



# Folding and Metric Entropies for Extended Shifts

Neemias Martins<sup>1</sup> · Pedro G. Mattos<sup>2</sup> · Régis Varão<sup>2</sup>

Received: 11 August 2025 / Revised: 21 November 2025 / Accepted: 21 December 2025  
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2026

## Abstract

In this paper we calculate the metric and folding entropies for a family of non-invertible symbolic dynamical systems  $(\Sigma_{m_-, m_+}, \sigma_\phi)$  which generalizes the standard bilateral Bernoulli shifts. The space  $\Sigma_{m_-, m_+}$  consists of symbolic sequences over two distinct finite alphabets, with dynamics governed by a shift map  $\sigma_\phi$  incorporating a non-invertible function  $\phi$  that maps one of the alphabets to the other one. These systems are, for instance, particularly useful for encoding the many-to-one baker's transformation endomorphisms, and they can also be seen as a skew product with a unilateral Bernoulli shift on the base.

**Keywords** Metric entropy · Folding entropy · Extended shifts · Disintegration

**Mathematics Subject Classification** Primary: 37A35; Secondary: 37B10

## Contents

1	Introduction	.....
1.1	Classical Theory of Bernoulli Shifts	.....
1.2	The Case of Endomorphisms	.....
1.3	Shifts and Extended Shifts	.....
1.4	Main Results	.....
2	Preliminaries	.....
2.1	Measure Spaces	.....
2.2	Metric Entropy	.....
2.3	Disintegration of Measure	.....
2.4	Conditional and Folding Entropies	.....
3	Extended Shifts	.....
3.1	Measurable Structure	.....
3.2	Measure Structure	.....
4	Metric Entropy of Extended Shifts	.....

---

✉ Régis Varão  
varao@unicamp.br

Neemias Martins  
neemias.martins@ufv.br

Pedro G. Mattos  
pedrogmattos@ime.unicamp.br

<sup>1</sup> Department of Mathematics, Universidade Federal de Viçosa, Viçosa, MG 36570-000, Brazil

<sup>2</sup> Department of Mathematics, IMECC, Unicamp, Campinas, SP 13083-970, Brazil

4.1	Proof Sketch of Theorem A	.....
4.2	Partitions by Cylinders	.....
4.3	Technical Lemmas for Theorem A	.....
4.4	Proof of Theorem A	.....
5	Folding Entropy of Extended Shifts	.....
5.1	Proof Sketch of Theorem B	.....
5.2	The Quotient Measure	.....
5.3	Disintegration and Calculating the Folding Entropy	.....
5.4	Proof of Theorem B	.....
	References	.....

## 1 Introduction

### 1.1 Classical Theory of Bernoulli Shifts

Bernoulli shifts play a central role in the study of dynamical systems, serving as a prototype for understanding chaotic and stochastic behavior. Since Hadamard's study of the geodesic flow on negatively curved manifolds [6], their significance is underscored by their connections to other fundamental results in the field.

For instance, Anosov flows and geodesic flows on manifolds with negative curvature exhibit strong mixing properties and are often modeled using Bernoulli shifts due to their similar chaotic characteristics [1, 11, 18, 20]. Similarly, hyperbolic toral automorphisms, such as Arnold's cat map, share ergodicity and mixing properties with Bernoulli shifts, making them discrete-time analogues in the study of hyperbolic dynamics [9, 23]. More generally, Bowen used symbolic dynamics to study Axiom A diffeomorphisms [3].

The profound impact of Bernoulli shifts is further highlighted by Ornstein's classification theorem [17], which established that metric entropy, as defined by Kolmogorov and Sinai [10, 24], is a complete invariant for isomorphisms of bilateral Bernoulli shifts, a result that has influenced the classification of a wide range of dynamical systems. These connections demonstrate the unifying role of Bernoulli shifts in linking geometric, topological, and measure-theoretic aspects of dynamics.

### 1.2 The Case of Endomorphisms

Although the classification of invertible dynamical systems, such as Bernoulli shifts, has seen significant progress — most notably through Ornstein's isomorphism theorem — the classification of non-invertible maps (endomorphisms) presents a far more complex and nuanced challenge. Unlike invertible systems, where entropy serves as a complete invariant for Bernoulli shifts, non-invertible systems lack a similarly comprehensive classification framework. Entropy alone is insufficient to classify one-sided Bernoulli shifts [5, Section 4.5].

An endomorphism has a Bernoulli extension if its natural extension in the limit space is isomorphic to a Bernoulli shift. However, two non-isomorphic endomorphisms can have isomorphic natural extensions. Some characterizations and applications of non-invertible Bernoulli transformations are presented by Ashley, Marcus, and Tuncel [2], Bruin and Hawkins [4], and Hoffman and Rudolph [7, 8].

Anosov endomorphisms are also a rich family of examples to once more illustrate how non-invertibility provides new behavior if compared to their invertible analogue. It is well known that Anosov diffeomorphisms are structurally stable, but this fails for Anosov endo-

morphisms, even though from the inverse limit perspective they are structurally stable. In fact, not even a linear Anosov endomorphism is structurally stable [19].

The mathematical physicist David Ruelle introduced the folding entropy as a tool to measure the complexity of the preimage branches of an endomorphism [22]. Ruelle was interested in investigating entropy production in non-equilibrium statistical mechanics. Considering any state  $\mu$  as a limit of a sequence of absolutely continuous measures with probability densities, Ruelle expressed the entropy production by a limiting process from which the folding entropy emerged as a conditional entropy measuring the complexities of preimages of the system in terms of the state  $\mu$ . Some recent applications of the folding entropy are given by Liao and Wang [14], and Wu and Zhu [26, 27]. For non-invertible endomorphisms like  $\sigma_\phi$  (where  $m_+ > m_-$ ), as will be presented in the next subsection, folding entropy quantifies the ‘branching complexity’ of preimages — unlike metric entropy, it vanishes for invertible systems (e.g., standard bilateral shifts with  $m_+ = m_-$ ), making it a natural measure of non-invertibility.

As we will properly present in the next subsection, our work deals with a certain case of endomorphism for which our two main results, Theorems A and B, are the calculation of the metric and folding entropy, respectively. Our system could also be seen as a skew product with a one-sided Bernoulli system on the base, but we do not want to take this approach and we deliberately see it as an extended shift, as it will become clear soon, since we work with two alphabets.

### 1.3 Shifts and Extended Shifts

Before we present our systems, we establish some notation for standard unilateral and bilateral shifts. The symbolic space in  $m \in \mathbb{Z}_{>0}$  symbols can be either  $\Sigma_m^+ := \{0, \dots, m - 1\}^{\mathbb{Z}_{\geq 0}}$ , the space of all *unilateral sequences*  $x = (x_0, x_1, \dots)$ , or  $\Sigma_m := \{0, \dots, m - 1\}^{\mathbb{Z}}$ , the space of all *bilateral sequences*  $x = (\dots, x_{-1}; x_0, x_1, \dots)$ . In both cases, the entries of the sequences are elements of the alphabet  $\{0, \dots, m - 1\}$  (we will only consider finite alphabets). We also denote  $\Sigma_m^- := \{0, \dots, m - 1\}^{\mathbb{Z}_{< 0}}$ , the space of unilateral sequences  $(\dots, x_{-2}, x_{-1})$  on  $m$  symbols indexed by the negative integers (this is not standard notation but we will use this ahead). For unilateral sequences, the shift map  $\sigma : \Sigma_m^+ \rightarrow \Sigma_m^+$  is defined by  $\sigma(x) := (x_1, x_2, \dots)$ , whereas for bilateral sequences we define  $\sigma : \Sigma_m \rightarrow \Sigma_m$  by  $\sigma(x) := (\dots, x_0; x_1, x_2, \dots)$ . This map captures the behavior of the orbits of points in the system being encoded.

We will define an extension of bilateral Bernoulli shifts which can also be interpreted as a skew product of unilateral shifts with finite alphabets. Given two strictly positive integers  $m_-, m_+ \in \mathbb{Z}_{>0}$ , we define the product space  $\Sigma_{m_-, m_+} := \Sigma_{m_-}^- \times \Sigma_{m_+}^+$ . The elements of  $\Sigma_{m_-, m_+}$  are pairs  $x = (x^-, x^+)$ , with  $x^- = (\dots, x_{-2}, x_{-1})$  and  $x^+ = (x_0, x_1, \dots)$ , which can be seen as bilateral sequences

$$x = (\dots, x_{-1}; x_0, x_1, \dots).$$

It is important to notice that the symbols indexed by strictly negative numbers (entries of  $x^- \in \Sigma_{m_-}^-$ ) are elements of the *backward alphabet*  $\{0, \dots, m_- - 1\}$ , while the symbols indexed by positive numbers are elements of the *forward alphabet*  $\{0, \dots, m_+ - 1\}$ . This means that, when  $m_- < m_+$ , we cannot define the usual shift map on these sequences. To overcome this problem, we use a surjective function  $\phi : \{0, \dots, m_+ - 1\} \rightarrow \{0, \dots, m_- - 1\}$  and define an alternative ‘‘shift’’ map  $\sigma_\phi : \Sigma_{m_-, m_+} \rightarrow \Sigma_{m_-, m_+}$  which shifts the sequence  $x$  to the left as usual, but translates the symbol  $x_0$  from the forward alphabet  $\{0, \dots, m_+ - 1\}$

into the symbol  $\phi(x_0)$  from the backward alphabet  $\{0, \dots, m_- - 1\}$ ; that is,  $\sigma_\phi(x)$  is given by

$$\sigma_\phi(\dots, x_{-2}, x_{-1}; x_0, x_1, \dots) := (\dots, x_{-1}, \phi(x_0); x_1, x_2, \dots)$$

If for each symbol  $s \in \{0, \dots, m_- - 1\}$  we define the map  $\sigma_s^{-1}: \Sigma_{m_-}^- \rightarrow \Sigma_{m_-}^-$ ,  $\sigma_s^{-1}(x^-) = (\dots, x_{-1}, s)$ , then the action of  $\sigma_\phi$  on  $x = (x^-, x^+) \in \Sigma_{m_-, m_+}$  is

$$\sigma_\phi(x^-, x^+) = (\sigma_{\phi(x_0)}^{-1}(x^-), \sigma(x^+)). \tag{1}$$

We summarize these definitions here for the sake of reference.

**Definition 1** (Extended shift) Let  $m_-, m_+ \in \mathbb{Z}_{>0}$  be strictly positive integers and  $\phi: \{0, \dots, m_+ - 1\} \rightarrow \{0, \dots, m_- - 1\}$  a surjective function. The  $(m_-, m_+)$ -extended shift dynamical system (with transition function  $\phi$ ) is the system  $(\Sigma_{m_-, m_+}, \sigma_\phi)$  consisting of

1. the  $(m_-, m_+)$ -symbolic space  $\Sigma_{m_-, m_+} := \Sigma_{m_-}^- \times \Sigma_{m_+}^+$ , whose elements are pairs  $x = (x^-, x^+)$  with  $x^- = (\dots, x_{-2}, x_{-1})$  and  $x^+ = (x_0, x_1, \dots)$ , which can be identified with bilateral sequences

$$x = (\dots, x_{-2}, x_{-1}; x_0, x_1, \dots)$$

with  $x_{-1}, x_{-2}, \dots \in \{0, \dots, m_- - 1\}$  and  $x_0, x_1, \dots \in \{0, \dots, m_+ - 1\}$ ;

2. the  $(m_-, m_+)$ -extended shift  $\sigma_\phi: \Sigma_{m_-, m_+} \rightarrow \Sigma_{m_-, m_+}$ , which acts on each  $x \in \Sigma_{m_-, m_+}$  by

$$\sigma_\phi(x) := (\sigma_{\phi(x_0)}^{-1}(x^-), \sigma(x^+)) = (\dots, x_{-1}, \phi(x_0); x_1, x_2, \dots).$$

When  $m_+ = m_-$ , this is isomorphic to  $\Sigma_{m_+} = \{0, \dots, m_+ - 1\}^{\mathbb{Z}}$  with the standard bilateral shift  $\sigma$ , since  $\phi$  will just be a permutation of the symbols.

