ENTROPY FORMULAE OF CONDITIONAL ENTROPY IN MEAN METRICS

PING HUANG
School of Mathematical Sciences and Institute of Mathematics
Nanjing Normal University
Nanjing 210023, Jiangsu, China

ERCAI CHEN*
School of Mathematical Sciences and Institute of Mathematics
Nanjing Normal University
Nanjing 210023, Jiangsu, China
Center of Nonlinear Science, Nanjing University
Nanjing 210093, Jiangsu, China

CHENWEI WANG
College of Mathematical and Physical Sciences, Taizhou University
Taizhou 225300, Jiangsu, China

(Communicated by Xiangdong Ye)

Abstract. In this paper, we construct the Brin-Katok formula of conditional entropy for invariant measures of continuous maps on a compact metric space by replacing the Bowen metrics with the corresponding mean metrics. Additionally, this paper is also devoted to establishing the Katok’s entropy formula of conditional entropy for ergodic measures in the case of mean metrics.

1. Introduction. Let a triple \((X, d, T)\) (or pair \((X, T)\) for short) be a topological dynamical system (TDS for short) in the sense that \(T : X \to X\) is a continuous map on the compact metric space \(X\) with metric \(d\). The terms \(M(X)\), \(M(X, T)\) and \(E(X, T)\) represent the sets of all Borel probability measures, \(T\)-invariant Borel probability measures, and \(T\)-invariant ergodic Borel probability measures, respectively.

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

\[
d_n(x, y) = \max\{d(T^i(x), T^i(y)) : i = 0, 1, \ldots, n-1\}.
\]

Given \(\varepsilon > 0\), let \(B_{d_n}(x, \varepsilon) = \{y \in X : d_n(x, y) < \varepsilon\}\) denote the \(d_n\)-ball centered at \(x\) with radius \(\varepsilon\). We also write \(B_n(x, \varepsilon)\) for convenience, when there is no confusion.

In classical ergodic theory, measure-theoretic entropy and topological entropy are important determinants of complexity in dynamical system. The relationship between these two quantities is the well-known variational principle. Brin-Katok

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

\[\text{Key words and phrases. Conditional entropy, Brin-Katok formula, Katok’ entropy formula, mean metrics, measure decomposition.}
\]

The second author was supported by NNSF of China (11671208, 11431012). The third author was supported by NNSF of China (11401581).

* Corresponding author: Ercai Chen.
In 1983, Brin and Katok [1] introduced the Brin-Katok formula: suppose \( \mu \in M(X, T) \), for \( \mu \)-a.e. \( x \in X \),

\[
\lim_{\varepsilon \to 0} \lim_{n \to \infty} \frac{- \log \mu(B_n(x, \varepsilon))}{n} = \lim_{\varepsilon \to 0} \limsup_{n \to \infty} \frac{- \log \mu(B_n(x, \varepsilon))}{n} = h_\mu(T, x),
\]

where \( h_\mu(T, x) \) is \( T \)-invariant, \( \int h_\mu(T, x) d\mu = h_\mu(T) \), and \( h_\mu(T) \) is the measure theoretic entropy. Particularly, if \( T \) is ergodic with respect to \( \mu \), then for \( \mu \)-a.e. \( x \in X \), \( h_\mu(T, x) = h_\mu(T) \).

It was shown in [19, 20] that the Brin-Katok formula above in the case of random dynamical systems. The Brin-Katok formula for the measure theoretic \( r \)-entropy was constructed in [17].

Given a TDS \( (X, d, T) \). Let \( B_X \) be the Borel \( \sigma \)-algebra of \( X \). Then each \( \mu \in M(X, T) \) induces a measure preserving dynamical system \( (X, B_X, \mu, T) \). Let \( B_\mu \) be the completion of \( B_X \), and \( I = \{ E \in B_\mu : T^{-1}E = E \} \) is the \( \sigma \)-algebra of \( T \)-invariant sets of \( B_\mu \). Then \( \mu \) can be decomposed into a generalized combination of ergodic measures \( \mu = \int \mu_x^T d\mu(x) \), where \( \mu_x^T \) denotes the conditional measure of \( \mu \) at \( x \) with respect to \( I \). There is a well-known result that the ergodic decomposition of entropy [6], i.e.

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

where \( h_\mu(T | I) \) denotes the conditional entropy with respect of \( I \) (see section 2 for details).

Let \( A \) be a \( T \)-invariant sub-\( \sigma \)-algebra of \( B_\mu \), i.e.

\[
T^{-1}A = A \mod \mu,
\]

and \( \mu = \int \mu_x^T d\mu(x) \) is the disintegration of \( \mu \) over \( A \), where \( \mu_x^A \) denotes the conditional measure of \( \mu \) at \( x \) with respect to \( A \). Nevertheless, \( \mu_x^A \) is only Borel probability measure but maybe not \( T \)-invariant for \( \mu \)-a.e. \( x \in X \). Furthermore, for any \( f \in L^1(X, B_X, \mu) \), the following equation holds [4]

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

Particularly, for any \( B \in B_X \), we have

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

In 2016, Zhou [18] established a conditional version of Brin-Katok formula and obtained an analogous conclusion with (1) for \( A \).

In 1980, Katok [11] introduced the Katok’s entropy formula: for any \( \mu \in E(X, T) \), \( 0 < \delta < 1 \),

\[
\lim_{\varepsilon \to 0} \limsup_{n \to \infty} \frac{\log N_\mu(n, \varepsilon, \delta)}{n} = \lim_{\varepsilon \to 0} \liminf_{n \to \infty} \frac{\log N_\mu(n, \varepsilon, \delta)}{n} = h_\mu(T),
\]

where \( N_\mu(n, \varepsilon, \delta) \) denotes the minimal number of \( d_n \)-balls with radius \( \varepsilon \) which cover the set of \( \mu \)-measure more than or equal to \( 1 - \delta \). Katok’s entropy formula is an equivalent definition of the measure-theoretic entropy in a manner analogous to the definition of the topological entropy.

In 2004, using spanning sets, He, Lv and Zhou [9] introduced a definition of measure-theoretic pressure of additive potentials for ergodic measures, and obtained
a pressure version of Katok’s entropy formula. In 2009, Zhao and Cao [15] gave a
definition of measure-theoretic pressure of sub-additive potentials for ergodic mea-
ures, and generalized the above results in [11] and [9]. Moreover, we refer to [3, 2]
for more pressure versions of Katok’s entropy formula. In 2009, Zhu [20] established
Katok’s entropy formula in the case of random dynamical systems. Very recently,
in order to establish large deviations bounds for countable discrete amenable group
actions, Zheng, Chen and Yang [16] introduced an amenable version of Katok’s
entropy formula.

All of the above studies were carried out in the case of Bowen metrics. However,
in this paper, we consider the mean metrics. For any \( x, y \in X, \ n \in \mathbb{N} \), the mean
metric \( \tilde{d}_n \) [8] is given by

\[
\tilde{d}_n(x, y) := \frac{1}{n} \sum_{i=0}^{n-1} d(T^i x, T^i y).
\]

For \( x \in X, \ \varepsilon > 0 \), let

\[
B_{\tilde{d}_n}(x, \varepsilon) := \{ y \in X : \tilde{d}_n(x, y) < \varepsilon \}
\]
denote the \( \tilde{d}_n \)-ball centered at \( x \) with radius \( \varepsilon \).

Recently, the mean metrics attract a lot of attention. In 2016, using separated
sets, Gröger and Jäger [7] gave a definition of topological entropy of the whole system
in mean metrics, and proved that the topological entropy defined by mean metrics is
equivalent to the topological entropy defined by Bowen metrics. Huang, Wang and
Ye [10] introduced the notion of measure complexity for a TDS. They established
Katok’s entropy formula for ergodic measures in the case of mean metrics and this
result is exactly the corollary of our main results (see Corollary 2). Meanwhile, this
formula showed that one can replace the \( d_n \)-balls with \( \tilde{d}_n \)-balls when studying the
measure complexity for a TDS. Since the advantage to use mean metrics is that it
is an isomorphic invariant [10], they discussed the measure complexity for a TDS
by mean metrics and showed that Sarnak’s Möbius disjointness conjecture holds for
any system for which every invariant Borel probability measure has sub-polynomial
measure complexity.

