PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 143, Number 1, January 2015, Pages 141–150
S 0002-9939(2014)12175-7
Article electronically published on August 22, 2014

TOPOLOGICAL CONDITIONAL ENTROPY FOR AMENABLE GROUP ACTIONS

XIAOYAO ZHOU, YAQING ZHANG, AND ERCAI CHEN

(Communicated by Yingfei Yi)

Abstract. We introduce the topological conditional entropy for countable discrete amenable group actions and establish a variational principle for it.

1. Introduction

The theory related to the entropy and variational principle plays a fundamental role in ergodic theory and dynamical systems. The first variational principle that reveals the relationship between topological entropy and measure-theoretic entropy was obtained by L. Goodwyn [6] and T. Goodman [5]. M. Misiurewicz gave a short proof in [11]. R. Bowen [2] established the variational principle of entropy for non-compact sets in 1973. Y. Pesin and B. Pitskel [15] studied the variational principle of pressure for non-compact sets.

The topological conditional entropy is an invariant introduced by M. Misiurewicz [12], which is bigger than the defect of upper semi-continuity of the measure theoretical entropy regarded as a function of invariant regular probability measures. The variational principle of the topological conditional entropy in the case of $\mathbb{Z}_+$ was established by F. Ledrappier [8]. There is a close relationship between topological conditional entropy and tail entropy. The reader is referred to [3, 4, 9] for details. Also, we refer the reader to [1, 7, 10, 14] for entropy theory of amenable group actions.

This article is devoted to investigating the topological conditional entropy for the actions of countable discrete amenable groups and establishing the variational principle for it.

2. Preliminaries

2.1. **Amenable group.** This subsection presents some backgrounds of countable discrete amenable infinite groups. First, we present a remark about some notations to be used in this article.

Remark 2.1. Let G be a countable discrete infinite group.

- Let $\mathcal{F}(G)$ be the set of all non-empty finite subsets of G.
- $|\cdot|$ is the counting measure.
- For two sets A, B, $A\Delta B := (A \setminus B) \cup (B \setminus A)$.

Received by the editors December 4, 2012 and, in revised form, February 19, 2013.
2010 *Mathematics Subject Classification.* Primary 37D35, 37A35.
Key words and phrases. Topological conditional entropy, amenable group, variational principle.

- For $K, F \in \mathcal{F}(G), KF := \{kf : k \in K, f \in F\}, K^{-1} := \{k^{-1} : k \in K\}, B(F,K) := \{g \in G : Kg \cap F \neq \emptyset \text{ and } Kg \cap (G \setminus F) \neq \emptyset\} = K^{-1}F \cap K^{-1}(G \setminus F)$.
- For $E \in \mathcal{F}(G)$ and $\epsilon > 0$, write $M[E,\epsilon] = \{F \in \mathcal{F}(G) : F \text{ is } [E,\epsilon]\text{-invariant}\}$, where F is $[E,\epsilon]$-invariant means $|\{s \in F : Es \subset F\}| \geq (1-\epsilon)|F|$.

We say that G is an amenable group (or G is amenable for short) if for every $K \in \mathcal{F}(G)$ and $\delta > 0$, there exists $F \in \mathcal{F}(G)$ such that

$$\frac{|F \Delta KF|}{|F|} < \delta.$$

A set $A \in \mathcal{F}(G)$ is (K,δ)-invariant means that

$$\frac{|B(A,K)|}{|A|} < \delta.$$

A sequence $\{F_n\}_{n\in\mathbb{N}} \subset \mathcal{F}(G)$ called a Følner sequence, refers to that for every $K \in \mathcal{F}(G)$ and $\delta > 0, F_n$ is (K,δ)-invariant for sufficiently large n. In fact, a countable group G is amenable if and only if there exists a Følner sequence.

Let $f : \mathcal{F}(G) \to \mathbb{R}$ be a real-valued function. We say that f is

(i) monotone, if $f(E) \leq f(F)$, for any $E, F \in \mathcal{F}(G)$ with $E \subset F$;
(ii) non-negative, if for every $F \in \mathcal{F}(G), f(F) \geq 0$;
(iii) G-invariant, if $f(Fg) = f(F)$ for any $F \in \mathcal{F}(G)$ and $g \in G$;
(iv) sub-additive, if $f(E \cup F) \leq f(E) + f(F)$ for any disjoint $E, F \in \mathcal{F}(G)$;
(v) strongly sub-additive, if $f(E \cup F) + f(E \cap F) \leq f(E) + f(F)$ for any $E, F \in \mathcal{F}(G)$, where we set $f(\emptyset) = 0$ by convention.

The following limit theorem plays a central role in the definition of some dynamical invariants such as measure-theoretic entropy.

Lemma 2.1 (Ornstein-Weiss). *Let $f : \mathcal{F}(G) \to \mathbb{R}$ be a monotone, G-invariant, sub-additive function. Then there exists $\lambda = \lambda(G,f) \in [-\infty, +\infty)$ depending only on G and f such that*

$$\lim_{n\to\infty} \frac{f(F_n)}{|F_n|} = \lambda$$

for all Følner sequences $\{F_n\}_{n\in\mathbb{N}}$ of G.

