Topological entropy of switched nonlinear systems

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

ABSTRACT
This paper studies the topological entropy of continuous-time switched nonlinear systems. We construct a general upper bound for the topological entropy in terms of an average of the asymptotic supremum of the measures of Jacobian matrices of individual modes, weighted by the corresponding active rates. A general lower bound is constructed 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 construct 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 construct more conservative upper bounds that 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:

1 INTRODUCTION
In systems theory, topological entropy describes information generation rate in terms of the growth rate 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. Adler et al. first defined topological entropy as an extension of Kolmogorov’s metric entropy [17], 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, 27] and references therein). In general, a switched system does not inherit the stability properties of its individual modes. For example, a switched system with two stable modes may still be unstable [19, p. 19]. 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]. A simplest case of pairwise commuting modes is when the system functions are simultaneously diagonalizable. This motivates us to consider 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 a linear time-invariant control system, 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 [19, 25, 30]), which is equal to the topological entropy 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 for feedback stabilization of switched linear systems were established in [21, 35]. Similar data-rate conditions were constructed in [28] by extending the estimation entropy from [23]. In [36, 37], formulae...
and bounds for the topological entropy of switched linear systems were constructed in terms of the active rates of individual modes, an approach which is also adopted in the current work. Relations between topological entropy and stability 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 Joint Spectral Radius [3], while a formula for estimation entropy was derived under additional regularity conditions [31].

For switched nonlinear systems, topological entropy has not been explored so far. This paper’s main contribution is to construct upper and lower bounds for the topological entropy of switched nonlinear systems, which generalizes previous results for switched linear systems in [36, 37]. In Section 2, we present the definition of topological entropy for switched systems, and provide a standard construction of spanning and separated sets using a notion of grid. We also define switching-related quantities such as active rates of individual modes, and construct upper and lower bounds for the distance between two 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 construct 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 the corresponding active rates and maximized over the $\omega$-limit set. A general lower bound is constructed in terms of an active-rate-weighted average of the traces of these Jacobian matrices, minimised over the $\omega$-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 construct 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 construct 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.

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 $\text{Re}(a)$ 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 $A \in \mathbb{R}^{n \times m}$, denote by $\text{tr}(A)$, $\text{det}(A)$, and $\text{spec}(A)$ its trace, determinant, and spectrum (as a multiset) respectively. For a finite set $E$, denote by $|E|$ its cardinality. For a set $K \subset \mathbb{R}^n$, denote by $\text{vol}(K)$ and $\text{co}(K)$ its volume (Lebesgue measure) and convex hull, respectively. Denote by $\|v\|_\infty := \max_{1 \leq i \leq n} |v_i|$ the $\infty$-norm of a vector $v \in \mathbb{R}^n$, and by $\|A\|_\infty := \max_{1 \leq i \leq n} \sum_{j=1}^m |a_{ij}|$ the induced $\infty$-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 $J_x f(r, \varphi) \in \mathbb{R}^{k \times n}$ the Jacobian matrix of $f(r, \varphi)$ with respect to $x$ at $(r, \varphi)$.

2 PRELIMINARIES

Consider a family of continuous-time nonlinear dynamical systems

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

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 functions $f_p : \mathbb{R}^n \to \mathbb{R}^n$ are continuously differentiable. We are interested in the corresponding switched system defined by

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

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 $x$ and switched signal $\sigma$, which, under the above assumptions, is absolutely continuous in $t$, Differentiable in $x$, and satisfies the differential equation (2) away from discontinuities of $\sigma$, 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 $\sigma$ 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 $\mathbb{R}^n$ with center $x$:

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

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

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

or equivalently, for each $x \in E$, there is a point $x \in E$ such that $\|\xi_\sigma(t, x) - \xi_\sigma(t, x)\| < \varepsilon$ for all $t \in [0, T]$. Let $S(f_\sigma, \varepsilon, T, K)$ denote the minimal cardinality of a $(T, \varepsilon)$-spanning set, which is nondecreasing in $T$ and nonincreasing in $\varepsilon$. The topological entropy of the switched system (2) with initial set $K$ and switching signal $\sigma$ is defined in terms of the exponential growth rate of $S(f_\sigma, \varepsilon, T, K)$ by

$$h(f_\sigma, K) := \lim_{T \to \infty} \limsup_{\varepsilon \to 0} \frac{1}{T} \log S(f_\sigma, \varepsilon, T, K) \geq 0.$$ (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. Unless otherwise specified, we take $\| \cdot \|$ to be the $\infty$-norm on $\mathbb{R}^n$ or the induced $\infty$-norm on $\mathbb{R}^{n \times n}$.

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

$$\xi \notin B_{f_\sigma}(x, \varepsilon, T).$$ (6)
for all distinct points \( x, \hat{x} \in E \), or equivalently, there is a time \( t \in [0, T] \) such that \( \| x(t) - \hat{x}(t) \| \geq \varepsilon \). Let \( N(f_\varepsilon, \varepsilon, T, K) \) denote the maximal cardinality of a \((T, \varepsilon)\)-separated set, which is also nondecreasing in \( T \) and nonincreasing in \( \varepsilon \). The entropy of (2) can be equivalently formulated in terms of the exponential growth rate of \( N(f_\varepsilon, \varepsilon, T, K) \); cf. [15, p. 110].

