

# STATE-ROBUST OBSERVABILITY MEASURES FOR SENSOR SELECTION IN NONLINEAR DYNAMIC SYSTEMS

Mohamad H. Kzma<sup>1</sup>, Sebastian A. Nugroho<sup>2</sup>, Aleksandar Haber<sup>3</sup>, and Ahmad F. Taha<sup>1</sup>

**Abstract**—This paper explores the problem of selecting sensor nodes for a general class of nonlinear dynamical networks. In particular, we study the problem by utilizing altered definitions of observability and open-loop lifted observers. The approach is performed by discretizing the system's dynamics using the implicit Runge-Kutta method and by introducing a state-averaged observability measure. The observability measure is computed for a number of perturbed initial states in the vicinity of the system's true initial state. The sensor node selection problem is revealed to retain the submodular and modular properties of the original problem. This allows the problem to be solved efficiently using a greedy algorithm with a guaranteed performance bound while showing an augmented robustness to unknown or uncertain initial conditions. The validity of this approach is numerically demonstrated on a  $H_2/O_2$  combustion reaction network.

**Index Terms**—Nonlinear Systems, sensor selection, nonlinear observability, discrete systems, greedy algorithm

## I. INTRODUCTION AND PAPER CONTRIBUTIONS

THE sensor selection problem is one of the fundamental control engineering problems. The problem is crucial for the control, monitoring, and safe operation of a large number of engineered systems, such as electric power grids [1], municipal water networks [2], and transportation systems [3]. From a control- and observability-based formulation, this problem aims to find the optimal combination of sensor nodes (graph nodes whose local states should be observed) that optimize appropriate observability measures. The goal is to make the system *as observable as possible* using a limited number of sensors to be placed on select nodes in the network.

Sensor selection problems have gained considerable research interest in recent years as a plethora of methods have been proposed in the literature, especially for *linear* systems. These methods can be categorized based on underlying mathematical approaches, such as network and graph theory [4], [5], sparsity promoting algorithms [6], [7], semidefinite approximations and relaxations [8], heuristic optimization under convex relaxations [9], greedy approach under submodular set maximization [10], and mixed-integer optimization [1], [11]. Regardless, methods for solving sensor selection problems for nonlinear dynamic networks are significantly less developed. Only a handful of methods have been proposed so far to address this problem for nonlinear dynamic systems.

A sensor selection algorithm for target tracking in nonlinear dynamic networks based on a generalized information gain is proposed in [12]. Next, an empirical observability Gramian approach is utilized in [13] for placing phasor measurement units in transmission power networks. Another approach based on an open-loop moving horizon estimation for sensor selection and state observation is proposed in [14]. The approach presented in [14] is more numerically tractable than the approaches based on empirical observability Gramians. A new randomized algorithm is presented in [15] in which theoretical bounds for eigenvalues and condition numbers of observability Gramians are developed. A novel framework is proposed in [16] for sensor selection and observer design. This approach is developed by using the Lyapunov stability theory and mixed-integer semidefinite optimization. Lastly, methods to place actuators in nonlinear networks that are based on heuristically solving mixed-integer nonlinear optimization problems have been recently developed in [17].

Here it should be emphasized that most of the developed approaches for solving the sensor selection problems, especially the ones involving mixed-integer programs, are *not* necessarily efficient and scalable for large nonlinear dynamic networks. The computational burden of the developed approaches becomes significant even for small or medium-sized nonlinear networks. Another issue with sensor selection problems for nonlinear networks is that, in practice, the initial states of the system are usually not known *a priori*. This creates model uncertainties and difficulties in formulating and solving the sensor selection problem since the numerically tractable observability-based approach [14] for nonlinear systems involves a dependency on initial states. This implies that under such state-dependency any perturbation to the initial state tends to yield in most cases different sensor node selections for the same nonlinear network modeled under similar system parameters.

To partly address the aforementioned limitations, we extend observability-based sensor selection method introduced in [14] by introducing state-averaged observability measures for nonlinear networks. That is, instead of utilizing the observability measures associated with a single guess on the initial state, we consider a state-averaged observability metric that relies on several points located around the actual initial state. This allows the constructed observability-based measures to take into account the variabilities resulting from initial conditions perturbations on the sensor selection measures.

Note that due to the structure preserving operations that yield the state-averaged observability Gramian, we show that the observability metrics for sensor selection retain modularity and submodularity properties. This consequently allows the

<sup>1</sup>Department of Civil and Environmental Engineering, Vanderbilt University, 2201 West End Ave, Nashville, Tennessee 37235, USA (mohamad.h.kzma@vanderbilt.edu, ahmad.taha@vanderbilt.edu.)

<sup>2</sup>Cummins Technical Center, Cummins Inc., Columbus, IN 47201 USA (sebastian.nugroho@cummins.com)

<sup>3</sup>Department of Manufacturing and Mechanical Engineering Technology, College of Engineering Technology, Rochester Institute of Technology, 1 Lomb Memorial Dr, Rochester, NY 14623, USA (ml.mecheng@gmail.com).

This work is partially supported by National Science Foundation (NSF) under Grants 2152450 and 2151571.

sensor selection problem to be solved through greedy heuristic, and thus make it suitable to solve sensor selection for large-scale nonlinear dynamic networks.

Accordingly, the main contributions of this paper are:

- We introduce a *state-averaged observability measure* for sensor selection in nonlinear networks. We use a number of points located around the nonlinear system's initial state. By relying on such local state-averaged observability measure, we attain an optimal sensor selection that is robust against unknown or uncertain initial conditions.
- We provide theoretical and numerical validation that under such state-averaged observability measures the submodularity and modularity of the sensor selection objective function is retained. In particular, we leverage the modularity and submodularity of the trace and log-det measures of the constructed observability Gramian to perform the sensor selection. Under such formulation, greedy algorithms are employed to solve the combinatorial set optimization problem and as a consequence the selection problem is rendered scalable to large-scale nonlinear dynamic networks.
- We evaluate the validity and robustness of the proposed approach by providing descriptive numerical experiments that showcase the proposed sensor selection strategy. The method is tested on a nonlinear  $H_2/O_2$  combustion reaction network.

This rest of the paper is organized as follows. Section II introduces the problem formulation. Section III presents some theoretical results pertaining to the state-averaged observability measures. Numerical results are presented in Section IV, and Section V concludes the paper.

*Paper's Notation:* Let  $\mathbb{R}$ ,  $\mathbb{R}^n$ , and  $\mathbb{R}^{p \times q}$  denote the set of real numbers, and real-valued row vectors with size of  $n$ , and  $p$ -by- $q$  real matrices. The cardinality of the a set  $\mathcal{N}$  is denoted by  $|\mathcal{N}|$ . The symbol  $\otimes$  denotes the Kronecker product. The identity matrix of size  $n$  is denoted by  $I_n$ . The operators  $\log\text{-det}(\mathbf{A})$  returns the logarithmic-determinant of matrix  $\mathbf{A}$ ,  $\text{trace}(\mathbf{A})$  returns the trace of matrix of matrix  $\mathbf{A}$ . The operator  $\text{diag}\{a_i\}_{i=1}^n \in \mathbb{R}^{n \times n}$  constructs a block diagonal matrix with scalar  $a_i$  as the diagonal entries for all  $i \in \{1, \dots, N\}$ . The operator  $\text{col}\{\mathbf{x}_i\}_{i=0}^N \in \mathbb{R}^{N \cdot n}$  constructs a column vector that concatenates vectors  $\mathbf{x}_i \in \mathbb{R}^n$  for all  $i \in \{0, \dots, N\}$ . For any vector  $\mathbf{x} \in \mathbb{R}^n$ ,  $\|\mathbf{x}\|_2$  denotes the Euclidean norm of  $\mathbf{x}$ , defined as  $\|\mathbf{x}\|_2 := \sqrt{\mathbf{x}^\top \mathbf{x}}$ , where  $\mathbf{x}^\top$  is the transpose of  $\mathbf{x}$ .

