Topological entropy of switched nonlinear systems

Guosong Yang
Center for Control, Dynamical Systems, and Computation
University of California
Santa Barbara, CA, USA
guosongyang@ucsb.edu

Daniel Liberzon
Coordinated Science Laboratory
University of Illinois at Urbana-Champaign
Urbana, IL, USA
liberzon@illinois.edu

João P. Hespanha
Center for Control, Dynamical Systems, and Computation
University of California
Santa Barbara, CA, USA
hespanha@ucsb.edu

ABSTRACT

This paper studies topological entropy of switched nonlinear systems. We construct a general upper bound for the topological entropy in terms of an average of the asymptotic suprema of the measures of Jacobian matrices of individual modes, weighted by the corresponding active rates. A general lower bound is also established in terms of an active-rate-weighted average of the asymptotic infima of the traces of these Jacobian matrices. For switched systems with diagonal modes, we establish upper and lower bounds that only depend on the eigenvalues of Jacobian matrices, their relative order among individual modes, and the active rates. For both cases, we also establish upper bounds that are more conservative but require less information on the switching, with their relations illustrated by numerical examples of a switched Lotka–Volterra ecosystem model.

CCS CONCEPTS

• **Computing methodologies** → **Uncertainty quantification**; **Discrete-event simulation**; *Systems theory*; Continuous models.

KEYWORDS

Topological entropy; Switched systems; Nonlinear systems

ACM Reference Format:
Guosong Yang, Daniel Liberzon, and João P. Hespanha. . Topological entropy of switched nonlinear systems. In *HSCC '21: 24th ACM International Conference on Hybrid Systems: Computation and Control, May 19–21, 2021, Nashville, TN, USA*. ACM, New York, NY, USA, 11 pages.

1 INTRODUCTION

In systems theory, topological entropy describes the information generation rate in terms of the number of trajectories distinguishable with a finite precision, or the complexity growth rate of a system acting on a set with finite measure. The latter idea corresponds to Kolmogorov's original definition in [17], and bears a striking resemblance to Shannon's information entropy [27]. Adler et al. first defined topological entropy as an extension of Kolmogorov's metric entropy, quantifying the expansion of a map via the minimal cardinality of a subcover refinement [1]. A different definition in terms of the maximal number of trajectories separable with a finite precision was introduced by Bowen [5] and independently by Dinaburg [12]. An equivalence between these two notions was established in [6]. Most existing results on topological entropy are for time-invariant systems, as time-varying dynamics introduce complexities that require new methods to understand [16, 18]. This work on the topological entropy of switched systems aims at contributing to our understanding of some of these complexities.

Entropy also plays a prominent role in control theory, where information flow occurs between sensors and controllers for generating feedback controls. First, a notion of topological feedback entropy was introduced for discrete-time systems [26], following the construction in [1]. Its definition extended the classical entropy concepts and described the growth rate of control complexity as time evolves. Later, a notion of invariance entropy was proposed for continuous-time systems [9], which is closer in spirit to the trajectory-counting formulation in [5, 12]. An equivalence between these two notions was established in [10]. Results from [9] were extended from set invariance to exponential stabilization in [8].

This paper studies the topological entropy of continuous-time switched nonlinear systems. Switched systems have become a popular topic in recent years (see, e.g., [19, 28] and references therein). In general, a switched system does not inherit the stability properties of individual modes. For example, a switched system with two stable modes may still be unstable [19, p. 19]. However, it is well known that a switched linear system generated by a finite family of pairwise commuting Hurwitz matrices is globally exponentially stable under arbitrary switching (see, e.g., [19, 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 [24, 33]. One of the simplest sufficient conditions for pairwise commuting modes is that the system functions are simultaneously diagonalizable, which motivates us to study the topological entropy of switched systems with diagonal modes in addition to the general case.

Our interest in the topological entropy of switched systems is strongly motivated by its relation to the data-rate requirements in control problems. For linear time-invariant control systems, the minimal data rate for feedback stabilization is given by the sum of the positive real parts of eigenvalues of the system matrix [13] (or, for discrete-time systems, the sum of their logarithms [13, 25, 31]), which equals the topological entropy of these systems in open-loop [5, 9]. Data-rate conditions and entropy notions for nonlinear time-invariant control systems were established in [8, 22, 26]. For switched systems, however, neither the minimal data rate nor the topological entropy are completely understood. Sufficient data rates

HSCC '21, May 19–21, 2021, Nashville, TN, USA

for feedback stabilization of switched linear systems were established in [21, 35]. Similar data-rate conditions were constructed in [29] by extending the estimation entropy from [23]. In [36, 37], formulae and upper and lower bounds for the topological entropy of switched linear systems were constructed in terms of the active rates of individual modes, an approach which is adopted in the current work. Relations between topological entropy and stability properties of switched linear systems were established in [34, 37]. For discrete-time switched linear systems, the topological entropy under worse-case switching sequences was obtained based on the notion of Joint Spectral Radius [3].

The topological entropy of switched nonlinear systems has not been explored so far. The main contribution of this paper is to establish upper and lower bounds for the topological entropy of switched nonlinear systems, which generalize previous results for switched linear systems in [36, 37] and further our understanding of how switching affects topological entropy. In Section 2, after presenting the entropy definition for switched systems, we provide a standard construction of spanning and separated sets using a notion of grid. Moreover, we define switching-related quantities such as the active rates of individual modes, and construct upper and lower bounds for the distance between solutions and a lower bound for the volume of a reachable set. These bounds are essential to the computation of topological entropy and are also of independent interest.

In Section 3, we establish a general upper bound for the topological entropy of switched nonlinear systems, in terms of an average of the measures of Jacobian matrices of individual modes, weighted by their corresponding active rates and maximized over the ω-limit set. A general lower bound is also established in terms of an active-rate-weighted average of the traces of these Jacobian matrices, minimized over the ω-limit set. In Section 4, we consider the case with diagonal modes (i.e., each scalar component of the nonlinear system functions only depends on the corresponding scalar component of the state) and establish improved upper and lower bounds that only depend on the eigenvalues of Jacobian matrices, their relative order among individual modes, and the active rates. In both the general case and the case with diagonal modes, we also establish upper bounds that are more conservative but require less information on the switching, with their relations illustrated by numerical examples motivated by a switched Lotka–Volterra ecosystem model. Section 5 concludes the paper with a brief summary and remarks on future research directions.

Notations: Let $\mathbb{R}_{\geq 0} := [0, \infty)$, $\mathbb{R}_{>0} := (0, \infty)$, and $\mathbb{N} := \{0, 1, \ldots\}$. Denote by I_n the identity matrix in $\mathbb{R}^{n\times n}$; the subscript is omitted when the dimension is implicit. For a complex number $a \in \mathbb{C}$, denote by $\mathrm{Re}(a)$ its real part. For a vector $v \in \mathbb{R}^n$, denote by v_i its i-th scalar component and write $v = (v_1, \ldots, v_n)$. For a matrix $A \in \mathbb{R}^{n\times n}$, denote by $\mathrm{spec}(A)$, $\mathrm{tr}(A)$, and $\det(A)$ its spectrum, trace, and determinant, respectively. For a finite set E, denote by $|E|$ its cardinality. For a set $K \subset \mathbb{R}^n$, denote by $\mathrm{vol}(K)$ and $\mathrm{co}(K)$ its volume (Lebesgue measure) and convex hull, respectively. Denote by $\|v\|_\infty := \max_{1\leq i\leq n} |v_i|$ the ∞-norm of a vector $v \in \mathbb{R}^n$, and by $\|A\|_\infty := \max_{1\leq i\leq n} \sum_{j=1}^n |a_{ij}|$ the induced ∞-norm of a matrix $A = [a_{ij}] \in \mathbb{R}^{n\times n}$. By default, all logarithms are natural logarithms. For a function $f : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^k$, denote by $J_x f(r, v) \in \mathbb{R}^{k\times n}$ the Jacobian matrix of $f(r, x)$ with respect to x at (r, v).

2 PRELIMINARIES

Consider a 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 from a finite *index set* $\mathcal{P}$, and all the functions $f_p : \mathbb{R}^n \to \mathbb{R}^n$ are continuously differentiable. We are interested in the corresponding switched system defined by

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

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

2.1 Entropy definitions

In this subsection, we define the topological entropy of the switched system (2) with a 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 chosen norm on $\mathbb{R}^n$ and 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 K with a center $x \in K$:

$$B_{f_\sigma}(x, \varepsilon, T) := \left\{\bar{x} \in K : \max_{t\in[0,T]} \|\xi_\sigma(t, \bar{x}) - \xi_\sigma(t, x)\| < \varepsilon\right\}. \tag{3}$$

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

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

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 $t \in [0, T]$. Let $S(f_\sigma, \varepsilon, T, K)$ denote the minimal cardinality of a (T, ε)-spanning set, which is nondecreasing in T and nonincreasing in ε. The *topological entropy* of the switched system (2) with initial set K and switching signal σ is defined in terms of the exponential growth rate of $S(f_\sigma, \varepsilon, T, K)$ by

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

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

Remark 2.1. In view of the equivalence of norms on a finite-dimensional vector space, the values of $h(f_\sigma, K)$ are the same for all norms $\|\cdot\|$ on $\mathbb{R}^n$. In particular, the topological entropy is invariant under a change of basis. For convenience and concreteness, we take $\|\cdot\|$ to be the ∞-norm on $\mathbb{R}^n$ or the induced ∞-norm on $\mathbb{R}^{n\times n}$ unless otherwise specified.

