

CHARACTERIZATION FOR ENTROPY OF SHIFTS OF FINITE TYPE ON CAYLEY TREES

JUNG-CHAO BAN AND CHIH-HUNG CHANG*

ABSTRACT. The notion of tree-shifts constitutes an intermediate class in between one-sided shift spaces and multidimensional ones. This paper proposes an algorithm for computing of the entropy of a tree-shift of finite type. Meanwhile, the entropy of a tree-shift of finite type is $\frac{1}{p} \ln \lambda$ for some $p \in \mathbb{N}$, where λ is a Perron number. This extends Lind's work on one-dimensional shifts of finite type. As an application, the entropy minimality problem is investigated, and we obtain the necessary and sufficient condition for a tree-shift of finite type being entropy minimal with some additional conditions.

1. INTRODUCTION

A one-dimensional shift space is a set consisting of right-infinite or bi-infinite words which avoid words in a so-called *forbidden set* \mathcal{F} . Such a shift space is denoted by $X_{\mathcal{F}}$. It is no doubt that the most interesting $X_{\mathcal{F}}$ which has been extensively investigated is the shift of finite type (SFT). That is, the shift space $X_{\mathcal{F}}$ such that \mathcal{F} is a finite set.

SFT plays an important role in symbolic dynamical systems and some other fields. For example, Shannon used it as a model of discrete communication channels [28] in information theory. In dynamical systems, due to the work of Adler and Weiss [1] and Bowen [8] that any expansive map admits a Markov partition, SFTs are used to study the hyperbolic dynamics and the classification theory of the Anosov and Axiom A diffeomorphisms. Meanwhile, SFT is also a useful tool for investigating the Lipschitz equivalence and thermodynamic properties in fractal geometry [11, 27].

A significant invariant of the SFTs is its topological entropy, which measures the growth rate of the number of the admissible patterns. Such an invariant reflects the complexity on its own right, we refer readers to [2]

Date: May 1, 2017.

1991 Mathematics Subject Classification. 37E05, 11A63.

Key words and phrases. Shifts of finite type; Cayley tree; entropy; Perron number; entropy minimality.

*To whom correspondence should be addressed.

for more details. For a one-dimensional SFT, since its topological entropy is the logarithm of the spectral radius of a certain non-negative integral matrix [21], the topological entropy are thus being easily calculated. Such calculation method also leads to the further classification of the SFTs. The algebraic characterization of the topological entropy for SFTs is given by D. Lind [20], which reveals that such numbers are Perron numbers. (Recall that a Perron number is a real algebraic integer greater than 1 and greater than the modulus of its algebraic conjugates.) More precisely, the entropies of one-dimensional SFTs are exactly the non-negative rational multiples of logarithms of Perron numbers.

The scenario for multidimensional cases is dramatically different. For example, unlike the one-dimensional case, the computations of the topological entropies for multidimensional SFTs are difficult and there is no general method. Very few models one can compute their rigorous value of entropy. Some approximation algorithms can be found in [7, 14, 22, 23, 24]. For the entropy classification theory, the celebrated result of Hochman and Meyerovitch [17] indicated that the topological entropy of a multidimensional SFT is right recursively enumerable (RRE for short). Roughly speaking, it is the infimum of a monotonic recursive sequence of rational numbers. Such result has been improved by Hochman [16] to the multidimensional effective dynamical systems. Later, Pavlov and Schraudner [25] showed that, for every $d \geq 3$ and every \mathbb{Z}^d full shift, there is a block gluing \mathbb{Z}^d SFT which shares identical topological entropy.

It is worth pointing out that the reason which causes these differences between one- and multidimensional shift spaces are the spatial structure. One-dimensional lattice \mathbb{Z} is a free group with one generator while multidimensional lattice \mathbb{Z}^d , $d \geq 2$, is an abelian group with d generators. Thus, the investigation of symbolic dynamics on Cayley trees arises naturally. Aubrun and Béal [3, 4] introduced the notion of tree-shifts, which is a special kind of Cayley trees, and then studied the classification theory up to conjugacy, languages, and its application to automaton theory. An important point to note here is that such a shift constitutes an intermediate class in between one-sided shifts and multidimensional ones. Thus, it sheds some new light on the study of multidimensional SFTs. In [5], the formal definition of the entropy of a tree-shift was given and the authors demonstrated that the computation of entropy of a tree-shift of finite type (TSFT) is equivalent to solving a system of nonlinear recursive equations (SNRE), and vice versa.

Note that the computation of the rigorous value of entropy is not easy due to the doubly exponential growth rate of the patterns for a TSFT (see Section 2 for more details). Some partial results can be found in [5].

The aim of this paper is to investigate classification theory of entropy. Our theorem below provides a natural and complete characterization of the entropy of TSFTs in an algebraic viewpoint.

Theorem 1.1. *The set of entropies of tree-shifts of finite type is*

$$(1) \quad \mathcal{E} = \left\{ \frac{1}{p} \ln \lambda : \lambda \in \mathcal{P}, p \geq 1 \right\},$$

where \mathcal{P} stands for the set of Perron numbers.

One may ask whether there is some computation method for the entropy of TSFTs. The affirmative solution is given in Theorem 1.2, which says that the entropy is equal to the maximal value of the spectral radius of the adjacency matrices induced from $X_{\mathcal{F}}$, we refer the reader to Section 2.3 for more details of the (reduced) system nonlinear recursive equations and the corresponding adjacency matrix.

Theorem 1.2. *Let X be a TSFT and let F be the corresponding system of nonlinear recursive equations which is defined in (6). Then*

$$(2) \quad h(X) = \max\{\ln \lambda_{ME} : E \text{ is a reduced SNRE of } F\}.$$

Roughly speaking, the entropy of X is attained on the entropy of some subsystems of itself. This makes the differences between the classical one-dimensional SFTs and TSFTs. For the convenience of the reader, we give a table for the computation method and characterization of the entropy. Let Ω be a 1-dimensional SFT with the adjacency matrix M and let λ be its spectral radius.

Entropy \ dimension	1-d SFT	TSFT	r -d SFT
Formula	$\ln \lambda_M$	Theorem 1.2	None
Algebraic criterion	Perrons	Theorem 1.1	None
Computational criterion	Computable	Computable	RRE

From the table one can see that for the algebraic characterization of the entropy for multidimensional SFT is still lacking since there is no general method for the entropy computation. Our method herein sheds some new light on it due to the fact that TSFT is an intermediate class between 1-d and r -d SFTs. After revealing the entropy computation algorithm, we use it to investigate which TSFT is entropy minimal.