In this paper, inspired by the ideas of Brin and Katok [1], Zhou, Zhou and Chen
[17] and Zhou [18], we construct the Brin-Katok formula of conditional entropy
\( h_\mu(T|A) \) for invariant measures of continuous maps on a compact metric space by
replacing the Bowen metrics with the corresponding mean metrics. Furthermore,
following the idea of Katok [11], we establish the Katok’s entropy formula of con-
ditional entropy \( h_\mu(T|A) \) for ergodic measures in the case of mean metrics.

The following theorems present the main results of this paper.

**Theorem 1.1.** (Brin-Katok formula of conditional entropy in mean metrics)

Let \( (X, d, T) \) be a TDS, \( \mu \in \mathcal{M}(X, T) \) and \( \mathcal{B}_\mu \) be the completion of \( \mathcal{B}_X \)
under \( \mu \). Suppose \( A \) is a T-invariant sub-\( \sigma \)-algebra of \( \mathcal{B}_\mu \), i.e. \( T^{-1}A = A \ (mod \ \mu) \) and the
measure disintegration of \( \mu \) over \( A \) is

\[
\mu = \int \mu_A^x \, d\mu(x).
\]

Then, for \( \mu \)-a.e. \( x \in X \),

\[
\tilde{h}_\mu(T) = \tilde{h}_\mu(T),
\]
and
\[ h_\mu(T|A) = \int \tilde{h}_{\mu^A}(T) \, d\mu(x) = \int \tilde{h}_{\mu^A}(T) \, d\mu(x). \]

Particularly, if \( \mu \in E(X,T) \) then for \( \mu \)-a.e. \( x \in X \), \( h_\mu(T|A) = \tilde{h}_{\mu^A}(T) = \tilde{h}_{\mu^A}(T) \).

See Definition 2.3 in section 2 for the definitions of \( \tilde{h}_{\mu^A}(T) \) and \( \tilde{h}_{\mu^A}(T) \).

**Theorem 1.2.** (Katok’s entropy formula of conditional entropy in mean metrics)

Let \( (X,d,T) \) be a TDS, \( \mu \in E(X,T) \) and \( B_\mu \) be the completion of \( B_X \) under \( \mu \).

Suppose \( A \) is a T-invariant sub-\( \sigma \)-algebra of \( B_\mu \) i.e. \( T^{-1}A = A \text{ (mod } \mu \text{)} \) and the measure disintegration of \( \mu \) over \( A \) is
\[ \mu = \int \mu^A_x \, d\mu(x). \]

Then for any \( 0 < \delta < 1 \), we have
\[ \lim_{\varepsilon \to 0} \liminf_{n \to \infty} \frac{\log \tilde{N}_\mu^A(n, \varepsilon, \delta)}{n} = \lim_{\varepsilon \to 0} \limsup_{n \to \infty} \frac{\log \tilde{N}_\mu^A(n, \varepsilon, \delta)}{n} = h_\mu(T|A) \]
for \( \mu \)-a.e. \( x \in X \), where \( \tilde{N}_\mu^A(n, \varepsilon, \delta) \) denotes the minimal number of \( \tilde{d}_n \)-balls with radius \( \varepsilon \) whose union has \( \mu^A_x \)-measure more than or equal to \( 1 - \delta \).

Let \( (X,d,T) \) be a TDS, \( B_X \) be the Borel \( \sigma \)-algebra of \( X \), and \( \mu \in M(X,T) \).

Suppose \( B_\mu \) is the completion of \( B_X \), and \( A = \{ \emptyset, X \} \). Clearly \( A \) is a T-invariant sub-\( \sigma \)-algebra of \( B_\mu \), and for each \( x \in X \), \( \mu^A_x = \mu \), where \( \mu^A_x \) denotes the conditional measure of \( \mu \) at \( x \) with respect to \( A \). Also, we observe that \( h_\mu(T|A) = h_\mu(T) \), where \( h_\mu(T) \) is the measure theoretic entropy. Obviously, we have the following two corollaries of Theorem 1.1 and Theorem 1.2, respectively.

**Corollary 1.** (Brin-Katok formula in mean metrics)

Let \( (X,d,T) \) be a TDS. Suppose \( \mu \in M(X,T) \). Then for \( \mu \)-almost every \( x \in X \), we have
\[ \lim_{\varepsilon \to 0} \liminf_{n \to \infty} \frac{-\log \mu(B_{\tilde{d}_n}(x, \varepsilon))}{n} = \lim_{\varepsilon \to 0} \limsup_{n \to \infty} \frac{-\log \mu(B_{\tilde{d}_n}(x, \varepsilon))}{n} = \tilde{h}_\mu(T,x), \]
where \( \int \tilde{h}_\mu(T,x) \, d\mu = h_\mu(T) \). Particularly, if \( \mu \in E(X,T) \), then for \( \mu \)-a.e. \( x \in X \), \( \tilde{h}_\mu(T,x) = h_\mu(T) \).

**Corollary 2.** (see [10], Theorem 2.4 ) (Katok’s entropy formula in mean metrics)

Let \( (X,d,T) \) be a TDS. For each \( \mu \in E(X,T) \), \( 0 < \delta < 1 \), we have
\[ \lim_{\varepsilon \to 0} \liminf_{n \to \infty} \frac{\log \tilde{N}_\mu(n, \varepsilon, \delta)}{n} = \lim_{\varepsilon \to 0} \limsup_{n \to \infty} \frac{\log \tilde{N}_\mu(n, \varepsilon, \delta)}{n} = h_\mu(T), \]
where \( \tilde{N}_\mu(n, \varepsilon, \delta) \) denotes the minimal number of \( \tilde{d}_n \)-balls with radius \( \varepsilon \) whose union has \( \mu \)-measure more than or equal to \( 1 - \delta \).

The remainder of this paper is organized as follows. Section 2 gives some preliminaries. Section 3 provides the proofs of the main results. Finally, in section 4, we consider the one-sided symbolic space as an example of Theorem 1.1.
2. Preliminaries. Given a TDS \((X,T)\). A partition of \(X\) is a disjoint collection of elements of \(\mathcal{B}_X\) whose union is \(X\). Let \(\mathcal{P}_X\) denote the collection of all finite Borel partitions of \(X\). Suppose \(\xi, \eta \in \mathcal{P}_X\). We write \(\xi \leq \eta\) to mean that each element of \(\xi\) is a union of elements of \(\eta\) (i.e. \(\eta\) is a refinement of \(\xi\)). Let \(\xi = \{A_1, \cdots, A_k\}, \eta = \{B_1, \cdots, B_m\}\) be two finite partitions of \((X,T)\). Their join is the partition

\[
\xi \bigvee \eta = \{A_i \cap B_j : 1 \leq i \leq k, 1 \leq j \leq m\}.
\]

For a measurable partition \(\xi\) of \(X\) and \(x \in X\), denote by \(\xi(x)\) the element of \(\xi\) containing \(x\), and set

\[
\xi_0^{n-1} = \xi \bigvee T^{-1} \xi \bigvee \cdots \bigvee T^{-(n-1)} \xi.
\]

Let \(\mu \in M(X,T)\). Given a \(T\)-invariant sub-\(\sigma\)-algebra \(\mathcal{A}\) of \(\mathcal{B}_\mu\) and \(\xi \in \mathcal{P}_X\). Define (see \([14]\), Definition 4.8)

\[
H_\mu(\xi|\mathcal{A}) = -\int \sum_{B \in \xi} E(\chi_B|\mathcal{A}) \log E(\chi_B|\mathcal{A}) \, d\mu,
\]

where \(\chi_B\) is the characteristic function of set \(B\) and \(E(\chi_B|\mathcal{A})\) is the conditional expectation of \(\chi_B\) with respect to \(\mathcal{A}\). Since \(H_\mu(\bigvee_{i=0}^{n-1} T^{-i} \xi|\mathcal{A})\) is a non-additive sequence, the conditional entropy of \(\xi\) with respect to \(\mathcal{A}\) is given by

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

Furthermore, the conditional entropy of \(T\) with respect to \(\mathcal{A}\) is defined by

\[
h_\mu(T|\mathcal{A}) = \sup_{\xi \in \mathcal{P}_X} h_\mu(T,\xi|\mathcal{A}).
\]

If \(\{\xi_i\}_{i=1}^\infty\) is a family of finite Borel partitions of \(X\) satisfying \(\xi_1 \leq \xi_2 \leq \xi_3 \leq \cdots\) and \(\text{diam}(\xi_i) \to 0\) as \(i \to \infty\), where \(\text{diam}(\xi_i) = \max\{\text{diam}(B), B \in \xi_i\}\). Then the conditional entropy of \(T\) with respect to \(\mathcal{A}\) can be computed by

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

The following theorem and lemma will be used in proving the main results.

