MIXING VIA SEQUENCE ENTROPY

Wen Huang, Song Shao and Xiangdong Ye

University of Science and Technology of China

June 29, 2004

Abstract. We investigate systematically several mixing concepts via sequence entropy both in measure-theoretical dynamical systems and in topological dynamical systems, and obtain some new theorems or new proofs of some known theorems. Particularly, characterizations of rigidity and mild mixing via sequence entropy are obtained. Moreover, we show that if the topological entropy of an open cover is positive, then the sequence entropy of the cover with respect to any sequence is positive. As applications we prove that minimal topological Kolmogorov system is strongly mixing and any transitive diagonal flow is mildly mixing.

§1 Introduction

It exhibits a remarkable parallelism between measure-theoretical dynamical systems (MDS for short) and topological dynamical systems (TDS for short). There are many concepts which have counterparts in both theories, e.g., ergodicity - minimality, discrete spectrum - equicontinuity, relatively discrete spectrum - almost periodic extension, measure-theoretical entropy - topological entropy etc. These have led to the formulation of analogous theorems in both theories. In this paper we study several mixing concepts using the notion of the sequence entropy both in MDS and TDS. We will find the results in both theories are analogous, but the proofs involved are entirely different and neither directly deducible from the other.

Let $(X, \mathcal{B}, \mu, T)$ be an invertible MDS. Ergodicity and all kinds of mixing properties are discussed by many authors from different viewpoints. It is found that to characterize different mixing properties using entropy concept is effective and fruitful. This study was originated by Kushnirenko [Ku] in 1965. Let $\mathcal{IF}$ be the set of all infinite sequences of $\mathbb{Z}_+$. Kushnirenko introduced the notion of sequence entropy along a given $A \in \mathcal{IF}$ for a MDS and proved that an invertible MDS has discrete

Key words and phrases. mixing, sequence entropy, sequence entropy pair.

2000 *Mathematics Subject Classification:* Primary: 37A25, 37B40.

spectrum iff the sequence entropy of the system with respect to any $A \in \mathcal{IF}$ is zero. Later Saleski gave a characterization of weakly mixing and strongly mixing MDS via sequence entropy [S]. Let $\mathcal{P}_X$ be the set of finite measurable partitions of X. He showed that T is weakly mixing iff $\sup_{A\in\mathcal{IF}} h_\mu^A(T,\alpha) = H_\mu(\alpha)$ for all $\alpha \in \mathcal{P}_X$; T is strongly mixing iff $\sup\{h_\mu^B(T,\alpha) : B \subseteq A\} = H_\mu(\alpha)$ for any $A \in \mathcal{IF}$ and any $\alpha \in \mathcal{P}_X$. Hulse [Hu1] improved some of the Saleski's results and showed that T is weakly mixing iff there exists $A \in \mathcal{IF}$ such that $h_\mu^A(T,\alpha) = H_\mu(\alpha)$ for any $\alpha \in \mathcal{P}_X$; T is strongly mixing iff for any $A \in \mathcal{IF}$ there exists an infinite sequence $B \subseteq A$ such that $h_\mu^B(T,\alpha) = H_\mu(\alpha)$ for all $\alpha \in \mathcal{P}_X$. Moreover, Hulse [Hu2] gave the characterizations of the compact and weakly mixing extensions of MDS via conditional sequence entropy.

Between weak mixing and strong mixing concepts there is an important mixing property, known as mild mixing in ergodic theory. It is well known that an invertible MDS $(X,\mathcal{B},\mu.T)$ is weakly mixing if and only if for any ergodic MDS $(Y,\mathcal{A},\nu,S)$ with $\nu(Y) < \infty$, the product X and Y is ergodic. But if we drop the assumption that $\nu(Y) < \infty$, then the property is strictly stronger than weak mixing, which is called mild mixing by Furstenberg and Weiss [FW]. In [F2,F3] many equivalence conditions of mild mixing were obtained. Moreover, Furstenberg and Kateznelson [FK] introduced the relative version of the notion. The characterizations of mild mixing and mildly mixing extension via sequence entropy can be found in [Z1,Z2]. For example, Zhang [Z1] showed that T is mildly mixing iff for every $A \in \mathcal{IF}$, there is a sequence F_n of pairwise disjoint finite subsets of A such that for any $\alpha \in \mathcal{P}_X$, $h_\mu^{\{s_i\}}(T,\alpha) = H_\mu(\alpha)$, where $s_i = \sum_{a\in F_i} a$.

It is well known that the notion of discrete spectrum is contrary to that of weak mixing which can be characterized as having no nontrivial factors with discrete spectrum for an invertible MDS. Similarly, the notion of rigidity is contrary to that of mild mixing which can be characterized as having no nontrivial rigid factors for an invertible MDS. We have mentioned Kushnirenko' theorem on the characterization of discrete spectrum via sequence entropy. Now we recall that Zhang [Z1] showed that an invertible MDS $(X,\mathcal{B},\mu,T)$ is rigid iff there exists $A \in \mathcal{IF}$ such that if $\{F_n\}$ is any sequence of pairwise disjoint finite subsets of A and $s_i = \sum_{a\in F_i} a$, then $h_\mu^{\{s_i\}}(T) = 0$.

In section 3, we will give a new proof of Kushnirenko's theorem and some proofs of the above mentioned theorems using different approach (see Theorems 3.5, 3.6, 3.7, 3.11) and obtain characterizations of mixing and rigidity via sequence entropy (Theorems 3.6, 3.10).

In the category of topological dynamical system, in 1974 Goodman [G] introduced the notion of topological sequence entropy and studied some properties of null systems which are defined as having zero topological sequence entropy for any infinite sequence. It is a natural question whether we have similar characterizations of topological mixing properties using topological sequence entropy. In [HLSY], the first characterization of topological weak mixing was obtained using sequence entropy.

Namely, the authors localized the notion of sequence entropy by defining sequence entropy pairs and proved that a system is topologically weakly mixing iff any pair not in the diagonal is a sequence entropy pair. Moreover, they showed that for a minimal system Kushnirenko' statement remains true modulo an almost one to one extension, i.e. if a minimal system is null, then it is an almost one to one extension of a topological system with discrete spectrum .

Recently, in [HY2] (see also [GW]) the notion of topological mild mixing was introduced. A TDS (X,T) is topologically mildly mixing if for any transitive system (Y,S), $(X\times Y, T\times S)$ is transitive. Similar to the case in ergodic theory, this property is strictly between weak mixing and strong mixing. A characterization of topological mild mixing via sequence entropy will be obtained in section 4 and new proofs of known theorems are given (see Theorem 4.11, 4.12 and 4.13):

The notion of Kolmogorov system in ergodic theory is useful and is extensively studied. The notion of *disjointness* of two TDS was introduced in [F1] to show how great the two systems are differ from each other both in MDS and TDS. It is known that any K-system in MDS is disjoint from zero entropy system and is strongly mixing. Topological analogues (u.p.e. in [B1], and topological K in [HY1]) of K-system were introduced in [B1] and [HY1]. In [B2], it was proved that diagonal flows (which is weaker than u.p.e.) are disjoint from all minimal TDS with zero entropy. It is an open question if minimal u.p.e. implies strong mixing. In the last section we show that if the topological entropy of an open cover is positive, then the sequence entropy of the cover with respect to any sequence is positive (a similar result in MDS was obtained in [S] and we supply a direct proof for completeness). As applications we prove that any minimal topological K-system is strongly mixing and any transitive diagonal flow [B2] is mildly mixing.

§2 Preliminaries

We use $\mathbb{Z}$ (resp. $\mathbb{R}$, $\mathbb{C}$ and $\mathbb{N}$) to denote the set of integers (resp. real numbers, complex numbers and natural numbers) and $\mathbb{Z}_+$ the set of the non-negative integers. For convenience when we speak $A \in \mathcal{IF}$ we mean it is ordered in the natural way, i.e., $A = \{a_1 < a_2 < \cdots\}$.

Let $(X, \mathcal{B}, \mu)$ be a standard Borel space, μ a regular probability measure on X and $T : X \to X$ a measure-preserving transformation. The quadruple $(X, \mathcal{B}, \mu, T)$ is said to be a *measure-theoretical dynamical system* (for short MDS) if $T\mu = \mu$, i.e. $\mu(B) = \mu(T^{-1}B)$ for all $B \in \mathcal{B}$. If T is measure-preserving, bijective, and T^{-1} is also measure-preserving, we say $(X, \mathcal{B}, \mu, T)$ *invertible.* A MDS $(X, \mathcal{B}, \mu, T)$ is *ergodic* if the only measurable sets A for which $\mu(A\Delta T^{-1}A) = 0$ satisfy $\mu(A) = 0$ or $\mu(A) = 1$. A MDS $(X, \mathcal{B}, \mu, T)$ is *weakly mixing* if $(X \times X, \mathcal{B} \times \mathcal{B}, \mu \times \mu, T \times T)$ is ergodic. $(X, \mathcal{B}, \mu, T)$ is *strongly mixing* if $\lim_{n\to\infty} \mu(T^{-n}A \cap B) = \mu(A)\mu(B)$ for any $A, B \in \mathcal{B}$.

An *eigenfunction* for T is some non-zero function $f \in L^2(X, \mathcal{B}, \mu) = L^2(\mu)$ such that $Uf = Tf = \lambda f$ for some $\lambda \in \mathbb{C}$. λ is called the *eigenvalue corresponding to f.*

It is easy to see every eigenvalue has norm one, i.e. $|\lambda| = 1$. It is well known that T is ergodic iff 1 is a simple eigenvalue of T. If $f \in L^2(\mu)$ is an eigenfunction, then $cl\{U^n f : n \in \mathbb{Z}\}$ is a compact subset of $L^2(\mu)$. Generally, we say f *almost periodic* if $cl\{U^n f : n \in \mathbb{Z}\}$ is compact in $L^2(\mu)$. It is well known that the set of all bounded almost periodic functions forms a U-invariant and conjugated-invariant algebra of $L^2(\mu)$ (denoted by A_c). The set of all almost periodic functions is just the closure of A_c (denoted by H_c), and is also spanned by the set of eigenfunctions. The following proposition is a classical result (see for example [Zi]).

Proposition 2.1. *Let $(X, \mathcal{B}, \mu, T)$ be a MDS and H be a conjugated-invariant algebra of $L^2(\mu)$ consisting of bounded functions. Then there exists a sub-σ-algebra $\mathcal{A}$ of $\mathcal{B}$ such that $cl(H) = L^2(X, \mathcal{A}, \mu)$. Moreover if H is U-invariant, then $\mathcal{A}$ is T-invariant.*

By Proposition 2.1, there exists a T-invariant sub-σ-algebra $\mathcal{K}$ of $\mathcal{B}$ such that $H_c = L^2(X, \mathcal{K}, \mu)$. We call $\mathcal{K}$ the *Kronecker algebra of* $(X, \mathcal{B}, \mu, T)$. It is well known that T is weakly mixing iff H_c is trivial iff the constants are the only eigenfunctions for T, i.e. the Korenecker algebra $\mathcal{K}$ is trivial. T is said to be *compact* or to have *discrete spectrum* if $L^2(\mu)$ can be spanned by the set of eigenfunctions, i.e. $H_c = L^2(X, \mathcal{B}, \mu)$ or equivalently $\mathcal{K} = \mathcal{B}$.

A function $f \in L^2(\mu)$ is called *rigid* if there is $\{t_n\} \in \mathcal{IF}$ with $\lim T^{t_n} f = f$ in L^2-norm. For a fixed $F = \{t_n\} \in \mathcal{IF}$, it is easy to see that the set of all bounded functions $f \in L^2(\mu)$ with $\lim T^{t_n} f = f$ in L^2-norm forms a U-invariant and conjugated-invariant algebra of $L^2(\mu)$ (denoted by A_F). The set of all functions $f \in L^2(\mu)$ with $\lim T^{t_n} f = f$ in L^2-norm is just the closure of A_F (denoted by H_F). By Proposition 2.1, there exists T-invariant sub-σ-algebra $\mathcal{K}_F$ of $\mathcal{B}$ such that $H_F = L^2(X, \mathcal{K}_F, \mu)$.

A MDS $(X, \mathcal{B}, \mu, T)$ is called *rigid* if there is $F = \{t_n\} \in \mathcal{IF}$ with $H_F = L^2(\mu)$, i.e. $\lim T^{t_n} f = f$ in L^2-norm for all $f \in L^2(\mu)$. A MDS $(X, \mathcal{B}, \mu, T)$ is called *mild mixing* if it has no nonconstant rigid function (this definition is equivalent to that $\mathcal{K}_F$ is trivial for each $F \in \mathcal{IF}$, and is also equivalent to the condition described in the first section [FW, F3]). Obviously, for a strongly mixing MSD $\mathcal{K}_F$ is trivial for each $F \in \mathcal{IF}$, i.e. strong mixing implies mild mixing. On the other hand, since every eigenfunction is rigid, mild mixing implies weak mixing. In fact, it can be showed that mild mixing is strictly between weak mixing and strong mixing [FW].