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

$$\bar{x} \notin B_{f_\sigma}(x, \varepsilon, T) \tag{6}$$

for all distinct points $x, \bar{x} \in E$, or equivalently, there is a time $t \in [0, T]$ such that $\|\xi_\sigma(t, \bar{x}) - \xi_\sigma(t, x)\| \geq \varepsilon$. Let $N(f_\sigma, \varepsilon, T, K)$ denote the maximal cardinality of a (T, ε)-separated set, which is also nondecreasing in T and nonincreasing in ε. As stated in the following result, the entropy of (2) can be equivalently formulated in terms of the exponential growth rate of $N(f_\sigma, \varepsilon, T, K)$; the proof is along the lines of [15, p. 110] and thus omitted here.

Proposition 2.2. *The topological entropy of the switched system* (2) *satisfies*

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

2.2 Standard spanning and separated sets

Given a time horizon $T \geq 0$ and a radius $\varepsilon > 0$, we provide a standard construction of (T, ε)-spanning and (T, ε)-separated sets by extending the notion of grid in [37]. Given a vector $\theta = (\theta_1, \dots, \theta_n) \in \mathbb{R}^n_{>0}$ which may depend on T and ε, we define the following grid on the initial set K:

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

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

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

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

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

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

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

By 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 [37, Lemma 2]; the proof can be found in Appendix A.

Lemma 2.3. *1. If the vector θ is selected so that $R(x) \subset B_{f_\sigma}(x, \varepsilon, T)$ for all $x \in G(\theta)$, then the grid $G(\theta)$ is (T, ε)-spanning; additionally, if all θ_i are nonincreasing in T, then the topological entropy of* (2) *satisfies*

$$h(f_\sigma, K) \leq \lim_{\varepsilon \searrow 0} \limsup_{T \to \infty} \sum_{i=1}^{n} \frac{\log(1/\theta_i)}{T}. \tag{10}$$

2. If the vector θ is selected so that $B_{f_\sigma}(x, \varepsilon, T) \subset R(x)$ for all $x \in G(\theta)$, then the grid $G(\theta)$ is (T, ε)-separated; additionally, if all θ_i are nonincreasing in T, then the topological entropy of (2) *satisfies*

$$h(f_\sigma, K) \geq \lim_{\varepsilon \searrow 0} \limsup_{T \to \infty} \sum_{i=1}^{n} \frac{\log(1/\theta_i)}{T}. \tag{11}$$

2.3 Active times and active rates

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

$$\tau_p(t) := \int_0^t \mathbb{1}_p(\sigma(s))\, \mathrm{d}s, \qquad p \in \mathcal{P} \tag{12}$$

with the indicator function

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

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

$$\rho_p(t) := \tau_p(t)/t, \qquad p \in \mathcal{P} \tag{13}$$

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

$$\hat{\rho}_p := \limsup_{t \to \infty} \rho_p(t), \qquad p \in \mathcal{P}. \tag{14}$$

Clearly, the active times τ_p are nonnegative and nodecreading, and satisfy $\sum_{p \in \mathcal{P}} \tau_p(t) = t$ for all $t \geq 0$; the active rates ρ_p take values in $[0, 1]$ and satisfy $\sum_{p \in \mathcal{P}} \rho_p(t) = 1$ for all $t \geq 0$. In contrast, due to the limit supremum in (14), it is possible that $\sum_{p \in \mathcal{P}} \hat{\rho}_p > 1$ for the asymptotic active rates $\hat{\rho}_p$, as illustrated in [37, Example 1].

In [37, Lemma 1], it was shown that for a family of scalars $\{c_p \in \mathbb{R} : p \in \mathcal{P}\}$, 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 present a technical lemma that generalizes this result to the case with a family $\{a_p : p \in \mathcal{P}\}$ of integrable functions $a_p : \mathbb{R}_{\geq 0} \to \mathbb{R}$; the proof can be found in Appendix B.

Lemma 2.4. *We have*

$$\begin{aligned} &\limsup_{T \to \infty} \frac{1}{T} \max_{t \in [0,T]} \sum_{p \in \mathcal{P}} \int_0^t a_p(s)\, \mathbb{1}_p(\sigma(s))\, \mathrm{d}s \\ &\quad = \max\left\{ \limsup_{t \to \infty} \sum_{p \in \mathcal{P}} \frac{1}{t} \int_0^t a_p(s)\, \mathbb{1}_p(\sigma(s))\, \mathrm{d}s,\ 0 \right\}. \end{aligned} \tag{15}$$

Moreover, the first term in the maximum on the right-hand side of (15) *satisfies*

$$\begin{aligned} \limsup_{t \to \infty} \sum_{p \in \mathcal{P}} \check{a}_p \rho_p(t) &\leq \limsup_{t \to \infty} \sum_{p \in \mathcal{P}} \frac{1}{t} \int_0^t a_p(s)\, \mathbb{1}_p(\sigma(s))\, \mathrm{d}s \\ &\leq \limsup_{t \to \infty} \sum_{p \in \mathcal{P}} \hat{a}_p \rho_p(t) \end{aligned} \tag{16}$$

with

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

Note that the sum on the left-hand side of (15) is in fact the integral of a_σ over $[0, t]$, as

$$\begin{aligned} \int_0^t a_{\sigma(s)}(s)\, \mathrm{d}s &= \sum_{k=0}^{N_\sigma(t,0)} \int_{t_k}^{t_{k+1}} a_{\sigma(t_k)}(s)\, \mathrm{d}s \\ &= \sum_{p \in \mathcal{P}} \int_0^t a_p(s)\, \mathbb{1}_p(\sigma(s))\, \mathrm{d}s, \end{aligned} \tag{17}$$

where we denote by $0 < t_1 < \cdots < t_{N_\sigma(t,0)} \le t$ the sequence of switching times and let $t_0 := 0$ and $t_{N_\sigma(t,0)+1} := t$.

2.4 Bounds for distance between solutions and volume of reachable set

In preparation for the computation of topological entropy, we construct upper and lower bounds for the distance between 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) at time t from initial set K.

Following [11, p. 30], for an induced norm $\|\cdot\|$ on $\mathbb{R}^{n\times n}$, the *matrix measure* $\mu : \mathbb{R}^{n\times n} \to \mathbb{R}$ is defined by

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

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

$$\mu(A) = \max_{1\le i\le n} \left(a_{ii} + \sum_{j\ne i} |a_{ij}| \right) \tag{18}$$

for a matrix $A = [a_{ij}] \in \mathbb{R}^{n\times n}$. For all induced norms $\|\cdot\|$ on $\mathbb{R}^{n\times n}$, the function μ is convex and satisfies [11, Th. 5, p. 31]

$$-\mu(-A) \le \operatorname{Re}(\lambda) \le \mu(A) \le \|A\| \qquad \forall\, A \in \mathbb{R}^{n\times n}, \forall\, \lambda \in \operatorname{spec}(A). \tag{19}$$

Proposition 2.5. *For all initial states $x, \bar{x} \in K$, the corresponding solutions to the switched system* (2) *satisfy*

$$e^{\underline{\eta}_\sigma(t)}\|\bar{x} - x\| \le \|\xi_\sigma(t,\bar{x}) - \xi_\sigma(t,x)\| \le e^{\overline{\eta}_\sigma(t)}\|\bar{x} - x\| \qquad \forall\, t \ge 0 \tag{20}$$

with

$$\begin{aligned}\underline{\eta}_\sigma(t) &:= \sum_{p\in\mathcal{P}} \int_0^t \left(\min_{v\in\operatorname{co}(\xi_\sigma(s,K))} -\mu(-J_x f_p(v)) \right) \mathbb{1}_p(\sigma(s))\,\mathrm{d}s, \\ \overline{\eta}_\sigma(t) &:= \max_{v\in\operatorname{co}(K)} \sum_{p\in\mathcal{P}} \int_0^t \mu(J_x f_p(\xi_\sigma(s,v)))\,\mathbb{1}_p(\sigma(s))\,\mathrm{d}s.\end{aligned} \tag{21}$$

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

$$\operatorname{vol}(\xi_\sigma(t,K)) \ge e^{\gamma_\sigma(t)} \operatorname{vol}(K) \qquad \forall\, t \ge 0 \tag{22}$$

with

$$\gamma_\sigma(t) := \min_{v\in K} \sum_{p\in\mathcal{P}} \int_0^t \operatorname{tr}(J_x f_p(\xi_\sigma(s,v)))\,\mathbb{1}_p(\sigma(s))\,\mathrm{d}s. \tag{23}$$

Note that $\underline{\eta}_\sigma$, $\overline{\eta}_\sigma$, and γ_σ are in fact constructed in terms of the integrals of the measure (or its minimum) and the trace of the *active* Jacobian matrix over $[0,t]$, rewritten via a similar transformation to (17).

