

Classification of backward filtrations and factor filtrations: examples from cellular automata

PAUL LANTHIER and THIERRY DE LA RUE 

*Laboratoire de Mathématiques Raphaël Salem, Université de Rouen, CNRS,
Avenue de l'Université, F-76801 Saint Étienne du Rouvray, France
(e-mail: planthier76@outlook.fr, Thierry.de-la-Rue@univ-rouen.fr)*

(Received 16 March 2021 and accepted in revised form 11 June 2021)

Abstract. We consider backward filtrations generated by processes coming from deterministic and probabilistic cellular automata. We prove that these filtrations are standard in the classical sense of Vershik's theory, but we also study them from another point of view that takes into account the measure-preserving action of the shift map, for which each sigma-algebra in the filtrations is invariant. This initiates what we call the dynamical classification of factor filtrations, and the examples we study show that this classification leads to different results.

Key words: filtrations, I-cosiness, measure-theoretic factors, cellular automata

2020 Mathematics Subject Classification: 60J05, 37A35 (Primary); 37B15 (Secondary)

1. Introduction

The topic of this paper is the study of some backward filtrations, or filtrations in discrete negative time, which are non-decreasing sequences of the form $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ where each \mathcal{F}_n is a sub-sigma-algebra on a given probability space. In the sequel, we will simply refer to them as *filtrations*.

In this work, all measure spaces are implicitly assumed to be Polish spaces equipped with their Borel sigma-algebra, and we only consider random variables taking their values in such spaces. If X is a random variable defined on some probability space (Ω, \mathbb{P}) , taking values in a Polish space E , the *law* of X is the probability measure $\mathcal{L}(X)$ on E which is the pushforward measure of \mathbb{P} by the measurable map X . We call a *copy* of X any random variable X' , possibly defined on another probability space, such that $\mathcal{L}(X') = \mathcal{L}(X)$ (in particular, X' takes its values in the same space as X). Given a family $(X_m)_{m \in I}$ of random variables defined on a given probability space (Ω, \mathbb{P}) , we denote by $\Sigma(X_m : m \in I)$ the sub-sigma algebra generated by the random variables X_m , $m \in I$ and the negligible sets. All sigma-algebras are supposed to be complete and *essentially*

separable: they are generated (modulo negligible sets) by countably many events (or, equivalently, by countably many random variables).

Therefore, each filtration $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ we consider can be generated by a *process in negative time* (that is to say, a family $X = (X_n)_{n \leq 0}$ of random variables), which means that for each $n \leq 0$, $\mathcal{F}_n = \Sigma(X_m : m \leq n)$. The filtration \mathcal{F} is then the mathematical object that describes the acquisition of information as the process $(X_n)_{n \leq 0}$ evolves from $-\infty$ up to the present time. Different processes can generate the same filtration, and the classification of filtrations is roughly equivalent to answering the following question: given a filtration \mathcal{F} , can we find a ‘nice’ process generating it?

In this direction, the simplest structure we can have is called a *filtration of product type*, which means a filtration which can be generated by a sequence $X = (X_n)_{n \leq 0}$ of *independent* random variables. Then there is the slightly more general notion of *standard filtration*: a filtration which is *immersible* into a filtration of product type (see Definition 2.3 below).

Classification of backward filtrations was initiated by Vershik in the 1970s [16]. His work, which was written in the language of ergodic theory, did not receive much attention until a later publication in the 1990s [17]. Then some authors started to bring the subject into probability theory where they found nice applications (see, in particular, [5, 6, 15]). In the present paper, among other things we propose a return trip to ergodic theory by considering special families of filtrations on measure-theoretic dynamical systems (such systems are given by the action of a measure-preserving transformation T on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$). Filtrations in measure-theoretic dynamical systems have already been considered by many authors (see, for example, [2, 3, 8, 9]), but in situations where the transformation acts ‘in the direction of the filtration’: by this we mean that, in these works, the filtrations satisfy $\mathcal{F}_{n-1} = T^{-1}\mathcal{F}_n \subsetneq \mathcal{F}_n$ for each $n \leq 0$. Here we adopt a transverse point of view: we consider examples where the sigma-algebras forming the filtration are invariant with respect to the underlying transformation: $\mathcal{F}_n = T^{-1}\mathcal{F}_n$ for each $n \leq 0$. We get what we call *factor filtrations*, and we would like to classify these objects by taking into account the transverse action of the transformation. To emphasize the role of the underlying dynamical system, we will call the classification of factor filtrations the *dynamical classification* of factor filtrations. In contrast, we will refer to the usual classification of filtrations as the *static* classification.

The examples that we consider in this paper are all derived from the theory of cellular automata, where the underlying measure-preserving transformation is the shift of coordinates. We provide in this context natural examples of factor filtrations, some of which show a different behaviour depending on whether we look at them from a static or a dynamical point of view.

1.1. *Examples of filtrations built from cellular automata.* We define in this section the two filtrations which will be studied in the paper. The first is built from an algebraic deterministic cellular automaton τ , and the second from a probabilistic cellular automaton (PCA) which is a random perturbation of τ , depending on a parameter ε and denoted by τ_ε . Both automata are one-dimensional (cells are indexed by \mathbb{Z}) and the state of each cell is an element of a fixed finite Abelian group $(\mathbb{A}, +)$, which we assume non-trivial: $|\mathbb{A}| \geq 2$.

Throughout the paper the index n is viewed as the time during which the cellular automata evolve (represented vertically from top to bottom in the figures), whereas the index i should be interpreted as the position within a given configuration (represented horizontally). For a given configuration $x = (x(i) : i \in \mathbb{Z})$ and $j \leq k \in \mathbb{Z}$, we denote by $x[j, k]$ the restriction $(x(i) : j \leq i \leq k)$ of x to the sites between j and k .

1.1.1. *The deterministic cellular automaton τ and the filtration $\mathcal{F}^{\tau^{-1}}$.* We first define the deterministic cellular automaton $\tau : \mathbb{A}^{\mathbb{Z}} \rightarrow \mathbb{A}^{\mathbb{Z}}$. Given a configuration $x = (x(i))_{i \in \mathbb{Z}}$, the image configuration τx is given by

$$\text{for all } i \in \mathbb{Z}, \quad \tau x(i) := x(i) + x(i + 1). \tag{1}$$

Observe that τ can be written as $\tau = \sigma + \text{Id}$, where σ is the left shift on $\mathbb{A}^{\mathbb{Z}}$.

For any finite set S , we denote by \mathcal{U}_S the uniform probability measure on S . We consider on $\mathbb{A}^{\mathbb{Z}}$ the product probability measure $\mu := \mathcal{U}_{\mathbb{A}}^{\otimes \mathbb{Z}}$, which is the normalized Haar measure on the compact Abelian group $\mathbb{A}^{\mathbb{Z}}$. The measure μ is invariant by τ .

On the probability space $(\mathbb{A}^{\mathbb{Z}}, \mu)$, we construct a negative-time process $(X_n)_{n \leq 0}$ as follows: we first consider the random variable X_0 taking values in $\mathbb{A}^{\mathbb{Z}}$ and simply defined by the identity on $\mathbb{A}^{\mathbb{Z}}$. Then we set for each integer $n \leq 0$,

$$X_n := \tau^{|n|} X_0.$$

Since the law μ of X_0 is preserved by τ , the process $(X_n)_{n \leq 0}$ defined in this way is stationary. In particular, all the random variables $X_n, n \leq 0$, have the same law μ .

We denote by $\mathcal{F}^{\tau^{-1}}$ the filtration generated by this process $(X_n)_{n \leq 0}$: for each $n \leq 0$, $\mathcal{F}_n^{\tau^{-1}} = \Sigma(X_m : m \leq n) = \Sigma(X_m)$. We use τ^{-1} instead of τ in this notation to emphasize the fact that the time of the process generating this filtration goes in the other direction than that of the cellular automaton τ .

1.1.2. *The probabilistic cellular automaton τ_ε and the filtration $\mathcal{F}^{\tau_\varepsilon}$.* We also introduce a random perturbation of the deterministic cellular automaton τ , depending on a parameter ε which is the probability of making an error in the computation of the new state of a given cell. More precisely, we fix $\varepsilon > 0$ and we define τ_ε as the Markov kernel on $\mathbb{A}^{\mathbb{Z}}$ given by the following: for each $x \in \mathbb{A}^{\mathbb{Z}}$, $\tau_\varepsilon(x, \cdot)$ is the law of a random variable of the form $\tau x + \xi$, where ξ is a random error, with law

$$\mathcal{L}(\xi) = \bigotimes_{i \in \mathbb{Z}} \left((1 - \varepsilon)\delta_{0_{\mathbb{A}}} + \frac{\varepsilon}{|\mathbb{A}| - 1} \sum_{a \in \mathbb{A} \setminus \{0_{\mathbb{A}}\}} \delta_a \right).$$

In the sequel, we will always assume that

$$\varepsilon < \frac{|\mathbb{A}| - 1}{|\mathbb{A}|}, \tag{2}$$

which is equivalent to $1 - \varepsilon > \varepsilon / (|\mathbb{A}| - 1)$. In other words, we give more weight to $0_{\mathbb{A}}$. Setting

$$\tilde{\varepsilon} := \varepsilon \frac{|\mathbb{A}|}{|\mathbb{A}| - 1} < 1, \tag{3}$$

we can also write the law of the random error ξ as

$$\mathcal{L}(\xi) = \bigotimes_{i \in \mathbb{Z}} \left((1 - \tilde{\varepsilon})\delta_{0_{\mathbb{A}}} + \frac{\tilde{\varepsilon}}{|\mathbb{A}|} \sum_{a \in \mathbb{A}} \delta_a \right). \tag{4}$$

The following lemma shows that the probability measure μ on $\mathbb{A}^{\mathbb{Z}}$ defined in §1.1.1 is invariant by this Markov kernel (in fact, as we will prove later, it is the only one; see Corollary 2.17).

LEMMA 1.1. *Let $X = (X(i) : i \in \mathbb{Z})$ and $\xi = (\xi(i) : i \in \mathbb{Z})$ be two independent random variables taking values in $\mathbb{A}^{\mathbb{Z}}$, where $\mathcal{L}(X) = \mu$. Then $\mathcal{L}(\tau X + \xi) = \mu$.*

Proof. We already know that $\mathcal{L}(\tau X) = \mu$. By independence of τX and ξ , and since μ is the Haar measure on $\mathbb{A}^{\mathbb{Z}}$, for $a \in \mathbb{A}^{\mathbb{Z}}$ we have

$$\mathcal{L}(\tau X + \xi | \xi = a) = \mathcal{L}(\tau X + a) = \mu. \quad \square$$

(Note that the proof of the lemma is valid regardless of the law of the error ξ , and that it also proves the independence of $\tau X + \xi$ and ξ .)

By invariance of μ under τ_{ε} , we can construct on some probability space (Ω, \mathbb{P}) a stationary Markov chain $(X_n)_{n \in \mathbb{Z}}$, where for each $n \in \mathbb{Z}$, X_n takes its values in $\mathbb{A}^{\mathbb{Z}}$, the law of X_n is $\mathcal{L}(X_n) = \mu$, and for each $x \in \mathbb{A}^{\mathbb{Z}}$, the conditional distribution of X_{n+1} given $X_n = x$ is

$$\mathcal{L}(X_{n+1} | X_n = x) = \tau_{\varepsilon}(x, \cdot).$$

We will denote by $\mathcal{F}^{\tau_{\varepsilon}}$ the filtration generated by the negative-time part of such a process: for each $n \leq 0$, $\mathcal{F}_n^{\tau_{\varepsilon}} = \Sigma(X_m : m \leq n)$. If we want to define $\mathcal{F}^{\tau_{\varepsilon}}$ in a canonical way, we can always assume that in this construction, $\Omega = (\mathbb{A}^{\mathbb{Z}})^{\mathbb{Z}}$ and \mathbb{P} is the law of the stationary Markov chain on Ω .

2. Usual (static) classification of filtrations

2.1. *A very short abstract of the theory.* We recall the main points in the usual theory of classification of filtrations. We refer to [6, 11] for details.

We start with the following definition of isomorphic filtrations, which is of course equivalent to the definition provided in the cited references.

Definition 2.1. (Isomorphism of filtrations) Let \mathcal{F} and \mathcal{F}' be two filtrations, possibly defined on two different probability spaces. We say that they are *isomorphic* if we can find two processes $X = (X_n)_{n \leq 0}$ and $X' = (X'_n)_{n \leq 0}$ with the same law generating respectively \mathcal{F} and \mathcal{F}' . We write in this case $\mathcal{F} \sim \mathcal{F}'$, or $\mathcal{F} \stackrel{X, X'}{\sim} \mathcal{F}'$ to specify the processes involved in the isomorphism.

It may not seem obvious in the above formulation that this notion of isomorphism is transitive. For details on this point, see [10, §1.2.1].

Let \mathcal{F} and \mathcal{F}' be two isomorphic filtrations, and let X and X' be two processes such that $\mathcal{F} \stackrel{X, X'}{\sim} \mathcal{F}'$. If Y is an \mathcal{F}_0 -measurable random variable, it is a classical result in

measure theory that there exists some measurable map ϕ satisfying $Y = \phi(X)$ almost surely (a.s.). Then the isomorphism implicitly provides an \mathcal{F}'_0 -measurable copy of Y , namely the random variable $Y' := \phi(X')$. Of course, this copy depends on the choice of the processes X and X' used in the isomorphism, but once these processes are fixed, the copy Y' of Y is a.s. unique (because the map ϕ is a.s. unique with respect to the law of X).

Let \mathcal{A} , \mathcal{B} and \mathcal{C} be three sub-sigma-algebras on the same probability space, with $\mathcal{C} \subset \mathcal{A} \cap \mathcal{B}$. We recall that \mathcal{A} and \mathcal{B} are *independent conditionally on \mathcal{C}* if, whenever X and Y are bounded real random variables, respectively \mathcal{A} and \mathcal{B} measurable, we have

$$\mathbb{E}[XY|\mathcal{C}] = \mathbb{E}[X|\mathcal{C}] \mathbb{E}[Y|\mathcal{C}] \quad (\text{a.s.}).$$

We note in this case that

$$\mathcal{A} \overset{\mathcal{C}}{\perp\!\!\!\perp} \mathcal{B}.$$

Definition 2.2. (Immersion) Let \mathcal{F} and \mathcal{G} be two filtrations on the same probability space. We say that \mathcal{F} is *immersed* in \mathcal{G} if the following conditions are satisfied:

- for each $n \leq 0$, $\mathcal{F}_n \subset \mathcal{G}_n$ (\mathcal{F} is included in \mathcal{G});
- for each $n \leq -1$, $\overset{\mathcal{F}_n}{\mathcal{F}_{n+1}} \perp\!\!\!\perp \mathcal{G}_n$.

We write in this case $\mathcal{F} \prec \mathcal{G}$.

We say that \mathcal{F} and \mathcal{G} are *jointly immersed* if both \mathcal{F} and \mathcal{G} are immersed in the filtration $\mathcal{F} \vee \mathcal{G}$.

Definition 2.3. (Immersibility) Let \mathcal{F} and \mathcal{G} be two filtrations, possibly defined on different probability spaces. We say that \mathcal{F} is *immersible* in \mathcal{G} if there exists a filtration \mathcal{F}' immersed in \mathcal{G} such that \mathcal{F}' and \mathcal{F} are isomorphic. We write in this case $\mathcal{F} \prec\sim \mathcal{G}$.

We can now define the following basic hierarchy of properties for filtrations.

