Approximate Supermodularity of Kalman Filter Sensor Selection

Luiz F. O. Chamon, *Student Member, IEEE*, George J. Pappas, *Fellow, IEEE*, and Alejandro Ribeiro

***Abstract*—This article considers the problem of selecting sensors in a large-scale system to minimize the error in estimating its states, more specifically, the state estimation mean-square error (MSE) and worst-case error for Kalman filtering and smoothing. Such selection problems are in general NP-hard, i.e., their solution can only be approximated in practice even for moderately large problems. Due to its low complexity and iterative nature, greedy algorithms are often used to obtain these approximations by selecting one sensor at a time choosing at each step the one that minimizes the estimation performance metric. When this metric is supermodular, this solution is guaranteed to be $(1 - 1/e)$-optimal. This is, however, not the case for the MSE or the worst-case error. This issue is often circumvented by using supermodular surrogates, such as the $\log\det$, despite the fact that minimizing the $\log\det$ is not equivalent to minimizing the MSE. Here, this issue is addressed by leveraging the concept of approximate supermodularity to derive near-optimality certificates for greedily minimizing the estimation mean-square and worst-case error. In typical application scenarios, these certificates approach the $(1 - 1/e)$ guarantee obtained for supermodular functions, thus demonstrating that no change to the original problem is needed to obtain guaranteed good performance.**

***Index Terms*—Discrete optimization, greedy algorithms, kalman filters, sensor selection, submodularity.**

I. Introduction

OFTEN, the outputs of a dynamical system can only be partially observed due to a limited sensing budget and/or because the system scale makes instrumenting all outputs impractical. This issue is critical for distributed systems, where power and communication constraints further limit the number of sensors available [2], [3], and is found in applications such as target tracking, field monitoring, power allocation, and biological systems analysis [4]–[9]. In these cases, we must choose which outputs to observe in order to best estimate the system internal states.

Manuscript received May 8, 2018; revised May 9, 2018, March 29, 2019, and September 18, 2019; accepted February 7, 2020. Date of publication February 13, 2020; date of current version December 24, 2020. This work was supported by ARL DCIST CRA W911NF-17-2-0181. Part of the results in this paper was presented in IEEE Conference on Decision and Control, 2017. Recommended by Associate Editor Q.-S. Jia. *(Corresponding author: Luiz F. O. Chamon.)*

The authors are with the Department of Electrical and Systems Engineering, University of Pennsylvania, Philadelphia, PA 19104 USA (e-mail: luizf@seas.upenn.edu; pappasg@seas.upenn.edu; aribeiro@seas.upenn.edu).

Color versions of one or more of the figures in this article are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TAC.2020.2973774

In general, solving this constrained output selection problem, often known as *sensor selection*, is NP-hard [10]–[16]. Still, there exist instances for which its solution can be approximated. For instance, when the goal is to find the smallest set of outputs that makes the system observable, greedy sensor selection is near-optimal [17]. When only structural descriptions are available, the problem is also known to be NP-hard, although a variant in which the output sensing can be designed freely admits a polynomial time solution [18]–[20]. Nevertheless, observability alone may not guarantee good state estimates because it does not take into account how hard each system mode is to observe. To do so, we turn to measures of the state estimation error.

In this setting, natural choices of objectives, especially in Kalman filtering applications, are the state estimation mean-square error (MSE) and the worst-case estimation error. However, approximation guarantees for these problems are scarce. In particular, typical guarantees for greedy search do not apply since these functions are not supermodular, a diminishing returns property used to obtain these results (see [13]–[16] for counterexamples in the context of control theory). Moreover, though convex relaxations of these problems have been proposed, they too lack near-optimality certificates [15], [21]–[26]. In fact, system-independent guarantees cannot be provided for optimizing the state estimation MSE unless P = NP [16].

A typical way to address this issue is to use surrogate supermodular objectives, such as the log-determinant of the error covariance matrix or some information-theoretic measure relating the selected output set and its complement [10], [12], [17], [23], [27], [28]. Though effective, there is no direct relation between the estimation MSE or the worst-case estimation error and these cost functions. As a matter of fact, despite its usefulness in minimizing the volume of the confidence ellipsoid, the log-determinant is a poor proxy for the MSE (see Remark 1).

In this article, we take a page out of Tukeys book and try to solve the "right problem" [29], i.e., *selecting sensors so as to directly minimize the estimation MSE or the worst-case estimation error*. To make this selection scalable and obtain performance guarantees, we prove that greedy sensor selection is near-optimal for minimizing the estimation MSE and the worst-case estimation error in both filtering and smoothing settings. Our results show that when the measurement noise is large compared to the process noise, a scenario of practical relevance in Kalman filtering, greedy solutions to these problems approach the classical $(1 - 1/e)$-optimal certificate for supermodular objectives [30].

0018-9286 © 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See https://www.ieee.org/publications/rights/index.html for more information.

Authorized licensed use limited to: University of Pennsylvania. Downloaded on October 29,2026 at 15:10:48 UTC from IEEE Xplore. Restrictions apply.

To derive these guarantees, we leverage two notions of *approximate supermodularity* in which limited violations of the diminishing return property are allowed (Section III). In the case of α-supermodular functions, violations are allowed up to a multiplicative factor α, whereas ϵ-supermodular functions violate supermodularity up to an additive constant ϵ. We prove that these functions can be minimized near-optimally using greedy search and that α and ϵ not only quantify how much supermodularity is violated, but also the loss in performance guarantee due to these violations (Theorems 1 and 2). We proceed to show that the estimation MSE is α-supermodular and that the worst-case estimation error—i.e., the spectral norm of the error covariance matrix—is ϵ-supermodular (Section IV). Combining these results yields performance guarantees for greedy sensor selection using the MSE and the worst-case error in filtering (Theorem 5) and smoothing (Theorem 6) state estimation problems (Section V). Note that the α-supermodularity and corresponding guarantees for the MSE in filtering and smoothing problems were previously established in [1]. A field monitoring example is provided to illustrate these results (Section VI).

Notation: Lowercase boldface letters represent vectors ($\boldsymbol{x}$), uppercase boldface letters are matrices ($\boldsymbol{X}$), and calligraphic letters denote sets ($\mathcal{X}$). We write $|\mathcal{X}|$ for the cardinality of $\mathcal{X}$ and let $\emptyset$ denote the empty set. To say $\boldsymbol{X}$ is a positive semidefinite (PSD) matrix, we write $\boldsymbol{X} \succeq 0$, so that for $\boldsymbol{X}, \boldsymbol{Y} \in \mathbb{R}^{n\times n}$, $\boldsymbol{X} \preceq \boldsymbol{Y} \Leftrightarrow \boldsymbol{b}^T\boldsymbol{X}\boldsymbol{b} \leq \boldsymbol{b}^T\boldsymbol{Y}\boldsymbol{b}$ for all $\boldsymbol{b} \in \mathbb{R}^n$. Similarly, we write $\boldsymbol{X} \succ 0$ when $\boldsymbol{X}$ is positive definite. For a square matrix $\boldsymbol{X}$, we denote its spectral norm (ℓ_2 induced norm) by $\|\boldsymbol{X}\|$ and its maximum and minimum eigenvalues by $\lambda_{\max}(\boldsymbol{X})$ and $\lambda_{\min}(\boldsymbol{X})$, respectively.

II. Problem Formulation

Consider an autonomous dynamical system whose internal states $\boldsymbol{x}_k \in \mathbb{R}^n$ evolve as

$$\boldsymbol{x}_{k+1} = \boldsymbol{F}\boldsymbol{x}_k + \boldsymbol{w}_k \tag{1}$$

and a set of sensors $\mathcal{O}$ such that the output $\boldsymbol{y}_{u,k} \in \mathbb{R}^{p_u}$ of sensor $u \in \mathcal{O}$ is given by

$$\boldsymbol{y}_{u,k} = \boldsymbol{H}_u\boldsymbol{x}_k + \boldsymbol{v}_{u,k} \tag{2}$$

where $\boldsymbol{F} \in \mathbb{R}^{n\times n}$ is the state transition matrix and $\boldsymbol{H}_u \in \mathbb{R}^{p_u\times n}$ are the sensor output matrices. The process noise $\boldsymbol{w}_k$ and measurement noises $\boldsymbol{v}_{u,k}$ are zero-mean Gaussian random vectors with arbitrary covariances $\mathbb{E}\,\boldsymbol{w}_k\boldsymbol{w}_k^T = \boldsymbol{R}_w \succ 0$ and $\mathbb{E}\,\boldsymbol{v}_{u,k}\boldsymbol{v}_{u,k}^T = \boldsymbol{R}_{v,u} \succ 0$ for all k. We assume $\{\boldsymbol{v}_{u_1,i}, \boldsymbol{v}_{u_2,j}, \boldsymbol{w}_i, \boldsymbol{w}_j\}$ are independent for all $i \neq j$ and $u_1, u_2 \in \mathcal{O}$. The initial states $\boldsymbol{x}_0 \sim \mathcal{N}(\bar{\boldsymbol{x}}_0, \boldsymbol{\Pi}_0)$ are assumed to be a Gaussian random vector with mean $\bar{\boldsymbol{x}}_0$ and covariance $\boldsymbol{\Pi}_0 \succ 0$.

We seek to estimate the states $\boldsymbol{x}_k$ from the observations $\{\boldsymbol{y}_{u,i}\}_{u\in\mathcal{X}}$ for a sensing set $\mathcal{X} \subseteq \mathcal{O}$ and $i \leq k$. Depending on the observation history used to estimate the states at time k, we define the following estimation problems.

i) *Filtering*, where the current states are estimated using past observations, i.e., we seek $\hat{\boldsymbol{x}}_k(\mathcal{X}) = \mathbb{E}[\boldsymbol{x}_k \mid \{\boldsymbol{y}_{u,i}\}_{u\in\mathcal{X},\, i\leq k}]$.

ii) *Smoothing*, where all states up to time k are estimated based on all observations. Formally, we seek $\tilde{\boldsymbol{x}}_k(\mathcal{X}) = \mathbb{E}[\bar{\boldsymbol{x}}_k \mid \{\boldsymbol{y}_{u,i}\}_{u\in\mathcal{X},\, i\leq k}]$, where $\bar{\boldsymbol{x}}_k = [\,\boldsymbol{x}_0^T\ \boldsymbol{x}_1^T\ \cdots\ \boldsymbol{x}_k^T\,]^T$. As opposed to i), states are estimated based on both past and future observations.

Note that we do not consider the problem of *prediction* in which the states at time k is estimated based on observations up to $j < k$, i.e., in which we evaluate $\check{\boldsymbol{x}}_k(\mathcal{X}) = \mathbb{E}[\boldsymbol{x}_k \mid \{\boldsymbol{y}_{u,i}\}_{u\in\mathcal{X},\, i\leq j<k}]$. Although filtering and prediction are different estimation problems, they do not differ from the viewpoint of sensor selection. Indeed, prediction only extrapolate the jth filtered estimate using $\check{\boldsymbol{x}}_k = \boldsymbol{F}^{k-j}\hat{\boldsymbol{x}}_j$. Since it does not use additional measurements after the jth iteration, the sensor selection problem for prediction is in fact the same as the one for filtering.

The performance of the estimation problems i) and ii) for a given sensing set $\mathcal{X} \subseteq \mathcal{O}$ is fully characterized by the error covariance matrix of the estimators $\hat{\boldsymbol{x}}$ and $\tilde{\boldsymbol{x}}$. Explicitly

$$\boldsymbol{P}_k(\mathcal{X}) = \mathbb{E}\left[\hat{\boldsymbol{x}}_k(\mathcal{X}) - \boldsymbol{x}_k\right]\left[\hat{\boldsymbol{x}}_k(\mathcal{X}) - \boldsymbol{x}_k\right]^T \tag{3a}$$

$$\boldsymbol{Q}_k(\mathcal{X}) = \mathbb{E}\left[\tilde{\boldsymbol{x}}_k(\mathcal{X}) - \bar{\boldsymbol{x}}_k\right]\left[\tilde{\boldsymbol{x}}_k(\mathcal{X}) - \bar{\boldsymbol{x}}_k\right]^T. \tag{3b}$$

Observe that since the estimators depend on the sensing set $\mathcal{X}$, so do their error covariance matrices. In the propositions below, we show that the matrices in (3) can be written in a unified manner as a set function $\boldsymbol{Y}_k : 2^{\mathcal{O}} \to \mathbb{S}_+$ of the form

$$\boldsymbol{Y}_k(\mathcal{X}) = \left(\boldsymbol{M}_{\emptyset,k} + \sum_{u\in\mathcal{X}} \boldsymbol{M}_{u,k}\right)^{-1} \tag{4}$$

where $\boldsymbol{M}_{\emptyset,k} \succ 0$ is the *a priori* error covariance matrix that represents the error in the absence of sensor measurements and $\boldsymbol{M}_{u,k} \succeq 0$ carries the information obtained by using sensor $u \in \mathcal{O}$ at time k.

