Sensor Selection and Attack for State Estimation of Linear Systems with Unknown Inputs

Nathaniel Woodford and Shreyas Sundaram

***Abstract*— An unknown input observer provides perfect asymptotic tracking of the state of a system affected by unknown inputs. Such an observer exists (possibly requiring a delay in estimation) if and only if the system satisfies a property known as strong detectability. In this paper, we consider the problem of selecting (at design-time) a minimum cost subset of sensors from a given set in order to make a given system strongly detectable. We show that this problem is NP-hard even when the system is stable. Furthermore, we show that it is not possible to approximate the minimum cost within a factor of $\log n$ in polynomial-time (unless $P = NP$). However, we show that if a given system (with a selected set of sensors) is already strongly detectable, finding the smallest set of additional sensors to install in order to obtain a zero-delay observer can be done in polynomial time. Finally, we consider the problem of attacking a set of deployed sensors in order to remove the property of strong detectability. We show that finding the smallest number of sensors to remove is NP-hard.**

I. Introduction

There is an increasing need to design controllers and estimators for large-scale systems in a variety of application domains, including computational biology, system of systems, intelligent traffic systems, communication networks, and power grids [1]–[4]. The states of such systems can be (partially) measured by sensors deployed at various locations. However, there are many instances in which it would be difficult or impractical to measure all the states of the system. This could be due to initial implementation cost or runtime energy cost of the sensors [5]. Therefore, a key challenge is to find a subset of sensors with minimum cost to deploy on the system in order to achieve certain performance objectives.

The problem of determining the minimal cost selection of sensors has been studied extensively in recent years. Existing approaches can be broadly separated into dynamically switching (or scheduling) between different sensors at runtime (e.g., [6]–[9]), and choosing sensors at design time (e.g., [10]–[17]). For instance, [10] considered the problem of selecting the smallest number of sensors to make a system observable, and showed that this problem is NP-hard to approximate within a factor of $\log n$. In the context of sensor selection for Kalman filtering, the papers [13], [14] showed that selecting a set of sensors (within a budget constraint) to minimize the trace of the steady state mean square estimation error (MSEE) is NP-hard, and furthermore, the minimum MSEE cannot be approximated within any constant factor in polynomial-time (unless $P = NP$). Similarly, [15] sought to minimize the number of sensors to achieve a certain estimation error, and to minimize the estimation error with a given number of sensors. The paper [16] studied minimal actuator placement for structural controllability, and [17] took a geometric approach to optimal sensor design.

In this paper, we consider the problem of (design-time) sensor selection for linear time-invariant systems that are affected by unknown (and arbitrary) inputs. Such inputs can be used to represent faults, disturbances, model reduction errors, or malicious attacks [18]–[20]. For instance, in large-scale critical infrastructure and industrial plants, cyber-attacks can be injected at various points in the system, and the characteristics of those attacks may not be known a priori; such attacks can thus be modeled as unknown inputs [19]–[21]. In this case, the system operator's task is to place sensors on the system in order to estimate the state despite the attacks injected by the adversary. As another example, consider a diffusive process such as a gas spreading over a given area [22], or temperature dynamics across a Multi-Processor-System-on-Chip [23]. These diffusive dynamics are driven by source injections at various locations, whose characteristics may not be known. In such cases, a limited number of sensors must be carefully deployed to estimate the gas concentrations or temperatures at all points in the space, despite lack of knowledge of the injected quantities.

In order to obtain perfect (asymptotic) estimation of the state of systems driven by unknown inputs (such as those described above), one must construct an *unknown input observer* which monitors the output of the system (provided by the deployed sensors) and maintains an estimate of the state (possibly with some delay) [24]. Such observers also find applications in fault-detection and robust estimation [25], [26]. For an unknown input observer to exist, the system must be *strongly detectable* (i.e., all invariant zeros of the system must be stable) [27], [28]. As a necessary condition for strong detectability is detectability, and it was shown in [10] that it is NP-hard to determine the minimum set of sensors to make a system detectable, the problem that we consider in this paper is trivially NP-hard as well. However, the fundamental question that motivates this paper is the following: is the NP-hardness of the sensor selection problem for strong detectability solely due to the need to obtain detectability? In other words, **do the unknown inputs contribute to the computational complexity of the problem?**

We answer this latter question in the affirmative by showing that it is NP-hard to find a minimum cost selection of sensors to make a given system strongly detectable,

The authors are with the School of Electrical and Computer Engineering at Purdue University. This research was supported by NSF grant CMMI-1635014 and by the Purdue Military Research Initiative. Email: {nwoodfor,sundara2}@purdue.edu

even when the system is stable. In particular, by restricting our attention to stable systems, we ensure that all sensor selections cause the system to be detectable, and thereby eliminate the complexity of choosing sensors to satisfy that property. Our proof of NP-hardness relies on carefully constructed instances of stable LTI systems affected by unknown inputs, along with sets of available sensors. Additionally, we show the stronger result that the minimum cost cannot be approximated within a factor of $\log n$. This inapproximability result mirrors the corresponding result for minimal sensor selection for observability provided in [10], but again arises from the need to handle the unknown inputs (as opposed to ensuring detectability as in [10]). However, we show that once a set of sensors is selected to make the system strongly detectable, the problem of finding the lowest cost set of additional sensors to obtain a zero-delay unknown input observer can be solved in polynomial time.

