On topological entropy of switched linear systems with diagonal, triangular, and general matrices

Guosong Yang, A. James Schmidt, and Daniel Liberzon

Abstract— This paper introduces a notion of topological entropy for switched systems, formulated using the minimal number of initial states needed to approximate all initial states within a finite precision. We show that it can be equivalently defined using the maximal number of initial states separable within a finite precision, and introduce switching-related quantities such as the active time of each mode, which prove to be useful in calculating the topological entropy of switched linear systems. For general switched linear systems, we show that the topological entropy is independent of the set of initial states, and establish upper and lower bounds using the active-time-weighted averages of the norms and traces of system matrices in individual modes, respectively. For switched linear systems with scalar-valued state or simultaneously diagonalizable matrices, we derive formulae for the topological entropy using active-time-weighted averages of eigenvalues, which can be extended to the case with simultaneously triangularizable matrices to obtain an upper bound. In these three cases with special matrix structure, we also provide more general but more conservative upper bounds for the topological entropy.

I. INTRODUCTION

In systems theory, topological entropy describes the information accumulation needed to approximate trajectories within a finite precision, or the complexity growth of a system acting on sets with finite measure. The latter idea corresponds to Kolmogorov's original definition in [1], and shares a striking resemblance to Shannon's information entropy [2]. Adler first defined topological entropy as an extension of Kolmogorov's metric entropy, quantifying a map's expansion by the minimal cardinality of subcover refinements [3]. An alternative definition using the maximal number of trajectories separable within a finite precision was introduced by Bowen [4] and independently Dinaburg [5]. Equivalence between the two definitions above was established in [6]. Most results on topological entropy are for time-invariant systems, as time-varying dynamics introduce complexities which require new methods to understand [7], [8]. This work on the topological entropy of switched linear systems provides an initial study on some of these complexities.

Entropy has played a prominent role in control theory, in which information flow appears between sensors and actuators for maintaining or inducing desired properties.

Nair et al. first introduced topological feedback entropy for discrete-time systems [9], following the construction in [3]. Their definition extended the classical entropy notions, notably in allowing for non-compact state spaces, but still described the uncertainty growth as time evolves. Colonius and Kawan later proposed a notion of invariance entropy for continuous-time systems [10], which is closer in spirit to the trajectory-counting formulation in [4], [5]. In [11], the two notions above were summarized and an equivalence was established between them. The results of [10] were extended from set invariance to exponential stabilization in [12].

This paper studies the topological entropy of switched linear systems. Switched systems have become a popular topic in recent years (see, e.g., [13] and references therein). It is well-known that, in general, a switched system does not inherit stability properties of the individual modes. In [14], it was shown that a switched linear system generated by a finite family of pairwise-commuting Hurwitz matrices is globally uniformly exponentially stable. This result has been generalized to the cases where the Lie algebra generated by the system matrices is nilpotent [15], solvable [16], [17], or has a compact semisimple part [18], [19]. In particular, a nilpotent or solvable Lie algebra implies that the system matrices are simultaneously triangularizable, which motivates us to study the topological entropy of such switched linear systems.

Our interest in studying entropy of switched systems is strongly motivated by its relation to the data-rate requirements in control systems. For a linear time-invariant control system, it has been shown that the minimal data rate for stabilization equals the topological entropy in open-loop [20], [21]. For switched systems, however, neither the minimal data rate nor the topological entropy is well-understood. Sufficient data rates for feedback stabilization of switched linear systems were established in [22], [23]. The paper [24] extended the estimation entropy from [25] to switched systems and formulated similar data-rate conditions. In this work we aim at contributing to these efforts.

The main contribution of this paper is the construction of formulae and bounds for the topological entropy of switched linear systems. Section II introduces the entropy definition and switching-related quantities such as the active time of each mode, which prove to be useful in calculating the topological entropy. Section III proves that the topological entropy of a switched linear system is independent of the set of initial states, and establishes upper and lower bounds for the topological entropy of a general switched linear system using active-time-weighted averages of the norms and traces of systems matrices in individual modes, respectively.

Sections IV–VI provide formulae and tighter bounds for the topological entropy of switched linear systems with special matrix structure. Section IV considers the case with a scalar-valued state, and derives a formula for the entropy—which equals the previous lower bound for the general case—and also general upper bounds. Section V examines the case with simultaneously diagonalizable matrices by deriving a formula for the entropy using the component-wise active-time-weighted averages of eigenvalues, relating it to the entropy in each scalar component or in each mode, and establishing general upper bounds. Section VI studies the more general case with simultaneously triangularizable matrices, and derives an upper bound for the entropy using the active-time-weighted averages of eigenvalues in the first scalar component and eigenvalue differences between adjacent scalar components, as well as more general upper bounds. In the proof, we derive a formula for solutions of switched triangular systems, which is of independent interest. Section VII remarks on a future research topic.

Notations: By default, all logarithms are natural logarithms. Let $\mathbb{R}_+ := [0, \infty)$ and $\mathbb{N} := \{0, 1, \ldots\}$. For a scalar $a \in \mathbb{C}$, denote by $\mathrm{Re}(a)$ and $\mathrm{Im}(a)$ its real and imaginary parts, respectively. For a vector $v \in \mathbb{C}^n$, denote by v_i its i-th scalar component. For a matrix $A \in \mathbb{C}^{n \times n}$, denote by $\mathrm{spec}(A)$, $\mathrm{tr}(A)$, and $\det(A)$ its spectrum, trace, and determinant, respectively. For a set $E \subset \mathbb{C}^n$, denote by $|E|$ and $\mathrm{vol}(E)$ its cardinality and volume (Lebesgue measure), respectively. Denote by $\|v\|_\infty := \max_i |v_i|$ the ∞-norm of a vector v, and by $\|A\|_\infty := \max_i \sum_j |a_{ij}|$ the (induced) ∞-norm of a matrix $A = [a_{ij}]$.

