Topological entropy of switched nonlinear and interconnected systems

Guosong Yang, Daniel Liberzon, and João P. Hespanha

Abstract

This paper studies topological entropy of switched nonlinear and interconnected systems. We construct a general upper bound for the topological entropy of a switched nonlinear system, in terms of an average of the asymptotic suprema of measures of Jacobian matrices of system functions of individual modes, weighted by their active rates. We also construct a general lower bound in a similar form, in terms of the asymptotic infima of traces of these Jacobian matrices. The general upper bound is readily used to construct an upper bound for entropy in the interconnected case. We also construct less conservative upper bounds for entropy in diagonal and block-diagonal cases, as well as more conservative upper bounds that require less information on the switching in all cases. The upper bounds for entropy and their relations are illustrated by numerical examples of a switched Lotka–Volterra ecosystem model.

I. Introduction

In systems theory, topological entropy describes the information generation rate of a dynamical system in terms of the number of separable trajectories for a finite precision, or the complexity growth rate of a system acting on a set with finite measure. Adler et al. first defined topological entropy as an extension of Kolmogorov's metric entropy [1], quantifying the expansion of a map via the minimal cardinality of a subcover refinement [2]. A different definition in terms of the number of separable trajectories for a finite precision was introduced by Bowen [3] and independently by Dinaburg [4]. 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 methods to understand [6], [7]. This work on topological entropy of switched systems aims at contributing to our understanding of these complexities.

Entropy is also a fundamental notion in control theory. Topological feedback entropy was first defined for discrete-time systems in [8] following the construction in [2], which extends classical entropy concepts and describes the growth rate of control complexity for set invariance. A definition of invariance entropy was later introduced for continuous-time systems in [9], which is conceptually closer to the trajectory-counting formulation in [3], [4]. An equivalence between these two definitions was established in [10]. Entropy definitions have also been proposed and studied for stabilization [11], state estimation [12]–[14] and model detection [14].

This paper studies topological entropy of continuous-time switched nonlinear systems. Switched systems have been a popular topic in recent years (see, e.g., [15], [16] and references therein). In general, a switched system does not inherit stability properties of its individual modes. For example, a switched system with two stable modes may still be unstable [15, p. 19]. Meanwhile, 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., [15, Th. 2.5, p. 31]). This result has been generalized to global uniform asymptotic stability for switching nonlinear systems with pairwise commuting, globally asymptotically stable modes [17], [18]. A simplest case of pairwise commuting modes is when the system functions are simultaneously diagonalizable, which motivates us to study switched systems with diagonal modes in addition to the general case.

The study of interconnected systems plays a significant role in systems theory, as it allows one to represent a complex system as an interconnection of simpler subsystems, and establish properties of the overall system by analyzing the individual components. A standard example in this context is the use of small-gain theorems in establishing stability of interconnected linear, nonlinear, and switched systems (see, e.g., [19]–[21]).

Our interest in topological entropy of switched systems is strongly motivated by its characterization of data-rate requirements for control problems. For a linear time-invariant control system, the minimal data rate for feedback

G. Yang and J. P. Hespanha are with the Center for Control, Dynamical Systems, and Computation, University of California, Santa Barbara, Santa Barbara, CA 93106 USA (e-mail: {guosongyang,hespanha}@ucsb.edu).

D. Liberzon is with the Coordinated Science Laboratory, University of Illinois Urbana–Champaign, Urbana, IL 61801 USA (e-mail: liberzon@illinois.edu).

stabilization is given by the sum of the positive real parts of eigenvalues of the system matrix [22] (or, for discrete-time systems, the sum of their logarithms [22]–[24]), which is equal to the topological entropy in open-loop [3], [9]. Data-rate conditions and entropy definitions for nonlinear and interconnected time-invariant control systems have been studied in [8], [11], [25]–[28]. For switched systems, however, neither the minimal data rate nor the topological entropy are completely understood. Sufficient data rates for feedback stabilization of switched linear systems were established in [29], [30]. Similar data-rate conditions were constructed in [31] by extending the estimation entropy from [14]. In [32], [33], formulae and bounds for topological entropy of switched linear systems were constructed in terms of the active rates of individual modes, an approach that is also adopted in the current work. Relations between topological entropy and stability of switched linear systems were studied in [33], [34]. For discrete-time switched linear systems, the topological entropy under worse-case switching sequences was obtained based on joint spectral radius [35], while a formula for estimation entropy was derived under additional regularity conditions [36].

For switched nonlinear systems, topological entropy has not been explored closely so far. This paper's main contribution is to construct upper and lower bounds for topological entropy of switched nonlinear systems with general and special structures, which generalizes and extends previous results for switched linear systems from [32], [33]. We develop the necessary preliminaries in Section II, including basic definitions for topological entropy of switched systems, a universal construction of spanning and separated sets, and switching-related quantities such as active rates of individual modes. In Section III, we establish upper and lower bounds for the distance between two solutions and a lower bound for the volume of a reachable set, which are essential to the construction of bounds for topological entropy, and are also of independent interest.

In Section IV, we construct a general upper bound for the topological entropy of a switched nonlinear system, in terms of an asymptotic average of the measures of Jacobian matrices of system functions of individual modes. The matrix measures are maximized over the ω-limit set, and the average is weighted by the active rates. We also construct a general lower bound in a similar form, in terms of the traces of Jacobian matrices of system functions of individual modes, minimized over the ω-limit set. In Section V, we consider special cases with diagonal, interconnected, and block-diagonal structure. We show in Section V-B that the general upper bound for entropy from Section IV can be readily used to construct an upper bound for entropy in the interconnected case that only depend on "network-level" information, and construct upper bounds for entropy in the diagonal and block-diagonal cases in Sections V-A and V-C that are less conservative than or equivalent to those in Sections IV and V-B, respectively. For each general or special case, we also construct upper bounds for entropy that are more conservative but require less information about the switching signal, with their relations illustrated by numerical examples of a switched Lotka–Volterra ecosystem model in Section VI. Section VII concludes the paper with a brief summary and some remarks on future research directions.

A preliminary version of some of the results in this paper was presented in the paper [37]. The current paper improves upon [37] by considering also the interconnected and block-diagonal cases, and providing more extensive numerical examples, analysis details, and explanatory remarks.

Notations: Denote by $\mathbf{1}_n$ the vector of ones in $\mathbb{R}^n$, and by I_n and $\mathbf{1}_{n\times n}$ the identity matrix and matrix of ones in $\mathbb{R}^{n\times n}$, respectively; the subscript is omitted when the dimension is clear from context. Denote by $\mathrm{Re}(c)$ the real part of a complex number $c \in \mathbb{C}$. Denote by $(v_1, \dots, v_k) := [v_1^\top \cdots v_k^\top]^\top$ the concatenation of vectors $v_1 \in \mathbb{R}^{n_1}, \dots, v_k \in \mathbb{R}^{n_k}$. Denote by $\mathrm{tr}(A)$, $\det(A)$, and $\mathrm{spec}(A)$ the trace, determinant, and spectrum (as a multiset in which each eigenvalue has a number of instances equal to its algebraic multiplicity) of a matrix $A \in \mathbb{R}^{n\times n}$, respectively. Denote by $\#E$ the cardinality of a finite set E. Denote by $\mathrm{vol}(S)$ and $\mathrm{co}(S)$ the volume (Lebesgue measure) and convex hull of a set $S \subset \mathbb{R}^n$, respectively. Denote by $|v|_\infty = \max_{1\le i\le n} |v_i|$ the ∞-norm of a vector $v = (v_1, \dots, v_n) \in \mathbb{R}^n$, and by $\|A\|_\infty = \max_{1\le i\le n} \sum_{j=1}^n |a_{ij}|$ the induced ∞-norm of a matrix $A = [a_{ij}] \in \mathbb{R}^{n\times n}$. An inequality between two vectors or matrices of the same size, or between a vector or matrix and a scalar, is to be interpreted entrywise (e.g., $A \ge 0$ means that A is a nonnegative matrix). By default, all logarithms are natural logarithms (in order to avoid generating an extra multiplicative factor $\ln 2$ when computing entropy).

II. Preliminaries

Consider a finite family of continuous-time nonlinear dynamical systems

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

with the state $x \in \mathbb{R}^n$, in which each system is labeled with an index p, and $\mathcal{P}$ is the finite set of these indices. We assume that the functions $f_p : \mathbb{R}^n \to \mathbb{R}^n$ are continuously differentiable, and the systems in (1) are forward complete. We are interested in the corresponding *switched system* defined by

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

with a right-continuous, piecewise-constant *switching signal* $\sigma : \mathbb{R}_{\geq 0} \to \mathcal{P}$. We call the system with index p in (1) *mode* p of the switched system (2), and $\sigma(t)$ the *active mode* at time t. We denote by $\xi_\sigma(t, x)$ the solution to (2) with switched signal σ at time t with initial state x. Under the assumptions above, $\xi_\sigma(t, x)$ is unique, absolutely continuous in t, continuously differentiable in x, 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 in each finite time interval (i.e., the set of switches contains no accumulation point).

A. Entropy definitions

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

$$B_{f_\sigma}(x, \varepsilon, T) := \Big\{ \bar{x} \in \mathbb{R}^n : \max_{t\in[0,T]} |\xi_\sigma(t, \bar{x}) - \xi_\sigma(t, x)| < \varepsilon \Big\}. \tag{3}$$

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

$$K \subset \bigcup_{x\in E} B_{f_\sigma}(x, \varepsilon, T), \tag{4}$$

or equivalently, for each $\bar{x} \in K$, there is a point $x \in E$ such that $|\xi_\sigma(t, \bar{x}) - \xi_\sigma(t, x)| < \varepsilon$ for all $t \in [0, T]$. We denote by $S(f_\sigma, \varepsilon, T, K) \geq 1$ 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 switching signal σ and initial set K is defined in terms of the exponential growth rate of $S(f_\sigma, \varepsilon, T, K)$ by

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

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

Remark 1. In view of the equivalence of norms on a finite-dimensional vector space, the value of $h(f_\sigma, K)$ is the same for every norm $|\cdot|$ on $\mathbb{R}^n$. In particular, it is invariant under a change of basis. (More generally, one can define topological entropy on a metric space (X, d) instead of the normed space $(\mathbb{R}^n, |\cdot|)$, in which case its value depends on the given metric, but is a topological invariant for an initial set contained in a compact positively invariant set; see [38, Prop. 3.1.2, p. 109] and [9, p. 1703] for more details.) For convenience and concreteness, we take $|\cdot|$ to be the ∞-norm on $\mathbb{R}^n$ and $\|\cdot\|$ the induced ∞-norm on $\mathbb{R}^{n\times n}$ unless otherwise specified.

Next, we provide an equivalent definition for the entropy of (2). With $T \geq 0$ and $\varepsilon > 0$ given as before, we say that a finite set $E \subset K$ is (T, ε)-*separated* if

$$\bar{x} \notin B_{f_\sigma}(x, \varepsilon, T) \qquad \forall\, x, \bar{x} \in E : \bar{x} \neq x, \tag{6}$$

or equivalently, for each pair of distinct points $x, \bar{x} \in E$, there is a time $t \in [0, T]$ such that $|\xi_\sigma(t, \bar{x}) - \xi_\sigma(t, x)| \geq \varepsilon$. We denote by $N(f_\sigma, \varepsilon, T, K) \geq 1$ 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 defined in terms of the exponential growth rate of $N(f_\sigma, \varepsilon, T, K)$; the proof is along the lines of [38, p. 110] and thus omitted here.

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

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

Remark 2. Following [38, pp. 109–110], for a time-invariant system $\dot{x} = f(x)$ and an initial set K contained in a compact positively invariant set, the value of $h(f, K)$ remains the same when the upper limits in (5) and (7) are replaced with lower limits. However, this is not necessarily true for the switched system (2), since the subadditivity required in the proof of [38, Lemma 3.1.5, p. 109] may not hold.

B. Universal spanning and separated sets

Given a time horizon $T \geq 0$ and a radius $\varepsilon > 0$, we provide a universal construction of (T, ε)-spanning and (T, ε)-separated sets by extending a notion of grid from [33]. Given a vector $\theta = (\theta_1, \ldots, \theta_n) \in \mathbb{R}^n_{>0}$ which may depend on T and ε, we define the following grid on the initial set K:

$$G(\theta) := \{(k_1\theta_1, \ldots, k_n\theta_n) \in K : k_1, \ldots, k_n \in \mathbb{Z}\}. \tag{8}$$

As K is a compact set with nonempty interior, there exist closed hypercubes B_1 with radius $r_1 > 0$ and B_2 with radius $r_2 > 0$ such that $B_1 \subset K \subset B_2$. Then the cardinality of the grid $G(\theta)$ satisfies

$$\prod_{i=1}^{n} \left\lfloor \frac{2r_1}{\theta_i} \right\rfloor \leq \#G(\theta) \leq \prod_{i=1}^{n} \left(\left\lfloor \frac{2r_2}{\theta_i} \right\rfloor + 1 \right).$$

For a point $x = (x_1, \ldots, x_n) \in G(\theta)$, let $R(x)$ be the open hyperrectangle in $\mathbb{R}^n$ with center x and sides $2\theta_1, \ldots, 2\theta_n$, that is,

$$R(x) := \{\bar{x}_1, \ldots, \bar{x}_n) \in \mathbb{R}^n : |\bar{x}_1 - x_1| < \theta_1, \ldots, |\bar{x}_n - x_n| < \theta_n\}. \tag{9}$$

Then the points in $G(\theta)$ adjacent to x are on the boundary of $R(x)$, and the union of all $R(x)$ covers K, that is,

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

Comparing the hyperrectangle $R(x)$ to the open ball $B_{f_\sigma}(x, \varepsilon, T)$ defined by (3), we obtain the following result, which extends [33, Lemma 2]; see Appendix A for the proof.

Lemma 2. *1) If the vector θ is selected so that*

$$R(x) \subset B_{f_\sigma}(x, \varepsilon, T) \qquad \forall\, x \in G(\theta), \tag{10}$$

then the grid $G(\theta)$ is (T, ε)-spanning. If (10) holds for all $T \geq 0$ and $\varepsilon > 0$, and

$$\limsup_{\varepsilon \searrow 0} \limsup_{T \to \infty} \frac{\log \theta_i}{T} \leq 0 \qquad \forall\, i \in \{1, \ldots, n\}, \tag{11}$$

then the topological entropy of the switched system (2) is upper bounded by

$$h(f_\sigma, K) \leq \limsup_{\varepsilon \searrow 0} \limsup_{T \to \infty} \sum_{i=1}^{n} \frac{\log(1/\theta_i)}{T}. \tag{12}$$

2) If the vector θ is selected so that

$$B_{f_\sigma}(x, \varepsilon, T) \subset R(x) \qquad \forall\, x \in G(\theta), \tag{13}$$

then the grid $G(\theta)$ is (T, ε)-separated. If (13) holds for all $T \geq 0$ and $\varepsilon > 0$, then the topological entropy of the switched system (2) is lower bounded by

$$h(f_\sigma, K) \geq \liminf_{\varepsilon \searrow 0} \limsup_{T \to \infty} \sum_{i=1}^{n} \frac{\log(1/\theta_i)}{T}. \tag{14}$$

In particular, (11) holds if all θ_i are nonincreasing in T.

C. Active times and active rates

In this subsection, we introduce some switching-related quantities that will be useful for computing the topological entropy of a switched system. Given a switching signal σ, we define the *active time* of mode p of the switched system (2) over an interval $[0,t]$ by

$$\tau_p(t) := \int_0^t \mathbb{1}_p(\sigma(s))\,\mathrm{d}s, \qquad p \in \mathcal{P} \tag{15}$$

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 mode p over $[0,t]$ by

$$\rho_p(t) := \frac{\tau_p(t)}{t}, \qquad p \in \mathcal{P} \tag{16}$$

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

$$\hat{\rho}_p := \limsup_{t\to\infty} \rho_p(t), \qquad p \in \mathcal{P}. \tag{17}$$

Clearly, the active times $\tau_p(t) \geq 0$ are nondecreasing and satisfy $\sum_{p\in\mathcal{P}} \tau_p(t) = t$ for all $t \geq 0$; the active rates $\rho_p(t) \in [0,1]$ satisfy $\sum_{p\in\mathcal{P}} \rho_p(t) = 1$ for all $t \geq 0$. In contrast, due to the upper limit in (17), it is possible that $\sum_{p\in\mathcal{P}} \hat{\rho}_p > 1$ for the asymptotic active rates $\hat{\rho}_p \geq 0$; see [33, Example 1] for an illustration.

It was shown in [33, Lemma 1] that, for a family of constants $c_p \in \mathbb{R}$ with $p \in \mathcal{P}$, we have

$$\limsup_{T\to\infty} \frac{1}{T} \max_{t\in[0,T]} \sum_{p\in\mathcal{P}} c_p \tau_p(t) = \max\left\{ \limsup_{t\to\infty} \sum_{p\in\mathcal{P}} c_p \rho_p(t),\, 0 \right\}.$$

Next, we generalize this result to the case with a family of locally integrable functions $a_p : \mathbb{R}_{\geq 0} \to \mathbb{R}$ with $p \in \mathcal{P}$; see Appendix B for the proof.

