RESEARCH-ARTICLE

Worst-case topological entropy and minimal data rate for state observation of switched linear systems

GUILLAUME O BERGER

RAPHAËL M JUNGERS

Published: 22 April 2020

HSCC '20: 23rd ACM International Conference on Hybrid Systems: Computation and Control
April 22 - 24, 2020
New South Wales, Sydney, Australia

Conference Sponsors:
SIGBED

https://doi.org/10.1145/3365365.3382195
ISBN: 9781450370189

Worst-case topological entropy and minimal data rate for state observation of switched linear systems

Guillaume O. Berger*
UCLouvain
Louvain-la-Neuve, Belgium
guillaume.berger@uclouvain.be

Raphaël M. Jungers†
UCLouvain
Louvain-la-Neuve, Belgium
raphael.jungers@uclouvain.be

ABSTRACT

We introduce and study the concept of worst-case topological entropy of switched linear systems under arbitrary switching. It is shown that this quantity is equal to the minimal data rate (number of bits per second) required for the state observation of the switched linear system with any switching signal. A computable closed-form expression is presented for the worst-case topological entropy of switched linear systems. Finally, a practical coder–decoder, operating at a data rate arbitrarily close to the worst-case topological entropy, is described.

CCS CONCEPTS

• **Computing methodologies** → **Modeling methodologies**; **Uncertainty quantification**; **Systems theory**.

KEYWORDS

Topological entropy, Switched linear systems, Network-based control, Joint Spectral Radius

ACM Reference Format:
Guillaume O. Berger and Raphaël M. Jungers. 2020. Worst-case topological entropy and minimal data rate for state observation of switched linear systems. In *23rd ACM International Conference on Hybrid Systems: Computation and Control (HSCC '20), April 22–24, 2020, Sydney, NSW, Australia.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3365365.3382195

1 INTRODUCTION

In recent years, the outbreak of cyber-physical systems and communication technologies has opened the door to a new generation of systems, in which the different agents (plants, sensors, actuators, controllers, etc.) are spatially distributed and communicate through a shared, band-limited, digital communication network. This new configuration of systems, coined as Networked Control Systems (NCSs), offers several major advantages compared to traditional centralized or wired systems, such as increased flexibility, lower cost and power, ease of maintenance, etc. [32]. As a consequence, NCSs have found applications in a broad range of areas, like intelligent transportation, remote surgery, haptics collaboration over the Internet, etc. [14]. On the other hand, controlling systems through digital/band-limited/delayed/error-prone communication networks poses many challenges for the control theoretist, which cannot be solved with the traditional approach where control and communication issues are treated separately. This motivated the development of a new chapter in control theory, where the two frameworks are integrated [3, 30].

Digital communication networks can carry only a finite amount of information per unit of time. Moreover, high communication capacity generally comes at the cost of increased power consumption, infrastructure, etc…when not physically infeasible (e.g., applications to underwater vehicles [35]). Consequently, a major challenge in NCSs design is to determine the minimal communication data rate between the plant and the controller required to achieve a given control objective. This fundamental question has attracted a significant research effort in the last decades, covering a large variety of settings (e.g., observability, stabilizability, optimal control; for linear, nonlinear, hybrid systems); see, e.g., [8, 12, 31, 34].

Inspired by Shannon's work on the link between information entropy of a data source and the minimal communication capacity required to carry the information reliably, it was soon realized that the question of data rate requirement for NCSs had strong connections with the notion of topological entropy of dynamical systems. This quantity, introduced in the late 60's [1, 7] and now ubiquitous in dynamical system theory, measures the rate at which information about the initial condition is generated by the system as time evolves. Topological entropy can also be defined as a measure of the growth rate of the smallest number of trajectories necessary to approximate the state of the system with arbitrary finite accuracy [7]. More recently, variants of topological entropy have been proposed to address further aspects of NCSs design; e.g., model uncertainties [34], time-varying systems [19, 22], exponentially decreasing estimation error [27], feedback invariance and feedback stabilization [10, 11, 31] and their modifications [9, 13, 17].

We are interested in the problem of determining the minimal data rate at which a coder needs to send information to a decoder to be able to estimate the state of the system with exponentially decreasing error, and its connection with topological entropy. The problem of state observation with limited data rate is depicted in Figure 1. For LTI systems, the situation is well understood: comprehensive eigenvalue-based expressions exist for the topological entropy of the system, and this quantity is shown to be equal to the minimal data rate for state observation. Moreover, practical

*G. Berger is a FNRS/FRIA Fellow (F.R.S.–FNRS).
†R. Jungers is an FNRS Research Associate. He is supported by the French Community of Belgium, the Walloon Region and the Innoviris Foundation.

HSCC '20, April 22–24, 2020, Sydney, NSW, Australia

ACM ISBN 978-1-4503-7018-9/20/04…$15.00
https://doi.org/10.1145/3365365.3382195

(i.e., implementable) coders–decoders for the state observation of LTI systems, operating at a data rate as close as desired to the topological entropy, are available; see, e.g., [30, §2.4].

Beyond the LTI realm, the situation is unfortunately much more elusive: topological entropy is only known to be a lower bound on the minimal data rate for state observation. In some cases (e.g., autonomous systems with compact domain), it also provides a theoretical upper bound on the minimal data-rate [20, 22, 29, 34]; however, the implementability of a practical coder–decoder approaching the optimal data rate is not guaranteed (e.g., the definition of topological entropy relies on the existence of a minimal number of trajectories needed to approximate the state of the system, disregarding the computability of these trajectories). Furthermore, negative results for the computability of the topological entropy have been proved, even for very simple systems [21]; and the exact value of the topological entropy of most nonlinear systems (even widely-studied ones, such as the Hénon map, the van Der Pol oscillator, etc. [29]) is still unknown. On the other hand, constructive lower bounds and upper bounds on the topological entropy and the minimal data rate for state observation of nonlinear systems have been proposed in the literature (see, e.g., [27, 29]); the major advantage of these constructive data rate upper bounds is that they are generally accompanied by practical coders–decoders that work whenever the channel capacity fits this bound (see, e.g., [27, 29]).

This paper focuses on discrete-time Switched Linear Systems (SLSs). These are systems described by a finite set of linear modes, among which the system can switch in time. As a paradigmatic class of cyber-physical systems, SLSs have attracted much attention from the control community in recent years [15, 25, 28]. These systems turn out to be extremely challenging in terms of control and analysis, even for basic questions like stability or stabilizability [15]. In particular, neither the topological entropy nor the data rate requirements for state observation are well understood for SLSs. These questions have been addressed, e.g., in [38, 40], where upper bounds and lower bounds on the topological entropy, *when the sequence of modes is fixed a priori*, are derived for different classes of SLSs (diagonal, commuting, general, etc.); and also in [26, 39], where sufficient data rates for feedback stabilization of SLSs are established.

In this paper, we introduce and study the concept of *worst-case topological entropy* of SLSs, which is the maximal topological entropy that can be reached by the system among all sequences of modes (aka. switching sequences). The motivation of this concept is to study the questions of topological entropy and data rate requirements for state observation of SLSs in a *worst-case scenario*. The worst-case scenario approach is a popular approach is the study of switched systems as it provides formal guarantees that the system will satisfy the specifications in every situations. In our case, this amounts to find coders–coders that are able to observe the state of the system for every switching sequences.

The contribution of the paper is twofold. First, we present a closed-form expression for the worst-case topological entropy of SLSs. More precisely, the worst-case topological entropy is expressed as the Joint Spectral Radius (a ubiquitous measure of stability of SLSs) of a higher-dimensional SLS representing the action of the original system on elements of volume. The definition of this

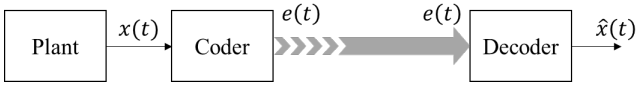

Figure 1: State observation with limited data rate.

SLS acting on volume elements will require tools from multilinear algebra. The main asset of this closed-form expression is that it can be computed numerically via well-established algorithms for the computation of the Joint Spectral Radius (see, e.g., [37]). Consequently, it allows for a systematic analysis of the worst-case topological entropy of SLSs.

The second contribution is to provide a practical coder–decoder that estimates the state of the SLS with exponentially decreasing estimation error, and operating at a data rate arbitrarily close to the worst-case topological entropy of the system. In particular, compared to other bounds on the topological entropy of SLSs available in the literature—for which no practical coders–decoders have been proposed—, this demonstrates the practical relevance of worst-case topological entropy for the problem of state observation of SLSs with limited data rate.

In our analysis, it is assumed that the switching sequence is known by the coder–decoder during the state estimation process. This framework is motivated by the fact that the switching sequence is not always known at the time of the coder–decoder's *design*, but is available to the coder–decoder during its *operation*. An example of application is when one has to design the communication infrastructure between the plant (e.g., of a factory) and the observer (placed at a remote location, e.g., headquarter of the company), and the switching signal is not known at the time of the infrastructure design or might change with time. For instance, a given sequence of modes (periodic in many applications, but this is not a requirement) might be used by the plant of a factory for a certain amount of time. But after a few days or months, another sequence might be needed (e.g. to meet the ever-changing consumer demand). In this case, it is desirable that only the plant and the observer need to be reconfigured, while the communication infrastructure remains unchanged (as it is easier/cheaper to reconfigure a device than rebuilding a whole infrastructure). To meet these requirements, the communication channel will need to satisfy constraints driven by the worst-case topological entropy. Other examples of applications involving the control of SLSs with data rate constraints are discussed in the conclusions.

The paper is organized as follows. In Section 2, we introduce the notions of topological entropy and SLSs. In Section 3, we present the closed-from expression for the worst-case topological entropy of SLSs and discuss the computability aspects. In Section 4, we present a practical coder–decoder for the state observation of SLSs, operating at data rate as close as desired to the worst-case topological entropy of the system. Finally, in Section 5, we demonstrate the applicability of our results on numerical examples. Proofs of the main results can be found in Sections 6 and 7.