By a *topological dynamical system* (for short TDS) we mean a pair (X, T), where X is a compact metric space and T is a continuous surjective map from X to X. When T is a homeomorphism, we say T *invertible.* The set of T-invariant probability measures defined on Borel sets of X, $\mathcal{B}(X)$, is denoted by $\mathcal{M}_T(X)$. In context measurability will be always related to $\mathcal{B}(X)$. Every invariant probability measure $\mu \in \mathcal{M}_T(X)$ induces a MDS $(X, \mathcal{B}(X), \mu, T)$.

(X, T) is *transitive* if for each pair of opene (i.e. open and nonempty) subsets U and V of X, there is $n \in \mathbb{N}$ such that $U \cap T^{-n}V \neq \emptyset$, and (X, T) is (topologically) *weakly mixing* if $(X \times X, T \times T)$ is transitive. (X, T) is (topologically) *strongly mixing* if for any opene subsets U and V of X, there is $N \in \mathbb{N}$ such that $U \cap T^{-n}V \neq \emptyset$

for each $n \geq N$. A TDS (X,T) is (topologically) *mildly mixing* if for any transitive system (Y,S), $(X \times Y, T \times S)$ is transitive.

Let $\mathcal{F}$ be a collection of subsets of $\mathbb{Z}_+$. If it is hereditary upward, i.e. $F_1 \subseteq F_2$ and $F_1 \in \mathcal{F}$ imply $F_2 \in \mathcal{F}$, then $\mathcal{F}$ is said to be *a family.* If a family $\mathcal{F}$ is closed under intersection and satisfies $\emptyset \notin \mathcal{F}$, then it is called a *filter*. For a family $\mathcal{F}$ its *dual* is $\mathcal{F}^* = \{F \subseteq \mathbb{Z}_+ | F \cap F' \neq \emptyset \ for\ all\ F' \in \mathcal{F}\}$. The dual family of $\mathcal{IF}$ is the set of cofinite subsets of $\mathbb{Z}_+$. The *upper density of* $A \in \mathcal{IF}$ is $\bar{d}(A) = \limsup_{N\to\infty} |A \cap \{0,1,\cdots,N-1\}|/N$. The *lower density* $\underline{d}(A)$ are similarly defined. If $\bar{d}(A) = \underline{d}(A)$, then we say A has *density* $d(A)$. Let $\mathcal{D} = \{A \subseteq \mathbb{Z}_+ | d(A) = 1\}$. It is easy to see $\mathcal{D}$ is a filter and $\mathcal{D}^* = \{A \subseteq \mathbb{Z}_+ | \bar{d}(A) > 0\}$. Let $\{b_i\}_{i=1}^I$ with $b_i \in \mathbb{N}$, where $I \in \mathbb{N}$ or $I = +\infty$, we define $FS(\{b_i\}_{i=1}^I) = \{\sum_{i\in\alpha} b_i : \alpha$ is a finite non-empty subset of $\{1,2,\cdots,I\}$ or $\mathbb{N}$ $\}$. F is an *IP set* if there exists sequence of natural number $\{b_i\}_{i=1}^\infty$ such that $F = FS(\{b_i\}_{i=1}^\infty)$. Denote the set of all IP sets by $\mathcal{IP}$.

Let $\{x_n\}$ be a sequence of a metric space (X,d), $x \in X$ and $\mathcal{F}$ be a family. We say x_n *$\mathcal{F}$-converges to* x, denoted by $\mathcal{F} - \lim x_n = x$, if for any neighborhood U of x, $\{n : x_n \in U\} \in \mathcal{F}$. The following is a well known result concerning mixing in ergodic theory (see [F2,F3,W1]).

Theorem 2.2. *Let $(X,\mathcal{B},\mu,T)$ be a MDS. Then*
1. T is weak mixing iff $\mathcal{D} - \lim_{n\to\infty} \mu(T^{-n}A \cap B) = \mu(A)\mu(B)$ for any $A,B \in \mathcal{B}$;
2. T is mild mixing iff $\mathcal{IP}^* - \lim_{n\to\infty} \mu(T^{-n}A \cap B) = \mu(A)\mu(B)$ for any $A,B \in \mathcal{B}$;
3. T is strong mixing iff $\mathcal{IF}^* - \lim_{n\to\infty} \mu(T^{-n}A \cap B) = \mu(A)\mu(B)$ for any $A,B \in \mathcal{B}$.

In the topological setting we have a similar description. Let (X,T) be a TDS and $U,V \subseteq X$. We define the *return times set* $N(U,V) = \{n \in \mathbb{Z}_+ : U \cap T^{-n}(V) \neq \emptyset\}$. Then we have ([F1,HY2,GW])

Theorem 2.3. *Let (X,T) be a TDS. Then*
1. (X,T) is weakly mixing iff $N(U,V)$ is thick (i.e. containing arbitrarily long intervals of $\mathbb{Z}_+$) for any opene sets U,V of X.
2. (X,T) is mildly mixing iff $N(U,V)$ has non-empty intersection with $F-F$ for any opene sets U,V of X and any IP set F, where $F-F = \{a-b : a,b \in F$ and $a \geq b\}$.
3. (X,T) is strongly mixing iff $N(U,V)$ is cofinite for any opene sets U,V of X.

Remark 2.4. 1. Assume (X,T) is a TDS and there is some $\mu \in \mathcal{M}_T(X)$ with full support. Then it is easy to see that if T is weakly mixing (resp. mildly mixing, strongly mixing) with respect to μ, then it is topologically weakly mixing (resp. mildly mixing, strongly mixing).

2. Note that if (X,T) is minimal then it is weakly mixing iff $N(U,V) \in \mathcal{D}$, and it is mildly mixing iff $N(U,V) \in \mathcal{IP}^*$ [HY2].

Now we define sequence entropy. Let $S = \{0 \leq t_1 < t_2 < \cdots\} \in \mathcal{IF}$ and $\mathcal{U}$ be a finite open cover of X. The *topological sequence entropy of $\mathcal{U}$ with respect to (X,T) along S* is defined by $h^S_{\text{top}}(T,\mathcal{U}) = \limsup_{n\to\infty} \frac{1}{n} \log N(\bigvee_{i=1}^n T^{-t_i}\mathcal{U})$, where

$N(\bigvee_{i=1}^n T^{-t_i}\mathcal{U})$ is the minimal cardinality among all cardinalities of subcovers of $\bigvee_{i=1}^n T^{-t_i}\mathcal{U}$. The *topological sequence entropy of* (X,T) *along sequence* S is $h^S_{\text{top}}(T) = \sup_{\mathcal{U}} h^S_{\text{top}}(T,\mathcal{U})$, where supremum is taken over all finite open covers of X. If $S = \mathbb{Z}_+$ we recover standard topological entropy. In this case we omit the superscript $\mathbb{Z}_+$.

Let $(X, \mathcal{B}, \mu, T)$ be a MDS and suppose ξ and η are two finite partitions of X. The *entropy of* ξ, written $H(\xi)$, is defined by the formula $H_\mu(\xi) = -\sum_{A\in\xi}\mu(A)\log\mu(A)$ and the *entropy of* ξ *given* η, written $H(\xi|\eta)$, is defined by the formula

$$H_\mu(\xi|\eta) = H_\mu(\xi \vee \eta) - H_\mu(\eta) = -\sum_{B\in\eta}\sum_{A\in\xi}\mu(A\cap B)\log\frac{\mu(A\cap B)}{\mu(B)}.$$

The *sequence entropy of* ξ *with respect to* (X,μ,T) *along* S is defined by

$$h^S_\mu(T,\xi) = \limsup_{n\longrightarrow+\infty}\frac{1}{n}H_\mu(\bigvee_{i=1}^n T^{-t_i}\xi)(= \limsup_{n\longrightarrow+\infty}\frac{1}{n}\sum_{j=2}^n H_\mu(T^{-t_i}\xi|\bigvee_{i=1}^{j-1}T^{-t_i}\xi)).$$

And the *sequence entropy of* $(X, \mathcal{B}, T, \mu)$ *along* S is $h^S_\mu(T) = \sup_\alpha h^S_\mu(T,\alpha)$, where supremum is taken over all finite measurable partitions. As in the topological case, when $S = \mathbb{Z}_+$ we recover entropy of T with respect to μ. In this case we omit the superscript $\mathbb{Z}_+$.

§3 MIXING AND SEQUENCE ENTROPY FOR A MEASURE

In this section, we will give some characterizations of mixing properties via sequence entropy for a MDS. First we give a new proof of a well known result by Kushnirenko [K]. In the process to do it we obtain a lemma which is crucial for the sequel.

Let $(X, \mathcal{B}, T, \mu)$ be an invertible MDS. We define a unitary operator $U : H \to H$ by $U(f) = f\circ T$, where $H = L^2(X,\mathcal{B},\mu)$. Let $F \in \mathcal{IF}$. Since H is a separable metric space, it is easy to see that there exists $S = \{s_1 < s_2 < \cdots\} \subseteq F$ such that $\lim_{i\to\infty} < g, U^{s_i}f >$ exists for any $f, g \in H$. Fixed $f \in H$, we define $J_f : H \to \mathbb{C}$ with $J_f(g) = \lim_{i\to\infty} < g, U^{s_i}f >$. Then J_f is a continuous linear functional on H. By the Riesz representation theorem, there exists $S(f) \in H$ such that $J_f(g) =< g, S(f) >$. Clearly, if $f \geq 0$ then $S(f) \geq 0$. First we have

Lemma 3.1. *Given* $\alpha = \{A_1, ..., A_k\} \in \mathcal{P}_X$ *there exists an infinite subsequence* $S' \subseteq S$ *such that* $h^{S'}_\mu(T,\alpha) \geq \sum_{A\in\alpha}\int_X -S(1_A)\log S(1_A)d\mu$.

Proof. First, we have

Claim: For any $\beta = \{B_1, ..., B_l\} \in \mathcal{P}_X$ and $\epsilon > 0$, there exists $M \in \mathbb{N}$ such that when $m \in S$ and $m \geq M$, $H_\mu(T^{-m}\alpha|\beta) \geq \sum_{A\in\alpha}\int_X -S(1_A)\log S(1_A)d\mu - \epsilon$.

Proof of the claim. Since $\lim_{n\to\infty}\mu(T^{-s_n}A_i\cap B_j) =< S(1_{A_i}), 1_{B_j} >$ for $1\le i\le k$ and $1\le j\le l$,

$$\lim_{n\to+\infty} H_\mu(T^{-s_n}\alpha|\beta) = \lim_{n\to+\infty}\sum_{i,j} -\mu(T^{-s_n}A_i\cap B_j)\log(\frac{\mu(T^{-s_n}A_i\cap B_j)}{\mu(B_j)}),$$
$$= \sum_{i,j} - < S(1_{A_i}), 1_{B_j} > \log(\frac{< S(1_{A_i}), 1_{B_j} >}{\mu(B_j)}).$$

Let $a_{ij} = - < S(1_{A_i}), 1_{B_j} > \log(< S(1_{A_i}), 1_{B_j} > /\mu(B_j))$. Since $\varphi(x) = -x\log x$ is concave and $< S(1_{A_i}), 1_{B_j} > /\mu(B_j) = \int_{B_j} S(1_{A_i})d\mu_{B_j}$, where $\mu_{B_j}(\cdot) = \mu(\cdot\cap B_j)/\mu(B_j)$, we deduce that

$$a_{ij} = \mu(B_j)\varphi(\int_{B_j} S(1_{A_i})d\mu_{B_j}) \ge \mu(B_j)\int_{B_j}\varphi(S(1_{A_i}))d\mu_{B_j} = \int_{B_j}\varphi(S(1_{A_i}))d\mu.$$

We conclude that

$$\sum_{i,j} a_{ij} \ge \sum_{i,j}\int_{B_j}\varphi(S(1_{A_i}))d\mu = \sum_i\int_X \varphi(S(1_{A_i}))d\mu$$

and $\lim_{n\to+\infty} H_\mu(T^{-s_n}\alpha|\beta) \ge \sum_i\int_X\varphi(S(1_{A_i}))d\mu$. This finishes the proof of the claim.

