POLITECNICO DI TORINO
Repository ISTITUZIONALE

Comprehending Complexity: Data-Rate Constraints in Large-Scale Networks

Original
Comprehending Complexity: Data-Rate Constraints in Large-Scale Networks / Matveev, Alexey S.; Proskurnikov, Anton V.; Pogromsky, Alexander; Fridman, Emilia. - In: IEEE TRANSACTIONS ON AUTOMATIC CONTROL. - ISSN 0018-9286. - 64:10(2019), pp. 4252-4259. [10.1109/TAC.2019.2894369]

Availability:
This version is available at: 11583/2756012 since: 2019-09-27T16:41:00Z

Publisher:
IEEE

Published
DOI:10.1109/TAC.2019.2894369

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright
IEEE postprint/Author's Accepted Manuscript

©2019 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collecting works, for resale or lists, or reuse of any copyrighted component of this work in other works.

(Article begins on next page)

12 October 2024

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2019.2894369, IEEE Transactions on Automatic Control

Comprehending complexity: Data-rate constraints in large-scale networks

Alexey S. Matveev, Anton V. Proskurnikov, Alexander Pogromsky, Emilia Fridman

***Abstract*—The paper is concerned with the rate at which a discrete-time, deterministic, and possibly large network of nonlinear systems generates information, and so with the minimum rate of data transfer under which the addressee can maintain the level of awareness about the current state of the network. While being aimed at development of tractable techniques for estimation of this rate, the paper advocates benefits from directly treating the dynamical system as a set of interacting subsystems. To this end, a novel estimation method is elaborated that is alike in flavor to the small gain theorem on input-to-output stability. The utility of this approach is demonstrated by rigorously justifying an experimentally discovered phenomenon: The topological entropy of nonlinear time-delay systems stays bounded as the delay grows without limits. This is extended on the studied observability rates and appended by constructive 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.**

***Index Terms*—Observability, Nonlinear systems, Entropy, Second Lyapunov method, Data rate estimates**

I. Introduction

A fundamental issue in the area of control of networked systems is about constraints on communication among the network agents. Some key aspects of such constraints are captured in the concept of communication channel with a limited data transmission bit-rate, and lead to inquiry about the minimal rate needed to attain a specific control objective. Recent extensive studies of this issue (see e.g., [1]–[6] and the literature therein) have shown that this data-rate threshold is alike in spirit to the topological entropy (TE) [7] of the system at hand, but is not always identical, and various relevant analogs of TE were introduced [2], [4], [8]–[17]. In effect, these thresholds evaluate the complexity of the system's temporal behavior by assessing the rate at which the system generates new information and so the minimum rate at which an observer must be supplied with data in order that its level of awareness about the network state can be maintained.

A. Pogromsky is with Department of Mechanical Engineering, Eindhoven University of Technology, Eindhoven, The Netherlands. He is also with Faculty of Control Systems and Robotics, Saint-Petersburg National Research University of Information Technologies Mechanics and Optics (ITMO), St. Petersburg, Russia (a.pogromsky@tue.nl).

A.S. Matveev is with Department of Mathematics and Mechanics, Saint Petersburg University, St. Petersburg, Russia. He is also with Faculty of Control Systems and Robotics, Saint-Petersburg National Research University of Information Technologies Mechanics and Optics (ITMO), St. Petersburg, Russia (almat1712@yahoo.com).

A.V. Proskurnikov is with Delft Center for Systems and Control, Delft University of Technology, Delft, The Netherlands. He is also with Institute for Problems of Mechanical Engineering of the Russian Academy of Sciences (IPME RAS). (anton.p.1982@ieee.org).

E. Fridman is with Department of Electrical Engineering and Systems, Tel Aviv University, Tel Aviv, Israel (emilia@eng.tau.ac.il).

Feasible computation or fine estimation of those thresholds is an intricate matter even for low dimensional nonlinear systems [18]–[20]. This intricacy rises fast at increase of the system's dimension, which typically makes that matter highly complex for even medium-scale nonlinear networks. Meanwhile, the thresholds proposed in [13], [14] were computed in closed form for several classic nonlinear chaotic systems (the bouncing ball system, Hénon, logistic, and Lozy maps [13], [21], among others) via the techniques elaborated in [13], [22]. They turn off the classic road of the first Lyapunov approach in study of TE and the likes towards his second method.

The goal of this paper is to develop the approach of [13], [14] into tractable techniques of handling networks of interconnected nonlinear dynamic agents with inputs and outputs. This focus is partly motivated by ubiquity of such interconnections, which are therefore a classic subject of study in control theory. The stated goal is to be approached by following the lines of the famous small-gain theorem on the input-to-output stability of a nonlinear plant (for generalizations concerned with networks, see, e.g., [23], [24]). To this end, we disclose individual input-to-output characteristics of the agents and relations among them that enable feasibly estimating the rate at which the entire network generates information. In this respect, the paper generalizes our preliminary results from [25].

To illustrate the utility of these developments, we use them to rigorously prove the fact previously discovered via numerical studies of a few particular chaotic delayed systems: their TE remains bounded as the delay grows without limits [26], [27]. We show that this phenomenon is common and extends on the studied observability rate thresholds. We also offer explicit upper bounds on them that are independent of the delay, and show that these bounds are asymptotically tight for a time-delay analog of the bouncing ball dynamics.

The paper is organized as follows. Section II presents background information. The problem setup and the main result are given in Sect. III. Section IV deals with time-delay systems, its findings are illustrated in Sect. V via an example.

The following notations are used throughout the paper:

- $[k_1 : k_2]$ denotes set of integers $j \in [k_1, k_2]$;
- $^\top$ stands for transposition;
- I_m is identity $m \times m$-matrix;
- $\|\cdot\|$ denotes the Euclidean norm of a vector $v \in \mathbb{R}^l$ and the operator norm of a matrix $M \in \mathbb{R}^{m\times l}$, i.e., $\|M\| := \max_{v\in\mathbb{R}^l:\|v\|=1} \|Mv\| = \max_{v\in\mathbb{R}^l:v\neq 0} \|Mv\|/\|v\|$ is the square root of the maximal eigenvalue of $M^\top M$;
- $\mathbf{stack}(p_i) \in \mathbb{R}^{r_1+\cdots+r_N}$ is the result of stacking vectors $p_i \in \mathbb{R}^{r_i}, i \in [1 : N]$ on top of one another.

0018-9286 (c) 2018 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2019.2894369, IEEE Transactions on Automatic Control

II. Observation via Bit-Rate Limited Channels

In this section, we introduce general concepts that are concerned with observation via finite capacity communication channels and will be employed in our main results.

A. Observation Problem Setup and Topological Entropy

This paper is concerned with building, in real time, an effective estimate of the current state $x(t) \in \mathbb{R}^n$ of a discrete-time invariant nonlinear system

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

The compact set K of feasible initial states and the continuous map $\phi : \mathbb{R}^n \to \mathbb{R}^n$ are known to the designer of an estimator. Data on (perfect) measurements of the state can reach the estimator only through a finite capacity communication channel. Per unit time, it can transmit only a finite part of the infinity of bits embodying the full knowledge of the current state. So at the estimator, the information about the state is inevitably inexact. Our main interest is in the case where due to unstable dynamics of the plant, this inaccuracy tends to grow as time passes, unless extra data arrive in course of time and are used to compensate for this growth. Success of such a compensation depends on the content of the messages and the transmission rate, with the latter being the main subject of our interest.

