J. Korean Math. Soc. **55** (2018), No. 5, pp. 1157–1175
https://doi.org/10.4134/JKMS.j170620
pISSN: 0304-9914 / eISSN: 2234-3008

TOPOLOGICAL ENTROPY OF SWITCHED SYSTEMS

Yu Huang and Xingfu Zhong

Abstract. For a switched system with constraint on switching sequences, which is also called a subshift action, on a metric space not necessarily compact, two kinds of topological entropies, average topological entropy and maximal topological entropy, are introduced. Then we give some properties of those topological entropies and estimate the bounds of them for some special systems, such as subshift actions generated by finite smooth maps on p-dimensional Riemannian manifold and by a family of surjective endomorphisms on a compact metrizable group. In particular, for linear switched systems on $\mathbb{R}^p$, we obtain a better upper bound, by joint spectral radius, which is sharper than that by Wang et al. in [42,43].

1. Introduction

A switched system consists of a family of subsystems and a rule that governs the switching among them. More precisely, let X be a metric space not necessarily compact and $\mathcal{G} = \{f_1, \ldots, f_k\}$ a family of continuous self-maps of X, we consider the discrete-time dynamical system in the form of

$$x_{n+1} = f_{\omega_n}(x_n), \tag{1}$$

where $x_n \in X$, ω_n takes a value in the finite-symbolic set $\mathcal{I} \triangleq \{1, 2, \ldots, k\}$. Let $\mathbb{N}$ denote the non-negative integers. If we denote the set (also called symbolic space) of all mappings $\mathbb{N} \to \mathcal{I}$ by

$$\mathcal{I}^{\mathbb{N}} = \{\omega : \ \mathbb{N} \to \mathcal{I}\},$$

then switching can be classified into two situations: (i) arbitrary switching; i.e., the switching rule can be taken arbitrarily from $\mathcal{I}^{\mathbb{N}}$; (ii) switching is subject to certain constraints; i.e., the switching rule is characterized by a subset of $\mathcal{I}^{\mathbb{N}}$.

Received September 26, 2017; Accepted January 16, 2018.
2010 *Mathematics Subject Classification.* 37A35, 37B10, 37B40, 37B55.
Key words and phrases. maximal topological entropy, average topological entropy, switched system, joint spectral radius.
Project supported by National Nature Science Funds of China (11371380, 11771459 and 11471125), and International Program for Ph.D. Candidates, Sun Yat-Sen University, and project of distinctive innovation of ordinary university of Guangdong Province (No.2016KTSCX093).

Switched systems are found in many practical systems, see [10] and [40]. When every element f_i in $\mathcal{G}$ is a linear continuous map from a finite dimensional linear space into itself, we call (1) a switched linear system. Many control properties for switched linear systems, such as stability, stabilizability, controllability, observability and so on, have been widely studied in the last two decades. The books [10] and [40] contain many the basic theory, examples and many results on switched linear systems.

For a topological dynamical system (X, f) where X is a compact topological space and f is a continuous map from X into itself, Adler, Konheim and McAndrew [1] in 1965 introduced a quantity named topological entropy to measure the complexity for the system. Late, Dinaburg [15] and Bowen [7] respectively gave equivalent quantities when X is a metrizable space, which is known as Bowen-metric entropy. Since this entropy is invariant under topological conjugacy, many researchers extended this notion to other kinds of systems, such as non-autonomous dynamical systems, group and semigroup actions, foliations, graphs and so on [18,21,27,31–34,36,44]. Similar to topological entropy, measure theoretic entropy, which is first established by Kolmogorov [26] and Sinai [39], is another important notion to character the complexity of measure-preserving systems. Recently, there are also some works devoted to the measure theoretic entropy for non-autonomous dynamical systems [9,23,24,44].

Our aim is to study the topological entropy of the switched system (1). When the switchings are arbitrary, one can take the switched system (1) as a free semigroup action G generated by $\mathcal{G}$, i.e., $G = \bigcup_{n\in\mathbb{N}} G^n$, $G^n = \{f_{\omega_{n-1}} \cdots f_{\omega_0} \mid \omega_i \in \mathcal{I},\ i = 1, \dots, n\}$. We remark that $G^0 = \{id_X\}$, where id_X is the identity map of X. In this situation, Biś [5] in 2004 introduced the notion of topological entropy on the semigroup G in the case that the state space X is a compact metric space, which we call the maximal entropy of the semigroup G. Earlier, Bufetov in [8] gave another definition of topological entropy of G, which is called the average entropy of G. Recently, Wang et al. in [42] and [43], respectively, extend the notions of the maximal entropy and the average entropy of G to the case that the state space X is not necessarily compact. We will consider the maximal entropy and average entropy of the switched system (1) when the switchings are subject to a subset Λ of $\mathcal{I}^{\mathbb{N}}$ that can be identified as a subshift of the full shift $(\mathcal{I}^{\mathbb{N}}, \sigma)$, where σ is the classical shift operator on $\mathcal{I}^{\mathbb{N}}$, that is, for $\omega \in \mathcal{I}^{\mathbb{N}}$ $(\sigma\omega)_i = (\omega)_{i+1}$ and $\sigma(\Lambda) \subset \Lambda$. Let us call such system a subshift action on Λ. We remark that a proper subshift action has no semigroup structure. On the other hand, one can also view in this context the system as a skew-product transformation (or a cocycle) $F : \Lambda \times X \to \Lambda \times X$ which is defined by

$$F(\omega, x) = (\sigma(\omega), f_{\omega_0}(x)). \tag{2}$$

Thus, for the switched system (1), there are three kinds of different topological entropies: the maximal entropy, average entropy and the classical topological entropy of F.

The switched system (1), in some sense, can be regarded as a random system. Consider the following system:

$$x_{n+1} = f_{\overline{\sigma}^n \omega} x(n),$$

where $(\Lambda, \mathcal{B}, \mu, \overline{\sigma})$ is a measurable dynamical system, $\overline{\sigma} : \Lambda \to \{1, 2, \ldots, k\}$ is defined as $\overline{\sigma}^n(\omega) = \omega_n$ and $x(n) \in X$. The solution of the above equation can be written as $x_n = \phi(n, x_0, \omega)$ (see [2, 25] for more details about random dynamical systems). Entropies for random dynamical systems have been wildly studied. We refer the readers to Mihailescu and Urbanski [35], Bogenschütz [6], Dooley and Zhang [16], Huang [22], Gary [19].

In the present paper, we extend the notion of the maximal entropy and average entropy defined by Wang et al. in [42] and [43] to a subshift action on a metric space not necessarily compact. Then we give some properties of those entropies, including the topological analogue of the famous Abramov-Rokhlin formula and estimate the bounds of the entropies for some particular systems, such as subshift actions on a Riemanian manifold and switched linear systems with constraint. The results obtained extend the related results in [5], [8], [42] and [43]. In particular, for switched linear systems, we get a lower and upper bound of the maximal entropy via the joint spectral radius of $\mathcal{G}$, which is sharper than that obtained in [42]. It is well known that joint spectral radius play an important role in stability for switched linear systems.

This paper is organized as follows. In Section 2, we give the definitions of maximal entropy and average entropy of a subshift action on a metric space not necessarily compact and the relationship among them. Then some basic properties of the entropies are given in Section 3. In Section 4, we shall give some estimates of the entropies for three kinds of particular systems, including a subshift action on finite-dimensional Riemannian manifold, a subshift action on a compact metrizable group and a linear subshift action on finite-dimensional linear space, respectively. Finally in Section 5, we shall give an alternative lower bound for linear subshift actions with periodic points and a sharper upper bound, by joint spectral radius, for the full shift action on linear systems, which is sharper than that by Wang et al. in [43] and [42].

2. Topological entropy of a switched system

Let us recall some notions on symbolic spaces. Following [30], any finite set $\mathcal{I}$ with at least two elements is called an *alphabet*. And we write $\mathcal{I} = \{0, 1, \ldots, k-1\}$. For $n \in \mathbb{N}$ we denote by $\mathcal{I}^n$ the set of words of $\mathcal{I}$ of length n, i.e., $\mathcal{I}^n = \{u = (\omega_0 \cdots \omega_{n-1}) \mid \omega_i \in \mathcal{I},\ \ i = 0, \ldots, n-1\}$. Let $\mathcal{I}^* = \cup_{n \geq 1} \mathcal{I}^n$ be the set of all words of $\mathcal{I}$. For two words $u,\ v \in \mathcal{I}^*$, the concatenation of u and v is defined as uv. The word u is called a subword of v (denoted by $u \sqsubseteq v$) if there exists words x, y such that $xuy = v$. Furthermore, if x is the empty word, then u is called a prefix of v.