This system has the structure of an iterated function system (IFS) given by a skew product of base  $\Sigma_{m_+}^+$  and fiber  $\Sigma_{m_-}^-$ . (Notice that this inverts the usual order of the product in the notation for skew products, since we denote the fiber to the left and the base to the right; we do so because our perspective is primarily that of  $\Sigma_{m_-, m_+}$  as a space of bilateral sequences.) Using the maps  $\sigma_s^{-1}: \Sigma_{m_-}^- \rightarrow \Sigma_{m_-}^-$ ,  $\sigma_s^{-1}(x^-) = (\dots, x_{-1}, s)$ , Equation (1) shows the action on the fiber and the maps  $\sigma_s^{-1}: \Sigma_{m_-}^- \rightarrow \Sigma_{m_-}^-$  are easily shown to be contractions.

Apart from trivial cases, these systems are non-invertible. They can be used to encode non-invertible dynamics, which is not possible with the standard bilateral shift because it is invertible. An example is the  $m$ -to-1 baker’s transformation, which represents a non-invertible model of deterministic chaos and is a measure-preserving generalization of the classical (1-to-1) baker’s transformation. They are also mixing, ergodic and have chaotic behavior: they are transitive maps with dense periodic points [16].

### 1.4 Main Results

Given a probability distribution  $p^+ = (p_0^+, \dots, p_{m_+ - 1}^+)$  on the forward alphabet  $\{0, \dots, m_+ - 1\}$ , we define a measure  $\mu$  on  $\Sigma_{m_-, m_+}$  with respect to which the extended shift  $\sigma_\phi$  is measure-preserving and ergodic (see Subsection 3.2 for details). The first main result we obtain relates the Kolmogorov-Sinai metric entropy  $h_\mu(\sigma_\phi)$  of the extended shift  $\sigma_\phi$  with the metric entropy  $H_\mu(\mathcal{C}_0)$  of the partition  $\mathcal{C}_0$  by cylinders  $\mathcal{C}_0[s] = \{x \in \Sigma_{m_-, m_+} \mid x_0 = s\}$ ,  $0 \leq s < m_+$  (see Subsection 3.1), which is just the Shannon entropy of  $p^+$ . In particular, it shows  $h_\mu(\sigma_\phi)$  is not dependent on the transition function  $\phi$ .

**Theorem A** *The metric entropy of the  $(m_-, m_+)$ -extended shift  $\sigma_\phi$  (with respect to the measure  $\mu$  induced by the probability distribution  $p^+$ ) is*

$$h_\mu(\sigma_\phi) = H_\mu(\mathcal{C}_0) = \sum_{s=0}^{m_+-1} -p_s^+ \log(p_s^+).$$

The probability distribution  $p^+$  also induces a probability distribution  $p^- = (p_0^-, \dots, p_{m_- - 1}^-)$  on the backward alphabet  $\{0, \dots, m_- - 1\}$ , and these probability distributions are used to define “quotient distributions”  $q^{s-}$  for each  $s_- \in \{0, \dots, m_- - 1\}$  (these are related to disintegrations, cf. Subsection 5.3 for details).

Our second main result relates the folding entropy  $\mathcal{F}(\sigma_\phi)$  of the extended shift with the metric entropy of the partitions  $\mathcal{C}_0$  and  $\mathcal{C}_{-1}$  by cylinders  $\mathcal{C}_0[s]$  and  $\mathcal{C}_{-1}[z]$ , respectively, and explicitly to a formula involving the quotient distributions  $q^{s-}$ .

**Theorem B** *The folding entropy of the  $(m_-, m_+)$ -extended shift  $\sigma_\phi$  (with respect to the measure  $\mu$  induced by the probability distribution  $p^+$ ) is*

$$\mathcal{F}(\sigma_\phi) = H_\mu(\mathcal{C}_0) - H_\mu(\mathcal{C}_{-1}) = \sum_{s_-=0}^{m_- - 1} \left( \sum_{s_+ \in \phi^{-1}(s_-)} -q_{s_+}^{s_-} \log q_{s_+}^{s_-} \right) p_{s_-}^-.$$

When  $m_+ = m_-$ , the extended shift is in fact the standard bilateral Bernoulli shift, which is invertible and has zero folding entropy. That illustrates how the folding entropy measures the non-invertibility of the system.

**Structure of the paper.** Section 2 contains basic nomenclature, notation and definitions from measure theory and ergodic theory that we will use throughout the paper. In Section 3 we present the formal definition of the extended shift  $\sigma_\phi: \Sigma_{m_-, m_+} \rightarrow \Sigma_{m_-, m_+}$  and define their measure structure. In Section 4 we calculate the Kolmogorov–Sinai entropy of the extended shifts. We find the general form for cylinder sets pulled-back by the extended shift dynamics, which involves proving a technically difficult Lemma 4.4, then use the Kolmogorov–Sinai theorem to prove Theorem A. Finally, in Section 5 we calculate the folding entropy of the extended shifts. For this we calculate the disintegration of the measure  $\mu$  with respect to the dynamical pullback of the atomic partition. After the technical work of Subsections 5.2 and 5.3, we obtain Theorem B.

## 2 Preliminaries

We will denote the natural numbers (including 0) by  $\mathbb{N}$ , the integers by  $\mathbb{Z}$  and the real numbers by  $\mathbb{R}$ . We denote the strictly positive, positive, strictly negative, and negative integers by  $\mathbb{Z}_{>0}$ ,  $\mathbb{Z}_{\geq 0}$ ,  $\mathbb{Z}_{<0}$ , and  $\mathbb{Z}_{\leq 0}$ , respectively (and likewise for the other number sets).

### 2.1 Measure Spaces

Let  $X$  be a set. A  $\sigma$ -algebra over  $X$  is a family  $\mathcal{M}$  of subsets of  $X$ , whose elements are called *measurable sets*, that contains the empty set and is closed under set complements and countable unions. The pair  $(X, \mathcal{M})$  is called a *measurable space*. Given any family  $\mathcal{S}$  of subsets of  $X$ , the  $\sigma$ -algebra *generated* by  $\mathcal{S}$  is the smallest (relative to  $\subseteq$ )  $\sigma$ -algebra over  $X$  that contains  $\mathcal{S}$ .

A *measure* on  $(X, \mathcal{M})$  is a function  $\mu: \mathcal{M} \rightarrow \mathbb{R}_{\geq 0}$  that assigns the value 0 to the empty set and is *countably additive*, meaning that, for every pairwise disjoint countable family of measurable sets  $(M_i)_{i \in \mathbb{N}}$ ,

$$\mu\left(\bigcup_{i \in \mathbb{N}} M_i\right) = \sum_{i \in \mathbb{N}} \mu(M_i).$$

The triple  $(X, \mathcal{M}, \mu)$  is called a *measure space*. A *probability measure* is a measure such that  $\mu(X) = 1$ , and the respective measure space is called a *probability space*. We say that a property is valid for *almost every point of X* when it is valid for every point of a subset of  $X$  whose complement has measure 0.

A *measurable transformation* from a measure space  $(X, \mathcal{M})$  to another  $(X', \mathcal{M}')$  is a transformation  $f: X \rightarrow X'$  such that, for every measurable set  $M' \in \mathcal{M}'$ , its inverse image by  $f$  is measurable:  $f^{-1}(M') \in \mathcal{M}$ . A *measure-preserving transformation* from a measure space  $(X, \mathcal{M}, \mu)$  to another  $(X', \mathcal{M}', \mu')$  is a measurable transformation  $f: X \rightarrow X'$  such that, for every measurable set  $M' \in \mathcal{M}'$ ,  $\mu(f^{-1}(M')) = \mu'(M')$ .

On a measure space, the *integral* can be defined for functions  $f: X \rightarrow \mathbb{R}$ . We will denote the integral of  $f$  with respect to  $\mu$  over a measurable set  $M \subseteq X$  by  $\int_M f \mu$ , or by  $\int_{x \in M} f(x) \mu(dx)$  when it is necessary to make the argument  $x$  of  $f$  explicit.

## 2.2 Metric Entropy

Kolmogorov and Sinai first defined metric entropy, and used it as an invariant for dynamical systems over measure spaces [10, 24]. Here we briefly define it and state the main theorem we will use in this work, the Kolmogorov–Sinai theorem (Theorem 2.1). We refer the reader to [25] for the following definitions and any further information on metric entropy.

Let  $(X, \mathcal{M}, \mu)$  be a probability space. We will refer to any finite or countable family  $\mathcal{P}$  of pairwise disjoint measurable sets whose union has measure 1 by a *partition* of  $X$ . (This is similar to the usual definition of a partition, but weakened by the measure structure of the space). This defines, for almost every point  $x \in X$ , a unique set  $\pi_{\mathcal{P}}(x) \in \mathcal{P}$  such that  $x \in \pi_{\mathcal{P}}(x)$ , and hence a (almost everywhere defined) projection  $\pi_{\mathcal{P}}: X \rightarrow \mathcal{P}$ .

A partition  $\mathcal{P}$  is *coarser* than a partition  $\mathcal{P}'$  (or  $\mathcal{P}'$  is *finer* than  $\mathcal{P}$ ) when, for every element  $P' \in \mathcal{P}'$ , there exists an element  $P \in \mathcal{P}$  such that  $\mu(P' \setminus P) = 0$  (which means that almost every point of  $P'$  is contained in  $P$ ). This is denoted by  $\mathcal{P} \preceq \mathcal{P}'$ . We can also define an operation on the partitions: to each (finite or countable) family of partitions  $(\mathcal{P}_n)_{n \in \mathbb{N}}$ , its *correfinement* is

$$\bigvee_{i \in \mathbb{N}} \mathcal{P}_n := \left\{ \bigcap_{n \in \mathbb{N}} P_n \mid P_n \in \mathcal{P}_n \text{ for each } n \in \mathbb{N} \right\}.$$

When we have only 2 (of finitely many) partitions, we denote their correfinement by  $\mathcal{P} \vee \mathcal{P}'$ . The correfinement of a family of partitions is the smallest partition, relative to  $\preceq$ , that is larger than every partition of the family.

The *entropy* of  $\mathcal{P}$  is defined as

$$H_{\mu}(\mathcal{P}) := \sum_{P \in \mathcal{P}} -\mu(P) \log(\mu(P)). \tag{2}$$

(Here and in what follows, we always assume that  $0 \log 0 = 0$ .)

Now let  $f : M \rightarrow M$  be a measure-preserving transformation on  $(X, \mathcal{M}, \mu)$ . We can define the *pullback* of a partition  $\mathcal{P}$  by  $f$  as

$$f^{-1}(\mathcal{P}) := \{f^{-1}(P) \mid P \in \mathcal{P}\}.$$

This is also a partition in our specific sense. Then, for each  $n \in \mathbb{N}$ , the  $n$ -th *dynamical correfinement* of  $\mathcal{P}$  is

$$\mathcal{P}^n := \bigvee_{i=0}^{n-1} f^{-i}(\mathcal{P}) \tag{3}$$

and the  $n$ -th *bilateral dynamical correfinement* of  $\mathcal{P}$  is

$$\mathcal{P}^{\pm n} := \bigvee_{i=-n}^{n-1} f^{-i}(\mathcal{P}).$$

An element  $Q \in \mathcal{P}^n$  is of the form  $Q = P_0 \cap f^{-1}(P_1) \cap \dots \cap f^{-(n-1)}(P_{n-1})$ , for  $P_i \in \mathcal{P}$ , and a point  $x \in X$  belongs to  $Q$  if, and only if, for every  $0 \leq i \leq n - 1$ ,  $f^i(x) \in P_i$ . This shows that the elements of  $\mathcal{P}^n$  partition the space into points which have the same orbit under  $f$  for  $n$  units of time.

The *entropy* of  $f$  relative to  $\mathcal{P}$  is the limit

$$h_\mu(f, \mathcal{P}) := \lim_{n \rightarrow \infty} \frac{1}{n} H_\mu(\mathcal{P}^n).$$

(Notice that  $\mathcal{P}^n$  depends on  $f$  even though the notation does not make it explicit). The *entropy* of  $f$  is then the supremum of the entropies relative to all partitions with finite entropy (or, equivalently, finite partitions):

$$h_\mu(f) := \sup_{\mathcal{P}} h_\mu(f, \mathcal{P}).$$

This definition is very abstract and requires information about every finite partition, but there is a way to calculate the entropy of a transformation using only a sequence of partitions that have a special property. This is the content of the following theorem, which we are going to use to obtain Theorem A. The proof can be found in [25].

