

Topological entropy of switched linear systems

General matrices and matrices with commutation relations

Guosong Yang · A. James Schmidt · Daniel Liberzon ·
João P. Hespanha

the date of receipt and acceptance should be inserted later

Abstract This paper studies a notion of topological entropy for switched systems, formulated in terms of the minimal number of trajectories needed to approximate all trajectories with a finite precision. For general switched linear systems, we prove that the topological entropy is independent of the set of initial states, and construct an upper bound in terms of an average of the measures of system matrices of individual modes, weighted by their corresponding active times. We also construct a lower bound in terms of an active-time-weighted average of their traces. For switched linear systems with scalar-valued state and those with pairwise commuting matrices, we establish formulae for the topological entropy in terms of active-time-weighted averages of the eigenvalues of system matrices of individual modes. For the more general case with simultaneously triangularizable matrices, we construct an upper bound for the topological entropy that only depends on the eigenvalues, their order in a simultaneous triangularization, and the active times. In each case above, we also establish upper bounds that are more conservative but require less information on the system matrices or on the switching, with their relations illustrated by numerical examples. Stability conditions inspired by the upper bounds for the topological entropy are presented as well.

Keywords Topological entropy · Switched systems · Exponential stability

1 Introduction

Since its introduction for dynamical systems by Kolmogorov [20], entropy has been an invaluable tool for understanding system behavior. The Ornstein isomorphism theorem [32], which characterizes Bernoulli shifts entirely according to their entropy, further solidified its importance. Broadly, the entropy of a dynamical system captures the rate at which the uncertainty about the state grows as time evolves, which intuitively corresponds to entropy notions in other disciplines such as thermodynamics and information theory [9].

In systems theory, topological entropy describes the information accumulation required to approximate trajectories with 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 [20], and shares a striking resemblance to Shannon's information entropy [35]. Adler first defined topological entropy as an extension of Kolmogorov's metric entropy, quantifying the expansion of a map via the minimal cardinality of a subcover refinement [1]. An alternative definition in terms of the maximal number of trajectories separable with a finite precision was introduced by Bowen [4] and independently by Dinaburg [11]. An equivalence between these two definitions was established in [5]. Most existing results on topological entropy are for time-invariant systems, as time-varying dynamics introduce complexities that require new

G. Yang · J. P. Hespanha

Center for Control, Dynamical Systems, and Computation, University of California, Santa Barbara, CA 93106, USA
E-mail: {guosongyang, hespanha}@ucsb.edu

A. J. Schmidt · D. Liberzon

Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA
E-mail: {ajschmd2, liberzon}@illinois.edu

methods to understand [21, 19]. This work on the topological entropy of switched systems provides an initial investigation into some of these complexities.

Entropy has also played a prominent role in control theory, in which information flow appears between sensors and actuators for achieving control objectives. Nair et al. first introduced topological feedback entropy for discrete-time systems [31], following the construction in [1]. Their definition extended the classical entropy concepts, 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 [7], which is closer in spirit to the trajectory-counting formulation in [4, 11]. An equivalence between these two notions was established in [8]. The results of [7] were extended from set invariance to exponential stabilization in [6]. Entropy has also been studied in the dual problem of state estimation in, e.g., [33, 27].

This paper studies the topological entropy of switched systems. Switched system has become a popular topic in recent years (see, e.g., [25, 36] and references therein). In general, a switched system does not inherit the stability properties of individual modes. For example, a switched system with two stable modes may still be unstable [25, p. 19]. However, it is well known that a switched linear system generated by a finite family of pairwise commuting Hurwitz matrices is globally exponentially stable under arbitrary switching (see, e.g., [25, Th. 2.5, p. 31]). This result has been generalized to the cases where the Lie algebra generated by the system matrices is nilpotent [13], solvable [22, 28], or has a compact semisimple part [23, 2]. 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 switched linear systems with such matrices. See [3] and [14] for related results on robustness to perturbations and on state feedbacks that induce simultaneous triangularizability, respectively.

Our interest in the topological entropy of switched systems is strongly motivated by its relation to the data-rate requirements in control problems. For a linear time-invariant control system, it has been shown that the minimal data rate for feedback stabilization equals the topological entropy in open-loop [15, 30, 38]. For switched systems, however, neither the minimal data rate nor the topological entropy are well understood. Sufficient data rates for feedback stabilization of switched linear systems were established in [26, 40]. Similar data-rate conditions were constructed in [37] by extending the estimation entropy from [27] to switched systems. In this work, we seek to contribute to these efforts.

The main contribution of this paper is the construction of formulae and bounds for the topological entropy of switched linear systems. In Section 2, we introduce a notion of topological entropy for switched systems, and define switching-related quantities such as the active time of each individual mode, which prove useful in calculating the topological entropy. In Section 3, after proving that the topological entropy of a switched linear system is independent of the set of initial states, we construct a general upper bound for the topological entropy in terms of an active-time-weighted average of the measures of system matrices of individual modes, and a general lower bound in terms of an active-time-weighted average of their traces.

Sections 4–6 provide formulae and improved upper bounds for the topological entropy of switched linear systems generated by matrices with commutation relations. In Section 4, we consider the case with scalar-valued state, in which the general upper and lower bounds from Section 3 coincide and become a formula for the topological entropy. Section 5 studies the case with pairwise commuting matrices, by establishing a formula for the topological entropy in terms of component-wise active-time-weighted averages of the eigenvalues of system matrices of individual modes. In Section 6, we investigate the more general case with simultaneously triangularizable matrices, and construct an upper bound for the topological entropy that only depends on the eigenvalues, their order in a simultaneous triangularization, and the corresponding active times. The upper bound is obtained by first establishing a formula for the solution to a switched triangular system and an upper bound for its norm, which are also of independent interest. In the last two cases, the overall topological entropy is related to the topological entropy in each individual scalar component and to that of each individual mode. In each case above, we also establish upper bounds that are more conservative but require less information on the system matrices or on the switching, with their relations illustrated by numerical examples. Stability conditions inspired by the upper bounds for topological entropy are presented in Section 7. Section 8 concludes the paper with a brief summary and remarks on future research directions.

Notations: Let $\mathbb{R}_{\geq 0} := [0, \infty)$, $\mathbb{R}_{> 0} := (0, \infty)$, and $\mathbb{N} := \{0, 1, \dots\}$. Denote by I_n the identity matrix in $\mathbb{R}^{n \times n}$; the subscript is omitted when the dimension is implicit. For a complex number $a \in \mathbb{C}$, denote by $\text{Re}(a)$ its real part. For a vector $v \in \mathbb{C}^n$, denote by v_i its i -th scalar component and write $v = (v_1, \dots, v_n)$.

For a matrix $A \in \mathbb{C}^{n \times n}$, denote by $\text{spec}(A)$, $\text{tr}(A)$, and $\det(A)$ its spectrum, trace, and determinant, respectively. For a set $E \subset \mathbb{C}^n$, denote by $|E|$ and $\text{vol}(E)$ its cardinality and volume (Lebesgue measure), respectively. Denote by $\|v\|_\infty := \max_{1 \leq i \leq n} |v_i|$ the ∞ -norm of a vector $v \in \mathbb{C}^n$, and by $\|A\|_\infty := \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|$ the induced ∞ -norm of a matrix $A = [a_{ij}] \in \mathbb{C}^{n \times n}$. By default, all logarithms are natural logarithms.¹

2 Preliminaries

2.1 Entropy definitions

Consider a family of continuous-time dynamical systems

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

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

$$\dot{x} = f_\sigma(x), \quad (2)$$

where $\sigma : \mathbb{R}_{\geq 0} \rightarrow \mathcal{P}$ is a right-continuous, piecewise constant *switching signal*. We call the system with index p in (1) the p -th *mode* of the switched system (2), and $\sigma(t)$ the *active mode* at time t . Denote by $\xi_\sigma(x, t)$ the solution to (2) with initial state x at time t . For a fixed 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). Denote by $N_\sigma(t, \tau)$ the number of switches on an interval $(\tau, t]$.

In the following, we define a notion of topological entropy for the switched system (2). Denote by $\|\cdot\|$ some chosen norm on \mathbb{R}^n and the corresponding induced norm on $\mathbb{R}^{n \times n}$. Consider a switching signal σ and initial states drawn from a compact set with nonempty interior $K \subset \mathbb{R}^n$ called the *initial set*. Given a time horizon $T \geq 0$ and a radius $\varepsilon > 0$, we define the following open ball in K with center x :

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

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

$$K = \bigcup_{\hat{x} \in E} B_{f_\sigma}(\hat{x}, \varepsilon, T), \quad (4)$$

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, or equivalently, the cardinality of a minimal (T, ε) -spanning set, which is nondecreasing in T and nonincreasing in ε . 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

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

For brevity, we at times refer to $h(f_\sigma, K)$ simply as the (topological) entropy of the switched system (2).

Remark 1. In view of the equivalence of norms on a finite-dimensional vector space, the values of $h(f_\sigma, K)$ are the same for all norms $\|\cdot\|$ on \mathbb{R}^n ; see [18, Prop. 3.1.2, p. 109] for a slightly stronger statement for the case with a compact invariant set. In particular, the topological entropy is invariant under a change of basis. For convenience and concreteness, we take $\|\cdot\|$ to be the ∞ -norm on \mathbb{R}^n or the induced ∞ -norm on $\mathbb{R}^{n \times n}$ unless otherwise specified.

¹ In information theory, entropy notions are often formulated using binary logarithms due to their connection with binary signals. In this paper, we use natural logarithms to avoid generating additional multiplicative constants $\ln 2$ when calculating the topological entropy.

Next, we introduce an equivalent definition for the entropy of the switched system (2). With T and ε given as before, we say that a finite set $E \subset K$ is (T, ε) -separated if

$$\hat{x}' \notin B_{f_\sigma}(\hat{x}, \varepsilon, T) \quad (6)$$

for each pair of distinct points $\hat{x}, \hat{x}' \in E$, or equivalently, 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 and nonincreasing in ε . As stated in the following result, the entropy of (2) can be equivalently formulated in terms of the exponential growth rate of $N(f_\sigma, \varepsilon, T, K)$; the proof is along the lines of [18, p. 110] and thus omitted here.

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

$$h(f_\sigma, K) = \lim_{\varepsilon \searrow 0} \limsup_{T \rightarrow \infty} \frac{1}{T} \log N(f_\sigma, \varepsilon, T, K). \quad (7)$$

Remark 2. Following [18, 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 (5) and (7) are replaced with limit infima. However, this is not necessarily the case for a time-varying system such as the switched system (2), for which the subadditivity required in the proof of [18, Lemma 3.1.5, p. 109] does not necessarily hold.

2.2 Active times, active rates, and weighted averages

In this subsection, we introduce switching-related quantities that will be useful in calculating the entropy of a switched linear system. The *active time* of the p -th mode over an interval $[0, t]$ is defined by

$$\tau_p(t) := \int_0^t \mathbb{1}_p(\sigma(s)) \, ds, \quad p \in \mathcal{P} \quad (8)$$

with the indicator function

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

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

$$\rho_p(t) := \tau_p(t)/t, \quad p \in \mathcal{P} \quad (9)$$

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

$$\hat{\rho}_p := \limsup_{t \rightarrow \infty} \rho_p(t), \quad p \in \mathcal{P}. \quad (10)$$

Clearly, the active times τ_p are nonnegative and nondecreasing, and satisfy $\sum_{p \in \mathcal{P}} \tau_p(t) = t$ for all $t \geq 0$; the active rates ρ_p take values 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 (10), it is possible that $\sum_{p \in \mathcal{P}} \hat{\rho}_p > 1$ for the asymptotic active rates $\hat{\rho}_p$, as illustrated in the following example.

Example 1. Consider the index set $\mathcal{P} = \{1, 2\}$. We construct a switching signal σ_* as follows²:

- **CONVERGING SET-POINTS:** Let $t_1 := 1$. For $k \geq 1$, let $t_{2k} := \min\{t > t_{2k-1} : \rho_2(t) \geq 1 - 2^{-2k}\}$ 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 $\hat{\rho}_1 = \hat{\rho}_2 = \limsup_{k \rightarrow \infty} 1 - e^{-2^k} = 1$.

The switching signal σ_* (purple) and the active rates ρ_1 (blue) and ρ_2 (orange) are plotted in Fig. 1 below, with the asymptotic active rates $\hat{\rho}_1$ and $\hat{\rho}_2$ indicated by the green dashed line (as the intervals between consecutive switches grow superexponentially, logarithmic scale is used for the long-range plot).

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

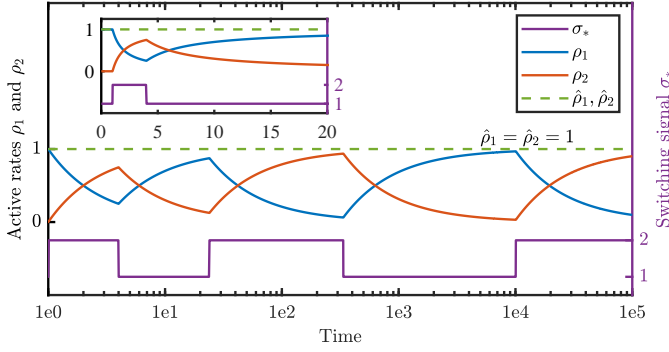


Fig. 1 A switching signal σ_* with converging set-points: the sum of the active rates $\rho_1 + \rho_2 = 1$ at all times, whereas both asymptotic active rates $\hat{\rho}_1 = \hat{\rho}_2 = 1$.

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

$$\hat{a} := \limsup_{t \rightarrow \infty} \sum_{p \in \mathcal{P}} a_p \rho_p(t) = \limsup_{t \rightarrow \infty} \frac{1}{t} \sum_{p \in \mathcal{P}} a_p \tau_p(t), \quad (11)$$

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

$$\bar{a}(T) := \frac{1}{T} \max_{t \in [0, T]} \sum_{p \in \mathcal{P}} a_p \tau_p(t). \quad (12)$$

As $\tau_p(0) = 0$ for all $p \in \mathcal{P}$, the maximal weighted average \bar{a} is nonnegative. In the following lemma, we establish a relation between these two notions; the proof can be found in Appendix A.

