

Topological Entropy of Nonlinear Time-Varying Systems

Guosong Yang, *Member, IEEE* and Daniel Liberzon, *Fellow, IEEE*

Abstract—Two general upper bounds on the topological entropy of nonlinear time-varying systems are established: one using the matrix measure of the system Jacobian, the other using the largest real part of the eigenvalues of the Jacobian matrix with off-diagonal entries replaced by their absolute values. A general lower bound is constructed using the trace of the Jacobian matrix. For interconnected systems, an upper bound is first derived by adapting one of the general upper bounds, using the matrix measure of an interconnection matrix function. A new upper bound is then developed using the largest real part of the eigenvalues of this function. This new bound is closely related to the individual upper bounds for subsystems and implies each of the two general upper bounds when the system is viewed as one of two suitable interconnections. These entropy bounds all depend only on upper or lower limits of the Jacobian matrix along trajectories.

Index Terms—Interconnected systems, nonlinear systems, time varying systems, topological entropy.

I. INTRODUCTION

IN systems theory, topological entropy characterizes the rate at which information is generated in deterministic systems, captured by the increase in distinguishable behaviors at finite resolution, or by the complexity growth of a system acting on a set with finite measure. Adler et al. [1] introduced the concept of topological entropy as an extension of Kolmogorov's metric entropy [2], quantifying the exponential growth rate of the minimal number of sets in iterated open-cover refinements. An alternative definition, formulated in terms of the maximal number of separated trajectories over finite time horizons, was proposed by Dinaburg [3] and independently by Bowen [4]. The equivalence between these two formulations was established in [5]. The variational principle, which identifies topological entropy as the supremum of metric entropy over invariant measures, was established through the combined works of Dinaburg [3], [6], Goodman [7], and Goodwyn [8]. For a comprehensive introduction to topological entropy, see, e.g., [9], [10] and the references therein.

This work was supported in part by the National Science Foundation under Grants CMMI-2452534 and CMMI-2106043, and by the Air Force Office of Scientific Research under MURI Grant FA9550-23-1-0337.

Guosong Yang is with the Department of Electrical and Computer Engineering at Rutgers University, Piscataway, NJ 08854 USA (email: guosong.yang@rutgers.edu).

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

Most existing results on topological entropy focus on time-invariant systems, as time-varying dynamics introduce additional complexities that require new analytical methods. Kolyada and Snoha [11] defined topological entropy for discrete-time systems modeled by a sequence of selfmaps on a compact topological space, and established basic properties which strengthen under assumptions such as equicontinuity or uniform convergence. Their definition and results were extended to time-varying maps defined on a sequence of compact metric spaces in [12]. A variational principle based on this extended definition of topological entropy and a time-varying notion of metric entropy was established in [13]. In the continuous-time setting, an upper bound on the topological entropy of a time-varying system on a compact invariant set was derived via the direct Lyapunov method in [14].

Topological entropy notions play a fundamental role in control theory, where feedback depends on the flow of information between sensors and actuators. Nair et al. [15] introduced topological feedback entropy for discrete-time systems, building on the cover-refinement framework of [1]. Their definition extends classical topological entropy to quantify the exponential growth rate of control complexity required to ensure set invariance and stabilization. For continuous-time systems, Colonius and Kawan [16] proposed invariance entropy, which is more closely aligned with the trajectory-counting formulation of [3], [4]. The equivalence between topological feedback entropy and invariance entropy was established in [17]. Suitable entropy notions have also been developed for exponential stabilization [18], state estimation [19]–[21], and model detection [21].

Topological entropy is closely linked to data-rate requirements for control. For continuous-time linear time-invariant systems, the minimal data rate for feedback stabilization is given by the sum of the positive real parts of the eigenvalues of the system matrix [22]; in the discrete-time case, it is the sum of their logarithms [22]–[24]. This quantity coincides with the topological entropy of the open-loop system [4]. While this connection has also been extensively studied for nonlinear time-invariant systems (e.g., [15], [16], [18], [25]), its counterpart for time-varying dynamics remains largely unexplored, except in certain cases such as switched systems, discussed below.

In systems and control, a common paradigm is to decompose a complex system into an interconnection of simpler subsystems and derive system-level properties from those of the components. Entropy-based approaches have also been

applied in this context, particularly for discrete-time systems. The invariance entropy of a network of control systems was related to the entropy of its subsystems and associated data rates in [26]. Observability rates for networked systems were linked to topological entropy and auxiliary constructs such as storage functions and supply rates derived from subsystem linearization in [27]. Invariance feedback entropy for a network of uncertain systems was investigated in [28]. In the continuous-time setting, an upper bound on the topological entropy of interconnected systems was derived in [29], which can be viewed as a prototype for some of the results in the present paper. However, much remains to be understood about when and how explicit bounds on the topological entropy of an interconnected nonlinear time-varying system can be deduced from those of its subsystems.

Topological entropy and data-rate requirements for control have been studied for switched systems, an important class of time-varying systems. Sufficient data rates for feedback stabilization of switched linear systems were established in [30], [31]. Similar data-rate conditions were derived in [32] by extending a notion of estimation entropy from [21]. Formulae and bounds on the topological entropy of continuous-time switched linear systems were derived in [33]–[35], with their relationship to stability conditions analyzed in [35], [36]. In an extension to switched nonlinear and interconnected systems [37], bounds on entropy were constructed using upper or lower limits of the matrix measure or the trace of the system Jacobian. These quantities also appear in the bounds developed in the present paper. For discrete-time switched linear systems, the topological entropy under worst-case switching sequences was studied using the joint spectral radius in [38], and the connection between estimation entropy, Lyapunov exponents, and average data rate was examined under suitable regularity conditions in [39].

The main objective of this paper is to establish upper and lower bounds on the topological entropy of general and interconnected nonlinear time-varying systems. We introduce the definition of topological entropy in Section II and examine its dependence on the initial time and the set of initial states. We show that changing the initial time does not affect the entropy if the system is reinitialized from the corresponding reachable set, but it may differ if initialized from the original initial set. In Section III, we present two general upper bounds on the entropy of nonlinear time-varying systems. One upper bound is constructed using the matrix measure of the system Jacobian, while the other is based on the largest real part of the eigenvalues of the Jacobian matrix with off-diagonal entries replaced by their absolute values. A general lower bound is also constructed using the trace of the Jacobian matrix. Moreover, we specialize these results to linear time-varying systems and provide examples illustrating that neither upper bound is uniformly tighter than the other.

In Section IV, we consider the interconnected case and first derive an upper bound on entropy by adapting one of the general upper bounds from Section III, using the matrix measure of an interconnection matrix function. We then develop a new upper bound based on the largest real part of the eigenvalues of this function. Notably, this new bound is closely related

to the individual upper bounds for subsystems. Moreover, it implies each of the two general upper bounds from Section III when the system is viewed as an interconnection composed of either a single subsystem or scalar subsystems. These entropy bounds all depend only on the values of the Jacobian matrix over the ω -limit set or its convex variations.

In Section V, we provide additional upper bounds on entropy for both general and interconnected systems, along with a remark on constructing analytically more tractable but potentially less tight bounds. The proofs of the main results are presented in Section VI, following preliminary lemmas including a key componentwise bound on the separation between trajectories for interconnected systems. Section VII concludes the paper with a brief summary and remarks on future research directions.

Notations: For vectors $v_1 \in \mathbb{R}^{n_1}, \dots, v_m \in \mathbb{R}^{n_m}$, let $(v_1, \dots, v_m) \in \mathbb{R}^{n_1 + \dots + n_m}$ denote their concatenation. Let $\text{tr}(A)$ denote the trace of a matrix $A \in \mathbb{R}^{n \times n}$. For a Metzler matrix M (i.e., a matrix with nonnegative off-diagonal entries), let $\lambda_{\max}(M)$ denote its eigenvalue with the largest real part; then $\lambda_{\max}(M) \in \mathbb{R}$ (see, e.g., [40, Th. 9.4, p. 129]). Let $\#E$ denote the cardinality of a finite set E . For a set $S \subset \mathbb{R}^n$, let $\text{vol}(S)$ and $\text{co}(S)$ denote its volume (Lebesgue measure) and convex hull, respectively. For a vector $v = (v_1, \dots, v_n) \in \mathbb{R}^n$, let $|v|_{\infty} = \max_{1 \leq i \leq n} |v_i|$ denote its ∞ -norm. For a matrix $A = [a_{ij}] \in \mathbb{R}^{n \times n}$, let $\|A\|_{\infty} = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|$ denote its induced ∞ -norm. By default, all logarithms are natural logarithms (to avoid an extra multiplicative factor $\ln 2$ in entropy computations).

II. PRELIMINARIES

Consider a continuous-time *nonlinear time-varying* dynamical system

$$\dot{x} = f(t, x), \quad t \geq t_0, \quad (1)$$

where $x \in \mathbb{R}^n$ is the state and $t_0 \in \mathbb{R}$ is the initial time. Assume that $f(t, x)$ is piecewise continuous in t , continuously differentiable in x , and that (1) is forward complete. Let $\xi(t, t_0, x)$ denote the solution of (1) at time t with initial state x at time t_0 . Under these assumptions, $\xi(t, t_0, x)$ is unique, continuously differentiable in x , absolutely continuous in t , and satisfies (1) away from discontinuities of $f(t, x)$ in t . Moreover, its Jacobian matrix $J_x \xi(t, t_0, x)$ is absolutely continuous and piecewise continuously differentiable in t (see, e.g., [41, Th. 7, p. 24]). For brevity, we write

$$J_x f(t, v) := J_x f(t, x)|_{x=v}, \quad J_x \xi(t, t_0, v) := J_x \xi(t, t_0, x)|_{x=v}$$

for the Jacobian matrices of $f(t, x)$ and $\xi(t, t_0, x)$ with respect to x evaluated at $x = v$, respectively.

A. Entropy definition

We define the topological entropy of the nonlinear time-varying system (1) with initial states drawn from a compact set $K \subset \mathbb{R}^n$ with nonempty interior, referred to as the *initial set*. Let $|\cdot|$ be a norm on \mathbb{R}^n , and $\|\cdot\|$ the corresponding induced norm on $\mathbb{R}^{n \times n}$. Given a time horizon $T \geq 0$ and a

radius $\varepsilon > 0$, define the following open ball in \mathbb{R}^n centered at x :

$$B_{f,t_0}(x, \varepsilon, T) := \left\{ \bar{x} \in \mathbb{R}^n : \max_{t \in [t_0, t_0+T]} |\xi(t, t_0, \bar{x}) - \xi(t, t_0, x)| < \varepsilon \right\}. \quad (2)$$

A subset $E \subset K$ is said to be (T, ε) -spanning (for the initial set K at initial time t_0) if

$$K \subset \bigcup_{x \in E} B_{f,t_0}(x, \varepsilon, T), \quad (3)$$

that is, for every $\bar{x} \in K$, there exists $x \in E$ such that $|\xi(t, t_0, \bar{x}) - \xi(t, t_0, x)| < \varepsilon$ for all $t \in [t_0, t_0 + T]$. Let $S(f, t_0, \varepsilon, T, K) \geq 1$ denote the minimal cardinality of a (T, ε) -spanning set; equivalently, it is the cardinality of a minimal (T, ε) -spanning set. This function is nondecreasing in T and nonincreasing in ε . The *topological entropy* of the system (1) with initial set K at initial time t_0 is defined in terms of the exponential growth rate of $S(f, t_0, \varepsilon, T, K)$:

$$h(f, t_0, K) := \lim_{\varepsilon \searrow 0} \limsup_{T \rightarrow \infty} \frac{1}{T} \log S(f, t_0, \varepsilon, T, K). \quad (4)$$

For brevity, we occasionally refer to $h(f, t_0, K)$ simply as the (topological) entropy of (1).

Remark 1. By the equivalence of norms on finite-dimensional vector spaces, the value of $h(f, t_0, K)$ is independent of the choice of norm $|\cdot|$ on \mathbb{R}^n . In particular, it is invariant under a change of basis. More generally, topological entropy can be defined on a metric space (X, d) rather than the normed space $(\mathbb{R}^n, |\cdot|)$, in which case its value may depend on the chosen metric. However, if the initial set is contained in a compact positively invariant set, the topological entropy is a topological invariant. See [9, Prop. 3.1.2, p. 109] [12, p. 165], and [16, p. 1703] for further discussion.

