Variational principles for topological entropies of subsets

De-Jun Feng\textsuperscript{a}, Wen Huang\textsuperscript{b,*}

\textsuperscript{a} Department of Mathematics, The Chinese University of Hong Kong Shatin, Hong Kong
\textsuperscript{b} Department of Mathematics, University of Science and Technology of China, Hefei, Anhui 230026, PR China

Received 5 January 2012; accepted 11 July 2012
Available online 24 July 2012
Communicated by D. Voiculescu

Abstract

Let \((X, T)\) be a topological dynamical system. We define the measure-theoretical lower and upper entropies \(h_\mu(T), \bar{h}_\mu(T)\) for any \(\mu \in M(X)\), where \(M(X)\) denotes the collection of all Borel probability measures on \(X\). For any non-empty compact subset \(K\) of \(X\), we show that

\[
h^B_{\text{top}}(T, K) = \sup \{ h_\mu(T) : \mu \in M(X), \mu(K) = 1 \},
\]

\[
h^P_{\text{top}}(T, K) = \sup \{ \bar{h}_\mu(T) : \mu \in M(X), \mu(K) = 1 \},
\]

where \(h^B_{\text{top}}(T, K)\) denotes the Bowen topological entropy of \(K\), and \(h^P_{\text{top}}(T, K)\) the packing topological entropy of \(K\). Furthermore, when \(h_{\text{top}}(T) < \infty\), the first equality remains valid when \(K\) is replaced by any analytic subset of \(X\). The second equality always extends to any analytic subset of \(X\).

© 2012 Elsevier Inc. All rights reserved.

Keywords: Topological entropies; Measure-theoretical entropies; Variational principles

1. Introduction

Throughout this paper, by a topological dynamical system (TDS) \((X, T)\) we mean a compact metric space \(X\) together with a continuous self-map \(T : X \to X\). Let \(M(X), M(X, T),\)
and $E(X, T)$ denote respectively the sets of all Borel probability measures, $T$-invariant Borel probability measures, and $T$-invariant ergodic Borel probability measures on $X$. By a measure theoretical dynamical system (m.t.d.s.) we mean $(Y, C, v, T)$, where $Y$ is a set, $C$ is a $\sigma$-algebra over $Y$, $v$ is a probability measure on $C$ and $T$ is a measure preserving transformation. A probability measure $\mu \in M(X, T)$ induces a m.t.d.s. $(X, B_X, \mu, T)$ or just $(X, \mu, T)$, where $B_X$ is the $\sigma$-algebra of Borel subsets of $X$.

In 1958 Kolmogorov [21] associated to any m.t.d.s. $(Y, C, v, T)$ an isomorphic invariant, namely the measure-theoretical entropy $h_v(T)$. Later on in 1965, Adler, Konheim and McAndrew [1] introduced for any TDS $(X, T)$ an analogous notion of topological entropy $h_{\text{top}}(T)$, as an invariant of topological conjugacy. There is a basic relation between topological entropy and measure-theoretic entropy: if $(X, T)$ is a TDS, then $h_{\text{top}}(T) = \sup \{ h_\mu(T) : \mu \in M(X, T) \}$. This variational principle was proved by Goodwyn, Dinaburg and Goodman [14, 8, 13], and plays a fundamental role in ergodic theory and dynamical systems (cf. [28, 31]). Recently, Kerr and Li obtained a variational principle of topological entropy and measure-theoretic entropy for actions of sofic group [19, 20].

In 1973, Bowen [4] introduced the topological entropy $h_{\text{top}}^B(T, Z)$ for any set $Z$ in a TDS $(X, T)$ in a way resembling Hausdorff dimension, which we call the Bowen topological entropy (see Section 2 for the definition). In particular, $h_{\text{top}}^B(T, X) = h_{\text{top}}(T)$. Later in 1984, inspired by Bowen’s approach, Pesin and Pitskel’ [29] extended the notion of topological pressure to arbitrary subsets of $X$. The notion of topological entropy and topological pressure of (arbitrary) subsets play an important role in topological dynamics and dimension theory [28].

A question arises naturally whether there is certain variational relation between Bowen topological entropy and measure-theoretic entropy for arbitrary non-invariant compact set, or Borel set in general. However, when $K \subseteq X$ is $T$-invariant but not compact, or $K$ is compact but not $T$-invariant, it may happen that $h_{\text{top}}^B(T, K) > 0$ but $\mu(K) = 0$ for any $\mu \in M(X, T)$ (see Example 1.5). Hence we don’t expect to have such variational principle on the class $M(X, T)$. For our purpose, we need to define the measure-theoretic entropy for elements in $M(X)$.

Fix a compatible metric $d$ on $X$. For any $n \in \mathbb{N}$, the $n$-th Bowen metric $d_n$ on $X$ is defined by

$$d_n(x, y) = \max\{ d(T^k(x), T^k(y)) : k = 0, \ldots, n - 1 \}. \quad (1.1)$$

For every $\epsilon > 0$ we denote by $B_n(x, \epsilon), \overline{B}_n(x, \epsilon)$ the open (resp. closed) ball of radius $\epsilon$ in the metric $d_n$ around $x$, i.e.,

$$B_n(x, \epsilon) = \{ y \in X : d_n(x, y) < \epsilon \}, \quad \overline{B}_n(x, \epsilon) = \{ y \in X : d_n(x, y) \leq \epsilon \}. \quad (1.2)$$

Following the idea of Brin and Katok [6], we give the following.

**Definition 1.1.** Let $\mu \in M(X)$. The measure-theoretical lower and upper entropies of $\mu$ are defined respectively by

$$\underline{h}_\mu(T) = \int h_\mu(T, x) \, d\mu(x), \quad \overline{h}_\mu(T) = \int \overline{h}_\mu(T, x) \, d\mu(x),$$

where
\[
\begin{align*}
\underline{h}_\mu(T, x) &= \lim_{\epsilon \to 0} \liminf_{n \to +\infty} -\frac{1}{n} \log \mu(B_n(x, \epsilon)) , \\
\overline{h}_\mu(T, x) &= \lim_{\epsilon \to 0} \limsup_{n \to +\infty} -\frac{1}{n} \log \mu(B_n(x, \epsilon)).
\end{align*}
\]

Brin and Katok [6] proved that for any \( \mu \in M(X, T) \), \( \underline{h}_\mu(T, x) = \overline{h}_\mu(T, x) \) for \( \mu \)-a.e. \( x \in X \), and \( \int \overline{h}_\mu(T, x) \, d\mu(x) = h_\mu(T) \). Hence for \( \mu \in M(X, T) \),

\[ \underline{h}_\mu(T) = \overline{h}_\mu(T) = h_\mu(T). \]

To formulate our results, we need to introduce an additional notion. A set in a metric space is said to be \textit{analytic} if it is a continuous image of the set \( \mathcal{N} \) of infinite sequences of natural numbers (with its product topology). It is known that in a Polish space, the analytic subsets are closed under countable unions and intersections, and any Borel set is analytic (cf. Federer [12, 2.2.10]).

The main results of this paper are the following two theorems.

**Theorem 1.2.** Let \((X, T)\) be a TDS.

(i) If \( K \subseteq X \) is non-empty and compact, then

\[ h^B_{\text{top}}(T, K) = \sup \{ \underline{h}_\mu(T): \mu \in M(X), \mu(K) = 1 \}. \]

(ii) Assume that \( h_{\text{top}}(T) < \infty \). If \( Z \subseteq X \) is analytic, then

\[ h^B_{\text{top}}(T, Z) = \sup \{ h^B_{\text{top}}(T, K): K \subseteq Z \text{ is compact} \}. \tag{1.3} \]

**Theorem 1.3.** Let \((X, T)\) be a TDS.

(i) If \( K \subseteq X \) is non-empty and compact, then

\[ h^P_{\text{top}}(T, K) = \sup \{ \overline{h}_\mu(T): \mu \in M(X), \mu(K) = 1 \}, \]

where \( h^P_{\text{top}}(T, K) \) denotes the packing topological entropy of \( K \) (see Section 2 for the definition).

(ii) If \( Z \subseteq X \) is analytic, then

\[ h^P_{\text{top}}(T, Z) = \sup \{ h^P_{\text{top}}(T, K): K \subseteq Z \text{ is compact} \}. \tag{1.4} \]

The above two theorems establish the variational principles for Bowen and packing topological entropies of arbitrary Borel sets in a dual manner. They provide as a kind of extension of the classical variational principle for topological entropy of compact invariant sets. In the remainder of this section, we give two examples which motivated this paper.

**Example 1.4.** Let \((X, T)\) denote the one-sided full shift over a finite alphabet \( \{1, 2, \ldots, \ell\} \), where \( \ell \) is an integer \( \geq 2 \). Endow \( X \) with the metric \( d(x, y) = e^{-n} \) for \( x = (x_j)_{j=1}^\infty \) and \( y = (y_j)_{j=1}^\infty \),
where $n$ is the largest integer such that $x_j = y_j$ ($1 \leq j \leq n$). It is easy to check by definition that for any $E \subseteq X$,
\[
h^B_{\text{top}}(T, E) = \dim_H E, \quad h^P_{\text{top}}(T, E) = \dim_P E,
\]
where $\dim_H E, \dim_P E$ denote respectively the Hausdorff dimension and the packing dimension of $E$ in the ultra-metric space $(X, d)$ (cf. [27]). It is a well-known fact in geometric measure theory (cf. [27]) that, for any analytic set $Z \subseteq X$ with $\dim_H Z > 0$, and any $0 \leq s < \dim_H Z$, $0 \leq t < \dim_P Z$, there exist compact sets $K_1, K_2 \subset Z$ such that
\[
0 < \mathcal{H}^s(K_1) < \infty, \quad 0 < \mathcal{P}^t(K_2) < \infty,
\]
where $\mathcal{H}^s, \mathcal{P}^t$ denote respectively the $s$-dimensional Hausdorff measure and packing measure, and hence $\dim_H K_1 = s, \dim_P K_2 = t$. Furthermore, for $\mathcal{H}^s$-a.e. $x \in K_1$, and $\mathcal{P}^t$-a.e. $y \in K_2$,
\[
\liminf_{r \to 0} \frac{\log \mathcal{H}^s(K_1 \cap B_r(x))}{\log r} = s, \quad \limsup_{r \to 0} \frac{\log \mathcal{P}^t(K_2 \cap B_r(x))}{\log r} = t,
\]
where $B_r(x)$ denotes the open ball centered at $x$ of radius $r$. This can be used to derive Theorems 1.2–1.3 in the full shift case with some additional density arguments as in [27, p. 99, Exercises 6–7].

**Example 1.5.** Again let $(X, T)$ denote the one-sided full shift over a finite alphabet $\{1, 2, \ldots, \ell\}$. Define $\varphi : X \to \mathbb{R}$ as
\[
\varphi(x) = \begin{cases} 
1 & \text{if } x_1 = 1, \\
0 & \text{otherwise}
\end{cases}
\]
for $x = (x_i)_{i=1}^\infty \in X$. Let $E$ denote the set of “non-typical points” associated with the Birkhoff average of $\varphi$, i.e.,
\[
E = \left\{ x \in X : \liminf_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} \varphi(T^i x) \neq \limsup_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} \varphi(T^i x) \right\}.
\]
It is easy to see that $E$ is $T$-invariant and Borel. By the Birkhoff ergodic theorem, $\mu(E) = 0$ for any $\mu \in M(X, T)$. However $h^B_{\text{top}}(T, E) = h_{\text{top}}(T) = \log \ell$ (cf. [2]). Furthermore, as we mention in Example 1.4 that for any $0 \leq s < \log \ell$, there exists a compact set $K \subset E$ such that $h^B_{\text{top}}(T, K) = \dim_H K = s$.

