Canad. J. Math. Vol. **45 (2)** 1993, pp. 429–448.

CONDITIONAL SEQUENCE ENTROPY AND MILD MIXING EXTENSIONS

QING ZHANG

ABSTRACT For a measure preserving system $(X, \mathcal{B}, \mu, T)$ with a factor $(Y, \mathcal{C}, \nu, T)$ and an infinite sequence $\{t_n\}$, one can define conditional sequence entropy We present two theorems which characterize rigid and mildly mixing extensions by conditional sequence entropy Properties of IP-systems are used to prove our main theorems

1. **Introduction.** Given an infinite subset $\Gamma = \{t_n\}$ of $\mathbf{N}$ and a dynamical system $(X, \mathcal{B}, \mu, T)$, one can define sequence entropy along Γ. ([7, 9, 12]) A. G. Kushnirenko [7] and A. Saleski [9] used this notion to characterize the transformations with discrete spectrum and mixing properties respectively. Later P. Hulse [6] gave conditional sequence entropy characterization of compact and weakly mixing extensions of dynamical systems. Hulse's results can be viewed as extensions of Kushnirenko's and Saleski's results. To quote his results we need the following definition.

DEFINITION 1.1. Suppose that $\mathcal{Y} = (Y, \mathcal{C}, \nu, T)$ is a factor of dynamical system $\mathcal{X} = (X, \mathcal{B}, \mu, T)$ and T is invertible. Let $\Gamma = \{t_n\,;\, n = 1, 2, \ldots\} \subset \mathbf{N}$ and ξ be a finite partition of X. The Γ-entropy of T relative to $\mathcal{Y}$ is defined as follows:

$$h_\Gamma(T, \xi|\mathcal{Y}) = \limsup_{n\to+\infty} \frac{1}{n} H\Big(\bigvee_{i=1}^{n} T^{t_i}\xi|\mathcal{Y}\Big)$$

and

$$h_\Gamma(T|\mathcal{Y}) = \sup_{\xi} h_\Gamma(T, \xi|\mathcal{Y})$$

for any finite partition ξ of X. Sometimes Γ-entropy relative to a certain factor is also called *conditional sequence entropy*.

THEOREM 1.1 (HULSE [6]). *Let $\mathcal{Y} = (Y, \mathcal{C}, \nu, T)$ be a factor of a dynamical system $\mathcal{X} = (X, \mathcal{B}, \mu, T)$ and T be invertible. Then:*

(i) $\mathcal{X}$ is a compact extension of $\mathcal{Y}$ if and only if $h_\Gamma(T|\mathcal{Y}) = 0$ for all $\Gamma \subset \mathbf{N}$.

(ii) $\mathcal{X}$ is a weakly mixing extension of $\mathcal{Y}$ if and only if for any $\Gamma_0 \subset \mathbf{N}$ with density one, there exists $\Gamma \subset \Gamma_0$ such that

$$h_\Gamma(T, \xi|\mathcal{Y}) = H(\xi|\mathcal{Y})$$

for all finite partition ξ.

Received by the editors April 17, 1991
AMS subject classification 28D20

Definitions of compact and weakly mixing extensions can be found in [6, 13, 14].

In [4] H. Furstenberg and B. Weiss introduced a new kind of mixing property of dynamical system which they called *mild mixing*. A function $f \in L^2(X, \mathcal{B}, \mu)$ is rigid if there exists $\{t_n\}$ such that $T^{t_n} \longrightarrow f$ in strong sense. A dynamical system $\mathcal{X} = (X, \mathcal{B}, \mu, T)$ is mild mixing if there is no nonconstant rigid function in $L^2(X, \mathcal{B}, \mu)$. Later Furstenberg and Katznelson introduced the relative version of this notion (see [3] or the Definitions 3.3 and 3.4.) Sequence entropy characterizations of the rigidity and mild mixing can be found in [12]. This note is a sequel of [12] and gives conditional sequence entropy characterizations of mild mixing and rigidity relative to a T-invariant sub-σ-algebra.

In §2 we bring some facts about IP-sets, IP-systems and IP-limits. This will be followed by a brief account of some results concerning factors of measure preserving systems. In §3 we use $\mathcal{Y}$-kernels and $\mathcal{Y}$-sequences to decompose a dynamical IP-system into relative rigid part and relative mixing part. In §4 the rigidity and mild mixing relative to a T-invariant sub-σ-algebra will be defined and their properties will be discussed. In the last section, we obtain the following sequence entropy characterizations of rigidity and mild mixing relative to a factor. (For the notions "I-representation" and "φ-monotone" see Definitions 4.1 and 4.2)

THEOREM 1.2. *Let $\mathcal{Y} = (Y, \mathcal{C}, \nu, T)$ be a factor of $\mathcal{X} = (X, \mathcal{B}, \mu, T)$. Then the following statements are equivalent.*

(i) T is $\mathcal{Y}$-rigid;

(ii) There exist an IP-*set I and I-representation φ such that for any φ-monotone subset $\Gamma \subset I$, $h_\Gamma(T|\mathcal{Y}) = 0$;*

(iii) There exists a subset $\Gamma \subset \mathbf{N}$ such that for any sequence $\{\Gamma_i\}$ of pairwise disjoint finite subsets of Γ, $h_{\tilde{\Gamma}}(T|\mathcal{Y}) = 0$. Here $\tilde{\Gamma} = \{t_n = \sum_{a\in\Gamma_n} a\}$.

THEOREM 1.3. *The following statements are equivalent:*

(i) T is mildly $\mathcal{Y}$-mixing;

(ii) For every IP-*set I and I-representation φ, there is an φ-monotone subset $\Gamma \subset I$ such that $h_\Gamma(T, \xi|\mathcal{Y}) = H(\xi|\mathcal{Y})$ for all ξ satisfying $H(\xi|\mathcal{Y}) < \infty$;*

(iii) For any subset $\Gamma \subset \mathbf{N}$, there is a sequence $\{\Gamma_i\}$ of pairwise disjoint finite subsets of Γ such that for any partition ξ of X, $h_{\tilde{\Gamma}}(T, \xi|\mathcal{Y}) = H(\xi|\mathcal{Y})$. Here $\tilde{\Gamma} = \{t_n = \sum_{a\in\Gamma_n} a\}$.

We will use $\mathbf{N}$, $\mathbf{Z}$ and $\mathcal{F}$ to denote respectively the set of all positive integers, the set of all integers and the set of all finite nonempty subsets of $\mathbf{N}$ respectively. For a subset (or a linear subspace) V of a topological space, $\bar{V}$ denote the closure of V.

This note is finished under the guidance of Professor V. Bergelson in The Ohio State University. The author wishes to thank him.

2. **Preliminaries.** We begin with several definitions which are mainly taken from [1, 2, 3].

DEFINITION 2.1. A homomorphism $\psi\colon \mathcal{F} \longrightarrow \mathcal{F}$ is a map such that $\alpha \cap \beta = \emptyset$ implies $\psi(\alpha) \cap \psi(\beta) = \emptyset$ and $\psi(\alpha \cup \beta) = \psi(\alpha) \cup \psi(\beta)$.

The following proposition is an immediate consequence from the definition above.

PROPOSITION 2.1. *Let ψ_1, ψ_2 be two homomorphisms and $\mathcal{F}_1 = \psi_1(\mathcal{F})$, $\mathcal{F}_2 = \psi_2(\mathcal{F})$. If $\mathcal{F}_1 \subset \mathcal{F}_2$, then there is a homomorphism $\psi\colon \mathcal{F} \longrightarrow \mathcal{F}$ such that $\psi_2 = \psi_1 \circ \psi$.*

An $\mathcal{F}$-sequence is a sequence $\{x_\alpha \,;\, \alpha \in \mathcal{F}\}$ indexed by elements $\alpha \in \mathcal{F}$. Given a semigroup X and a sequence $\{x_i\}$ of elements of X, one can define an $\mathcal{F}$-sequence by $x_{i_1, i_2, \ldots, i_k} = x_{i_1} x_{i_2} \cdots x_{i_k}$ where $i_1 < i_2 < \cdots < i_k$. Such an $\mathcal{F}$-sequence will be called an IP-*system*. An important examples of IP-system is an IP-set of **N** which consists of a sequence of real numbers $p_1, p_2, \ldots$ together with all finite sums $p_{i_1} + p_{i_2} + \cdots + p_{i_k}$ with $i_1 < i_2 < \cdots < i_k$.

Given an $\mathcal{F}$-sequence $\{x_\alpha\}$ and a homomorphism $\psi\colon \mathcal{F} \longrightarrow \mathcal{F}$, one can define $\mathcal{F}$-subsequence $\{y_\alpha = x_{\psi(\alpha)}\}$. In particular, if $\{x_\alpha\}$ is an IP-system, then we call $\{y_\alpha\}$ a *sub*-IP-*system*.

DEFINITION 2.2. An IP-*ring* is the range of a homomorphism. Let $\mathcal{F}_1$, $\mathcal{F}_2$ be IP-rings and $\mathcal{F}_1 \subset \mathcal{F}_2$. Then $\mathcal{F}_1$ is called a *sub*-IP-*ring* of $\mathcal{F}_2$.

REMARK. IP-ring is just an IP-system generated by a family of disjoint subsets of **N**. It is a sub-IP-system of $\mathcal{F}$, seen as a semigroup with set-theoretical union as the operation.

DEFINITION 2.3. Let $\{x_\alpha\}$ be an $\mathcal{F}$-sequence in a topological space X and $x \in X$. x is a limit of $\{x_\alpha\}$, $\lim_{\alpha \in \mathcal{F}} x_\alpha = x$ if for every neighborhood V of x there exists an index β so that $\alpha \cap \beta = \emptyset$ implies $x_\alpha \in V$.

REMARK. If X is a Hausdorff topological space, the limit is unique. In this paper, all the spaces which we deal with are Hausdorff.

A proof of the following Bolzano-Weierstrass type theorem can be found in [1, p. 155].

THEOREM 2.2. *If $\{x_\alpha\}$ is an $\mathcal{F}$-sequence with values in a compact metric space, then there exists an $\mathcal{F}$-subsequence $\{y_\alpha\}$ such that* $\lim_\alpha y_\alpha$ *exists.*

Another kind of sets we like to mention here is IP*-sets. A set $A \in \mathbf{N}$ is IP* if A intersects all IP-sets. This notion will be used in proofs later on.

Let $\mathcal{X} = (X, \mathcal{B}, \mu)$, $\mathcal{Y} = (Y, \mathcal{C}, \nu)$ be probability measure spaces. Let θ be a measure preserving map from X to Y. Then there exists a family of conditional probability measures $\{\mu_y \,;\, y \in Y\}$ on $(X, \mathcal{B})$ with following properties:

(i) $\mu_y\big(\theta^{-1}(y)\big) = 1$ for almost all $y \in Y$;

(ii) For every $f \in L^1(X, \mathcal{B}, \mu)$ and for a.e. $y \in Y$, the function

$$y \longrightarrow \int f \, d\mu_y$$

is measurable and $\int f \, d\mu = \int \{\int f \, d\mu_y\} \, d\nu$.