**Theorem 2.1 (Kolmogorov-Sinai)** *Let  $(X, \mathcal{M}, \mu)$  be a probability space,  $f : X \rightarrow X$  a measure-preserving transformation and  $(\mathcal{P}_n)_{n \in \mathbb{N}}$  be an increasing sequence of partitions<sup>1</sup> with finite entropy such that  $\bigcup_{i \in \mathbb{N}} \mathcal{P}_i$  generates  $\mathcal{M}$  (up to measure 0). Then*

$$h_\mu(f) = \lim_{n \rightarrow \infty} h_\mu(f, \mathcal{P}_n).$$

### 2.3 Disintegration of Measure

Given a probability space  $(X, \mathcal{M}, \mu)$  and a partition  $\mathcal{P}$  (we do not require the partition to be countable here), we have the (almost everywhere defined) natural projection  $\pi_{\mathcal{P}} : X \rightarrow \mathcal{P}$ . Using  $\pi_{\mathcal{P}}$  we can pushforward a probability space structure onto  $\mathcal{P}$ , namely  $(\mathcal{P}, \hat{\mathcal{M}}, \hat{\mu})$ , in which

$$\hat{\mathcal{M}} := \left\{ \mathcal{Q} \subseteq \mathcal{P} \mid \pi_{\mathcal{P}}^{-1}(\mathcal{Q}) \in \mathcal{M} \right\}$$

<sup>1</sup> That is, for every  $n, m \in \mathbb{N}$ , if  $n \leq m$  then  $\mathcal{P}_n \leq \mathcal{P}_m$ .

is the pushforward  $\sigma$ -algebra (or quotient  $\sigma$ -algebra) and

$$\hat{\mu}(\mathcal{Q}) := \mu(\pi_{\mathcal{P}}^{-1}(\mathcal{Q})) \quad \text{for every } \mathcal{Q} \in \hat{\mathcal{M}}$$

is the pushforward measure (or quotient measure).

**Definition 2** Let  $(X, \mathcal{M}, \mu)$  be a probability space and  $\mathcal{P}$  a partition of  $X$ . A *disintegration* of  $\mu$  with respect to  $\mathcal{P}$  is a family of probability measures  $(\mu_P)_{P \in \mathcal{P}}$  on  $X$  such that

1. For almost every  $P \in \mathcal{P}$ ,  $\mu_P(P) = 1$ ;
2. For every measurable set  $M \in \mathcal{M}$ , the transformation  $\mathcal{P} \rightarrow \mathbb{R}$ ,  $P \mapsto \mu_P(M)$  is measurable;
3. For every measurable set  $M \in \mathcal{M}$ ,

$$\mu(M) = \int_{P \in \mathcal{P}} \mu_P(M) \hat{\mu}(dP). \quad (4)$$

Intuitively, this describes the way we can relate the Lebesgue measure on a square with the Lebesgue measure on each of its vertical sections by integration using Fubini's theorem. Under certain conditions on the partition  $\mathcal{P}$ , the disintegration of a measure is unique up to measure zero and always exists [25].

## 2.4 Conditional and Folding Entropies

Besides defining the entropy of a partition as in Equation (2), we can also define the conditional entropy of a partition  $\mathcal{P}$  relative to a partition  $\mathcal{P}'$ . We follow the approach of Rokhlin [21]. First we define, for each  $P' \in \mathcal{P}'$ , the *partition induced by  $\mathcal{P}$  on  $P'$*  as

$$\mathcal{P}|_{P'} := \{P \cap P' \mid P \in \mathcal{P}\}.$$

Then the *conditional entropy of  $\mathcal{P}$  with respect to  $\mathcal{P}'$*  is defined [21, Section 5.1] using the disintegration of the measure  $\mu$  with respect to  $\mathcal{P}'$  by

$$H_{\mu}(\mathcal{P} \mid \mathcal{P}') = \int_{P' \in \mathcal{P}'} H_{\mu_{P'}}(\mathcal{P}|_{P'}) \mu_{P'}(dP'). \quad (5)$$

This is a more general definition that works for non-countable partitions. In the case the partitions are countable, we obtain the simplified formula presented in [25].

In [22] the author introduces the *folding entropy* for  $\mathcal{C}^1$  transformations. It can be defined [15, 26] as the conditional entropy of the *atomic partition*

$$\epsilon := \{\{x\} \mid x \in X\} \quad (6)$$

with respect to its *dynamical pullback*

$$f^{-1}(\epsilon) = \{f^{-1}(x) \mid x \in X\}.$$

**Definition 3** Let  $X$  be a probability space and  $f: X \rightarrow X$  a measure-preserving transformation. The *folding entropy* of  $f$  with respect to  $\mu$  is

$$\mathcal{F}(f) := H_{\mu}(\epsilon \mid f^{-1}(\epsilon)).$$

### 3 Extended Shifts

The extended shifts, also known as zip shifts, are a generalization of bilateral symbolic shifts. They were first introduced in [12, 13] and expanded on in [16]. Instead of a single set of symbols ranging in  $\{0, \dots, m - 1\}$ , used to compose a bilateral symbolic sequence  $x = (x_i)_{i \in \mathbb{Z}} \in \{0, \dots, m - 1\}^{\mathbb{Z}}$ , we consider bilateral sequences that have symbols ranging in  $\{0, \dots, m_+ - 1\}$  on their positive indices part and in  $\{0, \dots, m_- - 1\}$  on their negative indices part. To be able to still define the shift transformation, a function that translates one type of symbols to the other is needed. Definition 1 formalizes this construction.

Let us first recall that, for  $m \in \mathbb{Z}_{>0}$ , we define the *symbolic spaces*  $\Sigma_m := \{0, \dots, m - 1\}^{\mathbb{Z}}$ ,  $\Sigma_m^+ := \{0, \dots, m - 1\}^{\mathbb{Z}_{\geq 0}}$  and  $\Sigma_m^- := \{0, \dots, m - 1\}^{\mathbb{Z}_{< 0}}$ , the *bilateral shift map*  $\sigma : \Sigma_m \rightarrow \Sigma_m, (\dots, x_{-1}; x_0, x_1, \dots) \mapsto (\dots, x_{-1}, x_0; x_1, \dots)$ , the *unilateral shift map*  $\sigma : \Sigma_m^+ \rightarrow \Sigma_m^+, (x_0, x_1, \dots) \mapsto (x_1, x_2, \dots)$ , and, for each  $s \in \{0, \dots, m - 1\}$ , the *inverted unilateral shift map with inserted symbol s* as  $\sigma_s^{-1} : \Sigma_m^+ \rightarrow \Sigma_m^+, (x_0, x_1, \dots) \mapsto (s, x_0, x_1, \dots)$ . The unilateral shift and inverted unilateral shift are defined analogously on  $\Sigma_m^-$ . We now recall Definition 1.

**Definition 1** (Extended shift). Let  $m_-, m_+ \in \mathbb{Z}_{>0}$  be strictly positive integers and  $\phi : \{0, \dots, m_+ - 1\} \rightarrow \{0, \dots, m_- - 1\}$  a surjective function. The  $(m_-, m_+)$ -*extended shift* dynamical system (with *transition function*  $\phi$ ) is the system  $(\Sigma_{m_-, m_+}, \sigma_\phi)$  consisting of

1. the  $(m_-, m_+)$ -*symbolic space*  $\Sigma_{m_-, m_+} := \Sigma_{m_-}^- \times \Sigma_{m_+}^+$ , whose elements are pairs  $x = (x^-, x^+)$  with  $x^- = (\dots, x_{-2}, x_{-1})$  and  $x^+ = (x_0, x_1, \dots)$ , which can be identified with bilateral sequences

$$x = (\dots, x_{-2}, x_{-1}; x_0, x_1, \dots)$$

with  $x_{-1}, x_{-2}, \dots \in \{0, \dots, m_- - 1\}$  and  $x_0, x_1, \dots \in \{0, \dots, m_+ - 1\}$ ;

2. the  $(m_-, m_+)$ -*extended shift*  $\sigma_\phi : \Sigma_{m_-, m_+} \rightarrow \Sigma_{m_-, m_+}$ , which acts on each  $x \in \Sigma_{m_-, m_+}$  by

$$\sigma_\phi(x) := (\sigma_{\phi(x_0)}^{-1}(x^-), \sigma(x^+)) = (\dots, x_{-1}, \phi(x_0); x_1, x_2, \dots).$$

We can also define a notion of an extended Bernoulli transformation.

**Definition 4** A measure-preserving map  $f : X \rightarrow X$  defined on a Lebesgue space is an  $(m_-, m_+)$ -Bernoulli transformation if it is isomorphic (mod 0) to an  $(m_-, m_+)$ -extended shift  $\sigma_\phi$ .

The 2-to-1 baker’s transformation defined in [16] exemplifies a  $(2, 4)$ -Bernoulli transformation. We omit the formal definition here, but Figure 1 shows how this transformation is defined on the square  $Q$  in three steps, and Figure 2 shows the partitions of the square that are used to encode the system and obtain the isomorphism to a  $(2, 4)$ -extended shift and how these partitions iterate under the action of the dynamics over time.

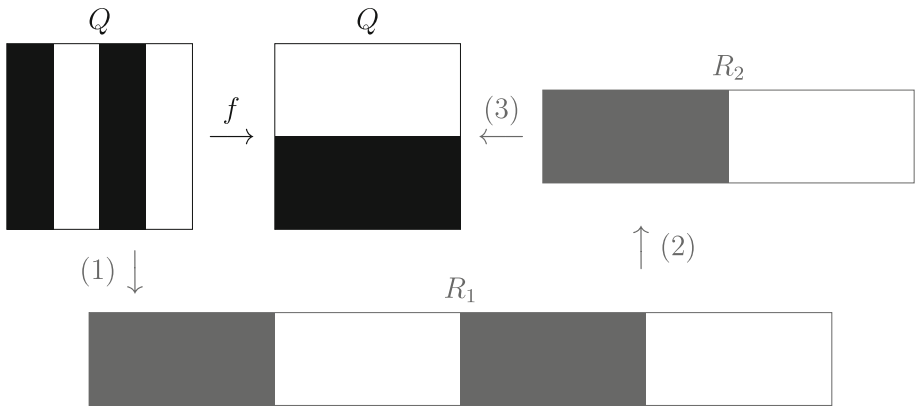
### 3.1 Measurable Structure

The  $\sigma$ -algebra  $\mathcal{B}$  of the space  $\Sigma_{m_-, m_+}$  is the one generated by cylinder sets: for each  $i \in \mathbb{Z}$ , and for each  $s \in \{0, \dots, m_- - 1\}$  if  $i < 0$  or  $s \in \{0, \dots, m_+ - 1\}$  if  $i \geq 0$ , we define the *cylinder*

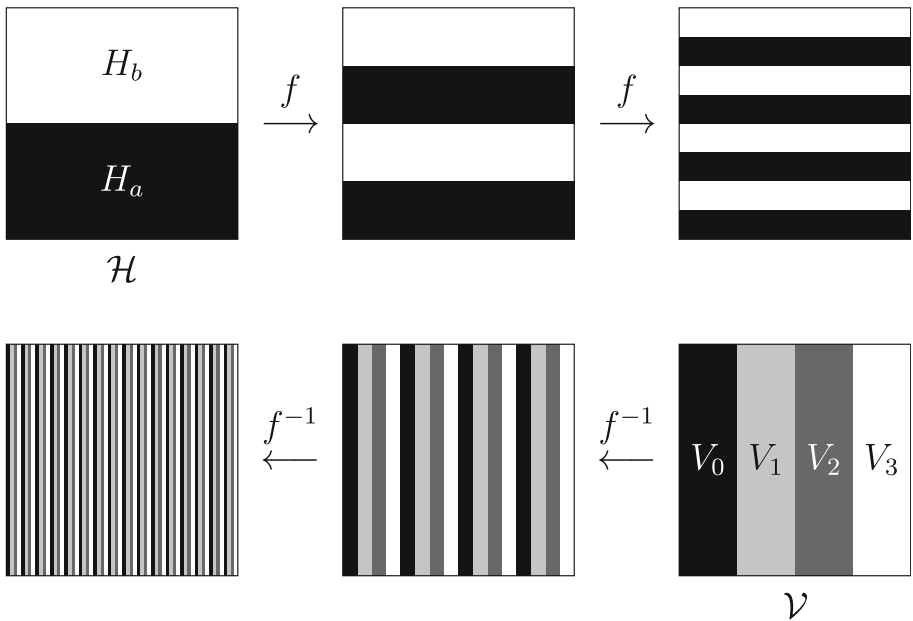
$$C_i[s] := \{x \in \Sigma_{m_-, m_+} \mid x_i = s\}$$

and denote

$$C_{i_1, \dots, i_k}[s_1, \dots, s_k] := \{x \in \Sigma_{m_-, m_+} \mid x_{i_1} = s_1, \dots, x_{i_k} = s_k\} = C_{i_1}[s_1] \cap \dots \cap C_{i_k}[s_k].$$



