On topological entropy of interconnected nonlinear systems

Daniel Liberzon

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.

I. INTRODUCTION

The object of study in this paper is entropy—more specifically, topological entropy—of nonlinear systems. Entropy concepts are classical and fundamental in dynamical system theory, as well documented, e.g., in [8], [6]. In the recent systems and control literature, suitable entropy notions have been introduced and studied in the context of controlled invariance [14], [4], stabilization [3], state estimation [16], [12], [11], and model detection [11]. 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 ([5], [7]) can be mentioned as standard examples of this line of reasoning.

Recently there has been some work on applying entropy concepts in the context of interconnected systems: [9] studies invariance feedback entropy of networks of control systems; [19] extends these results to uncertain systems; [13] analyzes observability rates of networked systems and relates them to topological entropy. However, none of these papers addresses the basic question considered here, namely: given an interconnection of deterministic nonlinear systems, under what conditions can we derive explicit bounds on the topological entropy, and relate them to bounds on the entropy 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. 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 the paper.

II. PRELIMINARIES

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

\[
\dot{x}_i = f_i(x_1, \ldots, x_k), \quad i = 1, \ldots, k
\]  

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 \( \infty \)-norm in \( \mathbb{R}^d \) (i.e., \( |x| := \max_{1 \leq i \leq d} |x_i| \)); here \( d \) can
be either the total system dimension \( n \) or the number of subsystems \( k \) or one of the subsystem dimensions \( n_i, i = 1, \ldots, k \). We write \( \| \cdot \| \) for the corresponding induced matrix norm on \( \mathbb{R}^{d \times d} \). The matrix measure \( \mu : \mathbb{R}^{d \times d} \to \mathbb{R} \) (with respect to the infinity norm) is defined by \( \mu(A) := \lim_{\varepsilon \to 0} \frac{\| A + \varepsilon I \|}{\varepsilon} \) (see, e.g., [20]). We let \( J \) denote the Jacobian matrix with blocks

\[
J_{ij}(x) := \frac{\partial f_i}{\partial x_j}(x), \quad 1 \leq i, j \leq k.
\]

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

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

### A. Entropy background

The notion of topological entropy that we use here is standard; see, e.g., [8], [16] for discrete-time versions and [11] 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 \( x \in K \) evaluated at time \( t \geq 0 \). For a given time horizon \( T > 0 \) and precision \( \varepsilon > 0 \), we say that a finite set of points \( S = \{ x_1, \ldots, x_N \} \subset K \) is \( (T, \varepsilon, K) \)-spanning if for every initial state \( x \in K \) there exists some point \( x_i \in S \) such that the corresponding solutions satisfy

\[
|\xi(x,t) - \xi(x_i,t)| < \varepsilon \quad \forall t \in [0,T]. \tag{2}
\]

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

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

where \( \log \) denotes the natural logarithm.

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

\[
\dot{x}_i = f_i(0, \ldots, 0, x_i, 0, \ldots, 0). \tag{3}
\]

Its entropy, which we label as \( h(f_i,K) \), satisfies the following upper bound which is essentially well-known (cf. [2], [3], [11]).

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

### B. Separation between trajectories

The following separation bound was derived in [1], building on earlier work such as [2] and [18]. 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 \( x_1(\cdot), \ldots, x_k(\cdot) \) and \( z_1(\cdot), \ldots, z_k(\cdot) \), respectively. Let \( A \) be the matrix with elements \( a_{ij} \) from Assumption 1. Then for all \( t \geq 0 \) we have

\[
\begin{pmatrix}
|x_1(t) - z_1(t)| \\
:\ \\
|x_k(t) - z_k(t)|
\end{pmatrix} \leq e^{At} \begin{pmatrix}
|x_1(0) - z_1(0)| \\
:\ \\
|x_k(0) - z_k(0)|
\end{pmatrix} \tag{4}
\]

where the inequality holds element-wise.

**Proof:** (sketch) For \( \lambda \in [0,1] \), 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

\[
\frac{\partial}{\partial \lambda} \xi(\lambda, t) = f(\xi(\lambda, t)). \tag{5}
\]

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 \( \xi_i(\lambda, t), i = 1, \ldots, k \) the components corresponding to the individual subsystems. We have

\[
x(t) - z(t) = \xi(1, t) - \xi(0, t) = \int_0^1 \frac{\partial \xi}{\partial \lambda}(\lambda, t) d\lambda = \frac{\partial \xi}{\partial \lambda}(\lambda', t)
\]

for some \( \lambda' \in [0,1] \). The time evolution of \( \frac{\partial \xi}{\partial \lambda}(\lambda, t) \) is given by\(^1\)

\[
\frac{\partial}{\partial t} \frac{\partial \xi}{\partial \lambda}(\lambda, t) = \frac{\partial}{\partial \lambda} \frac{\partial \xi}{\partial t}(x(t)) = \frac{\partial}{\partial \lambda} f(\xi(\lambda, t))
\]

\[
= J(\xi(\lambda, t)) \cdot \frac{\partial \xi}{\partial \lambda}(\lambda, t).
\]

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

\[
\begin{pmatrix}
|\frac{\partial \xi_1}{\partial \lambda}(\lambda, t)| \\
:\ \\
|\frac{\partial \xi_k}{\partial \lambda}(\lambda, t)|
\end{pmatrix} \leq e^{At} \begin{pmatrix}
|\frac{\partial \xi_1}{\partial \lambda}(\lambda, 0)| \\
:\ \\
|\frac{\partial \xi_k}{\partial \lambda}(\lambda, 0)|
\end{pmatrix}
\]

