Limit measures for affine cellular automata on topological Markov subgroups

Alejandro Maass, Servet Martínez and Marcelo Sobottka

Centro de Modelamiento Matemático and Departamento de Ingeniería Matemática, Facultad de Ciencias Físicas y Matemáticas, Universidad de Chile, Casilla 170/3-Correo 3, Santiago, Chile

E-mail: amaass@dim.uchile.cl, smartine@dim.uchile.cl and sobottka@dim.uchile.cl

Abstract
Consider a topological Markov subgroup which is p^s-torsion (with p prime) and an affine cellular automaton defined on it. We show that the Cesàro mean of the iterates, by the automaton of a probability measure with complete connections and summable memory decay that is compatible with the topological Markov subgroup, converges to the Haar measure.

1. Introduction

Let $(G, +)$ be a finite Abelian group and $(G^{\mathbb{Z}}, +)$ be the product group with the componentwise addition. Let $\mathfrak{G} \subseteq G^{\mathbb{Z}}$ be a subgroup shift, which without loss of generality can be considered to use all elements of G. It is well known that $\mathfrak{G}$ can be seen as a topological Markov chain (see [4]).

One says $\Phi : \mathfrak{G} \to \mathfrak{G}$ is an *affine cellular automaton* if it is given by $\Phi = \mathbf{a} \cdot id + \mathbf{b} \cdot \sigma + \mathbf{c}$, where $\mathbf{a}, \mathbf{b} \in \mathbb{Z}$ are such that the maps $g \mapsto \mathbf{a} \cdot g$ and $g \mapsto \mathbf{b} \cdot g$ are both automorphisms of G, $\mathbf{c} \in \mathfrak{G}$ is a constant sequence and $\sigma : \mathfrak{G} \to \mathfrak{G}$ is the shift map.

Cellular automata have been used to model several physical and biological systems presenting self-organization. Consequently, the study of their dynamics has attracted increasing attention over the last three decades. One approach to studying the dynamic behaviour of a cellular automaton is to consider a shift-invariant probability measure μ of full support in $\mathfrak{G}$, which represents a realistic amount of initial conditions, and the sequence of iterated measures $(\mu \circ \Phi^{-n} : n \in \mathbb{N})$. Even in the simplest case the limit of $\mu \circ \Phi^{-n}$ as $n \to \infty$ does not exist, thus one studies the convergence of the Cesàro mean distribution $N^{-1} \sum_{n=0}^{N-1} \mu \circ \Phi^{-n}$ as $n \to \infty$ instead. In this spirit, Lind [5] considered the case $G = \mathbb{Z}_2$, $\mathfrak{G} = G^{\mathbb{Z}}$, $\Phi = \sigma^{-1} + \sigma$ and μ a Bernoulli measure, and proved that the Cesàro mean distribution always converges to the Haar measure (in this case the uniform Bernoulli measure). Later, in [1] and [7], the same result was shown for μ a Markov measure. In [2], using regeneration theory of stochastic processes with finite state space, it was proved that the Cesàro mean distribution also converges to the Haar measure when μ has complete connections and summable memory decay, $G = \mathbb{Z}_{p^s}$ for some p prime, and $\Phi = \mathbf{a} \cdot id + \mathbf{b} \cdot \sigma$. By using harmonic analysis ideas Pivato and Yassawi [10, 11] found the same result for the case where: $G = \bigoplus_{i=1}^{k} \mathbb{Z}_{m_i}$ for any collection of positive integers $\{m_i : 1 \leqslant i \leqslant k\}$, the initial measure is harmonically mixing and Φ is a diffusive in density cellular automaton. Indeed, this case includes affine cellular automata and measures with complete connections and summable memory decay (see [2, 3]). Notice that all previous results are stated for the case $\mathfrak{G} = G^{\mathbb{Z}}$.

Recently in [8] it was shown that if $(G, +)$ is p^s-torsion with p prime, $\mathfrak{G}$ is a Markov field verifying a 'filling property', $\Phi = id + \sigma$ and μ is a Markov measure, then the attractiveness property of the Haar measure also holds. In [9] the same result was proven in the one-dimensional case. In this case the Haar measure corresponds to the uniform Markov measure on the allowed words.

In [4] Kitchens showed that any irreducible subgroup shift $(\mathfrak{G}, +)$ is isomorphic and topologically conjugated to a full shift group $(A^{\mathbb{Z}}, *)$, where $*$ is not necessarily a 1-block operation. Let $\pi : \mathfrak{G} \to A^{\mathbb{Z}}$ be this isomorphism. Hence, from the proof of Kitchens' representation theorem (theorem 1(ii) in [4]) one can get that if μ is a probability measure with complete connections and summable memory decay on $\mathfrak{G}$, then it is projected by π onto the probability measure $\mu \circ \pi^{-1}$ on $G^{\mathbb{Z}}$ having these same properties. Moreover, when $\Phi = \mathbf{a} \cdot id_G + \mathbf{b} \cdot \sigma_G + \mathbf{c}$ is an affine cellular automaton on $\mathfrak{G}$, then it is topologically conjugated to $\varphi = \mathbf{a} \cdot id_{A^{\mathbb{Z}}} * \mathbf{b} \cdot \sigma_{A^{\mathbb{Z}}} * \pi(\mathbf{c})$ through π. Therefore, whenever $*$ is a 1-block operation (e.g. if the topological entropy of $(\mathfrak{G}, \sigma_G)$ is $\log N$ with $N = p_1 \cdot p_2 \cdots p_q$, $p_1, \ldots, p_q$ distinct prime numbers) one can conclude the Cesàro mean convergence directly from [3]. However, this occurs only if $*$ has some type of regularity, and thus for most of affine cellular automata this method cannot be used (see also remark 1 below).

In this work we assume $(G, +)$ is p^s-torsion with p prime and we prove the convergence of the Cesàro mean distribution to the Haar measure when $\mathfrak{G} \subseteq G^{\mathbb{Z}}$ is any subgroup shift, μ is a probability measure with complete connections compatible with $\mathfrak{G}$ and summable memory decay and Φ is any affine cellular automaton on $\mathfrak{G}$. The elements of our proof share techniques with [2] and [9]. More precisely, the regenerative construction of the measure and the combinatorics of the binomial coefficients leading the dynamics.

In section 2 we develop the background and state the main result (theorem 1). In section 3, for a fixed left-infinite sequence $w = (w_{-1}, w_{-2}, \ldots)$ allowed in $\mathfrak{G}$, we construct a random sequence $\mathbf{x} = (x_n : n \in \mathbb{N})$ in G distributed as the conditional probability measure μ_w and present the renewal process associated with $\mathbf{x}$. Finally, in section 4 we prove that for sufficiently large n the finite-dimensional distribution of $\Phi^n(\mathbf{x})$, conditioned to some renewal times, coincides with the finite-dimensional law of the Haar measure. This implies theorem 1.

2. Background and main result

The Abelian group $(G, +)$ is p^s-torsion with p prime and $s \geqslant 1$ if: (i) $p^s g = 0$ for any $g \in G$; (ii) given $1 < m < p^s$ there exists $g \in G$ such that $mg \neq g$.