After establishing the above complexity results for the sensor selection problem, we turn our attention to the problem of *attacking* a deployed set of sensors in order to *remove* the property of strong detectability from the system. Specifically, we consider a scenario where an attacker can remove a given number of deployed sensors in an attempt to cause the remaining system to not be strongly detectable. We prove that it is NP-hard for the attacker to find the minimum number of sensors to remove to achieve this.

Throughout the paper, the set of real numbers and integers are denoted as $\mathbb{R}$ and $\mathbb{Z}$ respectively. We denote restrictions of those sets via subscripts (e.g., $\mathbb{R}_{\geq 0}$ denotes all nonnegative real numbers). Matrices are denoted in bold (e.g., $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$). The identity matrix of dimension $r \times r$ is denoted $\mathbf{I}_r$ and the zero matrix is denoted as $\mathbf{0}$ (with subscripts to denote the dimensions, as needed). The notation $\mathbf{A}(i,j)$ indicates the i^{th} row and j^{th} column of the matrix $\mathbf{A}$. We denote the transpose of a matrix $\mathbf{A}$ by $\mathbf{A}'$. The notation $\mathrm{diag}()$ indicates a diagonal matrix with the values in the parentheses along the diagonal. A binary indicator vector μ is a vector where each element is either a 1 or a 0. The complement of an indicator vector is denoted μ^c, where each 1 in μ becomes a 0 and vice versa. All vectors are column vectors, unless otherwise noted.

II. Background

Consider the linear time invariant system:

$$\mathbf{x}[t+1] = \mathbf{A}\mathbf{x}[t] + \mathbf{B}\mathbf{u}[t] \tag{1}$$

$$\mathbf{y}[t] = \mathbf{C}\mathbf{x}[t], \tag{2}$$

where $t \in \mathbb{Z}_{\geq 0}$ is the discrete-time index, $\mathbf{x}[t] \in \mathbb{R}^n$ is the state vector, $\mathbf{u}[t] \in \mathbb{R}^m$ is the unknown input vector, $\mathbf{y}[t] \in \mathbb{R}^p$ is the output vector, $\mathbf{A} \in \mathbb{R}^{n\times n}$ is the system dynamics matrix, $\mathbf{B} \in \mathbb{R}^{n\times m}$ is the input matrix, and $\mathbf{C} \in \mathbb{R}^{p\times n}$ is the output matrix. We assume without loss of generality throughout that $\mathbf{B}$ has full column rank.

Theorem 1 ([27], [28]): An unknown input observer (UIO) exists for system (1)-(2) if and only if

$$\mathrm{rank}\begin{bmatrix}\mathbf{A} - z_0\mathbf{I_n} & \mathbf{B}\\ \mathbf{C} & \mathbf{0}\end{bmatrix} = n + m, \quad \forall z_0 \in \mathbb{C}, |z_0| \geq 1. \tag{3}$$

Furthermore, if this condition is satisfied, the observer will estimate the state with a delay of at most $n-1$ time-steps. A zero-delay observer exists if and only if condition (3) holds, and in addition,

$$\mathrm{rank}(\mathbf{CB}) = \mathrm{rank}(\mathbf{B}). \tag{4}$$

□

We will refer to condition (3) as the **strong detectability condition**, and to (4) as the **matching condition**. A complex number z_0 reducing the rank of (3) below $n+m$ is said to be an **invariant zero** of the system.

III. The Strong Detectability Sensor Selection Problem (SDSS)

A. Problem Formulation

Consider again system (1), and suppose that there are no sensors deployed on the system (i.e., the output equation (2) is not initially given). Instead suppose that we have a set $\mathcal{S} = \{S_1, \ldots, S_p\}$ of available sensors and a cost vector $\mathbf{b} \in \mathbb{R}^p_{\geq 0}$ assigning a nonnegative cost to each sensor. In other words, the i^{th} element of $\mathbf{b}$ denotes the cost of sensor S_i, for each $1 \leq i \leq p$.

Each sensor $S_i \in \mathcal{S}$ provides a scalar measurement of the state given by

$$\mathbf{y}_i[t] = \mathbf{C}_i\mathbf{x}[t], \tag{5}$$

for a row vector $\mathbf{C}_i$. Let $\mathbf{C} = \begin{bmatrix}\mathbf{C}_1' & \mathbf{C}_2' & \cdots & \mathbf{C}_p'\end{bmatrix}'$. Given an indicator vector $\mu \in \{0,1\}^p$, we denote $\mathbf{C}(\mu)$ to be the submatrix of $\mathbf{C}$ consisting of the rows corresponding to the sensors indicated by μ.

We consider the following problem.

Problem 1 (Strong Detectability Sensor Selection (SDSS)): Suppose we are given the system matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$, the input matrix $\mathbf{B} \in \mathbb{R}^{n\times m}$, a set of p available sensors $\mathcal{S}$ whose measurements are given by the rows of matrix $\mathbf{C} \in \mathbb{R}^{p\times n}$, and a cost vector $\mathbf{b} \in \mathbb{R}^p_{\geq 0}$. The Strong Detectability Sensor Selection Problem (SDSS) is to solve

$$\min_{\mu\in\{0,1\}^p} \mathbf{b}'\mu$$

$$s.t. \quad \mathrm{rank}\begin{bmatrix}\mathbf{A} - z_0\mathbf{I_n} & \mathbf{B}\\ \mathbf{C}(\mu) & \mathbf{0}\end{bmatrix} = n+m, \quad \forall z_0 \in \mathbb{C},\ |z_0| \geq 1.$$

□

B. Complexity of SDSS

We start by showing the SDSS problem is NP-hard by providing a reduction from *Set Cover*, stated below.