A *full shift* is any dynamical system $(\mathcal{I}^{\mathbb{N}}, \sigma)$, where $\mathcal{I}^{\mathbb{N}} = \{x = (x_0, x_1, \ldots) \mid x_n \in \mathcal{I},\ \forall n \geq 0\}$ is the symbol space endowed with the distance $d_1(x, y) =$

$\max_n\{\frac{1}{n+1} \mid x_n \neq y_n\}$ and σ is the shift map defined by $\sigma(x)_i = x_{i+1}$. A subshift is any subsystem of $(\mathcal{I}^{\mathbb{N}}, \sigma)$, i.e., a closed σ-invariant subset $\Lambda \subset \mathcal{I}^{\mathbb{N}}$.

By language we mean a subset $\mathcal{L} \subset \mathcal{I}^*$. Let Λ be a closed subset of $\mathcal{I}^{\mathbb{N}}$. The language and pre-language of Λ are defined by $\mathcal{L}(\Lambda) = \{u \in \mathcal{I}^* : \exists x \in \Lambda, u \sqsubseteq x\}$ and $\mathcal{L}_0(\Lambda) = \{u \in \mathcal{I}^* : \exists x \in \Lambda, u \text{ is a prefix of } x\}$, respectively. Note that $\mathcal{L}(\Lambda) = \mathcal{L}_0(\Lambda)$ when Λ is a closed σ-invariant subset of $\mathcal{I}^{\mathbb{N}}$. Thus a subshift can be characterized by its language. We denote by $\mathcal{L}^n(\Lambda) = \mathcal{L}_0(\Lambda) \cap \mathcal{I}^n$ the language of words of length n.

Let X be a metric space with metric d, which is not necessarily compact. Consider a family of finite continuous self-maps $\mathcal{G} = \{f_0, f_1, \ldots, f_{k-1}\}$ of X. Given a subshift $(\Lambda, \sigma_\Lambda)$, we consider the switched system (1) under the switching sequences subjected to Λ. This system can be characterized by the subshift action $G = \cup_n G_n$ on X where $G_n = \{f_\omega = f_{\omega_{n-1}} \circ \cdots \circ f_{\omega_0} \mid \omega = (\omega_{n-1} \cdots \omega_0) \in \mathcal{L}^n(\Lambda)\}$. For any n and $\omega \in G_n$, we denote $f_\omega^{-1} = (f_{w_{n-1}} \circ \cdots \circ f_{w_0})^{-1}$. We remark that the subshift G is not a free semigroup acting on X when Λ is a proper subset of $\mathcal{I}^{\mathbb{N}}$.

Now let us define two different entropies for the subshift G. For any $n \in \mathbb{N}$ and any $w \in \mathcal{L}^n(\Lambda)$, define two dynamical metrics d_n and d_w on X, respectively by

$$d_n(x, y) = \max_{0 \leq i < n} \{d(g(x), g(y)) : g \in G_i\}, \tag{3}$$

and

$$d_w(x, y) = \max_{w' \sqsubseteq w} d(f_{w'}(x), f_{w'}(y)). \tag{4}$$

Let K be a compact subset of X and $\varepsilon > 0$. A subset E of X is said to be an $(n, \varepsilon, K, \Lambda)$-spanning set of K if for every $x \in K$, there exists $y \in E$ such that $d_n(x, y) \leq \varepsilon$. A subset F of K is called $(n, \varepsilon, K, \Lambda)$-separated set of K if for any $x, y \in F$ with $x \neq y$, $d_n(x, y) > \varepsilon$. Let $r(n, \varepsilon, K, \Lambda)$ and $s(n, \varepsilon, K, \Lambda)$ denote the smallest cardinality of any $(n, \varepsilon, K, \Lambda)$-spanning set of K and the largest cardinality of any $(n, \varepsilon, K, \Lambda)$-separated set of K, respectively.

Likewise, for a compact subset K of X, $w \in \mathcal{L}^n(\Lambda)$ and $\varepsilon > 0$, we can define the $(w, \varepsilon, K, \Lambda)$-spanning set and $(w, \varepsilon, K, \Lambda)$-separated set of K with respect to the dynamical metric d_w. The smallest cardinality of any $(w, \varepsilon, K, \Lambda)$-spanning set of K and the largest cardinality of any $(w, \varepsilon, K, \Lambda)$-separated set of K are denoted by $N_{span}(w, \varepsilon, K, \Lambda)$ and $N_{sep}(w, \varepsilon, K, \Lambda)$, respectively. For every $n \in \mathbb{N}$, let

$$N_{span}(n, \varepsilon, K, \Lambda) = \frac{1}{|\mathcal{L}^n(\Lambda)|} \sum_{w \in \mathcal{L}^n(\Lambda)} N_{span}(w, \varepsilon, K, \Lambda),$$
$$N_{sep}(n, \varepsilon, K, \Lambda) = \frac{1}{|\mathcal{L}^n(\Lambda)|} \sum_{w \in \mathcal{L}^n(\Lambda)} N_{sep}(w, \varepsilon, K, \Lambda).$$

It is easy to see that

$$r(n,\varepsilon,K,\Lambda) \leq s(n,\varepsilon,K,\Lambda) \leq r(n,\frac{\varepsilon}{2},K,\Lambda)$$

and

$$N_{span}(n,\varepsilon,K,\Lambda) \leq N_{sep}(n,\varepsilon,K,\Lambda) \leq N_{span}(n,\frac{\varepsilon}{2},K,\Lambda).$$

Definition 2.1. Let X be a metric space with metric d and G be a subshift generated by a finite numbers of continuous self-maps $\mathcal{G} = \{f_0, f_1, \ldots, f_{m-1}\}$ on X and a subshift $(\Lambda, \sigma_\Lambda)$. For a compact subset K of X, define

$$h_M(K,\Lambda) = \lim_{\varepsilon\to 0} r(\varepsilon,K,\Lambda) = \lim_{\varepsilon\to 0} s(\varepsilon,K,\Lambda)$$

and

$$h_A(K,\Lambda) = \lim_{\varepsilon\to 0} N_{span}(\varepsilon,K,\Lambda) = \lim_{\varepsilon\to 0} N_{sep}(\varepsilon,K,\Lambda),$$

where

$$r(\varepsilon,K,\Lambda) = \limsup_{n\to\infty} \frac{1}{n} \log r(n,\varepsilon,K,\Lambda),$$

$$s(\varepsilon,K,\Lambda) = \limsup_{n\to\infty} \frac{1}{n} \log s(n,\varepsilon,K,\Lambda),$$

and

$$N_{span}(\varepsilon,K,\Lambda) = \limsup_{n\to\infty} \frac{1}{n} \log N_{span}(n,\varepsilon,K,\Lambda),$$

$$N_{sep}(\varepsilon,K,\Lambda) = \limsup_{n\to\infty} \frac{1}{n} \log N_{sep}(n,\varepsilon,K,\Lambda).$$

We call the quantity $h_M(\mathcal{G}|_\Lambda)$ defined by

$$h_M(\mathcal{G}|_\Lambda) = \sup\{h_M(K,\Lambda) : K \subset X \text{ is compact}\}$$

the *maximal entropy of subshift action* G *on* X and the quantity $h_A(\mathcal{G}|_\Lambda)$ defined by

$$h_A(\mathcal{G}|_\Lambda) = \sup\{h_A(K,\Lambda) : K \subset X \text{ is compact}\}$$

the *average entropy of subshift action* G on X.

When $\Lambda = \mathcal{I}^{\mathbb{N}}$, we write $h_M(\mathcal{G}) = h_M(\mathcal{G}|_\Lambda)$ and $h_A(\mathcal{G}) = h_A(\mathcal{G}|_\Lambda)$.

Remark 2.2. (1) If $\Lambda = \mathcal{I}^{\mathbb{N}}$ is the whole symbol space, then the definitions of the maximal entropy and average entropy are the same as the topological entropy defined by Wang et al. in [42], where $\mathcal{G} = \{id_X, f_1, \ldots, f_{m-1}\}$, and [43], respectively.

(2) It is easy to see that $h_M(\mathcal{G}) \geq h_M(\mathcal{G}|_\Lambda)$ and $h_A(\mathcal{G}) \geq h_A(\mathcal{G}|_\Lambda)$.

Remark 2.3. For any $\mathcal{G} = \{f_0, f_1, \ldots, f_{m-1}\}$, a finite numbers of continuous self-maps on X, and any subshift $(\Lambda, \sigma_\Lambda)$, it is obvious that $h_A(\mathcal{G}|_\Lambda) \leq h_M(\mathcal{G}|_\Lambda)$.

Moreover, we have:

Proposition 2.4. *There exists a system such that* $h_A(\mathcal{G}) = 0 < h_M(\mathcal{G}) = \log 2$.

