Sensor Selection for Kalman Filtering of Linear Dynamical Systems: Complexity, Limitations and Greedy Algorithms [⋆]

Haotian Zhang [a], Raid Ayoub [b], Shreyas Sundaram [c]

[a] *Department of Electrical and Computer Engineering at the University of Waterloo, Waterloo, ON N2L 3G1, Canada*

[b] *Strategic CAD Labs, Intel Corporation, Hillsboro, Oregon 97124, USA*

[c] *School of Electrical and Computer Engineering at Purdue University, West Lafayette, IN 47907, USA*

Abstract

We consider the problem of selecting an optimal set of sensors to estimate the states of linear dynamical systems. Specifically, the goal is to choose (at design-time) a subset of sensors (satisfying certain budget constraints) from a given set in order to minimize the trace of the steady state *a priori* or *a posteriori* error covariance produced by a Kalman filter. We show that the *a priori* and *a posteriori* error covariance-based sensor selection problems are both NP-hard, even under the additional assumption that the system is stable. We then provide bounds on the worst-case performance of sensor selection algorithms based on the system dynamics, and show that greedy algorithms are optimal for a certain class of systems. However, as a negative result, we show that certain typical objective functions are not submodular or supermodular in general. While this makes it difficult to evaluate the performance of greedy algorithms for sensor selection (outside of certain special cases), we show via simulations that these greedy algorithms perform well in practice.

Key words: Sensor selection; Kalman filters; Complexity; Greedy algorithms; Multi-sensor estimation.

1 Introduction

One of the key problems in control system design is to select an appropriate set of actuators or sensors (either at design-time or at run-time) in order to achieve certain performance objectives [33]. For the objective of estimating the state of a given linear Gauss-Markov system, there has been a growing literature in the past few years that studies how to dynamically select sensors at run-time to minimize certain metrics of the error covariance of the corresponding Kalman filter. This is known as the *sensor scheduling problem*, due to the fact that a different set of sensors can be chosen at each time-step (e.g., see [9,12]).

The design-time sensor selection problem (where the set of chosen sensors is not allowed to change over time) has been studied in various forms, including cases where the objective is to guarantee a certain structural property of the system [26], to optimize energy or information theoretic metrics [29,17], or to compute the optimal sensing matrix under a norm constraint [2]. [1] Various sensor selection heuristics have also been proposed for estimation of *static* random variables (e.g., see [13,6,24]); however, the corresponding results do not directly translate to the case of estimating the (vector) state of dynamical systems.

In [8], the authors studied the design-time actuator/sensor selection problem for continuous-time linear dynamical systems using the sparsity-promoting framework from [20,27]. For the sensor selection problem, the objective is to design a Kalman gain matrix to minimize the resulting H_2 norm from the noise to the predicted estimation error. Sparsity is achieved by adding a penalty function for non-zero columns of the gain matrix. In contrast to the formulation in [8], in this paper, we directly

⋆ This paper was not presented at any IFAC meeting. Some of the results in this paper appeared in preliminary form in [36]. This material is based upon work supported by a grant from Intel Corporation. Corresponding author: Shreyas Sundaram. Telephone of the corresponding author: 1-(765)-496-0406.

Email addresses: h223zhan@uwaterloo.ca (Haotian Zhang), raid.ayoub@intel.com (Raid Ayoub), sundara2@purdue.edu (Shreyas Sundaram).

[1] There have also been various recent studies of the dual design-time actuator placement problem (e.g., see [32,27]).

focus on minimizing functions of the steady state error covariances of discrete-time Kalman filters, and impose a hard constraint on the set of sensors to be chosen.

In [30,31], the authors studied the design-time sensor selection problem for discrete-time linear time-varying systems over a finite horizon. They assumed that each sensor directly measures one component of the state vector, and the objective is either to minimize the estimation error with a cardinality constraint or to minimize the number of chosen sensors while guaranteeing a certain level of performance. Different from the formulation in [30,31], we consider general measurement matrices and focus on minimizing the steady state estimation error of the Kalman filter.

In [34], the authors considered the same problem as the one we considered here, namely the design-time sensor selection problem for Kalman filtering in discrete-time linear dynamical systems with hard constraints. They showed that the sensor selection problem can be expressed as a semidefinite program (SDP). However, the results in [34] can only be applied to systems where the sensor noise terms are uncorrelated, and no theoretical guarantees were provided on the performance of the proposed heuristics.

In this paper, we consider the design-time sensor selection problem for optimal filtering of discrete-time linear dynamical systems. Specifically, we study the problem of choosing a set of sensors (under certain constraints) to optimize either the *a priori* or the *a posteriori* error covariance of the corresponding Kalman filter; we will refer to these problems as the priori and posteriori *Kalman filtering sensor selection (KFSS)* problems, respectively. Note that the priori KFSS problem is applicable for settings where a prediction of system states is needed and the posteriori KFSS problem is suitable for applications where the estimation can be conducted after receiving up-to-date measurements [1].

Our contributions are threefold. First, we show that it is NP-hard to find the optimal solution of cost-constrained priori and posteriori KFSS problems, even under the assumption that the system is stable. It is often claimed in the literature that sensor selection problems are intractable [13,11]; however, except for certain problems with utility or energy based cost functions (e.g., see [32,3]), to the best of our knowledge, there is still no explicit characterization of the complexity of the optimal-filtering based sensor selection problems considered in this paper.

Our second contribution is to provide insights into what factors of the system affect the performance of sensor selection algorithms by using the concept of the *sensor information matrix* [11]. For the priori KFSS problem, we show that when the system is stable, the worst-case performance can be bounded by a parameter that depends only on the system dynamics matrix, and that the performance of a sensor selection algorithm cannot be arbitrarily bad if the system matrix is well conditioned, even under very large noise. For the posteriori KFSS problem, we show that for a given system, the worst-case performance of any selection of sensors can be upper-bounded in terms of the eigenvalues of the system noise covariance matrix and the corresponding sensor information matrix.

Since it is intractable to find the optimal selection of sensors in general, a reasonable tradeoff is to design appropriate approximation algorithms. In [12], the authors considered various cost functions for the (run-time) sensor scheduling problem. They showed that one of these considered cost functions is submodular while the others are not; for the submodular cost function, a certain greedy algorithm can be applied to obtain guaranteed performance. Greedy algorithms have also drawn much attention for other forms of sensor selection problems, e.g., see [28,29,17,31]. Thus, our third contribution is the study of greedy algorithms for the priori and posteriori KFSS problems. We first show that greedy algorithms are optimal (with respect to the corresponding KFSS problems) for a certain class of systems. However, for general systems, as a negative result, we show that the cost functions of both the priori and posteriori KFSS problems (and the other cost functions studied in [12]) do not necessarily have certain modularity properties. This precludes the direct application of classical results from the theory of combinatorial optimization and implies that the underlying structures of the KFSS problems are different from the other types of sensor selection problems. Nevertheless, we show via simulations that greedy algorithms perform well in practice. Moreover, compared to the algorithms in [34], the greedy algorithms provided in this paper can be applied to a more general class of systems (where the sensor noises are correlated), are more efficient and (in simulations) provide comparable performance.

The rest of the paper is organized as follows. In Section 2, we formulate the (design-time) sensor selection problems. In Section 3, we analyze the complexity of the priori and posteriori KFSS problems. In Section 4, we provide worst-case guarantees on the performance of sensor selection algorithms. In Section 5, we propose and study two greedy algorithms for sensor selection, and illustrate their performance and complexity in Section 6. We conclude in Section 7.

1.1 Notation and Terminology

The set of integers, real numbers and complex numbers are denoted as $\mathbb{Z}$, $\mathbb{R}$ and $\mathbb{C}$, respectively. For a square matrix $M \in \mathbb{R}^{n\times n}$, let M^T, trace(M), det(M), $\{\lambda_i(M)\}$ and $\{\sigma_i(M)\}$ be its transpose, trace, determinant, set of eigenvalues and set of singular values, respectively. The set of eigenvalues $\{\lambda_i(M)\}$ of M are ordered with nondecreasing magnitude (i.e., $|\lambda_1(M)| \geq \cdots \geq |\lambda_n(M)|$); the same order applies to the set of singular values $\{\sigma_i(M)\}$. A positive semi-definite matrix M is denoted by $M \succeq 0$ and $M \succeq N$ if $M - N \succeq 0$; the set of n by n positive semi-definite (resp. positive definite) matrices is denoted by $\mathbb{S}_+^n$ (resp. $\mathbb{S}_{++}^n$). The identity matrix with dimension n is denoted by $I_{n\times n}$. For a vector v, let $\text{diag}(v)$ be the diagonal matrix with diagonal entries being the elements of v; for a set of matrices $\{M_i\}_{i=1}^q$, let $\text{diag}(M_1, \cdots, M_q)$ be the block diagonal matrix with the i-th diagonal block being M_i. For a random variable w, denote $\mathbb{E}[w]$ as its expectation.

