Automatica 103 (2019) 189–199

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Observation of nonlinear systems via finite capacity channels, Part II: Restoration entropy and its estimates ☆

Alexey S. Matveev [a,c,*], Alexander Yu. Pogromsky [b,c]

[a] Department of Mathematics and Mechanics, Saint Petersburg University, St. Petersburg, Russia
[b] Department of Mechanical Engineering, Eindhoven University of Technology, Eindhoven, The Netherlands
[c] Department of Control Systems and Informatics, Saint-Petersburg National Research University of Information Technologies, Mechanics, and Optics (ITMO), Russia

ARTICLE INFO

Article history:
Received 1 November 2017
Received in revised form 29 October 2018
Accepted 21 December 2018
Available online 19 February 2019

Keywords:
Entropy
Nonlinear systems
Observability
Reliable state estimation
Lyapunov function

ABSTRACT

The paper deals with the state estimation problem for nonlinear dynamical systems via communication channels with limited data rate. We introduce several minimum data-rate limits associated with various types of observability. A notion of the restoration entropy (RE) is also introduced and its relevance to the problem is outlined by a corresponding Data Rate Theorem. Theoretical lower and upper estimates for the RE are proposed in the spirit of the first and second Lyapunov methods, respectively. For three classic chaotic multi-dimensional systems, it is demonstrated that the lower and upper estimates for the RE coincide for one of them and are nearly the same for the others.

1. Introduction

The multidisciplinary area of networked control systems lies at the crossroad of control, communication, and computer sciences and integrates their classical topics into a whole, while responding to new special challenges born out of their union. One of them is caused by bottlenecks in the process of information transmission among the network nodes that may be due to constraints on the data transmission bit-rate allocated to every particular transmitter/receiver pair within a shared fieldbus device (Ge, Yang, & Han, 2017; Gupta & Chow, 2010). Regarding this issue, a substantial research effort was to reveal the minimal requirements to the rate and quality of data transmission that are needed for efficient, stable, and robust control (Antsaklis & Baillieul, 2007; Colonius, 2012; Colonius & Kawan, 2009, 2011; Hagihara & Nair, 2013; Kawan, 2011a; Liberzon & Hespanha, 2005; Mahmoud, 2014; Matveev & Savkin, 2009; Nair, Evans, Mareels, & Moran, 2004; Nair, Fagnani, Zampieri, & Evans, 2007; de Persis, 2005; Postoyan, van de Wouw, Nešić, & Heemels, 2014; Savkin, 2006; Yüksel & Basar, 2013).

An essential focus in these regards was on building a reliable realtime estimate of the state of a dynamical plant based on data communicated across a bit-rate constrained channel, with leaving purely control issues behind the scenes. This does not come as a surprise not only since a reliable state observer is often of value in its own right but also since many control problems can be routinely solved if such an observer is available. Moreover, the very concept of a networked feedback assumes that in the circular process of inter-influence, the sensory data get to the decision-maker/regulator via a communication channel, which inevitably makes consistency between data at its transmitter and receiver ends, respectively, a real concern for control.

It is this observability problem that is addressed: What is the minimal bit-rate of data transfer that is needed to build a reliable and efficient state estimate. This work is a continuation and extension of our previous work (Matveev & Pogromsky, 2016; Pogromsky & Matveev, 2016a), where the raised issues are addressed in the context of discrete time. The current paper is focused on continuous time case and deals with the concerns discussed below.

As is now known, the issue of the communication bit-rate minimally required for solvability of various control and observation tasks is conceptually and computationally related to the classic concept of the topological entropy (TE) (Downarowicz, 2011; Katok, 2007) and its recent control-oriented analogs (Colonius, 2012; Colonius & Kawan, 2009, 2011; Colonius, Kawan, & Nair, 2013;

☆ A.S. Matveev and research in Sects. 2–6 were supported by the Russian Science Foundation under the Project No 14-21-00041p. A. Pogromsky acknowledges his support by the UCoCoS project (EU Horizon 2020 programme under the Marie Skłodowska-Curie grant agreement No 675080). The material in this paper was partially presented at the 20th World Congress of the International Federation of Automatic Control, July 9–14, 2017, Toulouse, France (Matveev and Pogromsky (2017)). This paper was recommended for publication in revised form by Associate Editor Romain Postoyan under the direction of Editor Daniel Liberzon.

* Corresponding author at: Department of Mathematics and Mechanics, Saint Petersburg University, St. Petersburg, Russia.
E-mail addresses: almat@am1540.spb.edu (A.S. Matveev), A.Pogromsky@tue.nl (A.Y. Pogromsky).

https://doi.org/10.1016/j.automatica.2019.01.019
0005-1098/

Hagihara & Nair, 2013; Kawan, 2011a; Liberzon & Mitra, 2016; Nair et al., 2004; Savkin, 2006; Sibai & Mitra, 2017). Meanwhile, computation or even estimation of the TE of nonlinear systems has earned the reputation of an intricate matter; this intricacy briskly grows up with the system's dimension. Constructive lower and upper bounds on recent descendants of the TE are given in Colonius (2012), Kawan, (2011a, 2011b, 2011c), Liberzon and Mitra (2016) and Sibai and Mitra (2017); their conservatism is basically an open issue, though closed-form formulas are obtained in some cases (Kawan, 2011a, 2011b, 2011c; Nair et al., 2004; Savkin, 2006; Silva & Kawan, 2016).

This paper aims at tractable two-sided estimates capable to join up into a closed-form formula or at least to form a minor gap; to demonstrate such an ability, we shall use three multi-dimensional chaotic systems that are classically employed as testbeds in dynamical systems theory. The above objective is achieved via focusing on a new dynamical invariant of the system called the *restoration entropy* (RE). Meanwhile, RE is introduced in this paper as a response to a purely control-oriented inquiry: What is the minimal rate at which sensory data should be transferred to an estimator in order that the initial estimation accuracy can be reproduced somewhere in the future and then be maintained and, moreover, exponentially improved. The proposed approach to estimation of RE, similarly to Matveev and Pogromsky (2016) and Pogromsky and Matveev (2016a), integrates the strengths of the two Lyapunov's methods: lower and upper estimation follow the lines of the first and second (direct) Lyapunov methods, respectively, whereas the latter much benefits from special design of non-quadratic Lyapunov functions recently elaborated in stability analysis of nonlinear oscillations (Pogromsky & Matveev, 2016b).

In the first part of this paper (Matveev & Pogromsky, 2016), such an approach is elaborated for a discrete time model based on the concepts of regular and fine observability rates. Its extension to the continuous time case requires, first of all, elaboration of these concepts in this case, which also calls for decoupling the bit-rate issue from those of sampling and encoding. The key point is relevant advancement of the direct-Lyapunov-method part of the approach, whose critical trait is deducing asymptotic behavior of the system from its local (moreover, infinitesimal) properties at any point in the region of interest. The long-term experience with the direct Lyapunov method shows that its full strength stems from directly handling the ODE (or other equation) that embodies the continuous-time model. In this paper, we elaborate an analytical machinery that is based on ideas from Matveev and Pogromsky (2016) and Pogromsky and Matveev (2016a, 2016b) and endows them with the desired capacity of such handling. The potential of the thus developed techniques is exposed via analysis of three celebrated samples of chaotic dynamical systems.

In fact, this paper continues studying the control-oriented characteristics of a dynamical plant called observability rates in Matveev and Pogromsky (2016) and Pogromsky and Matveev (2016a). Their original definitions are built upon existence of an external to the plant observation scheme with certain properties. In this paper, we show that in the important case of observation within the scope of an invariant compact set and full-state measurement, these characteristics are equal to the RE, which is a purely intrinsic dynamical invariant of the plant. This quantity addresses the uncertainty that is injected by the system's dynamics into the knowledge of the state with the same accuracy as initially or, more specifically, the rate at which this uncertainty grows over time. We also disclose relations of the RE with the instruments of the classic dynamical systems theory such as finite- and infinite-time Lyapunov exponents and the TE, which opens the door for using classic tools and associated results for study and computation of the RE and throws an extra light on relationships between the RE and classic TE.

The material in this paper was partially presented at the 20th World Congress of the International Federation of Automatic Control, July 9–14, 2017, Toulouse, France (Matveev and Pogromsky (2017)).

The body of the paper is organized as follows. Section 2 sets up the problem of state estimation, whereas Section 3 provides basic definitions concerned with performance of respective observers. Assumptions about the plant are given in Sections 4, 5 presents the definition of the restoration entropy and the data rate theorem, which links the RE with the observability rates from Section 3. Relations of the RE with the Lyapunov exponents and TE, as well as its lower estimates, are discussed in Section 6, whereas upper estimates based on the direct Lyapunov approach are presented in Section 7. Section 8 is devoted to study of the Lorentz, Rössler, and Sprott systems; Section 9 offers brief conclusions. The proofs of most theoretical results are placed in appendices.

The following notations are adopted in the paper:

- $\lceil s \rceil$, $\lfloor s \rfloor$, integer ceiling and floor of s, respectively;
- $[n_1 : n_2]$, set of integers $j \in [n_1, n_2]$;
- $f(t+)$, limit of the function $f(\cdot)$ from the right;
- **int** E, interior of the set E;
- $\overline{E}$, closure of the set E;
- **co**E, closed convex hull of E;
- B_a^δ, open ball with a radius of δ centered at a;
- $\|\cdot\|$, Euclidean norm of a vector in $\mathbb{R}^n$ and the spectral norm of a square matrix;
- I, identity matrix.

2. Continuous-time state estimation problem

We consider a nonlinear time-invariant plant

$$\dot{x} = f(x), \ \ x = x(t) \in \mathbb{R}^n, \ \ t \in [0, \infty), \ \ x(0) \in K, \tag{1}$$

where x is the state, the map $f : \mathbb{R}^n \to \mathbb{R}^n$ is C^1-smooth, and $K \neq \emptyset$ is a compact set of the feasible initial states. As time t runs from 0, a valid estimate $\widehat{x}(t)$ of $x(t)$ should be continuously generated in real time t at a certain faraway site S_{est}, where direct observation of x is impossible. Meanwhile, $x(t)$ is fully accessible at time t in another location $S_{\text{sen}} \neq S_{\text{est}}$ (typically that of the plant) on the basis of sensor data. The problem to be treated stems from the bottleneck caused by a finite bit-rate of data communication from S_{sen} to S_{est}. *How many bits need to be minimally transmitted per unit time in order that valid estimation be feasible?*

More formally, at every time t_k from a certain sequence $t_0 = 0 < t_1 < \cdots$, the communication channel from S_{sen} to S_{est} accepts for transmission of a message $e(t_k)$ of a finite bit size. The transmission consumes time $\tau_k \geq 0$ and $e(t_k)$ fully reaches S_{est} at time $t_k+\tau_k$. The size of $e(t_k)$ may be manipulable and may affect the transmission delay τ_k. The time of departure t_k is recognizable at S_{est}, which is true if, e.g., τ_k is a known function of the bit size of $e(t_k)$.

The departure times t_k and messages $e(t_k)$ are generated by a special device, located at the sensor site S_{sen} and called the *coder*. It uses the knowledge of the preceding measurements, as well as the initial estimate $\widehat{x}(0)$ and its accuracy $\delta > 0$

$$\|x(0) - \widehat{x}(0)\| < \delta, \quad x(0), \widehat{x}(0) \in K, \tag{2}$$

and is described by equations of the following forms

$$e(t_k) = \mathcal{C}\big[t_k, x|_{[0,t_k]}; \widehat{x}(0), \delta\big], \tag{3}$$

$$t_{k+1} = \mathcal{T}\big[k, t_0, \ldots, t_k, x|_{[0,t_k]}; \widehat{x}(0), \delta\big], \quad t_0 := 0. \tag{4}$$

Here $s|_\Delta$ is the restriction of the signal $s(\cdot)$ on $\Delta \subset \mathbb{R}$. An estimate $\widehat{x}(t)$ of the current state $x(t)$ is generated at the current time t at the site S_{est} by a *decoder* based on the messages $e(t_k)$ received prior to this time:

$$\widehat{x}(t) = \mathcal{D}\big[t, \{e(t_k)\}_{k:t_k+\tau_k \leq t}; \widehat{x}(0), \delta\big] \quad t \geq 0. \tag{5}$$

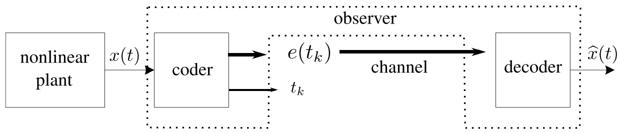

Fig. 1. State estimation over a finite capacity channel.

