Epidemic automaton and the Eden model: various aspects of robustness

Epidemic automaton and the Eden model: various aspects of robustness

Lucas Gerin

CMAP, École Polytechnique

November 5, 2014

Abstract

The two-dimensional probabilistic cellular automaton *Epidemic* models the spread of an epidemic without recovering on graph. We discuss some well-known and less well-known properties of Epidemic on a finite grid and its analogous on the infinite square lattice: the Eden model.

This survey is intended for non-probabilists and gives a detailed study of the robustness of a cellular automaton with respect to several sourcess of randomness.

1 Introduction

We discuss here several random perturbations of a particular (yet very interesting) 2D probabilistic cellular automaton: *Epidemic*. This is a "toy model" for the propagation in a graph of an epidemic without recovering. The goal of this article is to analyze the behaviour of Epidemic with respect to several sources of randomness:

- randomness in the updating scheme,
- randomness in the initial configuration,
- randomness (or defaults) of the graph.

We will discuss theses aspects on two variants of Epidemic: on a finite square grid (in Section 2) and the analogous rule on the infinite square lattice (Section 3). The latter variant is often called *Eden model* by physicists and probabilists.

The main questions that we will discuss are

- On a finite grid: how long does it take for the spread to occupy the whole grid?
- On $\mathbb{Z}^2$: what is the *typical* shape of the spread after a large time?
- For both models: how do these behaviours depend on the different parameters of the models?

Beyond its own interest, we believe that Epidemic is a good candidate to study robustness of Cellular Automata with respect to randomness. Its behaviour is rich enough to reveal some interesting phenomena and simple enough to allow rigorous analysis. Some results stated here are all more or less *folklore*, but the statements are not so easy to find in litterature, our goal was to present them in a self-contained way. The results on Epidemic are all rigorously proved, the discussion on the Eden model is more thought as a reading-guide in the probabilistic litterature.

2 Epidemic on a finite grid

2.1 The model

Let $L \geq 3$ be an integer, we denote by Λ the square grid $L \times L$, with torical boundary conditions (*i.e* we identify Λ with $\mathbb{Z}/L\mathbb{Z} \times \mathbb{Z}/L\mathbb{Z}$). Let $n = L^2$ be the number of *cells* in Λ, n will be our scale for the asymptotics.

We endow Λ with the usual graph distance, the *ball* $B(c, r)$ with center c and radius r being the set of cells c' such that there exists a *path* of $0 \leq \ell \leq r$ neighbouring cells

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

A *configuration* σ is one of the 2^n elements of $\{\mathtt{0}, \mathtt{1}\}^\Lambda$. For $c \in \Lambda$, $\sigma_c \in \{\mathtt{0}, \mathtt{1}\}$ is the *state* of cell c in configuration σ.

For c in Λ, $\mathcal{N}(c)$ is the so-called *Von Neumann* neighbourhood of c:

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

where $+$ stands for addition modulo L. In other words $\mathcal{N}(c) = B(c, 1)$.

We now can define Epidemic as a stochastic dynamical system. Each cell in state $\mathtt{0}$ (healthy), when it is updated, turns into state $\mathtt{1}$ (infected) if one at least of its neighbours is infected. There is no recovering: a $\mathtt{1}$ remains $\mathtt{1}$ forever. Besides, updating are random and independent from each other.

More formally:

Definition 1 (Epidemic in the α-synchronous dynamics)**.** *Let $\alpha \in (0, 1)$, the α-synchronous Epidemic cellular automaton is the stochastic process $\left(\sigma^t\right)_{t\geq 0}$ with values in $\{\mathtt{0}, \mathtt{1}\}^\Lambda$ such that $\sigma^0 \in \{\mathtt{0}, \mathtt{1}\}^\Lambda$ and whose evolution is given in the following way.*

For every $t \geq 0$, given σ^t at time t, the configuration σ^{t+1} is defined as follows: each cell in Λ is updated independently with probability α (independently from the past and from the $n - 1$ other cells) ;

- *If c is updated, $\sigma_c^{t+1} = \mathtt{1}$ if and only if at least one cell in $\mathcal{N}(c)$ is in state $\mathtt{1}$ at time t;*
- *Otherwise, $\sigma_c^{t+1} = \sigma_c^t$.*

The sequence (σ^t) is then a discrete-time Markov chain with values in $\{\mathtt{0}, \mathtt{1}\}^\Lambda$. Obviously it eventually reaches one of the two fixed configurations $\mathtt{0}^\Lambda$ our $\mathtt{1}^\Lambda$. The configuration $\mathtt{0}^\Lambda$ is *isolated*:

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

From now on we the trivial case $\sigma^0 = \mathtt{0}^\Lambda$.

