SEQUENCE ENTROPY AND RIGID $\sigma-$ALGEBRAS

ALVARO CORONEL, ALEJANDRO MAASS, AND SONG SHAO

Abstract. In this paper we study relationships between sequence entropy and the Kronecker and rigid algebras. Let $(Y, \mathcal{Y}, \nu, T)$ be a factor of the measure-theoretical dynamical system $(X, \mathcal{X}, \mu, T)$ and S be a sequence of positive integers with positive upper density. We prove there exists a subsequence $A \subseteq S$ such that $h_\mu^A(T, \xi|\mathcal{Y}) = H_\mu(\xi|\mathcal{K}(X|Y))$ for all finite partitions ξ, where $\mathcal{K}(X|Y)$ is the Kronecker algebra over $\mathcal{Y}$. A similar result holds for the rigid algebras over $\mathcal{Y}$. As an application, we characterize compact, rigid and mixing extensions via relative sequence entropy.

1. Introduction

Sequence entropy for a measure was introduced as an isomorphism invariant by Kushnirenko [Ku], who used it to distinguish between transformations with the same entropy. It is also a spectral invariant. He proved that an invertible measure preserving transformation has discrete spectrum if and only if the sequence entropy of the system is zero for any sequence [Ku]. Later, sequence entropy was mainly used to characterize different kinds of mixing properties in [S, Hu1, Hu2, Z1, Z2, HSY]. Also, in the articles [BD, KY] the relation between large sets of integers and mixing properties was considered.

The purpose of this paper is to study the relationship between sequence entropy and some important σ-algebras associated to a measure-theoretical dynamical system. More precisely, the Kronecker and the rigid algebras, and its relative versions.

Let $(X, \mathcal{X}, \mu, T)$ be a measure-theoretical dynamical system. It is not difficult to prove by using standard properties of the entropy and the Pinsker σ-algebra that given a finite measurable partition ξ

$$\lim_{n\to\infty} h_\mu(T^n, \xi) = H_\mu(\xi|\Pi(X)),$$

where $\Pi(X)$ is the Pinsker σ-algebra of the system (see [P] for general properties and [B-R] for an explicit proof). One can restate this result using the language of sequence entropy as follows

$$\sup_{A=n\mathbb{Z}_+, n\in\mathbb{N}} h_\mu^A(T, \xi) = H_\mu(\xi|\Pi(X)).$$

Date: 24.02.2009.

2000 *Mathematics Subject Classification.* Primary: 37A35, 37A25.

Key words and phrases. sequence entropy, compact and rigid factors.

The authors are supported by Nucleus Millennium Information and Randomness P04-069-F and Grant Basal-CMM. And the third author is supported by National Natural Science Foundation of China(10871186).

What happens if one takes the supremum over another increasing sequence of positive integers A ? It was shown in [HMY] that

$$\max_{A\subseteq \mathbb{Z}_+} h_\mu^A(T,\xi) = H_\mu(\xi|\mathcal{K}(X)), \tag{1.1}$$

where $\mathcal{K}(X)$ is the Kronecker algebra of the system and ξ is any finite measurable partition.

In this paper we address the previous question for conditinal sequence entropy with respect to a factor. Let $(Y,\mathcal{Y},\nu,T)$ be a factor of $(X,\mathcal{X},\mu,T)$.

First we show in section 3 (Theorem 3.4) that for any given increasing sequence of positive integers S with positive upper density

$$\max_{A\subseteq S} h_\mu^A(T,\xi|\mathcal{Y}) = H_\mu(\xi|\mathcal{K}(X|Y))$$

for any finite measurable partition ξ, where $\mathcal{K}(X|Y)$ is the Kronecker algebra relative to $\mathcal{Y}$. As a corollary (Corollary 3.5) we sligthly extend (1.1) by proving that

$$\max_{A\subseteq S} h_\mu^A(T,\xi) = H_\mu(\xi|\mathcal{K}(X)).$$

Then in section 4 we consider rigid algebras associated to the system relative to the factor. We prove (Theorem 4.11) that for any IP set F' there exists an IP-subset F of F' such that

$$\max\left\{h_\mu^A(T,\xi|\mathcal{Y}) : A\subseteq F \text{ is } \mathcal{F}\text{-monotone}\right\} = H_\mu(\xi|\mathcal{K}_F(X|Y)), \tag{1.2}$$

for all finite measurable partitions ξ, where $\mathcal{K}_F(X|Y)$ is the rigid algebra relative to $\mathcal{Y}$ along F (refer to Section 4 for related concepts). The analogous result of (1.1) for rigid algebras is given in Corollary 4.13.

Two applications of previous results are presented.

A first one is the characterization of compact and weakly mixing extensions and rigid and mild mixing extensions via conditional sequence entropy, providing new proofs and slightly more general statements to results in [Hu1, Hu2] and [Z1, Z2] respectively.

Another application is given in section 5. We show that $\max_A\{h_\mu^A(T|\mathcal{Y})\} \in \{\log k : k\in\mathbb{N}\}\cup\{\infty\}$.

In Section 2 we give some basic concepts and results in ergodic theory and entropy theory.

2. Preliminaries

In this article, integers, nonnegative integers, natural numbers and complex numbers are denoted by $\mathbb{Z}$, $\mathbb{Z}_+$, $\mathbb{N}$ and $\mathbb{C}$ respectively.

2.1. **Basic concepts.** Let $(X,\mathcal{X},\mu)$ be a standard Borel space with μ a regular probability measure on X and let $T : X\to X$ be an invertible measure-preserving transformation. The quadruple $(X,\mathcal{X},\mu,T)$ is called *measure-theoretical dynamical system*, or just *system*, if $T\mu=\mu$, that is, $\mu(B)=\mu(T^{-1}B)$ for all $B\in\mathcal{X}$. For simplicity, in the sequel all transformations of a system are called T.

A system $(X, \mathcal{X}, \mu, T)$ is *ergodic* if any measurable set $A \in \mathcal{X}$ for which $\mu(A \Delta T^{-1}A) = 0$ has $\mu(A) = 0$ or $\mu(A) = 1$. A system $(X, \mathcal{X}, \mu, T)$ is *weakly mixing* if $(X \times X, \mathcal{X} \otimes \mathcal{X}, \mu \otimes \mu, T \times T)$ is ergodic; and it is *strongly mixing* if $\lim_{n\to\infty} \mu(T^{-n}A \cap B) = \mu(A)\mu(B)$ for any $A, B \in \mathcal{X}$.
A system $(Y, \mathcal{Y}, \nu, T)$ is a *factor* of $(X, \mathcal{X}, \mu, T)$ if there exists a measurable map $\pi : X \to Y$ such that $\pi\mu = \nu$ and $\pi \circ T = T \circ \pi$. Equivalently one says that $(X, \mathcal{X}, \mu, T)$ is an *extension* of $(Y, \mathcal{Y}, \nu, T)$.
Let $(Y, \mathcal{Y}, \nu, T)$ be a factor of $(X, \mathcal{X}, \mu, T)$. One can identify $L^2(Y, \mathcal{Y}, \nu)$ with the subspace $L^2(X, \pi^{-1}(\mathcal{Y}), \mu)$ of $L^2(X, \mathcal{X}, \mu)$ via $f \mapsto f \circ \pi$. By using this identification it is possible to define the projection of $L^2(X, \mathcal{X}, \mu)$ into $L^2(Y, \mathcal{Y}, \nu)$: $f \mapsto \mathbb{E}(f|\mathcal{Y})$. The *conditional expectation* $\mathbb{E}(f|\mathcal{Y})$ is characterized as the unique $\mathcal{Y}$-measurable function in $L^2(Y, \mathcal{Y}, \nu)$ such that

$$\int_Y g\mathbb{E}(f|\mathcal{Y})d\nu = \int_X g \circ \pi f d\mu \tag{2.1}$$

for all $g \in L^2(Y, \mathcal{Y}, \nu)$.
The *disintegration* of μ over ν is given by a measurable map $y \mapsto \mu_y$ from Y to the space of probability measures on X such that

$$\mathbb{E}(f|\mathcal{Y})(y) = \int_X f d\mu_y \tag{2.2}$$

ν-almost everywhere.
The *self-joining of* $(X, \mathcal{X}, \mu, T)$ *relatively independent over the factor* $(Y, \mathcal{Y}, \nu, T)$ is the system $(X \times X, \mathcal{X} \otimes \mathcal{X}, \mu \times_Y \mu, T \times T)$, where the measure $\mu \times_Y \mu$ is defined by

$$(\mu \times_Y \mu)(B) = \int_Y \mu_y \times \mu_y(B) d\nu(y), \ \forall B \in \mathcal{X} \otimes \mathcal{X}. \tag{2.3}$$

This measure is characterized by

$$\int_{X\times X} f_1 \otimes f_2 d\mu \times_Y \mu = \int_Y \mathbb{E}(f_1|\mathcal{Y})\mathbb{E}(f_2|\mathcal{Y})d\nu \tag{2.4}$$

for all $f_1, f_2 \in L^2(X, \mathcal{X}, \mu)$, where $f_1 \otimes f_2(x_1, x_2) = f_1(x_1)f_2(x_2)$.
For details about the concepts in this subsection see [F1, G].

2.2. **Kronecker systems and rigid systems.** Let $(X, \mathcal{X}, \mu, T)$ be a system. An *eigenfunction* of T is a non-zero complex valued function $f \in L^2(X, \mathcal{X}, \mu)$ such that $Tf = \lambda f$ for some $\lambda \in \mathbb{C}$, where $Tf = f \circ T$. The complex number λ is called *eigenvalue of* T *associated to* f. If $f \in L^2(X, \mathcal{X}, \mu)$ is an eigenfunction of T, then $\mathrm{cl}\{T^n f : n \in \mathbb{Z}\}$ is a compact subset of $L^2(X, \mathcal{X}, \mu)$. In general, one says that f is *compact* if $\mathrm{cl}\{T^n f : n \in \mathbb{Z}\}$ is compact in $L^2(X, \mathcal{X}, \mu)$. Let $H_c(T)$ be the set of all compact functions in $L^2(X, \mathcal{X}, \mu)$. It is well known that $H_c(T)$ is the closure of the set spanned by all the eigenfunctions of T.
The following proposition is a classical result (see for example [Zi]).

Proposition 2.1. *Let* $(X, \mathcal{X}, \mu, T)$ *be a system and* H *be an algebra of bounded functions in* $L^2(X, \mathcal{X}, \mu)$ *which is invariant under complex conjugation. Then there* *exists a sub-σ-algebra $\mathcal{A}$ of $\mathcal{X}$ such that* $\mathrm{cl}(H) = L^2(X, \mathcal{A}, \mu)$. *Moreover, if H is T-invariant, then $\mathcal{A}$ is T-invariant.*

One easily deduces from Proposition 2.1 that there exists a T-invariant sub-σ-algebra $\mathcal{K}(X)$ of $\mathcal{X}$ such that $H_c(T) = L^2(X, \mathcal{K}(X), \mu)$. $\mathcal{K}(X)$ is called the *Kronecker algebra of* $(X, \mathcal{X}, \mu, T)$. The system $(X, \mathcal{X}, \mu, T)$ is said to be *compact* or to have *discrete spectrum* if $H_c(T) = L^2(X, \mathcal{X}, \mu)$ or equivalently $\mathcal{K}(X) = \mathcal{X}$.

A function $f \in L^2(X, \mathcal{X}, \mu)$ is *rigid* if there exists an increasing sequence $\{t_n\}_{n\in\mathbb{N}} \subseteq \mathbb{Z}_+$ with $\lim_{n\to\infty} T^{t_n} f = f$ in $L^2(X, \mathcal{X}, \mu)$. For a fixed sequence $F = \{t_n\}_{n\in\mathbb{N}} \subseteq \mathbb{Z}_+$, $H_F(T)$ denotes the set of all functions $f \in L^2(X, \mathcal{X}, \mu)$ with $\lim_{n\to\infty} T^{t_n} f = f$ in $L^2(X, \mathcal{X}, \mu)$. It is easy to see that the set of all bounded functions in $H_F(T)$ forms a T-invariant and invariant under complex conjugation algebra of $L^2(X, \mathcal{X}, \mu)$. Thus from Proposition 2.1 one deduces that there exists a T-invariant sub-σ-algebra $\mathcal{K}_F(X)$ of $\mathcal{X}$ such that $H_F(T) = L^2(X, \mathcal{K}_F(X), \mu)$. A system $(X, \mathcal{X}, \mu, T)$ is called *rigid* if there is $F = \{t_n\}_{n\in\mathbb{N}} \subseteq \mathbb{Z}_+$ such that $H_F(T) = L^2(X, \mathcal{X}, \mu)$.

