

On \bar{d} -approachability, entropy density and \mathcal{B} -free shifts

JAKUB KONIECZNY[†], MICHAL KUPSA[‡] and DOMINIK KWIETNIAK[§]

[†] *Camille Jordan Institute, Claude Bernard University Lyon 1,
43 Boulevard du 11 novembre 1918, 69622 Villeurbanne Cedex, France
(e-mail: jakub.konieczny@gmail.com)*

[‡] *The Czech Academy of Sciences, Institute of Information Theory and Automation,
CZ-18208 Prague 8, Czech Republic
(e-mail: kupsa@utia.cas.cz)*

[§] *Faculty of Mathematics and Computer Science, Jagiellonian University in Krakow,
ul. Łojasiewicza 6, 30-348 Kraków, Poland
(e-mail: dominik.kwietniak@uj.edu.pl)*

(Received 29 July 2021 and accepted in revised form 2 December 2021)

Abstract. We study approximation schemes for shift spaces over a finite alphabet using (pseudo)metrics connected to Ornstein's \bar{d} metric. This leads to a class of shift spaces we call \bar{d} -approachable. A shift space is \bar{d} -approachable when its canonical sequence of Markov approximations converges to it also in the \bar{d} sense. We give a topological characterization of chain-mixing \bar{d} -approachable shift spaces. As an application we provide a new criterion for entropy density of ergodic measures. Entropy density of a shift space means that every invariant measure μ of such a shift space is the weak* limit of a sequence μ_n of ergodic measures with the corresponding sequence of entropies $h(\mu_n)$ converging to $h(\mu)$. We prove ergodic measures are entropy-dense for every shift space that can be approximated in the \bar{d} pseudometric by a sequence of transitive sofic shifts. This criterion can be applied to many examples that were beyond the reach of previously known techniques including hereditary \mathcal{B} -free shifts and some minimal or proximal systems. The class of symbolic dynamical systems covered by our results includes also shift spaces where entropy density was established previously using the (almost) 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

We study approximation schemes for shift spaces over a finite alphabet \mathcal{A} (subshifts of the full shift \mathcal{A}^∞ over \mathcal{A}). For every shift space X there exists a canonically defined sequence of shifts of finite type (topological Markov approximations) converging to X in a natural topology on the space of all subshifts of \mathcal{A}^∞ . This fact, however, is of little practical use, because dynamical properties usually do not extend from a sequence of shift spaces to its limit. Here, we consider another, stronger topology on the powerset of \mathcal{A}^∞ which is induced by the \bar{d} pseudometric on \mathcal{A}^∞ or one of its relatives. The pseudometric \bar{d} is given for $x = (x_j)_{j=0}^\infty, y = (y_j)_{j=0}^\infty \in \mathcal{A}^\infty$ by

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

It is an analogue of Ornstein's metric $\bar{d}_{\mathcal{M}}$ on \mathcal{A} -valued stationary stochastic processes. Since \bar{d} is bounded, it induces a Hausdorff pseudometric \bar{d}^H on the space of all non-empty subsets of \mathcal{A}^∞ . We examine the existence of an approximating sequence in the \bar{d}^H sense and study its consequences. In particular, we study shift spaces which we call \bar{d} -approachable, which are \bar{d}^H -limits of their own topological Markov approximations. We provide a topological characterization of chain-mixing \bar{d} -approachable shift spaces using the \bar{d} -shadowing property. This can be considered as an analogue for Friedman and Ornstein's characterization of Bernoulli processes as totally ergodic $\bar{d}_{\mathcal{M}}$ -limits of their own Markov approximations [23]. We prove that many specification properties imply chain mixing and \bar{d} -approachability. It follows that all shift spaces with the almost specification property (this class includes all β -shifts and all mixing sofic shifts, in particular all mixing shifts of finite type) are \bar{d} -approachable. In a forthcoming paper [30], we also construct minimal and proximal examples of \bar{d} -approachable shift spaces, thus proving \bar{d} -approachability is a more general phenomenon than specification. Regarding consequences of \bar{d}^H approximation, we show that it implies convergence of simplices of invariant measures in the Hausdorff metric $\bar{d}_{\mathcal{M}}^H$ induced by Ornstein's $\bar{d}_{\mathcal{M}}$ metric on the space $\mathcal{M}_\sigma(\mathcal{A}^\infty)$ of all shift invariant measures on \mathcal{A}^∞ . To be more precise, write $\mathcal{M}_\sigma(X)$ for the space of all Borel invariant probability measures on a shift space X and $\mathcal{M}_\sigma^e(X)$ for the set of ergodic measures in $\mathcal{M}_\sigma(X)$. We prove that if a sequence of shift spaces $(X_n)_{n=1}^\infty$ converges to X in \bar{d}^H , then $(\mathcal{M}_\sigma(X_n))_{n=1}^\infty$ and $(\mathcal{M}_\sigma^e(X_n))_{n=1}^\infty$ converge in $\bar{d}_{\mathcal{M}}^H$ to $\mathcal{M}_\sigma(X)$, respectively to $\mathcal{M}_\sigma^e(X)$. This allows us to show that certain geometric features of the simplices of invariant measures of shift spaces in the approximating sequence are inherited by the simplex of the limit. The geometry of the space $\mathcal{M}_\sigma(X)$ is, in turn, an important feature of a shift space $X \subseteq \mathcal{A}^\infty$, and besides being interesting *per se*, it has profound connections with dynamical properties of X . It turns out that for many important shift spaces and, more generally, dynamical systems and continuous time flows on compact metric spaces, $\mathcal{M}_\sigma^e(X)$ is a non-trivial dense subset of $\mathcal{M}_\sigma(X)$ endowed with the weak* topology; see [24, 41, 45, 54, 57]. A stronger property is entropy density of ergodic measures introduced by Orey in 1986 [43] and Föllmer and Orey in 1988 [22]. A shift space X has entropy density of ergodic measures if for every $\nu \in \mathcal{M}_\sigma(X)$ there exists a sequence $\mu_n \in \mathcal{M}_\sigma^e(X)$ such that $\mu_n \rightarrow \nu$ in the weak* topology and, in addition, the measure-theoretic entropies also converge, that is, $h(\mu_n) \rightarrow h(\nu)$ as $n \rightarrow \infty$.

There are examples of shift spaces with dense, but not entropy-dense sets of ergodic measures (see [24]). In addition to being of interest in their own right, density of ergodic measures and entropy density are strongly related to various results on large deviations and multifractal analysis [12, 20, 48, 50]. We refer the reader to Comman's article [13] and references therein for more information about that connection and more large-deviations results.

We show that entropy density of ergodic measures is a \bar{d}^H -closed property. This allows for a new method of proof of entropy density. So far, specification-like conditions were invoked to prove entropy density (see [20, 48]), and our results extends that to \bar{d}^H limits of these shift spaces. In particular, mixing shifts of finite type are entropy-dense, hence our result shows that all \bar{d} -approachable chain-mixing shift spaces are also entropy-dense. This leads to a new proof of entropy density of systems with almost specification, which is now a direct consequence of entropy density of mixing shifts of finite type and \bar{d} -approachability. Since we know there are examples of minimal or proximal \bar{d} -approachable shifts, we see that our technique yields entropy density for examples which were beyond the reach of methods based on specification properties. Furthermore, we prove that there are cases where \bar{d} -approachability does not necessarily hold, but it is still possible to approximate in \bar{d}^H with a sequence of entropy-dense shift spaces. We apply this technique to hereditary closures of \mathcal{B} -free shifts (a class including many interesting \mathcal{B} -free shifts; see Corollary 35 below). These shift spaces are not weakly mixing, hence they do not have the \bar{d} -shadowing property (we show examples that are not \bar{d} -approachable; see Example 34). Nevertheless, using the Davenport–Erdős theorem, they are easily seen to be approximated by naturally defined sequences of transitive sofic shifts, and this implies entropy density. Note that \mathcal{B} -free shifts do not have even the weakest version of the specification property (see [38] for a list of possibilities). Recall that an integer is \mathcal{B} -free if it has no factor in a given set $\mathcal{B} \subseteq \mathbb{N}$. For example, the set of \mathcal{B}_{sq} -free integers where $\mathcal{B}_{\text{sq}} = \{p^2 : p \text{ prime}\}$ is just the set of *square-free integers*. Sets of \mathcal{B} -free integers were studied by Chowla, Davenport and Erdős, to name but a few. Recently, Sarnak [52] initiated the study of dynamics of a shift space X_{sq} associated to square-free integers. El Abdalauoi, Lemańczyk, and de la Rue [21] and Dymek *et al* [19] extended Sarnak's approach to \mathcal{B} -free integers for an arbitrary \mathcal{B} . These systems were also investigated by Avdeeva and her co-authors [1, 2], Cellarosi and Sinai [7], Kułaga-Przymus and her co-authors [31–33], Peckner [47], and Keller and his co-authors [27–29]. Higher-dimensional analogues of \mathcal{B} -free shifts have also attracted attention; see [3, 8, 18]).

Furthermore, it is known that ergodic measures are dense for hereditary \mathcal{B} -free shifts by [32], but a proof of this fact is a consequence of a non-trivial reasoning presented in [31] (more precisely, in [31] the analysis of invariant measures was performed under some additional technical assumptions on the set \mathcal{B} and it was later extended to the general case in [19]). We extend the main result of [32] in two directions: we add entropy to the picture and broaden the class of shift spaces for which this result holds (see Remark 37). Our proof is independent of [19, 31, 32]. As a matter of fact our methods lead to new, often shorter, proofs of many results from [19, 31, 32]. Furthermore, our methods can be adapted to shift spaces with an action of \mathbb{Z}^d for $d \geq 2$ or a countable amenable residually finite group. This will be the subject of a future paper.

The idea of using $\bar{d}_{\mathcal{M}}$ approximation was independently considered by Thompson [56], who used it in the setting of [10].

This paper is organized as follows. In the next section we recall some notation and basic definitions. In §3 we discuss Hausdorff (pseudo)metrics \bar{d}^H and $\bar{d}_{\mathcal{M}}^H$ induced by \bar{d} and $\bar{d}_{\mathcal{M}}$. Section 4 contains the definition and characterization of \bar{d} -approachability. In §5 we study consequences of \bar{d}^H convergence of shift spaces and $\bar{d}_{\mathcal{M}}^H$ convergence of simplices of invariant measures. The last three sections contain applications and examples illustrating our approach. In §6 we show that our results contain previous results on entropy density of shift spaces with a variant of the specification property. Section 7 discusses approximation schemes and entropy density of \mathcal{B} -free shifts and their hereditary closures. In §8 we show that topologically mixing S -gap shifts are another example of a shift space allowing for a natural \bar{d}^H -approximation by a sequence of mixing shifts of finite type.

2. Definitions

2.1. *General.* We let \mathbb{N} denote the set of positive integers and $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$.

2.2. *Shift spaces and languages.* Let \mathcal{A} be a finite set, henceforth referred to as the *alphabet*. The *full shift* \mathcal{A}^∞ is the set of all \mathcal{A} -valued infinite sequences indexed by non-negative integers. We endow \mathcal{A}^∞ with the product topology coming from the discrete topology on \mathcal{A} . A metric on \mathcal{A}^∞ compatible with the product topology is

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

We consider \mathcal{A}^∞ a dynamical system under the action of the *shift* transformation $\sigma: \mathcal{A}^\infty \rightarrow \mathcal{A}^\infty$, where $\sigma(x)_j = x_{j+1}$ for every $x = (x_i)_{i=0}^\infty \in \mathcal{A}^\infty$ and $j \geq 0$. A *shift space* over \mathcal{A} is a non-empty, closed and shift invariant subset of \mathcal{A}^∞ . A *block* or a *word* over \mathcal{A} is a finite sequence of symbols from \mathcal{A} . The *length* of a word w over \mathcal{A} is denoted by $|w|$. The *concatenation* of words u and v is denoted simply as uv . We agree that the *empty word* has length 0. Given $x \in X$ and $i, j \in \mathbb{N}_0$ with $i < j$, we set $x_{[i,j]}$ to be the word $w = w_1, \dots, w_n \in \mathcal{A}^n$ such that $n = j - i$ and $x_{i+k-1} = w_k$ for all $1 \leq k \leq n$. We say that a word $\omega \in \mathcal{A}^n$ *appears* in a shift space $X \subseteq \mathcal{A}^\infty$ if there exist some $x \in X$ and $\ell \in \mathbb{N}_0$ such that $\omega = x_{[\ell, \ell+n]}$. For $n \in \mathbb{N}$, we write $\mathcal{B}_n(X) \subseteq \mathcal{A}^n$ for the set of blocks of length n appearing in X . The *language* of a shift space $X \subseteq \mathcal{A}^\infty$ is the set $\mathcal{B}(X)$ of finite words over \mathcal{A} that appear in X . A shift space X is *transitive* if for every word $u, w \in \mathcal{B}(X)$ there is a word v such that $uvw \in \mathcal{B}(X)$. We say that a shift space X is *topologically mixing* if for every word $u, w \in \mathcal{B}(X)$ there is $N \in \mathbb{N}_0$ such that for every $n \geq N$ there is a word v satisfying $uvw \in \mathcal{B}(X)$ and $|v| = n$.

2.3. *Shifts of finite type.* Recall that every family \mathcal{F} of finite words over \mathcal{A} determines the collection $X_{\mathcal{F}}$ consisting of sequences $x = (x_i)_{i=0}^\infty \in \mathcal{A}^\infty$ such that no word from \mathcal{F} appears in x is either empty or a shift space. Conversely, for every shift space X there is a family \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 $X = X_{\mathcal{F}}$.

