Modular Computation of Restoration Entropy for Networks of Systems: A Dissipativity Approach

Mahendra Singh Tomar and Majid Zamani, *Senior Member, IEEE*

***Abstract*—The problem of state estimation based on information received over a finite bit rate channel gives rise to the study of minimal bit rate above which state can be estimated with any desired accuracy. In the past few years, researchers have studied the minimal average bit rate which is sufficient enough for state estimation such that the estimation error stays within a given factor of its initial value. The notion of restoration entropy characterizes this type of bit rate. Recent results proposed numerical schemes to estimate restoration entropy by the computation of singular values of the linearized systems. Such schemes are either complex to implement or suffer severely from computational complexity and the size of the state dimension. In this letter, we describe a *modular* approach to compute an upper bound of the restoration entropy of a large network by decomposing the network to an interconnection of smaller subsystems. Then, we formulate a distributed optimization problem which is solved for each subsystem separately and then their optimization results are composed to get an upper bound of the restoration entropy for the overall network. We illustrate the effectiveness of our results by two examples.**

***Index Terms*—Networked control systems, entropy, estimation.**

I. Introduction

RECENT progresses in digital fabrication and technology are promoting great reduction in size and cost of sensors and actuators and, hence, contribute to increasing spatial distribution of the system components including plants, sensors, controllers, and actuators. Such networks of distributed components exchanging information over some communication channels are referred to as networked control systems. A networked control system such as an autonomous vehicle, transportation network, a swarm of drones or robotic-assisted surgery often involves a shared digital communication channel for the transfer of information from a sensor to a decoder located close to the controller. The finite data rate of the communication channel results in the inexactness of information about the system state at the location of the decoder. For unstable dynamics, this inexactness/uncertainty increases over time. To ensure the satisfaction of a control specification, it is *necessary* to transmit state information at a rate which is larger than the rate of the growth of this uncertainty. This gives rise to the study of the minimal data rate that permits satisfaction of the desired control objectives; see for example [1], [2], [3] or the surveys on control under communication constraints in [4], [5].

For many control tasks, the minimal data rates have been characterized in terms of some notions of entropy that are intrinsic quantities of the system dynamics and are independent of any particular choice of coder-decoder. Various notions of entropy have been introduced for different control tasks; for example, see [6] for exponential stabilization, [7], [8] for controlled invariance, [9] for reachability, [3] for invariance in networks, [10] for stochastic stabilization, and [2], [11], [12] for state estimation.

The smallest data rate or channel capacity, above which the state of a dynamical system can be estimated with an arbitrary precision, can be described by the classical notion of topological entropy [13]. Estimation of topological entropy for nonlinear systems is a hard problem (see, e.g., [14], [15]). This, along with the lack of robustness of the topological entropy with respect to the system parameters resulted in the study of three types of observation criteria [16] (characterized based on the way estimation error progresses over time) and the introduction of restoration entropy in [17] for continuous-time systems which was adapted to discrete-time systems in [18]. Roughly speaking, for a forward invariant set K, the restoration entropy quantifies the minimal data rate above which the state of a system can be estimated so that the estimation accuracy is not just preserved but can also be improved over time.

For a compact set satisfying certain properties, a closed-form expression in terms of the singular values of the linearized system, described in [19] as an upper bound for the topological entropy, also specifies the restoration entropy. A numerical scheme to compute the upper bound utilizing semidefinite optimization techniques was described in [20]. A different formulation of the restoration entropy that does

Manuscript received 18 March 2022; revised 12 May 2022; accepted 5 June 2022. Date of publication 21 June 2022; date of current version 30 June 2022. This work was supported in part by NSF under Grant CMMI-2013969 and DFG under Grant ZA 873/5-2. Recommended by Senior Editor J. Daafouz. *(Corresponding author: Mahendra Singh Tomar.)*

Mahendra Singh Tomar is with the Computer Science Department, University of Colorado Boulder, Boulder, CO 80309 USA (e-mail: mahendra.tomar@colorado.edu).

Majid Zamani is with the Computer Science Department, University of Colorado Boulder, Boulder, CO 80309 USA, and also with the Computer Science Department, LMU Munich, 80539 Munich, Germany (e-mail: majid.zamani@colorado.edu).

Digital Object Identifier 10.1109/LCSYS.2022.3184824

not involve any temporal limit was proposed in [21] which characterizes its estimation with any desired accuracy by a suitable choice of the Riemannian metric. A subgradient algorithm to estimate an upper bound of restoration entropy by searching for a suitable Riemannian metric was described in [22]. For interconnections of nonlinear subsystems, the results in [23] describe a compositional computation of upper bounds of restoration entropies by considering a small gain criterion as in [24, Sec. 2.2]. Particularly, for each subsystem and using the linearized dynamics, a quantity is computed. By assuming that a small gain criterion hold, the summation of those quantities over all the subsystems gives an upper bound for the restoration entropy.