Proof. Let $X = \mathbb{R}$ and $\mathcal{G} = \{f_0, f_1\}$ with $f_0 = 0$ the constant map and $f_1 = 2x$. We claim that

$$0 = h_A(\mathcal{G}) < h_M(\mathcal{G}).$$

It is clear that $h_M(\mathcal{G}) > 0$. In fact by Proposition 5.1, $h_M(\mathcal{G}) = \log 2 = h(f_1)$. So it suffices to show that $h_A(\mathcal{G}) = 0$. Let $[a, b] \subset \mathbb{R}$ be an interval. For any $\varepsilon > 0$, $n \geq 1$ and $j = 0, 1, \ldots, n$, $E_j = \{a + \frac{i\varepsilon(b-a)}{2^j},\ i = 0, 1, \ldots, \frac{2^j}{\varepsilon}\}$. Then $|E_j| \leq \lfloor \frac{2^j}{\varepsilon} \rfloor + 1 \leq \frac{2^{j+1}}{\varepsilon}$, where $\lfloor \frac{2^j}{\varepsilon} \rfloor$ denote the largest integer less than or equal to $\frac{2^j}{\varepsilon}$. If $|w| = n$ and $w_{[0,j]} = \overbrace{1 \cdots 1}^{j}0$, then $N_{span}(w, \varepsilon, [a, b], \mathcal{I}^{\mathbb{N}}) \leq |E_j| \leq \frac{2^{j+1}}{\varepsilon}$. So

$$\sum_{|w| = n,\, w_{[0,j]} = \overbrace{1\cdots1}^{j}0} N_{span}(w, \varepsilon, [a, b], \mathcal{I}^{\mathbb{N}}) \leq 2^{n-j+1} \frac{2^{j+1}}{\varepsilon} = \frac{2^n}{\varepsilon},$$

$$\begin{aligned}
\sum_{|w|=n} N_{span}(w, \varepsilon, [a, b], \mathcal{I}^{\mathbb{N}}) &\leq \sum_{j=0}^{n} \sum_{|w| = n,\, w_{[0,j]} = \overbrace{1\cdots1}^{j}0} N_{span}(w, \varepsilon, [a, b], \mathcal{I}^{\mathbb{N}}) + 2^n \\
&\leq \sum_{j=0}^{n} \frac{2^n}{\varepsilon} + 2^n = \frac{(n+2)2^n}{\varepsilon}.
\end{aligned}$$

Therefore

$$N_{span}(n, \varepsilon, [a, b], \mathcal{I}^{\mathbb{N}}) = \frac{1}{2^n} \sum_{|w|=n} N_{span}(w, \varepsilon, [a, b], \mathcal{I}^{\mathbb{N}}) \leq \frac{n+2}{\varepsilon}.$$

Thus

$$N_{span}(\varepsilon, [a, b], \mathcal{I}^{\mathbb{N}}) = 0$$

and

$$h_A([a, b], \mathcal{I}^{\mathbb{N}}) = 0,$$

which implies

$$h_A(\mathcal{G}) = 0. \qquad \square$$

Remark 2.5. From the above proof, we also have $h_M(\mathcal{G}|_\Lambda) = \log 2$ for any Λ with $1^\infty \in \Lambda$.

Let $X = \mathbb{R}$, $\mathcal{G}_1 = \{2x, \frac{1}{4}x\}$ and $\mathcal{G}_2 = \{4x, \frac{1}{2}x\}$. It is obvious that $\mathcal{G}_2$ generates the same semigroup as $\mathcal{G}_1$. By Proposition 5.1, we know $h_M(\mathcal{G}_1) = \log 2$ and $h_M(\mathcal{G}_2) = \log 4$. Suppose that G are generated by $\mathcal{G}_1$ and $\mathcal{G}_2$. It is interesting to know the relation between $h_M(\mathcal{G}_1)$ and $h_M(\mathcal{G}_2)$.

Proposition 2.6. *Let X be a metric space and $\mathcal{G}_1 = \{f_0, f_1, \ldots, f_m\}$ and $\mathcal{G}_2 = \{g_0, g_1, \ldots, g_n\}$, where f_i and g_k are continuous maps from X into itself, $i = 0, 1, \ldots, n$, $k = 0, 1, \ldots, m$. Suppose that $\mathcal{G}_1$ and $\mathcal{G}_2$ generate the same* *semigroup. If $f_i = g_{w_i}$ for some $w_i \in \{0,1,\ldots,n\}^{l_i}$, for any $i \in \{0,1,\ldots,m\}$, then*

$$h_M(\mathcal{G}_1) \le L_1 h_M(\mathcal{G}_2),$$

where L_1 is the lowest common multiple of $\{l_i\}_{i=0}^m$.

Proof. Let $\mathcal{I}_1 = \{0,1,\ldots,n\}$ and $\mathcal{I}_2 = \{0,1,\ldots,m\}$. For any compact set $K \subset X$. It is clear that $r(n,\varepsilon,K,{\mathcal{I}_1}^{\mathbb{N}}) \le r(L_1 n,\varepsilon,K,{\mathcal{I}_2}^{\mathbb{N}})$ and $r(n,\varepsilon,K,{\mathcal{I}_2}^{\mathbb{N}}) \le r(L_2 n,\varepsilon,K,{\mathcal{I}_1}^{\mathbb{N}})$. It follows that

$$\limsup_{n\to\infty} \frac{\log r(n,\varepsilon,K,{\mathcal{I}_1}^{\mathbb{N}})}{n} \le \limsup_{n\to\infty} \frac{\log r(L_1 n,\varepsilon,K,{\mathcal{I}_2}^{\mathbb{N}})}{n},$$
$$\limsup_{n\to\infty} \frac{\log r(n,\varepsilon,K,{\mathcal{I}_2}^{\mathbb{N}})}{n} \le \limsup_{n\to\infty} \frac{\log r(L_2 n,\varepsilon,K,{\mathcal{I}_1}^{\mathbb{N}})}{n}.$$

Hence

$$\limsup_{n\to\infty} \frac{\log r(n,\varepsilon,K,{\mathcal{I}_1}^{\mathbb{N}})}{n} \le L_1 \limsup_{n\to\infty} \frac{\log r(n,\varepsilon,K,{\mathcal{I}_2}^{\mathbb{N}})}{n},$$
$$\limsup_{n\to\infty} \frac{\log r(n,\varepsilon,K,{\mathcal{I}_2}^{\mathbb{N}})}{n} \le L_2 \limsup_{n\to\infty} \frac{\log r(n,\varepsilon,K,{\mathcal{I}_1}^{\mathbb{N}})}{n}.$$

It completes the proof. □

Remark 2.7. Let G be the semigroup generate by $\mathcal{G} = \{f_0,\ldots,f_n\}$. If $\hat{\mathcal{G}}$ generates the same semigroup, then $h_M(\mathcal{G}) > 0$ is equivalent to $h_M(\hat{\mathcal{G}}) > 0$. Let $\mathscr{G} = \{\hat{\mathcal{G}} : \hat{\mathcal{G}}$ is a generator of the semigroup$\}$. Suppose $\mathcal{G} \subset \hat{\mathcal{G}}$ for any $\hat{\mathcal{G}} \in \mathscr{G}$. Then $h_M(\mathcal{G}) = \inf_{\hat{\mathcal{G}}} h_M(\hat{\mathcal{G}})$.

The subshift action G can be also viewed as a skew-product transformation (or a cocycle) $F : \Lambda \times X \to \Lambda \times X$ which is defined by

$$F(\omega, x) = (\sigma(\omega), f_{\omega_0}(x)).$$

See (2). The relation between the classical topological entropy of F and the average entropy of the subshift action G is stated as follows, which is the topological analogue of the famous Abramov-Rokhlin formula.

Theorem 2.8. *Let (X,d) be a metric space and $\mathcal{G} = \{f_0, f_1, \ldots, f_{m-1}\}$ be a family of finite number continuous self-maps of X. Then the classical topological entropy of the skew-product transformation F satisfies*

$$h_D(F) = h(\Lambda,\sigma) + h_A(\mathcal{G}|_\Lambda),$$

where the metric D on $\Lambda \times X$ is defined by

$$D((\omega,x),(\omega',x')) = \max\{d'(\omega,\omega'), d(x,x')\}.$$

Here $h(\Lambda,\sigma)$ is the classical topological entropy of (Λ,σ), d and d' are metrics on Λ and X, respectively.

Remark 2.9. Recall that $h(\Lambda,\sigma) = \lim_{n\to\infty} \log \frac{|\mathcal{L}^n(\Lambda)|}{n}$. See e.g. [30, Proposition 3.24].

For the proof of the theorem, we need the following two lemmas, which are as analogous as those of Wang, Ma and Lin [43].

