

Coupling from the past for exponentially ergodic one-dimensional probabilistic cellular automata

Jean Bérard

▶ **To cite this version:**

Jean Bérard. Coupling from the past for exponentially ergodic one-dimensional probabilistic cellular automata. 2021. hal-03258335v2

HAL Id: hal-03258335
https://hal.science/hal-03258335v2

Preprint submitted on 30 Aug 2021 (v2), last revised 1 Sep 2021 (v3)

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

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

COUPLING FROM THE PAST FOR EXPONENTIALLY ERGODIC ONE-DIMENSIONAL PROBABILISTIC CELLULAR AUTOMATA

JEAN BÉRARD

Abstract. For every exponentially ergodic one-dimensional probabilistic cellular automaton with positive rates, we construct a locally defined coupling-from-the-past flow whose coalescence time has a finite exponential moment. This construction leads to a finite-size necessary and sufficient condition for exponential ergodicity of one-dimensional cellular automata. As a corollary, we prove that every sufficiently small perturbation of an exponentially ergodic one-dimensional cellular automaton is exponentially ergodic.

1. Introduction

1.1. **Definitions and the main result.** Probabilistic Cellular Automata (PCA) form a class of discrete-time Markov processes on spaces of the form $\mathscr{A}^{\mathbb{L}}$, where $\mathbb{L}$ is a lattice (typically, $\mathbb{L} = \mathbb{Z}^d$ for some $d \geq 1$), and $\mathscr{A}$ is a finite set called an alphabet (see e.g. [10] for a seminal reference on the subject, and [3] for a recent overview). In the present paper, we consider one-dimensional PCAs, that is, $\mathbb{L} = \mathbb{Z}$. An element of $\mathscr{A}^{\mathbb{Z}}$ is a bi-infinite sequence $(v(x))_{x\in\mathbb{Z}}$, where $v(x) \in \mathscr{A}$ for all $x \in \mathbb{Z}$, and we equip the set $\mathscr{A}^{\mathbb{Z}}$ with the product topology and product $\sigma-$algebra. The dynamics of the PCA is specified through a transition kernel $\mathcal{K}$ from $\mathscr{A}^{\{-1,0,1\}}$ to $\mathscr{A}$, so that for every $\mathbf{v} = (v_{-1}, v_0, v_1) \in \mathscr{A}^{\{-1,0,1\}}$, $\mathcal{K}(\mathbf{v}, \cdot)$ is a probability measure on $\mathscr{A}$. Formally, a probabilistic cellular automaton with kernel $\mathcal{K}$ is a discrete-time Markov process $(X_t)_t$ on $\mathscr{A}^{\mathbb{Z}}$, such that:

$$\text{The random variables } (X_t(x))_{x\in\mathbb{Z}} \text{ are independent given } X_{t-1}, \tag{1}$$

$$\forall x \in \mathbb{Z},\ X_t(x) \sim \mathcal{K}(\pi_{[\![x-1,x+1]\!]}(X_{t-1}), \cdot), \tag{2}$$

where $[\![x, y]\!]$ denotes the discrete interval $[\![x, y]\!] = \{z \in \mathbb{Z};\ x \leq z \leq y\}$, and where, given two sets $I \subset J$, we denote by π_I the canonical projection from $\mathscr{A}^J$ to $\mathscr{A}^I$.

Moreover, we say that our PCA satisfies the *positive rates* condition when there exists a $w \in \mathscr{A}$ such that

$$\min_{\mathbf{v}\in\mathscr{A}^{\{-1,0,1\}}} \mathcal{K}(\mathbf{v}, \{w\}) > 0. \tag{3}$$

One key question about the long-term dynamics of PCAs is that of ergodicity: we say that a PCA is *ergodic* when there exists a (necessarily unique) probability distribution μ on $\mathscr{A}^{\mathbb{Z}}$ such that, for every initial condition $X_0 = \xi \in \mathscr{A}^{\mathbb{Z}}$, one has the convergence $X_t \xrightarrow[t\to+\infty]{d} \mu$. For an ergodic PCA, an important additional question is that of the convergence speed: we say that a PCA is *exponentially ergodic* when there exist positive constants $a, b > 0$ such that, for all $t \geq 0$, all $\xi \in \mathscr{A}^{\mathbb{Z}}$, and all $I \subset \mathbb{Z}$,

$$d_{\mathrm{TV}}\left(\mathrm{Law}(\pi_I(X_t)), \pi_I(\mu)\right) \leq a|I|e^{-bt}, \tag{4}$$

where d_{TV} denotes the total variation distance between probability measures (see Section 2), and where, for a probability measure ν on $\mathscr{A}^J$ with $I \subset J$, we denote by $\pi_I(\nu)$ the corresponding image probability measure on $\mathscr{A}^I$.

Among the various methods that may be used to prove ergodicity, we focus on the so-called *coupling from the past* (CFTP) approach, which has become a popular tool in the context of Markov-chain based numerical

The author would like to thank Y. Spinka for useful discussions, and in particular for pointing the references [8, 9].

methods (see [7]), but had already been used earlier (not under this specific name) to establish ergodicity for a variety of processes – see e.g. [10] in the context of PCAs, or [4] in the context of continuous-time interacting particle systems.

To formalize this approach, we define a CFTP *flow* to be a family of random functions[1] $\left(\Phi_{t_n}^{t_{n+1}}\right)_{n\geq 0}$, where $(t_n)_{n\geq 0}$ is a decreasing integer-valued sequence such that $t_0 = 0$, and where $\Phi_{t_n}^{t_{n+1}} : \mathscr{A}^{\mathbb{Z}} \to \mathscr{A}^{\mathbb{Z}}$ is such that, for all $n \geq 1$, and all $\xi \in \mathscr{A}^{\mathbb{Z}}$, the sequence $\xi, \Phi_{t_{n-1}}^{t_n}(\xi), \ldots, \Phi_{t_0}^{t_1} \circ \cdots \circ \Phi_{t_{n-1}}^{t_n}(\xi)$, has the same distribution as $X_{t_n}, X_{t_{n-1}}, \ldots, X_{t_0}$, starting from $X_{t_n} = \xi$. The *coalescence time* of the flow at a site $x \in \mathbb{Z}$ is then defined as

$$T_x = \inf\{n \geq 1;\ \pi_x \circ \Phi_{t_0}^{t_1} \circ \cdots \circ \Phi_{t_{n-1}}^{t_n} \text{ is a constant function}\},$$

with the convention $\inf \emptyset = +\infty$, and where we use the notation π_x instead of $\pi_{\{x\}}$ when $I = \{x\}$. If for all x, one has that $T_x < +\infty$ a.s., then the PCA is ergodic. Moreover, if the tail of T_x satisfies an inequality of the form $\mathbb{P}(T_x > t) \leq ae^{-bt}$ for all $t \geq 0$ (with a and b not depending on x), one gets the bound (4) with the same constants a and b.

Our main result is a converse to this property. It states that, whenever a PCA is exponentially ergodic and has positive rates, it is possible to define a CFTP flow for which T_x has a finite exponential moment (uniformly bounded over x), and which is, in a precise sense, locally defined.

Theorem 1. *Consider an exponentially ergodic one-dimensional PCA with positive rates. Then there is a CFTP flow with $t_n = -n \cdot L$ for a certain integer L, enjoying the following properties:*

(i) for all $x \in \mathbb{Z}$ and $t \geq 0$, $\mathbb{P}(T_x > t) \leq ce^{-dt}$ where $c > 0$ and $d > 0$ do not depend on x;

(ii) the family of random functions $\left(\Phi_{t_n}^{t_{n+1}}\right)_{n\geq 0}$ is i.i.d.;

(iii) there exists an i.i.d. family of random variables $(V_y)_{y\in \frac{1}{2}\mathbb{Z}}$, and a (measurable) function F, such that, for all $x \in \mathbb{Z}$ and $\xi \in \mathscr{A}^{\mathbb{Z}}$, one can write the value of $\Phi_{t_0}^{t_1}(\xi)$ within $[\![x(2L) - L, x(2L) + L]\!]$ as:

$$\pi_{[\![x(2L)-L,x(2L)+L]\!]}\left(\Phi_{t_0}^{t_1}(\xi)\right) = F\left(V_{x-1/2}, V_x, V_{x+1/2}, \pi_{[\![(x-1)(2L),(x+1)2L]\!]}(\xi)\right).$$

Property (i) merely states the exponential bound on the tail of the coalescence time. Property (ii) is a locality property of the flow with respect to time: the flow is defined on a regular time-grid with mesh L, with an i.i.d. structure over distinct time cells. Finally, property (iii) is a locality property with respect to space: over a grid with mesh $2L$, the flow only involves the value of the initial condition and an auxiliary i.i.d. structure within a bounded window.