Proof of Proposition 2.5. Consider a linear time-varying (LTV) system

$$\dot{x} = A(t)\,x \tag{24}$$

with a piecewise continuous matrix-valued function $A : \mathbb{R}_{\ge 0} \to \mathbb{R}^{n\times n}$. For all $v \in \mathbb{R}^n$, its solution $\xi(t,v)$ with initial state v and state-transition matrix $\Phi_A(t,0)$ satisfy

$$\begin{aligned} e^{\int_0^t -\mu(-A(s))\,\mathrm{d}s}\|v\| &\le \|\xi(t,v)\| \\ &= \|\Phi_A(t,0)\,v\| \le e^{\int_0^t \mu(A(s))\,\mathrm{d}s}\|v\| \qquad \forall\, t \ge 0 \end{aligned} \tag{25}$$

[11, Th. 27, p. 34] and

$$\det(\Phi_A(t,0)) = e^{\int_0^t \operatorname{tr}(A(s))\,\mathrm{d}s} \qquad \forall\, t \ge 0 \tag{26}$$

(Liouville's formula [7, Prop. 2.18, p. 152]).

First, we write the Jacobian matrix $J_x\xi_\sigma(t,v)$ of a solution $\xi_\sigma(t,v)$ to the switched system (2) as a matrix solution to the LTV system (24) with an appropriate system matrix $A(t)$, based on a common procedure in nonlinear systems analysis (see, e.g., [20, 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 $t \ge 0$ that are not switching times. Hence for each fixed $v \in \mathbb{R}^n$, the Jacobian matrix $J_x\xi_\sigma(t,v)$ is the *principal fundamental matrix solution* [7, Def. 2.12, p. 150] to the LTV system (24) with the system matrix $A(t) = J_x f_{\sigma(t)}(\xi_\sigma(t,v))$ and thus equals the state-transition matrix $\Phi_A(t,0)$. Let $\nu(\rho) := \rho\bar{x} + (1-\rho)\,x$ for $\rho \in [0,1]$. Then $\nu(\rho) \in \operatorname{co}(K)$ for all $\rho \in [0,1]$. Hence the upper bound in (25) with $A(t) = J_x f_{\sigma(t)}(\xi_\sigma(t,v))$ and $\Phi_A(t,0) = J_x\xi_\sigma(t,v)$ implies[1]

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

for all $t \ge 0$, that is, the upper bound in (20) holds.

Second, we write the difference between solutions $\xi_\sigma(t,\bar{x}) - \xi_\sigma(t,x)$ to the switched system (2) as a solution to the LTV system (24) with an appropriate system matrix $A(t)$, based on a similar procedure to the one in the first part. Let $\bar{\nu}(t,\rho) := \rho\xi_\sigma(t,\bar{x}) + (1-\rho)\xi_\sigma(t,x)$ for $\rho \in [0,1]$. Then $\bar{\nu}(t,\rho) \in \operatorname{co}(\xi_\sigma(t,K))$ for all $\rho \in [0,1]$ and $t \ge 0$, and

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

for all $t \ge 0$ that are not switching times. Hence $\xi_\sigma(t,\bar{x}) - \xi_\sigma(t,x)$ is the solution to the LTV system (24) with the system matrix $A(t) = \int_0^1 J_x f_{\sigma(t)}(\bar{\nu}(t,\rho))\,\mathrm{d}\rho$, with initial state $\bar{x} - x$, and thus the lower bound in (25) implies[2]

$$\begin{aligned} \|\xi_\sigma(t,\bar{x}) - \xi_\sigma(t,x)\| &\ge e^{\int_0^t -\mu\left(-\int_0^1 J_x f_{\sigma(s)}(\bar{\nu}(s,\rho))\,\mathrm{d}\rho\right)\mathrm{d}s}\|\bar{x} - x\| \\ &= e^{\sum_{p\in\mathcal{P}} \int_0^t -\mu\left(-\int_0^1 J_x f_p(\bar{\nu}(s,\rho))\,\mathrm{d}\rho\right)\mathbb{1}_p(\sigma(s))\,\mathrm{d}s}\|\bar{x} - x\| \end{aligned}$$

[1]In the first part of the proof, the arguments based on integrating $J_x\xi_\sigma(t,v)$ along the line segment connecting x and $\bar{x}$ are inspired by similar ones in the proofs of [4, Th. 4.2] for time-invariant systems with a compact state space and [30, Th. 1] for contractive systems.

[2]In the second part of the proof , the arguments based on integrating $J_x f_\sigma(v)$ along the line segment connecting $\xi_\sigma(t,x)$ and $\xi_\sigma(t,\bar{x})$ are inspired by similar ones in [32, Sec. 2.5] for time-varying systems.

for all $t \geq 0$. Moreover, as the function μ is convex, for all $p \in \mathcal{P}$, we have

$$-\mu\left(-\int_0^1 J_x f_p(\bar{v}(t,\rho))\,\mathrm{d}\rho\right) \geq \int_0^1 -\mu(-J_x f_p(\bar{v}(t,\rho))\,\mathrm{d}\rho \geq \min_{v\in\mathrm{co}(\xi_\sigma(t,K))} -\mu(-J_x f_p(v)) \qquad \forall\, t \geq 0.$$

Hence the lower bound in (20) holds.

Finally, Liouville's formula (26) with $A(t) = J_x f_{\sigma(t)}(\xi_\sigma(t,v))$ and $\Phi_A(t,0) = J_x \xi_\sigma(t,v)$ from the first part implies

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

for all $t \geq 0$, that is, (22) holds. □

Additionally, we provide the following upper and lower bounds for the distance between solutions in terms of the active times τ_p, which are more conservative than those in (20) but illustrate the effect of switching and will be useful in establishing bounds for the entropy of switched systems with diagonal modes.

Corollary 2.6. *For all initial states $x, \bar{x} \in K$, the corresponding solutions to the switched system* (2) *satisfy*

$$e^{\sum_{p\in\mathcal{P}} \underline{\mu}_p(t)\tau_p(t)}\|\bar{x}-x\| \leq \|\xi_\sigma(t,\bar{x}) - \xi_\sigma(t,x)\| \leq e^{\sum_{p\in\mathcal{P}} \overline{\mu}_p(t)\tau_p(t)}\|\bar{x}-x\| \qquad \forall\, t \geq 0 \tag{27}$$

with

$$\begin{aligned}\underline{\mu}_p(t) &:= \min_{s\in[0,t],\,\sigma(s)=p,\,v\in\mathrm{co}(\xi_\sigma(s,K))} -\mu(-J_x f_p(v)),\\ \overline{\mu}_p(t) &:= \max_{s\in[0,t],\,\sigma(s)=p,\,v\in\mathrm{co}(K)} \mu(J_x f_p(\xi_\sigma(s,v))),\end{aligned} \qquad p \in \mathcal{P}, \tag{28}$$

where the active times τ_p are defined by (12).

Proof. The upper bound in (27) follows from the upper bound in (20) and the property that

$$\begin{aligned}\overline{\eta}_\sigma(t) &\leq \sum_{p\in\mathcal{P}} \int_0^t \left(\max_{v\in\mathrm{co}(K)} \mu(J_x f_p(\xi_\sigma(s,v)))\right)\mathbb{1}_p(\sigma(s))\,\mathrm{d}s \\ &\leq \sum_{p\in\mathcal{P}} \overline{\mu}_p(t)\int_0^t \mathbb{1}_p(\sigma(s))\,\mathrm{d}s = \sum_{p\in\mathcal{P}} \overline{\mu}_p(t)\tau_p(t) \qquad \forall\, t \geq 0.\end{aligned}$$

The lower bound in (27) follows from the lower bound in (20) and the property that

$$\underline{\eta}_\sigma(t) \geq \sum_{p\in\mathcal{P}} \underline{\mu}_p(t)\int_0^t \mathbb{1}_p(\sigma(s))\,\mathrm{d}s = \sum_{p\in\mathcal{P}} \underline{\mu}_p(t)\tau_p(t) \qquad \forall\, t \geq 0.$$

□

Remark 2.7. Suppose the switched system (2) satisfies that

1. the measure of each Jacobian matrix $\mu(J_x f_p(v))$ has a global upper bound $\overline{\mu}_p^*$, or
2. the convex hull of initial set $\mathrm{co}(K)$ is a subset of a compact positively invariant set S, and let $\overline{\mu}_p^* := \max_{v\in S} \mu(J_x f_p(v))$.

