Article

Entropy Increase in Switching Systems

José M. Amigó [1], **Peter E. Kloeden** [2,*] **and Ángel Giménez** [1]

[1] Centro de Investigación Operativa, Universidad Miguel Hernández. Avda. de la Universidad s/n, Elche 03202, Spain; E-Mails: jm.amigo@umh.es (J.M.A.); a.gimenez@umh.es (Á.G.)

[2] Fachbereich Mathematik, Johan Wolfgang Goethe Universität. Frankfurt am Main 60054, Germany

* Author to whom correspondence should be addressed; E-Mail: kloeden@math.uni-frankfurt.de; Tel.: +49-069-798-28622; Fax: +49-069-798-28846.

Abstract: The relation between the complexity of a time-switched dynamics and the complexity of its control sequence depends critically on the concept of a non-autonomous pullback attractor. For instance, the switched dynamics associated with scalar dissipative affine maps has a pullback attractor consisting of singleton component sets. This entails that the complexity of the control sequence and switched dynamics, as quantified by the topological entropy, coincide. In this paper we extend the previous framework to pullback attractors with nontrivial components sets in order to gain further insights in that relation. This calls, in particular, for distinguishing two distinct contributions to the complexity of the switched dynamics. One proceeds from trajectory segments connecting different component sets of the attractor; the other contribution proceeds from trajectory segments within the component sets. We call them "macroscopic" and "microscopic" complexity, respectively, because only the first one can be measured by our analytical tools. As a result of this picture, we obtain sufficient conditions for a switching system to be more complex than its unswitched subsystems, *i.e.*, a complexity analogue of Parrondo's paradox.

1. Introduction

The time-switched dynamics of two one-dimensional, dissipative affine maps,

$$f_\pm(x) = \theta_\pm x \pm 1 \ \ 0 < \theta_-, \theta_+ < 1, \ \theta_- \neq \theta_+ \tag{1}$$

was studied by the authors in [1], where they showed that the topological entropy of the resulting non-autonomous dynamical system coincides with the topological entropy of the autonomous dynamical system generating the switching (or control) sequence. In other words, the switched dynamics cannot be more complex than the switching sequence generator.

The situation envisaged in [1] is special in several regards.

(i) The state space is $\mathbb{R}$.

(ii) Both the forward and backward dynamics are amenable to detailed analysis.

(iii) The non-autonomous pullback attractor consists of singleton component sets.

In this follow-up paper we continue studying this question in an extended setting, namely, we consider this time switching between difference equations (called constituent maps) that have pullback attractors with nontrivial component sets. Our scope is to separate the particular results from the general ones, and so better understand the intricacies of switching and non-autonomous dynamics [2–4]. Specifically, in this paper:

(i') The state space is $\mathbb{R}^d$ or a proper subset of it.

(ii') The constituent maps are not specified, except for the fact that they are supposed to have attractors.

(iii') The component sets of the pullback attractor of the systems under switching are supposed to be uniformly bounded.

Both from an instrumental and a conceptual point of view, the main challenge is introduced by the generalization (**iii'**). To begin with, working with attractors with nontrivial component sets instead of point-valued components sets requires using the Hausdorff distance. Contrarily to what happens in the latter case, the Hausdorff distance between component sets of a set-valued attractor is not continuous, in general, but only upper-continuous when a control sequence converges to another one [2]. This technical shortcoming is behind Assumption 1 in Section 3 ("Results using measurability"), and Assumption 3 in Section 5 ("Results using continuity"). Also the generality of (**ii'**) will be limited in different ways (injectivity of a certain map Φ in Section 3, or Assumption 2 in Section 4) in order to derive sharper results.

On the conceptual side, new features related to the extended geometry of set-valued attractors manifest themselves. Indeed, the observable complexity is going to be the result of two contributions. One, the "macroscopic" complexity, comes from the trajectory segments connecting different component sets of the pullback attractor (think that each component set has been lumped to a point). The other one, the "microscopic" complexity, comes from trajectory segments within the component sets. No geometrical meaning should be attached to this denomination since the (in general, uncountably many) component

sets might be packed in a state space region of a size comparable to the attractor itself. The rationale for distinguishing two sorts of complexity is rather that our analytical tools will only be able to resolve the trajectories up to the precision set by the component sets. An offshoot of this picture is that the entropy of the coarse-grained dynamics (or "macroscopic entropy") is a lower bound of the entropy of the switched dynamics. This fact will again allow us to relate the entropy of trajectories and control sequences, once we have shown that the latter coincide with the macroscopic entropy.

Thus, the main purpose of the present paper is the study of the macroscopic complexity in switched dynamics, as measured by the macroscopic topological entropy. In Sections 3 and 5 we prove that, under certain provisos, the macroscopic topological entropy coincides with the topological entropy of the switching sequence generator. In this case, the complexity of the switched dynamics (a mixture of macroscopic and microscopic complexities) is certainly greater than the complexity of the switching sequence. This result can be brought in connection with Parrondo's paradox, *i.e.*, the emergence of new properties via switching [5–8]. Indeed, it suffices that the entropy of the switching sequence generator is higher than the entropies of the constituent maps for the switched dynamics to be more complex than the constituent dynamics.

This paper is organized as follows. Section 2 sets the mathematical framework (switching systems, Hausdorff distance and pullback attractors) of the paper. Additional materials on the Hausdorff distance have been collected in the Appendix for the reader's convenience. A few results for further reference, but also interesting on their own, are proved as well in this section and in the Appendix. The main theoretical results of the paper are derived in Section 3 (Theorem 3) and Section 5 (Theorem 4) depending on whether the switched dynamics complies with Assumption 1 (Section 3), or with Assumption 3 (Section 5). In between, Section 4 scrutinizes a property needed in Theorem 3 and reformulates it as Assumption 2 to be used in the proof of Theorem 4. Some numerical simulations illustrating our theoretical results make up Section 6. The highlights of all these sections are summarized in the Conclusion, followed by the references and the Appendix.

2. Preliminaries

This section introduces the background on switching and non-autonomous dynamical systems needed to make this paper self-contained. The interested readers are referred to the books [2,3], the review [4], and the papers [9–13].

2.1. Switching Systems

Let $\mathbf{s} = (\ldots, s_{-2}, s_{-1}, s_0, s_1, s_2, \ldots) \in \mathcal{S} := \{-1, 1\}^{\mathbb{Z}}$ be the collection of all switching controls, which is a compact metric space with the metric

$$\text{dist}_{\mathcal{S}}(\mathbf{s}, \mathbf{s}') = \sum_{n \in \mathbb{Z}} 2^{-|n|} |s_n - s'_n| \tag{2}$$

In calculations though, the following, equivalent distance might be more convenient.

$$\text{dist}_{\mathcal{S}}(\mathbf{s}, \mathbf{s}') = \begin{cases} 0 & \text{if } s_n = s'_n \text{ for all } n \in \mathbb{Z}, \\ 2^{-N} & \text{if } s_n = s'_n \text{ for } |n| < N, \text{ while } s_N \neq s'_N \text{ and/or } s_{-N} \neq s'_{-N} \end{cases} \tag{3}$$

The metric space $(\mathcal{S}, \operatorname{dist}_{\mathcal{S}})$ is compact; see, e.g., [14] (p. 207).

Consider the left shift operator σ defined on $\mathcal{S}$,

$$\sigma : (\cdots, s_n, s_{n+1}, \cdots) \mapsto (\cdots, s_{n+1}, s_{n+2}, \cdots)$$

Then σ is continuous with respect to the product topology on $\mathcal{S}$, *i.e.*, generates an autonomous dynamical system (group under composition) on $\mathcal{S}$.

We consider discrete time dynamical systems generated by switching between two continuous mappings $f_{\pm 1} : \mathbb{R}^d \to \mathbb{R}^d$, where the autonomous dynamical systems generated by each mapping are dissipative and have global attractors $A^{\pm}$. In other words, the (time) switched dynamics (or switching system) we are going to study in the following is described by the non-autonomous difference equation

$$x_{n+1} = f_{s_n}(x_n) \tag{4}$$

in $\mathbb{R}^d$ for different sequences $\mathbf{s} = (s_n)_{n \in \mathbb{Z}} \in \mathcal{S}$. Sometimes one says that $f_{\pm 1}$ are the constituent maps of the switching system (4). For brevity, we shorten $f_{\pm 1}$ to $f_{\pm}$, $x_{\pm 1}$ to $x_{\pm}$, *etc*.