Definition 2 (Convergence time)**.** *For an initial configuration $\sigma^0 \neq \mathtt{0}^\Lambda$, the* convergence time *$T_n(\sigma^0)$ is the first time at which all cells are infected:*

$$T_n(\sigma^0) = \min\left\{t \geq 0 \textit{ such that } \sigma^t = \mathtt{1}^\Lambda\right\}.$$

In this section we will focus on the asymptotic behaviour of the expectation of this random variable: $\mathbb{E}\left[T_n(\sigma^0)\right]$, in *worst* and *typical* cases.

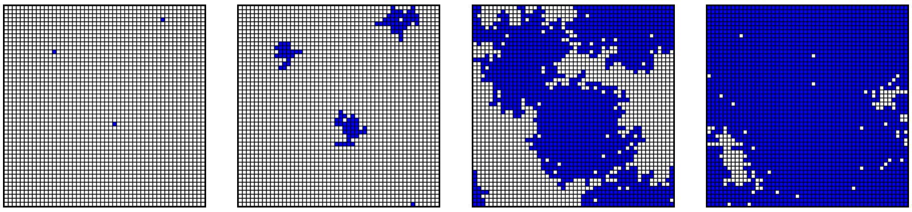

Figure 1: A simulation of α-synchronous Epidemic with $\alpha = 0.1$ and $L = 50$, σ^0 has only 3 cells in state `1`. Simulations were performed with FiatLux: `http://fiatlux.loria.fr/`.

2.2 The Worst Expected Convergence Time

We first consider the *worst* expected convergence time (WECT) for Epidemics , *i.e.* the mean convergence time when $\sigma^0 = \sigma^{\mathrm{w}}$ where σ^{w} is such that

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

Obviously such σ^{w}'s are exactly the n configurations with a single `1`.

Before stating the result, let us motivate the analysis of the WECT:

1. In dimension one, Fatès *et al.* [FMST06] have studied the WECT of the Elementary Cellular Automata with two quiescent states (see their article for the definitions). Their work revealed that the asymptotic behaviour of the WECT provides a relevant classification of 1D cellular automata. Precisely, they have shown that these rules may be classified into 5 families, according to whether the WECT is $\Theta(n \log n)$, $\Theta(n^2)$, $\Theta(n^3)$, $\Theta(n2^n)$ or infinite[1]. This approach was extended in [FG09] for a family of 2D cellular automata (in particular for Epidemic).
2. Another motivation comes from algorithmic complexity theory, since cellular automata are often thought as model in computability theory. With this point of view, it is natural to ask what happens when the system starts from the "worse" configuration.
3. Alternatively, if we think of cellular automata as (simplified!) models of physical or biological systems, studying the WECT provides us with an estimation of the maximum time needed to go back to equilibrium when a perturbation is applied.

Theorem 1 (Worst Expected Convergence Time)**.** *For Epidemics on a finite grid with n cells, for all $\alpha \in (0,1)$, if n is large enough,*

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

Remark.
- *In [FG09], a very similar result (but less precise, because of* $\log(n)$ *terms in both sides of the inequality) was proved in* asynchronous *dynamics.*
- *We believe that $\mathbb{E}\left[T_n(\sigma^{\mathrm{w}})\right]/\sqrt{n}$ converges to a constant, and more precisely that the sequence of random variables $T_n(\sigma^{\mathrm{w}})/\sqrt{n}$ converges in probability. We have not been able to prove so, and the usual tool (* subadditivity *theory, see [BS10] for instance) to deal with similar problems does not seem to work here.*

[1]We write $f_n = \Theta(g_n)$ when there exist two positive numbers c_1, c_2 such that, for n large enough, $c_1 g_n \leq f_n \leq c_2 g_n$.

