VARIATIONAL PRINCIPLE FOR WEIGHTED TOPOLOGICAL PRESSURE

DE-JUN FENG AND WEN HUANG

Abstract. Let $\pi : X \to Y$ be a factor map, where (X, T) and (Y, S) are topological dynamical systems. Let $\mathbf{a} = (a_1, a_2) \in \mathbb{R}^2$ with $a_1 > 0$ and $a_2 \geq 0$, and $f \in C(X)$. The $\mathbf{a}$-weighted topological pressure of f, denoted by $P^{\mathbf{a}}(X, f)$, is defined by resembling the Hausdorff dimension of subsets of self-affine carpets. We prove the following variational principle:

$$P^{\mathbf{a}}(X, f) = \sup \left\{ a_1 h_\mu(T) + a_2 h_{\mu\circ\pi^{-1}}(S) + \int f \, d\mu \right\},$$

where the supremum is taken over the T-invariant measures on X. It not only generalizes the variational principle of classical topological pressure, but also provides a topological extension of dimension theory of invariant sets and measures on the torus under affine diagonal endomorphisms. A higher dimensional version of the result is also established.

1. Introduction

Inspired by the theory of Gibbs states in statistical mechanics, Ruelle [33] introduced the notion of topological pressure to the theory of dynamical systems and established a variational principle for it. Ruelle only considered the case when the underlying dynamical systems satisfy expansiveness and specification. Later Walters [36] generalized these results to general topological dynamical systems. Topological pressure, and the associated variational principle and equilibrium measures constitute the main components of the thermodynamic formalism [34]. They play important roles in dimension theory of dynamical systems. Indeed they provide as a basic tool in studying dimension of invariant sets and measures for conformal dynamical systems (see e.g. [9, 35, 31]).

In this paper we aim to introduce a generalized notion of pressure for factor maps between general topological dynamical systems, and establish a variational principle for it. To be more precise, let us introduce some notation first. We say that (X, T)

2000 *Mathematics Subject Classification.* 37D35; 37C45; 37B40.

Key words and phrases. Variational principle, topological pressure, factor maps, Hausdorff dimension, self-affine carpets.

is a *topological dynamical system* (TDS) if X is a compact metric space and T is a continuous map from X to X. Let (X,T) and (Y,S) be two topological dynamical systems. Suppose that (Y,S) is a factor of (X,T), in the sense that there exists a continuous surjective map $\pi : X \to Y$ such that $\pi \circ T = S \circ \pi$. The map π is called a *factor map* from X to Y. Let f be a real-valued continuous function on X, and let $a_1 > 0$, $a_2 \geq 0$. The main purpose of this paper is to consider the following.

Question 1.1. *How can one define a meaningful term $P^{(a_1,a_2)}(T,f)$ such that the following variational principle holds?*

$$P^{(a_1,a_2)}(T,f) = \sup\left\{a_1 h_\mu(T) + a_2 h_{\mu\circ\pi^{-1}}(S) + \int f \, d\mu\right\}, \tag{1.1}$$

where the supremum is taken over the set of all T-invariant Borel probability measures μ on X, and $h_\mu(T), h_{\mu\circ\pi^{-1}}(S)$ stand for the measure-theoretic entropies of μ and $\mu\circ\pi^{-1}$ with respect to T and S, respectively (cf. [37]*).*

According to the variational principle of Ruelle and Walters, the left-hand side of (1.1) equals $a_1 P(T, \frac{1}{a_1} f)$ in the particular case when $a_2 = 0$, where $P(T,\cdot)$ stands for the classic topological pressure of continuous functions (cf. [37]). Our interest is on the general case that $a_2 \neq 0$. This project is motivated from the study of dimension of invariant sets and measures on the tori under diagonal affine expanding maps.

Let T be the endmorphism on the 2-dimensional torus $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$ represented by an integral diagonal matrix $A = \operatorname{diag}(m_1, m_2)$, where $2 \leq m_1 < m_2$. That is, $Tu = Au \pmod 1$ for $u \in \mathbb{T}^2$. In their seminal works, Bedford [5] and McMullen [27] independently determined the Hausdorff dimension of the so-called *self-affine Sierpinski gaskets*, which are a particular class of T-invariant subsets of $\mathbb{T}^2$ defined as follows:

$$K(T,\mathcal{D}) := \left\{\sum_{n=1}^{\infty} A^{-n} u_n : \; u_n \in \mathcal{D} \text{ for all } n \geq 1\right\},$$

where $\mathcal{D}$ runs over the non-empty subsets of

$$\left\{\begin{pmatrix} i \\ j \end{pmatrix} : \; i = 0,1,\ldots,m_1, \; j = 0,1,\ldots,m_2 - 1\right\}.$$

Moreover, McMullen [27] exhibited explicitly that for each $\mathcal{D}$, there exists an ergodic T-invariant measure μ supported on $K(T,\mathcal{D})$ with $\dim_H \mu = \dim_H K(T,\mathcal{D})$, where $\dim_H$ denotes the Hausdorff dimension of a set or measure (cf. [13]). Later Kenyon and Peres [20] extended this result to any compact T-invariant set $K \subseteq \mathbb{T}^2$, that is, there is an ergodic T-invariant measure μ supported on K so that $\dim_H \mu = \dim_H K$. Furthermore Kenyon and Peres [20] established the following variational principle for the Hausdorff dimension of K:

$$\dim_H K = \sup\left\{\frac{1}{\log m_2} h_\eta(T) + \left(\frac{1}{\log m_1} - \frac{1}{\log m_2}\right) h_{\eta\circ\pi^{-1}}(S)\right\}, \tag{1.2}$$

where the supremum is taken over the collection of T-invariant Borel probability measures η supported on K, $\pi : \mathbb{T}^2 \to \mathbb{T}^1$ denotes the projection $(x, y) \mapsto x$, and $S : \mathbb{T}^1 \to \mathbb{T}^1$ denotes the map $x \mapsto m_1 x (\text{mod } 1)$. It is easy to check that $(\mathbb{T}^1, S)$ is a factor of $(\mathbb{T}^2, T)$ with the factor map π. We emphasize that for any ergodic T-invariant measure η on $\mathbb{T}^2$, the sum in the bracket of (1.2) just equals $\dim_H \eta$ (cf. [20, Lemma 3.1]); i.e.

$$\dim_H \eta = \frac{1}{\log m_2} h_\eta(T) + \left(\frac{1}{\log m_1} - \frac{1}{\log m_2}\right) h_{\eta\circ\pi^{-1}}(S). \tag{1.3}$$

This is a version of Ledrappier-Young dimension formula for ergodic measures on $\mathbb{T}^2$. We remark that an extension of the variational relation (1.2) to higher dimensional tori was also established by Kenyon and Peres [20].

Let us turn back to Question 1.1. According to (1.2), if π is the factor map $(x, y) \mapsto x$ between the toral dynamics (K, T) and $(\pi(K), S)$ as in the above paragraph, and if $f \equiv 0$ on K, and $a_1 = \frac{1}{\log m_2}$, $a_2 = \frac{1}{\log m_1} - \frac{1}{\log m_2}$, then we can just define $P^{(a_1,a_2)}(f)$ to be the Hausdorff dimension of K. The problem arises how can we extend this to general factor maps between topological dynamical systems, as well as to general continuous functions f and vectors (a_1, a_2).

In [2, 15], Barral and the first author defined $P^{(a_1,a_2)}(f)$ (and called it *weighted topological pressure*) via relative thermodynamic formalism and subadditive thermodynamic formalism, in the particular case when the underlying dynamical systems (X, T) and (Y, S) are subshifts over finite alphabets. They also studied the dynamical properties of weighted equilibrium measures (i.e. the invariant measures μ which attain the supremum in (1.1)) and gave the applications to the multifractal analysis on Sirpinski gaskets/sponges [2], and to the uniqueness of invariant measures of full dimension supported on affine-invariant subsets of tori [15]. Independently, in this subshift case Yayama [38] defined $P^{(a_1,a_2)}(f)$ for the particular case $f \equiv 0$, along the similar way.

However, the approach of [2, 15] in defining $P^{(a_1,a_2)}(f)$ relies on certain special property of subshifts and does not extend to general topological dynamical systems (see Section 7.1 for details). Moreover, the variational principle established therein does not give a new proof of Kenyon and Peres' variational relation (1.2) for the Hausdorff dimension.

In the paper, we define $P^{(a_1,a_2)}(f)$ in a new way, which is inspired from the dimension theory of affine invariant subsets of tori, and from the "dimension" approaches of Bowen [8] and Pesin-Pitskel' [32] in defining the topological entropy and topological pressure for arbitrary subsets.

We will present our definition under a more general setting. Let $k \geq 2$. Assume that (X_i, d_i), $i = 1, \ldots, k$, are compact metric spaces, and (X_i, T_i) are topological dynamical systems. Moreover, assume that for each $1 \leq i \leq k-1$, (X_{i+1}, T_{i+1}) is a factor of (X_i, T_i) with a factor map $\pi_i : X_i \to X_{i+1}$; in other words, $\pi_1, \ldots, \pi_{k-1}$ are continuous maps so that the following diagrams commute.

$$\begin{array}{ccccccc} X_1 & \xrightarrow{\pi_1} & X_2 & \xrightarrow{\pi_2} & \cdots & \xrightarrow{\pi_{k-1}} & X_k \\ {\scriptstyle T_1}\downarrow & & \downarrow{\scriptstyle T_2} & & & & \downarrow{\scriptstyle T_k} \\ X_1 & \xrightarrow{\pi_1} & X_2 & \xrightarrow{\pi_2} & \cdots & \xrightarrow{\pi_{k-1}} & X_k \end{array}$$

For convenience, we use π_0 to denote the identity map on X_1. Define $\tau_i : X_1 \to X_{i+1}$ by $\tau_i = \pi_i \circ \pi_{i-1} \circ \cdots \circ \pi_0$ for $i = 0, 1, \ldots, k-1$.

Let $\mathcal{M}(X_i, T_i)$ denote the set of all T_i-invariant Borel probability measures on X_i, endowed with the weak-star topology. Fix $\mathbf{a} = (a_1, a_2, \ldots, a_k) \in \mathbb{R}^k$ with $a_1 > 0$ and $a_i \geq 0$ for $i \geq 2$. For $\mu \in \mathcal{M}(X_1, T_1)$, we call

$$h_\mu^{\mathbf{a}}(T_1) := \sum_{i=1}^{k} a_i h_{\mu \circ \tau_{i-1}^{-1}}(T_i)$$

the $\mathbf{a}$-*weighted measure-theoretic entropy of* μ *with respect to* T_1, or simply, the $\mathbf{a}$-*weighted entropy of* μ, where $h_{\mu \circ \tau_{i-1}^{-1}}(T_i)$ denotes the measure-theoretic entropy of $\mu \circ \tau_{i-1}^{-1}$ with respect to T_i.

Definition 1.2 ($\mathbf{a}$-weighted Bowen ball)**.** *For $x \in X_1$, $n \in \mathbb{N}$, $\epsilon > 0$, denote*

$$B_n^{\mathbf{a}}(x, \epsilon) := \left\{ y \in X_1 : \ d_i(T_i^j \tau_{i-1} x, T_i^j \tau_{i-1} y) < \epsilon \text{ for } 0 \leq j \leq \lceil (a_1 + \ldots + a_i) n \rceil - 1, \right. \\ \left. i = 1, \ldots, k \right\},$$

where $\lceil u \rceil$ denotes the least integer $\geq u$. We call $B_n^{\mathbf{a}}(x, \epsilon)$ the n-th $\mathbf{a}$*-weighted Bowen ball of radius ϵ centered at x.*

Following the approaches of Bowen [8] and Pesin-Pitskel' [32] in defining topological entropies and topological pressures of non-compact subsets [8], and in which replacing Bowen balls by $\mathbf{a}$-weighted Bowen balls, we can define the notions of $\mathbf{a}$-weighted topological entropy and $\mathbf{a}$-weighted topological pressure, respectively. To be concise, in this section we only give the definition of **a**-weighted topological entropy. The definition of **a**-weighted topological pressure will be given in Section 3.1.

Let $Z \subset X_1$ and $\epsilon > 0$. We say that an at most countable collection of **a**-weighted Bowen balls $\Gamma = \{B^{\mathbf{a}}_{n_j}(x_j, \epsilon)\}_j$ *covers* Z if $Z \subset \bigcup_j B^{\mathbf{a}}_{n_j}(x_j, \epsilon)$. For $\Gamma = \{B^{\mathbf{a}}_{n_j}(x_j, \epsilon)\}_j$, put $n(\Gamma) = \min_j n_j$. Let $s \geq 0$ and define

$$\Lambda^{\mathbf{a},s}_{N,\epsilon}(Z) = \inf \sum_j \exp(-sn_j),$$

where the infinum is taken over all collections $\Gamma = \{B^{\mathbf{a}}_{n_j}(x_j, \epsilon)\}$ covering Z, such that $n(\Gamma) \geq N$. The quantity $\Lambda^{\mathbf{a},s}_{N,\epsilon}(Z)$ does not decrease with N, hence the following limit exists:

$$\Lambda^{\mathbf{a},s}_{\epsilon}(Z) = \lim_{N\to\infty} \Lambda^{\mathbf{a},s}_{N,\epsilon}(Z).$$

There exists a critical value of the parameter s, which we will denote by $h^{\mathbf{a}}_{\text{top}}(T_1, Z, \epsilon)$, where $\Lambda^{\mathbf{a},s}_{\epsilon}(Z)$ jumps from ∞ to 0, i.e.

$$\Lambda^{\mathbf{a},s}_{\epsilon}(Z) = \begin{cases} 0, & s > h^{\mathbf{a}}_{\text{top}}(T_1, Z, \epsilon), \\ \infty, & s < h^{\mathbf{a}}_{\text{top}}(T_1, Z, \epsilon). \end{cases}$$

It is clear to see that $h^{\mathbf{a}}_{\text{top}}(T_1, Z, \epsilon)$ does not decrease with ϵ, and hence the following limit exists,

$$h^{\mathbf{a}}_{\text{top}}(T_1, Z) = \lim_{\epsilon\to 0} h^{\mathbf{a}}_{\text{top}}(T_1, Z, \epsilon).$$

Definition 1.3. *We call $h^{\mathbf{a}}_{\text{top}}(T_1, Z)$ the* **a***-weighted topological entropy of T_1 restricted to Z or, simply, the* **a***-weighted topological entropy of Z, when there is no confusion about T_1. In particular we write $h^{\mathbf{a}}_{\text{top}}(T_1)$ for $h^{\mathbf{a}}_{\text{top}}(T_1, X_1)$.*

Similarly we will define the **a**-weighted topological pressure $P^{\mathbf{a}}(T_1, f)$ of continuous functions f on X_1 (see Section 3.1). In the particular case when $f \equiv 0$, we have $P^{\mathbf{a}}(T_1, 0) = h^{\mathbf{a}}_{\text{top}}(T_1)$. The main result of this paper is the following variational principle for weighted topological pressure.

Theorem 1.4. *Let $f \in C(X_1)$. Then*

$$P^{\mathbf{a}}(T_1, f) = \sup\left\{\int f d\mu + h^{\mathbf{a}}_{\mu}(T_1) : \ \mu \in \mathcal{M}(X_1, T_1)\right\}. \tag{1.4}$$

In Section 6, we will extend the above theorem to the case that f is a sub-additive potential. As a corollary, taking $f \equiv 0$ in Theorem 1.4, we obtain the following variational principle for weighted topological entropy.

Corollary 1.5. $h^{\mathbf{a}}_{\text{top}}(T_1) = \sup\{h^{\mathbf{a}}_{\mu}(T_1) : \ \mu \in \mathcal{M}(X_1, T_1)\}$.

Theorem 1.4 and Corollary 1.5 provide as weighted versions of Ruelle-Walters' variational principle for topological pressure, and Goodwyn-Dinaburg-Goodman's variational principle for topological entropy (cf. [37]). They are also topological extensions of Kenyon-Peres' variational principle for Hausdorff dimension of toral affine invariant sets. Indeed, consider the aforementioned factor map π between the toral dynamics (K, T) and $(\pi(K), S)$ and let $a_1 = \frac{1}{\log m_2}$, $a_2 = \frac{1}{\log m_1} - \frac{1}{\log m_2}$. It is easy to see from our definition that $h_{\text{top}}^{(a_1,a_2)}(T, K)$ simply coincides with $\dim_H K$, and hence Corollary 1.5 recovers (1.2) and its higher dimensional versions given in [20]. Moreover, by Corollary 1.5, we can generalize (1.2) to a class of skew-product expanding maps on the k-torus (see Section 7.2 for details).

The proof of Theorem 1.4 is quite sophisticated. Besides adopting some ideas from [36, 28] and [20], we also introduce substantially new ideas in the proof. For the convenience of the readers, in the following we illustrate a rough outline of our proof.

To see the lower bound in (1.4), we first prove that for each ergodic measure $\mu \in \mathcal{M}(X_1, T_1)$,

$$\lim_{\epsilon \to 0} \liminf_{n \to +\infty} \frac{-\log \mu(B_n^{\mathbf{a}}(x, \epsilon))}{n} = \lim_{\epsilon \to 0} \limsup_{n \to +\infty} \frac{-\log \mu(B_n^{\mathbf{a}}(x, \epsilon))}{n} = h_\mu^{\mathbf{a}}(T_1) \tag{1.5}$$

for μ-a.e. $x \in X_1$. The above formula is not only a weighted version of Brin-Katok's Theorem [7] on local entropy, but also a topological extension of the Ledrappier-Young dimension formula (1.3). The justification of (1.5) is mainly adapted from Kenyon-Peres' proof of (1.3) in [20] and Brin-Katok's argument in [7]. Based on (1.5), the lower bound in (1.4) follows from a simple covering argument.

The proof of the upper bound in (1.4) is more complicated. First we apply the techniques in geometric measure theory to prove the following "dynamical" Frostman lemma: for any $0 < s < P^{\mathbf{a}}(T_1, f)$ and small enough $\epsilon > 0$, there exists a Borel probability measure ν on X_1 and $N \in \mathbb{N}$ such that

$$\nu(B_n^{\mathbf{a}}(x, \epsilon)) \leq \exp\left(-sn + \frac{1}{a_1} S_{\lceil a_1 n \rceil} f(x)\right), \quad \forall x \in X_1,\ n \geq N, \tag{1.6}$$

where $S_n f(x) := \sum_{i=0}^{n-1} f(T_1^i x)$. This is a key part in our proof. Notice that there exists a small $\tau \in (0, \epsilon)$ such that for any Borel partition α_i of X_i with $\text{diam}(\alpha_i) < \tau$, $i = 1, \ldots, k$, we have

$$\bigvee_{i=1}^{k} \bigvee_{j=t_{i-1}(n)}^{t_i(n)-1} T_1^{-j} \pi_{i-1}^{-1} \alpha_i(x) \subseteq B_n^{\mathbf{a}}(x, \epsilon), \quad \forall x \in X_1,\ n \geq N,$$

where $t_0(n) = 0$, $t_i(n) = \lceil (a_1 + \ldots + a_i)n \rceil$, and $\vee$ stands for the join of partitions. Hence (1.6) implies that

$$\sum_{i=1}^{k} H_\nu\Big(\bigvee_{j=t_{i-1}(n)}^{t_i(n)-1} T_1^{-j} \pi_{i-1}^{-1} \alpha_i \Big) \geq sn - \int \frac{1}{a_1} S_{\lceil a_1 n \rceil} f(x) d\nu(x). \tag{1.7}$$

Then, as another key part, we use (1.7) and entropy theory to show the existence of a T_1-invariant measure μ on X_1 such that $h_\mu^{\mathbf{a}}(T_1) > s - \int f d\mu$, from which the upper bound follows. In the proof of this part, a combinatoric lemma (see Lemma 5.4) established by Kenyon-Peres [20] plays an important role; besides this, we also use a delicate compactness argument based on the upper semi-continuity of certain entropy functions, and adopt some ideas from [36, 28] as well. Reducing back to the aforementioned toral dynamics, our approach provides a new proof for the upper bound in Kenyon-Peres' variational principle (1.2).

The paper is organized as follows. In Section 2, we prove the upper semi-continuity of certain entropy functions. In Section 3, we define weighted topological pressure for continuous functions and more generally for sub-additive potentials; we also establish a dynamical Frostman lemma for the weighted topological pressure. In Sections 4-5, we prove respectively the lower and upper bounds of Theorem 1.4. In Section 6, we extend Theorem 1.4 to the sub-additive case. In Section 7, we give some remarks, examples and questions. In Appendix A, we prove the formula (1.5).

2. Upper semi-continuity of certain entropy functions

In this section, we prove the upper semi-continuity of certain entropy functions (see Lemma 2.3), which is needed in our proof of the upper bound part of Theorem 1.4. We begin with the following.

Definition 2.1. *Let Z be a compact metric space. A function $f : Z \to [-\infty, +\infty)$ is called upper semi-continuous if one of the following equivalent conditions holds:*

- (C1) $\limsup_{z_N \to z} f(z_N) \leq f(z)$ *for each $z \in Z$;*
- (C2) *for each $r \in \mathbb{R}$ the set $\{z \in Z : f(z) \geq r\}$ is closed.*

By (C2), the infimum of any family of upper semi-continuous functions is again an upper semi-continuous function; both the sum and supremum of finitely many upper semi-continuous functions are upper semi-continuous functions.

Lemma 2.2. *Let Z be a compact metric space and $f : Z \to [-\infty, +\infty)$ be an upper semi-continuous function. Then for any $\mu \in \mathcal{M}(Z)$,*

$$\inf_{g\in C(Z), g\geq f} \int_Z g(z)d\mu(z) = \int_Z f(z)d\mu(z). \tag{2.1}$$

Proof. It is well known that the equality (2.1) holds when f is a real-valued upper semi-continuous function (see e.g. [12, Appendix (A7)] for a proof). In the following we assume that f is an upper semi-continuous function taking values in $[-\infty, +\infty)$.

By the upper semi-continuity of f, we have $\sup_{z\in Z} f(z) = \max_{z\in Z} f(z) < +\infty$. Thus $\int_Z f(z)d\mu(z)$ is well defined and $\int_Z f(z)d\mu(z) \in [-\infty, +\infty)$.