**Fig. 1** The 2-to-1 baker's transformation. On step (1), the square  $Q$  is dilated by 4 in the horizontal direction and by  $\frac{1}{2}$  in the vertical direction, resulting in a rectangle  $R_1$ . On step (2), the right-most bands of the rectangle  $R_1$  are glued over the left-most ones, resulting in a smaller rectangle  $R_2$ . On step (3), the left right band of the rectangle  $R_2$  is glued above the left band, resulting again in the square  $Q$



**Fig. 2** The first 2 iterations of the horizontal partition  $\mathcal{H} = \{H_a, H_b\}$  and the vertical partition  $\mathcal{V} = \{V_0, V_1, V_2, V_3\}$  of the 2-to-1 baker's transformation. In the limit, the iterations of the horizontal partition is the partition of the square by horizontal line, and the iteration of the vertical partition is the partition by vertical lines

For simplicity, when the  $i_j$ 's increase by 1, we just denote this by  $C_{i_1}[s_1, \dots, s_k]$ . We also define, for  $s_+ \in \{0, \dots, m_+ - 1\}$ , the *extended cylinder*

$$C_i[\phi^{-1}(s_+)] := \bigcup_{s_- \in \phi^{-1}(s_+)} C_i[s_-].$$

The next proposition shows how the dynamics acts backwards and forwards on cylinders.

**Proposition 3.1** *Let  $k \in \mathbb{Z}_{\geq 0}$  and  $0 \leq s < \max\{m_-, m_+\}$ . Then*

$$\sigma_\phi^{-k}(C_i[s]) = \begin{cases} C_{i+k}[s], & i \notin [-k, -1] \cap \mathbb{Z} \\ C_{i+k}[\phi^{-1}(s)], & i \in [-k, -1] \cap \mathbb{Z}. \end{cases}$$

and

$$\sigma_\phi^k(C_i[s]) = \begin{cases} C_{i-k}[s], & i \notin [0, k-1] \cap \mathbb{Z} \\ C_{i-k}[\phi(s)], & i \in [0, k-1] \cap \mathbb{Z}. \end{cases}$$

**Proof** For the inverse image, it holds that

$$\sigma_\phi^{-1}(C_i[s]) = \begin{cases} C_{i+1}[s], & i \neq -1 \\ C_0[\phi^{-1}(s)], & i = -1 \end{cases}$$

and, for the direct image, it holds that

$$\sigma_\phi(C_i[s]) = \begin{cases} C_{i-1}[s], & i \neq 0 \\ C_{-1}[\phi(s)], & i = 0. \end{cases}$$

Then, by induction on  $k$ , we obtain the statement. □

### 3.2 Measure Structure

In order to define a measure on  $(\Sigma_{m_-, m_+}, \mathcal{B})$ , it is sufficient to define it on the cylinders  $C_i[s]$ . We start with a probability measure  $\mu^+$  on the symbol set  $\{0, \dots, m_+ - 1\}$ . Since  $\{0, \dots, m_+ - 1\}$  is a finite set with atomic  $\sigma$ -algebra, this probability measure can be identified with a discrete probability distribution  $p^+ = (p_0^+, \dots, p_{m_+ - 1}^+)$  (that is, for every  $s \in \{0, \dots, m_+ - 1\}$  we have  $p_s^+ \in \mathbb{R}_{\geq 0}$ , and  $\sum_{s=0}^{m_+ - 1} p_s^+ = 1$ ) by defining, for each  $0 \leq s < m_+$ ,

$$p_s^+ := \mu^+(\{s\}).$$

Using the surjective transition function  $\phi: \{0, \dots, m_+ - 1\} \rightarrow \{0, \dots, m_- - 1\}$ , we can pushforward this probability measure  $\mu^+$  to the probability measure  $\mu^- := \phi_*\mu^+$  on  $\{0, \dots, m_- - 1\}$ . This is done by considering the partition  $\{\phi^{-1}(z)\}_{z \in \{0, \dots, m_- - 1\}}$  of  $\{0, \dots, m_+ - 1\}$  by the inverse images of elements of  $\{0, \dots, m_- - 1\}$ . The pushforward measure of  $\{z\} \subseteq \{0, \dots, m_- - 1\}$  is then the sum of the measure of all the elements of  $\phi^{-1}(z)$  on  $\{0, \dots, m_+ - 1\}$ , given for each  $z \in \{0, \dots, m_- - 1\}$  by

$$\mu^-(\{z\}) = \phi_*\mu^+(\{z\}) = \mu^+(\phi^{-1}(\{z\})) = \sum_{s \in \phi^{-1}(z)} \mu^+(\{s\}).$$

In the same way as we did for  $p^+$ , we can identify the measure  $\mu^-$  with a probability distribution  $p^- = (p_0^-, \dots, p_{m_- - 1}^-)$  by setting, for each  $0 \leq z < m_-$ ,

$$p_z^- := p^-(\{z\}).$$

Then, for a cylinder  $C_i[s]$ , we can define its measure as  $p_s^+$  if  $i \geq 0$  and  $p_s^-$  if  $i < 0$ .

**Definition 5** Let  $(\Sigma_{m_-, m_+}, \sigma_\phi)$  be an  $(m_-, m_+)$ -extended shift space,  $p^+$  a probability measure on  $\{0, \dots, m_+ - 1\}$  and  $p^- = \phi_* p^+$  the pushforward probability measure on  $\{0, \dots, m_- - 1\}$ . The *probability measure* on  $(\Sigma_{m_-, m_+}, \sigma_\phi)$  induced by  $p^+$  is the probability measure  $\mu: \mathcal{B} \rightarrow [0, 1]$  defined on cylinders by

$$\mu(C_i[s]) := \begin{cases} p_s^-, & i < 0 \\ p_s^+, & i \geq 0 \end{cases} = \begin{cases} \sum_{s' \in \phi^{-1}(s)} p_{s'}^+, & i < 0 \\ p_s^+, & i \geq 0. \end{cases}$$

From the way we defined the measure on  $\{0, \dots, m_- - 1\}$  by the pushforward, it is easy to show that the extended shift dynamics is measure-preserving. We just need to be careful considering the different cases.

**Proposition 3.2** Let  $(\Sigma_{m_-, m_+}, \sigma_\phi)$  be an  $(m_-, m_+)$ -extended shift space and  $p^+$  a probability measure on  $\{0, \dots, m_+ - 1\}$ . The dynamics  $\sigma_\phi$  preserves the measure  $\mu$ .

**Proof** It suffices to show that, for every basic cylinder  $C_i[s]$ ,

$$\mu(\sigma_\phi^{-1}(C_i[s])) = \mu(C_i[s]).$$

We consider three cases:

- $(i \geq 0)$  In this case,  $\sigma_\phi^{-1}(C_i[s]) = C_{i+1}[s]$  (Proposition 3.1). Since  $i + 1 \geq 1$ , it follows from Definition 5 that

$$\mu(\sigma_\phi^{-1}(C_i[s])) = \mu(C_{i+1}[s]) = p_s^+ = \mu(C_i[s]).$$

- $(i < -1)$  In this case, it also holds that  $\sigma_\phi^{-1}(C_i[s]) = C_{i+1}[s]$  (Proposition 3.1). Since  $i + 1 < 0$ , it follows from Definition 5 that

$$\mu(\sigma_\phi^{-1}(C_i[s])) = \mu(C_{i+1}[s]) = p_s^- = \mu(C_i[s]).$$

- $(i = -1)$  In this case,  $\sigma_\phi^{-1}(C_{-1}[s]) = C_0[\phi^{-1}(s)] = \bigcup_{s' \in \phi^{-1}(s)} C_0[s']$  (Proposition 3.1). Since  $i + 1 = 0$ , it follows from Definition 5 that

$$\mu(\sigma_\phi^{-1}(C_i[s])) = \mu\left(\bigcup_{s' \in \phi^{-1}(s)} C_0[s']\right) = \sum_{s' \in \phi^{-1}(s)} \mu(C_0[s']) = \sum_{s' \in \phi^{-1}(s)} p_{s'}^+ = \mu(C_i[s]).$$

□

## 4 Metric Entropy of Extended Shifts

This section is dedicated to prove Theorem A. We will calculate the metric entropy of  $(\Sigma_{m_-, m_+}, \sigma_\phi)$  and relate it to the entropy of the probability distributions  $p^+$  and  $p^-$ . We initially present a proof sketch to guide the reader through the lemmas needed to complete the proof.

### 4.1 Proof Sketch of Theorem A

**Proof sketch of Theorem A** We will use the Kolmogorov-Sinai theorem (Theorem 2.1) to calculate the entropy of  $(\Sigma_{m_-, m_+}, \sigma_\phi)$ . In order to do that, we consider the sequence of partitions by cylinders  $\mathcal{P}_n := \mathcal{C}_{-n, \dots, n-1}$  ( $n \in \mathbb{N}$ ), which satisfies the hypotheses of Theorem 2.1 as a

consequence of Lemmas 4.2 and 4.3. Then Kolmogorov-Sinai implies that the metric entropy of the system is

$$h_\mu(\sigma_\phi) = \lim_{n \rightarrow \infty} h_\mu(\sigma_\phi, \mathcal{P}_n).$$

We thus have to calculate  $h_\mu(\sigma_\phi, \mathcal{P}_n)$ , which is, by definition,

$$h_\mu(\sigma_\phi, \mathcal{P}_n) := \lim_{k \rightarrow \infty} \frac{1}{k} H_\mu(\mathcal{P}_n^k),$$

which in turn shows we have to calculate  $H_\mu(\mathcal{P}_n^k)$ .

In Lemma 4.4 we show that  $\mathcal{P}_n^k = \mathcal{C}_{-n, \dots, n+k-2}$  (for large  $k$ ) and in Lemma 4.3 we show that  $H_\mu(\mathcal{C}_{-n, \dots, 0, \dots, n'-1}) = nH_\mu(\mathcal{C}_{-1}) + n'H_\mu(\mathcal{C}_0)$ , so we can conclude in Lemma 4.5 that  $h_\mu(\sigma_\phi, \mathcal{P}_n) = H_\mu(\mathcal{C}_0)$ . From Lemma 4.2 we have  $H_\mu(\mathcal{C}_0) = \sum_{s=0}^{m_+-1} -p_s^+ \log p_s^+$ , therefore we obtain

$$h_\mu(\sigma_\phi) = H_\mu(\mathcal{C}_0) = \sum_{s=0}^{m_+-1} -p_s^+ \log p_s^+.$$

□

### 4.2 Partitions by Cylinders

We begin by defining some basic partitions of our space.

**Definition 6** Let  $i \in \mathbb{Z}$ . The *partition by cylinders of index  $i$*  is the partition

$$\mathcal{C}_i := \begin{cases} \{\mathcal{C}_i[s] \mid s \in \{0, \dots, m_+ - 1\}\} & i \geq 0 \\ \{\mathcal{C}_i[s] \mid s \in \{0, \dots, m_- - 1\}\} & i < 0. \end{cases}$$

Let  $n, n' \in \mathbb{Z}$ . The *partition by cylinders of indices from  $n$  to  $n'$*  is the partition

$$\mathcal{C}_{n, \dots, n'} := \bigvee_{i=n}^{n'} \mathcal{C}_i.$$

The following simple lemma sums up how the dynamics of the shift  $\sigma_\phi$  acts on these partitions.

