

Weakly mixing proximal topological models for ergodic systems and applications

by

Zhengxing Lian, Song Shao and Xiangdong Ye (Hefei)

Abstract. It is shown that every non-periodic ergodic system has two topologically weakly mixing, fully supported models: one is non-minimal but has a dense set of minimal points, and the other one is proximal. Also, for a given Kakutani–Rokhlin tower with relatively prime column heights, it is demonstrated how to get a new tall Kakutani–Rokhlin tower with the same property, which can be used in Weiss’s proof of Jewett–Krieger’s theorem and in the proofs of our theorems. Applications of the results are given.

1. Introduction. A *measurable dynamical system* is a quadruple (X, \mathcal{X}, μ, T) , where (X, \mathcal{X}, μ) is a Lebesgue probability space and $T : X \rightarrow X$ is an invertible measure preserving transformation. A *topological dynamical system* is a pair (X, T) , where X is a compact metric space and $T : X \rightarrow X$ is a homeomorphism.

Let (X, \mathcal{X}, μ, T) be an ergodic measurable dynamical system. We say that $(\hat{X}, \hat{\mathcal{X}}, \hat{\mu}, \hat{T})$ is a *topological model* (or just a *model*) for (X, \mathcal{X}, μ, T) if (\hat{X}, \hat{T}) is a topological dynamical system, $\hat{\mu}$ is an invariant Borel probability measure on \hat{X} and the systems (X, \mathcal{X}, μ, T) and $(\hat{X}, \hat{\mathcal{X}}, \hat{\mu}, \hat{T})$ are measure-theoretically isomorphic.

The theory of topological models is an important part of the theory of dynamical systems and has many applications. The well known Jewett–Krieger theorem asserts that every non-periodic ergodic system has a topological model which is strictly ergodic. Lehrer [9] showed that we can further require the model to be topologically (strongly) mixing. We refer to [2, 3, 13, 15] for surveys and nice results on this topic. We note that topological models can also be used to obtain the pointwise convergence of non-conventional ergodic averages [6].

2010 *Mathematics Subject Classification*: Primary 37B05; Secondary 37A05.

Key words and phrases: topological model, weakly mixing, proximal, minimal point.

Received 10 June 2015; revised 20 February 2016.

Published online 26 August 2016.

We mention that the models obtained above are minimal. In this paper we study non-minimal models for a given ergodic system, and obtain their applications. Here are our main results. Note that an ergodic system is non-periodic if it has no atom.

THEOREM 1.1.

- (1) *Every non-periodic ergodic system has a topological model which is a non-minimal topologically weakly mixing system with a fully supported invariant measure and a dense set of minimal points.*
- (2) *Every non-periodic ergodic system has a topological model which is a topologically weakly mixing system with a fully supported invariant measure and a unique fixed point as its only minimal point.*

Note that a topological system (X, T) with a unique fixed point as its only minimal point is *proximal*, i.e. $\inf_n d(T^n x, T^n y) = 0$ for all $x, y \in X$. Hence Theorem 1.1(2) means that every non-periodic ergodic system has a topological weakly mixing and proximal model with a fully supported invariant measure.

In Weiss's new proof of the Jewett–Krieger theorem ([12, Theorem 2] and [15, Chapter 4]) and in the proof of Weiss' theorem on the doubly minimal model [14, Theorem 5], one needs a technical lemma to get a new and tall enough Kakutani–Rokhlin tower from a given one. When one uses this lemma, one hopes that the column heights of the new Kakutani–Rokhlin tower are relatively prime, which is not guaranteed in Weiss's lemma. Hence in Weiss's new proof of the Jewett–Krieger theorem, one first assumes that the system has no rational spectrum, in which case automatically the column heights of every Kakutani–Rokhlin tower are relatively prime. Then one deals with other cases.

In this paper, we avoid this by showing that for a given Kakutani–Rokhlin tower with relatively prime column heights, one can get a new and tall enough Kakutani–Rokhlin tower with the same property (Lemma 3.2). This lemma will be used in the proofs of our theorems and we believe that it will be useful in other settings.

We find two applications of our results. One gives an affirmative answer to a question in [10] by showing that if (X, T) is a topological dynamical system and $(M(X), T_M)$ is the induced system on the probability space, then the density of minimal points of $(M(X), T_M)$ does not imply (X, T) has the same property. The other one concerns the existence of a proximal topological K-system. A topological analogue of K-systems, called a topological K-system, was studied in [8]. In [5] the authors constructed a proximal topological K-system which is weakly mixing. Using Theorem 1.1, we can get a lot of such examples which are strongly mixing, i.e. there exist strongly mixing proximal topological K-systems.

2. Preliminaries. In this section we recall some notions to be used later.

2.1. Measurable systems. A measurable system is *ergodic* if all T -invariant sets have measure 0 or 1. For an ergodic system, either the space X consists of a finite set of points on which μ is equidistributed, or the measure μ is atomless. In the former case the system is called *periodic*, and *non-periodic* in the latter.

A *homomorphism* from (X, \mathcal{X}, μ, T) to a system (Y, \mathcal{Y}, ν, S) is a measurable map $\pi : X_0 \rightarrow Y_0$, where X_0 is a T -invariant subset of X and Y_0 is an S -invariant subset of Y , both of full measure, such that $\pi_*\mu = \nu \circ \pi^{-1} = \nu$ and $S \circ \pi(x) = \pi \circ T(x)$ for $x \in X_0$. When we have such a homomorphism we say that (Y, \mathcal{Y}, ν, S) is a *factor* of (X, \mathcal{X}, μ, T) . If the factor map $\pi : X_0 \rightarrow Y_0$ can be chosen to be bijective and π^{-1} is also measurable, then we say that the systems (X, \mathcal{X}, μ, T) and (Y, \mathcal{Y}, ν, S) are (*measure-theoretically*) *isomorphic*.

2.2. Topological systems. A topological system (X, T) is *transitive* if for any non-empty open sets U, V there is some $n \in \mathbb{Z}_+$ such that $U \cap T^{-n}V \neq \emptyset$. When X has no isolated points, (X, T) is transitive if and only if there exists $x \in X$ whose orbit $\mathcal{O}(x, T) = \{T^n x : n \in \mathbb{Z}_+\}$ is dense in X ; we call such a point a *transitive point*. The system is *minimal* if the orbit of any point is dense in X . A point $x \in X$ is called a *minimal point* if $(\overline{\mathcal{O}(x, T)}, T)$ is minimal. (X, T) is (*topologically*) *weakly mixing* if the product system $(X \times X, T \times T)$ is transitive.

A *factor* of a topological system (X, T) is another topological system (Y, S) such that there exists a continuous surjective map $\phi : X \rightarrow Y$ satisfying $S \circ \phi = \phi \circ T$. In this case, (X, T) is called an *extension* of (Y, S) . The map ϕ is called a *factor map*.

2.3. Rokhlin towers. We need some basic knowledge of Kakutani–Rokhlin towers. We will use the notation from [2, 3, 15].

Let (X, \mathcal{X}, μ, T) be a dynamical system. Let $B \in \mathcal{X}$. An array

$$\mathbf{c} = \{B, TB, \dots, T^{N-1}B\}$$

with $\{T^j B\}_{j=0}^{N-1}$ pairwise disjoint is called a *Rokhlin tower* or a *column over B of height N* . The set B is called the *base* of the tower, and $T^{N-1}B$ is its *roof*. Let $|\mathbf{c}| = \bigcup_{j=0}^{N-1} T^j B$ be the *carrier* of \mathbf{c} . A collection \mathbf{t} of disjoint columns \mathbf{c}_k (with bases B_k and heights N_k) is called a *tower* and we let $|\mathbf{t}| = \bigcup_k |\mathbf{c}_k|$. The union $B = \bigcup_k B_k$ of the bases is the *base* of \mathbf{t} , and the union of the roofs is the *roof* of \mathbf{t} . The sets $\{T^i x : 0 \leq i < N_k\}$ for $x \in B_k$ are called the *fibers* of \mathbf{t} .

Here is the well known Rokhlin Lemma.

THEOREM 2.1 (Rokhlin’s Lemma). *Let (X, \mathcal{X}, μ, T) be an ergodic system. Given an $\epsilon > 0$ and a natural number N , there exists a Rokhlin tower \mathfrak{c} of height N with base $B \in \mathcal{X}$ such that $\mu(|\mathfrak{c}|) > 1 - \epsilon$.*

2.4. Partitions. Let (X, \mathcal{X}, μ, T) be a measurable system. A *partition* α of X is a family of disjoint measurable subsets of X whose union is X . Let α and β be two partitions of (X, \mathcal{X}, μ, T) . One says that α *refines* β , denoted $\alpha \succ \beta$ or $\beta \prec \alpha$, if each element of β is a union of elements of α . Then $\alpha \succ \beta$ is equivalent to $\sigma(\beta) \subseteq \sigma(\alpha)$, where $\sigma(\mathcal{A})$ is the σ -algebra generated by the family \mathcal{A} .

Let α and β be two partitions. Their *join* is the partition $\alpha \vee \beta = \{A \cap B : A \in \alpha, B \in \beta\}$ and we extend this definition naturally to a finite number of partitions. For $m \leq n$, define

$$\alpha_m^n = \bigvee_{i=m}^n T^{-i}\alpha = T^{-m}\alpha \vee T^{-m+1}\alpha \vee \dots \vee T^{-n}\alpha,$$

where $T^{-i}\alpha = \{T^{-i}A : A \in \alpha\}$.

2.5. Refining a tower according to a partition. Let \mathfrak{t} be a tower with columns $\{\mathfrak{c}_k : k \in K\}$ (K is finite or countable) and base $B = \bigcup_{k \in K} B_k$. Given a partition (finite or countable) α , we define an equivalence relation on B as follows: $x \sim y$ iff x and y are in the same base B_k and for every $0 \leq j < N_k$, $T^j x$ and $T^j y$ are in the same element of α , i.e. x and y have the same (α, N_k) -name. Now we consider each equivalence class $B_{k,\mathfrak{a}}$, with \mathfrak{a} an (α, N_k) -name, as a base of the column $\mathfrak{c}_{k,\mathfrak{a}} = \{B_{k,\mathfrak{a}}, TB_{k,\mathfrak{a}}, \dots, T^{N_k-1}B_{k,\mathfrak{a}}\}$ and say that the resulting tower $\mathfrak{t}_\alpha = \{\mathfrak{c}_{k,\mathfrak{a}} : \mathfrak{a} \in \alpha^{N_k}, k \in K\}$ is the *tower \mathfrak{t} refined according to α* .

2.6. Kakutani–Rokhlin towers. For an ergodic system (X, \mathcal{X}, μ, T) , let $B \in \mathcal{X}$ with positive measure. It is clear that $\bigcup_{n \geq 0} T^n B = X \pmod{\mu}$. Define the *return time function* $r_B : B \rightarrow \mathbb{N} \cup \{\infty\}$ by

$$r_B(x) = \min\{n \geq 1 : T^n x \in B\}$$

when this minimum is finite, and $r_B(x) = \infty$ otherwise. Let $B_k = \{x \in B : r_B(x) = k\}$ and note that by Poincaré’s recurrence theorem, B_∞ is a null set. Let \mathfrak{c}_k be the column $\{B_k, TB_k, \dots, T^{k-1}B_k\}$. We call the tower

$$\mathfrak{t} = \mathfrak{t}(B) = \{\mathfrak{c}_k : k = 1, 2, \dots\}$$

the *Kakutani tower over B* . If the Kakutani tower over B has finitely many columns (i.e. the function r_B is bounded), we say that B has a *finite height* and we call the Kakutani tower over B a *Kakutani–Rokhlin tower* or a *K-R tower*. The number $\max r_B$ is called the *height* of B or of the K-R tower.