Definition 2.4. The filtration \mathcal{F} is said to be of *product type* if it can be generated by a process formed of independent random variables.

The filtration \mathcal{F} is said to be *standard* if it is immersible in a filtration of product type.

The filtration \mathcal{F} is said to be *Kolmogorovian* if its tail sigma-algebra

$$\mathcal{F}_{-\infty} := \bigcap_{n \leq 0} \mathcal{F}_n$$

is trivial (it only contains events of probability 0 or 1).

It is a direct consequence of the definitions and of Kolmogorov’s 0–1 law that the following chain of implications holds:

$$\text{Product type} \implies \text{Standard} \implies \text{Kolmogorovian}.$$

A simple example of a standard filtration which is not of product type has been provided by Vinokurov (see [6]). The construction of a Kolmogorovian filtration which is not standard was one of the first spectacular achievements of Vershik in this theory.

2.2. *Filtrations of product type: example of $\mathcal{F}^{\tau^{-1}}$.* We now return to the filtration $\mathcal{F}^{\tau^{-1}}$ defined in §1.1.1, and we use here the same notation as in that section. In particular, the process $X = (X_n)_{n \leq 0}$ generating $\mathcal{F}^{\tau^{-1}}$ is a process where each coordinate X_n takes its values in $\mathbb{A}^{\mathbb{Z}}$, $\mathcal{L}(X_n) = \mu$, and $X_{n-1} = \tau X_n$.

The purpose of this section is to prove the following result.

THEOREM 2.5. *The filtration $\mathcal{F}^{\tau^{-1}}$ is of product type.*

The above theorem is a direct consequence of the two following lemmas.

LEMMA 2.6. *For each sequence $(i_n)_{n \leq 0}$ of integers, the random variables $X_n(i_n)$, $n \leq 0$, are independent.*

LEMMA 2.7. *Consider the sequence $(i_n)_{n \leq 0}$ of integers defined by $i_n := -\lfloor |n|/2 \rfloor$. Then the process $(X_n(i_n))_{n \leq 0}$ generates the filtration $\mathcal{F}^{\tau^{-1}}$.*

Proof of Lemma 2.6. Let us first consider, on a probability space (Ω, \mathbb{P}) , a random variable $Y = (Y(i) : i \in \mathbb{Z})$ taking values in $\mathbb{A}^{\mathbb{Z}}$, and such that $\mathcal{L}(Y) = \mu$. Let us fix some integers i, j, k such that $j \leq k$ and $i \in \{j, \dots, k, k + 1\}$. For each block $w \in \mathbb{A}^{\{j, \dots, k\}}$ and each $a \in \mathbb{A}$, it is straightforward to check from the construction of τ that there exists a unique block $w' \in \mathbb{A}^{\{j, \dots, k, k+1\}}$ such that:

- $w'(i) = a$;
- $(w'(h) + w'(h + 1) : j \leq h \leq k) = w$.

Let us denote this block w' by $\tau_{a,i}^{-1}(w)$. We then have the following equivalence:

$$\tau Y[j, k] = w \iff \text{there exists } a \in \mathbb{A}, Y[j, k + 1] = \tau_{a,i}^{-1}(w).$$

But by construction of μ , all the $|\mathbb{A}|$ events $(Y[j, k + 1] = \tau_{a,i}^{-1}(w))$, $a \in \mathbb{A}$, have the same probability $|\mathbb{A}|^{-(k-j+2)}$ (and are of course disjoint). It follows that, for each $a \in \mathbb{A}$,

$$\mathbb{P}(Y(i) = a \mid \tau Y[j, k] = w) = \mathbb{P}(Y[j, k + 1] = \tau_{a,i}^{-1}(w) \mid \tau Y[j, k] = w) = \frac{1}{|\mathbb{A}|}.$$

Since this holds for any block $w \in \mathbb{A}^{\{j, \dots, k\}}$, this proves that $Y(i)$ is independent of $\tau Y[j, k]$. Then, since this is true for each j, k such that $i \in \{j, \dots, k, k + 1\}$, $Y(i)$ is independent of τY .

Let us apply this with $Y = X_n$ for some $n \leq 0$: we get that, for each $i_n \in \mathbb{Z}$, $X_n(i_n)$ is independent of $\tau X_n = X_{n-1}$. Now if we have an arbitrary sequence $(i_n)_{n \leq 0}$ of integers, we note that for each $n \geq 0$, the random variables $X_m(i_m)$, $m \leq n - 1$, are all $\mathcal{F}_{n-1}^{\tau^{-1}}$ -measurable, hence $X_n(i_n)$ is independent of $\Sigma(X_m(i_m) : m \leq n - 1)$. □

Lemma 2.7 will be derived from the following result (see Figure 1).

LEMMA 2.8. *Consider integers $m \leq n \leq 0$, and let $(i_m, i_{m+1}, \dots, i_n)$ be a finite sequence of integers such that, for each $m \leq \ell \leq n - 1$,*

$$i_{\ell+1} = i_{\ell} \quad \text{or} \quad i_{\ell+1} = i_{\ell} + 1. \tag{5}$$

Then $X_n[i_m, i_m + n - m]$ is measurable with respect to $\Sigma(X_m(i_m), \dots, X_n(i_n))$.

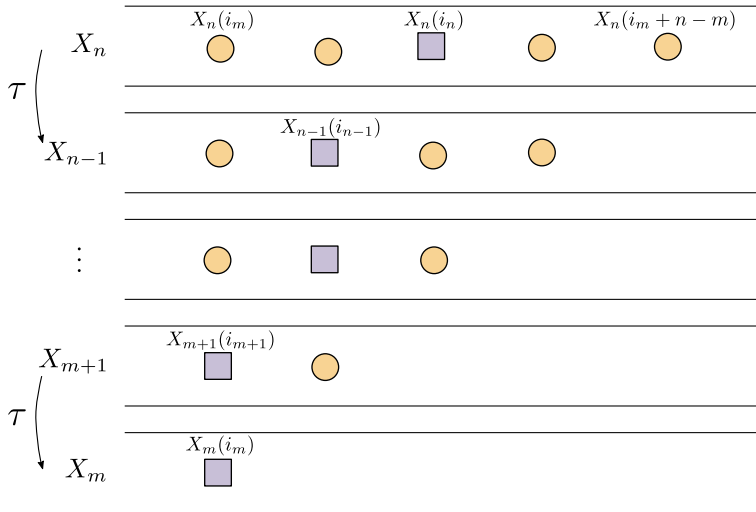


FIGURE 1. Illustration of Lemma 2.8: the variables marked with a circle can be computed from the ones marked with a square.

Proof. For a fixed m , we prove the result by induction on n . If $n = m$ the result is obvious. Assume now that, for some $m + 1 \leq n \leq 0$, the result holds up to $n - 1$. Then $X_{n-1}[i_m, i_m + n - 1 - m]$ is measurable with respect to $\Sigma(X_m(i_m), \dots, X_{n-1}(i_{n-1}))$. Remembering the notation $\tau_{a,i}^{-1}(w)$ from the proof of Lemma 2.6, and since by assumption we have $i_n \in \{i_m, \dots, i_m + n - m\}$, we can then write

$$X_n[i_m, i_m + n - m] = \tau_{X_n(i_n), i_n}^{-1}(X_{n-1}[i_m, i_m + n - 1 - m]),$$

hence $X_n[i_m, i_m + n - m]$ is measurable with respect to $\Sigma(X_m(i_m), \dots, X_n(i_n))$ as claimed. \square

Proof of Lemma 2.7. The sequence (i_n) defined in the statement of Lemma 2.7 obviously satisfies (5). Moreover, we have $i_m \rightarrow -\infty$ as $m \rightarrow -\infty$, and for any fixed $n \leq 0$, $i_m + n - m \rightarrow +\infty$ as $m \rightarrow -\infty$. By application of Lemma 2.8, it follows that X_n is measurable with respect to $\Sigma(X_m(i_m) : m \leq n)$. But conversely, all the random variables $X_m(i_m), m \leq n$, are $\mathcal{F}_n^{\tau^{-1}}$ -measurable. Hence these variables generate $\mathcal{F}_n^{\tau^{-1}}$. \square

2.3. *Standardness and I-cosiness.* The example of the filtration $\mathcal{F}^{\tau^{-1}}$ is very special, as it is not so hard to explain a process with independent coordinates which generates the filtration. In general, when we have a standard filtration, it may be very hard to find such a process from which the filtration is built. This is one of the reasons why several criteria of standardness have been developed, allowing the standardness of a filtration to be proved without explicitly giving the process with independent coordinates. The first such criterion was given by Vershik [17], but here we will be interested in another one, called *I-cosiness*, introduced by Émery and Schachermayer [6] and strongly inspired by ideas of Tsirelson [15] and Smorodinsky [14].

We first have to define the concept of *real-time coupling* for a filtration.

Definition 2.9. (Real-time coupling) Let $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ a filtration on a probability space (Ω, \mathbb{P}) . A *real-time coupling* of \mathcal{F} is a pair $(\mathcal{F}', \mathcal{F}'')$ of filtrations, both defined on the same probability space (but possibly different from (Ω, \mathbb{P})), such that:

- $\mathcal{F}' \sim \mathcal{F}$;
- $\mathcal{F}'' \sim \mathcal{F}$;
- \mathcal{F}' and \mathcal{F}'' are jointly immersed.

Such a real-time coupling of \mathcal{F} is said to be *independent in the distant past* if there exists some integer $n_0 \leq 0$ such that \mathcal{F}'_{n_0} and \mathcal{F}''_{n_0} are independent. In this case, we also say *n₀-independent* if we want to highlight n_0 .

In practice, if we want to construct a real-time coupling of a filtration generated by a process $X = (X_n)_{n \leq 0}$, we have to build on the same probability space two copies $X' = (X'_n)_{n \leq 0}$ and $X'' = (X''_n)_{n \leq 0}$ of X . The joint immersion of the filtrations they generate amounts to the following conditions for each $n \leq -1$:

$$\begin{aligned} \mathcal{L}(X'_{n+1} | \mathcal{F}'_n \vee \mathcal{F}''_n) &= \mathcal{L}(X'_{n+1} | \mathcal{F}'_n), \\ \mathcal{L}(X''_{n+1} | \mathcal{F}'_n \vee \mathcal{F}''_n) &= \mathcal{L}(X''_{n+1} | \mathcal{F}''_n). \end{aligned}$$

So, we can construct such a coupling step by step: assuming that we have already defined X'_m and X''_m for each $m \leq n$, the construction continues at time $n + 1$ with the realization, conditionally on $\mathcal{F}'_n \vee \mathcal{F}''_n$, of a coupling of the two conditional laws $\mathcal{L}(X'_{n+1} | \mathcal{F}'_n)$ and $\mathcal{L}(X''_{n+1} | \mathcal{F}''_n)$. This explains the name *real-time coupling*.

Definition 2.10. (I-cosiness) Let \mathcal{F} be a filtration, and let Y be an \mathcal{F}_0 -measurable random variable taking values in some Polish metric space (E, d) . Then Y is said to be *I-cosy* (with respect to \mathcal{F}) if, for each real number $\delta > 0$, we can find a real-time coupling $(\mathcal{F}', \mathcal{F}'')$ of \mathcal{F} which is independent in the distant past, and such that the two copies Y' and Y'' of Y in \mathcal{F}' and \mathcal{F}'' respectively satisfy

$$d(Y', Y'') < \delta \text{ with probability } > 1 - \delta. \tag{6}$$

The filtration itself is said to be *I-cosy* if each \mathcal{F}_0 -measurable random variable Y is I-cosy with respect to \mathcal{F} .

The variable Y whose copies we want to be close together is called a *target*. Classical arguments of measure theory allow us to considerably reduce the number of targets to test when we want to establish I-cosiness for a given filtration (see [11] or [10, §1.2.4]). In particular, we will use the following proposition.

PROPOSITION 2.11. *Assume that \mathcal{F}_0 is generated by a countable family $(Z_i)_{i \in I}$ of random variables taking values in finite sets. Assume also that we have written the countable set I as an increasing union of finite sets $I_k, k \geq 0$. For each integer $k \geq 0$, denote by Y_k the random variable $(Z_i)_{i \in I_k}$. If Y_k is I-cosy for each k , then the filtration is I-cosy.*

Note that, in the context of the above proposition, since the set of all possible values of Y_k is finite, we can endow it with the discrete metric and replace condition (6) with

$$Y'_k = Y''_k \text{ with probability } > 1 - \delta.$$

The following result is proved in [6] using ideas from [17] (see also [11]).

THEOREM 2.12. *A filtration \mathcal{F} is standard if and only if it is I-cosy.*

2.4. Standardness of the filtration $\mathcal{F}^{\tau_\varepsilon}$. We will apply the I-cosiness criterion to prove the standardness of the filtration $\mathcal{F}^{\tau_\varepsilon}$ defined in §1.1.1, from which we will also be able to derive that this filtration is of product type. We use now the notation introduced in that section: $X = (X_n)_{n \in \mathbb{Z}}$ is a stationary Markov process with transitions given by the Markov kernel τ_ε , and for each n , X_n takes its values in $\mathbb{A}^{\mathbb{Z}}$ and follows the law μ . The filtration $\mathcal{F}^{\tau_\varepsilon}$ is generated by the negative-time part of this process.

THEOREM 2.13. *For each $0 < \varepsilon < (|\mathbb{A}| - 1)/|\mathbb{A}|$, the filtration $\mathcal{F}^{\tau_\varepsilon}$ is I-cosy, hence it is standard.*

Proof. Since $\mathcal{F}^{\tau_\varepsilon}$ is generated by the countable family of random variables $(X_n(i) : n \leq 0, i \in \mathbb{Z})$, it is sufficient by Proposition 2.11 to check that, for each integer $k \geq 0$, the random variable $Y_k := (X_n[-k, k] : -k \leq n \leq 0)$ is I-cosy. In fact, we will see at the end of the proof that it is enough by stationarity of the process X to consider simpler targets, which are the random variables of the form $X_0[-k, k]$, ($k \geq 0$).

So we fix an integer $k \geq 0$, we consider the target $X_0[-k, k]$, and we fix a real number $\delta > 0$. To check the I-cosiness of $X_0[-k, k]$, we have to construct on some probability space (Ω, \mathbb{P}) two copies X' and X'' of the process X , such that:

- for some $n_0 \leq 0$,

$$\mathcal{F}'_{n_0} = \Sigma(X'_m : m \leq n_0) \text{ and } \mathcal{F}''_{n_0} = \Sigma(X''_m : m \leq n_0) \text{ are independent} \quad (7)$$

(to ensure that the filtrations \mathcal{F}' and \mathcal{F}'' generated by the negative parts of these process are independent in the distant past);

- for each $n_0 + 1 \leq n \leq 0$,

$$\mathcal{L}(X'_n | \mathcal{F}'_{n-1} \vee \mathcal{F}''_{n-1}) = \tau_\varepsilon(X'_{n-1}, \cdot) \text{ and } \mathcal{L}(X''_n | \mathcal{F}'_{n-1} \vee \mathcal{F}''_{n-1}) = \tau_\varepsilon(X''_{n-1}, \cdot) \quad (8)$$

(to get the joint immersion of the filtrations \mathcal{F}' and \mathcal{F}''); and

- the copies $(X'_0[-k, k])$ and $(X''_0[-k, k])$ of the target random variable satisfy

$$\mathbb{P}(X'_0[-k, k] = X''_0[-k, k]) > 1 - \delta. \quad (9)$$

Here is how we will proceed. We fix some $n_0 \leq 0$ and consider a probability space (Ω, \mathbb{P}) in which we have two independent copies $(X'_n : n \leq n_0)$ and $(X''_n : n \leq n_0)$ of $(X_n : n \leq n_0)$, and an independent family of i.i.d. random variables $(U_n(i) : n \geq n_0 + 1, i \in \mathbb{Z})$ which are all uniformly distributed on the interval $[0, 1]$. These random variables will be used to construct inductively the error processes $(\xi'_n(i) : n \geq n_0 + 1, i \in \mathbb{Z})$ and $(\xi''_n(i) : n \geq n_0 + 1, i \in \mathbb{Z})$ and the rest of the Markov processes $(X'_n : n \geq n_0 + 1)$ and

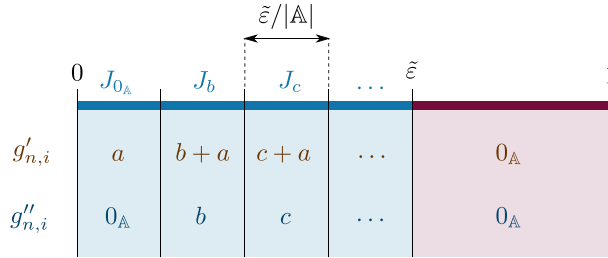


FIGURE 2. Choice of $g'_{n,i}$ and $g''_{n,i}$ when we apply the strategy S_a .