## II. PRELIMINARIES AND PROBLEM FORMULATION

In this section, we introduce mathematical preliminaries and define the problem of selecting sensor nodes. We consider a general nonlinear dynamic network defined in (1) under a continuous-time representation.

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t)), \quad (1a)$$

$$\mathbf{y}(t) = \mathbf{\Gamma} \mathbf{C} \mathbf{x}(t), \quad (1b)$$

where  $\mathbf{x} \in \mathbb{R}^{n_x}$  is the global state and  $\mathbf{y} \in \mathbb{R}^{n_y}$  is the global output vector. The nonlinear mapping function  $\mathbf{f} : \mathbb{R}^{n_x} \rightarrow \mathbb{R}^{n_x}$  is assumed to be smooth and at least twice continuously differentiable. The measurement matrix  $\mathbf{C} \in \mathbb{R}^{n_y \times n_x}$  is assumed to be known. The matrix  $\mathbf{\Gamma} := \text{diag}\{\gamma_j\}_{j=1}^{n_y} \in \mathbb{R}^{n_y \times n_y}$

determines the configuration of the sensors—that is, a node  $j$  is equipped with a sensor if  $\gamma_j = 1$ . Otherwise, we simply set  $\gamma_j = 0$ . We define the parameterize vector  $\boldsymbol{\gamma}$  that represents the sensor selection, i.e.,  $\boldsymbol{\gamma} = \text{col}\{\gamma_j\}_{j=1}^{n_y}$ . Without the loss of generality, we have assumed that the inputs are not affecting the system dynamics.

The objective of the sensor selection problem for the nonlinear dynamics (1) is to determine the combination of sensors (the 1 and 0 patterns in  $\boldsymbol{\gamma}$ ) such that an observability-based metric is maximized under a sensor ratio constraint. As such, in order to formulate the binary selection problem, we refer to utilizing a discrete-time representation of the nonlinear state model (1a).

There exists several methods that can be utilized to obtain a discrete-time model. The choice of discretization method must rely upon the system's *stiffness*, desired accuracy, and the performance of computation resources. In this paper, we consider the discretization of (1) using the *implicit Runge-Kutta* (IRK) method [18]. The main advantage of IRK method is that it can be applied to a wider class of nonlinear dynamic networks with various degree of stiffness. Readers can refer to [19] for the discrete-time modeling techniques of nonlinear systems. The methodology herein results in the following implicit discrete-time state-space model

$$\begin{aligned} \zeta_{1,k+1} &= \mathbf{x}_k + \frac{T}{4} (\mathbf{f}(\zeta_{1,k+1}) - \mathbf{f}(\zeta_{2,k+1})), \\ \zeta_{2,k+1} &= \mathbf{x}_k + \frac{T}{12} (3\mathbf{f}(\zeta_{1,k+1}) + 5\mathbf{f}(\zeta_{2,k+1})), \\ \mathbf{x}_{k+1} &= \mathbf{x}_k + \frac{T}{4} (\mathbf{f}(\zeta_{1,k+1}) + 3\mathbf{f}(\zeta_{2,k+1})), \end{aligned} \quad (2)$$

where  $T > 0$  denotes the discretization period,  $k \in \mathbb{N}$  is the discrete-time index such that  $\mathbf{x}_k = \mathbf{x}(kT)$ , and  $\zeta_{1,k+1}, \zeta_{2,k+1} \in \mathbb{R}^{n_x}$  are auxiliary vectors for computing  $\mathbf{x}_{k+1}$  provided that  $\mathbf{x}_k$  is given. Notice that in order to compute  $\mathbf{x}_{k+1}$ , we first need to solve a system of nonlinear equations that consists of the first two equations in (2). The unknowns in this system are  $\zeta_{1,k+1}, \zeta_{2,k+1}$ . This layer of complexity is necessary since the introduced discrete-time model can accurately and in a numerically stable manner represent a broad class of nonlinear networks, including networks with stiff dynamics.

### A. Initial State Estimation

Taking into account the model (2), the discrete-time equivalent of nonlinear dynamic network (1) can be compactly written in the following form

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \tilde{\mathbf{f}}(\mathbf{x}_{k+1}, \mathbf{x}_k), \quad (3a)$$

$$\mathbf{y}_k = \mathbf{\Gamma} \mathbf{C} \mathbf{x}_k, \quad (3b)$$

where the function  $\tilde{\mathbf{f}}(\cdot)$  in (3a) represents the implicit dynamics in (2). The proposed approach for sensor selection is developed using the concept of an open-loop lifted observer framework. To that end, we introduce the lifted vector  $\tilde{\mathbf{y}} \in \mathbb{R}^{N \cdot n_y}$  that is constructed as  $\tilde{\mathbf{y}} = \text{col}\{\tilde{\mathbf{y}}_i\}_{i=1}^{N-1}$ . The positive integer  $N$  is the observation window. For the sake of simplicity, it is assumed temporarily that  $\mathbf{\Gamma}$  is fixed such that the output measurement equation (3b) is reduced to  $\tilde{\mathbf{y}}_k = \tilde{\mathbf{C}} \mathbf{x}_k$ , where  $\tilde{\mathbf{C}}$  is obtained by compressing the zero rows of  $\mathbf{\Gamma} \mathbf{C}$ . Now, define the vector function  $\mathbf{h} : \mathbb{R}^{n_x} \rightarrow \mathbb{R}^{N \cdot n_y}$  as

$$\underbrace{\begin{bmatrix} h_0(x_0) \\ h_1(x_0) \\ \vdots \\ h_{N-1}(x_0) \end{bmatrix}}_{\mathbf{h}(x_0)} := \underbrace{\begin{bmatrix} \tilde{y}_0 \\ \tilde{y}_1 \\ \vdots \\ \tilde{y}_{N-1} \end{bmatrix}}_{\tilde{\mathbf{y}}} - \underbrace{\begin{bmatrix} g_0(x_0) \\ g_1(x_0) \\ \vdots \\ g_{N-1}(x_0) \end{bmatrix}}_{\mathbf{g}(x_0)}, \quad (4)$$

where  $\mathbf{h}(x_0) = \text{col}\{\mathbf{h}_i(x_0)\}_{i=0}^{N-1}$ . The function  $\mathbf{g} : \mathbb{R}^{n_x} \rightarrow \mathbb{R}^{N \cdot n_y}$  is constructed as  $\mathbf{g}(x_0) = \text{col}\{\mathbf{g}_i(x_0)\}_{i=0}^{N-1}$ , where  $\mathbf{g}_i : \mathbb{R}^{n_x} \rightarrow \mathbb{R}^{n_y}$  and  $\mathbf{g}_i := \tilde{\mathbf{C}}\mathbf{x}_i$  for all  $i \in \{0, 1, 2, \dots, N-1\}$ . It is understood from (3a) that  $\mathbf{g}_i$  is a function of only of the initial state  $\mathbf{x}_0$  due to the fact that  $\mathbf{x}_i$  is a recursive function of  $\mathbf{x}_0$  for each  $i$ . Consequently, we can write

$$\mathbf{h}(x_0) = 0 \Leftrightarrow \tilde{\mathbf{y}} = \mathbf{g}(x_0). \quad (5)$$

Since in practice the actual initial state is unknown *a priori*, then for a fixed selection of sensors,  $\mathbf{x}_0$  can be estimated by solving the following nonlinear state estimation optimization problem with a predefined weighting matrix  $\mathbf{Q} \succ 0$  such that

$$(\mathbf{P1}) \quad \underset{\hat{\mathbf{x}}_0 \in \mathcal{X}}{\text{minimize}} \quad \mathbf{h}(\hat{\mathbf{x}}_0)^\top \mathbf{Q} \mathbf{h}(\hat{\mathbf{x}}_0) \quad (6a)$$

$$\text{subject to} \quad \hat{\mathbf{x}}_0^l \leq \hat{\mathbf{x}}_0 \leq \hat{\mathbf{x}}_0^u, \quad (6b)$$

where  $\hat{\mathbf{x}}_0^l$  and  $\hat{\mathbf{x}}_0^u$  are respectively the lower and upper bounds of  $\hat{\mathbf{x}}_0$  and  $\mathbf{Q} \in \mathbb{R}^{N \cdot n_y \times N \cdot n_y}$  is the weighting matrix. The weight matrix  $\mathbf{Q}$  assigns weights to the measured states such that **P1**, the initial state estimation problem, efficiently reaches a solution.