The conclusion of Theorem 1 is already known to hold, under a stronger form, in the case of a *monotone* PCA (i.e. when the kernel $\mathcal{K}$ is stochastically monotone with respect to a total order on $\mathscr{A}$ and the corresponding partial product order on $\mathscr{A}^{\{-1,0,1\}}$), as observed in [11]. In such a case, ergodicity alone is enough to guarantee the existence of a CFTP flow, one can take $t_n = -n$, and $\pi_x((\Phi_0^{-1}(\xi))$ can be written as $F(V_x, \pi_{[\![x-1,x+1]\!]}(\xi))$; moreover, the tail of the coalescence time precisely matches the actual speed of convergence to the limiting distribution.

Still, to our knowledge, a result as general as Theorem 1 – where no other assumption beyond exponential ergodicity and positive rates is needed – is new, and, except in the monotone case just discussed, only sufficient (but not necessary) conditions for the existence of such a CFTP flow were known (see e.g. [2, 5]). Moreover, it is still an open question (see Problems 6.1 and 6.2 in [5]) whether ergodic but not exponentially ergodic PCA exist, so Theorem 1 can in fact be applied to every known example of an ergodic one-dimensional PCA.

[1] We use the space $\mathscr{S}$ of functions $\phi : \mathscr{A}^Z \to \mathscr{A}^{\mathbb{Z}}$ for which there exists an $r \geq 1$ such that the value of $\phi(\xi)$ at site x is a function of those values of $\xi(y)$ for which $|y - x| \leq r$ only. Measurability on $\mathscr{S}$ is then defined by viewing elements of $\mathscr{S}$ as countable collections of functions from $\mathscr{A}^{[\![x-r,x+r]\!]}$ to $\mathscr{A}$.

1.2. **Consequences.** A direct consequence of Theorem 1 is the existence of an algorithm to perfectly sample from the invariant distribution μ of any exponentially ergodic one-dimensional PCA with positive rates. Also, using the results[2] in [8], we deduce that, if $X \sim \mu$, the joint distribution of $\left(\pi_{x(2L),(x+1)2L}(X)\right)_{x\in\mathbb{Z}}$ admits a representation as a finite factor of a finite-valued i.i.d. process (it is unclear whether this can be strengthened to prove that μ itself enjoys this property).

Next, we observe that the flow constructed in the proof of Theorem 1 leads to a *finite-size* necessary and sufficient condition for exponential ergodicity (with positive rates). Specifically, the proof shows that, assuming positive rates, exponential ergodicity is equivalent to the existence of an integer $L \geq 1$ such that

$$\rho = (4L+1)\mathbb{P}\left(\pi_{[\![-L,L]\!]} \circ \Phi_0^{-L} \text{ is a constant function}\right) < 1. \tag{5}$$

As a consequence, at least in principle, the property of being an exponentially ergodic PCA can always be checked using an algorithm that explores larger and larger values of L (and of the other relevant parameters used in the construction), and stops when a value of $\rho < 1$ has been found. Another consequence of this finite-size condition is that exponential ergodicity (with positive rates) is a robust property with respect to sufficiently small perturbations of the dynamics, as stated in the following corollary.

Corollary 1. *If the kernel $\mathcal{K}$ defines an exponentially ergodic one-dimensional PCA with positive rates, it is also the case of any kernel $\mathcal{K}'$ that is a sufficiently small perturbation of $\mathcal{K}$.*

1.3. **Discussion.** Theorem 1 holds for exponentially ergodic one-dimensional PCAs, and it is indeed a natural question whether an analogous result holds in dimension $d \geq 2$.

One place[3] where the proof of Theorem 1 seems to rely heavily on the one-dimensional setting is Lemma 8, where we show that exponential ergodicity implies the existence of a coupling with good coalescence properties for the dynamics with boundary conditions. The proof of the lemma uses the fact that the number of sites within a fixed distance of the boundary of a $d-$dimensional box does not grow with the size of the box, which is specific to $d = 1$. (This is reminiscent of the proof in [6] that "weak mixing implies strong mixing for squares" in the context of two-dimensional spin systems, where here we have one dimension of space and one of time instead of two dimensions of space.) Using stronger mixing conditions (involving the dynamics with boundary conditions) may allow to extend the conclusion of Theorem 1 to dimensions $d \geq 2$, but it is unclear how such mixing conditions could be related to more familiar ones in the context of PCAs such as (4). Note that, in the distinct but related context of Markov random fields, the use of "strong" mixing conditions to build CFTP structures and/or perfect simulation algorithms is an active research topic (see e.g. [9, 1], and the references therein).

Another interesting extension would be to the case of (continuous-time) interacting particle systems, for which the deterministic bound on the speed of propagation of information in PCA dynamics does not hold.

1.4. **Organization of the paper.** The paper is essentially self-contained. Section 2 contains definitions and simple but useful results on couplings (no claim at originality is made there). Section 3 is devoted to definitions related to PCA dynamics within trapezoids, which are heavily used in the subsequent proofs. Section 4 contains a succession of lemmas leading to the proof of Theorem 1 and Corollary 1.

2. Couplings

Given a finite set S and a finite family $(\nu_e)_{e\in E}$ of probability distributions on S, a *coupling* of $(\nu_e)_{e\in E}$ is a family of $S-$valued random variables $(Z_e)_{e\in E}$, such that $Z_e \sim \nu_e$ for all $e \in E$. Alternatively, we may view such a coupling as a random map $\Psi : E \to S$, where $\Psi(e) = Z_e$.

Given two probabilities ν_1, ν_2 on S, remember the definition of the total variation distance $d_{\mathrm{TV}}(\nu_1, \nu_2) = \frac{1}{2}\sum_{s\in S} |\nu_1(s) - \nu_2(s)|$. It is a classical result that the total variation distance is the minimum value of

[2]Note that, in [8], the term "exponentially ergodic PCA" is used to refer to the existence of a suitable CFTP flow, whereas in the present paper, exponential ergodicity is a mixing property from which we have to deduce the existence of the CFTP flow.

[3]But not necessarily the only place, see also Lemma 5.

$\mathbb{P}(Z_1 = Z_2)$ over all couplings of ν_1, ν_2. The following two lemmas provide two useful variations over this kind of result.

Lemma 1. *Assume that, for a certain $0 < \epsilon < 1$, there exists an $e_0 \in E$ such that one has $d_{\mathrm{TV}}(\nu_e, \nu_{e_0}) \leq \epsilon/|S|$ for all $e \in E$. Then, for all $\gamma \in]\epsilon, 1[$, there exists a coupling of $(\nu_e)_{e\in E}$ such that $\mathbb{P}(Z_e = Z_{e_0}$ for all $e \in E) \geq (1-\gamma)(1-\epsilon/\gamma)$.*

Proof. Let $A = \{s \in S;\ \nu_{e_0}(s) \leq \epsilon/(\gamma|S|)\}$. One has that $\nu_{e_0}(A) = \sum_{s\in A} \nu_{e_0}(s) \leq |A| \cdot \epsilon/(\gamma|S|) \leq \epsilon/\gamma$. Now, if $s \in A^c$, one has that $\nu_{e_0}(s) \geq \epsilon/(\gamma|S|)$, so that, since $|\nu_e(s) - \nu_{e_0}(s)| \leq d_{\mathrm{TV}}(\nu_e, \nu_{e_0}) \leq \epsilon/|S|$, one has $\nu_e(s) \geq \nu_{e_0}(s) - \epsilon/|S| \geq (1-\gamma)\nu_{e_0}(s)$. Now consider a pairwise disjoint family $(\mathfrak{I}(s))_{s\in A^c}$ of subintervals of $[0,1]$, with respective lengths $(1-\gamma)\nu_{e_0}(s)$, Then, for every $e \in E$, complete these intervals into a partition of $[0,1]$ by adding pairwise disjoint intervals $(\mathfrak{L}_e(s))_{s\in A^c}$, with respective lengths $\nu_e(s) - (1-\gamma)\nu_{e_0}(s)$, and pairwise disjoint intervals $(\mathfrak{K}_e(s))_{s\in A}$, with respective lengths $\nu_e(s)$. Now consider a random variable U with uniform distribution on $[0,1]$. Whenever U belongs to the interval $\mathfrak{I}(s)$, we set $Z_e = s$ for all $e \in E$. When U does not belong to $\bigcup_{s\in A^c} \mathfrak{I}(s)$, for a given e, either U belongs to a (unique) interval $\mathfrak{L}_e(s)$, or to a (unique) interval $\mathfrak{K}_e(s)$, and we define Z_e as precisely the corresponding s. It is now apparent that each Z_e has ν_e as its distribution, while $\mathbb{P}(Z_{e_0} = Z_e$ for all $e \in E) \geq \mathbb{P}(U \in \bigcup_{s\in A^c} \mathfrak{I}(s)) = \sum_{s\in A^c}(1-\gamma)\nu_{e_0}(s) = (1-\gamma)\nu_{e_0}(A^c) \geq (1-\gamma)(1-\epsilon/\gamma)$. □