Notation. $\mathbb{N}$ is the set of nonnegative integers $\{0, 1, 2, \dots\}$. d is a positive integer representing the dimension of the system. For vectors, $\|\cdot\|$ denotes the Euclidean norm in $\mathbb{R}^d$, and for matrices, it denotes the associated matrix norm (i.e., $\|M\|$ = largest singular

value of M). $B(\xi, r)$ is the Euclidean closed ball in $\mathbb{R}^d$ centered at $\xi \in \mathbb{R}^d$, with radius $r \geq 0$. $\lceil \alpha \rceil$ is the smallest integer larger than or equal to $\alpha \in \mathbb{R}$ (aka. *ceil* of α). In this paper, we consider dynamical systems in discrete time; therefore, if $[T_1, T_2]$ (resp. $[T_1, T_2)$) refers to an interval of *times* (in particular, $T_1, T_2 \in \mathbb{N}$), then it is understood to contain only the integers from T_1 to T_2 (resp. $T_2 - 1$) inclusive. By convention, an "empty product" of real numbers is equal to 1; e.g., an expression like $\prod_{i=1}^{k} \beta_i$, $\beta_i \in \mathbb{R}$, is equal to 1 if $k \leq 0$. Similarly, an "empty product" of matrices is equal to I.

2 PRELIMINARIES

2.1 Topological entropy

Consider a discrete-time *switched system*

$$x(t+1) = f_{\sigma(t)}(x(t)), \quad x(t) \in \mathbb{R}^d, \quad t \in \mathbb{N}, \tag{1}$$

where $\sigma(t) \in \Sigma := \{1, \ldots, N\}$ and $f_i : \mathbb{R}^d \to \mathbb{R}^d$ for all $i \in \Sigma$. The function $\sigma : \mathbb{N} \to \Sigma$ is called the *switching signal*[1] of the system and specifies which *mode*, i.e., which transition map f_i, is used by the switched system at each time. We denote by $x_\sigma(t, \xi)$ the solution, at time t, of (1) with switching signal σ and initial state $\xi \in \mathbb{R}^d$. Let $K \subset \mathbb{R}^d$ be a compact set of initial states with nonempty interior. For the ease of notation, we will sometimes write (f_σ, K) to denote system (1) with switching signal σ and initial set K. In particular, $x_\sigma(\cdot, \xi)$ is a trajectory of (f_σ, K) if and only if $\xi \in K$.

We use the definition of topological entropy introduced by Bowen [7] (see, e.g., [29, 34] for more recent treatment), extended to the case of non-autonomous systems by Kolyada and Snoha [22]. The definition relies on the notion of minimal sets of trajectories necessary to approximate the state of the system with arbitrary finite accuracy for every time $t \in [0, T]$.

More precisely, let σ be a switching signal for system (1). For $\varepsilon > 0$ and $T \in \mathbb{N}$, we say that $E \subset K$ is an (ε, T)-*spanning set* for (f_σ, K) if for every $\xi \in K$, there is $\eta \in E$ such that $\|x_\sigma(t, \xi) - x_\sigma(t, \eta)\| \leq \varepsilon$ for all $t \in [0, T]$. This means that for every trajectory $x_\sigma(\cdot, \xi)$ of (f_σ, K), there is a trajectory of f_σ starting in $E \subset K$ that is ε-close to $x_\sigma(\cdot, \xi)$ for all $t \in [0, T]$. See Figure 2 for an illustration. We let $s_{\text{span}}(\varepsilon, T; f_\sigma, K)$ be the smallest cardinality of an (ε, T)-spanning set for (f_σ, K).

Definition 2.1. The *topological entropy* of (1) with switching signal σ and initial set K is defined as

$$h(f_\sigma, K) = \lim_{\varepsilon \searrow 0} \limsup_{T \to \infty} \frac{1}{T} \log_2 s_{\text{span}}(\varepsilon, T; f_\sigma, K). \tag{2}$$

The limit on the left is well defined because $s_{\text{span}}(\varepsilon, T; f_\sigma, K)$ is non-increasing in ε.

An equivalent definition of topological entropy can be obtained if instead of considering minimal spanning sets of trajectories, we consider the dual problem, which consists in finding maximal separated sets of trajectories: a set $F \subset K$ is (ε, T)-*separated* for (f_σ, K) if for any two points $\xi, \eta \in F$, there is $t \in [0, T]$ such that

[1] In our framework (worst-case scenario analysis), the switching signal is an external input on which the user has no control, and the objective is to deduce properties of the systems that will be valid for every switching signals.

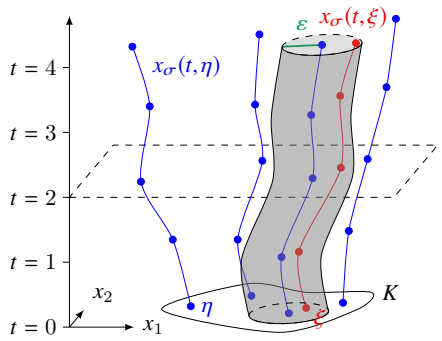

Figure 2: The set of trajectories in blue is (ε, T)-spanning for (f_σ, K) if every trajectory $x_\sigma(\cdot, \xi)$ (e.g., the trajectory represented in red) is contained in the "ε-tube" around at least one of the trajectories in blue for all $t \in [0, T]$.

$\|x_\sigma(t, \xi) - x_\sigma(t, \eta)\| > \varepsilon$. Letting $s_{\text{sep}}(\varepsilon, T; f_\sigma, K)$ be the largest cardinality of an (ε, T)-separated set for (f_σ, K), we obtain an equivalent definition of $h(f_\sigma, K)$:[2]

Proposition 2.2. *The topological entropy of* (1) *with switching signal σ and initial set K satisfies*

$$h(f_\sigma, K) = \lim_{\varepsilon \searrow 0} \limsup_{T \to \infty} \frac{1}{T} \log_2 s_{\text{sep}}(\varepsilon, T; f_\sigma, K). \tag{3}$$

The proof is along the same lines as [27, Theorem 1], and thus omitted here.

2.2 Switched linear systems

In this paper, we are interested in the topological entropy of discrete-time *Switched Linear Systems* (SLSs):

$$x(t+1) = A_{\sigma(t)} x(t), \quad x(t) \in \mathbb{R}^d, \quad t \in \mathbb{N}, \tag{4}$$

where σ is the switching signal as in (1) and $A_i \in \mathbb{R}^{d \times d}$ for all $i \in \Sigma$. SLSs are thus particular instances of switched systems where each mode is linear. Following the lines of Subsection 2.1, we let $x_\sigma(t, \xi)$ be the solution, at time t, of (4) with switching signal σ and initial state $\xi \in \mathbb{R}^d$. We will also use (A_σ, K) to denote system (4) with switching signal σ and initial set K. The system being linear, the transition of the state from a time t_1 to a time t_2 can be represented by a matrix: for $t_1, t_2 \in \mathbb{N}$, $t_1 \leq t_2$, we denote the *fundamental matrix solution* of (4) from t_1 to t_2 with switching signal σ by

$$\Phi_{\sigma, t_1, t_2} = A_{\sigma(t_2-1)} A_{\sigma(t_2-2)} \cdots A_{\sigma(t_1)}. \tag{5}$$

In particular, we have that the trajectories $x_\sigma(\cdot, \xi)$ of (A_σ, K) satisfy $x_\sigma(t_2, \xi) = \Phi_{\sigma, t_1, t_2} x_\sigma(t_1, \xi)$ for all $t_1, t_2 \in \mathbb{N}$, $t_1 \leq t_2$.

The topological entropy of (A_σ, K), denoted by $h(A_\sigma, K)$, is defined in the same way as for switched systems (see Definition 2.1). However, in the case of SLSs, $h(A_\sigma, K)$ does not depend on a particular choice of the initial set $K \subset \mathbb{R}^d$, as long as it is compact with nonempty interior; see, e.g., [40, Proposition 2] for a proof. Therefore, we omit the initial set in the notation and denote by $h(A_\sigma)$ the topological entropy of A_σ.

[2] In general, $s_{\text{span}}(\varepsilon, T; f_\sigma, K) \neq s_{\text{sep}}(\varepsilon, T; f_\sigma, K)$.

We study the worst-case topological entropy of (4), that is, the maximal topological entropy that can be reached by the system among all switching signals. (In formulas, it is convenient to identify SLSs by their sets of modes $\mathcal{M} = \{A_1, \dots, A_N\}$.)

Definition 2.3. The *worst-case topological entropy* of (4) is defined as

$$h_*(\mathcal{M}) = \sup_{\sigma} h(A_\sigma), \tag{6}$$

where the supremum is over all switching signals $\sigma : \mathbb{N} \to \Sigma$.

Computational aspects of worst-case topological entropy are discussed in Section 3. It should also be noted that at this point, there is a priori no link between the topological entropy (which is defined as a topological property of the system) and the minimal data rate for state observation of the system (which involves the notion of coder–decoder). In Section 4, we show that the worst-case topological entropy *is* in fact equal to the minimal data rate for state observation of the system under arbitrary switching signals, and that this data rate limit can be approached as closed as desired by *practical* coders–decoders.

To conclude this section, the example below illustrates the notions of SLS, spanning and separated sets, topological entropy, and worst-case topological entropy.

Example 2.4. Consider the 1-dimensional SLS with two modes $\mathcal{M} = \{A_1, A_2\} \subset \mathbb{R}^{1\times 1}$, $A_1 = 1$ and $A_2 = 2$. Let σ be the switching signal alternating modes 1 and 2: $\sigma = (1, 2, 1, 2, \dots)$. Trajectories of the systems are thus given by $x_\sigma(t, \xi) = 2^{t/2}\xi$ if t is even, and $x_\sigma(t, \xi) = 2^{(t-1)/2}\xi$ if t is odd. We will show that $h(A_\sigma) = 1/2$. For SLSs, as long as topological entropy is concerned, the choice of the initial set is not important. Hence, we fix $K = [0, 1]$.