### B. Observability-based Sensor Node Selection

Observability of nonlinear discrete-time systems can be quantified using the concept of *uniform observability* [20]. That is, the system (3a) with the measurement model (3b) is said to be uniformly observable in  $\mathcal{X}$  ( $\mathcal{X}$  is the subset representing a local operating region of (3a)) if there exists a finite  $N \in \mathbb{N}$  such that the relation  $\tilde{\mathbf{y}} = \mathbf{g}(x_0)$  is injective (one-to-one) with respect to  $x_0 \in \mathcal{X}$  for any given set of measured outputs  $\tilde{\mathbf{y}}$ .

Accordingly, if  $\mathbf{g}(\cdot)$  is injective with respect to  $x_0$ , then  $x_0$  can be uniquely determined from the set of measurements  $\tilde{\mathbf{y}}$ .

As such, let  $\mathbf{J}_g(\cdot)$  be a Jacobian matrix of the function  $\mathbf{g}(\cdot)$  around  $x_0$ . A *sufficient* condition for the mapping  $\mathbf{g}(\cdot)$  to be injective is that the Jacobian matrix of  $\mathbf{g}(\cdot)$  is of full rank [20].

The Jacobian matrix  $\mathbf{J}_g(\cdot) \in \mathbb{R}^{N \cdot n_y \times n_x}$  is given as

$$\mathbf{J}_g(\gamma, x_0) := \frac{d\mathbf{g}(x_0)}{dx_0} = \text{col} \left\{ \frac{\partial \mathbf{g}_i(x_0)}{\partial x_0} \right\}_{i=0}^{N-1}. \quad (7)$$

For each  $i \in \{0, 1, 2, \dots, N-1\}$ , the term  $\frac{\partial \mathbf{g}_i(x_0)}{\partial x_0}$  in (7) is equivalent to

$$\frac{\partial \mathbf{g}_i(x_0)}{\partial x_0} = \frac{\partial}{\partial x_0} \tilde{\mathbf{C}}\mathbf{x}_i = \tilde{\mathbf{C}} \frac{\partial \mathbf{x}_i}{\partial x_0} = \tilde{\mathbf{C}} \prod_{j=0}^{i-1} \frac{\partial \mathbf{x}_{j+1}}{\partial \mathbf{x}_j}. \quad (8)$$

It is important to mention that the computation of  $\frac{\partial \mathbf{x}_i}{\partial x_0}$  in (8) requires the knowledge of  $\mathbf{x}_j$  for all  $j$ . The value of  $\mathbf{x}_j$  can be obtained by simulating (3a) with the initial condition  $\mathbf{x}_0$ . Taking into account the parameterized measurement equation (3b), the Jacobian matrix  $\mathbf{J}_g(\cdot)$  in (7) around a specific initial state  $\hat{\mathbf{x}}_0$  is given as

$$\mathbf{J}_g(\gamma, \hat{\mathbf{x}}_0) := \mathbf{J}_g(\hat{\mathbf{x}}_0) = [\mathbf{I} \otimes \mathbf{\Gamma} \mathbf{C}] \times \boldsymbol{\xi}(\hat{\mathbf{x}}_0), \quad (9)$$

where  $\boldsymbol{\xi} : \mathbb{R}^{n_x} \rightarrow \mathbb{R}^{N \cdot n_x}$ ,  $\boldsymbol{\xi}(\hat{\mathbf{x}}_0) = \text{col}\{\boldsymbol{\xi}_i(\hat{\mathbf{x}}_0)\}_{i=0}^{N-1}$ , and  $\boldsymbol{\xi}_i := \frac{\partial \mathbf{x}_i}{\partial \hat{\mathbf{x}}_0}$ . Next, we define the matrix function  $\mathbf{W}(\cdot) : \mathbb{R}^{n_x} \rightarrow \mathbb{R}^{n_x \times n_x}$  as the following

$$\mathbf{W}(\gamma, x_0) := \mathbf{J}_g^\top(x_0) \mathbf{J}_g(x_0). \quad (10)$$

The matrix  $\mathbf{W}(\cdot)$  is fundamental for the analysis and solving the system of nonlinear equations as well as for the development of methods presented in this paper. Namely, the spectral properties of the matrix  $\mathbf{W}(\cdot)$  determine the convergence properties of the Newton's method used for solving the system of nonlinear equations (5) [21]. Note that, in a general case, this matrix is *not* equal to the observability Gramian for linear systems, since constructing it involves the computation of partial derivatives of the IRK equations (2). We note here that we have referred to the use of implicit IRK method since it accounts for a wide class of nonlinear networks, however other implicit discretization methods can be utilized to formulate the observability-based sensor selection problem. In our previous work [22], we perform optimal sensor selection for a class of differential algebraic equations under the trapezoidal implicit method [19] discretization.

Motivated by the fact that this matrix is closely related to the empirical observability Gramian [13], [14], we will refer to this matrix as the observability Gramian of the discrete-time system (3a)-(3b). Notice that the Gramian matrix (10) contains the matrix  $\tilde{\mathbf{C}}$ , which is a function of the vector  $\gamma$ .

To that end, the sensor selection problem can be mathematically formulated as follows. Let  $X = \{\hat{\mathbf{x}}_0^{(1)}, \hat{\mathbf{x}}_0^{(2)}, \dots, \hat{\mathbf{x}}_0^{(q)}\}$  be a set of initial conditions of the dynamics (3a)-(3b). This set of initial conditions is chosen by the user. Furthermore, let  $r$  be the final number of sensor nodes that is also specified by the user. Then, the sensor nodes are selected as the solution of the following integer optimization problem

$$(\mathbf{P2}) \quad \underset{\gamma}{\text{maximize}} \quad \mathcal{O}(\gamma, X) \quad (11a)$$

$$\text{subject to} \quad \sum_{i=1}^{n_y} \gamma_i = r, \quad \gamma \in \{0, 1\}^{n_y}, \quad (11b)$$

where  $\mathcal{O}(\gamma, X)$  is a user-selected function that quantifies the observability of the system.

The main idea of our approach is to incorporate a number of initial conditions into the function  $\mathcal{O}(\cdot)$  that quantifies the system observability. This is because the "exact" initial condition of the system is usually uncertain. By relying on a state-averaged observability matrix that is constructed under several predictions/perturbations of the initial state, the sensor selection procedure becomes less sensitive to uncertainties on initial states that are necessary to solve the system of nonlinear equations.

