Ergod. Th. & Dynam. Sys. (2017), **37**, 129–145
doi:10.1017/etds.2015.39

Measure evolution of cellular automata and of finitely anticipative transformations

PIERRE COLLET† and SERVET MARTÍNEZ‡

† *Centre de Physique Théorique, CNRS UMR 7644, Ecole Polytechnique F-91128 Palaiseau Cedex, France (e-mail: collet@cpht.polytechnique.fr)*

‡ *Departamento Ingeniería Matemática and Centro Modelamiento Matemático, Universidad de Chile, UMI 2807 CNRS, Casilla 170-3, Correo 3, Santiago, Chile (e-mail: smartine@dim.uchile.cl)*

(*Received* 10 *October* 2014 *and accepted in revised form* 13 *April* 2015)

Abstract. The evolution of cellular automata and of finitely anticipative transformations is studied by using right sets. These are the sets of symbols that are compatible with a past of a position and the respective coordinate of the transformation. Our main result shows, under some suitable conditions, that if the entropy converges to zero then the right sets increase towards the whole alphabet. We discuss these concepts with Wolfram automata.

1. *Main concepts*

Here we study explicit and computable necessary conditions in order that the entropy of iterated cellular automata—or more generally the iterated of finitely anticipative transformations—converges to zero.

The evolution of cellular automata has been widely studied. Many of these works dealt with the limit set, which is the intersection of all the iterated sets of the cellular automaton, see for example [**11**].

When the iterated sets of a cellular automaton starting from a fullshift stabilize after a finite number of steps, the automaton is called stable, and the limit set is a factor of a fullshift. Hence, it is a sofic system and the set of its words is a regular language. In [**9**] it was shown that any subshift of finite type with a fixed point can be the limit set of a stable cellular automaton and an analogous result is proved for a sofic system of almost of finite type.

We focus on the case of unstable limit sets, that is when the sequence of iterated sets of a cellular automaton is strictly decreasing. In this case the language of the limit set cannot be the one of a subshift of finite type, see [**6**]. In [**5**] there are given examples of limit sets having non-regular-context-free, or non-context-free context-sensitive, or non-recursive enumerable languages. This complexity was also studied in [**2**]. Several works have been devoted to study the limit sets of particular rules, thus in [**12**] it is shown that the limit set of Wolfram rule 122 is neither regular nor context-free.

In §§1.1 and 1.2 we introduce cellular automata and finitely anticipative shift-commuting measurable transformations. For these last mappings there exists some fixed number r such that, for all q, the past of the image point up to coordinate q only depends on the past of the point up to coordinate $q+r$. This class of mappings is considerably larger than cellular automata because they are not necessarily continuous.

In §2 we introduce right sets and right events for finitely anticipative shift-commuting measurable transformations. Right sets are the sets of symbols that are compatible with the past of a point and the respective coordinate of the automaton. They correspond to the compatible right extensions of length 1, introduced in [**3**] for surjective automata, see [**4**, §2].

For finitely anticipative mappings starting from a shift-invariant probability μ having lower bounded transition probabilities, we show in Theorem 3.1 that a necessary condition in order that the entropy converges to zero is that the right set converges to the whole alphabet μ-a.e. (almost everywhere), that is if any letter is compatible with the past of a point and the respective symbol of the mapping μ-a.e.

We note that the right set is a useless notion for right permutative cellular automata because the right set remain fixed along the iterations; in fact at any time it is a singleton. So, for these transformations our result does not apply.

1.1. *Shift-commuting, continuous and measurable transformations.* Let A be a finite non-empty alphabet; to avoid trivial situations we assume it is not a singleton. We fix $\Lambda := A^{\mathbb{Z}}$ the fullshift space and $F : \Lambda \to \Lambda$ a transformation, which is assumed to be *non-constant*, that there it does not exist $y^* \in \Lambda$ such that $F(x) = y^*$ for all $x \in \Lambda$.

The set Λ endowed with the usual metric is a compact metric space. We denote by $\mathcal{B}$ its Borel $\boldsymbol{\sigma}$-field. An element of Λ is denoted by $x = (x(n) : n \in \mathbb{Z})$. Let $\sigma : \Lambda \to \Lambda$, $x \to (\sigma x)(n) = x(n+1)$, for all $n \in \mathbb{Z}$, be the shift transformation which is an homeomorphism, so it is bimeasurable with respect to $\mathcal{B}$.

From now on, the transformation $F : \Lambda \to \Lambda$ is assumed to be *shift-commuting*, that is it satisfies $F \circ \sigma = \sigma \circ F$. Note that this implies $F^i \circ \sigma^j = \sigma^j \circ F^i$ for all $i \geq 1$ and $j \in \mathbb{Z}$, in particular $F^i : \Lambda \to \Lambda$ is shift-commuting for all $i \geq 1$.

For $-\infty \leq p \leq q \leq \infty$ define $Z_p^q = \{m \in \mathbb{Z} : p \leq m \leq q\}$ and for $x \in \Lambda$ denote $x_p^q = (x(n) : n \in Z_p^q) \in A^{Z_p^q}$. When $p \leq q$ are finite, we identify $A^{Z_p^q}$ with A^{q-p+1}.

The transformation F is a *cellular automaton* (*c.a.*) if it is continuous. In [**3**] this is shown to be equivalent to the existence of integers $l \leq r$ and $f : A^{Z_l^r} \to A$, called a local rule, that satisfies

$$\text{for all } x \in \Lambda, \quad \text{for all } n \in \mathbb{Z}: \quad Fx(n) = f(x_{n+l}^{n+r}).$$

Let us discuss measurability of F in the following framework. Let μ be a fixed σ-invariant probability measure on $(\Lambda, \mathcal{B})$, so it satisfies $\mu = \mu \circ \sigma^{-1}$. We will always assume that $\mathcal{B}$ has been μ-completed with the μ-negligible sets and by an abuse of notation we will continue to denote it by $\mathcal{B}$. From Lusin's theorem, see [**1**, p. 69], we have that $F : \Lambda \to \Lambda$ is a measurable transformation with respect to the μ-completed σ-field, if and only if, for all $m > 0$, there exists a compact set $K(m) \subseteq \Lambda$ such that $\mu(\Lambda \setminus K(m)) < 1/m$ and the restriction $F : K(m) \to \Lambda$ is continuous.

Note that $\mu(\bigcup_{m>0} K(m)) = 1$. Since $F : K(m) \to \Lambda$ is continuous, the shift-commuting property implies $F : \sigma^j K(m) \to \Lambda$ is also continuous. Then, in our framework the set $\bigcup_{m>0} K(m)$ can be assumed to be a σ-invariant set (of full probability).

Let $\Lambda_0 \in \mathcal{B}$. We have that the restriction $F : \Lambda_0 \to \Lambda$ is continuous at some point $x \in \Lambda_0$ if and only if, for all $k \geq 0$, there exists $l_x(k)$ and $r_x(k)$ with $l_x(k) \leq r_x(k)$ and such that

$$\begin{aligned} &\text{for all } y \in \Lambda_0, \quad y(j) = x(j) \quad \text{for all } j \in \{l_x(k), \dots, r_x(k)\} \\ &\quad \text{implies} \quad Fy(j) = Fx(j) \quad \text{for all } |j| \leq k. \end{aligned} \tag{1}$$

If $F : \Lambda_0 \to \Lambda$ is continuous for some $\Lambda_0 \in \mathcal{B}$ with $\mu(\Lambda_0) = 1$, then F is said to be μ-almost surely (a.s.) continuous. In this case F is measurable because Λ_0 can be interiorly approximated by compact sets. In §1.2.1, we will describe the μ-a.s. continuous transformations by using trees. We note that if $F : \Lambda_0 \to \Lambda$ is continuous, then $F : \bigcup_{j \in \mathbb{Z}} \sigma^j(\Lambda_0) \to \Lambda$ is also continuous, and so Λ_0 can be assumed to be σ-invariant.

From now on we assume that

$$F : \Lambda \to \Lambda \text{ is a shift-commuting measurable transformation.}$$

1.2. *Finitely anticipative transformations.* The transformation F is said to be *finitely anticipative* (*f.a.*) if there exists $r \in \mathbb{Z}$, called an anticipation (of F), such that

$$\text{for all } x \in \Lambda, \text{ for all } n \in \mathbb{Z}: \quad \text{the coordinate } (Fx)(n) \text{ only depends on } x_{-\infty}^{n+r}. \tag{2}$$

We also say that F has anticipation r.

Remark 1.1. In the definition of anticipation we are not requiring that it satisfies some sort of minimality. That is, an anticipation is not a unique number but any of the numbers that satisfies (2). Indeed, if r is an anticipation of F then any $r' \geq r$ is also an anticipation.