Lemma 3. *For a family of locally integrable functions $a_p : \mathbb{R}_{\geq 0} \to \mathbb{R}$ with $p \in \mathcal{P}$, we have*

$$\begin{aligned} &\limsup_{T\to\infty} \frac{1}{T} \max_{t\in[0,T]} \sum_{p\in\mathcal{P}} \int_0^t a_p(s)\, \mathbb{1}_p(\sigma(s))\,\mathrm{d}s \\ &\qquad = \max\left\{ \limsup_{t\to\infty} \frac{1}{t} \sum_{p\in\mathcal{P}} \int_0^t a_p(s)\, \mathbb{1}_p(\sigma(s))\,\mathrm{d}s,\, 0 \right\}. \end{aligned} \tag{18}$$

Moreover, the first term in the maximum on the right-hand side of (18) satisfies

$$\limsup_{t\to\infty} \sum_{p\in\mathcal{P}} \check{a}_p \rho_p(t) \leq \limsup_{t\to\infty} \frac{1}{t} \sum_{p\in\mathcal{P}} \int_0^t a_p(s)\, \mathbb{1}_p(\sigma(s))\,\mathrm{d}s \leq \limsup_{t\to\infty} \sum_{p\in\mathcal{P}} \hat{a}_p \rho_p(t) \tag{19}$$

with

$$\check{a}_p := \liminf_{t\to\infty:\,\sigma(t)=p} a_p(t), \quad \hat{a}_p := \limsup_{t\to\infty:\,\sigma(t)=p} a_p(t), \qquad p \in \mathcal{P}.$$

In this paper, a supremum or infimum with a quantifier is taken to be zero if the quantifier does not hold anywhere; for example, we have

$$\begin{aligned} &\sup_{s\in[0,t]:\,\sigma(s)=p} a_p(s) \\ &\quad := \begin{cases} \sup\{a_p(s) : s \in [0,t],\, \sigma(s) = p\}, & \text{if } \{s \in [0,t] : \sigma(s) = p\} \neq \emptyset; \\ 0, & \text{otherwise.} \end{cases} \end{aligned}$$

Similarly, an upper or lower limit with a quantifier is taken to be zero if the quantifier only holds on a bounded set; for example, we have

$$\begin{aligned}\limsup_{t\to\infty:\,\sigma(t)=p} a_p(t) &:= \lim_{t\to\infty} \sup_{s\geq t:\,\sigma(s)=p} a_p(s)\\ &= \begin{cases}\lim\limits_{t\to\infty} \sup\{a_p(s) : s\geq t,\, \sigma(s)=p\}, & \text{if } \{t\geq 0 : \sigma(t)=p\} \text{ is unbounded;}\\ 0, & \text{otherwise.}\end{cases}\end{aligned}$$

Note that the sum in (18) is in fact an integral of the function $t \mapsto a_{\sigma(t)}(t)$ over $[0,t]$, that is,

$$\int_0^t a_{\sigma(s)}(s)\,\mathrm{d}s = \sum_{p\in\mathcal{P}} \int_0^t a_p(s)\,\mathbb{1}_p(\sigma(s))\,\mathrm{d}s. \tag{20}$$

Clearly, if the limits $\lim_{t\to\infty:\,\sigma(t)=p} a_p(t)$ exist for all $p\in\mathcal{P}$, then both inequalities in (19) hold with equality, that is,

$$\limsup_{t\to\infty} \frac{1}{t} \sum_{p\in\mathcal{P}} \int_0^t a_p(s)\,\mathbb{1}_p(\sigma(s))\,\mathrm{d}s = \limsup_{t\to\infty} \sum_{p\in\mathcal{P}} \Big(\lim_{t\to\infty:\,\sigma(t)=p} a_p(t)\Big)\rho_p(t).$$

III. Bounds for distance between solutions and volume of reachable set

In preparation for the construction of bounds for topological entropy, we establish several upper and lower bounds for the distance between two solutions to the switched system (2) and a lower bound for the volume of its reachable set. For brevity, we denote by $\xi_\sigma(t,K) := \{\xi_\sigma(t,x) : x\in K\}$ the *reachable set* of (2) with switching signal σ at time t from initial set K.

For a norm $\|\cdot\|$ on $\mathbb{R}^{n\times n}$ induced by a norm $|\cdot|$ on $\mathbb{R}^n$, its one-sided directional derivative at I in a direction $A\in\mathbb{R}^{n\times n}$ is called the *measure of the matrix* A and denoted by $\mu(A)$, that is,

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

The matrix measure $\mu(A)$ can also be seen as the right derivative of the functions $t\mapsto \|e^{At}\|$ and $t\mapsto \log\|e^{tA}\|$ at $t=0$ [39, Fact 11.15.7, p. 690]. Following [19, Th. (5), p. 31], the matrix measure $\mu(\cdot)$ is a convex function and satisfies

$$-\mu(-A) \leq \mathrm{Re}(\lambda) \leq \mu(A) \leq \|A\| \qquad \forall\, A\in\mathbb{R}^{n\times n}, \lambda\in\mathrm{spec}(A), \tag{22}$$

and

$$\mu(A+B) \leq \mu(A)+\mu(B), \quad \mu(cA) = c\mu(A) \qquad \forall\, A,B\in\mathbb{R}^n, c\geq 0. \tag{23}$$

For standard norms, there are explicit formulae for the matrix measure [19, Th. (24), p. 33]; for example, for the induced ∞-norm, the measure of a matrix $A=[a_{ij}]$ is given by

$$\mu(A) = \max_{1\leq i\leq n} \Big(a_{ii} + \sum_{j\neq i} |a_{ij}|\Big). \tag{24}$$

Lemma 4. *For all initial states $x,\bar{x}\in K$, the corresponding solutions to the switched system* (2) *satisfy*

$$|\xi_\sigma(t,\bar{x}) - \xi_\sigma(t,x)| \leq e^{\overline{\eta}_\sigma(t)} |\bar{x}-x| \qquad \forall\, t\geq 0 \tag{25}$$

with

$$\overline{\eta}_\sigma(t) := \max_{v\in\mathrm{co}(K)} \sum_{p\in\mathcal{P}} \int_0^t \mu(J_x f_p(\xi_\sigma(s,v)))\,\mathbb{1}_p(\sigma(s))\,\mathrm{d}s.$$

Also, the reachable set of (2) *satisfies*

$$\mathrm{vol}(\xi_\sigma(t,K)) \geq e^{\gamma_\sigma(t)}\,\mathrm{vol}(K) \qquad \forall\, t\geq 0 \tag{26}$$

with

$$\gamma_\sigma(t) := \min_{v\in K} \sum_{p\in\mathcal{P}} \int_0^t \mathrm{tr}(J_x f_p(\xi_\sigma(s,v)))\,\mathbb{1}_p(\sigma(s))\,\mathrm{d}s.$$

Note that the functions $\overline{\eta}_\sigma(t)$ and $\gamma_\sigma(t)$ are in fact in terms of integrals of the measure and trace of the Jacobian matrix of the system function of the active mode over $[0,t]$, rewritten using the transformation (20). (A similar observation can be made about the functions $\underline{\eta}_\sigma(t)$ and $\overline{\eta}_\sigma^{\mathrm{alt}}(t)$ in Lemma 5 below.) The upper bound (25) extends a similar one in [40, Th. 1] to the case of non-contractive switched systems, as well as one in the proof of [41, Th. 4.2] to the case of switched systems without a compact invariant set; its proof is inspired by the variational constructions in these results.

The following properties of linear time-varying (LTV) systems are useful for establishing Lemma 4 as well as Lemma 5 below. Consider an LTV system

$$\dot{x} = A(t)\, x \tag{27}$$

with a piecewise-continuous, matrix-valued function $A : \mathbb{R}_{\geq 0} \to \mathbb{R}^{n\times n}$. The solution to (27) at time t with initial state $v \in \mathbb{R}^n$ is given by $\xi(t,v) = \Phi_A(t,0)\, v$, where $\Phi_A(t,0)$ is the state-transition matrix and satisfies [19, Th. (27), p. 34]

$$e^{\int_0^t -\mu(-A(s))\,\mathrm{d}s}|v| \leq |\Phi_A(t,0)\, v| \leq e^{\int_0^t \mu(A(s))\,\mathrm{d}s}|v| \qquad \forall\, t \geq 0, v \in \mathbb{R}^n \tag{28}$$

and Liouville's formula [42, Th. 4.1, p. 28]

$$\det(\Phi_A(t,0)) = e^{\int_0^t \mathrm{tr}(A(s))\,\mathrm{d}s} \qquad \forall\, t \geq 0. \tag{29}$$

Note that the upper bound (28), and therefore the upper bound (25) and the bounds in (30) and (31) below, hold for all norms on $\mathbb{R}^n$ and the corresponding matrix measures.

Proof of Lemma 4. We prove Lemma 4 by writing the Jacobian matrix of a solution to the switched nonlinear system (2) with respect to the initial state $J_x\xi_\sigma(t,x)$ as a matrix solution to the LTV system (27) with a suitable function $A(\cdot)$, using variational arguments from nonlinear systems analysis (see, e.g., [43, Sec. 4.2.4]). For all $v \in \mathbb{R}^n$, we have $J_x\xi_\sigma(0,v) = I$ and

$$\partial_t J_x\xi_\sigma(t,v) = J_x\dot{\xi}_\sigma(t,v) = J_x f_{\sigma(t)}(\xi_\sigma(t,v))\, J_x\xi_\sigma(t,v)$$

for all $t \geq 0$ that are not switches. Hence for each fixed $v \in \mathbb{R}^n$, the matrix $J_x\xi_\sigma(t,v)$ is equal to the state-transition matrix $\Phi_A(t,0)$ of (27) with $A(t) = J_x f_{\sigma(t)}(\xi_\sigma(t,v))$.

First, given arbitrary initial states $x, \bar{x} \in K$, let

$$\nu(\rho) := \rho\bar{x} + (1-\rho)\, x, \qquad \rho \in [0,1].$$

Then $\nu(\rho) \in \mathrm{co}(K)$ and $\nu'(\rho) = \bar{x} - x$ for all $\rho \in [0,1]$. Hence

$$\begin{aligned} |\xi_\sigma(t,\bar{x}) - \xi_\sigma(t,x)| &= |\xi_\sigma(t,\nu(1)) - \xi_\sigma(t,\nu(0))| = \left| \int_0^1 J_x\xi_\sigma(t,\nu(\rho))(\bar{x}-x)\,\mathrm{d}\rho \right| \\ &\leq \int_0^1 |J_x\xi_\sigma(t,\nu(\rho))(\bar{x}-x)|\,\mathrm{d}\rho \leq \int_0^1 e^{\int_0^t \mu(J_x f_{\sigma(s)}(\xi_\sigma(s,\nu(\rho))))\,\mathrm{d}s}|\bar{x}-x|\,\mathrm{d}\rho \\ &\leq \Big(\max_{v\in \mathrm{co}(K)} e^{\int_0^t \mu(J_x f_{\sigma(s)}(\xi_\sigma(s,v)))\,\mathrm{d}s} \Big) |\bar{x}-x| = e^{\overline{\eta}_\sigma(t)}|\bar{x}-x| \end{aligned}$$

for all $t \geq 0$, where the second inequality follows from the second inequality in (28) with $A(t) = J_x f_{\sigma(t)}(\xi_\sigma(t,\nu(\rho)))$ and $\Phi_A(t,0) = J_x\xi_\sigma(t,\nu(\rho))$, and the last equality follows from the transformation (20). Therefore, the upper bound (25) holds.

Second, Liouville's formula (29) with $A(t) = J_x f_{\sigma(t)}(\xi_\sigma(t,v))$ and the state-transition matrix $\Phi_A(t,0) = J_x\xi_\sigma(t,v)$ implies that

$$\det(J_x\xi_\sigma(t,v)) = e^{\int_0^t \mathrm{tr}(J_x f_{\sigma(s)}(\xi_\sigma(s,v)))\,\mathrm{d}s} \qquad \forall\, t \geq 0, v \in \mathbb{R}^n.$$

Hence

$$\begin{aligned} \mathrm{vol}(\xi_\sigma(t,K)) &= \int_K |\det(J_x\xi_\sigma(t,v))|\,\mathrm{d}v \geq \Big(\min_{v\in K} |\det(J_x\xi_\sigma(t,v))| \Big)\, \mathrm{vol}(K) \\ &= \Big(\min_{v\in K} e^{\int_0^t \mathrm{tr}(J_x f_{\sigma(s)}(\xi_\sigma(s,v)))\,\mathrm{d}s} \Big)\, \mathrm{vol}(K) = e^{\gamma_\sigma(t)}\, \mathrm{vol}(K) \end{aligned}$$

for all $t \geq 0$, where the last equality follows from the transformation (20). Therefore, the lower bound (26) holds. $\square$

The following result provides alternative upper and lower bounds for the distance between two solutions to (2).

Lemma 5. *For all initial states $x, \bar{x} \in K$, the corresponding solutions to the switched system (2) satisfy*

$$e^{\underline{\eta}_\sigma(t)}|\bar{x} - x| \leq |\xi_\sigma(t, \bar{x}) - \xi_\sigma(t, x)| \leq e^{\overline{\eta}_\sigma^{\text{alt}}(t)}|\bar{x} - x| \qquad \forall\, t \geq 0 \tag{30}$$

with

$$\underline{\eta}_\sigma(t) := \sum_{p \in \mathcal{P}} \int_0^t \Big(\min_{v \in \text{co}(\xi_\sigma(s,K))} -\mu(-J_x f_p(v)) \Big) \mathbb{1}_p(\sigma(s))\, \text{d}s,$$

$$\overline{\eta}_\sigma^{\text{alt}}(t) := \sum_{p \in \mathcal{P}} \int_0^t \Big(\max_{v \in \text{co}(\xi_\sigma(s,K))} \mu(J_x f_p(v)) \Big) \mathbb{1}_p(\sigma(s))\, \text{d}s.$$

Lemma 5 extends similar upper and lower bounds in [44, Th. (22), p. 52] by taking the maximum and minimum over the convex hull of reachable set instead of the entire state space in $\underline{\eta}_\sigma(t)$ and $\overline{\eta}_\sigma^{\text{alt}}(t)$; see Appendix C for the proof, which is inspired by the variational construction in [44, Sec. 2.5], and is similar to the proof of the upper bound (25) in Lemma 4.[1]

The results in Lemmas 4 and 5 are compared in the following remark.

Remark 3. 1) The upper bound (25) and the one given by the second inequality in (30) are both useful in the sense that neither is less conservative, partially because the relation between the sets $\xi_\sigma(t, \text{co}(K))$ and $\text{co}(\xi_\sigma(t, K))$ is undetermined in general for the switched nonlinear system (2). However, (25) is less conservative or equivalent if the initial set K is convex, or if all modes of (2) are linear, as $\xi_\sigma(t, \text{co}(K)) \subset \text{co}(\xi_\sigma(t, K))$ for the former case and $\xi_\sigma(t, \text{co}(K)) = \text{co}(\xi_\sigma(t, K))$ for the latter, and the maximum is a subadditive function.

2) The first inequality in (22) implies that $\gamma_\sigma(t)$ in (26) and $\underline{\eta}_\sigma(t)$ in (30) satisfy $\gamma_\sigma(t) \geq n\underline{\eta}_\sigma(t)$ for all $t \geq 0$. Hence volume-based arguments using (26) usually lead to less conservative lower bounds for entropy. However, the lower bound given by the first inequality in (30) will be useful for constructing less conservative lower bounds for entropy for the case with diagonal modes, as it can be applied to each scalar component separately.

Based on (25) and the first inequality in (30), we construct the following upper and lower bounds for the distance between two solutions in terms of the active times $\tau_p(t)$, which are more conservative but illustrate the effect of switching and will be useful for constructing bounds for entropy for the case with diagonal modes in Section V-A.

Corollary 1. *For all initial states $x, \bar{x} \in K$, the corresponding solutions to the switched system (2) satisfy*

$$e^{\sum_{p \in \mathcal{P}} \underline{\mu}_p(t)\tau_p(t)}|\bar{x} - x| \leq |\xi_\sigma(t, \bar{x}) - \xi_\sigma(t, x)| \leq e^{\sum_{p \in \mathcal{P}} \overline{\mu}_p(t)\tau_p(t)}|\bar{x} - x| \qquad \forall\, t \geq 0 \tag{31}$$

with

$$\left.\begin{aligned} \underline{\mu}_p(t) &:= \inf_{s \in [0,t] \colon \sigma(s) = p} \; \min_{v \in \text{co}(\xi_\sigma(s,K))} -\mu(-J_x f_p(v)), \\ \overline{\mu}_p(t) &:= \sup_{s \in [0,t] \colon \sigma(s) = p} \; \max_{v \in \xi_\sigma(s,\text{co}(K))} \mu(J_x f_p(v)), \end{aligned}\right. \qquad p \in \mathcal{P},$$

where the active times $\tau_p(t)$ are defined by (15).

Proof. The second inequality in (31) follows from (25) and the property that

$$\begin{aligned} \overline{\eta}_\sigma(t) &\leq \sum_{p \in \mathcal{P}} \int_0^t \Big(\max_{v \in \xi_\sigma(t,\text{co}(K))} \mu(J_x f_p(v)) \Big) \mathbb{1}_p(\sigma(s))\, \text{d}s \\ &\leq \sum_{p \in \mathcal{P}} \overline{\mu}_p(t) \int_0^t \mathbb{1}_p(\sigma(s))\, \text{d}s = \sum_{p \in \mathcal{P}} \overline{\mu}_p(t)\tau_p(t) \end{aligned}$$

for all $t \geq 0$, where the first inequality holds as the maximum is a subadditive function.

[1] Their main difference is that, in the proof of Proposition 5, the variational arguments are applied to the line segment connecting two solutions instead of the one connecting two initial states, which results in the different locations of the convex hulls in $\overline{\eta}_\sigma^{\text{alt}}(t)$ and $\overline{\eta}_\sigma(t)$.