For $M \in \mathbb{N}$, let $f_M(z) = \max\{f(z), -M\}$ for $z \in Z$. Then f_M is an upper semi-continuous real-valued function, and thus

$$\inf_{g\in C(Z), g\geq f_M} \int_Z g(z)d\mu(z) = \int_Z f_M(z)d\mu(z).$$

Since

$$\sup_{M\in\mathbb{N}} \sup_{z\in Z} f_M(z) \leq \max\left\{\max_{z\in Z} f(z), 0\right\} < +\infty$$

and $f_M(z) \searrow f(z)$ as $M \to +\infty$ for any $z \in Z$, one has

$$\lim_{M\to+\infty} \int_Z f_M(z)d\mu(z) = \int_Z \lim_{M\to+\infty} f_M(z)d\mu(z) = \int_Z f(z)d\mu(z)$$

by Lebesgue's monotone convergence theorem. Moreover

$$\begin{aligned}
\inf_{g\in C(Z), g\geq f} \int_Z g(z)d\mu(z) &= \inf_{M\in\mathbb{N}} \left\{\inf_{g\in C(Z), g\geq f_M} \int_Z g(z)d\mu(z)\right\} \\
&= \inf_{M\in\mathbb{N}} \int_Z f_M(z)d\mu(z) \\
&= \lim_{M\to+\infty} \int_Z f_M(z)d\mu(z) \\
&= \int_Z f(z)d\mu(z).
\end{aligned}$$

This completes the proof of the lemma. □

Let (X, T) be a TDS with a compatible metric d. For $\epsilon > 0$ and $M \in \mathbb{N}$, we define

$$\mathcal{P}_X(\epsilon, M) = \{\alpha : \alpha \text{ is a finite Borel partition of } X \text{ with } \operatorname{diam}(\alpha) < \epsilon, \#(\alpha) \leq M\}, \tag{2.2}$$

where $\operatorname{diam}(\alpha) := \max_{A\in\alpha} \operatorname{diam}(A)$, and $\#(\alpha)$ stands for the cardinality of α. Then we define

$$\mathcal{P}_X(\epsilon) = \{\alpha : \alpha \text{ is a finite Borel partition of } X \text{ with } \operatorname{diam}(\alpha) < \epsilon\}.$$

It is clear that for any $\epsilon > 0$, there exists $N := N(\epsilon) \in \mathbb{N}$ such that $\mathcal{P}_X(\epsilon, M) \neq \emptyset$ for any $M \geq N$. The main result of this section is the following.

Lemma 2.3. *Let (X, T) be a TDS and $\epsilon > 0$. Then*

(1) *If $M \in \mathbb{N}$ with $\mathcal{P}_X(\epsilon, M) \neq \emptyset$, then the map*

$$\theta \in \mathcal{M}(X) \mapsto H_\theta(\epsilon, M; \ell) := \inf_{\alpha\in\mathcal{P}_X(\epsilon,M)} \frac{1}{\ell} H_\theta\left(\bigvee_{i=0}^{\ell-1} T^{-i}\alpha\right) \tag{2.3}$$

is upper semi-continuous from $\mathcal{M}(X)$ to $[0, \log M]$ for each $\ell \in \mathbb{N}$.

(2) *The map*

$$\theta \in \mathcal{M}(X) \mapsto H_\theta(\epsilon; \ell) := \inf_{\alpha\in\mathcal{P}_X(\epsilon)} \frac{1}{\ell} H_\theta\left(\bigvee_{i=0}^{\ell-1} T^{-i}\alpha\right)$$

is a bounded upper semi-continuous non-negative function for each $\ell \in \mathbb{N}$.

(3) *The map*

$$\mu \in \mathcal{M}(X, T) \mapsto h_\mu(T, \epsilon) := \inf_{\alpha\in\mathcal{P}_X(\epsilon)} h_\mu(T, \alpha)$$

is a bounded upper semi-continuous non-negative function.

Proof. We first prove (1). Let $M \in \mathbb{N}$ with $\mathcal{P}_X(\epsilon, M) \neq \emptyset$, and $\ell \in \mathbb{N}$. Clearly, the map $H_\bullet(\epsilon, M; \ell)$ is defined from $\mathcal{M}(X)$ to $[0, \log M]$. Let $\theta_0 \in \mathcal{M}(X)$. It is sufficient to show that the map $H_\bullet(\epsilon, M; \ell)$ is upper semi-continuous at θ_0.

Let $\delta > 0$. Then there exists $\alpha \in \mathcal{P}_X(\epsilon, M)$ such that

$$\frac{1}{\ell} H_{\theta_0}\left(\bigvee_{i=0}^{\ell-1} T^{-i}\alpha\right) \leq H_{\theta_0}(\epsilon, M; \ell) + \delta. \tag{2.4}$$

Let $\alpha = \{A_1, \ldots, A_u\}$. Then $u \leq M$ and $\operatorname{diam}(A_i) < \epsilon$ for $i = 1, 2, \ldots, u$. By Lemma 4.15 in [37], there exists $\delta_1 = \delta_1(u, \delta) > 0$ such that whenever $\gamma_1 = \{E_1, \ldots, E_u\}, \gamma_2 = \{F_1, \ldots, F_u\}$ are two Borel partitions of X with $\sum_{j=1}^{u} \sum_{i=0}^{\ell-1} \theta_0 \circ T^{-i}(E_j \Delta F_j) < \delta_1$, then

$$\begin{aligned}(2.5)\qquad &\frac{1}{\ell}\left|H_{\theta_0}\left(\bigvee_{i=0}^{\ell-1}T^{-i}\gamma_1\right)-H_{\theta_0}\left(\bigvee_{i=0}^{\ell-1}T^{-i}\gamma_2\right)\right| \\ &\le \frac{1}{\ell}\sum_{i=0}^{\ell-1}|H_{\theta_0\circ T^{-i}}(\gamma_1|\gamma_2)+H_{\theta_0\circ T^{-i}}(\gamma_2|\gamma_1)|<\delta.\end{aligned}$$

Write $\eta=\sum_{i=0}^{\ell-1}\theta_0\circ T^{-i}$. Next, we are going to construct a Borel partition $\beta=\{B_1,\dots,B_u\}$ of X so that $\mathrm{diam}(\beta)<\epsilon$, $\sum_{j=1}^{u}\eta(A_j\Delta B_j)<\delta_1$ and $\eta(\partial\beta)=0$.

In fact, note that $\eta(X)=\ell<\infty$, hence η is regular on X. Thus there exist open subsets V_j of X such that $A_j\subseteq V_j$, $\mathrm{diam}(V_j)<\epsilon$ and $\eta(V_j\setminus A_j)<\frac{\delta_1}{u^2}$ for $j=1,\dots,u$. Clearly, $\mathcal{V}:=\{V_1,\dots,V_u\}$ is an open cover. Let $t>0$ be a Lebesgue number of $\mathcal{V}$. For any $x\in X$, there exists $0<t_x\le\frac{t}{3}$ such that $\eta(\partial B(x,t_x))=0$. Thus $\{B(x,t_x):x\in X\}$ forms an open cover of X. Take its finite subcover $\{B(x_i,t_{x_i})\}_{i=1}^{r}$, that is, $\bigcup_{i=1}^{r}B(x_i,t_{x_i})=X$. Obviously, each $B(x_i,t_{x_i})$ is a subset of some $V_{j(i)}$, $j(i)\in\{1,\dots,u\}$ since $t_{x_i}\le\frac{t}{3}$.

Let $I_j=\{i\in\{1,\dots,r\}:B(x_i,t_{x_i})\subset V_j\}$ for $j=1,\dots,u$. Then $\bigcup_{j=1}^{u}I_j=\{1,\dots,r\}$. Put $B_1=\bigcup_{i\in I_1}B(x_i,t_{x_i})$ and $B_j=\left(\bigcup_{i\in I_j}B(x_i,t_{x_i})\right)\setminus\bigcup_{m=1}^{j-1}B_m$ inductively for $j=2,\dots,u$. It is clear that $\beta=\{B_1,\dots,B_u\}$ is a Borel partition of X with $B_j\subseteq V_j$ and $\eta(\partial B_j)=0$ for $j=1,\dots,u$. Now for each $j\in\{1,\dots,u\}$,

$$\begin{aligned}A_j\Delta B_j&=(B_j\setminus A_j)\cup(A_j\cap(X\setminus B_j))\subseteq(V_j\setminus A_j)\cup\bigcup_{k\ne j}(A_j\cap B_k)\\ &\subseteq(V_j\setminus A_j)\cup\bigcup_{k\ne j}(A_j\cap V_k)\subseteq(V_j\setminus A_j)\cup\bigcup_{k\ne j}(A_j\cap(V_k\setminus A_k))\\ &\subseteq\bigcup_{k=1}^{u}(V_k\setminus A_k).\end{aligned}$$

Thus $\sum_{j=1}^{u}\eta(A_j\Delta B_j)\le u\sum_{k=1}^{u}\eta(V_k\setminus A_k)<\delta_1$.

Summing up, we have constructed a Borel partition $\beta=\{B_1,\dots,B_u\}\in\mathcal{P}_X(\epsilon,M)$ so that $\sum_{j=1}^{u}\eta(B_j\Delta A_j)<\delta_1$ and $\eta(\partial\beta)=0$. Now on the one hand, by (2.5) and (2.4), we have

$$\frac{1}{\ell}H_{\theta_0}\left(\bigvee_{i=0}^{\ell-1}T^{-i}\beta\right)\le\frac{1}{\ell}H_{\theta_0}\left(\bigvee_{i=0}^{\ell-1}T^{-i}\alpha\right)+\delta\le H_{\theta_0}(\epsilon,M;\ell)+2\delta.$$

On the other hand, since $\eta(\partial\beta)=0$, one has $\theta_0(T^{-i}\partial\beta)=0$ for $i=0,1,\cdots,\ell-1$. As $\partial T^{-i}A\subseteq T^{-i}\partial A$ for any $A\subseteq X$, one has $\theta_0(\partial T^{-i}\beta)=0$ for $i=0,1,\cdots,\ell-1$. Moreover note that $\partial(A\cap B)\subseteq(\partial A)\cap(\partial B)$ for any $A,B\subseteq X$, we have $\theta_0(\partial(\bigvee_{i=0}^{\ell-1}T^{-i}\beta))=0$. Thus the map $\theta\in\mathcal{M}(X)\mapsto\frac{1}{\ell}H_\theta(\bigvee_{i=0}^{\ell-1}T^{-i}\beta)$ is continuous at the point θ_0. Therefore

$$\begin{aligned}\limsup_{\theta\to\theta_0}H_\theta(\epsilon,M;\ell)&\leq\limsup_{\theta\to\theta_0}\frac{1}{\ell}H_\theta\left(\bigvee_{i=0}^{\ell-1}T^{-i}\beta\right)\\&=\frac{1}{\ell}H_{\theta_0}\left(\bigvee_{i=0}^{\ell-1}T^{-i}\beta\right)\\&\leq H_{\theta_0}(\epsilon,M;\ell)+2\delta.\end{aligned}$$

Finally letting $\delta\searrow 0$, we see that the map $H_\bullet(\epsilon,M;\ell)$ is upper semi-continuous at θ_0. This completes the proof of (1).

Now we turn to the proof of (2). Let $\ell\in\mathbb{N}$. Since $\mathcal{P}_X(\epsilon)=\bigcup_{M\in\mathbb{N},\mathcal{P}_X(\epsilon,M)\neq\emptyset}\mathcal{P}_X(\epsilon,M)$, we have

$$H_\theta(\epsilon;\ell)=\inf_{M\in\mathbb{N},\mathcal{P}_X(\epsilon,M)\neq\emptyset}H_\theta(\epsilon,M;\ell)$$

for $\theta\in\mathcal{M}(X)$. Moreover, by (1) and the fact that the infimum of any family of upper semi-continuous functions is again an upper semi-continuous one, we know that the map

$$\theta\in\mathcal{M}(X)\mapsto H_\theta(\epsilon;\ell):=\inf_{\alpha\in\mathcal{P}_X(\epsilon)}\frac{1}{\ell}H_\theta\left(\bigvee_{i=0}^{\ell-1}T^{-i}\alpha\right)$$

is a bounded upper semi-continuous non-negative function. This proves (2).

In the end we prove (3). Note that

$$\begin{aligned}h_\mu(T,\epsilon)&=\inf_{\alpha\in\mathcal{P}_X(\epsilon)}h_\mu(T,\alpha)=\inf_{\alpha\in\mathcal{P}_X(\epsilon)}\inf_{\ell\geq1}\frac{1}{\ell}H_\mu\left(\bigvee_{i=0}^{\ell-1}T^{-i}\alpha\right)\\&=\inf_{\ell\geq1}\inf_{\alpha\in\mathcal{P}_X(\epsilon)}\frac{1}{\ell}H_\mu\left(\bigvee_{i=0}^{\ell-1}T^{-i}\alpha\right)=\inf_{\ell\geq1}H_\mu(\epsilon;\ell)\end{aligned}$$

for $\mu\in\mathcal{M}(X,T)$. Using (2) and the fact that the infimum of any family of upper semi-continuous functions is again an upper semi-continuous one, we know that the map

$$\mu\in\mathcal{M}(X,T)\mapsto h_\mu(T,\epsilon)$$

is a bounded upper semi-continuous non-negative function. This completes the proof of the lemma. □

3. Weighted topological pressures and a dynamical Frostman lemma

In this section we introduce the definition of weighted topological pressure for (asymptotically) sub-additive potentials for general topological dynamical systems. Moreover, using some ideas from geometric measure theory, we establish a dynamical Frostman lemma (see Lemma 3.3) for weighted topological pressure, which plays a key role in our proof of Theorem 1.4.

3.1. Weighted topological pressures for sub-additive potentials.

Assume that (X, T) is a TDS. We say that a sequence $\Phi = \{\log \phi_n\}_{n=1}^{\infty}$ of functions on X is a *sub-additive potential* if each ϕ_n is an upper semi-continuous nonnegative-valued function on X such that

$$0 \le \phi_{n+m}(x) \le \phi_n(x)\phi_m(T^n x), \qquad \forall\, x \in X,\ m, n \in \mathbb{N}. \tag{3.1}$$

In particular, Φ is called *additive* if each ϕ_n is a continuous positive-valued function so that $\phi_{n+m}(x) = \phi_n(x)\phi_m(T^n x)$ for all $x \in X$ and $m, n \in \mathbb{N}$; in this case, there is a continuous real function g on X such that $\phi_n(x) = \exp(\sum_{i=0}^{n-1} g(T^i x))$ for each n.

Let $k \ge 2$. Assume that (X_i, d_i), $i = 1, \ldots, k$, are compact metric spaces, and (X_i, T_i) are TDS's. Moreover, assume that for each $1 \le i \le k-1$, (X_{i+1}, T_{i+1}) is a factor of (X_i, T_i) with a factor map $\pi_i : X_i \to X_{i+1}$.

Let $\mathbf{a} = (a_1, \ldots, a_k) \in \mathbb{R}^k$ with $a_1 > 0$ and $a_i \ge 0$ for $2 \le i \le k$. For any $n \in \mathbb{N}$ and $\epsilon > 0$, define

$$\mathcal{T}^{\mathbf{a}}_{n,\epsilon} := \{A \subset X_1 :\ A \text{ is Borel subset of } B^{\mathbf{a}}_n(x, \epsilon) \text{ for some } x \in X_1\}, \tag{3.2}$$

where $B^{\mathbf{a}}_n(x, \epsilon)$ is defined as in Definition 1.2.

Let $\Phi = \{\log \phi_n\}_{n=1}^{\infty}$ be a sub-additive potential on X_1. Let $Z \subseteq X_1$, $s \ge 0$ and $N \in \mathbb{N}$, define

$$\Lambda^{\mathbf{a},s}_{\Phi,N,\epsilon}(Z) = \inf \sum_j \exp\left(-s n_j + \frac{1}{a_1} \sup_{x \in A_j} \phi_{\lceil a_1 n_j \rceil}(x)\right),$$

where the infimum is taken over all countable collections $\Gamma = \{(n_j, A_j)\}_j$ with $n_j \ge N$, $A_j \in \mathcal{T}^{\mathbf{a}}_{n_j,\epsilon}$ and $\bigcup_j A_j \supseteq Z$. The quantity $\Lambda^{\mathbf{a},s}_{\Phi,N,\epsilon}(Z)$ does not decrease with N, hence the following limit exists:

$$\Lambda^{\mathbf{a},s}_{\Phi,\epsilon}(Z) = \lim_{N \to \infty} \Lambda^{\mathbf{a},s}_{\Phi,N,\epsilon}(Z).$$

There exists a critical value of the parameter s, which we will denote by $P^{\mathbf{a}}(T_1, \Phi, Z, \epsilon)$, where $\Lambda^{\mathbf{a},s}_{\Phi,\epsilon}(Z)$ jumps from ∞ to 0, i.e.

$$\Lambda^{\mathbf{a},s}_{\Phi,\epsilon}(Z) = \begin{cases} 0, & s > P^{\mathbf{a}}(T_1, \Phi, Z, \epsilon), \\ \infty, & s < P^{\mathbf{a}}(T_1, \Phi, Z, \epsilon). \end{cases}$$

Clearly $P^{\mathbf{a}}(T_1, \Phi, Z, \epsilon)$ does not decrease with ϵ, and hence the following limit exists,

$$P^{\mathbf{a}}(T_1, \Phi, Z) = \lim_{\epsilon \to 0} P^{\mathbf{a}}(T_1, \Phi, Z, \epsilon).$$

Definition 3.1. *We call $P^{\mathbf{a}}(T_1, \Phi) := P^{\mathbf{a}}(T_1, \Phi, X_1)$ the $\mathbf{a}$-weighted topological pressure of Φ with respect to T_1 or, simply, the $\mathbf{a}$-weighted topological pressure of Φ, when there is no confusion about T_1.*

Definition 3.2. *Let $f \in C(X_1)$. Define $\Phi = \{\log \phi_n\}_{n=1}^{\infty}$ by $\phi_n(x) = \exp(\sum_{j=0}^{n-1} f(T_1^j x))$. In this case, Φ is additive. We just define $P^{\mathbf{a}}(T_1, f) := P^{\mathbf{a}}(T_1, \Phi)$.*

Taking $f \equiv 0$, one can see that $P^{\mathbf{a}}(T_1, 0) = h^{\mathbf{a}}_{\text{top}}(T_1)$. Let $\Phi = \{\log \phi_n\}_{n=1}^{\infty}$ be a sub-additive potential on X_1. For any $\mu \in \mathcal{M}(X_1, T_1)$, define

$$\Phi_*(\mu) := \lim_{n \to \infty} \int \frac{\log \phi_n(x)}{n} \, d\mu(x). \tag{3.3}$$

This limit always exists and takes values in $\mathbb{R} \cup \{-\infty\}$ (cf. [36, Theorem 10.1]).

In our proof of Theorem 1.4, we need the following dynamical Frostman lemma.

Lemma 3.3. *Let $\Phi = \{\log \phi_n\}_{n=1}^{\infty}$ be a sub-additive potential on X_1. Suppose that $P^{\mathbf{a}}(T_1, \Phi) > 0$. Then for any $0 < s < P^{\mathbf{a}}(T_1, \Phi)$, there exist a Borel probability measure ν on X_1 and $\epsilon > 0$, $N \in \mathbb{N}$ such that for any $x \in X_1$ and $n \geq N$ we have*

$$\nu(B^{\mathbf{a}}_n(x, \epsilon)) \leq \exp(-sn) \sup_{y \in B^{\mathbf{a}}_n(x,\epsilon)} (\phi_{\lceil a_1 n \rceil}(y))^{1/a_1}. \tag{3.4}$$

A non-weighted version of the above lemma was first proved by the authors in the particular case when $\phi_n \equiv 1$ (see [17, Lemma 3.4]), using some ideas and techniques in geometric measure theory. In the remainder of this section, we will give the detailed proof of Lemma 3.3, by adapting and elaborating the approach in [17]. A key ingredient of our proof is the notion of average weighted topological pressure, which is an analogue of weight Hausdorff measure in geometric measure theory. The definition of this notion and some of its properties will be given in next subsection. In Subsection 3.3, we prove Lemma 3.3.

3.2. **Average weighted topological pressures.** Let $\Phi = \{\log \phi_n\}_{n=1}^{\infty}$ be a sub-additive potential on X_1. For any function $f : X_1 \to [0, \infty)$, for $s \geq 0$ and $N \in \mathbb{N}$, define

$$\mathcal{W}^{\mathbf{a},s}_{\Phi,N,\epsilon}(f) = \inf \sum_j c_j \exp\left(-sn_j + \frac{1}{a_1} \sup_{x \in A_j} \log \phi_{\lceil a_1 n_j \rceil}(x)\right), \tag{3.5}$$

where the infimum is taken over all countable collections $\Gamma = \{(n_j, A_j, c_j)\}_j$ with $n_j \geq N$, $A_j \in \mathcal{T}^{\mathbf{a}}_{n_j,\epsilon}$, $0 < c_j < \infty$, and

$$\sum_j c_j \chi_{A_j} \geq f,$$

where χ_A denotes the characteristic function of A, i.e., $\chi_A(x) = 1$ if $x \in A$ and 0 if $x \in X_1 \backslash A$.