2.3. **Mixing properties and filters.** A system $(X, \mathcal{X}, \mu, T)$ is *mild mixing* if it does not have non-constant rigid functions. In the strongly mixing case $\mathcal{K}_F(X)$ is trivial for any sequence F, thus strong mixing implies mild mixing. Also, since every eigenfunction is rigid, mild mixing implies weak mixing. It can be proved that mild mixing is strictly between weak mixing and strong mixing [FW]. For more details on mixing properties see [F1, F2, W2].
A hereditary upward collection of subsets $\mathcal{G}$ of $\mathbb{Z}_+$ is said to be a *family*. That is, subsets of $\mathbb{Z}_+$ containing elements of $\mathcal{G}$ are in $\mathcal{G}$ too. If a family $\mathcal{G}$ is closed under finite intersections and satisfies $\emptyset \notin \mathcal{G}$, then it is called a *filter*. The *dual* of a family $\mathcal{G}$ is $\mathcal{G}^* = \{F \subseteq \mathbb{Z}_+ | F \cap F' \neq \emptyset \ for\ all\ F' \in \mathcal{F}\}$.

Now some important families are introduced. Let A be a subset of either $\mathbb{Z}_+$ or $\mathbb{Z}$. The *upper Banach density of A* is

$$d^*(A) = \limsup_{|I|\to\infty} \frac{|A \cap I|}{|I|} \ ,$$

where I ranges over all intervals of $\mathbb{Z}_+$ or $\mathbb{Z}$ and $|\cdot|$ denotes the cardinality of the set. The *upper density* of a subset A of $\mathbb{Z}_+$ is

$$\bar{d}(A) = \limsup_{N\to\infty} \frac{|A \cap \{0, \dots, N-1\}|}{N}.$$

(If A is subset of $\mathbb{Z}$, then $\bar{d}(A) = \limsup_{N\to\infty} \dfrac{|A \cap \{-N, \dots, N\}|}{2N+1}$). The *lower Banach density* $d_*(A)$ and the *lower density* $\underline{d}(A)$ are defined analogously, with lim inf. If $\bar{d}(A) = \underline{d}(A)$, then one says A has density $d(A)$. Let $\mathcal{D} = \{A \subseteq \mathbb{Z}_+ : d(A) = 1\}$ and $\mathcal{BD} = \{A : d_*(A) = 1\}$. It is easy to see that $\mathcal{D}$ and $\mathcal{BD}$ are filters with duals $\mathcal{D}^* = \{A \subseteq \mathbb{Z}_+ | \bar{d}(A) > 0\}$ and $\mathcal{BD}^* = \{A : d^*(A) > 0\}$ respectively.
Let $\{b_i\}_{i\in I}$ be a finite or infinite sequence in $\mathbb{N}$. Define

$$\mathrm{FS}(\{b_i\}_{i\in I}) = \Big\{ \sum_{i\in\alpha} b_i : \alpha \text{ is a finite non-empty subset of } I \Big\}.$$

F is an IP *set* if there exists a sequence of natural numbers $\{b_i\}_{i\in\mathbb{N}}$ such that $F = \mathrm{FS}(\{b_i\}_{i\in\mathbb{N}})$. Denote the set of all IP sets by $\mathcal{IP}$.

Let $\{x_n\}_{n\in\mathbb{Z}_+}$ be a sequence in a metric space (X,d), $x\in X$ and $\mathcal{G}$ be a family. One says that x_n *$\mathcal{G}$-converges to* x, which is denoted by $\mathcal{G}-\lim x_n = x$, if for any neighborhood U of x, $\{n\in\mathbb{Z}_+ : x_n\in U\}\in\mathcal{G}$. The following is a well known result concerning mixing in ergodic theory (see [F1, F2]).

Theorem 2.2. *Let $(X,\mathcal{X},\mu,T)$ be a system. Then,*

(1) *T is weak mixing if and only if $\mathcal{D}-\lim\mu(T^{-n}A\cap B)=\mu(A)\mu(B)$ for any $A,B\in\mathcal{X}$;*
(2) *T is mild mixing if and only if $\mathcal{IP}^*-\lim\mu(T^{-n}A\cap B)=\mu(A)\mu(B)$ for any $A,B\in\mathcal{X}$.*

For more discussion about various kinds of mixing and families, refer to [BD, KY].

2.4. **Sequence entropy and conditional sequence entropy.** Let $(X,\mathcal{X},\mu,T)$ be a system and $S=\{t_i\}_{i\in\mathbb{N}}$ be an increasing sequence of non-negative integers. Let ξ be a measurable partition and $\mathcal{A}$ a sub-σ-algebra of $\mathcal{X}$. The *(Shannon) entropy of* ξ and the *(Shannon) entropy of ξ given $\mathcal{A}$* are given respectively by

$$H_\mu(\xi) = -\sum_{A\in\xi}\mu(A)\log\mu(A)$$

and

$$H_\mu(\xi|\mathcal{A}) = \sum_{A\in\xi}\int_X -\mathbb{E}(1_A|\mathcal{A})\log\mathbb{E}(1_A|\mathcal{A})d\mu.$$

One also uses the notation $\mu(A|\mathcal{A})=\mathbb{E}(1_A|\mathcal{A})$.
Let ξ, η be measurable partitions with $H_\mu(\xi|\mathcal{A})<\infty, H_\mu(\eta|\mathcal{A})<\infty$ and identify (when necessary) η with the σ-algebra it induces. It is known that $H_\mu(\xi|\mathcal{A})$ increases with respect to ξ and decreases with respect to $\mathcal{A}$ and

$$H_\mu(\xi\vee\eta|\mathcal{A}) = H_\mu(\xi|\eta\vee\mathcal{A}) + H_\mu(\eta|\mathcal{A}).$$

Definition 2.3. The *conditional sequence entropy along S of ξ given $\mathcal{A}$ in the system* $(X,\mathcal{X},\mu,T)$ is defined by

$$h_\mu^S(T,\xi|\mathcal{A}) = \limsup_{n\to\infty}\frac{1}{n}H_\mu(\bigvee_{i=1}^n T^{-t_i}\xi|\mathcal{A})\left(=\limsup_{n\to\infty}\frac{1}{n}\sum_{j=2}^n H_\mu(T^{-t_j}\xi|\bigvee_{i=1}^{j-1}T^{-t_i}\xi\bigvee\mathcal{A})\right).$$

The *conditional sequence entropy along S given $\mathcal{A}$ in the system* $(X,\mathcal{X},\mu,T)$ is

$$h_\mu^S(T|\mathcal{A}) = \sup_\xi\left\{h_\mu^S(T,\xi|\mathcal{A}) : H_\mu(\xi|\mathcal{A})<\infty\right\}.$$

When $S=\mathbb{Z}_+$ and $\mathcal{A}$ is trivial one recovers the entropy of T with respect to μ.

Let $(Y,\mathcal{Y},\nu,T)$ be a factor of $(X,\mathcal{X},\mu,T)$ and $\{\mu_y\}_{y\in Y}$ be the disintegration of μ over ν. Then, the conditional (Shannon) entropy of ξ given $\mathcal{Y}$ can be represented as

$$H_\mu(\xi|\mathcal{Y}) = \int_Y H_y(\xi)d\nu, \tag{2.5}$$

where $H_y(\cdot)$ denotes the entropy with respect to μ_y and $\mathcal{Y}$ is seen as a sub-σ-algebra of $\mathcal{X}$. The following two lemmas come from [Hu2].

Lemma 2.4. *Let ξ and η be measurable partitions of X with $H_\mu(\xi|\mathcal{Y}) < \infty, H_\mu(\eta|\mathcal{Y}) < \infty$. Then*

$$\left|h_\mu^S(T,\xi|\mathcal{Y}) - h_\mu^S(T,\eta|\mathcal{Y})\right| \leq \int_Y \Big(H_y(\xi|\eta) + H_y(\eta|\xi)\Big) d\nu$$

for any increasing sequence $S \subseteq \mathbb{Z}_+$.

Lemma 2.5. *There exists a countable set $\{\xi_n\}_{n\in\mathbb{N}}$ of finite measurable partitions of X such that*

$$\inf_n \left\{ \int_Y \left(H_y(\xi|\xi_n) + H_y(\xi_n|\xi)\right) d\nu \right\} = 0$$

for all measurable partitions ξ with $H_\mu(\xi|\mathcal{Y}) < \infty$.

Hence one has

$$h_\mu^S(T|\mathcal{Y}) = \sup_\xi \Big\{ h_\mu^S(T,\xi|\mathcal{Y}) : \xi \text{ is finite} \Big\}.$$

Thus $h_\mu^S(T|\mathcal{Y}) = 0$ if and only if $h_\mu^S(T,\xi|\mathcal{Y}) = 0$ for all two-set measurable partitions ξ, since every finite partition is a refinement of two-set partitions.
For more information about entropy theory refer to [G, P, W1].

3. Sequence entropy relative to the compact extension

Throughout this section we fix a system $(X,\mathcal{X},\mu,T)$ and a factor $(Y,\mathcal{Y},\nu,T)$.

3.1. **Almost periodic functions.** The $L^2(X,\mathcal{X},\mu)$ norm is denoted by $||\cdot||$ and the $L^2(X,\mathcal{X},\mu_y)$ norm by $||\cdot||_y$ for ν-almost every $y \in Y$. The corresponding inner products are denoted by $\langle\cdot,\cdot\rangle$ and $\langle\cdot,\cdot\rangle_y$. Recall $\{\mu_y\}_{y\in Y}$ is the disintegration of μ over ν.

Definition 3.1. A function $f \in L^2(X,\mathcal{X},\mu)$ is *almost periodic over* $\mathcal{Y}$ if for every $\varepsilon > 0$ there exist $g_1,\ldots,g_l \in L^2(X,\mathcal{X},\mu)$ such that for all $n \in \mathbb{Z}$

$$\min_{1\leq j\leq l} ||T^n f - g_j||_y < \varepsilon$$

for ν almost every $y \in Y$. One writes $f \in AP(\mathcal{Y})$.

Remark 3.2.

(1) The almost periodic functions over $\mathcal{Y}$ form a subspace of $L^2(X,\mathcal{X},\mu)$. Using Proposition 2.1 one can verify that there exists a sub-$\sigma-$algebra $\mathcal{K}(X|Y)$ of $\mathcal{X}$ such that
$$\overline{AP(\mathcal{Y})} = L^2(X,\mathcal{K}(X|Y),\mu).$$
(2) One calls $\mathcal{K}(X|Y)$ *the Kronecker algebra over* $\mathcal{Y}$. Any function $f \in \overline{AP(\mathcal{Y})}$ is called a *compact function over* $\mathcal{Y}$.

The following theorem will be used later.

Theorem 3.3. [Hu2] *Let $f \in L^2(X, \mathcal{X}, \mu)$. Then $f \in L^2(X, \mathcal{K}(X|Y), \mu)^\perp$ if and only if*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \Big| \mathbb{E}(gT^i f|\mathcal{Y}) \Big| = 0$$

in $L^1(Y, \mathcal{Y}, \nu)$ for all $g \in L^2(X, \mathcal{X}, \mu)$.

3.2. **Conditional sequence entropy and $\mathcal{K}(X|Y)$.** In this section we will prove the following result.

Theorem 3.4. *Let $(X, \mathcal{X}, \mu, T)$ be a system and $(Y, \mathcal{Y}, \nu, T)$ be its factor. Then for every increasing sequence $S \in \mathcal{D}^*$*

$$\max_{A\subseteq S} h_\mu^A(T, \xi|\mathcal{Y}) = H_\mu(\xi|\mathcal{K}(X|Y)) \tag{3.1}$$

for all measurable partitions ξ of X with $H_\mu(\xi|\mathcal{Y}) < \infty$.