Lemma 1. *The asymptotic weighted average \hat{a} defined by (11) and the maximal weighted average \bar{a} defined by (12) satisfy*

$$\limsup_{T \rightarrow \infty} \bar{a}(T) = \max\{\hat{a}, 0\}. \quad (13)$$

3 Entropy of general switched linear systems

The main objective of this paper is to study the topological entropy of the switched linear system

$$\dot{x} = A_{\sigma} x \quad (14)$$

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 (14) with initial set K and switching signal σ . In this section, we first prove that the entropy of a switched linear system is independent of the choice of the initial set, and provide standard constructions of spanning and separated sets based on a notion of grid. Second, we present a result for the non-switched case. Finally, we construct upper and lower bounds for the entropy of a general switched linear system.

3.1 Initial set and grid

Proposition 2. *The topological entropy of the switched linear system (14) is independent of the choice of the initial set K .*

Proof. For every initial state $x \in \mathbb{R}^n$, the solution to (14) satisfies

$$\xi_{\sigma}(x, t) = \Phi_{\sigma}(t, 0) x \quad \forall t \geq 0,$$

where the state-transition matrix $\Phi_{\sigma}(t, 0)$ is independent of the initial state x .

First, we prove that the entropy of (14) is invariant under translation and uniform scaling of the initial set. Let $K_1 \subset \mathbb{R}^n$ be an initial set, and define the translated and uniformly scaled set $K_2 := \{sx + v : x \in K_1\}$ for some scalar $s > 0$ and vector $v \in \mathbb{R}^n$. Given a time horizon $T \geq 0$ and a radius $\varepsilon > 0$, let E_1 be a minimal (T, ε) -spanning set of K_1 . For each $x_2 \in K_2$, the point $x_1 := (x_2 - v)/s \in K_1$; thus there is a point $\hat{x}_1 \in E_1$ such that $\|\xi_\sigma(x_1, t) - \xi_\sigma(\hat{x}_1, t)\| = \|\Phi_\sigma(t, 0)(x_1 - \hat{x}_1)\| < \varepsilon$ for all $t \in [0, T]$. Then the point $\hat{x}_2 := s\hat{x}_1 + v$ satisfies $\|\xi_\sigma(x_2, t) - \xi_\sigma(\hat{x}_2, t)\| = \|\Phi_\sigma(t, 0)(x_2 - \hat{x}_2)\| = s\|\Phi_\sigma(t, 0)(x_1 - \hat{x}_1)\| < s\varepsilon$ for all $t \in [0, T]$. Therefore, the set $E_2 := \{s\hat{x} + v : \hat{x} \in E_1\}$ is a $(T, s\varepsilon)$ -spanning set of K_2 . As $|E_2| = |E_1|$, we have $S(A_\sigma, s\varepsilon, T, K_2) \leq S(A_\sigma, \varepsilon, T, K_1)$ and thus $h(A_\sigma, K_2) \leq h(A_\sigma, K_1)$. Replacing s and ε with $1/s$ and $s\varepsilon$ in the arguments above, we obtain $S(A_\sigma, s\varepsilon, T, K_2) \geq S(A_\sigma, \varepsilon, T, K_1)$ and thus $h(A_\sigma, K_2) \geq h(A_\sigma, K_1)$. Hence $h(A_\sigma, K_2) = h(A_\sigma, K_1)$. Therefore, the entropy of (14) is invariant under translation and uniform scaling of the initial set.

Second, we establish that the entropy of (14) is independent of the choice of the initial set. Let $K \subset \mathbb{R}^n$ be an initial set. As K is a compact set with nonempty interior, there exist closed balls $B_1, B_2 \subset \mathbb{R}^n$ such that $B_1 \subset K \subset B_2$; thus $h(A_\sigma, B_1) \leq h(A_\sigma, K) \leq h(A_\sigma, B_2)$ by construction. As the entropy of (14) is invariant under translation and uniform scaling of the initial set, we have $h(A_\sigma, B_1) = h(A_\sigma, B_2)$. Hence $h(A_\sigma, B_1) = h(A_\sigma, K) = h(A_\sigma, B_2)$. Therefore, the entropy of (14) is independent of the choice of K . \square

Following Proposition 2, we omit the initial set and denote by $h(A_\sigma)$ the entropy of the switched linear system (14). For convenience and concreteness, we take the initial set to be the closed unit hypercube at the origin, that is, $K := \{x \in \mathbb{R}^n : \|x\| \leq 1\}$ (recall that we take $\|\cdot\|$ to be the ∞ -norm; see Remark 1) in calculating the entropy of (14).

Next, given a time horizon $T \geq 0$ and a radius $\varepsilon > 0$, we construct standard (T, ε) -spanning and (T, ε) -separated sets based on a notion of grid. Given a vector $\theta = (\theta_1, \dots, \theta_n) \in \mathbb{R}_{>0}^n$, we define the following *grid* on the closed unit hypercube K at the origin:

$$G(\theta) := \{(k_1\theta_1, \dots, k_n\theta_n) \in K : k_1, \dots, k_n \in \mathbb{Z}\}. \quad (15)$$

Simple calculation yields that the cardinality of the grid $G(\theta)$ satisfies

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

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

$$R(\hat{x}) := \{x \in K : |x_1 - \hat{x}_1| < \theta_1, \dots, |x_n - \hat{x}_n| < \theta_n\}. \quad (16)$$

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 , that is,

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

Comparing the hyperrectangle $R(\hat{x})$ to the open ball $B_{A_\sigma}(\hat{x}, \varepsilon, T)$ defined by (3), we obtain the following result; the proof can be found in Appendix B.

Lemma 2. *Consider the switched linear system (14).*

1. *If the vector θ is selected so that $R(\hat{x}) \subset B_{A_\sigma}(\hat{x}, \varepsilon, T)$ for all $\hat{x} \in G(\theta)$, then the grid $G(\theta)$ is (T, ε) -spanning; thus the minimal cardinality of a (T, ε) -spanning set satisfies*

$$\log S(A_\sigma, \varepsilon, T, K) \leq \sum_{i=1}^n \log(2/\theta_i + 1). \quad (17)$$

2. *If the vector θ is selected so that $B_{A_\sigma}(\hat{x}, \varepsilon, T) \subset R(\hat{x})$ for all $\hat{x} \in G(\theta)$, then the grid $G(\theta)$ is (T, ε) -separated; thus the maximal cardinality of a (T, ε) -separated set satisfies*

$$\log N(A_\sigma, \varepsilon, T, K) \geq \sum_{i=1}^n \log(2/\theta_i - 1). \quad (18)$$

Compared with the inequality (18), the inequality (17) seems more conservative due to the substantial overlaps between the hyperrectangles $R(\hat{x})$. However, as will be clear from the proofs of formulae and upper bounds for the entropy of switched linear systems below, such conservativeness does not affect the exponential growth rate of $S(A_\sigma, \varepsilon, T, K)$, and thus does not affect the entropy $h(A_\sigma)$.

3.2 Entropy of linear time-invariant systems

Before analyzing the entropy of the switched linear system (14), we present here a result for the non-switched case. Consider a linear time-invariant (LTI) system

$$\dot{x} = Ax \quad (19)$$

with a matrix $A \in \mathbb{R}^{n \times n}$. The following well-known result provides a formula for the entropy $h(A)$ of (19). The proof is along the lines of those of the corresponding discrete-time results (e.g., [4, Th. 15] and [33, Th. 4.1]) and thus omitted here; a complete proof can be found in [34, Ch. 4].

Lemma 3. *The topological entropy of the LTI system (19) satisfies*

$$h(A) = \sum_{\lambda \in \text{spec}(A)} \max\{\text{Re}(\lambda), 0\}. \quad (20)$$

3.3 Entropy of general switched linear systems

In this subsection, we construct upper and lower bounds for the entropy of the general switched linear system (14). The upper bound is formulated in terms of a notion of matrix measure for the system matrices of individual modes, and the lower bound is formulated in terms of their traces.

Following [10, p. 30], for an induced matrix norm $\|\cdot\|$, the *matrix measure* $\mu : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$ is defined by

$$\mu(A) := \lim_{t \searrow 0} \frac{\|I + tA\| - 1}{t}. \quad (21)$$

For standard norms, there are explicit formulae for the matrix measure; for example, for the ∞ -norm, the matrix measure satisfies

$$\mu(A) = \max_{1 \leq i \leq n} \left(a_{ii} + \sum_{j \neq i} |a_{ij}| \right)$$

for a matrix $A = [a_{ij}] \in \mathbb{R}^{n \times n}$. For all induced matrix norms $\|\cdot\|$ on \mathbb{R}^n , we have [10, Th. 5, p. 31]

$$\text{Re}(\lambda) \leq \mu(A) \leq \|A\| \quad \forall A \in \mathbb{R}^{n \times n}, \forall \lambda \in \text{spec}(A). \quad (22)$$

Moreover, an upper bound for the norm of the solution to the switched linear system (14) can be constructed in terms of the matrix measures of A_p , which is a direct consequence of [10, Th. 27, p. 34].³

Lemma 4. *For every initial state $x \in \mathbb{R}^n$, the solution to the switched linear system (14) satisfies*

$$\|\xi_\sigma(x, t)\| \leq e^{\sum_{p \in \mathcal{P}} \mu(A_p) \tau_p(t)} \|x\| \quad \forall t \geq 0$$

with the active times τ_p defined by (8).

Theorem 3. *The topological entropy of the switched linear system (14) is upper bounded by⁴*

$$h(A_\sigma) \leq \max \left\{ \limsup_{t \rightarrow \infty} \sum_{p \in \mathcal{P}} n \mu(A_p) \rho_p(t), 0 \right\} \quad (23)$$

and lower bounded by

$$h(A_\sigma) \geq \max \left\{ \limsup_{t \rightarrow \infty} \sum_{p \in \mathcal{P}} \text{tr}(A_p) \rho_p(t), 0 \right\} \quad (24)$$

with the active rates ρ_p defined by (9).

³ We can apply [10, Th. 27, p. 34] to the switched linear system (14) as the switching signal σ is piecewise constant.

⁴ Following (22), the upper bound (23) is tighter than the one in the previous result [41, eq. (19)].

Proof. First, we establish the upper bound (23). For all initial states $x, x' \in K$, the corresponding solutions to (14) satisfy

$$\|\xi_\sigma(x', t) - \xi_\sigma(x, t)\| = \|\xi_\sigma(x' - x, t)\| \leq e^{\sum_{p \in \mathcal{P}} \mu(A_p) \tau_p(t)} \|x' - x\| \quad \forall t \geq 0,$$

where the inequality follows from Lemma 4. Given a time horizon $T \geq 0$ and a radius $\varepsilon > 0$, we have

$$\max_{t \in [0, T]} \|\xi_\sigma(x', t) - \xi_\sigma(x, t)\| \leq e^{\max_{t \in [0, T]} \sum_{p \in \mathcal{P}} \mu(A_p) \tau_p(t)} \|x' - x\|. \quad (25)$$

Consider the grid $G(\theta)$ defined by (15) with

$$\theta_i := e^{-\max_{t \in [0, T]} \sum_{p \in \mathcal{P}} \mu(A_p) \tau_p(t)} \varepsilon, \quad i \in \{1, \dots, n\},$$

and the corresponding hypercubes $R(\hat{x})$ defined by (16). Comparing (16) and (25) to (3), we see that $R(\hat{x}) \subset B_{A_\sigma}(\hat{x}, \varepsilon, T)$ for all $\hat{x} \in G(\theta)$. Then Lemma 2 implies that $G(\theta)$ is (T, ε) -spanning, and substituting (17) into (5) yields

$$\begin{aligned} h(A_\sigma) &\leq \lim_{\varepsilon \searrow 0} \limsup_{T \rightarrow \infty} \sum_{i=1}^n \frac{\log(2/\theta_i + 1)}{T} \\ &= \lim_{\varepsilon \searrow 0} \limsup_{T \rightarrow \infty} \sum_{i=1}^n \frac{\log(1/\theta_i)}{T} + \lim_{\varepsilon \searrow 0} \limsup_{T \rightarrow \infty} \sum_{i=1}^n \frac{\log(2 + \theta_i)}{T} \\ &= \lim_{\varepsilon \searrow 0} \limsup_{T \rightarrow \infty} \sum_{i=1}^n \frac{\log(1/\theta_i)}{T} \\ &= \limsup_{T \rightarrow \infty} \frac{1}{T} \max_{t \in [0, T]} \sum_{p \in \mathcal{P}} n \mu(A_p) \tau_p(t) + \lim_{\varepsilon \searrow 0} \limsup_{T \rightarrow \infty} \frac{n \log(1/\varepsilon)}{T} \\ &= \limsup_{T \rightarrow \infty} \frac{1}{T} \max_{t \in [0, T]} \sum_{p \in \mathcal{P}} n \mu(A_p) \tau_p(t). \end{aligned}$$

Then the upper bound (23) follows from (13) with $a_p = n \mu(A_p)$ in (11) and (12).

