Electronic Journal of Probability

Electron. J. Probab. **18** (2013), no. 51, 1–22.
ISSN: 1083-6489 DOI: 10.1214/EJP.v18-2325

Density classification on infinite lattices and trees*

Ana Bušić† Nazim Fatès‡ Jean Mairesse§ Irène Marcovici§

Abstract

Consider an infinite graph with nodes initially labeled by independent Bernoulli random variables of parameter p. We address the density classification problem, that is, we want to design a (probabilistic or deterministic) cellular automaton or a finite-range interacting particle system that evolves on this graph and decides whether p is smaller or larger than $1/2$. Precisely, the trajectories should converge to the uniform configuration with only 0's if $p < 1/2$, and only 1's if $p > 1/2$. We present solutions to the problem on the regular grids of dimension d, for any $d \geq 2$, and on the regular infinite trees. For the bi-infinite line, we propose some candidates that we back up with numerical simulations.

Keywords: Cellular automata ; interacting particle systems ; density classification.
AMS MSC 2010: Primary 60K35 ; 68Q80, Secondary 37B15 ; 60J05.
Submitted to EJP on September 18, 2012, final version accepted on April 10, 2013.
Supersedes arXiv:1111.4582.
Supersedes HAL:hal-00643468.

1 Introduction

Consider a finite or a countably infinite set of cells, which are spatially arranged according to a group structure G. The *density classification problem* consists in deciding, in a decentralised way, if an initial configuration on G contains more 0's or more 1's. More precisely, the goal is to design a deterministic or probabilistic dynamical system that evolves in the configuration space $\{0,1\}^G$ with a local and homogeneous updating rule and whose trajectories converge to 0^G or to 1^G if the initial configuration contains more 0's or more 1's, respectively. To attack the problem, two natural instantiations of dynamical systems are considered, one with synchronous updates of the cells, and one with asynchronous updates. In the first case, time is discrete, all cells are updated at

*This work was partially supported by the ANR project MAGNUM (ANR-2010-BLAN-0204).
†Inria and ENS, France.
E-mail: Ana.Busic@inria.fr
‡Inria Nancy – Grand Est, LORIA, Nancy Université, CNRS, France.
E-mail: Nazim.Fates@loria.fr
§CNRS, UMR 7089, LIAFA, Univ Paris Diderot, Sorbonne Paris Cité, France.
E-mail: Jean.Mairesse@liafa.univ-paris-diderot.fr
E-mail: Irene.Marcovici@liafa.univ-paris-diderot.fr (corresponding author)

each time step, and the model is known as a *Probabilistic Cellular Automaton (PCA)* [5]. A *Cellular Automaton (CA)* is a PCA in which the updating rule is deterministic. In the second case, time is continuous, cells are updated at random instants, at most one cell is updated at any given time, and the model is known as a (finite range) *Interacting Particle System (IPS)* [20].

The general spirit of the problem is that of distributed computing: gathering a global information by exchanging only local information. The difficulty is twofold: first, it is impossible to centralise the information (cells are indistinguishable); second, it is impossible to use classical counting techniques (cells contain only binary information).

The density classification problem was originally introduced for synchronous models and rings of finite size ($G = \mathbb{Z}/n\mathbb{Z}$) [21]. After experimentally observing that finding good rules to perform this task was difficult, it was shown that perfect classification with CA is impossible, that is, there exists no given CA that solves the density classification problem for all values of n [19]. However, this result did not stop the quest for the best – although imperfect – models as nothing was known about how well CA could perform. The use of PCA opened a new path [8, 24] and it was shown that there exist PCA that can classify with an arbitrary precision [6]. In the present paper, we complement in Prop. 2.1 the results from Ref. [19, 6] by showing that there exists no PCA that perfectly solves the density classification problem for all values of n.

The challenge is now to extend the research to infinite groups whose Cayley graphs are lattices or regular trees. First, we need to specify the meaning of "having more 0's or more 1's" in this context. Consider a random configuration on $\{0,1\}^G$ obtained by assigning independently to each cell a value 1 with probability p and a value 0 with probability $1-p$. We say that a model "classifies the density" if the trajectories converge weakly to 1^G for $p > 1/2$, and to 0^G for $p < 1/2$. A couple of conjectures and negative results exist in the literature. Density classification on $\mathbb{Z}^d$ is considered in Ref. [3] under the name of "bifurcation". The authors study variants of the famous voter model IPS [20, Ch. V] and they propose two instances that are conjectured to bifurcate. The density classification question has also been addressed for the Glauber dynamics associated to the Ising model at temperature 0, both for lattices and for trees [7, 15, 16]. The Glauber dynamics defines an IPS or PCA having 0^G and 1^G as invariant measures. Depending on the cases, there is either a proof that the Glauber dynamics does not classify the density, or a conjecture that it does with a proof only for densities sufficiently close to 0 or 1.

The density classification problem has been approached with different perspectives on finite and infinite groups, as emphasised by the results collected above. For finite groups, the problem is studied *per se*, as a benchmark for understanding the power and limitations of cellular automata as a computational model. The community involved is rather on the computer science side. For infinite groups, the goal is to understand the dynamics of specific models that are relevant in statistical mechanics. The community involved is rather on the theoretical physics and probability theory side.

The aim of the present paper is to investigate how to generalise the finite group approach to the infinite group case. We want to build models of PCA and IPS, as simple as possible, that correct random noise in the initial configuration, even if the density of errors is close to $1/2$. We consider the groups $\mathbb{Z}^d$, whose Cayley graphs are lattices (Sec. 3), and the free groups, whose Cayley graphs are infinite regular trees (Sec. 4). In all cases, except for $\mathbb{Z}$, we obtain both PCA and IPS models that classify the density. To the best of our knowledge, they constitute the first known such examples. The case of $\mathbb{Z}$ is more complicated and still open. We provide some potential candidates for density classification together with simulation experiments (Sec. 5).

A preliminary version of the paper has been presented at the Conference LATIN'2012 [2].

2 Defining the density classification problem

Let $(G, \cdot)$ be a finite or countable set of *cells* equipped with a group structure. Set $\mathcal{A} = \{0, 1\}$, the *alphabet*, and $X = \mathcal{A}^G$, the set of *configurations*. For $x \in X$ and $u \in \{0, 1\}$, denote by $|x|_u$ the number of occurrences of u in x.

Let us equip $X = \mathcal{A}^G$ with the product topology. A *cylinder* is a subset of X having the form $[y] = \{x \in X \mid \forall k \in K, x_k = y_k\}$ for a given finite $K \subset G$ and a given $y = (y_k)_{k \in K} \in \mathcal{A}^K$. Let $\mathcal{M}(X)$ be the set of probability measures on X for the σ-algebra generated by all cylinder sets, which corresponds to the Borelian σ-algebra.

2.1 PCA and IPS

Given a finite set $\mathcal{N} \subset G$, a *transition function* of *neighbourhood* $\mathcal{N}$ is a function $f : \mathcal{A}^{\mathcal{N}} \to \mathcal{A}$. The *cellular automaton (CA)* F of transition function f is the function $F : X \to X$ defined by:

$$\forall x \in X, \forall g \in G, \quad F(x)_g = f((x_{g \cdot v})_{v \in \mathcal{N}}).$$

When the group G is $\mathbb{Z}^d$ or $\mathbb{Z}_n = \mathbb{Z}/n\mathbb{Z}$, we denote as usual the law of G by the sign $+$, so that the definition can be written: $\forall x \in X, \forall k \in \mathbb{Z}^d$ (resp. $\mathbb{Z}_n$), $F(x)_k = f((x_{k+v})_{v \in \mathcal{N}})$.

Probabilistic cellular automata (PCA) are an extension of classical CA: the transition function is now a function $\varphi : \mathcal{A}^{\mathcal{N}} \to \mathcal{M}(\mathcal{A})$, where $\mathcal{M}(\mathcal{A})$ denotes the set of probability measures on $\mathcal{A}$. At each time step, the cells are updated synchronously and independently, according to a distribution depending on a finite neighbourhood [5]. Formally, the PCA associated with φ is the function $F : \mathcal{M}(X) \to \mathcal{M}(X),\ \mu \mapsto \mu F$, defined on cylinders by: for $K \subset G$, $y = (y_k)_{k \in K}$,

$$\mu F([y]) = \sum_{x \in \mathcal{A}^{K \cdot \mathcal{N}}} \mu([x]) \prod_{u \in K} \varphi((x_{u \cdot v})_{v \in \mathcal{N}})(y_u)\,,$$

where $K \cdot \mathcal{N} = \{g \in G \mid g = g_1 \cdot g_2, g_1 \in K, g_2 \in \mathcal{N}\}$.

The analog of PCA in continuous time are *(finite-range) interacting particle systems (IPS)* [20]. IPS are characterised by a finite neighbourhood $\mathcal{N} \subset G$, and a transition function $f : \mathcal{A}^{\mathcal{N}} \to \mathcal{A}$ (or $\varphi : \mathcal{A}^{\mathcal{N}} \to \mathcal{M}(\mathcal{A})$). We attach random and independent clocks to the cells of G. For a given cell, the instants of $\mathbb{R}_+$ at which the clock rings form a Poisson process of parameter 1. Let x^t be the configuration at time $t \geq 0$ of the process. If the clock at cell g rings at instant t, the state of the cell g is updated into $f((x^t_{g \cdot v})_{v \in \mathcal{N}})$ (or according to the probability measure $\varphi((x^t_{g \cdot v})_{v \in \mathcal{N}})$). This defines a transition semigroup $F = (F^t)_{t \in \mathbb{R}_+}$, with $F^t : \mathcal{M}(X) \to \mathcal{M}(X)$. If the initial measure is μ, the distribution of the process at time t is given by μF^t.

In a PCA, all cells are updated at each time step, in a "synchronous" way. On the other hand, for an IPS, the updating is "fully asynchronous". Indeed, the probability of having two clocks ringing at the same instant is 0.

Observe that PCA are discrete-time Markov chains, while IPS are continuous-time Markov processes. A measure μ is said to be an *invariant measure* of a process F or $(F_t)_{t \in \mathbb{R}_+}$, if $\mu F = \mu$ or $\mu F_t = \mu$ for all $t \in \mathbb{R}_+$, respectively.

