

# INTERPLAY BETWEEN FINITE TOPOLOGICAL RANK MINIMAL CANTOR SYSTEMS, $\mathcal{S}$ -ADIC SUBSHIFTS AND THEIR COMPLEXITY

SEBASTIÁN DONOSO, FABIEN DURAND, ALEJANDRO MAASS, AND SAMUEL PETITE

**ABSTRACT.** Minimal Cantor systems of finite topological rank (that can be represented by a Bratteli-Vershik diagram with a uniformly bounded number of vertices per level) are known to have dynamical rigidity properties. We establish that such systems, when they are expansive, define the same class of systems, up to topological conjugacy, as primitive and recognizable  $\mathcal{S}$ -adic subshifts. This is done establishing necessary and sufficient conditions for a minimal subshift to be of finite topological rank. As an application, we show that minimal subshifts with non-superlinear complexity (like all classical zero entropy examples) have finite topological rank. Conversely, we analyze the complexity of  $\mathcal{S}$ -adic subshifts and provide sufficient conditions for a finite topological rank subshift to have a non-superlinear complexity. This includes minimal Cantor systems given by Bratteli-Vershik representations whose tower levels have proportional heights and the so called left to right  $\mathcal{S}$ -adic subshifts. We also exhibit that finite topological rank does not imply non-superlinear complexity. In the particular case of topological rank 2 subshifts, we prove their complexity is always subquadratic along a subsequence and their automorphism group is trivial.

## 1. INTRODUCTION

Using tools originally appearing in the theory of operator algebras, Herman *et al.* [HPS92] introduced a description of any minimal Cantor system by means of a Bratteli diagram and a Vershik map acting on it. This method proved to be powerful to solve ergodic, dynamical and operator algebras problems, like the computation of a complete invariant for topological orbit equivalence [GPS95] which is one of the most remarkable achievements.

A way to describe the simplest minimal Cantor systems is provided by the notion of finite topological rank. Analogously to the eponymous notion for measurable dynamics [ORW82], these minimal Cantor systems can be represented by Kakutani-Rohlin partitions defining a Bratteli-Vershik diagram with a uniformly bounded number of vertices per level. All the classical examples of zero-entropy minimal Cantor systems (odometers, substitution subshifts, linearly recurrent subshifts or symbolic codings of interval exchange maps) have a finite topological rank. The

---

*Date:* August 21, 2020.

*2010 Mathematics Subject Classification.* Primary: 54H20; Secondary: 37B10.

*Key words and phrases.*  $\mathcal{S}$ -adic subshifts, minimal Cantor systems, finite topological rank, recognizability, complexity.

This research was partially supported by grants PIA-ANID AFB 170001 & Fondap 15090007, and grant ANID/MEC/80180045 hosted by University of O'Higgins. The first and the third author thank the hospitality of the LAMFA UMR 7352 CNRS-UPJV and the "poste rouge" CNRS program.

study of finite topological rank systems from decades exhibits rigidity properties. Let us mention some of them. There is a dichotomy: these systems are either expansive or equicontinuous [DM08]. Hence, non-equicontinuous finite topological rank systems are conjugate to subshifts and it is a folklore result that they have zero entropy. The topological rank is an upper bound for the number of ergodic invariant probability measures (see [BKMS13]), for the rational rank of the dimension group [GPS95, GHH18] and for the rational rank of the continuous spectrum of the system [BDM10]. To understand the (typical) properties of these systems, it is natural to try to characterize their dynamical properties through their combinatorial structure given by the Bratteli-Vershik diagrams. Their combinatorial structures have been exploited in the characterization of Toeplitz subshifts [GJ00], linearly recurrent subshifts [CDHM03] or unique ergodicity [BKMS13]. It has been valuable too in the study of the continuous or measurable spectrum of minimal Cantor systems (see [BDM10, DFM15, DFM19]) and to understand the incidence of the orbit equivalence on the spectral properties (see [GPS95, CDP16, GHH18, BCB<sup>+</sup>]).

Another effective way to represent a minimal subshift is through an  $\mathcal{S}$ -adic subshift, that is a subshift generated by the iteration of a succession of morphisms (substitutions) on eventually different alphabets, see Section 2.2.3 for definitions. When the alphabets are bounded in size, the  $\mathcal{S}$ -adic subshift is said to have bounded alphabet rank. The constructive description of such subshifts leads to consider the same problem: the description of its dynamical properties through the combinatorial data provided by the morphisms.

Few classes of subshifts can be characterized in combinatorial terms, as substitution subshifts [Dur98] and more generally linearly recurrent subshifts [Dur00]. As a main issue in this paper, we want to state sufficient and/or necessary conditions for a minimal subshift to be of finite topological rank or conjugate to a bounded alphabet rank  $\mathcal{S}$ -adic subshift.

One of our principal results states, under natural conditions, a topological equivalence between minimal  $\mathcal{S}$ -adic subshifts with bounded alphabet rank and finite topological rank minimal Cantor system (Theorem 4.1). Positive evidence of this fact was shown in [BSTY19], where this result was proved in the measure theoretical category. To show it, we provide a necessary and sufficient condition for a minimal subshift to have a finite topological rank (Theorem 4.4). The essential key notion for the proofs is the recognizability of morphisms together with the remarkable result in [BSTY19] for  $\mathcal{S}$ -adic subshifts extending the seminal recognizability results for substitutions by B. Mossé in [Mos92] and the related ones in [BKM08, DM08].

As another application of this criterion, we get that any minimal subshift with non-superlinear word complexity function has finite topological rank (Theorem 5.5). Hence it is conjugate to an  $\mathcal{S}$ -adic subshift with bounded alphabet rank and has all the aforementioned dynamical restrictions. This also recovers, in the minimal setting, sufficient conditions of [Bos85, CK19, FM10] on the complexity to ensure a finite number of ergodic invariant probability measures. Analogous results were obtained in relation to the  $\mathcal{S}$ -adic conjecture, in the measurable setting [Fer96] and in [DL12] for minimal subshifts whose complexity difference is at most two. This also recovers one-dimensional partial results of [Jul10] linking the complexity and the rank of rational cohomology groups of tiling spaces. Along the way, and following an idea due to T. Monteil, we provide a characterization of Rauzy graphs associated to minimal subshifts of non-superlinear complexity (Proposition 5.3).

We observe that Theorem 5.5 mentioned above does not admit a converse since we explicitly present a minimal subshift of topological rank two with a superlinear complexity in Section 6.4. We go deeper into the relationship between the finite topological rank property and the complexity of words by studying the complexity function within the class of  $\mathcal{S}$ -adic subshifts. Upper bounds for it are given in terms of a notion of relative complexity of morphisms in Section 6. We deduce several applications, among them, conditions for a finite topological rank subshift to have a non-superlinear complexity. Concerning the complexity of a topological rank two minimal subshift, or of a minimal  $\mathcal{S}$ -adic subshift on 2 letters alphabets, we show it is always sub-quadratic along a subsequence (Corollary 6.6). Concerning the orbit equivalence theory, we obtain that any finite topological rank minimal Cantor system is strongly orbit equivalent to a subshift of sub-linear complexity (Corollary 6.14). We also exhibit, thanks to recognizability results, a rigid dynamical property of topological rank two systems: they have only one asymptotic component (Corollary 7.4). It follows, by using arguments developed in [DDMP16], that their automorphism groups are always trivial, *i.e.* generated by powers of the shift map.

**1.1. Organization.** In the next section we give the basic background in topological and symbolic dynamics needed in this article. Section 3 is devoted to review and study recognizability properties of a sequence of morphisms. The equivalence between finite topological rank and  $\mathcal{S}$ -adic subshifts is established in Section 4 and in Section 5 we show that non-superlinear subshifts have a finite topological rank. We investigate the complexity of  $\mathcal{S}$ -adic subshifts in Section 6 and provide conditions to have a non-superlinear/linear complexity. In the last but one section we analyze topological properties of topological rank two subshifts, focusing on the number of asymptotic components. The definitions needed are included at the beginning of this section, so that the interested reader can read it independently of the others. We conclude with some remarks and questions in the last section.

In this article, we let  $\mathbb{N}, \mathbb{N}^*$  and  $\mathbb{Z}$  denote the set of non-negative integers, the set of positive integers and the set of integer numbers respectively.

## 2. BASICS IN TOPOLOGICAL AND SYMBOLIC DYNAMICS

**2.1. Basics in topological dynamics.** A *topological dynamical system* (or just a system) is a pair  $(X, T)$  where  $X$  is a compact metric space and  $T: X \rightarrow X$  is a homeomorphism. We denote by  $\text{Orb}_T(x)$  the orbit  $\{T^n x : n \in \mathbb{Z}\}$  of  $x \in X$ . A point  $x \in X$  is *periodic* if  $\text{Orb}_T(x)$  is a finite set and *aperiodic* otherwise. A topological dynamical system is *aperiodic* if any point  $x \in X$  is aperiodic and is *minimal* if the orbit of every point is dense in  $X$ .

A *factor map* between the topological dynamical systems  $(X, T)$  and  $(Y, S)$  is a continuous onto map  $\pi: X \rightarrow Y$  such that  $\pi \circ T = S \circ \pi$ . We say that  $(Y, S)$  is a *factor* of  $(X, T)$  and that  $(X, T)$  is an *extension* of  $(Y, S)$ . We use the notation  $\pi: (X, T) \rightarrow (Y, S)$  to indicate the factor map. If in addition  $\pi$  is a bijective map we say that  $(X, T)$  and  $(Y, S)$  are *topologically conjugate* and we call  $\pi$  a topological conjugacy.

**2.2. Basics in symbolic dynamics.**

**2.2.1. Subshifts and their languages.** Let  $\mathcal{A}$  be a finite set that we call *alphabet*. Elements in  $\mathcal{A}$  are called *letters* or *symbols*. The set of finite sequences or *words* of length  $\ell \in \mathbb{N}$  with letters in  $\mathcal{A}$  is denoted by  $\mathcal{A}^\ell$ , the set of onesided sequences  $(x_n)_{n \in \mathbb{N}}$  in  $\mathcal{A}$  is denoted by  $\mathcal{A}^{\mathbb{N}}$  and the set of two-sided sequences  $(x_n)_{n \in \mathbb{Z}}$  in  $\mathcal{A}$  is denoted by  $\mathcal{A}^{\mathbb{Z}}$ . Also, a word  $w = w_1 \cdots w_\ell \in \mathcal{A}^\ell$  can be seen as an element of the free monoid  $\mathcal{A}^*$  endowed with the operation of concatenation. The integer  $\ell$  is the *length* of  $w$  and is denoted by  $|w| = \ell$ . For two finite words  $u$  and  $v$ , we say they are respectively a *prefix* and a *suffix* of the word  $uv$ .

The *shift map*  $S: \mathcal{A}^{\mathbb{Z}} \rightarrow \mathcal{A}^{\mathbb{Z}}$  is defined by  $S((x_n)_{n \in \mathbb{Z}}) = (x_{n+1})_{n \in \mathbb{Z}}$ . To simplify notations we denote the shift map by  $S$ , independent of the alphabet, which will be clear from the context. A *subshift* is a topological dynamical system  $(X, S)$  where  $X$  is a closed and  $S$ -invariant subset of  $\mathcal{A}^{\mathbb{Z}}$  (we consider the product topology in  $\mathcal{A}^{\mathbb{Z}}$ ). Classically one identifies  $(X, S)$  with  $X$ , so one says that  $X$  itself is a subshift. When we say that a sequence  $x$  in a subshift is aperiodic, we implicitly mean that  $x$  is aperiodic for the action of the shift. For convenience, when we state general results about topological dynamical systems we use the notation  $(X, T)$  and to state specific results about subshifts we use  $(X, S)$ .

Let  $(X, S)$  be a subshift. The *language* of  $(X, S)$  is the set  $\mathcal{L}(X)$  containing all words  $w \in \mathcal{A}^*$  such that  $w = x_m \cdots x_{m+|w|-1} = x_{[m, m+|w|]}$  for some  $x = (x_n)_{n \in \mathbb{Z}} \in X$  and  $m \in \mathbb{Z}$ . In this case we say that  $w$  *appears* or *occurs* in the sequence  $x$  and that the index  $m$  is an *occurrence* of  $w$  in  $x$ . We also say that  $w$  *starts* (resp. *ends*) in a word  $u$  whenever  $u = x_{[m-i, m+j]}$  (resp.  $u = x_{[m+|w|-1-i, m+|w|-1+j]}$ ) for some indices  $i \geq 0, j > 0$ . We use the same name for the similar notions for finite words  $x$ . We denote by  $\mathcal{L}_\ell(X)$  the set of words of length  $\ell$  in  $\mathcal{L}(X)$  and by  $\mathcal{L}_{\geq \ell}(X)$  the set of words of length greater than or equal to  $\ell$  in  $\mathcal{L}(X)$ . Given  $x \in X$  the language  $\mathcal{L}(x)$  is the set of all words appearing in  $x$ . As before we define  $\mathcal{L}_\ell(x)$  and  $\mathcal{L}_{\geq \ell}(x)$ . For two words  $u, v \in \mathcal{L}(X)$ , the *cylinder set*  $[u.v]$  is the set  $\{x \in X : x_{[-|u|, |v|]} = uv\}$ . When  $u$  is the empty word we only write  $[v]$ , erasing the dot. Cylinder sets are *clopen sets* (closed and open sets). They form a base for the topology of the subshift.

A word  $w \in \mathcal{L}(X)$  is said to be *right special* if there exist at least two distinct letters  $a$  and  $b$  such that  $wa$  and  $wb$  belong to  $\mathcal{L}(X)$ . In the same way we define *left special* words. We denote by  $\mathcal{LS}(X)$  and  $\mathcal{RS}(X)$  the set of left and right special words in  $X$  respectively. It is well known that for any infinite subshift  $X$  admits left and right special words of any length.

The non decreasing map  $p_X: \mathbb{N} \rightarrow \mathbb{N}$  defined by  $p_X(n) = |\mathcal{L}_n(X)|$  is called the *complexity function* of  $X$ . If  $\limsup_{n \rightarrow +\infty} p_X(n)/n < +\infty$  we say that  $X$  has *sublinear* complexity, which is equivalent to  $p_X(n) \leq cn$  for all  $n \in \mathbb{N}$  for some positive constant  $c$ . We say that  $X$  has *non-superlinear* complexity if  $\liminf_{n \rightarrow +\infty} p_X(n)/n < +\infty$ . In the contrary, we say  $X$  has *superlinear* complexity.

**2.2.2. Morphisms and substitutions.** For an alphabet  $\mathcal{A}$ , let  $\mathcal{A}^*$  denote the free monoid generated by concatenations of letters in  $\mathcal{A}$ . The neutral element is called the *empty word*. Let  $\mathcal{A}$  and  $\mathcal{B}$  be finite alphabets and  $\tau: \mathcal{A}^* \rightarrow \mathcal{B}^*$  be a morphism. We say that  $\tau$  is *erasing* whenever there exists  $a \in \mathcal{A}$  such that  $\tau(a)$  is the empty word. Otherwise we say it is *non-erasing*. The morphism  $\tau$  is *proper* if each word  $\tau(a)$  starts and ends with the same letter independently of  $a$ . When it is non-erasing it extends naturally to maps from  $\mathcal{A}^{\mathbb{N}}$  to itself and from  $\mathcal{A}^{\mathbb{Z}}$  to itself in the obvious way by concatenation (in the case of a twosided sequence we apply  $\tau$  to

positive and negative coordinates separately and we concatenate the results at the coordinate zero). To the morphism  $\tau$  we associate an *incidence matrix*  $M_\tau$  indexed by  $\mathcal{A} \times \mathcal{B}$  such that its entry at position  $(a, b)$  is the number of occurrences of  $b$  in  $\tau(a)$  for every  $a \in \mathcal{A}$  and  $b \in \mathcal{B}$ . When all the entries of  $M_\tau$  are positive, we say that  $\tau$  is *positive*. Note that any function  $\tau: \mathcal{A} \rightarrow \mathcal{B}^*$  can be naturally extended to a morphism (that we also denote  $\tau$ ) from  $\mathcal{A}^*$  to  $\mathcal{B}^*$  by concatenation, and we use this convention throughout the document. So from now on, all functions between finite alphabets are considered to be morphisms between their free monoids.

**2.2.3.  $\mathcal{S}$ -adic subshifts.** We recall the definition of an  $\mathcal{S}$ -adic subshift as stated in [BSTY19]. A *directive sequence*  $\tau = (\tau_n: \mathcal{A}_{n+1}^* \rightarrow \mathcal{A}_n^*)_{n \geq 0}$  is a sequence of morphisms. From now on, we only consider morphisms that are non-erasing. When all morphisms  $\tau_n$ ,  $n \geq 1$ , are proper we say that  $\tau$  is *proper* and when all incidence matrices of  $\tau_n$ ,  $n \geq 1$ , are positive, we say that  $\tau$  is *positive*. Here we stress the fact that in the definition there is no assumption (more than non-erasingness) on the first morphism  $\tau_0$ , since in many cases of interest this morphism is not proper nor positive.

For  $0 \leq n < N$ , we denote by  $\tau_{[n, N]}$  or  $\tau_{[n, N-1]}$ , the morphism  $\tau_n \circ \tau_{n+1} \circ \dots \circ \tau_{N-1}$ . We say  $\tau$  is *primitive* if for any  $n \in \mathbb{N}$  there exists  $N > n$  such that  $M_{\tau_{[n, N]}} > 0$ , i.e.,  $\tau_{[n, N]}(a)$  contains all letters in  $\mathcal{A}_N$  for all  $a \in \mathcal{A}_n$ .

For  $n \in \mathbb{N}$ , the *language*  $\mathcal{L}^{(n)}(\tau)$  of level  $n$  associated with  $\tau$  is defined by

$$\mathcal{L}^{(n)}(\tau) = \{w \in \mathcal{A}_n^* : w \text{ occurs in } \tau_{[n, N]}(a) \text{ for some } a \in \mathcal{A}_N \text{ and } N > n\}$$

and  $X_\tau^{(n)}$  is the set of points  $x \in \mathcal{A}_n^{\mathbb{Z}}$  such that  $\mathcal{L}(x) \subseteq \mathcal{L}^{(n)}(\tau)$ . This set clearly defines a subshift that we call the *subshift generated by*  $\mathcal{L}^{(n)}(\tau)$ . It may happen that  $\mathcal{L}(X_\tau^{(n)})$  is strictly contained in  $\mathcal{L}^{(n)}(\tau)$ , but in the primitive case both sets coincide and  $X_\tau^{(n)}$  is a minimal subshift. By definition we have  $\tau_n(\mathcal{L}(X_\tau^{(n+1)})) = \mathcal{L}(X_\tau^{(n)})$ . We set  $X_\tau = X_\tau^{(0)}$  and call  $(X_\tau, S)$  the  $\mathcal{S}$ -adic subshift generated by the directive sequence  $\tau$ .

We define the *alphabet rank* of  $\tau$  as

$$AR(\tau) = \liminf_{n \rightarrow +\infty} |\mathcal{A}_n|.$$

When  $AR(\tau)$  is finite this means  $(X_\tau, S)$  can be defined by a directive sequence where the morphisms are endomorphisms on a free monoid with  $AR(\tau)$  generators.

We define the *alphabet rank* of a subshift  $X$  (or  $(X, S)$ ) to be

$$AR(X) = \inf_{\tau, X_\tau = X} AR(\tau),$$

where the infimum runs over all sequences of morphisms  $\tau$  such that  $X_\tau = X$ . As a convention the infimum is equal to  $+\infty$  when this set is empty.

**2.3. Bratteli-Vershik systems.** Let  $(X, T)$  be a minimal Cantor system, i.e.,  $X$  is a Cantor set (has a countable basis of closed and open sets and it has no isolated points). It can be represented by an ordered Bratteli diagram together with the Vershik transformation acting on it. We give a brief outline of this construction, for details on this theory see [HPS92] or [Dur10].

**2.3.1. Bratteli diagrams.** A Bratteli diagram is an infinite graph which consists of a vertex set  $V$  and an edge set  $E$ , both of which are divided into levels  $V = V_0 \cup V_1 \cup \dots$  and  $E = E_1 \cup E_2 \cup \dots$ , where all levels are pairwise disjoint. The set  $V_0$  is a singleton  $\{v_0\}$  and for all  $n \geq 1$  edges in  $E_n$  join vertices in  $V_{n-1}$  to vertices in  $V_n$ . If  $e \in E$  connects  $u \in V_{n-1}$  with  $v \in V_n$  we write  $s(e) = u$  and  $r(e) = v$ , where  $s: E_n \rightarrow V_{n-1}$  and  $r: E_n \rightarrow V_n$  are the source and range maps, respectively. It is also required that  $s^{-1}(v) \neq \emptyset$  for all  $u \in V$  and that  $r^{-1}(v) \neq \emptyset$  for all  $v \in V \setminus V_0$ . For all  $n \geq 1$  we set  $|V_n| = d_n$  and we write  $V_n = \{1, \dots, d_n\}$  to simplify notation.