In this letter, we also focus on large-scale interconnected systems as described in [23] and generalize its results by using the most general quadratic expression for the so-called supply rate similar to [24, eq. (2.4)]. This new structure expands the class of systems for which upper bounds for the restoration entropy can be computed compositionally. Unlike the results in [23] that use a fixed dissipativity property for each subsystem, we combine the overall compositionality criterion with a simultaneous search over compatible subsystem dissipativity properties. We employ a distributed optimization based on the Alternating Direction Method of Multipliers (ADMM) algorithm [25] to decompose and solve this problem. In particular, we alternately solve a local and a global feasibility problem. The local problems are solved in a parallel fashion for each subsystem and the values obtained are then utilized in the global problem which checks for the satisfaction of a compositionality criterion. To solve local feasibility problems, we resort to a sum-of-squares (SOS) optimization formulation; to this end we restrict the dynamics describing the subsystems to be of polynomial structure. Compared to the numerical schemes in [20] and [22], our approach breaks down the computational complexity to the level of subsystems and, hence, can be applied to large-scale interconnected systems. Finally, we illustrate the effectiveness of our approach by two examples for which either the exact value of the restoration entropy is known to us or its upper bound.

II. Preliminaries and System Definitions

Notation: By $[k_1; k_2]$ we denote the set of integers $\{j \mid k_1 \leq j \leq k_2\}$. We use A^T to denote the transpose of matrix A. I_m denotes the $m \times m$ identity matrix. The induced 2-norm is denoted by $\|\cdot\|$ and the Frobenius norm by $\|\cdot\|_F$. For vectors $p_i \in \mathbb{R}^{r_i}$, $i \in [1; N]$, the result of stacking them on top of each other is denoted by $\mathbf{stack}(p_i) \in \mathbb{R}^{r_1+\cdots+r_N}$.

We consider interconnections of discrete-time nonlinear subsystems Σ_i, $i \in [1; N]$. The ith subsystem is described by:

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

where $x_i(t) \in X_i \subseteq \mathbb{R}^{n_i}$ is the state, $u_i(t) \in U_i \subseteq \mathbb{R}^{m_i}$ is the input to the system, and $y_i(t) \in Y_i \subseteq \mathbb{R}^{k_i}$ is the output. The interconnection constraint is given by

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

for some $V_{ij} \in \mathbb{R}^{m_i \times k_j}$.

The interconnected system can be written as

$$x(t) := \mathbf{stack}(x_i(t)), \quad \phi(x(t)) := \mathbf{stack}(\phi_i(x_i^{\leftarrow\rho}(x(t)))), \tag{2}$$

where

$$x_i^{\leftarrow\rho}(x) := \left(x_i, \sum_{j=1}^{N} V_{ij} h_j(x_j) \right), \quad \text{and } n := \sum_{i=1}^{N} n_i.$$

Let $X := \prod_{i=1}^{N} X_i$. We consider only those trajectories that start in a compact set $K \subset X$. The map $\phi : X \to X$ is assumed to be continuous. The state estimator is composed of a coder, located near the sensor, and a decoder. The decoder has access to the state measurements through a finite capacity communication channel, i.e., the state information received by the decoder is inaccurate (not perfect). For systems with unstable dynamics and insufficient channel data rate, the inaccuracy in the state estimation may worsen with time.

Now, we describe the definition of fine state observation, for which the required minimal average bit rate is equal to the restoration entropy for the case of a forward invariant set. We assume that both coder and decoder have access to a common initial estimate $\hat{x}(0)$ and its accuracy $\delta > 0$:

$$\|x(0) - \hat{x}(0)\| < \delta. \tag{3}$$

Definition 1: An observer is said to finely observe the system in (2) if the observation error exponentially decays to zero with time, i.e., there exists $\delta_* > 0$, $G > 0$, and $g > 0$ such that for all $\delta \leq \delta_*$ (δ defined in (3)), the following holds

$$\|x(t) - \hat{x}(t)\| < G\delta e^{-gt} \quad \forall t \geq 0, x(0), \hat{x}(0) \in K.$$

We follow [17] and assume that there exists a uniform upper bound $b_+(\Delta t)$ and a lower bound $b_-(\Delta t)$ on the number of bits that can be transferred in any interval of duration Δt. We further assume that the average number of transmittable bits converges, i.e., $b_\pm(\Delta t)/(\Delta t) \to c$ as $\Delta t \to \infty$. Then we denote by $\mathcal{R}_{\text{fo}}$ the infimum channel bit rate c needed to finely observe the system in (2). Now, we present the definition of restoration entropy.

Definition 2: Consider the interconnected system in (2). Let B_a^δ denote the open ball with radius δ centered at a and ϕ^n denote the n-th iterate of ϕ, with $\phi^0(a) := a$ and $\phi^{n+1} := \phi \circ \phi^n$. For every $a \in K$, $n \in \mathbb{N}$ and $\epsilon > 0$, let $p(n, a, \epsilon)$ denote the smallest number of open ϵ-balls required to cover $\phi^n(B_a^\delta \cap K)$. Then the restoration entropy of the system in (2) on K is[1]

$$H_{\text{res}}(\phi, K) := \overline{\lim_{n\to\infty}} \frac{1}{n} \overline{\lim_{\epsilon\to 0}} \sup_{a\in K} \log_2 p(n, a, \epsilon).$$

From [17, Th. 8], one has

$$H_{\text{res}}(\phi, K) \leq \mathcal{R}_{\text{fo}}(\phi, K),$$

[1] We use $\overline{\lim}$ to denote the limit superior.

and $H_{\rm res}(\phi, K) = \mathcal{R}_{\rm fo}(\phi, K)$ for a forward invariant set K (i.e., $\{\phi(x) \mid x \in K\} \subseteq K$).

