CONFIDENTIAL. Limited circulation. For review only.

Multilinear Extensions in Submodular Optimization for Optimal Sensor Scheduling in Nonlinear Networks

Mohamad H. Kazma and Ahmad F. Taha[◇]

***Abstract*—Optimal sensing nodes selection in dynamic systems is a combinatorial optimization problem that has been thoroughly studied in the recent literature. This problem can be formulated within the context of set optimization. For high-dimensional nonlinear systems, the problem is extremely difficult to solve. It scales poorly too. Current literature poses combinatorial submodular set optimization problems via maximizing observability performance metrics subject to matroid constraints. Such an approach is typically solved using greedy algorithms that require lower computational effort yet often yield sub-optimal solutions. In this paper, we address the sensing node selection problem for nonlinear dynamical networks using a variational form of the system dynamics, that basically perturb the system physics. As a result, we show that the observability performance metrics under such system representation are indeed submodular. The optimal problem is then solved using the *multilinear continuous extension*. This extension offers a computationally scalable and approximate continuous relaxation with a performance guarantee. The effectiveness of the extended submodular program is studied and compared to greedy algorithms. We demonstrate the proposed set optimization formulation for sensing node selection on nonlinear natural gas combustion networks.**

***Index Terms*—Sensing node selection, nonlinear dynamical networks, observability, submodularity, multilinear extension**

I. Introduction and Paper Contributions

For nonlinear dynamical systems, in particular high-dimensional systems, it is often costly or impractical to monitor the full state-space [1]. Thus a trade-off between cost and performance warrants solving the sensing node selection (SNS) problem posed as a constraint set optimization problem. Optimal SNS in dynamic systems is considered a combinatorial optimization problem that has been thoroughly studied in the recent literature [2]–[10]. Such problems can be formulated within the context of *set optimization*, where the decision variables are discrete sets and not vectors or matrices. There are two prominent approaches that quantify the objective function of the SNS problem: (i) objectives that are based on minimizing the state estimation error given a subset of the measured states [4], [8], [11], [12], and (ii) objectives that are based on metrics related to the system's observability Gramian [1], [3], [5], [13], [14].

A simple approach to solve the SNS problems in the aforementioned brief literature is through brute force, that is, testing all combinations of sensor nodes and selecting the subset that minimizes or maximizes the objective. This approach is infeasible and thus there are various optimization methods developed that can be used to solve the combinatorial SNS problem. Such methods can be categorized accordingly: convex relaxation heuristics [9], [15], mixed-integer semi-definite programs [2], [16], and greedy algorithms [3], [6], [17], [18].

One key property that is relevant to combinatorial set optimization problems is that of submodularity; it is a diminishing returns property which provides provable performance guarantees that allow to solve the combinatorial SNS problem via simple greedy algorithms [19]. This requires posing the SNS problem as a submodular set optimization problem of objective function f in the form

$$f^* := \underset{\mathcal{S}\subseteq\mathcal{V},\, \mathcal{S}\in\mathcal{I}}{\text{maximize}} \quad f(\mathcal{S}), \tag{1}$$

where the set $\mathcal{V}$ represent a set of available sensor nodes to choose from and the set $\mathcal{I}$ represents *matroid* constraints. A typical matroid constraint for SNS is the cardinality constraint $\mathcal{I}_c = \{\mathcal{S} \subseteq \mathcal{V} : |\mathcal{S}| = r\}$ for some r that represents the feasible number of sensing nodes. Such an approach is typically solved for $f^*_{\mathcal{S}}$ using greedy algorithms that require lower computational effort while achieving a $(1-1/e)$ performance guarantee. Given that $(1-1/e) \approx 0.63$, that is, $f^*_{\mathcal{S}}$ is at least 0.63 times the optimal value f^*.

In this paper, we address solving the SNS problem for nonlinear dynamical systems via quantifying submodular observability-based metrics while posing it as a submodular maximization problem (1). For linear systems, Gramian based observability quantification allows for scalable SNS by exploiting modular and submodular observability notions as demonstrated in [3]. As compared to nonlinear systems, observability-based SNS for nonlinear systems remains a topic of ongoing research [14]. Typically, quantifying nonlinear system observability can be approached by considering an empirical Gramian approach [20] or a Lie derivative approach [21]. Both methods can become infeasible for large scale systems when considering solving the SNS problem. A *variational* approach can be considered to handle system nonlinearities. This system can be viewed as a linear-time varying model along the system trajectory and thus an observability Gramian can be computed more efficiently [22].

Furthermore, the application of the aforementioned greedy algorithm has been studied well-studied and investigated throughout the literature for the cardinality constrained set problem (1). Recently, there has been gained interest towards extending such algorithm to handle other matroid constraints.

[◇]Corresponding author. This work is supported by National Science Foundation under Grants 2152450 and 2151571. The authors are with the Civil & Environmental Engineering and Electrical & Computer Engineering Departments, Vanderbilt University, 2201 West End Ave, Nashville, Tennessee 37235. Emails: mohamad.h.kazma@vanderbilt.edu, ahmad.taha@vanderbilt.edu.

Manuscript 1927 submitted to 2024 63rd IEEE Conference on Decision and Control (CDC). Received March 22, 2024.

In particular, the multilinear relaxation offers a powerful continuous extension to handle various matroids and non-monotone set functions [23]. The application of this extension for submodular set maximization is introduced in [24]; its applicability has gained recent interest within the literature [23], [25], [26]. The relaxed submodular maximization in general can be solved using a *continuous greedy* algorithm along with a *pipage rounding* algorithm while achieving a $(1-1/e)$ performance guarantee. Several algorithms have been developed for this continuous view point, this includes parallelization algorithms that can extensively reduce the computational effort of the maximization problem [26]. Accordingly, in this paper we apply the multilinear extension to solve the optimal SNS problem. To the best the authors' knowledge, the application of such an extension has not been applied within the context of observability-based SNS in linear and nonlinear dynamical systems.

With that in mind, the main contributions of this paper are as follows. (i) We show that the observability Gramian which is based on the variational form of a nonlinear dynamical system is a modular set function under the parametrized sensor selection formulation. (ii) We show that SNS problem that is based on observability metrics under the action of the variational Gramian are submodular and modular. In particular, we show that the trace is modular, and both the rank and $\log\det$ are monotone submodular. This is analogous to the case for a linear observability Gramian as demonstrated in [3]. (iii) We introduce and demonstrate a multilinear continuous relaxation to the SNS maximization problem (1). A continuous greedy algorithm along with a pipage rounding algorithm are used to solve the optimal problem under the cardinality constraint while achieving a worst case bound. The validity of the proposed method is demonstrated on a nonlinear combustion reaction network.

Broader Impacts. The multilinear continuous extension enables posing (1) under different constraints while solving using efficient algorithms. There are many other types of constraints besides the cardinality constraint. For example, Knapsack constraints can be considered in the context of SNS, that is, we can assign budget constraints to the cardinality set maximization problem as $\max_S f(S)$ s.t. $\sum_{s\in\mathcal{V}} c(s) \leq B$ where B is a non-negative budget. The applicability of the multilinear extension for such problem given parameter $\varepsilon > 0$ results in $(1-1/e-\epsilon)$ performance bounds; see [26]. Accordingly, such an extension enables guaranteed performance under various matroid constraints that can arise when considering SNS applications.

