J. Eur. Math. Soc. 27, 279–331 (2025) DOI 10.4171/JEMS/1379



Eli Glasner · Wen Huang · Song Shao · Benjamin Weiss · Xiangdong Ye

# Topological characteristic factors and nilsystems

Received May 20, 2021; revised December 8, 2021

**Abstract.** We prove that the maximal infinite step pro-nilfactor $X_\infty$ of a minimal dynamical system $(X, T)$ is the topological characteristic factor in a certain sense. Namely, we show that by an almost one-to-one modification of $\pi : X \to X_\infty$, the induced open extension $\pi^* : X^* \to X^*_\infty$ has the following property: for $x$ in a dense $G_\delta$ subset of $X^*$, the orbit closure $L_x = \overline{\mathcal{O}}((x, \dots, x), T \times T^2 \times \cdots \times T^d)$ is $(\pi^*)^{(d)}$-saturated, i.e., $L_x = ((\pi^*)^{(d)})^{-1}(\pi^*)^{(d)}(L_x)$. Using results derived from the above fact, we are able to answer several open questions: (1) if $(X, T^k)$ is minimal for some $k \geq 2$, then for any $d \in \mathbb{N}$ and any $0 \leq j < k$ there is a sequence $\{n_i\}$ of $\mathbb{Z}$ with $n_i \equiv j \pmod k$ such that $T^{n_i} x \to x, T^{2n_i} x \to x, \dots, T^{dn_i} x \to x$ for $x$ in a dense $G_\delta$ subset of $X$; (2) if $(X, T)$ is totally minimal, then $\{T^{n^2} x : n \in \mathbb{Z}\}$ is dense in $X$ for $x$ in a dense $G_\delta$ subset of $X$; (3) for any $d \in \mathbb{N}$ and any minimal t.d.s. which is an open extension of its maximal distal factor, $\mathbf{RP}^{[d]} = \mathbf{AP}^{[d]}$, where the former is the regionally proximal relation of order $d$ and the latter is the regionally proximal relation of order $d$ along arithmetic progressions.

**Keywords.** Multiple recurrence, maximal equicontinuous factor

## 1. Introduction

In this introductory section we will provide some background related to the notion of a characteristic factor of a topological dynamical system, present some open questions, state our main results, and finally explain the main ideas of the proofs. In the paper, $\mathbb{Z}$ is the set of integers and $\mathbb{N} = \{1, 2, \dots\}$ is the set of positive integers.

Eli Glasner: Department of Mathematics, Tel Aviv University, Ramat Aviv, 69978 Tel Aviv, Israel; glasner@math.tau.ac.il

Wen Huang: CAS Wu Wen-Tsun Key Laboratory of Mathematics, University of Science and Technology of China, Hefei, 230026, P.R. China; wenh@mail.ustc.edu.cn

Song Shao: Department of Mathematics, University of Science and Technology of China, Hefei, 230026, P.R. China; songshao@ustc.edu.cn

Benjamin Weiss: Institute of Mathematics, Hebrew University of Jerusalem, Edmond J. Safra Campus, Givat Ram, 91904 Jerusalem, Israel; weiss@math.huji.ac.il

Xiangdong Ye: CAS Wu Wen-Tsun Key Laboratory of Mathematics, University of Science and Technology of China, Hefei, 230026, P.R. China; yexd@ustc.edu.cn

*Mathematics Subject Classification (2020)*: Primary 37B05; Secondary 37B20, 37A17, 37A46

### *1.1. Backgrounds*

*1.1.1. Characteristic factors.* A connection between ergodic theory and additive combinatorics was established in the 1970's with Furstenberg's elegant proof of Szemerédi's theorem via ergodic theory. Furstenberg [18] proved Szemerédi's theorem by means of the following theorem: Let $T$ be a measure preserving transformation (m.p.t. for short) on the Borel probability space $(X, \mathcal{X}, \mu)$. Then for every $d \geq 1$ and $A \in \mathcal{X}$ with positive measure,

$$\liminf_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \mu(A \cap T^{-n}A \cap T^{-2n}A \cap \cdots \cap T^{-dn}A) > 0. \tag{1.1}$$

In view of this theorem it is natural to ask about the convergence of these averages; or more generally, about the convergence, either in $L^2(X, \mu)$ or pointwise, of the *multiple ergodic averages*

$$\frac{1}{N} \sum_{n=0}^{N-1} f_1(T^n x) \cdots f_d(T^{dn} x), \tag{1.2}$$

where $f_1, \ldots, f_d \in L^\infty(X, \mu)$. After nearly 30 years' efforts of many researchers, this problem (for $L^2$ convergence) was finally solved in [28, 43].

In the study of the averages (1.2), the idea of characteristic factors plays an important role. For the origin of these ideas and this terminology, see [18, 21]. To be more precise, let $(X, \mathcal{X}, \mu, T)$ be a measure preserving system (m.p.s.) and $(Y, \mathcal{Y}, \nu, T)$ be a factor of $X$. For $d \geq 1$, we say that $Y$ is a *characteristic factor* of $X$ if for all $f_1, \ldots, f_d \in L^\infty(X, \mu)$,

$$\frac{1}{N} \sum_{n=0}^{N-1} f_1(T^n x) \cdots f_d(T^{dn} x) - \frac{1}{N} \sum_{n=0}^{N-1} \mathbb{E}(f_1|\mathcal{Y})(T^n x) \cdots \mathbb{E}(f_d|\mathcal{Y})(T^{dn} x) \to 0$$

in $L^2(X, \mu)$.

Finding a good characteristic factor for certain schemes of averages often yields a useful reduction of the problem of evaluating their limit behavior. For example, Furstenberg [18] proved that for each $d \geq 2$, the $(d-1)$-step measurable distal factor (in the structure theorem of an ergodic m.p.s.) is a characteristic factor for (1.2). The result in [28, 43] improves the result of Furstenberg significantly, i.e., they show that for each $d \geq 2$, a $(d-1)$-step pro-nilsystem is a characteristic factor for (1.2).

By a *topological dynamical system* $(X, T)$ (t.d.s. for short) we mean a homeomorphism $T$ from a compact metric space $X$ to itself. A counterpart of the notion of characteristic factors in a t.d.s. was first studied in 1994 by Glasner [23]. There, the author studied the characteristic factors for the transformation $\tau_d = T \times T^2 \times \cdots \times T^d$ in the sense of *saturation*: Let $\pi : X \to Y$ be a map between two sets $X$ and $Y$. A subset $L$ of $X$ is called *$\pi$-saturated* if $\{x \in L : \pi^{-1}(\pi(x)) \subseteq L\} = L$, i.e., $L = \pi^{-1}(\pi(L))$. Given a factor map $\pi : (X, T) \to (Y, T)$ and $d \geq 2$, the t.d.s. $(Y, T)$ is said to be a *$d$-step topological characteristic factor (for $\tau_d$) of $(X, T)$* if there exists a dense $G_\delta$ subset $\Omega$ of $X$ such that for each $x \in \Omega$ the orbit closure $L_x = \overline{\mathcal{O}}((x, \ldots, x), \tau_d)$ is $\pi \times \cdots \times \pi$ ($d$ times) saturated.

In [23], it was shown that for minimal systems, up to a canonically defined proximal extension, a characteristic family for $\tau_d$ is the family of canonical PI flows of class $d-1$. In particular, if $(X,T)$ is distal, then its largest class $d-1$ distal factor (in the structure theorem of Furstenberg [17]) is its topological characteristic factor for $\tau_d$. Moreover, if $(X,T)$ is weakly mixing, then the trivial system is its topological characteristic factor.

As in the ergodic situation, in topological dynamics one expects that the largest class of $d-1$ distal factor can be replaced by the $(d-1)$-step pro-nilfactor. So, based on the result of [23] and the parallelism between ergodic theory and topological dynamical systems, one naturally asks:

**Question 1.** *Assume that $(X,T)$ is a minimal t.d.s. and $d \geq 2$. Is it true that its maximal $(d-1)$-step pro-nilfactor is a topological characteristic factor for $\tau_d$?*

*1.1.2. Odd recurrence.* It is easy to see that one consequence of (1.1) is the following *multiple ergodic recurrence theorem* (MERT for short): if $(X,\mathcal{X},\mu,T)$ is a m.p.s., then for each $d \geq 1$ and $A \in \mathcal{X}$ with $\mu(A) > 0$ there is $n \in \mathbb{N}$ such that

$$\mu(A \cap T^{-n}A \cap \cdots \cap T^{-dn}A) > 0. \tag{1.3}$$

As an immediate application of MERT, if $(X,T)$ is minimal then for each $d \geq 1$ and each non-empty open subset $U$ of $X$, there is $n \in \mathbb{N}$ such that

$$U \cap T^{-n}U \cap \cdots \cap T^{-dn}U \neq \emptyset. \tag{1.4}$$

We will refer to this property as the *topological multiple recurrence theorem* (TMRT, for short). For topological proofs of the TMRT see [2, 4, 19, 20]. It is easy to see that TMRT is equivalent to the following statement: if $(X,T)$ is minimal and $d \geq 1$, then there is a dense $G_\delta$ subset $\Omega$ of $X$ such that for each $x \in \Omega$ there is an increasing sequence $\{n_i\}$ in $\mathbb{N}$ with

$$T^{n_i}x \to x, \quad T^{2n_i}x \to x, \quad \ldots, \quad T^{dn_i}x \to x. \tag{1.5}$$

We note that TMRT, or (1.5), is also equivalent to the well known van der Waerden theorem: if $r \geq 1$ and $\mathbb{N} = N_1 \cup \cdots \cup N_r$ then one of the sets $N_i$ contains arbitrarily long arithmetic progressions.

There are several ways in which one can generalize (1.3) and (1.4). The first one is to extend these properties to nilpotent group actions (there are counterexamples for solvable groups [3]). For this type of results we refer to [2, 33, 34] and the references therein.

Another way is to restrict $n$ to a particular congruence class: $n \equiv j \pmod{k}$ for a given $k \geq 2$ and $0 \leq j < k$; or to other subsets of $\mathbb{N}$, for example to the set of primes. Host and Kra [27] (for $d \leq 3$) and Frantzikinakis [14, Corollary 6.5] (for the general $d$) showed that if $(X,\mathcal{X},\mu,T)$ is a m.p.s. and $T^k$ is ergodic for some $k \geq 2$, then for any $d \geq 1$, any $A \in \mathcal{X}$ with $\mu(A) > 0$ and any $0 \leq j < k$, we have $\mu(A \cap T^{-n}A \cap \cdots \cap T^{-dn}A) > 0$ for some $n \equiv j \pmod{k}$.

In view of the results of Host, Kra and Frantzikinakis the following well known question has been open till now.

**Question 2.** *Let $(X, T^k)$ be minimal for some $k \geq 2$ and let $d \geq 1$. Is it true that for any non-empty open subset $U$ of $X$ and $0 \leq j < k$ one has*

$$U \cap T^{-n}U \cap \cdots \cap T^{-dn}U \neq \emptyset \tag{1.6}$$

*for some $n \equiv j \pmod k$?*

Since the fact that $(X, T^k)$ is minimal for some $k \geq 2$ does not imply that there is a Borel invariant probability measure $\mu$ with $(X, \mathcal{X}, T^k, \mu)$ ergodic, Question 2 cannot be answered by using ergodic results. We remark that if $(X, T)$ is minimal and weakly mixing then Question 2 has an affirmative answer [23, 32].

*1.1.3. Density problems.* In ergodic theory there are many results stating that the time averages are equal to the spatial averages under various ergodicity assumptions. For example, the von Neumann mean ergodic theorem tells us that if $(X, \mathcal{X}, \mu, T)$ is ergodic, then for each $f \in L^2(X, \mu)$, one has $\frac{1}{N}\sum_{n=1}^{N} f(T^n x) \to \int f\, d\mu$ in $L^2(X, \mu)$ as $N \to \infty$. The corresponding topological statement is the following: if $(X, T)$ is a transitive t.d.s., then there is a dense $G_\delta$ subset $\Omega$ of $X$ such that each $x \in \Omega$ has a dense orbit.

Furstenberg [19] (for $L^2$) and Bourgain [6] (pointwise for general $p$) have shown that if $(X, \mathcal{X}, \mu, T)$ is totally ergodic, then for each $f \in L^p(X, \mu)$ with $p > 1$ and each non-constant integral polynomial $P(n)$, we have

$$\frac{1}{N}\sum_{n=1}^{N} f(T^{P(n)})x \to \int f\, d\mu \quad \text{in } L^p(X, \mu). \tag{1.7}$$

As not every minimal t.d.s. admits a totally ergodic measure, the following question is natural.

**Question 3.** *Let $(X, T)$ be totally minimal and $P(n)$ be a non-constant integral polynomial. Is it true that $\{T^{P(n)}x : n \in \mathbb{Z}\}$ is dense in $X$ for $x$ in a dense $G_\delta$ subset of $X$?*

We note that the total minimality assumption is necessary for the above question. Let $X$ be a periodic orbit of period 3; then $(X, T^2)$ is minimal but it is easy to check that $\{T^{n^2}x : n \in \mathbb{Z}\}$ is not dense in $X$ for any $x \in X$.

A more challenging problem is whether one can replace the polynomial times by the set of primes in the above question. A convergence similar to (1.7) has been proved to be true in ergodic theory due to Vinogradov [41]: under the total ergodicity assumption, for all $f \in L^2(X, \mu)$,

$$\lim_{N\to\infty} \frac{1}{\pi(N)} \sum_{p \leq N,\, p \text{ prime}} T^p f = \int f\, d\mu \quad \text{in } L^2(X, \mu),$$

where $\pi(N)$ denotes the number of primes less than or equal to $N$. See [5, 15] for more information.

*1.1.4. Regionally proximal relations of higher order.* Finally, we proceed to give the background of the last problem.

The notion of the regionally proximal relation $\mathbf{RP}^{[1]}$ is an important tool in the study of a t.d.s. For a minimal t.d.s., when the acting group is amenable, it is known that it is a closed equivalence relation and that $X/\mathbf{RP}^{[1]}$ is the maximal equicontinuous factor [35]. In [28] Host and Kra introduced very useful new tools, like the so-called Gowers–Host–Kra seminorms, the $\mathcal{G}^{[d]}$-actions ($d \geq 1$) etc., to construct a (pro-nilsystem) factor $Z_{d-1}$, and to show that it is the characteristic factor for the averages (1.2).

To get the corresponding factors in a t.d.s., in the pioneering work [29] Host, Kra and Maass introduced the notion of the *regionally proximal relation of order d* ($d \geq 1$), denoted by $\mathbf{RP}^{[d]}$, and proved that for a minimal distal $\mathbb{Z}$-system, $\mathbf{RP}^{[d]}$ is an equivalence relation and $X/\mathbf{RP}^{[d]}$ is a pro-nilsystem of order $d$. Later, Shao and Ye [37] showed that $\mathbf{RP}^{[d]}$ is an equivalence relation for arbitrary minimal systems of abelian groups. See Glasner, Gutman and Ye [25] for the case of general group actions.

In [29] the authors study $\mathbf{RP}^{[d]}$ through the so-called *dynamical parallelepiped of dimension d*, $\mathbf{Q}^{[d+1]}(X)$. Since the average in (1.2) is only related to $\tau_d$, it is natural to define a kind of regionally proximal relation of higher order, by using $\tau_d$ directly. In [26] the authors followed this direction by introducing a notion, called regionally proximal relation of order $d$ along arithmetic progressions, denoted by $\mathbf{AP}^{[d]}$. Among other things, the authors proved that under some additional assumptions, for a uniquely ergodic minimal distal system, one has $\mathbf{RP}^{[d]} = \mathbf{AP}^{[d]}$ for every $d \geq 1$. They posed the following conjecture:

**Conjecture 1.1 of [26].** *Let $(X,T)$ be a minimal distal system. Then $\mathbf{RP}^{[d]} = \mathbf{AP}^{[d]}$ for $d \geq 1$.*

### *1.2. The main results*

In this subsection we state our main results. For a minimal t.d.s. $(X,T)$ and $d \geq 1$ we use $\mathbf{RP}^{[d]}$ to denote the regionally proximal relation of order $d$, and set $\mathbf{RP}^{[\infty]} = \bigcap_{d\geq 1} \mathbf{RP}^{[d]}$. Let $X_d = X/\mathbf{RP}^{[d]}$ for $d \in \mathbb{N} \cup \{\infty\}$. Then $X_1$ is the maximal equicontinuous factor of $X$.

The following theorem says that for a minimal t.d.s. and $d \geq 2$, up to almost one-to-one extensions, its maximal $(d-1)$-step pro-nilfactor is a topological characteristic factor for $\tau_d$.

**Theorem A.** *Let $(X,T)$ be a minimal t.d.s. and $X_\infty$ be the maximal $\infty$-step pro-nilfactor of $X$, and let $\pi : X \to X_\infty$ be the factor map. Then there are minimal t.d.s. $X^*$ and $X_\infty^*$ which are almost one-to-one extensions of $X$ and $X_\infty$ respectively, and a commuting diagram below such that $X_\infty^*$ is a $d$-step topological characteristic factor of $X^*$ for all $d \geq 2$.*

$$\begin{array}{ccc} X & \xleftarrow{\ \sigma^*\ } & X^* \\ \Big\downarrow{\scriptstyle \pi} & & \Big\downarrow{\scriptstyle \pi^*} \\ X_\infty & \xleftarrow{\ \tau^*\ } & X_\infty^* \end{array}$$

Using Theorem A, we can show that if a minimal t.d.s. $(X,T)$ is an open extension of its maximal distal factor, then for each $d\geq 2$, the $d$-step topological characteristic factor of $X$ is $X_{d-1}=X/\mathbf{RP}^{[d-1]}$ (Theorem 4.3). This fact emphasizes the analogy with the ergodic situation. We point out that the number $d-1$ is the sharp result, since $T^{n_i}x\to x,\dots,T^{(d-1)n_i}x\to x$ and $T^{dn_i}x\to y$ for some $y$ implies $(x,y)\in\mathbf{RP}^{[d-1]}$ (see Lemma 2.9).

Moreover, Theorem A answers Question 1 in the best possible way. This is so because in structure theorems of a general minimal t.d.s. $(X,T)$ there may appear proximal extensions (in one sense or another), Theorem A is the best result we can expect, meaning that the almost one-to-one modifications are actually needed. One may find an example of a minimal t.d.s. $(X,T)$ which is an almost one-to-one extension of an equicontinuous t.d.s. $(Z,T)$, and yet $(Z,T)$ is not a characteristic factor for $T\times T^2$ [23,42].

Assume that $(X,T)$ is minimal and $x\in X$. The orbit closure of $(x,\dots,x)$ under the action $\langle\sigma_d,\tau_d\rangle$ is denoted by $N_d(X,T,x)$, where

$$\tau_d(T)=T\times T^2\times\cdots\times T^d\quad\text{and}\quad\sigma_d(T)=T^{(d)}=T\times\cdots\times T.$$

It is easy to see that $N_d(X,T,x)$ is independent of $x$, so $N_d(X,T,x)$ will be denoted by $N_d(X,T)$ or $N_d(T)$ or $N_d(X)$. A basic result proved by Glasner [23] is that $N_d(X)$ is minimal under the $\langle\sigma_d,\tau_d\rangle$ action. We note that the minimality of $N_d(X)$ implies van der Waerden's theorem; see [24, Theorem 1.56].

We further investigate the dynamical properties of $N_d(X)$, and one consequence of this study, Theorem C, will be used in proving Theorems D and E.

**Theorem B.** *Let $(X,T)$ be a minimal t.d.s. and $d\geq 1$. Then the maximal equicontinuous factor of $(N_d(X,T),\langle\sigma_d,\tau_d\rangle)$ is $(N_d(X_1,T),\langle\sigma_d,\tau_d\rangle)$, where as above $X_1$ is the maximal equicontinuous factor of $(X,T)$.*

In fact, we will show more: see Theorems 5.7 and 5.8. Namely, it is proved that for each $d,k\in\mathbb{N}$, the maximal $k$-step pro-nilfactor of $N_d(X)$ is the same as the one of $N_d(X_\infty)$, and that there is a dense $G_\delta$ set $\Omega\subseteq X$ such that for each $x\in\Omega$, the maximal $k$-step pro-nilfactor of $\overline{\mathcal{O}}(x^{(d)},\tau_d)$ is the same as the one of $\overline{\mathcal{O}}((\pi_\infty x)^{(d)},\tau_d)$, where $\pi_\infty: X\to X_\infty$ is the canonical factor map. Applying Theorem B we obtain

**Theorem C.** *Let $(X,T)$ be a minimal t.d.s. and $k\geq 2$. Then $(X,T^k)$ is minimal if and only if $N_d(X,T)=N_d(X,T^k)$ for each $d\geq 1$.*

As applications of these results we get an affirmative answer to Question 2 and state it in its equivalent form:

**Theorem D.** *Let $(X,T^k)$ be minimal for some $k\geq 2$. Then for any $d\geq 1$ and any $0\leq j<k$ there is a sequence $\{n_i\}\subseteq\mathbb{Z}$ with $n_i\equiv j \pmod k$ such that $T^{n_i}x\to x, T^{2n_i}x\to x$, $\dots,T^{dn_i}x\to x$ for $x$ in a dense $G_\delta$ subset of $X$.*

It is shown in [18] that if $r\in\mathbb{N}$ and $\mathbb{N}=N_1\cup\cdots\cup N_r$, then there is $i$ such that $N_i$ contains a piecewise syndetic set. Then the orbit closure of the characteristic function

$1_{N_i} \in \{0,1\}^{\mathbb{N}}$ contains a point $\omega$ which is not $(0,0,\ldots)$, and such that each word appearing in $\omega$ appears syndetically. If in addition there is some $k \geq 2$ such that each word appearing in $\omega$ also appears in position $nk+1$ for some $n \in \mathbb{N}$, then we say that the partition $\mathbb{N} = N_1 \cup \cdots \cup N_r$ is *irreducible of type* $k$.

Using this terminology Theorem D can be restated as follows:

If $\mathbb{N} = N_1 \cup \cdots \cup N_r$ is an irreducible partition of type $k$, then there is an $i$ such that for each $l \in \mathbb{N}$ and $0 \leq j < k$ there are $a, b \in \mathbb{N}$ with $a, a+b, \ldots, a+lb \in N_i$ and $b \equiv j \pmod k$.

The following is an affirmative answer to Question 3 for polynomials of degree 2.

**Theorem E.** *Let* $(X,T)$ *be a totally minimal t.d.s., and* $P(n) = an^2 + bn + c$ *be an integral polynomial with* $a \neq 0$. *Then there is a dense* $G_\delta$ *subset* $\Omega$ *of* $X$ *such that for every* $x \in \Omega$, *the set* $\{T^{P(n)}(x) : n \in \mathbb{Z}\}$ *is dense in* $X$.

We note that Theorems D and E cannot be obtained by using ergodic results, since $(X, T^k)$ minimal for some $k \geq 2$ does not imply that there is a Borel invariant probability measure $\mu$ with $(X, \mathcal{X}, T^k, \mu)$ ergodic.

Finally, we confirm Conjecture 1.1 of [26]. In fact, we show more:

**Theorem F.** *Let* $(X,T)$ *be a minimal t.d.s. which is an open extension of its maximal distal factor. Then for any* $d \geq 1$, $\mathbf{AP}^{[d]} = \mathbf{RP}^{[d]}$.

We remark that Theorem F is sharp in some sense; see Conjecture 4 in the last section.

### *1.3. The main ideas of the proofs*

We start from the proof of Theorem A. In a deep sense Theorem A is similar to the ergodic case: one wants to reduce questions regarding the $\tau_d$-action from a general system (meaning ergodic m.p.s. or minimal t.d.s.) to the same questions in a pro-nilsystem.

Now unlike the ergodic situation where the structure theorem for ergodic systems involves only two kinds of extensions, namely isometric and weakly mixing extensions, in the structure theorem of the general minimal t.d.s. [11,40], proximal extensions (which in general need not be open) necessarily appear. This fact causes great difficulties when one wants to apply this structure theorem. To overcome these difficulties, we slightly modify the structure by introducing various kinds of auxiliary extensions. If all we need is openness of the maps then the price we pay is the introduction of an auxiliary almost one-to-one modification of the original extension. Fortunately, such a modification exists in a canonical way by the classical construction called the *O-diagram*.

The second difficulty we face is more essential: there are no tools like Gowers–Host–Kra seminorms or the van der Corput lemma in topological dynamics, whereas these tools are frequently used in [28]. The two main ingredients we use instead are: a simplified version of a construction used by Glasner [23], and the characterizations of the regionally proximal relation of order $d$ obtained by Huang, Shao and Ye [30], which involves Poincaré and Birkhoff sets, introduced by Furstenberg [18], and their higher order versions by Frantzikinakis, Lesigne and Wierdl [16].