For $Z \subseteq X_1$, we set $\mathcal{W}^{\mathbf{a},s}_{\Phi,N,\epsilon}(Z) = \mathcal{W}^{\mathbf{a},s}_{\Phi,N,\epsilon}(\chi_Z)$. The quantity $\mathcal{W}^{\mathbf{a},s}_{\Phi,N,\epsilon}(Z)$ does not decrease with N, hence the following limit exists:

$$\mathcal{W}^{\mathbf{a},s}_{\Phi,\epsilon}(Z) = \lim_{N \to \infty} \mathcal{W}^{\mathbf{a},s}_{\Phi,N,\epsilon}(Z).$$

There exists a critical value of the parameter s, which we will denote by $P^{\mathbf{a}}_W(T_1, \Phi, Z, \epsilon)$, where $\mathcal{W}^{\mathbf{a},s}_{\Phi,\epsilon}(Z)$ jumps from ∞ to 0, i.e.

$$\mathcal{W}^{\mathbf{a},s}_{\Phi,\epsilon}(Z) = \begin{cases} 0, & s > P^{\mathbf{a}}_W(T_1, \Phi, Z, \epsilon), \\ \infty, & s < P^{\mathbf{a}}_W(T_1, \Phi, Z, \epsilon). \end{cases}$$

Clearly $P^{\mathbf{a}}_W(T_1, \Phi, Z, \epsilon)$ does not decrease with ϵ, and hence the following limit exists,

$$P^{\mathbf{a}}_W(T_1, \Phi, Z) = \lim_{\epsilon \to 0} P^{\mathbf{a}}_W(T_1, \Phi, Z, \epsilon).$$

Definition 3.4. *We call $P^{\mathbf{a}}_W(T_1, \Phi) := P^{\mathbf{a}}_W(T_1, \Phi, X_1)$ the average $\mathbf{a}$-weighted topological pressure of Φ with respect to T_1 or, simply, the average $\mathbf{a}$-weighted topological pressure of Φ, when there is no confusion about T_1.*

The main result of this subsection is the following.

Proposition 3.5. *Let $Z \subseteq X_1$. Then for any $s \geq 0$ and $\epsilon, \delta > 0$, we have*

$$\Lambda^{\mathbf{a},s+\delta}_{\Phi,N,6\epsilon}(Z) \leq \mathcal{W}^{\mathbf{a},s}_{\Phi,N,\epsilon}(Z) \leq \Lambda^{\mathbf{a},s}_{\Phi,N,\epsilon}(Z),$$

when N is large enough. As a consequence, $P^{\mathbf{a}}_W(T_1, \Phi) = P^{\mathbf{a}}(T_1, \Phi)$.

Before giving the proof of Proposition 3.5, we first state some lemmas.

Lemma 3.6. *For any $s \geq 0$, $N \in \mathbb{N}$ and $\epsilon > 0$, both $\Lambda^{\mathbf{a},s}_{\Phi,N,\epsilon}$ and $\mathcal{W}^{\mathbf{a},s}_{\Phi,N,\epsilon}$ are outer measures on X.*

Proof. It follows directly from the definitions $\Lambda^{\mathbf{a},s}_{\Phi,N,\epsilon}$ and $\mathcal{W}^{\mathbf{a},s}_{\Phi,N,\epsilon}$. □

The following combinatoric lemma plays an important role in the proof of Proposition 3.5.

Lemma 3.7. *Let (X,d) be a compact metric space and $\epsilon>0$. Let $(E_i)_{i\in\mathcal{I}}$ be a finite or countable family of subsets of X with diameter less than ϵ, and $(c_i)_{i\in\mathcal{I}}$ a family of positive numbers. Let $t>0$. Assume that $F\subseteq X$ such that*

$$F\subseteq\left\{x\in X:\ \sum_i c_i\chi_{E_i}>t\right\}.$$

Then F can be covered by no more than $\frac{1}{t}\sum_i c_i$ balls with centers in $\bigcup_{i\in\mathcal{I}}E_i$ and radius 6ϵ.

To prove Lemma 3.7, we need the following well known covering lemma.

Lemma 3.8 (cf. Theorem 2.1 in [26])**.** *Let (X,d) be a compact metric space and $\mathcal{B}=\{B(x_i,r_i)\}_{i\in\mathcal{I}}$ be a family of open balls in X. Then there exists a finite or countable subfamily $\mathcal{B}'=\{B(x_i,r_i)\}_{i\in\mathcal{I}'}$ of pairwise disjoint balls in $\mathcal{B}$ such that*

$$\bigcup_{B\in\mathcal{B}}B\subseteq\bigcup_{i\in\mathcal{I}'}B(x_i,5r_i).$$

Proof of Lemma 3.7. Without loss of generality, assume that $\mathcal{I}\subseteq\mathbb{N}$. For any $i\in\mathcal{I}$, pick $x_i\in E_i$ and write $B_i=B(x_i,\epsilon)$ and $5B_i=B(x_i,5\epsilon)$ for short. Clearly $E_i\subseteq B_i$. Define

$$Z=\left\{x\in X:\ \sum_i c_i\chi_{B_i}>t\right\}.$$

We have $F\subset Z$. To prove the lemma, it suffices to show that Z can be covered by no more than $\frac{1}{t}\sum_i c_i$ balls with centers in $\{x_i: i\in\mathcal{I}\}$ and radius 6ϵ. To avoid triviality, we assume that $\sum_i c_i<\infty$; otherwise there is nothing left to prove.

For $k\in\mathbb{N}$, define

$$\mathcal{I}_k=\{i\in\mathcal{I}:\ i\le k\}\quad\text{and}\quad Z_k=\Big\{x\in Z:\ \sum_{i\in\mathcal{I}_k}c_i\chi_{B_i}(x)>t\Big\}.$$

We divide the remaining proof into two small steps.

Step 1. For each $k\in\mathbb{N}$, there exists a finite set $\mathcal{J}_k\subseteq\mathcal{I}_k$ such that the balls B_i $(i\in\mathcal{J}_k)$ are pairwise disjoint, $Z_k\subseteq\bigcup_{i\in\mathcal{J}_k}5B_i$ and

$$\#(\mathcal{J}_k)\le\frac{1}{t}\sum_{i\in\mathcal{I}_k}c_i.$$

To prove the above result, we adopt the argument from Federer [14, 2.10.24] in the study of weighted Hausdorff measures (see also Mattila [26, Lemma 8.16]). Since $\mathcal{I}_k$ is finite, by approximating the c_i's from above, we may assume that each c_i is a positive rational, and then multiplying c_i and t with a common denominator we may assume that each c_i is a positive integer. Let m be the least integer with $m \geq t$. Denote $\mathcal{B} = \{B_i,\ i \in \mathcal{I}_k\}$ and define $u : \mathcal{B} \to \mathbb{N}$ by $u(B_i) = c_i$. We define by induction integer-valued functions $v_0, v_1, \ldots, v_m$ on $\mathcal{B}$ and sub-families $\mathcal{B}_1, \ldots, \mathcal{B}_m$ of $\mathcal{B}$ starting with $v_0 = u$. Using Lemma 3.8 we find a pairwise disjoint subfamily $\mathcal{B}_1$ of $\mathcal{B}$ such that $\bigcup_{B\in\mathcal{B}} B \subseteq \bigcup_{B\in\mathcal{B}_1} 5B$, and hence $Z_k \subseteq \bigcup_{B\in\mathcal{B}_1} 5B$. Then by repeatedly using Lemma 3.8, we can define inductively for $j = 1, \ldots, m$, disjoint subfamilies $\mathcal{B}_j$ of $\mathcal{B}$ such that

$$\mathcal{B}_j \subseteq \{B \in \mathcal{B} :\ v_{j-1}(B) \geq 1\}, \quad Z_k \subseteq \bigcup_{B\in\mathcal{B}_j} 5B$$

and the functions v_j such that

$$v_j(B) = \begin{cases} v_{j-1}(B) - 1 & \text{for } B \in \mathcal{B}_j, \\ v_{j-1}(B) & \text{for } B \in \mathcal{B}\backslash\mathcal{B}_j. \end{cases}$$

This is possible since for $j < m$, $Z_k \subseteq \left\{x : \sum_{B\in\mathcal{B}:\ B\ni x} v_j(B) \geq m - j\right\}$, whence every $x \in Z_k$ belongs to some ball $B \in \mathcal{B}$ with $v_j(B) \geq 1$. Thus

$$\begin{aligned} \sum_{j=1}^{m} \#(\mathcal{B}_j) &= \sum_{j=1}^{m} \sum_{B\in\mathcal{B}_j} (v_{j-1}(B) - v_j(B)) = \sum_{B\in\mathcal{B}_j} \sum_{j=1}^{m} (v_{j-1}(B) - v_j(B)) \\ &\leq \sum_{B\in\mathcal{B}} \sum_{j=1}^{m} (v_{j-1}(B) - v_j(B)) \leq \sum_{B\in\mathcal{B}} u(B) = \sum_{i\in\mathcal{I}_k} c_i. \end{aligned}$$

Choose $j_0 \in \{1, \ldots, m\}$ so that $\#(\mathcal{B}_{j_0})$ is the smallest. Then

$$\#(\mathcal{B}_{j_0}) \leq \frac{1}{m} \sum_{i\in\mathcal{I}_k} c_i \leq \frac{1}{t} \sum_{i\in\mathcal{I}_k} c_i.$$

Hence $\mathcal{J}_k := \{i \in \mathcal{I}_k :\ B_i \in \mathcal{B}_{j_0}\}$ is desired.

Step 2. There exists $\mathcal{I}' \subset \mathcal{I}$ with $\#(\mathcal{I}') \leq \frac{1}{t} \sum_{i\in\mathcal{I}} c_i$ so that $Z \subseteq \bigcup_{i\in\mathcal{I}'} 6B_i$.

Since $Z_k \uparrow Z$, $Z_k \neq \emptyset$ when k is large enough. Let $\mathcal{J}_k$ be constructed as in Step 1. Then $\mathcal{J}_k \neq \emptyset$ when k is large enough. Define $G_k = \{x_i :\ i \in \mathcal{J}_k\}$. Then

$$\#(G_k) = \#(\mathcal{J}_k) \leq \frac{1}{t} \sum_{i\in\mathcal{I}_k} c_i \leq \frac{1}{t} \sum_{i\in\mathcal{I}} c_i.$$

Since the space of non-empty compact subsets of X is compact with respect to the Hausdorff distance (cf. Federer [14, 2.10.21]), there is a subsequence (k_j) of natural numbers and a non-empty compact set $G \subset X$ such that G_{k_j} converges to G in the

Hausdorff distance as $j \to \infty$. As any two different points in G_k have a distance not less than ϵ, so do the points in G. Thus G is a finite set, moreover, $\#(G_{k_j}) = \#(G)$ when j is large enough. Hence

$$\bigcup_{x \in G} B(x, 5.5\epsilon) \supseteq \bigcup_{x \in G_{k_j}} B(x, 5\epsilon) = \bigcup_{i \in \mathcal{J}_{k_j}} 5B_i \supseteq Z_{k_j}$$

when j is large enough, and thus $\bigcup_{x \in G} B(x, 5.5\epsilon) \supseteq Z$. On the other hand, when j is large enough, we have

$$\bigcup_{x' \in G_{k_j}} B(x', 6\epsilon) \supseteq \bigcup_{x \in G} B(x, 5.5\epsilon),$$

hence we have $\bigcup_{x' \in G_{k_j}} B(x', 6\epsilon) \supseteq Z$, with $\#(G_{k_j}) \le \frac{1}{t} \sum_{i \in \mathcal{I}} c_i$. □

Return back to the metric spaces (X_i, d_i) and TDS's (X_i, T_i), $i = 1, \ldots, k$. For $n \in \mathbb{N}$, define a metric $d_n^{\mathbf{a}}$ on X_1 by

$$d_n^{\mathbf{a}}(x, y) = \sup \left\{ d_i(T_i^j \tau_{i-1} x, T_i^j \tau_{i-1} y) : \ 1 \le i \le k, \ 0 \le j \le \lceil (a_1 + \ldots + a_i) n \rceil - 1 \right\}.$$

Lemma 3.9. *Let $\epsilon > 0$. Then there exist $\gamma > 0$ such that for any $n \in \mathbb{N}$, X_1 can be covered by no more than $\exp(n\gamma)$ balls of radius ϵ in metric $d_n^{\mathbf{a}}$.*

Proof. By compactness, for each $1 \le i \le k$, we can find a finite open cover α_i of X_i with $\mathrm{diam}(\alpha_i) < \epsilon$ (in metric d_1). Let $n > 0$. Define

$$\beta = \bigvee_{i=1}^{k} \left(\bigvee_{j=0}^{\lceil (a_1 + \cdots + a_i) n \rceil - 1} T_1^{-j} \tau_{i-1}^{-1} \alpha_i \right).$$

Then β is an open cover of X_1 with diameter less than ϵ (with respect to the metric $d_n^{\mathbf{a}}$). Hence X_1 can be covered by at most $\#(\beta)$ many balls of radius ϵ in metric $d_n^{\mathbf{a}}$. Let $\gamma > 0$ so that $\exp(\gamma) = \prod_{i=1}^{k} (\#(\alpha_i))^{a_1 + \cdots + a_i + 1}$. Then

$$\#(\beta) \le \prod_{i=1}^{k} (\#(\alpha_i))^{\lceil (a_1 + \cdots + a_i) n \rceil} \le \exp(n\gamma),$$

which implies the result of the lemma. □

Proof of Proposition 3.5. Let $Z \subseteq X_1$, $s \ge 0$, $\epsilon, \delta > 0$. Taking $f = \chi_Z$ and $c_i \equiv 1$ in the definition (3.5), we see that $\mathcal{W}_{\Phi, N, \epsilon}^{\mathbf{a}, s}(Z) \le \Lambda_{\Phi, N, \epsilon}^{\mathbf{a}, s}(Z)$ for each $N \in \mathbb{N}$. In the following, we prove that $\Lambda_{\Phi, N, 6\epsilon}^{\mathbf{a}, s+\delta}(Z) \le \mathcal{W}_{\Phi, N, \epsilon}^{\mathbf{a}, s}(Z)$ when N is large enough.

Let $\gamma > 0$ be given as in Lemma 3.9. Assume that $N \ge 2$ such that

$$n^2 (n+1) e^{\gamma - n\delta} \le 1, \quad \forall \, n \ge N. \tag{3.6}$$

Let $\{(n_i, A_i, c_i)\}_{i\in\mathcal{I}}$ be a family so that $\mathcal{I} \subseteq \mathbb{N}$, $A_i \in \mathcal{T}^{\mathbf{a}}_{n_i,\epsilon}$, $0 < c_i < \infty$, $n_i \geq N$ and

$$\sum_{i\in\mathcal{I}} c_i \chi_{A_i} \geq \chi_Z. \tag{3.7}$$

We show below that

$$\Lambda^{\mathbf{a},s+\delta}_{\Phi,N,6\epsilon}(Z) \leq \sum_{i\in\mathcal{I}} c_i \exp\left(-n_i s + \frac{1}{a_1} \sup_{x\in A_j} \log \phi_{\lceil a_1 n_j\rceil}(x)\right), \tag{3.8}$$

which implies $\Lambda^{\mathbf{a},s+\delta}_{\Phi,N,6\epsilon}(Z) \leq \mathcal{W}^{\mathbf{a},s}_{\Phi,N,\epsilon}(Z)$.

To prove (3.8), we write $\mathcal{I}_n = \{i \in \mathcal{I} :\ n_i = n\}$,

$$g_n(x) = (\phi_{\lceil a_1 n\rceil}(x))^{1/a_1}, \quad g_n(E) = \sup_{x\in E} g_n(x)$$

for $n \in \mathbb{N}$, $x \in X_1$ and $E \subseteq X_1$. Moreover set

$$Z_{n,t} = \Big\{x \in Z :\ \sum_{i\in\mathcal{I}_n} c_i \chi_{A_i}(x) > t\Big\}.$$

We claim that

$$\Lambda^{\mathbf{a},s+\delta}_{\Phi,N,\epsilon}(Z_{n,t}) \leq \frac{1}{tn^2} \sum_{i\in\mathcal{I}_n} c_i \exp(-ns) g_n(A_i), \quad \forall\, n \geq N,\ 0 < t < 1. \tag{3.9}$$

To prove the claim, assume that $n \geq N$ and $0 < t < 1$. Set $D = \frac{1}{n}\log g_n(Z_{n,t})$. For $\ell = 1, \ldots, n$ and $i \in \mathcal{I}_n$, write

$$Z^\ell_{n,t} = \left\{x \in Z_{n,t} : \frac{1}{n}\log g_n(x) \in \left(D - \frac{\gamma\ell}{n}, D - \frac{\gamma(\ell-1)}{n}\right]\right\}, \quad A_{i,\ell} := A_i \cap Z^\ell_{n,t},$$

and

$$Z^0_{n,t} = \left\{x \in Z_{n,t} : \frac{1}{n}\log g_n(x) \leq D - \gamma\right\}, \quad A_{i,0} = A_i \cap Z^0_{n,t}.$$

For $\ell = 0, 1, \ldots, n$, write $\mathcal{I}_{n,\ell} = \{i \in \mathcal{I}_n :\ A_{i,\ell} \neq \emptyset\}$; then

$$Z^\ell_{n,t} = \Big\{x \in X_1 :\ \sum_{i\in\mathcal{I}_{n,\ell}} c_i \chi_{A_{i,\ell}}(x) > t\Big\}.$$

Hence by Lemma 3.7, $Z^\ell_{n,t}$ can be covered by no more than $\frac{1}{t}\sum_{i\in\mathcal{I}_{n,\ell}} c_i$ balls with center in $\bigcup_{i\in\mathcal{I}_n} A_{i,\ell}$ and radius 6ϵ (in metric $d^{\mathbf{a}}_n$). It follows that for $\ell = 1, \ldots, n$,

$$\begin{aligned}\Lambda^{\mathbf{a},s+\delta}_{\Phi,N,6\epsilon}(Z^\ell_{n,t}) &\leq e^{-n(s+\delta)}\Big(\frac{1}{t}\sum_{i\in\mathcal{I}_{n,\ell}} c_i\Big) g_n(Z^\ell_{n,t}) \leq e^{-n(s+\delta)} e^{\gamma} \frac{1}{t}\sum_{i\in\mathcal{I}_{n,\ell}} c_i g_n(A_{i,\ell}) \\ &\leq e^{\gamma - n\delta}\frac{1}{t}\sum_{i\in\mathcal{I}_n} c_i e^{-ns} g_n(A_i).\end{aligned} \tag{3.10}$$

We still need to estimate $\Lambda^{\mathbf{a},s+\delta}_{\Phi,N,6\epsilon}(Z^0_{n,t})$. By Lemma 3.9, X_1 (and thus $Z^0_{n,t}$) can be covered by no more than $\exp(n\gamma)$ balls of radius 6ϵ (in metric $d^{\mathbf{a}}_n$). Hence

$$\begin{aligned}\Lambda^{\mathbf{a},s+\delta}_{\Phi,N,6\epsilon}(Z^0_{n,t}) &\leq \exp(n\gamma)e^{-n(s+\delta)}g_n(Z^0_{n,t}) \leq \exp(n\gamma)e^{-n(s+\delta)}\exp(n(D-\gamma)) \\ &\leq e^{-n(s+\delta)}\exp(nD) \leq e^{-n\delta}\frac{1}{t}\sum_{i\in\mathcal{I}_n} c_i e^{-ns} g_n(A_i),\end{aligned} \tag{3.11}$$

where the last inequality uses the following arguments: since $\exp(nD) = g_n(Z_{n,t})$, for any $u < \exp(nD)$, there exists $x \in Z_{n,t}$ so that $g_n(x) \geq u$; however since $x \in Z_{n,t}$ we have $\sum_{i\in\mathcal{I}_n:\, A_i\ni x} c_i \geq t$, which implies

$$\frac{1}{t}\sum_{i\in\mathcal{I}_n} c_i g_n(A_i) \geq \frac{1}{t}\sum_{i\in\mathcal{I}_n: A_i\ni x} c_i g_n(A_i) \geq \frac{1}{t}\sum_{i\in\mathcal{I}_n: A_i\ni x} c_i u \geq u.$$

Combining (3.10)-(3.11), we have

$$\begin{aligned}\Lambda^{\mathbf{a},s+\delta}_{\Phi,N,6\epsilon}(Z_{n,t}) \leq \sum_{\ell=0}^{n} \Lambda^{\mathbf{a},s+\delta}_{\Phi,N,6\epsilon}(Z^\ell_{n,t}) &\leq (n+1)e^{\gamma-n\delta}\frac{1}{t}\sum_{i\in\mathcal{I}_n} c_i e^{-ns} g_n(A_i) \\ &\leq \frac{1}{n^2 t}\sum_{i\in\mathcal{I}_n} c_i e^{-ns} g_n(A_i),\end{aligned} \tag{3.12}$$

where in the last inequality we use (3.6). This finishes the proof of (3.9).

To complete the proof of Proposition 3.5, notice that $\sum_{n=N}^{\infty} n^{-2} \leq \sum_{n=2}^{\infty} n^{-2} \leq 1$; hence if $x \notin \bigcup_{n\geq N} Z_{n,n^{-2}t}$, then

$$\sum_{i\in\mathcal{I}} c_i \chi_{A_i}(x) = \sum_{i\in\bigcup_{n=N}^{\infty}\mathcal{I}_n} c_i\chi_{A_i}(x) \leq \sum_{n=N}^{\infty}\sum_{i\in\mathcal{I}_n} c_i\chi_{A_i}(x) \leq \sum_{n=N}^{\infty} n^{-2}t \leq t < 1,$$

thus $x \notin Z$ by (3.7). Therefore $Z \subseteq \bigcup_{n\geq N} Z_{n,n^{-2}t}$. By (3.12),

$$\Lambda^{\mathbf{a},s+\delta}_{\Phi,N,6\epsilon}(Z) \leq \sum_{n=N}^{\infty}\Lambda^{\mathbf{a},s+\delta}_{\Phi,N,6\epsilon}(Z_{n,n^{-2}t}) \leq \frac{1}{t}\sum_{n=N}^{\infty}\sum_{i\in\mathcal{I}_n} c_i e^{-ns} g_n(A_i) \leq \frac{1}{t}\sum_{i\in\mathcal{I}} c_i e^{-n_i s} g_{n_i}(A_i).$$

Letting $t \uparrow 1$, we have

$$\Lambda^{\mathbf{a},s+\delta}_{\Phi,N,6\epsilon}(Z) \leq \sum_{i\in\mathcal{I}} c_i e^{-n_i s} g_{n_i}(A_i),$$

that is, (3.8) holds. This finishes the proof of Proposition 3.5. □

3.3. **Proof of Lemma 3.3.** It is easy to see that Lemma 3.3 follows directly from Proposition 3.5 and the following lemma.