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

\[
h(f_\varepsilon, K) = \lim_{T \to \infty} \frac{1}{T} \log N(f_\varepsilon, \varepsilon, T, K). \tag{7}
\]

### 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 \((T, \varepsilon)\)-spanning and \((T, \varepsilon)\)-separated sets by extending the notion of grid in [37]. Given a vector \( \theta = (\theta_1, \ldots, \theta_n) \in \mathbb{R}_+^n \), which may depend on \( T \) and \( \varepsilon \), 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

\[
\left| \bigcap_{i=1}^n \left[ \frac{2r_1}{\theta_i} \right] \right| \leq |G(\theta)| \leq \left| \bigcap_{i=1}^n \left( \left[ \frac{2r_2}{\theta_i} \right] + 1 \right) \right|.
\]

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

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

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,

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

Comparing the hyperrectangle \( R(x) \) with the open ball \( B_{f_{\varepsilon}}(x, \varepsilon, T) \) defined by (3), we obtain the following result, which extends [37, Lemma 2]; the proof is along the same lines and thus omitted here.

**Lemma 2.3.** 1. If the vector \( \theta \in \mathbb{R}_+^n \) is selected so that

\[
R(x) \subset B_{f_{\varepsilon}}(x, \varepsilon, T) \quad \forall x \in G(\theta), \tag{10}
\]

then the grid \( G(\theta) \) is \((T, \varepsilon)\)-spanning. If (10) holds for all \( T \geq 0 \) and \( \varepsilon > 0 \), and all \( \theta_i \) are nonincreasing in \( T \), then the topological entropy of the switched system (2) satisfies

\[
h(f_\varepsilon, K) \leq \lim_{T \to \infty} \frac{1}{T} \sum_{i=1}^n \log \frac{1}{\theta_i}. \tag{11}
\]

2. If the vector \( \theta \in \mathbb{R}_+^n \) is selected so that

\[
B_{f_{\varepsilon}}(x, \varepsilon, T) \subset R(x) \quad \forall x \in G(\theta), \tag{12}
\]

then the grid \( G(\theta) \) is \((T, \varepsilon)\)-separated. If (12) holds for all \( T \geq 0 \) and \( \varepsilon > 0 \), and all \( \theta_i \) are nonincreasing in \( T \), then the topological entropy of the switched system (2) satisfies

\[
h(f_\varepsilon, K) \geq \lim_{T \to \infty} \frac{1}{T} \sum_{i=1}^n \log \frac{1}{\theta_i}. \tag{13}
\]

\[\text{The upper bound (11) holds whenever the limit on the right-hand side exists, which is always the case for the grids in this paper; the same holds for the lower bound (13).}\]

### 2.3 Active times and active rates

In this subsection, we introduce switching-related quantities that are useful for computing the entropy of switched systems. The active time of the \( p \)-th mode over an interval \([0, T]\) is defined by

\[
\tau_p(t) := \int_0^t \mathbb{I}_p(\sigma(s)) \, ds, \quad p \in \mathcal{P} \tag{14}
\]

with the indicator function

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

We also define the active rate of the \( p \)-th mode over \([0, T]\) by

\[
\rho_p(t) := \frac{\tau_p(t)}{t}, \quad p \in \mathcal{P} \tag{15}
\]

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

\[
\hat{\rho}_p := \limsup_{t \to \infty} \rho_p(t), \quad p \in \mathcal{P}. \tag{16}
\]

Clearly, the active times \( \tau_p \) are nonnegative and nondecreasing, and satisfy \( \sum_{p \in \mathcal{P}} \tau_p(t) = t \) for all \( t \geq 0 \); the active rates \( \rho_p \) 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 (16), 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

\[
\lim_{T \to \infty} \frac{1}{T} \max_{t \in [0, T]} \sum_{p \in \mathcal{P}} c_p \tau_p(t) = \max \left\{ \limsup_{t \to \infty} \sum_{p \in \mathcal{P}} c_p \rho_p(t), 0 \right\}. \tag{17}
\]

Next, we present a technical lemma that generalizes this result to the case with a family of integrable functions.

**Lemma 2.4.** For a family \( \{a_p : p \in \mathcal{P}\} \) of integrable functions \( a_p : \mathbb{R}_{\geq 0} \to \mathbb{R} \), we have

\[
\lim_{T \to \infty} \frac{1}{T} \max_{t \in [0, T]} \sum_{p \in \mathcal{P}} a_p(s) \mathbb{I}_p(\sigma(s)) \, ds = \max \left\{ \limsup_{t \to \infty} \sum_{p \in \mathcal{P}} \frac{1}{t} \int_0^t a_p(s) \mathbb{I}_p(\sigma(s)) \, ds, 0 \right\}. \tag{17}
\]

Moreover, the first term in the maximum on the right-hand side of (17) satisfies

\[
\limsup_{t \to \infty} \sum_{p \in \mathcal{P}} a_p \rho_p(t) \leq \limsup_{t \to \infty} \sum_{p \in \mathcal{P}} \frac{1}{t} \int_0^t a_p(s) \mathbb{I}_p(\sigma(s)) \, ds \leq \limsup_{t \to \infty} \sum_{p \in \mathcal{P}} \hat{a}_p \rho_p(t) \tag{18}
\]

with

\[
\hat{a}_p := \liminf_{t \to \infty} a_p(t), \quad \tilde{a}_p := \limsup_{t \to \infty} a_p(t), \quad p \in \mathcal{P}.
\]

Before proving Lemma 2.4, we note that the sum on the left-hand side of (17) is in fact the integral of \( a_\sigma \) over \([0, T]\), that is,

\[
\int_0^T a_\sigma(s) \, ds = \sum_{p \in \mathcal{P}} \int_0^T a_p(s) \mathbb{I}_p(\sigma(s)) \, ds. \tag{19}
\]

\[\text{If \( \{t \geq 0 : \sigma(t) = p\} \) is a bounded set, then the corresponding limit supremum and limit infimum are taken to be 0.}\]
Proof of Lemma 2.4. For brevity, we define the following functions \( \eta, \hat{a} : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R} \) and constant \( \hat{a} \in \mathbb{R} \cup \{\infty\} \):

\[
\eta(t) := \sum_{p \in \mathcal{P}} \int_{0}^{t} a(p) \mu_p(\sigma(s)) \, ds, \quad \hat{a}(t) := \frac{1}{t} \max_{t \in [0,T]} \eta(t),
\]

with \( \hat{a}(0) := \max\{a_{\sigma(0)}(0), 0\} \), and

\[
\hat{a} := \limsup_{t \to \infty} \frac{\eta(t)}{t}.
\]

First, we establish (17), that is, \( \limsup_{T \to \infty} \hat{a}(T) = \max\{\hat{a}, 0\} \).