\(^1\)See, e.g., [10, Section 4.2.4] for a more rigorous argument that can be used to arrive at the same conclusion.
where the inequality holds element-wise. Using (6) and the fact that \( \xi(\lambda, 0) = \lambda x(0) + (1 - \lambda) z(0) \), we arrive at (4). 

\[\theta := \frac{\varepsilon}{c \max\{M_\delta, e^{(\lambda_{\max}(A) + \delta)T}\}}.\]

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 \( M_\delta \), we have
\[\|e^{At}\| \leq \max\{M_\delta, e^{(\lambda_{\max}(A) + \delta)T}\}\]
for all \( t \geq 0 \). It follows that
\[\max_{t \in [0, T]} \|e^{At}\| \leq \max\{M_\delta, e^{(\lambda_{\max}(A) + \delta)T}\}.\]

Combining this with (8) and (9), applying Lemma 2, and recalling that \( \cdot \) is the \( \infty \)-norm, we can easily conclude that \( G \) is a \((T, \varepsilon)\)-spanning set.

The cardinality of the grid \( G \) is
\[\#G = (2[R/\theta] + 1)^n\]
which gives
\[\log(\#G) \leq n\log(2R/\theta + 1) = n\log((2R + \theta)/\theta)\]
and hence
\[\frac{1}{T}\log(\#G) \leq \frac{1}{T}n\log(1/\theta) + \frac{1}{T}n\log(2R + \theta).\]

Since \( \theta \) converges to 0 as \( T \to \infty \), upon taking the limsup as \( T \to \infty \) only the first term will remain:
\[\limsup_{T \to \infty} \frac{1}{T}\log(\#G) \leq \limsup_{T \to \infty} \frac{1}{T}n\log(1/\theta).\]

Recall that, by definition of \( \theta \),
\[\frac{1}{\theta} = c \max\{M_\delta, e^{(\lambda_{\max}(A) + \delta)T}\}\]
and thus
\[\frac{1}{T}\log(1/\theta) = \frac{1}{T}\log\left( \max\{M_\delta, e^{(\lambda_{\max}(A) + \delta)T}\}\right) + \frac{1}{T}\log(e/\varepsilon).\]

It is now convenient to note that for \( T \geq (\log M_\delta)/(\lambda_{\max}(A) + \delta) \) we have
\[\max\{M_\delta, e^{(\lambda_{\max}(A) + \delta)T}\} = 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 (7).
Remark 1 It is worth noting that the Metzler property of $A$ also implies that $\lambda_{\text{max}}(A) \geq a_{ii}$ for all $i$ (because the corresponding eigenvector can be selected to have non-negative components; see again (15)). Thus we have $n\lambda_{\text{max}}^+(A) \geq \sum_{i=1}^{k} n_i a_{ii}^+$, which in view of Lemma 1 implies that the right-hand side of (7) is at least as large as the sum of the entropies of the disconnected individual subsystems (3).

A. Cascade connection

We now consider the special case of the cascade system

$$\begin{align*}
\dot{x}_1 &= f_1(x_1) \\
\dot{x}_2 &= f_2(x_1, x_2)
\end{align*}$$

(10)

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

$$A = \begin{pmatrix} a_{11} & 0 \\ a_{21} & a_{22} \end{pmatrix}. \quad (11)$$

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).

Proposition 1 The entropy of the cascade system (10) satisfies

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

(12)

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

Proof: The bound (4) from Lemma 2 on the separation between two system trajectories specializes, in this case, to

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

(13)

and

$$|x_2(t) - z_2(t)| \leq e^{a_{22}t}|x_2(0) - z_2(0)|$$

$$+ \frac{a_{21}}{a_{11} - a_{22}} (e^{a_{11}t} - e^{a_{22}t})|x_1(0) - z_1(0)|. \quad (14)$$

(This can be verified by computing the exponential of the matrix (11), or by direct analysis of the LTV system that arises in the proof of Lemma 2.)

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

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

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 $(T, \varepsilon)$-spanning follows directly from (13) and (14). 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 $t \in [0, T]$,

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

and

$$|x_2(t) - z_2(t)| < e^{a_{22}t} \frac{\varepsilon}{2e^{a_{22}T}}$$

$$+ \frac{a_{21}}{a_{11} - a_{22}} (e^{a_{11}t} - e^{a_{22}t}) \frac{\varepsilon |a_{11} - a_{22}|}{2a_{21} e^{\max\{a_{11}, a_{22}\}T}} \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$  

Next, the cardinality of the grid $G$ is

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

which gives

$$\log(\#G) \leq \sum_{i=1}^{2} n_i \log(2R/\theta_i + 1)$$

$$= \sum_{i=1}^{2} n_i \log \left( (2R + \theta_i)/\theta_i \right)$$

and hence

$$\frac{1}{T} \log(\#G) \leq \sum_{i=1}^{2} \frac{1}{T} n_i \log(1/\theta_i)$$

$$+ \sum_{i=1}^{2} \frac{1}{T} n_i \log(2R + \theta_i).$$
Since $\theta_i$ 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:

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

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

$$\frac{1}{T} \log(1/\theta_1) = \max \left\{ \frac{1}{T} \log \varepsilon, \max\{a_{11}^+, a_{22}^+\} - \frac{1}{T} \log(2a_{21}/|a_{11} - a_{22}|) \right\}.$$

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 (12).

\textbf{Remark 2} Sontag [18] 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., [11, Proposition 2]. However, the resulting entropy bound appears to be slightly more conservative than the one derived above.

IV. CONCLUSIONS

This paper 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 a 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 [21] treat switched linear systems in triangular form based on ideas similar to the ones explored here.

\textbf{Acknowledgment.} The author thanks Sayan Mitra, Necmiye Ozay and Guosong Yang for helpful discussions and pointers to the literature.

\textbf{REFERENCES}


