2018 Annual American Control Conference (ACC)
June 27–29, 2018. Wisconsin Center, Milwaukee, USA

On the Complexity and Approximability of Optimal Sensor Selection for Kalman Filtering

Lintao Ye, Sandip Roy and Shreyas Sundaram

***Abstract*— Given a linear dynamical system, we consider the problem of selecting (at design-time) an optimal set of sensors (subject to certain budget constraints) to minimize the trace of the steady state error covariance matrix of the Kalman filter. Previous work has shown that this problem is NP-hard for certain classes of systems and sensor costs; in this paper, we show that the problem remains NP-hard even for the special case where the system is stable and all sensor costs are identical. Furthermore, we show the stronger result that there is no constant-factor (polynomial-time) approximation algorithm for this problem. This contrasts with other classes of sensor selection problems studied in the literature, which typically pursue constant-factor approximations by leveraging greedy algorithms and submodularity of the cost function. Here, we provide a specific example showing that greedy algorithms can perform arbitrarily poorly for the problem of design-time sensor selection for Kalman filtering.**

I. Introduction

Selecting an appropriate set of actuators or sensors in order to achieve certain performance requirements is an important problem in control system design (e.g., [1], [2], [3]). For instance, in the case of linear Gauss-Markov systems, researchers have studied techniques to select sensors dynamically (at run-time) or statically (at design-time) in order to minimize certain metrics of the error covariance of the corresponding Kalman filter. These are known as sensor scheduling problems (e.g., [4], [5], [6]) and design-time sensor selection problems (e.g., [7], [8], [9], [10]), respectively. These problems are NP-hard in general (e.g., [10]), and various approximation algorithms have been proposed to solve them. For example, the concept of submodularity [11] has been widely used to analyze the performance of greedy algorithms for sensor scheduling and selection (e.g., [12], [13], [6], [14]).

In this paper, we consider the design-time sensor selection problem for optimal filtering of discrete-time linear dynamical systems. We study the problem of choosing a subset of sensors (under given budget constraints) to optimize the steady state error covariance of the corresponding Kalman filter. We refer to this problem as the *Kalman filtering sensor selection (KFSS)* problem. We summarize some related work as follows.

In [7], the authors considered the design-time sensor selection problem of a sensor network for discrete-time linear dynamical systems, also known as dynamic data-reconciliation problems. Their objective was to minimize the estimation error subject to network defined mass-balance equations. In contrast, we consider the problem of minimizing the estimation error under a cardinality constraint on the chosen sensors and analyze the complexity of the problem.

In [9], the authors studied the design-time sensor selection problem for discrete-time linear time-varying systems over a finite time horizon. The objective is to minimize the number of chosen sensors while guaranteeing a certain level of performance (or alternatively, to minimize the estimation error with a cardinality constraint on the chosen sensors). In contrast, we focus on minimizing the steady state estimation error of the Kalman filter. The same problem was considered in [15] and [10]. In [15], the authors expressed the problem as a semidefinite program (SDP) without theoretical guarantees on the performance of the proposed algorithm. The paper [10] showed that the problem is NP-hard and the cost function is not submodular in general. Upper bounds (which are functions of system parameters) were provided on the performance of algorithms for the problem. Although [10] showed via simulations that greedy algorithms performed well for randomly generated systems, the question of whether such algorithms (or other polynomial-time algorithms) could provide constant-factor approximation ratios for the problem was left open.

Our contributions to this problem are as follows. First, we show that the KFSS problem is NP-hard *even* for the special case when the system is stable and all sensors have the same cost. This complements and strengthens the complexity results in [10]. Our second (and most significant) contribution is to show that *there is no constant factor approximation algorithm for this problem* (unless $P = NP$). This stands in stark contrast to other sensor selection problems studied in the literature, which leveraged submodularity of their associated cost functions to provide greedy algorithms with constant-factor approximation ratios [9]. Our inapproximability result above immediately implies that greedy algorithms cannot provide constant-factor guarantees for our problem. Our third contribution in this paper is to explicitly show how greedy algorithms can provide arbitrarily poor performance even for very small instances of the KFSS problem.

The rest of this paper is organized as follows. In Section II, we formulate the KFSS problem. In Section III, we analyze the complexity of the KFSS problem. In Section IV, we study a greedy algorithm for the KFSS problem and analyze its performance. In Section V, we conclude the paper.

This research was supported by NSF grant CMMI-1635014. Lintao Ye and Shreyas Sundaram are with the School of Electrical and Computer Engineering at Purdue University. Email: {ye159,sundara2}@purdue.edu. Sandip Roy is with the School of Electrical and Computer Engineering and Computer Science at Washington State University. Email: sroy@eecs.wsu.edu.

A. Notation and terminology

The set of natural numbers, integers, real numbers, rational numbers, and complex numbers are denoted as $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{R}$, $\mathbb{Q}$ and $\mathbb{C}$, respectively. For any $x \in \mathbb{R}$, denote $\lceil x \rceil$ as the least integer greater than or equal to x. For a matrix $P \in \mathbb{R}^{n\times n}$, let P^T be its transpose. Denote P_{ij} as the element in the ith row and jth column of P. The set of n by n positive-semidefinite matrices is denoted by $\mathbb{S}_+^n$. The identity matrix with dimension n is denoted as $I_{n\times n}$. For a vector v, denote v_i as the ith element of v and let $\text{supp}(v)$ be its support, where $\text{supp}(v) = \{i : v_i \neq 0\}$. Denote the Euclidean norm of v by $\|v\|_2$. Define $\mathbf{e}_i$ to be a row vector where the ith element is 1 and all the other elements are zero; the dimension of the vector can be inferred from the context. For a random variable ω, let $\mathbb{E}(\omega)$ be its expectation. For a set $\mathcal{A}$, let $|\mathcal{A}|$ be its cardinality.