Problem 2 (Set Cover): Consider a tuple $(\mathcal{U}, \mathcal{H}, k)$, where $\mathcal{U}$ is a finite set of r elements, $\mathcal{H}$ is a collection of sets $\{H_1, H_2, \ldots, H_q\}$ such that $H_i \subset \mathcal{U}$ for all $i \in \{1, 2, \ldots, q\}$, and k is a nonnegative integer.

Question: Do at most k sets from $\mathcal{H}$ exist whose union is equal to $\mathcal{U}$? □

Set Cover is NP-hard [29]. We will now provide a reduction from Set Cover to SDSS, prove certain useful properties of the created instance of SDSS, and consequently prove that SDSS is NP-hard.

1) Polynomial-Time Reduction from Set Cover to SDSS: Given an instance of Set Cover (with r elements in the set $\mathcal{U}$, and a collection $\mathcal{H}$ containing q subsets of $\mathcal{U}$), we will create an instance of SDSS as follows. Define the matrices

$$\begin{aligned} \mathbf{A} &= \begin{bmatrix} \mathbf{0}_{r\times r} & \mathbf{A_1} \\ \mathbf{0}_{r\times r} & \mathbf{0}_{r\times r} \end{bmatrix}, \ \mathbf{A_1} = \mathrm{diag}(1,2,\ldots,r), \\ \mathbf{B} &= \begin{bmatrix} \mathbf{I}_r \\ -\mathbf{I}_r \end{bmatrix}. \end{aligned} \tag{6}$$

Next, we define a $q \times r$ matrix $\mathbf{S}$ to encode the set $\mathcal{H}$. Each column of $\mathbf{S}$ corresponds to an element in $\mathcal{U}$, and each row encodes one of the sets H_i in the collection $\mathcal{H}$. Specifically, element (i, j) of $\mathbf{S}$ is equal to 1 if set $H_i \in \mathcal{H}$ contains the element $j \in \mathcal{U}$, and zero otherwise.

Now we define the set of sensors for SDSS. Specifically, we create $p = r + q$ sensors to choose from. Each sensor i's measurement matrix $\mathbf{C}_i$ consists of a single row; the collection of the measurement matrices for all sensors is given by

$$\mathbf{C} = \begin{bmatrix} \mathbf{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{S} \end{bmatrix}. \tag{7}$$

Each sensor's cost is defined by an element of the column vector $\mathbf{b}$. In this instance, we set the first r elements of $\mathbf{b}$ to '0', and the remaining elements to '1'.

2) Properties of the Created Instance: Consider the set of available sensors in the created instance, given by the rows of the matrix $\mathbf{C}$ in (7). Since the first r sensors all have zero cost (as specified in $\mathbf{b}$), they can always be included in any sensor selection without increasing the total cost. Thus, the indicator vector μ for the sensor selection will be assumed to have a '1' in each of its first r elements. The matrix $\mathbf{C}(\mu)$ containing the rows of all sensors selected by μ is given by

$$\mathbf{C}(\mu) = \begin{bmatrix} \mathbf{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{S}(\mu) \end{bmatrix}, \tag{8}$$

where $\mathbf{S}(\mu)$ is the matrix containing those rows of $\mathbf{S}$ that are included in the sensor selection μ.

Lemma 1: Consider a sensor selection μ and the corresponding tuple $(\mathbf{A}, \mathbf{B}, \mathbf{C}(\mu))$, where $\mathbf{A}$ and $\mathbf{B}$ are given by (6) and $\mathbf{C}(\mu)$ is given by (8). All invariant zeros of this tuple (if they exist) are unstable. □

Proof: Suppose z_0 is an invariant zero of the tuple $(\mathbf{A}, \mathbf{B}, \mathbf{C}(\mu))$, i.e., there exists a nonzero vector $\begin{bmatrix} \mathbf{X}_0' & \mathbf{U}_0' \end{bmatrix}'$ such that

$$\begin{bmatrix} \mathbf{A} - z_0\mathbf{I}_{2r} & \mathbf{B} \\ \mathbf{C}(\mu) & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{X_0} \\ \mathbf{U_0} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}. \tag{9}$$

By substituting the expressions for $\mathbf{A}, \mathbf{B}$ and $\mathbf{C}(\mu)$ from (6) and (8) into equation (9) and partitioning $\mathbf{X}_0$ as $\begin{bmatrix} \mathbf{X_1}' & \mathbf{X_2}' \end{bmatrix}'$, where $\mathbf{X}_1$ and $\mathbf{X}_2$ each have r elements, the expression (9) becomes

$$\left[\begin{array}{cc|c} -z_0\mathbf{I}_r & \mathbf{A_1} & \mathbf{I}_r \\ \mathbf{0} & -z_0\mathbf{I}_r & -\mathbf{I}_r \\ \hline \mathbf{I}_r & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{S}(\mu) & \mathbf{0} \end{array}\right] \begin{bmatrix} \mathbf{X_1} \\ \mathbf{X_2} \\ \mathbf{U_0} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}. \tag{10}$$

