2016 American Control Conference (ACC)
Boston Marriott Copley Place
July 6-8, 2016. Boston, MA, USA

# Sensor Placement for Optimal Kalman Filtering: Fundamental Limits, Submodularity, and Algorithms

V. Tzoumas, A. Jadbabaie, G. J. Pappas*

***Abstract*—In this paper, we focus on sensor placement in linear dynamic estimation, where the objective is to place a small number of sensors in a system of interdependent states so to design an estimator with a desired estimation performance. In particular, we consider a linear time-variant system that is corrupted with process and measurement noise, and study how the selection of its sensors affects the estimation error of the corresponding Kalman filter over a finite observation interval. Our contributions are threefold: First, we prove that the minimum mean square error of the Kalman filter decreases only linearly as the number of sensors increases. That is, adding extra sensors so to reduce this estimation error is ineffective, a fundamental design limit. Similarly, we prove that the number of sensors grows linearly with the system's size for fixed minimum mean square error and number of output measurements over an observation interval; this is another fundamental limit, especially for systems where the system's size is large. Second, we prove that the $\log\det$ of the error covariance of the Kalman filter, which captures the volume of the corresponding confidence ellipsoid, with respect to the system's initial condition and process noise is a supermodular and non-increasing set function in the choice of the sensor set. Therefore, it exhibits the diminishing returns property. Third, we provide an efficient approximation algorithm that selects a small number sensors so to optimize the Kalman filter with respect to this estimation error —the worst-case performance guarantees of this algorithm are provided as well.**

***Index Terms*—Least-Squares Linear Estimator, Minimal Sensor Placement, Greedy Algorithms.**

## I. Introduction

In this paper, we consider a linear time-variant system corrupted with process and measurement noise. Our first goal is to study how the placement of their sensors affects the minimum mean square error of their Kalman filter over a finite observation interval [1]. Moreover, we aim to select a small number of sensors so to minimize the volume of the corresponding confidence ellipsoid of this estimation error. Thereby, this study is an important distinction in the minimal sensor placement literature [2]–[13], since the Kalman filter is the optimal linear estimator —in the minimum mean square sense— given a sensor set [14].

Our contributions are threefold:

*Fundamental limits:* First, we identify fundamental limits in the design of the Kalman filter with respect to its sensors. In particular, given any finite number of output measurements over an observation interval, we prove that the minimum mean square error of the Kalman filter decreases only linearly as the number of sensors increases. That is, adding extra sensors so to reduce this estimation error of the Kalman filter is ineffective, a fundamental design limit. Similarly, we prove that the number of sensors grows linearly with the system's size for fixed minimum mean square error; this is another fundamental limit, especially for systems where the system's size is large. Overall, our novel results quantify the trade-off between the number of sensors and that of output measurements so to achieve a specified value for the minimum mean square error.

These results are the first to characterize the effect of the sensor set on the minimum mean square error of the Kalman filter. In particular, in [6], the authors quantify only the trade-off between the total energy of the consecutive output measurements and the number of its selected sensors. Similarly, in [12], the authors consider only the maximum-likelihood estimator for the system's initial condition and only for a special class of stable linear time-invariant systems. Moreover, they consider systems that are corrupted merely with measurement noise, which is white and Gaussian. Finally, they also assume an infinite observation interval, that is, infinite number of consecutive output measurements. Nonetheless, we assume a finite observation interval and study the Kalman estimator both for the system's initial condition and for the system's state at the time of the last output measurement. In addition, we consider general linear time-variant systems that are corrupted with both process and measurement noise, of any distribution (with zero mean and finite variance). Overall, our results characterize the effect of the cardinality of the sensor set on the minimum mean square error of the Kalman filter, that is, the optimal linear estimator.

*Submodularity:* Second, we identify properties for the $\log\det$ of the error covariance of the Kalman filter, which captures the volume of the corresponding confidence ellipsoid, with respect to the system's initial condition and process noise over a finite observation interval as a sensor set function — the design of an optimal Kalman filter with respect to the system's initial condition and process noise implies the design of an optimal Kalman filter with respect to the system's state. Specifically, we prove that it is a supermodular and non-increasing set function in the choice of the sensor set.

In contrast, in [15], the authors study sensor placement for monitoring static phenomena with only spatial correlations. To this end, they prove that the mutual information between the chosen and non-chosen locations is submodular. Notwith-

*All authors are with the Department of Electrical and Systems Engineering, University of Pennsylvania, Philadelphia, PA 19104-6228 USA (email: {vtzoumas, jadbabai, pappasg}@seas.upenn.edu).