2.4. *Ergodic properties of shift spaces.* The set of all Borel probability measures on a shift space X is denoted by $\mathcal{M}(X)$. We equip $\mathcal{M}(X)$ with the weak* topology, which is known to be metrizable and compact. The set of shift invariant probability measures supported on a shift space $X \subseteq \mathcal{A}^\infty$ is denoted by $\mathcal{M}_\sigma(X)$, and $\mathcal{M}_\sigma^e(X)$ stands for the set of ergodic σ -invariant measures on X .

A point $x \in \mathcal{A}^\infty$ is *generic* for an ergodic measure $\mu \in \mathcal{M}_\sigma^e(\mathcal{A}^\infty)$, if for every continuous function $f: \mathcal{A}^\infty \rightarrow \mathbb{R}$ the sequence

$$\frac{1}{N} \sum_{j=0}^{N-1} f(\sigma^j(x))$$

converges as $N \rightarrow \infty$ to $\int_{\mathcal{A}^\infty} f \, d\mu$. Every ergodic measure has a generic point. Given $\mu \in \mathcal{M}_\sigma(\mathcal{A}^\infty)$, we let $h(\mu)$ denote the Kolmogorov–Sinai entropy of μ .

The ergodic measures are *entropy-dense* if for every measure $\mu \in \mathcal{M}_\sigma(X)$, every neighbourhood U of μ in $\mathcal{M}_\sigma(X)$ and every $\varepsilon > 0$ there is an ergodic measure $\nu \in U$ with $|h(\nu) - h(\mu)| < \varepsilon$.

Let us point out that property of having entropy-dense ergodic measures is preserved by conjugacy. In other words, two conjugated dynamical systems either both have entropy-dense ergodic measures or both do not have this property.

2.5. *Hereditary shifts.* Suppose that \mathcal{A} is additionally equipped with a (total) order \leq . In particular, if $\mathcal{A} = \{0, 1\}$ then \leq is the usual order. We equip \mathcal{A}^∞ with a coordinatewise partial order also denoted by \leq . This means that for $x, y \in \mathcal{A}^\infty$ we have $x \leq y$ if and only if $x_i \leq y_i$ for all $i \in \mathbb{N}_0$. A shift space $X \subseteq \mathcal{A}^\infty$ is *hereditary* if for every $x \in X$ and $y \in \mathcal{A}^\infty$ with $y \leq x$ we have $y \in X$. The hereditary closure \tilde{X} of X is the smallest hereditary shift containing X and consists of those $y \in \mathcal{A}^\infty$ for which there exists $x \in X$ with $y \leq x$. For more details, see [37].

2.6. *Sofic shifts.* An (oriented) \mathcal{A} -labelled (multi)graph $G = (V, E, \tau)$ consists of a vertex set V , an edge set E and a label map $\tau: E \rightarrow \mathcal{A}$. Each edge $e \in E$ has two endpoints $i(e), t(e) \in V$, sometimes respectively called the initial vertex and the terminal vertex of e . A path of length ℓ (finite or infinite) in G is a sequence of ℓ edges e_1, e_2, \dots such that for each $i < \ell$, the terminal vertex of e_i is the same as the initial vertex of e_{i+1} ; equivalently, a sequence of edges e_1, e_2, \dots is a path if there exists a sequence of vertices $v_1, v_2, \dots \in V$ such that $i(e_i) = v_i$ and $t(e_i) = v_{i+1}$ for every $i < \ell$.

The shift $X_G \subseteq \mathcal{A}^\infty$ corresponding to G consists of all sequences $x \in \mathcal{A}^\infty$ such that there exists an infinite path e_1, e_2, \dots on G such that $x_i = \tau(e_{i+1})$ for each $i \in \mathbb{N}_0$. That is, the shift space X_G is obtained by reading off labels of all infinite paths on G . A shift space is a *sofic shift* if there exists a labelled graph $G = (V, E, \tau)$ with V finite such that $X = X_G$. Then we also say that X is presented by 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 by a path); see [42, Proposition 3.3.11]. A sofic shift is topologically mixing if and only if it can be presented by a (strongly) connected graph such that there are two closed paths on G of coprime lengths.

2.7. *Markov approximations.* We recall the topological approximation scheme of shift spaces by canonically defined sequences of shifts of finite type (see [9], [17, p. 111], or [36] for more details).

Given a shift space $X \subseteq \mathcal{A}^\infty$ and a family \mathcal{F} of finite words over \mathcal{A} such that $X = X_{\mathcal{F}}$, one sets $\mathcal{F}[n]$ to be the set of all words w in \mathcal{F} with length $|w| \leq n + 1$. Clearly, the shift space $X_n^M = X_{\mathcal{F}[n]}$ is a shift of finite type whose language has the same words of length at most $n + 1$ as X . The shift space X_n^M is called the *n*th (topological) Markov approximation of X or *finite type approximation of order n* to X .

An alternative description of X_n^M uses Rauzy graphs. The *n*th Rauzy graph of X is a labelled graph $G_n = (V_n, E_n, \tau_n)$, where $V_n = \mathcal{B}_n(X)$, $E_n = \mathcal{B}_{n+1}(X)$, and for $w = w_0, w_1, \dots, w_n \in E_n$ we set the initial vertex $i(w)$ of w to be $w_0, \dots, w_{n-1} \in V_n$, the terminal vertex $t(w)$ of w is $w_1, \dots, w_n \in V_n$, and the label $\tau_n(w) = w_0 \in \mathcal{A}$. The shift space X_n presented by G_n is clearly sofic, and by Proposition 3.62 in [36], it satisfies $\mathcal{B}_j(X_n) = \mathcal{B}_j(X)$ for $j = 1, \dots, n + 1$. It is now easy to see that X_n is the *n*th topological Markov approximations for X .

3. *Alternative (pseudo)metrics for \mathcal{A}^∞ and $\mathcal{M}_\sigma(\mathcal{A}^\infty)$ and their hyperspaces*

3.1. *The functions \underline{d} , \bar{d} and $\bar{d}_{\mathcal{M}}$.* We now discuss the pre- and pseudometrics \underline{d} and \bar{d} for the space \mathcal{A}^∞ and Ornstein’s metric $\bar{d}_{\mathcal{M}}$ on $\mathcal{M}_\sigma(\mathcal{A}^\infty)$. These functions are not compatible with the natural topologies on \mathcal{A}^∞ and $\mathcal{M}_\sigma(\mathcal{A}^\infty)$, which are the product topology and the weak* topology, respectively. Note that usually $\bar{d}_{\mathcal{M}}$ is also denoted by \bar{d} , but as we will juggle between \bar{d} and $\bar{d}_{\mathcal{M}}$ a lot, we decided to change the notation to avoid confusion. We call $\bar{d}_{\mathcal{M}}$ the ‘d-bar distance for measures’, and its pseudometric version for \mathcal{A}^∞ we call ‘pointwise d-bar’.

Given $x = (x_n)_{n=0}^\infty, y = (y_n)_{n=0}^\infty \in \mathcal{A}^\infty$, we define

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

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

The function \underline{d} is only a *premetric*, that is, \underline{d} is a real-valued, non-negative, symmetric function on $\mathcal{A}^\infty \times \mathcal{A}^\infty$ vanishing on the diagonal $\{(x, y) \in \mathcal{A}^\infty \times \mathcal{A}^\infty : x = y\}$. It is easy to see that the triangle inequality fails for \underline{d} on \mathcal{A}^∞ . The function \bar{d} is a *pseudometric*, that is, \bar{d} is a premetric satisfying the triangle inequality and the implication $\bar{d}(x, y) = 0 \implies x = y$ fails, so \bar{d} is not a metric on \mathcal{A}^∞ if \mathcal{A} has at least two elements. Furthermore, $\bar{d}, \underline{d}: \mathcal{A}^\infty \times \mathcal{A}^\infty \rightarrow [0, 1]$ are both shift invariant, that is, $\underline{d}(x, y) = \underline{d}(\sigma(x), \sigma(y))$ and $\bar{d}(x, y) = \bar{d}(\sigma(x), \sigma(y))$ for all $x, y \in \mathcal{A}^\infty$.

The distance function $\bar{d}_{\mathcal{M}}$ is defined using joinings. We say that a measure λ on $\mathcal{A}^\infty \times \mathcal{A}^\infty$ is a *joining* of shift invariant measures μ and ν , if λ is $\sigma \times \sigma$ -invariant and μ and ν are the marginal measures for λ under the projection to the first, respectively the second, coordinate. We denote by $J(\mu, \nu)$ the set of all joinings of μ and ν . It is a non-empty set because the product measure $\mu \times \nu$ is always a joining of μ and ν . We can now define $\bar{d}_{\mathcal{M}}$ distance on $\mathcal{M}_\sigma(\mathcal{A}^\infty)$ as follows:

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

where $d_0(x, y) = 1$ if $x_0 \neq y_0$ and $d_0(x, y) = 0$ otherwise. It is well known that \bar{d}_M is a metric on $\mathcal{M}_\sigma(\mathcal{A}^\infty)$ and the convergence in this metric implies weak* convergence. This metric was introduced by Ornstein (for more details, see [44]). Furthermore, the space $\mathcal{M}_\sigma(\mathcal{A}^\infty)$ under the \bar{d}_M metric is complete but not separable; in particular, the space $\mathcal{M}_\sigma(\mathcal{A}^\infty)$ endowed with \bar{d}_M is not compact. 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 strong-mixing and Bernoulli measures on \mathcal{A}^∞ . The entropy function $\mu \mapsto h(\mu)$ is continuous on $\mathcal{M}_\sigma(\mathcal{A}^\infty)$ under \bar{d}_M .

3.2. *Hausdorffification of \underline{d} , \bar{d} , and \bar{d}_M .* In order to introduce the notion of approximation of one shift with another, we extend the functions \bar{d}_M , \bar{d} and \underline{d} so that they make sense for pairs of subsets of \mathcal{A}^∞ and $\mathcal{M}_\sigma(\mathcal{A}^\infty)$ following the well-known construction of the Hausdorff metric.

In general, a Hausdorff premetric may be defined on the powerset of any bounded premetric (or pseudometric, or metric) space. If (Z, ρ) is a set equipped with a bounded premetric, then we define the *Hausdorff premetric* ρ^H induced by ρ on the space of all non-empty subsets of Z . For a point $z \in Z$ and non-empty $A, B \subseteq Z$, we put

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

The function ρ^H is a premetric, which becomes a pseudometric whenever Z is a bounded pseudometric space. Even if ρ is a bounded metric on Z , its Hausdorff counterpart induced on the powerset of Z might still be only a pseudometric. This is because for non-empty $A, B \subset Z$ we have $\rho(A, B) = \rho(\bar{A}, \bar{B})$, where $(\bar{\cdot})$ is the closure operator naturally defined on the powerset of (Z, ρ) . Nevertheless, if ρ is a bounded metric, then the Hausdorff pseudometric ρ^H induced by ρ becomes a metric when restricted to the set $\text{CL}(Z, \rho)$ of closed non-empty subsets of (Z, ρ) . We apply these ideas to the metric space $(\mathcal{M}_\sigma(\mathcal{A}^\infty), \bar{d}_M)$, the pseudometric space $(\mathcal{A}^\infty, \bar{d})$, and the premetric space $(\mathcal{A}^\infty, \underline{d})$, obtaining a metric \bar{d}_M^H on the space $\text{CL}(\mathcal{M}_\sigma(\mathcal{A}^\infty), \bar{d}_M)$ of non-empty closed subsets of $(\mathcal{M}_\sigma(\mathcal{A}^\infty), \bar{d}_M)$ together with a pseudometric \bar{d}^H and a premetric \underline{d}^H on the set of all non-empty subsets of \mathcal{A}^∞ .

Below, we discuss some properties of the convergence of sets with respect to the Hausdorff metric ρ^H in the case when (Z, ρ) is not necessarily compact bounded pseudometric space. In this setting some properties, well known in the compact case, fail. For example, equivalent metrics, that is, metrics $\rho, \tilde{\rho}$ inducing the same topology on Z , may induce non-homeomorphic spaces $(\text{CL}(Z), \rho^H)$ and $(\text{CL}(Z), \tilde{\rho}^H)$.

The following propositions gather some properties of the Hausdorff metric. Proofs can be found in the literature (see [25, §2.15]).

PROPOSITION 1. *Let (Z, ρ) be a bounded metric space. If (Z, ρ) is a complete metric space, then so is $(\text{CL}(Z), \rho^H)$.*

It follows from Proposition 1 that the Hausdorff metric \bar{d}_M^H induced for $\text{CL}(\mathcal{M}_\sigma(\mathcal{A}^\infty))$ by \bar{d}_M is complete. Therefore the usual Cauchy condition provides a criterion for convergence, 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

$\mathcal{M} = \mathcal{M}_\sigma(X)$. Nevertheless, if $X_1 \supseteq X_2 \supseteq \dots$, then we can set X to be the intersection of all X_n .

