A small-gain-theorem-like approach to nonlinear observability via finite capacity channels

Pogromsky, A.; Matveev, A.; Proskurnikov, A.; Fridman, E.

A small-gain-theorem-like approach to nonlinear observability via finite capacity channels [★]

A. Pogromsky [**,***] **A. Matveev** [*] **A. Proskurnikov** [****] **E. Fridman** [†]

[*] *Department of Mathematics and Mechanics, Saint Petersburg University, St. Petersburg, Russia (e-mail: almat1712@yahoo.com).*
[**] *Department of Mechanical Engineering, Eindhoven University of Technology, Eindhoven, The Netherlands, (e-mail: A.Pogromsky@tue.nl).*
[***] *Department of Control Systems and Industrial Robotics, Saint-Petersburg National Research University of Information Technologies Mechanics and Optics (ITMO), Russia.*
[****] *Delft Center for Systems and Control, Delft University of Technology, Delft, The Netherlands (e-mail: anton.p.1982@ieee.org).*
[†] *Department of Electrical Engineering and Systems, Tel Aviv University, Tel Aviv 69978, Israel, (e-mail: emilia@eng.tau.ac.il).*

Abstract: The paper is concerned with observation of discrete-time, nonlinear, deterministic, and maybe chaotic systems via communication channels with finite data rates, with a focus on minimum data-rates needed for various types of observability. With the objective of developing tractable techniques to estimate these rates, the paper discloses benefits from regard to the operational structure of the system in the case where the system is representable as a feedback interconnection of two subsystems with inputs and outputs. To this end, a novel estimation method is elaborated, which is alike in flavor to the celebrated small gain theorem on input-to-output stability. The utility of this approach is demonstrated for general nonlinear time-delay systems by rigorously justifying an experimentally discovered phenomenon: Their topological entropy stays bounded as the delay grows without limits. This is extended on the studied observability rates and appended by constructive finite upper bounds independent of the delay. It is shown that these bounds are asymptotically tight for a time-delay analog of the bouncing ball dynamics.

Keywords: Observability, Nonlinear systems, Entropy, Second Lyapunov method, Data rate estimates

1. INTRODUCTION

One of the fundamental issues in the rapidly emerging area of control of networked systems is about constraints on communication among the network agents. Some key aspects of such constraints are commonly modeled based on a concept of communication channel with a limited data transmission rate. In this framework, a primary inquiry is about the minimal rate needed to achieve a desired control objective. This momentous threshold has been a subject of recent extensive studies; see e.g., (Baillieul, 2004; De Persis and Isidori, 2004; Nair et al., 2004; Liberzon and Hespanha, 2005; De Persis, 2005; Savkin, 2006; Nair et al., 2007; Matveev and Savkin, 2009) and literature therein. This data-rate threshold has appeared to be alike in spirit to the topological entropy (Donarowicz, 2011) of the system at hands, but is not always identical; and these studies have introduced various analogs of this entropy (Nair et al., 2004; Savkin, 2006; Colonius and Kawan, 2009, 2011; Kawan, 2011; Hagihara and Nair, 2013; Colonius et al., 2013; Matveev and Pogromsky, 2016).

Unlike linear plants, computation or even fine estimation of these thresholds is an intricate matter for nonlinear systems so that, e.g., the exact value of the topological entropy is still unknown even for many prototypical low-dimensional chaotic systems. This is consonant with rigorous uncomputability facts; e.g., even for piece-wise affine continuous maps and $\varepsilon \approx 0$, no program can generate a rational number in a finite time that approximates this entropy with precision ε (Koiran, 2001).

Nevertheless, it was recently shown that data rate thresholds concerned with observability and introduced in (Matveev and Pogromsky, 2016) can yet be computed in closed form for some of the above nonlinear prototypical systems, e.g., the bouncing ball system, Hénon system and logistic and Lozy maps (Matveev and Pogromsky, 2016; Pogromsky and Matveev, 2016a). This is thanks to the novel techniques for upper estimation of these thresholds that are elaborated in (Pogromsky et al., 2013; Pogromsky and Matveev, 2013, 2011, 2016b; Matveev and Pogromsky, 2016) and turn off the classic road of the first Lyapunov approach in study of topological entropy and the likes towards the second Lyapunov method.

The objective of this paper is to add more functionality to the just discussed approach of (Matveev and Pogromsky, 2016; Pogromsky and Matveev, 2016a) via its further elaboration in a situation where the observed dynamics result from a feedback interconnection of two subsystems with inputs and outputs. Such interconnection is very common in engineering practice so that certain whole chapters of control theory assume that the plant is given in this form. Among them, there are studies of absolute and robust stability, see, e.g. (Willems, 1972; Yakubovich, 2000, 2002; Megretski and Rantzer, 1997), where

interconnection customarily joins a nominal system to uncertainty. By following the lines of the famous small-gain theorem, where input-to-output stability of the system is guaranteed in terms of input-to-output features of open-loop subsystems, this paper discloses their input-to-output characteristics and relations among them that enable observability of the overall plant via communication channel with a given capacity.

To illustrate utility of these developments, we use them to rigorously prove a fact previously discovered via numerical studies for a few particular chaotic delayed systems: Their topological entropy remains bounded as the delay grows without limits (Manffra, 2002; Manffra et al., 2001). We prove that this phenomenon is common and extends on the studied observability rates, and also offer explicit upper bounds on them that are uniform over all delays.

The paper is organized as follows. Section 2 presents necessary background information. The problem statement and the main result are formulated in Sect. 3. Section 4 deals with time-delay systems, whereas Section 5 illustrates respective findings via an example.

2. OBSERVATION VIA FINITE-RATE CHANNELS

This section sets up key concepts to be used in our main results.