One approach for tackling sensor selection problems within networks, is posing such combinatorial problem as a constraint set maximization problem [10], [23]. The rationality behind such approach is later evident when solving to the sensor selection problem, where underlying set function properties allow for a scalable solution to the optimization problem. As such, the sensor node selection problem **P2** can be rewritten as a set maximization problem **P3** by defining the *set function*  $\mathcal{O}(\mathcal{S}) : 2^{\mathcal{V}} \rightarrow \mathbb{R}$  with  $\mathcal{V} := \{i \in \mathbb{N} \mid 0 < i \leq n_y\}$ . Herein, the

set  $\mathcal{V}$  denotes the set of all possible combinations of sensor locations.

$$(\mathbf{P3}) \quad \mathcal{O}_{\mathcal{S}}^* := \max_{\mathcal{S} \subseteq \mathcal{V}} f(\mathcal{S}), \text{ subject to } |\mathcal{S}| = r. \quad (12)$$

In the context of sensor selection, **P3** translates to the problem of finding the best sensor configuration  $\mathcal{S}$  containing  $r$  number of sensors such that a particular observability metric is maximized. The variable  $\Gamma$  is encoded in the set  $\mathcal{S}$ , such that for each sensor node a value of  $\gamma_j$  is attributed to the set  $\mathcal{S}$  at location  $j$ .

### III. OBSERVABILITY-BASED SENSOR SELECTION

In this section, we introduce several observability measures, quantify their properties, and present our approach for solving the problem (11). Our approach is based on defining a state-averaged observability measure and using a greedy algorithm to efficiently solve the sensor selection problem. The justification of using the greedy algorithm will be established by showing that the introduced set function measures retain set function properties—modularity or submodularity. For the development of our approach we need to obtain a closed-form expression for (10). The following proposition establishes this expression.

**Proposition 1.** *The parametrized observability Gramian (10) for the nonlinear discrete-time dynamic networks (3a) with parametrized measurement model (3b) around a particular initial state  $\hat{x}_0$  can be expressed as follows*

$$\mathbf{W}(\gamma, \hat{x}_0) = \sum_{j=1}^{n_y} \gamma_j \left( \sum_{i=0}^{N-1} \left( \frac{\partial \hat{x}_i}{\partial \hat{x}_0} \right)^\top \mathbf{c}_j^\top \mathbf{c}_j \frac{\partial \hat{x}_i}{\partial \hat{x}_0} \right), \quad (13)$$

where  $\mathbf{c}_j \in \mathbb{R}^{1 \times n_x}$  is the  $j$ -th row of  $\mathbf{C}$ .

*Proof.* From (9) and (10), it follows that

$$\begin{aligned} \mathbf{W}(\gamma, \hat{x}_0) &= \boldsymbol{\xi}^\top(\hat{x}_0) [\mathbf{I} \otimes \boldsymbol{\Gamma} \mathbf{C}]^\top [\mathbf{I} \otimes \boldsymbol{\Gamma} \mathbf{C}] \boldsymbol{\xi}(\hat{x}_0) \\ &= \sum_{i=0}^{M-1} \boldsymbol{\xi}_i^\top(\hat{x}_0) \mathbf{C}^\top \boldsymbol{\Gamma}^2 \mathbf{C} \boldsymbol{\xi}_i(\hat{x}_0) \\ &= \sum_{i=0}^{M-1} \boldsymbol{\xi}_i^\top(\hat{x}_0) \left( \sum_{j=1}^{n_y} \gamma_j \mathbf{c}_j^\top \mathbf{c}_j \right) \boldsymbol{\xi}_i(\hat{x}_0) \\ &= \sum_{i=0}^{M-1} \sum_{j=1}^{n_y} \gamma_j \boldsymbol{\xi}_i^\top(\hat{x}_0) \mathbf{c}_j^\top \mathbf{c}_j \boldsymbol{\xi}_i(\hat{x}_0), \end{aligned} \quad (14)$$

which holds since  $\boldsymbol{\Gamma}^2 = \boldsymbol{\Gamma}$ . Since (14) is equivalent to (13), then the proof is complete. ■

#### A. Modular & Submodular Set Functions

There exist several observability measures and metrics. Usually, these measures are expressed on the basis of the rank, smallest eigenvalue, condition number, trace, and determinant of an appropriate matrix—see [13] and the references therein. Such measures have set function properties, modularity and submodularity, that allow greedy algorithm to solve the optimal sensor selection problem. The following definition characterizes modular and submodular set functions [10], [24].

**Definition 1.** *A set function  $\mathcal{O} : 2^{\mathcal{V}} \rightarrow \mathbb{R}$  is said to be modular if and only if for any  $\mathcal{S} \subseteq \mathcal{V}$  and weight function  $w : \mathcal{V} \rightarrow \mathbb{R}$  it holds that*

$$\mathcal{O}(\mathcal{S}) = w(\emptyset) + \sum_{s \in \mathcal{S}} w(s), \quad (15a)$$

*and  $\mathcal{O}(\cdot)$  is said to be submodular if and only if for any  $\mathcal{A}, \mathcal{B} \subseteq \mathcal{V}$  given that  $\mathcal{A} \subseteq \mathcal{B}$ , it holds that for all  $s \notin \mathcal{B}$*

$$\mathcal{O}(\mathcal{A} \cup \{s\}) - \mathcal{O}(\mathcal{A}) \geq \mathcal{O}(\mathcal{B} \cup \{s\}) - \mathcal{O}(\mathcal{B}). \quad (15b)$$

As seen from (15b), for any submodular function, the addition of an element  $s$  to a smaller subset  $\mathcal{A}$  yields a greater reward compared to adding the same element to a bigger subset  $\mathcal{B}$ . This notion is normally termed as *diminishing return property* [10]. Aside from modularity and submodularity, the notion of monotone increasing and decreasing functions are also important to achieve scalable sensor selection.

**Definition 2.** *A set function  $\mathcal{O} : 2^{\mathcal{V}} \rightarrow \mathbb{R}$  is called monotone increasing if, for  $\mathcal{A}, \mathcal{B} \subseteq \mathcal{V}$ ,  $\mathcal{A} \subseteq \mathcal{B}$  implies  $\mathcal{O}(\mathcal{B}) \geq \mathcal{O}(\mathcal{A})$  and called monotone decreasing if  $\mathcal{A} \subseteq \mathcal{B}$  implies  $\mathcal{O}(\mathcal{A}) \geq \mathcal{O}(\mathcal{B})$ .*

In retrospect with the sensor selection problem posed in **P3**, the parametrized observability Gramian associated with  $\mathcal{S} \subseteq \mathcal{V}$  around a presumed initial state  $\hat{x}_0$  is defined as

$$\tilde{\mathbf{W}}(\mathcal{S}, \hat{x}_0) := \sum_{j \in \mathcal{S}} \left( \sum_{i=0}^{N-1} (\xi_i)^\top \mathbf{c}_j^\top \mathbf{c}_j \xi_i \right). \quad (16)$$

It is worthwhile to note that the notation  $j \in \mathcal{S}$  corresponds to every activated sensor such that  $\gamma_j = 1$ . If the chosen form of the observability measure function renders **P3** to be submodular and monotone increasing, then the greedy algorithm can be used to efficiently determine sensor locations. The greedy algorithm is summarized in Algorithm 1. If the function  $\mathcal{O}(\cdot)$  is submodular and monotone increasing, and if the set of sensor locations computed using the greedy algorithm is  $\mathcal{S}$ , then we have the following performance guarantee [25]

$$\frac{\mathcal{O}^* - \mathcal{O}(\mathcal{S})}{\mathcal{O}^* - \mathcal{O}(\emptyset)} \leq \left( \frac{r-1}{r} \right)^r \leq \frac{1}{e},$$

where  $\mathcal{O}^*$  is the optimal value of **P3** and  $e \approx 2.71828$ . Note that the above worst-case bound is merely theoretical. For submodular set maximization it has been shown that an accuracy of 99% is achieved [10].