2.7. Symbolic dynamics. Let Σ be a finite alphabet with l symbols, $l \geq 2$. We usually suppose that $\Sigma = \{1, \dots, l\}$. Let $\Omega = \Sigma^{\mathbb{Z}}$ be the set of all sequences $x = \dots x_{-1}x_0x_1\dots = (x_i)_{i \in \mathbb{Z}}$, $x_i \in \Sigma$, $i \in \mathbb{Z}$, with the product topology. A compatible metric is given by $d(x, y) = 1/(1 + k)$, where $k = \min\{|n| : x_n \neq y_n\}$, $x, y \in \Omega$. The shift map $\sigma : \Omega \rightarrow \Omega$ is defined by $(\sigma x)_n = x_{n+1}$ for all $n \in \mathbb{Z}$. The pair (Ω, σ) is called a *shift dynamical system*. Any subsystem of (Ω, σ) is called a *subshift system*. Similarly we can replace \mathbb{Z} by $\mathbb{Z}_+ = \{0, 1, 2, \dots\}$; then σ will not be a homeomorphism but a surjective map.

Each element of $\Sigma^* = \bigcup_{k \geq 1} \Sigma^k$ is called a *word* or a *block* (over Σ). We use $|A| = n$ to denote the length of A if $A = a_1 \dots a_n$. If $\omega = (\dots \omega_{-1}\omega_0\omega_1 \dots) \in \Omega$ and $a \leq b \in \mathbb{Z}$, then $\omega[a, b] =: \omega_a\omega_{a+1} \dots \omega_b$ is a $(b - a + 1)$ -word occurring in ω starting at place a and ending at place b . Similarly we define $A[a, b]$ when A is a word. A word A *appears* (or *occurs*) in the word B if there are some $a \leq b$ such that $B[a, b] = A$.

For $n \in \mathbb{N}$ and words A_1, \dots, A_n , we denote by $A_1 \dots A_n$ the concatenation of A_1, \dots, A_n . When $A_1 = \dots = A_n = A$ we denote $A_1 \dots A_n$ by A^n . If (X, σ) is a subshift system, let $[i] = [i]_X = \{x \in X : x(0) = i\}$ for $i \in \Sigma$, and $[A] = [A]_X = \{x \in X : x_0x_1 \dots x_{|A|-1} = A\}$ for any word A .

2.8. Symbolic representation. Let (X, \mathcal{X}, μ, T) be an ergodic system and $\alpha = \{A_j\}_{1 \leq j \leq l}$ a finite partition (we usually assume $\mu(A_j) > 0$ for all j). We sometimes think of the partition α as a function $\xi_0 : X \rightarrow \Sigma = \{1, \dots, l\}$ defined by $\xi_0(x) = j$ for $x \in A_j$. The pair (X, α) is traditionally called a *process*. Let $\Omega = \Omega(l) = \{1, \dots, l\}^{\mathbb{Z}}$ and let S be the shift. One can define a homomorphism ϕ_α from X to Ω by $\phi_\alpha(x) = \omega \in \Omega$, where

$$\omega_n = \xi_n(x) = \xi_0(T^n x).$$

We denote the distribution of the stochastic process $(\phi_\alpha)_*(\mu)$ by $\rho = \rho(X, \alpha)$ and call it the *symbolic representation measure* of (X, α) . Let

$$X_\alpha = \text{supp}(\phi_\alpha)_*\mu = \text{supp} \rho.$$

Then we get a homomorphism $\phi_\alpha : (X, \mathcal{X}, \mu, T) \rightarrow (X_\alpha, \mathcal{X}_\alpha, \rho, S)$. This homomorphism is called the *symbolic representation* of the process (X, α) . This will not be a model for (X, \mathcal{X}, μ, T) unless $\bigvee_{i=-\infty}^{\infty} T^{-i}\alpha = \mathcal{X}$ modulo null sets, but in any case this does give a model for a non-trivial factor of X .

2.9. Copying names. An important way to produce partitions is by copying or painting names on towers. If $\mathfrak{c} = \{T^j B\}_{j=0}^{N-1}$ is a column and $\mathbf{a} = (a_0, \dots, a_{N-1}) \in \Sigma^N$ then *copying the name \mathbf{a} on the column \mathfrak{c}* means that on $|\mathfrak{c}| = \bigcup_{j=0}^{N-1} T^j B$ we define a partition (maybe not on the whole space) by letting

$$A_k = \bigcup \{T^j B : \mathbf{a}_j = a_j = k\}, \quad k \in \Sigma = \{1, \dots, l\}.$$

If $\mathbf{c} = \{T^j B\}_{j=0}^{h-1}$ is a column with $N \leq h$, then copying the name $\mathbf{a} \in \Sigma^N$ on the column \mathbf{c} means that we copy the name \mathbf{a} on the first N levels of \mathbf{c} .

If there is a tower \mathbf{t} with q columns $\mathbf{c}_i = \{T^j B_i\}_{j=0}^{N_i-1}$, and q names $\mathbf{a}(i) \in \Sigma^{N_i}, i = 1, \dots, q$, then copying these names on \mathbf{t} means we copy each name $\mathbf{a}(i)$ on the column \mathbf{c}_i , i.e. we define a partition on $|\mathbf{t}|$ by

$$A_k = \bigcup \{T^j B_i : \mathbf{a}(i)_j = k, i = 1, \dots, q\}, \quad k \in \Sigma.$$

These partitions can be extended to a partition $\alpha = \{A_1, \dots, A_l\}$ of the whole space by assigning, for example, the value 1 to the rest of the space. We will do so in the following.

2.10. A metric on partitions. For the set of all finite partitions with the same cardinality, there is a complete metric.

DEFINITION 2.2. Let (X, \mathcal{X}, μ, T) be a system. Let $\alpha = \{A_1, \dots, A_l\}$ and $\beta = \{B_1, \dots, B_l\}$ be two l -set partitions ($l \geq 2$), and define

$$d(\alpha, \beta) = \mu(\alpha \triangle \beta) = \frac{1}{2} \sum_{j=1}^l \mu(A_j \triangle B_j).$$

Note that $d(\alpha, \beta)$ will be different when the partitions are indexed in different ways.

2.11. Some important subsets of integers. A subset S of \mathbb{Z}_+ is *syndetic* if it has bounded gaps, i.e. there is $N \in \mathbb{N}$ such that $\{i, i+1, \dots, i+N\} \cap S \neq \emptyset$ for every $i \in \mathbb{Z}_+$. And S is *thick* if it contains arbitrarily long runs of positive integers, i.e. there is a strictly increasing subsequence $\{n_i\}$ of \mathbb{Z}_+ such that $S \supset \bigcup_{i=1}^\infty \{n_i, n_i+1, \dots, n_i+i\}$. Some dynamical properties can be interpreted by using the notions of syndetic or thick subsets. For example, a classical result of Gottschalk and Hedlund [4] states that x is a minimal point if and only if the set

$$N(x, U) = \{n \in \mathbb{Z}_+ : T^n x \in U\}$$

is syndetic for any neighborhood U of x , and by Furstenberg [1] a topological system (X, T) is weakly mixing if and only if the set

$$N(U, V) = \{n \in \mathbb{Z}_+ : U \cap T^{-n}V \neq \emptyset\}$$

is thick for any non-empty open subsets U, V of X .

A set S is called *thickly syndetic* if for every N the positions where length N runs begin form a syndetic set. A subset S of \mathbb{Z}_+ is *piecewise syndetic* if it is the intersection of a syndetic set with a thick set. It is known that a topological system (X, T) is an *M-system* (i.e. the set of minimal points of (X, T) is dense) if and only if there is a transitive point x such that $N(x, U)$ is piecewise syndetic for any neighborhood U of x (see for example [7, Lemma 2.1]).

In a symbolic system, it is useful to use these subsets of integers to characterize the dynamical properties. For example, a sequence $\omega \in \Omega$ is recurrent if and only if every word occurring in ω occurs in ω a second time; and ω is a minimal point if and only if every word that occurs in ω occurs along a syndetic set.

3. An improvement of a technical lemma in Weiss's proof of the Jewett–Krieger theorem. In this section we will prove a lemma which is an improvement of a technical lemma in Weiss's proof of the Jewett–Krieger theorem. Using this lemma one may simplify Weiss's arguments in some sense.

In Weiss's new proof of the Jewett–Krieger theorem [12, 15, 2] and in the proof of Weiss's theorem on the doubly minimal model [14], one needs the following technical lemma to get a new K-R tower from a given one:

LEMMA 3.1 ([2, 3, 15]). *Let (X, \mathcal{B}, μ, T) be a non-periodic ergodic system, and let $\mathfrak{t}(C_0)$ be a K-R tower (i.e. $\max r_{C_0} < \infty$). Then for all N sufficiently large, there exists a set $C_1 \subset C_0$ such that*

$$N \leq r_{C_1}(y) \leq N + 4 \max r_{C_0}, \quad \forall y \in C_1.$$

That is, the corresponding K-R tower $\mathfrak{t}(C_1)$ satisfies $\text{range } r_{C_1} \subset [N, N + 4 \max r_{C_0}]$.

When one uses this lemma, one hopes that the column heights of the K-R tower are relatively prime, which is not guaranteed in Lemma 3.1. Hence in Weiss's new proof of the Jewett–Krieger theorem, one first assumes that the system has no rational spectrum, in which case automatically the column heights of every K-R tower are relatively prime. Then one deals with other cases. The following lemma will avoid this kind of discussion.

LEMMA 3.2. *Let (X, \mathcal{B}, μ, T) be a non-periodic ergodic system. Let \mathfrak{t} be a K-R tower with bases C_i and heights h_i , $1 \leq i \leq k$, and let $N = \max_i \{h_i\}$ and $C = \bigcup_{i=1}^k C_i$. Assume that h_1, \dots, h_k are relatively prime. Then for any n large enough, there is a K-R tower \mathfrak{t} with base D such that:*

- (1) $D \subset C$;
- (2) $r_D(y) \in [n, n + 6N]$ for all $y \in D$;
- (3) the column heights of $\mathfrak{t}(D)$ are relatively prime.

Proof. First we will find a set $\hat{D} \subset C$ with the following two properties:

- (i) $n + N \leq r_{\hat{D}}(y) \leq n + 5N$ for all $y \in \hat{D}$;
- (ii) $\mu(\hat{D} \cap C_i) > 0$ for each $1 \leq i \leq k$.

Then according to the second property of \hat{D} , we adjust some part of \hat{D} to get D such that the column heights of $\mathfrak{t}(D)$ are relatively prime.

STEP 1: *The construction of \hat{D} .* We first construct a Kakutani tower $\mathfrak{t}(\hat{B})$ with height larger than $10(n + 3N)^2$ and $\mu(\hat{B} \cap C_i) > 0$ for each i . But at this point we may have $\max r_{\hat{B}} = \infty$ (i.e. $\mathfrak{t}(\hat{B})$ may not be a K-R tower). So we will have to modify it to get a K-R tower.

By the Rokhlin Lemma, there is a $B \subset C$ such that the Rokhlin tower

$$\mathfrak{c} = \{B, TB, \dots, T^{M-1}B\}$$

satisfies $M > 20(n + 3N)^2$ and $\mu(B) < \frac{\min_i \{\mu(C_i)\}}{10k(n+3N)^2+k}$.