An equivalent definition can be formulated using the maximal cardinality of a (T, ε) -separated set, which is a subset $E \subset K$ such that for all distinct $x, \bar{x} \in E$, there exists $t \in [t_0, t_0 + T]$ with $|\xi(t, t_0, \bar{x}) - \xi(t, t_0, x)| \geq \varepsilon$. This alternative formulation is standard but omitted here, as it is not used in this paper. See [37] for a version of this definition applied to switched nonlinear systems.

In the special case where (1) is time-invariant, we adopt the convention $t_0 = 0$, that is, we consider the *nonlinear time-invariant* system

$$\dot{x} = f(x), \quad t \geq 0 \quad (5)$$

with a continuously differentiable function $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$. For brevity, when discussing time-invariant systems, we omit t_0 from the entropy-related notations introduced above.

B. Initial set and initial time

We examine how the topological entropy $h(f, t_0, K)$ of the nonlinear time-varying system (1) depends on the choice of initial set K and initial time t_0 . We begin with a lemma characterizing entropy over a finite cover of K . The proof is analogous to that of [9, Prop. 3.1.7, p. 111] and is omitted.

Lemma 1. Let $K_1, \dots, K_k \subset K$ form a finite cover of the initial set K , that is, $K = \bigcup_{1 \leq l \leq k} K_l$, where each K_l is a compact set with nonempty interior. The topological entropy of the system (1) satisfies

$$h(f, t_0, K) = \max_{1 \leq l \leq k} h(f, t_0, K_l).$$

Let $\xi(t, t_0, K) := \{\xi(t, t_0, x) : x \in K\}$ denote the *reachable set* at time t of solutions starting from the set K at time t_0 . By definition, $\xi(t_0, t_0, K) = K$. The dependence of the entropy $h(f, t_0, K)$ on the initial time t_0 raises the following question: for a new initial time t_1 , should K be interpreted as the reachable set at t_0 or at t_1 ? In the former case, the entropy remains unchanged, as stated in the following lemma. The proof is provided in Appendix I-A.

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

$$h(f, t_0, K) = h(f, t_1, \xi(t_1, t_0, K)) \quad \forall t_1 \geq t_0. \quad (6)$$

Similar time-varying results for discrete-time systems with equicontinuous functions and for continuous-time systems on compact invariant sets appeared in [13, Prop. 2.1] and [14, Lemma 2.8], respectively. Both results build on the definition and properties of topological entropy developed in [11], [12].

However, the entropy of (1) may differ for different initial sets or for the same initial set at different initial times, as illustrated in the following example. The first case is demonstrated for the nonlinear time-invariant system (5), which is a special case of (1).

Example 1. Consider the function $g : \mathbb{R} \rightarrow \mathbb{R}$ defined by

$$g(x) := \begin{cases} 0, & x < -1, \\ \sqrt{3} - \sqrt{3 - (x+1)^2}, & -1 \leq x \leq 1/2, \\ \sqrt{3}x, & x > 1/2. \end{cases}$$

- 1) (Different initial sets.) Consider the nonlinear time-invariant system (5) with $f(x) = g(x)$. For an initial set $K_1 \subset (-\infty, -1)$, we have $h(g, K_1) = 0$. For an initial set $K_2 \subset (1, \infty)$, we have $h(g, K_2) = \sqrt{3}$.
- 2) (The same initial set at different initial times.) Consider the nonlinear time-varying system (1) with

$$f(t, x) := \begin{cases} g(x+4), & t < 0, \\ g(x), & t \geq 0, \end{cases}$$

and fix the initial set $K := [-3, -2]$. For an initial time $t_0 \geq 0$, we have $h(f, t_0, K) = h(g, K) = 0$. For an initial time $t_1 \leq -1$, the solution satisfies $\xi(t, t_1, x) = e^{\sqrt{3}(t-t_1)}(x+4) - 4$ for $t \in [t_1, 0]$ and $x \in K$. Hence, the reachable set at $t = 0$ is the interval $\xi(0, t_1, K) = [e^{-\sqrt{3}t_1} - 4, 2e^{-\sqrt{3}t_1} - 4] \subset (1, \infty)$. By (6), we have $h(f, t_1, K) = h(f, 0, \xi(0, t_1, K)) = h(g, \xi(0, t_1, K)) = \sqrt{3}$.

Consider the special case where (1) is a *linear time-varying (LTV)* system

$$\dot{x} = A(t)x, \quad t \geq t_0 \quad (7)$$

with a piecewise-continuous matrix-valued function $A : \mathbb{R} \rightarrow \mathbb{R}^{n \times n}$. The following lemma extends [35, Prop. 2] from switched linear systems to general LTV systems.

Lemma 3. *The topological entropy of the LTV system (7) is independent of the choices of initial set and initial time.*

Proof. Independence from the initial set follows from arguments analogous to those in the proof of [35, Prop. 2]. Independence from the initial time then follows from (6). \square

In light of Lemma 3, we omit the initial set K and initial time t_0 and let $h(A)$ denote the entropy of the LTV system (7).

III. ENTROPY OF NONLINEAR TIME-VARYING SYSTEMS

In this section, we present upper and lower bounds on the topological entropy of the nonlinear time-varying system (1). We begin by recalling the notion of matrix measure, which will be used to construct upper bounds.

Given an induced norm $\|\cdot\|$ on $\mathbb{R}^{n \times n}$, the *matrix measure* (or *logarithmic norm*) of a matrix $A \in \mathbb{R}^{n \times n}$, denoted by $\mu(A)$, is the right-sided directional derivative of the norm at the identity matrix I in the direction of A , and is defined by

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

As shown in [42, Th. 2.2.16, p. 22], the matrix measure satisfies

$$-\mu(-A) \leq \operatorname{Re}(\lambda) \leq \mu(A) \leq \|A\| \quad \forall A \in \mathbb{R}^{n \times n}, \quad (8)$$

for any eigenvalue λ of A , where $\operatorname{Re}(\lambda)$ denotes its real part. Moreover,

$$\mu(A + B) \leq \mu(A) + \mu(B), \quad \mu(cA) = c\mu(A)$$

for all $A, B \in \mathbb{R}^{n \times n}$ and $c \geq 0$. In particular, $\mu(\cdot)$ is a convex function. For standard induced norms, explicit expressions for $\mu(A)$ are available. For example, for the induced ∞ -norm, the matrix measure of $A = [a_{ij}] \in \mathbb{R}^{n \times n}$ is

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

See [42, Sec. 2.2.2] and [43, Fact 11.15.7, p. 690] for further properties of the matrix measure.

The following theorem provides upper and lower bounds on the entropy of the system (1). The proof, which partly builds on results for interconnected systems in the next section, is given in Section VI-D.

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

$$h(f, t_0, K) \leq n \max\{\hat{\mu}, 0\} \quad (9)$$

with

$$\hat{\mu} := \limsup_{t \rightarrow \infty} \max_{v \in \operatorname{co}(\xi(t, t_0, K))} \mu(J_x f(t, v)), \quad (10)$$

and lower bounded by

$$h(f, t_0, K) \geq \max\{\tilde{\chi}, 0\} \quad (11)$$

with

$$\tilde{\chi} := \liminf_{t \rightarrow \infty} \min_{v \in \xi(t, t_0, K)} \operatorname{tr}(J_x f(t, v)). \quad (12)$$

The upper and lower limits in (10) and (12) have useful implications for the bounds on entropy. The upper bound in (9) depends only on the matrix measure of the Jacobian $J_x f(t, x)$ over the convex hull of the ω -limit set of solutions starting from the initial set K . Similarly, the lower bound in (11) depends only on the trace of $J_x f(t, x)$ over the ω -limit set. These formulations allow flexibility in deriving tighter or analytically more tractable bounds, depending on the complexity of characterizing the ω -limit set or its convex hull. See Remark 2 in Section V for further discussion.

As a preview of results for interconnected systems, we provide an alternative upper bound on the entropy of (1) that involves neither an induced norm $\|\cdot\|$ nor a matrix measure $\mu(\cdot)$. The proof is given in Section VI-D. This bound is obtained by viewing (1) as n interconnected scalar subsystems. For each $i \in \{1, \dots, n\}$, let x_i and $f_i(t, x)$ denote the i -th scalar components of the state x and the function $f(t, x)$ in (1), respectively.

Proposition 1. *Suppose the function $f(t, x)$ in (1) admits a Metzler matrix $\hat{A} = [\hat{a}_{ij}] \in \mathbb{R}^{n \times n}$ such that*

$$\begin{aligned} \hat{a}_{ii} &\geq \limsup_{t \rightarrow \infty} \max_{v \in \operatorname{co}(\xi(t, t_0, K))} \partial_{x_i} f_i(t, v), \\ \hat{a}_{ij} &\geq \limsup_{t \rightarrow \infty} \max_{v \in \operatorname{co}(\xi(t, t_0, K))} |\partial_{x_j} f_i(t, v)| \end{aligned} \quad (13)$$

for all $i, j \in \{1, \dots, n\}$ such that $i \neq j$. Then the topological entropy of the system (1) is upper bounded by

$$h(f, t_0, K) \leq n \max\{\lambda_{\max}(\hat{A}), 0\}. \quad (14)$$

Comparing the upper bounds on entropy in Theorem 1 and Proposition 1, we observe that the bound in (9) holds for all induced norms and the corresponding matrix measures, whereas the bound in (14) does not involve either. Both bounds are useful, as neither is uniformly tighter than the other. On the one hand, by the second inequality in (8), the bound in (14) is tighter when, for example, (1) is a *linear time-invariant (LTI)* system

$$\dot{x} = Ax, \quad t \geq 0 \quad (15)$$

with a constant Metzler matrix $A \in \mathbb{R}^{n \times n}$. On the other hand, the bound in (9) may be tighter due to the entrywise upper limits in (13), as illustrated in Example 2 below. Additional upper bounds on the entropy of (1) are given in Section V.

A. Entropy of LTV systems

We now specialize the bounds on entropy from Theorem 1 and Proposition 1 to the LTV system (7).

Corollary 1. *The topological entropy of the LTV system (7) is upper bounded by*

$$h(A) \leq n \max\left\{\limsup_{t \rightarrow \infty} \mu(A(t)), 0\right\} \quad (16)$$

and lower bounded by

$$h(A) \geq \max\left\{\liminf_{t \rightarrow \infty} \operatorname{tr}(A(t)), 0\right\}.$$

Moreover, suppose the matrix-valued function $A(t) = [a_{ij}(t)] \in \mathbb{R}^{n \times n}$ in (7) admits a Metzler matrix $\hat{A} = [\hat{a}_{ij}] \in \mathbb{R}^{n \times n}$ such that

$$\hat{a}_{ii} \geq \limsup_{t \rightarrow \infty} a_{ii}(t), \quad \hat{a}_{ij} \geq \limsup_{t \rightarrow \infty} |a_{ij}(t)| \quad (17)$$

for all $i, j \in \{1, \dots, n\}$ such that $i \neq j$. Then the topological entropy of (7) is upper bounded by

$$h(A) \leq n \max\{\lambda_{\max}(\hat{A}), 0\}. \quad (18)$$

As discussed following Proposition 1, the upper bounds in (16) and (18) are both useful, since neither is uniformly tighter. The bound in (18) is tighter when the matrix-valued function $A(t)$ in (7) is a constant Metzler matrix, whereas the bound in (16) is tighter in the following example.