**Lemma 4.1** For every  $i \geq 0$ ,

1.  $\sigma_\phi^i(\mathcal{C}_0) = \mathcal{C}_{-i}$ ;
2.  $\sigma_\phi^{-i}(\mathcal{C}_0) = \mathcal{C}_i$ ;
3.  $\sigma_\phi^{-i}(\mathcal{C}_{-(i+1)}) = \mathcal{C}_{-1}$ ;
4.  $\mathcal{C}_0^n = \mathcal{C}_{0, \dots, n-1}$ .
5.  $\mathcal{C}_0^{\pm n} = \mathcal{C}_{-n, \dots, n-1}$ .

**Proof** This is a consequence of Proposition 3.1.

1. Since  $\sigma_\phi(\mathcal{C}_0[s]) = \mathcal{C}_{-1}[\phi(s)]$  and  $\phi$  is surjective, it follows that  $\sigma_\phi(\mathcal{C}_0) = \mathcal{C}_{-1}$ . By induction,  $\sigma_\phi^i(\mathcal{C}_0) = \mathcal{C}_{-i}$ .
2. Since  $\sigma_\phi^{-1}(\mathcal{C}_0[s]) = \mathcal{C}_1[s]$ , it follows that  $\sigma_\phi^{-1}(\mathcal{C}_0) = \mathcal{C}_1$ . By induction,  $\sigma_\phi^{-i}(\mathcal{C}_0) = \mathcal{C}_i$ .
3. Since  $\sigma^{-1}(\mathcal{C}_{-2}[s]) = \mathcal{C}_{-1}[s]$ , it follows that  $\sigma_\phi^{-1}(\mathcal{C}_{-2}) = \mathcal{C}_{-1}$ . By induction,  $\sigma_\phi^{-i}(\mathcal{C}_{-(i+1)}) = \mathcal{C}_{-1}$ .

4. It follows that

$$C_0^n = \bigvee_{i=0}^{n-1} \sigma_\phi^{-i}(C_0) = \bigvee_{i=0}^{n-1} C_i.$$

5. It follows that

$$C_0^{\pm n} = \bigvee_{i=-n}^{n-1} \sigma_\phi^{-i}(C_0) = \bigvee_{i=-n}^{n-1} C_i.$$

□

### 4.3 Technical Lemmas for Theorem A

In this subsection, we present a series of technical lemmas that will be essential in the subsequent proof of Theorem A. We start our lemmas by calculating the entropy of the simple partitions  $C_0$  and  $C_{-1}$ .

**Lemma 4.2**  $H_\mu(C_0) = \sum_{s=0}^{m_+-1} -p_s^+ \log p_s^+$  and  $H_\mu(C_{-1}) = \sum_{s=0}^{m_--1} -p_s^- \log p_s^-$ .

**Proof** It follows directly from Equation (2) from the simple calculations

$$H(C_0) = \sum_{s=0}^{m_+-1} -\mu(C_0[s]) \log(\mu(C_0[s])) = \sum_{s=0}^{m_+-1} -p_s^+ \log p_s^+.$$

and

$$H(C_{-1}) = \sum_{s=0}^{m_--1} -\mu(C_{-1}[s]) \log(\mu(C_{-1}[s])) = \sum_{s=0}^{m_--1} -p_s^- \log p_s^-.$$

□

This shows, as could be expected, that the entropy of the partition  $C_0$  is related to  $p^+$ , the distribution of the positive part of the extended shift  $\Sigma_{m_-,m_+}$ , while the entropy of the partition  $C_{-1}$  is related to  $p^-$ , the distribution of the negative part of  $\Sigma_{m_-,m_+}$ . We can now calculate the metric entropy of a partition by cylinders other than the basic  $C_0$  and  $C_{-1}$ .

**Lemma 4.3**  $H_\mu(C_{-n,\dots,0,\dots,n'-1}) = nH_\mu(C_{-1}) + n'H_\mu(C_0)$ .

**Proof** For every  $i \geq 1$ , it holds that  $C_i = \sigma_\phi^{-i}(C_0)$  and  $\sigma_\phi^{-i}(C_{-(i+1)}) = C_{-1}$  (Lemma 4.1). Since  $\sigma_\phi$  preserves the measure  $\mu$  (Proposition 3.2), it follows that  $H_\mu(C_i) = H_\mu(C_0)$  and  $H_\mu(C_{-(i+1)}) = H_\mu(C_{-1})$ .

Besides that, for any integers  $i < i'$ , the partitions  $C_i$  and  $C_{i'}$  are independent, because  $C_i[s] \cap C_{i'}[s'] = C_{i,i'}[s, s']$  and  $\mu(C_{i,i'}[s, s']) = \mu(C_i[s])\mu(C_{i'}[s'])$ . Thus it follows that

$$H_\mu(C_{-n,\dots,0,\dots,n'-1}) = H_\mu\left(\bigvee_{i=-n}^{n'-1} C_i\right) = \sum_{i=-n}^{n'-1} H_\mu(C_i) = nH_\mu(C_{-1}) + n'H_\mu(C_0).$$

□

In particular, since  $C_0^n = C_{0,\dots,n-1}$  (Lemma 4.1), this implies that

$$h_\mu(\sigma_\phi, C_0) = \lim_{n \rightarrow \infty} \frac{1}{n} H_\mu(C_0^n) = \lim_{n \rightarrow \infty} \frac{1}{n} nH_\mu(C_0) = H_\mu(C_0).$$

To calculate the metric entropy of the system, we will use the Kolmogorov-Sinai theorem (Theorem 2.1), as explained in the proof sketch of Theorem A (Subsection 4.1). To that end we define the following sequence of partitions.

**Definition 7** For every  $n \in \mathbb{N}$ ,  $\mathcal{P}_n := \mathcal{C}_0^{\pm n} = \mathcal{C}_{-n, \dots, n-1}$ .

We will eventually need to use the metric entropy of  $\mathcal{P}_n^k$ , the  $k$ -th dynamical correfinement of the partition  $\mathcal{P}_n$  (check Equation (3)), so the following lemma shows that it is just a partition by cylinders.

**Lemma 4.4** Let  $n \geq 1$  and  $k \geq 2n$ . Then  $\mathcal{P}_n^k = \mathcal{C}_{-n, \dots, n+k-2}$ .

**Proof** The dynamical correfinement of  $\mathcal{P}_n$  is defined by  $\mathcal{P}_n^k = \bigvee_{j=0}^{k-1} \sigma_\phi^{-j}(\mathcal{P}_n)$ , so let us first calculate a generic element of the pullback partition

$$\sigma_\phi^{-j}(\mathcal{P}_n) = \left\{ \sigma_\phi^{-j}(C) \mid C \in \mathcal{P}_n \right\}.$$

Each cylinder of  $\mathcal{P}_n = \mathcal{C}_{-n, \dots, n-1}$  has the form

$$\mathcal{C}_{-n, \dots, n-1}[s_{-n}, \dots, s_{n-1}] = \bigcap_{i=-n}^{n-1} C_i[s_i],$$

with  $s_i \in \{0, \dots, m_- - 1\}$  if  $i < 0$  and  $s_i \in \{0, \dots, m_+ - 1\}$  if  $i \geq 0$ . Then

$$\sigma_\phi^{-j}(\mathcal{C}_{-n, \dots, n-1}[s_{-n}, \dots, s_{n-1}]) = \sigma_\phi^{-j} \left( \bigcap_{i=-n}^{n-1} C_i[s_i] \right) = \bigcap_{i=-n}^{n-1} \sigma_\phi^{-j}(C_i[s_i]).$$

Based on Proposition 3.1, we can separate this in 3 intersections<sup>2</sup> as follows:

$$\begin{aligned} \sigma_\phi^{-j}(\mathcal{C}_{-n, \dots, n-1}[s_{-n}, \dots, s_{n-1}]) &= \\ &= \bigcap_{i=-n}^{-(j+1)} \sigma_\phi^{-j}(C_i[s_i]) \cap \bigcap_{i=-j}^{-1} \sigma_\phi^{-j}(C_{-1}[s_{-1}]) \cap \bigcap_{i=0}^{n-1} \sigma_\phi^{-j}(C_i[s_i]) \\ &= \bigcap_{i=-n}^{-(j+1)} C_{i+j}[s_i] \cap \bigcap_{i=-j}^{-1} C_{i+j}[\phi^{-1}(s_i)] \cap \bigcap_{i=0}^{n-1} C_{i+j}[s_i]. \end{aligned} \tag{7}$$

Notice that in Equation (7), for  $-n \leq i \leq -(j+1)$  and  $0 \leq i \leq n-1$  we have basic cylinders of the form  $C_{i+j}[s_i]$  and, for  $-j \leq i \leq -1$ , we have extended cylinders (unions of cylinders) of the form

$$C_{i+j}[\phi^{-1}(s_i)] = \bigcup_{s \in \phi^{-1}(s_i)} C_{i+j}[s].$$

This shows that  $\sigma_\phi^{-j}(\mathcal{P}_n)$  is not a partition by cylinders (unless  $\phi$  is bijective and hence the sets  $\phi^{-1}(s_i^j)$  are singletons, but this is just a regular shift, not the usual case for extended shifts).

<sup>2</sup> In order to simplify notation, we define that intersections that have the top index strictly smaller than the bottom index should be consider to be the whole space  $\Sigma_{m_-, m_+}$ , so that they can be ignored. In Equation (7), this happens for the first intersection in the case  $j > n-1$  (or equivalently  $-(j+1) < -n$ ) and for the second intersection in the case  $j = 0$  (or equivalently  $-1 < -j$ ).

We must now calculate a generic element of  $\mathcal{P}_n^k = \bigvee_{j=0}^{k-1} \sigma_\phi^{-j}(\mathcal{P}_n)$ . To that end, for each  $0 \leq j \leq k - 1$  we take cylinders  $C^j \in \mathcal{P}_n$ , defined by

$$C^j := C_{-n, \dots, n-1}[s_{-n}^j, \dots, s_{n-1}^j] = \bigcap_{i=-n}^{n-1} C_i[s_i^j]$$

with  $s_i^j \in \{0, \dots, m_- - 1\}$  if  $i < 0$  and  $s_i^j \in \{0, \dots, m_+ - 1\}$  if  $i \geq 0$ . An element of  $\mathcal{P}_n^k$  is a non-empty set of the form  $\bigcap_{j=0}^{k-1} \sigma_\phi^{-j}(C^j)$ . From Equation (7), it follows that this set is given by

$$\begin{aligned} & \bigcap_{j=0}^{k-1} \sigma_\phi^{-j}(C_{-n, \dots, n-1}[s_{-n}^j, \dots, s_{n-1}^j]) = \\ & = \bigcap_{j=0}^{k-1} \bigcap_{i=-n}^{k-1-(j+1)} C_{i+j}[s_i^j] \cap \bigcap_{j=0}^{k-1} \bigcap_{i=-j}^{-1} C_{i+j}[\phi^{-1}(s_i^j)] \cap \bigcap_{j=0}^{k-1} \bigcap_{i=0}^{n-1} C_{i+j}[s_i^j]. \end{aligned} \tag{8}$$

This shows that a generic element of  $\mathcal{P}_n^k$  (as in Equation (8)) is an intersection of basic cylinders and extended cylinders (which are unions of basic cylinders). These cylinders on the right-hand side of Equation (8) are indexed by  $l := i + j$ , which varies between  $-n$  and  $n - k - 2$  since  $j$  varies between 0 and  $k - 1$ , and  $i$  varies between  $-n$  and  $n - 1$ .

We wish to find conditions on the symbols  $s_i^j$  that guarantee the intersections in Equation (8) is non-empty. For that, we will reorganize the intersections based on the indices  $l$  and  $j$ . Define  $B_l$  to be the intersection of every cylinder and extended cylinder in Equation (8) that has index  $l$ . Thus

$$\bigcap_{j=0}^{k-1} \sigma_\phi^{-j}(C_{-n, \dots, n-1}[s_{-n}^j, \dots, s_{n-1}^j]) = \bigcap_{l=-n}^{n+k-2} B_l, \tag{9}$$

and each set  $B_l$  is an intersection that depends on a range of values of  $j$ .

Since the intersection of a cylinder or extended cylinder with another cylinder or extended cylinder is non-empty if they have different indices, the intersection on the right-hand side of Equation (9) is non-empty if, and only if, each  $B_l \neq \emptyset$ . In what follows we shall determine the range of  $j$  for each  $l$  and find conditions on the symbols  $s_i^j$ . We separate our analysis in many cases.