For $\varepsilon > 0$ and $T \in \mathbb{N}$, let $n = \lceil \varepsilon^{-1} 2^{T/2-1} \rceil$ and $E = \{i/n : 0 \le i \le n\}$. We show that E is (ε, T)-spanning for (A_σ, K). To do this, let $\xi \in K$, and let $\eta \in E$ such that $|\xi - \eta| = \min_{\zeta \in E} |\xi - \zeta|$. Then, by definition of E, $|\xi - \eta| \le 1/(2n) \le \varepsilon 2^{-T/2}$. This implies that for every $t \in [0, T]$, $|x_\sigma(t, \xi) - x_\sigma(t, \eta)| \le 2^{t/2}|\xi - \eta| \le \varepsilon$. Hence, E is (ε, T)-spanning for (A_σ, K) and thus

$$s_{\mathrm{span}}(\varepsilon, T; A_\sigma, K) \le |E| = n + 1 \le \varepsilon^{-1} 2^{T/2-1} + 2.$$

Injecting in (2), we get that $1/2$ is an upper bound on $h(A_\sigma)$:

$$h(A_\sigma) \le \lim_{\varepsilon \searrow 0} \limsup_{T\to\infty} \frac{1}{T} \log_2 \left(\varepsilon^{-1} 2^{T/2-1} + 2\right) = \lim_{\varepsilon \searrow 0} \frac{1}{2} = \frac{1}{2}.$$

Now, in order to show that $1/2$ is also a lower bound on $h(A_\sigma)$, we rely on (ε, T)-separated sets. Let $m = \lceil \varepsilon^{-1} 2^{T/2} \rceil - 1$. Without loss of generality, we may assume $m > 0$ (because we take the limit when $\varepsilon \to 0$ and $T \to \infty$). Define $F = \{i/m : 0 \le i \le m\}$. Then, any two points $\xi, \eta \in F$ satisfy $|\xi - \eta| > \varepsilon 2^{-T/2}$. Hence, for $T \in \mathbb{N}$ even, $|x_\sigma(T, \xi) - x_\sigma(T, \eta)| > \varepsilon$, showing that F is (ε, T)-separated for (A_σ, K). This implies that $s_{\mathrm{sep}}(\varepsilon, T; A_\sigma, K) \ge |F| = m + 1 \ge \varepsilon^{-1} 2^{T/2} - 1$. Injecting in (3), this finally gives that $h(A_\sigma) \ge \log_2 2^{1/2} = 1/2$. Thus, $h(A_\sigma) = 1/2$.

As for the worst-case topological entropy, it is quite intuitive that σ above is not the switching signal that maximizes the topological entropy of the SLS. It is also not difficult to see that the entropy-maximizing sequence is given by using only mode 2: $\sigma = (2, 2, 2, \dots)$. In this case, $x_\sigma(t, \xi) = 2^t \xi$ and we deduce that $h(A_\sigma) = \log_2 2 = 1$. Hence, $h_*(\mathcal{M}) = 1$. □

3 CLOSED-FORM EXPRESSION FOR THE WORST-CASE TOPOLOGICAL ENTROPY OF SWITCHED LINEAR SYSTEMS

We start by presenting a closed-form expression for the worst-case topological entropy of SLSs. This will require concepts from stability analysis of SLSs (namely, the Joint Spectral Radius) and from multilinear algebra (namely, exterior algebras). We also discuss the algorithmic aspects of computing the worst-case topological entropy of SLSs with this expression. Connections with related results in the literature are discussed at the end of this section.

3.1 Joint Spectral Radius

The Joint Spectral Radius (JSR) of a set of matrices measures the asymptotic growth rate of the maximal norm of products of matrices in the set, when the size of the product goes to ∞. More precisely, for a finite set of matrices $\mathcal{M} = \{A_1, \dots, A_N\} \subset \mathbb{R}^{d\times d}$, the *Joint Spectral Radius* of $\mathcal{M}$ is defined as

$$\rho(\mathcal{M}) = \limsup_{t\to\infty} \max_{A_{i_1}, \dots, A_{i_t} \in \mathcal{M}} \|A_{i_1} A_{i_2} \cdots A_{i_t}\|^{1/t}. \tag{7}$$

This quantity was introduced by Rota and Strang in 1960 [33] in order to characterize the stability of SLSs. In particular, the JSR has the following property (for a proof, we refer the reader to [15, Theorem 1.2]): every trajectory $x_\sigma(\cdot, \xi)$ (i.e., for any switching signal σ and any $\xi \in \mathbb{R}^d$) of the SLS associated to $\mathcal{M}$ converges to zero as $t \to \infty$, if and only if the JSR of $\mathcal{M}$ satisfies $\rho(\mathcal{M}) < 1$.

3.2 Exterior algebras

The exterior algebra of a vector space V is an algebraic construction used to study the notions of areas, volumes, and their higher-dimensional analogues in V. In finite dimension, exterior algebras can be constructed from the exterior products of vectors in V. From now on, we assume that $V = \mathbb{R}^d$, and we consider an arbitrary set of k vectors $\{v_1, \dots, v_k\} \subset V$. The *exterior product* of $v_1, \dots, v_k$, denoted by $v_1 \wedge \dots \wedge v_k$, is the k-linear map from V^k to $\mathbb{R}$, defined by

$$(v_1 \wedge \dots \wedge v_k)(w_1, \dots, w_k) = \det \left[w_i^\top v_j\right]_{\substack{1\le i\le k,\\ 1\le j\le k}}, \quad w_i \in V.$$

If $\{v_1, \dots, v_d\}$ is a basis of V, the *kth exterior power* of V, denoted by $\Lambda^k V$, is the vector space spanned by the exterior products of the form $v_{i_1} \wedge v_{i_2} \wedge \dots \wedge v_{i_k}$, $1 \le i_1 < i_2 < \dots < i_k \le d$. In particular, the dimension of $\Lambda^k V$ is $C(k, d) = d!/(k!(d-k)!)$. Note that $\Lambda^k V = \{0\}$ if $k > d$, and by convention we let $\Lambda^0 V = \mathbb{R}$. In numerical computations, it is convenient to treat $\Lambda^k V$ as the coordinate space $\mathbb{R}^{C(k,d)}$. This can be done by fixing a basis $\mathcal{B}$ for $\Lambda^k V$: e.g., $\mathcal{B} = \{e_{i_1} \wedge e_{i_2} \wedge \dots \wedge e_{i_k} : 1 \le i_1 < i_2 < \dots < i_k \le d\}$, where $\{e_1, \dots, e_d\}$ is the canonical basis of $\mathbb{R}^d$. If the elements of $\mathcal{B}$ are ordered with respect to the lexicographical order of their indices $(i_1, \dots, i_k)$, then $\mathcal{B}$ is called the *canonical basis* of $\Lambda^k V$.

If $A \in \mathbb{R}^{d\times d}$, the *$k$th exterior power* of A is the unique linear map $A^{\wedge k} : \Lambda^k V \to \Lambda^k V$ satisfying

$$A^{\wedge k}(v_1 \wedge \dots \wedge v_k) = Av_1 \wedge \dots \wedge Av_k, \quad v_i \in V.$$

By convention, $A^{\wedge 0} = 1$. Using the canonical basis $\mathcal{B}$, $A^{\wedge k}$ can be represented by a $C(k,d) \times C(k,d)$ matrix. Finally, the *exterior power* of A is the $2^d \times 2^d$ matrix $A^\wedge = \operatorname{diag}\{A^{\wedge 0}, \dots, A^{\wedge d}\}$.

The following proposition, whose proof can be found in [2, §3.2.3], summarizes all the properties of exterior powers of maps that we will need in this work.

Proposition 3.1. *Let $k \in \{0, \dots, d\}$ and $A, B \in \mathbb{R}^{d\times d}$.*

(1) $I^{\wedge k} = I$, $(AB)^{\wedge k} = A^{\wedge k}B^{\wedge k}$, $(A^*)^{\wedge k} = (A^{\wedge k})^*$.
(2) *If A is upper-/lower-triangular/ diagonal/ orthogonal, then so is $A^{\wedge k}$ (in the canonical basis of $\Lambda^k V$).*
(3) *The eigenvalues (with multiplicity) of $A^{\wedge k}$ are given by*
$$\lambda_{i_1}(A)\lambda_{i_2}(A)\cdots\lambda_{i_k}(A), \quad 1 \le i_1 < i_2 < \dots < i_k \le d,$$
where $\lambda_i(A)$ is the ith eigenvalue (with multiplicity) of A.
(4) *The singular values of $A^{\wedge k}$ are given by*
$$\bar\sigma_{i_1}(A)\bar\sigma_{i_2}(A)\cdots\bar\sigma_{i_k}(A), \quad 1 \le i_1 < i_2 < \dots < i_k \le d,$$
where $\bar\sigma_i(A)$ is the ith singular value of A.

3.3 Main result and consequences

The main contribution of this section is the following theorem which provides a closed-form expression for the worst-case topological entropy of SLSs.

Theorem 3.2. *The worst-case topological entropy of (4) satisfies*
$$h_*(\mathcal{M}) = \log_2 \rho(\mathcal{M}^\wedge) \tag{8}$$
where $\mathcal{M}^\wedge = \{A_1^\wedge, \dots, A_N^\wedge\}$.

Proof. The proof is presented in Subsection 6.1. □

A straightforward consequence of Theorem 3.2 is that the worst-case topological entropy depends continuously on the SLS: that is, if $\mathcal{M}$ and $\mathcal{N}$ are two finite sets of matrices, that are close to each other (with respect to the Hausdorff distance between sets), then $h_*(\mathcal{M})$ and $h_*(\mathcal{N})$ will also be close to each other. This fact, which is not obvious from the definition of worst-case topological entropy, follows from the continuity of $\mathcal{M}^\wedge$ with respect to $\mathcal{M}$ (easy to see from the definition of $A^\wedge$) and the continuity of the JSR with respect to bounded set of matrices [15, Proposition 1.10].

A wide range of methods, of very different natures, have been proposed in the last decades to evaluate the JSR of a set of matrices; see, e.g., [15, §2.3]. While theoretical discouraging results exist for the computation of the JSR in general, these methods turn out to be extremely powerful in practice and to provide JSR approximation algorithms of high accuracy. Any of these algorithms can be used to approximate the right-hand term of (8). The computation of the exterior power of a matrix is straightforward from its definition, and thus $\mathcal{M}^\wedge$ can be computed in a systematic way. However, it should be noted that the dimension of $\mathcal{M}^\wedge$ increases exponentially with the dimension of the system, and so will the complexity of approximating the JSR of $\mathcal{M}^\wedge$ (this is the curse of dimensionality). In this regard, we note that a simple and algorithm-independent way to substantially speed up the approximation of $\rho(\mathcal{M}^\wedge)$—although not sufficient to fight the curse of dimensionality—is to observe that because the matrices $A_i^\wedge$ are block diagonal, the computation of the JSR of $\mathcal{M}^\wedge$ can be decoupled among the different diagonal blocks [15, Proposition 1.5]: $\rho(\mathcal{M}^\wedge) = \max_{0\le k\le d} \rho(\mathcal{M}^{\wedge k})$.