---

#### Algorithm 1: Greedy Algorithm [10]

---

```

1 input:  $r, \mathcal{V}$ 
2 initialize:  $\mathcal{S} \leftarrow \emptyset, k \leftarrow 1$ 
3 while  $k \leq r$  do
4   compute:  $\mathcal{G}_k = \mathcal{O}(\mathcal{S} \cup \{a\}) - \mathcal{O}(\mathcal{S}), \forall a \in \mathcal{V} \setminus \mathcal{S}$ 
5   assign:  $\mathcal{S} \leftarrow \mathcal{S} \cup \left\{ \arg \max_{a \in \mathcal{V} \setminus \mathcal{S}} \mathcal{G}_k \right\}$ 
6   update:  $k \leftarrow k + 1$ 
7 output:  $\mathcal{S}$ 

```

---

#### B. State-Averaged Observability Sensor Selection

Ideally, the parametrized Gramian (13) should be constructed using the system's actual initial state  $\mathbf{x}_0$ . Nonetheless,

this state is usually unknown or only some vector entries are known *a priori*. In practice, we only have a guess of the initial state, that is denoted by  $\hat{x}_0$ . To minimize the variability from quantifying the observability around  $\hat{x}_0$ , we opt to use a state-averaged observability metric which, instead of computing the observability Gramian around a single guess of initial state  $\hat{x}_0$ , alternatively it is computed by taking into account several points of presumed initial states  $\hat{x}_0^{(\kappa)}$  for  $\kappa \in \{1, 2, \dots, q\}$ . Using this concept of state-averaged observability, we introduce the following metric

$$\mathcal{O}(\mathcal{S}, X) = \frac{1}{q} \sum_{\kappa=1}^q \mathcal{L}\left(\tilde{\mathbf{W}}^{(\kappa)}(\mathcal{S}, \hat{x}_0^{(\kappa)})\right), \quad (17)$$

where  $\mathcal{L}(\cdot)$  is an appropriate function mapping matrix into a scalar

$$\tilde{\mathbf{W}}^{(\kappa)}(\mathcal{S}, \hat{x}_0^{(\kappa)}) := \sum_{j \in \mathcal{S}} \left( \sum_{i=0}^{M-1} \left( \xi_i^{(\kappa)} \right)^\top \mathbf{c}_j^\top \mathbf{c}_j \xi_i^{(\kappa)} \right). \quad (18)$$

This form of the Gramian matrix is established on the basis of Proposition 1. Using this new measure, **P3** is developed further into the following set optimization problem

$$(\text{P4}) \quad \underset{\mathcal{S}}{\text{maximize}} \quad \mathcal{O}(\mathcal{S}) := \frac{1}{q} \sum_{\kappa=1}^q \mathcal{L}\left(\tilde{\mathbf{W}}^{(\kappa)}(\mathcal{S})\right), \quad (19a)$$

$$\text{subject to} \quad |\mathcal{S}| = r, \mathcal{S} \subseteq \mathcal{V}. \quad (19b)$$

### C. Modularity & Submodularity of the Proposed Measures

In the sequel we will analyze the modularity and submodularity properties of the average observability metrics (17). We will analyze the cases when the function  $\mathcal{L}(\cdot)$  is trace, and log-det. The following Lemma provides support to the analysis on the modularity, submodularity, and monotonicity properties of the average observability metric  $\mathcal{O}(\cdot)$ , when the function  $\mathcal{L}(\cdot)$  in **P4** takes the form of the trace and log-det.

**Lemma 1.** For set functions  $\mathcal{L}_1, \mathcal{L}_2, \dots, \mathcal{L}_k : 2^{\mathcal{V}} \rightarrow \mathbb{R}$  that are submodular. Any conic combination, that is, any weighted non-negative sum defined as

$$\mathcal{O}(\mathcal{S}) := \sum_{\kappa=1}^q w_k \mathcal{L}_k, \quad (20)$$

is submodular, such that  $w_k \geq 0 \forall k$ .

*Proof.* We prove the submodularity of a non-negative weighted sum from the definition of submodularity. As such, from Def. 1, we have  $\mathcal{A}, \mathcal{B} \subseteq \mathcal{V}$  given that  $\mathcal{A} \subseteq \mathcal{B}$ , and that for all  $s \notin \mathcal{B}$

$$\mathcal{L}_k(\mathcal{A} \cup \{s\}) - \mathcal{L}_k(\mathcal{A}) \geq \mathcal{L}_k(\mathcal{B} \cup \{s\}) - \mathcal{L}_k(\mathcal{B}),$$

then under a conic combination and based on Def. 1 the following holds true

$$\begin{aligned} & \sum_{\kappa=1}^q w_k \left( \mathcal{L}_k(\mathcal{A} \cup \{s\}) - \mathcal{L}_k(\mathcal{A}) \right) \\ & \geq \sum_{\kappa=1}^q w_k \left( \mathcal{L}_k(\mathcal{B} \cup \{s\}) - \mathcal{L}_k(\mathcal{B}) \right), \end{aligned}$$

for any  $\mathcal{A}, \mathcal{B} \subseteq \mathcal{V}$  given that  $\mathcal{A} \subseteq \mathcal{B}$ , and for all  $s \notin \mathcal{B}$ . ■

Conic combinations along with set restrictions and contractions are submodularity preserving operations [26]. Lemma 1

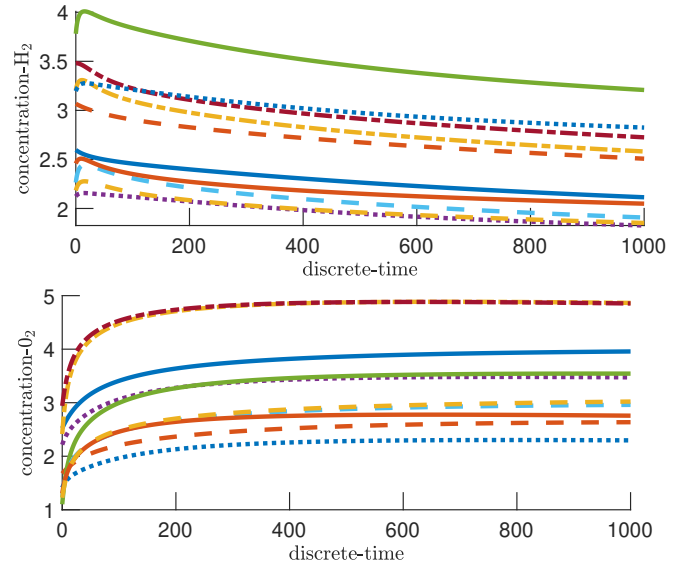


Figure 1. State trajectories of the simulated  $H_2/O_2$  combustion network under perturbed initial conditions. The states are concentrations of  $H_2$  and  $O_2$  chemical species.

shows that submodularity of the original submodular functions is retained under a non-negative weighted sum and thus formulates the rationale behind developing a state-averaged observability metric. As such, the following proposition shows that the state-averaged  $\text{trace}(\cdot)$  metric is modular.

**Proposition 2.** A set function  $\mathcal{O} : 2^{\mathcal{V}} \rightarrow \mathbb{R}$  defined by

$$\mathcal{O}(\mathcal{S}) := \frac{1}{q} \sum_{\kappa=1}^q \text{trace}\left(\tilde{\mathbf{W}}^{(\kappa)}(\mathcal{S})\right), \quad (21)$$

for  $\mathcal{S} \subseteq \mathcal{V}$  is modular.