Proof.
Lower bound. Let c be the only cell in state $\mathtt{1}$ in σ^{w}, and let c' be one of the cells of Λ which is at distance $\lfloor L/2 \rfloor$ from c, where $\lfloor x \rfloor$ is the integer part of x. It is enough to prove that $\mathbb{E}[\tau_{c'}] \geq \frac{L}{8\alpha}$, with

$$\tau_{c'} = \min\left\{t \geq 0,\ \sigma^t_{c'} = \mathtt{1}\right\},$$

since obviously $T_n(\sigma^{\mathrm{w}}) \geq \tau_{c'}$.

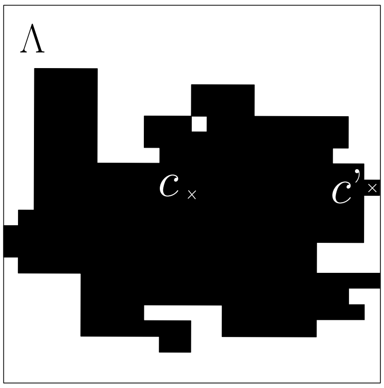

Figure 2: The configuration at time $\tau_{c'}$.

Set $k = \lfloor L/4\alpha \rfloor$, we will prove that if L is large enough

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

This proves the upper bound since then

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

We focus on (1). The proof is not very difficult but quite technical, here is the general strategy. As often in this kind of optimization problem there is a trade-off between two phenomena:

- There is a huge number of paths (we will bound this number by 3^j in the proof) going from c to c' along which successive updatings would infect cell c';
- On the other hand, if we fix such a path, it is very unlikely if k is well-chosen (this will be our Lemma 1 below) that its cells are updated in the proper order before k.

We now go into the details. Assume that $\tau_{c'} \leq k$, then there is a j with $\lfloor L/2 \rfloor \leq j \leq k$ and a path $\mathcal{P}$ made of j disjoint cells, going from c to c':

$$\mathcal{P} = \left\{c = c_0 \to c_1 \to c_2 \to \cdots \to c_j = c'\right\}$$

such that, during the k first time units, cells $c_1, \ldots, c_j$ are updated *in this order.* This gives

$$\mathbb{P}(\tau_{c'} \leq k) = \mathbb{P}\left(\bigcup_{j=\lfloor L/2 \rfloor}^{k} \ \bigcup_{\mathcal{P}, |\mathcal{P}|=j} \{c_1, c_2, \ldots, c_j \text{ are updated in this order and before } k\}\right)$$

$$\leq \sum_{j=\lfloor L/2 \rfloor}^{k} \ \sum_{\mathcal{P}, |\mathcal{P}|=j} \mathbb{P}\left(c_1, c_2, \ldots, c_j \text{ are updated in this order and before } k\right),$$

(here we used $\mathbb{P}(\cup) \leq \sum \mathbb{P}$, often called *union bound*). The second sum runs over all paths of j cells going from c to c'. Fix such a $\mathcal{P}$ and bound the last probability. Among times $\{1, 2, \dots, k\}$, there are times $t_1 < t_2 < \cdots < t_j$ such that, at time t_j, cell c_j is updated. Each updating being independent we have

$$\mathbb{P}\left(c_1, c_2, \dots, c_j \text{ are updated in this order and before } k\right) = \mathbb{P}\left(\mathrm{Binom}(k, \alpha) \geq j\right),$$

where $\mathrm{Binom}(k, \alpha)$ has the binomial distribution with parameters k, α and has expectation $k\alpha$. If $k\alpha \ll j$ then this last probability is small, the following lemma is useful, this is for instance (2.5) in [JŁR00].

Lemma 1 (Right-deviations for the binomial)**.** *For all $j \geq k\alpha$,*

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

There are less that 3^j paths of length j going from c to c' (this is a rough bound but sufficient here), we have

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

One can check that $j \mapsto 3^j \exp\left(-3\frac{(j-k\alpha)^2}{2k\alpha+j}\right)$ is non-increasing for $L/2 \leq j \leq k$ (recall $k\alpha \approx L/8$) and thus, skipping the integer parts in order to lighten notations,

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

and therefore is less than 1/2 if L is large (depending on α). We have proved (1).

Upper bound. We will prove that for L large enough and $k \geq 2L/\alpha$

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

This yields the upper bound since

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

for large L.