2 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 $\mathbb{E}\left[w[k](w[k])^T\right] = W$ for all $k \in \mathbb{N}$, and $A \in \mathbb{R}^{n\times n}$ is the system dynamics matrix. We assume throughout that the pair $(A, W^{\frac{1}{2}})$ is stabilizable.

The set of sensors to be installed on the system must come from a given set $\mathcal{Q}$ consisting of q sensors. Each sensor $i \in \mathcal{Q}$ provides a measurement of 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 that sensor, and $v_i[k] \in \mathbb{R}^{s_i}$ is a zero-mean white Gaussian noise process. For convenience, we 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 $v[k] \triangleq \begin{bmatrix}(v_1[k])^T & \cdots & (v_q[k])^T\end{bmatrix}^T$. Then the measurement equation corresponding to the output of all sensors is

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

Denote $\mathbb{E}\left[v[k](v[k])^T\right] = V$ and take $\mathbb{E}\left[v[k](w[j])^T\right] = 0$ for all $j, k \in \mathbb{N}$. We assume that the pair (A, C) is detectable.

Each sensor $i \in \mathcal{Q}$ has an associated *cost* $b_i \in \mathbb{R}_{\geq 0}$, representing, for example, monetary costs of purchasing and installing that sensor or the energy consumption of that sensor. Define the cost vector $b \triangleq [b_1 \cdots b_q]^T$. We also assume there is a *sensor budget* $B \in \mathbb{R}_{\geq 0}$, representing the total cost that can be spent on sensors from $\mathcal{Q}$.

For any given subset of sensors that is installed, the Kalman filter provides the optimal estimate of the state using the measurements from those sensors (in the sense of minimizing mean square estimation error, under the stated assumptions on the noise processes). Let $z \in \{0,1\}^q$ be the indicator vector of the installed sensors, i.e., $z_i = 1$ if and only if sensor $i \in \mathcal{Q}$ is installed. Define the selection matrix $Z \triangleq \text{diag}(z_1 I_{s_1\times s_1}, \cdots, z_q I_{s_q\times s_q})$ and denote $\tilde{C} \triangleq ZC$ and $\tilde{V} \triangleq ZVZ^T$. Let $\Sigma_{k|k-1}(z)$ and $\Sigma_{k|k}(z)$ be the *a priori* error covariance matrix and the *a posteriori* error covariance matrix (at time-step k) of the Kalman filter when the set of sensors indicated by the vector z are installed, respectively. If the pair $(A, \tilde{C})$ is detectable (and given the stabilizability of $(A, W^{\frac{1}{2}})$), both $\Sigma_{k|k-1}(z)$ and $\Sigma_{k|k}(z)$ will converge to unique limits [1]; denote the limits of $\Sigma_{k|k-1}(z)$ and $\Sigma_{k|k}(z)$ by $\Sigma(z)$ and $\Sigma^*(z)$, respectively. We will also use $\Sigma(\mathcal{S})$ and $\Sigma^*(\mathcal{S})$ to denote these quantities for a specific set of sensors $\mathcal{S} \subseteq \mathcal{Q}$.

The limit $\Sigma(z)$ of the *a priori* error covariance satisfies the following *discrete algebraic Riccati equation (DARE)* [1]:

$$\begin{aligned}\Sigma(z) = {} & A\Sigma(z)A^T + W - \\ & A\Sigma(z)\tilde{C}^T(\tilde{C}\Sigma(z)\tilde{C}^T + \tilde{V})^{-1}\tilde{C}\Sigma(z)A^T. \end{aligned} \tag{4}$$

Define $\tilde{\Sigma}(z) \triangleq A\Sigma^*(z)A^T + W$. The limit $\Sigma^*(z)$ of the *a posteriori* error covariance satisfies the following equation [5]:

$$\Sigma^*(z) = \tilde{\Sigma}(z) - \tilde{\Sigma}(z)\tilde{C}^T\left(\tilde{C}\tilde{\Sigma}(z)\tilde{C}^T + \tilde{V}\right)^{-1}\tilde{C}\tilde{\Sigma}(z). \tag{5}$$

Using the matrix inversion lemma [10], the DARE (4) and the equation (5) can also be written as

$$\Sigma(z) = W + A(\Sigma^{-1}(z) + R(z))^{-1}A^T, \tag{6}$$

and

$$\Sigma^*(z) = \left((A\Sigma^*(z)A^T + W)^{-1} + R(z)\right)^{-1}, \tag{7}$$

respectively, where the matrix $R(z) \triangleq \tilde{C}^T\tilde{V}^{-1}\tilde{C}$ is the so-called sensor information matrix corresponding to the indicator vector z. [2] Note that $\Sigma(z)$ and $\Sigma^*(z)$ are coupled as follows [5]:

$$\Sigma^*(z) = (\Sigma^{-1}(z) + R(z))^{-1}. \tag{8}$$

Further note that the inverses in the equations (4)-(8) are interpreted as pseudo-inverses if the arguments are not invertible. [3]

[2] Note that the sensor information matrix is different from the *Fisher information matrix*, which is the inverse of the error covariance matrix [1].

[3] For the special case of $V = 0$, the matrix inversion lemma does not hold under pseudo-inverses (unless $z = 0$) and thus we compute $\Sigma(z)$ and $\Sigma^*(z)$ by equations (4) and (5).

Definition 1 (Feasible Sensor Selection) *The sensor selection $z \in \{0,1\}^q$ is said to be **feasible** if the cost constraint is satisfied (i.e., $b^T z \leq B$) and both $\Sigma_{k|k-1}(z)$ and $\Sigma_{k|k}(z)$ converge to finite limits (denoted by $\Sigma(z)$ and $\Sigma^*(z)$, respectively) which do not depend on $\Sigma_{0|0}(z)$ as $k \to \infty$.*

We will use the following well-known result on Kalman filtering [1].

Lemma 1 *When the pair $(A, W^{\frac{1}{2}})$ is stabilizable, the indicator vector z satisfying $b^T z \leq B$ is feasible if and only if the pair $(A, \tilde{C})$ is detectable.*

We now propose the following priori and posteriori Kalman filtering sensor selection (KFSS) problems. Denote $s = \sum_{i=1}^{q} s_i$.

Problem 1 (Priori KFSS Problem) *Given a system dynamics matrix $A \in \mathbb{R}^{n\times n}$, a measurement matrix $C \in \mathbb{R}^{s\times n}$, 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 priori KFSS problem is to find a sensor selection vector z that solves*

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

where $\Sigma(z)$ is given by equation (4), or else determine that no feasible sensor selection exists.

Problem 2 (Posteriori KFSS Problem) *Given a system dynamics matrix $A \in \mathbb{R}^{n\times n}$, a measurement matrix $C \in \mathbb{R}^{s\times n}$, 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 posteriori KFSS problem is to find a sensor selection vector z that solves*

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

where $\Sigma^*(z)$ is given by equation (5), or else determine that no feasible sensor selection exists.

Note that the only difference between Problem 1 and Problem 2 is the objective function (the former is to minimize $\text{trace}(\Sigma(z))$ and the latter is to minimize $\text{trace}(\Sigma^*(z))$). Further note that other types of sensor selection problems also consider the same constraints (e.g., [13,24,23]), but with different objective functions.

In the following sections, we will discuss the complexity of the two KFSS problems and investigate approaches to address these problems.

3 Complexity of the Priori and Posteriori KFSS Problems

In this section, we show that the priori and posteriori KFSS problems are both NP-hard. To do this, we will relate them to the problems described below.