PROPOSITION 2. *If $(X_n)_{n=1}^\infty$ is a decreasing sequence of shift spaces over \mathcal{A} , then $X = \bigcap_{n=1}^\infty X_n$ is a non-empty shift space such that $\mathcal{M}_\sigma(X) = \bigcap_{n=1}^\infty \mathcal{M}_\sigma(X_n)$.*

Proof. Since $X_1 \supseteq X_2 \supseteq \dots \supseteq X$, we have

$$\mathcal{M}_\sigma(X) = \mathcal{M}_\sigma\left(\bigcap_{n=1}^\infty X_n\right) \subseteq \bigcap_{n=1}^\infty \mathcal{M}_\sigma(X_n).$$

On the other hand, if $\mu \in \bigcap_{n=1}^\infty \mathcal{M}_\sigma(X_n)$, then $\mu(X_n) = 1$ for every $n \geq 1$. Therefore $\mu(X) = 1$. It follows that $\mu \in \mathcal{M}_\sigma(X)$. □

LEMMA 3. *If $(X_n)_{n=1}^\infty$ is a decreasing sequence of shift spaces over \mathcal{A} such that the sequence $(\mathcal{M}_\sigma(X_n))_{n=1}^\infty$ is $\bar{d}_{\mathcal{M}}^H$ -Cauchy, then $X = \bigcap_{n=1}^\infty X_n$ is a shift space such that $\bar{d}_{\mathcal{M}}^H(\mathcal{M}_\sigma(X_n), \mathcal{M}_\sigma(X)) \rightarrow 0$ as $n \rightarrow \infty$.*

Proof. It follows from Proposition 1 that $(\mathcal{M}_\sigma(X_k))_{k=1}^\infty$ converges to some $\mathcal{M} \in \text{CL}(\mathcal{A}^\infty, \bar{d}_{\mathcal{M}})$ with respect to $\bar{d}_{\mathcal{M}}^H$. Using Proposition 2, we can easily identify this limit with $\mathcal{M}_\sigma(X)$. □

4. \bar{d} -approachability

In this section we introduce \bar{d} -approachable shift spaces which are limits of their own canonical finite type approximations not only in the ‘usual’ Hausdorff metric topology, but also in the topology given by the \bar{d}^H pseudometric.

Definition 4. We say that a shift space $X \subseteq \mathcal{A}^\infty$ is \bar{d} -approachable if the sequence X_1^M, X_2^M, \dots of its Markov approximations of X satisfies $\bar{d}^H(X_n^M, X) \rightarrow 0$ as $n \rightarrow \infty$.

Clearly, every shift of finite type is \bar{d} -approachable. More examples of \bar{d} -approachable shift spaces will follow from the characterization of chain-mixing \bar{d} -approachable shift spaces by a notion we call \bar{d} -shadowing. The \bar{d} -shadowing property is closely related to the *average shadowing property* introduced by Blank [6] and discussed in more detail in [35]. Actually, following the ideas presented in [35], one can prove that a shift space X with the average shadowing property also has \bar{d} -shadowing. Furthermore, for a surjective shift space X (i.e. when $\sigma(X) = X$), \bar{d} -shadowing implies the average shadowing property. We leave the details to the interested reader. We decided to use \bar{d} -shadowing because it is much easier to use for symbolic systems and allows us to keep our paper self-contained by avoiding a very technical detour into topological dynamics. We will apply a similar strategy to several other notions: instead of presenting a general (i.e. stated for continuous maps acting on compact metric spaces) definition of some properties (e.g. chain transitivity, chain mixing, specification and its variants), we will present equivalent definitions adapted to the symbolic dynamics setting.

Definition 5. A shift space X has the \bar{d} -shadowing property if for every $\varepsilon > 0$ there is $N \in \mathbb{N}$ such that for every sequence $(w^{(j)})_{j=1}^\infty$ of words in $\mathcal{B}(X)$ satisfying $|w^{(j)}| \geq N$ for $j = 1, 2, \dots$ there is $x' \in X$ such that setting $x = w^{(1)}w^{(2)}w^{(3)} \dots$, we have $\bar{d}(x, x') < \varepsilon$.

It is easy to see that \bar{d} -shadowing implies \bar{d} -approachability, but to prove the converse we need an additional assumption. We say that a shift space is *chain transitive* (respectively, *chain mixing*) if X_n^M is transitive (respectively, topologically mixing) for all except finitely many n . For shift spaces this definition is equivalent to the usual one phrased in terms of δ -chains (see [36]).

Our main result in this section characterizes chain-mixing \bar{d} -approachable shift spaces. As mentioned above, this characterization is a topological counterpart of the result saying that a totally ergodic shift invariant probability measure is Bernoulli if and only if its canonical k -step Markov approximations converge to μ with respect to \bar{d}_M .

THEOREM 6. *For a shift space $X \subseteq \mathcal{A}^\infty$ the following conditions are equivalent.*

- (1) *There exists a descending sequence $(X_n)_{n=1}^\infty$ of mixing sofic shifts such that $X = \bigcap_{n=1}^\infty X_n$ and $\bar{d}^H(X, X_n) \rightarrow 0$ as $n \rightarrow \infty$.*
- (2) *$\sigma(X) = X$ and X has the \bar{d} -shadowing property.*
- (3) *X is chain mixing and \bar{d} -approachable.*

As a consequence of Theorem 6 we obtain that a transitive but not topologically mixing shift of finite type is an example of a \bar{d} -approachable shift space which is not chain mixing, hence it does not have the \bar{d} -shadowing property. In Example 11 we show that there are non- \bar{d} -approachable transitive sofic shifts. It follows that condition (1) in Theorem 6 cannot be relaxed.

The proof of the following lemma is straightforward. We will later generalize it in §6; see Lemma 22 there.

LEMMA 7. *Every mixing sofic shift space has the \bar{d} -shadowing property.*

Proof. Recall that the condition saying that X is a mixing sofic shift space implies that there exists $k \geq 0$ such that for every two words $u, w \in \mathcal{B}(X)$ one can find (this property is called the *specification property* and we discuss it in more detail in §6) a word v of length k such that $uvw \in \mathcal{B}(X)$. Fix $\varepsilon > 0$. Let N be such that $k/N < \varepsilon$. Let $(w^{(j)})_{j=1}^\infty$ be a sequence of words in $\mathcal{B}(X)$ satisfying $|w^{(j)}| \geq N$ for $j = 1, 2, \dots$. For $j = 1, 2, \dots$, we set $\ell(j) = |w^{(j)}|$ and let $u^{(j)} = w_1^{(j)} \dots w_{\ell(j)-k}^{(j)}$ be the prefix of $w^{(j)}$ of length $\ell(j) - k$. Using the above-mentioned property of mixing sofic shifts, we can find a sequence $(v^{(j)})_{j=1}^\infty$ of words of length k such that

$$x = u^{(1)}v^{(1)}u^{(2)}v^{(2)} \dots u^{(n)}v^{(n)} \dots$$

belongs to X . We also have

$$\bar{d}(x, w^{(1)}w^{(2)} \dots w^{(n)} \dots) \leq k/N < \varepsilon. \quad \square$$

The proof of the next lemma is adapted from the proof of [39, Lemma 25].

LEMMA 8. If $X \subseteq \mathcal{A}^\infty$ is a shift space such that there exists a sequence $(X_n)_{n=1}^\infty$ of shift spaces with the \bar{d} -shadowing property such that $X \subseteq X_n$ for every $n \in \mathbb{N}$ and $\bar{d}^H(X, X_n) \rightarrow 0$ as $n \rightarrow \infty$, then X also has \bar{d} -shadowing.

Proof. Fix $\varepsilon > 0$. Let $N \in \mathbb{N}$ be such that for every $n \geq N$ we have

$$\bar{d}^H(X, X_n) < \varepsilon/2. \tag{1}$$

Use \bar{d} -shadowing of X_N to find $M \geq 0$ such that if $y = w^{(1)}w^{(2)}w^{(3)} \dots \in \mathcal{A}^\infty$ satisfies $w^{(j)} \in \mathcal{B}(X_N)$ and $|w^{(j)}| \geq M$ for every $j \in \mathbb{N}$, then one can find $\bar{x} \in X_N$ satisfying

$$\bar{d}(\bar{x}, y) < \varepsilon/2. \tag{2}$$

Since $\mathcal{B}(X) \subseteq \mathcal{B}(X_N)$, we can apply (1) and (2) to $y = w^{(1)}w^{(2)}w^{(3)} \dots \in \mathcal{A}^\infty$ satisfying $w^{(j)} \in \mathcal{B}(X)$ and $|w^{(j)}| \geq M$ for every $j \in \mathbb{N}$. It follows that there is $x \in X$ such that $\bar{d}(x, y) < \varepsilon$. □

The proof of the following lemma is adapted from the proof of [35, Lemma 3.1].

LEMMA 9. If a shift space $X \subseteq \mathcal{A}^\infty$ has the \bar{d} -shadowing property and $\sigma(X) = X$, then X is chain mixing.

Proof. Assume that $\sigma(X) = X$. Recall that X is chain mixing if and only if $X \times X$ is chain transitive. Note that it is enough to prove that X is chain transitive, because if X has \bar{d} -shadowing and $\sigma(X) = X$, then the same holds for $X \times X$. Fix n . We will show that the Markov approximation X_n is transitive, that is, for every $u, w \in \mathcal{B}_{n-1}(X)$ there is a word v such that every word of length n appearing in uvw belongs to $\mathcal{B}_n(X)$. Fix $0 < \varepsilon < 1/(4n)$. Use \bar{d} -shadowing to find N for that ε . Set $n_0 := \max\{N, n\}$ and $n_j = 2^j n_0$ for $j \in \mathbb{N}$. For $j \geq 0$ let $w^{(2j)}$ be any word in $\mathcal{B}(X)$ such that $|w^{(2j)}| = n_{2j}$ and u is the prefix of $w^{(2j)}$, and let $w^{(2j+1)}$ be any word in $\mathcal{B}(X)$ such that $|w^{(2j+1)}| = n_{2j+1}$ and w is the suffix of $w^{(2j+1)}$ (the latter word exists because $\sigma(X) = X$). Thus for every $j \in \mathbb{N}_0$ we have words $u^{(2j)}, v^{(2j+1)}$ such that

$$w^{(2j)} = uu^{(2j)} \quad \text{and} \quad w^{(2j+1)} = v^{(2j+1)}w.$$

Let $y = w^{(0)}w^{(1)}w^{(2)} \dots$. Write $\xi(0) = 0$ and

$$\xi(k + 1) = |w^{(0)}| + \dots + |w^{(k)}|.$$

Therefore

$$y_{[\xi(2j), \xi(2j+1)]}y_{[\xi(2j+1), \xi(2j+2)]} = w^{(2j)}w^{(2j+1)} = uu^{(2j)}v^{(2j+1)}w.$$

By the \bar{d} -shadowing property there is $x \in X$ such that $\bar{d}(x, y) < \varepsilon$. We claim that there exists $j \geq 0$ such that for some $\xi(2j) + n \leq p < \xi(2j + 1) - n$ and $\xi(2j + 1) \leq q < \xi(2j + 2) - 2n$ we have

$$x_{[p, p+n]} = y_{[p, p+n]} \quad \text{and} \quad x_{[q, q+n]} = y_{[q, q+n]}.$$

It follows that $y_{[\xi(2j), p]}x_{[p, q+n]}y_{[q+n, \xi(2j+2)]} \in \mathcal{B}(X_n)$. If the claim is false, then easy computations show that $\bar{d}(x, y) \geq \varepsilon$, which would be a contradiction. This completes our proof. □

LEMMA 10. Every shift space with the \bar{d} -shadowing property is \bar{d} -approachable.

Proof. Let X be a shift space with the \bar{d} -shadowing property. We need to show that $\bar{d}^H(X_n^M, X) \rightarrow 0$ as $n \rightarrow \infty$. Fix $\varepsilon > 0$. For this ε we choose N using the \bar{d} -shadowing property. Let $n \geq N$. Since $X \subseteq X_n^M$ it is enough to check that if $x \in X_n^M$ then one can find $y \in X$ with $\bar{d}(x, y) < \varepsilon$. Hence we fix $x \in X_n^M$ and define the sequence $w^{(j+1)} = x_{[nj, n(j+1))} = x_{nj}x_{nj+1} \dots x_{n(j+1)-1}$ for $j = 0, 1, 2, \dots$. Since for each $k \geq 1$ we have $w^{(k)} \in \mathcal{B}_n(X)$, because $\mathcal{B}_n(X_n) = \mathcal{B}_n(X)$, we can apply \bar{d} -shadowing to find $y \in X$ such that

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

It follows $\bar{d}^H(X_n^M, X) < \varepsilon$ for $n \geq N$ as needed. □

We finish this section with the proof of the characterization theorem.

Proof of Theorem 6. (1) \implies (2) First, note that $\sigma(X) = X$ because $\sigma(X_n) = X_n$ whenever X_n is a mixing sofic shift and σ is finite-to-one. Note that for every $n \in \mathbb{N}$ the shift space X_n has the \bar{d} -shadowing property by Lemma 7. Then we apply Lemma 8.

(2) \implies (3) This follows from Lemmas 10 and 9.

(3) \implies (1) This is straightforward consequence of the fact that Markov approximations for a chain-mixing shift space are mixing sofic shifts. □

In the next section we will see that if a shift space $X \subseteq \mathcal{A}^\infty$ is \bar{d} -approachable and chain transitive, then its ergodic measures are entropy-dense.

Example 11. We show that even very simple transitive sofic shifts may have Markov approximations that are away in \bar{d} . This example also shows that we cannot replace mixing sofic shifts in Theorem 6 by transitive sofic shifts.

The shift space X is defined as the set of all sequences $x \in \{0, 1\}^\infty$ such that $x \in X$ if and only if $x_{2i} = 0$ for every $i \geq 0$ or $x_{2i+1} = 0$ for every $i \geq 0$. Clearly, $X = \tilde{X}$ is a hereditary transitive sofic shift, which is not topologically mixing. Given $k \in \mathbb{N}$, we define a periodic point

$$y^{(k)} = y_0^{(k)}y_1^{(k)} \dots = ((10)^{k+1}0^{2k+1})^\infty.$$

We note that $y^{(k)} \in X_{2k}^M$ for every $k \in \mathbb{N}$. By definition, for every $k \in \mathbb{N}$ and $j > 0$, the symbol 1 occurs $2j(k + 1)$ times in the block $y_{[0, 2j(4k+3))}^{(k)}$. These occurrences are evenly distributed among even and odd positions. Therefore for every $x \in \tilde{X} = X$ and $j \in \mathbb{N}$ we have

$$d_{\text{Ham}}(y_{[0, 2j(4k+3))}^{(k)}x_{[0, 2j(4k+3)}) \geq \frac{j(k + 1)}{j(8k + 6)} \geq \frac{1}{8},$$

It follows immediately that for every $k \in \mathbb{N}$ we have

$$\underline{d}(y^{(k)}, x) \geq \frac{1}{8}.$$

Therefore for each $k \in \mathbb{N}$,

$$\underline{d}^H(X, X_{2k}^M) \geq \frac{1}{8},$$

that is, X is an example of a non- \bar{d} -approachable transitive sofic shift. In addition, since every point $y^{(k)}$ is generic for an ergodic measure $\nu^{(k)}$ on X_{2k}^M we can conclude that for each $k \in \mathbb{N}$,

$$\bar{d}_{\mathcal{M}}^H(\mathcal{M}_\sigma(X), \mathcal{M}_\sigma(X_{2k}^M)) \geq \frac{1}{8}.$$

This means that we can replace Markov approximation by the approximation by sofic shifts only if these sofic shift are mixing as in condition (1) of Theorem 6.