Once we have these tools, the real difficulty is in checking one specific condition in the construction. Namely, we need to verify that if $O$ is a relatively open subset of $N_d(X)$, then the orbit closure of $O$ under the $\tau_d$-action is "saturated" in the sense that if it contains some point in a fiber, then it already contains the full fiber (see Lemma 4.1). In trying to do this for a while, we realized that this can be done only when all the generators of the group $\langle \tau_d, \sigma_d \rangle$ are used. For example, when $d = 3$ and $\langle \sigma_3, \tau_3 \rangle = \langle T \times T \times T, T \times T^2 \times T^3 \rangle$, in the proof we have to use the generators

$$\{\mathrm{id} \times T \times T^2, \sigma_3\}, \quad \{T \times \mathrm{id} \times T^{-1}, \sigma_3\}, \quad \{T^2 \times T \times \mathrm{id}, \sigma_3\}.$$

In previous works we never expected that the last two generators may become useful. The idea to use all the generators is crucial in the current paper, and we also believe that this phenomenon will become useful in other settings as well.

Now we turn to the proof of Theorem B. By Theorem A, it is relatively easy to see that the maximal equicontinuous factor of $N_d(X)$ is the same as the one of $N_d(X_\infty)$. So, it remains to show this for the higher order pro-nilsystems. This is done by using Glasner's result (Lemma 3.4), a recent result proved by Qiu and Zhao (Lemma 2.10), and Lemma 2.9.

Theorem C is obtained as an application of Theorem B, together with a result for equicontinuous systems (Proposition 5.10), and a discussion of the decomposition for minimal t.d.s. under the iterations of $T$.

With the preparations we have outlined so far it is not hard to get Theorems D–F, except that we need to develop a tool in order to switch results for $N_d(X)$ under the $\langle \tau_d, \sigma_d \rangle$ action to $N_d(X)$ under the $\tau_d$ action. We provide such a tool in Lemma 6.1.

To finish, we note that Theorem A opens a window for the possibility to explore some further natural questions which we will discuss in the last section of this paper.

### *1.4. The organization of the paper*

In Section 2, we present some preliminaries. In Section 3 we provide the two main tools for the proof of Theorem A. Section 4 is devoted to proving Theorem A. The proofs for Theorems B and C are expounded in Section 5. In Section 6 we give some applications of Theorems B and C; more specifically, we prove Theorems D–F there. Some open questions are discussed in the final section.

## 2. Preliminaries

In this section we give some necessary notions and some known facts which we will use later.

### *2.1. General topological dynamics*

A *topological dynamical system* (t.d.s. for short) is a triple $\mathcal{X} = (X, \Gamma, \Pi)$, where $X$ is a compact Hausdorff space, $\Gamma$ is a Hausdorff topological group and $\Pi : \Gamma \times X \to X$ is a continuous map such that $\Pi(e,x) = x$ and $\Pi(s,\Pi(t,x)) = \Pi(st,x)$, where $e$ is the unit of $\Gamma$, $s,t \in \Gamma$ and $x \in X$. We shall fix $\Gamma$ and suppress the action symbol. Thus for $x \in X$ and $t \in \Gamma$, write $tx$ for $\Pi(t,x)$.

We always assume that $X$ *is a compact metric space with metric* $\rho(\cdot,\cdot)$*, and* $\Gamma$ *is a discrete countable group*. When $\Gamma = \mathbb{Z}$, we will write the t.d.s. as $(X,T)$ with $T$ being a homeomorphism on $X$. So in this notation $\Gamma = \{T^n : n \in \mathbb{Z}\}$.

Let $(X,\Gamma)$ be a t.d.s. and $x \in X$. Then $\mathcal{O}(x,\Gamma) = \{gx : g \in \Gamma\}$ denotes the *orbit* of $x$, which is also denoted by $\Gamma x$. We usually denote the closure of $\mathcal{O}(x,\Gamma)$ by $\overline{\mathcal{O}}(x,\Gamma)$, or $\overline{\Gamma x}$. For $A \subseteq X$, the orbit of $A$ is given by $\mathcal{O}(A,\Gamma) = \{tx : x \in A, t \in \Gamma\}$, and $\overline{\mathcal{O}}(A,\Gamma) = \overline{\mathcal{O}(A,\Gamma)}$.

A subset $A \subseteq X$ is called *invariant* (or $\Gamma$-invariant) if $ga \subseteq A$ for all $a \in A$ and $g \in \Gamma$. When $Y \subseteq X$ is a closed and invariant subset of the system $(X,\Gamma)$, we say that the system $(Y,\Gamma)$ is a *subsystem* of $(X,\Gamma)$. If $(X,\Gamma)$ and $(Y,\Gamma)$ are two t.d.s., their *product system* is the system $(X \times Y,\Gamma)$, where $g(x,y) = (gx,gy)$ for any $g \in \Gamma$ and $x,y \in X$.

A t.d.s. $(X,\Gamma)$ is called *minimal* if $X$ contains no proper non-empty closed invariant subsets. It is easy to verify that a t.d.s. is minimal if and only if every orbit is dense. In a general system $(X,\Gamma)$ we say that a point $x \in X$ is minimal if $(\overline{\mathcal{O}}(x,\Gamma),\Gamma)$ is minimal.

A *factor map* $\pi : X \to Y$ between the t.d.s. $(X,\Gamma)$ and $(Y,\Gamma)$ is a continuous onto map which intertwines the actions; we say that $(Y,\Gamma)$ is a *factor* of $(X,\Gamma)$ and that $(X,\Gamma)$ is an *extension* of $(Y,\Gamma)$. The systems are said to be *isomorphic* if $\pi$ is bijective. Let $\pi : (X,\Gamma) \to (Y,\Gamma)$ be a factor map. Then

$$R_\pi = \{(x_1,x_2) : \pi(x_1) = \pi(x_2)\}$$

is a closed invariant equivalence relation, and $Y = X/R_\pi$.

Let $X,Y$ be compact metric spaces and $T : X \to Y$ be a map. For $n \geq 2$ let

$$T^{(n)} = \underbrace{T \times \cdots \times T}_{n \text{ times}} : X^n \to Y^n.$$

Thus we write $(X^n,T^{(n)})$ for the $n$-fold product system $(X \times \cdots \times X, T \times \cdots \times T)$. The diagonal of $X^n$ is

$$\Delta_n(X) = \{(x,\ldots,x) \in X^n : x \in X\}.$$

When $n = 2$ we write $\Delta(X) = \Delta_2(X)$.

### 2.2. *Proximal, distal and regionally proximal relations*

Let $(X,\Gamma)$ be a t.d.s. Fix $(x,y) \in X^2$. It is a *proximal* pair if $\liminf_{g\in\Gamma} \rho(gx,gy) = 0$; it is a *distal* pair if it is not proximal. Denote by $\mathbf{P}(X,\Gamma)$ the set of proximal pairs of $(X,\Gamma)$. It is also called the *proximal relation*. A well known theorem of Auslander–Ellis states that for a t.d.s. $(X,\Gamma)$, any $x \in X$ is proximal to some minimal point $y$ in $\overline{\mathcal{O}}(x,\Gamma)$ (see e.g. [19, Theorem 8.7]).

A t.d.s. $(X,\Gamma)$ is *equicontinuous* if for every $\varepsilon > 0$ there exists $\delta > 0$ such that $\rho(x,y) < \delta$ implies $\rho(gx,gy) < \varepsilon$ for every $g \in \Gamma$. It is *distal* if $\mathbf{P}(X,\Gamma) = \Delta(X)$. Any equicontinuous system is distal.

Let $S_{\text{distal}}$ (respectively $S_{\text{eq}}$) be the smallest closed invariant equivalence relation $S$ on $X$ for which the factor $X/S$ is a distal (respectively equicontinuous) system. The equivalence relation $S_{\text{distal}}$ (resp. $S_{\text{eq}}$) is called the *distal* (resp. *equicontinuous*) *structure relation* of $X$. It is well known that $S_{\text{distal}}$ is the smallest closed invariant equivalence relation on $X$ which includes $\mathbf{P}(X)$ and $X/S_{\text{distal}}$ is the largest distal factor of $X$ (see e.g. [9, Chapter V, (1.5)]).

In the study of t.d.s., one of the first problems was to characterize $S_{\text{eq}}$. A natural candidate for $S_{\text{eq}}$ is the so-called *regionally proximal relation* $\mathbf{RP}(X)$ introduced by Ellis and Gottschalk [12]. Let $(X, \Gamma)$ be a minimal t.d.s. The regionally proximal relation $\mathbf{RP}(X, \Gamma)$ is defined as: $(x, y) \in \mathbf{RP}(X, \Gamma)$ if for any $\varepsilon > 0$ and for any neighborhood $U \times V$ of $(x, y)$ there are $(x', y') \in U \times V$ and $g \in \Gamma$ such that $\rho(gx', gy') < \varepsilon$. It is well known that $\mathbf{RP}(X, \Gamma)$ is an invariant closed relation and this relation defines the *maximal equicontinuous factor* $X_{\text{eq}} = X/S_{\text{eq}}$ of $(X, \Gamma)$ (see e.g. [9, Chapter V, (1.6)]).

It is a difficult problem to find conditions under which $\mathbf{RP}(X)$ is an equivalence relation (i.e., $\mathbf{RP}(X) = S_{\text{eq}}$). Starting with Veech [38], various authors, including Peterson, Ellis–Keynes [13], McMahon [35] and Bronstein [7], came up with various sufficient conditions for $\mathbf{RP}(X)$ to be an equivalence relation. What we will use is the following result (see e.g. [9, Chapter V, (1.17), (2.12)]).

**Theorem 2.1.** *Let $(X, \Gamma)$ be a minimal t.d.s., where $\Gamma$ is an amenable group. Then we have the following statements:*

(1) $\mathbf{RP}(X)$ *is an invariant closed equivalence relation which induces the maximal equicontinuous factor $X_{\text{eq}}$.*
(2) *If $(Y, \Gamma)$ is a factor of $(X, \Gamma)$ and we let $\pi : X \to Y$ be a factor map, then*

$$\pi \times \pi(\mathbf{RP}(X, \Gamma)) = \mathbf{RP}(Y, \Gamma).$$

An extension $\pi : X \to Y$ is said to be *proximal* if $R_\pi \subseteq \mathbf{P}(X)$. The following lemma is well known (see e.g. [9, Chapter V, (1.17), (2.9)]).

**Lemma 2.2.** *Let $\pi_i : (X_i, \Gamma) \to (Y_i, \Gamma)$ be proximal extensions for $1 \le i \le n$. Then*

$$\prod_{i=1}^{n} \pi_i = \pi_1 \times \cdots \times \pi_n : \Big(\prod_{i=1}^{n} X_i, \Gamma\Big) \to \Big(\prod_{i=1}^{n} Y_i, \Gamma\Big)$$

*is proximal.*

### 2.3. $\mathbf{RP}^{[d]}$ *and* $\mathbf{AP}^{[d]}$

If $\mathbf{n} = (n_1, \ldots, n_d) \in \mathbb{Z}^d$ and $\boldsymbol{\varepsilon} \in \{0, 1\}^d$, we define

$$\mathbf{n} \cdot \boldsymbol{\varepsilon} = \sum_{i=1}^{d} n_i \varepsilon_i.$$

For a t.d.s. $(X, T)$, Host, Kra and Maass [29] introduced the following definition.

**Definition 2.3.** Let $(X,T)$ be a t.d.s. and let $d \in \mathbb{N}$. The points $x, y \in X$ are said to be *regionally proximal of order* $d$ if for any $\delta > 0$, there exist $x', y' \in X$ and a vector $\mathbf{n} = (n_1, \dots, n_d) \in \mathbb{Z}^d$ such that $\rho(x,x') < \delta$, $\rho(y,y') < \delta$, and

$$\rho(T^{\mathbf{n}\cdot\boldsymbol{\varepsilon}}x', T^{\mathbf{n}\cdot\boldsymbol{\varepsilon}}y') < \delta \quad \text{for any } \boldsymbol{\varepsilon} \in \{0,1\}^d \setminus \{\mathbf{0}\},$$

where $\mathbf{0} = (0,\dots,0) \in \{0,1\}^d$. The set of regionally proximal pairs of order $d$ is denoted by $\mathbf{RP}^{[d]}$ (or by $\mathbf{RP}^{[d]}(X,T)$ in case of ambiguity), and is called the *regionally proximal relation of order* $d$.

Similarly we can define $\mathbf{RP}^{[d]}(X,\Gamma)$ for a system $(X,\Gamma)$ with abelian group $\Gamma$. We note that $\mathbf{RP}^{[1]} = \mathbf{RP}$. The regionally proximal relation of order $d$ was introduced by Host, Kra and Maass [29]. It is easy to see that $\mathbf{RP}^{[d]}$ is a closed and invariant relation. Observe that

$$\mathbf{P}(X) \subseteq \cdots \subseteq \mathbf{RP}^{[d+1]} \subseteq \mathbf{RP}^{[d]} \subseteq \cdots \subseteq \mathbf{RP}^{[2]} \subseteq \mathbf{RP}^{[1]} = \mathbf{RP}.$$

**Definition 2.4.** Let $G$ be a group. For $A, B \subseteq G$, we write $[A,B]$ for the subgroup spanned by $\{[a,b] = aba^{-1}b^{-1} : a \in A, b \in B\}$. The commutator subgroups $G_j$, $j \geq 1$, are defined inductively by setting $G_1 = G$ and $G_{j+1} = [G_j, G]$. Let $d \geq 1$ be an integer. We say that $G$ is *$d$-step nilpotent* if $G_{d+1}$ is the trivial subgroup.

Let $G$ be a $d$-step nilpotent Lie group and $\Lambda$ be a discrete cocompact subgroup of $G$. The compact manifold $X = G/\Lambda$ is called a *$d$-step nilmanifold*. The group $G$ acts on $X$ by left translations and we write this action as $(g,x) \mapsto gx$. The Haar measure $\mu$ of $X$ is the unique probability measure on $X$ invariant under this action. Let $\tau \in G$ and $T$ be the transformation $x \mapsto \tau x$ of $X$. Then $(X,\mu,T)$ is called a *$d$-step nilsystem*. An inverse limit of $d$-step nilsystems is called a $d$-step *pro-nilsystem* or a *system of order $d$*.

Host, Kra and Maass [29] showed that if a t.d.s. $(X,T)$ is minimal and distal then $\mathbf{RP}^{[d]}$ is an equivalence relation, and a very deep result stating that $(X/\mathbf{RP}^{[d]}, T)$ is the maximal $d$-step pro-nilfactor of the system. Shao and Ye [37] showed that all these results in fact hold for arbitrarily minimal t.d.s. of abelian group actions. See Glasner, Gutman and Ye [25] for similar results regarding general group actions.

The following theorems proved in [29] (for minimal distal systems) and in [37] (for general minimal t.d.s.) give us conditions under which the pair $(x,y)$ belongs to $\mathbf{RP}^{[d]}$ and the relation between $\mathbf{RP}^{[d]}$ and $d$-step pro-nilsystems. We state them for $\mathbb{Z}$-actions and they hold for minimal t.d.s. under abelian group actions.

**Theorem 2.5.** *Let $(X,T)$ be a minimal t.d.s. and let $d \geq 1$ be an integer. Then*

(1) $\mathbf{RP}^{[d]}$ *is an equivalence relation.*

(2) $(X,T)$ *is a $d$-step pro-nilsystem if and only if* $\mathbf{RP}^{[d]} = \Delta(X)$.

**Theorem 2.6.** *Let $\pi : (X,T) \to (Y,S)$ be a factor map between minimal t.d.s. and let $d \geq 1$ be an integer. Then:*

(1) $\pi \times \pi(\mathbf{RP}^{[d]}(X,T)) = \mathbf{RP}^{[d]}(Y,S)$.

(2) *$(Y,T)$ is a $d$-step pro-nilsystem if and only if* $\mathbf{RP}^{[d]}(X,T) \subseteq R_\pi$.

*In particular, the quotient of $(X,T)$ under* $\mathbf{RP}^{[d]}(X,T)$ *is the maximal $d$-step pro-nilfactor of $X$ (i.e., the maximal factor which is a $d$-step pro-nilsystem).*

Let $X_d = X/\mathbf{RP}^{[d]}(X,T)$ and $\pi_d : (X,T) \to (X_d,T)$ be the factor map. The system $X_0$ is the trivial system and the system $X_1$ is the maximal equicontinuous factor $X_{\text{eq}}$.

The following lemma is an easy consequence of the definition.

**Lemma 2.7.** *Let $(X,T)$ be a t.d.s. and $d \in \mathbb{N}$. Then*

$$\mathbf{RP}^{[d]}(X,T) = \mathbf{RP}^{[d]}(X,T^n), \quad \forall n \in \mathbb{N}.$$

*Proof.* Let $d,n \in \mathbb{N}$. It is clear that $\mathbf{RP}^{[d]}(X,T^n) \subseteq \mathbf{RP}^{[d]}(X,T)$. Now we show the opposite inclusion.

Let $(x,y) \in \mathbf{RP}^{[d]}(X,T)$ and $\delta > 0$. There is $\delta' > 0$ such that if $\rho(x_1,x_2) < \delta'$ then $\rho(T^j x_1, T^j x_2) < \delta$ for all $j \in \{1,\dots,dn\}$. Now since $(x,y) \in \mathbf{RP}^{[d]}(X,T)$, there exist $x', y' \in X$ and a vector $\mathbf{n}' = (n'_1,\dots,n'_d) \in \mathbb{Z}^d$ such that $\rho(x,x'), \rho(y,y') < \delta'$, and

$$\rho(T^{\mathbf{n}'\cdot\boldsymbol{\varepsilon}}x', T^{\mathbf{n}'\cdot\boldsymbol{\varepsilon}}y') < \delta' \quad \text{for any } \boldsymbol{\varepsilon} \in \{0,1\}^d \setminus \{\mathbf{0}\}.$$

Thus by the choice of $\delta'$, one has

$$\rho(T^{\mathbf{n}'\cdot\boldsymbol{\varepsilon}+j}x', T^{\mathbf{n}'\cdot\boldsymbol{\varepsilon}+j}y') < \delta \quad \text{for any } \boldsymbol{\varepsilon} \in \{0,1\}^d \setminus \{\mathbf{0}\} \text{ and } 1 \le j \le dn.$$

Let $n_k = n'_k + j_k \equiv 0 \pmod n$, where $0 \le j_k \le n-1$ for $k = 1,\dots,d$. Note that

$$\mathbf{n}'\cdot\boldsymbol{\varepsilon} \le \mathbf{n}\cdot\boldsymbol{\varepsilon} \le \mathbf{n}'\cdot\boldsymbol{\varepsilon} + dn.$$

It follows that

$$\rho(T^{\mathbf{n}\cdot\boldsymbol{\varepsilon}}x', T^{\mathbf{n}\cdot\boldsymbol{\varepsilon}}y') < \delta \quad \text{for any } \boldsymbol{\varepsilon} \in \{0,1\}^d \setminus \{\mathbf{0}\}.$$

Since $n_k \equiv 0 \pmod n$ for all $k = 1,\dots,d$, we have $(x,y) \in \mathbf{RP}^{[d]}(X,T^n)$. ■

Now we give the definition of $\mathbf{AP}^{[d]}$.

**Definition 2.8.** Let $(X,T)$ be a t.d.s. and $d \in \mathbb{N}$. We say that $(x,y) \in X \times X$ is a *regionally proximal pair of order $d$ along arithmetic progressions* if for each $\delta > 0$ there exist $x', y' \in X$ and $n \in \mathbb{Z}$ such that $\rho(x,x'), \rho(y,y') < \delta$ and

$$\rho(T^{in}x', T^{in}y') < \delta \quad \text{for each } 1 \le i \le d.$$

The set of all such pairs is denoted by $\mathbf{AP}^{[d]}(X,T)$ or $\mathbf{AP}^{[d]}(X)$ and called the *regionally proximal relation of order $d$ along arithmetic progressions*.

It follows easily that $\mathbf{AP}^{[d]}(X,T) \subseteq \mathbf{RP}^{[d]}(X,T)$ for each $d \in \mathbb{N}$. The following simple observation will be used in the sequel. Let $(X,T)$ be a t.d.s., $x \in X$, and $d \in \mathbb{N}$. Set $x^{(d)} = (x,\dots,x) \in X^d$.

**Lemma 2.9.** *Let $(X, T)$ be a minimal t.d.s. Then for each $d \geq 3$, $(x^{(d-1)}, y) \in N_d(X)$ for some $x, y \in X$ implies that*

$$(x, y) \in \mathbf{AP}^{[d-2]}(X, T) \subseteq \mathbf{RP}^{[d-2]}(X, T).$$

*Moreover, for $d \geq 3$ and $x \in X$, the conditions $T^{n_i} x \to x, \ldots, T^{(d-1)n_i} x \to x, T^{dn_i} x \to y$ for some $y$ imply that $(x, y) \in \mathbf{AP}^{[d-1]}(X, T) \subseteq \mathbf{RP}^{[d-1]}(X, T)$.*

### 2.4. $\infty$-step nilsystems

For any minimal t.d.s. $(X, T)$, $\mathbf{RP}^{[\infty]} = \bigcap_{d=1}^{\infty} \mathbf{RP}^{[d]}$ is a closed invariant equivalence relation (we write $\mathbf{RP}^{[\infty]}(X, T)$ in case of ambiguity). The following notion first appeared in [10].

A minimal t.d.s. $(X, T)$ is an *$\infty$-step pro-nilsystem*, or a *system of order $\infty$*, if the equivalence relation $\mathbf{RP}^{[\infty]}$ is trivial, i.e., coincides with the diagonal. Similarly, one can show that the quotient of a minimal t.d.s. $(X, T)$ under $\mathbf{RP}^{[\infty]}$ is the maximal $\infty$-step pro-nilfactor of $(X, T)$.

Let $(X, T)$ be a minimal t.d.s. It is easy to see that if $(X, T)$ is an inverse limit of minimal nilsystems, then $(X, T)$ is an $\infty$-step pro-nilsystem. Conversely, if $(X, T)$ is a minimal $\infty$-step pro-nilsystem, i.e., $\mathbf{RP}^{[\infty]} = \Delta(X)$, then $(X, T) = \varprojlim (X_d, T)_{d \in \mathbb{N}}$ since $\Delta(X) = \mathbf{RP}^{[\infty]} = \bigcap_{d \geq 1} \mathbf{RP}^{[d]}$. In fact, a minimal t.d.s. is an $\infty$-step pro-nilsystem if and only if it is an inverse limit of minimal nilsystems [10].

Since minimal pro-nilsystems are uniquely ergodic, it is easy to see that minimal $\infty$-step pro-nilsystems are also uniquely ergodic.

### 2.5. Furstenberg's tower for minimal distal systems

Let $\pi : (X, \Gamma) \to (Y, \Gamma)$ be a factor map. We define the notion of *regionally proximal relation relative to* $\pi$ (denoted by $\mathbf{RP}_\pi(X)$ or $\mathbf{RP}_\pi$) as follows: $(x, y) \in \mathbf{RP}_\pi$ if for any neighborhood $U \times V$ of $(x, y)$ and any $\varepsilon > 0$ there are $(x', y') \in U \times V$ with $\pi(x') = \pi(y')$ and $g \in \Gamma$ such that $\rho(gx', gy') < \varepsilon$. Thus for $(Y, \Gamma)$ the trivial one-point system, we retrieve the regionally proximal relation. We say that $\pi$ is an *equicontinuous* or *isometric* extension if for any $\varepsilon > 0$ there exists $\delta > 0$ such that $\pi(x_1) = \pi(x_2)$ and $\rho(x_1, x_2) < \delta$ imply $\rho(gx_1, gx_2) < \varepsilon$ for any $g \in \Gamma$. A factor map $\pi$ is equicontinuous if and only if $\mathbf{RP}_\pi(X) = \Delta(X)$.

Furstenberg's structure theorem for minimal distal systems [17] says that any minimal distal system can be constructed by equicontinuous extensions. Furstenberg showed that if $\pi : X \to Y$ is a factor map with $(X, \Gamma)$ minimal and distal, then $\mathbf{RP}_\pi$ is a closed invariant equivalence relation and $X/\mathbf{RP}_\pi$ the largest equicontinuous extension of $Y$ within $X$. This gives a structure theorem for minimal distal systems: For every minimal distal system $(X, \Gamma)$ there is an ordinal $\eta$ (which is countable when $X$ is metrizable) and a family $\{(F_n, \Gamma)\}_{n \leq \eta}$ of systems such that

(i) $F_0$ is a one-point trivial system,

(ii) for every $n < \eta$ there exists a homomorphism $\phi_{n+1} : F_{n+1} \to F_n$ which is equicontinuous,

(iii) for a limit ordinal $\nu \le \eta$ the system $F_\nu$ is the inverse limit of the systems $\{F_\iota\}_{\iota<\nu}$,

(iv) $F_\eta = X$.