II. PRELIMINARIES

A. Entropy definitions

Consider a family of continuous-time dynamical systems

$$\dot{x} = f_p(x), \qquad p \in \mathcal{P} \tag{1}$$

with the state $x \in \mathbb{R}^n$, in which each system is labeled by an index p from a finite *index set* $\mathcal{P}$, and all the functions $f_p : \mathbb{R}^n \to \mathbb{R}^n$ are locally Lipschitz. We are interested in the corresponding switched system defined by

$$\dot{x} = f_\sigma(x), \qquad x(0) \in K, \tag{2}$$

where $\sigma : \mathbb{R}_+ \to \mathcal{P}$ is a right-continuous, piecewise constant *switching signal*, and $K \subset \mathbb{R}^n$ is a compact set of initial states with a nonempty interior. We call the system with index p in (1) the p-th *mode* of (2), and $\sigma(t)$ the active mode at time t. Denote by $\xi_\sigma(x, t)$ the solution of (2) at time t with switching signal σ and initial state x. For fixed σ and x, the trajectory $\xi_\sigma(x, \cdot)$ is absolutely continuous and satisfies the differential equation (2) away from discontinuities of σ, which are called switching times, or simply *switches*. We assume there is at most one switch at a time, and finitely many switches on a finite time interval (i.e., the set of switches has no accumulation point). Denote by $N_\sigma(t, \tau)$ the number of switches on an interval $(\tau, t]$.

Let $\|\cdot\|$ be some chosen norm on $\mathbb{R}^n$ and the corresponding induced norm on $\mathbb{R}^{n \times n}$, and σ a switching signal. Given a time horizon $T \geq 0$ and a radius $\varepsilon > 0$, we define the following open ball in K with center x:

$$B_{f_\sigma}(x, \varepsilon, T) := \Big\{ x' \in K : \max_{t \in [0,T]} \|\xi_\sigma(x', t) - \xi_\sigma(x, t)\| < \varepsilon \Big\}. \tag{3}$$

We say a finite set of points $E \subset K$ is (T, ε)-*spanning* if

$$K = \bigcup\nolimits_{\hat{x} \in E} B_{f_\sigma}(\hat{x}, \varepsilon, T), \tag{4}$$

or equivalently, for each $x \in K$, there is a point $\hat{x} \in E$ such that $\|\xi_\sigma(x, t) - \xi_\sigma(\hat{x}, t)\| < \varepsilon$ for all $t \in [0, T]$. Denote by $S(f_\sigma, \varepsilon, T, K)$ the minimal cardinality of a (T, ε)-spanning set. The *topological entropy* of the switched system (2) with initial set K and switching signal σ is defined as

$$h(f_\sigma, K) := \lim_{\varepsilon \searrow 0} \limsup_{T \to \infty} \frac{1}{T} \log S(f_\sigma, \varepsilon, T, K). \tag{5}$$

Note that $h(f_\sigma, K)$ is nonnegative. For brevity, we will refer to $h(f_\sigma, K)$ simply as the (topological) entropy of the switched system (2) in the remainder of the paper.

Remark 1. In light of [26, Prop. 3.1.2], the value of $h(f_\sigma, K)$ is the same for all metrics defining the same topology. Hence $\|\cdot\|$ can be arbitrary. For concreteness, we take $\|\cdot\|$ to be the ∞-norm of a vector or the (induced) ∞-norm of a matrix.

Next, we introduce an equivalent definition for the entropy of the switched system (2). With T and ε given as before, we say a finite set of points $E \subset K$ is (T, ε)-*separated* if $\hat{x}' \notin B_{f_\sigma}(\hat{x}, \varepsilon, T)$ for all $\hat{x}, \hat{x}' \in E$, or equivalently, for all distinct points $\hat{x}, \hat{x}' \in E$, there is a time $t \in [0, T]$ such that $\|\xi_\sigma(\hat{x}', t) - \xi_\sigma(\hat{x}, t)\| \geq \varepsilon$. Denote by $N(f_\sigma, \varepsilon, T, K)$ the maximal cardinality of a (T, ε)-separated set. The entropy $h(f_\sigma, K)$ can be equivalently formulated as follows:

Proposition 1. *The topological entropy of* (2) *satisfies*

$$h(f_\sigma, K) = \lim_{\varepsilon \searrow 0} \limsup_{T \to \infty} \frac{1}{T} \log N(f_\sigma, \varepsilon, T, K). \tag{6}$$

B. Active time, active rates, and weighted averages

For each mode, we define the *active time* over $[0, t]$ by

$$\tau_p(t) := \int_0^t \mathbb{1}_p(\sigma(s)) \, \mathrm{d}s, \qquad p \in \mathcal{P} \tag{7}$$

with the indicator function $\mathbb{1}_p(\sigma(s)) = 1$ if $\sigma(s) = p$ and $\mathbb{1}_p(\sigma(s)) = 0$ if $\sigma(s) \neq p$, the *active rate* over $[0, t]$ by

$$\rho_p(t) := \tau_p(t)/t, \qquad p \in \mathcal{P} \tag{8}$$

with $\rho_p(0) := \mathbb{1}_p(\sigma(0))$, and the *asymptotic active rate* by

$$\hat{\rho}_p := \limsup_{t \to \infty} \rho_p(t), \qquad p \in \mathcal{P}. \tag{9}$$

Clearly, τ_p are nondecreasing and satisfy $\sum_p \tau_p(t) = t$ for all $t \geq 0$, and ρ_p satisfy $\sum_p \rho_p(t) = 1$ for all $t \geq 0$. In contrast, it is possible that $\sum_p \hat{\rho}_p > 1$.