Let $n_0 = 0$. Find the smallest $n_1 \in \mathbb{N}$ with

- (a₁) $n_1 - n_0 \geq 10(n + 3N)^2$;
- (b₁) $\mu(T^{n_1}B \cap (\bigcup_{j=1}^k C_j \setminus \bigcup_{j=0}^{10(n+3N)^2} T^j B)) > 0$.

Hence there is some $d_1 \in \{1, \dots, k\}$ such that

$$\mu\left(T^{n_1}B \cap \left(C_{d_1} \setminus \bigcup_{j=0}^{10(n+3N)^2} T^j B\right)\right) > 0.$$

Let

$$B_1 = T^{n_1}B \cap \left(C_{d_1} \setminus \bigcup_{j=0}^{10(n+3N)^2} T^j B\right).$$

Inductively, assume that for $1 \leq i \leq k - 1$ we have obtained n_1, \dots, n_i , distinct numbers $d_1, \dots, d_i \in \{1, \dots, k\}$ and measurable sets B_1, \dots, B_i .

Let n_{i+1} be the smallest natural number satisfying

- (a _{$i+1$}) $n_{i+1} - n_i \geq 10(n + 3N)^2$;
- (b _{$i+1$}) $\mu(T^{n_{i+1}}B \cap (\bigcup_{j=1}^k C_j \setminus \bigcup_{j=1}^i C_{d_j} \cup \bigcup_{s=0}^i \bigcup_{j=n_s}^{n_s+10(n+3N)^2} T^j B)) > 0$.

Hence there is some $d_{i+1} \in \{1, \dots, k\} \setminus \{d_1, \dots, d_i\}$ such that

$$\mu\left(T^{n_{i+1}}B \cap \left(C_{d_{i+1}} \setminus \left(\bigcup_{j=1}^i C_{d_j} \cup \bigcup_{s=0}^i \bigcup_{j=n_s}^{n_s+10(n+3N)^2} T^j B\right)\right)\right) > 0.$$

Let

$$B_{i+1} = T^{n_{i+1}}B \cap \left(C_{d_{i+1}} \setminus \left(\bigcup_{j=1}^i C_{d_j} \cup \bigcup_{s=0}^i \bigcup_{j=n_s}^{n_s+10(n+3N)^2} T^j B\right)\right).$$

Note that $B_{i+1} = T^{n_{i+1}}B \cap (C_{d_{i+1}} \setminus \bigcup_{s=0}^i \bigcup_{j=n_s}^{n_s+10(n+3N)^2} T^j B)$. This inductive process can be done for $i = 2, \dots, k$ since $\mu(B) < \frac{\min_i \{\mu(C_i)\}}{10k(n+3N)^2+k}$, which means $\mu(\bigcup_{i=0}^{k-1} \bigcup_{j=0}^{10(n+3N)^2} T^{n_i+j} B) < \mu(C_s)$ for $1 \leq s \leq k$.

Now by induction we obtain subsets B_1, \dots, B_k . Let $\hat{B} = \bigcup_{i=1}^k B_i$. We claim that:

The height of each column in the Kakutani tower $\mathfrak{t}(\hat{B})$ is larger than $10(n+3N)^2$, i.e. $r_{\hat{B}}(y) \geq 10(n+3N)^2$ for all $y \in \hat{B}$.

To prove the claim, we need to prove that for any $l > 0$ and $1 \leq u, v \leq k$, $\mu(T^l B_u \cap B_v) > 0$ implies $l \geq 10(n+3N)^2$. Since $\mu(T^l B_u \cap B_v) > 0$, there is a subset $P \subset T^l B_u \cap B_v$ with positive measure. If $u = v$, then $l \geq 10(n+3N)^2$ since $B_u \subset B$. If $u < v$, then $l \geq 10(n+3N)^2$ since $T^{-l}P \subset B_u \subset B$, $P \subset B_v$ and $\mu(\bigcup_{j=n_u}^{n_u+10(n+3N)^2} T^j B) \cap B_v = 0$. Finally, assume $u > v$. Since n_u is the first number satisfying the inductive condition (a_u), we have $\mu(\bigcup_{j=n_{v-1}+10(n+3N)^2+1}^{n_u-1} T^j B \cap B_u) = 0$. Moreover $\mu(\bigcup_{j=n_{v-1}}^{n_{v-1}+10(n+3N)^2} B \cap B_u) = 0$, so $\mu(\bigcup_{j=n_{v-1}}^{n_{v-1}-1} T^j B) \cap B_u = 0$. Since $n_u - n_{u-1} \geq 10(n+3N)^2$ and $T^{-l}P \subset B_u \cap T^{-l}B_v$, we conclude that $l \geq 10(n+3N)^2$.

By the construction we also see that $\mu(\hat{B} \cap C_i) > 0$ for each $i \in \{1, \dots, k\}$. Since $n+3N$ and $n+3N+1$ are relatively prime, we may partition each column of $\mathfrak{t}(\hat{B})$ into blocks of sizes $n+3N$ and $n+3N+1$. Then we move the base level of each block to the nearest level that belongs to C . Taking the union of the base level and \hat{B} , we get a set $\hat{D} \subset C$ satisfying

- (I) the height of $\mathfrak{t}(\hat{D})$ is in $[n+N, n+5N]$;
- (II) $\mu(\hat{D} \cap C_i) > 0$ for each $1 \leq i \leq k$.

The set of heights of $\mathfrak{t}(\hat{D})$ may not be relatively prime, so we need to modify it.

STEP 2: *The construction of D .* For each $i \in \{1, \dots, k\}$, let $E_i \subset \hat{D} \cap C_i$ be a measurable subset with positive measure. Thus we get k sets E_1, \dots, E_k with respective heights $\hat{h}_1, \dots, \hat{h}_k$. Since T is non-periodic and ergodic, $\mu(E_i \setminus T^{\hat{h}_i} E_i) > 0$ for each i . Let $\epsilon_0 = \min_i \mu(E_i \setminus T^{\hat{h}_i} E_i)$.

Let $F_1 \subset E_1 \setminus T^{\hat{h}_1} E_1$ be a subset satisfying $0 < \mu(F_1) < \epsilon_0/2^{k+1}$. Inductively assume for $1 \leq i \leq k-1$ we have constructed subsets F_1, \dots, F_i satisfying $2\mu(F_j) \leq \mu(F_{j+1}) < \epsilon_0/2^{k-j+1}$ for each $1 \leq j \leq i-1$. Note that

$$(3.1) \quad \sum_{j=1}^i \mu(F_j) \leq \sum_{j=1}^i \frac{\epsilon_0}{2^{k-j+2}} = \frac{\epsilon_0}{2^{k-i+2}} \frac{1-1/2^i}{1-1/2} < \frac{\epsilon_0}{2^{k-i+1}}.$$

Thus $\mu((E_{i+1} \setminus T^{\hat{h}_{i+1}} E_{i+1}) \setminus \bigcup_{j=1}^i T^{\hat{h}_j} F_j) > \epsilon_0(1-1/2^{k-i+1})$. Hence one can find

$$F_{i+1} \subset (E_{i+1} \setminus T^{\hat{h}_{i+1}} E_{i+1}) \setminus \bigcup_{j=1}^i T^{\hat{h}_j} F_j$$

satisfying $2\mu(F_i) \leq \mu(F_{i+1}) < \epsilon_0/2^{k-i+1}$.

In such a way by induction we get k sets F_1, \dots, F_k . For each i , we have the following properties:

- (i) $F_i \subset C_i$, which implies $T^{h_i} F_i \subset C$;
- (ii) $F_i \subset E_i \setminus T^{\hat{h}_i} E_i$, which implies $T^{\hat{h}_i} F_i \subset \hat{D} \setminus F_i$;
- (iii) $T^{\hat{h}_i} F_i \cap F_j = \emptyset$ for $j \geq i$ and $\mu(T^{\hat{h}_i} F_i) = \mu(F_i) > \sum_{s=1}^{i-1} \mu(F_s)$;
- (iv) $\mu(T^{\hat{h}_i} F_i \cap (\hat{D} \setminus \bigcup_{j=1}^k F_j)) > 0$.

Here (i), (ii) and the first part of (iii) follow from the definition of F_i . The second part of (iii) follows from the inequality $\mu(F_{j+1}) \geq 2\mu(F_j)$:

$$\mu(T^{\hat{h}_i} F_i) = \mu(F_i) \geq 2\mu(F_{i-1}) \geq \mu(F_{i-1}) + 2\mu(F_{i-2}) \geq \dots > \sum_{s=1}^{i-1} \mu(F_s).$$

And (iv) is deduced from (iii) readily.

Finally, we set $D = (\hat{D} \setminus \bigcup_{i=1}^k F_i) \cup \bigcup_{i=1}^k T^{h_i} F_i$. By the properties of $\{F_i\}$, we conclude:

- (1) $D \subset C$;
- (2) $\mathfrak{t}(D)$ is a K-R tower, and its height is in $[n, n + 6N]$;
- (3) the collection of the column heights of the K-R tower $\mathfrak{t}(D)$ contains $\{\hat{h}_i, \hat{h}_i - h_i\}_{i=1}^k$, which are relatively prime since $\{h_i\}_{i=1}^k$ are relatively prime.

(1) follows from the definition of D , and (2) from (i) above. By (iv) and the definition of D , for each $i \in \{1, \dots, k\}$ there is some column of $\mathfrak{t}(D)$ with height $\hat{h}_i - h_i$. By (3.1), we have $\mu(E_i)/2 \geq \epsilon_0/2 > \sum_{i=1}^k \mu(F_i)$, which implies that for each $i \in \{1, \dots, k\}$ there is some column of $\mathfrak{t}(D)$ with height \hat{h}_i . Hence we have (3).

The tower $\mathfrak{t}(D)$ is as required. ■

4. Proof of Theorem 1.1(1). To prove Theorem 1.1(1), we begin with the following observation.

LEMMA 4.1. *Let (X, \mathcal{X}, μ, T) be a non-periodic ergodic system. Then there is a tower whose set of column heights is infinite.*

Proof. Given a tower with base C_1 , if the set of column heights is infinite, then we are done; otherwise let this set be $\{h_1, \dots, h_{n_1}\}$. Let C_1^i be the corresponding column-base with height h_i ; we may assume that $h_1 < \dots < h_{n_1}$ (by putting the column-bases with the same height together to form a new column-base). Choose a measurable set $E_1 \subset T^{h_{n_1}} C_1^{n_1}$ such that $0 < \mu(E_1) < \frac{1}{2} \min_{1 \leq i \leq n_1} \mu(C_1^i)$.

Let $C_2 = C_1 \setminus E_1$ and suppose we have a tower with base C_2 . If the set of its column heights is infinite, then we are done. Otherwise we get a bigger height set than for the tower with base C_1 , say $\{h_1, \dots, h_{n_1}, h_{n_1+1}, \dots, h_{n_2}\}$.

Let C_2^i be the corresponding column-base with height h_i ; we assume that $h_1 < \dots < h_{n_1} < \dots < h_{n_2}$. Choose a measurable set $E_2 \subset T^{h_{n_2}} C_2^{n_2}$ such that $0 < \mu(E_2) < \frac{1}{2^2} \min_{1 \leq i \leq n_2} \mu(C_2^i)$.