This work was supported in part by TerraSwarm, one of six centers of STARnet, a Semiconductor Research Corporation program sponsored by MARCO and DARPA, in part by AFOSR Complex Networks Program and in part by AFOSR MURI CHASE.

standing, we consider dynamic phenomena with both spatial and temporal correlations, and as a result, we characterize as submodular a richer class of estimation performance metrics. Furthermore, in the sensor scheduling literature [16], the $\log\det$ of the error covariance of the Kalman filter has been proven submodular but only for special cases of systems with zero process noise [17] and [18]. Nevertheless, we consider the presence of process noise, and prove our supermodularity result for the general case.[1]

*Algorithms:* Third, we consider the problem of sensor placement so to optimize the $\log\det$ of the error covariance of the Kalman filter with respect to the system's initial condition and process noise over a finite observation interval —henceforth, we refer to this error as $\log\det$ *error*, and to the latter problem as $\mathcal{P}_1$. Naturally, $\mathcal{P}_1$ is combinatorial, and in particular, it involves the minimization of a supermodular set function, that is, the minimum mean square error. Because the minimization of a general supermodular function is NP-hard [19], we provide efficient approximation algorithms for their general solution, along with their worst-case performance guarantees. Specifically, we provide an efficient algorithm for $\mathcal{P}_1$ that returns a sensor set that satisfies the estimation guarantee of $\mathcal{P}_1$ and has cardinality up to a multiplicative factor from the minimum cardinality sensor sets that meet the same estimation bound. Moreover, this multiplicative factor depends only logarithmically on the problem's $\mathcal{P}_1$ parameters.[2]

In contrast, the related literature has focused either on the optimization of average estimation performance metrics, such as the $\log\det$ of the error's covariance, or on heuristic algorithms that provide no worst-case performance guarantees. In particular, in [26], the authors minimize the $\log\det$ of the error's covariance matrix of the Kalman estimator for the case where there is no process noise in the system's dynamics —in contrast, in our framework we assume both process and measurement noise. Moreover, to this end they use convex relaxation techniques that provide no performance guarantees. Furthermore, in [27] and [28], the authors design an $H_2$-optimal estimation gain with a small number of non-zero columns. To this end, they also use convex relaxation techniques that provide no performance guarantees. Finally, in [29], the author designs an output matrix with a desired norm so to minimize the minimum mean square error of the corresponding Kalman estimator; nonetheless, the author does not minimize the number of selected sensors. Overall, with this paper we are the first to optimize the minimum mean square error of the Kalman filter using a small number of sensors and to provide worst-case performance guarantees.

The remainder of this paper is organized as follows. In Section II, we introduce our model, and our estimation and sensor placement framework, along with our sensor placement problems. In Section III, we provide a series of design and performance limits and characterize the properties of the Kalman estimator with respect to its sensor set; in Section IV, we prove that the $\log\det$ estimation error of the Kalman filter with respect to the system's initial condition and process noise is a supermodular and non-increasing set function in the choice of the sensor set; and in Section V, we provide efficient approximation algorithms for selecting a small number of sensors so to design an optimal Kalman filter with respect to its $\log\det$ error —the worst-case performance guarantees of these algorithms are provided as well. Finally, Section VI concludes the paper. Due to space limitations, the proofs of all of our results, as well as, the corresponding simulations, are omitted; they can be found in the full version of this paper, located at our websites.

## II. Problem Formulation

*Notation:* We denote the set of natural numbers $\{1, 2, \ldots\}$ as $\mathbb{N}$, the set of real numbers as $\mathbb{R}$, and the set $\{1, 2, \ldots, n\}$ as $[n]$, where $n \in \mathbb{N}$. Given a set $\mathcal{X}$, $|\mathcal{X}|$ is its cardinality. Matrices are represented by capital letters and vectors by lower-case letters. For a matrix $A$, $A^\top$ is its transpose and $A_{ij}$ its element located at the $i-$th row and $j-$th column. $\|A\|_2 \equiv \sqrt{A^\top A}$ is its spectral norm, and $\lambda_{\min}(A)$ and $\lambda_{\max}(A)$ its minimum and maximum eigenvalues, respectively. Moreover, if $A$ is positive semi-definite or positive definite, we write $A \succeq 0$ and $A \succ 0$, respectively. Furthermore, $I$ is the identity matrix —its dimension is inferred from the context; similarly for the zero matrix $0$. Finally, for a random variable $x \in \mathbb{R}^n$, $\mathbb{E}(x)$ is its expected value, and $\mathbb{C}(x) \equiv \mathbb{E}\left([x - \mathbb{E}(x)]\,[x - \mathbb{E}(x)]^\top\right)$ its covariance. The rest of our notation is introduced when needed.

### A. Model and Estimation Framework

For $k \geq 0$, we consider the linear time-variant system

$$\begin{aligned} x_{k+1} &= A_k x_k + w_k, \\ y_k &= C_k x_k + v_k, \end{aligned} \tag{1}$$

where $x_k \in \mathbb{R}^n$ ($n \in \mathbb{N}$) is the state vector, $y_k \in \mathbb{R}^c$ ($c \in [n]$) the output vector, $w_k$ the process noise and $v_k$ the measurement noise —without loss of generality, the input vector is assumed zero. The initial condition is $x_0$.