($X''_n : n \geq n_0 + 1$) through the formula

$$X'_n := \tau X'_{n-1} + \xi'_n \quad \text{and} \quad X''_n := \tau X''_{n-1} + \xi''_n.$$

We will use the auxiliary process $Z := X' - X''$, and we note that for $n \geq n_0 + 1$ we have

$$Z_n = \tau Z_{n-1} + \xi'_n - \xi''_n. \tag{10}$$

Rewriting (9) with this notation, we want to achieve the coupling in such a way that, provided $|n_0|$ is large enough,

$$\mathbb{P}(Z_0[-k, k] = (0_{\mathbb{A}}, \dots, 0_{\mathbb{A}})) > 1 - \delta. \tag{11}$$

We explain now how we construct the error processes inductively. Assuming that, for some $n \geq n_0 + 1$, we have already defined the processes X' and X'' up to time $n - 1$, we choose for each $i \in \mathbb{Z}$ two random maps $g'_{n,i}$ and $g''_{n,i}$ from $[0, 1]$ to \mathbb{A} and set $\xi'_n(i) := g'_{n,i}(U_n(i))$ and $\xi''_n(i) := g''_{n,i}(U_n(i))$. These maps are random in the sense that they depend on the realizations of the processes up to time $n - 1$. However, they have to comply with the required law for the errors, which we ensure by choosing them among the family of maps $(g_a)_{a \in \mathbb{A}}$ defined as follows. Recalling the definition (3) of $\tilde{\varepsilon}$ and the formulation (4) of the law of the error, we cut the interval $[0, 1]$ into $|\mathbb{A}| + 1$ subintervals: the first $|\mathbb{A}|$ of them are of the same length $\tilde{\varepsilon}/|\mathbb{A}|$, and we denote them by J_b , $b \in \mathbb{A}$ (they correspond to the uniform measure part of the error); the last one is $[\tilde{\varepsilon}, 1]$. Then, for each $a \in \mathbb{A}$, we define the map $g_a : [0, 1] \rightarrow \mathbb{A}$ by setting

$$g_a(u) := \begin{cases} a + b & \text{if } u \in J_b \ (b \in \mathbb{A}), \\ 0_{\mathbb{A}} & \text{if } u \in [\tilde{\varepsilon}, 1]. \end{cases}$$

By choosing $g'_{n,i}$ and $g''_{n,i}$ in the set $\{g_a : a \in \mathbb{A}\}$, we get the correct conditional law of the errors ξ'_n and ξ''_n knowing $\mathcal{F}'_{n-1} \vee \mathcal{F}''_{n-1}$, and we have (8). It remains to explain how we make this choice for our purposes.

Given a site (n, i) , for each $a \in \mathbb{A}$ we define the strategy S_a as the choice $(g'_{n,i}, g''_{n,i}) := (g_a, g_{0_{\mathbb{A}}})$ (see Figure 2). In this way, when we apply the strategy S_a , we obtain

$$\xi'_n(i) - \xi''_n(i) = \begin{cases} a & \text{if } U_n(i) \in [0, \tilde{\varepsilon}), \\ 0_{\mathbb{A}} & \text{otherwise.} \end{cases}$$

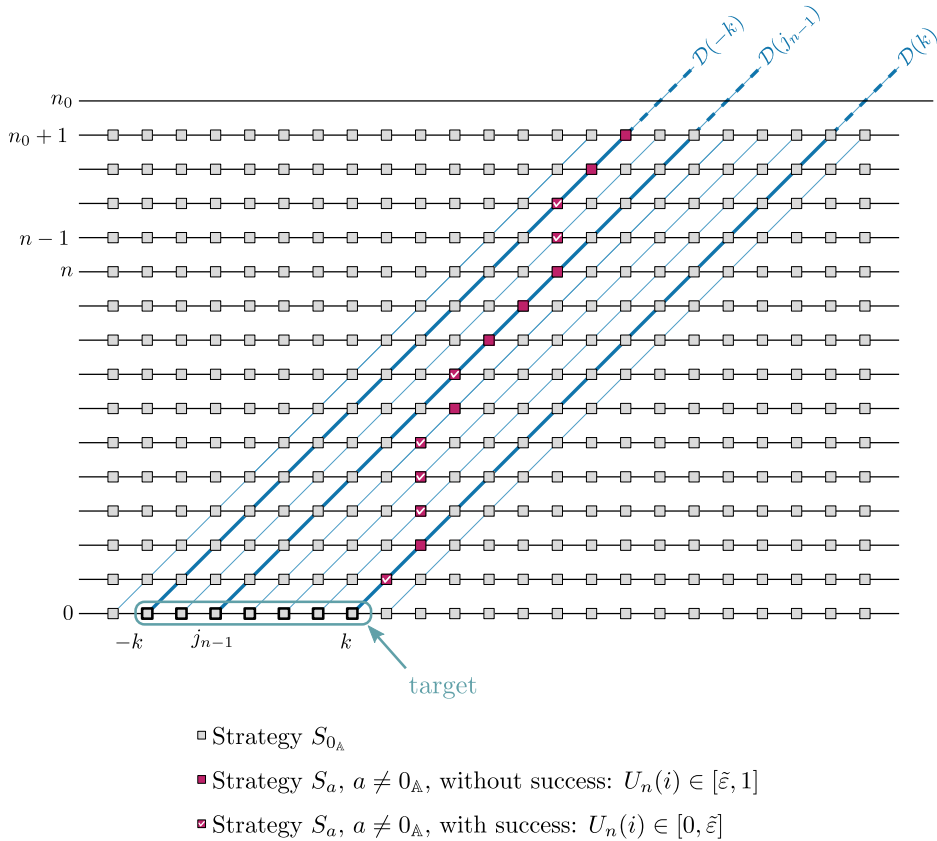


FIGURE 3. The choice of the local strategies to construct the real-time coupling.

The special case $a = 0_{\mathbb{A}}$ gives $\xi''_n(i) - \xi'_n(i) = 0_{\mathbb{A}}$ with probability 1, and thus the choice of the strategy $S_{0_{\mathbb{A}}}$ on a given site (n, i) ensures that, locally, the evolution of the Z process is given by the deterministic action of the cellular automaton τ (remember (10)).

The set of sites (n, i) for which we have to define $g'_{n,i}$ and $g''_{n,i}$ is partitioned into ‘diagonals’ $\mathcal{D}(j)$, $j \in \mathbb{Z}$, where

$$\mathcal{D}(j) := \{(n, i) / n \geq n_0 + 1, i = j - n\}$$

(see Figure 3). We observe that, for $i < j \in \mathbb{Z}$, any error added on some site of $\mathcal{D}(j)$ has no influence on $Z_0(i)$.

On all the diagonals $\mathcal{D}(j)$, $j < -k$ or $j > k$, we systematically choose the strategy $S_{0_{\mathbb{A}}}$, so that Z follows the deterministic evolution of the cellular automaton τ in these areas. It remains to explain the choice of the strategies on the sites of the diagonals $\mathcal{D}(j)$, $-k \leq j \leq k$.

Consider an integer n , $n_0 < n \leq 0$, and assume that we know the Z process up to time $n - 1$. We compute $\tau^{|n-1|}(Z_{n-1})$, and according to what we get we make the following choices.

- If $\tau^{|n-1|}(Z_{n-1})[-k, k] = (0_{\mathbb{A}}, \dots, 0_{\mathbb{A}})$, we can easily ensure that $Z_0[-k, k] = (0_{\mathbb{A}}, \dots, 0_{\mathbb{A}})$ by choosing the strategy $S_{0_{\mathbb{A}}}$ for all the following sites. In this case we set $j_{n-1} := k + 1$.
- Otherwise, we define j_{n-1} as the smallest integer $j \in \{-k, \dots, k\}$ such that

$$\tau^{|n-1|}(Z_{n-1})(j) \neq 0_{\mathbb{A}},$$

and we apply the strategy $S_{0_{\mathbb{A}}}$ at all sites (n, i) except when (n, i) falls on $\mathcal{D}_{j_{n-1}}$ (that is, when $i = j_{n-1} - n$), where we apply $S_{-a_{n-1}}$ with $a_{n-1} := \tau^{|n-1|}(Z_{n-1})(j_{n-1})$.

With this method, if $j_{n-1} \in \{-k, \dots, k\}$, the only i on line n for which $Z_{n,i}$ is not determined is $i = j_{n-1} - n$, and the value of $Z_{n,j_{n-1}-n}$ now depends only on the random variable $U_n(j_{n-1} - n)$:

- If $U_n(j_{n-1} - n) \in [0, \tilde{\varepsilon})$, the difference $\xi'_n(j_{n-1} - n) - \xi''_n(j_{n-1} - n)$ takes the value $-a_{n-1}$. This exactly compensates the deterministic action of τ for the target site $(0, j_{n-1})$, which ensures that, at the next step, we get $\tau^{|n|}(Z_n)(j_{n-1}) = 0_{\mathbb{A}}$, and thus $j_n \geq j_{n-1} + 1$.
- Otherwise, we have $\xi'_n(j_{n-1} - n) - \xi''_n(j_{n-1} - n) = 0$, so that $\tau^{|n|}(Z_n)(j_{n-1}) = a_{n-1}$ and $j_n = j_{n-1}$.

We construct the real-time coupling inductively by applying the above method for $n_0 + 1 \leq n \leq 0$. How we build the coupling for times $n \geq 0$ is irrelevant; we can, for example, choose always the strategy $S_{0_{\mathbb{A}}}$ for these positive times. Now we have to prove that (11) is achieved for $|n_0|$ large enough.

We inductively define random times $T(-k), \dots, T(k)$. We start with $T(-k)$, which only depends on the random variables $U_n(i), (n, i) \in \mathcal{D}(-k)$:

$$T(-k) := \min\{n \geq n_0 + 1 : U_n(-k - n) \in [0, \tilde{\varepsilon})\}.$$

(With the usual convention that $\min \emptyset = +\infty$, but observe that $T(-k) < \infty$ with probability 1.) The random variable $T(-k) - n_0$ then follows a geometric law of parameter $\tilde{\varepsilon}$. And by construction of the coupling, for each $n_0 + 1 \leq n \leq 0$ we have

$$T(-k) \leq n \implies j_n > -k.$$

Then, if we have already defined $T(j)$ for some $-k \leq j \leq k - 1$, we set

$$T(j + 1) := \min\{n \geq T(j) + 1 : U_n(j + 1 - n) \in [0, \tilde{\varepsilon})\}.$$

Note that $T(j + 1)$ only depends on $T(j)$ and on the random variables $U_n(i), (n, i) \in \mathcal{D}(j + 1)$. And by construction of the coupling, we get by induction that, for each $j \in \{-k, \dots, k\}$ and each $n \leq 0$,

$$T(j) \leq n \implies j_n > j.$$

In particular,

$$T(k) \leq 0 \implies j_0 > k \implies Z_0[-k, k] = (0_{\mathbb{A}}, \dots, 0_{\mathbb{A}}). \tag{12}$$

For $j \in \{-k, \dots, k - 1\}$, the difference $T(j + 1) - T(j)$ is distributed according to a geometric law of parameter $\tilde{\varepsilon}$, and is independent of $T(-k), \dots, T(j)$. Hence $T(k) - n_0$

is distributed as the sum of $2k + 1$ independent geometric random variables of parameter $\tilde{\varepsilon}$. Therefore, as $\tilde{\varepsilon}$ and k are fixed, we get

$$\mathbb{P}(T(k) \leq 0) \xrightarrow[n_0 \rightarrow -\infty]{} 1. \tag{13}$$

By (13) and (12), we get that (11) is satisfied for $|n_0|$ large enough.

This proves that for each $k \geq 0$, the target $X_0[-k, k]$ is I-cosy. Now, the same argument works also to prove the I-cosiness of $X_{-k}[-2k, 2k]$ for each $k \geq 0$. This gives us a real-time coupling up to time $-k$ for which

$$\mathbb{P}(Z_{-k}[-2k, 2k] = (0_{\mathbb{A}}, \dots, 0_{\mathbb{A}})) > 1 - \delta.$$

Note that

$$Z_{-k}[-2k, 2k] = (0_{\mathbb{A}}, \dots, 0_{\mathbb{A}}) \implies \text{for all } 0 \leq \ell \leq k, \tau^\ell Z_k[-k, k] = (0_{\mathbb{A}}, \dots, 0_{\mathbb{A}}).$$

So we can continue this coupling by always using the strategy $S_{0_{\mathbb{A}}}$ on lines $-k + 1, \dots, 0$ to prove the I-cosiness of $Y_k = (X_n[-k, k] : -k \leq n \leq 0)$. □

We further observe that the filtration $\mathcal{F}^{\tau_\varepsilon}$ is *homogeneous*: for each $n \leq 0$, $\mathcal{F}_n^{\tau_\varepsilon} = \mathcal{F}_{n-1}^{\tau_\varepsilon} \vee \Sigma(\xi_n)$, where

- ξ_n is independent of $\mathcal{F}_{n-1}^{\tau_\varepsilon}$,
- the law of ξ_n is diffuse.

Then a direct application of Theorem A in [11] yields the following corollary.

COROLLARY 2.14. *For each $0 < \varepsilon < (|\mathbb{A}| - 1)/|\mathbb{A}|$, the filtration $\mathcal{F}^{\tau_\varepsilon}$ is of product type.*

2.5. Uniform cosiness and ergodicity of the Markov kernel. If we look carefully at the proof that the filtration $\mathcal{F}^{\tau_\varepsilon}$ is I-cosy, we see that the probability that the two copies of the target coincide in the real-time coupling converges to 1 as $|n_0| \rightarrow \infty$ *uniformly with respect to the states of the two copies at time n_0* . We in fact get a stronger property than I-cosiness, which we call *uniform cosiness*, and which implies not only the I-cosiness of the filtration, but also the ergodicity of the Markov kernel.

Recall that a Markov kernel on a compact metric space (\mathcal{X}, d) is *Feller* if $x \mapsto P(x, \cdot)$ is continuous for the weak* topology. Any Feller Markov kernel on a compact metric space admits an invariant probability distribution on X . The Markov kernel is said to be *ergodic in the Markov sense* if there exists a probability distribution μ on X such that, for each probability measure ν on \mathcal{X} ,

$$\nu Q^n \xrightarrow[n \rightarrow \infty]{w^*} \mu.$$

In this case, μ is the unique Q -invariant probability measure.

