Worst-Case Topological Entropy and Minimal Data Rate for State Estimation of Switched Linear Systems

By Guillaume O. Berger and Raphaël M. Jungers

Abstract

In this paper, we study the problem of estimating the state of a switched linear system (SLS), when the observation of the system is subject to communication constraints. We introduce the concept of worst-case topological entropy of such systems, and we show that this quantity is equal to the minimal data rate (number of bits per second) required for the state estimation of the system under arbitrary switching. Furthermore, we provide a closed-form expression for the worst-case topological entropy of switched linear systems, showing that its evaluation reduces to the computation of the joint spectral radius (JSR) of some lifted switched linear system obtained from the original one by using tools from multilinear algebra, and thus can benefit from well-established algorithms for the stability analysis of switched linear systems. Finally, drawing on this expression, we describe a practical coder–decoder that estimates the state of the system and operates at a data rate arbitrarily close to the worst-case topological entropy.

1. INTRODUCTION

Modern control systems (such as IoT, cyber-physical systems, etc.) often involve spatially distributed components that communicate through a shared, digital communication network. Due to the digital nature of the network, all data must be quantized before transmission, resulting in quantization error that can affect the performance of the observing/controlling scheme. Furthermore, in applications, the capacity of the network is often limited by cost, power, physical, and/or security constraints. Consequently, a major challenge in the design of such networked systems is to determine the minimal communication data rate required to achieve a given control task. This fundamental question has motivated a lot of research efforts from the control community in recent years, with great theoretical and practical advances; as surveyed in the works of Hespanha et al.[8] and Matveev and Savkin[17].

Inspired by Shannon's work on the link between information entropy and minimal data rate for reliable communication, it was soon realized that the question of data rate requirements for networked systems has strong connections with the notion of topological entropy of these systems; see, for example, Matveev[16] and Pogromsky and Savkin.[19] This quantity, introduced in the late 1960s and now ubiquitous in systems theory, accounts for the growth rate of the smallest number of functions necessary to approximate the trajectories of the system with arbitrary finite accuracy on bounded time intervals (with respect to the length of the interval). It can also be seen as a measure of the rate at which information about the initial condition is generated by the system as time evolves.[19] It is known for instance that for some classes of dynamical systems (such as time-invariant systems with forward invariant initial set), the topological entropy coincides with the minimal data rate for state estimation.[16]

In this paper, we study discrete-time switched linear systems (SLSs). We are interested in 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 (see also Figure 1 for an illustration). SLSs are systems described by a finite set of linear modes among which the system can switch over time. As a paradigmatic class of cyber-physical systems, SLSs have attracted much attention from the control community in recent years. These systems turn out to be extremely challenging in terms of control and analysis, even for basic questions such as stability or stabilizability.[9] In particular, in contrast to LTI systems for which comprehensive eigenvalue-based expressions for the topological entropy and the minimal data rate for state estimation are available, none of these quantities are well understood for SLSs. Several theoretical and practical advances have nevertheless been achieved; for instance, upper bounds and lower bounds on the topological entropy of SLSs, *when the sequence of modes is fixed a priori*, are derived in the work of Yang et al.,[22] and sufficient data rate bounds for feedback stabilization of SLSs, when

Figure 1. State estimation with limited data rate.
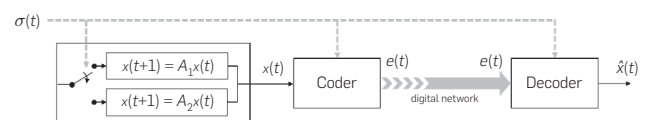
the switching signal is unobserved by the decoder, are established in the work of Liberzon.[14]

In this paper, we study the questions of topological entropy and minimal data rate for state estimation of SLSs in a *worst-case scenario* approach; that is, we aim to determine sharp upper bounds on the topological entropy and find coders–decoders that can estimate the state of the system, for every switching sequence. The worst-case scenario is a popular approach in the study of switched systems, as it provides formal guarantees that the system satisfies the specifications in every situation. Our contribution is twofold. First, we introduce and study the concept of *worst-case topological entropy* of SLSs, defined as the maximal topological entropy that can be reached by the system among all sequences of modes (aka. switching signals). We present a closed-form expression for the worst-case topological entropy of SLSs. More precisely, we show that it can be expressed as the joint spectral radius (JSR, a ubiquitous measure of stability of SLSs) of a "lifted" SLS representing the action of the original system on elements of volume (obtained by leveraging tools from multilinear algebra). The main asset of this expression is that it can be computed numerically via well-established algorithms for the computation of the joint spectral radius. 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 operates 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 the worst-case topological entropy for the problem of state estimation of SLSs under data-rate constraints.

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 estimation 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.