Example 2. Let

$$A_1 = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}.$$

Consider the LTV system (7) with $A(t) = A_1 \sin(t) + A_2 \cos(t)$, and let $\mu(\cdot)$ be the matrix measure defined using the induced ∞ -norm $\|\cdot\|_\infty$. Then

$$\mu(A(t)) = \max\{\sin(t) + |\cos(t)|, |\sin(t)| + \cos(t)\},$$

and thus $\limsup_{t \rightarrow \infty} \mu(A(t)) = \sqrt{2}$. Hence, the upper bound in (16), or equivalently (9), yields $h(A) \leq 2\sqrt{2}$. On the other hand, for all matrices \hat{A} satisfying (17), we have $\hat{a}_{11}, \hat{a}_{12}, \hat{a}_{21}, \hat{a}_{22} \geq 1$, and thus

$$\lambda_{\max}(\hat{A}) = (\hat{a}_{11} + \hat{a}_{22})/2 + \sqrt{\hat{a}_{12}\hat{a}_{21} + (\hat{a}_{11} - \hat{a}_{22})^2/4} \geq 2.$$

Hence, the upper bound in (18), or equivalently (14), yields $h(A) \leq 4$. Therefore, the bound in (16) or (9) is tighter than that in (18) or (14) for this example.

IV. ENTROPY OF INTERCONNECTED SYSTEMS

In this section, we consider the case where the nonlinear time-varying system (1) is composed of $m \geq 2$ interconnected subsystems. For each $i \in \{1, \dots, m\}$, let $x_i \in \mathbb{R}^{n_i}$ denote the state of the i -th subsystem, and let $f_i(t, x) \in \mathbb{R}^{n_i}$ denote the corresponding component function of $f(t, x)$ in (1). Then $n_1 + \dots + n_m = n$, and (1) can be written as the *interconnected* nonlinear time-varying system

$$\dot{x}_i = f_i(t, x_1, \dots, x_m), \quad t \geq t_0, \quad i \in \{1, \dots, m\}. \quad (19)$$

The overall state and system function are given by $x = (x_1, \dots, x_m)$ and $f(t, x) = (f_1(t, x), \dots, f_m(t, x))$, respectively. For brevity, we write

$$J_{x_j} f_i(t, v) := J_{x_j} f_i(t, x) \Big|_{x=v}$$

for the Jacobian matrix of $f_i(t, x)$ with respect to x_j evaluated at $x = v$.

We first derive an upper bound on the topological entropy of the interconnected system (19) by adapting the upper bound in Theorem 1 for the general system (1), using only information captured by an interconnection matrix function. Following [44], we assume that the following norms are given:

- 1) for each $i \in \{1, \dots, m\}$, a “local” norm $|\cdot|_i$ on \mathbb{R}^{n_i} , and
- 2) a “network” norm $|\cdot|_N$ on \mathbb{R}^m that is *monotone*: for all nonnegative vectors $v, w \in \mathbb{R}_{\geq 0}^m$,

$$v \geq w \implies |v|_N \geq |w|_N.$$

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

For a vector $v = (v_1, \dots, v_m) \in \mathbb{R}^n$ with $v_i \in \mathbb{R}^{n_i}$, we define the “global” norm $|\cdot|_G$ by

$$|v|_G := (|v_1|_1, \dots, |v_m|_m)|_N. \quad (20)$$

As $|\cdot|_N$ is monotone, one can verify that $|\cdot|_G$ satisfies the triangle inequality and is indeed a norm. Let $\|\cdot\|_i$, $\|\cdot\|_N$, and $\|\cdot\|_G$ denote the corresponding induced norms on $\mathbb{R}^{n_i \times n_i}$, $\mathbb{R}^{m \times m}$, and $\mathbb{R}^{n \times n}$, respectively, and let $\mu_i(\cdot)$, $\mu_N(\cdot)$, and $\mu_G(\cdot)$ denote the corresponding matrix measures. We also define the corresponding mixed induced norms $\|\cdot\|_{ij}$ on $\mathbb{R}^{n_i \times n_j}$ by

$$\|A\|_{ij} := \max_{|v|_j=1} |Av|_i, \quad i, j \in \{1, \dots, m\}.$$

As $|\cdot|_N$ is monotone, the induced norm $\|\cdot\|_N$ satisfies a monotonicity property for nonnegative matrices. Moreover, the induced norm of matrix exponential and the matrix measure $\mu_N(\cdot)$ also satisfy monotonicity properties for Metzler matrices, as stated in the following lemma.

Lemma 4. 1) For all nonnegative matrices $A, B \in \mathbb{R}_{\geq 0}^{m \times m}$, we have

$$A \geq B \implies \|A\|_N \geq \|B\|_N. \quad (21)$$

2) For all Metzler matrices $A, B \in \mathbb{R}^{m \times m}$, we have

$$A \geq B \implies \|e^A\|_N \geq \|e^B\|_N, \quad (22)$$

and

$$A \geq B \implies \mu_N(A) \geq \mu_N(B). \quad (23)$$

Proof. The implications in (21) and (23) are established in [37, Lemma 2]. To prove (22), let A and B be Metzler matrices with $A \geq B$. Then there exists a sufficiently large integer $k_0 \geq 0$ such that, for all integers $k \geq k_0$, we have $I + A/k \geq I + B/k \geq 0$ and thus $(I + A/k)^k \geq (I + B/k)^k \geq 0$. Therefore, by the definition $e^A := \lim_{k \rightarrow \infty} (I + A/k)^k$, we have $e^A \geq e^B \geq 0$, and (22) follows from (21). \square

Consider the *interconnection matrix function* $A_N(t, x) = [a_{ij}(t, x)] \in \mathbb{R}^{m \times m}$ defined by

$$a_{ii}(t, x) := \mu_i(J_{x_i} f_i(t, x)), \quad a_{ij}(t, x) := \|J_{x_j} f_i(t, x)\|_{ij} \quad (24)$$

for $i, j \in \{1, \dots, m\}$ such that $i \neq j$. By adapting the upper bound in (9) on the entropy of the general system (1) to the interconnected system (19), we obtain the following result.

Corollary 2. The topological entropy of the interconnected system (19) is upper bounded by

$$h(f, t_0, K) \leq n \max\{\hat{\mu}_N, 0\} \quad (25)$$

¹Inequalities between vectors or matrices are interpreted entrywise (e.g., $A \geq 0$ means that A is a nonnegative matrix).

with

$$\hat{\mu}_N := \limsup_{t \rightarrow \infty} \max_{v \in \text{co}(\xi(t, t_0, K))} \mu_N(A_N(t, v)). \quad (26)$$

Proof. The upper bound in (25) follows from the general upper bound in (9), as the constant $\hat{\mu}$ defined in (10) using the “global” matrix measure $\mu_G(\cdot)$ satisfies $\hat{\mu} \leq \hat{\mu}_N$, which can be shown using Lemma 5 below. \square

Lemma 5 ([44, Th. 2]). *For a block matrix $A = [A_{ij}] \in \mathbb{R}^{n \times n}$ with $A_{ij} \in \mathbb{R}^{n_i \times n_j}$ for $i, j \in \{1, \dots, m\}$, suppose there exists a matrix $A_N = [a_{ij}] \in \mathbb{R}^{m \times m}$ such that*

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

Then

$$\mu_G(A) \leq \mu_N(A_N).$$

We now present a new upper bound on the entropy of the interconnected time-varying system (19), obtained by extending a result for interconnected time-invariant systems from [29]. The proof is provided in Section VI-C, following some preliminary lemmas in Sections VI-A and VI-B.

Theorem 2. *Consider the interconnection matrix function $A_N(t, x) = [a_{ij}(t, x)] \in \mathbb{R}^{m \times m}$ defined in (24). Suppose there exists a Metzler matrix $\hat{A}_N = [\hat{a}_{ij}] \in \mathbb{R}^{m \times m}$ such that*

$$\hat{a}_{ij} \geq \limsup_{t \rightarrow \infty} \max_{v \in \text{co}(\xi(t, t_0, K))} a_{ij}(t, v) \quad \forall i, j \in \{1, \dots, m\}. \quad (27)$$

Then the topological entropy of the interconnected system (19) is upper bounded by

$$h(f, t_0, K) \leq n \max\{\lambda_{\max}(\hat{A}_N), 0\}. \quad (28)$$

The upper bounds on entropy in Corollary 2 and Theorem 2 both depend only on the values of the Jacobian matrix $J_x f(t, x)$ over the convex hull of the ω -limit set of solutions starting from the initial set K , as reflected in the upper limits in (26) and (27). Theorem 2 extends [29, Th. 1] by accommodating time-varying dynamics and by avoiding maximization over all states or even all reachable states.

The bounds in (25) and (28) are both constructed using the interconnection matrix function $A_N(t, x)$ defined in (24). Furthermore, they both hold for all “local” norms and the corresponding matrix measures. The bound in (25) also holds for all monotone “network” norms and the corresponding matrix measures, whereas the bound in (28) does not involve either. Both bounds are useful, as neither is uniformly tighter. On the one hand, by the second inequality in (8), the bound in (28) is tighter when, for example, the interconnection matrix function $A_N(t, x)$ defined in (24) is constant. On the other hand, the bound in (25) may be tighter due to the entrywise upper limits in (27). This is illustrated by Example 2 in Section III-A, where the LTV system can be viewed as an interconnected system (19) with $m = 2$ and $n_1 = n_2 = 1$. Additional upper bounds on the entropy of (19) are presented in Section V.

By comparing (27) and the definition of interconnection matrix function (24) to (10), we observe that each diagonal entry \hat{a}_{ii} of \hat{A}_N in the overall bound in (28) coincides with $\hat{\mu}$ in the bound in (9) for the i -th subsystem. As a result,

for systems with a cascade (i.e., triangular) interconnection, the overall bound in (28) reduces to the maximum among the applications of the bound in (9) to each individual subsystem, normalized by the corresponding dimensions.

V. ADDITIONAL UPPER BOUNDS ON ENTROPY

In this section, we present alternative upper bounds on the topological entropy of general and interconnected nonlinear time-varying systems.

We begin by providing an upper bound on the entropy of the general system (1), followed by one for the interconnected system (19). Sketches of the proofs are given in Appendix II.

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

$$h(f, t_0, K) \leq n \max\{\hat{\mu}^*, 0\} \quad (29)$$

with

$$\hat{\mu}^* := \inf_{t_1 \geq t_0} \limsup_{t \rightarrow \infty} \max_{v \in \text{co}(\xi(t_1, t_0, K))} \mu(J_x f(t, \xi(t, t_1, v))). \quad (30)$$

The upper bound in (29) depends only on the values of the Jacobian matrix $J_x f(t, x)$ over the ω -limit set of solutions starting from the convex hull of the reachable set $\xi(t_1, t_0, K)$, for an arbitrary $t_1 \geq t_0$, as reflected in the infimum and upper limit in (30). As a result, it is tighter than the upper bound in (9) if there exists $t_1 \geq t_0$ such that $\xi(t_1, t_0, K)$ is convex. Moreover, (29) holds for all induced norms and the corresponding matrix measures.

Proposition 3. *Consider the interconnection matrix function $A_N(t, x) = [a_{ij}(t, x)] \in \mathbb{R}^{m \times m}$ defined in (24). Suppose there exists a Metzler matrix $\bar{A}_N^* = [\bar{a}_{ij}^*] \in \mathbb{R}^{m \times m}$ such that*

$$\bar{a}_{ij}^* \geq \inf_{t_1 \geq t_0} \sup_{t \geq t_1} \max_{v \in \text{co}(\xi(t_1, t_0, K))} a_{ij}(t, \xi(t, t_1, v)) \quad (31)$$

for all $i, j \in \{1, \dots, m\}$. Then the topological entropy of the interconnected system (19) is upper bounded by

$$h(f, t_0, K) \leq n \max\{\lambda_{\max}(\bar{A}_N^*), 0\}. \quad (32)$$

The upper bound in (32) depends only on the values of the Jacobian matrix $J_x f(t, x)$ along solutions starting from the convex hull of the reachable set $\xi(t_1, t_0, K)$, for an arbitrary $t_1 \geq t_0$, as reflected in the infimum and supremum in (31). The supremum over $t \geq t_1$, rather than the upper limit as $t \rightarrow \infty$, is used in (31) due to differences in the application of variational arguments and the use of the maximum over $[t_0, t]$ in Lemma 11 (see Appendix II). Moreover, (32) holds for all “local” norms and the corresponding matrix measures, and does not involve the “network” norm or the corresponding matrix measure.