It is known that an irreducible \mathbb{Z} SFT is entropy minimal; that is, any proper subshift $Y \subset X$ has smaller entropy than that of an irreducible SFT X . For $r \geq 2$, every \mathbb{Z}^r SFT having the mixing property called *uniform filling property* is entropy minimal while there is a non-trivial *block gluing* \mathbb{Z}^r SFT which is not entropy minimal. Readers are referred to [9, 21, 26] for more details. Recently, it is demonstrated that the dimension minimality of self-affine sets holds for a generic choice in arbitrary dimension. More specifically, let $E_{\mathbf{A}, \mathbf{v}} = \bigcup_{i=1}^K A_i(E_{\mathbf{A}, \mathbf{v}}) + v_i$ be a self-affine set corresponding to $\mathbf{A} = (A_1, \dots, A_K) \in GL_r(\mathbb{R})^K$ and $\mathbf{v} = (v_1, \dots, v_K) \in (\mathbb{R}^r)^K$ with $\|A_i\| < 1$ for all i . A folklore conjecture asserts that $\dim_H E_{\mathbf{A}', \mathbf{v}'} < \dim_H E_{\mathbf{A}, \mathbf{v}}$, where $\mathbf{A}' = (A_1, \dots, A_{K-1})$ and $\mathbf{v}' = (v_1, \dots, v_{K-1})$. There exist simple counterexamples showing that this cannot be the case for all self-affine sets; however, the conjecture holds in arbitrary dimension for a generic choice of the matrix tuple. See [10, 12, 13, 18, 19] and the references therein.

Suppose that $X_{\mathcal{F}}$ is a Markov tree-shift (defined later) with $\mathcal{F} = \{u_1, \dots, u_K\}$ for some $K \in \mathbb{N}$. Proposition 2.10 reveals the necessary and sufficient condition for $X_{\mathcal{F}}$ being entropy minimal. More precisely, $h(X_{\mathcal{F}'}) < h(X_{\mathcal{F}})$, where $\mathcal{F}' = \{u_1, \dots, u_{K-1}\}$.

The rest of this paper is structured as follows. In Section 2, we set up the notation and terminology of the TSFTs; previous results in the computation of the entropy of TSFTs which are useful for the proof of Theorem 1.1 are also presented therein. Section 2.4 applies Theorem 1.2 to investigate some restricted entropy minimality problem and reveals the necessary and sufficient condition. The proofs of Theorem 1.1 and Theorem 1.2 are presented in Section 3.

2. DEFINITIONS AND PREVIOUS RESULTS

This section collects some basic definitions of symbolic dynamics on Cayley trees.

2.1. Basic definitions. Let $\Sigma = \{0, 1, \dots, d-1\}$ and let $\Sigma^* = \bigcup_{n \geq 0} \Sigma^n$ be the union of finite words over Σ , where $\Sigma^n = \{w_1 w_2 \cdots w_n : w_i \in \Sigma \text{ for } 1 \leq i \leq n\}$ is the collection of words of length n for $n \in \mathbb{N}$ and $\Sigma^0 = \{\epsilon\}$ consists of the empty word ϵ . An *infinite tree* t over a finite alphabet \mathcal{A} is a function from Σ^* to \mathcal{A} . Denote by a *node of an infinite tree* a word of Σ^* and the empty word relates to the root of the tree. Suppose x is a node of a tree. x has children xi with $i \in \Sigma$. A sequence of words $(x_k)_{1 \leq k \leq n}$ is called a *path*

if, for all $k \leq n-1$, $x_{k+1} = x_k i_k$ for some $i_k \in \Sigma$. Suppose t is a tree and let x be a node, we refer t_x to $t(x)$ for simplicity. A subset of words $L \subset \Sigma^*$ is called *prefix-closed* if each prefix of L belongs to L . A function u defined on a finite prefix-closed subset L with codomain \mathcal{A} is called a *pattern*, and L is called the *support* of the pattern. A subtree of a tree t rooted at a node x is the tree t' satisfying $t'_y = t_{xy}$ for all $y \in \Sigma^*$ such that xy is a node of t , where $xy = x_1 \cdots x_m y_1 \cdots y_n$ means the concatenation of $x = x_1 \cdots x_m$ and $y = y_1 \cdots y_n$.

Suppose $n \in \mathbb{N} \cup \{0\}$, $\Sigma_n = \bigcup_{k=0}^n \Sigma^k$ denotes the set of words of length at most n . We say that a pattern u is a *block of height n* (or *n -block*) if the support of u is Σ_{n-1} , denoted by $\text{height}(u) = n$. Furthermore, u is a pattern of a tree t if there exists $x \in \Sigma^*$ such that $u_y = t_{xy}$ for every node y of u , and say that u is a pattern of t rooted at the node x in this case. A tree t is said to *avoid* u if u is not a pattern of t . If u is a pattern of t , then u is called an *allowed pattern* of t .

Denote by \mathcal{T} the set of all infinite trees over \mathcal{A} . For $i \in \Sigma$, the shift transformations σ_i from \mathcal{T} to itself are defined as follows. For every tree $t \in \mathcal{T}$, $\sigma_i(t)$ is the tree rooted at the i th child of t , that is, $(\sigma_i(t))_x = t_{ix}$ for all $x \in \Sigma^*$. For the simplification of the notation, we omit the parentheses and denote $\sigma_i(t)$ by $\sigma_i t$. The set \mathcal{T} equipped with the shift transformations σ_i is called the *full tree-shift* of infinite trees over \mathcal{A} . Suppose $w = w_1 \cdots w_n \in \Sigma^*$. Define $\sigma_w = \sigma_{w_n} \circ \sigma_{w_{n-1}} \circ \cdots \circ \sigma_{w_1}$. It follows immediately that $(\sigma_w t)_x = t_{wx}$ for all $x \in \Sigma^*$.

Given a collection of patterns \mathcal{F} , let $X_{\mathcal{F}}$ denote the set of all trees avoiding any element of \mathcal{F} . A subset $X \subseteq \mathcal{T}$ is called a *tree-shift* if $X = X_{\mathcal{F}}$ for some \mathcal{F} . We say that \mathcal{F} is a *set of forbidden patterns* of X . A tree-shift $X = X_{\mathcal{F}}$ is called a *tree-shift of finite type* if the forbidden set \mathcal{F} is finite. Denote by $B_n(X)$ the set of all blocks of height n of X , and $B(X)$ the set of all blocks of X . Suppose $u \in B_n(X)$ for some $n \geq 2$. Let $\sigma_i u$ be the block of height $n-1$ such that $(\sigma_i u)_x = u_{ix}$ for $x \in \Sigma_{n-2}$. The block u is written as $u = (u_\epsilon, \sigma_0 u, \sigma_1 u, \dots, \sigma_{d-1} u)$.

Suppose X and Y are two one-dimensional shift spaces, the Curtis-Lyndon-Hedlund theorem (see [15]) indicates that a map $\phi : X \rightarrow Y$ is a sliding block code if and only if ϕ is continuous and $\phi \circ \sigma_x = \sigma_Y \circ \phi$. A similar discussion extends to tree-shifts; in other words, ϕ is a sliding block code (between tree-shifts) if and only if ϕ is continuous and commutes with all tree-shift maps σ_i for $i \in \Sigma$.