*Proof.* For any  $\mathcal{S} \subseteq \mathcal{V}$ , observe that

$$\begin{aligned} & \frac{1}{q} \sum_{\kappa=1}^q \text{trace}\left(\tilde{\mathbf{W}}^{(\kappa)}(\mathcal{S})\right) \\ &= \frac{1}{q} \sum_{\kappa=1}^q \text{trace}\left(\sum_{j \in \mathcal{S}} \left( \sum_{i=0}^{M-1} \left( \xi_i^{(\kappa)} \right)^\top \mathbf{c}_j^\top \mathbf{c}_j \xi_i^{(\kappa)} \right)\right) \\ &= \sum_{j \in \mathcal{S}} \left( \frac{1}{q} \sum_{\kappa=1}^q \text{trace}\left(\sum_{i=0}^{M-1} \left( \xi_i^{(\kappa)} \right)^\top \mathbf{c}_j^\top \mathbf{c}_j \xi_i^{(\kappa)} \right) \right), \end{aligned}$$

thus showing that  $\text{trace}(\cdot)$  is a linear matrix function and therefore is modular. ■

The state-averaged  $\log\text{-det}(\cdot)$  observability metric is submodular and monotone increasing.

**Proposition 3.** A set function  $\mathcal{O} : 2^{\mathcal{V}} \rightarrow \mathbb{R}$  characterized by

$$\mathcal{O}(\mathcal{S}) := \frac{1}{q} \sum_{\kappa=1}^q \log \det\left(\tilde{\mathbf{W}}^{(\kappa)}(\mathcal{S})\right), \quad (22)$$

for  $\mathcal{S} \subseteq \mathcal{V}$  is submodular and monotone increasing.

*Proof.* For brevity we do not provide the full proof regarding the submodularity and the increasing monotonicity of the  $\log\text{-det}(\cdot)$ . Such metric is well studied in the field of submodular optimization and is proved therein—readers are referred to [10], [27]–[29]. For the state-averaged observability metric

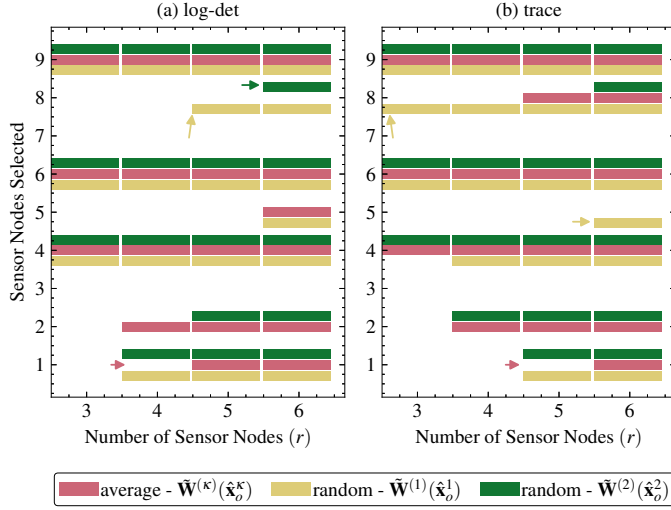


Figure 2. Selected sensor nodes resulting from state-averaged observability measures and observability measures that are based on a single randomly selected initial condition. Arrows represent the changes relative to state-averaged metrics.

$\mathcal{O}(\mathcal{S})$  in (22) and based on Lemma 1, the submodularity of the set function  $\log\text{-det}(\cdot)$  under a non-negative weighted sum is preserved and thus it is submodular. ■

The following section showcases the robustness of the sensor selection problem that is based on the proposed state-averaged observability metrics and that it is solved via scalable greedy heuristics.

#### IV. NUMERICAL STUDIES

In this section, we numerically validate and investigate the effectiveness of the averaged-observability based sensor selection framework. To numerically test our methods, we choose a general nonlinear model of a combustion reaction network. Consider the following list of  $N_r$  chemical reactions

$$\sum_{i=1}^{n_x} q_{ji} \mathcal{R}_i \rightleftharpoons \sum_{i=1}^{n_x} w_{ji} \mathcal{R}_i, \quad j = 1, 2, \dots, N_r, \quad (23)$$

where  $q_{ji}$  and  $w_{ji}$  are stoichiometric coefficients and  $\mathcal{R}_i$ ,  $i \in \{1, 2, \dots, n_x\}$ , are chemical species (notice that the number of chemical species is equal to the global state dimension). With the chemical reactions described in (23), we associate a state-space model. In this representation, the states are concentrations of chemical species. The resulting state equation has the following form [30], [31]

$$\dot{\mathbf{x}}(t) = \Theta \psi(\mathbf{x}(t)), \quad (24)$$

where  $\psi(\mathbf{x}) = [\psi_1(\mathbf{x}), \psi_2(\mathbf{x}), \dots, \psi_{n_r}(\mathbf{x})]^T$ , and  $\Theta = [w_{ji} - q_{ji}] \in \mathbb{R}^{n_x \times N_r}$ , and  $\mathbf{x} = [x_1, x_2, \dots, x_{n_x}]$ , where  $x_i$ ,  $i \in \{1, 2, \dots, n_x\}$  are concentrations of chemical species, and finally  $\psi_j$ ,  $j = 1, 2, \dots, N_r$  are the polynomial functions of concentrations defined as follows

$$\psi_j(\mathbf{x}) = v_j \prod_{i=1}^{n_x} x_i^{q_{ji}} - b_j \prod_{i=1}^{n_x} x_i^{w_{kj}}, \quad r = 1, 2, \dots, N_r, \quad (25)$$

where  $v_j, b_j \in \mathbb{R}_+$  are the forward and backward reaction rates that are computed on the basis of the Arrhenius law.

In this paper, we consider an  $H_2/O_2$  combustion network. This network has 27 reactions and 9 chemical species. The

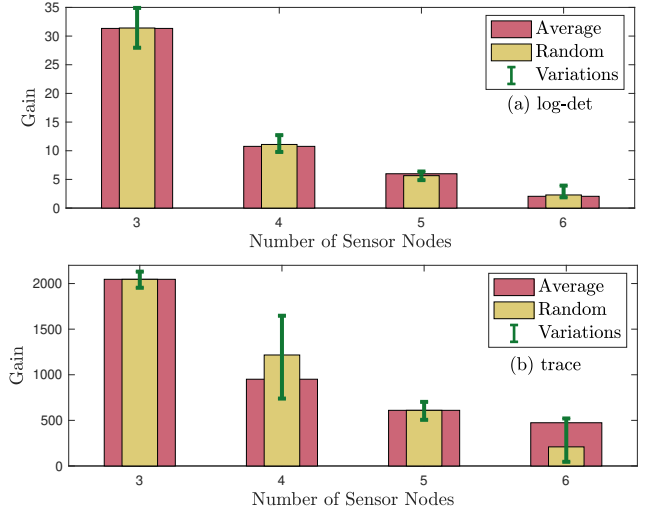


Figure 3. Gain on observability measures with additional sensor node selections. The average (red) represents that of state-averaged observability and that of perturbed single guesses is represented by an average gain (yellow) and the max and min variations on that gain (green).

reaction rates are computed using the Cantera software [32]. We use a chemical reaction network model described in the Cantera database file “h2o2.cti”. In our computations, we assume a temperature of 2500 [K] and an initial pressure equal to the atmospheric pressure. We have chosen a smaller combustion network in order to be able to compare our methods with randomized sensor node placements. To discretize the dynamics we use a discretization constant of  $T = 1 \cdot 10^{-12}$  and we assume the observation window of  $N = 1000$ . We have chosen such value of the discretization constant by analyzing an initial condition response of the system.