Given a family of scalars $\{a_p \in \mathbb{R} : p \in \mathcal{P}\}$, we define the *asymptotic weighted average* by

$$\hat{a} := \limsup_{t \to \infty} \sum\nolimits_{p \in \mathcal{P}} a_p \rho_p(t), \tag{10}$$

and the *maximal weighted average* over $[0, T]$ by

$$\bar{a}(T) := \max_{t \in [0,T]} \sum\nolimits_{p \in \mathcal{P}} a_p \tau_p(t)/T. \quad (11)$$

Lemma 1. *The asymptotic weighted average $\hat{a}$ and maximal weighted average $\bar{a}$ satisfy $\limsup_{T \to \infty} \bar{a}(T) = \max\{\hat{a}, 0\}$.*

III. ENTROPY OF GENERAL SWITCHED LINEAR SYSTEMS

Our main goal is to study the entropy of a switched linear system with a family of matrices $\{A_p \in \mathbb{R}^{n \times n} : p \in \mathcal{P}\}$:

$$\dot{x} = A_\sigma x, \qquad x(0) \in K. \quad (12)$$

Viewing matrices as linear operators, we denote the entropy of (12) by $h(A_\sigma, K)$. In this section, we first prove $h(A_\sigma, K)$ is independent of K, and provide standard constructions of spanning and separated sets using grids. Next, we present a result for the non-switched case. Finally, we establish general upper and lower bounds for the entropy of (12).

A. Initial set and grid

Proposition 2. *The topological entropy of (12) is independent of the choice of the initial set K.*

Proof. First, the solution of (12) at time t with switching signal σ and initial state x satisfies $\xi_\sigma(x, t) = \Phi_\sigma(t, 0)\, x$ with the state-transition matrix $\Phi_\sigma(t, 0)$.

Second, we show the entropy of (12) is invariant under translation and uniform scaling of the initial set. Let K_1 be an initial set, and define $K_2 := \{sx + v : x \in K_1\}$ for some scalar $s > 0$ and vector $v \in \mathbb{R}^n$. Given $T, \varepsilon > 0$, let E_1 be a minimal (T, ε)-spanning set of K_1. For an $x_2 \in K_2$, let $x_1 := (x_2 - v)/s \in K_1$. Then there is an $\hat{x}_1 \in E_1$ such that $\max_{t \in [0,T]} \|\Phi_\sigma(t, 0)(x_1 - \hat{x}_1)\| < \varepsilon$; thus $\max_{t \in [0,T]} \|\Phi_\sigma(t, 0)(x_2 - \hat{x}_2)\| < s\varepsilon$ with $\hat{x}_2 := s\hat{x}_1 + v$. Hence $E_2 := \{s\hat{x} + v : \hat{x} \in E_1\}$ is a $(T, s\varepsilon)$-spanning set of K_2. As $|E_2| = |E_1|$, we get $S(A_\sigma, s\varepsilon, T, K_2) \leq S(A_\sigma, \varepsilon, T, K_1)$ and thus $h(A_\sigma, K_2) \leq h(A_\sigma, K_1)$. Similar arguments with separated sets imply $h(A_\sigma, K_2) \geq h(A_\sigma, K_1)$. Hence $h(A_\sigma, K_2) = h(A_\sigma, K_1)$.

Finally, we show the entropy of (12) is independent of the choice of the initial set. Let K be an initial set. As K is a compact set with a nonempty interior, there exist closed balls $B_1, B_2 \subset \mathbb{R}^n$ such that $B_1 \subset K \subset B_2$; thus $h(A_\sigma, B_1) \leq h(A_\sigma, K) \leq h(A_\sigma, B_2)$. The result from the second step implies $h(A_\sigma, B_1) = h(A_\sigma, B_2)$ and thus $h(A_\sigma, B_1) = h(A_\sigma, K) = h(A_\sigma, B_2)$. Therefore, $h(A_\sigma, K)$ is independent of the choice of K. $\square$

Following Proposition 2, we omit the initial set K and denote by $h(A_\sigma)$ the entropy of (12). For concreteness, we take $K := \{x \in \mathbb{R}^n : \|x\| \leq 1\}$ in the following analysis.

Next, given $T, \varepsilon > 0$, we construct standard (T, ε)-spanning and (T, ε)-separated sets using grids. Given a vector $\theta \in \mathbb{R}^n_{>0}$, we define the following grid on K:

$$G(\theta) := \{(k_1\theta_1, \ldots, k_n\theta_n) \in K : k_1, \ldots, k_n \in \mathbb{Z}\}. \quad (13)$$

Its cardinality satisfies $|G(\theta)| = \prod_i (2\lfloor 1/\theta_i \rfloor + 1)$. For an $\hat{x} \in G(\theta)$, denote by $R(\hat{x})$ the open hyperrectangle in K with center $\hat{x}$ and sides $2\theta_1, \ldots, 2\theta_n$, that is,

$$R(\hat{x}) := \{x \in K : |x_i - \hat{x}_i| < \theta_i \text{ for } i = 1, \ldots, n\}. \quad (14)$$

Then the points in $G(\theta)$ adjacent to $\hat{x}$ are on the boundary of the closure of $R(\hat{x})$, and the union of all $R(\hat{x})$ covers K.