Lemma 2. *Assume that, for a certain $0 < \epsilon < 1$, there exists an $e_0 \in E$ such that one has $d_{\mathrm{TV}}(\nu_e, \nu_{e_0}) \leq \epsilon$ for all $e \in E$. Then there exists a coupling of $(\nu_e)_{e\in E}$ such that, for all $J \subset E$, $\mathbb{P}(Z_e = Z_{e_0}$ for all $e \in J) \geq 1 - |J|\epsilon$.*

Proof. We recycle the classical coupling construction leading to the probability of equality between a pair of random variables being equal to the total variation distance. First consider a partition of the interval $[0,1]$ into a pairwise disjoint family $(\mathfrak{J}_{e_0}(s))_{s\in S}$ of subintervals, with respective lengths $\nu_{e_0}(s)$. For $e \in E \setminus \{e_0\}$, let $A_e = \{s \in S;\ \nu_{e_0}(s) \geq \nu_e(s))\}$. For $s \in A_e$, let $\mathfrak{J}_e(s)$ be a subinterval of $[0,1]$ with length $\nu_e(s)$ such that $\mathfrak{J}_e(s) \subset \mathfrak{J}_{e_0}(s)$. Then, for $s \in A_e^c$, let $\mathfrak{J}_e(s)$ be the union of a finite number of disjoint subintervals of $[0,1]$, in such a way that $\mathfrak{J}_{e_0}(s) \subset \mathfrak{J}_e(s)$, that the total length of $\mathfrak{J}_e(s)$ equals $\nu_e(s)$, and that the family $(\mathfrak{J}_e(s))_{s\in S}$ forms a partition of $[0,1]$. Using a random variable U with uniform distribution on $[0,1]$, and defining Z_e as the unique s such that $\mathfrak{J}_e(s)$ contains U, we have that $Z_e \sim \nu_e$ for all $e \in E$, and, for all $e \in E \setminus \{e_0\}$, $\mathbb{P}(Z_{e_0} = Z_e) = d_{\mathrm{TV}}(\nu_e, \nu_{e_0})$. Thus, $\mathbb{P}(\exists\ e \in J$ such that $Z_{e_0} \neq Z_e) \leq \sum_{e\in J} \mathbb{P}(Z_{e_0} \neq Z_e) = \sum_{e\in J} d_{\mathrm{TV}}(\nu_e, \nu_{e_0}) \leq |J|\epsilon$. □

In the sequel, a coupling provided by Lemma 1 (resp. Lemma 2) will be called a type I (resp. type II) coupling.

3. Trapezoids

In this paper, we use the generic term *trapezoid* to refer to discrete isoceles trapezoids drawn on the space-time lattice $\mathbb{Z} \times (-\mathbb{N})$ whose lateral sides have their respective slopes equal either to $-1, +1$ or to $+1, -1$, as shown in Fig. 1. We distinguish between *downward* trapezoids (when the top is longer than the base), and *upward* trapezoids (when the top is shorter than the base), with time flowing from top to bottom.

We define the *outer boundary* of an upward trapezoid $\mathbf{T}$ as the union, on both sides, of the two discrete segments parallel to the lateral sides of $\mathbf{T}$, at horizontal distance respectively 1 and 2 from $\mathbf{T}$, starting at the ordinate of the top, and stopping one unit above the ordinate of the base. The outer boundary is denoted by $\partial_+\mathbf{T}$. We also use the notation $\mathbf{T}(m) = \mathbf{T} \cap (\mathbb{Z} \times \{m\})$.

3.1. **Dynamics within a trapezoid.**

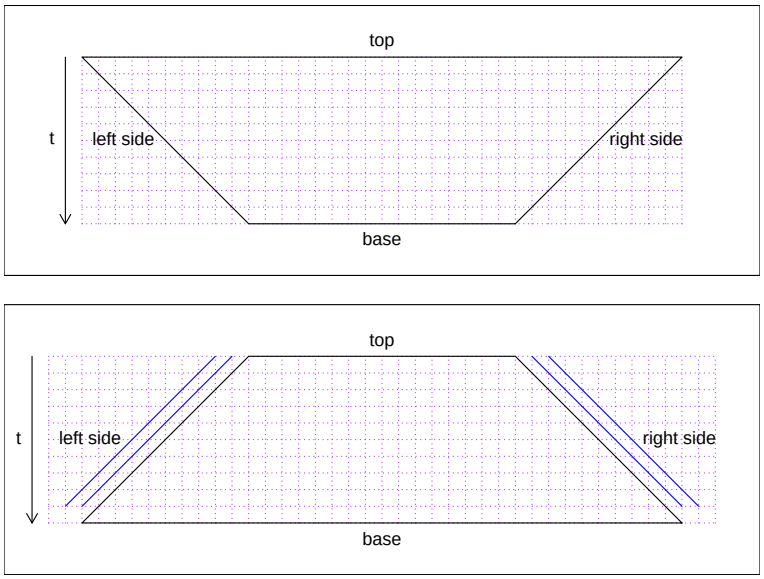


FIGURE 1. A downward trapezoid (above) and an upward trapezoid (below). In the upward case, the outer boundary is shown in blue. The integer lattice is drawn using purple dotted lines.

3.1.1. *Downward case.* Consider a downward trapezoid $\mathbf{T}$ with height L and base-length K, with $\mathrm{top}(\mathbf{T}) = [\![z-L, z+K+L]\!] \times \{\tau\} = \mathbf{T}(\tau)$, and $\mathrm{base}(\mathbf{T}) = [\![z, z+K]\!] \times \{\tau+L\} = \mathbf{T}(\tau+L)$. Starting from a configuration ζ consisting of an element of $\mathscr{A}$ at each site of the top of $\mathbf{T}$, we define the PCA dynamics within $\mathbf{T}$ as a Markov process on the successive state spaces $\mathscr{A}^{\mathbf{T}(\tau)}, \dots, \mathscr{A}^{\mathbf{T}(\tau+L)}$ in which, given the configurations within $\mathbf{T}(\tau+m)$, where $0 \leq m \leq L-1$, the configuration within $\mathbf{T}(\tau+m+1)$ is obtained by following (1)-(2), for $x \in [\![z-L+m+1, z+K+L-m-1]\!]$. We denote by $G(\zeta, \mathbf{T})$ the resulting overall distribution on $\mathscr{A}^{\mathbf{T}}$.

The following restriction property shows that the dynamics within $\mathbf{T}$ we have just defined, coincides with the restriction of the overall dynamics of the PCA within $\mathbf{T}$, conditional upon a suitably defined "outside" of $\mathbf{T}$. The proof is omitted, and is an easy consequence of e.g. the basic coupling described in Subsection 3.2.3 below.

Lemma 3. *Consider $s \leq \tau$, and define $out_s(\mathbf{T})$ as the set of (x,t) such that either $s \leq t < \tau$, or $\tau \leq t \leq \tau+L$ and the horizontal distance from (x,t) to the boundary of $\mathbf{T}$ is $\geq t-\tau$. Starting from $X_s = \xi$ at a time $s \leq \tau$, the distribution of $(X_t(x))_{(x,t)\in\mathbf{T}}$, conditional upon $(X_t(x))_{(x,t)\in out_s(\mathbf{T})}$, is $G(\zeta, \mathbf{T})$, with $\zeta = (X_t(x), (x,t) \in base(\mathbf{T}))$.*

3.1.2. *Upward case.* Consider an upward trapezoid $\mathbf{T}$ with height L and top-length M, $\mathrm{top}(\mathbf{T}) = [\![z, z+M]\!] \times \{\tau\} = \mathbf{T}(\tau)$ and $\mathrm{base}(\mathbf{T}) = [\![z-L, z+M+L]\!] \times \{\tau+M\} = \mathbf{T}(\tau+L)$. Starting from a configuration ζ consisting of an element of $\mathscr{A}$ at each site of the top of $\mathbf{T}$, and a boundary condition χ consisting of an element of $\mathscr{A}$ at each site of the outer boundary $\partial_+\mathbf{T}$, we can define the PCA dynamics within $\mathbf{T}$ as in the previous case: given the configurations within $\mathbf{T}(\tau+m)$, where $0 \leq m \leq L-1$, the configuration within $\mathbf{T}(\tau+m+1)$ is obtained by following (1)-(2), for $x \in [\![z-m-1, z+M+m+1]\!]$, using the boundary condition to make sense of (2) when $x \in \{z-m-1, z-m, z+M+m, z+M+m+1\}$. We denote by $G_\chi(\zeta, \mathbf{T})$ the resulting overall distribution on $\mathscr{A}^{\mathbf{T}}$.