Let F be *f.a.* having anticipation r. We define the mapping

$$f_r^+ : A^{Z_{-\infty}^r} \to A,\ x_{-\infty}^r \to f_r^+(x_{-\infty}^r) = (Fx)(0). \tag{3}$$

For any other anticipation r' of F we have

$$f_{r'}^+(x_{-\infty}^{r'}) = (Fx)(0) = f_r^+(x_{-\infty}^r). \tag{4}$$

The shift-commuting property implies

$$(Fx)(n) = (F\sigma^n x)(0) = f_r^+((\sigma^n x)_{-\infty}^r) = f_r^+(x_{-\infty}^{n+r}).$$

For all $q \in \mathbb{Z}$, we define a one-sided map denoted F_r^+ by

$$F_r^+ : A^{Z_{-\infty}^{q+r}} \to A^{Z_{-\infty}^q} \quad \text{with } (F_r^+ x_{-\infty}^{q+r})(n) = f_r^+(x_{-\infty}^{n+r}) \quad \text{for } n \leq q. \tag{5}$$

This definition strongly depends on the fixed anticipation r. Note that the notation F_r^+ does not mention explicitly the dependence on q, which is through the domain of definition and the range $A^{Z_{-\infty}^{q+r}}$ and $A^{Z_{-\infty}^{q}}$, respectively.

An inductive argument implies that, for $i \geq 1$, the iterated transformation $F^i : \Lambda \to \Lambda$ is also *f.a.* having ir as an anticipation (which may not be a minimal anticipation for F^i). As before we can define

$$f_{ir}^{(i)+} : A^{Z_{-\infty}^{ir}} \to A \quad \text{by } f_{ir}^{(i)+}(x_{-\infty}^{ir}) = (F^i x)(0).$$

We have the recursive relation

$$f_{ir}^{(i)+}(x_{-\infty}^{ir}) = f_r^+(f^{(i-1)+}(x_{-\infty}^{q+(i-1)r}) : q \leq r). \tag{6}$$

The associated one-sided map (similar to (5)), denoted by F_{ir}^{i+}, is given by

$$F_{ir}^{i+} : A^{Z_{-\infty}^{q+ir}} \to A^{Z_{-\infty}^{q}} \quad \text{with } (F_{ir}^{i+} x_{-\infty}^{q+ir})(n) = f_{ir}^{(i)+}(x_{-\infty}^{n+ir}) \text{ for } n \leq q. \tag{7}$$

Relation (6) can be written as

$$F_{ir}^{i+}(x_{-\infty}^{m+ir-1}) = F_r^+(F_{(i-1)r}^{(i-1)+}(x_{-\infty}^{m+ir-1})). \tag{8}$$

Similarly, we say F is *finite memory* (*f.m.*) if there exists $l \in \mathbb{Z}$ ($-l$ is called a memory) such that, for all $n \in \mathbb{Z}$, the coordinate $Fx(n)$ only depends on x_{n+l}^{∞}. A similar observation to Remark (1.1) can be stated. In the previous definitions the anticipation and memory, r and $-l$ respectively, have intuitive meaning when $l \leq 0$ and $r \geq 0$, but we do not impose these conditions.

We will only focus on *f.a.* transformations because all the notions and results we shall obtain for them can be also rephrased for *f.m.* transformations.

We will choose the minimal anticipation r for the transformation F (that is r is an anticipation and $r-1$ is not). The anticipation of F^i will be fixed as ir. Hence, in all the previous notions we will not mention explicitly the dependence on the anticipation, so we write f^+, F^+ instead of f_r^+, F_r^+ and $f^{(i)+}$, F^{i+} instead of $f_{ir}^{(i)+}$, F_{ir}^{i+}.

Remark 1.2. Let F be an *f.a.* transformation. One could ask why, if we have fixed the minimal anticipation r for F, we do not fix the minimal anticipation for F^i instead of taking it as ir. This question is at the core of our work. In fact, the anticipation of F^i could be strictly smaller than ir for some i. More precisely, when μ is an ergodic σ-invariant probability measure having lower bounded transition probabilities, in our main result, which is Theorem 3.1, we show that if the entropy of the shift with respect to $\mu \circ F^{-i}$ goes to zero, then for almost all points the anticipation of F^i is smaller than ir for i big enough (but this is not necessarily fulfilled in a uniform way).

Let F be a cellular automaton, so there exists integers $l \leq r$ defining a local rule $f : A^{Z_l^r} \to A$. If we fix $-l$ and r as the minimal ones, F is *f.a.* and *f.m.* with anticipation r and memory $-l$. For $q \in \mathbb{Z}$, the mapping $F^+ : A^{Z_{-\infty}^{q}} \to A^{Z_{-\infty}^{q-r}}$ satisfies $(F^+ x_{-\infty}^{q})(n) = f(x_{n+l}^{n+r})$ for all $n \leq q-r$.

Let $p, q \in \mathbb{Z}$ with $p \leq q$ and such that $q - p \geq r - l$, then we can define

$$f(x_p^q) = (f(x_p^{p+r-l}), \ldots, f(x_{q-(r-l)}^{q})) \in A^{Z_{p-l}^{q-r}}.$$

Let $i \geq 1$. The iterated $F^i : \Lambda \to \Lambda$ is also a *c.a.* with local rule $f^{(i)} : A^{i(r-l)+1} \to A$. Here, ir and $-il$ are the anticipation and memory fixed for F^i. These rules satisfy the recursive relation

$$f^{(1)} = f \quad \text{and} \quad f^{(i)}(a_1^{i(r-l)+1}) = f^{(i-1)}(b_1^{(i-1)(r-l)+1}) \quad \text{for } i \geq 1, \text{ where}$$
$$b(j(r-l)+k) = f(a_{(j-1)(r-l)+k}^{j(r-l)+k}), \quad k = 1, \ldots, r-l;\ j = 0, \ldots, i-1.$$

1.2.1. *Description of μ-a.s. continuous transformations by trees.* Let μ be a σ-invariant probability measure on $(A^{\mathbb{Z}}, \mathcal{B})$. We will characterize μ-a.s. continuous *f.a.* transformations by means of trees. Only to have simple notation, we assume that μ has complete support, that is it satisfies $\mu\{x \in A^{\mathbb{Z}} : x_l = i_l, l \in J\} > 0$ for all finite subsets $J \subset \mathbb{Z}$.

For convenience, the nodes of the trees will be indexed by finite sequences of negative numbers. Denote $-\mathbb{N} = Z_{-\infty}^{-1}$. Let μ^- be the induced σ^{-1}-invariant probability measure on $A^{-\mathbb{N}}$ that satisfies $\mu^-\{x \in A^{-\mathbb{N}} : x_l = i_l, l \in J\} = \mu\{x \in A^{\mathbb{Z}} : x_l = i_l, l \in J\}$ for every finite subset $J \subset -\mathbb{N}$.

For $k \geq 1$ the product space $A^{\{-k,\ldots,-1\}}$ is simply denoted by A^{-k}. Let $\vec{v}_{-k} = (v_{-k}, \ldots, v_{-1}) \in A^{-k}$. The set A^0 is defined as the singleton containing the empty word $\vec{v}_0 = \emptyset$. For $\vec{v}_{-k} \in A^{-k}$ and $a \in A$, we set $a\vec{v}_{-k} = (a, v_{-k}, \ldots, v_{-1}) \in A^{-(k+1)}$. For $\mathbf{v} = (v_n : n \leq -1) \in A^{-\mathbb{N}}$ and $k \leq -1$, we put $\vec{v}_{-k} = (v_{-k}, \ldots, v_{-1}) \in A^{-k}$.

Let $T \subset \bigcup_{k \leq 0} A^{-k}$ be a rooted tree graph with root $\emptyset$. We assume that T satisfies

$$\text{for all } \vec{v}_{-k} \in T \quad \text{either } [\forall a \in A a\vec{v}_{-k} \in T] \quad \text{or} \quad [\forall a \in A a\vec{v}_{-k} \notin T]. \tag{9}$$

When the second case occurs $\vec{v}_{-k} \in T$ is called a leaf. We denote by $\mathcal{L}$ the set of all leaves. We assume that $\vec{v}_0 = \emptyset$ is not a leaf:

$$\vec{v}_0 \notin \mathcal{L}. \tag{10}$$

Note that we can have sequences $(v_n : n \leq -1) \in A^{-\mathbb{N}}$ for which no $\vec{v}_{-k} = (v_{-k}, \ldots, v_{-1}) \in A^{-k}$ is a leaf. We will assume that

$$\mu^-\{\mathbf{v} = (v_n : n \leq -1) \in A^{-\mathbb{N}} : \exists \ell(\mathbf{v}) \leq -1 \text{ such that } \vec{v}_{\ell(\mathbf{v})} \in \mathcal{L}\} = 1. \tag{11}$$

From (9), the integer $\ell(\mathbf{v})$ is necessarily unique. Let $\vec{v}_{-k} \in T$. We denote by $\text{Succ}^*(\vec{v}_{-k}) = \{\vec{v}_{-(k+j)} \in T : j \geq 0\}$ the set of nodes of T hanging from $\vec{v}_{-k}$. The above assumptions, (11) and μ with complete support, imply that $\mathcal{L} \cap \text{Succ}^*(\vec{v}_{-k})$ is a singleton if and only if $\vec{v}_{-k} \in \mathcal{L}$.