Definition 2.15. (Uniform cosiness) Let Q be a Markov kernel on a compact metric space (\mathcal{X}, d) . We say that Q is *uniformly cosy* if, for each $\delta > 0$, there exists $M = M(\delta) > 0$ such that, whenever n_0 is a negative integer with $|n_0| \geq M$, for each $x', x'' \in \mathcal{X}$, there exists a probability measure $m_{x', x''}$ on $(\mathcal{X}^{n_0, \dots, 0})^2$, depending measurably on (x', x'')

such that, denoting by $(X'_n)_{n_0 \leq n \leq 0}$ and $(X''_n)_{n_0 \leq n \leq 0}$ the canonical processes defined by the coordinates on $(\mathcal{X}^{\{n_0, \dots, 0\}})^2$, the following conditions hold.

- The starting points of the processes are given by $X'_{n_0} = x'$ and $X''_{n_0} = x''$, $m_{x',x''}$ -a.s.
- Under $m_{x',x''}$, for each $n_0 \leq n \leq -1$, the conditional distribution

$$\mathcal{L}((X'_{n+1}, X''_{n+1}) \mid \Sigma(X'_{n_0}, \dots, X'_n, X''_{n_0}, \dots, X''_n))$$

is a.s. a coupling of the two probability distributions $Q(X'_n, \cdot)$ and $Q(X''_n, \cdot)$ (we thus realize from n_0 a real-time coupling of two Markov processes of transitions probabilities given by Q , one starting at x' and the other at x'').

- We have

$$m_{x',x''}(d(X'_0, X''_0) > \delta) < \delta.$$

For example, the proof of Theorem 2.13 shows that the Markov kernel defined by the PCA τ_ε is uniformly cosy.

THEOREM 2.16. *Let Q be a Feller Markov kernel on the compact metric space (\mathcal{X}, d) . If Q is uniformly cosy, then Q is ergodic in the Markov sense.*

Proof. Let μ be a Q -invariant probability measure on \mathcal{X} , and let ν be any probability measure on \mathcal{X} . We want to prove that, for each continuous function $f : \mathcal{X} \rightarrow \mathbb{R}$ and each $\varepsilon > 0$, for each n large enough we have

$$\left| \int_{\mathcal{X}} f d(\nu Q^n) - \int_{\mathcal{X}} f d\mu \right| \leq \varepsilon.$$

Given f and ε , by uniform continuity of f on the compact \mathcal{X} , we get $\delta > 0$ such that,

$$\text{for all } x, y \in \mathcal{X}, \quad d(x, y) \leq \delta \implies |f(x) - f(y)| < \varepsilon/2.$$

We can also assume that $\delta \|f\|_\infty < \varepsilon/4$.

By uniform cosiness, we can take $M > 0$ such that, given any integer $n_0 \leq -M$, there exists a family of probability measures $(m_{x',x''})_{x',x'' \in \mathcal{X}}$ associated to this δ as in the statement of the definition of uniform cosiness. We then define the probability measure $m_{\mu,\nu}$ on $(\mathcal{X}^{\{n_0, \dots, 0\}})^2$ by

$$m_{\mu,\nu} := \int_{\mathcal{X}} \int_{\mathcal{X}} m_{x',x''} d\mu(x') d\nu(x'').$$

Under $m_{\mu,\nu}$, the processes $(X'_n)_{n_0 \leq n \leq 0}$ and $(X''_n)_{n_0 \leq n \leq 0}$ are two Markov chains with transitions given by Q , with respective initial laws $\mathcal{L}(X'_{n_0}) = \mu$ and $\mathcal{L}(X''_{n_0}) = \nu$. We then have $\mathcal{L}(X'_0) = \mu Q^{|n_0|} \mu = \mu$ by invariance of μ , and $\mathcal{L}(X''_0) = \nu Q^{|n_0|}$. The uniform cosiness yields

$$m_{\mu,\nu}(d(X'_0, X''_0) > \delta) = \int_{\mathcal{X}} \int_{\mathcal{X}} \underbrace{m_{x',x''}(d(X'_0, X''_0) > \delta)}_{< \delta} d\mu(x') d\nu(x'') < \delta.$$

We then have

$$\begin{aligned} \left| \int_{\mathcal{X}} f d(\nu \mathcal{Q}^{|\mathfrak{n}_0|}) - \int_{\mathcal{X}} f d\mu \right| &\leq \mathbb{E}_{m_{\mu,\nu}}[|f(X''_0) - f(X'_0)|] \\ &\leq \int_{d(X'_0, X''_0) > \delta} |f(X''_0) - f(X'_0)| dm_{\mu,\nu} + \varepsilon/2 \\ &\leq 2\delta \|f\|_{\infty} + \varepsilon/2 \leq \varepsilon. \end{aligned} \quad \square$$

COROLLARY 2.17. *The Markov kernel given by the cellular automaton τ_ε is ergodic in the Markov sense, and $\mu = \mathcal{U}_{\mathbb{A}}^{\otimes \mathbb{Z}}$ is the unique τ_ε -invariant probability measure.*

3. Dynamical classification of factor filtrations

3.1. Factor filtrations. In the two cases we have studied above, we observe that the filtrations $\mathcal{F}^{\tau^{-1}}$ and $\mathcal{F}^{\tau_\varepsilon}$ both enjoy an important property: if we consider their canonical construction, which is on $\mathbb{A}^{\mathbb{Z}}$ for $\mathcal{F}^{\tau^{-1}}$ and on $(\mathbb{A}^{\mathbb{Z}})^{\mathbb{Z}}$ for τ_ε , we have on the ambient probability space the action of the left shift σ , which is a measure-preserving transformation, and moreover all the sigma-algebras in those filtrations are invariant with respect to this automorphism. Our purpose in this section is to formalize such a situation, and adapt the study of the filtration to this context.

Here a *dynamical system* is any system of the form (Ω, \mathbb{P}, T) , where (Ω, \mathbb{P}) is a probability space, and $T : \Omega \rightarrow \Omega$ is an invertible, bi-measurable transformation which preserves the probability measure \mathbb{P} . Given such a system, a *factor sigma-algebra* of (Ω, \mathbb{P}, T) is any sub-sigma algebra \mathcal{A} of the Borel sigma algebra of Ω which is invariant by T : for each $A \in \mathcal{A}$, $T^{-1}A$ and TA are also in \mathcal{A} . For any random variable X_0 defined on (Ω, \mathbb{P}) , as T preserves \mathbb{P} , the process $X = (X(i))_{i \in \mathbb{Z}}$ defined by

$$X(i) := X_0 \circ T^i, \quad \text{for all } i \in \mathbb{Z}, \tag{14}$$

is stationary. Such a stationary process will be called a *T-process*. Whenever X is a T -process, the sigma-algebra generated by X is clearly a factor sigma-algebra of (Ω, \mathbb{P}, T) . Conversely, any factor sigma-algebra of (Ω, \mathbb{P}, T) is generated by some T -process (remember that all sigma-algebras are assumed to be essentially separable).

Definition 3.1. A *factor filtration* is any pair (\mathcal{F}, T) where

- $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ is a filtration on some probability space (Ω, \mathbb{P}) ,
- (Ω, \mathbb{P}, T) is a dynamical system,
- for each $n \leq 0$, \mathcal{F}_n is a factor sigma-algebra of (Ω, \mathbb{P}, T) .

In view of the above discussion, in a factor filtration (\mathcal{F}, T) the filtration \mathcal{F} is always generated by a process $(X_n)_{n \leq 0}$, where for each n , $X_n = (X_n(i))_{i \in \mathbb{Z}}$ is a T -process.

The dynamical classification of factor filtrations that we want to introduce now aims at distinguishing these objects up to the following notion of isomorphism, which is the adaptation of Definition 2.1 to the dynamical context.

Definition 3.2. (Dynamical isomorphism of factor filtrations) Let (\mathcal{F}, T) and (\mathcal{F}', T') be two factor filtrations, possibly defined on two different dynamical systems. We say that they

are *dynamically isomorphic* if we can find two processes $X = (X_n)_{n \leq 0}$ and $X' = (X'_n)_{n \leq 0}$ generating respectively \mathcal{F} and \mathcal{F}' , and such that:

- for each n , $X_n = (X_n(i))_{i \in \mathbb{Z}}$ is a T -process, and $X'_n = (X'_n(i))_{i \in \mathbb{Z}}$ is a T' -process;
- $\mathcal{L}(X) = \mathcal{L}(X')$.

We write in this case $(\mathcal{F}, T) \sim (\mathcal{F}', T')$, or $(\mathcal{F}, T) \overset{X, X'}{\sim} (\mathcal{F}', T')$ to specify the processes involved in the isomorphism.

We note the following specificity of dynamical isomorphism: if $(\mathcal{F}, T) \overset{X, X'}{\sim} (\mathcal{F}', T')$, and if Y' is the copy of an \mathcal{F}_0 random variable Y provided by this isomorphism, then $Y' \circ T'$ is the corresponding copy of $Y \circ T$.

The notion of immersion is unchanged: (\mathcal{F}, T) and (\mathcal{G}, T) being two factor filtrations in the same dynamical system, we simply say that (\mathcal{F}, T) is immersed in (\mathcal{G}, T) if \mathcal{F} is immersed in \mathcal{G} . However, the notion of immersibility takes into account the above definition of dynamical isomorphism.

Definition 3.3. (Dynamical immersibility of factor filtrations) Let (\mathcal{F}, T) and (\mathcal{G}, S) be two factor filtrations, possibly defined on two different dynamical systems. We say that (\mathcal{F}, T) is dynamically immersible in (\mathcal{G}, S) if there exists a factor filtration (\mathcal{F}', S) immersed in (\mathcal{G}, S) which is dynamically isomorphic to (\mathcal{F}, T) .

We can now adapt the notions of product type and standardness to factor filtrations.

Definition 3.4. The factor filtration (\mathcal{F}, T) is said to be *dynamically of product type* if there exists a family $(\mathcal{G}_n)_{n \leq 0}$ of independent factor sigma-algebras such that for each $n \leq 0$, $\mathcal{F}_n = \bigvee_{m \leq n} \mathcal{G}_m$.

Equivalently, (\mathcal{F}, T) is of product type if \mathcal{F} can be generated by a process $X = (X_n)_{n \leq 0}$ whose coordinates X_n are independent T -processes. Of course, if the factor filtration (\mathcal{F}, T) is of product type, then the filtration \mathcal{F} is itself of product type, but the converse is not true (see the example in §3.2).

Definition 3.5. The factor filtration (\mathcal{F}, T) is said to be *dynamically standard* if it is dynamically immersible in some factor filtration dynamically of product type.

It is natural also to translate the notions of real-time coupling and I-cosiness to the dynamical context. The corresponding notions are formally the same as in the static case, except that we have to replace the isomorphism of filtrations by the dynamical isomorphism of factor filtrations.

Definition 3.6. (Dynamical real-time coupling) Let (\mathcal{F}, T) be a factor filtration on a dynamical system (Ω, \mathbb{P}, T) . A *dynamical real-time coupling* of (\mathcal{F}, T) is a pair $((\mathcal{F}', S), (\mathcal{F}'', S))$ of factor filtrations, both defined on the same dynamical system (but possibly different from (Ω, \mathbb{P}, T)), such that:

- $(\mathcal{F}', S) \sim (\mathcal{F}, T)$;
- $(\mathcal{F}'', S) \sim (\mathcal{F}, T)$;
- \mathcal{F}' and \mathcal{F}'' are jointly immersed.

Definition 3.7. (Dynamical I-cosiness) Let (\mathcal{F}, T) be a factor filtration, and let Y be an \mathcal{F}_0 -measurable random variable taking values in some Polish metric space (E, d) . Then Y is said to be *dynamically I-cosy with respect to (\mathcal{F}, T)* if, for each real number $\delta > 0$, we can find a dynamical real-time coupling $((\mathcal{F}', S), (\mathcal{F}'', S))$ of (\mathcal{F}, T) which is independent in the distant past, and such that the two copies Y' and Y'' of Y in \mathcal{F}' and \mathcal{F}'' respectively satisfy

$$d(Y', Y'') < \delta \text{ with probability } > 1 - \delta. \tag{15}$$

The factor filtration (\mathcal{F}, T) is said to be *dynamically I-cosy* if each \mathcal{F}_0 -measurable random variable Y is dynamically I-cosy with respect to (\mathcal{F}, T) .

We expect of course some relationship between dynamical I-cosiness and dynamical standardness. In the static case, showing that standardness implies I-cosiness is not very difficult, and in fact exactly the same arguments apply in the dynamical case. We provide them below.

LEMMA 3.8. *Let (\mathcal{F}, T) be a factor filtration which is dynamically of product type. Then (\mathcal{F}, T) is dynamically I-cosy.*

Proof. Let (Ω, \mathbb{P}, T) be the dynamical system where the factor filtration is defined, and let $X = (X_n)_{n \leq 0}$ be a process generating \mathcal{F} , where the coordinates X_n are independent T -processes: $X_n = (X_n(i))_{i \in \mathbb{Z}} = (X_n(0) \circ T^i)_{i \in \mathbb{Z}}$.

We have to check the dynamical I-cosiness with respect to (\mathcal{F}, T) of a target random variable Y which is \mathcal{F}_0 -measurable. It is enough to consider the case where Y is measurable with respect to $\Sigma(X_{n_0+1}, \dots, X_0)$ for some $n_0 \leq -1$.

On the product dynamical system $(\Omega \times \Omega, \mathbb{P} \otimes \mathbb{P}, T \times T)$, we have two independent copies $X^{(1)}$ and $X^{(2)}$ of X , whose coordinates are independent $(T \times T)$ processes. Let us now consider the two copies X' and X'' of X defined by $X' := X^{(1)}$, and

$$X'' := \begin{cases} X_n^{(2)} & \text{if } n \leq n_0, \\ X_n^{(1)} & \text{if } n \geq n_0 + 1. \end{cases}$$

Define \mathcal{F}' (respectively, \mathcal{F}'') as the filtration generated by X' (respectively, X''). Then $((\mathcal{F}', T \times T), (\mathcal{F}'', T \times T))$ is an n_0 -independent dynamical coupling of (\mathcal{F}, T) . Moreover, since by construction we have $X'_n = X''_n$ for $n \geq n_0 + 1$, the corresponding copies of Y' and Y'' of Y satisfy $Y' = Y''$. This concludes the proof. \square

LEMMA 3.9. *Let (\mathcal{F}, T) and (\mathcal{G}, S) be two factor filtrations. Assume that (\mathcal{G}, S) is dynamically I-cosy, and that (\mathcal{F}, T) is dynamically immersible in (\mathcal{G}, S) . Then (\mathcal{F}, T) is dynamically I-cosy.*

Proof. Since cosiness is preserved by isomorphism, we can assume without loss of generality that (\mathcal{F}, T) is *immersed* in some dynamically I-cosy factor filtration (\mathcal{G}, T) . Let $X = (X_n)_{n \leq 0}$ be a process generating \mathcal{F} , where the coordinates X_n are T -processes. Let Y be an \mathcal{F}_0 -measurable random variable, taking values in the Polish metric space (E, d) , and let $\delta > 0$. Since $\mathcal{F}_0 \subset \mathcal{G}_0$, and since (\mathcal{G}, T) is dynamically I-cosy, there exists

a dynamical real-time coupling $((\mathcal{G}, S), (\mathcal{G}', S))$ of (\mathcal{G}, T) , independent in the distant past, such that the corresponding copies Y' and Y'' of Y satisfy $d(Y', Y'') < \delta$ with probability at least $1 - \delta$. But the generating process X is \mathcal{G}_0 -measurable, so we can consider its copies X' and X'' in \mathcal{G} and \mathcal{G}' , respectively. Let \mathcal{F}' (respectively, \mathcal{F}'') be the filtration generated by X' (respectively, X''). Since these copies are provided by a dynamical isomorphism, the coordinates of X' and X'' are S -processes, hence we get two dynamically isomorphic copies (\mathcal{F}', S) and (\mathcal{F}'', S) of the factor filtration (\mathcal{F}, T) . Since \mathcal{G} and \mathcal{G}' are jointly immersed, and \mathcal{F}' (respectively, \mathcal{F}'') is immersed in \mathcal{G} (respectively, \mathcal{G}'), \mathcal{F}' and \mathcal{F}'' are jointly immersed. Hence we get a dynamical real-time coupling $((\mathcal{F}', S), (\mathcal{F}'', S))$ of the factor filtration (\mathcal{F}, T) . Moreover this coupling is independent in the distant past because this is the case for $((\mathcal{G}, S), (\mathcal{G}', S))$. But Y' and Y'' are respectively \mathcal{F}'_0 - and \mathcal{F}''_0 -measurable, as Y is \mathcal{F}_0 -measurable, so they are also the corresponding copies of Y in the dynamical real-time coupling $((\mathcal{F}', S), (\mathcal{F}'', S))$. □

From Lemmas 3.8 and 3.9, we immediately derive the following theorem.