Thus, only a finite-bit message $e(t)$ can be sent via the channel at time t. So there is a need in a *coder* that converts sensor readings $x(t)$ into such messages. Based on prior messages, a *decoder* at time t produces an estimate $\widehat{x}(t) \in \mathbb{R}^n$ of the state $x(t)$. The coder and decoder form an *observer* and are described by the following respective equations:

$$\begin{aligned} e(t) &= \mathfrak{C}[t, x(0), \ldots, x(t) | \widehat{x}(0), \delta], \quad t \geqslant 0, \\ \widehat{x}(t) &= \mathfrak{D}\left[t, e(0), \ldots, e(t-1) | \widehat{x}(0), \delta\right], \quad t \geqslant 1. \end{aligned} \quad (2)$$

It is assumed here that both coder and decoder have access to a common initial estimate $\widehat{x}(0)$ and its accuracy δ

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

We borrow the concept of channel capacity from [6, Sect. 3.4] by assuming that no less/more than $b_-(r)/b_+(r)$ bits of data can be transferred across the channel within any time interval of duration r, and that the respective averaged rates are close to a common value c (*channel capacity*) for $r \approx \infty$:

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

As discussed in [6, Sect. 3.4], this model admits unsteady instant rates, transmission delays, and dropouts.

Definition 2.1 ([13]): The system (1) is said to be *observable via a given communication channel* if for any $\varepsilon > 0$, there exists $\delta(\varepsilon, K) > 0$ and an observer (2) that operates via the channel at hand and ensures $\|x(t) - \widehat{x}(t)\| \leqslant \varepsilon \ \forall t \geqslant 0$ whenever (3) holds with $\delta := \delta(\varepsilon, K)$, $x(0), \widehat{x}(0) \in K$.

The associated demand to the channel capacity is related to the *topological entropy* (TE) [7], [28] of the system (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 $a \in K$:

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

Specifically, the following claim holds.

Theorem 2.2 ([13]): For observability via a communication channel, it is necessary that its capacity $c \geqslant H(\phi, K)$. Conversely, if K is positively invariant, the system is observable via any channel with capacity $c > H(\phi, K)$.

B. Regular and Fine Observability

Definition 2.1 allows critical regress of the estimation accuracy over time: $\varepsilon \gg \delta(\varepsilon, K)$. This is excepted by the next definition: The accuracy stays proportional to its initial value.

Definition 2.3 ([13]): The observer (2) is said to *regularly observe the system* (1) if there exist δ_* and $G > 0$ such that the estimation accuracy $\|x(t) - \widehat{x}(t)\| \leqslant G\delta \ \forall t \geqslant 0$ 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 eventually restored and then exponentially improved.

Definition 2.4 ([13]): The observer (2) is said to *finely observe the plant* (1) if there are $\delta_*, G > 0, g \in (0, 1)$ such that $\|x(t) - \widehat{x}(t)\| \leqslant G\delta g^t \ \forall t \geqslant 0$ if $x(0), \widehat{x}(0) \in K, \delta < \delta_*$.

Definition 2.5 ([13]): The system (1) is said to be *regularly/finely observable via a given communication channel* if there exists an observer (2) that regularly/finely observes the system (1) and operates via the channel at hand.

What channel capacity c is needed for every kind of observability? Since the larger the capacity the better [13], this question in fact addresses the infimum $\mathcal{R}(\phi, K)$ of the needed c's. Here $\mathcal{R}$ is equipped with the index $_{\mathrm{o}}$, $_{\mathrm{ro}}$, $_{\mathrm{fo}}$ in the cases from Definitions 2.1, 2.3, 2.4, respectively, is called the *observability rate*, and is fully determined by the system (1).

Lemma 2.6 ([13]): For any positively invariant $\phi(K) \subset K$ compact set K, the following relations hold:

$$H(\phi, K) = \mathcal{R}_{\mathrm{o}}(\phi, K) \leqslant \mathcal{R}_{\mathrm{ro}}(\phi, K) = \mathcal{R}_{\mathrm{fo}}(\phi, K). \quad (7)$$

The results of, e.g., [6, Sect. 3.5] imply that for any linear $x(t+1) = Ax(t)$ system (1) and K with nonempty interior,

$$\begin{aligned} \mathcal{R}_{\mathrm{o}}(\phi, K) &= \mathcal{R}_{\mathrm{ro}}(\phi, K) = \mathcal{R}_{\mathrm{fo}}(\phi, K) \\ &= H(A) := \sum_{i=1}^{n} \log_2 \max\{|\lambda_i|; 1\}, \end{aligned} \quad (8)$$

where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A. In nonlinear case, computation or even fine estimation of the TE is an intricate matter [18]–[20] so that its exact value is unknown even for many prototypical low-dimensional chaotic systems, like the Hénon map, Dufing oscillator, or bouncing ball system.

Whereas positivity of TE is classically associated with chaotic behavior, there is not enough evidence that the positivity of the regular and fine observability rates can serve as an onset of chaos in general. Meanwhile, Lemma 2.6 shows that these rates give an upper bound on TE.

0018-9286 (c) 2018 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2019.2894369, IEEE Transactions on Automatic Control

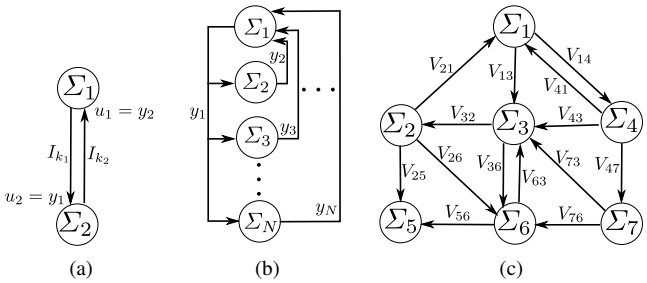


Fig. 1. Examples of networks of interconnected dynamical systems Σ_i.

C. Comprehending complexity of network temporal behavior

This paper is focused on systems (1) that represent networks of interconnected subsystems. Because of multiplicity and heterogeneity of aspects and factors that contribute to complexity, there exists a whole variety of notions of network complexity; see, e.g., [29]. We are concerned with complexity of the temporal behavior, which is understood in terms of predictability, chaoticity, and the likes. Then the values of either TE, or observability rates, or their upper estimates can serve as measures of complexity since they characterize the rate at which the network generates a new information.

Meanwhile, large dimensionality of networked systems carries a good potential to hamper direct application of available techniques for evaluation of these measures. A way to cope with this trouble is to disintegrate analysis into tractable portions in accordance with the network structure.

We focus on the techniques from [13], [22], [30]–[32], whose efficacy in the non-networked case has been proved by closed-form computation of the regular and fine observability rates for a number of prototypical nonlinear chaotic systems, e.g., the bouncing ball system, logistic map and, under certain circumstances, Hénon system [13], Lozy and Lorenz maps [33]. To acquire a tractable method of dealing with networked setting, we carry out the above disintegrated analysis along the avenue of [13], [22], [30]–[32] via the following steps:

1) The plant is directly treated as an interconnection of subsystems with inputs and outputs;
2) The individual input-to-output properties of the linearized subsystems are characterized via inequalities on quadratic "storage-" and "supply-" like functions in a fashion portrayed in, e.g., [34], [35];
3) The final data-rate estimate is built on an argument in the spirit of the celebrated small gain theorem.

III. Topological entropy and observability rates of a nonlinear networked system

A. Problem statement

From now on, we consider a network of interconnected discrete time invariant nonlinear systems Σ_i, labeled 1 through N; see Fig. 1. The ith system is described by the equations:

$$x_i(t+1) = \phi_i[x_i(t), u_i(t)], \quad y_i(t) = h_i[x_i(t)]. \tag{9}$$

Here $x_i \in \mathbb{R}^{n_i}$ is the state, $u_i \in \mathbb{R}^{m_i}$ is the system's input, and $y_i \in \mathbb{R}^{k_i}$ is the output. The interconnection is given by

$$u_i(t) = \sum_{j=1}^{N} V_{ij} y_j(t), \tag{10}$$

where the given $m_i \times k_j$-matrix V_{ij} quantifies the impact of jth subsystem on the ith one; $V_{ij} = 0$ in the case of no impact.