Lemma 3.10. *Let $s \geq 0$, $N \in \mathbb{N}$ and $\epsilon > 0$. Suppose that $c := \mathcal{W}^{\mathbf{a},s}_{\Phi,N,\epsilon}(X_1) > 0$. Then there is a Borel probability measure μ on X_1 such that for any $n \geq N$, $x \in X_1$, and any compact $K \subset B^{\mathbf{a}}_n(x,\epsilon)$,*

$$\mu(K) \leq \frac{1}{c} e^{-ns} g_n(K),$$

where

$$g_n(z) = (\phi_{\lceil a_1 n \rceil}(z))^{1/a_1}, \quad g_n(K) = \sup_{z \in K} g_n(z).$$

Proof. Here we adopt the idea employed by Howroyd in his proof of the Frostman lemma in compact metric spaces (cf. [19, Theorem 2]). Clearly $c < \infty$. We define a function p on the space $C(X_1)$ of continuous real-valued functions on X_1 by

$$p(f) = (1/c)\mathcal{W}^{\mathbf{a},s}_{\Phi,N,\epsilon}(f).$$

Let $\mathbf{1} \in C(X_1)$ denote the constant function $\mathbf{1}(x) \equiv 1$. It is easy to verify that

(1) $p(f+g) \leq p(f) + p(g)$ for any $f, g \in C(X_1)$.
(2) $p(tf) = tp(f)$ for any $t \geq 0$ and $f \in C(X_1)$.
(3) $p(\mathbf{1}) = 1$, $0 \leq p(f) \leq \|f\|_\infty$ for any $f \in C(X_1)$, and $p(g) = 0$ for $g \in C(X_1)$ with $g \leq 0$.

By the Hahn-Banach theorem, we can extend the linear functional $t \mapsto tp(\mathbf{1})$, $t \in \mathbb{R}$, from the subspace of the constant functions to a linear functional $L : C(X_1) \to \mathbb{R}$ satisfying

$$L(\mathbf{1}) = p(\mathbf{1}) = 1 \text{ and } -p(-f) \leq L(f) \leq p(f) \text{ for any } f \in C(X_1).$$

If $f \in C(X_1)$ with $f \geq 0$, then $p(-f) = 0$ and so $L(f) \geq 0$. Hence combining the fact $L(\mathbf{1}) = 1$, we can use the Riesz representation theorem to find a Borel probability measure μ on X_1 such that $L(f) = \int f d\mu$ for $f \in C(X_1)$.

Now let $x \in X_1$ and $n \geq N$. Suppose that K is a compact subset of $B^{\mathbf{a}}_n(x,\epsilon)$. Let $\delta > 0$. Since g_n is upper semi-continuous, there exists an open set $B^{\mathbf{a}}_n(x,\epsilon) \supset V \supset K$ such that $g_n(V) \leq g_n(K) + \delta$.

By the Uryson lemma, there exists $f \in C(X_1)$ such that $0 \leq f \leq 1$, $f(y) = 1$ for $y \in K$, and $f(y) = 0$ for $y \in X_1 \backslash V$. Then $\mu(K) \leq L(f) \leq p(f)$. Since $f \leq \chi_V$ and $n \geq N$, we have $\mathcal{W}^{\mathbf{a},s}_{\Phi,N,\epsilon}(f) \leq e^{-ns} g_n(V)$ and thus $p(f) \leq \frac{1}{c} e^{-sn} g_n(V)$. Therefore

$$\mu(K) \leq \frac{1}{c} e^{-ns} g_n(V) \leq \frac{1}{c} e^{-ns}(g_n(K) + \delta).$$

Letting $\delta \to 0$, we have $\mu(K) \leq \frac{1}{c} e^{-ns} g_n(K)$. This completes the proof of the lemma. □

4. The proof of Theorem 1.4: Lower bound

In this section, we prove the lower bound part of Theorem 1.4. The following weighted version of Brin-Katok theorem plays a key role in our proof.

Theorem 4.1. *For each ergodic measure $\mu \in \mathcal{M}(X_1, T_1)$, we have*

$$\lim_{\epsilon \to 0} \liminf_{n \to +\infty} \frac{-\log \mu(B_n^{\mathbf{a}}(x, \epsilon))}{n} = \lim_{\epsilon \to 0} \limsup_{n \to +\infty} \frac{-\log \mu(B_n^{\mathbf{a}}(x, \epsilon))}{n} = h_\mu^{\mathbf{a}}(T_1)$$

for μ-a.e. $x \in X_1$.

We shall postpone the proof of Theorem 4.1 to Appendix A. In the following we prove the lower bound part of Theorem 1.4 for sub-additive potentials rather than additive potentials.

Proposition 4.2. *Let $\Phi = \{\log \phi_n\}_{n=1}^\infty$ be a sub-additive potential on X_1. Then*

$$P^{\mathbf{a}}(T_1, \Phi) \geq \sup \left\{ \Phi_*(\mu) + h_\mu^{\mathbf{a}}(T_1) : \ \mu \in \mathcal{M}(X_1, T_1), \Phi_*(\mu) \neq -\infty \right\}.$$

Proof. By Jacobs' theorem (cf. [37, Theorem 8.4]) and Proposition A.1.(3) in [16], if $\mu = \int_{\mathcal{E}(X_1,T_1)} m \, d\tau(m)$ is the ergodic decomposition of an element μ in $\mathcal{M}(X_1, T_1)$, then

$$h_\mu^{\mathbf{a}}(T_1) = \int_{\mathcal{E}(X_1,T_1)} h_m^{\mathbf{a}}(T_1) \, d\tau(m), \quad \Phi_*(\mu) = \int_{\mathcal{E}(X_1,T_1)} \Phi_*(m) \, d\tau(m).$$

Hence to prove the proposition, it suffices to show that

$$P^{\mathbf{a}}(T_1, \Phi) \geq \Phi_*(\mu) + \min\{\delta^{-1}, \ h_\mu^{\mathbf{a}}(T_1) - \delta\} - \delta \tag{4.1}$$

for any $\delta > 0$ and any ergodic $\mu \in \mathcal{M}(X_1, T_1)$ with $\Phi_*(\mu) \neq -\infty$.

For this purpose, we fix $\delta > 0$ and an ergodic measure μ on X_1 with $\Phi_*(\mu) \neq -\infty$. Write

$$H := \min\{\delta^{-1}, \ h_\mu^{\mathbf{a}}(T_1) - \delta\}.$$

By Theorem 4.1, we can choose $\epsilon > 0$ so that

$$\liminf_{n \to \infty} \frac{-\log \mu(B_n^{\mathbf{a}}(x, \epsilon))}{n} > H \quad \text{for } \mu\text{-a.e. } x \in X_1. \tag{4.2}$$

Since Φ is sub-additive, by Kingman's subadditive ergodic theorem (cf. [37, p. 231] and [16, Proposition A.1.]), we have

$$\lim_{n \to \infty} \frac{1}{n} \log \phi_n(x) = \Phi_*(\mu)$$

for μ-a.e. $x \in X_1$. Hence there exists a large $N \in \mathbb{N}$ and a Borel set $E_N \subset X_1$ with $\mu(E_N) > 1/2$ such that for any $x \in E_N$ and $n \geq N$,

$$\mu(B_n^{\mathbf{a}}(x,\epsilon)) < \exp(-nH), \quad \log \phi_{\lceil a_1 n \rceil}(x) \geq a_1 n \Phi_*(\mu) - a_1 n\delta. \tag{4.3}$$

Now assume that $\Gamma = \{(n_j, A_j)\}_i$ is a countable collection so that $n_j \geq N$, $A_j \in \mathcal{T}^{\mathbf{a}}_{n_j,\epsilon/2}$ (cf. (3.2) for the definition) and $\bigcup_j A_j = X_1$. By definition, for each j, there exists $x_j \in X$ so that $A_j \subseteq B^{\mathbf{a}}_{n_j}(x_j, \epsilon/2)$. Set

$$\mathcal{I} := \{j : \ A_j \cap E_N \neq \emptyset\}.$$

For $j \in \mathcal{I}$, pick $y_j \in A_j \cap E_N$; then we have

$$A_j \subseteq B^{\mathbf{a}}_{n_j}(x_j, \epsilon/2) \subseteq B^{\mathbf{a}}_{n_j}(y_j, \epsilon)$$

and thus

$$\mu(A_j) \leq \mu(B^{\mathbf{a}}_{n_j}(y_j,\epsilon)) \leq \exp(-n_j H);$$

moreover,

$$\frac{1}{a_1}\sup_{x\in A_j} \log \phi_{\lceil a_1 n_j\rceil}(x) \geq \frac{1}{a_1}\log \phi_{\lceil a_1 n_j\rceil}(y_j) \geq n_j\Phi_*(\mu) - n_j\delta.$$

Set $s := \Phi_*(\mu) + H - \delta$. Then for any $j \in \mathcal{I}$,

$$\exp\left(-sn_j + \frac{1}{a_1}\sup_{x\in A_j}\phi_{\lceil a_1 n_j\rceil}(x)\right) \geq \mu(A_j)\exp\left(n_j\left(-s + \Phi_*(\mu) + H - \delta\right)\right) = \mu(A_j).$$

Summing over $j \in \mathcal{I}$, we have

$$\sum_{j\in\mathcal{I}}\exp\left(-sn_j + \frac{1}{a_1}\sup_{x\in A_j}\phi_{\lceil a_1 n_j\rceil}(x)\right) \geq \sum_{j\in\mathcal{I}}\mu(A_j) \geq \mu\left(\bigcup_{j\in\mathcal{I}} A_j\right) \geq \mu(E_N) \geq 1/2.$$

It follows that $\Lambda^{\mathbf{a},s}_{\Phi,\epsilon}(X_1) \geq \Lambda^{\mathbf{a},s}_{\Phi,N,\epsilon}(X_1) \geq 1/2$, and thus

$$P^{\mathbf{a}}(T_1,\Phi) \geq P^{\mathbf{a}}(T_1,\Phi,X_1,\epsilon/2) \geq s = \Phi_*(\mu) + \min\{\delta^{-1},\ h^{\mathbf{a}}_\mu(T_1) - \delta\} - \delta,$$

as desired. □

5. The proof of Theorem 1.4: upper bound

In this section, we prove the upper bound in Theorem 1.4, that is, for any $f \in C(X_1)$ and $\delta > 0$, there exists $\mu \in \mathcal{M}(X_1, T_1)$ such that

$$P^{\mathbf{a}}(T_1, f) \leq h^{\mathbf{a}}_\mu(T_1) + \int_{X_1} f d\mu + \delta.$$

Before proving the above result, we first give some lemmas.

Lemma 5.1. *Let (X,T) be a TDS and $\mu \in \mathcal{M}(X)$. Let $\alpha = \{A_1, \dots, A_M\}$ be a Borel partition of X with cardinality M. Write for brevity*

$$h(n) := H_{\frac{1}{n}\sum_{i=0}^{n-1}\mu\circ T^{-i}}(\alpha), \quad h(n,m) := H_{\frac{1}{m}\sum_{i=n}^{m+n-1}\mu\circ T^{-i}}(\alpha).$$

for $n, m \in \mathbb{N}$. Then

(i) $h(n) \le \log M$ *and* $h(n,m) \le \log M$ *for* $n, m \in \mathbb{N}$.
(ii) $|h(n+1) - h(n)| \le \frac{1}{n+1}\log\left(3M^2(n+1)\right)$ *for all* $n \in \mathbb{N}$.
(iii) $\left|h(n+m) - \frac{n}{n+m}h(n) - \frac{m}{n+m}h(n,m)\right| \le \log 2$ *for all* $n, m \in \mathbb{N}$.

Proof. (i) is obvious. Now we turn to the proof of (ii). It is well known (see e.g. [37, Theorem 8.1] and the proof therein) that for any finite Borel partition β of X, $\nu_1, \nu_2 \in \mathcal{M}(X)$ and $p \in [0,1]$,

$$\begin{aligned} 0 &\le H_{p\nu_1+(1-p)\nu_2}(\beta) - pH_{\nu_1}(\beta) - (1-p)H_{\nu_2}(\beta) \\ &\le -(p\log p + (1-p)\log(1-p)) \\ &\le \log 2. \end{aligned} \tag{5.1}$$

Let $n \in \mathbb{N}$. Applying (5.1) and (i), we have

$$\begin{aligned} &|h(n+1) - h(n)| \\ &\quad = \left|h(n+1) - \frac{n}{n+1}h(n) - \frac{1}{n+1}H_{\mu\circ T^{-n}}(\alpha) - \frac{1}{n+1}h(n) + \frac{1}{n+1}H_{\mu\circ T^{-n}}(\alpha)\right| \\ &\quad \le \left|h(n+1) - \frac{n}{n+1}h(n) - \frac{1}{n+1}H_{\mu\circ T^{-n}}(\alpha)\right| + \frac{2}{n+1}\log M \\ &\quad \le -\frac{n}{n+1}\log\frac{n}{n+1} - \frac{1}{n+1}\log\frac{1}{n+1} + \frac{2}{n+1}\log M \\ &\quad \le \frac{1}{n+1}\log\left(3M^2(n+1)\right), \end{aligned}$$

where we use the fact $(1+1/n)^n < e < 3$ in the last inequality. This proves (ii).

Finally, since

$$\frac{1}{n+m}\sum_{i=0}^{n+m-1}\mu\circ T^{-i} = \frac{n}{n+m}\left(\frac{1}{n}\sum_{i=0}^{n-1}\mu\circ T^{-i}\right) + \frac{m}{n+m}\left(\frac{1}{m}\sum_{i=n}^{n+m-1}\mu\circ T^{-i}\right)$$

for $n, m \in \mathbb{N}$, (iii) follows from (5.1). □

Lemma 5.2. *Let (X,T) be a TDS and $\mu \in \mathcal{M}(X)$. For $\epsilon > 0$ and $\ell, M \in \mathbb{N}$, let $H_\bullet(\epsilon, M; \ell)$ be defined as in (2.3). Then the following statements hold.*

(1) *For all* $n \in \mathbb{N}$,

$$\left| H_{\frac{1}{n}\sum_{i=0}^{n-1}\mu\circ T^{-i}}(\epsilon, M; \ell) - H_{\frac{1}{n+1}\sum_{i=0}^{n}\mu\circ T^{-i}}(\epsilon, M; \ell) \right| \leq \frac{1}{\ell(n+1)} \log\left(3M^{2\ell}(n+1)\right).$$

(2) *For all* $n, m \in \mathbb{N}$,

$$\begin{aligned} &\frac{n}{n+m} H_{\frac{1}{n}\sum_{i=0}^{n-1}\mu\circ T^{-i}}(\epsilon, M; \ell) + \frac{m}{n+m} H_{\frac{1}{m}\sum_{i=n}^{n+m-1}\mu\circ T^{-i}}(\epsilon, M; \ell) \\ &\leq H_{\frac{1}{n+m}\sum_{i=0}^{m+n-1}\mu\circ T^{-i}}(\epsilon, M; \ell) + \frac{\log 2}{\ell}. \end{aligned} \tag{5.2}$$

Proof. The statements directly follow from the definition of $H_\bullet(\epsilon, M; \ell)$ and Lemma 5.1. $\square$

Lemma 5.3 (Lemma 2.4 of [10]). *Let* $\nu \in \mathcal{M}(X)$ *and* $M \in \mathbb{N}$. *Suppose* $\xi = \{A_1, \ldots, A_j\}$ *is a Borel partition of* X *with* $j \leq M$. *Then for any positive integers* n, ℓ *with* $n \geq 2\ell$, *we have*

$$\frac{1}{n} H_\nu\left(\bigvee_{i=0}^{n-1} T^{-i}\xi\right) \leq \frac{1}{\ell} H_{\nu_n}\left(\bigvee_{i=0}^{\ell-1} T^{-i}\xi\right) + \frac{2\ell}{n}\log M,$$

where $\nu_n = \frac{1}{n}\sum_{i=0}^{n-1} \nu \circ T^{-i}$.

The following lemma is a slight variant of [20, Lemma 4.1] by Kenyon and Peres.

Lemma 5.4. *Let* $p \in \mathbb{N}$. *Let* $u_j : \mathbb{N} \to \mathbb{R}$ *(*$j = 1, \ldots, p$*) be bounded functions with*

$$\lim_{n\to\infty} |u_j(n+1) - u_j(n)| = 0.$$

Then for any positive numbers $c_1, \ldots, c_p$ *and* $r_1, \ldots, r_p$,

$$\limsup_{n\to+\infty} \sum_{j=1}^{p} (u_j(\lceil c_j n\rceil) - u_j(\lceil r_j n\rceil)) \geq 0.$$

Proof. For the convenience of reader, we give a proof by adapting the argument of Kenyon and Peres in [20].

For $j = 1, \ldots, p$, extend u_j in a piecewise linear fashion to a bounded continuous function on $[1, +\infty)$. Then for each $1 \leq j \leq p$,

$$\lim_{t\to+\infty} \sup\left\{ |u_j(x) - u_j(y)| : x, y \geq t,\ |x-y| \leq \max_{1\leq i\leq p} \max\{c_i, r_i, 1\} \right\} = 0. \tag{5.3}$$

Take a positive number M so that

$$M > \max_{1\le j\le p}\{|\log c_j| + |\log r_j| + 1\}. \tag{5.4}$$

Then for every $w > M$,

$$\begin{aligned}
&\Big|\int_M^w \sum_{j=1}^p \big(u_j(e^{x+\log c_j}) - u_j(e^{x+\log r_j})\big)dx\Big| \\
&\quad= \left|\sum_{j=1}^p \left[\int_{M+\log c_j}^{w+\log c_j} u_j(e^x)dx - \int_{M+\log r_j}^{w+\log r_j} u_j(e^x)dx\right]\right| \\
&\quad\le \sum_{j=1}^p \left|\int_{M+\log c_j}^{w+\log c_j} u_j(e^x)dx - \int_{M+\log r_j}^{w+\log r_j} u_j(e^x)dx\right| \\
&\quad= \sum_{j=1}^p \Big|\int_{w+\log r_j}^{w+\log c_j} u_j(e^x)dx - \int_{M+\log r_j}^{M+\log c_j} u_j(e^x)dx\Big|,
\end{aligned}$$

Since each u_j is bounded, the sum in the right-hand side of the last '=' above is uniformly bounded. It follows that

$$\limsup_{x\to+\infty}\sum_{j=1}^p \big(u_j(e^{x+\log c_i}) - u_j(e^{x+\log r_j})\big) \ge 0.$$

Setting $t = e^x$, one has

$$\limsup_{t\to+\infty}\sum_{j=1}^p (u_j(c_i t) - u_j(r_j t)) \ge 0.$$

Combining the above inequality with (5.3), we have

$$\begin{aligned}
&\limsup_{n\to+\infty}\sum_{j=1}^p (u_j(\lceil c_j n\rceil) - u_j(\lceil r_j n\rceil) \\
&\qquad= \limsup_{n\to+\infty}\sum_{j=1}^p (u_j(c_j n) - u_j(r_j n)) \\
&\qquad= \limsup_{t\to+\infty}\sum_{j=1}^p (u_j(c_j\lceil t\rceil) - u_j(r_j\lceil t\rceil)) \\
&\qquad= \limsup_{t\to+\infty}\sum_{j=1}^p (u_j(c_j t) - u_j(r_j t)) \ge 0,
\end{aligned}$$

which completes the proof of the lemma. □

Proof of Theorem 1.4: upper bound. Suppose that $P^{\mathbf{a}}(T_1,f)>0$. Fix $0<s<s'<P^{\mathbf{a}}(T_1,f)$. Let $\Phi=\{\log\phi_n\}_{n=1}^{\infty}$ be the additive potential generated by f, that is, $\phi_n(x)=\exp(S_nf(x))$ where $S_nf(x):=\sum_{i=0}^{n-1}f(T_1^ix)$. Take $\epsilon_0>0$ such that

$$\sup\{|f(x)-f(y)|:\ x,y\in X_1,\ d_1(x,y)\le\epsilon_0\}<(s'-s)a_1/(1+a_1). \tag{5.5}$$

By Lemma 3.3, there exist $\nu\in\mathcal{M}(X_1)$, $\epsilon\in(0,\epsilon_0)$, and $N\in\mathbb{N}$ such that

$$\begin{aligned}\nu(B_n^{\mathbf{a}}(x,\epsilon))&\le\sup_{y\in B_n^{\mathbf{a}}(x,\epsilon)}\exp\left(-s'n+\frac{1}{a_1}S_{\lceil a_1n\rceil}f(y)\right)\\&\le\exp\left(-sn+\frac{1}{a_1}S_{\lceil a_1n\rceil}f(x)\right)\end{aligned} \tag{5.6}$$

for any $n\ge N$ and $x\in X_1$, where in the last inequality we use (5.5).