Lemma 2. *1) If $R(\hat{x}) \subset B_{A_\sigma}(\hat{x}, \varepsilon, T)$ for all $\hat{x} \in G(\theta)$, then the grid $G(\theta)$ is (T, ε)-spanning; thus*

$$\log S(A_\sigma, \varepsilon, T, K) \leq \sum\nolimits_{i=1}^n \log(2/\theta_i + 1). \quad (15)$$

2) If $B_{A_\sigma}(\hat{x}, \varepsilon, T) \subset R(\hat{x})$ for all $\hat{x} \in G(\theta)$, then the grid $G(\theta)$ is (T, ε)-separated; thus

$$\log N(A_\sigma, \varepsilon, T, K) \geq \sum\nolimits_{i=1}^n \log(2/\theta_i - 1). \quad (16)$$

B. Entropy of linear time-invariant systems

Consider a linear time-invariant (LTI) system

$$\dot{x} = Ax, \qquad x(0) \in K \quad (17)$$

with a matrix $A \in \mathbb{R}^{n \times n}$. The next result provides a formula for the entropy $h(A)$. The proof is along the lines of those of the corresponding discrete-time results such as [4, Th. 15] or [27, Th. 2.4.2]; see [28, Ch. 4] for a complete proof.

Proposition 3. *The topological entropy of (17) satisfies*

$$h(A) = \sum\nolimits_{\lambda \in \mathrm{spec}(A)} \max\{\mathrm{Re}(\lambda), 0\}. \quad (18)$$

C. Entropy of general switched linear systems

Now we establish general upper and lower bounds for the entropy of the switched linear system (12).

Theorem 4. *The topological entropy of (12) satisfies*

$$h(A_\sigma) \leq \limsup_{t \to \infty} \sum\nolimits_{p \in \mathcal{P}} n\|A_p\|\rho_p(t) \quad (19)$$

and also

$$h(A_\sigma) \geq \max\left\{ \limsup_{t \to \infty} \sum\nolimits_{p \in \mathcal{P}} \mathrm{tr}(A_p)\rho_p(t), 0 \right\} \quad (20)$$

with the active rates ρ_p defined by (8).

Proof. First, we establish (19). Let $t_1 < \cdots < t_{N_\sigma(t,0)}$ be the sequence of switches on $(0, t] =: (t_0, t_{N_\sigma(t,0)+1}]$. For initial states $x, x' \in K$, the corresponding solutions at time t with switching signal σ satisfy $\|\xi_\sigma(x', t) - \xi_\sigma(x, t)\| \leq e^{\sum_{l=0}^{N_\sigma(t,0)} \|A_{\sigma(t_l)}\|(t_{l+1}-t_l)}\|x' - x\| = e^{\sum_p \|A_p\|\tau_p(t)}\|x' - x\|$ with the active times τ_p defined by (7). As τ_p are non-decreasing functions, $\max_{t \in [0,T']} \|\xi_\sigma(x', t) - \xi_\sigma(x, t)\| \leq e^{\sum_p \|A_p\|\tau_p(T')}\|x' - x\|$ for all $T' \geq 0$. Given $T, \varepsilon > 0$, consider the grid $G(\theta)$ and hypercubes $R(\hat{x})$ defined by (13) and (14) with $\theta_i := e^{-\sum_p \|A_p\|\tau_p(T)}\varepsilon$, respectively. Clearly, $R(\hat{x}) \subset B_{A_\sigma}(\hat{x}, \varepsilon, T)$ for all $\hat{x} \in G(\theta)$. Then Lemma 2 implies $G(\theta)$ is (T, ε)-spanning, and substituting (15) into (5) yields $h(A_\sigma) \leq \limsup_{T \to \infty} \sum_p n\|A_p\|\rho_p(T)$.

Second, we prove (20) via volume-based arguments. Given $T, \varepsilon > 0$, an open ball $B_{A_\sigma}(x, \varepsilon, T)$ defined by (3) satisfies $B_{A_\sigma}(x, \varepsilon, T) \subset \{x' \in K : \|\xi_\sigma(x', T) - \xi_\sigma(x, T)\| < \varepsilon\} =$

$\{\Phi_\sigma(T,0)^{-1}v + x \in K : \|v\| < \varepsilon\}$. Thus its volume satisfies $\mathrm{vol}(B_{A_\sigma}(x,\varepsilon,T)) \leq (2\varepsilon)^n/\det(\Phi_\sigma(0,T)) = e^{-\sum_{l=0}^{N_\sigma(t,0)} \mathrm{tr}(A_{\sigma(t_l)})(t_{l+1}-t_l)}(2\varepsilon)^n = e^{-\sum_p \mathrm{tr}(A_p)\tau_p(T)}(2\varepsilon)^n$, where the first equality uses Liouville's trace formula. For a (T,ε)-spanning set E, the previous volume bound and (4) implies $\mathrm{vol}(K) \leq |E|e^{-\sum_p \mathrm{tr}(A_p)\tau_p(T)}(2\varepsilon)^n$. Hence the minimal cardinality of a (T,ε)-spanning set satisfies $S(A_\sigma, \varepsilon, T) \geq e^{\sum_p \mathrm{tr}(A_p)\tau_p(T)}\mathrm{vol}(K)/(2\varepsilon)^n$, and thus $h(A_\sigma) \geq \lim_{\varepsilon \searrow 0} \limsup_{T\to\infty} \log(e^{\sum_p \mathrm{tr}(A_p)\tau_p(T)})/T = \limsup_{T\to\infty} \sum_p \mathrm{tr}(A_p)\rho_p(T)$. The proof of (20) is completed by recalling the entropy is nonnegative. $\square$

Remark 2. The upper bound (19) has been formulated using ∞-norms but actually holds for all induced matrix norms (which is consistent with Remark 1), as all norms on $\mathbb{R}^n$ are equivalent up to a multiplicative factor [29, Cor. 5.4.5] (which would disappear after dividing by T and taking the limit supremum). Hence it can be improved by taking the infimum over all induced matrix norms.

There is in general a gap between the upper bound (19) and lower bound (20) in Theorem 4 (e.g., consider the LTI system (17) with a matrix of one positive and one negative eigenvalue). For a general switched linear system, due to the lack of "alignment" between eigenspaces of different modes, we are not able to construct an exact formula for the entropy $h(A_\sigma)$. In Sections V–VI, we examine switched linear systems with special matrix structure, and establish formulae and tighter bounds for the entropy.