**Theorem 2.1.** \([18]\) Let \((X,T)\) be a TDS, \(\mu \in M(X,T)\), \(\xi \in \mathcal{P}_X\) and \(\mathcal{A}\) be a \(T\)-invariant sub-\(\sigma\)-algebra of \(\mathcal{B}_\mu\). Then there exists a \(T\)-invariant function \(h_\mu(\xi|\mathcal{A},x) \in L^1(\mu)\) such that

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

and

\[
\lim_{n \to \infty} \frac{1}{n} I_\mu(\xi_0^{n-1}|\mathcal{A})(x) = h_\mu(\xi|\mathcal{A},x),
\]

for \(\mu\)-a.e. \(x \in X\) and in \(L^1(\mu)\), where

\[
I_\mu(\xi_0^{n-1}|\mathcal{A})(x) = -\sum_{B \in \xi_0^{n-1}} \chi_B(x) \log E(\chi_B|\mathcal{A})(x)
\]

is the conditional informational function of \(\xi_0^{n-1}\) with respect to \(\mathcal{A}\). Moreover, if \(\mu \in E(X,T)\) then \(h_\mu(\xi|\mathcal{A},x) = h_\mu(T,\xi|\mathcal{A})\) for \(\mu\)-a.e. \(x \in X\).
Lemma 2.2. [18] Let \((X,T)\) be a TDS, \(\mu \in M(X,T)\), \(\xi \in \mathcal{P}_X\) and \(\mathcal{A}\) be a \(T\)-invariant sub-\(\sigma\)-algebra of \(\mathcal{B}_\mu\). Then for \(\mu\text{-a.e. } x \in X\), there exists \(W_x \in \mathcal{B}_X\) with \(\mu_x^x(W_x) = 1\) such that
\[
\lim_{n \to \infty} \frac{- \log \mu_x^x(\xi_0^{n-1}(y))}{n} = h_\mu(\xi,\mathcal{A},x)
\]
for each \(y \in W_x\), where \(\mu = \int \mu_x^x d\mu(x)\) is the measure disintegration of \(\mu\) over \(\mathcal{A}\) and \(h_\mu(\xi,\mathcal{A},x)\) is the function obtained in Theorem 2.1.

Following the ideas of Brin and Katok [1] as well as Feng and Huang [5], we replace the Bowen metric \(d_n\) with the mean metric \(\tilde{d}_n\) and give a definition of measure-theoretic lower and upper entropies.

Definition 2.3. Let \((X,d,T)\) be a TDS and \(\nu \in M(X)\). Set
\[
\tilde{h}_\nu(T,y) = \lim_{\varepsilon \to 0} \liminf_{n \to \infty} \frac{- \log \nu(B_{\tilde{d}_n}(y,\varepsilon))}{n},
\]
\[
\tilde{h}_\nu(T,y) = \lim_{\varepsilon \to 0} \limsup_{n \to \infty} \frac{- \log \nu(B_{\tilde{d}_n}(y,\varepsilon))}{n}.
\]
Then, the measure-theoretical lower and upper entropies of \(\nu\) in mean metrics are defined respectively by
\[
\tilde{h}_\nu(T) = \int \tilde{h}_\nu(T,y) d\nu(y),
\]
\[
\tilde{h}_\nu(T) = \int \tilde{h}_\nu(T,y) d\nu(y).
\]

In order to prove the main theorems, we need the relationship between the \(\tilde{d}_n\)-ball \(B_{\tilde{d}_n}(x,\varepsilon)\) and the mistake dynamical ball \(B_n(g;x,\varepsilon)\), where the mistake function \(g\) and mistake dynamical ball \(B_n(g;x,\varepsilon)\) are defined in Definition 2.4 and Definition 2.5, respectively.

Definition 2.4. Define a function \(g : \mathbb{N} \times (0,1] \to \mathbb{R}\). For any \(\varepsilon \in (0,1]\), \(n \in \mathbb{N}\),
\[
g(n,\varepsilon) := n\varepsilon.
\]
The function \(g\) is called a mistake function. Note that \(g\) is a special mistake function and the definition of the function \(g\) is different from the definition of the mistake function in [12] or [13]. In fact, the mistake function \(\hat{g}(n)\) in [12] should satisfy \(\lim_{n \to \infty} \hat{g}(n) = 0\), while the mistake function \(g'(n,\varepsilon)\) in [13] should satisfy \(\lim_{n \to \infty} g'(n,\varepsilon) = 0\). However, since in this paper \(\lim_{n \to \infty} \frac{g(n,\varepsilon)}{n} = \varepsilon\), we only have \(\lim_{\varepsilon \to 0} \lim_{n \to \infty} \frac{g(n,\varepsilon)}{n} = 0\).

Definition 2.5. Let \(\Lambda_n = \{0,1,\ldots,n-1\}\), \(0 < \varepsilon < 1\) and \(n \geq 1\). The mistake dynamical ball \(B_n(g;x,\varepsilon)\) centered at \(x\) with radius \(\varepsilon\) and length \(n\) associated to function \(g\) is defined as follows:
\[
B_n(g;x,\varepsilon) = \left\{ y \in X : \# \{ i \in \Lambda_n : d(T^ix,T^iy) < \varepsilon \} > n - g(n,\varepsilon) \right\}
\]
\[
= \left\{ y \in X : \# \left\{ i \in \Lambda_n : d(T^ix,T^iy) < \varepsilon \right\} > 1 - \varepsilon \right\},
\]
where \(#A\) denotes the cardinality of the set \(A\).

Then, we can have the following relationship which plays a crucial role in our proofs.
Lemma 2.6. For any \( x \in X, n \in \mathbb{N}^+ \) and \( 0 < \varepsilon < 1 \), we have
\[
B_n(x, \varepsilon) \subset B_{d_n}(x, \varepsilon) \subset B_n(g; x, \sqrt{\varepsilon}).
\]

Proof. Given \( y \in X \). If \( d_n(x, y) < \varepsilon \), then \( \tilde{d}_n(x, y) < \varepsilon \). So \( B_n(x, \varepsilon) \subset B_{\tilde{d}_n}(x, \varepsilon) \).
Set
\[
I_1 = \{ i \in \Lambda_n : d(T^i x, T^i y) < \sqrt{\varepsilon} \},
I_2 = \{ i \in \Lambda_n : d(T^i x, T^i y) \geq \sqrt{\varepsilon} \}.
\]
Noting that
\[
\frac{1}{n} \sum_{i=0}^{n-1} d(T^i x, T^i y) = \frac{\sum_{i \in I_1} d(T^i x, T^i y) + \sum_{i \in I_2} d(T^i x, T^i y)}{n} \geq \frac{\sum_{i \in I_1} d(T^i x, T^i y) + \sqrt{\varepsilon} \# I_2}{n} \geq \frac{\sqrt{\varepsilon}}{n} \# I_2,
\]
if \( y \in B_{d_n}(x, \varepsilon) \), i.e. \( \tilde{d}_n(x, y) < \varepsilon \), we have
\[
\# I_2 < n \sqrt{\varepsilon}.
\]
Thus, we obtain
\[
\frac{\# \{ i \in \Lambda_n : d(T^i x, T^i y) \geq \sqrt{\varepsilon} \}}{n} < \sqrt{\varepsilon},
\]
i.e.
\[
\frac{\# \{ i \in \Lambda_n : d(T^i x, T^i y) < \sqrt{\varepsilon} \}}{n} > 1 - \sqrt{\varepsilon},
\]
then \( y \in B_n(g; x, \sqrt{\varepsilon}) \). Therefore, we have \( B_{d_n}(x, \varepsilon) \subset B_n(g; x, \sqrt{\varepsilon}) \). \( \square \)

3. Proofs of main results. This section gives the proofs of the main results. Firstly, we give the proof of Theorem 1.1 in Subsection 3.1. Secondly, we show the proof of Theorem 1.2 in Subsection 3.2.