The following result is now immediate.

Corollary 3.5. [HMY] *Let $(X, \mathcal{X}, \mu, T)$ be a system. Then for every increasing sequence $S \in \mathcal{D}^*$*

$$\max_{A\subseteq S} h_\mu^A(T, \xi) = H_\mu(\xi|\mathcal{K}(X)) \tag{3.2}$$

for all measurable partitions ξ of X with $H_\mu(\xi) < \infty$.

Observe that our result is more general than the one in [HMY], where it is only proved that $\max_{A\subseteq\mathbb{Z}_+} h_\mu^A(T, \xi) = H_\mu(\xi|\mathcal{K}(X))$.
The proof of Theorem 3.4 follows directly from the following series of lemmas.

Lemma 3.6. *For any increasing sequence $S \in \mathcal{D}^*$ there exists a subsequence $A \subseteq S$ such that*

$$h_\mu^A(T, \xi|\mathcal{Y}) \geq H_\mu(\xi|\mathcal{K}(X|Y)) \tag{3.3}$$

for any measurable partition ξ of X with $H_\mu(\xi|\mathcal{Y}) < \infty$.

Proof. To simplify the notation $\mathcal{K}(X|Y)$ is denoted by $\mathcal{K}$. First, we prove the following claim.
Claim: *Given finite measurable partitions ξ and η of X and $\epsilon > 0$, there exist a sequence $D \in \mathcal{D}$ and $M \in \mathbb{N}$ such that for any $m \geq M$ in D one has,*

$$\int_Y H_y(T^{-m}\xi|\eta) d\nu \geq H_\mu(\xi|\mathcal{K}) - \epsilon. \tag{3.4}$$

Proof of the claim. Let $\xi = \{A_1, \ldots, A_k\}$ and $\eta = \{B_1, \ldots, B_l\}$. For any $A, B \in \mathcal{X}$, since $1_A - \mathbb{E}(1_A|\mathcal{K}) \in L^2(X, \mathcal{K}, \mu)^\perp$, by Theorem 3.3 one has that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n} \Big| \int_X T^i \Big(1_A - \mathbb{E}(1_A|\mathcal{K})\Big) 1_B d\mu_y \Big| = 0$$

in $L^1(Y, \mathcal{Y}, \nu)$. Equivalently, $\exists D' = D'(A,B) \in \mathcal{D}$ such that

$$\lim_{D' \ni n \to \infty} \int_Y \Big| \int_X T^n \Big(1_A - \mathbb{E}(1_A|\mathcal{K})\Big) 1_B d\mu_y \Big| d\nu = 0.$$

Since $\mathcal{D}$ is a filter, there exists $D \in \mathcal{D}$ such that for any $1 \le i \le k, 1 \le j \le l$

$$\lim_{D \ni n \to \infty} \int_Y \Big| \mu_y(T^{-n}A_i \cap B_j) - \int_X T^n \mathbb{E}(1_{A_i}|\mathcal{K}) 1_{B_j} d\mu_y \Big| d\nu = 0. \tag{3.5}$$

Let $\varphi(x) = -x \log x$. Choose $0 < \delta < \frac{\epsilon}{4} \frac{1}{kl \log l}$ such that

$$|u - v| < \delta \Rightarrow |\varphi(u) - \varphi(v)| < \frac{\epsilon}{4} \frac{1}{kl}.$$

By (3.5) there exists $M > 0$ such that for every $m > M$ in D: there is $E_m \subseteq Y$ with $\nu(E_m) > 1 - \frac{\epsilon}{2 \log k}$ such that

$$\Big| \mu_y(T^{-m}A_i \cap B_j) - \int_X T^m \mathbb{E}(1_{A_i}|\mathcal{K}) 1_{B_j} d\mu_y \Big| < \delta$$

for all $1 \le i \le k, 1 \le j \le l$ and $y \in E_m$.
For any $y \in E_m$

$$\begin{aligned}
& \Big| \sum_{i,j} -\mu_y(T^{-m}A_i \cap B_j) \log \frac{\mu_y(T^{-m}A_i \cap B_j)}{\mu_y(B_j)} \\
& + \left(\int_X T^m \mathbb{E}(1_{A_i}|\mathcal{K}) 1_{B_j} d\mu_y \right) \log \frac{\int_X T^m \mathbb{E}(1_{A_i}|\mathcal{K}) 1_{B_j} d\mu_y}{\mu_y(B_j)} \Big| \\
\le \; & \sum_{i,j} \Big| \mu_y(T^{-m}A_i \cap B_j) \log \mu_y(T^{-m}A_i \cap B_j) \\
& - \int_X T^m \mathbb{E}(1_{A_i}|\mathcal{K}) 1_{B_j} d\mu_y \log \int_X T^m \mathbb{E}(1_{A_i}|\mathcal{K}) 1_{B_j} d\mu_y \Big| \\
& + \sum_i \Big| \sum_j \left(\mu_y(T^{-m}A_i \cap B_j) - \int_X T^m \mathbb{E}(1_{A_i}|\mathcal{K}) 1_{B_j} d\mu_y \right) \log \mu_y(B_j) \Big| \\
\le \; & kl \frac{\epsilon}{4} \frac{1}{kl} + k\delta \Big| \log(\prod_j \mu_y(B_j)) \Big| \le \frac{\epsilon}{4} + k \frac{\epsilon}{4} \frac{1}{kl \log l} \Big| \log \left(\frac{\sum_j \mu_y(B_j)}{l} \right)^l \Big| = \frac{\epsilon}{2}
\end{aligned}$$

Hence

$$\begin{aligned}
H_y(T^{-m}\xi|\eta) &= \sum_{i,j} -\mu_y(T^{-m}A_i \cap B_j) \log \frac{\mu_y(T^{-m}A_j \cap B_j)}{\mu_y(B_j)} \\
&\ge \sum_{i,j} \left(- \int_X T^m \mathbb{E}(1_{A_i}|\mathcal{K}) 1_{B_j} d\mu_y \right) \log \frac{\int_X T^m \mathbb{E}(1_{A_i}|\mathcal{K}) 1_{B_j} d\mu_y}{\mu_y(B_j)} - \frac{\epsilon}{2}
\end{aligned}$$

Let

$$a_{ij} = - \left(\int_X T^m \mathbb{E}(1_{A_i}|\mathcal{K}) 1_{B_j} d\mu_y \right) \log \frac{\int_X T^m \mathbb{E}(1_{A_i}|\mathcal{K}) 1_{B_j} d\mu_y}{\mu_y(B_j)}.$$

Then $a_{ij} = \mu_y(B_j)\varphi(\int_{B_j} T^m\mathbb{E}(1_{A_i}|\mathcal{K})d\mu_{B_j,y})$, where $\mu_{B_j,y} = \dfrac{\mu_y(\cdot \cap B_j)}{\mu_y(B_j)}$. Since φ is concave, one deduces

$$\begin{aligned} a_{ij} &\geq \mu_y(B_j)\int\limits_{B_j} -T^m\mathbb{E}(1_{A_i}|\mathcal{K})\log T^m\mathbb{E}(1_{A_i}|\mathcal{K})d\mu_{B_j,y} \\ &= \int\limits_{B_j} -T^m\mathbb{E}(1_{A_i}|\mathcal{K})\log T^m\mathbb{E}(1_{A_i}|\mathcal{K})d\mu_y. \end{aligned}$$

Thus

$$H_y(T^{-m}\xi|\eta) \geq \sum_{i,j} a_{ij} - \frac{\epsilon}{2} \geq \sum_i \int_X -T^m\mathbb{E}(1_{A_i}|\mathcal{K})\log T^m\mathbb{E}(1_{A_i}|\mathcal{K})d\mu_y - \frac{\epsilon}{2}.$$

Integrating with respect to ν one obtains

$$\begin{aligned} \int\limits_Y H_y(T^{-m}\xi|\eta)d\nu &\geq \int\limits_{E_m} H_y(T^{-m}\xi|\eta)d\nu \\ &\geq \sum_i \int\limits_{E_m}\int_X -T^m\mathbb{E}(1_{A_i}|\mathcal{K})\log T^m\mathbb{E}(1_{A_i}|\mathcal{K})d\mu_y d\nu - \frac{\epsilon}{2} \\ &= \sum_i \int\limits_{Y}\int_X -T^m\mathbb{E}(1_{A_i}|\mathcal{K})\log T^m\mathbb{E}(1_{A_i}|\mathcal{K})d\mu_y d\nu \\ &\quad -\sum_i \int\limits_{Y\setminus E_m}\int_X -T^m\mathbb{E}(1_{A_i}|\mathcal{K})\log T^m\mathbb{E}(1_{A_i}|\mathcal{K})d\mu_y d\nu - \frac{\epsilon}{2} \\ &\geq \sum_i \int\limits_X -T^m\mathbb{E}(1_{A_i}|\mathcal{K})\log T^m\mathbb{E}(1_{A_i}|\mathcal{K})d\mu - \frac{\epsilon}{2\log k}\log k - \frac{\epsilon}{2} \\ &= H_\mu(\xi|\mathcal{K}) - \epsilon\ . \end{aligned}$$

In the last inequality we use $\int_X \sum_i -T^m\mathbb{E}(1_{A_i}|\mathcal{K})\log T^m\mathbb{E}(1_{A_i}|\mathcal{K})d\mu_y = H_y(T^{-m}\xi|\mathcal{K})$ $\leq \log k$. This completes the proof of the claim.

Let $\{\xi_k\}_{k\in\mathbb{N}}$ be as in Lemma 2.5. For any increasing sequence $S \in \mathcal{D}^*$, since $\mathcal{D}$ is a filter, one can choose $A = \{t_1 < t_2 < \ldots\} \subseteq S$ such that

$$\int_Y H_y(T^{-t_n}\xi_j|\bigvee_{i=1}^{n-1} T^{-t_i}\xi_j)d\nu \geq H_\mu(\xi_j|\mathcal{K}) - \frac{1}{2^n},\ \text{for any } n \geq 2 \text{ and } 1 \leq j \leq n.$$

Fix $k \in \mathbb{N}$. One has

$$\begin{aligned}
h_\mu^A(T,\xi_k|\mathcal{Y}) &= \limsup_{n\to\infty}\frac{1}{n}H_\mu(\bigvee_{i=1}^n T^{-t_i}\xi_k|\mathcal{Y}) = \limsup_{n\to\infty}\frac{1}{n}\int_Y H_y(\bigvee_{i=1}^n T^{-t_i}\xi_k)d\nu \\
&= \limsup_{n\to\infty}\frac{1}{n}\int_Y \Big[H_y(\bigvee_{i=1}^k T^{-t_i}\xi_k) + \sum_{i=k+1}^n H_y(T^{-t_i}\xi_k|\bigvee_{j=1}^{i-1}T^{-t_j}\xi_k)\Big]d\nu \\
&\geq \limsup_{n\to\infty}\frac{1}{n}\left[\int_Y H_y(\bigvee_{i=1}^k T^{-t_i}\xi_k)d\nu + (n-k)H_\mu(\xi_k|\mathcal{K}) - \sum_{i=k+1}^n \frac{1}{2^i}\right] \\
&= H_\mu(\xi_k|\mathcal{K})
\end{aligned}$$

Therefore $h_\mu^A(T,\xi_k|\mathcal{Y}) \geq H_\mu(\xi_k|\mathcal{K})$ for any $k \in \mathbb{N}$. Now let ξ be any partition with $H(\xi|\mathcal{Y}) < \infty$. Given $\delta > 0$, by Lemma 2.5 one can choose ξ_k such that $\int_Y H_y(\xi|\xi_k) + H_y(\xi_k|\xi)d\nu < \delta$. Then, $|h_\mu^A(T,\xi|\mathcal{Y}) - h_\mu^A(T,\xi_k|\mathcal{Y})| < \delta$ and $|H_\mu(\xi|\mathcal{Y}) - H_\mu(\xi_k|\mathcal{Y})| < \delta$. So

$$h_\mu^A(T,\xi|\mathcal{Y}) \geq h_\mu^A(T,\xi_k|\mathcal{Y}) - \delta \geq H_\mu(\xi|\mathcal{Y}) - 2\delta.$$