The above expression shows that $\mathbf{X_1} = 0$ and $\mathbf{U_0} = -z_0\mathbf{X_2}$. The first block row in (10) then leads to:

$$\mathbf{A_1}\mathbf{X_2} = z_0\mathbf{X_2}. \tag{11}$$

This implies that either $\mathbf{X}_2$ is the zero vector, or that z_0 and $\mathbf{X_2}$ are an eigenvalue and corresponding eigenvector of $\mathbf{A_1}$. The former case is impossible, since then $\mathbf{U}_0$ would also be zero in (10), contradicting the fact that all three of $\mathbf{X}_1$, $\mathbf{X}_2$ and $\mathbf{U}_0$ cannot be zero. Thus, z_0 and $\mathbf{X}_2$ must be an eigenvalue and eigenvector of $\mathbf{A}_1$ respectively. Since all eigenvalues of $\mathbf{A}_1$ in (6) are unstable, the claim follows. ■

Lemma 2: The tuple $(\mathbf{A}, \mathbf{B}, \mathbf{C}(\mu))$ has no invariant zeros if and only if all columns of $\mathbf{S}(\mu)$ are nonzero. □

Proof: Suppose all columns of $\mathbf{S}(\mu)$ are nonzero, and assume by way of contradiction that there exists an invariant zero z_0. Thus, there exists a nonzero vector $\begin{bmatrix} \mathbf{X}_1' & \mathbf{X}_2' & \mathbf{U}_0' \end{bmatrix}'$ satisfying (10). Furthermore, from the proof of Lemma 1, we know that $\mathbf{X}_2$ is an eigenvector of $\mathbf{A}_1$. Since each eigenvector of $\mathbf{A}_1$ has exactly one nonzero element (by design, from (6)), the quantity $\mathbf{S}(\mu)\mathbf{X}_2$ will be a scaled version of a column of $\mathbf{S}(\mu)$. However, if all columns of $\mathbf{S}(\mu)$ are nonzero, then this contradicts the last row of (9) and thus there cannot be any invariant zeros of the tuple $(\mathbf{A}, \mathbf{B}, \mathbf{C}(\mu))$.

Conversely, suppose there is a column of $\mathbf{S}(\mu)$ that is zero. Select $\mathbf{X}_2$ to be the indicator vector with a '1' in the position corresponding to that column, and zeros everywhere else. Set z_0 to be the eigenvalue of $\mathbf{A}_1$ corresponding to the eigenvector $\mathbf{X}_2$, $\mathbf{X}_1 = 0$, and $\mathbf{U}_0 = -z_0\mathbf{X}_2$. We see that this choice of z_0, $\mathbf{X}_1$, $\mathbf{X}_2$, and $\mathbf{U}_0$ satisfy (9). Thus, the tuple $(\mathbf{A}, \mathbf{B}, \mathbf{C}(\mu))$ will have an unstable invariant zero. ■

3) NP-hardness of SDSS: Using the reduction from Set Cover given by the system (6), the set of sensors (7) and cost vector $\mathbf{b}$, along with the properties of such instances given above, we obtain the following result.

Theorem 2: Given an instance of Set Cover and the associated instance of SDSS (given by (6), (7) and the cost vector $\mathbf{b}$), there exists a sensor selection of cost k that makes the system strongly detectable if and only if a set cover of size k or less exists. Thus, SDSS is NP-hard. □

Proof: Suppose there exists a set cover of size k or less. Let μ be the sensor selection vector that selects the first r sensors from (7) and the k sensors from the bottom q rows of $\mathbf{b}$ to correspond to the elements in the set cover instance. By the definition of the costs in $\mathbf{b}$, this selection has total cost k. Since each row of $\mathbf{S}$ encodes a different subset in the given instance of Set Cover, we see that $\mathbf{S}(\mu)$ will have no empty columns. From Lemma 2, if there are no empty columns in $\mathbf{S}(\mu)$ the tuple $(\mathbf{A}, \mathbf{B}, \mathbf{C}(\mu))$ will have no invariant zeros, and thus will be strongly detectable.

Now suppose that there is no set cover of size k. Then, for any sensor selection μ of cost k or less, there will be at least one column of $\mathbf{S}(\mu)$ that is zero. From Lemmas 1 and 2, we see that the tuple $(\mathbf{A}, \mathbf{B}, \mathbf{C}(\mu))$ will have an unstable invariant zero, and thus will not be strongly detectable.

Thus, we see that given any instance of Set Cover, we can create an instance of SDSS in polynomial-time, and solve the

Set Cover instance by solving the sensor selection instance. Since Set Cover is NP-hard, SDSS is as well. ■

When dealing with NP-hard problems, it is of interest to find polynomial-time *approximation algorithms* which provide solutions that are within a certain factor of the optimal. The following result provides a bound on the ability to approximate the minimum sensor cost in polynomial time.

Corollary 1: For all $\epsilon > 0$, SDSS cannot be approximated within a factor of $(1-\epsilon)\log n$ where n is the number of states, unless $P = NP$. □

Proof: By contradiction, suppose there is some $\epsilon > 0$ and an approximation algorithm for SDSS that always finds a set of sensors within a factor of $(1-\epsilon)\log n$ of the minimum cost. By running this algorithm on the constructed instance of SDSS given by (6), (7) and $\mathbf{b}$, we would obtain a set of sensors that provide strong detectability with a cost $\mathcal{B}$ that is within a factor $(1-\epsilon)\log n$ of the optimal cost. However the optimal cost is precisely equal to the smallest size of a set cover (by construction), and since the set of sensors yielded by the algorithm must be a set cover (by Lemma 2), we see that the algorithm would yield an approximation to Set Cover as well. Since Set Cover cannot be approximated within a factor of $(1-\epsilon)\log n$ of the optimal solution for any $\epsilon > 0$ (unless $P = NP$) [30], the result follows. ■