THEOREM 3.10. *If the factor filtration (\mathcal{F}, T) is dynamically standard, then it is dynamically I-cosy.*

Whether the converse of the above theorem is true remains an open question (see a partial result in this direction in §3.4). However, the fact that dynamical I-cosiness is necessary for dynamical standardness can already be used to establish that some factor filtrations are not dynamically standard, as in the following section.

3.2. Dynamical versus static classification of filtrations: the example of $\mathcal{F}^{\tau^{-1}}$. We return again to the filtration $\mathcal{F}^{\tau^{-1}}$ associated to the deterministic cellular automaton τ , and defined in §1.1.1. The probability measure $\mu = \mathcal{U}_{\mathbb{A}}^{\otimes \mathbb{Z}}$ on $\mathbb{A}^{\mathbb{Z}}$ is invariant by the shift map σ . Furthermore, since the transformations τ and σ commute, each coordinate $X_n = (X_n(i))_{i \in \mathbb{Z}}$ of the process generating the filtration $\mathcal{F}^{\tau^{-1}}$ is itself a stationary σ -process: for each $i \in \mathbb{Z}$, $X_n(i) = X_n(0) \circ \sigma^i$. Thus, $(\mathcal{F}^{\tau^{-1}}, \sigma)$ is a factor filtration in the dynamical system $(\mathbb{A}^{\mathbb{Z}}, \mu, \sigma)$.

We recall that, if we look at the filtration $\mathcal{F}^{\tau^{-1}}$ from the static point of view, it is of product type (Theorem 2.5). However the following result shows that dynamical classification of factor filtrations may lead to different results than the static one.

THEOREM 3.11. *The factor filtration $(\mathcal{F}^{\tau^{-1}}, \sigma)$ is not dynamically standard.*

Proof. The strategy is to show that the factor filtration $(\mathcal{F}^{\tau^{-1}}, \sigma)$, and more precisely the target $X_0(0)$, is not dynamically I-cosy. For this, we consider a real-time dynamical coupling of this factor filtration in a dynamical system (Ω, \mathbb{P}, T) , so that we have two copies $X' = (X'_n)_{n \leq 0}$ and $X'' = (X''_n)_{n \leq 0}$ of the process (X_n) , where for each $n \leq 0$,

- $X'_n = (X'_n(i))_{i \in \mathbb{Z}} = \tau^{|n|} X'_0$,
- $X''_n = (X''_n(i))_{i \in \mathbb{Z}} = \tau^{|n|} X''_0$,

and for each $i \in \mathbb{Z}$,

- $X'_n(i) = X'_n(0) \circ T^i$,
- $X''_n(i) = X''_n(0) \circ T^i$.

For each $n \leq 0$, set $Z_n := X'_n - X''_n$. Then, for each n , $Z_n = (Z_n(i))_{i \in \mathbb{Z}} = (Z_n(0) \circ T^i)_{i \in \mathbb{Z}}$ is a stationary T -process taking values in \mathbb{A} . And since τ is an endomorphism of the group $\mathbb{A}^{\mathbb{Z}}$, we also have $Z_n = \tau^{|n|} Z_0$ for each $n \leq 0$.

Assume that this coupling is n_0 -independent for some $n_0 \leq 0$; in other words, X'_{n_0} and X''_{n_0} are independent. Then the probability distribution of Z_{n_0} is nothing more than $\mu = \mathcal{U}_{\mathbb{A}}^{\otimes \mathbb{Z}}$. Therefore, the Kolmogorov–Sinai entropy of this T -process is

$$h(Z_{n_0}, T) = \log |\mathbb{A}|.$$

But we have

$$Z_{n_0} = \tau^{|n_0|} Z_0,$$

so that the T -process Z_{n_0} is a factor of the T -process Z_0 . In particular, since the Kolmogorov–Sinai entropy cannot increase when we pass to a factor, and observing that Z_0 is a T process taking values in \mathbb{A}

$$\log |\mathbb{A}| = h(Z_{n_0}, T) \leq h(Z_0, T) \leq \log |\mathbb{A}|.$$

However the only T -process taking values in \mathbb{A} and whose entropy is $\log |\mathbb{A}|$ is the uniform Bernoulli process. It follows that the probability distribution of Z_0 is necessarily also μ . Therefore,

$$\mathbb{P}(X'_0(0) = X''_0(0)) = \mathbb{P}(Z_0(0) = 0) = \frac{1}{|\mathbb{A}|},$$

which prevents the dynamical I-cosiness of the target $X_0(0)$. □

3.2.1. An alternative proof for the case of an infinite group. In this section only we allow the non-trivial Abelian group $(\mathbb{A}, +)$ to be metrizable compact, not necessarily finite. The action of the deterministic cellular automaton τ on $\mathbb{A}^{\mathbb{Z}}$ can be defined by the same formula as in the finite case (1). We denote by $\mathcal{U}_{\mathbb{A}}$ the Haar measure on \mathbb{A} (which is the uniform measure in the case of a finite group). Still, $\mu = \mathcal{U}_{\mathbb{A}}^{\otimes \mathbb{Z}}$ is invariant by τ and by the shift map σ , and we can consider the filtration $\mathcal{F}^{\tau^{-1}}$ defined in §1.1.1, as well as the factor filtration $(\mathcal{F}^{\tau^{-1}}, \sigma)$ in this more general context. As in the finite case, we can show by exactly the same arguments that $\mathcal{F}^{\tau^{-1}}$ is of product type even in the non-finite case. But the above proof of Theorem 3.11 makes essential use of the fact that \mathbb{A} is a *finite* Abelian group. We propose below an alternative proof, which replaces the entropy argument by a coupling argument, and which is still valid in the infinite case.

We first need to recall some basic facts about joinings of stationary processes in ergodic theory.

Definition 3.12. (Joining) Let I be a finite or countable set of indices. For $i \in I$, let X_i be a T_i -process in a dynamical system $(\Omega_i, \mathbb{P}_i, T_i)$. A *dynamical coupling* or *joining* of the processes $(X_i)_{i \in I}$ is the simultaneous realization in the same dynamical system $(\bar{\Omega}, \bar{\mathbb{P}}, \bar{T})$ of \bar{T} -processes $\bar{X}_i, i \in I$, such that $\mathcal{L}(\bar{X}_i) = \mathcal{L}(X_i)$ for each $i \in I$.

Definition 3.13. (Disjointness) The stationary processes X_1 and X_2 are said to be *disjoint* if, for each joining (\bar{X}_1, \bar{X}_2) of X_1 and X_2 , \bar{X}_1 and \bar{X}_2 are independent.

Disjointness was introduced by Furstenberg in 1967 in his famous paper [7], where he also provided one of the first examples of disjointness.

THEOREM 3.14. *Let X_1 be a Bernoulli process (that is, a stationary process with independent coordinates), and let X_2 be a stationary process with zero entropy. Then X_1 and X_2 are disjoint.*

The following result is a direct application of the construction of a relatively independent joining over a common factor (see, for example, [4]).

LEMMA 3.15. *Let X_1 and Y_1 be two T_1 -processes in a dynamical system $(\Omega_1, \mathbb{P}_1, T_1)$, and let X_2 and Y_2 be two T_2 -processes in a second dynamical system $(\Omega_2, \mathbb{P}_2, T_2)$. Assume that $\mathcal{L}(Y_1) = \mathcal{L}(Y_2)$. Then there exists a joining $(\bar{X}_1, \bar{Y}_1, \bar{X}_2, \bar{Y}_2)$ of X_1, Y_1, X_2, Y_2 , such that:*

- $\mathcal{L}(\bar{X}_1, \bar{Y}_1) = \mathcal{L}(X_1, Y_1)$;
- $\mathcal{L}(\bar{X}_2, \bar{Y}_2) = \mathcal{L}(X_2, Y_2)$;
- $\bar{Y}_1 = \bar{Y}_2$ a.s.

The invariance of μ by τ tells us that, if a random variable X takes its values in $\mathbb{A}^{\mathbb{Z}}$ and follows the distribution μ , then the law of τX is also μ . The following lemma, which is the key ingredient in the alternative proof, can be viewed as the converse of this fact in the stationary case.

LEMMA 3.16. *Let X be a random variable taking values in $\mathbb{A}^{\mathbb{Z}}$, whose law is σ -invariant. If $\mathcal{L}(\tau X) = \mu$, then $\mathcal{L}(X) = \mu$.*

Proof. Let ρ be a σ -invariant probability measure on $\mathbb{A}^{\mathbb{Z}}$, whose pushforward measure by τ is $\tau_*(\rho) = \mu$. In the dynamical system $(\mathbb{A}^{\mathbb{Z}}, \rho, \sigma)$, we consider the σ -process X_1 defined by the coordinates, and we set $Y_1 := \tau(X_1)$, so that Y_1 is a σ -process with law μ . We also consider a second dynamical system $(\mathbb{A}^{\mathbb{Z}}, \mu, \sigma)$, where we define the σ -process X_2 by the coordinates, and we set $Y_2 := \tau(X_2)$. We have $\mathcal{L}(X_2) = \mathcal{L}(Y_2) = \mu$. By Lemma 3.15 there exists a joining $(\bar{X}_1, \bar{Y}_1, \bar{X}_2, \bar{Y}_2)$ realized in a dynamical system $(\bar{\Omega}, \bar{\mathbb{P}}, \bar{T})$, such that:

- $\mathcal{L}(\bar{X}_1) = \rho$;
- $\mathcal{L}(\bar{X}_2) = \mu$;
- $\tau(\bar{X}_1) = \bar{Y}_1 = \bar{Y}_2 = \tau(\bar{X}_2)$.

Set $Z := \bar{X}_1 - \bar{X}_2$. We have $\tau(Z) = \tau(\bar{X}_1) - \tau(\bar{X}_2) = 0$, therefore for each $i \in \mathbb{Z}$,

$$Z(i) = Z(0) \circ \bar{T}^i = (-1)^i Z(0).$$

The process Z being 2-periodic, its Kolmogorov–Sinai entropy vanishes:

$$h(Z, \bar{T}) = 0.$$

Moreover, \bar{X}_2 has independent coordinates since its law is μ . By Theorem 3.14, \bar{X}_2 is disjoint of Z , hence independent of Z . But we have $\bar{X}_1 = Z + \bar{X}_2$, and the law of \bar{X}_2 is the Haar measure μ on $\mathbb{A}^{\mathbb{Z}}$. It follows that the law ρ of \bar{X}_1 is also μ . □

Note that the assumption of σ -invariance in Lemma 3.16 cannot be removed, because we could take a random variable Y of law μ , and define X as the only preimage of Y by τ whose coordinate X_0 satisfies $X_0 = 0_{\mathbb{A}}$. Then we would have $\mathcal{L}(\tau X) = \mathcal{L}(Y) = \mu$, but $\mathcal{L}(X) \neq \mu$ since \mathbb{A} is not the trivial group.

We now have all the tools to modify the end of the proof of Theorem 3.11, so that it can still be valid in the infinite case.

Alternative proof of Theorem 3.11. We consider the same situation as in the first proof, but we just modify the final argument. We know that Z_{n_0} has law μ , and $\tau^{|n_0|}Z_0 = Z_{n_0}$. All processes are stationary, so we can apply Lemma 3.16 $|n_0|$ times to prove recursively that each $Z_n, n_0 \leq n \leq 0$, has law μ . Observe also that we can choose the metric d on \mathbb{A} to be invariant by translation (if we start with an arbitrary metric d' defining the topology on \mathbb{A} , we can set $d(a, b) := d'(a - b, b - a)$). Then, for each $\delta > 0$, denoting by B_δ the ball of center $0_{\mathbb{A}}$ and with radius δ , we get

$$\mathbb{P}(d(X'_0(0), X''_0(0)) < \delta) = \mathbb{P}(Z_0(0) \in B_\delta) = U_{\mathbb{A}}(B_\delta).$$

But, as $\delta \rightarrow 0, U_{\mathbb{A}}(B_\delta)$ converges to $U_{\mathbb{A}}(\{0_{\mathbb{A}}\})$, and we have $U_{\mathbb{A}}(\{0_{\mathbb{A}}\}) < 1$ since \mathbb{A} is not the trivial group. This shows that the target $X_0(0)$ is not dynamically I-cosy.

3.3. The factor filtration $(\mathcal{F}^{\tau_\varepsilon}, \sigma)$. We return now to the case of a finite Abelian group, and we wish to study the dynamical standardness of the factor filtration $(\mathcal{F}^{\tau_\varepsilon}, \sigma)$, where $\mathcal{F}^{\tau_\varepsilon}$ is the filtration associated to the probabilistic cellular automaton τ_ε . This section is essentially devoted to the proof of the following result.

THEOREM 3.17. *There exists $0 < \varepsilon_0 < (|\mathbb{A}| - 1)/|\mathbb{A}|$ such that, for each $\varepsilon_0 < \varepsilon \leq (|\mathbb{A}| - 1)/|\mathbb{A}|$, $(\mathcal{F}^{\tau_\varepsilon}, \sigma)$ is dynamically standard.*

For this we will construct the filtration $\mathcal{F}^{\tau_\varepsilon}$ on a bigger space than $(\mathbb{A}^{\mathbb{Z}})^{\mathbb{Z}}$, namely we will consider the probability space $\Omega := [0, 1]^{\mathbb{Z} \times \mathbb{Z}}$ equipped with the product probability measure $\mathbb{P} := \lambda^{\otimes (\mathbb{Z} \times \mathbb{Z})}$ (λ denotes the Lebesgue measure on $[0, 1]$ here). On (Ω, \mathbb{P}) we define the measure-preserving \mathbb{Z}^2 -action of shift maps: for $m, j \in \mathbb{Z}$, we define

$$\sigma_{m,j} : \omega = (\omega_{n,i})_{n \in \mathbb{Z}, i \in \mathbb{Z}} \mapsto \sigma_{m,j}\omega := (\omega_{n+m,i+j})_{n \in \mathbb{Z}, i \in \mathbb{Z}},$$

and we simply denote by σ the usual left-shift, $\sigma := \sigma_{0,1}$. For each $n \in \mathbb{Z}, i \in \mathbb{Z}$, we set $U_n(i)(\omega) := \omega_{n,i}$, so that $(U_n(i))_{n \in \mathbb{Z}, i \in \mathbb{Z}}$ is a family of independent variables, uniformly distributed in $[0, 1]$. For each $n \in \mathbb{Z}$, we denote by U_n the σ -process $U_n := (U_n(i))_{i \in \mathbb{Z}}$. Let $\mathcal{G} = (\mathcal{G}_n)_{n \leq 0}$ be the filtration generated by $(U_n)_{n \leq 0}$. Then (\mathcal{G}, σ) is a factor filtration, and as the random variables $U_n, n \in \mathbb{Z}$, are independent, (\mathcal{G}, σ) is dynamically of product type.