The above model admits "master" systems without an input

$$x_i(t+1) = \phi_i[x_i(t)], \quad y_i(t) = h_i[x_i(t)], \tag{11}$$

which influence the peers, being unaffected by them, as well as "slave" systems without an output

$$x_i(t+1) = \phi_i[x_i(t), u_i(t)], \tag{12}$$

which are influenced by the peers with no backward effect on them. To embed these cases into (9), it suffices to endow the "master" (11) with a "void" input $u_i(t) \in \mathbb{R}$ with no effect on the dynamics of Σ_i and put $V_{ij} = 0\,\forall j$ to set $u_i(t) \equiv 0$ for the sake of definiteness. Similarly, it suffices to endow the "slave" (12) with the "void" output map $h_i(x) \equiv 0 \in \mathbb{R}$.

The network at hand can be written in the form (1) with

$$x := \mathbf{stack}(x_i) \in \mathbb{R}^n, \qquad \phi(x) = \mathbf{stack}\left(\phi_i\left[x_i^{\leftarrow\rho}(x)\right]\right),$$
$$\text{where } x_i^{\leftarrow\rho}(x) := \left[x_i, \sum_{j=1}^{N} V_{ij} h_j(x_j)\right] \text{ and } n := \sum_{i=1}^{N} n_i. \tag{13}$$

We still consider only the trajectories that start in a given compact set $K \subset \mathbb{R}^n$. So the material of Sect. II-A and II-B is fully applicable to this network, which is attributed to the just introduced ϕ and K from now on. Our goal is to constructively estimate $\mathcal{R}_{\text{ro/fo}}(\phi, K)$ in this case.

B. Basic constructions and assumptions

These assumptions are distributed into three groups.

Assumptions about the interactions are motivated by the fact that direct use of the coupling matrices V_{ij}, which exhaustively describe the topology and strengths of interactions, may be troublesome. The reasons combine high dimensionality of the data represented by the totality of all V_{ij}'s with problems of their practical acquisition. So we admit that V_{ij}'s may be unavailable and only a less problematic upper estimate of the summary "strength" of actions on every subsystem is known.

Assumption 3.1: For any i, a bound M_i is known such that

$$\|V_{i*}\|^2 \leqslant M_i. \tag{14}$$

Here the block matrix $V_{i*} := [V_{i1}\, V_{i2} \ldots V_{iN}]$ of dimension $m_i \times (k_1 + \ldots + k_N)$ sets up actions of the peers on Σ_i.

There is a way to reduce the dimensions of the matrices whose operator norm $\|\cdot\|$ should be computed: it suffices to verify the inequality $\sum_{j=1}^{N} \|V_{ij}\|^2 \leqslant M_i$, which clearly implies (14). Certainly, this injects more conservatism in general.

Section III-E will discuss replacement of Assumption 3.1 by a less conservative though more involved requirement.

Assumptions about every subsystem Σ_i.

Assumption 3.2: In (9), the maps $\phi_i : \mathbb{R}^{n_i} \times \mathbb{R}^{m_i} \to \mathbb{R}^{n_i}$ and $h_i : \mathbb{R}^{n_i} \to \mathbb{R}^{k_i}$ are continuously differentiable for any i.

0018-9286 (c) 2018 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2019.2894369, IEEE Transactions on Automatic Control

In order to proceed, we denote by $x(t,a)$ the trajectory of the system (9), (10) that starts with $x(0)=a\in\mathbb{R}^n$, and put

$$X(t):=\{x=x(t,a):a\in K\},\quad X^\infty:=\bigcup_{t=0}^{\infty}X(t).\qquad(15)$$

Definition 3.3: A function $f(x)$ mapping $\mathbb{R}^n$ 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 $\delta>0$ providing the following:

$$\|f(x)-f(\chi)\|<\varepsilon\quad\forall x\in X_*,\chi\in\mathbb{R}^n,\text{ such that }\|x-\chi\|<\delta.$$

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

Assumption 3.4: For any subsystem Σ_i, the following functions (which are defined by using $x_i^{\leftarrow\rho}(x)$ from (13))

$$\frac{\partial\phi_i}{\partial x_i}\left[x_i^{\leftarrow\rho}(x)\right],\frac{\partial\phi_i}{\partial u_i}\left[x_i^{\leftarrow\rho}(x)\right],\frac{\partial h_i}{\partial x_i}[x_i]$$

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

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

Any trajectory of the networked system is associated with a particular process in every subsystem i, which is described by the time sequences $x_i(t),u_i(t),y_i(t),t=0,1,\ldots$. The further analysis will be much concerned with the first order approximation of every subsystem near a particular trajectory

$$z_i(t+1)=A_i(t)z_i(t)+B_i(t)w_i(t),\qquad \zeta_i(t)=C_i(t)z_i(t).\qquad(16)$$

Here z_i,w_i, and ζ_i stand for the "increments" of x_i,u_i, and y_i, respectively, and

$$A_i(t)=\frac{\partial\phi_i}{\partial x_i}[x_i(t),u_i(t)],\quad B_i(t)=\frac{\partial\phi_i}{\partial u_i}[x_i(t),u_i(t)],\quad C_i(t)=\frac{\partial h_i}{\partial x_i}[x_i(t)].\qquad(17)$$

A productive approach to characterization of input-to-output properties of linear systems is by using dissipation inequalities on certain quadratic "storage" and "supply" functions [34]. We follow these lines and associate the i-th subsystem with a "storage" $z_i^\top P_iz_i$ function and take the function giving the "supply" rate in the form $z_i^\top[Q_i-P_i]z_i-\frac{1}{\gamma_i}\|\zeta_i\|^2+\gamma_i\|w_i\|^2$. Description of the input-to-output properties addresses the incremental values of the input and output of the linearized system (16), and is stated in the following.

Assumption 3.5: There exist $n_i\times n_i$ matrices $P_i=P_i^\top>0$, $Q_i=Q_i^\top\geqslant 0$ and a number $\gamma_i>0$ such that the following inequality is true along all solutions of the networked system (9), (10) starting in the given compact set K:

$$[A_i(t)z_i+B_i(t)w_i]^\top P_i[A_i(t)z_i+B_i(t)w_i]\leqslant z_i^\top Q_iz_i-\frac{1}{\gamma_i}\|\zeta_i\|^2+\gamma_i\|w_i\|^2,\ \zeta_i=C_i(t)z_i\quad\forall z_i,w_i,t.\qquad(18)$$

If $Q_i\leqslant P_i$, this implies that γ_i upper bounds the l_2-gain of the system (16) from the input w_i to output ζ_i: for $z(0)=0$,

$$\sum_t\|\zeta_i(t)\|^2\leqslant\gamma_i^2\sum_t\|w_i(t)\|^2.$$

In this case, Assumption 3.5 gives upper bounds γ_i on the incremental l_2-gains of subsystems (9).