In our analysis, it is assumed that the switching signal is known by the coder–decoder during the state estimation process. This framework is motivated by the fact that the switching signal 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., the headquarter of the company), and the switching signal is not known at the time of the infrastructure's design or might change with time. For instance, a given sequence of modes might be used by the plant of a factory for a certain amount of time. But after a few days or months, another signal 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, whereas the communication infrastructure remains unchanged. 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 conclusion.

Notation. $\mathbb{N}$ is the set of nonnegative integers $\{0, 1, 2, \ldots\}$. For vectors, $\|\cdot\|$ denotes the Euclidean norm, and for matrices, it denotes the spectral norm. $B(\xi, r)$ is the Euclidean closed ball in $\mathbb{R}^d$ centered at $\xi \in \mathbb{R}^d$, with radius $r \geq 0$. $\lceil \cdot \rceil, \lfloor \cdot \rfloor$ are the ceil and floor functions respectively. We focus on 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.

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*[a] of the system and specifies which *mode*, that is, 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 Equation (1) with switching signal σ and initial state $\xi \in \mathbb{R}^d$. Let $K \subseteq \mathbb{R}^d$ be a compact set with nonempty interior, called the *initial set*. We will write (f_σ, K) to denote system (1) with switching signal σ and initial set K; that is, $x_\sigma(\cdot, \xi)$ is a trajectory of (f_σ, K) if and only if $\xi \in K$.

We introduce here the definition of topological entropy (firstly introduced by Bowen,[5] and extended to the case of nonautonomous systems by Kolyada and Snoha[12]). The definition relies on the notion of minimal sets of functions necessary to approximate the trajectories of the system with arbitrary finite accuracy on bounded intervals.

More precisely, let σ be a switching signal for system (1). For $\varepsilon > 0$ and $T \in \mathbb{N}$, we say that $E \subseteq K$ 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 $E \subseteq K$ that is ε-close to $x_\sigma(\cdot, \xi)$ for all $t \in [0, T]$. See Figure 2 for an illustration. We let $s_{span}(\varepsilon, T; f_\sigma, K)$ be the smallest cardinality of an (ε, T)-spanning set for (f_σ, K).
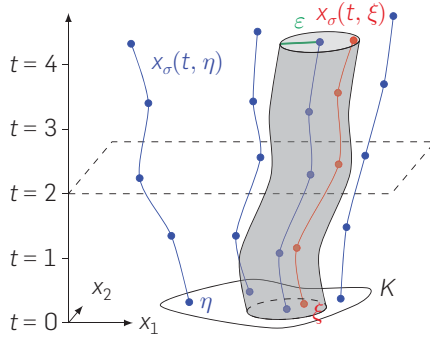
Definition 1. The *topological entropy* of system (1) with switching signal σ and initial set K is defined as

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

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

---

[a] 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 system that will be valid for all 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 $t \in [0, T]$.

2.2. Switched linear systems

In this paper, we are interested in the topological entropy of discrete-time *switched linear systems* (SLSs):

$$x(t+1) = A_{\sigma(t)}x(t), \quad x(t) \in \mathbb{R}^d, \quad t \in \mathbb{N}, \qquad (3)$$

where σ is the switching signal as in Equation (1) and $A_i \in \mathbb{R}^{d \times d}$ for all $i \in \Sigma$. SLSs are thus particular instances of system (1) where each mode is linear. Following the notation of the previous subsection, we let $x_\sigma(t, \xi)$ be the solution, at time t, of system (3) with switching signal σ and initial state $\xi \in \mathbb{R}^d$. In formulas, it is convenient to identify the SLS (3) by its set of modes $\mathcal{A}_\Sigma := \{A_1, ..., A_N\}$. We will also use (A_σ, K) to denote system (3) 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 system (3) from t_1 to t_2 with switching signal σ by

$$\Phi_{\sigma, t_1, t_2} = A_{\sigma(t_2-1)}A_{\sigma(t_2-2)} \cdots A_{\sigma(t_1)}. \qquad (4)$$

The fundamental matrix solution satisfies that for all $\xi \in \mathbb{R}^d$ and $t_1, t_2 \in \mathbb{N}$, $t_1 \leq t_2$, $x_\sigma(t_2, \xi) = \Phi_{\sigma, t_1, t_2}x_\sigma(t_1, \xi)$.

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 1). However, in the case of SLSs, $h(A_\sigma, K)$ does not depend on a particular choice of the initial set $K \subseteq \mathbb{R}^d$, as long as it is compact with nonempty interior; see, for example, Yang et al[23] (Proposition 2). Therefore, we omit the initial set in the notation and denote by $h(A_\sigma)$ the topological entropy of A_σ.

We study the worst-case topological entropy of system (3), that is, the maximal topological entropy that can be reached by the system among all its switching signals.