For the convenience of the readers, we illustrate some rough ideas in the proof of Theorem 1.2(i). To see the lower bound, let $\mu \in M(X)$ with $\mu(K) = 1$, and take $u \in \mathbb{R}$ with $u < h_H(T)$. Then we can show that there exist $\epsilon > 0, N \in \mathbb{N}$ and a subset $A$ of $K$ with $\mu(A) > 0$ so that
\[
\mu(B_n(x, \epsilon)) \leq e^{-nu}, \quad \forall x \in A, \ n \geq N.
\]
Now for any Bowen ball $B_n(y, \epsilon/2)$ with $n \geq N$ that intersects $A$, if taking $x \in B_n(y, \epsilon/2) \cap A$, then one has $B_n(y, \epsilon/2) \subseteq B_n(x, \epsilon)$ and thus

$$\mu(B_n(y, \epsilon/2)) \leq \mu(B_n(x, \epsilon)) \leq e^{-nu}.$$ 

According to this observation, if a countable collection of Bowen balls $\{B_{n_i}(x_i, \epsilon/2)\}_i$ covers $K$ and satisfies $n_i \geq N$, then

$$\sum_i e^{-n_i u} \geq \sum_i \mu(B_{n_i}(x_i, \epsilon/2) \cap A) \geq \mu\left(\bigcup_i B_{n_i}(x_i, \epsilon/2) \cap A\right) = \mu(A) > 0.$$ 

By Bowen’s definition, we can derive that $h^B_{\text{top}}(T, K) \geq u$.

To see the upper bound in Theorem 1.2(i), we need to show that for any $s < h^B_{\text{top}}(T, K)$, there exists some $\mu \in M(X)$ with $\mu(K) = 1$ such that $h_\mu(T) \geq s$. The proof of this step is inspired by Howroyd’s elegant proof [15] of Frostman’s lemma for compact metric spaces (which says, for any compact set $E \subset X$ with Hausdorff dimension greater than $t$, there exists a Borel probability measure $\mu$ on $X$ with $\mu(E) = 1$ so that $\mu(B(x, r)) < cr^t$ for some constant $c > 0$ and any $r > 0$, $x \in X$). Indeed, by extending some ideas in [15] we can manage to prove a dynamical version of Frostman’s lemma (see Lemma 3.4), which says that there exists a Borel probability measure $\mu$ supported on $K$ so that there exist $c, \epsilon > 0$ and $N \in \mathbb{N}$ such that

$$\mu(B_n(x, \epsilon)) \leq \frac{1}{c} e^{-ns}, \quad \forall x \in X, \ n \geq N.$$ 

This implies $h_\mu(T) \geq s$. A key part for proving the above dynamical Frostman’s lemma is to compare the Bowen topological entropy and a kind of weighted topological entropy introduced in Section 3.1. For this, we adopt some classical techniques in geometry measure theory that deal with weighted Hausdorff measures.

Theorem 1.3 provides a dual version of Theorem 1.2. However, its proof is quite different from that of Theorem 1.3. Our approach (see Section 4 for details) is inspired by the fundamental work of Joyce and Preiss [18] on the existence of subsets with finite packing measures.

The paper is organized as follows. In Section 2 we give the definitions and some basic properties of several topological entropies of subsets in a TDS: upper capacity topological entropy, Bowen topological entropy, packing topological entropy. In Section 3, we prove Theorem 1.2. In Section 4, we prove Theorem 1.3. In Section 5, we include a table of main notation and conventions used in the paper.

2. Topological entropies of subsets

In this section, we give the definitions and some basic properties of several topological entropies of subsets in a TDS: upper capacity topological entropy, Bowen topological entropy and packing topological entropy.

Let $(X, d)$ be a compact metric space and $T : X \to X$ a continuous transformation. Let $d_n$ and $B_n(x, \epsilon)$ be defined as in (1.1)–(1.2).
2.1. Upper capacity topological entropy

Let $Z \subseteq X$ be a non-empty set. For $\epsilon > 0$, a set $E \subseteq Z$ is called an $(n, \epsilon)$-separated set of $Z$ if $x, y \in E, x \neq y$ implies $d_n(x, y) > \epsilon$; $E \subseteq X$ is called $(n, \epsilon)$-spanning set of $Z$, if for any $x \in Z$, there exists $y \in E$ with $d_n(x, y) \leq \epsilon$. Let $r_n(Z, \epsilon)$ denote the largest cardinality of $(n, \epsilon)$-separated sets for $Z$, and $\tilde{r}_n(Z, \epsilon)$ the smallest cardinality of $(n, \epsilon)$-spanning sets of $Z$. The upper capacity topological entropy of $T$ restricted on $Z$, or simply, the upper capacity topological entropy of $Z$ is defined as

$$h_{top}^{UC}(T, Z) = \lim_{\epsilon \to 0} \limsup_{n \to \infty} \frac{1}{n} \log r_n(Z, \epsilon) = \lim_{\epsilon \to 0} \limsup_{n \to \infty} \frac{1}{n} \log \tilde{r}_n(Z, \epsilon).$$

We remark that the second equality holds for each $Z \subseteq X$ (cf. [31, p. 169]). The quantity $h_{top}^{UC}(T, Z)$ is the straightforward generalization of the Adler–Konheim–McAndrew definition [1] of the topological entropy to arbitrary subsets.

2.2. Bowen topological entropy

Suppose that $U$ is a finite open cover of $X$. Denote $\text{diam}(U) := \max\{\text{diam}(U): U \in \mathcal{U}\}$. For $n \geq 1$ we denote by $W_n(U)$ the collection of strings $U = U_1 \ldots U_n$ with $U_i \in \mathcal{U}$. For $U \in W_n(U)$ we call the integer $m(U) = n$ the length of $U$ and define

$$X(U) = U_1 \cap T^{-1}U_2 \cap \cdots \cap T^{-(n-1)}U_n$$
$$= \{x \in X: T^{j-1}x \in U_j \text{ for } j = 1, \ldots, n\}.$$

Let $Z \subseteq X$. We say that $\Lambda \subset \bigcup_{n \geq 1} W_n(U)$ covers $Z$ if $\bigcup_{U \in \Lambda} X(U) \supseteq Z$. For $s \in \mathbb{R}$, define

$$\mathcal{M}^s_N(\mathcal{U}, Z) = \inf_{\Lambda} \sum_{U \in \Lambda} \exp(-sm(U)),$$

where the infimum is taken over all $\Lambda \subset \bigcup_{j \geq N} W_j(\mathcal{U})$ that cover $Z$. Clearly $\mathcal{M}^s_N(\mathcal{U}, \cdot)$ is a finite outer measure on $X$, and

$$\mathcal{M}^s_N(\mathcal{U}, Z) = \inf\{\mathcal{M}^s_N(\mathcal{U}, G): G \supseteq Z, G \text{ is open}\}. \quad (2.1)$$

Note that $\mathcal{M}^s_N(\mathcal{U}, Z)$ increases as $N$ increases. Define $\mathcal{M}^s(\mathcal{U}, Z) = \lim_{N \to \infty} \mathcal{M}^s_N(\mathcal{U}, Z)$ and

$$h_{top}^{B}(T, \mathcal{U}, Z) = \inf\{s: \mathcal{M}^s(\mathcal{U}, Z) = 0\} = \sup\{s: \mathcal{M}^s(\mathcal{U}, Z) = +\infty\}.$$

Set

$$h_{top}^{B}(T, Z) = \sup_{\mathcal{U}} h_{top}^{B}(T, \mathcal{U}, Z), \quad (2.2)$$

where $\mathcal{U}$ runs over finite open covers of $Z$. We call $h_{top}^{B}(T, Z)$ the Bowen topological entropy of $T$ restricted to $Z$ or, simply, the topological entropy of $Z$. This quantity was first introduced by
Bowen in [4]. It is known (see, e.g. [28, Theorem 11.1]) that
\[
\sup_{\mathcal{U}} h^B_{\text{top}}(T, \mathcal{U}, Z) = \lim_{\text{diam}(\mathcal{U}) \to 0} h^B_{\text{top}}(T, \mathcal{U}, Z). \tag{2.3}
\]

The Bowen topological entropy of subsets can be defined in an alternative way. For \( Z \subseteq X \), \( s \geq 0 \), \( N \in \mathbb{N} \) and \( \epsilon > 0 \), define
\[
\mathcal{M}_{s,N,\epsilon}^s(Z) = \inf \sum_i \exp(-sn_i),
\]
where the infimum is taken over all finite or countable families \( \{B_{n_i}(x_i, \epsilon)\} \) such that \( x_i \in X \), \( n_i \geq N \) and \( \bigcup_i B_{n_i}(x_i, \epsilon) \supseteq Z \). The quantity \( \mathcal{M}_{s,N,\epsilon}^s(Z) \) does not decrease as \( N \) increases and \( \epsilon \) decreases, hence the following limits exist:
\[
\mathcal{M}_s^s(Z) = \lim_{N \to \infty} \mathcal{M}_{s,N,\epsilon}^s(Z), \quad \mathcal{M}^s(Z) = \lim_{\epsilon \to 0} \mathcal{M}_s^s(Z).
\]
The Bowen topological entropy \( h^B_{\text{top}}(T, Z) \) can be equivalently defined as a critical value of the parameter \( s \), where \( \mathcal{M}^s(Z) \) jumps from \( \infty \) to 0, i.e.
\[
\mathcal{M}^s(Z) = \begin{cases} 
0, & s > h^B_{\text{top}}(T, Z), \\
\infty, & s < h^B_{\text{top}}(T, Z).
\end{cases}
\]
For details, see [28, p. 74].

2.3. Packing topological entropy

Let \( Z \subseteq X \). For \( s \geq 0 \), \( N \in \mathbb{N} \) and \( \epsilon > 0 \), define
\[
P_{s,N,\epsilon}^s(Z) = \sup \sum_i \exp(-sn_i),
\]
where the supremum is taken over all finite or countable pairwise disjoint families \( \{\overline{B}_{n_i}(x_i, \epsilon)\} \) such that \( x_i \in Z \), \( n_i \geq N \) for all \( i \), where
\[
\overline{B}_n(x, \epsilon) := \{y \in X: d_n(x, y) \leq \epsilon\}.
\]
The quantity \( P_{s,N,\epsilon}^s(Z) \) does not decrease as \( N, \epsilon \) decrease, hence the following limit exists:
\[
P_{s,\epsilon}^s(Z) = \lim_{N \to \infty} P_{s,N,\epsilon}^s(Z).
\]
Define
\[
P_{s,\epsilon}^s(Z) = \inf \left\{ \sum_{i=1}^{\infty} P_{s,\epsilon}^s(Z_i): \bigcup_{i=1}^{\infty} Z_i \supseteq Z \right\}.
\]
Clearly, \( P^s \) satisfies the following property: if \( Z \subseteq \bigcup_{i=1}^{\infty} Z_i \), then \( P^s_e(Z) \leq \sum_{i=1}^{\infty} P^s_e(Z_i) \). There exists a critical value of the parameter \( s \), which we will denote by \( h_{\text{top}}^P(T, Z, \epsilon) \), where \( P^s_e(Z) \) jumps from \( \infty \) to 0, i.e.

\[
P^s_e(Z) = \begin{cases} 
0, & s > h_{\text{top}}^P(T, Z, \epsilon), \\
\infty, & s < h_{\text{top}}^P(T, Z, \epsilon).
\end{cases}
\]

Note that \( h_{\text{top}}^P(T, Z, \epsilon) \) increases when \( \epsilon \) decreases. We call

\[
h_{\text{top}}^P(T, Z) := \lim_{\epsilon \to 0} h_{\text{top}}^P(T, Z, \epsilon)
\]

the packing topological entropy of \( T \) restricted to \( Z \) or, simply, the packing topological entropy of \( Z \), when there is no confusion about \( T \). This quantity is defined in way which resembles the packing dimension. We remark that an equivalent definition of packing topological entropy was given earlier in [17].