Let $f_{\mathcal{L}} : \mathcal{L} \to A$ be a function satisfying the following condition: for all $\vec{v}_{-k} \in T$ the restriction $f_{\mathcal{L}} : \mathcal{L} \cap \text{Succ}^*(\vec{v}_{-k}) \to A$ is constant if and only if $\vec{v}_{-k} \in \mathcal{L}$ (that is when $\mathcal{L} \cap \text{Succ}^*(\vec{v}_{-k}) = \{\vec{v}_{-k}\}$). This is a minimality condition on the tree for representing $f_{\mathcal{L}}$, because when the restriction $f_{\mathcal{L}} : \mathcal{L} \cap \text{Succ}^*(\vec{v}_{-k}) \to A$ is constant then all the nodes of the tree hanging from $\vec{v}_{-k}$ can be collapsed into a unique leaf.

Define

$$\Lambda_0^- = \{y \in \Lambda : \exists k \geq 1, (y_{-k}, \ldots, y_{-1}) \in \mathcal{L}\}.$$

By (11) we have $\mu(\Lambda_0^-) = 1$. Then, the shift-invariant set $\Lambda_0 = \bigcap_{n \in \mathbb{Z}} \sigma^j \Lambda_0^-$ satisfies $\mu(\Lambda_0) = 1$.

Let r be a fixed number. We can define $F_{\mathcal{L}} : \Lambda_0 \to \Lambda$ by:

$$\text{for all } m \in \mathbb{Z}, \quad F_{\mathcal{L}} x(m) = f_{\mathcal{L}}(\vec{v}_{\ell(\mathbf{v})}) \quad \text{where}$$
$$\mathbf{v} = (v_n : n \leq -1) \in A^{-\mathbb{N}} \quad \text{is such that } v_n = x(m + r + 1 + n) \quad \text{for all } n \leq -1.$$

Let us prove that $F_{\mathcal{L}} : \Lambda_0 \to \Lambda$ is continuous. Let $k \geq 0$ be fixed and $x \in \Lambda_0$. We have $F_{\mathcal{L}} x(j) = f_{\mathcal{L}}(\vec{v}_{\ell(\mathbf{v}^j)})$, where $\mathbf{v}^j = (v_n^j : n \leq -1) \in A^{-\mathbb{N}}$ is such that $v_n^j = x(j + r + 1 + n)$ for $n \leq -1$ and $|j| \leq k$. Take

$$r_x(k) = k + r \quad \text{and} \quad l_x(k) = \min\{j + r + \ell(\mathbf{v}^j) : |j| \leq k\}.$$

By definition, when $y \in \Lambda_0$ satisfies $y(j) = x(j)$, for all $j \in \{l_x(k), \ldots, r_x(k)\}$, we get $F_{\mathcal{L}} x(j) = F_{\mathcal{L}} y(j)$, for all $|j| \leq k$. From (1) the continuity of $F_{\mathcal{L}} : \Lambda_0 \to \Lambda$ follows.

Since Λ_0 is shift-invariant, we can extend $F_{\mathcal{L}}$ to Λ by simply setting $F_{\mathcal{L}}(x) = y^*$, for all $x \in \Lambda \setminus \Lambda_0$, where $y^* \in \Lambda$ is a fixed element. This extension $F_{\mathcal{L}}$ is shift-commuting and μ-a.s. continuous, and so also measurable. From definition $F_{\mathcal{L}}$ is *f.a.* and the minimality condition of the tree implies that r is the minimal anticipation of F.

It can be shown that every *f.a.* measurable shift-commuting *f.a.* transformation $F : \Lambda \to \Lambda$ is equal, μ-a.s., to some $F_{\mathcal{L}} : \Lambda \to \Lambda$. We note that the assumption (10), $\vec{v}_0 = \emptyset$ is not a leaf, is consistent with the fact that the transformation F is not constant.

When μ does not have complete support, in the above construction we must only take care that all the branches $\mathbf{v} = (v_n : n \leq -1) \in A^{-\mathbb{N}}$ containing a cylinder whose μ measure vanishes, do not have leaves.

1.3. *Additional facts on measurability.* Let us make further considerations on measurability. As introduced, μ is a σ-invariant probability measure on $(\Lambda, \mathcal{B})$ and $\mathcal{B}$ is the μ-completed Borel $\boldsymbol{\sigma}$-field. As usual, all the sub-$\boldsymbol{\sigma}$-fields of $\mathcal{B}$ are assumed to contain all the μ-negligible sets.

Every sub-$\boldsymbol{\sigma}$-field $\mathcal{C}$ induces a measurable partition $\widehat{\Lambda}_{\mathcal{C}}$ of Λ whose elements are called fibers and denoted by $\xi_{\mathcal{C}}$. The fiber that contains a point $x \in \Lambda$ is denoted by $\xi_{\mathcal{C}}(x)$. We denote by $\widehat{\mu}$ the measure induced by μ on $\widehat{\Lambda}_{\mathcal{C}}$. Let $\mathbb{E}_\mu(\cdot|\mathcal{C})(x)$ be the mean expected value with respect to $\mathcal{C}$, which is defined μ-a.e. in x. We have that, $\widehat{\mu}$-a.e. in $\xi_{\mathcal{C}}$, there exist probability measures $\mu_{\xi_{\mathcal{C}}}$ supported on $\xi_{\mathcal{C}}$ such that, for every Borel set $B \subseteq \Lambda$, we have

$$E_\mu(1_B|\mathcal{C})(x) = \int_{\xi_{\mathcal{C}}(x)} 1_B(z)\, d\mu_{\xi_{\mathcal{C}}(x)}(z) \mu\text{-a.e.} \quad \text{in } x \in \Lambda. \tag{12}$$

See [**10**] and [**7**, §3.5]. The measure $\widehat{\mu}$ will be simply denoted by μ.

Let $\mathcal{Y}$ be a family of functions with domain Λ and with range in some complete separable metric space. We denote by $\boldsymbol{\sigma}(\mathcal{Y})$ the $\boldsymbol{\sigma}$-field generated by $\mathcal{Y}$, that is the smallest $\boldsymbol{\sigma}$-field on Λ such that all $\mathcal{Y}$ are measurable.

We define $X(n) : \Lambda \to A$, $x \to X(n)(x) = x(n)$ the nth coordinate function and, for $q \leq n$, we put $X_q^n = (X(m) : q \leq m \leq n)$. Also set $X_{-\infty}^n = (X(m) : m \leq n)$.

Let $F : \Lambda \to \Lambda$ be a measurable shift-commuting transformation. We define $(FX)_q^n = ((F \circ X)(m) : q \leq m \leq n)$ and $(FX)_{-\infty}^n = ((F \circ X)(m) : m \leq n)$. In these expressions $(F \circ X)(m)(x) = (Fx)(m)$ for $x \in \Lambda$. In particular, for the shift σ^j we have $(\sigma^j \circ X)(m)(x) = (\sigma^j x)(m) = x(m + j)$.

On the variables X we use all the notation already introduced for points x. For $q \le m$ we define the following sub-$\boldsymbol{\sigma}$-fields:

$$\mathbb{G}_q^m = \boldsymbol{\sigma}(X_q^m), \quad \mathbb{G}_{-\infty}^m = \boldsymbol{\sigma}(X_{-\infty}^m),$$
$$[F\mathbb{G}]_q^m = \boldsymbol{\sigma}((FX)_q^m), \quad [F\mathbb{G}]_{-\infty}^m = \boldsymbol{\sigma}((FX)_{-\infty}^m).$$

Note that

$$[F\mathbb{G}]_q^m = F^{-1}(\mathbb{G}_q^m), \quad [F\mathbb{G}]_{-\infty}^m = F^{-1}(\mathbb{G}_{-\infty}^m).$$

The fibers $\xi_{\mathcal{C}}(x)$ of these sub-$\boldsymbol{\sigma}$-fields $\mathcal{C}$, which contain the point $x \in \Lambda$, are identified with

$$\xi_{\mathcal{C}}(x) = \{z \in \Lambda : z_{-\infty}^m = x_{-\infty}^m\} \quad \text{when } \mathcal{C} = \mathbb{G}_{-\infty}^m,$$
$$\xi_{\mathcal{C}}(x) = \{z \in \Lambda : (Fz)_{-\infty}^m = (Fx)_{-\infty}^m\} \quad \text{when } \mathcal{C} = [F\mathbb{G}]_{-\infty}^m. \tag{13}$$