Proposition 1: The error covariance matrix $\boldsymbol{P}_k(\mathcal{X})$ of the filtered estimator $\hat{\boldsymbol{x}}$ from i) can be written recursively using (4) with

$$\boldsymbol{M}_{\emptyset,k} = \boldsymbol{P}_{k|k-1}^{-1} \text{ and } \boldsymbol{M}_{u,k} = \boldsymbol{V}_u \tag{5}$$

where $\boldsymbol{P}_{k|k-1} = \boldsymbol{F}\boldsymbol{P}_{k-1}\boldsymbol{F}^T + \boldsymbol{R}_w$ is the prediction error covariance matrix, $\boldsymbol{P}_{k-1}$ is the estimation error covariance matrix at time $k-1$, $\boldsymbol{P}_{0|-1} = \boldsymbol{\Pi}_0$, and $\boldsymbol{V}_u = \boldsymbol{H}_u^T\boldsymbol{R}_{v,u}^{-1}\boldsymbol{H}_u$.

Proof: This result is obtained directly from the Kalman filter (KF) recursion (see, e.g., [31]). For ease of reference, we provide a derivation in the Appendix. ■

Proposition 2: The error covariance matrix $\boldsymbol{Q}_k(\mathcal{X})$ of the smoothed estimator $\tilde{\boldsymbol{x}}$ from ii) can be written as in (4) with

$$\boldsymbol{M}_{\emptyset,k} = \operatorname{blkdiag}(\boldsymbol{\Pi}_0, \boldsymbol{I}\otimes\boldsymbol{R}_w)^{-1}$$

$$\boldsymbol{M}_{u,k} = \boldsymbol{\Phi}_k^T\left(\boldsymbol{I}\otimes\boldsymbol{V}_u\right)\boldsymbol{\Phi}_k \tag{6}$$

where $\boldsymbol{V}_u = \boldsymbol{H}_u^T\boldsymbol{R}_{v,u}^{-1}\boldsymbol{H}_u$, $\operatorname{blkdiag}(\boldsymbol{X}, \boldsymbol{Y})$ is the block diagonal matrix whose diagonal blocks are $\boldsymbol{X}$ and $\boldsymbol{Y}$, and

$$\boldsymbol{\Phi}_k = \begin{bmatrix} \boldsymbol{I} & & & \\ \boldsymbol{F} & \boldsymbol{I} & & \\ \vdots & \vdots & \ddots & \\ \boldsymbol{F}^k & \boldsymbol{F}^{k-1} & \cdots & \boldsymbol{I} \end{bmatrix}. \tag{7}$$

Authorized licensed use limited to: University of Pennsylvania. Downloaded on October 29,2026 at 15:10:48 UTC from IEEE Xplore. Restrictions apply.

Proof: The proof follows classical stochastic least-squares results (see, e.g., [31]). For ease of reference, we provide a derivation in the Appendix. ■

Propositions 1 and 2 show how the abstract set function (4) applies to both filtering and smoothing estimation problems. In the sequel, we formulate the sensor selection problem and derive theoretical guarantees directly in terms of the generic (4). Particular results for the estimation problems i) and ii) are provided in Section V.

A. Kalman Filter Sensor Selection

Our goal is to study sensor selection problems in which we seek a sensing set $\mathcal{X} \subseteq \mathcal{O}$ constrained by a budget $|\mathcal{X}| \leq s$ so that the estimation error covariance matrix $\boldsymbol{Y}_k(\mathcal{X})$ in (4) is minimized in some sense over a given time horizon N. Formally,

Problem (Kalman filter sensor selection): Take $N \in \mathbb{N}$ and let $\theta_k \geq 0$ for $k \in [0, N-1]$ be a set of nonnegative scalar weights and $h : \mathbb{S}_+ \to \mathbb{R}$ be a real-valued spectral function over the PSD cone. Kalman filter sensor selection can be written as the discrete optimization problem

$$\underset{\mathcal{X}\subseteq\mathcal{O}, |\mathcal{X}|\leq s}{\text{minimize}} \sum_{k=0}^{N-1} \theta_k \, h\left[\boldsymbol{Y}_{m+k}(\mathcal{X})\right] - C_\emptyset \tag{P*}$$

where $C_\emptyset \triangleq \sum_{k=0}^{N-1} \theta_k h[\boldsymbol{M}_{\emptyset,m+k}^{-1}]$ is a constant.

It is worth noting that the constant $C_\emptyset$ is included only so that the objective of (P*) vanishes when no sensor is selected—since $\boldsymbol{Y}_k(\emptyset) = \boldsymbol{M}_{\emptyset,k}^{-1}$, taking $\mathcal{X} = \emptyset$ in (P*) yields $\sum_{k=0}^{N-1} \theta_k \, h[\boldsymbol{Y}_{m+k}(\emptyset)] - C_\emptyset = 0$. This normalization has no effect on the optimization problem, but simplifies the statement of the approximation certificates presented in the following sections. Also, observe that Propositions 1 and 2 do not take control inputs into account. Hence, though a sensing set obtained from (P*) can be used to perform state estimation in the presence of control inputs, it need not be optimal for the joint sensing set-controller design problem. This formulation is left for future works.

Different choices of N, θ_k, and h in (P*) yield different sensor selection problems. For instance, (P*) becomes a myopic sensor selection problem when $N = 1$ [27]. When $N > 1$, we can choose to optimize the *final estimation error*—corresponding to $\theta_k = 0$ for all $k < N-1$ and $\theta_{N-1} = 1$—*the average error*—by making $\theta_k = 1$ for all k—or some *weighted average error*—e.g., a geometric discount using $\theta_k = \rho^{N-1-k}$ for $\rho < 1$. Observe that since the smoothing problem ii) estimates all states up to k, these multistep versions of (P*) may not be useful in practice. We still consider them for the sake of symmetry.

The scalarization h is necessary because matrix minimization in the PSD cone is not well-posed [32] and its choice affects both the complexity and the practical usefulness of (P*). For instance, (P*) with $h(\boldsymbol{X}) = -\operatorname{Tr}(\boldsymbol{X}^{-1})$ is combinatorial but modular and can therefore be solved in linear time [33], [34]. This objective, however, is of little relevance in practice inasmuch as it has no direct relation to the estimation MSE. In general, three scalarizations are widely agreed to be of practical significance.

i) The *trace* defined as

$$\mathcal{T}^\star \in \underset{\mathcal{T}\subseteq\mathcal{O}, |\mathcal{T}|\leq s}{\operatorname{argmin}} \; f_T(\mathcal{T}) \triangleq \sum_{k=0}^{N-1} \theta_k \operatorname{Tr}\left[\boldsymbol{Y}_{m+k}(\mathcal{T})\right] - C_\emptyset \tag{PI}$$

in which case the objective represents the estimation MSE.

ii) The *spectral norm* or the maximum eigenvalue

$$\mathcal{S}^\star \in \underset{\mathcal{S}\subseteq\mathcal{O}, |\mathcal{S}|\leq s}{\operatorname{argmin}} \; f_S(\mathcal{S}) \triangleq \sum_{k=0}^{N-1} \theta_k \left\|\boldsymbol{Y}_{m+k}(\mathcal{S})\right\| - C_\emptyset \tag{PII}$$

in which case the objective is a robust version of (PI), since it minimizes the worst-case estimation error.

iii) The *log-determinant*

$$\mathcal{D}^\star \in \underset{\mathcal{D}\subseteq\mathcal{O}, |\mathcal{D}|\leq s}{\operatorname{argmin}} \sum_{k=0}^{N-1} \theta_k \log\det\left[\boldsymbol{Y}_{m+k}(\mathcal{D})\right] - C_\emptyset \tag{PIII}$$

in which case the objective corresponds to the volume of the confidence ellipsoid [22], [28].

These formulations are natural because they minimize different estimation error metrics related to the MSE, which is a customary objective in traditional state estimation, i.e., problems that do not consider sensor selection ($\mathcal{X} = \mathcal{O}$). Still, problems (PI)–(PIII) are challenging to solve. They are known to be NP-hard [10], [12]–[15], [17], [33], so their solutions can only be approximated in practice, typically by means of greedy algorithms [15], [17].

In the case of problem (PIII), the log-determinant is a monotonically decreasing and supermodular set function [17], [23], [27], [28], i.e., it displays a "diminishing returns" property that yields near-optimal performance guarantees for its greedy minimization [30]. The MSE and the worst-case estimation error, on the other hand, are not supermodular in general [12]–[15], [33]. In fact, stringent conditions on the system matrices are needed to make the estimation MSE supermodular [35], [36]. These facts notwithstanding, greedy sensor selection has been observed to perform well when solving (PI) and (PII) in both control [13], [15], [17] and other contexts [37], [38].

Our goal in the remainder of this article is to show that in certain noise regimes of interest, these objectives are approximately submodular, so that the greedy solutions of (PI) and (PII) come with performance guarantees analogous to those of (PIII), thereby reconciling empirical observations with our theoretical understanding of these problems. To obtain these near-optimal certificates, the following section develops a theory of *approximately supermodular* functions, studying operations that preserve approximate supermodularity and showing that these functions can be near-optimally minimized. We then show that the trace and the spectral norm are approximately supermodular scalarizations of $\boldsymbol{Y}_k$ so as to finally provide explicit suboptimality bounds for greedy sensor selection for filtering/smoothing.

Remark 1: Besides its intrinsic value as a minimizer of the volume of the confidence ellipsoid [22], [28], the $\log\det$ is often put forward as a supermodular surrogate for the MSE. It is important to point out that this figure of merit is not directly

Authorized licensed use limited to: University of Pennsylvania. Downloaded on October 29,2026 at 15:10:48 UTC from IEEE Xplore. Restrictions apply.

related to the MSE or the worst-case error and that, in general, it is not a good surrogate for either of them. In fact, the log determinant is a good substitute for the MSE only when the problem has some inherent structure that constraints the confidence ellipsoid to be somewhat symmetric. Otherwise, since the volume of an ellipsoid can be reduced by decreasing the length of a single principal axis, using the $\log\det$ can lead to designs that perform well—in the MSE sense—along a few directions of the parameter space and poorly along all others. Formally, this can be seen by comparing the variation of the log-determinant and trace functions with respect to the eigenvalues of the PSD matrix $\boldsymbol{Y}$

$$\frac{\partial \log\det(\boldsymbol{Y})}{\partial \lambda_j(\boldsymbol{Y})} = \frac{1}{\lambda_j(\boldsymbol{Y})} \quad \text{and} \quad \frac{\partial \operatorname{Tr}(\boldsymbol{Y})}{\partial \lambda_j(\boldsymbol{Y})} = 1.$$

The gradient of the log-determinant is largest in the direction of the smallest eigenvalue of the error covariance matrix. In contrast, the MSE gives equal weight to all directions of the space. The latter yields balanced designs that are similar to the former only if those are forced to be balanced by the problem structure. Even for random dynamical systems with 10 states and 10 outputs (see Section VI for details), the MSE error achieved by a sensing set ($s = 3$) selected by minimizing the $\log\det$ was up to 13% larger than a sensing set selected by minimizing the MSE itself (100 realizations).

III. Approximately Supermodular Function Minimization

Supermodularity (*submodularity*) encodes a "diminishing returns" property of certain set functions that implies near-optimality bounds on their greedy minimization (maximization). Well-known representatives of this class include the rank or $\log\det$ of a sum of PSD matrices, the Shannon entropy, and the mutual information [33], [34]. Still, supermodularity is a stringent condition. In particular, it does not hold for the objectives of (PI) or (PII) [12]–[14], [33].

The purpose of *approximate supermodularity* (*submodularity*) is to relax the original "diminishing returns" property while controlling the magnitude of the violations. The rationale is that if a function is "almost" supermodular, then it should behave similar to a supermodular function. In what follows, we formalize and quantify these statements.