3.1. Proof of Theorem 1.1.

Proof. Let \( \{ \xi_i \}_{i=0}^\infty \) be a family of finite Borel partitions of \( X \) satisfying
\begin{itemize}
  \item \( \xi_1 \leq \xi_2 \leq \cdots \),
  \item \( \text{diam}(\xi_i) \to 0 \), as \( i \to \infty \),
  \item \( \mu(\partial \xi_i) = 0 \), for \( i = 1, 2, \cdots \),
\end{itemize}
where \( \partial \xi_i \) denotes the union of the boundaries \( \partial B \) of all elements \( B \in \xi_i \). In fact, using the Monotone Convergence Theorem it is sufficient to show that the following equation holds
\[
\bar{h}_{\mu A}(T) = \bar{h}_{\mu A}(T) = \sup_{i \geq 1} h_{\mu}(\xi_i|A, x)
\]
for \( \mu \)-a.e. \( x \in X \), where \( h_{\mu}(\xi_i|A, x) \) is the function obtained in Theorem 2.1.
(1) Firstly, we want to show that for \( \mu \)-a.e. \( x \in X \),
\[
\bar{h}_{\mu A}(T) \leq \sup_{i \geq 1} h_{\mu}(\xi_i|A, x).
\]
For any $\varepsilon > 0$, since $\operatorname{diam}(\xi_i) \rightarrow 0$, there exists $N = N(\varepsilon) \in \mathbb{N}$ such that $\operatorname{diam}(\xi_i) < \varepsilon$ when $i \geq N$. Then, by Lemma 2.6, when $i \geq N$, for each $y \in X$ and $n \geq 1$, we have $(\xi_i)_{\theta > 0}^{-1}(y) \subset B_n(y, \varepsilon) \subset B_{\tilde{d}_n}(y, \varepsilon)$. Hence

$$
\overline{\mu}^A_{\alpha}(T, y) = \lim_{n \to \infty} \limsup_{\varepsilon \to 0} -\frac{\log \mu^A_{\alpha}(B_{\tilde{d}_n}(y, \varepsilon))}{n}
= \sup_{\varepsilon > 0} \limsup_{n \to \infty} -\frac{\log \mu^A_{\alpha}(B_{\tilde{d}_n}(y, \varepsilon))}{n}
\leq \sup_{i \geq 1} \limsup_{n \to \infty} -\frac{\log \mu^A_{\alpha}((\xi_i)_{\theta > 0}^{-1}(y))}{n}.
$$

Therefore, according to Lemma 2.2, for $\mu$-a.e. $x \in X$,

$$
\overline{\mu}^A_{\alpha}(T) = \int \overline{\mu}^A_{\alpha}(T, y) d\mu^A_{\alpha}(y)
\leq \int \sup_{i \geq 1} \limsup_{n \to \infty} -\frac{\log \mu^A_{\alpha}((\xi_i)_{\theta > 0}^{-1}(y))}{n} d\mu^A_{\alpha}(y)
= \int \sup_{i \geq 1} \mu^A_{\alpha}((\xi_i)_{\theta > 0}^{-1}(y)) d\mu^A_{\alpha}(y)
= \sup_{i \geq 1} \mu^A_{\alpha}((\xi_i)_{\theta > 0}^{-1}(y)).
$$

(2) Secondly, we turn to prove that for $\mu$-a.e. $x \in X$,

$$
\overline{\mu}^A_{\alpha}(T) \geq \sup_{i \geq 1} \mu^A_{\alpha}((\xi_i)_{\theta > 0}^{-1}(y)).
$$

(i) Let $\xi \in \mathcal{P}_X$ with $\mu(\partial \xi) = 0$ and $0 < \varepsilon < \frac{1}{4\pi}$. For $\theta > 0$, let

$$
U_\theta(\xi) = \{x \in X : \text{the ball } B(x, \theta) \text{ is not contained in } \xi(x)\},
$$

where $\xi(x)$ denotes the element of the partition $\xi$ containing $x$. Since $\cap_{\theta > 0} U_\theta(\xi) = \partial \xi$, we have

$$
\mu(U_\theta(\xi)) \rightarrow 0, \text{ as } \theta \rightarrow 0.
$$

Therefore, there exists $0 < \delta < \varepsilon$ such that $\mu(U_\theta(\xi)) \leq \varepsilon$ for any $0 < \theta \leq \delta$.

For $l \in \mathbb{N}^+$, we define

$$
A_l = \left\{ y \in X : \frac{1}{k} \sum_{i=0}^{k-1} \chi_{U_i(\xi)}(T^i(y)) \leq 2\sqrt{\varepsilon} \text{ for any } k \geq l \right\},
$$

where $\chi_{U_i(\xi)}$ is the characteristic function of the set $U_i(\xi)$. Therefore, by the same method of the proof of Theorem 1.2 in [18], we can choose $N_1 \in \mathbb{N}$ such that for any $l \geq N_1$, we have

$$
\mu(A_l) > 1 - 2\sqrt{\varepsilon}.
$$

Define $Q_l = \{x \in X : \mu^A_{\alpha}(A_l) \geq 1 - 2\varepsilon^{\frac{1}{4}}\}$. Then also by the same method of the proof of Theorem 1.2 in [18], we have for any $l \geq N_1$, $\mu(Q_l) > 1 - \varepsilon^{\frac{1}{4}}$. Clearly, the sets $A_l$ are nested, i.e. $A_1 \subset A_2 \subset \cdots$. Then fix some $l_1 > N_1$, for any $x \in Q_{l_1}$, $l \geq l_1$ we have

$$
\mu^A_{\alpha}(A_l) \geq \mu^A_{\alpha}(A_{l_1}) \geq 1 - 2\varepsilon^{\frac{1}{4}}.
$$

According to Lemma 2.2, we can find a subset $X_1 \subset X$ with $\mu(X_1) = 1$ such that for any $x \in X_1$, there exists $W_x \in \mathcal{B}_X$ with $\mu^A_{\alpha}(W_x) = 1$ such that for any $y \in W_x$,

$$
\lim_{n \to \infty} -\frac{\log \mu^A_{\alpha}((\xi_i)_{\theta > 0}^{-1}(y))}{n} = \mu^A_{\alpha}(\xi_i | A, x).
$$
Let $I = X_1 \cap Q_{l_1}$. Clearly, $\mu(I) > 1 - \varepsilon^\frac{1}{3}$. Fix $\hat{x} \in I$. By the Egorov Theorem, we can find a number $l_2$ large enough such that $\mu_\delta^\xi(B_l) \geq 1 - \varepsilon^\frac{1}{3}$ for any $l \geq l_2$, where

$$B_l = \left\{ y \in W_{\hat{x}} : -\log \frac{\mu_\delta^\xi(\xi_0^{n-1}(y))}{n} \geq h_\mu(\xi|A, \hat{x}) - \varepsilon, \text{ for any } n \geq l \right\}.$$  

Fix $l \geq \max\{l_1, l_2\}$. Let $E = A_1 \cap B_l$. Obviously, $\mu_\delta^\xi(E) \geq 1 - 3\varepsilon^\frac{1}{3}$.

(ii) For $n \in \mathbb{N}$ and given a point $y \in X$, we call the collection

$$C(n, y) := (\xi(y), \xi(T(y)), \ldots, \xi(T^{n-1}(y)))$$

the $(\xi, n)$-name of $y$. Since each point in one element $V$ of $\xi_0^{n-1}$ has the same $(\xi, n)$-name, we can define

$$C(n, V) := C(n, y)$$

for any $y \in V$, which is called the $(\xi, n)$-name of $V$.