Lemma 2.10. *For any compact subset E of X, $n \geq 1$, $0 \leq \varepsilon \leq \frac{1}{2}$, it holds that*

$$N_{sep}(n, \varepsilon, \Lambda \times E, F) \geq \sum_{w \in \mathcal{L}^n(\Lambda)} N_{sep}(w, \varepsilon, E, \Lambda).$$

Proof. Let $\mathcal{L}^n(\Lambda)$ contain N distinct words, say, $\mathcal{L}^n(\Lambda) = \{w_1, \dots, w_N\}$. For every $1 \leq i \leq N$, pick $\omega(i) \in \Lambda$ such that $\omega(i)|_{[0,n-1]} = w_i$. It is clear that for $0 < \varepsilon < \frac{1}{2}$, the subset $\{\omega(i) : i = 1, \dots, N\}$ is a $(n, \varepsilon, \Lambda, \sigma)$-separated set of Λ. Let $N_i = N_{sep}(w(i), \varepsilon, E, \Lambda)$ and $\{x_1^i, \dots, x_{N_i}^i\}$ a $(w(i), \varepsilon, E, \Lambda)$-separated set of Λ. Then the points

$$(\omega(i), x_j^i) \in \Lambda \times X, \ i = 1, \dots, N, \ j = 1, \dots, N_i$$

form a $(n, \varepsilon, \Lambda \times E, F)$-separated set of $\Lambda \times E$. So $N_{sep}(n, \varepsilon, \Lambda \times E, F) \geq \sum_{w \in \mathcal{L}^n(\Lambda)} N_{sep}(w, \varepsilon, E, \Lambda)$. □

Lemma 2.11. *For any compact subset E of X, $n \geq 1$ and $\varepsilon > 0$, there exists a positive integer $C(\varepsilon)$ such that*

$$N_{span}(n, \varepsilon, \Lambda \times E, F) \leq |\mathcal{L}^{C(\varepsilon)}(\Lambda)| \sum_{w \in \mathcal{L}^n(\Lambda)} N_{span}(w, \varepsilon, E, \Lambda).$$

Proof. Given $\varepsilon > 0$, pick $C(\varepsilon)$ such that $\frac{1}{C(\varepsilon)+1} < \varepsilon$. Let $N = |\mathcal{L}^{n+C(\varepsilon)}(\Lambda)|$, then there are N distinct words of length $n + C(\varepsilon)$ in $\mathcal{L}^{n+C(\varepsilon)}$, which are denoted by $w_1, \dots, w_N$. For every $1 \leq i \leq N$, choose $\omega(i) \in \Lambda$ satisfying $\omega(i)|_{[0,n+C(\varepsilon)]} = w_i$. It is clear that $\{\omega(i) : i = 1, \dots, N\}$ is a (n, ε, σ)-spanning set of Λ. Let $\omega(i)|_{[0,n-1]} = w_i'$, $B_i = N_{span}(w_i, \varepsilon, E, \Lambda)$ and $\{x_1^i, \dots, x_{B_i}^i\}$ be a $(w_i, \varepsilon, E, \Lambda)$-spanning set of E. Then the set

$$\{(\omega(i), x_j^i) \in \Lambda \times X, i = 1, \dots, N, j = 1, \dots, B_i\}$$

is a $(n, \varepsilon, \Lambda \times E, F)$-spanning set of $\Lambda \times E$. Hence

$$N_{span}(n, \varepsilon, \Lambda \times E, F) \leq |\mathcal{L}^{C(\varepsilon)}(\Lambda)| \sum_{w \in \mathcal{L}^n(\Lambda)} N_{span}(w, \varepsilon, E, \Lambda).$$

□

The proof of Theorem 2.8. By Lemma 2.10, for any compact subset E of X, we have

$$N_{sep}(n, \varepsilon, \Lambda \times E, F) \geq |\mathcal{L}^n(\Lambda)| N_{sep}(n, \varepsilon, E, \Lambda).$$

It follows that

$$h_D(F) \geq h_D(\Lambda \times E, F) \geq h(\Lambda, \sigma) + h_A(E, \Lambda).$$

So

$$h_D(F) \geq h(\Lambda, \sigma) + h_A(\mathcal{G}|_\Lambda).$$

On the other hand, from Lemma 2.11, one has

$$N_{span}(n, \varepsilon, \Lambda \times E, F) \leq |\mathcal{L}^{C(\varepsilon)}(\Lambda)| |\mathcal{L}^n(\Lambda)| N_{span}(n, \varepsilon, E, \Lambda).$$

Hence

$$h_D(\Lambda \times E, F) \leq h(\Lambda, \sigma) + h_A(E, \Lambda) \leq h(\Lambda, \sigma) + h_A(\mathcal{G}|_\Lambda).$$

Since any compact subset of $\Lambda \times X$ is a subset of $\Lambda \times E$ for some compact subset of $E \subset X$, we have

$$h_D(F) = \sup\{h_D(\Lambda \times E, F) : E \text{ is a compact subset of } X\}.$$

It follows that

$$h_D(F) \leq h(\Lambda, \sigma) + h_A(\mathcal{G}|_\Lambda).$$ □

Remark 2.12. (1) If (X, d) is a compact metric space, Bufetov proved in [8] that

$$h_D(F) = \log m + h_A(\mathcal{G}).$$

(2) If (X, d) is a metric space not necessarily compact, Wang et al. proved in [43] that

$$h_D(F) = \log m + h_A(\mathcal{G}).$$

3. Basic properties of the entropies for subshift actions

In this section, we will give some basic properties of $h_M(\mathcal{G}|_\Lambda)$ and $h_A(\mathcal{G}|_\Lambda)$ for a subshift action on a metric space X generated by $\mathcal{G} = \{f_0, f_1, \ldots, f_{m-1}\}$ and a subshift Λ.

Recall that two metrics d and d' on X are said to be uniformly equivalent if both $id_X : (X, d) \to (X, d')$ and $id_X : (X, d') \to (X, d)$ are uniformly continuous.

By the analogous methods as that in [42] and [43], we have the following two theorems.

Theorem 3.1. *Let (X, d) be a metric space and $\mathcal{G} = \{f_0, f_1, \ldots, f_{m-1}\}$ be a family of finite number continuous self-maps of X. If d and d' are uniformly equivalent, then*

$$h_{M\,d}(\mathcal{G}|_\Lambda) = h_{M\,d'}(\mathcal{G}|_\Lambda)$$

and

$$h_{A\,d}(\mathcal{G}|_\Lambda) = h_{A\,d'}(\mathcal{G}|_\Lambda).$$

Remark 3.2. If X is compact, then any two equivalent metrics are uniformly equivalent. Therefore, the maximal entropy and average entropy of a subshift action does not rely on the metric on X.

Theorem 3.3. *Let (X, d) be a metric space and $\mathcal{G} = \{f_0, f_1, \ldots, f_{m-1}\}$ be a family of finite number continuous self-maps of X. Let $\delta > 0$. Then*

$$h_M(\mathcal{G}|_\Lambda) = \sup\{h_M(K, \Lambda), diam(K) < \delta\}$$

and

$$h_A(\mathcal{G}|_\Lambda) = \sup\{h_A(K, \Lambda), diam(K) < \delta\}.$$

The product rule for the two kinds of topological entropies is given as follows.

Theorem 3.4. *Let (X_i, d_i), $i = 1, 2$ be two metric spaces, $\mathcal{F}^{(1)} = \{f_0^{(1)}, \ldots, f_{m-1}^{(1)}\}$ a set of finite continuous self-maps of X_1 and $\mathcal{F}^{(2)} = \{f_0^{(2)}, \ldots, f_{k-1}^{(2)}\}$ a set of finite continuous self-maps of X_2. Define $\mathcal{F}^{(1)} \times \mathcal{F}^{(2)} = \{f_i \times g_j : i \in \{0, \ldots, m-1\}, j \in \{0, \ldots, k-1\}\}$, where $(f_i \times g_j)(x_1, x_2) = (f_i(x_1), g_j(x_2))$ for every $x_1 \in X_1$, $x_2 \in X_2$. A metric d on $X_1 \times X_2$ is defined by $d((x_1, x_2), (y_1, y_2)) = \max\{d_1(x_1, y_1), d_2(x_2, y_2)\}$. Given two subshifts Λ_i, $i = 1, 2$. Then*

$$h_{Md}(\mathcal{F}_1 \times \mathcal{F}_2|_{\Lambda_1 \times \Lambda_2}) \leq h_{Md_1}(\mathcal{F}_1|_{\Lambda_1}) + h_{Md_2}(\mathcal{F}_2|_{\Lambda_2})$$

and

$$h_{Ad}(\mathcal{F}_1 \times \mathcal{F}_2|_{\Lambda_1 \times \Lambda_2}) \leq h_{Ad_1}(\mathcal{F}_1|_{\Lambda_1}) + h_{Ad_2}(\mathcal{F}_2|_{\Lambda_2}).$$