Both the coder and decoder are given common values of $\widehat{x}(0) \in K$ and δ and are aware of $f(\cdot)$ and K. Overall, an *observer* is given by (3)–(5), see Fig. 1.

A *trivial observer* does not employ communication from S_{sen} to S_{est} and merely generates the solution $\widehat{x}(\cdot)$ of the ODE from (1) that starts with $\widehat{x}(0)$. This observer is not acceptable in the case of our main interest, where K is an invariant chaotic set since then even a minor initial estimation error δ may drastically grow as time progresses. Our main concern is *how to protect the estimation accuracy from a critical degradation or even to preserve or improve it, and how much data must be minimally communicated per unit time to achieve these ends.*

In this paper, we are concerned with the time-averaged amount of data and employ a concept of the channel that is similar to that from Matveev and Savkin (2009, Sect. 3.4) and views the "channel" as a communication facility with the following features.

(c1) The time of the transmission time may be restricted: $t_{k+1} \in \mathcal{N}[\eta_k]$, where $\eta_k := [t_0, \dots, t_k, e(t_0), \dots, e(t_k)]$ and $\mathcal{N}(\eta_k) \subset (t_k, \infty)$ is a multivalued map associated with the "channel"; "no restriction" is embodied by $\mathcal{N}(\eta_k) = (t_k, \infty)$.

The trait **(c1)** may catch, e.g., that only integer multiples of a given $\Delta > 0$ may serve as departure times (then $\mathcal{N}(\eta_k) \subset \{t = i\Delta > t_k : i \text{ is integer}\}$) or that the next departure cannot be commenced while the channel is busy transmitting $e(t_k)$ (then, e.g., $\mathcal{N}(\eta_k) \subset (t_k + \Theta[|e(t_k)|], \infty)$, where $|e(t_k)|$ is the bit-size of $e(t_k)$ and $\Theta[|e(t_k)|]$ is the duration of the channel being closed). Also, the times t_k may be pre-specified, then $\mathcal{N}(\eta_k)$ contains a single element t_{k+1}.

(c2) Message choice may be restricted:

$$e(t_{k+1}) \in \mathcal{E}[t_0, \dots, t_{k+1}, e(t_0), \dots, e(t_k)].$$

At its simplest, $\mathcal{E}(\cdot)$ is constantly a given finite set E. A somewhat converse situation is where $\mathcal{E}(\cdot)$ is the set of all binary sequences of arbitrary lengths.

Needless to say that the coder (4) is restricted to meet **(c1,2)**.

(c3) Messages are not corrupted but may arrive out of order: $t_k + \tau_k > t_{k+1} + \tau_{k+1}$. The message may be lost (then $\tau_k := \infty$); otherwise τ_k does not exceed a constant $\overline{\tau} \in (0, \infty)$.[1]

(c4) There is a uniform upper bound $b_+(\Delta t)$ on the total number of bits that can be transferred during any time interval of duration Δt subject to **(c1)–(c3)**.[2] Meanwhile, no less than $b_-(\Delta t)$ bits can be surely transferred during any such interval provided that t_k and $e(t_k)$ are properly chosen.

(c5) The respective per unit time upper and lower bounds converge to a common value c called the *channel capacity*:

$$b_\pm(\Delta t)/(\Delta t) \to c \quad \text{as} \quad \Delta t \to \infty. \tag{6}$$

For examples and more discussion on this notion of the channel, we refer the reader to Matveev and Savkin (2005) and Matveev and Savkin (2009, Sect. 3.4).

What channel capacity is minimally required in order that a certain degree of observation performance be achievable?

3. Definitions on the observation performance

From now on, we assume that $x(0), \widehat{x}(0) \in K$. The first three definitions follow the lines of Matveev and Pogromsky (2016).

Definition 1. An observer is said *to observe the system* (1) if for any $\varepsilon > 0$ there is $\delta(\varepsilon) > 0$ such that whenever (2) holds with $\delta = \delta(\varepsilon)$, the estimation error never exceeds ε:

$$\|x(t) - \widehat{x}(t)\| \le \varepsilon \qquad \forall t \ge 0. \tag{7}$$

This does not exclude that a disproportionally higher initial accuracy $\delta(\varepsilon) \ll \varepsilon$ is needed, which means that the initial accuracy is drastically degraded as time passes. The next definition introduces two classes of observers for which, in various degrees, such a decay of accuracy is excluded.

Definition 2. An observer is said

(i) *to regularly observe* the system (1) if the anytime error is uniformly proportional to the initial error: there exist $\delta_* > 0$ and $G > 0$ such that whenever $\delta \le \delta_*$ in (2),

$$\|x(t) - \widehat{x}(t)\| \le G\delta \qquad \forall t \ge 0; \tag{8}$$

(ii) *to finely observe* the system (1) if in addition to (i) the observation error exponentially decays to zero as time progresses: there exist $\delta_* > 0$, $G > 0$, and $g > 0$ such that whenever $\delta \le \delta_*$ in (2), the following holds

$$\|x(t) - \widehat{x}(t)\| \le G\delta e^{-gt} \qquad \forall t \ge 0. \tag{9}$$

Definition 3. The system (1) is said to be **(i)** *observable,* **(ii)** *regularly observable, or* **(iii)** *finely observable* via a given communication channel if there is an observer that firstly, operates through this channel and secondly, observes, regularly observes, or finely observes the system, respectively.

What channel rate c is minimally needed in order that the system (1) be observable in every of these senses?[3] In other words, what is the infimum value $\mathfrak{R}(f, K)$ of the rates c such that (1) is observable in the selected sense via any communication channel with capacity c?[4] Here the index $_{\text{o}}$, $_{\text{ro}}$, and $_{\text{fo}}$ is attached to $\mathfrak{R}$ in the cases (i), (ii), and (iii), respectively. The arguments f, K underscore that these quantities are entirely determined by the system (1). It is easy to see that

$$\mathfrak{R}_{\text{o}}(f, K) \le \mathfrak{R}_{\text{ro}}(f, K) \le \mathfrak{R}_{\text{fo}}(f, K). \tag{10}$$

If f or K is clear from the context, it can be dropped in these notations. In the remainder, we study $\mathfrak{R}_{\text{o}}, \mathfrak{R}_{\text{ro}}, \mathfrak{R}_{\text{fo}}$, disclose their relations with the classic topological entropy, and offer constructive separate necessary and sufficient bounds on them. These bounds may coincide, thus providing an exhaustive (necessary and sufficient) answer, as will be shown in Section 8.1 for a classic chaotic system. However, we cannot guarantee that these bounds coincide in general.

Since any two norms in $\mathbb{R}^n$ are equivalent $\mu_1\|x\|_1 \le \|x\|_2 \le \mu_2\|x\|_1$, the norm does not affect the quantities from (10).

4. Technical assumptions about the plant

From now on, the following assumptions are adopted.

Assumption 4. Any solution $x(t, a)$ of the ODE from (1) starting with $x(0) = a \in K$ is defined for all $t \ge 0$.

[1] This may result from the policy under which excessively delayed messages are ignored at S_{est}.

[2] A message is said to *be transferred during a time interval* if it both departs and fully arrives within this interval.

[3] The fact that apart from c, the other channel's features are of no concern here will be rigorously justified in Lemma 18.

[4] By Lemma 18, "any" can be replaced by "some" here.

Then the system (1) gives rise to a flow $a \in K \mapsto \phi^t(a) := x(t,a), t \geq 0$. We also put

$$K(t) := \phi^t(K), \quad K^\infty := \bigcup_{t\geq 0} K(t),$$
$$A(x) = \frac{\partial f}{\partial x}(x), \quad X(t,x) := \frac{\partial \phi^t}{\partial x}(x) \tag{11}$$

and recall that the η-vicinity of a set $E \subset \mathbb{R}^n$ is defined as

$$E^\eta := \{x : \inf_{y\in E} \|x-y\| < \eta\}, \qquad \eta > 0. \tag{12}$$

Assumption 5. The set K is compact and $A(\cdot)$ is bounded and uniformly continuous on $[\mathbf{co}K^\infty]^\eta$ with some $\eta > 0$.

Assumptions 4, 5 hold if K is compact and positively invariant (Hartman, 1964, Th. 3.1) or if the system (1) is ultimately bounded, i.e., there exist $\beta, \gamma, T > 0$ such that whenever $\|a\| \leq \gamma$, the solution $x(t,a)$ is defined for all $t \geq 0$ and $\|x(t,a)\| \leq \beta \ \forall t \geq T$; in the last case, one may take $K := \overline{B_0^\gamma}$.

5. Topological and restoration entropies, and associated data rate theorem

Now we relate the introduced notions with the classic concept of the topological entropy (TE). We first recall the latter and offer its modification, which definition is somewhat similar in flavor to that of the TE.

Definition 6. For $T, \varepsilon > 0$, a set Q is said to be *(T,ε)-spanning* (for the system (1)) if its members are solutions of (1) defined on $[0,T]$ and starting in K and for any $a \in K$, there exists $x^\star(\cdot) \in Q$ such that

$$\left\|x(t,a) - x^\star(t)\right\| \leq \varepsilon \qquad \forall t \in [0,T]. \tag{13}$$

The smallest possible number of elements in Q with the described property is denoted as $q(T,\varepsilon)$. The *topological entropy* of the system (1) on K is defined as (Downarowicz, 2011; Katok, 2007)

$$H(f,K) := \lim_{\varepsilon\to 0} \overline{\lim_{T\to\infty}} T^{-1} \log_2 q(T,\varepsilon). \tag{14}$$

Thus the TE answers the question: how many bits of data must be conveyed to an estimator per unit time in order that the error in the state estimate be constantly kept below a given level irrespective of how small this level might be.

As will be shown, the rates $\mathfrak{R}_{\rm ro}$ and $\mathfrak{R}_{\rm fo}$ respond to another inquiry. Specifically, let $p(T,a,\delta)$ be the minimal number of open δ-balls required to cover the image $\phi^T(B_a^\delta \cap K)$.

Let an estimate $a \in K$ of $x(0) \in K$ be known and let its accuracy be δ. How many bits must be conveyed to the estimator in order that at time T, it could enjoy awareness of $x(T)$ with the same accuracy? By treating the centers of the covering balls as candidates for the estimates, and vice versa, we get the answer: $\log_2 p(T,a,\delta)$ at the minimum.

In this answer, we consecutively sharpen the focus on

(i) the uniform bound $p_K^{\log}(T,\delta) := \sup_{a\in K} \log_2 p(T,a,\delta)$;
(ii) the highest accuracies by considering $\overline{\lim}_{\delta\to 0}\, p_K^{\log}(T,\delta)$;
(iii) the asymptotically average number of bits that must be sent to the observer per unit time provided that the estimate of $x(T)$ is needed at time $T \to \infty$ by considering

$$H_{\rm res}(f,K) := \overline{\lim_{T\to\infty}} \frac{1}{T} \overline{\lim_{\delta\to 0}} \sup_{a\in K} \log_2 p(T,a,\delta). \tag{15}$$

This quantity addresses the uncertainty that is injected by the system's dynamics into the knowledge of the state with the same accuracy as initially; more specifically, (15) characterizes the rate at which this uncertainty grows over time.

Definition 7. The quantity (15) is called the *restoration entropy* of the system (1) on K.

Definition 7 addresses restoring the accuracy somewhere in the future. Now we show that (15) is also responsible for the feasibility to constantly keep the accuracy proportional to its initial value and to quickly improve the accuracy by driving the error to zero at an exponential rate.