II. Problem Formulation

Consider the discrete-time linear system

$$x[k+1] = Ax[k] + w[k], \tag{1}$$

where $x[k] \in \mathbb{R}^n$ is the system state, $w[k] \in \mathbb{R}^n$ is a zero-mean white Gaussian noise process with $\mathbb{E}[w[k](w[k])^T] = W$ for all $k \in \mathbb{N}$, and $A \in \mathbb{R}^{n\times n}$ is the system dynamics matrix.

Consider a set $\mathcal{Q}$ consisting of q sensors. Each sensor $i \in \mathcal{Q}$ provides a measurement of the system in the form

$$y_i[k] = C_i x[k] + v_i[k], \tag{2}$$

where $C_i \in \mathbb{R}^{s_i\times n}$ is the state measurement matrix for sensor i, and $v_i[k] \in \mathbb{R}^{s_i}$ is a zero-mean white Gaussian noise process. We further define $y[k] \triangleq \begin{bmatrix}(y_1[k])^T & \cdots & (y_q[k])^T\end{bmatrix}^T$, $C \triangleq \begin{bmatrix}C_1^T & \cdots & C_q^T\end{bmatrix}^T$ and $v[k] \triangleq \begin{bmatrix}(v_1[k])^T & \cdots & (v_q[k])^T\end{bmatrix}^T$. Thus, the output provided by all sensors together is given by

$$y[k] = Cx[k] + v[k], \tag{3}$$

where $C \in \mathbb{R}^{s\times n}$ and $s = \sum_{i=1}^q s_i$. We denote $\mathbb{E}[v[k](v[k])^T] = V$ and consider $\mathbb{E}[v[k](w[j])^T] = \mathbf{0}$, $\forall k, j \in \mathbb{N}$.

Consider that there are no sensors initially deployed on the system. Instead, the system designer must select a subset of sensors from the set $\mathcal{Q}$ to install. Each sensor $i \in \mathcal{Q}$ has a cost $b_i \in \mathbb{R}_{\geq 0}$; define the cost vector $b \triangleq \begin{bmatrix}b_1 & \cdots & b_q\end{bmatrix}^T$. The designer has a budget $B \in \mathbb{R}_{\geq 0}$ that can be spent on selecting sensors from $\mathcal{Q}$.

After a set of sensors is selected and installed, the Kalman filter is then applied to provide an optimal estimate of the states using the measurements from the installed sensors in the sense of minimizing the mean square estimation error (MSEE). We define a vector $\mu \in \{0,1\}^q$ as the indicator vector of the selected sensors, where $\mu_i = 1$ if and only if sensor $i \in \mathcal{Q}$ is installed. Denote $C(\mu)$ as the measurement matrix of the installed sensors indicated by μ, i.e., $C(\mu) \triangleq \begin{bmatrix}C_{i_1}^T & \cdots & C_{i_p}^T\end{bmatrix}^T$, where $\text{supp}(\mu) = \{i_1, \ldots, i_p\}$. Similarly, denote $V(\mu)$ as the measurement noise covariance matrix of the installed sensors, i.e., $V(\mu) = \mathbb{E}[\tilde{v}[k](\tilde{v}[k])^T]$, where $\tilde{v}[k] = \begin{bmatrix}(v_{i_1}[k])^T & \cdots & (v_{i_p}[k])^T\end{bmatrix}^T$. Let $\Sigma_{k|k-1}(\mu)$ and $\Sigma_{k|k}(\mu)$ denote the *a priori* error covariance matrix and the *a posteriori* error covariance matrix of the Kalman filter at time step k, respectively, when the sensors indicated by μ are installed. We will use the following result [16].

Lemma 1: Suppose the pair $(A, W^{\frac{1}{2}})$ is stabilizable. For a given indicator vector μ, both $\Sigma_{k|k-1}(\mu)$ and $\Sigma_{k|k}(\mu)$ will converge to finite limits $\Sigma(\mu)$ and $\Sigma^*(\mu)$, respectively, as $k \to \infty$ if and only if the pair $(A, C(\mu))$ is detectable. $\square$

The limit $\Sigma(\mu)$ satisfies the *discrete algebraic Riccati equation (DARE)* [16]:

$$\Sigma(\mu) = A\Sigma(\mu)A^T + W - \\ A\Sigma(\mu)C(\mu)^T\big(C(\mu)\Sigma(\mu)C(\mu)^T + V(\mu)\big)^{-1}C(\mu)\Sigma(\mu)A^T. \tag{4}$$

Applying the matrix inversion lemma [17], we can rewrite Eq. (4) as

$$\Sigma(\mu) = W + A(\Sigma^{-1}(\mu) + R(\mu))^{-1}A^T, \tag{5}$$

where $R(\mu) \triangleq C(\mu)^T V(\mu)^{-1} C(\mu)$ is the sensor information matrix corresponding to sensor selection indicated by μ. Note that the inverses in Eq. (4) and Eq. (5) are interpreted as pseudo-inverses if the arguments are not invertible. For the case when $V = \mathbf{0}$, we compute $\Sigma(\mu)$ via Eq. (4).

The limits $\Sigma(\mu)$ and $\Sigma^*(\mu)$ are coupled as [18]:

$$\Sigma(\mu) = A\Sigma^*(\mu)A^T + W. \tag{6}$$

For the case when the pair $(A, C(\mu))$ is not detectable, we define the limit $\Sigma(\mu) = +\infty$. The Kalman filter sensor selection (KFSS) problem is defined as follows.

