2018 56th Annual Allerton Conference on Communication, Control, and Computing (Allerton)
Allerton Park and Retreat Center
Monticello, IL, USA, October 2-5, 2018

On topological entropy of switched linear systems with pairwise commuting matrices

Guosong Yang and João P. Hespanha

Abstract— **We study a notion of topological entropy for switched systems, formulated in terms of the minimal number of initial states needed to approximate all initial states within a finite precision. This paper focuses on the topological entropy of switched linear systems with pairwise commuting matrices. First, we prove there exists a simultaneous change of basis under which each of the matrices can be decomposed into a diagonal part and a nilpotent part, and all the diagonal and nilpotent parts are pairwise commuting. Then a formula for the topological entropy is established in terms of the component-wise averages of the eigenvalues, weighted by the active time of each mode, which indicates that the topological entropy is independent of the nilpotent parts above. We also present how the formula generalizes known results for the non-switched case and the case with simultaneously diagonalizable matrices, and construct more general but more conservative upper bounds for the entropy. A numerical example is provided to demonstrate properties of the formula and the upper and lower 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, which motivates us to study the topological entropy of switched linear systems with pairwise commuting matrices.

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 [15]–[17]. 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 [18], [19]. In [20], the notion of estimation entropy from [21] was extended to switched systems to formulate similar data-rate conditions. The paper [22] introduced a notion of topological entropy for switched systems, and established formulae and bounds for the topological entropy of switched linear systems with diagonal, triangular, and general matrices, which also serves as the basis for this work.

The main contribution of this paper is the construction of a formula and bounds for the topological entropy of switched linear systems with pairwise commuting matrices. In Section II, we present the notion of topological entropy for switched systems, define switching-related quantities such as the active time of each mode, which prove to be useful in calculating the topological entropy, and recall key results on the topological entropy of general switched linear systems. In Section III, we study the topological entropy of switched linear systems with pairwise commuting matrices. First, we prove there exists a simultaneous change of basis under which each of the matrices can be decomposed into a diagonal part and a nilpotent part, and all the diagonal and nilpotent parts are pairwise commuting. Then a formula for the topological entropy is established in terms of the component-

G. Yang and J. P. Hespanha are with the Center for Control, Dynamical Systems and Computation, University of California, Santa Barbara, CA 93106, USA. Email: {guosongyang, hespanha}@ucsb.edu. Their work was supported by the ONR MURI grant N00014-16-1-2710 and the NSF grants EPCN-1608880 and CNS-1329650.

wise active-time-weighted averages of the eigenvalues. We also present how the formula generalizes known results for the non-switched case and the case with simultaneously diagonalizable matrices, and derive more general but more conservative upper bounds for the entropy. Properties of the formula and the upper and lower bounds for the topological entropy are summarized in a remark and demonstrated via a numerical example. Section IV summarizes the paper and remarks on future research directions.

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 $\mathrm{Re}(a)$ its real part. For a vector $v \in \mathbb{C}^n$, denote by v_i its i-th scalar component and write $v = (v_1, \ldots, v_n)$. For a matrix $A \in \mathbb{C}^{n\times n}$, denote by $\mathrm{spec}(A)$ and $\mathrm{tr}(A)$ its spectrum and trace, respectively. For a set $E \subset \mathbb{C}^n$, denote by $|E|$ its cardinality. Denote by $|a|$ the absolute value of a scalar a, 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 $A = [a_{ij}]$. We call a set of pairwise commuting matrices a *commuting family*.

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. The system with index p in (1) is called the p-th *mode*, or mode p, of the switched system (2), and $\sigma(t) \in \mathcal{P}$ is called 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 that there is at most one switch at each time, and finitely many switches on each finite time interval (i.e., the set of switches contains no accumulation point).

Let $\|\cdot\|$ be some chosen norm on $\mathbb{R}^n$ and the corresponding induced norm on $\mathbb{R}^{n\times n}$. Fix an arbitrary switching signal σ. 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

$$K = \bigcup_{\hat{x}\in E} B_{f_\sigma}(\hat{x}, \varepsilon, T),$$

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, or equivalently, the cardinality of a minimal (T, ε)-spanning set. The *topological entropy* of the switched system (2) with initial set K and switching signal σ is defined in terms of the exponential growth rate of $S(f_\sigma, \varepsilon, T, K)$ by

$$h(f_\sigma, K) := \lim_{\varepsilon\searrow 0} \limsup_{T\to\infty} \frac{1}{T} \log S(f_\sigma, \varepsilon, T, K). \tag{4}$$

The entropy $h(f_\sigma, K)$ is nonnegative as $S(f_\sigma, \varepsilon, T, K)$ is nondecreasing in T, nonincreasing in ε, and at least 1. For brevity, we will refer to $h(f_\sigma, K)$ simply as the (topological) entropy of the switched system (2) in the rest of the paper.

Remark 1. In light of [23, p. 109, Prop. 3.1.2], the value of $h(f_\sigma, K)$ is the same for all metrics defining the same topology. Hence the norm $\|\cdot\|$ can be arbitrary. For convenience and 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

$$\hat{x}' \notin B_{f_\sigma}(\hat{x}, \varepsilon, T) \qquad \forall\, \hat{x}, \hat{x}' \in E$$

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, or equivalently, the cardinality of a maximal (T, ε)-separated set, which is also nondecreasing in T, nonincreasing in ε, and at least 1. The entropy of (2) can be equivalently formulated in terms of the exponential growth rate of $N(f_\sigma, \varepsilon, T, K)$ as follows; the proof is along the lines of [23, p. 110] and thus omitted here.

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

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

Remark 2. In light of [23, pp. 109–110], for a time-invariant system $\dot{x} = f(x)$, the value of $h(f, K)$ remains the same if the limit suprema in (4) and (5) are replaced with limit infima. However, this is not the case for a time-varying system, for which the subadditivity required in the proof of [23, p. 109, Lemma 3.1.5] does not necessarily hold.

B. Active time, active rates, and weighted averages

In this subsection, we defined several switching-related quantities which will be useful in calculating the entropy of switched linear systems.