For the system (11) (which has no input), (18) shapes into

$$A_i(t)^\top P_iA_i(t)+\gamma_i^{-1}C_i(t)^\top C_i(t)\leqslant Q_i\quad\forall t.$$

For the system (12) (which has no output), (18) takes the form

$$[A_i(t)z_i+B_i(t)w_i]^\top P_i[A_i(t)z_i+B_i(t)w_i]\leqslant\leqslant z_i^\top Q_iz_i+\gamma_i\|w_i\|^2\quad\forall z_i,w_i,t.$$

Assumption on the balance between the strengths of interactions and the input-to-output gains of the subsystems. Whereas these strengths are assessed by the constants M_i from Assumption 3.1, the concerned "gains" are characterized by γ_i from Assumption 3.5. It is worth noting that the "gains" γ_i from this assumption are not uniquely determined like classical input-output gains: indeed, selecting a larger Q_i one can, in turn, choose a smaller γ_i. Our last assumption may be viewed as an analog of the classical "small gain" inequality.

Assumption 3.6: For any i, the following inequality holds:

$$\gamma_i\sum_{j:V_{ji}\neq 0}\gamma_jM_j\leqslant 1.\qquad(19)$$

It assumes knowledge of the interaction graph illustrated in Fig. 1c. If this graph is unknown, the following stronger condition can be verified since it surely implies Assumption 3.6:

$$\gamma_i\sum_{j=1}^{N}\gamma_jM_j\leqslant 1\qquad\forall i.\qquad(20)$$

In turns, this holds whenever $\gamma_i\leqslant(M_1+\cdots+M_N)^{-1/2}\ \forall i$.

C. The main result

Let $P=P^\top>0$ and $Q=Q^\top\geqslant 0$ be square matrices of a common size. The roots of the algebraic equation

$$\det(Q-\lambda P)=0\qquad(21)$$

are nonnegative (since $\lambda=\frac{x^\top Qx}{x^\top Px}$, where $x\neq 0$ is any solution of the singular linear equation $(Q-\lambda P)x=0$) and equal to the eigenvalues of each of the matrices QP^{-1} and $P^{-1}Q$. Let us enumerate these roots in descending order, repeating any of them in accordance with its algebraic multiplicity. Partly inspired by (8), we introduce the following quantity

$$H_{\mathrm{L}}(P,Q)=\frac{1}{2}\sum_j\max\{0,\log_2\lambda_j\},\qquad(22)$$

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

Now we are in a position to state the main result.

Theorem 3.7: Suppose that Assumptions 3.1—3.6 hold. Then the observability rates of the networked system (9), (10) obey the following inequalities

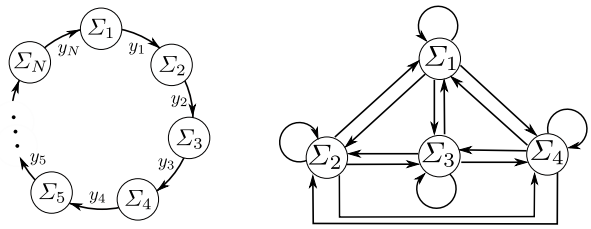
$$H(\phi,K)\leqslant\mathcal{R}_{\mathrm{ro}}(\phi,K)\leqslant\mathcal{R}_{\mathrm{fo}}(\phi,K)\leqslant\sum_{j=1}^{N}H_{\mathrm{L}}(P_j,Q_j),\qquad(23)$$

where the matrices P_i,Q_i are taken from Assumption 3.5.

The proof of this theorem is given in Appendix A.

0018-9286 (c) 2018 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2019.2894369, IEEE Transactions on Automatic Control



(a) Ring-like connection (b) All-to-all connection

Fig. 2. Examples of networks of interconnected dynamical systems Σ_i.

D. Examples: networks with special topologies

For networks with special topologies, we now specify Assumption 3.6, which is among the key conditions that imply the conclusion (23) of Theorem 3.7.

1) Feedback connection of two systems: is shown in Fig. 1a. In this case, $u_1 = \zeta_2 \in \mathbb{R}^{k_2}, u_2 = \zeta_1 \in \mathbb{R}^{k_1}$ and $V_{12} = I_{k_2}, V_{21} = I_{k_1}$, $V_{11} = 0, V_{22} = 0$. Thus $M_1 = M_2 = 1$ in (14), and Asm 3.6 shapes into $\gamma_1\gamma_2 \leqslant 1$. This is close to the key condition $\gamma_1\gamma_2 < 1$ from the classic small-gain theorem.

2) Feedback connection of N systems: is shown in Fig. 1b. In this case, $m_1 = k_2 + \ldots + k_N, m_2 = \ldots = m_N = k_1$,

$$u_1 = \mathbf{stack}(y_2, \ldots, y_N), \quad u_i = y_1 \quad \forall i \geqslant 2,$$
$$V_{1*} = \mathbf{diag}[0, I_{k_2}, \ldots, I_{k_N}], \quad V_{i*} = [I_{k_1}\, 0 \ldots 0] \quad \forall i \geqslant 2.$$

So $M_1 = \ldots = M_N = 1$ in (14) and (19) $\Leftrightarrow \gamma_1 \sum_{j=2}^N \gamma_j \leqslant 1$.

3) Ring-like connection: is illustrated in Fig. 2a. In this case, $m_2 = k_1, V_{21} = I_{k_1}, m_3 = k_2, V_{32} = I_{k_2}, \ldots m_N = k_{N-1}, V_{N,N-1} = I_{k_{N-1}}, m_1 = k_N, V_{1N} = I_{k_N}$, all other V_{ij}'s are zero, $M_1 = M_2 = \ldots = M_N = 1$ in (14). So (19) means that $\gamma_1\gamma_N \leqslant 1$ and $\gamma_j\gamma_{j+1} \leqslant 1$ whenever $1 \leqslant j \leqslant N-1$.

4) All-to-all connection via broadcasting communication: is illustrated in Fig. 2b. Every subsystem i affects its peers via broadcasting a signal $y_i(t)$. Any subsystem Σ_j averages the incoming signals via the classic nearest neighbors rule $N^{-1}\sum_i y_i(t)$ to form the input to the controller that drives Σ_j. So $m_i = k_i =: p\ \forall i$, $V_{ij} = N^{-1} I_p\ \forall i, j$, and hence

$$V_{i*} = N^{-1}[I_p\, I_p\, \ldots I_p], \quad M_i = \|V_{i*}\|^2 = N^{-1} \quad \forall i;$$
$$(20) \Leftrightarrow \gamma_i(\gamma_1 + \cdots + \gamma_N) \leqslant N \quad \forall i. \tag{24}$$

E. Relaxation of Assumptions 3.1 and 3.6

Theorem 3.7 remains true if a weaker assumption is put in place of the above two ones. The utility of this is somewhat subverted by more involved verification insomuch as the operator norm of a potentially much larger matrix is concerned.

To define it, we arrange V_{ij} from (10) into $m \times k$-matrix $V = (V_{ij})$, where $m := m_1 + \ldots + m_N$ and $k := k_1 + \ldots + k_N$, and introduce the following block-diagonal matrices

$$\Gamma_w = \mathbf{diag}(\sqrt{\gamma_1} I_{m_1}, \ldots, \sqrt{\gamma_N} I_{m_N}),$$
$$\Gamma_\zeta = \mathbf{diag}(\sqrt{\gamma_1} I_{k_1}, \ldots, \sqrt{\gamma_N} I_{k_N}). \tag{25}$$

Assumption 3.8: The following inequality holds

$$\|\Gamma_w V \Gamma_\zeta\| \leqslant 1. \tag{26}$$

Lemma 3.9: Theorem 3.7 remains true if Assumptions 3.1 and 3.6 are replaced with Assumption 3.8.

The proofs of this and next remark and the following lemma are given in Appendix A. This lemma shows that Lemma 3.9 does relax the assumptions of Theorem 3.7.

Lemma 3.10: Assumptions 3.1, 3.6 imply Assumption 3.8.