Denote by $\sigma : \mathfrak{G} \to \mathfrak{G}$ the shift map $(\sigma(\mathbf{g}))_n = g_{n+1}$ where $\mathbf{g} \in \mathfrak{G}$ and $n \in \mathbb{Z}$. Given $\mathbf{g} \in \mathfrak{G}$, $\mathbf{g} = (g_i : i \in \mathbb{Z})$ and $m \leqslant n$ set $\mathbf{g}_m^n = (g_m, \ldots, g_n)$. For $\ell \geqslant 1$ denote by $\mathfrak{G}_\ell$ the set of all allowed finite words of length ℓ in $\mathfrak{G}$. Given $g \in G$, $\mathcal{F}(g) = \{h \in G : (g, h) \in \mathfrak{G}_2\}$ is the set of followers of g. In the same way, one defines $\mathcal{P}(g)$ the set of predecessors of g.

For $g \in G$ put $\mathcal{F}^1(g) = \mathcal{F}(g)$ and for any $n > 1$ define recursively $\mathcal{F}^{n+1}(g) = \cup_{h \in \mathcal{F}(g)} \mathcal{F}^n(h)$. From [4], for every $n \geqslant 1$, $\mathcal{F}^n := \mathcal{F}^n(0)$ is a normal subgroup of G and for any $h \in G$ there exists $h' \in G$ such that $\mathcal{F}^n(h) = h' + \mathcal{F}^n$. In particular, one can choose an arbitrary map $f : G \to G$ such that $f(g) \in \mathcal{F}(g)$. It follows that $\mathcal{F}(g) = f(g) + \mathcal{F}$. Furthermore, the irreducibility of $\mathfrak{G}$ implies it is mixing and thus there exists $m \geqslant 0$ such that $\mathcal{F}^m = G$. From now on we assume $\mathfrak{G}$ is irreducible and we denote by r the smallest m satisfying the previous property.

Given $n \geqslant 1$ and $g, h \in G$, one defines

$$\mathbb{C}^n(g, h) = \{(g_1, \ldots, g_{n-1}) \in \mathfrak{G}_{n-1} : \quad (g, g_1, \ldots, g_{n-1}, h) \in \mathfrak{G}_{n+1}\}.$$

Notice that $\mathbb{C}^n(0, 0)$ is a normal subgroup of $\mathfrak{G}_{n-1}$ and thus for all $n \geqslant r$ and $g, h \in G$, one has that

$$|\mathbb{C}^n(g, h)| = |\mathbb{C}^n(0, 0)| = |G|^{-1}|\mathcal{F}|^n. \tag{1}$$

In fact, since $n \geqslant r$, given $g, h \in G$ the set $\mathbb{C}^n(g, h)$ is not empty and $\cup_{g,h\in G}\mathbb{C}^n(g, h) = \mathfrak{G}_{n+1}$. One deduces that $|G|^2|\mathbb{C}^n(0, 0)| = |\mathfrak{G}_{n+1}| = |G||\mathcal{F}|^n$.

Denote by $\mathbb{N}$ the set of all non-negative integers and by $\mathbb{N}^*$ the set of all positive integers. Let $\mathfrak{G}^-$ and $\mathfrak{G}^+$ be the projections of $\mathfrak{G}$ on $G^{-\mathbb{N}^*}$ and $G^{\mathbb{N}}$, respectively. Given $w \in \mathfrak{G}^-$ denote by $\mathfrak{G}^+_w$ the projection on $\mathfrak{G}^+$ of the set of all sequences $(g_i : i \in \mathbb{Z}) \in \mathfrak{G}$ with $g_i = w_i$ for $i \leqslant -1$. Notice that if $n \geqslant r$ then for any $w \in \mathfrak{G}^-$ one has $\sigma^n(\mathfrak{G}^+_w) = \mathfrak{G}^+$.

Let μ be any shift-invariant probability measure on $\mathfrak{G}$. For $w \in \mathfrak{G}^-$, $w = (w_{-1}, w_{-2}, \ldots)$ let μ_w be the conditional probability measure on $\mathfrak{G}^+_w$. We say μ has *complete connections* compatible with $\mathfrak{G}$ if given $a \in G$, for all $w \in \mathfrak{G}^-$ such that $a \in \mathcal{F}(w_{-1})$, one has $\mu_w(a) > 0$. If μ is a probability measure with complete connections, we define the quantities γ_m, for $m \geqslant 1$, by

$$\gamma_m := \sup\left\{\left|\frac{\mu_v(a)}{\mu_w(a)} - 1\right| : v, w \in \mathfrak{G}^-; v_{-i} = w_{-i}, 1 \leqslant i \leqslant m; a \in \mathcal{F}(v_{-1}) = \mathcal{F}(w_{-1})\right\}. \tag{2}$$

When $\sum_{m\geqslant 1}\gamma_m < \infty$, we say μ has *summable memory decay*, and this implies a uniform continuity condition on $\mu_w(a)$ as a function of w.

Denote by ν the Haar measure on $\mathfrak{G}$, which is the Markovian measure given by the stochastic matrix $\mathbf{L} = (L_{gh} : g, h \in G)$, where $L_{gh} = |\mathcal{F}^{-1}|\mathbf{1}_{\mathcal{F}(g)}(h)$ and the $\mathbf{L}$-stationary vector $\rho = (\rho_g = |G|^{-1} : g \epsilon G)$. Recall ν is the maximal entropy measure for the Markov shift $(\mathfrak{G}, \sigma)$ and it is the unique Φ-invariant Markov measure with full support. For all $m \leqslant n$, let $\ell = m - n + 1$ and $\mathbf{g} = (g_m, \ldots, g_n) \in \mathfrak{G}_\ell$, it holds that $\nu\{\mathbf{x} \in \mathfrak{G} : \quad \mathbf{x}^n_m = \mathbf{g}\} = |G|^{-1}|\mathcal{F}|^{-(\ell-1)}$.

Now, using the above notations we are able to state the main result.

Theorem 1. *Let G be p^s-torsion with p prime and $\mathfrak{G} \subseteq G^{\mathbb{Z}}$ be an irreducible subgroup shift. Let $\Phi : \mathfrak{G} \to \mathfrak{G}$ be an affine cellular automaton and μ be a shift-invariant probability measure with complete connections and summable memory decay compatible with $\mathfrak{G}$. Then, the Cesàro mean of μ under the action of Φ converges to the Haar measure.*

Notice that to prove theorem 1 it is enough to show that for any $w \in \mathfrak{G}^-$, $m \in \mathbb{N}$ and $\mathbf{g} \in \mathfrak{G}_{m+1}$:

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\mu_w((\Phi^n\mathbf{x})^m_0 = \mathbf{g}) = |G|^{-1}|\mathcal{F}|^{-m}. \tag{3}$$

Remark 1. *We point out that from Kitchens' theorem proof one can get that any probability measure with complete connections and summable memory decay μ on $\mathfrak{G}$ is projected by an isomorphism $\pi : \mathfrak{G} \to G^{\mathbb{Z}}$ onto the probability measure $\mu' := \mu \circ \pi^{-1}$ on $G^{\mathbb{Z}}$.*

Indeed, the isomorphism π given in theorem 1(ii) of [4] can be constructed with memory ℓ and anticipation 0, which implies that for any $w' \in G^{\mathbb{Z}^-}$ and $a' \in \mathcal{F}(w'_{-1})$ there exist unique $w \in \mathfrak{G}^-$ and $a \in \mathcal{F}(w_{-1})$ such that $aw = \pi^{-1}(a'w')$. Therefore, one gets $\mu'_{w'}(a') = \mu_w(a)$ that gives the complete connection property for μ'. Now, denote by γ_n and γ'_n the quantities defined by the expression (2) for the measures μ and μ', respectively. It can be shown that the way used to construct π in [4], allows us to deduce that $\gamma'_m \leqslant \gamma_{m-\ell}$, which implies the summable decay property of μ'. Now, if $\Phi = \mathbf{a} \cdot id_G + \mathbf{b} \cdot \sigma_G + \mathbf{c}$ is an affine cellular automaton on $\mathfrak{G}$, then it is topologically conjugated to $\varphi = \mathbf{a} \cdot id_{A^{\mathbb{Z}}} * \mathbf{b} \cdot \sigma_{A^{\mathbb{Z}}} * \pi(\mathbf{c})$ through π. Therefore, whenever $*$ is a 1-block operation one can conclude the Cesàro mean convergence of μ directly from [3]. However, for most of the cases $*$ will be a k-block operation for $k > 1$, and this argumentation cannot be used to conclude the Cesàro mean convergence of μ under the action of Φ.