C. Complexity of Satisfying the Matching Condition

As the SDSS problem is NP-hard (as shown in Theorem 2), it is also NP-hard to find a minimum cost selection of sensors in order to construct a UIO (by Theorem 1). However, suppose that we consider a system that already has a set of sensors deployed which make the system strongly detectable, but that the matching condition (4) is not satisfied (so that a zero-delay UIO cannot be created). Suppose that we wish to deploy additional sensors (from a given set) of lowest cost in order to obtain a zero-delay UIO. This requires that the total set of deployed sensors satisfy the matching condition (4). In this section, we show that when each sensor provides a scalar measurement of the state (i.e., $\mathbf{C}_i$ in (5) is a row vector), a minimum cost set of sensors to satisfy the matching condition can be found in polynomial time. We will use the following result.

Lemma 3 ([31]): Consider a set $\mathcal{V} = \{v_1, v_2, \ldots, v_p\}$ consisting of p vectors, and a weight $\mathbf{w}_i \in \mathbb{R}_{\geq 0}$ for each vector $v_i \in \mathcal{V}$. The problem of finding the lowest cost maximal linearly independent subset[1] of vectors can be solved in polynomial time via a greedy algorithm. □

We will start by considering the general problem of selecting a subset of sensors of lowest cost in order to satisfy only the matching condition (i.e., without considering the strong detectability condition).

Theorem 3: Consider a set of sensors $\mathcal{S} = \{S_1, S_2, \ldots, S_p\}$, where each sensor provides a scalar measurement of the form (5). Let the vector $\mathbf{b} \in \mathbb{R}^p_{\geq 0}$ contain the cost of each sensor. Let $\mathbf{C}$ be the matrix consisting of all of the individual sensor matrices. Then, the sensor selection vector $\mu \in \{0,1\}^p$ of lowest cost satisfying the matching condition $\mathrm{rank}(\mathbf{C}(\mu)\mathbf{B}) = \mathrm{rank}(\mathbf{B})$ (if such a selection exists) can be found in polynomial time via a greedy algorithm. □

Proof: Define the matrix $\mathbf{J} = \mathbf{C}\mathbf{B}$, where the i^{th} row of $\mathbf{J}$ is given by $\mathbf{C}_i\mathbf{B}$. Thus, define the cost of the i^{th} row of $\mathbf{J}$ to be $\mathbf{b}_i$, i.e., the cost of the corresponding sensor S_i.

For any sensor selection vector μ, define $\mathbf{J}(\mu) = \mathbf{C}(\mu)\mathbf{B}$, which implies $\mathrm{rank}(\mathbf{J}(\mu)) = \mathrm{rank}(\mathbf{C}(\mu)\mathbf{B})$. Thus, finding a set of rows of $\mathbf{C}$ of minimum cost such that $\mathrm{rank}(\mathbf{C}(\mu)\mathbf{B}) = \mathrm{rank}(\mathbf{B})$ (if it exists) is equivalent to finding a maximal linearly independent set of rows of $\mathbf{J}$ of lowest cost. By Lemma 3 this can be done in polynomial-time via a greedy algorithm. Thus, the lowest cost set of sensors to satisfy the matching condition can be found in polynomial time. ■

Algorithm 1 is an example of a greedy algorithm that takes matrices $\mathbf{B}$ and $\mathbf{C}$, along with a cost for each row of $\mathbf{C}$, and returns a lowest cost sensor selection μ satisfying $\mathrm{rank}(\mathbf{C}(\mu)\mathbf{B}) = \mathrm{rank}(\mathbf{B})$ (if such a selection exists).

Algorithm 1 Greedy Selection for Matching Condition

Notation:$\mu \cup \{i\}$ indicates setting the i^{th} element of μ to 1.
Input: Sensor matrix $\mathbf{C} \in \mathbb{R}^{p\times n}$, input matrix $\mathbf{B} \in \mathbb{R}^{n\times m}$, and a vector $\mathbf{b} \in \mathbb{R}^p_{\geq 0}$ indicating the cost of each row of $\mathbf{C}$.
Output: A sensor selection vector $\mu \in \{0,1\}^p$ that minimizes $\mathbf{b}'\mu$ while ensuring $\mathrm{rank}(\mathbf{C}(\mu)\mathbf{B}) = \mathrm{rank}(\mathbf{B})$.

```
Sort the rows of C to be in nondecreasing order by cost.
Initialize μ to the zero vector and i = 1
while rank(C(μ)B) < rank(B) do
    if rank(C(μ ∪ {i})B) > rank(C(μ)B) then
        μ = μ ∪ {i}
    end if
    i = i + 1
end while
return μ
```

Note that the sensor costs are allowed to be arbitrary nonnegative values in the above result. Thus, this captures (as a special case) the scenario where we already have a set of sensors installed on the system (e.g., to provide strong detectability), and we only need to select an additional set of sensors in order to satisfy the matching condition. Specifically, by setting the cost of all sensors that are already installed to be zero, the algorithm is guaranteed to select from the installed set of sensors first (as it checks the sensors in nondecreasing order of cost), and then to select the lowest cost subset of additional sensors to install. This is encapsulated in the following corollary.