Meanwhile, the converse is not true in general. For example, in the case from Subsect. III-D-4,

$$\Gamma_w = \Gamma_\zeta = \mathbf{diag}\left[\sqrt{\gamma_1} I_p, \sqrt{\gamma_2} I_p, \ldots, \sqrt{\gamma_N} I_p\right]$$
$$V = \frac{1}{N}\begin{pmatrix} I_p & I_p & \ldots & I_p \\ I_p & I_p & \ldots & I_p \\ \vdots & \vdots & \vdots & \vdots \\ I_p & I_p & \ldots & I_p \end{pmatrix}$$
$$\Gamma_w V \Gamma_\zeta = \frac{1}{N}\begin{pmatrix} \sqrt{\gamma_1\gamma_1} I_p & \sqrt{\gamma_1\gamma_2} I_p & \ldots & \sqrt{\gamma_1\gamma_N} I_p \\ \sqrt{\gamma_2\gamma_1} I_p & \sqrt{\gamma_2\gamma_2} I_p & \ldots & \sqrt{\gamma_2\gamma_N} I_p \\ \vdots & \vdots & \vdots & \vdots \\ \sqrt{\gamma_N\gamma_1} I_p & \sqrt{\gamma_N\gamma_2} I_p & \ldots & \sqrt{\gamma_N\gamma_N} I_p \end{pmatrix};$$
$$\|\Gamma_w V \Gamma_\zeta \zeta\|^2 = \frac{1}{N^2}\left\|\sum_{j=1}^N \sqrt{\gamma_j}\zeta_j\right\|^2 \sum_{j=1}^N \gamma_j,$$
$$\|\Gamma_w V \Gamma_\zeta\|^2 = \frac{1}{N^2}\sum_{j=1}^N \gamma_j \max_{\zeta_j : \sum_j \|\zeta_j\|^2 = 1}\left\|\sum_{j=1}^N \sqrt{\gamma_j}\zeta_j\right\|^2$$
$$= \frac{1}{N^2}\left[\sum_{j=1}^N \gamma_j\right]^2.$$

So Assumption 3.8 takes the form $\frac{1}{N}\sum_{j=1}^N \gamma_j \leqslant 1$ and is much weaker than the set of inequalities (24) that embodies Assumptions 3.1 and 3.6 in the case at hand.

IV. Entropy of systems with delays

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

$$x(t+1) = f[x(t), Cx(t-\tau)], \quad t = 0, 1, \ldots. \tag{27}$$

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 $d \times n$-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

$$x(0) \in \mathcal{K}, x(-1) \in \mathcal{K}, \ldots, x(-\tau) \in \mathcal{K}. \tag{28}$$

The standard state augmentation

$$\mathfrak{x}(t) := [x(t), x(t-1), \ldots, x(t-\tau)] \tag{29}$$

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 II are fully applicable to (27), (28).

We study the behavior of the TE $H(\tau)$ and the observability rates $\mathcal{R}_{\text{ro/fo}}(\tau)$ of the system (27), (28) as $\tau \to \infty$. A stimulus for this is given by the numerical studies in [26], [27], which have shown that $\overline{\lim}_{\tau\to\infty} H(\tau) < \infty$ for particular chaotic systems. Now we rigorously prove that this phenomenon is common and extends on $\mathcal{R}_{\text{ro/fo}}(\tau)$, and offer explicit upper bounds on $\mathcal{R}_{\text{ro/fo}}(\tau), H(\tau)$ that are uniform over τ.

We impose the following analog of Assumption 3.4.

0018-9286 (c) 2018 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2019.2894369, IEEE Transactions on Automatic Control

Assumption 4.1: There is $\mathcal{K}_* \subset \mathbb{R}^n$ such that the following statements hold:

i) Irrespective of the delay, any solution of (27) satisfying (28) lies in $\mathcal{K}_*$, i.e., $x(t) \in \mathcal{K}_*\ \forall t \geqslant 0$;

ii) The first derivatives of $f(\cdot,\cdot)$ are bounded on $\mathcal{K}_* \times C\mathcal{K}_*$ and are 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 $\mathbb{R}^n \times \mathbb{R}^d$.

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 (18) with $\gamma_i := 1$.

Assumption 4.2: There are symmetric $n \times n$-matrices $P > 0, Q \geqslant 0$ such that for any $\varpi \in \mathcal{K}_* \times C\mathcal{K}_*$,

$$\begin{aligned}[A(\varpi)z + B(\varpi)w]^\top P[A(\varpi)z + B(\varpi)w] &\leqslant z^\top Qz \\ - \zeta^\top\zeta + w^\top w, \quad \zeta = Cz, &\quad \forall z \in \mathbb{R}^n, w \in \mathbb{R}^d. \quad (30)\end{aligned}$$

Thanks to ii) in Assumption 4.1, such matrices do exist: it suffices to pick P and Q small and large enough, respectively, although a more refined choice may be also possible.

Theorem 4.3: Let Assumptions 4.1 and 4.2 hold. Then

$$H(\tau) \leqslant \mathcal{R}_{\text{ro}}(\tau) \leqslant \mathcal{R}_{\text{fo}}(\tau) \leqslant H_L(P, Q) \qquad \forall \tau.$$

Proof: We represent the system (27) as the interconnection (10) of the following two subsystems with $u_i(t) \in \mathbb{R}^d$

$$\Sigma_1 \ : \ \left\{ \begin{array}{rcl} x_1(t+1) & = & f[x_1(t), u_1(t)] \in \mathbb{R}^n, \\ y_1(t) & = & Cx_1(t) \in \mathbb{R}^d, \end{array} \right.$$

$$\Sigma_2 \ : \ \left\{ \begin{array}{rcl} x_2(t+1) & = & \mathfrak{A}\, x_2(t) + \mathfrak{B}\, u_2(t) \in \mathbb{R}^{d\tau}, \\ y_2(t) & = & \mathfrak{C}\, x_2(t) \in \mathbb{R}^d. \end{array} \right.$$

Here the second subsystem is a τ-step delay line:

$$\mathfrak{A} = \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 \\ I_d & 0 & \cdots & 0 & 0 \\ 0 & I_d & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & 0 & 0 \\ 0 & 0 & \cdots & I_d & 0 \end{pmatrix}, \mathfrak{B} = \begin{pmatrix} I_d \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

$$\mathfrak{C} = \begin{pmatrix} 0 & 0 & \cdots & 0 & I_d \end{pmatrix}.$$

For Σ_1, Asm 3.5 holds with $P_1 := P, Q_1 := Q, \gamma_1 := 1$ by i) in Asm 4.1 and Asm 4.2. For Σ_2, we have

$$x_2(t+1)^\top x_2(t+1) = x_2(t)^\top x_2(t) + u_2(t)^\top u_2(t) - y_2(t)^\top y_2(t);$$

so Asm 3.5 is true with $P_2 = Q_2 = I_{d\tau}, \gamma_2 = 1$. Thus Asm 3.6 holds; Asm 4.1 implies Asm 3.4. The proof is completed by Theorem 3.7 since $H_L(P_2, Q_2) = 0$. ■

The last equation is an epitome of the fact that the delay line Σ_2 does not produce uncertainty. An easily visible sign of this is that the knowledge $\mathcal{X}^\triangledown = [x^\triangledown_{t_0}, \ldots, x^\triangledown_{t_0-\tau}]$ of the state $\mathcal{X}(t_0) := [x(t_0), \ldots, x(t_0 - \tau)]$ of Σ_2 (with $u_2(t) \equiv 0$ for simplicity) up to the δ-uncertainty $\max_{t=t_0-\tau,\ldots,t_0} \|x(t) - x^\triangledown_t\| < \delta$ enables one to predict the subsequent states with the same accuracy δ (e.g., by running Σ_2 from $\mathcal{X}^\triangledown$). Meanwhile, the delay line is able to affect uncertainty production in a feedback interconnection so that the entropy of the overall system becomes dependent on the delay [26], [27].