Altogether, we have

$$F_0 \xleftarrow{\phi_1} F_1 \xleftarrow{\phi_2} \cdots \xleftarrow{\phi_n} F_n \xleftarrow{\phi_{n+1}} F_{n+1} \leftarrow \cdots \leftarrow F_\eta = X. \tag{2.1}$$

(2.1) is referred to as the *Furstenberg tower*. Note that in (2.1) for each $n < \eta$, the system $(F_{n+1}, \Gamma)$ can be chosen to be the largest equicontinuous extension of $F_n$ within $X$. That is, if we let $\psi_n : X \to F_n$ be the factor map, then $F_{n+1} = F_n/\mathbf{RP}_{\psi_n}$ for $n < \eta$, and we have

$$\begin{array}{ccc} & X & \\ {\scriptstyle \psi_{n+1}} \swarrow & & \searrow {\scriptstyle \psi_n} \\ F_{n+1} & \xrightarrow{\phi_{n+1}} & F_n \end{array}$$

We call $F_n$ the *largest distal factor of order* $n$.

An interesting result proved by Qiu and Zhao [36, Section 6] is that $F_n = X_n = X/\mathbf{RP}^{[n]}$ for pro-nilsystems.

**Lemma 2.10.** *Let $k, d \in \mathbb{N}$ with $k \le d$ and $(X, T)$ be a minimal $d$-step pro-nilsystem. Then $X_k = X/\mathbf{RP}^{[k]}$ coincides with $F_k$ for $1 \le k \le d$.*

By Lemma 2.10, we have

**Lemma 2.11.** *Let $(X, T)$ be a minimal t.d.s. and $n \ge 1$. Then $X_{n+1}$ is an equicontinuous extension of $X_n$.*

We remark that a simple argument shows the following result.

**Proposition 2.12.** *Let $(X, T)$ be a minimal t.d.s. and $n \in \mathbb{N}$. Then $X_n = X/\mathbf{RP}^{[n]}$ is a factor of $F_n$, the largest distal factor of order $n$.*

*Proof.* We use induction on $n \in \mathbb{N}$. When $n = 1$, we have $\mathbf{RP}^{[1]}(X) = \mathbf{RP}(X)$ and $X_1 = X/\mathbf{RP}^{[1]}(X) = X/\mathbf{RP}(X) = F_1$. Now we assume that $X_n$ is a factor of $F_n$ for $n \in \mathbb{N}$. Let $\psi_n : X \to F_n$ and $\pi_n : X \to X_n$ be the corresponding factor maps:

$$\begin{array}{ccc} & X & \\ {\scriptstyle \psi_n} \swarrow & & \searrow {\scriptstyle \pi_n} \\ F_n & \longrightarrow & X_n \end{array}$$

This induces the factor map

$$\phi : F_{n+1} = X/\mathbf{RP}_{\psi_n}(X) \to X'_{n+1} = X/\mathbf{RP}_{\pi_n}(X).$$

Note $X'_{n+1}$ is the largest equicontinuous extension of $X_n$ within $X$. As $X_{n+1}$ is an equicontinuous extension of $X_n$ within $X$, it follows that $X_{n+1}$ is a factor of $X'_{n+1}$. Thus $X_{n+1}$ is a factor of $F_{n+1}$. The proof is complete. ∎

## 2.6. *Some properties of $N_d(X, T)$*

Let $(X, T)$ be a t.d.s., and let $x \in X$, $A \subseteq X$ and $d \in \mathbb{N}$. Set $x^{(d)} = (x, \dots, x) \in X^d$ and

$$\Delta_d(A) = A^{(d)} = \{x^{(d)} = (x, \dots, x) : x \in A\} \subseteq X^d,$$
$$\sigma_d = \sigma(T) = T^{(d)} = T \times \cdots \times T \ (d \text{ times}),$$
$$\tau_d = \tau_d(T) = T \times T^2 \times \cdots \times T^d.$$

Note that $\Delta_d(X)$ is the diagonal of $X^d$. Let $\mathcal{G}_d = \langle \sigma_d, \tau_d \rangle$, the group generated by $\sigma_d$ and $\tau_d$. Let $\tau'_d = \tau'_d(T) = \mathrm{id} \times T \times \cdots \times T^{d-1} = \mathrm{id} \times \tau_{d-1}$. Note that $\mathcal{G}_d = \langle \sigma_d, \tau_d \rangle = \langle \sigma_d, \tau'_d \rangle$, which will be frequently used in this paper.

Let $X, Y$ be sets, and let $\pi : X \to Y$ be a map. A subset $L$ of $X$ is called *$\pi$-saturated* if

$$\{x \in L : \pi^{-1}(\pi(x)) \subseteq L\} = L,$$

i.e., $L = \pi^{-1}(\pi(L))$.

**Definition 2.13** ([23]). Let $\pi : (X, T) \to (Y, T)$ be a factor map of t.d.s. and $d \in \mathbb{N}$. Then $(Y, T)$ is said to be a *$d$-step topological characteristic factor (for $\tau_d$)* or *topological characteristic factor of order $d$* if there exists a dense $G_\delta$ subset $X_0$ of $X$ such that for each $x \in X_0$ the orbit closure

$$L_x = \overline{\mathcal{O}}(x^{(d)}, \tau_d)$$

is $\pi^{(d)} = \pi \times \cdots \times \pi$ ($d$ times) saturated. That is, $(x_1, \dots, x_d) \in L_x$ if and only if $(x'_1, \dots, x'_d) \in L_x$ whenever $\pi(x_i) = \pi(x'_i)$ for all $i \in \{1, \dots, d\}$.

**Theorem 2.14** ([23]). *If $(X, T)$ is a distal minimal t.d.s. and $d \geq 2$, then $F_{d-1}$ its largest distal factor of order $d - 1$, is a topological characteristic factor of order $d$.*

The following result follows from Theorem 2.14 and Lemma 2.10.

**Theorem 2.15.** *Let $(X, T)$ be a $d$-step nilsystem for some $d \in \mathbb{N}$. Then for each $1 \leq i \leq d - 1$, $X_i$ is an $(i + 1)$-step topological characteristic factor of $X$.*

Let $(X, T)$ be a t.d.s. and $d \in N$. Let

$$N_d(X, T) = N_d(X) = \overline{\mathcal{O}}(\Delta_d(X), \tau_d).$$

If $(X, T)$ is transitive and $x \in X$ is a transitive point, then $N_d(X) = \overline{\mathcal{O}}(x^{(d)}, \langle \sigma_d, \tau_d \rangle)$.

We want to emphasize that $N_d(X, T)$ also plays a key role in the study of the point-wise convergence of (1.2) for ergodic distal m.p.s. In particular, it is shown in [31] that each ergodic m.p.s. admits a uniquely ergodic minimal model $(X, T)$ such that $Z_d = X_d$

for $d \in \mathbb{N}$ and $(N_d(X), \langle \sigma_d, \tau_d \rangle)$ is minimal and uniquely ergodic, where $Z_d$ is the measurable pro-nilfactor defined in [28].

The next theorem is fundamental for the analysis carried throughout our work (for a short enveloping semigroup proof see [24, Proposition 1.55]):

**Theorem 2.16** (Glasner [23]). *Let $(X, T)$ be a minimal t.d.s. and $d \in \mathbb{N}$. Then the system $(N_d(X), \langle \sigma_d, \tau_d \rangle)$ is minimal and the $\tau_d$-minimal points are dense in $N_d(X)$.*

By the same proof of Theorem 2.16, we have

**Theorem 2.17.** *Let $(X, T)$ be a minimal t.d.s. and $a_1, \ldots, a_d$ be distinct integer numbers, where $d \in \mathbb{N}$. Let*

$$\tau = T^{a_1} \times \cdots \times T^{a_d}.$$

*Then the system $(\overline{\mathcal{O}}(\Delta_d(X), \tau), \langle \sigma_d, \tau \rangle)$ is minimal and the $\tau$-minimal points are dense in $\overline{\mathcal{O}}(\Delta_d(X), \tau)$.*

The following two lemmas follow from [23].

**Lemma 2.18.** *Let $(X, T)$ be a minimal t.d.s. and $d \in \mathbb{N}$. Let $U \subseteq X$ be a non-empty open subset and let $U^{(d)} = \Delta_d(U) = \{x^{(d)} : x \in U\}$. Then*

$$\mathrm{int}_{N_d(X)} \overline{\mathcal{O}}(U^{(d)}, \tau_d) \neq \emptyset.$$

*Proof.* Since $(X, T)$ is minimal, there is some $K \in \mathbb{N}$ such that $X = \bigcup_{k=1}^{K} T^{-k} U$. It follows that $\Delta_d(X) = \bigcup_{k=1}^{K} (T^{(d)})^{-k} U^{(d)}$. Thus

$$N_d(X) = \overline{\mathcal{O}}(\Delta_d(X), \tau_d) = \overline{\mathcal{O}}\Big(\bigcup_{k=1}^{K} (T^{(d)})^{-k} U^{(d)}, \tau_d\Big) = \bigcup_{k=1}^{K} (T^{(d)})^{-k} \overline{\mathcal{O}}(U^{(d)}, \tau_d).$$

Therefore $\mathrm{int}_{N_d(X)} \overline{\mathcal{O}}(U^{(d)}, \tau_d) \neq \emptyset$. ∎

### 2.7. *Open extensions*

Let $(X, T)$ and $(Y, S)$ be t.d.s. and let $\pi : X \to Y$ be a factor map. One says that

(1) $\pi$ is an *open extension* if it is open as a map;

(2) $\pi$ is an *almost one-to-one extension* if there exists a dense $G_\delta$ set $\Omega \subseteq X$ such that $\pi^{-1}(\{\pi(x)\}) = \{x\}$ for any $x \in \Omega$.

We will often use the following construction which is originally due to Veech (see [39, Theorem 3.1])

**Theorem 2.19.** *Given a factor map $\pi : X \to Y$ between minimal t.d.s. $(X, T)$ and $(Y, S)$, there exists a commutative diagram of factor maps* (*called $O$*-diagram)

$$\begin{array}{ccc} X & \xleftarrow{\ \sigma^*\ } & X^* \\ \pi \big\downarrow & & \big\downarrow \pi^* \\ Y & \xleftarrow{\ \tau^*\ } & Y^* \end{array}$$

*such that*

(a) $\sigma^*$ *and* $\tau^*$ *are almost one-to-one extensions;*

(b) $\pi^*$ *is an open extension;*

(c) $X^*$ *is the unique minimal set in* $R_{\pi\tau^*} = \{(x, y) \in X \times Y^* : \pi(x) = \tau^*(y)\}$, *and* $\sigma^*$ *and* $\pi^*$ *are the restrictions to* $X^*$ *of the projections of* $X \times Y^*$ *onto* $X$ *and* $Y^*$ *respectively.*

We note that this diagram is canonical. In particular, when the map $\pi$ is open, we have $X^* = X$.

### 2.8. *Some subsets of* $\mathbb{Z}$

A subset $S$ of $\mathbb{Z}$ is *syndetic* if it has a bounded gap, i.e., there is $N \in \mathbb{N}$ such that $\{i, i+1, \dots, i+N\} \cap S \neq \emptyset$ for every $i \in \mathbb{Z}$. A subset $S \subseteq \mathbb{Z}$ is *thick* if it contains arbitrarily long runs of integers, i.e., for every $n \in \mathbb{N}$ there exists some $a_n \in \mathbb{Z}$ such that $\{a_n, a_n + 1, \dots, a_n + n\} \subseteq S$.

A subset $S$ of $\mathbb{Z}$ is *piecewise syndetic* if it is the intersection of a syndetic set with a thick set; and it is *thickly syndetic* if for each $n \in \mathbb{N}$ there is a syndetic subset $\{w_1^n, w_2^n, \dots\}$ of $S$ such that $\{w_i^n, w_i^n + 1, \dots, w_i^n + n\} \subseteq S$ for each $i \in \mathbb{N}$. Denote by $\mathcal{F}_{\mathrm{ts}}$ the family of all thickly syndetic sets. It is clear that if $F_1, F_2 \in \mathcal{F}_{\mathrm{ts}}$ then so is $F_1 \cap F_2$. That is, $\mathcal{F}_{\mathrm{ts}}$ is a filter.

## 3. Some tools used in proving Theorem A

In this section we will introduce some tools that will be used in proving Theorem A. We start with the saturation theorem.

### 3.1. *A saturation theorem*

In [23] Glasner proved an auxiliary theorem in order to prove saturation with respect to the Furstenberg tower. By simplifying the assumptions of this theorem, we find that it applies to a more general setup. We will discuss this theorem, and other lemmas related to saturation properties.

**Lemma 3.1** ([22, Lemma 2.1]). *Let* $\phi : X \to Y$ *be an open map of compact metric spaces. Let* $\mathcal{V} = \{V \subseteq X : V$ *open and* $\phi(V) = Y\}$. *Then there exists a countable subset* $\{V_i\}_{i=1}^{\infty}$ *of* $\mathcal{V}$ *such that every element of* $\mathcal{V}$ *contains some* $V_i$.

*Proof.* First we show that for each $V \in \mathcal{V}$, there exists a closed subset $L_V \subseteq V$ with $\phi(L_V) = Y$. Let $V \in \mathcal{V}$. If $V^c = \emptyset$, then set $L_V = V$. If $V^c \neq \emptyset$, then for $\varepsilon > 0$ denote

$$V_\varepsilon = \{x \in V : d(x, V^c) \geq \varepsilon\}.$$

If for every $n \in \mathbb{N}$, $\phi(V_{1/n}) \neq Y$, then there exists $y_n \in Y \setminus \phi(V_{1/n})$. Let $\lim_{n\to\infty} y_n = y$ without loss of generality. By assumption there exists $x \in V$ with $\phi(x) = y$. Let $\delta = d(x, V^c)$. By our assumption, $\delta > 0$. Since $\phi$ is open we can find $x_n \in X$ with $\phi(x_n) = y_n$ such that $\lim_{n\to\infty} x_n = x$. But then $x_n$ is eventually in $V_{\delta/2}$ and $y_n \in \phi(V_{\delta/2})$, a contradiction. Thus we have proved that for every $V \in \mathcal{V}$ there exists a closed subset $L_V \subseteq V$ with $\phi(L_V) = Y$.

Let $\mathcal{U} = \{U_i\}_{i=1}^{\infty}$ be a countable basis for open sets on $X$. Then for each $V \in \mathcal{V}$, one can find a finite subset $\{U_{i_1}, \dots, U_{i_k}\}$ that covers $L_V$ and satisfies $\bigcup_{j=1}^{k} U_{i_j} \subseteq V$, which implies that $\phi(\bigcup_{j=1}^{k} U_{i_j}) = Y$. Thus

$$\mathcal{V}_0 = \Big\{ V = \bigcup_{j=1}^{k} U_{i_j} : U_{i_j} \in \mathcal{U} \text{ and } \phi(V) = Y \Big\}$$

is the required countable collection of subsets. ■

Based on results in [23, Section 3], we derive the following useful theorem. For completeness, we include a proof.

**Theorem 3.2** (Saturation theorem). *Let $I$ be some index set and for each $\zeta \in I$ let $\sigma_\zeta : (X_\zeta, \Gamma) \to (Z_\zeta, \Gamma)$ be an extension of t.d.s. (not necessarily minimal t.d.s.), where $\Gamma$ is a discrete countable group. Let*

$$(X, \Gamma) = \Big( \prod_{\zeta \in I} X_\zeta, \Gamma \Big), \quad (Z, \Gamma) = \Big( \prod_{\zeta \in I} Z_\zeta, \Gamma \Big),$$

*and let $\sigma : X \to Z$ be the product homomorphism. Let $N_Z$ be a non-empty closed $\Gamma$-invariant subset of $Z$, and let $N_X = \sigma^{-1}(N_Z)$.*

*Let $Q$ be a closed subset of $N_X$. Suppose that*

(1) *$\sigma$ is open, i.e., $\sigma_\zeta$ is open for each $\zeta \in I$;*

(2) *for every non-empty relatively open set $U \subseteq N_X$,*

$$\overline{\mathcal{O}}(U, \Gamma) = \sigma^{-1}(\sigma(\overline{\mathcal{O}}(U, \Gamma)));$$

(3) *$\overline{\mathcal{O}}(Q, \Gamma) = N_X$;*

(4) *for every non-empty relatively open subset $U$ of $Q$, $\mathrm{int}_{N_X} \overline{\mathcal{O}}(U, \Gamma) \neq \emptyset$.*

*Then there exists a dense $G_\delta$ subset $\Omega$ of $Q$ such that $x \in \Omega$ implies that $\overline{\mathcal{O}}(x, \Gamma)$ is $\sigma$-saturated.*

*Proof.* Let $S$ be a non-empty closed subset of $N_Z$ for which $\mathrm{cl}(\mathrm{int}_{N_Z} S) = S$, and let

$$Q_S = \{x \in Q : \mathcal{O}(x, \Gamma) \cap \sigma^{-1}(\mathrm{int}_{N_Z} S) \neq \emptyset\}.$$

Then by (3), $Q_S$ is a non-empty open subset of $Q$.

**Claim.** *There exists a dense $G_\delta$ subset $\Omega_S$ of $Q_S$ such that for each $x \in \Omega_S$, there exists $z \in S$ with*

$$\sigma^{-1}(z) \subseteq L_x = \overline{\mathcal{O}}(x, \Gamma).$$

*Proof of the Claim.* Let

$$\mathcal{V} = \{V \subseteq N_X : V \cap \sigma^{-1}(S) \text{ is relatively open in } \sigma^{-1}(S) \text{ and } \sigma(V) \supseteq S\}.$$

Since $\sigma(U \cap \sigma^{-1}(S)) = \sigma(U) \cap S$ for each open set $U$ of $N_X$, by openness of $\sigma$ it follows that $\sigma|_{\sigma^{-1}(S)} : \sigma^{-1}(S) \to S$ is open. Therefore by Lemma 3.1 there exists a countable subcollection $\{V_k\}_{k=1}^\infty$ of $\mathcal{V}$ such that each $V \in \mathcal{V}$ contains an element of $\{V_k\}_{k=1}^\infty$.

For each $k \in \mathbb{N}$, $V_k \cap \sigma^{-1}(S)$ is relatively open in $\sigma^{-1}(S)$ and $\sigma(V_k) \supseteq S$, so by the fact that $\operatorname{int}_{N_Z} S \neq \emptyset$, we conclude that $\operatorname{int}_{N_X} V_k \neq \emptyset$. By condition (2) and $\operatorname{cl}(\operatorname{int}_{N_Z} S) = S$ we get

$$\overline{\mathcal{O}}(\operatorname{int}_{N_X} V_k, \Gamma) = \sigma^{-1}(\sigma(\overline{\mathcal{O}}(\operatorname{int}_{N_X} V_k, \Gamma))) \supseteq \overline{\mathcal{O}}(\sigma^{-1}(S), \Gamma). \tag{3.1}$$

For each $k \in \mathbb{N}$, let

$$\Lambda_k = \{x \in Q_S : \mathcal{O}(x, \Gamma) \cap V_k \neq \emptyset\}.$$

Let $U$ be a non-empty open subset of $Q_S$. Then $U$ is a relatively open subset of $Q$ (since $Q_S$ is an open subset of $Q$), and by (4), $\operatorname{int}_{N_X} \overline{\mathcal{O}}(U, \Gamma) \neq \emptyset$. From this fact we find that

$$\overline{\mathcal{O}}(U, \Gamma) \cap \mathcal{O}(V_k, \Gamma) \neq \emptyset.$$

For each $x \in U \subseteq Q_S$, we have

$$\mathcal{O}(x, \Gamma) \cap \sigma^{-1}(\operatorname{int}_{N_Z} S) \neq \emptyset.$$

Thus $x \in \mathcal{O}(\sigma^{-1}(\operatorname{int}_{N_Z} S), \Gamma)$, i.e.,

$$U \subseteq \mathcal{O}(\sigma^{-1}(\operatorname{int}_{N_Z} S), \Gamma).$$

By (3.1),

$$\overline{\mathcal{O}}(U, \Gamma) \subseteq \overline{\mathcal{O}}(\sigma^{-1}(\operatorname{int}_{N_Z} S), \Gamma) \subseteq \overline{\mathcal{O}}(\operatorname{int}_{N_X} V_k, \Gamma).$$

Since $\operatorname{int}_{N_X} \overline{\mathcal{O}}(U, \Gamma) \neq \emptyset$, we have

$$\overline{\mathcal{O}}(U, \Gamma) \cap \mathcal{O}(V_k, \Gamma) \neq \emptyset.$$

Hence $\mathcal{O}(U, \Gamma) \cap V_k \neq \emptyset$, and it follows that $\Lambda_k$ is an open dense subset of $Q_S$. The set

$$\Omega_S = \bigcap_{k=1}^{\infty} \Lambda_k$$

is therefore a dense $G_\delta$ subset of $Q_S$. In particular, for $S = N_Z$, $\Omega_{N_Z}$ is a dense $G_\delta$ subset of $Q$.

Now for $x \in \Omega_S$ we show that there exists $z \in S$ with $\sigma^{-1}(z) \subseteq L_x = \overline{\mathcal{O}}(x, \Gamma)$. Let $V = N_X \setminus L_x$. Then $V$ is an open $\Gamma$-invariant subset of $N_X$, and if $\sigma(V) \supseteq S$, then $V \in \mathcal{V}$ and $V_k \subseteq V$ for some $k$. Since then, however, $x \in \overline{\mathcal{O}}(V_k, \Gamma) \subseteq V$, we get a contradiction. Thus there exists $z \in S$ with $\sigma^{-1}(z) \cap V = \emptyset$, i.e., $\sigma^{-1}(z) \subseteq L_x$.

The proof of the Claim is complete. ■

Let $\{B_0 = N_Z, B_1, B_2, \ldots\}$ be a basis for the topology of $N_Z$. Define $S_j = \overline{B_j}$ and, inductively, we define dense $G_\delta$ subsets $\Omega_j$ of $Q$ as follows. Let $\Omega_0 = \Omega_{S_0} = \Omega_{N_Z}$. We put

$$\Omega_{j+1} = (\Omega_j \cap \Omega_{S_{j+1}}) \cup \big(\Omega_j \cap (\overline{Q_{S_{j+1}}})^c\big).$$

Notice that this is a disjoint union. Finally, let $\Omega = \bigcap_{j=0}^{\infty} \Omega_j$.

Let $x_0 \in \Omega$ and denote $L = \overline{\mathcal{O}}(x_0, \Gamma)$. Put

$$L_\sigma = \{x \in L : \sigma^{-1}(\sigma(x)) \subseteq L\}.$$

Since $\sigma$ is open, $L_\sigma$ is closed and clearly $L_\sigma$ is $\sigma$-saturated. If $L_\sigma = L$, we are done. Otherwise, let $z_0 \in \sigma(L) \setminus \sigma(L_\sigma)$. Since $N_Z \setminus \sigma(L_\sigma)$ is open and $\{B_i\}_{i=0}^{\infty}$ is a basis of the topology of$N_Z$, there is some $j$ such that $z_0 \in B_j \subseteq S_j \subseteq N_Z \setminus \sigma(L_\sigma)$.

Now

$$x_0 \in \Omega \subseteq \Omega_j = (\Omega_{j-1} \cap \Omega_{S_j}) \cup \big(\Omega_{j-1} \cap (\overline{Q_{S_j}})^c\big).$$

Since $z_0 \in \sigma(L) \cap B_j$, we have

$$\mathcal{O}(x_0, \Gamma) \cap \sigma^{-1}(B_j) \neq \emptyset,$$

and hence $x_0 \in Q_{S_j}$. Therefore $x_0 \in \Omega_{j-1} \cap \Omega_{S_j} \subseteq \Omega_{S_j}$. By the Claim there exists $z \in S_j$ with $\sigma^{-1}(z) \subseteq L$, whence $\sigma^{-1}(z) \subseteq L_\sigma$; this contradicts $S_j \subseteq N_Z \setminus \sigma(L_\sigma)$. To sum up, for all $x_0 \in \Omega$ the set $\overline{\mathcal{O}}(x_0, \Gamma)$ is $\sigma$-saturated. The proof is complete. ■

The following lemma was implicitly proved in [23]; we give a proof for completeness.