For a switching signal σ, we define the *active time* of each mode over an interval $[0, t]$ by

$$\tau_p(t) := \int_0^t \mathbb{1}_p(\sigma(s))\, \mathrm{d}s, \qquad p \in \mathcal{P} \tag{6}$$

with the indicator function

$$\mathbb{1}_p(\sigma(s)) := \begin{cases} 1, & \sigma(s) = p, \\ 0, & \sigma(s) \neq p. \end{cases}$$

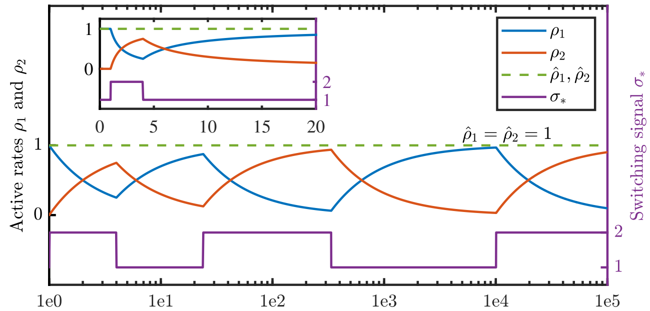

Fig. 1. A switching signal σ_* with converging set-points: the sum of the active rates $\rho_1+\rho_2 = 1$ at all times, whereas both asymptotic active rates $\hat{\rho}_1 = \hat{\rho}_2 = 1$.

We also define the *active rate* of each mode over $[0, t]$ by

$$\rho_p(t) := \tau_p(t)/t, \qquad p \in \mathcal{P} \tag{7}$$

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

$$\hat{\rho}_p := \limsup_{t\to\infty} \rho_p(t), \qquad p \in \mathcal{P}. \tag{8}$$

Clearly, the active times $\tau_p(t) \geq 0$ are nondecreasing in t and satisfy $\sum_{p\in\mathcal{P}} \tau_p(t) = t$ for all $t \geq 0$; the active rates $\rho_p(t) \in [0, 1]$ and satisfy $\sum_{p\in\mathcal{P}} \rho_p(t) = 1$ for all $t \geq 0$. In contrast, due to the limit supremum in (8), it is possible that $\sum_{p\in\mathcal{P}} \hat{\rho}_p > 1$ for the asymptotic active rates $\hat{\rho}_p$, as demonstrated in the following example.

Example 1. Consider the index set $\mathcal{P} = \{1, 2\}$ and the switching signal σ_* constructed as follows[1]:

- SWITCHING SIGNAL σ_* WITH CONVERGING SET-POINTS: Let $t_1 := 1$. For $k \geq 1$, set $t_{2k} := \min\{t > t_{2k-1} : \rho_2(t) \geq 1 - 2^{-2k}\}$ and $t_{2k+1} := \min\{t > t_{2k} : \rho_1(t) \geq 1 - 2^{-(2k+1)}\}$. Simple calculation yields $t_k = 2^k \prod_{l=1}^{k-1}(2^l - 1)$ for $k \geq 2$, and the asymptotic active rates $\hat{\rho}_1 = \hat{\rho}_2 = \limsup_{k\to\infty} 1 - e^{-2k} = 1$.

The switching signal σ_*, active rates ρ_1 and ρ_2, and asymptotic active rates $\hat{\rho}_1$ and $\hat{\rho}_1$ are plotted in Fig. 1 above (as the intervals between consecutive switches grow superexponentially, logarithmic scale is used for the long-range plot).

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

$$\hat{a} := \limsup_{t\to\infty} \sum_{p\in\mathcal{P}} a_p \rho_p(t) = \limsup_{t\to\infty} \frac{1}{t} \sum_{p\in\mathcal{P}} a_p \tau_p(t), \tag{9}$$

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

$$\bar{a}(T) := \frac{1}{T} \max_{t\in[0,T]} \sum_{p\in\mathcal{P}} a_p \tau_p(t). \tag{10}$$

[1] In all examples, we denote by $t_1 < t_2 < \cdots$ the sequence of switching times and set $\sigma = 1$ on $[t_{2k}, t_{2k+1})$ and $\sigma = 2$ on $[t_{2k+1}, t_{2k+2})$ with $t_0 := 0$.

Lemma 1. *The asymptotic weighted average $\hat{a}$ and maximal weighted average $\bar{a}$ satisfy*

$$\limsup_{T\to\infty} \bar{a}(T) = \max\{\hat{a}, 0\}. \tag{11}$$

Proof. See Appendix A. $\square$

C. Entropy of switched linear systems

In this subsection, we recall some known results on the topological entropy of the switched linear system

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

with a family of matrices $\{A_p \in \mathbb{R}^{n\times n} : p \in \mathcal{P}\}$. Thinking of matrices as linear operators, we denote by $h(A_\sigma, K)$ the entropy of (12).

First, it has been proved in [22] that the entropy $h(A_\sigma, K)$ is the same for all initial sets K.

Proposition 2 ([22, Prop. 2])**.** *The topological entropy of the switched linear system* (12) *is independent of the choice of the initial set K.*

Following Proposition 2, we omit the initial set K and denote by $h(A_\sigma)$ the entropy of (12). For convenience and concreteness, we take K to be the closed unit hypercube (recall that $\|\cdot\|$ is the ∞-norm) at the origin, that is, $K := \{x \in \mathbb{R}^n : \|x\| \leq 1\}$, in the following analysis.

Second, the entropy $h(A_\sigma)$ satisfies the following upper and lower bounds:

Proposition 3 ([22, Th. 4])**.** *The topological entropy of the switched linear system* (12) *is upper bounded by*

$$h(A_\sigma) \leq \limsup_{t\to\infty} \sum_{p\in\mathcal{P}} n\|A_p\|\rho_p(t) \tag{13}$$

and lower bounded by

$$h(A_\sigma) \geq \max\left\{ \limsup_{t\to\infty} \sum_{p\in\mathcal{P}} \mathrm{tr}(A_p)\rho_p(t),\, 0 \right\} \tag{14}$$

with the active rates ρ_p defined by (7)*.*

III. Entropy of Switched Linear Systems with Pairwise Commuting Matrices

The main objective of this paper is to examine the entropy of the switched linear system (12) for the case where $\{A_p : p \in \mathcal{P}\}$ is a commuting family.