Moreover, if X_i is compact and $h_{Md_i}(\Lambda_i) = \lim_{\varepsilon \to 0} \liminf_{n \to \infty} \frac{1}{n} s(n, \varepsilon, X_i, \Lambda_i)$ ($h_{Ad_i}(\Lambda_i) = \lim_{\varepsilon \to 0} \liminf_{n \to \infty} \frac{1}{n} N_{sep}(n, \varepsilon, X_i, \Lambda_i)$ respectively) for $i = 1$ or $i = 2$, then

$$h_{Md}(\mathcal{F}_1 \times \mathcal{F}_2|_{\Lambda_1 \times \Lambda_2}) = h_{Md_1}(\mathcal{F}_1|_{\Lambda_1}) + h_{Md_2}(\mathcal{F}_2|_{\Lambda_2})$$

and

$$h_{Ad}(\mathcal{F}_1 \times \mathcal{F}_2|_{\Lambda_1 \times \Lambda_2}) = h_{Ad_1}(\mathcal{F}_1|_{\Lambda_1}) + h_{Ad_2}(\mathcal{F}_2|_{\Lambda_2}).$$

Proof. First, we will show that $\Lambda_1 \times \Lambda_2$ can be regard as a subshift. Let $A = \{(i, j),\ i \in \{0, \ldots, m-1\},\ j \in \{0, \ldots, k-1\}\}$ and Λ_A be the full shift on A. Put $\Lambda = \{x \in \Lambda_A : x^1_{[n,n+1]} \in \mathcal{L}^2(\Lambda_1)$ and $x^2_{[n,n+1]} \in \mathcal{L}^2(\Lambda_2)$ for all $n\}$, where $x^1_{[n,n+1]}$ and $x^2_{[n,n+1]}$ denote the first and second components, respectively, of $x_{[n,n+1]}$. That is $(i_1\ j_2)(i_2\ j_2) \in \mathcal{L}^2(\Lambda)$ if and only if $i_1 i_2 \in \Lambda_1$, $j_1 j_2 \in \Lambda_2$. It is easy to check that $(\Lambda_1 \times \Lambda_2, \sigma_1 \times \sigma_2)$ and (Λ, σ) are conjugate each other.

Let $\nu = \nu_0 \cdots \nu_{n-1} \in \mathcal{L}^n(\Lambda_1 \times \Lambda_2)$, $\mu^{(i)} = \mu_0^{(i)} \cdots \mu_{n-1}^{(i)} \in \mathcal{L}^n(\Lambda_i)$, $i = 1, 2$. It is clear that, by the proof of [43, Theorem 3.8], the map $\nu = \mu^{(1)} \times \mu^{(2)} \to (\mu^{(1)}, \mu^{(2)})$ is a one to one correspondence.

The rest of the proof follows the proofs of [42, Theorem 3.11] and [43, Theorem 3.8]. $\square$

Finally in this section, we describe an appropriate version of topological conjugacy, which preserves the entropies of a subshift action.

Theorem 3.5. *Let (X, d) and $(\overline{X}, \overline{d})$ be metric spaces and $\pi : X \to \overline{X}$ a continuous surjection such that there is $\delta > 0$ with*

$$\pi|_{B_d(x,\delta)} : B_d(x, \delta) \to B_{\overline{d}}(\pi(\overline{x}), \delta)$$

being an isometric surjection for all $x \in X$. If $\mathcal{G} = \{f_0, f_1, \ldots, f_{m-1}\}$ are uniformly continuous transformations on X, and $\overline{\mathcal{G}} = \{\overline{f}_0, \overline{f}_1, \ldots, \overline{f}_{m-1}\}$ are uniformly continuous transformations on $\overline{X}$ satisfying $\pi f_i = \overline{f}_i \pi$ for any $0 \leq i \leq m-1$, then

$$h_{Md}(\mathcal{G}|_\Lambda) = h_{M\overline{d}}(\overline{\mathcal{G}}|_\Lambda)$$

and

$$h_{Ad}(\mathcal{G}|_\Lambda) = h_{A\overline{d}}(\overline{\mathcal{G}}|_\Lambda).$$

Proof. The proof is similar to the proof of [42, Theorem 4.4] and is omitted. □

4. Estimates of the entropies

In this section, by the methods used in [42] and [43], we will give some estimates of the topological entropies defined in Section 2 for three kinds of particular systems. First, we shall obtain an upper bound for the entropies of a subshift action generated by a family of differentiable maps on a finite-dimensional Riemannian manifold.

Theorem 4.1. *Let M' be a p-dimensional Riemannian manifold and $\mathcal{G} = \{f_0, f_1, \ldots, f_{k-1}\}$ be a family of C^1 maps on M'. Then*

$$h_{M\,d}(\mathcal{G}|_\Lambda) \leq \max\{0, p\log(\max_{0\leq i\leq k-1}\sup_{x\in M'}\|d_x f_i\|)\}$$

and

$$h_{A\,d}(\mathcal{G}|_\Lambda) \leq \limsup_{n\to\infty}\frac{1}{n}\log(\sum_{w\in\mathcal{L}^n(\Lambda)}(\prod_{i=1}^{n}\max\{1, \sup_{x\in M}\|d_x f_{w_i}\|\})^p) - h(\Lambda,\sigma),$$

where d denotes the metric on M' induced by the Riemannian metric.

Next result gives the entropies of a subshift action generated by a family of surjective endomorphisms on a compact metrizable group.

Theorem 4.2. *Let X be a compact metrizable group, $\mathcal{A} = \{A_0, \ldots, A_{m-1}\}$ surjective endomorphisms of X, $a_0, \ldots, a_{m-1} \in X$. Denote by μ the Haar measure on X and by d a left-invariant metric on X. Let $\overline{\mathcal{A}} = \{a_0\cdot A_0, \ldots, a_{m-1}\cdot A_{m-1}\}$. Then*

$$h_M(\mathcal{A}|_\Lambda) = h_M(\overline{\mathcal{A}}|_\Lambda),$$

$$h_M(\mathcal{A}|_\Lambda) = \lim_{\varepsilon\to 0}\limsup_{n\to\infty}[-\frac{1}{n}\log\mu(D_n(e,\varepsilon,\Lambda))]$$

and

$$h_A(\mathcal{A}|_\Lambda) = h_A(\overline{\mathcal{A}}|_\Lambda),$$

$$h_A(\mathcal{A}|_\Lambda) = \lim_{\varepsilon\to 0}\limsup_{n\to\infty}[\frac{1}{n}\log(\frac{1}{|\mathcal{L}^n(\Lambda)|}\sum_{w\in\mathcal{L}^n(\Lambda)}\frac{1}{\mu(D_w(e,\varepsilon,\Lambda))})],$$

where

$$D_n(e,\varepsilon,\Lambda) = \cap_{w'\sqsubseteq w, w\in\mathcal{L}^n(\Lambda)}A_{w'}^{-1}(B_d(e,\varepsilon)),$$

$$D_w(e,\varepsilon,\Lambda) = \cap_{w'\sqsubseteq w}A_{w'}^{-1}(B_d(e,\varepsilon)),$$

and e is the identity element of X and $B_d(e,\varepsilon)$ is the open ball with center e and radius ε.

Finally, we shall give the entropies of a linear subshift action on $\mathbb{R}^p$, which plays a key in next section.

Lemma 4.3. *Let $\mathcal{A} = \{A_0, \ldots, A_{m-1}\}$ be the linear transformations on $\mathbb{R}^p$, μ the Lebesgue measure on $\mathbb{R}^p$ and ρ a metric on $\mathbb{R}^p$ defined by a norm. Then*

$$h_{M\rho}(\mathcal{A}|_\Lambda) = \lim_{\varepsilon\to 0} \limsup_{n\to\infty} [-\frac{1}{n} \log \mu(D_n(0, \varepsilon, \Lambda))]$$

and

$$h_{A\rho}(\mathcal{A}|_\Lambda) = \lim_{\varepsilon\to 0} \limsup_{n\to\infty} [\frac{1}{n} \log(\frac{1}{|\mathcal{L}^n(\Lambda)|} \sum_{w\in\mathcal{L}^n(\Lambda)} \frac{1}{\mu(D_w(0, \varepsilon, \Lambda))})],$$

where

$$D_n(0, \varepsilon, \Lambda) = \cap_{w'\sqsubseteq w, w\in\mathcal{L}^n(\Lambda)} A_{w'}^{-1} B_d(0, \varepsilon),$$
$$D_w(0, \varepsilon, \Lambda) = \cap_{w'\sqsubseteq w} A_{w'}^{-1} B_d(0, \varepsilon),$$

and

$$B_\rho(0, \varepsilon) = \{x \in \mathbb{R}^p : \rho(x, 0) < \varepsilon\}.$$

$h_{M\rho}(\Lambda)$ and $h_{A\rho}(\Lambda)$ do not depend on the norm chosen.