Let us prove (2). First, for each $c' \neq c$, we choose (in a non-random way) a path $\mathcal{P}_{c'}$ among all shortest paths $c \to c'$: $\mathcal{P}_{c'}$ can be written

$$\mathcal{P}_{c'} = \left\{c = c_0 \to c_1 \to c_2 \to \cdots \to c_j = c'\right\}$$

where $j \leq L/2$ is the distance between c and c'.

If $T_n(\sigma^{\mathrm{w}}) > k$ then there is c' which is still in state $\mathtt{0}$ at time k. In particular there is $j \leq L/2$ and a cell c' at distance j such that cells $c_1, c_2, \ldots, c_j$ of its associated path $\mathcal{P}_{c'}$ have not been updated in this order before time k :

$$\mathbb{P}(T_n(\sigma^{\mathrm{w}}) > k) = \mathbb{P}\left(\bigcup_{c' \in \Lambda} \{c_1, c_2, \ldots, c_j \text{ are not updated in this order and before } k\}\right)$$
$$\leq \mathrm{card}(\Lambda) \max_{c' \in \Lambda} \mathbb{P}\left(c_1, c_2, \ldots, c_j \text{ are updated in this order and before } k\right).$$

and by the same argument as for the lower bound

$$\mathbb{P}\left(c_1, c_2, \ldots, c_j \text{ are updated in this order and before } k\right) = \mathbb{P}\left(\mathrm{Binom}(k, \alpha) < j\right),$$

now we need the following bound, this is (2.6) in [JŁR00].

Lemma 2 (Left-deviations for the binomial)**.** *For all $j \leq k\alpha$,*

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

For all $j \leq L/2$, we have $j \leq k\alpha$ (recall $k \geq 2L/\alpha$) and we can apply the lemma :

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

we have proved (2). □

2.3 Typical convergence time

We now estimate the *typical* expected convergence time, when σ^0 is drawn uniformly at random in $\{\mathtt{0}, \mathtt{1}\}^{\Lambda}$.

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

As expected, Typ_n is much smaller than in the worst case.

Theorem 2 (Typical expected convergence time)**.** *For n large enough,*

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

Proof. We closely follow ([Ger08], Chap.2).
Lower bound. The number of cells in state $\mathtt{0}$ in σ^0, which is a Binomial $(n, 1/2)$, is larger than $n/2$ with more than 50% chance. For such σ^0, the convergence takes more time than the time needed to update all these cells at least once. Thus

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

where $\mathcal{G}_i$ are i.i.d geometric random variables with mean $1/\alpha$. For large k we have (see [SR90] for instance)

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

and then, for large enough n, $\mathbb{E}\left[T_n(\sigma^0)\right] \geq \frac{2}{3\alpha}\log(n/2) \geq \frac{1}{2\alpha}\log(n)$ when σ^0 has more than $n/2$ cells in state $\texttt{0}$. Now,

$$\begin{aligned}\mathrm{Typ}_n = \frac{1}{2^n}\sum_{\sigma^0\in\{\texttt{0},\texttt{1}\}^\Lambda}\mathbb{E}\left[T_n(\sigma^0)\right] &\geq \frac{1}{2^n}\sum_{\substack{\sigma^0 \text{ with}\\ \text{more than } n/2\ \texttt{0}\text{'s}}}\mathbb{E}\left[T_n(\sigma^0)\right]\\ &\geq \frac{1}{2^n}\mathrm{card}\left\{\sigma^0 \text{ with more than } n/2\ \texttt{0}\text{'s }\right\}\frac{1}{2\alpha}\log(n)\\ &\geq \frac{1}{2^n}\frac{2^n}{2}\frac{1}{2\alpha}\log(n),\end{aligned}$$

hence the lower bound.

Upper bound. The first step is to show that with high probability there is no ball of radius $3\sqrt{\log(L)}$ which is full of $\texttt{0}$'s in the initial configuration σ^0.