Notation. Let $\mathbb{N}$, $\mathbb{R}^n$, and $\mathbb{R}^{n\times m}$ denote the set of natural numbers, real-valued row vectors with size of n, and n-by-m real matrices respectively. The symbol $\otimes$ denotes the Kronecker product. The cardinality of a set $\mathcal{N}$ is denoted by $|\mathcal{N}|$. The operators $\log\det(\boldsymbol{A})$ returns the logarithmic determinant of matrix $\boldsymbol{A}$, $\mathrm{trace}(\boldsymbol{A})$ returns the trace of matrix. The operator $\{\boldsymbol{x}_i\}_{i=1}^{\mathrm{N}} \in \mathbb{R}^{\mathrm{N}n}$ constructs a column vector that concatenates vectors $\boldsymbol{x}_i \in \mathbb{R}^n$ for all $i \in \{1,2,\ldots,\mathrm{N}\}$.

Paper Organization. The paper is organized as follows: Section II provides preliminaries on nonlinear observability. Section III introduces submodular set maximization functions and formulates the multilinear extension for the SNS problem. Section IV provides evidence regarding the submodularity of the observability measures. The numerical results are presented in Section V, and Section VI concludes this paper.

II. Preliminaries on Nonlinear Observability

In this section, we formulate the nonlinear discrete-time model of a dynamical system and present its variational model representation. Subsequently, we introduce the concept of variational observability Gramian, which is shown to be equivalent to the linear observability Gramian.

A. Nonlinear Dynamical System Setup

Consider the following continuous nonlinear dynamical system with parametrized measurement equation evolving on the smooth manifold $\mathcal{M}$ which represents the state-space under the action of system dynamics.

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{f}(\boldsymbol{x}(t)), \tag{2a}$$

$$\boldsymbol{y}(t) = \tilde{\boldsymbol{C}}\boldsymbol{x}(t), \tag{2b}$$

where the system state vector evolving in $\mathcal{M}$ is denoted as $\boldsymbol{x}(t) := \boldsymbol{x} \in \mathbb{R}^{n_x}$, and $\boldsymbol{y}(t) := \boldsymbol{y} \in \mathbb{R}^{n_y}$ is the global output measurement vector. The nonlinear mapping function $\boldsymbol{f}(\cdot) : \mathcal{M} \to \mathbb{R}^{n_x}$ and nonlinear mapping measurement function $\boldsymbol{h}(\cdot) : \mathcal{M} \to \mathbb{R}^{n_y}$ are smooth and at least twice continuously differentiable. The parametrized measurement matrix $\tilde{\boldsymbol{C}} \in \mathbb{R}^{n_y\times n_x}$ represents the mapping of output states under a configuration of the sensors.

Discrete-time dynamics are more general than that in continuous-time,such perspective is more natural when considering measurement from sensors. As such, the continuous-time nonlinear dynamics network system (2) can be equivalently represented in discrete-time and compactly written in the following form

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

$$\boldsymbol{y}_k = \tilde{\boldsymbol{C}}\boldsymbol{x}_k, \tag{3b}$$

where the discretization period is denoted as $T > 0$ and $k \in \mathbb{N}$ is the discrete-time index, such that $\boldsymbol{x}_k = \boldsymbol{x}(kT)$. The nonlinear mapping function $\tilde{\boldsymbol{f}}(\cdot) \in \mathbb{R}^{n_x}$ depicts the discretized mapping function $\boldsymbol{f}(\cdot)$ for a given discretization model. The nonlinear mapping function $\tilde{\boldsymbol{f}}(\cdot)$ depends on the choice of discretization method.

In this paper, we refer to the use of the *implicit Runge-Kutta* (IRK) method [27]; it offers a wide-range of application to systems with various degrees of stiffness. There are a plethora of discretization methods that can be utilized to discretize nonlinear systems and the choice of discretization method relies on the stiff dynamics and the desired accuracy of the discretization [28]. With that in mind, the nonlinear mapping function $\tilde{\boldsymbol{f}}(\cdot)$ is defined for the IRK method can be written as

$$\tilde{\boldsymbol{f}}(\cdot) := \tfrac{T}{4}\left(\boldsymbol{f}(\boldsymbol{\zeta}_{1,k+1}) + 3\boldsymbol{f}(\boldsymbol{\zeta}_{2,k+1})\right), \tag{4}$$

where auxiliary state vectors $\boldsymbol{\zeta}_{1,k+1}, \boldsymbol{\zeta}_{2,k+1} \in \mathbb{R}^{n_x}$ are auxiliary for computing $\boldsymbol{x}_{k+1}$ provided that $\boldsymbol{x}_k$ is given.

For brevity, we do not include auxiliary vectors $\boldsymbol{\zeta}_{1,k+1}$ and $\boldsymbol{\zeta}_{2,k+1}$; refer to [28, Section 2] for the full description of the IRK method. The following assumption states that $\boldsymbol{x}$ belongs to a compact set $\mathcal{X}$ along the trajectory of the system.

Assumption 1. *For a compact set $\mathcal{X}$ that contains the set of all feasible solutions of system (2), the system trajectory remains in $\mathcal{X} \subseteq \mathcal{M}$ for any $k \geq 0$ and given any initial state $\boldsymbol{x}_0 \in \mathcal{X}_0$.*

B. Nonlinear Systems Observability: A Variational Approach

There are a plethora of different approaches for sensor selection, and this is due to the fact that the quantification of nonlinear observability lacks universal construction methods. A typical approach to evaluate a dynamical system's observability is to compute the empirical observability Gramian of the system [29]–[31]. However, there is no clear methodology for scaling the internal states and outputs measurements so that the Gramian's eigenvalues captures the local variations in states [32]. Lie derivative on the other hand are typically not considered for assessing nonlinear observability under the context of sensor selection for two reasons. (i) Lie derivatives are computationally expensive and require the calculation of higher order derivatives [33], and (ii) the resulting observability measure is a rank condition that is qualitative in nature [21], [32] and thus not suitable for sensor selection. A moving horizon approach for discretized nonlinear dynamics is introduced in [14] and further developed in our previous work [28]; it offers a more tractable and robust solution than empirical Gramian. However, the proposed approach has no clear relation to the linear observability Gramian and the notions of Lyapunov stability. Recently, observability Gramians based on a variational system representation of the nonlinear dynamics have been developed; see [22], [34]. As claimed by [22], the proposed Gramians extends the linear Gramian to nonlinear systems.

Consider two nearby trajectories $\boldsymbol{x}_k$ and $\boldsymbol{x}_k + \boldsymbol{\delta x}_k$ resulting from initial states $\boldsymbol{x}_0$ and $\boldsymbol{x}_0 + \boldsymbol{\delta x}_0$. Note that $\boldsymbol{\delta x}_0 \in \mathbb{R}^{n_x}$ is an infinitesimal perturbation $\varepsilon > 0$ to initial conditions $\boldsymbol{x}_0$ and its exponential decay or growth for $k \in \{0\,, 1\,,\; \cdots\;, \mathrm{N}-1\}$ is denoted as $\boldsymbol{\delta x}_k \in \mathbb{R}^{n_x}$. Note that M is the discrete-time interval from the simulation, that is, $\mathrm{N} := t/T$. To that end, the variational system of the discrete-time nonlinear system (3) can be written as

$$\boldsymbol{\delta x}_k = \boldsymbol{\Phi}_0^k(\boldsymbol{x}_0)\boldsymbol{\delta x}_0, \tag{5a}$$

$$\boldsymbol{\delta y}_k = \boldsymbol{\Psi}_0^k(\boldsymbol{x}_0)\boldsymbol{\delta x}_0, \tag{5b}$$