First, we recall the following result from linear algebra:

Proposition 4 (Jordan–Chevalley decomposition [24, p. 17])**.** *For each matrix A, there exist polynomials f and g, without constant term, such that $f(A)$ is a diagonalizable matrix, $g(A)$ is a nilpotent matrix, and*[2]

$$A = f(A) + g(A).$$

Using the Jordan–Chevalley decomposition, we show that there exists a (possibly complex) simultaneous change of basis under which every matrix A_p can be written as the

[2] This result is also known as the S–N decomposition, where $f(A)$ and $g(A)$ are called the semisimple part and the nilpotent part, respectively.

sum of a diagonal matrix and a nilpotent matrix, and all the diagonal and nilpotent matrices are pairwise commuting.

Proposition 5. *For the commuting family $\{A_p : p \in \mathcal{P}\}$, there exists an invertible matrix $\Gamma \in \mathbb{C}^{n\times n}$ such that*

$$\Gamma A_p \Gamma^{-1} = D_p + N_p \qquad \forall\, p \in \mathcal{P},$$

where all $D_p \in \mathbb{C}^{n\times n}$ are diagonal matrices, all $N_p \in \mathbb{C}^{n\times n}$ are nilpotent matrices, and $\{D_p, N_p : p \in \mathcal{P}\}$ is a commuting family.

Proof. First, for each $p \in \mathcal{P}$, the Jordan–Chevalley decomposition implies there exist polynomials f_p and g_p such that $f_p(A_p)$ is a diagonalizable matrix, $g_p(A_p)$ is a nilpotent matrix, and

$$A_p = f_p(A_p) + g_p(A_p).$$

Next, as $f_p(A_p)$ and $g_p(A_p)$ are polynomials of A_p, they commute with all matrices that commute with A_p [25, p. 276, Th. 4.4.19]. Therefore, $\{f_p(A_p), g_p(A_p) : p \in \mathcal{P}\}$ is a commuting family. In particular, the subset $\{f_p(A_p) : p \in \mathcal{P}\}$ is a commuting family of diagonalizable matrices. Hence there exists an invertible matrix $\Gamma \in \mathbb{C}^{n\times n}$ such that

$$D_p := \Gamma f_p(A_p)\Gamma^{-1}, \qquad p \in \mathcal{P}$$

are all diagonal matrices [26, p. 52, Th. 1.3.19]. Moreover, as changing the basis preserves both matrix nilpotency and commutativity, it follows that

$$N_p := \Gamma g_p(A_p)\Gamma^{-1}, \qquad p \in \mathcal{P}.$$

are all nilpotent matrices, and $\{D_p, N_p : p \in \mathcal{P}\}$ is a commuting family. ☐

In view of Proposition 5, we assume, without loss of generality, that every matrix in the commuting family $\{A_p : p \in \mathcal{P}\}$ satisfes $A_p = D_p + N_p$ with a diagonal matrix $D_p := \operatorname{diag}(a_p^1, \dots, a_p^n) \in \mathbb{C}^{n\times n}$, that is, a_p^i is the i-th diagonal entry of D_p, and a nilpotent matrix $N_p \in \mathbb{C}^{n\times n}$, and that $\{D_p, N_p : p \in \mathcal{P}\}$ is a commuting family. Then (12) becomes the switched linear system in $\mathbb{C}^n$ defined by

$$\dot{x} = (D_\sigma + N_\sigma)x, \qquad x(0) \in K \tag{15}$$

with the commuting family $\{D_p, N_p : p \in \mathcal{P}\}$.

In the following theorem, we establish a formula for the entropy $h(D_\sigma + N_\sigma)$ of (15).

Theorem 6. *The topological entropy of the switched linear system with pairwise commuting matrices* (15) *satisfies*

$$h(D_\sigma + N_\sigma) = \limsup_{T\to\infty} \sum_{i=1}^{n} \bar{a}_i(T) \tag{16}$$

with the component-wise maximal weighted averages over $[0,T]$ defined by

$$\bar{a}_i(T) := \frac{1}{T}\max_{t\in[0,T]} \sum_{p\in\mathcal{P}} \operatorname{Re}(a_p^i)\tau_p(t) \geq 0, \qquad i = 1,\dots,n,$$

where the active times τ_p are defined by (6).

In particular, (16) implies that the entropy $h(D_\sigma + N_\sigma)$ is independent of the nilpotent part of (15). To prove Theorem 6, we first formulate an estimate for the effect of the nilpotent matrices N_p.

Lemma 2. *Consider the commuting family of nilpotent matrices $\{N_p : p \in \mathcal{P}\}$. For each $\delta > 0$, there is a constant $c_\delta > 0$ such that for all $v \in \mathbb{C}^n$,*

$$c_\delta^{-1} e^{-\delta t}\|v\| \leq \left\| e^{\sum_{p\in\mathcal{P}} N_p\tau_p(t)} v\right\| \leq c_\delta e^{\delta t}\|v\| \tag{17}$$

for all $t \geq 0$ with the active times τ_p defined by (6).

Proof. See Appendix B. ☐

Proof of Theorem 6. For all initial states $x, x' \in K$, as $\{D_p, N_p : p \in \mathcal{P}\}$ is a commuting family, the corresponding solutions of (15) at time t with switching signal σ satisfy

$$\begin{aligned}
&\|\xi_\sigma(x', t) - \xi_\sigma(x, t)\| \\
&= \left\| e^{\sum_{p\in\mathcal{P}}(D_p+N_p)\tau_p(t)}(x' - x)\right\| \\
&= \left\| e^{\sum_{p\in\mathcal{P}} N_p\tau_p(t)} e^{\sum_{p\in\mathcal{P}} D_p\tau_p(t)}(x' - x)\right\|.
\end{aligned}$$