The decomposition $\{\mu_y \,;\, y \in Y\}$ of the measure μ is essentially unique; that is, if $\{\mu_y\}$ and $\{\mu_y'\}$ both have above properties, then $\mu_y = \mu_y'$ for a.e. $y \in Y$. Conditional expectations and conditional measures are related by

$$\mathbf{E}(f|\mathcal{Y})(x) = \int f(t) \, d\mu_{\theta(x)}(t)$$

for all $f \in L^1(X, \mathcal{B}, \mu)$. For $(X \times X, \mathcal{B} \times \mathcal{B})$, we can define a new measure: for any set $B \in \mathcal{B} \times \mathcal{B}$,

$$\mu^{\mathcal{Y}}(B) = \int \mu_y \times \mu_y(B)\, d\nu(y).$$

$\mu^{\mathcal{Y}}$ is characterized by its effect on functions of the form $f \otimes g$, where

$$(f \otimes g)(x_1, x_2) = f(x_1)g(x_2).$$

Since the support of $\mu_y \times \mu_y$ is $\theta^{-1}(y) \times \theta^{-1}(y)$, the support of $\mu^{\mathcal{Y}}$ is

$$\bigcup_{y \in Y} \theta^{-1}(y) \times \theta^{-1}(y) = \{(x_1, x_2) \in X \times X\, ;\, \theta(x_1) = \theta(x_2)\}.$$

Let θ_1, θ_2 be two maps defined by: $\theta_i(x_1, x_2) = \theta(x_i)$ for $i = 1, 2$. It is easy to check that θ_1, θ_2 are measure preserving, *i.e.* $\mu^{\mathcal{Y}}(\theta_i^{-1}(C) = \nu(C)$ for all $C \in \mathcal{C}$. Moreover θ_i; $i = 1, 2$ agree on the set $\bigcup_{y \in Y} \theta^{-1}(y) \times \theta^{-1}(y)$ which has measure 1 with respect to $\mu^{\mathcal{Y}}$.

From now on, $\Sigma = \{T_\alpha\}$ will denote an IP-system of commuting, invertible, measure preserving transformations on $(X, \mathcal{B}, \mu)$ and we will call $\mathcal{X} = (X, \mathcal{B}, \mu, \Sigma)$ a *dynamical* IP-*system*. $\mathcal{Y} = (Y, \mathcal{C}, \mu, \Sigma)$ is a factor of $\mathcal{X}$ if there is a measure preserving map $\theta: X \longrightarrow Y$ such that $\theta T_\alpha = T_\alpha \theta$ for all $T_\alpha \in \Sigma$. From the uniqueness of the decomposition, we get $T_\alpha \mu_y = \mu_{T_\alpha y}$ for a.e. $y \in Y$. This is equivalent to $\mathbf{E}(T_\alpha f | \mathcal{Y}) = T_\alpha \mathbf{E}(f | \mathcal{Y})$. It is also easy to see that

$$\Sigma^2 = \{S_\alpha = T_\alpha \times T_\alpha\}$$

is an IP-system of commuting measure preserving transformations on $(X \times X, \mathcal{B} \times \mathcal{B}, \mu^{\mathcal{Y}})$ and $\mathcal{Y}$ is a factor of $(X \times X, \mathcal{B} \times \mathcal{B}, \mu^{\mathcal{Y}}, \Sigma^2)$. Therefore abusing the terminology, for any $F \in L^2(X \times X, \mathcal{B} \times \mathcal{B}, \mu^{\mathcal{Y}})$, we also use $\mathbf{E}(F|\mathcal{Y})$ to denote the conditional expectation with respect to $\mathcal{Y}$.

DEFINITION 2.4. Let $\mathcal{Y}$ be a factor of an IP-system $\mathcal{X}$. We shall say that $\Sigma = \{T_\alpha\}$ satisfies Condition A if

$$\lim_\alpha \langle S_\alpha H, G \rangle$$

exists for all $H, G \in L^2(X \times X, \mathcal{B} \times \mathcal{B}, \mu^{\mathcal{Y}})$.

REMARKS.

(i) Condition A implies that

$$\lim_\alpha \langle S_{-\alpha} H, G \rangle$$

exists for all $H, G \in L^2(X \times X, \mathcal{B} \times \mathcal{B}, \mu^{\mathcal{Y}})$. Here $S_{-\alpha}$ denotes S_α^{-1}.

(ii) $\{T_\alpha\}$ satisfies Condition A if and only if $\lim_\alpha \int |\mathbf{E}(T_\alpha f g | \mathcal{Y})|^2\, d\mu$ exists for all f, $g \in L^\infty(X, \mathcal{B}, \mu)$.

THEOREM 2.3. *Let* $(X, \mathcal{B}, \mu, \Sigma)$ *be a dynamical* IP-*system. Then there is a sub*-IP-*system* $\Sigma' \subset \Sigma$ *and an orthogonal projection* $\mathbf{P}$ *such that*

$$\lim_{T_\alpha \in \Sigma'} \langle T_\alpha f, g \rangle = \langle \mathbf{P} f, g \rangle$$

for all $f, g \in L^2(X, \mathcal{B}, \mu)$. *In particular, if* $\mathbf{P}f = f$, *then:*

$$\lim_{T_\alpha \in \Sigma'} \|T_\alpha f - f\|^2 = 0.$$

A proof of this theorem can be found in [3, p. 124].

COROLLARY 2.4. *There exists a sub-*IP*-system* $\Sigma' \subset \Sigma$ *satisfying Condition A.*

DEFINITION 2.5. Let $\mathcal{Y}$ be a factor of $\mathcal{X}$. We say that Σ is *$\mathcal{Y}$-mixing* if for any F, $G \in L^2(X \times X, \mathcal{B} \times \mathcal{B}, \mu^{\mathcal{Y}})$,

$$\lim_{\alpha \in \mathcal{F}} \int \left(S_\alpha FG - S_\alpha \mathbf{E}(F|\mathcal{Y})\mathbf{E}(G|\mathcal{Y}) \right) d\mu^{\mathcal{Y}} = 0.$$

THEOREM 2.5. *The following conditions are equivalent:*
(i) $\{T_\alpha\}$ *is* $\mathcal{Y}$*-mixing on* X;
(ii) For any $f, g \in L^\infty(X, \mathcal{B}, \mu)$,

$$\lim_{\alpha \in \mathcal{F}} \|\mathbf{E}(T_\alpha f g|\mathcal{Y}) - T_\alpha \mathbf{E}(F|\mathcal{Y})\mathbf{E}(G|\mathcal{Y})\|_2 = 0\ ;$$

(iii) For any f, $g \in L^2(X, \mathcal{B}, \mu)$,

$$\lim_{\alpha \in \mathcal{F}} \|\mathbf{E}(T_\alpha f g|\mathcal{Y}) - T_\alpha \mathbf{E}(f|\mathcal{Y})\mathbf{E}(g|\mathcal{Y})\|_1 = 0.$$

PROOF. (i) $\Rightarrow$ (ii): We first assume that $\mathbf{E}(f|\mathcal{Y}) = 0$. Let $F = f \otimes f$, $G = g \otimes g$. By Definition 2.5, one gets $\lim_{\alpha \in \mathcal{F}} \int S_\alpha (f \otimes \bar{f}) g \otimes \bar{g}\, d\mu^{\mathcal{Y}} = 0$. This gives:

$$\lim_{\alpha \in \mathcal{F}} \|\mathbf{E}(T_\alpha f g|\mathcal{Y})\|^2 = \lim_{\alpha \in \mathcal{F}} \int \left(\int T_\alpha f g\, d\mu_y \overline{\int T_\alpha f g\, d\mu_y} \right) d\nu = 0. \tag{1}$$

For any $f \in L^\infty(X, \mathcal{B}, \mu)$, define $f_1 = f - \mathbf{E}(f|\mathcal{Y})$. Then $\mathbf{E}(f_1|\mathcal{Y}) = 0$. By (1), $\|\mathbf{E}(T_\alpha f_1 g|\mathcal{Y})\|_2 \longrightarrow 0$ which implies (ii).

(ii) $\Rightarrow$ (i): Let f, f', g and $g' \in L^\infty(X, \mathcal{B}, \mu)$ and

$$M = \max\{\|f\|, \|f'\|, \|g\|, \|g'\|\}.$$

Then:

$$\begin{aligned}\left| \int \mathbf{E}(T_\alpha f g|\mathcal{Y})\mathbf{E}(T_\alpha f' g'|\mathcal{Y})\, d\mu - \int T_\alpha \mathbf{E}(f|\mathcal{Y}) T_\alpha \mathbf{E}(f'|\mathcal{Y})\mathbf{E}(g|\mathcal{Y})\mathbf{E}(g'|\mathcal{Y})\, d\mu \right| \\ \le M\|\mathbf{E}(T_\alpha f g|\mathcal{Y}) - T_\alpha \mathbf{E}(f|\mathcal{Y})\mathbf{E}(g|\mathcal{Y})\|_2 \\ + M\|\mathbf{E}(T_\alpha f' g'|\mathcal{Y}) - T_\alpha \mathbf{E}(f'|\mathcal{Y})\mathbf{E}(g'|\mathcal{Y})\|_2.\end{aligned}$$

Hence:

$$\lim_{\alpha \in \mathcal{F}} \left(\int S_\alpha (f \otimes f') g \otimes g'\, d\mu^{\mathcal{Y}} - \int S_\alpha \mathbf{E}(f \otimes f'|\mathcal{Y})\mathbf{E}(g \otimes g'|\mathcal{Y})\, d\mu^{\mathcal{Y}} \right) = 0.$$

Since all functions with the form $f \otimes f'$ span $L^2(X \times X, \mathcal{B} \times \mathcal{B}, \mu^{\mathcal{Y}})$, (i) follows.

(iii) $\Rightarrow$ (ii): Suppose $|f| \leq M$, $|g| \leq M$. Let

$$A_\varepsilon(\beta) = \{x\,;\, |\mathbf{E}(T_\alpha fg|\mathcal{Y})(x) - T_\alpha \mathbf{E}(f|\mathcal{Y})(x)\mathbf{E}(g|\mathcal{Y})(x)| \geq \varepsilon\}.$$

From (iii), we have that for every $\varepsilon > 0$ there is a β_0 such that: $\mu\big(A_\varepsilon(\beta)\big) < \varepsilon$ whenever $\beta \cap \beta_0 = \emptyset$. Then:

$$\begin{aligned}\int |\mathbf{E}(T_\alpha fg|\mathcal{Y}) - T_\alpha \mathbf{E}(f|\mathcal{Y})\mathbf{E}(g|\mathcal{Y})|^2\, d\mu &\leq \int_{A\ (\beta)} |\mathbf{E}(T_\alpha fg|\mathcal{Y}) - T_\alpha \mathbf{E}(f|\mathcal{Y})\mathbf{E}(g|\mathcal{Y})|^2\, d\mu \\ &\quad + \int_{A\ (\beta^c)} |\mathbf{E}(T_\alpha fg|\mathcal{Y})|^2\, d\mu \\ &\leq \varepsilon + 2M^2\varepsilon\end{aligned}$$

which implies (ii).