For $n \in \mathbb{N}$ and $\xi$, we give a metric $d^\xi_n$ between $(\xi, n)$-names of $y$ and $z$ as follows:

$$d^\xi_n(C(n, y), C(n, z)) = \frac{1}{n} \#\{0 \leq i \leq n - 1 : \xi(T^i(y)) \neq \xi(T^i(z))\}.$$ 

It can also be viewed as a semi-metric on $X$. If $z \in B(y, \delta)$, then either $y$ and $z$ belong to the same element of $\xi$ or $y \in U_\delta(\xi), z \notin \xi(y)$. Noting that $B_{d_\delta}(y, \delta^2) \subseteq B_n(y, \delta)$, hence if $y \in E, n > l$ and $z \in B_{d_\delta}(y, \delta^2)$, the distance $d^\xi_n$ between $(\xi, n)$-names of $y$ and $z$ does not exceed $\delta + 2\sqrt{\varepsilon}$, i.e.

$$d^\xi_n(C(n, y), C(n, z)) \leq \delta + 2\sqrt{\varepsilon}.$$ 

Furthermore, for $y \in E$ and $n > l$, $B_{d_\delta}(y, \delta^2)$ is contained in the set of points $z$ whose $(\xi, n)$-names are $(\delta + 2\sqrt{\varepsilon})$-close to the $(\xi, n)$-name of $y$, i.e.

$$B_{d_\delta}(y, \delta^2) \subseteq B_{d_\delta}(y, \delta + 2\sqrt{\varepsilon}) = \bigcup_{V \in \mathcal{M}_n} V,$$

where

$$\mathcal{M}_n = \{ V \in \xi_0^{n-1} : C(n, V) \text{ is } (\delta + 2\sqrt{\varepsilon}) - \text{close to } C(n, y) \}.$$  

By Stirling’s formula, there exists a large number $l_3 \in \mathbb{N}$ and for any $n \geq l_3$, it can be shown that the total number $L_n$ of such $(\xi, n)$-names consisting of $B_{d_\delta}(y, \delta + 2\sqrt{\varepsilon})$ admits the following estimate:

$$L_n = \#\mathcal{M}_n \leq \sum_{j=0}^{[n(\delta + 2\sqrt{\varepsilon})]} C_n^j(\#\xi - 1)^j \leq \sum_{j=0}^{[n(\delta + 2\sqrt{\varepsilon})]} C_n^j(\#\xi)^j \leq \exp\{(\delta + 2\sqrt{\varepsilon} + \Delta)n\},$$

where $[n(\delta + 2\sqrt{\varepsilon})]$ denotes the largest integer no larger than $n(\delta + 2\sqrt{\varepsilon})$, and $\Delta = (\delta + 2\sqrt{\varepsilon}) \log(\#\xi) - (\delta + 2\sqrt{\varepsilon}) \log(\delta + 2\sqrt{\varepsilon}) - (1 - \delta - 2\sqrt{\varepsilon}) \log(1 - \delta - 2\sqrt{\varepsilon}) - \log(\#\xi)$.

More precisely, we have shown that for any $y \in E$, $n \geq \max\{l, l_3\}$,

$$B_{d_\delta}(y, \delta^2) \subseteq B_{d_\delta}(y, \delta + 2\sqrt{\varepsilon}) = \bigcup_{V \in \mathcal{M}_n} V,$$  

$$L_n = \#\mathcal{M}_n \leq \exp\{(\delta + 2\sqrt{\varepsilon} + \Delta)n\}. \tag{2}$$

(iii) For $n \in \mathbb{N}$, let $E_n$ be the set of those points in $E$ whose $(\xi, n)$-names have an element of the partition $\xi_0^{n-1}$ of $\mu_\delta^\xi$-measure greater than $\exp\{-h_\mu(\xi|A, \hat{x}) + 2(\delta + 2\sqrt{\varepsilon} + \Delta)n\}$ in their $(\delta + 2\sqrt{\varepsilon})$-neighborhood in the semi-metric $d^\xi_n$. It is obvious that the total number of such elements does not exceed $\exp\{h_\mu(\xi|A, \hat{x}) -
2(\delta + 2\sqrt{\varepsilon} + \Delta)n}, since \(\mu^2_\varepsilon(X) = 1\). Denote by \(Q_n\) the set of all atoms \(B \in \xi_0^{-1}\) satisfying that
\[
d^2_\varepsilon(C(n, B), C(n, V)) \leq \delta + 2\sqrt{\varepsilon},
\]
for some \(V \in \xi_0^{-1}\) with \(\mu^2_\varepsilon(V) \geq \exp\{(h_\mu(\xi|A, \hat{x}) + 2(\delta + 2\sqrt{\varepsilon} + \Delta)n\} \). Clearly,
\[
#Q_n \leq \exp\{h_\mu(\xi|A, \hat{x}) - 2(\delta + 2\sqrt{\varepsilon} + \Delta)n\} \cdot L_n
\]
\[
\leq \exp\{h_\mu(\xi|A, \hat{x}) - (\delta + 2\sqrt{\varepsilon} + \Delta)n\}.
\]

Now for any \(n \geq \max\{l_1, l_3\}\), by the definition of \(E_n\), we have
\[
E_n \subset \bigcup_{B \in Q_n} B.
\]

For any \(y \in E_n\), since \(y \in B_l\),
\[
\mu^2_\varepsilon(\xi_0^{-1}(y)) \leq \exp\{(-h_\mu(\xi|A, \hat{x}) + \varepsilon)n\},
\]
we have
\[
\mu^2_\varepsilon(E_n) \leq \exp\{(-h_\mu(\xi|A, \hat{x}) + \varepsilon)n\} \cdot #Q_n
\]
\[
\leq \exp\{(-h_\mu(\xi|A, \hat{x}) + \varepsilon)n\} \cdot \exp\{h_\mu(\xi|A, \hat{x}) - (\delta + 2\sqrt{\varepsilon} + \Delta)n\}
\]
\[
= \exp\{\varepsilon - (\delta + 2\sqrt{\varepsilon} + \Delta)n\}
\]
\[
\leq \exp\{-\delta + \varepsilon\}. \quad (4)
\]

Therefore, putting (2), (3) and (4) together, we have
\[
\mu^2_\varepsilon(B_{d_\delta}(y, \delta^2)) \leq \mu^2_\varepsilon(B_{d_\delta}(y, \delta + 2\sqrt{\varepsilon}))
\]
\[
\leq \exp\{(-h_\mu(\xi|A, \hat{x}) + 2(\delta + 2\sqrt{\varepsilon} + \Delta)n\} \cdot L_n
\]
\[
\leq \exp\{(-h_\mu(\xi|A, \hat{x}) + 2(\delta + 2\sqrt{\varepsilon} + \Delta)n\} \cdot \exp\{(\delta + 2\sqrt{\varepsilon} + \Delta)n\}
\]
\[
= \exp\{(-h_\mu(\xi|A, \hat{x}) + 3(\delta + 2\sqrt{\varepsilon} + \Delta)n\}.
\]

Thus, for any \(y \in D\) and \(n \geq l_4\), we obtain
\[
-\frac{\log\mu^2_\varepsilon(B_{d_\delta}(y, \delta^2))}{n} \geq \frac{h_\mu(\xi|A, \hat{x}) - 3(\delta + 2\sqrt{\varepsilon} + \Delta)}{n}.
\]

Hence, for any \(y \in D\),
\[
\liminf_{n \to \infty} -\frac{\log\mu^2_\varepsilon(B_{d_\delta}(y, \delta^2))}{n} \geq h_\mu(\xi|A, \hat{x}) - 3(\delta + 2\sqrt{\varepsilon} + \Delta).
\]

Summing up (i)(ii)(iii), we had shown that for any \(\xi \in P_X\) with \(\mu(\partial\xi) = 0\) and \(0 < \varepsilon < \frac{1}{4}\), we can find a \(0 < \delta < \varepsilon\), a measurable subset \(I\) of \(X\) satisfying \(\mu(I) > 1 - \varepsilon^2\) such that for each \(\hat{x} \in I\), there exists a measurable subset \(D\) of \(X\) such that \(\mu^2_\varepsilon(D) > 1 - 4\varepsilon^2\) and for any \(y \in D\),
\[
\liminf_{n \to \infty} -\frac{\log\mu^2_\varepsilon(B_{d_\delta}(y, \delta^2))}{n} \geq h_\mu(\xi|A, \hat{x}) - 3(\delta + 2\sqrt{\varepsilon} + \Delta), \quad (5)
\]
where

\[ \Delta = (\delta + 2\sqrt{\varepsilon}) \log (\# \xi) - (\delta + 2\sqrt{\varepsilon}) \log (\delta + 2\sqrt{\varepsilon}) - (1 - \delta - 2\sqrt{\varepsilon}) \log (1 - \delta - 2\sqrt{\varepsilon}). \]