Let us denote by $\phi^t(a)$ the state of the network at time t which starts from $x(0) = a \in X$. Now we define sets:

$$X(t) := \{\phi^t(a) \mid a \in K\}, \text{ and } X^\infty := \cup_{t=0}^{\infty} X(t).$$

The set X^∞ denotes the set of all reachable states from the initial states in K over an infinite time horizon. The following assumptions are made to arrive at an upper bound for the restoration entropy.

Assumption 1: Maps $\phi_i : X_i \times U_i \to X_i$ and $h_i : X_i \to Y_i$ are continuously differentiable for any i.

This ensures that the derivatives of ϕ_i and h_i are also continuous. Next we define uniform continuity near a subset which is used in Assumption 2.

Definition 3: A function $f : \mathbb{R}^n \to \mathbb{R}^q$ is said to be *uniformly continuous near a subset* $X_* \subset \mathbb{R}^n$ if for any $\epsilon > 0$, there exists $\delta > 0$ such that:

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

Note that continuous functions are uniformly continuous near compact sets.

Assumption 2: For any subsystem Σ_i, functions

$$\frac{\partial \phi_i}{\partial x_i}(x_i^{\leftarrow\!\!\!\!\!\!\curvearrowleft}(x)), \quad \frac{\partial \phi_i}{\partial u_i}(x_i^{\leftarrow\!\!\!\!\!\!\curvearrowleft}(x))), \quad \frac{\partial h_i}{\partial x_i}(x_i),$$

are bounded on X^∞ and uniformly continuous near this set. Assumption 2 holds when the set X^∞ is bounded, which is true for a compact forward invariant set K of initial states.

Now consider the first-order approximation of subsystems near any trajectory:

$$z_i(t+1) = A_i(x(t))z_i(t) + B_i(x(t))w_i(t), \quad \zeta_i(t) = C_i(x(t))z_i(t). \tag{4}$$

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

$$A_i(x) := \frac{\partial \phi_i}{\partial x_i}(x_i^{\leftarrow\!\!\!\!\!\!\curvearrowleft}(x)), \quad B_i(x) := \frac{\partial \phi_i}{\partial u_i}(x_i^{\leftarrow\!\!\!\!\!\!\curvearrowleft}(x)), \quad C_i(x) := \frac{\partial h_i}{\partial x_i}(x_i),$$

where $x_i^{\leftarrow\!\!\!\!\!\!\curvearrowleft}(x) = (x_i, u_i)$ and $u_i = \sum_{j=1}^N V_{ij} h_j(x_j)$.

For the first-order approximation of a subsystem as in (4), we assume the following dissipation-like inequality [24].

Assumption 3: Consider a subsystem Σ_i as in (1). There exist $n_i \times n_i$ symmetric matrices $P_i > 0$, $Q_i \geq 0$ and $(m_i + k_i) \times (m_i + k_i)$ symmetric matrices $\underline{X}_i$ with conformal block partitions $\underline{X}_i^{jl}$, $j, l \in \{1, 2\}$, such that the following inequality is true along all solutions of the interconnected system in (1) starting in the given compact set K:

$$[A_i(x)z_i + B_i(x)w_i]^T P_i [A_i(x)z_i + B_i(x)w_i] \leq z_i^T Q_i z_i + \begin{bmatrix} w_i \\ \zeta_i \end{bmatrix}^T \begin{bmatrix} \underline{X}_i^{11} & \underline{X}_i^{12} \\ \underline{X}_i^{21} & \underline{X}_i^{22} \end{bmatrix} \begin{bmatrix} w_i \\ \zeta_i \end{bmatrix}, \ \zeta_i = C_i(x)z_i, \tag{5}$$

for all $z_i \in \mathbb{R}^{n_i}$, $w_i \in \mathbb{R}^{m_i}$ and $x \in X^\infty$.

The following compositionality assumption relates matrices $\underline{X}_i$ for subsystems, from the above inequality, via the interconnection matrix[2] V.

Assumption 4: Assume the following inequality holds:

$$\begin{bmatrix} V \\ I_k \end{bmatrix}^T \underline{X}^{comp} \begin{bmatrix} V \\ I_k \end{bmatrix} \leq 0 \tag{6}$$

where

$$\underline{X}^{\rm comp} := \begin{bmatrix} \underline{X}_1^{11} & & & \underline{X}_1^{12} & & \\ & \ddots & & & \ddots & \\ & & \underline{X}_N^{11} & & & \underline{X}_N^{12} \\ \underline{X}_1^{21} & & & \underline{X}_1^{22} & & \\ & \ddots & & & \ddots & \\ & & \underline{X}_N^{21} & & & \underline{X}_N^{22} \end{bmatrix},$$

and $k = \sum_{i=1}^N k_i$.

Inequality (6) is similar to [24, eq. (2.8)] in the context of compositional verification of stability for interconnected systems.