(ii) $\Rightarrow$ (iii): Since bounded functions are dense in $L^2(X, \mathcal{B}, \mu)$, it is enough to consider $f, g \in L^\infty(X, \mathcal{B}, \mu)$. Then (ii) and Holder inequality implies (iii) immediately. ∎

DEFINITION 2.6. Let $\mathcal{Y} = (Y, \mathcal{C}, \nu, \Sigma)$ be a factor of $\mathcal{X} = (X, \mathcal{B}, \mu, \Sigma)$. $f \in L^2(X, \mathcal{B}, \mu)$ is almost periodic with respect to $\mathcal{Y}$ along the IP-ring $\mathcal{F}_1$, written as $f \in \mathbf{AP}(\mathcal{Y}, \Sigma, \mathcal{F}_1)$, if for every $\varepsilon > 0$, there is $\alpha_0 \in \mathcal{F}_1$, a set of functions $g_1, g_2, \ldots, g_k \in L^2(X, \mathcal{B}, \mu)$ and a measurable subset $E \subset Y$ with $\nu(E) < \varepsilon$ such that

$$\inf_{1 \leq j \leq k} \|T_\alpha f - g_j\|_y \leq \varepsilon$$

for all $y \notin E$ and all $\alpha \in \mathcal{F}_1$ with $\alpha \cap \alpha_0 = \emptyset$.

REMARK. Actually one can choose $g_1, g_2, \ldots, g_k \in L^\infty(X, \mathcal{B}, \mu)$. We leave the details to the reader.

PROPOSITION 2.6. *Let $\mathcal{Y} = (Y, \mathcal{C}, \nu, \Sigma)$ be a factor of $\mathcal{X} = (X, \mathcal{B}, \mu, \Sigma)$. Then $\mathbf{AP}(\mathcal{Y}, \Sigma, \mathcal{F}_1)$ has following properties:*

(i) $\mathbf{AP}(\mathcal{Y}, \Sigma, \mathcal{F}_1)$ is a linear space;

(ii) If $\mathcal{F}_1$ and $\mathcal{F}_2$ are two IP*-rings and $\mathcal{F}_1 \subset \mathcal{F}_2$, then:*

$$\mathbf{AP}(\mathcal{Y}, \Sigma, \mathcal{F}_2) \subset \mathbf{AP}(\mathcal{Y}, \Sigma, \mathcal{F}_1)\ ;$$

(iii) $\overline{\mathbf{AP}(\mathcal{Y}, \Sigma, \mathcal{F}_1)} \supset L^\infty\big(X, \theta^{-1}(\mathcal{C}), \mu\big)$;

(iv) There exists a sub-σ-algebra $\mathcal{B}_0 \subset \mathcal{B}$ such that

$$\overline{\mathbf{AP}(\mathcal{Y}, \Sigma, \mathcal{F}_1)} = L^2(X, \mathcal{B}_0, \mu).$$

PROOF. (i) Let $f_1, f_2 \in \mathbf{AP}(\mathcal{Y}, \Sigma, \mathcal{F}_1)$ and $\alpha, \beta \in \mathbf{R}$. Then for every $\epsilon > 0$ and $i = 1, 2$, there are α_0^i, E_i and $g_1^i, \ldots, g_{k(i)}^i$ with $\nu(E_i) < \varepsilon/2$ such that

$$\inf_{1 \leq j \leq k(i)} \|T_\alpha f_i - g_j^i\|_y < \varepsilon/(|a| + |b|),$$

https://doi.org/10.4153/CJM-1993-022-2 Published online by Cambridge University Press

where $\alpha \cap \alpha_0^i = \emptyset$ and $y \notin E_i$. Let $\alpha_0 = \alpha_0^1 \cup \alpha_0^2$ and $E = E_1 \cup E_2$. For any $\alpha \cup \alpha_0 = \emptyset$ and $y \notin E$, we have

$$\inf_{1 \le j \le k(1)\ 1 \le m \le k(2)} \|T_\alpha(af_1 + bf_2) - (ag_j^i + bg_m^i)\|_y \le \varepsilon.$$

This means $af_1 + bf_2 \in \mathbf{AP}(\mathcal{Y}, \Sigma, \mathcal{F}_1)$.

(ii) Restricting $\mathcal{F}_1$ to a sub-IP-ring $\mathcal{F}_2$ will enlarge the set of functions satisfying the conditions in Definition 2.6. The details will be omitted in here.

(iii) Let $g \in L^\infty\big(X, \theta^{-1}(\mathcal{C}), \mu\big)$. There is a constant M such that: $|g| \le M$ a.e. Now for $\varepsilon > 0$ there is a partition $-M = a_0 < a_1, \ldots, < a_n = M$ such that $\max_{0 \le i \le n-1} |a_{i+1} - a_i| < \varepsilon$. Since g is measurable with respect to $\big(X, \theta^{-1}(\mathcal{C}), \mu\big)$, it is a constant almost every where with respect to μ_y for a.e. (ν) $y \in Y$. Then $\|T_\alpha f - a_j\|y = |T_\alpha f(x) - a_j|$ for any $x \in \theta^{-1}(y)$. Therefore

$$\inf_{0 \le j \le n} \|T_\alpha f - a_j\|_y < \varepsilon$$

which means $f \in \mathbf{AP}(\mathcal{Y}, \Sigma, \mathcal{F})$.

(iv) A proof can be found in [3, Lemma 7.3]. B_0 is the smallest Σ-invariant σ-algebra with respect to which the functions in $\mathbf{AP}(\mathcal{Y}, \Sigma, \mathcal{F}_1)$ are measurable. ∎

DEFINITION 2.7. A function $H(x, x')$ on $X \times X$ is a $\mathcal{Y}$-kernel if it is measurable with respect to $\mathcal{B} \times \mathcal{B}$ and:

(i) $\int H(x, x')\, d\mu_{\theta(x)}(x') = 0$ for almost all $x \in X$;

(ii) the function $h(y) = \int |H(x, x')|^2\, d\mu_y \times \mu_y$ belongs to $L^\infty(Y, \mathcal{C}, \nu)$.

For a $\mathcal{Y}$-kernel H, we define $H * f(x) = \int H(x, x') f(x')\, d\mu_y(x')$. Here $y = \theta(x)$ and $f \in L^2(X, \mathcal{B}, \mu)$. We leave it to the reader to show that H is a bounded operator on $L^2(X, \mathcal{B}, \mu)$.

LEMMA 2.7. *Let H be a $\mathcal{Y}$-kernel satisfying $\lim_{\alpha \in \mathcal{F}} S_\alpha H = H$. Then there exists a sub-IP-ring $\mathcal{F}_1$ of $\mathcal{F}$ such that for every $\varepsilon > 0$, there is $\beta_0 \in \mathcal{F}_1$, a subset $E \subset Y$ with $\nu(E) < \varepsilon$ and a finite set of functions $g_1, \ldots, g_k \in L^2(X, \mathcal{B}, \mu)$ such that for any $\beta \in \mathcal{F}_1$ with $\beta \cap \beta_0 = \emptyset$, any $y \notin E$ and any $f \in L^2(X, \mathcal{B}, \mu)$ with $\|f\|_y \le 1$ we get*

$$\inf_{1 \le j \le k} \|(S_\beta H) * f - g_j\|_y \le \varepsilon.$$

Moreover $H * f \in \overline{\mathbf{AP}(\mathcal{Y}, \Sigma, \mathcal{F}_1)}$ for every $f \in L^2(X, \mathcal{B}, \mu)$.

REMARK. $\lim_{\alpha \in \mathcal{F}} S_\alpha H = H$ in weak topology if and only if $\lim_{\alpha \in \mathcal{F}} S_\alpha H = H$ in strong topology. This is an immediate consequence of the Theorem 1.3.

PROOF. A proof of the first part of this lemma can be found in [3, Lemma 7.2]. We only need to show that for every $f \in L^2(X, \mathcal{B}, \mu)$, $H * f \in \overline{\mathbf{AP}(\mathcal{Y}, \Sigma, \mathcal{F}_1)}$. Since H is a bounded operator, it is enough to prove this argument for $f \in L^\infty(X, \mathcal{B}, \mu)$. Without loss of generality, suppose $|f| \le 1$. It follows from the first part of the lemma that for every $\varepsilon > 0$, there is a $\beta_0 \in \mathcal{F}$, a subset $E \subset Y$ with $\nu(E) < \varepsilon$ and a finite set of functions $g_1, \ldots, g_k \in L^2(X, \mathcal{B}, \mu)$ such that for arbitrary β with $\beta \cap \beta_0 = \emptyset$, $y \notin E$ and $f' \in L^\infty(X, \mathcal{B}, \mu)$ with $|f'| \leq 1$,

$$\inf_{1 \leq j \leq k} \|(S_\beta H) * f' - g_j\|_y \leq \varepsilon.$$

Taking $f' = T_\beta f$, we have

$$(S_\beta H) * f' = T_\beta(H * f)$$

which implies $H * f \in \mathbf{AP}(\mathcal{Y}, \Sigma, \mathcal{F}_1)$. ∎

3. **Hilbert space decomposition.** In this section we always assume that $\mathcal{Y} = (Y, C, \mu, \Sigma)$ is a factor of $X = (X, \mathcal{B}, \mu, \Sigma)$ and $\Sigma^2 = \{S_\alpha = T_\alpha \times T_\alpha\}$ is the IP-system of measure preserving transformations on $(X \times X, \mathcal{B} \times \mathcal{B}, \mu^{\mathcal{Y}})$.

LEMMA 3.1. *For any $g \in L^\infty(X, \mathcal{B}, \mu)$ satisfying that $\mathbf{E}(g|\mathcal{Y}) = 0$, let H denote the weak IP-limit of $S_\alpha(g \otimes \bar{g})$, i.e. for any $F \in L^2(X \times X, \mathcal{B} \times \mathcal{B}, \mu^{\mathcal{Y}})$,*

$$\langle H, F \rangle = \lim_{\alpha \in \mathcal{F}} \langle S_\alpha(g \otimes \bar{g}), F \rangle.$$

Then H is a $\mathcal{Y}$-kernel satisfying $H = \lim_{\alpha \in \mathcal{F}} S_\alpha H$.