Our purpose is to construct, in the dynamical system $(\Omega, \mathbb{P}, \sigma)$, a stationary Markov process $(X_n)_{n \in \mathbb{Z}}$ with transitions given by the Markov kernel τ_ε . We will still denote by $\mathcal{F}^{\tau_\varepsilon}$ the natural filtration of the negative-time part of this process, and we want to proceed in such a way that $(\mathcal{F}^{\tau_\varepsilon}, \sigma)$ is a factor filtration immersed in (\mathcal{G}, σ) . More precisely, we want the following five conditions to be realized:

- (a) $X_n = (X_n(i))_{i \in \mathbb{Z}} = (X_n(0) \circ \sigma^i)_{i \in \mathbb{Z}}$ is a σ -process;

- (b) $\mathcal{L}(X_n) = \mu$;
- (c) for each $x \in \mathbb{A}^{\mathbb{Z}}$, $\mathcal{L}(X_{n+1} | X_n = x) = \tau_\varepsilon(x, \cdot)$;
- (d) X_n is \mathcal{G}_n -measurable, namely, is a measurable function of $(U_m)_{m \leq n}$;
- (e) X_{n+1} is measurable with respect to $\Sigma(X_n, U_{n+1})$.

Conditions (a), (b) and (c) ensure that $(X_n)_{n \leq 0}$ generates the correct factor filtration $(\mathcal{F}^{\tau_\varepsilon}, \sigma)$ that we want to study. Condition (d) shows that $\mathcal{F}^{\tau_\varepsilon}$ is a subfiltration of \mathcal{G} , and condition (e) implies that $\mathcal{F}^{\tau_\varepsilon}$ is immersed in \mathcal{G} .

The key condition (d) is the most difficult to obtain. Our strategy will be strongly inspired by clever techniques introduced by Marcovici [13] to prove ergodicity in the Markov sense for some PCAs. Let us start by sketching the argument. It relies on the following fact, which is a straightforward consequence of (4): for each $n, i \in \mathbb{Z}$, the conditional distribution of $X_{n+1}(i)$ always satisfies

$$\mathcal{L}(X_{n+1}(i) | \mathcal{F}_n^{\tau_\varepsilon}) = \mathcal{L}(X_{n+1}(i) | X_n(i), X_n(i + 1)) \geq \tilde{\varepsilon} \mathcal{U}_{\mathbb{A}}. \tag{16}$$

Therefore, we are able to parametrize our Markov chain with the uniform random variables $U_n(i)$, through a recursion relation of the form

$$X_{n+1}(i) = \varphi_\varepsilon(X_n(i), X_n(i + 1), U_{n+1}(i)),$$

where the updating function $\varphi_\varepsilon(\cdot, \cdot, u)$ only depends on u provided $u \geq 1 - \tilde{\varepsilon}$; that is, we split the interval $[1 - \tilde{\varepsilon}, 1]$ into $|\mathbb{A}|$ subintervals I_a with length $\tilde{\varepsilon}/|\mathbb{A}|$ each, and set for each $a, b, c \in \mathbb{A}$ and $u \in I_a$,

$$\varphi_\varepsilon(b, c, u) := a.$$

(The complete construction of the updating function is given in §3.3.3.) In this way, the knowledge of the process U is sufficient to recover the values of the random variables $X_n(i)$ at some good sites (n, i) , directly if $U_n(i) > \tilde{\varepsilon}$, or indirectly through the recursion relation. Actually, these good sites are exactly those which are not connected to $-\infty$ in an oriented site percolation process introduced in the next section. We will see that, if $\tilde{\varepsilon}$ is not too small, then with probability 1 there is no site connected to $-\infty$, hence all sites are good and we get condition (d).

The whole construction of the process also borrows from [13] the concept of *envelope PCA*. This is an auxiliary probabilistic cellular automaton, presented in §3.3.2, which will be coupled with the percolation process in §3.3.4.

3.3.1. Oriented site percolation. The percolation model that we will use is built on an infinite oriented graph whose vertices (sites) are the elements of \mathbb{Z}^2 , and the oriented edges connect each site (n, i) to the sites $(n + 1, i)$ and $(n + 1, i - 1)$. We fix $p \in [0, 1]$ (this parameter will be specified later), and we use the random variables $U_n(i)$ to define a random configuration of the graph: the site (n, i) is declared to be *open* if $U_n(i) \in [0, p[$, and *closed* otherwise. In this random configuration, we say that a site (n, i) *leads to* a site (m, j) if $n \leq m$ and there exists in the oriented graph a path starting from (n, i) , ending in (m, j) and passing only through open sites, which in particular requires that both sites (n, i) and (m, j) be open (see Figure 4).

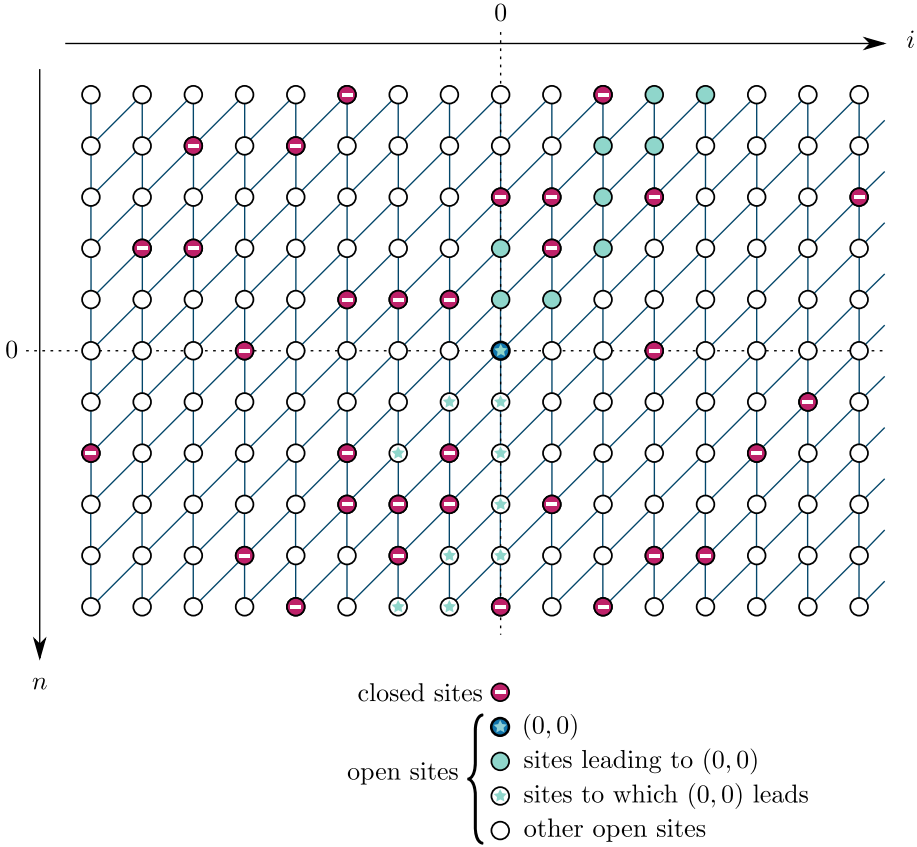


FIGURE 4. The oriented site percolation model.

For $n \geq 0$, we denote by O_n the event ‘there exists $i \in \mathbb{Z}$ such that $(0, 0)$ leads to (n, i) ’, and similarly, when $n \leq 0$, O_n stands for the events ‘there exists $i \in \mathbb{Z}$ such that (n, i) leads to $(0, 0)$ ’. We also set

$$O_\infty := \bigcap_{n \geq 0} O_n \quad \text{and} \quad O_{-\infty} := \bigcap_{n \leq 0} O_n.$$

We note that $\mathbb{P}(O_n) = \mathbb{P}(O_{-n})$ since exchanging each random variable $U_n(i)$ with $U_{-n}(-i)$ does not affect the law of the process, but exchanges O_n and O_{-n} . It follows that $\mathbb{P}(O_\infty) = \mathbb{P}(O_{-\infty})$.

The essential question in percolation theory asks whether the probability of O_∞ is positive or not. In the model presented here, this question has been addressed by Liggett [12], who proved the following theorem.

THEOREM 3.18. (Liggett) *There exists a critical value $p_c \in [\frac{2}{3}, \frac{3}{4}]$ such that:*

- $p < p_c \implies \mathbb{P}(O_\infty) = 0$;
- $p > p_c \implies \mathbb{P}(O_\infty) > 0$.

COROLLARY 3.19. *If $0 \leq p < p_c$, the integer-valued random variable*

$$N_0 := \min\{n \leq 0 : O_n \text{ does not hold}\}$$

is a.s. well defined, and it is measurable with respect to $\Sigma(U_n, n \leq 0)$.

3.3.2. *Envelope automaton.* The envelope automaton associated to τ_ε is an auxiliary PCA acting on configurations built on a larger alphabet: $\mathbb{A}_{\text{env}} := \mathbb{A} \cup \{?\}$. We associate to each finite word $w \in (\mathbb{A}_{\text{env}})^*$ the set \bar{w} of all words in \mathbb{A}^* obtained from w by replacing each question mark with any symbol from \mathbb{A} . For example, if $\mathbb{A} = \{0, 1\}$ and $w = ?1?$, we have $\bar{w} = \{010, 011, 110, 111\}$. Of course, if $w \in \mathbb{A}^*$, we have $\bar{w} = \{w\}$.

To define the envelope automaton we first have to introduce the *local transition rules* of τ_ε : if (X_n) is a Markov chain evolving according the Markov kernel τ_ε , we define, for $a, b, c \in \mathbb{A}$,

$$f_{\tau_\varepsilon}(a \mid bc) := \mathbb{P}(X_{n+1}(i) = a \mid X_n(i) = b, X_n(i+1) = c) \\ = \begin{cases} 1 - \varepsilon & \text{if } a = b + c, \\ \frac{\varepsilon}{|\mathbb{A}| - 1} & \text{otherwise.} \end{cases}$$

Now we define the local transition rules of the envelope PCA: if w is a word of length 2 over the alphabet \mathbb{A}_{env} , and $a \in \mathbb{A}$, we set

$$f_{\tau_\varepsilon}^{\text{env}}(a \mid w) := \min_{bc \in \bar{w}} f_{\tau_\varepsilon}(a \mid bc)$$

and

$$f_{\tau_\varepsilon}^{\text{env}}(? \mid w) := 1 - \sum_{a \in \mathbb{A}} \min_{bc \in \bar{w}} f_{\tau_\varepsilon}(a \mid bc).$$

Note that, if w contains at least one $?$, for each $a \in \mathbb{A}$ there exists $bc \in \bar{w}$ such that $b + c \neq a$. Recalling the assumption (2) we made on ε and the definition (3) of $\tilde{\varepsilon}$, we thus get the following rules: for each $a, b, c \in \mathbb{A}_{\text{env}}$,

$$f_{\tau_\varepsilon}^{\text{env}}(a \mid bc) = \begin{cases} f_{\tau_\varepsilon}(a \mid bc) & \text{if } a, b, c \in \mathbb{A}, \\ 0 & \text{if } a = ? \text{ and } b, c \in \mathbb{A}, \\ \frac{\varepsilon}{|\mathbb{A}| - 1} & \text{if } a \in \mathbb{A} \text{ and } ? \in \{b, c\}, \\ 1 - \tilde{\varepsilon} & \text{if } a = ? \text{ and } ? \in \{b, c\}. \end{cases}$$

In particular, the symbol $?$ cannot appear if the two preceding symbols are in \mathbb{A} , and in this case the local transition rules of the envelope PCA coincide with those of τ_ε .

For each $b, c \in \mathbb{A}_{\text{env}}$, $f_{\tau_\varepsilon}^{\text{env}}(\cdot \mid bc)$ can be viewed as a probability measure on \mathbb{A}_{env} . Then the envelope PCA $\tau_\varepsilon^{\text{env}}$ is defined as the Markov kernel on $\mathbb{A}_{\text{env}}^{\mathbb{Z}}$ given by

$$\tau_\varepsilon^{\text{env}}(\cdot \mid x) := \bigotimes_{i \in \mathbb{Z}} f_{\tau_\varepsilon}^{\text{env}}(\cdot \mid x(i)x(i+1)), \quad \text{for all } x \in \mathbb{A}_{\text{env}}^{\mathbb{Z}}.$$

3.3.3. *Simultaneous implementation of the two PCAs.* We will realize a coupling of τ_ε and $\tau_\varepsilon^{\text{env}}$ on our space Ω , using the independent uniform random variables $(U_n(i))_{n,i \in \mathbb{Z}}$.

	0	1 - $\tilde{\varepsilon}$	1
position of $u \in [0, 1]$	$I_?$	I_{a_1}	$I_{a_{ \mathbb{A} }}$
$b, c \in \mathbb{A}$: value of $\varphi_\varepsilon^{\text{env}}(b, c, u) = \varphi_\varepsilon(b, c, u)$	$b + c$	a_1	$a_{ \mathbb{A} }$
b or c is ?: value of $\varphi_\varepsilon^{\text{env}}(b, c, u)$?	a_1	$a_{ \mathbb{A} }$
state of site (n, i) for $u = U_n(i)$	open	closed	

FIGURE 5. Values of the updating functions and coupling with the percolation model.