Fix  $n \geq 1$ . We call *level  $n$*  of the diagram the subgraph consisting of the vertices in  $V_{n-1} \cup V_n$  and the edges  $E_n$  between these vertices. Level 1 is called the *hat* of the Bratteli diagram. We describe the edge set  $E_n$  using a  $V_n \times V_{n-1}$  incidence matrix  $M_n$  for which its  $(v, u)$  entry is the number of edges in  $E_n$  joining vertex  $u \in V_{n-1}$  with vertex  $v \in V_n$ . We also consider several levels at the same time. For integers  $0 \leq n < N$  we denote by  $E_{n,N}$  the set of all paths in the graph joining vertices of  $V_n$  with vertices of  $V_N$ .

A Bratteli diagram  $(V, E)$  is *simple* or *primitive* if for any  $n \geq 1$  there exists  $N > n$  such that each pair of vertices  $u \in V_n$  and  $v \in V_N$  is connected by a finite path.

**2.3.2. Ordered Bratteli diagrams and Bratteli-Vershik representations.** An *ordered* Bratteli diagram is a triple  $B = (V, E, \preceq)$ , where  $(V, E)$  is a Bratteli diagram and  $\preceq$  is a partial ordering on  $E$  such that: edges  $e$  and  $e'$  in  $E$  are comparable if and only if  $r(e) = r(e')$ . This partial ordering naturally defines maximal and minimal edges. The partial ordering of  $E$  induces another one on paths of  $E_{n,N}$  for all  $0 \leq n < N$ :  $(e_{n+1}, \dots, e_N) \preceq (f_{n+1}, \dots, f_N)$  if and only if there is  $n+1 \leq i \leq N$  such that  $e_i \preceq f_i$  and  $e_j = f_j$  for  $i < j \leq N$ .

Given a strictly increasing sequence of integers  $(n_k)_{k \geq 0}$  with  $n_0 = 0$  one defines the *contraction* or *telescoping* of  $B = (V, E, \preceq)$  with respect to  $(n_k)_{k \geq 0}$  by

$$\left( (V_{n_k})_{k \geq 0}, (E_{n_k, n_{k+1}})_{k \geq 0}, \preceq \right),$$

where  $\preceq$  is the order induced in each set of edges  $E_{n_k, n_{k+1}}$ . The converse operation is called *microscoping* (see [HPS92] and [GPS95] for more details).

Given an ordered Bratteli diagram  $B = (V, E, \preceq)$  one defines  $X_B$  as the set of infinite paths  $(x_1, x_2, \dots)$  starting in  $v_0$  such that  $r(x_n) = s(x_{n+1})$  for all  $n \geq 1$ . We topologize  $X_B$  by postulating a basis of open sets, namely the family of *cylinder sets*

$$[e_1, e_2, \dots, e_n] = \{(x_1, x_2, \dots) \in X_B : x_i = e_i, \text{ for } 1 \leq i \leq n\}.$$

Each  $[e_1, e_2, \dots, e_n]$  is also closed, as is easily seen, and so  $X_B$  is a compact, totally disconnected metrizable space. If  $(V, E)$  is simple and  $X_B$  infinite then  $X_B$  is Cantor.

When there is a unique point  $(x_1, x_2, \dots) \in X_B$  such that  $x_n$  is (locally) maximal for any  $n \geq 1$  and a unique point  $(y_1, y_2, \dots) \in X_B$  such that  $y_n$  is (locally) minimal for any  $n \geq 1$ , one says that  $B = (V, E, \preceq)$  is a *properly ordered* Bratteli diagram. We call these particular points  $x_{\max}$  and  $x_{\min}$  respectively. In this case, we define the map  $V_B$  on  $X_B$  called the *Vershik map* as follows. Let  $x = (x_1, x_2, \dots) \in X_B \setminus \{x_{\max}\}$  and let  $n \geq 1$  be the smallest integer so that  $x_n$  is not a maximal edge. Let  $y_n$  be the successor of  $x_n$  for the corresponding local order and  $(y_1, \dots, y_{n-1})$

be the unique minimal path in  $E_{0,n-1}$  connecting  $v_0$  with the initial vertex of  $y_n$ . We set  $V_B(x) = (y_1, \dots, y_{n-1}, y_n, x_{n+1}, \dots)$  and  $V_B(x_{\max}) = x_{\min}$ .

The couple  $(X_B, V_B)$  is a topological dynamical system called the *Bratteli-Vershik system* generated by  $B = (V, E, \preceq)$ . The dynamical system induced by any telescoping of  $B$  is topologically conjugate to  $(X_B, V_B)$ .

In [HPS92] it is proved that the system  $(X_B, V_B)$  is minimal whenever the associated Bratteli diagram  $(V, E)$  is simple. Conversely, it is also proved that any minimal Cantor system  $(X, T)$  is topologically conjugate to a Bratteli-Vershik system  $(X_B, V_B)$  where  $(V, E)$  is simple. We say that a simple and properly ordered Bratteli diagram  $B = (V, E, \preceq)$  is a *Bratteli-Vershik representation* of the minimal Cantor system  $(X, T)$  if  $(X, T)$  and  $(X_B, V_B)$  are topologically conjugate. We also say that  $(X, T)$  is given by the Bratteli-Vershik representation  $B$ .

Let  $B = (V, E, \preceq)$  be an ordered Bratteli diagram. We associate different subshifts to the dynamics of  $(X_B, V_B)$ . First, for each  $n \geq 1$  set

$$X_B^{(n)} = \{(\pi_n(V_B^k x))_{k \in \mathbb{Z}} : x = (x_i)_{i \geq 1} \in X_B\} \subseteq V_n^{\mathbb{Z}},$$

where  $\pi_n : X_B \rightarrow E_n$  is the projection map given by  $\pi_n(x) = x_n$ . We call this subshift the *coding of the orbits of  $(X_B, V_B)$  at level  $n$* . The subshift  $(X_B^{(n)}, S)$  is minimal whenever the system  $(X_B, V_B)$  is. Now, let  $X_B^{(0)} = \{(\pi_1(V_B^k x))_{k \in \mathbb{Z}} : x \in X_B\} \subseteq E_1^{\mathbb{Z}}$ . Denote by  $\phi_0 : X_B \rightarrow X_B^{(0)}$  the natural factor map between  $(X_B, V_B)$  and  $(X_B^{(0)}, S)$ . It is clear that the subshifts  $X_B^{(0)}$  and  $X_B^{(1)}$  coincide, up to the alphabets, whenever there is a unique edge between the vertex in  $V_0$  and each vertex in  $V_1$ .

Define  $\tau_0^B : V_1^* \rightarrow E_1^*$  by  $\tau_0^B(v) = e_1(v) \dots e_\ell(v)$ , where the edges  $e_1(v), \dots, e_\ell(v)$  correspond to those connecting  $V_0$  with  $v$  and are ordered following the partial order of  $B$ . In a similar way, for  $n \geq 1$  consider the morphisms  $\tau_n^B : V_{n+1}^* \rightarrow V_n^*$  defined for a vertex  $v \in V_{n+1}$  by  $\tau_n^B(v) = s(e_1(v)) \dots s(e_\ell(v))$ , where edges  $e_1(v), \dots, e_\ell(v)$  are all the edges in  $E_{n+1}$  whose range is  $v$  ordered as indicated by the partial order of  $B$ . We say  $\tau_n^B$  is the morphism *read at level  $n+1$  on  $B$*  and that  $\boldsymbol{\tau}^B = (\tau_n^B)_{n \geq 0}$  is the *sequence of morphisms read on  $B$* . Following these definitions, an ordered Bratteli diagram  $B$  defines a unique sequence of morphisms  $(\tau_n^B : \mathcal{A}_{n+1}^* \rightarrow \mathcal{A}_n^*)_{n \geq 0}$  with  $\mathcal{A}_0 = E_1$  and  $\mathcal{A}_n = V_n$  for  $n \geq 1$ . In the sequel it will be useful to observe that for the first morphism  $\tau_0^B$  all the letters occurring in  $\tau_0^B(v)$ ,  $v \in V_1$ , are pairwise distinct.

The sequence  $\boldsymbol{\tau}^B$  is primitive whenever the Bratteli diagram is simple. Observe moreover that  $\tau_n^B(X_B^{(n+1)})$  is included in  $X_B^{(n)}$ , and  $\cup_{k \in \mathbb{Z}} S^k \tau_n^B(X_B^{(n+1)}) = X_B^{(n)}$ . Hence  $X_B^{(0)}$  is the  $\mathcal{S}$ -adic subshift generated by  $\boldsymbol{\tau}^B$ .

It is standard to check that the unicity of the maximal and minimal paths in a properly ordered Bratteli diagram does imply, up to telescoping, that the morphisms  $\tau_n^B$  are proper for all  $n \geq 1$ .

**2.3.3. Kakutani-Rohlin partitions and Bratteli-Vershik representations.** The construction of a Bratteli-Vershik representation for a minimal Cantor system  $(X, T)$  relies on the notion of *clopen Kakutani-Rohlin partition (CKR partition)*. A CKR partition  $\mathcal{T}$  of  $X$  is given by:

$$(1) \quad \mathcal{T} = \{T^{-j}B(k) : 1 \leq k \leq d, 0 \leq j < h(k)\},$$

where  $d, h(1), \dots, h(d)$  are positive integers and  $B(1), \dots, B(d)$  are clopen subsets of  $X$  such that

$$\bigcup_{k=1}^d T^{-h(k)} B(k) = \bigcup_{k=1}^d B(k).$$

The *base* of  $\mathcal{T}$  is the set  $B(\mathcal{T}) = \bigcup_{1 \leq k \leq d} B(k)$  and the set  $\mathcal{T}(k) = \{T^{-j} B(k) : 0 \leq j < h(k)\}$  is the  $k$ -th *tower* of  $\mathcal{T}$  and  $h(k)$  its *height*.

To construct a Bratteli-Vershik representation of  $(X, T)$  we need a sequence of *nested* CKR partitions. For each  $n \geq 0$  let

$$(2) \quad \mathcal{T}_n = \{T^{-j} B_n(k) : 1 \leq k \leq d_n, 0 \leq j < h_n(k)\}$$

such that  $\mathcal{T}_0$  is the trivial partition, (i.e.,  $d_0 = 1$ ,  $h_0(1) = 1$  and  $B_0(1) = X$ ). We say that the sequence of CKR partitions  $(\mathcal{T}_n)_{n \geq 0}$  is *nested* if for any  $n \geq 0$ :

- (KR1)  $B(\mathcal{T}_{n+1}) \subseteq B(\mathcal{T}_n)$ ;
- (KR2)  $\mathcal{T}_{n+1} \succeq \mathcal{T}_n$ , i.e., for any atom  $A \in \mathcal{T}_{n+1}$  there exists an atom  $A' \in \mathcal{T}_n$  such that  $A \subseteq A'$ ;
- (KR3)  $\bigcap_{n \geq 0} B(\mathcal{T}_n)$  consists of a unique point;
- (KR4) the sequence of partitions spans the topology of  $X$ .

These conditions imply in particular that for all  $n \geq 1$  and  $1 \leq k \leq d_{n+1}$  all points in  $T^{-(h_{n+1}(k)-1)} B_{n+1}(k)$  (the *roof* of the  $k$ -th tower) iterated by  $T$  cross completely and in the same order several towers of level  $n$  until reaching  $B_{n+1}(k)$ . Recursively one has that for all  $1 \leq n < N$  and  $1 \leq k \leq d_N$  all points in  $T^{-(h_N(k)-1)} B_N(k)$  iterated by  $T$  cross completely, and in the same order, several towers of level  $n$  until reaching the base  $B_N(k)$ . Not necessarily all towers of level  $n$  are visited. We say that the nested sequence of CKR partitions  $(\mathcal{T}_n)_{n \geq 0}$  is *primitive* if for all  $n \geq 1$  there exists  $N > n$  such that for all  $1 \leq k \leq d_N$  all points in  $T^{-(h_N(k)-1)} B_N(k)$  iterated by  $T$  cross completely and in the same order all towers of level  $n$  until reaching the base  $B_N(k)$ .

Given a primitive sequence of nested CKR partitions of  $(X, T)$  it is possible to construct a Bratteli-Vershik representation of  $(X, T)$  and vice versa. We will briefly describe the construction of the representation and the converse will be clear from this construction too. First, define the associated ordered Bratteli diagram (we follow notations of Section 2.3.1). Fix  $n \geq 0$ . The vertices  $V_n$  will represent the towers in  $\mathcal{T}_n$  so we set  $V_n = \{1, \dots, d_n\}$ , in particular  $V_0 = \{1\}$ . In the set of edges  $E_{n+1}$  we put an arrow from vertex  $u \in V_n$  to vertex  $v \in V_{n+1}$  each time the tower  $u$  of level  $n$  is crossed by points in tower  $v$  of level  $n+1$  as discussed before. The local order at the vertex  $v \in V_{n+1}$  is given by the order points in  $B_{n+1}(v)$  cross the towers of level  $n$  when iterated by  $T^{-1}$ . In this way we have defined an ordered Bratteli diagram. By construction and primitivity assumption the constructed diagram is simple. The fact that it is properly ordered follows from (KR3). Indeed, the unique maximal point  $x_{\max}$  is given by (KR3) and the unique minimal point  $x_{\min}$  is its image by  $T$ . The conjugacy between  $(X_B, V_B)$  and  $(X, T)$  is described as follows. A point in  $x \in X$  is determined uniquely by the intersection  $\{x\} = \bigcap_{n \geq 0} T^{-j_{n,x}} B_n(u_{n,x})$ , where  $1 \leq u_{n,x} \leq d_n$ ,  $0 \leq j_{n,x} < h_n(u_{n,x})$ . Then for each  $n \geq 0$  we select the edge  $e_{n+1}$  with extremities  $u_{n,x} \in V_n$  and  $u_{n+1,x} \in V_{n+1}$  whose local order corresponds to the apparition of the  $u_{n,x}$ -th tower of level  $n$  containing  $T^{-j_{n+1,x}} B_{n+1}(u_{n+1,x})$ . The map  $x \rightarrow (e_1, e_2, \dots)$  is a conjugacy. Thus we have defined a Bratteli-Vershik representation for  $(X, T)$ .



**2.3.4. Minimal Cantor systems of finite topological rank.** A minimal Cantor system is of finite topological rank if it admits a Bratteli-Vershik representation such that the number of vertices per level is uniformly bounded by some integer  $d$ . The minimum possible value of  $d$  is called the *topological rank* of the system.

Equivalently (due to previous section), a minimal Cantor system is of finite topological rank if it admits a primitive sequence of nested CKR partitions such that the number of towers per level is uniformly bounded by some  $d \geq 1$ . When  $d = 1$  the system is called an *odometer*.

A fundamental result concerning finite topological rank systems states the following dichotomy.

**Theorem 2.1.** [DM08] *A finite topological rank minimal Cantor system is either an odometer or expansive (so a subshift).*

It is important to notice that in the expansive case the Cantor system is topologically conjugate to one of the level subshifts  $(X_B^{(n)}, S)$  associated to a Bratteli-Vershik representation  $B$  of finite topological rank.

For other proofs of this result and extensions to non minimal contexts see [BKM08, Høy17, BSTY19, Shi17].

### 3. RECOGNIZABILITY FOR MORPHISMS AND LANGUAGES

The notion of recognizability in symbolic dynamics is mainly known because of the result by B. Mossé stating that all primitive substitutions or the corresponding subshifts are recognizable [Mos92]. A possible way to state this property is that points in a substitutive subshift can be uniquely factorized using words of a given finite set or the same for families of finite sets of words. In the case of substitutions those families correspond to the words produced by iterating letters. This notion seen in this way can be naturally extended to many other systems, in particular to  $\mathcal{S}$ -adic subshifts [BSTY19] or subshifts defined from a Bratteli-Vershik system [DM08]. It is worth mentioning that Theorem 2.1 can be seen as a recognizability result. Now, we make precise the notions of factorization and recognizability we will need in this article. For definitions of recognizability we mainly follow [BSTY19].

**3.1. Factorization of sequences.** Let  $\mathcal{A}$  be a finite alphabet and  $\mathcal{W}$  be a finite set of non-empty words in  $\mathcal{A}^*$ . A  $\mathcal{W}$ -factorization of  $x \in \mathcal{A}^{\mathbb{Z}}$  is a map  $F_x: \mathbb{Z} \rightarrow \mathcal{W} \times \mathbb{Z}$  such that:

- for each  $k \in \mathbb{Z}$ , if  $F_x(k) = (w, i)$  and  $F_x(k+1) = (w', j)$  then  $x_{[i,j]} = w$ ;
- if  $F_x(0) = (w, i)$  and  $F_x(1) = (w', j)$  then  $i \leq 0 < j$ .

For  $F_x(k) = (w, i)$  we say  $i \in \mathbb{Z}$  is a *starting position* of  $w$  in  $x$  for  $F_x$ . The starting positions are also called *cuts* (see [BSTY19, DM08]). In other words,  $x$  admits a  $\mathcal{W}$ -factorization if there exists a strictly increasing sequence of integers  $(n_k)_{k \in \mathbb{Z}}$  such that:  $n_0 \leq 0 < n_1$  and  $x_{[n_k, n_{k+1})} \in \mathcal{W}$  for all  $k \in \mathbb{Z}$ . Of course, this sequence corresponds to the sequence of starting positions of the factorization:  $F_x(k) = (x_{[n_k, n_{k+1})}, n_k)$  for all  $k \in \mathbb{Z}$ . Two factorizations  $F_x, F'_x$  of  $x$  are *disjoint* whenever the starting positions of  $F_x$  and  $F'_x$  are distinct, so they do not have common cuts.

We recall a theorem that is very useful when studying factorization of finite or infinite words. It is a consequence of the so called “defect theorems” in language theory. We will use it to improve some results concerning recognizability of  $\mathcal{S}$ -adic systems in the next section.

**Theorem 3.1** ([KM02, Corollary 1]). *Let  $\mathcal{W} \subset \mathcal{A}^*$  be a finite set of non-empty words and  $x \in \mathcal{A}^{\mathbb{Z}}$  be a non-periodic point. Then,  $x$  possesses at most  $|\mathcal{W}|$  disjoint  $\mathcal{W}$ -factorizations.*

It is interesting to mention that in the proof of Theorem 2.1 one can replace the so called “infection lemma” by Theorem 3.1. The original proof of Theorem 2.1 is by contradiction, reducing the study to the case where there are more than  $|\mathcal{W}|^{|\mathcal{W}|}$  disjoint factorizations of a given point for some finite set of non-empty words  $\mathcal{W}$ . Using Theorem 3.1 the constant  $|\mathcal{W}|^{|\mathcal{W}|}$  can be reduced to  $|\mathcal{W}| + 1$  answering the last question in [DM08]. We will see in the next section that the related result, Theorem 5.1 from [BSTY19], can also be improved using the same theorem.

**3.2. Recognizability for morphisms.** Let  $\tau: \mathcal{A}^* \rightarrow \mathcal{B}^*$  be a morphism and  $x \in \mathcal{B}^{\mathbb{Z}}$ . If  $x = S^k \tau(y)$  for some  $y \in \mathcal{A}^{\mathbb{Z}}$  and  $k \in \mathbb{Z}$ , then  $(k, y)$  is said to be a  $\tau$ -representation of  $x$ . If  $y$  belongs to some  $Y \subseteq \mathcal{A}^{\mathbb{Z}}$  we say  $(k, y)$  is a  $\tau$ -representation of  $x$  in  $Y$ . Moreover, if  $0 \leq k < |\tau(y_0)|$  then  $(k, y)$  is a *centered  $\tau$ -representation* of  $x$  in  $Y$ . Following the notion introduced in [BSTY19], we say that  $\tau$  is *recognizable in  $Y$*  if each  $x \in \mathcal{B}^{\mathbb{Z}}$  has at most one centered  $\tau$ -representation in  $Y$ . If any aperiodic point  $x \in \mathcal{B}^{\mathbb{Z}}$  has at most one centered  $\tau$ -representation in  $Y$ , we say that  $\tau$  is *recognizable in  $Y$  for aperiodic points*.

The following lemma is proven in [BSTY19] using a contradiction argument that does not give any information on the constant  $R$  below. We provide a different and direct proof showing the origin of such constant  $R$ .

**Lemma 3.2.** *Let  $Y \subseteq \mathcal{A}^{\mathbb{Z}}$  be a subshift and  $\tau: \mathcal{A}^* \rightarrow \mathcal{B}^*$  be a non-erasing morphism. Let  $X \subseteq \mathcal{B}^{\mathbb{Z}}$  be the smallest subshift containing  $\tau(Y)$ . Then, the following are equivalent:*

- (1)  $\tau$  is recognizable in  $Y$ ;
- (2)  $\tau([a])$  is a non-empty clopen subset of  $X$  for each  $a \in \mathcal{A}$ , where  $[a] = \{y = (y_n)_{n \in \mathbb{Z}} \in Y : y_0 = a\}$ ;
- (3) there exists a positive integer  $R$  such that if  $(k, y)$  and  $(k', y')$  are centered  $\tau$ -representations of  $x, x' \in X$  respectively and  $x_{[-R, R]} = x'_{[-R, R]}$ , then  $k = k'$  and  $y_0 = y'_0$ .

The constant  $R$  provided by Lemma 3.2 is called a *recognizability constant* for  $\tau$ .

*Proof.* Since  $\tau$  is non erasing, for every  $a \in \mathcal{A}$  the set  $\tau([a])$  is clearly a non-empty closed subset of  $X$ . From the hypothesis  $X$  is equal to  $\cup_{a \in \mathcal{A}} \cup_{0 \leq k < |\tau(a)|} S^k \tau([a])$ .