Now we may define an increasing sequence $S' = \{0\le t_1 < t_2 < ...\}\subseteq S$ such that

$$H_\mu(T^{-t_r}\alpha|\bigvee_{i=1}^{r-1} T^{-t_i}\alpha) \ge \sum_{A\in\alpha}\int_X -S(1_A)\log S(1_A)d\mu - \frac{1}{2^r},\ \forall r\in\mathbb{N}.$$

Therefore,

$$h_\mu^{S'}(T,\alpha) = \limsup_{r\to+\infty}\frac{1}{r}\sum_{k=2}^{r} H_\mu(T^{-t_k}\alpha|\bigvee_{i=1}^{k-1}T^{-t_i}\alpha)$$
$$\ge \limsup_{r\to+\infty}\frac{1}{r}\sum_{k=2}^{r}(\sum_{A\in\alpha}\int_X -S(1_A)\log S(1_A)d\mu - \frac{1}{2^k})$$
$$= \sum_{A\in\alpha}\int_X -S(1_A)\log S(1_A)d\mu.$$

$\square$

The following lemma is well known (see for example Lemma 4.15 in [W1]).

Lemma 3.2. *Let $r \geq 1$ be a fixed integer. For each $\epsilon > 0$ there exists $\delta > 0$ such that if $\xi = \{A_1, \cdots, A_r\}, \eta = \{C_1, \cdots, C_r\}$ are two partitions of $(X, \mathcal{B}, \mu)$ into r sets with $\sum_{i=1}^{r} \mu(A_i \Delta C_i) < \delta$ then the Rohlin metric $\rho(\xi, \eta) = H_\mu(\xi|\eta) + H_\mu(\eta|\xi) < \epsilon$.*

Lemma 3.3. *Let $(X, \mathcal{B}, \mu, T)$ be an invertible MDS, $B \in \mathcal{B}$ and $R \in \mathcal{IF}$. Then cl($\{U^n 1_B : n \in R\}$) is a compact set of $L^2(\mu)$ iff for each infinite sequence $A \subseteq R$, $h_\mu^A(T, \{B, B^c\}) = 0$.*

Proof. Assume cl($\{U^n 1_B : n \in R\}$) is compact in $L^2(\mu)$. Let $A = \{a_1 < a_2 < \cdots\} \subseteq R$ and $\alpha = \{B, B^c\}$. Note that $\|U^n 1_B - U^m 1_B\|_{L^2(\mu)} = \mu(T^{-n}B \triangle T^{-m}B)$. By Lemma 3.2 for any $\epsilon > 0$ there is some $N \in \mathbb{N}$ such that for any $n > N$ we can find some $i(n) \leq N$ with $\rho(T^{-a_n}\alpha, T^{-a_{i(n)}}\alpha) < \epsilon$.

Hence for any $n > N$ we have

$$H(T^{-a_n}\alpha | \bigvee_{i=1}^{n-1} T^{-a_i}\alpha) \leq H(T^{-a_n}\alpha | T^{-a_{i(n)}}\alpha) \leq \rho(T^{-a_n}\alpha, T^{-a_{i(n)}}\alpha) < \epsilon.$$

Then $h_\mu^A(T, \alpha) = \limsup_{n\to+\infty} \frac{1}{n} \sum_{k=2}^{n} H_\mu(T^{-a_k}\alpha | \bigvee_{i=1}^{k-1} T^{-a_i}\alpha) \leq \epsilon$. As ϵ is arbitrary, $h_\mu^A(T, \alpha) = 0$.

Now we show the converse. If cl($\{U^n 1_B : n \in R\}$) is not a compact set of $L^2(\mu)$, then there exist $\epsilon > 0$ and an infinite sequence $F \subseteq R$ such that for $a, b \in F, a \neq b$ one has $\|U^a 1_B - U^b 1_B\|_{L^2(\mu)} \geq \epsilon$. By the discussion at the beginning of the section, there exists $S = \{s_1 < s_2 < \cdots\} \subseteq F$ such that for any $f, g \in L^2(\mu)$,

$$\lim_{i\to\infty} < g, U^{s_i} f > = < g, S(f) > . \tag{3.3.1}$$

By Lemma 3.1, there exists an infinite subsequence $A \subseteq S$ such that

$$h_\mu^A(T, \{B, B^c\}) \geq \int_X (-S(1_B) \log S(1_B) - S(1_{B^c}) \log S(1_{B^c})) d\mu.$$

Since $h_\mu^A(T, \{B, B^c\}) = 0$, one has $-S(1_B)(x) \log S(1_B)(x) = 0$ for $x \in X$ μ-a.e. Therefore $S(1_B)$ is a characteristic function, so that $< 1_X, S(1_B) > = ||S(1_B)||_{L^2(\mu)}^2$. Take $f = 1_B$, $g = 1_X$ in (3.3.1), one has $\lim_{i\to\infty} < 1_X, U^{s_i} 1_B > = < 1_X, S(1_B) >$, which implies $||S(1_B)||_{L^2(\mu)}^2 = ||1_B||_{L^2(\mu)}^2$. Combining the fact with equation (3.3.1), one has $\lim_{i\to\infty} U^{s_i} 1_B = S(1_B)$. Hence for large enough $i > j$, one has $\|U^{s_i} 1_B - U^{s_j} 1_B\|_{L^2(\mu)} \leq \frac{\epsilon}{2}$ which contradicts the choose of F. $\square$

Taking $R = \mathbb{Z}_+$ in Lemma 3.3 we have

Kushnirenko's Theorem. *[K] An invertible MDS $(X, \mathcal{B}, \mu, T)$ has discrete spectrum iff $h_\mu^A(T) = 0$ for any $A \in \mathcal{IF}$.*

Theorem 3.5. *Let $(X, \mathcal{B}, \mu, T)$ be an invertible MDS. Then the following statements are equivalent*
1. $(X, \mathcal{B}, \mu, T)$ is weakly mixing;
2. for any $B \in \mathcal{B}$ with $0 < \mu(B) < 1$ and $F \in \mathcal{D}^*$, there exists an infinite sequence $A \subseteq F$ such that $h_\mu^A(T, \{B, B^c\}) > 0$.
3. for any finite non-trivial partition α of X and $F \in \mathcal{D}^*$, there exists an infinite sequence $A \subseteq F$ such that $h_\mu^A(T, \alpha) > 0$.

Proof. (1) $\Rightarrow$ (2) Assume there is some $B \in \mathcal{B}$ with $0 < \mu(B) < 1$ and $F = \{a_1 < a_2 < \cdots\} \in \mathcal{D}^*$ such that $h_\mu^A(T, \{B, B^c\}) = 0$ for any infinite sequence $A \subseteq F$. Then by Lemma 3.3 $\mathrm{cl}(\{U^n 1_B : n \in F\})$ is a compact set of $L^2(\mu)$. Thus for any $\epsilon > 0$ there is some $N \in \mathbb{N}$ such that for any $n > N$ we can find some $i(n) \le N$ with $\mu(T^{-a_n} B \Delta T^{-a_{i(n)}} B) = \|U^{a_n} 1_B - U^{a_{i(n)}} 1_B\|_{L^2(\mu)} < \epsilon$.

Since $(X, \mathcal{B}, \mu, T)$ is weakly mixing, $\{n \in \mathbb{Z}_+ : |\mu(T^{-n} B \cap T^{-a_i} B) - (\mu(B))^2| < \epsilon, \forall 1 \le i \le N\} \in \mathcal{D}$. Hence there is some $m > N$ such that $|\mu(T^{-a_m} B \cap T^{-a_i} B) - (\mu(B))^2| < \epsilon$ for $1 \le i \le N$. Particularly, $|\mu(T^{-a_m} B \cap T^{-a_{i(m)}} B) - (\mu(B))^2| < \epsilon$ and $\mu(T^{-a_m} B \Delta T^{-a_{i(m)}} B) < \epsilon$. Thus

$$\mu(B) = \mu(T^{-a_m} B) \le \mu(T^{-a_m} B \cap T^{-a_{i(m)}} B) + \mu(T^{-a_m} B \Delta T^{-a_{i(m)}} B) < (\mu(B))^2 + 2\epsilon.$$

Since ϵ is arbitrary, we have $\mu(B) \le (\mu(B))^2$ which contradicts the assumption.

(2) $\Rightarrow$ (1) If $(X, \mathcal{B}, \mu, T)$ is not weakly mixing, then the Korenecker algebra $\mathcal{K}$ of $(X, \mathcal{B}, \mu, T)$ is non-trivial. Hence there is some $B \in \mathcal{K}$ with $0 < \mu(B) < 1$ and $\mathrm{cl}(\{U^n 1_B : n \in \mathbb{Z}_+\})$ is a compact set of $L^2(\mu)$. By Lemma 3.3 this contradicts the assumption.

(2) $\Leftrightarrow$ (3) is clear, as if $\alpha = \{A_1, \ldots, A_n\} \in \mathcal{P}_X$, then $\bigvee_{i=1}^n \{A_i, A_i^c\} \succeq \alpha$. $\square$

Theorem 3.6. *Let $(X, \mathcal{B}, \mu, T)$ be an invertible MDS. Then the following statements are equivalent*
1. $(X, \mathcal{B}, \mu, T)$ is mildly mixing.
2. For any $B \in \mathcal{B}$ with $0 < \mu(B) < 1$ and IP-set F, there exists an infinite sequence $A \subseteq F$ such that $h_\mu^A(T, \{B, B^c\}) > 0$.
3. For any finite non-trivial partition α of X and IP-set F, there exists an infinite sequence $A \subseteq F$ such that $h_\mu^A(T, \alpha) > 0$.
4. For any $B \in \mathcal{B}$ with $0 < \mu(B) < 1$ and any infinite set F of $\mathbb{Z}_+$, there exists an infinite sequence $A \subseteq F$ such that $h_\mu^A(T, \{B, B^c\}) > 0$.
5. For any finite non-trivial partition α of X and any infinite set F of $\mathbb{Z}_+$, there exists an infinite sequence $A \subseteq F$ such that $h_\mu^A(T, \alpha) > 0$.

Proof. (2) $\Leftrightarrow$ (3), (4) $\Leftrightarrow$ (5) and (4) $\Rightarrow$ (2) are clear.

(1) $\Rightarrow$ (4) Let $(X, \mathcal{B}, \mu, T)$ be mildly mixing. Assume there exist a non-trivial partition $\alpha = \{B, B^c\}$ of X and $F \in \mathcal{IF}$ such that for any infinite sequence $A \subseteq F$, one has $h_\mu^A(T, \alpha) = 0$. By Lemma 3.3, $\mathrm{cl}(\{U^n 1_B : n \in F\})$ is a compact set of metric space $L^2(\mu)$. Hence, there exists $\{n_1 < n_2 < \cdots\} \subset F$ such that $\lim_{i \to \infty} T^{n_i} 1_B = g$

for some g. Now for any $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that when $i, j \geq N$ and $i > j$ then $||T^{n_i}1_B - T^{n_j}1_B||_{L^2(\mu)} \leq \epsilon$, that is, $||T^{n_i-n_j}1_B - 1_B||_{L^2(\mu)} \leq \epsilon$ for $i, j \geq N$ and $i > j$. Therefore we can find $F' = \{s_1 < s_2 < \cdots\}$ such that $\lim_{k\to\infty} T^{s_k}1_B = 1_B$, this shows that $\mathcal{K}_{F'}$ is non-trivial, a contradiction as $(X, \mathcal{B}, \mu, T)$ is mildly mixing.

(2) $\Rightarrow$ (1) Assume that for any finite non-trivial partition α of X and IP-set F, there exists infinite sequence $A \subseteq F$ such that $h^A_\mu(T, \alpha) > 0$. If $(X, \mathcal{B}, \mu, T)$ is not mildly mixing, then $F = \{n_1 < n_2 < \cdots\} \in \mathcal{IF}$ such that $\mathcal{K}_F$ is non-trivial. Hence there exists $B \in \mathcal{K}_F$ with $0 < \mu(B) < 1$ such that $\lim_{i\to\infty} T^{n_i}1_B = 1_B$. Without loss of generality, we assume $||T^{n_i}1_B - 1_B||_{L^2(\mu)} < \frac{1}{2^i}$ for any $i \in \mathbb{N}$(or we take subsequence of F).

Let $F = FS(\{n_i\}_{i=1}^\infty)$, then F is an IP-set. Since $||T^{n_i}1_B - 1_B||_{L^2(\mu)} < \frac{1}{2^i}$ for any $i \in \mathbb{N}$, it is easy to see that $\mathrm{cl}(\{U^n 1_B : n \in F\})$ is a compact set of $L^2(\mu)$. By Lemma 3.3 for each infinite sequence $A \subseteq F$, $h^A_\mu(T, \{B, B^c\}) = 0$, a contradiction. $\square$

Call a MDS $(X, \mathcal{B}, \mu, T)$ *intermixing* if $\liminf_{n\to\infty} \mu(T^{-n}A \cap B) > 0$ for any $A, B \in \mathcal{B}$ with $\mu(A) > 0$ and $\mu(B) > 0$ ([FO,W2]). It is easy to see that strong mixing is stronger than intermixing, and the example in [FO] showed the converse is not true.