Furthermore, there are cases for which the computation of the JSR is straightforward. For instance, if $\mathcal{M}$ is a set of normal (or upper-/lower-triangular) matrices, then the JSR is equal to the largest spectral radius of the matrices in $\mathcal{M}$ [15, Propositions 2.2–2.3]. Combining these observations with the properties of the exterior powers (Proposition 3.1), this gives efficient ways to compute the worst-case topological entropy of such sets of matrices.

Corollary 3.3. *Let $\mathcal{M} = \{A_1, \dots, A_N\} \subset \mathbb{R}^{d\times d}$ be a set of normal matrices. For $A_i \in \mathcal{M}$, let $\lambda_1(A_i), \dots, \lambda_d(A_i)$ be its eigenvalues values (with multiplicity) ordered by decreasing modulus. Then*
$$h_*(\mathcal{M}) = \log_2 \max_{\substack{0\le k\le d \\ A_i\in\mathcal{M}}} |\lambda_1(A_i)\cdots\lambda_k(A_i)|.$$
The same holds for sets of upper-/lower-triangular matrices. Moreover, in this case, the eigenvalues are on the diagonal of the A_i's.

Proof. The proof is presented in Subsection 6.2. □

Numerical illustrative examples of the computation of the worst-case topological entropy of SLSs, using Theorem 3.2 and Corollary 3.3, are presented in Subsection 5.1.

To conclude this subsection, we would like to mention another consequence of Theorem 3.2, stated in Corollary 3.4 below. This result is motivated by the fact that in some applications, the switching signal is forced to be periodic (see, e.g., [24]), even if the signal itself or its period is not known beforehand. The corollary states that as long as worst-case topological entropy is concerned, there is no loss of generality in restricting to periodic sequences.

Corollary 3.4. *The worst-case topological entropy of (4) satisfies*
$$h_*(\mathcal{M}) = \sup_{\sigma \text{ periodic}} h(A_\sigma).$$

Proof. The proof is presented in Subsection 6.3. □

3.4 Related works

The worst-case topological entropy provides an upper bound on the topological entropy of A_σ for any switching signal σ. The question of estimating $h(A_\sigma)$ has been addressed, e.g., in [38, 40]. Because the focus is put on a particular matrix sequence A_σ (disregarding other sequences), the bounds on $h(A_\sigma)$ obtained in [38, 40] are in general better than the worst-case topological entropy. However, in "ill-conditioned" cases (e.g., triangular systems with large values at the $(1,1)$-entries), the bounds available in [38, 40] can be quite conservative. In these cases, it can be beneficial to use $h_*(\mathcal{M})$ as an upper bound on $h(A_\sigma)$.

As already mentioned, the Joint Spectral Radius is a cornerstone of SLS theory, and has attracted a lot of attention in the last decades [15, 33, 37]. As a measure of the *worst-case* asymptotic growth rate of the trajectories of the system, it is not surprising to encounter this quantity in the characterization of the worst-case topological entropy.

Exterior algebras have also received attention in dynamical systems theory; in particular, in the study of the Lyapunov exponents [2, 4] and entropy-related properties [18, 23] of dynamical systems and control systems. For instance, we note the remarkable formula

by Kozlovski [23] for the topological entropy of a discrete-time autonomous dynamical system, described by a C^∞ map $f : X \to X$, on a compact Riemannian manifold:

$$h(f) = \lim_{n\to\infty} \frac{1}{n} \log_2 \int_X \|(Df_x^n)^\wedge\| \, \mathrm{d}x.$$

4 PRACTICAL CODER–DECODER

We investigate the problem of state observation of SLSs through communication networks with limited data rate. The situation is depicted in Figure 1. At specific sampling/transmission times, $0 = T_0 < T_1 < T_2 < \dots$ ($T_j \in \mathbb{N}$), a *coder* measures the state $x(T_j)$ of the system, and is connected to a *decoder* via a noiseless digital channel which can carry one discrete-valued symbol $e(T_j)$ per time epoch $[T_j, T_{j+1}]$, selected from a coding alphabet $\mathcal{E}_j$ of time-varying size. Neglecting the propagation delay and transmission errors, each symbol takes at most one epoch duration $T_{j+1} - T_j$ to be completely transmitted. Hence, at time T_{j+1}, the decoder has $e(T_0), \dots, e(T_j)$ available and generates estimates $\hat{x}(t)$ of the state of the system for the ongoing epoch $[T_{j+1}, T_{j+2})$.

More precisely, the coder is a family of functions $C[\cdot,\cdot,\cdot \mid j, \sigma]$ (parameterized by j the index of the epoch and σ the switching signal of the system):

$$\begin{aligned} &C[\cdot,\cdot,\cdot \mid j,\sigma] : \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}_{>0} \to \mathcal{E}_j, \\ &e(T_j) = C[x(T_j), \hat{x}(T_j), \delta_j \mid j, \sigma], \quad j \in \mathbb{N}, \end{aligned} \tag{9}$$

where $\hat{x}(T_j) \in \mathbb{R}^d$ is an estimate of the current state $x(T_j)$, and δ_j satisfies $\|x(T_j) - \hat{x}(T_j)\| \le \delta_j$. The output is $e(T_j) \in \mathcal{E}_j \subset \mathbb{R}^d$ where $\mathcal{E}_j$ is a finite set depending on j and σ. The symbol $e(T_j)$ will be transmitted to the decoder at most at T_{j+1}. The decoder is a family of functions $\mathcal{D}[\cdot,\cdot,\cdot \mid j, \sigma]$ (also parameterized by j and σ):

$$\begin{aligned} &C[\cdot,\cdot,\cdot \mid j,\sigma] : \mathcal{E}_{j-1} \times \mathbb{R}^d \times \mathbb{R}_{>0} \to (\mathbb{R}^d)^{T_{j+1}-T_j}, \\ &[\hat{x}(T_j), \dots, \hat{x}(T_{j+1}-1)] = \\ &\qquad \mathcal{D}[e(T_{j-1}), \hat{x}(T_{j-1}), \delta_{j-1} \mid j, \sigma], \quad j \in \mathbb{N}, \end{aligned} \tag{10}$$

where $\hat{x}(T_{j-1})$ satisfies $\|x(T_{j-1}) - \hat{x}(T_{j-1})\| \le \delta_{j-1}$, and $e(T_{j-1})$ is the symbol transmitted at T_{j-1} and received by the decoder at T_j. If $j = 0$, take $e(T_{-1}) = \hat{x}(T_{-1}) = 0$. The decoder outputs estimates of the state for the ongoing epoch $[T_j, T_{j+1})$. The sampling/transmission times T_j, the error bounds δ_j and the coding alphabets $\mathcal{E}_j$ depend only on the switching signal σ and thus they can be computed by both the coder and the decoder independently. The *data rate* R (in bits per unit of time) of the coder–decoder is defined as

$$R = \sup_\sigma \sup_{j\in\mathbb{N}} \frac{\log_2 |\mathcal{E}_j|}{T_{j+1} - T_j}. \tag{11}$$

We want to build coders–decoders that estimate the state of the system with exponentially decreasing error.

Definition 4.1. The coder–decoder (9)–(10) is said to *observe* the SLS (4) with initial set K if there exist $C > 0$ and $g \in (0, 1)$ such that for every switching signal σ and initial state $\xi \in K$, it holds that

$$\|x_\sigma(t, \xi) - \hat{x}(t)\| \le C g^t, \quad \forall\, t \in \mathbb{N}. \tag{12}$$

Remark 4.1. (9)–(10) assume that the whole switching signal is known by the coder and the decoder during its operation (see also the Introduction for the relevance of this assumption). In fact, as it will be clear from the implementation of the coder and decoder (see paragraphs below), it is sufficient that only the $T_{j+1} - T_j$ modes that are being used during the ongoing epoch $[T_j, T_{j+1}]$ are known by the coder and the decoder.

We describe a family of practical coders–decoders that observe (4) and whose data rate can be as close as desired to the worst-case topological entropy of the system. More precisely, for any compact set $K \subset \mathbb{R}^d$ and $R' > h_*(\mathcal{M})$, there is such a coder–decoder that observes (4) with initial set K and whose data rate satisfies $R \le R'$ (see also Theorem 4.2 at the end of this section). The construction relies on the properties of the Joint Spectral Radius and exterior powers of matrices to build coders–decoders with data rate as close as desired to the right-hand term of (8).

Coder–decoder's implementation. For $r > 0$, we let

$$I(r) = \begin{cases} \mathbb{Z}_{\text{even}} \cap [1 - \lceil r \rceil, \lceil r \rceil - 1] & \text{if } \lceil r \rceil \text{ is odd,} \\ \mathbb{Z}_{\text{odd}} \cap [1 - \lceil r \rceil, \lceil r \rceil - 1] & \text{if } \lceil r \rceil \text{ is even,} \end{cases}$$

where $\mathbb{Z}_{\text{even}}$ ($\mathbb{Z}_{\text{odd}}$) is the set of even (odd) integers. By construction, for every $\xi \in [-r, r]$, there is $\eta \in I(r)$ such that $|\xi - \eta| \le 1$, and we have that $|I(r)| \le \lceil r \rceil$. If $r = 0$, we let $I(0) = \{0\}$. Now, for $r_1, \dots, r_d \ge 0$, we define

$$\mathrm{Grid}(r_1, \dots, r_d) = I(r_1) \times \dots \times I(r_d).$$

Its cardinality satisfies $\hat{\mu} := |\mathrm{Grid}(r_1, \dots, r_d)| \le \prod_{i=1}^d \lceil r_i \rceil^*$ where $\lceil \alpha \rceil^* = \max\{\lceil \alpha \rceil, 1\}$. Finally, for $\xi \in \mathbb{R}^d$, we let $Q(\xi)$ be the closest point to ξ in $\mathrm{Grid}(r_1, \dots, r_d)$. Hence, $Q(\cdot)$ is a $\hat{\mu}$-points quantizer and satisfies

$$\|\xi - Q(\xi)\| \le d^{1/2}, \quad \forall\, \xi \in [-r_1, r_1] \times \dots \times [-r_d, r_d].$$

Let $K \subset \mathbb{R}^d$ be a compact set and fix a target data rate R' strictly larger than the right-hand side of (8). We will build a coder–decoder that observes the SLS (4) with initial set K and whose data rate satisfies $R \le R'$. (The reader may find useful to refer to Figure 3, where the different quantities appearing in the definition of the coder–decoder are represented.)