where $\boldsymbol{\Phi}_0^k(\boldsymbol{x}_0) := \left(\boldsymbol{I}_n + \frac{\partial \tilde{\boldsymbol{f}}(\boldsymbol{x}_k, \boldsymbol{x}_{k-1})}{\partial \boldsymbol{x}_k}\right)\frac{\partial \boldsymbol{x}_k}{\partial \boldsymbol{x}_0} \in \mathbb{R}^{n_x \times n_x}$ defines the variational mapping function, such that $\boldsymbol{\Phi}_0^0(\boldsymbol{x}_0) = \boldsymbol{I}_n$ and matrix $\boldsymbol{I}_{n_x} \in \mathbb{R}^{n_x \times n_x}$ is an identity matrix. The variational parametrized measurement mapping function is denoted as $\boldsymbol{\Psi}_0^k(\boldsymbol{x}_0) := \tilde{\boldsymbol{C}}\boldsymbol{\Phi}_0^k(\boldsymbol{x}_0) \in \mathbb{R}^{n_y \times n_x}$. Readers are referred to [28], [34] for the derivation of the variational system (5). For ease of notation, moving forward we remove the dependency of $\boldsymbol{\Phi}_0^k(\boldsymbol{x}_0) = \boldsymbol{\Phi}_0^k$ and $\boldsymbol{\Psi}_0^k(\boldsymbol{x}_0) = \boldsymbol{\Psi}_0^k$ on $\boldsymbol{x}_0$.

Remark 1. *Notice that $\boldsymbol{\Phi}_0^k$ represents the derivative of (3) with respect to $\boldsymbol{x}_0$ for $k \in \{0\,, 1\,,\; \cdots\;, \mathrm{N}-1\}$. This being said, the transition matrix $\boldsymbol{\Phi}_0^k$ requires the knowledge of $\boldsymbol{x}_k$ for all k. As such, we can apply the chain rule to evaluate $\boldsymbol{\Phi}_0^k$ for any discrete-time index k as*

$$\boldsymbol{\Phi}_0^k = \boldsymbol{\Phi}_{k-1}^k \boldsymbol{\Phi}_{k-2}^{k-1} \;\cdots\; \boldsymbol{\Phi}_0^1 \boldsymbol{\Phi}_0^0 = \prod_{1}^{i=k} \boldsymbol{\Phi}_{i-1}^i. \tag{6}$$

Assumption 2. *Consider the variational system (5), the state-transition matrix function $\boldsymbol{\Phi}_0^k$ is bounded, i.e.,*

$$\sup\left\{\, \left\|\boldsymbol{\Phi}_0^k\right\| : k \in \mathbb{N} \,\right\} < \infty.$$

The above assumption on the state-transition matrix is similar to having a bound on the Jacobian of the nonlinear dynamical system, that is, $\|\boldsymbol{J}(\tilde{\boldsymbol{f}}(\cdot))\| = \|\partial \tilde{\boldsymbol{f}}(\cdot)/\partial \boldsymbol{x}_k\| < \infty$.

Given the variational dynamics (5), the observability Gramian for the variational discrete-time system (5a) with parametrized measurement model around initial state $\boldsymbol{x}_0 \in \mathcal{X}_0$ satisfying assumptions 1and 2 can be expressed as

$$\boldsymbol{V}_o(\mathcal{S}) := \boldsymbol{\Psi}^\top \boldsymbol{\Psi} \in \mathbb{R}^{n_x \times n_x}, \tag{7}$$

where $\boldsymbol{\Psi} \in \mathbb{R}^{\mathrm{N}n_y \times n_x}$ denotes the observability matrix that concatenates the variational observations $\boldsymbol{\delta y}_k$ over measurement horizon the observation horizon N for $k \in \{0,\; 1,\; \cdots\;, \; \mathrm{N}-1\}$ and can be written as

$$\boldsymbol{\Psi} := \left[\boldsymbol{\Psi}_0^0\,, \boldsymbol{\Psi}_0^1\,, \boldsymbol{\Psi}_0^2\,,\; \cdots\;, \boldsymbol{\Psi}_0^{\mathrm{N}-1}\right]^\top, \tag{8}$$

where $\boldsymbol{\Psi}_0^k$ is the variational measurement mapping function defined in (5b).

Considering the SNS problem in (1), the parametrized variational observability Gramian around an initial state $\boldsymbol{x}_0$ for $\mathcal{S} \subseteq \mathcal{V}$ can be defined as

$$\boldsymbol{V}_o(\mathcal{S}) = \sum_{j \in \mathcal{S}} \left(\sum_{i=0}^{\mathrm{N}-1} \left(\boldsymbol{\varphi}_0^i\right)^\top \tilde{\boldsymbol{c}}_j^\top \tilde{\boldsymbol{c}}_j \boldsymbol{\varphi}_0^i \right), \tag{9}$$

where $\boldsymbol{\varphi}_0^i$ represents the column vectors of matrix $\boldsymbol{\Psi}_0^k$. Note that the notation $j \in \mathcal{S}$ corresponds to every activated sensor in $\mathcal{S} \subseteq \mathcal{V}$, such that $\tilde{\boldsymbol{c}}_j = 1$ for a selected sensing node, and $\tilde{\boldsymbol{c}}_j = 0$ otherwise.

Theorem 1. *The variational observability Gramian (7) reduces to the linear observability Gramian, denoted by $\boldsymbol{W}_o$, for a linear time-invariant system and is equivalent to the Empirical Gramian when considering a general nonlinear system.*

Proof. Refer to [22, Corollary 1 and Theorem 2] for the proof. □

The above equivalence relation indicates that the nonlinear system (2) is observable if and only if the $\boldsymbol{\Psi}_0^k$ is the variational measurement mapping function is full rank.

III. Submodular Set Maximization

In this section, we provide a brief review of submodularity, submodular set optimization, and present the multilinear continuous extension.

A. Submodularity and the Greedy Approach

Consider a finite set $\mathcal{V}$ and the set of all its subsets given by $2^{\mathcal{V}}$. Let a set function $f(\mathcal{S}) : 2^{\mathcal{V}} \rightarrow \mathbb{R}$ for any $\mathcal{S} \subseteq \mathcal{V}$, then the modularity and submodularity of f is defined as follows

Definition 1. *([35]) The set function f is said to be modular if for any $\mathcal{S} \subseteq \mathcal{V}$ and weight function $w : \mathcal{V} \rightarrow \mathbb{R}$ it holds true that*

$$f(\mathcal{S}) = w(\emptyset) + \sum_{s \in \mathcal{S}} w(s), \tag{10a}$$

and f is said to be submodular if for any $\mathcal{A}, \mathcal{B} \subseteq \mathcal{V}$ and $\mathcal{A} \subseteq \mathcal{B}$, it holds true that for all $s \notin \mathcal{B}$

$$f(\mathcal{A} \cup \{s\}) - f(\mathcal{A}) \geq f(\mathcal{B} \cup \{s\}) - f(\mathcal{B}). \tag{10b}$$

The above definition (10b) shows the diminishing returns property of a submodular set function. For additional definitions refer to [36, Theorem 2.2]. The set function f is said to be supermodular if the reverse inequality in (10b) holds true for all $s \notin \mathcal{B}$. Based on (10a), a function is said to modular if it is both super and submodular. We also say that a set function is normalized if $f(\emptyset) = 0$.

Definition 2. *([35]) A set function $f : 2^{\mathcal{V}} \rightarrow \mathbb{R}$ is called monotone increasing if, for $\mathcal{A}, \mathcal{B} \subseteq \mathcal{V}$, $\mathcal{A} \subseteq \mathcal{B}$ implies $f(\mathcal{B}) \geq f(\mathcal{A})$ and called monotone decreasing if $\mathcal{A} \subseteq \mathcal{B}$ implies $f(\mathcal{A}) \geq f(\mathcal{B})$.*

A submodular function is the discrete analog to concave functions. When a set function f is submodular, monotone increasing and normalized, it can be referred to as a *polymatroid* function [36]. Such functions are indicative of information, meaning that, the functions tend to give a high value from a set $\mathcal{A} \subseteq \mathcal{V}$ that has a large amount of information and give a lower value to a set $\mathcal{C} \subseteq \mathcal{V}$ of equal cardinality to $\mathcal{A}$ but has a smaller amount of information.

