CrossMark

Mean Li–Yorke chaos along some good sequences

Jian Li[1] · **Yixiao Qiao[2,3]**

Abstract If a topological dynamical system (X, T) has positive topological entropy, then it is multivariant mean Li–Yorke chaotic along a sequence $\{a_k\}_{k=1}^{\infty}$ of positive integers which is "good" for pointwise ergodic convergence with a mild condition; more specifically, there exists a Cantor subset K of X such that for every $n \geq 2$ and pairwise distinct points $x_1, x_2, \ldots, x_n$ in K we have

$$\liminf_{N \to \infty} \frac{1}{N} \sum_{k=1}^{N} \max_{1 \leq i < j \leq n} d\left(T^{a_k} x_i, T^{a_k} x_j\right) = 0$$

and

$$\limsup_{N \to \infty} \frac{1}{N} \sum_{k=1}^{N} \min_{1 \leq i < j \leq n} d\left(T^{a_k} x_i, T^{a_k} x_j\right) > 0.$$

Examples are given for the classic sequences of primes and generalized polynomials.

Communicated by A. Constantin.

✉ Yixiao Qiao
yxqiao@mail.ustc.edu.cn

Jian Li
lijian09@mail.ustc.edu.cn

1 Department of Mathematics, Shantou University, Shantou 515063, Guangdong, People's Republic of China

2 Department of Mathematics, University of Science and Technology of China, Hefei 230026, Anhui, People's Republic of China

3 Institute of Mathematics, Polish Academy of Sciences, ul. Śniadeckich 8, 00-656 Warsaw, Poland

🌀 Springer

1 Introduction

In investigations of topological dynamical systems, different versions of chaos, which represent complexity in various senses, attract a lot of attention over the past few decades. The most studied chaos are Devaney chaos, Li–Yorke chaos, distributional chaos, positive entropy and weak mixing. The relationship among them naturally became a central topic as well. It was known that weak mixing implies Li–Yorke chaos [24]. In 2002, Huang and Ye [20] showed that Devaney chaos implies Li–Yorke chaos. Later, using ergodic theoretic method, Blanchard, Glasner, Kolyada and Maass [3] proved that positive topological entropy also implies Li–Yorke chaos. A combinatorial proof was given by Kerr and Li [25].

Many of the classical notions in topological dynamical systems have an analogous version in the mean sense. Downarowicz [11] observed that mean Li–Yorke chaos is equivalent to the so-called DC2 chaos and proved that positive topological entropy implies mean Li–Yorke chaos, which strengthens the result of [3]. Huang, Ye and the first author [22] provided a different approach and showed that positive topological entropy implies a multivariant version of mean Li–Yorke chaos. For related topics on chaotic behaviours we refer to [19,23,26] for general group actions, [15] for a new condition implying mean Li–Yorke chaos, [2,13,18] for more careful discussions, and the survey article [28] for more aspects and details.

It is natural to study "sequence versions" of dynamical notions (e.g., sequence entropy). In order to make our statement explicit, we restate the main theorem of [22] in the following, which concerns the terminology of multivariant mean Li–Yorke chaos. By a *topological dynamical system*, we mean a pair (X, T), where X is a compact metric space equipped with a metric d and T is a homeomorphism of X onto itself.

Theorem 1.1 ([22, Theorem 1.1]). *If a topological dynamical system (X, T) has positive topological entropy, then it is multivariant mean Li–Yorke chaotic; namely, there exists a Cantor subset K of X such that for every integer $n \geq 2$ and pairwise distinct points $x_1, x_2, \ldots, x_n$ in K it holds that*

$$\liminf_{N \to \infty} \frac{1}{N} \sum_{k=1}^{N} \max_{1 \leq i < j \leq n} d\left(T^k x_i, T^k x_j\right) = 0$$

and

$$\limsup_{N \to \infty} \frac{1}{N} \sum_{k=1}^{N} \min_{1 \leq i < j \leq n} d\left(T^k x_i, T^k x_j\right) > 0.$$

Motivated by the above consideration and results we naturally ask if "sequence chaos" appears in positive entropy systems; more precisely, we investigate the following question.

Question 1.2 Let $\{a_k\}_{k=1}^{\infty}$ be a sequence of positive integers. Suppose that (X, T) is a topological dynamical system with positive entropy. Is there a Cantor (in particular, uncountable) subset K of X such that for every integer $n \geq 2$ and pairwise distinct points $x_1, x_2, \ldots, x_n$ in K it holds that

$$\liminf_{N \to \infty} \frac{1}{N} \sum_{k=1}^{N} \max_{1 \leq i < j \leq n} d\left(T^{a_k} x_i, T^{a_k} x_j\right) = 0$$

and

$$\limsup_{N \to \infty} \frac{1}{N} \sum_{k=1}^{N} \min_{1 \leq i < j \leq n} d\left(T^{a_k} x_i, T^{a_k} x_j\right) > 0 ?$$

As it was shown in previous results (e.g., Theorem 1.1), Birkhoff's pointwise ergodic theorem plays a key role. To study Question 1.2, we would follow this idea. However, for general sequences of positive integers the situation of pointwise ergodic convergence becomes much more complicated. Fortunately, there were classic pointwise ergodic theorems for subsequences of positive integers as well (see Sect. 3.3 for details), which are quite useful for our argument. It is worth mentioning that the proofs of these ergodic theorems are extremely nontrivial. All these facts lead to the following restrictions on sequences.

Definition 1.3 An increasing sequence $\{a_k\}_{k=1}^{\infty}$ of positive integers is *pointwise good* if for each measure preserving system $(X, \mathscr{B}, \mu, T)$ and $f \in L^2(\mu)$,

$$\frac{1}{N} \sum_{k=1}^{N} f(T^{a_k} x)$$

converges for μ-a.e. $x \in X$.

There are a lot of pointwise good sequences. In a series of papers, Bourgain [5–7] proved that $\{[p(k)]\}$ is pointwise good provided $p(x)$ is a polynomial of real coefficients, where $[\cdot]$ denotes the integer part of a real number. This result was generalized to some logarithmic–exponential subpolynomials [4]. The sequence of primes [5] and many return time sequences [6,8,27] are also pointwise good. We refer to [9,30] for a more comprehensive understanding of pointwise ergodic theorems.

If a sequence is pointwise good, then the ergodic average along this sequence converges almost everywhere. Nonetheless, we do not have a further understanding of the value of the limit. Following [27], we say that a sequence $\{a_k\}$ is *very good* if for any ergodic system $(X, \mathscr{B}, \mu, T)$ and $f \in L^2(\mu)$, it holds that

$$\lim_{N \to \infty} \frac{1}{N} \sum_{k=1}^{N} f(T^{a_k} x) = \int f \, d\mu \quad \text{for } \mu\text{-a.e. } x \in X.$$

In this terminology, Birkhoff's pointwise ergodic theorem states that the sequence of natural numbers is very good. Following the ideas in [22], it is routine to check that Theorem 1.1 also holds along any very good sequence; namely, the answer to Question 1.2 is affirmative for every very good sequence. For example, it was shown in [27] that the Morse sequence is very good and in [4] that sequences generalized by some logarithmic–exponential subpolynomials (including sequences $\{[k^r]\}$ for all non-integer numbers $r > 0$) are very good.

Unfortunately, there are only few pointwise good sequences which are verified to be very good. Comparing to the unsolved cases of Question 1.2, this is far from satisfying. To deal with the remaining cases, we follow the idea in [14] to study the limit of ergodic average along sequences.