2.4. Some basic properties

**Proposition 2.1.**

(i) For \( Z \subseteq Z' \),

\[
h_{\text{top}}^{UC}(T, Z) \leq h_{\text{top}}^{UC}(T, Z'), \quad h_{\text{top}}^{B}(T, Z) \leq h_{\text{top}}^{B}(T, Z'), \quad h_{\text{top}}^{P}(T, Z) \leq h_{\text{top}}^{P}(T, Z').
\]

(ii) For \( Z \subseteq \bigcup_{i=1}^{\infty} Z_i, s \geq 0 \) and \( \epsilon > 0 \), we have

\[
\mathcal{M}^s_e(Z) \leq \sum_{i=1}^{\infty} \mathcal{M}^s_e(Z_i),
\]

\[
h_{\text{top}}^{B}(T, Z) \leq \sup_{i \geq 1} h_{\text{top}}^{B}(T, Z_i), \quad h_{\text{top}}^{P}(T, Z) \leq \sup_{i \geq 1} h_{\text{top}}^{P}(T, Z_i).
\]

(iii) For any \( Z \subseteq X, h_{\text{top}}^{B}(T, Z) \leq h_{\text{top}}^{P}(T, Z) \leq h_{\text{top}}^{UC}(T, Z) \).

(iv) Furthermore, if \( Z \) is \( T \)-invariant and compact, then

\[
h_{\text{top}}^{B}(T, Z) = h_{\text{top}}^{P}(T, Z) = h_{\text{top}}^{UC}(T, Z).
\]

**Proof.** (i) and (ii) follow directly from the definitions of topological entropies. To see (iii), let \( Z \subseteq X \) and assume \( 0 < s < h_{\text{top}}^{B}(T, Z) \). For any \( n \in \mathbb{N} \) and \( \epsilon > 0 \), let \( R = R_n(Z, \epsilon) \) be the largest number so that there is a disjoint family \( \{ \overline{B}_n(x_i, \epsilon) \}_{i=1}^{R} \) with \( x_i \in Z \). Then it is easy to see that for any \( \delta > 0 \),

\[
\bigcup_{i=1}^{R} \overline{B}_n(x_i, 2\epsilon + \delta) \supseteq Z.
\]
which implies that \( M_{s,2\epsilon+\delta}(Z) \leq R \epsilon^{ns} \leq P_{n,\epsilon}(Z) \) for any \( s \geq 0 \), and hence \( M_{s,2\epsilon+\delta}(Z) \leq P_{\epsilon}(Z) \). By (ii), \( M_{s,2\epsilon+\delta}(Z) \leq P_{\epsilon}(Z) \). Since \( 0 < s < h_{\text{top}}^P(T, Z) \), we have \( M^s(Z) = \infty \) and thus \( M_{s,2\epsilon+\delta}(Z) \geq 1 \) when \( \epsilon \) and \( \delta \) are small enough. Hence \( P_{\epsilon}(Z) \geq 1 \) and \( h_{\text{top}}^P(T, Z, \epsilon) \geq s \) when \( \epsilon \) is small. Therefore \( h_{\text{top}}^P(T, Z) = \lim_{\epsilon \to 0} h_{\text{top}}^P(T, Z, \epsilon) \geq s \). This implies that \( h_{\text{top}}^P(T, Z) \leq h_{\text{top}}^{P}(T, Z) \).

Next we show that \( h_{\text{top}}^P(T, Z) \leq h_{\text{top}}^{UC}(T, Z) \). Our argument is modified slightly from the proof of [11, Lemma 3.7]. Assume that \( h_{\text{top}}^P(T, Z) > 0 \); otherwise there is nothing left to prove. Choose \( 0 < t < h_{\text{top}}^P(T, Z) \). Then there exists \( \delta > 0 \) such that for \( 0 < \epsilon < \delta \), \( h_{\text{top}}^P(T, Z, \epsilon) > s \) and thus \( P_{\epsilon}(Z) \geq P_{\epsilon}(Z) = \infty \). Thus for any \( N \), there exists a countable pairwise disjoint family \( \{B_{n_i}(x_i, \epsilon)\} \) such that \( x_i \in Z, n_i \geq N \) for all \( i \), and \( 1 < \sum_i e^{-n_i s} \). For each \( k \), let \( m_k \) be the number of \( i \) so that \( n_i = k \). Then we have

\[
1 < \sum_{k=N}^{\infty} m_k e^{-ks}.
\]

There must be some \( k \geq N \) with \( m_k > e^{kt}(1 - e^{-s}) \), otherwise the above sum is at most \( \sum_{k=1}^{\infty} e^{kt}(1 - e^{-s}) < 1 \). Let \( r_k(Z, \epsilon) \) denote the largest cardinality of \((k, \epsilon)\)-separated sets for \( Z \). Then \( r_k(Z, \epsilon) \geq m_k > e^{kt}(1 - e^{-s}) \). Hence \( \limsup_{n \to \infty} \frac{1}{n} \log r_n(Z, \epsilon) \geq t \). Letting \( \epsilon \to 0 \), we obtain \( h_{\text{top}}^{UC}(T, Z) \geq t \). This is true for any \( 0 < t < h_{\text{top}}^P(T, Z) \) so \( h_{\text{top}}^{UC}(T, Z) = h_{\text{top}}^P(T, Z) \).

When \( Z \subseteq X \) is \( T \)-invariant and compact, Bowen [4] proved that \( h_{\text{top}}^R(T, Z) = h_{\text{top}}^{UC}(T, Z) \); this together with (iii) yields (iv). \( \square \)

3. Variational principle for Bowen topological entropy of subsets

3.1. Weighted topological entropy

For any bounded function \( f : X \to \mathbb{R} \), \( N \in \mathbb{N} \) and \( \epsilon > 0 \), define

\[
W_{N,\epsilon}^s(f) = \inf \sum_i c_i \exp(-sn_i),
\]

where the infimum is taken over all finite or countable families \( \{(B_{n_i}(x_i, \epsilon), c_i)\} \) such that \( 0 < c_i < \infty, x_i \in X, n_i \geq N \) and

\[
\sum_i c_i \chi_{B_i} \geq f,
\]

where \( B_i := B_{n_i}(x_i, \epsilon) \), and \( \chi_A \) denotes the characteristic function of \( A \), i.e., \( \chi_A(x) = 1 \) if \( x \in A \) and 0 if \( x \in X \setminus A \).

For \( Z \subseteq X \) and \( f = \chi_Z \) we set \( W_{N,\epsilon}^s(Z) = W_{N,\epsilon}^s(\chi_Z) \). The quantity \( W_{N,\epsilon}^s(Z) \) does not decrease as \( N \) increases and \( \epsilon \) decreases, hence the following limits exist:

\[
W_{\epsilon}(Z) = \lim_{N \to \infty} W_{N,\epsilon}^s(Z), \quad W^s(Z) = \lim_{\epsilon \to 0} W_{\epsilon}^s(Z).
\]
We remark that $\mathcal{W}^s$ is defined in a way which resembles the weighted Hausdorff measure in geometric measure theory (cf. [12,27]). Clearly, there exists a critical value of the parameter $s$, which we will denote by $h_{\text{top}}^{\text{WB}}(T,Z)$, where $\mathcal{W}^s(Z)$ jumps from $\infty$ to 0, i.e.

$$\mathcal{W}^s(Z) = \begin{cases} 0, & s > h_{\text{top}}^{\text{WB}}(T,Z), \\ \infty, & s < h_{\text{top}}^{\text{WB}}(T,Z). \end{cases}$$

We call $h_{\text{top}}^{\text{WB}}(T,Z)$ the weighted Bowen topological entropy of $T$ restricted to $Z$ or, simply, the weighted Bowen topological entropy of $Z$.

### 3.2. Equivalence of $h_{\text{top}}^B$ and $h_{\text{top}}^{\text{WB}}$

The following properties about $M^s$ (cf. Section 2.2) and $\mathcal{W}^s$ can be verified directly from the definitions.

**Proposition 3.1.**

(i) For any $s \geq 0$, $N \in \mathbb{N}$ and $\epsilon > 0$, both $M_{N,\epsilon}^s$ and $\mathcal{W}_{N,\epsilon}^s$ are outer measures on $X$.

(ii) For any $s \geq 0$, both $M^s$ and $\mathcal{W}^s$ are metric outer measures on $X$.

We remark that $M^s$ and $\mathcal{W}^s$ depend on not only $s$ but also the TDS $(X,T)$. However, $M^s$ and $\mathcal{W}^s$ are purely topological and independent of the special choice of the metric $d$.

The main result of this subsection is the following.

**Proposition 3.2.** Let $Z \subseteq X$. Then for any $s \geq 0$ and $\epsilon, \delta > 0$, we have

$$M_{N,6\epsilon}^{s+\delta}(Z) \subseteq \mathcal{W}_{N,\epsilon}^s(Z) \subseteq M_{N,\epsilon}^s(Z),$$

when $N$ is large enough. As a result, $M^{s+\delta}(Z) \subseteq \mathcal{W}^s(Z) \subseteq M^s(Z)$ and $h_{\text{top}}^B(T,Z) = h_{\text{top}}^{\text{WB}}(T,Z)$.

To prove Proposition 3.2, we need the following lemma.

**Lemma 3.3.** (See [27, Theorem 2.1].) Let $(X,d)$ be a compact metric space and $B = \{B(x_i, r_i)\}_{i \in \mathcal{I}}$ be a family of closed (or open) balls in $X$. Then there exists a finite or countable subfamily $B' = \{B(x_i, r_i)\}_{i \in \mathcal{I}'}$ of pairwise disjoint balls in $B$ such that

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

**Proof of Proposition 3.2.** Let $Z \subseteq X$, $s \geq 0$, $\epsilon, \delta > 0$. Taking $f = \chi_Z$ and $c_i \equiv 1$ in the definition (3.1), we see that $\mathcal{W}_{N,\epsilon}^s(Z) \subseteq M_{N,\epsilon}^s(Z)$ for each $N \in \mathbb{N}$. In the following, we prove that $M_{N,6\epsilon}^{s+\delta}(Z) \subseteq \mathcal{W}_{N,\epsilon}^s(Z)$ when $N$ is large enough.
Assume that $N \geq 2$ is such that $n^2e^{-n\delta} \leq 1$ for $n \geq N$. Let $\{(B_{ni}(x_i, \epsilon), c_i)\}_{i \in \mathcal{I}}$ be a family so that $\mathcal{I} \subseteq \mathbb{N}, x_i \in X, 0 < c_i < \infty, n_i \geq N$ and
\[
\sum_{i} c_i \chi_{B_i} \geq \chi_Z,
\]
where $B_i := B_{ni}(x_i, \epsilon)$. We show below that
\[
\mathcal{M}_{N,6\epsilon}^{s+\delta}(Z) \leq \sum_{i \in \mathcal{I}} c_i e^{-n_i s},
\]
which implies $\mathcal{M}_{N,6\epsilon}^{s+\delta}(Z) \leq W_{s,6\epsilon}^{s}(Z)$.

Denote $\mathcal{I}_n := \{i \in \mathcal{I}: n_i = n\}$ and $\mathcal{I}_{n,k} = \{i \in \mathcal{I}_n: i \leq k\}$ for $n \geq N$ and $k \in \mathbb{N}$. Write for brevity $B_i := B_{ni}(x_i, \epsilon)$ and $5B_i := B_{ni}(x_i, 5\epsilon)$ for $i \in \mathcal{I}$. Obviously we may assume $B_i \neq B_j$ for $i \neq j$. For $t > 0$, set
\[
Z_{n,t} = \left\{ x \in Z: \sum_{i \in \mathcal{I}_n} c_i \chi_{B_i}(x) > t \right\}
\]
and
\[
Z_{n,k,t} = \left\{ x \in Z: \sum_{i \in \mathcal{I}_{n,k}} c_i \chi_{B_i}(x) > t \right\}.
\]

We divide the proof of (3.3) into the following three steps.