2.1 Observation Problem Statement and Topological Entropy

This paper is aimed at characterization of the minimal data rate needed to effectively observe, in real time, the current state $x(t) \in \mathbb{R}^n$ of a discrete-time invariant nonlinear system

$$x(t+1) = \phi[x(t)], \; t = 0, 1, \ldots, \quad x(0) \in K \subset \mathbb{R}^n. \quad (1)$$

Here the continuous map $\phi : \mathbb{R}^n \to \mathbb{R}^n$ and compact set K of initial states are known to the designer of observer. Whereas the full state is perfectly measured at the system's site, the state estimate is needed at a remote site, where sensory data can be delivered only over a discrete communication channel. It can transmit only a finite part of the infinity of bits embodying the perfect knowledge of the state. So it is preordained that at the remote site, the state estimate is inexact. Its accuracy depends on the content of the transmitted data and the transmission rate, with the latter being the main subject of interest to us.

At time t, only a discrete symbol $e(t)$ can be communicated across the channel. This necessitates a *coder*, which converts sensor readings $x(t) \in \mathbb{R}^n$ into such symbols. At the receiving end, a *decoder* produces (at time t) an estimate $\widehat{x}(t) \in \mathbb{R}^n$ of the current state $x(t)$ based on prior transmissions. The overall *observer* is composed of the coder and decoder, which are described by equations of the following respective forms:

$$e(t) = \mathfrak{C}[t, x(0), \ldots, x(t)|\widehat{x}(0), \delta], \quad t \geq 0,$$
$$\widehat{x}(t) = \mathfrak{D}[t, e(0), \ldots, e(t-1)|\widehat{x}(0), \delta], \quad t \geq 1. \quad (2)$$

They assume that the coder and decoder both have access to a common initial estimate $\widehat{x}(0)$ and its accuracy δ

$$\|x(0) - \widehat{x}(0)\| < \delta. \quad (3)$$

As for the transmission rate, we attend to its time-averaged value c by assuming that no less/more than $b_-(r)/b_+(r)$ bits of data can be transferred across the channel for any time interval of duration r, and the respective averaged rates are close to a common value c (the *channel capacity*) for $r \approx \infty$:

$$r^{-1}b_-(r) \to c \quad \text{and} \quad r^{-1}b_+(r) \to c \quad \text{as} \quad r \to \infty. \quad (4)$$

According to (Matveev and Savkin, 2009, Sect. 3.4), this model admits unsteady instant rate, transmission delays, and dropouts.

Our first definition addresses the possibility to merely keep the estimation error below an arbitrarily small level.

Definition 1. The system (1) is said to be *observable via a communication channel* if for any $\varepsilon > 0$, there exists an observer (2) and $\delta(\varepsilon, K) > 0$ such that

$$\|x(t) - \widehat{x}(t)\| \leq \varepsilon \quad \forall t \geq 0$$

whenever (3) holds with $\delta := \delta(\varepsilon, K)$, $x(0), \widehat{x}(0) \in K$.

The related demand to the channel rate refers to the *topological entropy* (Donarowicz, 2011) of the plant (1) on K

$$H(\phi, K) := \lim_{\varepsilon \to 0} \overline{\lim_{k \to \infty}} \frac{1}{k+1} \log_2 q(k, \varepsilon). \quad (5)$$

Here $q(k, \varepsilon)$ is the minimal number of elements in a set $Q \subset \mathbb{R}^{(k+1)n}$ that fits to approximate, with accuracy ε and for k steps, any trajectory $x(t, a)$ of (1) outgoing from $a \in K$:

$$\min_{\{x_0^\star, \ldots, x_k^\star\} \in Q} \max_{t=0, \ldots, k} \|x(t, a) - x_t^\star\| < \varepsilon \quad \forall a \in K. \quad (6)$$

Specifically, the following claim holds, which is well-known for linear systems and conforms to a number of nonlinear observability results, e.g., (Matveev and Savkin, 2009, Ch. 2).

Theorem 2. For observability via a communication channel, its capacity must be no less than the topological entropy of the system $c \geq H(\phi, K)$. Conversely, if the set K is positively invariant, the system is observable whenever $c > H(\phi, K)$.

2.2 Regular and Fine Observability

Definition 1 does not exclude critical degradation of accuracy over time: $\varepsilon \gg \delta(\varepsilon, K)$. The next definition disallows a violent regress: The accuracy stays proportional to its initial value.

Definition 3. The observer (2) is said to *regularly observe the system* (1) if there exist $\delta_* > 0$ and $G > 0$ such that the estimation accuracy $\|x(t) - \widehat{x}(t)\| \leq G\delta \; \forall t \geq 0$ whenever $x(0), \widehat{x}(0) \in K$ and in (3), δ is small enough $\delta < \delta_*$.

A stronger property is that the initial accuracy is also restored and exponentially improved.

Definition 4. The observer (2) is said to *finely observe the system* (1) if there exist $\delta_*, G > 0$, and $g \in (0, 1)$ such that

$$\|x(t) - \widehat{x}(t)\| \leq G\delta g^t \quad \forall t \geq 0$$

whenever $x(0), \widehat{x}(0) \in K$ and $\delta < \delta_*$ in (3).

Definition 5. The system (1) is said to be *regularly/finely observable via a communication channel* if there exists an observer (2) that regularly/finely observes the system (1) and operates via the channel at hands.

What channel capacity c is needed for observability in each of the above senses and how the answers relate to one another?