In any case our proof of theorem 1 does not use the properties on projections of complete connections and summable memory decay measures supplied by the construction of Kitchens.

3. Processes with infinite memory and regeneration times

As mentioned in the introduction we follow closely the construction made in [2]. Let μ be a probability measure with complete connections and summable memory decay compatible with $\mathfrak{G}$. In this section, for a fixed $w \in \mathfrak{G}^-$ we construct a stochastic process $\{x_n : n \geqslant 0)$ and a probability law $\mathbb{P}_w$ such that

$$\mathbb{P}_w\{x_n = g_n | x_{n-1} = g_{n-1}, \ldots, x_0 = g_0\} = \mu_{g_{n-1}\ldots g_0 w}(g_n). \tag{4}$$

Let $w \in \mathfrak{G}^-$ and $g \in \mathcal{F}(w_{-1})$. By assumption of complete connections $\mu_w(g) > 0$. Define,

$$a_{-1}(g|w) = \inf\{\mu_v(z) : v \in \mathfrak{G}^-, z \in \mathcal{F}(v_{-1})\},$$

and

$$a_0(g|w) = \inf\{\mu_v(g) : v \in \mathfrak{G}^- \textit{ such that } g \in \mathcal{F}(v_{-1})\}.$$

Notice that $a_{-1}(g|w)$ depends neither on g nor on w, while $a_0(g|w)$ does not depend on w. Furthermore, the function $\mu_{(\cdot)}(z) : \{v \in \mathfrak{G}^- : z \in \mathcal{F}(v_{-1})\} \to [0, 1]$ is continuous because μ has summable memory decay, and is non-zero because μ has complete connections. Thus, since $\mathfrak{G}$ is compact, $a_{-1}(g|w) > 0$. Let $\alpha > 0$ be such that $\alpha < a_{-1}(g|w)|\mathcal{F}|$. Now, for $k \geqslant 1$, set

$$a_k(g|w) = \inf\{\mu_v(g) : \quad v \in \mathfrak{G}^-, \quad v_i = w_i, \ -k \leqslant i \leqslant -1, \quad g \in \mathcal{F}(v_{-1})\},$$

which is non-decreasing in k and $\lim_{k\to\infty} a_k(g|w) = \mu_w(g)$.

We will use the above quantities to determine a partition of the interval [0, 1] as follows. Let $b_{-1}(g|w) = a_{-1}(g|w) - \alpha|\mathcal{F}|^{-1}$ and for $k \geqslant 0$ put $b_k(g|w) = a_k(g|w) - a_{k-1}(g|w)$. Then we can construct a partition of $(\alpha, 1]$ by intervals $B_k(g|w)$ of Lebesgue measure $b_k(g|w)$, respectively, disposed in increasing order with respect to g and k. That is, writing $g_1, \ldots, g_{|\mathcal{F}|} \in \mathcal{F}(w_{-1})$ with $g_i < g_{i+1}$, $i \in \{1, \ldots, |\mathcal{F}|\}$, then one orders the intervals: $B_{-1}(g_1|w), \ldots, B_{-1}(g_{|\mathcal{F}|}|w), B_0(g_1|w), \ldots, B_0(g_{|\mathcal{F}|}|w), \ldots$. From this construction

$$\left|\bigcup_{k\geqslant -1} Bk(g|w)\right| = \mu_w(g) - \alpha|\mathcal{F}|^{-1} \qquad \text{and} \qquad \left|\bigcup_{g\in\mathcal{F}(w_{-1})\bigcup_{k\geqslant -1}} B_k(g|w)\right| = 1 - \alpha. \tag{5}$$

Then [0, 1] can be written as the disjoint union $[0, \alpha] \cup \bigcup_{g\in\mathcal{F}(w_{-1})} \bigcup_{k\geqslant -1} B_k(g|w)$.

Consider $\mathbb{P}$ a probability law such that $(U_n : n \in \mathbb{Z})$ and $(V_n : n \in \mathbb{N})$ are two independent sequences of i.i.d. random variables such that U_n is uniformly distributed in $[0, 1]$ and V_n is uniformly distributed in $\mathcal{F}$, that is, $\mathbb{P}\{V_n = g\} = |\mathcal{F}|^{-1}$ for any $g \in \mathcal{F}$. Then, for each $w \in \mathfrak{G}^-$ one can construct recursively a stochastic process $(x_n : n \geqslant 0)$ by

$$x_n := (f(x_{n-1}) + V_n)\mathbf{1}_{[0,\alpha]}(U_n) + \sum_{g\in\mathcal{F}(x_{n-1})} g \sum_{k\geqslant -1} \mathbf{1}_{B_k(g|x_{n-1}\ldots x_0 w)}(U_n), \tag{6}$$

where $f : G \to G$ is an arbitrary map such that $f(g) \in \mathcal{F}(g)$.

From here, we shall denote $\mathbb{P}_w := \mathbb{P}$ the probability law of the process $(x_n : n \geqslant 0)$ which was constructed with respect to the past w. One checks that the equation (4) holds,

$$\begin{aligned}
&\mathbb{P}_w\{x_n = g_n | x_{n-1} = g_{n-1}, \ldots, x_0 = g_0\} \\
&\qquad = \mathbb{P}_w\{U_n \leqslant \alpha, V_n = g_n - f(g_{n-1})\} + \mathbb{P}_w\{U_n \in \cup_{k\geqslant -1} B_k(g_n|g_{n-1}\ldots g_0 w)\} \\
&\qquad = \alpha|F|^{-1} + |\cup_{k\geqslant -1} B_k(g_n|g_{n-1}\ldots g_0 w)| =_{(*)} \mu_{g_{n-1}\ldots g_0 w}(g_n),
\end{aligned}$$

where $=_{(*)}$ is because of equalities (5).

Now, define for $\ell \geqslant -1$,

$$a_\ell = \min_{w\in\mathfrak{G}^-} \left\{ \sum_{g\in\mathcal{F}(w_{-1})} a_\ell(g|w) \right\}.$$

It is a non-decreasing non-zero sequence such that for any $w \in \mathfrak{G}^-$

$$[0, a_k] \subseteq [0, \alpha] \cup \bigcup_{\ell=-1}^{k} \bigcup_{g\in\mathcal{F}(w_{-1})} B_\ell(g|w).$$

Hence, one has recovered lemma 2.4 of [2]. Namely, for $n \in \mathbb{N}$, in the event $\{U_n \leqslant a_k\}$ one only needs to look at $x_{n-1}, \ldots, x_{n-k}$ to decide the value of x_n.