Remark that (see Chapter 3 of [13]) if f is also strongly sub-additive, then

$$\lim_{n\to\infty} \frac{f(F_n)}{|F_n|} = \inf_{F\in\mathcal{F}(G)} \frac{f(F)}{|F|}.$$

We view the family of two-tuples $\{(E,\epsilon) : E \in \mathcal{F}(G), \epsilon > 0\}$ as a directed set by declaring $(E_1,\epsilon_1) \leq (E_2,\epsilon_2)$ if $E_2 \supset E_1, \epsilon_2 \leq \epsilon_1$.

Lemma 2.2. *Given a real-valued function $f : \mathcal{F}(G) \to \mathbb{R}$, then there exists a Følner sequence $\{F_n\}_{n\in\mathbb{N}}$ such that*

$$\lim_{n\to\infty} \frac{f(F_n)}{|F_n|} = \lim_{(E,\epsilon)} \sup_{F\in M[E,\epsilon]} \frac{f(F)}{|F|}.$$

Remark that if f is monotone, non-negative, G-invariant, and sub-additive, then

$$\lim_{(E,\epsilon)} \sup_{F\in M[E,\epsilon]} \frac{f(F)}{|F|} = \lim_{(E,\epsilon)} \inf_{F\in M[E,\epsilon]} \frac{f(F)}{|F|} = \lim_{n\to\infty} \frac{f(F_n)}{|F_n|},$$

for any Følner sequence $\{F_n\}_{n\in\mathbb{N}}$.

2.2. **Topological conditional entropy and measure conditional entropy.** Throughout this article, by a G-dynamical system (X,G), we mean a compact metric space X paired with a countable discrete amenable infinite group G. A remark is presented here for some notations of a G-dynamical system (X,G).

Remark 2.2. Let (X,G) be a G-dynamical system.

- Denote by $\mathcal{B}(X), \mathcal{C}_X, \mathcal{P}(X), \mathcal{C}_X^o$, the collection of all Borel subsets of X, the set of all finite Borel covers of X, the set of all finite Borel partitions of X, the set of all finite open covers of X, respectively.
- Let $M(X)$ be the set of all Borel probability measures on X and let $M(X,G)$ be the set of all G-invariant Borel probability measures on X. The amenability ensures that $M(X,G) \neq \emptyset$.
- Given two covers $\mathcal{U}, \mathcal{V} \in \mathcal{C}_X, \mathcal{U}$ is said to be finer than $\mathcal{V}$ (denoted by $\mathcal{U} \succeq \mathcal{V}$ or $\mathcal{V} \preceq \mathcal{U}$) if each element of $\mathcal{U}$ is contained in some element of $\mathcal{V}$.
- Given $F \in \mathcal{F}(G)$ and $\mathcal{U}, \mathcal{V} \in \mathcal{C}_X$, set $\mathcal{U} \vee \mathcal{V} = \{U \cap V : U \in \mathcal{U}, V \in \mathcal{V}\}$ and $\mathcal{U}_F = \bigvee_{g \in F} g^{-1}\mathcal{U}$ (letting $\mathcal{U}_\emptyset = \{X\}$).
- For a subset A of $X, \partial(A) = \overline{A} \cap \overline{X \setminus A}$ denote the boundary of A.

Given $\mathcal{A} \in \mathcal{C}_X$ and $X_1 \subset X$, $N(X_1, \mathcal{A})$ is given by

$$N(X_1,\mathcal{A}) = \min\left\{ |\mathcal{A}_1| : \mathcal{A}_1 \subset \mathcal{A}, X_1 \subset \bigcup_{B \in \mathcal{A}_1} B \right\}.$$

Given $\mathcal{A}, \mathcal{B} \in \mathcal{C}_X, E \in \mathcal{F}(G)$ and $\epsilon > 0$, set

$$\begin{aligned} N(\mathcal{A}|\mathcal{B}) &= \max\{N(B,\mathcal{A}) : B \in \mathcal{B}\}, \\ H(\mathcal{A}|\mathcal{B}) &= \log N(\mathcal{A}|\mathcal{B}), \\ h(\mathcal{A}|\mathcal{B}) &= \lim_{(E,\epsilon)} \sup_{F \in M[E,\epsilon]} \frac{H(\mathcal{A}_F|\mathcal{B}_F)}{|F|}, \\ h(G,\mathcal{B}) &= \sup_{\mathcal{A}} h(\mathcal{A}|\mathcal{B}), \\ h^*(G,X) &= \inf_{\mathcal{B}} h(G,\mathcal{B}). \end{aligned}$$

Remark that

- By Lemma 2.2, we can define $h(\mathcal{A}|\mathcal{B})$ by fixing a Følner sequence $\{F_n\}_{n\in\mathbb{N}}$ as follows:
$$h(\mathcal{A}|\mathcal{B}) = \lim_{n\to\infty} \frac{1}{|F_n|} H(\mathcal{A}_{F_n}|\mathcal{B}_{F_n}).$$
- Replacing supremum in the third equality with infimum, the variational principle still holds. This implies that $h^*(G,X)$ can be defined independent of the choice of Følner sequences.
- The supremum in the fourth equality is taken over all finite open covers of X.
- The infimum in the fifth equality runs over all finite open covers of X.
- $h^*(G,X)$ is called the topological conditional entropy of a G-dynamical system (X,G).