Consider a set function $f : 2^{\mathcal{O}} \to \mathbb{R}$ whose value for an arbitrary set $\mathcal{X} \subseteq \mathcal{O}$ is denoted by $f(\mathcal{X})$. We say f is *normalized* if $f(\emptyset) = 0$ and f is *monotone decreasing* if for all sets $\mathcal{A} \subseteq \mathcal{B} \subseteq \mathcal{O}$ it holds that $f(\mathcal{A}) \geq f(\mathcal{B})$. Observe that if a function is normalized and monotone decreasing, it must be that $f(\mathcal{X}) \leq 0$ for all $\mathcal{X} \subseteq \mathcal{O}$. Define

$$\Delta_u f(\mathcal{X}) = f(\mathcal{X}) - f(\mathcal{X} \cup \{u\}) \tag{8}$$

to be the variation in the value of f incurred by adding the element $u \in \mathcal{O} \setminus \mathcal{X}$ to the set $\mathcal{X}$. Then, a set function f is *supermodular* if for all sets $\mathcal{A} \subseteq \mathcal{B} \subseteq \mathcal{O}$ and elements $u \in \mathcal{O} \setminus \mathcal{B}$, it holds that

$$\Delta_u f(\mathcal{A}) \geq \Delta_u f(\mathcal{B}). \tag{9}$$

A function f is *submodular* if $-f$ is supermodular.

The relevance of supermodular functions in this article is due to the celebrated bound on the suboptimality of their greedy minimization [30]. Specifically, consider the generic cardinality constrained optimization problem

$$\mathcal{X}^\star \in \operatorname*{argmin}_{|\mathcal{X}| \leq s} f(\mathcal{X}) \tag{PIV}$$

and construct its greedy solution by starting with $\mathcal{G}_0 = \emptyset$ and incorporating the elements from $\mathcal{O}$ one at a time so as to maximize the gain at each step. Explicitly, at step j, we do

$$\mathcal{G}_{j+1} = \mathcal{G}_j \cup \{u\} \quad \text{with} \quad u = \operatorname*{argmin}_{w \in \mathcal{O} \setminus \mathcal{G}_j} f(\mathcal{G}_j \cup \{w\}). \tag{10}$$

The recursion in (10) is repeated for s steps to obtain a greedy solution with s elements. If f is monotone decreasing and supermodular [30], then

$$f(\mathcal{G}_s) \leq (1 - e^{-1}) f(\mathcal{X}^\star). \tag{11}$$

The guarantee in (11), however, no longer applies when f is not supermodular. To provide guarantees in these cases, we leverage two measures of approximate supermodularity and derive near-optimality bounds for each of them. It is worth noting that though intuitive, such results are not trivial. In fact, Horel and Singer[39] showed that for another measure of proximity, functions δ-close to supermodular cannot be optimized in polynomial time unless δ is small.

We start with a multiplicative relaxation of the supermodular property (9).

Definition 1 (α-Supermodularity): A set function $f : 2^{\mathcal{O}} \to \mathbb{R}$ is *α-supermodular*, for $\alpha \in \mathbb{R}$, if for all sets $\mathcal{A} \subseteq \mathcal{B} \subseteq \mathcal{O}$ and all $u \in \mathcal{O} \setminus \mathcal{B}$ it holds that

$$\Delta_u f(\mathcal{A}) \geq \alpha\, \Delta_u f(\mathcal{B}). \tag{12}$$

For $\alpha \geq 1$, (12) reduces the original definition of supermodularity (9), in which case we refer to the function simply as supermodular [33], [34]. On the other hand, when $\alpha < 1$, f is said to be *approximately supermodular*. Notice that if f is decreasing, then (12) always holds for $\alpha = 0$. We are therefore interested in the largest α for which (12) holds, i.e.,

$$\alpha = \min_{\substack{\mathcal{A} \subset \mathcal{B} \subseteq \mathcal{O} \\ u \in \mathcal{O} \setminus \mathcal{B}}} \frac{\Delta_u f(\mathcal{A})}{\Delta_u f(\mathcal{B})}. \tag{13}$$

This concept first appeared in the context of auction design [40], although it has been rediscovered in the context of discrete optimization, estimation, and control [1], [41]–[43]. It is worth noting that α in Definition 1 is also related to the *submodularity ratio* γ introduced in [11]. However, the proposed bounds on γ depended on the sparse eigenvalues of a matrix, that are NP-hard to compute. The first explicit (P-computable) bounds on α were obtained in [1], [42], though the same bounds were more recently derived for γ for a special case of (PI) ($m = 1$, $\boldsymbol{M}_\emptyset = \sigma^2 \boldsymbol{I}$, and rank-one $\boldsymbol{M}_u$) [43]. These results were particularized for the problem of controlling linear systems in [44].

Authorized licensed use limited to: University of Pennsylvania. Downloaded on October 29,2026 at 15:10:48 UTC from IEEE Xplore. Restrictions apply.

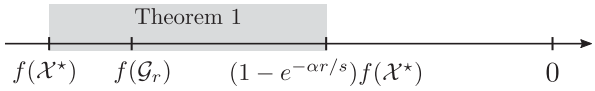


Fig. 1. Illustration of the near-optimal guarantee from Theorem 1.

Interestingly, α not only measures how much f violates supermodularity, but it also quantifies the loss in performance guarantee incurred from these violations.

Theorem 1: Let f be a normalized, monotone decreasing, and α-supermodular set function (i.e., $f(\mathcal{X}) \leq 0$ for all $\mathcal{X} \subseteq \mathcal{O}$). Then, the solution obtained after r steps of the greedy algorithm in (10) obeys

$$f(\mathcal{G}_r) \leq (1 - e^{-\alpha r/s}) f(\mathcal{X}^\star). \tag{14}$$

Proof: See Appendix. ■

Theorem 1 bounds the suboptimality of the greedy solution from (10) when its objective is α-supermodular. Indeed, since f is a nonpositive function, it guarantees that $f(\mathcal{G}_r)$ cannot be too far from the optimal value (Fig. 1). At the same time, it quantifies the effect of relaxing the supermodularity hypothesis typically used to provide performance guarantees in these settings. In fact, if f is supermodular ($\alpha = 1$) and $r = s$, we recover the guarantee (11) from [30]. On the other hand, for an approximately supermodular function ($\alpha < 1$), the result in (14) shows that the same 63% guarantee is recovered by greedily selecting a set of size s/α. Hence, α not only measures how much f violates supermodularity, but also gives a factor by which a solution set must increase to maintain supermodular-like near-optimality. It is worth noting that, as with the original bound in [30], (14) is not tight and that better results are typically obtained in practice (see Section VI).

Although α-supermodularity yields a multiplicative approximation factor, finding meaningful bounds on α can be challenging for certain set functions, such as the objective of (PII). It is therefore useful to look at approximate supermodularity from a different perspective, as proposed in [45].

Definition 2 (ϵ-supermodularity): A set function $f : 2^{\mathcal{O}} \to \mathbb{R}$ is *ϵ-supermodular*, for $\epsilon \in \mathbb{R}$, if for all multisets $\mathcal{A} \subseteq \mathcal{B} \subseteq \mathcal{O}$ and all $u \in \mathcal{O} \setminus \mathcal{B}$ it holds that

$$\Delta_u f(\mathcal{A}) \geq \Delta_u f(\mathcal{B}) - \epsilon. \tag{15}$$

Again, we say f is supermodular if $\epsilon \leq 0$, and approximately supermodular otherwise. As with α, we want the best ϵ that satisfies (15), which is given by

$$\epsilon = \max_{\substack{\mathcal{A} \subseteq \mathcal{B} \subseteq \mathcal{O} \\ u \in \mathcal{O} \setminus \mathcal{B}}} \Delta_u f(\mathcal{B}) - \Delta_u f(\mathcal{A}). \tag{16}$$

In contrast to α-supermodularity, we obtain an additive approximation guarantee for the greedy minimization of ϵ-supermodular functions.

Theorem 2: Let f be a normalized, monotone decreasing, and ϵ-supermodular set function (i.e., $f(\mathcal{X}) \leq 0$ for all $\mathcal{X} \subseteq \mathcal{O}$). Then, the solution obtained after r steps of the greedy algorithm in (10) obeys

$$f(\mathcal{G}_r) \leq (1 - e^{-r/s}) \left[f(\mathcal{X}^\star) + s \cdot \epsilon \right]. \tag{17}$$

Proof: See Appendix. ■

As before, ϵ quantifies the loss in performance guarantee due to relaxing supermodularity. Indeed, (17) reveals that ϵ-supermodular functions have the same guarantees as a supermodular function up to an additive factor of $\Theta(s\epsilon)$. In fact, if $\epsilon \leq (es)^{-1} |f(\mathcal{X}^\star)|$ (recall that $f(\mathcal{X}^\star) \leq 0$ due to normalization), then taking $r = 3$ s recovers the supermodular 63% approximation factor. This same factor is obtained for $(\alpha \geq 1/3)$-supermodular functions.

Although Theorems 1 and 2 characterize the loss in suboptimality incurred from violating supermodularity, their performance certificates depend on the specific values of α and ϵ. However, (13) and (16) reveal that finding α and ϵ for a general function is a combinatorial problem. To give actual near-optimal guarantees for greedy solution of the sensor selection problems (PI) and (PII), the next section bounds the values of α and ϵ for different scalarizations of (4). Notice that we do not need to tackle the objectives of (PI) and (PII) directly, thanks to the following lemma.

Lemma 1: Consider the set functions $f_i : 2^{\mathcal{O}} \to \mathbb{R}$, $i \in \mathbb{N}$. Then, for $\theta_i \geq 0$ and $b \in \mathbb{R}$,

i) if the f_i are α_i-supermodular, then $g = \sum_i \theta_i f_i + b$ is $\min(\alpha_i)$-supermodular;

ii) if the f_i are ϵ_i-supermodular, then $g = \sum_i \theta_i f_i + b$ is $(\sum_i \theta_i \epsilon_i)$-supermodular.

Proof: See Appendix. ■

IV. Approximately Supermodular Scalarizations

Theorems 1 and 2 apply to set functions that are i) normalized, ii) monotone decreasing, and iii) approximately supermodular. By construction, the objective of (P*) is normalized [i)]. In this section, we obtain properties ii) and iii) by leveraging Lemma 1 and focusing only on the scalarizations underlying the objectives of (PI) and (PII), namely the trace

$$t(\mathcal{X}) = \operatorname{Tr}[\boldsymbol{Y}_k(\mathcal{X})] \tag{18}$$

and the spectral norm

$$e(\mathcal{X}) = \|\boldsymbol{Y}_k(\mathcal{X})\| \tag{19}$$

for $\boldsymbol{Y}_k$ defined as in (4).

To establish that set functions t and e are monotone decreasing [ii)], we prove that $\boldsymbol{Y}_k(\mathcal{X})$ is a decreasing set function in the PSD cone (Lemma 2). The definition of Loewner order and the monotonicity of the trace [46] imply the desired result. The monotone decreasing property of the objectives of (PI) and (PII) then follows immediately from $\theta_j \geq 0$.

Lemma 2: The matrix-valued set function $\boldsymbol{Y}_k(\mathcal{X})$ in (4) is monotonically decreasing with respect to the PSD cone, i.e., $\mathcal{A} \subseteq \mathcal{B} \subseteq \mathcal{O} \Rightarrow \boldsymbol{Y}_k(\mathcal{A}) \succeq \boldsymbol{Y}_k(\mathcal{B})$.

Proof: The monotone decreasing nature of $\boldsymbol{Y}_k$ stems from the fact that matrix inversion is an operator antitone function, i.e., that for $\boldsymbol{X}, \boldsymbol{Y} \succeq 0$, it holds that $\boldsymbol{X} \succeq \boldsymbol{Y} \Leftrightarrow \boldsymbol{X}^{-1} \preceq \boldsymbol{Y}^{-1}$ [47]. To see this, write (4) as $\boldsymbol{Y}_k(\mathcal{X}) = \boldsymbol{R}(\mathcal{X})^{-1}$, where $\boldsymbol{R}(\mathcal{X}) = \boldsymbol{M}_{\emptyset,k} + \sum_{u \in \mathcal{X}} \boldsymbol{M}_{u,k}$. Then, notice that since $\boldsymbol{R}$ is a sum of PSD matrices, it holds that $\boldsymbol{R}(\mathcal{A}) \succeq 0$ for all $\mathcal{A} \subseteq \mathcal{O}$. Moreover, $\boldsymbol{R}$ is a modular (additive) function, i.e., $\boldsymbol{R}(\mathcal{A} \cup \mathcal{B}) = \boldsymbol{R}(\mathcal{A}) + \boldsymbol{R}(\mathcal{B})$. Hence, for $\mathcal{A} \subseteq \mathcal{B} \subseteq \mathcal{O}$, we obtain

$$\boldsymbol{R}(\mathcal{B}) = \boldsymbol{R}(\mathcal{A}) + \boldsymbol{R}(\mathcal{B} \setminus \mathcal{A}) \succeq \boldsymbol{R}(\mathcal{A}). \tag{20}$$

Authorized licensed use limited to: University of Pennsylvania. Downloaded on October 29,2026 at 15:10:48 UTC from IEEE Xplore. Restrictions apply.