Given a radius $\varepsilon > 0$, Lemma 2 with $\delta = \varepsilon$ and $v = e^{\sum_{p\in\mathcal{P}} D_p\tau_p(t)}(x' - x)$ implies there is a constant $c_\varepsilon > 0$ such that

$$\begin{aligned}
&c_\varepsilon^{-1} e^{-\varepsilon t}\left\| e^{\sum_{p\in\mathcal{P}} D_p\tau_p(t)}(x' - x)\right\| \\
&\leq \|\xi_\sigma(x', t) - \xi_\sigma(x, t)\| \\
&\leq c_\varepsilon e^{\varepsilon t}\left\| e^{\sum_{p\in\mathcal{P}} D_p\tau_p(t)}(x' - x)\right\|,
\end{aligned}$$

in which

$$\left\| e^{\sum_{p\in\mathcal{P}} D_p\tau_p(t)}(x' - x)\right\| = \max_{i=1,\dots,n} e^{\sum_{p\in\mathcal{P}} \operatorname{Re}(a_p^i)\tau_p(t)}|x_i' - x_i|$$

as D_p are diagonal matrices. Define $\bar{\eta}_i : \mathbb{R}_+ \to \mathbb{R}_+$ for $i = 1,\dots,n$ by

$$\bar{\eta}_i(T) := \max_{t\in[0,T]} \sum_{p\in\mathcal{P}} \operatorname{Re}(a_p^i)\tau_p(t).$$

Then for all $T' \geq 0$,

$$\begin{aligned}
&c_\varepsilon^{-1} e^{-\varepsilon T'} \max_{i=1,\dots,n} e^{\bar{\eta}_i(T')}|x_i' - x_i| \\
&\leq \max_{t\in[0,T']} \|\xi_\sigma(x', t) - \xi_\sigma(x, t)\| \\
&\leq c_\varepsilon e^{\varepsilon T'} \max_{i=1,\dots,n} e^{\bar{\eta}_i(T')}|x_i' - x_i|.
\end{aligned} \tag{18}$$

Fix a time horizon $T \geq 0$. First, consider the grid $G(\theta)$ defined by

$$G(\theta) := \{(k_1\theta_1, \dots, k_n\theta_n) \in K : k_1, \dots, k_n \in \mathbb{Z}\} \tag{19}$$

with the vector $\theta = (\theta_1, \dots, \theta_n)$ defined by

$$\theta_i := e^{-\varepsilon T} e^{-\bar{\eta}_i(T)}\varepsilon / c_\varepsilon, \qquad i = 1,\dots,n. \tag{20}$$

Recall that we take the initial set to be $K := \{x \in \mathbb{R}^n : \|x\| \leq 1\}$. Hence the cardinality of the grid $G(\theta)$ satisfies

$$|G(\theta)| = \prod_{i=1}^{n}(2\lfloor 1/\theta_i\rfloor + 1).$$

For each $\hat{x} \in G(\theta)$, denote by $R(\hat{x})$ the open hyperrectangle in K with center $\hat{x}$ and sides $2\theta_1, \dots, 2\theta_n$, that is,

$$R(\hat{x}) := \{x \in K : |x_i - \hat{x}_i| < \theta_i \text{ for } i = 1, \dots, n\}. \quad (21)$$

Then the union of all $R(\hat{x})$ covers the initial set K, that is,

$$K = \bigcup_{\hat{x} \in G(\theta)} R(\hat{x}).$$

By comparing (20), (21), and the upper bound in (18) to (3), we see that $R(\hat{x}) \subset B_{D_\sigma + N_\sigma}(\hat{x}, \varepsilon, T)$ for all $\hat{x} \in G(\theta)$. Hence the grid $G(\theta)$ is (T, ε)-spanning, and thus the minimal cardinality of a (T, ε)-spanning set satisfies

$$S(A_\sigma, \varepsilon, T, K) \leq |G(\theta)| \leq \prod_{i=1}^{n} (2/\theta_i + 1).$$

Then the definition (4) of the topological entropy implies

$$\begin{aligned}
&h(D_\sigma + N_\sigma) \\
&\leq \lim_{\varepsilon \searrow 0} \limsup_{T \to \infty} \sum_{i=1}^{n} \frac{\log(2/\theta_i + 1)}{T} \\
&= \lim_{\varepsilon \searrow 0} \limsup_{T \to \infty} \sum_{i=1}^{n} \frac{\log(1/\theta_i)}{T} + \limsup_{T \to \infty} \sum_{i=1}^{n} \frac{\log(2 + \theta_i)}{T} \\
&= \limsup_{T \to \infty} \sum_{i=1}^{n} \frac{\bar{\eta}_i(T)}{T} + \lim_{\varepsilon \searrow 0} n\varepsilon + \limsup_{T \to \infty} \frac{n \log(c_\varepsilon/\varepsilon)}{T} \\
&= \limsup_{T \to \infty} \sum_{i=1}^{n} \frac{1}{T} \max_{t \in [0,T]} \sum_{p \in \mathcal{P}} \mathrm{Re}(a_p^i) \tau_p(t).
\end{aligned}$$

Second, consider the grid $G(\theta)$ defined by (19) with the vector $\theta = (\theta_1, \dots, \theta_n)$ defined by

$$\theta_i := e^{\varepsilon T} e^{-\bar{\eta}_i(T)} \varepsilon c_\varepsilon, \qquad i = 1, \dots, n, \quad (22)$$

and the hyperrectangles $R(\hat{x})$ with center $\hat{x} \in G(\theta)$ and sides $2\theta_1, \dots, 2\theta_n$ defined by (21). By comparing (21), (22), and the lower bound in (18) to (3), we see that $B_{D_\sigma + N_\sigma}(\hat{x}, \varepsilon, T) \subset R(\hat{x})$ for all $\hat{x} \in G(\theta)$. As the points in $G(\theta)$ adjacent to $\hat{x}$ are on the boundary of the closure of $R(\hat{x})$, the grid $G(\theta)$ is (T, ε)-separated, and thus the maximal cardinality of a (T, ε)-separated set satisfies

$$N(A_\sigma, \varepsilon, T, K) \geq |G(\theta)| \geq \prod_{i=1}^{n} (2/\theta_i - 1).$$

Then the property (5) of the topological entropy implies

$$\begin{aligned}
&h(D_\sigma + N_\sigma) \\
&\geq \lim_{\varepsilon \searrow 0} \limsup_{T \to \infty} \sum_{i=1}^{n} \frac{\log(2/\theta_i - 1)}{T} \\
&= \lim_{\varepsilon \searrow 0} \limsup_{T \to \infty} \sum_{i=1}^{n} \frac{\log(1/\theta_i)}{T} + \limsup_{T \to \infty} \sum_{i=1}^{n} \frac{\log(2 - \theta_i)}{T} \\
&= \limsup_{T \to \infty} \sum_{i=1}^{n} \frac{\bar{\eta}_i(T)}{T} - \lim_{\varepsilon \searrow 0} n\varepsilon - \limsup_{T \to \infty} \frac{n \log(c_\varepsilon \varepsilon)}{T} \\
&= \limsup_{T \to \infty} \sum_{i=1}^{n} \frac{1}{T} \max_{t \in [0,T]} \sum_{p \in \mathcal{P}} \mathrm{Re}(a_p^i) \tau_p(t). \qquad \square
\end{aligned}$$

In the following subsections, we discuss the implications of Theorem 6 in certain special scenarios, and also use (16) to derive more general but more conservative upper bounds for the entropy of the switched linear system with pairwise commuting matrices (15).

