

The Measure-Theoretic Entropy of Linear Cellular Automata with Respect to a Markov Measure

HASAN AKIN

Department of Mathematics, Education Faculty,
Zirve University, 27260, Gaziantep, Turkey
akinhasan25@gmail.com

Abstract. The purpose of this short paper is to compute the measure-theoretic entropy of the one-dimensional linear cellular automata defined by bipermutative local rules on the ring \mathbb{Z}_m , $m \geq 2$, with respect to a Markov measure generated by a stochastic matrix P and a probability vector π such that $\pi P = \pi$.

2010 Mathematics Subject Classification: Primary: 28D20; Secondary: 37A35, 37B40

Keywords and phrases: Cellular automata, measure entropy, Markov measure.

1. Introduction

Cellular automata (CAs for short), begun by Stanislaw Ulam and John von Neumann, are discrete models studied extensively in mathematics, physics, theoretical biology, computability theory and microstructure modelling with different purposes. Hedlund's paper [11] started investigation of current problems in symbolic dynamics. CAs from the point of view of the ergodic theory have received remarkable attention in the last few years [1, 8, 13]. Shereshevsky [14] has defined n th iteration of a permutative CA and shown that if the local rule f is right (left) permutative, then its n th iteration also is right (left) permutative. For the definition and some properties of one-dimensional linear cellular automata (1-D LCA) we refer to [10, 11].

Mass and Martinez [13] have studied the dynamics of Markov measures by a particular linear cellular automata (LCA). They have reviewed some results on the evolution of probability measures under CA acting on a fullshift.

The concept of entropy has been studied in computer science, mathematics, physics, chemistry, information theory and social sciences. In the last decades, a lot of works are devoted to this subject (see [1–7] and [9, 12]). This notion first arose in thermodynamics as a measure of the heat absorbed (or emitted), when external work is done on a system. In mathematics, the notion of entropy of a dynamical system has been given by Kolmogorov [12]. This entropy is known as

Communicated by Lee See Keong.

Received: September 17, 2008; Revised: March 27, 2010.

measure-theoretic entropy, Kolmogorov entropy, Kolmogorov-Sinai entropy, or just KS entropy. In this paper, we deal with KS entropy.

In [1], the author has computed the measure-theoretic entropy with respect to uniform Bernoulli measure for the case $\lambda_i = 1$, for all $i \in \mathbb{Z}_m$. In [5], the author has computed the measure entropy of the 1-D LCA defined by a bipermutative local rule with respect to arbitrary Bernoulli measure. In this paper, we generalize the result obtained in [5] to any Markov measure generated by a stochastic matrix $P = (p_{ij})$ and probability vector $\pi = (p_i)$. We compute the measure-theoretic entropy of the 1-D LCA $T_{f[-l,r]}$ defined by a bipermutative local rule with respect to a Markov measure $\mu_{\pi P}$ generated by a stochastic matrix P and a probability vector π such that $\pi P = \pi$. We show that if the local rule f is bipermutative, then we have

$$h_{\mu_{\pi P}}(T_{f[-l,r]}) = -(l+r) \sum_{i,j=0}^{m-1} \pi_i p_{ij} \log p_{ij}.$$

2. Preliminaries and definitions

2.1. The 1-D linear cellular automata

Let $\mathbb{Z}_m = \{0, 1, \dots, m-1\}$ ($m \geq 2$) be a ring of the integers modulo m and $\mathbb{Z}_m^{\mathbb{Z}}$ be the space of all doubly-infinite sequences $x = (x_n)_{n=-\infty}^{\infty} \in \mathbb{Z}_m^{\mathbb{Z}}$ and $x_n \in \mathbb{Z}_m$. A CA can be defined as a homomorphism $\mathbb{Z}_m^{\mathbb{Z}}$ with product topology. A CA is a continuous map $T : \mathbb{Z}_m^{\mathbb{Z}} \rightarrow \mathbb{Z}_m^{\mathbb{Z}}$ defined by $(Tx)_i = f(x_{i-l}, \dots, x_{i+r})$, where $f : \mathbb{Z}_m^{r+l+1} \rightarrow \mathbb{Z}_m$ is a given local rule or map. Favati *et al.* [10] have stated that a local rule f is linear (additive) if and only if it can be written as