As an important example, let $f_{\pm 1}$ be affine maps,

$$f_{\pm}(x) = M_{\pm} x + b_{\pm} \tag{5}$$

where $M_{\pm}$ are $d \times d$-matrices, and $x, b_{\pm}$ are column vectors. Linear systems correspond to $b_{\pm} = 0$. Then

$$x_{n+k} = \prod_{i=0}^{k-1} M_{s_{n+i}} x_n + \prod_{i=1}^{k-1} M_{s_{n+i}} b_{s_n} + \prod_{i=2}^{k-1} M_{s_{n+i}} b_{s_{n+1}} \tag{6}$$

$$\cdots + M_{s_{n+k-1}} b_{s_{n+k-2}} + b_{s_{n+k-1}}$$

for $k \geq 1$. The order in the matrix products $\prod_{i=0}^{k-1} M_{s_{n+i}}$, *etc*., is right to left as the index i increases, *i.e.*,

$$\prod_{i=0}^{k-1} M_{s_{n+i}} = M_{s_{n+k-1}} \cdots M_{s_{n+1}} M_{s_n} \tag{7}$$

Likewise if the matrices $M_{\pm}$ are invertible, then from

$$x_{n-1} = f_{s_{n-1}}^{-1}(x_n) = M_{s_{n-1}}^{-1}(x_n - b_{s_{n-1}}) \tag{8}$$

we get

$$x_{n-k} = \prod_{i=1}^{k} M_{s_{n-i}}^{-1} x_n - \prod_{i=1}^{k} M_{s_{n-i}}^{-1} b_{s_{n-1}} - \prod_{i=2}^{k} M_{s_{n-i}}^{-1} b_{s_{n-2}} \tag{9}$$

$$\cdots - M_{s_{n-k}}^{-1} b_{s_{n-k}}$$

The same rule as in Equation (7) applies here to the products of inverse matrices.

As usual, let $|x|$ denote the norm of $x \in \mathbb{R}^d$ and $\|M\|$ a compatible norm of the $d \times d$ matrix M.

Proposition 1 *Consider the affine switched dynamics Equations (4) and (5). Furthermore, suppose that* $\mu := \max\{\|M_-\|, \|M_+\|\} < 1$, *and* $\beta := \max\{|b_+|, |b_-|\}$. *Then*

$$\limsup_{n \to \infty} |x_n| \leq \frac{\beta}{1 - \mu} \tag{10}$$

for any $x_0 \in \mathbb{R}$.

Proof. From Equation (6), we have

$$|x_{n+k}| \leq \mu^k |x_n| + (\mu^{k-1} + \mu^{k-2} + \cdots + \mu + 1)\beta$$

$$= \mu^k |x_n| + \frac{1 - \mu^k}{1 - \mu}\beta$$

where $\mu < 1$. Let now $k \to \infty$. $\square$

Linear systems with $\|M_-\|, \|M_+\| < 1$ are globally dissipative. The origin is a fixed point and, according to Equation (10) with $\beta = 0$, all other orbits converge to it for each $\mathbf{s} \in \mathcal{S}$.

2.2. Hausdorff Metric and Sequences of Compact Subsets

Let $\mathcal{K}$ be the space of nonempty compact subsets of $\mathbb{R}^d$, which is a complete metric space with the Hausdorff metric

$$\text{dist}_H(A, B) := \max\{\rho(A, B), \rho(B, A)\} \tag{11}$$

where $\rho(A, B)$ is the Hausdorff semi-distance defined by

$$\rho(A, B) := \max_{a \in A} \text{dist}(a, B) \qquad \text{dist}(a, B) := \min_{b \in B} |a - b|$$

At variance with $\text{dist}_H(A, B)$, $\rho(A, B)$ is not a metric because it is not always symmetric, and $\rho(A, B) = 0$ only implies $A \subset B$. Intuitively speaking, for the distance between A and B to be small, both sets have to almost overlap, *i.e.*, the difference set $A \Delta B = A \backslash B \cup B \backslash A$ has to be "small". Further results on the Hausdorff metric can be found in the Appendix.

Let $B_R(a)$ and $\bar{B}_R(a)$ be the open and closed balls of radius $R \geq 0$ and center $a \in \mathbb{R}^d$, respectively. Define $\mathcal{K}_R = \mathcal{K} \cap \bar{B}_R(0)$, *i.e.*, the family of nonempty compact subsets of $\mathbb{R}^d$ that are contained in $\bar{B}_R(0)$. Then, $(\mathcal{K}_R, \text{dist}_H)$ is a compact metric space.

The next proposition will be needed in Section 5. It is a particular case of Lemma A2, stated and proved in the Appendix.

Proposition 2 *Let $f : \bar{B}_R(0) \to \mathbb{R}^d$ be continuous, and $A, B \in \mathcal{K}_R$. Then*

$$\text{dist}_H(f(A), f(B)) \leq \omega(\text{dist}_H(A, B)), \tag{12}$$

where $\omega(\cdot)$ is a continuous function with $\omega(0) = 0$.

In order to apply Proposition 2 to our constituent maps $f_\pm$, whose definition domain is in principle all $\mathbb{R}^d$, it suffices to consider their restrictions to $\bar{B}_R(0)$.

Further, we define a metric on the space $\mathfrak{K}_R := \mathcal{K}_R{}^{\mathbb{Z}}$ of bi-infinite sequences $\mathbf{A} := (A_n)_{n \in \mathbb{Z}}$ of nonempty compact subsets of $\mathbb{R}^d$ in $\bar{B}_R(0)$ by

$$\text{dist}_{\mathfrak{K}_R}(\mathbf{A}, \mathbf{A}') := \sum_{n \in \mathbb{Z}} 2^{-|n|}\text{dist}_H(A_n, A'_n) \tag{13}$$

for $\mathbf{A} = (A_n)_{n \in \mathbb{Z}}$, and $\mathbf{A}' = (A'_n)_{n \in \mathbb{Z}}$.

2.3. Skew Product Flows and Pullback Attractors

Define $\varphi : \mathbb{N}_0 \times \mathcal{S} \times \mathbb{R}^d \to \mathbb{R}^d$ by $\varphi(0, \mathbf{s}, x_0) = x_0$ and

$$\varphi(n, \mathbf{s}, x_0) := f_{s_{n-1}} \circ \cdots \circ f_{s_1} \circ f_{s_0}(x_0) \qquad n \geq 1$$

Then φ is a continuous cocycle mapping with respect to σ, *i.e.*,

$$\varphi(n + k, \mathbf{s}, x_0) = \varphi(n, \sigma^k \mathbf{s}, \varphi(k, \mathbf{s}, x_0)) \tag{14}$$

for all $n, k \geq 0$, and (σ, φ) is a discrete time skew product flow (non-autonomous dynamical system) on $\mathcal{S} \times \mathbb{R}^d$.

An *entire solution* of a discrete time skew product flow (σ, φ) is a mapping $\chi : \mathcal{S} \to \mathbb{R}^d$ such that

$$\chi(\sigma^n \mathbf{s}) = \varphi(n - k, \sigma^k \mathbf{s}, \chi(\sigma^k \mathbf{s})) \tag{15}$$

for all $\mathbf{s} \in \mathcal{S}$ and $n, k \in \mathbb{Z}$ with $k \leq n$. In particular,

$$\chi(\sigma^n \mathbf{s}) = \varphi(n, \mathbf{s}, \chi(\mathbf{s})) \text{ for all } n \geq 0 \tag{16}$$

A *pullback attractor* is a family of nonempty compact subsets, $\mathfrak{A} = \{A(\mathbf{s}), \mathbf{s} \in \mathcal{S}\} \subset \mathcal{K}$ which is φ-invariant, *i.e.*,

$$\varphi(n, \mathbf{s}, A(\mathbf{s})) = A(\sigma^n \mathbf{s}) \qquad n \geq 0 \tag{17}$$

and pullback attracts, *i.e.*,

$$\mathrm{dist}_H \left(\varphi(n, \sigma^{-n} \mathbf{s}, D), A(\mathbf{s}) \right) \to 0 \qquad \text{for } n \to \infty \tag{18}$$

for every nonempty bounded subset $D \subset \mathbb{R}^d$. The $A(\mathbf{s})$ are called the *component sets* of the attractor $\mathfrak{A}$.

We will assume that the switched dynamics (4) has a pullback attractor $\mathfrak{A} = \{A(\mathbf{s}) : \mathbf{s} \in \mathcal{S}\}$ such that the $A(\mathbf{s})$ are nonempty, uniformly bounded compact subsets of $\mathbb{R}^d$. This means that there is $R > 0$ such that $A(\mathbf{s}) \subset \bar{B}_R(0)$ for every $\mathbf{s} \in \mathcal{S}$. If the constituent maps are both affine, then it is inferred from Proposition 1 that one may choose $R = \frac{\beta}{1-\mu}$. A sufficient condition ensuring the existence of such a pullback attractor is that the two unswitched systems have a common bounded, positively invariant absorbing set.

Proposition 3 *[2](Theorem 3.34) The set-valued mapping* $\mathbf{s} \mapsto A(\mathbf{s})$ *is upper semi-continuous in* $(\mathcal{K}_R, \mathrm{dist}_H)$, i.e.,

$$\rho\left(A(\mathbf{s}), A(\mathbf{s}^*)\right) \to 0 \quad as \quad \mathrm{dist}_{\mathcal{S}}(\mathbf{s}, \mathbf{s}^*) \to 0$$

Counterexamples show that, in general, we cannot replace the Hausdorff semi-distance here by the Hausdorff metric, but in special cases we can do that, e.g., when the pullback attractor consists of singleton sets as in [1] (see Example 1 below), since then obviously

$$\rho\left(A(\mathbf{s}), A(\mathbf{s}^*)\right) = \mathrm{dist}_H\left(A(\mathbf{s}), A(\mathbf{s}^*)\right) = |A(\mathbf{s}) - A(\mathbf{s}^*)|$$

See Case 2 of the numerical simulations (Section 4.2) for other example.

A general, sufficient condition for the continuity of the map $\mathbf{s} \mapsto A(\mathbf{s})$ in $(\mathcal{K}_R, \mathrm{dist}_H)$ is provided by the following result.