Then for all initial states $x, \bar{x} \in K$, we have

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

which is more conservative but simpler than the upper bounds in (20) and (27). Similarly, a more conservative but simpler lower bound than the ones in (20) and (27) can be constructed for the cases with globally lower-bounded measures $\mu(J_x f_p(v))$, or with a compact positively invariant set containing K. Similar results hold for the lower bound (22). On the other hand, without a global bound for each $\mu(J_x f_p(v))$ and without a compact positively invariant set, $\underline{\eta}_\sigma$, $\overline{\eta}_\sigma$, $\underline{\mu}_p$, and $\overline{\mu}_p$ defined in (21) and (28) may be unbounded.

3 ENTROPY OF GENERAL SWITCHED SYSTEMS

In this section, we establish upper and lower bounds for the entropy of the switching nonlinear system (2).

Theorem 3.1. *The topological entropy of the switched system* (2) *is upper-bounded by*

$$h(f_\sigma, K) \leq \max\left\{\limsup_{t\to\infty} \sum_{p\in\mathcal{P}} n\hat{\mu}_p \rho_p(t),\ 0\right\} \tag{29}$$

with

$$\hat{\mu}_p := \limsup_{s\to\infty,\,\sigma(s)=p} \max_{v\in\mathrm{co}(K)} \mu(J_x f_p(\xi_\sigma(s,v))), \qquad p \in \mathcal{P}, \tag{30}$$

and lower-bounded by

$$h(f_\sigma, K) \geq \max\left\{\limsup_{t\to\infty} \sum_{p\in\mathcal{P}} \check{\chi}_p \rho_p(t),\ 0\right\} \tag{31}$$

with

$$\check{\chi}_p := \liminf_{s\to\infty,\,\sigma(s)=p} \min_{v\in K} \mathrm{tr}(J_x f_p(\xi_\sigma(s,v))), \qquad p \in \mathcal{P}, \tag{32}$$

where the active rates ρ_p are defined by (13).

Proof. First, we establish the upper bound (29). Fix a time horizon $T \geq 0$ and a radius $\varepsilon > 0$. The upper bound in (20) implies that for all initial states $x, \bar{x} \in K$, the corresponding solutions to (2) satisfy

$$\max_{t\in[0,T]} \|\xi_\sigma(t,\bar{x}) - \xi_\sigma(t,x)\| \leq e^{\max_{t\in[0,T]} \overline{\eta}_\sigma(t)}\|\bar{x}-x\|. \tag{33}$$

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

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

and the corresponding hypercubes $R(x)$ defined by (9). Comparing (9) and (33) to (3), we see that $R(x) \subset B_{f_\sigma}(x,\varepsilon,T)$ for all $x \in G(\theta)$. Then Lemma 2.3 implies that $G(\theta)$ is (T,ε)-spanning and, as all θ_i are nonincreasing in T, the upper bound (10) yields

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

with

$$a_{1p}(s) := \max_{v \in \operatorname{co}(K)} n\mu(J_x f_p(\xi_\sigma(s, v))), \qquad p \in \mathcal{P}.$$

Then the upper bound (29) follows from (15) and the upper bound in (16) with the functions $a_p(t) := a_{1p}(t)$ for $p \in \mathcal{P}$.

Second, we establish the lower bound (31) using volume-based arguments. Fix a time horizon $T \geq 0$ and a radius $\varepsilon > 0$. The lower bound (22) implies that the reachable set $\xi_\sigma(T, K)$ of (2) satisfies

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

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

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

and thus the corresponding volumes satisfy (recall that we take $\|\cdot\|$ to be the ∞-norm; see Remark 2.1)

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

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

$$S(f_\sigma, \varepsilon, T, K) = |E| \geq \operatorname{vol}(\xi_\sigma(T, K))/(2\varepsilon)^n \geq e^{\gamma_\sigma(T)} \operatorname{vol}(K)/(2\varepsilon)^n,$$

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

$$\begin{aligned} h(f_\sigma, K) &\geq \lim_{\varepsilon \searrow 0} \limsup_{T \to \infty} \frac{1}{T} \log\bigl(e^{\gamma_\sigma(T)} \operatorname{vol}(K)/(2\varepsilon)^n\bigr) \\ &= \limsup_{T \to \infty} \frac{\gamma_\sigma(T)}{T} + \lim_{\varepsilon \searrow 0} \limsup_{T \to \infty} \frac{\log(\operatorname{vol}(K)/(2\varepsilon)^n)}{T} \\ &\geq \limsup_{T \to \infty} \frac{1}{T} \sum_{p \in \mathcal{P}} \int_0^T a_{2p}(s)\, \mathbb{1}_p(\sigma(s))\, \mathrm{d}s \end{aligned}$$

with

$$a_{2p}(s) := \min_{v \in K} \operatorname{tr}(J_x f_p(\xi_\sigma(s, v))), \qquad p \in \mathcal{P}.$$

Then the lower bound (31) follows from the lower bound in (16) with the functions $a_p(t) := a_{2p}(t)$ for $p \in \mathcal{P}$ and the property that $h(f_\sigma, K) \geq 0$. □

Thinking of the non-switched case as a switched system with a constant switching signal, Theorem 3.1 implies the following result for the entropy of nonlinear time-invariant systems.

Corollary 3.2. *The topological entropy of the p-th nonlinear time-invariant system in* (1) *is upper-bounded by*

$$h(f_p, K) \leq \max\{n\hat{\mu}_p, 0\} \tag{34}$$

with the constant $\hat{\mu}_p$ defined by (30)*, and lower-bounded by*

$$h(f_p, K) \geq \max\{\check{\chi}_p, 0\} \tag{35}$$

with the constant $\check{\chi}_p$ defined by (32).

Based on the upper bound (29), we construct the following upper bounds for the entropy of (2) that require less information on the switching signal.

Corollary 3.3. *The topological entropy of the switched system* (2) *is upper-bounded by*

$$h(f_\sigma, K) \leq \sum_{p \in \mathcal{P}} \max\{n\hat{\mu}_p, 0\}\hat{\rho}_p \tag{36}$$

with the asymptotic active rates $\hat{\rho}_p$ defined by (14)*, and also by*

$$h(f_\sigma, K) \leq \max_{p \in \mathcal{P}} \max\{n\hat{\mu}_p, 0\}, \tag{37}$$

where the constants $\hat{\mu}_p$ are defined by (30).

Proof. First, the upper bound (29) and the subadditivity of limit suprema imply

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

Second, the upper bound (29) also implies

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

Remark 3.4. Consider the case where all the functions f_p in (2) are linear, that is, there is a family of matrices $\{A_p \in \mathbb{R}^{n \times n} : p \in \mathcal{P}\}$ such that for all $p \in \mathcal{P}$, we have

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

Then the constants $\hat{\mu}_p$ and $\check{\chi}_p$ defined by (30) and (32) satisfy

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

Hence Theorem 3.1 generalizes [37, Th. 1], and Corollary 3.3 generalizes [37, Remark 5].

Remark 3.5. 1. The constants $\hat{\mu}_p$ and $\check{\chi}_p$ defined by (30) and (32) only depend on the measure and the trace of each Jacobian matrix $J_x f_p(v)$ over the ω-limit set from the convex hull of initial set $\operatorname{co}(K)$ and that from K, respectively, instead of over all reachable points from $\operatorname{co}(K)$ and K, respectively. In particular, (29), (36), and (37) will yield finite values for the case with unbounded Jacobian matrices but a compact global attractor.

2. In view of Remark 2.7, if each $\mu(J_x f_p(v))$ has a global upper bound $\hat{\mu}_p^*$, or an upper bound $\hat{\mu}_p^*$ over a compact positively invariant set containing $\operatorname{co}(K)$, then the upper bounds (29), (36), and (37) hold with $\hat{\mu}_p^*$ in place of $\hat{\mu}_p$. Similarly, a more conservative but simpler lower bound than (31) can be constructed for the case with globally lower-bounded traces $\operatorname{tr}(J_x f_p(v))$, or with a compact positively invariant set containing K.

3. For a fixed family of functions $\{f_p : p \in \mathcal{P}\}$, compared with the upper bound (29), the upper bound (36) depends on the asymptotic active rates $\hat{\rho}_p$ instead of the active rates ρ_p, and the upper bound (37) does not involve active rates at all. If a global upper bound $\hat{\mu}_p^*$ is used in place of $\hat{\mu}_p$ for each p, then (37) is independent of switching.

4. The upper bound (29) is tighter than the upper bounds (36) and (37). The upper bounds (36) and (37) are both useful in the sense that neither is more conservative than the other, as it is possible that $\sum_{p \in \mathcal{P}} \hat{\rho}_p > 1$. The relations between the upper bounds (29), (36), and (37) are illustrated numerically in Example 3.6 below.

Example 3.6. Consider the following switched nonlinear system on the nonnegative orthant $\mathbb{R}^n_{\geq 0}$ from [2]:

$$\dot{x}_i = f^i_\sigma(x) := \left(r^i_\sigma + \sum_{j=1}^{n} a^{ij}_\sigma x_j\right) x_i, \qquad i \in \{1, \dots, n\} \tag{38}$$

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 $\mathcal{P}$. Each individual mode p of (38) is a Lotka–Volterra ecosystem model that describes the population dynamics of n species in a biological community [14, Ch. 5], where x_i denotes the population density of the i-th species, the coefficient $r^i_p \in \mathbb{R}$ quantifies the intrinsic growth rate of the i-th population, the self-interaction term $a^{ii}_p < 0$ is justified by the limitation of resources in the environment, and the interaction term $a^{ij}_p \in \mathbb{R}$ for $j \neq i$ quantifies the influence of the j-th population on the i-th one. Switching in (38) may be due to seasonal changes or other environmental factors. Clearly, $\mathbb{R}^n_{\geq 0}$ is a positively invariant set for (38).