We obtain another upper bound on the entropy of the general system (1) by viewing it as an interconnected system (19) composed of n scalar subsystems and applying Proposition 3, similarly to the derivation of Proposition 1 from Theorem 2.

Corollary 3. Suppose the function $f(t, x)$ in (1) admits a Metzler matrix $\bar{A}^* = [\bar{a}_{ij}^*] \in \mathbb{R}^{n \times n}$ such that

$$\begin{aligned} \bar{a}_{ii}^* &\geq \inf_{t_1 \geq t_0} \sup_{t \geq t_1} \max_{v \in \text{co}(\xi(t_1, t_0, K))} \partial_{x_i} f_i(t, \xi(t, t_1, v)), \\ \bar{a}_{ij}^* &\geq \inf_{t_1 \geq t_0} \sup_{t \geq t_1} \max_{v \in \text{co}(\xi(t_1, t_0, K))} |\partial_{x_j} f_i(t, \xi(t, t_1, v))| \end{aligned} \quad (33)$$

for all $i, j \in \{1, \dots, n\}$ s.t. $i \neq j$. Then the topological entropy of the system (1) is upper bounded by

$$h(f, t_0, K) \leq n \max\{\lambda_{\max}(\bar{A}^*), 0\}. \quad (34)$$

We obtain another upper bound on the entropy of the interconnected system (19) by applying Proposition 2 together with Lemma 5, similarly to the derivation of Corollary 2 from Theorem 1.

Corollary 4. The topological entropy of the interconnected system (19) is upper bounded by

$$h(f, t_0, K) \leq n \max\{\hat{\mu}_N^*, 0\} \quad (35)$$

with

$$\hat{\mu}_N^* := \inf_{t_1 \geq t_0} \limsup_{t \rightarrow \infty} \max_{v \in \text{co}(\xi(t_1, t_0, K))} \mu_N(A_N(t, \xi(t, t_1, v))), \quad (36)$$

where the interconnection matrix function $A_N(t, x)$ is defined in (24).

Remark 2. In many scenarios, one can derive bounds on topological entropy that are analytically more tractable but potentially less tight than those in Theorems 1 and 2 and Propositions 2 and 3. For example, consider a set $S \subset \mathbb{R}^n$ such that one of the following conditions holds:

- 1) $S = \mathbb{R}^n$;
- 2) S is compact and contains the convex hull of the ω -limit set of solutions starting from the initial set K ; or
- 3) S is compact, positively invariant, and contains the convex hull of the reachable set $\xi(t_1, t_0, K)$ for some $t_1 \geq t_0$ (e.g., $\text{co}(K) \subset S$).

In this case, the bounds on entropy can be simplified by bounding relevant quantities over the set S . Suppose there exists a constant $\hat{\mu}^S$ such that $\hat{\mu}^S \geq \mu(J_x f(t, v))$ for all $v \in S$ and all sufficiently large $t \geq t_0$. Then the upper bound in (9) on the entropy of the system (1) remains valid with $\hat{\mu}^S$ replacing $\hat{\mu}$. Similar substitutions apply to all upper and lower bounds on entropy presented in this paper.

VI. PROOFS OF MAIN RESULTS

In this section, we present the proofs of the main results, following some necessary preliminary lemmas.

A. Bounds on the separation between trajectories and the volume of the reachable set

We begin by introducing a componentwise upper bound on the separation between trajectories for the interconnected system (19), which will be used to derive the upper bound on entropy in Theorem 2. For each $i \in \{1, \dots, m\}$, let $\xi_i(t, t_0, x)$ denote the solution of the i -th subsystem of (19) at time t with

initial state x at time t_0 . The overall solution of (19) is then given by $\xi(t, t_0, x) = (\xi_1(t, t_0, x), \dots, \xi_m(t, t_0, x))$.

Lemma 6. Consider the interconnection matrix function $A_N(t, x) = [a_{ij}(t, x)] \in \mathbb{R}^{m \times m}$ defined in (24), and the matrix-valued function $\bar{A}_N(t, t_0, K) = [\bar{a}_{ij}(t, t_0, K)] \in \mathbb{R}^{m \times m}$ defined by

$$\bar{a}_{ij}(t, t_0, K) := \max_{s \in [t_0, t]} \max_{v \in \text{co}(\xi(s, t_0, K))} a_{ij}(s, v) \quad (37)$$

for $i, j \in \{1, \dots, m\}$. For all initial states $x = (x_1, \dots, x_m)$, $\bar{x} = (\bar{x}_1, \dots, \bar{x}_m) \in K$ with $x_i, \bar{x}_i \in \mathbb{R}^{n_i}$ for all i , the solutions of the interconnected system (19) satisfy

$$\begin{aligned} &\begin{bmatrix} |\xi_1(t, t_0, \bar{x}) - \xi_1(t, t_0, x)|_1 \\ \vdots \\ |\xi_m(t, t_0, \bar{x}) - \xi_m(t, t_0, x)|_m \end{bmatrix} \\ &\leq e^{\bar{A}_N(t, t_0, K)(t-t_0)} \begin{bmatrix} |\bar{x}_1 - x_1|_1 \\ \vdots \\ |\bar{x}_m - x_m|_m \end{bmatrix} \quad \forall t \geq t_0. \end{aligned} \quad (38)$$

Lemma 6 extends [45, Prop. 1] by introducing the time-varying matrix-valued function $\bar{A}_N(t, t_0, K)$ in place of a constant matrix. This generalization is essential for deriving the upper bound on entropy in Theorem 2. The proof, provided in Appendix I-B, is inspired by the variational construction in the proof of [45, Prop. 1], which in turn builds on earlier work such as [46, Th. 1] and [47, Th. 4.2].

We next present a lower bound on the volume of the reachable set for the general system (1), which will be used to derive the lower bound on entropy in Theorem 1. The proof is provided in Appendix I-C.

Lemma 7. The reachable set of solutions of the system (1) starting from K satisfies

$$\text{vol}(\xi(t, t_0, K)) \geq e^{\gamma(t, t_0)} \text{vol}(K) \quad \forall t \geq t_0 \quad (39)$$

with

$$\gamma(t, t_0) := \min_{v \in K} \int_{t_0}^t \text{tr}(J_x f(s, \xi(s, t_0, v))) \, ds.$$

B. Universal spanning sets

Given a time horizon $T \geq 0$ and a radius $\varepsilon > 0$, we provide a universal formulation of (T, ε) -spanning sets by extending a notion of grid from [35]. Recall that the initial set K is a compact set with nonempty interior. Then there exists a closed hypercube B of radius $r > 0$ such that $K \subset B$. For a vector $\theta = (\theta_1, \dots, \theta_n) \in \mathbb{R}_{\geq 0}^n$, which may depend on T and ε , define the following grid in K centered at an arbitrary $\hat{x} \in K$:

$$G(\theta) := \{\hat{x} + (k_1 \theta_1, \dots, k_n \theta_n) \in K : k_1, \dots, k_n \in \mathbb{Z}\}. \quad (40)$$

The cardinality of $G(\theta)$ is upper bounded by

$$\#G(\theta) \leq \prod_{i=1}^n \left(\left\lceil \frac{2r}{\theta_i} \right\rceil + 1 \right).$$

For each $x = (x_1, \dots, x_n) \in G(\theta)$, let $R(x)$ denote the open hyperrectangle in \mathbb{R}^n centered at x with side lengths $2\theta_1, \dots, 2\theta_n$, that is,

$$R(x) := \{(\bar{x}_1, \dots, \bar{x}_n) \in \mathbb{R}^n : |\bar{x}_1 - x_1| < \theta_1, \dots, |\bar{x}_n - x_n| < \theta_n\}. \quad (41)$$

Then

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

By comparing the hyperrectangle $R(x)$ to the open ball $B_{f,t_0}(x, \varepsilon, T)$ defined in (2), we obtain the following lemma. The proof is provided in Appendix I-D.

Lemma 8. *If the vector θ is selected so that*

$$(R(x) \cap K) \subset B_{f,t_0}(x, \varepsilon, T) \quad \forall x \in G(\theta), \quad (42)$$

then the grid $G(\theta)$ is a (T, ε) -spanning set. Moreover, if (42) holds for all $T \geq 0$ and $\varepsilon > 0$, and

$$\limsup_{\varepsilon \searrow 0} \limsup_{T \rightarrow \infty} \frac{\log \theta_i}{T} \leq 0 \quad \forall i \in \{1, \dots, n\}, \quad (43)$$

then the topological entropy of the system (1) is upper bounded by

$$h(f, t_0, K) \leq \limsup_{\varepsilon \searrow 0} \limsup_{T \rightarrow \infty} \sum_{i=1}^n \frac{\log(1/\theta_i)}{T}. \quad (44)$$

Lemma 8 extends the results on spanning sets and upper bounds on entropy from [35, Lemma 2] and [37, Lemma 2.3], from switched linear and nonlinear systems, respectively, to general nonlinear time-varying systems. Similar extensions of their results on separated sets and lower bounds on entropy are possible but omitted here, as they are not used in this paper. Note that the condition (43) holds if all θ_i are nonincreasing in T .

C. Proof of Theorem 2

We begin by deriving an intermediate upper bound on entropy.

Lemma 9. *The topological entropy of the interconnected system (19) is upper bounded by*

$$h(f, t_0, K) \leq \limsup_{T \rightarrow \infty} \frac{n}{T} \max_{t \in [t_0, t_0+T]} \log \|e^{\bar{A}_N(t, t_0, K)(t-t_0)}\|_N, \quad (45)$$

where the matrix-valued function $\bar{A}_N(t, t_0, K)$ is defined in (37).

Proof. Consider arbitrary initial states $x = (x_1, \dots, x_m)$, $\bar{x} = (\bar{x}_1, \dots, \bar{x}_m) \in K$ with $x_i, \bar{x}_i \in \mathbb{R}^{n_i}$ for all i . By applying the componentwise separation bound in (38), together with the definition of the “global” norm $|\cdot|_G$ in (20), and using the

monotonicity of the “network” norm $|\cdot|_N$, we obtain

$$\begin{aligned} & |\xi(t, t_0, \bar{x}) - \xi(t, t_0, x)|_G \\ &= \left\| \begin{bmatrix} |\xi_1(t, t_0, \bar{x}) - \xi_1(t, t_0, x)|_1 \\ \vdots \\ |\xi_m(t, t_0, \bar{x}) - \xi_m(t, t_0, x)|_m \end{bmatrix} \right\|_N \\ &\leq \left\| e^{\bar{A}_N(t, t_0, K)(t-t_0)} \begin{bmatrix} |\bar{x}_1 - x_1|_1 \\ \vdots \\ |\bar{x}_m - x_m|_m \end{bmatrix} \right\|_N \\ &\leq \|e^{\bar{A}_N(t, t_0, K)(t-t_0)}\|_N \left\| \begin{bmatrix} |\bar{x}_1 - x_1|_1 \\ \vdots \\ |\bar{x}_m - x_m|_m \end{bmatrix} \right\|_N \end{aligned}$$

for all $t \geq t_0$. By the equivalence of norms on finite-dimensional vector spaces, there exist constants $r_1, \dots, r_m, r_N > 0$ such that

$$|v_i|_i \leq r_i |v_i|_\infty \quad \forall i \in \{1, \dots, m\}, v_i \in \mathbb{R}^{n_i}$$

for the “local” norms $|\cdot|_i$, and

$$|v|_N \leq r_N |v|_\infty \quad \forall v \in \mathbb{R}^m.$$

for the “network” norm $|\cdot|_N$. Consequently,

$$\begin{aligned} & |\xi(t, t_0, \bar{x}) - \xi(t, t_0, x)|_G \\ &\leq \|e^{\bar{A}_N(t, t_0, K)(t-t_0)}\|_N r_N \max_{1 \leq i \leq m} |\bar{x}_i - x_i|_i \\ &\leq \|e^{\bar{A}_N(t, t_0, K)(t-t_0)}\|_N r_N \max_{1 \leq i \leq m} r_i |\bar{x}_i - x_i|_\infty \end{aligned}$$

for all $t \geq t_0$. Therefore, for arbitrary $T \geq 0$ and $\varepsilon > 0$, we have

$$\begin{aligned} & \max_{t \in [t_0, t_0+T]} |\xi(t, t_0, \bar{x}) - \xi(t, t_0, x)|_G \\ &\leq \max_{1 \leq i \leq m} \left(\max_{t \in [t_0, t_0+T]} \|e^{\bar{A}_N(t, t_0, K)(t-t_0)}\|_N \right) r_N r_i |\bar{x}_i - x_i|_\infty. \end{aligned}$$