Let us now write the first-order approximation of the network as

$$z(t+1) = A(x(t))z(t) = \mathbf{stack}(A_i(x(t))z_i(t) + B_i(x(t))w_i(t))$$

for all $z(t) := \mathbf{stack}(z_i(t))$ with

$$w_i = \sum_{j=1}^N V_{ij}\zeta_j, \ \ \zeta_j = C_j(x)z_j, \ \text{ and } A(x) := \frac{\partial \phi}{\partial x}(x).$$

For given $n \times n$ symmetric matrices $P > 0$ and $Q \geq 0$, let $H_L(P, Q) := 0.5 \sum_{j=1}^n \max\{0, \log_2 \lambda_j\}$ where $\lambda_j, j \in [1; n]$, are the eigenvalues[3] of QP^{-1} (or $P^{-1}Q$).

Now we have all the ingredients to present the main result of this letter.

Theorem 1: For a network as in (2) and under Assumptions 1-4, the following holds

$$\mathcal{R}_{\rm fo}(\phi, K) \leq \sum_{i=1}^N H_L(P_i, Q_i). \tag{7}$$

Proof: The proof is inspired by that of [23, Th. 3.7]. From Assumption 2, we have that $A(x) = \frac{\partial \phi}{\partial x}(x)$ is bounded on X^∞ and is uniformly continuous near this set.

Let P and Q denote the following positive definite and positive semidefinite block-diagonal $n \times n$ matrices, respectively:

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

We have

$$z^T A(x)^T P A(x) z = \sum_{i=1}^N [A_i(x)z_i + B_i(x)w_i]^T P_i [A_i(x)z_i + B_i(x)w_i]$$

[2]Here V is a block matrix with the (i, j)th block given by the matrix V_{ij}.

[3]Note that eigenvalues λ_j are non-negative since they are the roots of $\det(Q - \lambda_j P) = 0$ and as a result $\lambda_j = \mathrm{x}^T Q\mathrm{x}/(\mathrm{x}^T P\mathrm{x})$ for some $\mathrm{x} \neq 0$.

$$
\begin{aligned}
&\overset{(5)}{\leq} \sum_{i=1}^{N}\left(z_i^T Q_i z_i + \begin{bmatrix} w_i \\ \zeta_i \end{bmatrix}^T \begin{bmatrix} \underline{X}_i^{11} & \underline{X}_i^{12} \\ \underline{X}_i^{21} & \underline{X}_i^{22} \end{bmatrix} \begin{bmatrix} w_i \\ \zeta_i \end{bmatrix}\right) \\
&= \sum_{i=1}^{N} z_i^T Q_i z_i \\
&\quad + \begin{bmatrix} w_1 \\ \vdots \\ w_N \\ \zeta_1 \\ \vdots \\ \zeta_N \end{bmatrix}^T \begin{bmatrix} \underline{X}_1^{11} & & & \underline{X}_1^{12} & & \\ & \ddots & & & \ddots & \\ & & \underline{X}_N^{11} & & & \underline{X}_N^{12} \\ \underline{X}_1^{21} & & & \underline{X}_1^{22} & & \\ & \ddots & & & \ddots & \\ & & \underline{X}_N^{21} & & & \underline{X}_N^{22} \end{bmatrix} \begin{bmatrix} w_1 \\ \vdots \\ w_N \\ \zeta_1 \\ \vdots \\ \zeta_N \end{bmatrix} \\
&= z^T Q z + \begin{bmatrix} V \begin{bmatrix} \zeta_1 \\ \vdots \\ \zeta_N \end{bmatrix} \\ \zeta_1 \\ \vdots \\ \zeta_N \end{bmatrix}^T \underline{X}^{\text{comp}} \begin{bmatrix} V \begin{bmatrix} \zeta_1 \\ \vdots \\ \zeta_N \end{bmatrix} \\ \zeta_1 \\ \vdots \\ \zeta_N \end{bmatrix} \\
&= z^T Q z + \begin{bmatrix} \zeta_1 \\ \vdots \\ \zeta_N \end{bmatrix}^T \begin{bmatrix} V \\ I_k \end{bmatrix}^T \underline{X}^{\text{comp}} \begin{bmatrix} V \\ I_k \end{bmatrix} \begin{bmatrix} \zeta_1 \\ \vdots \\ \zeta_N \end{bmatrix} \\
&\overset{(6)}{\leq} z^T Q z \quad \forall z. \qquad (8)
\end{aligned}
$$

In the rest of the proof, we leverage Theorem 2 and Lemma 2 given in the appendix. In Theorem 2, we have $\sum_{i=1}^{d} \log_2 \lambda_i(\mathrm{x}) \leq H_L(P, A(\mathrm{x})^T P A(\mathrm{x}))$ for all $d \in [1; n]$ and $\mathrm{x} \in X^\infty$. From (8) and Lemma 2, we have $H_L(P, A(\mathrm{x})^T P A(\mathrm{x})) \leq H_L(P, Q)$ for every $\mathrm{x} \in X^\infty$. Thus, by using Theorem 2 with $v_d(\cdot) := 0$ and $\Lambda_d := 2H_L(P, Q)$, $d \in [1; n]$, we get $\mathcal{R}_{\text{fo}}(\phi, K) \leq H_L(P, Q)$. Since P and Q are block diagonal, the set of roots of $\det(Q - \lambda P) = 0$ is the same as $\{\lambda \mid \det(Q_i - \lambda P_i) = 0, i \in [1; N]\}$. Thus $H_L(P, Q) = \sum_{i=1}^{N} H_L(P_i, Q_i)$ which completes the proof. ■

Next, we describe a distributed optimization approach to search for P_i, Q_i and $\underline{X}_i$ in (5) in order to compute the upper bound in (7).