Let $C_3 = C_2 \setminus E_2$ and continue the process above. If after finitely many steps we get a tower with infinitely many heights, then we are done. Otherwise we will have a sequence of towers with decreasing bases $\{C_k\}$, $n_1 < \dots < n_k$ and measurable sets E_j with $0 < \mu(E_j) < \frac{1}{2^j} \min_{1 \leq i \leq n_j} \mu(C_j^i)$ for $1 \leq j \leq k$ such that

$$\mu(C_{k+1}) \geq \mu(C_k) - \mu(E_k) > (1 - 1/2^k)\mu(C_k)$$

for all $k \in \mathbb{N}$. Let $C = \bigcap_{k=1}^{\infty} C_k$. Then $\mu(C) > 0$ and the tower with base C has infinitely many heights. ■

We now follow the standard procedure to prove Theorem 1.1(1). Namely, first for a given partition $\hat{\alpha}$ we construct a partition α close to $\hat{\alpha}$ such that the corresponding symbolic representation $(X_\alpha, \mathcal{X}_\alpha, \rho, S)$ is a non-minimal topologically weakly mixing system with a dense set of minimal points. Then we take the inverse limit, using a more delicate argument. Finally, we show that the resulting system is as desired.

PROPOSITION 4.2. *Let (X, \mathcal{X}, μ, T) be a non-periodic ergodic system and let $\hat{\alpha}$ be a finite partition of X . For each $\epsilon > 0$, there is a partition α such that the corresponding symbolic representation $(X_\alpha, \mathcal{X}_\alpha, \rho, S)$ is a non-minimal topologically weakly mixing system with a dense set of minimal points, and $d(\alpha, \hat{\alpha}) < \epsilon$.*

Proof. By Lemma 4.1, there is a tower consisting of infinitely many columns with different heights. More precisely, let $\mathfrak{t}(C)$ be a tower as in Lemma 4.1 with columns $\{\mathfrak{c}_k : k \in \mathbb{N}\}$ and base $C = \bigcup_{k \in \mathbb{N}} C_k$. Let h_k be the height of the column \mathfrak{c}_k , and assume that $h_1 < h_2 < \dots$. Note that for k large enough, $|\mathfrak{c}_k|$ will be very small. We will adjust some $|\mathfrak{c}_k|$ to get what we need.

Let $\alpha_0 = \hat{\alpha} = \{\hat{A}_1, \dots, \hat{A}_k\}$. For each $m \in \mathbb{N}$, let $\omega_m = v_1 \dots v_k^m$, where

$$v_{a_1 k^{m-1} + a_2 k^{m-2} + \dots + a_m} = (a_1, \dots, a_m)$$

for each $(a_1, \dots, a_m) \in \{1, \dots, k\}^m$. That is, each v_i is a word of length m and ω_m is a word which contains all the m -names in $\{1, \dots, k\}^m$. Note that $|\omega_m| = mk^m$.

Before going on, let us recall the notion of copying a name on the column. Let $\mathfrak{c} = \{T^j B\}_{j=0}^{h-1}$ be a column and $\mathbf{a} = (a_0, \dots, a_{N-1}) \in \{1, \dots, k\}^N$ with $N \leq h$. Then copying the name \mathbf{a} on the column \mathfrak{c} means that we copy the name \mathbf{a} on the first N levels of \mathfrak{c} . This means that for the new partition $\{A_1, \dots, A_k\}$ one has

$$T^i B \subset A_{a_i}, \quad 0 \leq i \leq N - 1.$$

STEP 1. Since $\sum_k |\mathbf{c}_k| < \infty$, there are columns $\mathbf{c}_{n_1^1}$ and $\mathbf{c}_{n_1^2}$ such that $\mu(|\mathbf{c}_{n_1^1} \cup \mathbf{c}_{n_1^2}|) < \epsilon/2$ and $h_{n_1^2} > h_{n_1^1} > 2k^2$. Let $\xi_{1_1^1} = \omega_1 1^{h_{n_1^1} - k} \in \{1, \dots, k\}^{h_{n_1^1}}$, where $1^j = (1, \dots, 1)$ with length j . And let $\xi_{1_1^2} = \omega_2 1^{h_{n_1^2} - k} \in \{1, \dots, k\}^{h_{n_1^2}}$. For $i = 1, 2$, copy the name $\xi_{1_1^i}$ to the column $\mathbf{c}_{n_1^i}$, to get a partition α_1 . Note that $d(\alpha_0, \alpha_1) = d(\hat{\alpha}, \alpha_1) < \epsilon/2$. The first step of adjustment is finished.

STEP m . In general, for each $m \in \mathbb{N}$, choose columns $\mathbf{c}_{n_m^1}, \dots, \mathbf{c}_{n_m^{m+1}}$ such that

$$\mu\left(\bigcup_{i=1}^{m+1} |\mathbf{c}_{n_m^i}|\right) < \frac{\epsilon}{2^m},$$

and assume that $h_{n_m^{m+1}} > \dots > h_{n_m^1} > 2m^2 k^{2m}$. For $1 \leq i \leq m$, let

$$\begin{aligned} \xi_{n_m^i} &= \underbrace{\omega_{2i-1} \dots \omega_{2i-1}}_{m \text{ times}} 1^{h_{n_m^i} - m(2i-1)k^{2i-1}} \\ &= (\omega_{2i-1})^m 1^{h_{n_m^i} - m(2i-1)k^{2i-1}} \in \{1, \dots, k\}^{h_{n_m^i}}, \end{aligned}$$

and let

$$\xi_{n_m^{m+1}} = \omega_{2m} 1^{h_{n_m^{m+1}} - 2mk^{2m}} \in \{1, \dots, k\}^{h_{n_m^{m+1}}}.$$

Now for $1 \leq i \leq m + 1$, copy the name $\xi_{n_m^i}$ to the column $\mathbf{c}_{n_m^i}$, to get a new partition α_m . Note that $d(\alpha_{m-1}, \alpha_m) < \epsilon/2^m$.

Moreover, note that we do copying on the m columns $\mathbf{c}_{n_m^1}, \dots, \mathbf{c}_{n_m^m}$ to make sure that the symbolic representation of the resulting partition is not minimal but has a dense set of minimal points, and we do copying on $\mathbf{c}_{n_m^{m+1}}$ to make sure the symbolic representation of the resulting partition is weakly mixing. Of course, we could do it on a single column, but this would cause complications when dealing with the situation in Proposition 4.4.

For all $m \in \mathbb{N}$, we make the above adjustment, and we obtain a new partition $\alpha = \{A_1, \dots, A_k\}$. It is clear that

$$d(\hat{\alpha}, \alpha) \leq \sum_{i=1}^{\infty} d(\alpha_{i-1}, \alpha_i) < \sum_{i=1}^{\infty} \frac{\epsilon}{2^i} = \epsilon.$$

Properties of α . Now we prove that $(X_\alpha, \mathcal{X}_\alpha, \rho, S)$ is non-minimal, weakly mixing, and the set of minimal points is dense.

To show (X_α, S) is weakly mixing, it suffices to show that for each $m \in \mathbb{N}$ and $E_1, F_1, E_2, F_2 \in \bigvee_{i=0}^{m-1} T^{-i} \alpha$ of positive measure, we have

$$\mu \times \mu((E_1 \times F_1) \cap (T \times T)^{-m}(E_2 \times F_2)) > 0.$$

This depends on the adjustment on $\mathbf{c}_{n_m^{m+1}}$.

Recall that $\omega_{2m} = v_1 \dots v_{k^{2m}}$, where $\{v_j\}_{j=1}^{k^{2m}} = \{1, \dots, k\}^{2m}$. Let the names of E_1, E_2, F_1, F_2 be $e_1, e_2, f_1, f_2 \in \{1, \dots, k\}^m$ respectively. Then

$e_1 e_2 = u_s$ and $f_1 f_2 = u_t$ for some $1 \leq t, s \leq k^{2m}$. By the construction of α ,

$$T^{2m(s-1)}C_{n_m^{m+1}} \subset E_1 \cap T^{-m}E_2 \quad \text{and} \quad T^{2m(t-1)}C_{n_m^{m+1}} \subset F_1 \cap T^{-m}F_2.$$

Thus,

$$(E_1 \times F_1) \cap (T \times T)^{-m}(E_2 \times F_2) \supset T^{2m(s-1)}C_{n_m^{m+1}} \times T^{2m(t-1)}C_{n_m^{m+1}}.$$

In particular,

$$\begin{aligned} \mu \times \mu((E_1 \times F_1) \cap (T \times T)^{-m}(E_2 \times F_2)) \\ \geq \mu \times \mu(T^{2m(s-1)}C_{n_m^{m+1}} \times T^{2m(t-1)}C_{n_m^{m+1}}) > 0. \end{aligned}$$

To see (X_α, S) is a non-minimal M -system, we show that each transitive point $w \in X_\alpha$ is piecewise syndetically but not syndetically recurrent. Let $x \in X$ be such that $\phi_\alpha(x) = w$. It is easy to see that $w \neq 1^\infty$.

Let $w = (\dots, a_{-2}, a_{-1}, a_0, a_1, a_2, \dots)$. Then for each $m \in \mathbb{N}$,

$$[w]_{-m+1}^{m-1} = \{p \in X_\alpha : p[-m+1, m-1] = (a_{-m+1}, \dots, a_0, a_1, \dots, a_{m-1})\}$$

is a neighborhood of w . Let $A \in \bigvee_{i=-m+1}^{m-1} T^{-i}\alpha$ with name $(a_{-m+1}, \dots, a_{m-1})$. Since $w \neq 1^\infty$, it is clear that for m large enough, $(a_{-m+1}, \dots, a_{m-1}) \neq 1^{2m-1}$.

As defined before, $\omega_{2m-1} = v_1 \dots v_{k^{2m-1}}$, where $\{v_i\}_{i=1}^{k^{2m-1}} = \{1, \dots, k\}^{2m-1}$. Then $(a_{-m+1}, \dots, a_{m-1}) = v_r$ for some r . For each $j \geq m$, by the definition of x , one can find l_j such that $T^{l_j}x \in C_{n_j^m}$. By the construction of α , for $1 \leq i \leq j$, $T^{i(2m-1)k^{2m-1} + (2m-1)(r-1) + m}C_{n_j^m} \subset A$. This means that for each $j \geq m$,

$$\{l_j + i(2m-1)k^{2m-1} + (2m-1)(r-1) + m\}_{1 \leq i \leq j} \subset N(w, [w]_{-m+1}^{m-1}),$$

which implies $N(w, [w]_{-m+1}^{m-1})$ is piecewise syndetic.

On the other hand, for each $j > m$ and $m-1 < i < h_{n_j^m} - j(2m-1)k^{2m-1}$, we have $T^{l_j + j(2m-1)k^{2m-1} + i}x \in \bigcap_{d=-m+1}^{m-1} T^{-d}A_1$. As $(a_{-m+1}, \dots, a_{m-1}) \neq 1^m$, it follows that $\bigcap_{d=0}^{m-1} T^{-d}A_1 \cap A = \emptyset$ and hence $T^{l_j + j(2m-1)k^{2m-1} + i}x \notin A$, which implies that for each $j > m$,

$$\{l_j + j(2m-1)k^{2m-1} + i\}_{m-1 < i < h_{n_j^m} - j(2m-1)k^{2m-1}} \cap N(w, [w]_{-m+1}^{m-1}) = \emptyset.$$

Since $h_{n_j^m} - j(2m-1)k^{2m-1} \geq 2j^2k^{2j} - j(2m-1)k^{2m-1}$ and

$$2j^2k^{2j} - j(2m-1)k^{2m-1} \xrightarrow{j \rightarrow \infty} \infty,$$

we conclude that $N(w, [w]_{-m+1}^{m-1})$ is not syndetic. ■

To prove Theorem 1.1(1), Proposition 4.2 is not enough. In fact we need to get an increasing sequence of partitions γ_n such that the inverse limit of the corresponding symbolic representations is what we need.