If Y is another compact Hausdorff space and (Y,G) is a G-dynamical system, then $(Y \times X, G)$ is a G-dynamical system.[1] Suppose m is a Borel probability measure on

[1]For $g \in G, x \in X, y \in Y$, set $g(x,y) = (gx, gy)$.

$Y\times X, \pi_X, \pi_Y$ are the projection maps on $Y\times X$ defined by $\pi_X(y,x)=x, \pi_Y(y,x)=y$, and $\xi=\{A_1, A_2, \cdots, A_n\}\in\mathcal{P}(Y\times X)$.

A function $\phi:[0,1]\to[0,\infty]$ is given by

$$\phi(t)=\begin{cases} -t\log t, & t>0; \\ 0, & t=0. \end{cases}$$

Define

$$H_m(\xi|Y)=\sum_{j=0}^{n-1}\int\phi(\mathbb{E}(1_{A_j}|\pi_Y^{-1}\mathcal{B}(Y)))\mathrm{d}m,$$

where $\mathbb{E}(1_{A_j}|\pi_Y^{-1}\mathcal{B}(Y))$ denotes the conditional expectation of 1_{A_j} with respect to $\pi_Y^{-1}\mathcal{B}(Y)$.

By a straightforward computation, we have that the function $F\in\mathcal{F}(G)\mapsto H_m(\xi_F|Y)$ is monotone, non-negative, G-invariant and strongly sub-additive when m is a G-invariant measure. Thus, for any Følner sequences $\{F_n\}_{n\in\mathbb{N}}$, and G-invariant measures m,

$$\lim_{n\to\infty}\frac{1}{|F_n|}H_m(\xi_{F_n}|Y)=\inf_{F\in\mathcal{F}(G)}\frac{1}{|F|}H_m(\xi_F|Y).$$

The conditional measure entropy of $m\in(Y\times X,G)$ is given by setting

$$h(m|Y)=\sup_\xi h_m(\xi|Y)=\sup_\xi\lim_{n\to\infty}\frac{1}{|F_n|}H_m(\xi_{F_n}|Y),$$

$$h^*(m|Y)=\begin{cases}\limsup\limits_{\mu\to m}h(\mu|Y)-h(m|Y), & h(m|Y)\neq\infty; \\ \infty, & h(m|Y)=\infty,\end{cases}$$

where the supremum in the first equality ranges over $\mathcal{P}(Y\times X)$.

3. Variational principle

Variational principles are beautiful results in a dynamical system. Establishing variational principles is an important topic in ergodic theory. The variational principle of this article reveals the relationship between the topological conditional entropy and the conditional measure entropy. We state our main result below:

Theorem 3.1. *Let X, Y be two compact Hausdorff spaces, and let $(X,G),(Y,G)$ be two G-dynamical systems; then*

$$h^*(G,X)\geq\sup\{h^*(m|Y): m\in M(Y\times X,G)\}.$$

In particular, if $(Y,G)=(X,G)$, then

$$h^*(G,X)=\max\{h^*(m|X): m\in M(X\times X,G)\}.$$

Proof. The theorem follows from Lemma 3.6 and Lemma 3.8. □

The following three lemmas can be seen in [8] for the case of $G=\mathbb{Z}_+$. The corresponding results of the case of amenable groups can be obtained by minor modification. So the proofs are omitted here.

Lemma 3.1. *Suppose $\{\mathcal{C}_\alpha\}$ is a finite sub-σ-algebra sequence increasing to $\mathcal{B}(Y), \xi$ is a Borel partition of $Y\times X$; then*

$$H_\mu(\xi|Y)=\inf_\alpha H_\mu(\xi|\pi_Y^{-1}\eta_\alpha),$$

where $\eta_\alpha=\xi(\mathcal{C}_\alpha)$ is the partition corresponding to sub-σ-algebra $\mathcal{C}_\alpha$.