The definition of \( \hat{a} \) implies \( \hat{a}(T) \geq \max(\eta(T)/T, 0) \) for all \( T > 0 \), and thus \( \limsup_{T \to \infty} \hat{a}(T) \geq \max\{\hat{a}, 0\} \) (in particular, if \( \hat{a} = \infty \) then \( \limsup_{T \to \infty} \hat{a}(T) = \infty \)). 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 \( t_\delta \geq 0 \) such that

\[
\eta(t) < (\hat{a} + \delta) t \quad \forall t > t_\delta.
\]

For a time \( T > t_\delta \), let

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

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

\[
\hat{a}(T) = \frac{\eta(t^*(T))}{t^*(T)} < \hat{a} + \delta.
\]

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

\[
\hat{a}(T) = \frac{\eta(t^*(T))}{T} = \frac{\eta(t^*(t_\delta))}{t^*(t_\delta)}.
\]

Combining the two cases above, we obtain

\[
\hat{a}(T) \leq \max\{\hat{a} + \delta, \frac{\eta(t^*(t_\delta))}{t^*(t_\delta)}\} \quad \forall T > t_\delta.
\]

Hence

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

As \( \delta > 0 \) is arbitrary, we have

\[
\limsup_{T \to \infty} \hat{a}(T) \leq \max\{\hat{a}, 0\},
\]

and thus (17) holds.

Second, we establish (18), that is,

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

Recall that the index set \( \mathcal{P} \) is finite. Following the definitions of \( \hat{a}_p \) and \( \bar{a}_p \), for an arbitrary \( \delta > 0 \), there is a large enough \( t_\delta \geq 0 \) such that for all \( p \in \mathcal{P} \), we have

\[
(\hat{a}_p - \delta) I_p(\sigma(t)) \leq a_p(t) I_p(\sigma(t)) \leq (\hat{a}_p + \delta) I_p(\sigma(t))
\]

for all \( t > t_\delta \). Therefore, we have

\[
\hat{a} \leq \limsup_{t \to \infty} \sum_{p \in \mathcal{P}} (\hat{a}_p + \delta) \rho_p(t) = \limsup_{t \to \infty} \sum_{p \in \mathcal{P}} \hat{a}_p \rho_p(t) + \delta,
\]

\[
\hat{a} \geq \limsup_{t \to \infty} \sum_{p \in \mathcal{P}} (\hat{a}_p - \delta) \rho_p(t) = \limsup_{t \to \infty} \sum_{p \in \mathcal{P}} \hat{a}_p \rho_p(t) - \delta.
\]

Thus (18) holds as \( \delta > 0 \) is arbitrary. \( \Box \)

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 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) at time \( t \) from initial set \( K \).

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

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

For standard norms, there are explicit formulae for the matrix measure; for example, for the induced \( \infty \)-norm, we have

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

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

\[
-\mu(-A) \leq \Re(\lambda) \leq \mu(A) \leq \|A\| \quad \forall A \in R^{n \times n}, \forall \lambda \in \text{spec}(A).
\]

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

\[
e^{\eta_\sigma(t)} \|\dot{x} - x\| \leq \|\xi_\sigma(t, x) - \xi_\sigma(t, x)\| \leq e^{\eta_\sigma(t)} \|\dot{x} - x\| \quad \forall t \geq 0
\]

with

\[
\eta_\sigma(t) := \max_{v \in \text{co}(K)} \sum_{p \in \mathcal{P}} \int_{0}^{t} \min_{\sigma_0} \mu(-J_x f_p(v)) I_p(\sigma(s)) \, ds,
\]

\[
\bar{\eta}_\sigma(t) := \max_{v \in \text{co}(K)} \sum_{p \in \mathcal{P}} \int_{0}^{t} \mu(J_x f_p(\xi_\sigma(s, v))) I_p(\sigma(s)) \, ds.
\]

Also, the reachable set of (2) satisfies

\[
\text{vol}(\xi_\sigma(t, K)) \geq e^{\gamma_\sigma(t)} \text{vol}(K) \quad \forall t \geq 0
\]

with

\[
\gamma_\sigma(t) := \min_{v \in \text{co}(K)} \sum_{p \in \mathcal{P}} \int_{0}^{t} \text{tr}(J_x f_p(\xi_\sigma(s, v))) I_p(\sigma(s)) \, ds.
\]

Note that \( \eta_\sigma, \bar{\eta}_\sigma, \) and \( \gamma_\sigma \) 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 the transformation (19).

Proof of Proposition 2.5. Let’s consider a linear time-varying (LTV) system

\[
\dot{x} = A(t) x
\]

with a piecewise continuous matrix-valued function \( A : \mathbb{R}_{\geq 0} \rightarrow R^{n \times n} \). 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 [11, Th. 27, p. 34]

\[
e^{\int_{t}^{t} \mu(-A(s)) \, ds} \|v\| \leq \|\xi(t, v)\|
\]

\[
\|\Phi_A(t, 0) v\| \leq e^{\int_{0}^{t} \mu(A(s)) \, ds} \|v\| \quad \forall t \geq 0
\]

and

\[
\det(\Phi_A(t, 0)) = e^{\int_{0}^{t} \text{tr}(A(s)) \, ds} \quad \forall t \geq 0
\]
(Liouville’s formula [7, Prop. 2.18, p. 152]).