**Lemma 3.3.** *Let $\pi : (X, T) \to (Y, T)$ be an open extension of minimal t.d.s. and $d \in \mathbb{N}$. If $Y$ is a $d$-step topological characteristic factor of $X$, then $N_{d+1}(X)$ is $\pi^{(d+1)}$-saturated, i.e.,*

$$(\pi^{(d+1)})^{-1}(N_{d+1}(Y)) = N_{d+1}(X).$$

*Proof.* By the hypothesis, there is a dense $G_\delta$ subset $\Omega_d$ of $X$ such that for any $x \in \Omega_d$, $\overline{\mathcal{O}}(x^{(d)}, \tau_d)$ is $\pi^{(d)}$-saturated. It is clear that

$$N_{d+1}(X) \subseteq (\pi^{(d+1)})^{-1}(N_{d+1}(Y)).$$

Now we prove the opposite inclusion. Let $y \in Y$ and $x \in \pi^{-1}(y)$. Since $\Omega_d$ is dense, choose $\{x_i\}_{i \in \mathbb{N}} \subseteq \Omega_d$ such that $x_i \to x$ as $i \to \infty$. Let $y_i = \pi(x_i)$; then $y_i \to y$ as $i \to \infty$. Since $\{x_i\}_{i \in \mathbb{N}} \subseteq \Omega_d$, for each $i \in \mathbb{N}$ we have

$$\overline{\mathcal{O}}((x_i)^{(d)}, \tau_d) = (\pi^{(d)})^{-1}(\overline{\mathcal{O}}({y_i}^{(d)}, \tau_d)).$$

It follows that (recall here $\tau'_{d+1} = \mathrm{id} \times \tau_d$)

$$\begin{aligned}\{x_i\} \times (\pi^{(d)})^{-1}(\overline{\mathcal{O}}({y_i}^{(d)}, \tau_d)) &= \{x_i\} \times \overline{\mathcal{O}}(x_i^{(d)}, \tau_d) \\ &= \overline{\mathcal{O}}(x_i^{(d+1)}, \mathrm{id} \times \tau_d) = \overline{\mathcal{O}}(x_i^{(d+1)}, \tau'_{d+1}) \\ &\subseteq \overline{\mathcal{O}}(\Delta_{d+1}(X), \tau'_{d+1}) = \overline{\mathcal{O}}(\Delta_{d+1}(X), \tau_{d+1}).\end{aligned}$$

In particular,

$$\{x_i\} \times (\pi^{(d)})^{-1}({y_i}^{(d)}) \subseteq \overline{\mathcal{O}}(\Delta_{d+1}(X), \tau_{d+1}).$$

Note that $\pi^{-1}$ is continuous as $\pi$ is open, and we have

$$\{x\} \times (\pi^{(d)})^{-1}(y^{(d)}) = \lim_{i\to\infty} \{x_i\} \times (\pi^{(d)})^{-1}({y_i}^{(d)}) \subseteq \overline{\mathcal{O}}(\Delta_{d+1}(X), \tau_{d+1}).$$

That is,

$$(\pi^{(d+1)})^{-1}(y^{(d+1)}) = \pi^{-1}(y) \times (\pi^{(d)})^{-1}(y^{(d)}) \subseteq \overline{\mathcal{O}}(\Delta_{d+1}(X), \tau_{d+1}).$$

Since $y \in Y$ is arbitrary, we have

$$(\pi^{(d+1)})^{-1}(\Delta_{d+1}(Y)) \subseteq \overline{\mathcal{O}}(\Delta_{d+1}(X), \tau_{d+1}).$$

By the continuity of $\pi^{-1}$, we have

$$\begin{aligned}(\pi^{(d+1)})^{-1}(N_{d+1}(Y)) &= (\pi^{(d+1)})^{-1}(\overline{\mathcal{O}}(\Delta_{d+1}(Y), \tau_{d+1})) \\ &= \overline{\mathcal{O}}\big((\pi^{(d+1)})^{-1}(\Delta_{d+1}(Y)), \tau_{d+1}\big) \subseteq \overline{\mathcal{O}}(\Delta_{d+1}(X), \tau_{d+1}) = N_{d+1}(X).\end{aligned}$$

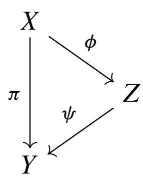
The proof is complete. ■

**Lemma 3.4.** *Let $d \in \mathbb{N}$, $(X, T)$ be a distal minimal t.d.s. and $\psi_d : X \to F_d$ be the factor map to its largest distal factor of order $d$. Then there is a dense $G_\delta$ subset $\Omega$ of $X$ such that if $x \in \Omega$, then $\overline{\mathcal{O}}(x^{(d+1)}, \tau_{d+1})$ is $\psi_d^{(d+1)}$-saturated and $(\psi_d^{(d+2)})^{-1} N_{d+2}(F_d) = N_{d+2}(X)$.*

*Proof.* The first part is from Theorem 2.14, and the second follows from Lemma 3.3. ■

The following lemma is easy to verify, by the definitions.

**Lemma 3.5.** *Let $X, Y, Z$ be compact metric spaces. Let $\pi : X \to Y$, $\phi : X \to Z$ and $\psi : Z \to Y$ be continuous surjective maps such that $\pi = \psi \circ \phi$:*




(1) *If $A \subseteq X$ is $\pi$-saturated, then $A$ is $\phi$-saturated.*

(2) *If $A \subseteq X$ is $\phi$-saturated and $\phi(A)$ is $\psi$-saturated, then $A$ is $\pi$-saturated.*

### 3.2. *The connection of* $\mathbf{RP}^{[d]}$ *with recurrence sets*

We now turn to the second tool we use in the proof of Theorem A. We need the notions of Poincaré and Birkhoff recurrence sets of higher order [16, 18]. To define them we appeal to the MERT and TMRT theorems stated in the introduction.

Let $(X, T)$ be a t.d.s., and let $x \in X$ and $U \subseteq X$. Set

$$N_T(x, U) = \{n \in \mathbb{Z} : T^n x \in U\}.$$

**Definition 3.6.** Let $d \in \mathbb{N}$.

(1) We say that $S \subseteq \mathbb{Z}$ is a set of *d-recurrence* if for every measure preserving system $(X, \mathcal{X}, \mu, T)$ and for every $A \in \mathcal{X}$ with $\mu(A) > 0$, there exists $n \in S$ such that

$$\mu(A \cap T^{-n} A \cap \cdots \cap T^{-dn} A) > 0.$$

(2) We say that $S \subseteq \mathbb{Z}$ is a set of *d-topological recurrence* if for every minimal t.d.s. $(X, T)$ and for every non-empty open subset $U$ of $X$, there exists $n \in S$ such that

$$U \cap T^{-n} U \cap \cdots \cap T^{-dn} U \neq \emptyset.$$

(3) We say that $S \subseteq \mathbb{Z}$ is a *Nil$_d$ Bohr$_0$-set* if there are a $d$-step nilsystem $(X, T)$, $x \in X$ and a neighborhood $U$ of $x$ such that $S \supset N_T(x, U)$.

Let $\mathcal{F}_{\mathrm{Poi}_d}$ (resp. $\mathcal{F}_{\mathrm{Bir}_d}$, $\mathcal{F}_d$) be the family of sets of $d$-recurrence (resp. sets of $d$-topological recurrence, Nil$_d$ Bohr$_0$-sets). It is obvious from the definitions above that $\mathcal{F}_{\mathrm{Poi}_d} \subseteq \mathcal{F}_{\mathrm{Bir}_d}$.

The following result plays an important role in the proof of Theorem A.

**Theorem 3.7** ([30, Theorem 7.2.7]). *Let $(X, T)$ be a minimal t.d.s., and let $d \in \mathbb{N}$ and $x, y \in X$. Then the following statements are equivalent:*

(1) $(x, y) \in \mathbf{RP}^{[d]}(X, T)$.

(2) *$N_T(x, U) \in \mathcal{F}_{\mathrm{Poi}_d}$ for each neighborhood $U$ of $y$.*

(3) *$N_T(x, U) \in \mathcal{F}_{\mathrm{Bir}_d}$ for each neighborhood $U$ of $y$.*

(4) *$N_T(x, U) \in \mathcal{F}_d^*$, i.e., $N_T(x, U) \cap A \neq \emptyset$ for each Nil$_d$ Bohr$_0$-set $A$ and each neighborhood $U$ of $y$.*

**Remark 3.8.** Theorem 3.7 still holds if $(X, T)$ is a t.d.s. consisting of finitely many minimal subsystems. In fact, if one of the statements in Theorem 3.7 holds, then it is easy to see that $x, y$ are in the same minimal subsystem of $(X, T)$, and using Theorem 3.7 in the minimal subsystem, we see that the rest of the statements hold. In particular, if $(X, T)$ is minimal and $n \in \mathbb{N}$, Theorem 3.7 holds for $(X, T^n)$.

## 4. Proof of Theorem A

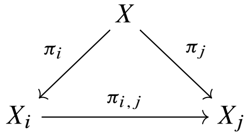
With the preparation in Section 3, we are in a position to show Theorem A.

### *4.1. Proof of Theorem A assuming a key lemma*

In this subsection we will give a very useful lemma, which allows us to reduce problems related to general minimal t.d.s. to their nilfactors. Recall that $X_d = X/\mathbf{RP}^{[d]}$ for $d \in \mathbb{N}$ and $\pi_d : (X, T) \to (X_d, T)$ is the corresponding factor map. For $j < i$, let

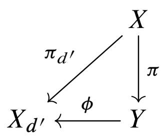
$$\pi_{i,j} : (X_i, T) \to (X_j, T).$$

Then we have




The following lemma plays a key role in the proof of Theorem A; the proof of the lemma will be given in later subsections since it is very long. We remark that perhaps the number $d'$ picked in the lemma is not sharp, but it is convenient for us to get all the information we need.

**Lemma 4.1.** *Let $\pi : (X, T) \to (Y, T)$ be an extension of minimal t.d.s., and let $d \in \mathbb{N}$.*

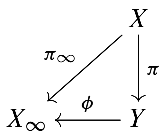



*If $\pi$ is open and $X_{d'}$ be a factor of $Y$ with $d' \geq 2d!(d-1)!$, then for every non-empty relatively open subset $O$ of $N_d(X)$, one has*

$$\overline{\mathcal{O}}(O, \tau_d) = (\pi^{(d)})^{-1}(\pi^{(d)}(\overline{\mathcal{O}}(O, \tau_d))).$$

Now we prove Theorem A assuming Lemma 4.1. In fact, Theorem A follows from the following theorem and the $O$-diagram construction.

**Theorem 4.2.** *Let $\pi : (X, T) \to (Y, T)$ be an extension of minimal t.d.s. If $\pi$ is open and $X_\infty$ is a factor of $Y$, then $Y$ is a $d$-step topological characteristic factor of $X$ for all $d \in \mathbb{N}$.*




*That is, for all $d \in \mathbb{N}$ there exists a dense $G_\delta$ subset $\Omega_d$ of $X$ such that for each $x \in \Omega_d$ the orbit closure $L_x = \overline{\mathcal{O}}(x^{(d)}, \tau_d)$ is $\pi^{(d)}$-saturated.*

*Proof of Theorem 4.2 assuming Lemma 4.1.* We use induction on $d$. The case $d = 1$ is trivial.

*Case* $d = 2$. In this case $N_2(X) = X \times X$, $N_2(Y) = Y \times Y$. Let $\pi : (X, T) \to (Y, T)$. Then we also have $\pi : (X, T^2) \to (Y, T^2)$. Let

$$\pi^{(2)} : (X \times X, T \times T^2) \to (Y \times Y, T \times T^2).$$

We need to verify the following:

$(1)_2$ $\pi^{(2)}$ is open;

$(2)_2$ for every non-empty relatively open subset $O$ of $N_2(X)$,

$$\overline{\mathcal{O}}(O, T \times T^2) = (\pi^{(2)})^{-1}(\pi^{(2)}(\overline{\mathcal{O}}(O, T \times T^2)));$$

$(3)_2$ $\overline{\mathcal{O}}(\Delta(X), T \times T^2) = N_2(X) = (\pi^{(2)})^{-1}(N_2(Y))$;

$(4)_2$ for every non-empty open subset $U$ of $X$, $\mathrm{int}_{N_2(X)} \overline{\mathcal{O}}(U^{(2)}, T \times T^2) \neq \emptyset$.

$(1)_2$ is from our assumption; $(3)_2$ is clear; by Lemma 2.18, we have $(4)_2$; and $(2)_2$ follows from Lemma 4.1.

Then by the saturation theorem (Theorem 3.2) there exists a dense $G_\delta$ subset $\Omega_2$ of $X$ such that $x \in \Omega_2$ implies that $\overline{\mathcal{O}}(x^{(2)}, T \times T^2)$ is $\pi^{(2)}$-saturated. That is, $Y$ is a 2-step topological characteristic factor of $X$.

*Case* $d + 1$. We assume that the result holds for $d \geq 2$, and we show it for $d + 1$. We will verify the conditions of the saturation theorem. That is, we will verify the following conditions:

$(1)_{d+1}$ $\pi^{(d+1)}$ is open;

$(2)_{d+1}$ for every non-empty relatively open subset $O$ of $N_{d+1}(X)$,

$$\overline{\mathcal{O}}(O, \tau_{d+1}) = (\pi^{(d+1)})^{-1}(\pi^{(d+1)}(\overline{\mathcal{O}}(O, \tau_{d+1})));$$

$(3)_{d+1}$ $\overline{\mathcal{O}}(\Delta_{d+1}(X), \tau_{d+1}) = N_{d+1}(X) = (\pi^{(d+1)})^{-1}(N_{d+1}(Y))$;

$(4)_{d+1}$ for every non-empty open subset $U$ of $X$, $\mathrm{int}_{N_{d+1}(X)} \overline{\mathcal{O}}(U^{(d+1)}, \tau_{d+1}) \neq \emptyset$.

Condition $(1)_{d+1}$ follows from our assumption that $\pi$ is open. $(2)_{d+1}$ follows from Lemma 4.1. By inductive hypothesis on $d$, $Y$ is a $d$-step topological characteristic factor of $X$; then by Lemma 3.3, $N_{d+1}(X)$ is $\pi^{(d+1)}$-saturated, i.e.,

$$(\pi^{(d+1)})^{-1}(N_{d+1}(Y)) = N_{d+1}(X).$$

Hence we have $(3)_{d+1}$. By Lemma 2.18, we have $(4)_{d+1}$.

So by the saturation theorem (Theorem 3.2) there exists a dense $G_\delta$ subset $\Omega_{d+1}$ of $X$ such that $x \in \Omega_{d+1}$ implies that $\overline{\mathcal{O}}(x^{(d+1)}, \tau_{d+1})$ is $\pi^{(d+1)}$-saturated. That is, $Y$ is a $d + 1$-step topological characteristic factor of $X$. The proof is complete. ■

We are now ready to show how Theorem A follows from Theorem 4.2.

*Proof of Theorem A.* Let $(X, T)$ be a minimal t.d.s., and $\pi : X \to X_\infty$ be the factor map. Then by the $O$-diagram (Theorem 2.19) there are minimal t.d.s. $X^*$ and $X_\infty^*$ which are almost one-to-one extensions of $X$ and $X_\infty$ respectively such that $\pi^*$ is open:

$$\begin{array}{ccc} X & \xleftarrow{\sigma^*} & X^* \\ \downarrow{\scriptstyle \pi} & & \downarrow{\scriptstyle \pi^*} \\ X_\infty & \xleftarrow{\tau^*} & X_\infty^* \end{array}$$

By Theorem 4.2, $X_\infty^*$ is a $d$-step topological characteristic factor of $X^*$ for all $d \geq 2$. The proof is complete. ■

As a corollary of Theorem 4.2, we have

**Theorem 4.3.** *Let $(X, T)$ be a minimal t.d.s. which is an open extension of its maximal distal factor, and $d \in \mathbb{N}$. Then $X_d$ is a $(d+1)$-step topological characteristic factor of $X$.*

*Proof.* We want to show that $X_d$ is a $d + 1$-step topological characteristic factor of $X$. Since $X_\infty$ is a factor of the maximal distal factor, by our assumption $\pi_\infty : X \to X_\infty$ is open. By the proof of Theorem 4.2, there is some $\widetilde{d}$ such that $X_{\widetilde{d}}$ is a $d + 1$-step topological characteristic factor of $X$:

$$\begin{array}{ccc} X & \xrightarrow{\pi_{\widetilde{d}}} & X_{\widetilde{d}} \\ & {\scriptstyle \pi_d} \searrow & \downarrow{\scriptstyle \pi_{\widetilde{d},d}} \\ & & X_d \end{array}$$

By Theorem 2.15, $X_d$ is a $d + 1$-step topological characteristic factor of $X_{\widetilde{d}}$. By Lemma 3.5, $X_d$ is a $d + 1$-step topological characteristic factor of $X$. ■

We now proceed to the proof of Lemma 4.1.

### 4.2. *Cases $d = 1$, $d = 2$ for Lemma 4.1*

In this subsection, to make the idea clear, we show the cases $d = 1$ and $d = 2$ of Lemma 4.1, and we show the general case after that.

Let $\rho$ be the metric of $X$ and $\rho_d$ the metric of $X^d$ defined by

$$\rho_d(\mathbf{x}, \mathbf{y}) = \max_{1 \leq j \leq d} \rho(x_j, y_j),$$

where $\mathbf{x} = (x_1, \ldots, x_d), \mathbf{y} = (y_1, \ldots, y_d) \in X^d$. For $\mathbf{x} \in X^d$ and $\delta > 0$, let

$$B_\delta(\mathbf{x}) = \{\mathbf{y} \in X^d : \rho_d(\mathbf{x}, \mathbf{y}) < \delta\}.$$

*Case $d = 1$.* In this case , we take $d' = 1$:

$$\begin{array}{ccc} & & X \\ & \swarrow{\scriptstyle \pi_1} & \downarrow{\scriptstyle \pi} \\ X_1 & \xleftarrow{\phi} & Y \end{array}$$

It is easy to see that $N_2(X) = X \times X$ and $N_2(Y) = Y \times Y$. Let $\pi : (X, T) \to (Y, T)$. Note that $R_\pi \subseteq \mathbf{RP}^{[1]}(X)$. Let

$$\pi^{(2)} : (X \times X, T \times T^2) \to (Y \times Y, T \times T^2).$$

For any non-empty open set $O \subseteq N_2(X)$, we need to show that

$$N := \overline{\mathcal{O}}(O, T \times T^2) = (\pi^{(2)})^{-1}(\pi^{(2)}(\overline{\mathcal{O}}(O, T \times T^2))).$$

To prove this, we first show the following claim:

**Claim.** *Let* $(x_1, x_2) \in O$. *Then* $\pi^{-1}(\pi(x_1)) \times \{x_2\} \subseteq N$ *and* $\{x_1\} \times \pi^{-1}(\pi(x_2)) \subseteq N$.

Let $(x_1, x_2) \in O$. Let $\varepsilon > 0$ be such that $B_\varepsilon(x_1) \times B_\varepsilon(x_2) \subseteq O$. Let $z_1 \in X$ be such that $(x_1, z_1) \in R_\pi \subseteq \mathbf{RP}^{[1]}(X)$.

Since $(x_1, z_1) \in R_\pi \subseteq \mathbf{RP}^{[1]}(X)$, we know by Theorem 3.7 that $N_T(x_1, B_\varepsilon(z_1)) \in \mathcal{F}_{\mathrm{Bir}_1}$. By the definition of $\mathcal{F}_{\mathrm{Bir}_1}$, there is $n \in \mathbb{Z}$ such that

$$T^n(x_1) \in B_\varepsilon(z_1) \quad \text{and} \quad A = B_\varepsilon(x_2) \cap T^{-2n} B_\varepsilon(x_2) \neq \emptyset.$$

Pick $x_2' \in A$, and we have $T^n x_1 \in B_\varepsilon(z_1)$ and $T^{2n} x_2' \in B_\varepsilon(x_2)$. It is clear that

$$(x_1, x_2') \in B_\varepsilon(x_1) \times B_\varepsilon(x_2) \subseteq O.$$

Thus

$$\rho_2((z_1, x_2), \overline{\mathcal{O}}(O, T \times T^2)) \leq \rho_2((z_1, x_2), \overline{\mathcal{O}}((x_1, x_2'), T \times T^2)) < \varepsilon.$$

Since $\varepsilon$ is arbitrary, we see that $(z_1, x_2) \in N = \overline{\mathcal{O}}(O, T \times T^2)$. This implies that

$$\pi^{-1}(\pi(x_1)) \times \{x_2\} \subseteq N.$$

Now let $z_2 \in X$ with $(x_2, z_2) \in R_\pi \subseteq \mathbf{RP}^{[1]}(X)$. Since $(x_2, z_2) \in R_\pi \subseteq \mathbf{RP}^{[1]}(X, T) = \mathbf{RP}^{[1]}(X, T^2)$, by Theorem 3.7 we have

$$N_{T^2}(x_2, B_\varepsilon(z_2)) \in \mathcal{F}_{\mathrm{Bir}_1}.$$

Thus by the definition of $\mathcal{F}_{\mathrm{Bir}_1}$, there is some $n$ such that

$$T^{2n} x_2 \in B_\varepsilon(z_2) \quad \text{and} \quad B_\varepsilon(x_1) \cap T^{-n} B_\varepsilon(x_1) \neq \emptyset.$$

Let $x_1' \in B_\varepsilon(x_1) \cap T^{-n} B_\varepsilon(x_1)$. Then

$$T^n x_1' \in B_\varepsilon(x_1), \quad T^{2n} x_2 \in B_\varepsilon(z_2).$$

Thus $\rho_2((T \times T^2)^n (x_1', x_2), (x_1, z_2)) < \varepsilon$. Note that $(x_1', x_2) \subseteq B_\varepsilon(x_1) \times B_\varepsilon(x_2) \subseteq O \subseteq N$. It follows that

$$\rho_2((x_1, z_2), \overline{\mathcal{O}}(O, T \times T^2)) \leq \rho_2\big((x_1, z_2), \overline{\mathcal{O}}((x_1', x_2), T \times T^2)\big) < \varepsilon.$$

Since $\varepsilon$ is arbitrary, $(x_1, x_2') \in N = \overline{\mathcal{O}}(O, T \times T^2)$. Thus

$$\{x_1\} \times \pi^{-1}(\pi(x_2)) \subseteq N.$$

The proof of the Claim is complete. ■

Let $(z_1, z_2) \in N = \overline{\mathcal{O}}(O, \tau_2)$. We will show that

$$(z_1', z_2') \in N \quad \text{whenever } \pi(z_i) = \pi(z_i'),\ i = 1, 2.$$

First we show that $(z_1', z_2) \in N$. Since $(z_1, z_2) \in N = \overline{\mathcal{O}}(O, \tau_2)$, there are some sequences $\{(x_1^i, x_2^i)\}_{i \in \mathbb{N}} \subseteq O$ and $\{n_i\}_{i \in \mathbb{N}} \subseteq \mathbb{Z}$ such that

$$\tau_2^{n_i}(x_1^i, x_2^i) \to (z_1, z_2), \quad i \to \infty.$$

By the Claim and since $(x_1^i, x_2^i) \in O$,

$$\pi^{-1}(\pi(x_1^i)) \times \{x_2^i\} \subseteq N.$$

Since $\pi$ is open, it follows that

$$\begin{aligned} \tau_2^{n_i}(\pi^{-1}(\pi(x_1^i)) \times \{x_2^i\}) &= \pi^{-1}(\pi(T^{n_i} x_1^i)) \times \{T^{2n_i} x_2^i\} \\ &\xrightarrow{i \to \infty} \pi^{-1}(\pi(z_1)) \times \{z_2\} \subseteq \overline{\mathcal{O}}(N, \tau_2) = N. \end{aligned}$$

Thus $(z_1', z_2) \in N$. Similarly, $(z_1', z_2') \in N$.

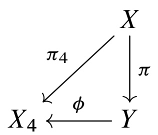
Thus we have finished the case $d = 1$.

*Case $d = 2$.* For every non-empty relatively open set $O \subseteq N_3(X)$, we want to show that

$$N = \overline{\mathcal{O}}(O, \tau_3) = (\pi^{(3)})^{-1}(\pi^{(3)}(\overline{\mathcal{O}}(O, \tau_3))),$$

where $\tau_3 = T \times T^2 \times T^3$.

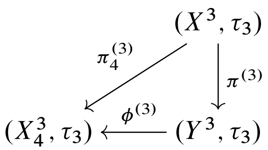
In this case, we take $d' = 4$:




Let

$$\pi^{(3)} : (X^3, \tau_3) \to (Y^3, \tau_3).$$

We have




Let $\tau_{3,1} = \mathrm{id} \times T \times T^2$, $\tau_{3,2} = T \times \mathrm{id} \times T^{-1}$ and $\tau_{3,3} = T^2 \times T \times \mathrm{id}$. It is easy to see that

$$\langle \tau_3, T^{(3)} \rangle = \langle \tau_{3,1,}, T^{(3)} \rangle = \langle \tau_{3,2}, T^{(3)} \rangle = \langle \tau_{3,3}, T^{(3)} \rangle.$$

**Step 1.** *Let* $(x_1, x_2, x_3) \in O$. *If* $y \in X$ *and* $\pi(x_1) = \pi(y)$, *then* $(y, x_2, x_3) \in N = \overline{\mathcal{O}}(O, \tau_3)$.