$$(2.1) \quad f(x_{-l}, \dots, x_r) = \sum_{i=-l}^r \lambda_i x_i \pmod{m},$$

where at least one between λ_{-l} and λ_r is nonzero. In this paper, we consider 1-D LCA $T_{f[-l,r]}$ defined by the local rule f in (2.1):

$$(2.2) \quad (Tx) = (y_n)_{n=-\infty}^{\infty}, y_n = f(x_{n-l}, \dots, x_{n+r}) = \sum_{i=-l}^r \lambda_i x_{n+i} \pmod{m},$$

where $\lambda_{-l}, \dots, \lambda_r \in \mathbb{Z}_m$.

We are going to use the notation $T_{f[-l,r]}$ for LCA-map defined in (2.2) to emphasize the local rule f and the numbers $-l$ and r .

Definition 2.1. *The local rule defined by (2.1) is permutative in x_j , $-l \leq j \leq r$, iff for any given finite sequence $\bar{x}_{-l}, \dots, \bar{x}_{j-1}, \bar{x}_{j+1}, \dots, \bar{x}_r \in \mathbb{Z}_m^{l+r}$ we have*

$$\{f(\bar{x}_{-l}, \dots, \bar{x}_{j-1}, x_j, \bar{x}_{j+1}, \dots, \bar{x}_r) : x_j \in \mathbb{Z}_m\} = \mathbb{Z}_m.$$

The linear local rule f defined by (2.1) is permutative in the j th variable if and only if $\gcd(\lambda_j, m) = 1$, where \gcd denotes the greatest common divisor. A local rule f is said to be right (respectively, left) permutative, if $\gcd(\lambda_r, m) = 1$ (respectively, $\gcd(\lambda_{-l}, m) = 1$). It is said that f is bipermutative if it is both left and right permutative (see [11] for a formal definition of permutivity).

2.2. Measure-theoretic entropy

Recall necessary definitions and Theorems (see [9] and [15] for details). Let X be a set and \mathcal{B} be a σ -algebra of subsets of X . The pair (X, \mathcal{B}) is called a measurable space. If the function $\mu : \mathcal{B} \rightarrow [0, 1]$ is a finite measure on the space (X, \mathcal{B}) , then (X, \mathcal{B}, μ) is called probability space (or measure space). Let $T : X \rightarrow X$ be a measure preserving transformation, then (X, \mathcal{B}, μ, T) is called measure theoretical dynamical system.

If $\xi = \{A_1, \dots, A_n\}$ and $\beta = \{B_1, \dots, B_m\}$ are two measurable partitions of X , their refinement $\xi \vee \beta$ is the partition $\{A_i \cap B_j : i = 1, \dots, n; j = 1, \dots, m\}$. Also, $T^{-1}\xi$ is the partition $\{T^{-1}A_1, \dots, T^{-1}A_n\}$.

Definition 2.2. Let (X, \mathcal{B}, μ, T) be a measure-theoretical dynamical system and ξ be a measurable partition of X . The partition ξ is called a strong generator if

$$\bigvee_{k=0}^{\infty} T^{-k}\xi = \mathcal{B}.$$

Definition 2.3. Let ξ be a measurable partition of X . The quantity

$$H_{\mu}(\xi) = -\sum_{A \in \xi} \mu(A) \log \mu(A)$$

is called the entropy of the partition ξ . Let ξ be a partition with finite entropy, then the quantity

$$h_{\mu}(\xi, T) = \lim_{n \rightarrow \infty} \frac{1}{n} H_{\mu}(\bigvee_{i=0}^{n-1} T^{-i}\xi)$$

is called the entropy of ξ with respect to T . The quantity

$$h_{\mu}(T) = \sup_{\xi} \{h_{\mu}(\xi, T) : \xi \text{ is a partition with } H_{\mu}(\xi) < \infty\}$$

is called the measure-theoretic entropy of (X, \mathcal{B}, μ, T) , the entropy of T (with respect to μ).