By continuity, there exists $\tau\in(0,\epsilon)$ such that for any $1\le i<j\le k$, if $x_i,y_i\in X_i$ satisfy $d_i(x_i,y_i)<\tau$, then

$$d_j(\pi_{j-1}\circ\cdots\circ\pi_i(x_i),\pi_{j-1}\circ\cdots\circ\pi_i(y_i))<\epsilon.$$

Take $M_0\in\mathbb{N}$ with $\mathcal{P}_{X_i}(\tau,M_0)\ne\emptyset$ for $i=1,\ldots,k$, where $\mathcal{P}_{X_i}(\tau,M_0)$ is defined as in (2.2). Now fix $M\in\mathbb{N}$ with $M\ge M_0$. Let $\alpha_i\in\mathcal{P}_{X_i}(\tau,M)$ for $i=1,\ldots,k$. Set $\beta_i=\tau_{i-1}^{-1}\alpha_i$ and write for brevity that

$$t_0(n)=0,\quad t_i(n)=\lceil(a_1+\ldots+a_i)n\rceil$$

for $n\in\mathbb{N}$ and $i=1,\ldots,k$. Then for any $n\in\mathbb{N}$ and $x\in X_1$, we have

$$\bigvee_{i=1}^{k}\bigvee_{j=t_{i-1}(n)}^{t_i(n)-1}T_1^{-j}\beta_i(x)\subseteq B_n^{\mathbf{a}}(x,\epsilon). \tag{5.7}$$

Now assume that $n\ge N$. By (5.6) and (5.7),

$$\nu\Big(\bigvee_{i=1}^{k}\bigvee_{j=t_{i-1}(n)}^{t_i(n)-1}T_1^{-j}\beta_i(x)\Big)\le\exp\left(-sn+\frac{1}{a_1}S_{\lceil a_1n\rceil}f(x)\right) \tag{5.8}$$

for any $x\in X_1$. It follows that

$$\begin{aligned}H_\nu\Big(\bigvee_{i=1}^{k}\bigvee_{j=t_{i-1}(n)}^{t_i(n)-1}T^{-j}\beta_i\Big)&=-\int\log\nu\Big(\bigvee_{i=1}^{k}\bigvee_{j=t_{i-1}(n)}^{t_i(n)-1}T_1^{-j}\beta_i(x)\Big)d\nu(x)\\&\ge sn-\int\frac{1}{a_1}S_{\lceil a_1n\rceil}f(x)d\nu(x).\end{aligned}$$

Hence

$$\sum_{i=1}^{k} H_\nu\Big(\bigvee_{j=t_{i-1}(n)}^{t_i(n)-1} T_1^{-j}\beta_i\Big) \geq sn - \int \frac{1}{a_1} S_{\lceil a_1 n\rceil} f(x)d\nu(x). \tag{5.9}$$

Now fix $\ell \in \mathbb{N}$. By Lemma 5.3, the left-hand side of (5.9) is bounded from above by

$$\sum_{i=1}^{k} \frac{t_i(n)-t_{i-1}(n)}{\ell} H_{w_{i,n}}\Big(\bigvee_{j=0}^{\ell-1} T_1^{-j}\beta_i\Big) + 2k\ell\log M,$$

where

$$w_{i,n} := \frac{\sum_{j=t_{i-1}(n)}^{t_i(n)-1} \nu\circ T_1^{-j}}{t_i(n)-t_{i-1}(n)}.$$

Hence by (5.9) and the definition of $H_\bullet(\tau, M;\ell)$ (cf. (2.3)), we have

$$\begin{aligned}\sum_{i=1}^{k}&(t_i(n)-t_{i-1}(n))H_{w_{i,n}\circ\tau_{i-1}^{-1}}(\tau, M;\ell)\\ &\geq sn - \frac{\lceil a_1 n\rceil}{a_1}\int f dw_{1,n} - 2k\ell\log M.\end{aligned} \tag{5.10}$$

Define $\nu_m = \frac{\sum_{j=0}^{m-1}\nu\circ T_1^{-j}}{m}$ for $m\in\mathbb{N}$. For $i=1,\dots,k$, we have

$$\nu_m\circ\tau_{i-1}^{-1} = \frac{\sum_{j=0}^{m-1}(\nu\circ\tau_{i-1}^{-1})\circ T_i^{-j}}{m},\ w_{i,n}\circ\tau_{i-1}^{-1} = \frac{\sum_{j=t_{i-1}(n)}^{t_i(n)-1}(\nu\circ\tau_{i-1}^{-1})\circ T_i^{-j}}{t_i(n)-t_{i-1}(n)}$$

and

$$\nu_{t_i(n)}\circ\tau_{i-1}^{-1} = \frac{t_{i-1}(n)}{t_i(n)}\nu_{t_{i-1}(n)}\circ\tau_{i-1}^{-1} + \frac{t_i(n)-t_{i-1}(n)}{t_i(n)} w_{i,n}\circ\tau_{i-1}^{-1}. \tag{5.11}$$

Applying Lemma 5.2(2) to the measure $\nu\circ\tau_{i-1}^{-1}$ (more precisely, in (5.2), we replace the terms T, μ, n, m by T_i, $\nu\circ\tau_{i-1}^{-1}$, $t_{i-1}(n)$, $t_i(n)-t_{i-1}(n)$, respectively), we have

$$\begin{aligned}\frac{t_{i-1}(n)}{t_i(n)}&H_{\nu_{t_{i-1}(n)}\circ\tau_{i-1}^{-1}}(\tau, M,\ell) + \frac{t_i(n)-t_{i-1}(n)}{t_i(n)}H_{w_{i,n}\circ\tau_{i-1}^{-1}}(\tau, M;\ell)\\ &\leq H_{\nu_{t_i(n)}\circ\tau_{i-1}^{-1}}(\tau, M;\ell) + \frac{\log 2}{\ell}.\end{aligned}$$

That is,

$$\begin{aligned}t_i(n)&H_{\nu_{t_i(n)}\circ\tau_{i-1}^{-1}}(\tau, M;\ell) - t_{i-1}(n)H_{\nu_{t_{i-1}(n)}\circ\tau_{i-1}^{-1}}(\tau, M,\ell)\\ &\geq (t_i(n)-t_{i-1}(n))H_{w_{i,n}\circ\tau_{i-1}^{-1}}(\tau, M;\ell) - \frac{t_i(n)\log 2}{\ell}.\end{aligned}$$

Combining the above inequality with (5.10), we have

$$\begin{aligned}\Theta_n :=& \sum_{i=1}^{k} \Big(t_i(n) H_{\nu_{t_i(n)}\circ\tau_{i-1}^{-1}}(\tau, M; \ell) - t_{i-1}(n) H_{\nu_{t_{i-1}(n)}\circ\tau_{i-1}^{-1}}(\tau, M, \ell) \Big) \\ \geq & sn - \frac{t_1(n)}{a_1} \int f d\nu_{t_1(n)} - 2k\ell \log M - \frac{k t_k(n) \log 2}{\ell}.\end{aligned} \tag{5.12}$$

Write $g_i(n) := H_{\nu_n \circ \tau_{i-1}^{-1}}(\tau, M; \ell)$. Then by Lemma 5.2(1),

$$|g_i(n) - g_i(n+1)| \leq \frac{1}{\ell(n+1)} \log \big(3M^{2\ell}(n+1)\big). \tag{5.13}$$

Set

$$\gamma(n) := \sum_{i=2}^{k} t_i(n)(g_i(t_i(n)) - g_i(t_1(n))) - \sum_{i=2}^{k} t_{i-1}(n)(g_i(t_{i-1}(n)) - g_i(t_1(n))).$$

Then we have

$$\Theta_n = \gamma(n) + \sum_{i=1}^{k} (t_i(n) - t_{i-1}(n)) g_i(t_1(n)),$$

where Θ_n is defined as in (5.12). Hence by (5.12), we have

$$\begin{aligned}&\sum_{i=1}^{k} \frac{t_i(n) - t_{i-1}(n)}{n} g_i(t_1(n)) + \frac{t_1(n)}{a_1 n} \int f d\nu_{t_1(n)} \\ &\qquad \geq -\frac{\gamma(n)}{n} + s - \frac{2k\ell \log M}{n} - \frac{k t_k(n) \log 2}{n\ell}.\end{aligned} \tag{5.14}$$

Define

$$w(n) = \sum_{i=2}^{k} (a_1 + \cdots + a_{i-1})(g_i(t_{i-1}(n)) - g_i(t_1(n))) - \sum_{i=2}^{k} (a_1 + \cdots + a_i)(g_i(t_i(n)) - g_i(t_1(n))).$$

Then we have $\limsup_{n\to\infty} w(n) \geq 0$ by applying Lemma 5.4, in which we take $p = 2k - 2$,

$$u_j(n) = \begin{cases} (a_1 + \cdots + a_j) g_{j+1}(n) & \text{if } 1 \leq j \leq k-1, \\ -(a_1 + \cdots + a_{j-k+2}) g_{j-k+2}(n) & \text{if } k \leq j \leq 2k-2, \end{cases}$$

and

$$c_j = \begin{cases} a_j & \text{if } 1 \leq j \leq k-1, \\ a_{j-k+2} & \text{if } k \leq j \leq 2k-2, \end{cases}$$

and $r_j = 1$ for all j; the condition $\lim_{n\to\infty} |u_j(n+1) - u_j(n)| = 0$ fulfils, thanks to (5.13).

Since g_i's are bounded functions, we have

$$\limsup_{n\to\infty} \frac{-\gamma(n)}{n} = \limsup_{n\to\infty} w(n) \geq 0.$$

Hence letting $n \to \infty$ in (5.14) and taking the upper limit, we obtain

$$\limsup_{n\to\infty} \left(\sum_{i=1}^{k} a_i g_i(t_1(n)) + \int f dv_{t_1(n)} \right) \geq s - \frac{k(a_1 + \cdots + a_k)\log 2}{\ell}. \tag{5.15}$$

Take a subsequence (n_j) of the natural numbers so that the left-hand side of (5.15) equals

$$\lim_{j\to\infty} \left(\sum_{i=1}^{k} a_i H_{\nu_{t_1(n_j)} \circ \tau_{i-1}^{-1}}(\tau, M; \ell) + \int f d\nu_{t_1(n_j)} \right)$$

and moreover, $\nu_{t_1(n_j)}$ converges to an element $\lambda \in \mathcal{M}(X_1, T_1)$ in the weak* topology. Since the map $H_\bullet(\tau, M; \ell)$ is upper semi-continuous on $\mathcal{M}(X_1)$ (see Lemma 2.3), we have

$$\sum_{i=1}^{k} a_i H_{\lambda\circ\tau_{i-1}^{-1}}(\tau, M; \ell) + \int f d\lambda \geq s - \frac{k(a_1 + \cdots + a_k)\log 2}{\ell}. \tag{5.16}$$

Define

$$\mathcal{E} := \left\{ (M, \ell, \delta) : M, \ell \in \mathbb{N}, \delta > 0 \text{ with } M \geq M_0, \ell \geq \frac{k(a_1 + \cdots + a_k)\log 2}{\delta} \right\}$$

and

$$\Omega_{M,\ell,\delta} := \left\{ \eta \in \mathcal{M}(X_1, T_1) : H_\eta^{\mathbf{a}}(\tau, M; \ell) + \int f d\eta \geq s - \delta \right\},$$

where $H_\eta^{\mathbf{a}}(\tau, M; \ell) := \sum_{i=1}^{k} a_i H_{\eta\circ\tau_{i-1}^{-1}}(\tau, M; \ell)$. Then by (5.16), $\Omega_{M,\ell,\delta}$ is a non-empty compact set whenever $(M, \ell, \delta) \in \mathcal{E}$. However

$$\Omega_{M_1,\ell_1,\delta_1} \cap \Omega_{M_2,\ell_2,\delta_2} \supseteq \Omega_{M_1+M_2,\ell_1\ell_2,\min\{\delta_1,\delta_2\}}$$

for any $(M_1, \ell_1, \delta_1), (M_2, \ell_2, \delta_2) \in \mathcal{E}$. It follows (by finite intersection property) that

$$\bigcap_{(M,\ell,\delta)\in\mathcal{E}} \Omega_{M,\ell,\delta} \neq \emptyset.$$

Take $\mu_s \in \bigcap_{(M,\ell,\delta)\in\mathcal{E}} \Omega_{M,\ell,\delta}$. Then

$$h_{\mu_s}^{\mathbf{a}}(T_1, \tau) + \int f d\mu_s \geq s,$$

where $h^{\mathbf{a}}_{\mu_s}(T_1,\tau) := \sum_{i=1}^k a_i h_{\mu_s\circ\tau_{i-1}^{-1}}(T_i,\tau)$. Since the map $\theta \in \mathcal{M}(X_1,T_1) \mapsto h^{\mathbf{a}}_\theta(T_1,\tau)$ is upper semi-continuous (see Lemma 2.3), we can find $\mu \in \mathcal{M}(X_1,T_1)$ such that

$$h^{\mathbf{a}}_\mu(T_1,\tau) + \int f d\mu \geq P^{\mathbf{a}}_W(T_1,f,\epsilon) - \omega_\epsilon(f)$$

by letting $s \nearrow P^{\mathbf{a}}_W(T_1,f,\epsilon)$. Since $h^{\mathbf{a}}_\mu(T_1) \geq h^{\mathbf{a}}_\mu(T_1,\tau)$, this completes the proof of the proposition. □

6. Sub-additive case

In this section, we extend Theorem 1.4 to sub-additive potentials, under the following two additional assumptions: (1) $h_{\text{top}}(T_1) < \infty$ and (2) the entropy maps $\theta \in \mathcal{M}(X_i,T_i) \mapsto h_\theta(T_i)$, $i = 1,2,\cdots,k$, are upper semi-continuous.

Definition 6.1. *Let $f : X_1 \to [-\infty,+\infty)$ be an upper semicontinuous function. Define $\Psi = \{\log\psi_n\}_{n=1}^\infty$ by $\psi_n(x) = \exp(\sum_{j=0}^{n-1} f(T_1^j x))$. In this case, Ψ is additive. We just define*

$$P^{\mathbf{a}}(T_1,f) := P^{\mathbf{a}}(T_1,\Psi).$$

Lemma 6.2. *Assume that $h_{\text{top}}(T_1) < \infty$ and the entropy maps $\theta \in \mathcal{M}(X_i,T_i) \mapsto h_\theta(T_i)$, $i = 1,2,\cdots,k$, are upper semi-continuous. Let $f : X_1 \to [-\infty,+\infty)$ be a upper semicontinuous function. Then there exists $\mu \in \mathcal{M}(X_1,T_1)$ such that*

$$h^{\mathbf{a}}_\mu(T_1) + \int_{X_1} f d\mu \geq P^{\mathbf{a}}(T_1,f).$$

Proof. For $g \in C(X_1)$ with $g \geq f$, we define

$$\mathcal{M}_g = \Big\{\nu \in \mathcal{M}(X_1,T_1) : h^{\mathbf{a}}_\nu(T_1) + \int_{X_1} g d\nu \geq P^{\mathbf{a}}(T_1,f)\Big\}.$$

Notice that, under the assumptions of the lemma, the entropy map $\nu \in \mathcal{M}(X_1,T_1) \mapsto h^{\mathbf{a}}_\nu(T_1)$ is a bounded upper semi-continuous function. Hence by Theorem 1.4, there exists $\mu_g \in \mathcal{M}(X_1,T_1)$ such that

$$h^{\mathbf{a}}_{\mu_g}(T_1) + \int_{X_1} g d\mu_g \geq P^{\mathbf{a}}(T_1,g) \geq P^{\mathbf{a}}(T_1,f).$$

Thus $\mu_g \in \mathcal{M}_g$. Since $\nu \in \mathcal{M}(X_1,T_1) \mapsto \int_{X_1} g d\nu$ is a bounded continuous non-negative valued function on $\mathcal{M}(X_1,T_1)$, the mapping $\nu \in \mathcal{M}(X_1,T_1) \mapsto h^{\mathbf{a}}_\nu(T_1) + \int_{X_1} g d\nu$ is a bounded upper semicontinuous non-negative valued function on $\mathcal{M}(X_1,T_1)$. Thus $\mathcal{M}_g$ is a non-empty closed subset of $\mathcal{M}(X_1,T_1)$.

Now put

$$\mathcal{M}_f := \bigcap_{g \in C(X_1), g \geq f} \mathcal{M}_g.$$

Note that $\mathcal{M}_{g_1} \cap \mathcal{M}_{g_2} \supseteq \mathcal{M}_{\min\{g_1, g_2\}}$ for any $g_1, g_2 \in C(X_1)$ with $g_1 \geq f$, $g_2 \geq f$, and each $\mathcal{M}_g$ is a non-empty closed subset of the compact metric space $\mathcal{M}(X_1, T_1)$. Hence $\mathcal{M}_f \neq \emptyset$, by the finite intersection property characterization of compactness. Take any $\mu \in \mathcal{M}_f$. Then

$$h_\mu^{\mathbf{a}}(T_1) + \int_{X_1} g d\mu \geq P^{\mathbf{a}}(T_1, f)$$

for any $g \in C(X_1)$ with $g \geq f$. Moreover, since $0 \leq h_\mu^{\mathbf{a}}(T_1) < \infty$, we have

$$h_\mu^{\mathbf{a}}(T_1) + \inf_{g \in C(X_1), g \geq f} \int_{X_1} g d\mu \geq P^{\mathbf{a}}(T_1, f).$$

Finally by Lemma 2.2, $\inf\limits_{g \in C(X_1), g \geq f} \int_{X_1} g d\mu = \int_{X_1} f d\mu$ and thus

$$h_\mu^{\mathbf{a}}(T_1) + \int_{X_1} f d\mu \geq P^{\mathbf{a}}(T_1, f).$$

This completes the proof of the lemma. □

Lemma 6.3. *Let $\Phi = \{\log \phi_n\}_{n=1}^\infty$ be a sub-additive potential on X_1. If for $\ell \in \mathbb{N}$ and $M \in \mathbb{N}$, let $f_{\ell,M}(x) = \max\{\frac{1}{\ell} \log \phi_\ell(x), -M\}$ for $x \in X_1$, then $f_{\ell,M} : X_1 \to \mathbb{R}$ is a bounded upper semi-continuous function and*

$$P^{\mathbf{a}}(T_1, f_{\ell,M}) \geq P^{\mathbf{a}}(T_1, \Phi).$$

Proof. Let $\ell \in \mathbb{N}$ and $M \in \mathbb{N}$. Let $f_{\ell,M} = \max\{\frac{1}{\ell} \log \phi_\ell, -M\}$. It is clear that $f_{\ell,M} : X_1 \to \mathbb{R}$ is a bounded upper semi-continuous function since $\frac{1}{\ell} \log \phi_\ell : X_1 \to [-\infty, +\infty)$ is upper semi-continuous.

Let $\phi_0(x) \equiv 1$ for $x \in X_1$ and

$$D := D(\ell) = \sup_{x \in X_1,\, i \in \{0,1,\cdots,\ell-1\}} \log \phi_i(x).$$

Then $0 \leq D < \infty$. For $x \in X_1$ and $n \geq 2\ell$, we have

$$\begin{aligned}
\log \phi_n(x) &\leq \log \phi_i(x) + \left(\sum_{j=0}^{[\frac{n-i}{\ell}]-1} \log \phi_\ell(T_1^{j\ell+i} x) \right) + \log \phi_{n-i-[\frac{n-i}{\ell}]\ell} \left(T_1^{i+[\frac{n-i}{\ell}]\ell} x \right) \\
&\leq 2D + \sum_{j=0}^{[\frac{n-i}{\ell}]-1} \log \phi_\ell(T_1^{j\ell+i} x)
\end{aligned}$$

for each $i \in \{0, 1, \ldots, \ell - 1\}$, using the sub-additivity of $\Phi = \{\log \phi_n\}_{n=1}^{\infty}$, where $[a]$ denotes the greatest integer $\leq a$. Summing i from 0 to $\ell - 1$, we obtain

$$\begin{aligned}\log \phi_n(x) &\leq 2D + \sum_{i=0}^{\ell-1} \sum_{j=0}^{[\frac{n-i}{\ell}]-1} \frac{1}{\ell} \log \phi_\ell(T^{j\ell+i}x) = 2D + \sum_{j=0}^{n-\ell} \frac{1}{\ell} \log \phi_\ell(T_1^j x) \\ &\leq 2D + \sum_{j=0}^{n-\ell} f_{\ell,M}(T_1^j x) \leq C + \sum_{j=0}^{n-1} f_{\ell,M}(T_1^j x)\end{aligned}$$

where $C = 2D + \ell M \in [0, +\infty)$.

Define $\Psi = \{\log \psi_n\}_{n=1}^{\infty}$ by $\psi_n(x) = \exp\left(\sum_{j=0}^{n-1} f_{\ell,M}(T_1^j x)\right)$. Then

$$\phi_n(x) \leq e^C \psi_n(x), \qquad \forall\, x \in X_1,\ n \geq 2\ell, \tag{6.1}$$

This implies that for any $\epsilon > 0$, $s \in \mathbb{R}$ and $N \geq 2a_1\ell$,

$$\mathcal{M}^{\mathbf{a},s}_{\Phi,N,\epsilon}(X_1) \leq e^{\frac{C}{a_1}} \cdot \mathcal{M}^{\mathbf{a},s}_{\Psi,N,\epsilon}(X_1).$$

Hence $\mathcal{M}^{\mathbf{a},s}_{\Phi,\epsilon}(X_1) \leq e^{\frac{C}{a_1}} \mathcal{M}^{\mathbf{a},s}_{\Psi,\epsilon}(X_1)$ for $\epsilon > 0$, $s \in \mathbb{R}$. It follows that

$$P^{\mathbf{a}}(T_1, \Phi, X_1, \epsilon) \leq P^{\mathbf{a}}(T_1, \Psi, X_1, \epsilon) = P^{\mathbf{a}}(T_1, f_{\ell,M}, X_1, \epsilon).$$

Letting $\epsilon \to 0$, we are done. □

Theorem 6.4. *Assume that $h_{\text{top}}(T_1) < \infty$ and the entropy maps $\theta \in \mathcal{M}(X_i, T_i) \mapsto h_\theta(T_i)$, $i = 1, 2, \cdots, k$, are upper semi-continuous. Let $\Phi = \{\log \phi_n\}_{n=1}^{\infty}$ be a sub-additive potential on X_1. Then*

$$P^{\mathbf{a}}(T_1, \Phi) = \sup\{h^{\mathbf{a}}_\mu(T_1) + \Phi_*(\mu) : \mu \in \mathcal{M}(X_1, T_1)\},$$

and moreover the supremum is attainable.

Proof. By Proposition 4.2, it is sufficient to show that there exists $\mu \in \mathcal{M}(X_1, T_1)$ such that $P^{\mathbf{a}}(T_1, \Phi) \leq h^{\mathbf{a}}_\mu(T_1) + \Phi_*(\mu)$.