That being said, for a chosen objective function $f(\mathcal{S})$ that is a polymatroid function, in the context of the SNS selection problem (1), the information gain is indicative of the measurement information from a subset of sensors $j \in \mathcal{S}$ regarding the full state-space of the system. For such a submodular set function, a greedy algorithm that is summarized in Algorithm 1, offers a theoretical worst-case bound according to the following theorem.

Theorem 2. *([37]) Let $f : 2^{V} \rightarrow \mathbb{R}$ be a polymatroid function, f^* be the optimal solution of SNS problem (1) and $f^*_{\mathcal{S}}$ be the solution computed using the greedy algorithm. Then, the following performance bound holds true*

$$f^*_{\mathcal{S}} - f(\emptyset) \geq \left(1 - \frac{1}{e}\right)(f^* - f(\emptyset)), \quad \textit{with } f(\emptyset) = 0.$$

We note that the above bound is theoretical, and generally a greedy approach performs better in practice. It has been shown that a 99% accuracy can be achieved for actuator placement; see [3] and the many references that cite this work.

Algorithm 1: Greedy Algorithm

input: r, $\mathcal{V}$
initialize: $\mathcal{S} \leftarrow \emptyset$, $k \leftarrow 1$
while $i \leq r$ **do**
  compute: $\mathcal{G}_i = f(\mathcal{S} \cup \{a\}) - f(\mathcal{S})$, $\forall a \in \mathcal{V} \setminus \mathcal{S}$
  if $\mathcal{G}_i \leq 0$ **then**
    return: $\mathcal{S}$
  else
    assign: $\mathcal{S} \leftarrow \mathcal{S} \cup \left\{\arg\max_{a \in \mathcal{V} \setminus \mathcal{S}} \mathcal{G}_i\right\}$
  update: $i \leftarrow i + 1$
output: $\mathcal{S}$

B. Submodular Maximization via a Multilinear Extension

As an alternative to the greedy algorithm approach described in the aforementioned section, it is sometimes useful to solve the submodular set maximization problem continuously. This can be done by applying continuous extensions to a submodular function, that is, extending $f(\mathcal{S})$ to a function $F : [0,1]^n \rightarrow \mathbb{R}$ that agrees with $f(\mathcal{S})$ on the hypercube vertices [38]. Extensions to submodular functions include: (i) the Lovász extension [39] which is equivalent to the exact convex closure of $f(\mathcal{S})$ and the multilinear extension [24] is equivalent to an approximate concave closure.

For the purpose of submodular maximization, the multilinear extension is shown to be useful [24], whereas the Lovász extension is applicable for submodular minimization problems. The application of the multilinear relaxation in the context of observability-based SNS problem, whether for linear or nonlinear systems, has to the best of our knowledge, not been applied. For a submodular function $f : 2^{\mathcal{V}} \rightarrow \mathbb{R}$, its multilinear extension $F : [0,1]^n \rightarrow \mathbb{R}$, where $n = |\mathcal{V}|$, in the continuous space can be written as

$$F(\mathbf{x}) = \sum_{\mathcal{S} \subset \mathcal{V}} f(\mathcal{S}) \prod_{s \in \mathcal{S}} [\mathbf{x}]_s \prod_{s \notin \mathcal{S}} (1 - [\mathbf{x}]_s), \quad \mathbf{x} \in [0,1]^n. \tag{11}$$

We define $\mathcal{S}_{\mathbf{x}}$ for any $\mathbf{x} \in [0,1]^n$ such that each element $s \in \mathcal{V}$ is included in $\mathcal{S}$ with probability $[\mathbf{x}]_s$ and not included with probability $1 - [\mathbf{x}]_s$. The multilinear extension $F(\mathbf{x})$ thus extends the function evaluation over the space between the vertices of the boolean hypercube $\{0,1\}^n$ to that of the vertices of hypercube $[0,1]^n$.

The computation of the multilinear extensions is not straightforward. That being said, the extension $F(\mathbf{x})$ for any submodular function $f(\mathcal{S})$ can be approximated by randomly sampling sets $\mathcal{S}$ to the probabilities in $[\mathbf{x}]_s$ [40]. With that in mind, $F(\mathbf{x})$ can be written as

$$F(\mathbf{x}) = \mathbb{E}\left[f(\mathcal{S}_{\mathbf{x}})\right], \tag{12}$$

where $\mathbb{E}[\cdot]$ indicates the expected value. Taking the derivatives of $F(\mathbf{x})$ we obtain the following

$$\frac{\partial F(\mathbf{x})}{\partial [\mathbf{x}]_s} = \mathbb{E}\left[f(\mathcal{S}_{\mathbf{x}} \cup \{s\}) - f(\mathcal{S}_{\mathbf{x}} \backslash \{s\})\right], \tag{13}$$

Algorithm 2: Continuous Greedy Algorithm

input: multilinear extension F, ground set $\mathcal{V}$, r
initialize: $\mathbf{x} \leftarrow \mathbf{0}$, $i \leftarrow 1$
while $i \leq r$ **do**
  sample: K times of $\mathcal{S}$ from $\mathcal{V}$ according to $\mathbf{x}$
  for $s \in \mathcal{V}$ **do**
    estimate: $w_s \sim \mathbb{E}\left[f\left(\mathcal{S}_{\mathbf{x}} \cup \{s\}\right) - f\left(\mathcal{S}_{\mathbf{x}} \backslash \{s\}\right)\right]$
  solve for: $\mathcal{S}^\star = \underset{\mathcal{S} \in \mathcal{I}}{\operatorname{argmax}} \sum_{s \in \mathcal{S}} w_s$
  update: $\mathbf{x} \leftarrow \mathbf{x}_{\mathcal{S}^\star}$
  $i \leftarrow i+1$
Use pipage rounding to convert the fractional solution $\mathbf{x}^*$ to a discrete solution.

and the second order derivative for and $a, b \in \mathcal{V}$ with $a \neq b$ can be written as

$$\begin{aligned} \frac{\partial^2 F(\mathbf{x})}{\partial[\mathbf{x}]_a \partial[\mathbf{x}]_b} =& \mathbb{E}\left[f\left(\mathcal{S}_{\mathbf{x}} \cup \{a, b\}\right) - f\left(\mathcal{S}_{\mathbf{x}} \cup \{b\} \backslash \{a\}\right)\right. \\ & \left. - f\left(\mathcal{S}_{\mathbf{x}} \cup \{a\} \backslash \{b\}\right) + f\left(\mathcal{S}_{\mathbf{x}} \backslash \{a, b\}\right)\right]. \end{aligned} \tag{14}$$

We note that, the partial derivative $\frac{\partial F(\mathbf{x})}{\partial[\mathbf{x}]_s} \geq 0$ if and only if f is monotone and $\frac{\partial^2 F(\mathbf{x})}{\partial[\mathbf{x}]_a \partial[\mathbf{x}]_b} \leq 0$ if and only if f is submodular.