Proof. Since all norms on $\mathbb{R}^p$ are equivalent, they induce uniformly equivalent metrics on $\mathbb{R}^p$. So by Theorem 3.1, we have $h_{M\rho}(\mathcal{A}|_\Lambda) = h_{Md}(\mathcal{A}|_\Lambda)$, where d is the Euclidean distance of $\mathbb{R}^p$. It is also clear that the expression given in the theorem is independent of the norm. Hence we suppose ρ is the Euclidean distance.

The rest of the proof follows the proofs of [42, Lemma 4.5] and [43, Lemma 5.5] and is omitted. □

5. Maximal entropy for linear switched systems

In this section we estimate the entropies of some linear subshift actions, including subshift with periodic points and finite subshift, on Euclidean spaces.

Recall Theorem 4.6 of [42], which gives a lower and upper bounds of maximal entropy for a full shift action.

Proposition 5.1 ([42, Theorem 4.6]). *Let G be a semigroup generated by $\mathcal{A} = \{id_{\mathbb{R}^d}, A_1, \ldots, A_k\}$ which is a set of linear transformations on $\mathbb{R}^d$. If all the eigenvalues of A_i are of modulus great than or equal to 1 for each $1 \le i \le k$, then*

$$\max_{1\le i\le k} \sum_{j=1}^{d} \log|\lambda_j^{(i)}| \le h_{M\rho}(\mathcal{A}) \le d \max_{1\le i\le k} \log\hat{\lambda}_i, \tag{5}$$

where $\lambda_1^{(i)}, \lambda_2^{(i)}, \ldots, \lambda_d^{(i)}$ are the eigenvalues of A_i, $1 \le i \le k$, counted with their multiplicities, and $\hat{\lambda}_i$ is the biggest eigenvalue of $\sqrt{A_i A_i^T}$, $1 \le i \le k$.

Particularly in the case $d = 1$ and $X = \mathbb{R}^1$, we have

$$h_{M\rho}(\mathcal{A}) = \max_{1\le i\le k} \log|\lambda_1^{(i)}|,$$

where $\lambda_1^{(i)}$ is the proportionality constant of $A_i : \mathbb{R} \to \mathbb{R} :\ x \mapsto \lambda_1^{(i)} x$, $1 \le i \le k$.

The bound of $h_{M\rho}(\mathcal{A})$ obtained by the proposition above is too conservative. We will give a sharper bound of the entropy without the condition that all the eigenvalues of A_i are of modulus great than or equal to 1 for each $1 \le i \le k$. First, for the lower bound, we have:

Theorem 5.2. *Let $\mathcal{A} = \{A_1, \dots, A_k\}$ be a set of linear transformations on $\mathbb{R}^d$ and Λ be a subshift. If $G_l = \{A_{\omega_{[0,l)}}, \omega \in Per_l(\Lambda)\} \neq \phi$, then*

$$h_{M\rho}(\mathcal{A}|_\Lambda) \ge \sup_{l\ge k} \max_{A\in G_l} \frac{1}{l} \sum_{\{i||\lambda_i^A|>1\}} \log|\lambda_i^A|, \tag{6}$$

where $k = \min\{l : Per_l(\Lambda) \neq \emptyset\}$, $Per_l(\Lambda) = \{x : \sigma^l(x) = x,\ x \in \Lambda\}$ and $\lambda_1^A, \dots, \lambda_d^A$ are the eigenvalues of A. (Some of λ_i^A can be equal.)

Proof. It is sufficient to show that for any $l \ge k$ and any $A_* \in G_l$ it holds

$$h_{M\rho}(\mathcal{A}|_\Lambda) \ge \frac{1}{l} \sum_{\{i||\lambda_i^{A_*}|>1\}} \log|\lambda_i^{A_*}|. \tag{7}$$

By Lemma 4.3, we have

$$\begin{aligned} h_{M\rho}(\mathcal{A}|_\Lambda) &= \lim_{\varepsilon\to 0} \limsup_{n\to\infty} \left[-\frac{1}{n}\log m(D_n(0,\varepsilon,G))\right] \\ &\ge \lim_{\varepsilon\to 0} \limsup_{n\to\infty} \left[-\frac{1}{nl}\log m(D_{nl}(0,\varepsilon,G))\right], \end{aligned}$$

where

$$D_{nl}(0,\varepsilon,G) = \bigcap_{A\in G_{nl}} A^{-1}B_\rho(0,\varepsilon).$$

Since

$$D_{nl}(0,\varepsilon,G) \subset \bigcap_{i=0}^{n-1} A_*^{-i}B_\rho(0,\varepsilon) \triangleq D_n^*(0,\varepsilon,A_*),$$

we further have

$$\begin{aligned} h_{M\rho}(\mathcal{A}|_\Lambda) &\ge \frac{1}{l} \lim_{\varepsilon\to 0} \limsup_{n\to\infty} \left[-\frac{1}{n}\log m(D_n^*(0,\varepsilon,A_*))\right] \\ &= \frac{1}{l} \sum_{\{i||\lambda_i^{A_*}|>1\}} \log|\lambda_i^{A_*}|. \end{aligned}$$

The last equality comes from Lemma 8.13 and Theorem 8.14 in [41]. So the inequality (7) holds and the proof is completed. □

We have several remarks on Theorem 5.2.

Remark 5.3. Roughly speaking, the topological entropy of a switched system is bounded below by the topological entropy of any periodic switching sequence.

Remark 5.4. When all the eigenvalues of A_i are of modulus great than or equal to 1 for each $1 \le i \le k$, it is obvious that

$$\max_{1\le i\le k}\sum_{j=1}^{d}\log|\lambda_j^{(i)}| = \sup_{l\ge 1}\max_{A\in G_l}\frac{1}{l}\sum_{\{i||\lambda_i^A|>1\}}\log|\lambda_i^A|.$$

Remark 5.5. If $\mathcal{A} = \{id, A\}$, then it is well known that

$$h_{M\rho}(\mathcal{A}) = \sum_{\{i||\lambda_i|>1\}}\log|\lambda_i|,$$

see *e.g.* Theorem 8.14 in [41]. So the lower bound is sharp and cannot be improved.

Next, we give a new lower and upper bound for $h_{M\rho}(\mathcal{A})$ by the joint spectral radius of $\mathcal{A}$. Recall that the joint spectral radius of $\mathcal{A}$, due to Rota and Strang [37], is defined by

$$\hat{\rho}(\mathcal{A}) = \limsup_{n\to\infty}\max_{(i_1,\dots,i_n)\in\mathcal{I}^n}\|A_{i_n}\cdots A_{i_1}\|^{\frac{1}{n}},$$

where we denote

$$\mathcal{I}^n = \overbrace{\mathcal{I}\times\cdots\times\mathcal{I}}^{n\ times}.$$

It is easy to see that $\hat{\rho}(\mathcal{A})$ can be rewritten as

$$\hat{\rho}(\mathcal{A}) = \lim_{n\to\infty}\max_{(i_1,\dots,i_n)\in\mathcal{I}^n}\|A_{i_n}\cdots A_{i_1}\|^{\frac{1}{n}} = \inf_{n\ge 1}\max_{(i_1,\dots,i_n)\in\mathcal{I}^n}\|A_{i_n}\cdots A_{i_1}\|^{\frac{1}{n}}, \tag{8}$$

by the subadditivity of the function $n \to \log\max_{(i_1,\dots,i_n)\in\mathcal{I}^n}\|A_{i_n}\cdots A_{i_1}\|$. It is well known that all infinite products of the matrices in $\mathcal{A}$ converge to zero if and only if $\hat{\rho}(\mathcal{A}) < 1$ (e.g. see Barabanov [3]; Shih, Wu, & Pang [38]). Moreover, the quantity $\hat{\rho}(\mathcal{A})$ is independent of the matrix norm $\|\cdot\|$ used here, and according to Berger and Wang [4] and Elsner [17], $\hat{\rho}(\mathcal{A})$ is equal to the generalized spectral radius $\rho(\mathcal{A})$ of $\mathcal{A}$, which was firstly introduced by Daubechies and Lagarias [14], given by

$$\rho(\mathcal{A}) = \sup_{n\ge 1}\max_{(i_1,\dots,i_n)\in\mathcal{I}^n}\rho(A_{i_n}\cdots A_{i_1})^{1/n}$$

or equivalently

$$\rho(\mathcal{A}) = \limsup_{n\to\infty}\max_{(i_1,\dots,i_n)\in\mathcal{I}^n}\rho(A_{i_n}\cdots A_{i_1})^{1/n}.$$