Problem 3 *Given a matrix $A \in \mathbb{R}^{n\times n}$, the problem of finding a diagonal matrix $M \in \mathbb{R}^{n\times n}$ with the fewest nonzero elements such that the pair (A, M) is controllable (resp. stabilizable, detectable) is referred to as the minimum controllability (resp. minimum stabilizability, minimum detectability) problem.*

Theorem 1 *The priori KFSS problem and the posteriori KFSS problem are NP-hard.*

PROOF. We first give a reduction from the minimum detectability problem to the priori KFSS problem (resp. posteriori KFSS problem). Given $A \in \mathbb{R}^{n\times n}$ for the minimum detectability problem and some $p \in \{1, \ldots, n\}$, the instance for the corresponding priori KFSS problem (resp. posteriori KFSS problem) with parameter p is the system matrix A, the set $\mathcal{Q}$ of n sensors with the measurement matrix $C = I_{n\times n}$, the system noise covariance matrix $W = I_{n\times n}$, the sensor noise covariance matrix $V = I_{n\times n}$, the cost vector $b = [1 \cdots 1]^T$ and the budget $B = p$.[4] Suppose there is an algorithm $\mathcal{A}$ that outputs a sensor selection vector z that minimizes $\text{trace}(\Sigma(z))$ (resp. $\text{trace}(\Sigma^*(z))$) and satisfies $b^T z \leq B$, or outputs a flag if there is no feasible sensor selection. If the output of algorithm $\mathcal{A}$ is a subset of sensors (rather than the flag), by Lemma 1, we know that the solution to the minimum detectability problem (i.e., the minimum number of nonzero entries of the diagonal matrix $M \in \mathbb{R}^{n\times n}$ such that (A, M) is detectable) is at most p. In order to solve the minimum detectability problem, we need to call algorithm $\mathcal{A}$ at most n times (i.e., increase p from 1 to n). Thus, if the minimum detectability problem is NP-hard, then the priori KFSS problem (resp. posteriori KFSS problem) is also NP-hard.

The NP-hardness of the minimum detectability problem follows from the proof of NP-hardness of the minimum controllability problem in [25]. Specifically, given $n_1, n_2 \in \mathbb{Z}_{\geq 1}$ and a collection $\mathcal{C}$ of n_1 nonempty subsets of $\{1, \cdots, n_2\}$, let $A(\mathcal{C}) = U^{-1}\,\text{diag}(1, \cdots, n_1+n_2+1)U$, where U is some invertible matrix related to $\mathcal{C}$.[5] In [25], the author proved that $\mathcal{C}$ has a hitting set with cardinality s if and only if there exists a diagonal matrix M with no more than s nonzero entries such that $(A(\mathcal{C}), M)$ is controllable. Since the hitting set problem

[4] Note that here $s = n$ (i.e., $s_i = 1, \forall i$).

[5] Note that U is constructed based on the incidence matrix of $\mathcal{C}$; we omit the construction details and refer to the proof of Theorem 1.1 in [25].

is NP-hard, the minimum controllability problem is also NP-hard. Note that the set of eigenvalues of $A(\mathcal{C})$ is $\{1, \cdots, n_1 + n_2 + 1\}$, which are all unstable. Thus, to find a matrix M such that $(A(\mathcal{C}), M)$ is stabilizable is equivalent to finding a matrix M such that $(A(\mathcal{C}), M)$ is controllable, which implies that the minimum stabilizability problem is NP-hard. By the duality of stabilizability and detectability, the minimum detectability problem is also NP-hard, completing the proof. □

Note that the above result shows that it is NP-hard to find a feasible solution for the priori and posterior KFSS problems, even when all sensors have identical costs. The following result shows that the priori and posterior KFSS problems are still NP-hard if the system dynamics matrix A is stable (so that *all* sensor selections satisfying the cost constraint are feasible), but when the sensor costs can be arbitrary.

Theorem 2 *The priori KFSS problem and posteriori KFSS problem are NP-hard even under the additional assumption that the system dynamics matrix A is stable.*

PROOF. We show the result by giving reductions from the optimization form of the 0-1 knapsack problem [14] to the priori and posteriori KFSS problems. Throughout this proof, we assume that the measurement of each sensor is a scalar (i.e., $s_i = 1, \forall i$). For the steady state *a priori* error covariance Σ (resp. *a posteriori* error covariance Σ^*), when $A = aI_{n\times n}$ with $0 < a < 1$ being some constant (resp. when $A = 0$), $C = I_{n\times n}$, $V = 0$, and $W = \text{diag}([w_1 \cdots w_n])$, we know that Σ (resp. Σ^*) is diagonal with $\Sigma_{ii} = w_i$ (resp. $\Sigma^*_{ii} = 0$) if $z_i = 1$ and $\Sigma_{ii} = \frac{w_i}{1-a^2}$ (resp. $\Sigma^*_{ii} = w_i$) if $z_i = 0$. Thus, the reduction of the *a priori* (resp. *a posteriori*) estimation error by adding sensor i is $\Sigma_{ii}(z_i = 0) - \Sigma_{ii}(z_i = 1) = \frac{a^2}{1-a^2} w_i$ (resp. $\Sigma^*_{ii}(z_i = 0) - \Sigma^*_{ii}(z_i = 1) = w_i$), $\forall i$.

Given the number of items n, the set of values $\{\alpha_i\}$, the set of weights $\{\beta_i\}$ and the weight budget $\bar{\beta}$ for the 0-1 knapsack problem, the corresponding instance for the priori (resp. posteriori) KFSS problem is the *stable* system matrix [6] $A = \frac{1}{2} I_{n\times n}$ (resp. $A = 0$), the set $\mathcal{Q}$ of n sensors with the measurement matrix $C = I_{n\times n}$, the system noise covariance matrix $W = \text{diag}([w_1 \cdots w_n])$ with $w_i = \frac{1-a^2}{a^2}\alpha_i = 3\alpha_i$ (resp. $w_i = \alpha_i$), the sensor noise covariance matrix $V = 0$, the cost vector $b = [\beta_1 \cdots \beta_n]^T$ and the budget $B = \bar{\beta}$. Then we can see that an indicator vector z for the 0-1 knapsack problem is optimal if and only if it is optimal for the corresponding priori (resp. posteriori) KFSS problem. Since the optimization form of the 0-1 knapsack problem is NP-hard, the priori and posteriori KFSS problems are NP-hard even under the additional assumption that the matrix A is stable, completing the proof. □

[6] Note that here we take the constant $a = \frac{1}{2}$.

Remark 1 *The above results complement the few works in the literature that explicitly characterize the complexities of sensor selection problems, such as [3] where a utility-based sensor selection problem is shown to be NP-hard, and [32] where the NP-hardness of an energy metric based sensor selection problem is established.*

In the rest of this paper, we focus on the case where the pair (A, C_i) is detectable, $\forall i \in \{1, \cdots, q\}$. Note that in this case, any choice of sensors satisfying the cost constraint (except $z = \mathbf{0}$ if A is unstable) is feasible.

4 Upper Bounds on the Performance of Sensor Selection Algorithms

In this section, we study worst-case bounds on the performance of sensor selection algorithms for the priori and posteriori KFSS problems. Specifically, we consider the following ratio $r(\Sigma)$:

$$r(\Sigma) \triangleq \frac{\text{trace}(\Sigma_{\text{worst}})}{\text{trace}(\Sigma_{\text{opt}})},$$

where Σ_{opt} and Σ_{worst} are the solutions of the DARE corresponding to the optimal selection of sensors and the worst-case feasible selection, respectively, and also the ratio

$$r(\Sigma^*) \triangleq \frac{\text{trace}(\Sigma^*_{\text{worst}})}{\text{trace}(\Sigma^*_{\text{opt}})},$$

which is defined similarly. Note that for any *a priori* covariance (resp. *a posteriori* covariance) based sensor selection algorithm, the performance of that algorithm is within $r(\Sigma)$ (resp. $r(\Sigma^*)$) times the optimal performance. In other words, the quantities $r(\Sigma)$ and $r(\Sigma^*)$ characterize the 'spectrum' of the performance of all feasible selections.

Note that since it is difficult in general to obtain the analytical solution of the DARE (4) (and also the steady state *a posteriori* error covariance from equation (8)), the problem of providing bounds for the DARE solution has been studied extensively in the literature; see [18] and the references therein. However, the existing upper bounds on the DARE solution typically assume that the system is stable [19] or that the corresponding sensor information matrix is nonsingular [15,16]; the latter assumption is restrictive in the context of sensor selection. Thus, in this section, we focus on the case where the system is stable to obtain more insights into the factors that affect the performance of sensor selection algorithms.