Considering the above, the multilinear relaxation of the submodular set maximization problem (1) under the application of the multilinear extension can be expressed as

$$F_{\mathcal{S}}^* := \underset{\mathcal{S} \subseteq \mathcal{V},\, \mathcal{S} \in \mathcal{I}_c}{\operatorname{maximize}} \quad F(\mathbf{x}). \tag{15}$$

Typically, $F(\mathbf{x})$ is concave in certain directions and convex in others, meaning that (1) is not easily solvable even under a simple cardinality constraint. In [23] a continuous greedy algorithm was developed to solve (15). The developed method solves for a fractional value of $F_{\mathcal{S}}^*$ and then utilizes a rounding algorithm to convert the fractional solution into a discrete solution.

That being said, consider the following continuous greedy algorithm that is detailed in Algorithm 2. The algorithm defines a path $\mathbf{x} : [0,1] \rightarrow \mathcal{S}_{\mathbf{x}}$, where $\mathbf{x}(0) = 0$ and $\mathbf{x}(1)$ is the output of the algorithm. In the continuous algorithm, $\mathbf{x}$ is defined by a differential equation, and the gradient of $\mathbf{x}$ is chosen greedily in $\mathcal{V}$ to maximize F, meaning that, we are maximizing $\frac{\mathrm{d}}{\mathrm{d}l}\mathbf{x}(l) = \operatorname{argmax}_{\mathbf{x} \in \mathcal{S}_{\mathbf{x}}} \frac{\partial F}{\partial[\mathbf{x}]_s}(\mathbf{x}(l))$. This is equivalent to solving for

$$\underset{\mathcal{S} \in \mathcal{I}}{\operatorname{argmax}} \sum_{s \in \mathcal{S}} w_s \sim \mathbb{E}\left[f\left(\mathcal{S}_{\mathbf{x}} \cup \{s\}\right) - f\left(\mathcal{S}_{\mathbf{x}} \backslash \{s\}\right)\right],$$

as a consequence of the equality defined in (13).

The result obtained is fractional and thus a rounding algorithm is employed to convert this fractional solution. Randomly rounding the solution does not preserve the feasibility of the constraints, in particular, equality constraints. A *pipage rounding algorithm* has been shown to efficiently round the fractional value to a discrete value without any loss in the objective value. For brevity, we do not include the pipage rounding algorithm; refer to [23, Section 3.2] for a formal description of the pipage rounding algorithm. The algorithm works by taking the fractional solution $\mathbf{x}^*$ from the continuous greedy algorithm and then gradually eliminating all the fractional variables. This is done by minimizing along the convex direction of the multilinear continuous set function. The result iterates until it agrees with the vertices of the hypercube $\{0,1\}^n$. The following Lemma 1 shows that a discrete solution is solved in polynomial time.

Lemma 1. *([23]) Given $\mathbf{x}^*$, the pipage rounding algorithm outputs in polynomial time a discrete solution $S \in \mathcal{I}_c$ of value $\mathbf{E}[f(S)] \geq F(\mathbf{x}^*)$.*

The following theorem ensures a performance bound for solving (15) via the continuous greedy algorithm.

Theorem 3. *([23]) Let $f : 2^V \rightarrow \mathbb{R}$ be a polymatroid function and $F : [0,1]^n \rightarrow \mathbb{R}$ be its multilinear extension. Let f^* be the optimal solution of SNS problem (1) and $F_{\mathcal{S}}^*$ be the solution computed using the continuous greedy algorithm. Then, the following performance bound holds true*

$$F_{\mathcal{S}}^* - f(\emptyset) \geq \left(1 - \frac{1}{e}\right)\left(f^* - f(\emptyset)\right), \quad \textit{with } f(\emptyset) = 0.$$

For both the presented submodular maximization frameworks the $1/e$ guarantee holds true regardless of the size of the initial set $\mathcal{V}$ and which polymatroid function f is being optimized. Given the aforementioned performance guarantees of the presented algorithms for submodular set maximization, the next section establishes the submodularity of certain observability measures. The observability measures are based on the parametrized variational observability Gramian (9) for nonlinear systems.

IV. Variational Observability Gramian & Submodularity

In this section, we show that certain observability metrics that are based on the variational Gramian (7) for nonlinear systems are indeed modular and submodular. This is analogous to the linear case, where certain observability metrics (i.e., trace, rank, and $\log\det$) based on the linear Gramian are shown to be modular and submodular. Such submodularity properties enable the use of a greedy algorithm along with its continuous extensions to solve the SNS problem with provable optimal error bound guarantees.

For SNS applications, network measures based on system observability are often considered for quantifying information gain from the allocation of sensor nodes within a dynamical network. Observability-based network centrality measures have key properties related to submodularity. Observability measures based on the linear observability Gramian have been shown to be submodular or modular, in particular, $\log\det$ and rank are submodular, while the trace is modular; see [3]. Nevertheless, other observability measures such as $\log\det\left(\boldsymbol{W}_o^{-1}\right)$ and $\lambda_{\min}$ are non submodular. Such important metrics have been shown to have provable guarantees when solved using greedy algorithms; refer to [5]. In this paper, we consider only the metrics that have shown to have submodular properties in the linear case. With that in mind, we show that such properties hold also true for nonlinear dynamical systems by considering such centrality measures under the action of the variational observability Gramian (9).

Accordingly, the following theorem establishes that the variational Gramian is linear matrix function with respect to the selected sensing node $j \in \mathcal{S}$. Such, property shows that the observability Gramian can be computed form the sum of the individual contributions from each sensing node. Theorem 4 is essential for the proofs related to the modularity or submodularity of the observability-based measures.

Theorem 4. *The parametrized variational observability Gramian $\boldsymbol{V}_o(\mathcal{S})$ for $\mathcal{S} \subseteq \mathcal{V}$ is modular.*

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

$$\boldsymbol{V}_o(\mathcal{S}) = \sum_{j\in\mathcal{S}} \left\{ {\boldsymbol{\Phi}_0^k}^\top \tilde{\boldsymbol{c}}_j^\top \tilde{\boldsymbol{c}}_j \boldsymbol{\Phi}_0^k \right\}_{k=0}^{\mathrm{N}-1} = \sum_{j\in\mathcal{S}} \boldsymbol{V}_o(j),$$

where $\boldsymbol{V}_o(\mathcal{S})$ is a linear matrix function of with respect to $\tilde{\boldsymbol{c}}_j$ satisfying modularity as $\boldsymbol{V}_o(\mathcal{S}) = \boldsymbol{V}_o(\emptyset) + \sum_{j\in\mathcal{S}} \boldsymbol{V}_o(j)$. ☐

Consequently, the following proposition shows that the trace of the variational Gramian (9) is modular set function.

Proposition 1. *Set function $f : 2^{\mathcal{V}} \to \mathbb{R}$ characterized by*

$$f(\mathcal{S}) = \mathrm{trace}\left(\boldsymbol{V}_o(\mathcal{S})\right), \tag{16}$$

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

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

$$\begin{aligned}\mathrm{trace}\left(\boldsymbol{V}_o(\mathcal{S})\right) &= \mathrm{trace}\left(\sum_{j\in\mathcal{S}}\left(\sum_{i=0}^{\mathrm{N}-1}\left(\boldsymbol{\varphi}_0^i\right)^\top \tilde{\boldsymbol{c}}_j^\top \tilde{\boldsymbol{c}}_j \boldsymbol{\varphi}_0^i\right)\right), \\ &= \sum_{j\in\mathcal{S}}\left(\mathrm{trace}\left(\sum_{i=0}^{\mathrm{N}-1}\left(\boldsymbol{\varphi}_0^i\right)^\top \tilde{\boldsymbol{c}}_j^\top \tilde{\boldsymbol{c}}_j \boldsymbol{\varphi}_0^i\right)\right), \\ &= \sum_{j\in\mathcal{S}}\left(\mathrm{trace}\left(\boldsymbol{V}_o(j)\right)\right),\end{aligned}$$

where the last equality is a due to the modularity of the parametrized variational Gramian. This shows that $\mathrm{trace}(\boldsymbol{V}_o(\mathcal{S}))$ is a linear matrix function and therefore is a modular set function. ☐

The following result shows that the rank of the variational Gramian (9) is submodular and monotone increasing.