Theorem 2.1. [15, Theorem 4.18] If T is a measure-preserving transformation (but not necessary invertible) of the probability space (X, \mathcal{B}, μ) and if ξ is a finite partition of \mathcal{B} with $\bigvee_{k=0}^{\infty} T^{-k}\xi = \mathcal{B}$ then $h_{\mu}(T) = h_{\mu}(\xi, T)$.

Theorem 2.1 is known as Kolmogorov-Sinai theorem. They proved that if a partition ξ is a strong generator for T then $h_{\mu}(T) = h_{\mu}(T, \xi)$.

3. The main result

In this section we study the measure-theoretic entropy of the 1-D LCA given in equation (2.2) with respect to Markov measure. In order to compute the entropy of the 1-D LCA over ring \mathbb{Z}_m ($m \geq 2$) we must define σ -algebra and the Markov measure. In symbolic dynamical system, it is well known that this σ -algebra \mathcal{B} is generated by thin cylinder sets

$$C =_a [j_0, j_1, \dots, j_s]_{s+a} = \{x \in \mathbb{Z}_m^{\mathbb{Z}} : x_a = j_0, \dots, x_{a+s} = j_s\},$$

where $j_0, j_1, \dots, j_s \in \mathbb{Z}_m$.

Recall that a subshift of finite type $\sigma : X \rightarrow X$ defined on a space

$$(3.1) \quad X = \{x \in \mathbb{Z}_m^{\mathbb{Z}} : M_{(x_n, x_{n+1})} = 1, n \in \mathbb{Z}\}$$

for some $m \times m$ irreducible matrix M with entries either zero or unity. Let $P = (p_{(i,j)})$ denote a $m \times m$ stochastic matrix ($p_{(i,j)} \geq 0, \sum_{j=0}^{m-1} p_{(i,j)} = 1$) with entries $p_{(i,j)} = 0$ iff $M_{(i,j)} = 0$, let $\pi = \{\pi_0, \pi_1, \dots, \pi_{m-1}\}$ be its left eigenvector such that $\pi P = \pi$.

Put a measure determined on cylinder sets by an m -dimensional probability vector π (the initial distribution) and a stochastic transition matrix P such that $\pi P = \pi$ on the space $\mathbb{Z}_m^{\mathbb{Z}}$ of sequences on an alphabet \mathbb{Z}_m with m symbols. A pair (π, P) defines a set function $\mu_{\pi P}$ on the cylinders of $\mathbb{Z}_m^{\mathbb{Z}}$. Recall that the associated Markov measure is defined as follows:

$$\mu_{\pi P}(0[i_0, \dots, i_k]_k) = \pi_{i_0} p_{(i_0, i_1)} \cdots p_{(i_{k-1}, i_k)} = \pi_{i_0} p_{i_0 i_1} \cdots p_{i_{k-1} i_k},$$

where the (i, j) -th element of the stochastic matrix P can be accepted as the transition probability from the state i to the state j . See [9, 13, 15] for the properties of the Markov measure. We take definitions given in [1, 5] again for the sake of the integrity of expression.

Let us consider a particular case. Assume that the local rule f is bipermutative, so, we have the following Lemma.

Lemma 3.1. [5, Lemma 1] *Suppose that $f(x_{-l}, \dots, x_r) = \sum_{i=-l}^r \lambda_i x_i \pmod{m}$ is a bipermutative local rule, and $\xi = \{0[i] : 0 \leq i < m\}$ is a partition of $\mathbb{Z}_m^{\mathbb{Z}}$, then the partition ξ is a strong generator for 1-D LCA defined by the local rule f .*

From Lemma 3.1, one can show that the partition $\bigvee_{k=0}^n T_{f[-l,r]}^{-k}(\xi)$ consists of all cylinder sets of length $(l+r)n+1$: $_{-ln}[j_{-ln}, j_{-ln+1}, \dots, j_{rn}]_{rn} = \{x \in \mathbb{Z}_m^{\mathbb{Z}} : x_{-ln} = j_{-ln}, \dots, x_{rn} = j_{rn}\}, j_{-ln}, \dots, j_{rn} \in \mathbb{Z}_m$.

The following theorem is the main result of this short paper. Our main purpose here is to compute the measure entropy of a bipermutative LCA with respect to an arbitrary Markov measure.