Let $T_0 = 0$. Also, let $\hat{x}(0)$ be an estimation of the initial state and $\delta_0 \ge 0$ be such that $K \subseteq B(\hat{x}(0), \delta_0)$. Fix $\alpha > 1$.

- At time T_j, the values of T_{j+1}, $\mathcal{E}_j$, and δ_{j+1} are computed as follows (these computations are carried out by both the coder and the decoder independently). For $T \in \mathbb{N}$, $T > T_j$, we let $\bar{\sigma}_i(\Phi_{\sigma,T_j,T})$, $i \in \{1, \dots, d\}$, be the singular values of the fundamental matrix solution $\Phi_{\sigma,T_j,T}$ (defined in (5)). We define T_{j+1} as the smallest $T \in \mathbb{N}$, $T > T_j$, satisfying

$$\prod_{i=1}^d \lceil \alpha d^{1/2} \bar{\sigma}_i(\Phi_{\sigma,T_j,T}) \rceil^* \le 2^{(T-T_j)R'}. \tag{13}$$

Finally, T_{j+1} being fixed, we let

$$\mathcal{E}_j = \mathrm{Grid}(r_1, \dots, r_d), \quad r_i = \alpha d^{1/2} \bar{\sigma}_i(\Phi_{\sigma,T_j,T_{j+1}}), \tag{14}$$

and[3]

$$\delta_{j+1} = \delta_j / \alpha. \tag{15}$$

[3] That is, the bound δ_j on the estimation error $\|x_\sigma(T_j, \xi) - \hat{x}(T_j)\|$ decreases by a factor $1/\alpha$ between two sampling/transmission times T_j and T_{j+1}. And thus the rate g of decay of the estimation error (see Definition 4.1) is given by $\alpha^{-1/\tau}$ where τ is an upper bound on $T_{j+1} - T_j$; see also Subsection 7.1.

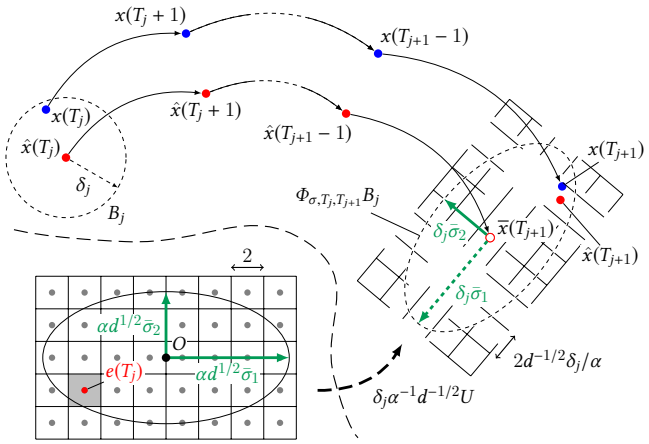

Figure 3: The different quantities appearing in the definition of the coder–decoder. The gray dots (bottom left) represent $\mathrm{Grid}(r_1, r_2)$**, where** $\bar{\sigma}_1$**,** $\bar{\sigma}_2$ **are the singular values of** $\Phi_{\sigma,T_j,T_{j+1}}$**. According to (17),** $\mathrm{Grid}(r_1, r_2)$ **is scaled and rotated by** $\delta_{j+1}d^{-1/2}U$**, and centered at** $\bar{x}(T_{j+1}) \coloneqq \Phi_{\sigma,T_j,T_{j+1}}\hat{x}(T_j)$**. The latter is the best available estimate of** $x(T_{j+1})$ **before the reception of the symbol** $e(T_j)$**. At reception of** $e(T_j)$**, the new estimate** $\hat{x}(T_{j+1})$ **is then given by the center of the square in which the state** $x(T_{j+1})$ **lies.**

- The coder is defined as follows. At time T_j, if $\hat{x}(T_j)$ is the current estimate of the state (stored in the memory of the coder) and $x(T_j)$ is the current state of the system (the coder has access to the plant), we let $\Delta\xi_j = x(T_j) - \hat{x}(T_j)$. Let USV^* be the Singular Value Decomposition of $\Phi_{\sigma,T_j,T_{j+1}}$, where the singular values on the diagonal of S are in the same order as in (14). The symbol sent by the coder at time T_j is then defined as

$$e(T_j) = Q(d^{1/2}SV^*\Delta\xi_j/\delta_{j+1}) \in \mathcal{E}_j. \tag{16}$$

where $Q(\cdot)$ is the quantizer with respect to $\mathcal{E}_j = \mathrm{Grid}(r_1, \ldots, r_d)$.

- The decoder is defined as follows. At time T_j, the decoder receives the symbol $e(T_{j-1})$ and has the last estimate $\hat{x}(T_{j-1})$ in memory. Then, the decoder computes the estimates $\hat{x}(\cdot)$ for the ongoing epoch $[T_j, T_{j+1})$ as follows:

$$\hat{x}(T_j) = \Phi_{\sigma,T_{j-1},T_j}\hat{x}(T_{j-1}) + \delta_j U e(T_{j-1})/d^{1/2}, \tag{17}$$

where USV^* is the Singular Value Decomposition of $\Phi_{\sigma,T_{j-1},T_j}$, as above. If $T_j = 0$, simply use the initial estimate $\hat{x}(0)$ given in the parameters of the coder–decoder. Next, define inductively

$$\hat{x}(t) = A_{\sigma(t-1)}\hat{x}(t-1), \quad t \in (T_j, T_{j+1}), \tag{18}$$

The implementation of the above coder–decoder is described in Algorithms 1–2. The proof of its correctness, i.e., that it satisfies (12), is presented in Subsection 7.1. Summarizing, we have the following result on the equivalence of the worst-case topological entropy and the minimal data rate for state observation of the system under arbitrary switching signals.

Theorem 4.2. *Let $K \subset \mathbb{R}^d$ be compact, and consider the SLS (4). If $R' < h_*(\mathcal{M})$, then there is no coder–decoder with data rate $R \le R'$ that observes (4). If $R' > h_*(\mathcal{M})$, then there is a practical coder–decoder with data rate $R \le R'$ that observes (4).*

Proof. The proof is presented in Subsection 7.2. □

Algorithm 1: Coder

Input: R, α, $\hat{x}(0)$, δ_0, and σ.
Let $j = 0$ and $T_j = 0$;
loop *wait* $t_{\text{real}} = T_j$ `/* t_real is the real time */`
  Measure current state $x(T_j)$ of the system;
  Compute T_{j+1}, $\mathcal{E}_j$, and δ_{j+1} as in (13)–(15);
  Compute $e(T_j)$ as in (16) and send $e(T_j)$ to the decoder;
  Let $j = j + 1$;
  Compute $\hat{x}(T_j)$ as in (17);
end loop

Algorithm 2: Decoder

Input: R, α, $\hat{x}(0)$, δ_0, and σ.
Let $j = 0$ and $T_j = 0$;
loop *wait* $t_{\text{real}} = T_j$ `/* t_real is the real time */`
  Receive symbol $e(T_{j-1})$; `// If j = 0, e(T_-1) = 0.`
  Compute T_{j+1}, $\mathcal{E}_j$, and δ_{j+1} as in (13)–(15);
  Compute $\hat{x}(T_j)$ as in (17); `// If j = 0, x̂(T_0) = x̂(0).`
  Compute $\hat{x}(t)$ for $t \in (T_j, T_{j+1})$ as in (18);
  Let $j = j + 1$;
end loop

5 NUMERICAL EXPERIMENTS

5.1 Worst-case topological entropy

We use the results of Section 3 to compute the worst-case topological entropy of SLSs with general and triangular matrices.

Example 5.1. Consider the set of 2×2 matrices

$$\mathcal{M} = \left\{ A_1 \coloneqq \begin{bmatrix} -0.1 & 0.7 \\ -1.3 & 0.2 \end{bmatrix}, \quad A_2 \coloneqq \begin{bmatrix} 0.1 & 1.3 \\ -0.7 & 0.5 \end{bmatrix} \right\}. \tag{19}$$

We use Theorem 3.2 to compute the worst-case topological entropy of the SLS associated with $\mathcal{M}$. Therefore, we compute the exterior powers of A_1 and A_2:

$$A_1^{\wedge} = \begin{bmatrix} 1 & & & \\ & -0.1 & 0.7 & \\ & -1.3 & 0.2 & \\ & & & 0.89 \end{bmatrix}, \quad A_2^{\wedge} = \begin{bmatrix} 1 & & & \\ & 0.1 & 1.3 & \\ & -0.7 & 0.5 & \\ & & & 0.96 \end{bmatrix}.$$

We have used the JSR Toolbox [37] (in MATLAB) to compute the JSR of $\mathcal{M}^{\wedge}$: this gives $\rho(\mathcal{M}^{\wedge}) = 1.2379$. Hence, we conclude that $h_*(\mathcal{M}) = 0.3079$. Observe that the worst-case topological entropy is nonzero even if A_1 and A_2 are both stable matrices (indeed their spectral radii are given by 0.94 and 0.98 respectively). □

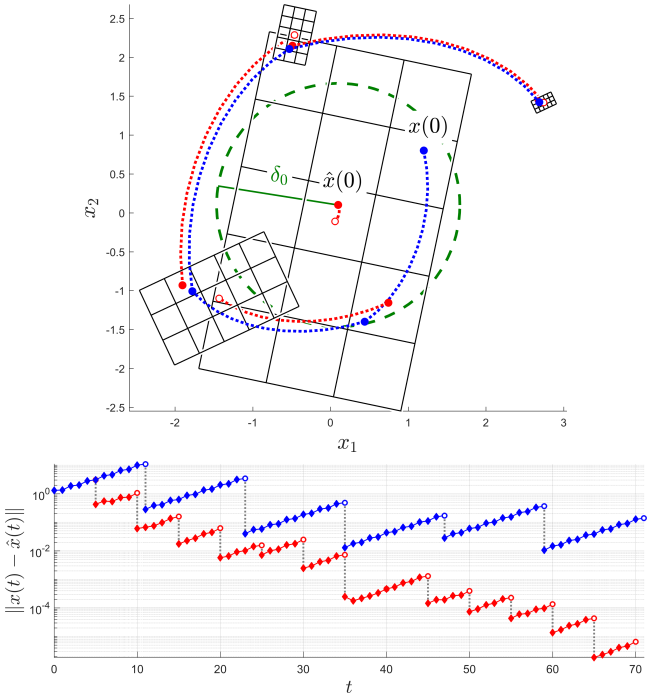

Figure 4: Top: Evolution of $x(t)$ and $\hat{x}(t)$ for a sample execution of the coder–decoder with data rate $R = 4$. Bottom: Evolution of the estimation error $\|x(t) - \hat{x}(t)\|$ for sample executions of the coder–decoder with data rates $R = 0.8$ (red) and $R = 0.5$ (blue). The vertical gray lines indicate the sampling/transmission times T_j (in general, the error decreases at these times because the decoder receives a new symbol).