The averaged rate (4) is a comprehensive figure of merit for channel evaluation regarding these issues, and the larger this rate the better (Matveev and Pogromsky, 2016). So the posed question in fact addresses the infimum $\mathcal{R}(\phi, K)$ of the needed c's, where $\mathcal{R}$ is equipped with the index either ₒ, or ᵣₒ, or fₒ, which refers to "observability", "regular" and "fine observability", respectively. These quantities are called the *observability rates* and are fully determined by the plant (1). It is clear that observability $\Leftarrow$ regular observability $\Leftarrow$ fine observability and

$$H(\phi, K) = \mathcal{R}_o(\phi, K) \leq \mathcal{R}_{ro}(\phi, K) \leq \mathcal{R}_{fo}(\phi, K). \quad (7)$$

The first equation holds by Theorem 2. The second inequality turns into an equality in an important particular case.

Lemma 6. (Matveev and Pogromsky, 2016) For any positively invariant $\phi(K) \subset K$ compact set K, the regular and fine observability rates are the same: $\mathcal{R}_{\mathrm{ro}}(\phi, K) = \mathcal{R}_{\mathrm{fo}}(\phi, K)$.

Theorem 2 and the results of, e.g., (Matveev and Savkin, 2009, Sect. 3.5) imply that for any linear $x(t+1) = Ax(t)$ system (1), all three rates from (7) coincide and are explicitly given by $\mathcal{R}_{\mathrm{o}}(\phi, K) = \mathcal{R}_{\mathrm{ro}}(\phi, K) = \mathcal{R}_{\mathrm{fo}}(\phi, K) = H(A)$, where the quantity $H(A)$ is defined by the eigenvalues $\lambda_1, \dots, \lambda_n$ of A

$$H(A) := \sum_{i=1}^{n} \log_2 \max\{|\lambda_i|; 1\}. \tag{8}$$

For nonlinear systems, computation or even fine estimation of the topological entropy is an intricate matter so that its exact value is still unknown even for many popular low-dimensional chaotic systems, like the Hénon map, Dufing and van der Pol oscillators, and bouncing ball system, etc.

Paradoxically, the regular and fine observability rates can be computed in closed form for a number of prototypical chaotic systems, e.g., the bouncing ball system and, under certain circumstances, Hénon system (Matveev and Pogromsky, 2016), as well as logistic and Lozy maps (Pogromsky and Matveev, 2016a). This potentiality stems from novel techniques for upper estimation of these rates that are elaborated in (Pogromsky and Matveev, 2013, 2011, 2016b; Matveev and Pogromsky, 2016) and follow the lines of the second Lyapunov method.

Meanwhile, these techniques inherit the main problem related to the second Lyapunov approach to stability analysis: the lack of systematic procedures to construct proper Lyapunov functions for nonlinear systems. Somewhat general methods are available only for particular classes of systems and kinds of stability. An example emerged in the absolute stability theory, which has reached a certain degree of exhaustiveness in the part concerned with the so called integral quadratic constraints in both time and frequency domains (Yakubovich, 2000, 2002; Megretski and Rantzer, 1997). The involved system's analysis via integral quadratic constraints (SAIQC) is backed by both efficient analytical results, like KYP lemma, and numerical procedures, like LMI or matrix Ricatti equation solvers.

This paper is aimed to inject these merits into the data-rate estimation techniques from (Pogromsky et al., 2013; Pogromsky and Matveev, 2013, 2011, 2016b; Matveev and Pogromsky, 2016) in order to acquire a method with improved efficacy. To this end, the following steps are undertaken:

(1) The plant is studied in the form of a feedback interconnection of two sub-systems;
(2) Each of them is in fact replaced by its linear approximation, which injects linearity into the matter at hands;
(3) Every linearized sub-system is characterized by the input-to-output gain properties termed as quadratic constraints by using quadratic "storage-" and "supply-"like functions in a fashion portrayed in, e.g., (Willems, 1972; Megretski and Rantzer, 1997);
(4) The final data-rate estimate is underpinned by an argument in the spirit of a small gain theorem.

Here the hints 1) and 3) are typical for SAIQC and H_∞-control, whereas the small gain theorem is generally attributed to them.

3. TOPOLOGICAL ENTROPY AND OBSERVABILITY RATES OF INTERCONNECTED SYSTEMS

3.1 Problem statement

We consider a system in the form of feedback interconnection of two subsystems Σ_1 and Σ_2 with inputs and outputs, see Fig. 1. The subsystem Σ_i is governed by the following equations:

$$\Sigma_i : \begin{cases} x_i(t+1) = \phi_i[x_i(t), u_i(t)], \\ y_i(t) = h_i[x_i(t)], \end{cases} \quad i = 1, 2. \tag{9}$$

Here $x_i \in \mathbb{R}^{n_i}$, $u_i \in \mathbb{R}^{m_i}, y_1 \in \mathbb{R}^{m_2}, y_2 \in \mathbb{R}^{m_1}$. The interconnection is described by the following static relation

$$u_1(t) = y_2(t), \qquad u_2(t) = y_1(t). \tag{10}$$

The maps $\phi_i(\cdot), h_i(\cdot)$ are continuously differentiable. The con-
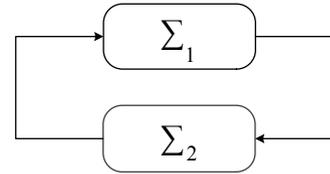

Fig. 1. A feedback interconnection

sidered interconnection does give a system of the form (1) with

$$\phi(x) = \begin{bmatrix} \phi_1(x_1, h_2(x_2)) \\ \phi_2(x_2, h_1(x_1)) \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Our goal is to provide estimates of its observability rates in terms of input-output properties of the subsystems.