V. Example

We consider an integer delay $\tau > 0$ and the τ-delayed analog of the "bouncing-ball dynamics" [13] (which is among the classic examples of low-dimensional chaotic behavior):

$$y(t+1) = (1+\alpha)y(t) - \beta \cos y(t) - \alpha y(t-\tau) \in \mathbb{R}. \quad (31)$$

Here α and $\beta > 0$ are parameters. Since (31) is invariant to the change $y \mapsto y \pm 2\pi$, this equation defines not only a dynamical system in $\mathbb{R}$ ($\mathbb{R}$-system) but also a system in the unit circle S^1_0 (S^1_0-system). By [13, Remark 5], the concepts from Section II are fully applicable to the S^1_0-system.

Proposition 5.1: For the S^1_0-system (31) with any delay τ,

$$H(\tau) \leqslant \mathcal{R}_{\text{ro}}(\tau) \leqslant \mathcal{R}_{\text{fo}}(\tau) \leqslant \mathcal{L} := \log_2(1 + 2\alpha + \beta). \quad (32)$$

Proof: The $\mathbb{R}$-system (31) has the form (27), (28) with

$$x = y \in \mathbb{R}, C = 1, f(x', x'') = (1+\alpha)x' - \beta \cos x' - \alpha x''$$

in (27) and $\mathcal{K} := [-\pi, \pi]$ in (28). Then Assumption 4.1 holds with $\mathcal{K}_* := \mathbb{R}$. To check Assumption 4.2, we note that in (30),

$$A(\varpi)z + B(\varpi)w = (1 + \alpha + \beta \sin y)z - \alpha w \quad \forall \varpi = (y, y_1).$$

Now $P = p \in (0, \infty), Q = q \in [0, \infty)$ and $z^\top Pz = pz^2, z^\top Qz = qz^2$. So the left-hand side L of (30)

$$L = p|(1 + \alpha + \beta \sin y)z - \alpha w|^2 \leqslant p[\gamma|z| + \alpha|w|]^2,$$

where $\gamma := 1 + \alpha + \beta$. Hence (30) does hold whenever

$$L_1 := (q-1)|z|^2 + |w|^2 - p[\gamma|z| + \alpha|w|]^2 \geqslant 0.$$

By treating L_1 as a quadratic form in $|z|, |w|$ and applying Sylvester's criterion, we see that (30) is true whenever

$$\begin{aligned} 1 - p\alpha^2 > 0 \text{ and } 0 \leqslant &\begin{vmatrix} q - 1 - p\gamma^2 & -p\gamma\alpha \\ -p\gamma\alpha & 1 - p\alpha^2 \end{vmatrix} \\ &= q(1 - p\alpha^2) - [1 + p(\gamma^2 - \alpha^2)]. \end{aligned}$$

Thus (30) holds with $p \in \left(0, \alpha^{-2}\right), q := \frac{1+p(\gamma^2-\alpha^2)}{1-p\alpha^2}$. We take the point $p = \frac{1}{\alpha(\gamma+\alpha)} \in (0, \alpha^{-2})$, where the root $\lambda = q/p$ of (21) attains its maximum $(\alpha+\gamma)^2$. Via elementary computation of (22), Theorem 4.3 implies (32) for the $\mathbb{R}$-system.

By (29), its phase space $\mathfrak{X}_\mathbb{R}$ is $\mathbb{R}^{\tau+1}$; such space of the S^1_0-system is the multidimensional torus $\mathfrak{X}_{S^1_0} = \{[s_0, \ldots, s_{-\tau}] : s_{-j} \in S^1_0\ \forall j\}$. The covering projection of $[x_0, \ldots, x_{-\tau}] \in \mathfrak{X}_\mathbb{R}$ into $[e^{\imath x_0}, \ldots, e^{\imath x_{-\tau}}] \in \mathfrak{X}_{S^1_0}$ clearly maps the next state of the $\mathbb{R}$-system into the state of the S^1_0-system that is next to the projection of the current state of the $\mathbb{R}$-system. It remains to note that the change of the $\mathbb{R}$-system to S^1_0-system does not increase the data rates in (32) by [13, Lemma 13]. ■

For $\tau = 0$, Theorem 15 from [13] offers an exact formula for $\mathcal{R}_{\text{ro/fo}}(0)$ and displays a gap in (32): $\mathcal{R}_{\text{ro}}(0) = \mathcal{R}_{\text{fo}}(0) < \mathcal{L}$. For $\tau > 0$, computation of $\mathcal{R}_{\text{ro/fo}}(\tau)$ is not easy. So we analyze that gap in the limit as $\tau \to \infty$ via lower estimating the associated limit of $\mathcal{R}_{\text{ro}}(\tau)$. To this end, we need the following.

Lemma 5.2: Whenever $\alpha > 0$ and $a > 1 + \alpha$, the equation

$$\chi_n(\lambda) = \lambda^n - a\lambda^{n-1} + \alpha = 0 \quad (33)$$

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

0018-9286 (c) 2018 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2019.2894369, IEEE Transactions on Automatic Control

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

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

It is easy to see that the eigenvalues of the Jacobian matrix of (31) at the equilibrium $\pi/2$ are the roots of (33) with $a := 1+\alpha+\beta$, $n := \tau+1$. One of them converges to a as $\tau \to \infty$ by Lemma 5.2, whereas the others do not lie on S^1_0 (since $|\lambda| = 1 \Rightarrow |\lambda^{n-1}(a-\lambda)| \geqslant a-1 > \alpha$). So the equilibrium is hyperbolic. Then Theorem 9 in [13] entails the following.

Corollary 5.3: The following holds for any $\alpha, \beta > 0$:

$$\liminf_{\tau\to\infty} \mathcal{R}_{\text{ro}}(\tau) \geqslant \log_2(1+\alpha+\beta). \tag{34}$$

Thus for $\alpha \ll \beta$, the estimates (32) become tight as $\tau \to \infty$. The smaller the ratio $\alpha/(\beta+1)$, the narrower the gap between the upper (32) and lower (34) bounds on $\mathcal{R}_{\text{ro/fo}}$.

VI. Conclusions and Future Work

For a discrete-time deterministic network of interacting nonlinear systems, an upper bound on the bit-rate at which the network generates information was given. This bound is based on separate estimates of the individual contributions of the subsystems and their integration in line with the network topology. The obtained results were used to show that the topological entropy of nonlinear delayed systems, as well as the above rate, stays bounded as the delay grows without limits. A delay-independent upper estimate of these quantities was provided; this estimate is shown to be asymptotically tight for a time-delay analog of the bouncing ball dynamics.

Future work includes study of decentralized observation schemes and systems with disturbances.

Acknowledgement

A. Matveev and research in Sects. II, III.A-C, and IV were supported by the Russian Science Foundation under project 14-21-00041p. Results of Sect. III.E were obtained by A.V. Proskurnikov under sole support of Russian Science Foundation project 16-19-00057. Results of Sect. V were partially supported by Russian Foundation for Basic Research grant 18-38-20037. A. Pogromsky acknowledges his support from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 675080 (UCoCoS). E. Fridman acknowledges her support from Israel Science Foundation grant 1128/14.

Appendix A
Proofs of the results from Section III.

These proofs use Theorem 12 from [13], which is reproduced here for the convenience of the reader.

