A FORMULA OF CONDITIONAL ENTROPY
AND SOME APPLICATIONS

XIAOMIN ZHOU*

Wu Wen-Tsun Key Laboratory of Mathematics
USTC, Chinese Academy of Sciences
School of Mathematics
University of Science and Technology of China
Hefei, Anhui 230026, China

(Communicated by Jairo Bochi)

Abstract. In this paper we establish a formula of conditional entropy and give two examples of applications of the formula.

1. Introduction. Throughout this paper, a topological dynamical system (TDS) is a triple \((X,d,T)\) (or pair \((X,T)\) for short) consisting of a compact metric space \((X,d)\) and a continuous self-map \(T : X \to X\). Let \(\mathcal{M}(X),\mathcal{M}(X,T),\mathcal{M}^e(X,T)\) denote respectively the sets of all Borel probability measures, \(T\)-invariant Borel probability measures, and \(T\)-invariant ergodic Borel probability measures on \(X\). Given a TDS \((X,T)\), a \(T\)-invariant Borel probability measure \(\mu\) on \(X\) (i.e. \(\mu \in \mathcal{M}(X,T)\)) induces a measure preserving dynamical system \((X,B_X,\mu,T)\) where \(B_X\) is the \(\sigma\)-algebra of \(T\)-invariant sets of \(B_X\), and \(\mu \in \mathcal{M}(X,T)\).

In 1958 Kolmogorov \cite{Kolmogorov} associated to any measure preserving dynamical system \((X,B_X,\mu,T)\) an isomorphic invariant, namely the measure-theoretical entropy \(h_\mu(T)\). Later on in 1965, Alder, Konheim and McAndrew \cite{Alden} 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 \(h_{\text{top}}(T) = \sup\{h_\mu(T) : \mu \in \mathcal{M}(X,T)\}\). This variational principle was proved Goodwyn, Dinaburg and Goodman \cite{Goodwyn, Dinaburg, Goodman}, and plays a fundamental role in ergodic theory and dynamical systems.

Given a TDS \((X,T)\) and \(\mu \in \mathcal{M}(X,T)\). Let \(B_\mu\) be the completion of \(B_X\), then \(\mu\) can be decomposed into a generalized combination of ergodic measures, i.e.

\[
\mu = \int_X \mu_x^E \, d\mu(x),
\]

where \(E = \{E \in B_\mu : T^{-1}E = E\}\) is the \(\sigma\)-algebra of \(T\)-invariant sets of \(B_\mu\), and \(\mu_x^E\) denotes the conditional measure of \(\mu\) at \(x\) with respect to \(E\). It is well known that

\[2010 \text{ Mathematics Subject Classification. Primary: 37B05, 37B20, 37A15.} \]

\[\text{Key words and phrases. Measure-theoretical entropies, measure decomposition, conditional entropies.} \]

The author was partially supported by NNSF for Distinguished Young Scholar (11225105).
the ergodic decomposition of entropy (see [10, Theorem 15.12]), i.e.

\[ h_\mu(T) = h_\mu(T|\mathcal{E}) = \int h_{\mu^T}(T)d\mu(x), \]  

(1)

where \( h_\mu(T|\mathcal{E}) \) denotes the conditional entropy with respect to \( \mathcal{E} \) (see Section 2 for details).

A question arises naturally whether there is analogous conclusion for arbitrary \( T \)-invariant sub-\( \sigma \)-algebra (may include non-invariant set) of \( B_\mu \). More precise, let \( \mathcal{A} \) be a \( T \)-invariant sub-\( \sigma \)-algebra of \( B_\mu \) (i.e. \( T^{-1}\mathcal{A} = \mathcal{A}( \bmod \mu) \)) and \( \mu = \int_X \mu^T_\mathcal{A}d\mu(x) \) is the disintegration of \( \mu \) over \( \mathcal{A} \), how can we extend the formula (1) to general conditional entropy \( h_\mu(T|\mathcal{A}) \)? However, conditional measure \( \mu^T_\mathcal{A} \) is only Borel probability measure but maybe not \( T \)-invariant for \( \mu \)-a.e. \( x \in X \). Therefore, we need to introduce the measure-theoretic entropy for general Borel probability measures defined by Feng and Huang in [9].

For \( n \in \mathbb{N} \), the Bowen metric \( d_n \) is given by

\[ d_n(x, y) = \max\{d(T^i x, T^i y) : i = 0, 1, 2, \cdots, n - 1\}, \text{ for } x, y \in X. \]

Given \( \epsilon > 0 \), let \( B_{d_n}(x, \epsilon) = \{y \in X : d_n(x, y) < \epsilon\} \) be the \( d_n \)-ball about \( x \) of radius \( \epsilon \). We also write \( B_n(x, \epsilon) \) for convenience, when there is no confusion. Feng and Huang in [9] introduced the measure-theoretical lower and upper entropies of Borel probability measures which follows the idea of Brin and Katok [4].

**Definition 1.1.** Given a TDS \((X, T)\) and \( \mu \in \mathcal{M}(X) \). The measure-theoretical lower and upper entropies of \( \nu \) are defined respectively by

\[ \underline{h}_\mu(T) = \int \underline{h}_\mu(T, y)d\mu(y), \quad \overline{h}_\mu(T) = \int \overline{h}_\mu(T, y)d\mu(y), \]

where

\[ \underline{h}_\mu(T, y) = \lim_{\epsilon \to 0} \liminf_{n \to \infty} -\frac{\log \mu(B_n(y, \epsilon))}{n}, \]

\[ \overline{h}_\mu(T, y) = \lim_{\epsilon \to 0} \limsup_{n \to \infty} -\frac{\log \mu(B_n(y, \epsilon))}{n}. \]

Brin and Katok [4] proved that for any \( T \)-invariant Borel probability measure \( \mu \), \( \underline{h}_\mu(T, x) = \overline{h}_\mu(T, x) \) for \( \mu \)-a.e. \( x \in X \), and

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

The main result of this paper is as follows:

**Theorem 1.2.** Let \((X, T)\) be a TDS, \( \mu \in \mathcal{M}(X, T) \) and \( B_\mu \) be the completion of \( B_X \) under \( \mu \). If \( \mathcal{A} \) is a \( T \)-invariant sub-\( \sigma \)-algebra of \( B_\mu \) and

\[ \mu = \int_X \mu^A_\mathcal{A}d\mu(x) \]

is the measure disintegration of \( \mu \) over \( \mathcal{A} \), then \( \underline{h}_{\mu^A_\mathcal{A}}(T) = \overline{h}_{\mu^A_\mathcal{A}}(T) \) for \( \mu \)-a.e. \( x \in X \) and

\[ h_\mu(T|\mathcal{A}) = \int \underline{h}_{\mu^A_\mathcal{A}}(T)d\mu(x) = \int \overline{h}_{\mu^A_\mathcal{A}}(T)d\mu(x). \]

In particular, if \( \mu \) is ergodic then \( h_\mu(T|\mathcal{A}) = \underline{h}_{\mu^A_\mathcal{A}}(T) = \overline{h}_{\mu^A_\mathcal{A}}(T) \) for \( \mu \)-a.e. \( x \in X \).
As applications of Theorem 1.2, we give below two examples. Let \((X, T)\) and \((Y, S)\) be two TDSs. Suppose that \((Y, S)\) is a factor of \((X, T)\) in the sense that there exists a continuous surjective map \(\pi : (X, T) \to (Y, S)\) such that \(\pi \circ T = S \circ \pi\). The map \(\pi\) is called a factor map from \(X\) to \(Y\). We always write \(h_{\text{top}}(T|\pi)\) for the conditional topological entropy relative to \(\pi\) (see Section 2 for details).

