

# Minimal and proximal examples of $\bar{d}$ -stable and $\bar{d}$ -approachable shift spaces

MELIH EMIN CAN<sup>†</sup>, JAKUB KONIECZNY<sup>‡</sup>, MICHAŁ KUPSA<sup>§</sup> and  
DOMINIK KWIETNIAK<sup>†</sup>

<sup>†</sup> *Faculty of Mathematics and Computer Science, Jagiellonian University in Krakow,  
ul. Łojasiewicza 6, 30-348 Kraków, Poland  
(e-mail: melih.can@im.uj.edu.pl; dominik.kwietniak@uj.edu.pl)*

<sup>‡</sup> *Department of Computer Science, University of Oxford, Wolfson Building, Parks Road,  
Oxford, OX1 3QD, UK  
(e-mail: jakub.konieczny@gmail.com)*

<sup>§</sup> *The Czech Academy of Sciences, Institute of Information Theory and Automation,  
CZ-18208 Prague 8, Czechia  
(e-mail: kupsa@utia.cas.cz)*

(Received 13 August 2023 and accepted in revised form 29 April 2024)

*Abstract.* We study shift spaces over a finite alphabet that can be approximated by mixing shifts of finite type in the sense of (pseudo)metrics connected to Ornstein's  $\bar{d}$  metric ( $\bar{d}$ -approachable shift spaces). The class of  $\bar{d}$ -approachable shifts can be considered as a topological analog of measure-theoretical Bernoulli systems. The notion of  $\bar{d}$ -approachability, together with a closely connected notion of  $\bar{d}$ -shadowing, was introduced by Konieczny, Kupsa, and Kwietniak [*Ergod. Th. & Dynam. Sys.* **43**(3) (2023), 943–970]. These notions were developed with the aim of significantly generalizing specification properties. Indeed, many popular variants of the specification property, including the classic one and the almost/weak specification property, ensure  $\bar{d}$ -approachability and  $\bar{d}$ -shadowing. Here, we study further properties and connections between  $\bar{d}$ -shadowing and  $\bar{d}$ -approachability. We prove that  $\bar{d}$ -shadowing implies  $\bar{d}$ -stability (a notion recently introduced by Tim Austin). We show that for surjective shift spaces with the  $\bar{d}$ -shadowing property the Hausdorff pseudodistance  $\bar{d}^H$  between shift spaces induced by  $\bar{d}$  is the same as the Hausdorff distance between their simplices of invariant measures with respect to the Hausdorff distance induced by Ornstein's metric  $\bar{d}$  between measures. We prove that without  $\bar{d}$ -shadowing this need not to be true (it is known that the former distance always bounds the latter). We provide examples illustrating these results, including minimal examples and proximal examples of shift spaces with the  $\bar{d}$ -shadowing property. The existence of such shift spaces was announced in the earlier paper mentioned above. It shows that  $\bar{d}$ -shadowing indeed generalizes the specification property.

Key words: specification property, topological entropy, shift space, Poulsen simplex, Besicovitch pseudometric

2020 Mathematics Subject Classification: 37B05 (Primary); 37A35, 37B10, 37B40, 37D20 (Secondary)

### 1. Introduction

Given a finite set (an *alphabet*)  $\mathcal{A}$  we let  $\mathcal{A}^\infty$  stand for the *full shift over  $\mathcal{A}$* , that is, a set of all  $\mathcal{A}$ -valued infinite sequences. To avoid trivialities, we assume that  $\mathcal{A}$  has at least two elements. We endow  $\mathcal{A}^\infty$  with the product topology induced by the discrete topology on  $\mathcal{A}$ , which turns  $\mathcal{A}^\infty$  into a compact metrizable space. Let  $\rho$  be a metric compatible with the topology on  $\mathcal{A}^\infty$ . The shift operator  $\sigma: \mathcal{A}^\infty \rightarrow \mathcal{A}^\infty$  turns  $\mathcal{A}^\infty$  into a non-invertible dynamical system. From the dynamical point of view, the most interesting objects are closed non-empty  $\sigma$ -invariant subsets of  $\mathcal{A}^\infty$  (*one-sided shift spaces* or *subshifts*). We also consider the space  $\mathcal{M}(\mathcal{A}^\infty)$  of all Borel probability measures on  $\mathcal{A}^\infty$  with the weak\* topology. The set of  $\sigma$ -invariant measures in  $\mathcal{M}(\mathcal{A}^\infty)$  concentrated on a shift space  $X \subseteq \mathcal{A}^\infty$  is denoted by  $\mathcal{M}_\sigma(X)$ . Each of these objects (invariant measures and subshifts) has a canonically defined sequence of Markov approximations converging to it in a natural topology. This fact, however, is of little practical use, because the convergence is too weak to allow for a transfer of dynamical properties from an approximating sequence to the properties of its limit.

Recall that the natural topology on the space of all subshifts of  $\mathcal{A}^\infty$  is the hyperspace (Vietoris) topology of non-empty closed subsets of a compact metric space. In other words, a sequence of shift spaces  $(X_n)_{n=1}^\infty \subseteq \mathcal{A}^\infty$  converges to a shift space  $X \subseteq \mathcal{A}^\infty$  in the hyperspace topology if  $\rho^H(X_n, X) \rightarrow 0$  as  $n \rightarrow \infty$  (here,  $\rho^H$  is the Hausdorff metric corresponding to  $\rho$ ). Similarly, we say that simplices of invariant measures of shift spaces  $(X_n)_{n=1}^\infty \subseteq \mathcal{A}^\infty$  approximate the simplex of invariant measures of a shift space  $X \subseteq \mathcal{A}^\infty$  if  $\mathcal{M}_\sigma(X_n)$  converges to  $\mathcal{M}_\sigma(X)$  as  $n \rightarrow \infty$  in the natural hyperspace topology of  $\mathcal{M}_\sigma(\mathcal{A}^\infty)$ , that is, if  $D^H(\mathcal{M}_\sigma(X_n), \mathcal{M}_\sigma(X)) \rightarrow 0$  as  $n \rightarrow \infty$ , where  $D^H$  is the Hausdorff metric corresponding to a metric  $D$  compatible with the weak\* topology on  $\mathcal{M}_\sigma(\mathcal{A}^\infty)$ .

Fortunately, for both measures and subshifts, stronger metrics than  $\rho$  and  $D$  are also available. A useful metric for  $\sigma$ -invariant measures is Ornstein's metric  $\bar{d}_M$ . In [23] we studied a topology on the powerset of  $\mathcal{A}^\infty$  induced by the Hausdorff pseudometric  $\bar{d}^H$  derived from  $\bar{d}$ -pseudometric on  $\mathcal{A}^\infty$ . A very similar idea of using  $\bar{d}$ -approximation was independently considered by Thompson [42], who used it in the settings of [4].

Recall that the pseudometric  $\bar{d}$  is given for  $x = (x_j)_{j=0}^\infty, y = (y_j)_{j=0}^\infty \in \mathcal{A}^\infty$  by

$$\bar{d}(x, y) = \limsup_{n \rightarrow \infty} \frac{1}{n} |\{0 \leq j < n : x_j \neq y_j\}|. \quad (1)$$

Since  $\bar{d}(x, y)$  can be zero for distinct  $x$  and  $y$ ,  $\bar{d}$  is not a metric. Nevertheless, after factorizing by the equivalence relation  $\sim$  on  $\mathcal{A}^\infty$ , where  $x \sim y$  if and only if  $\bar{d}(x, y) = 0$ , we obtain the factor space  $\mathcal{A}^\infty / \sim$  on which  $\bar{d}$  becomes a complete, non-separable metric. Since  $\bar{d}$  is bounded by 1 on  $\mathcal{A}^\infty$ , it induces a Hausdorff pseudometric  $\bar{d}^H$  on the space  $\text{CL}(\mathcal{A}^\infty, \bar{d})$  of all non-empty  $\bar{d}$ -closed subsets of  $\mathcal{A}^\infty$ . Similarly,  $\bar{d}_M$  is a complete

bounded non-separable metric on  $\mathcal{M}_\sigma(\mathcal{A}^\infty)$  inducing a Hausdorff metric  $\bar{d}_M^H$  on the space  $\text{CL}(\mathcal{M}_\sigma(\mathcal{A}^\infty), \bar{d}_M)$  of all non-empty  $\bar{d}_M$ -closed subsets of  $\mathcal{M}_\sigma(\mathcal{A}^\infty)$ . Note that  $\bar{d}_M$ -convergence implies weak\* convergence, so for each shift space  $X \subseteq \mathcal{A}^\infty$  the set  $\mathcal{M}_\sigma(X)$  is  $\bar{d}_M$ -closed. It is also known that the set of ergodic measures on  $X$ , denoted by  $\mathcal{M}_\sigma e(X)$ , is  $\bar{d}_M$ -closed.

Hence, we obtain two more ways to say that shift spaces  $(X_n)_{n=1}^\infty \subseteq \mathcal{A}^\infty$  approximate  $X \subseteq \mathcal{A}^\infty$ :

$$\lim_{n \rightarrow \infty} \bar{d}^H(X_n, X) = 0, \tag{2}$$

$$\lim_{n \rightarrow \infty} \bar{d}_M^H(\mathcal{M}_\sigma(X_n), \mathcal{M}_\sigma(X)) = 0, \tag{3}$$

Note that we do not assume that the approximation in (2) and (3) is monotone (meaning  $X_1 \supseteq X_2 \supseteq \dots$  and  $X = \bigcap X_n$ ), but in practice it is often the case.

In [23], we studied the consequences of the existence of an approximating sequence as in (2) and (3). We were especially interested in the case when the approximating sequence is the sequence of Markov approximations. We introduced  $\bar{d}$ -approachable shift spaces (subshifts that are approached by their topological Markov approximations not only in the ‘usual’ Hausdorff metric topology, but also in the  $\bar{d}^H$  sense). We also considered a condition that is ostensibly a relaxation of (3):

$$\lim_{n \rightarrow \infty} \bar{d}_M^H(\mathcal{M}_\sigma e(X_n), \mathcal{M}_\sigma e(X)) = 0. \tag{4}$$

We proved in [23] that for every shift space  $X$  and  $Y$  over  $\mathcal{A}$  we have

$$\bar{d}_M^H(\mathcal{M}_\sigma(X), \mathcal{M}_\sigma(Y)) = \bar{d}_M^H(\mathcal{M}_\sigma e(X), \mathcal{M}_\sigma e(Y)) \leq \bar{d}^H(X, Y), \tag{5}$$

hence

$$(2) \implies (3) \iff (4).$$

In other words,  $\bar{d}^H$  approximation (2) implies convergence of simplices of invariant measures in the Hausdorff metric  $\bar{d}_M^H$  induced by Ornstein’s  $\bar{d}_M$  metric on the space  $\mathcal{M}_\sigma(\mathcal{A}^\infty)$  as in (3). As a consequence, certain features of simplices of invariant measures of shift spaces in the approximating sequence are inherited by the simplex of the limit.

In analogy with Friedman and Ornstein’s result characterizing Bernoulli measures among all totally ergodic shift-invariant measures as  $\bar{d}_M$ -limits of their own Markov approximations (see [14]), we also characterized in [23] chain-mixing  $\bar{d}$ -approachable shift spaces using the newly introduced  $\bar{d}$ -shadowing property. (Note that Ornstein’s theory characterizing Bernoullicity works best for invertible measure-preserving systems. In particular, Markov shifts carrying Markov approximations mentioned here are two-sided, while we consider the one-sided setting.) In this way we obtained a large family of shift spaces that contains all  $\beta$ -shifts and all mixing sofic shifts, in particular all mixing shifts of finite type. This is because many specification properties imply chain mixing and  $\bar{d}$ -approachability (this is the case, for example, for all shift spaces with the almost specification property). We refer to [23] for more details.

We also showed in [23] that if every  $X_n$  has an entropy-dense set of ergodic measures and the sequence  $(X_n)_{n=1}^\infty$  converges to  $X$  in the sense defined by any of (2)–(4), then

ergodic measures of  $X$  are also entropy dense. This established a new method of proving entropy density. As a consequence, we obtained entropy density of ergodic measures for all surjective shift spaces with the  $\bar{d}$ -shadowing property. Entropy density of ergodic measures is a property introduced by Orey in 1986 [34] and Föllmer and Orey in 1988 [13]. Recall that ergodic measures of a shift space  $X$  are entropy dense if every invariant measure can be approximated with an ergodic one with respect to the weak\* topology and entropy at the same time. In particular,  $\mathcal{M}_\sigma e(X)$  is a dense subset of  $\mathcal{M}_\sigma(X)$ . Note that there are shift spaces with dense, but not entropy-dense, sets of ergodic measures (see [16]). Density of ergodic measures and entropy density are strongly related to the theory of large deviations and multifractal analysis [5, 12, 38, 39]; see Comman's article [6] and references therein for more information about that connection.

The results of [23] are also applicable to the study of the dynamics of the so-called  $\mathcal{B}$ -free shifts (or systems), a subject that has recently attracted considerable interest (see [10, 11, 19–22, 24–26]).

In the present paper, we study further properties of  $\bar{d}$ -approachable shift spaces. In particular, in §5 we construct minimal and proximal examples of chain-mixing  $\bar{d}$ -approachable shift spaces. These examples demonstrate that our technique yields entropy density for shift spaces that are beyond the reach of methods based on specification, as specification excludes both proximality and minimality. So far only specification-like conditions have been invoked to prove entropy density explicitly (see [12, 38]). We note that there exists a general theorem due to Downarowicz and Serafin [8] that guarantees existence of minimal shifts with entropy-dense ergodic measures, but due to its generality it is hard to see concrete examples. We also prove (see §3) that  $\bar{d}$ -approachability implies  $\bar{d}$ -stability, where  $\bar{d}$ -stability is a property recently introduced by Austin [1]. Austin combined one of Ornstein's conditions equivalent to Bernoullicity with equivariant analogs of some basic results in measure concentration to characterize Bernoullicity of the equilibrium measure of a continuous potential  $\varphi$  under the assumption that the equilibrium is unique. Austin formulated his main condition in terms of a stronger kind of differentiability of the pressure functional at  $\varphi$ . He proved that the condition is always necessary and he showed that it is sufficient if the shift space is ' $\bar{d}$ -stable'. Austin remarked that the class of ' $\bar{d}$ -stable' subshifts includes the full shift and several other examples with the specification property. He also suspected that  $\bar{d}$ -stability holds also for examples without any specification properties. Hence, our minimal and proximal examples of  $\bar{d}$ -approachable shifts confirm this suspicion. In §4 we also show that the implication (3)  $\implies$  (2) holds true if all shift spaces involved have the  $\bar{d}$ -shadowing property, hence chain mixing and  $\bar{d}$ -approachability suffice for this implication; see Theorem 4.1. We note that the implication (3)  $\implies$  (2) does not hold in general by producing a sequence of shift spaces  $(X_n)_{n=1}^\infty$  such that for some shift space  $X$  we have  $d_{\mathcal{M}}^H(\mathcal{M}_\sigma(X_n), \mathcal{M}_\sigma(X)) \rightarrow 0$  while  $d^H(X_n, X) \rightarrow 1$  as  $n \rightarrow \infty$ ; see Proposition 4.4.

We prove (Proposition 4.2) that it is possible to find a sequence of shift spaces  $(X_n)_{n=1}^\infty$  such that  $\mathcal{M}_\sigma(X_n)$  converges in  $\bar{d}_{\mathcal{M}}$  to a singleton set that is not a simplex of invariant measures for any shift space. Finally, we prove that the  $\bar{d}$ -shadowing on the measure center of a shift space (the smallest invariant subshift of full measure for every invariant measure) implies the same for the shift. We recall our notation and basic definitions in §2.

The results of the present paper, as well as the results of [23], are also applicable to two-sided shift spaces (shift-invariant subsets of  $\mathcal{A}^{\mathbb{Z}}$ ), provided that the definition of the pseudometric  $\bar{d}$  stays as in (1), that is, we average over the coordinates  $0, 1, \dots, n - 1$ .

2. Definitions

2.1. Hausdorff pseudometrics. Let  $Z$  be a set. A pseudometric on  $Z$  is a real-valued, non-negative, symmetric function  $\rho$  on  $Z \times Z$  vanishing on the diagonal  $\{(x, y) \in Z \times Z : x = y\}$  and satisfying the triangle inequality. Let  $\rho$  be a bounded pseudometric on  $Z$ . For  $z \in Z$  and non-empty  $A, B \subseteq Z$ , we define

$$\rho(z, B) = \inf_{b \in B} \rho(z, b) \quad \text{and} \quad \rho^H(A, B) = \max \left\{ \sup_{a \in A} \rho(a, B), \sup_{b \in B} \rho(b, A) \right\}.$$

We call  $\rho^H$  the Hausdorff pseudometric induced by  $\rho$  on the space of all non-empty subsets of  $Z$ . If  $\rho$  is a bounded metric, then  $\rho^H$  becomes a metric on the set  $\text{CL}(Z, \rho)$  of closed non-empty subsets of  $(Z, \rho)$ . Note that in our settings some properties, well known in the compact case, fail because we consider  $(Z, \rho)$  where  $\rho$  is not necessarily compact, but only a bounded pseudometric space. For example,  $\rho$  and another pseudometric  $\tilde{\rho}$  may induce the same topology on  $Z$  but the spaces  $(\text{CL}(Z), \rho^H)$  and  $(\text{CL}(Z), \tilde{\rho}^H)$  need not be homeomorphic.