**Assumption 1** (For all $k \geq 0$, the initial condition, the process noise and the measurement noise are uncorrelated random variables). *$x_0$ is a random variable with covariance $\mathbb{C}(x_0) = \sigma^2 I$, where $\sigma \geq 0$. Moreover, for all $k \geq 0$, $\mathbb{C}(w_k) = \mathbb{C}(v_k) = \sigma^2 I$ as well. Finally, for all $k, k' \geq 0$*

[1] In [18], the authors prove with a counterexample in the context of sensor scheduling that the minimum mean square error of the Kalman filter with respect to the system's state is not in general a supermodular set function. We can extend this counterexample in the context of minimal sensor placement as well: the minimum mean square error of the Kalman with respect to the system's state is not in general a supermodular set function with respect to the choice of the sensor set.

[2] Such algorithms, that involve the minimization of supermodular set functions, are also used in the machine learning [20], leader selection [2], [21], [22], sensor scheduling [17], [18], actuator placement [4], [7], [8], [11], [13], [23] and sensor placement in static environments [15], [24] literature. Their popularity is due to their simple implementation — they are greedy algorithms — and provable worst-case approximation factors, that are the best one can achieve in polynomial time for several classes of supermodular functions [19], [25].

*such that $k \neq k'$, $x_0$, $w_k$ and $v_k$, as well as, $w_k$, $w_{k'}$, $v_k$ and $v_{k'}$, are uncorrelated.*[3]

Moreover, for $k \geq 0$, consider the vector of measurements $\bar{y}_k$, the vector of process noises $\bar{w}_k$ and the vector of measurement noises $\bar{v}_k$, defined as follows: $\bar{y}_k \equiv (y_0^\top, y_1^\top, \ldots, y_k^\top)^\top$, $\bar{w}_k \equiv (w_0^\top, w_1^\top, \ldots, w_k^\top)^\top$, and $\bar{v}_k \equiv (v_0^\top, v_1^\top, \ldots, v_k^\top)^\top$, respectively; the vector $\bar{y}_k$ is known, while the $\bar{w}_k$ and $\bar{v}_k$ are not.

**Definition 1** (Observation interval and its length)**.** *The interval* $[0,k] \equiv \{0,1,\ldots,k\}$ *is called the* observation interval *of* (1). *Moreover,* $k+1$ *is its* length.

Evidently, the length of an observation interval $[0,k]$ equals the number of measurements $y_0, y_1, \ldots, y_k$.

In this paper, given an observation interval $[0,k]$, we consider the minimum mean square linear estimators for $x_{k'}$, for any $k' \in [0,k]$ [1]. In particular, (1) implies

$$\bar{y}_k = \mathcal{O}_k z_{k-1} + \bar{v}_k, \tag{2}$$

where $\mathcal{O}_k$ is the $c(k+1) \times n(k+1)$ matrix $[L_0^\top C_0^\top, L_1^\top C_1^\top, \ldots, L_k^\top C_k^\top]^\top$, $L_0$ the $n \times n(k+1)$ matrix $[I, 0]$, $L_i$, for $i \geq 1$, the $n \times n(k+1)$ matrix $[A_{i-1} \cdots A_0, A_{i-1} \cdots A_1, \ldots, A_{i-1}, I, 0]$, and $z_{k-1} \equiv (x_0^\top, \bar{w}_{k-1}^\top)^\top$. As a result, the minimum mean square linear estimate of $z_{k-1}$ is the $\hat{z}_{k-1} \equiv \mathbb{E}(z_{k-1}) + \mathcal{O}_k^\top \left(\mathcal{O}_k \mathcal{O}_k^\top + I\right)^{-1} (\bar{y}_k - \mathcal{O}_k \mathbb{E}(z_{k-1}) - \mathbb{E}(\bar{v}_k))$; its error covariance is

$$\begin{aligned} \Sigma_{z_{k-1}} &\equiv \mathbb{E}\left((z_{k-1} - \hat{z}_{k-1})(z_{k-1} - \hat{z}_{k-1})^\top\right) \\ &= \sigma^2 \left(I - \mathcal{O}_k^\top \left(\mathcal{O}_k \mathcal{O}_k^\top + I\right)^{-1} \mathcal{O}_k\right) \end{aligned} \tag{3}$$

and its minimum mean square error

$$\begin{aligned} \text{mmse}(z_{k-1}) &\equiv \mathbb{E}\left((z_{k-1} - \hat{z}_{k-1})^\top (z_{k-1} - \hat{z}_{k-1})\right) \\ &= \text{tr}\left(\Sigma_{z_{k-1}}\right). \end{aligned} \tag{4}$$

As a result, the corresponding minimum mean square linear estimator of $x_{k'}$, for any $k' \in [0,k]$, is

$$\hat{x}_{k'} = L_{k'} \hat{z}_{k-1}, \tag{5}$$

(since $x_{k'} = L_{k'} z_{k-1}$), with minimum mean square error

$$\text{mmse}(x_{k'}) \equiv \text{tr}\left(L_{k'} \Sigma_{z_{k-1}} L_{k'}^\top\right). \tag{6}$$

In particular, the recursive implementation of (5) results to the Kalman filtering algorithm [30].