The first inequality in (31) follows from the first inequality in (30) and the property that

$$\underline{\eta}_\sigma(t) \geq \sum_{p\in\mathcal{P}} \underline{\mu}_p(t) \int_0^t \mathbb{1}_p(\sigma(s))\,\mathrm{d}s = \sum_{p\in\mathcal{P}} \underline{\mu}_p(t)\tau_p(t)$$

for all $t \geq 0$. $\square$

IV. Entropy of general switched nonlinear systems

In this section, we construct upper and lower bounds for the topological entropy of a general switched nonlinear system.

Theorem 1. *The topological entropy of the switched system* (2) *is upper bounded by*

$$h(f_\sigma, K) \leq \max\left\{\limsup_{t\to\infty} \sum_{p\in\mathcal{P}} n\hat{\mu}_p\rho_p(t),\, 0\right\} \tag{32}$$

with

$$\hat{\mu}_p := \limsup_{t\to\infty:\, \sigma(t)=p}\; \max_{v\in\xi_\sigma(t,\mathrm{co}(K))} \mu(J_x f_p(v)), \qquad p\in\mathcal{P}, \tag{33}$$

and lower bounded by

$$h(f_\sigma, K) \geq \max\left\{\limsup_{t\to\infty} \sum_{p\in\mathcal{P}} \check{\chi}_p\rho_p(t),\, 0\right\} \tag{34}$$

with

$$\check{\chi}_p := \liminf_{t\to\infty:\, \sigma(t)=p}\; \min_{v\in\xi_\sigma(t,K)} \mathrm{tr}(J_x f_p(v)), \qquad p\in\mathcal{P}, \tag{35}$$

where the active rates $\rho_p(t)$ are defined by (16).

Note that the bounds (32) and (34) are in terms of asymptotic weighted averages of the constants $\hat{\mu}_p$ and $\check{\chi}_p$, with the active rates $\rho_p(t)$ as weights. Due to the upper limit in (33), the constants $\hat{\mu}_p$, and therefore the upper bound (32), depend on the Jacobian matrices of system functions $J_x f_p(x)$ over only the ω-limit set from the convex hull of initial set $\mathrm{co}(K)$, instead of all reachable points from $\mathrm{co}(K)$. In particular, (32) will yield a finite bound in the case with unbounded Jacobian matrices of system functions but a compact ω-limit set. A similar property holds for the lower bound (34) due to the lower limit in (35), and these properties of (32) and (34) are obtained using Lemma 3 from Section II-C. (In fact, similar properties hold for all bounds for entropy in this paper, except (41) and (43) for the case with diagonal modes in Section V-A.)

Proof of Theorem 1. First, we prove the upper bound (32). Given arbitrary time horizon $T \geq 0$ and radius $\varepsilon > 0$, the upper bound for the distance between two solutions (25) implies that, for all initial states $x, \bar{x} \in K$, the corresponding solutions to (2) satisfy

$$\max_{t\in[0,T]} |\xi_\sigma(t,\bar{x}) - \xi_\sigma(t,x)| \leq \max_{t\in[0,T]} e^{\overline{\eta}_\sigma(t)}|\bar{x} - x|. \tag{36}$$

Consider the grid $G(\theta)$ defined by (8) with

$$\theta_i := e^{-\max_{t\in[0,T]}\overline{\eta}_\sigma(t)}\varepsilon, \qquad i\in\{1,\dots,n\},$$

and the corresponding hypercubes $R(x)$ defined by (9). Comparing (9) and (36) to (3), we see that $R(x) \subset B_{f_\sigma}(x,\varepsilon,T)$ for all $x \in G(\theta)$. Then Lemma 2 implies that $G(\theta)$ is (T,ε)-spanning. As $T \geq 0$ and $\varepsilon > 0$ are arbitrary, and all θ_i are nonincreasing in T, the upper bound (12) implies that

$$\begin{aligned}
h(f_\sigma, K) &\leq \limsup_{\varepsilon\searrow 0}\limsup_{T\to\infty} \sum_{i=1}^n \frac{\log(1/\theta_i)}{T} \\
&= \limsup_{T\to\infty} \frac{1}{T}\max_{t\in[0,T]} n\overline{\eta}_\sigma(t) + \lim_{\varepsilon\searrow 0}\lim_{T\to\infty}\frac{n\log(1/\varepsilon)}{T} \\
&\leq \limsup_{T\to\infty} \frac{1}{T}\max_{t\in[0,T]} \sum_{p\in\mathcal{P}} \int_0^t \Big(\max_{v\in\xi_\sigma(s,\mathrm{co}(K))} n\mu(J_x f_p(v))\Big)\mathbb{1}_p(\sigma(s))\,\mathrm{d}s,
\end{aligned}$$

where the equality holds as the limits in the second term on its right-hand side exist, and the last inequality holds partially because the maximum is a subadditive function. Then (32) follows from (18) and the second inequality in (19) with the functions

$$a_p(t) = \max_{v \in \xi_\sigma(t, \mathrm{co}(K))} \mu(J_x f_p(v)), \qquad p \in \mathcal{P}.$$

Second, we prove the lower bound (34) using volume-based arguments. Given arbitrary time horizon $T \geq 0$ and radius $\varepsilon > 0$, the lower bound for the volume of reachable set (26) implies that the reachable set $\xi_\sigma(T, K)$ of (2) satisfies

$$\mathrm{vol}(\xi_\sigma(T, K)) \geq e^{\gamma_\sigma(T)} \mathrm{vol}(K).$$

Let E be a minimal (T, ε)-spanning set. Then (3) and (4) imply that

$$\xi_\sigma(T, K) \subset \bigcup_{x \in E} \{\bar{x} \in \mathbb{R}^n : |\bar{x} - \xi_\sigma(T, x)| < \varepsilon\},$$

and thus (recall that we take $|\cdot|$ to be the ∞-norm; see Remark 1)

$$\mathrm{vol}(\xi_\sigma(T, K)) \leq \sum_{x \in E} \mathrm{vol}(\{\bar{x} \in \mathbb{R}^n : |\bar{x} - \xi_\sigma(T, x)| < \varepsilon\}) = (2\varepsilon)^n \# E.$$

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

$$S(f_\sigma, \varepsilon, T, K) = \# E \geq \frac{\mathrm{vol}(\xi_\sigma(T, K))}{(2\varepsilon)^n} \geq \frac{e^{\gamma_\sigma(T)} \mathrm{vol}(K)}{(2\varepsilon)^n},$$

which, combined with the definition of entropy (5), implies that

$$\begin{aligned} h(f_\sigma, K) &\geq \liminf_{\varepsilon \searrow 0} \limsup_{T \to \infty} \frac{\log(e^{\gamma_\sigma(T)} \mathrm{vol}(K)/(2\varepsilon)^n)}{T} \\ &= \limsup_{T \to \infty} \frac{\gamma_\sigma(T)}{T} + \lim_{\varepsilon \searrow 0} \lim_{T \to \infty} \frac{\log(\mathrm{vol}(K)/(2\varepsilon)^n)}{T} \\ &\geq \limsup_{T \to \infty} \frac{1}{T} \sum_{p \in \mathcal{P}} \int_0^T \Big(\min_{v \in \xi_\sigma(s, K)} \mathrm{tr}(J_x f_p(v)) \Big) \mathbb{1}_p(\sigma(s)) \, \mathrm{d}s, \end{aligned}$$

where the equality holds as the limits in the second term on its right-hand side exist, and the last inequality holds partially because the minimum is a superadditive function. Then (34) follows from the first inequality in (19) with the functions

$$a_p(t) = \min_{v \in \xi_\sigma(t, K)} \mathrm{tr}(J_x f_p(v)), \qquad p \in \mathcal{P},$$

and the property that $h(f_\sigma, K) \geq 0$. □

Remark 4. In view of the equivalence of norms on a finite-dimensional vector space, the upper bound for entropy (32) holds when the matrix measures in all $\hat{\mu}_p$ are computed with a norm $\|\cdot\|_\alpha$ on $\mathbb{R}^{n \times n}$ induced by an arbitrary norm $|\cdot|_\alpha$ on $\mathbb{R}^n$. The proof remains the same, except we now consider $\theta_i := e^{-\max_{t \in [t_0, t_0 + T]} \overline{\eta}(t, t_0)} \varepsilon / r_\alpha$, where $r_\alpha > 0$ is a constant such that

$$|v|_\alpha \leq r_\alpha |v|_\infty \qquad \forall\, v \in \mathbb{R}^n$$

and vanishes after taking the upper limit as $T \to \infty$. (Similar properties hold for all upper bounds for entropy in this section.)

Remark 5. The upper bound for entropy (32) also holds with

$$\hat{\mu}_p^{\mathrm{alt}} = \limsup_{t \to \infty :\, \sigma(t) = p} \max_{v \in \mathrm{co}(\xi_\sigma(t, K))} \mu(J_x f_p(v)), \qquad p \in \mathcal{P}$$

in place of $\hat{\mu}_p$. The proof remains that same, except now the second inequality in (30) is used in place of (25). Following item 1 of Remark 3, the relation between the values of the upper bound (32) computed using $\hat{\mu}_p$ and $\hat{\mu}_p^{\mathrm{alt}}$ is undetermined in general for the switched nonlinear system (2). However, the former is less conservative than or equivalent to the latter if the initial set K is convex, and they are equivalent for all initial sets if all modes of (2) are linear. (Similar properties hold for all upper bounds for entropy in this paper.)

Remark 6. In many scenarios, we can construct simpler but more conservative bounds for entropy based on those in Theorem 1. For example, let $S \subset \mathbb{R}^n$ be a set such that one of the following holds:
1) $S = \mathbb{R}^n$;
2) S is compact and positively invariant for (2) and contains the convex hull of initial set $\mathrm{co}(K)$; or
3) S is compact and contains the ω-limit set from $\mathrm{co}(K)$.

Suppose there exists a family of constant $\hat{\mu}_p^*$ with $p \in \mathcal{P}$ such that $\hat{\mu}_p^* \geq \mu(J_x f_p(v))$ for all $p \in \mathcal{P}$ and $v \in S$. Then the upper bound (32) also holds with $\hat{\mu}_p^*$ in place of $\hat{\mu}_p$. In the numerical examples in Section VI, the computation will be simplified based on scenario 3 here. (Similar properties hold for all bounds for entropy in this paper.)

Thinking of the non-switched case as a switched system with a constant switched signal, Theorem 1 implies the following upper and lower bounds for the entropy of a time-invariant system.

Corollary 2. *The topological entropy of a time-invariant system $\dot{x} = f(x)$ with $x \in \mathbb{R}^n$ satisfies*

$$\max\{\check{\chi}, 0\} \leq h(f, K) \leq \max\{n\hat{\mu}, 0\} \tag{37}$$

with

$$\hat{\mu} := \limsup_{t\to\infty} \max_{v\in\xi(t,\mathrm{co}(K))} \mu(J_x f(v)), \qquad \check{\chi} := \liminf_{t\to\infty} \min_{v\in\xi(t,K)} \mathrm{tr}(J_x f(v)).$$

Based on the upper bound (32), we construct the following upper bounds for the entropy of (2), which require less information about the switching signal.

Corollary 3. *The topological entropy of the switched system (2) is upper bounded by*

$$h(f_\sigma, K) \leq \sum_{p\in\mathcal{P}} \max\{n\hat{\mu}_p, 0\}\, \hat{\rho}_p \tag{38}$$

with the asymptotic active rates $\hat{\rho}_p$ defined by (17), and also by

$$h(f_\sigma, K) \leq \max_{p\in\mathcal{P}} \max\{n\hat{\mu}_p, 0\}, \tag{39}$$

where the constants $\hat{\mu}_p$ are defined by (33).

Proof. First, as the upper limit is a subadditive function, the upper bound (32) implies that

$$\begin{aligned} h(f_\sigma, K) &\leq \max\left\{\sum_{p\in\mathcal{P}} \limsup_{t\to\infty} n\hat{\mu}_p \rho_p(t),\, 0\right\} \\ &\leq \sum_{p\in\mathcal{P}} \max\{n\hat{\mu}_p, 0\} \limsup_{t\to\infty} \rho_p(t) = \sum_{p\in\mathcal{P}} \max\{n\hat{\mu}_p, 0\}\, \hat{\rho}_p. \end{aligned}$$

Second, (32) also implies that

$$\begin{aligned} h(f_\sigma, K) &\leq \max\left\{\limsup_{t\to\infty} \left(\max_{p\in\mathcal{P}} n\hat{\mu}_p\right) \sum_{p\in\mathcal{P}} \rho_p(t),\, 0\right\} \\ &= \max\left\{\max_{p\in\mathcal{P}} n\hat{\mu}_p,\, 0\right\} = \max_{p\in\mathcal{P}} \max\{n\hat{\mu}_p, 0\}. \end{aligned}$$

□

The results in Theorem 1 and Corollary 3 are compared in the following remark and illustrated by Example 1 in Section VI.

Remark 7. 1) The upper bound (32) is less conservative than or equivalent to the upper bounds (38) and (39), while (38) and (39) are both useful in the sense that neither is less conservative, as it is possible that $\sum_{p\in\mathcal{P}} \hat{\rho}_p > 1$.
2) For a fixed family of modes, the upper bounds (38) and (39) require less information about the switching signal than the upper bound (32), as (38) depends on the asymptotic active rates $\hat{\rho}_p$ instead of the active rates $\rho_p(t)$,

and (39) does not involve active rates at all. If an upper bound $\hat{\mu}_p^*$ as in scenarios 1 and 2 in Remark 6 is used in place of each $\hat{\mu}_p$, then (39) is independent of switching.

Remark 8. Consider the case where all modes of (2) are linear, that is, there is a family of matrices $A_p \in \mathbb{R}^{n\times n}$ with $p \in \mathcal{P}$ such that

$$f_p(x) = A_p x \qquad \forall\, p \in \mathcal{P}, x \in \mathbb{R}^n.$$

Then the constants $\hat{\mu}_p$ and $\check{\chi}_p$ defined by (33) and (35) satisfy

$$\hat{\mu}_p = \mu(A_p), \quad \check{\chi}_p = \mathrm{tr}(A_p) \qquad \forall\, p \in \mathcal{P}.$$

Hence Theorem 1 and Corollary 3 generalize [33, Th. 1 and Remark 5] to the case with nonlinear modes, respectively.

V. Entropy of systems with special structure

A. Switched diagonal systems

Consider the case where all modes of the switched system (2) are diagonal. Specifically, for each $p \in \mathcal{P}$ and $i \in \{1, \dots, n\}$, the i-th scalar component $f_p^i(x)$ of the function $f_p(x)$ depends only on the corresponding scalar component x_i of the state x. For brevity, we regard $f_p^i(x)$ as a function on $\mathbb{R}$ and denote it by $f_p^i(x_i)$. Then (2) can be written as the *switched diagonal system*

$$\dot{x}_i = f_\sigma^i(x_i), \qquad i \in \{1, \dots, n\}. \tag{40}$$

Clearly, the i-th scalar component of the solution $\xi_\sigma(t, x)$ to (40) depends only on the corresponding scalar component x_i of the initial state x. For brevity, we denote it by $\xi_\sigma^i(t, x_i)$ and the corresponding reachable set by $\xi_\sigma^i(t, K_i)$, where $K_i \subset \mathbb{R}$ is the corresponding projection of the initial set $K \subset \mathbb{R}^n$.

In this subsection, we construct bounds for the topological entropy of (40) that are usually less conservative than the results of applying the general bounds for entropy from Section IV.

Proposition 1. *The topological entropy of the switched diagonal system* (40) *is upper bounded by*

$$h(f_\sigma, K) \le \limsup_{T\to\infty} \sum_{i=1}^{n} \frac{1}{T} \max_{t\in[0,T]} \sum_{p\in\mathcal{P}} \overline{a}_p^i(t)\tau_p(t) \tag{41}$$

with

$$\overline{a}_p^i(t) := \sup_{s\in[0,t]:\, \sigma(s)=p}\ \max_{v_i \in \xi_\sigma^i(s, \mathrm{co}(K_i))} (f_p^i)'(v_i), \qquad p \in \mathcal{P}, i \in \{1, \dots, n\}, \tag{42}$$

and lower bounded by

$$h(f_\sigma, K) \ge \limsup_{T\to\infty} \sum_{i=1}^{n} \frac{1}{T} \max_{t\in[0,T]} \sum_{p\in\mathcal{P}} \underline{a}_p^i(t)\tau_p(t) \tag{43}$$

with

$$\underline{a}_p^i(t) := \inf_{s\in[0,t]:\, \sigma(s)=p}\ \min_{v_i \in \mathrm{co}(\xi_\sigma^i(s, K_i))} (f_p^i)'(v_i), \qquad p \in \mathcal{P}, i \in \{1, \dots, n\}, \tag{44}$$

where the active times $\tau_p(t)$ are defined by (15).