3.2 Basic constructions and assumptions

Further analysis will address the subsystems (9) via their first order approximation in a vicinity of a particular solution. To this end, we introduce the following matrix sequences

$$A_i(t) = \frac{\partial \phi_i}{\partial x_i}[x_i(t), u_i(t)], \quad B_i(t) = \frac{\partial \phi_i}{\partial u_i}[x_i(t), u_i(t)],$$

$$C_i(t) = \frac{\partial h_i}{\partial x_i}[x_i(t)], \qquad i = 1, 2.$$

These approximations are described by the following linear time-varying systems of difference equations, where $i = 1, 2$

$$z_i(t+1) = A_i(t)z_i(t) + B_i(t)w_i(t),$$
$$\zeta_i(t) = C_i(t)z_i(t). \tag{11}$$

Here $z_i \in \mathbb{R}^{n_i}$, $w_i \in \mathbb{R}^{m_i}$, $\zeta_1 \in \mathbb{R}^{m_2}, \zeta_2 \in \mathbb{R}^{m_1}$, whereas the interconnection equations (10) entail that

$$w_1(t) = \zeta_2(t), \qquad w_2(t) = \zeta_1(t). \tag{12}$$

A productive approach to characterization of input-to-output properties of linear systems is by using quadratic dissipation inequalities (Willems, 1972). We shall follow these lines and associate (11) with a quadratic "storage" $z_i^\top P_i z_i$ function and a "supply" rate in the form $z_i^\top [Q_i - P_i] z_i - \frac{1}{\gamma_i} \zeta_i^\top \zeta_i + \gamma_i w_i^\top w_i$. More precisely, the following assumption is adopted.

Assumption 7. For $i = 1, 2$, there exist $n_i \times n_i$ matrices $P_i = P_i^\top > 0$, $Q_i = Q_i^\top \geq 0$ and a positive number $\gamma_i > 0$ such that

the following inequality is true along all solutions of (9), (10) starting in the given compact set K:

$$[A_i(t)z_i + B_i(t)w_i]^\top P_i[A_i(t)z_i + B_i(t)w_i] \leq z_i^\top Q_i z_i$$
$$- \frac{1}{\gamma_i}\zeta_i^\top\zeta_i + \gamma_i w_i^\top w_i, \quad \zeta_i = C_i(t)z_i \quad \forall z_i, w_i, t. \quad (13)$$

If $Q_i \leq P_i$, this yields the upper estimate γ_i of the l_2-gain of the system (11) from the input w_i to output ζ_i

$$\sum_t \zeta_i^2(t) \leq \gamma_i^2 \sum_t w_i^2(t)$$

and Assumption 7 thus gives an upper bound γ_i on the incremental l_2-gain for the subsystem (9). With this in mind, we proceed in the spirit of the small-gain theorem.

Assumption 8. In Assumption 7, $\gamma_1\gamma_2 \leq 1$.

We introduce the following subsets of $\mathbb{R}^n$, $n := n_1 + n_2$

$$X(t) := \{x = x(t,a) : a \in K\}, \quad X^\infty := \bigcup_{t=0}^\infty X(t), \quad (14)$$

where $x(t,a), t \geq 0$ is the solution of the primal interconnected system (9), (10) that starts with $x(0) = a$. Our last assumption is purely technical and employs the following.

Definition 9. A map $f(x)$ from $\mathbb{R}^n$ to an Euclidean space is said to be *uniformly continuous near a subset* $X_* \subset \mathbb{R}^n$ if for any $\varepsilon > 0$, there exists $\delta > 0$ such that

$$x \in X_* \wedge \|x - x'\| < \delta \Rightarrow \|f(x) - f(x')\| < \varepsilon.$$

As is well known, any continuous function is uniformly continuous near any compact set.

Assumption 10. Let j_i denote the element of the set $\{1,2\}$ complementary to $i \in \{1,2\}$. For $i = 1,2$, the derivatives of the associated functions from (9)

$$\frac{\partial\phi_i}{\partial x_i}[x_i, h_i(x_{j_i})], \frac{\partial\phi_i}{\partial u_i}[x_i, h_i(x_{j_i})], \frac{\partial h_i}{\partial x_i}[x_i]$$

are bounded on X^∞ and uniformly continuous near this set.

This holds if the set X^∞ is bounded, in particular, if the given compact set K of initial states is positively invariant.

In conclusion, we note that the afore-mentioned inequality $Q_i \leq P_i$ is not assumed and was discussed for an illustrative purpose only. So Assumption 7 can be always satisfied since $A_i(t), B_i(t)$ are bounded thanks to Assumption 10: for arbitrary γ_i, it suffices to pick P_i and Q_i "small" and "large" enough, respectively. This choice influences our estimate of the observability rates and will be discussed in more details further.

3.3 The main result

Let $P > 0$ and $Q \geq 0$ be square symmetric matrices of a common size. The roots of the algebraic equation

$$\det(Q - \lambda P) = 0 \quad (15)$$

are equal to the eigenvalues of each of the matrices QP^{-1}, $P^{-1}Q$ and $P^{-1/2}QP^{-1/2} \geq 0$. Since the latter matrix in symmetric and positive semi-definite, these roots are nonnegative.

We enumerate them $\lambda_1 = \lambda_1(P,Q) \geq \ldots \geq \lambda_n = \lambda_n(P,Q)$ in the descending order, repeating any root in accordance with its algebraic multiplicity. Partly inspired by (8), we introduce

$$H_L(P,Q) := \frac{1}{2}\sum_j \max\{0, \log_2\lambda_j\}, \quad (16)$$

where the sum is over all j's and $\log_2 0 = -\infty$.