Proposition 4 *Suppose that* $\mathrm{dist}_H\left(\varphi(n, \sigma^{-n}\mathbf{s}, D), A(\mathbf{s})\right) \to 0$ *uniformly in* $\mathbf{s}$ *for some nonempty, bounded set* $D \subset \mathbb{R}^d$. *Then the map* $\mathbf{s} \mapsto A(\mathbf{s})$ *is continuous in* $(\mathcal{K}_R, \mathrm{dist}_H)$, *i.e.,*

$$\mathrm{dist}_H\left(A(\mathbf{s}), A(\mathbf{s}^*)\right) \to 0 \quad as \quad \mathrm{dist}_{\mathcal{S}}(\mathbf{s}, \mathbf{s}^*) \to 0$$

Proof. By the property (18) of pullback attractors, for all $\varepsilon > 0$ there exists an N such that

$$\mathrm{dist}_H\left(\varphi(n, \sigma^{-n}\mathbf{s}, D), A(\mathbf{s})\right) < \frac{\varepsilon}{2} \quad \mathrm{dist}_H\left(\varphi(n, \sigma^{-n}\mathbf{s}^*, D), A(\mathbf{s}^*)\right) < \frac{\varepsilon}{2} \tag{19}$$

for all $n \geq N$, where $D \subset \mathbb{R}^d$ is bounded and fixed for the time being. By hypothesis, N depends on ε but not on either $\mathbf{s}$ or $\mathbf{s}^*$. The triangle inequality then yields

$$\mathrm{dist}_H(A(\mathbf{s}), A(\mathbf{s}^*)) \leq \mathrm{dist}_H\left(A(\mathbf{s}), \varphi(N, \sigma^{-N}\mathbf{s}, D)\right)$$

$$+ \mathrm{dist}_H\left(\varphi(N, \sigma^{-N}\mathbf{s}, D), \varphi(N, \sigma^{-N}\mathbf{s}^*, D)\right)$$

$$+ \mathrm{dist}_H\left(\varphi(N, \sigma^{-N}\mathbf{s}^*, D), A(\mathbf{s}^*)\right)$$

$$< \varepsilon + \mathrm{dist}_H\left(\varphi(N, \sigma^{-N}\mathbf{s}, D), \varphi(N, \sigma^{-N}\mathbf{s}^*, D)\right) \tag{20}$$

where

$$\varphi(N, \sigma^{-N}\mathbf{s}, D) = f_{s_{-1}} \circ \cdots \circ f_{s_{-N+1}} \circ f_{s_{-N}}(D)$$

$$\varphi(N, \sigma^{-N}\mathbf{s}^*, D) = f_{s_{-1}^*} \circ \cdots \circ f_{s_{-N+1}^*} \circ f_{s_{-N}^*}(D)$$

Take now $\mathrm{dist}_{\mathcal{S}}(\mathbf{s}, \mathbf{s}^*) < 2^{-(N+1)}$ so as $s_k = s_k^*$ for $0 \leq |k| \leq N$, hence $\varphi(N, \sigma^{-N}\mathbf{s}, D) = \varphi(N, \sigma^{-N}\mathbf{s}^*, D)$, and

$$\mathrm{dist}_H\left(\varphi(N, \sigma^{-N}\mathbf{s}, D), \varphi(N, \sigma^{-N}\mathbf{s}^*, D)\right) = 0$$

We conclude from Equation (20) that if $\mathrm{dist}_{\mathcal{S}}(\mathbf{s}, \mathbf{s}^*) < 2^{-(N+1)}$ [where $N = N(\varepsilon)$ is such that Equation (19) holds], then

$$\mathrm{dist}_H(A(\mathbf{s}), A(\mathbf{s}^*)) < \varepsilon$$

for all $\varepsilon > 0$. $\qquad\qquad\square$

Remark 1 *It follows from Proposition 3 that the pullback attractor is also a forward attractor, i.e., with*

$$\mathrm{dist}_H\left(\varphi(n, \mathbf{s}, D), A(\sigma^n\mathbf{s})\right) \to 0 \; as \; n \to \infty$$

3. Results Using Measurability

Remember that $\mathfrak{K}_R = \mathcal{K}_R^{\mathbb{Z}}$ is endowed with the metric (13). Let $\Phi : \mathcal{S} \to \mathfrak{K}_R$ be the map defined by $\Phi(\mathbf{s}) = (A(\sigma^n\mathbf{s}))_{n \in \mathbb{Z}}$. We show first that Φ is Borel measurable.

By definition the Borel sigma-algebras of the product spaces $\mathcal{S}$ and $\mathfrak{K}_R$ are generated by the corresponding cylinder sets or just cylinders. If $p, q \in \mathbb{Z}$, $p \leq q$, the cylinders of $\mathcal{S} = \{-1, +1\}^{\mathbb{Z}}$ have the form

$$[a_p, ..., a_q]_p^q = \{(s_n)_{n \in \mathbb{Z}} \in \mathcal{S} : s_n = a_n \text{ for } p \leq l \leq q\}$$

where $a_p, ..., a_q \in \{-1, +1\}$, and the cylinders of $\mathfrak{K}_R$ have the form

$$[K_p, ..., K_q]_p^q = \{(A_n)_{n \in \mathbb{Z}} \in \mathfrak{K}_R : A_n = K_n \text{ for } p \le n \le q\}$$

where $K_p, ...K_q \in \mathcal{K}_R$. If $p = q$, then we simplify the notation to $[a_p]_p$ or $[a]_p$, and analogously for the cylinders of $\mathfrak{K}_R$.

If, as before, σ is the (left) shift on $\mathcal{S}$, and Σ is the shift on $\mathfrak{K}_R$, then $\sigma^n [a_p, ..., a_q]_p^q = [a_p, ..., a_q]_{p-n}^{q-n}$ and, similarly, $\Sigma^n [K_p, ..., K_q]_p^q = [K_p, ..., K_q]_{p-n}^{q-n}$ for all $n \in \mathbb{Z}$.

Lemma 1 $\Phi : \mathcal{S} \to \mathfrak{K}_R$ *is Borel measurable.*

Proof. Since the cylinder sets build a semi-algebra of the product sigma-algebras they generate, it suffices to prove that $\Phi^{-1}[K_p, ..., K_q]_p^q$ is Borel measurable for every $p, q \in \mathbb{Z}$, $p \le q$, and every $K_p, ...K_q \in \mathcal{K}_R$ (see Theorem 1.1 in [15]). Note that $\Phi^{-1}[K_p, ..., K_q]_p^q = \emptyset$ unless [see Equation (17)] $K_p = A(\mathbf{s})$ for some $\mathbf{s} \in \mathcal{S}$, and $K_{p+n} = A(\sigma^n \mathbf{s}) = \varphi(n, \mathbf{s}, K_p)$ for $n = 1, ..., q - p$. In any case,

$$\begin{aligned} \Phi^{-1}[K_p, ..., K_q]_p^q &= \Phi^{-1}([K_p]_p \cap [K_{p+1}]_{p+1} \cap ... \cap [K_q]_q) \\ &= \Phi^{-1}[K_p]_p \cap \Phi^{-1}[K_{p+1}]_{p+1} \cap .... \cap \Phi^{-1}[K_q]_q \end{aligned}$$

Since the map $\mathbf{s} \mapsto A(\mathbf{s})$ is upper semi-continuous (hence Borel measurable), the sets $\Phi^{-1}[K_{p+n}]_{p+n}$, $0 \le n \le q - p$, are Borel measurable and so is their intersection as well. $\square$

Suppose now that Φ is one-to-one and let us explore when Φ^{-1} is also Borel measurable, resulting in a Borel bimeasurable mapping. For $p, q \in \mathbb{Z}$, $p \le q$,

$$\Phi[a_p, ..., a_q]_p^q = (A(\sigma^n [a_p, ..., a_q]_p^q))_{n \in \mathbb{Z}} \tag{21}$$

where

$$\begin{aligned} A(\sigma^n [a_p, ..., a_q]_p^q) &= \{A(\mathbf{s}) : \mathbf{s} \in \sigma^n [a_p, ..., a_q]_p^q\} \\ &= \{A(\mathbf{s}) : \mathbf{s} \in [a_p, ..., a_q]_{p-n}^{q-n}\} \\ &= \{A(\mathbf{s}) : \mathbf{s} \in [a_p]_{p-n} \cap [a_{p+1}]_{p-n+1} \cap ... \cap [a_q]_{q-n}\} \\ &= \{A(\mathbf{s}) : \mathbf{s} \in [a_p]_{p-n}\} \cap ... \cap \{A(\mathbf{s}) : \mathbf{s} \in [a_q]_{q-n}\} \\ &= A([a_p]_{p-n}) \cap ... \cap A([a_q]_{q-n}) \tag{22} \end{aligned}$$

Note that

$$A([a]_p) = \{A(\mathbf{s}) : s_p = a\}$$

is an uncountable union of compact sets.