Precisely, set $a(L) = \lfloor 3\sqrt{\log L}\rfloor$, introduce the vent

$$\begin{aligned}A &= \bigcup_{c\in\Lambda}\{\text{ the ball of center } c \text{ and radius } a(L) \text{ is full of } \texttt{0}\text{'s at time } 0.\ \}\\ &= \bigcup_{c\in\Lambda}\left\{\forall c' \in B(c,a(L)), \sigma^0_{c'} = \texttt{0}\right\}.\end{aligned}$$

By the union bound,

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

Each ball $B(c, a(L))$ contains more than $2a(L)^2$ cells, it is full of $\texttt{0}$'s with probability less than $(1/2)^{2a(L)^2}$. We get (we skip integer parts once more)

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

(for large n). Let us write

$$\begin{aligned}\frac{1}{2^n}\sum_{\sigma^0}\mathbb{E}[T(\sigma^0)] &= \frac{1}{2^n}\sum_{\substack{\sigma^0 \text{ such that}\\ A \text{ is true}}}\mathbb{E}[T(\sigma^0)] + \frac{1}{2^n}\sum_{\substack{\sigma^0 \text{ such that}\\ A \text{ is false}}}\mathbb{E}[T(\sigma^0)],\\ &\leq \mathbb{P}(A)\max_{\substack{\sigma^0 \text{ such that}\\ A \text{ is true}}}\mathbb{E}[T(\sigma^0)] + \mathbb{P}(\text{not } A)\max_{\substack{\sigma^0 \text{ such that}\\ A \text{ is false}}}\mathbb{E}[T(\sigma^0)],\\ &\leq 1/\sqrt{n}\max_{\substack{\sigma^0 \text{ such that}\\ A \text{ is true}}}\mathbb{E}[T(\sigma^0)] + 1\times\max_{\substack{\sigma^0 \text{ such that}\\ A \text{ is false}}}\mathbb{E}[T(\sigma^0)]. \qquad (4)\end{aligned}$$

We bound both terms :

- $\max_{\substack{\sigma^0 \text{ such that}\\ A \text{ is true}}}\mathbb{E}[T(\sigma^0)]$ is a $\mathcal{O}(\sqrt{n})$ by the upper bound of the WECT;
- If A is false then every $\texttt{0}$ is less than $a(L)$ away from a $\texttt{1}$. The configuration has thus converged before the time at which each cell has been updated $a(L)$ times. By (3), it takes less than $2\log(n)/\alpha$ in average to update the n cells at least once. Then

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

And (4) yields

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

for large enough n. □

Remark. *It is in fact possible (but tedious) using* (2) *to improve the upper bound from* $\mathcal{O}(\log(n)^{3/2})$ *to* $\mathcal{O}(\log(n))$*. The idea is that a ball of* $\texttt{0}$*'s of radius* $\log(L)$ *is filled with* $\texttt{1}$*'s in less than* $\mathcal{O}(\log(L))$ *time steps (with high probability).*

Discussion

Our aim here was to present with self-contained proofs some quantitative results that did not seem to appear in litterature. It is worth noting that many natural questions still remain open: in particular the order of magnitude of the variance of Variance($T_n(\sigma^{\mathrm{w}})$ is still unknown.

3 Epidemics in $\mathbb{Z}^2$: the Eden model

We now consider the analogous of Epidemics on the infinite lattice $\mathbb{Z}^2$, it is usually refered to as the *Eden model* [Ede61] or *Richardson's model* [Ric73]. Let $\alpha > 0$, we consider the stochastic process $\left(\sigma^t\right)_{t\geq 0}$ with values in $\{\mathtt{0},\mathtt{1}\}^{\mathbb{Z}^2}$ defined as follows:

- $\sigma^0_{\mathbf{0}} = \mathtt{1}$ and $\sigma^0_c = 0$ for $c \neq \mathbf{0}$, where $\mathbf{0}$ is the origin of $\mathbb{Z}^2$.
- At time $t+1$, each $\mathtt{0}$ that has a neighbour in state $\mathtt{1}$ in σ^t turns into $\mathtt{1}$ with probability α, independently from the past and the other cells.

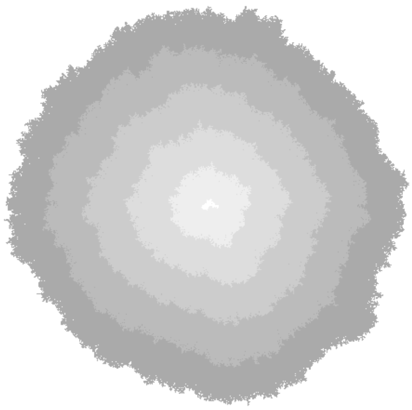

Figure 3: The Eden model with $\alpha = 0.02$, at different times up to 10^6.