4.1 Upper Bound for $r(\Sigma)$

We first derive an upper bound on the ratio $r(\Sigma)$ for the priori KFSS problem when the system is stable. We will be using the following results.

Lemma 2 ([16]) *For $\Sigma(z) \succeq 0$ satisfying the DARE (6) with $W \succ 0$, we have $\Sigma(z) \succeq A(W^{-1}+R(z))^{-1}A^T+W$.*

Lemma 3 ([35,10]) *For Hermitian matrices $M, N \in \mathbb{C}^{n\times n}$, we have $\lambda_n(M+N) \geq \lambda_n(M)+\lambda_n(N)$, $\lambda_1(M+N) \leq \lambda_1(M)+\lambda_1(N)$, and $\lambda_n(M)\,\mathrm{trace}(N) \leq \mathrm{trace}(MN) \leq \lambda_1(M)\,\mathrm{trace}(N)$.*

Lemma 4 ([21]) *A square matrix $A \in \mathbb{R}^{n\times n}$ is Schur stable if and only if there exists a nonsingular matrix P such that $\sigma_1(PAP^{-1}) < 1$.*

Remark 2 *Note that in the above lemma, the matrix P can be constructed by using the eigenvalues and (generalized) eigenvectors of A [21]. Thus, for any stable square matrix A, there exists some positive constant $\alpha_A \triangleq \frac{\sigma_1^2(P)}{\sigma_n^2(P)(1-\sigma_1^2(PAP^{-1}))}$ which only depends on A, where the matrix P is nonsingular and satisfies $\sigma_1(PAP^{-1}) < 1$. This constant α_A will be used to establish the performance upper bounds in the rest of this section.*

To incorporate the nature of the sensor set $\mathcal{Q}$, our results will use the sensor information matrix $R(z)$ from (6) which encapsulates both the measurement matrix $\tilde{C}$ and the sensor noise covariance matrix $\tilde{V}$ corresponding to the indicator vector z.

Theorem 3 *For a given cost vector b and budget B, let $\mathcal{R} = \{R(z)\}$ be the set of all sensor information matrices such that the constraint $b^Tz \leq B$ is satisfied. Denote $\lambda_1^{max} \triangleq \max\{\lambda_1(R) | R \in \mathcal{R}\}$. Then for the system (1) with stable A and $W \succ 0$,*

$$r(\Sigma) \leq \frac{\alpha_A(1+\lambda_1^{max}\lambda_n(W))\,\mathrm{trace}(W)}{n\sigma_n^2(A)\lambda_n(W)+(1+\lambda_1^{max}\lambda_n(W))\,\mathrm{trace}(W)}, \tag{9}$$

where α_A is some positive constant that only depends on A, as defined in Remark 2.

PROOF. We first provide an upper bound for $\mathrm{trace}(\Sigma_{\mathrm{worst}})$. Consider the case where $z = \mathbf{0}$ (i.e., no sensors are chosen). Note that since A is stable, $z = \mathbf{0}$ is feasible (in the sense of Definition 1). In this case, the DARE (4) becomes the Lyapunov equation $\Sigma(\mathbf{0}) = A\Sigma(\mathbf{0})A^T + W$. Define $\bar{\Sigma} = P\Sigma(\mathbf{0})P^T$ and $\bar{W} = PWP^T$, where P is nonsingular and satisfies $\sigma_1(PAP^{-1}) < 1$. Note that since the matrix A is stable, by Lemma 4, such a matrix P always exists.

Let $D = PAP^{-1}$. Then we get $\bar{\Sigma} = D\bar{\Sigma}D^T + \bar{W}$. By Lemma 3, we know that $\mathrm{trace}(D\bar{\Sigma}D^T) = \mathrm{trace}(D^TD\bar{\Sigma}) \leq \sigma_1^2(D)\,\mathrm{trace}(\bar{\Sigma})$ and thus $\mathrm{trace}(\bar{\Sigma}) \leq \frac{\mathrm{trace}(\bar{W})}{1-\sigma_1^2(D)}$. Since $W, \Sigma \succeq 0$ and the matrix P^TP is symmetric, by Lemma 3, we know that $\mathrm{trace}(\bar{\Sigma}) = \mathrm{trace}(P^TP\Sigma(\mathbf{0})) \geq \sigma_n^2(P)\,\mathrm{trace}(\Sigma(\mathbf{0}))$ and $\mathrm{trace}(\bar{W}) = \mathrm{trace}(P^TPW) \leq \sigma_1^2(P)\,\mathrm{trace}(W)$. Combining the above analysis, we obtain

$$\begin{aligned}\mathrm{trace}(\Sigma_{\mathrm{worst}}) &\leq \mathrm{trace}(\Sigma(\mathbf{0}))\\ &\leq \frac{\sigma_1^2(P)}{\sigma_n^2(P)}\frac{\mathrm{trace}(W)}{1-\sigma_1^2(D)} = \alpha_A\,\mathrm{trace}(W),\end{aligned}$$

where the last equality is due to Lemma 4 and Remark 2.

Next we derive a lower bound for $\mathrm{trace}(\Sigma_{\mathrm{opt}})$. Specifically, for any given z, we have

$$\begin{aligned}\mathrm{trace}(\Sigma(z)) &\geq \mathrm{trace}(A(W^{-1}+R(z))^{-1}A^T+W) \quad (10)\\ &\geq \lambda_n(A^TA)\,\mathrm{trace}((W^{-1}+R(z))^{-1})\\ &\quad + \mathrm{trace}(W) \quad (11)\\ &= \sigma_n^2(A)\sum_{i=1}^{n}\frac{1}{\lambda_i(W^{-1}+R(z))} + \mathrm{trace}(W)\\ &\geq \frac{n\sigma_n^2(A)}{\lambda_1(W^{-1}+R(z))} + \mathrm{trace}(W)\\ &\geq \frac{n\sigma_n^2(A)}{\lambda_1(W^{-1})+\lambda_1(R(z))} + \mathrm{trace}(W) \quad (12)\\ &\geq \frac{n\sigma_n^2(A)}{\frac{1}{\lambda_n(W)}+\lambda_1^{\max}} + \mathrm{trace}(W).\end{aligned}$$

Note that inequality (10) is due to Lemma 2 and inequalities (11) and (12) are due to Lemma 3. Further note that the derived lower bound for $\mathrm{trace}(\Sigma(z))$ holds for any sensor selection z and thus holds for $\mathrm{trace}(\Sigma_{\mathrm{opt}})$.

The result follows by combining the upper bound for $\mathrm{trace}(\Sigma_{\mathrm{worst}})$ and the lower bound for $\mathrm{trace}(\Sigma_{\mathrm{opt}})$. □

The above result also yields a simpler upper bound for $r(\Sigma)$ which highlights the role of the system dynamics matrix A.

Corollary 1 *If the given system (1) is stable, there exists a constant α_A (given by Remark 2) which only depends on the matrix A such that $r(\Sigma) \leq \alpha_A$. Furthermore, if the matrix A is stable and normal (i.e., $A^TA = AA^T$), then $r(\Sigma) \leq \frac{1}{1-\lambda_1^2(A)}$.*

PROOF. The proof of the first part (i.e., $r(\Sigma) \leq \alpha_A$) immediately follows by noting that the denominator in (9) is lower bounded by $(1+\lambda_1^{\max}\lambda_n(W))\,\mathrm{trace}(W)$. When A is normal, the set of singular values of A coincides with its eigenvalues [10] (i.e., $\sigma_i(A) = |\lambda_i(A)|, \forall i$). Since A is stable, we know that $\sigma_1(A) = |\lambda_1(A)| < 1$. Thus, in Lemma 4, we can choose the transformation matrix P to be the identity matrix (i.e., $P = I$). Then by Remark 2, we have $r(\Sigma) \leq \alpha_A = \frac{1}{1-\lambda_1^2(A)}$, completing the proof. $\square$

Remark 3 *Note that for fixed* A *and* W*, the upper bound of* $r(\Sigma)$ *in (9) approaches* α_A *as* λ_1^{max} *gets bigger. In other words, when the measurement (from the past time-step) is more accurate, the worst-case difference among all feasible sensor selection algorithms is mainly determined by the system dynamics. Since there exists an upper bound for* $r(\Sigma)$ *which only depends on the system dynamics matrix* A*, no sensor selection algorithm will provide arbitrarily bad performance as long as* A *is well conditioned, regardless of the statistics of the noise processes and the nature of the sensor set* $\mathcal{Q}$*. In particular, if* $A = 0$*, then the state* $x[k+1]$ *in (1) is uncorrelated with* $x[k]$*, and thus measurements of the current state are not useful in predicting the next state. This is corroborated by the fact that* $r(\Sigma) = 1$ *in this case.*

4.2 Upper Bound for $r(\Sigma^*)$

Next, we provide an upper bound on the ratio $r(\Sigma^*)$ for the posteriori KFSS problem when the system is stable. We will use the following result.