Problem 1: (KFSS) Given a system dynamics matrix $A \in \mathbb{R}^{n\times n}$, a measurement matrix $C \in \mathbb{R}^{s\times n}$ containing all of the individual sensor measurement matrices, a system noise covariance matrix $W \in \mathbb{S}_+^n$, a sensor noise covariance matrix $V \in \mathbb{S}_+^s$, a cost vector $b \in \mathbb{R}_{\geq 0}^q$ and a budget $B \in \mathbb{R}_{\geq 0}$, the Kalman filtering sensor selection problem is to find the sensor selection μ, i.e., the indicator vector μ of the selected sensors, that solves

$$\begin{aligned} &\min_{\mu} \ \text{trace}(\Sigma(\mu)) \\ &s.t.\ b^T\mu \leq B \\ &\mu \in \{0,1\}^q \end{aligned}$$

where $\Sigma(\mu)$ is given by Eq. (4) if the pair $(A, C(\mu))$ is detectable, and $\Sigma(\mu) = +\infty$, otherwise.

III. Complexity Analysis

As mentioned in the Introduction, the KFSS problem was shown to be NP-hard in [10] for two classes of systems and sensor costs. First, when the A matrix is unstable and sensor costs are identical, [10] provided a reduction from the "minimal controllability" (or minimal detectability) problem considered in [2] to KFSS. Second, when the A matrix is stable, [10] showed that when the sensor costs can be arbitrary, the $0-1$ knapsack problem can be encoded as a special case of KFSS, thereby again showing NP-hardness of the latter problem.

In this section, we provide a stronger result and show that KFSS is NP-hard even for the special case where the A matrix is stable *and* all sensors have the same cost. We will consider throughout this paper the case when $C_i \in \mathbb{R}^{1\times n}$, $\forall i \in \{1,\ldots,q\}$, i.e., each sensor corresponds to one row of matrix C, and the sensor selection cost vector is $b = [1 \ \cdots \ 1]^T$, i.e., each sensor has cost equal to 1.

We will use the following Lemmas, whose proofs can be found in [21].

Lemma 2: Consider a discrete-time linear system as defined in (1) and (3). Suppose the system dynamics matrix is of the form $A = \text{diag}(\lambda_1,\ldots,\lambda_n)$ with $0 \le |\lambda_i| < 1$, $\forall i \in \{1,\ldots,n\}$, the system noise covariance matrix W is diagonal, and the sensor noise covariance matrix is $V = \mathbf{0}$. Then, the following holds for all sensor selections μ.

(a) For all $i \in \{1,\ldots,n\}$, $(\Sigma(\mu))_{ii}$ satisfies

$$W_{ii} \le (\Sigma(\mu))_{ii} \le \frac{W_{ii}}{1-\lambda_i^2}. \tag{7}$$

(b) If $\exists i \in \{1,\ldots,n\}$ s.t. $W_{ii} = 0$, then $(\Sigma(\mu))_{ii} = 0$.

(c) If $\exists i \in \{1,\ldots,n\}$ s.t. $\lambda_i = 0$, then $(\Sigma(\mu))_{ii} = W_{ii}$.

(d) If $\exists i \in \{1,\ldots,n\}$ s.t. $W_{ii} \neq 0$ and the ith column of $C(\mu)$ is zero, then $(\Sigma(\mu))_{ii} = \frac{W_{ii}}{1-\lambda_i^2}$.

(e) If $\exists i \in \{1,\ldots,n\}$ s.t. $\mathbf{e}_i \in \text{rowspace}(C(\mu))$, then $(\Sigma(\mu))_{ii} = W_{ii}$.

□

Lemma 3: Consider a discrete-time linear system as defined in Eq. (1) and Eq. (3). Suppose the system dynamics matrix is of the form $A = \text{diag}(\lambda_1, 0, \ldots, 0) \in \mathbb{R}^{n\times n}$, where $0 < |\lambda_1| < 1$, the measurement matrix $C = [1 \ \gamma]$, where $\gamma \in \mathbb{R}^{1\times(n-1)}$, the system noise covariance matrix $W = I_{n\times n}$, and the sensor noise covariance matrix $V = \mathbf{0}$. Then, the (steady state) MSEE of state 1, i.e., Σ_{11}, satisfies

$$\Sigma_{11} = \frac{1+\alpha^2\lambda_1^2 - \alpha^2 + \sqrt{(\alpha^2 - \alpha^2\lambda_1^2 - 1)^2 + 4\alpha^2}}{2}, \tag{8}$$

where $\alpha^2 \triangleq \|\gamma\|_2^2$. Moreover, if we view Σ_{11} as a function of α^2, denoted as $\Sigma_{11}(\alpha^2)$, then $\Sigma_{11}(\alpha^2)$ is a strictly increasing function of $\alpha^2 \in \mathbb{R}_{\ge 0}$ with $\Sigma_{11}(0) = 1$ and $\lim_{\alpha\to\infty}\Sigma_{11}(\alpha^2) = \frac{1}{1-\lambda_1^2}$. □

A. NP-hardness of the KFSS problem

To prove the KFSS problem (Problem 1) is NP-hard, we relate it to the problem described below.

Definition 1: ($X3C$) Given a finite set D with $|D| = 3m$ and a collection $\mathcal{C} = \{c_1,\ldots,c_\tau\}$ of 3-element subsets of D, an *exact cover* for D is a subcollection $\mathcal{C}' \subseteq \mathcal{C}$ such that every element of D occurs in exactly one member of $\mathcal{C}'$. □

Remark 1: Note that if $\tau < m$, it is clear that there does not exist an exact cover for D. Hence, we assume $\tau \ge m$. Since each member in $\mathcal{C}$ is a subset of D with exactly 3 elements, if there exists an exact cover for D, then it must consist of exactly m members of $\mathcal{C}$. □

We will use the following result [19].

Lemma 4: Given a finite set D with $|D| = 3m$ and a collection $\mathcal{C}$ of 3-element subsets of D, the problem of determining whether $\mathcal{C}$ contains an exact cover for D is NP-complete. □

We are now in place to prove the following result.