5. Approximation and entropy density

In this section we discuss approximation schemes of shift spaces related to the pseudo-metric \bar{d} and the metric $\bar{d}_{\mathcal{M}}$. Then we show that these approximation schemes allow for a transfer of entropy density.

Recall that 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 usual hyperspace topology if $\rho^H(X_n, X) \rightarrow 0$ as $n \rightarrow \infty$, where ρ^H is the Hausdorff metric corresponding to any metric ρ compatible with the product topology on \mathcal{A}^∞ . We similarly define the convergence in the usual hyperspace topology of $\mathcal{M}_\sigma(\mathcal{A}^\infty)$, and we have that $\mathcal{M}_\sigma(X_n)$ converges to $\mathcal{M}_\sigma(X)$ as $n \rightarrow \infty$ if $D^H(X_n, X) \rightarrow 0$ as $n \rightarrow \infty$, where D^H is the Hausdorff metric corresponding to an arbitrary metric D compatible with the weak* topology on $\mathcal{M}_\sigma(\mathcal{A}^\infty)$ (we can replace D by any other compatible metric).

We will consider four new ways in which a sequence of shift spaces $(X_n)_{n=1}^\infty \subseteq \mathcal{A}^\infty$ can approximate a shift space $X \subseteq \mathcal{A}^\infty$. These ways are given by the conditions

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

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

$$\lim_{n \rightarrow \infty} \underline{d}^H(X_n, X) = 0, \tag{5}$$

$$\lim_{n \rightarrow \infty} \bar{d}^H(X_n, X) = 0. \tag{6}$$

Our aim is to show that

$$(6) \implies (5) \implies ((4) \iff (3)),$$

and any of these modes of convergence allows us to infer entropy density for X provided that entropy density holds for X_n for all n . We stress that our scheme does not require the approximation to be monotone, that is, we do not require $X_1 \supseteq X_2 \supseteq \dots$ and $X = \bigcap X_n$, although in practice these often hold. Bearing in mind future applications, we also include a result (Corollary 17) dealing with monotone limits of shift spaces.

Our first main result states that $\bar{d}_{\mathcal{M}}^H$ -convergence for simplices of invariant measures given by (3) preserves entropy density.

THEOREM 12. Let $(X_k)_{k=1}^\infty$ and X be shift spaces over \mathcal{A} such that

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

If ergodic measures are entropy-dense in $\mathcal{M}_\sigma(X_k)$ for each $k \in \mathbb{N}$, then ergodic measures are entropy-dense in $\mathcal{M}_\sigma(X)$.

A key component of the proof of Theorem 12 is the fact that for every shift space X , we have the equality

$$\bar{d}_{\mathcal{M}}^H(\mathcal{M}_\sigma(X_k), \mathcal{M}_\sigma(X)) = \bar{d}_{\mathcal{M}}^H(\mathcal{M}_\sigma^e(X_k), \mathcal{M}_\sigma^e(X)),$$

which implies the equivalence (4) \iff (3). This is established in the following two lemmas.

LEMMA 13. Let $Y \subseteq \mathcal{A}^\infty$ be a shift space and let $\mu \in \mathcal{M}_\sigma^e(\mathcal{A}^\infty)$ be an ergodic measure. Then $\bar{d}_{\mathcal{M}}(\mu, \mathcal{M}_\sigma^e(Y)) = \bar{d}_{\mathcal{M}}(\mu, \mathcal{M}_\sigma(Y))$.

Proof. It is enough to show that for every $\mu \in \mathcal{M}_\sigma^e(\mathcal{A}^\infty)$ and $\nu \in \mathcal{M}_\sigma(Y)$ there exists $\nu' \in \mathcal{M}_\sigma^e(Y)$ such that $\bar{d}_{\mathcal{M}}(\mu, \nu') \leq \bar{d}_{\mathcal{M}}(\mu, \nu)$. Let $\bar{\lambda}$ be a joining which realizes the $\bar{d}_{\mathcal{M}}$ distance between $\mu \in \mathcal{M}_\sigma^e(\mathcal{A}^\infty)$ and $\nu \in \mathcal{M}_\sigma(Y)$, that is,

$$\bar{d}_{\mathcal{M}}(\mu, \nu) = \int_{\mathcal{A}^\infty \times Y} d_0(x, y) \, d\bar{\lambda}(x, y).$$

Let $\bar{\xi}$ be the ergodic decomposition of $\bar{\lambda}$. We have

$$\int_{\mathcal{A}^\infty \times Y} d_0(x, y) \, d\bar{\lambda}(x, y) = \int_{\mathcal{M}_{\sigma \times \sigma}^e(\mathcal{A}^\infty \times Y)} \left(\int_{\mathcal{A}^\infty \times Y} d_0(x, y) \, d\lambda(x, y) \right) \, d\bar{\xi}(\lambda).$$

Let E be the set of all $\lambda \in \mathcal{M}_{\sigma \times \sigma}^e(\mathcal{A}^\infty \times Y)$ such that

$$\int_{\mathcal{A}^\infty \times Y} d_0(x, y) \, d\lambda(x, y) \leq \bar{d}_{\mathcal{M}}(\mu, \nu).$$

It is clear that E satisfies $\bar{\xi}(E) > 0$. Additionally, since μ is ergodic, for $\bar{\xi}$ -almost every $\lambda \in \mathcal{M}_{\sigma \times \sigma}^e(\mathcal{A}^\infty \times Y)$ the pushforward of λ through the projection onto the first coordinate is μ . Hence, there exists $\lambda' \in \mathcal{M}_{\sigma \times \sigma}^e(\mathcal{A}^\infty \times Y)$ such that $(\pi_{\mathcal{A}^\infty})_*(\lambda') = \mu$ and

$$\int_{\mathcal{A}^\infty \times Y} d_0(x, y) \, d\lambda'(x, y) \leq \bar{d}_{\mathcal{M}}(\mu, \nu).$$

Let $\nu' = (\pi_Y)_*(\lambda')$. Then ν' is an ergodic measure supported on Y and λ' is a joining of μ and ν' . Consequently, $\bar{d}_{\mathcal{M}}(\mu, \nu') \leq \bar{d}_{\mathcal{M}}(\mu, \nu)$. \square

LEMMA 14. If X and Y are shift spaces over \mathcal{A} , then

$$\bar{d}_{\mathcal{M}}^H(\mathcal{M}_\sigma(X), \mathcal{M}_\sigma(Y)) = \bar{d}_{\mathcal{M}}^H(\mathcal{M}_\sigma^e(X), \mathcal{M}_\sigma^e(Y)).$$

Proof. It follows from Lemma 13 that

$$\begin{aligned} & \bar{d}^H(\mathcal{M}_\sigma^e(X), \mathcal{M}_\sigma^e(Y)) \\ &= \max \left(\sup_{\mu \in \mathcal{M}_\sigma^e(X)} \bar{d}_{\mathcal{M}}(\mu, \mathcal{M}_\sigma(Y)), \sup_{\nu \in \mathcal{M}_\sigma^e(Y)} \bar{d}_{\mathcal{M}}(\nu, \mathcal{M}_\sigma(X)) \right). \end{aligned} \tag{7}$$

Since $\mathcal{M}_\sigma^e(X) \subset \mathcal{M}_\sigma(X)$ we have

$$\sup_{\mu \in \mathcal{M}_\sigma^e(X)} \bar{d}_{\mathcal{M}}(\mu, \mathcal{M}_\sigma(Y)) \leq \sup_{\mu \in \mathcal{M}_\sigma(X)} \bar{d}_{\mathcal{M}}(\mu, \mathcal{M}_\sigma(Y)),$$

and the same inequality with the roles of X and Y reversed. These inequalities, together with (7), give us

$$\begin{aligned} &\bar{d}^H(\mathcal{M}_\sigma^e(X), \mathcal{M}_\sigma^e(Y)) \\ &= \max \left(\sup_{\mu \in \mathcal{M}_\sigma^e(X)} \bar{d}_{\mathcal{M}}(\mu, \mathcal{M}_\sigma(Y)), \sup_{\nu \in \mathcal{M}_\sigma^e(Y)} \bar{d}_{\mathcal{M}}(\nu, \mathcal{M}_\sigma(X)) \right) \\ &\leq \bar{d}_{\mathcal{M}}^H(\mathcal{M}_\sigma(X), \mathcal{M}_\sigma(Y)). \end{aligned}$$

It remains to prove that for every $\mu \in \mathcal{M}_\sigma(X)$, there exists $\nu \in \mathcal{M}_\sigma(Y)$ such that $\bar{d}_{\mathcal{M}}(\mu, \nu) \leq \bar{d}_{\mathcal{M}}^H(\mathcal{M}_\sigma^e(X), \mathcal{M}_\sigma^e(Y))$. This is sufficient, since the roles of X and Y are interchangeable, meaning that

$$\bar{d}_{\mathcal{M}}^H(\mathcal{M}_\sigma(X), \mathcal{M}_\sigma(Y)) \leq \bar{d}^H(\mathcal{M}_\sigma^e(X), \mathcal{M}_\sigma^e(Y)).$$

Recall that finite convex combinations of ergodic measures are weak* dense in $\mathcal{M}_\sigma(X)$. Therefore we have

$$\mu = \lim_{n \rightarrow \infty} \sum_{j \in I(n)} \alpha_j^{(n)} \mu_j^{(n)},$$

where $I(n)$ is a finite set for every $n \in \mathbb{N}$; furthermore, for a fixed n and for each $j \in I(n)$ we have $\mu_j^{(n)} \in \mathcal{M}_\sigma^e(X)$, $\alpha_j^{(n)} > 0$ and $\sum_{j \in I(n)} \alpha_j^{(n)} = 1$. By our assumption, to every $\mu_j^{(n)}$ there correspond an ergodic measure $\nu_j^{(n)} \in \mathcal{M}_\sigma^e(Y)$ and a joining $\lambda_j^{(n)} \in J(\mu_j^{(n)}, \nu_j^{(n)})$ with

$$\bar{d}_{\mathcal{M}}(\mu_j^{(n)}, \nu_j^{(n)}) = \int_{X \times Y} d_0(x, y) \, d\lambda_j^{(n)}(x, y) \leq \bar{d}_{\mathcal{M}}(\mathcal{M}_\sigma^e(X), \mathcal{M}_\sigma^e(Y)). \tag{8}$$

For each $n \in \mathbb{N}$ define

$$\lambda^{(n)} = \sum_{j \in I(n)} \alpha_j^{(n)} \lambda_j^{(n)}.$$

Using compactness of $\mathcal{M}_{\sigma \times \sigma}(X \times Y)$. we may assume that the sequence $(\lambda^{(n)})_{n \in \mathbb{N}}$ weak* converges to some $\bar{\lambda} \in \mathcal{M}_{\sigma \times \sigma}(X \times Y)$. By the definition of weak* topology, it follows immediately that

$$\int_{X \times Y} d_0(x, y) \, d\bar{\lambda}(x, y) = \lim_{n \rightarrow \infty} \int_{X \times Y} d_0(x, y) \, d\lambda^{(n)}(x, y). \tag{9}$$

It is easy to see that $\bar{\lambda}$ is a joining of μ with $\nu = (\pi_Y)_*(\bar{\lambda}) \in \mathcal{M}_\sigma(Y)$. Definition of $\lambda^{(n)}$ yields that for every $n \in \mathbb{N}$ we have

$$\int_{X \times Y} d_0(x, y) \, d\lambda^{(n)}(x, y) = \sum_{j \in I(n)} \alpha_j^{(n)} \int_{X \times Y} d_0(x, y) \lambda_j^{(n)}(x, y). \tag{10}$$

Combining (8), (9), and (10), we obtain

$$\int_{X \times Y} d_0(x, y) \, d\bar{\lambda}(x, y) \leq \bar{d}_{\mathcal{M}}^H(\mathcal{M}_\sigma^e(X), \mathcal{M}_\sigma^e(Y)).$$

Hence, for each $\mu \in \mathcal{M}_\sigma(X)$ we have $\nu \in \mathcal{M}_\sigma(Y)$ such that $\bar{d}_{\mathcal{M}}(\mu, \nu) \leq \bar{d}_{\mathcal{M}}^H(\mathcal{M}_\sigma^e(X), \mathcal{M}_\sigma^e(Y))$. \square

Having Lemma 14 at our disposal, we now proceed with the proof of Theorem 12.