Definition 1.4 A sub-σ-algebra $\mathscr{F}$ of $\mathscr{B}$ is a *characteristic σ-algebra for the sequence* $\{a_k\}$ if for every $f \in L^2(\mu)$,

$$\frac{1}{N} \sum_{k=1}^{N} T^{a_k} f - \frac{1}{N} \sum_{k=1}^{N} T^{a_k} E(f|\mathscr{F}) \to 0$$

as $N \to \infty$ in $L^2(\mu)$, where $E(\cdot|\mathscr{F})$ is the conditional expectation with respect to $\mathscr{F}$.

As we will show in Theorem 3.2, the Pinsker σ-algebra of a system is characteristic for all sequences $\{a_k\}$ satisfying Condition (∗) below, which helps us overcome additional difficulties in the proof of our main theorem.

Condition (∗). For every $L > 0$,

$$\lim_{N \to \infty} \frac{1}{N^2} \#\{(i, j) \in [1, N]^2 : |a_i - a_j| \le L\} = 0.$$

We note here that Condition (∗) is quite mild. In fact, it is not hard to check that any strictly increasing sequence of positive integers satisfies Condition (∗) (see Lemma 3.1), and that all those pointwise good sequences we previously mentioned satisfy Condition (∗) as well (see Remark 3.15). Moreover, for any pointwise good sequence satisfying Condition (∗), the limit of ergodic average is a constant on each atom of the generating partition of the Pinsker σ-algebra.

We now state the main result of this paper in the following, solving Question 1.2 affirmatively for a large class of sequences.

Theorem 1.5 (Main theorem)*. Let $\{a_k\}$ be a pointwise good sequence satisfying Condition (∗). Suppose that (X, T) is a topological dynamical system with positive topological entropy. Then (X, T) is multivariant mean Li–Yorke chaotic along the sequence $\{a_k\}$; namely, there exists a Cantor subset K of X such that for every integer*

$n \geq 2$ *and pairwise distinct points* $x_1, x_2, \ldots, x_n$ *in* K *we have*

$$\liminf_{N \to \infty} \frac{1}{N} \sum_{k=1}^{N} \max_{1 \leq i < j \leq n} d\left(T^{a_k} x_i, T^{a_k} x_j\right) = 0$$

and

$$\limsup_{N \to \infty} \frac{1}{N} \sum_{k=1}^{N} \min_{1 \leq i < j \leq n} d\left(T^{a_k} x_i, T^{a_k} x_j\right) > 0.$$

The proof of Theorem 1.5 is based on the techniques used in [22]. A main new ingredient in this paper is the convergence argument for sequences given in Sect. 3.

Corollary 1.6 *Let* $\{a_k\}$ *be a pointwise good sequence satisfying Condition* (∗)*. Suppose that* (X, T) *is a topological dynamical system with positive topological entropy. Then* (X, T) *is Li–Yorke chaotic along the sequence* $\{a_k\}$*; namely, there exists a Cantor subset* K *of* X *such that for any* $x, y \in K$ *with* $x \neq y$ *we have*

$$\liminf_{N \to \infty} d(T^{a_k} x, T^{a_k} y) = 0 \quad \textit{and} \quad \limsup_{N \to \infty} d(T^{a_k} x, T^{a_k} y) > 0.$$

As we mentioned previously, the prime sequence and $\{[p(k)]\}$ (where $p(x)$ is a polynomial of real coefficients) are pointwise good and satisfy Condition (∗).

Corollary 1.7 *Any positive entropy system is multivariant mean Li–Yorke chaotic along the prime sequence.*

Corollary 1.8 *Any positive entropy system is multivariant mean Li–Yorke chaotic along the sequence* $\{[p(k)]\}$*, where* $p(x)$ *is a polynomial of real coefficients.*

This paper is organized as follows. In Sect. 2, we review some necessary notions and required properties. In Sect. 3, we study ergodic average along sequences. More precisely, we show that the Pinsker σ-algebra is a characteristic σ-algebra for any sequence satisfying Condition (∗) (Theorem 3.2). Moreover, for any pointwise good sequence satisfying Condition (∗) we have a decomposition for any invariant measure (Theorem 3.4); meanwhile, the limit of ergodic average is a constant on each atom of the generating partition of the Pinsker σ-algebra (Theorem 3.5). In Sect. 4, we prove Theorem 1.5, which holds for all continuous surjective maps as well (Theorem 4.5).

2 Preliminaries

For convenience, our notations will be as close to [22] as possible. For the reader who is not familiar with notions in ergodic theory and dynamical systems, we refer to [12,32].

2.1 Mycielski's theorem

Let X be a compact metric space and $C(X)$ the space of all real-valued continuous functions on X equipped with the supremum norm. For $n \geq 2$, denote $X^n = X \times X \times \cdots \times X$ (n-copies). Set $\Delta_n = \{(x, x, \ldots, x) \in X^n : x \in X\}$ and $\Delta^{(n)} = \{(x_1, x_2, \ldots, x_n) \in X^n : \text{ there exists } 1 \leq i < j \leq n \text{ such that } x_i = x_j\}$. We shall use the following version of Mycielski's theorem.

Theorem 2.1 (Cf. [29, Theorem 1]). *Assume that X is a perfect compact metric space. If for every integer $n \geq 2$, R_n is a dense G_δ subset of X^n, then there exists a dense subset K of X which is a union of countably many Cantor sets such that $K^n \subset R_n \cup \Delta^{(n)}$ holds for all integers $n \geq 2$.*

2.2 Conditional expectation and disintegration

Let $(X, \mathscr{B}, \mu)$ be a probability space and $\mathscr{A}$ a sub-σ-algebra of $\mathscr{B}$. The conditional expectation is a map

$$E(\cdot | \mathscr{A}) : L^1(X, \mathscr{B}, \mu) \to L^1(X, \mathscr{A}, \mu)$$

satisfying the following conditions:

(1) for every $f \in L^1(X, \mathscr{B}, \mu)$, $E(f|\mathscr{A})$ is $\mathscr{A}$-measurable;
(2) if g is $\mathscr{A}$-measurable and $fg \in L^1(X, \mathscr{B}, \mu)$, then $E(fg|\mathscr{A}) = gE(f|\mathscr{A})$;
(3) if $f \in L^p(X, \mathscr{B}, \mu)$ for some $p \geq 1$, then $E(f|\mathscr{A}) \in L^p(X, \mathscr{A}, \mu)$ and

$$\|E(f|\mathscr{A})\|_{L^p} \leq \|f\|_{L^p}.$$

The Martingale theorem is well known (see e.g., [16, Theorem 14.26], [12, Chapter 5.2]).

Theorem 2.2 (Martingale theorem). *Let $(X, \mathscr{B}, \mu)$ be a probability space. Suppose that $\{\mathscr{A}_n\}_{n=1}^{\infty}$ is a decreasing sequence (resp. an increasing sequence) of sub-σ-algebras of $\mathscr{B}$ and $\mathscr{A} = \bigcap_{n \geq 1} \mathscr{A}_n$ (resp. $\mathscr{A} = \bigvee_{n \geq 1} \mathscr{A}_n$). Then for any $f \in L^1(\mu)$,*

$$E(f|\mathscr{A}_n) \to E(f|\mathscr{A})$$

as $n \to \infty$ in $L^1(\mu)$ and μ-almost everywhere.