2.2. Shift spaces and languages. We let  $\mathbb{N}$  denote the set of positive integers. We also write  $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$ . Unless otherwise stated, the letters  $i, j, k, l, m, n$  always denote integers. An alphabet is a finite set  $\mathcal{A}$  endowed with the discrete topology. We refer to elements of  $\mathcal{A}$  as symbols or letters. The full shift  $\mathcal{A}^{\infty}$  is the Cartesian product of infinitely many copies of  $\mathcal{A}$  indexed by  $\mathbb{N}_0$ . We endow  $\mathcal{A}^{\infty}$  with the product topology. A compatible metric on  $\mathcal{A}^{\infty}$  is given for  $x, y \in \mathcal{A}^{\infty}$  by

$$\rho(x, y) = \begin{cases} 0 & \text{if } x = y, \\ 2^{-\min\{j : x_j \neq y_j\}} & \text{otherwise.} \end{cases}$$

The shift map  $\sigma : \mathcal{A}^{\infty} \rightarrow \mathcal{A}^{\infty}$  is given for  $x = (x_i)_{i=0}^{\infty} \in \mathcal{A}^{\infty}$  and  $j \geq 0$  by  $\sigma(x)_j = x_{j+1}$ . A shift space over  $\mathcal{A}$  is a non-empty, closed, and  $\sigma$ -invariant subset of  $\mathcal{A}^{\infty}$ . A word over  $\mathcal{A}$  is a finite sequence of elements of  $\mathcal{A}$ . The number of entries of a word  $w$  is called the length of  $w$  and is denoted by  $|w|$ . The empty sequence is called the empty word and is the only word of length 0. We denote it by  $\lambda$ . The concatenation of words  $u = u_1 \cdots u_k$  and  $v = v_1 \cdots v_m$  is the word  $u_1 \cdots u_k v_1 \cdots v_m$  denoted simply by  $uv$ . Given  $x \in \mathcal{A}^{\infty}$  and  $0 \leq i < j$ , we let  $x_{[i,j]}$  denote the word  $x_i x_{i+1} \cdots x_{j-1}$  over  $\mathcal{A}$  of length  $j - i$ . We say that a word  $w$  appears in  $x \in \mathcal{A}^{\infty}$  if there exist  $0 \leq i < j$  such that  $w = x_{[i,j]}$ . A word  $w$  appears in a shift space  $X \subseteq \mathcal{A}^{\infty}$  if there exists  $x \in X$  such that  $w$  appears in  $x$ . The language of a shift space  $X \subseteq \mathcal{A}^{\infty}$  is the set  $\mathcal{B}(X)$  of all finite words over  $\mathcal{A}$  appearing in  $X$ . We agree that the empty word appears in every sequence in  $\mathcal{A}^{\infty}$ . For  $n \in \mathbb{N}_0$ , we let  $\mathcal{B}_n(X) \subseteq \mathcal{A}^n$  be the set of all words  $w \in \mathcal{B}(X)$  with  $|w| = n$ . Given a set  $\mathcal{F}$  of finite words over  $\mathcal{A}$ , we define  $X_{\mathcal{F}}$  be the set of all  $x = (x_i)_{i=0}^{\infty} \in \mathcal{A}^{\infty}$  such that no word from  $\mathcal{F}$  appears in  $x$ . The resulting set  $X_{\mathcal{F}}$  is either empty or a shift space. Furthermore, for every shift space  $X$  over  $\mathcal{A}$  one can find a collection  $\mathcal{F}$  of finite words

such that  $X = X_{\mathcal{F}}$ . A shift space  $X$  is a *shift of finite type* if there exists a finite set  $\mathcal{F}$  such that  $X = X_{\mathcal{F}}$ . Every shift space  $X \subseteq \mathcal{A}^{\infty}$  is the intersection of a sequence  $(X_n^M)_{n \geq 0}$  of shifts of finite type. To construct that sequence, we define  $\mathcal{F}[n]$  to consist of all words  $w$  over  $\mathcal{A}$  with  $|w| = n + 1$  and  $w \notin \mathcal{B}_{n+1}(X)$ . In this way, we obtain for each  $n \geq 0$  a shift of finite type  $X_n^M = X_{\mathcal{F}[n]}$  such that  $\mathcal{B}_j(X) = \mathcal{B}_j(X_n^M)$  for  $0 \leq j \leq n + 1$ . We call the shift space  $X_n^M$  the *n*th (topological) Markov approximation of  $X$  or *finite type approximation of order  $n$  to  $X$* . We note that for every shift space  $X$ , its Markov approximation  $X_n^M$  can be conveniently described using a Rauzy graph. The *n*th Rauzy graph of  $X$  is a labeled graph  $G_n = (V_n, E_n, \tau_n)$ , where we set  $V_n = \mathcal{B}_n(X)$  and  $E_n = \mathcal{B}_{n+1}(X)$ , and for each  $w = w_0 w_1 \cdots w_n \in E_n$  we define  $i(w) = w_0 \cdots w_{n-1} \in V_n$ ,  $t(w) = w_1 \cdots w_n \in V_n$ , and  $\tau_n(w) = w_0 \in \mathcal{A}$ . The sofic shift space  $X_n$  presented by  $G_n$  satisfies  $\mathcal{B}_j(X_n) = \mathcal{B}_j(X)$  for  $j = 1, \dots, n + 1$ ; see Proposition 3.62 in [28]. It is now easy to see that  $X_n$  is the *n*th topological Markov approximations for  $X$ .

The following definitions of (chain) mixing and (chain) transitivity are stated only for shift spaces. We will do the same for several notions: instead of presenting a general definition for continuous maps acting on compact metric spaces (for the latter, see [28]), we will state an equivalent definition adapted to symbolic dynamics. This applies to (chain) transitivity, (chain) mixing, specification, and its variants.

A shift space  $X$  is *transitive* if for every  $u, w \in \mathcal{B}(X)$  there exists  $v$  with  $uvw \in \mathcal{B}(X)$ . A shift space  $X$  is *topologically mixing* if for any  $u, w \in \mathcal{B}(X)$  there exists  $N \in \mathbb{N}_0$  such that for each  $n \geq N$  there is  $v = v(n) \in \mathcal{B}_n(X)$  such that  $uvw \in \mathcal{B}(X)$ .

A shift space is *chain transitive* (respectively, *chain mixing*) if its topological Markov approximations  $X_n^M$  are transitive (respectively, topologically mixing) for all except finitely many  $n$ s.

2.3. *Ergodic properties of shift spaces.* Let  $\mathcal{M}(X)$  be the set of all Borel probability measures supported on a shift space  $X \subseteq \mathcal{A}^{\infty}$ . In particular,  $\mathcal{M}(\mathcal{A}^{\infty})$  stands for the space of all Borel probability measures on  $\mathcal{A}^{\infty}$ . We write  $\mathcal{M}_{\sigma}(X)$  and  $\mathcal{M}_{\sigma e}(X)$  to denote respectively the sets of  $\sigma$ -invariant and ergodic  $\sigma$ -invariant measures in  $\mathcal{M}(X)$ . We endow  $\mathcal{M}(\mathcal{A}^{\infty})$  with the weak\* topology, hence it becomes a compact metrizable space and  $\mathcal{M}_{\sigma}(X)$  is its closed subset for every shift space  $X \subseteq \mathcal{A}^{\infty}$ .

We say that  $x \in \mathcal{A}^{\infty}$  *generates*  $\mu \in \mathcal{M}_{\sigma}(X)$  along a strictly increasing sequence of integers  $(N_k)_{k=1}^{\infty}$ , if for every continuous function  $f : X \rightarrow \mathbb{R}$  the sequence of Cesàro averages of  $(f(\sigma^n(x)))_{n=0}^{\infty}$  along  $(N_k)_{k=1}^{\infty}$  converges and the limit satisfies

$$\lim_{k \rightarrow \infty} \frac{1}{N_k} \sum_{n=0}^{N_k-1} f(\sigma^n(x)) = \int_{\mathcal{A}^{\infty}} f \, d\mu.$$

Compactness implies that for every strictly increasing sequence of integers  $(N_k)_{k=1}^{\infty}$  and every point  $x \in X$  there is a subsequence of  $(N_k)_{k=1}^{\infty}$  such that  $x$  generates an invariant measure along that subsequence. In particular, every point always generates at least one measure. A point  $x \in \mathcal{A}^{\infty}$  is *generic* for  $\mu \in \mathcal{M}_{\sigma}(X)$  if  $\mu$  is the unique measure generated by  $x$ . Every ergodic measure has a generic point. We denote by  $h(\mu)$  the Kolmogorov–Sinai entropy of  $\mu \in \mathcal{M}_{\sigma}(X)$ . We say that ergodic measures of a shift space  $X$  are *entropy dense*

if for every measure  $\mu \in \mathcal{M}_\sigma(X)$ , every neighborhood  $U$  of  $\mu$  in  $\mathcal{M}_\sigma(X)$ , and every  $\varepsilon > 0$  there is  $\nu \in U \cap \mathcal{M}_\sigma e(X)$  with  $|h(\nu) - h(\mu)| < \varepsilon$ . Note that having entropy-dense ergodic measures is preserved by conjugacy.

2.4. *The functions  $\bar{d}$  and  $\bar{d}_M$ .* Given  $x = (x_n)_{n=0}^\infty, y = (y_n)_{n=0}^\infty \in \mathcal{A}^\infty$ , we set

$$\bar{d}(x, y) = \limsup_{n \rightarrow \infty} \frac{1}{n} |\{0 \leq j < n : x_j \neq y_j\}|.$$

The function  $\bar{d}$  is a pseudometric on  $\mathcal{A}^\infty$ , but  $\bar{d}$  is not a metric if  $\mathcal{A}$  has at least two elements, because the implication  $\bar{d}(x, y) = 0 \implies x = y$  fails. The function  $\bar{d}$  is not continuous in general. Furthermore,  $\bar{d}: \mathcal{A}^\infty \times \mathcal{A}^\infty \rightarrow [0, 1]$  is shift invariant (for all  $x, y \in \mathcal{A}^\infty$  we have  $\bar{d}(x, y) = \bar{d}(\sigma(x), \sigma(y))$ ).

Ornstein’s metric  $\bar{d}_M$  on  $\mathcal{M}_\sigma(\mathcal{A}^\infty)$  is usually defined with the help of joinings. (Ornstein’s metric  $\bar{d}_M$  is usually denoted by  $\bar{d}$ , but in [23] as well as in this paper the distinction between  $\bar{d}$  and  $\bar{d}_M$  is crucial. We refer to  $\bar{d}_M$  as the ‘ $d$ -bar distance for measures’ and we call  $\bar{d}$  on  $\mathcal{A}^\infty$  ‘pointwise  $d$ -bar’ or simply ‘ $d$ -bar’.) A  $\sigma \times \sigma$ -invariant measure  $\xi$  on  $\mathcal{A}^\infty \times \mathcal{A}^\infty$  is a *joining* of  $\mu, \nu \in \mathcal{M}_\sigma(\mathcal{A}^\infty)$  if  $\mu$  and  $\nu$  are the marginal measures for  $\xi$  under the projection to the first (respectively, the second) coordinate. We write  $J(\mu, \nu)$  for the set of all joinings of  $\mu$  and  $\nu$ . Note that  $J(\mu, \nu)$  is always non-empty because the product measure  $\mu \times \nu$  belongs to  $J(\mu, \nu)$ . Ornstein’s metric  $\bar{d}_M$  on  $\mathcal{M}_\sigma(\mathcal{A}^\infty)$  is given by

$$\bar{d}_M(\mu, \nu) = \inf_{\xi \in J(\mu, \nu)} \int_{\mathcal{A}^\infty \times \mathcal{A}^\infty} d_0(x, y) d\xi(x, y),$$

where  $d_0(x, y) = 1$  if  $x_0 \neq y_0$  and  $d_0(x, y) = 0$  otherwise. The space  $\mathcal{M}_\sigma(\mathcal{A}^\infty)$  endowed with the  $\bar{d}_M$ -metric becomes a complete but non-separable (hence, non-compact) metric space. The space  $\mathcal{M}_\sigma e(\mathcal{A}^\infty) \subseteq \mathcal{M}_\sigma(\mathcal{A}^\infty)$  of ergodic measures is  $\bar{d}_M$ -closed, as are the spaces of strongly mixing and Bernoulli measures on  $\mathcal{A}^\infty$ . The entropy function  $\mu \mapsto h(\mu)$  is continuous on  $\mathcal{M}_\sigma(\mathcal{A}^\infty)$  under  $\bar{d}_M$ . The convergence in  $\bar{d}_M$  implies weak\* convergence (for more details, see [35]).

Applying Hausdorff metric construction described in §2.1 to the bounded metric space  $(\mathcal{M}_\sigma(\mathcal{A}^\infty), \bar{d}_M)$ , we obtain a metric denoted by  $\bar{d}_M^H$  defined on the space  $\text{CL}(\mathcal{M}_\sigma(\mathcal{A}^\infty), \bar{d}_M)$  of non-empty closed subsets of  $(\mathcal{M}_\sigma(\mathcal{A}^\infty), \bar{d}_M)$ . For every shift space  $X$  the sets  $\mathcal{M}_\sigma(X)$  and  $\mathcal{M}_\sigma e(X)$  are closed sets in the Hausdorff metric  $\bar{d}_M^H$ . Similarly, starting from the pseudometric space  $(\mathcal{A}^\infty, \bar{d})$ , we get a pseudometric  $\bar{d}^H$  on the set of all non-empty subsets of  $\mathcal{A}^\infty$ .

2.5. *On  $\bar{d}$ -shadowing and  $\bar{d}$ -approachability.* A shift space  $X \subseteq \mathcal{A}^\infty$  is  $\bar{d}$ -approachable if its Markov approximations  $X_1^M, X_2^M, \dots$  satisfy  $\bar{d}^H(X_n^M, X) \rightarrow 0$  as  $n \rightarrow \infty$ . Every shift of finite type is  $\bar{d}$ -approachable.

We say that the shift space  $X$  has the  $\bar{d}$ -shadowing property if for every  $\varepsilon > 0$  there is  $N \in \mathbb{N}$  such that every sequence  $(w^{(j)})_{j=1}^\infty$  in  $\mathcal{B}(X)$  with  $|w^{(j)}| \geq N$  for  $j = 1, 2, \dots$  is  $\varepsilon$ -traced by some point  $x' \in X$ , that is, there is  $x' \in X$  such that  $\bar{d}(x, x') < \varepsilon$ , where  $x = w^{(1)}w^{(2)}w^{(3)} \dots$ . The  $\bar{d}$ -shadowing property was introduced in [23]. It is closely related to the *average shadowing property* introduced by Blank [3] and studied in [27].

Every mixing sofic shift space has the  $\bar{d}$ -shadowing property, and  $\bar{d}$ -shadowing is inherited by  $\bar{d}^H$ -limits of shift spaces with the  $\bar{d}$ -shadowing. Note that  $\bar{d}$ -shadowing implies  $\bar{d}$ -approachability, but to prove the converse we need to assume additionally that the shift space in question is chain mixing. The exact statement is Theorem 6 in [23], which says that a shift space  $X \subseteq \mathcal{A}^\infty$  is chain mixing and  $\bar{d}$ -approachable if and only if  $\sigma(X) = X$  and  $X$  has the  $\bar{d}$ -shadowing property. Theorem 6 in [23] lists a third condition, but we will not need it until §5, so we postpone the exact statement.

As we have already mentioned in the introduction, this characterization is a topological counterpart of the result saying that a totally ergodic shift-invariant probability measure is Bernoulli if and only if it is the  $\bar{d}_M$ -limit of the sequence of its canonical Markov approximations.

In §5.1, we also show how to apply this corollary even in the case when natural approximations of our shift are not comparable via inclusion, hence do not form a descending chain of shift spaces.

### 3. $\bar{d}$ -approachability vs. $\bar{d}$ -stability

We recall the notion of  $\bar{d}_M$ -stability that has been recently introduced by Austin [1]. Then we show that it follows from the  $\bar{d}$ -shadowing property. Later we discuss some further properties of  $\bar{d}_M$ -stable shifts.

*Definition 3.1.* A shift space  $X \subseteq \mathcal{A}^\infty$  is  $\bar{d}_M$ -stable if for every  $\varepsilon > 0$  there is an open neighborhood  $\mathcal{U}$  of  $\mathcal{M}_\sigma(X)$  in the weak\* topology on  $\mathcal{M}_\sigma(\mathcal{A}^\infty)$  such that if  $\nu \in \mathcal{U}$ , then there is  $\mu \in \mathcal{M}_\sigma(X)$  with  $\bar{d}_M(\mu, \nu) < \varepsilon$ .