Assumption 1. *The sets $A([+1]_p)$ and $A([-1]_p)$ are Borel measurable for every $p \in \mathbb{Z}$.*

Continuity or closedness of the mapping $\mathbf{s} \mapsto A(\mathbf{s})$ are obvious sufficient conditions for Assumption 1 to hold. (As a matter of fact, by the closed map lemma, if $\mathbf{s} \mapsto A(\mathbf{s})$ is continuous then it is closed because $\mathcal{S}$ is compact and $\mathcal{K}_R$ is a Hausdorff topological space.) For example, in the affine, one-dimensional case studied in [1] $A(\mathbf{s})$ is a singleton, *i.e.*, $A(\mathbf{s}) = \{\chi(\mathbf{s})\}$, for all $\mathbf{s} \in \mathcal{S}$. That all the sets $A([a]_p)$ are compact (hence Borel measurable) follows in this case from the continuity of $\mathbf{s} \mapsto \chi(\mathbf{s})$. In Section 5 we will study with more detail the consequences of assuming the mapping $\mathbf{s} \mapsto A(\mathbf{s})$ continuous.

Lemma 2 *If* $\Phi : \mathbf{s} \mapsto (A(\sigma^n \mathbf{s}))_{n \in \mathbb{Z}}$ *is one-to-one and Assumption 1 holds, then* Φ *is Borel bimeasurable.*

Proof. We only need to prove that Φ transforms cylinders into Borel measurable sets. In view of Equation (21) this boils down to showing that $A(\sigma^n[a_p, ..., a_q]_p^q)$ is Borel measurable for every $p, q \in \mathbb{Z}$, $p \leq q$.

From Equation (22) and Assumption 1, it follows that $A(\sigma^n[a_p, ..., a_q]_p^q)$ is a finite intersection of Borel measurable sets, hence it is Borel measurable. $\square$

Consider the diagram

$$
\begin{array}{ccc}
\mathcal{S} & \overset{\sigma}{\to} & \mathcal{S} \\
\Phi \downarrow & & \downarrow \Phi \\
\mathfrak{K}_R & \overset{\Sigma}{\to} & \mathfrak{K}_R
\end{array}
\tag{23}
$$

then it is straightforward to check that this diagram commutes, *i.e.*, $\Phi \circ \sigma = \Sigma \circ \Phi$. In order to derive that $h(\sigma) = h(\Sigma|_{\Phi(\mathcal{S})})$, it suffices that $\Phi : \mathcal{S} \to \Phi(\mathcal{S})$ is a bimeasurable bijection (see Corollary 8.6.1 (iv) in [15]). This being the case, Lemma 3.2 yields the following result.

Theorem 2 *Let* $\Phi : \mathcal{S} \to \Phi(\mathcal{S})$ *be a bijection and suppose that Assumption 1 holds. Then* $h(\sigma) = h(\Sigma|_{\Phi(\mathcal{S})})$.

Other conditions leading also to $h(\sigma) = h(\Sigma|_{\Phi(\mathcal{S})})$ will be discussed in Section 5.

Notice that $h(\Sigma|_{\Phi(\mathcal{S})})$ corresponds to what we called *macroscopic entropy* in the Introduction. Indeed, since (i) $\Phi(\mathbf{s})$ is a trajectory in $\mathfrak{A}$, a set whose "points" are the component sets $A(\mathbf{s})$ of $\mathfrak{A}$; and (ii) $\chi(\mathbf{s}) \in A(\mathbf{s})$ implies $\chi(\sigma^n \mathbf{s}) \in A(\sigma^n \mathbf{s})$ for all $n \in \mathbb{Z}$, it holds that $h(\Sigma|_{\Phi(\mathcal{S})})$ measures the complexity of the trajectories $(\chi(\sigma^n \mathbf{s}))_{n \in \mathbb{Z}}$ up to the precision set by the distinct component sets $A(\sigma^n \mathbf{s})$. To relate the macroscopic entropy, $h_{macro} \equiv h(\Sigma|_{\Phi(\mathcal{S})})$, to the entropy of the switched dynamics (referred to as *microscopic entropy* in the Introduction), h_{micro}, let $E \subset \mathfrak{K}_R$ be the set of entire orbits, and consider the commutative diagram

$$
\begin{array}{ccc}
E & \overset{\Sigma}{\to} & E \\
i \downarrow & & \downarrow i \\
\Phi(\mathcal{S}) & \overset{\Sigma}{\to} & \Phi(\mathcal{S})
\end{array}
\tag{24}
$$

where i is the inclusion $(\chi(\sigma^n \mathbf{s}))_{n \in \mathbb{Z}} \hookrightarrow (A(\sigma^n \mathbf{s}))_{n \in \mathbb{Z}}$. Thus $\Sigma|_{\Phi(\mathcal{S})}$ is a factor of (or semi-conjugate to) $\Sigma|_E$ on account of i being measure-preserving ($i^{-1}B = B$ for all measurable $B \subset \Phi(\mathcal{S})$) and onto. It follows that $h(\Sigma|_E) \geq h(\Sigma|_{\Phi(\mathcal{S})}) \equiv h_{macro}$, *i.e.*, h_{macro} is a lower bound of the topological entropy of the switched dynamics $h_{micro} \equiv h(\Sigma|_E)$. To keep with the physical identification of entropies, set $h(\sigma) \equiv h_{control}$ for the topological entropy (or complexity) of the control sequence generator. From

$$
h_{macro} \leq h_{micro}
\tag{25}
$$

and Theorem 2,

$$
h_{control} = h_{macro},
\tag{26}
$$

we conclude the following result on the relation between the topological entropies of the control sequence generator, $h(\sigma)$, and the ensuing switched dynamics, h_{micro}.

Theorem 3 *Under the hypotheses of Theorem 2, $h_{control} \leq h_{micro}$.*

This theorem provides sufficient conditions for a complexity version of Parrondo's paradox. Indeed, it suffices that $h(f_+), h(f_-) < h(\sigma)$ (along with the assumptions of Theorem 2), where $h(f_\pm)$ is the topological entropy of $f_\pm$, for the complexity of the switched dynamics to exceed the complexity of the constituent maps.

In the special case of pullback attractors consisting of singletons, like in [1], the inclusion i in the diagram (24) becomes the identity, a trivial isomorphism, and $h_{macro} = h_{micro}$.

4. Conditions for the Injectivity of Φ

According to Theorem 2, the injectivity of Φ (along with Assumption 1) is instrumental for $h(\sigma)$ and $h(\Sigma)$ to coincide. We explore next sufficient conditions for the injectivity of Φ.

If Φ is not injective, then there are control sequences $\mathbf{s} \neq \mathbf{s}^*$ such that $\Phi(\mathbf{s}) = \Phi(\mathbf{s}^*)$, *i.e.*, $A(\sigma^n \mathbf{s}) = A(\sigma^n \mathbf{s}^*)$ for all $n \in \mathbb{Z}$. Then

$$f_{s_n}(A(\sigma^n \mathbf{s})) = A(\sigma^{n+1} \mathbf{s}) = A(\sigma^{n+1} \mathbf{s}^*) = f_{s_n^*}(A(\sigma^n \mathbf{s}^*)) = f_{s_n^*}(A(\sigma^n \mathbf{s}))$$

and, similarly,

$$f_{s_n^*}(A(\sigma^n \mathbf{s}^*)) = f_{s_n}(A(\sigma^n \mathbf{s}^*))$$

for all $n \in \mathbb{Z}$. It follows that

$$f_+(A(\sigma^n \mathbf{s}) = f_-(A(\sigma^n \mathbf{s}), \ f_+(A(\sigma^n \mathbf{s}^*) = f_-(A(\sigma^n \mathbf{s}^*) \tag{27}$$

for those $n \in \mathbb{Z}$ such that $s_n \neq s_n^*$ (a nonempty set by hypothesis). This being the case, there are several conditions that prevent Equation (27) from occurring, thus guaranteeing the injectivity of Φ.

Assumption 2. There is no $A(\mathbf{s}) \in \mathfrak{A}$ such that $f_+(A(\mathbf{s})) = f_-(A(\mathbf{s}))$ or, equivalently, $f_+(A(\mathbf{s})) \neq f_-(A(\mathbf{s}))$ for all $A(\mathbf{s}) \in \mathfrak{A}$.

This condition can be reworded as follows: f_+ and f_- distinguish all pullback attractors of the switching dynamics.

Example 1 *Let $d = 1$, and $f_\pm$ as in Equation (1). In this case, the pullback attractor $\mathfrak{A} = \{A(\mathbf{s}), \mathbf{s} \in \mathcal{S}\}$ consists of singleton sets,* i.e. *$A(\mathbf{s}) = \{\chi(\mathbf{s})\}$, with $\chi(\mathbf{s}) \in [\frac{-1}{1-\theta_-}, \frac{1}{1-\theta_+}] = \mathfrak{A}$ (see [1]). Let us check whether Assumption 2 holds:*

$$f_+(\chi(\mathbf{s})) = f_-(\chi(\mathbf{s})) \ \Leftrightarrow \ \theta_- \neq \theta_+ \text{ and } \chi(\mathbf{s}) = \frac{2}{\theta_- - \theta_+}$$

but $\frac{2}{\theta_- - \theta_+} \notin [\frac{-1}{1-\theta_-}, \frac{1}{1-\theta_+}]$ because $\theta_- + \theta_+ < 2$. Since Assumption 2 holds, Φ is injective.