Lemma 3.2. *Let $\xi \in \mathcal{P}(Y \times X)$ with $\mu(\partial\xi) := \mu(\bigcup_{A\in\xi} \partial(A)) = 0$. Then the function $\mu \mapsto H_\mu(\xi|Y)$ defined on the set of Radon probability measures on $Y \times X$, is upper semi-continuous at μ.*

Lemma 3.3. *Assume that $\xi_1, \xi_2 \in \mathcal{P}(Y \times X)$; then*

$$H_\mu(\xi_1|Y) \leq H_\mu(\xi_2|Y) + \log N(\xi_1|\xi_2).$$

Lemma 3.4. *If $\zeta \in \mathcal{P}(Y), \xi \in \mathcal{P}(X)$ and $\mu \in M(Y \times X, G)$, then*

$$h_\mu(\zeta \times \xi|Y) = h_\mu(\pi_X^{-1}\xi|Y).$$

Proof. It follows from Lemma 3.1 that for any $F \in \mathcal{F}(G)$,

$$\begin{aligned}
H_\mu(\zeta_F \times \xi_F|Y) &= \inf_\alpha H_\mu(\zeta_F \times \xi_F|\pi_Y^{-1}\eta_\alpha) \\
&= \inf_\alpha H_\mu(\pi_Y^{-1}\zeta_F \vee \pi_X^{-1}\xi_F|\pi_Y^{-1}\eta_\alpha) \\
&= \inf_\alpha (H_\mu(\pi_X^{-1}\xi_F|\pi_Y^{-1}(\eta_\alpha \vee \zeta_F)) + H_\mu(\pi_Y^{-1}\zeta_F|\pi_Y^{-1}\eta_\alpha)) \\
&= \inf_\alpha H_\mu(\pi_X^{-1}\xi_F|\pi_Y^{-1}(\eta_\alpha \vee \zeta_F)) + \inf_\alpha H_\mu(\pi_Y^{-1}\zeta_F|\pi_Y^{-1}\eta_\alpha) \\
&= H_\mu(\pi_X^{-1}\xi_F|Y) + H_\mu(\pi_Y^{-1}\zeta_F|\pi_Y^{-1}\mathcal{B}(Y)) \\
&= H_\mu(\pi_X^{-1}\xi_F|Y).
\end{aligned}$$

Since $F \in \mathcal{F}(G)$ is arbitrary, we get $h_\mu(\zeta \times \xi|Y) = h_\mu(\pi_X^{-1}\xi|Y)$. □

Lemma 3.5. *If $\xi \in \mathcal{P}(X)$ and $\mu \in M(Y \times X, G)$, then*

$$h(\mu|Y) \leq h_\mu(\pi_X^{-1}\xi|Y) + h(G, \xi).$$

Proof. We claim that if $\alpha = \{A_1, A_2, \cdots, A_k\} \in \mathcal{P}(X)$ and $\epsilon > 0$, then there exists $\mathcal{U} = \{U_1, U_2, \cdots, U_k\} \in \mathcal{C}_X^o$ such that

$$H_{\pi_X\mu}(\alpha|\beta) \leq \epsilon$$

for each $\beta \in \mathcal{P}(X)$ with $\beta \succeq \mathcal{U}$. To see this, given $k \in \mathbb{N}$ and $\epsilon > 0$, there exists $\delta_1 = \delta_1(k, \epsilon)$ such that $H_{\pi_X\mu}(\alpha|\gamma) + H_{\pi_X\mu}(\gamma|\alpha) < \epsilon$ where $\gamma = \{C_1, \cdots, C_k\} \in \mathcal{P}(X)$ with $\pi_X\mu(\alpha\Delta\gamma) := \sum_{i=1}^{k} \pi_X\mu(A_i\Delta C_i) < \delta_1$ (see Lemma 4.15 of [16]). Since $\pi_X\mu$ is regular, there exist closed subsets $B_i \subset A_i$, such that $\pi_X\mu(A_i \setminus B_i) < \frac{\delta_1}{2k^2}$. Let $B_0 = X \setminus \bigcup_{i=1}^k B_i$; then $\pi_X\mu(B_0) < \frac{\delta_1}{2k}$. Let $U_i = B_0 \cup B_i, i = 1, 2, \cdots, k$; then $\mathcal{U} := \{U_1, \cdots, U_k\} \in \mathcal{C}_X^o$. If $\beta \in \mathcal{P}(X)$ and $\beta \succeq \mathcal{U}$, then there exists $\beta' = \{C_1, \cdots, C_k\} \in \mathcal{P}(X)$ such that $\beta \succeq \beta'$ and $C_i \subset U_i, i = 1, 2, \cdots, k$. Thus, $H_{\pi_X\mu}(\alpha|\beta) \leq H_{\pi_X\mu}(\alpha|\beta')$. Since $B_i = X \setminus \bigcup_{j\neq i} U_j \subset C_i \subset U_i$, we have

$$\pi_X\mu(C_i\Delta A_i) \leq \pi_X\mu(A_i \setminus B_i) + \pi_X\mu(B_0) < \frac{\delta_1}{2k} + \frac{\delta_1}{2k} = \frac{\delta_1}{k}.$$

This implies that $\pi_X\mu(\alpha\Delta\beta') < \delta_1$. Therefore, $H_{\pi_X\mu}(\alpha|\beta) \leq H_{\pi_X\mu}(\alpha|\beta') \leq \epsilon$. This completes the proof of the claim.