Consider the grid $G(\theta)$ defined in (40) with the vector $\theta = (\bar{\theta}_1 \mathbf{1}_{n_1}, \dots, \bar{\theta}_m \mathbf{1}_{n_m}) \in \mathbb{R}_{>0}^n$, where

$$\bar{\theta}_i := \varepsilon / \left(r_N r_i \max_{t \in [t_0, t_0+T]} \|e^{\bar{A}_N(t, t_0, K)(t-t_0)}\|_N \right)$$

for $i \in \{1, \dots, m\}$, and $\mathbf{1}_{n_i}$ denotes the vector of ones in \mathbb{R}^{n_i} . By comparing the corresponding hyperrectangles $R(x)$ defined in (41) to the open balls $B_{f,t_0}(x, \varepsilon, T)$ defined in (2) with the “global” norm $|\cdot|_G$, we obtain $(R(x) \cap K) \subset B_{f,t_0}(x, \varepsilon, T)$ for all $x \in G(\theta)$. Then Lemma 8 implies that $G(\theta)$ is a (T, ε) -spanning set. Moreover, the upper bound on entropy in (44) holds as $T \geq 0$ and $\varepsilon > 0$ are arbitrary and all $\bar{\theta}_i$ are nonincreasing in T . Therefore,

$$\begin{aligned} h(f, t_0, K) &\leq \limsup_{\varepsilon \searrow 0} \limsup_{T \rightarrow \infty} \sum_{i=1}^m \frac{n_i \log(1/\bar{\theta}_i)}{T} \\ &= \limsup_{T \rightarrow \infty} \sum_{i=1}^m \frac{n_i}{T} \max_{t \in [t_0, t_0+T]} \log \|e^{\bar{A}_N(t, t_0, K)(t-t_0)}\|_N \\ &\quad + \lim_{\varepsilon \searrow 0} \lim_{T \rightarrow \infty} \frac{C}{T}, \end{aligned}$$

where $C := \sum_{i=1}^m n_i \log(r_N r_i / \varepsilon)$ is independent of T and thus satisfies $C/T \rightarrow 0$ as $T \rightarrow \infty$. As $n_1 + \dots + n_m = n$, we conclude that the entropy $h(f, t_0, K)$ satisfies the upper bound in (45). \square

We now establish the upper bound on entropy in (28).

Proof of Theorem 2. The upper limit in (27) implies that, for each $\delta_1 > 0$, there exists a sufficiently large $t_1 \geq t_0$ such that

$$\sup_{t \geq t_1} \max_{v \in \text{co}(\xi(t, t_0, K))} a_{ij}(t, v) \leq \hat{a}_{ij} + \delta_1 \quad \forall i, j \in \{1, \dots, m\}.$$

As $\xi(t, t_0, K) = \xi(t, t_1, \xi(t_1, t_0, K))$, the matrix-valued function $\bar{A}_N(t, t_0, K)$ defined in (37) satisfies

$$\bar{A}_N(t, t_1, \xi(t_1, t_0, K)) \leq \hat{A}_N + \delta_1 \mathbf{1} \quad \forall t \geq t_1,$$

where $\mathbf{1}$ denotes the matrix of ones in $\mathbb{R}^{m \times m}$.

By Lemma 2, we have $h(f, t_0, K) = h(f, t_1, \xi(t_1, t_0, K))$. Then applying Lemma 9 with the new initial time t_1 yields

$$\begin{aligned} & h(f, t_1, \xi(t_1, t_0, K)) \\ & \leq \limsup_{T \rightarrow \infty} \frac{n}{T} \max_{t \in [t_1, t_1+T]} \log \|e^{\bar{A}_N(t, t_1, \xi(t_1, t_0, K))(t-t_1)}\|_N. \end{aligned}$$

Moreover, as $\bar{A}_N(t, t_1, \xi(t_1, t_0, K))$ and $\hat{A}_N + \delta_1 \mathbf{1}$ are both Metzler matrices, the monotonicity property in (22) implies that

$$\begin{aligned} \|e^{\bar{A}_N(t, t_1, \xi(t_1, t_0, K))(t-t_1)}\|_N & \leq \|e^{(\hat{A}_N + \delta_1 \mathbf{1})(t-t_1)}\|_N \\ & \leq \|e^{\hat{A}_N(t-t_1)}\|_N \|e^{\delta_1 \mathbf{1}(t-t_1)}\|_N \\ & \leq \|e^{\hat{A}_N(t-t_1)}\|_N e^{\delta_1 \|\mathbf{1}\|_N (t-t_1)} \end{aligned}$$

for all $t \geq t_1$. Therefore,

$$\begin{aligned} & h(f, t_0, K) \\ & \leq \limsup_{T \rightarrow \infty} \frac{n}{T} \max_{t \in [t_1, t_1+T]} (\log \|e^{\hat{A}_N(t-t_1)}\|_N + \delta_1 \|\mathbf{1}\|_N (t-t_1)) \\ & \leq \limsup_{T \rightarrow \infty} \frac{n}{T} \max_{t \in [0, T]} \log \|e^{\hat{A}_N t}\|_N + n \delta_1 \|\mathbf{1}\|_N. \end{aligned}$$

As $\delta_1 > 0$ is arbitrary, we have

$$h(f, t_0, K) \leq \limsup_{T \rightarrow \infty} \frac{n}{T} \max_{t \in [0, T]} \log \|e^{\hat{A}_N t}\|_N.$$

According to [43, Fact 11.15.7, p. 690], as \hat{A}_N is a Metzler matrix, its eigenvalue with the largest real part $\lambda_{\max}(\hat{A}_N)$ satisfies

$$\lambda_{\max}(\hat{A}_N) = \lim_{t \rightarrow \infty} \frac{1}{t} \log \|e^{\hat{A}_N t}\|_N.$$

Hence, for each $\delta_2 > 0$, there exists a sufficiently large $t_2 \geq 0$ such that

$$\log \|e^{\hat{A}_N t}\|_N \leq (\lambda_{\max}(\hat{A}_N) + \delta_2) t \quad \forall t \geq t_2.$$

Consequently,

$$\frac{1}{T} \max_{t \in [0, T]} \log \|e^{\hat{A}_N t}\|_N \leq \max \left\{ \lambda_{\max}(\hat{A}_N) + \delta_2, \frac{C_{t_2}}{T} \right\}$$

for all $T \geq 0$, where $C_{t_2} := \max_{t \in [0, t_2]} \log \|e^{\hat{A}_N t}\|_N$ is independent of T and thus satisfies $C_{t_2}/T \rightarrow 0$ as $T \rightarrow \infty$. Therefore,

$$\begin{aligned} h(f, t_0, K) & \leq \limsup_{T \rightarrow \infty} \frac{n}{T} \max_{t \in [0, T]} \log \|e^{\hat{A}_N t}\|_N \\ & \leq n \max \{ \lambda_{\max}(\hat{A}_N) + \delta_2, 0 \}. \end{aligned}$$

As $\delta_2 > 0$ is arbitrary, we conclude that the entropy $h(f, t_0, K)$ satisfies the upper bound in (28). \square

D. Proof of Theorem 1 and Proposition 1

Proof of Theorem 1. First, the general system (1) can be viewed as an interconnected system (19) with a single subsystem. Consequently, the upper bound in (9) on the entropy of (1) follows from the upper bound in (28) on the entropy of (19) by setting $m = 1$ and letting the “local” matrix measure $\mu_1(\cdot)$ be the given matrix measure $\mu(\cdot)$.

Second, we prove the lower bound in (11) using volume-based arguments. For arbitrary $T \geq 0$ and $\varepsilon > 0$, the lower bound in (39) on the volume of the reachable set implies

$$\text{vol}(\xi(t_0 + T, t_0, K)) \geq e^{\gamma(t_0+T, t_0)} \text{vol}(K).$$

Let E be a minimal (T, ε) -spanning set. By the equivalence of norms on finite-dimensional vector spaces, there exists a constant $r_\infty > 0$ such that

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

for the given norm $|\cdot|$. Hence

$$\begin{aligned} \xi(t_0 + T, t_0, K) & \subset \bigcup_{x \in E} \{ \bar{x} \in \mathbb{R}^n : |\bar{x} - \xi(t_0 + T, t_0, x)| < \varepsilon \} \\ & \subset \bigcup_{x \in E} \{ \bar{x} \in \mathbb{R}^n : |\bar{x} - \xi(t_0 + T, t_0, x)|_\infty < r_\infty \varepsilon \}, \end{aligned}$$

and thus

$$\begin{aligned} & \text{vol}(\xi(t_0 + T, t_0, K)) \\ & \leq \sum_{x \in E} \text{vol}(\{ \bar{x} \in \mathbb{R}^n : |\bar{x} - \xi(t_0 + T, t_0, x)|_\infty < r_\infty \varepsilon \}) \\ & = (2r_\infty \varepsilon)^n \#E. \end{aligned}$$

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

$$\begin{aligned} S(f, t_0, \varepsilon, T, K) = \#E & \geq \frac{\text{vol}(\xi(t_0 + T, t_0, K))}{(2r_\infty \varepsilon)^n} \\ & \geq \frac{e^{\gamma(t_0+T, t_0)} \text{vol}(K)}{(2r_\infty \varepsilon)^n}. \end{aligned}$$

Consequently, by the definition of entropy (4), we have

$$\begin{aligned} & h(f, t_0, K) \\ & \geq \liminf_{\varepsilon \searrow 0} \limsup_{T \rightarrow \infty} \frac{\gamma(t_0 + T, t_0) + \log(\text{vol}(K)/(2r_\infty \varepsilon)^n)}{T} \\ & = \limsup_{T \rightarrow \infty} \frac{\gamma(t_0 + T, t_0)}{T} + \lim_{\varepsilon \searrow 0} \lim_{T \rightarrow \infty} \frac{C}{T}, \end{aligned}$$

where $C := \log(\text{vol}(K)/(2r_\infty \varepsilon)^n)$ is independent of T and thus satisfies $C/T \rightarrow 0$ as $T \rightarrow \infty$. As the minimum in $\gamma(t_0 + T, t_0)$ is a superadditive function, we have

$$h(f, t_0, K) \geq \limsup_{T \rightarrow \infty} \frac{1}{T} \int_{t_0}^{t_0+T} \left(\min_{v \in K} \text{tr}(J_x f(s, \xi(s, t_0, v))) \right) ds,$$

The lower limit in (12) implies that, for each $\delta_1 > 0$, there exists a sufficiently large $t_1 \geq t_0$ such that

$$\min_{v \in K} \text{tr}(J_x f(t, \xi(t, t_0, v))) \geq \tilde{\chi} - \delta_1 \quad \forall t \geq t_1.$$

Consequently,

$$\frac{1}{T} \int_{t_0}^{t_0+T} \left(\min_{v \in K} \text{tr}(J_x f(s, \xi(s, t_0, v))) \right) ds \geq \tilde{\chi} - \delta_1 + \frac{C_{t_1}}{T}$$

for all $T \geq t_1 - t_0$, where

$$C_{t_1} := \int_{t_0}^{t_1} \left(\min_{v \in K} \text{tr}(J_x f(s, \xi(s, t_0, v))) - (\tilde{\chi} - \delta_1) \right) ds$$

is independent of T and thus satisfies $C_{t_1}/T \rightarrow 0$ as $T \rightarrow \infty$. Therefore,

$$\begin{aligned} h(f, t_0, K) &\geq \limsup_{T \rightarrow \infty} \frac{1}{T} \int_{t_0}^{t_0+T} \left(\min_{v \in K} \text{tr}(J_x f(s, \xi(s, t_0, v))) \right) ds \\ &\geq \tilde{\chi} - \delta_1. \end{aligned}$$

As $\delta_1 > 0$ is arbitrary and $h(f, t_0, K) \geq 0$, we conclude that the entropy $h(f, t_0, K)$ satisfies the lower bound in (11). \square

Proof of Proposition 1. The general system (1) can be viewed as an interconnected system (19) composed of n scalar subsystems. Consequently, the upper bound in (14) on the entropy of (1) follows from the upper bound in (28) on the entropy of (19) by setting $m = n$ and letting the “local” norms be the absolute value (in which case the “local” matrix measure of a scalar is simply the scalar itself). \square

VII. CONCLUSION

This paper studied the topological entropy of general and interconnected nonlinear time-varying systems. We analyzed the dependence of entropy on the initial set and initial time, and established both upper and lower bounds on entropy. For interconnected systems, we constructed an interconnection matrix function that captures information from the Jacobian matrices of each subsystem and their interactions. We then derived upper bounds on entropy in terms of either the matrix measure of this function or its eigenvalue with the largest real part, maximized over the ω -limit set or its convex variations. For general systems, we established upper bounds on entropy by viewing them as interconnected systems composed of either a single subsystem or scalar subsystems. This yielded two upper bounds on entropy, neither of which is uniformly tighter.