1. ( $-n \leq l \leq -1$ ) In this case  $0 \leq j \leq l + n$  and no extended cylinder occurs. In order to have  $B_l \neq \emptyset$ , all the relations in Table 1 must be satisfied, and hence

$$B_l = \bigcap_{j=0}^{l+n} C_l[s_{l-j}^j] = C_l[s_l^0]. \tag{10}$$

2. ( $0 \leq l \leq n - 1$ ) In this case, when  $0 \leq j \leq l$  we have basic cylinders and when  $l + 1 \leq j \leq l + n$  we have extended cylinders. In order to have  $B_l \neq \emptyset$ , all the relations in Table 1 must be satisfied, and hence

$$B_l = \bigcap_{j=0}^l C_l[s_{l-j}^j] \cap \bigcap_{j=l+1}^{l+n} C_l[\phi^{-1}(s_{l-j}^j)] = C_l[s_l^l]. \tag{11}$$

3.  $(n \leq l \leq k - n - 1)$  In this case, when  $l - n + 1 \leq j \leq l$  we have basic cylinders and when  $l + 1 \leq j \leq l + n$  we have extended cylinders. In order to have  $B_l \neq \emptyset$ , all the relations in Table 1 must be satisfied, and hence

$$B_l = \bigcap_{j=l-n+1}^l C_l[s_{l-j}^j] \cap \bigcap_{j=l+1}^{l+n} C_l[\phi^{-1}(s_{l-j}^j)] = C_l[s_0^l]. \tag{12}$$

4.  $(k - n \leq l \leq k - 2)$  In this case, when  $l - n + 1 \leq j \leq l$  we have basic cylinders and when  $l + 1 \leq j \leq k - 1$  we have extended cylinders. In order to have  $B_l \neq \emptyset$ , all the relations in Table 1 must be satisfied, and hence

$$B_l = \bigcap_{j=l-n+1}^l C_l[s_{l-j}^j] \cap \bigcap_{j=l+1}^{k-1} C_l[\phi^{-1}(s_{l-j}^j)] = C_l[s_0^l]. \tag{13}$$

5.  $(k - 1 \leq l \leq k + n - 2)$  In this case  $l - n + 1 \leq j \leq k - 1$  and no extended cylinder occurs. In order to have  $B_l \neq \emptyset$ , all the relations in Table 1 must be satisfied, and hence

$$B_l = \bigcap_{j=l-n+1}^{k-1} C_l[s_{l-j}^j] = C_l[s_{l-k+1}^{k-1}]. \tag{14}$$

Thus using Equation (10) to (15) on Equation (9), it follows that

$$\bigcap_{j=0}^{k-1} \sigma_\phi^{-j}(\mathcal{C}_{-n, \dots, n-1}[s_{-n}^j, \dots, s_{n-1}^j]) = \bigcap_{l=-n}^{-1} C_l[s_l^0] \cap \bigcap_{l=0}^{k-2} C_l[s_0^l] \cap \bigcap_{l=k-1}^{n-1+k-1} C_l[s_l^{k-1}],$$

that is, a generic element of  $\mathcal{P}_n^k$  is a cylinder of  $\mathcal{C}_{-n, \dots, n+k-2}$ , and every such cylinder can be formed in this way because the symbols  $s_{-n}^0, \dots, s_0^0, \dots, s_0^{k-1}, \dots, s_{n-1}^{k-1}$  can be chosen arbitrarily (see Table 1), so we conclude that  $\mathcal{P}_n^k = \mathcal{C}_{-n, \dots, n+k-2}$ .  $\square$

It is now trivial to conclude the following last results.

**Lemma 4.5**  $h_\mu(\sigma_\phi, \mathcal{P}_n) = H_\mu(\mathcal{C}_0)$ .

**Proof** From Lemmas 4.4 and 4.3 it follows that

$$H_\mu(\mathcal{P}_n^k) = H_\mu(\mathcal{C}_{-n, \dots, n-1+k-1}) = nH_\mu(\mathcal{C}_{-1}) + (n + k - 1)H_\mu(\mathcal{C}_0),$$

therefore

$$\begin{aligned} h_\mu(\sigma_\phi, \mathcal{P}_n) &= \lim_{k \rightarrow \infty} \frac{1}{k} H_\mu(\mathcal{P}_n^k) \\ &= \lim_{k \rightarrow \infty} \frac{1}{k} (nH_\mu(\mathcal{C}_{-1}) + (n + k - 1)H_\mu(\mathcal{C}_0)) \\ &= H_\mu(\mathcal{C}_0). \end{aligned}$$

$\square$

### 4.4 Proof of Theorem A

We are now ready to prove Theorem A.

**Table 1** Relations between the symbols  $s_l^j = s_{l-j}^j$  from Equation (8) for  $-n \leq l \leq n+k-2$ . For each  $l$ , the symbol in bold determines every other symbol in that line

$l$	Relations		
$-n$	$s_{-n}^0$		
$-n+1$	$s_{-n+1}^0 = s_{-n}^1$		
$\vdots$	$\vdots$		
$-1$	$s_{-n}^0 = \dots = s_{-n}^{n-1}$		
$0$	$s_0^0$	$\in$	$\phi^{-1}(s_{-1}^1) = \dots = \phi^{-1}(s_{-n}^n)$
$1$	$s_1^0 = s_0^1$	$\in$	$\phi^{-1}(s_{-1}^2) = \dots = \phi^{-1}(s_{-n}^{n+1})$
$\vdots$	$\vdots$	$\vdots$	
$n-1$	$s_{n-1}^0 = \dots = s_0^{n-1}$	$\in$	$\phi^{-1}(s_{-1}^n) = \dots = \phi^{-1}(s_{-n}^{2n-1})$
$n$	$s_{n-1}^1 = \dots = s_0^n$	$\in$	$\phi^{-1}(s_{-1}^{n+1}) = \dots = \phi^{-1}(s_{-n}^{2n})$
$\vdots$	$\vdots$	$\vdots$	
$k-1-n$	$s_{n-1}^{k-2n} = \dots = s_0^{k-1-n}$	$\in$	$\phi^{-1}(s_{-1}^{k-n}) = \dots = \phi^{-1}(s_{-n}^{k-1})$
$k-n$	$s_{n-1}^{k-2n+1} = \dots = s_0^{k-n}$	$\in$	$\phi^{-1}(s_{-1}^{k-n+1}) = \dots = \phi^{-1}(s_{-(n-1)}^{k-1})$
$\vdots$	$\vdots$	$\vdots$	
$k-2$	$s_{n-1}^{k-1-n} = \dots = s_0^{k-2}$	$\in$	$\phi^{-1}(s_{-1}^{k-1})$
$k-1$	$s_{n-1}^{k-n} = \dots = s_0^{k-1}$		
$k$	$s_{n-1}^{k-n+1} = \dots = s_1^{k-1}$		
$\vdots$	$\vdots$		
$n+k-2$	$s_{n-1}^{k-1}$		

**Proof of Theorem A** The sequence of partitions  $\mathcal{P}_n = \mathcal{C}_{-n, \dots, n-1}$  ( $n \in \mathbb{N}$ ) is increasing relative to the refinement order:

$$\mathcal{P}_0 \leq \mathcal{P}_1 \leq \dots \leq \mathcal{P}_n \leq \dots$$

Besides that, the union of  $\mathcal{P}_n$  generates the  $\sigma$ -algebra of the space  $\Sigma_{m_-, m_+}$ . Finally, the entropy of  $\mathcal{P}_n$  is finite, because the entropy of  $\mathcal{C}_{-1}$  and  $\mathcal{C}_0$  are finite (follows from Lemmas 4.2 and 4.3). Therefore, by the Kolmogorov-Sinai theorem (Theorem 2.1), the metric entropy of the system is

$$h_\mu(\sigma_\phi) = \lim_{n \rightarrow \infty} h_\mu(\sigma_\phi, \mathcal{P}_n).$$

Then Lemma 4.5 implies that

$$h_\mu(\sigma_\phi) = \lim_{n \rightarrow \infty} h_\mu(\sigma_\phi, \mathcal{P}_n) = H_\mu(\mathcal{C}_0),$$

which by Lemma 4.2 equals  $\sum_{s=0}^{m_+-1} -p_s^+ \log p_s^+$ . □

### 5 Folding Entropy of Extended Shifts

Let  $(\Sigma_{m_-, m_+}, \sigma_\phi)$  be an extended shift space. As a consequence of  $\phi$  being surjective, we have that  $m_+ \geq m_-$ . When  $m_+ > m_-$ , the extended shift  $\sigma_\phi$  is not invertible and, for any given  $x \in \Sigma_{m_-, m_+}$ , the set  $\sigma_\phi^{-1}(x)$  has more than one element. In the following discussion, we will need a way to refer to each element of this set  $\sigma_\phi^{-1}(x)$ , so, for each  $s \in \phi^{-1}(x_{-1})$ , we define

$$\hat{x}(s) := (\dots, x_{-2}; s, x_0, \dots). \tag{15}$$

We also denote

$$\hat{x} := \sigma_\phi^{-1}(x) = \{\hat{x}(s) \mid s \in \phi^{-1}(x_{-1})\} \tag{16}$$

and, for each set  $X \subseteq \Sigma_{m_-, m_+}$ ,

$$\hat{X} := \{\hat{x} \mid x \in X\} \subseteq \sigma_\phi^{-1}(\epsilon),$$

in which  $\epsilon = \{\{x\} \mid x \in X\}$  is the atomic partition defined in Equation (6).

This section is dedicated to prove Theorem B. Before proceeding to prove the propositions necessary to complete the proof, we present a proof sketch to guide the reader through the steps.

#### 5.1 Proof Sketch of Theorem B

*Proof sketch of Theorem B* We will first show that

$$\mathcal{F}(\sigma_\phi) = \sum_{s_-=0}^{m_- - 1} \left( \sum_{s_+ \in \phi^{-1}(s_-)} -q_{s_+}^{s_-} \log q_{s_+}^{s_-} \right) p_{s_-}^-.$$

The relation

$$H_\mu(\mathcal{C}_0) - H_\mu(\mathcal{C}_{-1}) = \sum_{s_-=0}^{m_- - 1} \left( \sum_{s_+ \in \phi^{-1}(s_-)} -q_{s_+}^{s_-} \log q_{s_+}^{s_-} \right) p_{s_-}^-$$

will then follow from a straightforward calculation which we present inside the proof itself.

From Definition 3, the folding entropy of  $\sigma_\phi$  is given by

$$\mathcal{F}(\sigma_\phi) = H_\mu(\epsilon \mid \sigma_\phi^{-1}(\epsilon))$$

and, from Equation (5), the conditional entropy of the atomic partition  $\epsilon$  with respect to the dynamical pullback  $\sigma_\phi^{-1}(\epsilon) = \{\hat{x} \mid x \in \Sigma_{m_-, m_+}\}$  can be calculated by

$$H_\mu(\epsilon \mid \sigma_\phi^{-1}(\epsilon)) = \int_{\hat{x} \in \sigma_\phi^{-1}(\epsilon)} H_{\mu_{\hat{x}}}(\epsilon \mid_{\hat{x}}) \hat{\mu}(d\hat{x}),$$

in which  $(\mu_{\hat{x}})_{\hat{x} \in \sigma_\phi^{-1}(\epsilon)}$  is the disintegration of  $\mu$  with respect to  $\sigma_\phi^{-1}(\epsilon)$  and  $\hat{\mu}$  is the quotient measure of  $\sigma_\phi^{-1}(\epsilon)$ .

So in order to calculate the folding entropy of the extended shift, we need to

1. find the quotient measure  $\hat{\mu}$ , which we do in Subsection 5.2 through the simple Propositions 5.1 and 5.2, and

2. disintegrate the measure  $\mu$  with respect to the dynamical pullback  $\sigma_\phi^{-1}(\epsilon)$  of the atomic partition  $\epsilon$  of  $\Sigma_{m_-,m_+}$  (defined in Section 2.4), which we do in Subsection 5.3 by a careful definition of conditional measures  $q_{s_+}^{s_-}$  (Definition 8) that are shown to be a disintegration of  $\mu$  in Proposition 5.3. □

### 5.2 The Quotient Measure

Let us denote the natural projection with respect to the partition  $\sigma_\phi^{-1}(\epsilon)$  by  $\pi : \Sigma_{m_-,m_+} \rightarrow \sigma_\phi^{-1}(\epsilon)$ . Let us first determine the quotient  $\sigma$ -algebra  $\hat{\mathcal{B}}$ , which is the pushforward of the cylinders  $\sigma$ -algebra of  $\Sigma_{m_-,m_+}$  by the natural projection  $\pi$ .