Proposition 3.7. *If an invertible MDS $(X, \mathcal{B}, \mu, T)$ is intermixing, then it is mildly mixing.*

Proof. Let $(X, \mathcal{B}, \mu, T)$ be intermixing. If it is not mildly mixing, then there is $B \in \mathcal{B}$ with $0 < \mu(B) < 1$ and $A \in \mathcal{IF}$ such that for each infinite sequence $S \subseteq A$, $h^S_\mu(T, \{B, B^c\}) = 0$ by Theorem 3.6. By Lemma 3.3 $\mathrm{cl}(\{U^n 1_B : n \in A\})$ is a compact set of $L^2(\mu)$. Following the argument in the second part of the proof of Lemma 3.3 we get that there is some infinite sequence $\{s_i\} \subseteq A$ and $D \in \mathcal{B}$ such that $0 < \mu(D) < 1$ and $\lim_{i\to\infty} U^{s_i}1_B = 1_D$. Hence $\lim_{i\to\infty} \mu(T^{-s_i}B \cap D^c) = 0$, which implies that $\liminf_{n\to\infty} \mu(T^{-n}B \cap D^c) = 0$, a contradiction. $\square$

Now we turn our attention to the characterization of rigidity. First we need some preparation. Let $\mathcal{K}_\mu(F) = \{A \in \mathcal{B} : h^S_\mu(T, \{A, A^c\}) = 0 \text{ for any } S \subseteq F\}$. It is easy to see that $\mathcal{K}_\mu(F)$ is a T-invariant σ-algebra of X.

Let F be an IP-set generated by $n_1, n_2, \cdots$, i.e. $F = FS(\{n_i\}_{i=1}^\infty)$. Let $\mathcal{F}$ be the set of all finite subsets of $\mathbb{N}$. For $\alpha \in \mathcal{F}$, set $n_\alpha = \sum_{i\in\alpha} n_i$, then we have $F = \{n_\alpha\}_{\alpha\in\mathcal{F}}$. In the sequel by $F - \lim a_n = a$ we mean for any $\epsilon > 0$ there is some $\beta \in \mathcal{F}$ such that $|a_{n_\alpha} - a| < \epsilon$ for any $\alpha \in \mathcal{F}$ with $\alpha \cap \beta = \emptyset$. A subset F' of F is called an *IP-subsystem*, if there exist $\{\alpha_i\}_{i=1}^\infty \subseteq \mathcal{F}$ with $\alpha_i \cap \alpha_j = \emptyset, i \neq j$ such that $F' = FS(\{n_{\alpha_i}\}_{i=1}^\infty)$. Obviously, F' is an IP-subset of F.

Lemma 3.8. *[F3] Let $(X, \mathcal{B}, \mu, T)$ be an invertible MDS and F be an IP-set. Then there is an IP-subsystem F' of F and an orthogonal projection $\mathbb{P}$ on $L^2(\mu)$ such that*

$$F' - \lim < T^n f, g > = < \mathbb{P}f, g >, \quad \forall f, g \in L^2(\mu).$$

In particular, if $\mathbb{P}f = f$, then $F' - \lim \|T^n f - f\|_{L^2(\mu)} = 0$. Moreover, there is a T-invariant sub-σ-algebra $\mathcal{B}_0$ of $\mathcal{B}$ such that $\mathbb{P}(L^2(X, \mathcal{B}, \mu)) = L^2(X, \mathcal{B}_0, \mu)$.

Lemma 3.9. *Let $(X, \mathcal{B}, \mu, T)$ be an invertible MDS. For a given IP-set F, there exists an IP-system F' of F such that*
1. $F' - \lim T^n f = \mathbb{E}(f|K_\mu(F'))$ with respect to the weak topology for $f \in L^2(\mu)$,
2. $F' - \lim T^n \mathbb{E}(f|K_\mu(F')) = \mathbb{E}(f|K_\mu(F'))$ in $L^2(\mu)$-norm for $f \in L^2(\mu)$.

Proof. Let $F', \mathcal{B}_0$ as in Lemma 3.8. It remains to show $\mathcal{B}_0 = K_\mu(F')$. Let $F' = FS(\{n_i\}_{i=1}^\infty)$ be an IP-set.

If $A \in \mathcal{B}_0$, then $F' - \lim \|U^n 1_A - 1_A\|_{L^2(\mu)} = 0$. For any $\epsilon > 0$, there is some $M \in \mathbb{N}$ such that $\|U^n 1_A - 1_A\|_{L^2(\mu)} < \epsilon$ for any $n \in FS(\{n_i\}_{i=M}^\infty)$. It is easy to see that if $a \in F'$ then $a = b + c$ for some $b \in FS(\{n_i\}_{i=M}^\infty)$ and $c \in FS(\{n_i\}_{i=1}^{M-1})$. We have

$$\|U^a 1_A - U^c 1_A\|_{L^2(\mu)} = \|U^b 1_A - 1_A\|_{L^2(\mu)} < \epsilon.$$

Since ϵ is arbitrary, $\mathrm{cl}(\{U^n 1_A : n \in F'\})$ is compact in $L^2(\mu)$. By Lemma 3.3, $A \in K_\mu(F')$.

Conversely, let $A \in K_\mu(F')$. By Lemma 3.8, $\lim_{i\to\infty} < U^{n_i} f, g > = < \mathbb{P}f, g >$ for each $f, g \in L^2(\mu)$. By Lemma 3.1 there exists an infinite subset $S' \subseteq \{n_1, n_2, \cdots\}$ such that $h_\mu^{S'}(T, \{A, A^c\}) \geq \int_X [-\mathbb{P}(1_A) \log \mathbb{P}(1_A) - \mathbb{P}(1_{A^c}) \log \mathbb{P}(1_A^c)] d\mu$.

Since $A \in K_\mu(F')$, $h_\mu^{S'}(T, \{A, A^c\}) = 0$. Moreover we have that $\mathbb{P}(1_A)$ is a characteristic function and $\lim_{i\to\infty} U^{n_i} 1_A = \mathbb{P}(1_A)$ (following the argument in the second part of the proof of Lemma 3.3). Since $\mathbb{P}(1_A) \in L^2(\mathcal{B}_0)$ and $\mathcal{B}_0$ is T-invariant, one has $U^m \mathbb{P}(1_A) \in L^2(\mathcal{B}_0)$, $\forall m \in \mathbb{Z}$ (see Lemma 3.8). Since $||1_A - U^{-n_i}\mathbb{P}(1_A)||_{L^2(\mu)} = ||U^{n_i} 1_A - \mathbb{P}(1_A)||_{L^2(\mu)} \to 0$ when $n \to \infty$ and $L^2(\mathcal{B}_0)$ is a closed subspace of $L^2(\mu)$, $1_A \in L^2(\mathcal{B}_0)$, i.e. $A \in \mathcal{B}_0$. Thus we have $\mathcal{B}_0 = K_\mu(F')$. □

Now we ready to show

Theorem 3.10. *An invertible MDS $(X, \mathcal{B}, \mu, T)$ is rigid iff there is some IP set F such that $h_\mu^A(T) = 0$ for any infinite $A \subseteq F$.*

Proof. Assume that $(X, \mathcal{B}, \mu, T)$ is rigid. Then there exists an infinite sequence $\{n_i\}$ such that $\lim_{i\to\infty} U^{n_i} f = f$ in $L^2(\mu)$-norm for $f \in L^2(\mu)$. Let $\{f_j\}_1^\infty$ be a countable dense subset of $L^2(\mu)$. Without loss of generality (if necessity we pass to a subsequence), we can assume that

$$||U^{n_i} f_j - f_j||_{L^2(\mu)} < \frac{1}{2^i} \text{ for } 1 \leq j \leq i. \tag{3.10.1}$$

Let F be an IP-set generated by $\{n_i\}$. By (3.10.1) for each $j \in \mathbb{N}$, $\mathrm{cl}(\{U^n f_j : n \in F\})$ is a compact subset of $L^2(\mu)$. Moreover, $\mathrm{cl}(\{U^n f : n \in F\})$ is a compact subset of $L^2(\mu)$ for each $f \in L^2(\mu)$. By Lemma 3.3 $h_\mu^A(T) = 0$ for any $A \subseteq F$.

Conversely assume there is some IP set F such that $h_\mu^A(T) = 0$ for any $A \subseteq F$. By Lemma 3.9 there exists an IP-system F' of F such that

$(*)$ $\quad F' - \lim T^n \mathbb{E}(f|K_\mu(F')) = \mathbb{E}(f|K_\mu(F'))$ in $L^2(\mu)$-norm for $f \in L^2(\mu)$.

Since $h_\mu^A(T) = 0$ for any $A \subseteq F$, $K_\mu(F') = \mathcal{B}$. Hence $\mathbb{E}(f|K_\mu(F')) = f$ for $f \in L^2(\mu)$. Combing the fact and $(*)$, we have $F' - \lim T^n f = f$ in $L^2(\mu)$-norm for any $f \in L^2(\mu)$. Let F' be generated by $\{n_i\}_{i=1}^\infty$, then $\lim_{i\to\infty} T^{n_i} f = f$ in $L^2(\mu)$-norm for $f \in L^2(\mu)$, i.e. $(X, \mathcal{B}, \mu, T)$ is rigid. □

Finally we give another description for several kinds of mixing using sequence entropy. Recall that $\mathcal{P}_X$ be the set of finite measurable partitions of X and it is a complete separable metric space with Rohlin metric ρ (see Lemma 3.2 for definition). We remark that (1) and (3) appeared in [Hu1].

Theorem 3.11. *Let $(X, \mathcal{B}, \mu, T)$ be an invertible MDS. Then*

1. *T is weakly mixing iff for any $A \in \mathcal{D}^*$ there is some $B \subseteq A$ such that $h_\mu^B(T, \xi) = H_\mu(\xi)$ for all $\xi \in \mathcal{P}_X$.*
2. *T is mildly mixing iff for any IP set A there is some $B \subseteq A$ such that $h_\mu^B(T, \xi) = H_\mu(\xi)$ for all $\xi \in \mathcal{P}_X$.*
3. *T is strongly mixing iff for any infinite sequence A of $\mathbb{Z}_+$ there is some $B \subseteq A$ such that $h_\mu^B(T, \xi) = H_\mu(\xi)$ for all $\xi \in \mathcal{P}_X$.*

Proof. Assume T is weakly mixing (resp. mild mixing, strong mixing). Then we have $\mathcal{F} - \lim \mu(T^{-n}C \cap D) = \mu(C)\mu(D)$, $\forall C, D \in \mathcal{B}$, where $\mathcal{F}$ is $\mathcal{D}$ (resp. $\mathcal{IP}^*$, $\mathcal{IF}^*$). Moreover since $\mathcal{F}$ is a filter, for any $\alpha, \beta \in \mathcal{P}_X$ one has

$$\begin{aligned}
&\mathcal{F} - \lim H_\mu(T^{-n}\alpha|\beta) = \mathcal{F} - \lim H_\mu(T^{-n}\alpha \vee \beta) - H_\mu(\beta) \\
&= \mathcal{F} - \lim\Big(\sum_{C\in\alpha, D\in\beta} -\mu(T^{-n}C \cap D) \log \mu(T^{-n}C \cap D)\Big) - H_\mu(\beta) \\
&= \sum_{C\in\alpha, D\in\beta} -(\mu(C)\mu(D))) \log(\mu(C)\mu(D)))) - H_\mu(\beta) \\
&= H_\mu(\alpha). \qquad (*)
\end{aligned}$$

Let $\{\xi_k\}_{k=1}^\infty$ be a dense set of finite partitions in $\mathcal{P}_X$ and $A \in \mathcal{F}^*$. By (*) and the fact that $\mathcal{F}$ is a filter, we can define an increasing sequence $B = \{t_n\}_{n=1}^\infty \subseteq A$ such that

$$H_\mu(T^{-t_n}\xi_k | \bigvee_{i=1}^{n-1} T^{-t_i}\xi_k) \geq H_\mu(\xi_k) - 2^{-n}, \text{ for any } n \geq 2, 1 \leq k \leq n$$

Fix $k \in \mathbb{N}$. Then for $n > k$,

$$\begin{aligned}
H_\mu(\bigvee_{i=1}^{n} T^{-t_i}\xi_k) &= H_\mu(\bigvee_{i=1}^{k} T^{-t_i}\xi_k) + \sum_{i=k+1}^{n} H_\mu(T^{-t_i}\xi_k | \bigvee_{j=1}^{i-1} T^{-t_j}\xi_k) \\
&\geq H_\mu(\bigvee_{i=1}^{k} T^{-t_i}\xi_k) + (n-k)H_\mu(\xi_k) - \sum_{i=k+1}^{n} 2^{-i}.
\end{aligned}$$

Therefore, we have $h^B_\mu(T,\xi_k)=H_\mu(\xi_k)$ holds for any k. As $\{\xi_k\}$ is dense in $\mathcal{P}_X$, for any $\xi\in\mathcal{P}_X$ and any $\epsilon>0$ there is some ξ_k such that $\rho(\xi,\xi_k)<\epsilon$. Using $|h^B_\mu(T,\xi)-h^B_\mu(T,\eta)|\le\rho(\xi,\eta)$, we have $h^B_\mu(T,\xi)\ge h^B_\mu(T,\xi_k)-\epsilon\ge H_\mu(\xi)-2\epsilon$. As ϵ is arbitrary, we have $h^B_\mu(T,\xi)=H_\mu(\xi)$ for each $\xi\in\mathcal{P}_X$.