Second, we establish the lower bound (24) using volume-based arguments. For every initial state $x \in K$, the solution to (14) satisfies

$$\xi_\sigma(x, t) = \Phi_\sigma(t, 0) x \quad \forall t \geq 0$$

with the state-transition matrix defined by

$$\Phi_\sigma(t, s) := e^{A_\sigma(t_{N_\sigma(t, s)})(t_{N_\sigma(t, s)+1} - t_{N_\sigma(t, s)})} \dots e^{A_\sigma(t_0)(t_1 - t_0)}, \quad t \geq s \geq 0,$$

where $t_1 < \dots < t_{N_\sigma(t, s)}$ is the sequence of switches on $(s, t]$, and $t_0 := s$ and $t_{N_\sigma(t, s)+1} := t$. Given a time horizon $T \geq 0$ and a radius $\varepsilon > 0$, the open ball $B_{A_\sigma}(x, \varepsilon, T)$ defined by (3) satisfies

$$\begin{aligned} B_{A_\sigma}(x, \varepsilon, T) &\subset \{x' \in K : \|\xi_\sigma(x', T) - \xi_\sigma(x, T)\| < \varepsilon\} \\ &= \{x' \in K : \|\Phi_\sigma(T, 0)(x' - x)\| < \varepsilon\} \\ &= \{\Phi_\sigma(T, 0)^{-1}v + x \in K : \|v\| < \varepsilon\}. \end{aligned}$$

Hence its volume satisfies

$$\text{vol}(B_{A_\sigma}(x, \varepsilon, T)) \leq \det(\Phi_\sigma(T, 0)^{-1}) (2\varepsilon)^n = \left(\prod_{i=0}^{N_\sigma(T, 0)} \det(e^{-A_\sigma(t_i)(t_{i+1} - t_i)}) \right) (2\varepsilon)^n.$$

Applying Liouville's trace formula, we obtain

$$\prod_{i=0}^{N_\sigma(T, 0)} \det(e^{-A_\sigma(t_i)(t_{i+1} - t_i)}) = e^{-\sum_{i=0}^{N_\sigma(T, 0)} \text{tr}(A_\sigma(t_i))(t_{i+1} - t_i)} = e^{-\sum_{p \in \mathcal{P}} \text{tr}(A_p) \tau_p(T)}$$

with the active times τ_p defined by (8); thus

$$\text{vol}(B_{A_\sigma}(x, \varepsilon, T)) \leq e^{-\sum_{p \in \mathcal{P}} \text{tr}(A_p) \tau_p(T)} (2\varepsilon)^n.$$

Combining the upper bound above with (4), we conclude that for all (T, ε) -spanning sets $E \subset K$,

$$\text{vol}(K) \leq \sum_{\hat{x} \in E} \text{vol}(B_{A_\sigma}(\hat{x}, \varepsilon, T)) \leq |E| e^{-\sum_{p \in \mathcal{P}} \text{tr}(A_p) \tau_p(T)} (2\varepsilon)^n.$$

Therefore, the minimal cardinality of a (T, ε) -spanning set satisfies

$$S(A_\sigma, \varepsilon, T) \geq |E| \geq e^{\sum_{p \in \mathcal{P}} \text{tr}(A_p) \tau_p(T)} \text{vol}(K) / (2\varepsilon)^n,$$

which, combined with (5), implies

$$\begin{aligned} h(A_\sigma) &\geq \lim_{\varepsilon \searrow 0} \limsup_{T \rightarrow \infty} \frac{1}{T} \log \left(e^{\sum_{p \in \mathcal{P}} \text{tr}(A_p) \tau_p(T)} \text{vol}(K) / (2\varepsilon)^n \right) \\ &= \limsup_{T \rightarrow \infty} \sum_{p \in \mathcal{P}} \frac{\text{tr}(A_p) \tau_p(T)}{T} + \lim_{\varepsilon \searrow 0} \limsup_{T \rightarrow \infty} \frac{\log(\text{vol}(K) / (2\varepsilon)^n)}{T} \\ &= \limsup_{T \rightarrow \infty} \sum_{p \in \mathcal{P}} \text{tr}(A_p) \rho_p(T), \end{aligned}$$

where the last equality follows from the definition (9) of the active rates ρ_p . The proof of (24) is completed by recalling that $h(A_\sigma) \geq 0$. \square

Note that the value of the matrix measure depends on the induced norm $\|\cdot\|$ in the definition (21). Therefore, the upper bound (23) can be improved by taking the infimum over all induced norms on $\mathbb{R}^{n \times n}$. See Remark 5 below for additional upper bounds that are more conservative but require less information on the switching.

In general, there is a gap between the upper bound (23) and lower bound (24) in Theorem 3 (e.g., consider an LTI system with a matrix with one positive and one negative eigenvalue). The formula (20) for the entropy of an LTI system, together with the property (22), implies that $\max\{\text{tr}(A_p), 0\} \leq h(A_p) \leq \max\{n\mu(A_p), 0\}$ for all $p \in \mathcal{P}$, and thus

$$\max \left\{ \limsup_{t \rightarrow \infty} \sum_{p \in \mathcal{P}} \text{tr}(A_p) \rho_p(t), 0 \right\} \leq \limsup_{t \rightarrow \infty} \sum_{p \in \mathcal{P}} h(A_p) \rho_p(t) \leq \max \left\{ \limsup_{t \rightarrow \infty} \sum_{p \in \mathcal{P}} n\mu(A_p) \rho_p(t), 0 \right\}.$$

However, for a general switched linear system, due to the lack of “alignment” between eigenspaces of different individual modes, the relation between $h(A_\sigma)$ and $\limsup_{t \rightarrow \infty} \sum_{p \in \mathcal{P}} h(A_p) \rho_p(t)$ is undetermined (examples where the former is smaller can be found in Example 3 below; an example where the former is larger can be seen from the unstable switched linear system generated by Hurwitz matrices in [25, p. 26]). In Sections 4–6, we will consider switched linear systems generated by matrices with commutation relations, and establish formulae and improved upper bounds for the topological entropy.

4 Entropy of switched scalar systems

In this section, we consider the case of switched linear systems with scalar-valued state. Then each A_p becomes a scalar $a_p \in \mathbb{R}$, and (14) becomes the *switched scalar system*

$$\dot{x} = a_\sigma x \tag{26}$$

with the family of scalars $\{a_p : p \in \mathcal{P}\}$. In this case, $\text{tr}(a_p) = a_p = \mu(a_p)$ for all $p \in \mathcal{P}$, and hence the upper bound (23) and lower bound (24) in Theorem 3 coincide and become the following formula for the entropy $h(a_\sigma)$.

Corollary 4. *The topological entropy of the switched scalar system (26) satisfies*

$$h(a_\sigma) = \max\{\hat{a}, 0\} \quad (27)$$

with the asymptotic weighted average \hat{a} defined by (11).

Based on the formula (27), we construct upper bounds for the entropy $h(a_\sigma)$ that require less information; the proof can be found in Appendix C.

Corollary 5. *The topological entropy of the switched scalar system (26) is upper bounded by*

$$h(a_\sigma) \leq \sum_{p \in \mathcal{P}} h(a_p) \hat{\rho}_p \quad (28)$$

with the asymptotic active rates $\hat{\rho}_p$ defined by (10), and also by

$$h(a_\sigma) \leq \max_{p \in \mathcal{P}} h(a_p), \quad (29)$$

where $h(a_p)$ denotes the topological entropy of the p -th mode and satisfies (20), that is, $h(a_p) = \max\{a_p, 0\}$. Moreover, if the limits $\lim_{t \rightarrow \infty} \rho_p(t)$ exist and $a_p \geq 0$ for all $p \in \mathcal{P}$, then (28) holds with equality.

- Remark 3.* 1. For a fixed family of scalars $\{a_p : p \in \mathcal{P}\}$, compared with the formula (27), the upper bound (28) depends only on the asymptotic active rates $\hat{\rho}_p$; the upper bound (29) is independent of switching.
2. The upper bounds (28) and (29) are both useful in the sense that neither is more conservative than the other, as illustrated in the following example.

Example 2. Consider the index set $\mathcal{P} = \{1, 2\}$ and the scalars $a_1 = 2$ and $a_2 = 1$. We construct switching signals σ_0 , σ_1 , and σ_2 as follows (see also footnote 2):

- NO SWITCH: Let $\sigma_0(t) := 1$ for all $t \geq 0$. Then the asymptotic active rates satisfy $\hat{\rho}_1 = 1$ and $\hat{\rho}_2 = 0$.
- PERIODIC SWITCHES: For $k \in \mathbb{N}$, let $t_k := 2k$. Then $\hat{\rho}_1 = \hat{\rho}_2 = 0.5$.
- CONSTANT SET-POINTS: Let $t_1 := 1$. For $k \geq 1$, let $t_{2k} := \min\{t > t_{2k-1} : \rho_2(t) \geq 0.9\}$ and $t_{2k+1} := \min\{t > t_{2k} : \rho_1(t) \geq 0.9\}$. Simple calculation yields $t_k = 9^{k-1} + 9^{k-2}$ for $k \geq 2$ and $\hat{\rho}_1 = \hat{\rho}_2 = 0.9$.

The values of $h(a_{\sigma_0})$, $h(a_{\sigma_1})$, and $h(a_{\sigma_2})$ calculated using the formula (27) and their upper bounds calculated using (28) and (29) are summarized in Table 1. In particular, $h(a_{\sigma_2})$ can be calculated as follows:

$$h(a_{\sigma_2}) = \limsup_{t \rightarrow \infty} (a_1 \rho_1(t) + a_2 (1 - \rho_1(t))) = a_2 + (a_1 - a_2) \hat{\rho}_1 = 1.9.$$

Table 1 Entropy values and bounds for the switched scalar systems in Example 2.

	$(\hat{\rho}_1, \hat{\rho}_2)$	(27)	(28)	(29)
σ_0	(1, 0)	2	2	2
σ_1	(0.5, 0.5)	1.5	1.5	2
σ_2	(0.9, 0.9)	1.9	2.7	2

Remark 4. 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 each a_p with its real part $\text{Re}(a_p)$ in (11) and (12) and noticing that (20) implies $h(a_p) = \max\{\text{Re}(a_p), 0\}$. More specifically, this can be seen from the fact that for all initial states $x, x' \in K$, the corresponding solutions to (26) satisfy $|\xi_\sigma(x', t) - \xi_\sigma(x, t)| = e^{\sum_{p \in \mathcal{P}} \text{Re}(a_p) \tau_p(t)} |x' - x|$ for all $t \geq 0$.

Remark 5. Comparing (23) and (24) to (27), we conclude that the entropy of the general switched linear system (14) is upper and lower bounded by those of the switched scalar systems (26) with $a_p = n\mu(A_p)$ and with $a_p = \text{tr}(A_p)$, respectively. Consequently, Corollary 5 implies that the entropy of (14) is upper bounded by

$$h(A_\sigma) \leq \sum_{p \in \mathcal{P}} \max\{n\mu(A_p) \hat{\rho}_p, 0\},$$

which only depends on the asymptotic active rates $\hat{\rho}_p$ defined by (10), and also by

$$h(A_\sigma) \leq \max \left\{ \max_{p \in \mathcal{P}} n\mu(A_p), 0 \right\},$$

which is independent of switching.

5 Entropy of switched commuting systems

In this section, we consider the case of switched linear systems with pairwise commuting matrices, that is, the family of matrices $\{A_p : p \in \mathcal{P}\}$ in (14) satisfies

$$A_p A_q = A_q A_p \quad \forall p, q \in \mathcal{P}.$$

We call such a set of pairwise commuting matrices a *commuting family*.

The following result shows that there exists a (possibly complex) simultaneous change of basis under which every matrix in the commuting family $\{A_p : p \in \mathcal{P}\}$ can be written as the sum of a diagonal part and a nilpotent part, and these diagonal and nilpotent parts are pairwise commuting.

Lemma 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*

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

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

Proof. Lemma 5 is a consequence of [16, Cor. 2.4.6.4, p. 115]. An alternative proof based on the Jordan–Chevalley decomposition can be found in [39]. \square

In view of Lemma 5 and Remark 1, we assume, without loss of generality, that every matrix in the commuting family $\{A_p : p \in \mathcal{P}\}$ satisfies $A_p = D_p + N_p$ with a diagonal matrix $D_p := \text{diag}(a_p^1, \dots, a_p^n) \in \mathbb{C}^{n \times n}$, 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.⁵ Then (14) becomes the *switched commuting system* in \mathbb{C}^n defined by

$$\dot{x} = (D_\sigma + N_\sigma) x \quad (30)$$

with the commuting family of diagonal and nilpotent matrices $\{D_p, N_p : p \in \mathcal{P}\}$.

Theorem 6. *The topological entropy of the switched commuting system (30) satisfies*

$$h(D_\sigma + N_\sigma) = \limsup_{T \rightarrow \infty} \sum_{i=1}^n \bar{a}_i(T) \quad (31)$$

with

$$\bar{a}_i(T) := \frac{1}{T} \max_{t \in [0, T]} \sum_{p \in \mathcal{P}} \text{Re}(a_p^i) \tau_p(t), \quad i \in \{1, \dots, n\}, \quad (32)$$

where the active times τ_p are defined by (8).

Here the functions \bar{a}_i are the component-wise maximal weighted averages of the real parts of the eigenvalues. Hence the entropy $h(D_\sigma + N_\sigma)$ is independent of the nilpotent matrices N_p in (30).

To prove Theorem 6, we first formulate upper and lower bounds for the effect of the nilpotent matrices N_p ; the proof can be found in Appendix D.

Lemma 6. *Consider the commuting family of nilpotent matrices $\{N_p : p \in \mathcal{P}\}$. For each $\delta > 0$, there exists a constant $c_\delta > 0$ such that for all $v \in \mathbb{C}^n$,*

$$c_\delta^{-1} e^{-\delta t} \|v\| \leq \left\| e^{\sum_{p \in \mathcal{P}} N_p \tau_p(t)} v \right\| \leq c_\delta e^{\delta t} \|v\| \quad \forall t \geq 0 \quad (33)$$

with the active times τ_p defined by (8).

⁵ In particular, for each $p \in \mathcal{P}$, the diagonal entries a_p^i are also the eigenvalues of the original system matrix A_p [16, Th. 2.4.8.1, p. 117].

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 to (30) satisfy (see, e.g., [25, p. 31])

$$\begin{aligned} \|\xi_\sigma(x', t) - \xi_\sigma(x, t)\| &= \left\| e^{\sum_{p \in \mathcal{P}} (D_p + N_p) \tau_p(t)} (x' - x) \right\| \\ &= \left\| e^{\sum_{p \in \mathcal{P}} N_p \tau_p(t)} e^{\sum_{p \in \mathcal{P}} D_p \tau_p(t)} (x' - x) \right\| \quad \forall t \geq 0. \end{aligned}$$

Given a time horizon $T \geq 0$ and a radius $\varepsilon > 0$, Lemma 6 with $\delta = \varepsilon$ and $v = e^{\sum_{p \in \mathcal{P}} D_p \tau_p(t)} (x' - x)$ implies that there is a constant $c_\varepsilon > 0$ such that

$$c_\varepsilon^{-1} e^{-\varepsilon t} \left\| e^{\sum_{p \in \mathcal{P}} D_p \tau_p(t)} (x' - x) \right\| \leq \|\xi_\sigma(x', t) - \xi_\sigma(x, t)\| \leq c_\varepsilon e^{\varepsilon t} \left\| e^{\sum_{p \in \mathcal{P}} D_p \tau_p(t)} (x' - x) \right\| \quad \forall t \geq 0,$$

in which

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

as D_p are diagonal (recall that we take $\|\cdot\|$ to be the ∞ -norm; see Remark 1). Taking the maximum over $t \in [0, T]$, we obtain

$$c_\varepsilon^{-1} \max_{1 \leq i \leq n} e^{(\bar{a}_i(T) - \varepsilon)T} |x'_i - x_i| \leq \max_{t \in [0, T]} \|\xi_\sigma(x', t) - \xi_\sigma(x, t)\| \leq c_\varepsilon \max_{1 \leq i \leq n} e^{(\bar{a}_i(T) + \varepsilon)T} |x'_i - x_i| \quad (34)$$

with \bar{a}_i defined by (32).