If a sliding block code $\phi : X \rightarrow Y$, herein X and Y are tree-shifts, is onto, then ϕ is called a *factor code* from X to Y . A tree-shift Y is a *factor* of X if there is a factor code from X onto Y . If ϕ is one-to-one, then ϕ is called an *embedding* of X into Y . A sliding block code $\psi : Y \rightarrow X$ is called an *inverse* of ϕ if $\psi(\phi(x)) = x$ for all $x \in X$ and $\phi(\psi(y)) = y$ for all $y \in Y$. In this case, we say that ϕ is *invertible* and write $\psi = \phi^{-1}$.

Definition 2.1. A sliding block code $\phi : X \rightarrow Y$ is a *conjugacy* from X to Y if it is invertible. Two tree-shifts X and Y are called *conjugate*, denoted by $X \cong Y$, if there is a conjugacy from X to Y .

A TSFT $X = X_{\mathcal{F}}$ is called a *Markov tree-shift* if the forbidden set \mathcal{F} consists of 2-blocks. In [6], Ban and Chang showed that every TSFT is conjugated to a Markov tree-shift. Therefore, it suffices to investigate Markov tree-shifts for characterizing the properties of TSFTs.

2.2. Entropy. Let X be a tree-shift. The *entropy* of X is defined as follows.

Definition 2.2. (1) The *entropy* of X , denoted by $h(X)$, is defined as

$$(3) \quad h(X) = \lim_{n \rightarrow \infty} \frac{\ln^2 |B_n(X)|}{n}$$

whenever the limit exists, where $|\cdot|$ stands for the cardinality of a set and $\ln^2 = \ln \circ \ln$.

(2) If $|B_n(X)|$ behaves like $\exp(\alpha \kappa^n)$, such as $|B_n(X)| \approx c \exp(\alpha \kappa^n)$ for instance, where c is a constant, then the value α is called the *hidden entropy* (or *sub-entropy*) of X .

This paper provides an algorithm for the computation of entropy, i.e., the value κ , and gives a complete characterization of such a value. One question still unanswered is whether the same results hold for the hidden entropy α . This question is at present far from being solved, even for the simplest cases. Namely, the case where $(d, k) = (2, 2)$, the description for the hidden entropy α of such X is still lacking. However, the computation of the exact values of $|B_n(X)|$ relies on the both values α and κ .

We introduce the notion of *system of nonlinear recursive equations* which is useful for the computation of the entropy.

Definition 2.3. Let $\mathcal{A} = \{a^{(1)}, a^{(2)}, \dots, a^{(k)}\}$ be the symbol set and suppose \mathcal{A}^d is an ordered set with respect to the lexicographic order, $d \in \mathbb{N}$.

- (1) Let $F = \sum_{\mathbf{a} \in \mathcal{A}^d} f_{\mathbf{a}} \mathbf{a}$ be a binary combination over \mathcal{A}^d , i.e., $f_{\mathbf{a}} \in \{0, 1\}$ for $\mathbf{a} \in \mathcal{A}^d$. The vector $v_F = (f_{\mathbf{a}})_{\mathbf{a} \in \mathcal{A}^d} \in \mathbb{R}^{k^d}$ is called the *indicator vector* of F .
- (2) A sequence $\{a_n^{(1)}, \dots, a_n^{(k)}\}_{n \in \mathbb{N}}$ is defined by a *system of nonlinear recursive equations (SNRE)* of degree (d, k) if

$$a_n^{(i)} = F^{(i)} \text{ for } n \geq 2, 1 \leq i \leq k,$$

and $a_1^{(i)} \in \mathbb{N}$ is given for $1 \leq i \leq k$, where $F^{(1)}, \dots, F^{(k)}$ are binary combinations over $\{a_{n-1}^{(1)}, a_{n-1}^{(2)}, \dots, a_{n-1}^{(k)}\}^d$, respectively.

- (3) A symbol $a^{(i)} \in \mathcal{A}$ is called *essential* if there exists an $m \in \mathbb{N}$ such that $a_m^{(i)} \geq 2$; otherwise, $a^{(i)} = 1$ is called *inessential*.
 - (4) Suppose $F = \{F^{(i)}\}_{i=1}^k$ defines an SNRE. The *indicator matrix* $I_F \in \mathbf{M}_{k \times k^d}$ of F is defined as
- $$(4) \quad I_F = (v_{F^{(i)}})_{i=1}^k.$$

It is remarkable that an SNRE defined by F induces a unique indicator matrix I_F , and vice versa (up to permutation). Furthermore, each $F^{(i)}$ is seen as an ordered binary combination. For example, consider the symbol set $\mathcal{A} = \{a^{(1)}, a^{(2)}\}$ and the following SNRE.

$$(5) \quad \begin{cases} a_n^{(1)} = F^{(1)} = \left(a_{n-1}^{(1)}\right)^2 + \left(a_{n-1}^{(2)}\right)^2, \\ a_n^{(2)} = F^{(2)} = a_{n-1}^{(1)} a_{n-1}^{(2)} + a_{n-1}^{(2)} a_{n-1}^{(1)}, \\ a_1^{(1)} = a_1^{(2)} = 2. \end{cases}$$

Then the corresponding indicator matrix is

$$I_F = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}.$$

For the rest of this paper, we simply use F to describe the SNRE of $\{a_n^{(1)}, \dots, a_n^{(k)}\}_{n=1}^\infty$. Suppose $X = \mathbf{X}_{\mathcal{F}}$ is a TSFT over $\mathcal{A} = \{a^{(1)}, a^{(2)}, \dots, a^{(k)}\}$. Let

$$X_{a^{(i)}} = \{t \in X : t_\epsilon = a^{(i)}\}$$

be the set of those trees whose roots are assigned with the symbol $a^{(i)}$, and let $a_n^{(i)} = |B_n(X_{a^{(i)}})|$, where $1 \leq i \leq k$. Theorem 2.4 follows immediately.

Theorem 2.4 (See [5]). *The values $\{a_n^{(1)}, \dots, a_n^{(k)}\}_{n=1}^\infty$ satisfies the following SNRE.*

$$(6) \quad \begin{cases} a_n^{(i)} = \sum_{(a^{(i)}, i_1, i_2, \dots, i_d) \notin \mathcal{F}} \prod_{j=1}^d a_{n-1}^{(i_j)}, & 1 \leq i \leq k, n \geq 2, \\ a_1^{(i)} = |B_2(X_{a^{(i)}})|, & 1 \leq i \leq k. \end{cases}$$

Notably, the initial condition $a_1^{(i)} = |B_2(X_{a^{(i)}})|$ in (6) is the number of items of $F^{(i)}$ while, generally, the initial condition of an SNRE can be arbitrary. Define the *entropy*, say $h(F)$, for an SNRE $F = \{F^{(i)}\}_{i=1}^k$ as

$$(7) \quad h(F) = \lim_{n \rightarrow \infty} \frac{\ln^2 \sum_{i=1}^k a_n^{(i)}}{n}.$$

Theorem 2.5 indicates that, for any TSFT X , there exists an SNRE F such that $h(X) = h(F)$.