Theorem 1: The KFSS problem is NP-hard when the system dynamics matrix A is stable and each sensor $i \in \mathcal{Q}$ has identical cost. □

Proof: We give a reduction from $X3C$ to KFSS. Consider an instance of $X3C$ as described in Definition 1. For each element $c_i \in \mathcal{C}$, define the column vector $g_i \in \mathbb{R}^{3m}$ to encode which elements of D are contained in c_i. Specifically, for $i \in \{1,2,\ldots,\tau\}$ and $j \in \{1,2,\ldots,3m\}$, $(g_i)_j = 1$ if element j of set D is in c_i, and $(g_i)_j = 0$ otherwise. Define the matrix $G = \begin{bmatrix} g_1 & \cdots & g_\tau \end{bmatrix}$ and the vector $d = [1 \ \cdots \ 1]^T \in \mathbb{R}^{3m}$. Thus $Gx = d$ has a solution $x \in \{0,1\}^\tau$ such that x has m nonzero entries if and only if the answer to the instance of $X3C$ is "yes" [20].

Given the above instance of $X3C$, we then construct an instance of KFSS as follows. We define the system dynamics matrix as $A = \text{diag}(\lambda_1, 0, \ldots, 0) \in \mathbb{R}^{(3m+1)\times(3m+1)}$, where $0 < |\lambda_1| < 1$.[1] The set $\mathcal{Q}$ is defined to contain $\tau + 1$ sensors with the collective measurement matrix

$$C = \begin{bmatrix} 1 & d^T \\ \mathbf{0} & G^T \end{bmatrix}, \tag{9}$$

where G and d are defined based on the given instance of $X3C$ as above. The system noise covariance matrix is set to be $W = I_{(3m+1)\times(3m+1)}$, and the measurement noise covariance matrix is set to be $V = \mathbf{0}_{(\tau+1)\times(\tau+1)}$. The sensor cost vector is set as $b = [1 \ \cdots \ 1]^T \in \mathbb{R}^{\tau+1}$, and the sensor selection budget is set as $B = m + 1$. Note that the sensor selection vector for this instance is denoted by $\mu \in \{0,1\}^{\tau+1}$. For the above construction, since the only nonzero eigenvalue of A is λ_1, we know from Lemma 2(c) that $\sum_{i=2}^{3m+1}(\Sigma(\mu))_{ii} = \sum_{i=2}^{3m+1} W_{ii} = 3m$ for all sensor selections μ.

We claim that the solution μ^* to the constructed instance of the KFSS problem satisfies $\text{trace}(\Sigma(\mu^*)) = \text{trace}(W) = 3m+1$ if and only if the answer to the given instance of the $X3C$ problem is "yes".

Suppose that the answer to the instance of the $X3C$ problem is "yes". Then $Gx = d$ has a solution such that x has m nonzero entries. Denote the solution as x^* and denote $\text{supp}(x^*) = \{i_1,\ldots,i_m\}$. Define $\tilde{\mu}$ as the sensor selection vector that indicates selecting the first and the (i_1+1)th to the (i_m+1)th sensors, i.e., sensors that correspond to rows $C_1, C_{i_1+1}, \ldots, C_{i_m+1}$ from (9). Since $Gx^* = d$, we have $[1 \ -x^{*T}]C = \mathbf{e}_1$ for C as defined in Eq. (9). Noting that $\text{supp}(x^*) = \{i_1,\ldots,i_m\}$, it then follows that $\mathbf{e}_1 \in \text{rowspace}(C(\tilde{\mu}))$. Hence, we know from Lemma 2(a) and Lemma 2(e) that $(\Sigma(\tilde{\mu}))_{11} = 1$, which is also the minimum value of $(\Sigma(\mu))_{11}$ among all possible sensor selections μ. Since $\sum_{i=2}^{3m+1}(\Sigma(\mu))_{ii} = 3m$ always holds as argued above, we have $\text{trace}(\Sigma(\tilde{\mu})) = \text{trace}(W) = 3m+1$ and $\tilde{\mu}$ is the optimal sensor selection, i.e., $\tilde{\mu} = \mu^*$.

Conversely, suppose that the answer to the $X3C$ problem is "no". Then, for any union of $l \le m$ ($l \in \mathbb{Z}$) subsets

[1] We take $\lambda_1 = \frac{1}{2}$ for the proof.

in $\mathcal{C}$, denoted as $\mathcal{C}_l$, there exist $\omega \geq 1$ ($\omega \in \mathbb{Z}$) elements in D that are not covered by $\mathcal{C}_l$, i.e., for any $l \leq m$ and $\mathcal{L} \triangleq \{i_1, \ldots, i_l\} \subseteq \{1, \ldots, \tau\}$, $G_{\mathcal{L}} \triangleq \begin{bmatrix} g_{i_1} & \cdots & g_{i_l} \end{bmatrix}$ has ω zero rows, for some $\omega \geq 1$. We then show that $\text{trace}(\Sigma(\mu)) > 3m+1$ for all sensor selections μ that satisfy the budget constraint. First, for any possible sensor selection μ that does not select the first sensor, we have the first column of $C(\mu)$ is zero (from the form of C as defined in Eq. (9)) and we know from Lemma 2(d) that $(\Sigma(\mu))_{11} = \frac{1}{1-\lambda_1^2} = \frac{4}{3}$, which implies that $\text{trace}(\Sigma(\mu)) = 3m + \frac{4}{3} > 3m+1$. Thus, consider sensor selections μ that select the first sensor, denote $\text{supp}(\mu) = \{1, i_1, \ldots, i_l\}$, where $l \leq m$ and define $G(\mu) = \begin{bmatrix} g_{i_1-1} & \cdots & g_{i_l-1} \end{bmatrix}$. We then have

$$C(\mu) = \begin{bmatrix} 1 & d^T \\ \mathbf{0} & G(\mu)^T \end{bmatrix}, \tag{10}$$

where $G(\mu)^T$ has ω zero columns, for some $\omega \geq 1$. As argued in Lemma 5 in the appendix, there exists an orthogonal matrix $T \in \mathbb{R}^{(3m+1)\times(3m+1)}$ of the form $T = \left[\begin{smallmatrix} 1 & 0 \\ 0 & N \end{smallmatrix}\right]$ such that

$$\tilde{C}(\mu) \triangleq C(\mu)T = \begin{bmatrix} 1 & \gamma & \beta \\ \mathbf{0} & \mathbf{0} & \tilde{G}(\mu)^T \end{bmatrix}.$$