In this paper, in addition to the minimum mean square error of $\hat{x}_{k'}$, we also consider per (5) the estimation error metric that is related to the $\eta$-confidence ellipsoid of $z_{k-1} - \hat{z}_{k-1}$ [26]. Specifically, this is the minimum volume ellipsoid that contains $z_{k-1} - \hat{z}_{k-1}$ with probability $\eta$, that is, the $\mathcal{E}_\epsilon \equiv \{z : z^\top \Sigma_{z_{k-1}} z \leq \epsilon\}$, where $\epsilon \equiv F^{-1}_{\chi^2_{n(k+1)}}(\eta)$ and $F_{\chi^2_{n(k+1)}}$ is the cumulative distribution function of a $\chi$-squared random variable with $n(k+1)$ degrees of freedom [31]. Therefore, the volume of $\mathcal{E}_\epsilon$,

$$\text{vol}(\mathcal{E}_\epsilon) \equiv \frac{(\epsilon\pi)^{n(k+1)/2}}{\Gamma\left(n(k+1)/2+1\right)} \det\left(\Sigma_{z_{k-1}}^{1/2}\right), \tag{7}$$

where $\Gamma(\cdot)$ denotes the Gamma function [31], quantifies the estimation's error of $\hat{z}_{k-1}$, and as a result, for any $k' \in [0,k]$, of $\hat{x}_{k'}$ as well, since per (5) the optimal estimator for $z_{k-1}$ defines the optimal estimator for $x_{k'}$.

Henceforth, we consider the logarithm of (7),

$$\log \text{vol}(\mathcal{E}_\epsilon) = \beta + 1/2 \log\det\left(\Sigma_{z_{k-1}}\right); \tag{8}$$

$\beta$ is a constant that depends only on $n(k+1)$ and $\epsilon$, in accordance to (7), and as a result, we refer to the $\log\det\left(\Sigma_{z_{k-1}}\right)$ as the $\log\det$ estimation error of the Kalman filter of (1):

**Definition 2** ($\log\det$ estimation error of the Kalman filter)**.** *Given an observation interval* $[0,k]$*, the* $\log\det\left(\Sigma_{z_{k-1}}\right)$ *is called the* $\log\det$ estimation error of the Kalman filter *of* (1).

In the following paragraphs, we present our sensor placement framework, that leads to our sensor placement problems.

### *B. Sensor Placement Framework*

In this paper, we study among others the effect of the selected sensors in (1) on mmse$(x_0)$ and mmse$(x_k)$. Therefore, this translates to the following conditions on $C_k$, for all $k \geq 0$, in accordance with the minimal sensor placement literature [4].

**Assumption 2** ($C$ is a full row-rank constant zero-one matrix)**.** *For all $k \geq 0$, $C_k = C \in R^{c \times n}$, where $C$ is a zero-one constant matrix. Specifically, each row of $C$ has one element equal to one, and each column at most one, such that $C$ has rank $c$.*

In particular, when for some $i$, $C_{ij}$ is one, the $j$-th state of $x_k$ is measured; otherwise, it is not. Therefore, the number of non-zero elements of $C$ coincides with the number of placed sensors in (1).

**Definition 3** (Sensor set and sensor placement)**.** *Consider a $C$ per Assumption 2 and define $\mathcal{S} \equiv \{i : i \in [n]$ and $C_{ji} = 1$, for some $j \in [r]\}$; $\mathcal{S}$ is called a* sensor set *or a* sensor placement *and each of its elements a* sensor.

### *C. Sensor Placement Problems*

We introduce three objectives, that we use to define the sensor placement problems we consider in this paper.

**Objective 1** (Fundamental limits in optimal sensor placement)**.** *Given an observation interval* $[0,k]$*,* $i \in \{0,k\}$ *and a desired* mmse$(x_i)$*, identify fundamental limits in the design of the sensor set.*

As an example of a fundamental limit, we prove that the number of sensors grows linearly with the system's size for fixed estimation error mmse$(x_i)$ —this is clearly a major limitation, especially when the system's size is large. This result, as well as, the rest of our contributions with respect to Objective 1, is presented in Section III.

[3]This assumption is common in the related literature [26], and it translates to a worst-case scenario for the problem we consider in this paper.

**Objective 2** ($\log\det$ estimation error as a sensor set function)**.** *Given an observation interval* $[0,k]$*, identify properties of the* $\log\det\left(\Sigma_{z_{k-1}}\right)$ *as a sensor set function.*

We address this objective in Section IV, where we prove that $\log\det\left(\Sigma_{z_{k-1}}\right)$ is a supermodular and non-increasing set function with respect to the choice of the sensor set —the basic definitions of supermodular set functions are presented in that section as well.