When F is *f.a.* with anticipation r we have the inclusion

$$\text{for all } m \in \mathbb{Z}, \quad [F\mathbb{G}]_{-\infty}^m \subseteq \mathbb{G}_{-\infty}^{m+r}.$$

For any sub-$\boldsymbol{\sigma}$-field $\mathcal{C}$ on Λ we put $\sigma^j(\mathcal{C}) = \{\sigma^j(B) : B \in \mathcal{C}\}$. By using that F is σ-commuting, we get

$$\sigma^j(\mathbb{G}_{-\infty}^{m+r}) = \mathbb{G}_{-\infty}^{m+r-j}, \quad \sigma^j([F\mathbb{G}]_{-\infty}^m) = [F\mathbb{G}]_{-\infty}^{m-j}. \tag{14}$$

Recall the notation $\mu(B|\mathbb{G}_{-\infty}^m) = \mathbb{E}_\mu(\mathbf{1}_B|\mathbb{G}_{-\infty}^m)$ for $B \in \mathcal{B}$. For $c \in A$ we put

$$\mu(c|x_{-\infty}^m) = \mu(X(m+1) = c|\mathbb{G}_{-\infty}^m)(x). \tag{15}$$

Then $\mu(x(m+1)|x_{-\infty}^m)$ is well defined by taking $c = x(m+1)$ in (15).

For every measurable bounded function $g : \Lambda \to \mathbb{R}$ and every sub-$\boldsymbol{\sigma}$-field $\mathcal{C}$ the following holds:

$$\mathbb{E}_\mu(g \circ F|F^{-1}(\mathcal{C}))(x) = \mathbb{E}_{\mu \circ F^{-1}}(g|\mathcal{C})(Fx)\mu\text{-a.e.} \quad \text{in } x \in \Lambda.$$

By using that $[F\mathbb{G}]_{-\infty}^{-1} = F^{-1}(\mathbb{G}_{-\infty}^{-1})$ we deduce

$$\mathbb{E}_{\mu \circ F^{-1}}(g|\mathbb{G}_{-\infty}^{-1}) \circ F = \mathbb{E}_\mu(g \circ F|[F\mathbb{G}]_{-\infty}^{-1})\mu\text{-a.e.}$$

When $y = Fx$, we have

$$\mu \circ F^{-1}(X(m) = y(m+1)|\mathbb{G}_{-\infty}^{m-1})(y) = \mu((FX)(m) = y(m)|[F\mathbb{G}]_{-\infty}^{m-1})(x). \tag{16}$$

1.4. *Measurable dynamics.* Let μ be a σ-invariant probability measure on $(\Lambda, \mathcal{B})$, so it satisfies $\mu = \mu \circ \sigma^{-1}$. The entropy of σ with respect to μ is

$$h_\mu(\sigma) = -\lim_{n\to\infty} \frac{1}{n} \sum_{x_0^{n-1} \in A^n} \mu(X_0^{n-1} = x_0^{n-1}) \cdot \log \mu(X_0^{n-1} = x_0^{n-1}).$$

The measure μ is ergodic if every σ-invariant function $g \in L^1(\mu)$ (so $g = g \circ \sigma$ μ-a.e.) is necessarily constant μ-a.e. Since F is measurable shift-commuting, we get that $\mu \circ F^{-1}$ is σ-invariant because

$$\mu \circ F^{-1} \circ \sigma^{-1} = \mu \circ (\sigma \circ F)^{-1} = \mu \circ (F \circ \sigma)^{-1} = \mu \circ \sigma^{-1} \circ F^{-1} = \mu \circ F^{-1}.$$

When μ is ergodic then $\mu \circ F^{-1}$ is also ergodic.

Let $i \geq 1$. Since $F^i : \Lambda \to \Lambda$ is also measurable shift-commuting, then $\mu \circ F^{-i}$ is σ-invariant and if μ is ergodic then $\mu \circ F^{-i}$ is also ergodic. Then F^i is a measurable factor from (Λ, μ, σ) on $(\Lambda, \mu \circ F^{-i}, \sigma)$. Moreover, for all $j \geq 1$, F^j is a measurable factor from $(\Lambda, \mu \circ F^{-i}, \sigma)$ on $(\Lambda, \mu \circ F^{-(i+j)}, \sigma)$. Then, the entropy decreases with the evolution of the automata,

$$\cdots h_{\mu \circ F^{-(i+1)}}(\sigma) \leq h_{\mu \circ F^{-i}}(\sigma) \leq \cdots \leq h_\mu(\sigma). \tag{17}$$

If F is a *c.a.*, the sets $F^i(\Lambda)$ are compact σ-invariant and they decrease with $i \geq 1$. The map $F : F^i(\Lambda) \to F^{i+1}(\Lambda)$ is onto and shift-commuting, so it is a topological factor between $(F^i(\Lambda), \sigma)$ and $(F^{i+1}(\Lambda), \sigma)$.

We shall be interested in understanding what kind of phenomenon occurs when $h_{\mu \circ F^{-i}}(\sigma)$ decreases to zero. To this purpose we will introduce some concepts. For a σ-invariant probability measure μ, define the lower bound of its transition probabilities given the whole past:

$$\chi_\mu^+ := \inf\{\mu(x(m+1) = a | x_{-\infty}^m) : a \in A, x \in \Lambda\}. \tag{18}$$

Since μ is σ-invariant, the above definitions do not depend on m. In our main result we assume $\chi_\mu^+ > 0$. In this case, when $\mu(X(m+1) \in D | x_{-\infty}^m)$ is arbitrary close to 1 then we have $D = A$. This last fact will be used in the proof of Theorem 3.1. The class of measures satisfying $\chi_\mu^+ > 0$ contains the Markov measures on finite states with strictly positive transition matrices and so the Bernoulli measures with strictly positive weights.

We note that if $\mu = \lambda$ is the uniform Bernoulli measure and F is surjective, then λ is F-invariant, so it is also F^i-invariant, $\lambda \circ F^{-i} = \lambda$ for $i \geq 1$, and our main result will not apply. For a deep study of entropy in the surjective case see [**4**].

2. *Right sets and events*

2.1. *Right sets.* Let F be an *f.a.* transformation. We denote by r the fixed anticipation for F, which is the minimal one. The right set of $x_{-\infty}^{r-1}$ and $c \in A$ is defined by

$$R^{f^+}(x_{-\infty}^{r-1}, c) = \{b \in A : f^+(x_{-\infty}^{r-1} b) = c\}, \tag{19}$$

which is an element in $\wp(A) = \{D : D \subseteq A\}$.

When $c = (Fx)(0)$, $R^{f^+}(x_{-\infty}^{r-1}, (Fx)(0))$ is the set of compatible right extensions of length 1, see [**3**] and [**4**]. We denote it by $R^{f^+}(x) = R^{f^+}(x_{-\infty}^{r-1}, (Fx)(0))$. Note that $R^{f^+}(x) \in \wp^*(A) = \{D : D \subseteq A, D \neq \emptyset\}$.

We have $R^{f^+}((\sigma^j x)_{-\infty}^{r-1}, c) = \{b \in A : f^+(x_{-\infty}^{j+r-1} b) = c\}$, and so

$$R^{f^+}(\sigma^j x) = \{b \in A : f^+(x_{-\infty}^{j+r-1} b) = (Fx)(j)\}. \tag{20}$$

Let $i \geq 1$. The transformation $F^i : \Lambda \to \Lambda$ is *f.a.* and has anticipation chosen to be ir. We have $R^{f^{(i)+}}(x_{-\infty}^{ir-1}, c) = \{b \in A : f^{(i)+}(x_{-\infty}^{ir-1} b) = c\}$ and we put $R^{f^{(i)+}}(x) = R^{f^{(i)+}}(x_{-\infty}^{ir-1}, (F^i x)(0))$. From (6) we get

$$f^{(i-1)+}(x_{-\infty}^{ir-1} a) = f^{(i-1)+}(x_{-\infty}^{ir}) \Rightarrow f^{i+}(x_{-\infty}^{ir-1} a) = f^{i+}(x_{-\infty}^{ir}). \tag{21}$$

From relations (20) and (21) we obtain the pointwise inclusion

$$\text{for all } i \geq 2, \quad R^{f^{(i-1)+}} \circ \sigma^r \subseteq R^{f^{(i)+}}.$$

Then,

$$\text{for all } i \geq 2, \quad R^{f^{(i-1)+}} \subseteq R^{f^{(i)+}} \circ \sigma^{-r}.$$

Hence, the following limit is increasing and defines an element of $\wp^*(A)$:

$$R^{\infty}(x) = \lim_{i\to\infty} R^{f^{(i)+}} \circ \sigma^{-ir}(x). \tag{22}$$

For each x, the limit is obviously attained after a finite number of steps.