Note that we use slightly different notation: Austin writes  $\bar{d}$  instead of  $\bar{d}_M$ . Equivalently, a shift space  $X$  is  $\bar{d}_M$ -stable if any shift-invariant measure which lives close enough to  $X$  in the weak\* topology is actually close in Ornstein’s  $\bar{d}_M$  metric on  $\mathcal{M}_\sigma(\mathcal{A}^\infty)$  to a shift-invariant measure supported on  $X$ . This observation (noted already in [1]) is formulated as the next lemma for future reference. We state it in terms of the natural basis  $(U_n)_{n=1}^\infty$  of open neighborhoods of a shift space  $X$  with respect to the Hausdorff topology induced by the usual (product) topology on  $\mathcal{A}^\infty$  (the topology of the Hausdorff metric  $\rho^H$ ), where for  $n \geq 1$  we have

$$U_n(X) = \bigcup \{[u] \mid u \in \mathcal{B}_n(X)\}.$$

**LEMMA 3.2.** *A shift space  $X \subseteq \mathcal{A}^\infty$  is  $\bar{d}_M$ -stable if, and only if, for any  $\varepsilon > 0$  there are  $\delta > 0$  and  $N \in \mathbb{N}$  such that  $\bar{d}_M(\nu, \mathcal{M}_\sigma(X)) < \varepsilon$  whenever  $\nu \in \mathcal{M}_\sigma(\mathcal{A}^\infty)$ ,  $n \geq N$ , and  $\nu(U_n(X)) > 1 - \delta$ .*

**PROPOSITION 3.3.** *If a shift space  $X \subseteq \mathcal{A}^\infty$  has the  $\bar{d}$ -shadowing property, then  $X$  is  $\bar{d}_M$ -stable.*

*Proof.* Fix  $1 > \varepsilon > 0$ . Use the definition of the  $\bar{d}$ -shadowing property to pick  $n$  such that every sequence  $(w^{(j)})_{j=1}^\infty$  in  $\mathcal{B}(X)$  with  $|w^{(j)}| \geq n$  for every  $j \geq 1$  is  $\varepsilon/2$ -traced in the  $\bar{d}$  pseudometric by some point in  $X$ . Let  $\nu$  be a shift-invariant measure on  $\mathcal{A}^\infty$  such that  $\nu(U_n(X)) > 1 - \varepsilon/2$ . We would like to show that  $\bar{d}_M(\nu, \mathcal{M}_\sigma(X)) < \varepsilon$ .



Let  $y = (y_i)_{i=0}^\infty \in \mathcal{A}^\infty$  be a generic point for  $\nu$  (such a point always exists in  $\mathcal{A}^\infty$ ). Hence, the frequency of visits of  $y$  to in  $U_n(X)$  satisfies

$$\lim_{N \rightarrow \infty} \frac{1}{N} \{0 \leq k < N : \sigma^k(y) \in U_n(X)\} = \nu(U_n(X)) > 1 - \varepsilon/2.$$

Define  $m_0 \geq 0$  as the smallest  $\ell \geq 0$  such that  $y_{[\ell, \ell+n)} \in \mathcal{B}_n(X)$ . Inductively, given  $m_k$  for some  $k \geq 0$ , we define  $m_{k+1}$  as the smallest integer  $\ell \geq m_k + n$  such that  $y_{[\ell, \ell+n)} \in \mathcal{B}_n(X)$ . Then the set

$$M = \mathbb{N}_0 \setminus \bigcup_{k=0}^\infty [m_k, m_{k+1})$$

is contained in the set

$$\{\ell \in \mathbb{N}_0 \mid y_{[\ell, \ell+n)} \notin \mathcal{B}_n(X)\},$$

so its upper density is less than  $\varepsilon/2$ . For  $k \in \mathbb{N}_0$ , we extend the word  $y_{[m_k, m_{k+1})} \in \mathcal{B}_n(X)$  to the right to form some word  $v^{(k)} \in \mathcal{B}(X)$  of length  $m_{k+1} - m_k$ . Then the sequence

$$z = y_{[0, m_0)} v^{(0)} v^{(1)} \dots$$

differs from  $y$  only at positions (indices) belonging to the set  $M$ . Hence,  $\bar{d}(y, z) \leq \varepsilon/2$ . The same is true for  $y' = \sigma^{m_0}(y)$  and  $z' = \sigma^{m_0}(z)$ . Furthermore,  $y'$  is still generic for  $\nu$ . By the  $\bar{d}$ -shadowing property,

$$z' = v^{(0)} v^{(1)} \dots$$

can be approximated by  $z'' \in X$  such that  $\bar{d}(z', z'') < \varepsilon/2$ . Therefore,

$$\bar{d}(y', z'') < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

It is a standard fact (see [17, Theorem 15.23]) that every measure generated by  $z''$  is  $\varepsilon$ -close with respect to the  $\bar{d}_M$  distance to the measure generated by  $y'$ . Hence  $\bar{d}_M(\nu, \mathcal{M}_\sigma(X)) \leq \varepsilon$ . □

Shift spaces with the specification property are primary examples of  $\bar{d}$ -stable shift spaces provided by [1]. Recall that a shift space  $X \subset \mathcal{A}^\infty$  has the *specification property* if there exists  $k \in \mathbb{N}$  such that for any  $u, w \in \mathcal{B}(X)$  there is  $v$  with  $|v| = k$  such that  $uvw \in \mathcal{B}(X)$ . The specification property is a very useful property with many consequences for a shift space. For a more extensive overview on the specification property and its relatives we refer the reader to [30]. Here we note that the specification property and even either one of its two weaker, incommensurable variants known as the *almost specification property* or the *weak specification property* (see [30]) imply  $\bar{d}$ -approachability and chain mixing, hence  $\bar{d}$ -shadowing [23]. Therefore, we obtain the following corollary.

**COROLLARY 3.4.** *Let  $X$  be a shift space satisfying one of the following conditions:*

- (1)  *$X$  has the specification property;*
- (2)  *$X$  has the weak specification property;*
- (3)  *$X$  has the almost specification property.*

*Then  $X$  is  $\bar{d}_M$ -stable.*

Note that the nomenclature is not fixed; we follow [30]. Recall that a *mistake function* is a non-decreasing function  $g: \mathbb{N} \rightarrow \mathbb{N}$  with  $g(n)/n \rightarrow 0$  as  $n \rightarrow \infty$ . We say that a shift space  $X$  has the *almost specification property* if there is a mistake function  $g: \mathbb{N} \rightarrow \mathbb{N}$  such that for any  $u, w \in \mathcal{B}(X)$  there is a word  $v \in \mathcal{B}(X)$  satisfying  $v = u'w'$ , where  $|u| = |u'|$ ,  $|v| = |v'|$ , and the following inequalities hold:

$$\begin{aligned} |\{1 \leq j \leq n : u_j \neq u'_j\}| &\leq g(|u|), \\ |\{1 \leq j \leq n : w_j \neq w'_j\}| &\leq g(|w|). \end{aligned}$$

We say that a shift space  $X$  has the *weak specification property* if there is a mistake function  $g$  such that for any  $u, w \in \mathcal{B}(X)$  there is  $v$  with  $uvw \in \mathcal{B}(X)$  satisfying  $|v| = g(|w|)$ . The weak and almost specification properties are independent of each other: neither one implies the other (see [31]). Additionally, in contrast to the classical specification property, the weaker versions do not imply the uniqueness of the measure of maximal entropy (see [31, 37]).

*Remark 3.5.* Using Proposition 3.3, we see that the proximal shift space constructed in Example 5.4 and the minimal shift space from §5.2 (see Theorem 5.11) are  $\bar{d}_M$ -stable. This answers Austin’s question affirmatively. We note that these examples do not have any of the specification properties mentioned in Corollary 3.4, because any of these specification properties implies that a shift having one of them and positive topological entropy has many disjoint minimal proper subsets. Hence, such a shift space is neither minimal nor proximal.

We list certain properties of  $\bar{d}_M$ -stable shift spaces. But first we recall some definitions. Let  $X \subseteq \mathcal{A}^\infty$  be a shift space. The *measure center* of  $X$  is the smallest shift space  $X^+ \subseteq X$  such that  $\mu(X^+) = 1$  for every  $\mu \in \mathcal{M}_\sigma(X)$ . In other words,  $X^+$  is the smallest subshift of  $X$  containing supports of all invariant measures on  $X$ . The measure center is determined by the language of all words in  $\mathcal{B}(X)$  whose cylinders have positive measure for at least one measure in  $\mathcal{M}_\sigma(X)$ , that is,

$$\mathcal{B}(X^+) = \{w \in \mathcal{B}(X) \mid \text{there exists } \mu \in \mathcal{M}_\sigma(X) : \mu[w] > 0\}.$$

The following observation follows directly from the definitions.

**PROPOSITION 3.6.** *A shift space  $X$  is  $\bar{d}_M$ -stable if and only if its measure center  $X^+$  is  $\bar{d}_M$ -stable.*

It is possible that  $X^+ = X$  holds true for every  $\bar{d}_M$ -stable shift space. In the next proposition, we note that  $\bar{d}_M$ -stability of  $X$  implies that the canonical Markov approximations of  $X$  form a sequence of shift spaces satisfying (3). That is,  $\bar{d}_M$ -stability implies that  $X$  has a property we call  *$\bar{d}_M$ -approachability*.

**PROPOSITION 3.7.** *If  $X \subseteq \mathcal{A}^\infty$  is a  $\bar{d}_M$ -stable shift space, then*

$$\bar{d}_M^H(\mathcal{M}_\sigma(X_n^M), \mathcal{M}_\sigma(X)) \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

*Proof.* Fix  $\varepsilon > 0$ . Let  $\delta > 0$  and  $n \geq 1$  be such that if  $\nu \in \mathcal{M}_\sigma(\mathcal{A}^\infty)$  and  $\nu(U_n(X)) > 1 - \delta$ , then  $\bar{d}_M(\mathcal{M}_\sigma(X), \nu) < \varepsilon$ . Fix  $m \geq n$  and  $\mu \in \mathcal{M}_\sigma(X_m^M)$ . Since  $X_m^M \subseteq U_n(X)$  we see that  $\bar{d}_M(\mu, \mathcal{M}_\sigma(X)) < \varepsilon$ , so  $\bar{d}_M^H(\mathcal{M}_\sigma(X_m^M), \mathcal{M}_\sigma(X)) < \varepsilon$  for  $m \geq n$ . Hence,  $\bar{d}_M^H(\mathcal{M}_\sigma(X_n^M), \mathcal{M}_\sigma(X)) \rightarrow 0$  as  $n \rightarrow \infty$ .  $\square$

The converse is not true, that is, the condition  $\bar{d}_M^H(\mathcal{M}_\sigma(X_n^M), \mathcal{M}_\sigma(X)) \rightarrow 0$  as  $n \rightarrow \infty$  does not imply  $\bar{d}_M$ -stability. Any shift of finite type  $X$  with non-transitive  $X^+$  is a counterexample. A concrete example is the shift space  $X$  over  $\{0, 1\}$  consisting of only two fixed points  $0^\infty$  and  $1^\infty$ . It is a binary shift of finite type with  $01$  and  $10$  as the forbidden words. Such a shift space (trivially) satisfies  $\bar{d}_M^H(\mathcal{M}_\sigma(X_n^M), \mathcal{M}_\sigma(X)) \rightarrow 0$  as  $n \rightarrow \infty$ , but is not  $\bar{d}_M$ -stable because of Proposition 3.9 below. Nevertheless, adding an assumption of chain mixing to  $\bar{d}_M$ -approachability, we obtain  $\bar{d}_M$ -stability.

**PROPOSITION 3.8.** *If  $X \subseteq \mathcal{A}^\infty$  is a chain-mixing shift space with  $\bar{d}_M^H(\mathcal{M}_\sigma(X_n^M), \mathcal{M}_\sigma(X)) \rightarrow 0$  as  $n \rightarrow \infty$ , then  $X$  is  $\bar{d}_M$ -stable.*

*Proof.* Fix  $\varepsilon > 0$ . Find  $N > 0$  such that  $\bar{d}_M^H(\mathcal{M}_\sigma(X_N^M), \mathcal{M}_\sigma(X)) \leq \varepsilon/2$ . Since  $X$  is chain mixing,  $X_N^M$  is a topologically mixing shift of finite type so it also has the  $\bar{d}$ -shadowing property (see [23]). Hence,  $X_N^M$  is  $\bar{d}_M$ -stable by Proposition 3.3. Let  $\delta > 0$  and  $m \geq 1$  be such that if  $\nu \in \mathcal{M}_\sigma(\mathcal{A}^\infty)$  and  $\nu(U_m(X_N^M)) > 1 - \delta$ , then  $\bar{d}_M(\mathcal{M}_\sigma(X_N^M), \nu) < \varepsilon/2$ . Without loss of generality we have  $m \geq N$ , so  $U_m(X) \subseteq U_m(X_N^M)$ . Let  $\mu \in \mathcal{M}_\sigma e(\mathcal{A}^\infty)$  be such that  $\mu(U_m(X)) > 1 - \delta$ . Then  $\mu(U_m(X_N^M)) \geq \mu(U_m(X)) > 1 - \delta$ . Hence,  $\bar{d}_M(\mu, \mathcal{M}_\sigma(X_N^M)) < \varepsilon/2$ , so there exists  $\mu' \in \mathcal{M}_\sigma e(X_N^M)$  such that  $\bar{d}_M(\mu, \mu') < \varepsilon/2$ . Since  $\bar{d}_M^H(\mathcal{M}_\sigma(X_N^M), \mathcal{M}_\sigma(X)) < \varepsilon/2$  there exists  $\xi \in \mathcal{M}_\sigma e(X)$  with  $\bar{d}_M(\xi, \mu) \leq \bar{d}_M(\xi, \mu') + \bar{d}_M(\mu', \mu) < \varepsilon$ .  $\square$

We do not know if  $\bar{d}_M$ -stability implies the stronger condition called  $\bar{d}$ -approachability (the sequence of canonical Markov approximations of our shift space is a sequence of shift spaces satisfying condition (2)), even if we assume that the shift space is chain transitive or chain mixing. The examples discussed in Proposition 4.4 suggest that it might not be the case.

Next, we note that  $\bar{d}_M$ -stability implies weak\* density of ergodic measures in  $\mathcal{M}_\sigma(X)$  and, as a consequence, entropy density (interestingly, we need to prove weak\* density first to obtain transitivity of  $X^+$  and obtain entropy density as a consequence of transitivity of  $X^+$ ). Hence, if  $X$  is a  $\bar{d}_M$ -stable shift space then the simplex of invariant measures  $\mathcal{M}_\sigma(X)$  is either a Poulsen simplex or a singleton. Note that the latter possibility can occur. As an example, take  $X = \{0^\infty\}$ .