First, we write the Jacobian matrix $J_\sigma \xi_\sigma(t, v)$ 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 $J_\sigma \xi_\sigma(0, v) = I$ and

$$
d J_\sigma \xi_\sigma(t, v) = J_\sigma \xi_\sigma(t, v) = J_\sigma f_\sigma(t) (\xi_\sigma(t, v)) J_\sigma \xi_\sigma(t, v)
$$

for all $t \geq 0$ that are not switching times. Hence for each fixed $v \in \mathbb{R}^n$, the Jacobian matrix $J_\sigma \xi_\sigma(t, v)$ is the principal fundamental matrix solution [7, Def. 2.12, p. 150] to the LTV system (24) with the system matrix $A(t) = J_\sigma f_\sigma(t) (\xi_\sigma(t, v))$, and thus equal to the state-transition matrix $\Phi_{\xi}(0)$. Let $v(\rho) := \rho x + (1 - \rho) \xi_\sigma$ for $\rho \in [0, 1]$. Then $v(\rho) \in \text{co}(\xi_\sigma(t, K))$ for all $\rho \in [0, 1]$ and $t \geq 0$, and $d_\beta v(\rho) = \xi_\sigma(t, x) - \xi_\sigma(t, x)$. Thus

$$
d_\beta (f_\sigma(t) (\xi_\sigma(t, x)) - f_\sigma(t) (\xi_\sigma(t, x))) = f_\sigma(t) (\xi_\sigma(t, x) - f_\sigma(t) (\xi_\sigma(t, x)))
$$

for all $t \geq 0$ that are not switching times. Hence $\xi_\sigma(t, x) - \xi_\sigma(t, x)$ is the solution with initial state $\hat{x} = x$ to the LTV system (24) with the system matrix $A(t) := \int_0^t J_\sigma f_\sigma(t) (\xi_\sigma(t, x)) \, dt$. Thus the lower bound in (25) implies

$$
\left\| \xi_\sigma(t, x) - \xi_\sigma(t, x) \right\| \geq e^\int_0^t J_\sigma f_\sigma(t) (\xi_\sigma(t, x)) \, dt \left\| \xi_\sigma(t, x) - \xi_\sigma(t, x) \right\|
$$

for all $t \geq 0$. Moreover, as the function $\mu$ is convex, for all $\rho \in P$, we have

$$
-\mu \left( -\int_0^t J_\sigma f_\sigma(\rho) \, d\rho \right) \geq \int_0^t -\mu(\xi_\sigma(t, K)) \, d\rho
$$

for all $t \geq 0$. Hence the lower bound in (22) holds.

Finally, Liouville’s formula (26) with $A(t) = J_\sigma f_\sigma(t) (\xi_\sigma(t, v))$ and $A_\sigma(t, 0) = J_\sigma \xi_\sigma(t, v)$ from the first part implies

$$
\text{vol}(\xi_\sigma(t, K)) = \int_K | \det(J_\sigma \xi_\sigma(t, v)) | \, dv
$$

and

$$
\text{vol}(K) = e^{\int_0^t \text{tr}(J_\sigma f_\sigma(s) (\xi_\sigma(s, v))) \, ds}
$$

for all $t \geq 0$, that is, (23) holds.

Additionally, we provide the following upper and lower bounds for the distance between two solutions in terms of the active times $\tau_{p}$, which are more conservative than those in [22] but illustrate the effect of switching and will be useful for establishing bounds for the topological entropy of switched systems with diagonal modes.

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

$$
e^{\sum_{p \in \mathcal{P}} \mu_p^{-1} (\tau_{p}) (\hat{x} - x)} \leq e^{\sum_{p \in \mathcal{P}} \mu_p^{-1} (\tau_{p}) (\hat{x} - x)} \forall t \geq 0
$$

with

$$
u_p := \min_{s \in [0, 1]} \mu(J_\sigma f_\sigma(s) (\xi_\sigma(s, v))), \quad p \in \mathcal{P},
$$

where the active times $\tau_{p}$ are defined by (14).

**Proof.** The upper bound in (27) follows from the upper bound in (22) and the property that

$$
\eta_a(t) \leq \sum_{p \in \mathcal{P}} \int_0^t \left( \max_{v \in \text{co}(\xi_\sigma(t, x))} \mu(J_\sigma f_\sigma(s) (\xi_\sigma(s, v))) \right) \mathbb{I}_p (s) \, ds
$$

and

$$
\sum_{p \in \mathcal{P}} \mu_p (\tau_{p}) \int_0^t \mathbb{I}_p (s) \, ds = \sum_{p \in \mathcal{P}} \mu_p (\tau_{p}) \forall t \geq 0.
$$

The lower bound in (27) follows from the lower bound in (22) and the property that

$$
\eta_a(t) \geq \sum_{p \in \mathcal{P}} \mu_p (\tau_{p}) \int_0^t \mathbb{I}_p (s) \, ds = \sum_{p \in \mathcal{P}} \mu_p (\tau_{p}) \forall t \geq 0.
$$

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

1. for each $p \in \mathcal{P}$, the measure of Jacobian matrix $\mu(J_\sigma f_\sigma(\nu))$ has a global upper bound $\overline{\mu}_p$, or

2. the convex hull of initial set $\text{co}(K)$ is a subset of a compact positively invariant set $S$, and let $\overline{\mu}_p := \max_{s \in S} \mu(J_\sigma f_\sigma(\nu))$.