We are interested in the asymptotic behaviour $\left(\sigma^t\right)_{t\geq 0}$. We observe on simulations (Fig. 3) that the component of $\mathtt{1}$'s seems to grow like a particular shape. Richardson [Ric73] has proved that it has indeed a *limiting shape*, in the following sense.

Theorem 3 (Limiting shape theorem for the Eden model)**.** *Let B_t be the set of cells in state* $\mathtt{1}$ *at time t. There is a non-random set $B_\star \subset \mathbb{R}^2$ which is compact, convex and non void such that for every $\varepsilon > 0$,*

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

This result was further improved by [CD81] into an almost-sure convergence.

3.1 The link with First-passage percolation

The Eden model is a dynamical model of growth process but in fact it can be seen as a static model. To do so, set as before, for $c' \in \mathbb{Z}^2$,

$$\tau_{c'} = \min\left\{t \geq 0,\ \sigma^t_{c'} = \mathtt{1}\right\}.$$

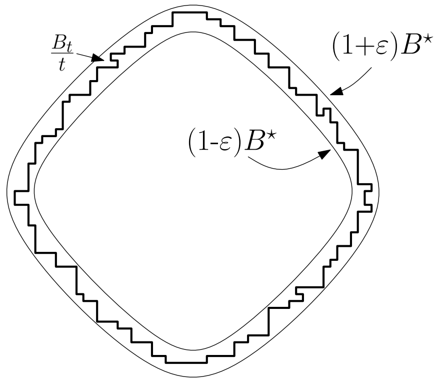

Figure 4: What Theorem 3 says, for large t.

As in the previous section, $\tau_{c'} \leq k$ if and only if there is a path $\mathcal{P}$ of $j \leq k$ neighbouring cells going from $\mathbf{0}$ to c'

$$\mathcal{P} = \left\{\mathbf{0} \to c_1 \to c_2 \to \cdots \to c_j = c'\right\}$$

such that successive updating along $\mathcal{P}$ turn c' into `1`. Then one can show that, for each fixed c',

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

where $\left\{g_c,\ c \in \mathbb{Z}^2\right\}$ is a family of i.i.d. geometric random variables with mean $1/\alpha$ and $\stackrel{\text{(d)}}{=}$ means "are equal in distribution". Here g_{c_i} is the time needed to update c_i, once one of its neighbours is `1`.

Thus, a way to construct $\tau_{c'}$ is to draw for each cell in $\mathbb{Z}^2$ some independent random times g_{c_i}, and then $\tau_{c'}$ is the sum of these times over the path $\mathbf{0} \to c'$ such that the sum is minimal. This model is known as *First-passage percolation* (FPP) and has been studied for the first time by Hammersley and Welsh [HW65]. We refer to [BS10] for a modern introduction to FPP and its connections with growth models.

The full connection between $\tau_{c'}$'s in the Eden model and first-passage percolation can be written as follows:

Proposition 1 (Eden model is FPP). *Let $\left\{g_c,\ c \in \mathbb{Z}^2\right\}$ be a family of i.i.d. geometric random variables with mean $1/\alpha$. Then*

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

where the min *is taken over all paths going from* $\mathbf{0}$ *to* c*:*

$$\mathcal{P} = \left(\mathbf{0} = c_0 \to c_1 \to \dots c_{|\mathcal{P}|} = c\right)$$

and $|\mathcal{P}|$ *is the number of cells of* $\mathcal{P}$.

The connection between Eden model and FPP is usually attributed to Richardson, even if first-passage percolation is not clearly mentioned in [Ric73]. Surprisingly enough, it seems that there is no rigorous proof of Proposition 1 available in litterature. It is often considered as folklore, but it is not so easy to write down a complete proof (the main difficulty is to establish the equality for the whole family of τ_c's and not only for a fixed c).

3.2 Influence of the lattice

The Eden model being a toy model for propagation, one might wonder if the properties proved in this particular model are *robust* under various perturbations of the lattice. This question is not clearly understood.

Since [Ric73], the following conjecture is attributed to Eden:

Conjecture (Eden conjecture)**.** *For the Eden model in continuous time, the set of cells in state* `1` *is asymptotically shaped as a disc:* $B_\star$ *is a euclidian ball.*

The Eden model in continuous time is defined as before, except that the updating times are exponentially distributed. There are some simulations in [Ede61], but the conclusions are not so clear:

> *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.*

In 1984, H.Kesten had the intuition that this conjecture should be false, at least in high dimensions, for geometrical reasons. He disproved the conjecture for $d > 600000$ (see [Kes86]), since then [Dha88] and [CEG11] improved the result up to dimension $d > 35$.