**PROPOSITION 3.9.** *If  $X \subseteq \mathcal{A}^\infty$  is a  $\bar{d}_M$ -stable shift space, then  $X^+$  is transitive and ergodic measures are entropy dense in  $\mathcal{M}_\sigma(X)$ .*

*Proof.* We first prove that  $\mathcal{M}_\sigma e(X)$  is weak\* dense in  $\mathcal{M}_\sigma(X)$ . To prove the density of ergodic measures it is enough to show that for every  $\mu_1, \mu_2 \in \mathcal{M}_\sigma e(X)$  the measure  $\frac{1}{2}(\mu_1 + \mu_2)$  is a limit of a sequence of ergodic measures in  $\mathcal{M}_\sigma e(X)$ . Fix  $\varepsilon > 0$ . Let  $\delta > 0$  and  $n \geq 1$  be such that if  $\mu(U_n(X)) > 1 - \delta$ , then  $\bar{d}_M(\mathcal{M}_\sigma(X), \mu) < \varepsilon$ . Since the ergodic measures of  $\mathcal{A}^\infty$  are dense in  $\mathcal{M}_\sigma(\mathcal{A}^\infty)$ , when the latter space is endowed

with the weak\* topology we can find  $\nu \in \mathcal{M}_\sigma(\mathcal{A}^\infty)$  with  $D(\frac{1}{2}(\mu_1 + \mu_2), \nu)$  as small as necessary. Here  $D$  stands for any metric on  $\mathcal{M}_\sigma(\mathcal{A}^\infty)$  compatible with the weak\* topology. In particular, we may assume that  $D(\frac{1}{2}(\mu_1 + \mu_2), \nu)$  is sufficiently small to guarantee  $\nu(U_n(X)) > 1 - \delta$ . By  $\bar{d}_M$ -stability, there is  $\xi \in \mathcal{M}_\sigma(X)$  such that  $\bar{d}_M(\nu, \xi) < \varepsilon$ . Since  $\nu$  is ergodic, we can assure that  $\xi$  is an ergodic measure. By the triangle inequality,  $D(\frac{1}{2}(\mu_1 + \mu_2), \xi) \leq D(\frac{1}{2}(\mu_1 + \mu_2), \nu) + D(\nu, \xi)$ . Since  $\nu$  and  $\xi$  can be arbitrarily close in  $\bar{d}_M$ , they can also be arbitrarily close in  $D$ . Hence,  $\xi$  can be arbitrarily close to  $\frac{1}{2}(\mu_1 + \mu_2)$  in  $D$  and the ergodic measures must be weak\* dense. Now, transitivity of  $X^+$  follows easily from weak\* density of ergodic measures (see Proposition 6.4 in [16] for details). By Proposition 3.6 the measure center  $X^+$  is a  $\bar{d}_M$ -stable shift space. Now, Proposition 3.7 implies  $\bar{d}_M^H(\mathcal{M}_\sigma((X^+)_n^M), \mathcal{M}_\sigma(X^+)) \rightarrow 0$  as  $n \rightarrow \infty$ . Transitivity of  $X^+$  implies that its Markov approximations are entropy-dense shift spaces. Hence, ergodic measures are entropy dense in  $\mathcal{M}_\sigma(X^+)$  by [23], but clearly  $\mathcal{M}_\sigma(X) = \mathcal{M}_\sigma(X^+)$ .  $\square$

**PROPOSITION 3.10.** *If  $X$  is a strictly ergodic  $\bar{d}_M$ -stable shift space, then the unique invariant measure on  $X$  is isomorphic to an odometer.*

*Proof.* Assume that  $X$  is a strictly ergodic infinite shift space. Let  $\nu$  be its unique ergodic invariant measure, that is,  $\mathcal{M}_\sigma(X) = \{\nu\}$ . Hence, for every  $n \geq 1$  the Markov approximation  $X_n^M$  of  $X$  is an uncountable shift of finite type. In particular, for every  $n \geq 1$  the simplex  $\mathcal{M}_\sigma(X_n^M)$  contains infinitely many periodic ergodic invariant measures (measures concentrated on periodic orbits). Now assume that  $X$  is  $\bar{d}_M$ -stable, so  $\bar{d}_M^H(\mathcal{M}_\sigma(X_n^M), \mathcal{M}_\sigma(X)) \rightarrow 0$  as  $n \rightarrow \infty$ . It follows that  $\lim_{n \rightarrow \infty} \bar{d}_M(\mu_n^{\text{per}}, \nu) = 0$  for any choice of periodic ergodic measures  $\mu_n^{\text{per}} \in \mathcal{M}_\sigma(X_n^M)$ . In particular, one may take measures on periodic points whose primary periods tend to infinity. A measure that is a  $\bar{d}_M$ -limit of such a sequence of periodic measures must be isomorphic to a Haar measure on some odometer (this result is implicit in [40] and follows directly from [2, Theorem 1.7]; it does not need the invertibility assumption).  $\square$

*Remark 3.11.* (Some open questions) The results in the present section do not provide a complete picture of connections between the notions of  $\bar{d}_M$ -stability and  $\bar{d}$ -approachability. For example, we were unable to answer the following questions. First, can a non-trivial periodic orbit be  $\bar{d}_M$ -stable shift space? Second, can a strictly ergodic infinite shift space be  $\bar{d}_M$ -stable? Third, is every  $\bar{d}_M$ -stable system topologically mixing on its measure center? Finally, can a shift space  $X$  such that  $X^+ \neq X$  be  $\bar{d}_M$ -stable?

#### 4. Comparing $\bar{d}_M^H$ with $\bar{d}^H$

If  $(Z, \rho)$  is a bounded complete metric space, then so is  $(\text{CL}(Z), \rho^H)$  (see [18, §2.15]). Hence, the Hausdorff metric  $\bar{d}_M^H$  induced on  $\text{CL}(\mathcal{M}_\sigma(\mathcal{A}^\infty))$  by  $\bar{d}_M$  is complete and the Cauchy condition provides a criterion for convergence of a sequence  $(\mathcal{M}_\sigma(X_k))_{k=1}^\infty$ , where  $X_k \subseteq \mathcal{A}^\infty$  is a shift space for every  $k \geq 1$ . But even if we know that  $(\mathcal{M}_\sigma(X_k))_{k=1}^\infty$  converges in  $\bar{d}_M$  to some  $\mathcal{M} \in \text{CL}(\mathcal{M}_\sigma(\mathcal{A}^\infty), \bar{d}_M)$ , it is not clear if there exists a shift space  $X \subseteq \mathcal{A}^\infty$  such that  $\mathcal{M} = \mathcal{M}_\sigma(X)$ . We provide an example showing that this need not be the case at the end of this section. But first we demonstrate that shift spaces  $X$  and  $Y$

with the  $\bar{d}$ -shadowing property are  $\bar{d}^H$  close if and only if they are  $\bar{d}_M^H$  close. On the other hand, without the  $\bar{d}$ -shadowing property the inequality in (5) can be strict. We use a variant of Oxtoby's construction of non-uniquely ergodic minimal Toeplitz subshift to show that for every  $\delta > 0$  there are shift spaces  $X$  and  $Y$  such that  $\bar{d}_M^H(\mathcal{M}_\sigma(X), \mathcal{M}_\sigma(Y)) < \delta$  but  $\bar{d}^H(X, Y) > 1 - \delta$ . Finally, we show that if the measure center of a shift space  $X$  has the  $\bar{d}$ -shadowing property, then so also does  $X$ .

**THEOREM 4.1.** *If  $X$  and  $Y$  are shift spaces over  $\mathcal{A}$  with the  $\bar{d}$ -shadowing property such that*

$$\bar{d}_M^H(\mathcal{M}_\sigma(X), \mathcal{M}_\sigma(Y)) < \varepsilon^2$$

for some  $\varepsilon > 0$ , then  $\bar{d}^H(X, Y) < 7\varepsilon$ .

*Proof.* Fix  $x \in X$ . Use  $\bar{d}$ -shadowing of  $Y$  to find  $s \in \mathbb{N}$  such that for every sequence  $\{w^{(j)}\}_{j=1}^\infty$  of words in  $\mathcal{B}(Y)$  with  $|w^{(j)}| \geq s$  for every  $j \geq 1$ , there exists  $y \in Y$  such that

$$\bar{d}(w^{(1)}w^{(2)}w^{(3)} \dots, y) < \varepsilon. \tag{6}$$

Pick  $m$  such that  $s < m\varepsilon$ . By [9, Theorem 3.4] we find  $l \geq m$  such that  $x$  can be decomposed into an infinite concatenation of blocks, that is, we can write

$$x = A^{(1)}B^{(1)}A^{(2)}B^{(2)} \dots, \tag{7}$$

and the blocks  $A^{(1)}, A^{(2)}, \dots$  and  $B^{(1)}, B^{(2)}, \dots$  satisfy the following properties.

- For every  $i \geq 1$  we have  $m \leq |B^{(i)}| \leq l$ .
- For every  $i \geq 1$  there exists an ergodic measure  $\mu^{(i)} \in \mathcal{M}_\sigma e(X)$  such that

$$d^*(B^{(i)}, \mu^{(i)}) = \sum_{k=1}^\infty 2^{-k} \sum_{w \in \mathcal{A}^k} |\text{freq}(w, B^{(i)}) - \mu^{(i)}([w])| < \frac{\varepsilon}{2^s}, \tag{8}$$

where

$$\text{freq}(w, B^{(i)}) = \begin{cases} \frac{|\{1 \leq j \leq |B^{(i)}| - l + 1 : B^{(i)}_{[j, j+l)} = w\}|}{|B^{(i)}|} & \text{if } |w| = l \leq |B^{(i)}|, \\ 0 & \text{otherwise.} \end{cases} \tag{9}$$

- The set of coordinates of  $x$  which belong to the block  $A^{(i)}$  in (7) for some  $i \geq 1$  has upper Banach density smaller than  $\varepsilon$ . In particular, we have

$$\limsup_{n \rightarrow \infty} \frac{|A^{(1)}| + \dots + |A^{(n)}|}{|A^{(1)}B^{(1)}| + \dots + |A^{(n)}B^{(n)}|} < \varepsilon. \tag{10}$$

We now use the assumption  $\bar{d}_M^H(\mathcal{M}_\sigma(X), \mathcal{M}_\sigma(Y)) < \varepsilon^2$  and for every  $i \geq 1$  we find an ergodic measure  $\nu^{(i)} \in \mathcal{M}_\sigma(Y)$  such that

$$\bar{d}_M(\mu^{(i)}, \nu^{(i)}) < \varepsilon^2. \tag{11}$$

Following Shields [41], for every  $\mu, \nu \in \mathcal{M}_\sigma(\mathcal{A}^\infty)$  and  $n \geq 1$  we define  $J_n = J_n(\mu, \nu)$  to be the set of measures  $\lambda_n$  on  $\mathcal{A}^n \times \mathcal{A}^n$  endowed with the powerset  $\sigma$ -algebra such that for every  $u, w \in \mathcal{A}^n$  we have  $\mu[u] = \lambda_n(\{u\} \times \mathcal{A}^n)$  and  $\nu[w] = \lambda_n(\mathcal{A}^n \times \{w\})$ . For  $\alpha > 0$

we let  $\Delta_n(\alpha) = \{(u, w) \in \mathcal{A}^n \times \mathcal{A}^n : d_{\text{Ham}}(u, w) \leq \alpha\}$ . Finally, for  $\mu, \nu \in \mathcal{M}_\sigma(\mathcal{A}^\infty)$  we write

$$d_n^*(\mu, \nu) = \max_{\lambda_n \in J_n(\mu, \nu)} \min\{\alpha > 0 : \lambda_n(\Delta_n(\alpha)) \geq 1 - \alpha\}.$$

By [41, §I.9], in particular [41, Lemma I.9.12], we see that  $\bar{d}_{\mathcal{M}}(\mu^{(i)}, \nu^{(i)}) < \varepsilon^2$  implies that for every  $n \geq 1$  we have  $d_n^*(\mu^{(i)}, \nu^{(i)}) < \varepsilon$ . Hence, for every  $n \geq 1$  and  $i \geq 1$  there exists  $\lambda_n^{(i)} \in J_n(\mu^{(i)}, \nu^{(i)})$  such that  $\lambda_n^{(i)}(\Delta_n(\varepsilon)) > 1 - \varepsilon$ . Consider the set

$$G_n^{(i)} = \{u \in \mathcal{B}_n(X) : \text{there exists } w \in \mathcal{B}_n(Y) \text{ with } \lambda_n^{(i)}(\{(u, w)\} \cap \Delta_n(\varepsilon)) > 0\}.$$

It follows that for every  $u \in G_n^{(i)}$  we can pick  $w_n^{(i)}(u)$  such that  $d_{\text{Ham}}(u, w_n^{(i)}(u)) < \varepsilon$  and  $\lambda_n^{(i)}(\{(u, w_n^{(i)}(u))\}) > 0$ . In particular,  $\nu^{(i)}[w_n^{(i)}(u)] > 0$  and hence  $w_n^{(i)}(u) \in \mathcal{B}_n(Y)$ . In addition, we clearly have

$$\lambda_n^{(i)}(\Delta_n(\varepsilon)) = \lambda_n^{(i)}(\Delta_n(\varepsilon) \cap (G_n^{(i)} \times \mathcal{A}^n)) > 1 - \varepsilon.$$

By an abuse of notation, for  $i \geq 1$  and  $n \geq 1$ , by  $\bigcup G_n^{(i)}$  we will understand  $\bigcup\{[u] : u \in G_n^{(i)}\}$ .

It follows that for every  $n \geq 1$  and  $i \geq 1$  we have

$$\mu^{(i)}\left(\bigcup G_n^{(i)}\right) = \lambda_n^{(i)}(G_n^{(i)} \times \mathcal{A}^n) \geq \lambda_n^{(i)}(\Delta_n(\varepsilon) \cap G_n^{(i)} \times \mathcal{A}^n) > 1 - \varepsilon. \tag{12}$$

For each  $i \geq 1$  we take  $n = s$  and consider  $G_s^{(i)} \subseteq \mathcal{B}_s(X)$ . Note that (12) implies that  $\mu^{(i)}(\bigcup G_s^{(i)}) > 1 - \varepsilon$ . In analogy with (9), we define  $\text{freq}(G_s^{(i)}, B^{(i)})$  to be number of coordinates in  $B^{(i)}$  where some word from  $G_s^{(i)}$  appears in  $B^{(i)}$  divided by the length of  $B^{(i)}$ , that is,

$$\text{freq}(G_s^{(i)}, B^{(i)}) = \frac{|\{1 \leq j \leq |B^{(i)}| - s + 1 : B_{[j, j+s]}^{(i)} \in G_s^{(i)}\}|}{|B^{(i)}|}.$$

We easily see that

$$\text{freq}(G_s^{(i)}, B^{(i)}) = \sum_{w \in G_s^{(i)}} \text{freq}(w, B^{(i)}).$$

Let  $P$  be the set of coordinates in  $B^{(i)}$  covered by occurrences of words from  $G_s^{(i)}$  in  $B^{(i)}$ , that is,

$$P = \{1 \leq p \leq |B^{(i)}| : \text{there exists } 1 \leq j \leq |B^{(i)}| - s + 1 \text{ with } B_{[j, j+s]}^{(i)} \in G_s^{(i)} \text{ and } j \leq p < j + s\}.$$

We clearly have

$$\text{freq}(G_s^{(i)}, B^{(i)}) \leq \frac{|P|}{|B^{(i)}|}. \tag{13}$$

Furthermore,

$$\left| \mu^{(i)}\left(\bigcup G_s^{(i)}\right) - \text{freq}(G_s^{(i)}, B^{(i)}) \right| \leq \sum_{w \in G_s^{(i)}} |\text{freq}(w, B^{(i)}) - \mu^{(i)}([w])|. \tag{14}$$

Using  $d^*(B^{(i)}, \mu^{(i)}) < \varepsilon/2^s$  (cf. (8)), we obtain

$$\sum_{w \in G_s^{(i)}} |\text{freq}(w, B^{(i)}) - \mu^{(i)}([w])| \leq \sum_{w \in \mathcal{A}^s} |\text{freq}(w, B^{(i)}) - \mu^{(i)}([w])| \tag{15}$$

$$\leq 2^s d^*(B^{(i)}, \mu^{(i)}) < \varepsilon. \tag{16}$$

Combining (14) and (15), we get

$$\left| \mu\left(\bigcup G_s^{(i)}\right) - \text{freq}(G_s^{(i)}, B^{(i)}) \right| \leq \varepsilon. \tag{17}$$

Combining (12) and (17), we see that

$$1 - 2\varepsilon \leq \frac{|P|}{|B^{(i)}|}.$$

It follows that there exists a decomposition of  $B^{(i)}$  such that

$$B^{(i)} = v^{(i,1)} u^{(i,1)} v^{(i,2)} u^{(i,2)} \dots v^{(i,\kappa(i))} u^{(i,\kappa(i))} v^{(i,\kappa(i)+1)}, \tag{18}$$

where  $\kappa(i)$  is some (large) positive integer,  $u^{(i,j)} \in G_s^{(i)}$ , and  $v^{(i,j)} \in \mathcal{B}(X) \setminus G_s^{(i)}$ . Furthermore,

$$\frac{|v^{(i,1)}| + |v^{(i,2)}| + \dots + |v^{(i,\kappa(i))}| + |v^{(i,\kappa(i)+1)}|}{|B^{(i)}|} \leq \frac{|B^{(i)}| - |P|}{|B^{(i)}|} \leq 2\varepsilon. \tag{19}$$

(Actually, (19) tells us that  $v^{(i,j)}$  is an empty word for many  $js$ .)

We claim that if for each  $i \geq 1$  we find blocks  $\hat{w}^{(i,j)} \in \mathcal{B}(Y)$  ( $j = 1, \dots, \kappa(i)$ ) with  $|\hat{w}^{(i,j)}| \geq s$  for each  $j$  and these blocks satisfy

$$|A^{(i)} B^{(i)}| = |\hat{w}^{(i,1)} \dots \hat{w}^{(i,\kappa(i))}| \tag{20}$$

and

$$d_{\text{Ham}}(A^{(i)} B^{(i)}, \hat{w}^{(i,1)} \dots \hat{w}^{(i,\kappa(i))}) \leq \frac{|A^{(i)}| + 2\varepsilon|B^{(i)}| + |v^{(i,1)}| + \dots + |v^{(i,\kappa(i)+1)}| + s}{|A^{(i)} B^{(i)}|}, \tag{21}$$

then the proof will be complete. Indeed, assume that for each  $i \geq 1$  we have found blocks  $\hat{w}^{(i,1)}, \dots, \hat{w}^{(i,\kappa(i))}$  satisfying (20) and (21), each of them of length at least  $s$ . We set  $\hat{y}$  to be infinite concatenation of  $\hat{w}^{(i,1)}, \dots, \hat{w}^{(i,\kappa(i))}$ , where  $i = 1, 2, \dots$ , that is,

$$\hat{y} = \hat{w}^{(1,1)} \dots \hat{w}^{(1,\kappa(1))} \hat{w}^{(2,1)} \dots \hat{w}^{(2,\kappa(2))} \dots \hat{w}^{(i,1)} \dots \hat{w}^{(i,\kappa(i))} \dots$$

By (7), (10), (19), (20), and (21) such  $\hat{y}$  satisfies

$$\bar{d}(x, \hat{y}) \leq 6\varepsilon. \tag{22}$$

Note that for every  $i \geq 1$  and for every  $1 \leq j \leq \kappa(i)$  we have  $\hat{w}^{(i,j)} \in \mathcal{B}(Y)$  and  $|\hat{w}^{(i,j)}| \geq s$ , so the  $\bar{d}$ -shadowing property guarantees there is  $y \in Y$  with

$$\bar{d}(y, \hat{y}) < \varepsilon. \tag{23}$$

By (22) and (23) we have  $\bar{d}(x, y) < 7\varepsilon$  as needed.