Lemma 5 ([10]) *For matrices* $M, N \in \mathbb{S}_+^n$*, if* $M \succeq N$*, we have* $M^{-1} \preceq N^{-1}$.

Theorem 4 *For given cost vector* b *and budget* B*, let* $\mathcal{R} = \{R(z)\}$ *be the set of all sensor information matrices such that the constraint* $b^T z \leq B$ *is satisfied. Denote* $\lambda_1^{max} \triangleq \max\{\lambda_1(R) | R \in \mathcal{R}\}$*. Then for the system (1) with stable* A *and* $W \succ 0$*,*

$$r(\Sigma^*) \leq \alpha_A \left(\frac{\lambda_1(W)}{\lambda_n(W)} + \lambda_1^{max} \lambda_1(W) \right), \tag{13}$$

where α_A *is some positive constant that only depends on* A*, as defined in Remark 2.*

PROOF. We first give an upper bound for $\text{trace}(\Sigma^*_{\text{worst}})$. Since $R(z) \succeq 0, \forall z$, by Lemma 5 and equation (8), we know that $\Sigma^*(z) \preceq \Sigma(z), \forall z$. Thus, a simple upper bound for $\text{trace}(\Sigma^*_{\text{worst}})$ is $\text{trace}(\Sigma^*_{\text{worst}}) \leq \text{trace}(\Sigma_{\text{worst}}) \leq \alpha_A \, \text{trace}(W) \leq n \alpha_A \lambda_1(W)$, where α_A is defined in Remark 2.

Next we give a lower bound for $\text{trace}(\Sigma^*_{\text{opt}})$. For convenience, define the following notation:

$$X_1(z) \triangleq (A(W^{-1} + R(z))^{-1} A^T + W)^{-1} + R(z),$$
$$X_2(z) \triangleq A(W^{-1} + R(z))^{-1} A^T + W.$$

Note that $X_2(z)$ is the matrix lower bound for $\Sigma(z)$ given in Lemma 2 and $X_1(z) = X_2^{-1}(z) + R(z)$. Thus, by Lemma 5 and equation (8), we have $\Sigma^*(z) = (\Sigma^{-1}(z) + R(z))^{-1} \succeq (X_2^{-1}(z) + R(z))^{-1} = X_1^{-1}(z)$. Moreover, it is easy to see that $X_2(z) \succeq W$ and thus $\lambda_n(X_2(z)) \geq \lambda_n(W)$.

Then for any given z, we have

$$\begin{aligned} \text{trace}(\Sigma^*(z)) &\geq \text{trace}(X_1^{-1}(z)) \\ &= \sum_{i=1}^{n} \frac{1}{\lambda_i(X_1(z))} \geq \frac{n}{\lambda_1(X_1(z))} \\ &\geq \frac{n}{\lambda_1(X_2^{-1}(z)) + \lambda_1(R(z))} \\ &\geq \frac{n}{\frac{1}{\lambda_n(X_2(z))} + \lambda_1^{\max}} \geq \frac{n}{\frac{1}{\lambda_n(W)} + \lambda_1^{\max}}. \end{aligned} \tag{14}$$

Note that inequality (14) is due to Lemma 3. Since the above lower bound holds for any sensor selection z, it also holds for $\text{trace}(\Sigma^*_{\text{opt}})$.

The result follows by combining the upper bound for $\text{trace}(\Sigma^*_{\text{worst}})$ and the lower bound for $\text{trace}(\Sigma^*_{\text{opt}})$. $\square$

Remark 4 *As argued in Remark 3, when the system is stable,* $r(\Sigma)$ *can be upper bounded by a constant which only depends on the system matrix* A*. However, the above result suggests that* $r(\Sigma^*)$ *depends on both the system noise covariance matrix* W *and the achievable 'quality' of measurements (which is characterized by* λ_1^{max}*). In particular, when* $C = I_{n \times n}, V = \bar{v} I_{n \times n}$ *and* $W = \bar{w} I_{n \times n}$*, where* $\bar{v}, \bar{w} > 0$ *are some constants, we have* $\lambda_1(W) = \lambda_n(W) = \bar{w}$*,* $\lambda_1^{max} = \frac{1}{\bar{v}}$ *and* $r(\Sigma^*) \leq \alpha_A(1 + \frac{\bar{w}}{\bar{v}})$*; thus, for a fixed matrix* A*, the worst-case difference among all feasible sensor selection algorithms becomes smaller if the system noise gets smaller (i.e.,* $\bar{w}$ *gets smaller) or the measurements become more inaccurate (i.e.,* $\bar{v}$ *gets bigger).*

5 Greedy Algorithms

In this section, we explore simple greedy algorithms to solve the priori and posteriori KFSS problems, given as Algorithm 1 and Algorithm 2, respectively. We focus on the case where the cost vector $b = [1 \cdots 1]^T$ and the budget $B = p$ for some $p \in \{1, \cdots, q\}$ (i.e., our goal is to choose p sensors out of the total q sensors to optimize the performance of the Kalman filter). In other words, an indicator vector z is valid if $z \in \mathcal{Z}_p$ where $\mathcal{Z}_p$ is defined to be the set of indicator vectors with no more than p nonzero elements. The basic idea of the greedy algorithms is to iteratively pick sensors that provide the largest incremental decrease in the steady state (*a priori* or *a posteriori*) error covariance.

It is easy to show that if the influence of each sensor is "separable", then greedy algorithms output the optimal solution. Specifically, if the system matrices A and C and

Algorithm 1 *A Priori* Covariance based Greedy Algorithm

Input: System dynamics matrix A, set of all sensors $\mathcal{Q}$, noise covariances W and V, and constant p
Output: A set $\mathcal{S}$ of chosen sensors

$k \leftarrow 0$, $\mathcal{S} \leftarrow \emptyset$
for $k \le p$ **do**
  for $i \in \mathcal{Q} \cap \bar{\mathcal{S}}$ **do**
    Calculate $\text{trace}(\Sigma_{i,\mathcal{S}}) \triangleq \text{trace}(\Sigma(\mathcal{S} \cup \{i\}))$
  end for
  Choose j with $\text{trace}(\Sigma_{j,\mathcal{S}}) = \min_i \text{trace}(\Sigma_{i,\mathcal{S}})$
  $\mathcal{S} \leftarrow \mathcal{S} \cup \{j\}$, $\mathcal{Q} \leftarrow \mathcal{Q} \setminus \{j\}$, $k \leftarrow k+1$
end for

Algorithm 2 *A Posteriori* Covariance based Greedy Algorithm

Input: System dynamics matrix A, set of all sensors $\mathcal{Q}$, noise covariances W and V, and constant p
Output: A set $\mathcal{S}$ of chosen sensors

$k \leftarrow 0$, $\mathcal{S} \leftarrow \emptyset$
for $k \le p$ **do**
  for $i \in \mathcal{Q} \cap \bar{\mathcal{S}}$ **do**
    Calculate $\text{trace}(\Sigma^*_{i,\mathcal{S}}) \triangleq \text{trace}(\Sigma^*(\mathcal{S} \cup \{i\}))$
  end for
  Choose j with $\text{trace}(\Sigma^*_{j,\mathcal{S}}) = \min_i \text{trace}(\Sigma^*_{i,\mathcal{S}})$
  $\mathcal{S} \leftarrow \mathcal{S} \cup \{j\}$, $\mathcal{Q} \leftarrow \mathcal{Q} \setminus \{j\}$, $k \leftarrow k+1$
end for

the covariance matrices W and V are of the following form:

$$A = \text{diag}(A_1, \cdots, A_q), W = \text{diag}(W_1, \cdots, W_q),$$
$$C = \text{diag}(C_1^d, \cdots, C_q^d), V = \text{diag}(V_1, \cdots, V_q),$$

where $A_i, W_i \in \mathbb{R}^{s'_i \times s'_i}$, $C_i^d \in \mathbb{R}^{s_i \times s'_i}$, and $V_i \in \mathbb{R}^{s_i \times s_i}, \forall i$, then Algorithm 1 and Algorithm 2 are optimal for the priori and posteriori KFSS problems, respectively. Note that $C_i \in \mathbb{R}^{s_i \times n}$ is the set of rows of the matrix C corresponding to sensor i and C_i^d contains s'_i columns of C_i. Further note that $\sum_i s'_i = n$.