III. Distributed Optimization

In this section, we describe a distributed optimization method, based on the alternating direction method of multipliers (ADMM) algorithm [25], for computation of the upper bound in (7). If Assumptions 1-4 hold, then our objective is to find, for each subsystem, matrices P_i, Q_i, and $\underline{X}_i$ that satisfy (5) and (6). For this, we alternately solve local feasibility problems and a global feasibility problem. By local feasibility problem, we refer to a search for the triple $(P_i, Q_i, \underline{X}_i)$, $i \in [1; N]$, that satisfy (5). And by global feasibility problem, we refer to a search for matrices $\underline{X}_1, \ldots, \underline{X}_N$ that satisfy (6).

Local feasibility problem:

$$\mathcal{S}_i = \{(P_i, Q_i, \underline{X}_i) \mid P_i, Q_i, \text{ and } \underline{X}_i \text{ satisfying the requirements in Assumption 3}\}. \qquad (9)$$

Global feasibility problem:

$$\mathcal{G} = \{(\underline{X}_1, \ldots, \underline{X}_N) \mid \text{condition (6) is satisfied}\}. \qquad (10)$$

Now consider the following indicator functions:

$$\mathbb{I}_{\mathcal{S}_i}(P_i, Q_i, \underline{X}_i) = \begin{cases} 0, & (P_i, Q_i, \underline{X}_i) \in \mathcal{S}_i, \\ \infty, & \text{otherwise}, \end{cases}$$

$$\mathbb{I}_{\mathcal{G}}(\underline{X}_1, \ldots, \underline{X}_N) = \begin{cases} 0, & (\underline{X}_1, \ldots, \underline{X}_N) \in \mathcal{G}, \\ \infty, & \text{otherwise}. \end{cases}$$

We introduce auxiliary variables Z_i for each subsystem to formulate an optimization problem

$$\min_{d} \sum_{i=1}^{N} \mathbb{I}_{\mathcal{S}_i}(P_i, Q_i, \underline{X}_i) + \mathbb{I}_{\mathcal{G}}(Z_1, \ldots, Z_N), \qquad (11)$$

$$\text{s.t.} \quad \underline{X}_i - Z_i = 0, \quad \forall i \in [1; N], \qquad (12)$$

where $d = (P_1, \ldots, P_N, Q_1, \ldots, Q_N, \underline{X}_1, \ldots, \underline{X}_N, Z_1, \ldots, Z_N)$.

Often large optimization problems can be converted into smaller sub-problems by separating the objective function across decision variables. In our objective function, the first part $\mathbb{I}_{\mathcal{S}_i}$ is separable by subsystems. This permits parallel computation of the local problems across subsystems. For this, we introduce dual variables Λ_i. The variables in the ADMM are updated as follows:

1) For each $i \in [1; N]$, we solve local problems:

$$P_i^{k+1}, Q_i^{k+1}, \underline{X}_i^{k+1} = \operatorname{argmin}_{(P_i, Q_i, \underline{X}_i) \in \mathcal{S}_i} \left\| \underline{X}_i - Z_i^k + \Lambda_i^k \right\|_F^2.$$

2) If $\underline{X}_{1:N}^{k+1} \in \mathcal{G}$, then the optimal solution is achieved and the algorithm terminates. Otherwise, we solve the global problem:

$$Z_{1:N}^{k+1} = \operatorname{argmin}_{(Z_{1:N}) \in \mathcal{G}} \sum_{i=1}^{N} \left\| \underline{X}_i^{k+1} - Z_i + \Lambda_i^k \right\|_F^2.$$

3) Now we update the dual variables:

$$\Lambda_i^{k+1} = \underline{X}_i^{k+1} - Z_i^{k+1} + \Lambda_i^k, \quad i \in [1; N],$$

and continue with the first step until a possible convergence.

The local problems are solved parallelly by sum-of-squares (SOS) optimization formulation as described in the following paragraphs. If required, these values are then used in the global problem which can be solved by semi-definite programming. The dual variables are updated using the values obtained from steps 1 and 2. The procedure repeats until convergence is achieved. Since the indicator functions in (11) are convex, the solutions are guaranteed to converge to the optimal ones [26].

We reformulate the condition in (5) as an SOS problem to search for suitable values of P_i, Q_i, and $\underline{X}_i$. For this we make the following assumption.

Assumption 5: Maps $\phi_i : X_i \times U_i \to X_i$ and $h_i : X_i \to Y_i$ in (1) are polynomial functions of x_i and u_i.

The SOS formulation is given by the following lemma.

Lemma 1: Let Assumption 5 holds and assume set X^∞ can be written in terms of a vector of polynomial inequalities[4]

[4]The inequalities are considered element-wise.

$X^\infty = \{x \in \mathbb{R}^n \mid b(x) \geq 0\}$. Further, assume that for each subsystem Σ_i, there exists a vector of sum-of-squares polynomial $\beta(x, \alpha)$, $\alpha \in \mathbb{R}^{m_i+n_i}$ of an appropriate dimension so that the expression in (13) is a sum-of-squares polynomial. Then Assumption 3 holds.