Example 5.2. Consider the set of 2×2 matrices

$$\mathcal{M} = \left\{ A_1 := \begin{bmatrix} 3 & 1 \\ & 0.1 \end{bmatrix}, \quad A_2 := \begin{bmatrix} 1.1 & 1 \\ & 2 \end{bmatrix} \right\}.$$

Because A_1 and A_2 are upper-triangular, we may apply Corollary 3.3 (note that the eigenvalues of A_1 and A_2 are on their diagonal). We deduce that the worst-case topological entropy of the SLS associated with $\mathcal{M}$ is equal to $\log_2 3 = 1.5850$. The reader will check that the same result can be obtained by applying directly Theorem 3.2; indeed the exterior powers of A_1 and A_2 are given by

$$A_1^{\wedge} = \begin{bmatrix} 1 & & & \\ & 3 & 1 & \\ & & 0.1 & \\ & & & 0.3 \end{bmatrix}, \quad A_2^{\wedge} = \begin{bmatrix} 1 & & & \\ & 1.1 & 1 & \\ & & 2 & \\ & & & 2.2 \end{bmatrix},$$

and the JSR of upper-triangular matrices is given by the largest absolute value of its diagonal entries [15, Proposition 2.3]. □

5.2 State observation with limited data rate

We apply the coder–decoder described in Section 4 for the state observation of the SLS (19). The parameters of the coder–decoder (see Algorithms 1–2) are set as follows: we fix the value $\alpha = 2.5$; the values of $\hat{x}(0)$ and δ_0 are given in Figure 4 (top); and we use different values for the maximal data rate R, as explained below.[4]

Firstly, to make a comprehensive visual illustration, we use a data rate of $R = 4$, which is much larger than the worst-case topological entropy (cf. Example 5.1). This ensures that each epoch lasts one unit of time, i.e., $T_{j+1} - T_j = 1$. A sample execution of the coder–decoder is presented in Figure 4 (top). In this picture, the states $x(t)$ of the true system are represented in blue. The estimates $\hat{x}(t)$ computed by the coder–decoder are represented in red.

In a second time, we simulate the execution of the coder–decoder with data rates that are closer to the worst-case topological entropy of the system, namely $R = 0.8$ and $R = 0.5$. For these values of R, the duration of the epochs are longer (between 5 and 12, in our simulation). The evolution of the estimation error $\|x(t) - \hat{x}(t)\|$ is represented in Figure 4 (bottom). As expected, we observe that the estimation error decreases more rapidly when the data rate is higher.

6 PROOFS OF SECTION 3

6.1 Proof of Theorem 3.2

Consider the SLS (4) with set of matrices $\mathcal{M}$ and denote by $R(\mathcal{M})$ the right-hand term of (8). The proof is divided into two parts: first we prove that $R(\mathcal{M})$ is an upper bound on $h_*(\mathcal{M})$ and then we prove that it is also a lower bound.

Part 1: Let $R' > R(\mathcal{M})$. We will show that $h_*(\mathcal{M}) \leq R'$. The most natural way to do this would probably be to use an argument based on the definition of topological entropy in terms of spanning sets (see Definition 2.1). However, we will present an alternative proof[1], which draws on the fact that we have described a coder–decoder for the state observation of the system (see Section 4). The proof is based on the result below (Lemma 6.1), which is well known (with slight modifications) for autonomous systems [30, Theorem 2.3.6], and certainly also for non-autonomous systems because the proof argument is identical. For the sake of completeness, the proof of this result is presented below.

Lemma 6.1. *Consider the SLS (4) with compact initial set K. Assume there is a coder–decoder (9)–(10) with data rate R that observes (4) with initial set K. Then, the worst-case topological entropy of (4) satisfies $h_*(\mathcal{M}) \leq R$.*

Proof of Lemma 6.1. Fix a switching signal σ. We will prove that $h(A_\sigma, K) \leq R$. The proof is based on the definition of topological entropy in terms of separated sets of trajectories (Proposition 2.2). Let $\varepsilon > 0$. We will show that there is $D > 0$ such that for all $T \in \mathbb{N}$, $s_{\text{sep}}(\varepsilon, T; A_\sigma, K) \leq D2^{TR}$. By (3), and because ε is arbitrary, this will prove that $h(A_\sigma, K) \leq R$.

[4] It is worth noticing the following effects of the parameters α and R on the output of the coder–decoder. By (15), α gives the rate of decrease of the worst-case estimation error at the sampling/transmission times T_j. On the other hand, by (13), for a fixed $\alpha > 0$, the maximal length $T_{j+1} - T_j$ of an epoch will depend on the maximal allowed data rate R of the coder–decoder; the smaller R, the longer $T_{j+1} - T_j$. Furthermore, if R is smaller than the worst-case topological entropy—and only in this case—, the maximal epoch length may be infinite.

[1] The two proofs are of course intrinsically related. With the coder–decoder, we build a set of functions approximating the trajectories of the system. Because these functions are not trajectories themselves, this set is not a spanning set as defined in Subsection 2.1, but the two concepts are akin.

To do this, let $S \in \mathbb{N}$ be such that $Cg^{S+1} \leq \varepsilon/2$ where C and g are as in (12). Define $D = s_{\text{sep}}(\varepsilon, S; f_\sigma, K)$. Now, assume for a contradiction there is $T \in \mathbb{N}$ and an (ε, T)-separated set F for (A_σ, K), with $|F| > D2^{TR}$. By definition (11) of the data rate, the sequence of symbols $\{e(T_0), \ldots, e(T_j)\}$ received by the decoder at time T (i.e., $T \in [T_{j+1}, T_{j+2})$) can take at most $2^{T_{j+1}R} \leq 2^{TR}$ different values. Thus, the sequence of estimates $\{\hat{x}(0), \ldots, \hat{x}(T)\}$, which depends only on $\hat{x}(0)$ and the past symbols, can also take at most 2^{TR} different values. Hence, since $|F| > D2^{TR}$, there is a subset $F' \subseteq F$ with cardinality $|F'| > D$, such that all trajectories of f_σ starting from F' have the same sequence of estimates on $[0, T]$, say $\{\hat{x}(0), \ldots, \hat{x}(T)\}$. By definition of S, this implies that for any two points $\xi, \eta \in F'$ and every $t \in (S, T]$, it holds that

$$\|x_\sigma(t, \xi) - x_\sigma(t, \eta)\| \leq \|x_\sigma(t, \xi) - \hat{x}(t)\| + \|x_\sigma(t, \eta) - \hat{x}(t)\| \leq \varepsilon.$$

Hence, because F' is (ε, T)-separated (as a subset of F), there must be $T \in [0, S]$ such that $\|x_\sigma(t, \xi) - x_\sigma(t, \eta)\| > \varepsilon$. This implies that F' is in fact (ε, S)-separated, a contradiction with the definition of D. Thus, $h(A_\sigma, K) \leq R$. Now, because the topological entropy of A_σ is independent from the initial set, we have that $h(A_\sigma) \leq R$, and since σ is arbitrary, it follows that $h_*(\mathcal{M}) \leq R$, concluding the proof of the lemma. □

By applying Lemma 6.1 to the coder–decoder described in Section 4 that observes the SLS (4) with initial set K (see Section 7.1 for the proof) and operating at data rate R', we deduce that $h(\mathcal{M}) \leq R'$. Finally, since R' is arbitrary, this proves that $h_*(\mathcal{M}) \leq R(\mathcal{M})$, concluding the proof of Part 1.

Part 2: We will show that $R(\mathcal{M}) \leq h_*(\mathcal{M})$. The proof is based on the the following result, known as the "Joint Spectral Radius Theorem", whose proof can be found in [15, Theorem 2.3].

Proposition 6.2. *If $\mathcal{M} = \{A_1, \ldots, A_N\} \subset \mathbb{R}^{d\times d}$ is a finite set of matrices, then*

$$\rho(\mathcal{M}) = \limsup_{t\to\infty} \max_{A_{i_1}, \ldots, A_{i_t} \in \mathcal{M}} \bar{\rho}(A_{i_1} A_{i_2} \cdots A_{i_t})^{1/t}$$

where $\bar{\rho}(M)$ is the spectral radius of M, i.e., the largest modulus of its eigenvalues.

We proceed with the proof of Part 2. Therefore, let $R' < R(\mathcal{M})$. From Proposition 6.2, there is $T \in \mathbb{N}_{>0}$ and $M = A_{i_1} \cdots A_{i_T}$ $(A_{i_1}, \ldots, A_{i_T} \in \mathcal{M})$ such that $\bar{\rho}(M^\wedge) \geq 2^{TR'}$. Because $M^\wedge$ is block diagonal, its spectral radius is given by the maximal spectral radius of the diagonal blocks. It follows that there is $k \in \{0, \ldots, d\}$ such that $\bar{\rho}(M^{\wedge k}) \geq 2^{TR'}$. By Proposition 3.1 (item 3), this is equivalent to $|\lambda_1(M) \cdots \lambda_k(M)| \geq 2^{TR'}$ where $\lambda_1(M), \ldots, \lambda_d(M)$ are the eigenvalues of M ordered by decreasing modulus. On the other hand, it is well known [30, Theorem 2.4.2] that the topological entropy of the LTI system $x(t+1) = Mx(t)$, denoted by $h(M)$, satisfies $h(M) = \sum_{j=1}^{d} \log_2 \max\{|\lambda_j(M)|, 1\}$. Hence, $h(M) \geq TR'$.