Theorem A.1 ([13]): Suppose that for the system (1), the Jacobian matrix $A(x) := \phi'(x)$ is bounded on the set X^∞ given by (15) and is uniformly continuous near this set. Let there exist continuous and bounded on X^∞ functions $v_d : \mathbb{R}^n \to \mathbb{R}$, constants $\Lambda_d \geqslant 0$, $d \in [1:n]$, and a positive definite $n\times n$-matrix $P = P^\top$ such that for any $d \in [1:n]$,

$$v_d[\phi(x)] - v_d[x] + \sum_{i=1}^{d} \log_2 \lambda_i(x) \leqslant \Lambda_d \quad \forall x \in X^\infty. \tag{A.1}$$

Here $\lambda_1(x) \geqslant \cdots \geqslant \lambda_n(x) \geqslant 0$ are the roots of the algebraic equation

$$\det\left[A(x)^\top P A(x) - \lambda P\right] = 0 \tag{A.2}$$

are repeated in accordance with their multiplicities. Then

$$\mathcal{R}_{\text{fo}}(\phi, K) \leqslant \Lambda_\star := 2^{-1}\max_d \Lambda_d. \tag{A.3}$$

We also need the following

Lemma A.2: If $P > 0$, Q_1, Q_2 are symmetric $n\times n$-matrices, then $H_{\text{L}}(P, Q_1) \leqslant H_{\text{L}}(P, Q_2)$ whenever $0 \leqslant Q_1 \leqslant Q_2$.

Proof: We put $S := P^{-1/2}$, where $P^{1/2}$ is the positively definite square root of P, and note that (21) means that λ is an eigenvalue $\det(SQS - \lambda I) = 0$ of the symmetric matrix SQS. Meanwhile, $x^\top SQ_iSx = (Sx)^\top Q_i(Sx)$ and so $0 \leqslant SQ_1S \leqslant SQ_2S$. The proof is completed by Weyl's inequality for eigenvalues of symmetric matrices [36, Cor. 4.3.3]. $\blacksquare$

Proof of Lemma 3.10: It is easy to see that

$$\|u_i\|^2 \overset{(10)}{=} \|V_{i*}y\|^2 \overset{(14)}{\leqslant} M_i\|y\|^2 = M_i\sum_{j=1}^{N}\|y_j\|^2, \tag{A.4}$$

where $u := \mathbf{stack}(u_i), y := \mathbf{stack}(y_i)$. Putting $y_j := 0$ whenever $V_{ij} = 0$ keeps (10) and so (A.4) true. Hence

$$\begin{aligned}
\|u_i\|^2 &\overset{(A.4)}{\leqslant} M_i \sum_{j:V_{ij}\neq 0}\|y_j\|^2 \\
&\xRightarrow{\times\gamma_i} \sum_{i=1}^{N}\gamma_i\|u_i\|^2 \leqslant \sum_{i=1}^{N}\gamma_i M_i \sum_{j:V_{ij}\neq 0}\|y_j\|^2 \\
&= \sum_{i,j\in[1:N]:V_{ij}\neq 0}\gamma_i M_i\|y_j\|^2 \leqslant \sum_{j=1}^{N}\|y_j\|^2 \sum_{i:V_{ij}\neq 0}\gamma_i M_i.
\end{aligned}$$

Now we put $y_j := \gamma_j^{1/2}\zeta_j$ here, denote $w_i := \gamma_i^{1/2}u_i$ and invoke (19) to see that

$$\sum_{i=1}^{N}\|w_i\|^2 \leqslant \sum_{j=1}^{N}\|\zeta_j\|^2 \underbrace{\gamma_j \sum_{i:V_{ij}\neq 0}\gamma_i M_i}_{\leqslant 1 \text{ due to (19)}} \leqslant \sum_{j=1}^{N}\|\zeta_j\|^2.$$

If $w \in \mathbb{R}^m, \zeta \in \mathbb{R}^k$ and $w = \Gamma_w V \Gamma_\zeta \zeta$, the vectors $u = \Gamma_w^{-1}w$ and $y = \Gamma_\zeta\zeta$ are related by $u = Vy$, which is equivalent to (10). So $\|\Gamma_w V\Gamma_\zeta\zeta\|^2 \leqslant \|\zeta\|^2\ \forall \zeta \in \mathbb{R}^k \Rightarrow$ (26). $\blacksquare$

Proof of Theorem 3.7 and Lemma 3.9: The first two inequalities in (23) are borrowed from Theorem 8 in [13] and formula (7) in [13]. So it remains to prove the third inequality. In view of Lemma 3.10, it suffices to prove Lemma 3.9.

Assumption 3.4 and (13) guarantee that $A(x) := \phi'(x)$ is bounded on X^∞ from (15) and is uniformly continuous near this set, as is required by Theorem A.1. Also, (13) yields that

$$A(x)z = \mathbf{stack}\left[A_i(x)z_i + B_i(x)w_i\right] \qquad \forall z := \mathbf{stack}\left(z_i\right),$$

$$\text{where } w_i = \sum_{j=1}^{N} V_{ij}\zeta_j, \quad \zeta_j = C_j(x)z_j, \text{ and} \tag{A.5}$$

$$A_i(x) := \frac{\partial\phi_i}{\partial x_i}[x_i^{\leftarrow\rho}(x)], \qquad B_i(x) := \frac{\partial\phi_i}{\partial u_i}[x_i^{\leftarrow\rho}(x)],$$

$$C_i(x) = \frac{\partial h_i}{\partial x_i}[x_i].$$

0018-9286 (c) 2018 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2019.2894369, IEEE Transactions on Automatic Control

Now we introduce the following positively and non-negatively definite block-diagonal $n \times n$-matrices, respectively

$$P = \mathbf{diag}(P_1, \ldots, P_N), \quad Q = \mathbf{diag}(Q_1, \ldots, Q_N). \quad \text{(A.6)}$$

For any $x \in X^\infty$, the vector $x_i^{\leftrightarrow\rho}(x)$ has the form $[x_i(t), u_i(t)]$ for some $t \geqslant 0$ and some trajectory of the networked system that starts in K. Hence, $A_i(x), B_i(x), C_i(x)$ coincide with matrices (17) that satisfy inequality (18), and so

$$\begin{aligned} & z^\top A(x)^\top P A(x) z \\ & \quad = \sum_{j=1}^{N} [A_i(t) z_i + B_i(t) w_i]^\top P_i [A_i(t) z_i + B_i(t) w_i] \\ & \quad \overset{(18)}{\leqslant} \sum_{j=1}^{N} \left[z_i^\top Q_i z_i - \frac{1}{\gamma_i} \|\zeta_i\|^2 + \gamma_i \|w_i\|^2 \right] \\ & \quad \overset{(25)}{=} z^\top Q z - \left\| \Gamma_\zeta^{-1} \zeta \right\|^2 + \|\Gamma_w w\|^2 . \end{aligned}$$

Here $\zeta := \mathbf{stack}(\zeta_i)$ and $w := \mathbf{stack}(w_i)$ are related by $w = V\zeta$ due to the first relation in (A.5). Whence

$$\begin{aligned} & z^\top A(x)^\top P A(x) z \leqslant z^\top Q z - \left\| \Gamma_\zeta^{-1} \zeta \right\|^2 + \left\| \Gamma_w V \Gamma_\zeta \Gamma_\zeta^{-1} \zeta \right\|^2 \\ & \leqslant z^\top Q z - \left[1 - \|\Gamma_w V \Gamma_\zeta\|^2 \right] \left\| \Gamma_\zeta^{-1} \zeta \right\|^2 \overset{\text{Ass. 3.8}}{\leqslant} z^\top Q z \quad \forall z. \end{aligned}$$

So by Lemma A.2, $H_{\mathrm{L}}[PA(x)^\top PA(x)] \leqslant H_{\mathrm{L}}(P, Q)$ $\forall x \in X^\infty$. So (A.1) holds with $v_d(\cdot) \equiv 0$ and $\Lambda_d := 2H_{\mathrm{L}}(P, Q)$ $\forall d$, whence $\mathcal{R}_{\mathrm{fo}}(\phi, K) \leqslant H_{\mathrm{L}}(P, Q)$ by Theorem A.1. Meanwhile, (A.6) implies that the roots of equation (21) are formed via the union of the sets of the roots of all equations $\det(Q_j - \lambda P_j), j \in [1 : N]$. So $H_{\mathrm{L}}(P, Q) = \sum_{j=1}^{N} H_{\mathrm{L}}(P_j, Q_j)$ by (22), which completes the proof of the third inequality in (23). ■

References

[1] C. De Persis and A. Isidori, "Stabilizability by state feedback implies stabilizability by encoded state feedback," *Syst. Control Lett.*, vol. 53, pp. 249–258, 2004.