It remains to find appropriate  $\hat{w}^{(i,1)}, \dots, \hat{w}^{(i,\kappa(i))}$  for each  $i \geq 1$ . To this end we fix  $i \geq 1$  and for each  $2 \leq j \leq \kappa(i)$  we take  $u^{(i,j)}$  in equation (18) to find  $w^{(i,j)} = w^{(i,j)}(u^{(i,j)}) \in \mathcal{B}_s(Y)$  with  $d_{\text{Ham}}(u^{(i,j)}, w^{(i,j)}) \leq \varepsilon$ . Now, for  $j = 1$  we set  $t(i) = |A^{(i)}| + |v^{(i,1)}| + |w^{(i,1)}| + |v^{(i,2)}| \geq s$  and we pick any  $\hat{w}^{(i,1)} \in \mathcal{B}_{t(i)}(Y)$ . For  $2 \leq j \leq \kappa(i)$  we simply extend each  $w^{(i,j)}$  to a word  $\hat{w}^{(i,j)} = w^{(i,j)}\hat{v}^{(i,j+1)} \in \mathcal{B}_{|w^{(i,j)}|+|v^{(i,j+1)}|}(Y)$ , where  $|\hat{v}^{(i,j+1)}| = |v^{(i,j+1)}|$ . We clearly have  $|\hat{w}^{(i,j)}| \geq |w^{(i,j)}| = s$  and

$$d_{\text{Ham}}(\hat{w}^{(i,j)}, u^{(i,j)}v^{(i,j+1)}) \leq \frac{|v^{(i,j+1)}| + \varepsilon|w^{(i,j)}|}{|\hat{w}^{(i,j)}|}.$$

It is now straightforward to see that the blocks  $\hat{w}^{(i,1)}, \dots, \hat{w}^{(i,\kappa(i))}$  satisfy (20) and (21). To finish the proof, reverse the roles of  $X$  and  $Y$ . □

**4.1. Examples.** In this subsection, we explore what happens if we abandon the assumption of  $\bar{d}$ -shadowing. First, we prove Proposition 4.2 showing that the  $\bar{d}_{\mathcal{M}}^H$  limit of a sequence of simplices of invariant measures need not be a simplex of all invariant measures of some subshift. In particular, the shift spaces  $X_k$  consisting of a single periodic orbit that have been constructed in the course of the proof of Proposition 4.2 do not converge in  $\bar{d}^H$ -distance to a shift space, as their convergence to a shift  $X$  would imply that the limit of the corresponding simplices would be  $\mathcal{M}_\sigma(X)$ ; see [23].

**PROPOSITION 4.2.** *For every alphabet  $\mathcal{A}$  there exists a sequence of transitive finite shifts  $(X_k)_{k=1}^\infty$  such that for some ergodic fully supported measure  $\mu \in \mathcal{M}_\sigma e(\mathcal{A}^\infty)$  we have*

$$\bar{d}_{\mathcal{M}}^H(\mathcal{M}_\sigma(X_k), \{\mu\}) \rightarrow 0 \quad \text{as } k \rightarrow \infty.$$

*In particular, there does not exist a shift space  $X$  such that  $\{\mu\} = \mathcal{M}_\sigma(X)$ .*

*Proof.* We order the non-empty words over  $\mathcal{A}$  into a sequence  $(W_k)_{k=0}^\infty$ . Let  $(\delta_k)_{k=1}^\infty$  be a sequence of positive reals such that

$$\sum_{k=1}^\infty \delta_k < \frac{1}{2}. \tag{24}$$

We inductively define words  $(V_k)_{k=0}^\infty$  by

$$\begin{aligned} V_0 &= W_0 \\ V_{k+1} &= V_k^{a_{k+1}} W_{k+1} 1^{b_{k+1}} \quad \text{for } k \geq 0, \end{aligned}$$

in such a way that  $b_{k+1} \geq 0$  is the smallest number such that  $|W_{k+1}|1^{b_{k+1}}$  is a multiple of  $|V_k|$  and  $a_{k+1} \geq 1$  is chosen so that the following inequality holds true:

$$\frac{|W_{k+1}| + b_{k+1}}{|V_{k+1}|} < \delta_{k+1}. \tag{25}$$



This implies that  $c_k = |V_{k+1}|/|V_k|$  is a positive integer. For  $k \geq 1$ , let  $x^{(k)} = V_k^\infty \in \mathcal{A}^\infty$  be a periodic point,  $X_k$  be its orbit, and  $\mu_k$  be the unique ergodic measure of the shift space  $X_k$ .

Note that  $\mu_n[W_k] \geq 1/|V_k| > 0$ , and for every  $n > k$  we have

$$\mu_n[W_k] \geq \frac{1}{|V_k|}((1 - \delta_{k+1})(1 - \delta_{k+2}) \cdots (1 - \delta_n)) > 0, \tag{26}$$

because the infinite product  $(1 - \delta_{k+1})(1 - \delta_{k+2}) \cdots (1 - \delta_n) \cdots$  converges to a non-zero limit by (24). We constructed the words  $V_k$  in such a way that  $|V_{k+1}|$  is a multiple of  $|V_k|$  and that  $V_k^{a_{k+1}}$  is a prefix of  $V_{k+1}$  for every  $k \geq 0$ ; hence, using (25), we see that for every  $k \geq q$  we have

$$\bar{d}(x^{(k+1)}, x^{(k)}) = d_{\text{Ham}}(V_k^{a_{k+1}}W_{k+1}1^{b_{k+1}}, V_k^{c_k}) \leq \delta_{k+1}.$$

It follows that  $x^{(k)}$  is a Cauchy sequence in the  $\bar{d}$  pseudometric. Since  $\bar{d}$  is a complete pseudometric the sequence  $x^{(k)}$  converges, so there is  $x \in \mathcal{A}^\infty$  such that

$$\lim_{k \rightarrow \infty} \bar{d}(x, x^{(k)}) = 0.$$

This point  $x$  must then be generic for an ergodic shift-invariant measure  $\mu$  such that  $\mu$  is the  $\bar{d}_{\mathcal{M}}$ -limit of the measures  $\mu_k$ . Since  $\bar{d}_{\mathcal{M}}$ -convergence implies weak\* convergence, the portmanteau theorem and (26) imply that

$$\mu[W_k] = \lim_{n \rightarrow \infty} \mu_n[W_k] \geq \frac{1}{|V_k|} \prod_{n=1}^{\infty} (1 - \delta_n) > 0.$$

Since  $\mu[W_k] > 0$  for every  $k \geq 0$ , the only subshift  $X$  such that  $\mu \in \mathcal{M}_\sigma(X)$  is the full shift. On the other hand,  $\mathcal{M}_\sigma(X_k) = \{\mu_k\}$  and  $\bar{d}_{\mathcal{M}}^H(\mathcal{M}_\sigma(X_k), \{\mu\}) = \bar{d}(\mu_k, \mu) \rightarrow 0$  as  $k \rightarrow \infty$ . We see that  $\mathcal{M}_\sigma(X_k) = \{\mu_k\}$  converges to  $\{\mu\}$  in  $\bar{d}_{\mathcal{M}}^H$  but there is no subshift  $X$  of  $\mathcal{A}^\infty$  such that  $\mathcal{M}_\sigma(X) = \{\mu\}$ . □

Our next goal is to show an example of a sequence  $(X_k)_{k=1}^\infty$  of shift spaces such that the  $\bar{d}_{\mathcal{M}}^H$ -limit of simplices  $\mathcal{M}_\sigma(X_k)$  exists and is a simplex of invariant measures of some shift space  $X$ , but the shift spaces  $X_k$  do not converge to  $X$  with respect to the  $\bar{d}$  pseudometric.

To find our examples we will adapt the construction of one-sided Oxtoby sequences. The original Oxtoby sequence generates a minimal non-uniquely ergodic Toeplitz subshift; see [7, 36, 43]. As parameters of this construction we need a sequence of positive integers  $(p_k)_{k=0}^\infty$ .

**Definition 4.3.** Let  $\mathbf{p} = (p_k)_{k=0}^\infty$  be a sequence of positive integers such that  $p_0 = 1$ , and for each  $k \geq 0$  we have that  $p_k$  divides  $p_{k+1}$  and  $p_{k+1}/p_k \geq 3$ . Let  $M_0 = \emptyset$ , and for  $k \geq 1$  define  $M_k = ([-p_k, p_k] + p_{k+1}\mathbb{N}) \cap \mathbb{N}$ . Note that for every  $i \in \mathbb{N}$  there exists a unique  $k = k(i) \geq 1$  such that  $i \in M_k \setminus \bigcup_{\ell=0}^{k-1} M_\ell$ . We define the *Oxtoby sequence with the scale  $\mathbf{p}$*  to be a binary sequence  $x(\mathbf{p}) \in \{0, 1\}^\infty$  such that  $x(\mathbf{p})_i = k(i) \bmod 2$ .

By Lemma 3.2 in [43], if  $x(\mathbf{p}) \in \{0, 1\}^\infty$  is an Oxtoby sequence with scale  $\mathbf{p}$  satisfying

$$\sum_{k=0}^\infty \frac{p_k}{p_{k+1}} < \infty,$$

then the orbit closure of  $x(\mathbf{p})$  in  $\{0, 1\}^\infty$  is a minimal shift space  $X(\mathbf{p})$  with exactly two ergodic invariant measures.

PROPOSITION 4.4. *If  $\mathbf{p} = (p_k)_{k \in \mathbb{N}}$  is a sequence of positive integers satisfying*

$$\sum_{k=1}^\infty \frac{2p_k}{p_{k+1}} < \delta, \tag{27}$$

*for some  $0 < \delta < \frac{1}{2}$ , then the minimal shift  $X(\mathbf{p})$  obtained as the orbit closure of the Oxtoby sequence with scale  $\mathbf{p}$  satisfies*

$$\bar{d}^H(X(\mathbf{p}), \{0^\infty, 1^\infty\}) > 1 - \delta \quad \text{and} \quad \bar{d}_M^H(\mathcal{M}_\sigma(X(\mathbf{p})), \mathcal{M}_\sigma(\{0^\infty, 1^\infty\})) < \delta.$$

*Proof.* Fix  $0 < \delta < 1$  and a sequence of positive integers  $\mathbf{p}$  as above. For simplicity we write  $x$  for the Oxtoby sequence  $x(\mathbf{p})$  defined taking  $\mathbf{p}$  as its scale and  $X$  for the associated minimal subshift  $X(\mathbf{p})$  (see Definition 4.3). By Lemma 3.2 in [43],  $\mathcal{M}_\sigma e(X) = \{\mu', \nu'\}$ . Let  $(M_k)_{k=1}^\infty$  be a sequence of sets as in Definition 4.3. Fix  $k \geq 1$  and consider the prefix  $x_{[0, p_{k+1}]}$ . Since  $p_{k+1}$  is a multiple of  $p_\ell$  for every  $\ell \leq k$ , using the structure of the sets  $M_k$  for  $k \geq \ell$  we get that

$$|M_\ell \cap [0, p_{k+1}]| = \frac{p_{k+1}}{p_{\ell+1}} \cdot 2p_\ell.$$

Hence,

$$\frac{|(\bigcup_{\ell=0}^k M_\ell) \cap [0, p_{k+1}]|}{p_{k+1}} \leq \sum_{\ell=0}^k \frac{2p_\ell}{p_{\ell+1}} \leq \delta.$$

But for every  $i \in [0, p_{k+1}] \setminus \bigcup_{\ell=0}^k M_\ell$  we have  $x_i = k + 1 \pmod 2$ . In other words, the Oxtoby sequence is constant for all indices  $i$  in  $[0, p_{k+1}] \setminus \bigcup_{\ell=0}^k M_\ell$  with the constant depending only on the parity of  $k$ . Since

$$\frac{|([0, p_{k+1}] \cap \mathbb{N} \setminus \bigcup_{\ell=0}^k M_\ell)|}{p_{k+1}} \geq 1 - \delta,$$

we see that for both  $\alpha = 0$  and  $\alpha = 1$  we have

$$\limsup_{k \rightarrow \infty} \frac{| \{0 \leq i < p_{k+1} : x_i = \alpha \} |}{p_{k+1}} \geq 1 - \delta.$$

Hence, for some invariant measure  $\mu, \nu \in \mathcal{M}_\sigma(X)$  we have  $\mu([0]) > 1 - \delta$  and  $\nu([1]) > 1 - \delta$ . By ergodic decomposition, these measures are convex combinations of the ergodic measures  $\mu'$  and  $\nu'$ . This implies that for one ergodic measure, say  $\mu'$ , we have  $\mu'([0]) > 1 - \delta$ , while for the other one we have  $\nu'([1]) > 1 - \delta$ .

A generic point for  $\mu'$  has density of 1 at most  $\delta$ , so it is at most  $\delta$  far away from  $0^\infty$ . Hence,  $\bar{d}_M(\mu', \delta_{0^\infty}) < \delta$ . Similarly,  $\bar{d}_M(\nu', \delta_{1^\infty}) < \delta$ . Therefore, the

$\bar{d}_{\mathcal{M}}^H$ -distance between sets of ergodic measures on  $X$  and  $\{0^\infty, 1^\infty\}$  is bounded by  $\delta$ . By [23, Lemma 14] (see (5) in the introduction) we have  $\bar{d}_{\mathcal{M}}^H(\mathcal{M}_\sigma(X(\mathbf{p})), \mathcal{M}_\sigma(\{0^\infty, 1^\infty\})) = \bar{d}_{\mathcal{M}}^H(\mathcal{M}_\sigma e(X(\mathbf{p})), \mathcal{M}_\sigma e(\{0^\infty, 1^\infty\})) < \delta$ . On the other hand, by the above calculations, we see that the Oxtoby sequence  $x$  satisfies  $\bar{d}(x, 1^\infty) > 1 - \delta$  and  $\bar{d}(x, 0^\infty) > 1 - \delta$ , which means that  $\bar{d}^H(X(\mathbf{p}), \{0^\infty, 1^\infty\}) > 1 - \delta$ .  $\square$

**COROLLARY 4.5.** *There exists a sequence  $(X_k)_{k=1}^\infty$  of minimal shift spaces such that for some shift space  $X$  we have  $\bar{d}_{\mathcal{M}}^H(\mathcal{M}_\sigma(X_n), \mathcal{M}_\sigma(X)) \rightarrow 0$  while  $\bar{d}^H(X_n, X) \rightarrow 1$  as  $n \rightarrow \infty$ .*

*Proof.* One can take  $X = \{0^\infty, 1^\infty\}$  and the sequence of minimal shift spaces  $X_n$  generated by Oxtoby sequences  $x^{(n)}$  constructed in the previous proposition for a sequence of  $\delta$ s going to zero.  $\square$

**4.2. On  $\bar{d}$ -shadowing on the measure center of  $X$ .** Let us recall that the measure center  $X^+$  of a shift space  $X$  is the smallest subshift of  $X$  containing supports of all invariant measures on  $X$ , that is,  $X^+$  is the smallest closed set such that  $\mu(X^+) = 1$  for every  $\mu \in \mathcal{M}_\sigma(X)$ .

Fix a word  $u$  over  $\mathcal{A}$  with  $|u| \geq k$ . Given a word  $w$  over  $\mathcal{A}$ , we define  $\gamma_w(u)$  to be the number of occurrences of  $w$  in  $u$ , that is,

$$\gamma_w(u) = |\{1 \leq j \leq |u| - k + 1 \mid u_j u_{j+1} \cdots u_{j+k-1} = w\}|.$$

Furthermore, for  $n \in \mathbb{N}$  with  $n \geq |w|$  we set  $\Gamma_w(n)$  to be largest number of occurrences of  $w$  among all words  $u$  of length  $n$ , that is,

$$\Gamma_w(n) = \max\{\gamma_w(u) \mid u \in \mathcal{B}_n(X)\}.$$

It is a straightforward consequence of the definition that  $\Gamma_w(uv) \leq \Gamma_w(u) + \Gamma_w(v) + |w| - 1$  for every  $u, v, w \in \mathcal{A}^*$ . Thus,

$$\Gamma_w(n + m) \leq \Gamma_w(n) + \Gamma_w(m) + |w| - 1 \tag{28}$$

for all  $n, m$ .

We define the maximum limiting frequency of  $w$  in  $X$  as

$$\Lambda_X(w) = \lim_{n \rightarrow \infty} \frac{1}{n} \Gamma_w(n). \tag{29}$$

The existence of the limit follows from the subadditivity of the function  $\Gamma'_w(n) = \Gamma_w(n) + |w| - 1$  and the fact that the difference between ratios  $\Gamma_w(n)/n$  and  $\Gamma'_w(n)/n$  goes to zero.