Consider the switched system (38) in $\mathbb{R}^2_{\geq 0}$ with the index set $\mathcal{P} = \{1, 2\}$, the coefficients $(r^1_1, r^2_1) = (-1, 2)$ and $(r^1_2, r^2_2) = (3, -1)$, the self-interaction terms $a^{11}_1 = a^{22}_1 = a^{11}_2 = a^{22}_2 = -1$, and the interaction terms $a^{12}_1 = a^{21}_1 = a^{12}_2 = a^{21}_2 = 0.1$. Clearly, in $\mathbb{R}^2_{\geq 0}$, 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 $\{0\} \times \mathbb{R}_{\geq 0}$. Based on [2], (38) is *uniformly ultimately bounded (UUB)* in $\mathbb{R}^n_{\geq 0}$ and its ω-limit set is a subset of the set $\Omega := [0, 3.04] \times [0, 2.03]$. Note that Ω is not a positively invariant set for (38).

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

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

Typical trajectories of the individual modes 1 and 2 and the switched system (38) with switching signals σ_1 and σ_2 are plotted in Fig. 1.

The Jacobian matrices of individual modes of (38) are given by

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

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

As (38) is UUB and its ω-limit set is a subset of Ω, for all initial sets $K \subset \mathbb{R}^n_{\geq 0}$, the constants $\hat{\mu}_p$ defined by (30) satisfy

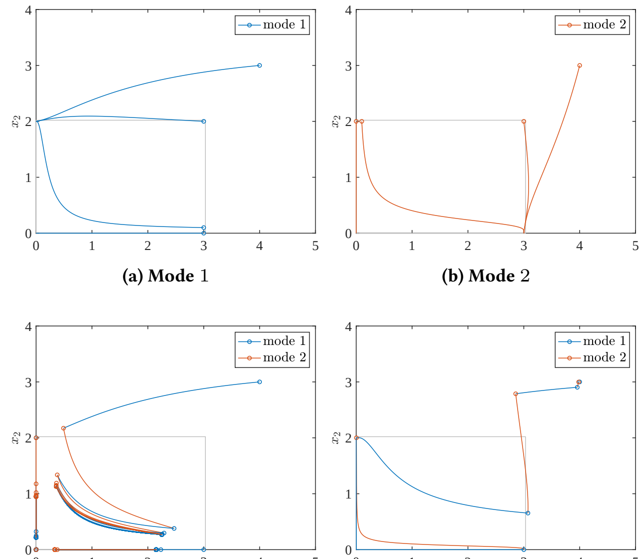
$$\hat{\mu}_1 \leq \max_{v \in \Omega} \max\{-1 - 1.9v_1 + 0.1v_2,\ 2 + 0.1v_1 - 1.9v_2\} \leq 2.31,$$

$$\hat{\mu}_2 \leq \max_{v \in \Omega} \max\{3 - 1.9v_1 + 0.1v_2,\ -1 + 0.1v_1 - 1.9v_2\} \leq 3.21.$$

The upper bounds for $h(f_{\sigma_1}, K)$ and $h(f_{\sigma_2}, K)$ computed using (29), (36), and (37) for all $K \subset \mathbb{R}^n_{\geq 0}$ are summarized in Table 1. In particular, the upper bound (29) for $h(f_{\sigma_2}, K)$ can be computed as follows:

$$\begin{aligned} h(f_{\sigma_2}, K) &\leq \limsup_{t \to \infty} \left(2\hat{\mu}_1(1 - \rho_2(t)) + 2\hat{\mu}_2\rho_2(t)\right) \\ &= 2\hat{\mu}_1 + 2(\hat{\mu}_2 - \hat{\mu}_1)\hat{\rho}_2 \leq 6.23. \end{aligned}$$

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



(a) Mode 1 (b) Mode 2

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

Figure 1: Typical trajectories of the switched system (38) for (a) mode 1 with initial states $(4, 3)$, $(3, 2)$, $(3, 0.1)$, and $(3, 0)$, (b) mode 2 with initial states $(4, 3)$, $(3, 2)$, $(0.1, 2)$, and $(0, 2)$, (c) switching signal σ_1 with initial states $(4, 3)$, $(3, 0)$, $(0, 2)$, and $(0, 0)$, and (d) switching signal σ_2 with initial state $(4, 3)$. The circles mark the beginning of segments after switching. The gray rectangle represents the set $\Omega = [0, 3.04] \times [0, 2.03]$.

Table 1: Upper bounds for the entropy of the switched system (38).

	$(\hat{\rho}_1, \hat{\rho}_2)$	(29)	(36)	(37)
σ_1	$(0.5, 0.5)$	5.51	5.51	6.41
σ_2	$(0.9, 0.9)$	6.23	9.91	6.41

4 ENTROPY OF SWITCHED DIAGONAL SYSTEMS

Consider the case where for each $p \in \mathcal{P}$ and $i \in \{1, \dots, n\}$, the i-th scalar component f^i_p of the function f_p only depends on the corresponding scalar component x_i of the state x. For brevity, we consider f^i_p as a function on $\mathbb{R}$ and denote by $f^i_p(x_i)$ the i-th scalar component of $f_p(x)$. Then (2) becomes the *switched diagonal system* defined by

$$\dot{x}_i = f^i_\sigma(x_i), \qquad i \in \{1, \dots, n\}. \tag{39}$$

Clearly, the i-th scalar component of the solution $\xi_\sigma(t, x)$ also only depends on the corresponding scalar component x_i of the initial state x. For brevity, we denote by $\xi^i_\sigma(t, x_i)$ the i-th scalar component of the solution $\xi_\sigma(t, x)$, and by $\xi^i_\sigma(t, K) := \{\xi^i_\sigma(t, x_i) : x \in K\}$ the projection of the reachable set $\xi_\sigma(t, K)$ onto the i-th dimension. In this section, we establish upper and lower bounds for the entropy of the switched diagonal system (39) that are generally tighter than the results of simply applying the bounds from Section 3 to (39).

Theorem 4.1. *The topological entropy of the switched diagonal system* (39) *satisfies*

$$\limsup_{T\to\infty}\sum_{i=1}^{n}\frac{1}{T}\max_{t\in[0,T]}\sum_{p\in\mathcal{P}}\underline{a}_p^i\tau_p(t)\le h(f_\sigma,K)$$
$$\le\limsup_{T\to\infty}\sum_{i=1}^{n}\frac{1}{T}\max_{t\in[0,T]}\sum_{p\in\mathcal{P}}\overline{a}_p^i\tau_p(t) \tag{40}$$

with

$$\begin{aligned}\underline{a}_p^i &:= \inf_{s\ge0,\,\sigma(s)=p}\ \min_{v_i\in\mathrm{co}(\xi_\sigma^i(s,K))}(f_p^i)'(v_i),\\ \overline{a}_p^i &:= \sup_{s\ge0,\,\sigma(s)=p}\ \max_{v\in\mathrm{co}(K)}(f_p^i)'(\xi_\sigma^i(s,v_i))\end{aligned} \tag{41}$$

for $i\in\{1,\dots,n\}$ and $p\in\mathcal{P}$, where the active times τ_p are defined by (12)*; it is also upper-bounded by*

$$h(f_\sigma,K)\le\sum_{i=1}^{n}\max\left\{\limsup_{t\to\infty}\sum_{p\in\mathcal{P}}\hat{a}_p^i\rho_p(t),\ 0\right\} \tag{42}$$

with

$$\hat{a}_p^i := \limsup_{s\to\infty,\,\sigma(s)=p}\ \max_{v\in\mathrm{co}(K)}(f_p^i)'(\xi_\sigma^i(s,v_i)) \tag{43}$$

for $i\in\{1,\dots,n\}$ and $p\in\mathcal{P}$, where the active rates ρ_p are defined by (13).