Since δ is arbitrary, the proof is complete. □

Lemma 3.7. *Let $B \in \mathcal{X}$. Then $B \in \mathcal{K}(X|Y)$ if and only if $h_\mu^A(T,\{B,B^c\}|\mathcal{Y}) = 0$ for any increasing sequence $A \subseteq \mathbb{Z}_+$.*

Proof. For necessity we refer to [Hu2, Theorem 1]. To prove sufficiency we will use Lemma 3.6. If $B \notin \mathcal{K}(X|Y)$ then $H_\mu(\{B,B^c\}|\mathcal{K}(X|Y)) > 0$. Thus, by Lemma 3.6, there exists $A \subseteq \mathbb{Z}_+$ such that

$$h_\mu^A(T,\{B,B^c\}|\mathcal{Y}) \geq H_\mu(\{B,B^c\}|\mathcal{K}(X|Y)) > 0\ .$$

This completes the proof of the lemma. □

Lemma 3.8. *For any measurable partition ξ of X with $H_\mu(\xi|\mathcal{Y}) < \infty$ and any increasing sequence $A \subseteq \mathbb{Z}_+$,*

$$h_\mu^A(T,\xi|\mathcal{Y}) \leq H_\mu(\xi|\mathcal{K}(X|Y)). \tag{3.6}$$

Proof. Let $\{\xi_k\}_{k\in\mathbb{N}}$ be a countable set of finite $\mathcal{K}(X|Y)$-measurable partitions such that $\xi_k \nearrow \mathcal{K}(X|Y)$. Let $A = \{t_1 < t_2 < \ldots\}$. Since ξ_k is $\mathcal{K}(X|Y)$-measurable, by

Lemma 3.7 one has that $h_\mu^A(T,\xi_k|\mathcal{Y})=0$. So

$$\begin{aligned}
h_\mu^A(T,\xi|\mathcal{Y}) &= \limsup_{n\to\infty}\frac{1}{n}\int_Y H_y(\bigvee_{i=1}^n T^{-t_i}\xi)d\nu - h_\mu^A(T,\xi_k|\mathcal{Y})\\
&\le \limsup_{n\to\infty}\frac{1}{n}\int_Y H_y(\bigvee_{i=1}^n T^{-t_i}(\xi\vee\xi_k))d\nu - \lim_{n\to\infty}\frac{1}{n}\int_Y H_y(\bigvee_{i=1}^n T^{-t_i}\xi_k)d\nu\\
&= \limsup_{n\to\infty}\frac{1}{n}\int_Y \left(H_y(\bigvee_{i=1}^n T^{-t_i}(\xi\vee\xi_k)) - H_y(\bigvee_{i=1}^n T^{-t_i}\xi_k)\right)d\nu\\
&= \limsup_{n\to\infty}\frac{1}{n}\int_Y H_y(\bigvee_{i=1}^n T^{-t_i}\xi|\bigvee_{i=1}^n T^{-t_i}\xi_k)d\nu\\
&\le \limsup_{n\to\infty}\frac{1}{n}\int_Y \sum_{i=1}^n H_y(T^{-t_i}\xi|T^{-t_i}\xi_k)d\nu\\
&= \limsup_{n\to\infty}\frac{1}{n}\int_Y \sum_{i=1}^n H_y(\xi|\xi_k)d\nu\\
&= \int_Y H_y(\xi|\xi_k)d\nu = \int_Y \big(H_y(\xi\vee\xi_k)-H_y(\xi_k)\big)d\nu\\
&= H_\mu(\xi\vee\xi_k|\mathcal{Y}) - H_\mu(\xi_k|\mathcal{Y}) = H_\mu(\xi|\xi_k\vee\mathcal{Y}) \le H_\mu(\xi|\xi_k)\ .
\end{aligned}$$

One concludes by martingale's theorem. □

3.3. **Compact and weakly mixing extensions.** Now as a corollary of the last subsection we recover some results of Hulse in [Hu1, Hu2]. First let us recall some notations.

Definition 3.9.

(1) $(X,\mathcal{X},\mu,T)$ is called a *compact extension* of $(Y,\mathcal{Y},\nu,T)$ if $L^2(X,\mathcal{K}(X|Y),\mu) = L^2(X,\mathcal{X},\mu)$, that is $\mathcal{K}(X|Y)=\mathcal{X}$;
(2) $(X,\mathcal{X},\mu,T)$ is called a *weakly mixing extension* of $(Y,\mathcal{Y},\nu,T)$ if $\mathcal{K}(X|Y)=\mathcal{Y}$.

Remark 3.10. For a more complete discussion on compact and weakly mixing extensions see [F1, Hu2]. For example, in [Hu2] it is proved that $(X,\mathcal{X},\mu,T)$ is a weakly mixing extension of $(Y,\mathcal{Y},\nu,T)$ if and only if $(X\times X,\mathcal{X}\otimes\mathcal{X},\mu\times_Y\mu,T\times T)$ is ergodic relative to $(Y,\mathcal{Y},\nu,T)$ if and only if $\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}|\mathbb{E}(gT^if|\mathcal{Y})|=0$ in $L^1(Y,\mathcal{Y},\nu)$ for all $f\in L^2(X,\mathcal{X},\mu)$ and $g\in L^2(X,\mathcal{Y},\mu)^\perp$ (here we see $\mathcal{Y}$ as a sub-σ-algebra of $\mathcal{X}$).

By the results of the last subsection it follows immediately that,

Corollary 3.11. [Hu2]

(1) *$(X,\mathcal{X},\mu,T)$ is a compact extension of $(Y,\mathcal{Y},\nu,T)$ if and only if $h_\mu^A(T|\mathcal{Y})=0$ for any increasing sequence $A\subseteq\mathbb{Z}_+$;*

(2) $(X,\mathcal{X},\mu,T)$ *is a weakly mixing extension of* $(Y,\mathcal{Y},\nu,T)$ *if and only if for any increasing sequence* $S\in\mathcal{D}^*$*, there exists an increasing subsequence* $A\subseteq S$ *such that*
$$h^A_\mu(T,\xi|\mathcal{Y})=H_\mu(\xi|\mathcal{Y})$$
for all measurable partitions ξ *of* X *with* $H_\mu(\xi|\mathcal{Y})<\infty$.

Observe that the second statement is a little stronger than the corresponding result in [Hu2]. The case when $\mathcal{Y}$ is trivial can be found in [Hu1].

4. Sequence entropy relative to the rigid extension

4.1. **$\mathcal{F}$-sequence and IP-systems.** Let $\mathcal{F}$ denote the collection of all non-empty finite subsets of $\mathbb{N}$. Given $\alpha,\beta\in\mathcal{F}$, $\alpha<\beta$ (or $\beta>\alpha$) if $\max\alpha<\min\beta$. The set
$$\mathcal{F}^{(1)}=\mathrm{FU}(\{\alpha_i\}_{i\in\mathbb{N}}):=\Big\{\bigcup_{i\in\beta}\alpha_i:\beta\in\mathcal{F}\Big\}$$
is called an IP-*ring*, where $\alpha_1<\alpha_2<\ldots<\alpha_n<\ldots$. The following theorem will be useful.

Theorem 4.1. *(Hindman's Theorem) For any finite partition* $\{C_1,\ldots,C_r\}$ *of* $\mathcal{F}$ *one of the* C_i*'s contains an* IP*-ring.*

Let $\{n_i\}_{i\in\mathbb{N}}\subseteq\mathbb{Z}_+$. Denoting $n_\alpha=\sum_{i\in\alpha}n_i$ for $\alpha\in\mathcal{F}$, then $\mathrm{FS}(\{n_i\}_{i\in\mathbb{N}})=\{n_\alpha\}_{\alpha\in\mathcal{F}}$ and the IP *set* generated by this sequence is $\mathrm{FS}(\{n_i\}_{i\in\mathbb{N}})$. Observe that we do not require the elements of $\{n_i\}_{i\in\mathbb{N}}$ to be distinct. If $\mathcal{F}^{(1)}$ is an IP-ring of $\mathcal{F}$, then $\{n_\alpha\}_{\alpha\in\mathcal{F}^{(1)}}$ is an IP subset of $\{n_\alpha\}_{\alpha\in\mathcal{F}}$; and conversely any IP subset of $\{n_\alpha\}_{\alpha\in\mathcal{F}}$ has this form.
A sequence in any space Y indexed by the set $\mathcal{F}$ is called an $\mathcal{F}$- *sequence*. If Y is a (multiplicative) semigroup, then an $\mathcal{F}$-sequence $\{y_\alpha\}_{\alpha\in\mathcal{F}}$ on Y defines an IP-*system* if for any $\alpha=\{i_1,\ldots,i_k\}\in\mathcal{F}$ with $i_1<\ldots<i_k$ one has $y_\alpha=y_{i_1}\cdots y_{i_k}$. An IP-system should be viewed as a generalized semigroup. Indeed, if $\alpha\cap\beta=\emptyset$ and $\alpha<\beta$ then $y_{\alpha\cup\beta}=y_\alpha y_\beta$.
Let $\mathcal{F}^{(1)}$ be an IP-ring. Then the map $\xi:\mathcal{F}\to\mathcal{F}^{(1)},\xi(\alpha)=\bigcup_{i\in\alpha}\alpha_i$ is bijective and structure preserving in the sense that $\xi(\alpha\cup\beta)=\xi(\alpha)\cup\xi(\beta)$. In particular, any sequence $\{y_\alpha\}_{\alpha\in\mathcal{F}^{(1)}}$ can be naturally identified with a particular $\mathcal{F}$-sequence, namely the $\mathcal{F}$-sequence $\{x_\alpha\}_{\alpha\in\mathcal{F}}$ where $x_\alpha=y_{\xi(\alpha)}$.

Definition 4.2. Assume that $\{x_\alpha\}_{\alpha\in\mathcal{F}}$ is an $\mathcal{F}$-sequence in a topological space X. Let $x\in X$ and $\mathcal{F}^{(1)}$ be an IP-ring. Write
$$\mathrm{IP}-\lim_{\alpha\in\mathcal{F}^{(1)}}x_\alpha=x$$
if for any neighborhood U of x there exists $\alpha_0\in\mathcal{F}^{(1)}$ such that for any $\alpha\in\mathcal{F}^{(1)}$ with $\alpha>\alpha_0$ one has $x_\alpha\in U$.

Theorem 4.3. [FK] *Let* $\{U_\alpha\}_{\alpha\in\mathcal{F}}$ *be an* IP*-system of unitary operators on a separable Hilbert space* $\mathcal{H}$*. Then there is an* IP*-subsystem* $\{U_\alpha\}_{\alpha\in\mathcal{F}^{(1)}}$*, with* $\mathcal{F}^{(1)}$ *an* IP*-ring, such that*
$$\mathrm{IP-}\lim_{\alpha\in\mathcal{F}^{(1)}}U_\alpha=P$$

weakly, where P is an orthogonal projection onto a subspace of $\mathcal{H}$.

4.2. **Almost periodicity over $\mathcal{Y}$ along $\mathcal{F}^{(1)}$.** Let $(X, \mathcal{X}, \mu, T)$ be a system and $\{n_\alpha\}_{\alpha\in\mathcal{F}}$ be an IP set. Then $\{T^{n_\alpha}\}_{\alpha\in\mathcal{F}}$ and $\{(T\times T)^{n_\alpha}\}_{\alpha\in\mathcal{F}}$ define IP-systems. One writes $T_\alpha = T^{n_\alpha}$ and $(T\times T)_\alpha = (T\times T)^{n_\alpha}$.
Now consider a factor $(Y, \mathcal{Y}, \nu, T)$ of $(X, \mathcal{X}, \mu, T)$ and $\pi : X \to Y$ the corresponding factor map. By Theorem 4.3, there exists an IP-ring $\mathcal{F}^{(1)} \subseteq \mathcal{F}$ such that for all $K \in L^2(X\times X, \mathcal{X}\otimes\mathcal{X}, \mu\times_Y \mu)$

$$\text{IP} - \lim_{\alpha\in\mathcal{F}^{(1)}} (T\times T)_\alpha K = PK \tag{4.1}$$

exists in the weak topology and P is an orthogonal projection.