It is known that

$$\Lambda_X(w) = \max_{\mu \in \mathcal{M}_\sigma(X)} \mu[w] = \max_{\nu \in \mathcal{M}_\sigma e(X)} \nu[w]; \tag{30}$$

see [15, Ch. 3]. This means that  $w \in \mathcal{B}(X) \setminus \mathcal{B}(X^+)$  if and only if  $\Lambda_X(w) = 0$ , that is, for every  $\varepsilon > 0$  there exists  $N \in \mathbb{N}$  such that for all  $n \geq N$  and for all  $u \in \mathcal{B}_n(X)$  we have  $\gamma_w(u) \leq n\varepsilon$ .

**THEOREM 4.6.** *If  $X^+$  has the  $\bar{d}$ -shadowing property then so does  $X$ .*

*Proof.* Fix  $\varepsilon > 0$ . For  $\varepsilon/3$  we use the  $\bar{d}$ -shadowing property of  $X^+$  to find  $N_1$  such that for any sequence of words  $\{a^{(j)}\}_{j=1}^\infty$  in  $\mathcal{B}(X^+)$  with  $|a^{(j)}| \geq N_1$  there exists  $x \in X^+$  such that  $\bar{d}(x, a^{(1)}a^{(2)} \dots) < \varepsilon/3$ .

Now fix  $m \geq N_1$ . For  $w \notin \mathcal{B}_m(X^+)$  let  $N_w > 0$  be such that for all  $n \geq N_w$  and for all  $u \in \mathcal{B}_n(X)$  we have

$$\gamma_w(u) \leq \frac{n\varepsilon}{|A|^m 3m}. \tag{31}$$

Set  $N_0 = \max_{w \in \mathcal{B}_m(X)} N_w$ . We take  $N \in \mathbb{N}$  such that  $m/N < \varepsilon/6$  and  $N \geq \max\{N_0, N_1\}$ .

Let us take any  $j \geq 1$  and any  $w^{(j)} \in \mathcal{B}(X)$  such that  $|w^{(j)}| \geq N$ . Each  $w^{(j)}$  can be written as a concatenation of finite blocks as follows:

$$w^{(j)} = u_1^{(j)} u_2^{(j)} \dots u_{k(j)-1}^{(j)} u_{k(j)}^{(j)},$$

where  $|u_i^{(j)}| = m$  for  $1 \leq i < k(j)$  and  $m \leq |u_{k(j)}^{(j)}| < 2m$ . Using (31), we see that for  $1 \leq i \leq k(j) - 1$  the number of  $u_i^{(j)}$  which are not in  $\mathcal{B}_m(X^+)$  is bounded from above by  $\varepsilon|w^{(j)}|/3m$ .

For each  $j \geq 1$ , we create  $\bar{w}^{(j)}$  by replacing each  $u_i^{(j)} \notin \mathcal{B}(X^+)$  by some word  $\bar{v} \in \mathcal{B}_m(X^+)$  for  $1 \leq i < k(j)$  and replacing  $u_{k(j)}^{(j)}$  by some word  $\bar{v}^{(j)} \in \mathcal{B}(X^+)$  with  $|u_{k(j)}| = |\bar{v}^{(j)}|$  if  $u_{k(j)}^{(j)} \notin \mathcal{B}(X^+)$ . Therefore, we have

$$\bar{d}(\bar{w}^{(1)}\bar{w}^{(2)} \dots, w^{(1)}w^{(2)} \dots) < \frac{m\varepsilon}{3m} + \frac{2\varepsilon}{6} = \frac{2\varepsilon}{3}. \tag{32}$$

Notice that for each  $j \geq 1$  the word  $\bar{w}^{(j)}$  is a concatenation of words from  $\mathcal{B}(X^+)$  whose lengths are greater than or equal to  $m$ , so the same applies to  $\bar{w}^{(1)}\bar{w}^{(2)} \dots$ . Now we use the  $\bar{d}$ -shadowing property of  $X^+$  and we find  $x \in X^+ \subseteq X$  such that

$$\bar{d}(x, \bar{w}^{(1)}\bar{w}^{(2)} \dots) < \frac{\varepsilon}{3}. \tag{33}$$

It follows from (32) and (33) that  $\bar{d}(x, w^{(1)}w^{(2)} \dots) < \varepsilon$ , which concludes the proof.  $\square$

### 5. $\bar{d}$ -approachable examples of proximal and minimal shift spaces

Before presenting the details of our constructions, we first recall the necessary background.

An (oriented)  $\mathcal{A}$ -labeled (multi)graph is a triple  $G = (V, E, \tau)$ , where  $V$  is the (finite) set of vertices,  $E \subset V \times V$  is the edge set, and  $\tau: E \rightarrow \mathcal{A}$  is the label map. For each  $e \in E$  we write  $i(e), t(e) \in V$ , to denote, respectively, the initial vertex and the terminal vertex of  $e$ . We say that a sequence (finite or infinite) consisting of  $\ell \in \mathbb{N}_0 \cup \{\infty\}$  edges  $e_1, e_2, \dots$  in  $E$  is a *path* of length  $\ell$  in  $G$  if for every  $i < \ell$  we have that  $t(e_i) = i(e_{i+1})$ . A path  $e_1, e_2, \dots, e_\ell$  is *closed* if  $t(e_\ell) = i(e_1)$ .

Given an oriented  $\mathcal{A}$ -labeled graph  $G = (V, E, \tau)$ , we define the shift  $X_G \subseteq \mathcal{A}^\infty$  by reading off labels of all infinite paths in  $G$ . In other words,  $X_G$  is the set of all  $x \in \mathcal{A}^\infty$  such that  $x_i = \tau(e_{i+1})$  for each  $i \geq 0$  for some path  $e_1, e_2, \dots$  in  $G$ . We say that  $X$  is a *sofic shift* if there exists a labeled graph  $G = (V, E, \tau)$  such that  $X$  is presented by  $G$ , meaning that  $X = X_G$ . Every shift of finite type is sofic. A sofic shift is transitive if and only if it can be presented by a (strongly) connected graph (each pair of vertices can be connected

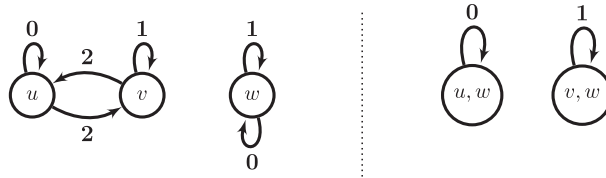


FIGURE 1. Two graphs (left) whose coupling (right) is disconnected. The first graph has no safe symbol.

by a path); see [32, Proposition 3.3.11]. A sofic shift is topologically mixing if and only if it can be presented by a (strongly) connected aperiodic graph, that is, the graph with two closed paths of coprime lengths.

To prove the properties of shift spaces resulting from our constructions we will use the following result, which is a direct corollary of a combination of Theorem 6 and Corollary 17 in [23].

**THEOREM 5.1.** *Let  $(X_n)_{n=1}^\infty$  be a decreasing sequence of mixing sofic shift spaces over  $\mathcal{A}$  such that*

$$\sum_{n=1}^\infty \bar{d}^H(X_n, X_{n+1}) < \infty.$$

*Then  $X = \bigcap_{n=1}^\infty X_n$  is a  $\bar{d}$ -approachable and chain-mixing shift space such that  $\bar{d}_M^H(\mathcal{M}_\sigma e(X_n), \mathcal{M}_\sigma e(X)) \rightarrow 0$  as  $n \rightarrow \infty$ . In particular,  $X$  satisfies  $\sigma(X) = X$  and has the  $\bar{d}$ -shadowing property.*

Let us recall that [23, Corollary 17] is stated using the lower density. We do not need such flexibility here, so we stated our result in terms of the  $\bar{d}$ -pseudodistance.

In fact, we would like to apply Theorem 5.1 to a sequence of mixing sofic shifts that is not decreasing. A natural way to apply Theorem 5.1 is to replace the sequence of sofic shifts  $(X_n)_{n=1}^\infty$  with the decreasing sequence of shift spaces  $(Y_n)_{n=1}^\infty$ , where  $Y_n := X_1 \cap \dots \cap X_n$  for  $n \in \mathbb{N}$ . It is easy to see that the shift  $Y_n$  thus defined is also sofic for  $n \in \mathbb{N}$ . Indeed, if for  $m = 1, 2, \dots, n$  a labeled graph  $G_m = (V_m, E_m, \tau_m)$  presents the sofic shift space  $X_m$ , then  $Y_n$  is a sofic shift presented by the graph  $G = (V, E, \tau)$ , where  $V = \prod_{1 \leq k \leq n} V_k$  and there is an edge from  $(v_1, \dots, v_n) \in V$  to  $(v'_1, \dots, v'_n) \in V$  with label  $\ell \in \mathcal{A}$  if and only if for every  $1 \leq k \leq n$  in the graph  $G_k$  there is an edge from  $v_k$  to  $v'_k$  labeled with  $\ell$ . We say that the graph  $G = (V, E, \tau)$  is the *coupling* of graphs  $G_k$ ,  $1 \leq k \leq n$ . Unfortunately,  $Y_n$  need not be mixing even if the shift spaces  $X_1, \dots, X_n$  are. The problem is that strong connectedness of the graphs  $G_k$  for  $1 \leq k \leq n$  need not ensure strong connectedness of their coupling  $G$ . Hence, the induced sofic shift  $Y_n$  need not be transitive and its ergodic measures need not be dense. Indeed, Figure 1 shows two graphs whose coupling is a sofic shift that is not transitive and whose ergodic measures are not dense in the set of all invariant measures. The same situation occurs for the sofic shifts represented in Figure 2.

However, under mild additional assumptions we can ensure that the sofic shifts  $Y_n$  are transitive. Let  $X$  be a sofic shift over the alphabet  $\mathcal{A}$  and pick a labeled graph

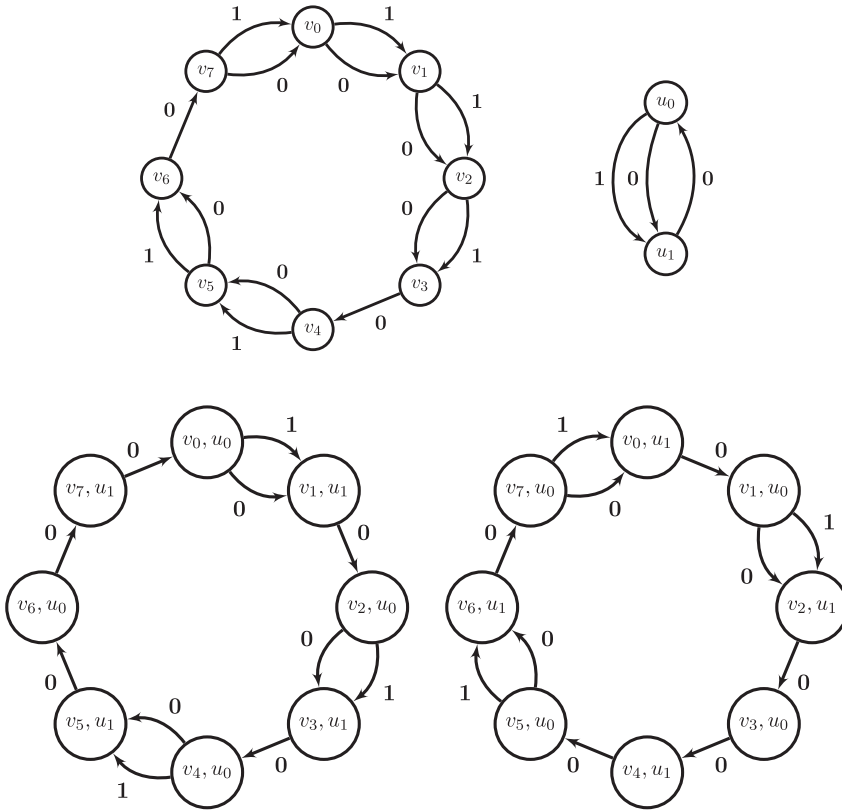


FIGURE 2. Two graphs whose periods are not coprime (above) and their coupling (below).

$G = (V, E, \tau)$  presenting  $X$ . We call a symbol  $b \in \mathcal{A}$  a *safe symbol* for  $X$  if for every edge  $e \in E$  that goes from a vertex  $v \in V$  to a vertex  $v' \in V$ , there is an edge  $e' \in E$  from  $v$  to  $v'$  with label  $\tau(e') = b$ . The *period* of a graph  $G = (V, E)$  is the greatest common divisor of the lengths of all cycles. Every sofic shift also has a period, which is the greatest common divisor of periods of its presentations through labeled graphs. A graph is *aperiodic* if it has period 1. Every sofic shift with period 1 has an aperiodic presentation. We will use a standard fact from graph theory, stating that if  $G = (V, E)$  is a strongly connected graph with period  $m$  then the set of lengths of paths between any pair of vertices  $u, v \in V$  is the set-theoretic difference of an infinite arithmetic progression with step  $m$  and a finite set.

Note that the two shifts in Figure 1 are aperiodic but lack a common safe symbol, while the two shifts in Figure 2 share a safe symbol 0 but have positive periods 8 and 2. Hence, neither assumption in the following proposition can be removed.

**PROPOSITION 5.2.** *If  $G_k = (V_k, E_k, \tau_k)$  for  $1 \leq k \leq n$  are strongly connected labeled graphs with a common safe symbol and pairwise coprime periods, then their coupling  $G$  is strongly connected.*

*Proof.* Let  $(v_1, v_2, \dots, v_k)$  and  $(v'_1, v'_2, \dots, v'_n)$  be any two vertices in  $G$ . For any  $1 \leq k \leq n$ , there exists a walk from  $v_k$  to  $v'_k$  in the graph  $G_k$ . Let  $\ell_k$  denote the length of one such walk and let  $m_k$  denote the period of  $G$ . For all sufficiently large  $j \in \mathbb{N}_0$  there exists a walk of length  $jm_k + \ell_k$  from  $v_k$  to  $v'_k$ . Since the graphs under consideration have a common safe symbol  $b$ , we may additionally assume that the edges in the aforementioned paths are all labeled with  $b$ . Since the integers  $m_k$  are coprime, there exist infinitely many integers  $\ell$  such that  $\ell \equiv \ell_k \pmod{m_k}$  for each  $1 \leq k \leq n$ . Consequently, we can find  $\ell_0$  such that for each  $1 \leq k \leq n$  there exists in  $G_k$  a walk from  $v_k$  to  $v'_k$  of length  $\ell_0$ , consisting only of edges labeled with  $b$ . This walk induces a walk of length  $\ell_0$  from  $(v_1, v_2, \dots, v_k)$  to  $(v'_1, v'_2, \dots, v'_n)$  in  $G$ .  $\square$

**COROLLARY 5.3.** *Let  $n \in \mathbb{N}$ . If  $X_k$ , for  $1 \leq k \leq n$ , are transitive sofic shifts with a common safe symbol and pairwise coprime periods, then  $Y = X_1 \cap \dots \cap X_n$  is a non-empty transitive sofic shift with a safe symbol. Furthermore, if for each  $1 \leq k \leq n$  the shift space  $X_k$  is topologically mixing, then  $Y$  is also topologically mixing.*

**5.1. A  $\bar{d}$ -approachable proximal shift space.** We will construct a  $\bar{d}$ -approachable and topologically mixing proximal shift. Furthermore, our example is hereditary and has positive topological entropy, hence its ergodic measures are entropy dense and its simplex of invariant measures is the Poulsen simplex. Assume that  $\mathcal{A} \subset \mathbb{N}_0$ . A shift space  $X \subseteq \mathcal{A}^\infty$  is *hereditary* if for every  $x \in X$  and  $y \in \mathcal{A}^\infty$  with  $y_i \leq x_i$  for all  $i \geq 0$  we have  $y \in X$ . The *hereditary closure*  $\tilde{X}$  of a shift space  $X$  is the smallest hereditary shift containing  $X$ . That is,  $\tilde{X}$  consists of all  $y \in \mathcal{A}^\infty$  such that there exists  $x = (x_i)_{i \geq 0} \in X$  with  $y_i \leq x_i$  for all  $i \geq 0$ . For more on hereditary shifts, see [29]. For some special properties of the simplex of invariant measures of hereditary shifts, see Remark 5.6 below. Recall that a hereditary shift space  $X$  is proximal if for every  $N > 0$  and  $x \in X$  the word  $0^N$  appears with bounded gaps in  $x$  (see the discussion of the Theorem B and the proof of Proposition 3.3 in [10]).

*Example 5.4.* For  $n \in \mathbb{N}$  we consider a sofic shift  $Z_n$  presented by a labeled graph  $G_n = (V_n, E_n, \tau_n)$  where  $V_n = \{v_0, v_1, \dots, v_{10^n-1}\}$  and edges and their labels are as follows:

- for every  $0 \leq k < 10^n$ , there is an edge from  $v_k$  to  $v_{k+1}$  with label 0, where  $v_{10^n} = v_0$ ;
- for every  $1 \leq k \leq 10^n - 2^n$ , there is an edge from  $v_k$  to  $v_{k+1}$  with label 1;
- there is an edge from  $10^n - 2^n$  to  $10^n - 2^n + 2$  with label 0.