For the proof in the other direction, (1),(2) see Theorem 3.5, 3.6 and (3) by Saleski's Theorem [S]. □

§4 Topological mixing and topological sequence entropy

In this section, we will give the characterizations of several topological mixing properties via sequence entropy. First, we start with some notations. Let X be a topological space. By an *admissible cover* $\mathcal{U}$ we mean that $\mathcal{U}$ is finite and if $\mathcal{U}=\{U_1,U_2,\cdots,U_n\}$, then $(\bigcup_{j\neq i}U_j)^c$ has nonempty interior for each $i\in\{1,2,\cdots,n\}$. By an *admissible cover* $\mathcal{U}$ *separating* (x_1,x_2) we mean that if $\mathcal{U}=\{U_1,U_2\}$, then for each U_i $(1\le i\le 2)$ there exists x_{j_i} $(j_i=1,2)$ such that $x_{j_i}\notin cl(U_i)$. Similarly, we define *admissible partition* α with respect to (x_1,x_2).

Definition 4.1. *Let (X,T) be a TDS and $F\in\mathcal{IF}$. Then*

1. *(x_1,x_2) is called an F-sequence entropy pair if $x_1\neq x_2$ and for any admissible open cover $\mathcal{U}$ separating (x_1,x_2) there exists an infinite sequence $A\subseteq F$ such that $h^A_{\text{top}}(T,\mathcal{U})>0$. Denote by $SE(X,T,F)$ the set of all F-sequence entropy pairs. When $F=\mathbb{Z}_+$, we write simply $SE(X,T,F)$ as $SE(X,T)$.*
2. *(X,T) is said to have F-sequence uniform positive entropy ($F-$s.u.p.e. for short) if $SE(X,T,F)=X\times X\setminus\Delta_X$.*
3. *(X,T) is F-null if $SE(X,T,F)=\emptyset$.*

The following proposition states the basic properties of sequence entropy pairs with respect to a given $F\in\mathcal{IF}$. The proof is similar to the ones of the corresponding results in [B2].

Proposition 4.2. *Let (X,T) be a TDS and $F\in\mathcal{IF}$.*

a. *If $\mathcal{U}=\{U_1,U_2\}$ is an open cover of X with $h^A_{\text{top}}(T,\mathcal{U})>0$ for some infinite sequence $A\subseteq F$, then there exist $x_1\in U_1^c, x_2\in U_2^c$ with $(x_1,x_2)\in SE(X,T,F)$.*
b. *$SE(X,T,F)\cup\Delta_X$ is a closed $T\times T$-invariant subset of $X\times X$.*
c. *Let $\pi:(Y,S)\longrightarrow(X,T)$ be a factor map of TDS.*
   1. *If $(x_1,x_2)\in SE(X,T,F)$, then there exist $y_1,y_2\in Y$ such that $\pi(y_1)=x_1,\pi(y_2)=x_2$ and $(y_1,y_2)\in SE(Y,S,F)$.*
   2. *If $(y_1,y_2)\in SE(Y,S,F)$ and $\pi(y_1)\neq\pi(y_2)$, then one has $(\pi(y_1),\pi(y_2))\in SE(X,T,F)$.*
d. *Suppose W is a closed T-invariant subset of (X,T). If (x_1,x_2) is an F-sequence entropy pair of $(W,T|_W)$, then it is also an F-sequence entropy pair of (X,T).*

Remark. *In [B2] Blanchard showed that entropy pairs have the same properties. By Proposition 4.2 (a), we know that (X,T) is F-null iff for each infinite sequence*

$A \subseteq F$, $h^A_{\text{top}}(T) = 0$. *In the next section we will show that any F-null system has zero topological entropy.*

Using Proposition 4.2 one easily gets

Theorem 4.3. *For a TDS (X, T) and $F \in \mathcal{IF}$ the smallest closed invariant equivalence relation containing $SE(X, T, F)$ induces the maximal F-null factor (X_F, T_F).*

Disjointness of two TDS is defined in [F1]. If (X, T) and (Y, S) are two TDS we say $J \subseteq X \times Y$ is a *joining* of X and Y if J is a non-empty closed invariant set and is projected onto X and Y. If each joining is equal to $X \times Y$ we then say that (X, T) and (Y, S) are *disjoint* or $(X, T) \perp (Y, S)$ or $X \perp Y$. Following [B2] it is easy to prove

Theorem 4.4. *Each F-s.u.p.e. system is disjoint from any minimal F-null system.*

For any $F \in \mathcal{IF}$ we define F-transitive and F-mixing as follows.

Definition 4.5. *Let (X, T) be a TDS and $F \in \mathcal{IF}$. (X, T) is called F-transitive if for each pair of opene sets U, V, there exists $n \in F$ such that $T^{-n}V \cap U \neq \emptyset$. If $(X \times X, T \times T)$ is F-transitive, then (X, T) is said to be F-mixing.*

Clearly, an F-mixing system is weakly mixing. When $F = \mathbb{Z}_+$ we omit F in the notation. The following is the classical description for weak mixing which we will use in the section.

Theorem 4.6. *([F1,Ak]) Let (X, T) be a TDS, then the following conditions are equivalent.*
1. *(X, T) is weakly mixing.*
2. *$N(U, U) \cap N(U, V) \neq \emptyset$ for every opene sets U, V in X.*
3. *For every opene sets U_1, U_2, V_1, V_2 in X, there exist opene sets U, V in X such that $N(U, V) \subseteq N(U_1, V_1) \cap N(U_2, V_2)$.*
4. *$\{N(U, V) | U, V$ are opene sets in $X\}$ generates a filter.*

By this theorem we can get the following result readily:

Proposition 4.7. *Let (X, T) be a TDS. Then the following statements are equivalent:*
1. *(X, T) is F-mixing;*
2. *(X, T) is weakly mixing and F-transitive.*
3. *$N(U, U) \cap N(U, V) \cap F \neq \emptyset$ for any opene sets U, V.*

For an $F \in \mathcal{IF}$, F-s.u.p.e. and mixing are connected by the following theorem. Note that $F - F = \{a - b \geq 0 : a, b \in F\}$.

Lemma 4.8. *Let (X, T) be a TDS and $F \in \mathcal{IF}$. If (X, T) has F-s.u.p.e., then it is $(F - F)$-mixing.*

Proof. Let (X, T) have F-s.u.p.e.. If (X, T) is not $(F - F)$-mixing, then by Proposition 4.7 there exist opene sets U_1, U_2 with $N(U_1, U_1) \cap N(U_1, U_2) \cap (F - F) = \emptyset$.

As $0 \in (F-F)$, $U_1 \cap U_2 = \emptyset$. Take $x_i \in U_i$ and a closed neighborhood $V_i \subseteq U_i$ of x_i, $i = 1,2$. Then $\mathcal{V} = \{V_1^c, V_2^c\}$ is an open cover of X. Since for any $n \in (F-F)$ we have $U_1 \cap T^{-n}U_1 = \emptyset$ or $U_1 \cap T^{-n}U_2 = \emptyset$, there is a sequence $\{W_n\}_{n\in(F-F)}$ and $W_n = V_1^c$ or $W_n = V_2^c$ such that $V_1 \subseteq T^{-n}W_n$ for each $n \in (F-F)$.

For any sequence $A = \{0 < t_1 < t_2 < \cdots\} \subseteq F$ and $n \in \mathbb{N}$, consider for each $x \in X$ the first $i \in \{1,2,\cdots,n\}$ such that $T^{t_i}x \in V_1$, when there exists one. We get that the $\mathcal{V}_n = \bigvee_{i=1}^n T^{-t_i}\mathcal{V}$ admits a subcover by the sets

$$T^{-t_1}V_1^c \cap \cdots \cap T^{-t_{i-1}}V_1^c \cap T^{-t_i}W_0 \cap T^{-t_{i+1}}W_{t_{i+1}-t_i} \cap \cdots \cap T^{-t_n}W_{t_n-t_i},$$

$i = 1,2,\cdots,n$, and $\bigcap_{i=1}^n T^{-t_i}V_1^c$. Hence for all $n \in \mathbb{N}$, $H(\mathcal{V}_n) \le n+1$. So we have $h^A_{\text{top}}(T,\mathcal{V}) = 0$. This shows that $(x_1,x_2) \notin SE(X,T,F)$ which contradicts the fact that (X,T) has F-s.u.p.e.. Thus (X,T) is $(F-F)$-mixing. $\square$

Remark. *By the above proof, it is easy to see that if $(x_1,x_2) \in SE(X,T,F)$ then $N(U_1,U_1) \cap N(U_1,U_2) \cap (F-F) \neq \emptyset$ for any open neighborhood U_i of x_i, $i = 1,2$.*

Let $\mathcal{U}$ be a finite cover of X. Define $H(\mathcal{U}) = \log N(\mathcal{U})$. We say $\mathcal{U}$ *topologically non-trivial*, if each element of $\mathcal{U}$ is not dense in X.

Lemma 4.9. *Let (X,T) be a TDS, $F = FS(\{p_i\}_{i=1}^\infty)$ and $F_n = FS(\{p_i\}_{i=n}^\infty)$, $n \in \mathbb{N}$. If for each $n \in \mathbb{N}$, (X,T) is $(F_n - F_n)$-mixing, then for each admissible open cover $\mathcal{U}$ (resp. non-trivial finite open cover $\mathcal{V}$) there exists an infinite sequence $A \subseteq F$ such that $h^A_{\text{top}}(T,\mathcal{U}) = H(\mathcal{U})$ (resp. $h^A_{\text{top}}(T,\mathcal{V}) > 0$).*

Proof. Let (X,T) be $(F_n - F_n)$-mixing for each $n \in \mathbb{N}$, and let $\mathcal{U} = \{U_1,U_2,\cdots,U_l\}$ be an admissible open cover. Put $W_i = \text{int}(U_i \setminus \bigcup_{j\neq i} U_j), i = 1,2,\cdots,l$, then $\{W_i\}$ are pairwise disjoint opene sets. First we have:

Claim: Fix $M \in \mathbb{N}$. For any $n \in \mathbb{N}$, there exists $C_n = \{M \le t_1^n < t_2^n < \cdots < t_n^n\} \subseteq F$ such that for any $s \in \{1,2,\cdots,l\}^n$, $\bigcap_{i=1}^n T^{-t_i^n}W_{s(i)} \neq \emptyset$.

Proof of Claim. We use induction on n. Obviously, if $n = 1$ the claim holds. Suppose that when $n = k$ the claim holds, and we want to show the claim holds when $n = k+1$. Thus by this assumption, there exists $C_k = \{M \le t_1^k < t_2^k < \cdots < t_k^k\} \subseteq F$ such that for any $s \in \{1,2,\cdots,l\}^k$, $\bigcap_{i=1}^k T^{-t_i^k}W_{s(i)} \neq \emptyset$.

Take $m_k \in \mathbb{N}$ larger enough such that $C_k \subseteq FS(\{p_i\}_{i=1}^{m_k-1})$. By Theorem 4.6 and Proposition 4.7, we have $[\bigcap_{s\in\{1,2,\cdots,l\}^k} \bigcap_{j=1}^{l} N(\bigcap_{i=1}^{k} T^{-t_i^k}W_{s(i)}, W_j)] \cap (F_{m_k} - F_{m_k})$ is infinite, and then there exist $a_1, a_2 \in F_{m_k}$ such that

$$(a_1 - a_2) \in \bigcap_{s\in\{1,2,\cdots,l\}^k} \bigcap_{j=1}^{l} N(\bigcap_{i=1}^{k} T^{-t_i^k}W_{s(i)}, W_j) \text{ and } a_1 - a_2 > t_k^k.$$

Let $t_i^{k+1} = t_i^k + a_2, i = 1,2,\cdots,k$ and $t_{k+1}^{k+1} = a_1$. Then $C_{k+1} = \{M \le t_1^{k+1} < t_2^{k+1} < \cdots < t_{k+1}^{k+1}\} \subseteq F$ and for any $s \in \{1,2,\cdots,l\}^k, j = 1,2,\cdots,l$, $\bigcap_{i=1}^k T^{-t_i^k}W_{s(i)} \cap$

$T^{-(a_1-a_2)}W_j \neq \emptyset$, i.e. $\bigcap_{i=1}^{k+1} T^{-t_i^{k+1}} W_{r(i)} \neq \emptyset$ for any $r \in \{1, 2, \cdots, l\}^{k+1}$. This shows when $n = k+1$ the claim holds. This ends the proof.