REMARK 4.3. In the proof of Proposition 4.4, the following simple fact will be used. Let $\alpha = \{A_1, \dots, A_a\}$, $\alpha' = \{A'_1, \dots, A'_a\}$, and $\beta = \{B_1, \dots, B_b\}$ be partitions with $\alpha \succ \beta$. Then there is a natural way to get a partition β' such that $\alpha' \succ \beta'$. Moreover, if $d(\alpha, \alpha') < \epsilon$, then also $d(\beta, \beta') < \epsilon$.

To see this, we note that the relation $\alpha \succ \beta$ defines a function

$$\phi : \{1, \dots, b\} \rightarrow 2^{\{1, \dots, a\}} \setminus \emptyset$$

such that $A_x \subset B_y$ iff $x \in \phi(y)$. Let

$$\beta' = \{B'_1, \dots, B'_b\}, \quad B'_s = \bigcup_{t \in \phi(s)} A'_t.$$

Notice that if $d(\alpha, \alpha') < \epsilon$, then also $d(\beta, \beta') < \epsilon$, since it is easy to check that $(A_1 \cup B_1) \triangle (A_2 \cup B_2) \subset A_1 \triangle A_2 \cup B_1 \triangle B_2$.

PROPOSITION 4.4. *Let (X, \mathcal{X}, μ, T) be a non-periodic ergodic system. Then there exists an increasing sequence $\{\gamma_n\}$ of finite partitions such that $\sigma(\gamma_n) \nearrow \mathcal{X}$ and for each $n \in \mathbb{N}$ the corresponding symbolic representation $(X_{\gamma_n}, \mathcal{X}_{\gamma_n}, \rho_n, S)$ is a non-minimal topologically weakly mixing system with a dense set of minimal points.*

Proof. The basic idea of the proof is the same as in the proof of Proposition 4.2. Since we have to deal with countably many partitions, we need to make some small modifications to the proof.

Let $\{\beta_n\}$ be an increasing sequence of finite partitions such that $\sigma(\beta_n) \nearrow \mathcal{X}$. First we fix the same tower $t(C)$ as in the proof of Proposition 4.2 and let $\{\epsilon_n\}$ be a sequence with $\sum_{n=1}^\infty \epsilon_n < \infty$.

For β_1 , we adjust α_1 as in Step 1 of the proof of Proposition 4.2 to get a new partition γ_1^1 . We replace β_2 by $\beta_2 \vee \gamma_1^1$, and thus we have $\gamma_1^1 \prec \beta_2$. Then we continue our induction. To be precise, we rewrite Step m .

STEP m' . We replace β_m by $\beta_m \vee \gamma_{m-1}^{m-1}$ (still denoting it by β_m), and thus we have $\gamma_{m-1}^{m-1} \prec \beta_m$. Let $\beta_j = \{B_1^j, \dots, B_{k_j}^j\}$ for $1 \leq j \leq m$. Since $\beta_1 \prec \dots \prec \beta_m$, we may assume that $B_i^j \subset B_i^{j-1}$ for all $2 \leq j \leq m$ and $1 \leq i \leq k_{j-1}$.

Recall that the word ω_m depends on the cardinality of the partition in the proof of Proposition 4.2. Unlike the situation there, now the cardinalities of partitions are increasing. Let $\omega_m = \omega_m(k)$ as in the proof of Proposition 4.2, and denote $\omega_{m,j} = \omega_m(k_j)$ for $1 \leq j \leq m$. That is, $\omega_{m,j}$ is a word which contains all the m -names in $\{1, \dots, k_j\}^m$.

For each $m \in \mathbb{N}$, choose columns $\mathbf{c}_{n_m^1}, \dots, \mathbf{c}_{n_m^{m(m+1)}}$ such that

$$\mu \left(\bigcup_{i=1}^{m(m+1)} |\mathbf{c}_{n_m^i}| \right) < \epsilon_m.$$

The columns should be disjoint from the columns in Step k for all $k < m$. We assume that $h_{n_m^{j(m+1)}} > \dots > h_{n_m^{(j-1)(m+1)+1}} > 2m^2k_j^{2m}$ for each $1 \leq j \leq m$. For each $1 \leq i, j \leq m$ and $s = (m+1)(j-1) + i$, let

$$\begin{aligned} \xi_{n_m^s} &= \underbrace{\omega_{2i-1,j} \dots \omega_{2i-1,j}}_{m-j \text{ times}} 1^{h_{n_m^s} - (m-j)(2i-1)k_j^{2i-1}} \\ &= (\omega_{2i-1,j})^{m-j} 1^{h_{n_m^s} - (m-j)(2i-1)k_j^{2i-1}} \in \{1, \dots, k_j\}^{h_{n_m^s}}, \end{aligned}$$

and

$$\xi_{n_m^{j(m+1)}} = \omega_{2m,j} 1^{h_{n_m^{j(m+1)}} - 2mk_j^{2m}} \in \{1, \dots, k_j\}^{h_{n_m^{j(m+1)}}}.$$

Now for $1 \leq i \leq m(m+1)$, copy the name $\xi_{n_m^i}$ to the column $\epsilon_{n_m^i}$, to obtain a new partition γ_m^m with $d(\beta_m, \gamma_m^m) < \epsilon_m$.

Inductively, we construct a sequence $\{\gamma_n^n\}_n$ of partitions with $d(\beta_m, \gamma_m^m) < \epsilon_m$ for each $m \in \mathbb{N}$. Now we need to build the required partition $\{\gamma_n\}$ from $\{\gamma_n^n\}_n$. First we construct partitions $\{\gamma_k^n\}_{n \in \mathbb{N}, 1 \leq k \leq n}$ via $\{\gamma_n^n\}_n$. Then $\gamma_k = \lim_n \gamma_k^n$ is what we are looking for.

$$\begin{array}{cccc} \gamma_1^1 & & & \\ \gamma_1^2 & \gamma_2^2 & & \\ \gamma_1^3 & \gamma_2^3 & \gamma_3^3 & \\ \dots & \dots & \dots & \dots \\ \downarrow & \downarrow & \downarrow & \downarrow \\ \gamma_1 & \gamma_2 & \gamma_3 & \dots \end{array}$$

Applying Remark 4.3 to β_2, γ_2^2 , and γ_1^1 we obtain γ_1^2 . Similarly, applying Remark 4.3 to β_3, γ_3^3 , and γ_2^2 we obtain γ_2^3 , and we get γ_1^3 by applying Remark 4.3 to β_3, γ_3^3 and γ_1^2 . Inductively, we construct γ_k^n for $k < n$ by applying Remark 4.3 and $\beta_n \prec \gamma_k^{n-1}$, for $k < n$. Since $\beta_1 \prec \dots \prec \beta_n$, we have $\gamma_1^n \prec \dots \prec \gamma_n^n$ accordingly.

Since $d(\gamma_n^n, \beta_n) < \epsilon_n$ and $\beta_n \succ \gamma_{n-1}^{n-1}$, we know that $d(\gamma_k^n, \gamma_k^{n-1}) < \epsilon_n$ for each $1 \leq k \leq n-1$. Thus for each k , $\{\gamma_k^n\}_{n \geq k}$ is a Cauchy sequence. So there is a partition γ_k such that $\gamma_k^n \rightarrow \gamma_k$ as $n \rightarrow \infty$. Let X_k^n denote the corresponding symbolic system of γ_k^n . The array shows the induction.

$$\begin{array}{cccc} X_1^1 & & & \\ X_1^2 & X_2^2 & & \\ X_1^3 & X_2^3 & X_3^3 & \\ \dots & \dots & \dots & \dots \end{array}$$

By the construction, for each k , the sequence $\{\gamma_k^n\}_{n \geq k}$ has the same property as $\{\alpha_n\}$ in Proposition 4.2. According to the proof of Proposition 4.2, the corresponding symbolic system $X_k = X_{\gamma_k}$ of γ_k is non-minimal topologically weakly mixing with a dense set of minimal points.

Since $\gamma_1^n \prec \dots \prec \gamma_n^n$ for each $n \in \mathbb{N}$, we conclude that $\{\gamma_k\}_k$ is increasing. As $\sigma(\beta_k) \nearrow \mathcal{X}$ and $d(\gamma_k, \beta_k) < \sum_{s=k}^\infty \epsilon_s$, we deduce that $\sigma(\gamma_k) \nearrow \mathcal{X}$ too. ■

Now Theorem 1.1(1) follows from Proposition 4.4 and the following lemma.

LEMMA 4.5. *Let (X, T) be the inverse limit of $\{(X_n, T_n)\}_n$, where each (X_n, T_n) is a non-minimal topologically weakly mixing system with a dense set of minimal points. Then (X, T) is also a non-minimal topologically weakly mixing system with a dense set of minimal points.*

Proof. By the definition of the inverse limit, it is easy to see that (X, T) is not minimal, as the factor of a minimal system is minimal.

To show the density of minimal points in X assume U is a non-empty open set. Let $\pi_n : X \rightarrow X_n$ be the projection. Then by the definition of the topology of X , there are $n \in \mathbb{N}$ and an open non-empty set $U_n \subset X_n$ such that $\pi_n^{-1}U_n \subset U$. Let $x_n \in U_n$ be a minimal point and A be its orbit closure. Then there is a minimal set B of X such that $\pi_n(B) = A$. This implies that there is a minimal point x of X such that $\pi_n(x) = x_n$, which implies that $x \in \pi_n^{-1}(x_n) \subset \pi_n^{-1}(U_n) \subset U$, and hence the set of minimal points of X is dense. A similar argument can be applied to show that (X, T) is weakly mixing. ■

5. Proof of Theorem 1.1(2). First we will construct a model which is weakly mixing with full support but its set of minimal points is not dense. Since in this case the closure of the set of minimal points has measure zero, we collapse it to a point and get the system required.

First we need the following lemma (see [2, 3, 15] for a proof).

LEMMA 5.1. *Let X be a non-periodic ergodic system. For any positive integers N_1, N_2 with $(N_1, N_2) = 1$, there exists a set C of finite height such that the K - R tower $\mathbf{t}(C)$ satisfies $\text{range } r_C \subset \{N_1, N_2\}$.*

To show Theorem 1.1(2) we start with the following proposition and then follow the standard procedure. To control the thickly syndetic sets, the construction here is more involved than that in Proposition 4.2.

PROPOSITION 5.2. *Let (X, \mathcal{X}, μ, T) be a non-periodic ergodic system and $\hat{\alpha}$ a finite partition of X . Then for each $\epsilon > 0$, there is a partition α such that the corresponding symbolic representation $(X_\alpha, \mathcal{X}_\alpha, \rho, S)$ is a weakly mixing system whose set of minimal points is not dense, and $d(\alpha, \hat{\alpha}) < \epsilon$.*

Proof. The proof will be conducted by an inductive procedure. We first choose a sequence $\{\epsilon_n\}_{n=0}^\infty$ of positive real numbers with $\sum_{n=0}^\infty \epsilon_n < \epsilon$. Then we start from $\alpha_{-1} = \hat{\alpha}$ and construct $\{\alpha_n\}$ so that $d(\alpha_n, \alpha_{n+1}) < \epsilon_{n+1}$ for $n \geq -1$. Then we define α as the limiting partition; it is easy to see that $d(\hat{\alpha}, \alpha) < \epsilon$. On the one hand, α is constructed so that (X_α, T) is

topologically weakly mixing. On the other hand, almost every point will enter A_2 thickly syndetically so that the set of minimal points is not dense. Now we begin our construction.