For $c \in A$ the random variable $R^{f^+}(X_{-\infty}^{r-1}, c)$ is well defined, as well as $R^{f^+}(X)$, and they take values in $\wp(A)$. The function $R^{f^+}((\sigma^j \circ X)_{-\infty}^{r-1}, c)$ is $\mathbb{G}_{-\infty}^{j+r-1}$-measurable and $R^{f^+}(\sigma^j \circ X)$ is $\mathbb{G}_{-\infty}^{j+r-1} \vee \boldsymbol{\sigma}((FX)(j))$-measurable.

Note that from the definition we have

$$(R^{f^+}(x) = A) \Leftrightarrow (\forall z \in \Lambda \text{ with } z_{-\infty}^{r-1} = x_{-\infty}^{r-1} : (Fz)(0) = (Fx)(0)).$$

Let F be a *c.a.* given by a local function f having $-l, r$ as the minimal memory and anticipation values. Recall that $f^{(i)}$ denotes the local rule of the *c.a.* F^i.

LEMMA 2.1. *Let F be a c.a. Assume that, for some integer $i \geq 1$, $R^{f^{(i)}}(\sigma^{-ir}z) = A$. Then, for any $x_{-\infty}^{i(l-r)-1}$, the right sets satisfy $R^{f^{(i)}}(\sigma^{-ir}(x_{-\infty}^{i(l-r)-1} z_{i(l-r)}^{\infty})) = A$. Moreover, $\{R^{\infty} = A\}$ is an open set.*

Proof. The first statement follows immediately from the definition. Let us show $\{R^{\infty} = A\}$ is open. Take $z \in \Lambda$ with $R^{\infty}(z) = A$. Then there exists $\widehat{i} \geq 1$ such that $R^{f^{(i)}} \circ \sigma^{-ir}(z) = A$ for all $i \geq \widehat{i}$. Note that, for all $x \in \Lambda$ such that $x_{l-i(r-l)}^{r} = z_{l-i(r-l)}^{r}$, we have

$$\text{for all } j \leq i: \quad R^{f^{(j)}} \circ \sigma^{-jr}(x) = R^{f^{(j)}} \circ \sigma^{-jr}(z).$$

Then $V(z) = \{x : x_{l-\widehat{i}(r-l)}^{r} = z_{l-\widehat{i}(r-l)}^{r}\}$ is a open neighborhood of z such that $R^{f^{(\widehat{i})}} \circ \sigma^{-\widehat{i}r}(x) = A$ for all $x \in V(z)$. □

2.2. *Right events.* We define the right event at coordinate m by

$$\begin{aligned}\mathcal{R}_m^F(x_{-\infty}^{m+r-1}, c) &= \{z \in \Lambda : z_{-\infty}^{m+r-1} = x_{-\infty}^{m+r-1}, z_m \in R^{f^+}(x_{-\infty}^{m+r-1}, c)\} \\ &= \{z \in \Lambda : z_{-\infty}^{m+r-1} = x_{-\infty}^{m+r-1}, (Fz)(m) = c\}.\end{aligned} \tag{23}$$

So $\mathcal{R}_m^F(x_{-\infty}^{m+r-1}, c)$ is an event belonging to $\mathcal{B}$. If $c = (Fx)(m)$ we write

$$\mathcal{R}_m^F(x) = \mathcal{R}_m^F(x_{-\infty}^{m+r-1}, (Fx)(m)).$$

For every $x_{-\infty}^{m+r-1} \in A_{-\infty}^{m+r-1}$ and $c \in A$, and for all $x \in \Lambda$, we have that $\{X \in \mathcal{R}_m^F(x_{-\infty}^{m+r-1}, c)\}$ is $\mathbb{G}_{-\infty}^{m+r-1}$-measurable and $\{X \in \mathcal{R}_m^F(x)\}$ is $\mathbb{G}_{-\infty}^{m+r-1} \vee \boldsymbol{\sigma}((FX)(m))$-measurable. Note that

$$\{X \in \mathcal{R}_m^F(x)\} = \{\mathcal{R}_m^F(X) = \mathcal{R}_m^F(x)\} \subseteq \{(FX)(m) = (Fx)(m)\}. \tag{24}$$

The function $\Lambda \to \mathcal{B}$, $x \to \mathcal{R}_m^F(x)$, satisfies

$$\text{for all } j \in \mathbb{Z}: \quad \mathcal{R}_m^F \circ \sigma^j = \mathcal{R}_{m+j}^F. \tag{25}$$

Let $i \geq 1$. The transformation $F^i : \Lambda \to \Lambda$ has anticipation ir. We have

$$\mathcal{R}_m^{F^i}(x_{-\infty}^{m+r-1}, c) = \{z \in \Lambda : z_{-\infty}^{m+ir-1} = x_{-\infty}^{m+ir-1}, (F^i z)(m) = c\},$$

and when $c = (F^i x)(m)$ we put

$$\mathcal{R}_m^{F^i}(x) = \mathcal{R}_m^{F}(x_{-\infty}^{m+ir-1}, (F^i x)(m)).$$

Let μ be a σ-invariant probability measure. From (25) we have, for all $m, q \in \mathbb{Z}$ and all $B \in \mathcal{B}$,

$$\mu(\mathcal{R}_m^{F^i}(X) = B) = \mu(\mathcal{R}_q^{F^i}(X) = B).$$

Let us state a relation between right sets and right events.

LEMMA 2.2. *Let* $\mathcal{C} = \mathbb{G}_{-\infty}^{m+ir-1}$. *Then*

$$\mathcal{R}_m^{F^i}(x) = \{z \in \xi_{\mathcal{C}}(x) : z(m+ir) \in R^{f^{(i)+}}(x)\}. \tag{26}$$

Proof. From (13) we have $\xi_{\mathcal{C}}(x) = \{z \in \Lambda : z_{-\infty}^{m+ir-1} = x_{-\infty}^{m+ir-1}\}$. Then,

$$\begin{aligned}
\mathcal{R}_m^{F^i}(x) &= \mathcal{R}_m^{F^i}(x_{-\infty}^{m+ir-1}, (F^i x)(m)) \\
&= \{z \in \xi_{\mathcal{C}}(x) : f^{(i)+}(z_{-\infty}^{m+ir}) = (F^i x)(m)\} \\
&= \{z \in \xi_{\mathcal{C}}(x) : z(m+ir) \in R^{f^{(i)+}}(x_{-\infty}^{m+ir-1}, (F^i x)(m))\} \\
&= \{z \in \xi_{\mathcal{C}}(x) : z(m+ir) \in R^{f^{(i)+}}(x)\},
\end{aligned}$$

and the result follows. □

In a similar way as done with right sets and events, we could define the left sets and left events for *f.m.* transformations, and we should obtain similar relations.

3. *Main result*

Let us state and prove our main result.

THEOREM 3.1. *Let* μ *be an ergodic* σ*-invariant probability measure. Let* $F : \Lambda \to \Lambda$ *be an f.a. measurable shift-invariant transformation and* $\chi_\mu^+ > 0$. *Then*

$$\lim_{i\to\infty} h_{\mu\circ F^{-i}}(\sigma) = 0 \quad \textit{implies} \quad R^\infty = A \mu\text{-a.e.}$$

Proof. Let r be the minimal anticipation of F. Let $c \in A$ and $\mathcal{C} = \mathbb{G}_{-\infty}^{r+m-1}$. From (12) and (24) we have

$$\begin{aligned}
\mu((FX)(m) = c|\mathcal{C})(z) &= \int_{\xi_{\mathcal{C}}(z)} \mathbf{1}_{\{(Fu)(m)=c\}}\, d\mu_{\xi_{\mathcal{C}}(z)}(u) \\
&= \int_{\{u\in\Lambda:u_{-\infty}^{m+r-1}=z_{-\infty}^{m+r-1},(Fu)(m)=c\}} d\mu_{\xi_{\mathcal{C}}(z)}(u) = \int_{\mathcal{R}_m^F(z,c)} d\mu_{\xi_{\mathcal{C}}(z)}(u) \\
&= \mu_{\xi_{\mathcal{C}}(z)}(\mathcal{R}_m^F(z,c)) = \mu(\mathcal{R}_m^F(z,c)|\mathcal{C})(z).
\end{aligned} \tag{27}$$