PROOF. For any $f(x) \in L^\infty(X, \mathcal{B}, \mu)$,

$$\int S_\alpha(g \otimes \bar{g}) f \otimes 1 \, d\mu^{\mathcal{Y}} = \int f T_\alpha g \overline{\mathbf{E}(T_\alpha g|\mathcal{Y})} \, d\mu = 0$$

which implies $\int Hf \otimes 1 \, d\mu^{\mathcal{Y}} = 0$. Hence H satisfies the first condition in Definition 2.7. Since $g \otimes \bar{g} \in L^\infty(X \times X, \mathcal{B} \times \mathcal{B}, \mu^{\mathcal{Y}})$ which, as a dual space of $L^1(X \times X, \mathcal{B} \times \mathcal{B}, \mu^{\mathcal{Y}})$, is closed with respect to *-weak topology, (see [11]) $H \in L^\infty(X \times X, \mathcal{B} \times \mathcal{B}, \mu^{\mathcal{Y}})$. Hence H is a $\mathcal{Y}$-kernel. The identity $H = \lim_{\alpha \in \mathcal{F}} S_\alpha H$ comes from Theorem 2.3 directly. ∎

The following lemma is an immediate corollary of Theorem 2.3.

LEMMA 3.2. *If $\{T_\alpha\}$ satisfies Condition A, then $\lim_{\alpha \in \mathcal{F}} T_\alpha^{-1} f = f_0$ in strong (weak) topology iff $\lim_{\alpha \in \mathcal{F}} T_\alpha f = f_0$ in strong (weak) topology.*

Let $\{g_n\} \subset L^\infty(X, \mathcal{B}, \mu) \cap L^2(X, C, \mu)^\perp$ be a dense set in $L^2(X, C, \mu)^\perp$ and $H_n = \lim_{\alpha \in \mathcal{F}} S_\alpha(g_n \otimes \bar{g}_n)$. By Lemma 3.1, H_n is a $\mathcal{Y}$-kernel satisfying $H_n = \lim_{\alpha \in \mathcal{F}} S_\alpha H_n$. By Lemma 2.7 one can inductively construct a sequence $\mathcal{F}_1 \supset \mathcal{F}_2 \supset \cdots$ of IP-rings such that $H_n * f \in \overline{\mathbf{AP}(\mathcal{Y}, \Sigma, \mathcal{F}_n)}$ for any $f \in L^2(X, \mathcal{B}, \mu)$ and any $n \in \mathbf{N}$. We shall call any sequence $\{\mathcal{F}_n\}$ of the kind described above an *$\mathcal{Y}$-sequence* corresponding to $\{g_n\}$, or just an $\mathcal{Y}$-sequence. For any $\mathcal{Y}$-sequence $\{\mathcal{F}_n\}$, we define

$$K(\mathcal{Y}, \Sigma) = \overline{\bigcup_{n=1}^{\infty} \mathbf{AP}(\mathcal{Y}, \Sigma, \mathcal{F})}.$$

Later, in Theorem 3.6, we will prove that $K(\mathcal{Y}, \Sigma)$ does not depend on the choice of $\mathcal{Y}$-sequence. A proof of the following lemma is omitted.

LEMMA 3.3. *There exists a sub-σ-algebra $\mathcal{D}$ of $\mathcal{B}$ such that*

$$K(\mathcal{Y}, \Sigma) = L^2(X, \mathcal{D}, \mu).$$

Moreover, $(X, \mathcal{D}, \mu, \Sigma)$ is a factor of X.

PROPOSITION 3.4. *Suppose that $\mathcal{Y}$ is a factor of* IP-*system X and Σ satisfies Condition A. Let $f \in L^2(X, \mathcal{B}, \mu)$ and $\{\mathcal{F}_n\}$ be any $\mathcal{Y}$-sequence corresponding to a dense set $\{g_n\}$ in $L^2\big(X, \theta(C), \mu\big)^{\perp}$. If $f \perp$ $\mathbf{AP}(\mathcal{Y}, \Sigma, \mathcal{F}_n)$ for all n, then*

$$\lim_{\alpha \in \mathcal{F}} \int |\mathbf{E}(gT_\alpha f | \mathcal{Y})| \, d\mu = 0$$

for any $g \in L^2(X, \mathcal{B}, \mu)$.

PROOF. Let $H_n = \lim_{\alpha\in\mathcal{F}} S_\alpha(g_n \otimes \bar{g}_n)$. Since $H_n * \bar{f} \in \overline{\mathbf{AP}(\mathcal{Y}, \Sigma, \mathcal{F}_n)}$, by Lemma 3.2, we have:

$$\begin{aligned}
\limsup_{\alpha\in\mathcal{F}} \Big(\int |\mathbf{E}(g_n T_\alpha f | \mathcal{Y})| \, d\mu \Big)^2 &\leq \lim_{\alpha\in\mathcal{F}} \int |\mathbf{E}(g_n T_\alpha f | \mathcal{Y})|^2 \, d\mu \\
&= \lim_{\alpha\in\mathcal{F}} \int S_\alpha(f \otimes \bar{f}) g_n \otimes \bar{g}_n \, d\mu^{\mathcal{Y}} \\
&= \int f \otimes \bar{f} H_n \, d\nu^{\mathcal{Y}} = \int f H_n * \bar{f} \, d\mu = 0
\end{aligned}$$

Since $\{g_n\}$ is dense in $L^2\big(X, \theta(C), \mu\big)^{\perp}$ and $\mathbf{E}(f | \mathcal{Y}) = 0$ the proposition follows. ∎

PROPOSITION 3.5. *Suppose that $f \in L^2(X, \mathcal{B}, \mu)$ and*

$$\lim_{\alpha \in \mathcal{F}} \int |\mathbf{E}(gT_\alpha f | \mathcal{Y})| \, d\mu = 0$$

for all $g \in L^2(X, \mathcal{B}, \mu)$. Then for any $\mathcal{Y}$-sequence $\{\mathcal{F}_n\}$, $f \perp$ $\mathbf{AP}(\mathcal{Y}, \Sigma, \mathcal{F}_n)$.

PROOF. Let $\{\mathcal{F}_n\}$ be a $\mathcal{Y}$-sequence and $K(\mathcal{Y}, \Sigma) = \overline{\bigcup_{n=1}^{\infty} \mathbf{AP}(\mathcal{Y}, \Sigma, \mathcal{F}_n)}$. For any $f \in L^2(X, \mathcal{B}, \mu)$, we have the unique decomposition $f = f_1 + f_2$ such that $f_1 \perp K(\mathcal{Y}, \Sigma)$ and $f_2 \in K(\mathcal{Y}, \Sigma)$. By Proposition 3.4, it will be enough to show that: if

$$\lim_{\alpha \in \mathcal{F}} \int |\mathbf{E}(gT_\alpha f | \mathcal{Y})| \, d\mu = 0$$

and $f \in K(\mathcal{Y}, \Sigma)$, then $f = 0$.

For any $\varepsilon > 0$, there is a positive $\delta < \varepsilon$ such that for any set F with $\nu(F) < \delta$, $\int_F |f|^2 \, d\mu < \varepsilon$. Since $f \in K(\mathcal{Y}, \Sigma)$ and $\mathbf{AP}(\mathcal{Y}, \Sigma, \mathcal{F}_{n+1}) \subset \mathbf{AP}(\mathcal{Y}, \Sigma, \mathcal{F}_n)$, there exists $f_0 \in \mathbf{AP}(\mathcal{Y}, \Sigma, \mathcal{F}_k)$ such that $\|f - f_0\|_{L^2(d\mu)} < \delta$. For ε and δ same as above, there exists $\beta_0 \in \mathcal{F}_k$, $\{g_j \in L^2(X, \mathcal{B}, \mu)\,;\, j = 1, \ldots, q\}$ and a set $E \subset \mathcal{Y}$ with $\nu(E) < \delta/2$ such that $\inf_{1\leq j\leq q} \|T_\beta f_0 - g_j\|_y < \varepsilon$ for $y \notin E$ and $\beta \in \mathcal{F}_k$ with $\beta \cap \beta_0 = \emptyset$. Now we claim: for $\beta \cap \beta_0 = \emptyset$

$$\|f\|^2 - 3\sqrt{\varepsilon}\|f\| - \varepsilon \leq \sum_{i=1}^{q} \int |\mathbf{E}(\bar{g}_i T_\beta f | \mathcal{Y})| \, d\mu. \tag{2}$$

Actually for a fixed $\beta \in \mathcal{F}_k$ satisfying $\beta \cap \beta_0 = \emptyset$, let $j(y)$ be the smallest number such that $||T_\beta f_0 - g_j||_y < \varepsilon$. Since

$$\{y\,;\, j(y) = 1\} = \{y\,;\, ||T_\beta f_0 - g_j||_y < \varepsilon\} \setminus \bigcup_{m=1}^{i-1} \{y\,;\, ||T_\beta f_0 - g_m||_y < \varepsilon\} \in \mathcal{C},$$

$j(y)$ is a measurable function on E^c. Notice that the inequality

$$||T_\beta f_0 - T_\beta f|| = ||f_0 - f|| < \delta$$

implies: there exists F_β with $\nu(F_\beta) < \delta/2$ such that $||T_\beta f_0 - T_\beta f||_y^2 < 2\delta$ for $y \notin F_\beta$. Now let $E_\beta = F_\beta \cup E$, then $\nu(E_\beta) < \delta$. For $y \notin E_\beta$, we have:

$$||T_\beta f - g_{j(y)}||_y \leq ||T_\beta f - T_\beta f_0||_y + ||T_\beta f_0 - g_{j(y)}||_y < \sqrt{2\delta} + 3\sqrt{\epsilon}.$$

This implies $\langle T_\beta f, T_\beta f - g_{j(y)}\rangle_y < 3\sqrt{\varepsilon}||T_\beta f||_y$ which means

$$\int T_\beta f \bar{g}_{j(y)}\, d\mu_y \geq ||T_\beta||_y^2 - 3\sqrt{\varepsilon}||T_\beta f||_y.$$

Define:

$$\xi_i(y) = \begin{cases} 1 & \text{if } y \notin E_\beta \text{ and } j(y) = i \\ 0 & \text{otherwise.} \end{cases}$$

and $h(y) = \sum_{i=1}^q \xi_i g_i$, then for $y \notin E_\beta$, $\int T_\beta f \bar{h}\, d\mu_y \geq ||T_\beta f||_y^2 - 3\sqrt{\varepsilon}||T_\beta f||_y$. Hence:

$$\begin{aligned}\int T_\beta \bar{h} &\geq \int_{Y\setminus E_\beta} ||T_\beta f||_y^2\, d\nu - 3\sqrt{\varepsilon}||T_\beta f||_y\, d\nu \\ &\geq ||T_\beta f||^2 - 3\sqrt{\varepsilon}||T_\beta f|| - \int_{E_\beta} |T_\beta f|^2\, d\nu.\end{aligned}$$

This implies (2). $f = 0$ follows from (2) and the equation

$$\lim_{\alpha\in\mathcal{F}} \int |\mathbf{E}(\bar{g}_i T_\beta f|\mathcal{Y})|\, d\mu = 0. \qquad \blacksquare$$

Now we are in position to formulate the main theorems in this section which are immediate corollaries of Propositions 3.4 and 3.5.