**Step 1.** For each $n \geq N$, $k \in \mathbb{N}$ and $t > 0$, there exists a finite set $\mathcal{J}_{n,k,t} \subseteq \mathcal{I}_{n,k}$ such that the balls $B_i$ ($i \in \mathcal{J}_{n,k,t}$) are pairwise disjoint, $Z_{n,k,t} \subseteq \bigcup_{i \in \mathcal{J}_{n,k,t}} 5B_i$ and
\[
\#(\mathcal{J}_{n,k,t}) e^{-ns} \leq \frac{1}{t} \sum_{i \in \mathcal{I}_{n,k}} c_i e^{-ns}.
\]

To prove the above result, we adopt the method of Federer [12, 2.10.24] used in the study of weighted Hausdorff measures (see also Mattila [27, Lemma 8.16]). Since $\mathcal{I}_{n,k}$ is finite, by approximating the $c_i$’s from above, we may assume that each $c_i$ is a positive rational, and then multiplying 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 $B = \{B_i: i \in \mathcal{I}_{n,k}\}$ and define $u : B \to \mathbb{Z}$ by $u(B_i) = c_i$. We define by induction integer-valued functions $v_0, v_1, \ldots, v_m$ on $B$ and subfamilies $B_1, \ldots, B_m$ of $B$ starting with $v_0 = u$. Using Lemma 3.3 (in which we take the metric $d_{n_i}$ instead of $d$) we find a pairwise disjoint subfamily $B_1$ of $B$ such that $\bigcup_{B \in B_1} B \subseteq \bigcup_{B \in B_1} 5B$, and hence $Z_{n,k,t} \subseteq \bigcup_{B \in B_1} 5B$. Then by repeatedly using Lemma 3.3, we can define inductively for $j = 1, \ldots, m$, disjoint subfamilies $B_j$ of $B$ such that
\[
B_j \subseteq \{ B \in B: v_j(B) \geq 1 \}, \quad Z_{n,k,t} \subseteq \bigcup_{B \in B_j} 5B
\]
and the functions $v_j$ such that...
This is possible since for \( j < m \), \( Z_{n,k,t} \subseteq \{ x: \sum_{B \in \mathcal{B}: B \ni x} v_j(B) \geq m - j \} \), whence every \( x \in Z_{n,k,t} \) belongs to some ball \( B \in \mathcal{B} \) with \( v_j(B) \geq 1 \). Thus

\[
\sum_{j=1}^{m} \#(\mathcal{B}_j)e^{-ns} = \sum_{j=1}^{m} \sum_{B \in \mathcal{B}_j} (v_{j-1}(B) - v_j(B))e^{-ns} \leq \sum_{B \in \mathcal{B}} \sum_{j=1}^{m} (v_{j-1}(B) - v_j(B))e^{-ns} \leq \sum_{B \in \mathcal{B}} u(B)e^{-ns} = \sum_{i \in \mathcal{I}_{n,k}} c_i e^{-ns}.
\]

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

\[
\#(\mathcal{B}_{j_0})e^{-ns} \leq \frac{1}{m} \sum_{i \in \mathcal{I}_{n,k}} c_i e^{-ns} \leq \frac{1}{t} \sum_{i \in \mathcal{I}_{n,k}} c_i e^{-ns}.
\]

Hence \( \mathcal{J}_{n,k,t} = \{ i \in \mathcal{I}: B_i \in \mathcal{B}_{j_0} \} \) is as desired.

**Step 2.** For each \( n \geq N \) and \( t > 0 \), we have

\[
\mathcal{M}_{n,6t}^{e+\delta}(Z_{n,t}) \leq \frac{1}{n^2t} \sum_{i \in \mathcal{I}_n} c_i e^{-ns}.
\]  

(3.4)

To see this, assume \( Z_{n,t} \neq \emptyset \); otherwise there is nothing to prove. Since \( Z_{n,k,t} \uparrow Z_{n,t} \), \( Z_{n,k,t} \neq \emptyset \) when \( k \) is large enough. Let \( \mathcal{J}_{n,k,t} \) be the sets constructed in Step 1. Then \( \mathcal{J}_{n,k,t} \neq \emptyset \) when \( k \) is large enough. Define \( E_{n,k,t} = \{ x_i: i \in \mathcal{J}_{n,k,t} \} \). Note that the family of all non-empty compact subsets of \( X \) is compact with respect to the Hausdorff distance (cf. Federer [12, 2.10.21]). It follows that there is a subsequence \( (k_j) \) of natural numbers and a non-empty compact set \( E_{m,t} \subseteq X \) such that \( E_{n,k_j,t} \) converges to \( E_{m,t} \) in the Hausdorff distance as \( j \to \infty \). Since any two points in \( E_{n,k,t} \) have a distance (with respect to \( d_n \)) not less than \( \epsilon \), so do the points in \( E_{m,t} \). Thus \( E_{m,t} \) is a finite set, moreover, \( \#(E_{n,k_j,t}) = \#(E_{m,t}) \) when \( j \) is large enough. Hence

\[
\bigcup_{x \in E_{n,t}} B_n(x, 5.5\epsilon) \supseteq \bigcup_{x \in E_{n,k_j,t}} B_n(x, 5\epsilon) = \bigcup_{i \in \mathcal{J}_{n,k_j,t}} 5B_i \supseteq Z_{n,k_j,t}
\]

when \( j \) is large enough, and thus \( \bigcup_{x \in E_{n,t}} B_n(x, 6\epsilon) \supseteq Z_{n,t} \). By the way, since \( \#(E_{n,k_j,t}) = \#(E_{m,t}) \) when \( j \) is large enough, we have \( \#(E_{m,t})e^{-ns} \leq \frac{1}{t} \sum_{i \in \mathcal{I}_n} c_i e^{-ns} \). This forces

\[
\mathcal{M}_{n,6t}^{e+\delta}(Z_{n,t}) \leq \#(E_{n,t})e^{-n(s+\delta)} \leq \frac{1}{e^\delta t} \sum_{i \in \mathcal{I}_n} c_i e^{-ns} \leq \frac{1}{n^2t} \sum_{i \in \mathcal{I}_n} c_i e^{-ns}.
\]

**Step 3.** For any \( t \in (0, 1) \), we have \( \mathcal{M}_{n,6t}^{e+\delta}(Z) \leq \frac{1}{t} \sum_{i \in \mathcal{I}} c_i e^{-ns} \). As a result, (3.3) holds.

To see this, fix \( t \in (0, 1) \). Note that \( \sum_{n=N}^{\infty} \frac{1}{n^2t} < 1 \). It follows that \( Z \subseteq \bigcup_{n=N}^{\infty} Z_{n,n^{-2t}} \) from (3.2). Hence by Proposition 3.1(i) and (3.4), we have
which finishes the proof of the proposition. □

3.3. A dynamical Frostman’s lemma and the proof of Theorem 1.2(i)

To prove Theorem 1.2(i), we need the following dynamical Frostman’s lemma, which is an analogue of the classical Frostman’s lemma for compact metric spaces. Our proof is adapted from Howroyd’s elegant argument (cf. [15, Theorem 2], [27, Theorem 8.17]).

Lemma 3.4. Let $K$ be a non-empty compact subset of $X$. Let $s \geq 0$, $N \in \mathbb{N}$ and $\epsilon > 0$. Suppose that $c := W_{N,\epsilon}^{s}(K) > 0$. Then there is a Borel probability measure $\mu$ on $X$ such that $\mu(K) = 1$ and

$$\mu(\mathcal{B}_{n}(x, \epsilon)) \leq \frac{1}{c}e^{-ns}, \quad \forall x \in X, n \geq N.$$ 

Proof. Clearly $c < \infty$. We define a function $p$ on the space $C(X)$ of continuous real-valued functions on $X$ by

$$p(f) = \frac{1}{c}W_{N,\epsilon}^{s}(\chi_{K} \cdot f),$$

where $W_{N,\epsilon}^{s}$ is defined as in (3.1).

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

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

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

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

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

Now we show that $\mu(K) = 1$. To see this, for any compact set $E \subseteq X \setminus K$, by the Uryson lemma there is $f \in C(X)$ such that $0 \leq f \leq 1$, $f(x) = 1$ for $x \in E$ and $f(x) = 0$ for $x \in K$. Then $f \cdot \chi_{K} \equiv 0$ and thus $p(f) = 0$. Hence $\mu(E) \leq L(f) \leq p(f) = 0$. This shows $\mu(X \setminus K) = 0$, i.e. $\mu(K) = 1$.

In the end, we show that $\mu(\mathcal{B}_{n}(x, \epsilon)) \leq (1/c)e^{-ns}$ for any $x \in X$ and $n \geq N$. To see this, for any compact set $E \subseteq \mathcal{B}_{n}(x, \epsilon)$, by the Uryson lemma, there exists $f \in C(X)$ such that $0 \leq f \leq 1$, $f(y) = 1$ for $y \in E$ and $f(y) = 0$ for $y \in X \setminus \mathcal{B}_{n}(x, \epsilon)$. Then $\mu(E) \leq L(f) \leq p(f)$. Since
\( f \cdot \chi_K \leq \chi_{B_n(x, \epsilon)} \) and \( n \geq N \), we have \( \mathcal{W}_{N, \epsilon}^N(\chi_K \cdot f) \leq e^{-ns} \) and thus \( p(f) \leq \frac{1}{c} e^{-sn} \). Therefore \( \mu(E) \leq \frac{1}{c} e^{-sn} \). It follows that

\[
\mu(B_n(x, \epsilon)) = \sup \{ \mu(E) : E \text{ is a compact subset of } B_n(x, \epsilon) \} \leq \frac{1}{c} e^{-sn}.
\]

**Remark 3.5.** There is a related known result (see, e.g. [26,30]) that, for any Borel set \( E \subset X \) and any Borel probability measure \( \mu \) on \( E \), if \( h_{\mu}(T, x) \leq s \) for all \( x \in E \), then \( h^{B_{\text{top}}}(T, E) \leq s \); conversely if \( h_{\mu}(T, x) \geq s \) for all \( x \in E \), then \( h^{B_{\text{top}}}(T, E) \geq s \), where \( h_{\mu}(T, x) \) is defined as in Section 1.

Now we are ready to prove Theorem 1.2(i).

**Proof of Theorem 1.2(i).** We first show that \( h^{B_{\text{top}}}(T, K) \geq h^\mu(T) \) for any \( \mu \in M(X) \) with \( \mu(K) = 1 \). Let \( \mu \) be a given such measure. Write

\[
h^\mu(T, x, \epsilon) = \lim \inf_{n \to \infty} -\frac{1}{n} \log \mu(B_n(x, \epsilon))
\]

for \( x \in X, n \in \mathbb{N} \) and \( \epsilon > 0 \). Clearly \( h^\mu(T, x, \epsilon) \) is nonnegative and increases as \( \epsilon \) decreases. Hence by the monotone convergence theorem and Definition 1.1,

\[
\lim_{\epsilon \to 0} \int h^\mu(T, x, \epsilon) \, d\mu = \int h^\mu(T, x) \, d\mu = h^\mu(T).
\]

Thus to show \( h^{B_{\text{top}}}(T, K) \geq h^\mu(T) \), it is sufficient to show \( h^{B_{\text{top}}}(T, K) \geq \int h^\mu(T, x, \epsilon) \, d\mu \) for each \( \epsilon > 0 \).