For $i \leqslant j$, denote by $\mathbf{U}_i^j \leqslant \alpha$ the event $\{U_i \leqslant \alpha, U_{i+1} \leqslant \alpha, \ldots, U_j \leqslant \alpha\}$. Given $m \geqslant 1$ define times $(T_i^{(m)} : i \geqslant 1)$ by

$$T_1^{(m)} = \min\{n \geqslant 0 : \mathbf{U}_n^{n+m} \leqslant \alpha, U_{n+m+j+1} \leqslant a_{j-1}, j \geqslant 0\}$$

and for $i \geqslant 2$

$$T_i^{(m)} = \min\{n > T_{i-1}^{(m)} : \mathbf{U}_n^{n+m} \leqslant \alpha, U_{n+m+j+1} \leqslant a_{j-1}, j \geqslant 0\}.$$

Let $\mathbf{N}^{(m)}$ be the counting measure on $\mathbb{N}$ induced by $(T_i^{(m)} : i \geqslant 1)$. That is, for any $A \subseteq \mathbb{N}$,

$$\mathbf{N}^{(m)}(A) = \sum_{i\geqslant 1} \mathbf{1}_A(T_i^{(m)}), \qquad \mathbf{N}^{(m)}(\{n\}) = \mathbf{N}^{(m)}(n).$$

In particular,

$$\{\mathbf{N}^{(m)}(n) = 1\} \Leftrightarrow \{\mathbf{U}_n^{n+m} \leqslant \alpha, U_{n+m+j+1} \leqslant a_{j-1}, j \geqslant 0\}.$$

Let $\mathbb{P}\{\mathbf{U}_n^{n+m} \leqslant \alpha, U_{n+m+j+1} \leqslant a_{j-1}, \forall j \geqslant 0\} = \alpha^{m+1} a_0 a_1 \cdots = \beta$. We claim $\beta > 0$. Indeed, since μ has summable memory decay $a_k(g|v) \geqslant (1 - \gamma_k)\mathbb{P}_w\{x_0 = g\}$ if $v_i = w_i$ for $i \in \{-k, \ldots, -1\}$. Therefore, $a_k \geqslant (1 - \gamma_k)$ and $\sum_{k\geqslant 0}(1 - a_k) \leqslant \sum_{k\geqslant 0} \gamma_k < \infty$, which implies $\beta > 0$. This is a crucial fact to show that $\mathbf{N}^{(m)}$ is a renewal stationary process on $\mathbb{N}$, with finite inter-renewal mean as was done in lemma 2.5 [2]. Then, by lemma 3.1 in [2],

$$\exists \varepsilon : \mathbb{N} \to \mathbb{R}^+ \text{ non-increasing, } \varepsilon(n) \to_{n\to\infty} 0 \text{ and } \mathbb{P}\{\mathbf{N}^{(m)}(A) = 0\} \leqslant \varepsilon(|A|). \tag{7}$$

4. The Cesàro limit

In this section we prove theorem 1. To achieve it one needs to prove the following two results, which are infinite memory measure versions of lemmas 3.1 and 3.2 in [9].

Recall we fixed r as the smallest integer such that $\mathcal{F}^r = G$. Given $w \in \mathfrak{G}^-$, let $\mathbf{x} = (x_n : n \in \mathbb{N})$ be the stochastic process associated to w.

Lemma 1. *Let $k \geqslant r$, $m \geqslant 0$, $l \geqslant 1$, $\mathbf{g} \in \mathfrak{G}_{m+1}$, $\mathbf{h} \in \mathfrak{G}_{k-r+1}$ such that $h_0 \in \mathcal{F}(w_{-1})$, and $\mathbf{y} \in \mathfrak{G}_l$, then*

(i) $\mathbb{P}_w\{\mathbf{x}_k^{k+m} = \mathbf{g} | \mathbf{U}_{k-r+1}^{k+m} \leqslant \alpha, \mathbf{x}_0^{k-r} = \mathbf{h}\} = |G|^{-1}|\mathcal{F}|^{-m}$,

(ii) $\mathbb{P}_w\{\mathbf{x}_k^{k+m} = \mathbf{g} | \mathbf{N}^{(m+2r-1)}(k-r+1) = 1, \mathbf{x}_0^{k-r} = \mathbf{h}, \mathbf{x}_{k+m+r}^{k+m+r+l-1} = \mathbf{y}\} = |G|^{-1}|\mathcal{F}|^{-m}$.

Proof.(I) Let $\mathbf{h} = (h_0, h_1, \ldots, h_{k-r}) \in \mathfrak{G}_{k-r+1}$ such that $h_0 \in \mathcal{F}(w_{-1})$, $g = (g_k, g_{k+1}, \ldots, g_{k+m}) \in \mathfrak{G}_{m+1}$, then

$$\begin{aligned}\mathbb{P}_w\{\mathbf{x}_k^{k+m} = \mathbf{g} | \quad &\mathbf{U}_{k-r+1}^{k+m} \leqslant \alpha, \quad \mathbf{x}_0^{k-r} = \mathbf{h}\} \\ &= \sum_{\mathbf{z} \in \mathbb{C}^r(h_{k-r}, g_k)} \mathbb{P}_w\{\mathbf{x}_{k-r+1}^{k-1} = \mathbf{z}, \mathbf{x}_k^{k+m} = \mathbf{g} | \mathbf{U}_{k-r+1}^{k+m} \leqslant \alpha, \mathbf{x}_0^{k-r} = \mathbf{h}\} \\ &=_{(*)} \sum_{\mathbf{z} \in \mathbb{C}^r(h_{k-r}, g_k)} |\mathcal{F}|^{-(r+m)} = |G|^{-1}|\mathcal{F}|^{-m},\end{aligned}$$

where $=_{(*)}$ is because the values $(x_n : k-r+1 \leqslant n \leqslant k+m)$ are distributed as $(V_n : k-r+1 \leqslant n \leqslant k+m)$, where each component is uniformly distributed on $\mathcal{F}$, and $|\mathbb{C}^r(h_{k-r}, g_k)| = |\mathbb{C}^r(0,0)| = |G|^{-1}|\mathcal{F}|^r$.

(ii) Fix $N = k+m+r+l-1$. Let $\mathbf{h} = (h_0, h_1, \ldots, h_{k-r}) \in \mathfrak{G}_{k-r+1}$, $\mathbf{g} = (g_k, \ldots, g_{k+m}) \in \mathfrak{G}_{m+1}$ and $\mathbf{y} = (y_{k+m+r}, \ldots, y_N) \in \mathfrak{G}_{N-k-m-r+1}$. Put $\tilde{m} = m + 2r - 1$ and $\tilde{k} = k - r + 1$, then

$$\begin{aligned}\mathbb{P}_w\{\mathbf{x}_k^{k+m} = \mathbf{g} | \quad &\mathbf{N}^{(\tilde{m})}(\tilde{k}) = 1, \quad \mathbf{x}_0^{k-r} = \mathbf{h}, \quad \mathbf{x}_{k+m+r}^N = \mathbf{y}\} \\ &= \frac{\mathbb{P}_w\{\mathbf{x}_{k+m+r}^N = \mathbf{y} | \mathbf{N}^{(\tilde{m})}(\tilde{k}) = 1, \mathbf{x}_0^{k-r} = \mathbf{h}, \mathbf{x}_k^{k+m} = \mathbf{g}\}}{\mathbb{P}_w\{\mathbf{x}_{k+m+r}^N = \mathbf{y} | \mathbf{N}^{(\tilde{m})}(\tilde{k}) = 1, \mathbf{x}_0^{k-r} = \mathbf{h}\}} \\ &\cdot \mathbb{P}_w\{\mathbf{x}_k^{k+m} = \mathbf{g} | \mathbf{N}^{(\tilde{m})}(\tilde{k}) = 1, \mathbf{x}_0^{k-r} = \mathbf{h}\}.\end{aligned} \tag{8}$$