First, consider the grid $G(\theta)$ defined by (15) with

$$\theta_i := e^{-(\bar{a}_i(T) + \varepsilon)T} / c_\varepsilon, \quad i \in \{1, \dots, n\},$$

and the corresponding hyperrectangles $R(\hat{x})$ defined by (16). Comparing (16) and the upper bound in (34) to (3), we see that $R(\hat{x}) \subset B_{D_\sigma + N_\sigma}(\hat{x}, \varepsilon, T)$ for all $\hat{x} \in G(\theta)$. Then Lemma 2 implies that $G(\theta)$ is (T, ε) -spanning, and substituting (17) into (5) yields

$$\begin{aligned} h(D_\sigma + N_\sigma) &\leq \lim_{\varepsilon \searrow 0} \limsup_{T \rightarrow \infty} \sum_{i=1}^n \frac{\log(2/\theta_i + 1)}{T} \\ &= \lim_{\varepsilon \searrow 0} \limsup_{T \rightarrow \infty} \sum_{i=1}^n \frac{\log(1/\theta_i)}{T} + \lim_{\varepsilon \searrow 0} \limsup_{T \rightarrow \infty} \sum_{i=1}^n \frac{\log(2 + \theta_i)}{T} \\ &= \lim_{\varepsilon \searrow 0} \limsup_{T \rightarrow \infty} \sum_{i=1}^n \frac{\log(1/\theta_i)}{T} \\ &= \limsup_{T \rightarrow \infty} \sum_{i=1}^n \bar{a}_i(T) + \lim_{\varepsilon \searrow 0} n\varepsilon + \lim_{\varepsilon \searrow 0} \limsup_{T \rightarrow \infty} \frac{n \log(c_\varepsilon/\varepsilon)}{T} \\ &= \limsup_{T \rightarrow \infty} \sum_{i=1}^n \bar{a}_i(T). \end{aligned}$$

Second, consider the grid $G(\theta)$ defined by (15) with

$$\theta_i := e^{-(\bar{a}_i(T) - \varepsilon)T} / c_\varepsilon, \quad i \in \{1, \dots, n\},$$

and the corresponding hyperrectangles $R(\hat{x})$ defined by (16). Comparing (16) and the lower bound in (34) to (3), we see that $B_{D_\sigma + N_\sigma}(\hat{x}, \varepsilon, T) \subset R(\hat{x})$ for all $\hat{x} \in G(\theta)$. Then Lemma 2 implies that $G(\theta)$ is

(T, ε) -separated, and substituting (18) into (5) yields

$$\begin{aligned}
h(D_\sigma + N_\sigma) &\geq \lim_{\varepsilon \searrow 0} \limsup_{T \rightarrow \infty} \sum_{i=1}^n \frac{\log(2/\theta_i - 1)}{T} \\
&= \lim_{\varepsilon \searrow 0} \limsup_{T \rightarrow \infty} \sum_{i=1}^n \frac{\log(1/\theta_i)}{T} + \lim_{\varepsilon \searrow 0} \limsup_{T \rightarrow \infty} \sum_{i=1}^n \frac{\log(2 - \theta_i)}{T} \\
&= \lim_{\varepsilon \searrow 0} \limsup_{T \rightarrow \infty} \sum_{i=1}^n \frac{\log(1/\theta_i)}{T} \\
&= \limsup_{T \rightarrow \infty} \sum_{i=1}^n \bar{a}_i(T) - \lim_{\varepsilon \searrow 0} n\varepsilon - \lim_{\varepsilon \searrow 0} \limsup_{T \rightarrow \infty} \frac{n \log(c_\varepsilon \varepsilon)}{T} \\
&= \limsup_{T \rightarrow \infty} \sum_{i=1}^n \bar{a}_i(T). \quad \square
\end{aligned}$$

Based on the formula (31), we establish upper bounds for the entropy $h(D_\sigma + N_\sigma)$ that require less information.

First, we construct an upper bound in terms of the entropy in each individual scalar component.

Proposition 7. *The topological entropy of the switched commuting system (30) is upper bounded by*

$$h(D_\sigma + N_\sigma) \leq \sum_{i=1}^n \max\{\hat{a}_i, 0\} \quad (35)$$

with

$$\hat{a}_i := \limsup_{t \rightarrow \infty} \sum_{p \in \mathcal{P}} \operatorname{Re}(a_p^i) \rho_p(t), \quad i \in \{1, \dots, n\}, \quad (36)$$

where the active rates ρ_p are defined by (9). Moreover, if the limits $\lim_{t \rightarrow \infty} \rho_p(t)$ exist for all $p \in \mathcal{P}$, then (35) holds with equality.⁶

Here the constants \hat{a}_i are the component-wise asymptotic weighted averages of the real parts of the eigenvalues. Combining (35) with (27) and Remark 4, we conclude that the entropy of the switching commuting system (30) is upper bounded by the sum of the entropy of the switched scalar system (26) with $a_p = a_p^i$ for each $i \in \{1, \dots, n\}$. For the case where all the active rates ρ_p converge, as (35) holds with equality, (20) implies that $h(D_\sigma + N_\sigma)$ equals the entropy of the LTI system (19) with the asymptotic weighted average matrix $A := \sum_{p \in \mathcal{P}} (D_p + N_p) \lim_{t \rightarrow \infty} \rho_p(t)$.

Proof of Proposition 7. Following (31) and the subadditivity of limit suprema, we have

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

where the last equality follows from (13) with $a_p = \operatorname{Re}(a_p^i)$ in (11) and (12). For the case where the limits $\lim_{t \rightarrow \infty} \rho_p(t)$ exist for all $p \in \mathcal{P}$, the inequality in the derivation above becomes an equality due to the additivity of limits. \square

Second, we construct an upper bound in terms of the entropy of each individual mode.

Proposition 8. *The topological entropy of the switched commuting system (30) is upper bounded by*

$$h(D_\sigma + N_\sigma) \leq \limsup_{t \rightarrow \infty} \sum_{p \in \mathcal{P}} h(D_p + N_p) \rho_p(t) \quad (37)$$

with the active rates ρ_p defined by (9), where $h(D_p + N_p)$ denotes the topological entropy of the p -th mode and satisfies (20) with $A = D_p + N_p$. Moreover, if $\operatorname{Re}(a_p^i) \geq 0$ for all $i \in \{1, \dots, n\}$ and $p \in \mathcal{P}$, then (37) holds with equality.

⁶ For example, when the switching signal σ is periodic; see [34, Sec. 3.2.1] for more conditions.

Combining (37) with (27), we conclude that the entropy of the switching commuting system (30) is upper bounded by the entropy of the switched scalar system (26) with $a_p = h(D_p + N_p)$. For the case where all the eigenvalues of system matrices a_p^i have nonnegative real parts, the general lower bound (24) coincides with (37) and thus also holds with equality.

Proof. For each $p \in \mathcal{P}$, let

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

Following (20), the entropy of the p -th mode of (30) satisfies $h(D_p + N_p) = \sum_{i=1}^n \bar{a}_p^i$; see also footnote 5. Consequently, (31) and (32) imply

$$\begin{aligned} h(D_\sigma + N_\sigma) &= \limsup_{T \rightarrow \infty} \sum_{i=1}^n \frac{1}{T} \max_{t \in [0, T]} \sum_{p \in \mathcal{P}} \operatorname{Re}(a_p^i) \tau_p(t) \\ &\leq \limsup_{T \rightarrow \infty} \sum_{i=1}^n \frac{1}{T} \sum_{p \in \mathcal{P}} \bar{a}_p^i \tau_p(T) = \limsup_{t \rightarrow \infty} \sum_{p \in \mathcal{P}} \left(\sum_{i=1}^n \bar{a}_p^i \right) \rho_p(t) = \limsup_{t \rightarrow \infty} \sum_{p \in \mathcal{P}} h(D_p + N_p) \rho_p(t). \end{aligned}$$

For the case where $\operatorname{Re}(a_p^i) \geq 0$ for all $i \in \{1, \dots, n\}$ and $p \in \mathcal{P}$, the inequality in the derivation above becomes an equality as $\bar{a}_p^i = \operatorname{Re}(a_p^i)$. \square

The upper bounds (35) and (37) can be further relaxed to obtain the following upper bounds for $h(D_\sigma + N_\sigma)$; the proof is along the lines of that of Corollary 5 and thus omitted here.

Corollary 9. *The topological entropy of the switched commuting system (30) is upper bounded by*

$$h(D_\sigma + N_\sigma) \leq \sum_{p \in \mathcal{P}} h(D_p + N_p) \hat{\rho}_p \quad (38)$$

with the asymptotic active rates $\hat{\rho}_p$ defined by (10), and also by

$$h(D_\sigma + N_\sigma) \leq \max_{p \in \mathcal{P}} h(D_p + N_p), \quad (39)$$

where $h(D_p + N_p)$ denotes the topological entropy of the p -th mode and satisfies (20) with $A = D_p + N_p$. Moreover, if the limits $\lim_{t \rightarrow \infty} \rho_p(t)$ exist and $\operatorname{Re}(a_p^i) \geq 0$ for all $i \in \{1, \dots, n\}$ and $p \in \mathcal{P}$, then (38) holds with equality.

The relations between the formula in Theorem 6 and the upper bounds in Propositions 7 and 8 and Corollary 9 are summarized in Fig. 2 and Remark 6, and illustrated numerically in Example 3.

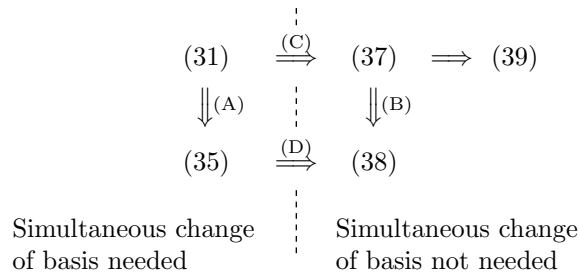


Fig. 2 Relations between the formula (31) and the upper bounds (35), (37), (38), and (39). The implications (A) and (B) become equivalences if all the active rates ρ_p converge; the implication (C) becomes an equivalence if all the eigenvalues of system matrices a_p^i have nonnegative real parts; the implication (D) becomes an equivalence if both of these conditions hold. The relations between these upper bounds that are not specified in this diagram are undetermined.

Remark 6. 1. Unlike the formula (31) and the upper bound (35), the upper bounds (37), (38), and (39) are independent of the order of eigenvalues (i.e., in which scalar component each eigenvalue of the system matrices U_p is), and thus can be calculated for a switched linear system with pairwise commuting matrices without knowledge of the simultaneous change of basis in Lemma 5.

2. For a fixed family of matrices $\{D_p : p \in \mathcal{P}\}$, compared with the formula (31), the upper bound (35) depends only on the component-wise asymptotic weighted averages \hat{a}_i ; the upper bound (37) depends only on the asymptotic weighted average of the entropy of each individual mode $h(D_p + N_p)$; the upper bound (38) depends only on the asymptotic active rates $\hat{\rho}_p$; the upper bound (39) is independent of switching.
3. The upper bounds (35) and (37) are both useful in the sense that neither is more conservative than the other; the same conclusion holds for the upper bounds (38) and (39).

Example 3. Consider the index set $\mathcal{P} = \{1, 2\}$ and the switching signals σ_0 , σ_1 , and σ_2 defined in Example 2. As the entropy of the switched commuting system (30) is independent of its nilpotent part, we 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 (31) and their upper bounds calculated using (35), (37), (38), and (39), as well as the general upper and lower bounds (23) and (24), are summarized in Table 2. For the case with σ_2 , the calculation using (35), (37), (23), and (24) is along the lines of calculating $h(a_{\sigma_2})$ in Example 2; see Appendix E for the calculation using (31).

Table 2 Entropy values and bounds for the switched commuting systems in Example 3.

	$(\hat{\rho}_1, \hat{\rho}_2)$	(31)	(35)	(37)	(38)	(39)	(23)	(24)
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

6 Entropy of switched triangular systems

In this section, we consider the case of switched linear systems with simultaneously triangularizable matrices, that is, there exists a (possibly complex) change of basis under which the matrices A_p in (14) are all upper triangular.⁷ Hence and in view of Remark 1, we assume, without loss of generality, that every A_p is upper triangular, and denote it by

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

Then (14) becomes the *switched triangular system* in \mathbb{C}^n defined by

$$\dot{x} = U_\sigma x \tag{40}$$

with the family of upper-triangular matrices $\{U_p : p \in \mathcal{P}\}$.

Theorem 10. *The topological entropy of the switched triangular system (40) is upper bounded by*

$$h(U_\sigma) \leq \limsup_{T \rightarrow \infty} \left(n \bar{a}_1(T) + \sum_{i=2}^n (n+1-i) \bar{d}_i(T) \right) \tag{41}$$

⁷ A sufficient condition for simultaneous triangularizability is that the matrices A_p are pairwise commuting (see, e.g., [16, Th. 2.3.3, p. 103]). More sufficient conditions can be found in [24]. A necessary and sufficient condition is that the Lie algebra $\{A_p : p \in \mathcal{P}\}_{LA}$ is solvable (see, e.g., [17, pp. 10, 16]). More necessary and sufficient conditions can be found in [29, 12].

with

$$\bar{a}_1(T) := \frac{1}{T} \max_{t \in [0, T]} \sum_{p \in \mathcal{P}} \operatorname{Re}(a_p^1) \tau_p(t) \quad (42)$$

and

$$\bar{d}_i(T) := \frac{1}{T} \max_{t \in [0, T]} \sum_{p \in \mathcal{P}} \operatorname{Re}(a_p^i - a_p^{i-1}) \tau_p(t), \quad i \in \{2, \dots, n\}, \quad (43)$$

where the active times τ_p are defined by (8).