IV. Entropy of Switched Scalar Systems

In this section, we consider switched linear systems with a scalar-valued state. Then each A_p becomes a scalar $a_p \in \mathbb{R}$, and (12) becomes a switched scalar system

$$\dot{x} = a_\sigma x, \qquad x(0) \in K \subset \mathbb{R}. \tag{21}$$

The next theorem provides a formula for the entropy $h(a_\sigma)$.

Theorem 5. *The topological entropy of* (21) *satisfies*

$$h(a_\sigma) = \max\{\hat{a}, 0\} \tag{22}$$

with the asymptotic weighted average $\hat{a}$ defined by (10).

Theorem 5 follows from Lemma 1 and the next result.

Lemma 3. *The entropy satisfies $h(a_\sigma) = \limsup_{T\to\infty} \bar{a}(T)$ with the maximal weighted average $\bar{a}$ defined by* (11).

Proof. For all initial states $x, x' \in K$, the corresponding solutions at time t with switching signal σ satisfy $|\xi_\sigma(x',t) - \xi_\sigma(x,t)| = e^{\sum_p a_p \tau_p(t)}|x' - x|$ with the active times τ_p defined by (7). Let $\bar{\eta}(T) := \max_{t\in[0,T]} \sum_p a_p \tau_p(t) = \bar{a}(T)T$. Then $\max_{t\in[0,T']} |\xi_\sigma(x',t) - \xi_\sigma(x,t)| = e^{\bar{\eta}(T')}|x' - x|$ for all $T' \geq 0$. Given $T, \varepsilon > 0$, consider the grid $G(\theta)$ and hypercubes $R(\hat{x})$ defined by (13) and (14) with $\theta := e^{-\bar{\eta}(T)}\varepsilon$, respectively. Clearly, $R(\hat{x}) = B_{a_\sigma}(\hat{x}, \varepsilon, T)$ for all $\hat{x} \in G(\theta)$. Then Lemma 2 implies $G(\theta)$ is both (T,ε)-spanning and (T,ε)-separated, and substituting (15) and (16) into (5) yields $h(a_\sigma) \leq \limsup_{T\to\infty} \bar{a}(T)$ and $h(a_\sigma) \geq \limsup_{T\to\infty} \bar{a}(T)$, respectively. $\square$

For a switched scalar system, the formula (22) coincides with the general lower bound (20). However, for non-scalar cases, the volume-based approach of Theorem 4 generally yields a strictly more conservative bound than the separated-set-based approach of Lemma 3.

Based on the formula (22), we derive the following upper bounds for the entropy of the switched scalar system (21).

Corollary 6. *The topological entropy of* (21) *satisfies*

$$h(a_\sigma) \leq \sum_{p\in\mathcal{P}} h(a_p)\hat{\rho}_p \tag{23}$$

with the asymptotic active rates $\hat{\rho}_p$ defined by (9), *and also*

$$h(a_\sigma) \leq \max_{p\in\mathcal{P}} h(a_p), \tag{24}$$

where each $h(a_p)$ denotes the topological entropy of the p-th mode and satisfies (18), *that is, $h(a_p) = \max\{a_p, 0\}$.*

Remark 3. When the scalars a_p are complex and state space is extended from $\mathbb{R}$ to $\mathbb{C}$, the results in this section still hold after replacing a_p with the real parts $\mathrm{Re}(a_p)$ in (10) and (11) and noticing that (18) implies $h(a_p) = \max\{\mathrm{Re}(a_p), 0\}$, since the imaginary parts $\mathrm{Im}(a_p)$ do not affect the entropy. This can be seen from $|\xi_\sigma(x',t) - \xi_\sigma(x,t)| = e^{\sum_p \mathrm{Re}(a_p)\tau_p(t)}|x' - x|$ for the solutions of (21).

Compared with (22), the upper bound (23) is the same for all switching signals with the same asymptotic active rates $\hat{\rho}_p$; the upper bound (24) is independent of the switching. In general, the relation between (23) and (24) is unknown.

V. Entropy of Switched Diagonal Systems

In this section, we consider switched linear systems with simultaneously diagonalizable matrices, that is, there exists a (possibly complex) change of basis under which the matrices A_p in (12) all become diagonal (or equivalently, A_p are all diagonalizable and commute pairwise [29, Th. 1.3.19]). Hence we assume, without loss of generality, that every A_p is diagonal, and denote it by $D_p := \mathrm{diag}(a_p^1, \ldots, a_p^n) \in \mathbb{C}^{n\times n}$, that is, a_p^i is the i-th diagonal entry of D_p. Then (12) becomes the switched diagonal system in $\mathbb{C}^n$

$$\dot{x} = D_\sigma x, \qquad x(0) \in K. \tag{25}$$

The next theorem provides a formula for the entropy $h(D_\sigma)$.

Theorem 7. *The topological entropy of* (25) *satisfies*

$$h(D_\sigma) = \limsup_{T\to\infty} \sum_{i=1}^n \bar{a}_i(T) \tag{26}$$

with the component-wise maximal weighted averages over $[0,T]$ defined by $\bar{a}_i(T) := \max_{t\in[0,T]} \sum_p \mathrm{Re}(a_p^i)\tau_p(t)/T \geq 0$ for $i = 1, \ldots, n$, where τ_p are defined by (7).