If the morphism  $\tau$  is recognizable in  $Y$ , then the closed sets  $S^k \tau([a])$  are pairwise disjoint. It follows that they are also open and hence they are clopen sets.

If the set  $\tau([a])$  is a non-empty clopen set, then for each integer  $k$  the set  $S^k \tau([a])$  is also a non-empty clopen set. This means that each  $S^k \tau([a])$  is a finite union of cylinders  $[u.v] = \{x \in X : x_{[-|u|, |v|]} = uv\}$  for words  $u, v$  in  $\mathcal{B}^*$ . Without loss of generality we can assume  $|u| = |v| = R$  for some integer  $R$ , for all the words  $u$  and  $v$  and  $0 \leq k < |\tau(a)|$ . Hence, if  $x_{[-R, R]} = x'_{[-R, R]}$  for  $x, x' \in X$ , then  $x$  and  $x'$  belong to the same clopen set  $S^k \tau([a])$  which proves (3).

Clearly (3) implies the recognizability of  $\tau$  in  $Y$ . □

Observe that Lemma 3.2 states that the notion of recognizability for a morphism is local. Historically, a first notion of recognizability for substitutions appears with

the work of Mossé [Mos92]. Actually the recognizability of a non-erasing morphism does imply the recognizability in the sense of Mossé [BSTY19].

For composition of morphisms, in [BSTY19] the following result is proved.

**Lemma 3.3.** [BSTY19, Lemma 3.5] *Let  $\tau: \mathcal{A}^* \rightarrow \mathcal{B}^*$  and  $\sigma: \mathcal{B}^* \rightarrow \mathcal{C}^*$  be non-erasing morphisms,  $Y \subseteq \mathcal{A}^{\mathbb{Z}}$  be a subshift and  $X$  be the smallest subshift containing  $\tau(Y)$ . Then  $\sigma \circ \tau$  is recognizable in  $Y$  if and only if  $\tau$  is recognizable in  $Y$  and  $\sigma$  is recognizable in  $X$ .*

In the next lemma we show that a non-recognizable morphism is the composition of a recognizable one with a letter-to-letter morphism.

**Lemma 3.4.** *Let  $Y \subseteq \mathcal{A}^{\mathbb{Z}}$  be a subshift and  $\tau: \mathcal{A}^* \rightarrow \mathcal{B}^*$  be a non-erasing morphism. Then there exist an alphabet  $\mathcal{C}$ , a morphism  $\sigma: \mathcal{A}^* \rightarrow \mathcal{C}^*$  recognizable in  $Y$  for aperiodic points and a map  $\psi: \mathcal{C} \rightarrow \mathcal{B}$  such that*

$$\tau = \psi \circ \sigma.$$

*Proof.* Set  $\mathcal{C} = \mathcal{B} \times \mathcal{A}$ . Let  $\sigma$  and  $\psi$  be the morphisms defined by

$$\sigma(a) = (1, a) \cdots (|\tau(a)|, a) \text{ and } \psi(b, a) = \tau(a)_b, \text{ for } b \in \mathcal{B}, a \in \mathcal{A}.$$

These maps clearly satisfy the required relation.  $\square$

Inspired by the definitions on  $\mathcal{S}$ -adic subshifts in [BSTY19], let us define a last notion of recognizability for a family of words with respect to a subshift.

Let  $\mathcal{W} \subset \mathcal{B}^*$  be a finite set of non-empty words and let  $X \subseteq \mathcal{B}^{\mathbb{Z}}$  be a subshift. Consider a subshift  $Y \subseteq \mathcal{A}^{\mathbb{Z}}$  and a morphism  $\tau: \mathcal{A}^* \rightarrow \mathcal{B}^*$  such that  $\tau(\mathcal{A}) = \mathcal{W}$ . In particular  $\tau$  is non-erasing.

We say  $\mathcal{W}$  is *recognizable in  $X$  with respect to  $(Y, \tau)$*  if  $X$  is the smallest subshift containing  $\tau(Y)$ , (i.e.,  $X = \bigcup_{k \in \mathbb{Z}} S^k(\tau(Y))$ ) and  $\tau$  is recognizable in  $Y$ , (i.e., every  $x \in X$  has a unique centered  $\tau$ -representation in  $Y$ ). We say that  $\mathcal{W}$  is *recognizable in  $X$*  if it is recognizable in  $X$  with respect to some subshift  $Y$  and a morphism  $\tau$  as before.

We say that a  $\mathcal{W}$ -factorization  $F_x: \mathbb{Z} \rightarrow \mathcal{W} \times \mathbb{Z}$  of  $x \in X$  is *adapted to  $(Y, \tau)$*  if there exists  $y = (y_i)_{i \in \mathbb{Z}} \in Y$  such that  $\tau(y_i)$  is the first coordinate of  $F_x(i)$  for all  $i \in \mathbb{Z}$ .

The following lemma follows straightforwardly from definitions.

**Lemma 3.5.** *Let  $\mathcal{W} \subset \mathcal{B}^*$  be a finite set of non-empty words and  $X \subseteq \mathcal{B}^{\mathbb{Z}}$  be a subshift. Consider a subshift  $Y \subseteq \mathcal{A}^{\mathbb{Z}}$  and a morphism  $\tau: \mathcal{A}^* \rightarrow \mathcal{B}^*$  such that  $\tau(\mathcal{A}) = \mathcal{W}$ . The following statements are equivalent:*

- (1)  $\mathcal{W}$  is recognizable in  $X$  with respect to  $(Y, \tau)$ ;
- (2)  $X$  is the smallest subshift containing  $\tau(Y)$  and every  $x \in X$  has a unique  $\mathcal{W}$ -factorization adapted to  $(Y, \tau)$ .

Lemma 3.2 together with Lemma 3.5 imply the following local property for this last notion of recognizability.

**Lemma 3.6.** *Let  $\mathcal{W} \subset \mathcal{B}^*$  be a finite set of non-empty words and let  $X \subseteq \mathcal{B}^{\mathbb{Z}}$  be a subshift. Consider a subshift  $Y \subseteq \mathcal{A}^{\mathbb{Z}}$  and a morphism  $\tau: \mathcal{A}^* \rightarrow \mathcal{B}^*$  such that  $\tau(\mathcal{A}) = \mathcal{W}$ . Assume  $\mathcal{W}$  is recognizable in  $X$  with respect to  $(Y, \tau)$  and let  $R$  be a recognizability constant for  $\tau$ . Then, for all  $x \in X$  the set  $\{x' \in X : F_{x'}(0) = F_x(0)\}$  is non-empty and clopen. Moreover, if  $x, x' \in X$  are such that  $x_{[-R, R)} = x'_{[-R, R)}$  then  $F_x(0) = F_{x'}(0)$ .*

**3.3. Recognizability for directive sequences.** A directive sequence  $\tau = (\tau_n: \mathcal{A}_{n+1}^* \rightarrow \mathcal{A}_n^*)_{n \geq 0}$  is *recognizable at level  $n$*  if  $\tau_n$  is recognizable in  $X_\tau^{(n+1)}$ . The directive sequence  $\tau$  is *recognizable* if it is recognizable at level  $n$  for each  $n \geq 0$ . If there is an  $N \geq 0$  such that  $\tau$  is recognizable at level  $n$  for each  $n \geq N$ , then we say that  $\tau$  is *eventually recognizable*.

It is not difficult to prove using the fact that  $\tau_n(X_\tau^{(n+1)})$  is included in  $X_\tau^{(n)}$  that the directive sequence  $\tau$  is recognizable if and only if for all  $n \geq 0$  and  $x \in X_\tau (= X_\tau^{(0)})$  there is at most one  $y \in X_\tau^{(n)}$  and  $0 \leq k < |\tau_{[0,n)}(y_0)|$  such that  $x = S^k(\tau_{[0,n)}(y))$ . In fact if  $\tau$  is recognizable, then all  $x \in X_\tau$  has a unique such couple  $(k, y)$ . The same result follows for the notion of eventually recognizable directive sequences.

Given an  $\mathcal{S}$ -adic subshift with bounded alphabet rank, in [BSTY19] (inspired by Theorem 2.1) the authors show that there is a uniform bound for the number of morphisms in its directed sequence that are not recognizable for aperiodic points. Thus any  $\mathcal{S}$ -adic subshift is forced to be eventually recognizable for aperiodic points. The bound in [BSTY19] can be improved by following the same steps of the original proof but considering the bound in Theorem 3.1.

**Theorem 3.7.** *Let  $\tau = (\tau_n: \mathcal{A}_{n+1}^* \rightarrow \mathcal{A}_n^*)_{n \geq 0}$  be a sequence of morphisms with  $d = \liminf_{n \rightarrow \infty} |\mathcal{A}_n| < \infty$ . Then, there are at most  $\log_2 d$  morphisms  $\tau_n$  that are not recognizable in  $X_\tau^{(n+1)}$  for aperiodic points. In particular,  $\tau$  is eventually recognizable for aperiodic points.*

#### 4. FROM FINITE TOPOLOGICAL RANK MINIMAL CANTOR SYSTEMS TO $\mathcal{S}$ -ADIC SUBSHIFTS AND VICE VERSA

The relation between the classes of expansive minimal Cantor systems of finite topological rank and  $\mathcal{S}$ -adic subshifts with bounded alphabet rank is not fully understood, at least from a topological perspective. Nevertheless, it is implicit in the works of [DM08, BSTY19, DL12] that both classes somehow coincide. Indeed, in [BSTY19] it is proved under mild conditions that measure theoretically this holds. Also, in [DL12] conditions on a directive sequence are provided so that it generates an  $\mathcal{S}$ -adic subshift of finite topological rank.

The main purpose of this section is to prove this correspondence in the topological setting modulo the recognizability of the directive sequence defining the  $\mathcal{S}$ -adic system. In order to do that, we provide canonical constructions allowing to build a finite topological rank Bratteli-Vershik representation of a given recognizable  $\mathcal{S}$ -adic subshift and on the other hand to define the directive sequence of an  $\mathcal{S}$ -adic subshift conjugate to a given expansive finite topological rank minimal Cantor system.

More precisely, the main theorem of this section is the following.

**Theorem 4.1.**

- (1) *An expansive finite topological rank minimal Cantor system is topologically conjugate to a primitive recognizable  $\mathcal{S}$ -adic subshift with bounded alphabet rank. Moreover, the alphabet rank of the corresponding directive sequence is bounded by the rank of the minimal Cantor system.*
- (2) *Any primitive  $\mathcal{S}$ -adic subshift (so minimal) with bounded alphabet rank is a topological factor of a finite topological rank minimal Cantor system whose rank is bounded by the square of the alphabet rank of the subshift. Moreover,*

*when the  $\mathcal{S}$ -adic subshift is recognizable then the factor map is invertible and the subshift has finite topological rank.*

By results in [BSTY19] one has that all primitive  $\mathcal{S}$ -adic subshifts with bounded alphabet rank are eventually recognizable, so the factor map in (2) is just the one projecting the recognizable level to the given  $\mathcal{S}$ -adic subshift. It is also interesting to comment that there are several  $\mathcal{S}$ -adic subshifts, like Chacon subshift, that are not primitive but one can prove easily that they have finite topological rank (for Chacon example the topological rank is two).

**4.1. Combinatorial lemmas to get finite topological rank systems.** Before proving the main theorem of the section we develop two general combinatorial ideas that we think can be useful in different contexts.

**4.1.1. Combinatorial constructions using recognizable sets of words.** In this section we give a sufficient condition for a minimal subshift to have finite topological rank. Let  $X \subseteq \mathcal{B}^{\mathbb{Z}}$  be a subshift and  $\mathcal{W} \subset \mathcal{B}^*$  be a finite set of non-empty words that is recognizable in  $X$  with respect to  $(Y, \tau)$ , where  $Y \subseteq \mathcal{A}^{\mathbb{Z}}$  is a subshift and  $\tau : \mathcal{A}^* \rightarrow \mathcal{B}^*$  is a morphism with  $\tau(\mathcal{A}) = \mathcal{W}$ . Then for each  $x \in X$  there exists a unique  $\mathcal{W}$ -factorization  $F_x : \mathbb{Z} \rightarrow \mathcal{W} \times \mathbb{Z}$  adapted to  $(Y, \tau)$ .

For  $w = w^- w^+ \in \mathcal{W}$  with  $w^-, w^+ \in \mathcal{B}^*$  we define the set

$$\llbracket w^-.w^+ \rrbracket = \{x \in X : F_x(0) = (w, -|w^-|)\}.$$

In words,  $\llbracket w^-.w^+ \rrbracket$  is the set of points in  $X$  whose  $\mathcal{W}$ -factorizations adapted to  $(Y, \tau)$  have  $w$  as a centered word following the decomposition  $w^-.w^+$ , where the dot indicates the separation between negative and non-negative coordinates. If  $w^-$  is the empty word we just write  $\llbracket w \rrbracket$ . Clearly  $\llbracket w^-.w^+ \rrbracket = S^{|w^-|} \llbracket w \rrbracket$ . Given another word  $v \in \mathcal{W}$  we also set

$$\llbracket v, w^-.w^+ \rrbracket = \{x \in X : F_x(-1) = (v, -(|v| + |w^-|)), F_x(0) = (w, -|w^-|)\}.$$

The following result follows from Lemma 3.6.

**Lemma 4.2.** *Consider  $\mathcal{W} \subset \mathcal{B}^*$ ,  $X \subseteq \mathcal{B}^{\mathbb{Z}}$ ,  $Y \subseteq \mathcal{A}^{\mathbb{Z}}$  and  $\tau : \mathcal{A}^* \rightarrow \mathcal{B}^*$  as before. Assume  $\mathcal{W}$  is recognizable in  $X$  with respect to  $(Y, \tau)$  and let  $R$  be a recognizability constant for  $\tau$ . Then, for all  $w = w^- w^+, v \in \mathcal{W}$  the sets  $\llbracket w^-.w^+ \rrbracket$  and  $\llbracket v, w^-.w^+ \rrbracket$  are clopen, thus finite unions of cylinder sets of  $X$ :*

$$\llbracket v, w^-.w^+ \rrbracket = \bigcup_{x \in \llbracket v, w^-.w^+ \rrbracket} [x_{[-R-|v|, 0)} \cdot x_{[0, |w^+|+R)}].$$

The next theorem is the main construction of this subsection.

**Theorem 4.3** (Sufficient condition for finite topological rank). *Let  $(X, S)$  be a minimal subshift. Then its topological rank is bounded from above by  $K(X)^2$ , where*

$$K(X) = \lim_{n \rightarrow +\infty} \inf_{\substack{\mathcal{W} \subset \bigcup_{k \geq n} \mathcal{L}_k(X) \\ \mathcal{W} \text{ is recognizable in } X}} |\mathcal{W}|.$$

*Proof.* We assume that  $K = K(X)$  is finite, which is the unique relevant case. Under this assumption there exists a sequence  $(\mathcal{W}_n)_{n \geq 1}$  of subsets of  $\mathcal{B} = \mathcal{L}(X)$  such that  $|\mathcal{W}_n| = K$ ,  $\lim_{n \rightarrow +\infty} \inf_{w \in \mathcal{W}_n} |w| = +\infty$ , and  $\mathcal{W}_n$  is recognizable in  $X$  with respect to a subshift  $Y_n \subseteq \mathcal{A}_n^{\mathbb{Z}}$  and a morphism  $\tau_n : \mathcal{A}_n^* \rightarrow \mathcal{B}^*$  where  $\tau_n(\mathcal{A}_n) = \mathcal{W}_n$ . For all  $n \geq 1$  let  $R_n$  be a recognizability constant for  $\tau_n$ . We can assume that the sequence  $(R_n)_{n \geq 1}$  is increasing.

Now we will suppose, taking a subsequence of  $(\mathcal{W}_n)_{n \geq 1}$  if needed, that all  $w \in \bigcup_{n \geq 1} \mathcal{W}_n$  decomposes as  $w = w^- w^+$  in such a way that:

- (1) for all  $n \geq 2$ ,  $\max_{w \in \mathcal{W}_{n-1}}(|w^-|, |w^+|) < \min_{w \in \mathcal{W}_n}(|w^-|, |w^+|)/2$ ;
- (2) for all  $n \geq 2$ ,

$$\bigcup_{w \in \mathcal{W}_n} \llbracket w^-.w^+ \rrbracket \subseteq \llbracket v_{n-1}, w_{n-1}^-.w_{n-1}^+ \rrbracket \cap [l_{n-1}v_{n-1}w_{n-1}^-.w_{n-1}^+r_{n-1}]$$

for some  $w_{n-1}, v_{n-1} \in \mathcal{W}_{n-1}$  and  $l_{n-1}, r_{n-1} \in \mathcal{L}_{R_{n-1}}(X)$ ;

- (3) for all  $n \geq 1$ , the set

$$\mathcal{T}_n = \{S^{-i} \llbracket v, w^-.w^+ \rrbracket : v, w \in \mathcal{W}_n, 0 \leq i < |w^-| + |w^+|\}$$

is a partition of  $X$ .

First we check that once we have a sequence  $(\mathcal{W}_n)_{n \geq 1}$  with these properties, then the sequence of partitions  $(\mathcal{T}_n)_{n \geq 0}$ , with  $\mathcal{T}_0 = \{X\}$ , is a nested sequence of CKR partitions, (*i.e.*, it satisfies (KR1)-(KR4)). This will prove the theorem. After we will show how to build such a sequence  $(\mathcal{W}_n)_{n \geq 1}$ .

Let  $n \geq 1$ . We order in a particular way the sets in  $\mathcal{T}_n$  to get the structure of Kakutani-Rohlin towers. Each tower is indexed by pairs of words  $v, w \in \mathcal{W}_n$  such that  $vw$  appears in a  $\mathcal{W}_n$ -factorization in  $X$  adapted to  $(Y_n, \tau_n)$ . For such  $v, w \in \mathcal{W}_n$  the associated tower is given by:

$$\mathcal{T}_n(v, w) = \{S^{-i} \llbracket v, w^-.w^+ \rrbracket : 0 \leq i < |w^-| + |w^+|\}.$$

In the nomenclature of Kakutani-Rohlin partitions the base of this tower is given by  $\llbracket v, w^-.w^+ \rrbracket$ . Observe that this set is clopen by Lemma 4.2. Thus we have defined at most  $K^2$  towers per level.

By Property (2) the sequence of towers satisfies (KR1).

Now we prove (KR2) holds. Let  $n \geq 2$ ,  $v, w \in \mathcal{W}_n$  and  $i \in [0, |w^-| + |w^+|]$  be an integer. From property (2), since the length of  $l_{n-1}$  and  $r_{n-1}$  are  $R_{n-1}$ , all elements  $x, x' \in S^{-i} \llbracket v, w^-.w^+ \rrbracket$  satisfy  $x_{[-R_{n-1}, R_{n-1})} = x'_{[-R_{n-1}, R_{n-1})}$ . Hence the unicity of the factorization into  $\mathcal{W}_{n-1}$  words (Lemma 3.6) and Property (1) imply there is a sequence of words  $u_1, \dots, u_\ell \in \mathcal{W}_{n-1}$  with  $u_1 = u_\ell = w_{n-1}$  such that any set  $S^{-i} \llbracket v, w^-.w^+ \rrbracket$  in  $\mathcal{T}_n(v, w)$  is contained in some  $S^{-j} \llbracket u_k, u_{k+1}^-.u_{k+1}^+ \rrbracket \in \mathcal{T}_{n-1}(u_k, u_{k+1})$  for some  $0 \leq k \leq \ell$ , with  $u_0 = v_{n-1}$  (Property (2)). This proves (KR2) and provides  $v^+w^- = u_1^+u_2 \cdots u_{\ell-1}u_\ell^-$ .

Property (KR3) is deduced from properties (1) and (2). Indeed, Property (2) implies that for  $n \geq 2$  and  $v, w \in \mathcal{W}_n$  as before,

$$\llbracket v, w^-.w^+ \rrbracket \subseteq \llbracket w^-.w^+ \rrbracket \subseteq \llbracket v_{n-1}, w_{n-1}^-.w_{n-1}^+ \rrbracket,$$

for fixed words  $v_{n-1}, w_{n-1} \in \mathcal{W}_{n-1}$ , and Property (1) implies that the lengths of  $w_{n-1}^-$  and  $w_{n-1}^+$  grow to infinity with  $n$ .

Finally, to show (KR4) it is enough to observe from Property (2) that each atom of  $\mathcal{T}_n$  of the form  $S^{-i} \llbracket v, w^-.w^+ \rrbracket$  is included into a classical cylinder set of the form  $[u.u']$  for words  $u$  and  $u'$  whose lengths are greater than  $\min\{|w_{n-1}^+|, |w_{n-1}^-|\}$  (going to infinity with  $n$  by Property (1)). So the diameter of each atom of  $\mathcal{T}_n$  decreases to 0 which shows (KR4).

Now we describe how to construct from the sequence of finite sets of words  $(\mathcal{W}_n)_{n \geq 1}$  at the beginning of the proof a sub-sequence  $(\mathcal{W}_{n_i})_{i \geq 1}$  satisfying Properties (1)-(3). This is done recursively.