Example 2 *Let $A^\pm$ be the attractors of $f_\pm$, i.e., $A^+ = A((+1)_{n \in \mathbb{Z}})$, and $A^- = A((-1)_{n \in \mathbb{Z}})$. If $A^+ = A^-$, then Assumption 2 does not hold since $f_+(A^+) = A^+$, and $f_-(A^-) = A^-$. Think, for instance, of two linear dynamics $f_\pm(x) = M_\pm x$ having only 0 as a fixed point. Therefore, in this case the topological entropy of the switched dynamics may be different from the topological entropy of the control sequence.*

Now suppose that at least one of the maps $f_\pm$ is invertible and define $g := f_+^{-1} \circ f_-$ if f_+ is invertible, or $g := f_-^{-1} \circ f_+$ otherwise; if both maps are invertible, then either choice is the inverse of the other one. Then

$$f_+(A(\mathbf{s})) = f_-(A(\mathbf{s})) \iff A(\mathbf{s}) = g(A(\mathbf{s})).$$

Therefore, Assumption 2 can be particularized in the following way.

Assumption 2 for Invertible Mappings. In case that one of the maps $f_\pm$ is invertible, there is no g-invariant $A(\mathbf{s}) \in \mathfrak{A}$.

Example 3 *Suppose that $M_\pm$ are two $d \times d$-matrices, with M_+ being invertible, and $f_\pm(x) = M_\pm x$, so that $g(x) = Gx$ with $G = M_+^{-1} M_-$. Thus, when applying Assumption 2 for invertible mappings, we are asking about the existence of invariant sets of known linear maps. Linear maps can be decomposed as actions of scalings (expansion/contraction, projection, reflection) and/or rotations. Therefore, the existence of G-invariant pullback attractors $A(\mathbf{s})$ boils down to well-known geometrical facts.*

- *The only invariant set of an (in general anisotropic) expansion/contraction is $\{0\}$.*

- *If a linear map is a projection onto a subspace V, then any set contained in V is invariant.*

- *If a linear map is a reflection, then any set symmetric under reflections (like a star-like shaped object) is invariant.*

- *If a linear map is a rotation by an angle α, then any set symmetric under that rotation is invariant, in particular any ball centered at the origin of radius $r \geq 0$ and dimension n or $0 \leq n \leq d - 1$ if contained in the orthogonal complement of the rotation axis.*

Since the origin is invariant in any case, it might happen that $A(\mathbf{s}) = \{0\}$ for some $\mathbf{s} \in \mathcal{S}$ (for example, for $(+1)_{n\in\mathbb{Z}}$ or $(-1)_{n\in\mathbb{Z}}$, since $\{0\} \subset A^+, A^-$, where $A^\pm$ are the attractors of $f_\pm$). In this case, the version of Assumption 2 for invertible mappings would not hold.

From the preceding discussion we conclude the following result.

Lemma 3 *Suppose that the maps $f_\pm$ are such that Assumption 2 holds. Then the map $\Phi : \mathcal{S} \to \mathfrak{K}_R$ defined by $\Phi(\mathbf{s}) := (A(\sigma^n \mathbf{s}))_{n\in\mathbb{Z}}$ is one-to-one.*

5. Results Using Continuity

The main technical difference between the setting of [1] and the present general setting is that the map $\mathbf{s} \mapsto A(\mathbf{s})$ is not necessarily continuous (Proposition 3). It is, however, worth following the same path as in [1] to see how far we can go and under which conditions. The first condition is obvious.

Assumption 3. *The set-valued mapping $\mathbf{s} \mapsto A(\mathbf{s})$ is continuous in $(\mathcal{K}_R, \mathrm{dist}_H)$, i.e.,*

$$\mathrm{dist}_H\left(A(\mathbf{s}), A(\mathbf{s}^*)\right) \to 0 \quad \text{as} \quad \mathrm{dist}_{\mathcal{S}}(\mathbf{s}, \mathbf{s}^*) \to 0 \tag{28}$$

Proposition 4 provides a sufficient condition for Assumption 3 to hold.

We are going to show in this section that if Assumptions 2 (Section 4) and 3 are fulfilled, then $\Phi|_{\Phi(\mathcal{S})}$ is a homeomorphism.

Lemma 4 *Under Assumption 3, the map* $\Phi : \mathcal{S} \to \mathfrak{K}_R$ *defined by* $\Phi(\mathbf{s}) = (A(\sigma^n \mathbf{s}))_{n \in \mathbb{Z}}$ *is uniformly continuous.*

Proof. We have to show that for any $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that

$$\mathrm{dist}_{\mathfrak{K}_R}\left((A(\sigma^n \mathbf{s}))_{n \in \mathbb{Z}}, (A(\sigma^n \mathbf{s}^*))_{n \in \mathbb{Z}}\right) < \varepsilon$$

whenever $\mathrm{dist}_{\mathcal{S}}(\mathbf{s}, \mathbf{s}^*) < 2^{-(N+1)}$ (*i.e.*, whenever $s_n = s_n^*$ for $|n| \leq N$). The uniformity follows then from the compactness of $\mathcal{S}$.

Consider first the second term in the decomposition

$$\mathrm{dist}_{\mathfrak{K}_R}((A(\sigma^n \mathbf{s}))_{n \in \mathbb{Z}}, (A(\sigma^n \mathbf{s}^*))_{n \in \mathbb{Z}} \tag{29}$$

$$= \sum_{|n| \leq N} \frac{\mathrm{dist}_H \left(A(\sigma^n \mathbf{s}), A(\sigma^n \mathbf{s}^*)\right)}{2^{|n|}} + \sum_{|n| \geq N+1} \frac{\mathrm{dist}_H \left(A(\sigma^n \mathbf{s}), A(\sigma^n \mathbf{s}^*)\right)}{2^{|n|}}$$

Since $A(\sigma^n \mathbf{s}), A(\sigma^n \mathbf{s}^*) \subset \bar{B}_R(0)$ for any $n \in \mathbb{Z}$, we have

$$\mathrm{dist}_H \left(A(\sigma^n \mathbf{s}), A(\sigma^n \mathbf{s}^*)\right) \leq 2R,$$

hence

$$\sum_{|n| \geq N+1} \frac{\mathrm{dist}_H \left(A(\sigma^n \mathbf{s}), A(\sigma^n \mathbf{s}^*)\right)}{2^{|n|}} \leq 4 \sum_{n=N+1}^{\infty} \frac{1}{2^{n-1}} = \frac{4}{2^{N-1}} < \frac{\varepsilon}{2} \tag{30}$$

if $N \geq N_{II} := \left\lfloor \frac{\log(8/\varepsilon)}{\log 2} \right\rfloor + 2$

We use next Assumption 1 to show that the first term in Equation (29) with $N = N_{II}$ can also be made smaller than $\varepsilon/2$. Indeed, for each n with $|n| \leq N_{II}$ there exist by Equation (28) $N_n \in \mathbb{N}$ such that

$$\mathrm{dist}_{\mathcal{S}}(\sigma^n \mathbf{s}, \sigma^n \mathbf{s}^*) < 2^{-(N_n+1)} \implies \mathrm{dist}_H \left(A(\sigma^n \mathbf{s}), A(\sigma^n \mathbf{s}^*)\right) < \frac{\varepsilon}{6}$$

Set $N_I := \max\{N_n : |n| \leq N_{II}\}$. Then

$$\mathrm{dist}_{\mathcal{S}}(\mathbf{s}, \mathbf{s}^*) < 2^{-(N_I+N_{II}+1)} \implies \mathrm{dist}_{\mathcal{S}}(\sigma^n \mathbf{s}, \sigma^n \mathbf{s}^*) < 2^{-(N_n+1)}$$

for $|n| \leq N_{II}$ since

$$(\sigma^n \mathbf{s})_k = (\sigma^n \mathbf{s}^*)_k \text{ for } |k| \leq N_n \iff s_{k+n} = s_{k+n}^* \text{ for } |k| \leq N_n$$

where $|n| \leq N_{II}$ and $|k| \leq N_n \leq N_I$. Therefore if $\mathrm{dist}_{\mathcal{S}}(\mathbf{s}, \mathbf{s}^*) < 2^{-(N_I+N_{II}+1)}$, then

$$\sum_{|n| \leq N_{II}} \frac{\mathrm{dist}_H \left(A(\sigma^n \mathbf{s}), A(\sigma^n \mathbf{s}^*)\right)}{2^{|n|}} < \frac{\varepsilon}{6} \sum_{|n| \leq N_{II}} \frac{1}{2^{|n|}} < \frac{\varepsilon}{2} \tag{31}$$

All in all, we conclude from Equations (29), (31) and (30) with $N = N_{II}$ that if $\mathrm{dist}_{\mathcal{S}}(\mathbf{s}, \mathbf{s}^*) < 2^{-(N_I+N_{II}+1)}$, then $\mathrm{dist}_{\mathfrak{K}_R}\left((A(\sigma^n \mathbf{s}))_{n \in \mathbb{Z}}, (A(\sigma^n \mathbf{s}^*))_{n \in \mathbb{Z}}\right) < \varepsilon$. $\square$

Remark. The above proof exploits the special structure of the model and provides insight into it. The result, in fact, follows from the continuity of the mapping Φ and the compactness of the metric space $(\mathcal{S}, \mathrm{dist}_{\mathcal{S}})$.

If the correspondence $\mathbf{s} \mapsto \Phi(\mathbf{s}) := (A(\sigma^n \mathbf{s}))_{n \in \mathbb{Z}}$ is invertible, then given a sequence $(A_n)_{n \in \mathbb{Z}} \in \Phi(\mathcal{S})$ there exists a unique control sequence $\mathbf{s} = (s_n)_{n \in \mathbb{Z}}$ such that $A_{n+1} = f_{s_n}(A_n)$ for every n. Therefore, s_n can be determined from the knowledge of A_n and A_{n+1}.