Here, the spectral radius of a single matrix $A \in \mathbb{R}^{d\times d}$ is defined by

$$\rho(A) = \max\{|\lambda| \mid \lambda \text{ is an eigenvalue of } A\}.$$

Though the computation of $\rho(\mathcal{A})$ in a closed formula form is generally impossible, there are some approximate computation method to evaluate it. Furthermore, several different techniques have been developed for computing the exact value of $\rho(\mathcal{A})$ for some particular kinds of $\mathcal{A}$. See [12, 13, 29] and references therein. Recently Dai generalized in [11] the Berger-Wang formula to a matrix multiplicative semigroup $\mathcal{A} = \{A_0, \dots, A_{m-1}\}$ restricted to a subset that need not carry the algebraic structure of $\mathcal{A}$, by ergodic theory; and later Kozyak ([28]) extended this concept to product of matrices on the sliding block relative frequencies of occurrences of different factors. For a subshift Λ. We write

$$\hat{\rho}_n = \max_{w \in \mathcal{L}^n(\Lambda)} \|A_w\|.$$

According to [11], the joint spectral radius of $\mathcal{A}$ restricted on Λ is defined by

$$\hat{\rho}(\mathcal{A}|_\Lambda) = \limsup_{n\to\infty} \sqrt[n]{\hat{\rho}_n}.$$

Theorem 5.6. *Let V be a p-dimensional vector space, ρ a metric on V induced by a norm on V, $\mathcal{A} = \{A_0, \dots, A_{m-1}\}$ the linear transformations on V. If for every $i \in \{0, \dots, m-1\}$, all eigenvalues of A_i are of modulus greater than or equal to 1, then*

$$\limsup_{n\to\infty} \frac{1}{n} \max_{w \in \cup_{i=0}^n \mathcal{L}^n(\Lambda)} \{\sum_{i=0}^{|w|} \sum_{j=1}^{p} \log|\lambda_j^{(w_i)}|\} \le h_{M\rho}(\mathcal{A}|_\Lambda) \le \max\{0, p\log\hat{\rho}(\mathcal{A}|_\Lambda)\},$$

where $\lambda_1^{(i)}, \dots, \lambda_p^{(i)}$ are eigenvalues of A_i, $0 \le i \le m-1$, counted with their multiplicities.

Proof. Without loss of generality, we can suppose $V = \mathbb{R}^p$. Let μ be the Lebesgue measure of $\mathbb{R}^p$. Since all norms on $\mathbb{R}^p$ are equivalent they induce uniformly equivalent metrics on $\mathbb{R}^p$ and by Theorem 3.1, we have $h_{M\rho}(\mathcal{A}|_\Lambda) = h_{Md}(\mathcal{A}|_\Lambda)$ where d is the Euclidean distance.

By Lemma 4.3, one has

$$h_{M\rho}(\mathcal{A}|_\Lambda) = \lim_{\varepsilon\to 0} \limsup_{n\to\infty} [-\frac{1}{n}\log\mu(D_n(0,\varepsilon,\Lambda))].$$

For any $0 \le i \le m-1$, we have $\mu(A_i(B)) = |\det A_i|\mu(B)$ for any $B \in \mathcal{B}(\mathbb{R}^p)$. Choose $w^{(n)} \in \cup_{i=0}^n \mathcal{L}^i(\Lambda)$ such that

$$\sum_{i=0}^{|w^{(n)}|} \sum_{j=1}^{p} \log|\lambda_j^{(w_i^{(n)})}| = \max_{w \in \cup_{i=0}^n \mathcal{L}^i(\Lambda)} \{\sum_{i=0}^{|w|} \sum_{j=1}^{p} \log|\lambda_j^{(w_i)}|\}.$$

So

$$\begin{aligned} \mu(D_n(0,\varepsilon,\Lambda)) &\le \mu(A_{w^{(n)}}^{-1} B_\rho(0,\varepsilon)) \\ &= |\det(A_{w^{(n)}}^{-1})| \cdot \mu(B_\rho(0,\varepsilon)) \\ &= \frac{1}{\prod_{i=0}^{|w^{(n)}|} \prod_{j=1}^{p} |\lambda_j^{(w_i^{(n)})}|} \cdot \mu(B_\rho(0,\varepsilon)). \end{aligned}$$

Thus

$$-\frac{1}{n}\log\mu(D_n(0,\varepsilon,\Lambda)) \ge \frac{1}{n} \sum_{i=0}^{|w^{(n)}|} \sum_{j=1}^{p} \log|\lambda_j^{(w_i^{(n)})}| - \frac{1}{n}\log\mu(B_\rho(0,\varepsilon)).$$

It follows that

$$h_{M\rho}(\mathcal{A}|_\Lambda) \geq \limsup_{n\to\infty} \frac{1}{n} \sum_{i=0}^{|w^{(n)}|} \sum_{j=1}^{p} \log |\lambda_j^{(w_i^{(n)})}|$$

$$= \limsup_{n\to\infty} \frac{1}{n} \max_{w\in\cup_{i=0}^{n}\mathcal{L}^i(\Lambda)} \{\sum_{i=0}^{|w|} \sum_{j=1}^{p} \log |\lambda_j^{(w_i)}|\}.$$

The proof for the right inequality includes two steps. Step1, if $\hat{\rho}(\mathcal{A}|_\Lambda) < 1$, then there exists $N > 0$ such that $\hat{\rho}_n < 1$ for every $n \geq N$. That is $\|A_w\| < 1$ for any $w \in \mathcal{L}^n(\Lambda)$, $n \geq N$. For any compact subset K of $\mathbb{R}^d$, it is easy to see that if E is a $(N, \varepsilon, K, \Lambda)$-spanning set of K, then E is also a $(n, \varepsilon, K, \Lambda)$-spanning set of K for any $n \geq N$, which implies that $h_M(\mathcal{A}|_\Lambda) = 0$.

Step2, if $\rho \triangleq \hat{\rho}(\mathcal{A}|_\Lambda) \geq 1$, then for any $\delta > 0$, there exists $N > 0$ such that $\|A_w\| \leq (\rho+\delta)^n$ for any $n \geq N$ and $w \in \mathcal{L}^n(\Lambda)$. We have, for any $n \geq N$, that

$$B_\rho(0, \frac{1}{\rho_n}\varepsilon) \subset (A_w)^{-1} B_\rho(0, \varepsilon),$$

where

$$\rho_n = \max\{\max_{w\in\cup_{i=0}^{N}\mathcal{L}^i(\Lambda)} \|A_w\|, (\rho+\delta)^n\}.$$

Since $(\rho+\delta)^n \to \infty$ as $n \to \infty$. We can suppose that

$$\rho_n = (\rho+\delta)^n > 1,$$

when we take $n \in \mathbb{N}$ enough large.

So

$$B_\rho(0, \frac{1}{\rho_n}\varepsilon) \subset D_n(0, \varepsilon, \Lambda) = \cap_{w\in\cup_{i=0}^{n}\mathcal{L}^i(\Lambda)} (A_w)^{-1} B_\rho(0, \varepsilon).$$

Thus

$$-\frac{1}{n} \log \mu(D_n(0, \varepsilon, \Lambda)) \leq \frac{1}{n} p \cdot \log \rho_n - \frac{1}{n} \log \mu(B(0, \varepsilon)).$$

Hence

$$h_{M\rho}(\mathcal{A}|_\Lambda) \leq \limsup_{n\to\infty} \frac{1}{n} p \log \rho_n$$

$$= \limsup_{n\to\infty} \frac{1}{n} p \log(\rho+\delta)^n$$

$$= p \log(\rho+\delta).$$

Let $\delta \to 0$, we have

$$h_{M\rho}(\mathcal{A}|_\Lambda) \leq p \log \rho = p \log \hat{\rho}(\mathcal{A}|_\Lambda). \qquad \square$$

Remark 5.7. From the above proof, we see that the upper bound by joint spectral radius works, no matter whether the eigenvalues of A_i are of modulus greater than or equal to 1 or not.

Corollary 5.8. *Under the assumptions of Proposition 5.1, we have*

$$h_{M\rho}(\mathcal{A}) \le d \log \hat{\rho}(\mathcal{A}),$$

where $\hat{\rho}(\mathcal{A})$ *is the joint spectral radius of* $\mathcal{A}$.

Remark 5.9. If we take the matrix norm $\|\cdot\|$ induced by the Euclidean norm in $\mathbb{R}^d$, then $\|A\| \le \lambda_A$, where λ_A is the largest eigenvalue of $\sqrt{AA^T}$. Thus by the definition of $\hat{\rho}(\mathcal{A})$, we have

$$\hat{\rho}(\mathcal{A}) \le \max_{A\in\mathcal{A}} \lambda_A.$$

The following example illustrates the upper bound of $h_{M\rho}(\mathcal{A})$ obtained by the above theorem is strictly less than that in Proposition 5.1.