Let X be a compact metric space. Denote by $\mathscr{B}_X$ the Borel σ-algebra of X and $\mathscr{M}(X)$ the set of all Borel probability measures on X. For $\mu \in \mathscr{M}(X)$, let $\mathscr{B}_\mu$ be the completion of $\mathscr{B}_X$ under the measure μ. Then $(X, \mathscr{B}_\mu, \mu)$ is a Lebesgue space. A finite partition of X is a finite family of pairwise distinct measurable subsets of X whose union is X. If $\{\alpha_i\}_{i \in I}$ is a countable family of finite partitions of X, then we say that $\alpha = \bigvee_{i \in I} \alpha_i$ is a *measurable partition*. The sets $A \in \mathscr{B}_\mu$, which are unions of atoms of α, form a sub-σ-algebra $\mathscr{B}_\mu$, which we denote by $\widehat{\alpha}$ or α if there is no

ambiguity. Every sub-σ-algebra of $\mathscr{B}_\mu$ coincides with a σ-algebra constructed in this way (mod μ).

Let $\mathscr{F}$ be a sub-σ-algebra of $\mathscr{B}_\mu$ and α a measurable partition of X with $\widehat{\alpha} = \mathscr{F}$ (mod μ). Then μ can be disintegrated over $\mathscr{F}$ as

$$\mu = \int_X \mu_x d\mu(x),$$

where $\mu_x \in \mathscr{M}(X)$ and $\mu_x(\alpha(x)) = 1$ for μ-a.e. $x \in X$. The disintegration can be characterized by the properties (2.1) and (2.2) as follows:

for every $f \in L^1(X, \mathscr{B}_X, \mu)$, $f \in L^1(X, \mathscr{B}_X, \mu_x)$ for μ-a.e. $x \in X$, $\qquad$ (2.1)

and the map $x \mapsto \int_X f(y)\, d\mu_x(y)$ is in $L^1(X, \mathscr{F}, \mu)$;

for every $f \in L^1(X, \mathscr{B}_X, \mu)$, $E(f|\mathscr{F})(x) = \int_X f\, d\mu_x$ for μ-a.e. $x \in X$. $\quad$ (2.2)

Then for any $f \in L^1(X, \mathscr{B}_X, \mu)$, one has

$$\int_X \left(\int_X f(y)\, d\mu_x(y) \right) d\mu(x) = \int_X f\, d\mu.$$

2.3 Ergodic theory

By a *measure-preserving system*, we mean a quadruple $(X, \mathscr{B}, \mu, T)$, where $(X, \mathscr{B}, \mu)$ is a probability space and $T \colon X \to X$ is an invertible measure-preserving transformation. A measure-preserving system $(X, \mathscr{B}, \mu, T)$ is called *ergodic* if the only members $B \in \mathscr{B}$ with $T^{-1}B = B$ satisfy $\mu(B) = 0$ or $\mu(B) = 1$.

Let α be a finite partition of X. The measure-theoretic entropy of μ relative to α is denoted by $h_\mu(T, \alpha)$, and the measure-theoretic entropy of μ is defined as

$$h_\mu(T) = \sup_\alpha h_\mu(T, \alpha),$$

where the supremum ranges over all finite partitions of X.

The *Pinsker σ-algebra* of a system $(X, \mathscr{B}, \mu, T)$ is defined as

$$P_\mu(T) = \{A \in \mathscr{B} \colon h_\mu(T, \{A, X \backslash A\}) = 0\}.$$

It is easy to see that $P_\mu(T)$ is T-invariant. The Rokhlin–Sinai theorem identifies the Pinsker σ-algebra as the "remote past" of a generating partition.

Theorem 2.3 ([31]). *For a measure-preserving system $(X, \mathscr{B}, \mu, T)$, there exists a sub-σ-algebra $\mathscr{P}$ of $\mathscr{B}$ such that $T^{-1}\mathscr{P} \subset \mathscr{P}$, $\bigvee_{k=0}^{\infty} T^k \mathscr{P} = \mathscr{B}$ and $\bigcap_{n=0}^{\infty} T^{-k}\mathscr{P} = P_\mu(T)$.*

2.4 Topological dynamics

Let (X, T) be a topological dynamical system. For a point $x \in X$, the *stable set* of x is defined as

$$W^s(X, T) = \left\{ y \in X \colon \lim_{k \to \infty} d(T^k x, T^k y) = 0 \right\},$$

and the *unstable set* of x is defined as

$$W^u(X, T) = \left\{ y \in X \colon \lim_{k \to \infty} d(T^{-k} x, T^{-k} y) = 0 \right\}.$$

Let $\mathcal{M}(X, T)$ (resp. $\mathcal{M}^e(X, T)$) be the collection of all T-invariant (resp. ergodic T-invariant) Borel probability measures on X. For every $\mu \in \mathcal{M}(X, T)$, $(X, \mathcal{B}_X, \mu, T)$ is a measure-preserving system. The variational principle is famous:

$$h_{\text{top}}(X, T) = \sup_{\mu \in \mathcal{M}(X, T)} h_\mu(X, T) = \sup_{\mu \in \mathcal{M}^e(X, T)} h_\mu(X, T),$$

where $h_{\text{top}}(X, T)$ denotes the topological entropy of (X, T).

3 The Pinsker σ-algebra and pointwise good sequences

3.1 Characteristic σ-algebras

Lemma 3.1 *Any strictly increasing sequence of positive integers satisfies Condition* (∗).

Proof Fix a strictly increasing sequence $\{a_k\}$ of positive integers and $L > 0$. It is clear that $|a_i - a_j| \le L$ implies $|i - j| \le L$. For every integer $N > L$, we have

$$\#\{(i, j) \in [1, N]^2 \colon |a_i - a_j| \le L\} \le \#\{(i, j) \in [1, N]^2 \colon |i - j| \le L\} \le 2NL.$$

Thus,

$$\lim_{N \to \infty} \frac{1}{N^2} \#\{(i, j) \in [1, N]^2 \colon |a_i - a_j| \le L\} = 0.$$

Theorem 3.2 *Suppose that $\{a_k\}$ is a sequence of positive integers satisfying Condition* (∗) *and $(X, \mathcal{B}, \mu, T)$ is a measure-preserving system. Then the Pinsker σ-algebra $P_\mu(T)$ is a characteristic σ-algebra for the sequence $\{a_k\}$.*

Proof Fix $f \in L^2(\mu)$ and $\varepsilon > 0$. Choose a partition $\mathcal{P}$ as in the Rohlin–Sinai theorem. By Theorem 2.2, there exist $m > 0$ and $g_m = E(f|T^m \mathcal{P})$ such that

$$\|g_m - f\|_{L^2(\mu)} < \varepsilon.$$

Thus, it holds that

$$\left\|\frac{1}{N}\sum_{k=1}^{N}\left(T^{a_k}f - T^{a_k}g_m\right)\right\|_{L^2(\mu)} \le \frac{1}{N}\sum_{k=1}^{N}\left\|T^{a_k}f - T^{a_k}g_m\right\|_{L^2(\mu)}$$

$$= \frac{1}{N}\sum_{k=1}^{N}\|f - g_m\|_{L^2(\mu)} < \varepsilon$$

and

$$\left\|\frac{1}{N}\sum_{k=1}^{N}\left(T^{a_k}E(g_m|P_\mu(T)) - T^{a_k}E(f|P_\mu(T))\right)\right\|_{L^2(\mu)}$$

$$\le \frac{1}{N}\sum_{k=1}^{N}\|T^{a_k}E(g_m|P_\mu(T)) - T^{a_k}E(f|P_\mu(T))\|_{L^2(\mu)}$$

$$= \frac{1}{N}\sum_{k=1}^{N}\|E(g_m|P_\mu(T)) - E(f|P_\mu(T))\|_{L^2(\mu)}$$

$$= \frac{1}{N}\sum_{k=1}^{N}\|E(g_m - f|P_\mu(T))\|_{L^2(\mu)}$$

$$\le \frac{1}{N}\sum_{k=1}^{N}\|g_m - f\|_{L^2(\mu)} < \varepsilon.$$

By Theorem 2.2 again, there exist $n > 0$ and $h_n = E(g_m|T^{-n}\mathcal{P})$ such that

$$\|h_n - E(g_m|P_\mu(T))\|_{L^2(\mu)} < \varepsilon.$$

It follows that

$$\left\|\frac{1}{N}\sum_{k=1}^{N}\left(T^{a_k}h_n - T^{a_k}E(g_m|P_\mu(T))\right)\right\|_{L^2(\mu)} < \varepsilon.$$