Let  $n_1 > 0$  be an integer such that  $\inf_{w \in \mathcal{W}_{n_1}} |w| > 2$ . For any word  $w \in \mathcal{W}_{n_1}$ , we choose an arbitrary decomposition  $w = w^- w^+$  such that  $|w^+|, |w^-| \geq |w|/2 - 1$ . We only have Property (3) to check. It is obviously done by the recognizability of the set  $\mathcal{W}_{n_1}$  with respect to  $(Y_{n_1}, \tau_{n_1})$ .

Let us assume we have constructed  $\mathcal{W}_{n_1}, \dots, \mathcal{W}_{n_\ell}$  with the desired properties. We choose two arbitrary words  $v_{n_\ell}$  and  $w_{n_\ell}$  in  $\mathcal{W}_{n_\ell}$  such that  $v_{n_\ell} w_{n_\ell}$  appears in a  $\mathcal{W}_{n_\ell}$ -factorization of  $X_{n_\ell}$  adapted to  $(Y_{n_\ell}, \tau_{n_\ell})$ . Let  $l_{n_\ell}$  and  $r_{n_\ell}$  be two words of length  $R_{n_\ell}$  such that  $l_{n_\ell} v_{n_\ell} w_{n_\ell} r_{n_\ell}$  belongs to  $\mathcal{L}(X)$  and  $[l_{n_\ell} v_{n_\ell} w_{n_\ell}^+ r_{n_\ell}] \subset \llbracket v_{n_\ell}, w_{n_\ell}^- \cdot w_{n_\ell}^+ \rrbracket$  (Lemma 3.6).

By minimality of  $(X, S)$ , since the lengths of the words of  $\mathcal{W}_n$  grow to infinity, for any large enough integer  $n$  (greater than some  $N_\ell$ ), any word  $w$  in  $\mathcal{W}_n$  contains an occurrence of  $l_{n_\ell} v_{n_\ell} w_{n_\ell} r_{n_\ell}$ . Choosing such an occurrence leads to decompose  $w = w^- w^+$ , where the words  $w^-$  and  $w^+$  contain respectively the word  $l_{n_\ell} v_{n_\ell} w_{n_\ell}^-$  and  $w_{n_\ell}^+ r_{n_\ell}$  as a suffix and as a prefix respectively. Moreover, taking  $n$  large enough, we can also assume that the occurrence is far away from the border of  $w$  so that

$$\max_{w \in \mathcal{W}_{n_\ell}} (|w^-|, |w^+|) < \min_{w \in \mathcal{W}_n} (|w^-|, |w^+|)/2.$$

We choose one integer  $n_{\ell+1}$  realizing this. The last inequality ensures Property (1) and the construction ensures Property (2).

To show  $\mathcal{T}_{n_{\ell+1}}$  is a partition observe that given  $x \in X$ , the recognizability of  $\mathcal{W}_{n_{\ell+1}}$  in  $X$  with respect to  $(Y_{n_{\ell+1}}, \tau_{n_{\ell+1}})$  implies there are unique words in  $v, w \in \mathcal{W}_{n_{\ell+1}}$  and  $0 \leq i < |w^-| + |w^+|$  such that  $S^i x$  belongs to  $\llbracket v, w^- \cdot w^+ \rrbracket$ . Hence  $\mathcal{T}_{n_{\ell+1}}$  covers the space  $X$ . By construction all the sets  $S^{-i} \llbracket v, w^- \cdot w^+ \rrbracket$  are pairwise disjoint.  $\square$

We deduce the following theorem.

**Theorem 4.4.** *Let  $(X, S)$  be a minimal subshift. It has finite topological rank if and only if  $K(X)$  is finite.*

*Proof.* The sufficient condition follows directly from Theorem 4.3.

To prove the necessity of the condition assume  $(X, S)$  has finite topological rank and that it is given by a Bratteli-Vershik representation  $B = (V, E, \preceq)$  where the number of vertices per level is uniformly bounded by the rank. Moreover, by Theorem 2.1, after telescoping the diagram and because  $(X_B, V_B)$  is expansive, we can assume  $(X_B, V_B)$  is topologically conjugate to the subshift  $(X_B^{(0)}, S)$ ,  $X_B^{(0)} \subseteq E_1^{\mathbb{Z}}$ , defined by coding orbits of the Bratteli-Vershik representation on the first level of the diagram via the map  $\phi_0$  (we follow the notations in Section 2.3.2).

We claim it is enough to show that  $K(X_B^{(0)})$  is finite. Indeed, assume there is a sequence  $(\mathcal{W}_n)_{n \geq 1}$  of subsets of  $\mathcal{B} = \mathcal{L}(X_B^{(0)})$  such that for all  $n \geq 1$ ,  $|\mathcal{W}_n| \leq K(X_B^{(0)}) < +\infty$ ,  $\lim_{n \rightarrow +\infty} \inf_{w \in \mathcal{W}_n} |w| = +\infty$  and  $\mathcal{W}_n$  is recognizable in  $X_B^{(0)}$  with respect to a subshift  $Y_n \subseteq \mathcal{A}_n^{\mathbb{Z}}$  and a morphism  $\tau_n : \mathcal{A}_n^* \rightarrow \mathcal{B}^*$  with  $\tau_n(\mathcal{A}_n) = \mathcal{W}_n$ . By the Curtis-Hedlund-Lyndon theorem, the conjugacy map  $\phi : X_B^{(0)} \rightarrow X$  is given by a block code  $\hat{\phi} : \mathcal{L}_{2r+1}(X^{(0)}) \rightarrow \mathcal{L}_1(X)$  of some radius  $r$ : for all  $i \in \mathbb{Z}$

$$\phi(x)_i = \hat{\phi}(x_{[i-r, i+r]}).$$

For any large enough  $n$ , the length of any word in  $\mathcal{W}_n$  is larger than  $2r + 1$ . Consider a triple of words  $u, w, v \in \mathcal{W}_n$  such that  $uwv$  appears in a  $\mathcal{W}_n$ -factorization in  $X_B^{(0)}$  adapted to  $(Y_n, \tau_n)$ . Its image  $\hat{\phi}(uwv)$  by the local map is of the form

$\hat{u}\hat{w}\hat{v} \in \mathcal{L}_{|u|-r+|w|+|v|-r}(X)$  with  $|\hat{w}| = |w|$ ,  $|\hat{u}| = |u| - r$  and  $|\hat{v}| = |v| - r$ . Thus we define  $\hat{\mathcal{W}}_n$  to be the collection of words  $\hat{w}$  built in this way. It is clear from construction that such sets of words have cardinality at most  $K(X_B^{(0)})^3$  and since the inverse of the map  $\phi$  is also given by a block code, the set  $\hat{\mathcal{W}}_n$  is recognizable in  $X$  with respect to  $(Y_n, \tau_n)$ . It follows that  $K(X)$  is finite.

Now we produce from the Bratteli-Vershik representation  $B$  the sets of words  $(\mathcal{W}_n)_{n \geq 1}$  as described before. Recall that  $X_B^{(0)}$  is the  $\mathcal{S}$ -adic subshift generated by the sequence  $\tau^B = (\tau_n^B : V_{n+1}^* \rightarrow V_n^*)_{n \geq 0}$  of morphisms read on  $B$ . For each  $n \geq 1$  we define  $\mathcal{W}_n$  to be the set of words  $\tau_{[0,n)}^B(v) \in E_1^{\mathbb{Z}}$  with  $v \in V_n$ . Clearly the cardinality of each  $\mathcal{W}_n$  is bounded by the topological rank of  $(X, \mathcal{S})$ . Recall that the map  $\phi_0 : X_B \rightarrow X_B^{(0)}$  is a homeomorphism. It is simple to check that the preimage by  $\phi_0$  of the sets  $\tau_{[0,n)}^B([v])$ ,  $v \in V_n$ , are the roofs of towers of level  $n$  in the Bratteli-Vershik representation of the system and in particular are clopen sets. Thus, by Lemmas 3.2 and 3.5, the set  $\mathcal{W}_n$  is recognizable in  $X_B^{(0)}$  with respect to  $(X_B^{(n)}, \tau_{[0,n)}^B)$ . Since the sequence  $\tau^B$  is primitive, the lengths of the words in  $\mathcal{W}_n$  grow with  $n$  to infinity. This construction shows that  $K(X_B^{(0)})$  is finite and thus the theorem holds.  $\square$

**4.2. Proof of Theorem 4.1.** We have all the elements to prove Theorem 4.1. Actually Part (1) of Theorem 4.1 comes from a property (Proposition 4.6) available for Cantor minimal systems of eventually infinite rank. This proposition provides a necessary condition on a Bratteli-Vershik system to be expansive in terms of recognizability of the sequence of morphisms read on the diagram. This condition is not too far from being sufficient, but the morphisms need to be proper to get the converse (see Proposition 4.5). The results are presented in Subsection 4.2.1. The second subsection is devoted to the second part of Theorem 4.1 stated as Proposition 4.7.

**4.2.1. From finite topological rank systems to  $\mathcal{S}$ -adic systems.** First, we recall a useful property to identify a Bratteli-Vershik system with an  $\mathcal{S}$ -adic subshift. We call a *hat-morphism* any morphism  $\tau : \mathcal{A}^* \rightarrow \mathcal{B}^*$  such that for all  $a, b \in \mathcal{A}$ ,  $a \neq b$ , the letters occurring in  $\tau(a)$  and  $\tau(b)$  are all distinct. For instance the morphism read at the first level on an ordered Bratteli diagram is a hat-morphism (see Section 2.3.2). The next result is a slight modification of [DL12, Proposition 2.2].

**Proposition 4.5** ([DL12, Proposition 2.2]). *Let  $\tau = (\tau_n : \mathcal{A}_{n+1} \rightarrow \mathcal{A}_n)_{n \geq 0}$  be a recognizable directive sequence of morphisms such that  $\tau_0$  is a hat-morphism and  $(\tau_n)_{n \geq 1}$  is proper. Then, the  $\mathcal{S}$ -adic subshift generated by  $\tau$  is topologically conjugate to the Bratteli-Vershik system  $(X_B, V_B)$  where  $B$  is the Bratteli diagram such that  $\tau$  is the sequence of morphisms read on  $B$ .*

Here is a converse to Proposition 4.5.

**Proposition 4.6.** *Let  $(X, T)$  be a minimal Cantor system given by a Bratteli-Vershik representation  $B$ . If  $(X, T)$  is expansive, then the sequence of morphisms  $\tau^B = (\tau_n^B : V_{n+1}^* \rightarrow V_n^*)_{n \geq 0}$  read on  $B$  is eventually recognizable. Moreover,  $(X, T)$  is topologically conjugate to the  $\mathcal{S}$ -adic subshift generated by a proper recognizable directive sequence obtained from  $\tau^B$  after a telescoping of  $B$ .*



*Proof.* Let us assume that  $(X, T)$  is an expansive minimal Cantor system. Then, the system is topologically conjugate to  $(X_B^{(N)}, S)$  for some  $N \geq 0$ . Thus, by telescoping finitely many levels of the diagram  $B$ , we can assume that  $N = 0$ , so  $(X, T)$  is topologically conjugate to  $(X_B^{(0)}, S)$  (see Section 2.3.2 for notations). In addition, by simplicity of  $B$ , after another telescoping we can assume that the directive sequence  $\tau^B = (\tau_n^B : V_{n+1}^* \rightarrow V_n^*)_{n \geq 0}$  of morphisms read on  $B$  is proper. To conclude we only need to prove that  $\tau^B$  is recognizable. The result follows after interpreting the telescoping process.

Recall that under our assumptions for each  $n \geq 0$  the subshift  $(X_B^{(n)}, S)$  is minimal and  $X_B^{(n)}$  is the smallest subshift containing  $\tau_n^B(X_B^{(n+1)})$ . This implies that  $X_B^{(0)}$  is the smallest subshift containing  $\tau_{[0, n+1)}^B(X_B^{(n+1)})$ . Since  $\tau_{[0, n+1)}^B([a])$  is a clopen set for any letter  $a \in V_{n+1}$  (just consider its preimage by the homeomorphism  $\phi_0 : X_B \rightarrow X_B^{(0)}$ ), then by Lemma 3.2 we get that the morphism  $\tau_{[0, n+1)}^B$  is recognizable in  $X_B^{(n+1)}$ . This property and a recursive use of Lemma 3.3 proves that  $\tau^B$  is recognizable.  $\square$

**4.2.2. From recognizable  $\mathcal{S}$ -adic subshifts to finite topological rank minimal Cantor systems.** For  $\mathcal{S}$ -adic subshifts generated by non necessarily proper morphisms, we have the following dynamical property.

**Proposition 4.7.** *Let  $\tau$  be a primitive directive sequence with finite alphabet rank  $AR(\tau) < +\infty$ . Then, the  $\mathcal{S}$ -adic subshift generated by  $\tau$  is a factor of a finite topological rank minimal Cantor system whose rank is bounded by  $AR(\tau)^2$ . Moreover, when the  $\mathcal{S}$ -adic subshift is recognizable then the factor map is invertible.*

*Proof.* Let  $\tau = (\tau_n : \mathcal{A}_{n+1}^* \rightarrow \mathcal{A}_n^*)_{n \geq 0}$ . If the  $\mathcal{S}$ -adic subshift  $X_\tau$  generated by  $\tau$  is finite, the result is classical since any minimal periodic system occurs as the factor of an odometer (a system with topological rank one). When  $X_\tau$  is infinite, Theorem 3.7 implies that  $\tau$  is eventually recognizable. This means that for some  $N \geq 1$ , the directive sequence  $(\tau_n)_{n \geq N}$  is recognizable. They generate the  $\mathcal{S}$ -adic subshift  $X_\tau^{(N)}$ . Let  $\psi$  and  $\sigma$  be the morphisms provided by Lemma 3.4 when applied to  $\tau = \tau_{[0, N)}$ . Hence the subshift  $X_\tau$  is a factor, via the map induced by  $\psi$ , of the  $\mathcal{S}$ -adic subshift  $X_{\tau'}$  generated by the recognizable directive sequence  $\tau' = (\sigma, \tau_N, \tau_{N+1}, \dots)$ . Thus the set of words  $\mathcal{W}_n = \{\tau'_{[0, n+1)}(a) : a \in \mathcal{A}_{n+1}\}$  for all  $n \geq 0$  is recognizable in  $X_{\tau'} = X_{\tau'}^{(0)}$  with respect to  $(X_{\tau'}^{(n+1)}, \tau'_{[0, n+1)})$ . Since  $\tau'$  is primitive, the lengths of words in  $\mathcal{W}_n$  grow to infinity with  $n$ , so that  $K(X_{\tau'})$  defined in Theorem 4.3 is bounded by  $AR(\tau)$ . This implies it has finite topological rank and that its topological rank is bounded by  $K(X_{\tau'})^2$ , showing the result.  $\square$

## 5. NON-SUPERLINEAR COMPLEXITY SUBSHIFTS HAVE FINITE TOPOLOGICAL RANK

In this section we show that subshifts with non-superlinear complexity have finite topological rank. Even better, we prove this also holds under the weaker assumption that the complexity function satisfies  $\liminf_{n \rightarrow +\infty} p_X(n+1) - p_X(n) < +\infty$  (Theorem 5.5). To do this, we follow an idea originally due to T. Monteil to interpret complexity conditions in terms of graph properties. Thus we characterize minimal subshifts with non-superlinear complexity in terms of deconnectability of the associated Rauzy graphs. Next we use the characterization of subshifts of finite topological rank provided in Theorem 4.4 to prove Theorem 5.5. Notice that the

converse of Theorem 5.5 does not hold as illustrated by the example in Section 6. Let us also mention that in Section 6 we present sufficient conditions so that a finite topological rank subshift has a non-superlinear complexity.

**5.1. Subshifts with non-superlinear complexity and Rauzy Graphs.** We start with a combinatorial lemma on graphs. Next we apply this lemma to Rauzy graphs associated to subshifts to get bounds for the complexity of the subshift. For a directed graph  $G = (V, E)$  we denote by  $d^+(v)$  (resp.  $d^-(v)$ ) the outdegree (resp. indegree) of a vertex  $v \in V$ . Recall that a (directed) *path* in  $G$  is a finite sequence of distinct edges  $e_1, \dots, e_\ell \in E$  (oriented in the same direction) which joins a sequence of distinct vertices ( $e_i = (v_i, v_{i+1}) \in E$  and  $e_{i+1} = (v_{i+1}, v_{i+2}) \in E$ ). Given a subset  $V' \subseteq V$  of vertices, we say that the path is *in the complement of  $V'$  except the border* whenever  $v_1, v_{\ell+1}$  are in  $V'$  and  $v_i \notin V'$  for  $i \in \{2, \dots, \ell\}$ . A *tree* is an unoriented acyclic connected graph and a *forest* is a disjoint union of trees. Recall that a *directed forest* is a directed acyclic graph whose underlying undirected graph is a forest.

**Lemma 5.1.** *Let  $G = (V, E)$  be a strongly connected finite directed graph and let  $V'$  be the set of vertices  $v \in V$  such that  $d^+(v) \geq 2$ . Assume  $V'$  is non-empty. Then,*

- *the subgraph induced by the set of vertices  $V \setminus V'$  is a directed forest;*
- *the number of directed paths in  $G$  in the complementary of  $V'$  except the border, is bounded from above by  $|V'|^2 \cdot (\max_{v \in V} d^+(v)) \cdot (\max_{v \in V} d^-(v))$ .*

*Proof.* Let  $G[V \setminus V']$  denote the subgraph induced by the set of vertices  $V \setminus V'$ . The graph  $G[V \setminus V']$  has no directed cycles. This follows noticing that if there is a directed cycle in it, then the strong connectivity of  $G$  implies that one of the vertex in the cycle should have at least two outgoing edges: this contradicts the definition of  $G[V \setminus V']$ . There is actually no undirected cycles in it. Otherwise, a simple undirected cycle which is not a directed cycle necessarily has a vertex whose outdegree is at least two, leading again to a contradiction. To prove this fact, observe that restricted to such cycle, the sum of the outdegrees and the sum of the indegrees (for the induced orientation) are both equal to the number of edges which is equal to the number of vertices. It follows that each indegree and outdegree equals 1 and the cycle is a directed cycle: a contradiction. Hence the graph  $G[V \setminus V']$  is a directed forest.

To bound the number of directed paths in  $G$  in the complementary of  $V'$  except the border, observe that such a path is uniquely determined by its starting edge (an outgoing edge of one element of  $V'$ ) and its ending edge (an ingoing edge of an element of  $V'$ ). Indeed, given these choices, the path determines a path  $\gamma$  in  $G[V \setminus V']$ , eventually empty if the starting and ending edges are the same. Since  $G[V \setminus V']$  is a directed forest, the path  $\gamma$  is uniquely determined by the starting and ending edges.  $\square$

Let  $(X, S)$  be a subshift. The *Rauzy graph* of  $(X, S)$  at level  $n \geq 2$  is the directed graph  $G_n = (V_n, E_n)$  with vertex set  $V_n = \mathcal{L}_n(X)$ , where the ordered pair of words  $(u_0 \dots u_{n-1}, v_0 \dots v_{n-1})$  is an edge in  $E_n$  if  $u_1 \dots u_{n-1} = v_0 \dots v_{n-2}$ . The indegree and outdegree of each vertex are bounded from above by the cardinality of the alphabet. Moreover, when the subshift  $(X, S)$  is minimal the Rauzy graph  $G_n$  is strongly connected.

For a subset  $V'_n \subset V_n = \mathcal{L}_n(X)$  let  $G_n[V_n \setminus V'_n]$  denote the subgraph induced by the set of vertices  $V_n \setminus V'_n$ . Our target for  $V'_n$  will be the set of vertices in  $V_n$  with outdegree greater than 1, *i.e.*, the set of right special words of length  $n$ .

Using Lemma 5.1 we propose a combinatorial characterization of subshifts with non-superlinear complexity in terms of Rauzy graphs. It uses a refinement of the notion of  $K$ -deconnectability introduced by Monteil and developed together with Ferenczi in [FM10] to improve results of [Bos85]. We call this notion *strong  $K$ -deconnectability*.

**Definition 5.2.** *Let  $K \geq 1$  be an integer. We say that a minimal subshift  $(X, S)$  is  $K$ -deconnectable if there exist a sequence of non-negative integers  $(n_i)_{i \geq 1}$  and a constant  $K' \geq 1$  such that for all  $i \geq 1$  there exists a subset  $V'_{n_i} \subset \mathcal{L}_{n_i}(X)$  of at most  $K$  vertices such that every path in  $G_{n_i}[\mathcal{L}_{n_i}(X) \setminus V'_{n_i}]$  has a length at most  $K'n_i$ . If we further require that  $G_{n_i}[\mathcal{L}_{n_i}(X) \setminus V'_{n_i}]$  is a directed forest, we say that  $(X, S)$  is strongly  $K$ -deconnectable.*

The next proposition links the combinatorial properties of the Rauzy graph with the asymptotic behavior of the complexity function.