STEP 0. Let $\hat{\alpha} = \{\hat{A}_1, \dots, \hat{A}_k\}$. Let ω_0 be the name containing all pairs of names of non-trivial elements in $\bigvee_{i=0}^1 T^{-i}\hat{\alpha}$, where “non-trivial elements” in this proof means elements of positive measure.

Let $M_0 = \min\{\mu(B) : B \in \bigvee_{i=0}^1 T^{-i}\hat{\alpha}\}$ and $0 < \epsilon_0 < \min\{\epsilon/3, M_0/3\}$. Choose $l_0 \in \mathbb{N}$ such that $l_0 > \max\{6/\epsilon_0, 2k^2\}$. Now for a fixed $N_0 > \max\{6l_0/\epsilon_0, 6/M_0\}$, by Lemma 5.1 there is a tower $t(C_0) = \{c_0^1, c_0^2\}$ such that the heights of the columns c_0^1, c_0^2 are $N_0, N_0 + 1$ respectively and the corresponding bases are C_0^1, C_0^2 . It is clear $C_0 = C_0^1 \cup C_0^2$. Set $e_0 = \mu(C_0)$. Copy the name ω_0 on the partial column $\{T^i C_0\}_{0 \leq i \leq |\omega_0| - 1}$. Then in the column c_0^1 , copy 2 to position il_0 for all $0 \leq i \leq (N_0 - 1)/l_0$, and in the column c_0^2 , copy 2 to position il_0 for all $0 \leq i \leq N_0/l_0$.

In such a way we have constructed a new partition $\alpha_0 = \{A_1^0, A_2^0, \dots, A_k^0\}$. Note that $d(\alpha_0, \hat{\alpha}) < \epsilon_0/3$, since the measure change is less than

$$\left(2k^2 + \frac{N_0}{l_0}\right)\mu(C_0) < \left(2k^2 + \frac{N_0}{l_0}\right)\frac{1}{N_0} < \frac{\epsilon_0}{6} + \frac{\epsilon_0}{6} = \frac{\epsilon_0}{3}.$$

Let $A_{i_1}^0, A_{i_2}^0, A_{i_3}^0, A_{i_4}^0 \in \alpha_0$. Assume that the positions of the 2-names (i_1, i_3) and (i_2, i_4) appearing in ω_0 are s and t . Then $T^s C_0 \subseteq A_{i_1}^1$, $T^{s+1} C_0 \subseteq A_{i_3}^1$, $T^r C_0 \subseteq A_{i_2}^0$ and $T^{r+1} C_0 \subseteq A_{i_4}^0$. Hence

$$\begin{aligned} T^s C_0 \times T^r C_0 &\subset (A_{i_1}^0 \cap T^{-1} A_{i_3}^0) \times (A_{i_2}^0 \cap T^{-1} A_{i_4}^0) \\ &= (A_{i_1}^0 \times A_{i_2}^0) \cap (T \times T)^{-1} (A_{i_3}^0 \times A_{i_4}^0). \end{aligned}$$

In particular,

$$\mu \times \mu((A_{i_1}^0 \times A_{i_2}^0) \cap (T \times T)^{-1} (A_{i_3}^0 \times A_{i_4}^0)) \geq \mu(C_0)^2 = e_0^2 > 0.$$

Now assume that inductively we have constructed partitions $\{\alpha_i\}_{i=0}^n$, two sequences $\{l_i\}_{0 \leq i \leq n}$, $\{s_i\}_{0 \leq i \leq n}$ of positive integers, and two sequences $\{\epsilon_i\}_{0 \leq i \leq n}$, $\{e_i\}_{0 \leq i \leq n}$ of positive numbers with $\epsilon_{i+1} < \min\{\epsilon_i/3, e_i^2/3\}$. Also assume that we have obtained a sequence $\{t(C_i^{(j)})\}_{1 \leq i \leq j \leq n}$ of K-R towers with relatively prime heights such that $C_0 \supset C_1^{(1)} \supset \dots \supset C_n^{(n)}$, and the height of $t(C_i^{(n)})$ is in $[N_i, N_i + 6N_{i-1}]$ with some positive integers $\{N_j\}_{1 \leq j \leq n}$.

Let $\alpha_i = \{A_1^i, \dots, A_k^i\}$ for $1 \leq \alpha_i \leq n$. Then $\{\alpha_i\}_{i=1}^n$ has the following properties for each $i \leq n$:

- (1)_i $d(\alpha_{i-1}, \alpha_i) < \epsilon_i$. Let $\bigvee_{j=0}^i T^{-j}\alpha_{i-1} = \{U_1, \dots, U_\eta\}$ with U_j non-trivial. Then there is a subset $\{B_1, \dots, B_\eta\} \subset \bigvee_{j=0}^i T^{-j}\alpha_i$ such that the α_{i-1} -name of U_h and the α_i -name of B_h are the same for each

$1 \leq h \leq \eta$. Moreover, for all $E_1, F_1, E_2, F_2 \in \{U_1, \dots, U_\eta\}$,

$$\mu \times \mu((T \times T)^{s_i}(E_1 \times F_1) \cap (E_2 \times F_2)) > e_i^2 > 0.$$

In particular, for all $E_1, F_1, E_2, F_2 \in \bigvee_{j=0}^{i-1} T^{-j} \alpha_{i-1}$,

$$\mu \times \mu((T \times T)^{s_i}(E_1 \times F_1) \cap (E_2 \times F_2)) > e_i^2 > 0.$$

(2)_i $C_i^{(i)} \subset C_{i-1}^{(i)} \subset \dots \subset C_1^{(i)}$ and for $j \leq i - 1$, $\mu(|\mathfrak{t}(C_j^{(i-1)})| \Delta |\mathfrak{t}(C_j^{(i)})|) < \epsilon$. Refine the towers $\mathfrak{t}(C_j^{(i)})$ according to α_i for each $1 \leq j \leq i$. For each $1 \leq j \leq i$, if a column \mathfrak{c} with base C in the resulting tower $\mathfrak{t}(C_j^{(i)})$ has α_i -name $(a_1, \dots, a_h) \in \Sigma^h$, then the name satisfies

$$(5.1) \quad a_{sl_j+t} = 2 \quad \text{for all } 0 \leq t \leq j \text{ and } 1 \leq s \leq (h - j - 2)/l_j,$$

i.e. $T^{sl_j+t}C \subset A_2^i$.

$$\begin{array}{ccccccc} \mathfrak{t}(C_1^{(1)}) & \mathfrak{t}(C_1^{(2)}) & \mathfrak{t}(C_1^{(3)}) & \cdots & \rightarrow & \mathfrak{t}(C_1^*) & \\ & \mathfrak{t}(C_2^{(2)}) & \mathfrak{t}(C_2^{(3)}) & \cdots & \rightarrow & \mathfrak{t}(C_2^*) & \\ & & \mathfrak{t}(C_3^{(3)}) & \cdots & \rightarrow & \mathfrak{t}(C_3^*) & \\ & & & \cdots & \cdots & \cdots & \end{array}$$

Note that (1)_i will be used to show that X_α is weakly mixing, and (2)_i will be used to show that the minimal points are not dense in X_α .

STEP $n + 1$. Now we explain the inductive step. First we need to define a word ω_{n+1} which contains all pairs of names of non-trivial elements in $\bigvee_{i=0}^{n+1} T^{-i} \alpha_n$. We do it as follows.

Refine the tower $\mathfrak{t}(C_n^{(n)})$ according to α_n , and let the resulting tower be $\mathfrak{t}(C_n^{(n)}) = \{\mathfrak{c}_n^j\}_j$. Note that the height of each column is in $[N_n, N_n + 6N_{n-1}]$. Let $W_{n+2} = \{B_1, \dots, B_t\} \subset \{1, \dots, k\}^{n+2}$ be the set of all names of non-trivial elements of $\bigvee_{i=0}^{n+1} T^{-i} \alpha_n$. Each $(n + 2)$ -word B_j ($j \in \{1, \dots, t\}$) in W_{n+2} either appears in some column $\mathfrak{c}_n^{i_j}$ of $\mathfrak{t}(C_n^{(n)})$, or appears in the concatenation of two columns of $\mathfrak{t}(C_n^{(n)})$ (i.e. there are $\mathfrak{c}_a, \mathfrak{c}_b$ in $\mathfrak{t}(C_n^{(n)})$ such that the name appears in $\mathfrak{c}_a \mathfrak{c}_b$). In the second case we also use $\mathfrak{c}_n^{i_j}$ to denote the concatenation of the two columns. Let \tilde{B}_j be the name of $\mathfrak{c}_n^{i_j}$.

Now fix a large number $s_{n+1} > 10N_n^2$, and construct the word ω_{n+1} as follows: For each pair $(j_1, j_2) \in \{1, \dots, t\}^2$, make sure that the words $\mathfrak{c}_n^{i_{j_1}}$ and $\mathfrak{c}_n^{i_{j_2}}$ appear in ω_{n+1} , and the distance from B_{j_1} to B_{j_2} is s_{n+1} . Since the column heights of $\mathfrak{t}(C_n^{(n)})$ are relatively prime and s_{n+1} is large enough, one can use the α_n -names of the columns $\{\mathfrak{c}_n^i\}_i$ to fill gaps between each pair \tilde{B}_i, \tilde{B}_j .

Let $M_{n+1} < \frac{1}{2} \min_{B \in \bigvee_{i=0}^{n+2} T^{-i} \alpha_n} \mu(B)$ and $\epsilon_{n+1} < \min\{\epsilon_n/3, e_n^2/3\}$. Then let

$$l_{n+1} > \max\{|\omega_{n+1}| + 10N_n^2 + 3N_n, 6n/\epsilon_{n+1}\}$$

and

$$N_{n+1} > \max\{6(n+1)l_{n+1}/\epsilon_{n+1}, (n+3)/M_{n+1}\}.$$

By Proposition 3.2, we have a new K-R tower $\mathfrak{t}(C_{n+1}^{(n+1)})$ with relatively prime column heights and $C_{n+1}^{(n+1)} \subset C_n^{(n)}$, and its height is in $[N_{n+1}, N_{n+1} + 6N_n]$. Refine $\mathfrak{t}(C_{n+1}^{(n+1)})$ according to α_n , and let the resulting tower be $\{\mathfrak{c}_{n+1}^j\}_j$. Let the base of \mathfrak{c}_{n+1}^j be C_{n+1}^j , and let its height be H_j . Let $e_{n+1} = \min_i \mu(C_{n+1}^i)$. Now we make the following adjustment for each column \mathfrak{c}_{n+1}^j .

Denote the name of \mathfrak{c}_{n+1}^j by $(c_1, \dots, c_{H_j}) \in \Sigma^{H_j}$. First, copy the name ω_{n+1} to $(c_h, \dots, c_{|\omega_{n+1}|+h-1})$, where $h > n+3$ is the first number such that $T^{h-1}C_{n+1}^j \subset C_n^{(n)}$. Secondly, choose $R \in \mathbb{N}$ such that

$$l_{n+1} - 2N_n \leq R < l_{n+1}, \quad R - (|\omega_{n+1}| + h - 1) > 10N_n^2, \quad T^{R-1}C_{n+1}^j \subset C_n^{(n)}.$$

Since the column heights of the tower $\mathfrak{t}(C_n^{(n)})$ are relatively prime, we can replace $(c_{|\omega_{n+1}|+1}, \dots, c_{R-1})$ by the names encountered in the tower $\mathfrak{t}(C_n^{(n)})$.