Proof. Fix a time horizon $T\ge0$ and a radius $\varepsilon>0$. Applying the upper and lower bounds (27) and the upper bound in (20) for the distance between solutions to each scalar component of (39), we obtain that for all initial states $x,\bar{x}\in K$, the corresponding solutions satisfy (recall that we take $\|\cdot\|$ to be the ∞-norm; see Remark 2.1)

$$\max_{1\le i\le n} e^{\max_{t\in[0,T]}\sum_{p\in\mathcal{P}}\underline{\mu}_p^i(t)\tau_p(t)}|\bar{x}_i-x_i|\le\max_{t\in[0,T]}\|\xi_\sigma(t,\bar{x})-\xi_\sigma(t,x)\|$$
$$\le\max_{1\le i\le n} e^{\max_{t\in[0,T]}\sum_{p\in\mathcal{P}}\overline{\mu}_p^i(t)\tau_p(t)}|\bar{x}_i-x_i| \tag{44}$$

with

$$\begin{aligned}\underline{\mu}_p^i(t) &:= \min_{s\in[0,t],\,\sigma(s)=p,\,v_i\in\mathrm{co}(\xi_\sigma^i(s,K))}(f_p^i)'(v_i)\ge\underline{a}_p^i,\\ \overline{\mu}_p^i(t) &:= \max_{s\in[0,t],\,\sigma(s)=p,\,v\in\mathrm{co}(K)}(f_p^i)'(\xi_\sigma^i(s,v_i))\le\overline{a}_p^i\end{aligned}$$

for $i\in\{1,\dots,n\}$ and $p\in\mathcal{P}$, and also

$$\max_{t\in[0,T]}\|\xi_\sigma(t,\bar{x})-\xi_\sigma(t,x)\|\le\max_{1\le i\le n}e^{\max_{t\in[0,T]}\overline{\eta}_\sigma^i(t)}|\bar{x}_i-x_i|, \tag{45}$$

with

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

for $i\in\{1,\dots,n\}$.

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

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

and the corresponding hyperrectangles $R(x)$ defined by (9). Comparing (9) and the upper bound in (44) to (3), we see that $R(x)\subset B_{f_\sigma}(x,\varepsilon,T)$ for all $x\in G(\theta)$. Then Lemma 2.3 implies that $G(\theta)$ is (T,ε)-spanning and, as all θ_i are nonincreasing in T, the upper bound (10) yields

$$\begin{aligned}h(f_\sigma,K)&\le\lim_{\varepsilon\searrow0}\limsup_{T\to\infty}\sum_{i=1}^{n}\frac{\log(1/\theta_i)}{T}\\ &=\limsup_{T\to\infty}\sum_{i=1}^{n}\frac{1}{T}\max_{t\in[0,T]}\sum_{p\in\mathcal{P}}\overline{\mu}_p^i(t)\tau_p(t)+\lim_{\varepsilon\searrow0}\limsup_{T\to\infty}\frac{n\log(1/\varepsilon)}{T}\\ &\le\limsup_{T\to\infty}\sum_{i=1}^{n}\frac{1}{T}\max_{t\in[0,T]}\sum_{p\in\mathcal{P}}\overline{a}_p^i\tau_p(t),\end{aligned}$$

that is, the upper bound in (40) holds.

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

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

we can show that $G(\theta)$ is (T,ε)-separated and, as all θ_i are nonincreasing in T, the lower bound (11) yields the one in (40).

Finally, following similar arguments to those in the first part of the proof of Theorem 3.1 while considering

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

we can show that $G(\theta)$ is (T,ε)-spanning and, as all θ_i are nonincreasing in T, the upper bound (10) yields

$$\begin{aligned}h(f_\sigma,K)&\le\limsup_{T\to\infty}\sum_{i=1}^{n}\frac{1}{T}\max_{t\in[0,T]}\sum_{p\in\mathcal{P}}\int_0^t a_p^i(s)\,\mathbb{1}_p(\sigma(s))\,\mathrm{d}s\\ &\le\sum_{i=1}^{n}\limsup_{T\to\infty}\frac{1}{T}\max_{t\in[0,T]}\sum_{p\in\mathcal{P}}\int_0^t a_p^i(s)\,\mathbb{1}_p(\sigma(s))\,\mathrm{d}s\end{aligned}$$

with

$$a_p^i(s) := \max_{v\in\mathrm{co}(K)}(f_p^i)'(\xi_\sigma^i(s,v_i)),\qquad p\in\mathcal{P},$$

where the last inequality is due to the subadditivity of limit suprema. Then we obtain the upper bound (42) by applying (15) and the upper bound in (16) in each scalar component with the functions $a_p(t) := a_p^i(t)$ for $p\in\mathcal{P}$. □

Based on the upper bound (42), we construct the following upper bounds for the entropy of (39) that require less information on the switching signal; the proof is along the lines of that of Corollary 3.3 and thus omitted here.

Corollary 4.2. *The topological entropy of the switched diagonal system* (39) *is upper-bounded by*

$$h(f_\sigma,K)\le\sum_{p\in\mathcal{P}}\left(\sum_{i=1}^{n}\max\{\hat{a}_p^i,\ 0\}\right)\hat{\rho}_p \tag{46}$$

with the asymptotic active rates $\hat{\rho}_p$ defined by (14)*, and also by*

$$h(f_\sigma,K)\le\max_{p\in\mathcal{P}}\left(\sum_{i=1}^{n}\max\{\hat{a}_p^i,\ 0\}\right), \tag{47}$$

where the constants $\hat{a}_p^i$ are defined by (43).

Remark 4.3. Consider the case where all the functions f_p^i in (39) are linear, that is, there is a family of diagonal matrices $\{D_p = \mathrm{diag}(a_p^1, \ldots, a_p^n) \in \mathbb{R}^{n\times n} : p \in \mathcal{P}\}$ such that

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

Then the constants $\overline{a}_p^i$, $\underline{a}_p^i$, and $\hat{a}_p^i$ defined by (41) and (43) satisfy

$$\overline{a}_p^i = \underline{a}_p^i = \hat{a}_p^i = a_p^i \qquad \forall\, i \in \{1, \ldots, n\}, \forall\, p \in \mathcal{P}.$$

Hence Theorem 4.1 generalizes [36, Th. 7 and Prop. 8], and Corollary 4.2 generalizes [36, Cor. 10].

Remark 4.4. 1. The constants $\underline{a}_p^i$ and $\overline{a}_p^i$ defined by (41) depend on the partial derivatives $(f_p^i)'(v_i)$ over the convex hull of all reachable points from the initial set K and over all reachable points from the convex hull $\mathrm{co}(K)$, respectively, whereas $\hat{a}_p^i$ defined by (43) only depends on $(f_p^i)'(v_i)$ over the ω-limit set from $\mathrm{co}(K)$. Their difference is due to the different constructions of the upper and lower bounds in (20) and the upper bound in (27). In particular, (42), (46), and (47) will yield finite values for the case with unbounded partial derivatives but a compact global attractor.

2. In view of Remark 2.7, if each $(f_p^i)'(v_i)$ has a global upper bound $\hat{a}_p^{i*}$, or an upper bound $\hat{a}_p^{i*}$ over a compact positively invariant set containing $\mathrm{co}(K)$, then the upper bound in (40) and the upper bounds (42), (46), and (47) hold with $\hat{a}_p^{i*}$ in place of $\overline{a}_p^i$ and $\hat{a}_p^i$. Similarly, a more conservative but simpler lower bound than the one in (40) can be constructed for the case with globally lower-bound $(f_p^i)'(v_i)$, or with a compact positively invariant set containing K.

3. For a fixed family of functions $\{f_p : p \in \mathcal{P}\}$, compared with the upper bound in (40) and the upper bound (42), the upper bound (46) depends on the asymptotic active rates $\hat{\rho}_p$ instead of the active rates ρ_p, and the upper bound (47) does not involve active rates at all. If a global upper bound $\hat{a}_p^{i*}$ is used in place of $\hat{a}_p^i$ for each p and i, then (47) is independent of switching.

4. The upper bound (42) is tighter than the upper bounds (29), (46), and (47), while (46) and (47) are tighter than the upper bounds (36) and (37), respectively. The upper bound in (40) and the upper bound (42) are both useful in the sense that neither is more conservative than the other, due to their difference explained in item 1; however, if the ω-limit set from $\mathrm{co}(K)$ contains all reachable points from $\mathrm{co}(K)$, then the former is tighter than the latter. The same conclusion holds between the upper bound in (40) and the upper bounds (36) and (37). The upper bounds (46) and (47) are both useful in the same sense, as it is possible that $\sum_{p\in\mathcal{P}} \hat{\rho}_p > 1$. The relations between the upper bounds (29), (42), (46), and (47), and the one in (40) are illustrated numerically in Example 4.5 below.

Example 4.5. Consider the switched nonlinear system on the nonnegative orthant (38) in Example 3.6. In this example, we consider the case where $a_p^{ij} = 0$ for all $j \neq i$ and $p \in \mathcal{P}$. Then (38) becomes the switched diagonal system

$$\dot{x}_i = f_\sigma^i(x_i) := (r_\sigma^i + a_\sigma^{ii} x_i)\, x_i, \qquad i \in \{1, \ldots, n\}. \tag{48}$$

Note that for each $p \in \mathcal{P}$ and $i \in \{1, \ldots, n\}$, we have $f_p^i(x_i) < 0$ if $x_i > \max\{-r_p^i/a_p^{ii}, 0\}$. Thus (48) is UUB in $\mathbb{R}^n_{\geq 0}$ and its ω-limit set is a subset of the positively invariant set [2]

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

Consider the switched diagonal system (48) in $\mathbb{R}^2_{\geq 0}$ with the same index set and parameters as those in Example 3.6 except no interaction terms. Clearly, the individual modes have the same attractors and saddle points as those in Example 3.6, and the positively invariant set $\Omega = [0,3] \times [0,2]$. Typical trajectories of the individual modes 1 and 2 and the switched diagonal system (48) with the switching signals σ_1 and σ_2 defined in Example 3.6 are plotted in Fig. 2.