Since $(x_1, x_2, x_3) \in O$ and $O$ is a relatively open subset of $N_3(X)$, there is some $\delta > 0$ such that

$$B_\delta((x_1, x_2, x_3)) \cap N_3(X) \subseteq O.$$

By Theorem 2.17, $(N_3(X), \langle \mathrm{id} \times T \times T^2, T^{(3)} \rangle)$ is minimal and the $\mathrm{id} \times T \times T^2$-minimal points in $N_3(X)$ are dense in $N_3(X)$.

Let $\varepsilon > 0$ with $\varepsilon < \delta/2$. Choose an $\mathrm{id} \times T \times T^2$-minimal point $(z_1, z_2, z_3) \in N_3(X)$ such that

$$\rho_3((z_1, z_2, z_3), (x_1, x_2, x_3)) < \varepsilon < \delta/2.$$

By the openness of $\pi$, we may assume that there is some $y_1 \in X$ such that

$$\rho(y_1, y) < \varepsilon \quad \text{and} \quad \pi(z_1) = \pi(y_1).$$

Let

$$A = \overline{\mathcal{O}}((z_1, z_2, z_3), \mathrm{id} \times T \times T^2).$$

Since $(z_1, z_2, z_3)$ is $\mathrm{id} \times T \times T^2$-minimal, $(A, \mathrm{id} \times T \times T^2)$ is a minimal t.d.s. Note that by the definition of $A$, for all $(w_1, w_2, w_3) \in A$, one has $w_1 = z_1$.

Let

$$U = A \cap (B_\varepsilon(z_1) \times B_\varepsilon(z_2) \times B_\varepsilon(z_3))$$

be a non-empty open subset of $A$.

By the assumption $(z_1, y_1) \in R_\pi \subseteq \mathbf{RP}^{[4]}(X, T) = \mathbf{RP}^{[4]}(X, T^2)$, we have

$$N_{T^2}(z_1, B_\varepsilon(y_1)) \in \mathcal{F}_{\mathrm{Bir}_4}.$$

Thus, according to Theorem 3.7 there is some $n \in \mathbb{Z}$ such that $T^{2n} z_1 \in B_\varepsilon(y_1)$ and

$$\begin{aligned} B = U &\cap (\mathrm{id} \times T \times T^2)^{-n} U \cap (\mathrm{id} \times T \times T^2)^{-2n} U \\ &\cap (\mathrm{id} \times T \times T^2)^{-3n} U \cap (\mathrm{id} \times T \times T^2)^{-4n} U \neq \emptyset. \end{aligned}$$

Let $(z_1, y_2, y_3) \in B$. Then from $(z_1, y_2, y_3) \in (\mathrm{id} \times T \times T^2)^{-4n} U$, we get $T^{4n} y_2 \in B_\varepsilon(z_2)$, and from $(z_1, y_2, y_3) \in (\mathrm{id} \times T \times T^2)^{-3n} U$ we get $T^{6n} y_3 \in B_\varepsilon(z_3)$ (this explains why we need to use $\mathbf{RP}^{[4]}$ instead of $\mathbf{RP}^{[2]}$ for our method). Hence, we have

$$T^{2n} z_1 \in B_\varepsilon(y_1), \quad T^{4n} y_2 \in B_\varepsilon(z_2), \quad T^{6n} y_3 \in B_\varepsilon(z_3). \tag{4.1}$$

Thus

$$\begin{aligned} &\rho_3\big(\tau_3^{2n}(z_1, y_2, y_3), (y_1, z_2, z_3)\big) < \varepsilon, \\ &\rho_3\big((y_1, z_2, z_3), \overline{\mathcal{O}}((z_1, y_2, y_3), \tau_3)\big) < \varepsilon. \end{aligned}$$

Since $\rho_3((z_1, z_2, z_3), (x_1, x_2, x_3)) < \varepsilon$ and $\rho(y_1, y) < \varepsilon$, we have

$$\rho_3\big((y_1, z_2, z_3), (y, x_2, x_3)\big) < \varepsilon.$$

Thus

$$\rho_3\big((y, x_2, x_3), \overline{\mathcal{O}}((z_1, y_2, y_3), \tau_3)\big) < \varepsilon + \varepsilon = 2\varepsilon. \tag{4.2}$$

Since $(z_1, y_2, y_3) \in U$ and $\rho_3((z_1, y_2, y_3), (z_1, z_2, z_3)) < \varepsilon$, it follows that

$$\begin{aligned}&\rho_3((z_1, y_2, y_3), (x_1, x_2, x_3))\\ &\qquad < \rho_3((z_1, y_2, y_3), (z_1, z_2, z_3)) + \rho_d((z_1, z_2, z_3), (x_1, x_2, x_3)) < \varepsilon + \varepsilon < \delta.\end{aligned}$$

Hence

$$(z_1, y_2, y_3) \in B_\delta((x_1, x_2, x_3)) \cap N_3(X) \subseteq O.$$

So $\overline{\mathcal{O}}((z_1, y_2, y_3), \tau_3) \subseteq \overline{\mathcal{O}}(O, \tau_3)$, and we have

$$\rho_3((y, x_2, x_3), \overline{\mathcal{O}}(O, \tau_3)) < 2\varepsilon,$$

by (4.2). As $\varepsilon$ is arbitrary, we see that indeed $(y, x_2, x_3) \in \overline{\mathcal{O}}(O, \tau_3)$.

**Step 2.** *Let $(x_1, x_2, x_3) \in O$. If $y \in X$ and $\pi(x_2) = \pi(y)$, then $(x_1, y, x_3) \in N$.*

Since $(x_1, x_2, x_3) \in O$ and $O$ is a relatively open subset of $N_3(X)$, there is some $\delta > 0$ such that

$$B_\delta((x_1, x_2, x_3)) \cap N_3(X) \subseteq O.$$

By Theorem 2.17, $(N_3(X), \langle T \times \mathrm{id} \times T^{-1}, T^{(3)}\rangle)$ is minimal and the $T \times \mathrm{id} \times T^{-1}$-minimal points in $N_3(X)$ are dense in $N_3(X)$.

Let $\varepsilon > 0$ with $\varepsilon < \delta/2$. Choose an $T \times \mathrm{id} \times T^{-1}$-minimal point $(z_1, z_2, z_3) \in N_3(X)$ such that

$$\rho_3((z_1, z_2, z_3), (x_1, x_2, x_3)) < \varepsilon < \delta/2.$$

By the openness of $\pi$, we may assume that there is some $y_2 \in X$ such that

$$\rho(y_2, y) < \varepsilon \quad \text{and} \quad \pi(z_2) = \pi(y_2).$$

Let

$$A = \overline{\mathcal{O}}((z_1, z_2, z_3), T \times \mathrm{id} \times T^{-1}).$$

Since $(z_1, z_2, z_3)$ is $T \times \mathrm{id} \times T^{-1}$-minimal, $(A, T \times \mathrm{id} \times T^{-1})$ is a minimal t.d.s. Note that by the definition of $T \times \mathrm{id} \times T^{-1}$, for all $(w_1, w_2, w_3) \in A$, one has $w_2 = z_2$.

Let

$$U = A \cap (B_\varepsilon(z_1) \times B_\varepsilon(z_2) \times B_\varepsilon(z_3))$$

be a non-empty open subset of $A$.

By Theorem 3.7 and the assumption $(z_2, y_2) \in R_\pi \subseteq \mathbf{RP}^{[4]}(X, T) = \mathbf{RP}^{[4]}(X, T^2)$, we have

$$N_{T^2}(z_2, B_\varepsilon(y_2)) \in \mathcal{F}_{\mathrm{Bir}_4}.$$

Thus by the definition of $\mathcal{F}_{\mathrm{Bir}_4}$ there is some $n \in \mathbb{Z}$ such that $T^{2n} z_2 \in B_\varepsilon(y_2)$ and

$$\begin{aligned}B &= U \cap (T \times \mathrm{id} \times T^{-1})^{-n} U \cap (T \times \mathrm{id} \times T^{-1})^{-2n} U\\ &\quad \cap (T \times \mathrm{id} \times T^{-1})^{-3n} U \cap (T \times \mathrm{id} \times T^{-1})^{-4n} U \neq \emptyset.\end{aligned}$$

Let $(y_1, z_2, y_3) \in B$. Then we have

$$T^n y_1 \in B_\varepsilon(z_1), \quad T^{2n} z_2 \in B_\varepsilon(y_2), \quad T^{3n} y_3 \in B_\varepsilon(z_3). \tag{4.3}$$

Thus

$$\rho_3\big(\tau_3^n(y_1, z_2, y_3), (z_1, y_2, z_3)\big) < \varepsilon,$$
$$\rho_3\big((z_1, y_2, z_3), \overline{\mathcal{O}}((y_1, z_2, y_3), \tau_3)\big) < \varepsilon.$$

Since $\rho_3((z_1, z_2, z_3), (x_1, x_2, x_3)) < \varepsilon$ and $\rho(y_2, y) < \varepsilon$, we have

$$\rho_3\big((z_1, y_2, z_3), (x_1, y, x_3)\big) < \varepsilon.$$

Thus

$$\rho_3\big((x_1, y, x_3), \overline{\mathcal{O}}((y_1, z_2, y_3), \tau_3)\big) < \varepsilon + \varepsilon = 2\varepsilon. \tag{4.4}$$

Since $(y_1, z_2, y_3) \in U$ and $\rho_3((y_1, z_2, y_3), (z_1, z_2, z_3)) < \varepsilon$, it follows that

$$\begin{aligned}&\rho_3((y_1, z_2, y_3), (x_1, x_2, x_3))\\ &\qquad < \rho_3((y_1, z_2, y_3), (z_1, z_2, z_3)) + \rho_d((z_1, z_2, z_3), (x_1, x_2, x_3)) < \varepsilon + \varepsilon < \delta.\end{aligned}$$

Hence

$$(y_1, z_2, y_3) \in B_\delta((x_1, x_2, x_3)) \cap N_3(X) \subseteq O.$$

So $\overline{\mathcal{O}}((y_1, z_2, y_3), \tau_3) \subseteq \overline{\mathcal{O}}(O, \tau_3)$, and we have

$$\rho_3((x_1, y, x_3), \overline{\mathcal{O}}(O, \tau_3)) < 2\varepsilon,$$

by (4.4). As $\varepsilon$ is arbitrary, we have $(x_1, y, x_3) \in \overline{\mathcal{O}}(O, \tau_3)$.

**Step 3.** *Let $(x_1, x_2, x_3) \in O$. If $y \in X$ and $\pi(x_3) = \pi(y)$, then $(x_1, x_2, y) \in N$.*

Since $(x_1, x_2, x_3) \in O$ and $O$ is a relatively open subset of $N_3(X)$, there is some $\delta > 0$ such that

$$B_\delta((x_1, x_2, x_3)) \cap N_3(X) \subseteq O.$$

By Theorem 2.17, $(N_3(X), \langle T^2 \times T \times \mathrm{id}, T^{(3)}\rangle)$ is minimal and the $T^2 \times T \times \mathrm{id}$-minimal points in $N_3(X)$ are dense in $N_3(X)$.

Let $\varepsilon > 0$ with $\varepsilon < \delta/2$. Choose an $T^2 \times T \times \mathrm{id}$-minimal point $(z_1, z_2, z_3) \in N_3(X)$ such that

$$\rho_3((z_1, z_2, z_3), (x_1, x_2, x_3)) < \varepsilon < \delta/2.$$

By the openness of $\pi$, we may assume that there is some $y_3 \in X$ such that

$$\rho(y_3, y) < \varepsilon \quad \text{and} \quad \pi(z_3) = \pi(y_3).$$

Let

$$A = \overline{\mathcal{O}}((z_1, z_2, z_3), T^2 \times T \times \mathrm{id}).$$

Since $(z_1, z_2, z_3)$ is $T^2 \times T \times \mathrm{id}$-minimal, $(A, T^2 \times T \times \mathrm{id})$ is a minimal t.d.s. Note that by the definition of $T^2 \times T \times \mathrm{id}$, for all $(w_1, w_2, w_3) \in A$, one has $w_3 = z_3$.

Let

$$U = A \cap (B_\varepsilon(z_1) \times B_\varepsilon(z_2) \times B_\varepsilon(z_3))$$

be a non-empty open subset of $A$.

By Theorem 3.7 and the assumption $(z_3, y_3) \in R_\pi \subseteq \mathbf{RP}^{[4]}(X, T) = \mathbf{RP}^{[4]}(X, T^6)$, we have

$$N_{T^6}(z_3, B_\varepsilon(y_3)) \in \mathcal{F}_{\mathrm{Bir}_4}.$$

Thus by the definition of $\mathcal{F}_{\mathrm{Bir}_4}$ there is some $n \in \mathbb{Z}$ such that $T^{6n} z_3 \in B_\varepsilon(y_3)$, and

$$B = U \cap (T^2 \times T \times \mathrm{id})^{-n} U \cap (T^2 \times T \times \mathrm{id})^{-2n} U \\ \cap (T^2 \times T \times \mathrm{id})^{-3n} U \cap (T^2 \times T \times \mathrm{id})^{-4n} U \neq \emptyset.$$

Let $(y_1, y_2, z_3) \in B$. Then we have

$$T^{2n} y_1 \in B_\varepsilon(z_1), \quad T^{4n} y_2 \in B_\varepsilon(z_2), \quad T^{6n} z_3 \in B_\varepsilon(y_3). \tag{4.5}$$

Thus

$$\rho_3\big(\tau_3^{2n}(y_1, y_2, z_3), (z_1, z_2, y_3)\big) < \varepsilon, \\ \rho_3\big((z_1, z_2, y_3), \overline{\mathcal{O}}((y_1, y_2, z_3), \tau_3)\big) < \varepsilon.$$

Since $\rho_3((z_1, z_2, z_3), (x_1, x_2, x_3)) < \varepsilon$ and $\rho(y_3, y) < \varepsilon$, we have

$$\rho_3\big((z_1, z_2, y_3), (x_1, x_2, y)\big) < \varepsilon.$$

Thus

$$\rho_3\big((x_1, x_2, y), \overline{\mathcal{O}}((y_1, y_2, z_3), \tau_3)\big) < \varepsilon + \varepsilon = 2\varepsilon. \tag{4.6}$$

Since $(y_1, y_2, z_3) \in U$ and $\rho_3((y_1, y_2, z_3), (z_1, z_2, z_3)) < \varepsilon$, it follows that

$$\rho_3((y_1, y_2, z_3), (x_1, x_2, x_3)) \\ < \rho_3((y_1, y_2, z_3), (z_1, z_2, z_3)) + \rho_d((z_1, z_2, z_3), (x_1, x_2, x_3)) < \varepsilon + \varepsilon < \delta.$$

Hence

$$(y_1, y_2, z_3) \in B_\delta((x_1, x_2, x_3)) \cap N_3(X) \subseteq O.$$

So $\overline{\mathcal{O}}((y_1, y_2, z_3), \tau_3) \subseteq \overline{\mathcal{O}}(O, \tau_3)$, and we have

$$\rho_3((x_1, x_2, y), \overline{\mathcal{O}}(O, \tau_3)) < 2\varepsilon,$$

by (4.6). As $\varepsilon$ is arbitrary, we have $(x_1, x_2, y) \in \overline{\mathcal{O}}(O, \tau_3)$.

**Step 4.** $N = (\pi^{(3)})^{-1}(\pi^{(3)} N)$.

Let $(z_1, z_2, z_3) \in N = \overline{\mathcal{O}}(O, \tau_3)$. We will show that

$$(z_1', z_2', z_3') \in N \quad \text{whenever } \pi(z_i) = \pi(z_i'),\ i = 1, 2, 3.$$

First we show that $(z_1', z_2, z_3) \in N$. Since $(z_1, z_2, z_3) \in N = \overline{\mathcal{O}}(O, \tau_3)$, there are some sequences $\{(x_1^i, x_2^i, x_3^i)\}_{i \in \mathbb{N}} \subseteq O$ and $\{n_i\}_{i \in \mathbb{N}} \subseteq \mathbb{Z}$ such that

$$\tau_3^{n_i}(x_1^i, x_2^i, x_3^i) \to (z_1, z_2, z_3), \quad i \to \infty.$$

By Step 1 and $(x_1^i, x_2^i, x_3^i) \in O$,

$$\pi^{-1}(\pi(x_1^i)) \times \{x_2^i\} \times \{x_3^i\} \subseteq N.$$

Since $\pi$ is open, it follows that

$$\begin{aligned}\tau_3^{n_i}(\pi^{-1}(\pi(x_1^i)) \times \{x_2^i\} \times \{x_3^i\}) &= \pi^{-1}(\pi(T^{n_i}x_1^i)) \times \{T^{2n_i}x_2^i\} \times \{T^{3n_i}x_3^i\} \\ &\xrightarrow{i\to\infty} \pi^{-1}(\pi(z_1)) \times \{z_2\} \times \{z_3\} \subseteq \overline{\mathcal{O}}(N, \tau_3) = N.\end{aligned}$$

Thus we have

$$(z_1', z_2, z_3) \in N.$$

Similarly, using Step 2 we have

$$(z_1', z_2', z_3) \in N.$$

And then using Step 3 we have

$$(z_1', z_2', z_3') \in N.$$

Thus we have finished the proof for the case $d = 2$.

### *4.3. Proof of Lemma 4.1 in the general case*

Recall that $\rho$ is the metric of $X$ and $\rho_d$ is the metric of $X^d$ defined by

$$\rho_d(\mathbf{x}, \mathbf{y}) = \max_{1\le j\le d} \rho(x_j, y_j),$$

where $\mathbf{x} = (x_1, \dots, x_d), \mathbf{y} = (y_1, \dots, y_d) \in X^d$.

Let $O$ be a non-empty relatively open subset of $N_d(X)$. First we show the following claim.

**Claim.** *Let $\mathbf{x} = (x_1, \dots, x_d) \in O$ and $j \in \{1, \dots, d\}$. Then for each $y \in X$ with $\pi(y) = \pi(x_j)$, one has*

$$(x_1, \dots, x_{j-1}, y, x_{j+1}, \dots, x_d) \in \overline{\mathcal{O}}(O, \tau_d).$$

*Proof of the Claim.* Since $\mathbf{x} \in O$ and $O$ is a relatively open subset of $N_d(X)$, there is some $\delta > 0$ such that

$$B_\delta(\mathbf{x}) \cap N_d(X) \subseteq O.$$

For $j \in \{1, \dots, d\}$, let

$$S_j = \tau_d^{-1}(T^{(d)})^j = T^{j-1} \times T^{j-2} \times \cdots \times T \times \mathrm{id} \times T^{-1} \times \cdots \times T^{-(d-j)}. \tag{4.7}$$

Note that

$$\langle \tau_d, T^{(d)} \rangle = \langle S_j, T^{(d)} \rangle.$$

By Theorem 2.17, $(N_d(X), \langle S_j, T^{(d)} \rangle)$ is minimal and the $S_j$-minimal points in $N_d(X)$ are dense in $N_d(X)$.

Let $\varepsilon > 0$ with $\varepsilon < \delta/2$. Choose an $S_j$-minimal point $\mathbf{z} = (z_1, \dots, z_d) \in N_d(X)$ such that

$$\rho_d(\mathbf{z}, \mathbf{x}) < \varepsilon < \delta/2.$$

By the openness of $\pi$, we may assume that there is some $y_j \in X$ such that

$$\rho(y_j, y) < \varepsilon \quad \text{and} \quad \pi(z_j) = \pi(y_j).$$

Let

$$A = \overline{\mathcal{O}}(\mathbf{z}, S_j).$$

Since $\mathbf{z}$ is $S_j$-minimal, $(A, S_j)$ is a minimal t.d.s. Note that by the definition of $S_j$, for all $\mathbf{w} = (w_1, \dots, w_d) \in A$, one has $w_j = z_j$.

Let

$$U = A \cap (B_\varepsilon(z_1) \times B_\varepsilon(z_2) \times \cdots \times B_\varepsilon(z_d))$$

be a non-empty open subset of $A$.

Since $X_{d'}$ is a factor of $Y$, we have $(z_j, y_j) \in \mathbf{RP}^{[d']}(X, T)$, where $d' = 2d!(d-1)!$. By Lemma 2.7, $\mathbf{RP}^{[d']}(X, T) = \mathbf{RP}^{[d']}(X, T^{j(d-1)!})$. Consequently, by Theorem 3.7, $N_{T^{j(d-1)!}}(z_j, V) \in \mathcal{F}_{\mathrm{Bir}_{d'}}$ for each neighborhood $V$ of $y_j$. Together with the definition of $\mathcal{F}_{\mathrm{Bir}_{d'}}$, there is some $n \in \mathbb{Z}$ such that

$$T^{j(d-1)!n} z_j \in B_\varepsilon(y_j), \tag{4.8}$$

and

$$U \cap S_j^{-n} U \cap S_j^{-2n} U \cap \cdots \cap S_j^{-d'n} U \neq \emptyset. \tag{4.9}$$

Let

$$\mathbf{w}' \in U \cap S_j^{-n} U \cap S_j^{-2n} U \cap \cdots \cap S_j^{-d'n} U \neq \emptyset,$$

and

$$\mathbf{w} = (w_1, \dots, w_d) = S_j^{(d'/2)n} \mathbf{w}' \in U.$$

Then by (4.9),

$$S_j^{in} \mathbf{w} = S_j^{(d'/2+i)n} \mathbf{w}' \in U \quad \text{for all } i \text{ with } -d'/2 \le i \le d'/2.$$

That is, for all $i$ with $-d'/2 \le i \le d'/2$, we have

$$\begin{aligned} T^{(j-1)in} w_1 &\in B_\varepsilon(z_1), \\ T^{(j-2)in} w_2 &\in B_\varepsilon(z_2), \\ &\dots, \\ T^{in} w_{j-1} &\in B_\varepsilon(z_{j-1}), \\ w_j &= z_j, \\ T^{-in} w_{j+1} &\in B_\varepsilon(z_{j+1}), \\ &\dots, \\ T^{-(d-j)in} w_d &\in B_\varepsilon(z_d). \end{aligned} \tag{4.10}$$

Since $d' = 2d!(d-1)!$, by (4.10) and (4.8) we have

$$\begin{aligned} T^{(d-1)!n} w_1 &\in B_\varepsilon(z_1), \\ T^{(d-1)!2n} w_2 &\in B_\varepsilon(z_2), \\ &\dots, \\ T^{(d-1)!(j-1)n} w_{j-1} &\in B_\varepsilon(z_{j-1}), \\ T^{(d-1)!jn} w_j &\in B_\varepsilon(y_j), \\ T^{(d-1)!(j+1)n} w_{j+1} &\in B_\varepsilon(z_{j+1}), \\ &\dots, \\ T^{(d-1)!dn} w_d &\in B_\varepsilon(z_d). \end{aligned}$$

It follows that

$$\rho_d\big(\tau_d^{(d-1)!n}\mathbf{w}, (z_1, \dots, z_{j-1}, y_j, z_{j+1}, \dots, z_d)\big) < \varepsilon,$$

and hence

$$\rho_d\big((z_1, \dots, z_{j-1}, y_j, z_{j+1}, \dots, z_d), \overline{\mathcal{O}}(\mathbf{w}, \tau_d)\big) < \varepsilon.$$

Since $\rho_d(\mathbf{z}, \mathbf{x}) < \varepsilon$ and $\rho(y_j, y) < \varepsilon$, we have

$$\rho_d\big((z_1, \dots, z_{j-1}, y_j, z_{j+1}, \dots, z_d), (x_1, \dots, x_{j-1}, y, x_{j+1}, \dots, x_d)\big) < \varepsilon.$$

Thus

$$\rho_d\big((x_1, \dots, x_{j-1}, y, x_{j+1}, \dots, x_d), \overline{\mathcal{O}}(\mathbf{w}, \tau_d)\big) < \varepsilon + \varepsilon = 2\varepsilon. \tag{4.11}$$

Since $\mathbf{w} \in U$ and $\rho_d(\mathbf{w}, \mathbf{z}) < \varepsilon$, it follows that

$$\rho_d(\mathbf{w}, \mathbf{x}) < \rho_d(\mathbf{w}, \mathbf{z}) + \rho_d(\mathbf{z}, \mathbf{x}) < \varepsilon + \varepsilon < \delta.$$

Hence $\mathbf{w} \in B_\delta(\mathbf{x}) \cap N_d(X) \subseteq O$.

By (4.11), we have

$$\rho_d\big((x_1, \dots, x_{j-1}, y, x_{j+1}, \dots, x_d), \overline{\mathcal{O}}(O, \tau_d)\big) < 2\varepsilon.$$

As $\varepsilon$ is arbitrary, we have

$$(x_1, \dots, x_{j-1}, y, x_{j+1}, \dots, x_d) \in \overline{\mathcal{O}}(O, \tau_d).$$

The proof of the Claim is complete. ∎

Now we will use the Claim to show that the orbit closure $L = \overline{\mathcal{O}}(O, \tau_d)$ is $\pi^{(d)}$-saturated.