Notice that

$$\left\|\frac{1}{N}\sum_{k=1}^{N}\left(T^{a_k}g_m - T^{a_k}h_n\right)\right\|_{L^2(\mu)}^2 = \int\left|\frac{1}{N}\sum_{k=1}^{N}\left(T^{a_k}g_m - T^{a_k}h_n\right)\right|^2 d\mu$$

$$= \frac{1}{N^2}\sum_{i,j=1}^{N}\int\left(T^{a_i}g_m - T^{a_i}h_n\right)\left(T^{a_j}g_m - T^{a_j}h_n\right)d\mu$$

$$= \frac{1}{N^2}\sum_{i,j=1}^{N}\left(A_{ij} - B_{ij} + C_{ij} - D_{ij}\right),$$

where

$$A_{ij} = \int T^{a_i} g_m \cdot T^{a_j} g_m d\mu, \quad B_{ij} = \int T^{a_i} h_n \cdot T^{a_j} g_m d\mu,$$

$$C_{ij} = \int T^{a_i} h_n \cdot T^{a_j} h_n d\mu, \quad D_{ij} = \int T^{a_i} g_m \cdot T^{a_j} h_n d\mu.$$

Claim 1 *If $a_j - m \geq n + a_i$, then we have $A_{ij} = B_{ij}$ and $C_{ij} = D_{ij}$.*

Proof of Claim 1 First, we have

$$B_{ij} = \int h_n \circ T^{a_i} \cdot g_m \circ T^{a_j} d\mu = \int \left(h_n \cdot g_m \circ T^{a_j - a_i} \right) \circ T^{a_i} d\mu$$

$$= \int h_n \cdot g_m \circ T^{a_j - a_i} d\mu.$$

Recall that $g_m = E(f|T^m \mathscr{P})$. Then g_m is $T^m \mathscr{P}$-measurable and hence $g_m \circ T^{a_j - a_i}$ is $T^{-(a_j - a_i - m)} \mathscr{P}$-measurable. As $a_j - m \geq n + a_i$, $a_j - a_i - m \geq n$, we know that $g_m \circ T^{a_j - a_i}$ is $T^{-n} \mathscr{P}$-measurable. Recall that $h_n = E(g_m|T^{-n} \mathscr{P})$. By (2) of Sect. 2.2, we have

$$B_{ij} = \int E(g_m \cdot g_m \circ T^{a_j - a_i}|T^{-n} \mathscr{P}) d\mu = \int g_m \cdot g_m \circ T^{a_j - a_i} d\mu$$

$$= \int \left(g_m \cdot g_m \circ T^{a_j - a_i} \right) \circ T^{a_i} d\mu = A_{ij}.$$

Now we consider the term C_{ij}:

$$C_{ij} = \int h_n \circ T^{a_i} \cdot h_n \circ T^{a_j} d\mu = \int h_n \cdot h_n \circ T^{a_j - a_i} d\mu.$$

As $a_j - m \geq n + a_i$, $a_j \geq a_i$, we have that $E(g_m|T^{-n} \mathscr{P}) \circ T^{a_j - a_i}$ is $T^{-n} \mathscr{P}$-measurable. Recall that $h_n = E(g_m|T^{-n} \mathscr{P})$. By (2) of Sect. 2.2, we have

$$C_{ij} = \int E\left(g_m \cdot h_n \circ T^{a_j - a_i}|T^{-n} \mathscr{P} \right) d\mu = \int g_m \cdot h_n \circ T^{a_j - a_i} d\mu$$

$$= \int g_m \circ T^{a_i} \cdot h_n \circ T^{a_j} d\mu = D_{ij}.$$

This ends the proof of Claim 1. □

Similarly, we have the following

Claim 2 *If $a_i - m \geq n + a_j$, then we have $A_{ij} = D_{ij}$ and $B_{ij} = C_{ij}$.*

By Hölder's inequality, it is easy to see that

$$\max \left\{ |A_{ij}|, |B_{ij}|, |C_{ij}|, |D_{ij}| \right\} \leq \|g_m\|_{L^2(\mu)}^2.$$

Note that the sequence $\{a_k\}$ satisfies Condition $(*)$. For $n + m - 1$ there exists $N_0 > 0$ such that whenever $N \geq N_0$ it holds that

$$\frac{1}{N^2} \#\{(i, j) \in [1, N]^2 : |a_j - a_i| \leq n + m - 1\} < \frac{\varepsilon^2}{4\|g_m\|^2_{L^2(\mu)}}.$$

Therefore, when $N \geq N_0$ we have

$$\left\| \frac{1}{N} \sum_{k=1}^{N} \left(T^{a_k} g_m - T^{a_k} h_n \right) \right\|^2_{L^2(\mu)}$$

$$= \frac{1}{N^2} \sum_{i,j=1}^{N} \left(A_{ij} - B_{ij} + C_{ij} - D_{ij} \right)$$

$$= \frac{1}{N^2} \sum_{\substack{i,j=1 \\ |a_j - a_i| \leq n+m-1}}^{N} \left(A_{ij} - B_{ij} + C_{ij} - D_{ij} \right)$$

$$\leq 4\|g_m\|^2_{L^2(\mu)} \cdot \frac{1}{N^2} \#\{(i, j) \in [1, N]^2 : |a_j - a_i| \leq n + m - 1\} < \varepsilon^2.$$

To sum up, for $N \geq N_0$ we have

$$\left\| \frac{1}{N} \sum_{k=1}^{N} \left(T^{a_k} f - T^{a_k} E(f|P_\mu(T)) \right) \right\|_{L^2(\mu)}$$

$$\leq \left\| \frac{1}{N} \sum_{k=1}^{N} (T^{a_k} f - T^{a_k} g_m) \right\|_{L^2(\mu)} + \left\| \frac{1}{N} \sum_{k=2}^{N} \left(T^{a_k} g_m - T^{a_k} h_n \right) \right\|_{L^2(\mu)}$$

$$+ \left\| \frac{1}{N} \sum_{k=1}^{N} \left(T^{a_k} h_n - T^{a_k} E(g_m|P_\mu(T)) \right) \right\|_{L^2(\mu)}$$

$$+ \left\| \frac{1}{N} \sum_{k=1}^{N} \left(T^{a_k} E(g_m|P_\mu(T)) - T^{a_k} E(f|P_\mu(T)) \right) \right\|_{L^2(\mu)}$$

$$< \varepsilon + \varepsilon + \varepsilon + \varepsilon = 4\varepsilon.$$

This ends the proof. □

Remark 3.3 We also refer to [10] for a variant of Theorem 3.2 which deals with polynomial iterates and pointwise convergence.

3.2 Good sequences for pointwise convergence

Similar to the ergodic decomposition theorem, we have the following decomposition of an invariant measure with respect to any pointwise good sequence.