**Proposition 5.3.** *Let  $(X, S)$  be a minimal subshift. Then,  $X$  has non-superlinear complexity if and only if  $(X, S)$  is strongly  $K$ -deconnectable for some  $K \in \mathbb{N}^*$ .*

*Proof.* One implication is implicit in [FM10, Section 7.3.2] and we repeat the argument here for completeness. Suppose that  $\liminf_{n \rightarrow +\infty} p_X(n)/n < +\infty$ . We can find a constant  $K$  and a sequence of positive integers  $(n_i)_{i \geq 1}$  such that simultaneously  $p_X(n_i)/n_i \leq K$  and  $s(n_i) = p_X(n_i + 1) - p_X(n_i) \leq K$  (see for instance Boshernitzan [Bos85, Theorem 2.2]). So the number of right special words of length  $n_i$  is bounded by  $K$  for all  $i \geq 1$ . We may remove all these special words  $V'_{n_i}$  from the Rauzy graph  $G_{n_i} = (\mathcal{L}_{n_i}(X), E_{n_i})$  and we obtain the graph  $G_{n_i}[\mathcal{L}_{n_i}(X) \setminus V'_{n_i}]$ . By Lemma 5.1 it is a directed forest and  $(X, S)$  is strongly  $K$ -deconnectable.

Conversely, let  $(n_i)_{i \geq 1}$  be a sequence of positive integers and let  $V'_{n_i} \subseteq \mathcal{L}_{n_i}(X)$  be given by the strong  $K$ -deconnectability of  $G_{n_i}$ . Recall that the indegree and outdegree of any vertex in all Rauzy graphs  $G_{n_i}$  are bounded by the cardinality  $|\mathcal{A}|$  of the alphabet of  $X$ . As a consequence, the number of extremal vertices (that is, vertices with outdegree or indegree equal to 0) in the subgraph  $G[\mathcal{L}_{n_i}(X) \setminus V'_{n_i}]$  is not greater than  $2K|\mathcal{A}|$ . Observe that in  $G[\mathcal{L}_{n_i}(X) \setminus V'_{n_i}]$ , between two given extremal vertices there is at most one path because there are no cycles and all outdegrees are equal to 0 or 1. Hence the number of paths between two extremal vertices in  $G[\mathcal{L}_{n_i}(X) \setminus V'_{n_i}]$  is bounded by  $2K|\mathcal{A}| - 1$ , uniformly in  $i$ . Then, a word of length  $n_i$  being either of outdegree at least 2 or being a vertex of some path between two extremal vertices, one obtains:

$$p_X(n_i) = |\mathcal{L}_{n_i}(X)| \leq K + \sum_{\substack{\gamma: \text{path between extremal vertices} \\ \text{in } G[\mathcal{L}_{n_i}(X) \setminus V'_{n_i}]}} \text{length}(\gamma) \leq K + (2K|\mathcal{A}| - 1) \cdot K'n_i,$$

where  $\text{length}(\gamma)$  stands for the length of the path  $\gamma$ . This finishes the proof.  $\square$

**5.2. Subshift of return words.** Let  $(X, S)$  be a minimal subshift. Let  $\mathcal{W} \subseteq \mathcal{L}(X)$  be a finite set of non-empty words. In this paper we will only consider the case where all words in  $\mathcal{W}$  have the same length  $\ell_{\mathcal{W}}$ . A *first return word* between  $w_1$  and  $w_2 \in \mathcal{W}$  (eventually identical) is a word  $u \in \mathcal{L}(X)$  such that:

- i)  $uw_2$  belongs to  $\mathcal{L}(X)$ ;
- ii) the word  $w_1$  is a prefix of  $uw_2$  and  $uw_2$  has no other occurrences of words in  $\mathcal{W}$  than these two.

Let  $\mathcal{R}(\mathcal{W})$  be the collection of all first return words between words in  $\mathcal{W}$ . The minimality of  $(X, S)$  implies that  $\mathcal{R}(\mathcal{W})$  is a finite set.

It will be convenient to label the return words. For this, consider  $R(\mathcal{W}) = \{1, \dots, |\mathcal{R}(\mathcal{W})|\}$  and let  $\tau_{\mathcal{W}}: R(\mathcal{W}) \rightarrow \mathcal{R}(\mathcal{W})$  be a bijection. The map  $\tau_{\mathcal{W}}$  induces a morphism  $R(\mathcal{W})^* \rightarrow \mathcal{R}(\mathcal{W})^*$  also denoted by  $\tau_{\mathcal{W}}$ . We call it a *coding morphism associated to  $\mathcal{W}$* .

Consider the set  $Y_{\mathcal{R}(\mathcal{W})}$  of all points  $y \in R(\mathcal{W})^{\mathbb{Z}}$  such that  $\tau_{\mathcal{W}}(y)$  belongs to  $X$  and

$$\begin{aligned} & \{i \in \mathbb{Z} : \tau_{\mathcal{W}}(y)_{[i, i+\ell_{\mathcal{W}}]} \in \mathcal{W}\} \\ & = \{|\tau_{\mathcal{W}}(y_0 \cdots y_{\ell})| : \ell \geq 0\} \cup \{-|\tau_{\mathcal{W}}(y_{\ell} \cdots y_{-1})| : \ell \leq -1\} \cup \{0\}. \end{aligned}$$

It is closed and shift-invariant. We call  $(Y_{\mathcal{R}(\mathcal{W})}, S)$  the *subshift of return words to  $\mathcal{W}$  in  $X$* .

Actually, the morphism  $\tau_{\mathcal{W}}$  is recognizable in  $Y_{\mathcal{R}(\mathcal{W})}$  (see Section 3.2 for background on recognizability).

**Lemma 5.4.** *Let  $(X, S)$  be a minimal subshift and  $\mathcal{W} \subset \mathcal{L}(X)$  be a finite set of non-empty words of the same length  $\ell_{\mathcal{W}}$ . Then,  $\tau_{\mathcal{W}}$  is recognizable in  $Y_{\mathcal{R}(\mathcal{W})}$ .*

*Proof.* Let  $R$  be greater than  $2\ell_{\mathcal{W}} + 2\sup_{u \in \mathcal{R}(\mathcal{W})} |u|$ . According to Lemma 3.2, it is enough to show for  $x, x' \in X$  and  $(k, y), (k', y')$  two centered  $\tau_{\mathcal{W}}$ -representations of  $x, x'$  that  $k = k'$  and  $y_0 = y'_0$  whenever  $x_{[-R, R]} = x'_{[-R, R]}$ . By the choice of the constant  $R$ , there exist at least three occurrences of words belonging to  $\mathcal{W}$  in  $x_{[-R, R]}$ . The definition of elements of  $Y_{\mathcal{R}(\mathcal{W})}$  together with item ii) in the definition of first return words enable us to conclude.  $\square$

Now we are ready to deduce that a minimal subshift with non-superlinear complexity has finite topological rank. In fact we prove a stronger result. We only need to consider systems where  $\liminf_{n \rightarrow +\infty} p_X(n+1) - p_X(n)$  is finite, which ensures the conclusions of Lemma 5.1 for infinitely many Rauzy graphs of the system. This last property is related but weaker than the strong  $K$ -deconnectability one. It is nevertheless enough to give a uniform bound on the number of return words to the (right) special words.

**Theorem 5.5.** *Let  $(X, S)$  be an infinite minimal subshift on the alphabet  $\mathcal{A}$  such that  $K_1 = \liminf_{n \rightarrow +\infty} p_X(n+1) - p_X(n)$  is finite. Then, the topological rank of  $(X, S)$  is bounded from above by  $(1 + |\mathcal{A}|^2 K_1^2)^{2(K_1+2)}$ . In particular, minimal subshifts with non-superlinear complexity have finite topological rank.*

*Proof.* To get the bound on the topological rank of the system  $(X, S)$ , thanks to Theorem 4.3 it is enough to exhibit a sequence  $(\mathcal{W}_n)_{n \geq 1}$  of finite sets of words in  $\mathcal{A}^*$  such that  $\sup_{n \geq 1} |\mathcal{W}_n| \leq (1 + |\mathcal{A}|^2 K_1^2)^{K_1+2}$ ,  $\min_{w \in \mathcal{W}_n} |w|$  goes to infinity with  $n$  and each  $\mathcal{W}_n$  is recognizable in  $X$ .

Let  $(m_i)_{i \geq 1}$  be an increasing sequence of positive integers such that  $p_X(m_i + 1) - p_X(m_i)$  converges to  $K_1$  when  $i$  goes to infinity. It follows that there are at most  $K_1$  right special words of length  $m_i$  for every  $i \geq 1$ .

Fix  $i \geq 1$ . Let  $\mathcal{RS}_{m_i}(X)$  be the set of right special words of length  $m_i$  in  $\mathcal{L}(X)$ . Since any first return word between  $w_1$  and  $w_2$  in  $\mathcal{RS}_{m_i}(X)$  corresponds to a path

in the Rauzy graph  $G_{m_i}$  from  $w_1$  to  $w_2$  in the complementary of  $\mathcal{RS}_{m_i}(X)$  except the border, by Lemma 5.1 the cardinality of the set of first return words between all the pairs of special words  $\mathcal{R}(\mathcal{RS}_{m_i}(X))$  is bounded from above by  $|\mathcal{A}|^2 K_1^2$ .

Let  $\ell_i$  be the largest positive integer such that any two occurrences of the same special word of length  $m_i$  are at least  $\ell_i$  appart. Then, a standard compactness argument gives that  $\ell_i$  goes to infinity as  $i$  goes to infinity, otherwise we would have a periodic point.

For  $n \in \mathbb{N}$  write  $\mathcal{R}(\mathcal{RS}_{m_i}(X)) = \mathcal{R}_i^{\geq n} \cup \mathcal{R}_i^{< n}$ , where  $\mathcal{R}_i^{\geq n}$  is the set of all words in  $\mathcal{R}(\mathcal{RS}_{m_i}(X))$  of length larger than or equal to  $n$  and  $\mathcal{R}_i^{< n}$  contains the words of length smaller than  $n$ . We claim that any concatenation of  $K_1 + 1$  words in  $\mathcal{R}(\mathcal{RS}_{m_i}(X))$  occurring in a  $\mathcal{R}(\mathcal{RS}_{m_i}(X))$ -factorization has a word in  $\mathcal{R}_i^{\geq \lfloor \ell_i / (K_1 + 1) \rfloor}$ . To prove this statement, for  $x \in X$  let  $(p_j)_{j \in \mathbb{Z}}$  be the consecutive occurrences of the words of  $\mathcal{RS}_{m_i}(X)$  in  $x$ , *i.e.*,  $\{p_j : j \in \mathbb{Z}\} = \{\ell \in \mathbb{Z} : x_{[\ell, \ell + m_i)} \in \mathcal{RS}_{m_i}(X)\}$ . For each index  $j \in \mathbb{Z}$  set  $u_j = x_{[p_j, p_{j+1})} \in \mathcal{R}(\mathcal{RS}_{m_i}(X))$ . We have,

$$(K_1 + 1) \max_{j \leq \ell \leq j + K_1} |u_j| \geq |u_j \cdots u_{j+K_1}| \geq \ell_i.$$

This proves that for any  $j \in \mathbb{Z}$  there is  $\ell \in \{j, \dots, j + K_1\}$  such that  $u_\ell$  belongs to  $\mathcal{R}_i^{\geq \lfloor \ell_i / (K_1 + 1) \rfloor}$ , thus our claim holds.

For  $n \geq 1$  let  $i = i(n) \geq 1$  be large enough so that  $\ell_{i(n)} \geq (K_1 + 1)n$  and thus, by previous claim,  $\mathcal{R}_{i(n)}^{\geq n} \neq \emptyset$ . Let  $(Y_i, S) = (Y_{\mathcal{R}(\mathcal{RS}_{m_i}(X))}, S)$  be the subshift of return words to  $\mathcal{RS}_{m_i}(X)$  in  $X$  and  $\tau_i$  an associated coding morphism. For convenience, denote by  $\mathcal{B}_i$  the alphabet of  $Y_i$ , so  $\tau_i$  is defined from  $\mathcal{B}_i^*$  to  $\mathcal{R}(\mathcal{RS}_{m_i}(X))^*$ . Let  $Z_i$  be the return word subshift associated to the set (of letters)  $\{a \in \mathcal{B}_i : \tau_i(a) \in \mathcal{R}_i^{\geq n}\}$  in  $Y_i$ . Let  $\mathcal{C}_i$  denote the alphabet of  $Z_i$  and  $\sigma_i : \mathcal{C}_i^* \rightarrow \mathcal{B}_i^*$  denote an associated coding morphism. Finally we set  $\mathcal{W}_n = \tau_i \circ \sigma_i(\mathcal{C}_i) \subset \mathcal{L}(X)$ .

Observe that by previous claim any word in  $\mathcal{W}_n$  is of the form  $u_0 u_1 \dots u_k$  with  $1 \leq k \leq K_1 + 1$ ,  $u_0 \in \mathcal{R}_i^{\geq n}$  and  $u_j \in \mathcal{R}_i^{< n} \cup \{\epsilon\}$  for  $0 < j \leq k$ . It follows that  $|\mathcal{W}_n|$  is bounded by  $|\mathcal{R}_i^{\geq n}|(1 + |\mathcal{R}_i^{< n}|)^{K_1}$ , which is bounded by  $|\mathcal{A}|^2 K_1^2 (1 + |\mathcal{A}|^2 K_1^2)^{K_1}$  uniformly in  $i$  and  $n$ . Also, by definition, we have  $\min_{w \in \mathcal{W}_n} |w| \geq n$ . Moreover, Lemma 5.4 implies that  $\sigma_i$  and  $\tau_i$  are recognizable in  $Z_i$  and in  $Y_i$  respectively. By Lemma 3.3,  $\tau_i \circ \sigma_i$  is recognizable in  $Z_i$ . Finally, Lemma 3.5 gives that  $\mathcal{W}_n$  is recognizable in  $X$ . This shows the result.  $\square$

**Remark 5.6.** In the previous proof, it is worth to notice that if  $\min_{w \in \mathcal{R}(\mathcal{RS}_{m_i}(X))} |w|$  goes to infinity as  $i$  goes to infinity, then for  $n \in \mathbb{N}^*$  and large enough  $i$ ,  $\mathcal{R}_i^{< n}$  is empty and therefore the topological rank of  $X$  can actually be bounded by  $|\mathcal{R}(\mathcal{RS}_{m_i}(X))| \leq |\mathcal{A}|^2 K_1^2$ .

## 6. COMPLEXITIES OF FINITE TOPOLOGICAL RANK $\mathcal{S}$ -ADIC SUBSHIFTS

In previous section we provide conditions on the complexity function of a minimal subshift (namely of non-superlinear complexity) to ensure finite topological rank (see Theorem 5.5). This section is devoted to study (partially) the converse. That is, the complexity function of a finite topological rank subshift and more generally of an  $\mathcal{S}$ -adic subshift. We introduce notions of relative complexity for the morphisms defining the subshift and derive upper bounds for the complexity function. We recover the well known result that finite topological rank Cantor systems have zero entropy (see *e.g.* [BH94]) even if for an arbitrary subexponential

function  $g: \mathbb{N} \rightarrow \mathbb{N}$  there are examples of finite topological rank subshifts whose complexity function is not dominated by  $g$  [DDMP16]. In addition, we give several sufficient conditions guaranteeing that a finite topological rank subshift has non-superlinear complexity and show that subshifts with alphabet rank equal to 2 have non-superquadratic complexity. Next we propose bounds for the complexity function in terms of combinatorial complexities. Thus we deduce a result related to topological orbit equivalence: any strong orbit equivalence class of a finite topological rank minimal subshift contains a subshift of sub-linear complexity (Corollary 6.14). Finally in Subsection 6.4 we provide a subshift of topological rank 2 with superlinear complexity, showing that the converse of Theorem 5.5 does not hold.

**6.1. Relative complexity of morphisms.** For a morphism  $\tau: \mathcal{A}^* \rightarrow \mathcal{B}^*$  set

$$\|\tau\| = \max_{a \in \mathcal{A}} |\tau(a)| \text{ and } \langle \tau \rangle = \min_{a \in \mathcal{A}} |\tau(a)|.$$

**Definition 6.1.** Let  $\sigma: \mathcal{B}^* \rightarrow \mathcal{C}^*$  be a morphism. For  $b \in \mathcal{B}$  and  $\mathcal{L} \subseteq \mathcal{B}^*$ , define the following subset of words that extends  $b$  to the right.

$$\mathcal{F}(b, n, \sigma, \mathcal{L}) = \{w \in \mathcal{L} : bw \in \mathcal{L}, w_1 \neq b, |\sigma(w_{[1, |w|-1]})| < n \leq |\sigma(w)|\}.$$

That is, each word in  $\mathcal{F}(b, n, \sigma, \mathcal{L})$  extends to the right some appearance of  $b$  in  $\mathcal{L}$ , starts with a letter different from  $b$  and has the shortest image under  $\sigma$  of length at least  $n$ .

**Definition 6.2.** Let  $\tau: \mathcal{A}^* \rightarrow \mathcal{B}^*$  and  $\sigma: \mathcal{B}^* \rightarrow \mathcal{C}^*$  be two morphisms. For  $i \in \mathbb{N}^*$ , we define the  $i$ -th complexity map of  $\sigma$  with respect to  $\tau$  as

$$P_{\sigma|\tau}^{(i)}(n) = \sum_{b \in \mathcal{B}} |\mathcal{F}(b, n, \sigma, \tau(\mathcal{A}^i))|.$$

The complexity map of  $\sigma$  with respect to  $\tau$  is defined by

$$P_{\sigma|\tau}(n) = \sum_{b \in \mathcal{B}} |\mathcal{F}(b, n, \sigma, \mathcal{L}(\tau(\mathcal{A}^{\mathbb{Z}})))|.$$

Note that  $P_{\sigma|\tau}^{(i)}(n) \leq P_{\sigma|\tau}^{(i+1)}(n) \leq P_{\sigma|\tau}(n)$  for all  $i, n \in \mathbb{N}^*$ . Also, observe that  $(P_{\sigma|\tau}(n))_n$  is increasing, whereas  $(P_{\sigma|\tau}^{(i)}(n))_n$  is not. Actually it is zero for all large enough  $n \in \mathbb{N}$ . We summarize some elementary properties of  $P_{\sigma|\tau}(\cdot)$  for later use.

**Lemma 6.3.** For morphisms  $\tau: \mathcal{A}^* \rightarrow \mathcal{B}^*$  and  $\sigma: \mathcal{B}^* \rightarrow \mathcal{C}^*$  the following properties hold.

- (1)  $P_{\sigma|\tau}(\|\sigma\|) \leq |\mathcal{B}|^{\frac{\|\sigma\|}{\langle \sigma \rangle} + 1}$ .
- (2) If  $|\mathcal{B}| = 2$ , then  $P_{\sigma|\tau}(\|\sigma\|) \leq \frac{\|\sigma\|}{\langle \sigma \rangle} + 1$ .

*Proof.* (1) It suffices to observe that for any  $b \in \mathcal{B}$  there are at most  $|\mathcal{B}|^{\frac{\|\sigma\|}{\langle \sigma \rangle}}$  words  $w$  in  $\mathcal{F}(b, \|\sigma\|, \sigma, \mathcal{L}(\tau(\mathcal{A}^{\mathbb{Z}})))$ .

(2) Write  $\mathcal{B} = \{b_1, b_2\}$  with  $|\sigma(b_1)| = \|\sigma\|$  and  $|\sigma(b_2)| = \langle \sigma \rangle$ . Note that if  $w$  belongs to  $\mathcal{F}(b_1, \|\sigma\|, \sigma, \mathcal{L}(\tau(\mathcal{A}^{\mathbb{Z}})))$ , then  $w$  contains at most one  $b_1$  as a letter. Hence  $|\mathcal{F}(b_1, \|\sigma\|, \sigma, \mathcal{L}(\tau(\mathcal{A}^{\mathbb{Z}})))| \leq \frac{\|\sigma\|}{\langle \sigma \rangle}$ . On the other hand,  $\mathcal{F}(b_2, \|\sigma\|, \sigma, \mathcal{L}(\tau(\mathcal{A}^{\mathbb{Z}}))) = \{b_1\}$ .

We conclude that  $P_{\sigma|\tau}(\|\sigma\|) \leq \frac{\|\sigma\|}{\langle \sigma \rangle} + 1$ .  $\square$

**Proposition 6.4.** *Let  $\tau: \mathcal{A}^* \rightarrow \mathcal{B}^*$  and  $\sigma: \mathcal{B}^* \rightarrow \mathcal{C}^*$  be two morphisms and suppose that  $\tau$  is positive. Then we have that*

$$p_{\sigma \circ \tau(\mathcal{A}^{\mathbb{Z}})}(n) \leq (|\mathcal{B}| + P_{\sigma|_{\tau}}^{(2)}(n)) \cdot n$$

for all  $n \in [\|\sigma\|, \langle \sigma \circ \tau \rangle]$ .

*Proof.* By definition, any word in  $\sigma \circ \tau(\mathcal{A}^{\mathbb{Z}})$  of length  $n \in [\|\sigma\|, \langle \sigma \circ \tau \rangle]$  is a subword of  $\sigma(b_{i_1})^{l_1} \dots \sigma(b_{i_k})^{l_k}$ , where  $b_{i_1}^{l_1} \dots b_{i_k}^{l_k} \in \tau(\mathcal{A}^2)$ ,  $b_{i_j} \in \mathcal{B}$  and  $l_j \in \mathbb{N}^*$ . We may consider two disjoint types of words of length  $n$ .

**Type 1:** Words of length  $n$  that start and end in a same block  $\sigma(b)^l$  for some  $b \in \mathcal{B}$  and  $l \in \mathbb{N}$ . For each  $b \in \mathcal{B}$ , such words are no more than  $|\sigma(b)|$  and therefore there are no more than  $|\mathcal{B}|\|\sigma\|$  words of this type.