In the PCA case, consider a realization $(X^n)_{n \in \mathbb{N}}$ of the Markov chain. As an extension of the usual notion of space-time diagrams in the deterministic context, the sequence

$(X^n)_{n\in\mathbb{N}} = (X^n_k)_{k\in\mathbb{Z},n\in\mathbb{N}}$ is called a *space-time diagram* (the space-coordinate is k, and the time-coordinate is n). The space-time diagram of an IPS is defined similarly, in continuous-time.

2.2 The density classification problem on $\mathbb{Z}_n$

The density classification problem was originally stated as follows: find a finite neighbourhood $\mathcal{N} \subset \mathbb{Z}$ and a transition function $f : \mathcal{A}^{\mathcal{N}} \to \mathcal{A}$ such that for any integer $n \geq 1$ and any configuration $x \in \mathcal{A}^{\mathbb{Z}_n}$, when applying the CA F of transition function f to x, the sequence of iterates $(F^k(x))_{k\geq 0}$ reaches the fixed point $\mathbf{0} = 0^n$ if $|x|_0 > |x|_1$ and the fixed point $\mathbf{1} = 1^n$ if $|x|_1 > |x|_0$. The problem can be extended to PCA by requiring the measure $(\delta_x F^t)_{t\geq 0}$ to converge to $\delta_{\mathbf{0}}$, resp. $\delta_{\mathbf{1}}$. (Or equivalently, by requiring the space-time diagram to converge almost surely to $\mathbf{0}$, resp. $\mathbf{1}$.)

Land and Belew have proved that there exists no CA that perfectly performs this density classification task for all values of n [19]. We now prove that this negative result can be extended to PCA. It provides at the same time a new proof for CA as a particular case.

Denote by δ_x the probability measure corresponding to a Dirac distribution centered on x.

Proposition 2.1. *There exists no PCA or IPS that solves perfectly the density classification problem on $\mathbb{Z}_n$, that is, for any integer $n \geq 1$, and for any configuration $x \in \mathcal{A}^{\mathbb{Z}_n}$, $(\delta_x F^t)_{t\geq 0}$ converges to $\delta_{\mathbf{0}}$ if $|x|_0 > n/2$ and to $\delta_{\mathbf{1}}$ if $|x|_1 > n/2$.*

Proof. We carry out the proof for PCA. For IPS, the argument is similar and even simpler. Let us assume that F is a PCA that solves perfectly the density classification problem on $\mathbb{Z}_n$. Let $\mathcal{N}$ be the neighbourhood of F, and let ℓ be such that $\mathcal{N} \subset [\![-\ell+1, \ell-1]\!]$. We will prove that for any $x \in \mathcal{A}^{\mathbb{Z}_n}$ (with $n \geq 2\ell$), the number of occurrences of 1 after application of F to x is almost surely equal to $|x|_1$. Let us assume that it is not the case. Then, we have:

$$\exists x, y \in \mathcal{A}^{\mathbb{Z}_n},\ |x|_1 \neq |y|_1, \qquad \delta_x F(y) > 0\,. \tag{2.1}$$

Assume for instance that $|y|_1 > |x|_1$ (the case $|y|_1 < |x|_1$ is treated similarly). We first assume that $|x|_1 = a > n/2$. We will construct a configuration z of density smaller than $1/2$, from which there is a positive probability to reach a configuration w of density larger than $1/2$. For integers $k \geq 2, m \geq 2\ell$, let us consider the configuration $z = x^k 0^m \in \mathcal{A}^{\mathbb{Z}_{kn+m}}$. We have $|z|_1 = ka$. Let y_s be the suffix of length $n-\ell$ of y, and let y_p be the prefix of length $n-\ell$ of y. By applying Eq. 2.1, it follows that:

$$\exists u, v, u', v' \in \mathcal{A}^{\ell}, \qquad \delta_z F(u y_s y^{k-2} y_p v u' 0^{m-2\ell} v') > 0\,.$$

Set $w = u y_s y^{k-2} y_p v u' 0^{m-2\ell} v'$.

$$z = \overbrace{\underset{\leftarrow\ \ kn\ \ \rightarrow}{x\ x\ \ldots\ x}}^{k}\ \overbrace{\underset{\leftarrow\ \ m\ \ \rightarrow}{0\ 0\ \ldots\ 0}}^{m}$$

$$w = \underset{\ell}{u}\ \ \underset{n-\ell}{y_s}\ \ \overbrace{\underset{\leftarrow\ (k-2)n\ \rightarrow}{y\ y\ \ldots\ y}}^{k-2}\ \ \underset{n-\ell}{y_p}\ \ \underset{\ell}{v}\ \ \underset{\ell}{u'}\ \ \overbrace{\underset{\leftarrow\ (m-2\ell)\ \rightarrow}{0\ 0\ \ldots\ 0}}^{m-2\ell}\ \ \underset{\ell}{v'}$$

We have $|w|_1 \geq k|y|_1 - 2\ell \geq k(a+1) - 2\ell$.

For large enough m, if we set k to be the largest integer such that $k(a-n/2) < m/2$ (implying that $(k+1)(a-n/2) \geq m/2$, so that $ka \geq (kn+m)/2 + n/2 - a$), we have:

$$|z|_1 = ka < \frac{kn+m}{2}, \qquad |w|_1 \geq k(a+1) - 2\ell \geq \frac{kn+m}{2} + \frac{n}{2} - a + k - 2\ell > \frac{kn+m}{2}\ ,$$

the last inequality coming from the fact that for large enough m, $k > a + 2\ell$. So, with a positive probability, we can reach a configuration with more ones than zeros starting from a configuration with more zeros than ones. Since F classifies the density with probability 1, the new configuration can be considered as an initial condition that needs to be classified and will thus almost surely evolve to the fixed point $\mathbf{1}$, that is, a bad classification will occur, which contradicts our hypothesis.

The case $|x|_1 < n/2$ can be handled by swapping the roles of 0 and 1.

We have proved that for any $x \in \mathcal{A}^{\mathbb{Z}_n}$ (with $n \geq \ell$), the number of occurrences of ones after application of F to x is almost surely equal to $|x|_1$. This is in contradiction with the fact that F classifies the density. □

The proof can be adapted to larger dimensions and we obtain the following.

Proposition 2.2. *For any $d \geq 1$, there is no d-dimensional PCA or IPS such that for any integers $n_1, \ldots, n_d \geq 1$, and for any configuration $x \in \mathcal{A}^{\mathbb{Z}_{n_1} \times \cdots \times \mathbb{Z}_{n_d}}$, $(\delta_x F^t)_{t \geq 0}$ converges to $\delta_{\mathbf{0}}$ if $|x|_0 > (n_1 \ldots n_d)/2$ and to $\delta_{\mathbf{1}}$ if $|x|_1 > (n_1 \ldots n_d)/2$.*

2.3 The density classification problem on infinite groups

Let us define formally the density classification problem on infinite groups.

We denote by μ_p the Bernoulli measure of parameter p, that is, the product measure of density p on $X = \mathcal{A}^G$. A realisation of μ_p is obtained by assigning independently to each element of G a label 1 with probability p and a label 0 with probability $1 - p$. Set $\mathbf{0} = 0^G$ and $\mathbf{1} = 1^G$.

The *density classification problem* consists in finding a PCA or an IPS F, such that:

$$\begin{cases} p < 1/2 \implies \mu_p F^t \xrightarrow[t \to \infty]{w} \delta_{\mathbf{0}} \\ p > 1/2 \implies \mu_p F^t \xrightarrow[t \to \infty]{w} \delta_{\mathbf{1}} \end{cases} . \tag{2.2}$$

The notation $\xrightarrow{w}$ stands for the weak convergence of measures. In our case, the interpretation of this convergence is that for any *finite* subset $K \subset G$, the probability that all the cells of K are labelled by 0 (resp. by 1) tends to 1 if $p < 1/2$ (resp. if $p > 1/2$). Or, equivalently, that for any single cell, the probability that it is labelled by 0 (resp. by 1) tends to 1 if $p < 1/2$ (resp. if $p > 1/2$).

From subgroups to groups. Next result will be used several times.

Proposition 2.3. *Let H be a subgroup of G, and let F_H be a process (PCA or IPS) of neighbourhood $\mathcal{N}$ and transition function f that classifies the density on $\mathcal{A}^H$. We denote by F_G the process on $\mathcal{A}^G$ having the same neighbourhood $\mathcal{N}$ and the same transition function f. Then, F_G classifies the density on $\mathcal{A}^G$.*

Proof. Since H is a subgroup, the group G is partitioned into a union of classes $g_1H, g_2H, \ldots$ We have $\mathcal{N} \subset H$, so that if an element $g \in G$ is in some class g_iH, then for any $v \in \mathcal{N}$, the element $g \cdot v$ is also in g_iH. Since F_H classifies the density, on each class g_iH, the process F_G satisfies Eq. 2.2. Thus for any cell of G, the probability that it is labelled by 0 (resp. by 1) tends to 1 if $p < 1/2$ (resp. if $p > 1/2$). □

3 Classifying the density on $\mathbb{Z}^d, d \geq 2$

According to Prop. 2.3, given a process that classifies the density on $\mathbb{Z}^2$, we can design a new process that classifies on $\mathbb{Z}^d$ for $d > 2$. Below, we concentrate on $\mathbb{Z}^2$.

To classify the density on $\mathbb{Z}^2$, a first natural idea is to apply the majority rule on a cell and its four direct neighbours. Unfortunately, this does not work, neither in the CA nor in the IPS version. Indeed, a 2×2 square of four cells in state 1 (resp. 0) remains in state 1 (resp. 0) forever. For $p \in (0,1)$, monochromatic elementary squares of both colors appear almost surely in the initial configuration which makes the convergence to $\mathbf{0}$ or $\mathbf{1}$ impossible. We prove more generally that on $\mathbb{Z}^d$, the majority rule over a symmetric neighborhood that contains the cell itself has a finite stable pattern (Fig. 1 represents two examples on $\mathbb{Z}^2$). Classification of the density is thus impossible. We recover the "forbidden symmetry" of Pippenger [23].