Theorem 3.4 *Let $\{a_k\}$ be a pointwise good sequence of positive integers. Suppose that (X, T) is a topological dynamical system and $\mu \in M(X, T)$. Then there exists a disintegration of μ,*

$$\mu = \int \tau_x d\mu(x),$$

in the sense that there exists a Borel subset X_0 of X with $\mu(X_0) = 1$ such that for any $x \in X_0$ and $f \in C(X)$, it holds that

$$\lim_{N \to \infty} \frac{1}{N} \sum_{k=1}^{N} f(T^{a_k} x) = \int f d\tau_x$$

and

$$\int \int f d\tau_x d\mu(x) = \int f d\mu.$$

Proof Fix a countable dense subset $\{f_n\}$ of $C(X)$. As $\{a_k\}$ is pointwise good, there exists a Borel subset X_0 of X with $\mu(X_0) = 1$ such that for every $x \in X_0$ and f_n, we have

$$\lim_{N \to \infty} \frac{1}{N} \sum_{k=1}^{N} f_n(T^{a_k} x) = \overline{f}_n(x).$$

Fix $x \in X_0$. Define

$$L_x \colon C(X) \to \mathbb{R}, \quad f \mapsto \lim_{N \to \infty} \frac{1}{N} \sum_{k=1}^{N} f(T^{a_k} x).$$

Since $\{f_n\}$ is dense in $C(X)$ endowed with the supremum norm, L_x is well defined. Moreover, L_x is a positive linear functional with $L_x(1) = 1$. By the Riesz Representation Theorem, there exists $\tau_x \in M(X)$ such that for any $f \in C(X)$,

$$\lim_{N \to \infty} \frac{1}{N} \sum_{k=1}^{N} f(T^{a_k} x) = \int f d\tau_x.$$

By Lebesgue's dominated convergence theorem,

$$\int \left(\int f d\tau_x \right) d\mu(x) = \int \left(\lim_{N \to \infty} \frac{1}{N} \sum_{k=1}^{N} f(T^{a_k} x) \right) d\mu(x)$$

$$= \lim_{N \to \infty} \frac{1}{N} \sum_{k=1}^{N} \int f(T^{a_k} x) d\mu(x) = \int f d\mu.$$

$\square$

Theorem 3.5 *Let $\{a_k\}$ be a pointwise good sequence, (X, T) a topological dynamical system and $\mu \in M(X, T)$. If the Pinsker σ-algebra $P_\mu(T)$ is a characteristic σ-algebra for the sequence $\{a_k\}$, then there exists a Borel subset X_1 of X with $\mu(X_1) = 1$ such that for any $x \in X_1$ and $f \in C(X)$, it holds that*

$$\int f d\tau_x = \int \left(\int f d\tau_y \right) d\mu_x(y),$$

where

$$\mu = \int \tau_x d\mu(x) \quad and \quad \mu = \int \mu_x d\mu(x)$$

are the disintegrations of μ as in Theorem 3.4 and over $P_\mu(T)$ respectively.

Proof As $\{a_k\}$ is pointwise good, for every $f \in C(X)$ there exist $\bar{f}$ and f^* in $L^2(\mu)$ such that

$$\frac{1}{N} \sum_{k=1}^{N} T^{a_k} f \to \bar{f} \quad \text{and} \quad \frac{1}{N} \sum_{k=1}^{N} T^{a_k} E(f|P_\mu(T)) \to f^*$$

μ-almost everywhere. As the Pinsker σ-algebra $P_\mu(T)$ is a characteristic σ-algebra for the sequence $\{a_k\}$, by the definition we have

$$\frac{1}{N} \sum_{k=1}^{N} T^{a_k} f - \frac{1}{N} \sum_{k=1}^{N} T^{a_k} E(f|P_\mu(T)) \to 0$$

as $N \to \infty$ in $L^2(\mu)$. So there exists a strictly increasing sequence $\{N_i\}$ of positive integers such that

$$\frac{1}{N_i} \sum_{k=1}^{N_i} T^{a_k} f - \frac{1}{N_i} \sum_{k=1}^{N_i} T^{a_k} E(f|P_\mu(T)) \to 0$$

as $i \to \infty$ μ-almost everywhere. Hence $\bar{f}(x) = f^*(x)$ for μ-a.e. $x \in X$. Clearly, f^* is $P_\mu(T)$-measurable, so is $\bar{f}$. Let

$$\mu = \int \tau_x d\mu(x)$$

be the disintegration of μ as in Theorem 3.4. Then $\bar{f}(x) = \int f d\tau_x$ for μ-a.e. $x \in X$. Let

$$\mu = \int \mu_x d\mu(x)$$

be the disintegration of μ over the Pinsker σ-algebra $P_\mu(T)$. As $\bar{f}$ is $P_\mu(T)$-measurable, by (2.2) we have $\bar{f}(x) = \int \bar{f} d\mu_x$ for μ-a.e. $x \in X$.

Now fix a countable dense subset $\{f_n\}$ of $C(X)$. By the above discussion, there exists a Borel subset X_1 of X_0 with $\mu(X_1) = 1$ such that for every $x \in X_1$ and $n \in \mathbb{N}$,

$$\int f_n d\tau_x = \int \left(\int f_n d\tau_y \right) d\mu_x(y). \tag{3.1}$$

As any function in $C(X)$ can be uniformly approximated by $\{f_n\}$, (3.1) holds for all $f \in C(X)$. $\square$

Remark 3.6 Let α be a measurable partition generating $P_\mu(T)$. Then $\mu_x(\alpha(x)) = 1$ for μ-a.e. $x \in X$, where $\alpha(x)$ is the atom of α containing x. If a function f is $P_\mu(T)$-measurable, then we know from (2.2) that for μ-a.e. $x \in X$, f is a constant almost everywhere with respect to μ_x on $\alpha(x)$. By Theorem 3.5, if a sequence is pointwise good and satisfies Condition $(*)$, then the limit of the ergodic average along this sequence is a constant on any atom of the generating partition of $P_\mu(T)$.

If $(X, \mathscr{B}, \mu, T)$ is a Kolmogorov system, then the Pinsker factor is trivial. This yields the following.

Corollary 3.7 *Let $\{a_k\}$ be a pointwise good sequence satisfying Condition $(*)$ and $(X, \mathscr{B}, \mu, T)$ a Kolmogorov system. Then for every $f \in L^2(\mu)$,*

$$\lim_{N \to \infty} \frac{1}{N} \sum_{k=1}^{N} f(T^{a_k} x) = \int f d\mu$$

holds for μ-a.e. $x \in X$.

3.3 Examples of pointwise good sequences

There are many results on pointwise good/bad sequences, we refer the reader to [9,30] for recent work on this topic. Here let us list some sequences of positive integers which are both pointwise good and satisfying Condition $(*)$.

Theorem 3.8 ([5, Theorem 1]). *The sequence of prime numbers is pointwise good.*

Theorem 3.9 ([6, Theorem 2]). *For any polynomial*

$$p(x) = b_0 + b_1 x + \cdots + b_m x^m,$$

where $m \geq 1$, $b_0, \cdots, b_m \in \mathbb{R}$, and $b_m > 0$, the sequence $\{[p(k)]\}$ is pointwise good.

Remark 3.10 There are very good random sequences of integers that are not too sparse, including ones that grow faster than any polynomial (see [7, Proposition 8.2]).