**Corollary 1.** Let \(\pi : (X, T) \to (Y, S)\) be a factor map between two TDSs where \(S\) is homeomorphism. The following hold

1. if \(\mu \in \mathcal{M}(X, T)\) such that \(\pi^{-1}(\pi(x))\) is countable for \(\mu\)-a.e \(x \in X\), then \(h_{\mu}(T|\pi^{-1}\mathcal{B}_Y) = 0\);
2. if \(\pi^{-1}(y)\) is countable for each \(y \in Y\), then \(h_{\text{top}}(T|\pi) = 0\).

We remark that it is shown in [16, Theorem 1.1] that two semi-conjugate random dynamical systems (RDSs) on Polish spaces have the same entropy if the cardinal number of the pre-image of a point under the semi-conjugacy is finite almost everywhere. Here Corollary 1.3 gives an extension of this result to the case of countable fibers when \(X\) is compact.

To state the next corollary we need some notions. Let \((X, T)\) be a TDS with a homeomorphism \(T\), and \(\mu \in \mathcal{M}(X, T)\). The Pinsker \(\sigma\)-algebra \(\mathcal{P}_\mu(T)\) is defined as the smallest sub-\(\sigma\)-algebra of \(\mathcal{B}_\mu\) containing \(\{B \in \mathcal{B}_X : h_{\mu}(T, \{B, X \setminus B\}) = 0\}\). It is easy to see that the Pinsker \(\sigma\)-algebra \(\mathcal{P}_\mu(T)\) is a \(T\)-invariant \(\sigma\)-algebra of \(\mathcal{B}_\mu\).

**The stable set** of a point \(x \in X\) is defined as (see [14])

\[W^s(x, T) = \{y \in X : \lim_{n \to +\infty} d(T^n x, T^n y) = 0\}\]

and the **unstable set** of \(x\) is defined as

\[W^u(x, T) = \{y \in X : \lim_{n \to +\infty} d(T^{-n} x, T^{-n} y) = 0\}\].

The Bowen topological entropy \(h_{\text{top}}^{B}(T, Z)\) for any set \(Z\) in \((X, T)\) was introduced by Bowen in [3] (see Section 4 for details).

**Corollary 2.** Let \((X, T)\) be a TDS with a homeomorphism \(T\) from \(X\) onto itself. Then

1. if \(\mu \in \mathcal{M}^e(X, T)\) and \(\mu = \sum x \mu^x d\mu(x)\) is the measure decomposition of \(\mu\) over the Pinsker \(\sigma\)-algebra \(\mathcal{P}_\mu(T)\), then

\[h_{\mu^x}(T) \leq h_{\text{top}}^{B}(T, W^s(x, T) \cap W^u(x, T))\]

for \(\mu\)-a.e \(x \in X\).
2. \(h_{\text{top}}(T) = \sup_{x \in X} h_{\text{top}}^{B}(T, W^s(x, T) \cap W^u(x, T))\).

We remark that it is shown in [8, Theorem 1.2] that if \((X, T)\) is a finite entropy TDS with a homeomorphism \(T\) from \(X\) onto itself, and \(\mu\) is an invariant ergodic measure of positive entropy \(h_{\mu}(T) > 0\), then for \(\mu\)-a.e. \(x \in X\) we have the following lower bound on the Bowen dimension entropy of the closure of the stable and unstable sets of \(x\):

\[h_{\text{top}}^{B}(T, W^s(x, T) \cap W^u(x, T)) \geq h_{\mu}(T)\].

The part (2) of Corollary 1.4 gives a little improvement because it does not require that the system have a finite entropy.

This paper is organized as follows. In Section 2 we give the definitions and some basic properties of the measure-theoretic conditional entropy and measure decomposition, and also introduce a conditional version of Shannon-McMillan-Breiman theorem. In Section 3, we prove Theorem 1.2. In Section 4, we give proofs of Corollary 1 and Corollary 2.
2. Preliminaries. Let \((X, T)\) be a TDS. A partition of \(X\) is a family of subsets of \(X\) with union \(X\) and all elements of the family are disjoint; Denote by \(P_X\) the collection of all finite Borel partitions of \(X\). For any \(\alpha, \beta \in P_X\), \(\alpha\) is said to be finer than \(\beta\) (write \(\beta \preceq \alpha\)) if each atom of \(\alpha\) is contained in some atom of \(\beta\). Given a partition \(\alpha\) of \(X\) and \(x \in X\), denote by \(\alpha(x)\) the atom of \(\alpha\) containing \(x\). If \(\{\alpha_i\}_{i \in I}\) is a countable family of finite Borel partition of \(X\), the partition \(\alpha = \vee_{i \in I} \alpha_i\), is called a measurable partition. For a measurable partition \(\alpha\), put \(\alpha_0^{-1} = \bigvee_{i=0}^{n-1} T^{-i} \alpha\) and \(\alpha^- = \bigvee_{n=1}^{\infty} T^{-n} \alpha\). It is known that \(\mathcal{M}(X)\) and \(\mathcal{M}(X, T)\) (defined before) are convex, compact metric spaces when endowed with the weak*-topology.