Proof of Theorem 12. Fix $\varepsilon > 0$ and $\nu \in \mathcal{M}_\sigma(X)$. Let D be any metric compatible with the weak* topology on $\mathcal{M}_\sigma(\mathcal{A}^\infty)$. Since the entropy function $\mu \mapsto h(\mu)$ is uniformly continuous on $\mathcal{M}_\sigma(\mathcal{A}^\infty)$ endowed with the metric $\bar{d}_{\mathcal{M}}$ [51], we can find $\delta_1 > 0$ such that $|h(\mu) - h(\mu')| < \varepsilon/3$ for any $\mu, \mu' \in \mathcal{M}_\sigma(\mathcal{A}^\infty)$ with $\bar{d}_{\mathcal{M}}(\mu, \mu') \leq \delta_1$. By [51, Theorem 7.7], there exists $\delta_2 > 0$ such that $D(\mu, \mu') < \varepsilon/3$ for any $\mu, \mu' \in \mathcal{M}_\sigma(\mathcal{A}^\infty)$ with $\bar{d}_{\mathcal{M}}(\mu, \mu') \leq \delta_2$. Put $\delta = \min(\delta_1, \delta_2)$ and fix $k \in \mathbb{N}$ such that (cf. Lemma 14)

$$\bar{d}_{\mathcal{M}}^H(\mathcal{M}_\sigma(X_k), \mathcal{M}_\sigma(X)) = \bar{d}_{\mathcal{M}}^H(\mathcal{M}_\sigma^e(X_k), \mathcal{M}_\sigma^e(X)) < \delta. \tag{11}$$

By (11), there exists $\mu \in \mathcal{M}_\sigma(X_k)$ such that $\bar{d}_{\mathcal{M}}(\nu, \mu) < \delta$, whence

$$D(\nu, \mu) < \varepsilon/3 \quad \text{and} \quad |h(\nu) - h(\mu)| < \varepsilon/3. \tag{12}$$

Since $\mathcal{M}_\sigma^e(X_k)$ is entropy-dense in $\mathcal{M}_\sigma(X_k)$, there exists $\mu' \in \mathcal{M}_\sigma^e(X_k)$ that satisfies

$$D(\mu, \mu') < \varepsilon/3 \quad \text{and} \quad |h(\mu) - h(\mu')| < \varepsilon/3. \tag{13}$$

By another application of (11), there exists $\nu' \in \mathcal{M}_\sigma^e(X)$ such that $\bar{d}_{\mathcal{M}}(\nu', \mu') < \delta$, so

$$D(\nu', \mu') < \varepsilon/3 \quad \text{and} \quad |h(\nu') - h(\mu')| < \varepsilon/3. \tag{14}$$

Combining (12), (13) and (14), we conclude that

$$D(\nu, \nu') < \varepsilon \quad \text{and} \quad |h(\nu) - h(\nu')| < \varepsilon. \tag{15} \quad \square$$

Note that Theorem 12 only transfers entropy density from sequences of simplices to their $\bar{d}_{\mathcal{M}}^H$ -limits. It turns out that very useful bounds for $\bar{d}_{\mathcal{M}}^H$ are provided by \underline{d}^H and \bar{d}^H . These bounds yield immediately the implications (6) \implies (5) \implies (3). This allows us to work directly with shift spaces bypassing the need to determine their simplices of invariant measures.

PROPOSITION 15. *If X and Y are shift spaces over \mathcal{A} , then*

$$\bar{d}_{\mathcal{M}}^H(\mathcal{M}_\sigma^e(X), \mathcal{M}_\sigma^e(Y)) \leq \underline{d}^H(X, Y) \leq \bar{d}^H(X, Y).$$

Proof. The second inequality is obvious. We turn to the proof of the first. Let $\mu \in \mathcal{M}_\sigma^e(X)$ and fix $\Delta > \underline{d}^H(X, Y)$. Pick a μ -generic point \bar{x} . By our assumptions, there is a point $\bar{y} \in Y$ such that $\underline{d}(\bar{x}, \bar{y}) < \Delta$. Let $(n_k)_{k \geq 1}$ be a strictly increasing sequence of integers such that

$$\underline{d}(\bar{x}, \bar{y}) = \lim_{k \rightarrow \infty} \frac{1}{n_k} |\{0 \leq j < n_k : \bar{x}_j \neq \bar{y}_j\}| < \Delta.$$

Passing to a subsequence if necessary, we assume that the point $(\bar{x}, \bar{y}) \in X \times Y$ generates a measure $\bar{\lambda} \in \mathcal{M}_{\sigma \times \sigma}(X \times Y)$, that is, the sequence of measures

$$\bar{\lambda}_k = \frac{1}{n_k} \sum_{j=0}^{n_k-1} \delta_{(\sigma \times \sigma)^j(\bar{x}, \bar{y})}$$

weak* converges to $\bar{\lambda}$ as $k \rightarrow \infty$. Let $\nu = (\pi_Y)_*(\bar{\lambda}) \in \mathcal{M}_\sigma(Y)$. Then $\bar{\lambda}$ is a joining of μ and ν . It follows from Lemma 14 that

$$\begin{aligned} \bar{d}_{\mathcal{M}}(\mu, \mathcal{M}_\sigma^e(Y)) &= \bar{d}_{\mathcal{M}}(\mu, \mathcal{M}_\sigma(Y)) \leq \bar{d}_{\mathcal{M}}(\mu, \nu) \leq \int_{X \times Y} d_0(x, y) \, d\bar{\lambda}(x, y) \\ &= \lim_{k \rightarrow \infty} \int_{X \times Y} d_0(x, y) \, d\bar{\lambda}_k(x, y) = \underline{d}(\bar{x}, \bar{y}) < \Delta. \end{aligned}$$

Since $\Delta > \underline{d}(X, Y)$ and $\mu \in \mathcal{M}_\sigma^e(X)$ were arbitrary, we have

$$\sup_{\mu \in \mathcal{M}_\sigma^e(X)} \bar{d}_{\mathcal{M}}(\mu, \mathcal{M}_\sigma^e(Y)) \leq \underline{d}^H(X, Y).$$

Noting that the roles of X and Y are interchangeable, we see that

$$\begin{aligned} &\bar{d}_{\mathcal{M}}^H(\mathcal{M}_\sigma^e(X), \mathcal{M}_\sigma^e(Y)) \\ &= \max \left(\sup_{\mu \in \mathcal{M}_\sigma^e(X)} \bar{d}_{\mathcal{M}}(\mu, \mathcal{M}_\sigma^e(Y)), \sup_{\nu \in \mathcal{M}_\sigma^e(Y)} \bar{d}_{\mathcal{M}}(\nu, \mathcal{M}_\sigma^e(X)) \right) \leq \underline{d}^H(X, Y). \quad \square \end{aligned}$$

In practice, we will not apply Theorem 12 directly, instead we will invoke one of the corollaries obtained from Theorem 12 and Proposition 15. The first of these is immediate.

COROLLARY 16. *Let $(X_n)_{n=1}^\infty$ be a sequence of shift spaces over \mathcal{A} with entropy density such that for a shift space $X \subseteq \mathcal{A}^\infty$ we have*

$$\lim_{n \rightarrow \infty} \underline{d}^H(X_n, X) = 0.$$

Then $\bar{d}_{\mathcal{M}}^H(\mathcal{M}_\sigma^e(X_n), \mathcal{M}_\sigma^e(X)) \rightarrow 0$ as $n \rightarrow \infty$ and the ergodic measures are entropy-dense for X .

Before stating the second practical corollary to Theorem 12, note that if $(X_n)_{n=1}^\infty$ is a sequence of shift spaces satisfying the Cauchy condition with respect to \bar{d}^H or \underline{d}^H , then using Proposition 15 we get that $\mathcal{M}_\sigma(X_n)_{n=1}^\infty$ satisfies the Cauchy condition for $\bar{d}_{\mathcal{M}}^H$, so the latter sequence converges in $\bar{d}_{\mathcal{M}}^H$ to some non-empty compact subset of $\mathcal{M}_\sigma(\mathcal{A}^\infty)$. Yet, it is not easy to identify the limit with $\mathcal{M}_\sigma(X)$ for some shift space X , unless $X_1 \supseteq X_2 \supseteq \dots$. Under this additional assumption the following result immediately follows from Lemma 3 and Corollary 16.

COROLLARY 17. *Let $(X_n)_{n=1}^\infty$ be a decreasing sequence of shift spaces over \mathcal{A} such that*

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

Then $X = \bigcap_{n=1}^{\infty} X_n$ is a shift space such that $\bar{d}_{\mathcal{M}}^H(\mathcal{M}_{\sigma}^e(X_n), \mathcal{M}_{\sigma}^e(X)) \rightarrow 0$ as $n \rightarrow \infty$. Furthermore, if ergodic measures are entropy-dense for X_n and each $n \geq 1$, then entropy density holds also for X .

We also find the following observation useful. We stress that neither \underline{d} nor \underline{d}^H obeys the triangle inequality, so the assumptions of Corollary 18 do not guarantee that $\underline{d}^H(X, Y) = 0$.

COROLLARY 18. *Suppose that $(X_k)_{k=1}^{\infty}$ and X, Y are shift spaces over \mathcal{A} such that $\underline{d}^H(X_k, X)$ and $\underline{d}^H(X_k, Y)$ tend to zero as $k \rightarrow \infty$. Then*

$$\mathcal{M}_{\sigma}(X) = \mathcal{M}_{\sigma}(Y).$$

Proof. This is a straightforward consequence of Proposition 15. □

In order to apply Theorem 12 or one of its corollaries we still need to single out a class of shift spaces for which entropy density is easily verifiable and then describe a family of shift spaces X such that for some sequence $(X_k)_{k=1}^{\infty}$ of shift spaces in the aforementioned class of shift spaces we have

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

A class of shifts where entropy density is easy to demonstrate is the family of transitive sofic shifts. The result is known even in a greater generality; see [20, 48]. The proofs given in [20, 48] simplify in the special case needed here (cf. the proof of Theorem 7.12 in [51]).

PROPOSITION 19. *Every transitive sofic shift space over a finite alphabet has an entropy-dense set of ergodic measures.*

In applications, we will use either Corollary 16 or Corollary 17 together with Proposition 19. For further reference we formulate our observation as a corollary.

COROLLARY 20. *Let $(X_k)_{k=1}^{\infty}$ be a sequence of transitive sofic shifts over \mathcal{A} . If X is a shift space over \mathcal{A} such that $\underline{d}^H(X_k, X) \rightarrow 0$ as $k \rightarrow \infty$ then ergodic measures are entropy-dense in $\mathcal{M}_{\sigma}(X)$.*

As the first application of Corollary 20 we may now prove entropy density of all \bar{d} -approachable and chain transitive shift spaces.

PROPOSITION 21. *If a shift space $X \subseteq \mathcal{A}^{\infty}$ is \bar{d} -approachable and chain transitive, then its ergodic measures are entropy-dense.*

Proof. Apply Corollary 20 with $X_k = X_k^M$. □

6. \bar{d} -approachability and specification

We prove that shift spaces satisfying popular variants of the specification property (specification and almost specification) are \bar{d} -approachable. Note that the definitions below are adapted to the symbolic setting. These formulations are equivalent to but slightly different than the ones usually used for general compact dynamical systems. For more

background on the specification property and its relatives we refer the reader to the survey paper [38].

A shift space $X \in \mathcal{A}^\infty$ has the *specification property* if there is $k \in \mathbb{N}$ such that for any words $u, w \in \mathcal{B}(X)$ one can find a word v of length k such that $uvw \in \mathcal{B}(X)$. A shift space X has the *almost specification property* if there is a mistake function $g: \mathbb{N} \rightarrow \mathbb{N}$, which is a non-decreasing function with $g(n)/n \rightarrow 0$ as $n \rightarrow \infty$ and such that for any words $u, w \in \mathcal{B}(X)$ there are words u', w' satisfying $|u| = |u'|$, $|v| = |v'|$, $d_{\text{Ham}}(u, u') \leq g(|u|)$ and $d_{\text{Ham}}(w, w') \leq g(|w|)$, such that $u'w' \in \mathcal{B}(X)$. Here and elsewhere, d_{Ham} stands for the normalized Hamming distance, that is,

$$d_{\text{Ham}}(u, w) = \frac{1}{n} |\{1 \leq j \leq n : u_j \neq w_j\}|,$$

where $n = |u| = |w|$.

A prototype for the almost specification property was the *g-almost product property* introduced in the context of general (not necessarily symbolic) dynamical systems by Pfister and Sullivan in [49]. Later, Thompson [55] proposed to slightly modify this notion and renamed it the *almost specification property*. We follow Thompson, hence our almost specification property is logically weaker (less restrictive) than the notion introduced by Pfister and Sullivan. The standard examples of shift spaces with the almost specification property are β -shifts (see [48]). For shift spaces it is easy to see that the specification property implies almost specification. For a generic $\beta > 1$ the β -shift X_β is an example of a shift space with almost specification but without specification (see [53]).

Both specification properties defined above are known to imply entropy density. For shifts with the specification property this was first proved in [20]. The almost specification property implies entropy density because it implies the approximate product property, and the latter implies entropy density by Theorem 2.1 of [48].

It is also easy to see that each of the specification properties considered here implies the \bar{d} -shadowing property (actually, the proof of Lemma 7 applies almost verbatim).

LEMMA 22. *For a shift space $X \subseteq \mathcal{A}^\infty$ specification implies almost specification. If a shift space $X \subseteq \mathcal{A}^\infty$ has the almost specification property, then it has also the \bar{d} -shadowing property.*

By Theorem 6, every shift space X with the almost specification property such that $\sigma(X) = X$ is \bar{d} -approachable. Unfortunately, this the latter condition (i.e. surjectivity) is not necessarily satisfied by a shift space with almost specification. For example, the shift space $X = \{0^\infty, 10^\infty\} \subseteq \{0, 1\}^\infty$ has the almost specification property, but $\sigma(X) \neq X$. Therefore we need to assume $\sigma(X) = X$ in our next result.