In the above expression, $\tilde{G}(\mu)^T \in \mathbb{R}^{l\times r}$ is of full column rank, where $r = \text{rank}(G(\mu)^T)$. Furthermore, $\gamma \in \mathbb{R}^{1\times(3m-r)}$ and ω of its elements are 1's, and $\beta \in \mathbb{R}^{1\times r}$. We perform a similarity transformation on the system with T (which does not affect the trace of the error covariance matrix in general and does not change A and W in this case), and perform additional elementary row operations to transform $\tilde{C}(\mu)$ into the matrix

$$\tilde{C}'(\mu) = \begin{bmatrix} 1 & \gamma & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \tilde{G}(\mu)^T \end{bmatrix}. \tag{11}$$

Since A and W are both diagonal, and $V = \mathbf{0}$, we can obtain from Eq. (4) that the steady state error covariance $\tilde{\Sigma}'(\mu)$ corresponding to the sensing matrix $\tilde{C}'(\mu)$ is of the form

$$\tilde{\Sigma}'(\mu) = \begin{bmatrix} \tilde{\Sigma}_1'(\mu) & \mathbf{0} \\ \mathbf{0} & \tilde{\Sigma}_2'(\mu) \end{bmatrix},$$

where $\tilde{\Sigma}_1'(\mu) \in \mathbb{R}^{(3m+1-r)\times(3m+1-r)}$, denoted as Σ for simplicity, satisfies

$$\Sigma = A_1 \Sigma A_1^T + W_1 - A_1 \Sigma C_1^T \left(C_1 \Sigma C_1^T\right)^{-1} C_1 \Sigma A_1^T,$$

where $A_1 = \text{diag}(\lambda_1, 0, \ldots, 0) \in \mathbb{R}^{(3m+1-r)\times(3m+1-r)}$, $C_1 = [1 \ \gamma]$ and $W_1 = I_{(3m+1-r)\times(3m+1-r)}$. We then know from Lemma 3 that $(\Sigma(\mu))_{11} = (\tilde{\Sigma}'(\mu))_{11} > 1$ since $\|\gamma\|_2^2 \geq \omega \geq 1 > 0$. Hence, we have $\text{trace}(\Sigma(\mu)) > 3m+1$.

This completes the proof of the claim above. Suppose that there is an algorithm $\mathcal{A}$ that outputs the optimal solution μ^* to the instance of the KFSS problem defined above. We can call algorithm $\mathcal{A}$ to solve the $X3C$ problem. Specifically, if the algorithm $\mathcal{A}$ outputs a solution μ^* such that $\text{trace}(\Sigma(\mu^*)) = \text{trace}(W)$, then the answer to the instance of $X3C$ is "yes"; otherwise, the answer is "no".

Hence, we have a reduction from $X3C$ to KFSS. Since $X3C$ is NP-complete and KFSS $\notin NP$, we conclude that the KFSS problem is NP-hard. ■

B. Inapproximability of the KFSS Problem

In this section, we analyze the achievable performance of algorithms for the KFSS problem. Specifically, consider any given instance of KFSS. For any given algorithm $\mathcal{A}$, we define the following ratio:

$$r_{\mathcal{A}}(\Sigma) \triangleq \frac{\text{trace}(\Sigma_{\mathcal{A}})}{\text{trace}(\Sigma_{opt})}, \tag{12}$$

where Σ_{opt} is the optimal solution to KFSS and $\Sigma_{\mathcal{A}}$ is the solution to KFSS given by algorithm $\mathcal{A}$.

In [10], the authors showed that there is an upper bound for $r_{\mathcal{A}}(\Sigma)$ for any sensor selection algorithm $\mathcal{A}$, in terms of the system matrices. However, the question of whether it is possible to find an algorithm $\mathcal{A}$ that is guaranteed to provide an approximation ratio $r_{\mathcal{A}}(\Sigma)$ that is *independent* of the system parameters has remained open up to this point. It is typically desirable to find *constant-factor* approximation algorithms, where the ratio $r_{\mathcal{A}}(\Sigma)$ is upper-bounded by some (system-independent) constant. Here, we provide a strong negative result showing that for the KFSS problem, there is no constant-factor approximation algorithm in general, i.e., for all polynomial-time algorithms $\mathcal{A}$ and $\forall K \in \mathbb{R}_{\geq 1}$, there are instances of KFSS where $r_{\mathcal{A}}(\Sigma) > K$.

Theorem 2: If $P \neq NP$, then there is no polynomial-time constant-factor approximation algorithm for the KFSS problem. □

Proof: Suppose that there exists such a (polynomial-time) approximation algorithm $\mathcal{A}$, i.e., $\exists K \in \mathbb{R}_{\geq 1}$ such that $r_{\mathcal{A}}(\Sigma) \leq K$ for all instances of the KFSS problem, where $r_{\mathcal{A}}(\Sigma)$ is as defined in Eq. (12). We will show that $\mathcal{A}$ can be used to solve the $X3C$ problem. Given an arbitrary instance of the $X3C$ problem as described in Definition 1, we construct a corresponding instance of KFSS in a similar way to that described in the proof of Theorem 1. Specifically, the system dynamics matrix is set as $A = \text{diag}(\lambda_1, 0, \ldots, 0) \in \mathbb{R}^{(3m+1)\times(3m+1)}$, where $\lambda_1 = \frac{K(3m+1)-3m-1/2}{K(3m+1)-3m}$. The set $\mathcal{Q}$ contains $\tau+1$ sensors with collective measurement matrix

$$C = \begin{bmatrix} 1 & \varepsilon d^T \\ \mathbf{0} & G^T \end{bmatrix}, \tag{13}$$

where G, d depend on the given instance of $X3C$ and are as defined in the proof of Theorem 1. The constant ε is chosen as $\varepsilon = 2[(K(3m+1)-3m)]\left\lceil \sqrt{K(3m+1)-3m-1} \right\rceil + 1$. The system noise covariance matrix W, the measurement noise covariance matrix V, the sensor cost vector b, and the selection budget B are set to be the same as in the proof of Theorem 1.