**Objective 3** (Algorithms for optimal sensor placement)**.** *Given an observation interval* $[0,k]$*, identify a sensor set* $\mathcal{S}$ *that solves the* minimal sensor placement problem:

$$\begin{aligned} &\underset{\mathcal{S}\subseteq[n]}{\text{minimize}} && |\mathcal{S}| \\ &\text{subject to} && \log\det\left(\Sigma_{z_{k-1}}\right)\leq R. \end{aligned} \qquad (\mathcal{P}_1)$$

That is, with $\mathcal{P}_1$ we design an estimator that guarantees a specified error and uses a minimal number of sensors. The corresponding algorithm is provided in Section V.

All of our contributions with respect to the Objectives 1, 2 and 3 are presented in the following sections.

## III. Fundamental Limits in Optimal Sensor Placement

In this section, we present our contributions with respect to Objective 1. In particular, given any finite observation interval, we prove that the minimum mean square error decreases only linearly as the number of sensors increases. That is, adding extra sensors so to reduce the minimum mean square estimation error of the Kalman filter is ineffective, a fundamental design limit. Similarly, we prove that the number of sensors grows linearly with the system's size for fixed minimum mean square error; this is another fundamental limit, especially for systems where the system's size is large. On the contrary, given a sensor set of fixed cardinality, we prove that the length of the observational interval increases only logarithmically with the system's size for fixed minimum mean square error. Overall, our novel results quantify the trade-off between the number of sensors and that of output measurements so to achieve a specified value for the minimum mean square error.

To this end, given $i\in\{0,k\}$, we first determine a lower and upper bound for $\text{mmse}(x_i)$.[4]

**Theorem 1** (A lower and upper bound for the estimation error with respect to the number of sensors and the length of the observation interval)**.** *Consider a sensor set* $\mathcal{S}$*, any finite observation interval* $[0,k]$ *and a non-zero* $\sigma$*. Moreover, let* $\mu\equiv\max_{m\in[0,k]}\|A_m\|_2$ *and assume* $\mu\neq 1$*. Given* $i\in\{0,k\}$*,*

$$\frac{n\sigma^2 l_i}{|\mathcal{S}|\left(1-\mu^{2(k+1)}\right)/\left(1-\mu^2\right)+1}\leq\text{mmse}(x_i)\leq n\sigma^2 u_i, \qquad (9)$$

*where* $l_0=1$*,* $u_0=1$*,* $l_k=\lambda_{\min}\left(L_k^\top L_k\right)$ *and* $u_k=(k+1)\lambda_{\max}\left(L_k^\top L_k\right)$*.*

[4] The extension of Theorem 1 to the case $\mu=1$ is straightforward, yet notationally involved; as a result, we omit it.

The upper bound corresponds to the case where no sensors have been placed ($C=0$). On the other hand, the lower bound corresponds to the case where $|\mathcal{S}|$ sensors have been placed.

As expected, the lower bound in (9) decreases as the number of sensors or the length of the observational interval increases; the increase of either should push the estimation error downwards. Overall, this lower bound quantifies fundamental limits in the design of the Kalman estimator: first, this bound decreases only inversely proportional to the number of sensors. Therefore, the estimation error of the optimal linear estimator decreases only linearly as the number of sensors increases. That is, adding extra sensors so to reduce the minimum mean square estimation error of the Kalman filter is ineffective, a fundamental design limit. Second, this bound increases linearly with the system's size. This is another fundamental limit, especially for systems where the system's size is large. Finally, for fixed and non-zero $\lambda_{\min}\left(L_k^\top L_k\right)$, these scaling extend to the $\text{mmse}(x_k)$ as well, for any finite $k$.

**Corollary 1** (Trade-off among the number of sensors, estimation error and the length of the observation interval)**.** *Consider any finite observation interval* $[0,k]$*, a non-zero* $\sigma$*, and for* $i\in\{0,k\}$*, that the desired value for* $\text{mmse}(x_i)$ *is* $\alpha$ *(*$\alpha>0$*). Moreover, let* $\mu\equiv\max_{m\in[0,k]}\|A_m\|_2$ *and assume* $\mu\neq 1$*. Any sensor set* $\mathcal{S}$ *that achieves* $\text{mmse}(x_i)=\alpha$ *satisfies:*

$$|\mathcal{S}|\geq\left(n\sigma^2 l_i/\alpha-1\right)\frac{1-\mu^2}{1-\mu^{2(k+1)}}. \qquad (10)$$

*where* $l_0=1$ *and* $l_k=\lambda_{\min}\left(L_k^\top L_k\right)$*.*

The above corollary shows that the number of sensors increases as the minimum mean square error or the number of output measurements decreases. More importantly, it shows that the number of sensors increases linearly with the system's size for fixed minimum mean square error. This is again a fundamental design limit, especially when the system's size is large.[5]

## IV. Submodularity in Optimal Sensor Placement

In this section, we present our contributions with respect to Objective 2. In particular, we first derive a closed formula for $\log\det\left(\Sigma_{z_{k-1}}\right)$ and then prove that it is a supermodular and non-increasing set function in the choice of the sensor set.