Theorem 2.5 (See [5]). *The entropy of a tree-shift of finite type is realized as a system of nonlinear recurrence equations of degree (d, k) for some $d, k \geq 2$. Conversely, every system of nonlinear recurrence equations of degree (d, k) is corresponding to the entropy of some tree-shifts of finite type.*

Let A and $B \in \mathbf{M}_{m \times n}(\mathbb{Z})$. We say that $A \leq B$ if $A(i, j) \leq B(i, j)$ for $1 \leq i \leq m$ and $1 \leq j \leq n$. Define the reduced SNRE as follows.

2.3. Reduced SNRE. This subsection introduces the notion of *reduced SNRE* which enables us to build up a computational method for the entropy of a TSFT (Theorem 2.4). Let us rewrite the SNRE (6) in the following form.

$$(8) \quad \begin{cases} a_n^{(i)} = F^{(i)} = \sum_{j=1}^{k^d} \alpha_j^{(i)} F_j^{(i)}; \\ a_1^{(i)} = \sum_{j=1}^{k^d} \alpha_j^{(i)}, 1 \leq i \leq k. \end{cases}$$

Definition 2.6 (Reduced SNRE). Suppose X is a TSFT. Let F be the SNRE according to Theorem 2.4 and let $I_F \in \mathbf{M}_{k \times k^d}$ be its indicator matrix. We call E a *reduced SNRE* of F if E is the SNRE defined by some indicator matrix I_E which satisfies the following conditions.

- (i) $I_E \leq I_F$;
- (ii) I_E has exactly one 1's in each row;
- (iii) the initial condition of E is the same as F .

For example, consider the SNRE $F = \{F^{(i)}\}_{i=1}^2$ defined in (5); recall that the indicator matrix is a 2×4 matrix

$$I_F = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}.$$

Then

$$I_E = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

defines a reduced SNRE $E = \{E^{(i)}\}_{i=1}^2$ as follows.

$$\begin{cases} a_n^{(1)} = E^{(1)} = \left(a_{n-1}^{(1)}\right)^2, \\ a_n^{(2)} = E^{(2)} = \left(a_{n-1}^{(1)}\right) \left(a_{n-1}^{(2)}\right), \\ a_1^{(1)} = a_1^{(2)} = 2. \end{cases}$$

We remark here that the initial condition $a_1^{(i)}$ of the reduced SNRE E is no longer the number of the items of $E^{(i)}$ for $i = 1, \dots, k$. If an SNRE E which is defined by some indicator matrix I_E satisfying only (ii), then we also call E a reduced SNRE.

Let $F = \{F^{(i)}\}_{i=1}^k$ be a reduced SNRE. A $k \times k$ non-negative integral matrix M , called the *weighted adjacency matrix* of F , is defined as

$$(9) \quad M(i, j) = \begin{cases} m, & \text{if } a^{(j)} \text{ appears in } F^{(i)} \text{ and the degree of } a_{n-1}^{(j)} \text{ is } m; \\ 0, & \text{otherwise.} \end{cases}$$

2.4. Entropy Minimality Problem. The well-known entropy minimality problem investigates when the entropy of any proper subshift space is strictly smaller than the entropy of the original shift space. This subsection reveals the necessary and sufficient condition for the entropy minimality problem under some additional conditions.

Proposition 2.7. *Suppose $X = \mathbf{X}_{\mathcal{F}}$ is a tree-shift of finite type over $\mathcal{A} = \{a^{(1)}, a^{(2)}, \dots, a^{(k)}\}$ with an SNRE F of degree (d, k) . If every symbol in \mathcal{A} is essential, then $h(X) = \ln d$.*

Proof. It suffices to show that there exists a reduced SNRE E of F such that $h(E) = \ln d$ since $h(X) \leq \ln d$ (cf. [5]). Let E be a reduced SNRE of F . Then the weighted adjacency matrix M_E satisfies $\sum_{j=1}^k M_E(i, j) = d$ for $1 \leq i \leq d$. Since every symbol is essential, Theorem 3.3 infers that the entropy of E is $h(E) = \ln \rho_{M_E}$, where ρ_{M_E} is the spectral radius of M_E . This completes the proof since $\rho_{M_E} = d$. \square

Recall that a TSFT $X = \mathbf{X}_{\mathcal{F}}$ is called a Markov tree-shift if the height of each pattern in \mathcal{F} is less than or equal to two. In [6], Ban and Chang demonstrated that every TSFT is topologically conjugated to a Markov tree-shift. For the rest of this subsection, without loss of generality, we consider those Markov tree-shifts $X = \mathbf{X}_{\mathcal{F}}$ over symbol set \mathcal{A} such that every symbol is essential. Proposition 2.7 indicates that $h(X) = \ln d$.

This subsection investigates the entropy minimality problem described as follows. Let $Y = \mathbf{X}_{\mathcal{F}'}$ be a proper subspace of X such that

- (H1) $\mathcal{F} \subsetneq \mathcal{F}'$ and $\mathcal{F}' \setminus \mathcal{F}$ consists of only one pattern;
 (H2) if $\mathcal{A}' \subsetneq \mathcal{A}$, then Y is not a TSFT over \mathcal{A}' .

In other words, the forbidden set of Y is obtained by adding a pattern to the forbidden set of X , and every symbol which is seen in X remains to be used in Y .

Problem 2.8. Under the above conditions, what can we say if $h(Y) < h(X)$?

Definition 2.9. Suppose $X = \mathbf{X}_{\mathcal{F}}$ is a TSFT over \mathcal{A} . A symbol $a \in \mathcal{A}$ is called a *saving symbol* for X if, for each pattern $(\alpha, \alpha_1, \alpha_2, \dots, \alpha_d) \notin \mathcal{F}$ such that $\alpha \neq a$, there exists $1 \leq i \leq d$ such that $\alpha_i = a$.

Proposition 2.10. Suppose $X = \mathbf{X}_{\mathcal{F}}$ is a TSFT over \mathcal{A} with a saving symbol a . If $Y = \mathbf{X}_{\mathcal{F}'}$ is a proper subspace of X satisfying (H1) and (H2), then $h(Y) < h(X)$ if and only if a is an inessential symbol for Y .

Remark 2.11. Proposition 2.10 can be rephrased as follows. $h(Y) < h(X)$ if and only if there is exact two patterns in \mathcal{F} which start with a , $\mathcal{F}' = \mathcal{F} \cup \{(a, a_1, \dots, a_d)\}$ with $a_i \neq a$ for some i , and $(a, a, \dots, a) \notin \mathcal{F}'$. In other words, we can only remove the pattern (of height 2) that starts with a saving symbol and make it an inessential saving symbol.