Proof. Given arbitrary time horizon $T \ge 0$ and radius $\varepsilon > 0$, by applying (31) to each scalar component of (40), we obtain that, for all initial states $x = (x_1, \dots, x_n), \bar{x} = (\bar{x}_1, \dots, \bar{x}_n) \in K$, the corresponding solutions satisfy (recall that we take $|\cdot|$ to be the ∞-norm; see Remark 1)

$$\begin{aligned}\max_{t\in[0,T]} \max_{1\le i\le n} e^{\sum_{p\in\mathcal{P}} \underline{a}_p^i(t)\tau_p(t)} |\bar{x}_i - x_i| &\le \max_{t\in[0,T]} |\xi_\sigma(t, \bar{x}) - \xi_\sigma(t, x)| \\ &\le \max_{t\in[0,T]} \max_{1\le i\le n} e^{\sum_{p\in\mathcal{P}} \overline{a}_p^i(t)\tau_p(t)} |\bar{x}_i - x_i|. \end{aligned} \tag{45}$$

First, consider the grid $G(\theta)$ defined by (8) with

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

and the corresponding hyperrectangles $R(x)$ defined by (9). Comparing (9) and the second inequality in (45) to (3), we see that $R(x) \subset B_{f_\sigma}(x,\varepsilon,T)$ for all $x \in G(\theta)$. Then Lemma 2 implies that $G(\theta)$ is (T,ε)-spanning. As $T \geq 0$ and $\varepsilon > 0$ are arbitrary, and all θ_i are nonincreasing in T, the upper bound (12) implies that

$$\begin{aligned} h(f_\sigma, K) &\leq \limsup_{\varepsilon \searrow 0} \limsup_{T \to \infty} \sum_{i=1}^{n} \frac{\log(1/\theta_i)}{T} \\ &= \limsup_{T\to\infty} \sum_{i=1}^{n} \frac{1}{T} \max_{t\in[0,T]} \sum_{p\in\mathcal{P}} \overline{a}_p^i(t)\tau_p(t) + \lim_{\varepsilon\searrow 0} \lim_{T\to\infty} \frac{n\log(1/\varepsilon)}{T} \\ &= \limsup_{T\to\infty} \sum_{i=1}^{n} \frac{1}{T} \max_{t\in[0,T]} \sum_{p\in\mathcal{P}} \overline{a}_p^i(t)\tau_p(t), \end{aligned}$$

where the first equality holds as the limits in the second term on its right-hand side exist. Hence (41) holds.

Second, following arguments similar to those in the first part while considering

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

we can show that $G(\theta)$ is (T,ε)-separated, and the lower bound (14) implies (43). □

Based on the upper bound (41), we construct the following upper bound for the entropy of (40), which is an asymptotic weighted average over individual modes, with the active rates $\rho_p(t)$ as weights.

Corollary 4. *The topological entropy of the switched diagonal system* (40) *is upper bounded by*

$$h(f_\sigma, K) \leq \limsup_{t\to\infty} \sum_{p\in\mathcal{P}} \left(\sum_{i=1}^{n} \max\{\overline{a}_p^i(t),\, 0\} \right) \rho_p(t) \tag{46}$$

with the functions $\overline{a}_p^i(t)$ defined by (42) *and the active rates $\rho_p(t)$ defined by* (16).

Proof. Note that the functions $\overline{a}_p^i(t)$ are nondecreasing. Hence the upper bound (41) implies that

$$h(f_\sigma, K) \leq \limsup_{T\to\infty} \sum_{i=1}^{n} \frac{1}{T} \sum_{p\in\mathcal{P}} \max\{\overline{a}_p^i(T),\, 0\}\tau_p(T),$$

that is, (46) holds. □

Alternatively, we construct the following upper bound for the entropy of (40) which is a sum over scalar components, and each summand is in terms of an asymptotic weighted average over individual modes, with the active rates $\rho_p(t)$ as weights.

Proposition 2. *The topological entropy of the switched diagonal system* (40) *is upper bounded by*

$$h(f_\sigma, K) \leq \sum_{i=1}^{n} \max\left\{ \limsup_{t\to\infty} \sum_{p\in\mathcal{P}} \hat{a}_p^i \rho_p(t),\, 0 \right\} \tag{47}$$

with

$$\hat{a}_p^i := \limsup_{t\to\infty:\, \sigma(t)=p} \max_{v_i \in \xi_\sigma^i(t, \mathrm{co}(K_i))} (f_p^i)'(v_i), \qquad p \in \mathcal{P}, i \in \{1,\dots,n\} \tag{48}$$

and the active rates $\rho_p(t)$ defined by (16).

Proof. Given arbitrary time horizon $T \geq 0$ and radius $\varepsilon > 0$, by applying the upper bound (25) to each scalar component of (40), we obtain that, for all initial states $x = (x_1,\dots,x_n), \bar{x} = (\bar{x}_1,\dots,\bar{x}_n) \in K$, the corresponding solutions satisfy (recall that we take $|\cdot|$ to be the ∞-norm; see Remark 1)

$$\max_{t\in[0,T]} |\xi_\sigma(t,\bar{x}) - \xi_\sigma(t,x)| \leq \max_{t\in[0,T]} \max_{1\leq i\leq n} e^{\overline{\eta}_\sigma^i(t)} |\bar{x}_i - x_i|, \tag{49}$$

with

$$\overline{\eta}_\sigma^i(t) := \max_{v_i\in\mathrm{co}(K_i)} \sum_{p\in\mathcal{P}} \int_0^t (f_p^i)'(\xi_\sigma^i(s,v_i))\, \mathbb{1}_p(\sigma(s))\, \mathrm{d}s, \qquad i \in \{1,\dots,n\}.$$

Consider the grid $G(\theta)$ defined by (8) with

$$\theta_i := e^{-\max_{t\in[0,T]}\overline{\eta}_\sigma^i(t)}\varepsilon, \qquad i \in \{1,\dots,n\},$$

and the corresponding hyperrectangles $R(x)$ defined by (9). Comparing (9) and (49) to (3), we see that $R(x) \subset B_{f_\sigma}(x,\varepsilon,T)$ for all $x \in G(\theta)$. Then Lemma 2 implies that $G(\theta)$ is (T,ε)-spanning. As $T \geq 0$ and $\varepsilon > 0$ are arbitrary, and all θ_i are nonincreasing in T, the upper bound (12) implies that

$$\begin{aligned} h(f_\sigma, K) &\leq \limsup_{\varepsilon\searrow 0}\limsup_{T\to\infty}\sum_{i=1}^n \frac{\log(1/\theta_i)}{T} \\ &= \limsup_{T\to\infty}\sum_{i=1}^n \frac{1}{T}\max_{t\in[0,T]}\overline{\eta}_\sigma^i(t) + \lim_{\varepsilon\searrow 0}\lim_{T\to\infty}\frac{n\log(1/\varepsilon)}{T} \\ &\leq \sum_{i=1}^n \limsup_{T\to\infty}\frac{1}{T}\max_{t\in[0,T]}\sum_{p\in\mathcal{P}}\int_0^t \Big(\max_{v_i\in\xi_\sigma^i(s,\mathrm{co}(K_i))}(f_p^i)'(v_i)\Big)\mathbb{1}_p(\sigma(s))\,\mathrm{d}s, \end{aligned}$$

where the equality holds as the limits in the second term on its right-hand side exist, and the last inequality holds partially because the upper limit and maximum are subadditive functions. Then we obtain (47) by applying (18) and the second inequality in (19) to each scalar component i with the functions

$$a_p^i(t) = \max_{v_i\in\xi_\sigma^i(t,\mathrm{co}(K_i))}(f_p^i)'(v_i), \qquad p\in\mathcal{P}, i\in\{1,\dots,n\}$$

in place of $a_p(t)$. □

Thinking of the non-switched case as a switched system with a constant switched signal, Proposition 2 implies the following upper bound for the entropy of a time-invariant diagonal system, which is less conservative than or equivalent to the result of applying the upper bound given by the second inequality in (37).

Corollary 5. *The topological entropy of a time-invariant diagonal system $\dot{x}_i = f^i(x_i)$, $i \in \{1,\dots,n\}$ with $x_i \in \mathbb{R}$ is upper bounded by*

$$h(f, K) \leq \sum_{i=1}^n \max\Big\{\limsup_{t\to\infty}\max_{v_i\in\xi^i(t,\mathrm{co}(K_i))}(f^i)'(v_i),\, 0\Big\}. \tag{50}$$

Based on the upper bounds (46) and (47), we construct the following upper bounds for the entropy of (40), which require less information about the switching signal; the proof is along the lines of that of Corollary 3 and thus omitted here.

Corollary 6. *The topological entropy of the switched diagonal system (40) is upper bounded by*

$$h(f_\sigma, K) \leq \max_{p\in\mathcal{P}}\left(\sum_{i=1}^n \max\Big\{\lim_{t\to\infty}\overline{a}_p^i(t),\, 0\Big\}\right) \tag{51}$$

with the functions $\overline{a}_p^i(t)$ defined by (42), by

$$h(f_\sigma, K) \leq \sum_{p\in\mathcal{P}}\left(\sum_{i=1}^n \max\{\hat{a}_p^i,\, 0\}\right)\hat{\rho}_p \tag{52}$$

with the asymptotic active rates $\hat{\rho}_p$ defined by (17), and also by

$$h(f_\sigma, K) \leq \sum_{i=1}^n \max_{p\in\mathcal{P}}\max\{\hat{a}_p^i,\, 0\}, \tag{53}$$

where the constants $\hat{a}_p^i$ are defined by (48).

The results in Theorem 1, Propositions 1 and 2, and Corollaries 3, 4, and 6 are compared in the following remark and illustrated by Example 2 in Section VI.

Remark 9. 1) The upper bound (41) is less conservative than or equivalent to the upper bound (46), and they are equivalent if $\overline{a}_p^i(t) \geq 0$ for all $p \in \mathcal{P}$, $i \in \{1,\dots,n\}$, and $t \geq 0$. The upper bound (46) is less conservative than or equivalent to the upper bound (51), and the upper bound (47) is less conservative than or equivalent to the upper bounds (52) and (53), while (51), (52), and (53) are all useful in the sense that neither is less conservative, as $\hat{a}_p^i \leq \lim_{t\to\infty} \overline{a}_p^i(t)$ for all $p \in \mathcal{P}$ and $i \in \{1,\dots,n\}$, the maximum is a subadditive function, and it is possible that $\sum_{p\in\mathcal{P}} \hat{\rho}_p > 1$. The upper bounds (41) and (47) are both useful in the same sense. On one hand, (47) may be less conservative as $\hat{a}_p^i$ depends on $(f_p^i)'(v_i)$ over only the ω-limit set from $\mathrm{co}(K_i)$, while $\overline{a}_p^i(t)$ depends on $(f_p^i)'(v_i)$ over all reachable points up to T from $\mathrm{co}(K_i)$. On the other hand, (41) is less conservative or equivalent for the case with linear modes; see [32, Th. 7 and Prop. 8] and [33, Example 3]. The relations between (41) and the upper bounds (52) and (53), as well as that between (47) and (51), are undetermined for similar reasons.

2) For the switched diagonal system (40), the upper bound (47) is less conservative than or equivalent to the general upper bound (32), the upper bounds (51) and (53) are less conservative than or equivalent to the upper bound (39), and the upper bound (52) is less conservative than or equivalent to the upper bound (38).
3) For a fixed family of modes, the upper bounds (51), (52), and (53) require less information about the switching signal than the upper bounds (41), (46), and (47), as (52) depends on the asymptotic active rates $\hat{\rho}_p$ instead of the active times $\tau_p(t)$ and active rates $\rho_p(t)$, and (51) and (53) do not involve active rates at all. If an upper bound $(\hat{a}_p^i)^*$ as in scenarios 1 and 2 in Remark 6 is used in place of each $\overline{a}_p^i(t)$ and $\hat{a}_p^i$, then (51) and (53) are independent of switching.

Remark 10. Consider the case where all modes of (40) are linear, that is, there is a family of diagonal matrices $D_p = \mathrm{diag}(a_p^1,\dots,a_p^n) \in \mathbb{R}^{n\times n}$ with $p \in \mathcal{P}$ such that

$$f_p(x) = D_p x \qquad \forall\, p \in \mathcal{P}, x \in \mathbb{R}^n.$$

Then the functions $\underline{a}_p^i(t)$ and $\overline{a}_p^i(t)$ defined by (42) and (44) and the constants $\hat{a}_p^i$ defined by (48) satisfy

$$\underline{a}_p^i(t) = \overline{a}_p^i(t) = \hat{a}_p^i = a_p^i \qquad \forall\, p \in \mathcal{P}, i \in \{1,\dots,n\}, t \geq 0.$$

Hence Propositions 1 and 2 and Corollaries 4 and 6 generalize [32, Th. 7, Prop. 8 and 9, and Cor. 10] for diagonal matrices in $\mathbb{R}^{n\times n}$ to the case with nonlinear modes, respectively.

B. *Switched interconnected systems*

Consider the case where the switched system (2) is a network of $k \geq 2$ interconnected subsystems. For $p \in \mathcal{P}$ and $j \in \{1,\dots,k\}$, we denote by $x_j \in \mathbb{R}^{n_j}$ the state of the j-th subsystem and by $f_p^j(x_1,\dots,x_k)$ the system function of its mode p. Then (2) can be written as the *switched interconnected system*

$$\dot{x}_j = f_\sigma^j(x_1,\dots,x_k), \qquad j \in \{1,\dots,k\} \tag{54}$$

with $n = n_1 + \dots + n_k$, $x = (x_1,\dots,x_k)$, and $f_p(x) = (f_p^1(x),\dots,f_p^k(x))$ for $p \in \mathcal{P}$.

In this section, we show that the general upper bound (32) for the topological entropy of (2) can be readily used to construct upper bounds for the entropy of (54) that depend only on "network-level" information.

Following [45], we assume that the following norms are given:

1) a "local" norm $|\cdot|_j$ on $\mathbb{R}^{n_j}$ for $j \in \{1,\dots,k\}$, and
2) a monotone "network" norm $|\cdot|_{\mathrm{N}}$ on $\mathbb{R}^k$, that is, for all nonnegative vectors $v, w \in \mathbb{R}^k_{\geq 0}$, we have

$$v \geq w \implies |v|_{\mathrm{N}} \geq |w|_{\mathrm{N}}.$$

In particular, all p-norms with $p \geq 1$ are monotone.

For a vector $v = (v_1,\dots,v_k) \in \mathbb{R}^n$ with $v_j \in \mathbb{R}^{n_j}$ for $j \in \{1,\dots,k\}$, we define a "global" norm $|\cdot|_{\mathrm{G}}$ by (one can show that $|\cdot|_{\mathrm{G}}$ is indeed a norm as $|\cdot|_{\mathrm{N}}$ is monotone)

$$|v|_{\mathrm{G}} := |(|v_1|_1,\dots,|v_k|_k)|_{\mathrm{N}}. \tag{55}$$

We denote by $\|\cdot\|_j$, $\|\cdot\|_{\mathrm{N}}$, and $\|\cdot\|_{\mathrm{G}}$ the corresponding induced norms on $\mathbb{R}^{n_j\times n_j}$, $\mathbb{R}^{k\times k}$, and $\mathbb{R}^{n\times n}$, and by μ_j, μ_{N}, and μ_{G} the corresponding matrix measures. We also denote by $\|\cdot\|_{i,j}$ the induced norm on $\mathbb{R}^{n_i\times n_j}$ defined by

$$\|A\|_{i,j} := \max_{|v|_j=1} |Av|_i.$$

As the "network" norm $|\cdot|_{\mathrm{N}}$ is monotone, the corresponding induced norm $\|\cdot\|_{\mathrm{N}}$ and matrix measure μ_{N} satisfy similar monotonicity properties for nonnegative and Metzler matrices, respectively, as shown in the following result.

Lemma 6. *1) For all nonnegative matrices $A, B \in \mathbb{R}^{k\times k}_{\geq 0}$, we have*

$$A \geq B \implies \|A\|_{\mathrm{N}} \geq \|B\|_{\mathrm{N}}. \tag{56}$$

2) For all Metzler matrices $A, B \in \mathbb{R}^{k\times k}$, that is, matrices with nonnegative off-diagonal entries, we have

$$A \geq B \implies \mu_{\mathrm{N}}(A) \geq \mu_{\mathrm{N}}(B). \tag{57}$$

Proof. 1) The implication in (56) follows from the definition of induced norm and the monotonicity of $|\cdot|_{\mathrm{N}}$. More specifically, let $\bar{v} \in \arg\max_{|v|_{\mathrm{N}}=1} |Bv|_{\mathrm{N}}$. Then $\bar{v} \geq 0$ as $B \geq 0$ and $|\cdot|_{\mathrm{N}}$ is monotone. Hence $\|A\|_{\mathrm{N}} \geq |A\bar{v}|_{\mathrm{N}} \geq |B\bar{v}|_{\mathrm{N}} = \|B\|_{\mathrm{N}}$ as $A \geq B \geq 0$ and $|\cdot|_{\mathrm{N}}$ is monotone.