Theorem 8 (*Data Rate Theorem*)**.** *The following holds:*

$$H \leq \mathfrak{R}_{\rm o}, \qquad H_{\rm res} \leq \mathfrak{R}_{\rm ro}. \tag{16}$$

If the set K is positively invariant, then moreover,

$$\mathfrak{R}_{\rm o} = H, \qquad \mathfrak{R}_{\rm ro} = \mathfrak{R}_{\rm fo} = H_{\rm res}. \tag{17}$$

The proof of this theorem and Remark 10 will be given in Appendix C. Basic details of design of observers underlying the proof of Theorem 8 are elucidated in Appendix B. There are systems for which the strict inequality $H_{\rm res} > H$ holds true, as follows from Example 5.1 in Pogromsky and Matveev (2016a).

Remark 9. There are known several definitions of the TE of a flow, which yet produce the same number (14); see, e.g., Downarowicz (2011) and Katok (2007). For example, this notion is sometimes defined as the TE of the map $\phi^1(\cdot)$. Another idea is to form $q(T,\varepsilon)$ in (14) by using a larger variety of sets Q which still retain the approximation property (13) but whose elements are allowed not only to disobey (1) but also to loose continuity.

Remark 10. Definition 7 makes sense for any continuous flow $\{\phi^t(\cdot)\}_{t\geq 0}$ in $K \subset \mathbb{R}^n$, not only for ones given by ODE's. Consider two flows $\{\phi_i^t(\cdot)\}_{t\geq 0}$ in compacts sets $K_i \subset \mathbb{R}^n$, $i = 1,2$. Let these flows be Lipschitz conjugate: there exists an invertible map $h : \mathbb{R}^n \to \mathbb{R}^n$ such that $\phi_2^t(\cdot) = h \circ \phi_1^t(\cdot) \circ h^{-1} \ \forall t \geq 0, h(K_1) = K_2$, whereas h and h^{-1} are Lipschitz continuous on K_1^∞ and K_2^∞, respectively. Then these flows have a common value of the RE.

6. Restoration entropy and Lyapunov exponents

Now we relate the restoration entropy with the *finite-time Lyapunov exponents* (Haller, 2000, 2001), which are given (strictly speaking, modulo the logarithm base change $2 \mapsto e$) by

$$\lambda_i(T,a) := T^{-1} \log_2 \alpha_i[X], \text{ where } X := X(T,a) \tag{18}$$

is defined in (11) and $\alpha_1(X) \geq \cdots \geq \alpha_n(X)$ are the singular values of X. The relation to be established also provides a method for practical (mainly, lower) estimation of $H_{\rm res}$.

We introduce the sum of the positive Lyapunov exponents:

$$\Sigma(T,a) := \sum_{i=1}^n \{\lambda_i(T,a)\}_+, \ \{s\}_+ := \max\{s;0\}. \tag{19}$$

Theorem 11. *Let Assumptions 4 and 5 hold. Then*

$$\overline{\lim_{T\to\infty}} \max_{a\in\mathbf{int}K} \Sigma(T,a) \leq H_{\rm res} \leq \overline{\lim_{T\to\infty}} \max_{a\in K} \Sigma(T,a), \tag{20}$$
$$\mathfrak{R}_{\rm fo} \leq \overline{\lim_{T\to\infty}} \sup_{a\in K^\infty} \Sigma(T,a). \tag{21}$$

The proof of this theorem is given in Appendix D.

If $\mathbf{int}K = \emptyset$, the first inequality in (20) is void. If $\overline{\mathbf{int}K} = K$, (20) is true with $=$ put in place of $\leq$.

For any linear $f(x) = Ax + b$ system (1),

$$\lambda_i(T,a) = \frac{\log_2 \alpha_i\left[e^{AT}\right]}{T} \to \frac{1}{\ln 2}\mathrm{Re}\,\beta_i \text{ as } T \to \infty, \tag{22}$$

where $\beta_1, \ldots, \beta_n$ are the eigenvalues of A repeated in accordance with their algebraic multiplicities and ordered so that their real parts $\operatorname{Re}\beta_i$ decay as i grows (Barreira & Valls, 2017). Hence for any compact set K with a nonempty interior,

$$H_{\text{res}} = \frac{1}{\ln 2}\sum_{j=1}^{n}\{\operatorname{Re}\beta_j\}_+ \tag{23}$$

and so $H_{\text{res}} = H$ according to Savkin (2006).

Now let K be positively invariant. The limit $\lambda_i(a) := \lim_{T\to\infty} \lambda_i(T,a)$ is known under the name of the *Lyapunov exponent* (modulo $\log_2 \mapsto \ln$ in (18)). Let K_L be the set of all $a \in K$ for which this limit (i.e., $\lambda_i(a)$) does exist. For any invariant Borel measure μ on K, the multiplicative ergodic theorem (Ruelle, 1979) yields that μ-almost all points $a \in K$ are in K_L, the Lyapunov exponents are invariant with respect to the flow $\phi^T(\cdot)$ μ-almost everywhere; also, if μ is ergodic, they are constant up to a set of μ-zero measure. By (20),

$$\begin{aligned}
&\sup_{a\in K_L\cap \mathbf{int}K}\sum_{i=1}^{n}\{\lambda_i(a)\}_+ \le H_{\text{res}}\\
&\le \overline{\lim_{T\to\infty}}\sup_{a\in K}\sum_{i=1}^{n}\{\lambda_i(T,a)\}_+\\
&\overset{\maltese}{=} \sup_{a\in K}\overline{\lim_{T\to\infty}}\sum_{i=1}^{n}\{\lambda_i(T,a)\}_+ ,
\end{aligned}\tag{24}$$

where ($\maltese$) holds by Schreiber (1998, Thm. 1) (applied to the subadditive function $F(a,t) := \sum_{i=1}^{n}\{\log_2 \alpha_i[X(t,a)]\}_+$).[5]

The definition of H_{res} and (20) can be extended on flows on a compact Riemannian manifold $\mathfrak{M}$, in which case the differential $X(t,a)$ in (11) and (18) acts between the tangent spaces. This remark throws an extra light on the inequality $H \le H_{\text{res}}$ that holds for the invariant set $K := \mathfrak{M}$ by Theorem 8. Indeed, by the variational principle for entropy (Dinaburg, 1971, Cor. 2.16) and the Ruelle inequality (Katok & Hasselblatt, 1995, S.2.13) $H \le \sup_\mu \int_{\mathfrak{M}}\sum_{i=1}^{n}\{\lambda_i(a)\}_+\,d\mu$, where sup is over all invariant Borel probability measures μ on $\mathfrak{M}$. Meanwhile, H_{res} is the uniform "μ-almost sure" upper bound of the integrand by the first inequality from (24) (for all of the above μ's).

For any equilibrium point a of the nonlinear system (1), the r.h.s. in (23) with $A := \frac{\partial f}{\partial x}(a)$ is denoted as $H_L(f,a)$ and called the *proximate topological entropy around a*.

Corollary 12. *The following entailment holds:*

$$a \in \mathbf{int}\,K \;\text{ is an equilibrium } \Rightarrow H_{res} \ge H_L(f,a). \tag{25}$$

This is implied by (22) and the first inequality in (20).

Now we sketch (without proofs) two ways of weakening the assumptions of Corollary 12: relaxation of the condition that (1) a is an equilibrium or (2) $a \in \mathbf{int}\,K$, respectively.

(1) Let an orbit $x(t,a)$ be T-periodic: $x(T,a) = a$. Then clearly $X(t+T,a) = X(t,a)X(T,a)\ \forall t \ge 0$. So $\Sigma(t,a) = t^{-1}\sum_{i=1}^{n}\{\log_2 \alpha_i[X(t - T\lfloor t/T\rfloor, a)X[T,a]^{\lfloor t/T\rfloor}]\}_+$ and

$$\overline{\lim_{t\to\infty}}\,\Sigma(t,a) = \lim_{t\to\infty}\Sigma(t,a) = \sum_{i=1}^{n}\frac{\{\log_2|\gamma_i[X(T,a)]|\}_+}{T}.$$

Here $X(T,a)$ is the monodromy matrix of the periodic orbit at hands and $\gamma_i[X(T,a)]$ are its eigenvalues repeated in accordance with their algebraic multiplicity. If additionally $a \in \mathbf{int}\,K$, we thus arrive at another lower estimate

$$H_{\text{res}} \ge \sum_{i=1}^{n}\frac{\{\log_2|\gamma_i[X(T,a)]|\}_+}{T}.$$

[5] We thank an anonymous reviewer for the remark ($\maltese$).

(2) Let $M \ni a$ be a smooth manifold whose intersection with a sufficiently small vicinity of a (in $\mathbb{R}^n$) lies in K. (An example is the manifold of dimension 0 that consists of one point a.) Let $\mathfrak{M}(a)$ be the set of all such manifolds. Let $\alpha_i(T,a,M)$ be the singular values of $X(T,a)\pi_{T_a(M)}$. Here $T_a(M)$ is the tangent space of M at a and $\pi_{T_a(M)}$ is the orthogonal projection of $\mathbb{R}^n$ onto this space. Then

$$H_{\text{res}}(f,K) \ge \overline{\lim_{T\to\infty}}\sup_{a\in K}\Sigma_{\text{sm-incl}}(T,a), \quad \text{where} \tag{26}$$

$$\Sigma_{\text{sm-incl}}(T,a) := \sup_{M\in\mathfrak{M}(a)}\sum_{i=1}^{n}\left\{\frac{\log_2\alpha_i(T,a,M)}{T}\right\}_+ .$$

The proof of (26) is by combining the ideas of the proofs of Theorem 9 in Matveev and Pogromsky (2016) and Theorem 11, respectively.

7. The direct Lyapunov approach to upper estimation of the fine observability rate and restoration entropy

By (17) and (20), the studied observability bit-rates are mixed up with the TE and the long-horizon limits of the Lyapunov exponents. Their computation or even estimation still remains an intricate matter in general (Skokos, 2010; Thiffeault, 2002; Wolf, Swift, Swinney, & Vastano, 1985), with prevalence of numerical methods. Now we offer an analytical approach to upper estimation of $\mathfrak{R}_{\text{fo}}$ and simultaneously, thanks to (10) and (16), of the RE. This approach follows the lines of the second Lyapunov method and uses a non-traditional design of a Lyapunov-like function.

The components of this design are introduced in the following Assumption 13. Their proposed use for building such a function and detailed discussion of this and very close assumptions are offered in Pogromsky and Matveev (2011, 2013, 2016b) and Pogromsky, Matveev, Chaillet, and Rüffer (2013) in relation to the study of various dynamical properties, including the TE (Pogromsky & Matveev, 2011). As a brief reminder, we recall that this design relies on a positively definite state-dependent quadratic forms that, unlike Lyapunov functions, interchange growth and decay along the system's trajectories so that decay prevails " on average". The tractability of the advocated approach in the current context will be demonstrated in Section 8.

For a smooth function $g(x)$ of x, we put $\dot g(x) := \frac{\partial g}{\partial x}(x)f(x)$.