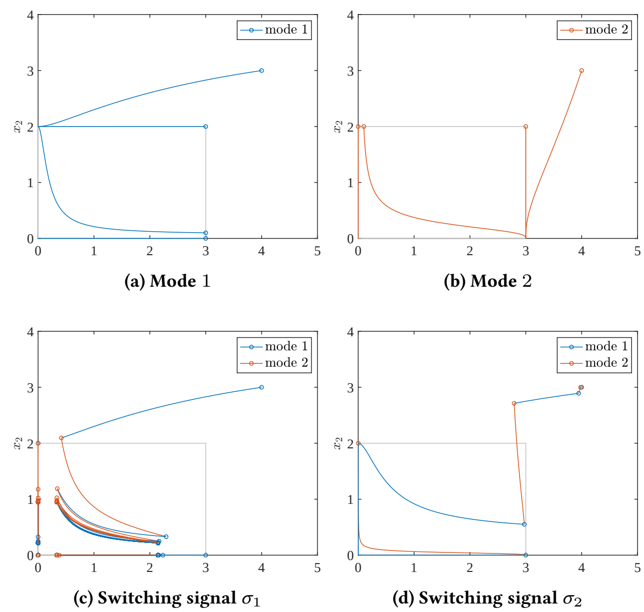


Figure 2: Typical trajectories of the switched diagonal system (48) for (a) mode 1 with initial states $(4,3)$, $(3,2)$, $(3,0.1)$, and $(3,0)$, (b) mode 2 with initial states $(4,3)$, $(3,2)$, $(0.1,2)$, and $(0,2)$, (c) switching signal σ_1 with initial states $(4,3)$, $(3,0)$, $(0,2)$, and $(0,0)$, and (d) switching signal σ_2 with initial state $(4,3)$. The circles mark the beginning of segments after switching. The gray rectangle represents the positively invariant set $\Omega = [0,3] \times [0,2]$.

The Jacobian matrices of individual modes of (48) are given by

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

As (48) is UUB and its ω-limit set is a subset of Ω, for all initial sets $K \subset \mathbb{R}^n_{\geq 0}$, the constants $\hat{\mu}_p$ and $\hat{a}_p^i$ defined by (30) and (43) satisfy

$$\hat{\mu}_1 \leq \max_{v\in\Omega} \max\{-1-2v_1,\ 2-2v_2\} = 2,$$
$$\hat{\mu}_2 \leq \max_{v\in\Omega} \max\{3-2v_1,\ -1-2v_2\} = 3,$$

and $\hat{a}_1^1 \leq -1$, $\hat{a}_1^2 \leq 2$, $\hat{a}_2^1 \leq 3$, and $\hat{a}_2^2 \leq -1$. Moreover, as Ω is a positively invariant set for (48), if $K \subset \Omega$ then the constants $\overline{a}_p^i$ defined by (41) satisfy $\overline{a}_1^1 \leq -1$, $\overline{a}_1^2 \leq 2$, $\overline{a}_2^1 \leq 3$, and $\overline{a}_2^2 \leq -1$. The upper bounds for $h(f_{\sigma_1}, K)$ and $h(f_{\sigma_2}, K)$ computed using (29), (42), (46), and (47) for all $K \subset \mathbb{R}^n_{\geq 0}$, as well as (40) for all $K \subset \Omega$, are summarized in Table 2. For the case with σ_2, the upper bounds (29) and (42) are computed along the lines of computing (29) in Example 3.6; the upper bound in (40) is computed along the lines of computing $h(D_{\sigma_2})$ in [37, Example 3 and Appendix E].

Table 2: Upper bounds for the entropy of the switched diagonal system (48).

	$(\hat\rho_1, \hat\rho_2)$	$K \subset \mathbb{R}^n_{\geq 0}$ (29)	(42)	(46)	(47)	$K \subset \Omega$ (40)
σ_1	$(0.5, 0.5)$	5	1.5	2.5	3	1.5
σ_2	$(0.9, 0.9)$	5.8	4.3	4.5	3	2.79

5 CONCLUSION

We established upper and lower bounds for the topological entropy of switched nonlinear systems, which generalized previous results for switched linear systems in [36, 37] and furthered 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 a finite value for the case with unbounded Jacobian matrices but a compact global attractor.

Future research directions include analyzing the complexity of computing the upper bounds for topological entropy in this paper, studying the relation between these upper bounds and existing stability conditions for switched nonlinear systems, and establishing bounds for the topological entropy of switched nonlinear systems with more general commutation relations than diagonal modes.

ACKNOWLEDGMENTS

G. Yang and J. P. Hespanha's work was supported by the Office of Naval Research MURI grant N00014-16-1-2710, and by the National Science Foundation grants CNS-1329650 and EPCN-1608880. D. Liberzon's work was supported by the National Science Foundation grant CMMI-1662708, and by the Air Force Office of Scientific Research grant FA9550-17-1-0236.

A PROOF OF LEMMA 2.3

First, as $R(x) \subset B_{f_\sigma}(x, \varepsilon, T)$ for all $x \in G(\theta)$, we have

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

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

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

Consequently, the definition of entropy (5) implies

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

where the last term equals zero due to the assumption that all $\theta_i > 0$ are nonincreasing in T and the fact that r_2 is independent of T.

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

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

Consequently, (7) implies

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

where the last term equals zero due to the assumption that all θ_i are nonincreasing in T and the fact that r_1 is independent of T.

B PROOF OF LEMMA 2.4

For brevity, we define the following functions $\bar{a}, a : \mathbb{R}_{\geq 0} \to \mathbb{R}$ and constant $\hat{a} \in \mathbb{R} \cup \{\infty\}$:

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

$$a(t) := \sum_{p \in \mathcal{P}} \frac{1}{t} \int_0^t a_p(s)\, \mathbb{1}_p(\sigma(s))\, \mathrm{d}s$$

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

$$\hat{a} := \limsup_{t \to \infty} a(t).$$

First, we establish (15), that is,

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

The definition of $\bar{a}$ implies $\bar{a}(T) \geq \max\{a(T), 0\}$ for all $T > 0$, and thus $\limsup_{T \to \infty} \bar{a}(T) \geq \max\{\hat{a}, 0\}$ (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 definition of $\hat{a}$ implies that for an arbitrary $\delta > 0$, there is a large enough $t_\delta \geq t_0$ such that

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

For a $T > t_\delta$, let

$$t^*(T) \in \operatorname*{arg\,max}_{t \in [0,T]} a(t),$$

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

$$\bar{a}(T) = \frac{1}{T} \sum_{p \in \mathcal{P}} \int_0^{t^*(T)} a_p(s)\, \mathbb{1}_p(\sigma(s))\, \mathrm{d}s \leq a(t^*(T)) < \hat{a} + \delta.$$

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

$$\bar{a}(T) = \frac{1}{T} \sum_{p \in \mathcal{P}} \int_0^{t^*(t_\delta)} a_p(s)\, \mathbb{1}_p(\sigma(s))\, \mathrm{d}s.$$

Combining the two cases above, we obtain

$$\bar{a}(T) \leq \max\left\{ \hat{a} + \delta,\ \frac{1}{T} \sum_{p \in \mathcal{P}} \int_0^{t^*(t_\delta)} a_p(s)\, \mathbb{1}_p(\sigma(s))\, \mathrm{d}s \right\}$$

for all $T > t_\delta$. Hence

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

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

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

and thus (15) holds.

Second, we prove (16), that is,

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

As the index set $\mathcal{P}$ is finite, the definitions of $\hat{a}_p$ and $\check{a}_p$ imply that for an arbitrary $\bar{\delta} > 0$, there is a large enough $\bar{t}_\delta \ge 0$ such that

$$(\check{a}_p - \bar{\delta})\, \mathbb{1}_p(\sigma(t)) \le a_p(t)\, \mathbb{1}_p(\sigma(t)) \le (\hat{a}_p + \bar{\delta})\, \mathbb{1}_p(\sigma(t))$$

for all $t > \bar{t}_\delta$ and $p \in \mathcal{P}$. Therefore, we have

$$\hat{a} \le \limsup_{t\to\infty} \sum_{p\in\mathcal{P}} (\hat{a}_p + \bar{\delta})\rho_p(t) = \limsup_{t\to\infty} \sum_{p\in\mathcal{P}} \hat{a}_p \rho_p(t) + \bar{\delta},$$

and

$$\hat{a} \ge \limsup_{t\to\infty} \sum_{p\in\mathcal{P}} (\check{a}_p - \bar{\delta})\rho_p(t) = \limsup_{t\to\infty} \sum_{p\in\mathcal{P}} \check{a}_p \rho_p(t) - \bar{\delta}.$$

Thus (16) holds as $\bar{\delta} > 0$ is arbitrary.