Fix \( \epsilon > 0 \) and \( \ell \in \mathbb{N} \). Denote \( u_\ell = \min(\ell, \int h^\mu(T, x, \epsilon) \, d\mu(x) - \frac{1}{\ell}) \). Then there exist a Borel set \( A_\ell \subset X \) with \( \mu(A_\ell) > 0 \) and \( N \in \mathbb{N} \) such that

\[
\mu(B_n(x, \epsilon)) \leq e^{-nu_\ell}, \quad \forall x \in A_\ell, n \geq N.
\]

Now let \( \{B_{n_i}(x_i, \epsilon/2)\} \) be a countable or finite family so that \( x_i \in X, n_i \geq N \) and \( \bigcup_i B_{n_i}(x_i, \epsilon/2) \supset K \cap A_\ell \). We may assume that for each \( i \), \( B_{n_i}(x_i, \epsilon/2) \cap (K \cap A_\ell) \neq \emptyset \), and choose \( y_i \in B_{n_i}(x_i, \epsilon/2) \cap (K \cap A_\ell) \). Then by (3.5),

\[
\sum_i e^{-n_iu_\ell} \geq \sum_i \mu(B_{n_i}(y_i, \epsilon)) \geq \sum_i \mu(B_{n_i}(x_i, \epsilon/2))
\]

\[
\geq \mu(K \cap A_\ell) = \mu(A_\ell) > 0.
\]

It follows that \( \mathcal{M}^{\mu}(K) \geq \mathcal{M}^{\mu}_{N, \epsilon/2}(K) \geq \mathcal{M}^{\mu}_{N, \epsilon/2}(K \cap A_\ell) \geq \mu(A_\ell) \). Therefore \( h^{B_{\text{top}}}(T, K) \geq u_\ell \). Letting \( \ell \to \infty \), we have the desired inequality \( h^{B_{\text{top}}}(T, K) \geq \int h^\mu(T, x, \epsilon) \, d\mu \). Hence \( h^{B_{\text{top}}}(T, K) \geq h^\mu(T) \).

We next show that \( h^{B_{\text{top}}}(T, K) \leq \sup \{ h^\mu(T) : \mu \in M(X), \mu(K) = 1 \} \). We can assume that \( h^{B_{\text{top}}}(T, K) > 0 \), otherwise we have nothing to prove. By Proposition 3.2, \( h^{B_{\text{top}}}(T, K) = h^{B_{\text{top}}}(T, K) \). Let \( 0 < s < h^{B_{\text{top}}}(T, K) \). Then there exist \( \epsilon > 0 \) and \( N \in \mathbb{N} \) such that \( c :=
By Proposition 3.4, there exists $\mu \in M(X)$ with $\mu(K) = 1$ such that $\mu(B_n(x, \epsilon)) \leq \frac{1}{e} e^{-sn}$ for any $x \in X$ and $n \geq N$. Clearly $h^B_{\mu}(T, x) \geq h^B_{\mu}(T, x, \epsilon) \geq s$ for each $x \in X$ and hence $h^B_{\mu}(T) \geq \int h^B_{\mu}(T, x) d\mu(x) \geq s$. This finishes the proof of Theorem 1.2(i).

### 3.4. The proof of Theorem 1.2(ii)

To prove Theorem 1.2(ii), we first prove the following.

**Theorem 3.6.** Let $(X, T)$ be a TDS. Assume that $X$ is zero-dimensional, i.e., for any $\delta > 0$, $X$ has a closed–open partition with diameter less than $\delta$. Then for any analytic set $Z \subset X$,

$$h^B_{\top}(T, Z) = \sup \{ h^B_{\top}(T, K) : K \subset Z, K \text{ is compact} \}.$$

The following proposition is needed for the proof of Theorem 3.6.

**Proposition 3.7.** Assume $\mathcal{U}$ is a closed–open partition of $X$. Let $N \in \mathbb{N}$. Then

(i) If $E_i \uparrow E$, i.e., $E_{i+1} \supseteq E_i$ and $\bigcup_i E_i = E$, then

$$\mathcal{M}^s_N(\mathcal{U}, E) = \lim_{i \to \infty} \mathcal{M}^s_N(\mathcal{U}, E_i).$$

(ii) Assume $Z \subset X$ is analytic. Then

$$\mathcal{M}^s_N(\mathcal{U}, Z) = \sup \{ \mathcal{M}^s_N(\mathcal{U}, K) : K \subset Z, K \text{ is compact} \}.$$

**Proof.** We first show that (i) implies (ii). Assume that (i) holds. Let $Z$ be analytic, i.e., there exists a continuous surjective map $\phi : N \to Z$. Let $\Gamma_{n_1,n_2,\ldots,n_p}$ be the set of $(m_1, m_2, \ldots) \in N$ such that $m_1 \leq n_1, m_2 \leq n_2, \ldots, m_p \leq n_p$ and let $Z_{n_1,\ldots,n_p}$ be the image of $\Gamma_{n_1,\ldots,n_p}$ under $\phi$. Let $(\epsilon_p)$ be a sequence of positive numbers. Due to (i), we can pick a sequence $(n_p)$ of positive integers recursively so that $\mathcal{M}^s_N(\mathcal{U}, Z_{n_1,\ldots,n_p}) \geq \mathcal{M}^s_N(\mathcal{U}, Z_{n_1,\ldots,n_p-1}) - \epsilon_p$, $p = 2, 3, \ldots$.

Hence $\mathcal{M}^s_N(\mathcal{U}, Z_{n_1,\ldots,n_p}) \geq \mathcal{M}^s_N(\mathcal{U}, Z_{n_1,\ldots,n_p-1}) - \sum_{i=1}^\infty \epsilon_i$ for any $p \in \mathbb{N}$. Let

$$K = \bigcap_{p=1}^\infty Z_{n_1,\ldots,n_p}.$$

Since $\phi$ is continuous, we can show that $\bigcap_{p=1}^\infty Z_{n_1,\ldots,n_p} = \bigcap_{p=1}^\infty Z_{n_1,\ldots,n_p}$ by applying Cantor’s diagonal argument. Hence $K$ is a compact subset of $Z$. If $A \subset \bigcup_{j \geq N} \mathcal{W}_j(\mathcal{U})$ is a cover of $K$ (of course it is an open cover), then it is a cover of $Z_{n_1,\ldots,n_p}$ when $p$ is large enough, which implies

$$\sum_{U \in A} e^{-s\mathcal{M}(U)} \geq \lim_{p \to \infty} \mathcal{M}^s_N(\mathcal{U}, Z_{n_1,\ldots,n_p}) \geq \mathcal{M}^s_N(\mathcal{U}, Z) - \sum_{i=1}^\infty \epsilon_i.$$
Hence $\mathcal{M}_N^r(U, K) \geq \mathcal{M}_N^r(U, Z) - \sum_{i=1}^{\infty} \epsilon_i$. Since $\sum_{i=1}^{\infty} \epsilon_i$ can be chosen arbitrarily small, we have proven (ii).

Now we turn to prove (i). Our argument is modified from the classical proof of the “increasing sets lemma” for Hausdorff outer measures (cf. [7, Sect. II] and [10, Lemma 5.3]). Note that any two non-empty elements in $\mathcal{W}_n(U)$ are disjoint, and each element in $\mathcal{W}_{n+1}(U)$ is a subset of some element in $\mathcal{W}_n(U)$. We call this the net property of $(\mathcal{W}_n(U))$.

Let $E_i \uparrow E$ be given. Let $(\delta_i)$ be a sequence of positive numbers to be specified later and for each $i$, choose a covering $\Lambda_i \subset \bigcup_{j \geq n} \mathcal{W}_j(U)$ of $E_i$ such that

$$\sum_{U \in A_i} e^{-sm(U)} \leq \mathcal{M}_N^r(U, E_i) + \delta_i. \quad (3.6)$$

By the net property of $(\mathcal{W}_n(U))$, we may assume that for each $i$, the elements in $A_i$ are disjoint.

For any $x \in E$, choose $U_x \in \bigcup_{i=1}^{\infty} A_i$ containing $x$ such that $m(U_x)$ is the smallest. By the net property of $(\mathcal{W}_n(U))$, the collection $\{U_x : x \in E\}$ consists of countable many disjoint elements. Relabel these elements as $U'_i$. Clearly $E \subset \bigcup_i U'_i$.

We now choose an integer $k$. Use $A_1$ to denote the collection of those $U'_i$’s that are taken from $A_1$. They cover a certain subset $Q_1$ of $E_k$. The same subset is covered by a certain sub-collection of $A_k$, denoted as $A_{k,1}$. Since $A_{k,1}$ also covers the smaller set $Q_1 \cap E_1$, by (3.6),

$$\sum_{U \in A_1} e^{-sm(U)} \leq \sum_{U \in A_{k,1}} e^{-sm(U)} + \delta_1. \quad (3.7)$$

To see this, assume that (3.7) is false. Then by (3.6),

$$\sum_{U \in (A_1 \setminus A_1) \cup A_{k,1}} e^{-sm(U)} < \mathcal{M}_N^r(U, E_1),$$

which contradicts the fact that $(A_1 \setminus A_1) \cup A_{k,1} \subset \bigcup_{j \geq n} \mathcal{W}_j(U)$ is an open cover of $E_1$. Next we use $A_2$ to denote the collection of those $U'_i$’s that are taken from $A_2$ but not from $A_1$. Define $A_{k,2}$ similarly. As above, we find

$$\sum_{U \in A_2} e^{-sm(U)} \leq \sum_{U \in A_{k,2}} e^{-sm(U)} + \delta_2. \quad (3.8)$$

We repeat the argument until all coverings $A_n$, $n \leq k$, have been considered. Note that $\bigcup_{i \leq k} U \subset \bigcup_{i \in I_k} U$ for $i \leq k$. For different $i, i' \leq k$, the elements in $A_{k,i}$ are disjoint from those in $A_{k,i'}$. The $k$ inequalities (3.7), (3.8), . . . , are added which yield

$$\sum_{U \in \bigcup_{n=1}^{k} A_n} e^{-sm(U)} \leq \sum_{U \in \bigcup_{n=1}^{k} A_{k,n}} e^{-sm(U)} + k \sum_{n=1}^{k} \delta_n + \delta_k.$$

Letting $k \to \infty$, we have

$$\sum_i e^{-sm(U_i)} \leq \lim_{k \to \infty} \mathcal{M}_N^r(U, E_k) + \sum_{n=1}^{\infty} \delta_n.$$
Since $\sum_{n=1}^{\infty} \delta_n$ can be chosen arbitrarily small we have
\[
\mathcal{M}^s_N(\mathcal{U}, E) \leq \lim_{k \to \infty} \mathcal{M}^s_N(\mathcal{U}, E_k).
\]

Since the opposite inequality is trivial we have proved (i). \qed

**Proof of Theorem 3.6.** Let $Z$ be an analytic subset of $X$ with $h^B_{\text{top}}(T, Z) > 0$. Let $0 < s < h^B_{\text{top}}(T, Z)$. By (2.2) and (2.3), there exists a closed–open partition $\mathcal{U}$ such that $h^B_{\text{top}}(\mathcal{U}, Z) > s$, and thus $\mathcal{M}^s(\mathcal{U}, Z) = \infty$. Hence $\mathcal{M}^s_N(\mathcal{U}, Z) > 0$ for some $N \in \mathbb{N}$. By Proposition 3.7, we can find a compact set $K \subset Z$ such that $\mathcal{M}^s_N(\mathcal{U}, K) > 0$. This implies $h^B_{\text{top}}(T, K) \geq h^B_{\text{top}}(T, \mathcal{U}, K) \geq s$. \qed

Before we prove Theorem 1.2(ii), we still need some notation and additional results in topological dynamical systems.