In the rest of this section, we will show that the greedy algorithms are optimal (with respect to the corresponding KFSS problem) for another class of systems where the set of information matrices is totally ordered. However, for general systems, we provide a negative result showing that the trace of the steady state *a priori* error covariance and *a posteriori* error covariance (and other related metrics) do not satisfy certain modularity properties in general, which precludes the direct application of classical results on submodular function optimization.

5.1 Optimality of Greedy Algorithms for a Class of Systems

First note that when the sensor noises are uncorrelated (i.e., $\mathbb{E}\left[v_i[k_1](v_j[k_2])^T\right] = 0, \forall i \neq j, k_1, k_2$), then the sensor noise covariance matrix V is block diagonal; in this case, let $V = \text{diag}(V_1, \cdots, V_q)$ where $V_i = \mathbb{E}\left[v_i[k](v_i[k])^T\right]$. Then the sensor information matrix $R(z)$ can be written as $R(z) = \sum_{i=1}^q z_i R_i$ where $R_i \triangleq C_i^T V_i^{-1} C_i$ is the sensor information matrix associated with sensor i. The following result characterizes the relationship between the partial orders on information matrices to the partial orders on the corresponding *a priori* and *a posteriori* error covariances.

Lemma 6 ([11,34]) *For two selections of sensors z and z', if $R(z) \succeq R(z')$, then we have $\Sigma(z) \preceq \Sigma(z')$ and $\Sigma^*(z) \preceq \Sigma^*(z')$.*

In other words, when $R(z) \succeq R(z')$, then $\text{trace}(\Sigma(z)) \le \text{trace}(\Sigma(z'))$ and $\text{trace}(\Sigma^*(z)) \le \text{trace}(\Sigma^*(z'))$, and thus the sensor selection associated with z is better than the one associated with z'. The following result shows that when the sensor noises are uncorrelated and the set of information matrices $\{R_i\}$ is totally ordered, then Algorithm 1 and Algorithm 2 are optimal (with respect to the corresponding KFSS problems).

Proposition 1 *If the sensor noises are uncorrelated and the set of information matrices $\{R_i\}$ is totally ordered with respect to the order relation of positive semidefiniteness, then the optimal solution of the priori and posteriori KFSS problems with $b = [1 \cdots 1]^T$ and $B = p$ is the set of sensors $\mathcal{P} \subseteq \mathcal{Q}$ such that $|\mathcal{P}| = p$ and $R_i \succeq R_j, \forall i \in \mathcal{P}, j \in \mathcal{Q} \setminus \mathcal{P}$. Furthermore, both Algorithm 1 and Algorithm 2 output this optimal set of sensors.*

PROOF. We first show that the optimal solution of the priori and posteriori KFSS problems is the specified set of sensors $\mathcal{P}$. Denote $z_{\mathcal{P}}$ as the indicator vector associated with set $\mathcal{P}$. Since the set of information matrices $\{R_i\}$ is totally ordered, we have $R(z_{\mathcal{P}}) \succeq R(z), \forall z \in \mathcal{Z}_p$. Thus, by Lemma 6, we know that $\text{trace}(\Sigma(z_{\mathcal{P}})) \le \text{trace}(\Sigma(z))$ and $\text{trace}(\Sigma^*(z_{\mathcal{P}})) \le \text{trace}(\Sigma^*(z)), \forall z \in \mathcal{Z}_p$, which implies that the set of sensors $\mathcal{P}$ is the optimal solution of the priori and posteriori KFSS problems.

Next we show by induction that the output of Algorithm 1 and Algorithm 2 is the set of sensors $\mathcal{P}$. Without loss of generality, let $R_1 \succeq \cdots \succeq R_q$. Then $\mathcal{P} = \{1, \cdots, p\}$. By Lemma 6, we know that $\Sigma(\{1\}) \preceq \Sigma(\{i\}), \forall i$ (resp. $\Sigma^*(\{1\}) \preceq \Sigma^*(\{i\}), \forall i$); thus, after the first loop, the output of Algorithm 1 (resp. Algorithm 2) is to choose the first sensor. Assume that Algorithm 1 (resp. Algorithm 2) outputs the first k sensors after the k-th loop. By Lemma 6, we know that $\Sigma(\{1, \cdots, k, k+1\}) \preceq \Sigma(\{1, \cdots, k, i\}), \forall i > k$ (resp. $\Sigma^*(\{1, \cdots, k, k+1\}) \preceq \Sigma^*(\{1, \cdots, k, i\}), \forall i > k$); thus, the output of Algorithm 1 (resp. Algorithm 2) is $\{1, \cdots, k, k+1\}$ after the $(k+1)$-th loop. Thus, after the p-th loop, the final output of both algorithms is the set of sensors $\mathcal{P}$, completing the proof. □

Remark 5 *Note that in [34], the authors showed that for a given p, under the same conditions as in Proposition 1, the optimal solution of the posteriori KFSS problem is the set of p sensors with 'largest' information matrices. However, their algorithm is a special case of Algorithm 2 and they do not consider the priori KFSS problem.*

5.2 Lack of Submodularity of the Cost Functions

Outside of the special case discussed in Proposition 1, there are few tools available to give performance guarantees on greedy algorithms. One such tool is the concept of submodularity, which has been used in the analysis of greedy algorithms for the sensor scheduling problem, as mentioned in the beginning of this section. Specifically, in order to solve the problem of *maximizing* a submodular cost function (which also satisfies certain other properties), one can obtain an approximation of the optimal solution within a factor of $1-\frac{1}{e}$ by using a greedy algorithm; see [22] for a comprehensive discussion. The definition of submodularity is as follows.

Definition 2 (Submodularity) *Consider a set E and a set function $f : 2^E \to \mathbb{R}$. The set function f is submodular if for every $X, Y \subseteq E$ with $X \subseteq Y$ and every $x \in E \setminus Y$, $f(X \cup \{x\}) - f(X) \geq f(Y \cup \{x\}) - f(Y)$, and is supermodular if $-f$ is submodular.*

For the priori or posteriori covariance matrices induced by the set of indicator vectors in $\mathcal{Z}_p$, we will consider the problem of *maximizing* three different performance metrics: $F_1(\cdot) = -\operatorname{trace}(\cdot)$, $F_2(\cdot) = -\log\det(\cdot)$ and $F_3(\cdot) = -\lambda_1(\cdot)$, where F_2 captures the volume of the confidence ellipsoid and F_3 captures the worst-case error covariance. Note that maximizing F_1 is equivalent to minimizing $-F_1$ as in the priori and posteriori KFSS problems. In [12], the authors showed that the metric F_2 is submodular for the single-step sensor scheduling problem while F_1 and F_3 are neither submodular nor supermodular. One question of interest is whether any of these metrics is submodular or supermodular for the priori and posteriori KFSS problems. However, the following counterexamples show that these metrics are neither supermodular nor submodular in general.