This framework provides a flexible method for constructing bounds on entropy, with several notable features:

- It enables the derivation of tighter or analytically more tractable bounds, depending on the complexity of characterizing the ω -limit set or its convex variations.

- It is modularized into three steps: constructing a bound on the separation between trajectories (Section VI-A), formulating spanning sets (Section VI-B), and deriving a bound on entropy (Section VI-C). Each step can be adapted individually for different applications, as illustrated in part by the alternative bounds in Section V.
- Beyond decomposing a system into scalar subsystems, the method supports arbitrary decompositions, which can potentially be optimized based on the system structure.

Future work will focus on improving the bounds by further analyzing the effect of time-varying dynamics (cf. [37]), extending the framework to systems with special structure such as cascade interconnection, and studying the computational complexity of the proposed bounds.

APPENDIX I PROOFS OF TECHNICAL LEMMAS

A. Proof of Lemma 2

Fix $t_1 \geq t_0$ and let $K_1 := \xi(t_1, t_0, K)$. Consider arbitrary $T \geq t_1 - t_0$ and $\varepsilon > 0$.

First, let E_0 be a minimal (T, ε) -spanning set for K at t_0 . We show that $E_1 := \xi(t_1, t_0, E_0)$ is a $(t_0 + T - t_1, \varepsilon)$ -spanning set for K_1 at t_1 . By definition, for each $\bar{z} \in K_1$, there exist $\bar{x} \in K$ and $x \in E_0$ such that $\bar{z} = \xi(t_1, t_0, \bar{x})$ and $\bar{x} \in B_{f, t_0}(x, \varepsilon, T)$. Then for $z := \xi(t_1, t_0, x) \in E_1$, we have

$$|\xi(t, t_1, \bar{z}) - \xi(t, t_1, z)| = |\xi(t, t_0, \bar{x}) - \xi(t, t_0, x)| < \varepsilon$$

for all $t \in [t_1, t_0 + T]$, that is, $\bar{z} \in B_{f, t_1}(z, \varepsilon, t_0 + T - t_1)$. Therefore, $K_1 \subset \bigcup_{z \in E_1} B_{f, t_1}(z, \varepsilon, t_0 + T - t_1)$. Consequently, the minimal cardinality of such a set satisfies

$$S(f, t_1, \varepsilon, t_0 + T - t_1, K_1) \leq \#E_1 = \#E_0 = S(f, t_0, \varepsilon, T, K).$$

By the definition of entropy (4), we conclude that $h(f, t_1, K_1) \leq h(f, t_0, K)$.

Second, let \bar{E}_1 be a minimal $(t_0 + T - t_1, \varepsilon)$ -spanning set for K_1 at t_1 . Then

$$\#\bar{E}_1 = S(f, t_1, \varepsilon, t_0 + T - t_1, K_1) \leq S(f, t_1, \varepsilon, T, K_1)$$

as $S(f, t_1, \varepsilon, T, K_1)$ is nondecreasing in T . By definition, for each $\bar{x} \in K$, there exists $z \in \bar{E}_1$ such that $\bar{z} := \xi(t_1, t_0, \bar{x}) \in B_{f, t_1}(z, \varepsilon, t_0 + T - t_1)$, that is,

$$|\xi(t, t_0, \bar{x}) - \xi(t, t_1, z)| = |\xi(t, t_1, \bar{z}) - \xi(t, t_1, z)| < \varepsilon$$

for all $t \in [t_1, t_0 + T]$. Hence, for

$$B_0(z) := \left\{ \bar{x} \in K : \max_{t \in [t_1, t_0 + T]} |\xi(t, t_0, \bar{x}) - \xi(t, t_1, z)| < \varepsilon \right\},$$

we have $K \subset \bigcup_{z \in \bar{E}_1} B_0(z)$. For each $z \in \bar{E}_1$, let $E_0(z)$ be a minimal $(t_1 - t_0, 2\varepsilon)$ -spanning set for $B_0(z)$ at t_0 . Then

$$\begin{aligned} \#E_0(z) &= S(f, t_0, 2\varepsilon, t_1 - t_0, B_0(z)) \\ &\leq S(f, t_0, 2\varepsilon, t_1 - t_0, K) \leq S(f, t_0, \varepsilon, t_1 - t_0, K), \end{aligned}$$

where the first inequality holds as $B_0(z) \subset K$, and the second one holds as $S(f, t_0, \varepsilon, t_1 - t_0, K)$ is nonincreasing in ε . We now show that $E_0(z)$ is also a $(T, 2\varepsilon)$ -spanning set for $B_0(z)$

at t_0 , using the triangle inequality. By definition, for each $\bar{x} \in B_0(z)$, there exists $x \in E_0(z) \subset B_0(z)$ such that

$$|\xi(t, t_0, \bar{x}) - \xi(t, t_0, x)| < 2\varepsilon \quad \forall t \in [t_0, t_1].$$

Moreover, for all $t \in [t_1, t_0 + T]$, as $x, \bar{x} \in B_0(z)$, we also have

$$\begin{aligned} & |\xi(t, t_0, \bar{x}) - \xi(t, t_0, x)| \\ & \leq |\xi(t, t_0, \bar{x}) - \xi(t, t_1, z)| + |\xi(t, t_0, x) - \xi(t, t_1, z)| < 2\varepsilon. \end{aligned}$$

Hence $\bar{x} \in B_{f, t_0}(x, 2\varepsilon, T)$, and thus $B_0(z) \subset \bigcup_{x \in E_0(z)} B_{f, t_0}(x, 2\varepsilon, T)$. Therefore,

$$K \subset \bigcup_{z \in \bar{E}_1} B_0(z) \subset \bigcup_{z \in \bar{E}_1} \bigcup_{x \in E_0(z)} B_{f, t_0}(x, 2\varepsilon, T),$$

that is, $\bigcup_{z \in \bar{E}_1} E_0(z)$ is a $(T, 2\varepsilon)$ -spanning set for K at t_0 . Consequently, the minimal cardinality of such a set satisfies

$$\begin{aligned} S(f, t_0, 2\varepsilon, T, K) & \leq \sum_{z \in \bar{E}_1} \#E_0(z) \\ & \leq S(f, t_1, \varepsilon, T, K_1) S(f, t_0, \varepsilon, t_1 - t_0, K). \end{aligned}$$

By the definition of entropy (4), we conclude that

$$\begin{aligned} h(f, t_0, K) & = \lim_{\varepsilon \searrow 0} \limsup_{T \rightarrow \infty} \frac{1}{T} \log S(f, t_0, 2\varepsilon, T, K) \\ & \leq \lim_{\varepsilon \searrow 0} \limsup_{T \rightarrow \infty} \frac{1}{T} \log S(f, t_1, \varepsilon, T, K_1) \\ & \quad + \lim_{\varepsilon \searrow 0} \lim_{T \rightarrow \infty} \frac{1}{T} \log S(f, t_0, \varepsilon, t_1 - t_0, K) \\ & \leq \lim_{\varepsilon \searrow 0} \limsup_{T \rightarrow \infty} \frac{1}{T} \log S(f, t_1, \varepsilon, T, K_1) \\ & = h(f, t_1, K_1). \end{aligned}$$

B. Proof of Lemma 6

Lemma 6 is established based on the following property of the *interconnected LTV system*

$$\dot{x}_i = \sum_{j=1}^m A_{ij}(t) x_j, \quad t \geq t_0, \quad i \in \{1, \dots, m\}, \quad (46)$$

where $A_{ij} : \mathbb{R} \rightarrow \mathbb{R}^{n_i \times n_j}$ are piecewise-continuous matrix-valued functions.

Lemma 10. Consider the matrix-valued function $\bar{A}_N^{\text{lin}}(t, t_0) = [\bar{a}_{ij}^{\text{lin}}(t, t_0)] \in \mathbb{R}^{m \times m}$ defined by

$$\begin{aligned} \bar{a}_{ii}^{\text{lin}}(t, t_0) & := \max_{s \in [t_0, t]} \mu_i(A_{ii}(s)), \\ \bar{a}_{ij}^{\text{lin}}(t, t_0) & := \max_{s \in [t_0, t]} \|A_{ij}(s)\|_{ij}, \quad i, j \in \{1, \dots, m\} : i \neq j. \end{aligned} \quad (47)$$

For all initial states $x = (x_1, \dots, x_m) \in \mathbb{R}^n$ with $x_i \in \mathbb{R}^{n_i}$ for all i , the solution of the interconnected LTV system (46) satisfies

$$\begin{bmatrix} |\xi_1(t, t_0, x)|_1 \\ \vdots \\ |\xi_m(t, t_0, x)|_m \end{bmatrix} \leq e^{\bar{A}_N^{\text{lin}}(t, t_0)(t-t_0)} \begin{bmatrix} |x_1|_1 \\ \vdots \\ |x_m|_m \end{bmatrix} \quad \forall t \geq t_0. \quad (48)$$

Proof. For all $i \in \{1, \dots, m\}$, $t_1 \geq t_0$, and $t \in [t_0, t_1]$, the right-sided derivative of the function $t \mapsto |\xi_i(t, t_0, x)|_i$, denoted by $\partial_t^+ |\xi_i(t, t_0, x)|_i$, satisfies

$$\begin{aligned} & \partial_t^+ |\xi_i(t, t_0, x)|_i \\ & = \lim_{s \searrow 0} \frac{|\xi_i(t+s, t_0, x)|_i - |\xi_i(t, t_0, x)|_i}{s} \\ & = \lim_{s \searrow 0} \frac{\left| \xi_i(t, t_0, x) + s \sum_{j=1}^m A_{ij}(t) \xi_j(t, t_0, x) \right|_i - |\xi_i(t, t_0, x)|_i}{s} \\ & \leq \lim_{s \searrow 0} \frac{\|I + s A_{ii}(t)\|_i - 1}{s} |\xi_i(t, t_0, x)|_i \\ & \quad + \sum_{j \neq i} \|A_{ij}(t)\|_{ij} |\xi_j(t, t_0, x)|_j \\ & = \mu_i(A_{ii}(t)) |\xi_i(t, t_0, x)|_i + \sum_{j \neq i} \|A_{ij}(t)\|_{ij} |\xi_j(t, t_0, x)|_j \\ & \leq \bar{a}_{ii}^{\text{lin}}(t_1, t_0) |\xi_i(t, t_0, x)|_i + \sum_{j \neq i} \bar{a}_{ij}^{\text{lin}}(t_1, t_0) |\xi_j(t, t_0, x)|_j. \end{aligned}$$

Therefore,

$$\partial_t^+ \begin{bmatrix} |\xi_1(t, t_0, x)|_1 \\ \vdots \\ |\xi_m(t, t_0, x)|_m \end{bmatrix} \leq \bar{A}_N^{\text{lin}}(t_1, t_0) \begin{bmatrix} |\xi_1(t, t_0, x)|_1 \\ \vdots \\ |\xi_m(t, t_0, x)|_m \end{bmatrix}$$

for all $t_1 \geq t_0$ and $t \in [t_0, t_1]$. As $\bar{A}_N^{\text{lin}}(t_1, t_0)$ defined in (47) is a Metzler matrix, the upper bound in (48) follows from standard comparison arguments analogous to those in the proof of [45, Prop. 1]. \square

Proof of Lemma 6. We prove Lemma 6 by constructing an LTV system (7) with a suitable matrix-valued function $A(t)$, using variational arguments from nonlinear systems analysis (see, e.g., [42, Sec. 2.5]). We then reformulate it as the interconnected LTV system (46) and apply Lemma 10.