Proposition 2. *Set function $f : 2^{\mathcal{V}} \to \mathbb{R}$ characterized by*

$$f(\mathcal{S}) = \mathrm{rank}\left(\boldsymbol{V}_o(\mathcal{S})\right), \tag{17}$$

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

Proof. For any $\mathcal{S} \subseteq \mathcal{V}$, first we will show that $f(\mathcal{S})$ in (17) is submodular. First, define the derived set function $f_s : 2^{\mathcal{V}\backslash\{s\}} \to \mathbb{R}$ for a $a \in \mathcal{V}$ as

$$\begin{aligned}f_s(\mathcal{S}) &= f\left(\mathcal{S}\cup\{s\}\right) - f\left(\{s\}\right), \\ &= \mathrm{rank}\left(\boldsymbol{V}_o(\mathcal{S}\cup\{s\})\right) - \mathrm{rank}\left(\boldsymbol{V}_o(\{s\})\right), \\ &= \mathrm{rank}\left(\boldsymbol{V}_o(\mathcal{S}) + \boldsymbol{V}_o(\{s\})\right) - \mathrm{rank}\left(\boldsymbol{V}_o(\{s\})\right), \\ &= \mathrm{rank}\left(\boldsymbol{V}_o(\{s\})\right) \\ &\quad - \dim\left(\mathrm{image}\left(\boldsymbol{V}_o(\mathcal{S})\right) \cap \mathrm{image}\left(\boldsymbol{V}_o(\{s\})\right)\right),\end{aligned}$$

This indicates that $f_a(\cdot)$ is monotone decreasing since $\mathrm{rank}\left(\boldsymbol{V}_o\left(\{a\}\right)\right)$ is constant while the dimension of $\mathrm{image}\left(\boldsymbol{V}_o(\mathcal{S})\right)$ is increasing with $\mathcal{S}$. This implies that $f(\cdot)$ in (17) is submodular [3], [39]. Second, it is straightforward to show that $f(\cdot)$ is also monotone increasing since for $\mathcal{A} \subseteq \mathcal{B}$ provided that $\mathcal{A}, \mathcal{B} \subseteq \mathcal{V}$ implies $f(\mathcal{B}) \geq f(\mathcal{A})$. ☐

The following result shows that the log det of the variational Gramian (9) is submodular and monotone increasing.

Proposition 3. *Set function $f : 2^{\mathcal{V}} \to \mathbb{R}$ characterized by*

$$f(\mathcal{S}) := \log\det\left(\boldsymbol{V}(\mathcal{S})\right), \tag{18}$$

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

Proof. Let $f_s : 2^{V-\{s\}} \to \mathbb{R}$ denote a derived set function defined as

$$\begin{aligned}f_s(\mathcal{S}) &= \log\det\boldsymbol{V}_o\left(\mathcal{S}\cup\{s\}\right) - \log\det\boldsymbol{V}_o(\mathcal{S}), \\ &= \log\det\left(\boldsymbol{V}_o(\mathcal{S}) + \boldsymbol{V}_o(\{s\})\right) - \log\det\boldsymbol{V}_o(\mathcal{S}).\end{aligned}$$

We first show $f_s(\mathcal{S})$ that is monotone decreasing for any $s \in V$. That being said, let $\mathcal{A} \subseteq \mathcal{B} \subseteq \mathcal{V} - \{s\}$, and let $\boldsymbol{V}_o(\tilde{\boldsymbol{c}}) = \boldsymbol{V}_o(\mathcal{A}) + \tilde{\boldsymbol{c}}\left(\boldsymbol{V}_o(\mathcal{B}) - \boldsymbol{V}_o(\mathcal{A})\right)$ for $\tilde{\boldsymbol{c}} \in [0,1]$. Then for

$$\tilde{f}_s(\boldsymbol{V}_o(\tilde{\boldsymbol{c}})) = \log\det\left(\boldsymbol{V}_o(\tilde{\boldsymbol{c}}) + \boldsymbol{V}_o(\mathcal{S})\right) - \log\det\left(\boldsymbol{V}_o(\tilde{\boldsymbol{c}})\right),$$

we obtain the following

$$\begin{aligned}&\frac{\mathrm{d}}{\mathrm{d}\tilde{\boldsymbol{c}}}\tilde{f}_s(\boldsymbol{V}_o(\tilde{\boldsymbol{c}})) \\ &\quad = \mathrm{trace}\Big[\left(\left(\boldsymbol{V}_o(\tilde{\boldsymbol{c}}) + \boldsymbol{V}_o(\mathcal{S})\right)^{-1} - \boldsymbol{V}_o(\tilde{\boldsymbol{c}})^{-1}\right) \\ &\qquad \left(\boldsymbol{V}_o(\mathcal{B}) - \boldsymbol{V}_o(\mathcal{A})\right)\Big] \leq 0.\end{aligned}$$

Such that $\left(\left(\boldsymbol{V}_o(\tilde{\boldsymbol{c}}) + \boldsymbol{V}_o(\mathcal{S})\right)^{-1} - \boldsymbol{V}_o(\tilde{\boldsymbol{c}})^{-1}\right)^{-1} \preceq 0$, and $\left(\boldsymbol{V}_o(\mathcal{B}) - \boldsymbol{V}_o(\mathcal{A})\right) \succeq 0$, then the above inequality holds. Thus, we have f_s is monotone decreasing, and $f(\mathcal{S})$ is submodular. Then, by the additive property of $\boldsymbol{V}_o(\mathcal{S})$ (see [3]) we have $f(\mathcal{S})$ being monotone increasing. The proof is analogous to [3, Theorem 6] and [7, Lemma 3] ☐

The validity of the performance guarantees for both the greedy and continuous greedy algorithms are contingent on the modularity and monotone submodularity properties shown in the aforementioned propositions and (16)–(18).

Remark 2. *Notice that, for the* log det *to be submodular and monotone increasing, the variational observability Gramian can have zero eigenvalues.*

In the study [7], the considered observability measures are based on the Lie derivative matrix $\boldsymbol{O}_l$, that being said, the above submodular properties hold true if and only if $\boldsymbol{O}_l$ is full rank. In this case when considering the variational Gramian, there is no such restriction. The submodularity of the log det still holds in rank deficient situations. Such situations can

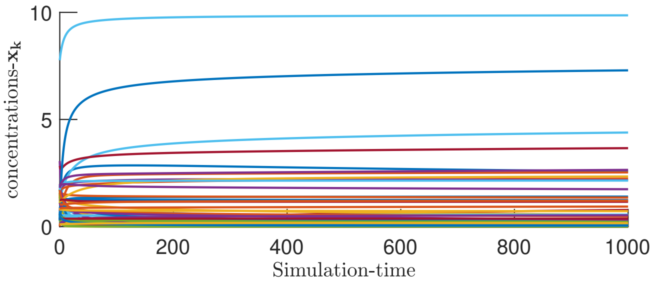

Figure 1. State Trajectories (53 states) of the combustion reaction network after disturbance.