PROPOSITION 23. *Let $X \subseteq \mathcal{A}^\infty$ be a shift space with the almost specification property. If $\sigma(X) = X$, then X is \bar{d} -approachable and chain mixing.*

Proof. By Lemma 22, X has the \bar{d} -shadowing property. Hence X satisfies condition (2) of Theorem 6, so we conclude that X is \bar{d} -approachable. \square

Remark 24. One may wonder if \bar{d} -shadowing or \bar{d} -approachability implies uniqueness of the measure of maximal entropy. It follows from Proposition 23 and examples of surjective shift spaces with the almost specification property and multiple measures of maximal entropy presented in [40, 46] that this is not the case.

We can now show that the entropy density results from [20, 48] are simple consequences of entropy density of transitive sofic shifts combined with \bar{d} -approachability.

COROLLARY 25. *If $X \subseteq \mathcal{A}^\infty$ is a shift space with the almost specification property, then X has an entropy-dense set of ergodic measures.*

Proof. Let X^+ be the measure center of X , that is,

$$X^+ = X \setminus \bigcup \{U \subseteq \mathcal{A}^\infty : U \text{ open and } \mu(U) = 0 \text{ for every } \mu \in \mathcal{M}_\sigma(X)\}.$$

Clearly, $\mathcal{M}_\sigma(X) = \mathcal{M}_\sigma(X^+)$. Furthermore, a shift space X has the almost specification property if and only if X^+ does (see [58, Theorem 6.7] and [35, Theorem 5.1]). To finish the proof, note that $\sigma(X^+) = X^+$ because recurrent points are dense in X^+ by the Poincaré recurrence theorem. □

7. Entropy density of \mathcal{B} -free shifts

In this section we demonstrate the utility of our approach by showing that for every \mathcal{B} -free shift $X_{\mathcal{B}}$, its hereditary closure $\tilde{X}_{\mathcal{B}}$ has entropy-dense set of ergodic measures. Since it is quite common that $X_{\mathcal{B}} = \tilde{X}_{\mathcal{B}}$, we obtain many examples of entropy-dense \mathcal{B} -free shifts. In fact, we show more, namely that the invariant measures on $\tilde{X}_{\mathcal{B}}$ are also the space of shift invariant measures on a potentially larger shift space $\tilde{X}_{\mathcal{B}}^*$.

Throughout this section we work with the alphabet $\mathcal{A} = \{0, 1\}$. Recall that the hereditary closure of a shift space $X \subseteq \{0, 1\}^\infty$ is defined as

$$\tilde{X} = \{y \in \{0, 1\}^\infty \mid \text{there exists } x \in X, y \leq x\},$$

where the order \leq on $\{0, 1\}^\infty$ is defined coordinatewise, meaning that $y \leq x$ if $y_i \leq x_i$ for all $i \in \mathbb{N}_0$. A shift space is *hereditary* if it coincides with its own hereditary closure. Given a set $\mathcal{B} \subseteq \mathbb{N}$, we say that a positive integer number is \mathcal{B} -free if it is not a multiple of any member of \mathcal{B} . The set of all \mathcal{B} -free numbers is denoted by $\mathcal{F}_{\mathcal{B}} \subseteq \mathbb{N}_0$, that is,

$$\mathcal{F}_{\mathcal{B}} = \mathbb{N}_0 \setminus \bigcup_{b \in \mathcal{B}} b\mathbb{N}_0.$$

We say that \mathcal{B} is *primitive* if for each $b, b' \in \mathcal{B}$ with $b \neq b'$ we have that b does not divide b' . A set $\mathcal{B} \subseteq \mathbb{N}$ is *taut* if for every $b_0 \in \mathcal{B}$ we have

$$\underline{d}\left(\bigcup_{b \in \mathcal{B} \setminus \{b_0\}} b\mathbb{N}_0\right) < \underline{d}\left(\bigcup_{b \in \mathcal{B}} b\mathbb{N}_0\right).$$

The characteristic sequence of $\mathcal{F}_{\mathcal{B}}$ is denoted by $\eta_{\mathcal{B}} := 1_{\mathcal{F}_{\mathcal{B}}}$, and we think of it as an element of the full-shift $\{0, 1\}^\infty$. The orbit closure $X_{\mathcal{B}} \subseteq \{0, 1\}^\infty$ of $\eta_{\mathcal{B}}$ is called the *\mathcal{B} -free shift*. Given an enumeration of $\mathcal{B} = \{b_1, b_2, \dots\}$ with $b_i < b_j$ for $i < j$ and $k \in \mathbb{N}$, we let $\mathcal{B}|k = \{b_1, \dots, b_k\}$ denote the set of the k smallest elements of \mathcal{B} and

let $\mathcal{F}_{\mathcal{B}|k} \subseteq \mathbb{N}_0$ stand for the set $\mathcal{B}|k$ -free integers. Let us write $\eta_{\mathcal{B}|k} \in \{0, 1\}^\infty$ for the characteristic function of $\mathcal{F}_{\mathcal{B}|k}$. Note that $\eta_{\mathcal{B}|k}$ is a periodic point in $\{0, 1\}^\infty$, hence the $\mathcal{B}|k$ -free shift $X_{\mathcal{B}|k}$ is just the orbit of $\eta_{\mathcal{B}|k}$. We clearly have $\mathcal{F}_{\mathcal{B}} \subseteq \mathcal{F}_{\mathcal{B}|k}$, hence $\eta_{\mathcal{B}} \leq \eta_{\mathcal{B}|k}$ and $\tilde{X}_{\mathcal{B}} \subseteq \tilde{X}_{\mathcal{B}|k}$. It also turns out that for every $k \in \mathbb{N}$ the hereditary shifts $\tilde{X}_{\mathcal{B}|k}$ are transitive and sofic. This was first noticed in [19] (transitivity in Proposition 3.17 and soficity in §3.4.1). We provide an independent proof of a slightly more general fact.

PROPOSITION 26. *Let $x \in \{0, 1\}^\infty$ be such that $\sigma^n(x) = x$ for some $n \geq 1$. If $X = \{\sigma^j(x) : j = 0, 1, \dots, n - 1\}$ is the orbit of x , then its hereditary closure \tilde{X} is a transitive sofic shift. Furthermore, $\tilde{X} = \{y \in \{0, 1\}^\infty : y \leq \sigma^j(x) \text{ for some } 0 \leq j < n\}$. If $\sigma(x) \neq x$, then \tilde{X} is not mixing.*

Proof. It is enough to notice that \tilde{X} is presented by a labelled graph with vertices denoted v_0, v_1, \dots, v_{n-1} and edges and their labels defined as follows: for each $0 \leq j < n$ we put one edge from v_j to $v_{j+1 \bmod n}$ labelled with 0, and if $x_j = 1$ we add one more edge labelled with 0. □

A key fact about \mathcal{B} -free integers which allows our argument to work is that the periodic sets $\mathcal{F}_{\mathcal{B}|k}$ approximate $\mathcal{F}_{\mathcal{B}}$ with respect to the premetric \underline{d} . This is a consequence of a classical result of Davenport and Erdős.

THEOREM 27. (Davenport and Erdős; see [14, 15]) *Let $\mathcal{B} = \{b_1, b_2, \dots\} \subseteq \mathbb{N}$. Then*

$$\underline{d}(\mathcal{F}_{\mathcal{B}|k} \setminus \mathcal{F}_{\mathcal{B}}) \rightarrow 0 \quad \text{as } k \rightarrow \infty.$$

Observing that $\underline{d}(\mathcal{F}_{\mathcal{B}|k} \setminus \mathcal{F}_{\mathcal{B}}) = \underline{d}(\eta_{\mathcal{B}|k}, \eta_{\mathcal{B}})$, we see that the characteristic function $\eta_{\mathcal{B}}$ is the \underline{d} -limit of a sequence of periodic points $\eta_{\mathcal{B}|k}$. Furthermore, we can reformulate the Davenport–Erdős theorem in terms of the premetric \underline{d}^H .

COROLLARY 28. *Let $\mathcal{B} = \{b_1, b_2, \dots\} \subseteq \mathbb{N}$. Then*

$$\underline{d}^H(\tilde{X}_{\mathcal{B}|k}, \tilde{X}_{\mathcal{B}}) \rightarrow 0 \quad \text{as } k \rightarrow \infty.$$

Proof. Fix $k \in \mathbb{N}$. We claim that $\underline{d}^H(\tilde{X}_{\mathcal{B}|k}, \tilde{X}_{\mathcal{B}}) \leq \underline{d}(\eta_{\mathcal{B}|k}, \eta_{\mathcal{B}})$. Since $\tilde{X}_{\mathcal{B}} \subseteq \tilde{X}_{\mathcal{B}|k}$ it is enough to show that for every $x \in \tilde{X}_{\mathcal{B}|k}$ there is $y \in \tilde{X}_{\mathcal{B}}$ satisfying $\underline{d}(x, y) \leq \underline{d}(\eta_{\mathcal{B}|k}, \eta_{\mathcal{B}})$. To this end, take $x \in \tilde{X}_{\mathcal{B}|k}$. By Proposition 26 there exists $m \geq 0$ such that $\sigma^m(\eta_{\mathcal{B}|k}) \geq x$ coordinatewise. Consider $y \in \{0, 1\}^\infty$ defined by the formula

$$y_i = \begin{cases} x_i & \text{if } (\eta_{\mathcal{B}|k})_{i+m} = (\eta_{\mathcal{B}})_{i+m}, \\ 0 & \text{otherwise, i.e. if } (\eta_{\mathcal{B}|k})_{i+m} = 1 \neq 0 = (\eta_{\mathcal{B}})_{i+m}. \end{cases}$$

We immediately see that $y \leq \sigma^m(\eta_{\mathcal{B}})$ coordinatewise, hence $y \in \tilde{X}_{\mathcal{B}}$. It is also clear that $\underline{d}(x, y) \leq \underline{d}(\eta_{\mathcal{B}|k}, \eta_{\mathcal{B}}) = \underline{d}(\mathcal{F}_{\mathcal{B}|k} \setminus \mathcal{F}_{\mathcal{B}})$. We use the Davenport–Erdős theorem to conclude that $\underline{d}^H(\tilde{X}_{\mathcal{B}|k}, \tilde{X}_{\mathcal{B}}) \rightarrow 0$ as $k \rightarrow \infty$. □

Before we state our main result regarding \mathcal{B} -free shifts, let us first discuss some technical issues caused by the fact that the ‘ \underline{d}^H -approximation’ appearing in (5) and Corollary 28 does not uniquely determine its ‘limit’.

Remark 29. Observe that we can consider a whole spectrum of intermediate shift spaces associated to \mathcal{B} . Namely, for $k \in \mathbb{N}$ we put

$$\tilde{X}_{\mathcal{B}}^{(k)} = \bigcap_{\substack{\mathcal{B}' \subseteq \mathcal{B} \\ \#\mathcal{B}'=k}} \tilde{X}_{\mathcal{B}'}$$

These sets are again shift invariant and hereditary. Moreover, setting

$$\tilde{X}_{\mathcal{B}}^* = \bigcap_{k=1}^{\infty} \tilde{X}_{\mathcal{B}}^{(k)},$$

we have the sequence of inclusions

$$X_{\mathcal{B}} \subseteq \tilde{X}_{\mathcal{B}} \subseteq \tilde{X}_{\mathcal{B}}^* \subseteq \dots \subseteq \tilde{X}_{\mathcal{B}}^{(2)} \subseteq \tilde{X}_{\mathcal{B}}^{(1)}. \tag{15}$$

We will see later (see Remark 31) that the second inclusion in (15) may be strict. As far as we know, only the sets $X_{\mathcal{B}}$, $\tilde{X}_{\mathcal{B}}$ and $\tilde{X}_{\mathcal{B}}^{(1)}$ have appeared in the literature before. The shift space $\tilde{X}_{\mathcal{B}}^{(1)}$ is the largest shift space related to \mathcal{B} -free constructions. It is usually called the *\mathcal{B} -admissible shift* and its elements are *\mathcal{B} -admissible sequences* (see [19]). Let us also point out that the hierarchy introduced above is still rather rough: given any countable family \mathcal{C} of subsets of \mathcal{B} , the intersection $\bigcap_{\mathcal{B}' \in \mathcal{C}} \tilde{X}_{\mathcal{B}'}$ is a hereditary shift space that contains $X_{\mathcal{B}}$.