Fix arbitrary initial states $x = (x_1, \dots, x_m)$, $\bar{x} = (\bar{x}_1, \dots, \bar{x}_m) \in K$ with $x_i, \bar{x}_i \in \mathbb{R}^{n_i}$ for all i , and define

$$\nu(t, \theta) := \theta \xi(t, t_0, \bar{x}) + (1 - \theta) \xi(t, t_0, x), \quad t \geq t_0, \theta \in [0, 1].$$

Then

$$\begin{aligned} & \partial_t (\xi(t, t_0, \bar{x}) - \xi(t, t_0, x)) \\ & = f(t, \xi(t, t_0, \bar{x})) - f(t, \xi(t, t_0, x)) \\ & = f(t, \nu(t, 1)) - f(t, \nu(t, 0)) \\ & = \int_0^1 J_x f(t, \nu(t, \theta)) \partial_\theta \nu(t, \theta) d\theta \\ & = \left(\int_0^1 J_x f(t, \nu(t, \theta)) d\theta \right) (\xi(t, t_0, \bar{x}) - \xi(t, t_0, x)) \end{aligned}$$

for all $t \geq t_0$, except on a set of measure zero where $\xi(t, t_0, \bar{x}) - \xi(t, t_0, x)$ is not continuously differentiable in t (see, e.g., [48, Sec. 4.2.4] for a more rigorous justification). Hence, the difference $\xi(t, t_0, \bar{x}) - \xi(t, t_0, x)$ solves the LTV system (7) with $A(t) := \int_0^1 J_x f(t, \nu(t, \theta)) d\theta$ and initial state $\bar{x} - x$ at t_0 .

Next, we write $A(t) = [A_{ij}(t)] \in \mathbb{R}^{n \times n}$ with $A_{ij}(t) := \int_0^1 J_{x_j} f_i(t, \nu(t, \theta)) d\theta \in \mathbb{R}^{n_i \times n_j}$ for $i, j \in \{1, \dots, m\}$. Then

the difference $(\xi_1(t, t_0, \bar{x}) - \xi_1(t, t_0, x), \dots, \xi_m(t, t_0, \bar{x}) - \xi_m(t, t_0, x))$ solves the corresponding interconnected LTV system (46) with initial state $(\bar{x}_1 - x_1, \dots, \bar{x}_m - x_m)$ at t_0 . By (48), we have

$$\begin{bmatrix} |\xi_1(t, t_0, \bar{x}) - \xi_1(t, t_0, x)|_1 \\ \vdots \\ |\xi_m(t, t_0, \bar{x}) - \xi_m(t, t_0, x)|_m \end{bmatrix} \leq e^{\bar{A}_N^{\text{lin}}(t, t_0)(t-t_0)} \begin{bmatrix} |\bar{x}_1 - x_1|_1 \\ \vdots \\ |\bar{x}_m - x_m|_m \end{bmatrix} \quad \forall t \geq t_0,$$

where the matrix-valued function $\bar{A}_N^{\text{lin}}(t, t_0) = [\bar{a}_{ij}^{\text{lin}}(t, t_0)] \in \mathbb{R}^{m \times m}$ is defined in (47).

Finally, we compare $\bar{A}_N^{\text{lin}}(t, t_0)$ to the matrix-valued function $\bar{A}_N(t, t_0, K) = [\bar{a}_{ij}(t, t_0, K)] \in \mathbb{R}^{m \times m}$ defined in (37). For all $t \geq t_0$, as the matrix measures $\mu_i(\cdot)$ and induced norms $\|\cdot\|_{ij}$ are convex functions and $\nu(t, \theta) \in \text{co}(\xi(t, t_0, K))$ for all $\theta \in [0, 1]$, we have

$$\begin{aligned} \bar{a}_{ii}^{\text{lin}}(t, t_0) &= \max_{s \in [t_0, t]} \mu_i \left(\int_0^1 J_{x_i} f_i(s, \nu(s, \theta)) d\theta \right) \\ &\leq \max_{s \in [t_0, t]} \int_0^1 \mu_i(J_{x_i} f_i(s, \nu(s, \theta))) d\theta \\ &\leq \max_{s \in [t_0, t]} \max_{\theta \in [0, 1]} a_{ii}(s, \nu(s, \theta)) \leq \bar{a}_{ii}(t, t_0, K), \\ \bar{a}_{ij}^{\text{lin}}(t, t_0) &= \max_{s \in [t_0, t]} \left\| \int_0^1 J_{x_j} f_i(s, \nu(s, \theta)) d\theta \right\|_{ij} \\ &\leq \max_{s \in [t_0, t]} \int_0^1 \|J_{x_j} f_i(s, \nu(s, \theta))\|_{ij} d\theta \\ &\leq \max_{s \in [t_0, t]} \max_{\theta \in [0, 1]} a_{ij}(s, \nu(s, \theta)) \leq \bar{a}_{ij}(t, t_0, K) \end{aligned}$$

for all $i, j \in \{1, \dots, m\}$ with $i \neq j$. Therefore, $\bar{A}_N^{\text{lin}}(t, t_0) \leq \bar{A}_N(t, t_0, K)$, and (38) follows. \square

C. Proof of Lemma 7

According to [41, Th. 7, p. 24], for each $v \in \mathbb{R}^n$, the Jacobian matrix $J_x \xi(t, t_0, v)$ is equal to the state-transition matrix $\Phi_A(t, t_0)$ of the LTV system (7) with the matrix-valued function $A(t) := J_x f(t, \xi(t, t_0, v))$. By Liouville's formula (see, e.g., [49, Th. 4.1, p. 28]), its determinant satisfies

$$\det(J_x \xi(t, t_0, v)) = e^{\int_{t_0}^t \text{tr}(J_x f(s, \xi(s, t_0, v))) ds} \quad \forall t \geq t_0.$$

Therefore,

$$\begin{aligned} \text{vol}(\xi(t, t_0, K)) &= \int_K |\det(J_x \xi(t, t_0, v))| dv \\ &\geq \left(\min_{v \in K} |\det(J_x \xi(t, t_0, v))| \right) \text{vol}(K) \\ &\geq \left(\min_{v \in K} e^{\int_{t_0}^t \text{tr}(J_x f(s, \xi(s, t_0, v))) ds} \right) \text{vol}(K) \\ &= e^{\gamma(t, t_0)} \text{vol}(K) \end{aligned}$$

for all $t \geq t_0$, that is, (39) holds. \square

D. Proof of Lemma 8

If (42) holds, then

$$K \subset \bigcup_{x \in G(\theta)} (R(x) \cap K) \subset \bigcup_{x \in G(\theta)} B_{f, t_0}(x, \varepsilon, T).$$

Hence, $G(\theta)$ is a (T, ε) -spanning set. Consequently, the minimal cardinality of such a set satisfies

$$\begin{aligned} S(f, t_0, \varepsilon, T, K) &\leq \#G(\theta) \\ &\leq \prod_{i=1}^n \left(\left\lfloor \frac{2r}{\theta_i} \right\rfloor + 1 \right) \leq \prod_{i=1}^n \left(\frac{2r}{\theta_i} + 1 \right). \end{aligned}$$

If (42) holds for all $T \geq 0$ and $\varepsilon > 0$, then the definition of entropy (4) implies that

$$\begin{aligned} h(f, t_0, K) &\leq \limsup_{\varepsilon \searrow 0} \limsup_{T \rightarrow \infty} \sum_{i=1}^n \frac{\log(2r/\theta_i + 1)}{T} \\ &\leq \limsup_{\varepsilon \searrow 0} \limsup_{T \rightarrow \infty} \sum_{i=1}^n \frac{\log(1/\theta_i)}{T} \\ &\quad + \sum_{i=1}^n \limsup_{\varepsilon \searrow 0} \limsup_{T \rightarrow \infty} \frac{\log(\theta_i + 2r)}{T}, \end{aligned} \quad (49)$$

where the last inequality holds as the upper limit is a subadditive function. Moreover, each summand in the last term of (49) satisfies

$$\begin{aligned} &\limsup_{\varepsilon \searrow 0} \limsup_{T \rightarrow \infty} \frac{\log(\theta_i + 2r)}{T} \\ &\leq \limsup_{\varepsilon \searrow 0} \limsup_{T \rightarrow \infty} \frac{\max\{\log(2\theta_i), \log(4r)\}}{T} \\ &= \max \left\{ \limsup_{\varepsilon \searrow 0} \limsup_{T \rightarrow \infty} \frac{\log \theta_i}{T}, 0 \right\}, \end{aligned}$$

where the last equality holds in part because r is independent of T . Therefore, (49) implies (44) under the condition (43). \square

APPENDIX II

PROOF SKETCHES OF PROPOSITIONS 2 AND 3

First, the proof of Proposition 3 is analogous to that of Theorem 2. The main difference lies in the construction of bounds on the separation between trajectories. To establish Proposition 3, we apply variational arguments to the line segment connecting two initial states, rather than the one connecting two solutions as in the proof of Lemma 6. This leads to the following lemma.

Lemma 11. *For the interconnection matrix function $A_N(t, x) = [a_{ij}(t, x)] \in \mathbb{R}^{m \times m}$ defined in (24), consider the matrix-valued function $\bar{A}_N^{\text{ini}}(t, t_0, K) = [\bar{a}_{ij}^{\text{ini}}(t, t_0, K)] \in \mathbb{R}^{m \times m}$ defined by*

$$\bar{a}_{ij}^{\text{ini}}(t, t_0, K) := \max_{s \in [t_0, t]} \max_{v \in \text{co}(K)} a_{ij}(s, \xi(s, t_0, v))$$

for $i, j \in \{1, \dots, m\}$. For all initial states $x = (x_1, \dots, x_m)$, $\bar{x} = (\bar{x}_1, \dots, \bar{x}_m) \in K$ with $x_i, \bar{x}_i \in \mathbb{R}^{n_i}$ for

all i , the solutions of the interconnected system (19) satisfy

$$\begin{bmatrix} |\xi_1(t, t_0, \bar{x}) - \xi_1(t, t_0, x)|_1 \\ \vdots \\ |\xi_m(t, t_0, \bar{x}) - \xi_m(t, t_0, x)|_m \end{bmatrix} \leq e^{\bar{A}_N^{\text{ini}}(t, t_0, K)(t-t_0)} \begin{bmatrix} |\bar{x}_1 - x_1|_1 \\ \vdots \\ |\bar{x}_m - x_m|_m \end{bmatrix} \quad \forall t \geq t_0.$$

Applying Lemma 11 in place of Lemma 6 yields the condition (31) for constants \bar{a}_{ij}^* in Proposition 3, in contrast to the condition (27) for constants \hat{a}_{ij} in Theorem 2.

Second, the proof of Proposition 2 is analogous to that of Theorem 2 with $m = 1$. As in the proof sketch of Proposition 3 above, the main difference lies in the construction of bounds on the separation between trajectories. In addition to the different application of variational arguments, we use Coppel's inequality (see, e.g., [42, Th. 2.5.3, p. 47]) instead of Lemma 10. This leads to the following lemma.

Lemma 12. *For all initial states $x, \bar{x} \in K$, the solutions of the system (1) satisfy*

$$|\xi(t, t_0, \bar{x}) - \xi(t, t_0, x)| \leq e^{\eta^{\text{ini}}(t, t_0, K)} |\bar{x} - x| \quad \forall t \geq t_0$$

with

$$\eta^{\text{ini}}(t, t_0, K) := \max_{v \in \text{co}(K)} \int_{t_0}^t \mu(J_x f(s, \xi(s, t_0, v))) \, ds.$$

Applying Lemma 12 in place of Lemma 6 yields the constant $\hat{\mu}^*$ defined by (30) in Proposition 2, in contrast to the constant $\hat{\mu}$ defined by (10) in Theorem 1.