[2] G. N. Nair, R. J. Evans, I. Mareels, and W. Moran, "Topological feedback entropy and nonlinear stabilization," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1585–1597, Sep. 2004.

[3] D. Liberzon and J. P. Hespanha, "Stabilization of nonlinear systems with limited information feedback," *IEEE Transactions on Automatic Control*, vol. 50, no. 6, pp. 910–915, Jun. 2005.

[4] 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, 2006.

[5] G. N. Nair, F. Fagnani, S. Zampieri, and R. J. Evans, "Feedback control under data rate constraints: an overview," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 108–137, 2007.

[6] A. S. Matveev and A. V. Savkin, *Estimation and Control over Communication Networks*. Boston: Birkhäuser, 2009.

[7] T. Downarowicz, *Entropy in Dynamical Systems*. NY: Cambridge University Press, 2011.

[8] F. Colonius and C. Kawan, "Invariance entropy for control systems," *SIAM J. Cont. Optim.*, vol. 48, no. 3, pp. 1701–1721, 2009.

[9] ——, "Invariance entropy for outputs," *Math. Cont., Signals, and Sys.*, vol. 22, no. 3, pp. 203–227, 2011.

[10] C. Kawan, "Invariance entropy of control sets," *SIAM Journal on Control and Optimization*, vol. 49, no. 2, pp. 732–751, 2011.

[11] R. Hagihara and G. Nair, "Two extensions of topological feedback entropy," *Math. of Cont., Signals, and Systems*, vol. 25, no. 4, pp. 473–490, 2013.

[12] F. Colonius, C. Kawan, and G. Nair, "A note on topological feedback entropy and invariance entropy," *Systems & Control Letters*, vol. 62, no. 5, pp. 377–381, 2013.

[13] A. Matveev and A. Y. Pogromsky, "Observation of nonlinear systems via finite capacity channels: constructive data rate limits," *Automatica*, vol. 70, pp. 217–229, 2016.

[14] A. Matveev and A. Pogromsky, "Observation of nonlinear systems via finite capacity channels. Part ii: Restoration entropy and its estimates," *Automatica*, accepted, 2019.

[15] D. Liberzon and S. Mitra, "Entropy and minimal data rates for state estimation and model detection," in *Proceedings of the 19th International Conference on Hybrid Systems: Computation and Control*, Vienna, Austria, April 2016, pp. 247–256.

[16] H. Sibai and S. Mitra, "Optimal data rate for state estimation of switched nonlinear systems," in *Proceedings of the 20th International Conference on Hybrid Systems: Computation and Control*, Pittsburgh, PA, April 2017, pp. 71–80.

[17] C. Kawan and J.-C. Delvenne, "Network entropy and data rates required for networked control," *IEEE Transactions on Control of Network Systems*, vol. 3, no. 1, pp. 57–66, 2016.

[18] J. Simonsen, "On the computability of the topological entropy of subshifts," *Discrete Math. and Theor. Comp. Science*, vol. 8, pp. 83–96, 2006.

[19] J. Delvenne and V. Blondel, "Quasi-periodic configurations and undecidable dynamics for tilings, infinite words and Turing machines," *Theoretical Computer Science*, vol. 319, pp. 127–143, 2004.

[20] P. Koiran, "The topological entropy of iterated piecewise affine maps is uncomputable," *Discrete Mathematics and Theoretical Computer Science*, vol. 4, no. 2, pp. 351–356, 2001.

[21] A. Y. Pogromsky and A. S. Matveev, "Data rate limitations for observability of nonlinear systems," *IFAC-PapersOnLine*, vol. 49, no. 14, pp. 119–124, 2016.

[22] A. Y. Pogromsky and A. Matveev, "Estimation of topological entropy via the direct Lyapunov method," *Nonlinearity*, vol. 24, pp. 1937–1959, 2011.

[23] S. Dashkovskiy, B. Rüffer, and F. Wirth, "Small gain theorems for large scale systems and construction of iss lyapunov functions," *SIAM J Control Optim*, vol. 48, no. 6, pp. 4089–4118, 2010.

[24] R. Geiselhart, M. Lazar, and F. Wirth, "A relaxed small-gain theorem for interconnected discrete-time systems," *IEEE Transactions Aut. Contr.*, vol. 60, no. 3, pp. 812–817, 2015.

[25] A. Pogromsky, A. Matveev, A. Proskurnikov, and E. Fridman, "A small-gain-theorem-like approach to nonlinear observability via finite capacity channels," *IFAC PapersOnLine*, vol. 50, pp. 15 397–15 402, 2017.

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

[27] E. Manffra, H. Kantz, and W. Just, "Periodic orbits and topological entropy of delayed maps," *Phys Rev E Stat Nonlin Soft Matter Phys.*, vol. 63, p. 046203, 2001.

[28] A. Katok, "Fifty years of entropy in dynamics: 1958–2007," *Journal of Modern Dynamics*, vol. 1, no. 4, pp. 545–596, 2007.

[29] S. Lloyd, "Measures of complexity: a nonexhaustive list," *IEEE Control Systems Magazine*, vol. 21, no. 4, pp. 7–8, 2001.

[30] A. Pogromsky, A. Matveev, A. Chaillet, and B. Rüffer, "Input-dependent stability analysis of systems with saturation in feedback," in *Proc. IEEE Conf. Decision and Control*, 2013, pp. 5903–5908.

[31] A. Pogromsky and A. Matveev, "A non-quadratic criterion for stability of forced oscillations," *Systems and Control Letters*, vol. 62, no. 5, pp. 408–412, 2013.

[32] A. Y. Pogromsky and A. Matveev, "Stability analysis via averaging functions," *IEEE Transactions on Automatic Control*, vol. 61, no. 4, pp. 1081–1086, 2016.

[33] A. Pogromsky and A. Matveev, "Data rate limitations for observability of nonlinear systems," *IFAC-PapersOnLine*, vol. 49, no. 14, pp. 119–124, 2016.

[34] J. Willems, "Dissipative dynamical systems – Part I: General theory; Part II: Linear systems with quadratic supply rates," *Archive for Rational Mechanics and Analysis*, vol. 45, no. 5, pp. 321–393, 1972.

[35] A. Megretski and A. Rantzer, "System analysis via integral quadratic constraints," *IEEE Trans. Automatic Control*, vol. 42, no. 6, pp. 819–830, 1997.

[36] R. Horn and C. Johnson, *Matrix Analysis*. Cambridge, UK: Cambridge University Press, 1990.

0018-9286 (c) 2018 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.