Let us define the natural extension $(\hat{X}, \hat{T})$ of a TDS $(X, T)$ with a metric $d$ and a surjective map $T$ where $\hat{X} = \{(x_1, x_2, \ldots): T(x_{i+1}) = x_i, x_i \in X, i \in \mathbb{N}\}$ is a subspace of the product space $X^\mathbb{N} = \prod_{i=1}^{\infty} X$ endowed with the compatible metric $d_T$ as
\[
d_T((x_1, x_2, \ldots), (y_1, y_2, \ldots)) = \sum_{i=1}^{\infty} \frac{d(x_i, y_i)}{2^i},
\]

$\tilde{T} : \hat{X} \to \hat{X}$ is the shift homeomorphism with $\tilde{T}(x_1, x_2, \ldots) = (T(x_1), x_1, x_2, \ldots)$, and $\pi_i : \hat{X} \to X$ is the projection to the $i$-th coordinate. Clearly, $\pi_i : (\hat{X}, \tilde{T}) \to (X, T)$ is a factor map.

**Lemma 3.8.** Let $(X, T)$ be a TDS with a metric $d$ and a surjective map $T$, $(\hat{X}, \tilde{T})$ be the natural extension of $(X, T)$ and $\pi_1 : \hat{X} \to X$ be the projection to the first coordinate. Then $\sup_{x \in \hat{X}} h^1_{\text{top}}(\tilde{T}, \pi_1^{-1}(x)) = 0$.

**Proof.** Fix $x \in X$. For any $\epsilon > 0$, take $N \in \mathbb{N}$ large enough such that $\sum_{i=1}^{\infty} \frac{\text{diam}(X)}{2^i} < \epsilon$.

Let $E_N \subseteq \pi_1^{-1}(x)$ be a finite $(N, \epsilon)$-spanning set of $\pi_1^{-1}(x)$. Next we are to show that $E_N$ is also an $(n, \epsilon)$-spanning set of $\pi_1^{-1}(x)$ for $n > N$.

Fix $n \in \mathbb{N}$ with $n > N$. For any $\hat{y} \in \pi_1^{-1}(x)$, since $E_N$ is an $(N, \epsilon)$-spanning set of $\pi_1^{-1}(x)$ there exists $\hat{x} \in E_N$ such that $d_T(\tilde{T}^i\hat{x}, \tilde{T}^i\hat{y}) < \epsilon$ for $i = 0, 1, \ldots, N - 1$. Now for $k \in \{N, N + 1, \ldots, n - 1\}$, we have $\pi_j(\tilde{T}^k\hat{x}) = \pi_j(\tilde{T}^k\hat{y}) = T^{k-j+1}(x)$ for $j = 1, \ldots, k, k + 1$. Thus
\[
d_T(\tilde{T}^k\hat{x}, \tilde{T}^k\hat{y}) = \sum_{j=1}^{\infty} \frac{d(\pi_j(\tilde{T}^k\hat{x}), \pi_j(\tilde{T}^k\hat{y}))}{2^j} = \sum_{j=k+2}^{\infty} \frac{d(\pi_j(\tilde{T}^k\hat{x}), \pi_j(\tilde{T}^k\hat{y}))}{2^j} \leq \sum_{j=k+2}^{\infty} \frac{\text{diam}(X)}{2^j} < \epsilon.
\]

This implies $(d_T)_n(\hat{x}, \hat{y}) < \epsilon$. Hence $E_N$ is also an $(n, \epsilon)$-spanning set of $\pi_1^{-1}(x)$ for $n > N$. Let $\tilde{r}_n(\pi^{-1}(x), \epsilon)$ denote the smallest cardinality of $(n, \epsilon)$-spanning sets of $\pi^{-1}(x)$. Then
\[ \tilde{r}_n(\pi^{-1}(x), \epsilon) \leq \#(E_N). \]

Hence
\[ h_{U^{C}}^{\text{top}}(\tilde{T}, \pi^{-1}(x)) = \lim_{\epsilon \to 0} \limsup_{n \to \infty} \frac{1}{n} \log \tilde{r}_n(\pi^{-1}(x), \epsilon) \leq \lim_{\epsilon \to 0} \limsup_{n \to \infty} \frac{1}{n} \log \#(E_N) = 0. \]

This ends the proof of the lemma. \( \square \)

In the following part we will lift general TDSs having finite topological entropy to zero-dimensional TDSs by the so-called principal extensions.

**Definition 3.9.** (See [22].) An extension \( \pi : (Z, R) \to (X, T) \) between two TDSs is a principal extension if \( h_{\nu}(R) = h_{\nu \circ \pi^{-1}}(T) \) for every \( \nu \in M(Z, R) \).

The following general result is needed in our proof of Theorem 1.2(ii).

**Proposition 3.10.** (See Proposition 7.8 in [5].) Every invertible TDS \((X, T)\) with \( h_{\text{top}}(T) < \infty \) has a zero-dimensional principal extension \((Z, R)\) with \( R \) being invertible.

Let \( \pi : (Y, S) \to (X, T) \) be a factor map between two TDSs. Bowen proved that \( h_{\text{top}}(S) \leq h_{\text{top}}(T) + \sup_{x \in X} h_{U^{C}}^{\text{top}}(S, \pi^{-1}(x)) \) (cf. [3, Theorem 17]). In fact, Bowen’s proof is also valid for the following result (see, e.g. Theorem 7.3 in [16] for a detailed proof).

**Theorem 3.11.** Let \( \pi : (Y, S) \to (X, T) \) be a factor map between two TDSs. Then for any \( E \subseteq Y \) one has
\[
0 \leq h_{\text{top}}^{B}(T, \pi(E)) \leq h_{\text{top}}^{B}(S, E) \leq h_{\text{top}}^{B}(T, \pi(E)) + \sup_{x \in X} h_{U^{C}}^{\text{top}}(S, \pi^{-1}(x)). \tag{3.9}
\]

We also need the following variational principle of conditional entropies.

**Proposition 3.12.** Let \( \pi : (Y, S) \to (X, T) \) be a factor map between two TDSs with \( h_{\text{top}}(S) < \infty \). Then we have
\[
\sup_{x \in X} h_{U^{C}}^{\text{top}}(S, \pi^{-1}(x)) = \sup_{\mu \in M(Y, S)} \left( h_\mu(S) - h_{\mu \circ \pi^{-1}}(T) \right). \tag{3.10}
\]

**Proof.** It is the direct combination of [9, Theorem 3] and [23, Theorem 2.1]. \( \square \)

**Lemma 3.13.** Let \((X, T)\) be a TDS with \( h_{\text{top}}(T) < \infty \). Then there exists a factor map \( \pi : (H, \Gamma) \to (X, T) \) such that \((H, \Gamma)\) is zero-dimensional and
\[
\sup_{x \in X} h_{U^{C}}^{\text{top}}(\Gamma, \pi^{-1}(x)) = 0.
\]

**Proof.** First, we take \( D = \{(1/n)_{n \in \mathbb{N}} \} \cup \{0\} \) and let \( Z = X \times D \). Define \( R : Z \to Z \) satisfying \( R(x, 1/n+1) = (x, 1/n) \) for \( n \in \mathbb{N} \); \( R(x, 1) = (Tx, 1) \) and \( R(x, 0) = (x, 0) \) for \( x \in X \). Then \((Z, R)\) is a TDS and \( R \) is surjective. If we identity \((x, 1)\) with \( x \) for each \( x \in X \), then \( X \) can be viewed as a closed subset of \( Z \) and \( R|_X = T \). It is also clear that \( h_{\text{top}}(R) = h_{\text{top}}(T) < \infty \).
Let \((\tilde{Z}, \tilde{R})\) be the natural extension of \((Z, R)\) and \(\pi_1 : \tilde{Z} \to Z\) be the projection to the first coordinate. Then
\[
\sup_{z \in Z} h_{\text{top}}^\text{UC} (\tilde{R}, \pi_1^{-1}(z)) = 0 \tag{3.11}
\]
by Lemma 3.8, and, so \(h_{\text{top}}(\tilde{R}) = h_{\text{top}}(R) < \infty\). Since \(\tilde{R}\) is a homeomorphism on \(\tilde{Z}\), by Proposition 3.10, there exists a factor map \(\psi : (W, G) \to (\tilde{Z}, \tilde{R})\) such that \((W, G)\) is a zero-dimensional TDS and \(\psi\) is principal extension.

Since \(h_{\text{top}}(\tilde{R}) < \infty\) and \(\psi\) is principal extension, we have the following variational principle of condition entropy
\[
\sup_{z \in \tilde{Z}} h_{\text{top}}^\text{UC} (G, \psi^{-1}(z)) = \sup_{\theta \in M(W, G)} \left( h_\theta(G) - h_{\theta \circ \psi^{-1}}(\tilde{R}) \right) = 0 \tag{3.12}
\]
The first equality in (3.12) follows from (3.10).

Let \(H = \psi^{-1}(\pi_1^{-1}X)\), \(\Gamma = G\vert_H\) and \(\pi = \pi_1 \circ \psi\vert_H\). Then \((H, \Gamma)\) is a zero-dimensional TDS and \(\pi : (H, \Gamma) \to (X, T)\) is a factor map. Applying Proposition 3.12 to the factor map \(\pi : (H, \Gamma) \to (X, T)\), we obtain
\[
\sup_{x \in X} h_{\text{top}}^\text{UC} (\Gamma, \pi^{-1}(x)) = \sup_{\mu \in M(H, \Gamma)} \left( h_\mu(\Gamma) - h_{\mu \circ \pi^{-1}}(T) \right) \leq \sup_{\mu \in M(W, G)} \left( h_\mu(G) - h_{\mu \circ \psi^{-1}}(T) \right) + \sup_{\nu \in (\tilde{Z}, \tilde{R})} \left( h_\nu(\tilde{R}) - h_{\nu \circ \pi_1^{-1}}(R) \right)
\]
\[
= \sup_{z \in Z} h_{\text{top}}^\text{UC} (G, \psi^{-1}(z)) + \sup_{z \in Z} h_{\text{top}}^\text{UC} (\tilde{R}, \pi_1^{-1}(z))
\]
\[
= 0 \quad \text{(by (3.12), (3.11)).}
\]

This shows \(\sup_{x \in X} h_{\text{top}}^\text{UC} (\Gamma, \pi^{-1}(x)) = 0\). \(\square\)

**Proof of Theorem 1.2(ii).** By Lemma 3.13, there exists a factor map \(\pi : (Y, S) \to (X, T)\) such that \((Y, S)\) is zero-dimensional and \(\sup_{x \in X} h_{\text{top}}^\text{UC} (S, \pi^{-1}(x)) = 0\). By Theorem 3.11, we have that for any \(F \subset Y\),
\[
h_{\text{top}}^B (S, F) = h_{\text{top}}^B (T, \pi(F)). \tag{3.13}
\]

Let \(Z\) be an analytic subset of \(X\). Then \(\pi^{-1}(Z)\) is also an analytic subset of \(Y\) (cf. Federer [12, 2.2.10]). By (3.13) and Theorem 3.6,
\[
h_{\text{top}}^B (T, Z) = h_{\text{top}}^B (S, \pi^{-1}(Z)) = \sup \{ h_{\text{top}}^B (S, E) : E \subseteq \pi^{-1}(Z), \ E \text{ is compact} \}
\]
\[
= \sup \{ h_{\text{top}}^B (T, \pi(E)) : E \subseteq \pi^{-1}(Z), \ E \text{ is compact} \}
\]
\[
\leq \sup \{ h_{\text{top}}^B (T, K) : K \subseteq Z, \ K \text{ is compact} \}.
\]
The reverse inequality is trivial, so
\[ h^B_{\text{top}}(T, Z) = \sup \{ h^B_{\text{top}}(T, K) : K \subseteq Z, K \text{ is compact} \}. \]

This finishes the proof. \qed

**Remark 3.14.** For an invertible TDS \((X, T)\), Lindenstrauss and Weiss [25] introduced the mean dimension \(mdim(X, T)\) (an idea suggested by Gromov). It is well known that for an invertible TDS \((X, T)\), if \(h^\text{top}(T) < \infty\) or the topological dimension of \(X\) is finite, then \(mdim(X, T) = 0\) (see [25, Definition 2.6 and Theorem 4.2]).