Theorem 11. Let Assumptions 7—10 hold and let P_i, Q_i be taken from Assumption 7. Then the observability rates of the system (9)—(10) obey the following inequalities

$$H(\phi, K) \leq \mathcal{R}_{\text{ro}}(\phi, K)$$
$$\leq \mathcal{R}_{\text{fo}}(\phi, K) \leq H_L(P_1, Q_1) + H_L(P_2, Q_2). \quad (17)$$

Proof: We put $z = (z_1^\top, z_2^\top)^\top$, $P := \mathbf{diag}\{P_1, P_2\}, Q := \mathbf{diag}\{Q_1, Q_2\}$. By using zero averaging functions $v_i(\cdot)$, Theorem 12 from (Matveev and Pogromsky, 2016) yields that

$$\mathcal{R}_{\text{fo}}(\phi, K) \leq \sup_{x \in X^\infty} H_L[P, \mathcal{Q}(x)], \ \mathcal{Q}(x) := \phi'(x)^\top P\phi'(x).$$

By invoking (13), we observe that

$$z^\top\mathcal{Q}(x)z = (A_1z_1 + B_1w_1)^\top P_1(A_1z_1 + B_1w_1)$$
$$+ (A_2z_2 + B_2w_2)^\top P_2(A_2z_2 + B_2w_2)$$
$$\leq z_1^\top Q_1 z_1 + z_2^\top Q_2 z_2 + \gamma_1\zeta_2^\top\zeta_2 - \frac{1}{\gamma_2}\zeta_2^\top\zeta_2 + \gamma_2\zeta_1^\top\zeta_1 - \frac{1}{\gamma_1}\zeta_1^\top\zeta_1$$
$$\leq z_1^\top Q_1 z_1 + z_2^\top Q_2 z_2 = z^\top Q z,$$

or briefly, $\mathcal{Q}(x) \leq Q$. Then $P^{-1/2}\mathcal{Q}(x)P^{-1/2} \leq P^{-1/2}QP^{-1/2}$ and thus $\lambda_j[P, \mathcal{Q}(x)] \leq \lambda_j[P, Q]$ due to Weyl's inequalities for eigenvalues of symmetric matrices (Franklin, 1993, Sect. 6.7). Hence $H_L[P, \mathcal{Q}(x)] \leq H_L[P, Q]$ and so $\mathcal{R}_{\text{fo}}(\phi, K) \leq H_L[P, Q]$. Since both P and Q are block-diagonal, it is easy to see that $H_L[P, Q] = H_L[P_1, Q_1] + H_L[P_2, Q_2]$. The proof is completed by (7). ∎

This theorem provides an estimate of the observability rates for the feedback interconnection of two subsystems via their input-output properties. These properties are captured by the pair $[H_L(P_i, Q_i), \gamma_i]$, which is not unique for a given subsystem. For example, (13) survives the change $\gamma_i := \tau\gamma_i, P_i := \tau P_i, Q_i := Q_i + \frac{1-\tau^2}{\tau\gamma_i}C_i^\top C_i$ with any $\tau \in (0,1)$, which decreases γ_i but increases $H_L(P_i, Q_i)$. By picking τ small enough, γ_i can be made as small as desired, in particular, such that Assumption 8 is satisfied. However, the accompanying trend in $H_L(P_i, Q_i)$ worsens the right-hand side of (17) and so larger γ_i's carry a potential of a finer estimate.

The next section illustrates the utility of Theorem 11 via its application to time delay systems. To this end, such a system is represented as a feedback interconnection of a controlled undelayed system and a pure delay line. Irrespective of the delay value, this line is linear, its incremental l_2-gain does not exceed 1, and due to stability, the related quantity $H_L = 0$. So Theorem 11 entails an estimate of the observability rates that is independent of the delay and shows that these rates stay bounded as the delay increases without limits.

4. ENTROPY OF SYSTEMS WITH LARGE DELAYS

Now we turn to delayed discrete-time systems of the form

$$x(t+1) = f[x(t), Cx(t-\tau)], t = 0, 1, \ldots. \quad (18)$$

Here $\tau > 0$ is an integer delay, $x(t) \in \mathbb{R}^n$, the smooth function $f(x,r) \in \mathbb{R}^n$ of $x \in \mathbb{R}^n, r \in \mathbb{R}^d$ is given, and the $r \times n$-matrix C typically "cuts out" a certain part of the state x. The initial states are restricted by a given compact set $\mathcal{K} \subset \mathbb{R}^n$ as follows

$$x(0) \in \mathcal{K}, x(-1) \in \mathcal{K}, \ldots, x(-\tau) \in \mathcal{K}. \quad (19)$$

The standard state augmentation

$$\mathfrak{x}(t) := [x(t), x(t-1), \ldots, x(t-\tau)] \quad (20)$$

shapes this system into (1) with

$$\phi(\mathfrak{x}) = [f(x_0, Cx_{-\tau}), x_0, \ldots, x_{-\tau+1}] \forall \mathfrak{x} = [x_0, \ldots, x_{-\tau}]$$

and $K := \{\mathfrak{x} : x_j \in \mathcal{K} \; \forall j\}$. So all concepts from Section 2 are fully applicable to (18), (19).

Given $f(\cdot, \cdot)$ and $\mathcal{K}$, we are interested in behavior of the topological entropy $H(\tau)$ and the observability rates $\mathcal{R}_{\mathrm{ro}}(\tau)$ and $\mathcal{R}_{\mathrm{fo}}(\tau)$ of the system (18), (19) as the delay τ increases to ∞. This interest is partly inspired by (Manffra, 2002; Manffra et al., 2001) where it is discovered via numerical studies that for particular chaotic delayed systems, $H(\tau)$ remains bounded as $\tau \to \infty$. Now we rigorously prove that this phenomenon is common and also extends on $\mathcal{R}_{\mathrm{ro/fo}}$, and offer explicit upper bounds on these quantities that are uniform over all delays.

We impose the following analog of Assumption 10.