Choose C_{n_i} such that $\max C_{n_i} < \min C_{n_{i+1}}$ and $|C_{n_i}|(= n_i) = (i+1)! - i!$ for $i = 1, 2, \cdots$. Let $A = \bigcup_{i=1}^{\infty} C_{n_i} \subseteq F$ and $A_m = \bigcup_{i=1}^{m} C_{n_i}$. Then

$$\begin{aligned} h_{top}^A(T,\mathcal{U}) &\geq \liminf_{m\to\infty} \frac{\log N(\bigvee_{j\in A_m} T^{-j}\mathcal{U})}{\sum_{i=1}^m n_i} \geq \liminf_{m\to\infty} \frac{\log N(\bigvee_{j\in C_{n_m}} T^{-j}\mathcal{U})}{\sum_{i=1}^m n_i} \\ &= \liminf_{m\to\infty} \frac{\log l^{n_m}}{\sum_{i=1}^m n_i} = \lim_{m\to\infty} \frac{n_m}{\sum_{i=1}^m n_i} \log l = \log l = H(\mathcal{U}). \end{aligned}$$

As $h_{top}^A(T,\mathcal{U}) \leq H(\mathcal{U})$, we have $h_{top}^A(T,\mathcal{U}) = H(\mathcal{U})$.

Now let $\mathcal{V} = \{V_1, \cdots, V_k\}$ be a non-trivial finite open cover. Take $x_j \in \mathrm{int}V_j^c$, $j = 1, 2, \cdots, k$. Assume $\{y_1, y_2, \cdots, y_l\} = \{x_1, x_2, \cdots, x_k\}$ with $y_s \neq y_t$ for $1 \leq s < t \leq l$. Since $\bigcap_{j=1}^k \mathrm{int}V_j^c = \emptyset$, we have $l \geq 2$. Moreover we may take pairwise disjoint closed neighborhood W_i of $y_i, i = 1, 2, \cdots, l$ such that the open cover $\mathcal{U} = \{W_1^c, W_2^c, \cdots, W_l^c\}$ is coarser than $\mathcal{V}$. Let $P = \{W_1, W_2, \cdots, W_l\}$. Note that $(l-1)^{|E|} N(\bigvee_{j\in E} T^{-j}\mathcal{U}) \geq |\{A : A \in \bigcap_{j\in E} T^{-j}P\}|$ where $E \subset \mathbb{Z}_+$ is a finite set. Similar to the case when $\mathcal{U}$ is an admissible cover, we can find an infinite sequence $A \subseteq F$ such that

$$\begin{aligned} h_{top}^A(T,\mathcal{U}) &\geq \liminf_{m\to\infty} \frac{\log N(\bigvee_{j\in A_m} T^{-j}\mathcal{U})}{\sum_{i=1}^m n_i} \geq \liminf_{m\to\infty} \frac{\log N(\bigvee_{j\in C_{n_m}} T^{-j}\mathcal{U})}{\sum_{i=1}^m n_i} \\ &\geq \liminf_{m\to\infty} \frac{\log \frac{1}{(l-1)^{n_m}} |\{A : A \in \bigcap_{j\in E} T^{-j}P\}|}{\sum_{i=1}^m n_i} = \liminf_{m\to\infty} \frac{\log \frac{1}{(l-1)^{n_m}} l^{n_m}}{\sum_{i=1}^m n_i} \\ &= \lim_{m\to\infty} \frac{n_m}{\sum_{i=1}^m n_i} \log \frac{l}{l-1} = \log \frac{l}{l-1} > 0. \end{aligned}$$

Hence $h_{\mathrm{top}}^A(T,\mathcal{V}) \geq h_{\mathrm{top}}^A(T,\mathcal{U}) > 0$. ∎

Remark. *Using the idea of the above proof, we could get easily that if (X,T) is F-mixing for an $F \in \mathcal{IF}$, then for each admissible open cover $\mathcal{U}$ (resp. non-trivial finite open cover $\mathcal{V}$) there exists an infinite sequence $A \subseteq F$ such that $h_{\mathrm{top}}^A(T,\mathcal{U}) = H(\mathcal{U})$ (resp. $h_{\mathrm{top}}^A(T,\mathcal{V}) > 0$).*

Lemma 4.10. *Let (X,T) be a TDS, $F = FS(\{p_i\}_{i=1}^\infty)$ and $F_n = FS(\{p_i\}_{i=n}^\infty)$, $n \in \mathbb{N}$. Then (X,T) has F-s.u.p.e. iff (X,T) is $(F_n - F_n)$-mixing for each $n \in \mathbb{N}$.*

Proof. By Lemma 4.9, it is left to show that if (X,T) has $F-$s.u.p.e., then (X,T) is $(F_n - F_n)$-mixing for each $n \in \mathbb{N}$. By Lemma 4.8. it remains to show (X,T) has F_n-s.u.p.e.. Let $\mathcal{U} = \{U_1, U_2\}$ be an admissible open cover. Since (X,T) has F-s.u.p.e., there exists an infinite sequence $A \subseteq F$ such that $h_{\mathrm{top}}^A(T,\mathcal{U}) > 0$.

Put $B_n = FS(\{p_i\}_{i=1}^n) = \{p_{i_1} + p_{i_2} + \cdots + p_{i_k} : 1 \le i_1 < i_2 < \cdots < i_k \le n\}$. For each $a \in B_n$, let $A_a = A \cap (a + F_n)$. Since $h_{\mathrm{top}}^A(T,\mathcal{U}) \le \sum_{a\in B_n} h_{\mathrm{top}}^{A_a}(T,\mathcal{U})$, there exists $a \in B_n$ such that $h_{\mathrm{top}}^{A_a}(T,\mathcal{U}) > 0$. Put $A' = A_a - a$. Then $A' \subseteq F_n$ and $h_{\mathrm{top}}^{A'}(T,\mathcal{U}) = h_{\mathrm{top}}^{A_a}(T,\mathcal{U}) > 0$. This shows that (X,T) has F_n-s.u.p.e.. □

By Lemma 4.10 and Theorem 2.3 (2), it is easy to see that

Theorem 4.11. *Let (X,T) be a TDS. Then the following statements are equivalent:*
1. *(X,T) is mildly mixing.*
2. *For each IP-set F, (X,T) has F-s.u.p.e..*
3. *For each topologically non-trivial finite open cover $\mathcal{U}$ and IP-set F, there exists an infinite sequence $A \subseteq F$ such that $h_{\mathrm{top}}^A(T,\mathcal{U}) > 0$.*

Theorem 4.12. *Let (X,T) be a TDS. Then the following statements are equivalent:*
1. *(X,T) is weakly mixing.*
2. *(X,T) has $\mathbb{Z}_+$-s.u.p.e..*
3. *For each topologically non-trivial finite open cover $\mathcal{U}$ there exists an infinite sequence $A \subseteq \mathbb{Z}_+$ such that $h_{\mathrm{top}}^A(T,\mathcal{U}) > 0$.*

Proof. The result follows by taking F to be the the IP-set generated by $1, 1, \cdots$ in Lemmas 4.9 and 4.10. Or see [HLSY] for a little short and direct proof. □

We remark that since for any $F \in \mathcal{IF}$, $SE(X,T,F) \subseteq SE(X,T)$, F-s.u.p.e. implies weak mixing by Theorem 4.12.

Li [L] showed that (X,T) is weakly mixing iff for any admissible open cover $\mathcal{U}$ of X there is some $A \subseteq \mathbb{Z}_+$ such that $h_{top}^A(T,\mathcal{U}) = H(\mathcal{U})$. We have

Theorem 4.13. *Let (X,T) be a TDS. Then*
1. *(X,T) is weakly mixing iff for any admissible open cover $\mathcal{U}$ of X there is some $A \subseteq \mathbb{Z}_+$ such that $h_{top}^A(T,\mathcal{U}) = H(\mathcal{U})$.*
2. *(X,T) is mildly mixing iff for any admissible open cover $\mathcal{U}$ of X and any IP set F there is some $A \subseteq F$ such that $h_{top}^A(T,\mathcal{U}) = H(\mathcal{U})$.*
3. *If (X,T) is strongly mixing, then for any admissible open cover $\mathcal{U}$ of X and any $F \in \mathcal{IF}$ there is some infinite $A \subseteq F$ such that $h_{top}^A(T,\mathcal{U}) = H(\mathcal{U})$. The converse is not true.*

Proof. (1) and (2) follow from Lemma 4.8, 4.9 and 4.10. The first part of (3) comes from the Remark of Lemma 4.9. Example 5 of [B1] shows the second part of (3) (see also Example A of section 6 in [HY2]). □

To end the section we remark that the topological counterpart of rigidity is *uniformly rigid*, i.e. if there are $\{n_i\} \to \infty$ with $\lim_{i\to\infty} \sup_{x\in X} d(x, T^{n_i}(x)) = 0$. By Lemma 4.1 and Proposition 3.3 of [HY2] we know that if (X,T) is uniformly rigid then there is an IP-set F such that it is F-null. But the converse is not true even under the minimality assumptions. In fact a minimal system is uniformly rigid iff it is A-equicontinuous for some IP-set A, see [HY2].

§5 Positive entropy implies positive sequence entropy for covers and applications

In the section, we aim to prove that if the topological entropy of a finite open cover is positive, then the topological entropy with respect to the cover and each infinite sequence is positive. The main tools are the so called local variational principle developed in [BGH], [HY1], [R] and [GW], see also [HMY], and [HMRY]. As applications, we show that an entropy pair is a sequence entropy pair with respect to any sequence. Moreover, we obtain that any transitive diagonal flow is mildly mixing and a minimal topological K-system is strongly mixing.

To do this we start with some notations. Let (X,T) be a TDS. In this section, a *cover* of X is a finite family of Borel subsets of X, whose union is X. We denote the set of covers by $\mathcal{C}_X$ and the set of open covers by $\mathcal{C}_X^0$.

In [R], Romagnoli introduced a notion of measure theoretical entropy for covers that extended definition of partition to covers. Let (X,T) be a TDS and $\mu \in \mathcal{M}_T(X)$. For $\mathcal{U} \in \mathcal{C}_X$ of X define $H_\mu(\mathcal{U}) = \inf_{\substack{\alpha \in \mathcal{P}_X \\ \alpha \succeq \mathcal{U}}} H_\mu(\alpha)$. It is not hard to see that many of the properties of $H_\mu(\alpha)$ can be extended to $H_\mu(\mathcal{U})$ from partitions to covers. The following lemma is proved in [R].

Lemma 5.1. *[R, Lemma 3.8] Let (X,T) be a TDS and $\mu \in \mathcal{M}_T(X)$. If $\mathcal{U}, \mathcal{V} \in \mathcal{C}_X$, then*
1) $0 \leq H_\mu(\mathcal{U}) \leq \log N(\mathcal{U})$.
2) If $\mathcal{U} \succeq \mathcal{V}$, then $H_\mu(\mathcal{U}) \geq H_\mu(\mathcal{V})$.
3) $H_\mu(\mathcal{U} \vee \mathcal{V}) \leq H_\mu(\mathcal{U}) + H_\mu(\mathcal{V})$.
4) $H_\mu(T^{-1}\mathcal{U}) = H_\mu(\mathcal{U})$.

One gets easily that $H_\mu(\mathcal{U}_0^{n-1})$ is a sub-additive function of $n \in \mathbb{N}$ from (3) and (4) of Lemma 5.1, and we may define the μ-entropy of $\mathcal{U}$ as

$$h_\mu(T,\mathcal{U}) = \lim_{n\to+\infty} \frac{1}{n} H_\mu(\mathcal{U}_0^{n-1}) = \inf_{n\geq 1} \frac{1}{n} H_\mu(\mathcal{U}_0^{n-1}).$$

Using (1) of Lemma 5.1 we have $h_\mu(T,\mathcal{U}) \leq h_{\mathrm{top}}(T,\mathcal{U})$. We remark that by the same way we may define $h_\mu^A(T,\mathcal{U})$ for any $A \in \mathcal{IF}$ and we also have $h_\mu^A(T,\mathcal{U}) \leq h_{\mathrm{top}}^A(T,\mathcal{U})$ for any $A \in \mathcal{IF}$.

