 :401–410, 1990.