Full proofs of Propositions 2 and 3 are available in [50].

REFERENCES

- [1] R. L. Adler, A. G. Konheim, and M. H. McAndrew, "Topological entropy," *Trans. Am. Math. Soc.*, vol. 114, no. 2, pp. 309–319, Feb. 1965.
- [2] A. N. Kolmogorov, "A new metric invariant of transitive dynamical systems and automorphisms of Lebesgue spaces," *Dokl. Akad. Nauk SSSR*, vol. 119, no. 5, pp. 861–864, 1958, in Russian.
- [3] E. I. Dinaburg, "The relation between topological entropy and metric entropy," *Dokl. Akad. Nauk SSSR*, vol. 190, no. 1, pp. 19–22, 1970, in Russian.
- [4] R. Bowen, "Entropy for group endomorphisms and homogeneous spaces," *Trans. Am. Math. Soc.*, vol. 153, pp. 401–414, Jan. 1971.
- [5] —, "Periodic points and measures for axiom A diffeomorphisms," *Trans. Am. Math. Soc.*, vol. 154, pp. 377–397, Feb. 1971.
- [6] E. I. Dinaburg, "On the relations among various entropy characteristics of dynamical systems," *Math. USSR-Izvestiya*, vol. 5, no. 2, pp. 337–378, Apr. 1971.
- [7] T. N. T. Goodman, "Relating topological entropy and measure entropy," *Bull. London Math. Soc.*, vol. 3, no. 2, pp. 176–180, Jul. 1971.
- [8] L. W. Goodwyn, "The product theorem for topological entropy," *Trans. Am. Math. Soc.*, vol. 158, no. 2, pp. 445–452, Aug. 1971.
- [9] A. Katok and B. Hasselblatt, *Introduction to the Modern Theory of Dynamical Systems*. Cambridge University Press, 1995.
- [10] T. Downarowicz, *Entropy in Dynamical Systems*. Cambridge University Press, 2011.
- [11] S. Kolyada and L. Snoha, "Topological entropy of nonautonomous dynamical systems," *Random & Comput. Dyn.*, vol. 4, no. 2-3, pp. 205–233, 1996.
- [12] S. Kolyada, M. Misiurewicz, and L. Snoha, "Topological entropy of nonautonomous piecewise monotone dynamical systems on the interval," *Fundam. Math.*, vol. 160, no. 2, pp. 161–181, 1999.
- [13] C. Kawan and Y. Latushkin, "Some results on the entropy of non-autonomous dynamical systems," *Dyn. Syst.*, vol. 31, no. 3, pp. 251–279, Jul. 2016.
- [14] A. Y. Pogromsky and A. S. Matveev, "Estimation of topological entropy via the direct Lyapunov method," *Nonlinearity*, vol. 24, no. 7, pp. 1937–1959, Jul. 2011.
- [15] G. N. Nair, R. J. Evans, I. M. Y. Mareels, and W. Moran, "Topological feedback entropy and nonlinear stabilization," *IEEE Trans. Automat. Contr.*, vol. 49, no. 9, pp. 1585–1597, Sep. 2004.
- [16] F. Colonius and C. Kawan, "Invariance entropy for control systems," *SIAM J. Control Optim.*, vol. 48, no. 3, pp. 1701–1721, May 2009.
- [17] F. Colonius, C. Kawan, and G. N. Nair, "A note on topological feedback entropy and invariance entropy," *Syst. & Control Lett.*, vol. 62, no. 5, pp. 377–381, May 2013.
- [18] F. Colonius, "Minimal bit rates and entropy for exponential stabilization," *SIAM J. Control Optim.*, vol. 50, no. 5, pp. 2988–3010, Oct. 2012.
- [19] A. V. Savkin, "Analysis and synthesis of networked control systems: topological entropy, observability, robustness and optimal control," *Automatica*, vol. 42, no. 1, pp. 51–62, Jan. 2006.
- [20] A. S. Matveev and A. Y. Pogromsky, "Observation of nonlinear systems via finite capacity channels: Constructive data rate limits," *Automatica*, vol. 70, pp. 217–229, Aug. 2016.
- [21] D. Liberzon and S. Mitra, "Entropy and minimal bit rates for state estimation and model detection," *IEEE Trans. Automat. Contr.*, vol. 63, no. 10, pp. 3330–3344, Oct. 2018.
- [22] J. P. Hespanha, A. Ortega, and L. Vasudevan, "Towards the control of linear systems with minimum bit-rate," in *15th Int. Symp. Math. Theory Networks Syst.*, 2002, 15 pages.
- [23] G. N. Nair and R. J. Evans, "Exponential stabilisability of finite-dimensional linear systems with limited data rates," *Automatica*, vol. 39, no. 4, pp. 585–593, Apr. 2003.
- [24] S. Tatikonda and S. Mitter, "Control under communication constraints," *IEEE Trans. Automat. Contr.*, vol. 49, no. 7, pp. 1056–1068, Jul. 2004.
- [25] D. Liberzon and J. P. Hespanha, "Stabilization of nonlinear systems with limited information feedback," *IEEE Trans. Automat. Contr.*, vol. 50, no. 6, pp. 910–915, Jun. 2005.
- [26] C. Kawan and J.-C. Delvenne, "Network entropy and data rates required for networked control," *IEEE Trans. Control Netw. Syst.*, vol. 3, no. 1, pp. 57–66, Mar. 2016.
- [27] A. S. Matveev, A. V. Proskurnikov, A. Pogromsky, and E. Fridman, "Comprehending complexity: Data-rate constraints in large-scale networks," *IEEE Trans. Automat. Contr.*, vol. 64, no. 10, pp. 4252–4259, Oct. 2019.
- [28] M. S. Tomar and M. Zamani, "Compositional quantification of invariance feedback entropy for networks of uncertain control systems," *IEEE Control Syst. Lett.*, vol. 4, no. 4, pp. 827–832, Oct. 2020.
- [29] D. Liberzon, "On topological entropy of interconnected nonlinear systems," *IEEE Control Syst. Lett.*, vol. 5, no. 6, pp. 2210–2214, Dec. 2021.
- [30] —, "Finite data-rate feedback stabilization of switched and hybrid linear systems," *Automatica*, vol. 50, no. 2, pp. 409–420, Feb. 2014.
- [31] G. Yang and D. Liberzon, "Feedback stabilization of switched linear systems with unknown disturbances under data-rate constraints," *IEEE Trans. Automat. Contr.*, vol. 63, no. 7, pp. 2107–2122, Jul. 2018.
- [32] H. Sibai and S. Mitra, "Optimal data rate for state estimation of switched nonlinear systems," in *20th Int. Conf. Hybrid Syst. Comput. Control*, 2017, pp. 71–80.
- [33] G. Yang, A. J. Schmidt, and D. Liberzon, "On topological entropy of switched linear systems with diagonal, triangular, and general matrices," in *57th IEEE Conf. Decis. Control*, Dec. 2018, pp. 5682–5687.
- [34] G. Yang and J. P. Hespanha, "On topological entropy of switched linear systems with pairwise commuting matrices," in *56th Annu. Allert. Conf. Commun. Control. Comput.*, Oct. 2018, pp. 429–436.
- [35] 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, Syst.*, vol. 32, no. 3, pp. 411–453, Sep. 2020.
- [36] G. Yang, J. P. Hespanha, and D. Liberzon, "On topological entropy and stability of switched linear systems," in *22nd ACM Int. Conf. Hybrid Syst. Comput. Control*, Apr. 2019, pp. 119–127.
- [37] G. Yang, D. Liberzon, and J. P. Hespanha, "Topological entropy of switched nonlinear and interconnected systems," *Math. Control. Signals, Syst.*, vol. 35, no. 3, pp. 641–683, Sep. 2023.
- [38] G. O. Berger and R. M. Jungers, "Worst-case topological entropy and minimal data rate for state observation of switched linear systems," in *23rd Int. Conf. Hybrid Syst. Comput. Control*, Apr. 2020, 11 pages.

- [39] G. S. Vicinansa and D. Liberzon, “Estimation entropy, Lyapunov exponents, and quantizer design for switched linear systems,” *SIAM J. Control Optim.*, vol. 61, no. 1, pp. 198–224, Feb. 2023.
- [40] F. Bullo, *Lectures on Network Systems*. Kindle Direct Publishing, 2022.
- [41] W. A. Coppel, *Stability and Asymptotic Behavior of Differential Equations*. Heath, 1965.
- [42] M. Vidyasagar, *Nonlinear Systems Analysis*, 2nd ed. SIAM, 2002.
- [43] D. S. Bernstein, *Matrix Mathematics*, 2nd ed. Princeton University Press, 2009.
- [44] G. Russo, M. di Bernardo, and E. D. Sontag, “A contraction approach to the hierarchical analysis and design of networked systems,” *IEEE Trans. Automat. Contr.*, vol. 58, no. 5, pp. 1328–1331, May 2013.
- [45] M. Arcak and J. Maidens, “Simulation-based reachability analysis for nonlinear systems using componentwise contraction properties,” in *Princ. Model.* Springer, 2018, ch. 4, pp. 61–76.
- [46] E. D. Sontag, “Contractive systems with inputs,” in *Perspect. Math. Syst. Theory, Control. Signal Process.* Springer, 2010, pp. 217–228.
- [47] V. A. Boichenko and G. A. Leonov, “Lyapunov’s direct method in estimates of topological entropy,” *J. Math. Sci.*, vol. 91, no. 6, pp. 3370–3379, Oct. 1998.
- [48] D. Liberzon, *Calculus of Variations and Optimal Control Theory*. Princeton University Press, 2012.
- [49] R. W. Brockett, *Finite Dimensional Linear Systems*. SIAM, 1970.
- [50] G. Yang and D. Liberzon, “Topological entropy of nonlinear time-varying systems,” Rutgers University, Tech. Rep., 2025, available at <https://guosong-yang.github.io/publications>.

PLACE
PHOTO
HERE

Guosong Yang (S’11–M’17) received the B.Eng. degree in Electronic Engineering from the Hong Kong University of Science and Technology, Hong Kong, in 2011, and the M.S. and Ph.D. degrees in Electrical and Computer Engineering from the University of Illinois at Urbana-Champaign, Urbana, IL, USA, in 2013 and 2017, respectively. He was a Postdoctoral Scholar at the University of California, Santa Barbara, Santa Barbara, CA, USA from 2017 to 2022, and a Postdoctoral Research Associate

at the University of Illinois at Urbana-Champaign, Urbana, IL, USA in 2022. He is currently an Assistant Professor in the Department of Electrical and Computer Engineering at Rutgers University, Piscataway, NJ, USA. His research interests span systems and control theory, optimization, and game theory, with emphasis on switched and hybrid systems, information-constrained control, nonconvex optimization, game-theoretic learning, and applications to cyber-physical systems and system resilience and security. Dr. Yang’s work has won the ACM SIGBED HSCC Best Paper Award in 2019.

PLACE
PHOTO
HERE

Daniel Liberzon (M’98–SM’04–F’13) received the Ph.D. degree in mathematics from Brandeis University in 1998 (under the supervision of Roger W. Brockett of Harvard University). Following a postdoctoral position in the Department of Electrical Engineering at Yale University from 1998 to 2000, he joined the University of Illinois Urbana-Champaign, where he is currently a Richard T. Cheng Professor of Electrical and Computer Engineering and a professor in the Coordinated Science Laboratory. His research

interests include switched and hybrid systems, nonlinear control theory, control with limited information, and uncertain and stochastic systems. He is the author of the books *Switching in Systems and Control* (Birkhäuser, 2003) and *Calculus of Variations and Optimal Control Theory: A Concise Introduction* (Princeton Univ. Press, 2012). His work has received several recognitions, including the 2002 IFAC Young Author Prize and the 2007 Donald P. Eckman Award. He was a plenary speaker at the IEEE Conference on Decision and Control in 2022 and served as a Senior Editor for AUTOMATICA from 2017 to 2022.