For $j \in \{1, \dots, d\}$, let

$$\mathbf{z} = (z_1, \dots, z_d) \in L = \overline{\mathcal{O}}(O, \tau_d).$$

We show that

$$\mathbf{z}' = (z_1, z_2, \dots, z_{j-1}, z_j', z_{j+1}, \dots, z_d) \in L \quad \text{whenever } \pi(z_j') = \pi(z_j).$$

Since $\mathbf{z} = (z_1, \ldots, z_d) \in L = \overline{\mathcal{O}}(O, \tau_d)$, there are sequences $\{\mathbf{x} = (x_1^i, \ldots, x_d^i)\}_{i\in\mathbb{N}} \subseteq O$, $\{n_i\}_{i\in\mathbb{N}}$ such that

$$\tau_d^{n_i}\mathbf{x}^i \to \mathbf{z}, \quad i \to \infty.$$

By the Claim we have

$$\{x_1^i\} \times \cdots \times \{x_{j-1}^i\} \times \pi^{-1}(\pi(x_j^i)) \times \{x_{j+1}^i\} \times \cdots \times \{x_d^i\} \subseteq L = \overline{\mathcal{O}}(O, \tau_d).$$

Since $\pi$ is open, it follows that

$$\begin{aligned}
\tau_d^{n_i}(\{x_1^i\} \times \cdots \times \{x_{j-1}^i\} \times \pi^{-1}(\pi(x_j^i)) \times \{x_{j+1}^i\} \times \cdots \times \{x_d^i\})&\\
&= \{T^{n_i}x_1^i\} \times \cdots \times \{T^{(j-1)n_i}x_{j-1}^i\} \times \pi^{-1}(\pi(T^{jn_i}x_j^i))\\
&\quad \times \{T^{(j+1)n_i}x_{j+1}^i\} \times \cdots \times \{T^{dn_i}x_d^i\}\\
&\xrightarrow{i\to\infty} \{z_1\} \times \cdots \times \{z_{j-1}\} \times \pi^{-1}(\pi(z_j)) \times \{z_{j+1}\} \times \cdots \times \{z_d\}\\
&\subseteq \overline{\mathcal{O}}(L, \tau_d) = L.
\end{aligned}$$

To sum up, if $\mathbf{z} = (z_1, \ldots, z_d) \in L = \overline{\mathcal{O}}(O, \tau_d)$, then for each $j \in \{1, \ldots, d\}$ we have

$$(z_1, z_2, \ldots, z_{j-1}, z_j', z_{j+1}, \ldots, z_d) \in L \quad \text{whenever } \pi(z_j') = \pi(z_j).$$

Thus, $(z_1, \ldots, z_d) \in L$ if and only if $(z_1', \ldots, z_d') \in L$ whenever $\pi(z_i) = \pi(z_i')$ for all $i \in \{1, \ldots, d\}$. That is, $L = \overline{\mathcal{O}}(O, \tau_d)$ is $\pi^{(d)}$-saturated.

The proof is complete. ■

## 5. Proofs of Theorems B and C

In this section we prove Theorems B and C.

### *5.1. Proof of Theorem B for pro-nilsystems*

For a t.d.s. $(X, T)$ and subsets $U, V$ of $X$, put $N_T(U, V) = \{n \in \mathbb{Z} : U \cap T^{-n}V \neq \emptyset\}$. First recall two facts used in the proof of Lemma 5.2.

**Lemma 5.1.** (1) *If $(X, T)$ has a dense set of minimal points, so does $(X \times X, T \times T)$.*

(2) *If $(X, T)$ is a t.d.s with dense minimal points and $(x, y) \in \mathbf{RP}(X)$, then for each neighborhood $U$ of $(x, y)$ and each neighborhood $W$ of the diagonal $\Delta(X)$, $N_{T\times T}(U, W)$ is thickly syndetic.*

*Proof.* (1) Let $U, V$ be open sets of $X$. It suffices to show that we can find a $T \times T$-minimal point in $U \times V$. Since $(X, T)$ has a dense set of minimal points, there are $T$-minimal points $x_1 \in U$ and $x_2 \in V$. Let $X_1 = \overline{\mathcal{O}}(x_1, T)$ and $X_2 = \overline{\mathcal{O}}(x_2, T)$. Then $(X_1, T)$ and $(X_2, T)$ are minimal t.d.s. and $(X_1 \times X_2) \cap (U \times V) \neq \emptyset$. Since any t.d.s. has a minimal subsystem, let $(z_1, z_2)$ be a $T \times T$-minimal point of $X_1 \times X_2$. As $(X_1, T)$ and $(X_2, T)$

are minimal, there are some $n_1, n_2 \in \mathbb{Z}$ such that $T^{n_1}z_1 \in U$ and $T^{n_2}z_2 \in V$. Hence $(T^{n_1}z_1, T^{n_2}z_2) \in U \times V$ and $(T^{n_1}z_1, T^{n_2}z_2)$ is $T \times T$-minimal. Thus $(X \times X, T \times T)$ has a dense set of minimal points.

(2) Let $(x, y) \in \mathbf{RP}(X)$ and $U$ be a neighborhood of $(x, y)$. For a neighborhood $W$ of $\Delta(X)$, choose $\varepsilon > 0$ such that $\Delta_\varepsilon(X) = \{(z, z') \in X \times X : \rho(z, z') < \varepsilon\} \subseteq W$. For any $k \in \mathbb{N}$ there is some $0 < \delta_k < \varepsilon$ such that $\rho(z, z') < \delta_k$ implies that $\rho(T^j z, T^j z') < \varepsilon$ for all $j \in \{1, \dots, k\}$. By (1), $T \times T$-minimal points are dense in $X \times X$, and thus by the definition of $\mathbf{RP}(X)$ there is some $T \times T$-minimal point $(x_k, y_k) \in U$ such that $\overline{\mathcal{O}}((x_k, y_k), T \times T) \cap \Delta_{\delta_k}(X) \neq \emptyset$. Since $(x_k, y_k)$ is minimal, the set

$$N_{T\times T}((x_k, y_k), \Delta_{\delta_k}) = \{n \in \mathbb{Z} : (T \times T)^n(x_k, y_k) \in \Delta_{\delta_k}\}$$

is syndetic. Let $n \in N_{T\times T}((x_k, y_k), \Delta_{\delta_k})$. Then $\rho(T^{n+j}x_k, T^{n+j}y_k) < \varepsilon$ for every $j \in \{1, \dots, k\}$, that is,

$$\{[n, n+k] : n \in N_{T\times T}((x_k, y_k), \Delta_{\delta_k})\} \subseteq N_{T\times T}((x_k, y_k), \Delta_\varepsilon) \subseteq N_{T\times T}(U, W),$$

where $[n, n+k] = \{n, n+1, \dots, n+k\}$. Thus $N_{T\times T}(U, W)$ is thickly syndetic. The proof is complete. ■

**Lemma 5.2.** *Let $(X, T)$ and $(Y, S)$ be two minimal t.d.s. Then the maximal equicontinuous factor of $(X \times Y, T^n \times S^m)$ is $X_{\mathrm{eq}} \times Y_{\mathrm{eq}}$ for any $n, m \in \mathbb{N}$, where $X_{\mathrm{eq}} = X_1$ and $Y_{\mathrm{eq}} = Y_1$ are the maximal equicontinuous factors of $X$ and $Y$ respectively.*

*Proof.* Let $\pi : (X, T) \to (X_{\mathrm{eq}}, T)$ and $\phi : (Y, S) \to (Y_{\mathrm{eq}}, S)$ be the factor maps to the maximal equicontinuous factors of $(X, T)$ and $(Y, S)$ respectively. Fix $n, m \in \mathbb{N}$. Since by Lemma 2.7, $\mathbf{RP}(X, T) = \mathbf{RP}(X, T^n)$ and $\mathbf{RP}(Y, S) = \mathbf{RP}(Y, S^m)$, we find that $\pi : (X, T^n) \to (X_{\mathrm{eq}}, T^n)$ and $\phi : (Y, S^m) \to (Y_{\mathrm{eq}}, S^m)$ are the factor maps to the maximal equicontinuous factors of $(X, T^n)$ and $(Y, S^m)$ respectively. Since $(X_{\mathrm{eq}} \times Y_{\mathrm{eq}}, T^n \times S^m)$ is equicontinuous, it remains to show that $R_{\pi\times\phi} \subseteq \mathbf{RP}(T^n \times S^m)$.

To this end, we assume that $(x_1, y_1), (x_2, y_2) \in X \times Y$ with $(x_1, x_2) \in \mathbf{RP}(X, T^n)$ and $(y_1, y_2) \in \mathbf{RP}(Y, S^m)$. Then $\pi(x_1) = \pi(x_2)$ and $\phi(y_1) = \phi(y_2)$. Fix $\varepsilon > 0$. Let $U_1 \times V_1$ and $U_2 \times V_2$ be neighborhoods of $(x_1, y_1)$ and $(x_2, y_2)$ respectively. Moreover, let $W_1$ and $W_2$ be the $\varepsilon$-neighborhoods of $\Delta(X)$ and $\Delta(Y)$ respectively. Note that $(X, T^n)$ and $(Y, S^m)$ each have a dense set of minimal points. Consequently, by Lemma 5.1, both $N_{T^n\times T^n}(U_1 \times U_2, W_1)$ and $N_{S^m\times S^m}(V_1 \times V_2, W_2)$ are thickly syndetic. As the family $\mathcal{F}_{\mathrm{ts}}$ of all thickly syndetic sets is a filter,

$$N_{T^n\times T^n}(U_1 \times U_2, W_1) \cap N_{S^m\times S^m}(V_1 \times V_2, W_2) \neq \emptyset.$$

Pick $k$ in this intersection. Then there are $(x_1', y_1') \in U_1 \times V_1$ and $(x_2', y_2') \in U_2 \times V_2$ such that

$$T^{kn} \times T^{kn}(x_1', x_2') \in W_1, \quad S^{km} \times S^{km}(y_1', y_2') \in W_2.$$

Denote the metrics of $X$ and $Y$ by $\rho_X$ and $\rho_Y$, and let the metric of $X \times Y$ be

$$\rho_{X\times Y}((x_1, y_1), (x_2, y_2)) = \max\{\rho_X(x_1, x_2), \rho_Y(y_1, y_2)\}.$$

Then

$$\rho_{X\times Y}\big((T^n\times S^m)^k(x_1',y_1'),(T^n\times S^m)^k(x_2',y_2')\big) \\ \le \rho_X(T^{nk}x_1',T^{nk}x_2')+\rho_Y(S^{mk}y_1',S^{mk}y_2')<2\varepsilon.$$

As $\varepsilon$ is arbitrary, we have

$$((x_1,x_2),(y_1,y_2))\in \mathbf{RP}(X\times Y,T^n\times S^m).$$

This ends the proof. ∎

**Lemma 5.3.** *Let $\pi:(X,\Gamma)\to(Y,\Gamma)$ be a factor map between two minimal t.d.s. and $k\in\mathbb{N}$, where $\Gamma$ is abelian. If $\pi^{-1}(y)^2\subseteq \mathbf{RP}^{[k]}(X)$ for some $y\in Y$, then*

$$R_\pi\subseteq \mathbf{RP}^{[k]}(X).$$

*Proof.* Let $z\in Y$ and $x_1,x_2\in\pi^{-1}(z)$. Then by the Auslander–Ellis Theorem, $(x_1,x_2)$ is proximal to some minimal point $(y_1,y_2)$ with $\pi(y_1)=\pi(y_2)$. Since $(Y,\Gamma)$ is minimal, it is easy to see that

$$\emptyset\ne\overline{\mathcal{O}}((y_1,y_2),\Gamma)\cap\pi^{-1}(y)^2\subseteq \mathbf{RP}^{[k]}(X).$$

As $\overline{\mathcal{O}}((y_1,y_2),\Gamma)$ is minimal and $\mathbf{RP}^{[k]}(X)$ is a $\Gamma$-invariant closed subset of $X\times X$, we have

$$\overline{\mathcal{O}}((y_1,y_2),\Gamma)\subseteq \mathbf{RP}^{[k]}(X).$$

In particular, $(y_1,y_2)\in\mathbf{RP}^{[k]}(X)$. Since $\mathbf{P}(X)\subseteq\mathbf{RP}^{[k]}(X)$ we have

$$(x_1,y_1),(x_2,y_2)\in\mathbf{RP}^{[k]}(X).$$

We conclude that $(x_1,x_2)\in\mathbf{RP}^{[k]}(X)$ since $\mathbf{RP}^{[k]}(X)$ is an equivalence relation by Theorem 2.5. ∎

**Corollary 5.4.** *Let $\pi:(X,T)\to(Y,T)$ be a factor map between two minimal t.d.s. and $d,k\in\mathbb{N}$. Then*

$$\pi^{(d)}:(N_d(X),\langle\sigma_d,\tau_d\rangle)\to(N_d(Y),\langle\sigma_d,\tau_d\rangle)$$

*is a factor map. If for some $x\in X$,*

$$\{x^{(d)}\}\times(\pi^{(d)})^{-1}(\pi(x)^{(d)})\subseteq\mathbf{RP}^{[k]}(N_d(X),\langle\sigma_d,\tau_d\rangle),$$

*then*

$$R_{\pi^{(d)}}\subseteq\mathbf{RP}^{[k]}(N_d(X),\langle\sigma_d,\tau_d\rangle).$$

*Proof.* Let $y=\pi(x)$ and $\mathcal{G}_d=\langle\sigma_d,\tau_d\rangle$. Note that $(N_d(X,T),\mathcal{G}_d)$ is minimal. By Lemma 5.3, it suffices to show that

$$((\pi^{(d)})^{-1}(y^{(d)}))^2\subseteq\mathbf{RP}^{[k]}(N_d(X),\langle\sigma_d,\tau_d\rangle).$$

Let $\mathbf{x}, \mathbf{x}' \in (\pi^{(d)})^{-1}(y^{(d)})$. Since $x^{(d)} \in N_d(X)$, by our assumption

$$(x^{(d)}, \mathbf{x}), (x^{(d)}, \mathbf{x}') \in \{x^{(d)}\} \times (\pi^{(d)})^{-1}(y^{(d)}) \subseteq \mathbf{RP}^{[k]}(N_d(X), \langle \sigma_d, \tau_d \rangle).$$

Since $\mathbf{RP}^{[k]}$ is an equivalence relation (Theorem 2.5),

$$(\mathbf{x}, \mathbf{x}') \in \mathbf{RP}^{[k]}(N_d(X), \langle \sigma_d, \tau_d \rangle).$$

That is,

$$((\pi^{(d)})^{-1}(y^{(d)}))^2 \subseteq \mathbf{RP}^{[k]}(N_d(X), \langle \sigma_d, \tau_d \rangle).$$

The proof is complete. ∎

We will prove the following theorem, which is a special case of Theorem B.

**Theorem 5.5.** *Let $(X, T)$ be a minimal pro-nilsystem and $d \in \mathbb{N}$. Then the maximal equicontinuous factor of $(N_d(X), \langle \sigma_d, \tau_d \rangle)$ is $(N_d(X_{\text{eq}}), \langle \sigma_d, \tau_d \rangle)$.*

*Proof.* Let $\pi_1 : X \to X_1 = X_{\text{eq}}$ be the factor map to the maximal equicontinuous factor and $d \in \mathbb{N}$. Let $\mathcal{G}_d = \langle \sigma_d, \tau_d \rangle$. Then $\pi_1$ induces a factor map

$$\pi_1^{(d)} : (N_d(X), \mathcal{G}_d) \to (N_d(X_1), \mathcal{G}_d).$$

Since $(X_1, T)$ is equicontinuous, so is $(N_d(X_1), \mathcal{G}_d)$. By Theorem 2.1 it follows that

$$\mathbf{RP}((N_d(X), \mathcal{G}_d)) \subseteq R_{\pi_1^{(d)}}.$$

To show that the maximal equicontinuous factor of $(N_d(X), \mathcal{G}_d)$ is $(N_d(X_1), \mathcal{G}_d)$, it remains to show that

$$R_{\pi_1^{(d)}} \subseteq \mathbf{RP}(N_d(X), \mathcal{G}_d). \tag{5.1}$$

For simplicity, we use

$$(x_1, \ldots, x_d) \underset{\mathcal{G}_d}{\sim} (y_1, \ldots, y_d) \quad \text{and} \quad (x_1, \ldots, x_d) \underset{\tau_d}{\sim} (y_1, \ldots, y_d)$$

to denote $((x_1, \ldots, x_d), (y_1, \ldots, y_d)) \in \mathbf{RP}(N_d(X), \mathcal{G}_d)$ and $((x_1, \ldots, x_d), (y_1, \ldots, y_d)) \in \mathbf{RP}(N_d(X), \tau_d)$ respectively.

We prove (5.1) by induction on $d$.

**Step 1.** For $d = 1$ it is clear. For $d = 2$, $\mathcal{G}_2$ is generated by $T \times T$ and $\text{id} \times T$. It is clear that $N_2(X) = X \times X$ and $N_2(X_1) = X_1 \times X_1$. By Lemma 5.2, if $\pi_1(x_1) = \pi_1(x_2)$ and $\pi_1(y_1) = \pi_1(y_2)$ then $((x_1, y_1), (x_2, y_2))$ is regionally proximal for $T \times T^2$, and thus $((x_1, y_1), (x_2, y_2)) \in \mathbf{RP}(N_2(X), \mathcal{G}_2)$.

**Step 2.** Now we assume that (5.1) holds for $d - 1$, where $d \geq 3$. Let

$$p_1 : (N_d(X), \mathcal{G}_d) \to (N_{d-1}(X), \mathcal{G}_{d-1}), \quad (x_1, \ldots, x_{d-1}, x_d) \mapsto (x_1, \ldots, x_{d-1}),$$

be the projection to the first $d-1$ coordinates. By Lemma 5.3, we need to show that for some $z \in X_1$,

$$(\pi_1^{(d)})^{-1}(z^{(d)}) \times (\pi_1^{(d)})^{-1}(z^{(d)}) \subseteq \mathbf{RP}(N_d(X), \mathscr{G}_d).$$

Let

$$(x_1, \dots, x_d), (y_1, \dots, y_d) \in (\pi_1^{(d)})^{-1}(z^{(d)}).$$

Then via the mapping $p_1$, the points $(x_1, \dots, x_{d-1}), (y_1, \dots, y_{d-1})$ are in $N_{d-1}(X)$ and in $(\pi_1^{(d-1)})^{-1}(z^{(d-1)})$. By the inductive hypothesis, $((x_1, \dots, x_{d-1}), (y_1, \dots, y_{d-1})) \in \mathbf{RP}(N_{d-1}(X), \mathscr{G}_{d-1})$, i.e.,

$$(x_1, \dots, x_{d-1}) \underset{\mathscr{G}_{d-1}}{\sim} (y_1, \dots, y_{d-1}).$$

By Theorem 2.1, there are $x'_d, y'_d \in X$ such that $(x_1, \dots, x_{d-1}, x'_d), (y_1, \dots, y_{d-1}, y'_d) \in N_d(X)$ with

$$(x_1, \dots, x_{d-1}, x'_d) \underset{\mathscr{G}_d}{\sim} (y_1, \dots, y_{d-1}, y'_d).$$

If we can show that

$$(x_1, \dots, x_{d-1}, x_d) \underset{\mathscr{G}_d}{\sim} (x_1, \dots, x_{d-1}, x'_d), \quad (y_1, \dots, y_{d-1}, y_d) \underset{\mathscr{G}_d}{\sim} (y_1, \dots, y_{d-1}, y'_d),$$

then since **RP** is an equivalence relation, we have

$$(x_1, \dots, x_{d-1}, x_d) \underset{\mathscr{G}_d}{\sim} (y_1, \dots, y_{d-1}, y_d),$$

that is, (5.1) holds for $d$. So we only need to show $(x_1, \dots, x_{d-1}, x_d) \underset{\mathscr{G}_d}{\sim} (x_1, \dots, x_{d-1}, x'_d)$, and similarly we will have $(y_1, \dots, y_{d-1}, y_d) \underset{\mathscr{G}_d}{\sim} (y_1, \dots, y_{d-1}, y'_d)$.

Since $(x_1, \dots, x_{d-1}, x_d) \in N_d(X)$, for a fixed $x \in \pi_1^{-1}(z)$ there is some sequence $\{g_i\}_{i \in \mathbb{N}} \subseteq \mathscr{G}_d$ such that

$$g_i(x_1, \dots, x_{d-1}, x_d) \to x^{(d)}, \quad i \to \infty.$$

Without loss of generality, we assume that

$$g_i(x_1, \dots, x_{d-1}, x'_d) \to (x^{(d-1)}, y), \quad i \to \infty,$$

for some $y \in X$. Since $(X, T)$ is a pro-nilsystem, it is distal and so is $(N_d(X), \mathscr{G}_d)$. In particular, $((x_1, \dots, x_{d-1}, x_d), (x_1, \dots, x_{d-1}, x'_d))$ is $\mathscr{G}_d^{(2)}$-minimal, where $\mathscr{G}_d^{(2)} = \{(g, g) : g \in \mathscr{G}_d\}$. Hence

$$\big((x_1, \dots, x_{d-1}, x_d), (x_1, \dots, x_{d-1}, x'_d)\big) \in \overline{\mathcal{O}}\big((x^{(d)}, (x^{(d-1)}, y)), \mathscr{G}_d^{(2)}\big).$$

Now if $x^{(d)} \underset{\mathscr{G}_d}{\sim} (x^{(d-1)}, y)$, then by the $\mathscr{G}_d^{(2)}$-invariance of $\mathbf{RP}(N_d(X), \mathscr{G}_d)$ we find that $(x_1, \dots, x_{d-1}, x_d) \underset{\mathscr{G}_d}{\sim} (x_1, \dots, x_{d-1}, x'_d)$. Thus it is left to show that

$$x^{(d)} \underset{\mathscr{G}_d}{\sim} (x^{(d-1)}, y).$$

By Lemma 2.9 and since $(x^{(d-1)}, y) \in N_d(X)$, we have $(x, y) \in \mathbf{RP}^{[d-2]}(X, T)$. Thus $(x, y) \in \mathbf{RP}_{\pi_{d-3}}(X, T)$ by Lemma 2.10, where $\pi_{d-3} : X \to X_{d-3}$. Note that when $d = 3$, $X_0$ is the trivial system and $\mathbf{RP}_{\pi_0}(X, T) = \mathbf{RP}(X, T)$.

We proceed to prove that

$$x^{(d-1)} \underset{\tau_{d-1}}{\sim} (x^{(d-2)}, y).$$

If $d = 3$, then by Step 1, $(x, x) \underset{T\times T^2}{\sim} (x, y)$. If $d \geq 4$, then we argue as follows. Let $\varepsilon > 0$. Since $(x, y) \in \mathbf{RP}_{\pi_{d-3}}(X, T) = \mathbf{RP}_{\pi_{d-3}}(X, T^{d-1})$, there are $x', y' \in X$, $w \in X_{d-3}$ and $n \in \mathbb{Z}$ such that

$$\rho(x, x') < \varepsilon, \rho(y, y') < \varepsilon, \pi_{d-3}(x') = \pi_{d-3}(y') = w, \quad \rho(T^{(d-1)n}x', T^{(d-1)n}y') < \varepsilon.$$

By Lemma 3.4 we know that $(\pi_{d-3}^{(d-1)})^{-1}N_{d-1}(X_{d-3}) = N_{d-1}(X)$. Since $w^{(d-1)} \in N_{d-1}(X_{d-3})$, we deduce

$$(\pi_{d-3}^{-1}(w))^{d-1} = (\pi_{d-3}^{(d-1)})^{-1}(w^{(d-1)}) \subseteq N_{d-1}(X).$$

This implies that $(x')^{(d-1)}, ((x')^{(d-2)}, y') \in N_{d-1}(X)$. To sum up, for each $\varepsilon > 0$, there are $(x')^{(d-1)}, ((x')^{(d-2)}, y') \in N_{d-1}(X)$ and $n \in \mathbb{Z}$ such that $\rho_{d-1}(x^{(d-1)}, (x')^{(d-1)}) < \varepsilon$, $\rho_{d-1}((x^{(d-2)}, y), ((x')^{(d-2)}, y')) < \varepsilon$ and

$$\rho_{d-1}\big(\tau_{d-1}^n((x')^{(d-1)}), \tau_{d-1}^n((x')^{(d-2)}, y')\big) < \varepsilon,$$

which implies that

$$x^{(d-1)} \underset{\tau_{d-1}}{\sim} (x^{(d-2)}, y).$$

We continue our proof. For each $\varepsilon > 0$, let $U$ and $V$ be open neighborhoods of $x^{(d)}$ and $(x^{(d-1)}, y)$ in $N_d(X)$ with $\operatorname{diam}(U), \operatorname{diam}(V) < \varepsilon/2$. Let

$$p_2 : (N_d(X), \mathcal{G}_d) \to (N_{d-1}(X), \mathcal{G}_{d-1}), \quad (x_1, x_2, \dots, x_d) \mapsto (x_2, \dots, x_d),$$

be the projection to the last $d-1$ coordinates. Then $p_2(U)$ and $p_2(V)$ are open neighborhoods of $x^{(d-1)}$ and $(x^{(d-2)}, y)$ in $N_{d-1}(X)$ respectively. Since $x^{(d-1)} \underset{\tau_{d-1}}{\sim} (x^{(d-2)}, y)$, there are $\mathbf{y} \in p_2(U)$, $\mathbf{y}' \in p_2(V)$ and $n \in \mathbb{Z}$ such that

$$\rho_{d-1}(\tau_{d-1}^n(\mathbf{y}), \tau_{d-1}^n(\mathbf{y}')) < \varepsilon.$$

There are $y_1, y_1' \in X$ such that $(y_1, \mathbf{y}) \in U$ and $(y_1', \mathbf{y}') \in V$. As $\operatorname{diam}(U), \operatorname{diam}(V) < \varepsilon/2$, it follows that $\rho(y_1, y_1') < \varepsilon$. This implies that

$$\rho_d\big((\mathrm{id} \times \tau_{d-1}^n(y_1, \mathbf{y}), (\mathrm{id} \times \tau_{d-1}^n(y_1', \mathbf{y}')\big) < \varepsilon. \tag{5.2}$$

This shows that

$$x^{(d)} \underset{\mathcal{G}_d}{\sim} (x^{(d-1)}, y),$$

since $\mathcal{G}_d$ is also generated by $\sigma_d$ and $\mathrm{id} \times \tau_{d-1}$. The proof is complete. ∎

*5.2. Proof of Theorem B for general systems*

Actually, we will show more. First we need a lemma.