Finally, copy 2 to $c_{sl_{n+1}+r}$ for all $0 \leq r \leq n+1, 1 \leq s \leq (H-n-2)/l_{n+1}$. We define α_{n+1} as the partition corresponding to the new name obtained in the way described above.

Properties of α_{n+1} . Note that by the construction of α_{n+1} , if we refine the tower $\mathfrak{t}(C_{n+1}^{(n+1)})$ according to α_{n+1} , then the resulting tower will still be $\{\mathfrak{c}_{n+1}^j\}_j$. Since the column heights of $\mathfrak{t}(C_n^{(n)})$ are relatively prime, we have ensured that the first length l_{n+1} part of the name along the column in $\{\mathfrak{c}_{n+1}^j\}_j$ consists only of the name encountered in $\mathfrak{t}(C_n^{(n)})$. In the column of $\{\mathfrak{c}_{n+1}^j\}_j$, the levels where the bases of the $\mathfrak{t}(C_n^{(n)})$ name blocks occur may be different. This defines a new base which we call $C_n^{(n+1)}$, and therefore a new K-R tower $\mathfrak{t}(C_n^{(n+1)})$. Since $d(\alpha_n, \alpha_{n+1}) < \epsilon_{n+1}$, changes from $\mathfrak{t}(C_n^{(n)})$ to $\mathfrak{t}(C_n^{(n+1)})$ are very small (less than ϵ_{n+1}). Since we copy 2 to $c_{sl_{n+1}+r}$ for all $0 \leq r \leq n+1$ and $1 \leq s \leq (H-n-2)/l_{n+1}$, each α_{n+1} -name of $\mathfrak{t}(C_n^{(n+1)})$ either has the same name with some column in $\mathfrak{t}(C_n^{(n)})$, or has more 2's than some column name in $\mathfrak{t}(C_n^{(n)})$. Anyway, for each name of length h in $\mathfrak{t}(C_n^{(n+1)})$, at positions $sl_n + t$ with $0 \leq t \leq j, 1 \leq s \leq (h-n-2)/l_n$ the names are 2.

By $(2)_n, C_n^{(n)} \subset C_{n-1}^{(n)} \subset \dots \subset C_1^{(n)}$, the above changes from $\mathfrak{t}(C_n^{(n)})$ to $\mathfrak{t}(C_n^{(n+1)})$ will induce corresponding changes such that $\mathfrak{t}(C_j^{(n)})$ will become some new tower $\mathfrak{t}(C_j^{(n+1)})$ for each $1 \leq j \leq n-1$, where $C_{n+1}^{(n+1)} \subset C_n^{(n+1)} \subset \dots \subset C_1^{(n+1)}$. For the same reason as for $\mathfrak{t}(C_n^{(n+1)})$, equality (5.1) holds for each $j \leq n+1$. Thus we have $(2)_{n+1}$.

Now we verify that α_{n+1} satisfies $(1)_{n+1}$.

By the construction, the measure change from α_n to α_{n+1} is less than

$$\begin{aligned} \mu(C_{n+1}) & \left(|\omega_{n+1}| + (n+2) \frac{N_{n+1} + 6N_n}{l_{n+1}} \right) \\ & < \frac{1}{N_{n+1}} \left(l_{n+1} + (n+2) \frac{2N_{n+1}}{l_{n+1}} \right) \\ & < \frac{1}{6n+1} \epsilon_{n+1} + \frac{2(n+1)}{l_{n+1}} < \frac{\epsilon_{n+1}}{2} + \frac{\epsilon_{n+1}}{2} = \epsilon_{n+1}. \end{aligned}$$

Thus we conclude that $d(\alpha_n, \alpha_{n+1}) < \epsilon_{n+1}$. And the second part of $(1)_{n+1}$ is guaranteed by the construction of ω_{n+1} .

Let $D_{i_1}, D_{i_2}, D_{j_1}, D_{j_2} \in \bigvee_{i=0}^{n+1} T^{-i} \alpha_n$, and let their names be $B_{i_1}, B_{i_2}, B_{j_1}, B_{j_2} \in W_{n+2}$ respectively, where $1 \leq i_1, i_2, j_1, j_2 \leq t$. Then by the definition of ω_{n+1} , the pairs (B_{i_1}, B_{j_1}) and (B_{i_2}, B_{j_2}) appear in the word ω_{n+1} . Given an arbitrary column C_{n+1}^i with base C_{n+1}^i , let p be the position of B_{i_1} in this column and let r be the distance from the position of B_{i_1} to the position of B_{i_2} . Then

$$\begin{aligned} T^{p-1} C_{n+1}^i & \subset D_{i_1}, & T^{p-1+s_{n+1}} C_{n+1}^i & \subset D_{j_1}, \\ T^{p-1+r} C_{n+1}^i & \subset D_{i_2}, & T^{p-1+r+s_{n+1}} C_{n+1}^i & \subset D_{j_2}. \end{aligned}$$

It follows that

$$\begin{aligned} T^{p-1} C_{n+1}^i \times T^{p-1+r} C_{n+1}^i & \subset (D_{i_1} \cap T^{-s_{n+1}} D_{j_1}) \times (D_{i_2} \cap T^{-s_{n+1}} D_{j_2}) \\ & = (D_{i_1} \times D_{i_2}) \cap (T \times T)^{-s_{n+1}} (D_{j_1} \times D_{j_2}). \end{aligned}$$

Hence

$$\begin{aligned} \mu \times \mu((D_{i_1} \times D_{i_2}) \cap (T \times T)^{-s_{n+1}} (D_{j_1} \times D_{j_2})) \\ \geq \mu \times \mu(T^{p-1} C_{n+1}^i \times T^{p-1+r} C_{n+1}^i) \geq e_{n+1}^2 > 0. \end{aligned}$$

Thus $(1)_{n+1}$ holds.

Properties of α . Let $\alpha = \{A_1, \dots, A_k\}$ be the limiting partition of the sequence $\{\alpha_n\}$. It is clear that

$$d(\hat{\alpha}, \alpha) < \sum_{i=0}^{\infty} \epsilon_i < \epsilon.$$

Also by $(2)_n$, for $n \geq 1$ the sequence $\{t(C_n^{(j)})\}_{j \geq n}$ has a limit tower $t(C_n^*)$ with base C_n^* . And by $(2)_n$, $C_1^* \supset C_2^* \supset \dots$.

We will now show that α satisfies the assertion of the proposition. First we will show that:

(1) For each $m \geq 0$ and $E_1, F_1, E_2, F_2 \in \bigvee_{j=0}^{m-1} T^{-j} \alpha$, we have

$$\mu \times \mu((T \times T)^{sm} (E_1 \times F_1) \cap (E_2 \times F_2)) > 0.$$

- (2) Refine the towers $\mathfrak{t}(C_j^*)$ according to α for each $j \geq 1$. If a column \mathfrak{c} in the resulting tower $\mathfrak{t}(C_j^*)$ has α -name $(a_1, \dots, a_h) \in \Sigma^h$ and its base is C , then the name satisfies

$$(5.2) \quad a_{sl_j+t} = 2 \quad \text{for all } 0 \leq t \leq j \text{ and } 1 \leq s \leq (h-j-2)/l_j,$$

i.e. $T^{sl_j+t}C \subset A_2$.

Condition (2) is guaranteed by (2)_n. It remains to verify (1). By (1)_m there are $E'_1, F'_1, E'_2, F'_2 \in \bigvee_{j=0}^{m-1} T^{-j}\alpha_{m-1}$ that have the same names as E_1, F_1, E_2, F_2 respectively. By (1)_m,

$$\mu \times \mu((T \times T)^{sm}(E'_1 \times F'_1) \cap (E'_2 \times F'_2)) > \epsilon_m^2.$$

Then $d(\alpha_m, \alpha) < \sum_{j=m+1}^\infty \epsilon_j$, one has

$$\mu \times \mu((T \times T)^{sm}(E_1 \times F_1) \cap (E_2 \times F_2)) > \epsilon_m^2 - \sum_{j=m+1}^\infty \epsilon_j > 0.$$

Now using conditions (1) and (2) we will show that α is as required. Let X_α be the corresponding symbolic representation of α , and $\phi : X \rightarrow X_\alpha$ be the factor map. Let $[i]_0 = \{w \in X_\alpha : w_0 = i\}$ for $i \in \{1, \dots, k\}$. Let $w = \phi(x) \in [1]_0$ be a transitive point of (X_α, T) .

By property (1), (X_α, T) is weakly mixing. By (2), $N(w, [2]_0)$ is thickly syndetic, which implies that $N(w, [1]_0)$ is not piecewise syndetic. Hence the set of minimal points of (X_α, T) is not dense. ■

Similar to Lemma 4.5 we have the following easy observation.

LEMMA 5.3. *Let (X, T) be the inverse limit of $\{(X_n, T_n)\}_n$, where each (X_n, T_n) is a non-minimal topologically weakly mixing system whose set of minimal points is not dense. Then (X, T) is also a non-minimal topologically weakly mixing system whose set of minimal points is not dense.*

Using a similar argument to the one used to obtain Proposition 4.4 from Proposition 4.2, and adjusting the proof of Proposition 5.2, we deduce the following result.

PROPOSITION 5.4. *Every non-periodic ergodic system has a topological model which is a weakly mixing system with full support and the set of minimal points is not dense.*

Proof. The idea of the proof is similar to the one used in the proof of Proposition 4.4. We will show that there exists an increasing sequence $\{\gamma_n\}$ of finite partitions such that $\sigma(\gamma_n) \nearrow \mathcal{X}$ and for each $n \in \mathbb{N}$ the corresponding symbolic representation $(X_{\gamma_n}, \mathcal{X}_{\gamma_n}, \rho_n, S)$ is a weakly mixing system with full support and the set of minimal points is not dense. This by Lemma 5.3 will finish the proof.

Let (X, \mathcal{X}, μ, T) be an ergodic system. Let $\{\beta_n\}_{n \geq 0}$ be an increasing sequence of finite partitions such that $\sigma(\beta_n) \nearrow \mathcal{X}$, and let $\{\epsilon_n\}$ be a sequence of positive numbers with $\sum_{n=0}^{\infty} \epsilon_n < \infty$. We will modify the proof of Proposition 5.2 carefully to get what we need.

As in the proof of Proposition 5.2 we choose a tower $t(C_0^0)$, and adjust β_0 by Step 0 to get a new partition γ_0^0 . We replace β_1 by $\beta_1 \vee \gamma_0^0$ (still denoting it by β_1); it is clear that $\gamma_0^0 \prec \beta_1$. We assume that the first (resp. second) element of β_1 is a subset of the first (resp. second) element of γ_0^0 .

As in Step 1 of the proof of Proposition 5.2, we modify β_1 to get a new partition γ_1^1 . We then construct a tower $t(C_1^1)$ using Lemma 3.2, and form a new tower $t(C_0^1)$.

By the Fact in the proof of Proposition 4.4, we construct $\gamma_0^1 \prec \gamma_1^1$. Refining γ_0^1 to $t(C_0^1)$, we know that γ_0^1 satisfies $(1)_1, (2)_1$ in Step 1 of the proof of Proposition 5.2 since $\beta_1 \succ \gamma_0^0$.