Then (27) implies that for all initial states $x, \hat{x} \in K$, we have

$$
\left\| \xi_\sigma(t, x) - \xi_\sigma(t, x) \right\| \leq e^{\sum_{p \in \mathcal{P}} \overline{\mu}_p^{-1} (\tau_{p}) (\hat{x} - x)} \forall t \geq 0,
$$

which is more conservative but simpler than the upper bounds in (22) and (27). Similarly, a more conservative but simpler lower bound than the ones in (22) and (27) can be constructed for the cases with globally lower-bounded measures $\mu(J_\sigma f_\sigma(\nu))$, or with a compact positively invariant set containing $K$. Similar results hold for the lower bound (23). On the other hand, without a global bound
for each $\mu(JK_f^p(u))$ and without a compact positively invariant set, the functions $\eta_\sigma, \tilde{\eta}_\sigma, \mu_p$, and $\tilde{\mu}_p$ in (22) and (27) may be unbounded.

3 ENTROPY OF GENERAL SWITCHED NONLINEAR 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

$$
h(f_a, K) \leq \max \left\{ \limsup_{t \to \infty} \sum_{p \in P} n\tilde{\mu}_p \rho_p(t), 0 \right\}
$$

with

$$
\tilde{\mu}_p := \limsup_{s \to \infty} \max_{\sigma(s)=p, v \in E(K)} \mu(JK_f^p(\xi\sigma(s), v)), \quad p \in P, \tag{29}
$$

and lower-bounded by

$$
h(f_a, K) \geq \max \left\{ \limsup_{t \to \infty} \sum_{p \in P} \tilde{\chi}_p \rho_p(t), 0 \right\}
$$

with

$$
\tilde{\chi}_p := \liminf_{s \to \infty} \min_{\sigma(s)=p, v \in K} \text{tr}(JK_f^p(\xi\sigma(s), v)), \quad p \in P, \tag{31}
$$

where the active rates $\rho_p$ are defined by (15).

Proof. First, we prove the upper bound (28). Fix a time horizon $T \geq 0$ and a radius $\epsilon > 0$. The upper bound in (22) implies that for all initial states $x, \bar{x} \in K$, the corresponding solutions to (2) satisfy

$$
\max_{t \in [0,T]} \|\bar{x}(t) - \bar{\xi}(t, x)\| \leq \max_{t \in [0,T]} \tilde{\eta}_\sigma(t) \|\bar{x} - x\|. \tag{32}
$$

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

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

and the corresponding hypercubes $R(x)$ defined by (9). Comparing (9) and (32) to (3), we see that $R(x) \subset B_f^\epsilon(x, \epsilon, T)$ for all $x \in G(\theta)$. Then Lemma 2.3 implies that $G(\theta)$ is $(T, \epsilon)$-spanning. As $T \geq 0$ and $\epsilon > 0$ are arbitrary, and all $\theta_i$ are nonincreasing in $T$, the upper bound (11) implies

$$
h(f_a, K) \leq \limsup_{T \to \infty} \frac{n \log(1/\theta_i)}{T} = \limsup_{T \to \infty} \frac{1}{T} \max_{t \in [0,T]} \tilde{\eta}_\sigma(t) + \limsup_{T \to \infty} \frac{n \log(1/\epsilon)}{T} \leq \limsup_{T \to \infty} \frac{1}{T} \max_{t \in [0,T]} \sum_{p \in P} \int_0^T a_1p(s) \mathbb{L}_p(\sigma(s)) ds
$$

with

$$
a_1p(s) := \max_{v \in E(K)} n\mu(JK_f^p(\xi\sigma(s), v)), \quad p \in P. \tag{33}
$$

Then the upper bound (28) follows from (17) and the upper bound in (18) with the functions $a_p(t) := a_1p(t)$ for $p \in P$.

Second, we establish the lower bound (30) using volume-based arguments. Fix a time horizon $T \geq 0$ and a radius $\epsilon > 0$. The lower bound (23) implies that the reachable set $\xi\sigma(T, K)$ of (2) satisfies

$$
\text{vol}(\xi\sigma(T, K)) \geq e^{\epsilon\sigma(T)} \text{vol}(K).
$$

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

$$
\text{vol}(\xi\sigma(T, K)) \leq \sum_{x \in E} \text{vol}(x \in [\epsilon, \xi\sigma(T)]).
$$

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

$$
\text{vol}(\xi\sigma(T, K)) \leq \sum_{x \in E} \|x \in [\epsilon, \xi\sigma(T)]\| \leq |E|/2\epsilon^n.
$$

Therefore, the minimal cardinality of a $(T, \epsilon)$-spanning set satisfies

$$
S(f_a, \epsilon, T, K) = |E| \geq \text{vol}(\xi\sigma(T, K))/(2\epsilon)^n \geq e^{\epsilon\sigma(T)} \text{vol}(K)/(2\epsilon)^n,
$$

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

$$
h(f_a, K) \geq \limsup_{T \to \infty} \frac{1}{T} \log(e^{\epsilon\sigma(T)} \text{vol}(K)/(2\epsilon)^n) = \limsup_{T \to \infty} \frac{1}{T} \sum_{p \in P} \int_0^T a_2p(s) \mathbb{L}_p(\sigma(s)) ds
$$

with

$$
a_2p(s) := \min_{v \in E(K)} \text{tr}(JK_f^p(\xi\sigma(s), v)), \quad p \in P. \tag{34}
$$

Then the lower bound (30) follows from the lower bound in (18) with the functions $a_p(t) := a_2p(t)$ for $p \in P$ and the property that $h(f_a, K) \geq 0$. \hfill \Box

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

$$
h(f_p, K) \leq \max\{n\tilde{\mu}_p, 0\} \tag{33}
$$

with the constant $\tilde{\mu}_p$ defined by (29), and lower-bounded by

$$
h(f_p, K) \geq \max\{\tilde{\chi}_p, 0\} \tag{34}
$$

with the constant $\tilde{\chi}_p$ defined by (31).