Let \(\mu \in \mathcal{M}(X, T)\). Given a \(T\)-invariant sub-\(\sigma\)-algebra \(\mathcal{A}\) of \(\mathcal{B}_\mu\) and \(\alpha \in P_X\), the conditional informational function of \(\alpha\) with respect to \(\mathcal{A}\) is defined by

\[
I_\mu(\alpha | \mathcal{A})(x) = \sum_{A \in \alpha} -1_A(x) \log \mathbb{E}(1_A | \mathcal{A})(x),
\]

where \(1_A\) is the characteristic function of \(A\) and \(\mathbb{E}(1_A | \mathcal{A})\) is the conditional expectation of \(1_A\) with respect to \(\mathcal{A}\). Define

\[
H_\mu(\alpha | \mathcal{A}) = \int_X I_\mu(\alpha | \mathcal{A})(x) d\mu(x) = \sum_{A \in \alpha} \int_X -\mathbb{E}(1_A | \mathcal{A}) \log \mathbb{E}(1_A | \mathcal{A}) d\mu(x).
\]

Note that \(H_\mu(\bigvee_{i=0}^{n-1} T^{-i} \alpha | \mathcal{A})\) is a subadditive sequence, then define the conditional entropy of \(\alpha\) with respect to \(\mathcal{A}\) by

\[
h_\mu(T, \alpha | \mathcal{A}) = \lim_{n \to \infty} \frac{1}{n} H_\mu(\bigvee_{i=0}^{n-1} T^{-i} \alpha | \mathcal{A}) = \inf_{n \geq 1} \frac{1}{n} H_\mu(\bigvee_{i=0}^{n-1} T^{-i} \alpha | \mathcal{A}).
\]

Moreover, define the conditional entropy with respect to \(\mathcal{A}\) by

\[
h_\mu(T | \mathcal{A}) = \sup_{\alpha \in P_X} h_\mu(T, \alpha | \mathcal{A}).
\]

If \(\{\alpha_i\}_{i \geq 1}\) is a family of finite Borel partition with \(\alpha_1 \preceq \alpha_2 \preceq \alpha_3 \preceq \ldots\) and \(\text{diam}(\alpha_i) \to 0\) as \(i \to \infty\), we can compute the conditional entropy by

\[
h_\mu(T | \mathcal{A}) = \sup_{i \geq 1} h_\mu(T, \alpha_i | \mathcal{A}) = \lim_{i \to \infty} h_\mu(T, \alpha_i | \mathcal{A}).
\]

A cover of \(X\) is a finite family of subsets of \(X\) whose union is \(X\). Let \(C_X\) denote the collection of all finite open covers of \(X\). For \(\mathcal{U}, \mathcal{V} \in C_X\), we say that \(\mathcal{U}\) is finer than \(\mathcal{V}\) (write \(\mathcal{V} \preceq \mathcal{U}\)) if each elements of \(\mathcal{U}\) is contained in some element of \(\mathcal{V}\). Given a factor map \(\pi : (X, T) \to (Y, S)\) between two TDSs and a cover \(\mathcal{U} \in C_X\). For any \(E \subset X\), denote by \(N(\mathcal{U}, E)\) the minimal cardinality of any subcover of \(\mathcal{U}\) that covers \(E\). Let \(N(\mathcal{U} | \pi) = \sup_{y \in Y} N(\mathcal{U}, \pi^{-1}(y))\). Clearly, \(N(\mathcal{U} | \pi) \leq N(\mathcal{U}, X)\).

It is not hard to see that \(a_n = \log N(\bigvee_{i=0}^{n-1} T^{-i} \mathcal{U} | \pi)\) is a subadditive sequence, so we can define the conditional topological entropy of \(\mathcal{U}\) relative to \(\pi\) by

\[
h_{\text{top}}(\mathcal{U}, \mathcal{U} | \pi) = \lim_{n \to \infty} \frac{1}{n} \log N(\bigvee_{i=0}^{n-1} T^{-i} \mathcal{U} | \pi) = \inf_{n \geq 1} \frac{1}{n} \log N(\bigvee_{i=0}^{n-1} T^{-i} \mathcal{U} | \pi).
\]

Similarly, we can define the conditional topological entropy relative to \(\pi\) by

\[
h_{\text{top}}(T | \pi) = \sup_{\mathcal{U} \in C_X} h_{\text{top}}(T, \mathcal{U} | \pi).
\]

There is a well-known result which characterizes the relation between the measure-theoretic conditional entropy and conditional topological entropy relative to a factor.
map (see [6, 15]), i.e.

\[
h_{\text{top}}(T \mid \pi) = \sup_{\mu \in \mathcal{M}(X, T)} h_\mu(T \mid \pi^{-1}B_Y) = \sup_{\mu \in \mathcal{M}_+(X, T)} h_\mu(T \mid \pi^{-1}B_Y). \tag{2}
\]

The support of \(\mu \in \mathcal{M}(X, T)\) is defined to be the set of all points \(x\) in \(X\) for which every open neighborhood \(U\) of \(x\) has positive measure, that is

\[\text{supp}(\mu) = \{x \in X : \mu(U) > 0 \text{ for every open neighborhood } U \text{ of } x\}.\]

In the following we give the definitions and some properties of measure disintegration and conditional measures (see [7, Section 5]).

Let \((X, T)\) be a TDS, \(\mu \in \mathcal{M}(X, T)\) and \(\mathcal{B}_\mu\) be the completion of \(\mathcal{B}_X\) under the measure \(\mu\). Then \((X, \mathcal{B}_\mu, \mu, T)\) is a Lebesgue system. The sets \(A \in \mathcal{B}_\mu\), which are the unions of atoms of a measurable partition \(\alpha\), form a sub-\(\sigma\)-algebra of \(\mathcal{B}_\mu\) denoted by \(\hat{\alpha}\) or \(\alpha\) if there is no ambiguity. Every sub-\(\sigma\)-algebra of \(\mathcal{B}_\mu\) coincides with a \(\sigma\)-algebra constructed in this way (mod \(\mu\)). Let \(\mathcal{A}\) be a sub-\(\sigma\)-algebra of \(\mathcal{B}_\mu\) and \(\alpha\) be a measurable partition of \(X\) with \(\hat{\alpha} = \mathcal{A}\) (mod \(\mu\)). Then \(\mu\) can disintegrated over \(\mathcal{A}\) as

\[\mu = \int_X \mu_x^A \, d\mu(x),\]

where \(\mu_x^A\) is a Borel probability measure on \(X\) and \(\mu_x^A(\alpha(x)) = 1\) for \(\mu\)-almost every \(x \in X\). This disintegration is characterized by (2) and (3) below:

1. for every \(f \in L^1(X, \mathcal{B}_X, \mu)\), \(f \in L^1(X, \mathcal{B}_X, \mu_x^A)\) for \(\mu\)-a.e. \(x \in X\),
2. the function \(x \mapsto \int_X f(y) \, d\mu_x^A(y)\) is in \(L^1(X, \mathcal{A}, \mu)\),
3. for every \(f \in L^1(X, \mathcal{B}_X, \mu)\), \(\mathbb{E}_\mu(f \mid \mathcal{A})(x) = \int_X f \, d\mu_x^A\) for \(\mu\)-a.e. \(x \in X\).