Furthermore, a lot of sequences generated by logarithmic–exponential subpolynomials are pointwise good (for details see [4, Theorems 3.4, 3.5, 3.8]). In particular, one has the following:

Theorem 3.11 *For every non-integer number $r > 0$, the sequence $\{[k^r]\}$ is very good.*

Theorem 3.12 ([6, Theorem in Appendix]). *Let $(X, \mathscr{B}, \mu, T)$ be an ergodic measure-preserving system and $A \in \mathscr{B}$ with $\mu(A) > 0$. Then for μ-a.e. $x \in X$, the sequence $\{n \in \mathbb{N}: T^n x \in A\}$ is pointwise good.*

Theorem 3.13 ([8, Theorem 1]). *Let (X, T) be a minimal equicontinuous system with the unique invariant measure μ. For every $x \in X$ and every subset A of X with $\mu(A) > 0$ and $\mu(\partial A) = 0$ (where ∂A is the boundary of A), the sequence $\{n \in \mathbb{N}: T^n x \in A\}$ is pointwise good.*

Theorem 3.14 ([27, Theorem 3]). *Return time sequences for dynamical systems which are abelian extensions of translations on compact abelian groups are pointwise good. Moreover, the Morse sequence is very good.*

Remark 3.15 Similar to the proof of Lemma 3.1, it is not hard to check that the sequences in Theorems 3.9 and 3.11 satisfy Condition ($*$).

4 Proof of the main theorem

4.1 Dynamical systems with positive entropy

Let (X, T) be a topological dynamical system, $\mu \in \mathscr{M}(X, T)$, and $\mathscr{B}_\mu$ the completion of $\mathscr{B}_X$ under μ. Then $(X, \mathscr{B}_\mu, \mu, T)$ is a Lebesgue system. Let $P_\mu(T)$ be the Pinsker σ-algebra of $(X, \mathscr{B}_\mu, \mu, T)$ and

$$\mu = \int_X \mu_x d\mu(x)$$

the disintegration of μ over $P_\mu(T)$. Under the above settings, we first state the following lemma.

Lemma 4.1 ([22, Lemma 3.1]). *If $\mu \in \mathscr{M}^e(X, T)$ and $h_\mu(T) > 0$, then for μ-a.e. $x \in X$,*

$$\overline{W^s(x, T) \cap \text{supp}(\mu_x)} = \text{supp}(\mu_x) \quad \text{and} \quad \overline{W^u(x, T) \cap \text{supp}(\mu_x)} = \text{supp}(\mu_x).$$

By Theorem 3.2 and the variational principle, our main result (Theorem 1.5) follows from Theorem 4.2:

Theorem 4.2 *Let (X, T) be a topological dynamical system and $\mu \in \mathscr{M}^e(X, T)$ with $h_\mu(T) > 0$. If $\{a_k\}$ is a pointwise good sequence and the Pinsker σ-algebra $P_\mu(T)$ is a characteristic σ-algebra for the sequence $\{a_k\}$, then for μ-a.e. $x \in X$, there exists*

a Cantor subset $K_x \subset \overline{W^s(x, T)} \cap \overline{W^u(x, T)}$ satisfying that for every integer $n \geq 2$ and pairwise distinct points $x_1, x_2, \ldots, x_n$ in K_x we have

$$\liminf_{N \to \infty} \frac{1}{N} \sum_{k=1}^{N} \max_{1 \leq i < j \leq n} d\left(T^{a_k} x_i, T^{a_k} x_j\right) = 0$$

and

$$\limsup_{N \to \infty} \frac{1}{N} \sum_{k=1}^{N} \min_{1 \leq i < j \leq n} d\left(T^{a_k} x_i, T^{a_k} x_j\right) \geq \eta_{x,n},$$

where $\eta_{x,n}$ is a positive constant depending only on x and n.

Proof Let

$$\mu = \int_X \mu_x d\mu(x)$$

be the disintegration of μ over $P_\mu(T)$. By Lemma 4.1, there exists a Borel subset X_1 of X with $\mu(X_1) = 1$ such that for any $x \in X_1$,

$$\overline{W^s(x, T)} \cap \text{supp}(\mu_x) = \text{supp}(\mu_x) \quad \text{and} \quad \overline{W^u(x, T)} \cap \text{supp}(\mu_x) = \text{supp}(\mu_x).$$

Fix an integer $n \geq 2$. Set

$$MP_n(X, T) = \left\{(x_1, x_2, \ldots, x_n) \in X^n : \liminf_{N \to \infty} \frac{1}{N} \sum_{k=1}^{N} \max_{1 \leq i < j \leq n} d(T^{a_k} x_i, T^{a_k} x_j) = 0\right\}.$$

By a similar argument as in [22, Lemma 2.4], $MP_n(X, T)$ is a G_δ subset of X^n. It is clear that $(x_1, x_2, \ldots, x_n) \in MP_n(X, T)$ for any $x_1, x_2, \ldots, x_n$ in $W^s(x, T)$. Thus, for each $x \in X_1$, $MP_n(X, T) \cap \text{supp}(\mu_x)^n$ is a dense G_δ subset of $\text{supp}(\mu_x)^n$.

Define a measure λ_n on $(X^n, T^{(n)})$ by

$$\lambda_n = \int_X \mu_x^{(n)} d\mu(x),$$

where $\mu_x^{(n)} = \mu_x \times \mu_x \times \cdots \times \mu_x$ (n-times) and $T^{(n)} = T \times T \times \cdots \times T$ (n-times). As μ is ergodic and has positive entropy, μ_x is non-atomic for μ-a.e. $x \in X$ and λ_n is a $T^{(n)}$-invariant ergodic measure on X^n (see [21, Lemma 5.4]). Since μ_x is non-atomic for μ-a.e. $x \in X$, by the Fubini theorem, $\lambda_n(\Delta^{(n)}) = 0$. By Theorem 3.4, there exists a disintegration of λ_n with respect to the sequence $\{a_k\}$:

$$\lambda_n = \int \tau_{(x_1, x_2, \ldots, x_n)} d\lambda_n(x_1, x_2, \ldots, x_n).$$

Consider the continuous function

$$f\colon X^n \to \mathbb{R}, \quad (x_1, x_2, \ldots, x_n) \mapsto \min\{d(x_i, x_j)\colon 1 \le i < j \le n\}.$$

By Theorem 3.4, we have

$$\lim_{N \to \infty} \frac{1}{N} \sum_{k=1}^{N} f\left((T^{(n)})^{a_k}(x_1, x_2, \ldots, x_n)\right) = \lim_{N \to \infty} \frac{1}{N} \sum_{k=1}^{N} \min_{1 \le i < j \le n} d(T^{a_k} x_i, T^{a_k} x_j)$$

$$= \int f \, d\tau_{(x_1, x_2, \ldots, x_n)}.$$

Since $\lambda_n(\Delta^{(n)}) = 0$, $\tau_{(x_1, x_2, \ldots, x_n)}(\Delta^{(n)}) = 0$ for λ_n-a.e. $(x_1, x_2, \ldots, x_n)$. Note that for any $(x_1, x_2, \ldots, x_n) \notin \Delta^{(n)}$, $f(x_1, x_2, \ldots, x_n) > 0$. Thus, we have $\int f \, d\tau_{(x_1, x_2, \ldots, x_n)} > 0$ for λ_n-a.e. $(x_1, x_2, \ldots, x_n)$.