Corollary 2: Consider a linear system of the form (1). Suppose we are given a set of p sensors $\mathcal{S}$, where each sensor in the set provides a scalar measurement of the form (5). Let $\mathbf{C}$ be the matrix whose rows contain the measurement matrices of the sensors, and let $\mathbf{b} \in \mathbb{R}^p_{\geq 0}$ indicate the cost of each sensor. Suppose that some subset of the sensors in $\mathcal{S}$ is already installed on the system. Then, the lowest

[1]A maximal linearly independent subset of vectors is a linearly independent subset of $\mathcal{V}$ such that no additional vectors from $\mathcal{V}$ can be added to the subset without violating linear independence.

cost set of additional sensors to install so that the set of all installed sensors satisfies the matching condition can be found in polynomial time. □

IV. The Strong Detectability Sensor Attack Problem (SDSA)

Having characterized the complexity of the sensor selection problem, we now turn our attention to the problem of *attacking* a set of deployed sensors in order to *remove* the property of strong detectability. We formulate this problem next, and then characterize its complexity.

A. Problem Formulation

Consider again system (1), and suppose that there are sensors deployed on the system (i.e., the output equation (2) is initially given) such that the system is strongly detectable. Instead of adding sensors to the system suppose one (i.e., an attacker) desires to remove sensors. The cost vector $\mathbf{v} \in \mathbb{R}^p_{\geq 0}$ assigns a nonnegative removal cost for each sensor. In other words, the i^{th} element of $\mathbf{v}$ denotes the cost of removing the i^{th} row from the matrix $\mathbf{C}$ for each $1 \leq i \leq p$.

As before, given an indicator vector $\mu \in \{0,1\}^p$, we denote $\mathbf{C}(\mu)$ to be the submatrix of $\mathbf{C}$ consisting of the rows corresponding to the sensors indicated by μ. Furthermore, the indicator vector $\mu^c \in \{0,1\}^p$ is the complement of μ (i.e. a '1' in μ is denoted as a '0' in μ^c and vice versa).

We consider the following problem.

Problem 3 (Strong Detectability Sensor Attack (SDSA)): Suppose we are given the system matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$, the input matrix $\mathbf{B} \in \mathbb{R}^{n\times m}$, the output matrix $\mathbf{C} \in \mathbb{R}^{p\times n}$, and a cost vector $\mathbf{v} \in \mathbb{R}^p_{\geq 0}$. The Strong Detectability Sensor Attack Problem (SDSA) is to solve

$$\min_{\mu\in\{0,1\}^p} \mathbf{v}'\mu$$
$$s.t. \quad \exists\, |z_0| \geq 1 \quad \text{with rank} \begin{bmatrix} \mathbf{A} - z_0\mathbf{I_n} & \mathbf{B} \\ \mathbf{C}(\mu^c) & \mathbf{0} \end{bmatrix} < n + m.$$

□

B. Complexity of SDSA

In this section, we will show that the SDSA problem is NP-hard. To do so, we will provide a reduction from the *MAX FLS$^=$* problem, stated below.

Problem 4 (MAX FLS$^=$): Consider a set of d homogeneous equations with f variables denoted by matrix $\mathbf{T} \in \mathbb{R}^{d\times f}$, and a nonnegative integer k.

Question: Is there a nonzero vector $\mathbf{x} \in \mathbb{R}^f$ such that at least k equalities in the equation $\mathbf{Tx} = \mathbf{0}$ are satisfied? □

The MAX FLS$^=$ problem is NP-hard [32]. We reduce MAX FLS$^=$ to SDSA, state a useful property of the created instance of SDSA, and finally prove that that the SDSA is NP-hard.

1) Polynomial-Time Reduction from MAX FLS$^=$ to SDSA: Given an instance of the MAX FLS$^=$ problem (with d equations and f variables denoted by matrix $\mathbf{T}$), we will create an instance of SDSA as follows. Define the matrices

$$\mathbf{A} = \mathbf{0}_{f\times f}, \;\; \mathbf{B} = \mathbf{I}_f, \;\; \text{and} \;\; \mathbf{C} = \mathbf{T}. \tag{12}$$

The cost vector $\mathbf{v}$ will consist of d elements, all equal to '1'.

2) Properties of the Created Instance:

Lemma 4: Given system (1)-(2) with $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ defined in (12), consider a sensor selection vector $\nu \in \{0,1\}^d$. The tuple $(\mathbf{A}, \mathbf{B}, \mathbf{C}(\nu))$ has at least one unstable invariant zero if and only if $\mathbf{C}(\nu)$ is not full column rank. □

Proof: Consider a sensor indicator vector $\nu \in \{0,1\}^d$ with associated matrix $\mathbf{C}(\nu)$. The tuple $(\mathbf{A}, \mathbf{B}, \mathbf{C}(\nu))$ will have an unstable invariant zero if and only if there is a complex number z_0 with $|z_0| \geq 1$, and a nonzero vector $\begin{bmatrix}\mathbf{X}_0' & \mathbf{U}_0'\end{bmatrix}'$ satisfying

$$\begin{bmatrix} \mathbf{A} - z_0\mathbf{I}_f & \mathbf{B} \\ \mathbf{C}(\nu) & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{X_0} \\ \mathbf{U_0} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}. \tag{13}$$

Substituting (12) into (13) we obtain

$$z_0\mathbf{X_0} = \mathbf{U}_0 \tag{14}$$
$$\mathbf{C}(\nu)\mathbf{X_0} = \mathbf{0}. \tag{15}$$

If $\mathbf{C}(\nu)$ is not full column rank, there will exist a $\mathbf{X_0}$ vector in the null space of $\mathbf{C}(\nu)$, thereby satisfying (15). This $\mathbf{X_0}$ vector can be paired with any unstable z_0 value to form the vector $\mathbf{U}_0$ in (14).