We now state a restriction property for the dynamics with boundary conditions on $\mathbf{T}$. The proof is similar to that of Lemma 4.

Lemma 4. *Consider $s \leq \tau$, and define $out_s(\mathbf{T})$ as the set of $(x,t) \in [\![s, \tau + L]\!] \times \mathbb{Z}$ such that either $(x,t) \notin \mathbf{T}$ or $(x,t) \in base(\mathbf{T})$. Starting from $X_s = \xi$ at a time $s \leq \tau$, the distribution of $(X_t(x))_{(x,t)\in\mathbf{T}}$, conditional upon $(X_t(x))_{(x,t)\in out_s(\mathbf{T})}$, is $G_\chi(\zeta, \mathbf{T})$, with $\zeta = (X_t(x), (x,t) \in base(\mathbf{T}))$ and $\chi = (X_t(x), (x,t) \in \partial_+(\mathbf{T})$.*

3.2. **Coupling within a trapezoid.**

3.2.1. *Downward $(K,L)-$coupling.* Given a downward trapezoid $\mathbf{T}$ with height L and base-length K, we define a downward $(K,L)-$coupling to be a coupling of the dynamics within $\mathbf{T}$, for every possible initial configuration on the top, that is, a coupling of the family $G(\zeta, \mathbf{T})$, $\zeta \in \mathscr{A}^{\mathrm{top}(\mathbf{T})}$. Note that, given a coupling for the configuration at the base, i.e. a coupling for the family $\pi_{\mathrm{base}(\mathbf{T})}(G(\zeta, \mathbf{T}))$, $\zeta \in \mathscr{A}^{\mathrm{top}(\mathbf{T})}$, one can always build a full $(K,L)-$coupling by sampling from the distribution of the whole dynamics within $\mathbf{T}$ starting from ζ, conditional upon the random configuration at the base generated by the coupling. If Ψ denotes a random function from $\mathscr{A}^{\mathrm{top}(\mathbf{T})}$ to $\mathscr{A}^{\mathbf{T}}$ corresponding to a $(K,L)-$coupling, we say that *coalescence* occurs when $\pi_{\mathrm{base}(\mathbf{T})} \circ \Psi$ is a constant function, and we say that an $(x,t) \in \mathrm{base}(\mathbf{T})$ is *locked* when $\pi_{(x,t)} \circ \Psi$ is a constant function.

3.2.2. *Upward $(M,L)-$coupling.* For an upward trapezoid $\mathbf{T}$ with height L and top-length M, we define an upward $(M,L)-$coupling to be a coupling of the dynamics within $\mathbf{T}$ for every possible boundary condition, and every possible initial configuration on the top of $\mathbf{T}$, i.e. a coupling for the family $G_\chi(\zeta, \mathbf{T})$, $\zeta \in \mathscr{A}^{\mathrm{top}(\mathbf{T})}, \chi \in \mathscr{A}^{\partial_+\mathbf{T}}$. As above, a coupling for the configuration at the base is enough to define a full $(M,L)-$coupling. If Ψ denotes a random function corresponding to an $(M,L)-$coupling, we say that *coalescence* occurs for the boundary condition χ when $\zeta \mapsto \pi_{\mathrm{base}(\mathbf{T})}(\Phi(\zeta,\chi))$ is a constant function, and we say that $(x,t) \in \mathrm{base}(\mathbf{T})$ is *locked* for the boundary condition χ when $\zeta \mapsto \pi_{(x,t)}(\Phi(\zeta,\chi))$ is a constant function.

3.2.3. *The basic coupling.* The basic coupling provides a simple way of defining couplings for the PCA dynamics. It is defined through an i.i.d. family of random functions $(\Gamma_{x,t})_{x\in\mathbb{Z},t\in\mathbb{Z}}$, where $\Gamma_{x,t} : \mathscr{A}^{\{-1,0,1\}} \to \mathscr{A}$ is such that, for all $\mathbf{v} \in \mathscr{A}^{\{-1,0,1\}}$, the law of $\Gamma_{x,t}(\mathbf{v})$ is $K(\mathbf{v}, \cdot)$. Moreover, thanks to the positive rates property (3), we may assume that there is a $\kappa > 0$ and a $w \in \mathscr{A}$ such that

$$\mathbb{P}(\Gamma_{x,t}(\mathbf{v}) = w \text{ for all } \mathbf{v}) \geq \kappa. \tag{6}$$

(It is easy to explicitly design such functions, using a single random variable $U_{x,t}$ with uniform distribution on $[0,1]$ and a suitable partition of $[0,1]$ into sub-intervals for each $\mathbf{v}$). Conditions (1)-(2) are then implemented through the equation:

$$X_t(x) = \Gamma_{x,t}\left(\pi_{[\![x-1,x+1]\!]}(X_{t-1})\right).$$

Using the basic coupling, we can easily design downward $(K,L)-$ or upward $(M,L)-$couplings, but these may not enjoy the coalescence properties we are after. We shall nevertheless use the basic coupling on parts of the trapezoids we consider, using the restriction properties contained in Lemmas 3 and 4 to patch together couplings defined on different parts.

4. Proof of the main results

Our first lemma shows that, for downward trapezoids with a sufficiently large height-to-base ratio, one has a coupling with suitable control over the non-coalescence probability.

Lemma 5. *There exist constants $\alpha > 0$, $c_1 > 0$, $d_1 > 0$ such that, for all large enough K, and all $L \geq \alpha K$, one can define a downward $(K,L)-$coupling such that the probability of non-coalescence is bounded above by $c_1 \cdot e^{-d_1 K}$.*

Proof. For $K \geq 1$, and arbitrary z, τ, denote by $\mathbf{T}$ the downward trapezoid with top $= [\![z-L, z+K+L]\!] \times \{\tau\}$, and base $= [\![z, z+K]\!] \times \{\tau + L\}$. We shall apply Lemma 1 with $S = \mathscr{A}^{\mathrm{base}}$, $E = \mathscr{A}^{\mathrm{top}}$, $\mu_e = \pi_{\mathrm{base}}(G(e, \mathbf{T}))$, and e_0 an arbitrarily chosen element of E.

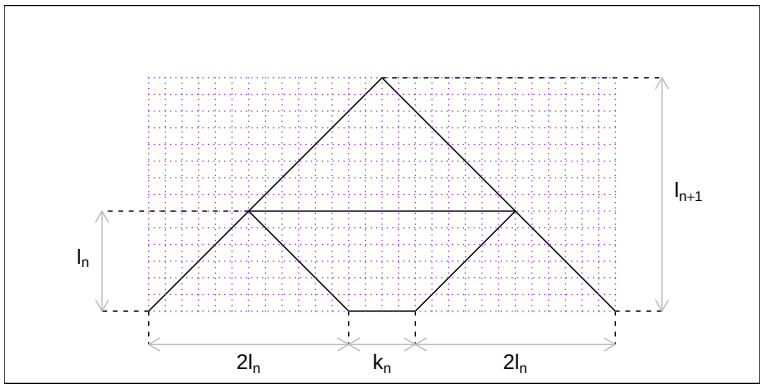


FIGURE 2. Fitting a trapezoid with base length k_n and height ℓ_n into a triangle with base length $2\ell_{n+1}$ and height ℓ_{n+1}. The integer lattice is drawn using purple dotted lines.

One has $|S| = |\mathscr{A}|^{K+1}$, and $d_{\mathrm{TV}}(\mu_e, \mu_{e_0}) \leq 2a(K+1) \cdot e^{-bL} \leq 2a(K+1) \cdot e^{-b\alpha K}$ for all e. As soon as $\alpha > \log(|\mathscr{A}|)/b$, we see that $\epsilon = 2a(K+1) \cdot e^{-b\alpha K} \cdot |\mathscr{A}|^{K+1}$ decays exponentially fast with K, so we can apply Lemma 1, with the value of ϵ just defined, and e.g. $\gamma = \epsilon^{1/2}$, to get the desired coupling. □

It turns out that, to prove Theorem 1, we need to extend Lemma 5 to allow for "flatter" trapezoids, at the price of a slightly worse bound on the coalescence probability. This is done in the following two lemmas, using as a key tool a family of nested self-similar trapezoids with a type I coupling (Lemma 6), followed by a type II coupling (Lemma 7).