Example 5.10. Let $\mathcal{A} = \{id, A_1, A_2\}$ with

$$A_1 = \frac{5}{4}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad A_2 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

Then $\max\{\lambda_{A_1}, \lambda_{A_2}\} = \frac{5}{4}\sqrt{\frac{3+\sqrt{5}}{2}}$. It follows from [20, Example 1, p. 63] that $\hat{\rho}(\mathcal{A}) = \frac{5}{4}(1 + \frac{1}{\sqrt{5}})$. So $\hat{\rho}(\mathcal{A}) < \max_{A\in\mathcal{A}} \lambda_A$. In fact, from Proposition 5.1 and Theorem 5.2 we obtain

$$2\log\frac{5}{4} \le h_{M\rho}(\mathcal{A}) \le 2\log\frac{5+\sqrt{5}}{4}.$$

Acknowledgment. The authors are grateful to the referees and Administration Editorial Committee for their valuable comments which have led to improvement of the paper. A part of this work was done when the second author was visiting University of Western Ontario. He would like to thank Prof. Xingfu Zou and Western University for their hospitality during his visit.

References

[1] R. L. Adler, A. G. Konheim, and M. H. McAndrew, *Topological entropy*, Trans. Amer. Math. Soc. **114** (1965), 309–319.

[2] L. Arnold, *Random Dynamical Systems*, Springer Monographs in Mathematics, Springer-Verlag, Berlin, 1998.

[3] N. E. Barabanov, *On Lyapunov indicators of discrete inclusions I*, Automat. Remote Control **49** (1988), no. 2, 152–157; translated from Avtomat. i Telemekh. **1988** (1988), no. 2, 40–46.

[4] M. A. Berger and Y. Wang, *Bounded semigroups of matrices*, Linear Algebra Appl. **166** (1992), 21–27.

[5] A. Biś, *Entropies of a semigroup of maps*, Discrete Contin. Dyn. Syst. **11** (2004), no. 2-3, 639–648.

[6] T. Bogenschütz, *Entropy, pressure, and a variational principle for random dynamical systems*, Random & Computational Dynamics **1** (1992), 99–116.

[7] R. Bowen, *Entropy for group endomorphisms and homogeneous spaces*, Trans. Amer. Math. Soc. **153** (1971), 401–414.

[8] A. Bufetov, *Topological entropy of free semigroup actions and skew-product transformations*, J. Dynam. Control Systems **5** (1999), no. 1, 137–143.

[9] J. S. Cánovas, *On entropy of non-autonomous discrete systems*, in Progress and challenges in dynamical systems, 143–159, Springer Proc. Math. Stat., **54**, Springer, Heidelberg, 2013.

[10] O. L. V. Costa, M. D. Fragoso, and R. P. Marques, *Discrete-time markov jump linear systems*, Automatic Control IEEE Transactions **51** (2006), no. 5, 916–917.

[11] X. Dai, *A Gel'fand-type spectral-radius formula and stability of linear constrained switching systems*, Linear Algebra Appl. **436** (2012), no. 5, 1099–1113.

[12] ———, *Some criteria for spectral finiteness of a finite subset of the real matrix space* $\mathbb{R}^{d\times d}$, Linear Algebra Appl. **438** (2013), no. 6, 2717–2727.

[13] X. Dai, Y. Huang, J. Liu, and M. Xiao, *The finite-step realizability of the joint spectral radius of a pair of $d \times d$ matrices one of which being rank-one*, Linear Algebra Appl. **437** (2012), no. 7, 1548–1561.

[14] I. Daubechies and J. C. Lagarias, *Sets of matrices all infinite products of which converge*, Linear Algebra Appl. **161** (1992), 227–263.

[15] E. I. Dinaburg, *A correlation between topological entropy and metric entropy*, Dokl. Akad. Nauk SSSR **190** (1970), 19–22.

[16] A. H. Dooley and G. Zhang, *Local entropy theory of a random dynamical system*, Mem. Amer. Math. Soc. **233** (2015), no. 1099, vi+106 pp.

[17] L. Elsner, *The generalized spectral-radius theorem: an analytic-geometric proof*, Linear Algebra Appl. **220** (1995), 151–159.

[18] S. Friedland, *Entropy of graphs, semigroups and groups*, in Ergodic theory of $\mathbf{Z}^d$ actions (Warwick, 1993–1994), 319–343, London Math. Soc. Lecture Note Ser., **228**, Cambridge Univ. Press, Cambridge, 1996.

[19] G. Froyland and O. Stancevic, *Metastability, Lyapunov exponents, escape rates, and topological entropy in random dynamical systems*, Stoch. Dyn. **13** (2013), no. 4, 1350004, 26 pp.

[20] N. Guglielmi and V. Protasov, *Exact computation of joint spectral characteristics of linear operators*, Found. Comput. Math. **13** (2013), no. 1, 37–97.

[21] K. H. Hofmann and L. N. Stojanov, *Topological entropy of group and semigroup actions*, Adv. Math. **115** (1995), no. 1, 54–98.

[22] W. Huang, *Entropy, chaos and weak horseshoe for infinite dimensional random dynamical systems*, Mathematics, 2015.

[23] C. Kawan, *Metric entropy of nonautonomous dynamical systems*, Nonauton. Dyn. Syst. **1** (2014), 26–52.

[24] C. Kawan and Y. Latushkin, *Some results on the entropy of non-autonomous dynamical systems*, Dyn. Syst. **31** (2016), no. 3, 251–279.

[25] Y. Kifer, *Ergodic Theory of Random Transformations*, Progress in Probability and Statistics, **10**, Birkhäuser Boston, Inc., Boston, MA, 1986.

[26] 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.

[27] S. Kolyada and L'. Snoha, *Topological entropy of nonautonomous dynamical systems*, Random Comput. Dynam. **4** (1996), no. 2-3, 205–233.

[28] V. Kozyakin, *Matrix products with constraints on the sliding block relative frequencies of different factors*, Linear Algebra Appl. **457** (2014), 244–260.

[29] ———, *Constructive stability and stabilizability of positive linear discrete-time switching systems*, arXiv:1511.05665v1, 2015.

[30] P. Kůrka, *Topological and Symbolic Dynamics*, Cours Spécialisés, **11**, Société Mathématique de France, Paris, 2003.

[31] R. Langevin and F. Przytycki, *Entropie de l'image inverse d'une application*, Bull. Soc. Math. France **120** (1992), no. 2, 237–250.

[32] R. Langevin and P. Walczak, *Entropie d'une dynamique*, C. R. Acad. Sci. Paris Sér. I Math. **312** (1991), no. 1, 141–144.

[33] E. Mihailescu and M. Urbański, *Inverse topological pressure with applications to holomorphic dynamics of several complex variables*, Commun. Contemp. Math. **6** (2004), no. 4, 653–679.
[34] ———, *Inverse pressure estimates and the independence of stable dimension for non-invertible maps*, Canad. J. Math. **60** (2008), no. 3, 658–684.
[35] ———, *Random countable iterated function systems with overlaps and applications*, Adv. Math. **298** (2016), 726–758.
[36] Z. H. Nitecki, *Topological entropy and the preimage structure of maps*, Real Anal. Exchange **29** (2003/04), no. 1, 9–41.
[37] G.-C. Rota and G. Strang, *A note on the joint spectral radius*, Nederl. Akad. Wetensch. Proc. Ser. A 63 = Indag. Math. **22** (1960), 379–381.
[38] M.-H. Shih, J.-W. Wu, and C.-T. Pang, *Asymptotic stability and generalized Gelfand spectral radius formula*, Linear Algebra Appl. **252** (1997), 61–70.
[39] J. Sinaǐ, *On the concept of entropy for a dynamic system*, Dokl. Akad. Nauk SSSR **124** (1959), 768–771.
[40] Z. Sun and S. S. Ge, *Switched Linear Systems*: *Control and Design*. Springer-Verlag, 2006.
[41] P. Walters, *An Introduction to Ergodic Theory*, Graduate Texts in Mathematics, **79**, Springer-Verlag, New York, 1982.
[42] Y. Wang and D. Ma, *On the topological entropy of a semigroup of continuous maps*, J. Math. Anal. Appl. **427** (2015), no. 2, 1084–1100.
[43] Y. Wang, D. Ma, and X. Lin, *On the topological entropy of free semigroup actions*, J. Math. Anal. Appl. **435** (2016), no. 2, 1573–1590.
[44] Y. Zhu, Z. Liu, X. Xu, and W. Zhang, *Entropy of nonautonomous dynamical systems*, J. Korean Math. Soc. **49** (2012), no. 1, 165–185.

Yu Huang
School of Mathematics
Sun Yat-sen University
Guangzhou, 510275 P. R. China
Email address: stshyu@mail.sysu.edu.cn

Xingfu Zhong
School of Mathematics
Sun Yat-sen University
Guangzhou, 510275 P. R. China
Email address: zhongxingfu224634@163.com