On the other hand, if $\mathbf{C}(\nu)$ is full column rank, then the only solution to (15) is $\mathbf{X}_0 = \mathbf{0}$, and thus, from (14), $\mathbf{U}_0 = \mathbf{0}$. In this case, the matrix pencil in (13) has no nontrivial nullspace for any z_0, and thus has no invariant zeros.

Therefore, the tuple $(\mathbf{A}, \mathbf{B}, \mathbf{C}(\nu))$ has an unstable invariant zero if and only if $\mathbf{C}(\nu)$ is not full column rank. ■

3) NP-hardness of SDSA: Using the reduction from MAX FLS$^=$ given by the system (12), and cost vector $\mathbf{v}$ consisting of all 1's, along with the property of such instances given above, we obtain the following result.

Theorem 4: Given an instance of MAX FLS$^=$ (with a $d \times f$ matrix $\mathbf{T}$ and integer k) and the associated instance of SDSA (given by (12)), it is possible to remove $d - k$ or fewer sensors from $\mathbf{C}$ to make the resulting system no longer strongly detectable if and only if the answer to the instance of MAX FLS$^=$ is "yes". Thus, SDSA is NP-hard. □

Proof: Suppose the answer to the instance of MAX FLS$^=$ is "yes" (i.e., there is a nonzero vector $\mathbf{x}$ satisfying at least k of the equalities in the equation $\mathbf{Tx} = 0$). Let μ^c be the indicator vector that selects the k satisfied equations from $\mathbf{Tx} = \mathbf{0}$. Consequently since $\mathbf{C} = \mathbf{T}$ in the created instance of SDSA (given by (12)), there must exist some nonzero vector $\mathbf{x}$ such that $\mathbf{C}(\mu^c)\mathbf{x} = \mathbf{0}$. Thus $\mathbf{x}$ is in the null space of $\mathbf{C}(\mu^c)$ and $\mathbf{C}(\mu^c)$ is not full rank. By Lemma 4, the tuple $(\mathbf{A}, \mathbf{B}, \mathbf{C}(\mu^c))$ is not strongly detectable if $\mathbf{C}(\mu^c)$ is not full rank. Thus, the conjugate μ of μ^c represents the sensors that once removed cause the system to lose strong detectability. Since at most $d - k$ sensors must be removed and each sensor has a cost of '1', there is a solution to the created SDSA instance that has cost at most $d - k$.

Now suppose that the answer to MAX FLS$^=$ is "no" (i.e., there are fewer than k equalities in $\mathbf{Tx} = \mathbf{0}$ that can be simultaneously satisfied). Therefore there must be more than $d - k$ sensors that must be removed for $\mathbf{C}$ to lose rank. Therefore, by only removing $d - k$ sensors the

system will remain strongly detectable. Thus, we see that given any instance of the MAX FLS$^=$ problem, we can create an instance of SDSA in polynomial-time, and solve the MAX FLS$^=$ instance by solving the sensor attack instance. Since MAX FLS$^=$ is NP-hard, SDSA is NP-hard. ■

Additionally this result indicates that it is NP-hard to minimally break the matching condition (4). This condition is satisfied if $\mathrm{rank}(\mathbf{C}(\mu^c)\mathbf{B}) = \mathrm{rank}(\mathbf{B})$. In the instance where $\mathbf{B}$ is full rank the only way for condition (4) to hold is for the $\mathbf{C}(\mu^c)$ matrix to be full column rank as well. Therefore, once again, the task is to remove the minimal amount of sensors from $\mathbf{C}$ such that it loses full column rank. Thus, as a positive result, it is NP-hard for an attacker to optimally select sensors to remove to break the matching condition (in contrast to the problem of selecting sensors to satisfy the matching condition, as shown in Corollary 2).

V. Summary

In this paper, we showed that it is NP-hard to select a set of sensors of minimum cost in order to make a system strongly detectable. Our proof shows that this result holds even for stable systems, and thus the computational complexity arises from the effects of the unknown inputs in the system, as opposed to the need to ensure system detectability. We also showed that it is not possible to approximate the minimum cost within a factor that is logarithmic in the size of the problem. However, we showed that if a set of sensors has already been chosen to make a system strongly detectable, finding an additional set of sensors of minimum cost in order to obtain zero-delay estimation can be done in polynomial time. Finally, we considered the problem of attacking a given strongly detectable system by removing a set of sensors to remove the strong detectability property. We showed that this problem is also NP-hard.

There are a variety of avenues for future research, including determining instances of the sensor selection and attack problems where optimal (or near-optimal) solutions can be found in polynomial time.

VI. Acknowledgements

The authors thank Lintao Ye for helpful discussions.

References

[1] F. Menolascina, V. Siciliano, and D. Di Bernardo, "Engineering and control of biological systems: a new way to tackle complex diseases," *FEBS letters*, vol. 586, no. 15, pp. 2122–2128, 2012.

[2] L. Figueiredo, I. Jesus, J. T. Machado, J. R. Ferreira, and J. M. De Carvalho, "Towards the development of intelligent transportation systems," in *Intelligent Transportation Systems, Proceedings.*, 2001, pp. 1206–1211.

[3] A. Chakrabortty and M. D. Ilić, *Control and optimization methods for electric smart grids*. Springer, 2011, vol. 3.