For $n, M \in \mathbb{N}$, let $f_n(x) = \frac{1}{n} \log \phi_n(x)$ and $f_{n,M}(x) = \max\{\frac{1}{n} \log \phi_n(x), -M\}$ for $x \in X_1$. Then $f_{n,M}$ is a bounded upper semi-continuous function. Define

$$\mathcal{M}_{n,M} = \left\{\nu \in \mathcal{M}(X_1, T_1) : h^{\mathbf{a}}_\nu(T_1) + \int_{X_1} f_{n,M}\, d\nu \geq P^{\mathbf{a}}(T_1, \Phi)\right\}.$$

By Lemma 6.2, there exists $\mu_{n,M} \in \mathcal{M}(X_1, T_1)$ such that

$$h^{\mathbf{a}}_{\mu_{n,M}}(T_1) + \int_{X_1} f_{n,M}\, d\mu_{n,M} \geq P^{\mathbf{a}}(T_1, f_{n,M}) \geq P^{\mathbf{a}}(T_1, \Phi),$$

where the last inequality comes from Lemma 6.3. Thus $\mu_{n,M} \in \mathcal{M}_{n,M}$. By the assumption, we know that the function $h^{\mathbf{a}}_{\bullet}(T_1)$ is bounded, upper semi-continuous and non-negative on $\mathcal{M}(X_1,T_1)$. Notice that $\nu\in\mathcal{M}(X_1,T_1)\mapsto\int_{X_1} f_{n,M}d\nu$ is also an upper semi-continuous function from $\mathcal{M}(X_1,T_1)$ to $\mathbb{R}$. Hence $\nu\in\mathcal{M}(X_1,T_1)\mapsto h_\nu^{\mathbf{a}}(T_1)+\int_{X_1} f_{n,M}d\nu$ is upper semi-continuous. Thus $\mathcal{M}_{n,M}$ is a non-empty closed subset of $\mathcal{M}(X_1,T_1)$. Moreover since $\mathcal{M}_{n,1}\supseteq\mathcal{M}_{n,2}\supseteq\cdots$ and $\inf_{M\in\mathbb{N}}\int_{X_1} f_{n,M}d\nu=\int_{X_1} f_n d\nu$ for any $\nu\in\mathcal{M}(X_1,T_1)$, one has $\mathcal{M}_n=\bigcap_{M\in\mathbb{N}}\mathcal{M}_{n,M}$ is a non-empty closed subset of $\mathcal{M}(X_1,T_1)$.

Now put

$$\mathcal{M}_\Phi:=\bigcap_{n\in\mathbb{N}}\mathcal{M}_n.$$

Since $\int_{X_1} f_{n_1n_2}d\nu\le\min\{\int_{X_1} f_{n_1}d\nu,\int_{X_1} f_{n_2}d\nu\}$ for $\nu\in\mathcal{M}(X_1,T_1)$, we have $\mathcal{M}_{n_1}\cap\mathcal{M}_{n_2}\supseteq\mathcal{M}_{n_1n_2}$ for any $n_1,n_2\in\mathbb{N}$. Moreover since each $\mathcal{M}_n$ is a non-empty closed subset of the compact metric space $\mathcal{M}(X_1,T_1)$, one has $\mathcal{M}_\Phi\neq\emptyset$ by the finite intersection property characterization of compactness. Take any $\mu\in\mathcal{M}_\Phi$. Then

$$h_\mu^{\mathbf{a}}(T_1)+\int_{X_1} f_n d\mu\ge P^{\mathbf{a}}(T_1,\Phi)$$

for any $n\in\mathbb{N}$. Moreover, since $0\le h_\mu^{\mathbf{a}}(T_1)<\infty$, we have

$$h_\mu^{\mathbf{a}}(T_1)+\inf_{n\in\mathbb{N}}\frac{1}{n}\int_{X_1}\log\phi_n d\mu\ge P^{\mathbf{a}}(T_1,\Phi).$$

Finally since $\inf\limits_{n\in\mathbb{N}}\frac{1}{n}\int_{X_1}\log\phi_n d\mu=\Phi_*(\mu)$ and thus

$$h_\mu^{\mathbf{a}}(T_1)+\Phi_*(\mu)\ge P^{\mathbf{a}}(T_1,\Phi).$$

This finishes the proof of the Theorem. □

7. Final remarks and examples

In this section we give some remarks, examples and questions.

7.1. In [2, 15], Barral and the first author defined weighted topological pressure for factor maps between subshifts in a different way, motivated from the study of multifractal analysis on affine Sierpinski gaskets [3, 4, 21, 29] and a question of Gatzouras and Peres [18] on the uniqueness of invariant measures of full dimension on certain affine invariant sets. The approach is based on the following lemma, which is derived from the relativized variational principle of Ledrappier and Walters [24] and its sub-additive extension [39].

Lemma 7.1. [2, 15] *Assume that (X,T) and (Y,S) are subshifts over finite alphabets and $\pi : X \to Y$ is a factor map. Let $f \in C(X)$ (or more general, a subadditive potential on X). Then there exists a sub-additive potential $\Phi_f = (\log \phi_n)_{n=1}^{\infty}$ on Y such that for any $\nu \in \mathcal{M}(Y,S)$,*

$$\sup_{\mu\in\mathcal{M}(X,T),\, \mu\circ\pi^{-1}=\nu} \left(\int f d\mu + h_\mu(T) - h_\nu(S) \right) = \Phi_*(\nu) := \lim_{n\to+\infty} \frac{1}{n} \int \log \phi_n d\nu.$$

According to above lemma, for given $a_1, a_2 > 0$, one has

$$\begin{aligned}
&\sup_{\mu\in\mathcal{M}(X,T),\, \mu\circ\pi^{-1}=\nu} \left(\int f d\mu + a_1 h_\mu(T) + a_2 h_\nu(S) \right) \\
&\quad = \sup_{\nu\in\mathcal{M}(Y,S)} \left\{ (a_1+a_2) h_\nu(S) + \sup_{\mu\in\pi^{-1}\nu} a_1 \left(\int \frac{1}{a_1} f d\mu + h_\mu(T) - h_\nu(S) \right) \right\} \\
&\quad = \sup_{\nu\in\mathcal{M}(Y,S)} \{(a_1+a_2) h_\nu(S) + (\Phi_{a_1^{-1} f})_*(\nu)\} \\
&\quad = (a_1+a_2) P\left(S, \frac{a_1}{a_1+a_2} \Phi_{a_1^{-1} f} \right).
\end{aligned}$$

where the last equality follows from the sub-additive thermodynamic formalism (see e.g. [10]). Hence in [2, 15], $P^{(a_1,a_2)}(T,f)$ was defined in terms of sub-additive topological pressure in the subshift case.

However, Lemma 7.1 does not extend to factor maps between general topological dynamical systems. Below we will give a counter example. Hence the approach in [2, 15] in defining weighted topological pressure does not extend to general topological dynamical systems.

Example 7.2. *Let $X = \{(x,y,z) \in \mathbb{R}^3 : -1 \le x \le 1, y^2 + z^2 = x^2\}$ be a cone surface. Define $T : X \to X$ by*

$$T((x, x\cos\theta, x\sin\theta)) = (x, x\cos(2\theta), x\sin(2\theta)), \quad x \in [-1,1].$$

Let $Y = [-1,1]$ and $S : Y \to Y$ be the identity. Set $\pi : X \to Y$ by $\pi((x,y,z)) = x$. Then (Y,S) is a factor of (X,T) associated with the factor map π. Take $f \in C(X)$ with $f \equiv 0$. Suppose that Lemma 7.1 extends to this case, that is, there exists a sub-additive potential Φ on Y such that for any $\nu \in \mathcal{M}(Y,S)$,

$$\sup_{\mu\in\pi^{-1}\nu} (h_\mu(T) - h_\nu(S)) = \Phi_*(\nu). \tag{7.1}$$

In what follows we derive a contradiction.

We first claim that the mapping

$$\nu \in \mathcal{M}(Y,S) \mapsto \sup_{\mu \in \pi^{-1}\nu} (h_\mu(T) - h_\nu(S)) \tag{7.2}$$

is not upper semi-continuous. To see this, for $t \in Y$, let $\nu_t = \delta_t$ (the Dirac measure at t). Clearly $\delta_t \in \mathcal{M}(Y,S)$ and when $t \to 0$, $\delta_t \to \delta_0$ in the weak-star topology. However one can check that

$$\sup_{\mu \in \pi^{-1}\delta_t} \big(h_\mu(T) - h_{\nu_t}(S)\big) = \begin{cases} \log 2, & \text{for } t \neq 0 \\ 0, & \text{if } t = 0 \end{cases}.$$

Hence the mapping in (7.2) *is not upper semi-continuous. Therefore by* (7.1)*, $\nu \mapsto \Phi_*(\nu)$ is not upper semi-continuous on $\mathcal{M}(Y,S)$. But this contradicts the fact that $\nu \mapsto \Phi_*(\nu)$ is always upper semi-continuous (see e.g.* [16, Proposition A.1.(2)]*).*

7.2. Using Corollary 1.5, we can extend Kenyon-Peres' variational principle (1.2) and its higher dimensional version to a particular class of skew product expanding maps on the k-torus $\mathbb{T}^k := \mathbb{R}^k/\mathbb{Z}^k$ ($k \geq 2$).

To see this, let $2 \leq m_1 \leq m_2 \leq \ldots \leq m_k$ be integers. For $i = 1, \ldots, k-1$, let ϕ_i be C^1 real-valued functions on $\mathbb{T}^i$. Define $T_1 : \mathbb{T}^k \to \mathbb{T}^k$ by

$$T_1((x_1, \ldots, x_k)) = (m_1 x_1, m_2 x_2 + \phi_1(x_1), \ldots, m_k x_k + \phi_{k-1}(x_1, \ldots, x_{k-1})).$$

This transformation can be viewed as a skew product of the maps

$$x_i \mapsto m_i x_i, \quad (i = 1, \ldots, k).$$

Let $K \subset \mathbb{T}^k$ be a T_1-invariant compact set. Let τ_i ($i = 1, \ldots, k-1$) be the canonical projection from $\mathbb{T}^k$ to $\mathbb{T}^{k-i}$, i.e.

$$\tau_i(x_1, \ldots, x_k) = (x_1, \ldots, x_{k-i}).$$

Set $X_1 = K$ and $X_i = \tau_{i-1}(K)$ for $2 \leq i \leq k$. Define $T_i : X_i \to X_i$ ($i = 2, \ldots, k$) by

$$T_i((x_1, \ldots, x_i)) = (m_1 x_1, m_2 x_2 + \phi_1(x_1), \ldots, m_i x_i + \phi_{i-1}(x_1, \ldots, x_{i-1})).$$

Then (X_{i+1}, T_{i+1}) is the factor of (X_i, T_i) associated with the factor map $\pi_i : X_i \to X_{i+1}$, which is defined by

$$(x_1, \ldots, x_{k+1-i}) \mapsto (x_1, \ldots, x_{k-i}).$$

Define $\mathbf{a} = (a_1, \ldots, a_k)$ with

$$a_1 = \frac{1}{\log m_k}, \quad a_i = \frac{1}{\log m_{k+1-i}} - \frac{1}{\log m_{k+2-i}} \quad \text{for } i = 2, \ldots, k.$$

It is direct to check that there exist two constants $C_1, C_2 > 0$ (depending on ϕ_i's) such that for any $\epsilon > 0$ and $x \in \mathbb{T}^k$,

$$C_2 B_{e^{-n}\epsilon}(x) \subset B_n^{\mathbf{a}}(x, \epsilon) \subset C_1 B_{e^{-n}\epsilon}(x). \tag{7.3}$$

Hence from the definition of $h_{\text{top}}^{\mathbf{a}}(\cdot)$, we see that $h_{\text{top}}^{\mathbf{a}}(T_1, K) = \dim_H K$. Applying Corollary 1.5, we have

$$\dim_H K = h_{\text{top}}^{\mathbf{a}}(T_1, K) = \sup_{\mu \in \mathcal{M}(X_1, T_1)} h_\mu^{\mathbf{a}}(T_1), \tag{7.4}$$

where the supremum is attainable at some ergodic $\mu \in \mathcal{M}(X_1, T_1)$. Moreover by (7.3) and Theorem A.1, we have $\dim_H \mu = h_\mu^{\mathbf{a}}(T_1)$ for each ergodic $\mu \in \mathcal{M}(X_1, T_1)$. Hence there exists an ergodic $\mu \in \mathcal{M}(X_1, T_1)$ of full Hausdorff dimension, i.e.

$$\dim_H \mu = \dim_H K. \tag{7.5}$$

This extends the work of Kenyon and Peres [20]. We remark that (7.5) was also proved by Luzia [25] for a more general class of skew product expanding maps on $\mathbb{T}^2$.

7.3. In [17], the authors proved a variational principle for topological entropies for arbitrary Borel subsets. We remark that this principle also holds for weighted topological entropies, by applying Lemma 3.10 and following the arguments in [17].

In the end we pose several questions about possible extensions of Theorem 1.4: does this result remain valid for $\mathbb{Z}^d$-actions? and moreover does it admit a relativized or randomized version? is there an analogous topological extension of the dimensional result on Gatzouras-Lalley self-affine carpets [23]?

Appendix A. A weighted version of the Brin-Katok theorem

The main result in this appendix is the following weighted version of the Brin-Katok theorem. It is needed in our proof of the lower bound of Theorem.

Theorem A.1. *For each ergodic measure $\mu \in \mathcal{M}(X_1, T_1)$, we have*

$$\lim_{\epsilon \to 0} \liminf_{n \to +\infty} \frac{-\log \mu(B_n^{\mathbf{a}}(x, \epsilon))}{n} = \lim_{\epsilon \to 0} \limsup_{n \to +\infty} \frac{-\log \mu(B_n^{\mathbf{a}}(x, \epsilon))}{n} = h_\mu^{\mathbf{a}}(T_1)$$

for μ-a.e. $x \in X_1$.

When $\mathbf{a} = (1, 0, \ldots, 0)$, the above result reduces to the Brin-Katok theorem on local entropy [7].

The proof of Theorem A.1 is based on the following weighted version of the Shannon-McMillan-Breiman theorem.

Proposition A.2. *Let $(X, \mathcal{B}, \mu, T)$ be a measure preserving dynamical system and $k \geq 1$. Let $\alpha_1, \ldots, \alpha_k$ be k countable measurable partitions of $(X, \mathcal{B}, \mu)$ with $H_\mu(\alpha_i) < \infty$ for each i, and $\mathbf{a} = (a_1, \ldots, a_k) \in \mathbb{R}^k$ with $a_1 > 0$ and $a_i \geq 0$ for $i \geq 2$. Then*

$$\lim_{N\to+\infty} \frac{1}{N} I_\mu\Big(\bigvee_{i=1}^{k} (\alpha_i)_0^{\lceil (a_1+\cdots+a_i)N\rceil - 1}\Big)(x) = \sum_{i=1}^{k} a_i \mathbb{E}_\mu(F_i|\mathcal{I}_\mu)(x) \tag{A.1}$$

almost everywhere, where

$$F_i(x) := I_\mu\Big(\bigvee_{j=i}^{k} \alpha_j \big| \bigvee_{n=1}^{\infty} T^{-n}(\bigvee_{j=i}^{k} \alpha_j)\Big)(x), \quad i = 1, \ldots, k$$

and $\mathcal{I}_\mu = \{B \in \mathcal{B} : \ \mu(B \triangle T^{-1}B) = 0\}$. In particular, if T is ergodic, we have

$$\lim_{N\to+\infty} \frac{1}{N} I_\mu\Big(\bigvee_{i=1}^{k} (\alpha_i)_0^{\lceil (a_1+\cdots+a_i)N\rceil - 1}\Big)(x) = \sum_{i=1}^{k} a_i h_\mu(T, \bigvee_{j=i}^{k} \alpha_j)$$

almost everywhere.

When $k = 1$ and $a_1 = 1$, Proposition A.2 reduces to the classical Shannon-McMillan-Breiman theorem (see e.g. [30, Theorem 7]). We remark that a variant of Proposition A.2, for certain particular partitions, was proved by Kenyon and Peres (cf. [20, Lemmas 3.1 and 4.4]) in the case that μ is ergodic. For completeness and for the convenience of the reader, we will provide a full proof of Proposition A.2 in the end of this section, by adapting the argument by Kenyon and Peres in [20].

The following result is a direct corollary of Proposition A.2.

Corollary A.3. *Let $(X, \mathcal{B}, \mu, T)$ be an ergodic measure preserving dynamical system and $k \geq 1$. If $\alpha_1, \ldots, \alpha_k$ are k countable measurable partitions of $(X, \mathcal{B}, \mu)$ with $\alpha_1 \succeq \alpha_2 \succeq \cdots \succeq \alpha_k$ and $H_\mu(\alpha_i) < \infty$, $i = 1, \ldots, k$, and $\mathbf{a} = (a_1, \ldots, a_k) \in \mathbb{R}^k$ with $a_1 > 0$ and $a_i \geq 0$ for $i \geq 2$, then*

$$\lim_{N\to+\infty} \frac{1}{N} I_\mu\Big(\bigvee_{i=1}^{k} \Big(\bigvee_{j=\lceil (a_0+\cdots+a_{i-1})N\rceil}^{\lceil (a_1+\cdots+a_i)N\rceil - 1} T^{-j}\alpha_i\Big)\Big)(x) = \sum_{i=1}^{k} a_i h_\mu(T, \alpha_i)$$

almost everywhere, where we make the convention $a_0 = 0$.

Proof of Theorem 4.1. We just adapt the proof of Brin and Katok [7] for their local entropy formula.

We first prove the upper bound. Let $\epsilon > 0$. Let α_i be a finite Borel partition of X_i, $i = 1, \ldots, k$, with $\mathrm{diam}(\alpha_i) < \epsilon$. Then

$$B_n^{\mathbf{a}}(x,\epsilon) \supseteq \bigcap_{i=1}^{k} (\tau_{i-1}^{-1}\alpha_i)_0^{\lceil (a_1+\cdots+a_i)n\rceil -1}(x)$$

for $x \in X_1$. Hence by Proposition A.2, for μ-a.e $x \in X_1$ we have

$$\begin{aligned}
\limsup_{n\to+\infty} \frac{-\log \mu(B_n^{\mathbf{a}}(x,\epsilon))}{n} &\le \limsup_{n\to+\infty} \frac{-\log \mu\Big(\bigcap_{i=1}^{k} (\tau_{i-1}^{-1}\alpha_i)_0^{\lceil (a_1+\cdots+a_i)n\rceil -1}(x)\Big)}{n} \\
&= \limsup_{n\to+\infty} \frac{I_\mu\Big(\bigvee_{i=1}^{k} (\tau_{i-1}^{-1}\alpha_i)_0^{\lceil (a_1+\cdots+a_i)n\rceil -1}\Big)(x)}{n} = \sum_{i=1}^{k} a_i h_\mu\Big(T_1, \bigvee_{j=i}^{k} \tau_{j-1}^{-1}\alpha_j\Big) \\
&= \sum_{i=1}^{k} a_i h_\mu\Big(T_1, \tau_{i-1}^{-1}\Big(\alpha_i \vee \bigvee_{j=i+1}^{k} \pi_i^{-1}\circ\cdots\circ\pi_{j-1}^{-1}\alpha_j\Big)\Big) \\
&= \sum_{i=1}^{k} a_i h_{\mu\circ\tau_{i-1}^{-1}}\Big(T_i, \alpha_i \vee \bigvee_{j=i+1}^{k} \pi_i^{-1}\circ\cdots\circ\pi_{j-1}^{-1}\alpha_j\Big) \\
&\le \sum_{i=1}^{k} a_i h_{\mu\circ\tau_{i-1}^{-1}}(T_i) = h_\mu^{\mathbf{a}}(T_1).
\end{aligned}$$

Letting $\epsilon \to 0$ in the above inequality, we have

$$\lim_{\epsilon\to 0}\limsup_{n\to+\infty} \frac{-\log \mu(B_n^{\mathbf{a}}(x,\epsilon))}{n} \le h_\mu^{\mathbf{a}}(T_1).$$

This completes the proof of the upper bound.

Next we prove the lower bound. It is sufficient to show that for any $\delta > 0$, there exist $\epsilon > 0$ and a measurable subset D of X_1 such that $\mu(D) > 1 - 3\delta$ and

$$\liminf_{n\to+\infty} \frac{-\log \mu(B_n^{\mathbf{a}}(x,\epsilon))}{n} \ge \min\left\{\frac{1}{\delta}, h_\mu^{\mathbf{a}}(T_1) - \delta\right\} - 2(1 + a_1 + \cdots + a_k)\delta$$

for any $x \in D$.

Fix $\delta > 0$. We are going to find such ϵ and D. First, we find a finite Borel partition $\alpha_i = \{A_1^i, A_2^i, \ldots, A_{u_i}^i\}$ of X_i, $i = 1, \ldots, k$, such that

(1) $\alpha_i \succeq \pi_i^{-1}(\alpha_{i+1})$ for $i = 1, \ldots, k-1$.
(2) $\sum_{i=1}^{k} a_i h_{\mu\circ\tau_{i-1}^{-1}}(T_i, \alpha_i) \ge \min\{\frac{1}{\delta}, h_\mu^{\mathbf{a}}(T_1) - \delta\}$.
(3) $\mu \circ \tau_{i-1}^{-1}(\partial\alpha_i) = 0$ for $i = 1, \ldots, k$.

Let $M = \max\{u_i : \ 1 \le i \le k\}$ and $\Lambda = \{1, \ldots, M\}$. Given $m \in \mathbb{N}$, for $\mathbf{s} = (s_i)_{i=0}^{m-1}, \mathbf{t} = (t_i)_{i=0}^{k-1} \in \Lambda^{\{0,1,\cdots,m-1\}}$, the *Hamming distance* between $\mathbf{s}$ and $\mathbf{t}$ is defined to be the following value

$$\frac{1}{m} \# \{i \in \{0, 1, \cdots, m-1\} : s_i \neq t_i\}.$$

For $\mathbf{s} \in \Lambda^{\{0,1,\cdots,m-1\}}$ and $0 < \tau \le 1$, let $Q(\mathbf{s}, \tau)$ be the total number of those $\mathbf{t} \in \Lambda^{\{0,1,\cdots,m-1\}}$ so that the Hamming distance between $\mathbf{s}$ and $\mathbf{t}$ does not exceed τ. Clearly,

$$Q_m(\tau) := \max_{\mathbf{s} \in \Lambda^{\{0,1,\cdots,m-1\}}} Q(\mathbf{s}, \tau) \le \binom{m}{\lceil m\tau \rceil} M^{\lceil m\tau \rceil}.$$

By the Stirling formula, there exists a small $\delta_1 > 0$ and a positive constant $C := C(\delta, M) > 0$ such that

$$\binom{m}{\lceil m\delta_1 \rceil} M^{\lceil m\delta_1 \rceil} \le e^{\delta m + C} \tag{A.2}$$

for all $m \in \mathbb{N}$.