Example 1 (Lack of submodularity of $F_i(\Sigma)$) *For metric $F_i(\Sigma)$ and two sets of sensors X and Y, let the change of utility by adding Y to X be $\Delta_{F_i}(Y|X)$, i.e., $\Delta_{F_i}(Y|X) = F_i(\Sigma(X \cup Y)) - F_i(\Sigma(X))$. Consider an instance of the priori KFSS problem with* $A = \begin{bmatrix} 0.3 & 0.4 \\ 0.2 & 0.6 \end{bmatrix}$,

$C = \begin{bmatrix} 1 & 0.5 & 0.7 & 0 & 0.3 \\ 0 & 0.5 & 0.3 & 0.7 & 0.7 \end{bmatrix}^T$, $W = I_{2\times 2}, V = I_{5\times 5}$, $s_i = 1, \forall i$, $b = [1 \cdots 1]^T$ *and* $B = 4$. *Note that A is stable and thus all selections of sensors are feasible. One can check that $\Delta_{F_i}(\{1\}|\{2,3\}) < \Delta_{F_i}(\{1\}|\{2,3,4\})$ and $\Delta_{F_i}(\{1\}|\{2\}) > \Delta_{F_i}(\{1\}|\{2,3\})$, $i \in \{1,2,3\}$, which contradicts the submodularity and supermodularity of the corresponding metrics, respectively.*

Example 2 (Lack of submodularity of $F_i(\Sigma^*)$) *For metric $F_i(\Sigma^*)$ and two sets of sensors X and Y, let the change of utility by adding Y to X be $\Delta^*_{F_i}(Y|X)$, i.e., $\Delta^*_{F_i}(Y|X) = F_i(\Sigma^*(X \cup Y)) - F_i(\Sigma^*(X))$. Consider the same instance in Example 1 for the posteriori KFSS problem. One can check that $\Delta^*_{F_i}(\{2\}|\{1,3\}) > \Delta^*_{F_i}(\{2\}|\{1,3,4\})$, $i \in \{1,2,3\}$, which contradicts the supermodularity of the corresponding metrics. Moreover, $\Delta^*_{F_i}(\{5\}|\{2,3\}) < \Delta^*_{F_i}(\{5\}|\{1,2,3\}), i \in \{1,3\}$, which contradicts the submodularity of the corresponding metrics. Finally, consider another instance as follows:*

$$A = \begin{bmatrix} 0.4 & 0.9 & 0.8 & 0.9 & 1.1 \\ 0.2 & 0.1 & -0.4 & 0.4 & -0.5 \\ -0.1 & 2.1 & 0.7 & 1.7 & 1.3 \\ -0.9 & -0.8 & -0.8 & -1.3 & -1 \\ 0.4 & -1.7 & 0.2 & -1.3 & -0.3 \end{bmatrix}, W = \begin{bmatrix} 6.3 & 1.6 & 1.6 & 0.9 & 0.6 \\ 1.6 & 9.2 & 1.1 & 1.6 & 1.2 \\ 1.6 & 1.1 & 6.6 & 1.5 & 1.8 \\ 0.9 & 1.6 & 1.5 & 5.5 & 1 \\ 0.6 & 1.2 & 1.8 & 1 & 5.9 \end{bmatrix},$$

$$C = \begin{bmatrix} -3 & 10 & 7 & -3 & 2 \\ -1 & -7 & 3 & -6 & 0 \\ -8 & 3 & 3 & 3 & 5 \\ -2 & 5 & -3 & 2 & -10 \\ 0 & 4 & -3 & 5 & 1 \end{bmatrix}, \qquad V = \begin{bmatrix} 2.3 & 0.2 & 0 & 0 & 0 \\ 0.2 & 0.7 & 0 & 0 & 0 \\ 0 & 0 & 2.3 & 0 & 0 \\ 0 & 0 & 0 & 2.5 & 0 \\ 0 & 0 & 0 & 0 & 2.3 \end{bmatrix},$$

$s_i = 1, \forall i$, $b = [1 \cdots 1]^T$ *and* $B = 4$. *Note that A is stable. One can check that $\Delta^*_{F_2}(\{1\}|\{2\}) < \Delta^*_{F_2}(\{1\}|\{2,3,4\})$, which contradicts the submodularity of $F_2(\Sigma^*)$.*

The above negative results imply that one may not be able to use classical results from combinatorial optimization to analyze Algorithm 1 and Algorithm 2; despite this, our simulations in Section 6 show that these greedy algorithms perform well in practice.

6 Simulation

In this section, we provide simulation results for the performance of the *a priori* covariance based greedy algorithm (Algorithm 1) and the *a posteriori* covariance based greedy algorithm (Algorithm 2).

In order to illustrate the performance of the greedy algorithms considered in this paper, we will compare them with the following sensor selection strategies:

- *Sparse optimization (abbr.: SparseOpt)* approach for the priori KFSS problem from [27,8], where sparsity is achieved by adding a penalty function on the columns of the gain matrix. Since there is in general no systematic method to choose the weight of the penalty function, we fix the weight in the simulations and select the sensors corresponding to the columns of the gain matrix with largest l_1 norm.
- *Priori convex relaxation (abbr.: PriConRe)* approach for the priori KFSS problem from [34]. Note that we

modify the algorithm in [34] for packet-dropping channels to handle the case of reliable channels (corresponding to the priori KFSS problem considered in this paper).

- *Posteriori convex relaxation (abbr.: PostConRe)* approach for the posteriori KFSS problem from [34].
- *A random selection (abbr.: Random)* of sensors for both the priori and posteriori KFSS problems. We use this as a benchmark.

We randomly generate 300 systems all having dimension 5 (i.e., $n = 5$). For each system, the goal is to choose 5 sensors out of a total of 20 (i.e., $p = 5$, $q = 20$, $b = [1 \cdots 1]^T$ and $B = p$) and the measurement of each sensor is a scalar (i.e., $s_i = 1, \forall i$). For each system, the system matrix A is unstable and the pair (A, C_i) is detectable, $\forall i \in \{1, \cdots, q\}$. The results are summarized in Table 1. From Table 1a, we see that for the priori KFSS problem, the priori convex relaxation approach from [34] provides a set of sensors with smaller trace than the other algorithms in a plurality of cases. However, this algorithm also exhibits larger variance than the other algorithms (with high worst case deviation from optimality). On average, the sparse optimization approach from [27,8] outperforms all the other algorithms, and this approach has the smallest variance. As illustrated in Table 1a, the greedy algorithm exhibits comparable average performance to the other algorithms. From Table 1b, we see that Algorithm 2 and the posteriori convex relaxation approach from [34] each outperforms the other in a comparable number of cases. However, once again, the greedy algorithm provides more consistent results with better average performance. To summarize, the greedy algorithms have comparable performance with the other sensor selection algorithms in general. Moreover, as we have argued in the Introduction, compared to the priori and posteriori convex relaxation approaches in [34], the greedy algorithms can be applied to a more general class of systems where the sensor noise covariance matrix V is not necessarily block-diagonal.

To compare the complexity of the previous algorithms, note that the complexity of solving the DARE is $O(n^3)$, where n is the number of states [7]. If we aim to choose p sensors from a set of q sensors, then the complexity of Algorithm 1 is $O(pqn^3)$. Since we can obtain Σ^* from Σ by equation (8), the complexity of Algorithm 2 is also $O(pqn^3)$. As argued in [8], when the weight of the sparsity penalty function is *fixed*, the complexity of the sparse optimization approach is $O((n + s)^6)$ by using the interior point method [7] (recall that $s = \sum_{i=1}^{q} s_i$ is the dimension of the combined output y); however, the process of choosing an appropriate weight for the sparsity penalty function (in order to obtain the desired level of sparsity) requires additional computation. Moreover, the complexities of the priori and the posteriori convex relaxation approaches from [34] are both $O((n+s)^6)$ by using the interior point method [4]. Thus, the complexity of the greedy algorithms is lower than those of the other SDP based approaches. Figure 1 shows simulations that support this conclusion. Note that the simulations are conducted on a typical 2.4-GHz personal computer, the goal is to choose 5 sensors out of 20 (i.e., $p = 5$, $q = 20$, $b = [1 \cdots 1]^T$ and $B = p$) and we take the measurement of each sensor to be a scalar (i.e., $s_i = 1, \forall i$). Further note that we found our solver ran out of memory when the number of states n exceeded 50 for the SDP based approaches.

Table 1
Performance comparison of different algorithms over 300 randomly generated systems with scalar measurements and diagonal V. For Algorithm $\mathcal{A}$, the table presents the **average**, **standard deviation** and **worst-case values** of $\frac{\text{trace}(\Sigma_{\mathcal{A}})}{\text{trace}(\Sigma_{\text{opt}})}$ (in 1a) and $\frac{\text{trace}(\Sigma^*_{\mathcal{A}})}{\text{trace}(\Sigma^*_{\text{opt}})}$ (in 1b) over the 300 runs. The last column presents the percentage of the 300 systems for which the corresponding algorithm **outperforms all the other algorithms**.

