On topological entropy of interconnected nonlinear systems

Daniel Liberzon, *Fellow, IEEE*

***Abstract*— We study topological entropy of a nonlinear system represented as an interconnection of smaller subsystems. Under suitable assumptions on the Jacobian matrices characterizing the interconnection, we obtain an explicit upper bound on the entropy of the overall system, and show that it can be related to upper bounds on the entropies of the subsystems. We also analyze in detail the special case of a cascade connection of two subsystems, establishing an upper bound on the entropy which is more tightly linked to individual entropy bounds for the subsystems.**

***Index Terms*— Information theory and control, quantized systems, stability of nonlinear systems**

I. Introduction

THE object of study in this letter is entropy—more specifically, topological entropy—of nonlinear continuous-time systems. Entropy concepts are classical and fundamental in dynamical system theory, as well documented, e.g., in [8], [10]. In the recent systems and control literature, suitable entropy notions have been introduced and studied in the context of controlled invariance [6], [17], stabilization [5], state estimation [13], [15], [19], and model detection [13]. The interest in entropy among control engineers stems at least in part from the fact that entropy characterizes data rates necessary for digital implementations of control and estimation algorithms, as discussed in the above references.

Another paradigm of paramount importance in system theory is that of representing a large system as an interconnection of smaller and simpler subsystems, and arguing about the behavior of the overall system on the basis of analyzing its individual components. Small-gain theorems for establishing stability of both linear and nonlinear systems (see [7], [9]) can be mentioned as standard examples of this line of reasoning.

Manuscript received January 4, 2021; revised March 9, 2021; accepted March 12, 2021. This work was supported in part by the NSF grant CMMI-1662708 and the AFOSR grant FA9550-17-1-0236. Recommended by Senior Editor C. Prieur.

The author is with the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA (e-mail: liberzon@uiuc.edu).

Recently there has been some work on applying entropy concepts in the context of interconnected systems (in discrete time): [11] studies invariance entropy of a network of control systems and relates it to entropies of the subsystems and to associated control data rates; [22] explores invariance feedback entropy for a network of uncertain systems; [16] analyzes observability rates of a networked system and relates them to the topological entropy of the entire system and to certain auxiliary quantities (storage functions and supply rates) defined for the linearizations of the subsystems. However, much still remains to be understood about the basic question considered here, namely: given an interconnection of deterministic nonlinear systems, under what conditions can we derive an explicit bound on its topological entropy and relate it to bounds on the entropies of the subsystems?

The entropy estimates that we derive involve upper bounds on the matrix measures and induced norms of the Jacobian matrices arising from the interconnection. (We note that entropy bounds of this form are very different from those obtained in [16].) For the case of a general interconnection, we establish an upper bound on the entropy that involves the largest eigenvalue of the interconnection matrix obtained from these Jacobians, multiplied by the system dimension. We then consider the special case of a cascade connection of two subsystems, and discover a rather simple upper bound on the entropy which involves, this time more transparently, upper bounds on the entropies of the two subsystems in terms of their Jacobians as well as their dimensions.

Section II introduces the class of systems under consideration and develops the necessary background, including basic definitions and facts on topological entropy and a bound on the separation between trajectories of the system which plays a key role in proving our results. The main results—an upper bound for the entropy of a general interconnection, followed by a more specific bound for the case of a cascade connection—are stated and proved in Section III. Section IV concludes this letter.

II. Preliminaries

Consider a system $\dot{x} = f(x)$, $x \in \mathbb{R}^n$ written as a collection of interconnected subsystems