A. Formulae for entropy in special scenarios

First, for the non-switched case, Theorem 6 becomes the following well-known result (see, e.g., [4, Th. 15] and [27, Th. 4.1]).

Corollary 7. *The topological entropy of the linear time-invariant (LTI) system*

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

equals the sum of the positive real parts of the eigenvalues of A, that is,

$$h(A) = \sum_{\lambda \in \mathrm{spec}(A)} \max\{\mathrm{Re}(\lambda), 0\}. \quad (23)$$

Proof. For the non-switched case, the matrix $D + N$ in (15) can be simply taken to be a Jordan canonical form of A, with a^i being the i-th diagonal entry, and thus $\mathrm{spec}(A) = \{a^1, \dots, a^n\}$.[3] For each i, the component-wise maximal weighted averages over $[0, T]$ in (16) become $\bar{a}_i(T) = \max\{\mathrm{Re}(a^i), 0\}$ for all $T \geq 0$, and thus (16) becomes (23). $\square$

Second, for the case where $\{A_p : p \in \mathcal{P}\}$ in (12) is a commuting family of diagonalizable matrices, Proposition 5 holds with just the diagonal matrices D_p. Then (15) becomes the switched diagonal system

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

and Theorem 6 becomes [22, Th. 7]. As a direct consequence, the results in [22, Prop. 8, 9 and Cor. 10] can be generalized to obtain more general but more conservative upper bounds for $h(D_\sigma + N_\sigma)$, as demonstrated in the next subsection.

B. More general upper bounds for entropy

First, we estimate $h(D_\sigma + N_\sigma)$ in terms of the entropy in each individual scalar component.

Proposition 8. *The topological entropy of the switched linear system with pairwise commuting matrices* (15) *is upper bounded by*

$$h(D_\sigma + N_\sigma) \leq \sum_{i=1}^{n} \max\{\hat{a}_i, 0\} \quad (24)$$

with the component-wise asymptotic weighted averages defined by

$$\hat{a}_i := \limsup_{t \to \infty} \sum_{p \in \mathcal{P}} \mathrm{Re}(a_p^i) \rho_p(t), \qquad i = 1, \dots, n, \quad (25)$$

where the active rates ρ_p are defined by (7).

[3] The property $\mathrm{spec}(A) = \{a^1, \dots, a^n\}$ holds even if $D + N$ is not a Jordan canonical form, as D and N commute and are thus simultaneously triangularizable.

Proof. From (16) and the subadditivity of limit supremum, it follows that

$$h(D_\sigma + N_\sigma) = \limsup_{T\to\infty} \sum_{i=1}^n \bar{a}_i(T) \leq \sum_{i=1}^n \limsup_{T\to\infty} \bar{a}_i(T) = \sum_{i=1}^n \max\{\hat{a}_i,\, 0\},$$

where the last equality follows from Lemma 1. □

If the limits $\lim_{t\to\infty} \rho_p(t)$ exist for all $p \in \mathcal{P}$ (e.g., when the switching signal σ is periodic; see [28, Sec. 3.2.1] for more conditions), then (24) holds with equality.

Corollary 9. *Consider the switched linear system with pariwise-commuting matrices* (15). *For a switching signal σ such that the active rates $\rho_p(t)$ converge as $t \to \infty$ for all $p \in \mathcal{P}$, the topological entropy of* (15) *satisfies*

$$h(D_\sigma + N_\sigma) = \sum_{i=1}^n \max\{\hat{a}_i,\, 0\},$$

with the component-wise asymptotic weighted averages $\hat{a}_i$ defined by (25). *Equivalently, $h(D_\sigma + N_\sigma)$ equals the topological entropy* (23) *of the LTI system defined by the asymptotic weighted average matrix $A := \sum_{p\in\mathcal{P}} D_p\hat{\rho}_p$ with $\hat{\rho}_p = \lim_{t\to\infty} \rho_p(t)$.*

Second, we estimate $h(D_\sigma + N_\sigma)$ in terms of the entropy of each individual mode.

Proposition 10. *The topological entropy of the switched linear system with pairwise commuting matrices* (15) *is upper bounded by*

$$h(D_\sigma + N_\sigma) \leq \limsup_{t\to\infty} \sum_{p\in\mathcal{P}} h(D_p)\rho_p(t) \tag{26}$$

with the active rates ρ_p defined by (7)*, where $h(D_p)$ denotes the topological entropy of the p-th mode, and satisfies* (23) *with $A = D_p$ and thus $h(D_p) = h(D_p + N_p)$.*

Proof. Consider the auxiliary switched diagonal system

$$\dot{x} = \bar{D}_\sigma x, \qquad x(0) \in K \tag{27}$$

with the family of positive-semidefinite diagonal matrices $\{\bar{D}_p := \operatorname{diag}(\bar{a}_p^1, \dots, \bar{a}_p^n) : p \in \mathcal{P}\}$ defined by

$$\bar{a}_p^i := \max\{\operatorname{Re}(a_p^i),\, 0\} \geq 0, \qquad i = 1, \dots, n.$$

Based on (23), the corresponding individual modes of (15) and (27) have the same entropy, that is, for each $p \in \mathcal{P}$,

$$h(\bar{D}_p) = \sum_{i=1}^n \bar{a}_p^i = \sum_{i=1}^n \max\{\operatorname{Re}(a_p^i),\, 0\} = h(D_p)$$