Assumption 12. There is $\mathcal{K}_* \subset \mathbb{R}^n$ such that the following holds:

  i. Irrespective of the delay, any solution of (18) satisfying (19) lies in $\mathcal{K}_*$, i.e., $x(t) \in \mathcal{K}_* \; \forall t \geq 0$;
  ii. The first derivatives of $f(\cdot, \cdot)$ are bounded on $\mathcal{K}_* \times C\mathcal{K}_*$ and uniformly continuous near this set.

This is true with $\mathcal{K}_* := \mathcal{K}$ if the compact set $\mathcal{K}$ is positively invariant for any τ, and with $\mathcal{K}_* := \mathbb{R}^n$ if the derivatives are bounded and uniformly continuous on the entire space $\mathbb{R}^n \times \mathbb{R}^d$.

For any $\varpi = (x, r), x \in \mathbb{R}^n, r \in \mathbb{R}^d$, we put
$$A(\varpi) := \frac{\partial f}{\partial x}(\varpi), \quad B(\varpi) := \frac{\partial f}{\partial r}(\varpi).$$
The next assumption is inspired by Assumption 7 with $\gamma_i := 1$.

Assumption 13. There are symmetric $n \times n$-matrices $P > 0, Q \geq 0$ such that for any $\varpi \in \mathcal{K}_* \times C\mathcal{K}_*$,
$$[A(\varpi)z + B(\varpi)w]^\top P[A(\varpi)z + B(\varpi)w] \leq z^\top Q z$$
$$- \zeta^\top \zeta + w^\top w, \quad \zeta = Cz, \quad \forall z \in \mathbb{R}^n, w \in \mathbb{R}^d. \quad (21)$$

Theorem 14. Suppose that Assumptions 12 and 13 hold. Then
$$H(\tau) \leq \mathcal{R}_{\mathrm{ro}}(\tau) \leq \mathcal{R}_{\mathrm{fo}}(\tau) \leq H_L(P, Q) \qquad \forall \tau.$$

Proof: We represent the system (18) as the interconnection (10) of the following two subsystems
$$\Sigma_1 : \begin{cases} x_1(t+1) = f[x_1(t), u_1(t)] \in \mathbb{R}^n, u_1(t) \in \mathbb{R}^d \\ y_1(t) = Cx_1(t) \in \mathbb{R}^d \end{cases}$$
$$\Sigma_2 : \begin{cases} x_2(t+1) = \mathfrak{A} x_2(t) + \mathfrak{B} u_2(t) \in \mathbb{R}^{d\tau}, u_2(t) \in \mathbb{R}^d \\ y_2(t) = \mathfrak{C} x_2(t) \in \mathbb{R}^d \end{cases}$$

Here the second subsystem is a τ-step delay line:
$$\mathfrak{A} = \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 \\ I_d & 0 & \cdots & 0 & 0 \\ 0 & I_d & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & 0 & 0 \\ 0 & 0 & \cdots & I_d & 0 \end{pmatrix}, \mathfrak{B} = \begin{pmatrix} I_d \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \mathfrak{C} = \begin{pmatrix} 0 & 0 & \cdots & 0 & I_d \end{pmatrix}$$

and I_s is the $s \times s$-identity matrix. For Σ_1, Assumption 7 holds with $P_1 := P, Q_1 := Q, \gamma_1 := 1$ thanks to Assumptions 12 (part **i**) and 13. For the second subsystem, it is easy to see that
$$x_1(t+1)^\top x_1(t+1) = x_1(t)^\top x_1(t) + u_1(t)^\top u_1(t) - y_1(t)^\top y_1(t)$$
and so Assumption 7 is true with $P_2 = Q_2 = I_{d\tau}$ and $\gamma_2 = 1$. Thus Assumption 8 holds, whereas Assumption 10 follows from Assumption 12. Theorem 11 completes the proof by noting that $H_L(P_2, Q_2) = 0$. ∎

Assumption 13 can be always satisfied in the most general case where $C = I_n$ and so (18) shapes into
$$x(t+1) = f[x(t), x(t-\tau)], t = 0, 1, \ldots. \quad (22)$$

Indeed, by **ii** in Assumption 12, there is $\eta \in (0, \infty)$ such that
$$\left[\frac{\partial f}{\partial x'}(\varpi), \frac{\partial f}{\partial x''}(\varpi)\right]^\top \left[\frac{\partial f}{\partial x'}(\varpi), \frac{\partial f}{\partial x''}(\varpi)\right] \leq \eta I_{2n}$$
for all $\varpi \in \mathcal{K}_* \times \mathcal{K}_*$. Then Assumption 13 clearly holds with $P := \eta^{-1} I_n, Q := 2I_n$ and so Theorem 14 yields the following

Corollary 15. For any system (22) satisfying Assumptions 12,
$$H(\tau) \leq \mathcal{R}_{\mathrm{ro}}(\tau) \leq \mathcal{R}_{\mathrm{fo}}(\tau) \leq \frac{n}{2} \max\{1 + \log_2 \eta; 0\} < \infty \quad \forall \tau.$$

5. EXAMPLE

We consider a delayed analog of the "bouncing-ball dynamics" (Matveev and Pogromsky, 2016), which is among celebrated prototypic examples of low-dimensional chaotic behavior,
$$y(t+1) = (1+\alpha)y(t) - \beta \cos y(t) - \alpha y(t-\tau), \quad t \geq 0. \quad (23)$$
Here $\tau > 0$ is an integer delay, $y \in \mathbb{R}$, and $\alpha, \beta > 0$ are given parameters. Since equation (23) is invariant under the transformation $y \mapsto y \pm 2\pi$, the phase space of the system can be viewed as the unit circle S_0^1. By (Matveev and Pogromsky, 2016, Remark 5), the concepts from Section 2 are fully applicable to this system (S_0^1-system). Theorem 14 immediately implies the following delay-independent estimate.