Then, for any \(f \in L^1(X, \mathcal{B}_X, \mu)\), the following holds:

\[\int_X \left(\int_X f \, d\mu_x^A\right) \, d\mu(x) = \int_X f \, d\mu.\]

Define for \(\mu\)-almost every \(x \in X\) the set \(\Gamma_x = \{y \in X : \mu_x^A = \mu_y^A\}\). Then \(\mu_x^A(\Gamma_x) = 1\) for \(\mu\)-almost every \(x \in X\). Hence given any \(f \in L^1(X, \mathcal{B}_X, \mu)\), for \(\mu\)-almost every \(x \in X\), one has

\[\mathbb{E}_\mu(f \mid \mathcal{A})(y) = \int_X f \, d\mu_x^A = \int_X f \, d\mu_y^A = \mathbb{E}_\mu(f \mid \mathcal{A})(x) \tag{3}\]

for \(\mu_x^A\)-almost every \(y \in X\). Particularly, if \(f\) is \(\mathcal{A}\)-measurable, then for \(\mu\)-almost every \(x \in X\), one has

\[f(y) = f(x) \quad \text{for } \mu_x^A\text{-almost every } y \in X. \tag{4}\]

To prepare the proof of the main result in the next section we also need the following result which is a conditional version of Shannon-McMillan-Breiman theorem. Its proof is completely similar to the proof of Shannon-McMillan-Breiman theorem (see for example [2, Theorem 4.2], [10] or [18]).

**Theorem 2.1.** Let \((X, T)\) be a TDS, \(\mu \in \mathcal{M}(X, T)\), \(\alpha \in \mathcal{P}_X\) and \(\mathcal{A}\) be a \(T\)-invariant sub-\(\sigma\)-algebra of \(\mathcal{B}_\mu\), i.e. \(T^{-1}\mathcal{A} = \mathcal{A}\) (mod \(\mu\)). Then there exists a \(T\)-invariant function \(h_\mu(\alpha \mid \mathcal{A}, x) \in L^1(\mu)\) such that

\[\int_X h_\mu(\alpha \mid \mathcal{A}, x) \, d\mu(x) = h_\mu(T, \alpha \mid \mathcal{A}) \quad \text{and} \quad \lim_{n \to \infty} \frac{L_n(\alpha_0^{n-1} \mid \mathcal{A})(x)}{n} = h_\mu(\alpha \mid \mathcal{A}, x),\]

where

\[L_n(\alpha_0^{n-1} \mid \mathcal{A})(x) = \frac{1}{n} \sum_{i=0}^{n-1} \mathbb{1}_{\mathcal{A}}(T^i x).\]
for $\mu$-a.e. $x \in X$ and in $L^1(\mu)$. Moreover, if $\mu$ is ergodic then $h_\mu(\alpha|\mathcal{A},x) = h_\mu(T,\alpha|\mathcal{A})$ for $\mu$-a.e. $x \in X$.

3. **Proof of main theorem.** In this section, we prove Theorem 1.2. The following lemma plays a key role in our proof.

**Lemma 3.1.** Let $(X,T)$ be a TDS, $\mu \in \mathcal{M}(X,T)$, $\alpha \in P_X$ and $\mathcal{A}$ a $T$-invariant sub-$\sigma$-algebra of $\mathcal{B}_\mu$. Then for $\mu$-a.e. $x \in X$, there exists $W_x \in \mathcal{B}_X$ with $\mu_x^A(W_x) = 1$ such that

$$\lim_{n \to \infty} \frac{-\log \mu_x^n(\alpha_0^{n-1}(y))}{n} = h_\mu(\alpha|\mathcal{A},x)$$

for each $y \in W_x$, where $\mu = \int_X \mu_x^A d\mu(x)$ is the measure disintegration of $\mu$ over $\mathcal{A}$ and $h_\mu(\alpha|\mathcal{A},x)$ is the function obtained in Theorem 2.1.

**Proof.** Note that $\lim_{n \to \infty} (I_\mu(\alpha_0^{n-1}|\mathcal{A})(x)/n) = h_\mu(\alpha|\mathcal{A},x)$ for $\mu$-a.e. $x \in X$ and the function $h_\mu(\alpha|\mathcal{A},x)$ is $\mathcal{A}$-measurable. Using (3) for the characteristic function $1_B$, $B \in \alpha_0^{n-1}$ and (4) for $h_\mu(\alpha|\mathcal{A},x)$, there exists $X_1 \in \mathcal{B}_X$ with $\mu(X_1) = 1$ such that for each $x \in X_1$, one can find $W_x \in \mathcal{B}_X$ with $\mu_x^A(W_x) = 1$, and for $y \in W_x$ we have

1. $\lim_{n \to \infty} (I_\mu(\alpha_0^{n-1}|\mathcal{A})(y)/n) = h_\mu(\alpha|\mathcal{A},y) = h_\mu(\alpha|\mathcal{A},x)$;
2. $E_\mu(1_B|\mathcal{A})(y) = E_\mu(1_B|\mathcal{A})(x) = \mu_x^A(B)$ for any $B \in \alpha_0^{n-1}$ and each $n \in \mathbb{N}$.

Moreover, for any $x \in X_1$ and $y \in W_x$ one has

$$\lim_{n \to \infty} \frac{-\log \mu_x^n(\alpha_0^{n-1}(y))}{n} = \lim_{n \to \infty} \frac{-\log E_\mu(1_{\alpha_0^{n-1}}|\mathcal{A})(y)}{n} = \frac{I_\mu(\alpha_0^{n-1}|\mathcal{A})(y)}{n} = h_\mu(\alpha|\mathcal{A},x).$$

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

We are going to prove the Theorem 1.2.

**Proof of Theorem 1.2.** Let $\{\alpha_i\}_{i=1}^\infty$ be a family of finite Borel partition of $X$ with $\alpha_1 \preceq \alpha_2 \preceq \alpha_3 \preceq \cdots$, $\text{diam}(\alpha_i) \to 0$ as $i \to \infty$ and $\mu(\partial \alpha_i) = 0$ for $i \in \mathbb{N}$. Actually by the Monotone Convergence Theorem it is sufficient to show that the following equation holds

$$h_{\mu^A}(T) = \tilde{h}_{\mu^A}(T) = \sup_{i \geq 1} h_\mu(\alpha_i|\mathcal{A},x)$$

for $\mu$-a.e. $x \in X$, where $\mu = \int_X \mu_x^A d\mu(x)$ is the measure disintegration of $\mu$ over $\mathcal{A}$ and $h_\mu(\alpha_i|\mathcal{A},x)$ is the function obtained in Theorem 2.1.