Based on the upper bound (28), 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

$$
h(f_a, K) \leq \sum_{p \in P} \max\{n\tilde{\mu}_p, 0\} \tilde{\rho}_p \tag{35}
$$

with the asymptotic active rates $\tilde{\rho}_p$ defined by (16), and also by

$$
h(f_a, K) \leq \max_{p \in P} \max\{n\tilde{\mu}_p, 0\}, \tag{36}
$$

where the constants $\tilde{\mu}_p$ are defined by (29).
Proof. First, the upper bound (28) and the subadditivity of limit suprema imply
\[
\begin{align*}
 h(f_0, K) & \leq \max \left\{ \sum_{p \in \mathcal{P}} \limsup_{t \to \infty} n \hat{\mu}_p \rho_p(t), 0 \right\} \\
 & \leq \sum_{p \in \mathcal{P}} \max \{n \hat{\mu}_p, 0\} \limsup_{t \to \infty} \rho_p(t) = \sum_{p \in \mathcal{P}} \max \{n \hat{\mu}_p, 0\} \hat{\rho}_p.
\end{align*}
\]
Second, the upper bound (28) also implies
\[
\begin{align*}
 h(f_0, K) & \leq \max \left\{ \limsup_{t \to \infty} \max_{p \in \mathcal{P}} n \hat{\mu}_p \sum_{p \in \mathcal{P}} \rho_p(t), 0 \right\} \\
 & = \max \left\{ \max_{p \in \mathcal{P}} n \hat{\mu}_p, 0\right\} = \max \max \{n \hat{\mu}_p, 0\}. \quad \square
\end{align*}
\]

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
\[
 f_p(x) = A_p x \quad \forall x \in \mathbb{R}^n.
\]

Then the constants \( \hat{\mu}_p \) and \( \hat{\chi}_p \) defined by (29) and (31) satisfy
\[
 \hat{\mu}_p = \mu(A_p), \quad \hat{\chi}_p = \text{tr}(A_p) \quad \forall p \in \mathcal{P}.
\]

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 \( \hat{\chi}_p \) defined by (29) and (31) only depend on the measure and the trace of each Jacobian matrix \( J_x f_p(v) \) over the \( \omega \)-limit set from the convex hull of initial set \( co(K) \) and that from \( K \), respectively, instead of over all reachable points from \( co(K) \) and \( K \), respectively. In particular, (28), (35), and (36) 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 \( co(K) \), then the upper bounds (28), (35), and (36) hold with \( \hat{\mu}_p \) in place of \( \hat{\rho}_p \). Similarly, a more conservative but simpler lower bound than (30) can be constructed for the case with globally lower-bounded traces \( \text{tr}(J_x f_p(v)) \), 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 (28), the upper bound (35) depends on the asymptotic active rates \( \hat{\rho}_p \) instead of the active rates \( \rho_p \), and the upper bound (36) does not involve active rates at all. If a global upper bound \( \hat{\mu}_p \) is used in place of \( \hat{\rho}_p \) for each \( p \), then (36) is independent of switching.

4. The upper bound (28) is tighter than the upper bounds (35) and (36). The upper bounds (35) and (36) 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 (28), (35), and (36) are illustrated numerically in Example 3.6 below.

Example 3.6. Consider the following switched nonlinear system in the nonnegative orthant \( \mathbb{R}^n_{\geq 0} \) from [2]:
\[
 \dot{x}_i = f^i_{\sigma}(x) := \left( \nu_i \sigma + \sum_{j=1}^{n} a_{ij} \sigma x_j \right) x_i, \quad i \in \{1, \ldots, n\} \quad (37)
\]
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 \(f_i^{\sigma}(t, x_i)\) the \(i\)-th scalar component of the solution \(\xi_\sigma(t, x)\), and by \(E_i^{\sigma}(t, K) := \{f_i^{\sigma}(t, x_i) : x \in K\}\) 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 (38) that are generally tighter than the results of simply applying the bounds from Section 3 to (38).

**Theorem 4.1.** The topological entropy of the switched diagonal system (38) satisfies

\[
\limsup_{T \to \infty} \frac{1}{T} \sum_{t=0}^{n} \max_{p \in P} \sum_{\rho_p(t)} d_p f_p(t) \leq h(f_\sigma, K)
\]

\[
\leq \limsup_{T \to \infty} \frac{1}{T} \sum_{t=0}^{n} \max_{p \in P} \sum_{\rho_p(t)} d_p f_p(t)
\]

with

\[
d^i_p := \inf_{s \in [0, \sigma(s) = p]} \min_{\rho(v, s) = p} (f^i_p(v, s))
\]

and

\[
\bar{d}^i_p := \sup_{s \in [0, \sigma(s) = p]} \max_{\rho(v, s) = p} (f^i_p(v, s))
\]

for \(i \in \{1, \ldots, n\}\) and \(p \in P\), where the active times \(\rho_p\) are defined by (14); it is also upper-bounded by

\[
\limsup_{T \to \infty} \frac{1}{T} \sum_{t=0}^{n} \max_{p \in P} \sum_{\rho_p(t)} d_p f_p(t)
\]

with

\[
d^i_p := \inf_{s \in [0, \sigma(s) = p]} \min_{\rho(v, s) = p} (f^i_p(v, s))
\]

and

\[
\bar{d}^i_p := \sup_{s \in [0, \sigma(s) = p]} \max_{\rho(v, s) = p} (f^i_p(v, s))
\]

for \(i \in \{1, \ldots, n\}\) and \(p \in P\), where the active rates \(\rho_p\) are defined by (15).