Assumption 13. There exist scalar $\upsilon_d(x)$, $d \in [1:n]$ and $n\times n$-matrix valued $P(x) = P(x)^\top$ functions defined and continuously differentiable in a vicinity (12) of K^∞ such that the following statements hold:

(i) $P(x)$ is bounded and uniformly positive definite:

$$\exists \mu_1,\mu_2 > 0:\ \mu_1^2 I_n \le P(x) \le \mu_2^2 I_n\ \forall x \in K^\infty; \tag{27}$$

(iii) $\upsilon_d(x)$ are bounded: $|\upsilon_d(x)| \le V < \infty\ \forall x \in K^\infty, d$;

(iii) Let $\sigma_1(x) \ge \sigma_2(x) \ge \cdots \ge \sigma_n(x)$ be the roots of the following algebraic equation

$$\det[A(x)^\top P(x) + P(x)A(x) + \dot P(x) - \sigma P(x)] = 0$$

repeated in accordance with their algebraic multiplicities. There exist constants $\Lambda_d \ge 0, d \in [1:n]$ such that

$$\sum_{i=1}^{d}\sigma_i(x) + \dot\upsilon_d(x) \le \Lambda_d \quad \forall d \in [1:n], x \in K^\infty. \tag{28}$$

Theorem 14. *Suppose that apart from Assumptions 4 and 5, the system (1) and the set K satisfy Assumption 13. Then*

$$\mathfrak{R}_{\text{fo}} \le \Lambda/(2\ln 2), \quad \text{where } \Lambda := \max_d \Lambda_d. \tag{29}$$

Proof. Claim (i) in Assumption 13 guarantees that for any $x \in K^\infty$, there is a positively definite $n \times n$-matrix $S(x)$ such that $S(x)S(x) = P(x)$. Since its eigenvalues are the square roots of the eigenvalues of $P(x)$, (27) yields that

$$\mu_1 I \leq S(x) \leq \mu_2 I. \tag{30}$$

By putting $\omega_d(A) := \prod_{i=1}^d \alpha_i(A)$ and using Lemma 8.1 and Proposition 8.6 in Pogromsky and Matveev (2011), we have for any $d \in [1:n]$, $a \in K$,

$$\begin{aligned}
&\omega_d \left\{S\left[x(T,a)\right]X(T,a)S(a)^{-1}\right\} \\
&\quad \leq \exp\left\{\frac{1}{2}\int_0^T \{\sigma_1[x(t,a)] + \dots + \sigma_d[x(t,a)]\}\,dt\right\} \\
&\quad \overset{(28),(29)}{\leq} \exp\left\{\frac{1}{2}\int_0^T \{\Lambda - \dot{v}_d[x(t,a)]\}\,dt\right\} \\
&\quad = \exp\left\{(\Lambda T - v_d[x(T,a)] + v_d[a])/2\right\} \overset{\text{(ii)}}{\leq} \exp(V + \Lambda T/2).
\end{aligned}$$

By Horn's inequality (Boichenko, Leonov, & Reitman, 2005, Prop. 2.3.1), $\omega_d(AB) \leq \omega_d(A)\omega_d(B)$ and so

$$\begin{aligned}
\omega_d[X(T,a)] &\leq \omega_d\left\{S\left[x(T,a)\right]X(T,a)S(a)^{-1}\right\} \\
&\quad \times \omega_d\left\{S\left[x(T,a)\right]^{-1}\right\}\omega_d\left\{S(a)\right\}.
\end{aligned}$$

The singular values of the positively definite matrices $S(x)$ and $S(x)^{-1}$ are identical to their eigenvalues, which lie in $[\mu_1, \mu_2]$ and $[\mu_2^{-1}, \mu_1^{-1}]$, respectively, due to (30). Thus

$$\begin{aligned}
\omega_d[X(T,a)] \quad &\leq \quad \left(\frac{\mu_2}{\mu_1}\right)^n \exp(V + \Lambda T/2); \\
\Sigma(T,a) \overset{(18),(19)}{=\!=} \quad & T^{-1}\sum_{i=1}^n \{\log_2 \alpha_i\left[X(T,a)\right]\}_+ \\
= \quad & T^{-1}\max_{d\in[1:n]}\left\{\sum_{i=1}^d \log_2 \alpha_i\left[X(T,a)\right]\right\}_+ \qquad (31)\\
= \quad & T^{-1}\max_{d\in[1:n]}\{\log_2 \omega_d\left[X(T,a)\right]\}_+ \\
\leq \quad & T^{-1}\max_{d\in[1:n]}\left\{\log_2\left[\left(\frac{\mu_2}{\mu_1}\right)^n \exp(V + \Lambda T/2)\right]\right\}_+ \\
= \quad & T^{-1}\{n(\log_2\mu_2 - \log_2\mu_1) \\
& + (V + \Lambda T/2)/\ln 2\}_+
\end{aligned}$$

The proof is completed by (21). ∎

Theorem 14 can be viewed as a continuous-time analog of an akin result from Matveev and Pogromsky (2016) and also enhances it by using state-dependent matrices P. In Pogromsky and Matveev (2011), the similar estimate $\frac{\Lambda}{2\ln 2} \geq H$ is derived for non-autonomous systems; this paper deals with autonomous systems only to ease technicalities.

We proceed to study of three celebrated chaotic systems.

8. Examples

8.1. Lorenz system

Consider the Lorenz system with the parameters $\sigma, r, b > 0$:

$$\dot{x} = -\sigma x + \sigma y, \quad \dot{y} = rx - y - xz, \quad \dot{z} = -bz + xy. \tag{32}$$

If $r \leq 1$, the equilibrium $x = y = z = 0$ is globally asymptotically stable. The main interest to this system is associated with $r > 1$. Then the origin is a hyperbolically unstable saddle and there are two more equilibria.

Theorem 15. *Let $r > 1$ and let all equilibria of (32) be hyperbolically unstable. Then for any compact set K whose interior contains the origin* $(0, 0, 0)$,

$$\mathfrak{R}_{\text{fo}}(K) = \frac{1}{2\ln 2}\left(\sqrt{(\sigma-1)^2 + 4r\sigma} - (\sigma+1)\right). \tag{33}$$

Proof. With a quadratic Lyapunov function, it is easy to see that the system (32) is ultimately bounded and all its solutions are extensible on $[0, \infty)$. Thus Assumptions 4 and 5 hold. Via applying (25) to $(0, 0, 0)$ and an elementary exercise in computation of the eigenvalues of the associated Jacobian, we see that the r.h.s. of (33) does not exceed H_{res}, where $H_{\text{res}} \leq \mathfrak{R}_{\text{ro}} \leq \mathfrak{R}_{\text{fo}}$ by Theorem 8 and (10). Meanwhile, Assumption 13 was justified in Pogromsky and Matveev (2011, Sect. 4) and the related quantity $\frac{\Lambda}{2\ln 2}$ from (29) was shown to be equal to the r.h.s. of (33). Theorem 14 completes the proof. ∎

In Pogromsky and Matveev (2011), the r.h.s. of (33) was shown to be no less than $H(f, K)$ for any positively invariant compact set K if all equilibria of (32) are unstable. This result was based on Leonov, Pogromsky, and Starkov (2011); whereas a certain flaw in Leonov et al. (2011) was later corrected in Leonov, Pogromsky, and Starkov (2012).

The assumption on instability of all equilibria excludes the case of the so called "transient chaos" (Lai & Tél, 2011), which can appear even if the Lorenz system possesses two stable equilibria along with the unstable saddle. Using a result from Leonov, Kuznetsov, Korzhemanova, and Kusakin (2016), that assumption can be relaxed to cope with a possible "transient chaos" so that (33) remains valid in this case as well. However, corresponding developments are very involved.

8.2. Rössler system

This is the following system with the parameters $a, b > 0$

$$\dot{x} = -y - z, \quad \dot{y} = x, \quad \dot{z} = -bz + a(y - y^2), \tag{34}$$

which is traditionally attributed to Rössler (Eq. (5) in Rössler (1979)).

Theorem 16. *For any compact positively invariant set K of the Rössler system (34), the following inequality holds:*

$$\mathfrak{R}_{\text{fo}}(K) \leq \frac{1}{2\ln 2}\left(\sqrt{2a + 4b + b^2} - b\right). \tag{35}$$

Proof. To check Assumption 13, we follow Leonov and Alexeeva (2015) and consider $P(x, y, z) = P := \mathbf{diag}(1, 1, p)$, where $p > 0$ will be specified later on. Then in **(iii)** from Assumption 13,

$$A(y)^\top P + PA(y) - \lambda P = \begin{pmatrix} -\lambda & 0 & -1 \\ 0 & -\lambda & ap(1-2y) \\ -1 & ap(1-2y) & -p(\lambda + 2b) \end{pmatrix}.$$

It is easy to see that the equation $\det(A^\top P + PA - \lambda P) = 0$ has three roots $\lambda_1 > \lambda_2 = 0 > \lambda_3$,

$$\lambda_{1,3} = -b \pm \sqrt{b^2 + p^{-1} + a^2 p(1-2y)^2}.$$

Hence $\lambda_1 \geq \sum_{i=1}^d \lambda_i$, $d = 2, 3$. So to ensure (28) (for all d's), it suffices to find $v_1(\cdot)$ and Λ_1 such that (28) holds with $d = 1$: then (28) is true with $v_3(\cdot) := v_2(\cdot) := v_1(\cdot)$ and $\Lambda_3 := \Lambda_2 := \Lambda_1$. For $v_1 := \gamma_1(z - bx)$, where γ_1 is a tunable parameter, we have $\dot{v}_1 = \gamma_1\left[(a+b)y - ay^2\right]$. So (28) (with $d = 1$) does hold if Λ_1 is any upper bound for

$$\lambda_1 + \dot{v}_1 = -b + \sqrt{b^2 + \frac{1}{p} + a^2 p(1-2y)^2} + \gamma_1\left[(a+b)y - ay^2\right].$$

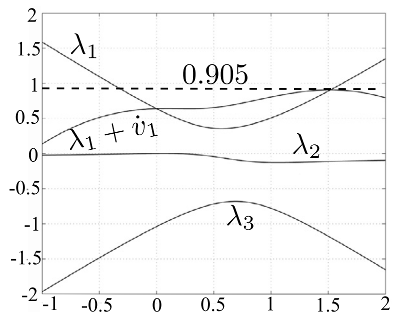

Fig. 2. The graphs of $\lambda_i,\ i=1,2,3$ and $\lambda_1+\dot{v}_1$ versus y.

To find such a bound, we use the elementary inequality $\sqrt{q} \le \gamma_2 q + (4\gamma_2)^{-1}\ \forall q \ge 0, \gamma_2 > 0$ to see that

$$\lambda_1+\dot{v}_1 \le \gamma_2\left[b^2+\frac{1}{p}+pa^2(1-2y)^2\right]+\frac{1}{4\gamma_2} \\ +\gamma_1\left[(a+b)y-ay^2\right]-b.$$

Maximizing the r.h.s. of this inequality over y and choosing consecutively $p := (a+2b)^{-1}$,

$$\gamma_2 := \frac{1}{2\sqrt{p(a+2b)^2+b^2+\frac{1}{p}}}, \gamma_1 := 4\gamma_2 ap\frac{a+2b}{a+b}$$

yields $\lambda_1+\dot{v}_1 \le \Lambda_1 := \sqrt{2a+4b+b^2}-b$. Theorem 14 completes the proof. ■

For $a=0.386$, $b=0.2$, as in Rössler (1979), the r.h.s. of (35) is equal to 0.7716. At the same time the proximate topological entropy around the saddle equilibrium $x_0=0,\ y_0=1+b/a,\ z_0=-1-b/a$ is smaller: $H_L(x_0,y_0,z_0)=0.6527$. Now let O be an orbit such that the solutions started in O are bounded in forward time and converge to x_0,y_0,z_0 in backward time, and let $K=\overline{O}$. Then (26) gives an evidence that $0.6527 \le \mathfrak{R}_{\rm fo}(K) \le 0.7716$ and so there is a gap between the upper and lower estimates in general.

Meanwhile, numerical analysis provides an evidence that the approach of this paper allows to shrink this gap to $\approx$ 0.0001. Indeed, for $a=0.386$, $b=0.2$, $v_1 := 1.494(z-bx)$, and

$$P=\begin{pmatrix}1.000 & 0.006 & -0.068\\ 0.006 & 0.811 & 0.484\\ -0.068 & 0.484 & 1.06\end{pmatrix},$$

one can numerically verify that

$$\lambda_1(y)+\dot{v}_1=\lambda_1(y)+1.494((a+b)y-ay^2)\lessapprox 0.905.$$

Numerical analysis also reveals that $\lambda_2(y) \lessapprox 0\ \ \forall y$ (see Fig. 2). Therefore, Theorem 14 gives the following estimate for an arbitrary positively invariant compact set K:

$$\mathfrak{R}_{\rm fo}(K) \le (2\ln 2)^{-1}[\lambda_1(y)+\dot{v}_1]\approx 0.6528.$$

This is a very accurate approximation of the lower estimate $\mathfrak{R}_{\rm fo}(K) \ge 0.6527$, which holds for $K=\overline{O}$ by (17) and (26). This approximation emphasizes the role of equilibrium points in providing the exact value of $\mathfrak{R}_{\rm fo}(K)$, for a relevant discussion related to the Rössler system, see also Kuznetsov, Mokaev, and Vasilyev (2014).