Of course we are far from a physical or biological model, yet this result says something interesting: the asymptotic properties of the Eden model strongly depend on the lattice on which it is constructed. We know that this is not the case for the position of a standard random walk on a regular lattice, whose asymptotic law does not depend on the lattice and is the normal distribution. It seems that the Eden model is sensible to the microscopic structure of the lattice.

3.3 Eden model and random defaults

What happens in Theorem 3 if some proportion of cells is *immunised* against infection? How does it change the growth of B_t?

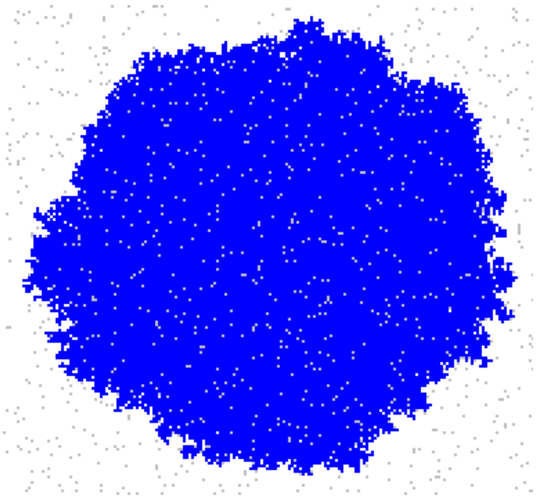

Figure 5: The Eden model with immunised cells (gray).

Assume that each cell is originally immunised independently with probability p. Obviously B_t cannot grow infinitely if the following event E occurs:

- either **0** is immunised,

- or there is a path of immunised cells surrounding $\mathbf{0}$.

Of course $\mathbb{P}(E) \geq \mathbb{P}(\mathbf{0} \text{ is immunised}) \geq p > 0$ but we can prove (see [Gri99] for instance) that if p is small enough then $\mathbb{P}(E) < 1$. In the case where E does not occur, it is possible than B_t grow infinitely and the growth is linear, as in the initial model. This has been proved rigorously by [GM04], we need a few notations to state the result.

Let $\vec{n}$ be the cell of coordinates $(n, 0)$. Let $\mathcal{A}$ be the (random) set of integers n such that there is a path of non-immunised cells going from the origin to $\vec{n}$. If E does not occur then there are infinitely many cells than can be infected and if $n \in \mathcal{A}$, the first time $\tau_{\vec{n}}$ at which $\vec{n}$ is in B_t is finite (almost-surely).

Theorem 4 (Linear growth for the Eden model with defaults)**.** *There exists $\mu > 0$ such that, if E does not occur,*

$$\lim_{\substack{n\to+\infty,\\ n\in\mathcal{A}}} \frac{\tau_{\vec{n}}}{n} = \mu \quad \textit{almost surely,}$$

where the limit is taken along the random subsequence $\{n \in \mathcal{A}\}$.

Discussion

Much is known now about the quantitative properties of the Eden model. In particular, many efforts have been made in order to understand the dependence of the limiting shape with respect to the different parameters of the model: dependence with respect to α ([CK81],[Mar02]) and to p [BEGG14].

We have just tried here to present a few results for non-probabilists, we refer to [BS10] for a nice and modern introduction to this topic. In particular, it is discussed of the variant in which there is a *competition* between two epidemics.

References

[BEGG14] A.-L. Basdevant, N. Enriquez, L. Gerin, and J.-B. Gouéré. The shape of large balls in highly supercritical percolation. *Electronic Journal of Probability*, 19:1–14, 2014.

[BS10] N.D. Blair-Stahn. First passage percolation and competition models. `arXiv: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.

[CK81] J.T. Cox and H. Kesten. On the continuity of the time constant of first-passage parcolation. *Journal of Applied Probability*, 18:809–819, 1981.

[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.

[JŁR00] S. Janson, T. Łuczak, and A.J. Rucinski. *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.

[Mar02] R. Marchand. Strict inequalities for the time constant in first passage percolation. *Annals of Applied Probability*, 12:1001–1038, 2002.

[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.