$$\dot{x}_i = f_i(x_1, \dots, x_k), \qquad i = 1, \dots, k \tag{1}$$

with $\dim(x_i) = n_i$ and $n_1 + \cdots + n_k = n$. We assume that the initial state $x(0)$ belongs to a known compact and convex set $K \subset \mathbb{R}^n$. We denote by $|\cdot|$ the ∞-norm in $\mathbb{R}^d$ (i.e., $|x| := \max_{1 \le i \le d} |x_i|$); here d can be either the total system dimension n or the number of subsystems k or one of the subsystem dimensions n_i, $i = 1, \dots, k$. We write $\|\cdot\|$ for the corresponding induced matrix ∞-norm on $\mathbb{R}^{d \times d}$ (or, occasionally, on $\mathbb{C}^{d \times d}$). The *matrix measure* $\mu : \mathbb{R}^{d \times d} \to \mathbb{R}$ (with respect to the infinity norm) is defined by $\mu(A) := \lim_{\varepsilon \searrow 0} \frac{\|I + \varepsilon A\| - 1}{\epsilon}$ (see, e.g., [23]). We assume that f is C^1 and let J denote the Jacobian matrix with blocks

$$J_{ij}(x) := \frac{\partial f_i}{\partial x_j}(x), \quad 1 \le i, j \le k. \tag{2}$$

We assume that the following holds for some known numbers a_{ij}, $1 \le i, j \le k$.

Assumption 1 $\mu(J_{ii}(x)) \le a_{ii}$ for all x and all i, and $\|J_{ij}(x)\| \le a_{ij}$ for all x and all $i \neq j$.

We note that one consequence of Assumption 1 (and the fact that f is C^1) is that all solutions starting in K exist globally in time, as can be seen from Lemma 2 below (see also the discussion in [13]).

Remark 1 Although we assumed the bounds in Assumption 1 to hold globally over $\mathbb{R}^n$ for simplicity, it is clearly sufficient for all our purposes if they hold over the set of all states reachable from the initial set K at some time $t \ge 0$. Moreover, if it so happens that all solutions of (1) starting from K remain in a compact set, then such finite bounds automatically exist; this will be the case in Example 1 below.

A. Entropy background

The notion of topological entropy that we use here is standard; see, e.g., [10], [19] for discrete-time versions and [13] for a continuous-time version which is slightly more general in that it also incorporates an exponential decay rate (which we set to 0 here for simplicity). Let us write $\xi(x, t)$ for the solution of our system (1) from initial state $x(0) = x \in K$ evaluated at time $t \ge 0$. For a given time horizon $T > 0$ and precision $\varepsilon > 0$, we say that a finite set of points $S = \{x_1, \dots, x_N\} \subset K$ is (T, ε, K)*-spanning* if for every initial state $x \in K$ there exists some point $x_i \in S$ such that the corresponding solutions satisfy

$$|\xi(x, t) - \xi(x_i, t)| < \varepsilon \qquad \forall\, t \in [0, T]. \tag{3}$$

Letting $s(T, \varepsilon, K)$ denote the minimal cardinality of such a (T, ε, K)-spanning set, we define the entropy as

$$h(f, K) := \lim_{\varepsilon \searrow 0} \limsup_{T \to \infty} \frac{1}{T} \log s(T, \varepsilon, K)$$

where $\log$ denotes the natural logarithm.

For each $i \in \{1, \dots, k\}$ we can consider the i-th subsystem in (1) disconnected from the other subsystems, i.e.,

$$\dot{x}_i = f_i(0, \dots, 0, x_i, 0, \dots, 0). \tag{4}$$

Its entropy, which we label as $h(f_i, K)$, satisfies the following upper bound which is essentially well-known (cf. [3], [5], [13]).

Lemma 1 *Under Assumption 1 we have* $h(f_i, K) \le n_i a_{ii}^+$, *where* $a_{ii}^+ := \max\{a_{ii}, 0\}$.

Note that, as a special case, we could always ignore the interconnection structure and just take $k = 1$ (and hence $n_1 = n$, $x_1 = x$, $f_1 = f$), in which case Lemma 1 simply gives the upper bound on $h(f, K)$ in terms of the bound on the matrix measure of the overall Jacobian.

B. Separation between trajectories

The following separation bound was derived in [2], building on earlier work such as [3] and [21]. We briefly sketch the main idea of the proof below for completeness.