**Proof.** Fix a time horizon \(T > 0\) and a radius \(\epsilon > 0\). Applying the upper and lower bounds (27) and the upper bound in (22) for the distance between solutions to each scalar component of (38), we obtain that for all initial states \(x, \hat{x} \in K\), the corresponding solutions satisfy (recall that we take \(\|\cdot\|\) to be the \(\infty\)-norm; see Remark 2.1)

\[
\max_{I \subseteq [0, T]} \epsilon \sum_{i=1}^{n} \max_{p \in P} \sum_{\rho_p(t)} d^i_p (\rho_p(t)) \|x_i - x_i\| \\
\leq \max_{I \subseteq [0, T]} \epsilon \sum_{i=1}^{n} \max_{p \in P} \sum_{\rho_p(t)} d^i_p (\rho_p(t)) \|x_i - x_i\|
\]

with

\[
\mu^i_p := \min_{s \in [0, \sigma(s) = p]} (f^i_p(v, s))
\]

and

\[
\bar{d}^i_p := \sup_{s \in [0, \sigma(s) = p]} (f^i_p(v, s))
\]

for \(i \in \{1, \ldots, n\}\) and \(p \in P\), and also

\[
\max_{t \in [0, T]} \|\xi_\sigma(t, x) - \xi_\sigma(t, \hat{x})\| \leq \max_{I \subseteq [0, T]} \epsilon \sum_{i=1}^{n} \max_{p \in P} \sum_{\rho_p(t)} d^i_p (\rho_p(t)) \|x_i - x_i\|
\]

with

\[
\epsilon \sum_{i=1}^{n} \max_{p \in P} \sum_{\rho_p(t)} (f^i_p(v, s)) \|x_i - x_i\|
\]

for \(i \in \{1, \ldots, n\}\) and \(p \in P\), and also

\[
\max_{t \in [0, T]} \|\xi_\sigma(t, x) - \xi_\sigma(t, \hat{x})\| \leq \max_{I \subseteq [0, T]} \epsilon \sum_{i=1}^{n} \max_{p \in P} \sum_{\rho_p(t)} d^i_p (\rho_p(t)) \|x_i - x_i\|
\]
for $i \in \{1, \ldots, n\}$.

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

$$
\theta_i := e^{-\max_{\epsilon \in [0, T]} \sum_{p \in P} \sum_{\epsilon > \theta_i} \frac{\mu^i_p (t)}{v_p (t)}}, \quad i \in \{1, \ldots, n\},
$$

and the corresponding hyperrectangles $R(x)$ defined by (9). Comparing (9) and (43) to (3), we see that $R(x) \subset B_p (x, \epsilon, T)$ for all $x \in G(\theta)$. Then Lemma 2.3 implies that $G(\theta)$ is $(T, \epsilon)$-spanning. As $T \geq 0$ and $\epsilon > 0$ are arbitrary, and all $\theta_i$ are nonincreasing in $T$, the upper bound (11) implies

$$
h(f_\delta, K) \leq \lim_{T \to \infty} \inf \frac{1}{T} \left( \sum_{i=1}^n \log \left( \frac{1}{\theta_i} \right) \right)
$$

that is, the upper bound in (39) holds.

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

$$
\theta_i := e^{-\max_{\epsilon \in [0, T]} \sum_{p \in P} \sum_{\epsilon > \theta_i} \frac{\mu^i_p (t)}{v_p (t)}}, \quad i \in \{1, \ldots, n\},
$$

we can show that $G(\theta)$ is $(T, \epsilon)$-spanning, and the lower bound (13) implies the lower bound in (39).

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

$$
\theta_i := e^{-\max_{\epsilon \in [0, T]} \sum_{p \in P} \sum_{\epsilon > \theta_i} \frac{\mu^i_p (t)}{v_p (t)}}, \quad i \in \{1, \ldots, n\},
$$

we can show that $G(\theta)$ is $(T, \epsilon)$-spanning, and the upper bound (11) implies

$$
h(f_\delta, K) \leq \lim_{T \to \infty} \inf \frac{1}{T} \left( \sum_{i=1}^n \log \left( \frac{1}{\theta_i} \right) \right)
$$

with the asymptotic active rates $\hat{\rho}_p$ defined by (16), and also by

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

where the constants $\hat{a}_p^i$ are defined by (42).

**Remark 4.3.** Consider the case where all the functions $f_p^i$ in (38) are linear, that is, there is a family of diagonal matrices $\{D_p = \text{diag}(a^i_p, \ldots, a^n_p) \in \mathbb{R}^{n \times n} : p \in P\}$ such that

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

Then the constants $\hat{a}_p^i$, $\hat{d}_p^i$, and $\hat{d}_p^i$ defined by (40) and (42) satisfy

$$
\hat{a}_p^i = \hat{d}_p^i = \hat{d}_p^i = a^i_p \quad \forall i \in \{1, \ldots, n\}, \forall p \in P.
$$

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 $\hat{a}_p^i$ and $\hat{a}_p^i$ defined by (40) depend on the partial derivatives $(f_p^i)^\prime (u_i)$ over the convex hull of all reachable points from the initial set $K$ and over all reachable points from the convex hull of all reachable points from $\text{co}(K)$, respectively, whereas $\hat{d}_p^i$ defined by (42) only depends on $(f_p^i)^\prime (u_i)$ over the $\omega$-limit set from $\text{co}(K)$. Their difference is due to the different constructions of the upper and lower bounds in (22) and the upper bound in (27). In particular, (41), (45), and (46) 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 $(f_p^i)^\prime (u_i)$ has a global upper bound $\hat{a}_p^i$, or an upper bound $\hat{a}_p^i$ over a compact positively invariant set containing $\text{co}(K)$, then the upper bound in (39) and the upper bounds (41), (45), and (46) hold with $\hat{a}_p^i$ in place of $\hat{a}_p^i$ and $\hat{d}_p^i$. Similarly, a more conservative but simpler lower bound than the one in (39) can be constructed for the case with globally lower-bounded $(f_p^i)^\prime (u_i)$, or with a compact positively invariant set containing $K$.