THEOREM 3.6. *$K(\mathcal{Y}, \Sigma)$ does not depend on the choice of $\mathcal{Y}$-sequence.*

THEOREM 3.7. *There is a unique orthogonal decomposition:*

$$L^2(X, \mathcal{B}, \mu) = K(\mathcal{Y}, \Sigma) \oplus M(\mathcal{Y}, \Sigma)$$

such that: $f \in M(\mathcal{Y}, \Sigma)$ if and only if

$$\lim_{\alpha\in\mathcal{F}} \int |\mathbf{E}(gT_\alpha f|\mathcal{Y})|\, d\mu = 0$$

for all $g \in L^2(X, \mathcal{B}, \mu)$.

we end this section with a proposition which will be used later.

PROPOSITION 3.8. *Let $\mathcal{Y} = (Y, \mathcal{C}, \mu, \Sigma)$ be a factor of a dynamical* IP-*system $X = (X, \mathcal{B}, \mu, \Sigma)$. Then, there exists a homomorphism $\psi\colon \mathcal{F} \longrightarrow \mathcal{F}$ satisfying: for any $\varepsilon > 0$ and any $f \in K(\mathcal{Y}, \Sigma)$, there is β_0 and bounded functions $g_1, g_2, \ldots, g_k$ such that if $\beta \cap \beta_0 = \emptyset$,*

$$\nu\Big\{y\,;\, \inf_{1\leq j\leq k} ||T_{\psi(\beta)} f - g_j||_y \geq \varepsilon\Big\} < \varepsilon.$$

PROOF. Let $\{\mathcal{F}_m\}$ be an $\mathcal{Y}$-sequence and let $K(\mathcal{Y},\Sigma) = \overline{\bigcup_m \mathbf{AP}(\mathcal{Y},\Sigma,\mathcal{F}_m)}$. We can choose a sequence $\{f_m \in \mathbf{AP}(\mathcal{Y},\Sigma,\mathcal{F}_m)\ ;\ m = 1,2,\ldots\}$ which is dense in $K(\mathcal{Y},\Sigma)$. Since $f_k \in \mathbf{AP}(\mathcal{Y},\Sigma,\mathcal{F}_m)$ for $k \leq m$, by the remark to Definition 2.6, there exist bounded functions $g_{1,k}, g_{2,k}, \ldots g_{r(m,k),k}$ and $\beta_{m,k} \in \mathcal{F}_m$ such that

$$\nu\Big\{y\ ;\ \inf_{1\leq j\leq r(m,k)} \|T_\beta f_k - g_{j,k}\|_y \geq \frac{1}{m}\Big\} < \frac{1}{m}$$

where $\beta \in \mathcal{F}_m$ and $\beta\cap\beta_{m,k} = \emptyset$ for $k \leq m$. Let $\beta'_m = \beta_{m,1}\cup\cdots\cup\beta_{m,m}$ and take a sequence $\{\alpha_m\}$ such that $\alpha_m \in \mathcal{F}_m$ and

$$\min\alpha_m > \max\{\beta'_1,\ldots,\beta'_{m-1},\alpha_1,\ldots,\alpha_{m-1}\}.$$

Define $\psi(\beta) = \bigcup_{j\in\beta}\alpha_j$. Then ψ is a homomorphism on $\mathcal{F}$.

For any $f \in K(\mathcal{Y},\Sigma)$ and any $\varepsilon > 0$, there is a f_m such that

$$\int |f - f_m|^2\, d\mu < (\varepsilon/2)^3.$$

Let $E' = \{y\ ;\ \int |f - f_m|^2\, d\mu_y \geq (\varepsilon/2)^2\}$, then $\nu(E') < \varepsilon/2$. Taking $n > m$ with $1/n < \varepsilon/2$, we have

$$\bigvee_{j\in\beta}\alpha_j \supset \psi(\beta)$$

for $\beta\cap\{1,2,\cdots,n\} = \emptyset$. Since $\alpha_j \in \mathcal{F}_n$ for $j > n$, $\varphi(\beta) \in \mathcal{F}_n$. Hence:

$$\nu\big\{y\ ;\ \inf_{1\leq j\leq r(n,m)} \|T_{\psi(\beta)}f - g_{j,m}\|_y \geq \varepsilon\big\}$$
$$\leq \nu(T^{-1}_{\psi(\beta)}E') + \nu\Big\{y\ ;\ \inf_{1\leq j\leq r(n,m)} \|T_{\psi(\beta)}f_m - g_{j,m}\|_y \geq \frac{1}{n}\Big\} < \varepsilon.$$

■

4. $\mathcal{Y}$-rigidity and mild $\mathcal{Y}$-mixing.

DEFINITION 4.1. Let I be an IP-set. A surjection $\varphi\colon \mathcal{F} \longrightarrow I$ is called an *I-representation* (of $\mathcal{F}$) if:

$$\varphi(\alpha\cup\beta) = \varphi(\alpha) + \varphi(\beta)$$

whenever $\alpha\cap\beta = \emptyset$.

It is clear that for any IP-set $I \subset \mathbf{N}$, at least one I-representation of $\mathcal{F}$ can be defined.

DEFINITION 4.2. Let I be an IP-set and φ be an I-representation (of $\mathcal{F}$). A set $A = \{a_i\} \subset I$ is φ-faithful if there exists a sequence $\{\alpha_i\} \subset \mathcal{F}$ with $\alpha_i\cap\alpha_j = \emptyset$ for $i \neq j$ such that $a_i = \varphi(\alpha_i)$ for all i. In particular, if $\max\alpha_i < \min\alpha_j$ for $i < j$, we will say that A is a *φ-monotone subset* of I.

Now we consider a single invertible measure preserving transformation T acting on a probability space $(X,\mathcal{B},\mu)$. Given an IP-set I and an I-representation φ, one can define an IP-system

$$\Sigma(I,\varphi) = \{T_\alpha = T^{\varphi(\alpha)}\ ;\ \alpha \in \mathcal{F}\}.$$

$\big(X,\mathcal{B},\mu,\Sigma(I,\varphi)\big)$ will be called a *dynamical* IP-*system* induced by (I,φ).

https://doi.org/10.4153/CJM-1993-022-2 Published online by Cambridge University Press

DEFINITION 4.3. Let $\mathcal{Y} = (Y, \mathcal{C}, \nu, T)$ be a factor of $\mathcal{X} = (X, \mathcal{B}, \mu, T)$, where T is a single measure preserving transformation. T is called *$\mathcal{Y}$-rigid* if there is an IP-set I and an I-representation φ such that:

(i) $\Sigma = \Sigma(I, \varphi)$ satisfies Condition A (for formulation of Condition A, see §2);
(ii) $K\big(\mathcal{Y}, \Sigma(I, \varphi)\big) = L^2(X, \mathcal{B}, \mu)$.

DEFINITION 4.4. Let $\mathcal{Y} = (Y, \mathcal{C}, \nu, T)$ be a factor of $\mathcal{X} = (X, \mathcal{B}, \mu, T)$, where T is a single measure preserving transformation. T is mildly $\mathcal{Y}$-mixing if for any $\varepsilon > 0$, the set

$$E_\varepsilon = \left\{ n \,;\, \int |\mathbf{E}(T^n f g|\mathcal{Y}) - T^n \mathbf{E}(f|\mathcal{Y})\mathbf{E}(g|\mathcal{Y})| \, d\mu < \varepsilon \right\}$$

is an IP*-set for all $f, g \in L^2(X, \mathcal{B}, \mu)$.

Now we give several equivalent statements for mild $\mathcal{Y}$-mixing and these statements will be used in next section to prove our main theorems.

THEOREM 4.1. *The following conditions are equivalent.*

(i) T is mildly $\mathcal{Y}$-mixing;

(ii) For any IP-*set I and I-representation φ, there is a sub*-IP-*system $\Sigma' \subset \Sigma = \Sigma(I, \varphi)$ such that*

$$\lim_{T_\alpha \in \Sigma'} \int |\mathbf{E}(T_\alpha f g|\mathcal{Y}) - T_\alpha \mathbf{E}(f|\mathcal{Y})\mathbf{E}(g|\mathcal{Y})| \, d\mu = 0$$

for all $f, g \in L^2(X, \mathcal{B}, \mu)$;

(iii) For every IP-*set I, there is an I-representation φ and a sub*-IP-*system $\Sigma' \subset \Sigma = \Sigma(I, \varphi)$ such that*

$$\lim_{T_\alpha \in \Sigma'} \int |\mathbf{E}(T_\alpha f g|\mathcal{Y}) - T_\alpha \mathbf{E}(f|\mathcal{Y})\mathbf{E}(g|\mathcal{Y})| \, d\mu = 0 \,;$$

(iv) For any IP-*set $I \subset \mathbf{N}$ and I-representation φ, there is a sub*-IP-*system $\Sigma' \subset \Sigma = \Sigma(I, \varphi)$ such that*

$$L^2\big(X, \theta^{-1}(\mathcal{C}), \mu\big) \supset K(\mathcal{Y}, \Sigma') \,;$$

(v) For any IP-*set $I \subset \mathbf{N}$, there is an I-representation φ and a sub*-IP-*system $\Sigma' \subset \Sigma = \Sigma(I, \varphi)$ such that*

$$L^2\big(X, \theta^{-1}(\mathcal{C}), \mu\big) \supset K(\mathcal{Y}, \Sigma') \,;$$

(vi) For every IP-*set I and I-representation φ,*

$$L^2\big(X, \theta^{-1}(\mathcal{C}), \mu\big) \supset \mathbf{AP}\big(\mathcal{Y}, \Sigma(I, \varphi), \mathcal{F}\big).$$

Before proving this theorem, we need following facts.

THEOREM 4.2 (HINDMAN [5]). *Let $\mathcal{F} = \bigcup_{j=1}^{r} D_j$. Then there exists a homomorphism $\psi \colon \mathcal{F} \longrightarrow \mathcal{F}$ and D_i with $1 \geq i \geq r$ such that $D_i \supset \psi(\mathcal{F})$.*

A proof of this theorem can be found in [5].