(iv) Now, let \( \varepsilon \to 0 \). Observe that \( 0 < \delta < \varepsilon \). Since \( \lim_{\varepsilon \to 0} \Delta = 0 \), \( \mu_x^A(D) > 1 - 4\varepsilon \), \( \hat{x} \in I \) and \( \mu(I) > 1 - \varepsilon^2 \). Hence by (5), for \( \mu \)-a.e. \( x \in X \), we have

\[ \lim_{\delta \to 0} \lim_{n \to \infty} \frac{-\log \mu_x^A(B_{\delta_n}(y, \delta^2))}{n} \geq h_\mu(\xi | A, x) \]

for \( \mu_x^A \)-a.e. \( y \in X \). Therefore, for \( \mu \)-a.e. \( x \in X \), this implies that

\[ \tilde{h}_{\mu_x^A}(T, y) = \lim_{\delta \to 0} \lim_{n \to \infty} \frac{-\log \mu_x^A(B_{\delta_n}(y, \delta))}{n} \geq h_\mu(\xi | A, x) \]

for \( \mu_x^A \)-a.e. \( y \in X \). Thus, for \( \mu \)-a.e. \( x \in X \),

\[ \tilde{h}_{\mu_x^A}(T) = \int \tilde{h}_{\mu_x^A}(T, y) d\mu_x^A(y) \geq h_\mu(\xi | A, x). \]

Furthermore, since \( \xi \in \mathcal{P}_X \) with \( \mu(\partial \xi) = 0 \) is arbitrary, we conclude that

\[ \tilde{h}_{\mu_x^A}(T) \geq \sup_{i \geq 1} h_\mu(\xi_i | A, x) \]

for \( \mu \)-a.e. \( x \in X \). \( \Box \)

3.2. Proof of Theorem 1.2.

Proof. (1) Firstly, we are going to show that for every \( 0 < \delta < 1 \), we have

\[ \lim_{\varepsilon \to 0} \limsup_{n \to \infty} \frac{\log \tilde{N}_{\mu_x^A}(n, \varepsilon, \delta)}{n} \leq h_\mu(T | A) \]

for \( \mu \)-a.e. \( x \in X \).

For \( 0 < \varepsilon < 1 \), let us choose a finite Borel partition \( \xi \) of \( X \) with \( \text{diam}(\xi) < \varepsilon \). Then, by Lemma 2.6, for each \( y \in X \) and \( n \geq 1 \), we have \( \xi_n^n(y) \subset B_n(y, \varepsilon) \subset B_{\delta_n}(y, \varepsilon) \). Observe that \( \mu \) is ergodic. According to Theorem 2.1 and Lemma 2.2, there exists a subset \( X_1 \subset X \) with \( \mu(X_1) = 1 \) such that for any \( x \in X_1 \), there exists \( W_x \in B_X \) with \( \mu_x^A(W_x) = 1 \) such that for any \( y \in W_x \),

\[ \lim_{n \to \infty} \frac{-\log \mu_x^A(\xi_n^n(y))}{n} = h_\mu(T, \xi | A). \]

Fix \( x \in X_1 \). For \( n \in \mathbb{N} \) and \( \gamma > 0 \), set

\[ Y_n = \{ y \in W_x : \mu_x^A(\xi_n^n(y)) > \exp(-(h_\mu(T, \xi | A) + \gamma)n) \} = \bigcup_{V \in J_n} V, \]

where \( J_n = \{ V \in \xi_n^n : \mu_x^A(V) > \exp(-(h_\mu(T, \xi | A) + \gamma)n) \} \). Then for each \( \gamma > 0 \), \( n \to \infty \), \( \mu_x^A(Y_n) = 1 \). Thus, for sufficiently large \( n \in \mathbb{N} \), we have \( \mu_x^A(Y_n) > 1 - \delta \). Since

\[ \# J_n = \# \{ V \in \xi_n^n : \mu_x^A(V) > \exp(-(h_\mu(T, \xi | A) + \gamma)n) \} \leq \exp((h_\mu(T, \xi | A) + \gamma)n), \]

the set \( Y_n \) contains at most \( \exp((h_\mu(T, \xi | A) + \gamma)n) \) elements of the partition \( \xi_n^n \).

And it can be covered by the same number of \( \delta_n \)-balls with radius \( \varepsilon \), so

\[ \tilde{N}_{\mu_x^A}(n, \varepsilon, \delta) \leq \exp((h_\mu(T, \xi | A) + \gamma)n). \]

Then for any \( \gamma > 0 \),

\[ \lim_{\varepsilon \to 0} \limsup_{n \to \infty} \frac{\log \tilde{N}_{\mu_x^A}(n, \varepsilon, \delta)}{n} \leq h_\mu(T, \xi | A) + \gamma. \]
Since $\gamma$ can be taken arbitrarily small and $h_\mu(T, \xi|A) \leq h_\mu(T|A)$, we obtain
\[
\lim_{\varepsilon \to 0} \limsup_{n \to \infty} \frac{\log \tilde{N}_{\mu^A}(n, \varepsilon, \delta)}{n} \leq h_\mu(T|A)
\]
for every $x \in X_1$. Noting that $\mu(X_1) = 1$, we have
\[
\lim_{\varepsilon \to 0} \limsup_{n \to \infty} \frac{\log \tilde{N}_{\mu^A}(n, \varepsilon, \delta)}{n} \leq h_\mu(T|A)
\]
for $\mu$-a.e. $x \in X$.

(2) Secondly, we will turn to prove the second part of the theorem: for every $0 < \delta < 1$, we have
\[
\lim_{\varepsilon \to 0} \liminf_{n \to \infty} \frac{\log \tilde{N}_{\mu^A}(n, \varepsilon, \delta)}{n} \geq h_\mu(T|A)
\]
for $\mu$-a.e. $x \in X$.

(i) Let $0 < \delta < 1$ be given. Let $\varepsilon > 0$, without of loss generality, we require additionally $\varepsilon^{\frac{1}{4}} < \frac{1-\delta}{\delta}$. There exists a finite Borel partition $\xi$ of $X$ satisfying:
\begin{itemize}
  \item $h_\mu(T, \xi|A) \geq h_\mu(T|A) - \varepsilon$,
  \item $\mu(\partial \xi) = 0$.
\end{itemize}
For $\theta > 0$, let
\[
U_\theta(\xi) = \{x \in X : \text{the ball } B(x, \theta) \text{ is not contained in } \xi(x)\},
\]
where $\xi(x)$ denotes the element of the partition $\xi$ containing $x$. Since $\int_{\theta > 0} U_\theta(\xi) = \partial \xi$, we have
\[
\mu(U_\theta(\xi)) \to 0, \quad \text{as } \theta \to 0.
\]
Therefore, there exists $0 < \gamma < \varepsilon$ such that $\mu(U_\theta(\xi)) \leq \varepsilon$ for any $0 < \theta \leq \gamma$. Using Birkhoff Ergodic Theorem, for $\mu$-a.e. $y \in X$ there exists $N_1(y) > 0$ such that for any $k \geq N_1(y)$,
\[
\frac{1}{k} \sum_{i=0}^{k-1} \chi_{U_\gamma(\xi)}(T^i(y)) \leq \varepsilon,
\]
where $\chi_{U_\gamma(\xi)}$ is the characteristic function of the set $U_\gamma(\xi)$. For $l \in \mathbb{N}^+$, we define
\[
D_l = \left\{y \in X : \frac{1}{k} \sum_{i=0}^{k-1} \chi_{U_\gamma(\xi)}(T^i(y)) \leq \varepsilon \text{ for any } k \geq l \right\}.
\]
Clearly, the sets $D_l$ are nested and exhaust $X$ up to a set of $\mu$-measure zero. Therefore, there exists $l_0 > 1$ such that $\mu(D_{l_0}) \geq 1 - 2\sqrt{\varepsilon}$ for any $l \geq l_0$.