Definition 2. The *worst-case topological entropy* of system (3) is defined as

$$h_*(\mathcal{A}_\Sigma) = \sup_\sigma h(A_\sigma), \qquad (5)$$

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

The computational aspects of the 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 estimation 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 estimation of the system under arbitrary switching, and that this data rate limit can be approached as close as desired by *practical* coders–decoders.

The example here illustrates the notions of SLS, spanning sets, topological entropy, and worst-case topological entropy.

Example 1. Consider the one-dimensional SLS (3) with $\Sigma = \{1, 2\}$, and $A_1 = 1$, and $A_2 = 2$. Let σ be the switching signal alternating modes 1 and 2: $\sigma = (1, 2, 1, 2, ...)$. The trajectories of the system are thus given by $x_\sigma(t, \xi) = 2^{\lfloor t/2 \rfloor}\xi$. We will show that $h(A_\sigma) = \frac{1}{2}$. As mentioned above, 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 $E = \{0, 1/n, 2/n, ..., 1\}$. We show that E is (ε, T)-spanning for (A_σ, K). To do this, let $\xi \in K$, and let $\eta \in E$ be such that $|\xi - \eta| = \min_{\eta' \in E}|\xi - \eta'|$. Then, by definition of E, it holds that $|\xi - \eta| \leq 1/(2n) \leq \varepsilon 2^{-T/2}$. This implies that for every $t \in [0, T]$, $|x_\sigma(t, \xi) - x_\sigma(t, \eta)| = 2^{\lfloor t/2 \rfloor}|\xi - \eta| \leq \varepsilon$. Hence, E is (ε, T)-spanning for (A_σ, K) and thus

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

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

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

Now, we show that $\frac{1}{2}$ is a lower bound on $h(A_\sigma)$. Let $m = \lceil \varepsilon^{-1}2^{T/2-2} \rceil - 1$ (without loss of generality, we may assume $m > 0$, because we take the limit when $\varepsilon \to 0$ and $T \to \yen$), and define $F = \{0, 1/m, 2/m, ..., 1\}$. Then, for any distinct $\xi, \eta \in F$, it holds that $|\xi - \eta| > \varepsilon 2^{2-T/2}$, so that $|x_\sigma(T, \xi) - x_\sigma(T, \eta)| = 2^{\lfloor T/2 \rfloor}|\xi - \eta| > 2\varepsilon$. This implies that

$$s_{\text{span}}(\varepsilon, T; A_\sigma, K) \geq |F| = m+1 \geq \varepsilon^{-1}2^{T/2-2} - 1$$

(see, for example, Liberzon and Mitra[15], Lemma 3, for details). Hence, injecting in Equation (2), it follows that $h(A_\sigma) \geq \log_2 2^{1/2} = \frac{1}{2}$.

As for the worst-case topological entropy, it is quite intuitive that a switching signal σ that maximizes the topological entropy is given for instance by using only mode 2: $\sigma = (2, 2, 2, ...)$. In this case, $x_\sigma(t, \xi) = 2^t\xi$, and we deduce that $h(A_\sigma) = \log_2 2 = 1$. Hence, $h_*(\mathcal{A}_\Sigma) = 1$.

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, the exterior power of a matrix). 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 $\yen$. More precisely, for a finite set of matrices $\mathcal{A}_\Sigma := \{A_1, ..., A_N\} \subseteq \mathbb{R}^{d \times d}$, the *joint spectral radius* of $\mathcal{A}_\Sigma$ is defined as

$$\rho(\mathcal{A}_\Sigma) = \limsup_{t \to \infty} \max_{A_{i_1}, ..., A_{i_t} \in \mathcal{A}_\Sigma} \left\| A_{i_1} A_{i_2} \cdots A_{i_t} \right\|^{1/t}. \tag{6}$$

This quantity was introduced by Rota and Strang in 1960[18] in order to characterize the stability of SLSs. In particular, the JSR has the following property (see, e.g., Jungers[9], 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 $\mathcal{A}_\Sigma$ converges to zero as $t \to \yen$, if and only if the JSR of $\mathcal{A}_\Sigma$ satisfies $\rho(\mathcal{A}_\Sigma) < 1$.

3.2. Exterior power of matrices

Exterior algebras are algebraic constructions used to study the notions of areas, volumes, and their higher-dimensional analogues, in general vector spaces. In particular, the notion of exterior power of linear operators is used to represent the action of these operators on such elements of areas, volumes, etc. The exterior power can be defined in a coordinate-free fashion; see, for example, Arnold[1] (Section 3.2.2). However, in this paper, due to space limitation, we will restrict our attention to the exterior powers of *matrices*, which are themselves matrices and thus allow for a coordinate-based definition.