2) The implication in (57) follows from (56) and the definition of matrix measure (21). More specifically, we have

$$\mu_{\mathrm{N}}(A) = \lim_{t\searrow 0} \frac{\|I+tA\|_{\mathrm{N}}-1}{t} \geq \lim_{t\searrow 0} \frac{\|I+tB\|_{\mathrm{N}}-1}{t} = \mu_{\mathrm{N}}(B),$$

where the inequality follows from (56) as $I + tA \geq I + tB \geq 0$ for the Metzler matrices $A \geq B$ and small enough $t > 0$. □

Corollary 7. *The topological entropy of the switched interconnected system (54) is upper bounded by*

$$h(f_\sigma, K) \leq \max\left\{\limsup_{t\to\infty} \sum_{p\in\mathcal{P}} n\hat{\mu}_p^{\mathrm{N}} \rho_p(t),\, 0\right\} \tag{58}$$

with the active rates $\rho_p(t)$ defined by (16) and

$$\hat{\mu}_p^{\mathrm{N}} := \limsup_{t\to\infty:\,\sigma(t)=p}\ \max_{v\in\xi_\sigma(t,\mathrm{co}(K))} \mu_{\mathrm{N}}(A_p^{\mathrm{N}}(v)), \qquad p\in\mathcal{P}, \tag{59}$$

where the matrix-valued functions $A_p^{\mathrm{N}}(v) = [a_p^{ij}(v)] \in \mathbb{R}^{k\times k}$ are defined by

$$a_p^{ii}(v) := \mu_i(J_{x_i} f_p^i(v)), \quad a_p^{ij}(v) := \|J_{x_j} f_p^i(v)\|_{i,j}, \qquad p\in\mathcal{P}, i,j\in\{1,\dots,k\} : i\neq j. \tag{60}$$

Proof. The upper bound (58) follows from the general upper bound (32) as the constants $\hat{\mu}_p$ defined by (33) with the matrix measure μ_{G} satisfy $\hat{\mu}_p \leq \hat{\mu}_p^{\mathrm{N}}$ for all $p \in \mathcal{P}$, which can be established using Lemma 7 below. □

Lemma 7 ([45, Th. 2])**.** *Consider a block matrix $A = [A_{ij}] \in \mathbb{R}^{n\times n}$ with $A_{ij} \in \mathbb{R}^{n_i\times n_j}$ for $i,j \in \{1,\dots,k\}$. Let $A_{\mathrm{N}} = [a_{ij}] \in \mathbb{R}^{k\times k}$ be a Metzler matrix such that*

$$a_{ii} \geq \mu_i(A_{ii}), \quad a_{ij} \geq \|A_{ij}\|_{i,j} \qquad \forall\, i,j\in\{1,\dots,k\} : i\neq j.$$

Then

$$\mu_{\mathrm{G}}(A) \leq \mu_{\mathrm{N}}(A_{\mathrm{N}}). \tag{61}$$

The next result shows that an upper bound similar to (58) holds when the upper limit and maximum in (59) are taken entrywise.

Corollary 8. *The topological entropy of the switched interconnected system (54) is upper bounded by*

$$h(f_\sigma, K) \leq \max\left\{\limsup_{t\to\infty} \sum_{p\in\mathcal{P}} n\mu_{\mathrm{N}}(\hat{A}_p^{\mathrm{N}})\rho_p(t),\, 0\right\} \tag{62}$$

with the active rates $\rho_p(t)$ defined by (16) and the matrices $\hat{A}_p^{\mathrm{N}} = [\hat{a}_p^{ij}] \in \mathbb{R}^{k\times k}$ defined by

$$\hat{a}_p^{ij} := \limsup_{t\to\infty:\,\sigma(t)=p}\ \max_{v\in\xi_\sigma(t,\mathrm{co}(K))} a_p^{ij}(v), \qquad p\in\mathcal{P}, i,j\in\{1,\dots,k\} \tag{63}$$

where the functions $a_p^{ij}(v)$ are defined by (60).

Proof. Note that the matrices $A_p^{\mathrm{N}}(v)$ defined by (60) and $\hat{A}_p^{\mathrm{N}}$ satisfying (63) are all Metzler matrices. Also, the upper limit and maximum in (63) imply that, for an arbitrary $\delta > 0$, there is a large enough $t_\delta \geq 0$ such that

$$\max_{v\in\xi_\sigma(t,\mathrm{co}(K))} a_p^{ij}(v) \leq \hat{a}_p^{ij} + \delta$$

for all $p \in \mathcal{P}$, $i, j \in \{1, \dots, k\}$, and $t > t_\delta$ such that $\sigma(t) = p$. Hence (57) implies that

$$\max_{v\in\xi_\sigma(t,\mathrm{co}(K))} \mu_{\mathrm{N}}(A_p^{\mathrm{N}}(v)) \leq \mu_{\mathrm{N}}(\hat{A}_p^{\mathrm{N}} + \delta\mathbf{1}) \leq \mu_{\mathrm{N}}(\hat{A}_p^{\mathrm{N}}) + \delta\mu_{\mathrm{N}}(\mathbf{1})$$

for all $p \in \mathcal{P}$ and $t > t_\delta$ such that $\sigma(t) = p$, where the last inequality follows from (23). Therefore, the constants $\hat{\mu}_p^{\mathrm{N}}$ defined by (59) satisfy $\hat{\mu}_p^{\mathrm{N}} \leq \mu_{\mathrm{N}}(\hat{A}_p^{\mathrm{N}})$ for all $p \in \mathcal{P}$ as $\delta > 0$ is arbitrary. Hence (62) follows from the upper bound (58). □

Thinking of the non-switched case as a switched system with a constant switched signal, Corollaries 7 and 8 imply the following upper bounds for the entropy of a time-invariant interconnected system.

Corollary 9. *The topological entropy of a time-invariant interconnected system*

$$\dot{x}_j = f^j(x_1, \dots, x_k), \qquad j \in \{1, \dots, k\} \tag{64}$$

with $x_j \in \mathbb{R}^{n_j}$, $n = n_1 + \cdots + n_k$, $x = (x_1, \dots, x_k)$, and $f(x) = (f^1(x), \dots, f^k(x))$ is upper bounded by

$$h(f, K) \leq \max\left\{\limsup_{t\to\infty} \max_{v\in\xi(t,\mathrm{co}(K))} n\mu_{\mathrm{N}}(A^{\mathrm{N}}(v)),\, 0\right\} \leq \max\{n\mu_{\mathrm{N}}(\hat{A}^{\mathrm{N}}),\, 0\} \tag{65}$$

with the matrix-valued function $A^{\mathrm{N}}(v) = [a^{ij}(v)] \in \mathbb{R}^{k\times k}$ defined by

$$a^{ii}(v) := \mu_i(J_{x_i} f^i(v)), \quad a^{ij}(v) := \|J_{x_j} f^i(v)\|_{i,j}, \qquad i, j \in \{1, \dots, k\} : i \neq j, \tag{66}$$

and the matrix $\hat{A}^{\mathrm{N}} = [\hat{a}^{ij}] \in \mathbb{R}^{k\times k}$ defined by

$$\hat{a}^{ij} := \limsup_{t\to\infty} \max_{v\in\xi(t,\mathrm{co}(K))} a^{ij}(v), \qquad i, j \in \{1, \dots, k\}. \tag{67}$$

A similar upper bound for the entropy of (64) was constructed in [46, Th. 1]:

$$h(f, K) \leq \max\{n\lambda_{\max}(\tilde{A}^{\mathrm{N}}),\, 0\}, \tag{68}$$

where $\lambda_{\max}(\tilde{A}^{\mathrm{N}})$ is the largest real part of the eigenvalues of a matrix $\tilde{A}^{\mathrm{N}} \in \mathbb{R}^{k\times k}$ satisfying

$$\tilde{A}^{\mathrm{N}} \geq A^{\mathrm{N}}(v) \qquad \forall\, v \in \mathbb{R}^n$$

with the function $A^{\mathrm{N}}(v) = [a^{ij}(v)]$ defined by (66).[2] The relations between the upper bounds given by the inequalities in (65) and the upper bound (68) are undetermined. On one hand, (68) may be less conservative due to the second inequality in (22). On the other hand, (68) may be more conservative as the matrix $\hat{A}^{\mathrm{N}} = [\hat{a}^{ij}]$ defined by (67) depends on $a^{ij}(v)$ over only the ω-limit set from $\mathrm{co}(K)$, while $\tilde{A}^{\mathrm{N}}$ depends on $a^{ij}(v)$ over entire $\mathbb{R}^n$.

Based on the upper bounds (58) and (62), we construct the following upper bounds for the entropy of (54), which require less information about the switching signal; the proof is along the lines of that of Corollary 3 and thus omitted here. The relations between them are similar to those between (32), (38), and (39) described in Remark 7.

Corollary 10. *The topological entropy of the switched interconnected system (54) is upper bounded by*

$$h(f_\sigma, K) \leq \sum_{p\in\mathcal{P}} \max\{n\hat{\mu}_p^{\mathrm{N}},\, 0\}\, \hat{\rho}_p \leq \sum_{p\in\mathcal{P}} \max\{n\mu_{\mathrm{N}}(\hat{A}_p^{\mathrm{N}}),\, 0\}\, \hat{\rho}_p \tag{69}$$

with the asymptotic active rates $\hat{\rho}_p$ defined by (17), and also by

$$h(f_\sigma, K) \leq \max_{p\in\mathcal{P}} \max\{n\hat{\mu}_p^{\mathrm{N}},\, 0\} \leq \max_{p\in\mathcal{P}} \max\{n\mu_{\mathrm{N}}(\hat{A}_p^{\mathrm{N}}),\, 0\}, \tag{70}$$

where the constants $\hat{\mu}_p^{\mathrm{N}}$ are defined by (59) and the matrices $\hat{A}_p^{\mathrm{N}}$ are defined by (63).

[2] Note that $\tilde{A}^{\mathrm{N}}$ is a Metzler matrix; thus its eigenvalue with the largest real part is real [**?**, Th. 10.2, p. 167].

C. Switched block-diagonal systems

Consider the case where the dynamics of the k subsystems of the switched interconnected system (54) are independent. Specifically, for each $p \in \mathcal{P}$ and $j \in \{1, \dots, k\}$, the system function $f_p^j(x_1, \dots, x_k)$ of the j-th subsystem depends only on its own state x_j. For brevity, we regard $f_p^j(x_1, \dots, x_k)$ as a function on $\mathbb{R}^{n_j}$ and denote it by $f_p^j(x_j)$. Then (54) can be written as the *switched block-diagonal system*

$$\dot{x}_j = f_\sigma^j(x_j), \qquad j \in \{1, \dots, k\} \tag{71}$$

with $f_p(x) = (f_p^1(x_1), \dots, f_p^k(x_k))$ for $p \in \mathcal{P}$. Clearly, the solution to the j-th subsystem depends only on its own initial state x_j. For brevity, we denote it by $\xi_\sigma^j(t, x_j)$ and the corresponding reachable set by $\xi_\sigma^j(t, K_j)$, where $K_j \subset \mathbb{R}^{n_j}$ is the corresponding projection of the initial set $K \subset \mathbb{R}^n$.

In this subsection, we construct bounds for the topological entropy of (71) that are usually less conservative than the results of applying the general bounds for entropy from Section V-B. Let the "local", "network", "global" norms $|\cdot|_j$ for $j \in \{1, \dots, k\}$, $|\cdot|_{\mathrm{N}}$, $|\cdot|_{\mathrm{G}}$, as well as the corresponding induced norms $\|\cdot\|_j$, $\|\cdot\|_{\mathrm{N}}$, and $\|\cdot\|_{\mathrm{G}}$ and matrix measures μ_j, μ_{N}, and μ_{G} be given as in Section V-B.

Proposition 3. *The topological entropy of the switched block-diagonal system* (71) *is upper bounded by*

$$h(f_\sigma, K) \le \sum_{j=1}^{k} \max\left\{ \limsup_{t\to\infty} \sum_{p\in\mathcal{P}} n_j \hat{a}_p^j \rho_p(t),\, 0 \right\} \tag{72}$$

with

$$\hat{a}_p^j := \limsup_{t\to\infty:\, \sigma(t)=p} \max_{v_j \in \xi_\sigma^j(t, \mathrm{co}(K_j))} \mu_j(J_{x_j} f_p^j(v_j)), \qquad p \in \mathcal{P}, j \in \{1, \dots, k\} \tag{73}$$

and the active rates $\rho_p(t)$ defined by (16).

Proof. Given arbitrary time horizon $T \ge 0$ and radius $\varepsilon > 0$, by applying the upper bound (25) to each subsystem of (71) with the corresponding "local" norm and matrix measure, we obtain that, for all initial states $x = (x_1, \dots, x_k), \bar{x} = (\bar{x}_1, \dots, \bar{x}_k) \in K$ with $x_j, \bar{x}_j \in \mathbb{R}^{n_j}$ for $j \in \{1, \dots, k\}$, the corresponding solutions satisfy

$$\max_{t\in[0,T]} |\xi_\sigma^j(t, \bar{x}_j) - \xi_\sigma^j(t, x_j)|_j \le \max_{t\in[0,T]} e^{\overline{\eta}_\sigma^j(t)} |\bar{x}_j - x_j|_j \qquad \forall\, j \in \{1, \dots, k\},$$

with

$$\overline{\eta}_\sigma^j(t) := \max_{v_j \in \mathrm{co}(K_j)} \sum_{p\in\mathcal{P}} \int_0^t \mu_j(J_{x_j} f_p^j(\xi_\sigma^j(s, v_j))\, \mathbb{1}_p(\sigma(s))\, \mathrm{d}s, \qquad j \in \{1, \dots, k\}.$$

Following the equivalence of norms on a finite-dimensional vector space, there are constants $r_1, \dots, r_k, r_{\mathrm{N}} > 0$ such that

$$|v_j|_j \le r_j |v_j|_\infty \qquad \forall\, j \in \{1, \dots, k\}, v_j \in \mathbb{R}^{n_j}$$

and

$$|v|_{\mathrm{N}} \le r_{\mathrm{N}} |v|_\infty \qquad \forall\, v \in \mathbb{R}^k.$$

Then the definition (55) of the "global" norm implies that

$$\begin{aligned} \max_{t\in[0,T]} |\xi_\sigma(t, \bar{x}) - \xi_\sigma(t, x)|_{\mathrm{G}} &\le \max_{t\in[0,T]} \max_{1\le j\le k} r_{\mathrm{N}} |\xi_\sigma^j(t, \bar{x}_j) - \xi_\sigma^j(t, x_j)|_j \\ &\le \max_{t\in[0,T]} \max_{1\le j\le k} r_{\mathrm{N}} e^{\overline{\eta}_\sigma^j(t)} |\bar{x}_j - x_j|_j \le \max_{t\in[0,T]} \max_{1\le j\le k} r_{\mathrm{N}} r_j e^{\overline{\eta}_\sigma^j(t)} |\bar{x}_j - x_j|_\infty. \end{aligned} \tag{74}$$

Consider the grid $G(\theta)$ defined by (8) with $\theta = (\theta_1 \mathbf{1}_{n_1}, \dots, \theta_k \mathbf{1}_{n_k}) \in \mathbb{R}^n_{>0}$ defined by

$$\theta_j := e^{-\max_{t\in[0,T]} \overline{\eta}_\sigma^j(t)} \varepsilon / (r_{\mathrm{N}} r_j), \qquad j \in \{1, \dots, k\},$$

and the corresponding hyperrectangles $R(x)$ defined by (9). Comparing (9) and (74) to (3), we see that $R(x) \subset B_{f_\sigma}(x, \varepsilon, T)$ for all $x \in G(\theta)$. Then Lemma 2 implies that $G(\theta)$ is (T, ε)-spanning. As $T \geq 0$ and $\varepsilon > 0$ are arbitrary, and all θ_i are nonincreasing in T, the upper bound (12) implies that

$$\begin{aligned} h(f_\sigma, K) &\leq \limsup_{\varepsilon \searrow 0} \limsup_{T \to \infty} \sum_{j=1}^{k} \frac{n_j \log(1/\theta_j)}{T} \\ &= \limsup_{T \to \infty} \sum_{j=1}^{k} \frac{1}{T} \max_{t \in [0,T]} n_j \overline{\eta}_\sigma^j(t) + \lim_{\varepsilon \searrow 0} \lim_{T \to \infty} \sum_{j=1}^{k} \frac{n_j \log(r_{\mathrm{N}} r_j/\varepsilon)}{T} \\ &\leq \sum_{j=1}^{k} \limsup_{T \to \infty} \frac{1}{T} \max_{t \in [0,T]} \sum_{p \in \mathcal{P}} \int_0^t \Big(\max_{v_j \in \xi_\sigma^j(s, \mathrm{co}(K_j))} n_j \mu_j (J_{x_j} f_p^j(v_j)) \Big) \mathbb{1}_p(\sigma(s)) \, \mathrm{d}s, \end{aligned}$$

where the equality holds as the limits in the second term on its right-hand side exist, and the last inequality holds partially because the upper limit and maximum are subadditive functions. Then we obtain (72) by applying (18) and the second inequality in (19) to each subsystem with the functions

$$a_p^j(t) = \max_{v_j \in \xi_\sigma^j(s, \mathrm{co}(K_j))} \mu_j(J_{x_j} f_p^j(v_j)), \qquad p \in \mathcal{P}, j \in \{1, \dots, k\}$$

in place of $a_p(t)$. $\square$

Thinking of the non-switched case as a switched system with a constant switched signal, Proposition 3 implies the following upper bound for the entropy of a time-invariant block-diagonal system, which is less conservative than or equivalent to the result of applying the upper bounds in (65).

Corollary 11. *The topological entropy of a time-invariant block-diagonal system $\dot{x}_j = f^j(x_j)$, $j \in \{1, \dots, k\}$ with $x_j \in \mathbb{R}^{n_j}$, $n = n_1 + \dots + n_k$, $x = (x_1, \dots, x_k)$, and $f(x) = (f^1(x_1), \dots, f^k(x_k))$ is upper bounded by*

$$h(f, K) \leq \sum_{j=1}^{k} \max \Big\{ \limsup_{t \to \infty} \max_{v_j \in \xi^j(t, \mathrm{co}(K_j))} n_j \mu_j (J_{x_j} f^j(v_j)), \, 0 \Big\}. \tag{75}$$

Based on the upper bound (72), we construct the following upper bounds for the entropy of (71), which require less information about the switching signal; the proof is along the lines of that of Corollary 3 and thus omitted here.