It is straightforward to obtain from (20) that for $\mathcal{A} \subseteq \mathcal{B} \subseteq \mathcal{O}$

$$\boldsymbol{R}(\mathcal{A}) \preceq \boldsymbol{R}(\mathcal{B}) \Leftrightarrow \boldsymbol{R}(\mathcal{A})^{-1} \succeq \boldsymbol{R}(\mathcal{B})^{-1} \Leftrightarrow \boldsymbol{Y}_k(\mathcal{A}) \succeq \boldsymbol{Y}_k(\mathcal{B}).$$

■

The remainder of this section is dedicated to showing that the trace and the spectral norm are approximately supermodular scalarizations of $\boldsymbol{Y}_k$, starting with the set trace function (18). In what follows, we omit the dependence of $\boldsymbol{Y}$, $\boldsymbol{M}_\emptyset$, and $\boldsymbol{M}_u$ on k for clarity.

Theorem 3: Let t be the set function in (18). Then, t is α-supermodular with

$$\alpha \geq \frac{\mu_{\min}}{\mu_{\max}} > 0 \tag{21}$$

where

$$0 < \mu_{\min} \leq \lambda_{\min}\left[\boldsymbol{M}_\emptyset\right] \leq \lambda_{\max}\left[\boldsymbol{M}_\emptyset + \sum_{u\in\mathcal{O}} \boldsymbol{M}_u\right] \leq \mu_{\max}.$$

Remark 2: Although there exist examples for which $t(\mathcal{X})$ is not supermodular, the general statement of Theorem 3 does not allow us to claim that $\alpha < 1$. A simple counterexample involves the case in which $\boldsymbol{M}_\emptyset = \mu_0 \boldsymbol{I}$ and $\boldsymbol{M}_u = \mu_u \boldsymbol{I}$, $\mu_0, \mu_u \geq 0$, so that t becomes effectively a scalar function of μ_0, μ_u. Since scalar convex functions of positive modular functions are supermodular [48], we have $\alpha \geq 1$ in this case.

The proof of Theorem 3 relies on the following bounds on the variation $\Delta_u t(\mathcal{X})$:

Lemma 3: For all $\mathcal{X} \subset \mathcal{O}$ and $u \in \mathcal{O} \setminus \mathcal{X}$, it holds that

$$\begin{aligned}\lambda_{\min}\left[\boldsymbol{Y}(\mathcal{X})\right] \operatorname{Tr}\left[\boldsymbol{M}_u \boldsymbol{Y}\left(\mathcal{X}\cup\{u\}\right)\right] &\leq \Delta_u t(\mathcal{X}) \\ &\leq \lambda_{\max}\left[\boldsymbol{Y}(\mathcal{X})\right] \operatorname{Tr}\left[\boldsymbol{M}_u \boldsymbol{Y}\left(\mathcal{X}\cup\{u\}\right)\right].\end{aligned} \tag{22}$$

Proof: See Appendix. ■

Theorem 3 then follows readily from this technical lemma.

Proof of Theorem 3: Notice from (8) and (13) that α can be written as

$$\alpha = \min_{\substack{\mathcal{A}\subset\mathcal{B}\subseteq\mathcal{O} \\ u\in\mathcal{O}\setminus\mathcal{B}}} \frac{\Delta_u t(\mathcal{A})}{\Delta_u t(\mathcal{B})}.$$

From Lemma 3, we then obtain

$$\alpha \geq \frac{\lambda_{\min}\left[\boldsymbol{Y}(\mathcal{A})\right]}{\lambda_{\max}\left[\boldsymbol{Y}(\mathcal{B})\right]} \times \frac{\operatorname{Tr}\left[\boldsymbol{M}_u \boldsymbol{Y}\left(\mathcal{A}\cup\{u\}\right)\right]}{\operatorname{Tr}\left[\boldsymbol{M}_u \boldsymbol{Y}\left(\mathcal{B}\cup\{u\}\right)\right]}. \tag{23}$$

To simplify (23), recall from Lemma 2 that $\boldsymbol{Y}$ is monotone decreasing in the PSD cone. Since $\mathcal{A} \subset \mathcal{B} \subseteq \mathcal{O}$, it holds that $\boldsymbol{Y}(\mathcal{A}\cup\{u\}) \succeq \boldsymbol{Y}(\mathcal{B}\cup\{u\})$. The ordering of the PSD cone (Loewner order) gives us that the second term in (23) is always greater than one, which yields

$$\alpha \geq \frac{\lambda_{\min}\left[\boldsymbol{Y}(\mathcal{A})\right]}{\lambda_{\max}\left[\boldsymbol{Y}(\mathcal{B})\right]}.$$

The lower bound in (21) is readily obtained by observing that the decreasing nature of $\boldsymbol{Y}$ implies that for any set $\mathcal{X} \subseteq \mathcal{O}$:

$$\lambda_{\min}\left[\boldsymbol{Y}(\mathcal{O})\right] \leq \lambda_{\min}\left[\boldsymbol{Y}(\mathcal{X})\right] \leq \lambda_{\max}\left[\boldsymbol{Y}(\mathcal{X})\right] \leq \lambda_{\max}\left[\boldsymbol{Y}(\emptyset)\right].$$

■

Theorem 3 gives a deceptively simple bound on the α-supermodularity of the set function t in (18) depending on the spectrum of the $\boldsymbol{M}_\emptyset, \boldsymbol{M}_i$. This bound can be interpreted geometrically in terms of the "range" of the error covariance matrix $\boldsymbol{Y}$. To see this, define the *numerical range* of the set function $\boldsymbol{Y}$ as

$$W_{\mathcal{O}}(\boldsymbol{Y}) = W\left[\bigoplus_{\mathcal{X}\subseteq\mathcal{O}} \boldsymbol{Y}(\mathcal{X})\right] \tag{24}$$

where $\boldsymbol{A} \oplus \boldsymbol{B} = \operatorname{blkdiag}(\boldsymbol{A}, \boldsymbol{B})$ is the direct sum of matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ and $W(\boldsymbol{M}) = \{\boldsymbol{x}^T \boldsymbol{M} \boldsymbol{x} \mid \|\boldsymbol{x}\|_2 = 1\}$ is the classical numerical range [46]. Since the numerical range is a convex set, we can define its relative diameter as

$$\Delta = \max_{\mu,\eta\in W_{\mathcal{O}}(\boldsymbol{Y})} \left|\frac{\mu-\eta}{\mu}\right|. \tag{25}$$

Then, the following holds:

Proposition 3: The set functions t in (18) is α-supermodular with

$$\alpha \geq 1 - \Delta$$

where Δ is the relative diameter of the numerical range of $\boldsymbol{Y}$ in (25).

Proof: Since $\boldsymbol{Y} \succ 0$, the numerical range in (24) is the bounded convex hull of the eigenvalues of $\boldsymbol{Y}(\mathcal{X})$ for all $\mathcal{X} \subseteq \mathcal{O}$ [46]. We can therefore simplify (25) using the fact that it is monotonically increasing in μ and decreasing in η. Explicitly,

$$\Delta = \max_{\mathcal{X},\mathcal{Y}\subseteq\mathcal{O}} \left|\frac{\lambda_{\max}\left[\boldsymbol{Y}(\mathcal{Y})\right] - \lambda_{\min}\left[\boldsymbol{Y}(\mathcal{X})\right]}{\lambda_{\max}\left[\boldsymbol{Y}(\mathcal{Y})\right]}\right|.$$

Using the fact that $\boldsymbol{Y}$ is monotonically decreasing (Lemma 2), this maximum is achieved for

$$\Delta = \frac{\lambda_{\max}\left[\boldsymbol{Y}(\emptyset)\right] - \lambda_{\min}\left[\boldsymbol{Y}(\mathcal{O})\right]}{\lambda_{\max}\left[\boldsymbol{Y}(\emptyset)\right]}.$$

The bound in (21) thus becomes

$$\alpha \geq \frac{\lambda_{\min}\left[\boldsymbol{Y}(\mathcal{O})\right]}{\lambda_{\max}\left[\boldsymbol{Y}(\emptyset)\right]} = 1 - \Delta.$$

■

Hence, (21) bounds how much t deviates from a supermodular function (as quantified by α) in terms of the numerical range of its underlying function $\boldsymbol{Y}$. The shorter the range of $\boldsymbol{Y}$, the more supermodular-like (18) will be.

Notice from Lemma 1, that evaluating the α-supermodularity of the objective of (PI) is a straightforward corollary of Theorem 3.

Corollary 1: The objective of (PI) is α-supermodular for

$$\alpha \geq \min_{0\leq k\leq N-1} \frac{\lambda_{\min}\left[\boldsymbol{M}_{\emptyset,m+k}\right]}{\lambda_{\max}\left[\boldsymbol{M}_{\emptyset,m+k} + \sum_{u\in\mathcal{O}} \boldsymbol{M}_{u,m+k}\right]}. \tag{26}$$

Proceeding, we now bound the ϵ-supermodularity of (19), once again temporarily omitting the dependences on k.

Theorem 4: Let e be defined as in (19). Then, e is ϵ-supermodular with

$$\epsilon \leq \frac{\lambda_{\max}\left(\sum_{u\in\mathcal{O}} \boldsymbol{M}_u\right)}{\lambda_{\min}\left[\boldsymbol{M}_\emptyset\right]^2}. \tag{27}$$

Proof: The proof follows a homotopy argument, i.e., we define a continuous map between $\Delta_u e(\mathcal{A})$ and $\Delta_u e(\mathcal{B})$ and

Authorized licensed use limited to: University of Pennsylvania. Downloaded on October 29,2026 at 15:10:48 UTC from IEEE Xplore. Restrictions apply.

bound its derivative using spectral bounds on the sum of Hermitian matrices. The inequality in (27) then follows from the fundamental theorem of calculus.

Start by defining the homotopy, with $t \in [0,1]$,

$$h_{\mathcal{AB}}(t) = \left\|\boldsymbol{Z}(t)^{-1}\right\| - \left\|\left(\boldsymbol{Z}(t)+\boldsymbol{M}_u\right)^{-1}\right\| \tag{28}$$

for $\mathcal{A} \subset \mathcal{B} \subseteq \mathcal{O}$, $\boldsymbol{Z}(t) = \boldsymbol{R}(\mathcal{A}) + t[\boldsymbol{R}(\mathcal{B}) - \boldsymbol{R}(\mathcal{A})]$, and $\boldsymbol{R}(\mathcal{X}) = \boldsymbol{M}_\emptyset + \sum_{u\in\mathcal{X}} \boldsymbol{M}_u$. Note that since $\boldsymbol{Y}(\mathcal{X}) = \boldsymbol{R}(\mathcal{X})^{-1}$, we have $h_{\mathcal{AB}}(0) = \Delta_u e(\mathcal{A})$ and $h_{\mathcal{AB}}(1) = \Delta_u e(\mathcal{B})$. If $\dot{h}_{\mathcal{AB}}(t)$ is the derivative of $h_{\mathcal{AB}}$ with respect to t, it therefore holds that

$$\Delta_u e(\mathcal{B}) = \Delta_u e(\mathcal{A}) + \int_0^1 \dot{h}_{\mathcal{AB}}(t)dt. \tag{29}$$

Comparing (29) to the definition of ϵ in (16), we obtain

$$\epsilon = \max_{\substack{\mathcal{A}\subset\mathcal{B}\subseteq\mathcal{O}\\ u\in\mathcal{O}\setminus\mathcal{B}}} \int_0^1 \dot{h}_{\mathcal{AB}}(t)dt. \tag{30}$$