In general, one can show that for an invertible TDS \((X, T)\), if \(mdim(X, T) = 0\) then \((X, T)\) has a zero-dimensional principal extension \((Z, R)\) with \(R\) being invertible. Indeed, let \((Y, S)\) be an irrational rotation on the circle. Then \((X \times Y, T \times S)\) admits a nonperiodic minimal factor \((Y, S)\) and \(mdim(X \times Y, T \times S) = 0\). Hence \((X \times Y, T \times S)\) has the so-called small boundary property [24, Theorem 6.2], which implies the existence of a basis of the topology consisting of sets whose boundaries have measure zero for every invariant measure. With these results it is easy to construct a refining sequence of small-boundary partitions for \((X \times Y, T \times S)\), where the partitions have small boundaries if their boundaries have measure zero for all \(\mu \in \mathcal{M}(X \times Y, T \times S)\). Then by a standard construction (see pp. 152–153 in [5]), we associate to this sequence a zero-dimensional principal extension \((Z, R)\) of \((X \times Y, T \times S)\) with \(R\) invertible. Finally note that since \((X \times Y, T \times S)\) is a principal extension of \((X, T)\), we know that \((Z, R)\) is also a zero-dimensional principal extension of \((X, T)\) since the composition of two principal extensions is still a principal extension.

**Remark 3.15.** By Remark 3.14, we may strengthen Theorem 1.2(ii) as follows: Let \((X, T)\) be a TDS with \(mdim(X, T) = 0\). Then for any analytic set \(Z \subseteq X\),
\[ h^B_{\text{top}}(T, Z) = \sup \{ h^B_{\text{top}}(T, K) : K \subseteq Z, K \text{ is compact} \}. \]

4. **Variational principle for the packing topological entropy**

In this section we prove Theorem 1.3. We first give a lemma.

**Lemma 4.1.** Let \(Z \subset X\) and \(s, \epsilon > 0\). Assume \(P_s^\epsilon(Z) = \infty\). Then for any given finite interval \((a, b) \subset \mathbb{R}\) with \(a \geq 0\) and any \(N \in \mathbb{N}\), there exists a finite disjoint collection \(\{B_{n_i}(x_i, \epsilon)\}\) such that \(x_i \in Z, n_i \geq N\) and \(\sum_i e^{-n_i s} \in (a, b)\).

**Proof.** Take \(N_1 > N\) large enough such that \(e^{-N_1 s} < b - a\). Since \(P_s^\epsilon(Z) = \infty\), we have \(P_{N_1, \epsilon}^s(Z) = \infty\). Thus there is a finite disjoint collection \(\{B_{n_i}(x_i, \epsilon)\}\) such that \(x_i \in Z, n_i \geq N_1\) and \(\sum_i e^{-n_i s} > b\). Since \(e^{-n_i s} < b - a\), we can discard elements in this collection one by one until we can have \(\sum_i e^{-n_i s} \in (a, b)\). \qed

**Proof of Theorem 1.3.** We divide the proof into two parts:

**Part 1.** \(h^P_{\text{top}}(T, Z) \geq \sup \{ h_\mu(T) : \mu \in \mathcal{M}(X), \mu(Z) = 1 \}\) for any Borel set \(Z \subseteq X\).

To see this, let \(\mu \in \mathcal{M}(X)\) with \(\mu(Z) = 1\) for some Borel set \(Z \subseteq X\). We need to show that \(h^P_{\text{top}}(T, Z) \geq h_\mu(T)\). For this purpose we may assume \(h_\mu(T) > 0\); otherwise we have nothing to
prove. Let $0 < s < \tilde{h}_\mu(T)$. Then there exist $\epsilon, \delta > 0$, and a Borel set $A \subset Z$ with $\mu(A) > 0$ such that

$$\tilde{h}_\mu(T, x, \epsilon) > s + \delta, \quad \forall x \in A,$$

where $\tilde{h}_\mu(T, x, \epsilon) := \limsup_{n \to \infty} -\frac{1}{n} \log \mu(B_n(x, \epsilon))$.

Next we show that $P_{\epsilon/5}(Z) = \infty$, which implies that $h_{\text{top}}(T, Z) \geq h_{\text{top}}(T, Z, \epsilon/5) \geq s$. To achieve this, it suffices to show that $P_{\epsilon/5}(E) = \infty$ for any Borel $E \subset A$ with $\mu(E) > 0$. Fix such a set $E$. Define

$$E_n = \{x \in E : \mu(B_n(x, \epsilon)) < e^{-n(s + \delta)}\}, \quad n \in \mathbb{N}.$$

Since $E \subset A$, we have $\bigcup_{n=N}^\infty E_n = E$ for each $N \in \mathbb{N}$. Fix $N \in \mathbb{N}$. Then $\mu(\bigcup_{n=N}^\infty E_n) = \mu(E)$, and hence there exists $n \geq N$ such that

$$\mu(E_n) \geq \frac{1}{n(n+1)} \mu(E).$$

Fix such $n$ and consider the family $\{B_n(x, \epsilon/5) : x \in E_n\}$. By Lemma 3.3 (in which we use $d_n$ instead of $d$), there exists a finite pairwise disjoint family $\{B_n(x_i, \epsilon/5)\}$ with $x_i \in E_n$ such that

$$\bigcup_i B_n(x_i, \epsilon) \supset \bigcup_{x \in E_n} B_n(x, \epsilon/5) \supset E_n.$$

Hence

$$P_{N, \epsilon/5}(E) \geq P_{N, \epsilon/5}(E_n) \geq \sum_i e^{-n\delta} \sum_i e^{-n(s + \delta)} \geq e^{n\delta} \sum_i \mu(B_n(x_i, \epsilon)) \geq e^{n\delta} \mu(E_n) \geq \frac{e^{n\delta}}{n(n+1)} \mu(E).$$

Since $\frac{e^{n\delta}}{n(n+1)} \to \infty$ as $n \to \infty$, letting $N \to \infty$ we obtain that $P_{\epsilon/5}(E) = \infty$.

**Part 2.** Let $Z \subset X$ be analytic with $h_{\text{top}}^P(T, Z) > 0$. For any $0 < s < h_{\text{top}}^P(T, Z)$, there exist a compact set $K \subset Z$ and $\mu \in M(K)$ such that $h_\mu(T) \geq s$.

Since $Z$ is analytic, there exists a continuous surjective map $\phi : \mathcal{N} \to Z$. Let $\Gamma_{n_1, n_2, \ldots, n_p}$ be the set of $(m_1, m_2, \ldots) \in \mathcal{N}$ such that $m_1 \leq n_1$, $m_2 \leq n_2$, $\ldots$, $m_p \leq n_p$ and let $Z_{n_1, \ldots, n_p}$ be the image of $\Gamma_{n_1, \ldots, n_p}$ under $\phi$.

Take $\epsilon > 0$ small enough so that $0 < s < h_{\text{top}}^P(T, Z, \epsilon)$. We are going to construct inductively a sequence of finite sets $(K_i)_{i=1}^{\infty}$ and a sequence of finite measures $(\mu_i)_{i=1}^{\infty}$ so that $K_i \subset Z$ and $\mu_i$ is supported on $K_i$ for each $i$. Together with these two sequences, we construct also a sequence of integers $(n_i)$, a sequence of positive numbers $(\gamma_i)$ and a sequence of integer-valued functions $(m_i : K_i \to \mathbb{N})$. The method of our construction is inspired to some extent by the work of Joyce and Preiss [18] on packing measures.
Here and afterwards, \( \mu \) denotes the Dirac measure at \( x \). Take a small \( \gamma_1 > 0 \) such that for any function \( z : K_1 \to X \) with \( d(x, z(x)) \leq \gamma_1 \), we have for each \( x \in K_1 \),

\[
(B(z(x), \gamma_1) \cup B_{m_1(x)}(z(x), \epsilon)) \cap \left( \bigcup_{y \in K_1 \setminus \{x\}} B(z(y), \gamma_1) \cup B_{m_1(y)}(z(y), \epsilon) \right) = \emptyset.
\]

Here and afterwards, \( B(x, \epsilon) \) denotes the closed ball \( \{ y \in X : d(x, y) \leq \epsilon \} \). Since \( K_1 \subset Z' \), \( P_{\epsilon}^l(Z \cap B(x, \gamma_1/4)) > P_{\epsilon}^l(Z' \cap B(x, \gamma_1/4)) > 0 \) for each \( x \in K_1 \). Therefore we can pick a large \( n_1 \in \mathbb{N} \) so that \( Z_{n_1} \supset K_1 \) and \( P_{\epsilon}^l(Z_{n_1} \cap B(x, \gamma_1/4)) > 0 \) for each \( x \in K_1 \).

Step 2. Construct \( K_2 \) and \( \mu_2 \), as well as \( m_2(\cdot) \), \( n_2 \) and \( \gamma_2 \).

By (4.1), the family of balls \( \{B(x, \gamma_1/4)\}_{x \in K_1} \) are pairwise disjoint. For each \( x \in K_1 \), since \( P_{\epsilon}^l(Z_{n_1} \cap B(x, \gamma_1/4)) > 0 \), we can construct as in Step 1, a finite set

\[
E_2(x) \subset Z_{n_1} \cap B(x, \gamma_1/4)
\]

and an integer-valued function

\[
m_2 : E_2(x) \to \mathbb{N} \cap \left[ \max\{m_1(y) : y \in K_1\}, \infty \right)
\]

such that

1. \( P_{\epsilon}^l(Z_{n_1} \cap G) > 0 \) for each open set \( G \) with \( G \cap E_2(x) \neq \emptyset \);
2. The elements in \( \{B_{m_2(y)}(y, \epsilon)\}_{y \in E_2(x)} \) are disjoint, and

\[
\mu_1([x]) < \sum_{y \in E_2(x)} e^{-m_2(y)x} < (1 + 2^{-2})\mu_1([x]).
\]
To see it, we fix \( x \in K \). Denote \( F = Z_{n_1} \cap B(x, \gamma_1/4) \). Let

\[
H_x := \bigcup \{ G \subset X : G \text{ is open } \mathcal{P}_t^e(F \cap G) = 0 \}.
\]

Set \( F' = F \setminus H_x \). Then as in Step 1, we can show that \( \mathcal{P}_t^e(F') = \mathcal{P}_t^e(F) > 0 \) and furthermore, \( \mathcal{P}_t^e(F' \cap G) > 0 \) for any open set \( G \) with \( G \cap F' \neq \emptyset \). Note that \( \mathcal{P}_t^e(F') = \infty \) (since \( s < t \)), as so by Lemma 4.1, we can find a finite set \( E_2(x) \subset F' \) and a map \( m_2: E_2(x) \to \mathbb{N} \cap \{ \max \{ m_1(y) : y \in K \}, \infty \} \) so that (2-b) holds. Observe that if an open set \( G \) satisfies \( G \cap E_2(x) \neq \emptyset \), then \( G \cap F' \neq \emptyset \), and hence \( \mathcal{P}_t^e(Z_{n_1} \cap G) > 0 \). Thus (2-a) holds.

Since the family \( \{ B(x, \gamma_1) \} \) is disjoint, \( E_2(x) \cap E_2(x') = \emptyset \) for different \( x, x' \in K \). Define \( K_2 = \bigcup_{x \in K_1} E_2(x) \) and \( \mu_2 = \sum_{y \in K_2} e^{-m_2(y)x} \delta_y \).