Meanwhile, the formula (16) implies

$$h(\bar{D}_\sigma) = \limsup_{T\to\infty} \sum_{i=1}^n \frac{1}{T} \max_{t\in[0,T]} \sum_{p\in\mathcal{P}} \bar{a}_p^i \tau_p(t) \geq \limsup_{T\to\infty} \sum_{i=1}^n \frac{1}{T} \max_{t\in[0,T]} \sum_{p\in\mathcal{P}} \operatorname{Re}(a_p^i)\tau_p(t) = h(D_\sigma + N_\sigma).$$

Moreover, for each $i \in \{1, \dots, n\}$, the sum $\sum_{p\in\mathcal{P}} \bar{a}_p^i \tau_p(t)$ is nondecreasing in t; thus

$$\frac{1}{T} \max_{t\in[0,T]} \sum_{p\in\mathcal{P}} \bar{a}_p^i \tau_p(t) = \frac{1}{T} \sum_{p\in\mathcal{P}} \bar{a}_p^i \tau_p(T) = \sum_{p\in\mathcal{P}} \bar{a}_p^i \rho_p(T).$$

Combining the results above, we obtain

$$h(D_\sigma + N_\sigma) \leq h(\bar{D}_\sigma) = \limsup_{T\to\infty} \sum_{i=1}^n \sum_{p\in\mathcal{P}} \bar{a}_p^i \rho_p(T) = \limsup_{T\to\infty} \sum_{p\in\mathcal{P}} h(D_p)\rho_p(T). \quad \square$$

The formula (23) for the entropy of LTI systems implies that, if all entries of all D_p in (15) have nonnegative real parts, that is, $\operatorname{Re}(a_p^i) \geq 0$ for all $i \in \{1, \dots, n\}$ and $p \in \mathcal{P}$, then $h(D_p) = \operatorname{tr}(D_p)$ for all $p \in \mathcal{P}$, and thus the upper bound (26) and the general lower bound (14) coincide.

Corollary 11. *Consider the switched linear system with pariwise-commuting matrices* (15). *If $\operatorname{Re}(a_p^i) \geq 0$ for all $i \in \{1, \dots, n\}$ and $p \in \mathcal{P}$, then the topological entropy of* (15) *satisfy*

$$h(D_\sigma + N_\sigma) = \limsup_{t\to\infty} \sum_{p\in\mathcal{P}} \operatorname{tr}(D_p)\rho_p(t).$$

Based on (26), we derive the following upper bounds for the entropy $h(D_\sigma + N_\sigma)$.

Corollary 12. *The topological entropy of the switched linear system with pairwise commuting matrices* (15) *is upper bounded by*

$$h(D_\sigma + N_\sigma) \leq \sum_{p\in\mathcal{P}} h(D_p)\hat{\rho}_p \tag{28}$$

with the asymptotic active rates $\hat{\rho}_p$ defined by (8)*, and also by*

$$h(D_\sigma + N_\sigma) \leq \max_{p\in\mathcal{P}} h(D_p), \tag{29}$$

where $h(D_p)$ denotes the topological entropy of the p-th mode, and satisfies (23) *with $A = D_p$ and thus $h(D_p) = h(D_p + N_p)$.*

Properties of the formula (16) and the upper bounds (24), (26), (28) and (29) are summarized in the following remark:

Remark 3. 1) Unlike the formula (16) and upper bound (24), the upper bounds (26), (28) and (29) are independent of the relative order of the scalar components between the matrices $D_p + N_p$ in (15), and can thus be calculated directly from the commuting family $\{A_p : p \in \mathcal{P}\}$, without the simultaneous change of basis in Proposition 5.

2) For a fixed family of matrices $\{D_p : p \in \mathcal{P}\}$, compared with the formula (16), the upper bounds (24) and (26) depend only on the active rates ρ_p, the upper bound (28) only on the asymptotic active rates $\hat{\rho}_p$; the upper bound (29) is independent of switching.
3) In general, the relation between the upper bounds (24) and (26), and that between the upper bounds (28) and (29), are both unknown. Meanwhile, both (24) and (26) imply

(28), whereas only (26) implies (29). These properties are demonstrated in the following example.

Example 2. Consider the index set $\mathcal{P} = \{1, 2\}$ and the switching signals σ_0, σ_1 and σ_2 defined as follows (see also footnote 1):

- SWITCHING SIGNAL σ_0 WITH NO SWITCH: Set $\sigma_0 \equiv 1$. Then the asymptotic active rates satisfy $\hat{\rho}_1 = 1$ and $\hat{\rho}_2 = 0$.
- SWITCHING SIGNAL σ_1 WITH PERIODIC SWITCHES: For $k \in N$, set $t_k = 2k$. Then $\hat{\rho}_1 = \hat{\rho}_2 = 0.5$.
- SWITCHING SIGNAL σ_2 WITH CONSTANT SET-POINTS: Set $t_1 := 1$. For $k \geq 1$, set $t_{2k} := \min\{t > t_{2k-1} : \rho_2(t) \geq 0.9\}$ and $t_{2k+1} := \min\{t > t_{2k} : \rho_1(t) \geq 0.9\}$. Then $t_k = 9^{k-1} + 9^{k-2}$ for $k \geq 2$, and $\hat{\rho}_1 = \hat{\rho}_2 = 0.9$.

As the entropy of the switched linear system with pairwise commuting matrices (15) is independent of the nilpotent part, consider the diagonal matrices

$$D_1 = \begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix}, \quad D_2 = \begin{bmatrix} 3 & 0 \\ 0 & -1 \end{bmatrix}.$$

The values of $h(D_{\sigma_0})$, $h(D_{\sigma_1})$ and $h(D_{\sigma_2})$ calculated using the formula (16), and their estimates using the upper bounds (24), (26), (28) and (29), as well as the general upper and lower bounds (13) and (14), are listed in Table I below. In particular, the calculation for the value and estimates of $h(D_{\sigma_2})$ can be found in Appendix C.