**Lemma 5.6.** *Let $\pi : (X, \Gamma) \to (Y, \Gamma)$ be a factor map between minimal t.d.s. and $d \in \mathbb{N}$, where $\Gamma$ is abelian. If $\pi$ is proximal, then the maximal $d$-step pro-nilfactor of $(X, \Gamma)$ is the same as the one of $(Y, \Gamma)$, i.e., $X_d(X) = X_d(Y)$.*

*Proof.* First we have the following commutative diagram:

$$\begin{array}{ccc} X & \xrightarrow{\pi_{d,X}} & X_d(X) \\ \pi\big\downarrow & & \big\downarrow\pi' \\ Y & \xrightarrow{\pi_{d,Y}} & X_d(Y) \end{array}$$

We need to show that $X_d(X) = X_d(Y)$. Otherwise there are $x_1' \neq x_2' \in X_d(X)$ such that $\pi'(x_1') = \pi'(x_2') = z$. Choose $x_1, x_2 \in X$ such that $\pi_{d,X}(x_1) = x_1'$ and $\pi_{d,X}(x_2) = x_2'$. Let $y_1 = \pi(x_1), y_2 = \pi(x_2)$. Since $\pi' \circ \pi_{d,X} = \pi_{d,Y} \circ \pi$, we have $\pi_{d,Y}(y_1) = \pi_{d,Y}(y_2) = z$, i.e., $(y_1, y_2) \in \mathbf{RP}^{[d]}(Y)$.

By Theorem 2.6, there are $(\tilde{x}_1, \tilde{x}_2) \in \mathbf{RP}^{[d]}(X)$ such that $\pi \times \pi(\tilde{x}_1, \tilde{x}_2) = (y_1, y_2)$. Since $\pi$ is proximal, $(\tilde{x}_1, x_1), (\tilde{x}_2, x_2) \in \mathbf{P}(X)$. Since $\mathbf{P} \subseteq \mathbf{RP}^{[d]}$, by Theorem 2.5 we have

$$(x_1, x_2) \in \mathbf{RP}^{[d]}(X).$$

It follows that $x_1' = \pi_{d,X}(x_1) = \pi_{d,X}(x_2) = x_2'$, a contradiction. ■

**Theorem 5.7.** *Let $(X, T)$ be minimal and $d \in \mathbb{N}$. Then for each $k \in \mathbb{N}$ the maximal $k$-step pro-nilfactor of $(N_d(X), \langle\sigma_d, \tau_d\rangle)$ is the same as the one of $(N_d(X_\infty), \langle\sigma_d, \tau_d\rangle)$. Moreover, the maximal equicontinuous factor of $(N_d(X), \langle\sigma_d, \tau_d\rangle)$ is $(N_d(X_1), \langle\sigma_d, \tau_d\rangle)$, where $X_1 = X_{\mathrm{eq}}$ is the maximal equicontinuous factor of $X$.*

*Proof.* Let $X_\infty$ be the $\infty$-step pro-nilfactor of $X$. Then by Theorem A we have the diagram

$$\begin{array}{ccc} X & \xleftarrow{\sigma^*} & X^* \\ \pi_\infty\big\downarrow & & \big\downarrow\pi^* \\ X_\infty & \xleftarrow{\tau^*} & X_\infty^* \end{array}$$

where $\sigma^*, \tau^*$ are almost one-to-one and $\pi^*$ is open, and there is a dense $G_\delta$ subset $\Omega$ of $X^*$ such that for each $x \in \Omega$ and each $l \in \mathbb{N}$ the orbit closure of $x^{(l)}$ under $\tau_l$ is $(\pi^*)^{(l)}$-saturated.

The diagram above induces the commutative diagram

$$\begin{array}{ccc} (N_d(X), \mathcal{G}_d) & \xleftarrow{(\sigma^*)^{(d)}} & (N_d(X^*), \mathcal{G}_d) \\ \pi_\infty^{(d)}\big\downarrow & & \big\downarrow(\pi^*)^{(d)} \\ (N_d(X_\infty), \mathcal{G}_d) & \xleftarrow{(\tau^*)^{(d)}} & (N_d(X_\infty^*), \mathcal{G}_d) \end{array}$$

where $(\sigma^*)^{(d)}, (\tau^*)^{(d)}$ are almost one-to-one, $\mathcal{G}_d = \langle \sigma_d, \tau_d \rangle$. Note that an almost one-to-one extension is proximal. By Lemma 5.6, for $k \in \mathbb{N}$, the maximal $k$-step pro-nilfactor of $(N_d(X), \mathcal{G}_d)$ is the same as the one of $(N_d(X^*), \mathcal{G}_d)$, and the maximal $k$-step pro-nilfactor of $(N_d(X_\infty), \mathcal{G}_d)$ is the same as the one of $(N_d(X_\infty^*), \mathcal{G}_d)$. Thus, for $k \in \mathbb{N}$, to show that the maximal $k$-step pro-nilfactor of $(N_d(X), \mathcal{G}_d)$ is the same as the one of $(N_d(X_\infty), \mathcal{G}_d)$, it suffices to show that the maximal $k$-step pro-nilfactor of $(N_d(X^*), \mathcal{G}_d)$ is the same as the one of $(N_d(X_\infty^*), \mathcal{G}_d)$. To that end, it suffices to show that

$$R_{(\pi^*)^{(d)}} \subseteq \mathbf{RP}^{[k]}(N_d(X^*), \mathcal{G}_d).$$

By Corollary 5.4, we show that for some $x \in X^*$,

$$\{x^{(d)}\} \times ((\pi^*)^{-1}(\pi^*(x)))^d \subseteq \mathbf{RP}^{[k]}(N_d(X^*), \mathcal{G}_d).$$

Let $x \in \Omega$ and let $(z_1, \dots, z_d) \in ((\pi^*)^{-1}(\pi^*(x)))^d$. We show $(x^{(d)}, (z_1, \dots, z_d)) \in \mathbf{RP}^{[k]}(N_d(X^*), \mathcal{G}_d)$.

For each $\varepsilon > 0$, let $U = B_\varepsilon(x)$ and $V_i = B_\varepsilon(z_i)$, $1 \le i \le d$. We claim that there is $n \in \mathbb{Z}$ such that

$$\mathbf{x} = (T^n x, T^{2n} x, \dots, T^{dn} x) \in U \times U \times \cdots \times U,$$
$$\tau_d^{dnj}(\mathbf{x}) \in V_1 \times \cdots \times V_d, \quad 1 \le j \le k+1.$$

Since $x \in \Omega$, for $l = (k+1)d^2 + d$ the orbit closure of $x^{(l)}$ under $\tau_l$ is $(\pi^*)^{(l)}$-saturated. Note that for $i_1, i_2 \in \{1, \dots, d\}$ and $j_1, j_2 \in \{1, \dots, k+1\}$, if $(i_1, j_1) \ne (i_2, j_2)$, then $(dj_1 + 1)i_1 \ne (dj_2 + 1)i_2$. Now we define

$$\mathbf{W} = W_1 \times \cdots \times W_l \subseteq X^l \quad \text{and} \quad \mathbf{w} = (w_1, \dots, w_l) \in X^l$$

as follows: for $r \in \{1, \dots, l\}$,

$$W_r = \begin{cases} V_i & \text{if } r = (dj+1)i \text{ for } 1 \le i \le d, 1 \le j \le k+1, \\ U & \text{else,} \end{cases}$$
$$w_r = \begin{cases} z_i & \text{if } r = (dj+1)i \text{ for } 1 \le i \le d, 1 \le j \le k+1, \\ x & \text{else.} \end{cases}$$

Thus $\mathbf{W}$ is an open neighborhood of $\mathbf{w}$ in $X^l$, and note that

$$(w_1, \dots, w_d) = x^{(d)},$$
$$(w_{dj+1}, w_{(dj+1)2}, \dots, w_{(dj+1)d}) = (z_1, z_2, \dots, z_d), \quad \forall 1 \le j \le k+1.$$

It is clear that $(\pi^*)^{(l)}(\mathbf{w}) = (\pi^*)^{(l)}(x^{(l)})$, and hence $\mathbf{w} \in ((\pi^*)^{(l)})^{-1}((\pi^*)^{(l)}(x^{(l)}))$. Since the orbit closure of $x^{(l)}$ under $\tau_l$ is $(\pi^*)^{(l)}$-saturated, there is $n \in \mathbb{Z}$ such that $\tau_l^n(x^{(l)}) \in \mathbf{W}$, that is, $T^{rn}x \in W_r$, $1 \le r \le l$. In particular,

$$\mathbf{x} = \tau_d^n(x^{(d)}) \in U \times \cdots \times U,$$

and for $1 \le j \le k+1$,

$$\tau_d^{dnj}(\mathbf{x}) = (T^{(dj+1)n}x, T^{(dj+1)2n}x, \dots, T^{(dj+1)dn}x) \in V_1 \times V_2 \times \cdots \times V_d.$$

Let $\mathbf{y} = \tau_d^{dn}(\mathbf{x})$. We have

$$\rho_d(\tau_d^{dnj}(\mathbf{x}), \tau_d^{dnj}(\mathbf{y})) \le d\varepsilon \quad \text{for } j = 1, \dots, k.$$

This implies that $(x^{(d)}, (z_1, \dots, z_d)) \in \mathbf{RP}^{[k]}(N_d(X^*), \mathcal{G}_d)$ by definition. Thus we have proved that the maximal $k$-step nilfactor factor of $(N_d(X^*), \mathcal{G}_d)$ is the same as the maximal $k$-step nilfactor of $(N_d(X^*_\infty), \mathcal{G}_d)$. It follows that the maximal $k$-step nilfactor of $(N_d(X), \mathcal{G}_d)$ is the same as the maximal $k$-step nilfactor of $(N_d(X_\infty), \mathcal{G}_d)$.

Moreover, the maximal equicontinuous factor of $(N_d(X), \mathcal{G}_d)$ is the same as the maximal equicontinuous factor of $(N_d(X_\infty), \mathcal{G}_d)$. By Theorem 5.5, the latter factor is $(N_d(X_1), \mathcal{G}_d)$. The proof is complete. ■

A similar proof yields the following result.

**Theorem 5.8.** *Let $(X, T)$ be a minimal t.d.s. with $d, k \in \mathbb{N}$. Then there is a dense $G_\delta$ subset $\Omega$ of $X$ such that for each $x \in \Omega$, the maximal $k$-step pro-nilfactor of $\overline{\mathcal{O}}(x^{(d)}, \tau_d)$ is the same as the one of $\overline{\mathcal{O}}((\pi_\infty x)^{(d)}, \tau_d)$.*

*Proof.* The proof is a modification of the previous one. Assume first we have the same diagram as in the proof of Theorem 5.7. By Theorem 4.2 there is a dense $G_\delta$ subset $\Omega^*$ of $X^*$ such that for each $x \in \Omega^*$ and each $l \in \mathbb{N}$ the orbit closure of $x^{(l)}$ under $\tau_l$ is $(\pi^*)^{(l)}$-saturated. It is clear that there is a dense $G_\delta$ set $\Omega$ of $X$ such that for each $x \in \Omega$, $(\sigma^*)^{-1}(x) \cap \Omega^*$ is not empty. By the same analysis as in the proof of Theorem 5.7, it remains to show that for each $x \in \Omega^*$, the maximal $k$-step pro-nilfactor of $\overline{\mathcal{O}}(x^{(d)}, \tau_d)$ is the same as the one of $\overline{\mathcal{O}}((\pi^* x)^{(d)}, \tau_d)$.

The result is clear for $d = 1$. We now assume that $d \ge 2$. Assume that $(x_1, \dots, x_d)$, $(y_1, \dots, y_d) \in \overline{\mathcal{O}}(x^{(d)}, \tau_d)$ with $\pi^*(x_i) = \pi^*(y_i)$ for a given $x \in \Omega^*$. We will show that

$$((x_1, \dots, x_d), (y_1, \dots, y_d)) \in \mathbf{RP}^{[k]}(\overline{\mathcal{O}}(x^{(d)}, \tau_d), \tau_d)$$

for each $k \in \mathbb{N}$.

To do this, for a given $k \in \mathbb{N}$ and each $\varepsilon > 0$, let $U_i = B_\varepsilon(x_i)$ and $V_i = B_\varepsilon(y_i)$, $1 \le i \le d$. Then by the argument in the proof of Theorem 5.7 there is $n \in \mathbb{Z}$ such that

$$\mathbf{x} = (T^n x, \dots, T^{dn} x) \in U_1 \times \dots \times U_d,$$
$$\tau_d^{dnj}(\mathbf{x}) \in V_1 \times \dots \times V_d, \quad 1 \le j \le k+1.$$

Let $\mathbf{y} = \tau_d^{dn}(\mathbf{x})$. Then $\mathbf{y} = \tau_d^{dn+n}(x^{(d)}) \in \overline{\mathcal{O}}(x^{(d)}, \tau_d)$ and

$$\rho_d(\tau_d^{dnj}(\mathbf{x}), \tau_d^{dnj}(\mathbf{y})) \le d\varepsilon \quad \text{for } j = 1, \dots, k.$$

This implies that $((x_1, \dots, x_d), (y_1, \dots, y_d)) \in \mathbf{RP}^{[k]}(\overline{\mathcal{O}}(x^{(d)}, \tau_d), \tau_d)$ by definition of $\mathbf{RP}^{[k]}$. The proof is complete. ■

We remark that in general $\overline{\mathcal{O}}(x^{(d)}, \tau_d)$ is not minimal. Thus, to show Theorem 5.8 we cannot use exactly the same arguments as for Theorem 5.7.

*5.3. Proof of Theorem C*

First we have the following simple observation.

**Lemma 5.9.** *Let $(X, T)$ be a minimal t.d.s. We have:*
(1) $N_d(X, T) = N_d(X, T^{-1})$ *and* $N_d(X, T^n) \subseteq N_d(X, T)$ *for any* $n \in \mathbb{Z} \setminus \{0\}$.
(2) $N_d(X, T) = \bigcup_{l,k=0}^{n-1} \sigma_d^l \tau_d^k N_d(X, T^n)$.

The following proposition will be used in the proof of Theorem C.

**Proposition 5.10.** *If $(X, T)$ is a minimal equicontinuous system and $(X, T^n)$ is minimal for some $n \in \mathbb{N}$, then $N_d(X, T) = N_d(X, T^n)$ for any $d \in \mathbb{N}$.*

*Proof.* Let $d \in \mathbb{N}$. It is clear that $N_d(X, T^n) \subseteq N_d(X, T)$. To show the opposite inclusion, let $x \in X$. It suffices to show that

$$\mathcal{O}(x^{(d)}, \langle \sigma_d, \tau_d \rangle) = \mathcal{O}((x, \dots, x), \langle \sigma_d, \tau_d \rangle) \subseteq N_d(X, T^n).$$

Since each point in $\mathcal{O}(x^{(d)}, \langle \sigma_d, \tau_d \rangle)$ has the form $(T^{k+l}x, T^{k+2l}x, \dots, T^{k+dl}x)$ for some $k, l \in \mathbb{Z}$, we need to show that

$$(T^{k+l}x, T^{k+2l}x, \dots, T^{k+dl}x) \in N_d(X, T^n).$$

By the assumption that $(X, T^n)$ is minimal, there are sequences $\{p_i\}_{i\in\mathbb{N}}, \{q_i\}_{i\in\mathbb{N}} \subseteq \mathbb{Z}$ such that

$$T^{np_i}x \to T^k x, \quad T^{nq_i} \to T^l x, \quad i \to \infty.$$

Let $\varepsilon > 0$. Since $(X, T)$ is equicontinuous, there is $\delta > 0$ such that if $\rho(x, y) < \delta$ then $\rho(T^i x, T^i y) < \varepsilon$ for all $i \in \mathbb{Z}$.

Since $T^{nq_i}x \to T^l x$ as $i \to \infty$, there is $N \in \mathbb{N}$ such that if $i \geq N$ then $\rho(T^{nq_i}x, T^l x) < \delta$. This implies that for any $1 \leq j \leq d$ and $i \geq N$ we have

$$\rho(T^{jnq_i}x, T^{(j-1)nq_i+l}x) < \varepsilon,$$
$$\rho(T^{(j-1)nq_i+l}x, T^{(j-2)nq_i+2l}x) < \varepsilon, \quad \dots,$$
$$\rho(T^{nq_i+(j-1)l}x, T^{jl}x) < \varepsilon,$$

which implies that

$$\rho(T^{jnq_i}x, T^{jl}x) < j\varepsilon \leq d\varepsilon, \quad \forall 1 \leq j \leq d.$$

Since $\varepsilon$ is arbitrary,

$$(T^{nq_i}x, T^{n2q_i}x, \dots, T^{ndq_i}x) \to (T^l x, T^{2l}x, \dots, T^{dl}x), \quad i \to \infty.$$

It follows that there are $p_i', q_i' \in \mathbb{Z}$ with

$$(T^{np_i'+nq_i'}x, T^{np_i'+n2q_i'}x, \dots, T^{np_i'+ndq_i'}x) \to (T^{k+l}x, T^{k+2l}x, \dots, T^{k+dl}x), \quad i \to \infty.$$

Thus

$$(T^{k+l}x, T^{k+2l}x, \dots, T^{k+dl}x) \in N_d(X, T^n).$$

The proof is complete. ■

**Theorem 5.11.** *Let $G$ be an abelian group and $\Gamma$ be a subgroup of finite index. Let $(X, G)$ be a minimal t.d.s., and let $\pi : (X, G) \to (X_{\text{eq}}, G)$ be the factor map to its maximal equicontinuous factor. Then $(X, \Gamma)$ is minimal if and only if $(X_{\text{eq}}, \Gamma)$ is minimal.*

*Proof.* Since $\pi : (X, \Gamma) \to (X_{\text{eq}}, \Gamma)$ is a factor map, the minimality of $(X, \Gamma)$ implies the minimality of $(X_{\text{eq}}, \Gamma)$. Now we show the converse.

Assume that $(X_{\text{eq}}, \Gamma)$ is minimal; we will show that $(X, \Gamma)$ is minimal. Note that $\pi : (X, G) \to (X_{\text{eq}}, G)$ is the factor map to the maximal equicontinuous factor. Since $[G : \Gamma] < \infty$, we have $\mathbf{RP}(X, G) = \mathbf{RP}(X, \Gamma)$ by the same proof as that of Lemma 2.7. It follows that $\pi : (X, \Gamma) \to (X_{\text{eq}}, \Gamma)$ is also the factor map to its maximal equicontinuous factor. Thus any equicontinuous factor of $(X, \Gamma)$ is also a factor of $(X_{\text{eq}}, \Gamma)$. In particular, any equicontinuous factor of $(X, \Gamma)$ is minimal as $(X_{\text{eq}}, \Gamma)$ is minimal.

If $(X, \Gamma)$ is not minimal, then there is a non-empty $\Gamma$-minimal subset $W$ of $(X, \Gamma)$ with $W \neq X$. Since $[G : \Gamma] < \infty$, there are $h_1, \dots, h_m \in G$ with $m \in \mathbb{N}$ such that

$$G = \bigcup_{i=1}^{m} h_i \Gamma.$$

Since $(W, \Gamma)$ is minimal and $G$ is abelian, $(h_i W, \Gamma)$ is also minimal for all $i \in \{1, \dots, m\}$. Note that $\bigcup_{i=1}^{m} h_i W$ is $G$-invariant, and we have $X = \bigcup_{i=1}^{m} h_i W$ as $(X, G)$ is minimal. Since minimal subsets are either identical or disjoint, there is a subset $\{g_1, \dots, g_r\} \subseteq \{h_1, \dots, h_m\}$, $2 \le r \le m$, such that

$$X = \bigsqcup_{i=1}^{r} g_i W,$$

where $\bigsqcup$ means disjoint union. Now define

$$\phi : (X, \Gamma) \to (\{1, \dots, r\}, \Gamma), \quad g_i W \mapsto \{i\}, \forall i \in \{1, \dots, r\}.$$

Since $(g_i W, \Gamma)$ is minimal for $i \in \{1, \dots, r\}$, $\Gamma = \text{id}$ on $\{1, \dots, r\}$. As $r \ge 2$, we see that $(\{1, \dots, r\}, \Gamma)$ is a non-minimal equicontinuous factor of $(X, \Gamma)$, which is a contradiction. Thus $(X, \Gamma)$ is minimal. The proof is complete. ■

Now we are ready to show Theorem C.

*Proof of Theorem C.* Let $(X, T)$ be minimal and $k \ge 2$. It suffices to show that if $(X, T^k)$ is minimal, then $N_d(X, T) = N_d(X, T^k)$ for each $d \in \mathbb{N}$. This is an application of Theorem B and some previous results.

Set $\mathcal{G}_d(T) = \langle \sigma_d(T), \tau_d(T) \rangle$. Then $\mathcal{G}_d(T^k)$ is a subgroup of $\mathcal{G}_d(T)$ of finite index. Since $(X, T^k)$ is minimal, we infer that $(X_{\text{eq}}, T^k)$ is minimal. Thus by Proposition 5.10, $N_d(X_{\text{eq}}, T) = N_d(X_{\text{eq}}, T^k)$, and hence $(N_d(X_{\text{eq}}, T), \mathcal{G}_d(T^k)) = (N_d(X_{\text{eq}}, T^k), \mathcal{G}_d(T^k))$ is minimal.

By Theorem B, the maximal equicontinuous factor of $(N_d(X, T), \mathcal{G}_d(T))$ is $(N_d(X_{\text{eq}}, T), \mathcal{G}_d(T))$. As $\mathcal{G}_d(T^k)$ is a subgroup of $\mathcal{G}_d(T)$ of finite index, by Theorem 5.11, the minimality of $(N_d(X_{\text{eq}}, T), \mathcal{G}_d(T^k))$ implies that $(N_d(X, T), \mathcal{G}_d(T^k))$ is minimal. Since $(X, T^k)$ is minimal, by Theorem 2.16, $(N_d(X, T^k), \mathcal{G}_d(T^k))$ is minimal. It follows that $N_d(X, T) = N_d(X, T^k)$ as $N_d(X, T^k) \subseteq N_d(X, T)$. The proof is complete. ∎

## 6. Proofs of Theorems D–F

In this section we prove Theorems D–F. First we need a lemma.

### 6.1. A key lemma

To prove Theorem D we need Lemma 6.1 below. Given a compact metric space $Z$ and a sequence of non-empty closed subsets $A_n \subseteq Z$, define

$$\liminf_{n\to\infty} A_n = \Big\{z \in Z : \exists z_n \in A_n \colon z = \lim_{n\to\infty} z_n\Big\},$$
$$\limsup_{n\to\infty} A_n = \Big\{z \in Z : \text{for some subsequence } \{n_i\}, \exists z_i \in A_{n_i} \colon z = \lim_{n\to\infty} z_i\Big\}.$$

When $A := \liminf_{n\to\infty} A_n = \limsup_{n\to\infty} A_n$, we write $A = \lim_{n\to\infty} A_n$ and call $A$ the *limit* of the sequence $\{A_n\}_{n\in\mathbb{N}}$. In fact, in this case the set $A$ is the limit of $\{A_n\}_{n\in\mathbb{N}}$ in the space $2^Z$ of non-empty closed subsets of $Z$ with the Hausdorff metric.