Lemma 5 *Suppose that Assumption 2 holds so that the mapping Φ is one-to-one. Then, $\Phi^{-1} : \Phi(\mathcal{S}) \to \mathcal{S}$ is continuous.*

Proof. Given $N > 0$, we have to prove that there exists $\delta > 0$ such that

$$\text{dist}_{\mathfrak{K}_R} \left((A_n)_{n \in \mathbb{Z}}, (A_n^*)_{n \in \mathbb{Z}} \right) < \delta \;\; \Rightarrow \;\; \text{dist}_{\mathcal{S}} \left(\mathbf{s}, \mathbf{s}^* \right) < 2^{-N}$$

where $(A_n)_{n \in \mathbb{Z}}, (A_n^*)_{n \in \mathbb{Z}} \in \Phi(\mathcal{S})$, and $\mathbf{s} = \Phi^{-1}((A_n)_{n \in \mathbb{Z}})$, $\mathbf{s}^* = \Phi^{-1}((A_n^*)_{n \in \mathbb{Z}})$.

From

$$\text{dist}_{\mathfrak{K}_R} \left((A_n)_{n \in \mathbb{Z}}, (A_n^*)_{n \in \mathbb{Z}} \right) = \sum_{|n| \leq N} \frac{\text{dist}_H \left(A_n, A_n^* \right)}{2^{|n|}} + \sum_{|n| > N} \frac{\text{dist}_H \left(A_n, A_n^* \right)}{2^{|n|}} < \delta$$

it follows

$$\text{dist}_H \left(A_n, A_n^* \right) < 2^N \delta \;\; \text{for} \;\; |n| \leq N \tag{32}$$

Consider now two neighboring sequences in $\Phi(\mathcal{S})$,

$$...A_{-N}, A_{-N+1}, ..., A_0, ..., A_{N-1}, A_N, ...,$$
$$...A_{-N}^*, A_{-N+1}^*, ..., A_0^*, ..., A_{N-1}^*, A_N^*, ...,$$

of distance δ apart. From these sequences we determine the corresponding switching sequence under Φ^{-1}:

$$...s_{-N}, s_{-N+1}, ..., s_0, ..., s_{N-1}, ...,$$
$$...s_{-N}^*, s_{-N+1}^*, ..., s_0^*, ..., s_{N-1}^*, ...$$

Note that $A_n = A(\sigma^n \mathbf{s})$, where $\mathbf{s} = (s_n)_{n \in \mathbb{Z}}$, and similarly $A_n^* = A(\sigma^n \mathbf{s}^*)$, where $\mathbf{s}^* = (s_n^*)_{n \in \mathbb{Z}}$.

Now suppose that, contrarily to what we need to prove, there is an n, $-N \leq n \leq N-1$, such that $s_n^* = -s_n$. For brevity assume $s_n = +1$, *i.e.*

$$A_{n+1} = f_+(A_n), \;\; \text{and} \;\; A_{n+1}^* = f_-(A_n^*). \tag{33}$$

Then, by the triangle inequality,

$$\text{dist}_H \left(f_+(A_n^*), f_-(A_n^*) \right) \leq \text{dist}_H \left(f_+(A_n^*), f_+(A_n) \right) + \text{dist}_H \left(f_+(A_n), f_-(A_n^*) \right). \tag{34}$$

We claim that the right hand side of Equation (34) can be made arbitrarily small by choosing δ in Equation (32) sufficiently small.

First of all, by Proposition 2 and Equation (32)

$$\text{dist}_H \left(f_+(A_n^*), f_+(A_n) \right) < \omega(\text{dist}_H \left(A_n^*, A_n \right))$$

where $\omega(\cdot)$ is continuous and $\omega(0) = 0$. As for the second term on the right hand side of Equation (34), use Equations (33) and (32) to derive

$$\text{dist}_H \left(f_+(A_n), f_-(A_n^*) \right) = \text{dist}_H \left(A_{n+1}, A_{n+1}^* \right) < 2^N \delta$$

Therefore,

$$\text{dist}_H\left(f_+(A_n^*), f_-(A_n^*)\right) < \omega(\text{dist}_H\left(A_n^*, A_n\right)) + 2^N \delta \tag{35}$$

where

$$\omega(\text{dist}_H\left(A_n^*, A_n\right)) \to 0 \ \text{ when } \ \delta \to 0 \tag{36}$$

This proves our claim.

It follows from Equations (35) and (36) that $f_+(A_n^*) = f_-(A_n^*)$, in contradiction to Assumption 2. Hence we conclude that $s_n^* = s_n$ for $-N \leq n \leq N-1$, *i.e.*, that $\text{dist}_{\mathcal{S}}\left(\mathbf{s}, \mathbf{s}^*\right) < 2^{-N}$, where $\mathbf{s} = \Phi^{-1}((A_n)_{n\in\mathbb{Z}})$, and $\mathbf{s}^* = \Phi^{-1}((A_n^*)_{n\in\mathbb{Z}})$. $\square$

Notice that we did not assume Φ to be continuous in Lemma 5. With Lemmas 4 and 5 we derive a different version of Theorem 2.

Theorem 4 *Suppose that Assumptions 2 and 3 hold. Then the mapping $\Phi : \mathcal{S} \to \Phi(\mathcal{S})$ defined as $\mathbf{s} \mapsto (A(\sigma^n \mathbf{s}))_{n\in\mathbb{Z}}$ is a homeomorphism, hence $h(\sigma) = h(\Sigma|_{\Phi(\mathcal{S})})$,* i.e., $h_{control} = h_{macro} \leq h_{micro}$.

6. Numerical Simulations

In this section we illustrate the previous theoretical results with a few numerical simulations. As in [1] we resort to two-state Markov processes with transition probability matrices

$$P = \begin{pmatrix} p & 1-p \\ q & 1-q \end{pmatrix} \tag{37}$$

to generate the control sequences. We recall that if the transition probability matrix is irreducible, then the process is ergodic (and even there is a unique invariant measure whose metric entropy coincides with the topological entropy of the corresponding topological Markov chain).

6.1. Case 1

Consider $f_\pm : \mathbb{R}^2 \to \mathbb{R}^2$ with

$$f_+(x, y) = (0.25x, 0.75y) \tag{38}$$

and the Hénon map

$$f_-(x, y) = (1 - 1.4x^2 + 0.3y, x) \tag{39}$$

A common bounded, positively invariant absorbing set for Equations (38) and (39) is the quadrilateral with vertices [16] (p. 614).

$$P_1 = (-1.33, 0.42), P_2 = (1.32, 0.133), P_3 = (1.245, -0.14), P_4 = (-1.06, -0.5) \tag{40}$$

Therefore, the switching system defined by Equations (38) and (39) has a pullback attractor, which is also a forward attractor (Remark 1). We take advantage of this fact in the estimation of the topological entropy by taking a sample of long forward trajectories and deleting the first 100 points.

The topological entropy of these two maps in bits per iteration (*i.e.*, taking logarithms to base 2) are

$$h(f_+) = 0, \quad h(f_-) = 0.67072 \pm 0.00004 \tag{41}$$

($h(f_-)$ taken from [17]). Table 1 summarizes the numerical results. $h_{control} \equiv h(\sigma)$ ($= \log \lambda$, where λ is the largest positive eigenvalue of the admissibility matrix [1,15] denotes the topological entropy of the switching sequence. As above, h_{micro} denotes the topological entropy of the switched dynamics. Conditioned on the assumptions of Theorem 3 or 4, the inequality $h_{control} \leq h_{micro}$ holds. Figure 1 shows the extrapolation technique we used based on entropy rates of order $3 \leq L \leq 7$ (see [1] for more details). A total of 40 orbits were generated for each choice of the probabilities $(p, q) = (0.7, 0.2)$, $(0.25, 1)$, and $(0, 1)$, with random initial conditions within the trapping region (40). Note that $h(f_+), h(f_-) < h_{control}$ for $(p, q) = (0.7, 0.2)$, $(0.25, 1)$.

Table 1. Numerical results of Case 1 ($h_{control} \leq h_{micro}$).

(p, q)	$h_{control}$ (bit/iter)	$h_{micro} \pm s.d.$ (bit/iter)
$(0.7, 0.2)$	1	1.237 ± 0.002
$(0.25, 1)$	0.694	0.702 ± 0.013
$(0, 1)$	0	0.000 ± 0.000

Figure 1. This plot illustrates the extrapolation technique used to estimate the values of h_{micro} in Table 1.
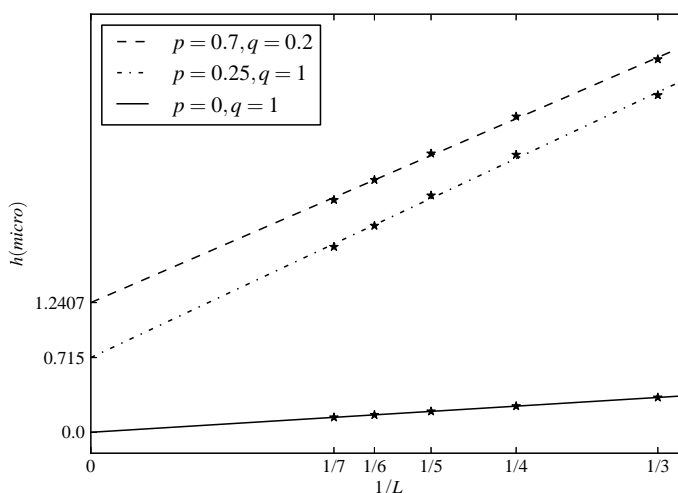
It turns out that, except for the periodic switching $(p, q) = (0, 1)$, the complexity of the switched dynamics is higher than the complexity of the control switching sequence, and also higher than the entropies of the constituent maps, see Equation (41). This provides an instance of Parrondo's paradox with regard to the complexity as measured by the topological entropy.