Here the function \bar{a}_1 is the maximal weighted average of the real parts of the eigenvalues in the first scalar component, and the functions \bar{d}_i are the maximal weighted averages of the differences between the real parts of the eigenvalues in adjacent scalar components. Hence the upper bound (41) is independent of the off-diagonal entries $b_p^{i,j}$ of the matrices U_p in (40).

To prove Theorem 10, we first establish a formula for the solution to the switched triangular system (40) and an upper bound for its norm, which are also of independent interest (e.g., the upper bound is used to establish a stability result in Section 7); the proofs can be found in Appendices F and G. Here we denote by $\xi_\sigma^k(x, t)$ the k -th scalar component of the solution $\xi_\sigma(x, t)$.

Lemma 7. *For every initial state $x \in \mathbb{R}^n$ and $k \in \{1, \dots, n\}$, the k -th scalar component of the solution $\xi_\sigma(x, t)$ of (40) satisfies*

$$\xi_\sigma^k(x, t) = e^{\eta_k(t)} \left(x_k + \sum_{l=k+1}^n \sum_{i=1}^{l-k} x_l \Psi(t, \mathcal{C}_{k,l,i}) \right) \quad \forall t \geq 0 \quad (44)$$

with

$$\eta_i(t) := \sum_{p \in \mathcal{P}} a_p^i \tau_p(t), \quad i \in \{1, \dots, n\}, \quad (45)$$

the sets

$$\mathcal{C}_{k,l,i} := \{(c_0, \dots, c_i) \in \mathbb{N}^{i+1} : k = c_0 < c_1 < \dots < c_{i-1} < c_i = l\} \quad (46)$$

for $l \in \{k+1, \dots, n\}$ and $i \in \{1, \dots, l-k\}$, and

$$\Psi(t, \mathcal{C}_{k,l,i}) := \sum_{(c_0, \dots, c_i) \in \mathcal{C}_{k,l,i}} \int_0^t \int_0^{s_1} \dots \int_0^{s_{i-1}} \prod_{j=1}^i \left(b_{\sigma(s_j)}^{c_{j-1}, c_j} e^{\eta_{c_j}(s_j) - \eta_{c_{j-1}}(s_j)} \right) ds_j.$$

Lemma 8. *For every initial state $x \in \mathbb{R}^n$, the solution $\xi_\sigma(x, t)$ of (40) satisfies*

$$\|\xi_\sigma(x, t)\| \leq e^{\operatorname{Re}(\eta_1(t))} \sum_{i=1}^n \left((b_M t + 1)^{i-1} e^{\sum_{j=2}^i \bar{d}_j(t) t} |x_i| \right) \quad \forall t \geq 0 \quad (47)$$

with the constant $b_M := \max_{p \in \mathcal{P}, 1 \leq i < j \leq n} |b_p^{i,j}|$, and η_1 and \bar{d}_i defined by (45) and (43), respectively.

From Lemma 8 and the proof of Theorem 10 below, we will see that the terms related to the off-diagonal entries $b_p^{i,j}$ of the matrices U_p in (40) are absorbed into the polynomials $(b_M t + 1)^{i-1}$, and thus do not appear in the bound (41).

Proof of Theorem 10. Following Lemma 8, for all initial states $x, x' \in K$, the corresponding solutions to (40) satisfy

$$\|\xi_\sigma(x', t) - \xi_\sigma(x, t)\| \leq e^{\operatorname{Re}(\eta_1(t))} \sum_{i=1}^n \left((b_M t + 1)^{i-1} e^{\sum_{j=2}^i \bar{d}_j(t) t} |x'_i - x_i| \right)$$

for all $t \geq 0$. Given a time horizon $T \geq 0$ and a radius $\varepsilon > 0$, following the definition (42) of \bar{a}_1 and the fact that $(b_M t + 1) \geq 1$ and $\bar{d}_i(t) t \geq 0$ are nondecreasing in t for all $i \in \{2, \dots, n\}$, we obtain

$$\max_{t \in [0, T]} \|\xi_\sigma(x', t) - \xi_\sigma(x, t)\| \leq \sum_{i=1}^n \left((b_M T + 1)^{i-1} e^{(\bar{a}_1(T) + \sum_{j=2}^i \bar{d}_j(T)) T} |x'_i - x_i| \right) \quad (48)$$

Consider the grid $G(\theta)$ defined by (15) with

$$\theta_i := e^{-(\bar{a}_1(T) + \sum_{j=2}^i \bar{d}_j(T))T} \varepsilon / (n(b_M T + 1)^{i-1}), \quad i \in \{1, \dots, n\},$$

and the corresponding hyperrectangles $R(\hat{x})$ defined by (16). Comparing (16) and (48) to (3), we see that $R(\hat{x}) \subset B_{U_\sigma}(\hat{x}, \varepsilon, T)$ for all $\hat{x} \in G(\theta)$. Then Lemma 2 implies that $G(\theta)$ is (T, ε) -spanning, and substituting (17) into (5) yields

$$\begin{aligned} h(U_\sigma) &\leq \lim_{\varepsilon \searrow 0} \limsup_{T \rightarrow \infty} \sum_{i=1}^n \frac{\log(2/\theta_i + 1)}{T} \\ &= \lim_{\varepsilon \searrow 0} \limsup_{T \rightarrow \infty} \sum_{i=1}^n \frac{\log(1/\theta_i)}{T} + \lim_{\varepsilon \searrow 0} \limsup_{T \rightarrow \infty} \sum_{i=1}^n \frac{\log(2 + \theta_i)}{T} \\ &= \lim_{\varepsilon \searrow 0} \limsup_{T \rightarrow \infty} \sum_{i=1}^n \frac{\log(1/\theta_i)}{T} \\ &= \limsup_{T \rightarrow \infty} \sum_{i=1}^n \left(\bar{a}_1(T) + \sum_{j=2}^i \bar{d}_j(T) \right) \\ &\quad + \limsup_{T \rightarrow \infty} \frac{n(n-1) \log(b_M T + 1)}{2T} + \lim_{\varepsilon \searrow 0} \limsup_{T \rightarrow \infty} \frac{n \log(n/\varepsilon)}{T} \\ &= \limsup_{T \rightarrow \infty} \left(n\bar{a}_1(T) + \sum_{i=2}^n (n+1-i) \bar{d}_i(T) \right), \end{aligned}$$

where in the last step, we change the order of summation by grouping terms in the same scalar component. \square

Based on the upper bound (41), we establish additional upper bounds for the entropy $h(U_\sigma)$ that are more conservative but require less information.

First, we construct an upper bound in terms of the entropy in the first scalar component and the entropy differences between adjacent scalar components; the proof is along the lines of that of Proposition 7 and thus omitted here.

Proposition 11. *The topological entropy of the switched triangular system (40) is upper bounded by*

$$h(U_\sigma) \leq n \max\{\hat{a}_1, 0\} + \sum_{i=2}^n (n+1-i) \max\{\hat{d}_i, 0\} \quad (49)$$

with

$$\hat{a}_1 := \limsup_{t \rightarrow \infty} \sum_{p \in \mathcal{P}} \operatorname{Re}(a_p^1) \rho_p(t) \quad (50)$$

and

$$\hat{d}_i := \limsup_{t \rightarrow \infty} \sum_{p \in \mathcal{P}} \operatorname{Re}(a_p^i - a_p^{i-1}) \rho_p(t), \quad i = 2, \dots, n, \quad (51)$$

where the active rates ρ_p are defined by (9).

Here the constant \hat{a}_1 is the asymptotic weighted average of the real parts of the eigenvalues in the first scalar component, and the constants \hat{d}_i are the asymptotic weighted averages of the differences between the real parts of the eigenvalues in adjacent scalar components.

Second, we construct two upper bounds in terms of two entropy-related quantities of each individual mode; the proof is along the lines of that of Proposition 8 and thus omitted here.

Proposition 12. *The topological entropy of the switched triangular system (40) is upper bounded by*

$$h(U_\sigma) \leq \limsup_{t \rightarrow \infty} \sum_{p \in \mathcal{P}} \tilde{h}(U_p) \rho_p(t) \quad (52)$$

$$\leq \limsup_{t \rightarrow \infty} \sum_{p \in \mathcal{P}} \tilde{h}^{\mathbb{S}}(U_p) \rho_p(t) \quad (53)$$

with

$$\begin{aligned}\tilde{h}(U_p) &:= n \max\{\operatorname{Re}(a_p^1), 0\} + \sum_{i=2}^n (n+1-i) \max\{\operatorname{Re}(a_p^i - a_p^{i-1}), 0\}, \\ \tilde{h}^{\mathbb{S}}(U_p) &:= \max_{(\lambda_1, \dots, \lambda_n) \in \mathbb{S}(\operatorname{spec}(U_p))} \left(n \max\{\operatorname{Re}(\lambda^1), 0\} + \sum_{i=2}^n (n+1-i) \max\{\operatorname{Re}(\lambda^i - \lambda^{i-1}), 0\} \right)\end{aligned}\quad (54)$$

for $p \in \mathcal{P}$, where the active rates ρ_p are defined by (9), and $\mathbb{S}(a_p^1, \dots, a_p^n)$ denotes the set of permutations of $\{a_p^1, \dots, a_p^n\}$.

The upper bounds (49), (52), and (53) can be further relaxed to establish the following upper bounds for $h(U_\sigma)$; the proof is along the lines of that of Corollary 5 and thus omitted here.

Corollary 13. *The topological entropy of the switched triangular system (40) is upper bounded by*

$$h(U_\sigma) \leq \sum_{p \in \mathcal{P}} \tilde{h}(U_p) \hat{\rho}_p \quad (55)$$

$$\leq \sum_{p \in \mathcal{P}} \tilde{h}^{\mathbb{S}}(U_p) \hat{\rho}_p \quad (56)$$

with the asymptotic active rates $\hat{\rho}_p$ defined by (10), and also by

$$h(U_\sigma) \leq \max_{p \in \mathcal{P}} \tilde{h}(U_p) \quad (57)$$

$$\leq \max_{p \in \mathcal{P}} \tilde{h}^{\mathbb{S}}(U_p), \quad (58)$$

where the quantities $\tilde{h}(U_p)$ and $\tilde{h}^{\mathbb{S}}(U_p)$ are defined by (54).

The relations between the upper bounds in Theorem 10, Propositions 11 and 12, and Corollary 13 are summarized in Fig. 3 and Remark 7, and illustrated numerically in Example 4.

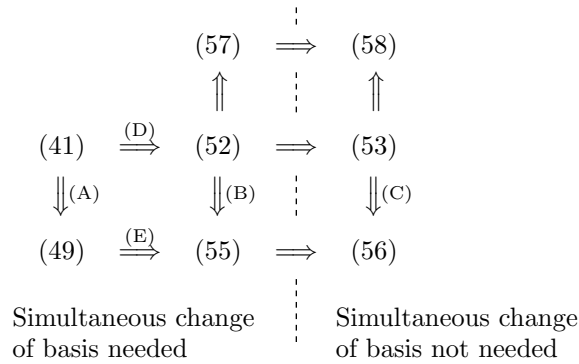


Fig. 3 Relations between the upper bounds (41), (49), (52), (53), (55), (56), (57), and (58). The implications (A), (B), and (C) become equivalences if all the active rates ρ_p converge; the implication (D) becomes an equivalence if the eigenvalues of all the system matrices U_p satisfy $0 \leq \operatorname{Re}(a_p^1) \leq \dots \leq \operatorname{Re}(a_p^n)$; the implication (E) becomes an equivalence if both of these conditions hold. The relations between these upper bounds that are not specified in this diagram are undetermined.

Remark 7. 1. Unlike the upper bounds (41), (49), (52), (55), and (57), the upper bounds (53), (56), and (58) are independent of the order of eigenvalues (i.e., in which scalar component each eigenvalue of the system matrices U_p is), and thus can be calculated for a switched linear system with simultaneously triangularizable matrices without knowledge of a basis for simultaneous triangularization.

2. For a fixed family of matrices $\{U_p : p \in \mathcal{P}\}$, compared with the upper bound (41), the upper bound (49) depends only on the asymptotic weighted averages \hat{a}_1 and \hat{d}_i ; the upper bounds (52) and (53) depend only on the asymptotic weighted averages of the entropy related quantities of each individual mode $\tilde{h}(U_p)$ and $\tilde{h}^{\mathbb{S}}(U_p)$; the upper bounds (55) and (56) depend only on the asymptotic active rates $\hat{\rho}_p$; the upper bounds (57) and (58) are independent of switching.

3. The upper bound (49) and (52) are both useful in the sense that neither is more conservative than the other; this is also true for the upper bounds (55) and (57). The same conclusion holds if the corresponding relaxed upper bounds (53), (56), and (58) are taken into consideration.

Example 4. Consider the index set $\mathcal{P} = \{1, 2\}$, the switching signals σ_0, σ_1 , and σ_2 defined in Example 2, and the upper-triangular matrices

$$U_1 = \begin{bmatrix} -1 & 1 \\ 0 & 2 \end{bmatrix}, \quad U_2 = \begin{bmatrix} 3 & 1 \\ 0 & -1 \end{bmatrix}.$$

The upper bounds of $h(U_{\sigma_0})$, $h(U_{\sigma_1})$, and $h(U_{\sigma_2})$ calculated using (41), (49), (52), (55), and (57), as well as the general upper and lower bounds (23) and (24), are summarized in Table 3. For the case with σ_2 , the calculation using (49), (52), (55), and (57) is along the lines of calculating $h(a_{\sigma_2})$ in Example 2; see Appendix H for the calculation using (41).

Table 3 Entropy bounds for the switched triangular systems in Example 4.