Notice that

$$\begin{aligned}&\mathbb{P}_w\{\mathbf{x}_{k+m+r}^N = \mathbf{y} | \mathbf{N}^{(\tilde{m})}(\tilde{k}) = 1, \mathbf{x}_0^{k-r} = \mathbf{h}, \mathbf{x}_k^{k+m} = \mathbf{g}\} \\ &= \sum_{\mathbf{z} \in \mathbb{C}^r(g_{k+m}, y_{k+m+r})} \mathbb{P}_w\{\mathbf{x}_{k+m+r}^N = \mathbf{y}, \mathbf{x}_{k+m+1}^{k+m+r-1} = \mathbf{z} | \mathbf{N}^{(\tilde{m})}(\tilde{k}) = 1, \mathbf{x}_0^{k-r} = \mathbf{h}, \mathbf{x}_k^{k+m} = \mathbf{g}\}.\end{aligned} \tag{9}$$

For each $\mathbf{z} \in \mathbb{C}^r(g_{k+m}, y_{k+m+r})$ one has

$$\begin{aligned}&\mathbb{P}_w\{\mathbf{x}_{k+m+r}^N = \mathbf{y}, \mathbf{x}_{k+m+1}^{k+m+r-1} = \mathbf{z} | \mathbf{N}^{(\tilde{m})}(\tilde{k}) = 1, \mathbf{x}_0^{k-r} = \mathbf{h}, \mathbf{x}_k^{k+m} = \mathbf{g}\} \\ &= \sum_{\mathbf{v} \in \mathbb{C}^r(h_{k-r}, g_k)} \mathbb{P}_w\{\mathbf{x}_{k+m+r}^N = \mathbf{y}, \mathbf{x}_{k+m+1}^{k+m+r-1} = \mathbf{z}, \mathbf{x}_{k-r+1}^{k-1} = \mathbf{v} | \mathbf{N}^{(\tilde{m})}(\tilde{k}) = 1, \mathbf{x}_0^{k-r} = \mathbf{h}, \mathbf{x}_k^{k+m} = \mathbf{g}\} \\ &= \sum_{\mathbf{v} \in \mathbb{C}^r(h_{k-r}, g_k)} \mathbb{P}_w\{\mathbf{x}_{k+m+r}^N = \mathbf{y} | \mathbf{N}^{(\tilde{m})}(\tilde{k}) = 1, \mathbf{x}_0^{k-r} = \mathbf{h}, \mathbf{x}_{k-r+1}^{k-1} = \mathbf{v}, \mathbf{x}_k^{k+m} = \mathbf{g}, \mathbf{x}_{k+m+1}^{k+m+r-1} = \mathbf{z}\} \\ &\qquad \cdot \mathbb{P}_w\{\mathbf{x}_{k+m+1}^{k+m+r-1} = \mathbf{z} | \mathbf{N}^{(\tilde{m})}(\tilde{k}) = 1, \mathbf{x}_0^{k-r} = \mathbf{h}, \mathbf{x}_{k-r+1}^{k-1} = \mathbf{v}, \mathbf{x}_k^{k+m} = \mathbf{g}\} \\ &\qquad \cdot \mathbb{P}_w\{\mathbf{x}_{k-r+1}^{k-1} = \mathbf{v} | \mathbf{N}^{(\tilde{m})}(\tilde{k}) = 1, \mathbf{x}_0^{k-r} = \mathbf{h}, \mathbf{x}_k^{k+m} = \mathbf{g}\} \\ &=_{(\sharp)} C|\mathcal{F}|^{-(r-1)} \sum_{\mathbf{v} \in \mathbb{C}^r(h_{k-r}, g_k)} \mathbb{P}_w\{\mathbf{x}_{k-r+1}^{k-1} = \mathbf{v} | \mathbf{N}^{(\tilde{m})}(\tilde{k}) = 1, \mathbf{x}_0^{k-r} = \mathbf{h}, \mathbf{x}_k^{k+m} = \mathbf{g}\},\end{aligned} \tag{10}$$

where $=_{(\sharp)}$ is because

- in the event $\{\mathbf{N}^{(\tilde{m})}(\tilde{k}) = 1, \quad \mathbf{x}_0^{k+m+r-1} = \mathbf{u}\}$ it follows that $\mathbf{x}_{k+m+r}^N = \mathbf{y}$ does not depend on the chosen $\mathbf{u} \in \mathbb{C}^{k+m+r+1}(w_{-1}, y_{k+m+r})$ which implies $\mathbb{P}\{\mathbf{x}_{k+m+r}^N = \mathbf{y} | \mathbf{N}^{(\tilde{m})}(\tilde{k}) = 1, \mathbf{x}_0^{k+m+r-1} = \mathbf{u}\} = C(\mathbf{y}) := C > 0$ a constant; and

- for all $k+m+1 \leqslant n \leqslant k+m+r-1$ the value of x_n only depends on the random variable V_n which is uniformly distributed on $\mathcal{F}$.

For each $\mathbf{v} \in \mathbb{C}^r(h_{k-r}, g_k)$ one has

$$
\begin{aligned}
&\mathbb{P}_w\{\mathbf{x}_{k-r+1}^{k-1} = \mathbf{v} | \mathbf{N}^{(\tilde{m})}(\tilde{k}) = 1, \mathbf{x}_0^{k-r} = \mathbf{h}, \mathbf{x}_k^{k+m} = \mathbf{g}\} \\
&= \frac{\mathbb{P}_w\{\mathbf{x}_k^{k+m} = \mathbf{g} | \mathbf{N}^{(\tilde{m})}(\tilde{k}) = 1, \mathbf{x}_0^{k-r} = \mathbf{h}, \mathbf{x}_{k-r+1}^{k-1} = \mathbf{v}\}}{\mathbb{P}_w\{\mathbf{x}_k^{k+m} = \mathbf{g} | \mathbf{N}^{(\tilde{m})}(\tilde{k}) = 1, \mathbf{x}_0^{k-r} = \mathbf{h}\}} \mathbb{P}_w\{\mathbf{x}_{k-r+1}^{k-1} = \mathbf{v} | \mathbf{N}^{(\tilde{m})}(\tilde{k}) = 1, \mathbf{x}_0^{k-r} = \mathbf{h}\} \\
&=_{(1)} \frac{|\mathcal{F}|^{-(m+1)}}{\mathbb{P}_w\{\mathbf{x}_k^{k+m} = \mathbf{g} | \mathbf{N}^{(\tilde{m})}(\tilde{k}) = 1, \mathbf{x}_0^{k-r} = \mathbf{h}\}} |\mathcal{F}|^{-(r-1)} =_{(2)} \frac{|\mathcal{F}|^{-(m+1)}}{|G|^{-1}|\mathcal{F}|^{-m}} |\mathcal{F}|^{-(r-1)}, \qquad (11)
\end{aligned}
$$

where $=_{(1)}$ reduces by the same 'renewal-epoch' argument as in equation (10) and $=_{(2)}$ by part (i).