We now proceed by evaluating $\dot{h}$. We omit the dependence on $\mathcal{A}$ and $\mathcal{B}$ for conciseness. First, recall from matrix analysis that for any $\boldsymbol{X} \succ 0$, we have

$$\begin{aligned}\frac{d}{dt}\left\|\boldsymbol{X}(t)^{-1}\right\| &= \boldsymbol{q}(t)^T\left[\frac{d}{dt}\boldsymbol{X}(t)^{-1}\right]\boldsymbol{q}(t)\\ &= -\boldsymbol{q}(t)^T\boldsymbol{X}(t)^{-1}\dot{\boldsymbol{X}}(t)\boldsymbol{X}(t)^{-1}\boldsymbol{q}(t)\end{aligned} \tag{31}$$

where $\boldsymbol{q}(t)$ is the eigenvector relative to the maximum eigenvalue of $\boldsymbol{X}(t)$. To obtain (31), we used the fact that $\|\boldsymbol{X}\| = \lambda_{\max}(\boldsymbol{X})$ for $\boldsymbol{X} \succeq 0$ and $\frac{d}{dt}\lambda_{\max}[\boldsymbol{X}(t)] = \boldsymbol{q}(t)^T\dot{\boldsymbol{X}}(t)\boldsymbol{q}(t)$. We then used $\frac{d}{dt}\boldsymbol{X}(t)^{-1} = -\boldsymbol{X}(t)^{-1}\dot{\boldsymbol{X}}(t)\boldsymbol{X}(t)^{-1}$ [47]. In view of (28) and (31), we obtain

$$\begin{aligned}\dot{h}(t) = {}& \boldsymbol{w}(t)^T\left[\boldsymbol{Z}(t)+\boldsymbol{M}_u\right]^{-1}\left[\boldsymbol{R}(\mathcal{B})-\boldsymbol{R}(\mathcal{A})\right]\\ &\times\left[\boldsymbol{Z}(t)+\boldsymbol{M}_u\right]^{-1}\boldsymbol{w}(t)\\ &-\boldsymbol{u}(t)^T\boldsymbol{Z}(t)^{-1}\left[\boldsymbol{R}(\mathcal{B})-\boldsymbol{R}(\mathcal{A})\right]\boldsymbol{Z}(t)^{-1}\boldsymbol{u}(t)\end{aligned} \tag{32}$$

where $\boldsymbol{u}(t)$ and $\boldsymbol{w}(t)$ are the eigenvectors relative to the maximum eigenvalues of $\boldsymbol{Z}(t)^{-1}$ and $[\boldsymbol{Z}(t)+\boldsymbol{M}_u]^{-1}$, respectively.

We now proceed by finding a bound for $\dot{h}$ in (32) that is independent of t. To do so, observe that $\boldsymbol{R}(\mathcal{B}) \succeq \boldsymbol{R}(\mathcal{A})$, since $\mathcal{A} \subseteq \mathcal{B}$ and $\boldsymbol{R}$ is monotone increasing (Lemma 2). By the Loewner ordering, the second term in (32) is therefore nonpositive. Thus, using Rayleigh's spectral inequality and the fact $\|\boldsymbol{w}(t)\| = 1$ for all t yields

$$\dot{h}(t) \leq \left\|\left(\boldsymbol{Z}(t)+\boldsymbol{M}_u\right)^{-1}\left[\boldsymbol{R}(\mathcal{B})-\boldsymbol{R}(\mathcal{A})\right]\left(\boldsymbol{Z}(t)+\boldsymbol{M}_u\right)^{-1}\right\|.$$

Next, we once again use the fact that $\boldsymbol{R}(\mathcal{B}) \succeq \boldsymbol{R}(\mathcal{A})$ to obtain $\boldsymbol{Z}(t) \succeq \boldsymbol{Z}(0) = \boldsymbol{R}(\mathcal{A})$, effectively removing the dependence on t. Using Cauchy–Schwartz then yields

$$\dot{h}(t) \leq \lambda_{\max}\left[\left(\boldsymbol{R}(\mathcal{A})+\boldsymbol{M}_u\right)^{-2}\right]\lambda_{\max}\left[\boldsymbol{R}(\mathcal{B})-\boldsymbol{R}(\mathcal{A})\right]$$

which can be used in (30) to get

$$\epsilon \leq \max_{\substack{\mathcal{A}\subset\mathcal{B}\subseteq\mathcal{O}\\ u\in\mathcal{O}\setminus\mathcal{B}}} \frac{\lambda_{\max}\left[\boldsymbol{R}(\mathcal{B})-\boldsymbol{R}(\mathcal{A})\right]}{\lambda_{\min}\left[\boldsymbol{R}(\mathcal{A})+\boldsymbol{M}_u\right]^2} \tag{33}$$

Algorithm 1: Cardinality r Greedy Set Selection for (P*).

Initialize $\mathcal{G}_0 \leftarrow \emptyset$
for $j = 1, \ldots, r$
$u \leftarrow \arg\min_{w\in\mathcal{O}\setminus\mathcal{G}_j} \sum_{k=0}^{N-1} \theta_k\, h[\boldsymbol{Y}_{m+k}(\mathcal{G}_j \cup \{w\})]$
$\mathcal{G}_{j+1} \leftarrow \mathcal{G}_j \cup \{u\}$
end

The inequality in (27) is then obtained using $\|\boldsymbol{R}(\mathcal{B}) - \boldsymbol{R}(\mathcal{A})\| \leq \lambda_{\max}(\sum_{u\in\mathcal{O}} \boldsymbol{M}_u)$ and $\lambda_{\min}[\boldsymbol{R}(\mathcal{A}) + \boldsymbol{M}_u] \geq \lambda_{\min}[\boldsymbol{M}_\emptyset]$. ■

From Theorem 4 and Lemma 1, we obtain Corollary 2.

Corollary 2: The objective of (PII) is ϵ-supermodular for

$$\epsilon \leq \sum_{k=0}^{N-1} \theta_k \frac{\lambda_{\max}(\sum_{u\in\mathcal{O}} \boldsymbol{M}_{u,m+k})}{\lambda_{\min}\left[\boldsymbol{M}_{\emptyset,m+k}\right]^2}. \tag{34}$$

In the sequel, we particularize these results for the estimation problems i) and ii) in Section II, showing that we obtain supermodular-like guarantees when a certain measure of signal-to-noise ratio (SNR) is small. We also argue that this is in fact the typical setting in which KFs are useful in practice.

V. Near-Optimal Sensor Selection

In this section, we obtain approximation guarantees for the greedy optimization of the MSE and worst-case error in filtering and smoothing problems. More specifically, we provide near-optimal certificates for the solution $\mathcal{G}_r$ obtained using Algorithm 1 with $\boldsymbol{Y}$ taken from Proposition 1 (filtering) or Proposition 2 (smoothing) and with the trace or spectral norm scalarizations h (as in Section II-A).

A. Greedy Sensor Selection for Filtering is Near-Optimal

The main result of this section stems from applying Proposition 1 to Corollaries 1 and 2.

Theorem 5: Let $\boldsymbol{Y}_k(\mathcal{A}) = \boldsymbol{P}_k(\mathcal{A})$, the filtering error covariance from Proposition 1, in (PI)–(PII) and denote their solutions as $\widehat{\mathcal{T}}^\star$ and $\widehat{\mathcal{S}}^\star$, respectively. Let $\widehat{\mathcal{T}}_g$ and $\widehat{\mathcal{S}}_g$ be the r-elements greedy solutions obtained by using Algorithm 1 on (PI) and (PII). Then,

$$\begin{aligned} f_T(\widehat{\mathcal{T}}_g) &\leq (1 - e^{-\hat{\alpha}r/s}) f_T(\widehat{\mathcal{T}}^\star)\\ f_S(\widehat{\mathcal{S}}_g) &\leq (1 - e^{-r/s})\left[f_S(\widehat{\mathcal{S}}^\star) + s\cdot\hat{\epsilon}\right]\end{aligned}$$

for f_T and f_S the objectives of (PI) and (PII), respectively, $\hat{\alpha} \geq \min_{k\leq N-1}\hat{\alpha}_k$, and $\hat{\epsilon} \leq \sum_{k=0}^{N-1}\theta_k\hat{\epsilon}_k$ with

$$\hat{\alpha}_k \triangleq \frac{\lambda_{\min}\left[\boldsymbol{P}^{-1}_{m+k|m+k-1}\right]}{\lambda_{\max}\left[\boldsymbol{P}^{-1}_{m+k|m+k-1} + \sum_{u\in\mathcal{O}}\boldsymbol{V}_u\right]} \tag{35}$$

$$\hat{\epsilon}_k \triangleq \lambda_{\max}\left(\sum_{u\in\mathcal{O}}\boldsymbol{V}_u\right)\lambda_{\max}\left(\boldsymbol{P}_{m+k|m+k-1}\right)^2 \tag{36}$$

where $\boldsymbol{V}_u = \boldsymbol{H}_u^T\boldsymbol{R}_{v,u}^{-1}\boldsymbol{H}_u$.

Theorem 5 shows that greedy sensor selection for state filtering is near-optimal. In fact, it gives explicit bounds on the

Authorized licensed use limited to: University of Pennsylvania. Downloaded on October 29,2026 at 15:10:48 UTC from IEEE Xplore. Restrictions apply.

suboptimality of greedy sensing sets as a function of parameters of the dynamical system. When we do not allow violations of the cardinality constraint, i.e., for $r = s$, and (35) is close to one or (36) is close to zero, the guarantees from Theorem 5 approach the $(1 - 1/e)$-optimality of greedy supermodular minimization.

To see when this occurs specifically, notice that (35) can be rewritten using the Riccati (48) as

$$\hat{\alpha}_k = \frac{\lambda_{\min}\left[\boldsymbol{P}^{-1}_{m+k|m+k-1}\right]}{\lambda_{\max}\left[\boldsymbol{P}^{-1}_{m+k}(\mathcal{O})\right]} = \frac{\lambda_{\min}\left[\boldsymbol{P}_{m+k}(\mathcal{O})\right]}{\lambda_{\max}\left[\boldsymbol{P}_{m+k|m+k-1}\right]}$$

i.e., as the ratio between the a posteriori error using all sensors and the a priori error. Hence, α_j in (35) is large when the a priori and a posteriori error covariance matrices are similar. To get additional insight, consider the particular case in which $\boldsymbol{\Pi}_0 = \boldsymbol{R}_w = \sigma_w^2 \boldsymbol{I}$ and $\boldsymbol{R}_v = \sigma_v^2 \boldsymbol{I}$. Also, suppose that each state is measured directly, i.e., that $\mathcal{O} = \{1, \dots, p\}$ and that $\boldsymbol{H}_u = \boldsymbol{e}_u^T$, $u \in \mathcal{O}$, where $\boldsymbol{e}_u$ is the uth column of a $p \times p$ identity matrix. Then, from (47) and (48), the matrices of interest take the form

$$\boldsymbol{P}_{k|k-1} = \boldsymbol{F}\boldsymbol{P}_{k-1}\boldsymbol{F}^T + \sigma_w^2 \boldsymbol{I} \tag{37a}$$

$$\boldsymbol{P}_k(\mathcal{O}) = \left(\boldsymbol{P}^{-1}_{k|k-1} + \sigma_v^{-2}\boldsymbol{I}\right)^{-1}. \tag{37b}$$

Notice that for (35) to approach one, it must be that $\boldsymbol{P}_k(\mathcal{O}) \approx \boldsymbol{P}_{k|k-1}$. Since $\boldsymbol{P}_{k|k-1}$ depends directly on σ_w^2, it is clear from (37b) that this occurs when $\sigma_v^2 \gg \sigma_w^2$. Additionally, we need $\boldsymbol{P}_{k|k-1}$ to be well-conditioned. When σ_w^2 is large, this stems directly from (37a). On the other hand, if $\sigma_w^2 \ll \|\boldsymbol{F}\|$, then $\boldsymbol{P}_{k|k-1} \approx \boldsymbol{F}\boldsymbol{P}_{k-1}\boldsymbol{F}^T$ and its condition number depends directly on $\boldsymbol{F}$, since $\kappa(\boldsymbol{P}_{k|k-1}) \leq \kappa(\boldsymbol{P}_{k-1})\kappa(\boldsymbol{F})^2$, where κ is the condition number with respect to the spectral norm [46]. A similar argument applies to (36), which in this simplified setting reduces to

$$\hat{\epsilon}_k = \sigma_v^{-2}\,\lambda_{\max}\left(\boldsymbol{F}\boldsymbol{P}_{m+k-1}\boldsymbol{F}^T + \sigma_w^2\boldsymbol{I}\right)^2.$$

Hence, similar conditions as above also guarantee that $\hat{\epsilon}$ is small. In particular, this occurs when σ_v^2 is much larger than σ_w^2 and $\|\boldsymbol{F}\|$.