Then, by using (13) and (16), we get, for $y = Fx$,

$$\begin{aligned}
&\mu \circ F^{-1}(X(m) = y(m)|y_{-\infty}^{m-1}) \\
&\quad = \mu \circ F^{-1}(Y(m) = y(m)|\mathbb{G}_{-\infty}^{m-1})(y) \\
&\quad = \mu((FX)(m) = (Fx)(m)|[F\mathbb{G}]_{-\infty}^{m-1})(x) \\
&\quad = \mu(\mu((FX)(m) = (Fx)(m)|\mathbb{G}_{-\infty}^{m+r-1})|[F\mathbb{G}]_{-\infty}^{m-1})(x) \\
&\quad = \mu(\mathcal{R}_m^F(X_{-\infty}^{r+m-1}, (Fx)(m))|[F\mathbb{G}]_{-\infty}^{m-1})(x).
\end{aligned} \tag{28}$$

It is known, for instance see [**8**, §4.3], that every ergodic σ-invariant probability measure ρ satisfies

$$h_\rho(\sigma) = -\int_\Lambda \log \rho(X(m) = x(m)|x_{-\infty}^{m-1})\, d\rho(x).$$

From the hypothesis we get that $\mu \circ F^{-1}$ is an ergodic σ-invariant probability measure. Then, from (28) we get, for all $m \in \mathbb{Z}$,

$$\begin{aligned}
h_{\mu\circ F^{-1}}(\sigma) &= -\int_\Lambda \log \mu \circ F^{-1}(X(m) = y(m)|y_{-\infty}^{m-1})\, d\mu \circ F^{-1}(y) \\
&= -\int_\Lambda \log \mu(\mathcal{R}_m^F(X_{-\infty}^{m+r-1}, (Fx)(m))|[F\mathbb{G}]_{-\infty}^{m-1})(x)\, d\mu(x).
\end{aligned}$$

Then

$$h_{\mu\circ F^{-i}}(\sigma) = -\int_\Lambda \log \mu(\mathcal{R}_m^{F^i}(X_{-\infty}^{m+ir-1}, (F^i x)(m))|[F^i\mathbb{G}]_{-\infty}^{m-1})(x)\, d\mu(x),$$

and so

$$\begin{aligned}
&\lim_{i\to\infty} h_{\mu\circ F^{-i}}(\sigma) \\
&\quad = -\lim_{i\to\infty} \int_\Lambda \log \mu(\mathcal{R}_m^{F^i}(X_{-\infty}^{m+ir-1}, (F^i x)(m))|[F^i\mathbb{G}]_{-\infty}^{m-1})(x)\, d\mu(x).
\end{aligned} \tag{29}$$

Assume now $\lim_{i\to\infty} h_{\mu\circ F^{-i}}(\sigma) = 0$. By (29), this implies

$$\lim_{i\to\infty} \int_\Lambda \log \mu(\mathcal{R}_m^{F^i}(X_{-\infty}^{m+ir-1}, (F^i x)(m))|[F^i\mathbb{G}]_{-\infty}^{m-1})(x)\, d\mu(x) = 0.$$

By Jensen's inequality we get

$$\liminf_{i\to\infty} \log \int_\Lambda \mu(\mathcal{R}_m^{F^i}(X_{-\infty}^{m+ir-1}, (F^i x)(m))|[F^i\mathbb{G}]_{-\infty}^{m-1})(x)\, d\mu(x) \geq 0.$$

Then

$$\liminf_{i\to\infty} \int_\Lambda \mu(\mathcal{R}_m^{F^i}(X_{-\infty}^{m+ir-1}, (F^i x)(m))|[F^i\mathbb{G}]_{-\infty}^{m-1})(x)\, d\mu(x) \geq 1. \tag{30}$$

Observe that

$$\begin{aligned}
&\liminf_{i\to\infty} \int_\Lambda \mu(\mathcal{R}_m^{F^i}(X_{-\infty}^{m+ir-1}, (F^i x)(m))|[F^i\mathbb{G}]_{-\infty}^{m-1})(x)\, d\mu(x) \\
&\quad = \liminf_{i\to\infty} \int_\Lambda \mu(\mathcal{R}_m^{F^i}(X_{-\infty}^{m+ir-1}, (F^i x)(m))|\mathbb{G}_{-\infty}^{m+ir-1})(x)\, d\mu(x).
\end{aligned}$$

Let $\mathcal{C} = \mathbb{G}_{-\infty}^{m+ir-1}$. From (13) and (26), $\xi_{\mathcal{C}}(x) = \{w \in \Lambda : z_{-\infty}^{m+ir-1} = x_{-\infty}^{m+ir-1}\}$ and $\mathcal{R}_m^{F^i}(x) = \{z \in \xi_{\mathcal{C}}(x) : z(m+ir) \in R^{f^{(i)+}}(x)\}$. Therefore,

$$\mu(\mathcal{R}_m^{F^i}(X_{-\infty}^{m+r-1}, (F^i x)(m)) | \mathbb{G}_{-\infty}^{m+ir-1})(x) = \mu_{\xi_{\mathcal{C}}(x)}(X(m+ir) \in R^{f^{(i)+}}(x)).$$

We have shown

$$\liminf_{i\to\infty} \int_\Lambda \mu(\mathcal{R}_m^{F^i}(X_{-\infty}^{m+r-1}, (F^i x)(m)) | [F^i \mathbb{G}]_{-\infty}^{m-1})(x)\, d\mu(x)$$
$$= \liminf_{i\to\infty} \int_\Lambda \mu_{\xi_{\mathcal{C}}(x)}(X(m+ir) \in R^{f^{(i)+}}(x))\, d\mu(x). \tag{31}$$

For $i \geq 1$, define

$$\Lambda_i = \{x \in \Lambda : \forall j \geq i\, R^{f^{(j)+}} \circ \sigma^{-jr}(x) = R^\infty(x)\}.$$

From (22) we have $\Lambda_i \nearrow \Lambda$, so for all $\epsilon > 0$ there exists $\tilde{i}(\epsilon)$ such that $\mu(\Lambda_{\tilde{i}(\epsilon)}) > 1 - \epsilon$. Since Λ_i is increasing with i,

$$\text{for all } i \geq \tilde{i}(\epsilon): \quad \mu(\Lambda_i) > 1 - \epsilon. \tag{32}$$

Let us denote

$$\Lambda^{\neq} = \{x \in \Lambda : R^\infty(x) \neq A\}.$$

Then

$$\Lambda^{\neq} \circ \sigma^j = \{x \in \Lambda : R^\infty \circ \sigma^{-j}(x) \neq A\}.$$

Assume that

$$\mu(\Lambda^{\neq}) > 0.$$

Take $\epsilon \in (0, \mu(\Lambda^{\neq})/2)$. From (32) and the shift invariance of μ we obtain, for all $i \geq \tilde{i}(\epsilon)$,

$$\mu(\Lambda_i \cap \sigma^j(\Lambda^{\neq})) \geq \mu(\Lambda_i) - \mu(\Lambda \backslash \sigma^j(\Lambda^{\neq}))$$
$$\geq 1 - \epsilon - (1 - \mu(\sigma^j(\Lambda^{\neq}))) = (\mu(\Lambda^{\neq}) - \epsilon). \tag{33}$$

On the other hand, χ_μ^+ defined in (18) can be written

$$\chi_\mu^+ = \inf\{\mu_{\xi_{\mathcal{C}}(x)}(X(q) = a) : a \in A, x \in \Lambda\}.$$

This does not depend on $q \in \mathbb{Z}$ because μ is σ-invariant. Then, the hypothesis $\chi_\mu^+ > 0$ is equivalent to

$$1 - \chi_\mu^+ = \sup\{\mu_{\xi_{\mathcal{C}}(x)}(X(q) \in D) : D \in \wp(A) \backslash \{A\}, x \in \Lambda\} < 1. \tag{34}$$

(Note that $\wp(A)\backslash\{A\} = \{D \subseteq A : D \neq A\}$.)