Proof: We use $E(x)$ to refer to the matrix appearing between α^T and α in (13). Using $\zeta_i = C_i(x)z_i$ and rearrangement of terms, the inequality in (5) can be rewritten as $E(x) \geq 0$ for all $x \in X^\infty$.

When the expression in (13) is an SOS polynomial, we get $\alpha^T E(x)\alpha - \beta(x, \alpha)^T b(x) \geq 0$ for all $\alpha \in \mathbb{R}^{m_i+n_i}$ and $x \in X^\infty$. Since $b(x) \geq 0$ for $x \in X^\infty$ and each term of the vector $\beta(x, \alpha)$ is an SOS polynomial, we get $\alpha^T E(x)\alpha \geq 0$ for all $\alpha \in \mathbb{R}^{m_i+n_i}$ and $x \in X^\infty$. This leads to $E(x) \geq 0$ for all $x \in X^\infty$ which completes the proof. ■

Remark 1: Note that X^∞ denotes the set of all reachable states starting from the initial set K over an infinite time horizon. If K is forward invariant, then we have $X^\infty = K$. Computation of X^∞ is a difficult task and in general undecidable. However, one can satisfy (13) either with an over-approximation of X^∞ or for any $x \in \mathbb{R}^n$ (i.e., $b(x) := 0$) (if successful) which still implies that Assumption 3 holds.

In the next section, we present a nonlinear and a linear example, and compute the upper bound in (7) by selecting appropriate values of P_i, Q_i, and $\underline{X}_i$ for their subsystems.

IV. Case Study

Example 1: Consider the Hénon system [27] with standard parameters $a = 1.4$ and $b = 0.3$:

$$\phi(x) = \begin{bmatrix} 1.4 - x_1^2 + 0.3x_2 \\ x_1 \end{bmatrix},$$

where $x = (x_1, x_2) \in \mathbb{R}^2$. The quadrilateral K with the following vertices is a compact forward invariant set [27]:

$$q_1 = (-1.862, 1.96), \quad q_2 = (1.848, 0.6267), \quad q_3 = (1.743, -0.6533), \quad q_4 = (-1.484, -2.3333).$$

From [16, Th. 16], we already know an upper bound of the restoration entropy $H_{\text{res}}(\phi, K) \leq 1.704793$.

Now, we consider the system as a network of two scalar subsystems:

$$\phi_1(x_1, u_1) := 1.4 - x_1^2 + 0.3u_1, \quad h_1(x_1) := x_1,$$
$$\phi_2(x_2, u_2) := u_2, \quad h_2(x_2) := x_2,$$

where $u_i = \sum_{j=1}^{2} v_{ij}h_j(x_j)$ and v_{ij} denote the entries of the matrix $V = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. Using the distributed optimization algorithm introduced in the previous section, we get $P_1 = 26$, $P_2 = 37$, $Q_1 = 421.9223$, $Q_2 = 36.99$, $\underline{X}_1^{11} = 36.9483$, $\underline{X}_1^{12} = 0.0356$, $\underline{X}_1^{22} = -37.029$, $\underline{X}_2^{11} = 37.015$, $\underline{X}_2^{12} = -0.0178$, and $\underline{X}_2^{22} = -36.9696$, satisfying conditions in (5) and (6), and thus we obtain $H_{\text{res}}(\phi, K) \leq \sum_{i=1}^{2} H_L(P_i, Q_i) = 2.0102$, which is close enough to the known upper bound 1.704793 [16, Th. 16].

Next we describe a linear system and compare the upper bound computed by our modular approach with the known exact value of the restoration entropy.

Example 2: Consider a five dimensional linear system $\phi(x) = Ax$. Let a_{ij} denote the (i, j)-th entry of A where $a_{11} = 2$, $a_{ii} = 0.5$ for $i \in [2; 5]$, $a_{ij} = 0$ for $i > j$, and $a_{ij} = 1$ for $i < j$. For linear systems, the restoration entropy coincides with the topological entropy which is equal to the sum of the logarithm of the absolute value of the unstable eigenvalues of matrix A [16]. Thus, $H_{\text{res}}(\phi, K) = 1$ for any compact set $K \subset \mathbb{R}^5$.

The system can be seen as a network of five scalar subsystems

$$\phi_1(x_1, u_1) := 2x_1 + u_1, \quad h_1(x_1) := x_1,$$
$$\phi_i(x_i, u_i) := 0.5x_i + u_i, \quad h_i(x_i) := x_i, \quad i \in [2; 4],$$
$$\phi_5(x_5) := 0.5x_5, \quad h_5(x_5) := x_5,$$

where $u_i = \sum_{j=1}^{5} v_{ij}h_j(x_j)$, $i \in [1 : 4]$, and the entries of the matrix V are given as $v_{ij} := 1$, for $i < j$, and $v_{ij} := 0$, for $i >= j$.