COROLLARY 4.3. *Let I contain an* IP-*set and φ be its representation. Then for any* IP*-*set I^*, there is an homomorphism $\psi \colon \mathcal{F} \longrightarrow \mathcal{F}$ such that $I \cap I^* \supset \varphi \circ \psi(\mathcal{F})$, and $\varphi \circ \psi(\mathcal{F})$ is an* IP-*set.*

PROOF. Let $D_1 = \{\alpha \in \mathcal{F} \ ; \ \varphi(\alpha) \in I \cap I^*\}$ and let $D_2 = \{\alpha \in \mathcal{F} \ ; \ \varphi(\alpha) \notin I^*\}$. Then $\mathcal{F} = D_1 \cup D_2$. By Theorem 4.2, there is a ψ such that $D_1 \supset \psi(\mathcal{F})$ or $D_2 \supset \psi(\mathcal{F})$. But I^* is an IP*-set, $\varphi(D_2)$ can not contain an IP-set. Hence $D_1 \supset \psi(\mathcal{F})$ which implies $I \cap I^* \supset \varphi \circ \psi(\mathcal{F})$. ∎

PROOF OF THEOREM 4.1. (ii) ⇒ (iii), (iii) ⇒ (i), (vi) ⇒ (iv) and (iv) ⇒ (v) are trivial. We only need to prove (i) ⇒ (ii), (ii) ⇒ (iv), (v) ⇒ (iii) and (iv) ⇒ (vi).

(i) ⇒ (ii) Let I be an IP-set and φ be an I-representation. For $f, g \in L^2(X, \mathcal{B}, \mu)$, let

$$E_\varepsilon = \Big\{ n \ ; \int |\mathbf{E}(T^n f g|\mathcal{Y}) - T^n \mathbf{E}(f|\mathcal{Y})\mathbf{E}(g|\mathcal{Y})| \, d\mu < \varepsilon \Big\}.$$

Now we inductively select $\{\alpha_n\}$ and $\{\psi_n : \mathcal{F} \longrightarrow \mathcal{F}\}$ such that:

- $\alpha_j \in \mathrm{Image}(\psi_j)$;
- $\alpha_i \cap \alpha_j = \emptyset$ for $i \neq j$;
- $\psi_1(\mathcal{F}) \supset \psi_2(\mathcal{F}) \supset \cdots$;
- $I \cap E_{1/2^n} \supset \varphi \circ \psi(\mathcal{F})$.

By Corollary 4.3, there is a $\psi_1 : \mathcal{F} \longrightarrow \mathcal{F}$ such that $I \cap E_{1/2^n} \supset \varphi \circ \psi_1(\mathcal{F})$. Take $\alpha_1 \in \psi(\mathcal{F})$ such that $\varphi(\alpha_1) \in E_{1/2^n}$. Suppose that we already have chosen $\alpha_1, \ldots, \alpha_{n-1}$ and $\psi_1, \ldots, \psi_{n-1}$. By Corollary 3.3, there is $\psi' : \mathcal{F} \longrightarrow \mathcal{F}$ such that

$$\phi \circ \psi_{n-1}(\mathcal{F}) \cap E_{1/2^{n-1}} \supset \varphi \circ \psi_{n-1} \circ \psi'(\mathcal{F}).$$

Now let $\psi_n = \psi_{n-1} \circ \psi'$. Since $\alpha_1 \cap \cdots \cap \alpha_{n-1}$ is a finite set, we can choose $\alpha_n \in \psi_n(\mathcal{F})$ such that $\alpha_i \cap \alpha_j = \emptyset$ for all $j < n$. Therefore α_n and ψ_n are selected.

Let $\psi : \mathcal{F} \longrightarrow \mathcal{F}$ defined by: $\psi(\beta) = \bigcap_{i \in \beta} \alpha_i$ and $I' = \varphi \circ \psi(\mathcal{F})$. We have a sub-IP-system

$$\Sigma_1 = \Sigma(I', \varphi \circ \psi) \subset \Sigma = \Sigma(I, \varphi \circ \psi).$$

Since $\alpha_m \in \psi_n(\mathcal{F})$ when $m \geq n$, $\psi(\beta) \in \psi_n(\mathcal{F})$ for all $\beta \cap 1, 2, \ldots, n = \emptyset$. Hence $\varphi \circ \psi(\beta) \in E_{1/2^n}$. This Σ_1 depends on f and g. Since $L^2(X, \mathcal{B}, \mu)$ is separable, we always can use diagonal method to obtain a subsystem $\Sigma' \subset \Sigma_1$ such that

$$\lim_{T_\alpha \in \Sigma'} \int |\mathbf{E}(T_\alpha f g|\mathcal{Y}) - T_\alpha \mathbf{E}(f|\mathcal{Y})\mathbf{E}(g|\mathcal{Y})| \, d\mu = 0$$

for all $f, g \in L^2(X, \mathcal{B}, \mu)$.

(ii) ⇒ (iv) Let I be an IP-set and let φ be an I-representation. By Corollary 2.4 and (ii), there is a sub-IP-system

$$\Sigma' \subset \Sigma = \Sigma(I, \varphi)$$

such that Σ' satisfies Condition A and

$$\lim_{T_\alpha \in \Sigma'} \int |\mathbf{E}(T_\alpha f g|\mathcal{Y}) - T_\alpha \mathbf{E}(f|\mathcal{Y})\mathbf{E}(g|\mathcal{Y})| \, d\mu = 0$$

for all $f, g \in L^2(X, \mathcal{B}, \mu)$. If $\mathbf{E}(f|\mathcal{Y}) = 0$, then

$$\lim_{T_\alpha \in \Sigma'} \int |\mathbf{E}(T_\alpha f g|\mathcal{Y})| \, d\mu = 0.$$

By Theorem 3.7, we have: if $f \perp L^2\big(X, \theta^{-1}(\mathcal{C}), \mu\big)$, then $f \perp K(\mathcal{Y}, \Sigma')$. This implies $L^2\big(X, \theta^{-1}(\mathcal{C}), \mu\big) \supset K(\mathcal{Y}, \Sigma')$.

(v) $\Rightarrow$ (iii) For any IP-set I, there is an I-representation φ and sub-IP-system $\Sigma' \subset \Sigma(I, \varphi)$ such that

$$L^2\big(X, \theta^{-1}(\mathcal{C}), \mu\big) \supset K(\mathcal{Y}, \Sigma).$$

Without loss of generality, we suppose that Σ' satisfies Condition A. By Theorem 3.7, we get that when $\mathbf{E}(f|\mathcal{Y}) = 0$

$$\lim_{T_\alpha \in \Sigma'} \int |\mathbf{E}(T_\alpha f g|\mathcal{Y})| \, d\mu = 0$$

This is equivalent to (iii).

(iv) $\Rightarrow$ (vi) For any IP-system $\Sigma(I, \varphi)$, by Corollary 2.4, there exists a sub-IP-system $\Sigma' \subset \Sigma(I, \varphi)$ satisfying Condition A. Then we have:

$$L^2\big(X, \theta^{-1}(\mathcal{C}), \mu\big) \supset K(\mathcal{Y}, \Sigma') \supset \mathbf{AP}(\mathcal{Y}, \Sigma', \mathcal{F}) \supset \mathbf{AP}\big(\mathcal{Y}, \Sigma(I, \varphi), \mathcal{F}\big). \qquad \blacksquare$$

5. Entropy characterizations of mild $\mathcal{Y}$-mixing and $\mathcal{Y}$-rigidity. In this section, we will prove the Theorems 1.2 and 1.3 which give conditional sequence entropy characterizations of $\mathcal{Y}$-rigidity and $\mathcal{Y}$-mild mixing for measure preserving transformations. As before, $\mathcal{X} = (X, \mathcal{B}, \mu, T)$ is a dynamical system and $\mathcal{Y} = (Y, \mathcal{C}, \nu, T)$ is a factor of $\mathcal{X}$. The notion of sequence entropy relative to a factor has been defined in the §1 (see Definition 1.1). Since conditional entropy for a finite partition has been used in Definition 1.1, we like to give the following definition for the sake of the completeness.

DEFINITION 5.1. If ξ is a finite and measurable partition of X, then entropy of ξ relative to $\mathcal{Y}$ is defined by:

$$H(\xi|\mathcal{Y}) = \int H_y(\xi) \, d\nu = \int \sum_{B \in \xi} -\mu_y(B) \log \mu_y(B) \, d\nu.$$

Proofs of the next two lemmas can be found in [6, p. 65].

LEMMA 5.1. *Let ξ, η be finite partitions. Then:*

$$|h_\Gamma(T, \xi|\mathcal{Y}) - h_\Gamma(T, \eta|\mathcal{Y})| \leq \int \big(H_y(\xi|\eta) + H_y(\eta|\xi)\big) \, d\nu$$

for any sequence $\Gamma \subset \mathbf{N}$.

LEMMA 5.2. *There exists a countable set $\{\xi_k\}$ of finite partitions such that* $\inf_k \int \big(H_y(\xi|\xi_k) + H_y(\xi_k|\xi)\big) \, d\nu = 0$ *for any finite partition ξ.*

Now we immediately have:

LEMMA 5.3. *There is a sequence $\big\{\xi_k = \{A_k, A_k^c\} \; ; k = 1, 2, \ldots\big\}$ of two cell partitions of X such that*

$$\inf_{i_1, \ldots, i_k} \int \big(H_y(\xi_{i_1} \vee \cdots \vee \xi_{i_k}|\xi) + H_y(\xi|\xi_{i_1} \vee \cdots \vee \xi_{i_k})\big) \, d\nu = 0$$

for any finite partition ξ.

The following result is Lemma 4.15 in [10].

LEMMA 5.4. *For each $\varepsilon > 0$ and $r \in \mathbf{N}$, there exists $\delta = \delta(\varepsilon, r) > 0$ (which depends only on ε and r) such that if $\xi = \{A_1, \ldots, A_r\}$, $\eta = \{C_1, \ldots, C_r\}$ are two r-cell partitions of any measurable space X with $\sum_{i=1}^{r} \mu(A_i \Delta C_i) < \delta$ then $H(\xi|\eta) + H(\eta|\xi) < \varepsilon$.*

Let $X_I = X \times [0, 1]$, $\mathcal{B}_I$ be the product σ-algebra of $\mathcal{B}$ and Borel σ-algebra on $[0, 1]$, and μ_λ be the product measure of μ and Lebesgue measure λ. Then we have a probability space $(X_I, \mathcal{B}_I, \mu_\lambda)$. Let $\pi: X_I \longrightarrow X$ defined by $(x, t) \longrightarrow x$. Then for any finite partition ξ of X, we have $H(\xi) = H\big(\pi^{-1}(\xi)\big)$. We abuse the terminology and also refer to H, H_y as the entropies on X_I with respect to $\mu \times \lambda$ and $\mu_y \times \lambda$ respectively. By direct computation, we have the following lemma.

LEMMA 5.5. *Let g be a measurable function from X to $[0, 1]$ and let B be a measurable subset of X. Define $B_I = \pi^{-1}(B)$ and $G_I = \{(x, y)\ ;\ 1 \geq y \geq g(x)\}$, then: $\mu_\lambda(B_I \Delta G_I) = \int |g - 1_B|\, d\mu$.*

We are now in position to prove Theorems 1.2 and 1.3.