This will be done by means of two updating functions,

$$\varphi_\varepsilon : \mathbb{A} \times \mathbb{A} \times [0, 1] \rightarrow \mathbb{A}$$

and

$$\varphi_\varepsilon^{\text{env}} : \mathbb{A}_{\text{env}} \times \mathbb{A}_{\text{env}} \times [0, 1] \rightarrow \mathbb{A}_{\text{env}},$$

adapted respectively to the local transition rules of τ_ε and $\tau_\varepsilon^{\text{env}}$. Here is how we define these updating functions (see Figure 5). We split the interval $[0, 1]$ into $|\mathbb{A}| + 1$ disjoint subintervals, labelled by symbols from \mathbb{A}_{env} . The first is

$$I_? := \left[0, 1 - \frac{\varepsilon|\mathbb{A}|}{|\mathbb{A}| - 1} \right[= [0, 1 - \tilde{\varepsilon}[,$$

and the complement of $I_?$ is itself split into $|\mathbb{A}|$ subintervals of equal length $\varepsilon/(|\mathbb{A}| - 1)$, denoted by $I_a, a \in \mathbb{A}$. Now for each $b, c \in \mathbb{A}, u \in [0, 1]$, we set

$$\varphi_\varepsilon(b, c, u) := \begin{cases} b + c & \text{if } u \in I_?, \\ a & \text{if } u \in I_a, a \in \mathbb{A}. \end{cases}$$

In this way, for a random variable U uniformly distributed in $[0, 1]$, we have, for each $b, c \in \mathbb{A}$,

$$\mathcal{L}(\varphi_\varepsilon(b, c, U)) = f_{\tau_\varepsilon}(\cdot | bc).$$

To define $\varphi_\varepsilon^{\text{env}}$, we use the same partition of $[0, 1]$ and we set, for $b, c \in \mathbb{A}_{\text{env}}$ and $u \in [0, 1]$,

$$\varphi_\varepsilon^{\text{env}}(b, c, u) := \begin{cases} b + c & \text{if } b, c \in \mathbb{A} \text{ and } u \in I_?, \\ ? & \text{if } b \text{ or } c \text{ is ? and } u \in I_?, \\ a & \text{if } u \in I_a, a \in \mathbb{A}. \end{cases}$$

Likewise, we see that for a random variable U uniformly distributed in $[0, 1]$, we have for each $b, c \in \mathbb{A}_{\text{env}}$

$$\mathcal{L}(\varphi_\varepsilon^{\text{env}}(b, c, U)) = f_{\tau_\varepsilon^{\text{env}}}(\cdot | bc).$$

Let $n_0 \leq 0$ be a fixed integer and $x \in \mathbb{A}^{\mathbb{Z}}$ be an arbitrary configuration. We realize simultaneously two Markov chains $(X_n^x)_{n \geq n_0}$ and $X^{\text{env}} = (X_n^{\text{env}})_{n \geq n_0}$ evolving respectively

according to τ_ε and $\tau_\varepsilon^{\text{env}}$. First, we initialize by setting $X_{n_0}^x := x$, and $X_{n_0}^{\text{env}} := \dots ????? \dots$ (the configuration with only symbols $?$). Later times of the processes are computed recursively through the formula

$$X_{n+1}^x(i) := \varphi_\varepsilon(X_n^x(i), X_n^x(i + 1), U_{n+1}(i)) \tag{17}$$

and

$$X_{n+1}^{\text{env}}(i) := \varphi_\varepsilon^{\text{env}}(X_n^{\text{env}}(i), X_n^{\text{env}}(i + 1), U_{n+1}(i)). \tag{18}$$

3.3.4. *Coupling with the percolation model.* We return now to the oriented site percolation model we defined in §3.3.1 from the *same* random variables $(U_n(i))_{n,i \in \mathbb{Z}}$. We just set the parameter p to be equal to $1 - \varepsilon$. In this way, the site (n, i) is open if and only if $U_n(i) \in I_?$ (see Figure 5).

The way we constructed the three processes together yields interesting properties which we describe in the two following lemmas.

LEMMA 3.20. *For each $n \geq n_0 + 1, i \in \mathbb{Z}$, we have $X_n^{\text{env}}(i) = ?$ if and only if in the percolation model there exists $j \in \mathbb{Z}$ such that $(n_0 + 1, j)$ leads to (n, i) .*

Proof. We prove the lemma by induction on $n \geq n_0 + 1$. Since $X_{n_0}^{\text{env}}(i) = ?$, we have $X_{n_0+1}^{\text{env}}(i) = ?$ if and only if $U_{n_0+1}(i) \in I_?$, that is, if and only if the site $(n_0 + 1, i)$ is open, which proves the result for $n = n_0 + 1$. Now assume that the result is true up to level $n \geq n_0 + 1$. By construction, $X_{n+1}^{\text{env}}(i) = ?$ if and only if the two following conditions hold:

- $X_n^{\text{env}}(i) = ?$ or $X_n^{\text{env}}(i + 1) = ?$, which by assumption means that there exists $j \in \mathbb{Z}$ such that $(n_0 + 1, j)$ leads to (n, i) or to $(n, i + 1)$;
- $U_{n+1}(i) \in I_?$, which means that the site $(n + 1, i)$ is open.

Therefore $X_{n+1}^{\text{env}}(i) = ?$ if and only if there exists $j \in \mathbb{Z}$ such that $(n_0 + 1, j)$ leads to $(n + 1, i)$, which ends the proof. □

The next lemma is a key result: it shows that wherever the process X^{env} displays symbols different from $?$, the process X^x displays the same symbols as X^{env} , *regardless of the initial configuration x .*

LEMMA 3.21. *For each $n \geq n_0, i \in \mathbb{Z}$ and $x \in \mathbb{A}^{\mathbb{Z}}$, if $X_n^{\text{env}}(i) \neq ?$ then $X_n^x(i) = X_n^{\text{env}}(i)$.*

Proof. Again we prove the result by induction on $n \geq n_0$. For $n = n_0$ this is obvious since $X_{n_0}^{\text{env}}(i) = ?$ for all $i \in \mathbb{Z}$. Now assume that the property holds up to some level $n \geq n_0$, and let us consider some site $(n + 1, i)$. There are two cases.

- First case: $X_n^{\text{env}}(i) = ?$ or $X_n^{\text{env}}(i + 1) = ?$. Then the only way to have $X_{n+1}^{\text{env}}(i) \neq ?$ is if $U_{n+1}(i) \in I_a$ for some $a \in \mathbb{A}$. By construction of the updating functions, this implies that $X_{n+1}^x(i) = a = X_{n+1}^{\text{env}}(i)$.
- Second case: $X_n^{\text{env}}(i) \in \mathbb{A}$ and $X_n^{\text{env}}(i + 1) \in \mathbb{A}$. Then by assumption this implies that we have both $X_n^x(i) = X_n^{\text{env}}(i)$ and $X_n^x(i + 1) = X_n^{\text{env}}(i + 1)$. By construction of the

updating functions, we deduce on the one hand that $X_{n+1}^{env}(i) \in \mathbb{A}$, and on the other hand that $X_{n+1}^x(i) = X_{n+1}^{env}(i)$.

In both cases the property also holds for the site $(n + 1, i)$. □

3.3.5. *Construction of the process (X_n) generating $\mathcal{F}^{\varepsilon}$.* We will now let n_0 vary in the preceding construction, and to make the dependence on n_0 more explicit we now denote $(X_n^{n_0,x})_{n \geq n_0}$ and $(X_n^{n_0,env})_{n \geq n_0}$ the processes respectively initialized at the configuration x and at the configuration \dots , and defined inductively by (17) and (18).

Since the initial configuration of the envelope processes is σ -invariant, the action of the shifts $\sigma_{n,i}$ on the envelope processes yields the following formula, valid for all $n_0 \leq n$, $i \in \mathbb{Z}$:

$$X_n^{n_0,env}(i) = X_0^{n_0-n,env}(0) \circ \sigma_{n,i}. \tag{19}$$

By a similar argument to the proof of Lemma (3.21), we get the following relations between all those processes.

LEMMA 3.22. *For each $n \geq n_0$, $i \in \mathbb{Z}$, if $X_n^{n_0,env}(i) \neq ?$, then for all $n_1 \leq n_0$ we have $X_n^{n_1,env}(i) = X_n^{n_0,env}(i)$.*

The next proposition will be used to define the process $(X_n)_{n \in \mathbb{Z}}$, evolving according to τ_ε , under the assumption that ε is not too close to 0. We recall that p_c stands for the critical value for the oriented site percolation process, defined in Theorem 3.18.

PROPOSITION 3.23. *Assume that*

$$(1 - p_c) \frac{|\mathbb{A}| - 1}{|\mathbb{A}|} < \varepsilon < \frac{|\mathbb{A}| - 1}{|\mathbb{A}|}, \quad \text{that is, } 0 < 1 - \tilde{\varepsilon} < p_c. \tag{20}$$

Then, with probability 1, for all $n, i \in \mathbb{Z}$, there exists a random variable $N_0(n, i) \leq n$, taking values in \mathbb{Z} , and satisfying the following properties.

- $N_0(n, i)$ is measurable with respect to $\mathcal{G}_n = \Sigma(U_m(j) : m \leq n, j \in \mathbb{Z})$.
- For each integer $n_1 \leq n_0 \leq N_0(n, i)$, we have

$$X_n^{n_1,env}(i) = X_n^{n_0,env}(i) \neq ?.$$

- For each $n, i \in \mathbb{Z}$, $N_0(n, i) = N_0(0, 0) \circ \sigma_{n,i}$.

Proof. Recall that the right-hand-side inequality, $\varepsilon < (|\mathbb{A}| - 1)/|\mathbb{A}|$, is just the reformulation of condition (2). The left-hand-side inequality, $(1 - p_c)((|\mathbb{A}| - 1)/|\mathbb{A}|) < \varepsilon$, is equivalent to $p < p_c$, where $p = 1 - \tilde{\varepsilon} = 1 - (\varepsilon|\mathbb{A}|/(|\mathbb{A}| - 1))$ is the parameter of the percolation process constructed in the coupling.

Consider first the case $(n, i) = (0, 0)$. For $p < p_c$, Corollary 3.19 ensures that the random variable $N_0 = N_0(0, 0) = \min\{m \leq 0 : \mathcal{O}_m \text{ does not hold}\}$ is a.s. well defined and \mathcal{G}_0 -measurable. Then Lemma 3.20 shows that, for each $n_0 \leq N_0(0, 0)$, $X_0^{n_0,env}(0) \neq ?$, and Lemma 3.22 gives

$$n_1 \leq n_0 \leq N_0(0, 0) \implies X_0^{n_1,env}(0) = X_0^{n_0,env}(0).$$

Then we can generalize this result to any site $(n, i) \in \mathbb{Z} \times \mathbb{Z}$, setting $N_0(n, i) := N_0 \circ \sigma_{n,i}$. Note that $N_0(n, i)$ is the greatest integer $n_0 \leq n$ such that there is no $j \in \mathbb{Z}$ with (n_0, j) leading to (n, i) . \square

We assume now that ε satisfies (20), and we explain how to construct the process (X_n) satisfying conditions (a)–(e) (see page 21). With the above proposition, we can a.s. define, for each $n, i \in \mathbb{Z}$,

$$X_n(i) := X_n^{N_0(n,i),\text{env}}(i) = \lim_{n_0 \rightarrow -\infty} X_n^{n_0,\text{env}}(i). \tag{21}$$

For each $n \leq n_0 \in \mathbb{Z}$, $X_n^{n_0,\text{env}}$ is measurable with respect to $\mathcal{G}_n = \Sigma(U_m(j))_{m \leq n, j \in \mathbb{Z}}$, so the same holds for X_n and we have condition (d).

Considering an integer $n_0 \leq \min\{N_0(n, i), N_0(n, i + 1), N_0(n + 1, i)\}$, we have simultaneously $X_n(i) = X_n^{n_0,\text{env}}(i)$, $X_n(i + 1) = X_n^{n_0,\text{env}}(i + 1)$ and $X_{n+1}(i) = X_{n+1}^{n_0,\text{env}}(i)$. As the process $(X_n^{n_0,\text{env}})$ satisfies the induction relation (18) at $(n + 1, i)$, it is therefore the same for (X_n) . This proves that the process $(X_n)_{n \in \mathbb{Z}}$ actually evolves according to the PCA τ_ε , and we have conditions (c) and (e). Moreover, from (19) we get, for each $n, i \in \mathbb{Z}$.

$$X_n(i) = X_0(0) \circ \sigma_{n,i}.$$

This gives on the one side that each row X_n is a σ -process, so we have condition (a), and on the other side we also get $X_n = X_0 \circ \sigma_{n,0}$. The rows X_n therefore all follow the same law, which has to be μ by Corollary 2.17, and we have condition (b). This concludes the proof of Theorem 3.17.

Remark 3.24. As already pointed out in the sketch of the proof, the argument relies on inequality (16), hence the result remains true for any PCA satisfying such inequality. In particular, the method can be applied to any probabilistic cellular automaton on $\mathbb{A}^{\mathbb{Z}}$ which is a random perturbation of a deterministic cellular automaton by addition of a random error ξ , provided the law of each $\xi_n(i)$ is bounded below by $\tilde{\varepsilon} \mathcal{U}_{\mathbb{A}}$.

We can even note that, in (16), the uniform measure $\mathcal{U}_{\mathbb{A}}$ does not play a particular role and can be replaced by any fixed probability measure on \mathbb{A} (we just have to adapt the length of the subintervals $(I_a)_{a \in \mathbb{A}}$ accordingly). Therefore the argument also applies to any PCA γ on $\mathbb{A}^{\mathbb{Z}}$ for which the conditional law of $X_{n+1}(i)$ is of the form

$$\mathbb{P}(X_{n+1}(i) = a \mid \Sigma(X_m : m \leq n)) = f_\gamma(a \mid X_n(i), X_n(i + 1)),$$

and the local transition rules f_γ satisfy

$$\text{for all } a, b, c \in \mathbb{A}, \quad f_\gamma(a \mid bc) \geq \tilde{\varepsilon} \rho(a),$$

for some fixed probability measure ρ on \mathbb{A} (here no algebraic assumption on \mathbb{A} is needed).

3.3.6. Questions about the factor filtration $(\mathcal{F}^{\tau_\varepsilon}, \sigma)$. When ε satisfies (20), we prove the dynamical standardness of $(\mathcal{F}^{\tau_\varepsilon}, \sigma)$ by immersing it into a factor filtration (\mathcal{G}, σ) dynamically of product type. The way we construct \mathcal{G} , it is clear that it carries more information than $\mathcal{F}^{\tau_\varepsilon}$, so that $\mathcal{F}^{\tau_\varepsilon}$ is a strict subfiltration of \mathcal{G} . Thus we may ask whether, in this case, $(\mathcal{F}^{\tau_\varepsilon}, \sigma)$ is itself a factor filtration dynamically of product type.

Another natural question is of course whether the dynamical standardness of $(\mathcal{F}^{\tau_\varepsilon}, \sigma)$ persists as ε gets close to 0, which prevents the percolation argument from applying. One could already ask whether, for ε close to 0, the criterion of dynamical I-cosiness is satisfied for the factor filtration $(\mathcal{F}^{\tau_\varepsilon}, \sigma)$. We do not have the answer to this question, but we provide below a somewhat naive argument suggesting that $(\mathcal{F}^{\tau_\varepsilon}, \sigma)$ might not be dynamically I-cosy when ε is close enough to 0.

We consider the case where the group \mathbb{A} is $\mathbb{Z}/2\mathbb{Z}$. In an attempt to establish dynamical I-cosiness, we try to realize a real-time dynamical coupling of $(\mathcal{F}^{\tau_\varepsilon}, \sigma)$, independent in the distant past, and for which the two copies X' and X'' of the process $X = (X_n)$ satisfy $X'_0(0) = X''_0(0)$ with probability close to 1. We want to construct this coupling in a dynamical system (Ω, \mathbb{P}, T) , in which for some $n_0 < 0$ we already have two independent copies $(X'_n)_{n \leq n_0}$ and $(X''_n)_{n \leq n_0}$ of $(X_n)_{n \leq n_0}$ (of course, each X'_n and each X''_n is supposed to be a T -process of law μ). We assume that we also have an independent family $(U_n)_{n \geq n_0+1} = (U_n(i) : i \in \mathbb{Z}_{n \geq n_0+1})$ of T -processes, each U_n being of law $\lambda^{\otimes \mathbb{Z}}$ (λ is the Lebesgue measure on $[0, 1]$), and the U_n being independent. We want to use this family to construct error processes $(\xi'_n)_{n \geq n_0+1}$ and $(\xi''_n)_{n \geq n_0+1}$, and then complete the processes X' and X'' inductively via the usual relations

$$X'_n = \tau X'_{n-1} + \xi'_n \quad \text{and} \quad X''_n = \tau X''_{n-1} + \xi''_n \quad (n \geq n_0 + 1).$$

As in §2.4, we still want to define the error processes in the form $\xi'_n(i) = g'_{n,i}(U_n(i))$ and $\xi''_n(i) = g''_{n,i}(U_n(i))$, where $g'_{n,i}$ and $g''_{n,i}$ are random updating functions, but we point out the essential difference with the construction in §2.4: we now want the law of the coupling to be invariant with respect to the left shift of coordinates. This means that these updating functions are still allowed to depend on the realizations of the processes up to time $n - 1$, but in an equivariant way. In other words, they can be chosen according to what we see around the site (n, i) , but the rule for the choice should be independent of i .