Lemma 3.1. *Let us consider a set $\mathcal{N} = \{e_0, e_1, \ldots, e_n, -e_1, \ldots, -e_n\}$ of $(2n+1)$ different elements of $\mathbb{Z}^d$, with $e_0 = (0, \ldots, 0)$. If the cells of the set $\mathcal{D} = \{\sum_{i \in S} e_i | S \subset \{0, \ldots, n\}\}$ are initially in the same state, then they remain in that same state when iterating the majority CA or IPS of neighborhood $\mathcal{N}$.*

Proof. Let us fix any subset S of $\{0, \ldots, n\}$, and consider the cell $c = \sum_{i \in S} e_i$. We want to prove that c has at least $n+1$ neighbors which belong to $\mathcal{D}$. First the cell c is in its own neighborhood. For $j \in S$, the cell $c - e_j = \sum_{i \in S \setminus \{j\}} e_i$ belongs to $\mathcal{D}$, and for $j \in \{1, \ldots, n\} \setminus S$, the cell $c + e_j = \sum_{i \in S \cup \{j\}} e_i$ belongs to $\mathcal{D}$. Therefore c has at least $n+1$ neighbors in $\mathcal{D}$. If all the cells of $\mathcal{D}$ are in the same state, when applying the majority rule, this state is preserved. □

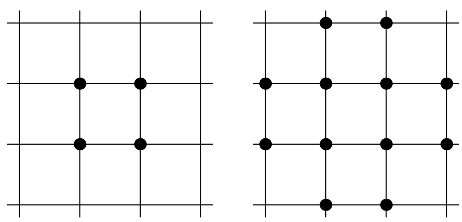


Figure 1: Stable patterns obtained respectively for the von Neumann neighborhood (the cell and its four nearest neighbors) and the Moore neighborhood (the cell and its eight surrounding neighbors).

On $\mathbb{Z}^2$, another idea is to apply the majority rule on the four nearest neighbours (excluding the cell itself) and to choose uniformly the new state of the cell in case of equality. In the IPS setting, this process is known as the Glauber dynamics associated to the Ising model. It has been conjectured to classify the density, but the result has been proved only for values of p that are sufficiently close to 0 or 1 [7].

To overcome the difficulty, we consider the majority CA but on the asymmetric neighborhood $\mathcal{N} = \{(0,0), (0,1), (1,0)\}$. This CA, known as Toom's rule [5, 25], has been introduced in connection with the positive rates problem, see Sec. 3.3. Here we prove that Toom's CA classifies the density on $\mathbb{Z}^2$. Our proof relies on the properties of the percolation clusters on the triangular lattice [14]. We then define an IPS inspired by this local rule and prove with the same techniques that it also classifies the density.

3.1 A cellular automaton that classifies the density

Let us denote by $\mathtt{maj} : \mathcal{A}^3 \to \mathcal{A}$, the majority function, so that

$$\mathtt{maj}(x,y,z) = \begin{cases} 0 & \text{if } x+y+z<2 \\ 1 & \text{if } x+y+z \geq 2 \end{cases}.$$

Theorem 3.2. *The cellular automaton $\mathcal{T} : \mathcal{A}^{\mathbb{Z}^2} \to \mathcal{A}^{\mathbb{Z}^2}$ defined by:*

$$\mathcal{T}(x)_{i,j} = \mathtt{maj}(x_{i,j}, x_{i,j+1}, x_{i+1,j})$$

for any $x \in \mathcal{A}^{\mathbb{Z}^2}, (i,j) \in \mathbb{Z}^2$, classifies the density.

Proof. By symmetry, it is sufficient to prove that if $p > 1/2$, then $(\mu_p \mathcal{T}^n)_{n\geq 0}$ converges weakly to $\delta_{\mathbf{1}}$.

Let us consider the graph defined with $\mathbb{Z}^2$ as the set of sites (vertices) and $\{\{(i,j),(i,j+1)\},\{(i,j),(i+1,j)\},\{(i+1,j),(i,j+1)\},(i,j) \in \mathbb{Z}^2\}$ as the set of bonds (edges). This graph is equivalent to a triangular lattice, on which our notion of connectivity is defined. We recall that a 0-*cluster* is a subset of connected sites labelled by 0 which is maximal for inclusion. The site percolation threshold on the triangular lattice is equal to $1/2$ so that, for $p > 1/2$, there exists almost surely no infinite 0-cluster [14]. Thus, if $S_{\mathbf{0}}$ denotes the set of sites labelled by 0, the set $S_{\mathbf{0}}$ consists almost surely of a countable union $S_{\mathbf{0}} = \cup_{k\in\mathbb{N}} S_k$ of finite 0-clusters. Moreover, the size of the 0-clusters decays exponentially: there exist some constants κ and γ such that the probability for a given site to be part of a 0-cluster of size larger than n is smaller than $\kappa e^{-\gamma n}$, see Ref. [14].

Let us describe how the 0-clusters are transformed by the action of the CA. For $S \subset \mathbb{Z}^2$, let 1_S be the configuration defined by $(1_S)_x = 1$ if $x \in S$ and $(1_S)_x = 0$ otherwise. Let $\mathcal{T}(S)$ be the subset S' of $\mathbb{Z}^2$ such that $\mathcal{T}(1_S) = 1_{S'}$. By a simple symmetry argument, this last equality is equivalent to $\mathcal{T}(1_{\mathbb{Z}^2\setminus S}) = 1_{\mathbb{Z}^2\setminus S'}$. We observe the following.

- The rule does not break up or connect different 0-clusters (proved by Gács [9, Fact 3.1]). More precisely, if S consists of the 0-clusters $(S_k)_k$, then the components of $\mathcal{T}(S)$ are the nonempty sets among $(\mathcal{T}(S_k))_k$.
- Any finite 0-cluster disappears in finite time: if S is a finite and connected subset of $\mathbb{Z}^2$, then there exists an integer $n \geq 1$ such that $\mathcal{T}^n(S) = \emptyset$. This is the *eroder* property [5].
- Let us consider a 0-cluster and a rectangle in which it is contained. Then the 0-cluster always remains within this rectangle. More precisely, if R is a rectangle set, that is, a set of the form $\{(x,y) \in \mathbb{Z}^2 \mid a_1 \leq x \leq a_2,\ b_1 \leq y \leq b_2\}$, and if $S \subset R$, then for all $n \geq 1$, $\mathcal{T}^n(S) \subset R$ (the proof follows from $\mathcal{T}(S) \subset \mathcal{T}(R) \subset R$).

Let us now consider all the 0-clusters for which the minimal enveloping rectangle contains the origin $(0,0)$. By the exponential decay of the size of the clusters, one can prove that the number of such 0-clusters is almost surely finite. Indeed, the probability that the point of coordinates (m,n) is a part of such a cluster is smaller than the probability for this point to belong to a 0-cluster of size larger than $\max(|m|,|n|)$. And since

$$\sum_{(m,n)\in\mathbb{Z}^2} \kappa e^{-\gamma \max(|m|,|n|)} < 4\kappa \sum_{m\in\mathbb{N}} (m e^{-\gamma m} + \sum_{n\geq m} e^{-\gamma n}) < \infty,$$

we can apply the Borel-Cantelli lemma to obtain the result. Let T_0 be the maximum of the time needed to erase these 0-clusters. The random variable T_0 is almost surely finite, and after T_0 time steps, the site $(0,0)$ will always be labelled by a 1. As the argument can be generalised to any site, it ends the proof. ☐

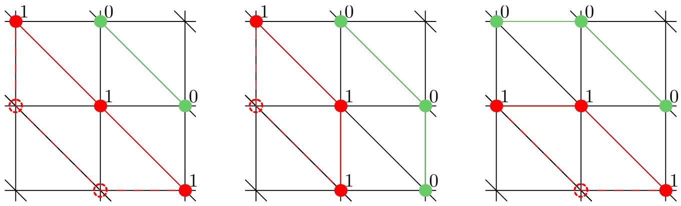


Figure 2: Illustration of the definition of the IPS.

We point out that Toom's CA classifies the density despite having many different invariant measures. For example:

- Any configuration x that can be decomposed into monochromatic North-East paths (that is, $x_{i,j} = x_{i,j+1}$ or $x_{i,j} = x_{i+1,j}$ for any i,j) is a fixed point and δ_x is an invariant measure.
- Let y be the checkerboard configuration defined by $y_{i,j} = 0$ if $i+j$ is even and $y_{i,j} = 1$ otherwise, and let z be defined by $z_{i,j} = 1 - y_{i,j}$. Since we have $\mathcal{T}(y) = z$ and $\mathcal{T}(z) = y$, the two configurations y and z form a periodic orbit and $(\delta_y + \delta_z)/2$ is an invariant measure.

3.2 An interacting particle system that classifies the density

We now define an IPS for which we use the same steps as above to prove that it classifies the density.

Note that the exact IPS analog of Toom's rule might classify the density but the above proof does not carry over since, in some cases, different 0-clusters may merge. To overcome the difficulty, we introduce a different IPS with a new neighbourhood of size 7: the cell itself and the six cells that are connected to it in the triangular lattice defined in the previous section.

For $\alpha \in \mathcal{A}$, set $\bar{\alpha} = 1 - \alpha$.

Theorem 3.3. *Let us consider the following IPS: for a configuration $x \in \mathcal{A}^{\mathbb{Z}^2}$, we update the state of the cell (i,j) by applying the majority rule on the North-East-Centre neighbourhood, except in the following cases (for which we keep the state unchanged):*

1. $x_{i,j} = x_{i-1,j+1} = x_{i+1,j-1} = \bar{x}_{i,j+1} = \bar{x}_{i+1,j}$

   and ($x_{i,j-1} = \bar{x}_{i,j}$ *or* $x_{i-1,j} = \bar{x}_{i,j}$),
2. $x_{i,j} = x_{i-1,j+1} = x_{i,j-1} = \bar{x}_{i,j+1} = \bar{x}_{i+1,j} = \bar{x}_{i+1,j-1}$ *and* $x_{i-1,j} = \bar{x}_{i,j}$,
3. $x_{i,j} = x_{i-1,j} = x_{i+1,j-1} = \bar{x}_{i,j+1} = \bar{x}_{i+1,j} = \bar{x}_{i-1,j+1}$ *and* $x_{i,j-1} = \bar{x}_{i,j}$.

This IPS classifies the density.

The three cases for which we always keep the state unchanged are illustrated below for $x_{i,j} = 1$ (central cell). In the first case, we allow to flip the central cell if and only if the two cells marked by a dashed circle are also labelled by 1. Otherwise, the updating could connect two different 0-clusters and break up the 1-cluster to which the cell (i,j) belongs to. The second and third cases are analogous.