PROOF OF THEOREM 1.2. (i) $\Rightarrow$ (ii). If T is $\mathcal{Y}$-rigid, there is an IP-set I and I-representation φ such that $K(\mathcal{Y}, \Sigma) = L^2(X, \mathcal{B}, \mu)$. Here $\Sigma = \Sigma(I, \varphi)$. By Proposition 3.8, there is a $\psi: \mathcal{F} \longrightarrow \mathcal{F}$ such that for any $\varepsilon > 0$ and $f \in L^2(X, \mathcal{B}, \mu)$, there is β_0 and $g_1, g_2, \ldots, g_k$ such that if $\beta \cap \beta_0 = \emptyset$,

$$\nu\Big\{y\ ;\ \inf_{1 \leq j \leq k} \|T_{\psi(\beta)} f - g_j\|_y \geq \varepsilon\Big\} < \varepsilon.$$

Now let $I' = \varphi \circ \psi(\mathcal{F})$ and let $\varphi' = \varphi \circ \psi$. We claim that for any φ'-monotone subset $\Gamma \subset I'$, $h_\Gamma(T|\mathcal{Y}) = 0$.

Suppose that $\{\alpha_n \in \mathcal{F}\}$ satisfies $\alpha_i \cap \alpha_j = \emptyset$ for $i \neq j$ and $\Gamma = \{t_n = \varphi'(\alpha_n)\}$. For any $\varepsilon > 0$ and any partition $\xi = \{A, A^c\}$, one can choose $\delta = \delta(\epsilon, 2)$ according to Lemma 5.4. Then by Proposition 3.8, there exist N_0 and functions $g_1, g_2, \ldots, g_k$ such that if $n > N_0$

$$\nu\Big\{y\ ;\ \inf_{1 \leq j \leq k} \|T^{t_n} 1_A - g_j\|_y \geq \frac{\delta}{2}\Big\} < \frac{\delta}{2}.$$

Let:

$$\hat{g}_i = \begin{cases} g_i & \text{for } 0 \geq g_i \geq 1 \\ 0 & \text{otherwise.} \end{cases}$$

Then it is clear that

$$\nu\Big\{y\ ;\ \inf_{1 \leq j \leq k} \|T^{t_n} 1_A - \hat{g}_j\|_y \geq \frac{\delta}{2}\Big\} < \frac{\delta}{2}.$$

Consider measure preserving transformation T_I on $X_I = X \times [0, 1]$ defined by $T_I(x, y) = (Tx, y)$. Let $G_{I,j} = \{(x, y)\ ;\ \hat{g} \geq y\}$ and partition $\eta_{I,j} = \{G_{I,j}, G^c_{I,j}\}$. Let

$$E_n = \Big\{y\ ;\ \inf_{1 \leq j \geq k} \|T^{t_n} 1_A - g_j\|_y \geq \delta/2\Big\},$$

then $\nu(E_n) < \delta/2$ for $n > N_0$. Thus by Lemma 5.5, for $y \notin E_n$

$$\inf_{1\leq J\leq k} \mu_y \times \lambda\big(T_I^{t_n}\pi^{-1}(A)\Delta G_{I,J}\big) = \inf_{1\leq J\leq k} \int |T^{t_n}1_A - g_J|\, d\mu_y \leq \inf_{1\leq J\leq k} \|T^{t_n}1_A - g_J\|_y < \delta/2.$$

By Lemma 5.4 and the definition of the δ, we have

$$\inf_{1\leq J\leq k} H_y\big(T_I^{t_n}\pi^{-1}(\xi)|\eta_{I,J}\big) < \varepsilon$$

for $n > N_0$ and $y \notin E_n$. Hence:

$$\begin{aligned}
\frac{1}{n}H\Big(\bigvee_{i=1}^{n} T^{t_n}\xi|\mathcal{Y}\Big) &= \frac{1}{n}\int H_y\Big(\bigvee_{i=1}^{n} T^{t_i}\xi\Big)\, d\nu = \frac{1}{n}\int H_y\Big(\bigvee_{i=1}^{n} T_I^{t_i}\pi^{-1}(\xi)\Big)\, d\nu \\
&\leq \frac{1}{n}\int H_y\Big(\bigvee_{i=1}^{n} T_I^{t_i}\pi^{-1}(\xi)\Big)\, d\nu \\
&\leq \frac{1}{n}\sum_{i=1}^{n}\int H_y\Big(T_I^{t_i}\pi^{-1}(\xi)\Big|\bigvee_{i=1}^{k}\eta_{I,J}\Big)\, d\nu \\
&\qquad + \frac{1}{n}\int H_y\Big(\bigvee_{i=1}^{k}\eta_{I,J}\Big)\, d\nu \\
&\leq \frac{1}{n}\sum_{i=1}^{n}\int \inf_{1\leq J\leq k} H_y\big(T_I^{t_i}\pi^{-1}(\xi)|\eta_{I,J}\big)\, d\nu + \frac{k}{n}\log 2 \\
&\leq \Big(\frac{N_0+k}{n} + \delta\Big)\log 2 + \varepsilon.
\end{aligned}$$

This gives $h_\Gamma(T,\xi|\mathcal{Y}) = 0$ for any 2-cell partition ξ. Let ξ be a partition of X and $\xi_1,\ldots,\xi_m$ are some 2-cell partitions of X. Then, by Lemma 5.1, we have:

$$\begin{aligned}
h_\Gamma(T,\xi|\mathcal{Y}) &\leq h_\Gamma\Big(T,\bigvee_{i=1}^{k}|\mathcal{Y}\Big) + \int\Big(H_y\Big(\bigvee_{i=1}^{k}\xi_i|\xi\Big) + H_y\Big(\xi|\bigvee_{i=1}^{k}\xi_i\Big)\Big)\, d\nu \\
&\leq \sum_{i=1}^{k} h_\Gamma(T,\xi_i|\mathcal{Y}) + \int\Big(H_y\Big(\bigvee_{i=1}^{k}\xi_i|\xi\Big) + H_y\Big(\xi|\bigvee_{i=1}^{k}\xi_i\Big)\Big)\, d\nu
\end{aligned}$$

By Lemma 5.3, we get $h_\Gamma(T,\xi|\mathcal{Y}) = 0$ for any finite partition ξ.

(ii) $\Rightarrow$ (iii) Let $\Gamma = \{\varphi(n)\}$. By definition one can get the result immediately.

(iii) $\Rightarrow$ (i) Let $\Gamma = \{a_n\}$. Define an IP-set $I_\Gamma = \{a_{n_1} + \cdots + a_{n_k}\ ;\ n_1 < \cdots < n_k\}$ and an I_Γ-representation $\varphi_\Gamma : \varphi_\Gamma(n) = a_n$. Then we have an IP-system $\Sigma = \Sigma(I_\Gamma, \varphi_\Gamma)$. Since every IP-system contains an IP-system satisfying Condition A (see Corollary 2.4), we can suppose without loss of generality that Σ_Γ satisfies Condition A. We need to prove $K(\mathcal{Y},\Sigma_\Gamma) = L^2(X,\mathcal{B},\mu)$. By Lemma 3.3, there exists a sub-σ-algebra $\mathcal{D}$ such that $K(\mathcal{Y},\Sigma_\Gamma) = L^2(X,\mathcal{D},\mu)$. If $K(\mathcal{Y},\Sigma_\Gamma) \neq L^2(X,\mathcal{B},\mu)$, there exists $B' \in \mathcal{B}$ with $B' \notin \mathcal{D}$. Take

$$D = \{x\ ;\ \varepsilon < \mathbf{E}(1_{B'}|\mathcal{D}) < 1-\varepsilon\} \in \mathcal{D}.$$

Then $\mu(D) > 0$ for some $\varepsilon > 0$ (Otherwise $\mathbf{E}(1_{B'}|\mathcal{D}) = 1$ or 0 which means $B' \in \mathcal{D}$). Let $B = B' \cap D$. Then $\varepsilon < \mathbf{E}(1_B|\mathcal{D}) < 1-\varepsilon$ on D; $\mathbf{E}(1_B|\mathcal{D}) = 0$ on D^c. Taking the partition

$\xi = \{B, B^c\}$, we will find an φ-monotone subset $\{t_n\}$ of I_Γ such that $h_{\{t_n\}}(T, \xi|\mathcal{Y}) > 0$ which gives a contradiction.

Suppose that $t_1 = \varphi_\Gamma(\alpha_2), t_2 = \varphi_\Gamma(\alpha_2), \ldots, t_{n-1} = \varphi_\Gamma(\alpha_{n-1})$ have been chosen such that

$$\mu_y(T^{t_j}B \cap C) < \big(1 - (1/2)\varepsilon\big)\mu_y(C)$$

for all $C \in \bigvee_{\iota=1}^{j-1} T^{t_\iota}\xi$, $y \notin E_j$ with $\nu(E_j) < 2^{-j}$ and $j \leq n-1$. Let $f_B = 1_B - \mathbf{E}(1_B|\mathcal{D})$, then $f_B \perp K(\mathcal{Y}, \Sigma_\Gamma)$. By Proposition 3.4, we get:

$$\lim_{\alpha \in \mathcal{F}} \int \left| \int T_\alpha f_B 1_C \, d\mu_y \right| d\nu = 0$$

for all $C \in \bigvee_{\iota=1}^{j-1} T^{t_\iota}\xi$ and $j \leq n-1$. We may choose $t_n = \varphi_\Gamma(\alpha_n)$ such that:

- $\alpha_n \cap \alpha_\iota = \emptyset$ for $1 \leq i \leq n-1$;
- $\int T^{t_n} f_B 1_C \, d\mu_y \leq (1/2)\varepsilon\mu_y(C)$ for all $C \in \bigvee_{\iota=1}^{n-1} T^{t_\iota}\xi$ and $y \notin E_n$ with $\nu(E_n) < 2^{-n}$.

Recall that $\varepsilon < \mathbf{E}(1_B|\mathcal{D}) < 1-\varepsilon$. Then for all $C \in \bigvee_{\iota=1}^{n-1} T^{t_\iota}\xi$ and $y \notin E_n$ with $\nu(E_n) < 2^{-n}$

$$\begin{aligned} \mu_y(T^{t_n}B \cap C) &= \int T^{t_n}\mathbf{E}(1_B|\mathcal{D})1_C \, d\mu_y + \int T^{t_n} f_B 1_C \, d\mu_y \\ &\leq (1-\varepsilon)\mu_y(C) + \frac{1}{2}\varepsilon\mu_y(C). \end{aligned}$$

Inductively, $\{t_n\}$ can be selected.