	$(\hat{\rho}_1, \hat{\rho}_2)$	(41)	(49)	(52)	(55)	(57)	(23)	(24)
σ_0	(1, 0)	3	3	3	3	6	4	1
σ_1	(0.5, 0.5)	2	2	4	4	6	6	1.5
σ_2	(0.9, 0.9)	5.46	7.5	5.7	8.1	6	7.6	1.9

7 Entropy and stability

In this section, we present stability conditions inspired by the upper bounds for topological entropy above. Suppose that the origin is a common equilibrium for all modes of the switched system (2), that is, $f_p(0) = 0$ for all $p \in \mathcal{P}$. The switched system (2) with switching signal σ is (*Lyapunov*) *stable* if for each $\varepsilon > 0$, there exists a $\delta > 0$ such that for every initial state $x \in \mathbb{R}^n$ satisfying $\|x\| \leq \delta$, the corresponding solution satisfy $\|\xi_\sigma(x, t)\| \leq \varepsilon$ for all $t \geq 0$; it is *globally exponentially stable (GES)* if there exist constants $c, \kappa > 0$ such that for all $x \in \mathbb{R}^n$,

$$\|\xi_\sigma(x, t)\| \leq ce^{-\kappa t} \|x\| \quad \forall t \geq 0.$$

Clearly, stability implies that the entropy $h(f_\sigma, K) = 0$ for a small enough initial set K , and GES implies $h(f_\sigma, K) = 0$ for all initial sets K .

For the general switched linear system (14), both stability and GES imply $h(A_\sigma) = 0$. However, it is possible that $h(A_\sigma) = 0$ while (2) is unstable (with the uncertainty about the state growing subexponentially); for example, the LTI system (19)—which can be viewed as a switched system with a constant switching signal—with

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

is unstable and yet $h(A) = 0$ because of (20). The upper bound (23) in Theorem 3 shows that the entropy $h(A_\sigma)$ can be upper bounded in terms of the asymptotic weighted average of the matrix measures $\mu(A_p)$, which can also be used to establish GES.

Proposition 14. *The switched linear system (14) is GES provided that the asymptotic weighted average of the matrix measures $\mu(A_p)$ defined by (21) satisfies*

$$\limsup_{t \rightarrow \infty} \sum_{p \in \mathcal{P}} \mu(A_p) \rho_p(t) < 0. \quad (59)$$

In particular, Proposition 14 implies that (14) is GES under arbitrary switching if the matrices measures $\mu(A_p)$ are all negative.

Corollary 15. *The switched linear system (14) is GES for all switching signals σ provided that the matrix measures defined by (21) satisfy $\mu(A_p) < 0$ for all $p \in \mathcal{P}$.*

Proof of Proposition 14. Following (59), there exists a constant $\kappa > 0$ such that

$$\kappa < -\frac{1}{2} \limsup_{t \rightarrow \infty} \sum_{p \in \mathcal{P}} \mu(A_p) \rho_p(t).$$

Then the limit supremum in (59) implies that there is a large enough $T_\kappa \geq 0$ such that

$$\sum_{p \in \mathcal{P}} \mu(A_p) \rho_p(t) < \kappa + \limsup_{s \rightarrow \infty} \sum_{p \in \mathcal{P}} \mu(A_p) \rho_p(s) < -\kappa \quad \forall t > T_\kappa.$$

Hence for every initial state $x \in \mathbb{R}^n$, Lemma 4 implies that the solution to (14) satisfies

$$\|\xi_\sigma(x, t)\| \leq e^{\sum_{p \in \mathcal{P}} \mu(A_p) \tau_p(t)} \|x\| \leq e^{-\kappa t} \|x\| \quad \forall t > T_\kappa.$$

Moreover, we have

$$\|\xi_\sigma(x, t)\| \leq e^{\sum_{p \in \mathcal{P}} \mu(A_p) \tau_p(t)} \|x\| \leq e^{\mu_m t} \|x\| \leq e^{\max\{\mu_m, 0\} T_\kappa} \|x\|$$

for all $t \in [0, T_\kappa]$ with the constant $\mu_m := \max_{p \in \mathcal{P}} \mu(A_p)$. Therefore,

$$\|\xi_\sigma(x, t)\| \leq e^{(\max\{\mu_m, 0\} + \kappa) T_\kappa} e^{-\kappa t} \|x\| \quad \forall t \geq 0,$$

that is, (14) is GES. \square

Similar to Proposition 14, the asymptotic weighted averages used in the upper bounds for topological entropy in Corollary 4 and Propositions 7 and 11 can also be used to establish GES for the corresponding switched linear systems generated by matrices with commutation relations.

Corollary 16. *The switched scalar system (26) is GES provided that the asymptotic weighted average \hat{a} defined by (11) satisfies $\hat{a} < 0$.*

Proposition 17. *The switched commuting system (30) is GES provided that \hat{a}_i defined by (36) satisfy $\hat{a}_i < 0$ for all $i \in \{1, \dots, n\}$.*

In particular, Proposition 17 implies that the switched commuting system (30) is GES for all switching signals σ if the diagonal matrices D_p are all Hurwitz; thus it generalizes the well-known result that a switched linear system generated by a finite family of pairwise commuting Hurwitz matrices is GES under arbitrary switching (see, e.g., [25, Th. 2.5, p. 31]).

Proof of Proposition 17. The proof is established by combining Lemma 6 with similar arguments to those in the proof of Proposition 14. For every initial state $x \in \mathbb{R}^n$, as $\{D_p, N_p : p \in \mathcal{P}\}$ is a commuting family, the solution to (30) satisfies

$$\|\xi_\sigma(x, t)\| = \left\| e^{\sum_{p \in \mathcal{P}} (D_p + N_p) \tau_p(t)} x \right\| = \left\| e^{\sum_{p \in \mathcal{P}} N_p \tau_p(t)} e^{\sum_{p \in \mathcal{P}} D_p \tau_p(t)} x \right\|$$

for all $t \geq 0$. As $\hat{a}_i < 0$ for all $i \in \{1, \dots, n\}$, there exists a constant $\kappa > 0$ such that $\kappa < -\hat{a}_i/3$ for all $i \in \{1, \dots, n\}$. Then the limit suprema in (36) imply that there is a large enough $T_\kappa \geq 0$ such that

$$\sum_{p \in \mathcal{P}} \operatorname{Re}(a_p^i) \rho_p(t) < \hat{a}_i + \kappa < -2\kappa \quad \forall t > T_\kappa, \forall i \in \{1, \dots, n\}.$$

Hence Lemma 6 with $\delta = \kappa$ and $v = e^{\sum_{p \in \mathcal{P}} D_p \tau_p(t)} x$ implies that there is a constant $c_\kappa > 0$ such that

$$\begin{aligned} \|\xi_\sigma(x, t)\| &\leq c_\kappa e^{\kappa t} \left\| e^{\sum_{p \in \mathcal{P}} D_p \tau_p(t)} x \right\| = c_\kappa e^{\kappa t} \max_{1 \leq i \leq n} e^{\sum_{p \in \mathcal{P}} \operatorname{Re}(a_p^i) \tau_p(t)} |x_i| \\ &\leq c_\kappa \max_{1 \leq i \leq n} e^{(\kappa + \sum_{p \in \mathcal{P}} \operatorname{Re}(a_p^i) \rho_p(t)) t} \|x\| < c_\kappa e^{-\kappa t} \|x\| \quad \forall t \geq T_\kappa, \end{aligned}$$

where the equality follows from the fact that D_p are diagonal. Moreover, we have

$$\|\xi_\sigma(x, t)\| \leq c_\kappa e^{(a_m + \kappa) t} \|x\| \leq e^{(\max\{a_m, 0\} + \kappa) T_\kappa} \|x\| \quad \forall t \in [0, T_\kappa]$$

with the constant $a_m := \max_{p \in \mathcal{P}, 1 \leq i \leq n} \operatorname{Re}(a_p^i)$. Therefore,

$$\|\xi_\sigma(x, t)\| \leq c_\kappa e^{(\max\{a_m, 0\} + 2\kappa) T_\kappa} e^{-\kappa t} \|x\| \quad \forall t \geq 0,$$

that is, (30) is GES. \square

Proposition 18. *The switched triangular system (40) is GES provided that \hat{a}_1 and \hat{d}_i defined by (50) and (51) satisfy $\hat{a}_1 < 0$ and $\hat{d}_i < 0$ for all $i \in \{2, \dots, n\}$.*

Proof. The proof is established by combining Lemmas 1 and 8 with similar arguments to those in the proof of Proposition 14. As $\hat{a}_1 < 0$, there exists a constant $\kappa > 0$ such that $\kappa < -\hat{a}_1/(n+2)$. Then the limit supremum in (50) implies that there is a large enough $T'_\kappa \geq 0$ such that

$$\sum_{p \in \mathcal{P}} \operatorname{Re}(a_p^1) \rho_p(t) < \hat{a}_1 + \kappa < -(n+1)\kappa \quad \forall t \geq T'_\kappa.$$

Also, following (13) with $a_p = \operatorname{Re}(a_p^i - a_p^{i-1})$ in (11) and (12) for $i = 2, \dots, n$, the maximal weighted averages \bar{d}_i defined by (43) satisfy

$$\limsup_{t \rightarrow \infty} \bar{d}_i(t) = \max\{\hat{d}_i, 0\} = 0 \quad \forall i \in \{2, \dots, n\},$$

in which the limit suprema imply that there is a large enough $T''_\kappa \geq 0$ such that

$$\bar{d}_i(t) < \kappa \quad \forall t > T''_\kappa, \forall i \in \{2, \dots, n\}.$$

Finally, there is a large enough $T_\kappa \geq \max\{T'_\kappa, T''_\kappa\}$ such that

$$(b_M t + 1)^{n-1} < e^{\kappa t} \quad \forall t > T_\kappa.$$

Hence for every initial state $x \in \mathbb{R}^n$, Lemma 8, together with the definition (45) of η_1 , implies that the solution to (40) satisfies

$$\begin{aligned} \|\xi_\sigma(x, t)\| &\leq e^{\sum_{p \in \mathcal{P}} \operatorname{Re}(a_p^1) \tau_p(t)} \sum_{i=1}^n \left((b_M t + 1)^{i-1} e^{\sum_{j=2}^i \bar{d}_j(t) t} |x_i| \right) \\ &\leq (b_M t + 1)^{n-1} e^{(\sum_{p \in \mathcal{P}} \operatorname{Re}(a_p^1) \rho_p(t) + \sum_{j=2}^n \bar{d}_j(t) t) t} \sum_{i=1}^n |x_i| < n e^{-\kappa t} \|x\| \end{aligned}$$

for all $t \geq T_k$. Moreover, we have

$$\begin{aligned} \|\xi_\sigma(x, t)\| &\leq (b_M t + 1)^{n-1} e^{(a_m^1 + \sum_{j=2}^n d_m^j) t} \sum_{i=1}^n |x_i| \\ &\leq n(b_M T_k + 1)^{n-1} e^{(\max\{a_m^1, 0\} + \sum_{j=2}^n \max\{d_m^j, 0\}) T_k} \|x\| \end{aligned}$$

for all $t \in [0, T_k]$ with the constants $a_m^1 := \max_{p \in \mathcal{P}} \operatorname{Re}(a_p^1)$ and $d_m^i := \max_{p \in \mathcal{P}} \operatorname{Re}(a_p^i - a_p^{i-1})$ for $i = 2, \dots, n$. Therefore,

$$\|\xi_\sigma(x, t)\| \leq n(b_M T_k + 1)^{n-1} e^{(\max\{a_m^1, 0\} + \sum_{j=2}^n \max\{d_m^j, 0\} + \kappa) T_k} e^{-\kappa t} \|x\|$$

for all $t \geq 0$, that is, (40) is GES. \square

Remark 8. Aside from the proofs above, Propositions 14, 17, and 18 can also be established using the *destabilizing-perturbation* method proposed in [42]. More specifically, they can be proved by combining the corresponding upper bounds (23), (35), and (49) with [42, Th. 5.1], respectively. The proofs presented here are more direct, whereas the results in [42] lead to additional upper bounds for topological entropy and additional stability conditions. In particular, Proposition 18 does not generalize the standard result that a switched linear system generated by upper-triangular Hurwitz matrices is GES under arbitrary switching, which is achieved in [42, Cor. 5.3].

8 Conclusion

In this paper, we studied a notion of topological entropy for switched systems. For general switched linear systems, we proved that the topological entropy is independent of the set of initial states, and constructed upper and lower bounds in terms of the measures and the traces of system matrices of individual modes, respectively. For switched linear systems with scalar-valued state and those with pairwise commuting matrices, we established formulae for the topological entropy in terms of the eigenvalues of systems matrices of individual modes. For the more general case with simultaneously triangularizable matrices, we constructed an upper bound for the topological entropy that only depends on the eigenvalues, their order in a simultaneous triangularization, and the active time of each individual mode. In each case above, we also established upper bounds that are more conservative but require less information on the system matrices or on the switching. Furthermore, we presented stability conditions inspired by the upper bounds for topological entropy.

The notion of topological entropy proposed in this paper depends on the switching signal. For switched systems with an uncertain switching signal, a different entropy notion is needed to capture the additional uncertainty about the trajectory and to quantify the extra information needed for stabilization. Sufficient data rates for feedback stabilization of switched linear systems were established in [26, 40]. A similar data-rate bound for state estimation was formulated in [37]. These data-rate bounds should be upper bounds for the entropy notion to be defined.

Another topic for future research is to reconcile the switching characterizations for entropy computation and for control design. More specifically, the entropy computation in this paper is based on the notion of active time (i.e., the accumulated time in which an individual 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. Some preliminary results on entropy-based stability conditions can be found in [42].

Acknowledgements The work of G. Yang and J. P. Hespanha was supported by the Office of Naval Research under the MURI grant N00014-16-1-2710, and by the National Science Foundation under the grant CNS-1329650. The work of D. Liberzon was supported by the National Science Foundation under the grant CMMI-1662708, and by the Air Force Office of Scientific Research under the grant FA9550-17-1-0236. The authors thank Raphaël M. Jungers for his comments on a preliminary version of the paper.