**Type 2:** Words of length  $n$  that start in a block  $\sigma(b)^l$  and end in another one. Any such word is a subword of  $\sigma(b^k w)$ , where  $k = \lfloor n/|\sigma(b)| \rfloor$  and  $w \in \mathcal{F}(b, n, \sigma, \tau(\mathcal{A}^2))$ . Noting that there are at most  $n$  subwords of length  $n$  in  $\sigma(b^k w)$  starting in the block  $\sigma(b^k)$  and ending in  $\sigma(w)$ , we conclude that words of **Type 2** are no more than

$$\sum_{b \in \mathcal{B}} n |\mathcal{F}(b, \sigma, n, \tau(\mathcal{A}^2))| = P_{\sigma|_{\tau}}^{(2)}(n) \cdot n$$

Combining the two types of words provides the result.  $\square$

Let us illustrate our approach by the following corollary that, even if not stated, can be deduced from [BH94]. A straightforward consequence is the folklore result asserting that finite topological rank systems have zero entropy.

**Corollary 6.5.** *Let  $(X, S)$  be an  $\mathcal{S}$ -adic subshift generated by the positive directed sequence  $\tau = (\tau_n: \mathcal{A}_{n+1}^* \rightarrow \mathcal{A}_n^*)_{n \geq 0}$ . If  $\liminf_{n \rightarrow +\infty} \frac{\log(|\mathcal{A}_{n+1}|)}{\langle \tau_{[0, n]} \rangle} = 0$ , then  $(X, S)$  has zero entropy, i.e.,  $\lim_{n \rightarrow 1} \log(p_X(n))/n = 0$ .*

*In particular, if  $(X, S)$  is a minimal Cantor system of finite topological rank, then it has zero entropy.*

*Proof.* Let  $\tau = (\tau_n: \mathcal{A}_{n+1}^* \rightarrow \mathcal{A}_n^*)_{n \geq 0}$  be a positive directed sequence of morphisms generating  $(X, S)$ . By Lemma 6.3-(1) we have that  $P_{\tau_{[0, n]}|_{\tau_{n+1}}}(\|\tau_{[0, n]}\|) \leq |\mathcal{A}_{n+1}|^{\frac{\|\tau_{[0, n]}\|}{\langle \tau_{[0, n]} \rangle} + 1}$ . Proposition 6.4 applied to the morphisms  $\tau_{[0, n]}$  and  $\tau_{n+1}$  implies that

$$\frac{\log(p_{\tau_{[0, n+1]}(\mathcal{A}_{n+2}^{\mathbb{Z}})}(\|\tau_{[0, n]}\|))}{\|\tau_{[0, n]}\|} \leq \frac{\log(2|\mathcal{A}_{n+1}|\|\tau_{[0, n]}\|)}{\|\tau_{[0, n]}\|} + \frac{\|\tau_{[0, n]}\|}{\|\tau_{[0, n]}\|\langle \tau_{[0, n]} \rangle} \log(|\mathcal{A}_{n+1}|),$$

which allows to conclude the first part of the statement. The second part can be deduced from Theorem 4.1 and Theorem 2.1.  $\square$

As another application, we deduce some facts about the complexity of  $\mathcal{S}$ -adic subshifts with an alphabet rank equal to 2.

**Corollary 6.6.** *Let  $(X, S)$  be the  $\mathcal{S}$ -adic subshift generated by the positive directive sequence  $\tau = (\tau_n: \mathcal{A}_{n+1}^* \rightarrow \mathcal{A}_n^*)_{n \geq 0}$ . If  $\liminf_{n \rightarrow +\infty} |\mathcal{A}_n| \leq 2$ , then  $(X, S)$  is sub-quadratic along a subsequence, i.e.,*

$$\liminf_{n \rightarrow +\infty} \frac{p_X(n)}{n^2} = 0.$$

*Proof.* Similarly to the previous proof, for  $n \in \mathbb{N}$  consider the morphisms  $\sigma = \tau_{[0,n]}$  and  $\tau_{n+1}$  and apply Proposition 6.4 and Lemma 6.3-(2) to obtain, for infinitely many  $n$ ,

$$\begin{aligned} p_X(\|\tau_{[0,n]}\|) &\leq p_{\tau_{[0,n]} \circ \tau_{n+1}}(\mathcal{A}_{n+2}^{\mathbb{Z}})(\|\tau_{[0,n]}\|), \\ &\leq (2 + P_{\tau_{[0,n]}|\tau_{n+1}}^{(2)}(\|\tau_{[0,n]}\|))\|\tau_{[0,n]}\|, \\ &\leq \left( \frac{\|\tau_{[0,n]}\|}{\langle \tau_{[0,n]} \rangle} + 3 \right) \|\tau_{[0,n]}\|. \end{aligned}$$

Dividing by  $\|\tau_{[0,n]}\|^2$  and taking the liminf in  $n$  give the conclusion.  $\square$

Concerning the growth rate of the complexity function of an  $\mathcal{S}$ -adic subshift of alphabet rank 2, Corollary 6.5 shows it is sub-exponential. Beyond this, examples in [DDMP16, Section 4] show that there are no more precise bounds for the growth rate of the complexity function. Actually, for any sub-exponential function  $\varphi: \mathbb{N} \rightarrow \mathbb{R}$  (i.e.,  $\limsup_{n \rightarrow +\infty} \frac{\varphi(n)}{\alpha^n} = 0$  for any  $\alpha > 1$ ), there are examples of  $\mathcal{S}$ -adic subshifts with alphabet rank equal to 2 with  $\limsup_{n \rightarrow +\infty} \frac{p_X(n)}{\varphi(n)} > 0$ .

**6.2.  $\mathcal{S}$ -adic subshifts with non-superlinear complexity.** We provide several criteria that ensure that an  $\mathcal{S}$ -adic subshift has a non-superlinear complexity.

**Corollary 6.7.** *Let  $(X, S)$  be the  $\mathcal{S}$ -adic subshift generated by the positive directive sequence  $\tau = (\tau_n: \mathcal{A}_{n+1}^* \rightarrow \mathcal{A}_n^*)_{n \geq 0}$ . Assume that  $\tau$  has the proportional towers property, i.e.,  $K = \liminf_{n \rightarrow +\infty} \frac{\|\tau_{[0,n]}\|}{\langle \tau_{[0,n]} \rangle} < +\infty$  and that  $(|\mathcal{A}_n|)_{n \geq 0}$  is bounded. Then we have that*

$$\liminf_{n \rightarrow +\infty} \frac{p_X(n)}{n} < +\infty.$$

*Proof.* Let  $(n_i)_{i \in \mathbb{N}}$  be a sequence in  $\mathbb{N}$  such that  $\frac{\|\tau_{[0,n_i]}\|}{\langle \tau_{[0,n_i]} \rangle} \leq K$  for all  $i \in \mathbb{N}$ . Then by Proposition 6.4 and Lemma 6.3-(1) we have that

$$\begin{aligned} p_X(\|\tau_{[0,n_i]}\|) &\leq p_{\tau_{[0,n_i]} \circ \tau_{n_i+1}}(\mathcal{A}_{n_i+2}^{\mathbb{Z}})(\|\tau_{[0,n_i]}\|), \\ &\leq (|\mathcal{A}_{n_i+1}|^{\frac{\|\tau_{[0,n_i]}\|}{\langle \tau_{[0,n_i]} \rangle} + 1} + |\mathcal{A}_{n_i+1}|)\|\tau_{[0,n_i]}\|, \\ &\leq (\max_{i \in \mathbb{N}} |\mathcal{A}_i|^{K+1} + \max_{i \in \mathbb{N}} |\mathcal{A}_i|)\|\tau_{[0,n_i]}\|, \end{aligned}$$

from where the result follows.  $\square$

In terms of clopen Kakutani-Rohlin partitions, as a direct corollary of Corollary 6.7 we obtain.

**Corollary 6.8.** *Let  $(X, T)$  be an expansive minimal Cantor system with a sequence of CKR partitions  $(\mathcal{T}_n)_{n \in \mathbb{N}}$  satisfying (KR1)-(KR4),  $\mathcal{T}_n$  having  $d_n$  towers and heights  $(h_n(k))_{1 \leq k \leq d_n}$  for each  $n \in \mathbb{N}$ . Suppose they satisfy  $\sup_{n \in \mathbb{N}} d_n < +\infty$  and  $\liminf_{n \rightarrow +\infty} \sup_{1 \leq k, k' \leq d_n} h_n(k)/h_n(k') < +\infty$ . Then the system  $(X, T)$  is topologically conjugate to a minimal subshift with a non-superlinear complexity.*

The preceding corollary applies to finite topological rank subshifts that admit tower partitions where all the towers of the same level have the same height (this is called the *equal row sum* property, or ERS for short). The class of subshifts having



the ERS property coincides with the one of Toeplitz subshifts [GJ00, Theorem 8]. Although finite topological rank Toeplitz subshifts have the ERS property, we do not know if they have a Bratteli-Vershik representation having simultaneously the ERS property and a bounded number of towers at each level. So we ask if within the class of Toeplitz subshifts, to have finite topological rank and non-superlinear complexity are equivalent conditions.

The following lemma gives a condition to deduce the hypothesis (hence the conclusion) of Corollary 6.7 by analyzing the adjacency matrices associated to the morphisms of the given directive sequence. Before stating it, we introduce and recall the following definitions. For a letter  $a \in \mathcal{A}$ ,  $|w|_a$  denotes the number of occurrences of the letter  $a$  in  $w$ , and  $\ell(w)$  is the vector  $(|w|_a)_{a \in \mathcal{A}}$ . For a morphism  $\tau: \mathcal{A}^* \rightarrow \mathcal{B}^*$ , its *incidence matrix* is  $M_\tau = (|\tau(a)|_b)_{b \in \mathcal{B}, a \in \mathcal{A}}$ . The following formulas are classical. For any word  $w \in \mathcal{A}^*$ ,  $\ell(\tau(w)) = M_\tau \ell(w)$  and for two morphisms  $\tau_0: \mathcal{A}_1^* \rightarrow \mathcal{A}_0^*$  and  $\tau_1: \mathcal{A}_2^* \rightarrow \mathcal{A}_1^*$ ,  $M_{\tau_0 \circ \tau_1} = M_{\tau_0} M_{\tau_1}$ . For a positive matrix  $M = (m_{i,j})_{(i,j) \in I \times J}$ , we set  $D(M) = \sup_{i \in I} \sup_{j,k \in J} \frac{m_{i,j}}{m_{i,k}}$ . Note that a bounded sequence of positive matrices  $(M_n)_{n \geq 0}$  leads to a bounded sequence  $(D(M_n))_{n \geq 0}$ . However the converse is not true, for instance consider the matrices  $M_n = \begin{pmatrix} n & n+1 \\ n^2 & (n+1)^2 \end{pmatrix}$ .

**Lemma 6.9.** *Let  $\tau = (\tau_n: \mathcal{A}_{n+1}^* \rightarrow \mathcal{A}_n^*)_{n \geq 0}$  be a sequence of positive and proper morphisms. Then, for any integer  $n \geq 0$  we have*

$$\frac{\|\tau_{[0,n]}\|}{\langle \tau_{[0,n]} \rangle} \leq D(M_n).$$

*In particular,  $\liminf_{n \rightarrow +\infty} \frac{\|\tau_{[0,n]}\|}{\langle \tau_{[0,n]} \rangle}$  is finite whenever  $\liminf_{n \rightarrow +\infty} D(M_n)$  is finite.*

*Proof.* Let  $M_n$  denotes the incidence matrix of the morphism  $\tau_n$ . For any letter  $a \in \mathcal{A}_{n+1}$ , notice that

$$\begin{aligned} |\tau_0 \circ \dots \circ \tau_n(a)| &= \langle (1, \dots, 1)^t, \ell(\tau_0 \circ \dots \circ \tau_n(a)) \rangle, \\ &= \langle (M_0 \cdots M_{n-1})^t (1, \dots, 1)^t, M_n \ell(a) \rangle, \end{aligned}$$

where  $\langle \cdot, \cdot \rangle$  denotes the usual scalar product. Let  $\{e_c : c \in \mathcal{A}_n\}$  be the canonical basis and let  $v_c > 0$ ,  $c \in \mathcal{A}_n$ , be a scalars such that  $(M_0 \cdots M_{n-1})^t (1, \dots, 1)^t = \sum_{c \in \mathcal{A}_n} v_c e_c$ . It follows that,

$$\langle (M_0 \cdots M_{n-1})^t (1, \dots, 1)^t, M_n \ell(a) \rangle = \sum_{c \in \mathcal{A}_n} v_c \langle e_c, M_n \ell(a) \rangle.$$

Since  $\langle e_c, M_n \ell(a) \rangle = |\tau_n(a)|_c$ , we get for any pair of letters  $a, b$

$$|\tau_0 \circ \dots \circ \tau_n(a)| \leq |\tau_0 \circ \dots \circ \tau_n(b)| \sup_{c \in \mathcal{A}_n} \sup_{i,j \in \mathcal{A}_{n+1}} \frac{|\tau_n(i)|_c}{|\tau_n(j)|_c}.$$

This shows the lemma.  $\square$

This lemma applies for instance to  $\mathcal{S}$ -adic subshifts given by morphisms  $(\tau_n)_{n \geq 0}$  defined on alphabets of bounded size and such that  $D(M_n)$  is bounded for infinitely many  $n \in \mathbb{N}$ .

**6.3. The repetition complexity.** The relative complexities  $P_{\sigma|\tau}^{(i)}(n)$  may not be easy to compute in explicit examples. We introduce a related notion that is easier to analyze and that provides an upper bound for  $P_{\sigma|\tau}^{(i)}(n)$  for values of  $n$  of interest.

Let  $\tau: \mathcal{A}^* \rightarrow \mathcal{B}^*$  be a morphism (we set  $\mathcal{B} = \{b_1, \dots, b_{|\mathcal{B}|}\}$ ). For  $a \in \mathcal{A}$ , write  $\tau(a) = b_{i_1}^{l_1} b_{i_2}^{l_2} \dots b_{i_{k(a)}}^{l_{k(a)}}$  for some  $i_1, \dots, i_{k(a)} \in \{1, \dots, |\mathcal{B}|\}$ , where  $b_{i_{j+1}} \neq b_{i_j}$  for all  $j = 1, \dots, k(a) - 1$ . The integer  $k(a)$  represents the number of times one needs to switch letters while writing  $\tau(a)$ , plus one. The repetition complexity of  $\tau$  is defined as

$$\text{r-comp}(\tau) = \sum_{a \in \mathcal{A}} k(a).$$

Note that if  $\tau$  is positive then  $k(a) \geq |\mathcal{B}|$  for all  $a \in \mathcal{A}$  and then  $\text{r-comp}(\tau) \geq |\mathcal{A}||\mathcal{B}|$ .

*Example:* We say that a morphism  $\tau: \mathcal{A}^* \rightarrow \mathcal{B}^*$  is *left to right* if there exists an ordering  $b_1, \dots, b_{|\mathcal{B}|}$  of the elements of  $\mathcal{B}$  such that  $\tau(a) = b_1^{l_1(a)} b_2^{l_2(a)} \dots b_{|\mathcal{B}|}^{l_{|\mathcal{B}|}(a)}$  for all  $a \in \mathcal{A}$  for some  $l_1(a), \dots, l_{|\mathcal{B}|}(a) \in \mathbb{N}^*$ . It is clear that if  $\tau$  is left to right then  $\text{r-comp}(\tau) = |\mathcal{A}||\mathcal{B}|$  and so a left to right morphism attains the minimum value of the repetition complexity for primitive morphisms.

**Lemma 6.10.** *Let  $\tau: \mathcal{A}^* \rightarrow \mathcal{B}^*$  and  $\sigma: \mathcal{B}^* \rightarrow \mathcal{C}^*$  be two morphisms. The following properties hold:*

- (1)  $P_{\sigma|\tau}^{(1)}(n) \leq \text{r-comp}(\tau)$ , for all  $n \in \mathbb{N}$ ;
- (2)  $P_{\sigma|\tau}^{(2)}(n) \leq |\mathcal{A}| \text{r-comp}(\tau) + P_{\sigma|\tau}^{(1)}(n) \leq (|\mathcal{A}| + 1) \text{r-comp}(\tau)$ , for all  $n \in \mathbb{N}$ .

*Proof.* By definition, for any word  $w \in \mathcal{F}(b, n, \sigma, \tau(\mathcal{A}))$ , the word  $bw$  appears in some  $\tau(a)$ ,  $a \in \mathcal{A}$ . We can associate to  $w$  an occurrence of  $b$  corresponding to the first letter of  $bw$  in  $\tau(a)$ . Clearly this association is injective and each such occurrence of  $b$  contributes by 1 to the value of  $k(a)$ . So

$$P_{\sigma|\tau}^{(1)}(n) = \sum_{b \in \mathcal{B}} |\mathcal{F}(b, n, \sigma, \tau(\mathcal{A}))| \leq \sum_{a \in \mathcal{A}} k(a) = \text{r-comp}(\tau).$$

This proves (1).

To prove (2), note that for any word  $w \in \mathcal{F}(b, n, \sigma, \tau(\mathcal{A}^2))$ , we have that  $bw$  is a subword of  $\tau(\mathcal{A}^2)$ . So either  $bw$  is a subword of  $\tau(a)$  for some  $a \in \mathcal{A}$  or it starts in  $\tau(a)$  and ends in  $\tau(a')$  for some  $a, a' \in \mathcal{A}$ . Putting together all the words  $w$  in the first case gives at most  $P_{\sigma|\tau}^{(1)}(n)$  words. In the second case, given a block  $b^l$  in  $\tau(a)$ , and given  $a' \in \mathcal{A}$ , there is a unique word  $w \in \mathcal{F}(b, \sigma, n, \tau(\mathcal{A}^2))$  such that  $bw$  is a subword of  $\tau(aa')$  (here the  $b$  is the last one of the block  $b^l$ ). Hence, there are at most  $\sum_{a \in \mathcal{A}} k(a)|\mathcal{A}| = |\mathcal{A}| \text{r-comp}(\tau)$  words  $w$  in this case. Combining the two cases gives the desired bound.  $\square$

**Theorem 6.11.** *Let  $\tau: \mathcal{A}^* \rightarrow \mathcal{B}^*$ ,  $\sigma: \mathcal{B}^* \rightarrow \mathcal{C}^*$  and  $\phi: \mathcal{C}^* \rightarrow \mathcal{D}^*$  be three morphisms and assume that  $\tau$  and  $\sigma$  are positive. Then, we have that*

$$P_{\phi \circ \sigma \circ \tau}(\mathcal{A}^{\mathbb{Z}})(n) \leq \begin{cases} (P_{\phi|\sigma}^{(2)}(n) + |\mathcal{C}|)n & \text{if } n \in [|\phi|, \langle \phi \circ \sigma \rangle) \\ (P_{\phi \circ \sigma|\tau}^{(2)}(n) + P_{\phi|\sigma}^{(1)}(n) + |\mathcal{B}| + |\mathcal{C}|)n & \text{if } n \in [\langle \phi \circ \sigma \rangle, \|\phi \circ \sigma\|] \end{cases}$$

*Proof.* Writing  $\mathcal{B} = \{b_1, \dots, b_{|\mathcal{B}|}\}$  so that  $|\sigma \circ \tau(b_i)| \leq |\sigma \circ \tau(b_{i+1})|$  for  $i = 1, \dots, |\mathcal{B}| - 1$ , one obtains

$$[\|\phi\|, \|\phi \circ \sigma\|] = [\|\phi\|, |\phi \circ \sigma(b_1)|] \cup \bigcup_{j=1}^{|\mathcal{B}|-1} [|\phi \circ \sigma(b_j)|, |\phi \circ \sigma(b_{j+1})|].$$

Let us analyze the number of words of length  $n$  depending on the interval of the right hand side where  $n$  belongs to.

Let  $n$  be in  $[\|\phi\|, |\phi \circ \sigma(b_1)|]$ . Proposition 6.4 gives

$$p_{\phi \circ \sigma \circ \tau(\mathcal{A}^{\mathbb{Z}})}(n) \leq p_{\phi \circ \sigma(\mathcal{B}^{\mathbb{Z}})}(n) \leq (|\mathcal{C}| + P_{\phi|\sigma}^{(2)}(n))n.$$

Now let  $n$  be in  $[|\phi \circ \sigma(b_i)|, |\phi \circ \sigma(b_{i+1})|]$  for  $1 \leq i < |\mathcal{B}|$ . Note that a word of length  $n$  is a subword of  $\phi \circ \sigma(b_{i_1}^{l_1} \dots b_{i_k}^{l_k})$ , where  $b_{i_1}^{l_1} \dots b_{i_k}^{l_k}$  belongs to  $\tau(\mathcal{A}^2)$ . We consider disjoint types of words of length  $n$ :

**Type 1:** Words of length  $n$  that start and end in a block of the form  $\phi \circ \sigma(b)^l$  and end in another block  $\phi \circ \sigma(b')^{l'}$ . By the same argument as in Proposition 6.4, there are no more than  $P_{\phi \circ \sigma|\tau}^{(2)}(n) \cdot n$  of such words.