Lemma 6. *For all $A > 0$, and for arbitrarily large L, there exists a downward $(K, L)-$coupling with the following properties as $L \to +\infty$:*

- $L/K \sim (\log L)^{-A}$
- *the probability that the number of unlocked sites exceeds $K \cdot (\log L)^{-A}$ is bounded above by $e^{-L^{1+o(1)}}$*

Proof. Let $\alpha > 0$, $c_1 > 0$ and $d_1 > 0$ be as in the statement of Lemma 5. Let ℓ_0 be a even integer number such that $\ell_0 \geq 4\alpha$, and define inductively the sequences $(\ell_n)_{n\geq 0}$ and $(k_n)_{n\geq 0}$ by $k_n = \lfloor \ell_n/\alpha \rfloor_2$, and $\ell_{n+1} = k_n/2 + 2\ell_n$, where $\lfloor m \rfloor_2$ stands for the largest even integer number less than or equal to m.

These definitions allow one to exactly fit a downward trapezoid with base length k_n and height ℓ_n into a discrete isoceles triangle with base length $2\ell_{n+1}$ and height ℓ_{n+1}, as shown in Fig. 2.

By definition, we have that, for all n, $\ell_{n+1} \geq \ell_n/(2\alpha) - 1 + 2\ell_n = (2 + 1/(2\alpha))\ell_n - 1$, and we deduce that the sequence $(\ell_n)_{n\geq 0}$ is increasing, and that, for all $n \geq 0$, $1/(2\alpha) \leq k_n/\ell_n \leq 1/\alpha$. On the other hand, we have that $(1 + 1/(2\alpha))^n \ell_0 \leq \ell_n \leq (2 + 1/(2\alpha))^n \ell_0$.

Now put side-by-side q downward trapezoids with base length k_n and height ℓ_n. These trapezoids form generation 0, and fit into a larger downward trapezoid $\mathbf{T}$ of height $L = \ell_n$ and base length $K = qk_n + (2q-2)\ell_n$. Between two consecutive trapezoids of generation 0 lies a triangle with base length $2\ell_n$ and height ℓ_n. Within every such triangle, we fit a trapezoid with base length k_{n-1} and height ℓ_{n-1}. These trapezoids form generation 1. We then iterate the following procedure for $i = 1, \ldots, n-1$. Between two consecutive trapezoids of generation $\leq i$ (two consecutive trapezoids may belong to distinct generations) lies a triangle with base length $2\ell_{n-i}$ and height ℓ_{n-i}. Within every such triangle, we fit a trapezoid with base length k_{n-i-1} and height ℓ_{n-i-1}. An illustration is provided in Fig. 3.

The $0-$th generation trapezoids cover a base of total length qk_n, and the triangles between them cover a base of total length $(2q-2)\ell_n$. For $i = 0, \ldots, n-1$, going from generation i to generation $i+1$ results in the addition of a new generation of trapezoids with heights ℓ_{n-i-1} and base lengths k_{n-i-1}, which multiplies the base length previously covered by triangles in generation i by a factor $1 - \frac{k_{n-i-1}}{k_{n-i-1}+4\ell_{n-i-1}} \leq \frac{8\alpha}{1+8\alpha} < 1$.

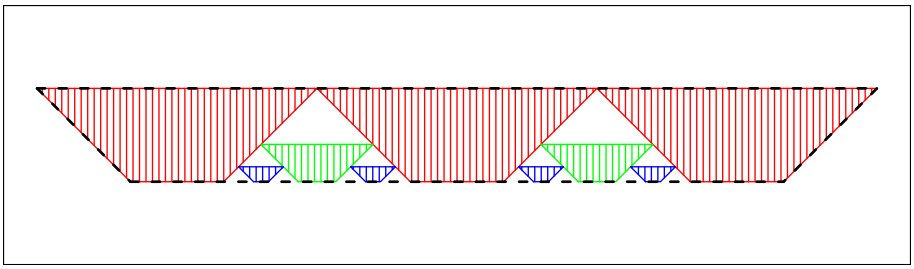


FIGURE 3. Illustration with $q = 3$ and $\alpha = 1$. The trapezoid $\mathbf{T}$ is drawn with dashed lines. Three generations of nested trapezoids are drawn: generation 0 (red), generation 1 (green), generation 2 (blue).

As a result, the total length in the base of $\mathbf{T}$ that is not covered by the base of a trapezoid of whichever generation, is less than $f = (2q-2)\ell_n \left(\frac{8\alpha}{1+8\alpha}\right)^n$.

There are q trapezoids in generation number 0, and $q-1$ in generation number 1. After generation 1, each further generation leads to twice as many trapezoids as in the previous one, so the total number of trapezoids is $r = q + (q-1)\cdot(1+2+\cdots+2^{n-2}) \leq q\cdot 2^n$.

We now define a downward $(K,L)-$coupling inside $\mathbf{T}$, for all large enough ℓ_0. We use within each trapezoid belonging to generation $n-j$ (with height ℓ_j and base length k_j), the $(k_j,\ell_j)-$coupling from Lemma 5, independently from other trapezoids (the fact that ℓ_0 is large enough, and that, by construction, $\ell_j \geq \alpha k_j$, ensures that the lemma can be applied for all $j = 0,\ldots,n$). In the part of $\mathbf{T}$ not belonging to any of the previous trapezoids, we just use the basic coupling. That this is a licit construction leading to a downward $(K,L)-$coupling is a consequence of Lemmas 3 and 4.

For a trapezoid of height ℓ_j and base length k_j, the probability of non-coalescence of the $(k_j,\ell_j)-$coupling is, according to Lemma 5, bounded above by $c_1 e^{-d_1 k_j}$. By the union bound, the probability that coalescence does not occur in at least one of the trapezoids, is less than $rc_1 e^{-d_1 k_0}$, and so less than $q\cdot 2^n e^{-d_1 k_0}$. When coalescence occurs in every trapezoid, every unlocked site of our overall $(K,L)-$coupling must belong to the complement of the bases of these trapezoids, whose total length does not exceed f.

Now let B be such that $B\cdot\log\left(1+\frac{1}{8\alpha}\right) > A$, let $q = \left\lceil \frac{1}{2+1/\alpha}\cdot(\log\ell_0)^A \right\rceil$ and let $n = \lceil B\cdot\log\log\ell_0\rceil$ (assuming that ℓ_0 is large enough so that $q\geq 2$ and $n\geq 1$).

From the bound $(2+1/\alpha)^n\ell_0 \leq \ell_n \leq (2+2/\alpha)^n\ell_0$, we see that, as $\ell_0\to+\infty$, $\log\ell_n \sim \log\ell_0$, and also $q\sim\frac{1}{2+1/\alpha}\cdot(\log\ell_n)^A$. Remembering that $L=\ell_n$ and that $K = qk_n + (2q-2)\ell_n$, we see that $K\sim(\log L)^A L$. Moreover, $f = (2q-2)\ell_n\left(\frac{8\alpha}{1+8\alpha}\right)^n \sim \frac{2}{2+1/\alpha}K(\log L)^{-B\log\left(1+\frac{1}{8\alpha}\right)} = o\left(K\cdot(\log L)^{-A}\right)$.

Now remember that the probability of having more than f unlocked sites is bounded above by $q\cdot 2^n e^{-d_1 k_0}$. We have $q\sim\frac{1}{2+1/\alpha}(\log L)^A$, $2^n\sim(\log L)^{B\log 2}$, and, writing $\ell_0 = e^{\log\ell_0}$, and using the fact that $\log L\sim\log\ell_0$, and $k_0 = \ell_0/\alpha + O(1)$, we may write $e^{-d_1 k_0}$ as $e^{-L^{1+o(1)}}$, and absorb both smaller order factors $(\log L)^A$ and $(\log L)^{B\log 2}$ into this expression, so that the probability of having more than f unlocked sites is bounded above by $e^{-L^{1+o(1)}}$.

□

Lemma 7. *For any $A>0$, and for arbitrarily large L, there exists a downward $(K,L)-$coupling with the following properties as $L\to+\infty$:*

- *$L/K \sim h(\log L)^{-A}$ for some constant $h>0$*
- *the probability that coalescence does not occur is bounded above by $e^{-L^{1+o(1)}}$.*

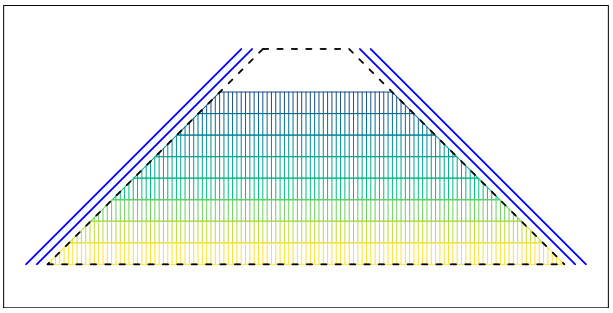


FIGURE 4. Slicing of $\mathbf{T}$ by consecutive trapezoids $\mathbf{T}_i$, depicted in various colours. Boundary conditions are shown in blue.