Lemma 2 *Let* $x(0), z(0) \in K$ *be two arbitrary initial states. Let* $x(\cdot) = \xi(x(0), \cdot)$ *and* $z(\cdot) = \xi(z(0), \cdot)$ *be the corresponding solutions of (1), with individual subsystem components* $x_1(\cdot), \dots, x_k(\cdot)$ *and* $z_1(\cdot), \dots, z_k(\cdot)$, *respectively. Let* A *be the matrix with elements* a_{ij} *from Assumption 1. Then for all* $t \ge 0$ *we have*

$$\begin{pmatrix} |x_1(t) - z_1(t)| \\ \vdots \\ |x_k(t) - z_k(t)| \end{pmatrix} \le e^{At} \begin{pmatrix} |x_1(0) - z_1(0)| \\ \vdots \\ |x_k(0) - z_k(0)| \end{pmatrix} \tag{5}$$

where the inequality holds element-wise.

Proof: (sketch) For $\lambda \in [0, 1]$, recalling that K is a convex set, let $\xi(\lambda, t)$ be the shorthand for $\xi(\lambda x(0) +$

$(1-\lambda)z(0), t)$, the solution of (1) at time t from initial condition $\lambda x(0) + (1-\lambda)z(0)$, so that

$$\frac{\partial}{\partial t}\xi(\lambda, t) = f(\xi(\lambda, t)). \tag{6}$$

Note that $\xi(1,t) = x(t)$ and $\xi(0,t) = z(t)$ are the solutions from initial conditions $x(0)$ and $z(0)$, respectively. We also denote by $\xi_i(\lambda, t)$, $i = 1, \dots, k$ the components corresponding to the individual subsystems. We have

$$x(t) - z(t) = \xi(1,t) - \xi(0,t) = \int_0^1 \frac{\partial \xi}{\partial \lambda}(\lambda, t) d\lambda. \tag{7}$$

The time evolution of $\frac{\partial \xi}{\partial \lambda}(\lambda, t)$ is given by[1]

$$\begin{aligned}\frac{\partial}{\partial t}\frac{\partial \xi}{\partial \lambda}(\lambda, t) &= \frac{\partial}{\partial \lambda}\frac{\partial}{\partial t}\xi(\lambda, t) \underset{(6)}{=} \frac{\partial}{\partial \lambda} f(\xi(\lambda, t)) \\ &= J(\xi(\lambda, t)) \cdot \frac{\partial \xi}{\partial \lambda}(\lambda, t).\end{aligned}$$

By our Assumption 1 and the proof of Proposition 1 in [2], the solution of this LTV system satisfies the bound

$$\begin{pmatrix} |\frac{\partial \xi_1}{\partial \lambda}(\lambda, t)| \\ \vdots \\ |\frac{\partial \xi_k}{\partial \lambda}(\lambda, t)| \end{pmatrix} \le e^{At} \begin{pmatrix} |\frac{\partial \xi_1}{\partial \lambda}(\lambda, 0)| \\ \vdots \\ |\frac{\partial \xi_k}{\partial \lambda}(\lambda, 0)| \end{pmatrix}$$

where the inequality holds element-wise. Using (7) and the fact that $\xi(\lambda, 0) = \lambda x(0) + (1-\lambda)z(0)$, we arrive at (5). □

III. Main results

Our goal is to relate the entropy of the interconnected system (1) to the entropies of its individual subsystems. To this end, since the latter are characterized via Lemma 1, we want to express (actually, upper-bound) the entropy of (1) in terms of the elements of the matrix A, particularly the diagonal ones. The general result stated as Theorem 1 below is loosely in this spirit (see also Remark 2), and afterwards Proposition 1 achieves this goal more precisely for a special case.

As before, let A be the matrix with elements a_{ij} from Assumption 1. Since it is a Metzler matrix ($a_{ij} \ge 0$ for $i \neq j$), its eigenvalue with the largest real part is real (see, e.g., [4, Theorem 10.2]); we denote this eigenvalue by $\lambda_{\max}(A)$.

Theorem 1 *The entropy of the interconnected system* (1) *satisfies*

$$h(f, K) \le n\lambda^+_{\max}(A) \tag{8}$$

where $\lambda^+_{\max}(A) := \max\{\lambda_{\max}(A), 0\}$.

[1] See, e.g., [12, Section 4.2.4] for a more rigorous argument that can be used to arrive at the same conclusion.