**Type 2:** Words of length  $n$  starting and ending in a same block  $(\phi \circ \sigma(b_j))^l$  for  $b_j \in \mathcal{B}$ , where  $b^l \in \tau(\mathcal{A}^{\mathbb{Z}})$ . We may further distinguish two disjoint types:

**Type 2.1:**  $j \leq i$  or equivalently  $n \geq |\phi \circ \sigma(b_j)|$ . In this case the number of words is at most  $|\phi \circ \sigma(b_j)|$ . Moreover, such words are the same for all the different concatenations of  $\phi \circ \sigma(b_j)$  among all the words  $\phi \circ \sigma \circ \tau(a)$ . It follows there are at most  $\sum_{b: |\phi \circ \sigma(b)| \leq n} |\phi \circ \sigma(b)| \leq in \leq |\mathcal{B}|n$  of such words.

**Type 2.2:**  $j > i$  or equivalently  $n < |\phi \circ \sigma(b_j)|$ . In this case we may write  $\sigma(b_j) = c_1^{l_1} \dots c_{e(b_j)}^{l_{e(b_j)}} \in \sigma(\mathcal{B})$  and then

$$\phi \circ \sigma(b_j) = \phi(c_1)^{l_1} \dots \phi(c_{e(b_j)})^{l_{e(b_j)}}.$$

We will use the same distinction of cases as before to bound the number of words, in this context, satisfying **Type 2.2**:

**Type 2.2.1:** The number of words in **Type 2.2** that start in some block  $\phi(c_k)^{l_k}$  and end in another one is bounded by  $n$ . There are at most  $P_{\phi|\sigma}^{(1)}(n) \cdot n$  words in this case.

**Type 2.2.2:** We bound the number of words starting and ending in a same block  $\phi(c_k)^{l_k}$ . As  $|\phi \circ \sigma(b_j)| > n \geq \|\phi\|$ , there are at most  $\|\phi\|$  of such words, whatever the power  $l_k$  is. Hence among all the letters  $c_k$  and  $b_j$  the number of such words is bounded by  $|\mathcal{C}|\|\phi\| \leq |\mathcal{C}|n$ .

Summing up all the bounds obtained above, we obtain the result.  $\square$

Combining Theorem 6.11 and Lemma 6.10 we get.

**Corollary 6.12.** *Let  $\tau: \mathcal{A}^* \rightarrow \mathcal{B}^*$ ,  $\sigma: \mathcal{B}^* \rightarrow \mathcal{C}^*$  and  $\phi: \mathcal{C}^* \rightarrow \mathcal{D}^*$  be three morphisms and assume that  $\tau$  and  $\sigma$  are positive. Then, we have that*

$$p_{\phi \circ \sigma \circ \tau(\mathcal{A}^{\mathbb{Z}})}(n) \leq (\max\{|\mathcal{A}|, |\mathcal{B}|, |\mathcal{C}|\} + 2)(\max\{\text{r-comp}(\tau), \text{r-comp}(\sigma)\} + 1) \cdot n$$

for all  $n \in [\|\phi\|, \|\phi \circ \sigma\|]$ .

As a direct consequence, we obtain the following.

**Corollary 6.13.** *Let  $(X, S)$  be the  $\mathcal{S}$ -adic subshift generated by the positive directive sequence  $\tau = (\tau_n: \mathcal{A}_{n+1}^* \rightarrow \mathcal{A}_n^*)_{n \geq 0}$ . Then, we have that*

$$p_X(n) \leq (\max_{i \in \mathbb{N}} |\mathcal{A}_i| + 2)(\limsup_{i \rightarrow +\infty} \text{r-comp}(\tau_i) + 1) \cdot n$$

for all large enough  $n \in \mathbb{N}$ .

*Proof.* Let  $K \in \mathbb{N}$  be such that  $\text{r-comp}(\tau_n) \leq \limsup_{i \rightarrow +\infty} \text{r-comp}(\tau_i)$  for all  $n \geq K$ . For  $n \in \mathbb{N}$ , let  $M \in \mathbb{N}$  be such that  $n \in [|\tau_{[0,M)}|, |\tau_{[0,M+1)}|]$ . If  $n$  is large enough, then as  $M$  is greater than  $K$ , Corollary 6.12 applied to  $\phi = \tau_{[0,M)}$ ,  $\sigma = \tau_M$  and  $\tau = \tau_{M+1}$  provides

$$p_X(n) \leq (\max_{i \in \mathbb{N}} |\mathcal{A}_i| + 2)(\limsup_{i \rightarrow +\infty} \text{r-comp}(\tau_i) + 1) \cdot n$$

which finishes the proof.  $\square$

The next result is an application of the bounds obtained above to topological orbit equivalence theory.

**Corollary 6.14.** *Let  $(X, S)$  be a finite topological rank minimal Cantor system. Then  $(X, S)$  is strongly orbit equivalent to a subshift of sub-linear complexity.*

*Proof.* We assume  $X$  is infinite, the finite case follows directly. Recall from [GPS95] that minimal Cantor systems are strong orbit equivalent whenever they can be represented using the same (unordered) Bratteli diagram (up to telescoping). Let  $(V, E \preceq)$  be a Bratteli-Vershik representation of  $(X, S)$ . Let us consider the (unordered) Bratteli diagram  $B = (V, E)$ . After telescoping and microscoping, we may assume that  $|V_n| \geq 3$  for all  $n \geq 1$ . We assign an ordering to  $B$ . To do this we consider morphisms  $\tau_n: V_{n+1}^* \rightarrow V_n^*$ ,  $n \geq 1$ , of the form

$$v \mapsto w_1^{l_1(v)} w_2^{l_2(v)} w_1^{l'_1(v)} w_2^{l'_2(v)} w_3^{l_3(v)} \dots w_{|V_n|}^{l_{|V_n|}(v)},$$

for all  $v \in V_{n+1}$ , where  $w_1, \dots, w_{|V_n|}$  is an ordering of  $V_n$  (and for  $n = 0$  just consider a left to right morphism  $\tau_0: V_1^* \rightarrow E_1^*$ ). It is proved in [GJ00, Lemma 10] that it is possible to adjust the lengths  $l_1, l'_1, l_2, l'_2$  to obtain an expansive system (there it is also required that the minimum number of edges connecting a vertex of level  $n$  with a vertex of level  $n + 1$  is at least  $6|V_n|$ , which can always be assumed after an appropriate telescoping). It is clear that the morphisms above are proper, recognizable and positive. Proposition 4.5 tells us that the Bratteli-Vershik system associated to these morphisms is topologically conjugate to the corresponding  $\mathcal{S}$ -adic subshift. It is straightforward to compute that the repetition complexity of  $\tau_n$  is equal to  $|V_{n+1}|(|V_n| + 2)$  (which is bounded in  $n$ ). Finally, Corollary 6.13 allows us to conclude that the complexity of the  $\mathcal{S}$ -adic subshift generated by  $(\tau_n)_{n \geq 0}$  is sub-linear.  $\square$

**6.4. Example: a rank 2 system with a superlinear complexity.** In this section we construct a subshift  $(X, S)$  such that  $\liminf_{n \rightarrow +\infty} p_X(n)/n = +\infty$  and whose topological rank is 2. Hence we show that finite topological rank does not imply the non-superlinear complexity property. Recall that by Theorem 5.5 non-superlinear complexity does imply finite topological rank.

We fix some increasing sequence of integers  $(a_n)_{n \geq 0}$  (specified later) and we consider the sequence of morphisms  $\tau = (\tau_n: \{0, 1\}^* \rightarrow \{0, 1\}^*)_{n \geq 0}$  defined by

$$\tau_n(0) = 011 \text{ and } \tau_n(1) = 0^2 10^3 1 \cdots 0^{a_n} 1.$$

Let  $(X_\tau, S)$  be the  $\mathcal{S}$ -adic subshift generated by  $\tau$ . Set  $A_n = |\tau_{[0, n]}(1)|$ ,  $B_n = |\tau_{[0, n]}(0)|$  and  $C_n = B_{n-1} + \cdots + B_1 + 1$ .

To show that  $\liminf_{n \rightarrow +\infty} p_{X_\tau}(n)/n = +\infty$  we indeed show the stronger condition  $\lim_{n \rightarrow +\infty} p_{X_\tau}(n+1) - p_{X_\tau}(n) = +\infty$ .

Recall that  $X_\tau^{(n)}$  denotes the subshift generated by  $(\tau_k)_{k \geq n}$ . We state some useful observations in the following lemmas. The details are left to the interested reader.

**Lemma 6.15.** *For all  $n \geq 1$ , the set of prefixes of length  $C_n + 1$  of the words  $\tau_{[0, n]}(0)$  and  $\tau_{[0, n]}(1)$  is  $\{\tau_{[0, n-1]}(0)\tau_{[0, n-2]}(0) \cdots \tau_0(0)0a : a \in \{0, 1\}\}$ .*

*Proof.* It follows by induction after noting that 00 and 01 belong to  $X_\tau^{(n+1)}$ .  $\square$

**Lemma 6.16.** *The directive sequence  $\tau$  is recognizable.*

*Proof.* The proof is left to the reader.  $\square$

**Lemma 6.17.** *For all  $n \geq 1$  and  $k \in [1, a_{n+1} - 1]$ , the words*

$$W_n(k) = \tau_{[0, n]}(0^{k-1}10^k)\tau_{[0, n-1]}(0) \cdots \tau_0(0)0$$

*are right special words in  $X_\tau^{(n)}$ .*

*Proof.* For all  $n \geq 1$  and  $k \in [1, a_{n+1} - 1]$ , the words  $0^{k-1}10^k0$  and  $0^{k-1}10^k1$  belong to  $\mathcal{L}(X_\tau^{(n+1)})$ . Hence,  $\tau_{[0, n]}(0^{k-1}10^k0)$  and  $\tau_{[0, n]}(0^{k-1}10^k1)$  belong to  $\mathcal{L}(X_\tau^{(n)})$ . Lemma 6.15 allows us to conclude.  $\square$

**Lemma 6.18.** *For all  $n \geq 1$ ,  $k \in [1, a_{n+1} - 1]$  and  $N \in [kB_n + C_n, (k+1)B_n + C_n]$  there are at least  $\min(k, A_n/2B_n)$  different right special words of length  $N$ .*

*Proof.* Let  $n \geq 1$ ,  $k \in [1, a_{n+1} - 1]$  and  $N \in [kB_n + C_n, (k+1)B_n + C_n]$ . For each  $i \in [k+1 - \min(k, A_n/2B_n), k]$  let  $w_i$  be the suffix of length  $N$  of  $W_n(i)$ . It is well-defined as the length of  $W_n(i)$  is greater than  $(k+1)B_n + C_n$ . One has,

$$(3) \quad w_i = p_i \tau_{[0, n]}(0^i) \tau_{[0, n-1]}(0) \cdots \tau_1(0)0,$$

where  $p_i$  is a suffix of  $\tau_{[0, n]}(0^{i-1}1)$ . From Lemma 6.17 these words are right special. It remains to show they are distinct. Let  $i$  and  $j$  be distinct integers belonging to the interval  $[k+1 - \min(k, A_n/2B_n), k]$ . As  $\tau_{[0, n]}(0)$  has a length smaller than the one of  $\tau_{[0, n]}(1)$  and is not a suffix of it, it is easy to check that  $w_i$  and  $w_j$  are distinct (see (3)).  $\square$

We have all the ingredients to construct a subshift of rank 2 with superlinear but subquadratic complexity.

**Proposition 6.19.** *There exists a minimal subshift  $(X, S)$  of topological rank 2 such that  $\liminf_{n \rightarrow +\infty} p_X(n)/n = +\infty$  and  $\lim_{n \rightarrow +\infty} p_X(n)/n^2 = 0$ .*

*Proof.* Let  $(X_\tau, S)$  be the  $\mathcal{S}$ -adic subshift generated by the previous morphisms  $\tau = (\tau_n: \{0, 1\}^* \rightarrow \{0, 1\}^*)_{n \geq 0}$ , where we choose an increasing integer valued sequence  $(a_n)_{n \geq 0}$  satisfying  $a_{n+1} > (3n+4)A_n/B_n$ .

The morphisms  $\tau_n$ ,  $n \geq 0$ , being positive, proper and defining a recognizable directive sequence (Lemma 6.16), generate a rank 2 minimal subshift  $(X_\tau, S)$ , thanks to Proposition 4.5.

To show that  $\liminf_{n \rightarrow +\infty} p_X(n)/n = +\infty$ , it is enough to prove that the number of right special factors goes to infinity. As  $\lim_{n \rightarrow +\infty} \min(n, A_n/2B_n) = +\infty$ , from Lemma 6.18 it suffices to show that the intervals  $[kB_n + C_n, (k+1)B_n + C_n)$  with  $n \leq k < a_{n+1}$ ,  $n \geq 0$ , cover a half line  $[N, +\infty)$  for some  $N$ . Thus we only need to check that  $(n+1)B_{n+1} + C_{n+1}$  is less than  $a_{n+1}B_n + C_n$ :

$$\begin{aligned} (n+1)B_{n+1} + C_{n+1} &\leq (n+2)B_n + 2(n+1)A_n + C_n \leq (3n+4)A_n + C_n \\ &< a_{n+1}B_n + C_n. \end{aligned}$$

This proves that  $p_X$  is superlinear.

For the subquadratic property we claim the following.

**Claim:** For all  $m, n \geq 0$  we have that  $P_{\tau_{[0,m]}|\tau_{m+1}}^{(2)}(n)$  is bounded by  $6n/B_m$ .

To see this, note that a word  $w$  in  $\mathcal{F}(\ell, \tau_{[0,m]}, n, \tau_{m+1}(\mathcal{A}_{m+2}^2))$ ,  $\ell = 0, 1$ , is either a subword of  $\tau_{m+1}(a)$ ,  $a \in \{0, 1\}$ , or it starts in  $\tau_{m+1}(a)$  and ends in  $\tau_{m+1}(a')$  for  $a, a' \in \{0, 1\}$ .

For the first type, any word  $w$  in  $\mathcal{F}(1, \tau_{[0,m]}, n, \tau_{m+1}(\mathcal{A}_{m+2}))$  is completely determined by its first occurrence of 1. Similarly for  $\ell = 0$ , any word  $w$  in the set  $\mathcal{F}(0, \tau_{[0,m]}, n, \tau_{m+1}(\mathcal{A}_{m+2}))$  is completely determined by its second occurrence of 1. Since  $|\tau_{[0,m]}(0)| = B_m$ , there are no more than  $n/B_m$  words  $w$  of this type in each case, for  $\ell = 0$  or 1.

For the second type, for any  $a, a' \in \{0, 1\}$  there are at most  $n/B_m$  words starting in  $a$  and ending in  $a'$ , since  $B_m = \langle \tau_{[0,m]} \rangle$ . Hence there are at most  $2^2 n/B_m$  of such words. This gives the conclusion of the Claim.

For  $n \in \mathbb{N}^*$ , let  $M = M(n) \in \mathbb{N}$  such that  $n \in [\|\tau_{[0,M]}\|, \|\tau_{[0,M+1]}\|]$ . Theorem 6.11 applied to  $\phi = \tau_{[0,M]}$ ,  $\sigma = \tau_M$  and  $\tau = \tau_{M+1}$ , together with the Claim, provide

$$p_X(n) \leq \left( 6 \frac{n}{B_{M(n)}} + 6 \frac{n}{B_{M(n)-1}} + 4 \right) \cdot n.$$

Dividing both sides by  $n^2$  enables us to conclude.  $\square$

## 7. ASYMPTOTIC COMPONENTS IN SYSTEMS OF TOPOLOGICAL RANK 2

This section is devoted to highlight a rigidity property in minimal Cantor systems of topological rank 2. For these systems and minimal  $\mathcal{S}$ -adic subshifts of alphabet rank equal to 2 we analyse a specific dynamical invariant, namely, their asymptotic components.

We show that minimal  $\mathcal{S}$ -adic subshifts on two letters have at most two asymptotic components and that a non-periodic minimal Cantor system of topological rank 2 has only one asymptotic component (Corollary 7.4). In particular, the example in Section 6.4 shows it exist non-periodic minimal subshifts on two letters with only one asymptotic component but at the same time having superlinear complexity. A relevant application of these results concerns the automorphisms group of a subshift, *i.e.*, the group of self-homeomorphisms of the subshift commuting with the shift map. A straightforward consequence of [DDMP16, Theorem 3.1] for minimal subshifts of topological rank equal to 2 is that any automorphism is a power of the shift map. The same result provides for a minimal  $\mathcal{S}$ -adic subshift

on two letters that the group generated by the shift map is a subgroup of index at most 2 within the one generated by the automorphisms of the subshift. To be more complete in this topic, let us recall that using the strategy in [CK15] and [QZ04, Lemma 5], it is possible to show that non-superquadratic complexity subshifts are *coalescent*, *i.e.*, any endomorphism of the subshift is invertible (see [DDMP17, Section 3] for details). Thus Corollary 6.6 provides that any  $\mathcal{S}$ -adic subshift on two letters is coalescent.

We recall some classical terminology in topological dynamics.

Let  $(X, T)$  be a topological dynamical system, where  $X$  is endowed with a metric  $\text{dist}(\cdot, \cdot)$ . Two points  $x, y \in X$  are said to be (left)<sup>1</sup> *asymptotic* whenever  $\lim_{n \rightarrow +\infty} \text{dist}(T^{-n}x, T^{-n}y) = 0$ . It is well known that an aperiodic subshift always admits (left) asymptotic points  $x, y$  where  $x \neq y$  (such pairs are also known as nontrivial asymptotic pairs), see [Aus88, Chapter 1]. For simplicity, in what follows we use the word asymptotic to mean left asymptotic.

Given  $x, y \in X$  we say that the orbits  $\text{Orb}_T(x)$  and  $\text{Orb}_T(y)$  are *asymptotic* if there exist points  $x' \in \text{Orb}_T(x)$  and  $y' \in \text{Orb}_T(y)$  that are asymptotic. This condition is equivalent to say that  $y$  is asymptotic to some  $T^n x$  or vice versa. Then, for each  $x' \in \text{Orb}_T(x)$  there is a point  $y' \in \text{Orb}_T(y)$  asymptotic to  $x'$ . It is clear that this relation is an equivalence relation on the collection of orbits. We call a non trivial equivalence class, (*i.e.*, a class not reduced to a point) an *asymptotic component*.

We start with a lemma, which is a weaker version of the famous Fine and Wilf theorem [Lot02, Section 8].

**Lemma 7.1.** *Let  $A$  and  $B$  be different non-empty words. Then, there exists a word  $u$  of length strictly less than  $|A| + |B|$  such that for any pair of different sequences  $x = x_0 \cdots x_L \in A\{A, B\}^*$  and  $y = y_0 \cdots y_M \in B\{A, B\}^*$ ,  $u$  is a common prefix of  $x$  and  $y$  and, when  $|u|$  is not greater than  $L$  and  $M$ ,  $x_{|u|} \neq y_{|u|}$ .*

*Proof.* We prove the lemma by induction on the length  $||B| - |A||$ . Assume first that  $|A| = |B|$ , so  $||B| - |A|| = 0$ . Since  $A \neq B$ ,  $u$  can be taken as the maximal common prefix of  $A$  and  $B$ , which can be the empty word.

Let us assume the lemma is true for all words  $A$  and  $B$  such that  $||B| - |A|| \leq n$  for all  $n \geq 0$ . Let  $A$  and  $B$  be two words such that  $||B| - |A|| = n + 1$ . Without loss of generality we may assume that  $|A| < |B|$ . If  $A$  is not a prefix of  $B$ , take  $u$  as the maximal common prefix of  $A$  and  $B$ . Otherwise,  $B$  has the form  $AB'$  for some non-empty word  $B'$ .

Let  $x$  and  $y$  be two sequences such that  $x = x_0 \cdots x_L \in A\{A, B\}^*$  and  $y = y_0 \cdots y_M \in B\{A, B\}^*$ . Changing  $B$  by  $AB'$ , we get that the sequence  $x' = x_{[|A|, L]}$ , if not empty, belongs to  $A\{A, B'\}^*$  and  $y' = y_{[|A|, M]}$  belongs to  $B'\{A, B'\}^*$ . The induction hypothesis applied to  $A$  and  $B'$  implies that it exists a word  $u'$  of length strictly less than  $|A| + |B'| = |B|$ , independent of  $x$  and  $y$ , which is a common prefix of  $x'$  and  $y'$  and  $x_{|A|+|u'|} \neq y_{|A|+|u'|}$ . The result follows by considering the word  $u = Au'$ .  $\square$

We recall the notion of recognizability of a family of words with respect to a subshift as defined in Section 3. Let  $\mathcal{W} \subset \mathcal{B}^*$  be a set of non-empty words and  $X \subseteq \mathcal{B}^{\mathbb{Z}}$  be a subshift. We say  $\mathcal{W}$  is recognizable in  $X$  if for a subshift  $Y \subseteq \mathcal{A}^{\mathbb{Z}}$  and a morphism  $\tau : \mathcal{A}^* \rightarrow \mathcal{B}^*$  with  $\tau(\mathcal{A}) = \mathcal{W}$  we have:  $X$  is the smallest subshift

<sup>1</sup>One can also consider *right* asymptotic pairs, but we do not need that here.

containing  $\tau(Y)$  and  $\tau$  is recognizable in  $Y$ , (i.e., every  $x \in X$  has a unique  $\mathcal{W}$ -factorization adapted to  $(Y, \tau)$ ).

The typical examples we will consider in this section are inspired by the following theorem that is a particular result of [BSTY19].