Combining equations (9), (10) and (11), and since $|\mathbb{C}^r(h_{k-r}, g_k)| = |\mathbb{C}^r(g_{k+m}, y_{k+m+r})| = |G|^{-1}|\mathcal{F}|^r$, one gets

$$
\mathbb{P}_w\{\mathbf{x}_{k+m+r}^N = \mathbf{y} | \mathbf{N}^{(\tilde{m})}(\tilde{k}) = 1, \mathbf{x}_0^{k-r} = \mathbf{h}, \mathbf{x}_k^{k+m} = \mathbf{g}\} = \frac{C|\mathcal{F}|}{|G|}. \qquad (12)
$$

On the other hand, one has

$$
\begin{aligned}
&\mathbb{P}_w\{\mathbf{x}_{k+m+r}^N = \mathbf{y} | \mathbf{N}^{(\tilde{m})}(\tilde{k}) = 1, \mathbf{x}_0^{k-r} = \mathbf{h}\} \\
&\qquad = \sum_{\mathbf{g}' \in \mathfrak{G}_{m+1}} \mathbb{P}_w\{\mathbf{x}_{k+m+r}^N = \mathbf{y} | \mathbf{N}^{(\tilde{m})}(\tilde{k}) = 1, \mathbf{x}_0^{k-r} = \mathbf{h}, \mathbf{x}_k^{k+m} = \mathbf{g}'\} \qquad (13) \\
&\qquad \cdot \mathbb{P}_w\{\mathbf{x}_k^{k+m} = \mathbf{g}' \, | \mathbf{N}^{(\tilde{m})}(\tilde{k}) = 1, \mathbf{x}_0^{k-r} = \mathbf{h}\} \\
&\qquad = \frac{C|\mathcal{F}|}{|G|} \sum_{\mathbf{g}' \in \mathfrak{G}_{m+1}} \mathbb{P}_w\{\mathbf{x}_k^{k+m} = \mathbf{g}' \, | \mathbf{N}^{(\tilde{m})}(\tilde{k}) = 1, \mathbf{x}_0^{k-r} = \mathbf{h}\} = \frac{C|\mathcal{F}|}{|G|}.
\end{aligned}
$$

Then, replacing (12) and (13) in (8), and using part (i) one concludes the result. □

From lemma 1 one can deduce the relevance of considering a measure with complete connections and summable memory decay. Indeed, for such a measure one has that there exists $0 < \alpha < a_{-1}(g|w)|\mathcal{F}|$, which can be interpreted as a positive transition probability to any state (allowed in the topological Markov chain $\mathfrak{G}$) independently of the past and w. This implies that the state of $(x_n : n \in \mathbb{N})$ between times k and $k+m$ is independent of the past and on w in the event $\mathbf{U}_{k-r+1}^{k+m} \leqslant \alpha$. Furthermore, the summable memory decay property is necessary to get that $\{U_{n+m+j+1} \leqslant a_{j-1}, \forall j \geqslant 0\}$ is an event with positive probability. This was used to conclude in part (ii) of previous lemma the independence of $\mathbf{x}_k^{k+m}$ with respect to the future between the times $k+m+r$ and $k+m+r+\ell-1$. To finish let us note that the summable memory decay property also is needed to get (7), which will be used later in the proof of theorem 1.

Now, consider $\Phi = \mathbf{a}\, id + \mathbf{b}\, \sigma + \mathbf{c}$ an affine cellular automaton on $\mathfrak{G}$, with $\mathbf{a}$ and $\mathbf{b}$ relatively prime to p and $\mathbf{c} = (\ldots, c, c, c, \ldots) \in \mathfrak{G}$. We will not distinguish the operations on G and $\mathfrak{G}$. Then given $\mathbf{x} \in \mathfrak{G}$, $i \in \mathbb{N}$ and $n \geqslant 1$ it follows

$$
(\Phi^n \mathbf{x})_i = \sum_{k=0}^{n} \binom{n}{k} \mathbf{a}^{n-k} \mathbf{b}^k x_{i+k} + c \sum_{k=0}^{n-1} \sum_{\ell=0}^{k} \binom{k}{\ell} \mathbf{a}^{k-\ell} \mathbf{b}^{\ell}.
$$

For every $m \in \mathbb{Z}$ denote by $m^{(s)}$ its equivalence class $mod(p^s)$ in $\mathbb{Z}_{p^s}$. Then

$$(\Phi^n \mathbf{x})_i = \sum_{k=0}^{n} \left(\binom{n}{k} \mathbf{a}^{n-k} \mathbf{b}^k \right)^{(s)} x_{i+k} + c \sum_{k=0}^{n-1} \sum_{\ell=0}^{k} \left(\binom{k}{\ell} \mathbf{a}^{k-\ell} \mathbf{b}^{\ell} \right)^{(s)} . \tag{14}$$

Suppose $\ell, m, k \geqslant 0$, $n \geqslant m + \ell + 1$ and $m \leqslant k \leqslant n - \ell$, then one says that k is (m, ℓ)-*isolated in* n if (i) $\binom{n}{k}^{(s)} \neq 0$; and (ii) for every $k' \in \{k - m, \dots, k + \ell\}$, $k' \neq k$, then $\binom{n}{k'}^{(s)} = 0$.

Lemma 2. *Let $m \geqslant 0$, $n \geqslant 2r + 2m + 1$, and assume k is $(r + m, r + m)$-isolated in n. Then, for every $i \in \mathbb{N}$ and $\mathbf{g} \in \mathfrak{G}_{m+1}$:*
$\mathbb{P}_w\{(\Phi^n \mathbf{x})_i^{i+m} = \mathbf{g} | \mathbf{U}_{i+k-r+1}^{i+k+m+r} \leqslant \alpha, \quad U_{i+k+m+r+j+1} \leqslant a_{j-1}, \quad j \geqslant 0\} = |G|^{-1} |\mathcal{F}|^{-m}$.

Proof. Define $\mathbf{X} = (X_i, \dots, X_{i+m})$ and $\mathbf{Y} = (Y_i, \dots, Y_{i+m})$ by

$$Y_{i+j} = \left(\binom{n}{k} \mathbf{a}^{n-k} \mathbf{b}^k \right)^{(s)} x_{i+j+k}, \;\; j \in \{0, \dots, m\} \tag{15}$$

and

$$X_{i+j} := (\Phi^n \mathbf{x})_{i+j} - Y_{i+j}, \;\; j \in \{0, \dots, m\}. \tag{16}$$

Notice that

$$\begin{aligned} X_{i+j} &= \sum_{\substack{k'=0 \\ k' \neq k}}^{n} \left(\binom{n}{k'} \mathbf{a}^{n-k'} \mathbf{b}^{k'} \right)^{(s)} x_{i+j+k'} + c \sum_{k'=0}^{n-1} \sum_{\ell=0}^{k'} \left(\binom{k'}{\ell} \mathbf{a}^{k'-\ell} \mathbf{b}^{\ell} \right)^{(s)} \\ &=_{(*)} \sum_{k'=0}^{k-r-m-1} \left(\binom{n}{k'} \mathbf{a}^{n-k'} \mathbf{b}^{k'} \right)^{(s)} x_{i+j+k'} + \sum_{k'=k+r+m+1}^{n} \left(\binom{n}{k'} \mathbf{a}^{n-k'} \mathbf{b}^{k'} \right)^{(s)} x_{i+j+k'} \\ &\quad + c \sum_{k'=0}^{n-1} \sum_{\ell=0}^{k'} \left(\binom{k'}{\ell} \mathbf{a}^{k'-\ell} \mathbf{b}^{\ell} \right)^{(s)}, \end{aligned} \tag{17}$$