We assume  $\kappa = 10$  and states are selected as random perturbations of the “true” state  $\mathbf{x} = [2, 0, 0, 1, 0, 0, 0, 0.2, 0]^T$ . The random perturbations are drawn from a uniform perturbation on the interval  $[0, p]$ . Fig. 1 depicts the state trajectories of  $H_2$  and  $O_2$  from the simulated  $H_2/O_2$  combustion network under uniform perturbation interval with  $p = 2$ . We note that for each state, different state trajectories are obtained when starting with different initial condition. That is, for each initial condition, the trajectory of the system tends to a different attractor. In dynamical systems, a basin of attractor is a state condition that the systems tends towards as it evolves over a time period [33]. This suggests that the nonlinear system (24) has several basins of attraction and as such, we investigate how such perturbed state trajectories affect the sensor node selection model and asses the robustness of the proposed state-averaged observability measures.

##### A. Robust Observability-based Sensor Node Selection

Our first goal is to determine optimal sensor locations using the greedy algorithm and the proposed state-averaged observability measures. These observability measures are computed on the basis of the perturbed states (the “true” state is not used to compute the observability measure). Fig. 2 represents the sensed node locations determined by solving P4 based on the state-averaged observability metrics (log-det, trace) and

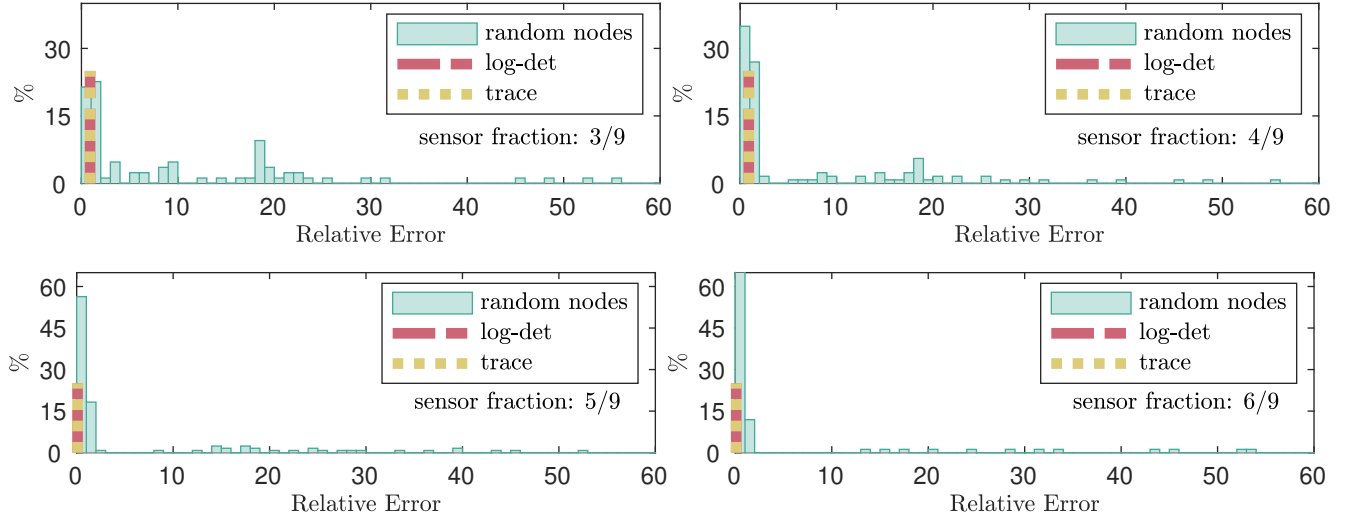


Figure 4. Relative error on initial state estimation. Histograms represent the error from randomly selected nodes with specified sensor fractions. The y-axis represents the percentage of sensor configurations that resulted in a certain relative error in the initial state estimation (histograms). The lines, having a percentage of 100%, represent the error under the optimal solutions to problem **P4**.

on the observability metrics associated with a single guess of the perturbed initial state. It can be pointed out that starting with different initial guesses, the optimal set of selected nodes is also different. The different selection as compared to that of the averaged metric are pointed by arrows. We note here that to evaluate the robustness of the state-averaged metric, we perform **P4** based on a random generation of  $\kappa = 10$  initial guesses, such that the same solution is obtained for each generated initial guesses. Thus, showing that the proposed metrics are robust to choice and perturbations of the initial conditions.

Understanding the underlying theory that allows us to optimally choose of  $\kappa$  for a specified perturbation  $p$  and ensure robustness on  $\mathcal{S}^*$  from initial state perturbations is outside the scope of this paper and will be investigated in future work. For now we note that for the presented general nonlinear network a choice of  $\kappa = 10$ , that is greater than the number of sensed node  $n_p = 9$ , results in robust optimal sensor node selections for a range of  $p \leq 20$ , which is a relatively high perturbation given the true state. As such, one can infer that the perturbation magnitude is not the critical factor for the choice of  $\kappa$ . The authors suggest that this is related to the number of states/sensor nodes and the stability of the state trajectories resulting from the perturbed states.

To further understand the performance of the state-averaged observability measure as compared with that based on a single initial guess, we investigate the variations on observability gain; Fig. 3 depicts the gain on the observability metrics resulting from an additional sensor selection relative to the prior number of sensed nodes. This variability in gain value for each of the log-det and trace metrics based on single initial conditions represents where the change in sensed nodes occurs. For instance, under states with different gain value a sensed node might be forfeited for another. It can be depicted that when the average of the gain resulting from a single perturbed initial conditions is different than that of the state-

average metric, a different sensed node is chosen. For example considering the log-det metric and when the number of nodes the be chosen is increasing from 4 to 5, we notice that this average differs. Referring back to Fig. 2 we realize that this is where the sensed nodes chosen is different than the state-averaged metric. Such change in observability degree or gain could be a result of local stability of the system dynamics associated with the perturbed initial conditions. We note, that the reason for this variability, that is postulated to be consequence of the perturbed state trajectories, is not fully investigated for that is out of the scope of this work. We here investigate the effects of the perturbations on the sensor selections and the robustness of the proposed state-averaged observability measures.

### B. Sensor Selections on Initial State Estimation

The second goal is to estimate the "true" state, using the optimally selected sensors and to show that an optimal solution is obtained using a greedy heuristic approach. By showing optimality of the greedy approach, we numerically validate the modularity and submodularity of the state-averaged observability measures.

We determine the optimal location of sensor nodes, then we compute the state estimate for this selection by solving **P1** with  $Q = I$ . State estimation results are also computed for randomly generated sensor locations under a fixed sensor fraction. The least-squares problem **P1** is solved using the MATLAB function `lsqnonlin` which implements the trust-region-reflective algorithm. We show the results for the two metrics, log-det and trace. We quantify the final estimation performance by computing the relative estimation error using the following formula  $e = \|\mathbf{x}_{\text{true}} - \hat{\mathbf{x}}\|_2 / \|\mathbf{x}_{\text{true}}\|_2$ , where  $\mathbf{x}_{\text{true}}$  is the true state that we want to estimate and  $\hat{\mathbf{x}}$  is its estimate computed by solving the nonlinear least squares problem for the fixed sensor location. The relative estimation

error is directly related to the degree of observability resulting from the placements. Meaning that for a given sensed node configuration that achieves a high observability degree, the relative error in state estimation would be minimal.

Fig. 4 shows computed relative errors for different fraction of sensor nodes. The errors represented by the histograms are computed for a random selection of sensor location for specified sensor fractions. At the same time we compute the relative error produced by the state-averaged approach for the two measures (red and yellow vertical lines). The log-det and trace measures produce optimal results for each of the specified sensor fraction. This shows that the greedy algorithm indeed results in optimal sensor placement under the state-averaged approach. Thereby providing numerical proof that submodularity and modularity is retained. And on this note, we conclude this section.