For $F \in \mathcal{F}(G)$, there exists $\beta \in \mathcal{P}(X)$ such that $\beta \succeq \mathcal{U}_F$ and $|\beta| = N(X,\mathcal{U}_F)$. Then

$$\begin{aligned}
H_\mu(\pi_X^{-1}\alpha_F|Y) &\le H_\mu(\pi_X^{-1}\alpha_F \vee \pi_X^{-1}\beta|Y)\\
&\le H_\mu(\pi_X^{-1}\beta|Y) + H_\mu(\pi_X^{-1}\alpha_F|\pi_X^{-1}\beta)\\
&\le H_\mu(\pi_X^{-1}\beta|Y) + \sum_{g\in F} H_\mu(\pi_X^{-1}(g^{-1}\alpha)|\pi_X^{-1}\beta)\\
&= H_\mu(\pi_X^{-1}\beta|Y) + \sum_{g\in F} H_{\pi_X\mu}(\alpha|g\beta)\\
&\le H_\mu(\pi_X^{-1}\beta|Y) + |F|\epsilon\\
&\le H_\mu(\pi_X^{-1}\xi_F|Y) + \log N(\beta|\xi_F) + |F|\epsilon\\
&\le H_\mu(\pi_X^{-1}\xi_F|Y) + \log N(\mathcal{U}_F|\xi_F) + |F|\epsilon.
\end{aligned}$$

Replacing F with the fixed Følner sequence $\{F_n\}_{n\in\mathbb{N}}$, letting $n \to \infty$ and $\epsilon \to 0$, we get

$$h(\mu|Y) \le h_\mu(\pi_X^{-1}\xi|Y) + h(G,\xi).$$ □

Lemma 3.6. *If $m \in M(Y \times X, G)$, then*

$$h^*(m|Y) \le h^*(G,X).$$

Proof. Given $\mathcal{B} \in \mathcal{C}_X^o$. Let $\xi \in \mathcal{P}(X)$ such that $\xi \succeq \mathcal{B}$ and $m \circ \pi_X^{-1}(\partial\xi) := m \circ \pi_X^{-1}(\bigcup_{A\in\xi} \partial A) = 0$. It follows from Lemma 3.5 that

$$h(\mu|Y) \le h_\mu(\pi_X^{-1}\xi|Y) + h(G,\xi) \le \frac{1}{|F_n|} H_\mu(\pi_X^{-1}\xi_{F_n}|Y) + h(G,\mathcal{B}).$$

Thus,

$$\begin{aligned}
\limsup_{\mu\to m} h(\mu|Y) &\le \frac{1}{|F_n|} \limsup_{\mu\to m} H_\mu(\pi_X^{-1}\xi_{F_n}|Y) + h(G,\mathcal{B})\\
&\le \frac{1}{|F_n|} H_m(\pi_X^{-1}\xi_{F_n}|Y) + h(G,\mathcal{B}).
\end{aligned}$$

Letting $n \to \infty$, we obtain $\limsup_{\mu\to m} h(\mu|Y) \le h(m|Y) + h^*(G,X)$. This implies that $h^*(m|Y) \le h^*(G,X)$. □

Lemma 3.7. *Let (X,G) be a G-dynamical system and let (X_1,G) and (X_2,G) be two copies of (X,G). Suppose $\mathcal{B} = \{B_1, B_2, \cdots, B_p\} \in \mathcal{C}_X^o$; then there exists $\mu_{\mathcal{B}} \in M(X_1 \times X_2, G)$ such that*

- **(1)** $h(\mu_{\mathcal{B}}|X_1) \ge h(G,\mathcal{B}) - 1/p$;
- **(2)** *supp* $\mu_{\mathcal{B}} \subset \bigcup_{i=1}^{p} \overline{B_i} \times \overline{B_i}$.

Proof. Let us choose $\mathcal{A} \in \mathcal{C}_X^o$ such that

$$h(\mathcal{A}|\mathcal{B}) \ge h(G,\mathcal{B}) - 1/p,$$

and there exists a Følner sequence $\{F_n\}_{n\in\mathbb{N}}$ (fix it) such that

$$\lim_{n\to\infty} \frac{1}{|F_n|} \log N(\mathcal{A}_{F_n}|\mathcal{B}_{F_n}) = h(\mathcal{A}|\mathcal{B}).$$

There is $C_{\mathcal{B}_{F_n}} \in \mathcal{B}_{F_n}$ such that $N(C_{\mathcal{B}_{F_n}}, \mathcal{A}_{F_n}) = N(\mathcal{A}_{F_n}|\mathcal{B}_{F_n})$. Choose a point $x \in C_{\mathcal{B}_{F_n}}$ and choose a neighbourhood L of the diagonal $X \times X$ such that for any $y \in X$ there exists $A_y \in \mathcal{A}$ with $\{z : (y,z) \in L\} \subset A_y$. In fact, we can construct L as follows. Assume that δ is a Lebesgue number of $\mathcal{A}$. For any $y \in X$, write $B_y = \{(y,z) : d(y,z) \leq \delta/3\}$, and set $L = \{B_y : y \in X\}$.