Thus, via the offered method, $\mathfrak{R}_{\rm fo}(K)$ can be computed with a high accuracy for the classic values $a=0.386$, $b=0.2$.

8.3. Sprott S-system

In Sprott (1994, Table 1), the following chaotic system is suggested:

$$\dot{x}=-x-4y,\quad \dot{y}=x+z^2,\quad \dot{z}=1+x. \tag{36}$$

It is easy to verify that this system is equivalent to

$$\dddot{z}+\ddot{z}+4\dot{z}+4z^2-4=0, \tag{37}$$

There are two equilibria in the system; for them, $z=\pm 1$, respectively; whereas the Jacobian matrix is given by

$$A(z)=\begin{pmatrix}-1 & -4 & 0\\ 1 & 0 & 2z\\ 1 & 0 & 0\end{pmatrix}.$$

By (37), $\dot{v}=z^2-1$ for $v := -\frac{1}{4}(\ddot{z}+\dot{z})-z$. With

$$P := \begin{pmatrix}1.0000 & 0.2876 & -0.1792\\ 0.2876 & 2.3047 & -1.5663\\ -0.1792 & -1.5663 & 3.7614\end{pmatrix},$$

numerical analysis shows that the roots λ_i from **(iii)** of Assumption 13 are such that $\lambda_{2,3}(z) < 0$ and the function $v_1=-0.735v$ obeys the inequality

$$\lambda_1(z)+\dot{v}_1=\lambda_1(z)-0.735\dot{v}\lessapprox 2.406. \tag{38}$$

Whence, for any positively invariant compact set K,

$$\mathfrak{R}_{\rm fo}(K) \le (2\ln 2)^{-1}\lambda_1(y)+\dot{v}_1 \approx 1.7356.$$

To lower estimate $\mathfrak{R}_{\rm fo}(K)$ based on (17) and (26), we observe that $2A(-1)$ has two complex conjugate eigenvalues with negative real parts and a real positive eigenvalue $\approx$ "the r.h.s. of (38)". Let O be an orbit such that solutions started in O are bounded in forward time and converge to the equilibrium with $z=-1$ in backward time. Then for $K=\overline{O}$,

$$\mathfrak{R}_{\rm fo}(K)\approx 1.7356.$$

We conclude this example by observing that via an affine change of coordinates, both the Rössler system (34) and the Sprott S-system (36) can be rewritten in the form

$$\dddot{z}+a_2\ddot{z}+a_1\dot{z}-a_0+bz^2=0,\quad a_0,b\ge 0. \tag{39}$$

Moreover, both the Sprott Q-system (Sprott, 1994, Table 1) and the Ermentrout system (Ermentrout, 1984, Eq.(4.1)) can be similarly reduced to (39), and so (39) covers four classic chaotic systems. There are two equilibria $z_{1,2}=\pm\sqrt{a_0/b}$ in (39). Numerical search of a matrix P with the averaging function v linear in $(z,\dot{z},\ddot{z})$ similar to that discussed in Sections 8.2, 8.3 supports the conjecture that for (39) and any positively invariant compact set K, $\mathfrak{R}_{\rm fo}(K)$ is upper bounded by the largest proximate TE around $z_{1,2}$. However, verification of this conjecture seems to be technically involved, and we leave this as a topic of further research. If this conjecture were true, then one could construct the set K similar to $K=\overline{O}$ as above so that the value of $\mathfrak{R}_{\rm fo}(K)$ can be found exactly.

9. Conclusion and future work

A new characteristic of nonlinear continuous time dynamical systems, called the restoration entropy (RE), was introduced. It was shown that it gives the minimal rate at which sensory data should be communicated to an observer to enable it to generate an effective estimate of the current state in real time. Relationships of the RE with the Lyapunov exponents and topological entropy were studied. An approach to two-sided estimation of the RE was elaborated that integrates the strengths of the two Lyapunov's methods. Its utility was demonstrated via study of three classic multi-dimensional chaotic systems. For the Lorentz system in its general (algebraic) form, a closed-form exact formula for the RE was obtained. For the Rössler system in its general form, separate closed-form upper and lower estimates of the RE were established; it was shown that the gap between them $\approx$ 0.0001 for the classic numerical values of the system's parameters. For the Sprott S-system, the RE was computed with nearly the same precision.

Our future research includes an insight into the issue of robustness in the face of dynamic uncertainties and an extension of the findings of the paper on networked scenarios.

Appendix A. Technical facts underlying the proofs in Appendix C, D

Lemma 17. *For any $T > 0$, there are $\varkappa_T, L_T > 0$ such that*

1. *the solution $x(t,a)$ of the ODE in (1) is defined on $[0,T]$ if $a \in [K^\infty]^{\varkappa_T}$, where $[K^\infty]^{\varkappa_T}$ is given by (11), (12);*
2. *$\|x(t,a) - x(t,b)\| \leq L_T\|a-b\|$ whenever $t \in [0,T], a \in [K^\infty]^{\varkappa_T}, b \in K^\infty$;*
3. *for $X(\cdot,\cdot)$ from (11), $\|X(t,a) - X(t,b)\| \to 0$ uniformly over $t \in [0,T], b \in K^\infty$ as $\|a-b\| \to 0$;*
4. *$x(t, a+\Delta a) = x(t,a) + X(t,a)\Delta a + \|\Delta a\|\zeta(a,\Delta a,t)$, where $\zeta(a,\Delta a,t) \to 0$ uniformly over $t \in [0,T], a \in [K^\infty]^{\varkappa_T}$ as $\Delta a \to 0$ if $\varkappa_T$ is small enough.*

Proof. 1. Let $\eta > 0$ be less than the eponymous value from Assumption 5 and let a, b be like in **2** (where $\varkappa_T$ will be specified later on). We denote by $[0, T_a]$ the maximal interval of the form $[0, T_*] \subset [0,T]$ such that $x(t,a)$ is defined for $t \in [0,T_*]$ and $\inf_{z\in K^\infty}\|x(t,a) - z\| \leq \eta\ \forall t \in [0,T_*]$. For $y(t) := x(t,a) - x(t,b), t \in [0,T_a]$, the mean value inequality (Schapira, 2011, Cor. 1.3.5, p. 22) and (1) yield that

$$\|\dot{y}(t)\| = \|f[x(t,a)] - f[x(t,b)]\|$$
$$\leq \max_{\theta\in[0,1]}\left\|\frac{\partial f}{\partial x}[\theta x(t,a) + (1-\theta)x(t,b)]y(t)\right\| \leq \Omega(\eta)\|y(t)\|,$$
$$\text{where}\quad \Omega(\eta) := \sup_{z\in[\mathbf{co}(K^\infty)]^\eta}\left\|\frac{\partial f}{\partial x}[z]\right\| \overset{\text{Assumption 5}}{<} \infty$$
$$\Rightarrow \|y(t)\| \leq \|a-b\| + \Omega(\eta)\int_0^t \|y(s)\|\,ds.$$

Then thanks to Grönwall's inequality,

$$\|y(t)\| \leq \exp[\Omega(\eta)t] \times \|a-b\|. \tag{A.1}$$

Since $a \in [K^\infty]^{\varkappa_T}$, there is $b \in K^\infty$ such that $\|a-b\| < \varkappa_T$. With this b, we get $\|y(t)\| \leq \zeta := \varkappa_T\exp[\Omega(\eta)T]\ \forall t \in [0,T_a]$. Now we pick $\varkappa_T > 0$ so small that $\zeta \leq \eta/2$. Then $\|x(t,a) - x(t,b)\| \leq \eta/2\ \forall t \in [0,T_a]$. So $T_a = T$ since otherwise $x(t,a)$ would exist and $\inf_{z\in K^\infty}\|x(t,a) - z\| \leq \eta$ for $t > T_a, t \approx T_a$, in violation of the definition of T_a.

2. This claim is immediate from (A.1).

3. By Hartman (1964, Cor. 3.1, Ch. 5) and Grönwall's inequality,

$$\frac{\partial X}{\partial t}(t,x_0) = \frac{\partial f}{\partial x}[x(t,x_0)]X(t,x_0),\ X(0,x_0) = I, \tag{A.2}$$
$$\Rightarrow \|X(t,b)\| \leq 1 + \int_0^t \left\|\frac{\partial f}{\partial x}[x(s,b)]X(s,b)\right\|\,ds$$
$$\leq 1 + \Omega(\eta)\int_0^t \|X(s,b)\|\,ds \Rightarrow \|X(t,b)\|$$
$$\leq \exp\big[t\Omega(\eta)\big].$$

Hence for $Y(t) := X(t,a) - X(t,b)$ and $t \in [0,T]$, we have

$$\|\dot{Y}(t)\| \leq \left\|\frac{\partial f}{\partial x}[x(t,a)]Y(t)\right\|$$
$$+ \left\|\left[\frac{\partial f}{\partial x}[x(t,a)] - \frac{\partial f}{\partial x}[x(t,b)]\right]X(t,b)\right\|$$
$$\leq \Omega(\eta)\|Y(t)\| + \left\|\frac{\partial f}{\partial x}[x(t,a)] - \frac{\partial f}{\partial x}[x(t,b)]\right\|\|X(t,b)\|$$
$$\leq \Omega(\eta)\|Y(t)\| + \left\|\frac{\partial f}{\partial x}[x(t,a)] - \frac{\partial f}{\partial x}[x(t,b)]\right\|e^{T\Omega(\eta)}. \tag{A.3}$$

Since $Y(0) = X(0,a) - X(0,b) = 0$, we have

$$\|Y(t)\| = \|Y(t) - Y(0)\| = \left\|\int_0^t \dot{Y}(s)\,ds\right\| \leq \int_0^t \|\dot{Y}(s)\|\,ds.$$

This and (A.3) imply that

$$\|Y(t)\| \leq \Omega(\eta)\int_0^t \|Y(s)\|\,ds + \alpha(t)$$
$$\alpha(t) := e^{T\Omega(\eta)}\int_0^t \left\|\frac{\partial f}{\partial x}[x(s,a)] - \frac{\partial f}{\partial x}[x(s,b)]\right\|\,ds.$$

By the integral version of Grönwall's inequality, we see that

$$\|Y(t)\| \leq \alpha(t) + \int_0^t \alpha(s)\Omega(\eta)e^{\Omega(\eta)(t-s)}ds.$$

Integrating by parts and invoking $\alpha(0) = 0$ yield that

$$\|Y(t)\| \leq e^{T\Omega(\eta)}\int_0^t e^{\Omega(\eta)(t-s)}\left\|\frac{\partial f}{\partial x}[x(s,a)] - \frac{\partial f}{\partial x}[x(s,b)]\right\|\,ds.$$

It remains to note that $x(s,a) - x(s,b) \to 0$ uniformly over $s \in [0,T], b \in K^\infty$ thanks to **2** and that $\partial f/\partial x$ is uniformly continuous on $[K^\infty]^\eta$ by Assumption 5.