Definition 4.4. Let $\{n_\alpha\}_{\alpha\in\mathcal{F}}$ be an IP set and $\mathcal{F}^{(1)} \subseteq \mathcal{F}$ an IP-ring. A function $f \in L^2(X, \mathcal{X}, \mu)$ is called *almost periodic over $\mathcal{Y}$ with respect to $\{n_\alpha\}_{\alpha\in\mathcal{F}}$ along $\mathcal{F}^{(1)}$* and one writes $f \in AP(\mathcal{Y}, \{n_\alpha\}, \mathcal{F}^{(1)})$, if for every $\varepsilon > 0$ there exists a set $D \in \mathcal{Y}$ with $\nu(D) < \varepsilon$ and functions $g_1, \dots, g_l \in L^2(X, \mathcal{X}, \mu)$, such that for every $\delta > 0$ there exists $\alpha_0 \in \mathcal{F}^{(1)}$ with the property that whenever $\alpha \in \mathcal{F}^{(1)}$ with $\alpha > \alpha_0$ there is a set $E_\alpha \in \mathcal{Y}$ with $\nu(E_\alpha) < \delta$ verifying for all $y \notin D\cup E_\alpha$ that

$$\min_{1\le j\le l} ||T_\alpha f - g_j||_y < \varepsilon.$$

Remark 4.5.

(1) The definition we use comes from [BM], which is different from [FK] and [Z2]. In [FK] a function is called almost periodic over $\mathcal{Y}$ with respect to $\{n_\alpha\}_{\alpha\in\mathcal{F}}$ along $\mathcal{F}^{(1)}$ if for every $\varepsilon > 0$ and $\alpha_0 \in \mathcal{F}^{(1)}$ there exist $g_1, \dots, g_l \in L^2(X, \mathcal{X}, \mu)$ and a set $E \subseteq \mathcal{Y}$ with $\nu(E) < \varepsilon$, such that for all $\alpha \in \mathcal{F}^{(1)}$ with $\alpha > \alpha_0$ and $y \notin E$ one has $\min_{1\le j\le l} ||T_\alpha f - g_j||_y < \varepsilon$. Refer to [BM] for the discussion of the difference between the two definitions.

(2) $AP(\mathcal{Y}, \{n_\alpha\}, \mathcal{F}^{(1)})$ need not to be closed, however, one has the following: if $f \in \overline{AP(\mathcal{Y}, \{n_\alpha\}, \mathcal{F}^{(1)})}$, then for every $\varepsilon > 0$ there exist $g_1, \dots, g_l \in L^\infty(X, \mathcal{X}, \mu)$ and $\alpha_0 \in \mathcal{F}^{(1)}$ such that for any $\alpha \in \mathcal{F}^{(1)}$ with $\alpha > \alpha_0$, there is a set $E_\alpha \in \mathcal{Y}$ with $\nu(E_\alpha) < \varepsilon$ and $\min_{1\le j\le l} ||T_\alpha f - g_j||_y < \varepsilon$ for all $y \notin E_\alpha$.

(3) $AP(\mathcal{Y}, \{n_\alpha\}, \mathcal{F}^{(1)}) \cap L^\infty(X, \mathcal{X}, \mu)$ is a $T-$invariant σ-algebra which contains $|g|$ and $\bar{g}$ whenever it contains g. Using Proposition 2.1, there exists a sub-σ-algebra $\mathcal{K}_F(X|Y)$, where $F = \{n_\alpha\}_{\alpha\in\mathcal{F}^{(1)}}$, such that (see also [FK, Lemma 7.3])

$$\overline{AP(\mathcal{Y}, \{n_\alpha\}, \mathcal{F}^{(1)})} = L^2(X, \mathcal{K}_F(X|Y), \mu). \tag{4.2}$$

(4) One calls $\mathcal{K}_F(X|Y)$ a *rigid algebra over $\mathcal{Y}$*, and any function $f \in \overline{AP(\mathcal{Y}, \{n_\alpha\}, \mathcal{F}^{(1)})}$ is called a *rigid function over $\mathcal{Y}$*.

To each $K \in L^2(X \times X, \mathcal{X} \otimes \mathcal{X}, \mu \times_Y \mu)$ and ν-almost every $y \in Y$ associate an operator (also called K) $K : L^2(X, \mathcal{X}, \mu_y) \to L^2(X, \mathcal{X}, \mu_y), f \mapsto K * f$ where

$$K * f(x) = \int_X K(x, x')f(x')d\mu_y(x'), \quad \text{where } y = \pi(x). \tag{4.3}$$

For ν-a.e. $y \in Y$, K is a Hilbert-Schmidt operator on $L^2(X, \mathcal{X}, \mu_y)$. In particular, it is a compact operator (i.e. the closure of the image of the unit ball is compact). See [F1, FK] for details.

For the next four results of the subsection fix $(X, \mathcal{X}, \mu, T)$ a system, $(Y, \mathcal{Y}, \nu, T)$ a factor of it, $\{n_\alpha\}_{\alpha \in \mathcal{F}}$ an IP set and $\mathcal{F}^{(1)}$ an IP-ring such that (4.1) holds. Let $F = \{n_\alpha\}_{\alpha \in \mathcal{F}^{(1)}}$ and $\mathcal{K}_F(X|Y)$ be the associated rigid algebra over $\mathcal{Y}$.

Lemma 4.6. *If $K \in L^2(X \times X, \mathcal{X} \otimes \mathcal{X}, \mu \times_Y \mu)$ with $PK = K$ and $f \in L^\infty(X, \mathcal{X}, \mu)$, then $K * f \in AP(\mathcal{Y}, \{n_\alpha\}, \mathcal{F}^{(1)})$.*

Proof. Let $\varepsilon > 0$. Since for ν-a.e. $y \in Y$ the operator K is compact on $L^2(X, \mathcal{X}, \mu_y)$, then there exists a number $M(y) \in \mathbb{N}$ such that $\{K * (T^j f) : -M(y) \le j \le M(y)\}$ is $\varepsilon/2$-dense in $\{K * (T^j f) : j \in \mathbb{Z}\}$ (in $L^2(X, \mathcal{X}, \mu_y)$). Let M be large enough such that $M > M(y)$ for all y outside of a set $D \in \mathcal{Y}$ with $\nu(D) < \varepsilon$ and let

$$\{g_1, \ldots, g_l\} = \{K * (T^j f) : -M \le j \le M\}.$$

Then for any $y \in D^c$ and any $n \in \mathbb{Z}$,

$$\inf_{1 \le j \le l} ||K * (T^n f) - g_j||_y < \varepsilon/2. \tag{4.4}$$

On the other hand, by (4.1) one has

$$\begin{aligned}
& \mathrm{IP} - \lim_{\alpha \in \mathcal{F}^{(1)}} ||T_\alpha(K * f) - K * (T_\alpha f)||^2 \\
= \; & \mathrm{IP} - \lim_{\alpha \in \mathcal{F}^{(1)}} \int_X \Big| \int_X \Big(K(T_\alpha x, T_\alpha x') - K(x, x')\Big) f(T_\alpha x') d\mu_{\pi(x)}(x')\Big|^2 d\mu(x) \\
\le \; & \mathrm{IP} - \lim_{\alpha \in \mathcal{F}^{(1)}} \int_X \int_X \Big|K(T_\alpha x, T_\alpha x') - K(x, x')\Big|^2 |f(T_\alpha x')|^2 d\mu_{\pi(x)}(x') d\mu(x) \\
\le \; & \mathrm{IP} - \lim_{\alpha \in \mathcal{F}^{(1)}} ||(T \times T)_\alpha K - K||^2 ||f||_\infty^2 = 0
\end{aligned}$$

So for any $\delta > 0$ there exists $\alpha_0 \in \mathcal{F}^{(1)}$ such that for any $\alpha \in \mathcal{F}^{(1)}$ with $\alpha > \alpha_0$ there is a set $E_\alpha \in \mathcal{Y}$ with $\nu(E_\alpha) < \delta$ and for any $y \notin E_\alpha$

$$||T_\alpha(K * f) - K * (T_\alpha f)||_y < \varepsilon/2. \tag{4.5}$$

Hence whenever $y \notin D \cup E_\alpha$, one has

$$\inf_{1 \le j \le l} ||T_\alpha(K * f) - g_j||_y < \varepsilon. \tag{4.6}$$

This completes the proof. $\square$

Theorem 4.7. *Let $(X, \mathcal{X}, \mu, T)$, $(Y, \mathcal{Y}, \nu, T)$, $\{n_\alpha\}_{\alpha \in \mathcal{F}}$ and $\mathcal{F}^{(1)}$ as fixed above. Then $f \in L^2(X, \mathcal{K}_F(X|Y), \mu)^\perp$ if and only if*

$$\mathrm{IP} - \lim_{\alpha \in \mathcal{F}^{(1)}} \int_Y \Big| \mathbb{E}(gT_\alpha f|\mathcal{Y}) \Big| d\nu = 0$$

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

Proof. Assume that $f \in L^2(X, \mathcal{K}_F(X|Y), \mu)^\perp$. Then

$$\begin{aligned}
& \mathrm{IP} - \lim_{\alpha \in \mathcal{F}^{(1)}} \Big(\int_Y |\mathbb{E}(gT_\alpha f|\mathcal{Y})| d\nu \Big)^2 \\
\leq\ & \mathrm{IP} - \lim_{\alpha \in \mathcal{F}^{(1)}} \int_Y \Big| \mathbb{E}(gT_\alpha f|\mathcal{Y}) \Big|^2 d\nu \\
=\ & \mathrm{IP} - \lim_{\alpha \in \mathcal{F}^{(1)}} \int_{X \times X} g \otimes \overline{g} \cdot (T \times T)_\alpha (f \otimes \overline{f}) d\mu \times_Y \mu \\
=\ & \int_{X \times X} g \otimes \overline{g} P(f \otimes \overline{f}) d\mu \times_Y \mu = \int_{X \times X} f \otimes \overline{f} P(g \otimes \overline{g}) d\mu \times_Y \mu
\end{aligned}$$

Let $K = P(g \otimes \overline{g})$. Then $PK = K$ and

$$\begin{aligned}
& \int_{X \times X} f \otimes \overline{f} P(g \otimes \overline{g}) d\mu \times_Y \mu \\
=\ & \int_{X \times X} f \otimes \overline{f} K d\mu \times_Y \mu \\
=\ & \int_Y \int_{X \times X} f(x) \overline{f}(x') K(x, x') d\mu_y(x) d\mu_y(x') d\nu(y) \\
=\ & \int_X f(x) \int_X K(x, x') \overline{f}(x') d\mu_{\pi(x)}(x') d\mu(x) \\
=\ & \int_X f(x) K * \overline{f}(x) d\mu(x) = \langle \overline{f}, K * \overline{f} \rangle.
\end{aligned}$$

By Lemma 4.6, $K * \overline{f} \in \overline{AP(\mathcal{Y}, \{n_\alpha\}, \mathcal{F}^{(1)})}$. So $\langle \overline{f}, K * \overline{f} \rangle = 0$. Hence

$$\mathrm{IP} - \lim_{\alpha \in \mathcal{F}^{(1)}} \int \Big| \mathbb{E}(gT_\alpha f|\mathcal{Y}) \Big| d\nu = 0.$$

Now we show the converse. It suffices to show that if $f \in L^2(X, \mathcal{K}_F(X|Y), \mu)$ and $\mathrm{IP} - \lim_{\alpha \in \mathcal{F}^{(1)}} \int_Y |\mathbb{E}(gT_\alpha f|\mathcal{Y})| d\nu = 0$ holds for any $g \in L^2(X, \mathcal{X}, \mu)$, then $f = 0$. The method follows from [BM, Lemma 3.13.] and [FK, Lemma 7.6.].

Let δ with $0 < \delta < 1$ be small enough such that for any $h \in L^2(X, \mathcal{X}, \mu)$ verifying $||f - h||^2 < 2\delta$ one has $|\langle f, h \rangle| > ||f||^2/2$. Let $\varepsilon > 0$ with $\varepsilon^2 < \delta$ and $\int_E ||f||_y^2 d\nu < \delta$ for all $E \in \mathcal{Y}$ with $\nu(E) < \varepsilon$.