**Proposition 5.1** *For every set  $X \subseteq \Sigma_{m_-,m_+}$ ,*

$$\pi^{-1}(\hat{X}) = \sigma_\phi^{-1}(X).$$

*Besides that, the quotient  $\sigma$ -algebra  $\hat{\mathcal{B}}$  is generated by the projected cylinder sets  $\hat{C}$  ( $C \in \mathcal{B}$  is a cylinder).*

**Proof** The first claim follows directly from

$$\pi^{-1}(\hat{X}) = \bigcup \hat{X} = \bigcup \left\{ \sigma_\phi^{-1}(x) \mid x \in X \right\} = \sigma_\phi^{-1}(X).$$

Now that  $\mathcal{Q} \subseteq \sigma_\phi^{-1}(\epsilon)$ . Since each element of  $\sigma_\phi^{-1}(\epsilon)$  is of the form  $\hat{x}$  for some  $x \in \Sigma_{m_-,m_+}$ , there exists a set  $X \subseteq \Sigma_{m_-,m_+}$  such that  $\mathcal{Q} = \{\hat{x} \mid x \in X\} = \hat{X}$ . This implies that its inverse image by the projection is of the form  $\pi^{-1}(\mathcal{Q}) = \pi^{-1}(\hat{X}) = \sigma_\phi^{-1}(X)$ . This shows that  $\hat{\mathcal{B}}$  is generated by sets  $\hat{C}$  such that  $\sigma_\phi^{-1}(C) \in \mathcal{B}$  is a cylinder, which means that  $C$  is also a cylinder. □

Since  $\hat{x} = \sigma_\phi^{-1}(x)$ , it may be confusing to understand the difference between the sets  $\hat{C}$  and  $\sigma_\phi^{-1}(C)$ . To better understand the notation, it is worth noticing that, if  $x \in C$ , then  $\hat{x} = \sigma_\phi^{-1}(x) \subseteq \sigma_\phi^{-1}(C)$ ; that is, for each  $s \in \phi^{-1}(x_{-1})$ , we have  $\hat{x}(s) \in \sigma_\phi^{-1}(C)$ . This shows that the elements of the set  $\hat{x}$  (which is an element of  $\hat{C}$ ) do not belong to the set  $\hat{C}$ , but instead to  $\sigma_\phi^{-1}(C)$ . To further avoid confusion, consider this example. Suppose  $x, y \in \Sigma_{m_-,m_+}$ ,  $\hat{x} = \{\hat{x}(0), \hat{x}(1)\}$  and  $\hat{y} = \{\hat{y}(0), \hat{y}(1)\}$ . If  $C = \{x, y\}$ , then

$$\hat{C} = \{\hat{x}, \hat{y}\} = \{\{\hat{x}(0), \hat{x}(1)\}, \{\hat{y}(0), \hat{y}(1)\}\},$$

while  $\sigma_\phi^{-1}(C) = \{\hat{x}(0), \hat{x}(1), \hat{y}(0), \hat{y}(1)\}$ .

In particular, it is worth noting that, for a cylinder  $C_i[s]$ ,

$$\pi^{-1}(\hat{C}_i[s]) = \sigma_\phi^{-1}(C_i[s]) = \begin{cases} C_{i+1}[s] & i \neq -1 \\ \bigcup_{s' \in \phi^{-1}(s)} C_0[s'] & i = -1. \end{cases}$$

The quotient measure  $\hat{\mu} := \pi_*\mu$  on  $\sigma_\phi^{-1}(\epsilon)$  is the pushforward of  $\mu$  by the natural projection  $\pi : \Sigma_{m_-,m_+} \rightarrow \sigma_\phi^{-1}(\epsilon)$  of the dynamical pullback of the atomic partition. The next proposition shows how we can easily calculate it using the original measure  $\mu$ .

**Proposition 5.2** *Let  $(\Sigma_{m_-,m_+}, \sigma_\phi)$  be an extended shift space. For every measurable set  $M \subseteq \Sigma_{m_-,m_+}$ ,*

$$\hat{\mu}(\hat{M}) = \mu(M).$$

**Proof** Since  $\pi^{-1}(\hat{M}) = \sigma_\phi^{-1}(M)$  (Proposition 5.1) and  $\sigma_\phi$  is measure-preserving (Proposition 3.2), it follows that

$$\hat{\mu}(\hat{M}) = \mu(\pi^{-1}(\hat{M})) = \mu(\sigma_\phi^{-1}(M)) = \mu(M).$$

□

### 5.3 Disintegration and Calculating the Folding Entropy

We wish to disintegrate the measure  $\mu$  on  $\Sigma_{m_-,m_+}$  with respect to the pullback partition  $\sigma_\phi^{-1}(\epsilon)$ . In order to do that, we must find, for each  $\hat{x} \in \sigma_\phi^{-1}(\epsilon)$ , the conditional measures  $\mu_{\hat{x}}$  on  $\Sigma_{m_-,m_+}$ , in such a way that, for every measurable set  $M \in \mathcal{B}$ , it holds that

$$\mu(M) = \int_{\hat{x} \in \sigma_\phi^{-1}(\epsilon)} \mu_{\hat{x}}(M) \hat{\mu}(d\hat{x}).$$

To define the conditional measures on  $\hat{x}$ , remember that  $\hat{x} = \{\hat{x}(s) \mid s \in \phi^{-1}(x_{-1})\}$  and that the conditional measure is supported on  $\hat{x}$ , so, for each measurable set  $M \in \mathcal{B}$ , it is given by  $\mu_{\hat{x}}(M) = \mu_{\hat{x}}(M \cap \hat{x})$ . Thus, since  $\hat{x}$  is finite, we can define it on each atom  $\{\hat{x}(s)\}$ .

Based on the probability distribution  $p^+$  on  $\{0, \dots, m_+ - 1\}$ , we have described how to induce a probability distribution  $p^-$  on  $\{0, \dots, m_- - 1\}$  by taking the pushforward of  $p^+$  by the transition function  $\phi$ . Using the two measures  $p^+$  on  $\{0, \dots, m_+ - 1\}$  and  $p^-$  on  $\{0, \dots, m_- - 1\}$ , we can define, for each  $s_- \in \{0, \dots, m_- - 1\}$ , a new probability measure  $q^{s_-}$  on the inverse image set  $\phi^{-1}(s_-)$  by setting, for each  $s_+ \in \phi^{-1}(s_-)$

$$q^{s_-}_{s_+} := \frac{p^{s_+}}{p_{s_-}}.$$

This is a probability measure because, for each  $s_- \in \{0, \dots, m_- - 1\}$ ,

$$\sum_{s_+ \in \phi^{-1}(s_-)} q^{s_-}_{s_+} = \sum_{s_+ \in \phi^{-1}(s_-)} \frac{p^{s_+}}{p_{s_-}} = \frac{\sum_{s_+ \in \phi^{-1}(s_-)} p^{s_+}}{p_{s_-}} = 1.$$

As a direct consequence of this definition,

$$p^+ = (p^{s_+}_{s_+})_{s_+=0}^{m_+-1} = ((p_{s_-}^- q^{s_-}_{s_+})_{s_+ \in \phi^{-1}(s_-)})_{s_-=0}^{m_- - 1}.$$

We use these measures  $q^{s_-}$  to define the conditional measures as follows, by identifying the set  $\hat{x}$  with the preimage  $\phi^{-1}(x_{-1})$ .

**Definition 8** Let  $(\Sigma_{m_-,m_+}, \sigma_\phi)$  be an extended shift space with measure  $\mu$  given by the probability distribution  $p^+$ , and let  $\hat{x} \in \sigma_\phi^{-1}(\epsilon)$ . The conditional measure  $\mu_{\hat{x}}$  on  $\hat{x}$  is the probability measure defined, for each  $s \in \phi^{-1}(x_{-1})$ , by

$$\mu_{\hat{x}}(\{\hat{x}(s)\}) := q^{x_{-1}}_s = \frac{p^s}{p_{x_{-1}}}.$$

Now we show this is the disintegration of  $\mu$ .

**Proposition 5.3** Let  $(\Sigma_{m_-,m_+}, \sigma_\phi)$  be an extended shift space. The family  $\{\mu_{\hat{x}}\}_{\hat{x} \in \sigma_\phi^{-1}(\epsilon)}$  is the disintegration of  $\mu$  with respect to  $\sigma_\phi^{-1}(\epsilon)$ .

**Proof** It suffices to show that, for each basic cylinder  $C_s[i]$ , it holds that

$$\mu(C_s[i]) = \int_{\hat{x} \in \sigma_\phi^{-1}(\epsilon)} \mu_{\hat{x}}(C_s[i] \cap \hat{x}) \hat{\mu}(d\hat{x}).$$

First let us calculate the sets  $C_s[i] \cap \hat{x}$ . For any set  $C \subseteq \Sigma_{m-,m+}$ , it holds that  $x \in C$  if, and only if,  $\hat{x} \subseteq \sigma_\phi^{-1}(C)$ . Because of this, we must consider the cases  $\hat{x} \in \widehat{\sigma_\phi(C_s[i])}$  and  $\hat{x} \notin \widehat{\sigma_\phi(C_s[i])}$ ; or equivalently,  $x \in \sigma_\phi(C_s[i])$  and  $x \notin \sigma_\phi(C_s[i])$ . According to Proposition 3.1, the expression for  $\sigma_\phi(C_s[i])$  depends on the value for  $i$ , so we consider 2 scenarios:

- ( $i \neq 0$ ) In this case, we have  $\sigma_\phi(C_s[i]) = C_s[i - 1]$ , hence

$$C_s[i] \cap \hat{x} = \begin{cases} \hat{x} & \hat{x} \in \hat{C}_s[i - 1] \\ \emptyset & \hat{x} \notin \hat{C}_s[i - 1]. \end{cases}$$

Since  $\mu_{\hat{x}}(\hat{x}) = 1$  e  $\mu_{\hat{x}}(\emptyset) = 0$ , it follows that

$$\begin{aligned} \mu(C_i[s]) &= \mu(C_{i-1}[s]) \\ &= \hat{\mu}(\hat{C}_{i-1}[s]) \\ &= \int_{\hat{x} \in \hat{C}_{i-1}[s]} 1 \hat{\mu}(d\hat{x}) + \int_{\hat{x} \in \sigma_\phi^{-1}(\epsilon) \setminus \hat{C}_{i-1}[s]} 0 \hat{\mu}(d\hat{x}) \\ &= \int_{\hat{x} \in \hat{C}_{i-1}[s]} \mu_{\hat{x}}(\hat{x}) \hat{\mu}(d\hat{x}) + \int_{\hat{x} \in \sigma_\phi^{-1}(\epsilon) \setminus \hat{C}_{i-1}[s]} \mu_{\hat{x}}(\emptyset) \hat{\mu}(d\hat{x}) \\ &= \int_{\hat{x} \in \sigma_\phi^{-1}(\epsilon)} \mu_{\hat{x}}(C_i[s] \cap \hat{x}) \hat{\mu}(d\hat{x}). \end{aligned}$$

- ( $i = 0$ ) In this case, we have that

$$C_0[s] \cap \hat{x} = \begin{cases} \{\hat{x}(s)\} & \hat{x} \in \hat{C}_{-1}[\phi(s)] \\ \emptyset & \hat{x} \notin \hat{C}_{-1}[\phi(s)]. \end{cases}$$

Since  $\mu_{\hat{x}}(\{\hat{x}(s)\}) = q_s^{x-1}$  and  $\mu_{\hat{x}}(\emptyset) = 0$  (and, for each  $x \in C_{-1}[\phi(s)]$ , it holds that  $x_{-1} = \phi(s)$ ), it follows that

$$\begin{aligned} \mu(C_0[s]) &= q_s^{\phi(s)} \mu(C_{-1}[\phi(s)]) \\ &= q_s^{\phi(s)} \hat{\mu}(\hat{C}_{-1}[\phi(s)]) \\ &= \int_{\hat{x} \in \hat{C}_{-1}[\phi(s)]} q_s^{x-1} \hat{\mu}(d\hat{x}) + \int_{\hat{x} \in \sigma_\phi^{-1}(\epsilon) \setminus \hat{C}_{-1}[\phi(s)]} 0 \hat{\mu}(d\hat{x}) \\ &= \int_{\hat{x} \in \hat{C}_{-1}[\phi(s)]} \mu_{\hat{x}}(C_0[s] \cap \hat{x}) \hat{\mu}(d\hat{x}) + \int_{\hat{x} \in \sigma_\phi^{-1}(\epsilon) \setminus \hat{C}_{-1}[\phi(s)]} \mu_{\hat{x}}(C_0[s] \cap \hat{x}) \hat{\mu}(d\hat{x}) \\ &= \int_{\hat{x} \in \sigma_\phi^{-1}(\epsilon)} \mu_{\hat{x}}(C_0[s] \cap \hat{x}) \hat{\mu}(d\hat{x}). \end{aligned}$$

□

### 5.4 Proof of Theorem B

We are finally ready to prove our main result on the folding entropy.