Let  $Z = \bigcap_{n=1}^\infty Z_n$ . Then for each  $n \geq 1$  we have  $Z_{n+1} \not\subset Z_n$  and  $Z_n \not\subset Z_{n+1}$ . Figure 3 shows the graph with  $n = 1$ .

**PROPOSITION 5.5.** *The shift space  $Z$  defined in Example 5.4 is hereditary, topologically mixing, proximal, has positive topological entropy, and the ergodic measures  $\mathcal{M}_\sigma e(Z)$  are entropy dense in  $\mathcal{M}_\sigma(Z)$ .*

*Proof.* Since all  $Z_n$  are hereditary, 0 is their common safe symbol and the shift  $Z$  is hereditary as well. Furthermore, it contains a sequence where 1s appear with positive density, hence  $Z$  has positive entropy. For every  $k \in \mathbb{N}$ , in the graph  $G_k$  presenting  $Z_k$  in Example 5.4 there are two closed walks of coprime lengths  $10^k$  and  $10^k - 1$ , whence

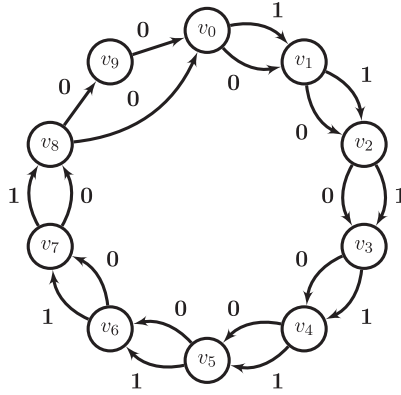


FIGURE 3. The graph from Example 5.4 with  $n = 1$ .

$G_k$  is aperiodic. By Corollary 5.3, for each  $n \in \mathbb{N}$  the intersection  $Y_n := Z_1 \cap \dots \cap Z_n$  is a topologically mixing sofic shift. In particular, the ergodic measures of  $Y_n$  are entropy dense [12]. By Corollary 5.1, in order to conclude that the ergodic measures on  $Z$  are entropy dense in  $\mathcal{M}_\sigma(Z)$ , it suffices to check that

$$\sum_{n=1}^{\infty} \bar{d}^H(Y_n, Y_{n+1}) < \infty.$$

Fix  $n \in \mathbb{N}$ . To bound  $\bar{d}^H(Y_n, Y_{n+1})$ , consider  $x \in Y_n$ . Define  $y \in \{0, 1\}^\infty$  by

$$y_j = \begin{cases} x_j & \text{if } j \bmod 10^{n+1} \in [0, 10^{n+1} - 2^{n+1}), \\ 0 & \text{if } j \bmod 10^{n+1} \in [10^{n+1} - 2^{n+1}, 10^{n+1}). \end{cases}$$

It is clear that  $y \in Z_{n+1}$ . Since  $Y_n$  is hereditary and  $x \in Y_n$ , we see that  $y$  belongs to  $Y_n$  as well. Thus,  $y \in Y_n \cap Z_{n+1} = Y_{n+1}$ . Furthermore,  $\bar{d}(x, y) \leq (\frac{1}{5})^{n+1}$ , and hence  $\bar{d}^H(Y_n, Y_{n+1}) \leq (\frac{1}{5})^{n+1}$ . This completes the proof of entropy density of ergodic measure.

Finally, we prove topological mixing for  $Z$ . Bearing in mind that  $Z$  and  $Z_n$  ( $n \in \mathbb{N}$ ) are hereditary, in order to show that  $Z$  is topologically mixing it is enough to show that for each  $u, v \in \mathcal{B}(Z_n)$  there exists  $M \in \mathbb{N}$  such that for all  $m \geq M$  and all  $n \in \mathbb{N}$  we have  $u0^m v \in \mathcal{B}(Z_n)$  (hence  $u0^m v \in \mathcal{B}(Z)$ ).

Fix  $u, v$ , and denote  $i = |u|$  and  $j = |v|$ . Let  $N \in \mathbb{N}$  be such that  $10^N > i + j + 2(2^n - 2)$ , for all  $n \geq N$ . Fix  $n \geq N$ . By the pigeonhole principle, for each  $m \in \mathbb{N}$  there exists  $t \in \mathbb{N}$  such that  $[t, t + i) \bmod 10^n \subset [0, 10^n - 2^n)$  and  $[t + i + m, t + i + m + j) \bmod 10^n \subset [0, 10^n - 2^n)$ . By the definition of  $Z_n$ , starting a path in the corresponding graph from  $v_t$  and ending in  $v_{t+i+m+j}$ , we can read  $u0^m v$  along the path, so the word belongs to  $\mathcal{B}_{i+m+j}(Z_n)$ . We have just proved that for all  $n \geq N$  and all  $m \in \mathbb{N}$ ,  $u0^m v \in \mathcal{B}(Z_n)$ . It remains to discuss the case when  $n \leq N$ .

Since for each  $n \in \mathbb{N}$  the system  $Z_n$  is mixing, there exists  $M_n \in \mathbb{N}$  such that  $u0^m v \in \mathcal{B}(Z_n)$  for all  $m \geq M_n$ . Set  $M$  as the maximum of  $M_n, n \leq N$ . Then for  $m \geq M, u0^m v \in \mathcal{B}(Z_n)$  for all  $n \leq N$ . But we have already proved the same conclusion for  $n \geq N$  too. This concludes the proof of topological mixing of  $Z$ .



Since  $Z$  is hereditary, to prove that it is proximal it is enough to show that for every  $N > 0$  and  $x \in Z$  the word  $0^N$  appears with bounded gaps in  $z$  (see [10]). But every  $z \in Z$  must belong to  $Y_n$  for every  $n \geq 1$ , so arbitrarily long blocks of 0s appear syndetically (i.e., with bounded gaps).  $\square$

*Remark 5.6.* Since  $Z$  is hereditary and proximal, some of the results from [23, 29] apply:  $Z$  is distributionally chaotic of type 2 [29, Theorem 23], but not of type 1 [29, Theorem 23] (cf. also [33]). Moreover, for each  $t > 0$ , the set of all ergodic invariant measures on  $Z$  with entropy not exceeding  $t$  is arcwise connected with respect to the  $\bar{d}$ -metric on the set of all invariant measures [23, Theorem 6].

5.2. *A  $\bar{d}$ -approachable minimal shift space.* We will construct a minimal shift space which is  $\bar{d}$ -approximable by a descending sequence of mixing sofic shifts. The sofic shift  $X_n$  in the sequence will be generated by a finite code  $\mathcal{B}_n \subset \{0, 1\}^+$ . The parameters of the construction are an initial finite non-empty set of words  $\mathcal{B}_1$  and a sequence of positive integers  $(t(n))_{n=1}^\infty$ . We assume that  $t(n) \geq 2$  for every  $n$ . We will impose some more conditions on the  $t(n)$  and  $\mathcal{B}_1$  later.

Assume we have defined the family of words  $\mathcal{B}_n$  for some  $n \geq 1$ . Write  $k(n)$  for the cardinality of  $\mathcal{B}_n$ . Enumerate the elements of  $\mathcal{B}_n$  as  $\beta_1^{(n)}, \dots, \beta_{k(n)}^{(n)}$ , and let  $\tau(n) = \beta_1^{(n)} \dots \beta_{k(n)}^{(n)}$  denote their concatenation. Let  $s(n)$  (respectively,  $\ell(n)$ ) be the length of the *shortest* (respectively, the *longest*) word in  $\mathcal{B}_n$ . Words belonging to  $\mathcal{B}_{n+1}$  are constructed as follows. First we concatenate  $t(n)$  arbitrarily chosen words from  $\mathcal{B}_n$ . Then we add the suffix  $\tau(n)$ :

$$\mathcal{B}_{n+1} = \{b_1 b_2 \dots b_{t(n)} \tau(n) : b_i \in \mathcal{B}_n \text{ for } 1 \leq i \leq t(n)\}.$$

By the construction,  $\ell(n) < \tau(n) < s(n + 1)$  for every  $n \geq 1$  and so  $s(n) \nearrow \infty$  as  $n \rightarrow \infty$ . Moreover, every word from  $\mathcal{B}_n$  is a subword of every word from  $\mathcal{B}_{n+1}$ . Recursively,

$$u \text{ is a subword of } v, \quad \text{for every } u \in \mathcal{B}_n, v \in \mathcal{B}_m, n \leq m. \tag{34}$$

For  $n \geq 1$ , let  $X_n$  be the coded shift generated by the code  $\mathcal{B}_n$ . That is,  $X_n$  consists of all concatenations of words from  $\mathcal{B}_n$  together with their shifts. Since  $\mathcal{B}_n$  is finite, the shift  $X_n$  is transitive and sofic. It follows from (34) that  $X_{n+1} \subseteq X_n$ . Hence,  $X = \bigcap_{n=1}^\infty X_n$  is a non-empty shift space.

**PROPOSITION 5.7.** *The shift  $X$  constructed above is minimal.*

*Proof.* If  $|\mathcal{B}_1| = 1$ , then  $X$  is an orbit of a periodic point. Let us assume that  $|\mathcal{B}_1| > 1$ , so  $|\mathcal{B}_n| > 1$  for every  $n$ . We need to prove that if  $u \in \mathcal{B}(X) = \bigcap_{n=1}^\infty \mathcal{B}(X_n)$  and  $x \in X$ , then  $u$  appears in  $x$ . Fix  $x \in X$  and  $u \in \mathcal{B}(X)$ . Take  $n$  large enough to imply  $|u| < s(n)$ . Since  $\mathcal{B}(X) \subseteq \mathcal{B}(X_n)$ , we see that  $u$  must appear in some  $x' \in X_n$ . As all words in  $\mathcal{B}_n$  are longer than  $u$  and  $x'$  is a shift (possibly trivial) of an infinite concatenation of words from  $\mathcal{B}_n$ , we conclude that  $u$  is a subword of some  $\bar{u} \in \mathcal{B}(X_n)$  which is the concatenation of two words  $v, w$  in  $\mathcal{B}_n$  (one of them might be empty). By the definition of the  $\mathcal{B}_n$ , every concatenation  $vw$  for words  $v$  and  $w$  from  $\mathcal{B}_n$  appears in some word from  $\mathcal{B}_{n+1}$ , therefore  $vw$  and, in

particular,  $u$  is a subword of a word  $w' \in \mathcal{B}_{n+1}$ . Hence, condition (34) ensures that  $u$  is a subword of all words from  $\mathcal{B}_{n+2}$ . But  $x \in X_{n+2}$ , so it is a shifted infinite concatenation of words from  $\mathcal{B}_{n+2}$ . In particular, some word from  $\mathcal{B}_{n+2}$  appears in  $x$ , and so does  $u$ .  $\square$

From now on, we set  $\mathcal{B}_1 = \{0, 11\}$ . For this choice of  $\mathcal{B}_1$  a simple inductive argument shows that for each  $n \geq 1$ , the set of lengths of all words in  $\mathcal{B}_n$  is an interval:

$$\text{for every } m \text{ with } s(n) \leq m \leq \ell(n), \text{ there exists } u \in \mathcal{B}_n \text{ with } |u| = m. \tag{35}$$

PROPOSITION 5.8. *For every  $n \geq 1$  the coded system  $X_n$  is a mixing sofic shift.*

*Proof.* Any coded system generated by a finite sequence of words is sofic. In the corresponding graphs, every word in  $\mathcal{B}_n$  is represented by a cycle, and all these cycles have a common vertex. In particular, the graph presenting  $X_n$  is strongly irreducible and  $X_n$  is transitive. In addition, it follows from (35) that there are two words in  $\mathcal{B}_n$  with coprime lengths, so the graph is aperiodic, thus  $X_n$  is mixing and has the specification property.  $\square$

In the rest of the section, it will be convenient to control the ratio  $s(n)/\ell(n)$ . Because of the identities

$$\begin{aligned} s(n + 1) &= t(n)s(n) + |\tau(n)|, \\ \ell(n + 1) &= t(n)\ell(n) + |\tau(n)|, \end{aligned}$$

we get that the ratio is increasing and

$$\frac{s(n)}{\ell(n)} \geq \frac{1}{2}, \quad n \geq 1.$$

On the other hand, we can ensure that

$$\frac{s(n)}{\ell(n)} < \frac{2}{3}, \quad n \geq 1 \tag{36}$$

by satisfying the equivalent condition (the equivalence follows from the inductive definition of  $s(n)$  and  $\ell(n)$  mentioned above)

$$t(n) > \frac{|\tau(n)|}{2\ell(n) - 3s(n)}, \quad n \geq 1. \tag{37}$$

Since  $|\tau(n)|$ ,  $\ell(n)$  and  $s(n)$  are determined by  $t(i)$ ,  $1 \leq i < n$ , we have enough freedom to construct the sequence  $t(n)$  satisfying condition (37) in an inductive way.

PROPOSITION 5.9. *Let  $\varepsilon > 0$  and  $t(n)$  be such a sequence that satisfies condition (37) and*

$$t(n) > \frac{|\tau(n)| + 3\ell(n)}{s(n)\varepsilon 2^{-n}}, \quad n \geq 1. \tag{38}$$

Then

$$\sum_{n=1}^{\infty} \bar{d}^H(X_n, X_{n+1}) < \varepsilon. \tag{39}$$

*Proof.* Put  $\varepsilon_n = \varepsilon 2^{-n}$ . We will show that  $\bar{d}^H(X_n, X_{n+1}) < \varepsilon_n$  for all  $n \geq 1$ , which directly implies (39). Fix  $n \geq 1$  and  $y \in X_n$ . Our goal is to find  $z \in X_{n+1}$  such that  $\bar{d}(y, z) < \varepsilon_n$ . Since  $\bar{d}$  is shift invariant, without loss of generality we assume that  $y$  is a concatenation of blocks from  $\mathcal{B}_n$ , that is, we have

$$y = b_1 b_2 b_3 \dots \quad \text{where } b_j \in \mathcal{B}_n \text{ for } j = 1, 2, \dots$$

We will construct  $z$  inductively. First, we note that the word

$$w = b_1 b_2 b_3 \dots b_{t(n)} \tau(n)$$

belongs to  $\mathcal{B}_{n+1}$ . Let  $j \geq t(n)$  be the index with

$$|b_1 b_2 \dots b_j| \leq |w| < |b_1 b_2 \dots b_{j+1}|,$$

and let  $a$  be the suffix of  $b_{j+1}$  such that  $|b_1 b_2 \dots b_j b_{j+1}| = |w| + |a|$ . We observe that there exist words  $b'_1, b'_2, b'_3 \in \mathcal{B}_n \cup \{\lambda\}$  such that  $|b'_1 b'_2 b'_3| = |ab_{j+2} b_{j+3}|$ . Indeed, if  $2s(n) \leq |ab_{j+2} b_{j+3}| \leq 2\ell(n)$  then it follows from (35) that we can find  $b'_1, b'_2 \in \mathcal{B}_n$  and  $b'_3 = \lambda$  with the required total length. Likewise, if  $3s(n) \leq |ab_{j+2} b_{j+3}| \leq 3\ell(n)$  then we can apply the same argument with  $b'_1, b'_2, b'_3 \in \mathcal{B}_n$ . Since  $|a| < \ell(n)$  and  $3s(n) < 2\ell(n)$  (we assumed (37), which is equivalent to (36)), these two cases cover all possibilities.

For  $i \geq 3$ , define  $b'_i = b_{j+i}$  and let  $y' = b'_1 b'_2 b'_3 \dots$ . Then  $y$  and  $wy'$  differ only on the positions where  $\tau(n)b'_1 b'_2 b'_3$  appears in  $wy'$ , that is, at most on positions between  $|w| - |\tau(n)|$  and  $|w| + 3\ell(n)$ . Let us point out that  $y'$  is again a concatenation of blocks from  $\mathcal{B}_n$ , even in the case when  $b'_3$  is the empty word. Hence, we can apply the same reasoning to  $y', y'', y'''$  and so on, to obtain the word  $z = ww'w'' \dots$ . Here, we adopt the convention that  $y'', w'$  are constructed from  $y'$  in the same way as  $y', w$  were constructed from  $y$ , and accordingly for further steps in the construction.

Since  $w, w', w'' \in \mathcal{B}_{n+1}$ , we have  $z \in X_{n+1}$ . Note that for each  $i \geq 0$  we have  $s(n)t(n) + \tau(n) \leq |w^{(i)}| \leq \ell(n)t(n) + \tau(n)$ , and consequently

$$\bar{d}(y, z) \leq \limsup_{i \rightarrow \infty} \frac{i\tau(n) + 3i\ell(n)}{|w| + |w'| + \dots + |w^{(i-1)}|} \leq \frac{\tau(n) + 3\ell(n)}{\tau(n) + s(n)t(n)} < \varepsilon_n.$$

In the last inequality we use condition (38), which is stronger.