Proof. For all initial states $x, x' \in K$, the corresponding solutions at time t with switching signal σ satisfy $\|\xi_\sigma(x',t) - \xi_\sigma(x,t)\| = \max_i e^{\sum_p \mathrm{Re}(a_p^i)\tau_p(t)}|x_i' - x_i|$. Let $\bar{\eta}_i(T) := \max_{t\in[0,T]} \sum_p \mathrm{Re}(a_p^i)\tau_p(t) = \bar{a}_i(T)T$ for $i = 1, \ldots, n$. Then $\max_{t\in[0,T']} \|\xi_\sigma(x',t) - \xi_\sigma(x,t)\| = \max_i e^{\bar{\eta}_i(T')}|x_i' - x_i|$ for all $T' \geq 0$. Given $T, \varepsilon > 0$, consider the grid $G(\theta)$ and hypercubes $R(\hat{x})$ defined by (13) and (14) with $\theta_i :=$

$\exp(-\bar\eta_i(T))\,\varepsilon$, respectively. Clearly, $R(\hat x) = B_{D_\sigma}(\hat x, \varepsilon, T)$ for all $\hat x \in G(\theta)$. Then Lemma 2 implies that $G(\theta)$ is both (T,ε)-spanning and (T,ε)-separated, and substituting (15) and (16) into (5) yields $h(D_\sigma) \le \limsup_{T\to\infty} \sum_i \bar a_i(T)$ and $h(D_\sigma) \ge \limsup_{T\to\infty} \sum_i \bar a_i(T)$, respectively. $\square$

While the formula (26) is rather complex, it gives the exact value of $h(D_\sigma)$. It also leads to the following upper bounds for the entropy of the switched diagonal system (25).

Proposition 8. *The topological entropy of* (25) *satisfies*

$$h(D_\sigma) \le \sum_{i=1}^{n} h(a_\sigma^i), \qquad (27)$$

where each $h(a_\sigma^i)$ satisfies $h(a_\sigma^i) = \max\{\hat a_i, 0\}$ with the component-wise asymptotic weighted averages defined by $\hat a_i := \limsup_{t\to\infty} \sum_p \mathrm{Re}(a_p^i)\rho_p(t)$ for $i = 1,\ldots,n$, where the active rates ρ_p are defined by (8). Moreover, (27) holds with equality for all switching signals σ such that the active rates $\rho_p(t)$ converge as $t\to\infty$ for all $p\in\mathcal{P}$.

Proposition 9. *The topological entropy of the switched diagonal system* (25) *is upper-bounded by*

$$h(D_\sigma) \le \limsup_{t\to\infty} \sum_{p\in\mathcal{P}} h(D_p)\rho_p(t) \qquad (28)$$

with the active rates ρ_p defined by (8), where each $h(D_p)$ denotes the topological entropy of the p-th mode and satisfies (18). Moreover, (28) holds with equality if $\mathrm{Re}(a_p^i) \ge 0$ for all $i \in \{1,\ldots,n\}$ and $p \in \mathcal{P}$.

Corollary 10. *The topological entropy of* (25) *satisfies*

$$h(D_\sigma) \le \sum_{p\in\mathcal{P}} h(D_p)\hat\rho_p \qquad (29)$$

with the asymptotic active rates $\hat\rho_p$ defined by (9), and also

$$h(D_\sigma) \le \max_{p\in\mathcal{P}} h(D_p), \qquad (30)$$

where each $h(D_p)$ denotes the topological entropy of the p-th mode and satisfies (18).

Unlike the formula (26) and upper bound (27), the upper bounds (28)–(30) are independent of the relative order of the scalar components between the matrices D_p in (25), and can thus be calculated without diagonalization. For a fixed family of matrices $\{D_p : p \in \mathcal{P}\}$, the upper bounds (27) and (28) depend only on the active rates ρ_p, the upper bound (29) only on the asymptotic active rates $\hat\rho_p$; the upper bound (30) is independent of switching. In general, the relation between the upper bounds (27) and (28), and that between the upper bounds (29) and (30), are both unknown. Meanwhile, both (27) and (28) imply (29), whereas only (28) implies (30).

VI. Entropy of switched triangular systems

In this section, we consider switched linear systems with simultaneously triangularizable matrices, that is, there exists a (possibly complex) change of basis under which the matrices A_p in (12) are all upper-triangular. Sufficient conditions for simultaneous triangularizability include that the matrices A_p commute pairwise [29, Th. 2.3.3], or, more generally, that the Lie algebra $\{A_p : p \in \mathcal{P}\}_{LA}$ is solvable (Lie's theorem [30, Th. A$''$]). Hence we assume, without loss of generality, that every A_p is upper-triangular, and denote it by

$$U_p := \begin{bmatrix} a_p^1 & b_p^{1,2} & \cdots & b_p^{1,n} \\ 0 & a_p^2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & b_p^{n-1,n} \\ 0 & \cdots & 0 & a_p^n \end{bmatrix} \in \mathbb{C}^{n\times n}.$$

Then (12) becomes the switched triangular system in $\mathbb{C}^n$:

$$\dot x = U_\sigma x, \qquad x(0) \in K. \qquad (31)$$

The next theorem provides an upper bound for $h(U_\sigma)$.

Theorem 11. *The topological entropy of* (31) *satisfies*

$$h(U_\sigma) \le \limsup_{T\to\infty} \left(n\bar a_1(T) + \sum_{i=2}^{n} (n+1-i)\bar d_i(T) \right) \quad (32)$$

with $\bar a_1(T) := \max_{t\in[0,T]} \sum_p \mathrm{Re}(a_p^1)\tau_p(t)/T \ge 0$ and $\bar d_i(T) := \max_{t\in[0,T]} \sum_p \mathrm{Re}(a_p^i - a_p^{i-1})\tau_p(t)/T \ge 0$ for $i = 2,\ldots,n$, where the active times τ_p are defined by (7).

Proof. The proof of Theorem 11 consists of three steps. In this proof, we omit the fixed switching signal σ and denote by $\xi_i(x,t)$ the i-th scalar component of the solution $\xi_\sigma(x,t)$.