Let σ be the periodic switching signal obtained by repeating the sequence of modes $(i_T, \ldots, i_1)$ where $i_1, \ldots, i_T$ are as in the definition of M above. Then, the topological entropy of A_σ satisfies $h(A_\sigma) \geq R'$. Indeed, if $K \subset \mathbb{R}^d$ is a compact initial set and $m \in \mathbb{N}$, then any (ε, mT)-spanning set for (A_σ, K) is an (ε, m)-spanning set for the LTI system given by M. Hence, $s_{\text{span}}(\varepsilon, mT; A_\sigma, K) \geq s_{\text{span}}(\varepsilon, m; M, K)$, implying that $h(A_\sigma, K) \geq h(M, K)/T \geq R'$. Finally, this implies that $h_*(\mathcal{M}) \geq R'$, and because R' is arbitrary, it follows that $h_*(\mathcal{M}) \geq R(\mathcal{M})$. This concludes the proof of Part 2.

6.2 Proof of Corollary 3.3

This follows from the following result, whose proof can be found in [15, Proposition 2.2–2.3].

Proposition 6.3. *If $\mathcal{M}$ is a set of normal (or upper-/lower-triangular) matrices, then $\rho(\mathcal{M})$ is equal to the largest spectral radius of the matrices in $\mathcal{M}$.*

From Proposition 3.1 (item 1–2), it holds that if A is normal, then $A^\wedge$ is normal as well. Indeed, if $A = UDU^*$ where U is orthogonal and D is diagonal, then $A^\wedge = U^\wedge D^\wedge (U^\wedge)^*$ and $U^\wedge$ is orthogonal and $D^\wedge$ is diagonal. Hence, by Proposition 6.3, $\rho(\mathcal{M}^\wedge) = \bar{\rho}(A_i^\wedge)$ for some $A_i \in \mathcal{M}$, where $\bar{\rho}(M)$ is the spectral radius of M, i.e., the largest modulus of its eigenvalues. Because $A_i^\wedge$ is block diagonal, its spectral radius is given by the maximal spectral radius of the diagonal blocks, i.e., $\bar{\rho}(A_i^\wedge) = \bar{\rho}(A_i^{\wedge k})$ for some $k \in \{0, \ldots, d\}$. Using Proposition 3.1 (item 3), this gives the desired result.

The proof for sets of upper-/lower-triangular matrices is identical, and thus omitted.

6.3 Proof of Corollary 3.4

The proof is straightforward from Part 2 in the proof of Theorem 3.2 (see Subsection 6.1). Indeed, for a given $R' < h_*(\mathcal{M})$, we showed the existence of a *periodic* switching signal σ satisfying $h(A_\sigma) \geq R'$. Because, R' is arbitrary, this concludes the proof.

7 PROOFS OF SECTION 4

7.1 Proof of the correctness of the coder–decoder presented in Section 4

The proof is divided into two parts: first, we prove that there is an absolute upper bound on the epoch duration $T_{j+1} - T_j$. This upper bound depends only on the SLS and the data rate, as long as it is strictly larger than the right-hand side of (8). Then, we prove that the estimation error at T_j is smaller than δ_j. Because δ_j decreases exponentially (with rate α^{-1}) with respect to j, this will imply that the estimation error $\|x(t) - \hat{x}(t)\|$ decreases exponentially with respect to t.

Part 1: Fix $\alpha > 1$. Let R' be strictly larger than the right-hand side of (8). Then, by the definition of the JSR, there is $\tau \in \mathbb{N}_{>0}$ large enough such that

$$\max_{A_{i_1}, \ldots, A_{i_\tau} \in \mathcal{M}} \|(A_{i_1} A_{i_2} \cdots A_{i_\tau})^\wedge\| \leq 2^{-d} \alpha^{-d} d^{-d/2} 2^{\tau R'}.$$

This τ will be an upper bound on $T_{j+1} - T_j$. Indeed, let σ be any switching signal and let $T = T_j + \tau$. The left-hand side of (13) satisfies

$$\prod_{i=1}^{d} \lceil \alpha d^{1/2} \bar{\sigma}_i(\Phi_{\sigma, T_j, T}) \rceil^* \leq \prod_{i=1}^{d} 2\alpha d^{1/2} \max\{\bar{\sigma}_i(\Phi_{\sigma, T_j, T}), 1\} \quad (20)$$

On the other hand, by definition of τ, we have that

$$\|\Phi_{\sigma, T_j, T}{}^\wedge\| \leq 2^{-d} \alpha^{-d} d^{-d/2} 2^{\tau R'}.$$

By Proposition 3.1 (item 4) and the block-diagonal structure of $\Phi_{\sigma,T_j,T}{}^{\wedge}$, the inequality above is equivalent to

$$\max_{0\le k\le d} \bar{\sigma}_1(\Phi_{\sigma,T_j,T})\cdots\bar{\sigma}_k(\Phi_{\sigma,T_j,T}) \le 2^{-d}\alpha^{-d}d^{-d/2}2^{\tau R'}.$$

With (20), this gives that the left-hand side of (13) satisfies

$$\prod_{i=1}^{d}\lceil \alpha d^{1/2}\bar{\sigma}_i(\Phi_{\sigma,T_j,T})\rceil^* \le 2^{\tau R'},$$

and thus τ is an upper bound of $T_{j+1}-T_j$.

Part 2: We prove, by induction on $j\in\mathbb{N}$, that the estimation error at T_j satisfies $\|x(T_j)-\hat{x}(T_j)\|\le\delta_j$ for all $j\in\mathbb{N}$. By definition of T_0, $\hat{x}(0)$ and δ_0, the claim is satisfied for $j=0$. Now, for the induction step, assume that it is satisfied for some $j\in\mathbb{N}$. This implies that $\Delta\xi_j := x(T_j)-\hat{x}(T_j)\in B(0,\delta_j)$. In the view of this, let us look at the definition (16) of $e(T_j)$. Therefore, let USV^* be the Singular Value Decomposition of $\Phi_{\sigma,T_j,T_{j+1}}$.

We have that $V^*\Delta\xi_j/\delta_{j+1}\in B(0,\alpha)$ because V^* is orthogonal and $\delta_{j+1}=\delta_j/\alpha$. It follows that $d^{1/2}SV^*\Delta\xi_j/\delta_{j+1}$ belongs to $[-r_1,r_1]\times\ldots\times[-r_d,r_d]$ where r_i are defined as in (14). By definition of $Q(\cdot)$, this implies that

$$\|d^{1/2}SV^*\Delta\xi_j/\delta_{j+1}-e(T_j)\|\le d^{1/2}. \tag{21}$$

Now, let us look at the definition (17) of $\hat{x}(T_{j+1})$ at time T_{j+1}. The difference between $x(T_{j+1})$ and $\hat{x}(T_{j+1})$ satisfies

$$\begin{aligned}
x(T_{j+1})-\hat{x}(T_{j+1}) &= \Phi_{\sigma,T_j,T_{j+1}}x(T_j)\\
&\quad -\Phi_{\sigma,T_j,T_{j+1}}\hat{x}(T_j)-\delta_{j+1}Ue(T_j)/d^{1/2}\\
&= \Phi_{\sigma,T_j,T_{j+1}}\Delta\xi_j-\delta_{j+1}Ue(T_j)/d^{1/2}\\
&= USV^*\Delta\xi_j-\delta_{j+1}Ue(T_j)/d^{1/2}\\
&= \delta_{j+1}U(d^{1/2}SV^*\Delta\xi_j/\delta_{j+1}-e(T_j))/d^{1/2}.
\end{aligned}$$

By (21) and the orthogonality of U, the latter implies that $\|x(T_{j+1})-\hat{x}(T_{j+1})\|\le\delta_{j+1}$, concluding the proof of Part 2.

Now, combining Part 1 and Part 2, we have that the estimates $\hat{x}(t)$ satisfy (12) with $g=\alpha^{-1/\tau}$ and $C=\delta_0 D$ where

$$D=\max_{\substack{0\le T\le\tau-1\\ A_{i_1},\ldots,A_{i_T}\in\mathcal{M}}} g^{-T}\|A_{i_1}\cdots A_{i_T}\|.$$

Indeed, if $t\in[T_j,T_{j+1})$ for some $j\in\mathbb{N}$, then

$$\begin{aligned}
\|x(t)-\hat{x}(t)\| &= \|\Phi_{\sigma,T_j,t}(x(T_j)-\hat{x}(T_j))\|\\
&\le \|\Phi_{\sigma,T_j,t}\|\,\|x(T_j)-\hat{x}(T_j)\| \le Dg^{t-T_j}\|x(T_j)-\hat{x}(T_j)\|\\
&\le Dg^{t-T_j}\delta_j = Dg^{t-T_j}\delta_0\alpha^{-j} \le Dg^{t-T_j}\delta_0\alpha^{-T_j/\tau}\\
&= Dg^{t-T_j}\delta_0 g^{T_j} = D\delta_0 g^t.
\end{aligned}$$

This concludes the proof of the correctness of the coder–decoder.

7.2 Proof of Theorem 4.2

The first implication is Lemma 6.1. The second one follows from the existence of the family of coders–decoders presented in Section 4, which can have data rates arbitrarily close to the worst-case topological entropy.

8 CONCLUSION

This paper introduced the concept of worst-case topological entropy for switched linear systems. It was shown that this quantity is relevant for the problem of state observation of these systems with limited data rate. More precisely, we constructed a practical coder–decoder, operating at a data rate as close as desired to the worst-case topological entropy, that estimates the state of the system for any switching signal and with exponentially decreasing estimation error. We also discussed the computational aspects of the worst-case topological entropy. In particular, we provided a closed-form formula expressing the worst-case topological entropy as the Joint Spectral Radius of some set of matrices obtained from the original one by taking exterior powers. Among other consequences, the computation of the worst-case topological entropy can thereby benefit from the numerous algorithmic tools developed in the last decades for the computation of the Joint Spectral Radius.