By (4.1) and (2-b), the elements in \( \{ B_{m_2(y)}(y, \epsilon) \} \) are pairwise disjoint. Hence we can take \( 0 < \gamma_2 < \gamma_1/4 \) such that for any function \( z : K_2 \to X \) with \( d(x, z(x)) < \gamma_2 \) for \( x \in K_2 \), we have

\[
(\overline{B}(z(x), \gamma_2) \cup \overline{B}_{m_2(x)}(z(x), \epsilon)) \cap \left( \bigcup_{y \in K_2 \setminus \{ x \}} \overline{B}(z(y), \gamma_2) \cup \overline{B}_{m_2(y)}(z(y), \epsilon) \right) = \emptyset \quad (4.2)
\]

for each \( x \in K_2 \). Choose a large \( n_2 \in \mathbb{N} \) such that \( Z_{n_1}, \ldots, n_2 \supset K_2 \) and \( \mathcal{P}_t^e(Z_{n_1}, n_2 \cap B(x, \gamma_2/4)) > 0 \) for each \( x \in K_2 \).

**Step 3.** Assume that \( K_i, \mu_i, m_i(\cdot), n_i \) and \( \gamma_i \) have been constructed for \( i = 1, \ldots, p \). In particular, assume that for any function \( z : K_p \to X \) with \( d(x, z(x)) < \gamma_p \) for \( x \in K_p \), we have

\[
(\overline{B}(z(x), \gamma_p) \cup \overline{B}_{m_p(x)}(z(x), \epsilon)) \cap \left( \bigcup_{y \in K_p \setminus \{ x \}} \overline{B}(z(y), \gamma_p) \cup \overline{B}_{m_p(y)}(z(y), \epsilon) \right) = \emptyset \quad (4.3)
\]

for each \( x \in K_p \); and \( Z_{n_1}, \ldots, n_p \supset K_p \) and \( \mathcal{P}_t^e(Z_{n_1}, \ldots, n_p \cap B(x, \gamma_p/4)) > 0 \) for each \( x \in K_p \). We construct below each term of them for \( i = p + 1 \) in a way similar to Step 2.

Note that the elements in \( \{ \overline{B}(x, \gamma_p) \} \) are pairwise disjoint. For each \( x \in K_p \), since \( \mathcal{P}_t^e(Z_{n_1}, \ldots, n_p \cap B(x, \gamma_p/4)) > 0 \), we can construct as in Step 2, a finite set

\[
E_{p+1}(x) \subset Z_{n_1}, \ldots, n_p \cap B(x, \gamma_p/4)
\]

and an integer-valued function

\[
m_{p+1}: E_{p+1}(x) \to \mathbb{N} \cap \{ \max \{ m_p(y) : y \in K_p \}, \infty \}
\]

such that
(3-a) \( P_{\epsilon}^t(Z_{n_1,...,n_p} \cap G) > 0 \) for each open set \( G \) with \( G \cap E_{p+1}(x) \neq \emptyset \); and

(3-b) \( \{ \overline{B}_{m_{p+1}}(y, \epsilon) \}_{y \in E_{p+1}(x)} \) are disjoint and satisfy

\[
\mu_p([x]) < \sum_{y \in E_{p+1}(x)} e^{-m_{p+1}(y)s} < (1 + 2^{-p-1}) \mu_p([x]).
\]

Clearly \( E_{p+1}(x) \cap E_{p+1}(x') = \emptyset \) for different \( x, x' \in K_p \). Define \( K_{p+1} = \bigcup_{x \in K_p} E_{p+1}(x) \) and

\[
\mu_{p+1} = \sum_{y \in K_{p+1}} e^{-m_{p+1}(y)s} \delta_y.
\]

By (4.3) and (3-b), \( \{ B_{m_{p+1}}(y, \epsilon) \}_{y \in K_{p+1}} \) are disjoint. Hence we can take \( 0 < \gamma_{p+1} < \gamma_p/4 \) such that for any function \( z : K_{p+1} \to X \) with \( d(x, z(x)) < \gamma_{p+1} \), we have for each \( x \in K_{p+1} \),

\[
(\overline{B}(z(x), \gamma_{p+1}) \cup \overline{B}_{m_{p+1}}(z(x), \epsilon)) \cap \left( \bigcup_{y \in K_{p+1} \setminus \{x\}} \overline{B}(z(y), \gamma_{p+1}) \cup \overline{B}_{m_{p+1}}(z(y), \epsilon) \right) = \emptyset. \tag{4.4}
\]

Choose a large \( n_{p+1} \in \mathbb{N} \) such that \( Z_{n_1,...,n_{p+1}} \supset K_{p+1} \) and

\[
P_{\epsilon}^t(Z_{n_1,...,n_{p+1}} \cap B(x, \gamma_{p+1}/4)) > 0
\]

for each \( x \in K_{p+1} \).

As in the above steps, we can construct by induction the sequences \( (K_i), (\mu_i), (m_i(\cdot)), (n_i) \) and \( (\gamma_i) \). We summarize some of their basic properties as follows:

(a) For each \( i \), the family \( \mathcal{F}_i := \{ \overline{B}(x, \gamma_i) : x \in K_i \} \) is disjoint. Each element in \( \mathcal{F}_{i+1} \) is a subset of \( \overline{B}(x, \gamma_i/2) \) for some \( x \in K_i \).

(b) For each \( x \in K_i \) and \( z \in \overline{B}(x, \gamma_i) \),

\[
\overline{B}_{m_i(x)}(z, \epsilon) \cap \bigcup_{y \in K_i \setminus \{x\}} \overline{B}(y, \gamma_i) = \emptyset \quad \text{and} \quad \mu_i(\overline{B}(x, \gamma_i)) = e^{-m_i(x)s} \leq \sum_{y \in E_{i+1}(x)} e^{-m_{i+1}(y)s} \leq (1 + 2^{-i-1}) \mu_i(\overline{B}(x, \gamma_i)),
\]

where \( E_{i+1}(x) = B(x, \gamma_i) \cap K_{i+1} \).

The second part in (b) implies,

\[
\mu_i(F_i) \leq \mu_{i+1}(F_i) = \sum_{F \in \mathcal{F}_{i+1}, F \subseteq F_i} \mu_{i+1}(F) \leq (1 + 2^{-i-1}) \mu_i(F_i), \quad F_i \in \mathcal{F}_i.
\]
Using the above inequalities repeatedly, we have for any $j > i$,
\[ \mu_i(F_i) \leq \mu_j(F_i) \leq \prod_{n = i+1}^{j} (1 + 2^{-n}) \mu_i(F_i) \leq C \mu_i(F_i), \quad \forall F_i \in \mathcal{F}_i, \quad (4.5) \]
where $C := \prod_{n=1}^{\infty} (1 + 2^{-n}) < \infty$.

Let $\tilde{\mu}$ be a limit point of $(\mu_i)$ in the weak-star topology. Let
\[ K = \bigcap_{n=1}^{\infty} \bigcup_{i \geq n} K_i. \]
Then $\mu$ is supported on $K$. Furthermore
\[ K = \bigcap_{n=1}^{\infty} \bigcup_{i \geq n} K_i \subset \bigcap_{p=1}^{\infty} Z_{n_1, \ldots, n_p}. \]
However by the continuity of $\phi$, we can show that $\bigcap_{p=1}^{\infty} Z_{n_1, \ldots, n_p} = \bigcap_{p=1}^{\infty} \overline{Z_{n_1, \ldots, n_p}}$ by applying Cantor’s diagonal argument. Hence $K$ is a compact subset of $Z$.

On the other hand, by (4.5),
\[ e^{-m_i(x)s} = \mu_i(B(x, \gamma_i)) \leq \tilde{\mu}(B(x, \gamma_i)) \leq C \mu_i(B(x, \gamma_i)) = C e^{-m_i(x)s}, \quad \forall x \in K_i. \]
In particular, $1 \leq \sum_{x \in K_i} \mu_1(B(x, \gamma_i)) \leq \tilde{\mu}(K) \leq \sum_{x \in K_i} C \mu_1(B(x, \gamma_i)) \leq 2C$. Note that $K \subset \bigcup_{x \in K_i} \overline{B(x, \gamma_i/2)}$. By the first part of (b), for each $x \in K_i$ and $z \in \overline{B(x, \gamma_i)}$,
\[ \tilde{\mu}(B_{m_i(x)}(z, \epsilon)) \leq \tilde{\mu}(B(x, \gamma_i/2)) \leq C e^{-m_i(x)s}. \]
For each $z \in K$ and $i \in \mathbb{N}$, $z \in \overline{B(x, \gamma_i/2)}$ for some $x \in K_i$. Hence
\[ \tilde{\mu}(B_{m_i(x)}(z, \epsilon)) \leq C e^{-m_i(x)s}. \]
Define $\mu = \tilde{\mu}/\tilde{\mu}(K)$. Then $\mu \in M(K)$, and for each $z \in K$, there exists a sequence $k_i \uparrow \infty$ such that $\mu(B_{k_i}(z, \epsilon)) \leq C e^{-k_i/s} \tilde{\mu}(K)$. It follows that $\tilde{h}_\mu(T) \geq s$. \qed

5. Main notation and conventions

For the reader’s convenience, we summarize in Table 1 the main notation and typographical conventions used in this paper.

Acknowledgments

The first author was partially supported by the RGC grant and the Focused Investments Scheme B in CUHK. The second author was partially supported by NSFC, Fok Ying Tung Education Foundation and the Fundamental Research Funds for the Central Universities (WK0010000014). The authors thank Hanfeng Li and the referee for their helpful comments that led to the improvement of the paper.
Table 1  Main notation and conventions.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>((X, T))</td>
<td>A topological dynamical system (Section 1)</td>
</tr>
<tr>
<td>(M(X))</td>
<td>Set of all Borel probability measures on (X)</td>
</tr>
<tr>
<td>(M(X, T), E(X, T))</td>
<td>Set of (T)-invariant (resp. ergodic) Borel probability measures on (X)</td>
</tr>
<tr>
<td>(d_n)</td>
<td>(n)-th Bowen metric (cf. (1.1))</td>
</tr>
<tr>
<td>(B(x, \epsilon), \overline{B}(x, \epsilon))</td>
<td>Open (resp. closed) ball in ((X, d)) centered at (x) of radius (\epsilon)</td>
</tr>
<tr>
<td>(B_n(x, \epsilon), \overline{B}_n(x, \epsilon))</td>
<td>Open (resp. closed) ball in ((X, d_n)) centered at (x) of radius (\epsilon)</td>
</tr>
<tr>
<td>(\overline{h}<em>\mu(T), \underline{h}</em>\mu(T))</td>
<td>Measure-theoretic upper (resp. lower) entropy of (T) with respect to (\mu \in M(X)) (Section 1)</td>
</tr>
<tr>
<td>(h^{UC}_{top}(T, Z))</td>
<td>Upper capacity topological entropy of (Z) (Section 2)</td>
</tr>
<tr>
<td>(h^{B}_{top}(T, Z))</td>
<td>Bowen topological entropy of (Z) (Section 2)</td>
</tr>
<tr>
<td>(h^{P}_{top}(T, Z))</td>
<td>Packing topological entropy of (Z) (Section 2)</td>
</tr>
<tr>
<td>(h^{top}(T))</td>
<td>Topological entropy of (T) (Section 2)</td>
</tr>
<tr>
<td>(M^s_{N, \epsilon}(Z), M^s_{\epsilon}(Z), M^s(Z))</td>
<td>(Section 2)</td>
</tr>
<tr>
<td>(W^s_{N, \epsilon}(Z), W^s_\epsilon(Z), W^s(Z))</td>
<td>(Section 2)</td>
</tr>
<tr>
<td>(P^s_{N, \epsilon}(Z), P^s_\epsilon(Z), P^s(Z))</td>
<td>(Section 2)</td>
</tr>
<tr>
<td>(M^u_{N}(U, Z), M^u_{\epsilon}(U, Z))</td>
<td>(Section 2)</td>
</tr>
<tr>
<td>(h^{top}_{P}(T, U, Z))</td>
<td>(Section 2)</td>
</tr>
<tr>
<td>(N^*)</td>
<td>the set of infinite sequences of natural numbers endowed with product topology</td>
</tr>
</tbody>
</table>

References