We now give the definition of a supermodular set function, as well as, that of an non-decreasing set function —we follow [32] for this material.

Denote as $2^{[n]}$ the power set of $[n]$.

**Definition 4** (Submodularity and supermodularity)**.** *A function* $h:2^{[n]}\mapsto\mathbb{R}$ *is* submodular *if for any sets* $\mathcal{S}$ *and* $\mathcal{S}'$*, with* $\mathcal{S}\subseteq\mathcal{S}'\subseteq[n]$*, and any* $a\notin\mathcal{S}'$*,*

$$h(\mathcal{S}\cup\{a\})-h(\mathcal{S})\geq h(\mathcal{S}'\cup\{a\})-h(\mathcal{S}').$$

[5] For fixed and non-zero $\lambda_{\min}\left(L_k^\top L_k\right)$, the comments of this paragraph extend to the $\text{mmse}(x_k)$ as well, for any finite $k$ —on the other hand, if $\lambda_{\min}\left(L_k^\top L_k\right)$ varies with the system's size, since $\lambda_{\min}\left(L_k^\top L_k\right)\leq 1$, the number of sensors can increase sub-linearly with the system's size for fixed $\text{mmse}(x_k)$.

*A function* $h : 2^{[n]} \mapsto \mathbb{R}$ *is* supermodular *if* $(-h)$ *is submodular.*

An alternative definition of a submodular function is based on the notion of non-increasing set functions.

**Definition 5** (Non-increasing and non-decreasing set function)**.** *A function* $h : 2^{[n]} \mapsto \mathbb{R}$ *is a* non-increasing set function *if for any* $\mathcal{S} \subseteq \mathcal{S}' \subseteq [n]$, $h(\mathcal{S}) \geq h(\mathcal{S}')$. *Moreover,* $h$ *is a* non-decreasing set function *if* $(-h)$ *is a non-increasing set function.*

Therefore, a function $h : 2^{[n]} \mapsto \mathbb{R}$ is submodular if, for any $a \in [n]$, the function $h_a : 2^{[n]\setminus\{a\}} \mapsto \mathbb{R}$, defined as $h_a(\mathcal{S}) \equiv h(\mathcal{S} \cup \{a\}) - h(\mathcal{S})$, is a non-increasing set function. This property is also called the *diminishing returns property*.

The first major result of this section follows, where we let

$$O_k \equiv \mathcal{O}_k^\top \mathcal{O}_k,$$

given an observation interval $[0, k]$.

**Proposition 1** (Closed formula for the $\log\det$ error as a sensor set function)**.** *Given any finite observation interval* $[0, k]$ *and non-zero* $\sigma$*, irrespective of Assumption 2,*

$$\log\det\left(\Sigma_{z_{k-1}}\right) = 2n(k+1)\log(\sigma) - \log\det(O_k + I). \quad (11)$$

Therefore, the $\log\det\left(\Sigma_{z_{k-1}}\right)$ depends on the sensor set through $O_k$. Now, the main result of this section follows, where we make explicit the dependence of $O_k$ on the sensor set $\mathcal{S}$.

**Theorem 2** (The $\log\det$ error is a supermodular and non-increasing set function with respect to the choice of the sensor set)**.** *Given any finite observation interval* $[0, k]$*, the*

$$\log\det\left(\Sigma_{z_{k-1}}, \mathcal{S}\right) = 2n(k+1)\log(\sigma) - \log\det(O_{k,\mathcal{S}} + I) : \mathcal{S} \in 2^{[n]} \mapsto \mathbb{R}$$

*is a supermodular and non-increasing set function with respect to the choice of the sensor set* $\mathcal{S}$.

The above theorem states that for any finite observation interval, the $\log\det$ error of the Kalman filter is a supermodular and non-increasing set function with respect to the choice of the sensor set for any finite $k$. Hence, it exhibits the diminishing returns property: its rate of reduction with respect to newly placed sensors decreases as the cardinality of the already placed sensors increases. On the one hand, this property implies another fundamental design limit, in accordance to that of Theorem 1: adding new sensors, after a first few, becomes ineffective for the reduction of the estimation error. On the other hand, it also implies that greedy approach for solving $\mathcal{P}_1$ is effective [19], [25]. Thereby, we next use the results from the literature on submodular function maximization [33] and provide an efficient algorithm for $\mathcal{P}_1$.

## V. Algorithms for Optimal Sensor Placement

In this section, we present our contributions with respect to Objective 3: $\mathcal{P}_1$ is combinatorial, and in Section IV we proved that it involves the minimization of the supermodular set function $\log\det$ error. In particular, because the minimization of a general supermodular function is NP-hard [19], in this section we provide efficient approximation algorithms for the general solution of $\mathcal{P}_1$, along with their worst-case performance guarantees.

Specifically, we provide an efficient algorithm for $\mathcal{P}_1$ that returns a sensor set that satisfies the estimation bound of $\mathcal{P}_1$ and has cardinality up to a multiplicative factor from the minimum cardinality sensor sets that meet the same estimation bound. More importantly, this multiplicative factor depends only logarithmically on the problem's $\mathcal{P}_1$ parameters.