REFERENCES

[1] Roy L. Adler, Alan G. Konheim, and M. H. McAndrew. 1965. Topological entropy. *Trans. Amer. Math. Soc.* 114, 2 (Feb. 1965), 309–319. https://doi.org/10.2307/1994177

[2] Alexander Yu Aleksandrov, Yangzhou Chen, Alexey V. Platonov, and Liguo Zhang. 2011. Stability analysis for a class of switched nonlinear systems. *Automatica* 47, 10 (2011), 2286–2291. https://doi.org/10.1016/j.automatica.2011.08.016

[3] Guillaume O. Berger and Raphaël M. Jungers. 2020. 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*. Sydney, Australia, 11 pages. https://doi.org/10.1145/3365365.3382195

[4] Vladimir A. Boichenko and Gennadij A. Leonov. 1998. Lyapunov's direct method in estimates of topological entropy. *Journal of Mathematical Sciences* 91, 6 (1998), 3370–3379. https://doi.org/10.1007/BF02434914

[5] Rufus Bowen. 1971. Entropy for group endomorphisms and homogeneous spaces. *Trans. Amer. Math. Soc.* 153 (Jan. 1971), 401–414. https://doi.org/10.1090/S0002-9947-1971-0274707-X

[6] Rufus Bowen. 1971. Periodic points and measures for axiom A diffeomorphisms. *Trans. Amer. Math. Soc.* 154 (Feb. 1971), 377–397. https://doi.org/10.2307/1995452

[7] Carmen Chicone. 2006. *Ordinary Differential Equations with Applications* (2 ed.). Vol. 34. Springer, New York, NY, USA. https://doi.org/10.1007/0-387-35794-7

[8] Fritz Colonius. 2012. Minimal bit rates and entropy for exponential stabilization. *SIAM Journal on Control and Optimization* 50, 5 (Oct. 2012), 2988–3010. https://doi.org/10.1137/110829271

[9] Fritz Colonius and Christoph Kawan. 2009. Invariance entropy for control systems. *SIAM Journal on Control and Optimization* 48, 3 (May 2009), 1701–1721. https://doi.org/10.1137/080713902

[10] Fritz Colonius, Christoph Kawan, and Girish N. Nair. 2013. A note on topological feedback entropy and invariance entropy. *Systems & Control Letters* 62, 5 (May 2013), 377–381. https://doi.org/10.1016/j.sysconle.2013.01.008

[11] Charles A. Desoer and Mathukumalli Vidyasagar. 2009. *Feedback Systems: Input-Output Properties*. SIAM, Philadelphia, PA, USA. https://doi.org/10.1137/1.9780898719055

[12] Efim I. Dinaburg. 1970. The relation between topological entropy and metric entropy. *Doklady Akademii Nauk SSSR* 190, 1 (1970), 19–22. In Russian.

[13] João P. Hespanha, Antonio Ortega, and Lavanya Vasudevan. 2002. Towards the control of linear systems with minimum bit-rate. In *15th International Symposium on Mathematical Theory of Networks and Systems*. Notre Dame, IN, USA, 1–15. https://www3.nd.edu/~mtns/papers/13040_3.pdf

[14] Josef Hofbauer and Karl Sigmund. 1998. *Evolutionary Games and Population Dynamics*. Cambridge University Press. https://doi.org/10.1017/CBO9781139173179

[15] Anatole Katok and Boris Hasselblatt. 1995. *Introduction to the Modern Theory of Dynamical Systems*. Cambridge University Press, Cambridge, UK. https://doi.org/10.1017/CBO9780511809187

[16] Christoph Kawan and Yuri Latushkin. 2016. Some results on the entropy of non-autonomous dynamical systems. *Dynamical Systems* 31, 3 (July 2016), 251–279. https://doi.org/10.1080/14689367.2015.1111299

[17] Andrei N. Kolmogorov. 1958. A new metric invariant of transitive dynamical systems and automorphisms of Lebesgue spaces. *Doklady Akademii Nauk SSSR* 119, 5 (1958), 861–864. In Russian.

[18] Sergiy Kolyada and L'ubomír Snoha. 1996. Topological entropy of nonautonomous dynamical systems. *Random & Computational Dynamics* 4, 2-3 (1996), 205–233.

[19] Daniel Liberzon. 2003. *Switching in Systems and Control*. Birkhäuser, Boston, MA, USA. https://doi.org/10.1007/978-1-4612-0017-8

[20] Daniel Liberzon. 2012. *Calculus of Variations and Optimal Control Theory: A Concise Introduction*. Princeton University Press, Princeton, NJ, USA. https://press.princeton.edu/books/hardcover/9780691151878/calculus-of-variations-and-optimal-control-theory

[21] Daniel Liberzon. 2014. Finite data-rate feedback stabilization of switched and hybrid linear systems. *Automatica* 50, 2 (Feb. 2014), 409–420. https://doi.org/10.1016/j.automatica.2013.11.037

[22] Daniel Liberzon and João P. Hespanha. 2005. Stabilization of nonlinear systems with limited information feedback. *IEEE Trans. Automat. Control* 50, 6 (June 2005), 910–915. https://doi.org/10.1109/TAC.2005.849258

[23] Daniel Liberzon and Sayan Mitra. 2018. Entropy and minimal bit rates for state estimation and model detection. *IEEE Trans. Automat. Control* 63, 10 (Oct. 2018), 3330–3344. https://doi.org/10.1109/TAC.2017.2782478

[24] José L. Mancilla-Aquilar. 2000. A condition for the stability of switched nonlinear systems. *IEEE Trans. Automat. Control* 45, 11 (2000), 2077–2079. https://doi.org/10.1109/9.887629

[25] Girish N. Nair and Robin J. Evans. 2003. Exponential stabilisability of finite-dimensional linear systems with limited data rates. *Automatica* 39, 4 (April 2003), 585–593. https://doi.org/10.1016/S0005-1098(02)00285-6

[26] Girish N. Nair, Robin J. Evans, Iven M. Y. Mareels, and William Moran. 2004. Topological feedback entropy and nonlinear stabilization. *IEEE Trans. Automat. Control* 49, 9 (Sept. 2004), 1585–1597. https://doi.org/10.1109/TAC.2004.834105

[27] Claude E. Shannon. 1948. A mathematical theory of communication. *The Bell System Technical Journal* 27, 3 (July 1948), 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

[28] Robert Shorten, Fabian R. Wirth, Oliver Mason, Kai Wulff, and Christopher King. 2007. Stability criteria for switched and hybrid systems. *SIAM Rev.* 49, 4 (Nov. 2007), 545–592. https://doi.org/10.1137/05063516X

[29] Hussein Sibai and Sayan Mitra. 2017. Optimal data rate for state estimation of switched nonlinear systems. In *20th ACM International Conference on Hybrid Systems: Computation and Control*. Pittsburgh, PA, USA, 71–80. https://doi.org/10.1145/3049797.3049799

[30] Eduardo D. Sontag. 2010. Contractive systems with inputs. In *Perspectives in Mathematical System Theory, Control, and Signal Processing*. Springer, 217–228. https://doi.org/10.1007/978-3-540-93918-4_20

[31] Sekhar Tatikonda and Sanjoy Mitter. 2004. Control under communication constraints. *IEEE Trans. Automat. Control* 49, 7 (July 2004), 1056–1068. https://doi.org/10.1109/TAC.2004.831187

[32] Mathukumalli Vidyasagar. 2002. *Nonlinear Systems Analysis*. SIAM. https://doi.org/10.1137/1.9780898719185

[33] Linh Vu and Daniel Liberzon. 2005. Common Lyapunov functions for families of commuting nonlinear systems. *Systems & Control Letters* 54, 5 (May 2005), 405–416. https://doi.org/10.1016/j.sysconle.2004.09.006

[34] Guosong Yang, João P. Hespanha, and Daniel Liberzon. 2019. On topological entropy and stability of switched linear systems. In *22nd ACM International Conference on Hybrid Systems: Computation and Control*. Montreal, Canada, 119–127. https://doi.org/10.1145/3302504.3311815

[35] Guosong Yang and Daniel Liberzon. 2018. Feedback stabilization of a switched linear system with an unknown disturbance under data-rate constraints. *IEEE Trans. Automat. Control* 63, 7 (July 2018), 2107–2122. https://doi.org/10.1109/TAC.2017.2767822

[36] Guosong Yang, A. James Schmidt, and Daniel Liberzon. 2018. On topological entropy of switched linear systems with diagonal, triangular, and general matrices. In *57th IEEE Conference on Decision and Control*. Miami Beach, FL, USA, 5682–5687. https://doi.org/10.1109/CDC.2018.8619087

[37] Guosong Yang, A. James Schmidt, Daniel Liberzon, and João P. Hespanha. 2020. Topological entropy of switched linear systems: General matrices and matrices with commutation relations. *Mathematics of Control, Signals, and Systems* 32, 3 (Sept. 2020), 411–453. https://doi.org/10.1007/s00498-020-00265-9