In our framework, it is assumed that the switching signal is known by the coder–decoder. However, it was noted in Remark 4.1 that only a few future values of the switching signal actually need to be known by the coder–decoder. Based on this observation, we plan to show that the concept of worst-case topological entropy is also relevant for the *control* of switched linear systems with limited data rate. We think for instance to control schemes involving the switching signal as control input (see, e.g., [36, Chapter 4] [16]).

We also plan to consider variants of this framework, involving for instance relaxations of the assumption that the switching signal is known by the coder–decoder (non-deterministic systems), or considering the control of switched linear systems under data rate constraints. Related questions have been addressed, e.g., in [34] (topological entropy for non-deterministic systems), [9, 13, 34] (control of autonomous systems with limited data rate), and [26] (stabilization of non-deterministic switched linear systems with limited data rate). However, the specific case of switched linear systems seems to have not received much attention yet. Another potential direction for further research is to combine our results for the computation of the worst-case topological entropy with other techniques for the analysis of switched linear systems, in order to fight the curse of dimensionality. We think for instance to p-dominance analysis techniques [5, 6] which allow one to decide whether the system has a low-dimensional dominant behavior.

ACKNOWLEDGMENTS

The authors would like to thank Daniel Liberzon (University of Illinois in Urbana-Champaign) for insightful discussions on the results presented in the paper.

REFERENCES

[1] Roy L Adler, Alan G Konheim, and M Harry McAndrew. 1965. Topological entropy. *Trans. Amer. Math. Soc.* 114, 2 (1965), 309–319. https://doi.org/10.2307/1994177
[2] Ludwig Arnold. 1998. *Random dynamical systems.* Springer-Verlag. https://doi.org/10.1007/978-3-662-12878-7
[3] John Baillieul and Panos J Antsaklis. 2007. Control and communication challenges in networked real-time systems. *Proc. IEEE* 95, 1 (2007), 9–28. https://doi.org/10.1109/JPROC.2006.887290
[4] Luís Barreira. 2017. *Lyapunov exponents.* Springer Science+Business Media. https://doi.org/10.1007/978-3-319-71261-1
[5] Guillaume O Berger, Fulvio Forni, and Raphaël M Jungers. 2018. Path-complete p-dominant switching linear systems. In *2018 IEEE 57th IEEE Conference on Decision and Control (CDC).* IEEE, 6446–6451. https://doi.org/10.1109/CDC.2018.8619703

[6] Guillaume O Berger and Raphaël M Jungers. 2019. A converse Lyapunov theorem for p-dominant switched linear systems. In *2019 18th European Control Conference (ECC)*. IEEE, 1263–1268. https://doi.org/10.23919/ECC.2019.8795923
[7] Rufus Bowen. 1971. Entropy for group endomorphisms and homogeneous spaces. *Trans. Amer. Math. Soc.* 153 (1971), 401–414.
[8] Roger W Brockett and Daniel Liberzon. 2000. Quantized feedback stabilization of linear systems. *IEEE Trans. Automat. Control* 45, 7 (2000), 1279–1289. https://doi.org/10.1109/9.867021
[9] Fritz Colonius. 2012. Minimal bit rates and entropy for exponential stabilization. *SIAM Journal on Control and Optimization* 50, 5 (2012), 2988–3010. https://doi.org/10.1137/110829271
[10] Fritz Colonius and Christoph Kawan. 2009. Invariance entropy for control systems. *SIAM Journal on Control and Optimization* 48, 3 (2009), 1701–1721. https://doi.org/10.1137/080713902
[11] Fritz Colonius, Christoph Kawan, and Girish N Nair. 2013. A note on topological feedback entropy and invariance entropy. *Systems & Control Letters* 62, 5 (2013), 377–381. https://doi.org/10.1016/j.sysconle.2013.01.008
[12] Nicola Elia and Sanjoy K Mitter. 2001. Stabilization of linear systems with limited information. *IEEE Trans. Automat. Control* 46, 9 (2001), 1384–1400. https://doi.org/10.1109/9.948466
[13] Rika Hagihara and Girish N Nair. 2013. Two extensions of topological feedback entropy. *Mathematics of Control, Signals, and Systems* 25, 4 (2013), 473–490. https://doi.org/10.1007/s00498-013-0113-7
[14] João P Hespanha, Payam Naghshtabrizi, and Yonggang Xu. 2007. A survey of recent results in networked control systems. *Proc. IEEE* 95, 1 (2007), 138–162. https://doi.org/10.1109/JPROC.2006.887288
[15] Raphaël M Jungers. 2009. *The joint spectral radius: theory and applications*. Lecture Notes in Control and Information Sciences, Vol. 385. Springer-Verlag Berlin Heidelberg. https://doi.org/10.1007/978-3-540-95980-9
[16] Raphaël M Jungers and Paolo Mason. 2017. On feedback stabilization of linear switched systems via switching signal control. *SIAM Journal on Control and Optimization* 55, 2 (2017), 1179–1198. https://doi.org/10.1137/15M1027802
[17] Christoph Kawan. 2011. Invariance entropy of control sets. *SIAM Journal on Control and Optimization* 49, 2 (2011), 732–751. https://doi.org/10.1137/100783340
[18] Christoph Kawan. 2013. *Invariance entropy for deterministic control systems*. Lecture Notes in Mathematics, Vol. 2089. Springer Science+Business Media. https://doi.org/10.1007/978-3-319-01288-9
[19] Christoph Kawan. 2014. Metric entropy of nonautonomous dynamical systems. *Nonautonomous Dynamical Systems* 1, 1 (2014). https://doi.org/10.2478/msds-2013-0003
[20] Christoph Kawan. 2017. Entropy of nonautonomous dynamical systems. In *Differential and Difference Equations with Applications*. Springer-Verlag, 179–191. https://doi.org/10.1007/978-3-319-75647-9_15
[21] Pascal Koiran. 2001. The topological entropy of iterated piecewise affine maps is uncomputable. *Discrete Mathematics & Theoretical Computer Science* 4, 2 (2001), 351–356.
[22] Sergiĭ Kolyada and Lubomìr Snoha. 1996. Topological entropy of nonautonomous dynamical systems. *Random and Computational Dynamics* 4, 2 (1996), 205–233.
[23] Oleg S Kozlovski. 1998. An integral formula for topological entropy of C^∞ maps. *Ergodic theory and Dynamical Systems* 18, 2 (1998), 405–424. https://doi.org/10.1017/S0143385798100391
[24] Ruggero Lanotte, Massimo Merro, and Fabio Mogavero. 2019. On the decidability of linear bounded periodic cyber-physical systems. In *Proceedings of the 22th International Conference on Hybrid Systems: Computation and Control*. ACM, 87–98. https://doi.org/10.1145/3302504.3311797
[25] Daniel Liberzon. 2003. *Switching in systems and control*. Springer Science+Business Media. https://doi.org/10.1007/978-1-4612-0017-8
[26] Daniel Liberzon. 2014. Finite data-rate feedback stabilization of switched and hybrid linear systems. *Automatica* 50, 2 (2014), 409–420. https://doi.org/10.1016/j.automatica.2013.11.037
[27] Daniel Liberzon and Sayan Mitra. 2017. Entropy and minimal bit rates for state estimation and model detection. *IEEE Trans. Automat. Control* 63, 10 (2017), 3330–3344. https://doi.org/10.1109/TAC.2017.2782478
[28] Hai Lin and Panos J Antsaklis. 2009. Stability and stabilizability of switched linear systems: a survey of recent results. *IEEE Trans. Automat. Control* 54, 2 (2009), 308–322. https://doi.org/10.1109/TAC.2008.2012009
[29] Alexey S Matveev and Alexander Pogromsky. 2016. Observation of nonlinear systems via finite capacity channels: constructive data rate limits. *Automatica* 70 (2016), 217–229. https://doi.org/10.1016/j.automatica.2016.04.005
[30] Alexey S Matveev and Andrey V Savkin. 2009. *Estimation and control over communication networks*. Springer Science+Business Media. https://doi.org/10.1007/978-0-8176-4607-3
[31] Girish N Nair and Robin J Evans. 2004. Stabilizability of stochastic linear systems with finite feedback data rates. *SIAM Journal on Control and Optimization* 43, 2 (2004), 413–436. https://doi.org/10.1137/S0363012902402116
[32] Reza S Raji. 1994. Smart networks for control. *IEEE Spectrum* 31, 6 (1994), 49–55. https://doi.org/10.1109/6.284793
[33] Gian-Carlo Rota and W Gilbert Strang. 1960. A note on the joint spectral radius. *Proceedings of the Netherlands Academy* 22 (1960), 379–381.
[34] Andrey V Savkin. 2006. Analysis and synthesis of networked control systems: topological entropy, observability, robustness and optimal control. *Automatica* 42, 1 (2006), 51–62. https://doi.org/10.1016/j.automatica.2005.08.021
[35] Daniel J Stilwell and Bradley E Bishop. 2000. Platoons of underwater vehicles. *IEEE Control Systems Magazine* 20, 6 (2000), 45–52. https://doi.org/10.1109/37.887448
[36] Zhendong Sun and Shuzhi Sam Ge. 2011. *Stability theory of switched dynamical systems*. Springer-Verlag London. https://doi.org/10.1007/978-0-85729-256-8
[37] Guillaume Vankeerberghen, Julien Hendrickx, and Raphaël M Jungers. 2014. JSR: a toolbox to compute the joint spectral radius. In *Proceedings of the 17th International Conference on Hybrid Systems: Computation and Control*. ACM, 151–156. https://doi.org/10.1145/2562059.2562124
[38] Guosong Yang, João P Hespanha, and Daniel Liberzon. 2019. On topological entropy and stability of switched linear systems. In *Proceedings of the 22th International Conference on Hybrid Systems: Computation and Control*. ACM, 119–127. https://doi.org/10.1145/3302504.3311815
[39] Guosong Yang and Daniel Liberzon. 2017. Feedback stabilization of switched linear systems with unknown disturbances under data-rate constraints. *IEEE Trans. Automat. Control* 63, 7 (2017), 2107–2122. https://doi.org/10.1109/TAC.2017.2767822
[40] Guosong Yang, A James Schmidt, and Daniel Liberzon. 2018. On topological entropy of switched linear systems with diagonal, triangular, and general matrices. In *2018 IEEE 57th IEEE Conference on Decision and Control (CDC)*. IEEE, 5682–5687. https://doi.org/10.1109/CDC.2018.8619087