Corollary 16. For the system (23) with any delay τ,
$$H(\tau) \leq \mathcal{R}_{\mathrm{ro}}(\tau) \leq \mathcal{R}_{\mathrm{fo}}(\tau) \leq \log_2(1 + 2\alpha + \beta). \quad (24)$$

Proof: At first, we treat (23) as a system (18), (19) in $\mathbb{R}$, where
$$x = y, C = 1, f(x', x'') = (1+\alpha)x' - \beta \cos x' - \alpha x''$$
in (18) and $\mathcal{K} := [-\pi, \pi]$ in (19). Then Assumption 12 holds with $\mathcal{K}_* := \mathbb{R}$. To check Assumption 13, we note that in (21),
$$A(\varpi)z + B(\varpi)w = (1 + \alpha + \beta \sin y)z - \alpha w \quad \forall \varpi = (y, y_1).$$
Now $P = p \in (0, \infty), Q = q \in [0, \infty)$ and $z^\top P z = pz^2, z^\top Q z = qz^2$. So the left-hand side L of (21)
$$L = p|(1 + \alpha + \beta \sin y)z - \alpha w|^2 \leq p[\gamma|z| + \alpha|w|]^2,$$
where $\gamma := 1 + \alpha + \beta$. Hence (21) does hold whenever
$$(q-1)|z|^2 + |w|^2 - p[\gamma|z| + \alpha|w|]^2 \geq 0.$$
By treating the left-hand side as a quadratic form in $|z|, |w|$ and applying Sylvester's criterion, we see that (21) is true whenever
$$1 - p\alpha^2 > 0 \text{ and } 0 \leq \begin{vmatrix} q - 1 - p\gamma^2 & -p\gamma\alpha \\ -p\gamma\alpha & 1 - p\alpha^2 \end{vmatrix}$$
$$= q(1 - p\alpha^2) - [1 + p(\gamma^2 - \alpha^2)].$$
Thus (21) holds with
$$p \in (0, \alpha^{-2}), \quad q := \frac{1 + p(\gamma^2 - \alpha^2)}{1 - p\alpha^2}.$$
Then the algebraic equation (15) has the unique root $\lambda = q/p$. From now on, we consider the point $p = \frac{1}{\alpha(\gamma+\alpha)} \in (0, \alpha^{-2})$, where λ attains its maximum value $\lambda = (\alpha + \gamma)^2$. Modulo elementary calculations, Theorem 14 and (16) then imply (24).

However, (24) has been established for the system (23) considered in $\mathbb{R}$ ($\mathbb{R}$-system), whereas the corollary addresses the S_0^1-system. Due to the state augmentation (20), the state space $\mathfrak{X}_\mathbb{R}$ of the former is $\mathbb{R}^{\tau+1} = \{\mathfrak{x} = [x_0, \ldots, x_{-\tau}]\}$, whereas the state space $\mathfrak{X}_{S_0^1}$ of the latter is the multidimensional torus $\{[s_0, \ldots, s_{-\tau}] : s_{-j} \in S_0^1 \; \forall j\}$. There is a covering projection
$$J(\mathfrak{x}) := [e^{\imath x_0}, \ldots, e^{\imath x_{-\tau}}]$$
of $\mathfrak{X}_\mathbb{R}$ onto $\mathfrak{X}_{S_0^1}$. For any one step, the projection of the next state of $\mathbb{R}$-system is clearly the next state of S_0^1-system starting

from the projection of the initial state of $\mathbb{R}$-system. Then for the S_0^1-system, any data rate from (24) does not exceed the respective data rate for the $\mathbb{R}$-system by Lemma 13 in (Matveev and Pogromsky, 2016), which completes the proof. ∎

Whenever only an upper bound on the quantities of interest is given, like in (24), a question arises how tight this bound is. Theorem 15 in (Matveev and Pogromsky, 2016) offers an exact formula for $\mathcal{R}_{\mathrm{ro/fo}}(0)$ and thus displays a gap in (24) for the zero delay: $\mathcal{R}_{\mathrm{ro}}(0) = \mathcal{R}_{\mathrm{fo}}(0) < \log_2(1 + 2\alpha + \beta)$. Whereas direct computation of $\mathcal{R}_{\mathrm{ro/fo}}(\tau)$ is not easy for $\tau > 0$, asymptotic (as $\tau \to \infty$) lower estimates can be found in closed form to assess the gap in (24) for very large delays.

To this end, we start with a simple lemma.

Lemma 17. Whenever $\alpha > 0$ and $a > 1 + \alpha$, the equation

$$\chi_n(\lambda) = \lambda^n - a\lambda^{n-1} + \alpha = 0 \qquad (25)$$

has a root $\lambda \in (a(n-1)/n; a)$ if the integer n is large enough.

Proof: It suffices to note that χ_n is continuous, $\chi_n(a) = \alpha > 0$,

$$\chi_n\left[a(n-1)/n\right] = \alpha - \frac{a^n}{n-1}\left[1 - \frac{1}{n}\right]^n \to -\infty \text{ as } n \to \infty. \ \blacksquare$$

Now we consider the linearization of (23) at the equilibrium point $\pi/2$. It is easy to see that the eigenvalues of the respective Jacobian matrix are the roots of (25), where $a = 1 + \alpha + \beta$ and $n = \tau + 1$. Due to Lemma 17, this Jacobian has an eigenvalue that converges to a as $\tau \to \infty$. Withal, (25) has no roots on the unit circle S_0^1 (since $|\lambda| = 1 \Rightarrow |\lambda^{n-1}(a-\lambda)| \geq a-1 > \alpha$) and so the equilibrium is hyperbolic. Then Theorem 9 in (Matveev and Pogromsky, 2016) entails the following.