The proof is similar to the one of Th. 3.2 but involves some additional technical points.

Proof. We assume as before that $p > 1/2$. Like the CA of the previous section, the new process that we have defined never breaks up a cluster or connects different ones. Furthermore, if we consider a 0-cluster and the smallest rectangle in which it is contained, we can check again that the 0-cluster will never go beyond this rectangle. As before, we only need to prove that any finite 0-cluster disappears almost surely in finite time to conclude the proof. We consider a realisation of the trajectory of the IPS with initial density μ_p. We associate to any finite 0-cluster $C \subset \mathbb{Z}^2$ the point $v(C) = \max\{(i,j) \in C\}$, using the lexicographic order on the coordinates (we set $v(\emptyset) = (-\infty, -\infty)$). In other words, the point $v(C)$ is the upmost point of C among its rightmost points. Let us consider at time 0 some finite 0-cluster C_0. We denote by C_t the state of this cluster at time t.

Claim. *The sequence $v(C_t)$ is nonincreasing. Moreover, if $t \geq 0$ is such that $C_t \neq \emptyset$, then there exists almost surely a time $t' > t$ such that $v(C_{t'}) < v(C_t)$.*

Let us prove the claim. Let us denote by $x \in \mathcal{A}^{\mathbb{Z}^2}$ a configuration attained at some time t, and let $(i,j) = v(C_t)$. By definition of $v(C_t)$, if a cell of coordinates $(i+1, j')$ is connected to a cell of C_t, then $x_{i+1,j'} = 1$. Either we have also $x_{i+1,j'+1} = 1$ and the cell $(i+1, j')$ will not flip, or $x_{i+1,j'+1} = 0$, but in this case, since $(i+1, j'+1)$ does not belong to C_t, $x_{i,j'+1} = 1$ and the cell of C_t to which is connected $(i+1, j')$ is necessarily (i, j'). So, $x_{i,j'} = 0$ and $x_{i+1,j'-1} = 1$, once again by definition of $v(C_t)$. Depending on the value of $x_{i+2,j'-1}$, either rule 1 or rule 2 forbids the cell $(i+1, j')$ to flip. In the same way, we can prove that if a cell of coordinates $(i, j'), j' > j$ is connected to C_t, then it is not allowed to flip. This proves that $v(C_t)$ is nonincreasing.

In order to prove the second part of the claim, we need to show that the cell (i,j) will almost surely be flipped in finite time. By definition of $(i,j) = v(C_t)$, we know that $x_{i,j+1} = x_{i+1,j} = x_{i+1,j-1} = 1$. The cell (i,j) will thus be allowed to flip, except if $x_{i-1,j+1} = x_{i,j-1} = 0$ and $x_{i-1,j} = 1$. But in that case, the cell $(i-1, j)$ will end up flipping, except if $x_{i-1,j-1} = x_{i-2,j+1} = 1, x_{i-2,j} = 0$, and so on. Let $W_n = \{(i-n, j), (i-1-n, j+1), (i-n, j-1)\}$. If for each n, the cells of W_n are in the state $(n \mod 2)$, then none of the cells $(i-n, j)$ is allowed to flip (see Fig. 3.a). But recall now that the initial measure is μ_p. There exists almost surely an integer $n \geq 0$ such that the initial state of the cell $(i-n, j)$ is *not* $(n \mod 2)$. Let $m(t)$ be the smallest integer n whose value at time t is not $n \mod 2$. Then, one can easily check that $m(t)$ is non-increasing, and that it reaches 0 in finite time. Thus, the cell (i,j) ends up flipping and we have proved the claim.

The example of Fig. 3.b illustrates how the proof works. Here, no cell of the cluster C_t is allowed to flip, but since the cells on the right and on the top of $v(C_t)$ cannot flip either, $v(C_t)$ does not increase. The cell at the left of $v(C_t)$ will end up flipping, and $v(C_t)$ will then be allowed to flip.

Since we know that a 0-cluster cannot go beyond its enveloping rectangle, a direct consequence of the claim is that any 0-cluster disappears in finite time. This allows us to conclude the proof in the same way as for the majority cellular automaton. □

3.3 The positive rates problem in $\mathbb{Z}^2$

Let us mention a connected problem and result. By definition, a PCA or an IPS of local function $\varphi : \mathcal{A} \longrightarrow \mathcal{M}(\mathcal{A})$ has *positive rates* if:

$$\forall u \in \mathcal{A}^{\mathcal{N}},\ \forall a \in \mathcal{A}, \qquad \varphi(u)(a) > 0\,. \tag{3.1}$$

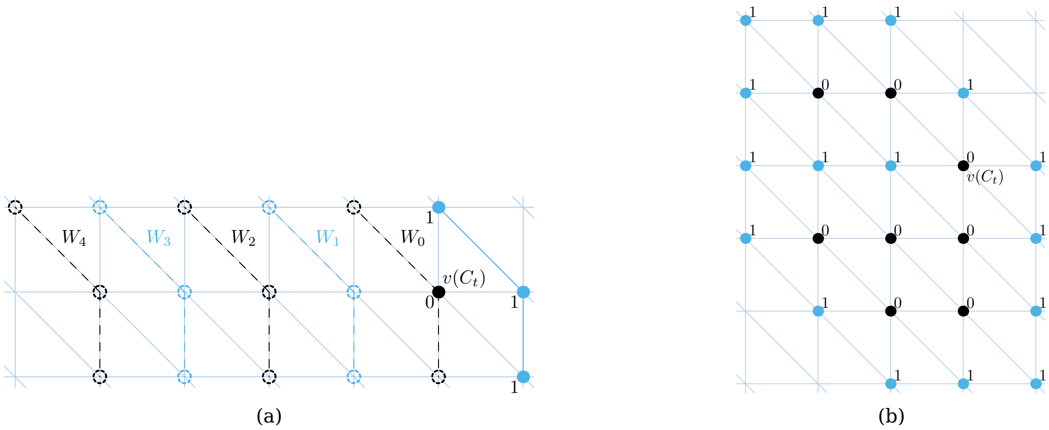


Figure 3: Illustration of the proof of Theorem 3.3

The *positive rates problem* consists in finding a positive rates model which is non-ergodic (with several invariant measures). This is a natural question, also relevant in the context of fault-tolerant models of computation, and which has been extensively studied.

In $\mathbb{Z}^2$, the positive rates problem is solved by a "perturbation" of Toom's CA. In fact, this was the motivation that led Toom to introduce the CA that bears his name. Let φ_0 be the local function of Toom's CA and define the positive rate PCA F with local function $\varphi : \mathcal{A}^{\mathcal{N}} \longrightarrow \mathcal{M}(\mathcal{A})$ given by:

$$\forall u \in \mathcal{A}^{\mathcal{N}},\ \forall a \in \mathcal{A}, \qquad \varphi(u) = (1-\varepsilon)\varphi_0(u) + \varepsilon\, \mathtt{Unif}\,, \tag{3.2}$$

where $\varepsilon \in (0,1)$ and where Unif is the uniform probability distribution on $\mathcal{A}$. The interpretation is that the computations are done according to Toom's rule, but, at each time and in each cell, an error may occur with probability ε in which case the new cell value is chosen uniformly. It is proved by Toom that for ε small enough, the positive rate PCA F has several invariant measures, with at least one close to "all 0", and one close to "all 1" [5, 25].

Intuitively and roughly, this non-ergodicity result and the one in Th. 3.2 can be viewed as being complementary, expressing the very strong "erasing" capacities of Toom's CA. Density classification amounts to erasing "errors" in the initial configuration (the symbols which are in minority), and non-ergodicity amounts to almost-erasing "errors" occurring in the whole space time diagram (the 1's if we are close to "all 0", or the 0's if we are close to "all 1").

4 Classifying the density on regular trees

Consider the finitely presented group $T_n = \langle a_1, \ldots, a_n \mid a_i^2 = 1 \rangle$. The Cayley graph of T_n is the infinite n-regular tree. For $n = 2k$, we also consider the free group with k generators, that is, $T'_{2k} = \langle a_1, \ldots, a_k \mid \cdot \rangle$. The groups T_{2k} and T'_{2k} are not isomorphic, but they have the same Cayley graph.

4.1 Shortcomings of the nearest neighbour majority rules

For odd values of n, a natural candidate for classifying the density is to apply the majority rule on the n neighbours of a cell. But it is proved that neither the CA (see Ref. [16] for $n = 3, 5$, and 7) nor the IPS (see Ref. [15] for $n = 3$) classify the density.

For $n = 4$, a natural candidate would be to apply the majority on the four neighbours and the cell itself. We now prove that it does not work either.

Proposition 4.1. *Consider the group $T_4' = \langle a, b \mid \cdot \rangle$. Consider the majority CA or IPS with neighbourhood $\mathcal{N} = \{1, a, b, a^{-1}, b^{-1}\}$. For $p \in (1/3, 2/3)$, the trajectories do not converge weakly to a uniform configuration.*

Proof. If $p \in (1/3, 2/3)$, then we claim that at time 0, there are almost surely infinite chains of zeros and infinite chains of ones that are fixed. Let us choose some cell labelled by 1. Consider the (finite or infinite) subtree of 1's originating from this cell viewed as the root. If we forget the root, the random tree is exactly a Galton-Watson process. The expected number of children of a node is $3p$ and since $3p > 1$, this Galton-Watson process survives with positive probability. Consequently, there exists almost surely an infinite chain of 1's at time 0 somewhere in the tree. In the same way, since $3(1 - p) > 0$, there exists almost surely an infinite chain of 0's. □

As for $\mathbb{Z}^2$, we get round the difficulty by keeping the majority rule but choosing a non-symmetrical neighbourhood.

4.2 A rule that classifies the density on T_4'

In this section, we consider the free group $T_4' = \langle a, b | \cdot \rangle$, see Fig. 4 (a).

Theorem 4.2. *The cellular automaton $F : \mathcal{A}^{T_4'} \to \mathcal{A}^{T_4'}$ defined by:*

$$F(x)_g = \mathtt{maj}(x_{ga}, x_{gab}, x_{gab^{-1}})$$

for any $x \in \mathcal{A}^{T_4'}, g \in T_4'$, classifies the density.