For $\eta > 0$, set

$$U_\eta^i(\alpha_i) = \{x \in X_1 : \ B(\tau_{i-1}x, \eta) \not\subseteq \alpha_i(\tau_{i-1}x)\}, \quad i = 1, \ldots, k.$$

Then $\bigcap_{\eta>0} U_\eta^i(\alpha_i) = \tau_{i-1}^{-1}(\partial \alpha_i)$, and hence $\mu(U_\eta^i(\alpha_i)) \to \mu(\tau_{i-1}^{-1}(\partial \alpha_i)) = 0$ as $\eta \to 0$. Therefore, we can choose $\epsilon > 0$ such that $\mu(U_\eta^i(\alpha_i)) < \delta_1$ for any $0 < \eta \le \epsilon$ and $i = 1, \ldots, k$.

By the Birkhoff ergodic theorem, for μ-a.e. $x \in X_1$, we have

$$\begin{aligned}
&\lim_{n\to+\infty} \frac{1}{\lceil (a_1 + \cdots + a_k)n \rceil} \sum_{i=1}^{k} \sum_{j=\lceil (a_0+\cdots+a_{i-1})n \rceil}^{\lceil (a_1+\cdots+a_i)n \rceil - 1} \chi_{U_\epsilon^i(\alpha_i)}(T_1^j x) \\
&= \frac{1}{(a_1 + \cdots + a_k)} \sum_{i=1}^{k} a_i \mu(U_\epsilon^i(\alpha_i)) < \delta_1,
\end{aligned}$$

where we take the convention $a_0 = 0$. Thus we can find a large natural number ℓ_0 such that $\mu(A_\ell) > 1 - \delta$ for any $\ell \ge \ell_0$, where

$$A_\ell = \left\{ x \in X_1 : \ \frac{1}{\lceil (a_1 + \cdots + a_k)n \rceil} \sum_{i=1}^{k} \sum_{j=\lceil (a_0+\cdots+a_{i-1})n \rceil}^{\lceil (a_1+\cdots+a_i)n \rceil - 1} \chi_{U_\epsilon^i(\alpha_i)}(T_1^j x) \le \delta_1 \text{ for all } n \ge \ell \right\}.$$

Since $\tau_0^{-1}\alpha_1 \succeq \tau_1^{-1}\alpha_2 \succeq \cdots \succeq \tau_{k-1}^{-1}\alpha_k$, we have

$$\lim_{n\to+\infty} \frac{-\log\mu\Big(\bigvee_{i=1}^{k}\Big(\bigvee_{j=\lceil (a_0+\cdots+a_{i-1})n\rceil}^{\lceil (a_1+\cdots+a_i)n\rceil-1} T_1^{-j}\tau_{i-1}^{-1}\alpha_i\Big)(x)\Big)}{n}$$
$$= \sum_{i=1}^{k} a_i h_\mu(T_1, \tau_{i-1}^{-1}\alpha_i) = \sum_{i=1}^{k} a_i h_{\mu\circ\tau_{i-1}^{-1}}(T_i, \alpha_i)$$

almost everywhere by Corollary A.3. Hence we can find a large natural number ℓ_1 such that $\mu(B_\ell) > 1-\delta$ for any $\ell \geq \ell_1$, where B_ℓ is the set of all points $x \in X_1$ such that

$$\text{(A.3)} \quad \frac{-\log\mu\Big(\bigvee_{i=1}^{k}\Big(\bigvee_{j=\lceil (a_0+\cdots+a_{i-1})n\rceil}^{\lceil (a_1+\cdots+a_i)n\rceil-1} T_1^{-j}\tau_{i-1}^{-1}\alpha_i\Big)(x)\Big)}{n} \geq \sum_{i=1}^{k} a_i h_{\mu\circ\tau_{i-1}^{-1}}(T_i, \alpha_i) - \delta$$

for all $n \geq \ell$.

Fix $\ell \geq \max\{\ell_0, \ell_1\}$. Let $E = A_\ell \cap B_\ell$. Then $\mu(E) > 1-2\delta$. For $x \in X_1$ and $n \in \mathbb{N}$, the unique element

$$C(n,x) = (C_j(n,x))_{j=0}^{\lceil (a_1+\cdots+a_k)n\rceil-1}$$

in $\Lambda^{\{0,1,\cdots,\lceil (a_1+\cdots+a_k)n\rceil-1\}}$ satisfying that $T_1^j x \in \tau_{i-1}^{-1}(A^i_{C_j(n,x)})$ for $\lceil (a_0+\cdots+a_{i-1})n\rceil \leq j \leq \lceil (a_1+\cdots+a_i)n\rceil - 1$, $i = 1,\ldots,k$, is called the *$(\{\alpha_i\}_{i=1}^k, \mathbf{a}; n)$-name of x*. Since each point in one atom A of $\bigvee_{i=1}^{k}\Big(\bigvee_{j=\lceil (a_0+\cdots+a_{i-1})n\rceil}^{\lceil (a_1+\cdots+a_i)n\rceil-1} T_1^{-j}\tau_{i-1}^{-1}\alpha_i\Big)$ has the same $(\{\alpha_i\}_{i=1}^k, \mathbf{a}; n)$-name, we define

$$C(n,A) := C(n,x)$$

for any $x \in A$, which is called the *$(\{\alpha_i\}_{i=1}^k, \mathbf{a}; n)$-name of A*.

Now if $y \in B_n^{\mathbf{a}}(x,\epsilon)$, then for $i = 1,\ldots,k$ and $\lceil (a_0+\cdots+a_{i-1})n\rceil \leq j \leq \lceil (a_1+\cdots+a_i)n\rceil - 1$, either $T_1^j x$ and $T_1^j y$ belong to the same element of $\tau_{i-1}^{-1}\alpha_i$ or $T_1^j x \in U_\epsilon^i(\alpha_i)$. Hence if $x \in E$, $n \geq \ell$ and $y \in B_n^{\mathbf{a}}(x,\epsilon)$, then the Hamming distance between $(\{\alpha_i\}_{i=1}^k, \mathbf{a}; n)$-name of x and y does not exceed δ_1. Furthermore, $B_n^{\mathbf{a}}(x,\epsilon)$ is contained in the set of points y whose $(\{\alpha_i\}_{i=1}^k, \mathbf{a}; n)$-name is δ_1-close to $(\{\alpha_i\}_{i=1}^k, \mathbf{a}; n)$-name of x. It is clear that the total number $L_n(x)$ of such $(\{\alpha_i\}_{i=1}^k, \mathbf{a}; n)$-names admits the following estimate:

$$\begin{aligned} L_n(x) &\leq \binom{\lceil (a_1+\cdots+a_k)n\rceil}{\lceil\lceil (a_1+\cdots+a_k)n\rceil\delta_1\rceil} M^{\lceil\lceil (a_1+\cdots+a_k)n\rceil\delta_1\rceil} \\ &\leq e^{\delta\lceil (a_1+\cdots+a_k)n\rceil + C} \\ &\leq e^{(a_1+\cdots+a_k)\delta n + C + \delta} \end{aligned}$$

where the second inequality comes from (A.2). More precisely, we have shown that for any $x \in E$ and $n \geq \ell$,

(A.4)

$$B_n^a(x,\epsilon) \subseteq \{y \in X_1 : C(n,y) \text{ is } \delta_1\text{-close to } C(n,x)\}$$
$$= \bigcup \Big\{ A \in \bigvee_{i=1}^{k} \Big(\bigvee_{j=\lceil (a_0+\cdots+a_{i-1})n \rceil}^{\lceil (a_1+\cdots+a_i)n \rceil -1} T_1^{-j}\tau_{i-1}^{-1}\alpha_i \Big) : C(n,A) \text{ is } \delta_1\text{-close to } C(n,x) \Big\}$$

and

$$\#\Big\{ A \in \bigvee_{i=1}^{k} \Big(\bigvee_{j=\lceil (a_0+\cdots+a_{i-1})n \rceil}^{\lceil (a_1+\cdots+a_i)n \rceil -1} T_1^{-j}\tau_{i-1}^{-1}\alpha_i \Big) : C(n,A) \text{ is } \delta_1\text{-close to } C(n,x) \Big\} \leq e^{(a_1+\cdots+a_k)\delta n + C + \delta}. \tag{A.5}$$

Now for $n \in \mathbb{N}$, let E_n denote the set of points x in E such that there exists an element A in $\bigvee_{i=1}^{k} \Big(\bigvee_{j=\lceil (a_0+\cdots+a_{i-1})n \rceil}^{\lceil (a_1+\cdots+a_i)n \rceil -1} T_1^{-j}\tau_{i-1}^{-1}\alpha_i \Big)$ with

$$\mu(A) > e^{\left(-\sum_{i=1}^{k} a_i h_{\mu\circ\tau_{i-1}^{-1}}(T_i,\alpha_i) + (2+a_1+\cdots+a_k)\delta\right)n}$$

and the $(\{\alpha_i\}_{i=1}^k, \mathbf{a}; n)$-name of A is δ_1-close to the $(\{\alpha_i\}_{i=1}^k, \mathbf{a}; n)$-name of x. It is clear that if $x \in E \setminus E_n$, then for each $A \in \bigvee_{i=1}^{k} \Big(\bigvee_{j=\lceil (a_0+\cdots+a_{i-1})n \rceil}^{\lceil (a_1+\cdots+a_i)n \rceil -1} T_1^{-j}\tau_{i-1}^{-1}\alpha_i \Big)$ whose $(\{\alpha_i\}_{i=1}^k, \mathbf{a}; n)$-name is δ_1-close to the $(\{\alpha_i\}_{i=1}^k, \mathbf{a}; n)$-name of x, one has

$$\mu(A) \leq e^{\left(-\sum_{i=1}^{k} a_i h_{\mu\circ\tau_{i-1}^{-1}}(T_i,\alpha_i) + (2+a_1+\cdots+a_k)\delta\right)n}.$$

In the following, we wish to estimate the measure of E_n for $n \geq \ell$.

Let $n \geq \ell$. Put

$$\mathcal{F}_n = \left\{ A \in \bigvee_{i=1}^{k} \Big(\bigvee_{j=\lceil (a_0+\cdots+a_{i-1})n \rceil}^{\lceil (a_1+\cdots+a_i)n \rceil -1} T_1^{-j}\tau_{i-1}^{-1}\alpha_i \Big) : \mu(A) > e^{\left(-\sum\limits_{i=1}^{k} a_i h_{\mu\circ\tau_{i-1}^{-1}}(T_i,\alpha_i) + (2+a_1+\cdots+a_k)\delta\right)n} \right\}.$$

Obviously,

$$\#\mathcal{F}_n \leq e^{\left(\sum_{i=1}^{k} a_i h_{\mu\circ\tau_{i-1}^{-1}}(T_i,\alpha_i) - (2+a_1+\cdots+a_k)\delta\right)n}$$

since $\mu(X_1) = 1$.

Let $x \in E_n$. On the one hand since $x \in B_\ell$,

$$\mu\Big(\bigvee_{i=1}^{k} \Big(\bigvee_{j=\lceil (a_0+\cdots+a_{i-1})n \rceil}^{\lceil (a_1+\cdots+a_i)n \rceil -1} T_1^{-j}\tau_{i-1}^{-1}\alpha_i \Big)(x) \Big) \leq e^{\left(-\sum_{i=1}^{k} a_i h_{\mu\circ\tau_{i-1}^{-1}}(T_i,\alpha_i) + \delta\right)n}$$

by (A.3). On the other hand by the definition of E_n, there exists $A \in \mathcal{F}_n$ with the $(\{\alpha_i\}_{i=1}^k, \mathbf{a}; n)$-name of A is δ_1-close to the $(\{\alpha_i\}_{i=1}^k, \mathbf{a}; n)$-name of x, that is the $(\{\alpha_i\}_{i=1}^k, \mathbf{a}; n)$-name of A is δ_1-close to the $(\{\alpha_i\}_{i=1}^k, \mathbf{a}; n)$-name of

$$\Big(\bigvee_{i=1}^{k} \bigvee_{j=\lceil (a_0+\cdots+a_{i-1})n\rceil}^{\lceil (a_1+\cdots+a_i)n\rceil -1} T_1^{-j}\tau_{i-1}^{-1}\alpha_i\Big)(x).$$

According to this, we have

$$E_n \subset \bigcup\{B : B \in \mathcal{G}_n\} \tag{A.6}$$

where $\mathcal{G}_n$ denotes the set all elements B in $\bigvee_{i=1}^k \Big(\bigvee_{j=\lceil (a_0+\cdots+a_{i-1})n\rceil}^{\lceil (a_1+\cdots+a_i)n\rceil -1} T_1^{-j}\tau_{i-1}^{-1}\alpha_i\Big)$ satisfying $\mu(B) \le e^{\big(-\sum_{i=1}^k a_i h_{\mu\circ\tau_{i-1}^{-1}}(T_i,\alpha_i)+\delta\big)n}$ and the $(\{\alpha_i\}_{i=1}^k, \mathbf{a}; n)$-name of B is δ_1-close to the $(\{\alpha_i\}_{i=1}^k, \mathbf{a}; n)$-name of A for some $A \in \mathcal{F}_n$.

Since for each $A \in \mathcal{F}_n$, the total number of B in $\bigvee_{i=1}^k \Big(\bigvee_{j=\lceil (a_0+\cdots+a_{i-1})n\rceil}^{\lceil (a_1+\cdots+a_i)n\rceil -1} T_1^{-j}\tau_{i-1}^{-1}\alpha_i\Big)$, whose $(\{\alpha_i\}, \mathbf{a}; n)$-name is δ_1-close to the $(\{\alpha_i\}, \mathbf{a}; n)$-name of A, is upper bounded by

$$\binom{\lceil (a_1+\cdots+a_k)n\rceil}{\lceil \lceil (a_1+\cdots+a_k)n\rceil \delta_1\rceil} M^{\lceil \lceil (a_1+\cdots+a_k)n\rceil \delta_1\rceil} \le e^{(a_1+\cdots+a_k)\delta n + C + \delta}.$$

Hence

$$\#\mathcal{G}_n \le e^{(a_1+\cdots+a_k)\delta n + C+\delta}\cdot (\#\mathcal{F}_n) \le e^{\big(\sum_{i=1}^k a_i h_{\mu\circ\tau_{i-1}^{-1}}(T_i,\alpha_i)-2\delta\big)n+C+\delta}.$$

Moreover

$$\mu(E_n) \le e^{\big(-\sum_{i=1}^k a_i h_{\mu\circ\tau_{i-1}^{-1}}(T_i,\alpha_i)+\delta\big)n}\cdot (\#\mathcal{G}_n) \le e^{-\delta n + C+\delta}$$

by (A.6) and the definition of $\mathcal{G}_n$.

Next we take $\ell_2 \ge \ell$ so that $\sum_{n=\ell_2}^\infty e^{-\delta n+C+\delta} < \delta$. Then $\mu(\bigcup_{n\ge \ell_2} E_n) < \delta$. Let $D = E \setminus \bigcup_{n\ge\ell_2} E_n$. Then $\mu(D) > 1-3\delta$. For $x \in D$ and $n \ge \ell_2$, since $x \in E\setminus E_n$, one has

$$\begin{aligned}\mu(B_n^{\mathbf{a}}(x,\epsilon)) &\le e^{(a_1+\cdots+a_k)n+C+\delta}\cdot e^{\big(-\sum_{i=1}^k a_i h_{\mu\circ\tau_{i-1}^{-1}}(T_i,\alpha_i)+(2+a_1+\cdots+a_k)\delta\big)n}\\ &= e^{\big(-\sum_{i=1}^k a_i h_{\mu\circ\tau_{i-1}^{-1}}(T_i,\alpha_i)+2(1+a_1+\cdots+a_k)\delta\big)n+C+\delta}\end{aligned}$$

by (A.4), (A.5) and the definition of E_n. Thus for $x \in D$,

$$\liminf_{n\to+\infty} \frac{-\log \mu(B_n^{\mathbf{a}}(x,\epsilon))}{n} \geq \sum_{i=1}^{k} a_i h_{\mu\circ\tau_{i-1}^{-1}}(T_i,\alpha_i) - 2(1+a_1+\cdots+a_k)\delta$$
$$\geq \min\left\{\frac{1}{\delta}, h_\mu^{\mathbf{a}}(T_1)-\delta\right\} - 2(1+a_1+\cdots+a_k)\delta.$$

This finishes the proof of Theorem A.1. □

In the remaining part of this section, we provide a full proof of Proposition A.2. First we give two lemmas.

Lemma A.4 (cf. [30])**.** *Let $(X,\mathcal{B},\mu,T)$ be a measure preserving dynamical system. Let α,β be two countable measurable partitions of $(X,\mathcal{B},\mu)$ with $H_\mu(\alpha)<\infty, H_\mu(\beta)<\infty$ and $\mathcal{A}$ a sub-σ-algebra of $\mathcal{B}$. Let $I_\mu(\cdot|\cdot)$ denote the conditional information of μ. Then we have the following:*

(i) $I_\mu(\alpha|\mathcal{A})\circ T = I_\mu(T^{-1}\alpha|T^{-1}\mathcal{A})$.
(ii) $I_\mu(\alpha\vee\beta|\mathcal{A}) = I_\mu(\alpha|\mathcal{A})+I_\mu(\beta|\alpha\vee\mathcal{A})$. *In particular,* $H_\mu(\alpha\vee\beta|\mathcal{A}) = H_\mu(\alpha|\mathcal{A})+H_\mu(\beta|\alpha\vee\mathcal{A})$.
(iii) *If $\mathcal{A}_1\subset\mathcal{A}_2\subset\cdots$ is an increasing sub-σ-algebra of $\mathcal{B}$ with $\mathcal{A}_n\uparrow\mathcal{A}$, then $I_\mu(\alpha|\mathcal{A}_n)$ converges almost everywhere and in L^1 to $I_\mu(\alpha|\mathcal{A})$. In particular, $\lim_{n\to+\infty} H_\mu(\alpha|\mathcal{A}_n) = H_\mu(\alpha|\mathcal{A})$.*

Lemma A.5. *Let $(X,\mathcal{B},\mu,T)$ be a measure preserving dynamical system and $F_n\in L^1(X,\mathcal{B},\mu)$ be a sequence that converges almost everywhere and in L^1 to $F\in L^1(X,\mathcal{B},\mu)$ and $\int_X \sup_k |F_n(x)|d\mu(x)<+\infty$. If $f:\mathbb{N}\to\mathbb{N}$ satisfies $f(n)\geq n$ for all $k\in\mathbb{N}$, then*

$$\lim_{n\to+\infty}\frac{1}{n}\sum_{j=0}^{n-1} F_{f(n)-j}(T^jx) = \mathbb{E}_\mu(F|\mathcal{I}_\mu)(x)$$

almost everywhere and in L^1, where $\mathcal{I}_\mu=\{B\in\mathcal{B}:\mu(B\Delta T^{-1}B)=0\}$ and $\mathbb{E}_\mu(F|\mathcal{I}_\mu)$ stands for the conditional expectation of F given $\mathcal{I}_\mu$.

Proof. This is a slight variant of Maker's ergodic theorem [22]. For the convenience of the reader, we give a detailed proof. Since $F\in L^1(X,\mathcal{B},\mu)$, by Birkhoff's ergodic theorem, we have

$$\lim_{n\to+\infty}\frac{1}{n}\sum_{j=0}^{n-1} F(T^jx) = \mathbb{E}_\mu(F|\mathcal{I}_\mu)(x)$$

almost everywhere and in L^1. Since

$$\frac{1}{n}\sum_{j=0}^{n-1} F_{f(n)-j}(T^j x) = \frac{1}{n}\sum_{j=0}^{n-1} F(T^j x) + \frac{1}{n}\sum_{j=0}^{n-1}(F_{f(n)-j}(T^j x) - F(T^j x)),$$

it is suffices to show that

$$\lim_{n\to+\infty} \frac{1}{n}\sum_{j=0}^{n-1} |F_{f(n)-j}(T^j x) - F(T^j x)| = 0$$

almost everywhere and in L^1. Set $Z_m(x) = \sup_{j\geq m} |F_j(x) - F(x)|$ for $m \in \mathbb{N}$. Then $0 \leq Z_m(x) \leq \sup_n |F_n(x)| + |F(x)|$ and $Z_m(x) \to 0$ as $m \to +\infty$ almost everywhere. Since $\sup_n |F_n(x)| + |F(x)| \in L^1(X, \mathcal{B}, \mu)$, we have $\lim_{m\to+\infty} \int Z_m(x) d\mu(x) = 0$ by Lebesgue's dominated convergence theorem. Then we have $\mathbb{E}_\mu(Z_m|\mathcal{I}_\mu) \to 0$ as $m \to +\infty$ almost everywhere and in L^1 (cf. [6, Theorem 34.2]).

Now let $m \in \mathbb{N}$. For $n > m + 1$,

$$\begin{aligned}
&\frac{1}{n}\sum_{j=0}^{n-1} |F_{f(n)-j}(T^j x) - F(T^j x)| \\
&\quad\leq \frac{1}{n}\sum_{j=n-m}^{n-1} |F_{f(n)-j}(T^j x) - F(T^j x)| + \frac{1}{n}\sum_{j=0}^{n-m-1} Z_m(T^j x) \\
&\quad\leq \frac{1}{n}\sum_{j=n-m}^{n-1} Z_1(T^j x) + \frac{n-m}{n}\Big(\frac{1}{n-m}\sum_{j=0}^{n-m-1} Z_m(T^j x)\Big).
\end{aligned}$$

Letting $n \to +\infty$ and using Birkhoff's ergodic theorem we have

$$\limsup_{n\to+\infty} \frac{1}{n}\sum_{j=0}^{n-1} |F_{f(n)-j}(T^j x) - F(T^j x)| \leq \mathbb{E}_\mu(Z_m|\mathcal{I}_\mu)(x)$$

almost everywhere. Since $\mathbb{E}_\mu(Z_m|\mathcal{I}_\mu) \to 0$ almost everywhere and in L^1 as $m \to \infty$, we have

$$\limsup_{n\to+\infty} \frac{1}{n}\sum_{j=0}^{n-1} |F_{f(n)-j}(T^j x) - F(T^j x)| = 0$$

almost everywhere and in L^1, as desired. □

Proof of Proposition A.2. Our proof is adapted from the arguments of Kenyon and Peres in [20, Lemmas 3.2, 4.4].