Remark 2 It is worth noting that due to the Metzler property of A, the eigenvector corresponding to $\lambda_{\max}(A)$ can be selected to have non-negative components (see, e.g., [4, Theorem 10.2]). If, in addition, we can choose all components of this eigenvector to be strictly positive—which can be done if A is irreducible (see again [4])—then it follows that $\lambda_{\max}(A) \ge a_{ii}$ for all i. Consequently we have $n\lambda^+_{\max}(A) \ge \sum_{i=1}^k n_i a^+_{ii}$, which in view of Lemma 1 implies that the right-hand side of (8) is at least as large as the sum of the entropy bounds of the disconnected individual subsystems (4).

Proof: To prove the theorem, we build a grid which we show serves as a spanning set, and then count its cardinality to upper-bound the entropy. Let R denote the radius (with respect to the infinity norm) of the initial set K, i.e., $R > 0$ is the smallest number such that K is contained in the hypercube $[-R, R]^n$. The value of R affects the grid construction but does not appear in the resulting entropy bound.

Let

$$A = P^{-1}\Lambda P \tag{9}$$

where Λ is a matrix in (complex) Jordan normal form. Let

$$c := \|P^{-1}\|\|P\|. \tag{10}$$

We need the following easy fact (see, e.g., [20, Chapter 4]).

Lemma 3 *For every* $\delta > 0$ *there exists a time* $T_\delta > 0$ *such that* $\|e^{\Lambda t}\| \le e^{(\lambda_{\max}(A)+\delta)t}$ *for all* $t \ge T_\delta$.

We omit the proof of this lemma but note that it follows by observing, first, that $\|e^{\Lambda t}\| = \max_{1\le i\le \ell} \|e^{\Lambda_i t}\|$ where $\Lambda_1, \dots, \Lambda_\ell$ are the Jordan blocks of Λ, and then applying [20, Lemma 4.1].

Continuing with the proof of the theorem, fix an arbitrary $\delta > 0$, and use Lemma 3 to define

$$M_\delta := \max_{t\in[0,T_\delta]} \|e^{\Lambda t}\|. \tag{11}$$

For $\varepsilon > 0$ and $T > 0$, let

$$\theta := \frac{\varepsilon}{c\max\{M_\delta, e^{(\lambda^+_{\max}(A)+\delta)T}\}}. \tag{12}$$

Consider the grid of points in $\mathbb{R}$ given by $\{k\theta : k \in \mathbb{Z}\}$. Using this grid on each scalar component of x, build a product grid on $\mathbb{R}^n$ and denote by G the set of all points in this grid that belong to the hypercube $[-R, R]^n$.

By Lemma 3 and by definition of M_δ, we have $\|e^{\Lambda t}\| \le \max\{M_\delta, e^{(\lambda_{\max}(A)+\delta)t}\}$ for all $t \ge 0$. It follows that

$$\max_{t\in[0,T]} \|e^{\Lambda t}\| \le \max\left\{M_\delta, e^{(\lambda^+_{\max}(A)+\delta)T}\right\}.$$

Combining this with (9) and (10), applying Lemma 2, and recalling that $|\cdot|$ is the ∞-norm, we can easily conclude that G is a (T,ε)-spanning set.

The cardinality of the grid G is

$$\#G = \big(2\lfloor R/\theta\rfloor + 1\big)^n$$

which gives

$$\log(\#G) \le n\log(2R/\theta+1) = n\log\big((2R+\theta)/\theta\big)$$

and hence

$$\frac{1}{T}\log(\#G) \le \frac{1}{T}n\log(1/\theta) + \frac{1}{T}n\log(2R+\theta).$$

Since θ converges to 0 as $T\to\infty$, upon taking the limsup as $T\to\infty$ only the first term will remain:

$$\limsup_{T\to\infty}\frac{1}{T}\log(\#G) \le \limsup_{T\to\infty}\frac{1}{T}n\log(1/\theta).$$