Figs. 2 and 3 illustrate these observations by showing the values of $\hat{\alpha}$ and $\hat{\epsilon}$, averaged over 100 dynamical systems realizations, for different values of σ_w^2/σ_v^2 and $\|\boldsymbol{F}\|$ (see Section VI for details). Indeed, the guarantees from Theorem 5 are stronger when $\sigma_v^2 \gg \sigma_w^2$, $\sigma_v^2 \gg \|\boldsymbol{F}\|$, and when the system has modes with similar rates ($\kappa(\boldsymbol{F}) \approx 1$). It is important to highlight that this is the scenario in which KFs are most useful: if the process noise dominates the estimation error, the system trajectory is mostly random and the choice of sensing subset has little impact on the estimation performance [31], [49]. Also notice from Fig. 2 that for $\sigma_v^2 > 10\sigma_w^2$ and $\|\boldsymbol{F}\| = 0.3$, we can obtain supermodular-like guarantees for greedy sensor selection by violating the cardinality constraint by 35%. Since guarantees for problem (PII) are additive, meaningful performance bounds are obtained on a narrower range of parameters (Fig. 3).

Remark 3: In [16], it was shown that it is impossible to obtain a nontrivial, universal (system-independent) bound on $\hat{\alpha}$ unless P = NP. To do so, they provide a reduction from the problem of *exact cover by 3-sets* (X3C) that relies on a dynamical system with noiseless outputs. Notice that this is in line with the results from Theorem 5 given that $\hat{\alpha} \to 0$ as $\sigma_v^2 \to 0$, i.e., the guarantees based on approximate supermodularity become vacuous for noise-free outputs. When this is not the case, however, Theorem 5 provides nontrivial near-optimal certificates for greedy sensor selection as a function of the dynamical system parameters.

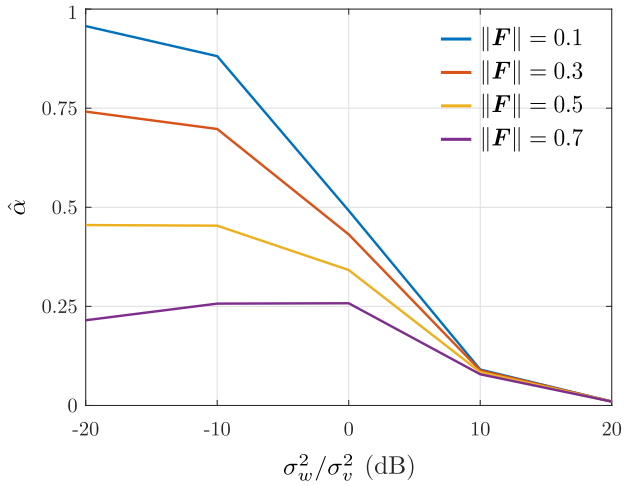


Fig. 2. α-Supermodularity bounds (Theorem 5) for different process noise variances and $\|\boldsymbol{F}\|_2$.

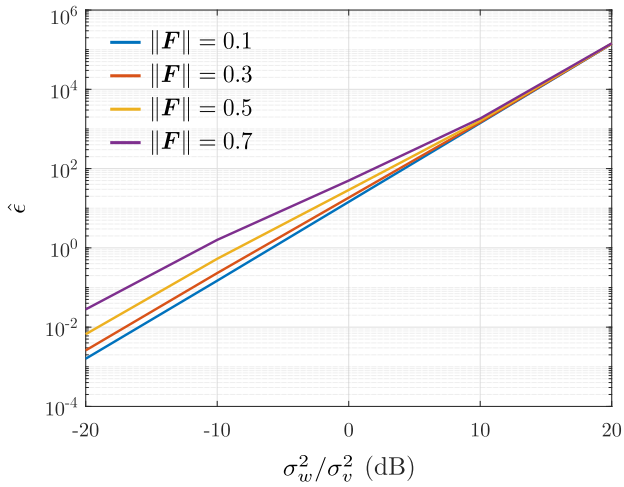


Fig. 3. ϵ-Supermodularity bounds (Theorem 5) for different process noise variances and $\|\boldsymbol{F}\|_2$.

B. Greedy Sensor Selection for Smoothing Is Near-Optimal

We obtain similar results for smoothing by taking the values of $\boldsymbol{M}_{\emptyset,k}$ and $\boldsymbol{M}_{u,k}$ from Proposition 2.

Theorem 6: Let $\boldsymbol{Y}_k(\mathcal{A}) = \boldsymbol{Q}_k(\mathcal{A})$, the smoothing error covariance matrix from Proposition 2, in (PI)–(PII) and take $\widetilde{\mathcal{T}}^\star$ and $\widetilde{\mathcal{S}}^\star$ to be their respective solutions. Let $\widetilde{\mathcal{T}}_g$ and $\widetilde{\mathcal{S}}_g$ be the r-elements greedy solutions obtained by using Algorithm 1

Authorized licensed use limited to: University of Pennsylvania. Downloaded on October 29,2026 at 15:10:48 UTC from IEEE Xplore. Restrictions apply.

on (PI) and (PII). Then,

$$f_T(\widetilde{\mathcal{T}}_g) \leq (1 - e^{-\tilde{\alpha} r/s}) f_T(\widetilde{\mathcal{T}}^\star) \tag{38}$$

$$f_S(\widetilde{\mathcal{S}}_g) \leq (1 - e^{-r/s}) \left[f_S(\widetilde{\mathcal{S}}^\star) + s \cdot \tilde{\epsilon} \right] \tag{39}$$

for f_T and f_S, the objectives of (PI) and (PII), respectively, $\tilde{\alpha} \geq \min_{k \leq N-1} \tilde{\alpha}_k$ and $\tilde{\epsilon} \leq \sum_{k=0}^{N-1} \theta_k \tilde{\epsilon}_k$ with

$$\tilde{\alpha}_k \triangleq \frac{1/\ell_{\max}}{\lambda_{\max}\left[\boldsymbol{C}^{-1} + \sum_{u \in \mathcal{O}} \boldsymbol{\Phi}_{m+k}^T (\boldsymbol{I} \otimes \boldsymbol{V}_u) \boldsymbol{\Phi}_{m+k} \right]} \tag{40}$$

$$\tilde{\epsilon}_k \triangleq \lambda_{\max} \left[\sum_{u \in \mathcal{O}} \boldsymbol{\Phi}_{m+k}^T (\boldsymbol{I} \otimes \boldsymbol{V}_u) \boldsymbol{\Phi}_{m+k} \right] \ell_{\max}^2 \tag{41}$$

where $\boldsymbol{V}_u = \boldsymbol{H}_u^T \boldsymbol{R}_{v,u}^{-1} \boldsymbol{H}_u$, $\boldsymbol{C} = \text{blkdiag}(\boldsymbol{\Pi}_0, \boldsymbol{I} \otimes \boldsymbol{R}_w)$, and $\ell_{\max} = \max[\lambda_{\max}(\boldsymbol{\Pi}_0), \lambda_{\max}(\boldsymbol{R}_w)]$.

Theorem 6 shows that greedy sensor selection is also near-optimal for state smoothing. The suboptimality guarantees, however, vary in strength depending on the parameters of the dynamical system. Nevertheless, as with the filtering case from Section V-A, the certificates improve in typical practical settings. For simplicity, let us again analyze the case where $\boldsymbol{\Pi}_0 = \boldsymbol{R}_w = \sigma_w^2 \boldsymbol{I}$, $\boldsymbol{R}_v = \sigma_v^2 \boldsymbol{I}$, and $\boldsymbol{H}_j = \boldsymbol{e}_j^T$, for $j \in \mathcal{O} = \{1, \ldots, p\}$. Immediately, $\ell_{\max} = \sigma_w^2$, $\boldsymbol{C} = \sigma_w^2 \boldsymbol{I}$, and (40) simplifies to

$$\tilde{\alpha}_k = \frac{1}{1 + \sigma_w^2/\sigma_v^2 \, \lambda_{\max}\left(\boldsymbol{\Phi}_{m+k}^T \boldsymbol{\Phi}_{m+k} \right)}. \tag{42}$$

Notice that a noise ratio appears explicitly in the denominator of (42). Hence, as for filtering, it is ready that $\tilde{\alpha} \approx 1$ when $\sigma_v^2 \gg \sigma_w^2$. Moreover, $\tilde{\alpha}_j$ depends on the system dynamics through the matrix $\boldsymbol{\Phi}$. In fact, this dependence can be made explicit by bounding the norm of $\boldsymbol{\Phi}$ in terms of its blocks. Explicitly, $\lambda_{\max}(\boldsymbol{\Phi}_{m+k}^T \boldsymbol{\Phi}_{m+k}) = \|\boldsymbol{\Phi}_{m+k}\| \leq \sqrt{\sum_{j=0}^{m+k} \|\boldsymbol{F}\|^{2j}}$. Hence, stronger guarantees are obtained when $\sigma_v^2 \gg \|\boldsymbol{F}\|$. Once again, similar arguments hold for (41), which reduces to

$$\tilde{\epsilon}_k = \sigma_w^4/\sigma_v^2 \, \lambda_{\max}\left[\boldsymbol{\Phi}_{m+k}^T \boldsymbol{\Phi}_{m+k} \right]. \tag{43}$$

We illustrate the values of $\tilde{\alpha}$ and $\tilde{\epsilon}$ for different values of σ_w^2/σ_v^2 and $\|\boldsymbol{F}\|$ in Figs. 4 and 5 (see Section VI for details). As (42) and (43) suggest, Theorem 6 provides better guarantees when $\|\boldsymbol{F}\| \sigma_w^2/\sigma_v^2 \ll 1$. In other words, when the measurement noise dominates over both the process noise and the decay rate of the system modes, a scenario of practical value in KF applications [31], [49]. Similar to the filtering case, Fig. 4 shows that if we violate the cardinality constraints by 35%, we can obtain $(1 - 1/e)$-optimality for $\|\boldsymbol{F}\| = 0.3$ and $\sigma_v^2 > 75\sigma_w^2$. As in the filtering case, good performance guarantees for problem (PII) are obtained for a stricter range of parameters (Fig. 5).

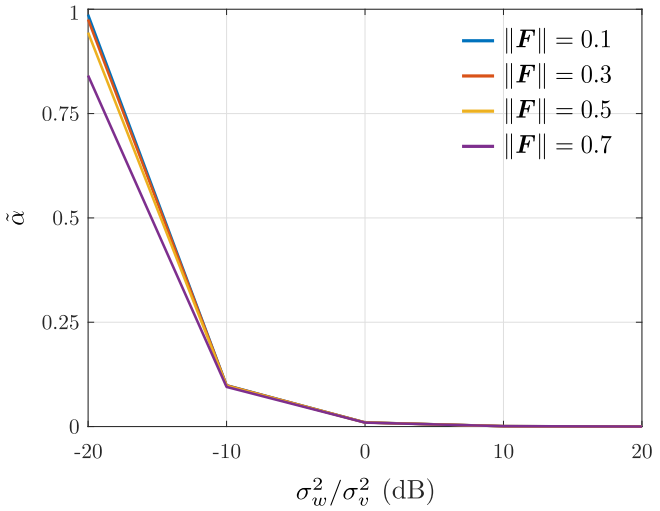


Fig. 4. α-Supermodularity bounds (Theorem 6) for different process noise variances and $\|\boldsymbol{F}\|_2$.

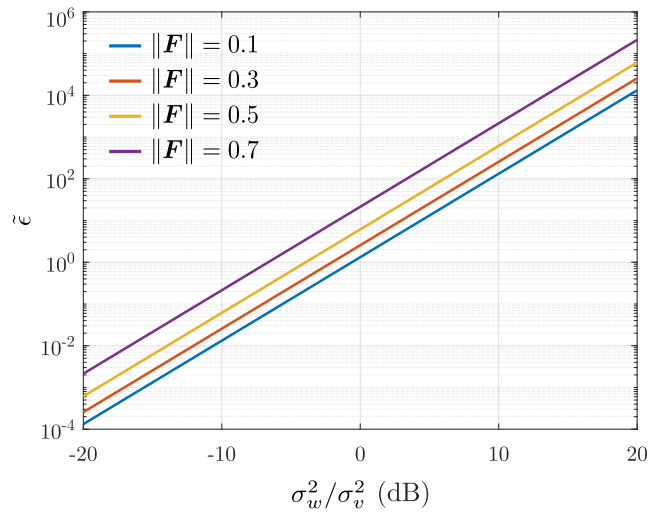


Fig. 5. ϵ-Supermodularity bounds (Theorem 6) for different process noise variances and $\|\boldsymbol{F}\|_2$.