Corollary 18. For any $\alpha, \beta > 0$ the following inequality holds

$$\liminf_{\tau \to \infty} \mathcal{R}_{\mathrm{ro}}(\tau) \geq \log_2(1 + \alpha + \beta). \qquad (26)$$

Thus for $\alpha \ll \beta$, the estimates (24) become tight as $\tau \to \infty$. The smaller the ratio $\alpha/(\beta + 1)$, the narrower the gap between the upper (24) and lower (26) bounds on the observability rates.

By applying the Kalman-Szegö lemma like in (Matveev and Pogromsky, 2016, Sect. 5), it can be shown that the bound (26) is tight and the observability rates converge to a common limit $\mathcal{R}_{\mathrm{ro}}(\infty) = \mathcal{R}_{\mathrm{fo}}(\infty) = \log_2(1 + \alpha + \beta)$ as $\tau \to \infty$. Due to the page limit, this analysis will be given in the full version of the paper; it is also available upon request.

REFERENCES

Baillieul, J. (2004). Data-rate requirements for nonlinear feedback control. In *Proc. 6th IFAC Symp. Nonlinear Control Syst.*, 1277–1282. Stuttgart, Germany.

Colonius, F. and Kawan, C. (2009). Invariance entropy for control systems. *SIAM J. Cont. Optim.*, 48(3), 1701–1721.

Colonius, F. and Kawan, C. (2011). Invariance entropy for outputs. *Math. Cont., Signals, and Sys.*, 22(3), 203–227.

Colonius, F., Kawan, C., and Nair, G. (2013). A note on topological feedback entropy and invariance entropy. *Systems & Control Letters*, 62(5), 377–381.

Donarowicz, T. (2011). *Entropy in Dynamical Systems*. Cambridge University Press, NY.

Franklin, J. (1993). *Matrix Theory*. Dover Publ.

Hagihara, R. and Nair, G. (2013). Two extensions of topological feedback entropy. *Math. of Cont., Signals, and Systems*, 25(4), 473–490.

Kawan, C. (2011). Invariance entropy of control sets. *SIAM Journal on Control and Optimization*, 49(2), 732–751.

Koiran, P. (2001). The topological entropy of iterated piecewise affine maps is uncomputable. *Discrete Mathematics and Theoretical Computer Science*, 4(2), 351–356.

Liberzon, D. and Hespanha, J.P. (2005). Stabilization of nonlinear systems with limited information feedback. *IEEE Transactions on Automatic Control*, 50(6), 910–915.

Manffra, E. (2002). *Properties of systems with time delayed feedback (Ph.D. Thesis)*. Univ. of Wuppertal.

Manffra, E., Kantz, H., and Just, W. (2001). Periodic orbits and topological entropy of delayed maps. *Phys Rev E Stat Nonlin Soft Matter Phys.*, 63, 046203.

Matveev, A.S. and Savkin, A.V. (2009). *Estimation and Control over Communication Networks*. Birkhäuser, Boston.

Matveev, A. and Pogromsky, A.Y. (2016). Observation of nonlinear systems via finite capacity channels: constructive data rate limits. *Automatica*, 70, 217–229.

Megretski, A. and Rantzer, A. (1997). System analysis via integral quadratic constraints. *IEEE Trans. Automatic Control*, 42(6), 819–830.

Nair, G.N., Evans, R.J., Mareels, I., and Moran, W. (2004). Topological feedback entropy and nonlinear stabilization. *IEEE Transactions on Automatic Control*, 49(9), 1585–1597.

Nair, G.N., Fagnani, F., Zampieri, S., and Evans, R.J. (2007). Feedback control under data rate constraints: an overview. *Proceedings of the IEEE*, 95(1), 108–137.

Pogromsky, A., Matveev, A., Chaillet, A., and Rüffer, B. (2013). Input-dependent stability analysis of systems with saturation in feedback. In *Proc. IEEE Conf. Decision and Control*, 5903–5908.

Pogromsky, A.Y. and Matveev, A.S. (2016a). Data rate limitations for observability of nonlinear systems. *IFAC-PapersOnLine*, 49(14), 119–124.

Pogromsky, A.Y. and Matveev, A. (2011). Estimation of topological entropy via the direct Lyapunov method. *Nonlinearity*, 24, 1937–1959.

Pogromsky, A.Y. and Matveev, A. (2016b). Stability analysis via averaging functions. *IEEE Transactions on Automatic Control*, 61(4), 1081–1086.

Pogromsky, A. and Matveev, A. (2013). A non-quadratic criterion for stability of forced oscillations. *Systems and Control Letters*, 62(5), 408–412.

Savkin, A.V. (2006). Analysis and synthesis of networked control systems: topological entropy, observability, robustness, and optimal control. *Automatica*, 42(1), 51–62.

De Persis, C. (2005). n-bit stabilization of n-dimensional nonlinear systems in feedforward form. *IEEE Trans. Aut. Cont.*, 50(3), 299–311.

De Persis, C. and Isidori, A. (2004). Stabilizability by state feedback implies stabilizability by encoded state feedback. *Syst. Control Lett.*, 53, 249–258.

Willems, J. (1972). Dissipative dynamical systems. Part I: General theory; Part II: Linear systems with quadratic supply rates. *Arch. Rational Mech. Analysis*, 45(5), 321–393.

Yakubovich, V. (2000). Necessity in quadratic criterion for absolute stability. *Int. J. Robust Nonlin. Control*, 10, 889–907.

Yakubovich, V. (2002). Popov's method and its subsequent development. *Eur. J. Control*, 8, 200–208.