Let $\pi\colon X^n \to X$, $(x_1, x_2, \ldots, x_n) \mapsto x_1$ be the canonical projection to the first coordinate. By [17, Theorem 4] (see also [23, Lemma 4.2]), we know that the Pinsker σ-algebra of $(X^n, \lambda_n, T^{(n)})$ equals $\pi^{-1}(P_\mu(T))$ (mod λ_n). So

$$\lambda_n = \int_X \mu_x^{(n)} d\mu(x)$$

can be also regarded as the disintegration of λ over the Pinsker σ-algebra of $(X^n, \lambda_n, T^{(n)})$. By Theorem 3.5, we have that for μ-a.e. $x \in X$, $\int f \, d\tau_{(x_1, x_2, \ldots, x_n)}$ is a constant for $\mu_x^{(n)}$-a.e. $(x_1, x_2, \ldots, x_n) \in X^n$. More specifically, there exists a Borel subset X_2 of X with $\mu(X_2) = 1$ such that for any $x \in X_2$,

$$\lim_{N \to \infty} \frac{1}{N} \sum_{k=1}^{N} \min_{1 \le i < j \le n} d(T^{a_k} x_i, T^{a_k} x_j) = \eta_{x,n}$$

for $\mu_x^{(n)}$-a.e. $(x_1, x_2, \ldots, x_n) \in X^n$ and some constant $\eta_{x,n} > 0$. Put

$$D_{n,\eta}(X, T) = \left\{ (x_1, x_2, \ldots, x_n) \in X^n \colon \limsup_{N \to \infty} \frac{1}{N} \sum_{k=1}^{N} \min_{1 \le i < j \le n} d(T^{a_k} x_i, T^{a_k} x_j) \ge \eta \right\}.$$

Similar to [22, Lemma 2.4], $D_{n,\eta}(X, T)$ is a G_δ subset of X^n. Then for each $x \in X_2$, $D_{n,\eta_{x,n}}(X, T) \cap \mathrm{supp}(\mu_x)^n$ is a dense G_δ subset of $\mathrm{supp}(\mu_x)^n$.

As $n \ge 2$ is arbitrary, there exists a Borel subset X_0 of X with $\mu(X_0) = 1$ such that for every integer $n \ge 2$ and $x \in X_0$, $MP_n(X, T) \cap D_{n,\eta_{x,n}}(X, T) \cap \mathrm{supp}(\mu_x)^n$ is a dense G_δ subset of $\mathrm{supp}(\mu_x)^n$. Since μ_x is non-atomic for μ-a.e. $x \in X$, we can also require that μ_x is non-atomic for every $x \in X_0$. Then for each $x \in X_0$, $\mathrm{supp}(\mu_x)$ is a perfect closed subset of X. By Mycielski's theorem, there exists a Cantor subset K_x of $\mathrm{supp}(\mu_x)$ such that for every integer $n \ge 2$,

$$K_x^n \subset (MP_n(X, T) \cap D_{n,\eta_{x,n}}(X, T)) \cup \Delta^{(n)}.$$

Thus, K_x is as required.   □

Recall that the Pinsker factor of a Kolmogorov system is trivial. Theorem 4.2 yields the following.

Corollary 4.3 *Let $\{a_k\}$ be a pointwise good sequence satisfying Condition $(*)$. Suppose that (X, T) is a topological dynamical system and there is an invariant measure μ such that $\mathrm{supp}(\mu) = X$ and $(X, \mathscr{B}, \mu, T)$ is a Kolmogorov system. Then there exists a Cantor subset K of X and a sequence $\{\eta_n\}$ of positive numbers such that for every integer $n \geq 2$ and pairwise distinct points $x_1, x_2, \ldots, x_n$ in K, it holds that*

$$\liminf_{N \to \infty} \frac{1}{N} \sum_{k=1}^{N} \max_{1 \leq i < j \leq n} d\left(T^{a_k} x_i, T^{a_k} x_j\right) = 0$$

and

$$\limsup_{N \to \infty} \frac{1}{N} \sum_{k=1}^{N} \min_{1 \leq i < j \leq n} d\left(T^{a_k} x_i, T^{a_k} x_j\right) \geq \eta_n.$$

Remark 4.4 One of the referees asked the following interesting question: Does there exist some subsequences which are bad for mean Li–Yorke chaos in dynamical systems with positive topological entropy? In particular, how about the sequence $\{2^n\}$?

It is proved by Bellow in [1] that very fast growing sequences, known as lacunary sequences, are bad for mean ergodic theorems. Recall that a positive sequence $\{a_n\}$ is called *lacunary* if there exists $\lambda > 1$ such that $\frac{a_{n+1}}{a_n} \geq \lambda$ for all $n \geq 1$. Clearly the sequence $\{2^n\}$ is lacunary. Note that the proof of Theorem 4.2 strongly depends on the sequences which are good for pointwise ergodic theorems. To solve the above question, new technique should be involved.

4.2 Non-invertible case

In this subsection, we generalize Theorem 4.2 to the non-invertible case. Let (X, T) be a non-invertible system, meaning that X is a compact metric space and the map $T : X \to X$ is continuous surjective but not one-to-one.

Theorem 4.5 *Let (X, T) be a non-invertible system and $\mu \in \mathscr{M}^e(X, T)$ with $h_\mu(T) > 0$. If $\{a_k\}$ is a pointwise good sequence and the Pinsker σ-algebra $P_\mu(T)$ is a characteristic σ-algebra for the sequence $\{a_k\}$, then for μ-a.e. $x \in X$, there exists a Cantor subset $K_x \subset \overline{W^s(x, T)}$ satisfying that for every integer $n \geq 2$ and pairwise distinct points $x_1, x_2, \ldots, x_n$ in K_x we have*

$$\liminf_{N \to \infty} \frac{1}{N} \sum_{k=1}^{N} \max_{1 \leq i < j \leq n} d\left(T^{a_k} x_i, T^{a_k} x_j\right) = 0$$

and

$$\limsup_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} \min_{1\le i<j\le n} d\left(T^{a_k}x_i, T^{a_k}x_j\right) \ge \eta_{x,n},$$

where $\eta_{x,n}$ is a positive constant depending only on x and n.

Proof We consider the natural extension $(\widetilde{X}, \widetilde{T})$ of (X, T); that is,

$$\widetilde{X} = \left\{(x_1, x_2, \cdots) \in X^{\mathbb{N}} : Tx_{i+1} = x_i,\ i \in \mathbb{N}\right\}$$

and $\widetilde{T}: \widetilde{X} \to \widetilde{X}$, $(x_1, x_2, \cdots) \mapsto (Tx_1, x_1, x_2, \cdots)$ is the shift homeomorphism. A compatible metric $\widetilde{d}$ on $\widetilde{X}$ is defined by

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

Let $\pi: \widetilde{X} \to X$, $(x_1, x_2, \ldots) \mapsto x_1$ be the projection to the first coordinate. Then $\pi: (\widetilde{X}, \widetilde{T}) \to (X, T)$ is a factor map. As μ is an ergodic invariant measure of (X, T), there exists an ergodic invariant measure $\widetilde{\mu}$ of $(\widetilde{X}, \widetilde{T})$ such that $\pi(\widetilde{\mu}) = \mu$. Clearly, $h_{\widetilde{\mu}}(\widetilde{T}) \ge h_\mu(T) > 0$. By Theorem 4.2, there is a Borel subset $\widetilde{X}_0$ of $\widetilde{X}$ with $\widetilde{\mu}(\widetilde{X}_0) = 1$ such that for each $\tilde{x} \in \widetilde{X}_0$, there exists a Cantor subset $K_{\tilde{x}} \subset \overline{W^s(\tilde{x}, \widetilde{T})}$ satisfying that for every integer $n \ge 2$ and pairwise distinct points $\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_n$ in $K_{\tilde{x}}$ one has

$$\liminf_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} \max_{1\le i<j\le n} \widetilde{d}\left(\widetilde{T}^{a_k}\tilde{x}_i, \widetilde{T}^{a_k}\tilde{x}_j\right) = 0$$

and

$$\limsup_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} \min_{1\le i<j\le n} \widetilde{d}\left(\widetilde{T}^{a_k}\tilde{x}_i, \widetilde{T}^{a_k}\tilde{x}_j\right) \ge \eta_{\tilde{x},n}.$$