[4] G. Bumiller, L. Lampe, and H. Hrasnica, "Power line communication networks for large-scale control and automation systems," *IEEE Communications Magazine*, vol. 48, no. 4, 2010.

[5] 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.

[6] V. Gupta, T. H. Chung, B. Hassibi, and R. M. Murray, "On a stochastic sensor selection algorithm with applications in sensor scheduling and sensor coverage," *Automatica*, vol. 42, no. 2, pp. 251–260, 2006.

[7] S. T. Jawaid and S. L. Smith, "Submodularity and greedy algorithms in sensor scheduling for linear dynamical systems," *Automatica*, vol. 61, pp. 282–288, 2015.

[8] M. P. Vitus, W. Zhang, A. Abate, J. Hu, and C. J. Tomlin, "On efficient sensor scheduling for linear dynamical systems," *Automatica*, vol. 48, no. 10, pp. 2482–2493, 2012.

[9] H. Rowaihy, S. Eswaran, M. Johnson, D. Verma, A. Bar-Noy, T. Brown, and T. La Porta, "A survey of sensor selection schemes in wireless sensor networks," in *Unattended Ground, Sea, and Air Sensor Technologies and Applications IX*, vol. 6562. International Society for Optics and Photonics, 2007, p. 65621A.

[10] A. Olshevsky, "Minimal controllability problems," *IEEE Transactions on Control of Network Systems*, vol. 1, no. 3, pp. 249–258, 2014.

[11] M. Van De Wal and B. De Jager, "A review of methods for input/output selection," *Automatica*, vol. 37, no. 4, pp. 487–510, 2001.

[12] X. Liu, B. M. Chen, and Z. Lin, "On the problem of general structural assignments of linear systems through sensor/actuator selection," *Automatica*, vol. 39, no. 2, pp. 233–241, 2003.

[13] 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.

[14] L. Ye, S. Roy, and S. Sundaram, "On the complexity and approximability of optimal sensor selection for Kalman filtering," in *American Control Conference*, 2018, pp. 5049–5054.

[15] V. Tzoumas, A. Jadbabaie, and G. J. Pappas, "Sensor placement for optimal Kalman filtering: Fundamental limits, submodularity, and algorithms," in *American Control Conference (ACC), 2016*. IEEE, 2016, pp. 191–196.

[16] S. Pequito, G. Ramos, S. Kar, A. P. Aguiar, and J. Ramos, "The robust minimal controllability problem," *Automatica*, vol. 82, pp. 261–268, 2017.

[17] M.-A. Belabbas, "Geometric methods for optimal sensor design," in *Proc. R. Soc. A*, vol. 472, no. 2185. The Royal Society, 2016, p. 20150312.

[18] R. J. Patton and J. Chen, "Robust model-based fault diagnosis for dynamic systems," 1999.

[19] S. Sundaram and C. N. Hadjicostis, "Distributed function calculation via linear iterative strategies in the presence of malicious agents," *IEEE Trans on Automatic Control*, vol. 56, no. 7, pp. 1495–1508, 2011.

[20] F. Pasqualetti, F. Dörfler, and F. Bullo, "Attack detection and identification in cyber-physical systems," *IEEE Transactions on Automatic Control*, vol. 58, no. 11, pp. 2715–2729, 2013.

[21] A. Teixeira, K. C. Sou, H. Sandberg, and K. H. Johansson, "Secure control systems: A quantitative risk management approach," *IEEE Control Systems*, vol. 35, no. 1, pp. 24–45, 2015.

[22] S. Roy and R. Dhal, "Situational awareness for dynamical network processes using incidental measurements," *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 2, pp. 304–316, 2015.

[23] S. Sharifi, D. Krishnaswamy, and T. Š. Rosing, "Prometheus: A proactive method for thermal management of heterogeneous MPSoCs," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 32, no. 7, pp. 1110–1123, 2013.

[24] Y. Guan and M. Saif, "A novel approach to the design of unknown input observers," *IEEE Transactions on Automatic Control*, vol. 36, no. 5, pp. 632–635, 1991.

[25] J. Chen, R. J. Patton, and H.-Y. Zhang, "Design of unknown input observers and robust fault detection filters," *International Journal of control*, vol. 63, no. 1, pp. 85–105, 1996.

[26] M. Saif and Y. Guan, "A new approach to robust fault detection and identification," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 29, no. 3, pp. 685–695, 1993.

[27] M. L. Hautus, "Strong detectability and observers," *Linear Algebra and its applications*, vol. 50, pp. 353–368, 1983.

[28] S. Sundaram and C. N. Hadjicostis, "Delayed observers for linear systems with unknown inputs," *IEEE Transactions on Automatic Control*, vol. 52, no. 2, pp. 334–339, 2007.

[29] U. Feige, "A threshold of ln n for approximating set cover," *Journal of the ACM (JACM)*, vol. 45, no. 4, pp. 634–652, 1998.

[30] I. Dinur and D. Steurer, "Analytical approach to parallel repetition," in *Proceedings of the forty-sixth annual ACM symposium on Theory of computing*. ACM, 2014, pp. 624–633.

[31] T. H. Cormen, *Introduction to algorithms*. MIT press, 2009.

[32] E. Amaldi and V. Kann, "The complexity and approximability of finding maximum feasible subsystems of linear relations," *Theoretical computer science*, vol. 147, no. 1-2, pp. 181–210, 1995.