Proof. We consider a realisation of the trajectory of the CA with initial distribution μ_p. Let us denote by X_g^n the random variable describing the state of the cell g at time n. Since the process is homogeneous, it is sufficient to prove that X_1^n converges almost surely to 0 if $p < 1/2$ and to 1 if $p > 1/2$. Let us denote by $h : [0, 1] \to [0, 1]$ the function that maps a given $p \in [0, 1]$ to the probability $h(p)$ that $\mathtt{maj}(X, Y, Z) = 1$ when X, Y, Z are three independent Bernoulli random variables of parameter p. An easy computation provides $h(p) = 3p^2 - 2p^3$, and one can check that the sequence $(h^n(p))_{n \geq 0}$ converges to 0 if $p < 1/2$ and to 1 if $p > 1/2$.

We prove by induction on $n \in \mathbb{N}$ that for any $k \in \mathbb{N}$, the family $\mathcal{E}_k(n) = \{X^n_{u_1 u_2 \dots u_k} \mid u_1, u_2, \dots, u_k \in \{a, ab, ab^{-1}\}\}$ consists of independent Bernoulli random variables of parameter $h^n(p)$. By definition of μ_p, the property is true at time $n = 0$. Let us assume that it is true at some time $n \geq 0$, and let us fix some $k \geq 0$. Two different elements of $\mathcal{E}_k(n+1)$ can be written as the majority on two disjoint triples of $\mathcal{E}_{k+1}(n)$. The fact that the triples are disjoint is a consequence of the fact that $\{a, ab, ab^{-1}\}$ is a code: a given word $g \in G$ written with the elementary patterns a, ab, ab^{-1} can be decomposed in only one way as a product of such patterns. By hypothesis, the family $\mathcal{E}_{k+1}(n)$ is made of i.i.d. Bernoulli variables of parameter $h^n(p)$, so the variables of $\mathcal{E}_k(n + 1)$ are independent Bernoulli random variables of parameter $h^{n+1}(p)$. Consequently, the process F classifies the density on T_4'. □

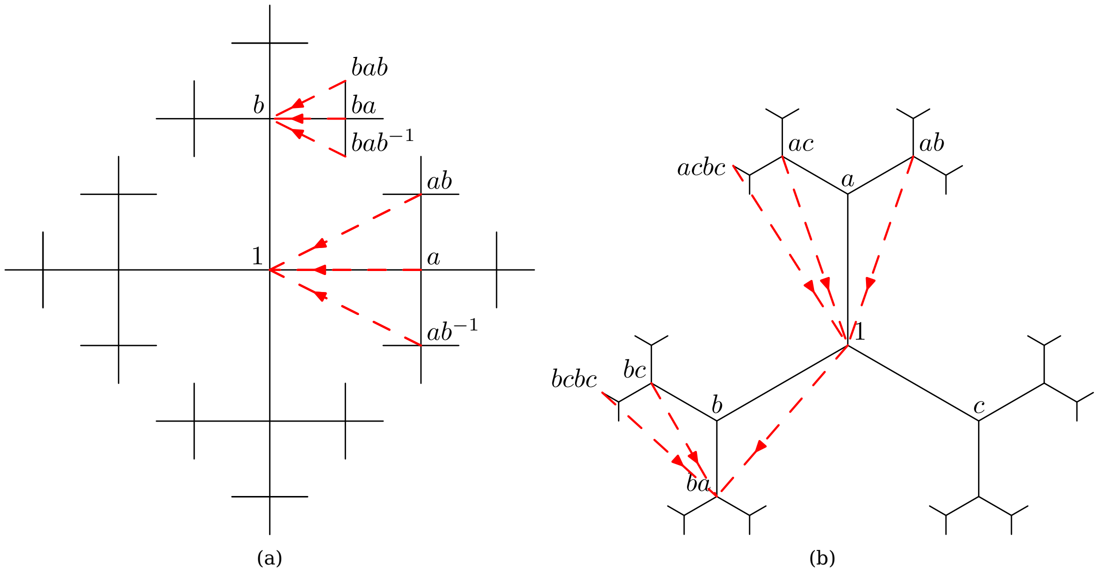


Figure 4: The cellular automata described by Theorem 4.2 and Theorem 4.3

Let us mention that from time $n \geq 1$, the field $(X_g^n)_{g \in G}$ is not i.i.d. For example, X_1^1 and $X^1_{ab^{-1}a^{-1}}$ are not independent since both of them depend on X_a^0.

On $T'_{2k} = \langle a_1, \dots, a_k | \cdot \rangle$, one can either apply Prop. 2.3 to obtain a cellular automaton that classifies the density, or define a new CA by the following formula: $F(x)_g = \mathtt{maj}(x_{ga_1}, x_{ga_1a_2}, x_{ga_1a_2^{-1}}, \dots, x_{ga_1a_k}, x_{ga_1a_k^{-1}})$ and check that it also classifies the density.

It is also possible to adapt the above proof to show that the IPS with the same local rule also classifies the density.

4.3 A rule that classifies the density on T_3

We now consider the group $T_3 = \langle a, b, c \mid a^2 = b^2 = c^2 = 1 \rangle$.

Theorem 4.3. *The cellular automaton* $F : \mathcal{A}^{T_3} \to \mathcal{A}^{T_3}$ *defined by:*

$$F(x)_g = \mathtt{maj}(x_{gab}, x_{gac}, x_{gacbc})$$

for any $x \in \mathcal{A}^{T_3}, g \in T_3$*, classifies the density.*

Proof. The proof is analogous to the previous case. We prove by induction on $n \in \mathbb{N}$ that for any $k \in \mathbb{N}$, that the family $\mathcal{E}_k(n) = \{X^n_{u_1u_2\dots u_k} \mid u_1, u_2, \dots, u_k \in \{ab, ac, acbc\}\}$ consists of independent Bernoulli random variables of parameter $h^n(p)$, the key point being that $\{ab, ac, acbc\}$ is a code. □

Once again, as explained in Prop. 2.3, since we have a solution on T_3, we obtain a CA that classifies the density for any $T_n, n \geq 3$, by applying exactly the same rule. The corresponding IPS on T_n also classifies the density.

The positive rates problem in regular trees. The positive rates problem is defined in Sec. 3.3. PCA solving the problem on regular trees appear in the literature, see Ref. [4]. Here, we obtain new examples by considering the CA of Th. 4.2 or the one of

Th. 4.3, and by defining its “perturbation” as in Eq. 3.2. It is not difficult to prove that for ε small enough, the resulting positive rates PCA is non-ergodic.

Again, this non-ergodicity result complements the density classification result, both of them reflecting strong erasing capacities of the CA (see the discussion at the end of Sec. 3.2).

5 Classifying the density on $\mathbb{Z}$

The density classification problem on $\mathbb{Z}$ appears as much more difficult than the other cases. We are not aware of any previous result in the literature (even partial), neither for (P)CA nor for IPS.

Below we focus on the synchronous version of the classification problem. First, we show that simple solutions do exist if we slightly relax the formulation of the problem (Sec. 5.1). Then we go back to the original problem. We first present a couple of naive (P)CA and show that they do not classify the density (Sec. 5.2). We then describe three models, two CA and one PCA, that are conjectured to classify the density (Sec. 5.3). We provide some preliminary analytical results (Sec. 5.4), as well as experimental investigations of the conjecture by using numerical simulations (Sec. 5.5).

In the examples below, the *traffic* cellular automaton, rule 184 according to Wolfram's notation, plays a central role. It is the CA with neighborhood $\mathcal{N} = \{-1, 0, 1\}$ and local function `traf` defined by:

x, y, z	111	110	101	100	011	010	001	000
$\texttt{traf}(x, y, z)$	1	0	1	1	1	0	0	0

This CA can be seen as a simple model of traffic flow on a single lane: the cars are represented by 1's moving one step to the right if and only if there are no cars directly in front of them. It is a density-preserving rule.

5.1 An exact solution with weakened conditions

On finite rings, several models have been proposed that solve relaxed variants of the density classification problem. We concentrate on one of these models introduced in Ref. [17]. The original setting is modified since the model operates on an extended alphabet, and the criterium for convergence is also weakened. Modulo this relaxation, it solves the problem on finite rings $\mathbb{Z}_n$. We show the same result on $\mathbb{Z}$.

Proposition 5.1. *Consider the cellular automaton F on the alphabet $\mathcal{B} = \mathcal{A}^2$, with neighbourhood $\mathcal{N} = \{-1, 0, 1\}$, and local function $f = (f_1, f_2)$ defined by:*

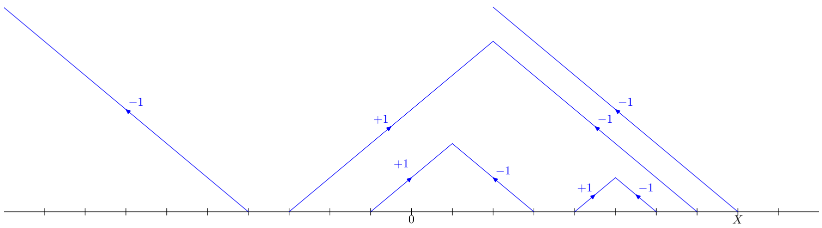
$$f_1(x, y, z) = \texttt{traf}(x_1, y_1, z_1) \quad ; \quad f_2(x, y, z) = \begin{cases} 0 & \text{if } x_1 = y_1 = 0 \\ 1 & \text{if } x_1 = y_1 = 1 \\ y_2 & \text{otherwise} \end{cases} \tag{5.1}$$

The projections $\mu_p F^n(\mathcal{A}^{\mathbb{Z}} \times \cdot)$ converge to $\delta_{\mathbf{0}}$ if $p < 1/2$ and to $\delta_{\mathbf{1}}$ if $p > 1/2$.

Intuitively, the CA operates on two tapes: on the first tape, it simply performs the traffic rule; on the second tape, what is recorded is the last occurrence of two consecutive zeros or ones in the first tape. If $p < 1/2$, then, on the first tape, there is a convergence to configurations which alternate between patterns of types 0^k and $(10)^\ell$. Consequently, on the second tape, there is convergence to the configuration $\delta_{\mathbf{0}}$. We formalise the argument below.