TABLE I

ENTROPY VALUES AND ESTIMATES IN EXAMPLE 2

	$(\hat{\rho}_1, \hat{\rho}_2)$	(16)	(24)	(26)	(28)	(29)	(13)	(14)
σ_0	$(1, 0)$	2	2	2	2	3	4	1
σ_1	$(0.5, 0.5)$	1.5	1.5	2.5	2.5	3	5	1.5
σ_2	$(0.9, 0.9)$	2.79	4.3	2.9	4.5	3	5.8	1.9

IV. CONCLUSION

In this paper, we studied the topological entropy of switched linear systems with pairwise commuting matrices. We devised a simultaneous change of basis under which each of the matrices can be decomposed into a diagonal part and a nilpotent part, and all the diagonal and nilpotent parts are pairwise commuting. A formula for the topological entropy is established in terms of the eigenvalues in each scalar component and the active time of each mode, and is thus independent of the nilpotent parts. We also demonstrated how the formula generalizes known results for the non-switched case and the case with simultaneously diagonalizable matrices, and constructed more general but more conservative upper bounds for the entropy.

An intriguing topic for future research is to reconcile the switching characterizations for entropy computation and for stability analysis and control design. For example, in stability and stabilization of switched systems, it is standard to impose slow-switching conditions such as the average dwell-time [29], which could also potentially be used in analyzing the entropy. Meanwhile, the entropy computation in this paper is based on the notion of active time (i.e., the accumulated time in which a mode is active). Such a quantity is rarely seen in the literature of switched control systems, and incorporating it into the control design procedure may lead to more precise data-rate bounds.

APPENDIX

A. Proof of Lemma 1

As a direct consequence of the definition (10), it holds that $\bar{a}(T) \geq \max\{\sum_{p\in\mathcal{P}} a_p\tau_p(T)/T,\, 0\}$ for all $T \geq 0$, and thus $\limsup_{T\to\infty} \bar{a}(T) \geq \max\{\hat{a}, 0\}$. Then it remains to prove that the reverse holds as well. The definition (9) of $\hat{a}$ implies that for each $\delta > 0$, there is a large enough $T'_\delta \geq 0$ such that

$$\sum_{p\in\mathcal{P}} a_p\rho_p(t) < \hat{a} + \delta \qquad \forall\, t > T'_\delta.$$

Consider an arbitrary $T > T'_\delta$, and let

$$t^*(T) := \operatorname*{arg\,max}_{t\in[0,T]} \sum_{p\in\mathcal{P}} a_p\tau_p(t).$$

If $t^*(T) \in (T'_\delta, T]$, then

$$\bar{a}(T) = \frac{1}{T}\sum_{p\in\mathcal{P}} a_p\tau_p(t^*(T)) \leq \sum_{p\in\mathcal{P}} a_p\rho_p(t^*(T)) < \hat{a} + \delta.$$

Otherwise $t^*(T) \in [0, T'_\delta]$, and thus

$$\bar{a}(T) = \frac{1}{T}\sum_{p\in\mathcal{P}} a_p\tau_p(t^*(T)) \leq \frac{a_m t^*(T)}{T} \leq \frac{|a_m|T'_\delta}{T}$$

with $a_m := \max_{p\in\mathcal{P}} a_p$. Combining the cases above yields $\bar{a}(T) \leq \max\{\hat{a}+\delta,\, |a_m|T'_\delta/T\}$ for all $T > T'_\delta$. Hence there is a large enough $T_\delta \geq 0$ (e.g., $T_\delta = \max\{T'_\delta,\, |a_m|T'_\delta/\delta\}$) such that $\bar{a}(T) \leq \max\{\hat{a},\, 0\} + \delta$ for all $T > T_\delta$. As $\delta > 0$ is arbitrary, it holds that $\limsup_{T\to\infty} \bar{a}(T) \leq \max\{\hat{a},\, 0\}$.

B. Proof of Lemma 2

First, we derive the upper bound in (17). For each $p \in \mathcal{P}$, as N_p is nilpotent, there is a positive integer k_p such that $N_p^{k_p} = 0$. Let $k_s := \sum_{p\in\mathcal{P}} k_p$. For an arbitrary $t \geq 0$, let

$$N(t) := \sum_{p\in\mathcal{P}} N_p\rho_p(t) \in \mathbb{C}^{n\times n}.$$

As $\{N_p : p \in \mathcal{P}\}$ is a commuting family and $\mathcal{P}$ is a finite set, it follows that k_s is a finite integer and $N(t)^{k_s} = 0$. Also, $\|N(t)\| \leq \max_{p\in\mathcal{P}} \|N_p\| =: \eta_m$. Hence for all $v \in \mathbb{C}^n$,

$$\begin{aligned}\left\|e^{N(t)t}v\right\| &= \left\|\left(\sum_{k=0}^{k_s-1}\frac{N(t)^k t^k}{k!}\right)v\right\| \\ &\leq \left(\sum_{k=0}^{k_s-1}\frac{\eta_m^k t^k}{k!}\right)\|v\| = \left(\sum_{k=0}^{k_s-1}\frac{\eta_m^k}{\delta^k}\cdot\frac{\delta^k t^k}{k!}\right)\|v\| \\ &\leq c_\delta\left(\sum_{k=0}^{k_s-1}\frac{\delta^k t^k}{k!}\right)\|v\| \leq c_\delta e^{\delta t}\|v\|\end{aligned}$$

with $c_\delta := \max\{(\eta_m/\delta)^{k_s-1},\, 1\} > 0$.

Second, we derive the lower bound in (17). For all $t \geq 0$, it holds that $\|-N(t)\| = \|N(t)\| \leq \eta_m$. Hence for all $v \in \mathbb{C}^n$, the proof above also implies

$$\|v\| = \left\|e^{-N(t)t}e^{N(t)t}v\right\| \leq c_\delta e^{\delta t}\left\|e^{N(t)t}v\right\|,$$

that is, $\|e^{N(t)t}v\| \geq c_\delta^{-1}e^{-\delta t}\|v\|$.