A Proof of Lemma 1

As a direct consequence of the definition (12), the maximal weighted average \bar{a} satisfies

$$\bar{a}(T) \geq \max \left\{ \frac{1}{T} \sum_{p \in \mathcal{P}} a_p \tau_p(T), 0 \right\} = \max \left\{ \sum_{p \in \mathcal{P}} a_p \rho_p(T), 0 \right\} \quad \forall T \geq 0,$$

and thus

$$\limsup_{T \rightarrow \infty} \bar{a}(T) \geq \max\{\hat{a}, 0\}.$$

It remains to prove that the reverse inequality holds as well. The definition (11) of the asymptotic weighted average \hat{a} implies that for each $\delta > 0$, there is a large enough $T'_\delta \geq 0$ such that

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

For a $T > T'_\delta$, let

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

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

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

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

$$\bar{a}(T) = \frac{1}{T} \sum_{p \in \mathcal{P}} a_p \tau_p(t^*(T)) \leq \frac{a_m t^*(T)}{T} \leq \frac{\max\{a_m, 0\} T'_\delta}{T}$$

with the constant $a_m := \max_{p \in \mathcal{P}} a_p$. Combining the two cases above yields

$$\bar{a}(T) \leq \max\{\hat{a} + \delta, \max\{a_m, 0\} T'_\delta / T\} \quad \forall T > T'_\delta.$$

Hence

$$\bar{a}(T) \leq \max\{\hat{a}, 0\} + \delta \quad \forall T > T_\delta$$

with $T_\delta = \max\{T'_\delta, \max\{a_m, 0\} T'_\delta / \delta\}$. As $\delta > 0$ is arbitrary, we have

$$\limsup_{T \rightarrow \infty} \bar{a}(T) \leq \max\{\hat{a}, 0\}.$$

B Proof of Lemma 2

1. As $R(\hat{x}) \subset B_{A_\sigma}(\hat{x}, \varepsilon, T)$ for all $\hat{x} \in G(\theta)$, we have

$$K = \bigcup_{\hat{x} \in G(\theta)} R(\hat{x}) \subset \bigcup_{\hat{x} \in G(\theta)} B_{A_\sigma}(\hat{x}, \varepsilon, T).$$

Then (4) implies that the grid $G(\theta)$ is (T, ε) -spanning, and thus

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

2. For all distinct points $\hat{x}, \hat{x}' \in G(\theta)$, as $\hat{x}' \notin R(\hat{x})$ and $B_{A_\sigma}(\hat{x}, \varepsilon, T) \subset R(\hat{x})$, we have $\hat{x}' \notin B_{A_\sigma}(\hat{x}, \varepsilon, T)$. Then (6) implies that the grid $G(\theta)$ is (T, ε) -separated, and thus

$$\log N(A_\sigma, \varepsilon, T, K) \geq \log |G(\theta)| = \sum_{i=1}^n \log(2 \lfloor 1/\theta_i \rfloor + 1) > \sum_{i=1}^n \log(2/\theta_i - 1).$$

C Proof of Corollary 5

First, the definition (11) of \hat{a} and the subadditivity of limit suprema imply

$$\hat{a} \leq \sum_{p \in \mathcal{P}} \limsup_{t \rightarrow \infty} a_p \rho_p(t) \leq \sum_{p \in \mathcal{P}} \max\{a_p, 0\} \limsup_{t \rightarrow \infty} \rho_p(t) = \sum_{p \in \mathcal{P}} h(a_p) \hat{\rho}_p;$$

which, combined with (27), implies the upper bound (28). For the case where the limits $\lim_{t \rightarrow \infty} \rho_p(t)$ exist and $a_p \geq 0$ for all $p \in \mathcal{P}$, the inequalities in the derivation above becomes equalities due to the additivity of limits and $\max\{a_p, 0\} = a_p$. Second, the definition (11) of \hat{a} implies

$$\hat{a} \leq \limsup_{t \rightarrow \infty} \left(\max_{p \in \mathcal{P}} a_p \right) \sum_{p \in \mathcal{P}} \rho_p(t) = \max_{p \in \mathcal{P}} a_p \leq \max_{p \in \mathcal{P}} h(a_p),$$

which, combined with (27), implies the upper bound (29).

D Proof of Lemma 6

First, we establish the upper bound in (33). 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 $k_s := \sum_{p \in \mathcal{P}} k_p$, which is finite as the index set \mathcal{P} is finite. Define the weighted average matrix over $[0, t]$ by

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

For all $t \geq 0$, as $\{N_p : p \in \mathcal{P}\}$ is a commuting family, we have $(N(t))^{k_s} = 0$. Also, $\|N(t)\| \leq \max_{p \in \mathcal{P}} \|N_p\| =: \eta_M$. Hence for all $v \in \mathbb{C}^n$,

$$\|e^{N(t)t}v\| = \left\| \left(\sum_{k=0}^{k_s-1} \frac{(N(t))^k t^k}{k!} \right) v \right\| \leq \left(\sum_{k=0}^{k_s-1} \frac{\eta_M^k t^k}{k!} \right) \|v\| \leq c_\delta \left(\sum_{k=0}^{k_s-1} \frac{\delta^k t^k}{k!} \right) \|v\| \leq c_\delta e^{\delta t} \|v\|$$

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

Second, we establish the lower bound in (33). As $\| -N(t) \| = \| N(t) \| \leq \eta_M$ for all $t \geq 0$, the proof above also implies that for all $v \in \mathbb{C}^n$,

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

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

E Calculation of $h(D_{\sigma_2})$ in Example 3

Recall from footnote 2 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

$$\begin{cases} \tau_1(t) = t - 0.9t_{2k}, & \tau_2(t) = 0.9t_{2k}, & t \in [t_{2k}, t_{2k+1}), \\ \tau_1(t) = 0.9t_{2k+1}, & \tau_2(t) = t - 0.9t_{2k+1}, & t \in [t_{2k+1}, t_{2k+2}), \end{cases} \quad (60)$$

and thus

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

Then \bar{a}_1 and \bar{a}_2 in (31) 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

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

Therefore,

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

F Proof of Lemma 7

We regard (40) as a family of scalar differential equations (recall that here ξ_σ^k denotes the k -th scalar component of ξ_σ)

$$\begin{aligned}\dot{\xi}_\sigma^1 &= a_\sigma^1 \xi_\sigma^1 + b_\sigma^{1,2} \xi_\sigma^2 + \cdots + b_\sigma^{1,n} \xi_\sigma^n, \\ \dot{\xi}_\sigma^2 &= a_\sigma^2 \xi_\sigma^2 + b_\sigma^{2,3} \xi_\sigma^3 + \cdots + b_\sigma^{2,n} \xi_\sigma^n, \\ &\vdots \\ \dot{\xi}_\sigma^{n-1} &= a_\sigma^{n-1} \xi_\sigma^{n-1} + b_\sigma^{n-1,n} \xi_\sigma^n, \\ \dot{\xi}_\sigma^n &= a_\sigma^n \xi_\sigma^n,\end{aligned}$$

and prove Lemma 7 by mathematical induction. For brevity, let

$$\psi_{i,j}(t) := b_{\sigma(t)}^{i,j} e^{\eta_j(t) - \eta_i(t)}, \quad i, j \in \{1, \dots, n\}.$$

Then Ψ in (44) can be written as

$$\Psi(t, \mathcal{C}_{k,l,i}) = \sum_{(c_0, \dots, c_i) \in \mathcal{C}_{k,l,i}} \int_0^t \int_0^{s_1} \cdots \int_0^{s_{i-1}} \prod_{j=1}^i (\psi_{c_{j-1}, c_j}(s_j) ds_j). \quad (61)$$

F.1 The basis of induction

For the n -th scalar differential equation $\dot{\xi}_\sigma^n = a_\sigma^n \xi_\sigma^n$, the state-transition function is defined by

$$\phi_n(t, s) := e^{\eta_n(t) - \eta_n(s)}, \quad t \geq s \geq 0.$$

Hence the n -th scalar component of $\xi_\sigma(x, t)$ satisfies $\xi_\sigma^n(x, t) = e^{\eta_n(t)} x_n$, that is, (44) holds for $k = n$.

F.2 The inductive step

For an arbitrary $m \in \{1, \dots, n-1\}$, suppose that $\xi_\sigma^k(x, t)$ satisfy (44) for all $k \in \{m+1, \dots, n\}$. For the m -th differential equation

$$\dot{\xi}_\sigma^m = a_\sigma^m \xi_\sigma^m + \sum_{k=m+1}^n b_\sigma^{m,k} \xi_\sigma^k,$$

the state-transition function is defined by

$$\phi_m(t, s) := e^{\eta_m(t) - \eta_m(s)}, \quad t \geq s \geq 0.$$

By variation of constants, the m -th scalar component of $\xi_\sigma(x, t)$ satisfies

$$\begin{aligned}\xi_\sigma^m(x, t) &= e^{\eta_m(t)} \left(x_m + \sum_{k=m+1}^n \int_0^t e^{-\eta_m(s_1)} b_{\sigma(s_1)}^{m,k} \xi_\sigma^k(x, s_1) ds_1 \right) \\ &= e^{\eta_m(t)} \left(x_m + \sum_{k=m+1}^n \int_0^t \psi_{m,k}(s_1) \left(x_k + \sum_{l=k+1}^n \sum_{i=1}^{l-k} x_l \Psi(s_1, \mathcal{C}_{k,l,i}) \right) ds_1 \right) \\ &= e^{\eta_m(t)} \left(x_m + \sum_{k=m+1}^n x_k \int_0^t \psi_{m,k}(s_1) ds_1 + \sum_{k=m+1}^n \sum_{l=k+1}^n \sum_{i=1}^{l-k} x_l \int_0^t \psi_{m,k}(s_1) \Psi(s_1, \mathcal{C}_{k,l,i}) ds_1 \right).\end{aligned}$$

Based on the definition (46) of $\mathcal{C}_{k,l,i}$ and the formula (61) of Ψ , we have

$$\int_0^t \psi_{m,k}(s_1) ds_1 = \Psi(t, \mathcal{C}_{m,k,1})$$

and

$$\begin{aligned}
& \int_0^t \psi_{m,k}(s_1) \Psi(s_1, \mathcal{C}_{k,l,i}) \, ds_1 \\
&= \sum_{(c_1, \dots, c_{i+1}) \in \mathcal{C}_{k,l,i}} \int_0^t \int_0^{s_1} \cdots \int_0^{s_i} \psi_{m,k}(s_1) \prod_{j=2}^{i+1} (\psi_{c_{j-1}, c_j}(s_j) \, ds_j) \, ds_1 \\
&= \sum_{(c_0, \dots, c_{i+1}) \in \{m\} \times \mathcal{C}_{k,l,i}} \int_0^t \int_0^{s_1} \cdots \int_0^{s_i} \prod_{j=1}^{i+1} (\psi_{c_{j-1}, c_j}(s_j) \, ds_j) \\
&= \Psi(t, \{m\} \times \mathcal{C}_{k,l,i}).
\end{aligned}$$

Changing the order of summation, we obtain

$$\begin{aligned}
\sum_{k=m+1}^n \sum_{l=k+1}^n \sum_{i=1}^{l-k} x_l \Psi(t, \{m\} \times \mathcal{C}_{k,l,i}) &= \sum_{l=m+2}^n \sum_{k=m+1}^{l-1} \sum_{i=1}^{l-k} x_l \Psi(t, \{m\} \times \mathcal{C}_{k,l,i}) \\
&= \sum_{l=m+2}^n \sum_{i'=2}^{l-m} \sum_{k=m+1}^{l-i'+1} x_l \Psi(t, \{m\} \times \mathcal{C}_{k,l,i'-1}),
\end{aligned}$$

where in the last step we also let $i' = i + 1$. Next, we prove that the family of sets $\{\{m\} \times \mathcal{C}_{k,l,i'-1} : k = m+1, \dots, l-i'+1\}$ forms a partition of $\mathcal{C}_{m,l,i'}$.

- For all $(c_1, \dots, c_{i'}) \in \mathcal{C}_{k_1,l,i'-1}$ and $(c'_1, \dots, c'_{i'}) \in \mathcal{C}_{k_2,l,i'-1}$ with $k_1 \neq k_2$, as $c_1 = k_1 \neq k_2 = c'_1$, we have $(c_1, \dots, c_{i'}) \neq (c'_1, \dots, c'_{i'})$. Hence the sets in $\{\{m\} \times \mathcal{C}_{k,l,i'-1} : k = m+1, \dots, l-i'+1\}$ are pairwise disjoint.
 - For all $(c_1, \dots, c_{i'}) \in \mathcal{C}_{k,l,i'-1}$, as $c_1 = k \geq m+1$ and $c_{i'} = l$, we have $(m, c_1, \dots, c_{i'}) \in \mathcal{C}_{m,l,i'}$. Hence $\bigcup_{k=m+1}^{l-i'+1} \{m\} \times \mathcal{C}_{k,l,i'-1} \subset \mathcal{C}_{m,l,i'}$.
 - For all $(c_0, \dots, c'_i) \in \mathcal{C}_{m,l,i'}$, as $c_1 \geq c_0 + 1 = m+1$ and $c_1 \leq c_{i'} - (i' - 1) = l - i' + 1$, we have $k := c_1$ satisfies $m+1 \leq k \leq l-i'+1$ and $(c_0, \dots, c_{i'}) \in \{m\} \times \mathcal{C}_{k,l,i'-1}$. Hence $\mathcal{C}_{m,l,i'} \subset \bigcup_{k=m+1}^{l-i'+1} \{m\} \times \mathcal{C}_{k,l,i'-1}$.
- Therefore,

$$\sum_{l=m+2}^n \sum_{i'=2}^{l-m} \sum_{k=m+1}^{l-i'+1} x_l \Psi(t, \{m\} \times \mathcal{C}_{k,l,i'-1}) = \sum_{l=m+2}^n \sum_{i'=2}^{l-m} x_l \Psi(t, \mathcal{C}_{m,l,i'}).$$

Combining the results above, we obtain

$$\xi_\sigma^m(x, t) = e^{\eta_m(t)} \left(x_m + \sum_{l=m+1}^n \sum_{i=1}^{l-m} x_l \Psi(t, \mathcal{C}_{m,l,i}) \right),$$

that is, (44) holds for $k = m$. Therefore, mathematical induction implies that (44) holds for all $k \in \{1, \dots, n\}$.