First we show that $\tilde{h}_{\mu^A}(T) \leq \sup_{i \geq 1} h_\mu(\alpha_i|\mathcal{A},x)$ for $\mu$-a.e. $x \in X$. For any $\epsilon > 0$, since $\text{diam}(\alpha_i) \to 0$, there exists $N = N(\epsilon) \in \mathbb{N}$ such that $\text{diam}(\alpha_i) \leq \epsilon$ when $i \geq N$. So, for every $x,y \in X$, we have $(\alpha_i)^{n-1}_0(y) \subset B_n(y,\epsilon)$ for $n \geq 1$ and

$$-\log \mu_x^n(B_n(y,\epsilon)) \leq -\log \mu_x^n((\alpha_i)^{n-1}_0(y)).$$

Thus

$$\tilde{h}_{\mu^A}(T,y) = \lim_{\epsilon \to 0} \limsup_{n \to +\infty} \frac{-\log \mu_x^n(B_n(y,\epsilon))}{n}$$

and
Step 1. We are going to find such \((\epsilon < \epsilon)\) and \(\alpha\) number of elements in \(M\) and \(X\) of it is sufficient to show the following property (#): for any finite Borel partition \(\alpha\)

Next we are going to show that for \(\mu\)-a.e. \(x \in X\)

It is sufficient to show the following property (#): for any finite Borel partition \(\alpha\) of \(X\) satisfying \(\mu(\partial \alpha) = 0\) and \(\epsilon > 0\), we can find a \(\delta > 0\), a measurable subset \(D\) of \(X\) such that

\[
\mu^A(D) > 1 - \epsilon^\frac{1}{2} \text{ and } \liminf_{n \to +\infty} \frac{-\log \mu^A(B_n(y, \delta))}{n} \geq h_\mu(\alpha|A, \hat{x}) - 3(\Delta + \epsilon)
\]  

(8)

for any \(y \in D\), where \(\Delta = 2\sqrt{\epsilon} \log(M - 1) - 2\sqrt{\epsilon} \log 2\sqrt{\epsilon} - (1 - 2\sqrt{\epsilon}) \log(1 - 2\sqrt{\epsilon})\) and \(M\) is the number of elements in \(\alpha\).

Let \(\alpha\) be a finite Borel partition of \(X\) with \(\mu(\partial \alpha) = 0\) and \(\epsilon > 0\). Let \(M\) be the number of elements in \(\alpha\). If \(M = 1\) then it is clear that property (#) holds for \(\alpha\) and \(\epsilon\). Now assume that \(M \geq 2\). Without loss generality, we require additionally \(\epsilon < \epsilon(n, \alpha)\).

We divide the remaining proof into the following three steps.

**Step 1.** We are going to find such \(\delta\) and \(I\) satisfying the property (#) for \(\alpha\) and \(\epsilon\).

For \(\tau > 0\) let

\[U_\tau(\alpha) = \{x \in X : \text{the ball } B(x, \tau) \text{ is not contained in } \alpha(x)\},\]

where \(\alpha(x)\) denotes the element of the partition \(\alpha\) containing \(x\). Since \(\bigcap_{\tau > 0} U_\tau(\alpha) = \partial \alpha\), we have

\[\mu(U_\tau(\alpha)) \to 0, \text{ as } \tau \to 0.\]

Therefore there exists \(\delta > 0\) such that \(\mu(U_\tau(\alpha)) < \epsilon\) for any \(0 < \tau \leq \delta\).

For \(n \in \mathbb{N}\) we define

\[A_n = \{y \in X : \frac{1}{k} \sum_{i=0}^{k-1} 1_{U_\delta(\alpha)}(T^i(y)) \leq 2\sqrt{\epsilon} \text{ for any } k \geq n\},\]

then we are going to show the measure of \(A_n\) is large enough for \(n \geq L\) for some \(L\).

By the Birkhoff ergodic theorem, the averages

\[\frac{1}{n} \sum_{i=0}^{n-1} 1_{U_\delta(\alpha)}(T^i(y))\]
converge almost everywhere and in $L^1_\mu$ as $n \to \infty$ to a $T$-invariant function $f^*$ and
\[
\int f^*(y) d\mu(y) = \mu(U_\beta(\alpha)) < \epsilon,
\]
where $1_{U_\beta(\alpha)}$ is the characteristic function of the set $U_\beta(\alpha)$. Thus by Egoroff Theorem we can find a large natural number $L$ such that
\[
\mu(\{y \in X : \frac{1}{K} \sum_{i=0}^{k-1} 1_{U_\beta(\alpha)}(T^i(y)) - f^*(y)\} < \sqrt{\epsilon} \text{ for any } k \geq L) > 1 - \sqrt{\epsilon}.
\]
By Chebyshev’s Inequality we have
\[
\mu(\{y \in X : f^*(y) \geq \sqrt{\epsilon}\}) \cdot \sqrt{\epsilon} < \int f^*(y) d\mu(y) < \epsilon.
\]
For $n \geq L$,
\[
\mu(A_n^c) \leq \mu(\{y \in X : \exists k \geq L, \frac{1}{K} \sum_{i=0}^{k-1} 1_{U_\beta(\alpha)}(T^i(y)) > 2\sqrt{\epsilon}\})
\leq \mu(\{y \in X : \exists k \geq L, \frac{1}{K} \sum_{i=0}^{k-1} 1_{U_\beta(\alpha)}(T^i(y)) - f^*(y)\} \geq \sqrt{\epsilon}\})
+ \mu(\{y \in X : f^*(y) \geq \sqrt{\epsilon}\})
< 2\sqrt{\epsilon}.
\]
Therefore we obtain that
\[
\mu(A_n) > 1 - 2\sqrt{\epsilon}
\]
for $n \geq L$.

Define $Q_n = \{x \in X : \mu_x^A(A_n) \geq 1 - 2\epsilon^\frac{1}{4}\}$, then $Q_n^c = \{x \in X : \mu_x^A(A_n^c) < 2\epsilon^\frac{1}{4}\}$.
Therefore for $n \geq L$
\[
\mu(Q_n^c) \cdot 2\epsilon^\frac{1}{4} \leq \int \mu_x^A(A_n^c) d\mu(x) = \mu(A_n^c) < 2\sqrt{\epsilon}.
\]
We have $\mu(Q_n) > 1 - \epsilon^\frac{1}{4}$, for $n \geq L$.

Obviously, the sets $A_n$ are nested, i.e. $A_{n-1} \subseteq A_n$ for $n \geq 1$. Thus there exists $l_0 > L$ such that for $x \in Q_{l_0}$,
\[
\mu_x^A(A_l) \geq \mu_x^A(A_{l_0}) \geq 1 - 3\epsilon^\frac{1}{4}
\]
for any $l \geq l_0$.

By Lemma 3.1 there exists $X_1 \subseteq X$ with $\mu(X_1) = 1$ such that for $x \in X_1$, there exists $W_x \in B_X$ with $\mu_x^A(W_x) = 1$ such that
\[
\lim_{n \to \infty} - \log \frac{\mu_x^A(\alpha_0^{n-1}(y))}{n} = h_\mu(\alpha|A, x)
\]
for each $y \in W_x$.