Proof of Proposition 2.10. Suppose that a is an inessential saving symbol for Y . It follows immediately that $(a, \alpha_1, \dots, \alpha_d) \in \mathcal{F}'$ if and only if $\alpha_i \neq a$ for some $1 \leq i \leq d$. Let \hat{F} be the corresponding SNRE of Y and let E be a reduced SNRE of \hat{F} . Remark 3.2 and Theorem 3.3 infers that $h(E) = \ln \rho_A$, where ρ_A is the spectral radius of A and A is the $(k-1) \times (k-1)$ matrix obtained by deleting the row and column indexed by a . Since a is a saving symbol, $\sum_{j=1}^{k-1} A(i, j) \leq d-1$ for $1 \leq i \leq k-1$. This demonstrates that $\rho_A \leq d-1$. Hence, $h(Y) \leq \ln(d-1) < h(X)$.

Conversely, $h(Y) < h(X)$ and Proposition 2.7 assert that there is a symbol $s \in \mathcal{A}$ such that s is inessential for Y . We claim that there are exactly two patterns of height 2 which start with s and are accessible in X . Indeed, the assumptions (H1) and (H2) infer that there are at least two accessible patterns (of height 2) in X which start with s . Furthermore, s is inessential for Y derives that there are at most two accessible patterns in X which start with s . The Claim then follows.

Suppose that $(s, \alpha_1, \dots, \alpha_d), (s, \beta_1, \dots, \beta_d) \notin \mathcal{F}$ and $(s, \alpha_1, \dots, \alpha_d) \in \mathcal{F}'$. Since s is an inessential symbol for Y , it is seen that $\beta_i = s$ for $1 \leq i \leq d$.

If $s \neq a$, then a being a saving symbol concludes that $\beta_i = a$ for some $1 \leq i \leq d$. The essentiality of a infers that s is essential for Y , which gets a contradiction. The proof is then complete. \square

3. PROOFS OF MAIN RESULTS

This section is dedicated to the proofs of Theorems 1.1 and 1.2. Some useful results are presented herein. Proposition 3.1 is a useful tool to compute $h(F)$.

3.1. Weighted adjacency matrix and its spectral radius.

Proposition 3.1. *Let F be an SNRE, then*

$$(10) \quad h(F) = \lim_{n \rightarrow \infty} \frac{\ln \sum_{i=1}^k \ln a_n^{(i)}}{n}.$$

Proof. Since for every $n, k \in \mathbb{N}$,

$$a_n^{(1)} a_n^{(2)} \cdots a_n^{(k)} \leq \left(\frac{\sum_{i=1}^k a_n^{(i)}}{k} \right)^k,$$

we derive that

$$\sum_{i=1}^k \ln a_n^{(i)} \leq k \left(\ln \sum_{i=1}^k a_n^{(i)} - \ln k \right)$$

and

$$\lim_{n \rightarrow \infty} \frac{\ln \sum_{i=1}^k \ln a_n^{(i)}}{n} \leq \lim_{n \rightarrow \infty} \frac{\ln^2 \sum_{i=1}^k a_n^{(i)}}{n} = h(F).$$

Conversely, let $a_n = \max_{1 \leq i \leq k} a_n^{(i)}$. The inequality

$$a_n \leq \sum_{i=1}^k a_n^{(i)} \leq k a_n,$$

yields that

$$(11) \quad \lim_{n \rightarrow \infty} \frac{\ln^2 \sum_{i=1}^k a_n^{(i)}}{n} = \lim_{n \rightarrow \infty} \frac{\ln^2 a_n}{n}.$$

On the other hand, we have

$$(12) \quad \sum_{i=1}^k \ln a_n^{(i)} = \ln \prod_{i=1}^k a_n^{(i)} \geq \ln \max_{1 \leq i \leq k} a_n^{(i)} = \ln a_n.$$

Combining (11) with (12) concludes that

$$\lim_{n \rightarrow \infty} \frac{\ln \sum_{i=1}^k \ln a_n^{(i)}}{n} \geq \lim_{n \rightarrow \infty} \frac{\ln^2 a_n}{n} = \lim_{n \rightarrow \infty} \frac{\ln^2 \sum_{i=1}^k a_n^{(i)}}{n} = h(F).$$

The proof is thus complete. \square

Remark 3.2. Suppose we partition the symbol set \mathcal{A} as

$$(13) \quad \mathcal{A} = \mathcal{A}_E \cup \mathcal{A}_I,$$

where \mathcal{A}_E is the collection of the essential symbols in \mathcal{A} and \mathcal{A}_I collects the inessential symbols (Definition 2.3), it follows from Proposition 3.1 that

$$h(F) = \lim_{n \rightarrow \infty} \frac{\ln \sum_{i=1}^k \ln a_n^{(i)}}{n} = \lim_{n \rightarrow \infty} \frac{\ln \sum_{a^{(i)} \in \mathcal{A}_E} a_n^{(i)}}{n}.$$

That is, the entropy $h(F)$ is the growth rate of the sum of all essential symbols. In this case, we say that $h(F)$ is *supported* on \mathcal{A}_E .

Let Ω be a one-dimensional subshift of finite type and let $A = A_\Omega$ be the corresponding adjacency matrix, the classical result in symbolic dynamics shows that the topological entropy of Ω is $h(\Omega) = \ln \lambda_A$, where λ_A is the maximal eigenvalue of A (cf. [21]). Theorem 3.3 is an analogous result for reduced SNREs.

Theorem 3.3. *Let $F = \{F^{(i)}\}_{i=1}^k$ be a reduced SNRE and let M be the corresponding weighted adjacency matrix which is defined in (9). If there exists $N \in \mathbb{N}$ such that $a_n^{(i)} > 1$ for all $i = 1, \dots, k$ and $n \geq N$, then*

$$h(F) = \ln \lambda_M,$$

where λ_M is the spectral radius of M .

Proof. Let F be a reduced SNRE. That is, F is defined by an indicator matrix I_F which satisfies the condition (ii) of Definition 2.6. We write the SNRE F in the following form.

$$F = \{F^{(i)} = \left(a_{n-1}^{(1)}\right)^{m_1^{(i)}} \left(a_{n-1}^{(2)}\right)^{m_2^{(i)}} \cdots \left(a_{n-1}^{(k)}\right)^{m_k^{(i)}}\},$$

where $(m_1^{(i)}, \dots, m_k^{(i)})$ is a non-negative integral k -tuple for all i . Define

$$b_n := (\ln a_n^{(1)}, \dots, \ln a_n^{(k)})^T.$$

It is seen that $b_n = M b_{n-1}$. Combining the facts of $b_n = M^{n-1} b_1$, (10), and $a_n^{(i)} > 1$ for n large enough yields that

$$h(F) = \lim_{n \rightarrow \infty} \frac{\ln \sum_{i=1}^k \ln a_n^{(i)}}{n} = \lim_{n \rightarrow \infty} \frac{\ln \sum_{i,j=1}^k M^{n-1}(i, j)}{n} = \ln \lambda_M.$$

This completes the proof. \square

3.2. Proof of Theorem 1.2. Theorem 3.3 reveals that the computation of the entropy of a reduced SNRE is analogous to the classical result of SFTs. Theorem 1.2 provides the method for the computation of $h(F)$ for general F ; that is, the entropy of $h(X)$ (Theorem 1.1). The proof of Theorem 1.2 is presented herein.