Define $M_l = \{x \in X : \mu^A(D_l) \geq 1 - 2\varepsilon^\frac{1}{2}\}$, then $M_l^C = \{x \in X : \mu^A(D_l^C) \geq 2\varepsilon^\frac{1}{2}\}$. Using Chebyshev’s Inequality, we obtain
\[
\mu(M_l^C) \leq \frac{\int \mu^A(D_l^C) \text{d}\mu(x)}{2\varepsilon^\frac{1}{2}} = \frac{\mu(D_l^C)}{2\varepsilon^\frac{1}{2}} < \frac{2\sqrt{\varepsilon}}{2\varepsilon^\frac{1}{2}} = \varepsilon^\frac{1}{4},
\]
for any $l \geq l_0$. Thus for any $l \geq l_0$, $\mu(M_l) > 1 - \varepsilon^\frac{1}{4}$. The sets $D_l$ are nested, i.e. $D_1 \subset D_2 \subset \cdots$. Then fix some $l_1 > l_0$, for any $x \in M_{l_1}$, $l \geq l_1$ we have
\[
\mu^A(D_l) \geq \mu^A(D_{l_1}) \geq 1 - 2\varepsilon^\frac{1}{4}.
\]
(6) According to Lemma 2.2, we can find a subset $X_1 \subset X$ with $\mu(X_1) = 1$ such that for any $x \in X_1$, there exists $W_x \in \mathcal{B}_X$ with $\mu^A(W_x) = 1$ such that for any $y \in W_x$,
\[
\lim_{n \to \infty} \frac{-\log \mu^A(\xi_{n-1}^0(y))}{n} = h_\mu(T, \xi|A).
\]
Let \( I = X_1 \cap M_1 \). Clearly, \( \mu(I) > 1 - \varepsilon^{\frac{1}{2}} \).

(ii) Fix \( \hat{x} \in I \) and \( l \geq l_1 \). Noting that \( B_{d_n}(y, \gamma^2) \subset B_n(y; y, \gamma) \), hence if \( y \in D_l, n \geq l \) and \( z \in B_{d_n}(y, \gamma^2) \), the distance \( d_n^2 \) between \((\xi, n)\)-names of \( y \) and \( z \) does not exceed \( \gamma + \varepsilon \), i.e.

\[
d_n^2(C(n, y), C(n, z)) \leq \gamma + \varepsilon.
\]

Furthermore, for \( y \in D_l, n \geq l, B_{d_n}(y, \gamma^2) \) is contained in the set of points \( z \) whose \((\xi, n)\)-names are \((\gamma + \varepsilon)\)-close to the \((\xi, n)\)-name of \( y \), i.e.

\[
B_{d_n}(y, \gamma^2) \subset B_{d_n}(y, \gamma + \varepsilon). \tag{7}
\]

By Stirling’s formula, there exists a large number \( l_2 \in \mathbb{N} \) and for any \( n \geq l_2 \), it can be shown that the total number \( K_n \) of such \((\xi, n)\)-names consisting of \( B_{d_n}(y, \gamma + \varepsilon) \) admits the following estimate:

\[
K_n \leq \sum_{j=0}^{[n(\gamma+\varepsilon)]} C_n^j (\# \xi - 1)^j \leq \sum_{j=0}^{[n(\gamma+\varepsilon)]} C_n^j (\# \xi)^j \leq \exp((\gamma + \varepsilon + \delta)n), \tag{8}
\]

where

\[
\delta = (\gamma + \varepsilon) \log(\# \xi) - (\gamma + \varepsilon) \log(\gamma + \varepsilon) - (1 - \gamma - \varepsilon) \log(1 - \gamma - \varepsilon).
\]

For \( n \geq \max\{l, l_2\} \), set

\[
\mathcal{U} := \left\{ B_{d_n}(y, \gamma^2) : i = 1, 2, \cdots, N_{\mu^A}(n, \frac{\gamma^2}{2}, \delta) \right\}
\]

with \( \mu^A_\varepsilon(F_n) > 1 - \delta \), where

\[
F_n := \bigcup_{i=1}^{S_{\mu^A}(n, \frac{\gamma^2}{2}, \delta)} B_{d_n}(y, \gamma^2).
\]

According to (6) and \( \varepsilon^{\frac{1}{2}} < \frac{1-\delta}{4} \), then \( \mu^A_\varepsilon(F_n \cap D_n) > 1 - \delta - 2\varepsilon^{\frac{1}{2}} > \frac{1-\delta}{4} \). For \( i = 1, 2, \cdots, N_{\mu^A}(n, \frac{\gamma^2}{2}, \delta) \), if \( B_{d_n}(y, \gamma^2) \cap D_n \neq \emptyset \), we choose any \( z_i \in B_{d_n}(y, \gamma^2) \cap D_n \). Then we apply the relation (7), so we have

\[
B_{d_n}(y, \gamma^2) \cap D_n \subset B_{d_n}(z_i, \gamma^2) \subset B_{d_n}(z_i, \gamma + \varepsilon).
\]

Thus,

\[
F_n \cap D_n \subset S_n,
\]

where

\[
S_n = \bigcup_{\{i : B_{d_n}(y, \gamma^2) \cap D_n \neq \emptyset\}} B_{d_n}(z_i, \gamma + \varepsilon).
\]

Let \( \mathcal{P}_n \) be the set of \( V \in \xi_0^{n-1} \) such that \( d_n^2(C(n, V), C(n, z_i)) < \gamma + \varepsilon \) for some \( i = 1, 2, \cdots, N_{\mu^A}(n, \frac{\gamma^2}{2}, \delta) \). It is clear that \( S_n = \bigcup_{V \in \mathcal{P}_n} V \) and

\[
\# \mathcal{P}_n \leq N_{\mu^A}(n, \frac{\gamma^2}{2}, \delta) \cdot K_n.
\]

By Lemma 2.2 and Egorov Theorem, there exists a large number \( l_3 > \max\{l, l_2\} \) such that, \( \mu^A_\varepsilon(T_n) \geq \frac{1-\delta}{4} \) for each \( n \geq l_3 \), where

\[
T_n = \{ y \in S_n : \mu^A_\varepsilon(\xi_0^{n-1}(y)) \leq \exp(-(h_\mu(T, \xi|A) - \varepsilon)n) \}.
\]
Write \( t_n := \# \{ \xi_0^{n-1}(y) : y \in T_n \} \). Then
\[
\frac{(1 - \delta) \exp((h_\mu(T, \xi|A) - \varepsilon)n)}{4} \leq t_n \leq \# P_n \leq \hat{N}_{\mu^A}(n, \frac{\varepsilon^2}{2}, \delta) \cdot K_n.
\]

Hence, we have
\[
\hat{N}_{\mu^A}(n, \frac{\varepsilon^2}{2}, \delta) \geq (1 - \delta) \exp((h_\mu(T, \xi|A) - \varepsilon)n) \frac{4}{K_n}.
\]

Noting that \( h_\mu(T, \xi|A) > h_\mu(T)|A| - \varepsilon \) and using (8), we obtain
\[
\liminf_{n \to \infty} \frac{\log \hat{N}_{\mu^A}(n, \frac{\varepsilon^2}{2}, \delta)}{n} \geq h_\mu(T, \xi|A) - \varepsilon - \limsup_{n \to \infty} \frac{\log K_n}{n} + \lim_{n \to \infty} \frac{1}{n} \log \frac{1 - \delta}{4} 
\]
\[
\geq h_\mu(T, \xi|A) - \varepsilon - (\gamma + \varepsilon + \delta)
\]
\[
\geq h_\mu(T|A) - 2\varepsilon - (\gamma + \varepsilon + \delta)
\]
\[
= h_\mu(T|A) - 3\varepsilon - \gamma - \delta.
\]

Let \( \varepsilon \to 0 \). Since \( \gamma < \varepsilon \), \( \lim_{\varepsilon \to 0} \delta = 0 \), \( \check{x} \in I \) and \( \mu(I) = 1 - \varepsilon \frac{1}{\varepsilon} \), we have
\[
\lim_{\gamma \to 0} \liminf_{n \to \infty} \frac{\log \hat{N}_{\mu^A}(n, \frac{\varepsilon^2}{2}, \delta)}{n} \geq h_\mu(T|A),
\]
for \( \mu \)-a.e. \( \check{x} \in X \).