Proof. Let $T : \mathcal{A}^{\mathbb{Z}} \to \mathcal{A}^{\mathbb{Z}}$ be the traffic CA, see above. Following an idea of Belitsky and Ferrari [1], we define the recoding $\psi : \mathcal{A}^{\mathbb{Z}} \to \{-1,0,1\}^{\mathbb{Z}}$ by $\psi(x)_i = 1 - x_i - x_{i-1}$. Consider $(\psi \circ T^n(x))_{n\geq 0}$, the recodings of the trajectory of the CA originating from $x \in \{0,1\}^{\mathbb{Z}}$. There is a convenient alternative way to describe $(\psi \circ T^n(x))_{n\geq 0}$. It corresponds to the trajectories in the so-called *Ballistic Annihilation* model: 1 and -1 are interpreted as particles that we call respectively positive and negative particles. Negative particles move one cell to the left at each time step while positive particles move one cell to the right; and when two particles of different types meet, they annihilate.

Consider the Ballistic Annihilation model with initial condition $\mu_p \psi$ for $p > 1/2$. The density of negative particles is p^2, while the density of positive particles is $(1-p)^2$. During the evolution, the density of positive particles decreases to 0, while the density of negative particles decreases to $2p-1$. In particular, the negative particles that will never disappear have density $2p-1$ (see [1] for details). We can track back the position at time 0 of the “eternal” negative particles. Let X be the (random) position at initial time of the first eternal particle on the right of cell 0. After time X, the column 0 in the space-time diagram contains only 0 or -1 values. This key point is illustrated on the figure below.



We now go back to the traffic CA with initial condition distributed according to μ_p for $p > 1/2$ and concentrate on two consecutive columns of the space-time diagram. The property tells us that after some almost surely finite time, the columns doe not contain the pattern 00.

For the CA defined by Eq. 5.1 with an initial condition distributed according to a measure μ satisfying $\mu(\cdot \times \mathcal{A}^{\mathbb{Z}}) = \mu_p$ for $p > 1/2$, the above key point gets translated as follows: in any given column of the space-time diagram, after some a.s. finite time, the column contains only the letters $(0,1)$ or $(1,1)$. In particular, $\mu_p F^t(\mathcal{A}^{\mathbb{Z}} \times \cdot)$ converges weakly to $\delta_{\mathbf{1}}$ if $p > 1/2$. □

5.2 Models that do not classify the density on $\mathbb{Z}$

The first natural idea is to consider the majority rule for some neighborhood of odd size.

Recall the situation in $\mathbb{Z}^2$: with a symmetric neighborhood, classification is impossible (Lemma 3.1); with a non-symmetric neighborhood, classification is possible (Th. 3.2). In $\mathbb{Z}$, Lemma 3.1 still holds, so classification is impossible with a symmetric neighborhood. We now show that it remains impossible even with a non-symmetric neighborhood.

Below, we denote by $[x_0 \cdots x_n]_k$ the cylinder of all configurations $y \in \mathcal{A}^{\mathbb{Z}}$ satisfying $y_{k+i} = x_i$ for $0 \leq i \leq n$.

Lemma 5.2. *Consider a cellular automaton F performing the majority rule over a neighborhood of odd size. Then there exists k, l such that $F([0^k]_0) \subset [0^k]_l$ and $F([1^k]_0) \subset [1^k]_l$. In particular, F does not classify the density.*

Proof. Let the neighborhood be $\mathcal{N} = \{e_0, \cdots, e_{2n}\}$ with $e_i \in \mathbb{Z}$ and $e_0 < e_1 < \cdots < e_{2n}$. Assume for simplicity that $e_n = 0$ (the general case is treated similarly). Set

$k = e_{2n} - e_0 + 1$ and consider $x \in [0^k]_{e_0}$. By definition, $F(x)_i = \texttt{maj}(x_{i+e_0}, \dots, x_{i+e_{2n}})$, and

$$\begin{aligned} \text{if } e_0 \le i \le 0, \quad F(x)_i &= \texttt{maj}(x_{i+e_0}, \dots, x_{i+e_{n-1}}, 0, \dots, 0) = 0, \\ \text{if } 0 < i \le e_{2n}, \quad F(x)_i &= \texttt{maj}(0, \dots, 0, x_{i+e_{n+1}}, \dots, x_{i+e_{2n}}) = 0. \end{aligned}$$

So we have $F([0^k]_{e_0}) \subset [0^k]_{e_0}$. Similarly $F([1^k]_{e_0}) \subset [1^k]_{e_0}$. For $p \in (0,1)$, under the probability measure μ_p, an initial configuration will contain both patterns 0^k and 1^k with probability 1. Therefore, the CA cannot classify the density. □

Another natural idea consists in having a model in which the interfaces between monochromatic regions evolve like random walks, leading to an homogenization of the configuration. Let us show that a direct implementation of this idea does not work.

Consider the PCA with neighborhood $\mathcal{N} = \{-1, 1\}$, and local function $\varphi(x, y) = (1/2)\delta_x + (1/2)\delta_y$. In words, at each time step, the value of a cell is updated to the value of its left neighbor with probability 1/2 and to the value of its right neighbor with probability 1/2. This is the synchronous version of the Glauber dynamic associated with the Ising model at temperature 0. (In $\mathbb{Z}^2$, the analogous dynamics is conjectured to classify, see the discussion in Sec. 3.)

More generally, consider the PCA F with neighborhood $\mathcal{N} = \{e_1, \dots e_k\}$, $e_i \in \mathbb{Z}$, parameters $p_1, \dots, p_k \in (0,1)$ such that $\sum_{i=1}^k p_i = 1$, and local function

$$\varphi(x_{e_1}, \dots, x_{e_k}) = p_1 \delta_{x_{e_1}} + \dots + p_k \delta_{x_{e_k}} .$$

Lemma 5.3. *The PCA F does not classify the density.*

Proof. Let $(U_n)_{n\in\mathbb{Z}}$ be a sequence of i.i.d. random variables valued in $\{e_1, \dots, e_k\}$ with common law: $p_1\delta_{e_1} + \dots + p_k\delta_{e_k}$. Let μ be a probability measure on $\mathcal{A}^{\mathbb{Z}}$ and consider a sequence of random variables $(X_n)_{n\in\mathbb{Z}}$ distributed according to μ, and independent of $(U_n)_{n\in\mathbb{Z}}$. Define $Y_n = X_{n+U_n}$ for all $n \in \mathbb{Z}$. By construction, the sequence $(Y_n)_{n\in\mathbb{Z}}$ is distributed according to μF. Assume now that μ is shift-invariant. (The value of $\mu[x]_k$ does not depend on the position k and we denote it by $\mu[x]$.) We have

$$\begin{aligned} \mu F[1] = \mathbb{P}\{Y_0 = 1\} &= \sum_{i=1}^k \mathbb{P}\{Y_0 = 1, U_0 = e_i\} = \sum_{i=1}^k \mathbb{P}\{X_{e_i} = 1, U_0 = e_i\} \\ &= \sum_{i=1}^k \mathbb{P}\{X_{e_i} = 1\}\mathbb{P}\{U_0 = e_i\} = \sum_{i=1}^k \mu[1] p_i = \mu[1] . \end{aligned}$$

So the density of 1 is preserved by the dynamic, and F does not classify the density. The expected behavior is that homogenization occurs, leading to:
$\mu_p F^n \xrightarrow[n\to\infty]{w} (1-p)\delta_{\mathbf{0}} + p\delta_{\mathbf{1}}$. □

The behavior is thus the same as for the one-dimensional voter model IPS.

5.3 Density classifier candidates on $\mathbb{Z}$

We now propose three models, two CA (GKL and Kari-traffic) and one PCA (majority-traffic), that are candidates to classify the density on $\mathbb{Z}$.

All three of them perform well with respect to the density classification on finite rings. Figures 5 and 6 illustrate this point with space-time diagrams for the ring $\mathbb{Z}/149\mathbb{Z}$.

All three of them have the *eroder* property: if the initial configuration contains only a finite number of ones (resp. zeros), then it reaches **0** (resp. **1**) in finite time (almost surely for the PCA). Proofs appear in Ref. [12] for GKL and in Ref. [17] for Kari-traffic. For majority-traffic, $\alpha < 1/2$, a proof could be worked out by considering the interfaces between regions (all-black, all-white, and checkerboard) as particles.

GKL cellular automaton. The Gács-Kurdyumov-Levin (GKL) cellular automaton [11] is the CA with neighbourhood $\mathcal{N} = \{-3, -1, 0, 1, 3\}$ defined by: for $x \in \mathcal{A}^{\mathbb{Z}}, i \in \mathbb{Z}$,

$$\mathtt{Gkl}(x)_i = \begin{cases} \mathtt{maj}(x_i, x_{i+1}, x_{i+3}) & \text{if } x_i = 1 \\ \mathtt{maj}(x_i, x_{i-1}, x_{i-3}) & \text{if } x_i = 0. \end{cases} \tag{5.2}$$

Kari-traffic cellular automaton. The Kari-traffic rule [17], denoted by Kari, is the CA of neighborhood $\mathcal{N} = \{-3, -2, -1, 0, 1, 2, 3\}$ defined by: for $x \in \mathcal{A}^{\mathbb{Z}}$,

$$\mathtt{Kari}(x) = \Phi \circ \mathtt{Traf}(x) ,$$

where Traf is the traffic CA, that is the global function associated with traf, and where Φ is the CA defined by: for $x \in \mathcal{A}^{\mathbb{Z}}, i \in \mathbb{Z}$,

$$\Phi(x)_i = \begin{cases} 0 & \text{if } (x_{i-2}, x_{i-1}, x_i, x_{i+1}) = 0010 \\ 1 & \text{if } (x_{i-1}, x_i, x_{i+1}, x_{i+2}) = 1011 \\ x_i & \text{otherwise.} \end{cases} \tag{5.3}$$

The Kari-traffic rule is closely related to Kůrka's modified version of GKL [18].

Both GKL and Kari-traffic are symmetric when swapping 0 and 1 and right and left simultaneously.

Majority-traffic probabilistic cellular automaton. The majority-traffic PCA of parameter $\alpha \in (0, 1)$ is the PCA of neighbourhood $\mathcal{N} = \{-1, 0, 1\}$ and local function:

$$(x, y, z) \mapsto \alpha\, \delta_{\mathtt{maj}(x,y,z)} + (1 - \alpha)\, \delta_{\mathtt{traf}(x,y,z)}.$$

In words, at each time step, we choose, independently for each cell, to apply the majority rule with probability α and the traffic rule with probability $1 - \alpha$ (see Fig. 6).