Corollary 12. *The topological entropy of the switched block-diagonal system (71) is upper bounded by*

$$h(f_\sigma, K) \leq \sum_{p \in \mathcal{P}} \Bigg(\sum_{j=1}^{k} \max\{n_j \hat{a}_p^j, \, 0\} \Bigg) \hat{\rho}_p \tag{76}$$

with the asymptotic active rates $\hat{\rho}_p$ defined by (17), and also by

$$h(f_\sigma, K) \leq \sum_{j=1}^{k} \max_{p \in \mathcal{P}} \max\{n_j \hat{a}_p^j, \, 0\}, \tag{77}$$

where the constants $\hat{a}_p^j$ are defined by (73).

Note that the upper bounds (72), (76), and (77) generalize the upper bounds (47), (52), and (53) to the case with block-diagonal modes, respectively. The relations between the former three are similar to those between the latter three described in Remark 9, and are illustrated by Example 3 in Section VI.

VI. Numerical examples

Consider the following switched nonlinear system in the nonnegative orthant $\mathbb{R}^n_{\geq 0}$ from [47]:

$$\dot{x} = f_\sigma(x) := (r_\sigma + A_\sigma x) \circ x \tag{78}$$

with the state $x \in \mathbb{R}^n_{\geq 0}$, a switching signal $\sigma : \mathbb{R}_{\geq 0} \to \mathcal{P}$, and a finite index set $\mathcal{P}$, where $\circ$ denotes the Hadamard (entrywise) product. Equivalently, (78) can be written as

$$\dot{x}_i = \left(r^i_\sigma + \sum_{j=1}^n a^{ij}_\sigma x_j \right) x_i, \qquad i \in \{1, \dots, n\}$$

with $x = (x_1, \dots, x_n)$, $r_p = (r^1_p, \dots, r^n_p)$, and $A_p = [a^{ij}_p]$ for $p \in \mathcal{P}$. Each mode p of (78) is a Lotka–Volterra ecosystem model that describes the population dynamics of n species in a biological community [48, Ch. 5], where x_i denotes the population density of the i-th species, $r^i_p \in \mathbb{R}$ quantifies its intrinsic growth rate, $a^{ii}_p < 0$ is a self-interaction term justified by the limitation of resources in the environment, and $a^{ij}_p \in \mathbb{R}$ for $j \neq i$ is an interaction term quantifying the influence of the j-th species on the i-th one. Switching in (78) may be due to seasonal changes or environment fluctuations. Clearly, $\mathbb{R}^n_{\geq 0}$ is a positively invariant set for (78).

We construct two switching signals σ_1 and σ_2 as follows[3]:

- σ_1 with periodic switching: Let $t_k := 1000k$ for $k \geq 1$. Then simple computation yields $\hat{\rho}_1 = \hat{\rho}_2 = 0.5$.
- σ_2 with constant set-points: Let $t_1 := 1$, and $t_{2k} := \min\{t > t_{2k-1} : \rho_2(t) \geq 0.9\}$ and $t_{2k+1} := \min\{t > t_{2k} : \rho_1(t) \geq 0.9\}$ for $k \geq 1$. Then simple computation yields $t_k = 9^{k-1} + 9^{k-2}$ for $k \geq 2$ and $\hat{\rho}_1 = \hat{\rho}_2 = 0.9$.

A. A general case

Suppose that the matrices $A_p = [a^{ij}_p]$ in the switched Lotka–Volterra system (78) satisfy

$$a^{ii}_p + \sum_{j \neq i} |a^{ij}_p| < 0, \quad a^{ii}_p + \sum_{j \neq i} |a^{ji}_p| < 0 \qquad \forall\, p \in \mathcal{P},\, i \in \{1, \dots, n\}. \tag{79}$$

Then [47, Th. 3] implies that (78) is *uniformly ultimately bounded (UUB)* in $\mathbb{R}^n_{\geq 0}$, and its ω-limit set is a subset of[4]

$$S := \prod_{i=1}^n \left[0, \max_{p \in \mathcal{P}} \max \left\{ -\frac{2r^i_p}{\lambda_{\max}(A_p + A_p^\top)}, 0 \right\} \right], \tag{80}$$

where $\lambda_{\max}(A_p + A_p^\top) < 0$ denotes the largest eigenvalue of the symmetric matrix $A_p + A_p^\top$, in which the inequality follows from (79), the formula (24), and the second inequality in (22).

Example 1. Consider the switched system (78) in $\mathbb{R}^2_{\geq 0}$ with the index set $\mathcal{P} = \{1,\, 2\}$ and the coefficients

$$r_1 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}, \qquad r_2 = \begin{bmatrix} 3 \\ -1 \end{bmatrix}, \qquad A_1 = A_2 = \begin{bmatrix} -1 & 0.1 \\ 0.1 & -1 \end{bmatrix}.$$

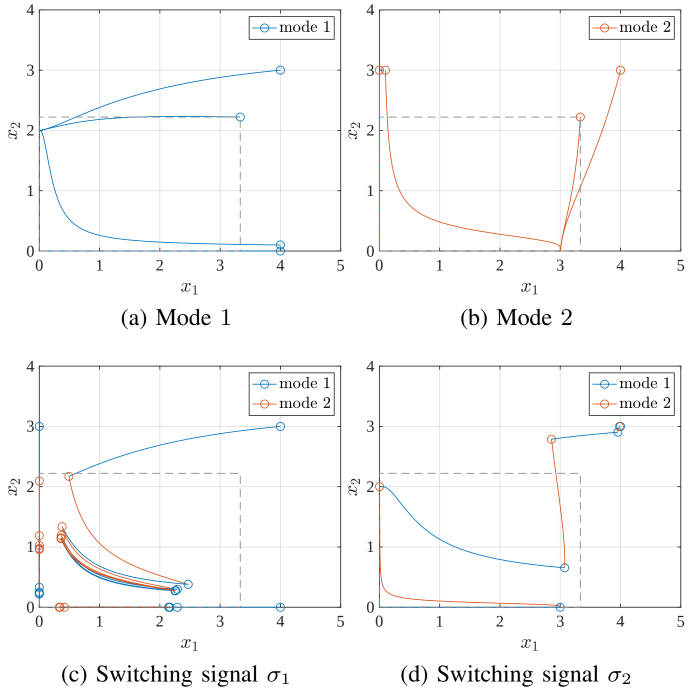
Clearly, mode 1 has an attractor $(0, 2)$ and a saddle point $(0, 0)$ with the stable manifold $\mathbb{R}_{\geq 0} \times \{0\}$, and mode 2 has an attractor $(3, 0)$ and a saddle point $(0, 0)$ with the stable manifold $\{0\} \times \mathbb{R}_{\geq 0}$. Moreover, the condition (79) holds and the set defined by (80) is given by $S = [0, 10/3] \times [0, 20/9]$. Typical trajectories for the individual modes 1 and 2 and the switching signals σ_1 and σ_2 are plotted in Fig. 1. In particular, S is not a positively invariant set.

The Jacobian matrices of individual modes are given by

$$J_x f_1(v) = \begin{bmatrix} -1 - 2v_1 + 0.1v_2 & 0.1v_1 \\ 0.1v_2 & 2 + 0.1v_1 - 2v_2 \end{bmatrix},$$
$$J_x f_2(v) = \begin{bmatrix} 3 - 2v_1 + 0.1v_2 & 0.1v_1 \\ 0.1v_2 & -1 + 0.1v_1 - 2v_2 \end{bmatrix}$$

[3] We denote by $0 < t_1 < t_2 < \cdots$ the sequence of switches and let $t_0 := 0$, with $\sigma(t) = 1$ on $[t_{2k}, t_{2k+1})$ and $\sigma(t) = 2$ on $[t_{2k+1}, t_{2k+2})$.

[4] Specifically, following the proof of [47, Th. 3], S contains the ω-limit set of (78) if for all $p \in \mathcal{P}$ and $x \in \mathbb{R}^n_{\geq 0} \backslash (S \cup \{0\})$, we have $\mathbf{1}_n^\top f_p(x) < 0$, i.e., $r_p^\top x + x^\top A_p x < 0$. Note that $r_p^\top x + x^\top A_p x = r_p^\top x + x^\top (A_p + A_p^\top) x/2 \leq (r_p + \lambda_{\max}(A_p + A_p^\top) x/2)^\top x$, which is negative if $2r^i_p + \lambda_{\max}(A_p + A_p^\top) x_i < 0$ for all $i \in \{1, \dots, n\}$ as $x \in \mathbb{R}^n_{\geq 0} \backslash \{0\}$.



(a) Mode 1 (b) Mode 2

(c) Switching signal σ_1 (d) Switching signal σ_2

Fig. 1. Trajectories of the switched system in Example 1, for (a) mode 1 with initial states $(4,3)$, $(4,0.1)$, $(4,0)$, and $(10/3,20/9)$; (b) mode 2 with initial states $(4,3)$, $(0.1,3)$, $(0,3)$, and $(10/3,20/9)$; (c) switching signal σ_1 with initial states $(4,3)$, $(4,0)$, and $(0,3)$; (d) switching signal σ_2 with initial state $(4,3)$. The circles mark the beginning of segments after switching. The dashed rectangles represent $S=[0,10/3]\times[0,20/9]$.

with $v=(v_1,v_2)\in\mathbb{R}^2_{\geq 0}$. As the switched system is UUB and its ω-limit set is a subset of S, for all initial sets $K\subset\mathbb{R}^2_{\geq 0}$, one can obtain upper bounds for the constants $\hat{\mu}_p$ defined by (33) by replacing the upper limit over $\{t\geq 0:\sigma(s)=p\}$ and maximum over $\xi_\sigma(t,\mathrm{co}(K))$ with a maximum over S, that is,

$$\hat{\mu}_1 \leq \max_{v\in S}\mu(J_x f_1(v)) = \max_{v\in S}\max\{-1-1.9v_1+0.1v_2,\ 2+0.1v_1-1.9v_2\}\leq 7/3,$$
$$\hat{\mu}_2 \leq \max_{v\in S}\mu(J_x f_2(v)) = \max_{v\in S}\max\{3-1.9v_1+0.1v_2,\ -1+0.1v_1-1.9v_2\}\leq 29/9,$$

where the equalities follow from the formula for matrix measure (24).

The upper bounds for $h(f_{\sigma_1},K)$ and $h(f_{\sigma_2},K)$ computed using (32), (38), and (39) for all initial sets $K\subset\mathbb{R}^2_{\geq 0}$ are summarized in Table I below.[5] In particular, the upper bound (32) for $h(f_{\sigma_2},K)$ can be computed as follows:

$$\begin{aligned} h(f_{\sigma_2},K) &\leq \limsup_{t\to\infty}\left(2\hat{\mu}_1\rho_1(t)+2\hat{\mu}_2\rho_2(t)\right)\\ &\leq \limsup_{t\to\infty} 2\left(\frac{7}{3}(1-\rho_2(t))+\frac{29}{9}\rho_2(t)\right) = 2\left(\frac{7}{3}+\left(\frac{29}{9}-\frac{7}{3}\right)\hat{\rho}_2\right)\approx 6.27.\end{aligned}$$

The values in Table I are consistent with the relations between (32), (38), and (39) described in Remark 7.

TABLE I
UPPER BOUNDS FOR THE ENTROPY OF THE SWITCHED SYSTEM IN EXAMPLE 1.

	$(\hat{\rho}_1,\hat{\rho}_2)$	(32)	(38)	(39)
σ_1	$(0.5,0.5)$	5.56	5.56	6.45
σ_2	$(0.9,0.9)$	6.27	10	6.45

[5]In this example, we use the set S defined by (80) which contains the ω-limit set for every switched system (78) satisfying (79). For a given family of coefficients, one can usually construct more precise over-approximation of the ω-limit set and thus obtain less conservative upper bounds for entropy, such as those in [37, Example 3.6].

B. A diagonal case

Consider the case where the switched Lotka–Volterra system (78) describes a biological community of n species with independent dynamics, that is, the interaction terms in (78) satisfy $a_p^{ij} = 0$ for all $p \in \mathcal{P}$ and $j \neq i$. Then (78) can be written as the switched diagonal system

$$\dot{x}_i = f_\sigma^i(x_i) := (r_\sigma^i + a_\sigma^{ii} x_i)\, x_i, \qquad i \in \{1, \ldots, n\} \tag{81}$$

with $f_p(x) = (f_p^1(x_1), \ldots, f_p^n(x_n))$ for $p \in \mathcal{P}$.

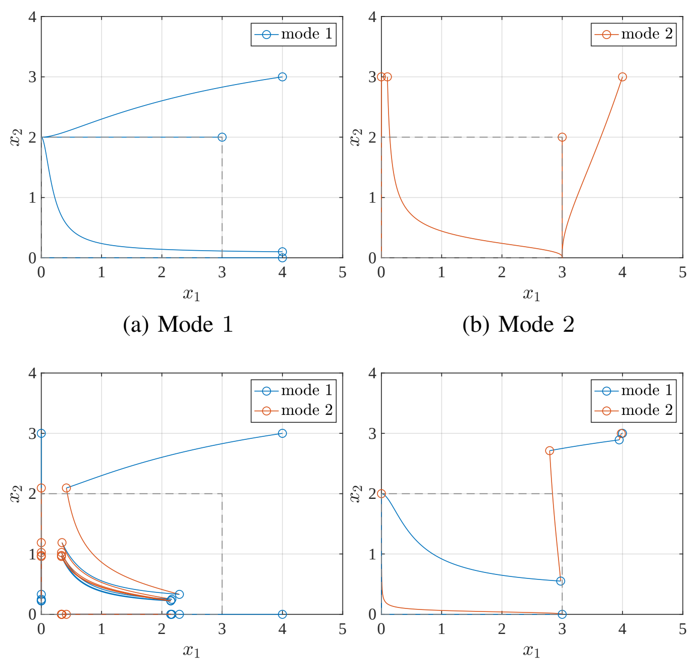
For each $p \in \mathcal{P}$ and $i \in \{1, \ldots, n\}$, we have $f_p^i(x_i) < 0$ if $x_i > \max\{-r_p^i / a_p^{ii}, 0\}$. Hence [47, Th. 3] implies that (81) is UUB in $\mathbb{R}^n_{\geq 0}$, and its ω-limit set is a subset of the positively invariant set

$$S := \prod_{i=1}^{n} \left[0, \max_{p \in \mathcal{P}} \max\left\{ -\frac{r_p^i}{a_p^{ii}}, 0\right\}\right]. \tag{82}$$

Example 2. Consider the switched diagonal system (81) in $\mathbb{R}^2_{\geq 0}$ with the index set $\mathcal{P} = \{1, 2\}$ and the coefficients

$$r_1 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}, \qquad r_2 = \begin{bmatrix} 3 \\ -1 \end{bmatrix}, \qquad A_1 = A_2 = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}.$$

Clearly, the individual modes have the same attractors and saddle points as those in Example 1, and the positively invariant set defined by (82) is given by $S = [0, 3] \times [0, 2]$. Typical trajectories for the individual modes 1 and 2 and the switching signals σ_1 and σ_2 are plotted in Fig. 2.



(a) Mode 1 (b) Mode 2

(c) Switching signal σ_1 (d) Switching signal σ_2

Fig. 2. Trajectories of the switched diagonal system in Example 2, for (a) mode 1 with initial states $(4, 3)$, $(4, 0.1)$, $(4, 0)$, and $(3, 2)$; (b) mode 2 with initial states $(4, 3)$, $(0.1, 3)$, $(0, 3)$, and $(3, 2)$; (c) switching signal σ_1 with initial states $(4, 3)$, $(4, 0)$, and $(0, 3)$; (d) switching signal σ_2 with initial state $(4, 3)$. The circles mark the beginning of segments after switching. The dashed rectangles represent the positively invariant set $S = [0, 3] \times [0, 2]$.

The Jacobian matrices of individual modes are given by

$$J_x f_1(v) = \begin{bmatrix} -1 - 2v_1 & 0 \\ 0 & 2 - 2v_2 \end{bmatrix}, \qquad J_x f_2(v) = \begin{bmatrix} 3 - 2v_1 & 0 \\ 0 & -1 - 2v_2 \end{bmatrix}$$

with $v = (v_1, v_2) \in \mathbb{R}^2_{\geq 0}$. As the switched diagonal system is UUB and its ω-limit set is a subset of S, for all initial sets $K \subset \mathbb{R}^2_{\geq 0}$, one can obtain upper bounds for the constants $\hat{\mu}_p$ and $\hat{a}_p^i$ defined by (33) and (48) by replacing the upper limits over $\{t \geq 0 : \sigma(s) = p\}$ and maxima over $\xi_\sigma(t, \mathrm{co}(K))$ with maxima over S, that is,

$$\hat{\mu}_1 \leq \max_{v \in S} \mu(J_x f_1(v)) = \max_{v \in S} \max\{-1 - 2v_1,\, 2 - 2v_2\} = 2,$$
$$\hat{\mu}_2 \leq \max_{v \in S} \mu(J_x f_2(v)) = \max_{v \in S} \max\{3 - 2v_1,\, -1 - 2v_2\} = 3,$$

where the equalities follow from the formula for matrix measure (24), and

$$\hat{a}_1^1 \leq \max_{v \in S} -1 - 2v_1 = -1, \qquad \hat{a}_1^2 \leq \max_{v \in S} 2 - 2v_2 = 2,$$
$$\hat{a}_2^1 \leq \max_{v \in S} 3 - 2v_1 = 3, \qquad \hat{a}_2^2 \leq \max_{v \in S} -1 - 2v_2 = -1.$$

Moreover, as S is a positively invariant set, for all initial sets $K \subset S$, one can obtain the same upper bounds for the functions $\overline{a}_p^i(t)$ defined by (42), that is,

$$\overline{a}_1^1(t) \leq -1, \quad \overline{a}_1^2(t) \leq 2, \quad \overline{a}_2^1(t) \leq 3, \quad \overline{a}_2^2(t) \leq -1 \qquad \forall\, t \geq 0.$$

The upper bounds for $h(f_{\sigma_1}, K)$ and $h(f_{\sigma_2}, K)$ computed using (32), (38), (39), (47), (52), and (53) for all initial sets $K \subset \mathbb{R}^2_{\geq 0}$, as well as (41), (46), and (51) for all $K \subset S$, are summarized in Table II below. In particular, for the case with σ_2, the upper bounds (32) and (47) are computed along the lines of the computation of (32) in Example 1; the upper bound in (41) is computed along the lines of the computation of $h(D_{\sigma_2})$ in [33, Example 3]. The values in Table II are consistent with the relations between (32), (38), (39), (41), (46), (47), (51), (52), and (53) described in Remark 9.