By Remark 4.5 part (2), for $f \in L^2(X, \mathcal{K}_F(X|Y), \mu) = \overline{AP(\mathcal{Y}, \{n_\alpha\}, \mathcal{F}^{(1)})}$, there exist $g_1, \ldots, g_l \in L^\infty(X, \mathcal{X}, \mu)$ and $\alpha_0 \in \mathcal{F}^{(1)}$ such that whenever $\alpha \in \mathcal{F}^{(1)}$ with $\alpha > \alpha_0$ there is a set $E_\alpha \in \mathcal{Y}$ with $\nu(E_\alpha) < \varepsilon$ verifying that: for all $y \notin E_\alpha$ there exists $j(\alpha, y)$ with $1 \leq j(\alpha, y) \leq l$ such that

$$||T_\alpha f - g_{j(\alpha,y)}||_y < \varepsilon.$$

For every $\alpha \in \mathcal{F}^{(1)}$ with $\alpha > \alpha_0$ and $i \in \{1, \dots, l\}$, let $\xi_i(y) = 1$ if $y \notin E_\alpha$ and $j(\alpha, y) = i$, and let $\xi_i(y) = 0$ otherwise. Write $h_\alpha = \sum_{i=1}^{l} \xi_i g_i$, that is, h_α is equal to $g_{j(\alpha,y)}$ on the fiber over y when $y \notin E_\alpha$, and equal to zero on fibers over $y \in E_\alpha$. Each h_α is measurable and one has

$$\begin{aligned} ||f - T_\alpha^{-1} h_\alpha||^2 &= ||T_\alpha f - h_\alpha||^2 \\ &= \int_Y \int_X \Big| T_\alpha f - h_\alpha \Big|^2 d\mu_y(x) d\nu \\ &= \int_{E_\alpha} ||T_\alpha f||_y^2 d\nu + \int_{Y \setminus E_\alpha} ||T_\alpha f - g_{j(\alpha,y)}||_y^2 d\nu \\ &\leq \delta + \varepsilon^2 \leq 2\delta. \end{aligned}$$

Hence $|\langle T_\alpha f, h_\alpha \rangle| = |\langle f, T_\alpha^{-1} h_\alpha \rangle| \geq ||f||^2/2$. Also,

$$\begin{aligned} |\langle T_\alpha f, h_\alpha \rangle| &= |\sum_{j=1}^{l} \int_Y \xi_j(y) \int_X T_\alpha f \cdot \overline{g_j} d\mu_y d\nu| \\ &\leq \sum_{j=1}^{l} \int_Y \Big| \int_X T_\alpha f \cdot \overline{g_j} d\mu_y \Big| d\nu \\ &= \sum_{j=1}^{l} \int_Y \Big| \mathbb{E}(T_\alpha f \cdot \overline{g_j} | \mathcal{Y}) \Big| d\nu \ . \end{aligned}$$

Since $\mathrm{IP} - \lim_{\alpha \in \mathcal{F}^{(1)}} \int_Y |\mathbb{E}(g T_\alpha f | \mathcal{Y})| d\nu = 0$ holds for any $g \in L^2(X, \mathcal{X}, \mu)$, it follows that $||f|| = 0$ and thus $f = 0$. The proof of the theorem is complete. □

Proposition 4.8. *A function $f \in L^2(X, \mathcal{K}_F(X|Y), \mu)$ if and only if for any $\varepsilon > 0$ and any* IP*-ring $\mathcal{F}^{(2)} = \mathrm{FU}(\{\alpha_i\}_{i \in \mathbb{N}}) \subseteq \mathcal{F}^{(1)}$, there is $M \in \mathbb{N}$ such that for every $\alpha \in \mathcal{F}^{(2)}$ with $\alpha > \alpha_M$ there exists $E_\alpha \in \mathcal{Y}$ with $\nu(E_\alpha) < \varepsilon$ verifying for any $y \notin E_\alpha$,*

$$\inf_{\beta \in \mathrm{FU}(\{\alpha_i\}_{i=1}^{M})} ||T_\alpha f - T_\beta f||_y < \varepsilon.$$

Proof. Let $f \in L^2(X, \mathcal{K}_F(X|Y), \mu)$. By Remark 4.5-(2), for any $\varepsilon > 0$ there exist $g_1, \dots, g_l \in L^\infty(X, \mathcal{X}, \mu)$ and $\alpha_0 \in \mathcal{F}^{(1)}$ such that for any $\alpha \in \mathcal{F}^{(1)}$ with $\alpha > \alpha_0$, there is $E'_\alpha \in \mathcal{Y}$ with $\nu(E'_\alpha) < \varepsilon/2$ verifying

$$\inf_{1 \leq j \leq l} ||T_\alpha f - g_j||_y < \varepsilon/2 \tag{4.7}$$

for any $y \notin E'_\alpha$. Without loss of generality, assume $\alpha_1 < \alpha_2 < \dots$. Let $M_1 \in \mathbb{N}$ such that $\alpha_{M_1} > \alpha_0$.

For $j \in \{1, \dots, l\}$ let $E_j = \{y \in Y : ||T_\alpha f - g_j||_y < \varepsilon/2 \text{ for some } \alpha \in \mathcal{F}^{(2)}\}$. Then we can assume that $\nu(E_j) = 1$ for all $1 \leq j \leq l$, otherwise we may modify g_j so that $g_j = T_{\alpha_{M_1}} f$ on $\pi^{-1}(E_j^c)$ without affecting (4.7).

For $j \in \{1, \dots, l\}$ let $E_{j,n} = \{y \in Y : ||T_\beta f - g_j||_y < \varepsilon/2 \text{ for some } \beta \in \mathrm{FU}(\{\alpha_i\}_{i=1}^{n})\}$. Then $\nu(\bigcup_{n \geq 1} E_{j,n}) = 1$ for all $1 \leq j \leq l$. Thus there is $M_2 \in \mathbb{N}$ such that

$\nu(\bigcap_{j=1}^{l} E_{j,n}) > 1 - \varepsilon/2$ for any $n \geq M_2$. Let $M = \max\{M_1, M_2\}$. Then for any $\alpha \in \mathcal{F}^{(2)}$ with $\alpha > \alpha_M$ (that is, $\alpha \in \mathrm{FU}(\{\alpha_i\}_{i>M})$)

$$\inf_{\beta \in \mathrm{FU}(\{\alpha_i\}_{i=1}^M)} ||T_\alpha f - T_\beta f||_y \leq \inf_{1\leq j\leq l} \{||T_\alpha f - g_j||_y + \inf_{\beta \in \mathrm{FU}(\{\alpha_i\}_{i=1}^M)} ||T_\beta f - g_j||_y\} < \varepsilon$$

for all y not in $E_\alpha = E'_\alpha \cup (\bigcap_{j=1}^{l} E_{j,M})^c$ which has measure less than ε.
Now we show the converse. If $f \notin L^2(X, \mathcal{K}_F(X|Y), \mu)$, then $f = f_1 + f_2$, where $f_1 \in L^2(X, \mathcal{K}_F(X|Y), \mu)$ and $f_2 \in L^2(X, \mathcal{K}_F(X|Y), \mu)^\perp$ with f_2 non trivial. One deduces that there is $\varepsilon > 0$ such that $||f_2||_y^2 \geq 3\varepsilon$ for any y in a set $E \in \mathcal{Y}$ of measure greater than 2ε.
By Theorem 4.7, $\mathrm{IP}-\lim_{\alpha\in\mathcal{F}^{(1)}} \int_Y |\mathbb{E}(gT_\alpha f_2|\mathcal{Y})| d\nu = 0$ for any $g \in L^2(X, \mathcal{X}, \mu)$. Fix any IP-ring $\mathcal{F}^{(2)} = \mathrm{FU}(\{\alpha_i\}_{i\in\mathbb{N}}) \subseteq \mathcal{F}^{(1)}$ and $M \in \mathbb{N}$. Since

$$\mathrm{IP} - \lim_{\alpha\in\mathcal{F}^{(1)}} \int_Y \left|\mathbb{E}(T_\beta \overline{f}_2 T_\alpha f_2|\mathcal{Y})\right| d\nu = 0$$

holds for any $\beta \in \mathrm{FU}(\{\alpha_i\}_{i=1}^M)$, then there is $\alpha_{M'} \in \mathcal{F}^{(2)}$, $M' > M$, such that for any $\alpha \in \mathcal{F}^{(2)}$ with $\alpha > \alpha_{M'}$, there exists $A_\alpha \in \mathcal{Y}$ with $\nu(A_\alpha) > 1 - \varepsilon$ such that $|\langle T_\alpha f_2, T_\beta f_2\rangle_y| < \varepsilon$ for any $y \in A_\alpha$. Then

$$||T_\alpha f_2 - T_\beta f_2||_y^2 = ||T_\alpha f_2||_y^2 + ||T_\beta f_2||_y^2 - 2\mathrm{Re}\langle T_\alpha f_2, T_\beta f_2\rangle_y > \varepsilon > \varepsilon^2$$

holds for any $y \in T_\alpha^{-1} E \cap A_\alpha$ the measure of which is $\nu(T_\alpha^{-1} E \cap A_\alpha) > \varepsilon$.
So for any $\alpha \in \mathcal{F}^{(2)}$ with $\alpha > \alpha_{M'}$, one has $\nu\{y \in Y : \inf_{\beta\in\mathrm{FU}(\{\alpha_i\}_{i=1}^M)} ||T_\alpha f - T_\beta f||_y^2 < \varepsilon^2\} \leq \nu\{y \in Y : \inf_{\beta\in\mathrm{FU}(\{\alpha_i\}_{i=1}^M)} ||T_\alpha f_2 - T_\beta f_2||_y^2 < \varepsilon^2\} < 1 - \varepsilon$. This contradicts the assumption of the proposition. □

Definition 4.9. Let $\{n_\alpha\}_{\alpha\in\mathcal{F}}$ be an IP set and $\mathcal{F}^{(1)}$ and IP-ring. Any subset having the form

$$\{n_{\beta_i}\}_{i\in\mathbb{N}} \subseteq \{n_\alpha\}_{\alpha\in\mathcal{F}^{(1)}} \tag{4.8}$$

with $\beta_1 < \beta_2 < \ldots$ is called an $\mathcal{F}-$*monotone subset* of $\{n_\alpha\}_{\alpha\in\mathcal{F}^{(1)}}$.

A similar proof of Proposition 4.8 yields to the following corollary.

Corollary 4.10. *$f \in L^2(X, \mathcal{K}_F(X|Y), \mu)$ if and only if for any any $\varepsilon > 0$ and any $\mathcal{F}-$monotone subset $\{n_{\beta_i}\}_{i\in\mathbb{N}} \subseteq \{n_\alpha\}_{\alpha\in\mathcal{F}^{(1)}}$ there exists $M \in \mathbb{N}$ such that for every $n \in \mathbb{N}$ with $n > M$ there is $E_n \in \mathcal{Y}$ with $\nu(E_n) < \varepsilon$ verifying for any $y \notin E_n$,*

$$\inf_{1\leq i\leq n-1} ||T_{\beta_n} f - T_{\beta_i} f||_y < \varepsilon.$$

4.3. **Conditional sequence entropy and rigid algebra $\mathcal{K}_F(X|Y)$.** As in the previous section in this subsection we study relations between conditional sequence entropy and the rigid algebra $\mathcal{K}_F(X|Y)$. The main result of the subsection is the following.