Therefore, for every \( 0 < \delta < 1 \), we obtain
\[
\liminf_{\varepsilon \to 0} \liminf_{n \to \infty} \frac{\log \hat{N}_{\mu^A}(n, \varepsilon, \delta)}{n} \geq h_\mu(T|A)
\]
for \( \mu \)-a.e. \( x \in X \). \( \square \)

4. Example. In this Section, we give an example of Theorem 1.1. Let \( m \geq 2 \) be a fixed integer and let \((p_0, p_1, \cdots, p_{m-1})\) be a probability vector with non-zero entries (i.e., \( p_i > 0 \) each \( i \) and \( \sum_{i=0}^{m-1} p_i = 1 \)). Let \( (Y, 2^Y, \nu) \) denote the measure space where \( Y = \{0, 1, \cdots, m-1\} \) and the point \( i \) has measure \( p_i \). Let \( (X, \mathcal{B}, \mu) = \prod_0^\infty (Y, 2^Y, \nu) \).

If \( n > 0 \) and \( a_j \in Y, j = 0, 1, \cdots, n \), the set
\[
[a_0, a_1, \cdots, a_n] := \{(x_i)_0^{\infty} : x_j = a_j, j = 0, 1, \cdots, n\}
\]
is the elementary rectangle, and we call it a block with end 0 and \( n \).

If we write points of \( X \) in the form \((x_0, x_1, \cdots)\), where \( x_i \in Y \), define \( T : X \to X \) by \( T(x_0, x_1, \cdots) = (x_1, x_2, \cdots) \). The space \((X, \mathcal{B}, \mu)\) is called the one-sided symbolic space, and the transformation \( T \) is called the one-sided \((p_0, p_1, \cdots, p_{m-1})\)-shift. It is well known that one-sided \((p_0, p_1, \cdots, p_{m-1})\)-shift is ergodic (see [14], Theorem 1.12) and \( h_\mu(T) = -\sum_{i=0}^{m-1} p_i \log p_i \) (see [14], Theorem 4.26).

Now we choose \( p_i = \frac{1}{m}, i = 0, 1, \cdots, m-1 \). Then \( h_\mu(T) = \log m \). Fix \( 0 < \lambda < 1 \), the metric \( d_\lambda \) on \((X, T)\) is given by
\[
d_\lambda(x,y) = \lambda^n |x-y|,
\]
where \( x = (x_0, x_1, \cdots), y = (y_0, y_1, \cdots) \in (X, T) \) and \( n(x,y) = \inf \{ i : x_i \neq y_i \} \).

For any \( x \in X \), and any \( \varepsilon > 0 \) there exists \( k := k(\varepsilon) \) such that
\[
[x_0, x_1, \cdots, x_{n+k+1}] \subset B_n(x, \varepsilon) \subset [x_0, x_1, \cdots, x_{n+k}].
\]
Then we have
\[
\liminf_{n \to \infty} -\frac{\log \mu(B_n(g; x, \varepsilon))}{n} \geq \lim_{n \to \infty} -\frac{\log(\frac{1}{m})^{n+k}}{n} + \liminf_{n \to \infty} -\frac{\log \sum_{j=0}^{[n\varepsilon]+k} C_{n+k}^j (m-1)^j}{n}.
\]

By Stirling’s formula, we obtain
\[
\liminf_{n \to \infty} -\frac{\log \sum_{j=0}^{[n\varepsilon]+k} C_{n+k}^j (m-1)^j}{n} \geq -\varepsilon - \varepsilon \log(m-1) + \varepsilon \log \varepsilon + (1 - \varepsilon) \log(1 - \varepsilon).
\]

Therefore
\[
\liminf_{n \to \infty} -\frac{\log \mu(B_n(g; x, \varepsilon))}{n} \geq \log m - \varepsilon - \varepsilon \log(m-1) + \varepsilon \log \varepsilon + (1 - \varepsilon) \log(1 - \varepsilon).
\]

Hence
\[
\lim_{\varepsilon \to 0} \liminf_{n \to \infty} -\frac{\log \mu(B_n(g; x, \varepsilon))}{n} \geq \log m.
\]

Also, we have
\[
\lim_{\varepsilon \to 0} \liminf_{n \to \infty} -\frac{\log \mu(B_n(g; x, \sqrt{\varepsilon}))}{n} \geq \log m.
\]

Obviously
\[
\lim_{\varepsilon \to 0} \lim_{n \to \infty} -\frac{\log \mu(B_n(x, \varepsilon))}{n} = \log m.
\]

Since
\[
B_n(x, \varepsilon) \subset B_{\tilde{d}_n}(x, \varepsilon) \subset B_n(g; x, \sqrt{\varepsilon}),
\]

we have
\[
\lim_{\varepsilon \to 0} \liminf_{n \to \infty} -\frac{\log \mu(B_{\tilde{d}_n}(x, \varepsilon))}{n} = \lim_{\varepsilon \to 0} \limsup_{n \to \infty} -\frac{\log \mu(B_{\tilde{d}_n}(x, \varepsilon))}{n} = \log m.
\]

Acknowledgments. We would like to thank Doctor Xiaoyao Zhou and Doctor Cao Zhao for valuable remarks and discussions. We also thank the anonymous referees for their helpful comments and suggestions that improved the manuscript.

REFERENCES

[1] M. Brin and A. Katok, On local entropy, in Lecture Notes in Mathematics, 1007, Springer, Berlin, (1983), 30–38.

[2] Y. Cao, H. Hu and Y. Zhao, Nonadditive measure-theoretic pressure and applications to dimensions of an ergodic Measure, Ergod. Theory Dyn. Syst., 33 (2013), 831–850.

[3] W. Cheng, Y. Zhao and Y. Cao, Pressures for asymptotically sub-additive potentials under a mistake function, Discrete Contin. Dyn. Syst., 32 (2012), 487–497.

[4] M. Einsiedler and T. Ward, Ergodic Theory with a View Towards Number Theory, Graduate Texts in Math., 259, Springer-Verlag, London, 2011.

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

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

[7] M. Gröger and T. Jäger, Some remarks on modified power entropy, Dynamics and Numbers, Contem. Math., 669 (2016), 105–122.

[8] B. Hasselblatt and A. Katok, Principal Structures, in Handbook of Dynamical Systems, 1A, North-Holland, Amsterdam, (2002), 1–203.

[9] L. He, J. Lv and L. Zhou, Definition of measure-theoretic pressure using spanning sets, Acta Math. Sinica, Engl. Ser., 20 (2004), 709–718.

[10] W. Huang, Z. Wang and X. Ye, Measure complexity and Möbius disjointness, arXiv:1707.06345.
[11] A. Katok, Lyapunov exponents, entropy and periodic orbits for diffeomorphisms, *Publ. Math. Inst. Hautes Études Sci.*, 51 (1980), 137–173.
[12] C. Pfister and W. Sullivan, On the topological entropy of saturated sets, *Ergod. Theory Dyn. Syst.*, 27 (2007), 929–956.
[13] D. Thompson, Irregular sets, the $\beta$-transformation and the almost specification property, *Trans. Amer. Math. Soc.*, 364 (2012), 5395–5414.
[14] P. Walters, *An Introduction to Ergodic Theory*, Graduate Texts in Math., 79, Springer, Berlin, 1982.
[15] Y. Zhao and Y. Cao, Measure-theoretic pressure for subadditive potentials, *Nonlinear Analysis*, 70 (2009), 2237–2247.
[16] D. Zheng, E. Chen E and J. Yang, On large deviations for amenable group actions, *Discrete Contin. Dyn. Syst. Ser. A.*, 36 (12), 2017, 7191–7206.
[17] X. Zhou, L. Zhou and E. Chen, Brin-Katok formula for the measure theoretic $r$–entropy, *C. R. Acad. Sci. Paris, Ser. I.*, 352 (2014), 473–477.
[18] X. Zhou, A formula of conditional entropy and some applications, *Discrete Contin. Dyn. Syst. Ser. A.*, 36 (2016), 4063–4075.
[19] Y. Zhu, On local entropy of random transformations, *Stoch. Dyn.*, 8 (2008), 197–207.
[20] Y. Zhu, Two notes on measure-theoretic entropy of random dynamic systems, *Acta Math. Sin.*, 25 (2009), 961–970.

Received January 2018; revised May 2018.

E-mail address: pinghuang1984@163.com
E-mail address: ecchen@njnu.edu.cn
E-mail address: chenweiwang010163.com