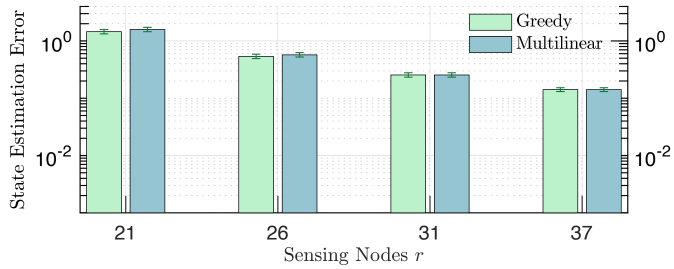

Figure 2. State estimation error based on the optimal selected node obtained from greedy algorithm (left) and continuous greedy algorithm (right).

arise when not enough sensing nodes are chosen and thereby the system is not yet fully observable. The ensuing section demonstrates the validity of the SNS problem under the action of the studied variational observability measures for nonlinear dynamical systems.

V. Numerical Case Study

In this paper, we consider a nonlinear system that represents a natural gas combustion reaction network of the form

$$\dot{\boldsymbol{x}}(t) = \Theta \boldsymbol{\psi}\left(\boldsymbol{x}(t)\right), \tag{19}$$

where the polynomial functions of concentrations ψ_j $j = \{1,\ 2,\ \ldots\ ,\ N_r\}$ are concatenated in vector $\boldsymbol{\psi}\left(\boldsymbol{x}\right) = \left[\psi_1\left(\boldsymbol{x}\right), \psi_2\left(\boldsymbol{x}\right), \ldots, \psi_{n_r}\left(\boldsymbol{x}\right)\right]^T$. The concentrations of n_x chemical species are denoted by vector $\boldsymbol{x} = [x_1,\ x_2,\ \cdots\ ,\ x_{n_x}]$. The stoichiometric coefficients q_{ji} and w_{ji} are defined by constant matrix $\Theta = [w_{ji} - q_{ji}] \in \mathbb{R}^{n_x \times N_r}$. We denote the number of chemical reactions by N_r and the list of chemical reactions can be expressed as

$$\sum_{i=1}^{n_x} q_{ji}\mathcal{R}_i \rightleftarrows \sum_{i=1}^{n_x} w_{ji}\mathcal{R}_i, \quad j \in \{1,\ 2\ ,\ \cdots\ ,\ N_r\},$$

where $\mathcal{R}_i$, $i \in \{1,\ 2,\ \cdots\ ,\ n_x\}$ represents the chemical species.

The considered network is a natural gas combustion reaction network GRI30 which has $N_r = 325$ reactions and $n_x = 53$ chemical species. For specifics regarding system parameters and definitions, we refer the readers to [14, Section V]. The discretization constant is $T = 1 \cdot 10^{-12}$ and observation window of $\mathrm{N} = 1000$ is chosen. The choice of discretization constant is a result of analyzing the system's initial condition response. The data required to calculate the reaction rates are taken from the reaction mechanisms database provided with Cantera software files. The actual initial state $\boldsymbol{x}_0 = [0, 0, 0, 2, \cdots, 1, \cdots, 7.52, \cdots, 0]$. Figure 1 depicts the state trajectory of the system discrete-time dynamics after applying a system disturbance. Meaning that, the system is simulated based on $\boldsymbol{x}_0 = \boldsymbol{x}_0 + \boldsymbol{x}_0 * \alpha_d$ where $\alpha_d \in \mathbb{R}$ is a random number between $(0, 0.2)$.

We assess the applicability and validity of the extended SNS problem (1) by comparing the state estimation error based on measurements from the optimally selected nodes to that obtained from the greedy algorithm. The optimality of the selected sensor nodes is directly related to the state estimation error. This is due to the underlying relations between observability and the ability to infer system states from limited measurement data.

Let $\mathcal{S}^*_M$ define the optimal sensor node location resulting from solving (1) using Algorithm 2 and let $\mathcal{S}^*$ define the optimal sensor node location resulting from solving (1) using Algorithm 1. That being said, let $\boldsymbol{x}_{actual}$ denote the state estimate resulting from solving the following nonlinear state estimation optimization problem expressed as

$$\underset{\tilde{\boldsymbol{x}}_0 \in \boldsymbol{\mathcal{X}}_0}{\text{minimize}} \quad \boldsymbol{g}(\tilde{\boldsymbol{x}}_0)^\top \boldsymbol{g}(\tilde{\boldsymbol{x}}_0) \tag{20}$$

$$\text{subject to} \quad \tilde{\boldsymbol{x}}_0^l \leq \tilde{\boldsymbol{x}}_0 \leq \tilde{\boldsymbol{x}}_0^u, \tag{21}$$

where $\tilde{\boldsymbol{x}}_0^l$ and $\tilde{\boldsymbol{x}}_0^u$ are respectively the lower and upper bounds of $\tilde{\boldsymbol{x}}_0$. The vector function $\boldsymbol{g}(\cdot) : \mathbb{R}^{n_x} \rightarrow \mathbb{R}^{\mathrm{N}n_y}$ that is defined as $\boldsymbol{g}(\tilde{\boldsymbol{x}}_0) := \tilde{\boldsymbol{y}} - \tilde{\boldsymbol{C}}\tilde{\boldsymbol{x}}$ represents the open-lifted observer. The measurement vector $\tilde{\boldsymbol{y}} := \{\tilde{\boldsymbol{y}}_i\}_{i=1}^{\mathrm{N}-1} \in \mathbb{R}^{\mathrm{N}n_y}$ and estimated state-vector $\tilde{\boldsymbol{x}} := \{\tilde{\boldsymbol{x}}_i\}_{i=1}^{\mathrm{N}-1} \in \mathbb{R}^{\mathrm{N}n_x}$. The above least squares optimization problem is based on an open observer framework introduced in [14]. It is solved using the trust-region-reflective algorithm on MATLAB. The state estimation error can we written as $\left\|\boldsymbol{x}_{\text{actual}} - \tilde{\boldsymbol{x}}\right\|_2 / \left\|\boldsymbol{x}_{\text{actual}}\right\|_2$. This validates the effectiveness of the solution obtained from the relaxed problem, thereby achieving the performance bound as indicated in Theorem 3.

The state estimation errors resulting from solving the SNS problems based on the aforementioned algorithms is depicted in Figure 2. The maximization problems are solved for sensor nodes cardinality constraint $\mathcal{S} = r$ with $r = [21, 26, 31, 37]$. The SNS problem is solved for both methods by considering 20 generated simulations based on $\boldsymbol{x}_0$ chosen randomly by applying perturbation α_d. The results show that the optimal solution $\mathcal{S}^*_M$ yields similar state estimation values when compared with the estimation values of optimal solution $\mathcal{S}^*$. Consequently, the application of multilinear extension for observability-based SNS can be further investigated under different matroid constraints; see Section I.

VI. Conclusion

In this paper, we showed that the SNS problem for nonlinear systems, modeled by considering the variational dynamics, can be solved as a submodular set optimization problem. In particular, we showed that metrics based on the parametrized variational Gramian (i.e., trace, rank, and $\log\det$), that extends the linear Gramian to nonlinear systems, are modular and submodular. Furthermore, we introduced a continuous extension to the submodular SNS problem. This multilinear extension presents a performance guarantee when $f(\mathcal{S})$ is monotone submodular. The resulting optimization problem is then solved using a continuous greedy algorithm and is compared to the well-known greedy algorithm that is typically used to solve submodular SNS problems.

REFERENCES

[1] K. Manohar, J. N. Kutz, and S. L. Brunton, "Optimal Sensor and Actuator Selection Using Balanced Model Reduction," *IEEE Transactions on Automatic Control*, vol. 67, no. 4, pp. 2108–2115, 2022.