Proof. Apply Lemma 6 to find a $(K_0, L_0)-$coupling for an arbitarily large L_0, with $L_0/K_0 \sim (\log L_0)^{-A}$ as $L_0 \to +\infty$, and let $t = \lceil (h-1)K_0 \cdot (\log L_0)^{-A} \rceil$ for a certain constant $h > 1 + (\log |\mathscr{A}|)/b$. We then let the resulting K_0 sites at the base evolve according to a type II $(K_0 - 2t, t)$-coupling (see Lemma 2), independent from the previous $(K_0, L_0)-$coupling.

Conditional upon the $(K_0, L_0)-$coupling, when the number of locked sites is less than $K_0 \cdot (\log L_0)^{-A}$, there are at most $|\mathscr{A}|^{K_0 \cdot (\log L_0)^{-A}}$ distinct initial configurations fed into the top of the type II $(K_0 - 2t, t)-$coupling. In such a case, by Lemma 2, the (conditional) probability that coalescence does not occur within the $(K_0 - 2t, t)-$coupling is bounded above by $|\mathscr{A}|^{K_0 \cdot (\log L_0)^{-A}} (K_0 - 2t) 2a e^{-bt}$, which rewrites as $e^{-L_0^{1+o(1)}}$ since $h > (\log |\mathscr{A}|)/b$ and $L_0/K_0 \sim (\log L_0)^{-A}$ On the other hand, by Lemma 6, the probability that the number of locked sites exceeds $K_0 \cdot (\log L_0)^{-A}$ in the $(K_0, L_0)-$coupling is also bounded above by $e^{-L_0^{1+o(1)}}$.

We have thus built a downward $(K, L)-$coupling with $K = K_0 - 2t$ and $L = L_0 + t$, with $K \sim K_0$ and $L \sim hL_0$, and so $L/K \sim h(\log L_0)^{-A} \sim h(\log L)^{-A}$. Moreover, the non-coalescence probability is bounded above by $e^{-L_0^{1+o(1)}}$, and so by $e^{-L^{1+o(1)}}$. □

We now consider couplings for the dynamics involving boundary conditions within an upward trapezoid.

Lemma 8. *There exists a constant $\theta > 0$ such that, for all large enough L and $M \leq L/(\log L)^\theta$, there is an $(M, L)-$coupling whose non-coalescence probability is bounded above, for any boundary condition, by $e^{-L^{1+o(1)}}$ as $L \to +\infty$, where the $o(1)$ is uniform over M and over the boundary condition.*

Proof. We define a coupling of the dynamics within an upward trapezoid $\mathbf{T}$ with height L and top length M, for a given boundary condition χ, assuming that $M \leq L$. Let $t = \lceil 2(\log L)/b \rceil$, and, for $\lceil (\log L)^3 \rceil \leq i \leq q$, where $q = \lfloor L/t \rfloor$, consider the slice of $\mathbf{T}$ formed by the upward trapezoid $\mathbf{T}_i$ with height t and base length $M + 2it$ (see Fig. 4).

For each i, we divide the base of $\mathbf{T}_i$ into seven consecutive intervals $I_1, \ldots, I_7$ (from left to right), whose lengths g_j are defined as follows: $g_1 = g_3 = g_5 = g_7 = k$, with $k = \lceil (6/b) \log \log L \rceil_2$ (here $\lceil m \rceil_2$ stands for the largest even integer number less than or equal to m), $g_2 = g_6 = \lceil (\log L)^2 \rceil$, $g_4 = 2it - (g_1 + g_2 + g_3 + g_5 + g_6 + g_7)$. (Condition $i \geq (\log L)^3$ ensures that, for all large enough L, we have $g_4 \geq 0$.)

We then put two downward trapezoids $\mathbf{T}_{i,1}$ and $\mathbf{T}_{i,3}$ of height $k/2$ on top of I_2 and I_6 respectively, and a downward trapezoid $\mathbf{T}_{i,2}$ of height t on top of I_4. Observe that these trapezoids do not intersect each other except on their boundaries, and, that for all large enough L, they do not touch the outer boundary of $\mathbf{T}$ (see Fig. 5).

We now define by induction the coupling within $\mathbf{T}$. To begin with, above $\mathbf{T}_{\lceil (\log L)^3 \rceil}$, we use the basic coupling. Then, assuming that the coupling has already been defined above $\mathbf{T}_i$, we do the following within $\mathbf{T}_i$. Outside $\mathbf{T}_{i,1} \cup \mathbf{T}_{i,2} \cup \mathbf{T}_{i,3}$, we use the basic coupling. Since these downward trapezoids do not touch the

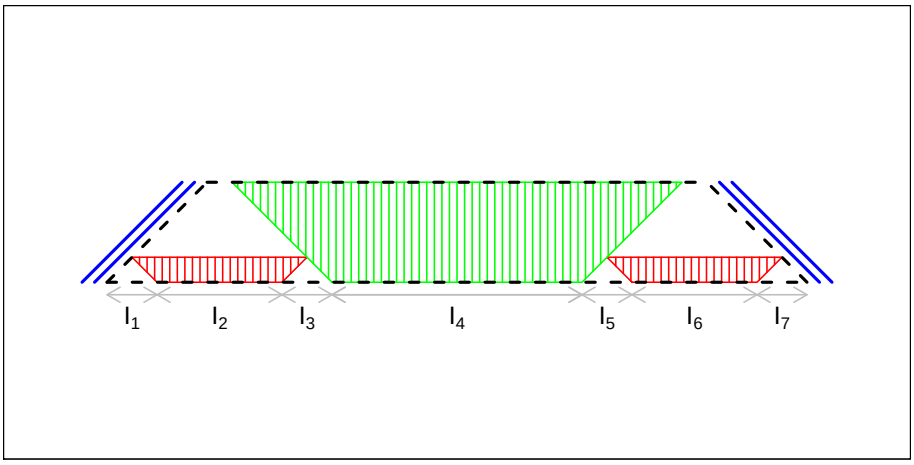


FIGURE 5. Within a trapezoid $\mathbf{T}_i$, intervals $I_1, \dots, I_7$, and trapezoids $\mathbf{T_{i,1}}$ (red), $\mathbf{T_{i,2}}$ (green), $\mathbf{T_{i,3}}$ (red). Boundary conditions are shown in blue.

outer boundary, the dynamics within them do not involve the boundary condition. Moreover, their bases and heights have been chosen in such a way that, for all large L, $2a \cdot \text{base length} \cdot e^{-b \cdot \text{height}} \leq 1/4$, say.

As a consequence, within $\mathbf{T}_{i,j}$ for $j = 1, 2, 3$, Lemma 2 provides a type II coupling such that, for any pair of configurations at the top of $\mathbf{T}_{i,j}$, the probability that they do not lead to the same configuration at the base of $\mathbf{T}_{i,j}$ is bounded above by $2 \cdot 1/4 = 1/2$. Since $\mathbf{T}_{i,j}$, for $j = 1, 2, 3$ do not touch each other except on their boundaries, we may use these couplings independently within $\mathbf{T}_{i,1}, \mathbf{T}_{i,2}, \mathbf{T}_{i,3}$. There remain less than $4k$ sites within $I_1 \cup I_3 \cup I_5 \cup I_7$ that do not belong the the bases of $\mathbf{T}_{i,1}, \mathbf{T}_{i,2}, \mathbf{T}_{i,3}$. Invoking the positive rates property of our PCA in conjunction with the basic coupling outside $\mathbf{T}_{i,1} \cup \mathbf{T}_{i,2} \cup \mathbf{T}_{i,3}$, see (6), the probability to have every such site in a certain state $w \in \mathscr{A}$, for every configuration at the top of $\mathbf{T}_i$, is bounded below by κ^{4k}, independently of what happens within $\mathbf{T}_{i,1} \cup \mathbf{T}_{i,2} \cup \mathbf{T}_{i,3}$. As a result, the probability of having the same pair of configurations on the base of $\mathbf{T}_i$ is bounded below by $(1/2)^3 \cdot \kappa^{4k}$.

The coupling is now defined on the whole of $\mathbf{T}$. Starting from a pair of configurations ζ_1, ζ_2 at the top of $\mathbf{T}$, the probability that all of the $q - \lceil (\log L)^3 \rceil$ trapezoids $\mathbf{T}_i$ fail to produce the same pair of configurations on their base, is bounded above by $\left(1 - \kappa^{4k}/8\right)^{q - \lceil (\log L)^3 \rceil}$. As soon as $\eta > 1 + \log(1/\kappa)(24/b)$, this quantity is bounded above by $e^{-L/(\log L)^\eta}$ for all large enough L.