Inductively, we replace β_n by $\beta_n \vee \gamma_{n-1}^{n-1}$, and we assume that the first (resp. second) element of β_n is a subset of the first (resp. second) element of γ_{n-1}^{n-1} .

We modify β_n by Step n to get a new partition γ_n^n such that $d(\beta_n, \gamma_n^n) < \epsilon_n$, and by the same argument we find that γ_n^n has the same properties listed in $(1)_n$ and $(2)_n$ for the tower $t(C_n^{(n)})$. Now construct $\gamma_k^n \prec \gamma_n^n$ by Lemma 4.3. Since the first and second elements of β_n are subsets of the first and second elements of γ_{n-1}^{n-1} respectively, and $\beta_n \succ \gamma_{n-1}^{n-1}$, we conclude that γ_k^n has the same properties listed in $(1)_n$ and $(2)_n$ for the tower $t(C_j^{(n)})$, $k \leq j \leq n$. By the proof of Proposition 5.2, the partition $\gamma_k = \lim_n \gamma_k^n$ has properties $(1), (2)$ there. Hence according to the proof of Proposition 5.2, X_{γ_k} is a weakly mixing system with full support and the set of minimal points is not dense.

Following the same discussion as in the proof of Proposition 4.4, we conclude that $\{\gamma_k\}$ is increasing and $\sigma(\gamma_k) \nearrow \mathcal{X}$. ■

Now using Proposition 5.4, we are able to finish the proof of Theorem 1.1(2).

Proof of Theorem 1.1(2). For a given ergodic system (X, \mathcal{X}, μ, T) , by Proposition 5.4, there is a topological model (Y, S) of X with an ergodic measure ρ , which is weakly mixing, non-minimal and the set $\text{Min}(Y)$ of minimal points is not dense in $\text{supp}(\rho) = Y$. Note that $\rho(\overline{\text{Min}(Y)}) = 0$, since $\overline{\text{Min}(Y)}$ is an S -invariant set.

Define an equivalence relation \sim in Y as follows: $x \sim y$ if $x, y \in \overline{\text{Min}(Y)}$. Then the quotient system $(\hat{X} = Y/\sim, \hat{T})$ is measure-theoretically isomorphic to (Y, S) since $\rho(\overline{\text{Min}(Y)}) = 0$. Hence (\hat{X}, \hat{T}) is also a topological model of (X, \mathcal{X}, μ, T) . Note that (\hat{X}, \hat{T}) is a topologically weakly mixing system with

full support and a unique fixed point as the only minimal point. Thus the proof is complete. ■

6. Applications. In this section we give two applications of the results we have obtained.

Let (X, T) be a topological dynamics and $M(X)$ the collection of all Borel probability measures on X with the weak* topology. Then T naturally induces a map T_M on $M(X)$ by sending $\mu \in M(X)$ to $T\mu$. Let $C(X, \mathbb{R})$ be the Banach space of continuous real-valued functions on X with the supremum norm $\|\cdot\|$. Let $\{f_n\}_{n=1}^\infty$ be a dense subset of $C(X, \mathbb{R})$. Then

$$d(\mu, \nu) = \sum_{n=1}^{\infty} \frac{|\int f_n d\mu - \int f_n d\nu|}{2^n(\|f_n\| + 1)}$$

is a metric on $M(X)$ giving the weak* topology. For more details about $(M(X), T_M)$, see [10].

An unsolved question in [10] is whether there is a weakly mixing proximal system (X, T) such that $(M(X), T_M)$ has dense minimal points. We give an affirmative answer to this question:

THEOREM 6.1. *There is a weakly mixing proximal system (X, T) such that $(M(X), T_M)$ has dense minimal points.*

To show this result we need a lemma from [10].

LEMMA 6.2. *Let X, Y be compact metric spaces, $\mu \in M(X)$ and $\nu \in M(Y)$.*

(1) *If $A = \bigcup_{i=1}^n A_i$, where A_1, \dots, A_n are Borel subsets of X with $\mu(A_i) > 0$ and $\mu(A_i \cap A_j) = 0$ for all $1 \leq i < j \leq n$, then*

$$\mu_A = \sum_{i=1}^n \frac{\mu(A_i)}{\mu(A)} \mu_{A_i}.$$

(2) *Let $\epsilon > 0$ and A be a Borel subset of X with $\mu(A) > 0$. If B is a Borel subset of X such that $\mu(B) > 0$ and $\mu(A \Delta B) < \mu(A) \cdot \epsilon$, then $d(\mu_A, \mu_B) \leq 2\epsilon$.*

(3) *If $\pi : (X, \mu) \rightarrow (Y, \nu)$ is measurable and $\pi\mu = \nu$, then $\pi\mu_{\pi^{-1}A} = \nu_A$ for each Borel subset A of Y .*

Proof of Theorem 6.1. Let (Σ_2, T) be the dyadic adding machine with a unique ergodic measure μ . By Theorem 1.1, (Σ_2, T, μ) is isomorphic to (Y, S, ν) , where (Y, S) is a weakly mixing proximal topological system and ν has full support. We now show that $P(S_M)$, the set of periodic points of $(M(Y), S_M)$, is dense.

Let $\pi : (\Sigma_2, T, \mu) \rightarrow (Y, S, \nu)$ be an isomorphism, that is, there are invariant Borel subsets $X_1 \subset X$ and $X_2 \subset Y$ with $\mu(X_1) = \nu(X_2) = 1$ and

an invertible measure-preserving transformation $\pi : X_1 \rightarrow X_2$ such that $\pi(Tx) = S\pi(x)$ for all $x \in X_1$.

Let $\epsilon > 0$ and let U be a non-empty open subset of Y . Since ν has full support, we have $\nu(U) > 0$. Thus, there are finitely many pairwise disjoint cylinders A_1, \dots, A_k of X such that $\mu(\pi^{-1}U \triangle A) < \nu(U) \cdot \epsilon$ with $A = \bigcup_{i=1}^k A_i$, which implies $\nu(U \triangle \pi(A \cap X_1)) < \nu(U) \cdot \epsilon$. Lemma 6.2(2) yields $d(\nu_U, \nu_{\pi(A \cap X_1)}) \leq 2\epsilon$. Since $T^{2|C|}C = C$ for each cylinder C of X , where $|C|$ stands for the length of C , we conclude that μ_C is periodic. In particular, each μ_{A_i} is periodic. By Lemma 6.2(3), each $\nu_{\pi(A_i \cap X_1)}$ is also periodic. By Lemma 6.2(1), $\nu_{\pi(A \cap X_1)} = \sum_{i=1}^k p_i \nu_{\pi(A_i \cap X_1)}$, where $p_i = \mu(A_i)/\mu(A)$. Thus, $\nu_{\pi(A \cap X_1)}$ is periodic. It follows that ν_U can be approximated by periodic points of $(M(Y), S_M)$.

Now take $y \in Y$ and let $\{U_n\}_{n=1}^\infty$ be a sequence of open neighborhoods of y such that $\text{diam}(U_n) \rightarrow 0$. For any $f \in C(Y, \mathbb{R})$, we have

$$\left| \int_Y f(z) d\nu_{U_n} - f(y) \right| \leq \int_{U_n} |f(z) - f(y)| d\nu_{U_n} \rightarrow 0.$$

A simple calculation shows $\nu_{U_n} \rightarrow \delta_y$, and hence δ_y is a limit point of $P(S_M)$. This implies that each element of $M_n(Y) = \{n^{-1} \sum_{i=1}^n \delta_{x_i} : x_i \in X\}$ can be approximated by elements of $P(S_M)$. Since $\bigcup_{n=1}^\infty M_n(Y)$ is dense in $M(Y)$, it follows that $P(S_M)$ is dense. ■

Another application of our result is the following. A topological analogue of K-systems, called a *topological K-system*, was studied in [8]. In [5] the authors constructed a proximal topological K-system which is weakly mixing. Using Theorem 1.1, we can get a lot of such examples which are strongly mixing.

THEOREM 6.3. *There exist strongly mixing proximal topological K-systems.*

Proof. Let (X, T, μ) be a measurable K-system. By Theorem 1.1, (X, T, μ) is isomorphic to a proximal system (Y, S) with a measure ν of full support. Thus (Y, S) is strongly mixing, since a K-system is strongly mixing in the measurable sense. At the same time we know that (Y, S) is a topological K-system by [8, Theorem 3.4]. ■

Acknowledgments. We would like to thank Wen Huang for useful suggestions. We would like to thank the referee for a thorough reading and many helpful corrections and suggestions concerning this paper.

This research is supported by NNSF of China (11171320, 11371339, 11431012, 11571335) and by the Fundamental Research Funds for the Central Universities.

References

- [1] H. Furstenberg, *Disjointness in ergodic theory, minimal sets, and a problem in Diophantine approximation*, Math. Systems Theory 1 (1967), 1–49.
- [2] E. Glasner, *Ergodic Theory via Joinings*, Math. Surveys Monogr. 101, Amer. Math. Soc., Providence, RI, 2003.
- [3] E. Glasner and B. Weiss, *On the interplay between measurable and topological dynamics*, in: Handbook of Dynamical Systems. Vol. 1B, Elsevier, Amsterdam, 2006, 597–648.
- [4] W. Gottschalk and G. Hedlund, *Topological Dynamics*, Amer. Math. Soc., Providence, 1955.
- [5] W. Huang, H. F. Li and X. D. Ye, *Family-independence for topological and measurable dynamics*, Trans. Amer. Math. Soc. 364 (2012), 5209–5242.
- [6] W. Huang, S. Shao and X. D. Ye, *Pointwise convergence of multiple ergodic averages and strictly ergodic models*, arXiv:1406.5930 [math.DS] (2014).
- [7] W. Huang and X. D. Ye, *Dynamical systems disjoint from all minimal systems*, Trans. Amer. Math. Soc. 357 (2005), 669–694.
- [8] W. Huang and X. D. Ye, *A local variational relation and applications*, Israel J. Math. 151 (2006), 237–280.
- [9] E. Lehrer, *Topological mixing and uniquely ergodic systems*, Israel J. Math. 57 (1987), 239–255.
- [10] J. Li, K. Yan and X. D. Ye, *Recurrence properties and disjointness on the induced spaces*, arXiv:1312.2056v1 [math.DS] (2013).
- [11] A. Rosenthal, *On strictly ergodic models for commuting ergodic transformations*, Ann. Inst. H. Poincaré Probab. Statist. 25 (1989), 73–92.
- [12] B. Weiss, *Strictly ergodic models for dynamical systems*, Bull. Amer. Math. Soc. (N.S.) 13 (1985), 143–146.
- [13] B. Weiss, *Countable generators in dynamics—Universal minimal models*, in: Contemp. Math. 94, Amer. Math. Soc., 1989, 321–326.
- [14] B. Weiss, *Multiple recurrence and doubly minimal systems*, in: Topological Dynamics and Applications (Minneapolis, MN, 1995), Contemp. Math. 215, Amer. Math. Soc., Providence, RI, 1998, 189–196.
- [15] B. Weiss, *Single Orbit Dynamics*, CBMS Reg. Conf. Ser. Math. 95, Amer. Math. Soc., Providence, RI, 2000.

Zhengxing Lian, Song Shao, Xiangdong Ye
Wu Wen-Tsun Key Laboratory of Mathematics
USTC, Chinese Academy of Sciences
and
Department of Mathematics
University of Science and Technology of China
Hefei, Anhui, 230026, P.R. China
E-mail: lianzx@mail.ustc.edu.cn
songshao@ustc.edu.cn
yexd@ustc.edu.cn