Recall that, by definition of θ,

$$\frac{1}{\theta} = \frac{c\max\{M_\delta, e^{(\lambda^+_{\max}(A)+\delta)T}\}}{\varepsilon}$$

and thus

$$\frac{1}{T}\log(1/\theta) = \frac{1}{T}\log\big(\max\{M_\delta, e^{(\lambda^+_{\max}(A)+\delta)T}\}\big) + \frac{1}{T}\log(c/\varepsilon).$$

It is now convenient to note that for $T \ge (\log M_\delta)/(\lambda^+_{\max}(A)+\delta)$ we have

$$\max\big\{M_\delta, e^{(\lambda^+_{\max}(A)+\delta)T}\big\} = e^{(\lambda^+_{\max}(A)+\delta)T}.$$

This yields

$$\limsup_{T\to\infty}\frac{1}{T}n\log(1/\theta) = n(\lambda^+_{\max}(A)+\delta)$$

which is an upper bound on the entropy. Since $\delta>0$ was arbitrary, we have established the claimed bound (8). □

Remark 3 If the matrix A is already in Jordan normal form, it is not hard to see that we could define a different grid spacing θ_i on each invariant subspace corresponding to a Jordan block Λ_i of dimension n_i with eigenvalue λ_i by using Λ_i instead of A (and taking $c=1$) in (11) and (12), which would result in an improved entropy bound of the form $h(f,K)\le\sum_{i=1}^{\ell} n_i\lambda_i^+$. For a general A, it is not immediately clear how to explore its Jordan normal form to refine the grid construction in the above proof, because a linear coordinate transformation on $\mathbb{R}^k$ that brings A to Jordan form need not correspond to a change of coordinates in $\mathbb{R}^n$ for the original system (1). In Section III-A below we examine a case when A has a special structure which indeed enables us to construct a more efficient (non-square) grid, leading to a tighter entropy bound.

Example 1 To illustrate the entropy bound from Theorem 1, we consider the well-known Lorenz system (which models atmospheric convection and exhibits chaotic behavior for certain parameter values [14]):

$$\begin{aligned}\dot x_1 &= \sigma x_2 - \sigma x_1\\ \dot x_2 &= -x_2 - x_1x_3 + \theta x_1\\ \dot x_3 &= -\beta x_3 + x_1x_2\end{aligned}\qquad(13)$$

where β, σ, θ are positive parameters. This system fits into the form (1) with $k=3$ and $n_1=n_2=n_3=1$. For the initial set K, let us take the ball of radius $r_0>0$ around the origin (with respect to the Euclidean norm). Note that each scalar subsystem disconnected from the others as in (4) is exponentially stable and has zero entropy. It is well known that all solutions of (13) are bounded; more specifically, it follows from the calculations given, e.g., in [1] that all solutions starting in K remain in the ball of radius

$$r := \max\Big\{\sigma+\theta+r_0, \frac{\sigma+\theta}{2}\Big(1+\sqrt{1+\beta\max\{1,\tfrac{1}{\sigma}\}}\Big)\Big\}$$

centered at $(0,0,\sigma+\theta)$. The Jacobian matrix (2) is

$$J = \begin{pmatrix}-\sigma & \sigma & 0\\ \theta-x_3 & -1 & -x_1\\ x_2 & x_1 & -\beta\end{pmatrix}$$

and hence we can take the matrix A, whose elements are the bounds from Assumption 1 but with respect to the r-ball around $(0,0,\sigma+\theta)$ rather than the entire $\mathbb{R}^n$ (see also Remark 1), to be

$$A = \begin{pmatrix}-\sigma & \sigma & 0\\ \sigma+r & -1 & r\\ r & r & -\beta\end{pmatrix}.$$

Following [14], we select the values $\beta=8/3$, $\sigma=10$ and $\theta=28$. Then we can numerically obtain that for $0<r_0<17.382$ we have $r\approx 55.382$, $\lambda_{\max}(A)\approx 62.048$, and Theorem 1 yields $h(f,K)<186.15$. We stress that the purpose of this example is just to demonstrate a simple application of Theorem 1, not to conduct a detailed and accurate analysis of the entropy of the Lorenz system (the reader interested in the latter can consult, e.g., [18]).