Since there are $|\mathscr{A}|^{M+1}$ distinct initial conditions, using the union bound exactly as in the proof of Lemma 2, the probability of non-coalescence of this coupling is bounded above by $|\mathscr{A}|^{M+1} e^{-L/(\log L)^\eta}$. Choosing any $\theta > \eta$, the inequality $M \leq L/(\log L)^\theta$ yields the desired bound on the coalescence probability, and this bound is uniform over χ. To get a coupling defined for every boundary condition, we use a version of the coupling just defined for every χ, drawn independently over the various values of χ.

□

Now consider the following construction (see Fig. 6): starting from an integer L, set $K = 2L$, and put side-by-side (from left to right) two downward trapezoids $\mathbf{T}_a, \mathbf{T}_c$ with height L and top length K, and put in between an upward trapezoid $\mathbf{T}_b$ with height $L - 1$ and base length $K - 2$. Since $K = 2L$, these three trapezoids are in fact triangles.

Lemma 9. *For arbitrarily large L, there exists a (K, L)-coupling within $\mathbf{T} = \mathbf{T}_a \cup \mathbf{T}_b \cup \mathbf{T}_c$ such that:*

- *The dynamics within $\mathbf{T}_a$ and $\mathbf{T}_c$ are given by two i.i.d. $(K, L)-$couplings.*

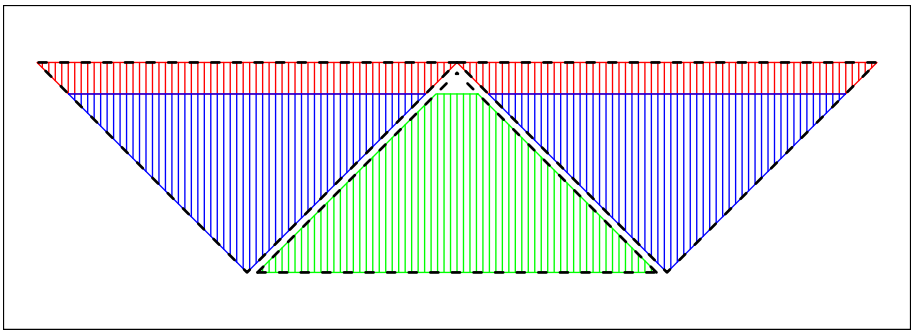


FIGURE 6. Here, the boundaries of $\mathbf{T}_a$ (left), $\mathbf{T}_b$ (middle), $\mathbf{T}_c$ (right) are drawn with dashed lines. Then $\mathbf{T}_1$ (left) and $\mathbf{T}_2$ (right) are drawn in red, while $\mathbf{T}_3$ is drawn in green, and $\mathscr{T}_1$ (left) and $\mathscr{T}_2$ (right) are drawn in blue.

- *The dynamics within $\mathbf{T}_b$ is given by an $(M, L-1)$ coupling with $M = 0$, independent from the above two $(K, L)-$couplings.*
- *The non-coalescence probability of the overall coupling is bounded above by $e^{-L^{1+o(1)}}$.*

Proof. Remember the constant θ from Lemma 8, and let $A > \theta$. Now consider an integer L_1 (which can be chosen to be arbitrarily large) to which we apply Lemma 7, yielding a downward $(K_1, L_1)-$coupling. Note that, by choosing q even in the proof of Lemma 6, we may assume that K_1 is an even number, and let $L = L_1 + K_1/2$ and $K = 2L$. Then let $L_2 = L - L_1$, and $M_2 = 2L_1 - 2$. For large L_1, we have that $K_1 \sim h^{-1} L_1 (\log L_1)^A$. We deduce that $L_2 \sim (h^{-1}/2) L_1 (\log L_1)^A$, and $\log L_2 \sim \log L_1$, so that $L_2/(\log L_2)^\theta \sim (h^{-1}/2) L_1 (\log L_1)^{A-\theta}$. Since $M_2 = 2L_1 - 2$ and $A > \theta$, we see that, for all large enough L_1, $M_2 \leq L_2/(\log L_2)^\theta$ so that we may use Lemma 8 to provide an upward $(M_2, L_2)-$coupling.

Now (see Fig. 6) put side-by-side two trapezoids downward $\mathbf{T}_1$, $\mathbf{T}_2$ with height L_1 and base length K_1. Then insert between $\mathbf{T}_1$ and $\mathbf{T}_2$ an upward trapezoid $\mathbf{T}_3$ whose top has the same ordinate as the base of $\mathbf{T}_1$, $\mathbf{T}_2$, with height L_2 and base length $K - 2$ (so that the top-length is M_2). Next, draw two triangles $\mathscr{T}_1$ and $\mathscr{T}_2$ just below respectively $\mathbf{T}_1$ and $\mathbf{T}_2$, so that the top of $\mathscr{T}_1$ (resp. $\mathscr{T}_2$) coincides with the base of $\mathbf{T}_1$ (resp. $\mathbf{T}_2$). Finally, let $\mathbf{T}_a = \mathbf{T}_1 \cup \mathscr{T}_1$, $\mathbf{T}_c = \mathbf{T}_2 \cup \mathscr{T}_2$, and let $\mathbf{T}_b$ denote the triangle located between $\mathbf{T}_a$ and $\mathbf{T}_c$, whose boundary is at horizontal distance 1 from these.

We use the downward $(K_1, L_1)-$coupling provided by Lemma 7, independently within $\mathbf{T}_1$ and $\mathbf{T}_2$. Within $\mathscr{T}_1$ and $\mathscr{T}_2$, and also within the triangle located between $\mathbf{T}_1$ and $\mathbf{T}_2$, we use the basic coupling. Finally, within $\mathbf{T}_3$, we use the upward $(M_2, L_2)-$coupling from Lemma 8, independently from the couplings used within $\mathbf{T}_1$ and $\mathbf{T}_2$.

The outer boundary of $\mathbf{T}_3$ is included in $\mathscr{T}_1 \cup \mathscr{T}_2$, so the boundary values for the dynamics in $\mathbf{T}_3$ are determined by the couplings we have already defined.

When there is coalescence within $\mathbf{T}_1$ and $\mathbf{T}_2$, only one (random) boundary condition appears at the outer boundary of $\mathbf{T}_3$, depending solely on the coupling within $\mathbf{T}_a$ and $\mathbf{T}_c$, so that, when in addition there is coalescence within $\mathbf{T}_3$ for this specific boundary condition, there is coalescence within $\mathbf{T}$. Since the coupling within $\mathbf{T}_3$ is independent from the couplings within $\mathbf{T}_a$ and $\mathbf{T}_b$, and since the bound on the non-coalescence probability of the coupling provided by Lemma 8 is uniform with respect to the boundary condition, the probability not to have coalescence, conditional upon the fact that there is coalescence within $\mathbf{T}_1$ and $\mathbf{T}_2$, is bounded above by $e^{-L_2^{1+o(1)}}$. Since the probability not to have coalescence within $\mathbf{T}_1$ or within $\mathbf{T}_2$ is also bounded above by $e^{-L_1^{1+o(1)}}$, and since both $\log L_1 \sim \log L$ and $\log L_2 \sim \log L$, we conclude that the overall probability of non-coalescence is bounded above by $e^{-L^{1+o(1)}}$. □

We are now ready to prove Theorem 1.

Proof of Theorem 1. Remember the definition of $\mathbf{T}$, $\mathbf{T}_a$, $\mathbf{T}_b$ from Lemma 9. We start with a triangle $\mathbf{T}_a$ whose top is $[\![-K,0]\!]\times\{L\}$, so that the base of $\mathbf{T}_b$ is $[\![-(L-1),(L-1)]\!]\times\{0\}$, and tile the whole lattice $\mathbb{Z}\times(-\mathbb{N})$ by translating $\mathbf{T}_a$ and $\mathbf{T}_b$ with vectors of the form $\lambda_1(2L,0)+\lambda_2(0,L)$, where $\lambda_1\in\mathbb{Z}$ and $\lambda_2\in-\mathbb{N}$. We then define a flow $\Phi_{t_n}^{t_{n+1}}$ for $t_n=-n\cdot L$ by using i.i.d. copies of the couplings provided by Lemma 9, respectively for every copy of $\mathbf{T}_a$, and every copy of $\mathbf{T}_b$. Properties (ii) and (iii) of the theorem are then direct consequences of the definition.