Hence,  $z$  has all of the required properties. □

PROPOSITION 5.10. *Let the sequence  $t(n), n \geq 1$ , satisfy (36) and*

$$t(n) \geq \frac{\ell(n)}{\ell(n) - s(n)} \quad \text{for all } n \geq 1; \tag{40}$$

$$t(n) \geq \frac{2s(n) + 2\ell(n) + 3|\tau(n)|}{\ell(n)} \quad \text{for all } n \geq 1. \tag{41}$$

*Then  $X$  is mixing.*

*Proof.* We need to show that for all  $u, v \in \mathcal{B}(X)$  there exists  $M$  such that for each  $m \geq M$  there exists a word  $w$  with  $|w| = m$  such that  $uwv \in \mathcal{B}(X)$ . Note that we can freely replace  $u, v$  with any other words  $u', v' \in \mathcal{B}(X)$  which contain  $u, v$  as subwords. Hence (repeating

an argument from the proof of Proposition 5.7), we may assume that  $u, v \in \mathcal{B}_n$  for some  $n \geq 1$ . Proceeding by induction on  $m$ , we will prove the following statement.

*Claim.* For all  $m \geq 0$ , for all  $n \geq 1$  such that  $2s(n) \leq m$ , for all  $u, v \in \mathcal{B}(X)$ , there exists  $w$  with  $|w| = m$  and  $uwv \in \mathcal{B}(X)$ .

We may assume that the claim above has been proved for all  $m' < m$ . We consider three cases depending on the magnitude of  $m$ .

*Case 1.* Suppose first that  $2s(n) \leq m \leq (t(n) - 2)\ell(n)$ . Note that for each  $j \geq 2$  it follows from (36) that

$$(j + 1)s(n) \leq \frac{3}{2}j \cdot \frac{2}{3}\ell(n) = j\ell(n).$$

As a consequence, the intervals  $[js(n), j\ell(n)]$  for  $2 \leq j \leq t(n) - 2$  fully cover the interval  $[2s(n), (t(n) - 2)\ell(n)]$ . Hence, we can find  $j$  with  $2 \leq j \leq t(n) - 2$  such that  $js(n) \leq m \leq j\ell(n)$ . It follows from (35) that there exists a word  $w$  with  $|w| = m$  of the form  $w = b_1b_2 \cdots b_j$  with  $b_1, b_2, \dots, b_j \in \mathcal{B}_n$ . Thus,  $uwv$  is a prefix of  $\mathcal{B}_{n+1}$  (specifically, of any word of the form  $ub_1b_2 \cdots b_jvb'_1b'_2 \cdots b'_i\tau(n)$ , where  $i = t(n) - j - 2$  and  $b'_1, b'_2, \dots, b'_i \in \mathcal{B}_n$ ). It follows that  $uwv \in \mathcal{B}(X)$ .

*Case 2.* Suppose next that  $(t(n) - 2)\ell(n) < m \leq (2t(n) - 2)\ell(n) + |\tau(n)|$ . Then by (41) we have  $m \geq 2s(n) + |\tau(n)|$ . Arguing similarly as in the first case, we can find a word  $w$  with  $|w| = m$  of the form

$$w = b_1b_2 \cdots b_j\tau(n)b'_1b'_2 \cdots b'_k,$$

where  $1 \leq j, k \leq t(n) - 1$  and  $b_1, b_2, \dots, b_j, b'_1, b'_2, \dots, b'_k \in \mathcal{B}_n$ . Thus,  $uwv$  is a subword of the concatenation  $c_1c_2$  of two words  $c_1, c_2 \in \mathcal{B}_{n+1}$ . Hence,  $uwv$  is a subword of a word in  $\mathcal{B}_{n+2}$  and thus  $uwv \in \mathcal{B}(X)$ .

*Case 3.* Suppose finally that  $m > (t(n) - 2)\ell(n) + \tau(n)$ . Put

$$\begin{aligned} u' &= b_1b_2 \cdots b_{t(n)-1}u\tau(n) \in \mathcal{B}_{n+1}, \\ v' &= b'_2b'_3 \cdots b'_{t(n)}\tau(n) \in \mathcal{B}_{n+1}, \end{aligned}$$

where  $b_1, b_2, \dots, b_{t(n)-1}, b'_2, b'_3, b'_{t(n)} \in \mathcal{B}_n$  are arbitrary. Put also  $n' = n + 1$  and  $m' = m - |\tau(n)|$ . By (40) we have

$$m' \geq 2(t(n)s(n) + |\tau(n)|) = 2s(n').$$

Hence, by the inductive assumption, there exists a word  $w'$  with  $|w'| = m$  such that  $u'w'v' \in \mathcal{B}(X)$ . It remains to observe that  $uwv$  is a subword of  $u'w'v'$ , where  $w = \tau(n)w$  has length  $|w| = |\tau(n)| + |w'| = m$ . □

**THEOREM 5.11.** *There exists a sequence of positive integers  $(t(n))_{n=1}^\infty$  such that  $X$  is minimal, mixing, and has positive entropy and  $\bar{d}$ -shadowing property. In particular, the shift space  $X$  has an entropy-dense and uncountable set of ergodic measures.*

*Proof.* Since  $X_1$  is a mixing sofic shift it has topological entropy  $h > 0$ . By the uniform continuity of the entropy function with respect to  $\bar{d}_M$ -distance, there is  $\varepsilon > 0$  such that for every  $\mu, \mu' \in \mathcal{M}_\sigma(\{0, 1\}^\infty)$  with  $\bar{d}_M(\mu, \mu') < \varepsilon$  we have  $|h(\mu) - h(\mu')| < h/3$ . Fix this  $\varepsilon$ .

All the conditions (37), (38) with respect to  $\varepsilon$  specified above, (40) and (41) can be satisfied simultaneously by one sequence  $t(n)$ ,  $n \geq 1$ . Indeed, it is enough to construct the sequence inductively and take  $t(n)$  large enough with respect to right-hand sides of the conditions which all depend only on the previously taken  $t(i)$ ,  $i < n$ . Such a sequence then satisfies all assumptions of Propositions 5.9 and 5.10. Hence,  $X$  is mixing and  $\bar{d}$ -approachable from above by mixing sofic shifts. In particular, the shift has an entropy-dense set of ergodic measure and is  $\bar{d}$ -shadowing.

Now it suffices to find two invariant measures on  $X$  with different entropies. By the variational principle, there is an ergodic invariant measure  $\nu_1$  on  $X_1$  with entropy  $h$ . Since  $X_1$  contains a periodic point, there is also an ergodic invariant measure  $\nu_2$  on  $X_1$  of entropy zero. The inequality  $\bar{d}_M^H(\mathcal{M}_\sigma e(X_n), \mathcal{M}_\sigma e(X)) < \varepsilon/2$  ensures that there is an invariant measure  $\nu'_1$  on  $X$  that is  $\varepsilon$ -close to  $\nu_1$  in  $\bar{d}_M$ -distance and so  $h/3$ -close to  $\nu_1$  in entropy. By the same argument, there is an invariant measure  $\nu'_2$  on  $X$  that is  $h/3$ -close to  $\nu_2$  in entropy. Since the difference between  $h(\nu_1)$  and  $h(\nu_2)$  equals  $h$ , the measures  $\nu'_1$  and  $\nu'_2$  have different entropy and, in particular, are distinct.  $\square$

*Acknowledgments.* The research cooperation between Michal Kupsa and Dominik Kwietniak was partially funded by the program Excellence Initiative—Research University under the Strategic Programme Excellence Initiative at Jagiellonian University in Kraków. The program supported the research stay of Michal Kupsa in Kraków in May–June 2023 during which the present paper was (almost) finished. The research of Melih Emin Can is part of the project no. 2021/43/P/ST1/02885 cofunded by the National Science Centre and the European Union’s Horizon 2020 research and innovation program under Marie Skłodowska-Curie grant agreement no. 945339. When this paper was being finished, D. Kwietniak was partially supported by the Flagship Project ‘Central European Mathematical Research Lab’ under the Strategic Programme Excellence Initiative at Jagiellonian University. J. Konieczny is supported by UKRI Fellowship EP/X033813/1, and during a significant part of the work on this project he was working within the framework of LABEX MILYON (ANR-10-LABX-0070) of Université de Lyon, within the program ‘Investissements d’Avenir’ (ANR-11-IDEX-0007) operated by the French National Research Agency (ANR). He also acknowledges support from the Foundation for Polish Science (FNP). The authors would like to express their gratitude to Tim Austin for sharing a very early version of his paper [1] and an enlightening discussion. We would like to thank Piotr Oprocha for allowing us to use his ideas that led to the construction of the minimal examples of  $\bar{d}$ -approachable shift spaces in §5.2. Thanks are also due to Alexandre Trilles, who helped us to find and fix numerous minor faults in §§5.1 and 5.2. Melih Emin Can would like to thank Damla Buldağ Can for her constant love and support. Finally, we are grateful to the referee for a careful reading of our paper and for helpful corrections and suggestions.

## REFERENCES

- [1] T. Austin. Private communication.
- [2] V. Bergelson, J. Kułaga-Przymus, M. Lemańczyk and F. K. Richter. Rationally almost periodic sequences, polynomial multiple recurrence and symbolic dynamics. *Ergod. Th. & Dynam. Sys.* **39**(9) (2019), 2332–2383.

- [3] M. L. Blank. Metric properties of  $\varepsilon$ -trajectories of dynamical systems with stochastic behaviour. *Ergod. Th. & Dynam. Sys.* **8**(3) (1988), 365–378.
- [4] V. Climenhaga and D. Thompson. Intrinsic ergodicity beyond specification:  $\beta$ -shifts,  $S$ -gap shifts, and their factors. *Israel J. Math.* **192**(2) (2012), 785–817.
- [5] H. Comman. Strengthened large deviations for rational maps and full shifts, with unified proof. *Nonlinearity* **22**(6) (2009), 1413–1429.
- [6] H. Comman. Criteria for the density of the graph of the entropy map restricted to ergodic states. *Ergod. Th. & Dynam. Sys.* **37**(3) (2017), 758–785.
- [7] T. Downarowicz. Survey on odometers and Toeplitz flows. *Algebraic and Topological Dynamics (Contemporary Mathematics, 385)*. Ed. S. Kolyada, Y. Manin and T. Ward. American Mathematical Society, Providence, RI, 2005, pp. 7–37.
- [8] T. Downarowicz and J. Serafin. Possible entropy functions. *Israel J. Math.* **135** (2003), 221–250.
- [9] T. Downarowicz and M. Więcek. Decomposition of a symbolic element over a countable amenable group into blocks approximating ergodic measures. *Groups Geom. Dyn.* **16**(3) (2022), 909–941.
- [10] A. Dymek, S. Kasjan, J. Kułaga-Przymus and M. Lemańczyk.  $\mathfrak{B}$ -free sets and dynamics. *Trans. Amer. Math. Soc.* **370**(8) (2018), 5425–5489.
- [11] A. Dymek, J. Kułaga-Przymus and D. Sell. Invariant measures for  $\mathfrak{B}$ -free systems revisited. *Ergod. Th. & Dynam. Sys.* doi:10.1017/etds.2024.7. Published online 8 March 2024.
- [12] A. Eizenberg, Y. Kifer and B. Weiss. Large deviations for  $\mathbb{Z}^d$ -actions. *Comm. Math. Phys.* **164**(3) (1994), 433–454.
- [13] H. Föllmer and S. Orey. Large deviations for the empirical field of a Gibbs measure. *Ann. Probab.* **16**(3) (1988), 961–977.
- [14] N. A. Friedman and D. S. Ornstein. On isomorphism of weak Bernoulli transformations. *Adv. Math.* **5** (1970), 365–394.
- [15] H. Furstenberg. *Recurrence in Ergodic Theory and Combinatorial Number Theory*, Vol. 14. Princeton University Press, Princeton, NJ, 2014.
- [16] K. Gelfert and D. Kwietniak. On density of ergodic measures and generic points. *Ergod. Th. & Dynam. Sys.* **38**(5) (2018), 1745–1767.
- [17] E. Glasner. *Ergodic Theory via Joinings (Mathematical Surveys and Monographs, 101)*. American Mathematical Society, Providence, RI, 2003.
- [18] A. Illanes and S. B. Nadler Jr. *Hyperspaces. Fundamentals and Recent Advances (Monographs and Textbooks in Pure and Applied Mathematics, 216)*. Marcel Dekker, New York, 1999.
- [19] S. Kasjan, G. Keller and M. Lemańczyk. Dynamics of  $\mathfrak{B}$ -free sets: a view through the window. *Int. Math. Res. Not. IMRN* **2019**(9) (2019), 2690–2734.
- [20] G. Keller. Tautness for sets of multiples and applications to  $\mathfrak{B}$ -free dynamics. *Studia Math.* **247**(2) (2019), 205–216.
- [21] G. Keller. Generalized heredity in  $\mathfrak{B}$ -free systems. *Stoch. Dyn.* **21**(3) (2021), 2140008.
- [22] J. Konieczny, M. Kupsa and D. Kwietniak. Arcwise connectedness of the set of ergodic measures of hereditary shifts. *Proc. Amer. Math. Soc.* **146**(8) (2018), 3425–3438.
- [23] J. Konieczny, M. Kupsa and D. Kwietniak. On  $\bar{d}$ -approachability, entropy density and  $\mathfrak{B}$ -free shifts. *Ergod. Th. & Dynam. Sys.* **43**(3) (2023), 943–970.
- [24] J. Kułaga-Przymus, M. Lemańczyk and B. Weiss. On invariant measures for  $\mathfrak{B}$ -free systems. *Proc. Lond. Math. Soc.* (3) **110** (2015), 1435–1474.
- [25] J. Kułaga-Przymus, M. Lemańczyk and B. Weiss. *Hereditary subshifts whose simplex of invariant measures is Poulsen. Ergodic Theory, Dynamical Systems, and the Continuing Influence of John C. Oxtoby (Contemporary Mathematics, 678)*. American Mathematical Society, Providence, RI, 2016, pp. 245–253.
- [26] J. Kułaga-Przymus and M. D. Lemańczyk. Hereditary subshifts whose measure of maximal entropy does not have the Gibbs property. *Colloq. Math.* **166**(1) (2021), 107–127.
- [27] M. Kulczycki, D. Kwietniak and P. Oprocha. On almost specification and average shadowing properties. *Fund. Math.* **224**(3) (2014), 241–278.
- [28] P. Kůrka. *Topological and Symbolic Dynamics (Cours Spécialisés [Specialized Courses], 11)*. Société Mathématique de France, Paris, 2003.
- [29] D. Kwietniak. Topological entropy and distributional chaos in hereditary shifts with applications to spacing shifts and beta shifts. *Discrete Contin. Dyn. Syst.* **33**(6) (2013), 2451–2467.
- [30] D. Kwietniak, M. Łącka and P. Oprocha. A panorama of specification-like properties and their consequences. *Dynamics and Numbers (Contemporary Mathematics, 669)*. American Mathematical Society, Providence, RI, 2016, pp. 155–186.
- [31] D. Kwietniak, P. Oprocha and M. Rams. On entropy of dynamical systems with almost specification. *Israel J. Math.* **213**(1) (2016), 475–503.

- [32] D. Lind and B. Marcus. *An Introduction to Symbolic Dynamics and Coding*. Cambridge University Press, Cambridge, 1995.
- [33] P. Oprocha. Families, filters and chaos. *Bull. Lond. Math. Soc.* **42**(4) (2010), 713–725.
- [34] S. Orey. Large deviations in ergodic theory. *Seminar on Stochastic Processes, 1984 (Evanston, IL, 1984) (Progress in Probability and Statistics, 9)*. Ed. E. Çinlar, K. L. Chung and R. K. Gettoor. Birkhäuser, Boston, 1986, pp. 195–249.
- [35] D. S. Ornstein. *Ergodic Theory, Randomness, and Dynamical Systems (Yale Mathematical Monographs, 5)*. Yale University Press, New Haven, CT, 1974.
- [36] J. C. Oxtoby. Ergodic sets. *Bull. Amer. Math. Soc. (N.S.)* **58**(2) (1952), 116–136.
- [37] R. Pavlov. On intrinsic ergodicity and weakenings of the specification property. *Adv. Math.* **295** (2016), 250–270.
- [38] C.-E. Pfister and W. G. Sullivan. Large deviations estimates for dynamical systems without the specification property. Applications to the  $\beta$ -shifts. *Nonlinearity* **18**(1) (2005), 237–261.
- [39] C.-E. Pfister and W. G. Sullivan. Weak Gibbs measures and large deviations. *Nonlinearity* **31**(1) (2018), 49–53.
- [40] D. J. Rudolph and G. Schwarz. The limits in  $\bar{d}$  of multi-step Markov chains. *Israel J. Math.* **28**(1–2) (1977), 103–109.
- [41] P. Shields. *The Ergodic Theory of Discrete Sample Paths (Graduate Studies in Mathematics, 13)*. American Mathematical Society, Providence, RI, 1991.
- [42] D. Thompson. A ‘horseshoe’ theorem in symbolic dynamics via single sequence techniques. Unpublished manuscript, 2017.
- [43] S. Williams. Toeplitz minimal flows which are not uniquely ergodic. *Z. Wahrscheinlichkeitstheorie Verw. Geb.* **67**(1) (1984), 95–107.