4. By Schapira (2011, Cor. 1.3.6, p. 22), we have

$$\|\zeta(a,\Delta a,t,f)\| \leq \max_{\theta\in[0,1]}\left\|\frac{\partial\phi^t}{\partial x}(t, a+\theta\Delta a) - \frac{\partial\phi^t}{\partial x}(t,a)\right\|$$
$$= \max_{\theta\in[0,1]}\|X(t,a+\theta\Delta a) - X(t,a)\|\,. \tag{A.4}$$

Let $a \in [K^\infty]^{\varkappa_T/2}$, $\|\Delta a\| < \varkappa_T/2$. Then $a + \theta\Delta a \in [K^\infty]^{\varkappa_T}\ \forall\theta \in [0,1]$ and so (A.2) $\Rightarrow \|X(t,a+\theta\Delta a)\| \leq e^{\Omega(\eta)T}$ for $t \in [0,T]$. By Schapira (2011, Cor. 1.3.5, p. 22),

$$\|x(t,a+\theta\Delta a) - x(t,a)\| \leq \max_{\theta'\in[0,\theta]}\left\|X(t,a+\theta'\Delta a)\right\|\|\Delta a\|$$
$$\leq \|\Delta a\|e^{\Omega(\eta)T} \to 0 \text{ as } \Delta a \to 0$$

uniformly over $a \in [K^\infty]^{\varkappa_T/2}, t \in [0,T], \theta \in [0,1]$. Hence $\frac{\partial f}{\partial x}[x(t,a+\theta\Delta a)] - \frac{\partial f}{\partial x}[x(t,a)] \to 0$ uniformly over $a \in [K^\infty]^{\varkappa_T/2}$, $t \in [0,T], \theta \in [0,1]$ as $\Delta a \to 0$. This convergence of the coefficients in the linear ODE (A.2) (considered for a and $a := a + \theta\Delta a$, respectively) implies that the solutions $X(t, a+\theta\Delta a) - X(t,a) \to 0$ uniformly over the same set as $\Delta a \to 0$. The proof is completed by (A.4). ■

Lemma 18. *Observability via some channel with capacity c in any of the senses from Definition 3 implies observability in the same sense via any channel with capacity $c' > c$.*

Proof. For all types of observability the proofs are similar; we focus on the fine observability. By Definition 3, there is an observer $\mathbf{O}_c$ that finely observes the system via a channel $\mathbf{CH}_c$ with capacity c. Let $b_\pm(\cdot)$ and $b'_\pm(\cdot)$ be taken from (6) for $\mathbf{CH}_c$ and a channel $\mathbf{CH}_{c'}$ with capacity $c' > c$, respectively, and let $\overline{\tau}$ be an upper bound (from **(c3)**) on the transmission time for $\mathbf{CH}_c$. Since $b'_-(T) \geq b_+(T+\overline{\tau})$ for $T \approx \infty$ by (6), any message transferred via $\mathbf{CH}_c$ within any time interval of duration $T + \overline{\tau}$ can be transferred within any time interval of duration T via $\mathbf{CH}_{c'}$.

We introduce the observer $\mathbf{O}_{c'}$ that operates via $\mathbf{CH}_{c'}$ as follows. At any time $\theta_j = jT, j = 0, 1, \ldots$, the coder of $\mathbf{O}_{c'}$ builds the prognosis $x(t), t \in [\theta_j, \theta_{j+1}]$ by integrating the ODE from (1). This prognosis is used to instantly generate the sequence of dispatch times (4) and messages (3) by running the coder of $\mathbf{O}_c$. At time θ_j, the coder of $\mathbf{O}_{c'}$ thus becomes aware of these sequences up to the future time θ_{j+1}. The total message that arrives at the decoder of $\mathbf{O}_c$ during $[\theta_j, \theta_{j+1}]$ is transferred via $\mathbf{CH}_c$ within the time interval $[\theta_j - \overline{\tau}, \theta_{j+1}]$ and so can be transmitted across $\mathbf{CH}_{c'}$ within the time interval $[\theta_j, \theta_{j+1}]$. This is exactly what the coder of $\mathbf{CH}_{c'}$ does, thus enabling the decoder $\mathbf{D}_c$ of $\mathbf{O}_c$ to work, without any alteration, on the basis of data transferred via $\mathbf{O}_{c'}$. The decoder of $\mathbf{O}_{c'}$ takes this opportunity to run $\mathbf{D}_c$, and thus acquires the estimate $\widehat{x}(\theta_{j+1})$ of $\mathbf{O}_c$ at time θ_{j+1}.

To build its own estimate $\widehat{x}'(t)$, the decoder of $\mathbf{O}_{c'}$ uses $\widehat{x}(t) =: \widehat{x}'(t+)$ at times $t = \theta_j$, and integrates the ODE from (1) with $x(0) = \widehat{x}'(\theta_j+)$ between them $t \in [\theta_j, \theta_{j+1})$.

$$\|x(t) - \widehat{x}'(t)\| \overset{\mathbf{2} \text{ in Lemma 17}}{\leq} L_T\|x(\theta_j) - \widehat{x}(\theta_j)\| \overset{(9)}{\leq} GL_T\delta e^{-g\theta_j}$$
$$= GL_T\delta e^{-gT\lfloor t/T\rfloor} \leq \left[GL_Te^{gT}\right]\delta e^{-gt}.$$

Definition 2 completes the proof. ■

Lemma 19. *Let $A \in \mathbb{R}^{n\times n}$ and let the operator of multiplication by A transform B_0^δ into a set that can be covered by balls $B_{x^1}^\varepsilon, \dots, B_{x^N}^\varepsilon$ with the radius of ε. Then their number*

$$N \geq \left[\delta/(\varepsilon\sqrt{n})\right]^d \omega_d(A) \qquad \forall d \in [1:n]. \tag{A.5}$$

Proof. If $\omega_d(A) = 0$, the inequality with this d is evident. Let $\omega_d(A) > 0$ and let $A = U\,\mathbf{diag}\,[\alpha_1(A), \dots, \alpha_n(A)]V$ be the singular value decomposition, where U and V are unitary $n \times n$-matrices. Since $O := \{x = \mathbf{col}\,(x_1, \dots, x_n) \in \mathbb{R}^n : |x_i|<\delta/\sqrt{n}\} \subset B_0^\delta = VB_0^\delta$, the image

$$Q := \mathbf{diag}\,[\alpha_1(A), \dots, \alpha_n(A)]O$$

is covered by the balls $U^{-1}B_{x^j}^\varepsilon = B_{U^{-1}x^j}^\varepsilon, j = 1, \dots, N$. Now we introduce the basis $e_i := \mathbf{col}(0, \dots, 1, \dots, 0), i = 1, \dots, n$, where 1 is in the ith position, the orthogonal projection π_d of $\mathbb{R}^n$ onto the linear subspace L_d spanned by $e_1, \dots, e_d$, and the Lebesgue measure $\mathbf{mes}_d$ in L_d. Since π_dQ is covered by $\pi_dB_{U^{-1}x^j}^\varepsilon, j = 1, \dots, N$, we have

$$\mathbf{mes}_d\,[\pi_dQ] \leq \sum_{j=1}^N \mathbf{mes}_d\left[\pi_dB_{U^{-1}x^j}^\varepsilon\right],$$
$$\pi_dQ = \left\{x = \sum_{i=1}^d \theta_ie_i : |\theta_i|<\alpha_i(A)\delta/\sqrt{n}\right\}.$$

So $\mathbf{mes}_d\,[\pi_dQ] = \omega_d(A)2^d\delta^dn^{-d/2}$, whereas $\pi_dB_{U^{-1}x^j}^\varepsilon$ is an open ball in L_d with a radius of ε and so $\mathbf{mes}_d\left[\pi_dB_{U^{-1}x^j}^\varepsilon\right] \leq \varepsilon^d2^d$. Summing up, we arrive at (A.5). ■

Appendix B. Observer underlying the proofs in Appendix C

Now we introduce an observer that backs our proofs. This observer is built of the following set of ingredients.

Definition 20. An *image covering structure* consists of $T, \delta_* > 0$, a δ-parametric sequence $\{r_j(\delta)\}_{j=0}^\infty \subset (0, \infty)$ (where $\delta \in (0, \delta_*)$, $r_0(\delta) = \delta$), and a mapping of any triplet (j, a, δ) (where $j = 0, 1, \dots$, $a \in K_j := K(jT)$, and $\delta \in (0, \delta_*)$) into a cover $\mathfrak{C}_{a,\delta}^j$ of $\phi^T[B_a^{r_j(\delta)} \cap K_j]$ with open $r_{j+1}(\delta)$-balls centered in K_{j+1}. The *size of the structure* is defined as the minimal N (either natural or ∞) such that the size of the cover $\mathfrak{C}_{a,\delta}^j$ does not exceed N for any j, a, δ.

Lemma 21. *For any image covering structure of a finite size N and a communication channel for which*

$$b_-(T) \geq \log_2 N + 1 \quad (b_-(\cdot) \text{ is taken from (6)}), \tag{B.1}$$

there is an observer (3)–(5) that acts via this channel and

$$\|x(t) - \widehat{x}(t)\| \leq L_Tr_{\lfloor t/T\rfloor}(\delta) \quad \forall t \geq 0, \delta \in (0, \delta_*) \tag{B.2}$$

whenever (2) holds. Here L_T is taken from **2** *in Lemma 17.*

Proof. Since the decoder is fully driven by the messages from the coder, the latter can and (in our design) does replicate the estimates $\widehat{x}(t)$ generated by the former. Let (2) hold with $\delta \in (0, \delta_*)$. At any time $t = \theta_j = jT, j = 0, 1, \dots$, the coder checks whether $x(t) \in B_{\widehat{x}(t)}^{r_j(\delta)}$. If not so, an "alarm" message ✠ is sent to the decoder. Otherwise $x(t+T) \in \phi^T\left[B_{\widehat{x}(t)}^{r_j(\delta)}\right]$ and the coder determines a ball B from the cover $\mathfrak{C}_{\widehat{x}(t),\delta}^j$ that contains $x(t+T)$ and then sends a binary code of the index of the ball B across the channel within the time interval $\Delta_j := [\theta_j, \theta_{j+1})$. This calls for communication of $\log_2 N$ bits since the size of the cover does not exceed N by Definition 20. The requested data transmission is executable by the definition of $b_-(T)$ and (B.1) (where one more bit is reserved for possible communication of ✠).

As t runs inside any interval Δ_j, the proposed decoder merely integrates the ODE from (1) $\frac{d\widehat{x}(t)}{dt} = f[\widehat{x}(t)]$. As t arrives at the end θ_{j+1} of Δ_j, the decoder checks whether an information about a ball B has arrived during Δ_{j-1}. If so, $\widehat{x}(\theta_{j+1})$ is instantly removed to the center of B.

It is easy to see that first, ✠ is never sent and second,

$$\|x(\theta_j) - \widehat{x}(\theta_j+)\| < r_j(\delta) \qquad \forall j = 0, 1, 2, \dots.$$

The proof is completed by **2** in Lemma 17. ■

Appendix C. Proof of Theorem 8 and Remark 10

The first relations in (16), (17) are established mostly in the same way as similar ones in Matveev and Pogromsky (2016). Now we present full proofs of (16), (17) for completeness and convenience of the reader.

An initial part of proving (16) is focused on the second inequality, with addressing the first one in brackets ⟨...⟩. We first pick $c > \mathfrak{R}_{\text{ro}}$ ⟨$c > \mathfrak{R}_\text{o}$⟩. By the definition of $\mathfrak{R}_{\text{ro}}$ ⟨$\mathfrak{R}_\text{o}$⟩, there is an observer that regularly observes ⟨observes⟩ the system via a channel with capacity $c' < c$. By Definition 2, there are $\delta_*, G > 0$ ⟨for any $\varepsilon > 0$, there is δ_*⟩ such that $\|x(t) - \widehat{x}(t)\| \leq G\delta\ \forall t \geq 0$ ⟨such that $\|x(t){-}\widehat{x}(t)\| \leq \varepsilon\ \forall t \geq 0$⟩ if $\delta < \delta_*$ and $a := \widehat{x}(0) \in K$, $x(0) \in B_a^\delta \cap K$. By (6), there also exists T_* such that $b_+(t)/t < c\ \ \forall t \geq T_*$. The closed ball $\overline{B_0^G}$ can be covered by a finite number L of open unit balls $B_{y_1}^1, \dots, B_{y_L}^1$ centered at some points $y_1, \dots, y_L$; then the closed ball $\overline{B_b^{G\delta}}$ is covered by the same number L of open balls $B_{\delta y_1+b}^\delta, \dots, B_{\delta y_L+b}^\delta$.