Now we compute the entropy along $\{t_n\}$. If $y \notin \bigcup_{\iota=n_0}^{\infty} E_\iota$, then for $n > n_0$

$$\begin{aligned} H_y\Big(T^{t_n}\xi | \bigvee_{\iota=1}^{n-1} T^{t_\iota}\xi\Big) &\geq -\sum_{B,C} \mu_y(T^{t_n}B \cap C)\big(\log \mu_y(T^{t_n}B \cap C) - \log \mu_y(C)\big) \\ &\geq -\sum_{C} \mu_y(T^{t_n}B \cap C)\Big(\log\Big(1 - \frac{1}{2}\varepsilon\Big)\mu_y(C) - \log \mu_y(C)\Big) \\ &\geq -\mu_y(T^{t_n}B)\log\Big(1 - \frac{1}{2}\varepsilon\Big). \end{aligned}$$

where the sums are taken over $C \in \bigvee_{\iota=1}^{n-1} T^{t_\iota}\xi$. Since $\nu(\bigcap_{n=1}^{\infty}\bigcup_{\iota=n}^{\infty} E_\iota) = 0$, we have

$$\liminf_{n\to\infty} H_y\Big(T^{t_n}\xi | \bigvee_{\iota=1}^{n-1} T^{t_\iota}\xi\Big) + \mu_y(T^{t_n}B)\log\Big(1 - \frac{1}{2}\varepsilon\Big) \geq 0$$

for a.e. (ν) $y \in Y$. By Fatou lemma,

$$\liminf_{n\to\infty} \int \Big(H_y\Big(T^{t_n}\xi | \bigvee_{\iota=1}^{n-1} T^{t_\iota}\xi\Big) + \mu_y(T^{t_n}B)\log\Big(1 - \frac{1}{2}\varepsilon\Big)\Big) d\nu \geq 0.$$

Therefore $h_{\{t_n\}}(T, \xi|\mathcal{Y}) \geq -\mu(B)\log(1 - \frac{1}{2}\varepsilon) > 0$. Since $\{t_n\}$ is an φ_Γ-faithful subset of I_Γ, we get a contradiction which implies T is $\mathcal{Y}$-rigid. ∎

PROOF OF THEOREM 1.3. (i) $\Rightarrow$ (ii) Let $\Sigma = \Sigma(I, \varepsilon)$. By Lemma 5.2, there is a countable set $\{\xi_k\}$ of finite partitions such that

$$\inf_k \int \big(H_y(\xi|\xi_k) + H_y(\xi_k|\xi)\big)\, d\nu = 0$$

for any finite partition ξ. We define inductively the sequence $\Gamma = \{t_n\}$ as follows. Let $t_1 = \varphi(\alpha_1)$ be any number and suppose that $t_1 = \varphi(\alpha_1), t_2 = \varphi(\alpha_2), \ldots, t_{n-1} = \varphi(\alpha_{n-1})$ have been defined. Let

$$N_n = \max_{1\le k\le n}\Big\{\#\Big\{\bigvee_{i=1}^{n-1} T^{t_i}\xi_k\Big\}\Big\}.$$

Choose δ_n such that for any $u, v \in [0,1]$ $|u-v| < \delta_n$ implies

$$|u\log u - v\log v| < \frac{1}{nN_n}.$$

Since T is mildly $\mathcal{Y}$-mixing, there is a sub-IP-system $\Sigma' \subset \Sigma$ such that

$$\lim_{T_\beta\in\Sigma'} \int \Big|\mu_y(T_\beta E\cap B) - \mu_y(T_\beta E)\mu_y(B)\Big|\, d\nu = 0$$

for all $E \in \xi_k$, $B \in \bigvee_{i=1}^{n-1} T^{t_i}\xi_k$ and $k \le n$. Thus for $\delta_n > 0$ there exists a β_0 such that

$$\nu\{y\,;\,|\mu_y(T_\beta E\cap B) - \mu_y(T_\beta E)\mu_y(B)| < \delta_n\} > 1 - \frac{1}{2^n}$$

for $\beta\cap\beta_0 = \emptyset$, $E \in \xi_k$, $B \in \bigvee_{i=1}^{n-1} T^{t_i}\xi_k$ and $k \le n$. Choose α_n such that $\alpha_n\cap\beta_0 = \emptyset$ and $\alpha_n\cap\alpha_j = \emptyset$ for $1 \le j \le n-1$.

Let

$$C_n = \{y\,;\,|\mu_y(T^{t_n}E\cap B) - \mu_y(T^{t_n}E)\mu_y(B)| < \delta_n\}$$

and note that $\nu(\bigcup_{j=1}^\infty \bigcap_{i=j}^\infty C_i) = 1$. Fix k and let $y \in \bigcap_{i=j}^\infty C_i$ for some j. Since

$$|\mu_y(T_{t_n}E\cap B) - \mu_y(T_{t_n}E)\mu_y(B)| < \delta_n$$

for $n > \max\{j,k\}$, we get:

$$\Big|-\sum_{E\,B}\mu_y(T^{t_n}E\cap B)\log\mu_y(T^{t_n}E\cap B) + \sum_{E\,B}\mu_y(T^{t_n}E)\mu_y(B)\log\mu_y(T^{t_n}E)\mu_y(B)\Big|$$

$$\le \sum_{E\,B}\frac{1}{nN_n} \le \sum_E \frac{1}{n} \tag{3}$$

where the sums are taken over all sets $E \in \xi_k$ and $B \in \bigvee_{i=1}^{n-1} T^{t_i}\xi_k$ for $k \le n$. Since:

$$\sum_{E\,B}\mu_y(T^{t_n}E)\mu_y(B)\log\mu_y(T^{t_n}E)\mu_y(B) = \sum_E \mu_y(T^{t_n}E)\log\mu_y(T^{t_n}E) + \sum_B \mu_y(B)\log\mu_y(B),$$

(3) implies

$$H_y\Big(\bigvee_{i=1}^n T^{t_i}\xi_k\Big) - H_y(T^{t_n}\xi_k) - H_y\Big(\bigvee_{i=1}^{n-1} T^{t_i}\xi_k\Big) < \sum_E \frac{1}{n} = \frac{\#\{\xi\}}{n} \longrightarrow 0.$$

One also can check that $|H_y(\bigvee_{i=1}^n T^{t_i}\xi_k) - H_y(T^{t_n}\xi_k) - H_y(\bigvee_{i=1}^{n-1} T^{t_i}\xi_k)|$ is a bounded function on Y with the bound $2\log\#\{\xi_k\}$. By Lebesgue Dominated Convergence Theorem:

$$H\Big(\bigvee_{i=1}^{n} T^{t_i}\xi_k|\mathcal{Y}\Big) - H(T^{t_n}\xi_k|\mathcal{Y}) - H\Big(\bigvee_{i=1}^{n-1} T^{t_i}\xi_k|\mathcal{Y}\Big) \to 0.$$

So for any k, $H(\bigvee_{i=1}^n T^{t_i}\xi_k|\mathcal{Y}) - H(\bigvee_{i=1}^{n-1} T^{t_i}\xi_k|\mathcal{Y}) \to H(\xi_k|\mathcal{Y})$. Therefore $h_\Gamma(T,\xi_k|\mathcal{Y}) = H(\xi_k|\mathcal{Y})$ for all k. By Lemmas 5.1 and 5.2,

$$h_\Gamma(T,\xi|\mathcal{Y}) = H(\xi|\mathcal{Y})$$

for all finite partitions ξ.

(ii) $\Rightarrow$ (iii) By the definition of the IP-set we can get the result immediately.

(iii) $\Rightarrow$ (i) Assume that T is not mildly $\mathcal{Y}$-mixing. According to Theorem 4.1, there is an IP-set I and I-representation φ such that $\Sigma = \Sigma(I,\varphi)$ satisfies Condition A and $K(\mathcal{Y},\Sigma) \supset L^2\big(X,\theta^{-1}(C),\mu\big)$ but $K(\mathcal{Y},\Sigma) \neq L^2\big(X,\theta^{-1}(C),\mu\big)$. By Lemma 3.3, there exists a sub-σ-algebra $\mathcal{D}$ such that $K(\mathcal{Y},\Sigma) = L^2(X,\mathcal{D},\mu)$. Then $(X,\mathcal{D},\mu,T)$ is a $\mathcal{Y}$-rigid dynamical system. Since $\mathcal{D} \supset \theta^{-1}(C)$ but $\mathcal{D} \neq \theta^{-1}(C)$, there is a $B \in \mathcal{D}$ but $B \notin \theta^{-1}(C)$. Take $\xi = \{B, B^c\}$, then $H(\xi|\mathcal{Y}) > 0$. By Theorem 1.2, there exists a subset Γ of $\mathbf{N}$ such that for any sequence $\{\Gamma_n\}$ of pairwise disjoint finite subsets of $\mathbf{N}$ $h_{\tilde\Gamma}(T,\xi|\mathcal{Y}) = 0$. Here $\tilde\Gamma = \{t_n = \sum_{\alpha\in\Gamma_n} a\}$. Hence $h_\Gamma(T,\xi|\mathcal{Y}) \neq H(\xi|\mathcal{Y})$. This contradiction finishes the proof of our theorem. ∎

REFERENCES

1. H Furstenberg, *IP-systems in Ergodic Theory*, Contemporary Mathematics, Conference in Modern Analysis and Prob **26** (1984), 131–148
2. ———, *Recurrence in Ergodic Theory and Combinatorial Number Theory*, Princeton University Press, Princeton, N J, 1981
3. H Furstenberg and Y Katznelson, *An Ergodic Szemeredi Theorem for IP-systems and Combinatorial Theory*, J d'Analyse Math **45**(1985), 117–169
4. H Furstenberg and B Weiss, *Topological Dynamics and Combinatorial Number Theory*, J d'Analyse Math **34** (1978), 161–185
5. N Hindman, *Finite Sums from Sequence within One Cell of a Partition of* **N**, J Combinatorial Theory (A) **17**(1974), 1–11
6. P Hulse, *Sequence Entropy Relative to an Invariant σ-algebra*, J London Math Soc (2) **33**(1986), 59–72
7. A G Kushnirenko, *On Metric Invariants of Entropy Type*, Russian Math Surveys (5) **22**(1967), 53–61
8. V A Rokhlin, *Lectures on the Entropy Theory of Transformations with Invariant Measure*, Russian Math Surveys (5) **22**(1967), 1–52
9. A Saleski, *Sequence Entropy and Mixing*, J Math Anal Appl **60**(1977), 58–66
10. P Walters, *An Introduction to Ergodic Theory*, Springer-Verlag, New York, 1982
11. K Yosida, *Functional Analysis*, Springer-Verlag, New York, 1968
12. Q Zhang, *Sequence Entropy and Mild Mixing*, Can J Math (1) **44**(1992), 215–224

13. R J Zimmer, *Extensions of ergodic group actions*, Illinois J Math **20**(1976), 373–409
14. ———, *Ergodic actions with generalized discrete spectrum*, Illinois J Math **20**(1976), 555–588

Department of Mathematics
The Ohio State University
Columbus, Ohio 43210
U S A