A. Cascade connection

We now consider the special case of the cascade system

$$\begin{aligned}\dot x_1 &= f_1(x_1)\\ \dot x_2 &= f_2(x_1,x_2)\end{aligned}\qquad(14)$$

where $\dim(x_1) = n_1$ and $\dim(x_2) = n_2$ as before. Here the matrix A is block-triangular:

$$A = \begin{pmatrix} a_{11} & 0 \\ a_{21} & a_{22} \end{pmatrix}. \tag{15}$$

This additional structure allows us to explicitly relate the entropy of the cascade system to entropy bounds for the individual subsystems (provided by Lemma 1). In the next result, $h(f,K)$ stands for the entropy of the system (14) as defined in Section II-A for the system (1) of which (14) is a special case.

Proposition 1 *The entropy of the cascade system* (14) *satisfies*

$$h(f,K) \le n_1 \max\{a_{11}^+, a_{22}^+\} + n_2 a_{22}^+ \tag{16}$$

with a_{ii}^+ as defined in Lemma 1.

Proof: The bound (5) from Lemma 2 on the separation between two system trajectories specializes, in this case, to

$$|x_1(t) - z_1(t)| \le e^{a_{11}t}|x_1(0) - z_1(0)| \tag{17}$$

and[2]

$$\begin{aligned}|x_2(t) - z_2(t)| &\le e^{a_{22}t}|x_2(0) - z_2(0)| \\ &+ \frac{a_{21}}{a_{11} - a_{22}}(e^{a_{11}t} - e^{a_{22}t})|x_1(0) - z_1(0)|. \end{aligned} \tag{18}$$

(This can be verified by computing the exponential of the matrix (15), or by directly solving the corresponding linear system.)

As before, let $R > 0$ be the smallest number such that the initial set K is contained in the hypercube $[-R,R]^n$. For $\varepsilon > 0$ and $T > 0$, let[3]

$$\theta_1 := \varepsilon \min\left\{\frac{1}{e^{a_{11}^+ T}}, \frac{|a_{11} - a_{22}|}{2a_{21}e^{\max\{a_{11}^+, a_{22}^+\}T}}\right\}.$$

Consider the grid of points in $\mathbb{R}$ given by $\{k\theta_1 : k \in \mathbb{Z}\}$. Using this grid on each scalar component of x_1, build a product grid on $\mathbb{R}^{n_1}$ and denote it by $G_1(\theta_1)$. Similarly, let

$$\theta_2 := \frac{\varepsilon}{2e^{a_{22}^+ T}}.$$

Using the grid $\{k\theta_2 : k \in \mathbb{Z}\}$ on each scalar component of x_2, build a product grid on $\mathbb{R}^{n_2}$ and denote it by $G_2(\theta_2)$. Finally, construct the product grid on $\mathbb{R}^n$ from $G_1(\theta_1)$ and $G_2(\theta_2)$, and call G the set of all points in this grid that belong to the hypercube $[-R,R]^n$.

The fact that G is (T,ε)-spanning follows directly from (17) and (18). Indeed, by construction, for every initial state $(x_1(0), x_2(0)) \in K$ we can pick an initial condition $(z_1(0), z_2(0)) \in G$ such that the corresponding trajectories satisfy, for all $t \in [0,T]$,

$$|x_1(t) - z_1(t)| < e^{a_{11}t}\frac{\varepsilon}{e^{a_{11}^+ T}} \le \varepsilon$$

and

$$\begin{aligned}&|x_2(t) - z_2(t)| < e^{a_{22}t}\frac{\varepsilon}{2e^{a_{22}^+ T}} \\ &+ \frac{a_{21}}{a_{11} - a_{22}}(e^{a_{11}t} - e^{a_{22}t})\frac{\varepsilon|a_{11} - a_{22}|}{2a_{21}e^{\max\{a_{11}^+, a_{22}^+\}T}} \\ &\le \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.\end{aligned}$$