First, we derive a formula for the solution $\xi_\sigma(x,t)$ of (31) at time t with switching signal σ and initial state x. Let $\eta_i(t) := \sum_p a_p^i \tau_p(t)$ for $i = 1,\ldots,n$, and $\nu_{i,j}(t) := \eta_j(t) - \eta_i(t)$ for $1 \le i < j \le n$.

Lemma 4. *For $k = 0,\ldots,n-1$, the $(n-k)$-th scalar component of $\xi_\sigma(x,t)$ satisfies* (33) *below with the sets $\mathcal{C}_{k,l,i} := \{(c_0,\ldots,c_i) \in \mathbb{N}^{i+1} : n-k = c_0 < c_1 < \cdots < c_{i-1} < c_i = n-l\}$ for $l = 0,\ldots,k-1$ and $i = 1,\ldots,k-l$.*

Second, for all initial states $x,x' \in K$, we estimate $\|\xi_\sigma(x',t) - \xi_\sigma(x,t)\|$ based on the formula (33). Let $\bar\eta_i(T) := \max_{t\in[0,T]} \mathrm{Re}(\eta_i(t)) = \bar a_i(T)\,T$ for $i = 1,\ldots,n$, and $\bar\nu_{i,j}(T) := \max_{t\in[0,T]} \mathrm{Re}(\nu_{i,j}(t))$ for $1 \le i < j \le n$. Then $\bar\nu_{i-1,i}(T') = \bar d_i(T')\,T'$ for all $T' \ge 0$ and $i \in \{2,\ldots,n\}$.

Lemma 5. *It holds that $\max_{t\in[0,T']} \|\xi_\sigma(x',t) - \xi_\sigma(x,t)\| \le \sum_i |x_i' - x_i| e^{\bar\eta_1(T') + \sum_{j=1}^{i-1} \bar\nu_{j,j+1}(T')} P_i(T')$ for all $T' \ge 0$ with polynomials P_i which are positive and increasing on $\mathbb{R}_+$.*

Finally, given $T,\varepsilon > 0$, consider the grid $G(\theta)$ and hypercubes $R(\hat x)$ defined by (13) and (14) with $\theta_i :=$

$$\xi_{n-k}(x,t) = e^{\eta_{n-k}(t)}\left(x_{n-k} + \sum_{l=0}^{k-1} \left(x_{n-l} \sum_{i=1}^{k-l} \sum_{(c_0,\ldots,c_i)\in\mathcal{C}_{k,l,i}} \int_0^t \int_0^{s_1} \cdots \int_0^{s_{i-1}} \prod_{j=1}^{i} b_{\sigma(s_j)}^{c_{j-1},c_j} e^{\nu_{c_{j-1},c_j}(s_j)} \, ds_j \right) \right) \qquad (33)$$

$\exp(-\bar{\eta}_1(T) - \sum_{j=1}^{i-1} \bar{\nu}_{j,j+1}(T))\, \varepsilon/(nP_i(T))$, respectively. Clearly, $R(\hat{x}) \subset B_{U_\sigma}(\hat{x})$ for all $\hat{x} \in G(\theta)$. Then Lemma 2 implies that $G(\theta)$ is (T, ε)-spanning, and substituting (15) into (5) yields $h(U_\sigma) \leq \limsup_{T \to \infty} \big(n\bar{a}_1(T) + \sum_{i=2}^{n}(n+1-i)\bar{d}_i(T)\big)$. In particular, the off-diagonal entries $b_p^{i,j}$ of the matrices U_p in (31) are absorbed into the polynomials P_i, and thus do not appear in the bound (32). $\qquad \square$

Based on (32), we derive the following upper bounds for the entropy of the switching triangular system (31).

Proposition 12. *The topological entropy of* (31) *satisfies*

$$h(U_\sigma) \leq nh(a_\sigma^1) + \sum_{i=2}^{n}(n+1-i)h(d_\sigma^i), \qquad (34)$$

where $h(a_\sigma^1) = \max\{\hat{a}_1, 0\}$ *and* $h(d_\sigma^i) = \max\{\hat{d}_i, 0\}$ *for* $i = 2, \dots, n$ *with* $\hat{a}_1 := \limsup_{t \to \infty} \sum_p \operatorname{Re}(a_p^1)\rho_p(t)$ *and* $\hat{d}_i := \limsup_{t \to \infty} \sum_p \operatorname{Re}(a_p^i - a_p^{i-1})\rho_p(t)$, *where the active rates* ρ_p *are defined by* (8).

Proposition 13. *The topological entropy of* (31) *satisfies*

$$h(U_\sigma) \leq \limsup_{t \to \infty} \sum_{p \in \mathcal{P}} \hat{h}(U_p)\rho_p(t) \qquad (35)$$

with $\hat{h}(U_p) := n\max\{\operatorname{Re}(a_p^1), 0\} + \sum_{i=2}^{n}(n+1-i) \times \max\{\operatorname{Re}(a_p^i - a_p^{i-1}), 0\}$ *for* $p \in \mathcal{P}$, *where the active rates* ρ_p *are defined by* (8).

Corollary 14. *The topological entropy of* (31) *satisfies*

$$h(U_\sigma) \leq \sum_{p \in \mathcal{P}} \hat{h}(U_p)\hat{\rho}_p \qquad (36)$$

with the asymptotic active rates $\hat{\rho}_p$ *defined by* (9)*, and by*

$$h(U_\sigma) \leq \max_{p \in \mathcal{P}} \hat{h}(U_p). \qquad (37)$$

For a fixed family of matrices $\{U_p : p \in \mathcal{P}\}$, the upper bounds (34) and (35) depend only on the active rates ρ_p, the upper bound (36) only on the same asymptotic active rates $\hat{\rho}_p$; the upper bound (37) is independent of switching. The upper bound (32) implies (34), (35), (36) and (37). In general, the relation between the upper bounds (34) and (35), and that between the upper bounds (36) and (37), are unknown. Meanwhile, both (34) and (35) imply (36), whereas only (35) implies (37).