Theorem 4.11. *Let $(X, \mathcal{X}, \mu, T)$ be a system and $(Y, \mathcal{Y}, \nu, T)$ one of its factors. For any IP set $\{n_\alpha\}_{\alpha\in\mathcal{F}}$ there exists an IP-ring $\mathcal{F}^{(1)}$ of $\mathcal{F}$ such that for any IP-ring $\mathcal{F}^{(2)} \subseteq \mathcal{F}^{(1)}$ there exists an $\mathcal{F}$-monotone subset $A \subseteq \{n_\alpha\}_{\alpha\in\mathcal{F}^{(2)}}$ such that*

$$h_\mu^A(T, \xi|\mathcal{Y}) = H_\mu(\xi|\mathcal{K}_F(X|Y)) \tag{4.9}$$

for all measurable partitions ξ of X with $H_\mu(\xi|\mathcal{Y}) < \infty$, where $F = \{n_\alpha\}_{\alpha\in\mathcal{F}^{(1)}}$.
In particular, for any IP *set $\{n_\alpha\}_{\alpha\in\mathcal{F}}$ there exists an* IP*-ring $\mathcal{F}^{(1)}$ of $\mathcal{F}$ such that*

$$\max\left\{h_\mu^A(T,\xi|\mathcal{Y}) : A \subseteq F \text{ is } \mathcal{F}\text{-monotone}\right\} = H_\mu(\xi|\mathcal{K}_F(X|Y)), \tag{4.10}$$

where $F = \{n_\alpha\}_{\alpha\in\mathcal{F}^{(1)}}$.

Remark 4.12. The reason why we need to consider the given IP set on an IP-ring $\mathcal{F}^{(1)}$ comes from the fact that in our proof we strongly use Theorem 4.3 and its consequence stated in (4.1). In fact the theorem works for any IP-ring $\mathcal{F}^{(1)}$ such that condition (4.1) holds.

By Theorem 4.11 with $\mathcal{Y}$ trivial, one obtains the following result immediately.

Corollary 4.13. *Let $(X, \mathcal{X}, \mu, T)$ be a system and $\{n_\alpha\}_{\alpha\in\mathcal{F}}$ be an* IP *set. Then there exists an* IP*-ring $\mathcal{F}^{(1)}$ of $\mathcal{F}$ such that for any* IP*-ring $\mathcal{F}^{(2)} \subseteq \mathcal{F}^{(1)}$ there exists an $\mathcal{F}$-monotone subset $A \subseteq \{n_\alpha\}_{\alpha\in\mathcal{F}^{(2)}}$ such that*

$$h_\mu^A(T,\xi) = H_\mu(\xi|\mathcal{K}_F(X)) \tag{4.11}$$

for all measurable partitions ξ of X with $H_\mu(\xi) < \infty$, where $F = \{n_\alpha\}_{\alpha\in\mathcal{F}^{(1)}}$.
In particular, for any IP *set $\{n_\alpha\}_{\alpha\in\mathcal{F}}$ there exists an* IP*-ring $\mathcal{F}^{(1)}$ of $\mathcal{F}$ such that*

$$\max\left\{h_\mu^A(T,\xi) : A \subseteq F \text{ is } \mathcal{F}\text{-monotone}\right\} = H_\mu(\xi|\mathcal{K}_F(X)), \tag{4.12}$$

where $F = \{n_\alpha\}_{\alpha\in\mathcal{F}^{(1)}}$.

The proof of Theorem 4.11 will follow directly from the following four lemmas.

Lemma 4.14. *Let $\{n_\alpha\}_{\alpha\in\mathcal{F}}$ be an* IP *set. Then there exists an* IP*-ring $\mathcal{F}^{(1)}$ of $\mathcal{F}$ such that for any* IP*-ring $\mathcal{F}^{(2)} \subseteq \mathcal{F}^{(1)}$ there exists an $\mathcal{F}-$monotone subset $A \subseteq \{n_\alpha\}_{\alpha\in\mathcal{F}^{(2)}}$ such that*

$$h_\mu^A(T,\xi|\mathcal{Y}) \geq H_\mu(\xi|\mathcal{K}_F(X|Y))$$

for all measurable partitions ξ of X with $H_\mu(\xi|\mathcal{Y}) < \infty$, where $F = \{n_\alpha\}_{\alpha\in\mathcal{F}^{(1)}}$.

Proof. The proof is similar to the proof of Lemma 3.6. We only point out the differences.
Let $\mathcal{F}^{(1)}$ be the IP-ring such that (4.1) holds. Consider $\mathcal{K}_F = \mathcal{K}_F(X|Y)$ to be the σ-algebra associated to $F = \{n_\alpha\}_{\alpha\in\mathcal{F}^{(1)}}$. All the results in the last subsection hold for this factor.
Let $\mathcal{G}_F$ be the family generated by $\{\{n_\alpha\}_{\alpha\in\mathcal{F}^{(2)}} : \mathcal{F}^{(2)}$ is an IP-ring of $\mathcal{F}^{(1)}\}$, that is, the family of all IP-subsets of F. Then by Hindman's Theorem $\mathcal{G}_F^*$ is a filter.
Claim: *For any measurable finite partitions ξ and η of X and $\varepsilon > 0$, there exists a sequence $S \in \mathcal{G}_F^*$ such that for any $m \in S$*

$$\int_Y H_y(T^{-m}\xi|\eta)d\nu(y) \geq H_\mu(\xi|\mathcal{K}_F) - \varepsilon. \tag{4.13}$$

Proof of Claim: It is easy to verify that $\mathcal{G}_F^* = \{S \subseteq \mathbb{Z}_+ :$ there exists $\alpha_0 \in \mathcal{F}^{(1)}$ such that for any $\alpha \in \mathcal{F}^{(1)}$ with $\alpha > \alpha_0$, $n_\alpha \in S\}$. For any $A, B \in \mathcal{X}$, since $1_A - \mathbb{E}(1_A|\mathcal{K}_F) \in L^2(X, \mathcal{K}_F, \mu)^\perp$, from Theorem 4.7 one deduces

$$\mathrm{IP} - \lim_{\alpha \in \mathcal{F}^{(1)}} \int_Y \Big| \int_X T_\alpha \Big(1_A - \mathbb{E}(1_A|\mathcal{K}_F)\Big) \cdot 1_B d\mu_y \Big| d\nu = 0.$$

Equivalently,

$$\mathcal{G}_F^* - \lim \int_Y \Big| \int_X T^n \Big(1_A - \mathbb{E}(1_A|\mathcal{K}_F)\Big) \cdot 1_B d\mu_y \Big| d\nu = 0.$$

Let $\xi = \{A_1, \ldots, A_k\}$ and $\eta = \{B_1, \ldots, B_l\}$. Since $\mathcal{G}_F^*$ is a filter, for any fixed $\delta_1, \delta_2 > 0$ there exists $S \in \mathcal{G}_F^*$ such that for any $m \in S$ there is a set $E_m \in \mathcal{Y}$ with $\nu(E_m) > 1 - \delta_1$ verifying

$$\Big|\mu_y(T^{-m}A_i \cap B_j) - \int_X T^m \mathbb{E}(1_{A_i}|\mathcal{K}_F) 1_{B_j} d\mu_y\Big| < \delta_2 \tag{4.14}$$

for all $1 \leq i \leq k, 1 \leq j \leq l$ and $y \in E_m$. One concludes in a similar way as in the the proof of the Claim in the proof of Lemma 3.6.

For any IP-ring $\mathcal{F}^{(2)} \subseteq \mathcal{F}^{(1)}$, $S = \{n_\alpha\}_{\alpha \in \mathcal{F}^{(2)}}$ is an IP set and hence $S \in \mathcal{G}_F$. Thus in a similar way than the proof of Lemma 3.6, one can complete this proof. We omit the details. □

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

Lemma 4.15. *Let $r \geq 1$ be a fixed integer. For each $\epsilon > 0$ there exists $\delta > 0$ such that if $\xi = \{A_1, \cdots, A_r\}$ and $\eta = \{C_1, \cdots, C_r\}$ are two measurable partitions of X with $\sum_{i=1}^r \mu(A_i \Delta C_i) < \delta$ then the Rohlin metric $\rho_\mu(\xi, \eta) = H_\mu(\xi|\eta) + H_\mu(\eta|\xi) < \epsilon$.*

Lemma 4.16. *Let $\{n_\alpha\}_{\alpha \in \mathcal{F}}$ be an* IP *set and $B \in \mathcal{X}$. Then there exists an* IP*-ring $\mathcal{F}^{(1)}$ of $\mathcal{F}$ such that $B \in \mathcal{K}_F(X|Y)$ if and only if $h_\mu^A(T, \{B, B^c\}|\mathcal{Y}) = 0$ holds for all $\mathcal{F}-$monotone subsets $A \subseteq F$, where $F = \{n_\alpha\}_{\alpha \in \mathcal{F}^{(1)}}$.*

Proof. Let $\mathcal{F}^{(1)}$ be the IP-ring such that (4.1) holds. Assume $B \in \mathcal{K}_F(X|Y)$. Let $A = \{n_{\beta_i}\}_{i \in \mathbb{N}}$ be an $\mathcal{F}-$monotone subset of $F = \{n_\alpha\}_{\alpha \in \mathcal{F}^{(1)}}$ and $\xi = \{B, B^c\}$.
Observe that $||T^n 1_B - T^m 1_B||_y^2 = \mu_y(T^{-n}B \Delta T^{-m}B)$. Since $1_B \in L^2(X, \mathcal{K}_F(X|Y), \mu)$, by Corollary 4.10, there is $M \in \mathbb{N}$ such that for any $n > M$ there exists $E_n \in \mathcal{Y}$ with $\nu(E_n) < \delta$ verifying for any $y \notin E_n$

$$\inf_{1 \leq j \leq n-1} ||T_{\beta_n} 1_B - T_{\beta_j} 1_B||_y < \delta,$$

where δ is chosen as in Lemma 4.15 small enough such that $\rho_y(\xi, \eta) := \rho_{\mu_y}(\xi, \eta) < \varepsilon$ for $r = 2$.
Hence for any $n > M$ there is $1 \leq j(n) \leq n-1$ such that $\rho_y(T_{\beta_n}^{-1}\xi, T_{\beta_{j(n)}}^{-1}\xi) < \varepsilon$. So

$$H_y(T_{\beta_n}^{-1}\xi | \bigvee_{i=1}^{n-1} T_{\beta_i}^{-1}\xi) \leq H_y(T_{\beta_n}^{-1}\xi | T_{\beta_{j(n)}}^{-1}\xi) \leq \rho_y(T_{\beta_n}^{-1}\xi, T_{\beta_{j(n)}}^{-1}\xi) < \varepsilon.$$

For any $y\in Y$, $H_y(T_{\beta_n}^{-1}\xi|T_{\beta_j}^{-1}\xi)\le\log 2$. Therefore

$$\begin{aligned}
h_\mu^A(T,\xi|\mathcal{Y}) &= \limsup_{n\to\infty}\frac{1}{n}\int_Y H_y(\bigvee_{i=1}^n T_{\beta_i}^{-1}\xi)d\nu(y)\\
&= \limsup_{n\to\infty}\frac{1}{n}\int_Y \sum_{i=2}^n H_y(T_{\beta_i}^{-1}\xi|\bigvee_{j=1}^{i-1}T_{\beta_j}^{-1}\xi)d\nu(y)\\
&= \limsup_{n\to\infty}\frac{1}{n}\Bigg(\int_{E_n}\sum_{i=2}^n H_y(T_{\beta_i}^{-1}\xi|\bigvee_{j=1}^{i-1}T_{\beta_j}^{-1}\xi)\\
&+ \int_{Y\setminus E_n}\sum_{i=2}^n H_y(T_{\beta_i}^{-1}\xi|\bigvee_{j=1}^{i-1}T_{\beta_j}^{-1}\xi)d\nu(y)\Bigg)\\
&\le \delta\log 2+\varepsilon.
\end{aligned}$$

Since ε and δ are arbitrary, it follows that $h_\mu^A(T,\xi|\mathcal{Y})=0$.
Now we show the converse. If $B\notin\mathcal{K}_F(X|Y)$ then $H_\mu(\{B,B^c\}|\mathcal{K}_F(X|Y))>0$. So by Lemma 4.14 there exists an $\mathcal{F}-$monotone subset A of F such that

$$h_\mu^A(T,\{B,B^c\}|\mathcal{Y})\ge H_\mu(\{B,B^c\}|\mathcal{K}_F(X|Y))>0$$

This completes the proof of Lemma 4.16. □

Lemma 4.17. *Let $\{n_\alpha\}_{\alpha\in\mathcal{F}}$ be an* IP *set. Then there exists an* IP*-ring $\mathcal{F}^{(1)}$ of $\mathcal{F}$ such that for any finite measurable partition ξ of X and any $\mathcal{F}-$monotone subset $A\subseteq F$, one has*

$$h_\mu^A(T,\xi|\mathcal{Y})\le H_\mu(\xi|\mathcal{K}_F(X|Y)), \tag{4.15}$$

where $F=\{n_\alpha\}_{\alpha\in\mathcal{F}^{(1)}}$.