We claim that algorithm $\mathcal{A}$ will return a sensor selection vector μ such that $\text{trace}(\Sigma(\mu)) \leq K(3m+1)$ if and only if the answer to the $X3C$ problem is "yes".

Suppose that the answer to the $X3C$ problem is "yes". We know from Theorem 1 that there exists a sensor selection μ^* such that $\text{trace}(\Sigma(\mu^*)) = 3m+1$. Since $\mathcal{A}$ has approximation ratio K, it returns a sensor selection μ such that $\text{trace}(\Sigma(\mu)) \leq K(3m+1)$.

Conversely, suppose that the answer to the $X3C$ problem is "no". We follow the discussion in Theorem 1. First, for any sensor selection μ that does not select the first sensor, we have $(\Sigma(\mu))_{11} = \frac{1}{1-\lambda_1^2}$. Hence, by our choice of λ_1, we have $(\Sigma(\mu))_{11} > K(3m+1) - 3m$, which implies $\text{trace}(\Sigma(\mu)) > K(3m+1)$ since $\sum_{i=2}^{3m+1}(\Sigma(\mu))_{ii} = 3m$ for all possible sensor selections. Thus, consider sensor selections μ that include the first sensor. As argued in the proof of Theorem 1 leading up to Eq. (11), we can perform an orthogonal similarity transformation on the system, along with elementary row operations on the measurement matrix $C(\mu)$ to obtain a measurement matrix of the form

$$\tilde{C}'(\mu) = \begin{bmatrix} 1 & \varepsilon\gamma & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \tilde{G}(\mu)^T \end{bmatrix}, \tag{14}$$

where $\omega \geq 1$ elements of $\gamma \in \mathbb{R}^{3m-r}$ are 1's and $r = \text{rank}(\tilde{G}(\mu)^T)$. Then, we have $\alpha^2 \triangleq \varepsilon^2\|\gamma\|_2^2 \geq \omega\varepsilon^2 \geq \varepsilon^2$. We then obtain from Lemma 3 that $\Sigma_{11}(\alpha^2) \geq \Sigma_{11}(\varepsilon^2)$, i.e.,

$$\Sigma_{11}(\alpha^2) \geq \frac{1+\varepsilon^2\lambda_1^2 - \varepsilon^2 + \sqrt{(\varepsilon^2 - \varepsilon^2\lambda_1^2 - 1)^2 + 4\varepsilon^2}}{2},$$

where we view $(\Sigma(\mu))_{11}$ as a function of α^2, denoted as $\Sigma_{11}(\alpha^2)$. By our choices of λ_1 and ε, we have $(\Sigma(\mu))_{11} > K(3m+1) - 3m$, which implies $\text{trace}(\Sigma(\mu)) > K(3m+1)$.

This completes the proof of the claim above. Hence, if algorithm $\mathcal{A}$ for the KFSS problem has $r_{\mathcal{A}}(\Sigma) \leq K$ for all instances, it is clear that $\mathcal{A}$ can be used to solve the $X3C$ problem by applying it to the above instance. Specifically, if the answer to the $X3C$ instance is "yes", then the optimal sensor selection μ^* would yield a trace of $\Sigma(\mu^*) = 3m + 1$, and thus the algorithm $\mathcal{A}$ would yield a trace no larger than $K(3m+1)$. On the other hand, if the answer to the $X3C$ instance is "no", all sensor selections would yield a trace larger than $K(3m+1)$, and thus so would the sensor selection provided by $\mathcal{A}$. In either case, the solution provided by $\mathcal{A}$ could be used to find the answer to the given $X3C$ instance. Since $X3C$ is NP-complete, there is no polynomial-time algorithm for it if $P \neq NP$, and we get a contradiction. This completes the proof of the theorem. ∎

IV. Greedy Algorithm

Our result in Theorem 2 indicates that no polynomial-time algorithm can be guaranteed to yield a solution that is within any constant factor of the optimal solution. In particular, this result applies to the greedy algorithms that are often studied for sensor selection in the literature [10]. In particular, it was shown via simulations in [10] that such algorithms work well in practice (e.g., for randomly generated systems). In this section, we provide an explicit example showing that greedy algorithms for KFSS can perform arbitrarily poorly, even for small systems. We will focus on the simple greedy algorithm for KFSS defined as Algorithm 1, for instances where all sensor costs are equal to 1, and the sensor selection budget $B = p_s$ for some $p_s \in \{1, \ldots, q\}$. For any such instance of KFSS, define $r_{gre}(\Sigma) = \frac{\text{trace}(\Sigma_{gre})}{\text{trace}(\Sigma_{opt})}$, where Σ_{gre} is the solution of the DARE corresponding to the sensors selected by Algorithm 1.