To this end, we first present a fact from the supermodular functions minimization literature that we use so to construct an approximation algorithm for $\mathcal{P}_1$ —we follow [32] for this material. In particular, consider the following problem, which is of similar structure to $\mathcal{P}_1$, where $h : 2^{[n]} \mapsto \mathbb{R}$ is a supermodular and non-increasing set function:

$$\begin{aligned} \underset{\mathcal{S}\subseteq[n]}{\text{minimize}} \quad & |\mathcal{S}| \\ \text{subject to} \quad & h(\mathcal{S}) \leq R. \end{aligned} \quad (\mathcal{P})$$

The following greedy algorithm has been proposed for its approximate solution, for which, the subsequent fact is true.

**Algorithm 1** Approximation Algorithm for $\mathcal{P}$.

**Input:** $h$, $R$.
**Output:** Approximate solution for $\mathcal{P}$.
$\mathcal{S} \leftarrow \emptyset$
**while** $h(\mathcal{S}) > R$ **do**
$a_i \leftarrow a' \in \arg\max_{a\in[n]\setminus\mathcal{S}}\left(h(\mathcal{S}) - h(\mathcal{S}\cup\{a\})\right)$
$\mathcal{S} \leftarrow \mathcal{S} \cup \{a_i\}$
**end while**

**Fact 1.** *Denote as* $\mathcal{S}^\star$ *a solution to* $\mathcal{P}$ *and as* $\mathcal{S}_0, \mathcal{S}_1, \ldots$ *the sequence of sets picked by Algorithm 1. Moreover, let* $l$ *be the smallest index such that* $h(\mathcal{S}_l) \leq R$. *Then,*

$$\frac{l}{|\mathcal{S}^\star|} \leq 1 + \log\frac{h([n]) - h(\emptyset)}{h([n]) - h(\mathcal{S}_{l-1})}.$$

For several classes of submodular functions, this is the best approximation factor one can achieve in polynomial time [19]. Therefore, we use this result to provide the approximation Algorithm 2 for $\mathcal{P}_1$, where we make explicit the dependence of $\log\det\left(\Sigma_{z_{k-1}}\right)$ on the selected sensor set $\mathcal{S}$. Moreover, its performance is quantified with Theorem 3.

**Algorithm 2** Approximation Algorithm for $\mathcal{P}_1$.

For $h(\mathcal{S}) = \log\det\left(\Sigma_{z_{k-1}}, \mathcal{S}\right)$, where $\mathcal{S} \subseteq [n]$, Algorithm 2 is the same as Algorithm 1.

**Theorem 3** (A Submodular Set Coverage Optimization for $\mathcal{P}_1$)**.** *Denote a solution to $\mathcal{P}_1$ as $\mathcal{S}^\star$ and the selected set by Algorithm 2 as $\mathcal{S}$. Then,*

$$\log\det\left(\Sigma_{z_{k-1}}, \mathcal{S}\right) \leq R, \tag{12}$$

$$\frac{|\mathcal{S}|}{|\mathcal{S}^\star|} \leq 1 + \log \frac{\log\det\left(\Sigma_{z_{k-1}}, \emptyset\right) - \log\det\left(\Sigma_{z_{k-1}}, [n]\right)}{R - \log\det\left(\Sigma_{z_{k-1}}, [n]\right)} \equiv F_i, \tag{13}$$

*where* $\log\det\left(\Sigma_{z_{k-1}}, \emptyset\right) \leq n(k+1)\log(\sigma^2)$*. Finally, the computational complexity of Algorithm 2 is* $O(n^2(nk)^3)$*.*

Therefore, Algorithm 2 returns a sensor set that meets the estimation bound of $\mathcal{P}_1$. Moreover, the cardinality of this set is up to a multiplicative factor of $F_i$ from the minimum cardinality sensor sets that meet the same estimation bound —that is, $F_i$ is a worst-case approximation guarantee for Algorithm 2. Additionally, $F_i$ depends only logarithmically on the problem's $\mathcal{P}_1$ parameters. Finally, the dependence of $F_i$ on $n$, $R$ and $\sigma^2$ is expected from a design perspective: increasing the network size $n$, requesting a better estimation guarantee by decreasing $R$, or incurring a noise of greater variance, should all push the cardinality of the selected sensor set upwards.

## VI. Concluding Remarks

We considered a linear time-variant system and studied the properties of its Kalman estimator given an observation interval and a sensor set. Our contributions were threefold. First, in Section III we presented several design limits. For example, we proved that the number of sensors grows linearly with the system's size for fixed minimum mean square error; this is a fundamental limit, especially for systems where the system's size is large. Second, in Section IV we proved that the $\log\det$ error is a supermodular and non-increasing set function with respect to the choice of the sensor set. Third, in Section V, we used this result to provide an efficient approximation algorithm for the solution of $\mathcal{P}_1$, along with its worst-case performance guarantees. Our future work is focused on extending the results of this paper to the problem of sensor scheduling.