C. Calculations of the value and estimates of $h(D_{\sigma_2})$ in Example 2

Recall from footnote 1 that $\sigma = 1$ on $[t_{2k}, t_{2k+1})$ and $\sigma = 2$ on $[t_{2k+1}, t_{2k+2})$, where $t_0 = 0$, $t_1 = 1$, and $t_k = 9^{k-1} + 9^{k-2}$ for all $k \geq 2$. Hence $\tau_1(t) = t - 0.9t_{2k}$ and $\tau_2(t) = 0.9t_{2k}$ for $t \in [t_{2k}, t_{2k+1})$, and $\tau_1(t) = 0.9t_{2k+1}$ and $\tau_2(t) = t - 0.9t_{2k+1}$ for $t \in [t_{2k+1}, t_{2k+2})$, and thus

$$\begin{aligned}
a_1^1\tau_1(t) + a_2^1\tau_2(t) &= 3\tau_2(t) - \tau_1(t) \\
&= \begin{cases} 3.6t_{2k} - t, & t \in [t_{2k}, t_{2k+1}), \\ 3t - 3.6t_{2k+1}, & t \in [t_{2k+1}, t_{2k+2}) \end{cases} \\
a_1^2\tau_1(t) + a_2^2\tau_2(t) &= 2\tau_1(t) \\
&= \begin{cases} 2t - 2.7t_{2k}, & t \in [t_{2k}, t_{2k+1}), \\ 2.7t_{2k+1} - t, & t \in [t_{2k+1}, t_{2k+2}). \end{cases}
\end{aligned}$$

Then $\bar{a}_1$ and $\bar{a}_2$ in (16) satisfy

$$\begin{aligned}
\bar{a}_1(T) &= \frac{1}{T}\max_{t\in[0,T]} a_1^1\tau_1(t) + a_2^1\tau_2(t) \\
&= \begin{cases} 2.6t_{2k}/T, & T \in [t_{2k}, t_{2k+1} + 8t_{2k}/3), \\ 3 - 3.6t_{2k+1}/T, & T \in [t_{2k+1} + 8t_{2k}/3, t_{2k+2}) \end{cases} \\
\bar{a}_2(T) &= \frac{1}{T}\max_{t\in[0,T]} a_1^2\tau_1(t) + a_2^2\tau_2(t) \\
&= \begin{cases} 1.7t_{2k+1}/T, & T \in [t_{2k+1}, t_{2k+2} + 4t_{2k+1}), \\ 2 - 2.7t_{2k+2}/T, & T \in [t_{2k+2} + 4t_{2k+1}, t_{2k+3}). \end{cases}
\end{aligned}$$

Hence

$$\begin{aligned}
&\bar{a}_1(T) + \bar{a}_2(T) \\
&= \begin{cases} 17.9t_{2k}/T, & T \in [t_{2k+1}, t_{2k+1} + 8t_{2k}/3), \\ 3 - 1.9t_{2k+1}/T, & T \in [t_{2k+1} + 8t_{2k}/3, t_{2k+2}), \\ 25.1t_{2k+1}/T, & T \in [t_{2k+2}, t_{2k+2} + 4t_{2k+1}), \\ 2 - 0.1t_{2k+2}/T, & T \in [t_{2k+2} + 4t_{2k+1}, t_{2k+3}). \end{cases}
\end{aligned}$$

Therefore, $h(D_{\sigma_2}) = \limsup_{T\to\infty} \bar{a}_1(T) + \bar{a}_2(T) = \max\{1.99, 2.79\} = 2.79$.

For the upper bound (24),

$$\begin{aligned}
h(D_{\sigma_2}) &= \max\{\hat{a}_1, 0\} + \max\{\hat{a}_2, 0\} \\
&= -1 + (3 - (-1))\hat{\rho}_2 + (2 - (-1))\hat{\rho}_1 - 1 = 4.3.
\end{aligned}$$

The bounds (26), (13) and (14) are calculated similarly.

References

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

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

[3] R. L. Adler, A. G. Konheim, and M. H. McAndrew, "Topological entropy," *Transactions of the American Mathematical Society*, vol. 114, no. 2, pp. 309–319, 1965.

[4] R. Bowen, "Entropy for group endomorphisms and homogeneous spaces," *Transactions of the American Mathematical Society*, vol. 153, pp. 401–414, 1971.

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

[6] R. Bowen, "Periodic points and measures for axiom A diffeomorphisms," *Transactions of the American Mathematical Society*, vol. 154, pp. 377–397, 1971.

[7] S. Kolyada and L. Snoha, "Topological entropy of nonautonomous dynamical systems," *Random & Computational Dynamics*, 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," *Dynamical Systems*, 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 Transactions on Automatic Control*, vol. 49, no. 9, pp. 1585–1597, 2004.

[10] F. Colonius and C. Kawan, "Invariance entropy for control systems," *SIAM Journal on Control and Optimization*, 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," *Systems & Control Letters*, vol. 62, no. 5, pp. 377–381, 2013.

[12] F. Colonius, "Minimal bit rates and entropy for exponential stabilization," *SIAM Journal on Control and Optimization*, 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 Transactions on Automatic Control*, vol. 39, no. 12, pp. 2469–2471, 1994.

[15] J. P. Hespanha, A. Ortega, and L. Vasudevan, "Towards the control of linear systems with minimum bit-rate," in *15th International Symposium on Mathematical Theory of Networks and Systems*, 2002, pp. 1–15.

[16] G. N. Nair and R. J. Evans, "Exponential stabilisability of finite-dimensional linear systems with limited data rates," *Automatica*, vol. 39, no. 4, pp. 585–593, 2003.

[17] S. Tatikonda and S. Mitter, "Control under communication constraints," *IEEE Transactions on Automatic Control*, vol. 49, no. 7, pp. 1056–1068, 2004.

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

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

[20] H. Sibai and S. Mitra, "Optimal data rate for state estimation of switched nonlinear systems," in *20th International Conference on Hybrid Systems: Computation and Control*, 2017, pp. 71–80.

[21] D. Liberzon and S. Mitra, "Entropy and minimal bit rates for state estimation and model detection," *IEEE Transactions on Automatic Control*, to appear.

[22] G. Yang, A. J. Schmidt, and D. Liberzon, "On topological entropy of switched linear systems with diagonal, triangular, and general matrices," in *57th IEEE Conference on Decision and Control*, to appear.

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

[24] J. E. Humphreys, *Introduction to Lie Algebras and Representation Theory*. Springer New York, 1972, vol. 9.

[25] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*. Cambridge University Press, 1991.

[26] ——, *Matrix Analysis*. Cambridge University Press, 1985.

[27] A. V. Savkin, "Analysis and synthesis of networked control systems: topological entropy, observability, robustness and optimal control," *Automatica*, vol. 42, no. 1, pp. 51–62, 2006.

[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] J. P. Hespanha and A. S. Morse, "Stability of switched systems with average dwell-time," in *38th IEEE Conference on Decision and Control*, vol. 3, 1999, pp. 2655–2660.