We also introduce the auxiliary process $Z := X' - X''$, which satisfies the same relation (10) for $n \geq n_0 + 1$. To get dynamical I-cosiness, we have to realize the coupling in such a way that

$$\mathbb{P}(Z_0(i) = 0) \xrightarrow[n_0 \rightarrow -\infty]{} 1$$

(note that the probability on the left-hand side does not depend on i).

Here is a natural idea to get what we are aiming at: we wish to encourage as much as possible the appearance of 0s in the Z process. For this, let us observe that we have, for each $n, i \in \mathbb{Z}$,

$$\begin{aligned} 0 \leq \mathbb{P}(Z_n(i) \neq \tau Z_{n-1}(i)) &= \mathbb{P}(\xi'_n(i) \neq \xi''_n(i)) \\ &\leq \mathbb{P}(\xi'_n(i) \neq 0 \text{ or } \xi''_n(i) \neq 0) \\ &\leq \mathbb{P}(\xi'_n(i) \neq 0) + \mathbb{P}(\xi''_n(i) \neq 0) \\ &\leq 2\varepsilon. \end{aligned} \tag{22}$$

Moreover, we have the possibility of making the inequality on the left-hand side an equality, by deciding to take $g'_{n,i} = g''_{n,i}$ (we call this *Option 0*). On the other hand, we can also choose to turn the right-hand-side inequalities into equalities, by choosing $g'_{n,i}$ and

$g''_{n,i}$ in such a way that the subsets $\{u \in [0, 1] : g'_{n,i}(u) = 1\}$ and $\{u \in [0, 1] : g''_{n,i}(u) = 1\}$ are disjoint (we call this *Option 1*). So a natural dynamical coupling is obtained by the following strategy. Knowing the processes up to time $n - 1$, we compute $\tau Z_{n-1}(i)$. If $\tau Z_{n-1}(i) = 0$, we decide to keep this 0 by choosing Option 0. If $\tau Z_{n-1}(i) = 1$, we maximize the probability that $Z_n(i)$ will still be 0 by choosing Option 1.

Applying systematically this strategy, we are left with a process $(Z_n)_{n \geq n_0}$ taking values in $\mathbb{Z}/2\mathbb{Z}$, which evolves according to the following PCA rule: for each n, i , if $\tau Z_{n-1}(i) = 0$ then $Z_n(i) = 0$, otherwise $Z_n(i) = 0$ with probability 2ε . Observe that the initial configuration Z_{n_0} has law $\mathcal{U}_{\mathbb{Z}/2\mathbb{Z}}^{\otimes \mathbb{Z}}$, so that the initial density of 1s is $1/2$. The question now is whether those 1s can survive over the long term.

It turns out that this type of evolution falls into the domain of a family of PCAs studied by Bramson and Neuhauser [1], who proved that if ε is small enough, then symbols 1 survive. More precisely, we can state the following consequence of their result.

THEOREM 3.25. (Bramson and Neuhauser) *In the coupling described above, for $\varepsilon > 0$ small enough, there exists $\rho = \rho(\varepsilon) > 0$ such that*

$$\liminf_{n_0 \rightarrow -\infty} \mathbb{P}(Z_0(0) = 1) \geq \rho.$$

This shows that this attempt to get dynamical I-cosiness for $(\mathcal{F}^\varepsilon, \sigma)$ fails if ε is small.

3.4. A partial converse to Theorem 3.10. The question whether dynamical I-cosiness of a factor filtration implies its dynamical standardness (as in the static case) is so far open. However, the construction given in §3.3 (for ε not too close to 0) provides an example where we are able to derive dynamical standardness from a particular form of dynamical I-cosiness. We want now to formalize this kind of situation, which can be viewed as the dynamic counterpart of the so-called *Rosenblatt self-joining criterion* in [11].

In a dynamical system (Ω, \mathbb{P}, T) , consider a factor filtration (\mathcal{F}, T) generated by a family $X = (X_n)_{n \leq 0}$ of T -processes, each $X_n(i)$ taking its values in a fixed finite set \mathbb{A} . We also assume the existence of a factor filtration (\mathcal{G}, T) , dynamically of product type, and generated by a family of independent T -processes $(U_n)_{n \leq 0}$ satisfying, for each $n \leq 0$,

- U_n is independent of $\mathcal{F}_{n-1} \vee \mathcal{G}_{n-1}$,
- $\mathcal{F}_n \subset \mathcal{F}_{n-1} \vee \Sigma(U_n)$.

(The two conditions above mean that $(U_n)_{n \leq 0}$ is a *superinnovation* of \mathcal{F} , according to [11].) In particular, there exists a measurable map ϕ_n such that

$$X_n = \phi_n(U_n, X_{n-1}, X_{n-2}, \dots).$$

We can also assume without loss of generality that there exists in (Ω, \mathbb{P}, T) a copy X' of X , whose coordinates X'_n are also T -processes, and which is independent of $\mathcal{G}_0 \vee \mathcal{F}_0$ (if not, we place ourselves on the product $(\Omega \times \Omega, \mathbb{P} \otimes \mathbb{P}, T \times T)$ as in the proof of Lemma 3.8). Now we build a family of dynamical real-time couplings of (\mathcal{F}, T) in the following way: for a fixed $n_0 \leq 0$, define $\bar{X}^{n_0} = (\bar{X}_n^{n_0})_{n \leq 0}$ by

- $\bar{X}_n^{n_0} = X'_n$ for $n \leq n_0$,
- then inductively, for $n_0 + 1 \leq n \leq 0$, $\bar{X}_n^{n_0} := \phi_n(U_n, \bar{X}_{n-1}^{n_0}, \bar{X}_{n-2}^{n_0}, \dots)$.

(The important point here is that, for $n \geq n_0 + 1$, the *same* random variable U_n is involved in the inductive relations giving X_n and $\bar{X}_n^{n_0}$.) Then the filtration $\bar{\mathcal{F}}^{n_0}$ generated by \bar{X}^{n_0} is dynamically isomorphic to \mathcal{F} , and the pair $((\mathcal{F}, T), (\bar{\mathcal{F}}^{n_0}, T))$ is a dynamical real-time coupling of (\mathcal{F}, T) , which is n_0 independent.

Definition 3.26. (Simple dynamical I-cosiness) If, in the above construction, we have, for each integer $m \leq 0$,

$$\mathbb{P}(\bar{X}_m^{n_0}(0) = X_m(0)) \xrightarrow{n_0 \rightarrow -\infty} 1, \tag{23}$$

then we say that the factor filtration (\mathcal{F}, T) is *simply dynamically I-cosy*.

Clearly, simple dynamical I-cosiness implies dynamical I-cosiness. Observe also that, when ε satisfies (20), the results of §3.3 prove that the factor filtration $(\mathcal{F}^{\varepsilon}, \sigma)$ is simply dynamically I-cosy.

The following theorem generalizes the conclusion of §3.3.

THEOREM 3.27. *If the factor filtration (\mathcal{F}, T) is simply dynamically I-cosy, then it is dynamically standard.*

Proof. We fix some integer $m \leq 0$. For each $n_0 \leq m - 1$, we denote by $U_{n_0+1}^m$ the collection of T -processes U_{n_0+1}, \dots, U_m . We have

$$\begin{aligned} \mathbb{P}(\bar{X}_m^{n_0}(0) = X_m(0)) &= \sum_{a \in \mathbb{A}} \mathbb{P}(\bar{X}_m^{n_0}(0) = X_m(0) = a) \\ &= \sum_{a \in \mathbb{A}} \mathbb{E}_{\mathbb{P}}[\mathbb{P}(\bar{X}_m^{n_0}(0) = X_m(0) = a \mid U_{n_0+1}^m)]. \end{aligned}$$

But $\bar{X}_m^{n_0}(0)$ and $X_m(0)$ are independent conditionally on $U_{n_0+1}^m$: indeed, once we know $U_{n_0+1}^m$, $\bar{X}_m^{n_0}(0)$ only depends on \bar{X}_{n_0} while $X_m(0)$ only depends on X_{n_0} . We thus get

$$\begin{aligned} \mathbb{P}(\bar{X}_m^{n_0}(0) = X_m(0)) &= \sum_{a \in \mathbb{A}} \mathbb{E}_{\mathbb{P}}[\mathbb{P}(\bar{X}_m^{n_0}(0) = a \mid U_{n_0+1}^m) \mathbb{P}(X_m(0) = a \mid U_{n_0+1}^m)] \\ &= \sum_{a \in \mathbb{A}} \mathbb{E}_{\mathbb{P}}[\mathbb{P}(X_m(0) = a \mid U_{n_0+1}^m)^2] \\ &= \mathbb{E}_{\mathbb{P}} \left[\sum_{a \in \mathbb{A}} \mathbb{P}(X_m(0) = a \mid U_{n_0+1}^m)^2 \right]. \end{aligned}$$

Let M_{n_0} be the random variable measurable with respect to $\Sigma(U_{n_0+1}^m)$ and defined by

$$M_{n_0} := \arg \max_{a \in \mathbb{A}} \mathbb{P}(X_m(0) = a \mid U_{n_0+1}^m).$$

(If the maximum is reached for several symbols in \mathbb{A} , we take the first of them with respect to some previously chosen total order on \mathbb{A} .) We get

$$\begin{aligned} \mathbb{P}(\bar{X}_m^{n_0}(0) = X_m(0)) &\leq \mathbb{E}_{\mathbb{P}} \left[\sum_{a \in \mathbb{A}} \mathbb{P}(X_m(0) = M_{n_0} \mid U_{n_0+1}^m) \mathbb{P}(X_m(0) = a \mid U_{n_0+1}^m) \right] \\ &= \mathbb{E}_{\mathbb{P}} \left[\mathbb{P}(X_m(0) = M_{n_0} \mid U_{n_0+1}^m) \underbrace{\sum_{a \in \mathbb{A}} \mathbb{P}(X_m(0) = a \mid U_{n_0+1}^m)}_{=1} \right] \\ &= \mathbb{P}(X_m(0) = M_{n_0}). \end{aligned}$$

But we assumed that (\mathcal{F}, T) is simply dynamically I-cosy. So we have (23), which yields

$$\mathbb{P}(X_m(0) = M_{n_0}) \xrightarrow[n_0 \rightarrow -\infty]{} 1.$$

For each integer $k \geq 1$, we can thus find $n_k \leq m - 1$ such that

$$\mathbb{P}(X_m(0) \neq M_{n_k}) < 2^{-k}.$$

By Borel–Cantelli we get $X_m(0) = \lim_{k \rightarrow \infty} M_{n_k}$ (\mathbb{P} -a.s.). Moreover, M_{n_k} is always \mathcal{G}_m -measurable. Therefore $X_m(0)$ is measurable with respect to \mathcal{G}_m , and since the sigma-algebras are factor sigma-algebras, the whole T -process X_m is \mathcal{G}_m -measurable.

This already proves that \mathcal{F} is a subfiltration of the natural filtration \mathcal{G} of the process $(U_n)_{n \leq 0}$, hence $(U_n)_{n \leq 0}$ is a *generating* superinnovation of \mathcal{F} . By [11], it follows that \mathcal{F} is immersed in \mathcal{G} . Since the factor filtration (\mathcal{G}, T) is dynamically of product type, we get the dynamical standardness of (\mathcal{F}, T) . □

Acknowledgements. The authors wish to thank Irène Marcovici for fruitful conversations which inspired the subject of this paper, and Séverin Benzoni, Christophe Leuridan, Alejandro Maass and an anonymous referee for their careful reading, relevant questions and valuable comments. Research partially supported by the Regional Research Project Moustic (Normandie).

REFERENCES

- [1] M. Bramson and C. Neuhauser. Survival of one-dimensional cellular automata under random perturbations. *Ann. Probab.* **22**(1) (1994), 244–263.
- [2] X. Bressaud, A. Maass, S. Martinez and J. San Martin. Stationary processes whose filtrations are standard. *Ann. Probab.* **34**(4) (2006), 1589–1600.
- [3] G. Ceillier and C. Leuridan. Sufficient conditions for the filtration of a stationary processes to be standard. *Probab. Theory Related Fields* **167**(3–4) (2017), 979–999.
- [4] T. de la Rue. An introduction to joinings in ergodic theory. *Discrete Contin. Dyn. Syst.* **15**(1) (2006), 121–142.
- [5] L. Dubins, J. Feldman, M. Smorodinsky and B. Tsirelson. Decreasing sequences of σ -fields and a measure change for Brownian motion. I. *Ann. Probab.* **24**(2) (1996), 882–904.
- [6] M. Émery and W. Schachermayer. On Vershik’s standardness criterion and Tsirelson’s notion of cosiness. *Séminaire de Probabilités XXXV*. Ed. J. Azéma, M. Émery, M. Ledoux and M. Yor. Springer, Berlin, 2001, pp. 265–305.
- [7] H. Furstenberg. Disjointness in ergodic theory, minimal sets, and a problem in Diophantine approximation. *Math. Syst. Theory* **1** (1967), 1–49.

- [8] C. Hoffman. A zero entropy T such that the $[T, \text{Id}]$ endomorphism is nonstandard. *Proc. Am. Math. Soc.* **128**(1) (2000), 183–188.
- [9] C. Hoffman and D. Rudolph. A dyadic endomorphism which is Bernoulli but not standard. *Israel J. Math.* **130** (2002), 365–379.
- [10] P. Lanthier. Aspects ergodiques et algébriques des automates cellulaires. *PhD Thesis*, Université de Rouen Normandie, 2020.
- [11] S. Laurent. On standardness and I-cosiness. *Séminaire de Probabilités XLIII, Poitiers, France, Juin 2009*. Ed. C. Donati-Martin, A. Lejay and A. Rouault. Springer, Berlin, 2011, pp. 127–186.
- [12] T. M. Liggett. Survival of discrete time growth models, with applications to oriented percolation. *Ann. Appl. Probab.* **5**(3) (1995), 613–636.
- [13] I. Marcovici. Ergodicity of noisy cellular automata: the coupling method and beyond. *Pursuit of the Universal. Proc. 12th Conf. on Computability in Europe, CiE 2016 (Paris, France, June 27–July 1, 2016)*. Ed. A. Beckmann, L. Bienvenu and N. Jonoska. Springer, Cham, 2016, pp. 153–163.
- [14] M. Smorodinsky. Processes with no standard extension. *Israel J. Math.* **107** (1998), 327–331.
- [15] B. Tsirelson. Triple points: from non-Brownian filtrations to harmonic measures. *Geom. Funct. Anal.* **7**(6) (1997), 1096–1142.
- [16] A. M. Vershik. Decreasing sequences of measurable partitions and their applications. *Sov. Math. Dokl.* **11** (1970), 1007–1011.
- [17] A. M. Vershik. The theory of decreasing sequences of measurable partitions. *St. Petersburg Math. J.* **6**(4) (1994), 1–68.