Let $I = X_1 \cap Q_{l_0}$. Obviously $\mu(I) > 1 - \epsilon^\frac{1}{4}$. Fix $\hat{x} \in I$. We can find a large number $l_1$ such that $\mu_x^A(B_l) \geq 1 - \epsilon^\frac{1}{4}$ for any $l \geq l_1$, where $B_l$ is the set of all points $y \in W_x$ such that
\[
- \log \frac{\mu_x^A(\alpha_0^{n-1}(y))}{n} \geq h_\mu(\alpha|A, \hat{x}) - \epsilon
\]
for all $n \geq l$.

Take $l \geq \max\{l_0, l_1\}$. Let $E = A_l \cap B_l$; Then $\mu_x^A(E) > 1 - 4\epsilon^\frac{1}{4}$.  

Step 2. For \( y \in X \) and \( n \in \mathbb{N} \), we call the collection
\[
C(n, y) = (\alpha(y), \alpha(T(y)), \ldots, \alpha(T^{n-1}(y)))
\]
the \((\alpha, n)\)-name of \( y \). Since each point in one atom \( V \) of \( \alpha_0^{n-1} \) has the same \((\alpha, n)\)-name, we define
\[
C(n, V) := C(n, x)
\]
for any \( x \in V \), which is called the \((\alpha, n)\)-name of \( V \). We give a metric \( d_n^\alpha \) between \((\alpha, n)\)-names of \( y \) and \( z \) as follows:
\[
d_n^\alpha(C(n, y), C(n, z)) = \frac{1}{n} \#\{0 \leq i \leq n - 1 : \alpha(T^i(y)) \neq \alpha(T^i(z))\}.
\]
Now if \( z \in B_\delta(y, \delta) \) then for any \( 0 \leq i \leq n - 1 \) either \( T^i y \) and \( T^i z \) belong to the same element of \( \alpha \) or \( T^i y \in U_\delta(\alpha) \). Hence when \( y \in E \), \( n \geq 1 \) and \( z \in B_\delta(y, \delta) \), we have that the \( d_n^\alpha \) distance between the \((\alpha, n)\)-names of \( y \) and \( z \) is less than \( 2\sqrt{\tau} \). Furthermore, Bowen’s ball \( B_\delta(y, \delta) \) is contained in the set of points \( z \) where \((\alpha, n)\)-names are \( 2\sqrt{\tau} \)-close to the \((\alpha, n)\)-name of \( y \).

Let \( B \in \alpha_0^{n-1} \) and \( L_n(B) \) be the total number of \( V \in \alpha_0^{n-1} \) such that \( C(n, V) \) is \( 2\sqrt{\tau} \)-close to \( C(n, B) \). By Stirling’s formula and also a combinatorial argument admits the following estimate:
\[
L_n(B) \leq \sum_{i=0}^{[2n\sqrt{\tau}]} \binom{n}{i} (M - 1)^i \leq \exp((\epsilon + \Delta)n) \tag{9}
\]
for each \( n \geq l_2 \) for some \( l_2 \in \mathbb{N} \), where \( M \) is the number of elements in \( \alpha \) and
\[
\Delta = 2\sqrt{\tau} \log(M - 1) - 2\sqrt{\tau} \log(2\sqrt{\tau} - (1 - 2\sqrt{\tau}) \log(1 - 2\sqrt{\tau})
\]
(see for example [17, Page 144]). More precisely, we had shown that for any \( y \in E \), \( n \geq \max\{l, l_2\} \) and \( B \in \alpha_0^{n-1} \),
\[
B_\delta(y, \delta) \subseteq \{z \in X : C(n, z) \text{ is } 2\sqrt{\tau}\text{-close to } C(n, y)\}
\]
\[
= \bigcup\{V \in \alpha_0^{n-1} : C(n, V) \text{ is } 2\sqrt{\tau}\text{-close to } C(n, y)\} \tag{10}
\]
and
\[
\#\{V \in \alpha_0^{n-1} : C(n, V) \text{ is } 2\sqrt{\tau}\text{-close to } C(n, B)\} \leq \exp((\epsilon + \Delta)n). \tag{11}
\]

Step 3. In the following we want to find suitable subset \( D \) of \( E \) satisfying the property (8).

For \( n \in \mathbb{N} \), let \( E_n \) be the set of points \( y \in E \) such that there exists an element \( V \in \alpha_0^{n-1} \) with
\[
\mu_\varepsilon^4(V) > \exp\{(-h_\mu(\alpha \mid A, \hat{x}) + 2(\Delta + \epsilon))n\}
\]
and the \((\alpha, n)\)-name of \( V \) is \( 2\sqrt{\tau}\)-close to the \((\alpha, n)\)-name of \( y \).

We wish to estimate the measure of \( E_n \) for \( n \geq \max\{l, l_2\} \). Put
\[
\mathcal{F}_n = \{V \in \alpha_0^{n-1} : \mu_\varepsilon^4(V) > \exp\{(-h_\mu(\alpha \mid A, \hat{x}) + 2(\Delta + \epsilon))n\}\}
\]
It is clear that
\[
\#\mathcal{F}_n \leq \exp\{(h_\mu(\alpha \mid A, \hat{x}) - 2(\Delta + \epsilon))n\}
\]
since \( \mu_\varepsilon^4(X) = 1 \).

Now for \( n \geq \max\{l, l_2\} \) and \( y \in E_n \), on one hand since \( y \in B_l \)
\[
\mu_\varepsilon^4(\alpha_0^{n-1}(y)) \leq \exp\{(-h_\mu(\alpha \mid A, \hat{x}) + \epsilon)n\}.
\]
On the other hand by the definition of $E_n$, there exists $V \in F_n$ with $(\alpha, n)$-name of $V$ is $2\sqrt{\epsilon}$-close to the $(\alpha, n)$-name of $\alpha_0^{-1}(y)$. Summing up, we had shown that

$$E_n \subseteq \bigcup \{B : B \in G_n\},$$

where $G_n$ is the set all elements $B \in \alpha_0^{-1}$ satisfying that

$$\mu^{A}(B) \leq \exp\{(-\mu(A, \hat{x}) + \epsilon)n\}$$

and $(\alpha, n)$-name of $B$ is $2\sqrt{\epsilon}$-close to the $(\alpha, n)$-name of $V$ for some $V \in F_n$. Thus by (11)

$$\#G_n \leq \exp\{(\Delta + \epsilon)n\} \cdot \#F_n.$$

Moreover

$$\mu^{A}(E_n) \leq \exp\{(-\mu(A, \hat{x}) + \epsilon)n\} \cdot \#G_n \leq \exp\{-\Delta \epsilon n\}.$$ 

Next we take $l_3 \geq \max\{l, l_2\}$ with $\sum_{n=1}^{\infty} \exp\{-\Delta \epsilon n\} < \epsilon^{\frac{1}{3}}$. Then $\mu^{A}(\bigcup_{n \geq l_3} E_n) < \epsilon^{\frac{1}{3}}$. Let $D = E \setminus \bigcup_{n \geq l_3} E_n$. Then $\mu^{A}(D) > 1 - 5\epsilon^{\frac{1}{3}}$.