First we show that for any $a > 0$, $b \geq 0$ and a countable measurable partition β of $(X, \mathcal{B}, \mu)$ with $H_\mu(\beta) < \infty$,

$$\lim_{N \to +\infty} \frac{1}{N} I_\mu\Big(\beta_{\lceil aN \rceil}^{\lceil (a+b)N \rceil - 1}\Big)(x) = b\mathbb{E}_\mu(G|\mathcal{I}_\mu)(x) \tag{A.7}$$

almost everywhere, where $G(x) := I_\mu\Big(\beta | \bigvee_{n=1}^{\infty} T^{-n}\beta\Big)(x)$.

If $b = 0$, then $\beta_{\lceil aN \rceil}^{\lceil (a+b)N \rceil - 1} = \{X, \emptyset\}$ (mod μ) for each $N \in \mathbb{N}$ and so (A.7) holds. Now assume that $b > 0$. Note that

$$I_\mu\Big(\bigvee_{n=\lceil aN \rceil}^{\lceil (a+b)N \rceil - 1} T^{-n}\beta\Big)(x) = I_\mu\Big(\bigvee_{n=0}^{\lceil (a+b)N \rceil - 1} T^{-n}\beta\Big)(x) - I_\mu\Big(\bigvee_{n=0}^{\lceil aN \rceil - 1} T^{-n}\beta | \bigvee_{n=\lceil aN \rceil}^{\lceil (a+b)N \rceil - 1} T^{-n}\beta\Big)(x).$$

By the Shannon-McMillan-Breiman theorem, (A.7) is equivalent to

$$\lim_{N \to +\infty} \frac{1}{N} I_\mu\Big(\bigvee_{n=0}^{\lceil aN \rceil - 1} T^{-n}\beta | \bigvee_{n=\lceil aN \rceil}^{\lceil (a+b)N \rceil - 1} T^{-n}\beta\Big)(x) = a\mathbb{E}_\mu(G|\mathcal{I}_\mu)(x) \tag{A.8}$$

almost everywhere.

Note that

$$\begin{aligned}
&I_\mu\Big(\bigvee_{n=0}^{\lceil aN \rceil - 1} T^{-n}\beta | \bigvee_{n=\lceil aN \rceil}^{\lceil (a+b)N \rceil - 1} T^{-n}\beta\Big)(x) \\
&= I_\mu\Big(\beta | \bigvee_{n=1}^{\lceil (a+b)N \rceil - 1} T^{-n}\beta\Big)(x) + I_\mu\Big(\bigvee_{n=1}^{\lceil aN \rceil - 1} T^{-n}\beta | \bigvee_{n=\lceil aN \rceil}^{\lceil (a+b)N \rceil - 1} T^{-n}\beta\Big)(x) \\
&= I_\mu\Big(\beta | \bigvee_{n=1}^{\lceil (a+b)N \rceil - 1} T^{-n}\beta\Big)(x) + I_\mu\Big(\bigvee_{n=0}^{\lceil aN \rceil - 2} T^{-n}\beta | \bigvee_{n=\lceil aN \rceil - 1}^{\lceil (a+b)N \rceil - 2} T^{-n}\beta\Big)(Tx) \\
&\qquad \vdots \\
&= \sum_{j=0}^{\lceil aN \rceil - 1} I_\mu\Big(\beta | \bigvee_{n=1}^{\lceil (a+b)N \rceil - 1 - j} T^{-n}\beta\Big)(T^j x).
\end{aligned}$$

Write $G_k(x) = I_\mu(\beta | \bigvee_{n=1}^{k-1} T^{-n}\beta)(x)$ for $k \in \mathbb{N}$ and $x \in X$. Then

$$I_\mu\Big(\bigvee_{n=0}^{\lceil aN \rceil - 1} T^{-n}\beta | \bigvee_{n=\lceil aN \rceil}^{\lceil (a+b)N \rceil - 1} T^{-n}\beta\Big)(x) = \sum_{j=0}^{\lceil aN \rceil - 1} G_{\lceil (a+b)N \rceil - j}(T^j x). \tag{A.9}$$

Since $\bigvee_{n=1}^{k-1} T^{-n}\beta \uparrow \bigvee_{n=1}^{\infty} T^{-n}\beta$ when $k \to +\infty$, $G_k \in L^1(X, \mathcal{B}, \mu)$ is a sequence that converges almost everywhere and in L^1 to $G \in L^1(X, \mathcal{B}, \mu)$ by Lemma A.4. As

$H_\mu(\beta) < \infty$, we have $\int_X \sup_k |G_k(x)| d\mu(x) \le H_\mu(\beta) + 1 < \infty$ by Chung's lemma [11]. By (A.9) and Lemma A.5,

$$\begin{aligned}
\lim_{N\to+\infty} \frac{1}{N} I_\mu\Big(\bigvee_{n=0}^{\lceil aN\rceil -1} T^{-n}\beta \Big| \bigvee_{n=\lceil aN\rceil}^{\lceil (a+b)N\rceil -1} T^{-n}\beta\Big)(x) \\
&= a \lim_{N\to+\infty} \frac{1}{\lceil aN\rceil} \sum_{j=0}^{\lceil aN\rceil -1} G_{\lceil (a+b)N\rceil - j}(T^j x) \\
&= a\mathbb{E}_\mu(G|\mathcal{I}_\mu)(x)
\end{aligned}$$

almost everywhere. Hence (A.8) holds, so does (A.7).

Now we are ready to prove (A.1), by induction on k. For $k = 1$, (A.1) reduces to the Shannon-McMillan-Breiman theorem. Assume that (A.1) holds for $k = \ell$ ($\ell \ge 1$). We show below that it holds for $k = \ell + 1$.

Let $k = \ell + 1$. Write $\beta_i = \bigvee_{j=i}^{\ell+1} \alpha_i$ for $i = 1, \ldots, \ell + 1$. Then $\beta_1 \succeq \beta_2 \succeq \cdots \succeq \beta_{\ell+1}$ and $F_i(x) = I_\mu(\beta_i | \bigvee_{n=1}^{+\infty} T^{-n}\beta_i)(x)$ for $i = 1, \ldots, \ell + 1$. Note that

$$\text{(A.10)} \quad \bigvee_{i=1}^{\ell+1} (\alpha_i)_0^{\lceil (a_1+\cdots+a_i)N\rceil -1} = \Big(\bigvee_{i=1}^{\ell} (\beta_i)_0^{\lceil (a_1+\cdots+a_i)N\rceil -1} \Big) \vee (\beta_{\ell+1})_{\lceil (a_1+\cdots+a_\ell)N\rceil}^{\lceil (a_1+\cdots+a_\ell+a_{\ell+1})N\rceil -1}.$$

By the induction assumption and (A.7), we have

$$\text{(A.11)} \quad \begin{aligned}
&\lim_{N\to+\infty} \frac{1}{N} I_\mu\Big(\bigvee_{i=1}^{\ell} (\beta_i)_0^{\lceil (a_1+\cdots+a_i)N\rceil -1}\Big)(x) = \sum_{i=1}^{\ell} a_i \mathbb{E}_\mu(F_i|\mathcal{I}_\mu)(x) \text{ and} \\
&\lim_{N\to+\infty} \frac{1}{N} I_\mu\Big((\beta_{\ell+1})_{\lceil (a_1+\cdots+a_\ell)N\rceil}^{\lceil (a_1+\cdots+a_\ell+a_{\ell+1})N\rceil -1})\Big)(x) = a_{\ell+1}\mathbb{E}_\mu(F_{\ell+1}|\mathcal{I}_\mu)(x)
\end{aligned}$$

almost everywhere. Next we use the idea employed by Algoet and Cover [1] in their elegant "sandwich" proof of the Shannon-McMillan-Breiman theorem. For μ-a.e. $x \in X$, we define

$$Z_m(x) = \frac{\mu\Big(\bigvee_{i=1}^{\ell} (\beta_i)_0^{\lceil (a_1+\cdots+a_i)m\rceil -1}(x)\Big) \cdot \mu\Big((\beta_{\ell+1})_{\lceil (a_1+\cdots+a_\ell)m\rceil}^{\lceil (a_1+\cdots+a_\ell+a_{\ell+1})m\rceil -1}(x)\Big)}{\mu\Big(\Big(\bigvee_{i=1}^{\ell} (\beta_i)_0^{\lceil (a_1+\cdots+a_i)m\rceil -1} \vee (\beta_{\ell+1})_{\lceil (a_1+\cdots+a_\ell)m\rceil}^{\lceil (a_1+\cdots+a_\ell+a_{\ell+1})m\rceil -1}\Big)(x)\Big)}$$

for all $m \in \mathbb{N}$. Then for μ-a.e. $x \in X$, $Z_m(x) > 0$ for all $m \in \mathbb{N}$.

Since

$$\begin{aligned}\int_X Z_m(x)d\mu(x) &= \sum_{\substack{A\in \bigvee_{i=1}^{\ell}(\beta_i)_0^{\lceil (a_1+\cdots+a_i)m\rceil-1}\\ B\in(\beta_{\ell+1})_{\lceil (a_1+\cdots+a_\ell)m\rceil}^{\lceil (a_1+\cdots+a_\ell+a_{\ell+1})m\rceil-1}}} \int_{A\cap B}\frac{\mu(A)\mu(B)}{\mu(A\cap B)}d\mu(x)\\ &= \sum_{\substack{A\in \bigvee_{i=1}^{\ell}(\beta_i)_0^{\lceil (a_1+\cdots+a_i)m\rceil-1}\\ B\in(\beta_{\ell+1})_{\lceil (a_1+\cdots+a_\ell)m\rceil}^{\lceil (a_1+\cdots+a_\ell+a_{\ell+1})m\rceil-1}}} \mu(A)\mu(B)\\ &= 1,\end{aligned}$$

the series $\sum_{m=1}^{\infty}\mu(\{x\in X: Z_m(x)\ge e^{\epsilon m}\})$ converges for every $\epsilon>0$ and the Borel-Canteli Lemma implies that $\limsup_{N\to+\infty}\frac{1}{N}\log Z_N(x)\le 0$ for μ-a.e. $x\in X$. Using the definition of Z_m, (A.10) and (A.11), we obtain

$$\limsup_{N\to+\infty}\frac{1}{N}I_\mu\Big(\bigvee_{i=1}^{\ell+1}(\alpha_i)_0^{\lceil (a_1+\cdots+a_i)N\rceil-1}\Big)(x)\le\sum_{i=1}^{\ell+1}a_i\mathbb{E}_\mu(F_i|\mathcal{I}_\mu)(x)$$

for μ-a.e. $x\in X$.

Conversely, by (A.7) and the induction assumption, we have

(A.12)

$$\lim_{N\to+\infty}\frac{1}{N}I_\mu\Big((\beta_i)_{\lceil (a_1+\cdots+a_\ell)N\rceil}^{\lceil (a_1+\cdots+a_\ell+a_{\ell+1})N\rceil-1}\Big)(x)=a_{\ell+1}\mathbb{E}_\mu(F_i|\mathcal{I}_\mu)(x),\ i=\ell,\ \ell+1 \text{ and}$$

$$\begin{aligned}&\lim_{N\to+\infty}\frac{1}{N}I_\mu\Big(\bigvee_{i=1}^{\ell-1}(\beta_i)_0^{\lceil (a_1+\cdots+a_i)N\rceil-1}\vee(\beta_\ell)_0^{\lceil (a_1+\cdots+a_\ell+a_{\ell+1})N\rceil-1}\Big)(x)\\ &=(a_\ell+a_{\ell+1})\mathbb{E}_\mu(F_\ell|\mathcal{I}_\mu)(x)+\sum_{i=1}^{\ell-1}a_i\mathbb{E}_\mu(F_i|\mathcal{I}_\mu)(x)\end{aligned}$$

almost everywhere. Then for μ-a.e. $x\in X$, we define

$$\begin{aligned}R_m(x)=&\frac{\mu\Big(\Big(\bigvee_{i=1}^{\ell}(\beta_i)_0^{\lceil (a_1+\cdots+a_i)m\rceil-1}\vee(\beta_{\ell+1})_{\lceil (a_1+\cdots+a_\ell)m\rceil}^{\lceil (a_1+\cdots+a_\ell+a_{\ell+1})m\rceil-1}\Big)(x)\Big)}{\mu\Big(\Big(\bigvee_{i=1}^{\ell-1}(\beta_i)_0^{\lceil (a_1+\cdots+a_i)N\rceil-1}\vee(\beta_\ell)_0^{\lceil (a_1+\cdots+a_\ell+a_{\ell+1})N\rceil-1}\Big)(x)\Big)}\\ &\times\frac{\mu\Big((\beta_\ell)_{\lceil (a_1+\cdots+a_\ell)m\rceil}^{\lceil (a_1+\cdots+a_\ell+a_{\ell+1})m\rceil-1}(x)\Big)}{\mu\Big((\beta_{\ell+1})_{\lceil (a_1+\cdots+a_\ell)m\rceil}^{\lceil (a_1+\cdots+a_\ell+a_{\ell+1})m\rceil-1}(x)\Big)}\end{aligned}$$

for all $m\in\mathbb{N}$. Then for μ-a.e. $x\in X$, $R_m(x)>0$ for all $m\in\mathbb{N}$.

Since $\beta_\ell \succeq \beta_{\ell+1}$, we have

$$\int_X R_m(x)d\mu(x) = \sum_{\substack{A\in \bigvee_{i=1}^{\ell}(\beta_i)_0^{\lceil (a_1+\cdots+a_i)m\rceil -1} \\ B\in(\beta_{\ell+1})_{\lceil (a_1+\cdots+a_\ell)m\rceil}^{\lceil (a_1+\cdots+a_\ell+a_{\ell+1})m\rceil-1} \\ C\in(\beta_{\ell})_{\lceil (a_1+\cdots+a_\ell)m\rceil}^{\lceil (a_1+\cdots+a_\ell+a_{\ell+1})m\rceil-1}}} \int_{A\cap B\cap C} \frac{\mu(A\cap B)\mu(B\cap C)}{\mu(A\cap B\cap C)\mu(B)} d\mu(x)$$

$$= \sum_{\substack{A\in \bigvee_{i=1}^{\ell}(\beta_i)_0^{\lceil (a_1+\cdots+a_i)m\rceil -1} \\ B\in(\beta_{\ell+1})_{\lceil (a_1+\cdots+a_\ell)m\rceil}^{\lceil (a_1+\cdots+a_\ell+a_{\ell+1})m\rceil-1} \\ C\in(\beta_{\ell})_{\lceil (a_1+\cdots+a_\ell)m\rceil}^{\lceil (a_1+\cdots+a_\ell+a_{\ell+1})m\rceil-1}}} \frac{\mu(A\cap B)\mu(B\cap C)}{\mu(B)}$$

$$= 1$$

for $m \in \mathbb{N}$. Thus the series $\sum_{m=1}^{\infty} \mu(\{x \in X : R_m(x) \geq e^{\epsilon m}\})$ converges for every $\epsilon > 0$ and the Borel-Canteli Lemma implies that $\limsup_{N\to+\infty} \frac{1}{N}\log R_N(x) \leq 0$ for μ-a.e. $x \in X$. Using the definition R_N, (A.10) and (A.12), we have

$$\liminf_{N\to+\infty} \frac{1}{N} I_\mu\Big(\bigvee_{i=1}^{\ell+1}(\alpha_i)_0^{\lceil (a_1+\cdots+a_i)N\rceil -1}\Big)(x) \geq \sum_{i=1}^{\ell+1} a_i \mathbb{E}_\mu(F_i|\mathcal{I}_\mu)(x)$$

for μ-a.e. $x \in X$. for μ-a.e. $x \in X$. This completes the proof of Proposition A.2. □

Acknowledgements. The first author was partially supported by RGC grants in the Hong Kong Special Administrative Region, China (projects CUHK401112, CUHK401013). The second author was partially supported by NNSF (11225105, 11371339, 11431012).

References

[1] P. H. Algoet and T. M. Cover, A sandwich proof of the Shannon-McMillan-Breiman theorem. *Ann. Probab.* **16** (1988), 899–909.

[2] J. Barral and D. J. Feng, Weighted thermodynamic formalism on subshifts and applications. *Asian J. Math.* **16** (2012), 319–352.

[3] J. Barral and M. Mensi, Gibbs measures on self-affine Sierpinski carpets and their singularity spectrum. *Ergodic Theory Dynam. Systems* **27** (2007), 1419–1443.

[4] J. Barral and M. Mensi, Multifractal analysis of Birkhoff averages on 'self-affine' symbolic spaces. *Nonlinearity* **21** (2008), 2409–2425.

[5] T. Bedford, Crinkly curves, Markov partitions and box dimension in self-similar sets. Ph.D. Thesis, University of Warwick, 1984.

[6] P. Billingsley, *Probability and measure.* Third edition. Wiley Series in Probability and Mathematical Statistics. A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York, 1995.

[7] M. Brin and A. Katok, On local entropy. Geometric dynamics (Rio de Janeiro, 1981), 30–38, *Lecture Notes in Math.*, **1007**, Springer, Berlin, 1983.

[8] R. Bowen, Topological entropy for noncompact sets. *Trans. Amer. Math. Soc.* **184** (1973), 125–136.
[9] R. Bowen, Hausdorff dimension of quasicircles. *Inst. Hautes Études Sci. Publ. Math.* **50** (1979), 11–25.
[10] Y. L. Cao, D. J. Feng and W. Huang, The thermodynamic formalism for sub-additive potentials. *Discrete Contin. Dyn. Syst.* **20** (2008), 639–657.
[11] K. L. Chung, A note on the ergodic theorem of information theory. *Ann. Math. Stat.* **32** (1961), 612-614.
[12] T. Downarowicz and J. Serafin, Fiber entropy and conditional variational principles in compact non-metrizable spaces. *Fund. Math.* 172 (2002), no. 3, 217?-247.
[13] K.J. Falconer. *Fractal geometry. Mathematical foundations and applications.* Second edition. John Wiley & Sons, Inc., Hoboken, NJ, 2003.
[14] H. Federer, *Geometric measure theory.* Springer-Verlag, New York Inc., 1969.
[15] D. J. Feng, Equilibrium states for factor maps between subshifts. *Adv. Math.* **226** (2011), 2470–2502.
[16] D. J. Feng and W. Huang, Lyapunov spectrum of asymptotically sub-additive potentials. *Comm. Math. Phys.* **297** (2010), 1–43.
[17] D. J. Feng and W. Huang, Variational principles for topological entropies of subsets. *J. Funct. Anal.* **263** (2012), 2228–2254.
[18] D. Gatzouras and Y. Peres, The variational principle for Hausdorff dimension: a survey. Ergodic theory of Z^d actions (Warwick, 1993–1994), 113–125, London Math. Soc. Lecture Note Ser., 228, Cambridge Univ. Press, Cambridge, 1996.
[19] J. D. Howroyd, On dimension and on the existence of sets of finite positive Hausdorff measure. *Proc. London Math. Soc.*, **70** (1995), 581–604.
[20] R. Kenyon and Y. Peres, Measures of full dimension on affine-invariant sets. *Ergodic Theory Dynam. Systems* **16** (1996), 307–323.
[21] J. F. King, The singularity spectrum for general Sierpinski carpets. *Adv. Math.* **116** (1995), 1–8.
[22] P. T. Maker. The ergodic theorem for a sequence of functions. *Duke Math. J.* **6** (1940), 27–30.
[23] S. Lalley and D. Gatzouras, Hausdorff and Box Dimensions of Certain Self-Affine Fractals, *Indiana Univ. Math. J.* **41** (1992), 533–568.
[24] F. Ledrappier and P. Walters, A relativised variational principle for continuous transformations. *J. London Math. Soc.* **16** (1977), 568-576.
[25] N. Luzia, A variational principle for the dimension for a class of non-conformal repellers. *Ergodic Theory Dynam. Systems* **26** (2006), 821-845.
[26] P. Mattila, *Geometry of sets and measures in Euclidean spaces.* Cambridge University Press, 1995.
[27] C. McMullen, The Hausdorff dimension of general Sierpinski carpets. *Nagoya Math. J.* **96** (1984), 1–9.
[28] M. Misiurewicz, A short proof of the variational principle for a $\mathbb{Z}_+^{\mathbb{N}}$ action on a compact space. *Asterisque* **40** (1976), 147-187.
[29] L. Olsen, Self-affine multifractal Sierpinski sponges in $\mathbb{R}^d$. *Pacific J. Math.* **183** (1998), 143–199.
[30] W. Parry, *Topics in ergodic theory.* Cambridge Tracts in Mathematics, 75. Cambridge University Press, Cambridge-New York, 1981.
[31] Ya. Pesin, *Dimension theory in dynamical systems. Contemporary views and applications.* Chicago Lectures in Mathematics. University of Chicago Press, Chicago, IL, 1997.
[32] Ya. Pesin and B. S. Pitskel', Topological pressure and the variational principle for non- compact sets. *Functional Anal. Appl.* **18** (1984), 50–63.

[33] D. Ruelle, Statistical mechanics on a compact set with Z^v action satisfying expansiveness and specification. *Trans. Amer. Math. Soc.* **187** (1973), 237–251.
[34] D. Ruelle, *Thermodynamic formalism.* The mathematical structures of classical equilibrium statistical mechanics. Encyclopedia of Mathematics and its Applications, 5. Addison-Wesley Publishing Co., Reading, Mass., 1978.
[35] D. Ruelle, Repellers for real analytic maps. *Ergodic Theory Dynam. Systems* **2**(1982), 99–107.
[36] P. Walters, A variational principle for the pressure of continuous transformations. *Amer. J. Math.* **97** (1975), 937–971.
[37] P. Walters, *An introduction to ergodic theory*, Springer-Verlag, Berlin, Heidelberg, New York, 1982.
[38] Y. Yayama, Applications of a relative variational principle to dimensions of nonconformal expanding maps. *Stoch. Dyn.* **11** (2011), 643-679.
[39] Y. Zhao and Y. L. Cao, On the topological pressure of random bundle transformations in sub-additive case. *J. Math. Anal. Appl.* **342** (2008), 715-725.

Department of Mathematics, The Chinese University of Hong Kong, Shatin, Hong Kong,

E-mail address: djfeng@math.cuhk.edu.hk

Department of Mathematics, Sichuan University, Chengdu, Sichuan 610064, China, *and*, School of Mathematical Sciences, University of Science and Technology of China, Hefei, Anhui 230026, China,

E-mail address: wenh@mail.ustc.edu.cn