Algorithm 1 Greedy Algorithm for KFSS

Input: System dynamics matrix A, set of all candidate sensors $\mathcal{Q}$, noise covariances W and V, budget p_s

Output: A set $\mathcal{S}$ of selected sensors

```
k ← 1, S ← ∅
for k ≤ p_s do
    for i ∈ Q ∩ S̄ do
        Calculate trace(Σ(S ∪ {i}))
    end for
    j = arg min_i trace(Σ(S ∪ {i}))
    S ← S ∪ {j}, k ← k + 1
end for
```

Example 1: Consider an instance of KFSS with matrices $W = I_{3\times 3}$, $V = \mathbf{0}_{3\times 3}$, and A, C defined as

$$A = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, C = \begin{bmatrix} 1 & h & h \\ 1 & 0 & h \\ 0 & 1 & 1 \end{bmatrix},$$

where $0 < |\lambda_1| < 1$, $\lambda_1 \in \mathbb{R}$ and $h \in \mathbb{R}_{>0}$. In addition, we have the set of candidate sensors $\mathcal{Q} = \{1, 2, 3\}$, the selection budget $B = 2$ and the cost vector $b = [1\ 1\ 1]^T$. □

Theorem 3: For the instance of the KFSS problem defined in Example 1, the ratio $r_{gre}(\Sigma) = \frac{\text{trace}(\Sigma_{gre})}{\text{trace}(\Sigma_{opt})}$ satisfies

$$\lim_{h\to\infty} r_{gre}(\Sigma) = \frac{2}{3} + \frac{1}{3(1-\lambda_1^2)}. \tag{15}$$

□

We give a sketch of the proof here; the complete proof can be found in [21].

Sketch of the proof:

Since the only nonzero eigenvalue of A is λ_1, we know from Lemma 2(c) that $(\Sigma(\mu))_{22} = 1$ and $(\Sigma(\mu))_{33} = 1$, $\forall\mu$, which implies that $(\Sigma_{gre})_{22} = 1$ and $(\Sigma_{gre})_{33} = 1$. Hence, we focus on determining $(\Sigma_{gre})_{11}$.

Using Lemma 2 and Lemma 3, we prove that the greedy algorithm defined as Algorithm 1 selects sensor 2 and sensor 3 in its first and second iterations, and $(\Sigma(\mu))_{11}\big|_{\mu=[0\ 1\ 1]^T}$, denoted as σ_{23}, is given by

$$\sigma_{23} = \frac{2}{\sqrt{(1-\lambda_1^2 - \frac{2}{h^2})^2 + \frac{8}{h^2}} + 1 - \lambda_1^2 - \frac{2}{h^2}}.$$

Hence, we have $\text{trace}(\Sigma_{gre}) = \sigma_{23} + 2$.

If $\mu = [1\ 0\ 1]^T$, then $\mathbf{e}_1 \in \text{rowspace}(C(\mu))$ and we know from Lemma 2(a) and Lemma 2(e) that $\text{trace}(\Sigma(\mu)) = 3 = \text{trace}(W)$, which is also the minimum value of $\text{trace}(\Sigma(\mu))$ among all possible sensor selections μ. Combining the results above and taking the limit as $h \to \infty$, we obtain the result in Eq. (15). ∎

Examining Eq. (15), we see that for the given instance of KFSS, we have $r_{gre}(\Sigma) \to \infty$ as $h \to \infty$ and $\lambda_1 \to 1$. Thus, $r_{gre}(\Sigma)$ can be made arbitrarily large by choosing the parameters in the instance appropriately. It is also useful to note that the above behavior holds for any algorithm that outputs a sensor selection that contains sensor 2 for the above example.

V. Conclusions

In this paper, we studied the KFSS problem for linear dynamical systems. We showed that this problem is NP-hard and has no constant-factor approximation algorithms, even under the assumption that the system is stable and each sensor has identical cost. We provided an explicit example showing how a greedy algorithm can perform arbitrarily poorly on this problem, even when the system only has three states. Our results shed new insights into the problem of sensor selection for Kalman filtering and show, in particular, that this problem is more difficult than other variants of the sensor selection problem that have submodular cost functions. Future work on characterizing achievable (non-constant) approximation ratios, or identifying classes of systems that admit near-optimal approximation algorithms, would be of interest.

Appendix

Lemma 5: Consider an instance of $X3C$ as described in Definition 1. For each element $c_i \in \mathcal{C}$, define the column vector $g_i \in \mathbb{R}^{3m}$ such that for $i \in \{1,2,\ldots,\tau\}$ and $j \in \{1,2,\ldots,3m\}$, $(g_i)_j = 1$ if element j of set D is in c_i, and $(g_i)_j = 0$ otherwise. Define $G = \begin{bmatrix} g_1 & \cdots & g_\tau \end{bmatrix}$ and $d = [1 \cdots 1]^T \in \mathbb{R}^{3m}$. For any $l \leq m$ ($l \in \mathbb{Z}$) and $\mathcal{L} \triangleq \{i_1,\ldots,i_l\} \subseteq \{1,\ldots,\tau\}$, define $G_{\mathcal{L}} = \begin{bmatrix} g_{i_1} & \cdots & g_{i_l} \end{bmatrix}$ and denote rank$(G_{\mathcal{L}}) = r_{\mathcal{L}}$.[2]. If the answer to the $X3C$ problem is "no", then for all $\mathcal{L}$ with $|\mathcal{L}| \leq m$, there exists an orthogonal matrix $N \in \mathbb{R}^{3m\times 3m}$ such that

$$\begin{bmatrix} d^T \\ G_{\mathcal{L}}^T \end{bmatrix} N = \begin{bmatrix} \gamma & \beta \\ \mathbf{0} & \tilde{G}_{\mathcal{L}}^T \end{bmatrix},$$

where $\tilde{G}_{\mathcal{L}}^T \in \mathbb{R}^{l\times r}$ is of full column rank, $\gamma \in \mathbb{R}^{1\times(3m-r)}$ and $\omega \geq 1$ ($\omega \in \mathbb{Z}$) elements of γ are 1's , and $\beta \in \mathbb{R}^{1\times r}$. Further elementary row operations on $\begin{bmatrix} \gamma & \beta \\ \mathbf{0} & \tilde{G}_{\mathcal{L}}^T \end{bmatrix}$ transform it into the form $\begin{bmatrix} \gamma & \mathbf{0} \\ \mathbf{0} & \tilde{G}_{\mathcal{L}}^T \end{bmatrix}$. □