A subset E of X is called (F_n, L)-separated if for any two different points $x, x' \in E$, there exists $g \in F_n$ such that $(gx, gx') \notin L$. Let $E_n^{C_{\mathcal{B}_{F_n}}}$ be an (F_n, L)-separated subset of $C_{\mathcal{B}_{F_n}}$ with maximal cardinality. It follows from the maximality that

$$C_{\mathcal{B}_{F_n}} \subset \bigcup_{y \in E_n^{C_{\mathcal{B}_{F_n}}}} \bigcap_{g \in F_n} g^{-1} A_{gy}.$$

Consider the measures σ_n and μ_n on the space $X_1 \times X_2$ defined by

$$\sigma_n = \frac{1}{|E_n^{C_{\mathcal{B}_{F_n}}}|} \sum_{y \in E_n^{C_{\mathcal{B}_{F_n}}}} \delta_{(x,y)}$$

and

$$\mu_n = \frac{1}{|F_n|} \sum_{g \in F_n} g\sigma_n,$$

where δ_x is the Dirac measure at the point x.

Taking a limit point $\mu_{\mathcal{B}}$ of the sequence $\{\mu_n\}_{n \in \mathbb{N}}$, we have $\mu_{\mathcal{B}} \in M(X_1 \times X_2, G)$. Letting $\xi = \{C_1, C_2, \cdots, C_p\} \in \mathcal{P}(X)$ such that $\mu_{\mathcal{B}} \circ \pi_{X_2}^{-1}(\partial C_i) = 0$ and $C_i \times C_i \subset L$, then

$$\begin{aligned}
&H_{\sigma_n}((\pi_{X_2}^{-1}\xi)_{F_n}|X_1) \\
&\quad = \inf_\alpha H_{\sigma_n}((\pi_{X_2}^{-1}\xi)_{F_n}|\pi_{X_1}^{-1}\eta_\alpha) \\
&\quad = \inf_\alpha \sum_{D_j \in \eta_\alpha} \sum_{A_i \in \xi_{F_n}} -\sigma_n(X_1 \times A_i \cap D_j \times X_2) \log \frac{\sigma_n(X_1 \times A_i \cap D_j \times X_2)}{\sigma_n(D_j \times X_2)} \\
&\quad = \inf_\alpha \sum_{A_i \in \xi_{F_n}} -\frac{1}{|E_n^{C_{\mathcal{B}_{F_n}}}|} \log \frac{1}{|E_n^{C_{\mathcal{B}_{F_n}}}|} \\
&\quad = \log |E_n^{C_{\mathcal{B}_{F_n}}}|.
\end{aligned}$$

Since for $F \in \mathcal{F}(G)$, the map $A \in \mathcal{F}(G) \mapsto H_{\sigma_n}(\xi_A|X_1)$ is monotone, non-negative, strongly sub-additive and $1_{F_n}(h) = \frac{1}{|F|} \sum_{g \in F^{-1}F_n} 1_{Fg \cap F_n}(h)$, we have

$$\begin{aligned}
H_{\sigma_n}((\pi_{X_2}^{-1}\xi)_{F_n} \mid X_1) &\leq \sum_{g \in F^{-1}F_n} \frac{1}{\mid F \mid} H_{\sigma_n}((\pi_{X_2}^{-1}\xi)_{Fg \cap F_n} \mid X_1) \\
&\leq \frac{1}{\mid F \mid} \sum_{g \in F^{-1}F_n} H_{g\sigma_n}((\pi_{X_2}^{-1}\xi)_F \mid X_1) \\
&\leq \frac{1}{\mid F \mid} (\sum_{g \in F_n} H_{g\sigma_n}((\pi_{X_2}^{-1}\xi)_F \mid X_1) + \sum_{g \in F^{-1}F_n \setminus F_n} H_{g\sigma_n}((\pi_{X_2}^{-1}\xi)_F \mid X_1)) \\
&\leq \frac{1}{\mid F \mid} \sum_{g \in F_n} H_{g\sigma_n}((\pi_{X_2}^{-1}\xi)_F \mid X_1) + \mid F^{-1}F_n \setminus F_n \mid \log \mid \xi \mid
\end{aligned}$$

and

$$\begin{aligned}
\frac{1}{|F_n|}\log|E_n^{C_{\mathcal{B}_{F_n}}}| &= \frac{1}{|F_n|}H_{\sigma_n}((\pi_{X_2}^{-1}\xi)_{F_n}\mid X_1)\\
&\le \frac{1}{|F_n|}(\frac{1}{|F|}\sum_{g\in F_n}H_{g\sigma_n}((\pi_{X_2}^{-1}\xi)_F\mid X_1)+|F^{-1}F_n\setminus F_n|\log|\xi|)\\
&= \frac{1}{|F|}(\frac{1}{|F_n|}H_{g\sigma_n}((\pi_{X_2}^{-1}\xi)_F\mid X_1))+\frac{|F^{-1}F_n\setminus F_n|}{|F_n|}\log|\xi|\\
&\le \frac{1}{|F|}H_{\mu_n}((\pi_{X_2}^{-1}\xi)_F\mid X_1)+\frac{|F^{-1}F_n\setminus F_n|}{|F_n|}\log|\xi|.
\end{aligned}$$