[2] S. Joshi and S. Boyd, "Sensor selection via convex optimization," *IEEE Transactions on Signal Processing*, vol. 57, no. 2, pp. 451–462, 2009.

[3] 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, mar 2016.

[4] S. Liu, S. P. Chepuri, M. Fardad, E. Masazade, G. Leus, and P. K. Varshney, "Sensor selection for estimation with correlated measurement noise," *IEEE Transactions on Signal Processing*, vol. 64, no. 13, pp. 3509–3522, 2016.

[5] T. Summers and M. Kamgarpour, "Performance guarantees for greedy maximization of non-submodular controllability metrics," *2019 18th European Control Conference, ECC 2019*, pp. 2796–2801, 2019.

[6] V. Tzoumas, A. Jadbabaie, and G. J. Pappas, "Sensor placement for optimal Kalman filtering: Fundamental limits, submodularity, and algorithms," *Proceedings of the American Control Conference*, vol. 2016-July, pp. 191–196, 2016.

[7] 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.

[8] A. Kohara, K. Okano, K. Hirata, and Y. Nakamura, "Sensor placement minimizing the state estimation mean square error: Performance guarantees of greedy solutions," *Proceedings of the IEEE Conference on Decision and Control*, vol. 2020-Decem, pp. 1706–1711, 2020.

[9] A. Haber, "Joint Sensor Node Selection and State Estimation for Nonlinear Networks and Systems," *IEEE Transactions on Network Science and Engineering*, vol. 8, no. 2, pp. 1722–1732, 2021.

[10] A. P. Vinod, A. J. Thorpe, G. S. Member, P. A. Olaniyi, T. H. Summers, M. M. K. Oishi, and S. Member, "Sensor Selection for Dynamics-Driven User-Interface Design," *IEEE Transactions on Control Systems Technology*, vol. 30, no. 1, pp. 71–84, 2022.

[11] S. Joshi and S. Boyd, "Sensor selection via convex optimization," *IEEE Transactions on Signal Processing*, vol. 57, no. 2, pp. 451–462, 2009.

[12] V. Tzoumas, A. Jadbabaie, and G. J. Pappas, "Near-optimal sensor scheduling for batch state estimation: Complexity, algorithms, and limits," *2016 IEEE 55th Conference on Decision and Control, CDC 2016*, no. Cdc, pp. 2695–2702, 2016.

[13] N. Mehr and R. Horowitz, "A Submodular Approach for Optimal Sensor Placement in Traffic Networks," *Proceedings of the American Control Conference*, vol. 2018-June, pp. 6353–6358, 2018.

[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] T. Summers and I. Shames, "Convex relaxations and Gramian rank constraints for sensor and actuator selection in networks," *IEEE International Symposium on Intelligent Control - Proceedings*, vol. 2016-Septe, pp. 5–10, 2016.

[16] 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.

[17] K. Yamada, Y. Saito, K. Nankai, T. Nonomura, K. Asai, and D. Tsubakino, "Fast greedy optimization of sensor selection in measurement with correlated noise," *Mechanical Systems and Signal Processing*, vol. 158, p. 107619, 2021. [Online]. Available: https://doi.org/10.1016/j.ymssp.2021.107619

[18] A. Hashemi, M. Ghasemi, H. Vikalo, and U. Topcu, "Randomized Greedy Sensor Selection: Leveraging Weak Submodularity," *IEEE Transactions on Automatic Control*, vol. 66, no. 1, pp. 199–212, 2021.

[19] B. Guo, O. Karaca, T. Summers, and M. Kamgarpour, "Actuator Placement under Structural Controllability Using Forward and Reverse Greedy Algorithms," *IEEE Transactions on Automatic Control*, vol. 66, no. 12, pp. 5845–5860, 2021.

[20] S. Lall, J. E. Marsden, and S. Glavaški, "Empirical model reduction of controlled nonlinear systems," *IFAC Proceedings Volumes*, vol. 32, no. 2, pp. 2598–2603, 1999.

[21] A. J. Krener and A. Isidori, "Linearization by output injection and nonlinear observers," *Systems and Control Letters*, vol. 3, no. 1, pp. 47–52, 1983.

[22] M. H. Kazma and A. F. Taha, "Observability for Nonlinear Systems: Connecting Variational Dynamics, Lyapunov Exponents, and Empirical Gramians," *arXiv*, 2024. [Online]. Available: http://arxiv.org/abs/2402.14711

[23] G. Calinescu, C. Chekuri, M. Pál, and J. Vondrák, "Maximizing a monotone submodular function subject to a matroid constraint," *SIAM Journal on Computing*, vol. 40, no. 6, pp. 1740–1766, 2011.

[24] J. Vondrák, "Optimal approximation for the Submodular Welfare Problem in the value oracle model," *Proceedings of the Annual ACM Symposium on Theory of Computing*, pp. 67–74, 2008.

[25] C. Chekuri, J. Vondrák, and R. Zenklusen, "Submodular function maximization via the multilinear relaxation and contention resolution schemes," *SIAM Journal on Computing*, vol. 43, no. 6, pp. 1831–1879, 2014.

[26] C. Chekuri and K. Quanrud, "Submodular function maximization in parallel via the multilinear relaxation," *Proceedings of the Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 303–322, 2019.

[27] A. Iserles, *A First Course in the numerical analysis of differential equations*, 2nd ed. Cambridge University Press, 2009.

[28] M. H. Kazma, S. A. Nugroho, A. Haber, and A. F. Taha, "State-Robust Observability Measures for Sensor Selection in Nonlinear Dynamic Systems," *2023 62nd IEEE Conference on Decision and Control (CDC)*, no. Cdc, pp. 8418–8426, 2023.

[29] 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, jul 2015.

[30] N. L. Brace, N. B. Andrews, J. Upsal, and K. A. Morgansen, "Sensor Placement on a Cantilever Beam Using Observability Gramians," *Proceedings of the IEEE Conference on Decision and Control*, vol. 2022-Decem, no. Cdc, pp. 388–395, 2022.

[31] L. Kunwoo, Y. Umezu, K. Konno, and K. Kashima, "Observability Gramian for Bayesian Inference in Nonlinear Systems with Its Industrial Application," *IEEE Control Systems Letters*, vol. 7, pp. 871–876, 2023.

[32] A. J. Krener and K. Ide, "Measures of unobservability," *Proceedings of the IEEE Conference on Decision and Control*, pp. 6401–6406, 2009.

[33] A. J. Whalen, S. N. Brennan, T. D. Sauer, and S. J. Schiff, "Observability and controllability of nonlinear networks: The role of symmetry," *Physical Review X*, vol. 5, no. 1, pp. 1–40, 2015.

[34] Y. Kawano and J. M. Scherpen, "Empirical differential Gramians for nonlinear model reduction," *Automatica*, vol. 127, p. 109534, 2021.

[35] F. Bach, "Convex Analysis and Optimization with Submodular Functions: a Tutorial," 2010. [Online]. Available: http://arxiv.org/abs/1010.4207

[36] J. Bilmes, "Submodularity In Machine Learning and Artificial Intelligence," 2022.

[37] 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.

[38] W. Bai, W. S. Noble, and J. A. Bilmes, "Submodular maximization via gradient ascent: The case of deep submodular functions," *Advances in Neural Information Processing Systems*, vol. 2018-Decem, no. NeurIPS, pp. 7978–7988, 2018.

[39] L. Lovász, *Submodular functions and convexity*. Berlin, Heidelberg: Springer Berlin Heidelberg, 1983, pp. 235–257.

[40] A. Gupta and A. Roth, "Constrained Maximization of Non-Monotone Submodular Functions," pp. 1–10, 2009.