Given $y \in D$ and $n \geq l_3$. Since $y \in E \setminus E_n$, it is clear that for each $V \in \alpha_0^{-1}$ whose $(\alpha, n)$-name is $2\sqrt{\epsilon}$-close to the $(\alpha, n)$-name of $y$, one has

$$\mu^{A}(V) \leq \exp\{(-\mu(A, \hat{x}) + 2(\Delta + \epsilon))n\}.$$ 

Moreover using (10) and (11), we have

$$\mu^{A}(B_n(y, \delta)) \leq \exp\{(\Delta + \epsilon)n\} \cdot \exp\{(-\mu(A, \hat{x}) + 2(\Delta + \epsilon))n\}$$

$$= \exp\{(-\mu(A, \hat{x}) + 3(\Delta + \epsilon))n\}.$$ 

Thus for any $y \in D$ and $n \geq l_3$,

$$-\frac{\log \mu^{A}(B_n(y, \delta))}{n} \geq \mu(A, \hat{x}) - 3(\Delta + \epsilon).$$

Summing up, we obtain the property (#). This finishes the proof of the theorem.

\[\square\]

4. Applications of Theorem 1.2. In this section, we prove Corollary 1 and 2 using the formula of conditional entropy given by Theorem 1.2. To prepare the proofs we state some notions and useful lemmas.

Let $(X, T)$ be a TDS. Recall that $d_n$ and $B_n(x, \epsilon)$ are defined in Introduction. The Bowen topological entropy of subsets was first introduced by Bowen [3] in a way resembling Hausdorff dimension and can be defined as follows. For $Z \subseteq X$, $s \geq 0$, $N \in \mathbb{N}$ and $\epsilon > 0$, define

$$M^s_{N, \epsilon}(Z) = \inf \sum \exp(-sn_i),$$

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 $M^s_{N, \epsilon}(Z)$ does not decrease as $N$ increases and $\epsilon$ decreases, hence the following limits exists:

$$M^s\epsilon(Z) = \lim_{N \to \infty} M^s_{N, \epsilon}(Z), \quad M^s(Z) = \lim_{\epsilon \to 0} M^s\epsilon(Z).$$

The Bowen topological entropy $h_{top}^B(T, Z)$ can be defined as a critical value of the parameter $s$, where $M^s(Z)$ jumps from $\infty$ to 0, i.e.

$$M^s(Z) = \begin{cases} 0, & s > h_{top}^B(T, Z), \\ \infty, & s < h_{top}^B(T, Z). \end{cases}$$
The following results are elementary (see for example [3, Propositions 1 and 2]).

**Proposition 1.**
1. For $Z \subseteq Z'$, $h^B_{\text{top}}(T, Z) \leq h^B_{\text{top}}(T, Z')$.
2. For $Z \subseteq \bigcup_{i=1}^{\infty} Z_i$, $s \geq 0$ and $\varepsilon > 0$, we have
   $$M^s_\varepsilon(Z) \leq \sum_{i=1}^{\infty} M^s_\varepsilon(Z_i), \quad h^B_{\text{top}}(T, Z) \leq \sup_{i \geq 1} h^B_{\text{top}}(T, Z_i).$$
3. In particular, $h^B_{\text{top}}(T, X) = h_{\text{top}}(T)$.

Feng and Huang [9] establish the variational principles for Bowen topological entropy of subsets. We restate their result as follows.

**Theorem 4.1.** [9, Theorem 1.2 (i)] Let $(X, T)$ be a TDS. If $K \subseteq X$ is non-empty and compact, then
$$h^B_{\text{top}}(T, K) = \sup \{h_\mu(T) : \mu \in \mathcal{M}(X), \mu(K) = 1\}.$$  

**Proof of Corollary 1.** Let $\mu \in \mathcal{M}(X, T)$ such that $\pi^{-1}(\pi(x))$ is countable for $\mu$-a.e $x \in X$. Let $\mu = \int \mu_x d\mu(x)$ be the measure disintegration of $\mu$ over $\pi^{-1}B_Y$ where $B_Y$ is the Borel $\sigma$-algebra of $Y$. Note that $T^{-1}\pi^{-1}(B_Y) = \pi^{-1}S^{-1}(B_Y) = \pi^{-1}B_Y$. Hence by Theorem 1.2 we have
$$h_\mu(T|\pi^{-1}B_Y) = \int h_{\mu_x}(T)d\mu(x).$$

Since $\pi^{-1}(\pi(x))$ is countable for each $\mu$-a.e $x \in X$, $h^B_{\text{top}}(T, \pi^{-1}(\pi(x))) = 0$ by Proposition 1 (2). Now for $\mu$-a.e. $x \in X$, $\text{supp}(\mu_x) \subset \pi^{-1}(\pi(x))$, then by Theorem 4.1 and Proposition 1 (1) we have
$$h_{\mu_x}(T) \leq h^B_{\text{top}}(T, \text{supp}(\mu_x)) \leq h^B_{\text{top}}(T, \pi^{-1}(\pi(x))) = 0.$$

Thus $h_{\mu_x}(T) = 0$ for $\mu$-a.e. $x \in X$ and $h_\mu(T|\pi^{-1}B_Y) = \int h_{\mu_x}(T)d\mu(x) = 0$. Finally if $\pi^{-1}(y)$ is countable for each $y \in Y$, then
$$h_{\text{top}}(T|\pi) = \sup_{\mu \in \mathcal{M}(X, T)} h_\mu(T|\pi^{-1}B_Y) = 0$$
by the variational principle of condition entropy (2).

For a TDS $(X, T)$ with a homeomorphism $T$, the stable set of a point $x \in X$ is defined as
$$W^s(x, T) = \{y \in X : \lim_{n \to +\infty} d(T^nx, T^ny) = 0\}$$
and the unstable set of $x$ is defined as
$$W^u(x, T) = \{y \in X : \lim_{n \to +\infty} d(T^{-nx}, T^{-ny}) = 0\}.$$  

Clearly, $W^s(x, T) = W^u(x, T^{-1})$ and $W^u(x, T) = W^s(x, T^{-1})$ for each $x \in X$. Recently, there are many results related to the chaotic behavior and stable (unstable) sets in positive entropy systems (see [8, 13, 14]). So it indicates that stable and unstable sets are important in the study of topological dynamical systems.