VII. Future research

The topological entropy proposed in this paper depends on switching. For switched systems with unknown switching signal, a different entropy notion is needed to capture the additional uncertainty of the trajectory and to quantify the extra information needed for stabilization. Sufficient data rate for feedback stabilization of switched linear systems were established in [22], [23]. A similar data-rate bound for state-estimation was formulated in [24]. These data-rate bounds should be upper bounds for the entropy notion to be defined.

References

[1] A. N. Kolmogorov, "A new metric invariant of transitive dynamical systems and automorphisms of Lebesgue spaces," *Dokl. Akad. Nauk SSSR*, vol. 119, no. 5, pp. 861–864, 1958.

[2] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, no. 3, pp. 379–423, 1948.

[3] R. L. Adler, A. G. Konheim, and M. H. McAndrew, "Topological entropy," *Trans. Am. Math. Soc.*, vol. 114, no. 2, pp. 309–319, 1965.

[4] R. Bowen, "Entropy for group endomorphisms and homogeneous spaces," *Trans. Am. Math. Soc.*, vol. 153, pp. 401–414, 1971.

[5] E. I. Dinaburg, "The relation between topological entropy and metric entropy," *Dokl. Akad. Nauk SSSR*, vol. 190, no. 1, pp. 19–22, 1970.

[6] R. Bowen, "Periodic points and measures for axiom A diffeomorphisms," *Trans. Am. Math. Soc.*, vol. 154, pp. 377–397, 1971.

[7] S. Kolyada and L. Snoha, "Topological entropy of nonautonomous dynamical systems," *Random & Comput. Dyn.*, vol. 4, no. 2-3, pp. 205–233, 1996.

[8] C. Kawan and Y. Latushkin, "Some results on the entropy of non-autonomous dynamical systems," *Dyn. Syst.*, vol. 31, no. 3, pp. 251–279, 2016.

[9] G. N. Nair, R. J. Evans, I. M. Y. Mareels, and W. Moran, "Topological feedback entropy and nonlinear stabilization," *IEEE Trans. Automat. Contr.*, vol. 49, no. 9, pp. 1585–1597, 2004.

[10] F. Colonius and C. Kawan, "Invariance entropy for control systems," *SIAM J. Control Optim.*, vol. 48, no. 3, pp. 1701–1721, 2009.

[11] F. Colonius, C. Kawan, and G. N. Nair, "A note on topological feedback entropy and invariance entropy," *Syst. & Control Lett.*, vol. 62, no. 5, pp. 377–381, 2013.

[12] F. Colonius, "Minimal bit rates and entropy for exponential stabilization," *SIAM J. Control Optim.*, vol. 50, no. 5, pp. 2988–3010, 2012.

[13] D. Liberzon, *Switching in Systems and Control*. Birkhäuser Boston, 2003.

[14] K. S. Narendra and J. Balakrishnan, "A common Lyapunov function for stable LTI systems with commuting A-matrices," *IEEE Trans. Automat. Contr.*, vol. 39, no. 12, pp. 2469–2471, 1994.

[15] L. Gurvits, "Stability of discrete linear inclusion," *Linear Algebra Appl.*, vol. 231, pp. 47–85, 1995.

[16] S. A. Kutepov, "Algebraic criterion of absolute stability," *Kibern. i Vychislitel'naia Tekhnika*, vol. 54, pp. 52–55, 1982.

[17] D. Liberzon, J. P. Hespanha, and A. S. Morse, "Stability of switched systems: A Lie-algebraic condition," *Syst. & Control Lett.*, vol. 37, no. 3, pp. 117–122, 1999.

[18] S. A. Kutepov, "Absolute stability of bilinear systems with a compact Levi factor," *Kibern. i Vychislitel'naia Tekhnika*, vol. 62, no. 118, pp. 28–33, 1984.

[19] A. A. Agrachev and D. Liberzon, "Lie-algebraic stability criteria for switched systems," *SIAM J. Control Optim.*, vol. 40, no. 1, pp. 253–269, 2001.

[20] J. P. Hespanha, A. Ortega, and L. Vasudevan, "Towards the control of linear systems with minimum bit-rate," in *15th Int. Symp. Math. Theory Networks Syst.*, 2002, pp. 1–15.

[21] S. Tatikonda and S. Mitter, "Control under communication constraints," *IEEE Trans. Automat. Contr.*, vol. 49, no. 7, pp. 1056–1068, 2004.

[22] D. Liberzon, "Finite data-rate feedback stabilization of switched and hybrid linear systems," *Automatica*, vol. 50, no. 2, pp. 409–420, 2014.

[23] G. Yang and D. Liberzon, "Feedback stabilization of a switched linear system with an unknown disturbance under data-rate constraints," *IEEE Trans. Automat. Contr.*, vol. 63, no. 7, pp. 2107–2122, 2018.

[24] H. Sibai and S. Mitra, "Optimal data rate for state estimation of switched nonlinear systems," in *20th Int. Conf. Hybrid Syst. Comput. Control*, 2017, pp. 71–80.

[25] D. Liberzon and S. Mitra, "Entropy and minimal bit rates for state estimation and model detection," *IEEE Trans. Automat. Contr.*, to appear.

[26] A. Katok and B. Hasselblatt, *Introduction to the Modern Theory of Dynamical Systems*. Cambridge University Press, 1995.

[27] A. S. Matveev and A. V. Savkin, *Estimation and Control over Communication Networks*. Birkhäuser Boston, 2009.

[28] A. J. Schmidt, "Topological entropy bounds for switched linear systems with Lie structure," M.S. thesis, University of Illinois at Urbana-Champaign, 2016.

[29] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 1985.

[30] H. Samelson, *Notes on Lie Algebras*. Springer New York, 1990.