Next, the cardinality of the grid G is

$$\#G = \prod_{i=1}^{2}\big(2\lfloor R/\theta_i\rfloor + 1\big)^{n_i}$$

which gives

$$\begin{aligned}\log(\#G) &\le \sum_{i=1}^{2} n_i \log(2R/\theta_i + 1) \\ &= \sum_{i=1}^{2} n_i \log\big((2R + \theta_i)/\theta_i\big)\end{aligned}$$

and hence

$$\begin{aligned}\frac{1}{T}\log(\#G) &\le \sum_{i=1}^{2}\frac{1}{T}n_i\log(1/\theta_i) \\ &+ \sum_{i=1}^{2}\frac{1}{T}n_i\log(2R + \theta_i).\end{aligned}$$

Since θ_i converges to 0 (or remains constant) as $T \to \infty$, $i = 1,2$, upon taking the limsup as $T \to \infty$ only the first summation will remain:

$$\limsup_{T\to\infty}\frac{1}{T}\log(\#G) \le \limsup_{T\to\infty}\sum_{i=1}^{2}\frac{1}{T}n_i\log(1/\theta_i).$$

For the first term in the summation, we have

$$\frac{1}{\theta_1} = \max\left\{\frac{e^{a_{11}^+ T}}{\varepsilon}, \frac{2a_{21}e^{\max\{a_{11}^+, a_{22}^+\}T}}{|a_{11} - a_{22}|}\right\}$$

and thus

$$\begin{aligned}\frac{1}{T}\log(1/\theta_1) = \max\Big\{&a_{11}^+ - \frac{1}{T}\log\varepsilon, \\ &\max\{a_{11}^+, a_{22}^+\} - \frac{1}{T}\log(2a_{21}/|a_{11} - a_{22}|)\Big\}.\end{aligned}$$

[2] We do not consider separately the case $a_{11} = a_{22}$, as it can be recovered from (18) in the limit as $a_{11} \to a_{22}$ via L'Hôpital's rule and, with minor modifications to the construction below, leads to the same result.

[3] For the case $a_{21} = 0$ we follow the convention that $\min\{r, \infty\} = r$ for $r \in \mathbb{R}$.

We obtain

$$\limsup_{T\to\infty} \frac{1}{T} n_1 \log(1/\theta_1) = n_1 \max\{a_{11}^+, a_{22}^+\}.$$

For the second term in the summation, we have $1/\theta_2 = 2e^{a_{22}^+ T}/\varepsilon$, hence similarly to how the first term in the maximum above was handled, we have

$$\limsup_{T\to\infty} \frac{1}{T} n_2 \log(1/\theta_2) = n_2 a_{22}^+.$$

This proves the bound (16). □

To compare Theorem 1 with Proposition 1, we can consider the case when $n_1 = n_2 = 1$ and A is as in (15) with $a_{11} > a_{22} \geq 0$. Then Theorem 1 gives the bound $2a_{11}$ while Proposition 1 gives the smaller bound $a_{11} + a_{22}$.

Remark 4 Sontag [21] derives an upper bound on the matrix measure of the Jacobian of the overall cascade system in terms of the a_{ij}'s from Assumption 1. This can then also be used to easily obtain an upper bound on the entropy via, e.g., [13, Proposition 2]. However, the resulting entropy bound appears to be slightly more conservative than the one derived above.

IV. Conclusions

This letter was a preliminary investigation of topological entropy for interconnected nonlinear deterministic continuous-time systems. Two upper bounds for the entropy were derived, one for a general interconnection and another, more explicit one for the special case of a cascade connection. These bounds were shown to relate to entropy bounds for the individual subsystems comprising the interconnection. Many avenues for future work remain, such as using more general grids to obtain tighter entropy estimates and relaxing the assumptions on the Jacobian matrices. Obtaining entropy bounds for switched systems is another interesting direction of ongoing work, and some of the results on this topic recently reported in [24] treat switched linear systems in triangular form based on ideas similar to the ones explored here.

Acknowledgment. The author thanks Sayan Mitra, Necmiye Ozay and Guosong Yang for stimulating discussions and pointers to the literature, and the anonymous reviewers for their helpful comments.

References

[1] B. Andrievsky, A. L. Fradkov, and D. Liberzon. Robustness of Pecora-Carroll synchronization under communication constraints. *Systems Control Lett.*, 111:27–33, 2018.