**Lemma 6.1.** *Let $(X, T)$ be a t.d.s. Assume that for some $d, n \in \mathbb{N}$,*

$$N_{d+1}(X, T) = N_{d+1}(X, T^n).$$

*Then the set*

$$\Omega_d = \{x \in X : \overline{\mathcal{O}}(x^{(d)}, \tau_d(T)) = \overline{\mathcal{O}}(x^{(d)}, \tau_d(T^n))\}$$

*is a dense $G_\delta$ subset of $X$, where $x^{(d)} = (x, \dots, x) \in X^d$, $\tau_d(T) = T \times T^2 \times \cdots \times T^d$.*

*Proof.* Let $\tau'_{d+1}(T) = \mathrm{id} \times T \times T^2 \times \cdots \times T^d = \mathrm{id} \times \tau_d(T)$. Note that

$$N_{d+1}(T) = \overline{\mathcal{O}}(\Delta_{d+1}, \tau_{d+1}(T)) = \overline{\mathcal{O}}(\Delta_{d+1}, \tau'_{d+1}(T)).$$

For $x \in X$, let

$$C(x) = \overline{\mathcal{O}}(x^{(d+1)}, \tau'_{d+1}(T)) = \{x\} \times \overline{\mathcal{O}}(x^{(d)}, \tau_d(T)),$$
$$D(x) = \{(x, u_1, \dots, u_d) : \exists x_i \in X, n_i \in \mathbb{Z} \colon (\tau'_{d+1}(T))^{n_i}(x_i^{(d+1)}) \to (x, u_1, \dots, u_d)\}.$$

Then it is clear that $C(x) \subseteq D(x) = N_{d+1}(T) \cap (\{x\} \times X^d)$.

**Claim 1.** *The map $C : X \to 2^{N_{d+1}(T)}$, $x \mapsto C(x)$, is lower semicontinuous, that is, $x_i \to x$, $i \to \infty$ implies that* $\liminf_{i\to\infty} C(x_i) \supset C(x)$.

In fact, by definition it follows from $x_i \to x$ and $x_i^{(d+1)} \in C(x_i)$ for all $i$ that

$$x^{(d+1)} \in \liminf_{i\to\infty} C(x_i).$$

Now for each $k \in \mathbb{Z}$, since $(\tau'_{d+1}(T))^k(x_i^{(d+1)}) \in C(x_i)$ for all $i$, one has

$$(\tau'_{d+1}(T))^k(x^{(d+1)}) \in \liminf_{i\to\infty} C(x_i).$$

Thus

$$\mathcal{O}(x^{(d+1)}, \tau'_{d+1}(T)) \subseteq \liminf_{i\to\infty} C(x_i).$$

It follows that

$$\overline{\mathcal{O}}(x^{(d+1)}, \tau'_{d+1}(T)) \subseteq \liminf_{i\to\infty} C(x_i)$$

as claimed.

The following claim is a direct consequence of the definition of the map $D$.

**Claim 2.** *For every $x \in X$,*

$$D(x) \subseteq \bigcup \Big\{\liminf_{i\to\infty} C(x_i) : x_i \to x\Big\}.$$

Let $X_0 \subseteq X$ be the set of $C$-continuity points. Then it is well known that $X_0$ is a dense $G_\delta$ subset of $X$ [8]. Now for $x_0 \in X_0$, for every sequence $x_i \to x_0$ as $i \to \infty$, we have $\liminf_{i\to\infty} C(x_i) = \lim_{i\to\infty} C(x_i) = C(x_0)$. It follows that $D(x_0) \subseteq C(x_0) \subseteq D(x_0)$, whence

$$D(x_0) = C(x_0). \tag{6.1}$$

In the following discussion we consider both $\tau'_{d+1}(T)$ and $\tau'_{d+1}(T^n)$. We will use the symbols $C_{\tau'_{d+1}(T)}$, $D_{\tau'_{d+1}(T)}$, $D_{\tau'_{d+1}(T^n)}$, and $C_{\tau'_{d+1}(T^n)}$, whose meaning is clear.

By assumption,

$$\begin{aligned} N_{d+1}(T) &= \bigcup_{x\in X} N_{d+1}(T) \cap (\{x\} \times X^d) \\ &= \bigcup_{x\in X} D_{\tau'_{d+1}(T)}(x) = \bigcup_{x\in X} D_{\tau'_{d+1}(T^n)}(x) = N_{d+1}(T^n), \end{aligned}$$

and therefore $D_{\tau'_{d+1}(T)}(x) = D_{\tau'_{d+1}(T^n)}(x)$ for all $x \in X$. Now let $\Omega_d$ be the intersection of the sets of continuity points of the maps $C_{\tau'_{d+1}(T)}$ and $C_{\tau'_{d+1}(T^n)}$. Then by (6.1), for each $x \in \Omega_d$,

$$C_{\tau'_{d+1}(T)}(x) = D_{\tau'_{d+1}(T)}(x) = D_{\tau'_{d+1}(T^n)}(x) = C_{\tau'_{d+1}(T^n)}(x).$$

Let $\pi : X^{d+1} \to X^d$ be the projection on coordinates $\{2, \ldots, d+1\}$, and we have

$$\overline{\mathcal{O}}(x^{(d)}, \tau_d(T)) = \pi C_{\tau'_{d+1}(T)}.$$

Hence for each $x \in \Omega_d$,

$$\overline{\mathcal{O}}(x^{(d)}, \tau_d(T)) = \pi C_{\tau'_{d+1}(T)} = \pi C_{\tau'_{d+1}(T^n)} = \overline{\mathcal{O}}(x^{(d)}, \tau_d(T^n)).$$

The proof is complete. ■

By Lemma 6.1 we have

**Theorem 6.2.** *Let $(X,T)$ be a minimal t.d.s. and $n,d \geq 2$. Then the following statements are equivalent:*

(1) $N_{d+1}(X,T) = N_{d+1}(X,T^n)$.

(2) *There is a dense $G_\delta$ subset $X_0$ of $X$ such that for any $l \in \mathbb{Z}$ and $x \in X_0$, there is a sequence $\{q_i\} \subseteq \mathbb{Z}$ with*

$$T^{nq_i}x \to T^l x, \quad T^{2nq_i}x \to T^{2l}x, \quad \ldots, \quad T^{dnq_i}x \to T^{dl}x.$$

(3) $\{x \in X : \overline{\mathcal{O}}(x^{(d)}, \tau_d(T)) = \overline{\mathcal{O}}(x^{(d)}, \tau_d(T^n))\}$ *is a dense $G_\delta$ subset of $X$.*

### 6.2. *Proof of Theorem D*

Let $(X,T^k)$ be minimal for some $k \geq 2$ and $d \in \mathbb{N}$. We show that for any $d \in \mathbb{N}$ and any $0 \leq j < k$ there exists a sequence $\{n_i\}$ with $n_i \equiv j \pmod k$ such that $T^{n_i}x \to x$, $T^{2n_i}x \to x, \ldots, T^{dn_i}x \to x$ as $i \to \infty$ for $x$ in a dense $G_\delta$ subset of $X$.

By Theorem C and Lemma 6.1, there is a dense $G_\delta$ subset $X_0$ such that for any $x \in X_0$,

$$\overline{\mathcal{O}}(x^{(d)}, \tau_d) = \overline{\mathcal{O}}(x^{(d)}, \tau_d^k).$$

Now assume that $0 \leq j < k$. Note that

$$(T^{-j}x, T^{-2j}x, \ldots, T^{-dj}x) \in \overline{\mathcal{O}}(x^{(d)}, \tau_d).$$

Thus there is a sequence $m_i \nearrow \infty$ such that

$$T^{km_i}x \to T^{-j}x, T^{2km_i}x \to T^{-2j}x, \ldots, T^{dkm_i}x \to T^{-dj}x, \quad i \to \infty,$$

i.e.,

$$T^{km_i+j}x \to x,\ T^{2km_i+2j}x \to x, \ldots, T^{dkm_i+dj}x \to x, \quad i \to \infty.$$

Then $n_i = km_i + j$ is what we need. The proof is complete. ■

### 6.3. *Proof of Theorem E*

Let $U$ and $V$ be non-empty open subsets of $X$. We are going to show that there is $n \in \mathbb{N}$ such that

$$U \cap T^{-P(n)}V \neq \emptyset.$$

Since $(X,T)$ is minimal, there is $N \in \mathbb{N}$ such that

$$X = \bigcup_{i=1}^{N} T^i U.$$

Let $q(n) = an^2 + bn$. By Bergelson–Leibman's theorem [2, Theorem C] there are $n \in \mathbb{N}$ and $x \in V$ such that

$$T^{q(n)}x \in V, \quad T^{q(2n)}x \in V, \quad \ldots, \quad T^{q(Nn)}x \in V.$$

Thus there is an open neighborhood $V_1$ of $x$ such that $V_1 \subseteq V$ with

$$T^{q(n)}V_1 \subseteq V, \quad T^{q(2n)}V_1 \subseteq V, \quad \dots, \quad T^{q(Nn)}V_1 \subseteq V.$$

By Theorem C there are $k_i, l_i \in \mathbb{Z}$ such that

$$(T^{2an} \times \cdots \times T^{2aNn})^{k_i}(T^{2an} \times \cdots \times T^{2an})^{l_i}(x, \dots, x) \\ \to (Tx, \dots, T^N x) \in TV_1 \times \cdots \times T^N V_1.$$

This implies that there are $k, l \in \mathbb{Z}$ such that

$$T^{2ank+2aln-1}x \in V_1, \quad T^{4ank+2anl-2}x \in V_1, \quad \dots, \quad T^{2aNnk+2anl-N}x \in V_1,$$

i.e.,

$$T^{q(n)+2ank+2aln-1}x \in V, \quad T^{q(2n)+4ank+2anl-2}x \in V, \quad \dots, \\ T^{q(Nn)+2aNnk+2anl-N}x \in V.$$

For $j \in \{1, \dots, N\}$,

$$\begin{aligned} q(jn) &+ 2jank + 2anl - j \\ &= a(jn)^2 + b(jn) + 2jank + 2anl - j \\ &= a(jn+k)^2 + b(jn+k) + c - (ak^2 + bk + c) + 2aln - j \\ &= P(jn+k) - P(k) + 2aln - j. \end{aligned}$$

Let $y = T^{-P(k)+2aln}x$. Then

$$T^{P(jn+k)}(T^{-j}y) \in V, \quad \forall j \in \{1, \dots, N\}.$$

Since $X = \bigcup_{i=1}^N T^i U$, there is some $j_0 \in \{1, \dots, N\}$ such that $T^{-j_0}y \in U$. Thus

$$T^{-j_0}y \in U \cap T^{-P(j_0 n+k)}V.$$

In particular, $U \cap T^{-P(j_0 n+k)}V \neq \emptyset$. Then a standard argument by considering the basis of the topology of $X$ and taking the intersection yields the conclusion of the theorem. The proof is complete. ∎

As a corollary we have

**Corollary 6.3.** *Let $(X, T)$ be a totally minimal t.d.s., and let $k \geq 2$ and $0 \leq j < k$. Let $P(n) = an^2 + bn + c$ with $a, b, c \in \mathbb{Z}, a \neq 0$ be an integral polynomial. Then there is a dense $G_\delta$ subset $\Omega$ of $X$ such that for any $x \in \Omega$, $T^{P(n_i)}(x) \to x$ for some sequence $\{n_i\}$ with $n_i \equiv j \pmod{k}$.*

*Proof.* By putting $Q(n) = P(kn + j)$ and using Theorem E we see that there is a sequence $\{n_i\}_1 \subseteq \mathbb{N}$ with $n_i \equiv j \pmod{k}$ such that $T^{P(n_i)}x \to x$ for $x$ in a dense $G_\delta$ set. ∎

*6.4. Proof of Theorem F*

Let $(X, T)$ be a minimal t.d.s. which is an open extension of its maximal distal factor. We will prove that for any $d \in \mathbb{N}$, $\mathbf{AP}^{[d]} = \mathbf{RP}^{[d]}$. Since $\mathbf{AP}^{[d]} \subseteq \mathbf{RP}^{[d]}$, it suffices to show that $\mathbf{RP}^{[d]} \subseteq \mathbf{AP}^{[d]}$.

Let $d \in \mathbb{N}$ and $\pi_d : X \to X_d$ be the factor to $X_d$. Since $(X, T)$ is an open extension of its maximal distal factor, $\pi_d$ is open. By Theorem 4.3, there is a dense $G_\delta$ subset $\Omega$ of $X$ such that for each $x \in \Omega$, $\overline{\mathcal{O}}(x^{(d+1)}, \tau_{d+1})$ is $\pi_d^{(d+1)}$-saturated.

Let $x \in \Omega$ and $(x, y) \in \mathbf{RP}^{[d]}$. Then for each neighborhood $U$ of $y$, there is $n \in \mathbb{Z}$ such that $T^{jn}x \in U$ for each $1 \leq j \leq d+1$, which implies that $(x, y) \in \mathbf{AP}^{[d]}$ by taking $x' = x$ and $y' = T^n x$ in the definition of $\mathbf{AP}^{[d]}$.

Now let $(x, y) \in \mathbf{RP}^{[d]}$. In each neighborhood $W$ of $(x, y)$, there are $(x', y') \in W$ with $x' \in \Omega$ and $(x', y') \in \mathbf{RP}^{[d]}$ by the openness of $\pi_d$. By what we have just proved, $(x', y') \in \mathbf{AP}^{[d]}$, which implies that $(x, y) \in \mathbf{AP}^{[d]}$ since $\mathbf{AP}^{[d]}$ is closed. This ends the proof. ■

## 7. Some conjectures

With the help of Theorem A and its consequences we have been able to answer several open questions. In fact, Theorem A also opens a window for exploring other natural questions, which we will discuss now.

**Conjecture 1.** Let $(X, T)$ be a totally minimal t.d.s., and $P(n)$ be a non-constant integral polynomial. Then there is a dense $G_\delta$ subset $\Omega$ of $X$ such that for every $x \in \Omega$, the set $\{T^{P(n)}(x) : n \in \mathbb{Z}\}$ is dense in $X$.

Note that in Theorem E we have shown that this is true when $P(n) = an^2 + bn + c$. Also, by [32], Conjecture 1 holds for minimal weakly mixing systems.

**Conjecture 2.** Let $(X, T)$ be a totally minimal t.d.s., $\mathbb{P}$ be the set of prime numbers, and $P(n)$ be a non-constant integral polynomial. Then there is a dense $G_\delta$ subset $\Omega$ of $X$ such that for every $x \in \Omega$, the set $\{T^{P(n)}(x) : n \in \mathbb{P}\}$ is dense in $X$.

In particular, there is some point $x$ such that $\{T^n x : n \in \mathbb{P}\}$ is dense in $X$.

We think that it may be relatively easy to show that Theorem D holds for any finite collection of non-constant integral polynomials $P_i(n)$ under the total minimality assumption (see Corollary 6.3). But we believe that it needs a real work to verify the following conjecture.

**Conjecture 3.** Let $(X, T^k)$ be minimal for some $k \geq 2$ and $d \in \mathbb{N}$. Then for non-constant integral polynomials $P_m(n)$ with $P_m(0) = 0$, $1 \leq m \leq d$, and any $0 \leq j < k$, there is a sequence $\{n_i\}_{i \in \mathbb{N}}$ such that

$$T^{P_1(n_i)}x \to x, \ldots, T^{P_d(n_i)}x \to x, \quad i \to \infty, \tag{7.1}$$

where $n_i \equiv j \pmod{k}$ and $x$ is in a dense $G_\delta$ subset of $X$.

The last conjecture is related to the $\mathbf{AP}^{[d]}$ relation.

**Conjecture 4.** There is a minimal t.d.s. $(X, T)$ with $\mathbf{AP}^{[2]}(X, T) = \Delta_X$, yet $(X, T)$ is not distal.

*Acknowledgments.* We would like to thank V. Bergelson, N. Frantzikinakis and J.-P. Thouvenot for suggesting some of the questions we discuss in this work.

*Funding.* This research is supported by National Natural Science Foundation of China (12031019, 12090012, 11971455, 11731003) and by a grant of the Israel Science Foundation (ISF 1194/19).

## References

[1] Bergelson, V., Host, B., Kra, B.: Multiple recurrence and nilsequences. Invent. Math. **160**, 261–303 (2005) Zbl 1087.28007 MR 2138068

[2] Bergelson, V., Leibman, A.: Polynomial extensions of van der Waerden's and Szemerédi's theorems. J. Amer. Math. Soc. **9**, 725–753 (1996) Zbl 0870.11015 MR 1325795

[3] Bergelson, V., Leibman, A.: A nilpotent Roth theorem. Invent. Math. **147**, 429–470 (2002) Zbl 1042.37001 MR 1881925

[4] Błaszczyk, A., Plewik, S., Turek, S.: Topological multidimensional van der Waerden theorem. Comment. Math. Univ. Carolin. **30**, 783–787 (1989) Zbl 0691.54027 MR 1045910

[5] Bourgain, J.: An approach to pointwise ergodic theorems. In: Geometric aspects of functional analysis (1986/87), Lecture Notes in Mathematics 1317, Springer, Berlin, 204–223 (1988) Zbl 0662.47006 MR 950982

[6] Bourgain, J.: Pointwise ergodic theorems for arithmetic sets. Inst. Hautes Études Sci. Publ. Math. **69**, 5–45 (1989) Zbl 0705.28008 MR 1019960

[7] Bronšteĭn, I. U.: Extensions of minimal transformation groups. Sijthoff and Noordhoff, Alphen aan den Rijn (1979) Zbl 0431.54023 MR 550605

[8] Choquet, G.: Convergences. Ann. Univ. Grenoble Sect. Sci. Math. Phys. (N.S.) **23**, 57–112 (1948) Zbl 0031.28101 MR 0025716

[9] de Vries, J.: Elements of topological dynamics. Mathematics and its Applications 257, Kluwer Academic Publishers Group, Dordrecht (1993) Zbl 0783.54035 MR 1249063

[10] Dong, P., Donoso, S., Maass, A., Shao, S., Ye, X.: Infinite-step nilsystems, independence and complexity. Ergodic Theory Dynam. Systems **33**, 118–143 (2013) Zbl 1268.37019 MR 3009105

[11] Ellis, R., Glasner, S., Shapiro, L.: Proximal-isometric ($\mathcal{P}\mathcal{J}$) flows. Adv. Math. **17**, 213–260 (1975) Zbl 0304.54039 MR 380755

[12] Ellis, R., Gottschalk, W. H.: Homomorphisms of transformation groups. Trans. Amer. Math. Soc. **94**, 258–271 (1960) Zbl 0094.17401 MR 123635

[13] Ellis, R., Keynes, H.: A characterization of the equicontinuous structure relation. Trans. Amer. Math. Soc. **161**, 171–183 (1971) Zbl 0233.54023 MR 282357

[14] Frantzikinakis, N.: The structure of strongly stationary systems. J. Anal. Math. **93**, 359–388 (2004) Zbl 1073.60040 MR 2110334

[15] Frantzikinakis, N., Host, B., Kra, B.: Multiple recurrence and convergence for sequences related to the prime numbers. J. Reine Angew. Math. **611**, 131–144 (2007) Zbl 1126.37005 MR 2361086

[16] Frantzikinakis, N., Lesigne, E., Wierdl, M.: Sets of $k$-recurrence but not $(k+1)$-recurrence. Ann. Inst. Fourier (Grenoble) **56**, 839–849 (2006) Zbl 1123.37001 MR 2266880

[17] Furstenberg, H.: The structure of distal flows. Amer. J. Math. **85**, 477–515 (1963) Zbl 0199.27202 MR 157368

[18] Furstenberg, H.: Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions. J. Anal. Math. **31**, 204–256 (1977) Zbl 0347.28016 MR 498471

[19] Furstenberg, H.: Recurrence in ergodic theory and combinatorial number theory. Princeton University Press, Princeton, NJ (1981) Zbl 0459.28023 MR 603625

[20] Furstenberg, H., Weiss, B.: Topological dynamics and combinatorial number theory. J. Anal. Math. **34**, 61–85 (1979) (1978) Zbl 0425.54023 MR 531271

[21] 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: Convergence in ergodic theory and probability (Columbus, OH, 1993), Ohio State University Mathematical Research Institute Publications 5, de Gruyter, Berlin, 193–227 (1996) Zbl 0869.28010 MR 1412607

[22] Glasner, E.: A topological version of a theorem of Veech and almost simple flows. Ergodic Theory Dynam. Systems **10**, 463–482 (1990) Zbl 0712.54029 MR 1074314

[23] Glasner, E.: Topological ergodic decompositions and applications to products of powers of a minimal transformation. J. Anal. Math. **64**, 241–262 (1994) Zbl 0855.54048 MR 1303514

[24] Glasner, E.: Ergodic theory via joinings. Mathematical Surveys and Monographs 101, American Mathematical Society, Providence, RI (2003) Zbl 1038.37002 MR 1958753

[25] Glasner, E., Gutman, Y., Ye, X.: Higher order regionally proximal equivalence relations for general minimal group actions. Adv. Math. **333**, 1004–1041 (2018) Zbl 1394.37016 MR 3818095

[26] Glasner, E., Huang, W., Shao, S., Ye, X.: Regionally proximal relation of order $d$ along arithmetic progressions and nilsystems. Sci. China Math. **63**, 1757–1776 (2020) Zbl 1451.37015 MR 4145918

[27] Host, B., Kra, B.: An odd Furstenberg–Szemerédi theorem and quasi-affine systems. J. Anal. Math. **86**, 183–220 (2002) Zbl 1013.37003 MR 1894481

[28] Host, B., Kra, B.: Nonconventional ergodic averages and nilmanifolds. Ann. of Math. (2) **161**, 397–488 (2005) Zbl 1077.37002 MR 2150389

[29] Host, B., Kra, B., Maass, A.: Nilsequences and a structure theorem for topological dynamical systems. Adv. Math. **224**, 103–129 (2010) Zbl 1203.37022 MR 2600993

[30] Huang, W., Shao, S., Ye, X.: Nil Bohr-sets and almost automorphy of higher order. Mem. Amer. Math. Soc. **241**, no. 1143, v+83 pp. (2016) Zbl 1375.37029 MR 3476203

[31] Huang, W., Shao, S., Ye, X.: Pointwise convergence of multiple ergodic averages and strictly ergodic models. J. Anal. Math. **139**, 265–305 (2019) Zbl 1434.37001 MR 4041103

[32] Huang, W., Shao, S., Ye, X.: Topological correspondence of multiple ergodic averages of nilpotent group actions. J. Anal. Math. **138**, 687–715 (2019) Zbl 1446.37006 MR 3996054

[33] Leibman, A.: Multiple recurrence theorem for nilpotent group actions. Geom. Funct. Anal. **4**, 648–659 (1994) Zbl 0848.54025 MR 1302335

[34] Leibman, A.: Multiple recurrence theorem for measure preserving actions of a nilpotent group. Geom. Funct. Anal. **8**, 853–931 (1998) Zbl 0917.28017 MR 1650102

[35] McMahon, D. C.: Relativized weak disjointness and relatively invariant measures. Trans. Amer. Math. Soc. **236**, 225–237 (1978) Zbl 0388.54030 MR 467704

[36] Qiu, J., Zhao, J.: Top-nilpotent enveloping semigroups and pro-nilsystems. J. Anal. Math. **148**, 595–636 (2022) Zbl 1510.37021 MR 4525763

[37] Shao, S., Ye, X.: Regionally proximal relation of order $d$ is an equivalence one for minimal systems and a combinatorial consequence. Adv. Math. **231**, 1786–1817 (2012) Zbl 1286.37010 MR 2964624

[38] Veech, W. A.: The equicontinuous structure relation for minimal Abelian transformation groups. Amer. J. Math. **90**, 723–732 (1968) Zbl 0177.51204 MR 232377

[39] Veech, W. A.: Point-distal flows. Amer. J. Math. **92**, 205–242 (1970) Zbl 0202.55503 MR 267560

[40] Veech, W. A.: Topological dynamics. Bull. Amer. Math. Soc. **83**, 775–830 (1977) Zbl 0384.28018 MR 467705

[41] Vinogradov, I. M.: The method of trigonometrical sums in the theory of numbers. Interscience Publishers, London (1947) Zbl 0041.37002 MR 0062183

[42] Wu Q., Xu, H., Ye, X.: On structure theorems and non-saturated examples. Comm. Math. Statist. (to appear); arXiv:2201.00152

[43] Ziegler, T.: Universal characteristic factors and Furstenberg averages. J. Amer. Math. Soc. **20**, 53–97 (2007) Zbl 1198.37014 MR 2257397