Proof of Theorem 1.2. Set

$$h = \max\{\ln \lambda_{M_E} : E \text{ is reduced from } F\}.$$

Let $b_n^{(i)} = \ln a_n^{(i)}$ for all $1 \leq i \leq k$ and $F^{(i)}$ be arranged as the following form.

$$(14) \quad F^{(i)} = \sum_{j=1}^{r_i} \beta_j^{(i)} F_j^{(i)}, \quad \beta_j^{(i)} \neq 0, i = 1, \dots, k.$$

Since the computation of $h(F)$ is supported on those essential symbols (see Remark 3.2), without loss of generality, we assume that $b_1^{(i)} \geq 2$ for $i = 1, \dots, k$. The existence of the limit of $h(X) = h(F)$ infers that there is a subsequence $\{a_{n_\ell}^{(j_i)}\}_{i=1}^k$ satisfying

$$(15) \quad a_{n_\ell}^{(j_1)} \geq a_{n_\ell}^{(j_2)} \geq \dots \geq a_{n_\ell}^{(j_k)} \quad \text{for } \ell \in \mathbb{N}$$

and

$$h(F) = \lim_{n_\ell \rightarrow \infty} \frac{\ln \sum_{i=1}^k \ln a_{n_\ell}^{(j_i)}}{n_\ell}.$$

For simplicity, we may assume that $j_i = i$ for $i = 1, \dots, k$ and $n_\ell = \ell$ for $\ell \in \mathbb{N}$. Thus, (14) can be rewritten as follows.

$$F^{(i)} = F_1^{(i)} \left(\beta_1^{(i)} + \sum_{l=2}^{r_i} \beta_l^{(i)} \frac{F_l^{(i)}}{F_1^{(i)}} \right) = F_1^{(i)} c_{n-1}^{(i)},$$

where

$$c_{n-1}^{(i)} = \beta_1^{(i)} + \sum_{l=2}^{r_i} \beta_l^{(i)} \frac{F_l^{(i)}}{F_1^{(i)}}.$$

It follows from (15) that

$$(16) \quad 2 \leq c_{n-1}^{(i)} \leq \sum_{l=1}^{r_i} \beta_l^{(i)} \leq C \text{ for all } n \in \mathbb{N},$$

where

$$C = r\beta, r := \max_{1 \leq i \leq k} r_i \text{ and } \beta := \max_{1 \leq i \leq k, 1 \leq l \leq r_i} \beta_l^{(i)}.$$

That is,

$$(17) \quad \begin{cases} b_n^{(1)} = \ln a_n^{(1)} = \ln F_1^{(1)} + \ln c_{n-1}^{(1)}, \\ b_n^{(2)} = \ln a_n^{(2)} = \ln F_1^{(2)} + \ln c_{n-1}^{(2)}, \\ \vdots \\ b_n^{(k)} = \ln a_n^{(k)} = \ln F_1^{(k)} + \ln c_{n-1}^{(k)}. \end{cases}$$

Let $b_n = (b_n^{(1)}, \dots, b_n^{(k)})^T$. Notably, for $i = 1, \dots, k$, $F_1^{(i)}$ is of the form

$$F_1^{(i)} = (a_{n-1}^{(1)})^{m_1^{(i)}} \cdots (a_{n-1}^{(k)})^{m_k^{(i)}};$$

we have

$$\ln F_1^{(i)} = \sum_{l=1}^k m_l^{(i)} \ln a_{n-1}^{(l)} = \sum_{l=1}^k m_l^{(i)} b_{n-1}^{(l)}.$$

Thus (17) can be represented as

$$(18) \quad b_n = M b_{n-1} + \ln c_{n-1},$$

where

$$\ln c_n = (\ln c_n^{(1)}, \dots, \ln c_n^{(k)})^T,$$

and M is the weighted adjacency matrix of $E = \{F_1^{(i)}\}_{i=1}^k$. Iterate (18) we obtain

$$(19) \quad b_n = M^{n-1} b_1 + M^{n-2} c_1 + \cdots + c_{n-1}.$$

Let $\lambda = \lambda_M$ it follows from Proposition 4.2.1 of [21] and (16) that

$$\begin{aligned} \sum_{i=1}^k b_n^{(i)} &= \|b_n\| \leq d_0 \left\| \sum_{i=0}^{n-1} M^i \right\| \leq d_0 \sum_{i=0}^{n-1} \|M^i\| \\ &\leq d_1 \left(\sum_{i=0}^{n-1} \lambda^i \right) \leq d_2 \lambda^n, \end{aligned}$$

where d_0 , d_1 , and d_2 only depend on the dimension of M and k . Thus we have

$$(20) \quad \sum_{i=1}^k b_n^{(i)} \leq d_2 \lambda^n.$$

Combining Proposition 3.1 with (20) infers that

$$h(F) = \lim_{n \rightarrow \infty} \frac{\ln \sum_{i=1}^k b_n^{(i)}}{n} \leq \lim_{n \rightarrow \infty} \frac{\ln d_2 \lambda^n}{n} = \ln \lambda.$$

Similarly, combining (16), (18) with the fact that $b_1^{(i)} \geq 2$ for all i we have

$$\sum_{i=1}^k b_n^{(i)} \geq d_3 \lambda^n,$$

for some $d_3 > 0$, which implies

$$h(F) \geq \ln \lambda.$$

Thus $h(F) = \ln \lambda$. That is $h(F)$ is the logarithm of the spectral radius of some integral matrix M which is a weighted adjacency matrix of some SNRE E reduced from F . Thus we conclude that

$$(21) \quad h(F) \leq h.$$

For the converse, suppose E is a reduced SNRE of F with $h = \ln \lambda_{M_E}$. From the (i) of Definition 2.6 we have $E^{(i)} \leq F^{(i)}$ for all i . It implies that

$$(22) \quad h(F) \geq h(E) = \ln \lambda_{M_E} = h.$$

Combining (21) with (22) yields (2). The proof is thus completed. \square

3.3. Proof of Theorem 1.1. The proof of Theorem 1.1 is presented.

Proof of Theorem 1.1. Let X be a TSFT and let F be its SNRE. Since $h(X^B) = h(F)$ from Theorem 2.4, thus it follows from Theorem 1.2 that

$$h(X) = h(F) = h(E) = \ln \lambda_{M_E} =: \ln \lambda_E$$

for some reduced SNRE E of F . If the weighted adjacency matrix M_E is primitive, the Perron-Frobenius theorem concludes that $\lambda_E \in \mathcal{P}$. That is, $\ln \lambda_E \in \mathcal{E}$ (recall (1)). If M_E is irreducible or reducible, Corollary of Theorem 3 of [20] shows that that $\lambda_E = \lambda^{\frac{1}{p}}$ for some $\lambda \in \mathcal{P}$ and $1 < p \in \mathbb{N}$. Thus, $\ln \lambda_E \in \mathcal{E}$. Conversely, we set $h = \ln \lambda^{\frac{1}{p}} \in \mathcal{E}$. If $p = 1$, then Theorem 1 of [20] is applied to show that there is a primitive non-negative integral matrix with λ is its spectral radius. More precisely, there is an $m \times m$ primitive non-negative integral matrix M such that λ is its spectral radius. Construct a new matrix $V = V_M$ as follows. Denote by

$$(23) \quad d := d(M) = \max_{1 \leq i \leq k} \sum_{j=1}^k M(i, j).$$

Define a $(k+1) \times (k+1)$ non-negative integral matrix V as follows.