TABLE II
UPPER BOUNDS FOR THE ENTROPY OF THE SWITCHED DIAGONAL SYSTEM IN EXAMPLE 2.

		$K \subset \mathbb{R}^2_{\geq 0}$						$K \subset S$		
	$(\hat{\rho}_1, \hat{\rho}_2)$	(32)	(38)	(39)	(47)	(52)	(53)	(41)	(46)	(51)
σ_1	$(0.5, 0.5)$	5	5	6	1.5	2.5	5	1.5	2.5	3
σ_2	$(0.9, 0.9)$	5.8	9	6	4.3	4.5	5	2.79	2.9	3

C. A block-diagonal case

Consider the case where the switched Lotka–Volterra system (78) describes a biological community of $k \geq 2$ groups of species where the j-th group consists of n_j species for $j \in \{1, \dots, k\}$, and the dynamics of two species from different groups are independent. Regarding the groups as interconnected subsystems, we denote by $\bar{x}_j \in \mathbb{R}^{n_j}_{\geq 0}$ the state of the j-th subsystem and by $\bar{f}_p^j(\bar{x}_j)$, $\bar{r}_p^j$, and $\bar{A}_p^j$ its system function and coefficients for $p \in \mathcal{P}$ and $j \in \{1, \dots, k\}$ to avoid confusion. Then (78) can be written as the switched block-diagonal system

$$\dot{\bar{x}}_j = \bar{f}_\sigma^j(\bar{x}_j) := (\bar{r}_\sigma^j + \bar{A}_\sigma^j \bar{x}_j) \circ \bar{x}_j, \qquad j \in \{1, \dots, k\} \tag{83}$$

with $n = n_1 + \cdots + n_k$, $x = (\bar{x}_1, \dots, \bar{x}_k)$, $f_p(x) = (\bar{f}_p^1(\bar{x}_1), \dots, \bar{f}_p^k(\bar{x}_k))$, $r_p = (\bar{r}_p^1, \dots, \bar{r}_p^k)$, and $A_p = \mathrm{diag}(\bar{A}_p^1, \dots, \bar{A}_p^k)$ for $p \in \mathcal{P}$.

For (83), the formula (24) implies that the condition (79) can be written as

$$\mu(\bar{A}_p^j), \mu((\bar{A}_p^j)^\top) < 0 \qquad \forall\, p \in \mathcal{P}, j \in \{1, \dots, k\} \tag{84}$$

with the matrix measure computed using the induced ∞-norm. If (84) holds, then [47, Th. 3] implies that (83) is *uniformly ultimately bounded (UUB)* in $\mathbb{R}^n_{\geq 0}$, and its ω-limit set is a subset of

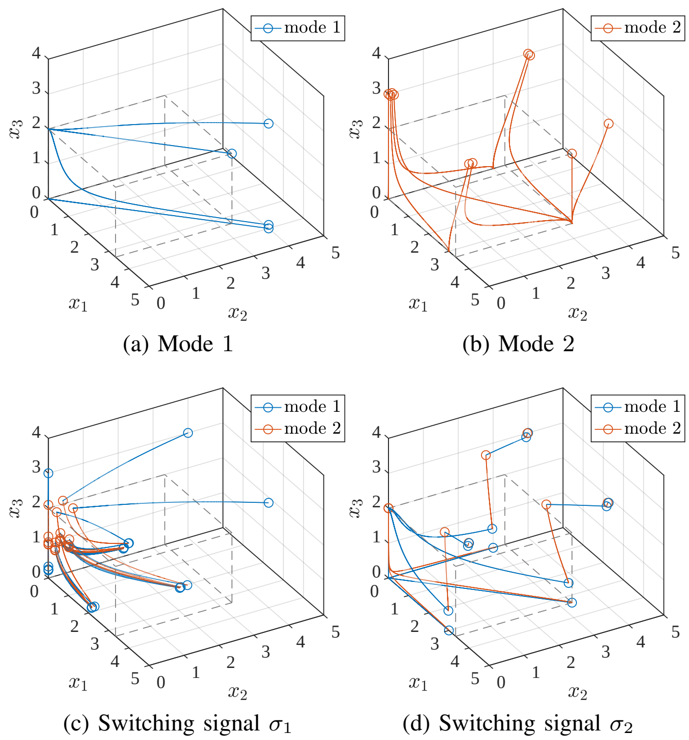
$$S := \prod_{i=1}^n \left[0, \max_{p \in \mathcal{P}} \max\left\{ -\frac{2r_p^i}{\lambda_{\max}(A_p^{j(i)} + (A_p^{j(i)})^\top)}, 0 \right\}\right], \tag{85}$$

where the i-th species is in the $j(i)$-th group, and $\lambda_{\max}(A_p^{j(i)} + (A_p^{j(i)})^\top) < 0$ denote the largest eigenvalue of the symmetric matrix $A_p^{j(i)} + (A_p^{j(i)})^\top$, in which the inequality follows from (84), (24), and the second inequality in (22).

Example 3. Consider the switched block-diagonal system (83) in $\mathbb{R}^3_{\geq 0}$ with the index set $\mathcal{P} = \{1, 2\}$ and the coefficients

$$\begin{aligned} \bar{r}_1^1 &= \begin{bmatrix} -1 \\ -1 \end{bmatrix}, & \bar{r}_2^1 &= \begin{bmatrix} 3 \\ 3 \end{bmatrix}, & \bar{A}_1^1 = \bar{A}_2^1 &= \begin{bmatrix} -1 & 0.1 \\ 0.1 & -1 \end{bmatrix}, \\ \bar{r}_1^2 &= 2, & \bar{r}_2^2 &= -1, & \bar{A}_1^2 = \bar{A}_2^2 &= -1. \end{aligned} \tag{86}$$

Simple computation shows that mode 1 has an attractor $(0, 0, 2)$ and a saddle point $(0, 0, 0)$ with the stable manifold $\mathbb{R}^2_{\geq 0} \times \{0\}$, and mode 2 has an attractor $(10/3, 10/3, 0)$ and three saddle points $(0, 0, 0)$, $(3, 0, 0)$, and $(0, 3, 0)$ with the stable manifolds $\{(0, 0)\} \times \mathbb{R}_{\geq 0}$, $\mathbb{R}_{>0} \times \{0\} \times \mathbb{R}_{\geq 0}$, and $\{0\} \times \mathbb{R}_{>0} \times \mathbb{R}_{\geq 0}$, respectively. Moreover, the condition (84) holds and the set defined by (85) is given by $S = [0, 10/3]^2 \times [0, 2]$. Typical trajectories for the individual modes 1 and 2 and the switching signals σ_1 and σ_2 are plotted in Fig. 3.



(a) Mode 1 (b) Mode 2

(c) Switching signal σ_1 (d) Switching signal σ_2

Fig. 3. Trajectories of the switched block-diagonal system in Example 3, for (a) mode 1 with initial states $(4, 4, 3)$, $(4, 4, 0.1)$, $(4, 4, 0)$, and $(10/3, 10/3, 2)$; (b) mode 2 with initial states $(4, 4, 3)$, $(4, 0.1, 3)$, $(0.1, 4, 3)$, $(4, 0, 3)$, $(0, 4, 3)$, $(0.1, 0.1, 3)$, $(0.1, 0, 3)$, $(0, 0.1, 3)$, $(0, 0, 3)$, and $(10/3, 10/3, 2)$; (c) switching signal σ_1 with initial states $(4, 4, 3)$, $(4, 0, 3)$, $(0, 4, 3)$, and $(0, 0, 3)$; (d) switching signal σ_2 with initial states $(4, 4, 3)$, $(4, 0, 3)$, and $(0, 4, 3)$. The circles mark the beginning of segments after switching. The dashed hyperrectangles represent $S = [0, 10/3]^2 \times [0, 2]$.

The Jacobian matrices of individual modes are given by

$$\begin{aligned} J_x f_1(v) &= \mathrm{diag}(J_{\bar{x}_1} \bar{f}_1^1(\bar{v}_1), J_{\bar{x}_2} \bar{f}_1^2(\bar{v}_2)), \\ J_x f_2(v) &= \mathrm{diag}(J_{\bar{x}_1} \bar{f}_2^1(\bar{v}_1), J_{\bar{x}_2} \bar{f}_2^2(\bar{v}_2)) \end{aligned}$$

with

$$\begin{aligned} J_{\bar{x}_1} \bar{f}_1^1(\bar{v}_1)) &= \begin{bmatrix} -1 - 2v_1 + 0.1v_2 & 0.1v_1 \\ 0.1v_2 & -1 + 0.1v_1 - 2v_2 \end{bmatrix}, & J_{\bar{x}_2} \bar{f}_1^2(\bar{v}_2)) &= 2 - 2v_3, \\ J_{\bar{x}_1} \bar{f}_2^1(\bar{v}_1)) &= \begin{bmatrix} 3 - 2v_1 + 0.1v_2 & 0.1v_1 \\ 0.1v_2 & 3 + 0.1v_1 - 2v_2 \end{bmatrix}, & J_{\bar{x}_2} \bar{f}_2^2(\bar{v}_2)) &= -1 - 2v_3, \end{aligned}$$

where $v = (\bar{v}_1, \bar{v}_2)$ with $\bar{v}_1 = (v_1, v_2) \in \mathbb{R}^2_{\geq 0}$, and $\bar{v}_2 = v_3 \in \mathbb{R}_{\geq 0}$. As the switched block-diagonal system is UUB and its ω-limit set is a subset of S, for all initial sets $K \subset \mathbb{R}^3_{\geq 0}$, one can obtain upper bounds for the constants $\hat{\mu}_p$ and $\hat{a}^j_p$ defined by (33) and (73) by replacing the upper limit over $\{t \geq 0 : \sigma(s) = p\}$ and maximum over $\xi_\sigma(t, \mathrm{co}(K))$ with a maximum over S, that is,

$$\begin{aligned}
\hat{a}^1_1 &\leq \max_{v \in S} \mu(J_{\bar{x}_1} \bar{f}^1_1(\bar{v}_1)) = \max_{v \in S} \max\{-1 - 1.9v_1 + 0.1v_2,\, -1 + 0.1v_1 - 1.9v_2\} \leq -2/3, \\
\hat{a}^2_1 &\leq \max_{v \in S} \mu(J_{\bar{x}_2} \bar{f}^2_1(\bar{v}_2)) = \max_{v \in S} 2 - 2v_3 = 2, \\
\hat{a}^1_2 &\leq \max_{v \in S} \mu(J_{\bar{x}_1} \bar{f}^1_2(\bar{v}_1)) = \max_{v \in S} \max\{3 - 1.9v_1 + 0.1v_2,\, 3 + 0.1v_1 - 1.9v_2\} \leq 10/3, \\
\hat{a}^2_2 &\leq \max_{v \in S} \mu(J_{\bar{x}_2} \bar{f}^2_2(\bar{v}_2)) = \max_{v \in S} -1 - 2v_3 = -1,
\end{aligned}$$

and

$$\begin{aligned}
\hat{\mu}_1 &\leq \max_{v \in S} \mu(J_x f_1(v)) = \max\{-2/3,\, 2\} = 2, \\
\hat{\mu}_2 &\leq \max_{v \in S} \mu(J_x f_2(v)) = \max\{3/10,\, -1\} = 10/3,
\end{aligned}$$

where the first equalities follow from the formula for matrix measure (24).

The upper bounds for $h(f_{\sigma_1}, K)$ and $h(f_{\sigma_2}, K)$ computed using (32), (38), (39), (72), (76), and (77) for all initial sets $K \subset \mathbb{R}^3_{\geq 0}$ are summarized in Table III below. In particular, for the case with σ_2, the upper bounds (32) and (72) are computed along the lines of the computation of (32) in Example 1.

TABLE III
UPPER BOUNDS FOR THE ENTROPY OF THE SWITCHED BLOCK-DIAGONAL SYSTEM IN EXAMPLE 3.

	$(\hat{\rho}_1, \hat{\rho}_2)$	(32)	(38)	(39)	(72)	(76)	(77)
σ_1	$(0.5, 0.5)$	8	8	10	3.17	4.34	8.67
σ_2	$(0.9, 0.9)$	9.6	14.4	10	7.57	7.8	8.67

VII. CONCLUSION

We constructed upper and lower bounds for topological entropy of switched nonlinear and interconnected systems, which generalize previous results for switched linear systems from [32], [33] and further our understanding of how switching affects topological entropy. A feature of most bounds presented here is that they only depend on the Jacobian matrices of system functions over the ω-limit set instead of all reachable points, and thus will yield finite values for the case with unbounded Jacobian matrices but a compact global attractor.

Future research directions include analyzing the computation complexities of the bounds for topological entropy in this paper, studying the relation between these bounds and stability conditions for switched nonlinear and interconnected systems, and constructing bounds for topological entropy of switched nonlinear systems with triangular and block-triangular modes.

APPENDIX A
PROOF OF LEMMA 2

1) If (10) holds, then

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

Hence $G(\theta)$ is (T, ε)-spanning following the definition (4), and thus

$$S(f_\sigma, \varepsilon, T, K) \leq \#G(\theta) \leq \prod_{i=1}^{n} \left(\left\lfloor \frac{2r_2}{\theta_i} \right\rfloor + 1 \right) \leq \prod_{i=1}^{n} \left(\frac{2r_2}{\theta_i} + 1 \right).$$

If (10) holds for all $T \geq 0$ and $\varepsilon > 0$, then the definition of entropy (5) implies that

$$h(f_\sigma, K) \leq \limsup_{\varepsilon \searrow 0} \limsup_{T \to \infty} \sum_{i=1}^{n} \frac{\log(2r_2/\theta_i + 1)}{T}$$
$$\leq \limsup_{\varepsilon \searrow 0} \limsup_{T \to \infty} \sum_{i=1}^{n} \frac{\log(1/\theta_i)}{T} + \sum_{i=1}^{n} \limsup_{\varepsilon \searrow 0} \limsup_{T \to \infty} \frac{\log(2r_2 + \theta_i)}{T},$$

where the last inequality holds as the upper limit is a subadditive function. Moreover, the summands in the last term satisfy

$$\limsup_{\varepsilon \searrow 0} \limsup_{T \to \infty} \frac{\log(2r_2 + \theta_i)}{T} \leq \limsup_{\varepsilon \searrow 0} \limsup_{T \to \infty} \frac{\max\{\log(2\theta_i),\, \log(4r_2)\}}{T}$$
$$= \max\left\{ \limsup_{\varepsilon \searrow 0} \limsup_{T \to \infty} \frac{\log(2\theta_i)}{T},\, \limsup_{T \to \infty} \frac{\log(4r_2)}{T} \right\}$$
$$= \max\left\{ \limsup_{\varepsilon \searrow 0} \limsup_{T \to \infty} \frac{\log \theta_i}{T},\, 0 \right\}$$

for all $i \in \{1, \ldots, n\}$, where the inequality holds as the logarithm is an increasing function and $r_2, \theta_i > 0$, and the last equality holds partially because r_2 is independent of T. Hence we obtain (12) if (11) holds.[6]

2) If (13) holds, then for all distinct points $x, \bar{x} \in G(\theta)$, we have $\bar{x} \notin B_{f_\sigma}(x, \varepsilon, T)$ as $\bar{x} \notin R(x)$. Hence $G(\theta)$ is (T, ε)-separated following the definition (6), and thus

$$N(f_\sigma, \varepsilon, T, K) \geq \#G(\theta) \geq \prod_{i=1}^{n} \left\lfloor \frac{2r_1}{\theta_i} \right\rfloor \geq \prod_{i=1}^{n} \max\left\{ \frac{2r_1}{\theta_i} - 1,\, 1 \right\}.$$

If (13) holds for all $T \geq 0$ and $\varepsilon > 0$, then the alternative definition of entropy (7) implies that

$$h(f_\sigma, K) \geq \liminf_{\varepsilon \searrow 0} \limsup_{T \to \infty} \sum_{i=1}^{n} \frac{\log(\max\{2r_1/\theta_i - 1,\, 1\})}{T}$$
$$\geq \liminf_{\varepsilon \searrow 0} \limsup_{T \to \infty} \sum_{i=1}^{n} \frac{\log(1/\theta_i)}{T}$$
$$+ \sum_{i=1}^{n} \liminf_{\varepsilon \searrow 0} \liminf_{T \to \infty} \frac{\log(\max\{2r_1 - \theta_i,\, \theta_i\})}{T},$$

where the last inequality holds as the lower limit is a superadditive function, and for two functions $f, g : \mathbb{R} \to \mathbb{R}$, we have

$$-\limsup_{T \to \infty} f(T) = \liminf_{T \to \infty} -f(T) = \liminf_{T \to \infty} \big(g(T) - (f(T) + g(T))\big)$$
$$\geq \liminf_{T \to \infty} g(T) + \liminf_{T \to \infty} -(f(T) + g(T))$$
$$= \liminf_{T \to \infty} g(T) - \limsup_{T \to \infty} (f(T) + g(T))$$

and thus

$$\limsup_{T \to \infty} (f(T) + g(T)) \geq \limsup_{T \to \infty} f(T) + \liminf_{T \to \infty} g(T).$$

Moreover, the summands in the last term satisfy

$$\liminf_{\varepsilon \searrow 0} \liminf_{T \to \infty} \frac{\log(\max\{2r_1 - \theta_i,\, \theta_i\})}{T} \geq \liminf_{\varepsilon \searrow 0} \liminf_{T \to \infty} \frac{\log(\max\{r_1,\, \theta_i\})}{T}$$
$$\geq \liminf_{\varepsilon \searrow 0} \liminf_{T \to \infty} \frac{\log r_1}{T} = 0$$

for all $i \in \{1, \ldots, n\}$, where the first inequality holds as the logarithm is an increasing function and $r_1, \theta_i > 0$, and the equality holds as r_1 is independent of T. Hence (14) holds. ☐

[6]There was a typo in [33, Lemma 2]: [33, eq. (17)] should be replaced with (11), as we need the inequality in (11) to hold for each $i \in \{1, \ldots, n\}$ instead of the upper limits of the sum. As an example, consider $n = 2$, $r_2 = 1$, $\theta_1 = e^T$, and $\theta_2 = e^{-2T}$. Then $\limsup_{T \to \infty} (\log \theta_1 + \log \theta_2)/T = -1 \leq 0$, while $\limsup_{T \to \infty} (\log(2 + \theta_1) + \log(2 + \theta_2))/T \geq \limsup_{T \to \infty} (\log \theta_1 + \log 2)/T = 1 > 0$.