Now, for any an invertible TDS (X,T), $\mu \in \mathcal{M}(X,T)$ and $\alpha \in \mathcal{P}_X$, we define $\Pi_\mu = \{A \in \mathcal{B}(X) : h_\mu(T,\{A,A^c\}) = 0\}$. We call Π_μ the *Pinsker σ-algebra of* (X,μ,T). The following lemma is a classic result.

Lemma 5.2. *Let (X,T) be an invertible TDS, $\mu \in \mathcal{M}_T(X)$ and $\alpha \in \mathcal{P}_X$. If Π_μ is the Pinsker σ-algebra of (X,μ,T), then $\lim_{k\to\infty} h_\mu(T^k,\alpha) = H_\mu(\alpha|\Pi_\mu)$.*

As a corollary of the above result, one has

Lemma 5.3. *Let (X,T) be an invertible TDS, $\mu \in \mathcal{M}_T(X)$ and $\alpha \in \mathcal{P}_X$. If Π_μ is the Pinsker σ-algebra of (X,μ,T), then $\lim_{k\to\infty} H_\mu(\alpha|\bigvee_{i=1}^\infty T^{-ik}\alpha \vee \Pi_\mu) = H_\mu(\alpha|\Pi_\mu)$.*

Proof. It follows from Lemma 5.2 and the fact (see for example Lemma 18.7.(1) in [G]) $h_\mu(T^k,\alpha) = H_\mu(\alpha|\bigvee_{i=1}^\infty T^{-ik}\alpha) = H_\mu(\alpha|\bigvee_{i=1}^\infty T^{-ik}\alpha \vee \Pi_\mu)$. □

The following proposition plays a crucial role in the proof of our main results. First let us introduce several notions. Let (X,T) be an invertible TDS, $\mu \in \mathcal{M}(X,T)$ and Π_μ be the Pinsker σ-algebra of (X,μ,T). For every $M \in \mathbb{N}$ define a measure λ_μ^M on the product space X^M ($X \times ... \times X$, M times) as the unique measure that satisfies for every $A_1,\dots,A_M \in \mathcal{B}(X)$, $\lambda_\mu^M(A_1\times\cdots\times A_M) = \int_X \prod_{m=1}^M \mathbb{E}(1_{A_m}|\Pi_\mu)d\mu$.

Proposition 5.4. *Let (X,T) be an invertible TDS, $\mu \in \mathcal{M}_T(X)$ and $\mathcal{U} \in \mathcal{C}_X^o$. If $h_\mu(T,\mathcal{U}) > 0$, then for each $A \in \mathcal{IF}$ one has $h_\mu^A(T,\mathcal{U}) > 0$.*

Proof. The proof follows the ideas in the proofs of [HY1] or [HMRY, Theorem 5.12]. Recall that the following two claims have been proved in the proof of Theorem 5.12 in [HMRY]. Let $\mathcal{U} = \{U_1,U_2,\dots,U_M\} \in \mathcal{C}_X^o$. First, we have

Claim 1: $\lambda_\mu^M(\prod_{m=1}^M U_m^c) > 0$.

Since $\lambda_\mu^M(\prod_{m=1}^M U_m^c) = \int_X \prod_{m=1}^M \mathbb{E}(1_{U_m^c}|\Pi_\mu)(x)d\mu(x) > 0$, there is $K \in \mathbb{N}$ such that $\mu(E_K) > 0$, where $E_K = \{x \in X \ : \ \min_{1\le m\le M} \mathbb{E}(1_{U_m^c}|\Pi_\mu)(x) \ge \frac{1}{K}\}$.

For any $s = (s(1),s(2),\cdots,s(M)) \in \{0,1\}^M$ set $A_s = \bigcap_{m=1}^M U_m(s(m))$, where $U_m(0) = U_m$ and $U_m(1) = U_m^c$, and put $\alpha = \{A_s : s \in \{0,1\}^M\}$.

Claim 2: $H_\mu(\alpha|\beta\vee\Pi_\mu) \le H_\mu(\alpha|\Pi_\mu) - \frac{\mu(E_K)}{K}\log(\frac{M}{M-1})$ for any finite Borel partition β which is finer than $\mathcal{U}$ as a cover.

Put $\varepsilon = \frac{\mu(E_K)}{K}\log(\frac{M}{M-1})$. By Lemma 5.3 there exists $l > 0$ such that

$$H_\mu(\alpha|\bigvee_{i=1}^\infty T^{-il}\alpha \vee \Pi_\mu) \ge H_\mu(\alpha|\Pi_\mu) - \frac{\varepsilon}{2}. \tag{5.4.1}$$

Now for a given sequence $A = \{a_1 < a_2 < \cdots\} \subset \mathbb{N}$, there exists $r \in \{0,1,\cdots,l-1\}$ such that

$$d^* = \limsup_{n\to\infty} \frac{|\{i \in \mathbb{N} : \ a_i \equiv r \pmod{l} \text{ and } i \le n\}|}{n} > 0.$$

Put $A_r = \{a_i : a_i \equiv r \pmod{l}, i \in \mathbb{N}\}$ and we rewrite $A_r = \{b_1 l + r < b_2 l + r < b_3 l + r \cdots\}$.

Now

$$\begin{aligned}
h^A_\mu(T,\mathcal{U}) &= \limsup_{n\to\infty} \frac{1}{n} H_\mu(\bigvee_{i=1}^{n} T^{-a_i}\mathcal{U}) \\
&\geq \limsup_{n\to\infty} \frac{1}{n} H_\mu(\bigvee_{a_i\in A_r, i\leq n} T^{-a_i}\mathcal{U}) \\
&\geq d^* \liminf_{n\to\infty} \frac{1}{|\{i\leq n : a_i \in A_r\}|} H_\mu(\bigvee_{a_i\in A_r, i\leq n} T^{-a_i}\mathcal{U}) \\
&= d^* \liminf_{m\to\infty} \frac{1}{m} H_\mu(\bigvee_{i=1}^{m} T^{-b_i l + r}\mathcal{U}) \\
&= d^* \liminf_{m\to\infty} \frac{1}{m} H_\mu(\bigvee_{i=1}^{m} T^{-b_i l}\mathcal{U}). \qquad (5.4.2)
\end{aligned}$$

Let $m \geq 1$ and let $\beta_m \in \mathcal{P}_X$ with $\beta_m \succeq \bigvee_{i=1}^{m} T^{-b_i l}\mathcal{U}$. Since for every $i \in \{1, ..., m\}$, $T^{b_i l}\beta_m \succeq \mathcal{U}$ (recall that T is a homeomorphism), from Claim 2 we deduce

$$\begin{aligned}
H_\mu(\beta_m) &\geq H_\mu(\beta_m|\Pi_\mu) \\
&= H_\mu(\beta_m \vee \bigvee_{i=1}^{m} T^{-b_i l}\alpha|\Pi_\mu) - H_\mu(\bigvee_{i=1}^{m} T^{-b_i l}\alpha|\beta_m \vee \Pi_\mu) \\
&\geq H_\mu(\bigvee_{i=1}^{m} T^{-b_i l}\alpha|\Pi_\mu) - \sum_{i=1}^{m} H_\mu(\alpha|T^{b_i l}\beta_m \vee \Pi_\mu) \\
&\geq H_\mu(\bigvee_{i=1}^{m} T^{-b_i l}\alpha|\Pi_\mu) - m(H_\mu(\alpha|\Pi_\mu) - \varepsilon) \text{ (by Claim 2).} \\
&= \sum_{i=1}^{m} H_\mu(\alpha| \bigvee_{j=i+1}^{m} T^{-(b_j - b_i)l}\alpha \vee \Pi_\mu) - m(H_\mu(\alpha|\Pi_\mu) - \varepsilon) \\
&\geq mH_\mu(\alpha|\bigvee_{i=1}^{\infty} T^{-il}\alpha \vee \Pi_\mu) - m(H_\mu(\alpha|\Pi_\mu) - \varepsilon) \\
&\geq m(H_\mu(\alpha|\Pi_\mu) - \frac{\varepsilon}{2}) - m(H_\mu(\alpha|\Pi_\mu) - \varepsilon) \text{ (by (5.4.1))} \\
&= \frac{\varepsilon}{2} m.
\end{aligned}$$

Hence $H_\mu(\bigvee_{i=1}^{m} T^{-b_i l}\mathcal{U}) \geq \frac{\varepsilon}{2} m$. Combining this inequality with (5.4.2), one obtains $h^A_\mu(T,\mathcal{U}) \geq d^* \cdot \frac{\varepsilon}{2} > 0$, which ends the proof. □

Let d be the compatible metric of X, we say (X_T, S) is the *natural extension* of (X,T), if $\widetilde{X} = \{(x_1, x_2, \cdots) : T(x_{i+1}) = x_i, x_i \in X, i \in \mathbb{N}\}$, which is a subspace of the product space $\Pi_{i=1}^{\infty} X$ with the compatible metric d_T defined by

$$d_T((x_1, x_2, \cdots), (y_1, y_2, \cdots)) = \sum_{i=1}^{\infty} \frac{d(x_i, y_i)}{2^i}.$$

Moreover, $\sigma_T : \widetilde{X} \to \widetilde{X}$ is the shift homeomorphism, i.e. $\sigma_T(x_1, x_2, \cdots) = (T(x_1), x_1, x_2, \cdots)$, and $\pi_i : \widetilde{X} \to X$ is the projection to the i-th coordinate. Particularly $\pi_1 : (\widetilde{X}, \sigma_T) \to (X,T)$ is a factor map. Now we are ready to prove

Theorem 5.5. *Let (X,T) be a TDS and $\mathcal{U} \in \mathcal{C}_X^o$. If $h_{\mathrm{top}}(T,\mathcal{U}) > 0$, then for each $A \in \mathcal{IF}$ one has $h_{\mathrm{top}}^A(T,\mathcal{U}) > 0$.*

Proof. First, we assume that T is a homeomorphism. By the local variational principle [HMRY,R] there exists $\mu \in \mathcal{M}_T(X)$ such that $h_\mu(T,\mathcal{U}) = h_{\mathrm{top}}(T,\mathcal{U}) > 0$. According to Proposition 5.4, one has $h_\mu^A(T,\mathcal{U}) > 0$ for each $A \in \mathcal{IF}$. Thus, $h_{\mathrm{top}}^A(T,\mathcal{U}) \geq h_\mu^A(T,\mathcal{U}) > 0$.

For the general case we may pass it to the natural extension. This ends the proof of the theorem. $\square$

Let (X,T) be a TDS. Recall that a pair (x_1, x_2) is called an *entropy pair* if $x_1 \neq x_2 \in X$ and for any admissible open cover $\mathcal{U}$ separating (x_1, x_2) we have $h_{\mathrm{top}}(T,\mathcal{U}) > 0$. Denote by $E(X,T)$ the set of entropy pairs of (X,T). As an application of Theorem 5.5, we have

Corollary 5.6. *Let (X,T) be a TDS. If $h_{\mathrm{top}}(T) > 0$, then for each $F \in \mathcal{IF}$, $h_{\mathrm{top}}^F(T) > 0$. Moreover for any $F \in \mathcal{IF}$, F-null system has zero entropy.*

Proof. Assume that $h_{\mathrm{top}}(T) > 0$. Then there is a finite open cover $\mathcal{U}$ such that $h_{\mathrm{top}}(T,\mathcal{U}) > 0$. By Theorem 5.5, for each $F \in \mathcal{IF}$ we have $h_{\mathrm{top}}^F(T,\mathcal{U}) > 0$, and thus $h_{\mathrm{top}}^F(T) > 0$. $\square$

Corollary 5.7. *Let (X,T) be a TDS. Then $E(X,T) \subseteq SE(X,T,F)$ for any $F \in \mathcal{IF}$.*

In [B2] Blanchard introduced the notion of diagonal flow and proved that any diagonal flow is disjoint from any minimal system with zero entropy. We say (X,T) has *u.p.e.* if $E(X,T) \cup \Delta_X = X^2$, and it is a *diagonal flow* if $E(X,T) \cup \Delta_X \supseteq \{(x, Tx) : x \in X\}$, and it has *c.p.e.* if the smallest closed invariant equivalence relation containing $E(X,T)$ is X^2.

Theorem 5.8. *Any transitive diagonal flow is mildly mixing.*

Proof. Let (X,T) be a transitive diagonal flow. Without loss of generality we assume that (X,T) is non-trivial. By Theorem 2.3 (2), it remains to show that for any opene U, V and IP-set F, $N(U,V) \cap (F - F) \neq \emptyset$.

Fix opene sets U, V and an IP-set F. Let $F = FS(\{p_i\}_{i=1}^{\infty})$. Since (X,T) is transitive, there exist opene subsets $V_0, V_1, \cdots, V_r$ with $V_0 \subseteq V, V_r = U$ and $TV_i \subseteq V_{i+1}, i = 0, 1, \cdots, r-1$.