The majority-traffic PCA has been introduced in Ref. [6] where the following is proved: for any $n \in \mathbb{N}$ and any $\varepsilon > 0$, there exists a value $\alpha_{n,\varepsilon}$ of the parameter such that on $\mathbb{Z}_n$, the PCA converges to the right uniform configuration with probability greater than $1 - \varepsilon$.

Conjecture 5.4. *The GKL CA, the Kari-traffic CA, and the majority-traffic PCA with* $0 < \alpha < \alpha_c$ *(for some* $0 < \alpha_c \leq 1/2$*) classify the density.*

5.4 Invariant Measures

Following ideas developed by Kůrka [18], we can give a precise description of the invariant measures of the three above models.

Let $x = (01)^{\mathbb{Z}}$ be the configuration defined by: $\forall n \in \mathbb{Z},\ x_{2n} = 0,\ x_{2n+1} = 1$. The configuration $(10)^{\mathbb{Z}}$ is defined similarly.

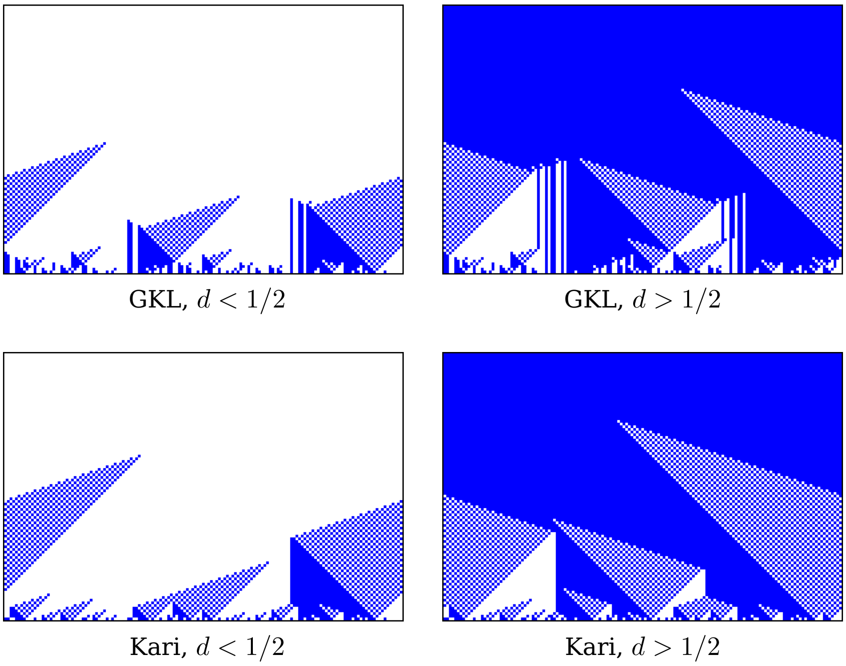


Figure 5: Two space-time diagrams of GKL (top) and Kari-traffic (bottom) on $\mathbb{Z}/149\mathbb{Z}$. The density of 1 in the initial condition is 70/149 (left) and 77/149 (right).

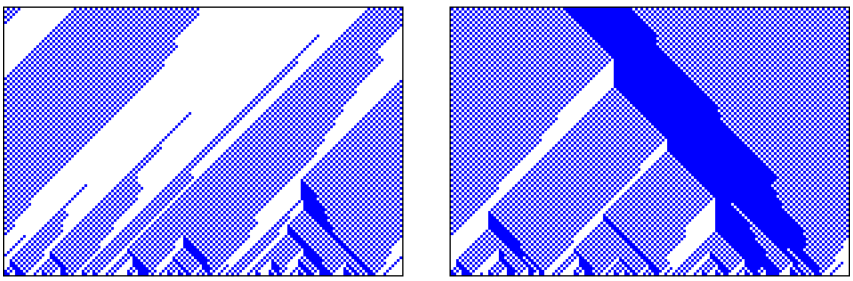


Figure 6: Two space-time diagrams of the majority-traffic PCA for $\alpha = 0.1$ on the ring $\mathbb{Z}/149\mathbb{Z}$. Both diagrams have the same initial condition with a density of 1 equal to 70/149. The right diagram corresponds to a rare event: evolution towards a configuration with only 1's, starting from a majority of 0's.

Proposition 5.5. *For the majority-traffic PCA and for the Kari-traffic CA, the extremal invariant measures are $\delta_{\mathbf{0}}, \delta_{\mathbf{1}}$, and $(\delta_{(01)^{\mathbb{Z}}} + \delta_{(10)^{\mathbb{Z}}})/2$. For GKL, on top of these three measures, there exist extremal invariant measures of density p for any $p \in [1/3, 2/3]$.*

Proof. **Majority-traffic.** Let us consider the majority-traffic PCA P of parameter $\alpha \in (0,1)$. Let μ be any shift-invariant measure. An exhaustive search shows that if at time 1, we observe the cylinder $[100]_0$ then there are only eight possible cylinders of size 5 at time 0, that are:

$$[01100]_{-1}, [10000]_{-1}, [10001]_{-1}, [10010]_{-1},$$
$$[10100]_{-1}, [11000]_{-1}, [11001]_{-1}, [11100]_{-1}.$$

Since the measure μ is shift-invariant, the probability $\mu([x_0 \cdots x_n]_k)$ does not depend on k and we denote it by $\mu[x_0 \cdots x_n]$. If we weight each of the above cylinder by the probability to reach $[100]_0$ from it, we obtain the following expression:

$$\mu P[100] = \alpha(1-\alpha)\mu[01100] + (1-\alpha)\mu[10000] + (1-\alpha)\mu[10001] + (1-\alpha)\mu[10010] + \alpha\mu[10100] + \alpha^2\mu[11000] + \alpha^2\mu[11001] + \alpha(1-\alpha)\mu[11100].$$

Gathering the terms with the same coefficient, we have:

$$\begin{aligned} \mu P[100] &= (1-\alpha)(\mu[100] - \mu[10011]) + \alpha\mu[10100] + \alpha(1-\alpha)\mu[1100] + \alpha^2\mu[1100] \\ &= (1-\alpha)(\mu[100] - \mu[10011]) + \alpha\mu[10100] + \alpha\mu[1100]. \end{aligned}$$

Some more rearrangements provide:

$$\begin{aligned} \mu P[100] &= (1-\alpha)(\mu[100] - \mu[10011]) + \alpha(\mu[100] - \mu[00100]) \\ &= \mu[100] - (1-\alpha)\mu[10011] - \alpha\mu[00100]. \end{aligned}$$

This proves that the sequence $(\mu P^n[100])_{n\geq 0}$ is non-increasing. From now on, let us assume that $\mu P = \mu$. Then, $\mu[10011] = \mu[00100] = 0$.

Let us consider the cylinder $[10^n0011]$ for some $n \geq 2$. If we apply the majority rule on each cell except on the second cell from the left, then after n iterations, we reach the cylinder $[10011]$. Since this occurs with a positive probability, we obtain that for any $n \geq 0, \mu[10^n0011] = 0$. This provides: $\mu[0011] = \mu[00011] = \mu[000011] = \ldots = \mu[0^n11]$ for any $n \geq 2$. Consequently, $\mu[0011] = 0$. From a cylinder of the form $[00(10)^n11]$, if we choose to apply the majority rule on each cell, then we reach the cylinder $[0011]$ in n steps. Thus, $\mu[00(10)^n11] = 0$ for any $n \geq 0$. It follows that μ can be written as the sum $\mu = \mu_0 + \mu_1$ of two invariant measures, where μ_0 charges only the subshift Σ_0 and μ_1 the subshift Σ_1 with

$$\Sigma_0 = \{x \in \mathcal{A}^{\mathbb{Z}} \mid \forall k \in \mathbb{Z}, x_k x_{k+1} \neq 00\}, \qquad \Sigma_1 = \{x \in \mathcal{A}^{\mathbb{Z}} \mid \forall k \in \mathbb{Z}, x_k x_{k+1} \neq 11\}. \quad (5.4)$$

Let us assume that $\mu[00] = 0$ (which is the case for μ_0). In the same way that we have computed $\mu P[110]$, we can compute $\mu P[11]$, and we obtain:

$$\begin{aligned} \mu P[11] &= \alpha\mu[0110] + \alpha\mu[1110] + \alpha\mu[1101] + \mu[1011] + \mu[0111] + \mu[1111] \\ &= \alpha\mu[110] + \alpha\mu[1101] + \mu[11] - \mu[0011] \\ &= \mu[11] + \alpha\mu[110] + \alpha\mu[1101]. \end{aligned}$$

By hypothesis, $\mu P = \mu$, so that the last equality implies that $\mu[110] = 0$.

In all cases, if μ is a shift-invariant measure such that $\mu P = \mu$, then $\mu[00] = \mu(\mathbf{0}), \mu[11] = \mu(\mathbf{1})$ and $\mu[01] = \mu[10] = \mu((01)^{\mathbb{Z}}) = \mu((10)^{\mathbb{Z}})$.

Kari-traffic. If at time 1, we observe the pattern 100 at position 0, then, at time 0, this same pattern was present at position -1. This can be checked by systematic inspection. In the same way, if, at time 1, we observe the pattern 110 at position 0, then, at time 0, this same pattern was present at position 1.

Let μ be a shift-invariant measure such that $\mu K = \mu$, where $K = \texttt{Kari}$. A consequence of the above results on the patterns 100 and 110 is that: $\mu K^{n+1}[110x100] = 0$ for any $n \geq 0$ and any $x \in \mathcal{A}^n$. But since $\mu K = \mu$, we obtain $\mu[110x100] = 0$ for any word x. Like for GKL, we can write $\mu = \mu_0 + \mu_1$ where μ_0 and μ_1 are two invariant measures defined on the subshifts Σ_0 and Σ_1, see Eq. 5.4.

Let us consider a configuration of Σ_0, that is, without the pattern 00. By the traffic rule, each 0 of the configuration will move one cell to the left. Then by rule Φ (see Eq. 5.3), if a 0 is at distance greater than 2 from the next 0 on its right, it is erased. The result follows.