Proof. Let $\mathcal{F}^{(1)}$ be the IP-ring such that (4.1) holds. In the proof of Lemma 3.8, replace $\mathcal{K}(X|Y)$ by $\mathcal{K}_F(X|Y)$ and use lemma 4.16. □

4.4. **Rigid and mild mixing extensions.** Let $(X,\mathcal{X},\mu,T)$ be a system and $(Y,\mathcal{Y},\nu,T)$ one of its factors.

Definition 4.18. [FK] Let $\{n_\alpha\}_{\alpha\in\mathcal{F}}$ be an IP set and $\mathcal{F}^{(1)}$ an IP-ring. The system $(X,\mathcal{X},\mu,T)$ is $\mathcal{Y}-$*mixing along* $\mathcal{F}^{(1)}$ if for each $f,g\in L^2(X\times X,\mathcal{X}\otimes\mathcal{X},\mu\times_Y\mu)$

$$\text{IP}-\lim_{\alpha\in\mathcal{F}^{(1)}}\Big\{\int_{X\times X}g\cdot(T\times T)_\alpha f d\mu\times_Y\mu-\int_Y\mathbb{E}(g|\mathcal{Y})\cdot(T\times T)_\alpha\mathbb{E}(f|\mathcal{Y})d\nu\Big\}=0. \tag{4.16}$$

Remark 4.19. It is easy to prove that [Z2, Theorem 2.5.] $(X,\mathcal{X},\mu,T)$ is $\mathcal{Y}-$mixing along $\mathcal{F}^{(1)}$ if and only if for each $f,g\in L^2(X,\mathcal{X},\mu)$

$$\text{IP}-\lim_{\alpha\in\mathcal{F}^{(1)}}\int_X\Big|\mathbb{E}(gT_\alpha f|\mathcal{Y})-\mathbb{E}(g|\mathcal{Y})T_\alpha\mathbb{E}(f|\mathcal{Y})\Big|d\mu=0.$$

Proposition 4.20. *Let $\{n_\alpha\}_{\alpha\in\mathcal{F}}$ be an* IP *set and $\mathcal{F}^{(1)}$ an IP-ring as in (4.1). Then the following conditions are equivalent:*

(1) *$(X,\mathcal{X},\mu,T)$ is $\mathcal{Y}-$mixing along $\mathcal{F}^{(1)}$;*

(2) *For any* $f \in L^2(X, \mathcal{Y}, \mu)^{\perp}$ *and* $g \in L^2(X, \mathcal{X}, \mu)$
$$\mathrm{IP} - \lim_{\alpha \in \mathcal{F}^{(1)}} \int_Y \big|\mathbb{E}(gT_\alpha f|\mathcal{Y})\big| d\nu = 0;$$
(3) $\mathcal{K}_F(X|Y) = \mathcal{Y}$, *where* $F = \{n_\alpha\}_{\alpha \in \mathcal{F}^{(1)}}$.

Proof. It follows from Theorem 4.7. □

Definition 4.21.

(1) $(X, \mathcal{X}, \mu, T)$ is a *rigid extension over* $\mathcal{Y}$ if there exists an IP set F such that $\mathcal{X} = \mathcal{K}_F(X|Y)$;
(2) $(X, \mathcal{X}, \mu, T)$ is a *mild mixing extension over* $\mathcal{Y}$ if for any IP set F, $\mathcal{K}_F(X|Y) = \mathcal{Y}$.

The proof of the following theorem follows easily from the work done in previous subsection. We left it to the reader.

Theorem 4.22.

(1) $(X, \mathcal{X}, \mu, T)$ *is a rigid extension over* $\mathcal{Y}$ *if and only if there exists an* IP *set* F *such that* $h^A_\mu(T|\mathcal{Y}) = 0$ *holds for all* $\mathcal{F}-$*monotone subsets* $A \subseteq F$*;*
(2) *Let* $\{n_\alpha\}_{\alpha \in \mathcal{F}}$ *be an* IP *set and* $\mathcal{F}^{(1)}$ *as in (4.1). Then* $(X, \mathcal{X}, \mu, T)$ *is* $\mathcal{Y}-$*mixing along* $\mathcal{F}^{(1)}$ *if and only if for any* IP*-ring* $\mathcal{F}^{(2)} \subseteq \mathcal{F}^{(1)}$ *there exists an* $\mathcal{F}-$*monotone subset* $A \subseteq \{n_\alpha\}_{\alpha \in \mathcal{F}^{(2)}}$ *such that*
$$h^A_\mu(T, \xi|\mathcal{Y}) = H_\mu(\xi|\mathcal{Y})$$
for all measurable partitions ξ *with* $H_\mu(\xi|\mathcal{Y}) < \infty$.
(3) $(X, \mathcal{X}, \mu, T)$ *is a mild mixing extension over* $\mathcal{Y}$ *if and only if for any* IP *set* F *there exists an* $\mathcal{F}-$*monotone subset* $A \subseteq F$ *such that*
$$h^A_\mu(T, \xi|\mathcal{Y}) = H_\mu(\xi|\mathcal{Y})$$
for all measurable partitions ξ *with* $H_\mu(\xi|\mathcal{Y}) < \infty$.

Remark 4.23. Parts (1) and (3) of Theorem 4.22 appear in [Z2], and the corresponding cases when $(Y, \mathcal{Y}, \nu, T)$ is the trivial factor appear in [Z1], but the results are stated in a slightly different language. Observe that the way to define $AP(\mathcal{Y}, \{n_\alpha\}, \mathcal{F}^{(1)})$ (Definition 4.4) and a rigid extension (Definition 4.21) differ from [Z2]. Also our method is different from that in [Z2].

5. An application

In this section we give an application of either Theorem 3.4 or Theorem 4.11.

Theorem 5.1. *Let* $(Y, \mathcal{Y}, \nu, T)$ *be a factor of the ergodic system* $(X, \mathcal{X}, \mu, T)$*. Then*
$$\max_A \{h^A_\mu(T|\mathcal{Y})\} \in \{\log k : k \in \mathbb{N}\} \cup \{\infty\}.$$

Proof. Denote by $(K, \mathcal{K}, \kappa, T)$ the factor of $(X, \mathcal{X}, \mu, T)$ associated to the T-invariant σ-algebra $\mathcal{K}(X|Y)$. Since $(X, \mathcal{X}, \mu, T)$ is ergodic, by Rohlin Theorem it is isomorphic to a skew-product over $(K, \mathcal{K}, \kappa, T)$. Explicitly, there exists a probability space $(U, \mathcal{U}, \rho)$, that may be a finite set with the uniform probability measure or the unit

interval with the Lebesgue measure, and a measurable function γ from Y to the automorphisms of $(U,\mathcal{U},\rho)$ such that $(X,\mathcal{X},\mu,T) \cong (K\times U,\mathcal{K}\otimes\mathcal{U},\kappa\times\rho,T_\gamma)$, where $T_\gamma(x,u)=(Tx,\gamma(x)u)$.
If ξ_1 and ξ_2 are finite partitions of K and U respectively, define partitions ξ_1' and ξ_2' of $K\times U$ by $\xi_1'=\xi_1\times U=\{B\times U : B\in\xi_1\}$ and $\xi_2'=K\times\xi_2=\{K\times B : B\in\xi_2\}$. By Theorem 3.4 for any sequence $S\in\mathcal{D}^*$, there exists a subsequence $A\subseteq S$ such that

$$h^A_{\kappa\times\rho}(T_\gamma,\xi_1\times\xi_2|\mathcal{Y}) = H_{\kappa\times\rho}(\xi_1\times\xi_2|\mathcal{K}) = H_{\kappa\times\rho}(\xi_1'\vee\xi_2'|\mathcal{K}) = H_{\kappa\times\rho}(\xi_2') = H_\rho(\xi_2).$$

Thus, $h^A_\mu(T|\mathcal{Y}) = \sup\limits_{\xi_1,\xi_2} h^A_{\kappa\times\rho}(T_\gamma,\xi_1\times\xi_2|\mathcal{Y})$ is $\log k$ for some $k\in\mathbb{N}$ if U is a finite set or ∞ if U is continuous. So

$$h^A_\mu(T|\mathcal{Y})\in\{\log k : k\in\mathbb{N}\}\cup\{\infty\}.$$

In particular, one gets the assertion. $\square$

Remark 5.2. (1) One can use Theorem 4.11 instead of Theorem 3.4 to prove Theorem 5.1.

(2) In fact, what is proved in the Theorem 5.1 is that for any sequence $S\in\mathcal{D}^*$, there exists a subsequence $A\subseteq S$ such that

$$h^A_\mu(T|\mathcal{Y})\in\{\log k : k\in\mathbb{N}\}\cup\{\infty\}.$$

And similarly using Theorem 4.11 one can show that for any IP set $\{n_\alpha\}_{\alpha\in\mathcal{F}}$ there exists an IP-ring $\mathcal{F}^{(1)}$ of $\mathcal{F}$ such that for any IP-ring $\mathcal{F}^{(2)}\subseteq\mathcal{F}^{(1)}$ there exists an $\mathcal{F}-$monotone subset $A\subseteq\{n_\alpha\}_{\alpha\in\mathcal{F}^{(2)}}$ such that

$$h^A_\mu(T|\mathcal{Y})\in\{\log k : k\in\mathbb{N}\}\cup\{\infty\}.$$

Acknowledgement: We would like to thank the referee for a careful reading and helpful suggestions.

References

[BD] V. Bergelson and T. Downarowicz, *Large sets of integers and hierarchy of mixing properties of measure preserving systems*, Colloq. Math. **110** (2008), no. 1, 117–150.

[BM] V. Bergelson and R. McCutcheon, *An ergodic IP polynomial Szemerédi theorem*, Mem. Amer. Math. Soc. **146** (2000), no. 695.

[B-R] F. Blanchard, B. Host, A. Maass, S. Martínez, D.J. Rudolph, *Entropy paris for a measure*, Ergodic Th. and Dyn. Systems **15** (1995), 621–632.

[F1] H. Furstenberg, *Recurrence in ergodic theory and combinatorial number theory*, Princeton Univ. Press, 1981.

[F2] 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.

[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.

[Hu1] P. Hulse, *Sequence entropy and subsequence generators*, J. London Math. Soc. **26** (1982), no. 2, 441–450.

[Hu2] P. Hulse, *Sequence entropy relative to an invariant σ-algebra*, J. London Math. Soc. **33** (1986), no. 2, 59–72.

[HMY] W. Huang, A. Maass and X. Ye, *Sequence entropy pairs and complexity pairs for a measure*, Annales de l'Institut Fourier **54** (2004), no. 4, 1005–1028.

[HSY] W. Huang, S. Shao and X. Ye, *Mixing via sequence entropy*, Algebraic and topological dynamics, 101–122, Contemp. Math. **385**, Amer. Math. Soc., Providence, RI, 2005.

[KY] R. Kuang and X.D. Ye, *Mixing via families for measure preserving transformations*, Colloq. Math. **110** (2008), no. 1, 151–165.

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

[P] W. Parry, *Topics in ergodic theory*, Cambridge Tracts in Mathematics **75**, Cambridge University Press, 1981.

[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.

Departamento de Ingeniería Matemática, Universidad de Chile, Av. Blanco Encalada 2120, Santiago, Chile.
E-mail address: `acoronel@dim.uchile.cl`

Centro de Modelamiento Matemático and Departamento de Ingeniería Matemática, Universidad de Chile, Av. Blanco Encalada 2120, Santiago, Chile.
E-mail address: `amaass@dim.uchile.cl`

Department of Mathematics, University of Science and Technology of China, Hefei, Anhui, 230026, P.R. China.
E-mail address: `songshao@ustc.edu.cn`