To do this, let $\mathcal{I}$ be the set of all multi-indexes of the form $(i_1, ..., i_k) \in \{1, ..., d\}^k$, with $k \in \{0, ..., d\}$ and $i_1 < i_2 < ... < i_k$. Hence, the size of $\mathcal{I}$ is 2^d.

Let $A \in \mathbb{R}^{d \times d}$. The *exterior power* of A, denoted by $A^\wedge$, is the $2^d \times 2^d$ matrix whose entries are indexed by the elements of $\mathcal{I}$, and is defined for any $I, J \in \mathcal{I}$ by

$$(A^\wedge)_{I,J} = \begin{cases} 0 & \text{if } |I| \neq |J| \\ 1 & \text{if } |I| = |J| = 0 \\ \det([A_{ij}]_{i \in I, j \in J}) & \text{otherwise.} \end{cases}$$

(See also Section 5 for a numerical example.)

The following proposition, whose proof can be found in Arnold[1] (Section 3.2.3), summarizes the properties of the exterior power of matrices that we will need in this work.

PROPOSITION 1. *Let $A, B \in \mathbb{R}^{d \times d}$.*
1. *$I^\wedge = I$, $(AB)^\wedge = A^\wedge B^\wedge$, $(A^\top)^\wedge = (A^\wedge)^\top$.*
2. *If A is upper-triangular (resp. lower-triangular, or orthogonal), then so is $A^\wedge$ (if the elements of $\mathcal{I}$ are ordered with the canonical "lexicographical" ordering).*
3. *$\|A^\wedge\| = \prod_{i=1}^{d} \max\{\overline{\sigma}_i(A), 1\}$ where $\overline{\sigma}_1(A), ..., \overline{\sigma}_d(A)$ are the singular values of A.*
4. *The eigenvalues of $A^\wedge$ are given by $\prod_{i \in I} \lambda_i(A), I \in \mathcal{I}$, where $\lambda_1(A), ..., \lambda_d(A)$ are the eigenvalues 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 1. *The worst-case topological entropy of* system (3) *satisfies*

$$h_*(\mathcal{A}_\Sigma) = \log_2 \rho(\mathcal{A}_\Sigma^\wedge) \tag{7}$$

where $\mathcal{A}_\Sigma^\wedge = \{A_1^\wedge, ..., A_N^\wedge\}$.

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, for example, Jungers[9] (Section 2.3). Although 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 side term of Equation (7). The computation of the exterior power of a matrix is straightforward from its definition, and thus $\mathcal{A}_\Sigma^\wedge$ can be computed in a systematic way. However, it should be noted that the dimension of $\mathcal{A}_\Sigma^\wedge$ increases exponentially with the dimension of the system, and thus so will the complexity of approximating the JSR of $\mathcal{A}_\Sigma^\wedge$ (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 $\rho(\mathcal{A}_\Sigma^\wedge)$—although not sufficient to fight the curse of dimensionality—is to observe that as the matrices $\mathcal{A}_i^\wedge$ are block diagonal (each diagonal block corresponding to a given size of the multi-indexes $I \in \mathcal{J}$), the computation of the JSR of $\mathcal{A}_\Sigma^\wedge$ can be decoupled among the different diagonal blocks (see Jungers[9], Proposition 1.5).

Furthermore, there are cases for which the computation of the JSR is straightforward; for instance, if $\mathcal{A}_\sigma$ is a set of normal (or triangular) matrices. Combining these observations with the properties of the exterior power of matrices (Proposition 1), this gives efficient ways to compute the worst-case topological entropy of such sets of matrices.

COROLLARY 1. *Let $\mathcal{A}_\Sigma := \{A_1, ..., A_N\} \subseteq \mathbb{R}^{d \times d}$ be a set of normal matrices. For each $i \in \Sigma$, let $\lambda_1(A_i), ..., \lambda_d(A_i)$ be the eigenvalues of A_i. Then*

$$h_*(\mathcal{A}_\Sigma) = \log_2 \max_{i \in \Sigma} \prod_{k=1}^{d} \max\{|\lambda_k(A_i)|, 1\}.$$

The same holds for sets of upper-triangular (resp. lower-triangular) matrices. Moreover, in this case, the eigenvalues are on the diagonal of the matrices.

Numerical illustrative examples of the computation of the worst-case topological entropy of SLSs, using Theorem 1 and Corollary 1, are presented in Subsection 5.1.