$$V(i, j) = \begin{cases} M(i, j), & \text{if } 1 \leq i \leq k, 1 \leq j \leq k; \\ d - \sum_{j=1}^k M(i, j), & \text{if } 1 \leq i \leq k, j = k+1; \\ 0, & \text{if } i = k+1, 1 \leq j \leq k; \\ d, & \text{if } i = k+1, j = k+1. \end{cases}$$

It can be easily checked that

$$\sum_{j=1}^{k+1} V(i, j) = d \text{ for } 1 \leq i \leq k+1.$$

Introduce the symbol set \mathcal{A} and an SNRE $F = \{F^{(i)}\}_{i=1}^{k+1}$ according to V as follows. Let

$$\mathcal{A} = \{a^{(i)}\}_{i=1}^k \cup \{a^{(k+1)}\}.$$

For $i = 1, \dots, k$, we define

$$\begin{aligned} F^{(i)} &= \left(a_{n-1}^{(1)}\right)^{V(i,1)} \cdots \left(a_{n-1}^{(k)}\right)^{V(i,k)} \left(a_{n-1}^{(k+1)}\right)^{d-V(i,k+1)} + \left(a_{n-1}^{(k+1)}\right)^d, \\ F^{(k+1)} &= \left(a_{n-1}^{(k+1)}\right)^d. \end{aligned}$$

Since $\sum_{j=1}^k V(i, j) = d$ for all i , each $F^{(i)}$ is a polynomial of degree d . It implies that $F = \{F^{(i)}\}_{i=1}^{k+1}$ is a $(d, k+1)$ -SNRE with the initial conditions of F is

$$(24) \quad a_1^{(i)} = 2 \text{ for } 1 \leq i \leq k \text{ and } a_1^{(k+1)} = 1.$$

Combining (24) with the fact that $a_n^{(k+1)} = 1$ for all n (since $a_n^{(k+1)}$ only connect to itself), we conclude that $a_n^{(k+1)}$ must be the least element with respect to the lexicographic order defined in the proof of Theorem 1.2 (since $a_1^{(i)} \geq 2 \geq a_1^{(1)}$ for $i = 1, \dots, k$ and $a_n^{(k+1)} = 1$). Therefore, we obtain that the entropy $h(F)$ is attained at the logarithm of the spectral radius of the weighted adjacency matrix M corresponding to a reduced SNRE E of F , where $E = \{E^{(i)}\}_{i=1}^{k+1}$ is as follows.

$$\begin{aligned} E^{(i)} &= \left(a_{n-1}^{(1)}\right)^{V(i,1)} \cdots \left(a_{n-1}^{(k)}\right)^{V(i,k)} \left(a_{n-1}^{(k+1)}\right)^{d-V(i,k+1)}, \quad 1 \leq i \leq k, \\ E^{(k+1)} &= \left(a_{n-1}^{(k+1)}\right)^d. \end{aligned}$$

Meanwhile, M is of the form

$$M_E = \begin{pmatrix} M & U \\ 0 & d \end{pmatrix},$$

where U is a $k \times 1$ matrix with entries are $U(i, k+1) = d - V(i, k+1)$ for $1 \leq i \leq k$.

Let

$$\mathcal{B} = \left\{ \left(a^{(i)}, \overbrace{a^{(1)}, \dots, a^{(1)}}^{V(i,1)\text{-times}}, \dots, \overbrace{a^{(k+1)}, \dots, a^{(k+1)}}^{d-V(i,k+1)\text{-times}} \right) \right\}_{i=1}^k \cup \left\{ \left(a^{(i)}, \overbrace{a^{(k+1)}, \dots, a^{(k+1)}}^{d\text{-times}} \right) \right\}_{i=1}^k \cup \left\{ a^{(k+1)}, \overbrace{a^{(k+1)}, \dots, a^{(k+1)}}^{d\text{-times}} \right\}$$

and let $\mathcal{F} = \mathcal{A}^d \setminus \mathcal{B}$. We claim that the TSFT $X = X_{\mathcal{F}}$ carries entropy $h(X) = \ln \lambda$. Indeed, it follows from Proposition 3.1, Theorems 1.2 and 2.4, and the fact of $a_n^{(k+1)} = 1$ for all n that

$$h(X) = h(F) = \lim_{n \rightarrow \infty} \frac{\ln \sum_{i=1}^k \ln a_n^{(i)}}{n} = h(E) = \ln \lambda = h.$$

This shows that h is a entropy of some tree SFT X . Hence, the claim holds. Finally, if $p > 1$, i.e., $h = \ln \lambda^{\frac{1}{p}} \in \mathcal{E}$, we first note that the above argument of the constructing the TSFT X is also applied to the case where M is a non-negative irreducible integral matrix. Theorem 3 of [20] says that a positive number is the spectral radius of an irreducible non-negative integral matrix if and only if some positive integral power of it is a Perron number. Since $\left(\lambda^{\frac{1}{p}}\right)^p = \lambda \in \mathcal{P}$, $\lambda^{\frac{1}{p}}$ must be a spectral radius of some non-negative irreducible integral matrix, say M . Using the same argument as above one could construct a TSFT X such that $h(X) = h(F) = \ln \lambda$. The proof is thus complete. \square

3.4. Example.

Example 3.4. Let $h = \ln \lambda$, where λ is the spectral radius of

$$M = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & 1 & 0 \end{pmatrix} \in \mathbf{M}_{3 \times 3}.$$

Since $d = 3$ (defined in (23)), we construct a 4×4 matrix V as

$$V = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 2 \\ 2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix} = \begin{pmatrix} M & U \\ 0 & 3 \end{pmatrix}.$$

The $(3, 4)$ -SNRE can be constructed as follows.