6.2. Case 2

Consider the constituent affine maps

$$f_{\pm}(x) = M_{\pm}x + b_{\pm} \tag{42}$$

where $x = (x_1, x_2) \in \mathbb{R}^2$,

$$M_+ = \frac{1}{4}\begin{pmatrix} 5 & -4 \\ 2 & -1 \end{pmatrix} \quad M_- = \frac{1}{4}\begin{pmatrix} 5 & -6 \\ 3 & -4 \end{pmatrix}$$

and

$$b_+ = \begin{pmatrix} 3 \\ 2 \end{pmatrix} \quad b_- = \begin{pmatrix} -3 \\ -2 \end{pmatrix}$$

A special feature of the matrices $M_{\pm}$ is that they are simultaneous diagonalizable. Indeed, in the common eigenvector base

$$v_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \quad v_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \tag{43}$$

the constituent maps are transformed to (we keep the notation $x = (x_1, x_2)$ for the coordinates in the new base)

$$f_+(x_1, x_2) = \frac{1}{4}\begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{3}{4}x_1 + 1 \\ \frac{1}{4}x_2 + 1 \end{pmatrix},$$

$$f_-(x_1, x_2) = \frac{1}{4}\begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} -1 \\ -1 \end{pmatrix} = \begin{pmatrix} \frac{1}{2}x_1 - 1 \\ \frac{-1}{4}x_2 - 1 \end{pmatrix},$$

Therefore, the two-dimensional switched dynamics (42) decomposes in the base (43) into two uncoupled, one-dimensional switched dynamics whose constituent maps are of the type (1) studied in [1]. Here the pullback attractors are singleton sets (hence the macroscopic and microscopic complexities coincide), Assumptions 2 and 3 are fulfilled, and $h_{micro} = h_{macro} = h_{control}$, the latter equality holding by Theorem 4. Furthermore, $h(f_+) = h(f_-) = 0$, thus $h(f_+), h(f_-) < h_{control}$ for $(p,q) = (0.7, 0.2)$, $(0.25, 1)$, as in Case 1. Table 2 shows numerical evidence of this statement. Figure 2 shows the extrapolation technique based on entropy rates of orders $3 \leq L \leq 7$.

Table 2. Numerical results of Case 2 ($h_{control} = h_{micro}$).

(p, q)	$h_{control}$ (bit/iter)	$h_{micro} \pm$ s.d. (bit/iter)
$(0.7, 0.2)$	1	1.002 ± 0.003
$(0.25, 1)$	0.694	0.687 ± 0.007
$(0, 1)$	0	0.000 ± 0.000

Figure 2. This plot illustrates the extrapolation technique used to estimate the values of h_{micro} in Table 2.
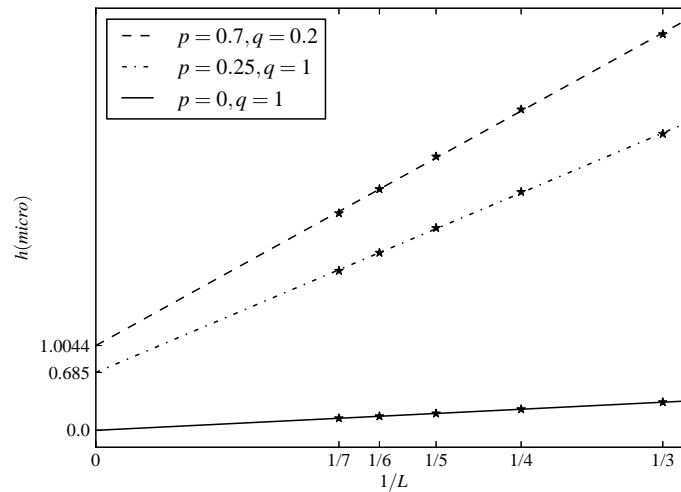
7. Conclusions

Consider the time-switched dynamics associated with two maps $f_{\pm} : \mathbb{R}^d \to \mathbb{R}^d$ having global attractors $A^{\pm}$. In this paper we studied the relation between the topological entropy of the control sequence generator and the topological entropy of the switched dynamics. Compared with the previous paper [1], we found some new technical obstacles.

First of all, we had to circumvent the shortcoming that the map $\mathbf{s} \mapsto A(\mathbf{s})$ is not necessarily continuous in $(\mathcal{K}_R, \mathrm{dist}_H)$; here $\mathbf{s}$ is a control sequence and $A(\mathbf{s})$ is the corresponding component set of the pullback attractor, $\mathfrak{A} = \{A(\mathbf{s}), \mathbf{s} \in \mathcal{S}\}$. To this end we assumed in Section 3 ("Results using measurability"), (i) the measurability of certain unions of component sets (Assumption 1); and (ii) the injectivity of $\Phi : \mathbf{s} \mapsto (A(\sigma^n \mathbf{s}))_{n \in \mathbb{Z}}$. It follows then, Theorem 2, that Φ is a topological conjugacy between the shift on the control sequences, $(\{+1, -1\}^{\mathbb{Z}}, \sigma)$, and the shift on the trajectories of the switched dynamics up to the precision set by the component sets, $(\mathfrak{K}_R, \Sigma)$. Being the entropy of coarse-grained trajectories, $h(\Sigma)$ is a lower bound of h_{micro}, the topological entropy of the switched dynamics. We call $h(\Sigma)$ the macroscopic entropy and write h_{macro}. Likewise, $h(\sigma)$ is the entropy of the control sequence generator, thus we write $h_{control}$. In sum, $h_{control} = h_{macro} \leq h_{micro}$ [Equations (25) and (26)] under the hypotheses (i) and (ii).

In Section 4 we surveyed sufficient conditions for the injectivity of Φ postulated in Section 3. We formulated our conclusions as Assumption 2, both in a general version, and in a special version when one of the maps $f_{\pm}$ is invertible.

In Section 5 ("Results using continuity") we hypothesized from the outset that the mapping $\mathbf{s} \mapsto A(\mathbf{s})$ is continuous (Assumption 3) and paralleled the approach of [1]. In Theorem 4 we proved that Φ is a homeomorphism (or topological conjugacy), Assumptions 2 and 3 granted. Once again we concluded that $h_{control} = h_{macro} \leq h_{micro}$.

The main results of this paper, Theorems 3 and 4, are in line with what we call the complexity version of Parrondo's paradox, namely, that the complexity of a switched dynamics may be higher than

the complexity of its constituent maps. Moreover, they provide sufficient conditions as well, to wit: $h(f_+), h(f_-) < h_{control}$.

Finally, both possibilities $h_{control} < h_{micro}$ (Case 1) and $h_{control} = h_{micro}$ (Case 2) have been numerically illustrated in Section 6.

Acknowledgements

Conflict of Interest

The authors declare no conflict of interest.

Appendix

Recall that the Hausdorff semi-distance and Hausdorff distance between two nonempty subsets X and Y of a metric space (M, dist) are denoted by $\rho(X, Y)$ and $\mathrm{dist}_H(X, Y)$, respectively, while

$$\mathrm{diam}\, X = \sup_{x, x' \in X} \mathrm{dist}(x, x') \tag{44}$$

is called the diameter of a set X.

Let

$$B_\varepsilon(x) = \{z \in M : \mathrm{dist}(z, x) < \varepsilon\}, \quad \bar{B}_\varepsilon(x) = \{z \in M : \mathrm{dist}(z, x) \leq \varepsilon\}$$

be the open and closed balls in (M, dist) with center at x and radius $\varepsilon > 0$, respectively. The set of all points within a distance ε of X, *i.e.*,

$$E_\varepsilon(X) = \bigcup_{z \in X} \bar{B}_\varepsilon(z)$$

is called the ε-inflation (or ε-expansion) of X. It follows trivially from this definition that

$$X \subset E_\varepsilon(Y) \text{ and } Y \subset E_\varepsilon(X) \;\Rightarrow\; \mathrm{dist}_H(X, Y) \leq \varepsilon \tag{45}$$

The converse of Equation (45) is in general false, except when X and Y are compact.