VI. Numerical Examples

We begin by giving details on the simulations in Figs. 2–5. We considered dynamical systems with $n = p = 100$ in which the elements of $\boldsymbol{F}$ were drawn randomly from a standard Gaussian distribution, $\boldsymbol{H}_u = \boldsymbol{e}_u^T$, where $\boldsymbol{e}_u \in \mathbb{R}^p$ is a vector of zeros except for the uth element, which is one, $\boldsymbol{\Pi}_0 = 10^{-2} \boldsymbol{I}$, $\boldsymbol{R}_w = \sigma_w^2 \boldsymbol{I}$, and $\boldsymbol{R}_v = \boldsymbol{I}$. The state transition matrix $\boldsymbol{F}$ was normalized so that its spectral norm is in [0.1,0.9] and the process noise σ_w^2 varied in $[0.01, 100]$. For Figs. 2 and 3 (filtering), we minimize the 10-steps average error [$N = 10$ and $\theta_k = 1$ in (P*)]. For Figs. 4 and 5 (smoothing), we minimize the error at the end of the window [$N = 10$, $\theta_9 = 1$, and $\theta_k = 0$ for $k < 9$ in (P*)].

Figs. 2–5, along with Theorems 5 and 6, illustrate the wide range of parameters over which good performance certificates can be provided for greedy sensor selection. It is worth noting, however, that these are worst-case guarantees and that better results are common in practice. To illustrate this point, we evaluate the relative suboptimality of greedily selected sensing sets for both filtering and smoothing over 1000 system realizations.

Authorized licensed use limited to: University of Pennsylvania. Downloaded on October 29,2026 at 15:10:48 UTC from IEEE Xplore. Restrictions apply.

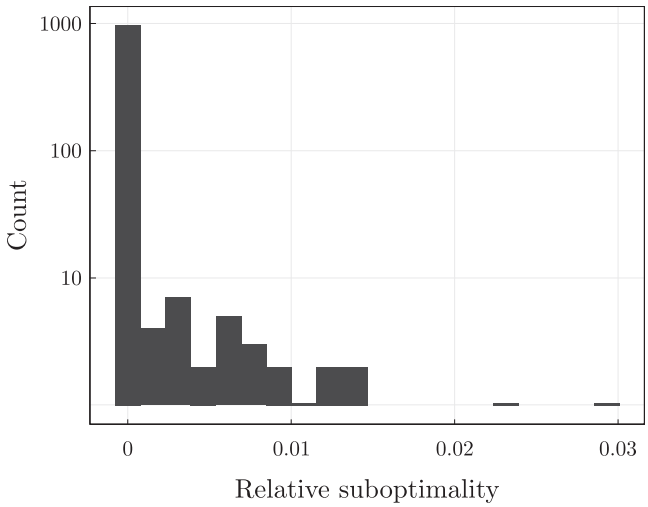


Fig. 6. Relative suboptimality of greedy sensor selection using the MSE [(PI)] for filtering (1000 system realizations).

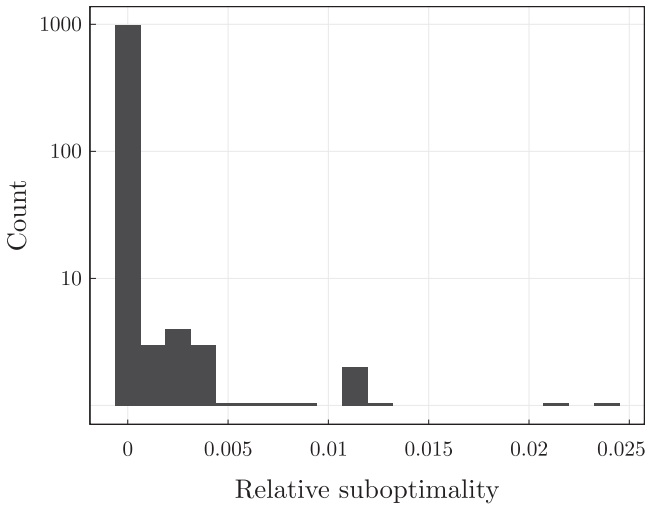


Fig. 7. Relative suboptimality of greedy sensor selection using the MSE [(PI)] for smoothing (1000 system realizations).

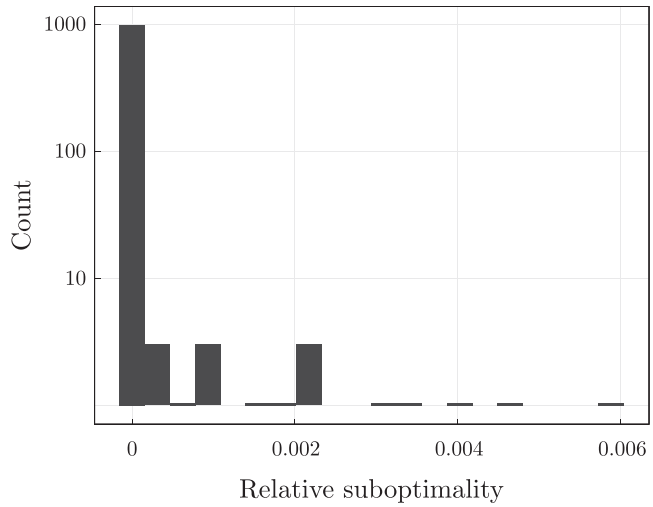


Fig. 8. Relative suboptimality of greedy sensor selection using the spectral norm [(PII)] for filtering (1000 system realizations).

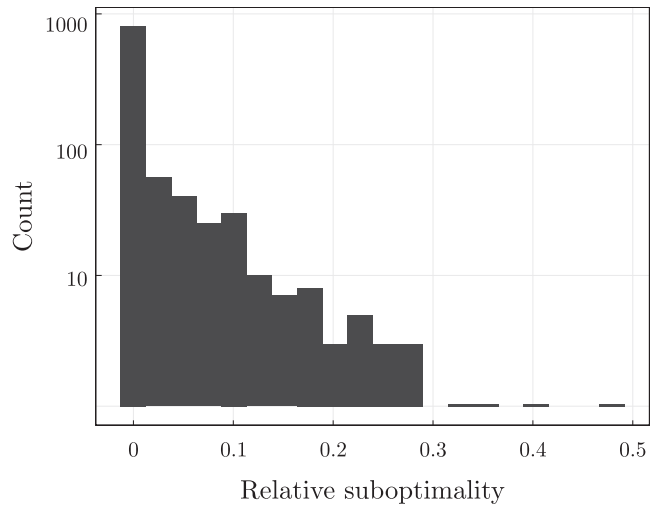


Fig. 9. Relative suboptimality of greedy sensor selection using the spectral norm [(PII)] for smoothing (1000 system realizations).

Explicitly, we evaluate

$$\nu^\star(\mathcal{G}) = \frac{f(\mathcal{X}^\star) - f(\mathcal{G})}{f(\mathcal{X}^\star)}$$

where $\mathcal{G}$ and $\mathcal{X}^\star$ are the greedy and optimal solutions of (P*), respectively. In all experiments, we take $r = s$, i.e., we do not consider violations of the cardinality constraint. Since $\nu^\star$ depends on the optimal sensing set, we restrict ourselves to small dynamical systems [$n = 10$ states, $p = 10$ outputs, and $s = 4$ in P*]. However, we now randomly draw the elements of both $\boldsymbol{F}$ and $\boldsymbol{H}$ from standard Gaussian distributions and normalize $\boldsymbol{F}$ so that its norm is 0.9. We keep $\sigma_w^2 = 10^{-2}$ fixed, but draw the measurement noise variances at each output uniformly at random from $[10^{-2}, 1]$.

In Figs. 6 and 7, we show the value of $\nu^\star$ for the result of greedily solving problem (PI) for filtering and smoothing. Notice that the values obtained are considerably lower than the guarantee given in Theorems 5 and 6. In fact, they are lower than the $1 - e^{-1} \approx 0.63$ bound for supermodular functions. For both problems, greedy sensor selection found the optimal sampling set in over 96% of the realizations.

In Figs. 8 and 9, we show the results for problem (PII). To make sure the system is observable over the time window selected, we use a smaller system [$n = 5$ states, $p = 10$ outputs, and $s = 5$ in (P*)]. Moreover, since the guarantees for the greedy solution of (PII) are additive and therefore not as strong as those for (PI), we take $\sigma_w^2 = 10^{-3}$. For the filtering problem, the results are similar to those observed in Fig. 6 for minimizing the MSE. In over 98% of the realizations, greedy sensor selection actually obtained the optimal sensing set. On the other hand, the smoothing problem is more challenging for the worst-case error and the results obtained are poorer. This is especially due to the fact that $f(\widetilde{\mathcal{S}}^\star) \ll s \cdot \tilde{\epsilon}$ in (39) for this example, so that the

Authorized licensed use limited to: University of Pennsylvania. Downloaded on October 29,2026 at 15:10:48 UTC from IEEE Xplore. Restrictions apply.

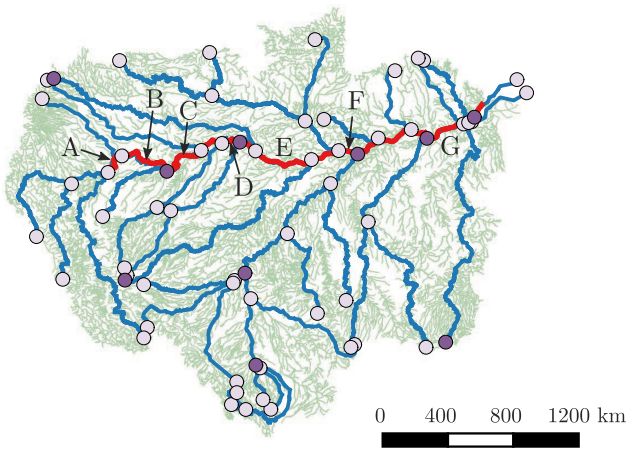


Fig. 10. Amazon basin, sensor pool (light circles), and greedy sensing set (dark circles).

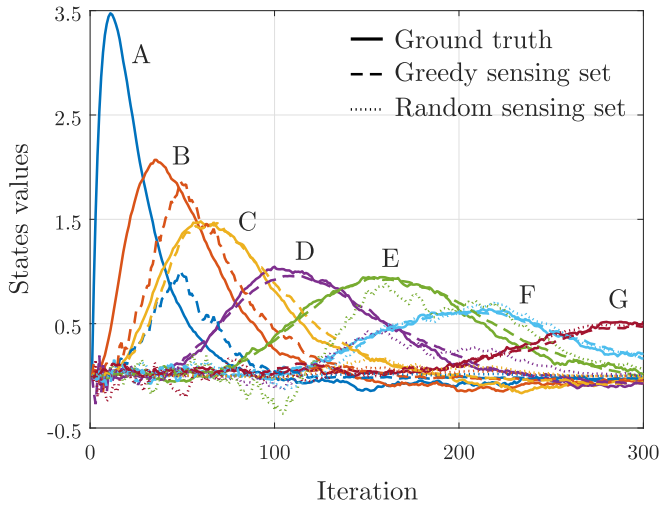


Fig. 11. Estimated unobserved states along the Amazon river. States positions are indicated in Fig. 10.

guarantees we obtain are less strict. Still, greedy sensor selection found the optimal sensing set in 77% of the realizations.

A. Application: Sensing for the Amazon Basin

We conclude these numerical examples by illustrating the use of greedy sensor selection for water monitoring in the Amazon basin. The Amazon drainage basin, shown in green in Fig. 10, covers 7.5 million km^2 and is composed of over 7000 tributaries. In our experiments, we use a simplified network of the basin (blue curve in Fig. 10) obtained by smoothing the original map [50].