Let $x(0)$ run over $B_a^\delta \cap K$ and let $T \geq T_*$. By (5), $\widehat{x}(T)$ takes no more values than the message transferred via the channel within $[0, T]$. The number of them does not exceed $2^{b_+(T)} < 2^{cT}$ by **(c4)**. The closed $(G\delta)$-balls centered at these $\widehat{x}(T)$'s cover $\phi^T(B_a^\delta \cap K)$; every of these balls is covered by L open δ-balls. Hence $p(T, a, \delta) \leq L2^{cT} \Rightarrow \log_2 p(T, a, \delta)/T \leq c + (\log_2 L)/T\ \forall a \in K, \delta \in (0, \delta_*), T \geq T_*$. Then by (15), $H_{\text{res}} \leq c\ \forall c > \mathfrak{R}_{\text{ro}} \Rightarrow$ "the second inequality in (16)".

To prove the first one, we cover the compact set K by finitely many open δ-balls $B_{a_1}^\delta, \dots, B_{a_M}^\delta$, $a_i \in K$. As $x(0)$ runs over $B_{a_i}^\delta$, the estimate $\widehat{x}(t)$, treated as a function of $t \in [0, T]$, ranges over a set Q_i with no more than 2^{cT} elements, as has been just shown. Since the set $Q := \cup_iQ_i$ is (T, ε)-spanning by Definition 6 (and Remark 9), $q(T, \varepsilon) \leq M2^{cT}\ \forall T \geq T_*$ and so $H(f, K) \leq c$ by (14). It remains to let $c \to \mathfrak{R}_\text{o}+$, which means that c approaches $\mathfrak{R}_\text{o}$ from above.

Now let K be positively invariant. Thanks to (10) and (16), it suffices to show that $\mathfrak{R}_\text{o} \leq H$ and $\mathfrak{R}_{\text{fo}} \leq H_{\text{res}}$ to prove (17). We start with the second inequality and pick $c > c' > H_{\text{res}}$. By (15), there exists T_* and a map $T \in [T_*, \infty) \mapsto \delta_T > 0$ such that for all $T \geq T_*, \delta \in (0, \delta_T), a \in K$,

$$T^{-1}\log_2 p(T, a, \delta) < c'.$$

Let M be the number of open $(1/4)$-balls needed to cover the unit closed ball $\overline{B_0^1}$. Since $T^{-1}\log_2 M \to 0$ as $T \to \infty$ and $c' < c$, properly increasing T_* (if necessary) ensures that for $T \geq T_*, \delta \in (0, \delta_T), a \in K$, we have

$$T^{-1}[\log_2 p(T, a, \delta) + \log_2 M] < c$$

So if $a \in K(jT) \subset K$, the image $\phi^T[B_a^\delta \cap K(jT)]$ can be covered by $M \cdot p(T,a,\delta)$ open η-balls with $\eta := \delta/4$; their centers can be removed to $K[(j+1)T]$ via altering $\eta := \delta/2$. By substituting $r_j(\delta) := 2^{-j}\delta \in (0,\delta_T)$ in place of δ here, defining $\mathfrak{C}^j_{a,\delta}$ as the resultant cover, we get an image covering structure of size $\leq N := \lfloor 2^{cT} \rfloor$ by Definition 20. Let $\mathfrak{C}$ be a channel with $b_-(T) = \log_2 N + 1$ and capacity $(\log_2 N + 1)/T$. By Lemma 21, there is an observer (3)–(5) that acts via $\mathfrak{C}$ and meets (B.2). Since the r.h.s. of (B.2) equals $L_T r_{\lfloor t/T \rfloor}(\delta) = L_T 2^{-\lfloor t/T \rfloor}\delta \leq 2L_T 2^{-t/T}\delta = Ge^{-gt}\delta$ with $g := \ln 2/T$ and $G := 2L_T$, we see that the system is finely observable via $\mathfrak{C}$ by Definition 3. Hence $\mathfrak{R}_{fo} \leq (\log_2 N + 1)/T = (\log_2 \lfloor 2^{cT} \rfloor + 1)/T \leq c + 1/T \ \forall T \geq T_*$. Consecutively letting $T \to \infty$ and $c \to H_{res}+$ yields that $\mathfrak{R}_{fo} \leq H_{res}$.

To show that $\mathfrak{R}_o \leq H$, we put $\varepsilon_k := 1/k$, and pick a channel $\mathfrak{C}$ with capacity $c > c' > H$. By (14) and Definition 6, for any large enough $T \geq 0$, there is a (T,ε_k)-spanning set $Q_k(T)$ whose size $\leq 2^{c'T}$. So within any time interval of a proper duration T_k, the coder can make the decoder aware of the choice among the elements of $Q_k(T_k)$ by sending messages via $\mathfrak{C}$. At any time $\theta_j := jT_k$, the coder computes $x(t), t \in [\theta_j, \theta_{j+2}]$ via integration of the ODE from (1), picks $x^\star(\cdot) \in Q_k(T_k)$ such that $\|x(\theta_{j+1}+\tau) - x^\star(\tau)\| \leq \varepsilon_k \ \forall \tau \in [0,T_k]$, and "communicates" this $x^\star(\cdot)$ to the decoder. On receiving $x^\star(\cdot)$ at $t = \theta_{j+1}$, the decoder puts $\widehat{x}(t) := x^\star(t-\theta_{j+1}) \ \forall t \in [\theta_{j+1}, \theta_{j+2}]$. Finally, we pick a decaying sequence $\{\delta_k > 0\}$ such that $\delta_k < L_{T_k}^{-1}\varepsilon_k$. Based on $\delta \in (0,\delta_1)$ in (3)–(5), the observer finds k such that $\delta_{k+1} < \delta \leq \delta_k$ and employs the above design with this k for the coder and decoder on $t \in [0,\infty)$ and $t \in [T_k,\infty)$, respectively. On $[0,T_k]$, the decoder integrates the ODE in (1) from the initial state $\widehat{x}(0)$. Definitions 1, 3 and letting $c \to H+$ assure that $\mathfrak{R}_o \leq H$. ■

Proof of Remark 10. The number $p(T,a,\delta)$ related to the ith flow $(i = 1,2)$ is marked with $_i$. There is $L_i > 0$ such that $\|h_i(x) - h_i(y)\| \leq L_i\|x-y\| \ \forall x,y \in K_i^\infty, i = 1,2$ for $h_1 := h, h_2 := h^{-1}$. For $\delta > 0, a \in K_2$, we have $h^{-1}(B_a^\delta \cap K_2) \subset B^{L_2\delta}_{h^{-1}a} \cap K_1$. So $\mathcal{K} := [\phi_1^T \circ h^{-1}](B_a^\delta \cap K_2) \subset K_1^\infty$ can be covered by $p_1(T, h^{-1}a, L_2\delta)$ open δ-balls. Via doubling δ, their centers can be removed to K_1^∞. Hence $\phi_2^T(B_a^\delta \cap K_2) = h(\mathcal{K})$ can be covered by $p_1(T,h^{-1}a, L_2\delta)$ open $(2L_1\delta)$-balls. Let M be the number of open 1-balls needed to cover $B_0^{2L_1}$. Then M open δ-balls are enough to cover any closed ball with a radius of $2L_1\delta$. Hence

$$p_2(T,a,\delta) \leq Mp_1(T,h^{-1}a, L_2\delta),$$

$$\overline{\lim_{T\to\infty}} \frac{1}{T} \overline{\lim_{\delta\to 0}} \sup_{a\in K_2} \log_2 p_2(T,a,\delta)$$

$$\leq \overline{\lim_{T\to\infty}} \frac{1}{T} \overline{\lim_{\delta\to 0}} \sup_{a\in K_1} \log_2 p_1(T,a,\delta)$$

and so the RE of the second flow does not exceed that of the first one by (15). The converse is established likewise. ■

Appendix D. Proof of Theorem 11

If $a \in K^\infty$ and $\|x-a\| < \delta$, (4) in Lemma 17 yields that

$$\phi^T(x) - X(T,a)x - b \in B_0^{\delta\mu_T(\delta)}, \tag{D.1}$$

where $b := \phi^T(a) - X(T,a)a$ and $\mu_T(\delta) \to 0$ as $\delta \to 0$. Then by Pogromsky and Matveev (2011, Prop. 8.5) and (41) in Pogromsky and Matveev (2011), the number $\mathcal{N}$ of open ε-balls needed to cover $\phi^T(B_a^\delta)$

$$\mathcal{N} \leq \prod_{i=1}^n \left\lceil \delta/\varepsilon \{\alpha_i[X(T,a)] + \mu_T(\delta)\}\sqrt{n}\, \right\rceil \tag{D.2}$$

$$\leq \max\left\{2^n n^{\frac{n}{2}} \max_{d\in[1:n]} (\delta/\varepsilon)^d \prod_{i=1}^d \{\alpha_i[X(T,a)] + \mu_T(\delta)\}\,;\, 1\right\}.$$

Taking $\varepsilon := \delta$ yields that

$$\overline{\lim_{\delta\to 0}} \max_{a\in K} p(T,a,\delta)$$

$$\leq 2^n n^{\frac{n}{2}} \max_{a\in K} \max\left\{\max_{d\in[1:n]} \prod_{i=1}^d \alpha_i[X(T,a)]\,;\,1\right\}.$$

Whence by invoking (15), we see that

$$H_{res} \leq \overline{\lim_{T\to\infty}} \frac{1}{T}\log_2 \max_{a\in K}\max\left\{\max_{d\in[1:n]}\prod_{i=1}^d \alpha_i[X(T,a)]\,;\,1\right\},$$

which implies the second inequality in (20) thanks to (31).

By taking $a \in K(jT)$ and $\varepsilon := \delta/4$ in (D.2), we see that to cover $\phi^T[B_a^\delta \cap K(jT)]$ by open $(\delta/4)$-balls, no more than

$$\max\left\{2^{3n} n^{\frac{n}{2}} \max_{d\in[1:n]} \prod_{i=1}^d \{\alpha_i[X(T,a)] + \mu_T(\delta)\}\,;\,1\right\}$$

balls are needed; via doubling their radius $\delta/4 \mapsto \delta/2$, their centers can be moved in $K[(j+1)T]$. Putting $r_j(\delta) := 2^{-j}\delta$ in place of δ here shows that there is a cover $\mathfrak{C}^j_{a,\delta}$ of $\phi^T[B_a^{r_j(\delta)} \cap K(jT)]$ with open $r_{j+1}(\delta)$-balls centered in $K[(j+1)T]$ whose size does not exceed

$$N := \sup_{a\in K^\infty} \max\left\{2^{3n+1} n^{\frac{n}{2}} \max_{d\in[1:n]} \prod_{i=1}^d \alpha_i[X(T,a)]\,;\,1\right\}$$

for $\delta \approx 0$. We thus acquire an image covering structure of size $\leq N$ by Definition 20. Retracing the arguments from the penultimate paragraph in the proof of Theorem 8 shows that $\mathfrak{R}_{fo} \leq (\log_2 N + 1)/T$. Letting $T \to \infty$ yields (21).