## V. PAPER SUMMARY AND FUTURE WORK

This paper investigates the robustness of an observability-based sensor selection problem towards unknown initial conditions. Specifically, our approach is built upon the open-loop lifted observer framework in which the parameterized observability Gramian is constructed. To accommodate the inaccuracy when quantifying the observability due to uncertain initial states, we introduce state-averaged observability measures. The proposed sensor selection problem posed under trace and log-det measures is shown to retain the modularity or submodularity properties. Greedy heuristics are employed to efficiently solve the optimization problem and render it scalable to larger nonlinear systems. Numerical results showcase the validity and effectiveness of proposed approach. For our future work, we will further investigate the relation between the proposed state-averaged observability metrics and empirical observability Gramian for discrete-time systems, and how to normalize nonlinear systems that have several basins of attraction such that the sensor placement is robust.

## REFERENCES

- [1] J. A. Taylor, N. Luangsomboon, and D. Fooladivanda, "Allocating sensors and actuators via optimal estimation and control," *IEEE Transactions on Control Systems Technology*, vol. 25, no. 3, pp. 1060–1067, 2017.
- [2] J. W. Berry, L. Fleischer, W. E. Hart, C. A. Phillips, and J.-P. Watson, "Sensor placement in municipal water networks," *Journal of Water Resources Planning and Management*, vol. 131, no. 3, pp. 237–243, 2005.
- [3] N. Mehr and R. Horowitz, "A submodular approach for optimal sensor placement in traffic networks," in *2018 Annual American Control Conference (ACC)*, 2018, pp. 6353–6358.
- [4] H. Zhang, R. Ayoub, and S. Sundaram, "Sensor selection for Kalman filtering of linear dynamical systems: Complexity, limitations and greedy algorithms," *Automatica*, vol. 78, pp. 202–210, 2017.
- [5] S. Pequito, S. Kar, and A. P. Aguiar, "A framework for structural input/output and control configuration selection in large-scale systems," *IEEE Transactions on Automatic Control*, vol. 61, no. 2, pp. 303–318, 2016.
- [6] N. K. Dhingra, M. R. Jovanović, and Z. Luo, "An ADMM algorithm for optimal sensor and actuator selection," in *53rd IEEE Conference on Decision and Control*, 2014, pp. 4039–4044.
- [7] A. Argha, S. W. Su, A. Savkin, and B. Celler, "A framework for optimal actuator/sensor selection in a control system," *International Journal of Control*, vol. 92, no. 2, pp. 242–260, 2019.
- [8] A. F. Taha, N. Gatsis, T. Summers, and S. A. Nugroho, "Time-varying sensor and actuator selection for uncertain cyber-physical systems," *IEEE Transactions on Control of Network Systems*, vol. 6, no. 2, pp. 750–762, 2019.
- [9] S. Joshi and S. Boyd, "Sensor selection via convex optimization," *IEEE Transactions on Signal Processing*, vol. 57, no. 2, pp. 451–462, 2009.
- [10] T. H. Summers, F. L. Cortesi, and J. Lygeros, "On submodularity and controllability in complex dynamical networks," *IEEE Transactions on Control of Network Systems*, vol. 3, no. 1, pp. 91–101, 2016.
- [11] S. A. Nugroho, A. F. Taha, N. Gatsis, T. H. Summers, and R. Krishnan, "Algorithms for joint sensor and control nodes selection in dynamic networks," *Automatica*, vol. 106, pp. 124–133, Aug. 2019.
- [12] X. Shen, S. Liu, and P. K. Varshney, "Sensor selection for nonlinear systems in large sensor networks," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 50, no. 4, pp. 2664–2678, 2014.
- [13] J. Qi, K. Sun, and W. Kang, "Optimal PMU placement for power system dynamic state estimation by using empirical observability Gramian," *IEEE Transactions on Power Systems*, vol. 30, no. 4, pp. 2041–2054, 2015.
- [14] A. Haber, F. Molnar, and A. E. Motter, "State observation and sensor selection for nonlinear networks," *IEEE Transactions on Control of Network Systems*, vol. 5, no. 2, pp. 694–708, 2018.
- [15] S. D. Bopardikar, O. Ennasr, and X. Tan, "Randomized sensor selection for nonlinear systems with application to target localization," *IEEE Robotics and Automation Letters*, vol. 4, no. 4, pp. 3553–3560, 2019.
- [16] S. A. Nugroho and A. F. Taha, "Sensor placement strategies for some classes of nonlinear dynamic systems via Lyapunov theory," in *2019 IEEE 58th Conference on Decision and Control (CDC)*, 2019, pp. 4551–4556.
- [17] A. Haber, S. A. Nugroho, P. Torres, and A. F. Taha, "Control node selection algorithm for nonlinear dynamic networks," *IEEE Control Systems Letters*, vol. 5, no. 4, pp. 1195–1200, 2021.
- [18] A. Iserles, *A First Course in the Numerical Analysis of Differential Equations*, ser. Cambridge Texts in Applied Mathematics. Cambridge University Press, 2008.
- [19] K. E. Atkinson, W. Han, and D. Stewart, *Numerical Solution of Ordinary Differential Equations*, 2011.
- [20] S. Hanba, "On the "uniform" observability of discrete-time nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 54, no. 8, pp. 1925–1928, 2009.
- [21] J. Nocedal and S. Wright, *Numerical Optimization*. Springer Science & Business Media, 2006.
- [22] M. H. Kazma and A. F. Taha, "Optimal Placement of PMUs in Power Networks: Modularity Meets A Priori Optimization," *Proceedings of the American Control Conference*, vol. 2023-June, pp. 4489–4494, 2023.
- [23] A. Krause, J. Leskovec, C. Guestrin, J. VanBriesen, and C. Faloutsos, "Efficient Sensor Placement Optimization for Securing Large Water Distribution Networks," *Journal of Water Resources Planning and Management*, vol. 134, no. 6, pp. 516–526, 2008.
- [24] L. Lovász, "Submodular functions and convexity," in *Mathematical programming the state of the art*. Springer, 1983, pp. 235–257.
- [25] G. L. Nemhauser, L. A. Wolsey, and M. L. Fisher, "An analysis of approximations for maximizing submodular set functions—i," *Mathematical programming*, vol. 14, no. 1, pp. 265–294, 1978.
- [26] F. Bach, "Learning with submodular functions: A convex optimization perspective," *Foundations and Trends in Machine Learning*, vol. 6, no. 2-3, pp. 145–373, 2013.
- [27] A. Krause and D. Golovin, "Submodular function maximization," pp. 71–104, 2011.
- [28] L. Zhou and P. Tokekar, "Sensor Assignment Algorithms to Improve Observability while Tracking Targets," *IEEE Transactions on Robotics*, vol. 35, no. 5, pp. 1206–1219, 2019.
- [29] J. Bilmes, "Submodularity In Machine Learning and Artificial Intelligence," 2022.
- [30] S. Turns, *An Introduction to Combustion: Concepts and Applications*, ser. McGraw-Hill series in mechanical engineering. McGraw-Hill, 1996.
- [31] N. Smirnov and V. Nikitin, "Modeling and simulation of hydrogen combustion in engines," *International Journal of Hydrogen Energy*, vol. 39, no. 2, pp. 1122 – 1136, 2014.
- [32] D. G. Goodwin, H. K. Moffat, and R. L. Speth, "Cantera: An object-oriented software toolkit for chemical kinetics, thermodynamics, and transport processes," <http://www.cantera.org>, accessed: 09-27-2020.
- [33] D. Dudkowski, S. Jafari, T. Kapitaniak, N. V. Kuznetsov, G. A. Leonov, and A. Prasad, *Hidden attractors in dynamical systems*. Elsevier B.V., 2016, vol. 637.