LEMMA 30. For $\mathcal{B} \subseteq \mathbb{N}$ we have

$$\tilde{X}_{\mathcal{B}}^* = \bigcap_{\substack{\mathcal{B}' \subseteq \mathcal{B} \\ \mathcal{B}' \text{ finite}}} \tilde{X}_{\mathcal{B}'} = \bigcap_{k=1}^{\infty} \tilde{X}_{\mathcal{B}|k}.$$

Proof. Fix $n \in \mathbb{N}$. We have

$$\bigcap_{\substack{\mathcal{B}' \subseteq \mathcal{B} \\ \mathcal{B}' \text{ finite}}} \tilde{X}_{\mathcal{B}'} \subseteq \tilde{X}_{\mathcal{B}}^* = \bigcap_{k=1}^{\infty} \tilde{X}_{\mathcal{B}}^{(k)} \subseteq \tilde{X}_{\mathcal{B}}^{(n)} = \bigcap_{\substack{\mathcal{B}' \subseteq \mathcal{B} \\ \#\mathcal{B}'=n}} \tilde{X}_{\mathcal{B}'} \subseteq \tilde{X}_{\mathcal{B}|n}.$$

It follows that

$$\bigcap_{\substack{\mathcal{B}' \subseteq \mathcal{B} \\ \mathcal{B}' \text{ finite}}} \tilde{X}_{\mathcal{B}'} \subseteq \tilde{X}_{\mathcal{B}}^* \subseteq \bigcap_{k=1}^{\infty} \tilde{X}_{\mathcal{B}|k}.$$

Assume that $x \in \tilde{X}_{\mathcal{B}|k}$ for every $k \in \mathbb{N}$. Take any finite set $\mathcal{B}' \subseteq \mathcal{B}$. Then there exists $n \in \mathbb{N}$ such that $\mathcal{B}' \subseteq (\mathcal{B} | n)$, hence $\tilde{X}_{\mathcal{B}|n} \subseteq \tilde{X}_{\mathcal{B}'}$. We conclude that $x \in \tilde{X}_{\mathcal{B}'}$ for every finite set $\mathcal{B}' \subseteq \mathcal{B}$. Therefore

$$\bigcap_{k=1}^{\infty} \tilde{X}_{\mathcal{B}|k} \subseteq \bigcap_{\substack{\mathcal{B}' \subseteq \mathcal{B} \\ \mathcal{B}' \text{ finite}}} \tilde{X}_{\mathcal{B}'},$$

which finishes the proof. □

Remark 31. The entire hierarchy (15) collapses to a single shift space ($X_{\mathcal{B}} = \tilde{X}_{\mathcal{B}}^{(1)}$) whenever \mathcal{B} is a taut set containing an infinite set of pairwise coprime integers (see [19, Theorem B] and [29, Corollary 2]; cf. the proof of Corollary 35 below). However, in general, each inclusion in (15) can be strict. For the example showing that the first inclusion may be strict see [19]. We show that $\tilde{X}_{\mathcal{B}}^*$ can differ from the \mathcal{B} -admissible shift $\tilde{X}_{\mathcal{B}}^{(1)}$. Indeed, if $\mathcal{B} = \{4, 6\}$, then $\tilde{X}_{\mathcal{B}}^*$ and $\tilde{X}_{\mathcal{B}}^{(1)}$ differ. More generally, if $\mathcal{B} = \{qp_1, qp_2, \dots, qp_k\}$ where $q \geq k$ and p_1, \dots, p_k are distinct primes then $\tilde{X}_{\mathcal{B}}^* = \tilde{X}_{\mathcal{B}}^{(k)}$ and $\tilde{X}_{\mathcal{B}}^{(k-1)}$ differ.

The following example shows that $\tilde{X}_{\mathcal{B}}$ and $\tilde{X}_{\mathcal{B}}^*$ can be different as well.

Example 32. Let p be a large prime number (e.g. $p = 107$), and put $l = (p - 2)(p - 1)$, and $\mathcal{B} = \{p - 2, p - 1, p\} \cup \{n \geq l : p \nmid n\}$. Consider the characteristic sequence $x = 1_A \in \{0, 1\}^\infty$ of the set $A \subseteq \mathbb{N}_0$ given by

$$A = \{j \in \mathbb{N}_0 : 0 \leq j < l\} \setminus ((p - 2)\mathbb{N} \cup (p - 1)\mathbb{N} \cup (p\mathbb{N} - 1)).$$

Then $x \in \tilde{X}_{\mathcal{B}}^*$, because for any finite $\mathcal{B}' \subseteq \mathcal{B}$ we can find $n \in \mathbb{N}$ with $n \equiv 1 \pmod p$ and $n \equiv 0 \pmod b$ for all $b \in \mathcal{B}' \setminus \{p\}$.

On the other hand, we claim that $x \notin \tilde{X}_{\mathcal{B}}$. For the sake of contradiction, suppose that $x \in \tilde{X}_{\mathcal{B}}$. Since $\tilde{X}_{\mathcal{B}}$ is finite, there exists $m \in \mathbb{N}_0$ such that $x \leq \sigma^m(\eta_{\mathcal{B}})$ coordinatewise. Since $n \notin \mathcal{F}_{\mathcal{B}}$ for $n \geq l$ and $l - 1 \in A$, it follows that $m = 0$. However, $p \in A$ while $p \notin \mathcal{F}_{\mathcal{B}}$ so $m \neq 0$, which is the sought contradiction.

THEOREM 33. *Let $\mathcal{B} \subseteq \mathbb{N}$. Then ergodic measures are entropy-dense for $\tilde{X}_{\mathcal{B}}$ and*

$$\mathcal{M}_\sigma(\tilde{X}_{\mathcal{B}}) = \mathcal{M}_\sigma(\tilde{X}_{\mathcal{B}}^*) = \bigcap_{k=1}^\infty \mathcal{M}_\sigma(\tilde{X}_{\mathcal{B}|k}).$$

Proof. The second equality is a consequence of Proposition 2 and Lemma 30. Fix $k \in \mathbb{N}$. We have

$$\tilde{X}_{\mathcal{B}} \subseteq \tilde{X}_{\mathcal{B}}^* \subseteq \tilde{X}_{\mathcal{B}|k}.$$

It follows that we have

$$\underline{d}(\tilde{X}_{\mathcal{B}|k}, \tilde{X}_{\mathcal{B}}^*) \leq \underline{d}(\tilde{X}_{\mathcal{B}|k}, \tilde{X}_{\mathcal{B}}).$$

In particular, by Corollary 28 all ‘distances’ above tend to 0 as $k \rightarrow \infty$ and $\mathcal{M}_\sigma(\tilde{X}_{\mathcal{B}}) = \mathcal{M}_\sigma(\tilde{X}_{\mathcal{B}}^*)$ by Corollary 18. We finish the proof applying Corollary 20. \square

Example 34. Note that for $\tilde{X}_{\mathcal{B}}$ we have only a d^H -approximation by transitive sofic shifts, which are not mixing. This raises the question whether $X_{\mathcal{B}}$ or $\tilde{X}_{\mathcal{B}}$ can be \bar{d} -approachable. Note that the shift space in Example 11 can be equivalently defined as $\tilde{X}_{\mathcal{B}}$, where $\mathcal{B} = \{2\}$, showing that hereditary closures of some \mathcal{B} -free shifts are non- \bar{d} -approachable. We will extend this example and prove that there exists an infinite set \mathcal{B} such that $\tilde{X}_{\mathcal{B}}$ is not

\bar{d} -approachable. Let $b_1 = 2$ and pick b_2, b_3, \dots such that

$$\sum_{k=2}^{\infty} \frac{1}{b_k} < \frac{1}{32}. \tag{16}$$

Note that (16) ensures that $\bar{d}^H(\tilde{X}_{\{2\}}, \tilde{X}_{\mathcal{B}}) < 1/32$. For n large enough, there is a word $u \in \mathcal{B}(\tilde{X}_{\mathcal{B}})$ with $|u| = 2n$ such that $d_{\text{Ham}}((01)^n, u) < 1/4$. In particular, at least half of the even entries in u are occupied by 1. In other words,

$$|\{1 \leq j \leq n : u_{2j} = 1\}| \geq \frac{1}{2}n.$$

By heredity of $\tilde{X}_{\mathcal{B}}$, the periodic point $y = (u0^{2n-1})^\infty$ belongs to $(\tilde{X}_{\mathcal{B}})_{2n}^M$. Reasoning as in Example 11, we obtain $\bar{d}(y, x) \geq 1/16$ for every $x \in \tilde{X}_{\{2\}}$. This implies that $\bar{d}^H((\tilde{X}_{\mathcal{B}})_{2n}^M, \tilde{X}_{\{2\}}) \geq 1/16$, yielding

$$\bar{d}^H((\tilde{X}_{\mathcal{B}})_{2n}^M, \tilde{X}_{\mathcal{B}}) > 1/32,$$

for n large enough. So the Markov approximations are \bar{d} -far from $\tilde{X}_{\mathcal{B}}$. Note that we can choose \mathcal{B} that consists of pairwise prime numbers and still satisfies (16). For such a set \mathcal{B} we have $\tilde{X}_{\mathcal{B}} = X_{\mathcal{B}}$ and the \mathcal{B} -free shift $X_{\mathcal{B}}$ is not \bar{d} -approachable.

So far we have discussed only the hereditary closure $\tilde{X}_{\mathcal{B}}$ of $X_{\mathcal{B}}$. In general, $X_{\mathcal{B}}$ does not need to be hereditary, as shown by the example $\mathcal{B} = \mathbb{N}_{\geq 3}$ where $X_{\mathcal{B}}$ is the (finite) orbit of the point 011000 . . . and hence does not contain the point 01000 . . . Nevertheless, for many examples of the sets $\mathcal{F}_{\mathcal{B}}$ studied in the literature the associated \mathcal{B} -free shift turns out to be hereditary (e.g. the square-free shift or shifts associated with abundant numbers; see [19]). For the record, we note a version of Theorem 33 for these shifts. But first, note that every hereditary shift contains the fixed point 0^∞ , and since every \mathcal{B} -free shift has a unique minimal subsystem (by Theorem A in [19]), the set $\{0^\infty\}$ must be the unique minimal subset for a hereditary \mathcal{B} -free shift. In other words, a hereditary \mathcal{B} -free shift must be proximal. For \mathcal{B} -free shifts proximality is equivalent to the fact that the set \mathcal{B} contains an infinite pairwise coprime subset. Therefore, hoping for $X_{\mathcal{B}} = \tilde{X}_{\mathcal{B}}$, we have to assume the latter condition.

COROLLARY 35. *If \mathcal{B} contains an infinite sequence of pairwise coprime integers, then ergodic measures are entropy-dense in $\mathcal{M}_\sigma(X_{\mathcal{B}})$.*

Proof. By Theorem C in [19], for every $\mathcal{B} \subseteq \mathbb{N}$ there is a unique taut $\mathcal{B}' \subseteq \mathbb{N}$ such that $\tilde{X}_{\mathcal{B}'} \subseteq \tilde{X}_{\mathcal{B}}$ and $\mathcal{M}_\sigma(\tilde{X}_{\mathcal{B}'}) = \mathcal{M}_\sigma(\tilde{X}_{\mathcal{B}})$. Therefore we may assume that \mathcal{B} is taut and proximal. Then we use a recent result of Keller [29, Corollary 2], who complemented some results from [19] by showing that if \mathcal{B} is taut and $X_{\mathcal{B}}$ is proximal, then $X_{\mathcal{B}}$ is hereditary, that is, $X_{\mathcal{B}} = \tilde{X}_{\mathcal{B}}$. We finish the proof by applying Theorem 33. □

Remark 36. Note that $\mathcal{M}_\sigma(X_{\mathcal{B}})$ may be trivial, that is, its unique element may be the Dirac measure concentrated on the fixed point 0^∞ . This is the case when \mathcal{B} is Behrend (see [19, p. 5437]).

Remark 37. Note that Theorem 33 remains true if we replace $\tilde{X}_{\mathcal{B}}$ by any shift space which is the hereditary closure of the orbit closure of a characteristic function of a set $A \subseteq \mathbb{N}_0$ such that there exists a sequence of periodic sets $(A_k)_{k=1}^\infty$ satisfying

$$A \subseteq A_k \quad \text{for every } k \in \mathbb{N},$$

$$\underline{d}(A_k, A) \rightarrow 0 \quad (k \rightarrow \infty).$$

Such shift spaces have been already considered in the literature [34, 39]. A closely related notion of a *rational set* (a subset of \mathbb{N} whose characteristic function can be \bar{d} -approximated by periodic characteristic functions) has also attracted attention; see [4, 5, 16].

8. *Inner \bar{d} -approachability*

So far we have considered shift spaces, which are approximated from the outside, that is, approximating subshifts are not contained in the approximated shift space. Inspired by an idea of approximation given by Dan Thompson in an unpublished manuscript [56], we will show that every topologically mixing S -gap shift is approximated in the \bar{d}^H sense by a sequence of shift spaces of finite type contained in X_S . We call this property *inner \bar{d} -approachability*.

We say that a shift space $X \subseteq \mathcal{A}^\infty$ is *inner \bar{d} -approachable* if it contains a sequence $(X_n)_{n=1}^\infty$ of mixing shifts of finite type such that

$$\overline{\bigcup_{n=1}^\infty X_n} = X \quad \text{and} \quad \lim_{n \rightarrow \infty} \bar{d}^H(X_n, X) = 0.$$

Given $S \subseteq \mathbb{N}$, we write $S = \{n_1, n_2, \dots\}$ with $n_i < n_{i+1}$ for $i < |S|$, where $|S|$ stands for the cardinality of S ($|S| = \infty$ if S is infinite). The S -gap shift X_S is a shift space over $\{0, 1\}$ consisting of all sequences such that the number of 0s between any two successive occurrences of the symbol 1 belongs to S . Equivalently, $\{10^n 1 : n \notin S\}$ is the collection of forbidden sequences for X_S . By [26, Example 3.4], X_S is topologically mixing if and only if $\gcd\{n + 1 : n \in S\} = 1$, and topologically mixing X_S has the specification property if and only if $\sup_i |n_i - n_{i+1}| < \infty$. Note that all sequences in X_S are labels of infinite paths in the labelled directed graph $G_S = (V, E_S)$, where $V = \{v_j : 0 \leq j \leq |S|\}$ is the set of vertices and there is an edge $v_i \rightarrow v_j$ in E_S if and only if $j = i + 1$ or $i \in S$ and $j = 0$. We label each edge $v_i \rightarrow v_{i+1}$ with 0 and all other edges with 1.