3. For a fixed family of functions $\{f_p : p \in P\}$, compared with the upper bound in (39) and the upper bound (41), the upper bound (45) depends on the asymptotic active rates $\hat{\rho}_p$ instead of the active rates $\rho_p$, and the upper bound (46) does not involve active rates at all. If a global upper bound $\hat{d}_p^i$ is used in place of $\hat{d}_p^i$ for each $p$ and $i$, then (46) is independent of switching.

4. The upper bound (41) is tighter than the upper bounds (28), (45), and (46), while (45) and (46) are tighter than the upper bounds (35) and (36), respectively. The upper bound in (39) and the upper bound (41) 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 $\omega$-limit set from $\text{co}(K)$ contains all reachable points from $\text{co}(K)$, then the former is tighter than the latter. The same conclusion holds between the upper bound in (39) and the upper bounds (35) and (36). The upper bounds (45) and (46) are both useful in the same sense, as it is possible that $\sum_{p \in P} \hat{\rho}_p > 1$. The relations between the upper bounds (28), (41), (45), and (46), and the one in (39) are illustrated numerically in Example 4.5 below.

**Example 4.5.** Consider the switched nonlinear system (37) in the nonnegative orthant $\mathbb{R}^n_{\geq 0}$ in Example 3.6. In this example, we...
consider the case where the interaction terms $a_{ij}^p = 0$ for all $j \neq i$ and $p \in \mathcal{P}$. Then (37) becomes the switched diagonal system

$$\dot{x}_i = f_i^p(x_i) := (r_i^p + a_{ii}^p x_i) x_i, \quad i \in \{1, \ldots, n\}. \quad (47)$$

Note that for each $p \in \mathcal{P}$ and $i \in \{1, \ldots, n\}$, we have $f_i^p(x_i) < 0$ if $x_i > \max(-r_i^p/a_{ii}^p, 0)$. Thus (47) is UUB in $\mathbb{R}^n_{\geq 0}$ and its $\omega$-limit set is a subset of the positively invariant set $\Omega$.

$$\Omega := \bigcap_{i=1}^n \left\{ 0, \max \left( \max_{p \in \mathcal{P}} -\frac{r_i^p}{a_{ii}^p}, 0 \right) \right\}. \quad (29)$$

Consider the switched diagonal system (47) 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 (47) with the switching signals $\sigma_1$ and $\sigma_2$ defined in Example 3.6 are plotted in Fig. 2 below.

![Figure 2: Trajectories of the switched diagonal system (47) 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 $\sigma_1$ with initial states (4, 3), (3, 0), (0, 2), and (0, 0); (d) switching signal $\sigma_2$ with initial state (4, 3). The circles mark the beginning of segments after switching. The gray rectangle represents the positively invariant set $\Omega = \{0, 3\} \times \{0, 2\}$.](image)

The Jacobian matrices of individual modes of (47) are given by

$$J_{f_1}(v) = \begin{bmatrix} -1 - 2v_1 & 0 \\ 0 & 2 - 2v_2 \end{bmatrix}, \quad J_{f_2}(v) = \begin{bmatrix} 3 - 2v_1 & 0 \\ 0 & -1 - 2v_2 \end{bmatrix}. \quad (48)$$

As (47) is UUB and its $\omega$-limit set is a subset of $\Omega$, for all initial sets $K \subset \mathbb{R}^n_{\geq 0}$, one can obtain upper bounds for the constants $\hat{a}_p$ and $\hat{a}'_p$ defined by (29) and (42) by replacing the limit suprema over $\{t \geq 0 : \sigma(s) = p\}$ and maxima over $\text{co}(K)$ in (29) and (42) with maxima over $\Omega$. Hence

$$\hat{a}_1 \leq \max_{v \in \Omega} \mu(J_{f_1}(v)) = \max_{v \in \Omega} \max(1 - 2v_1, 2 - 2v_2) = 2,$$

$$\hat{a}_2 \leq \max_{v \in \Omega} \mu(J_{f_2}(v)) = \max_{v \in \Omega} \max(3 - 2v_1, -1 - 2v_2) = 3,$$

and

$$\hat{a}'_1 \leq \max_{v \in \Omega} -1 - 2v_1 = -1,$$

$$\hat{a}'_2 \leq \max_{v \in \Omega} 3 - 2v_1 = 3.$$

Moreover, as $\Omega$ is a positively invariant set for (47), if $K \subset \Omega$, then one can obtain the same upper bounds for the constants $\hat{a}'_p$ defined by (40), that is, $\hat{a}'_1 \leq -1$, $\hat{a}'_2 \leq 2$, and $\hat{a}'_3 \leq -1$. The upper bounds for $h(f_{\sigma_1}, K)$ and $h(f_{\sigma_2}, K)$ computed using (28), (41), (45), and (46) for all $K \subset \mathbb{R}^n_{\geq 0}$, as well as (39) for all $K \subset \Omega$, are summarized in Table 2 below. For the case with $\sigma_2$, the upper bounds (28) and (41) are computed along the lines of computing (28) in Example 3.6; the upper bound in (39) is computed along the lines of computing $h(D_{\sigma_2})$ in [37, Example 3 and Appendix E]. The numerical results are consistent with the discussions on the relations between the upper bounds (28), (39), (41), (45), and (46) in Remark 4.4.4.

![Table 2: Upper bounds for the entropy of the switched diagonal system (47).](image)

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 $\omega$-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 grant 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.
REFERENCES


dons-and-optimal-control-theory


[12]ounces-and-optimal-control-theory