[2] M. Arcak and J. Maidens. Simulation-based reachability analysis for nonlinear systems using componentwise contraction properties. In M. Lohstroh, P. Derler, and M. Sirjani, editors, *Principles of Modeling*, volume 10760 of *Lecture Notes in Computer Science*, pages 61–76. Springer, 2018.

[3] V. A. Boichenko and G. A. Leonov. Lyapunov's direct method in estimates of topological entropy. *J. Math. Sci.*, 91:3370–3379, 1998.

[4] F. Bullo. *Lectures on Network Systems*. Kindle Direct Publishing, 2020. http://motion.me.ucsb.edu/book-lns.

[5] F. Colonius. Minimal bit rates and entropy for exponential stabilization. *SIAM J. Control Optim.*, 50:2988–3010, 2012.

[6] F. Colonius and C. Kawan. Invariance entropy for control systems. *SIAM J. Control Optim.*, 48:1701–1721, 2009.

[7] C. Desoer and M. Vidyasagar. *Feedback Systems: Input-Output Properties*. Academic Press, New York, 1975.

[8] T. Downarowicz. *Entropy in Dynamical Systems*. Cambridge University Press, 2011.

[9] Z.-P. Jiang, A. R. Teel, and L. Praly. Small-gain theorem for ISS systems and applications. *Math. Control Signals Systems*, 7:95–120, 1994.

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

[11] C. Kawan and J.-C. Delvenne. Network entropy and data rates required for networked control. *IEEE Trans. Control Netw. Syst.*, 3:57–66, 2016.

[12] D. Liberzon. *Calculus of Variations and Optimal Control Theory: A Concise Introduction*. Princeton University Press, 2012.

[13] D. Liberzon and S. Mitra. Entropy and minimal bit rates for state estimation and model detection. *IEEE Trans. Automat. Control*, 63:3330–3344, 2018.

[14] E. N. Lorenz. Deterministic nonperiodic flow. *J. Atmospheric Sciences*, 20:130–141, 1963.

[15] A. S. Matveev and A. Yu. Pogromsky. Observation of nonlinear systems via finite capacity channels: Constructive data rate limits. *Automatica*, 70:217–229, 2016.

[16] A. S. Matveev, A. V. Proskurnikov, A. Yu. Pogromsky, and E. Fridman. Comprehending complexity: data-rate constraints in large-scale networks. *IEEE Trans. Automat. Control*, 64:4252–4259, 2019.

[17] G. N. Nair, R. J. Evans, I. M. Y. Mareels, and W. Moran. Topological feedback entropy and nonlinear stabilization. *IEEE Trans. Automat. Control*, 49:1585–1597, 2004.

[18] A. Yu. Pogromsky and A. S. Matveev. Estimation of topological entropy via the direct Lyapunov method. *Nonlinearity*, 24:1937–1959, 2011.

[19] A. V. Savkin. Analysis and synthesis of networked control systems: Topological entropy, observability, robustness and optimal control. *Automatica*, 42:51–62, 2006.

[20] A. J. Schmidt. Topological entropy bounds for switched linear systems with Lie structure. MS Thesis, Univ. of Illinois at Urbana-Champaign, 2016. https://arxiv.org/abs/1610.02701.

[21] E. D. Sontag. Contractive systems with inputs. In J. Willems, S. Hara, Y. Ohta, and H. Fujioka, editors, *Perspectives in Mathematical System Theory, Control, and Signal Processing*, pages 217–228. Springer, 2010.

[22] M. S. Tomar and M. Zamani. Compositional quantification of invariance feedback entropy for networks of uncertain control systems. *IEEE Control Syst. Lett.*, 4:827–832, 2020.

[23] M. Vidyasagar. *Nonlinear Systems Analysis*. Prentice Hall, New Jersey, 2nd edition, 1993.

[24] G. Yang, A. J. Schmidt, D. Liberzon, and J. P. Hespanha. Topological entropy of switched linear systems: general matrices and matrices with commutation relations. *Math. Control Signals Systems*, pages 411–453, 2020.