Take a transitive point $x \in V_0$. Then $Tx \in V_1$ and $x \neq Tx$ as (X,T) is nontrivial. Thus $(x, Tx) \subseteq E(X,T) \subseteq SE(X,T,F)$ by Corollary 5.7. By the remark of Lemma 4.8, $N(V_1, V_0) \cap (F - F) \neq \emptyset$. Hence there exist $a_1 = p_{i_1^1} + p_{i_2^1} + \cdots + p_{i_{k_1}^1}$ and $b_1 = p_{j_1^1} + p_{j_2^1} + \cdots + p_{j_{m_1}^1}$ such that $W_1 = V_1 \cap T^{-(a_1 - b_1)} V_0 \neq \emptyset$, where $1 \leq i_1^1 < \cdots < i_{k_1}^1$ and $1 \leq j_1^1 < \cdots < j_{m_1}^1$. Let $n_1 = \max\{i_{k_1}^1, j_{m_1}^1\} + 1$ and $F_1 = FS(\{p_i\}_{i=n_1}^{\infty})$.

Replacing V_0, V_1 and F by W_1, V_2 and F_1 respectively, repeating the above discussion (as $T(W_1) \subseteq V_2$), we have that there exist $a_2 = p_{i_1^2} + p_{i_2^2} + \cdots + p_{i_{k_2}^2}$ and $b_2 = p_{j_1^2} + p_{j_2^2} + \cdots + p_{j_{m_2}^2}$ such that $W_2 = V_2 \cap T^{-(a_2 - b_2)} W_1 \neq \emptyset$, where $n_1 \leq i_1^2 < \cdots < i_{k_2}^2$ and $n_1 \leq j_1^2 < \cdots < j_{m_2}^2$. Let $n_2 = \max\{i_{k_2}^2, j_{m_2}^2\} + 1$ and $F_2 = FS(\{p_i\}_{i=n_2}^{\infty})$.

By induction for each $2 \leq t \leq r$ there exist $a_t = p_{i_1^t} + p_{i_2^t} + \cdots + p_{i_{k_t}^t}$ and $b_t = p_{j_1^t} + p_{j_2^t} + \cdots + p_{j_{m_t}^t}$ such that $W_t = V_t \cap T^{-(a_t - b_t)} W_{t-1} \neq \emptyset$, where $n_{t-1} \leq i_1^t < \cdots < i_{k_t}^t$ and $n_{t-1} \leq j_1^t < \cdots < j_{m_t}^t$. Let $n_t = \max\{i_{k_t}^t, j_{m_t}^t\} + 1$ and $F_t = FS(\{p_i\}_{i=n_t}^{\infty})$.

By the construction, one has

$$V_r \cap T^{-(a_r - b_r)} W_{r-1} \subseteq V_r \cap T^{-(a_r - b_r)} (T^{-(a_{r-1} - b_{r-1})} W_{r-2}) \cdots \subseteq V_r \cap T^{-(a-b)} V_0,$$

where $a = a_1 + \cdots + a_r, b = b_1 + \cdots + b_r$. This shows that $U \cap T^{-(a-b)} V \supseteq V_r \cap T^{-(a-b)} V_0 \supseteq W_r \neq \emptyset$. By the choose a_i, b_i, one has $a, b \in F$. Hence $a - b \in N(U, V) \cap F - F$. This finishes the proof of Theorem 5.8. $\square$

We remark that it is known that u.p.e. (which is strictly stronger than diagonality) implies mild mixing [HY1]. As a transitive c.p.e. system is not necessarily weak mixing [B1], transitive c.p.e. does not imply mild mixing in general.

In ergodic theory, the notions of Kolmogorov systems and Bernoulli systems play a great role. Thus it is a natural question how to introduce the corresponding notions in TDS. A systematic study of topological Kolmogorov systems was started by Blanchard [B1]. In [HY1] Huang and Ye introduced and studied the notion of *u.p.e. of order n* and the notion of *topological K-systems*, i.e., u.p.e. of order n for each $n \geq 2$. Let (X,T) be TDS. (X,T) is called

- *topological K-system* if each (topologically) non-trivial finite open cover of X has positive entropy;
- (topologically) *full-Bernoulli* if the topological entropy with respect to its each non-trivial finite open cover and each infinite sequence is positive;
- *full-scattering* if if for each $A = \{a_1 < a_2 < \dots\} \in \mathcal{IF}$ and each non-trivial finite open cover $\mathcal{U}$, $N(\bigvee_{i=1}^n T^{-a_i} \mathcal{U}) \longrightarrow \infty$;
- *quasi-Bernoulli* if for each $F \in \mathcal{IF}$, (X,T) is F-s.u.p.e.

It is clear that quasi-Bernoulli implies full scattering, which implies mild mixing [HY2].

Corollary 5.9. *Let (X, T) be TDS. Then (X, T) is a topological K-system if and only if (X, T) is full-Bernoulli. Moreover the product of two full-Bernoulli systems is again full-Bernoulli.*

Proof. By the definition a full-Bernoulli system is a topological K-system. By Theorem 5.5 the converse statement is true too. Finally, since the product of two topological K-systems is a topological K-system by [HY1], we know that the product of two full-Bernoulli systems is again full-Bernoulli. $\square$

It is well known that in ergodic theory a K-system is strongly mixing. Thus an immediate question is that if a topological K-system is strongly mixing. The answer is no if we do not assume minimality (see Example A in [HY2]). Recall that it is proved in [HY2] that a minimal full-scattering system is strongly mixing. So we have

Theorem 5.10. *Let (X, T) be a topological K-system. If (X, T) is minimal, then (X, T) is strongly mixing.*

To sum up we have

$$\begin{array}{ccc}
\textit{topo. } K = \textit{full Bernoulli} & & \\
+ & \searrow & \\
\textit{minimaliy} & & \textit{quasi Bernoulli} \Rightarrow \textit{full scattering} \Rightarrow \textit{mild mixing.} \\
\Downarrow & \nearrow & \\
\textit{strong mixing} & &
\end{array}$$

For completeness we give a direct proof of the following result which is firstly appeared in [S].

Theorem 5.11. *Let $(X, \mathcal{B}, \mu, T)$ be an ergodic invertible MDS and $\xi \in \mathcal{P}_X$. If $h_\mu(T, \xi) > 0$, then $h_\mu^A(T, \xi) \geq h_\mu(T, \xi) > 0$ for any $A \in \mathcal{IF}$.*

Proof. Let $A = \{0 \leq a_1 < a_2 < \cdots\}$. As $h_\mu(T, \xi) = H_\mu(\xi | \bigvee_{i=1}^\infty T^{-i}\xi)$, one has

$$\begin{aligned}
h_\mu^A(T, \xi) &= \limsup_{n\to\infty} \frac{1}{n} H_\mu(\bigvee_{i=1}^{n} T^{-a_i}\xi) = \limsup_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} H_\mu(T^{-a_j}\xi | \bigvee_{i=1}^{j-1} T^{-a_i}\xi) \\
&= \limsup_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} H_\mu(\xi | \bigvee_{i=1}^{j-1} T^{a_j - a_i}\xi) \geq \limsup_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} H_\mu(\xi | \bigvee_{i=1}^{\infty} T^{-i}\xi) \\
&= h_\mu(T, \xi).
\end{aligned}$$

This ends the proof. $\square$

Finally, we have some open problems:

Problem 1. Is a minimal u.p.e. system strongly mixing?

Problem 2. Is it true that for any IP-set F, F-null minimal system is an almost one to one extension of a uniformly rigid system?

Acknowledgement: The final version of the paper was written when the first author visits the Department of Mathematics, George Tech and the third author visits Max-Planck-Institut fur Mathematik, Bonn. The supports and the hospitality of the both institutes are very much appreciated.

References

[Ak] E. Akin, *Recurrence in topological dynamical systems: Furstenberg families and Ellis actions*, Plenum Press, New York, 1997.

[B1] F. Blanchard, *Fully positive topological entropy and topological mixing*, Symbolic dynamics and its applications, AMS Contemporary Mathematics, **135** (1992), 95-105.

[B2] F. Blanchard, *A disjointness theorem involving topological entropy*, Bull. de la Soc. Math. de France, **121** (1993), 465-478.

[BGH] F. Blanchard, E. Glasner and B. Host, *A variation on the variational principle and applications to entropy pairs*, Ergod. Th. and Dynam. Sys., **17** (1997), 29–43.

[BHM] F. Blanchard, B. Host, A. Maass, *Topological Complexity*, Ergod. Th. and Dynam. Sys., **20** (2000), 641-662.

[F1] H. Furstenberg, *Disjointness in ergodic theory, minimal sets, and a problem in Diophantine approximation*, Math. Systems Theory, **1** (1967), 1-49.

[F2] H. Furstenberg, *Recurrence in ergodic theory and combinatorial number theory*, Princeton Univ. Press, 1981.

[F3] H. Furstenberg, *IP-systems in ergodic theory*, Contemp. Math. **26** (1984), 131–148.

[FK] H. Furstenberg and Y. Katznelson, *An ergodic Szemerédi theorem for IP-systems and combinatorial theory*, J. Analyse Math. **45** (1985), 117–168.

[FO] N. Friedman and D. Ornstein, *On mixing and partial mixing*, Ill. J. Math. **16** (1972), 61-68.

[FW] H. Furstenberg and B. Weiss, *The finite multipliers of infinite ergodic transformations.*, The structure of attractors in dynamical systems (Proc. Conf., North Dakota State Univ., Fargo, N.D., 1977), Lecture Notes in Math., 668, Springer, Berlin, (1978), 127–132.

[G] E. Glasner, *Ergodic theory via joinings,*, Mathematical Surveys and Monographs, 101. American Mathematical Society, Providence, RI, 2003.

[GW] E. Glasner and B. Weiss, *On the interplay between measurable and topological dynamics*, preprint (2004).

[Go] T.N.T Goodman, *Topological sequence entropy*, Proc. London Math. Soc., **29** (1974), 331-350.

[Hu1] P. Hulse, *Sequence entropy and subsequence generators*, J. London Math. Soc., (2) **26** (1982), 441-450.

[Hu2] P. Hulse, *Sequence entropy relative to an invariant σ-algebra*, J. London Math. Soc. (2) **33** (1986), 59–72.

[HLSY] W. Huang, S. Li, S. Shao and X. Ye, *Null systems and sequence entropy pairs*, Ergod. Th. and Dynam. Sys., **23** (2003), 1505-1523.

[HMY] W. Huang, A. Maass and X. Ye, *Sequence entropy pairs and complexity pairs for a measure*, to appear in Annales de L'Institut Fourier.

[HMRY] W. Huang, A. Maass, P.P. Romagnoli and X. Ye, *Entropy pairs and a local abramov formula for measure theoretic entropy for a cover*, to appear in Ergodic theory and dynamical systems.

[HY1] W. Huang and X. Ye, *A local variational relation and applications*, submitted 2003..

[HY2] W. Huang and X. Ye, *Topological complexity, return times and weak disjointness*, Ergod. Th. and Dynam. Sys., **24** (2004), 825-846.

[Ku] A. G. Kushnirenko, *On metric invariants of entropy type*, Russian Math. Surveys, **22** (1967), 53-61.

[L] S. Li, *Topological sequence entropy and topologically weak mixing*, Bull. Austral. Math. Soc., **66** (2002), 207–211.

[R] P. Romagnoli, *A local variational principle for the topological entropy*, Ergod. Th. and Dynam. Sys., **23** (2003), 1601-1610.

[S] A. Saleski, *Sequence entropy and mixing*, J. of Math. Anal. and Appli., **60** (1977), 58-66.

[W1] P. Walters, *An introduction to ergodic theory*, Graduate Texts in Mathematics, 79. Springer - Verlag, New York-Berlin, 1982.

[W2] P. Walters, *Some invariant σ- algebras for measure-preserving transformations*, Trans. Amer. Math. Soc., **163** (1972), 357-368.

[Z1] Q. Zhang, *Sequence entropy and mild mixing*, Can. J. Math., **44** (1992), 215-224.

[Z2] Q. Zhang, *Conditional sequence entropy and mild mixing extensions*, Canad. J. Math., **45** (1993), 429–448.

[Zi] R. J. Zimmer, *Extensions of ergodic group actions*, Illinois J. Math., **20** (1976), 373-409.

[†]Department of Mathematics, University of Science and Technology of China, Hefei, Anhui, 230026, P.R. China

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

[†]Department of Mathematics, University of Science and Technology of China, Hefei, Anhui, 230026, P.R. China

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

[†]Department of Mathematics, University of Science and Technology of China, Hefei, Anhui, 230026, P.R. China

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