We have

$$\liminf_{i\to\infty} \int_\Lambda \mu_{\xi_{\mathcal{C}}(x)}(X(m+ir) \in R^{f^{(i)+}}(x))\, d\mu(x)$$
$$= \liminf_{i\to\infty} \int_\Lambda \mu_{\xi_{\mathcal{C}}(\sigma^{-ir}x)}(X(m+ir) \in R^{f^{(i)+}}(\sigma^{-ir}(x))\, d\mu(x)$$
$$\leq \liminf_{i\to\infty} \Big[\mu(\Lambda \backslash (\Lambda_i \cap \sigma^{ir}(\Lambda^{\neq})))$$
$$+ \int_{\Lambda_i \cap \sigma^{ir}(\Lambda^{\neq}))} \mu_{\xi_{\mathcal{C}}(\sigma^{-ir}x)}(X(m+ir) \in R^{f^{(i)+}}(\sigma^{-ir}(x)))\, d\mu(x) \Big].$$

Since $R^{f^{(i)+}}(\sigma^{-ir}(x)) \in \wp(A)\backslash\{A\}$, for all $x \in \Lambda_i \cap \sigma^{ir}(\Lambda^{\neq})$, we can use the bound (34) to obtain

$$\liminf_{i\to\infty} \int_{\Lambda} \mu_{\xi_C(x)}(X(m+ir) \in R^{f^{(i)+}}(x))\, d\mu(x)$$
$$\leq \liminf_{i\to\infty}[(1-\mu(\Lambda_i \cap \sigma^{ir}(\Lambda^{\neq}))) + (1-\chi_\mu^+)\mu(\Lambda_i \cap \sigma^{ir}(\Lambda^{\neq}))].$$

We have $\mu(\Lambda_i \cap \sigma^{ir}(\Lambda^{\neq})) \leq \mu(\Lambda^{\neq})$ and, from (33), $\mu(\Lambda_i \cap \sigma^{ir}(\Lambda^{\neq})) > (\mu(\Lambda^{\neq}) - \epsilon)$. This implies

$$\liminf_{i\to\infty} \int_{\Lambda} \mu_{\xi_C(x)}(X(m+ir) \in R^{f^{(i)+}}(x))\, d\mu(x) \leq 1 - \chi_\mu^+ \mu(\Lambda^{\neq}) + \epsilon.$$

Take $0 < \epsilon < \chi_\mu^+\mu(\Lambda^{\neq})$ to have $1 - \chi_\mu^+\mu(\Lambda^{\neq}) + \epsilon < 1$. Hence, from (31) we obtain

$$\liminf_{i\to\infty} \int_{\Lambda} \mu(\mathcal{R}_m^{F^i}(X_{-\infty}^{m+r-1}, (F^i x)(m)) | [F^i \mathbb{G}]_{-\infty}^{m-1})(x)\, d\mu(x) < 1,$$

which contradicts (30), hence $\mu(\Lambda^{\neq}) = 0$, and the proof of Theorem 3.1 is finished. □

Remark 3.2. It is easy to see that Theorem 3.1 holds trivially if, instead of choosing the minimal anticipation r, we use a strictly bigger anticipation $r' > r$. In fact, when we take $f^+ = f_{r'}^+$ (see (3)), we find $R^{f^+} = A$.

Let us write the statement of the main result for a finite sequence of symbols. Let F be an *f.a.* transformation with minimal anticipation r. Let k be a fixed integer greater than or equal to 1. We define the k-right set of $x_{-\infty}^{r-k}$ and $c = (c_1, \ldots, c_k) \in A^k$ by

$$R^{f^+;k}(x_{-\infty}^{r-k}, c) = \{b \in A^k : f^+(x_{-\infty}^{r-k} b) = c\}.$$

This is an element in $\wp(A^k) = \{D : D \subseteq A^k\}$.

If $c = (Fx)_{-(k-1)}^0$, then $R^{f^+;k}(x_{-\infty}^{r-1}, (Fx)_{-(k-1)}^0)$ is the set of compatible right extensions of length k; we denote it by $R^{f^+;k}(x) = R^{f^+;k}(x_{-\infty}^{r-1}, (Fx)_{-(k-1)}^0)$. We have

$$\text{for all } i \geq 2, \quad R^{f^{(i-1)+};k} \subseteq R^{f^{(i)+};k} \circ \sigma^{-r},$$

and so the following limit is increasing and defines an element of $\wp^*(A^k)$:

$$R^{\infty;k}(x) = \lim_{i\to\infty} R^{f^{(i)+};k} \circ \sigma^{-ir}(x).$$

The following result is satisfied.

THEOREM 3.3. *Let μ be an ergodic σ-invariant probability measure. Let $F : \Lambda \to \Lambda$ be an f.a. measurable shift-invariant transformation and $\chi_\mu^+ > 0$. Then*

$$\lim_{i\to\infty} h_{\mu\circ F^{-i}}(\sigma) = 0 \quad \textit{implies} \quad R^{\infty;k} = A^k \, \mu\text{-a.e.}$$

The proof of this result can be done by following similar steps to those of the proof of Theorem 3.1. Also it can be done by applying Theorem 3.1 to an increasing sequence of $f^{(i_q)}$.

Analogous statements to those of Theorems 3.1 and 3.3 can be done for *f.m.* transformations.

4. *Right sets for some Wolfram rules*

We recall that these are cellular automata with $l = -1$ and $r = 1$ on the alphabet $\{0, 1\}$, see [**11**]. There are 256 different Wolfram rules $f : \{0, 1\}^3 \to \{0, 1\}$. Each rule is coded by some $k \in \{0, \ldots, 255\}$ whose dyadic decomposition is the 8-tuple $(f(111), f(110), f(101), f(100), f(011), f(010), f(001), f(000))$.

In this section we treat several examples. We first give a general result useful for the computation of right sets (a similar result holds of course for left sets). Recall from Lemma 2.1 that, if for some integer $p \geq 1$, $R^{f^{(p)}}(\sigma^{-pr}x) = A$, then, for any $y_{-\infty}^{(l-r)p-1}$, $R^{f^{(p)}}(\sigma^{-pr}(y_{-\infty}^{(l-r)p-1} x_{(l-r)p}^{\infty})) = A$. Also, the set $(R^{f^{(p)}})^{-1}(A)$ is a union of cylinder sets. Since the sequence of sets $(R^{f^{(p)}}(\sigma^{-pr}x))$ is non-decreasing (see §2), we conclude that $(R^{\infty})^{-1}(A) = \{x \in \Lambda : R^{\infty}(x) = A\}$ is also a union of cylinder sets. It can be conveniently described as a finite union of trees.

In the case of Wolfram rules, since the alphabet has cardinality two, it is enough to compute the right set of sequences x with $x_0 = 0$ (or $x_0 = 1$). In order to construct the corresponding tree, one can use the following algorithm.

The root of the tree is labeled 0. Each node is labeled by a string of the form $y_{-2p}^{-1}0$. A node is either pending or finished. If it is finished it has no descendant and the string corresponds to a cylinder set belonging to $(R^{\infty})^{-1}(A)$. If the node with string $y_{-2p}^{-1}0$ is pending it has four descendants:

$$00y_{-2p}^{-1}0, \quad 01y_{-2p}^{-1}0, \quad 10y_{-2p}^{-1}0, \quad 11y_{-2p}^{-1}0.$$

For each of the four descendants $aby_{-2p}^{-1}0$ ($a, b \in \{0, 1\}$), we compare $F^{p+1}(aby_{-2p}^{-1}0)$ and $F^{p+1}(aby_{-2p}^{-1}1)$. If they are equal, the node $aby_{-2p}^{-1}0$ is set to be finished. Otherwise, it is set to be pending.

In practice we explore the tree by considering the nodes of successive depths.

In view of our main result, we are mostly interested in cases where $\mu(R^{\infty} \neq A) > 0$, which implies $\lim_{i\to\infty} h_{\mu\circ F^{-i}}(\sigma) > 0$.

4.1. *Examples.* In this section we give some examples of different behaviors in increasing order of complexity. We also show some (partial) pictures of trees where terminated nodes are drawn as squares while pending nodes are drawn as disks.

4.1.1. *Rule 108.* Certain rules give a finite tree. In this case $R^{\infty}(x) = A$ for all $x \in \Lambda$. This is the case of rules 108, 201 and several other ones. The proof of finiteness of the tree for these rules is left to the reader.

4.1.2. *Rule 44.* At each level, rule 44 has three new finished nodes and one pending node. The sequence of strings of pending nodes is obtained as follows. The first one is 100. One then prepends successively the prefixes 01, then 11, then 10, repeating this sequence of prepending infinitely many times. It is left to the reader to prove that this is indeed the tree of rule 44.

It follows at once that if μ is a non-atomic ergodic σ-invariant probability measure, the right set is equal to A almost certainly. Figure 1 shows a piece of the tree for this rule.

Rules 32, 50, 62, 128, 131, 176, 179, 203, 224, 242, 248, 251 and 254 behave similarly.