Using the distributed optimization algorithm introduced in the previous section, we get $P_1 = 0.01$, $P_2 = 1$, $P_3 = 6$, $P_4 = 32$, $Q_1 = 0.0432$, $Q_2 = 0.7584$, $Q_3 = 5.9853$, $Q_4 = 28.2875$, $\underline{X}_1^{11} = 0.0773$, $\underline{X}_1^{12} = 0.0093$, $\underline{X}_1^{22} = -0.00145$, $\underline{X}_2^{11} = 1.5495$, $\underline{X}_2^{12} = 0.1785$, $\underline{X}_2^{22} = -0.3203$, $\underline{X}_3^{11} = 12.8680$, $\underline{X}_3^{12} = 0.0261$, $\underline{X}_3^{22} = -3.1976$, $\underline{X}_4^{11} = 13377530.7867$, $\underline{X}_4^{12} = -7.9311$, $\underline{X}_4^{22} = -20.2875$, $\underline{X}_5^{11} = -132323.8913$, $\underline{X}_5^{12} = 0$, and $\underline{X}_5^{22} = -20251359.1959$, and thus we obtain $H_{\text{res}}(\phi, K) \leq \sum_{i=1}^{5} H_L(P_i, Q_i) = 1.0550$. Here, $H_L(P_5, Q_5) = 0$ for $P_5 = p$ and $Q_5 = 0.5^2 p$ for any $0 < p \in \mathbb{R}$. One can readily see that our computed upper bound is very close to the exact restoration entropy which is 1. ■

To demonstrate the advantage of the proposed modular approach here, we used the proposed algorithm in [22] for which the code is publicly available in Example 2. In the proposed algorithm in [22], the user needs to provide the number of grid points along each dimension. The grid points are used to tackle the involved nonlinear nonconvex maximization problem. By choosing a large number of grid points, one is expected to get sound results. Table I compares the upper estimates and the computation times from our approach with the algorithm proposed in [22]. The number of grid points along each dimension for columns [22](A), [22](B), and [22](C) are 10, 50, and 60, respectively, while the number of iterations for all of them is 15. The computations were performed on an Apple M1 with 8 cores and 16 GB RAM.

V. Conclusion

We provided a modular approach to compute upper bounds of the restoration entropies for large-scale networks by looking

$$\alpha^T \begin{bmatrix} -B_i(x)^T P_i B_i(x) + \underline{X}_i^{11} & -B_i(x)^T P_i A_i(x) + \underline{X}_i^{12} C_i(x) \\ -A_i(x)^T P_i B_i(x) + C_i(x)^T \underline{X}_i^{21} & -A_i(x)^T P_i A_i(x) + Q_i + C_i(x)^T \underline{X}_i^{22} C_i(x) \end{bmatrix} \alpha - \beta(x, \alpha)^T b(x), \quad \forall \alpha \in \mathbb{R}^{m_i+n_i}. \quad (13)$$

TABLE I
COMPARISON WITH THE ALGORITHM IN [22] FOR EXAMPLE 2

	Our approach	[22](A)	[22](B)	[22](C)
Time	0.8448 s	2.07 s	1.23 hr	2.93 hr
Estimate	1.0550	1.0076	1.0076	1.0076

at them as interconnections of smaller subsystems. In particular, we formulated a distributed optimization problem composed of local problems and a global problem. The local problems are solved separately for subsystems. We also described an SOS formulation to solve the local problems provided that the subsystem dynamics are of polynomial structure. The results from the local problems are then checked for the satisfaction of a compositionality criterion in the global problem. Finally, we applied the described approach on two case studies and compared the obtained estimates with the known values in the literature. Table I demonstrate the advantage of the modular approach over the proposed algorithm in [22]. As part of future work, we intend to formulate local feasibility problems in a way that will also minimize $H_L(P_i, Q_i)$ in order to arrive at the best possible upper bound given the proposed modular approach.

APPENDIX

Theorem 2: [16, Th. 12] For system in (2), let Assumption 2 holds and let there exist continuous and bounded on X^∞ functions $v_d : \mathbb{R}^n \to \mathbb{R}$, constants $\Lambda_d \geq 0$, $d \in [1; n]$, and an $n \times n$ symmetric matrix $P \geq 0$ such that

$$v_d(\phi(x)) - v_d(x) + \sum_{i=1}^{d} \log_2 \lambda_i(x) \leq \Lambda_d, \forall x \in X^\infty,$$

where $\lambda_i(x) \geq \lambda_{i+1}(x) \geq 0$ are the roots of $\det(A(x)^T PA(x) - \lambda P) = 0$ repeated in accordance with their algebraic multiplicities. Then $\mathcal{R}_{fo}(\phi, K) \leq 0.5 \max\{\Lambda_1, \dots, \Lambda_n\}$.

Lemma 2: For the symmetric matrices $P > 0$, $Q_2 \geq Q_1 \geq 0$ we have $H_L(P, Q_2) \geq H_L(P, Q_1)$.