Theorem 3.1. *Let $\mu_{\pi P}$ be a Markov measure given by the stochastic matrix $P = (p_{ij})$ and the probability vector $\pi = (p_i)$. Assume that l and r are positive integers and $\gcd(\lambda_{-l}, m) = 1, \gcd(\lambda_r, m) = 1$. Let $T_{f[-l,r]}$ be 1-D LCA defined by bipermutative local rule f . Then we have*

$$h_{\mu_{\pi P}}(T_{f[-l,r]}) = -(l+r) \sum_{i,j=0}^{m-1} \pi_i p_{ij} \log p_{ij}.$$

Proof. Now we can calculate the measure entropy of the 1-D LCA by means of the Kolmogorov-Sinai Theorem [15, p.95], namely, $h_{\mu_{\pi P}}(T_{f[-l,r]}) = h_{\mu_{\pi P}}(T_{f[-l,r]}, \xi)$. Let ξ be the zero-time partition of $\mathbb{Z}_m^{\mathbb{Z}}$: $\xi = \{0[i] : 0 \leq i < m\}$, where $0[i] = \{x \in \mathbb{Z}_m^{\mathbb{Z}} : x_0 = i\}$ is a cylinder set for all $i, 0 \leq i < m$. So, we can state the partition ξ as follows:

$$\xi = \{0[0]_{,0} [1], \dots, 0[m-1]\}.$$

From the definition of entropy we have

$$H_{\mu_{\pi P}}(\xi) = - \sum_{i=0}^{m-1} \mu_{\pi P}(0[i]) \log \mu_{\pi P}(0[i]) = - \sum_{i=0}^{m-1} \pi_i \log \pi_i < \infty.$$

So, we have

$$\begin{aligned} & H_{\mu_{\pi P}} \left(\bigvee_{k=0}^n T_{f[-l,r]}^{-k} \xi \right) \\ &= - \sum_{-nl, -nl+1, \dots, nl=0}^{m-1} \pi_{i_{-nl}} p_{i_{-nl} i_{-nl+1}} \cdots p_{i_{nr-1} i_{nr}} \log \pi_{i_{-nl}} p_{i_{-nl} i_{-nl+1}} \cdots p_{i_{nr-1} i_{nr}} \\ &= - \sum_{-nl, -nl+1, \dots, nl=0}^{m-1} \pi_{i_{-nl}} p_{i_{-nl} i_{-nl+1}} \cdots p_{i_{nr-1} i_{nr}} [\log \pi_{i_{-nl}} + \log p_{i_{-nl} i_{-nl+1}} + \cdots \\ &\quad + \log p_{i_{nr-1} i_{nr}}] \\ &= - \sum_{-nl=0}^{m-1} \pi_{-nl} \log \pi_{-nl} - ((n(l+r)+1) \sum_{i,j=0}^{m-1} \pi_i p_{ij} \log p_{ij}), \end{aligned}$$

where $\sum_{j=0}^{m-1} p_{ij} = 1$ and $\sum_{i=0}^{m-1} \pi_i p_{ij} = \pi_j$.

From Theorem 2.1 and Lemma 3.1, we have

$$\begin{aligned} h_{\mu_{\pi P}}(T_{f[-l,r]}) &= h_{\mu_{\pi P}}(T_{f[-l,r]}, \xi) = - \lim_{n \rightarrow \infty} \frac{1}{n} H_{\mu_{\pi P}} \left(\bigvee_{k=0}^n T_{f[-l,r]}^{-k} \xi \right) \\ &= -(l+r) \sum_{i,j=0}^{m-1} \pi_i p_{ij} \log p_{ij}. \end{aligned}$$

Therefore the proof is completed. ■

Example 3.1. Let us consider stochastic matrix $P = \begin{bmatrix} 1/2 & 1/2 \\ 2/3 & 1/3 \end{bmatrix}$. To compute π_0, π_1 we write the equation $\pi P = \pi$. So, the probability vector corresponding to the stochastic matrix P can be obtained as $\pi = (4/7, 3/7)$.