## References


[1] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear estimation*. Prentice Hall Upper Saddle River, NJ, 2000, vol. 1.

[2] A. Clark, L. Bushnell, and R. Poovendran, "On leader selection for performance and controllability in multi-agent systems," in *IEEE 51st Annual Conference on Decision and Control (CDC)*, Dec 2012, pp. 86–93.

[3] 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*, vol. 58, no. 9, pp. 2426–2431, 2013.

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

[5] S. Pequito, G. Ramos, S. Kar, A. P. Aguiar, and J. Ramos, "On the Exact Solution of the Minimal Controllability Problem," *arXiv preprint arXiv: 1401.4209*, 2014.

[6] F. Pasqualetti, S. Zampieri, and F. Bullo, "Controllability metrics, limitations and algorithms for complex networks," *IEEE Transactions on Control of Network Systems*, vol. 1, no. 1, pp. 40–52, March 2014.

[7] T. H. Summers, F. L. Cortesi, and J. Lygeros, "On submodularity and controllability in complex dynamical networks," *arXiv preprint arXiv:1404.7665*, 2014.

[8] 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*, 2015, in press.

[9] N. Matni and V. Chandrasekaran, "Regularization for design," in *IEEE 53rd Annual Conference on Decision and Control (CDC)*, 2014, pp. 1111–1118.

[10] S. Pequito, S. Kar, and A. Aguiar, "A framework for structural input/output and control configuration selection in large-scale systems," *IEEE Transactions on Automatic Control*, vol. PP, no. 99, pp. 1–1, 2015.

[11] V. Tzoumas, M. A. Rahimian, G. J. Pappas, and A. Jadbabaie, "Minimal actuator placement with optimal control constraints," in *Proceedings of the American Control Conference*, July 2015, pp. 2081 – 2086.

[12] G. Yan, G. Tsekenis, B. Barzel, J.-J. Slotine, Y.-Y. Liu, and A.-L. Barabási, "Spectrum of controlling and observing complex networks," *Nature Physics*, no. 11, pp. 779–786, 2015.

[13] Y. Zhao and J. Cortés, "Gramian-based reachability metrics for bilinear networks," *arXiv preprint arXiv:1509.02877*, 2015.

[14] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Journal of Fluids Engineering*, vol. 82, no. 1, pp. 35–45, 1960.

[15] 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*, vol. 9, pp. 235–284, 2008.

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

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

[18] 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.

[19] U. Feige, "A threshold of ln n for approximating set cover," *J. ACM*, vol. 45, no. 4, pp. 634–652, Jul. 1998.

[20] A. Krause and C. Guestrin, "Beyond convexity: Submodularity in machine learning."

[21] A. Clark, L. Bushnell, and R. Poovendran, "A supermodular optimization framework for leader selection under link noise in linear multi-agent systems," *IEEE Transactions on Automatic Control*, vol. 59, no. 2, pp. 283–296, Feb 2014.

[22] A. Clark, B. Alomair, L. Bushnell, and R. Poovendran, "Minimizing convergence error in multi-agent systems via leader selection: A supermodular optimization approach," *IEEE Transactions on Automatic Control*, vol. 59, no. 6, pp. 1480–1494, June 2014.

[23] V. Tzoumas, A. Jadbabaie, and G. J. Pappas, "Minimal reachability problems," in *54th IEEE Conference on Decision and Control (CDC)*, December 2015, to appear.

[24] C. Jiang, Y. Chai Soh, and H. Li, "Sensor placement by maximal projection on minimum eigenspace for linear inverse problem," *arXiv preprint arXiv: 1506.00747*, 2015.

[25] G. L. Nemhauser and L. A. Wolsey, "Best algorithms for approximating the maximum of a submodular set function," *Mathematics of operations research*, vol. 3, no. 3, pp. 177–188, 1978.

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

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

[28] U. Munz, M. Pfister, and P. Wolfrum, "Sensor and actuator placement for linear systems based on $h_2$ and $h_\infty$ optimization," *IEEE Transactions on Automatic Control*, vol. 59, no. 11, pp. 2984–2989, Nov 2014.

[29] M.-A. Belabbas, "Geometric methods for optimal sensor design," *arXiv preprint arXiv: 1503.05968*, 2015.

[30] D. P. Bertsekas, *Dynamic Programming and Optimal Control, Vol. I*, 3rd ed. Athena Scientific, 2005.

[31] S. Venkatesh, *The Theory of Probability: Explorations and Applications*. Cambridge University Press, 2012.

[32] L. A. Wolsey, "An analysis of the greedy algorithm for the submodular set covering problem," *Combinatorica*, vol. 2, no. 4, pp. 385–393, 1982.

[33] G. L. Nemhauser and L. A. Wolsey, *Integer and Combinatorial Optimization*. New York, NY, USA: Wiley-Interscience, 1988.