Proof: Note that both $\det(Q - \lambda P) = 0$ and $\det(P^{-1/2}QP^{-1/2} - \lambda I) = 0$ have identical roots. Since, $Q_2 \geq Q_1$ and $P^{-1/2}$ is symmetric, we have $P^{-1/2}Q_2P^{-1/2} \geq P^{-1/2}Q_1P^{-1/2}$. Then by Weyl's inequality [28, Th. 4.3.1], we obtain $H_L(P, Q_2) \geq H_L(P, Q_1)$. ∎

REFERENCES

[1] J. Hespanha, A. Ortega, and L. Vasudevan, "Towards the control of linear systems with minimum bit-rate," in *Proc. Int. Symp. Math. Theory Netw. Syst.*, 2002, p. 1.
[2] D. Liberzon and S. Mitra, "Entropy and minimal bit rates for state estimation and model detection," *IEEE Trans. Autom. Control*, vol. 63, no. 10, pp. 3330–3344, Oct. 2018.
[3] C. Kawan and J.-C. Delvenne, "Network entropy and data rates required for networked control," *IEEE Trans. Control Netw. Syst.*, vol. 3, no. 1, pp. 57–66, Mar. 2016.
[4] G. N. Nair, F. Fagnani, S. Zampieri, and R. J. Evans, "Feedback control under data rate constraints: An overview," *Proc. IEEE*, vol. 95, no. 1, pp. 108–137, Jan. 2007.
[5] B. R. Andrievsky, A. S. Matveev, and A. L. Fradkov, "Control and estimation under information constraints: Toward a unified theory of control, computation and communications," *Autom. Remote Control*, vol. 71, no. 4, pp. 572–633, 2010.
[6] F. Colonius, "Minimal bit rates and entropy for exponential stabilization," *SIAM J. Control Optim.*, vol. 50, no. 5, pp. 2988–3010, 2012.
[7] F. Colonius and C. Kawan, "Invariance entropy for control systems," *SIAM J. Control Optim.*, vol. 48, no. 3, pp. 1701–1721, 2009.
[8] G. N. Nair, R. J. Evans, I. M. Y. Mareels, and W. Moran, "Topological feedback entropy and nonlinear stabilization," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1585–1597, Sep. 2004.
[9] M. S. Tomar and M. Zamani, "On a notion of entropy for reachability properties," 2022, *arXiv:2203.00767*.
[10] C. Kawan and S. Yüksel, "Invariance properties of controlled stochastic nonlinear systems under information constraints," *IEEE Trans. Autom. Control*, vol. 66, no. 10, pp. 4514–4529, Oct. 2021. [Online]. Available: https://doi.org/10.1109/TAC.2020.3030846
[11] C. Kawan and S. Yüksel, "On optimal coding of non-linear dynamical systems," *IEEE Trans. Inf. Theory*, vol. 64, no. 10, pp. 6816–6829, Oct. 2018.
[12] 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.
[13] A. S. Matveev and A. V. Savkin, *Estimation and Control Over Communication Networks*. Boston, MA, USA: Birkhäuser, 2009.
[14] G. Froyland, O. Junge, and G. Ochs, "Rigorous computation of topological entropy with respect to a finite partition," *Physica D, Nonlinear Phenomena*, vol. 154, nos. 1–2, pp. 68–84, 2001.
[15] P. Koiran, "The topological entropy of iterated piecewise affine maps is uncomputable," *Discr. Math. Theor. Comput. Sci.*, vol. 4, no. 2, pp. 351–356, 2001.
[16] A. S. Matveev and A. Y. Pogromsky, "Observation of nonlinear systems via finite capacity channels: Constructive data rate limits," *Automatica*, vol. 70, pp. 217–229, Aug. 2016.
[17] A. S. Matveev and A. Y. Pogromsky, "Observation of nonlinear systems via finite capacity channels, part II: Restoration entropy and its estimates," *Automatica*, vol. 103, pp. 189–199, May 2019.
[18] C. Kawan, "On the relation between topological entropy and restoration entropy," *Entropy*, vol. 21, no. 1, p. 7, 2019.
[19] A. Y. Pogromsky and A. Matveev, "Estimation of topological entropy via the direct Lyapunov method," *Nonlinearity*, vol. 24, no. 7, p. 1937, 2011.
[20] S. Hafstein and C. Kawan, "Numerical approximation of the data-rate limit for state estimation under communication constraints," *J. Math. Anal. Appl.*, vol. 473, no. 2, pp. 1280–1304, 2019.
[21] C. Kawan, A. S. Matveev, and A. Y. Pogromsky, "Remote state estimation problem: Towards the data-rate limit along the avenue of the second Lyapunov method," *Automatica*, vol. 125, Mar. 2021, Art. no. 109467.
[22] C. Kawan, S. Hafstein, and P. Giesl, "A subgradient algorithm for data-rate optimization in the remote state estimation problem," *SIAM J. Appl. Dyn. Syst.*, vol. 20, no. 4, pp. 2142–2173, 2021.
[23] A. S. Matveev, A. V. Proskurnikov, A. Pogromsky, and E. Fridman, "Comprehending complexity: Data-rate constraints in large-scale networks," *IEEE Trans. Autom. Control*, vol. 64, no. 10, pp. 4252–4259, Oct. 2019.
[24] M. Arcak, C. Meissen, and A. Packard, *Networks of Dissipative Systems: Compositional Certification of Stability, Performance, and Safety*. Cham, Switzerland: Springer, 2016.
[25] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, 2011.
[26] C. Meissen, L. Lessard, M. Arcak, and A. K. Packard, "Compositional performance certification of interconnected systems using ADMM," *Automatica*, vol. 61, pp. 55–63, Nov. 2015.
[27] M. Hénon, "A two-dimensional mapping with a strange attractor," in *The Theory of Chaotic Attractors*. New York, NY, USA: Springer, 1976, pp. 94–102.
[28] R. Horn and C. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 2013.