Now, define local rule f as follows:

$$f(x_{-1}, x_0, x_1) = x_{-1} + x_0 + x_1 \pmod{2}.$$

Let $\mu_{\pi P}$ on $\mathbb{Z}_2^{\mathbb{Z}}$ be the Markov measure defined by stochastic matrix P and probability vector π associated to the matrix P . Then we have

$$h_{\mu_{\pi P}}(T_{f[-1,1]}) = -2 \sum_{i,j=0}^1 \pi_i p_{ij} \log p_{ij} = \frac{4}{7} + \log \sqrt[7]{\frac{27}{4}}.$$

In probability theory, it constitutes a measure of the uncertainty. In the information theory, it is known that this measure carries the amount of information of the system. Therefore, the value $(4/7) + \log \sqrt[7]{27/4}$ bits is the amount of information that can be carried (or the amount of information obtained).

Example 3.2. Let $f(x_{-2}, x_{-1}, x_0, x_1, x_2) = x_{-2} + 3x_{-1} + 2x_0 + 3x_2 \pmod{2^2}$. It is clear that this local rule is bipermutative.

Let us consider the stochastic matrix

$$(3.2) \quad P = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{8} & 0 & \frac{7}{8} & 0 \\ 0 & \frac{1}{16} & \frac{1}{16} & \frac{7}{8} \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \end{bmatrix}.$$

To compute $\pi_0, \pi_1, \pi_2, \pi_3$ we write the equations $\pi P = \pi$:

$$(3.3) \quad \begin{aligned} \frac{\pi_0}{2} + \frac{\pi_1}{8} + \frac{\pi_3}{2} &= \pi_0 \\ \frac{\pi_0}{2} + \frac{\pi_2}{16} &= \pi_1 \\ \frac{7\pi_1}{8} + \frac{\pi_2}{16} &= \pi_2 \\ \frac{7\pi_2}{8} + \frac{\pi_3}{2} &= \pi_3. \end{aligned}$$

From (3.3) we obtain $\pi_1 = 60\pi_0/113$, $\pi_2 = 56\pi_0/113$ and $\pi_3 = 98\pi_0/113$. Since $\pi_0 + \pi_1 + \pi_2 + \pi_3 = 1$, we have

$$\pi_0 = \frac{113}{327}, \pi_1 = \frac{60}{327}, \pi_2 = \frac{56}{327}, \pi_3 = \frac{98}{327}.$$

So, the probability vector corresponding to the matrix P given in (3.2) is $\pi = (113/327, 60/327, 56/327, 98/327)$. From Theorem 3.2, we have

$$\begin{aligned} h_{\mu_{\pi P}}(T_{f[-1,1]}) &= -(l+r) \sum_{i,j=0}^3 \pi_i p_{ij} \log p_{ij} = -4 \sum_{i=0}^3 \pi_i \sum_{j=0}^3 p_{ij} \log p_{ij} \\ &= 4[211 \cdot \log 2 + \frac{1}{656}(654 \cdot \log 2 - 203 \cdot \log 7)] \cong 848 - 1,23 \cdot \log 7. \end{aligned}$$

Therefore, we conclude that the value $848 - 1,23 \cdot \log 7$ is the amount of information that can be carried by the Markov measure $\mu_{\pi P}$ over $\mathbb{Z}_4^{\mathbb{Z}}$ defined by the stochastic matrix P and the probability vector π . The number $(848 - 1,23 \cdot \log 7)$ can be interpreted as the average uncertainty about what symbol will appear next, given the one it is seen at present.

A Markov measure on $\mathbb{Z}_m^{\mathbb{Z}}$ is uniform, if measure of any one-dimensional cylinder is equal to $1/m$, where m is the cardinality of \mathbb{Z}_m . A doubly stochastic matrix is a matrix P such that P and P^{tr} (transpose) are both stochastic. If a matrix P is a doubly stochastic then corresponding Markov measure is a uniform measure. Cardinality of \mathbb{Z}_m is equal to m , so that any doubly stochastic matrix P of $m \times m$ size will generate uniform Markov measure.