$$\begin{cases} a_n^{(1)} = a_{n-1}^{(1)} a_{n-1}^{(2)} a_{n-1}^{(4)} + \left(a_{n-1}^{(4)}\right)^3, a_2^{(1)} = 2 \\ a_n^{(2)} = a_{n-1}^{(3)} \left(a_{n-1}^{(4)}\right)^2 + \left(a_{n-1}^{(4)}\right)^3, a_2^{(2)} = 2 \\ a_n^{(3)} = \left(a_{n-1}^{(1)}\right)^2 a_{n-1}^{(2)} + \left(a_{n-1}^{(4)}\right)^3, a_2^{(3)} = 2 \\ a_n^{(4)} = \left(a_{n-1}^{(4)}\right)^3, a_2^{(4)} = 1. \end{cases}$$

Meanwhile, the corresponding set \mathcal{B} can also be defined as

$$\mathcal{B} = \left\{ (a^{(1)}, a^{(1)}, a^{(2)}, a^{(4)}), (a^{(1)}, a^{(4)}, a^{(4)}, a^{(4)}), (a^{(2)}, a^{(3)}, a^{(4)}, a^{(4)}), \right. \\ (a^{(2)}, a^{(4)}, a^{(4)}, a^{(4)}), (a^{(3)}, a^{(1)}, a^{(1)}, a^{(2)}), (a^{(3)}, a^{(4)}, a^{(4)}, a^{(4)}), \\ \left. (a^{(4)}, a^{(4)}, a^{(4)}, a^{(4)}) \right\}.$$

The same argument as seen in the proof of Theorem 1.1 shows that the TSFT $X = \mathbf{X}_{\mathcal{F}}$ with $\mathcal{F} = \mathcal{A}^4 \setminus \mathcal{B}$ is capable of the entropy $\ln \lambda$.

ACKNOWLEDGMENT

The author would like to thank Ms. Nai-Zhu Huang for the discussion during the preparation of this work.

REFERENCES

1. R. L. Adler, *Similarity of automorphisms of the torus*, vol. 98, American Mathematical Soc., 1970.
2. K. T. Alligood, T. D. Sauer, and J. A. Yorke, *Chaos*, Springer, 1997.
3. N. Aubrun and M.-P. Béal, *Tree-shifts of finite type*, Theor. Comput. Sci. **459** (2012), 16–25.
4. ———, *Sofic tree-shifts*, Theory Comput. Systems **53** (2013), 621–644.
5. J.-C. Ban and C.-H. Chang, *Tree-shifts: Entropy of tree-shifts of finite type*, arXiv:1509.08325, 2015.
6. ———, *Tree-shifts: Irreducibility, mixing, and chaos of tree-shifts*, Trans. Am. Math. Soc. (2016), accepted.
7. J.-C. Ban and S.-S. Lin, *Patterns generation and transition matrices in multi-dimensional lattice models*, Discrete Contin. Dyn. Syst. **13** (2005), 637–658.
8. R. Bowen, *Markov partitions for Axiom A diffeomorphisms*, Amer. J. Math. **92** (1970), no. 3, 725–747.

9. M. Boyle, R. Pavlov, and M. Schraudner, *Multidimensional sofic shifts without separation and their factors*, Trans. Am. Math. Soc. **362** (2010), 4617–4653.
10. K. Falconer, *The Hausdorff dimension of self-affine fractals*, Math. Proc. Cambridge Philos. Soc. **103** (1988), 339–350.
11. ———, *Fractal geometry: Mathematical foundations and applications.*, 2nd ed., John Wiley & Sons, New York-London-Sydney, 2003.
12. K. Falconer and J. Miao, *Dimensions of self-affine fractals and multifractals generated by upper-triangular matrices*, Fractals **15** (2007), 289–299.
13. D.-J. Feng and A. Käenmäki, *Equilibrium states of the pressure function for products of matrices*, Discrete Contin. Dyn. Syst. **30** (2011), 699–708.
14. S. Friedland, *Multi-dimensional capacity, pressure and Hausdorff dimension*, Mathematical Systems Theory in Biology, Communications, Computation, and Finance, Springer, 2003, pp. 183–222.
15. G. A. Hedlund, *Endomorphisms and automorphisms of full shift dynamical system*, Math. Systems Theory **3** (1969), 320–375.
16. M. Hochman, *On dynamics and recursive properties of multidimensional symbolic dynamics*, Invent. Math. **176** (2009), 131–167.
17. M. Hochman and T. Meyerovitch, *A characterization of the entropies of multidimensional shifts of finite type*, Ann. of Math. **171** (2010), 2011–2038.
18. A. Käenmäki and B. Li, *Genericity of dimension drop on self-affine sets*, arXiv:1611.00496, 2017.
19. A. Käenmäki and I. D. Morris, *Equilibrium states for singular value function and strict monotonicity of affinity dimension*, arXiv:1609.07360, 2016.
20. D. Lind, *The entropies of topological Markov shifts and a related class of algebraic integers*, Ergodic Theory Dynam. Systems **4** (1984), no. 02, 283–300.
21. D. Lind and B. Marcus, *An introduction to symbolic dynamics and coding*, Cambridge University Press, Cambridge, 1995.
22. B. Marcus and R. Pavlov, *Approximating entropy for a class of \mathbb{Z}^2 Markov random fields and pressure for a class of functions on \mathbb{Z}^2 shifts of finite type*, Ergodic Theory Dynam. Systems **33** (2013), no. 01, 186–220.

23. ———, *Computing bounds for entropy of stationary \mathbb{Z}^d markov random fields*, SIAM J. Discrete Math. **27** (2013), no. 3, 1544–1558.
24. N.G. Markley and M.E. Paul, *Maximal measures and entropy for z^ν subshift of finite type*, Classical Mechanics and Dynamical Systems, Lecture Notes in Pure and Appl. Math., vol. 70, 1979, pp. 135–157.
25. R. Pavlov and M. Schraudner, *Classification of sofic projective subdynamics of multidimensional shifts of finite type*, Trans. Am. Math. Soc. **367** (2015), 3371–3421.
26. A. Quas and A. Şahin, *Entropy gaps and locally maximal entropy in \mathbb{Z}_d subshifts*, Ergodic Theory Dynam. Systems **23** (2003), 1227–1245.
27. H. Rao, H.-J. Ruan, and Y. Wang, *Lipschitz equivalence of Cantor sets and algebraic properties of contraction ratios*, Trans. Amer. Math. Soc. **364** (2012), no. 3, 1109–1126.
28. C. E. Shannon, *Communication theory of secrecy systems*, Bell System Technical Journal **28** (1949), no. 4, 656–715.

(Jung-Chao Ban) DEPARTMENT OF APPLIED MATHEMATICS, NATIONAL DONG HWA UNIVERSITY, HUALIEN 970003, TAIWAN, R.O.C.

E-mail address: jcban@gms.ndhu.edu.tw

(Chih-Hung Chang) DEPARTMENT OF APPLIED MATHEMATICS, NATIONAL UNIVERSITY OF KAOHSIUNG, KAOHSIUNG 81148, TAIWAN, R.O.C.

E-mail address: chchang@nuk.edu.tw