3.4. Related works

The worst-case topological entropy provides an upper bound on the topological entropy of $\mathcal{A}_\sigma$ for any switching signal σ. The question of estimating $h(\mathcal{A}_\sigma)$ has been addressed, for example, in the works of Yang et al.[22, 23]. Because the focus is put on a particular matrix sequence A_σ (disregarding other sequences), the bounds on $h(\mathcal{A}_\sigma)$ obtained in the works of

Yang et al.[22, 23] are in general better than the worst-case topological entropy. However, in some "ill-conditioned" cases (e.g., triangular systems with large differences among the diagonal entries), the bounds available in the works of Yang et al.[22, 23] can be quite conservative. In these cases, it can be beneficial to use $h_\star(\mathcal{A}_\Sigma)$ 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.[9, 18, 21] 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[1, 2] and entropy-related properties[11, 13] of dynamical systems and control systems. For instance, we note the remarkable formula by Kozlovski[13] for the topological entropy of a discrete-time autonomous dynamical system, described by a $C^\yen$ map $f : X \to X$, on a compact Riemannian manifold:

$$h(f) = \lim_{n \to \infty} \frac{1}{n} \log_2 \int_X \left\| (Df_x^n)^\wedge \right\| dx.$$

Theorem 1 shows that the integral can be replaced by a maximum over all switching signals in the case of the worst-case topological entropy of SLSs, and enables practical computation using the stability theory of SLSs.

4. PRACTICAL CODER–DECODER
We investigate the problem of estimating the state of system (3) via a limited data-rate communication network. The observation process is as follows (see also Figure 1 for an illustration). At specific transmission times, $0 = T_0 < T_1 < T_2 < \ldots$ ($T_j \in \mathbb{N}$), a *coder* measures the state $x(T_j)$ of the system, and sends a symbol $e(T_j)$ to a *decoder* via a noiseless digital channel which can carry one discrete-valued symbol per time epoch $[T_j, T_{j+1}]$, selected from a coding alphabet $\mathcal{E}_j$. Neglecting 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 $e(T_0), \ldots, e(T_j)$ 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 $\mathcal{C}[\cdot, \cdot, \cdot \mid j, \sigma]$ (parameterized by j the index of the epoch and σ the switching signal of the system):

$$\mathcal{C}[\cdot, \cdot, \cdot \mid j, \sigma] : \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}_{>0} \to \mathcal{E}_j,$$
$$e(T_j) = \mathcal{C}[x(T_j), \hat{x}(T_j), \delta_j \mid j, \sigma], \quad j \in \mathbb{N}, \tag{8}$$

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 $e(T_j) \in \mathcal{E}_j \subseteq \mathbb{R}^d$, 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):

$$\mathcal{C}[\cdot, \cdot, \cdot \mid j, \sigma] : \mathcal{E}_{j-1} \times \mathbb{R}^d \times \mathbb{R}_{>0} \to (\mathbb{R}^d)^{T_{j+1} - T_j},$$
$$[\hat{x}(T_j), \ldots, \hat{x}(T_{j+1} - 1)] =$$
$$\mathcal{D}[e(T_{j-1}), \hat{x}(T_{j-1}), \delta_{j-1} \mid j, \sigma], \quad j \in \mathbb{N}, \tag{9}$$

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 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

$$R = \sup_\sigma \sup_{j \in \mathbb{N}} \frac{\log_2 |\mathcal{E}_j|}{T_{j+1} - T_j}. \tag{10}$$

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

Definition 3. The coder–decoder (Equations (8) and (9)) is said to *observe* the SLS (3) with initial set K if there exists $C > 0$ and $g \in (0, 1)$ such that for every switching signal σ and initial state $\xi \in K$, it holds that

$$\| x_\sigma(t, \xi) - \hat{x}(t) \| \le C g^t, \quad \forall t \in \mathbb{N}. \tag{11}$$

Remark 1. Equations (8) and (9) assume that the whole switching signal is known by the coder and the decoder during its operation (see also Section 1 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 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 system (3) 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 $K \subseteq \mathbb{R}^d$ and $R' > h_\star(\mathcal{A}_\Sigma)$, there is such a coder–decoder that observes system (3) with initial set K and whose data rate satisfies $R \le R'$ (see also Theorem 2 at the end of this section). This result relies on the properties of the joint spectral radius and the exterior power of matrices to build coders–decoders with data rate as close as desired to the right-hand side term of Equation (7).

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 $\mathbb{Z}_{\text{even}}$ (resp. $\mathbb{Z}_{\text{odd}}$) is the set of even (resp. 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 $r_1, \ldots, r_d \ge 0$, we define

$$\text{Grid}(r_1, \ldots, r_d) = I(r_1) \times \cdots \times I(r_d).$$

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

$$\| \xi - Q(\xi) \| \le d^{1/2}, \quad \forall \xi \in [-r_1, r_2] \times \cdots \times [-r_d, r_d].$$

Let $K \subseteq \mathbb{R}^d$ be a compact set and fix a target data rate R' strictly larger than the right-hand side term of Equation (7). We will build a coder–decoder that observes the SLS (3) with initial set K and whose data rate satisfies $R \leq R'$. (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 $\delta_0 \geq 0$ be such that $K \subseteq B(\hat{x}(0), \delta_0)$. Fix $\alpha > 1$. For each $j \in \mathbb{N}$, let $\delta_j = \delta_0 \alpha^j$ (this will imply that the bound on the estimation error $\left\| x_\sigma(T_j, \xi) - \hat{x}(T_j) \right\|$ decreases by a factor $1/\alpha$ between two transmission times T_j and T_{j+1}, and thus the rate g of decay of the estimation error (see Definition 3) is given by $\alpha^{-1/\tau}$ where τ is an upper bound on $T_{j+1} - T_j$).