To prove the first inequality in (20), we note that $\phi^T(B_a^\delta \cap K)$ can be covered by $p(T,a,\delta)$ open δ-balls by the definition of $p(T,a,\delta)$. Let $a \in \mathbf{int}K$ and $\delta_a > 0$ be so small that $B_a^{\delta_a} \subset K$. Then $\phi^T(B_a^\delta)$ can be covered by $p(T,a,\delta)$ open δ-balls if $\delta \in (0,\delta_a)$. Then by (D.1), we see that the image $X(T,a)B_0^\delta$ can be covered by $p(T,a,\delta)$ open ε-balls, where $\varepsilon := \delta[1+\mu_T(\delta)]$. So (A.5) yields that for all $\delta \in (0,\delta_a)$,

$$p(T,a,\delta) \geq \max_{d\in[1:n]} \omega_d[X(T,a)] \left(\frac{1}{[1+\mu_T(\delta)]\sqrt{n}}\right)^d.$$

Then for any $\delta_* > 0$, $\delta \in (0,\delta_*)$ and $K^-(\delta_*) := \{x : B_x^{\delta_*} \subset K\}$, we have

$$\max_{a\in K} p(T,a,\delta) \geq \max_{a\in K^-(\delta_*)} p(T,a,\delta)$$

$$\geq \sup_{a\in K^-(\delta_*)} \max_{d\in[1:n]} \max\left\{\omega_d[X(T,a)]\left(\frac{1}{[1+q_T(\delta)]\sqrt{n}}\right)^d ; 1\right\}$$

$$\Rightarrow \overline{\lim_{\delta\to 0}} \max_{a\in K} p(T,a,\delta)$$

$$\geq \sup_{a\in K^-(\delta_*)} \max_{d\in[1:n]} \max\left\{\omega_d[X(T,a)]\, n^{-\frac{d}{2}} ; 1\right\}.$$

By noting that $\bigcup_{\delta_*>0} K^-(\delta_*) = \mathbf{int}\, K$, we see that

$$\overline{\lim_{\delta\to 0}} \max_{a\in K} p(T,a,\delta)$$

$$\geq \sup_{a\in \mathbf{int}\,K} \max_{d\in[1:n]} \max\left\{\omega_d[X(T,a)]\, n^{-d/2} ; 1\right\}$$

Here $X(T,a)$ continuously depends on $a \in K$, whereas the singular values $\alpha_i(A)$ and so $\omega_d(A)$ continuously depend on the matrix A. It follows that the expression after $\sup_{a\in\mathbf{int}\,K}$ continuously depends on a and so this sup equals $\sup_{a\in\overline{\mathbf{int}\,K}}$. Hence by (15),

$$H_{res}(f,K) \geq \overline{\lim_{T\to\infty}} \max_{a\in\overline{\mathbf{int}\,K}} \max_{d\in[1:n]} \left\{\frac{\log_2 \omega_d[X(T,a)]}{T}\right\}_+,$$

which yields the first inequality in (20). ■

References

Antsaklis, P., & Baillieul, J. (2007). Special issue on the technology of networked control systems. *Proceedings of the IEEE*, *95*(1).
Barreira, L., & Valls, C. (2017). Lyapunov regularity via singular values. *Transactions of the American Mathematical Society*, published electronically, May 30.
Boichenko, V. A., Leonov, G. A., & Reitman, V. (2005). *Dimension theory for ordinary differential equations*. Wiesbaden, Germany: Teubner Verlag.
Colonius, F. (2012). Minimal bit rates and entropy for exponential stabilization. *SIAM Journal on Control and Optimization*, *50*(5), 2988–3010.
Colonius, F., & Kawan, C. (2009). Invariance entropy for control systems. *SIAM Journal on Control and Optimization*, *48*(3), 1701–1721.
Colonius, F., & Kawan, C. (2011). Invariance entropy for outputs. *Mathematics of Control, Signals, and Systems*, *22*(3), 203–227.
Colonius, F., Kawan, C., & Nair, G. (2013). A note on topological feedback entropy and invariance entropy. *Systems & Control Letters*, *62*(5), 377–381.
Dinaburg, E. (1971). On the relations among various entropy characteristics of dynamical systems. *Mathematics of the USSR-Izvestiya*, *5*(2), 337–378.
Downarowicz, T. (2011). *Entropy in dynamical systems*. NY: Cambridge University Press.
Ermentrout, G. B. (1984). Periodic doublings and possible chaos in neural models. *SIAM Journal of Applied Mathematics*, *44*, 80–95.
Ge, X., Yang, F., & Han, Q. (2017). Distributed networked control systems: A brief overview. *Information Sciences*, *380*, 117–131.
Gupta, R., & Chow, M. (2010). Networked control system: Overview and research trends. *IEEE Transactions on Industrial Electronics*, *57*(7), 2527–2535.
Hagihara, R., & Nair, G. N. (2013). Two extensions of topological feedback entropy. *Mathematics of Continuous, Signals, and Systems*, *25*(4), 473–490.
Haller, G. (2000). Finding finite-time invariant manifolds in two-dimensional velocity fields. *Chaos*, *10*, 99–108.
Haller, G. (2001). Distinguished material surfaces and coherent structures in 3d fluid flows. *Physica D*, *149*, 248–277.
Hartman, P. (1964). *Ordinary differential equations*. NY: Wiley.
Katok, A. (2007). Fifty years of entropy in dynamics: 1958–2007. *Journal of Modern Dynamics*, *1*(4), 545–596.
Katok, A., & Hasselblatt, B. (1995). *Introduction to the modern theory of dynamical systems*. Cambridge, UK: Cambridge Univ. Press.
Kawan, C. (2011a). Invariance entropy of control sets. *SIAM Journal on Control and Optimization*, *49*(2), 732–751.
Kawan, C. (2011b). Lower bounds for the strict invariance entropy. *Nonlinearity*, *24*(7), 1910–1936.
Kawan, C. (2011c). Upper and lower estimates for invariance entropy. *Discrete and Continuous Dynamical Systems*, *30*(1), 169–186.
Kuznetsov, N. V., Mokaev, T. N., & Vasilyev, P. A. (2014). Numerical justification of Leonov conjecture on Lyapunov dimension of Rossler attractor. *Communications in Nonlinear Science and Numerical Simulations*, *19*, 1027–1034.
Lai, Y., & Tél, T. (2011). *Transient chaos*. NY: Springer.
Leonov, G. A., & Alexeeva, T. A. (2015). Lyapunov functions in estimates of attractor dimensions for generalize Rössler systems. *Doklady Mathematics*, *91*, 5–8.
Leonov, G. A., Kuznetsov, N. V., Korzhemanova, N. A., & Kusakin, D. V. (2016). Lyapunov dimension formula for the global attractor of the Lorenz system. *Communications in Nonlinear Science and Numerical Simulation*, *41*(1), 84–103.
Leonov, G. A., Pogromsky, A. Y., & Starkov, K. E. (2011). The dimension formula for the Lorenz attractor. *Physics Letters A*, *375*, 1179–1182.
Leonov, G. A., Pogromsky, A. Yu., & Starkov, K. E. (2012). Erratum. The dimension formula for the Lorenz attractor. *Physics Letters A*, *376*(1), 34723474.
Liberzon, D., & Hespanha, J. P. (2005). Stabilization of nonlinear systems with limited information feedback. *IEEE Transactions on Automatic Control*, *50*(6), 910–915.
Liberzon, D., & Mitra, S. (2016). Entropy and minimal data rates for state estimation and model detection. In *Proceedings of the 19th international conference on hybrid systems: Computation and control* (pp. 247–256). Vienna, Austria.
Mahmoud, M. S. (2014). *Control and estimation methods over communication networks*. Heidelberg: Springer.
Matveev, A. S., & Pogromsky, A. Y. (2016). Observation of nonlinear systems via finite capacity channels: constructive data rate limits. *Automatica*, *70*, 217–229.
Matveev, A., & Pogromsky, A. (2017). Two lyapunov methods in nonlinear state estimation via finite capacity communication channels. *IFAC-PapersOnLine*, *50*(1), 4132–4137.
Matveev, A. S., & Savkin, A. V. (2005). Multi-rate stabilization of linear multiple sensor systems via limited capacity communication channels. *SIAM Journal on Optimization and Control*, *44*(2), 584–618.
Matveev, A. S., & Savkin, A. V. (2009). *Estimation and control over communication networks*. Boston: Birkhäuser.
Nair, G. N., Evans, R. J., Mareels, I. M. Y., & 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., & Evans, R. J. (2007). Feedback control under data rate constraints: An overview. *Proceedings of the IEEE*, *95*(1), 108–137.
de Persis, C. (2005). n-Bit stabilization of n-dimensional nonlinear systems in feedforward form. *IEEE Transactions on Automation Control*, *50*(3), 299–311.
Pogromsky, A. Y., & Matveev, A. S. (2011). Estimation of topological entropy via the direct Lyapunov method. *Nonlinearity*, *24*, 1937–1959.
Pogromsky, A. Yu., & Matveev, A. S. (2013). A non-quadratic criterion for stability of forced oscillations. *Systems & Control Letters*, *62*(5), 408–412.
Pogromsky, A., & Matveev, A. (2016a). Data rate limitations for observability of nonlinear systems. *IFAC-PapersOnLine*, *49*(14), 119–124.
Pogromsky, A. Y., & Matveev, A. S. (2016b). Stability analysis via averaging functions. *IEEE Transactions on Automatic Control*, *61*(4), 1081–1086.
Pogromsky, A., Matveev, A., Chaillet, A., & Rüffer, B. (2013). Input-dependent stability analysis of systems with saturation in feedback. In *Proc. 52nd IEEE conference on dcision and control* (pp. CD–ROM). Florence,.
Postoyan, R., van de Wouw, N., Nešić, Dragan, & Heemels, W. P. Maurice H. (2014). Tracking control for nonlinear networked control systems. *IEEE Transactions on Automatic Control*, *59*, 1539–1554.
Rössler, O. E (1979). Continuous chaos four prototype equations. *Annals of the New York Academy of Sciences*, *316*, 376–392.
Ruelle, D. (1979). Ergodic theory of differentiable dynamical systems. *Publications Mathématiques de l'Institut des Hautes Études Scientifiques*, *50*(1), 27–58.
Savkin, A. V. (2006). Analysis and synthesis of networked control systems: topological entropy, observability, robustness, and optimal control. *Automatica*, *42*(1), 51–62.
Schapira, P. (2011). *Differential calculus*. Paris, France: Universite Paris VI, https://webusers.imj-prg.fr/~pierre.schapira/lectnotes/CalD.pdf.
Schreiber, S. (1998). On growth rates of subadditive functions for semiflows. *Journal of Differential Equations*, *148*, 334–350.
Sibai, H., & Mitra, S. (2017). Optimal data rate for state estimation of switched nonlinear systems. In *Proceedings of the 20th international conference on hybrid systems: computation and control* (pp. 71–80). Pittsburgh, PA.
Silva, A. Da, & Kawan, C. (2016). Invariance entropy of hyperbolic control sets. *Discrete and Continuous Dynamical Systems*, *36*(1), 97–136.
Skokos, C. (2010). The Lyapunov characteristic exponents and their computation, arXiv:0811.0882.
Sprott, J. C. (1994). Some simple chaotic flows. *Physical Review E*, *50*, R647–R650.
Thiffeault, J. L. (2002). Derivatives and constraints in chaotic flows: asymptotic behaviour and a numerical method. *Physica D: Nonlinear Phenomena*, *172*(1–4), 139–161.
Wolf, A., Swift, J. B., Swinney, H. L., & Vastano, J. A. (1985). Determining Lyapunov exponents from a time series. *Physica D: Nonlinear Phenomena*, *16*(3), 285–317.
Yüksel, S., & Basar, T. (2013). *Stochastic networked control systems: Stabilization and optimization under information constraints*. Boston: Birkhäuser.

Alexey S. Matveev was born in Leningrad, USSR, in 1954. He received the M.S. and Ph.D. degrees in 1976 and 1980, respectively, both from the Leningrad University. Currently, he is a professor of the Department of Mathematics and Mechanics, Saint Petersburg University. His research interests include estimation and control over communication networks, hybrid dynamical systems, and navigation and control of mobile robots.

Alexander Yu. Pogromsky was born in Saint Petersburg, Russia, on March 1, 1970. He received the M.Sc degree (cum laude) from the Baltic State Technical University, Russia in 1991 and Ph.D (Candidate of Science) degree from the St.Petersburg Electrotechnical University in 1994. From 1995 till 1997 he was with the Laboratory "Control of Complex Systems" (Institute for Problems of Mechanical Engineering, St. Petersburg, Russia). From 1997 to 1998 he was a research fellow with the Department of Electrical Engineering, division on Automatic Control, Linkoping University, Sweden. In the period of 1999–2001 he was a research fellow with the department of Electrical Engineering of Eindhoven University and after that he joined the department of Mechanical Engineering of Eindhoven University of Technology, the Netherlands where he works till now.

He occupies visiting professor positions in St. Petersburg State University and ITMO University.

His research interests include theory of nonlinear, adaptive and robust control, nonlinear oscillations.