Proof: Assume without loss of generality that there are no identical subsets in $\mathcal{C}$. Since rank$(G_{\mathcal{L}}^T) = r$, the dimension of nullspace$(G_{\mathcal{L}}^T)$ is $3m-r$. We choose an orthonormal basis of nullspace$(G_{\mathcal{L}}^T)$ and let it form the first $3m-r$ columns of N, denoted as N_1. Then, we choose an orthonormal basis of rowspace$(G_{\mathcal{L}}^T)$ and let it form the rest of the r columns of N, denoted as N_2. Hence, $N = \begin{bmatrix} N_1 & N_2 \end{bmatrix} \in \mathbb{R}^{3m\times 3m}$ is an orthogonal matrix. Furthermore, since the answer to the $X3C$ problem is "no", for any union of $l \leq m$ ($l \in \mathbb{Z}$) subsets in $\mathcal{C}$, denoted as $\mathcal{C}_l$, there exist $\omega \geq 1$ ($\omega \in \mathbb{Z}$) elements in D that are not covered by $\mathcal{C}_l$, i.e., $G_{\mathcal{L}}^T$ has ω zero columns. Denote these as the j_1th, $\ldots$, j_ωth columns of $G_{\mathcal{L}}^T$, where $\{j_1,\ldots,j_\omega\} \subseteq \{1,\ldots,3m\}$. Hence, we can always choose $\mathbf{e}_{j_1},\ldots,\mathbf{e}_{j_\omega}$ to be in the orthonormal basis of nullspace$(G_{\mathcal{L}}^T)$, i.e., as columns of N_1. Constructing N in this way, we have $G_{\mathcal{L}}^T N_1 = \mathbf{0}$ and $G_{\mathcal{L}}^T N_2 = \tilde{G}_{\mathcal{L}}^T$, where $\tilde{G}_{\mathcal{L}}^T \in \mathbb{R}^{l\times r}$ is of full column rank since the columns of N_2 form an orthonormal basis of rowspace$(G_{\mathcal{L}}^T)$ and $r \leq l$. Moreover, we have $d^T N_1 = \gamma$ and $d^T N_2 = \beta$, where ω elements of γ are 1's (since $d^T \mathbf{e}_{j_s}^T = 1$, $\forall s \in \{1,\ldots,\omega\}$). Combining these results, we obtain Eq. (5). Since $\tilde{G}_{\mathcal{L}}^T \in \mathbb{R}^{l\times r}$ is of full column rank, we can perform elementary row operations on $\begin{bmatrix} \gamma & \beta \\ \mathbf{0} & \tilde{G}_{\mathcal{L}}^T \end{bmatrix}$ and obtain $\begin{bmatrix} \gamma & \mathbf{0} \\ \mathbf{0} & \tilde{G}_{\mathcal{L}}^T \end{bmatrix}$. ■

[2]We drop the subscript $\mathcal{L}$ on r for notational simplicity.

References

[1] 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.

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

[3] 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.

[4] 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.

[5] M. F. Huber, "Optimal pruning for multi-step sensor scheduling," *IEEE Transactions on Automatic Control*, vol. 57, no. 5, pp. 1338–1343, 2012.

[6] 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.

[7] D. J. Chmielewski, T. Palmer, and V. Manousiouthakis, "On the theory of optimal sensor placement," *AIChE journal*, vol. 48, no. 5, pp. 1001–1012, 2002.

[8] N. K. Dhingra, M. R. Jovanović, and Z.-Q. Luo, "An ADMM algorithm for optimal sensor and actuator selection," in *Decision and Control (CDC), 2014 IEEE 53rd Annual Conference on*. IEEE, 2014, pp. 4039–4044.

[9] 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.

[10] 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.

[11] G. L. Nemhauser, L. A. Wolsey, and M. L. Fisher, "An analysis of approximations for maximizing submodular set functions," *Mathematical Programming*, vol. 14, no. 1, pp. 265–294, 1978.

[12] A. Krause, A. Singh, and C. Guestrin, "Near-optimal sensor placements in Gaussian processes: Theory, efficient algorithms and empirical studies," *Journal of Machine Learning Research*, vol. 9, no. Feb, pp. 235–284, 2008.

[13] M. Shamaiah, S. Banerjee, and H. Vikalo, "Greedy sensor selection: Leveraging submodularity," in *Decision and Control (CDC), 2010 49th IEEE Conference on*. IEEE, 2010, pp. 2572–2577.

[14] 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.

[15] C. Yang, J. Wu, X. Ren, W. Yang, H. Shi, and L. Shi, "Deterministic sensor selection for centralized state estimation under limited communication resource," *IEEE transactions on signal processing*, vol. 63, no. 9, pp. 2336–2348, 2015.

[16] B. D. Anderson and J. B. Moore, *Optimal Filtering*. Dover Books, 1979.

[17] R. Horn and C. Johnson, *Matrix Analysis*. Cambridge University Press, 1985.

[18] D. E. Catlin, *Estimation, control, and the discrete Kalman filter*. Springer, 1989.

[19] M. R. Garey and D. S. Johnson, *Computers and intractability: a guide to the theory of NP-completeness*. Freeman, 1979.

[20] B. K. Natarajan, "Sparse approximate solutions to linear systems," *SIAM journal on computing*, vol. 24, no. 2, pp. 227–234, 1995.

[21] L. Ye, S. Roy, and S. Sundaram, "On the complexity and approximability of optimal sensor selection for Kalman filtering," *arXiv preprint arXiv:1711.01920*, 2017.