where $=_{(*)}$ is because k is $(r + m, r + m)$-isolated. From definition $(\Phi^n \mathbf{x})_i^{i+m} = \mathbf{X} + \mathbf{Y} \in \mathfrak{G}_{m+1}$. From $\mathbf{Y} = \left(\binom{n}{k} \mathbf{a}^{n-k} \mathbf{b}^k\right)^{(s)} \cdot \mathbf{x}_{i+k}^{i+k+m}$ one deduces that $\mathbf{Y} \in \mathfrak{G}_{m+1}$ and since $\mathfrak{G}_{m+1}$ is a group one concludes that $\mathbf{X}, \mathbf{Y} \in \mathfrak{G}_{m+1}$. By lemma 3.4 [9] $\binom{n}{k}^{(s)}$ is relatively prime to p. Since $\mathbf{a}$ and $\mathbf{b}$ are also relatively prime to p, it follows that $\mathbf{d} = \left(\binom{n}{k} \mathbf{a}^{n-k} \mathbf{b}^k\right)^{(s)}$ has a multiplicative inverse $\mathbf{d}^{-1}$ in $\mathbb{Z}_{p^s}$ and so $\mathbf{d}^{-1} \cdot \mathbf{Y} = \mathbf{x}_{i+k}^{i+k+m}$. Observe that if $\mathbf{x}_i^{i+k-r-1} = \mathbf{h}$ and $\mathbf{x}_{i+k+m+r+1}^{i+m+n} = \mathbf{z}$ then $\mathbf{X} = \kappa(\mathbf{h}, \mathbf{z})$. Hence, putting $\tilde{m} = m + 2r - 1$ and $\tilde{k} = i + k - r + 1$ one obtains

$$\begin{aligned} &\mathbb{P}_w\{(\Phi^n \mathbf{x})_i^{i+m} = \mathbf{g} | \mathbf{N}^{(\tilde{m})}(\tilde{k}) = 1\} \\ &= \sum_{\substack{\mathbf{z} \in \mathfrak{G}_{n-k-r} \\ \mathbf{h} \in \mathfrak{G}_{k-r}, \\ \mathbb{C}^{i+1}(w_{-1}, h_0) \neq \emptyset}} \mathbb{P}_w\{\mathbf{Y} = \mathbf{g} - \kappa(\mathbf{h}, \mathbf{z}) | \mathbf{x}_i^{i+k-r-1} = \mathbf{h}, \mathbf{x}_{i+k+m+r+1}^{i+m+n} = \mathbf{z}, \mathbf{N}^{(\tilde{m})}(\tilde{k}) = 1\} \\ &\qquad \cdot \mathbb{P}_w\{\mathbf{x}_i^{i+k-r-1} = \mathbf{h}, \mathbf{x}_{i+k+m+r+1}^{i+m+n} = \mathbf{z} | \mathbf{N}^{(\tilde{m})}(\tilde{k}) = 1\} \\ &=_{(\sharp)} |G|^{-1} |\mathcal{F}|^{-m} \sum_{\substack{\mathbf{z} \in \mathfrak{G}_{n-k-r} \\ \mathbf{h} \in \mathfrak{G}_{k-r}, \\ \mathbb{C}^{i+1}(w_{-1}, h_0) \neq \emptyset}} \mathbb{P}_w\{\mathbf{x}_i^{i+k-r-1} = \mathbf{h}, \mathbf{x}_{i+k+m+r+1}^{i+m+n} = \mathbf{z} | \mathbf{N}^{(\tilde{m})}(\tilde{k}) = 1\} = |G|^{-1} |\mathcal{F}|^{-m}, \end{aligned}$$

where $=_{(\sharp)}$ follows from lemma 1(ii). $\square$

4.1. Proof of theorem 1

This proof uses strongly the Pascal triangle properties showed in [9].

First, let us introduce the following notation. For $a, i, n \in \mathbb{N}$, let $n = \sum_{j\in\mathbb{N}} n_j p^j$ be the decomposition of n in base p (so $n_j \in \{0, ..., p-1\}$ for every j); $J_i(n) := \{j \geqslant i : n_j \neq 0\}$; $\xi_i(n) := |J_i(n)|$; and $p^a\mathbb{N} := \{p^a \cdot n : n \in \mathbb{N}\} = \{n \in \mathbb{N} : n_j = 0, \forall 0 \leqslant j < a\}$.

Suppose $a \geqslant 2s+1$, $p^a \geqslant 2m+2r+1$, and let

$$\mathcal{M} := \{n \in p^a\mathbb{N} : n \geqslant p^2(m+r)^2,\ \xi_{a+\lfloor \frac{1}{2}\log_p(n)\rfloor}(n) \geqslant \tfrac{1}{5}\log_p(n)\},$$

where $\lfloor c \rfloor$ denotes the integer part of any given real number c, and so given $n \in \mathcal{M}$ let $A(n) := \{k \leqslant n : k \text{ is } (m+r, m+r)\text{-}\textit{isolated in } n\}$. In lemma 3.4 and the proof of theorem 1.1 in [9] is shown that there exists $C > 0$ (it suffices to consider $C^{-1} = 5\log_2 p$) such that $|A(n)| \geqslant n^C - 1$. Thus, for any $i \geqslant 0$,

$$\mathbb{P}\{\exists k \in A(n)\ with\ \mathbf{U}_{i+k-r+1}^{i+k+m+r} \leqslant \alpha, \quad U_{i+k+m+r+j+1} \leqslant a_{j-1}, \quad j \geqslant 0\}$$

$$= \mathbb{P}\{\mathbf{N}^{(m+2r-1)}((i-r+1)+A(n)) \neq 0\} = 1 - \varepsilon(|A(n)|) \geqslant 1 - \varepsilon(\lfloor n^C - 1\rfloor),$$

where $\varepsilon : \mathbb{N} \to \mathcal{R}^+$ is a non-increasing function which converges to zero as n goes to infinity (see (7)).

Then, using lemma 2 we get for any $w \in \mathfrak{G}^-$, $\mathbf{g} \in \mathfrak{G}_{m+1}$, and $n \in \mathcal{M}$:

$$|\mathbb{P}_w\{(\Phi^n\mathbf{x})_i^{i+m} = \mathbf{g}\} - |G|^{-1}|\mathcal{F}|^{-m}| \leqslant \varepsilon(\lfloor n^C - 1\rfloor).$$

Lemma 3.5 in [9] says that the relative density of $\mathcal{M}$ is one in $p^a\mathbb{N}$; that is,

$$\lim_{n\to\infty} \frac{|\mathcal{M} \cap \{1, \ldots, n\}|}{|p^a\mathbb{N} \cap \{1, \ldots, n\}|} = 1.$$

This implies that the $(m+1)$-dimensional marginal of the Cesàro mean converges along $p^a\mathbb{N}$, that is,

$$\lim_{N\to\infty} \frac{1}{|p^a\mathbb{N} \cap \{0, \ldots, N-1\}|} \sum_{n\in p^a\mathbb{N}\cap\{0,\ldots,N-1\}} \mathbb{P}\{(\Phi^n\mathbf{x})_i^{i+m} = \mathbf{g}\} = |G|^{-1}|\mathcal{F}|^{-m}.$$

Therefore, since the Haar measure is invariant for powers of Φ, one has that for any $0 \leqslant j < p^a$, the $(m+1)$-dimensional marginal of the Cesàro mean also converges along $\mathcal{M}_j = \{n+j; n \in \mathcal{M}\}$. Hence, we conclude from the fact that

$$p^a \lim_{N\to\infty} \frac{1}{N} \sum_{n\leqslant N} \mathbb{P}\{(\Phi^n\mathbf{x})_i^{i+m} = \mathbf{g}\}$$

$$= \sum_{0\leqslant j<p^a} \frac{1}{|\mathcal{M}_j \cap \{0, \ldots, N-1\}|} \sum_{n\in\mathcal{M}_j\cap\{0,\ldots,N-1\}} \mathbb{P}\{(\Phi^n\mathbf{x})_i^{i+m} = \mathbf{g}\} = |G|^{-1}|\mathcal{F}|^{-m}.$$

We can use theorem 1 together with Kitchens' result to conclude about the convergence of Cesàro mean distribution of wider classes of algebraic cellular automata. In fact, if we consider an Abelian k-block group shift $(\mathfrak{A}, \tilde{+})$, that is, the group operation $\tilde{+}$ is a sliding block code from $\mathfrak{A} \times \mathfrak{A}$ to $\mathfrak{A}$ with a k-block local rule (see [4] and [6]), then we can take the following result.