We now prove (i). Given $x\in\mathbb{Z}$, let z denote the element of $\mathbb{Z}$ of the form $2kL$ closest to x, where $k\in\mathbb{Z}$, and define $I(x)=[\![z-2L,z+2L]\!]$. For $n\geq 0$, we let $J(x,n)=\emptyset$ if there is coalescence within $\mathbf{T}+(z,-Ln)$, and $J(x,n)=I(x)$ otherwise. The sets $J(x,n)$ have been defined in such a way that, for all $m\geq n+1$,

$$\pi_{J(x,n)}\circ\Phi_{t_{n+1}}^{t_m}\text{ is a constant function}\Rightarrow\pi_x\circ\Phi_{t_n}^{t_m}\text{ is a constant function }, \tag{7}$$

with the convention that $\Phi_{t_{n+1}}^{t_{n+1}}$ is the identity function.

We then define a random sequence $(P_n)_{n\geq 0}$ of finite subsets of $\mathbb{Z}$, in the following way. We start with $P_0=\{x\}$. Then, assuming $P_0,\dots,P_n$ have already been defined, we let $P_{n+1}=\bigcup_{x\in P_n}J(x,n)$.

One checks by induction using (7) that, if $P_n=\emptyset$, then $\pi_x\circ\Phi_{t_0}^{t_n}$ is a constant function, so that we have the bound $\mathbb{P}(T_x>nL)\leq\mathbb{P}(P_n\neq\emptyset)$.

Now observe that $|P_{n+1}|\leq\sum_{y\in\mathbb{Z}}\mathbf{1}(y\in P_n)|J(y,n)|$. Moreover, for fixed y, $\mathbf{1}(y\in P_n)$ is measurable with respect to $\sigma(\Phi_{t_{k-1}}^{t_k},1\leq k\leq n)$, while $|J(y,n)|$ is measurable with respect to $\sigma(\Phi_{t_n}^{t_{n+1}})$. Since the random functions $(\Phi_{t_{k-1}}^{t_k})_{k\geq 1}$ form an independent sequence, $\mathbf{1}(y\in P_n)$ and $|J(y,n)|$ are independent, so we have that $\mathbb{E}(|P_{n+1}|)\leq\sum_{y\in\mathbb{Z}}\mathbb{P}(y\in P_n)\mathbb{E}(|J(y,n)|)$.

In view of the definition of $J(y,n)$, we see that $\mathbb{E}(|J(y,n)|)=(4L+1)\cdot\mathbb{P}(\text{non-coalescence in }\mathbf{T})=\rho$. As a consequence, we deduce that $\mathbb{E}(|P_{n+1}|)\leq\rho\cdot\sum_{y\in\mathbb{Z}}\mathbb{P}(y\in P_n)=\rho\cdot\mathbb{E}(|P_n|)$. Iterating this inequality, we deduce that $\mathbb{E}(|P_n|)\leq\rho^n$.

Since the non-coalescence probability is bounded above by $e^{-L^{1+o(1)}}$, ρ can be made arbitrarily small by choosing a large enough value of L, and we indeed assume that ρ is <1. Using the Markov inequality and the fact that $|P_n|$ is an integer number, we have the following sequence of inequalities, which proves (i):

$$\mathbb{P}(T_x>nL)\leq\mathbb{P}(P_n\neq\emptyset)=\mathbb{P}(|P_n|>0)=\mathbb{P}(|P_n|\geq 1)\leq\mathbb{E}(|P_n|)\leq\rho^n.$$

□

Proof of Corollary 1. Assume that $\mathcal{K}$ is a transition kernel defining an exponentially ergodic PCA with positive rates, and let $\mathcal{K}'$ denote a transition kernel distinct from $\mathcal{K}$. Given $\epsilon\in]0,1[$, assume that $\mathcal{K}'$ is close enough to $\mathcal{K}$ so that, for any $\mathbf{v}\in\mathscr{A}^{\{-1,0,1\}}$ and $w\in\mathscr{A}$, $\mathcal{K}'(\mathbf{v},\{w\})\geq(1-\epsilon)\mathcal{K}(\mathbf{v},\{w\})$. Letting $\mathcal{K}''(\mathbf{v},\{w\})=\frac{1}{\epsilon}(\mathcal{K}'(\mathbf{v},\{w\})-(1-\epsilon)\mathcal{K}(\mathbf{v},\{w\}))$, we see that $\mathcal{K}''$ is a transition kernel.

We now reuse the tiling of $\mathbb{Z}\times(-\mathbb{N})$ with translated copies of $\mathbf{T}_a$ and $\mathbf{T}_b$ used to prove Theorem 1, and define a coupling for the dynamics of $\mathcal{K}'$ as follows. Within each copy of $\mathbf{T}_a$, declare each site in $\mathbf{T}_a\setminus\mathrm{top}(\mathbf{T}_a)$ to be blue with probability $1-\epsilon$ and red with probability ϵ, independently for each site. If all sites are blue, we use within $\mathbf{T}_a$ the coupling defined for the $\mathcal{K}-$dynamics in the proof of Theorem 1. If at least one site is red, we use a version of the basic coupling where red sites use $\mathcal{K}''$ while blue sites use $\mathcal{K}$. A similar construction is done for each copy of $\mathbf{T}_b$.

We now redo the construction of the sets $(P_n)_{n\geq 0}$ used in the proof of Theorem 1 with the following modification: $J(x,n)=\emptyset$ if there is coalescence within $\mathbf{T}+(z,Ln)$ and all sites in $\mathbf{T}+(z,Ln)$ are blue, and $J(x,n)=I(x)$ otherwise. As a consequence,

$$\mathbb{E}(|J(y,n)|)\leq(4L+1)\cdot((1-\epsilon)^m p+(1-(1-\epsilon)^m))\,,$$

where $m=|\mathbf{T}\setminus\mathrm{top}(\mathbf{T})|$ and $p=\mathbb{P}(\text{non-coalescence of the }\mathcal{K}-\text{dynamics in }\mathbf{T})$. For large enough L, we have that $(4L+1)p<1$. For such an L, noting that m depends only on L and not on ϵ, we see that, for all ϵ small

enough so that $\mathbb{E}(|J(y,n)|) < 1$, the same argument as in the proof of Theorem 1 leads to the conclusion that the coalescence time T'_x for the $\mathcal{K}'-$dynamics has a finite exponential moment uniformly bounded over x. □

References

[1] K. Anand and M. Jerrum, *Perfect sampling in infinite spin systems via strong spatial mixing*, arXiv:2106.15992, (2021).

[2] A. Bušić, J. Mairesse, and I. Marcovici, *Probabilistic cellular automata, invariant measures, and perfect sampling*, Adv. in Appl. Probab., 45 (2013), pp. 960–980.

[3] R. Fernández, P.-Y. Louis, and F. R. Nardi, *Overview: PCA models and issues*, in Probabilistic cellular automata, vol. 27 of Emerg. Complex. Comput., Springer, Cham, 2018, pp. 1–30.

[4] T. M. Liggett, *Interacting particle systems*, vol. 276 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], Springer-Verlag, New York, 1985.

[5] I. Marcovici, M. Sablik, and S. Taati, *Ergodicity of some classes of cellular automata subject to noise*, Electron. J. Probab., 24 (2019), pp. Paper No. 41, 44.

[6] F. Martinelli, E. Olivieri, and R. H. Schonmann, *For 2-D lattice spin systems weak mixing implies strong mixing*, Comm. Math. Phys., 165 (1994), pp. 33–47.

[7] J. G. Propp and D. B. Wilson, *Exact sampling with coupled Markov chains and applications to statistical mechanics*, in Proceedings of the Seventh International Conference on Random Structures and Algorithms (Atlanta, GA, 1995), vol. 9, 1996, pp. 223–252.

[8] Y. Spinka, *Finitary coding for the sub-critical Ising model with finite expected coding volume*, Electron. J. Probab., 25 (2020), pp. Paper No. 8, 27.

[9] ———, *Finitary codings for spatial mixing Markov random fields*, Ann. Probab., 48 (2020), pp. 1557–1591.

[10] A. Toom, N. Vasilyev, O. Stavskaya, L. Mityushin, G. Kurdyumov, and S. Pirogov, *Discrete local markov systems*, in Stochastic Cellular Systems: ergodicity, memory, morphogenesis, R. Dobrushin, V. Kryukov, and A. Toom, eds., Manchester University Press, 1990.

[11] J. van den Berg and J. E. Steif, *On the existence and nonexistence of finitary codings for a class of random fields*, Ann. Probab., 27 (1999), pp. 1501–1522.

Institut de Recherche Mathématique Avancée, UMR 7501, Université de Strasbourg et CNRS, 7 rue René Descartes, 67 000 Strasbourg, France

Email address: jberard@unistra.fr