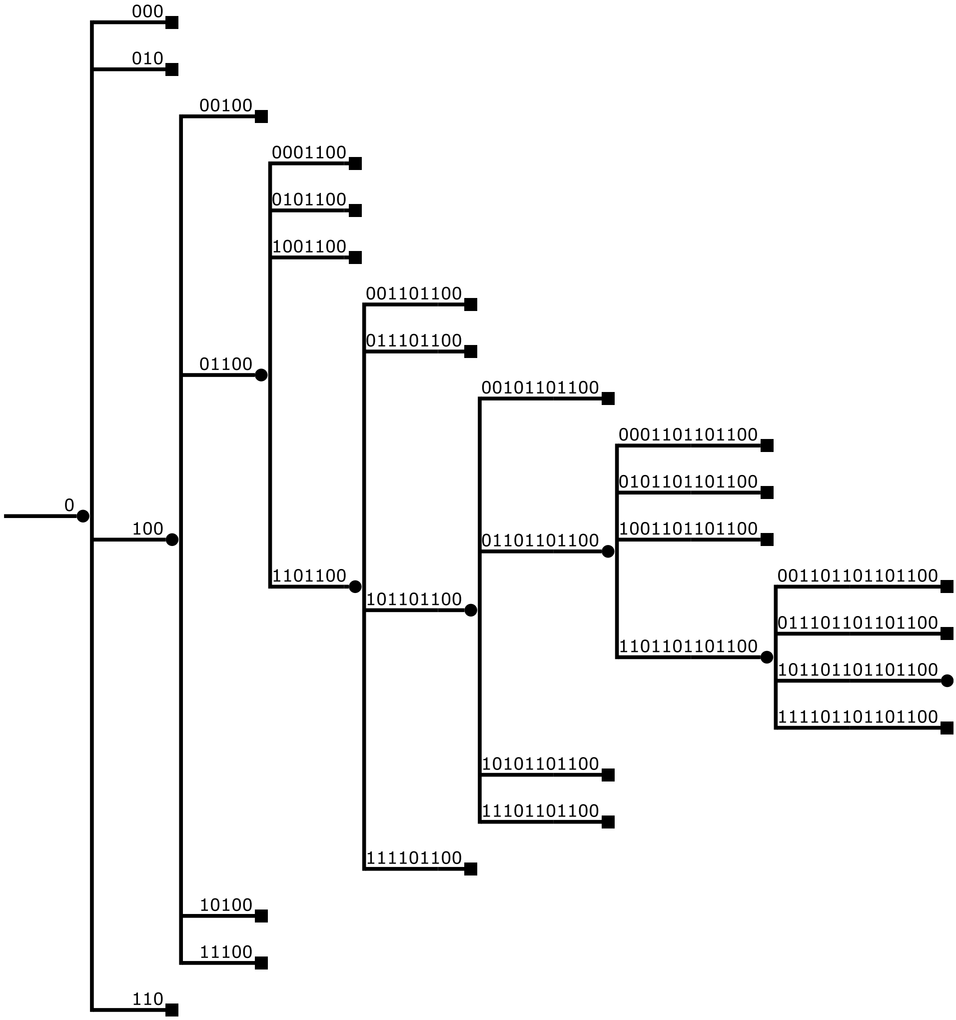

FIGURE 1. Part of the tree for rule 44.

4.1.3. *Rule 34.* It is easy to verify that if a string is of the form $y_{-2p}^{-2}00$, with $p \geq 2$, its image $F(y_{-2p}^{-2}00)$ is of the form $z_{-2p}^{-2}00$. This is because $f(a, b, 0) = 0$ for any $a, b \in \{0, 1\}$. Similarly, if a string is of the form $y_{-2p}^{-2}01$, with $p \geq 2$, its image $F(y_{-2p}^{-2}01)$ is of the form $z_{-2p}^{-2}01$. This is because $f(a, 0, 1) = 1$ for any $a \in \{0, 1\}$, and the previous observation. This implies that the strings terminating with 000 and 100 form complete subtrees of pending nodes. Therefore, the set $R^{\infty} \neq A$ has positive measure, and our main theorem implies that the limit of the sequences of entropies is strictly positive.

Figure 2 shows a piece of the tree for this rule.

4.1.4. *Right permutative rules.* Recall that a rule is said to be right permutative if, for any symbols a and b, the map $c \mapsto f(a, b, c)$ is a bijection of $\{0, 1\}$.

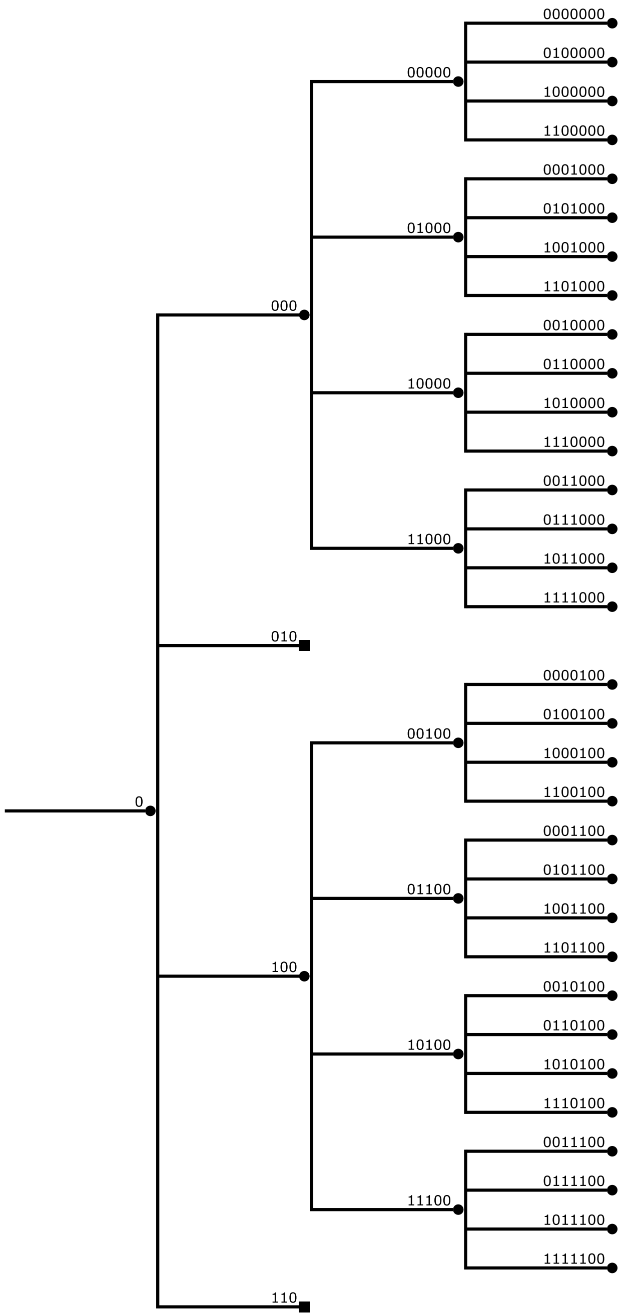

FIGURE 2. Part of the tree for rule 34.

It follows easily from this definition that, for right permutative rules, the right set of any sequence $x_{-\infty}^{0}$ is equal to x_0. It follows that the tree is the complete 4-ary tree.

One can check by direct computation that rules 85, 86, 89, 90, 101, 102, 105, 106, 149, 150, 153, 154, 165, 166, 169, 170 are the only right permutative rules.

Acknowledgements. The authors are indebted for the support of Program Basal CMM and MathAmsud from CONICYT. P. C. thanks hospitality of CMM University of Chile and S. M. thanks hospitality of Centre de Physique Théorique at École Polytechnique. We acknowledge an anonymous referee whose comments and suggestions allowed us to make precise the notion of anticipation given in §1.2, called our attention to the measurability of the transformation, and helped us to correct several mistakes and to improve the whole presentation of the paper.

REFERENCES

[**1**] J. L. Doob. *Measure Theory (Graduate Texts in Mathematics, 143)*. Springer, New York, 1994.

[**2**] E. Goles, A. Maass and S. Martínez. On the limit set of some universal cellular automata. *Theoret. Comput. Sci.* **110**(1) (1993), 53–78.

[**3**] G. A. Hedlund. Endomorphisms and automorphisms of the shift dynamical system. *Math. Syst. Theory* **3** (1969), 320–375.

[**4**] T. Helvik, K. Lindgren and M. G. Nordahl. Continuity of information transport in surjective cellular automata. *Comm. Math. Phys.* **272**(1) (2007), 53–74.

[**5**] L. Hurd. Formal language characterization of cellular automata limit sets. *Complex Systems* **1** (1987), 69–80.

[**6**] L. Hurd. Recursive cellular automata invariant sets. *Complex Systems* **4** (1990), 119–129.

[**7**] K. Itô. *Introduction to Probability Theory*. Cambridge University Press, Cambridge, 1984.

[**8**] S. Kalikow and R. McCutcheon. *An Outiline of Ergodic Theory (Cambridge Studies in Advanced Mathematics, 122)*. Cambridge University Press, Cambridge, 2010.

[**9**] A. Maass. On the sofic limit sets of cellular automata. *Ergod. Th. & Dynam. Sys.* **15** (1995), 663–684.

[**10**] V. A. Rohlin. On the fundamental ideas of measure theory. *Amer. Math. Soc. Transl.* **71** (1952), 1–54.

[**11**] S. Wolfram. Computation theory of cellular automata. *Comm. Math. Phys.* **96** (1984), 15–57.

[**12**] J. Zhisong. A complexity analysis of the elementary cellular automaton of rule 122. *Appl. Math. J. Chinese Univ. Ser. B* **20**(3) (2005), 268–276.