PROPOSITION 38. *Every mixing S -gap shift is inner \bar{d} -approachable.*

Proof. Assume that $S = \{s_1, s_2, \dots\}$ is infinite, $s_i < s_{i+1}$ for every $i \in \mathbb{N}$, and $\gcd\{n + 1 : n \in S\} = 1$ (if S is finite, then X_S is a shift of finite type). Let N be such that $\gcd\{s_1 + 1, \dots, s_N + 1\} = 1$. Let $L > 0$ be the smallest integer such that for every $\ell \geq L$ there are non-negative integers $\alpha_1, \dots, \alpha_N$ such that

$$\alpha_1(s_1 + 1) + \dots + \alpha_N(s_N + 1) = \ell.$$

It follows that for every $\ell \geq L$ there are $r \in \mathbb{N}$ and $q_1, \dots, q_r \in \{s_1, \dots, s_N\}$ such that the word

$$w = 0^{q_1} 10^{q_2} 1 \dots 0^{q_r} 1 \tag{17}$$

satisfies $w \in \mathcal{B}(X_S)$ and $|w| = \ell$. Set $S[n] = \{s_1, \dots, s_n\}$ and let X_n be the S -gap shift generated by $S[n]$. Clearly $X_1 \subseteq X_2 \subseteq X_3 \subseteq \dots$ and

$$\bigcup_{n=1}^{\infty} X_n = X_S.$$

Furthermore, for every $n \geq N$ the shift space X_n is a mixing shift of finite type. Fix $\varepsilon > 0$. Let $M \geq N$ be such that $1/s_M < \varepsilon/2$. Let $K \geq M$ satisfy $(L + s_M)/s_K < \varepsilon/2$. We claim that for every $k \geq K$ we have

$$\bar{d}^H(X_k, X) \leq \varepsilon.$$

Note that our proof is finished once we show that the claim holds. For a proof of our claim fix $k \geq K$. Let $x \in X$. Since \bar{d} is σ -invariant we assume that

$$x = 10^{t_1} 10^{t_2} 1 \dots 10^{t_j} 1 \dots,$$

that is, we assume that x begins with 1 and contains infinitely many 1s. The proof is similar, if the symbol 1 occurs in x only finitely many times. Note that $x \notin X_k$ means that for some $j \in \mathbb{N}$ we have $t_j \in S \setminus \{s_1, \dots, s_k\}$.

We will construct, for each j such that $t_j \in S \setminus \{s_1, \dots, s_k\}$, a word v_j with $|v_j| = t_j$ such that $d_{\text{Ham}}(v_j, 0^{t_j}) < \varepsilon$ and $1v_j1 \in \mathcal{B}(X_S[k])$. Put also $v_j = 0^{t_j}$ if $t_j \in \{s_1, \dots, s_k\}$. Once this is accomplished, we set $x' = 1v_11v_21 \dots$. It is now enough to observe that $x' \in X_{S[k]}$ and $\bar{d}(x, x') \leq \sup_j d_{\text{Ham}}(v_j, 0^{t_j})$.

Fix j such that $t_j \in S \setminus \{s_1, \dots, s_k\}$. Let ℓ be the unique integer such that $L \leq \ell < L + s_M$ and $t_j + 1 - \ell$ is divisible by $(s_M + 1)$. We replace the suffix $0^{\ell-1}1$ of $10^{t_j}1$ by a word w with $|w| = \ell$ having the form as in (17). The remaining prefix of $10^{t_j}1$ now has the form

$$10^{t_j+1-\ell}$$

where $t_j + 1 - \ell = p(s_M + 1)$ for some $p \geq 1$. Therefore we may replace $10^{t_j+1-\ell}$ by

$$(0^{s_M} 1)^p.$$

We have found a word v with $|v| = t_j$ such that

$$u = 1v1 = 1(0^{s_M} 1)^p 0^{q_1} 10^{q_2} 1 \dots 0^{q_r} 1 \in \mathcal{B}(X_S[k]).$$

Note that

$$d_{\text{Ham}}(0^{t_j}, v) \leq \frac{1}{t_j}(t_j/(s_M + 1) + (L + s_M)) < \varepsilon,$$

which completes the proof. □

Entropy density for mixing S -gap shifts follows from Theorems B and 3.5 in [11]. Using Proposition 38 and Corollary 16, we obtain a new proof of this fact.

COROLLARY 39. *Every mixing S -gap shift has entropy-dense set of ergodic measures.*

Acknowledgements. We thank the referees for their positive comments and corrections that helped us to improve this paper. We are grateful to Dan Thompson for sharing his unpublished manuscript [56]. We would like to thank Aurelia Dymek for reading the preprint and sharing with us her helpful and insightful remarks. D. Kwietniak was supported by the National Science Centre (NCN) Opus grant no. 2018/29/B/ST1/01340. J. Konieczny is working within the framework of the 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).

REFERENCES

- [1] M. Avdeeva. Limit theorems for B -free integers and the Moebius function. *PhD Thesis*, Princeton University, 2016.
- [2] M. Avdeeva, F. Cellarosi and Y. G. Sinai. Ergodic and statistical properties of B -free numbers. *Teor. Veroyatn. Primen.* **61**(4) (2016), 805–829; Engl. Transl. *Theory Probab. Appl.* **61**(4) (2017), 569–589.
- [3] M. Baake and C. Huck. Ergodic properties of visible lattice points. *Proc. Steklov Inst. Math.* **288**(1) (2015), 165–188.
- [4] 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.
- [5] V. Bergelson and I. Ruzsa. Squarefree numbers, IP sets and ergodic theory. *Paul Erdős and His Mathematics, I (Budapest, 1999) (Bolyai Society Mathematical Studies, 11)*. Eds. G. Halasz, L. Lovasz, M. Simonovits and V. T. Sós. Janos Bolyai Mathematical Society, Budapest, 2002, pp. 147–160.
- [6] M. L. Blank. Metric properties of ε -trajectories of dynamical systems with stochastic behaviour. *Ergod. Th. & Dynam. Sys.* **8**(3) (1988), 365–378.
- [7] F. Cellarosi and Y. G. Sinai. Ergodic properties of square-free numbers. *J. Eur. Math. Soc. (JEMS)* **15** (2013), 1343–1374.
- [8] F. Cellarosi and I. Vinogradov. Ergodic properties of k -free integers in number fields. *J. Mod. Dyn.* **7**(3) (2013), 461–488.
- [9] J.-R. Chazottes, L. Ramirez and E. Ugalde. Finite type approximations of Gibbs measures on sofic subshifts. *Nonlinearity* **18**(1) (2005), 445–463.
- [10] V. Climenhaga and D. Thompson. Intrinsic ergodicity beyond specification: β -shifts, S -gap shifts, and their factors. *Israel J. Math.* **192**(2) (2012), 785–817.
- [11] V. Climenhaga, D. Thompson and K. Yamamoto. Large deviations for systems with non-uniform structure. *Trans. Amer. Math. Soc.* **369**(6) (2017), 4167–4192.
- [12] H. Comman. Strengthened large deviations for rational maps and full shifts, with unified proof. *Nonlinearity* **22**(6) (2009), 1413–1429.
- [13] 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.
- [14] H. Davenport and P. Erdős. On sequences of positive integers. *Acta Arith.* **2** (1936), 147–151.
- [15] H. Davenport and P. Erdős. On sequences of positive integers. *J. Indian Math. Soc. (N.S.)* **15** (1951), 19–24.
- [16] K. Deka. Some properties of regular and rational sets. *Acta Arith.* **182**(3) (2018), 279–284.
- [17] M. Denker, C. Grillenberger and K. Sigmund. *Ergodic Theory on Compact Spaces (Lecture Notes in Mathematics, 527)*. Springer-Verlag, Berlin, 1976.
- [18] A. Dymek. Proximity of multidimensional B -free systems. *Discrete Contin. Dyn. Syst.* **41**(8) (2021), 3709–3724.
- [19] A. Dymek, S. Kasjan, J. Kułaga-Przymus and M. Lemańczyk. B -free sets and dynamics. *Trans. Amer. Math. Soc.* **370**(8) (2018), 5425–5489.
- [20] A. Eizenberg, Y. Kifer and B. Weiss. Large deviations for \mathbb{Z}^d -actions. *Comm. Math. Phys.* **164**(3) (1994), 433–454.
- [21] E. H. El Abdalaoui, M. Lemańczyk and T. de la Rue. A dynamical point of view on the set of B -free integers. *Int. Math. Res. Not. IMRN* **2015**(16) (2015), 7258–7286.

- [22] H. Föllmer and S. Orey. Large deviations for the empirical field of a Gibbs measure. *Ann. Probab.* **16**(3) (1988), 961–977.
- [23] N. A. Friedman and D. S. Ornstein. On isomorphism of weak Bernoulli transformations. *Adv. Math.* **5** (1970), pp. 365–394.
- [24] K. Gelfert and D. Kwietniak. On density of ergodic measures and generic points. *Ergod. Th. & Dynam. Sys.* **38**(5) (2018), 1745–1767.
- [25] 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.
- [26] U. Jung. On the existence of open and bi-continuing codes. *Trans. Amer. Math. Soc.* **363** (2011), 1399–1417.
- [27] S. Kasjan, G. Keller and M. Lemańczyk. Dynamics of \mathcal{B} -free sets: a view through the window. *Int. Math. Res. Not. IMRN* **2019**(9) (2019), 2690–2734.
- [28] G. Keller. Tautness for sets of multiples and applications to \mathcal{B} -free dynamics. *Studia Math.* **247**(2) (2019), 205–216.
- [29] G. Keller. Generalized heredity in \mathcal{B} -free systems. *Stoch. Dyn.* **21**(3) (2021), Paper no. 2140008.
- [30] J. Konieczny, M. Kupsa and D. Kwietniak. Minimal and proximal examples of \bar{d} -stable and \bar{d} -approachable shift spaces, in preparation.
- [31] J. Kułaga-Przymus, M. Lemańczyk and B. Weiss. On invariant measures for \mathcal{B} -free systems. *Proc. Lond. Math. Soc. (3)* **110** (2015), 1435–1474.
- [32] 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)*. Eds. J. Auslander, A. Johnson and C. E. Silva. American Mathematical Society, Providence, RI, 2016, pp. 245–253.
- [33] 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.
- [34] M. Kulczycki, D. Kwietniak and J. Li. Entropy of subordinate shift spaces. *Amer. Math. Monthly* **125**(2) (2018), 141–148.
- [35] M. Kulczycki, D. Kwietniak and P. Oprocha. On almost specification and average shadowing properties. *Fund. Math.* **224**(3) (2014), 241–278.
- [36] P. Kůrka. *Topological and Symbolic Dynamics (Cours Spécialisés [Specialized Courses], 11)*. Société Mathématique de France, Paris, 2003.
- [37] 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.
- [38] D. Kwietniak, M. Łącka and P. Oprocha. A panorama of specification-like properties and their consequences. *Dynamics and Numbers (Contemporary Mathematics, 669)*. Eds. S. Kolyada, M. Möller, P. Moree and Th. Ward. American Mathematical Society, Providence, RI, 2016, pp. 155–186.
- [39] D. Kwietniak, M. Łącka and P. Oprocha. Generic points for dynamical systems with average shadowing. *Monatsh. Math.* **183**(4) (2017), 625–648.
- [40] D. Kwietniak, P. Oprocha and M. Rams. On entropy of dynamical systems with almost specification. *Israel J. Math.* **213**(1) (2016), 475–503.
- [41] J. Li and P. Oprocha. Properties of invariant measures in dynamical systems with the shadowing property. *Ergod. Th. & Dynam. Sys.* **38**(6) (2018), 2257–2294.
- [42] D. Lind and B. Marcus. *An Introduction to Symbolic Dynamics and Coding*. Cambridge University Press, Cambridge, 1995.
- [43] S. Orey. Large deviations in ergodic theory. *Seminar on Stochastic Processes, 1984 (Evanston, Illinois, 1984) (Progress in Probability and Statistics, 9)*. Eds. E. Çinlar, K. L. Chung and R. K. Gettoor. Birkhäuser Boston, Boston, MA, 1986, pp. 195–249.
- [44] D. S. Ornstein. *Ergodic Theory, Randomness, and Dynamical Systems (Yale Mathematical Monographs, 5)*. Yale University Press, New Haven, CT, 1974.
- [45] K. R. Parthasarathy. On the category of ergodic measures. *Illinois J. Math.* **5** (1961), 648–656.
- [46] R. Pavlov. On intrinsic ergodicity and weakenings of the specification property. *Adv. Math.* **295** (2016), 250–270.
- [47] R. Peckner. Uniqueness of the measure of maximal entropy for the squarefree flow. *Israel J. Math.* **210**(1) (2015), 335–357.
- [48] C.-E. Pfister and W. G. Sullivan. Large deviations estimates for dynamical systems without the specification property. Applications to the β -shifts. *Nonlinearity* **18**(1) (2005), 237–261.
- [49] C.-E. Pfister and W. G. Sullivan. On the topological entropy of saturated sets. *Ergod. Th. & Dynam. Sys.* **27**(3) (2007), 929–956.
- [50] C.-E. Pfister and W. G. Sullivan. Weak Gibbs measures and large deviations. *Nonlinearity* **31**(1) (2018), 49–53.

- [51] D. Rudolph. *Fundamentals of Measurable Dynamics: Ergodic Theory on Lebesgue Spaces*. Clarendon Press, Oxford, 1990.
- [52] P. Sarnak. Three lectures on the Möbius function randomness and dynamics (Lecture 1). [http://publications.ias.edu/sites/default/files/MobiusFunctionsLectures\(2\).pdf](http://publications.ias.edu/sites/default/files/MobiusFunctionsLectures(2).pdf).
- [53] J. Schmeling. Symbolic dynamics for β -shifts and self-normal numbers. *Ergod. Th. & Dynam. Sys.* **17**(3) (1997), 675–694.
- [54] K. Sigmund. Generic properties of invariant measures for Axiom A-diffeomorphisms. *Invent. Math.* **11** (1970), 99–109.
- [55] D. Thompson. Irregular sets, the β -transformation and the almost specification property. *Trans. Amer. Math. Soc.* **364**(10) (2012), 5395–5414.
- [56] D. Thompson. A ‘horseshoe’ theorem in symbolic dynamics via single sequence techniques, unpublished manuscript, 2017.
- [57] J. Ville. Étude critique de la notion de collectif. *NUMDAM*, 1939, 116pp (in French).
- [58] X. Wu, P. Oprocha and G. Chen. On various definitions of shadowing with average error in tracing. *Nonlinearity* **29**(7) (2016), 1942–1972.