Proposition 1. *Suppose* $(\mathfrak{A}, \tilde{+})$ *is* p^s*-torsion for some prime number* p *and define* $\varphi := \mathbf{a} \cdot id\ \tilde{+}\ \mathbf{b} \cdot \sigma\ \tilde{+}\ \mathbf{c}$, *where* $\mathbf{a}, \mathbf{b} \in \mathbb{N}$ *are relatively prime to* p, *and* $\mathbf{c} \in \mathfrak{A}$ *is a constant sequence. If* μ *is a probability measure with complete connections and summable memory decay compatible with* $\mathfrak{A}$, *then the Cesàro mean distribution of the iterates of* μ *under the action of* φ *converges to the Haar measure.*

Proof. From proposition 4 in [4] we set that there exists a 1-block shift group $(\mathfrak{G}, +)$, such that $(\mathfrak{A}, \tilde{+}, \sigma_{\mathfrak{A}})$ is isomorphic to $(\mathfrak{G}, +, \sigma_{\mathfrak{G}})$, that is, there exists a map $\pi : \mathfrak{A} \to \mathfrak{G}$ which is a topological conjugacy between $(\mathfrak{A}, \sigma_{\mathfrak{A}})$ and $(\mathfrak{G}, \sigma_{\mathfrak{G}})$, and an isomorphism between $(\mathfrak{A}, \tilde{+})$ and $(\mathfrak{G}, +)$. Therefore, π is also a topological conjugacy between $(\mathfrak{A}, \varphi)$ and $(\mathfrak{G}, \Phi)$, where $\Phi = \mathbf{a} \cdot id_{\mathfrak{G}} + \mathbf{b} \cdot \sigma_{\mathfrak{G}} + \pi(\mathbf{c})$. In particular, $+$ is a 1-block operation on $\mathfrak{G}$, and so $(\mathfrak{G}, \Phi)$ is an affine cellular automaton as in theorem 1.

Suppose μ is a probability measure on $\mathfrak{A}$ with complete connections compatible with $\mathfrak{A}$ and summable memory decay. We need to show that $\mu' := \mu \circ \pi^{-1}$ also has complete connections compatible with $\mathfrak{G}$ and summable memory decay.

In fact, the construction of the proof of proposition 4 in [4] gives that π is a Higher Block Code (see [6]). Without loss of generality we can consider π has memory ℓ and anticipation 0. Thus, given a cylinder $[a'_i, \ldots, a'_m]$ of $\mathfrak{G}$ we have that $\pi^{-1}\big([a'_i, \ldots, a'_m]\big) = [a_{i-\ell}, \ldots, a_m]$ is a cylinder of $\mathfrak{A}$. Hence, given $a' \in G$, for all $w' \in \mathfrak{G}^-$ such that $a' \in \mathcal{F}(w'_{-1})$, set $aw = \pi^{-1}(a'w')$ with $a \in \mathfrak{A}$ and $w \in \mathfrak{A}^-$. One has that

$$\mu'_{w'}(a') = \mu_{\pi^{-1}(w')}(\pi^{-1}(a')) = \mu_w(a) > 0,$$

and

$$\gamma'_m := \sup\left\{\left|\frac{\mu_{v'}(a')}{\mu_{w'}(a')} - 1\right| : \ v', w' \in \mathfrak{G}^-;\ v'_{-i} = w'_{-i},\ 1 \leqslant i \leqslant m;\ a' \in \mathcal{F}(v'_{-1}) = \mathcal{F}(w'_{-1})\right\}$$

$$= \sup\left\{\left|\frac{\mu_{v}(a)}{\mu_{w}(a)} - 1\right| : \ v, w \in \mathfrak{A}^-;\ v_{-i} = w_{-i},\ 1 \leqslant i \leqslant m+\ell;\ a \in \mathcal{F}(v_{-1}) = \mathcal{F}(w_{-1})\right\}$$

$$= \gamma_{m+\ell}.$$

Now, since μ' has complete connections and summable memory decay, we use theorem 1 to get Φ randomizes μ' to ν', that is the Haar measure on $\mathfrak{G}$. Then

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \mu \circ \varphi^{-n} = \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \mu \circ \pi^{-1} \circ \Phi^{-n} \circ \pi = \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \mu' \circ \Phi^{-n} \circ \pi = \nu' \circ \pi,$$

and due to the uniqueness of the maximum-entropy measure we have $\nu' \circ \pi$ is the Haar measure on $\mathfrak{A}$.

Acknowledgments

The authors would like to thank Professor Marcus Pivato who read this paper carefully and made important suggestions.

This work was supported by Nucleus Millennium Information and Randomness ICM P04-069-F and MECESUP UCH0009.

References

[1] Athreya K B and Ney P 1978 A new approach to the limit theory of recurrent Markov chains *Trans. AMS* **245** 493–501

[2] Ferrari P A, Maass A, Martínez S and Ney P 2000 Cesàro mean distribution of group automata starting from measures with summable decay *Ergod. Theory Dyn. Syst.* **20** 1657–70

[3] Host B, Maass A and Martínez S 2003 Uniform Bernoulli measure in dynamics of permutative cellular automata with algebraic local rules *Discrete Contin. Dyn. Syst.* **9** 1423–46

[4] Kitchens B P 1987 Expansive dynamics on zero-dimensional groups *Ergod. Theory Dyn. Syst.* **7** 249–61

[5] Lind D A 1984 Applications of ergodic theory and sofic systems to cellular automata *Phys. D* **10** 36–44

[6] Lind D A and Marcus B 1995 *An Introduction to Symbolic Dynamics and Coding* (Cambridge: Cambridge University Press)

[7] Maass A and Martínez S 1998 On Cesàro limit distribution of a class of permutative cellular automata *J. Stat. Phys.* **90** 435–52

[8] Maass A, Martínez S, Pivato M and Yassawi R 2006 Asymptotic randomization of subgroup shifts by linear cellular automata *Ergod. Theory Dyn. Syst.* **26** 1203–24

[9] Maass A, Martínez S, Pivato M and Yassawi R 2006 Attractiveness of the Haar measure for linear cellular automata on Markov subgroups *IMS Lecture Notes — Monograph Series, Dynamics and Stochastics* **48** pp100–8 *Festschrift in honour of Michael Keane*

[10] Pivato M and Yassawi R 2002 Limit measures for affine cellular automata *Ergod. Theory Dyn. Syst.***22** 1269–87

[11] Pivato M and Yassawi R 2004 Limit measures for affine cellular automata II *Ergod. Theory Dyn. Syst.* **24** 1961–80