It follows from the fact that $\left\{\bigcap_{g\in F_n}g^{-1}A_{gy}:y\in E_n^{C_{\mathcal{B}_{F_n}}}\right\}\subset\mathcal{A}_{F_n}$ that

$$\frac{1}{|F_n|}\log N(\mathcal{A}_{F_n}\mid\mathcal{B}_{F_n})\le\frac{1}{|F|}H_{\mu_n}((\pi_{X_2}^{-1}\xi)_F\mid X_1)+\frac{|F^{-1}F_n\setminus F_n|}{|F_n|}\log|\xi|.$$

Since $\mu_{\mathcal{B}}\circ\pi_{X_2}^{-1}(\partial\xi_i)=0$, we get $\mu_{\mathcal{B}}\circ\pi_{X_2}^{-1}(\partial\xi_F)=0$. Letting $n\to\infty$,

$$\lim_{n\to\infty}H_{\mu_n}((\pi_{X_2}^{-1}\xi)_F\mid X_1)=H_{\mu_{\mathcal{B}}}((\pi_{X_2}^{-1}\xi)_F\mid X_1).$$

This implies that

$$h(\mathcal{A}\mid\mathcal{B})\le\frac{1}{|F|}H_{\mu_{\mathcal{B}}}((\pi_{X_2}^{-1}\xi)_F\mid X_1).$$

Since F is arbitrary, we get

$$h(\mathcal{A}\mid\mathcal{B})\le h_{\mu_{\mathcal{B}}}(\pi_{X_2}^{-1}\xi\mid X_1).$$

Then

$$h(\mu_{\mathcal{B}}\mid X_1)\ge h(G,\mathcal{B})-\frac{1}{p}.$$

Since

$$\begin{aligned}
\mu_n(\bigcup_{B_i\in\mathcal{B}}B_i\times B_i) &= \frac{1}{|F_n|}\sum_{g\in F_n}g\sigma_n(\bigcup_{B_i\in\mathcal{B}}B_i\times B_i)\\
&\ge \frac{1}{|F_n|}\sum_{g\in F_n}\sigma_n(C_{\mathcal{B}_{F_n}}\times C_{\mathcal{B}_{F_n}})=1,
\end{aligned}$$

we have supp $\mu_n\subset\bigcup_{B_i\in\mathcal{B}}\overline{B_i}\times\overline{B_i}$. This implies supp $\mu_{\mathcal{B}}\subset\bigcup_{i=1}^{p}\overline{B_i}\times\overline{B_i}$. □

Lemma 3.8. *Suppose*

$$(X_1,G)=(X_2,G)=(X,G);$$

then there exists $m\in M(X_1\times X_2,G)$ such that the support of m is carried by the diagonal and

$$h^*(m\mid X_1)=h^*(G,X_2).$$

Proof. For any finite open cover $\mathcal{B}$ of X, we choose the measure $\mu_{\mathcal{B}}$ as above, and take m adherent to the measures $\mu_{\mathcal{B}}$ as $\mathcal{B}$ becomes finer and finer. By the property (1) in Lemma 3.7, we have

$$\limsup_{\mu\to m} h(\mu \mid X_1) \geq \liminf_{\mathcal{B}} h(\mu_{\mathcal{B}} \mid X_1) \geq \inf_{\mathcal{B}} h(G,\mathcal{B}) = h^*(G,X_2),$$

and by the property (2), the measure m is carried by the diagonal of $X_1 \times X_2$. For the fixed Følner sequence $\{F_n\}_{n\in\mathbb{N}}$,

$$\begin{aligned} H_m((\pi_{X_2}^{-1}\xi)_{F_n} \mid X_1) &= \inf_{\alpha} H_m((\pi_{X_2}^{-1}\xi)_{F_n} \mid \pi_{X_1}^{-1}\eta_\alpha) \\ &= \inf_{\alpha} \sum_{D_j\in\eta_\alpha} \sum_{A_i\in\xi_{F_n}} -m(X_1\times A_i \cap D_j\times X_2)\log\frac{m(X_1\times A_i\cap D_j\times X_2)}{m(D_j\times X_2)} \\ &= \inf_{\alpha} \sum_{D_j\in\eta_\alpha} \sum_{A_i\in\xi_{F_n}} -m(D_j\times A_i)\log\frac{m(D_j\times A_i)}{m(D_j\times X_2)}. \end{aligned}$$

If $D_j \cap A_i = \emptyset$, then $m(D_j \times A_i) = 0$. If $D_j \cap A_i \neq \emptyset$, then there exists α large enough such that $\xi_{F_n} \preceq \eta_\alpha$ and $\log\frac{m(D_j\times A_i)}{m(D_j\times X_2)} = 0$. This implies that

$$H_m((\pi_{X_2}^{-1}\xi)_{F_n} \mid X_1) = \inf_{\alpha} H_m((\pi_{X_2}^{-1}\xi)_{F_n} \mid \pi_{X_1}^{-1}\eta_\alpha) = 0.$$

Thus $h_m(\xi|X_1) = 0$. By the arbitrariness of ξ, we have $h(m|X_1) = 0$, and

$$h^*(m \mid X_1) = \limsup_{\mu\to m} h(\mu \mid X_1) - h(m \mid X_1) \geq h^*(G, X_2).$$

This together with Lemma 3.6 implies that

$$h^*(m \mid X_1) = h^*(G, X_2).$$

□

Acknowledgements

The authors would like to thank an anonymous referee for useful comments. The work was supported by the National Natural Science Foundation of China (grant 11271191), the National Basic Research Program of China (grant 2013CB834100), and the Fundamental Research Funds for the Central Universities (grant WK0010000035) and China Postdoctoral Science Foundation (grant 2014M551806).