- At time T_j, the values of T_{j+1} and $\mathcal{E}_j$ 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 $\bar{\sigma}_i(\Phi_{\sigma, T_j, T})$, $i \in \{1, \ldots, d\}$, be the singular values of the fundamental matrix solution $\Phi_{\sigma, T_j, T}$. We define T_{j+1} as the smallest $T \in \mathbb{N}, T > T_j$, satisfying

$$\prod_{i=1}^{d} \left\lceil \alpha d^{1/2} \bar{\sigma}_i(\Phi_{\sigma, T_j, T}) \right\rceil^* \leq 2^{(T-T_j)R'}. \tag{12}$$

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

$$\mathcal{E}_j = \text{Grid}(r_1, \ldots, r_d), \quad r_i = \alpha d^{1/2} \bar{\sigma}_i(\Phi_{\sigma, T_j, T_{j+1}}). \tag{13}$$

- 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 Equation (13). The symbol sent by the coder at time T_j is then defined as

$$e(T_j) = Q(d^{1/2} SV^* \Delta \xi_j / \delta_{j+1}) \in \mathcal{E}_j. \tag{14}$$

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

- 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:

$$\hat{x}(T_j) = \Phi_{\sigma, T_{j-1}, T_j} \hat{x}(T_{j-1}) + \delta_j U e(T_{j-1}) / d^{1/2}, \tag{15}$$

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

$$\hat{x}(t) = A_{\sigma(t-1)} \hat{x}(t-1), \quad t \in (T_j, T_{j+1}), \tag{16}$$

The implementation of the above coder–decoder is described in Figure 4.

Summarizing, we have proved the following result on the equivalence of the worst-case topological entropy and the minimal data rate for state estimation of the system under arbitrary switching signals.

THEOREM 2. *Let $K \subseteq \mathbb{R}^d$ be compact, and consider the SLS (3). If $R' < h_*(\mathcal{A}_\Sigma)$, then there is no coder–decoder with data rate $R \leq R'$ that observes system (3). If $R' > h_*(\mathcal{A}_\Sigma)$, then there is a practical coder–decoder with data rate $R \leq R'$ that observes system (3).*

Remark 2. It is worth noticing the following effects of the parameters α and R on the output of the coder–decoder. By definition of $\delta_j = \delta_0 / \alpha^j$, α gives the rate of decrease of the

Figure 3. The different quantities appearing in the definition of the coder–decoder. The gray dots (bottom left) represent 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 Equation (15), Grid(r_1, r_2) is scaled and rotated by $\delta_{j+1} d^{-1/2} U$, and centered at $\bar{x}(T_{j+1}) := \Phi_{\sigma, T_j, T_{j+1}} \hat{x}(T_j)$. 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.
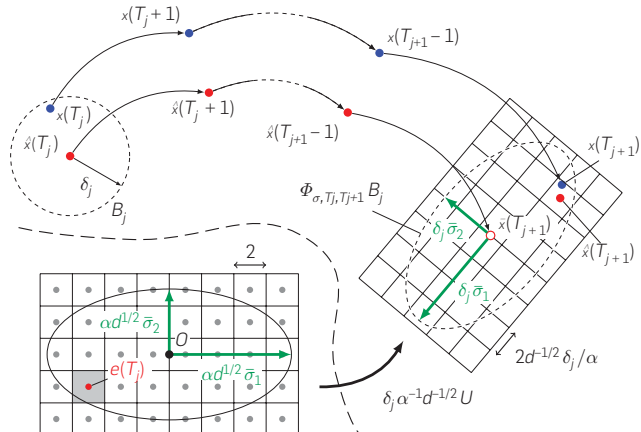

Figure 4. Coder and decoder implementations.