**Theorem 7.2** ([BSTY19, Theorem 4.6]). *Let  $\mathcal{A}$  be an alphabet with 2 letters. Let  $(X_\tau, S)$  be an aperiodic  $\mathcal{S}$ -adic subshift generated by the directive sequence  $\tau = (\tau_n : \mathcal{A}^* \rightarrow \mathcal{A}^*)_{n \geq 0}$ , with  $\tau_n$  non-erasing for all  $n \geq 0$ . Then  $\{\tau_{[0,n]}(a) : a \in \mathcal{A}\}$  is recognizable in  $X_\tau$  (with respect to  $(X_\tau^{(n)}, \tau_{n-1})$ ) for any  $n \geq 1$ .*

The next proposition particularly applies to Cantor systems with topological rank equal to 2 and to  $\mathcal{S}$ -adic subshifts obtained from recognizable morphisms on alphabets of 2 letters. For its proof we recall some notions and notations. Let  $\mathcal{W} \subset \mathcal{A}^*$  be a set of non-empty words and  $X \subseteq \mathcal{A}^\mathbb{Z}$  be a subshift. Recall that the cutting points of a  $\mathcal{W}$ -factorization  $F_x : \mathbb{Z} \rightarrow \mathcal{W} \times \mathbb{Z}$  of a point  $x \in X$  is the set

$$C(F_x) = \{i \in \mathbb{Z} : F_x(k) = (w, i) \text{ for some } k \in \mathbb{Z}, w \in \mathcal{W}\}.$$

We say the  $\mathcal{W}$ -factorizations  $F_x$  and  $F_{x'}$  of  $x$  and  $x'$  in  $X$  coincide on a set  $J \subseteq \mathbb{Z}$  if  $C(F_x) \cap J = C(F_{x'}) \cap J$ .

**Proposition 7.3.** *Let  $(A_n)_{n \geq 1}$  and  $(B_n)_{n \geq 1}$  be two sequences of non-empty words in  $\mathcal{A}^*$  with increasing length. Let  $X \subseteq \mathcal{A}^\mathbb{Z}$  be a subshift such that  $\{A_n, B_n\}$  is recognizable in  $X$  for all  $n \geq 1$ . Then,  $(X, S)$  admits at most 2 asymptotic components. Moreover, if the maximal common suffix of  $A_n$  and  $B_n$  has a length going to infinity with  $n$ , then  $(X, S)$  admits at most 1 asymptotic component.*

*Proof.* By hypothesis, for all  $n \geq 1$  there exist a morphism  $\tau_n : \mathcal{A}_n^* \rightarrow \mathcal{A}^*$  with  $\tau_n(\mathcal{A}_n) = \mathcal{W}_n = \{A_n, B_n\}$  and a subshift  $Y_n \subseteq \mathcal{A}_n^\mathbb{Z}$  such that  $X$  is the smallest subshift containing  $\tau_n(Y_n)$  and all  $x \in X$  has a unique  $\mathcal{W}_n$ -factorization adapted to  $(Y_n, \tau_n)$ . To avoid notations, in what follows the factorizations will be always adapted to the corresponding subshifts and morphisms, which will be clear from the context.

The proof is by contradiction. Let us assume that  $\{x^{(0)}, y^{(0)}\}$ ,  $\{x^{(1)}, y^{(1)}\}$ ,  $\{x^{(2)}, y^{(2)}\}$  are left asymptotic pairs representatives of three different asymptotic components of  $(X, S)$ . After shifting them, we can assume that  $x_{(-\infty, -1]}^{(j)} = y_{(-\infty, -1]}^{(j)}$  and  $x_0^{(j)} \neq y_0^{(j)}$  for each  $j \in \{0, 1, 2\}$ .

Fix  $n \in \mathbb{N}^*$  and  $j \in \{0, 1, 2\}$ . Since the distance between  $S^{-k}x^{(j)}$  and  $S^{-k}y^{(j)}$  goes to 0 as  $k$  goes to  $+\infty$ , then the fact that  $\mathcal{W}_n$  is recognizable in  $X$  is a local property (see Lemma 3.6) implies that there exists an integer  $\ell \leq 0$  such that the  $\mathcal{W}_n$ -factorizations of  $x^{(j)}$  and  $y^{(j)}$  coincide on the interval  $(-\infty, \ell) \cap \mathbb{Z}$ . Let  $\ell^{(j)}$  be the largest such integer and denote by  $F_j$  the  $\mathcal{W}_n$ -factorization of  $x^{(j)}$ . We set  $a^{(j)} = \max C(F_j) \cap (-\infty, \ell^{(j)})$ . In particular,  $S^{a^{(j)}}x^{(j)}$  and  $S^{a^{(j)}}y^{(j)}$  are in two different cylinders  $[A_n], [B_n]$ .

Let  $k_j$  be an integer and  $w^{(j)}$  be a word in  $\mathcal{W}_n$  such that  $F_j(k_j) = (w^{(j)}, a^{(j)} - |w^{(j)}|)$ . Then  $w^{(j)}$  is the word  $x_{[a^{(j)} - |w^{(j)}|, a^{(j)}]}^{(j)}$  before the cutting point  $a^{(j)}$ .

Let  $u_n$  be the word given by Lemma 7.1 associated to  $\mathcal{W}_n = \{A_n, B_n\}$ . The word  $u_n$  is the maximal common prefix of  $x_{[a^{(j)}, +\infty)}^{(j)}$  and  $y_{[a^{(j)}, +\infty)}^{(j)}$  and is independent of  $j$ . So its length  $|u_n|$  is equal to  $|a^{(j)}|$  for any  $j \in \{0, 1, 2\}$ . By the pigeonhole principle,



there are two indices  $j_1$  and  $j_2 \in \{0, 1, 2\}$  sharing the same word  $w^{(j_1)} = w^{(j_2)}$ . Hence the words  $x_{[-|u_n| - |w^{(j)}|, 0)}^{(j)} = w^{(j)}u_n$  are all the same for  $j \in \{j_1, j_2\}$ .

Repeating the same argument for each  $n \geq 1$  we obtain a sequence of pairs of indices  $\{j_1^{(n)}, j_2^{(n)}\}$  such that the word  $x_{[-|u_n| - |w^{(j)}|, 0)}^{(j)}$ , with  $w^{(j)}$  in  $\{A_n, B_n\}$ , is common to each  $j$  in  $\{j_1^{(n)}, j_2^{(n)}\}$ . Again the pigeonhole principle ensures that there is a pair of two indices  $\{j_1, j_2\}$  common to infinitely many integers  $n$ . Hence  $x_{(-\infty, 0)}^{(j_1)} = x_{(-\infty, 0)}^{(j_2)}$ . It follows that the sequences  $x^{(j_1)}$  and  $x^{(j_2)}$  are asymptotic, which is a contradiction.

In the case where the length  $\ell_n$  of the maximal common suffix of  $A_n$  and  $B_n$  goes to infinity with  $n$ , the words  $x_{[-|u_n| - \ell_n, 0)}^{(j)}$  are all equal for any  $j \in \{0, 1, 2\}$ . Making  $n$  converging to infinity allows to conclude that the sequences  $x^{(0)}$ ,  $x^{(1)}$  and  $x^{(2)}$  are asymptotic, proving that in this case there exists at most one asymptotic component.  $\square$

The next result is a direct application of Proposition 7.3 and Theorem 7.2. Recall that a topological rank 2 minimal Cantor system is always conjugate to an  $\mathcal{S}$ -adic subshift on two letters where the morphisms are proper (see Section 4).

**Corollary 7.4.** *Let  $\mathcal{A}$  be an alphabet with 2 letters. Suppose the  $\mathcal{S}$ -adic subshift  $(X_\tau, S)$  generated by the sequence  $\tau = (\tau_n: \mathcal{A}^* \rightarrow \mathcal{A}^*)_{n \geq 0}$  of non-erasing morphisms is aperiodic and minimal. Then,  $(X_\tau, S)$  has at most two asymptotic components.*

*Let  $(X, T)$  be a topological rank 2 minimal Cantor system, then it has a unique asymptotic component.*

As an example consider the Prouhet-Thue-Morse morphism (substitution)  $\tau$  defined by  $0 \mapsto 01$ ,  $1 \mapsto 10$ . The subshift generated by this substitution is an  $\mathcal{S}$ -adic subshift with only one substitution called *Prouhet-Thue-Morse subshift*. The sequences  $\lim_{n \rightarrow +\infty} \tau^{2n}(0)$ ,  $\lim_{n \rightarrow +\infty} \tau^{2n}(i)$  and  $\lim_{n \rightarrow +\infty} \tau^{2n}(1)$ ,  $\lim_{n \rightarrow +\infty} \tau^{2n}(i)$  are in the associated subshift and are asymptotic for each  $i \in \{0, 1\}$ . Hence, by Corollary 7.4, the subshift has exactly two asymptotic components but it is not of topological rank 2. Actually, using techniques of substitution theory on return words of [DHS99] and [DL12], a long but standard computation enables us to create an  $\mathcal{S}$ -adic subshift conjugate to the Prouhet-Thue-Morse subshift, with a proper directive sequence of alphabet rank 3. Thus its topological rank is not greater than 3, so that its topological rank is exactly 3.

## 8. FINAL REMARKS AND QUESTIONS

The bounds obtained on the topological rank (Theorem 4.1, Theorem 4.3 or Theorem 5.5) for a minimal subshift are far to be optimal. In addition, even for classical examples (*e.g.*, provided by substitutions), one observes that the computation of their topological rank is not an easy task. It makes us wonder about the calculability of the topological rank of a minimal subshift or about the alphabet rank of an  $\mathcal{S}$ -adic subshift. To be more precise, we may consider this problem for systems described by explicit algorithms, like substitutive subshifts or effective  $\mathcal{S}$ -adic subshifts. In this context, we ask

**Question 8.1.** *Do there exist algorithms that taking as input the data of a directive sequence of morphisms, compute the alphabet/topological rank of the corresponding  $\mathcal{S}$ -adic subshift ?*

A less ambitious objective is to obtain lower and upper bounds for the alphabet/topological rank. In this respect studying the complexity function of the subshift could provide partial answers. In light of Corollary 6.6 and Corollary 6.7, we ask:

**Question 8.2.** *For any  $r \geq 1$  does there exists  $d > 0$  such that any aperiodic minimal  $\mathcal{S}$ -adic subshift  $(X, S)$  generated by a positive directive sequence  $\tau = (\tau_n: \mathcal{A}_{n+1}^* \rightarrow \mathcal{A}_n^*)_{n \geq 0}$  with  $\liminf |\mathcal{A}_n| \leq r$ , satisfies  $\liminf_{n \rightarrow \infty} p_X(n)/n^d = 0$ ? If not, provide a counter example.*

The Prouhet-Thue-Morse subshift illustrates how the number of asymptotic components may provide a lower bound of the topological rank of a subshift. Moreover it is possible that results in Section 7 are available for minimal subshifts of higher topological or alphabetical ranks.

**Question 8.3.** *How much of the analysis made in Section 7 can be extended to minimal systems of finite topological/alphabet rank greater than two ? In particular, do finite topological rank minimal Cantor systems have finitely many asymptotic components?*

This question is of independent interest, since a positive answer to it implies for instance that the automorphism groups of such systems are virtually  $\mathbb{Z}$ , by a direct application of results in [DDMP16]. Right before finishing this article we got informed that a big progress has been made by B. Espinoza regarding Question 8.3 (personal communication).

In Theorem 4.1 part (2) we stated that minimal  $\mathcal{S}$ -adic subshifts of bounded alphabet rank are topological factors of finite topological rank minimal Cantor systems. It is natural to think that an  $\mathcal{S}$ -adic subshift itself is of finite topological rank, but this relation is still unclear. So it is natural to ask in general,

**Question 8.4.** *Is a topological factor of a minimal Cantor system of finite topological rank of finite topological rank too?*

We suspect this question has a positive answer since in the particular case of a subshift of non-superlinear complexity the subshifts factors also have a non-superlinear complexity, hence they have a finite topological rank.

It is interesting (and maybe more easy to manage) to restrict all the former questions to a rich class of systems with a specific dynamical property, like Toeplitz subshifts. Already for this particular case the questions are open since we do not know a suitable characterization of Toeplitz subshifts with finite topological rank. More precisely and motivated by the fact that any Toeplitz subshift admits a Bratteli-Vershik representation whose towers per level have the same height, we ask the following:

**Question 8.5.** *Let  $(X, S)$  be a Toeplitz subshift. Is it true that  $X$  has finite topological rank if and only if the complexity of  $X$  is non-superlinear?*

The subtle point of this question is that the representation of the Toeplitz subshift with a bounded number of vertices at each level does not need to satisfy the

ERS property (recall the discussion right after Corollary 6.7), and conversely, the representation satisfying the ERS property may not have a bounded number of vertices at each level. To answer Question 8.5 it suffices to give a positive answer to the following: does a Toeplitz subshift of finite topological rank admit a representation satisfying the ERS property with a bounded number of vertices at each level?

A positive answer to this question would allow us to better understand all previous questions within the family of Toeplitz subshifts.

## REFERENCES

- [Aus88] J. Auslander, *Minimal flows and their extensions*, North-Holland Mathematics Studies, vol. 153, North-Holland Publishing Co., Amsterdam, 1988, Notas de Matemática [Mathematical Notes], 122.
- [BCBD<sup>+</sup>] V. Berthé, P. Cecchi-Bernales, F. Durand, J. Leroy, D. Perrin, and S. Petite, *On the dimension group of unimodular  $\mathcal{S}$ -adic subshifts*, arXiv e-prints, arXiv:1911.07700.
- [BDM10] X. Bressaud, F. Durand, and A. Maass, *On the eigenvalues of finite rank Bratteli-Vershik dynamical systems*, Ergodic Theory Dynam. Systems **30** (2010), 639–664.
- [BH94] M. Boyle and D. Handelman, *Entropy versus orbit equivalence for minimal homeomorphisms*, Pacific J. Math. **164** (1994), 1–13.
- [BKM08] S. Bezuglyi, J. Kwiatkowski, and K. Medynets, *Aperiodic substitutional systems and their bratteli diagrams*, Ergodic Theory Dynam. Systems **29** (2008), 37–72.
- [BKMS13] S. Bezuglyi, J. Kwiatkowski, K. Medynets, and B. Solomyak, *Finite rank Bratteli diagrams: structure of invariant measures*, Trans. Amer. Math. Soc. **365** (2013), 2637–2679.
- [Bos85] M. Boshernitzan, *A unique ergodicity of minimal symbolic flows with linear block growth*, J. Analyse Math. **44** (1984/85), 77–96.
- [BSTY19] V. Berthé, W. Steiner, J. Thuswaldner, and R. Yassawi, *Recognizability for sequences of morphisms*, Ergodic Theory Dynam. Systems **39** (2019), 2896–2931.
- [CDHM03] M. I. Cortez, F. Durand, B. Host, and A. Maass, *Continuous and measurable eigenfunctions of linearly recurrent dynamical cantor systems*, J. of the London Math. Soc. **67** (2003), no. 3, 790–804.
- [CDP16] M. I. Cortez, F. Durand, and S. Petite, *Eigenvalues and strong orbit equivalence*, Ergodic Theory Dynam. Systems **36** (2016), 2419–2440.
- [CK15] V. Cyr and B. Kra, *The automorphism group of a shift of linear growth: beyond transitivity*, Forum Math. Sigma **3** (2015), e5, 27.
- [CK19] ———, *Counting generic measures for a subshift of linear growth*, J. Eur. Math. Soc. (JEMS) **21** (2019), no. 2, 355–380.
- [DDMP16] S. Donoso, F. Durand, A. Maass, and S. Petite, *On automorphism groups of low complexity subshifts*, Ergodic Theory Dynam. Systems **36** (2016), 64–95.
- [DDMP17] ———, *On automorphism groups of Toeplitz subshifts*, Discrete Anal. (2017), Paper No. 11, 19.
- [DFM15] F. Durand, A. Frank, and A. Maass, *Eigenvalues of Toeplitz minimal systems of finite topological rank*, Ergodic Theory Dynam. Systems **35** (2015), 2499–2528.
- [DFM19] ———, *Eigenvalues of minimal Cantor systems*, J. Eur. Math. Soc. (JEMS) **21** (2019), 727–775.
- [DHS99] F. Durand, B. Host, and C. Skau, *Substitutive dynamical systems, Bratteli diagrams and dimension groups*, Ergodic Theory Dynam. Systems **19** (1999), 953–993.
- [DL12] F. Durand and J. Leroy,  *$\mathcal{S}$ -adic conjecture and Bratteli diagrams*, C. R. Math. Acad. Sci. Paris **350** (2012), 979–983.
- [DM08] T. Downarowicz and A. Maass, *Finite-rank Bratteli-Vershik diagrams are expansive*, Ergodic Theory Dynam. Systems **28** (2008), 739–747.
- [Dur98] F. Durand, *A characterization of substitutive sequences using return words*, Discrete Math. **179** (1998), no. 1-3, 89–101.
- [Dur00] F. Durand, *Linearly recurrent subshifts have a finite number of non-periodic subshift factors*, Ergodic Theory Dynam. Systems **20** (2000), 1061–1078.

- [Dur10] ———, *Combinatorics on Bratteli diagrams and dynamical systems*, Combinatorics, automata and number theory, Encyclopedia Math. Appl., vol. 135, Cambridge Univ. Press, Cambridge, 2010, pp. 324–372.
- [Fer96] S. Ferenczi, *Rank and symbolic complexity*, Ergodic Theory Dynam. Systems **16** (1996), 663–682.
- [FM10] S. Ferenczi and T. Monteil, *Infinite words with uniform frequencies, and invariant measures*, Combinatorics, automata and number theory, Encyclopedia Math. Appl., vol. 135, Cambridge Univ. Press, Cambridge, 2010, pp. 373–409.
- [GHH18] T. Giordano, D. Handelman, and M. Hosseini, *Orbit equivalence of Cantor minimal systems and their continuous spectra*, Math. Z. **289** (2018), 1199–1218.
- [GJ00] R. Gjerde and Ø. Johansen, *Bratteli-Vershik models for Cantor minimal systems: applications to Toeplitz flows*, Ergodic Theory Dynam. Systems **20** (2000), 1687–1710.
- [GPS95] T. Giordano, I. F. Putnam, and C. F. Skau, *Topological orbit equivalence and  $C^*$ -crossed products*, J. Reine Angew. Math. **469** (1995), 51–111.
- [Høy17] S.-M. Høyne, *Finite-rank Bratteli-Vershik diagrams are expansive—a new proof*, Math. Scand. **120** (2017), 195–210.
- [HPS92] R. Herman, I. Putnam, and C. Skau, *Ordered Bratteli diagrams, dimension groups and topological dynamics*, Internat. J. Math. **3** (1992), 827–864.
- [Jul10] A. Julien, *Complexity and cohomology for cut-and-projection tilings*, Ergodic Theory and Dynamical Systems **30** (2010), no. 2, 489–523.
- [KM02] J. Karhumäki and J. Manuch, *Multiple factorizations of words and defect effect*, Theoret. Comput. Sci. **273** (2002), 81–97, WORDS (Rouen, 1999).
- [Lot02] M. Lothaire, *Algebraic combinatorics on words*, Encyclopedia of Mathematics and its Applications, vol. 90, Cambridge University Press, Cambridge, 2002.
- [Mos92] B. Mossé, *Puissances de mots et reconnaissabilité des points fixes d’une substitution*, Theoret. Comput. Sci. **99** (1992), 327–334.
- [ORW82] D. Ornstein, D. Rudolph, and B. Weiss, *Equivalence of measure preserving transformations*, Mem. Amer. Math. Soc. **37** (1982), no. 262, xii+116.
- [QZ04] A. Quas and L. Zamponi, *Periodicity and local complexity*, Theoret. Comput. Sci. **319** (2004), 229–240.
- [Shi17] T. Shimomura, *Finite-rank Bratteli-Vershik homeomorphisms are expansive*, Proc. Amer. Math. Soc. **145** (2017), 4353–4362.

DEPARTAMENTO DE INGENIERÍA MATEMÁTICA AND CENTRO DE MODELAMIENTO MATEMÁTICO,  
UNIVERSIDAD DE CHILE & UMI-CNRS 2807, BEAUCHEF 851, SANTIAGO, CHILE.

Email address: sdonoso@dim.uchile.cl

LABORATOIRE AMIÉNOIS DE MATHÉMATIQUES FONDAMENTALES ET APPLIQUÉES, CNRS-UMR 7352, UNIVERSITÉ DE PICARDIE JULES VERNE, 33 RUE SAINT LEU, 80039 AMIENS CEDEX 1, FRANCE.

Email address: fabien.durand@u-picardie.fr

DEPARTAMENTO DE INGENIERÍA MATEMÁTICA AND CENTRO DE MODELAMIENTO MATEMÁTICO,  
UNIVERSIDAD DE CHILE & UMI-CNRS 2807, BEAUCHEF 851, SANTIAGO, CHILE.

Email address: amaass@dim.uchile.cl

LABORATOIRE AMIÉNOIS DE MATHÉMATIQUES FONDAMENTALES ET APPLIQUÉES, CNRS-UMR 7352, UNIVERSITÉ DE PICARDIE JULES VERNE, 33 RUE SAINT LEU, 80039 AMIENS CEDEX 1, FRANCE.

Email address: samuel.petite@u-picardie.fr