GKL. Any word $x \in \mathcal{A}^{\mathbb{Z}}$ which is a concatenation of the patterns $u = 001$ and $v = 011$ is a fixed point of the GKL cellular automaton: if $x_n = 0$, then either $x_{n-1} = 0$ or $x_{n-3} = 0$ so that $F(x)_n = 0$ and if $x_n = 1$, then either $x_{n+1} = 1$ or $x_{n+3} = 1$ so that $F(x)_n = 1$. As a consequence, GKL has extremal invariant measures of density p for any $p \in [1/3, 2/3]$. □

To summarize, majority-traffic and Kari-traffic have a simpler set of invariant measures. It does not rule out GKL as a candidate for solving the density classification task, but rather indicates that it could be easier to prove the result for majority-traffic or Kari-traffic.

The positive rates problem in $\mathbb{Z}$. Recall that the positive rates problem is defined in Sec. 3.3. In $\mathbb{Z}$, it had been a long standing conjecture that all positive rates PCA and IPS are ergodic. The GKL CA, see Eq. 5.2, was originally introduced as a candidate to solve the positive rates problem, with the conjecture that its perturbed version may be non-ergodic [11]. But it is still unknown if it is the case or not [12, 22]. In 2001, Gács disproved the conjecture by exhibiting a very complex counter-example with several invariant measures, but with an alphabet of cardinality at least 2^{18} [10, 13]. It is the only known counter-example.

To summarize, in $\mathbb{Z}$, there is no known model that classifies the density, and there is no known “simple” model that solves the positive rates problem. This reflects the difficulty to build a model in $\mathbb{Z}$ with strong erasing properties.

5.5 Experimental results

Let us recall the arguments backing up Conjecture 5.4. First, the three models have the eroder property. Second, they classify reasonably well on a finite ring.

To go further, we perform some numerical experimentations. Our approach is to test if the proportion of good classification on a finite ring converges to one as the size of the ring increases. Indeed, it is reasonable to believe that there is a relationship between this last property and the ability to classify on $\mathbb{Z}$.

More precisely, we proceed as follows. We fix a rule (GKL, Kari-traffic, or Majority-traffic for $\alpha = 0.1$) and a parameter $p \in (0, 1/2)$. We consider different rings of odd sizes ranging from 101 to 2001. For each size, we perform 10^5 experiments, by choosing each time a new initial configuration according to the Bernoulli product measure μ_p, that is,

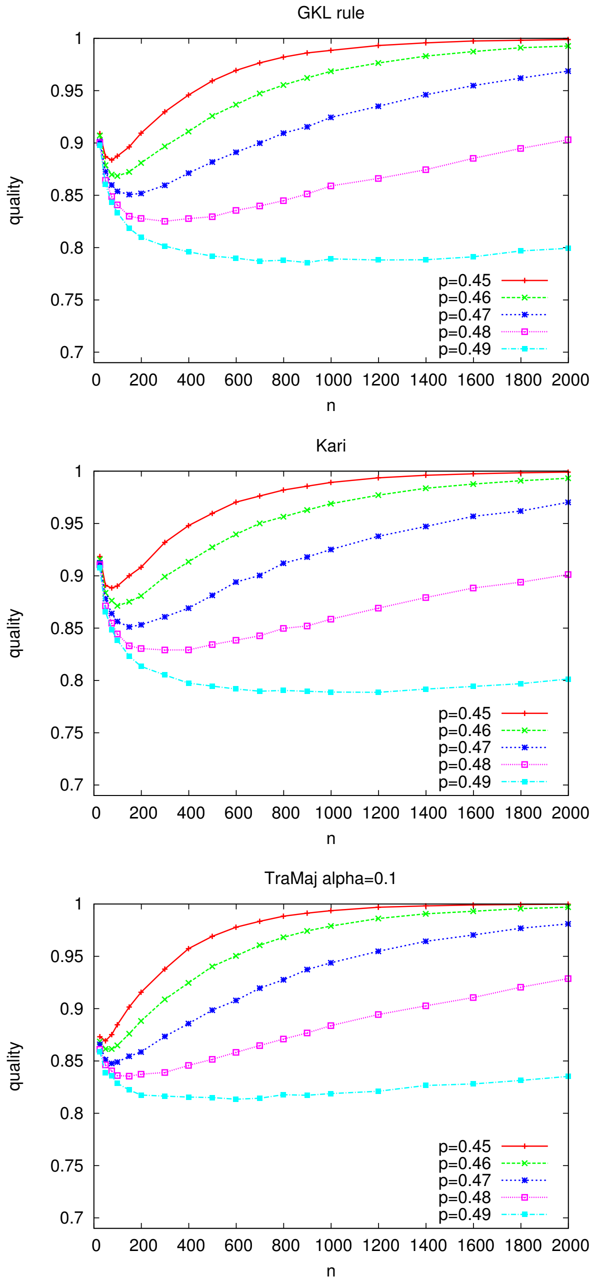


Figure 7: Experimental determination of the quality of classification $Q(n)$ as a function of the ring size n. Cells are initialised with a probability p to be in state 1. Each point represents an average computed on $100\,000$ experiments.

we assign to each cell the value 1 with a probability p and the value 0 with probability $1-p$. We record the proportion of good classifications among the 10^5 experiments. We denote this proportion by $Q(n)$ where n is the ring size. Let $d(x)$ be the proportion of 1 in the initial configuration x distributed according to μ_p. We may have $d(x) > 1/2$, although $E[d(x)] = p < 1/2$. We have a "good classification" for x if there is convergence to **0** when $d(x) < 1/2$ and to **1** when $d(x) > 1/2$.

The results are reported in Fig. 7. For each rule, we consider five different values for the parameter p, ranging from 0.45 to 0.49. For each rule and each value of the parameter, the plot is consistent with the hypothesis that $Q(n)$ converges to 1. However, when p approaches 1/2, the ring size n needed for $Q(n)$ to attain a certain quality level increases dramatically.

On each of the plots, we observe an initial decrease of $Q(n)$ followed by an increase for n large. For $p = 0.49$, the point of inflexion becomes hardly visible. Our explanation is that for small ring sizes, the dispersion of the actual density $d(x)$ is higher and covers values far from 1/2 for which the classification is easier.

References

[1] V. Belitsky and P. Ferrari. Invariant measures and convergence properties for cellular automaton 184 and related processes. *J. Stat. Phys.*, 118(3-4):589–623, 2005. MR-2123649

[2] A. Bušić, N. Fatès, J. Mairesse, and I. Marcovici. Density classification on infinite lattices and trees. In *LATIN 2012*, volume 7256 of *LNCS*, pages 109–120. Springer-Verlag, 2012.

[3] J. T. Cox and R. Durrett. Nonlinear voter models. In *Random walks, Brownian motion, and interacting particle systems*, volume 28 of *Progr. Probab.*, pages 189–201. Birkhäuser, 1991. MR-1146446

[4] D. Dawson. Stable states of probabilistic cellular automata. *Information and Control*, 34(2):93–106, 1977. MR-0446758

[5] R.L. Dobrushin, V.I. Kryukov, and A.L. Toom. *Stochastic cellular systems: ergodicity, memory, morphogenesis*. Nonlinear science. Manchester University Press, 1990.

[6] N. Fatès. Stochastic cellular automata solve the density classification problem with an arbitrary precision - when randomness helps computing. *Theory of Computing Systems*, OnLine first, DOI: 10.1007/s00224-012-9386-3, 2012. MR-2853436

[7] L. R. Fontes, R. H. Schonmann, and V. Sidoravicius. Stretched exponential fixation in stochastic Ising models at zero temperature. *Comm. Math. Phys.*, 228(3):495–518, 2002. MR-1918786

[8] H. Fukś. Nondeterministic density classification with diffusive probabilistic cellular automata. *Phys. Rev. E*, 66, 2002.

[9] P. Gács. A Toom rule that increases the thickness of sets. *J. Statist. Phys.*, 59(1-2):171–193, 1990. MR-1049966

[10] P. Gács. Reliable cellular automata with self-organization. *J. Statist. Phys.*, 103(1-2):45–267, 2001. MR-1828729

[11] P. Gács, G. Kurdyumov, and L. Levin. One-dimensional homogeneous media dissolving finite islands. *Problems of Information Transmission*, 14:92–96, 1978.

[12] P. Gonzaga de Sá and C. Maes. The Gacs-Kurdyumov-Levin automaton revisited. *J. Statist. Phys.*, 67(3-4):507–522, 1992. MR-1171143

[13] L. Gray. A reader's guide to P. Gács's "positive rates" paper: "Reliable cellular automata with self-organization". *J. Statist. Phys.*, 103(1-2):1–44, 2001. MR-1828728

[14] G. Grimmett. *Percolation*, volume 321 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, 2nd edition, 1999. MR-1707339

[15] C. Douglas Howard. Zero-temperature Ising spin dynamics on the homogeneous tree of degree three. *J. Appl. Probab.*, 37(3):736–747, 2000. MR-1782449

[16] Y. Kanoria and A. Montanari. Majority dynamics on trees and the dynamic cavity method. *Ann. Appl. Probab.*, 21(5):1694–1748, 2011. MR-2884049

[17] J. Kari and B. Le Gloannec. Modified traffic cellular automaton for the density classification task. *Fund. Inform.*, 116:1–16, 2012. MR-2977840

[18] P. Kůrka. Cellular automata with vanishing particles. *Fund. Inform.*, 58(3-4):203–221, 2003. MR-2073077

[19] M. Land and R. Belew. No perfect two-state cellular automata for density classification exists. *Phys. Rev. Lett.*, 74:5148–5150, 1995.

[20] T. Liggett. *Interacting particle systems*. Classics in Mathematics. Springer-Verlag, 2005. Reprint of the 1985 original. MR-2108619

[21] N. Packard. *Dynamic Patterns in Complex Systems*. World Scientific, 1988. MR-1037037

[22] K. Park. *Ergodicity and mixing rate of one-dimensional cellular automata*. Ph.D. Thesis, Boston University, 1997. MR-2694652

[23] N. Pippenger. Symmetry in self-correcting cellular automata. *J. Comput. Syst. Sci.*, 49(1):83 – 95, 1994. MR-1284592

[24] M. Schüle, T, Ott, and R. Stoop. Computing with probabilistic cellular automata. In *Proceedings of the 19th International Conference on Artificial Neural Networks: Part II*, ICANN '09, pages 525–533, 2009. Springer-Verlag.

[25] A. Toom. Stable and attractive trajectories in multicomponent systems. In *Multicomponent random systems*, volume 6 of *Adv. Probab. Related Topics*, pages 549–575. Dekker, 1980. MR-0599548