Huang, Li and Ye [14] studied the stable sets and unstable sets in positive entropy systems and proved the following lemma.

**Lemma 4.2.** [14, Lemma 3.1] Let $(X, T)$ be a TDS with a homeomorphism $T$ and $\mu \in \mathcal{M}^c(X, T)$ with $h_\mu(T) > 0$. If $\mu = \int X u_x^p d\mu(x)$ is the measure disintegration of $\mu$ over the Pinsker $\sigma$-algebra $\mathcal{P}_\mu(T)$, then for $\mu$-a.e. $x \in X$,
$$\overline{W^s(x, T)} \cap \text{supp}(\mu_x^P) = \text{supp}(\mu_x^P) \quad \text{and} \quad \overline{W^u(x, T)} \cap \text{supp}(\mu_x^P) = \text{supp}(\mu_x^P).$$
Proof of Corollary 2. Let $\mu \in \mathcal{M}^e(X, T)$ and $\mathcal{B}_\mu$ be the completion of $\mathcal{B}_X$ under $\mu$. Then $(X, \mathcal{B}_\mu, T)$ is a Lebesgue system. Let $\mathcal{P}_\mu(T)$ be the Pinsker $\sigma$-algebra of $(X, \mathcal{B}_\mu, \mu, T)$. It is well known that $\mathcal{P}_\mu(T)$ is $T$-invariant and $h_\mu(T) = h_\mu(T|\mathcal{P}_\mu(T))$.

If $\mu = \int_X \mu_x^P \, d\mu(x)$ is the disintegration of $\mu$ over $\mathcal{P}_\mu(T)$, applying Theorem 1.2 we have $h_\mu(T|\mathcal{P}_\mu(T)) = \int \mu_x^P(T) \, d\mu(x)$.

Suppose that $h_\mu(T) = 0$, then we know for $\mu$-a.e. $x \in X$ $\mathcal{B}_\mu^P(T) = 0 \leq h_{top}^B(T, W^s(x, T) \cap W^u(x, T))$.

Hence in the following we assume that $h_\mu(T) > 0$. By Lemma 4.2, for $\mu$-a.e. $x \in X$ $\text{supp}(\mu_x^P) \subseteq W^s(x, T) \cap W^u(x, T)$.

Applying Theorem 4.1 and Proposition 1, we have $h_{top}^B(T, \text{supp}(\mu_x^P)) \leq h_{top}^B(T, W^s(x, T) \cap W^u(x, T))$ for $\mu$-a.e. $x \in X$.

Hence we obtain $h_\mu(T) = h_\mu(T|\mathcal{P}_\mu(T)) = \int \mu_x^P(T) \, d\mu(x) \leq \text{essup}_{x \in X} h_{top}^B(T, W^s(x, T) \cap W^u(x, T))$.

Therefore by the variational principle of entropy, $h_{top}(T) \leq \sup_{x \in X} h_{top}^B(T, W^s(x, T) \cap W^u(x, T))$.

On the other hand, by Proposition 1 we have $\sup_{x \in X} h_{top}^B(T, W^s(x, T) \cap W^u(x, T)) \leq h_{top}^B(T, X) = h_{top}(T)$.

This completes the proof. \hfill $\square$

Acknowledgments. I would like to thank my supervisor Professor Wen Huang, Professor De-Jun Feng and Doctor Xiaoyao Zhou for valuable remarks and discussions. I also thank the anonymous referee for his helpful comments and suggestions that improved the manuscript.

REFERENCES

[1] R. L. Adler, A. G. Konheim and M. H. McAndrew, Topological entropy, Trans. Amer. Math. Soc., 114 (1965), 309–319.
[2] T. Bogenschütz, Entropy, pressure, and a variational principle for random dynamical systems, Random Comput. Dynam., 1 (1992/93), 99–116.
[3] R. Bowen, Topological entropy for noncompact sets, Trans. Amer. Math. Soc., 184 (1973), 125–136.
[4] M. Brin and A. Katok, On local entropy, in Geometric Dynamics, Rio de Janeiro, (1981), in Lecture Notes in Math., Springer, Berlin, 1007 (1983), 30–38.
[5] E. I. Dinaburg, A correlation between topological entropy and metric entropy, (Russian) Dokl. Akad. Nauk SSSR, 190 (1970), 19–22.
[6] T. Downarowicz and J. Serafin, Fiber entropy and conditional variational principles in compact non-metrizable spaces, Fund. Math., 172 (2002), 217–247.
[7] M. Einsiedler and T. Ward, Ergodic Theory with a View Towards Number Theory, Graduate Texts in Mathematics, 259, Springer-Verlag London, 2011.
A FORMULA OF CONDITIONAL ENTROPY AND SOME APPLICATIONS

[8] C. Fang, W. Huang, Y. Yi and P. Zhang, Dimensions of stable sets and scrambled sets in positive finite entropy systems, *Ergodic Theory Dynam. Systems*, 32 (2012), 599–628.

[9] D. Feng and W. Huang, Variational principles for topological entropies of subsets, *J. Funct. Anal.*, 263 (2012), 2228–2254.

[10] E. Glasner, *Ergodic Theory Via Joinings*, Mathematical Surveys and Monographs, 101, American Mathematical Society, Providence, RI, 2003.

[11] T. N. T. Goodman, Relating topological entropy and measure entropy, *Bull. London Math. Soc.*, 3 (1971), 176–180.

[12] L. W. Goodwyn, Topological entropy bounds measure-theoretic entropy, *Proc. Amer. Math. Soc.*, 23 (1969), 679–688.

[13] W. Huang, Stable sets and $\epsilon$-stable sets in positive-entropy systems, *Comm. Math. Phys.*, 279 (2008), 535–557.

[14] W. Huang, J. Li and X.D. Ye, Stable sets and mean Li-Yorke chaos in positive entropy systems, *J. Funct. Anal.*, 266 (2014), 3377–3394.

[15] F. Ledrappier and P. Walters, A relativised variational principle for continuous transformations, *J. London Math. Soc. (2)*, 16 (1977), 568–576.

[16] P. D. Liu, A note on the entropy of factors of random dynamical systems, *Ergodic Theory Dynam. Systems*, 25 (2005), 593–603.

[17] A. Katok, Lyapunov exponents, entropy and periodic orbits for diffeomorphisms, *Inst. Hautes Études Sci. Publ. Math.*, 51 (1980), 137–173.

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

[19] A. N. Kolmogorov, A new metric invariant of transient dynamical systems and automorphisms in Lebesgue spaces, (Russian) *Dokl. Akad. Nauk. SSSR (N.S.)*, 119 (1958), 861–864.

Received June 2015; revised November 2015.

E-mail address: zx12@mail.ustc.edu.cn