Proposition A1. *Let X and Y be compact subsets of (M, dist). Then*

$$\mathrm{dist}_H(X, Y) = \varepsilon \;\Rightarrow\; X \subset E_\varepsilon(Y) \text{ and } Y \subset E_\varepsilon(X)$$

Proof. If $\mathrm{dist}_H(X, Y) = \varepsilon$, then $\rho(X, Y) \leq \varepsilon$ and $\rho(Y, X) \leq \varepsilon$. This amounts to

$$\max_{x \in X} \mathrm{dist}(x, Y) \leq \varepsilon, \text{ where } \mathrm{dist}(x, Y) := \min_{y \in Y} \mathrm{dist}(x, y) \tag{46}$$

and

$$\max_{y \in Y} \mathrm{dist}(y, X) \leq \varepsilon, \text{ where } \mathrm{dist}(y, X) := \min_{x \in X} \mathrm{dist}(x, y) \tag{47}$$

respectively. Therefore, see Equation (46), for every $x \in X$, there exists $y \in Y$ such that $x \in \bar{B}_\varepsilon(y)$, and, see Equation (47), for every $y \in Y$, there exists $x \in X$ such that $y \in \bar{B}_\varepsilon(x)$. $\square$

It follows from Proposition A1 that

$$\operatorname{dist}_H(X, Y) \le \varepsilon \ \Rightarrow \ X \subset E_\varepsilon(Y) \text{ and } Y \subset E_\varepsilon(X) \tag{48}$$

when X and Y are compact.

Among the different properties of the Hausdorff distance, we mention here only two of them. Let $\mathcal{K}(M)$ be the set of all nonempty *compact* subsets of M.

1. If (M, dist) is complete, then $(\mathcal{K}(M), \operatorname{dist}_H)$ is also a complete.

2. If (M, dist) is compact, then $(\mathcal{K}(M), \operatorname{dist}_H)$ is also a compact.

In Section 2.2 we consider the metric space $M = \mathbb{R}^d$ endowed with the Euclidean distance $|\cdot|$, as well as the complete, compact metric space $(\mathcal{K}_R, \operatorname{dist}_H)$, where $\mathcal{K}_R = \mathcal{K}(\mathbb{R}^d) \cap \bar{B}_R(0)$.

Lemma A2. *Let $(M_1, \operatorname{dist}_1)$ be a compact metric space and $(M_2, \operatorname{dist}_2)$ a complete metric space. If $f : M_1 \to M_2$ is a continuous map, then there exists a real-valued function $\omega : [0, \infty) \to [0, \infty)$ with $\omega(0) = 0$ such that, for any $X, Y \in \mathcal{K}(M_1)$,*

$$\operatorname{dist}_{2,H}(f(X), f(Y)) \le \omega(\operatorname{dist}_{1,H}(X, Y)) \tag{49}$$

where $\operatorname{dist}_{1,H}$ (resp. $\operatorname{dist}_{2,H}$) is the Hausdorff distance in $\mathcal{K}(M_1)$ (resp. $\mathcal{K}(M_2)$).

The function ω here is called a modulus of continuity. Moduli of continuity are used to express in a convenient way both the continuity at a point and the uniform continuity of maps between metric spaces as, for instance, in Equation (49).

Proof of Lemma A2. Let X and Y be compact subsets of $(M_1, \operatorname{dist}_{1,H})$ such that $\operatorname{dist}_{1,H}(A, B) \le \delta$. By Equation (48),

$$X \subset \bigcup_{i=0}^{m} \bar{B}_\delta(y_i) \text{ and } Y \subset \bigcup_{j=0}^{n} \bar{B}_\delta(x_j)$$

where $x_j \in X$ and $y_i \in Y$. Thus,

$$f(X) \subset \bigcup_{i=0}^{m} f(\bar{B}_\delta(y_i)) \text{ and } f(Y) \subset \bigcup_{j=0}^{n} f(\bar{B}_\delta(x_j))$$

In turn, $f(\bar{B}_\delta(y_i))$ and $f(\bar{B}_\delta(x_j))$ are compact in M_2 because $f : M_1 \to M_2$ is continuous. This being the case,

$$f(\bar{B}_\delta(y_i)) \subset \bar{B}_{r_i}(f(y_i)) \text{ and } f(\bar{B}_\delta(x_j)) \subset \bar{B}_{s_j}(f(x_j))$$

where [see Equation (44)]

$$r_i \le \operatorname{diam} f(\bar{B}_\delta(y_i)) \text{ and } s_j \le \operatorname{diam} f(\bar{B}_\delta(x_j))$$

for $1 \leq i \leq m$ and $1 \leq j \leq n$. In sum,

$$f(X) \subset \bigcup_{i=0}^{m} \bar{B}_{r_i}(f(y_i)) \text{ and } f(Y) \subset \bigcup_{j=0}^{n} \bar{B}_{s_j}(f(x_j)) \tag{50}$$

Let $\omega_f : [0, +\infty] \to [0, +\infty]$ be the modulus of continuity of the uniformly continuous map f of M_1 into M_2, *i.e.*,

$$\text{dist}_2(f(z), f(z')) < \omega_f(\text{dist}_1(z, z'))$$

where

$$\lim_{t \to 0} \omega_f(t) = \omega_f(0) = 0 \tag{51}$$

Without loss of generality, ω_f may be assumed to be continuous and even differentiable. Then

$$r_i \leq \text{diam} \, f(\bar{B}_\delta(y_i)) \;\; = \;\; \max\{\text{dist}_2(f(z), f(z')) : z, z' \in \bar{B}_\delta(y_i)\}$$

$$\leq \;\; \max\{\omega_f(\text{dist}_1(z, z')) : z, z' \in \bar{B}_\delta(y_i)\}$$

for $1 \leq i \leq m$, and, analogously,

$$s_j \leq \max\{\omega_f(\text{dist}_1(z, z')) : z, z' \in \bar{B}_\delta(x_j)\}$$

for $1 \leq j \leq n$. From Equation (51) we conclude that r_i and s_j can be made arbitrarily small by choosing δ sufficiently small.

All in all, given $\varepsilon > 0$ and $X, Y \subset M_1$ compact such that $\text{dist}_H(A, B) < \delta$, we can choose δ sufficiently small so that $r_i, s_j \leq \varepsilon$ for $1 \leq i \leq m$, and $1 \leq j \leq n$. From Equation (50) we obtain

$$f(X) \subset \bigcup_{i=0}^{m} \bar{B}_\varepsilon(f(y_i)) \subset E_\varepsilon(f(Y)) \text{ and } f(Y) \subset \bigcup_{j=0}^{n} \bar{B}_\varepsilon(f(x_j)) \subset E_\varepsilon(f(X))$$

hence $\text{dist}_{2,H}(f(X), f(Y)) \leq \varepsilon$ by Equation (45).

This, in fact, proves the uniform continuity (since $\mathcal{K}(M_1)$ is compact) of the set-valued map $\mathcal{K}(M_1) \ni A \mapsto f(A) \in \mathcal{K}(M_2)$. Equation (49) is just an equivalent formulation of this fact. $\qquad\square$

References

1. Amigó, J.M.; Kloeden, P.E.; Giménez, A. Switching systems and entropy. *J. Differ. Equ. Appl.* **2013**, doi:10.1080/10236198.2013.788166.

2. Kloeden, P.E.; Rasmussen, M. *Nonautonomous Dynamical Systems*; American Mathematical Society: Providence, RI, USA, 2011.

3. Liberzon, D. *Switching in Systems and Control*; Birkhäuser: Boston, MA, USA, 2003.

4. Shorten, R.; Wirth, F.; Mason, O.; Wulff, K.; King, C. Stability criteria for switched and hybrid systems. *SIAM Rev.* **2007**, *49*, 545–592.

5. Almeida, J.; Peralta-Salas, D.; Romera, M. Can two chaotic systems give rise to order? *Phys. D: Nonlinear Phenom.* **2005**, *200*, 124–132.

6. Boyarsky, A.; Góra, P.; Islam, M.S. Randomly chosen chaotic maps can give rise to nearly ordered behavior. *Phys. D: Nonlinear Phenom.* **2005**, *210*, 284–294.

7. Cánovas, J.S.; Linero, A.; Peralta-Salas, D. Dynamic Parrondo's paradox. *Phys. D: Nonlinear Phenom.* **2006**, *218*, 177–184.

8. Harner, G.P.; Abbott, D. Parrondo's paradox. *Stat. Sci.* **1999**, *14*, 206–213.

9. Kloeden, P.E. Pullback attractors in nonautonomous difference equations. *J. Differ. Equ. Appl.* **2000**, *6*, 33–52.

10. Cheban, D.N.; Kloeden, P.E.; Schmalfuß, B. The relationship between pullback, forward and global attractors of nonautonomous dynamical systems. *Nonlinear Dyn. Syst. Theory* **2002**, *2*, 125–144.

11. Kloeden, P.E. Pullback attractors of nonautonomous semidynamical systems. *Stoch. Dyn.* **2003**, *3*, 101–112.

12. Kloeden, P.E.; Siegmund, S. Bifurcations and continuous transitions of attractors in autonomous and nonautonomous systems. *Int. J. Bifurcat. Chaos Appl. Sci. Eng.* **2005**, *15*, 743–762.

13. Kloeden, P.E. Nonautonomous attractors of switching systems. *Dyn. Syst.* **2006**, *21*, 209–230.

14. Amigó, J.M. *Permutation Complexity in Dynamical Systems*; Springer: Berlin, Germany, 2010.

15. Walters, P. *An Introduction to Ergodic Theory*; Springer: Berlin, Germany, 2000.

16. Peitgen, H.O.; Jürgens, H.; Saupe, D. *Chaos and Fractals: New Frontiers of Science*, 2nd ed.; Springer: Berlin, Germany, 2004.

17. Hirata, Y.; Mees, A.I. Estimating topological entropy via a symbolic data compression technique. *Phys. Rev. E* **2003**, *67*, 026205.