(a) Priori KFSS Problem

	Average	SD	Worst	Outperforms Other Algorithms
Algorithm 1	1.4	1.3	13.5	11.3%
PriConRe	1.5	2.7	41.3	47.0%
SparseOpt	1.1	0.3	4.7	41.7%
Random	11.3	26.0	307.3	0%

(b) Posteriori KFSS Problem

	Average	SD	Worst	Outperforms Other Algorithms
Algorithm 2	4.6	3.9	37.2	49.7%
PostConRe	17.8	52.4	600.7	50.3%
Random	55.0	46.5	456.7	0%

7 Conclusion

In this paper, we studied the priori and posteriori KFSS problems for linear dynamical systems. We showed that these problems are both NP-hard (even under the additional assumption that the system is stable). We also provided upper bounds for the performance of the worst-case selection of sensors and highlighted the factors that dominate the worst-case performance. Then we studied *a priori* covariance based and *a posteriori* covariance based greedy algorithms for sensor selection. We showed that these algorithms are optimal for a class of systems where the set of information matrices is totally ordered. For

[7] In [8], the authors present a customized algorithm to reduce the complexity to $O(n^6)$.

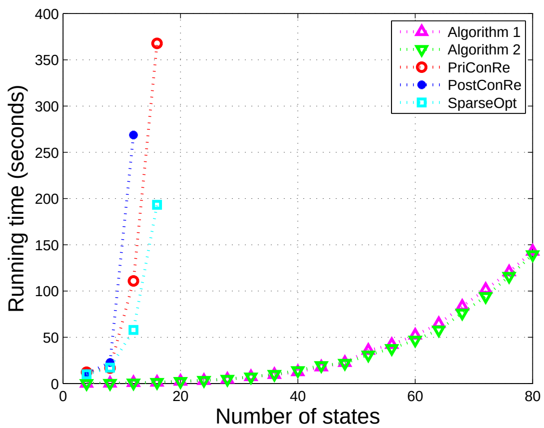

Figure 1. Complexity comparison of different algorithms. The x-axis is the number of states n and the y-axis is the running time of the algorithm.

general systems, we provided a negative result showing that the corresponding cost functions are neither supermodular nor submodular; however, simulations indicate that these algorithms perform well in practice. Further study on determining provable bounds on greedy algorithms for general systems and identifying (nontrivial) conditions under which the sensor selection problems become submodular are of interest.

Acknowledgements

The authors thank Stephen L. Smith and Jianghai Hu for helpful discussions during the course of obtaining the results in this paper.

References

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

[2] M.-A. Belabbas. Geometric methods for optimal sensor design. *Proc. of the Royal Society*, 472(2185), 2016.

[3] F. Bian, D. Kempe, and R. Govindan. Utility based sensor selection. In *Proc. of the 5th International Conference on Information Processing in Sensor Networks*, pages 11–18, 2006.

[4] S. Boyd and L. Vandenberghe. *Convex Optimization.* Cambridge University Press, 2004.

[5] D. E. Catlin. *Estimation, Control, and the Discrete Kalman Filter*, volume 71. Applied Mathmatical Sciences, 1989.

[6] S. P. Chepuri and G. Leus. Sparsity-promoting sensor selection for non-linear measurement models. *IEEE Transactions on Signal Processing*, 63(3):684–698, 2015.

[7] B. N. Datta. *Numerical Methods for Linear Control Systems: Design and Analysis.* Academic Press, 2004.

[8] N. K. Dhingra, M. R. Jovanović, and Z. Q. Luo. An ADMM algorithm for optimal sensor and actuator selection. In *Proc. of 53rd IEEE Conference on Decision and Control*, pages 4039–4044, 2014.

[9] 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*, 42(2):251–260, 2006.

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

[11] M. F. Huber. Optimal pruning for multi-step sensor scheduling. *IEEE Transactions on Automatic Control*, 57(5):1338–1343, 2012.

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

[13] S. Joshi and S. Boyd. Sensor selection via convex optimization. *IEEE Transactions on Signal Processing*, 57(2):451–462, 2009.

[14] H. Kellerer, U. Pferschy, and D. Pisinger. *Knapsack problems.* Springer Science & Business Media, 2004.

[15] N. Komaroff. Upper bounds for the solution of the discrete Riccati equation. *IEEE Transactions on Automatic Control*, 37(9):1370–1373, 1992.

[16] N. Komaroff. Iterative matrix bounds and computational solutions to the discrete algebraic Riccati equation. *IEEE Transactions on Automatic Control*, 39(8):1676–1678, 1994.

[17] A. Krause, A. Singh, and C. Guestrin. Near-optimal sensor placements in Gaussian processes: Theory, efficient algorithms and empirical studies. *The Journal of Machine Learning Research*, 9:235–284, 2008.

[18] W. H. Kwon, Y. S. Moon, and S. C. Ahn. Bounds in algebraic Riccati and Lyapunov equations: a survey and some new results. *International Journal of Control*, 64(3):377–389, 1996.

[19] C. H. Lee. Upper matrix bound of the solution for the discrete Riccati equation. *IEEE Transactions on Automatic Control*, 42(6):840–842, 1997.

[20] F. Lin, M. Fardad, and M. R. Jovanovic. Design of optimal sparse feedback gains via the alternating direction method of multipliers. *IEEE Transactions on Automatic Control*, 58(9):2426–2431, 2013.

[21] J. Liu and J. Zhang. The open question of the relation between square matrix's eigenvalues and its similarity matrix's singular values in linear discrete system. *International Journal of Control, Automation and Systems*, 9(6):1235–1241, 2011.

[22] L. Lovász. Submodular functions and convexity. In *Mathematical Programming The State of the Art*, pages 235–257. Springer, 1983.

[23] Y. Mo, R. Ambrosino, and B. Sinopoli. Sensor selection strategies for state estimation in energy constrained wireless sensor networks. *Automatica*, 47(7):1330–1338, 2011.

[24] A. Nordio, A. Tarable, F. Dabbene, and R. Tempo. Sensor selection and precoding strategies for wireless sensor networks. *IEEE Transactions on Signal Processing*, 63(16):4411–4421, 2015.

[25] A. Olshevsky. Minimal controllability problems. *IEEE Transactions on Control of Network Systems*, 1(3):249–258, 2014.

[26] S. Pequito, S. Kar, and A. Aguiar. A structured systems approach for optimal actuator-sensor placement in linear time-invariant systems. In *Proc. of American Control Conference*, pages 6108–6113, 2013.

[27] B. Polyak, M. Khlebnikov, and P. Shcherbakov. An LMI approach to structured sparse feedback design in linear

control systems. In *Proc. of European Control Conference*, pages 833–838, 2013.

[28] M. Shamaiah, S. Banerjee, and H. Vikalo. Greedy sensor selection: Leveraging submodularity. In *Proc. of 49th IEEE Conference on Decision and Control*, pages 2572–2577, 2010.

[29] T. Summers, F. Cortesi, and J. Lygeros. On submodularity and controllability in complex dynamical networks. *IEEE Transactions on Control of Network Systems*, 3(1):91–101, 2016.

[30] V. Tzoumas, A. Jadbabaie, and G. J. Pappas. Sensor placement for optimal Kalman filtering: Fundamental limits, submodularity, and algorithms. In *Proc. of American Control Conference*, 2016. to appear.

[31] V. Tzoumas, A. Jadbabaie, and G. J. Pappas. Sensor placement for optimal Kalman filtering: Fundamental limits, submodularity, and algorithms. *arXiv: 1509.08146*, 2016.

[32] V. Tzoumas, M. A. Rahimian, G. J. Pappas, and A. Jadbabaie. Minimal actuator placement with bounds on control effort. *IEEE Transactions on Control of Network Systems*, 3(1):67–78, 2015.

[33] M. Van De Wal and B. De Jager. A review of methods for input/output selection. *Automatica*, 37(4):487–510, 2001.

[34] 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*, 63(9):2336–2348, 2015.

[35] F. Zhang and Q. Zhang. Eigenvalue inequalities for matrix product. *IEEE Transactions on Automatic Control*, 51(9):1506–1509, 2006.

[36] H. Zhang, R. Ayoub, and S. Sundaram. Sensor selection for optimal filtering of linear dynamical systems: Complexity and approximation. In *Proc. of 54th IEEE Conference on Decision and Control*, pages 5002–5007, 2015.