References

[1] Lewis Bowen, *Sofic entropy and amenable groups*, Ergodic Theory Dynam. Systems **32** (2012), no. 2, 427–466, DOI 10.1017/S0143385711000253. MR2901354

[2] Rufus Bowen, *Topological entropy for noncompact sets*, Trans. Amer. Math. Soc. **184** (1973), 125–136. MR0338317 (49 #3082)

[3] Mike Boyle and Tomasz Downarowicz, *The entropy theory of symbolic extensions*, Invent. Math. **156** (2004), no. 1, 119–161, DOI 10.1007/s00222-003-0335-2. MR2047659 (2005d:37015)

[4] David Burguet, *A direct proof of the tail variational principle and its extension to maps*, Ergodic Theory Dynam. Systems **29** (2009), no. 2, 357–369, DOI 10.1017/S0143385708080425. MR2486774 (2010b:37088)

[5] L. Wayne Goodwyn, *Topological entropy bounds measure-theoretic entropy*, Proc. Amer. Math. Soc. **23** (1969), 679–688. MR0247030 (40 #299)

[6] T. N. T. Goodman, *Relating topological entropy and measure entropy*, Bull. London Math. Soc. **3** (1971), 176–180. MR0289746 (44 #6934)

[7] Wen Huang, Xiangdong Ye, and Guohua Zhang, *Local entropy theory for a countable discrete amenable group action*, J. Funct. Anal. **261** (2011), no. 4, 1028–1082, DOI 10.1016/j.jfa.2011.04.014. MR2803841

[8] F. Ledrappier, *A variational principle for the topological conditional entropy*, Ergodic theory (Proc. Conf., Math. Forschungsinst., Oberwolfach, 1978), Lecture Notes in Math., vol. 729, Springer, Berlin, 1979, pp. 78–88. MR550412 (80j:54011)

[9] Yuan Li, Ercai Chen, and Wen-Chiao Cheng, *Tail pressure and the tail entropy function*, Ergodic Theory Dynam. Systems **32** (2012), no. 4, 1400–1417, DOI 10.1017/S0143385711000204. MR2955319

[10] Bingbing Liang and Kesong Yan, *Topological pressure for sub-additive potentials of amenable group actions*, J. Funct. Anal. **262** (2012), no. 2, 584–601, DOI 10.1016/j.jfa.2011.09.020. MR2854714 (2012k:37021)

[11] M. Misiurewicz, *A short proof of the variational principle for a* Z_+^n *action on a compact space* (English, with Russian summary), Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys. **24** (1976), no. 12, 1069–1075. MR0430213 (55 #3220)

[12] Michał Misiurewicz, *Topological conditional entropy*, Studia Math. **55** (1976), no. 2, 175–200. MR0415587 (54 #3672)

[13] Jean Moulin Ollagnier, *Ergodic theory and statistical mechanics*, Lecture Notes in Mathematics, vol. 1115, Springer-Verlag, Berlin, 1985. MR781932 (86h:28013)

[14] Donald S. Ornstein and Benjamin Weiss, *Entropy and isomorphism theorems for actions of amenable groups*, J. Analyse Math. **48** (1987), 1–141, DOI 10.1007/BF02790325. MR910005 (88j:28014)

[15] Ya. B. Pesin and B. S. Pitskel′, *Topological pressure and the variational principle for noncompact sets* (Russian, with English summary), Funktsional. Anal. i Prilozhen. **18** (1984), no. 4, 50–63, 96. MR775933 (86i:28031)

[16] Peter Walters, *An introduction to ergodic theory*, Graduate Texts in Mathematics, vol. 79, Springer-Verlag, New York, 1982. MR648108 (84e:28017)

School of Mathematical Sciences and Institute of Mathematics, Nanjing Normal University, Nanjing 210023, Jiangsu, People's Republic of China
E-mail address: zhouxiaoyaodeyouxian@126.com

School of Mathematical Sciences and Institute of Mathematics, Nanjing Normal University, Nanjing 210023, Jiangsu, People's Republic of China
E-mail address: zhangyaqing45@126.com

School of Mathematical Sciences and Institute of Mathematics, Nanjing Normal University, Nanjing 210023, Jiangsu, People's Republic of China – and – Center of Nonlinear Science, Nanjing University, Nanjing 210093, Jiangsu, People's Republic of China
E-mail address: ecchen@njnu.edu.cn