G Proof of Lemma 8

For every $k \in \{1, \dots, n\}$, following the formula (44), the triangle inequality, and the definition of b_M , the k -th scalar component of $\xi_\sigma(x, t)$ satisfies

$$|\xi_\sigma^k(x, t)| \leq e^{\operatorname{Re}(\eta_k(t))} \left(|x_k| + \sum_{l=k+1}^n \sum_{i=1}^{l-k} |x_l| |\Psi(t, \mathcal{C}_{k,l,i})| \right),$$

in which

$$\begin{aligned}
|\Psi(t, \mathcal{C}_{k,l,i})| &\leq \sum_{(c_0, \dots, c_i) \in \mathcal{C}_{k,l,i}} \int_0^t \int_0^{s_1} \cdots \int_0^{s_{i-1}} \prod_{j=1}^i (b_M e^{\operatorname{Re}(\eta_{c_j}(s_j) - \eta_{c_{j-1}}(s_j))} ds_j) \\
&\leq \sum_{(c_0, \dots, c_i) \in \mathcal{C}_{k,l,i}} \prod_{j=1}^i \left(b_M t \max_{s_j \in [0, T]} e^{\operatorname{Re}(\eta_{c_j}(s_j) - \eta_{c_{j-1}}(s_j))} \right) \\
&= \sum_{(c_0, \dots, c_i) \in \mathcal{C}_{k,l,i}} b_M^i t^i e^{\sum_{j=1}^i \bar{d}_{c_j}(t) t} \\
&\leq \sum_{(c_0, \dots, c_i) \in \mathcal{C}_{k,l,i}} b_M^i t^i e^{\sum_{j=k+1}^l \bar{d}_j(t) t},
\end{aligned}$$

where the equality follows from the definitions (43) and (45) of \bar{d}_i and η_i , and the last inequality from the definition (46) of the sets $\mathcal{C}_{k,l,i}$. As $b_M^i t^i$ is independent of the choice of $(c_0, \dots, c_i) \in \mathcal{C}_{k,l,i}$, and $\sum_{j=k+1}^l \bar{d}_j(t) t$ is independent of the choices of $(c_0, \dots, c_i) \in \mathcal{C}_{k,l,i}$ and $i \in \{1, \dots, l-k\}$, we have

$$\sum_{l=k+1}^n \sum_{i=1}^{l-k} |x_l| |\Psi(t, \mathcal{C}_{k,l,i})| = \sum_{l=k+1}^n |x_l| \left(\sum_{i=1}^{l-k} |\mathcal{C}_{k,l,i}| b_M^i t^i \right) e^{\sum_{j=k+1}^l \bar{d}_j(t) t},$$

in which

$$\sum_{i=1}^{l-k} |\mathcal{C}_{k,l,i}| b_M^i t^i = \sum_{i=1}^{l-k} \binom{l-k-1}{i-1} b_M^i t^i = (b_M t + 1)^{l-k-1} b_M t \leq (b_M t + 1)^{l-k},$$

where the first equality follows from fact that the set $\mathcal{C}_{k,l,i}$ can be characterized by the combinations of $i-1$ increasing integers from $k+1$ to $l-1$ (in particular, in the definition (46) of $\mathcal{C}_{k,l,i}$, the first and last points $c_0 = k$ and $c_i = l$ are fixed), and the second one from the binomial formula. Hence

$$\begin{aligned}
|\xi_\sigma^k(x, t)| &\leq e^{\operatorname{Re}(\eta_k(t))} \left(|x_k| + \sum_{l=k+1}^n \left((b_M t + 1)^{l-k} e^{\sum_{j=k+1}^l \bar{d}_j(t) t} |x_l| \right) \right) \\
&\leq e^{\operatorname{Re}(\eta_1(t))} \sum_{l=k}^n \left((b_M t + 1)^{l-k} e^{\sum_{j=2}^l \bar{d}_j(t) t} |x_l| \right),
\end{aligned}$$

where the last inequality follows partially from the definitions (43) and (45) of \bar{d}_i and η_i . Note that the upper bound for $|\xi_\sigma^k(x, t)|$ above is nonincreasing in k . Indeed, the upper bound for $|\xi_\sigma^{k-1}(x, t)|$ satisfies

$$e^{\operatorname{Re}(\eta_1(t))} \sum_{l=k-1}^n \left((b_M t + 1)^{l-k+1} e^{\sum_{j=2}^l \bar{d}_j(t) t} |x_l| \right) \geq e^{\operatorname{Re}(\eta_1(t))} \sum_{l=k}^n \left((b_M t + 1)^{l-k} e^{\sum_{j=2}^l \bar{d}_j(t) t} |x_l| \right).$$

Hence we obtain (47) by taking the upper bound for $|\xi_\sigma^1(x, t)|$ (recall that we take $\|\cdot\|$ to be the ∞ -norm; see Remark 1).

H Calculation of $h(U_{\sigma_2})$ in Example 4

Following (60) in Appendix E, we have

$$\begin{aligned}
a_1^1 \tau_1(t) + a_2^1 \tau_2(t) &= 3\tau_2(t) - \tau_1(t) \\
&= \begin{cases} 3.6t_{2k} - t, & t \in [t_{2k}, t_{2k+1}), \\ 3t - 3.6t_{2k+1}, & t \in [t_{2k+1}, t_{2k+2}) \end{cases} \\
(a_1^2 - a_1^1) \tau_1(t) + (a_2^2 - a_2^1) \tau_2(t) &= 3\tau_1(t) - 4\tau_2(t) \\
&= \begin{cases} 3t - 6.3t_{2k}, & t \in [t_{2k}, t_{2k+1}), \\ 6.3t_{2k+1} - 4t, & t \in [t_{2k+1}, t_{2k+2}). \end{cases}
\end{aligned}$$

Then \bar{a}_1 and \bar{d}_2 in (41) satisfy

$$\begin{aligned}\bar{a}_1(T) &= \frac{1}{T} \max_{t \in [0, T]} a_1^1 \tau_1(t) + a_2^1 \tau_2(t) \\ &= \begin{cases} 2.6t_{2k}/T & T \in [t_{2k}, t_{2k+1} + 8t_{2k}/3); \\ 3 - 3.6t_{2k+1}/T, & T \in [t_{2k+1} + 8t_{2k}/3, t_{2k+2}) \end{cases} \\ \bar{d}_2(T) &= \frac{1}{T} \max_{t \in [0, T]} (a_1^2 - a_1^1) \tau_1(t) + (a_2^2 - a_2^1) \tau_2(t) \\ &= \begin{cases} 2.3t_{2k+1}/T, & T \in [t_{2k+1}, 2t_{2k+2} + 5t_{2k+1}/3); \\ 3 - 6.3t_{2k+2}/T, & T \in [2t_{2k+2} + 5t_{2k+1}/3, t_{2k+3}). \end{cases}\end{aligned}$$

Hence

$$2\bar{a}_1(T) + \bar{d}_2(T) = \begin{cases} 25.9t_{2k}/T, & T \in [t_{2k+1}, t_{2k+1} + 8t_{2k}/3); \\ 6 - 4.9t_{2k+1}/T, & T \in [t_{2k+1} + 8t_{2k}/3, t_{2k+2}); \\ 49.1t_{2k+1}/T, & T \in [t_{2k+2}, 2t_{2k+2} + 5t_{2k+1}/3); \\ 3 - 1.1t_{2k+2}/T, & T \in [2t_{2k+2} + 5t_{2k+1}/3, t_{2k+3}). \end{cases}$$

Therefore,

$$h(U_{\sigma_2}) = \limsup_{T \rightarrow \infty} 2\bar{a}_1(T) + \bar{d}_2(T) = \max\{2.88, 5.46\} = 5.46.$$

References

1. Adler RL, Konheim AG, McAndrew MH (1965) Topological entropy. *Transactions of the American Mathematical Society* 114(2):309–319
2. Agrachev AA, Liberzon D (2001) Lie-algebraic stability criteria for switched systems. *SIAM Journal on Control and Optimization* 40(1):253–269
3. Agrachev AA, Baryshnikov Y, Liberzon D (2012) On robust Lie-algebraic stability conditions for switched linear systems. *Systems & Control Letters* 61(2):347–353
4. Bowen R (1971) Entropy for group endomorphisms and homogeneous spaces. *Transactions of the American Mathematical Society* 153:401–414
5. Bowen R (1971) Periodic points and measures for axiom A diffeomorphisms. *Transactions of the American Mathematical Society* 154:377–397
6. Colonius F (2012) Minimal bit rates and entropy for exponential stabilization. *SIAM Journal on Control and Optimization* 50(5):2988–3010
7. Colonius F, Kawan C (2009) Invariance entropy for control systems. *SIAM Journal on Control and Optimization* 48(3):1701–1721
8. Colonius F, Kawan C, Nair GN (2013) A note on topological feedback entropy and invariance entropy. *Systems & Control Letters* 62(5):377–381
9. Cover TM, Thomas JA (2005) *Elements of Information Theory*, 2nd edn. John Wiley & Sons
10. Desoer CA, Vidyasagar M (2009) *Feedback Systems: Input-Output Properties*. SIAM
11. Dinaburg EI (1970) The relation between topological entropy and metric entropy. *Doklady Akademii Nauk SSSR* 190(1):19–22, in Russian
12. Drazin MP, Dungey JW, Gruenberg KW (1951) Some theorems on commutative matrices. *Journal of the London Mathematical Society* 26(3):221–228
13. Gurvits L (1995) Stability of discrete linear inclusion. *Linear Algebra and its Applications* 231:47–85
14. Haimovich H, Braslavsky JH (2013) Sufficient conditions for generic feedback stabilizability of switching systems via Lie-algebraic solvability. *IEEE Transactions on Automatic Control* 58(3):814–820
15. Hespanha JP, Ortega A, Vasudevan L (2002) Towards the control of linear systems with minimum bit-rate. In: 15th International Symposium on Mathematical Theory of Networks and Systems, pp 1–15
16. Horn RA, Johnson CR (2013) *Matrix Analysis*, 2nd edn. Cambridge University Press
17. Humphreys JE (1972) *Introduction to Lie Algebras and Representation Theory*. Springer
18. Katok A, Hasselblatt B (1995) *Introduction to the Modern Theory of Dynamical Systems*. Cambridge University Press
19. Kawan C, Latushkin Y (2016) Some results on the entropy of non-autonomous dynamical systems. *Dynamical Systems* 31(3):251–279
20. Kolmogorov AN (1958) A new metric invariant of transitive dynamical systems and automorphisms of Lebesgue spaces. *Doklady Akademii Nauk SSSR* 119(5):861–864, in Russian
21. Kolyada S, Snoha L (1996) Topological entropy of nonautonomous dynamical systems. *Random & Computational Dynamics* 4(2-3):205–233
22. Kutepov SA (1982) Algebraic criterion of absolute stability. *Kibernetika i Vychislitel'naia Tekhnika* 54:52–55, in Russian
23. Kutepov SA (1984) Absolute stability of bilinear systems with a compact Levi factor. *Kibernetika i Vychislitel'naia Tekhnika* 62(118):28–33, in Russian

24. Laffey TJ (1978) Simultaneous triangularization of matrices—low rank cases and the nonderogatory case. *Linear and Multilinear Algebra* 6(4):269–305
25. Liberzon D (2003) *Switching in Systems and Control*. Birkhäuser
26. Liberzon D (2014) Finite data-rate feedback stabilization of switched and hybrid linear systems. *Automatica* 50(2):409–420
27. Liberzon D, Mitra S (2018) Entropy and minimal bit rates for state estimation and model detection. *IEEE Transactions on Automatic Control* 63(10):3330–3344
28. Liberzon D, Hespanha JP, Morse AS (1999) Stability of switched systems: A Lie-algebraic condition. *Systems & Control Letters* 37(3):117–122
29. McCoy NH (1936) On the characteristic roots of matrix polynomials. *Bulletin of the American Mathematical Society* 42(8):592–601
30. Nair GN, Evans RJ (2003) Exponential stabilisability of finite-dimensional linear systems with limited data rates. *Automatica* 39(4):585–593
31. Nair GN, Evans RJ, Mareels IMY, Moran W (2004) Topological feedback entropy and nonlinear stabilization. *IEEE Transactions on Automatic Control* 49(9):1585–1597
32. Ornstein DS (1974) *Ergodic Theory, Randomness, and Dynamical Systems*. Yale University Press
33. Savkin AV (2006) Analysis and synthesis of networked control systems: topological entropy, observability, robustness and optimal control. *Automatica* 42(1):51–62
34. Schmidt AJ (2016) Topological entropy bounds for switched linear systems with Lie structure. M.s. thesis, University of Illinois at Urbana-Champaign
35. Shannon CE (1948) A mathematical theory of communication. *The Bell System Technical Journal* 27(3):379–423
36. Shorten R, Wirth FR, Mason O, Wulff K, King C (2007) Stability criteria for switched and hybrid systems. *SIAM Review* 49(4):545–592
37. Sibai H, Mitra S (2017) Optimal data rate for state estimation of switched nonlinear systems. In: 20th ACM International Conference on Hybrid Systems: Computation and Control, pp 71–80
38. Tatikonda S, Mitter S (2004) Control under communication constraints. *IEEE Transactions on Automatic Control* 49(7):1056–1068
39. Yang G, Hespanha JP (2018) On topological entropy of switched linear systems with pairwise commuting matrices. In: 56th Annual Allerton Conference on Communication, Control, and Computing, pp 429–436
40. Yang G, Liberzon D (2018) Feedback stabilization of a switched linear system with an unknown disturbance under data-rate constraints. *IEEE Transactions on Automatic Control* 63(7):2107–2122
41. Yang G, Schmidt AJ, Liberzon D (2018) On topological entropy of switched linear systems with diagonal, triangular, and general matrices. In: 57th IEEE Conference on Decision and Control, pp 5682–5687
42. Yang G, Hespanha JP, Liberzon D (2019) On topological entropy and stability of switched linear systems. In: 22nd ACM International Conference on Hybrid Systems: Computation and Control, pp 119–127