Let $X_0 = \pi(\widetilde{X}_0)$. Then X_0 is a μ-measurable set with $\mu(X_0) = 1$. For any $x \in X_0$, there exists $\tilde{x} \in \widetilde{X}_0$ such that $\pi(\tilde{x}) = x$. Let $K_x = \pi(K_{\tilde{x}})$. Then $K_x \subset \overline{W^s(x, T)}$. Similar to [22, Lemma 3.7], for every integer $n \ge 2$ there is a positive constant $\eta_{x,n}$ such that for any pairwise distinct points $x_1, x_2, \ldots, x_n$ in K_x one has

$$\liminf_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} \max_{1\le i<j\le n} d\left(T^{a_k}x_i, T^{a_k}x_j\right) = 0$$

and

$$\limsup_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} \min_{1\le i<j\le n} d\left(T^{a_k}x_i, T^{a_k}x_j\right) \ge \eta_{x,n}.$$

So K_x is as required. □

References

1. Bellow, A.: On "bad universal" sequences in ergodic theory (ii). In: Belley, J., Dubois, J., Morales, P. (eds.) Measure Theory and its Applications. Lecture Notes in Mathematics, vol. 1033, pp. 74–78. Springer, Berlin (1983)
2. Blanchard, F., Huang, W.: Entropy sets, weakly mixing sets and entropy capacity. Discrete Contin. Dyn. Syst. **20**(2), 275–311 (2008)
3. Blanchard, F., Glasner, E., Kolyada, S., Maass, A.: On Li–Yorke pairs. J. Reine Angew. Math. **547**, 51–68 (2002)
4. Boshernitzan, M., Kolesnik, G., Quas, A., Wierdl, M.: Ergodic averaging sequences. J. Anal. Math. **95**, 63–103 (2005)
5. Bourgain, J.: An Approach to Pointwise Ergodic Theorems. In: Lindenstrauss, J., Milman, V.D. (eds.) Geometric Aspects of Functional Analysis (1986/87), Lecture Notes of Mathematics, vol. 1317, pp. 204–223. Springer, Berlin (1988)
6. Bourgain, J.: Pointwise ergodic theorems for arithmetic sets. In: Voisin, C. (ed.) With an Appendix by the Author, Harry Furstenberg, Yitzhak Katznelson and Donald S. Ornstein. Institut des Hautes Études Scientifiques, vol. 69, pp. 5–45 (1989)
7. Bourgain, J.: On the maximal ergodic theorem for certain subsets of the integers. Isr. J. Math. **61**(1), 39–72 (1988)
8. Brunel, A., Keane, M.: Ergodic theorems for operator sequences. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **12**, 231–240 (1969)
9. del Junco, A.: Ergodic theorems. In: Mathematics of Complexity and Dynamical Systems, vol. 1–3, pp. 241–263. Springer, New York (2012)
10. Derrien, J., Lesigne, E.: Un théorème ergodique polynômial ponctuel pour les endomorphismes exacts et les K-systèmes. Annales de l'I.H.P. Probabilités et statistiques **32**(6), 765–778 (1996)
11. Downarowicz, T.: Positive topological entropy implies chaos DC2. Proc. Am. Math. Soc. **142**, 137–149 (2014)
12. Einsiedler, M., Ward, T.: Ergodic Theory with a View Towards Number Theory, Graduate Texts in Mathematics, vol. 259. Springer, London (2011)
13. Fang, C., Huang, W., Yi, Y., Zhang, P.: Dimension of stable sets and scrambled sets in positive finite entropy systems. Ergod. Theory Dyn. Syst. **32**(2), 599–628 (2012)
14. Furstenberg, H., Weiss, B.: A mean ergodic theorem for $(1/N) \sum_{n=1}^{N} f(T^n x)g(T^{n^2}x)$. In: Bergelson, V., March P., Rosenblatt, J.M. (eds.) Convergence in Ergodic Theory and Probability (Columbus, 1993), pp. 193–227. Ohio State Univ. Math. Res. Inst. Publ., 5, de Gruyter, Berlin (1996)
15. Garcia-Ramos, F., Jin, L.: Mean proximality and mean Li-Yorke chaos. Proc. Am. Math. Soc. **145**(7), 2959–2969 (2017)

16. Glasner, E.: Ergodic Theory via Joinings, Mathematical Surveys and Monographs, vol. 101. American Mathematical Society, Providence (2003)
17. Glasner, E., Thouvenot, J.-P., Weiss, B.: Entropy theory without a past. Ergod. Theory Dyn. Syst. **20**(5), 1355–1370 (2000)
18. Huang, W.: Stable sets and ε-stable sets in positive-entropy systems. Commun. Math. Phys. **279**(2), 535–557 (2008)
19. Huang, W., Jin, L.: Stable sets and mean Li-Yorke chaos in positive entropy actions of bi-orderable amenable groups. Ergod. Theory Dyn. Syst. **36**(8), 2482–2497 (2016)
20. Huang, W., Ye, X.: Devaney's chaos or 2-scattering implies Li–Yorke's chaos. Topol. Appl. **117**(3), 259–272 (2002)
21. Huang, W., Ye, X.: A local variational relation and applications. Isr. J. Math. **151**, 237–280 (2006)
22. Huang, W., Li, J., Ye, X.: Stable sets and mean Li–Yorke chaos in positive entropy systems. J. Funct. Anal. **266**, 3377–3394 (2014)
23. Huang, W., Xu, L., Yi, Y.: Asymptotic pairs, stable sets and chaos in positive entropy systems. J. Funct. Anal. **268**(4), 824–846 (2015)
24. Iwanik, A.: Independence and scrambled sets for chaotic mappings. In: Rassias, G.M. (ed.) The Mathematical Heritage of C. F. Gauss, pp. 372–378 . World Scientific Publishing, River Edge (1991)
25. Kerr, D., Li, H.: Independence in topological and C*-dynamics. Math. Ann. **338**(4), 869–926 (2007)
26. Kerr, D., Li, H.: Combinatorial independence and sofic entropy. Commun. Math. Stat. **1**, 213–257 (2013)
27. Lesigne, E.: Ergodic theorem along a return time sequence. In: Richter, U.K., Walstatt, V. (eds.) Ergodic Theory and Related Topics, III (Güstrow, 1990), Lecture Notes in Mathematics, vol. 1514, pp. 146–152. Springer, Berlin (1992)
28. Li, J., Ye, X.: Recent development of chaos theory in topological dynamics. Acta Math. Sin. **32**(1), 83–114 (2016)
29. Mycielski, J.: Independent sets in topological algebras. Fundam. Math. **55**, 139–147 (1964)
30. Petersen, K., Salama, I.: Ergodic Theory and its Connections with Harmonic Analysis. London Mathematical Society Lecture Note Series, vol. 205. Cambridge University Press, Cambridge (1995)
31. Rokhlin, V.A., Sinai, Y.G.: The construction and properties of invariant measurable partitions. Dokl. Akad. Nauk SSSR **141**(5), 1038–1041 (1961)
32. Walters, P.: An Introduction to Ergodic Theory. Graduate Texts in Mathematics, vol. 79. Springer, Berlin (1982)