Coder
Input: R, α, $\hat{x}(0)$, δ_0, and σ.
Let $j = 0$ and $T_j = 0$
Loop: wait until $t_{\text{real}} = T_j$ (t_{real} is the real time)
   Measure the current state $x(T_j)$ of the system
   Compute T_{j+1} and $\mathcal{E}_j$ as in (12)–(13)
   Compute $e(T_j)$ as in (14) and send $e(T_j)$ to the decoder
   Let $j = j + 1$
   Compute $\hat{x}(T_j)$ as in (15)
end loop

Decoder
Input: R, α, $\hat{x}(0)$, δ_0, and σ.
Let $j = 0$ and $T_j = 0$
Loop: wait until $t_{\text{real}} = T_j$ (t_{real} is the real time)
   Receive symbol $e(T_{j-1})$ (If $j = 0$, $e(T_{-1}) = 0$)
   Compute T_{j+1} and $\mathcal{E}_j$ as in (12)–(13)
   Compute $\hat{x}(T_j)$ as in (15)   (If $j = 0$, $\hat{x}(T_0) = \hat{x}(0)$)
   Compute $\hat{x}(t)$ for $t \in (T_j, T_{j+1})$ as in (16)
   Let $j = j + 1$
end loop

worst-case estimation error at the transmission times T_j. On the other hand, by Equation (12), 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.

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 2. Consider the two-dimensional SLS (3) with $\Sigma = \{1, 2\}$, and

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

We use Theorem 1 to compute the worst-case topological entropy of this system. Therefore, we compute the exterior power 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[21] (in MATLAB) to compute the JSR of $\mathcal{A}_\Sigma^\wedge$: This gives $\rho(\mathcal{A}_\Sigma^\wedge) = 1.2379$. Hence, we conclude that $h_*(\mathcal{A}_\Sigma) = 0.3079$.

Example 3. Consider the 2-dimensional SLS (3) with $\Sigma = \{1, 2\}$, and

$$A_1 = \begin{bmatrix} 3 & 1 \\ & 0.1 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 1.1 & 1 \\ & 2 \end{bmatrix}.$$

Because A_1 and A_2 are upper-triangular matrices, we may apply Corollary 1. We deduce that the worst-case topological entropy of the SLS associated with $\mathcal{A}_\Sigma$ is equal to $\log_2 3 = 1.5850$. The reader will check that the same result can be obtained by applying directly Theorem 1; indeed the exterior power 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 (see Jungers[9], Proposition 2.3).

5.2. State estimation with limited data rate
We apply the coder–decoder described in Section 4 for the state estimation of the SLS in Example 2. The parameters of the coder–decoder (see Figure 4) are set as follows: We fix the value $\alpha = 2.5$; the values of $\hat{x}(0)$ and δ_0 are given in Figure 5-top; and we

Figure 5. 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 transmission times T_j (in general, the error decreases at these times because the decoder receives a new symbol).
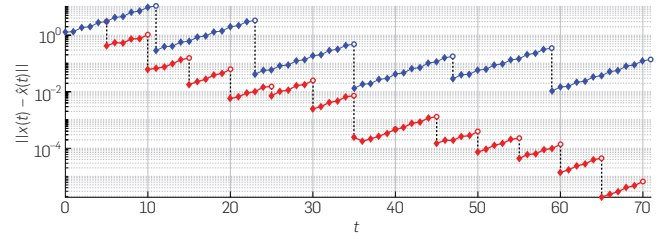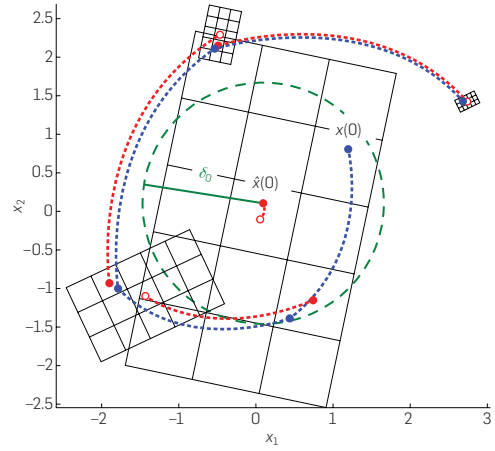
use different values for the maximal data rate R, as explained here.

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 (see Example 2). This ensures that each epoch lasts one unit of time, that is, $T_{j+1} - T_j = 1$ (see also Remark 2). A sample execution of the coder–decoder is presented in Figure 5—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.

Secondly, 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 is longer (between 5 and 12, in our simulation). The evolution of the estimation error $\|x(t) - \hat{x}(t)\|$ is represented in Figure 5-bottom. As expected, we observe that the estimation error decreases more rapidly when the data rate is higher.

6. 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 estimation 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, which 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 expression describing the worst-case topological entropy as the joint spectral radius of some set of matrices obtained from the original system by taking the exterior power of the individual modes. 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 of switched linear systems.

In our framework, it is assumed that the switching signal is known by the coder–decoder. However, it was noted in Remark 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., Jungers and Mason[10] and Sun and Ge[20]—Chapter 4).