Appendix B
Proof of Lemma 3

For brevity, we define the following functions $\eta, \bar{a} : \mathbb{R}_{\geq 0} \to \mathbb{R}$:

$$\eta(t) := \sum_{p \in \mathcal{P}} \int_0^t a_p(s)\, \mathbb{1}_p(\sigma(s))\, \mathrm{d}s, \qquad \bar{a}(T) := \frac{1}{T} \max_{t \in [0,T]} \eta(t),$$

with $\bar{a}(0) := \max\{a_{\sigma(0)}(0),\, 0\}$, and a constant

$$\hat{a} := \limsup_{t \to \infty} \frac{\eta(t)}{t} \in \mathbb{R} \cup \{\infty\}.$$

First, we prove (18), that is,

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

The definition of $\bar{a}$ implies that $\bar{a}(T) \geq 0$ and $\bar{a}(T) \geq \eta(T)/T$ for all $T > 0$. Hence

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

(in particular, if $\hat{a} = \infty$ then $\limsup_{T \to \infty} \bar{a}(T) = \infty$). It remains to prove that, when $\hat{a}$ is finite, the reverse inequality holds as well. The upper limit in the definition of $\hat{a}$ implies that, for each $\delta > 0$, there is a large enough $t_\delta \geq 0$ such that

$$\eta(t) < (\hat{a} + \delta)\, t \qquad \forall\, t > t_\delta.$$

For a $T > t_\delta$, let

$$t^*(T) \in \arg\max_{t \in [0,T]} \eta(t),$$

which exists as η is a continuous function. Then $\eta(t^*(T)) \geq \eta(0) = 0$. If $t^*(T) \in (t_\delta, T]$, then

$$\bar{a}(T) = \frac{\eta(t^*(T))}{T} \leq \frac{\eta(t^*(T))}{t^*(T)} < \hat{a} + \delta;$$

otherwise $t^*(T) \in [0, t_\delta]$, and thus $t^*(T) = t^*(t_\delta)$ and

$$\bar{a}(T) = \frac{\eta(t^*(T))}{T} = \frac{\eta(t^*(t_\delta))}{T}.$$

Combining the two cases above, we obtain

$$\bar{a}(T) \leq \max\left\{\hat{a} + \delta,\, \frac{\eta(t^*(t_\delta))}{T}\right\} \qquad \forall\, T > t_\delta.$$

Hence

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

As $\delta > 0$ is arbitrary, we have

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

and thus (18) holds.

Second, we prove (19), that is,

$$\limsup_{t \to \infty} \sum_{p \in \mathcal{P}} \check{a}_p \rho_p(t) \leq \hat{a} \leq \limsup_{t \to \infty} \sum_{p \in \mathcal{P}} \hat{a}_p \rho_p(t),$$

Recall that the index set $\mathcal{P}$ is finite. The upper and lower limits in the definitions of $\hat{a}_p$ and $\check{a}_p$ imply that, for each $\bar{\delta} > 0$, there is a large enough $t_{\bar{\delta}} \geq 0$ such that

$$\check{a}_p - \bar{\delta} \leq a_p(t) \leq \hat{a}_p + \bar{\delta} \qquad \forall\, p \in \mathcal{P}, t > t_{\bar{\delta}} : \sigma(t) = p.$$

Hence

$$\begin{aligned}\hat{a} &= \limsup_{t\to\infty} \frac{1}{t}\sum_{p\in\mathcal{P}}\int_0^t (a_p(s)-\hat{a}_p-\bar{\delta}+\hat{a}_p+\bar{\delta})\,\mathbb{1}_p(\sigma(s))\,\mathrm{d}s\\ &= \limsup_{t\to\infty}\sum_{p\in\mathcal{P}}\hat{a}_p\rho_p(t) + \limsup_{t\to\infty}\frac{1}{t}\sum_{p\in\mathcal{P}}\int_0^t (a_p(s)-\hat{a}_p-\bar{\delta})\,\mathbb{1}_p(\sigma(s))\,\mathrm{d}s+\bar{\delta}\\ &\le \limsup_{t\to\infty}\sum_{p\in\mathcal{P}}\hat{a}_p\rho_p(t) + \limsup_{t\to\infty}\frac{1}{t}\sum_{p\in\mathcal{P}}\int_0^{t_{\bar{\delta}}} (a_p(s)-\hat{a}_p-\bar{\delta})\,\mathbb{1}_p(\sigma(s))\,\mathrm{d}s+\bar{\delta}\\ &= \limsup_{t\to\infty}\sum_{p\in\mathcal{P}}\hat{a}_p\rho_p(t)+\bar{\delta}.\end{aligned}$$

and

$$\begin{aligned}\hat{a} &= \limsup_{t\to\infty} \frac{1}{t}\sum_{p\in\mathcal{P}}\int_0^t (a_p(s)-\check{a}_p+\bar{\delta}+\check{a}_p-\bar{\delta})\,\mathbb{1}_p(\sigma(s))\,\mathrm{d}s\\ &= \limsup_{t\to\infty}\sum_{p\in\mathcal{P}}\check{a}_p\rho_p(t) + \limsup_{t\to\infty}\frac{1}{t}\sum_{p\in\mathcal{P}}\int_0^t (a_p(s)-\check{a}_p+\bar{\delta})\,\mathbb{1}_p(\sigma(s))\,\mathrm{d}s-\bar{\delta}\\ &\ge \limsup_{t\to\infty}\sum_{p\in\mathcal{P}}\check{a}_p\rho_p(t) + \limsup_{t\to\infty}\frac{1}{t}\sum_{p\in\mathcal{P}}\int_0^{t_{\bar{\delta}}} (a_p(s)-\check{a}_p+\bar{\delta})\,\mathbb{1}_p(\sigma(s))\,\mathrm{d}s-\bar{\delta}\\ &= \limsup_{t\to\infty}\sum_{p\in\mathcal{P}}\check{a}_p\rho_p(t)-\bar{\delta}.\end{aligned}$$

Then (19) holds as $\bar{\delta}>0$ is arbitrary. □

Appendix C
Proof of Lemma 5

We prove Lemma 5 by writing the difference between two solutions to the switched nonlinear system (2) as a solution to the LTV system (27) with a suitable function $A(\cdot)$, using variational arguments similar to those in the proof of Lemma 4. Given arbitrary initial states $x,\bar{x}\in K$, let

$$\nu(t,\rho) := \rho\xi_\sigma(t,\bar{x}) + (1-\rho)\xi_\sigma(t,x), \qquad \rho\in[0,1], t\ge 0.$$

Then $\nu(t,\rho)\in \mathrm{co}(\xi_\sigma(t,K))$ and $\partial_\rho\nu(t,\rho)=\xi_\sigma(t,\bar{x})-\xi_\sigma(t,x)$ for all $\rho\in[0,1]$ and $t\ge 0$. Hence

$$\begin{aligned}\partial_t(\xi_\sigma(t,\bar{x})-\xi_\sigma(t,x)) &= f_{\sigma(t)}(\xi_\sigma(t,\bar{x})) - f_{\sigma(t)}(\xi_\sigma(t,x))\\ &= f_{\sigma(t)}(\nu(t,1)) - f_{\sigma(t)}(\nu(t,0))\\ &= \left(\int_0^1 J_x f_{\sigma(t)}(\nu(t,\rho))\,\mathrm{d}\rho\right)(\xi_\sigma(t,\bar{x})-\xi_\sigma(t,x))\end{aligned}$$

for all $t\ge 0$ that are not switches. Therefore, $\xi_\sigma(t,\bar{x})-\xi_\sigma(t,x)$ is the solution to (27) with $A(t)=\int_0^1 J_x f_{\sigma(t)}(\nu(t,\rho))\,\mathrm{d}\rho$, at time t with initial state $\bar{x}-x$. Consequently, (28) implies that

$$\begin{aligned}e^{\int_0^t -\mu\left(-\int_0^1 J_x f_{\sigma(s)}(\nu(s,\rho))\,\mathrm{d}\rho\right)\mathrm{d}s}|\bar{x}-x| &\le |\xi_\sigma(t,\bar{x})-\xi_\sigma(t,x)|\\ &\le e^{\int_0^t \mu\left(\int_0^1 J_x f_{\sigma(s)}(\nu(s,\rho))\,\mathrm{d}\rho\right)\mathrm{d}s}|\bar{x}-x|\end{aligned}$$

for all $t\ge 0$. Moreover, as the matrix measure $\mu(\cdot)$ is a convex function, Jensen's inequality implies that

$$\begin{aligned}\mu\left(\int_0^1 J_x f_{\sigma(t)}(\nu(t,\rho))\,\mathrm{d}\rho\right) &\le \int_0^1 \mu(J_x f_{\sigma(t)}(\nu(t,\rho))\,\mathrm{d}\rho\\ &\le \max_{v\in\mathrm{co}(\xi_\sigma(t,K))}\mu(J_x f_{\sigma(t)}(v))\end{aligned}$$

and

$$-\mu\Big(-\int_0^1 J_x f_{\sigma(t)}(\nu(t,\rho))\,\mathrm{d}\rho\Big) \geq \int_0^1 -\mu(-J_x f_{\sigma(t)}(\nu(t,\rho))\,\mathrm{d}\rho \\ \geq \min_{v\in\mathrm{co}(\xi_\sigma(t,K))} -\mu(-J_x f_{\sigma(t)}(v))$$

for all $t \geq 0$. Then (30) follows from the transformation (20). □

References

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

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

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

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

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

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

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

[8] G. N. Nair, R. J. Evans, I. M. Y. Mareels, and W. Moran, "Topological feedback entropy and nonlinear stabilization," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1585–1597, Sep. 2004.

[9] F. Colonius and C. Kawan, "Invariance entropy for control systems," *SIAM Journal on Control and Optimization*, vol. 48, no. 3, pp. 1701–1721, May 2009.

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

[11] F. Colonius, "Minimal bit rates and entropy for exponential stabilization," *SIAM Journal on Control and Optimization*, vol. 50, no. 5, pp. 2988–3010, Oct. 2012.

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

[13] A. S. Matveev and A. Y. Pogromsky, "Observation of nonlinear systems via finite capacity channels: Constructive data rate limits," *Automatica*, vol. 70, pp. 217–229, Aug. 2016.

[14] D. Liberzon and S. Mitra, "Entropy and minimal bit rates for state estimation and model detection," *IEEE Transactions on Automatic Control*, vol. 63, no. 10, pp. 3330–3344, Oct. 2018.

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

[16] R. Shorten, F. R. Wirth, O. Mason, K. Wulff, and C. King, "Stability criteria for switched and hybrid systems," *SIAM Review*, vol. 49, no. 4, pp. 545–592, Nov. 2007.

[17] J. L. Mancilla-Aguilar, "A condition for the stability of switched nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 45, no. 11, pp. 2077–2079, Nov. 2000.

[18] L. Vu and D. Liberzon, "Common Lyapunov functions for families of commuting nonlinear systems," *Systems & Control Letters*, vol. 54, no. 5, pp. 405–416, May 2005.

[19] C. A. Desoer and M. Vidyasagar, *Feedback Systems: Input-Output Properties*. SIAM, 2009.

[20] T. Liu, Z.-P. Jiang, and D. J. Hill, *Nonlinear Control of Dynamic Networks*. CRC Press, 2018.

[21] G. Yang and D. Liberzon, "A Lyapunov-based small-gain theorem for interconnected switched systems," *Systems & Control Letters*, vol. 78, pp. 47–54, Apr. 2015.

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

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

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

[25] D. Liberzon and J. P. Hespanha, "Stabilization of nonlinear systems with limited information feedback," *IEEE Transactions on Automatic Control*, vol. 50, no. 6, pp. 910–915, Jun. 2005.

[26] C. Kawan and J.-C. Delvenne, "Network entropy and data rates required for networked control," *IEEE Transactions on Control of Network Systems*, vol. 3, no. 1, pp. 57–66, Mar. 2016.

[27] A. S. Matveev, A. V. Proskurnikov, A. Pogromsky, and E. Fridman, "Comprehending complexity: Data-rate constraints in large-scale networks," *IEEE Transactions on Automatic Control*, vol. 64, no. 10, pp. 4252–4259, Oct. 2019.

[28] M. S. Tomar and M. Zamani, "Compositional quantification of invariance feedback entropy for networks of uncertain control systems," *IEEE Control Systems Letters*, vol. 4, no. 4, pp. 827–832, Oct. 2020.

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

[30] G. Yang and D. Liberzon, “Feedback stabilization of switched linear systems with unknown disturbances under data-rate constraints,” *IEEE Transactions on Automatic Control*, vol. 63, no. 7, pp. 2107–2122, Jul. 2018.

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

[32] G. Yang, A. J. Schmidt, and D. Liberzon, “On topological entropy of switched linear systems with diagonal, triangular, and general matrices,” in *57th IEEE Conference on Decision and Control*, 2018, pp. 5682–5687.

[33] G. Yang, A. J. Schmidt, D. Liberzon, and J. P. Hespanha, “Topological entropy of switched linear systems: General matrices and matrices with commutation relations,” *Mathematics of Control, Signals, and Systems*, vol. 32, no. 3, pp. 411–453, Sep. 2020.

[34] G. Yang, J. P. Hespanha, and D. Liberzon, “On topological entropy and stability of switched linear systems,” in *22nd ACM International Conference on Hybrid Systems: Computation and Control*, 2019, pp. 119–127.

[35] G. O. Berger and R. M. Jungers, “Worst-case topological entropy and minimal data rate for state observation of switched linear systems,” in *23rd International Conference on Hybrid Systems: Computation and Control*, Apr. 2020, 11 pages.

[36] G. S. Vicinansa and D. Liberzon, “Estimation entropy for regular linear switched systems,” in *58th IEEE Conference on Decision and Control*, 2019, pp. 5754–5759.

[37] G. Yang, D. Liberzon, and J. P. Hespanha, “Topological entropy of switched nonlinear systems,” in *24th ACM International Conference on Hybrid Systems: Computation and Control*, 2021, 11 pages.

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

[39] D. S. Bernstein, *Matrix Mathematics*. Princeton University Press, 2009.

[40] E. D. Sontag, “Contractive systems with inputs,” in *Perspectives in Mathematical System Theory, Control, and Signal Processing*, J. C. Willems, S. Hara, Y. Ohta, and H. Fujioka, Eds. Springer, 2010, pp. 217–228.

[41] V. A. Boichenko and G. A. Leonov, “Lyapunov’s direct method in estimates of topological entropy,” *Journal of Mathematical Sciences*, vol. 91, no. 6, pp. 3370–3379, Oct. 1998.

[42] R. W. Brockett, *Finite Dimensional Linear Systems*. SIAM, 1970.

[43] D. Liberzon, *Calculus of Variations and Optimal Control Theory*. Princeton University Press, 2012.

[44] M. Vidyasagar, *Nonlinear Systems Analysis*, 2nd ed. SIAM, 2002.

[45] G. Russo, M. di Bernardo, and E. D. Sontag, “A contraction approach to the hierarchical analysis and design of networked systems,” *IEEE Transactions on Automatic Control*, vol. 58, no. 5, pp. 1328–1331, May 2013.

[46] D. Liberzon, “On topological entropy of interconnected nonlinear systems,” *IEEE Control Systems Letters*, vol. 5, no. 6, pp. 2210–2214, Dec. 2021.

[47] A. Y. Aleksandrov, Y. Chen, A. V. Platonov, and L. Zhang, “Stability analysis for a class of switched nonlinear systems,” *Automatica*, vol. 47, no. 10, pp. 2286–2291, Oct. 2011.

[48] J. Hofbauer and K. Sigmund, *Evolutionary Games and Population Dynamics*. Cambridge University Press, 1998.