**Proof of Theorem B** As discussed in the beginning of the section, it follows from Definition 3 and Equation (5) that the folding entropy of  $\sigma_\phi$  is given by

$$\mathcal{F}(\sigma_\phi) = \int_{\hat{x} \in \sigma_\phi^{-1}(\epsilon)} H_{\mu_{\hat{x}}(\epsilon|\hat{x})} \hat{\mu}(d\hat{x}),$$

in which  $\hat{\mu} = \mu_{\sigma_\phi^{-1}(\epsilon)}$  is the quotient measure of  $\sigma_\phi^{-1}(\epsilon)$ .

Now notice that

$$\epsilon|_{\hat{x}} = \{\{y\} \cap \hat{x} \mid \{y\} \in \epsilon\} = \{\{\hat{x}(s_+)\} \mid s_+ \in \phi^{-1}(x_{-1})\},$$

hence from Equation (2) and Definition 8 it follows that

$$H_{\mu_{\hat{x}}(\epsilon|\hat{x})} = \sum_{s_+ \in \phi^{-1}(x_{-1})} -\mu_{\hat{x}}(\{\hat{x}(s_+)\}) \log \mu_{\hat{x}}(\{\hat{x}(s_+)\}) = \sum_{s_+ \in \phi^{-1}(x_{-1})} -q_{s_+}^{x_{-1}} \log q_{s_+}^{x_{-1}}.$$

This shows that this value depends only on  $x_{-1}$ , so it is constant on each set  $\hat{C}_{-1}[s_-]$ . The set

$$\hat{C}_{-1} := \left\{ \hat{C}_{-1}[s_-] \mid s_- \in \{0, \dots, m_- - 1\} \right\}$$

is a partition of  $\sigma_\phi^{-1}(\epsilon)$ , since (1)  $\hat{C}_{-1}[s_-] \neq \emptyset$ ; (2)  $\hat{C}_{-1}[s_-] \cap \hat{C}_{-1}[r^-] = \emptyset$  when  $s_- \neq r^-$ ; and (3)  $\sigma_\phi^{-1}(\epsilon) = \bigcup_{s_-=0}^{m_- - 1} \hat{C}_{-1}[s_-]$ .

Besides that, it follows from Proposition 5.2 and Definition 5 that  $\hat{\mu}(\hat{C}_{-1}[s_-]) = \mu(C_{-1}[s_-]) = p_{s_-}^-$ . Thus the folding entropy of  $\sigma_\phi$  is

$$\begin{aligned} \mathcal{F}(\sigma_\phi) &= \int_{\hat{x} \in \sigma_\phi^{-1}(\epsilon)} H_{\mu_{\hat{x}}(\epsilon|\hat{x})} \hat{\mu}(d\hat{x}) \\ &= \sum_{s_-=0}^{m_- - 1} \int_{\hat{x} \in \hat{C}_{-1}[s_-]} H_{\mu_{\hat{x}}(\epsilon|\hat{x})} \hat{\mu}(d\hat{x}) \\ &= \sum_{s_-=0}^{m_- - 1} \left( \sum_{s_+ \in \phi^{-1}(s_-)} -q_{s_+}^{s_-} \log q_{s_+}^{s_-} \right) \hat{\mu}(\hat{C}_{-1}[s_-]) \\ &= \sum_{s_-=0}^{m_- - 1} \left( \sum_{s_+ \in \phi^{-1}(s_-)} -q_{s_+}^{s_-} \log q_{s_+}^{s_-} \right) p_{s_-}^-. \end{aligned}$$

Noting that  $q_{s_+}^{s_-} p_{s_-}^- = p_{s_+}^+$  (Definition 8) and  $p_{s_-}^- = \sum_{s_+ \in \phi^{-1}(s_-)} p_{s_+}^+$ , it follows that

$$\begin{aligned}
\mathcal{F}(\sigma_\phi) &= \sum_{s_- = 0}^{m_- - 1} \sum_{s_+ \in \phi^{-1}(s_-)} -q_{s_+}^{s_-} p_{s_-}^- \log q_{s_+}^{s_-} \\
&= \sum_{s_- = 0}^{m_- - 1} \sum_{s_+ \in \phi^{-1}(s_-)} -p_{s_+}^+ (\log p_{s_+}^+ - \log p_{s_-}^-) \\
&= \sum_{s_+ = 0}^{m_+ - 1} -p_{s_+}^+ \log p_{s_+}^+ - \sum_{s_- = 0}^{m_- - 1} - \left( \sum_{s_+ \in \phi^{-1}(s_-)} p_{s_+}^+ \right) \log p_{s_-}^- \\
&= \sum_{s_+ = 0}^{m_+ - 1} -p_{s_+}^+ \log p_{s_+}^+ - \sum_{s_- = 0}^{m_- - 1} -p_{s_-}^- \log p_{s_-}^-.
\end{aligned}$$

Finally, since by Lemma 4.2 we have  $H_\mu(\mathcal{C}_0) = \sum_{s_+ = 0}^{m_+ - 1} -p_{s_+}^+ \log p_{s_+}^+$  and  $H_\mu(\mathcal{C}_{-1}) = \sum_{s_- = 0}^{m_- - 1} -p_{s_-}^- \log p_{s_-}^-$ , we conclude that

$$\mathcal{F}(\sigma_\phi) = H_\mu(\mathcal{C}_0) - H_\mu(\mathcal{C}_{-1}).$$

□

In particular, since the metric entropy is given by  $h_\mu(\sigma_\phi) = H_\mu(\mathcal{C}_0)$ , then

$$h_\mu(\sigma_\phi) = \mathcal{F}(\sigma_\phi) + H_\mu(\mathcal{C}_{-1}).$$

This can be interpreted as showing that the metric entropy is the sum of the entropy of the backwards alphabet with a component of non-invertibility, the folding entropy. In the trivial case  $m_+ = m_-$ , the folding entropy is zero, which shows that the metric entropy is equal to the entropy of the backward alphabet, which in this case equals the forward alphabet.

**Acknowledgements** The authors would like to thank Pouya Mehdipour for her comments. N. M. was partially financed by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior Brasil (CAPES) - grant 88887.645688/2021-00. P. M. was partially financed by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior Brasil (CAPES) - grant 141401/2020-6. R. V. was partially supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (grants 313947/2020-1 and 314978/2023-2), and partially supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) (grants 17/06463-3 and 18/13481-0).

**Author Contributions** All authors wrote the manuscript.

**Data Availability** No datasets were generated or analysed during the current study.

## Declarations

**Competing interests** The authors declare no competing interests.

## References

1. Anosov, D.V.: Geodesic flows on closed Riemannian manifolds of negative curvature. In: Trudy Matematicheskogo Instituta imeni V. A. Steklova 90 (1967). English translation in \*Proc. Steklov Inst. Math.\* **90** (1967), 1–235, 3–210
2. Ashley, J., Marcus, B., Tuncel, S.: The classification of one-sided Markov chains. In: Ergodic Theory and Dynamical Systems 17.2, 269–295 (1997)
3. Bowen, R.: Markov Partitions for Axiom A Diffeomorphisms. In: American Journal of Mathematics 92.3, 725–747 (1970)

4. Bruin, H., Hawkins, J.: Rigidity of smooth one-sided Bernoulli endomorphisms. *The New York Journal of Mathematics* [electronic only] **15**, 451–483 (2009)
5. Downarowicz, T.: *Entropy in dynamical systems*, vol. 18. *New Mathematical Monographs*. Cambridge University Press, Cambridge (2011)
6. Hadamard, J.: Les surfaces à courbures opposées et leurs lignes géodésiques. In: *Journal de Mathématiques Pures et Appliquées* 5th ser. **4**, 27–73 (1898)
7. Hoffman, C., Rudolph, D.: A dyadic endomorphism which is Bernoulli but not standard. English. In: *Israel J. Math.* **130**, 365–379 (2002)
8. Hoffman, C., Rudolph, D.: Uniform Endomorphisms Which Are Isomorphic to a Bernoulli Shift. In: *Annals of Mathematics* **156**(1), 79–101 (2002)
9. Katznelson, Y.: Ergodic automorphisms of  $T^n$  are Bernoulli shifts. *Israel J. Math.* **10**, 186–195 (1971)
10. Kolmogorov, A.N.: A New Metric Invariant of Transitive Dynamical Systems and Endomorphisms of Lebesgue Spaces. In: *Doklady Akademii Nauk SSSR* **119**(5), 861–864 (1958)
11. Kosloff, Z., Soo, T.: Sinai factors of nonsingular systems: Bernoulli shifts and Anosov flows. In: *Journal of Modern Dynamics* **20**(0), 597–634 (2024)
12. Lamei, S., Mehdipour, P.: An  $N$ -to-1 Smale Horseshoe. Preprint. (2025). [arXiv:2502.14441](https://arxiv.org/abs/2502.14441)
13. Lamei, S., Mehdipour, P.: Zip shift Space. Preprint. (2025). [arXiv:2502.11272](https://arxiv.org/abs/2502.11272)
14. Liao, G., Wang, S.: Continuity properties of folding entropy. In: *Israel Journal of Mathematics* (2024)
15. Liu, P.-D.: Ruelle Inequality Relating Entropy, Folding Entropy and Negative Lyapunov Exponents. *Commun. Math. Phys.* **240**, 531–538 (2003)
16. Mehdipour, P., Martins, N.: Encoding  $n$ -to-1 baker’s transformations. *Arch. Math.* **119**, 199–211 (2022)
17. Ornstein, D.: Bernoulli shifts with the same entropy are isomorphic. *Adv. Math.* **4**, 337–352 (1970)
18. Ornstein, D., Weiss, B.: Geodesic flows are Bernoullian. *Israel J. Math.* **14**, 184–198 (1973)
19. Przytycki, F.: Anosov endomorphisms. *Studia Math.* **58**(3), 249–285 (1976)
20. Ratner, M.: Anosov flows with Gibbs measures are also Bernoullian. *Israel J. Math.* **17**, 380–391 (1974)
21. Rokhlin, V.A.: Lectures on the Entropy Theory of Measure-Preserving Transformations. In: *Russian Mathematical Surveys* **22**(5), 1 (Oct. 1967)
22. Ruelle, D.: Positivity of entropy production in nonequilibrium statistical mechanics. *J. Stat. Phys.* **85**, 1–23 (1996)
23. Sinai, Y.G.: Classical dynamic systems with countably-multiple Lebesgue spectrum. II. In: *Izvestiya Mathematics* **30**(1) (1966)
24. Sinai, Y.G.: On the Notion of Entropy of a Dynamical System. *Dokl. Akad. Nauk SSSR* **124**, 768–771 (1959)
25. Viana, M., Oliveira, K.: *Foundations of Ergodic Theory*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, (2016)
26. Wu, W., Zhu, Y.: On preimage entropy, folding entropy and stable entropy. *Ergod. Th. & Dynam. Sys.* **41**, 1217–1249 (2021)
27. Wu, W., Zhu, Y.: Entropy via preimage structure. In: *Advances in Mathematics* **406**, 108483 (2022)

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.