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 (nondeterministic systems), and considering the stabilization of switched linear systems under data-rate constraints. Related questions have been addressed, for example, in the works of Savkin[19] (topological entropy for nondeterministic systems); Colonius[6], Hagihara and Nair[7], and Savkin[19] (control of autonomous systems with limited data rate), and Liberzon[14] (stabilization of nondeterministic switched linear systems with limited data rate). 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[3, 4] that allow one to decide whether the system has a low-dimensional dominant behavior.

Acknowledgments
⦿

References
1. Arnold, L. Random dynamical systems. In *Springer Monographs in Mathematics*. Springer-Verlag, Berlin, 1998.
2. Barreira, L. *Lyapunov exponents*. Springer Science+Business Media, Basel, 2017.
3. Berger, G.O., Forni, F., Jungers, R.M. Path-complete *p*-dominant switching linear systems. In *2018 IEEE 57th IEEE Conference on Decision and Control (CDC)* (2018), IEEE, Miami, 6446–6451.
4. Berger, G.O., Jungers, R.M. A converse Lyapunov theorem for *p*-dominant switched linear systems. In *2019 18th European Control Conference (ECC)* (2019), IEEE, Naples, 1263–1268.
5. Bowen, R. Entropy for group endomorphisms and homogeneous spaces. *Trans. Am. Math. Soc. 153* (1971), 401–414.
6. Colonius, F. Minimal bit rates and entropy for exponential stabilization. *SIAM J. Control Optim. 50*, 5 (2012), 2988–3010.
7. Hagihara, R., Nair, G.N. Two extensions of topological feedback entropy. *Math. Control Signals Syst. 25*, 4 (2013), 473–490.
8. Hespanha, J.P., Naghshtabrizi, P., Xu, Y. A survey of recent results in networked control systems. *Proc. IEEE 95*, 1 (2007), 138–162.
9. Jungers, R.M. The joint spectral radius: Theory and applications. In *Lecture Notes in Control and Information Sciences*. Volume 385. Springer-Verlag, Berlin Heidelberg, 2009.
10. Jungers, R.M., Mason, P. On feedback stabilization of linear switched systems via switching signal control. *SIAM J. Control Optim. 55*, 2 (2017), 1179–1198.
11. Kawan, C. Invariance entropy for deterministic control systems. In *Lecture Notes in Mathematics*. Volume 2089. Springer Science+Business Media, Cham, 2013.
12. Kolyada, S., Snoha, L. Topological entropy of nonautonomous dynamical systems. *Random Comput. Dyn. 4*, 2 (1996), 205–233.
13. Kozlovski, O.S. An integral formula for topological entropy of C^∞ maps. *Ergod. Theory Dyn. Syst. 18*, 2 (1998), 405–424.
14. Liberzon, D. Finite data-rate feedback stabilization of switched and hybrid linear systems. *Automatica 50*, 2 (2014), 409–420.
15. Liberzon, D., Mitra, S. Entropy and minimal bit rates for state estimation and model detection. *IEEE Trans. Autom. Control 63*, 10 (2018), 3330–3344.
16. Matveev, A.S., Pogromsky, A. Observation of nonlinear systems via finite capacity channels: Constructive data rate limits. *Automatica 70*, (2016), 217–229.
17. Matveev, A.S., Savkin, A.V. Estimation and control over communication networks. In *Control Engineering*. Springer Science+Business Media, Basel, 2009.
18. Rota, G.-C., Strang, W.G. A note on the joint spectral radius. *Proc. Netherlands Acad. 22*, (1960), 379–381.
19. Savkin, A.V. Analysis and synthesis of networked control systems: Topological entropy, observability, robustness and optimal control. *Automatica 42*, 1 (2006), 51–62.
20. Sun, Z., Ge, S.S. Stability theory of switched dynamical systems. In *Communications and Control Engineering*. Springer-Verlag, London, 2011.
21. Vankeerberghen, G., Hendrickx, J., Jungers, R.M. JSR: A toolbox to compute the joint spectral radius. In *Proceedings of the 17th International Conference on Hybrid Systems: Computation and Control* (2014), ACM, Berlin, 151–156.
22. Yang, G., Hespanha, J.P., Liberzon, D. On topological entropy and stability of switched linear systems. In *Proceedings of the 22nd International Conference on Hybrid Systems: Computation and Control* (2019), ACM, Montreal, 119–127.
23. Yang, G., Schmidt, A.J., Liberzon, D. On topological entropy of switched linear systems with diagonal, triangular, and general matrices. In *2018 IEEE 57th IEEE Conference on Decision and Control (CDC)* (2018), IEEE, Miami, 5682–5687.

Guillaume O. Berger and Raphaël M. Jungers ([guillaume.berger, raphael.jungers]@uclouvain.be), ICTEAM Institute, UCLouvain, Belgium.