Corollary 3.1. *Let $\mu_{\pi P}$ be the uniform Markov measure on $\mathbb{Z}_m^{\mathbb{Z}}$ and $f(x_{-l}, \dots, x_r) = \sum_{i=-l}^r \lambda_i x_i \pmod{m}$, where $f[-l, r]$ is bipermutative. Then measure-theoretic entropy of the 1-D LCA $T_{f[-l,r]}$ with respect to $\mu_{\pi P}$ is equal to $(l+r) \log m$.*

4. Conclusion

This paper contains the following results: We have found a generating partition for the 1-D LCA defined by a bipermutative local rule (Lemma 3.1). We have calculated the measure-theoretic entropy of the 1-D LCA with respect to any Markov measure (Theorem 3.1). We have illustrated the result by means of examples. This is the first step toward arbitrary Markov measure classification of multi-dimensional CA defined on alphabets of composite cardinality (or different ring [7]). In [1], the author has computed the measure-theoretic entropy with respect to uniform Bernoulli measure for the case $\lambda_i = 1$, for all $i \in \mathbb{Z}_m$. The author proved that the uniform Bernoulli measure is the maximal measure for these LCA. He also posed the question whether the maximal measure is unique. Thus, here a question raises:

Exist another Markov measure of maximal entropy there?

Let $f(x_{-2}, x_{-1}, x_0, x_1, x_2) = 2x_{-2} + 3x_{-1} + 2x_0 + 3x_1 + 2x_2 \pmod{2^2}$. It is clear that this local rule is not bipermutative. Consider the stochastic matrix in (3.2), can the measure entropy of the 1-D LCA $T_{f[-2,2]}$ with respect to Markov measure $\mu_{\pi P}$ be computed? This is an open problem.

Acknowledgement. The author would like to thank the anonymous referees for their careful reading and useful comments which greatly improved the writing and presentation of the paper.

References

- [1] H. Akin, On the measure entropy of additive cellular automata f_∞ , *Entropy* **5** (2003), no. 2, 233–238 (electronic).
- [2] H. Akin, On the directional entropy of \mathbb{Z}^2 -actions generated by additive cellular automata, *Appl. Math. Comput.* **170** (2005), no. 1, 339–346.
- [3] H. Akin, The topological entropy of n th iteration of an additive cellular automata, *Appl. Math. Comput.* **174** (2006), no. 2, 1427–1437.
- [4] H. Akin, On directional entropy of A \mathbb{Z}^2 -action, *Int. J. Appl. Math. Mech.* **2** (2006), no. 1, 94–101.
- [5] H. Akin, The entropy of linear cellular automata with respect to any Bernoulli measure, *Complex Systems* **18** (2009), no. 2, 237–244.
- [6] H. Akin, On the topological directional entropy, *J. Comput. Appl. Math.* **225** (2009), no. 2, 459–466.
- [7] H. Akin and I. Siap, On cellular automata over Galois rings, *Inform. Process. Lett.* **103** (2007), no. 1, 24–27.
- [8] G. Cattaneo, E. Formenti, G. Manzini and L. Margara, Ergodicity, transitivity, and regularity for linear cellular automata over \mathbb{Z}_m , *Theoret. Comput. Sci.* **233** (2000), no. 1–2, 147–164.
- [9] M. Denker, C. Grillenberger and K. Sigmund, *Ergodic Theory on Compact Spaces*, Lecture Notes in Mathematics, Vol. 527, Springer, Berlin, 1976.
- [10] P. Favati, G. Lotti and L. Margara, Additive one-dimensional cellular automata are chaotic according to Devaney’s definition of chaos, *Theoret. Comput. Sci.* **174** (1997), no. 1–2, 157–170.
- [11] G. A. Hedlund, Endomorphisms and automorphisms of the shift dynamical system, *Math. Systems Theory* **3** (1969), 320–375.
- [12] A. N. Kolmogorov, A new metric invariant of transient dynamical systems and automorphisms in Lebesgue spaces, *Dokl. Akad. Nauk SSSR (N.S.)* **119** (1958), 861–864.
- [13] A. Maass and S. Martinez, Evolution of probability measures by cellular automata on algebraic topological Markov chains, *Biol. Res.* **35** (2003), 113–118.

- [14] M. A. Shereshevsky, Ergodic properties of certain surjective cellular automata, *Monatsh. Math.* **114** (1992), no. 3–4, 305–316.
- [15] P. Walters, *An Introduction to Ergodic Theory*, Graduate Texts in Mathematics, 79, Springer, New York, 1982.