Using this river network, we construct a weighted directed tree $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. The set of vertices $\mathcal{V}$ is composed of the possible sensor positions shown in Fig. 10 and an additional node midway between each sensor ($|\mathcal{V}| = 127$). The set of edges $\mathcal{E}$ is composed of ordered pairs from $\mathcal{V} \times \mathcal{V}$ such that $(i, j) \in \mathcal{E}$ if water flows from node i to node j. We assume all river flow toward the ocean. The adjacency matrix of $\mathcal{G}$ is then given by

$$[\boldsymbol{A}]_{ij} = \begin{cases} \exp\left(\|\boldsymbol{z}_i - \boldsymbol{z}_j\|^2 / \sigma^2\right), & \text{if } (i,j) \in \mathcal{E} \\ 0, & \text{otherwise} \end{cases} \tag{44}$$

where $\boldsymbol{z}_i \in \mathbb{R}^2$ is the position of node i on the map and $\sigma^2 = 10$ is a smoothness parameter. We also define its Laplacian as $\boldsymbol{L} = \boldsymbol{D} - \boldsymbol{A}$, where $\boldsymbol{D}$ is a diagonal matrix whose elements are the sum of the columns of $\boldsymbol{A}$, and its symmetrized Laplacian as $\boldsymbol{L}' = \boldsymbol{D}' - (\boldsymbol{A} + \boldsymbol{A}^T)$, where $\boldsymbol{D}'$ is a diagonal matrix whose elements are the sum of the columns of $\boldsymbol{A} + \boldsymbol{A}^T$.

Using $\boldsymbol{A}$ from (44) and the Laplacians defined above, we construct a dynamical system whose states $\boldsymbol{x}_k$ evolve according to (1) with

$$\boldsymbol{F} = 0.9 \exp(-\boldsymbol{L}\Delta t) + 0.099 \exp(-\boldsymbol{L}'\Delta t) \tag{45}$$

where $\Delta t = 0.1$ is the sampling period. The dynamics in (45) are a combination of two processes: the first term corresponds to the advection process by which water flows to the ocean and the second term corresponds to a diffusion process. The combination coefficients are chosen so that the system is marginally stable ($\|\boldsymbol{F}\| < 1.02$). The process noise $\boldsymbol{w}_k$ is drawn from a zero-mean Gaussian random variable with covariances $\boldsymbol{R}_w = \sigma_w^2 \boldsymbol{I}$, for $\sigma_w^2 = 10^{-4}$. The true initial state of the system is a spike at the beginning of the Amazon river (the red curve in Fig. 10), i.e., $\boldsymbol{x}_0$ is a vector of zeros except at position 51, where it is 10. For the KF, however, we assume the initial state is a zero-mean Gaussian random variable with covariance $\boldsymbol{\Pi}_0 = \boldsymbol{I}$.

We greedily select $s = 10$ sensors (dark circles in Fig. 10) from a pool of $p = 64$ sensors whose positions are shown in Fig. 10. The sensors observe a noisy version of their respective states, i.e., $y_{u,k} = e_u^T \boldsymbol{x}_k + v_{u,k}$ for $u \in \mathcal{O}$, where $\boldsymbol{e}_u$ is a vector of zeros except for the uth element that is one and $v_{u,k}$ is a zero-mean Gaussian random variable independent of $v_{w,\ell}$ for all $w \neq u$ or $\ell \neq k$ with variance $\sigma_v^2 = 10^{-1}$.

The state trajectories estimated using Kalman filtering are shown in Fig. 11. We show only the values at seven unobserved portions along the Amazon river (indicated in Fig. 10). We compare the performance of a greedy sensing set selected to optimize the average 200-steps MSE and a sensing set containing one randomly selected sensor for each of the first ten levels of the Amazon basin tree. Notice from Fig. 11 that by portion C, the results using the greedy sensing set track closely the true states of the system, whereas the random sensing only detects the change in the water on the second half of the river (position E). In terms of error, the average 200-steps MSE for a full sensing solution ($\mathcal{X} = \mathcal{O}$) is 185.67. The greedy sensing set, which contains only 15% of the possible sensors, achieves a cost of 1005.03, an almost 50% reduction over the 1978.22 obtained by the random set. Since the modes of this system have similar decays (the condition number of $\boldsymbol{F}$ is 1.3), the sensing set obtained by greedily solving (PIII) differs in a single sensor from that obtained using (PI). Thus, they achieve similar MSE performances. As discussed in Remark 1, this is not necessarily the case in general.

Authorized licensed use limited to: University of Pennsylvania. Downloaded on October 29,2026 at 15:10:48 UTC from IEEE Xplore. Restrictions apply.

VII. Conclusion

This article studied the filtering/smoothing sensor selection problems for both the MSE and the worst-case error and provided near-optimal guarantees to their greedy solutions. To do so, it relied on two concepts of approximate supermodularity, α- and ϵ-supermodularity, and derived approximation bounds for the greedy minimization of these functions. By bounding α and ϵ for the MSE and the spectral norm of the error covariance matrix, it showed that filtering/smoothing sensor selection problem can be solved near-optimally. In fact, in typical application scenarios, the guarantees approach the $1-1/e$ certificate of supermodular functions. These results justify the use of greedy sensor selection without the need for surrogate cost functions, such as the $\log\det$. We expect that other cost functions have similar approximate supermodular behavior and that this theory can be used to give approximation certificates for other control problems. Of particular interest is a formulation that takes the controller inputs into account when designing the sensing set. The underlying theory developed here could also potentially be applied to problems involving joint state/input estimation and/or input and state constraints.

Appendix A
Proof of Proposition 1

Proof: The filtered estimate $\hat{\boldsymbol{x}}_k(\mathcal{X})$ can be obtained from a previous estimate $\hat{\boldsymbol{x}}_{k-1}$ using the KF iteration

$$\hat{\boldsymbol{x}}_k(\mathcal{X}) = \boldsymbol{F}\hat{\boldsymbol{x}}_{k-1} + \sum_{u\in\mathcal{X}} \boldsymbol{K}_{u,k}\left[\boldsymbol{y}_{u,k} - \boldsymbol{H}_u\boldsymbol{F}\hat{\boldsymbol{x}}_{k-1}\right] \tag{46}$$

where $\boldsymbol{K}_{u,k} = \boldsymbol{P}_{k|k-1}\boldsymbol{H}_u^T\boldsymbol{E}_{u,k}^{-1}$ is the Kalman gain of measurement $u\in\mathcal{O}$, with $\boldsymbol{E}_{u,k} = \boldsymbol{H}_u\boldsymbol{P}_{k|k-1}\boldsymbol{H}_u^T + \boldsymbol{R}_{v,u}$ denoting the innovation covariance matrix and

$$\boldsymbol{P}_{k|k-1} = \boldsymbol{F}\boldsymbol{P}_{k-1}\boldsymbol{F}^T + \boldsymbol{R}_w \tag{47}$$

denoting the a priori error covariance matrix. Then, $\boldsymbol{P}_k(\mathcal{X})$ is obtained directly from the KF recursion as

$$\boldsymbol{P}_k(\mathcal{X}) = \left[\boldsymbol{P}_{k|k-1}^{-1} + \sum_{u\in\mathcal{X}} \boldsymbol{H}_u^T\boldsymbol{R}_{v,u}^{-1}\boldsymbol{H}_u\right]^{-1} \tag{48}$$

where $\boldsymbol{P}_{k-1} = \mathbb{E}[\hat{\boldsymbol{x}}_{k-1} - \boldsymbol{x}_{k-1}][\hat{\boldsymbol{x}}_{k-1} - \boldsymbol{x}_{k-1}]^T$ is the error covariance matrix of the state estimate and $\boldsymbol{P}_{0|-1} = \boldsymbol{\Pi}_0$ [31]. ■

A. Proof of Proposition 2

Proof: Instead of relying on a Kalman smoother, we obtain the error covariance matrix by analyzing this problem in the batch setting. Note that both methods solve the same problem and therefore lead to the same solution. Proceeding as in [28], note from (1) that estimating $\bar{\boldsymbol{x}}_k$, i.e., all states $\boldsymbol{x}_j$ for $j\le k$, is equivalent to estimating the first state $\boldsymbol{x}_0$ and the process noises $\{\boldsymbol{w}_j\}_{j\le k-1}$. Indeed, we can recover the smoothed states using $\tilde{\boldsymbol{x}}_k = \boldsymbol{\Phi}_k\bar{\boldsymbol{z}}_k$, for $\bar{\boldsymbol{z}}_k = [\,\boldsymbol{x}_0^T\ \boldsymbol{w}_0^T\ \cdots\ \boldsymbol{w}_{k-1}^T\,]^T$ and $\boldsymbol{\Phi}_k$ as in (7).

The estimation of $\bar{\boldsymbol{z}}_k$ can be cast as a stochastic estimation problem by lifting. Formally, using the stacked $\bar{\boldsymbol{y}}_{u,k} = [\,\boldsymbol{y}_{u,0}^T\cdots\boldsymbol{y}_{u,k}^T\,]^T$ and $\bar{\boldsymbol{v}}_{u,k} = [\,\boldsymbol{v}_{u,0}^T\cdots\boldsymbol{v}_{u,k}^T\,]^T$, it is straightforward that

$$\begin{bmatrix}\bar{\boldsymbol{y}}_{u_1,k}\\ \vdots\\ \bar{\boldsymbol{y}}_{u_s,k}\end{bmatrix} = \underbrace{\begin{bmatrix}\boldsymbol{I}\otimes\boldsymbol{H}_{u_1}\\ \vdots\\ \boldsymbol{I}\otimes\boldsymbol{H}_{u_s}\end{bmatrix}}_{\boldsymbol{O}_k(\mathcal{X})}\boldsymbol{\Phi}_k\,\bar{\boldsymbol{z}}_k + \begin{bmatrix}\bar{\boldsymbol{v}}_{u_1,k}\\ \vdots\\ \bar{\boldsymbol{v}}_{u_s,k}\end{bmatrix} \tag{49}$$

where $\mathcal{X} = \{u_1,\dots,u_s\}$ is the observed output subset and $\otimes$ denotes the Kronecker product. The minimum MSE incurred from estimating $\bar{\boldsymbol{z}}$ from $\bar{\boldsymbol{y}}$ is then given by the trace of

$$\boldsymbol{K}(\mathcal{X}) = \left[\boldsymbol{C}^{-1} + \boldsymbol{O}_k(\mathcal{X})^T\,\mathrm{blkdiag}\left(\boldsymbol{I}\otimes\boldsymbol{R}_{v,u}\right)^{-1}\boldsymbol{O}_k(\mathcal{X})\right]^{-1} \tag{50}$$

for $u\in\mathcal{X}$, $\boldsymbol{O}_k$ in (49), $\boldsymbol{C} = \mathrm{blkdiag}(\boldsymbol{\Pi}_0, \boldsymbol{I}\otimes\boldsymbol{R}_w)$, and $\mathrm{blkdiag}(\boldsymbol{X},\boldsymbol{Y})$ denoting the block diagonal matrix whose diagonal blocks are $\boldsymbol{X}$ and $\boldsymbol{Y}$ [31].

To obtain the form in (4), use the fact that $\mathrm{blkdiag}(\boldsymbol{A}_i)^{-1} = \mathrm{blkdiag}(\boldsymbol{A}_i^{-1})$ and the mixed product property of the Kronecker product [46] to get

$$\begin{aligned}&\boldsymbol{O}_k(\mathcal{X})^T\,\mathrm{blkdiag}\left(\boldsymbol{I}\otimes\boldsymbol{R}_{v,u}\right)^{-1}\boldsymbol{O}_k(\mathcal{X})\\ &\quad= \sum_{u\in\mathcal{X}}\boldsymbol{\Phi}_k^T\left(\boldsymbol{I}\otimes\boldsymbol{H}_u^T\boldsymbol{R}_{v,u}^{-1}\boldsymbol{H}_u\right)\boldsymbol{\Phi}_k.\end{aligned} \tag{51}$$

Substituting (51) in (50) yields (6). ■

B. Proof of Theorem 1

Proof: Since f is monotone decreasing, it holds for every set $\mathcal{G}_j$ that

$$\begin{aligned}f(\mathcal{X}^\star) &\ge f(\mathcal{X}^\star\cup\mathcal{G}_j)\\ &= f(\mathcal{G}_j) - \sum_{i=1}^{s}\left[f(\mathcal{T}_{i-1}) - f(\mathcal{T}_{i-1}\cup v_i^\star)\right]\end{aligned} \tag{52}$$

where $\mathcal{T}_0 = \mathcal{G}_j$, $\mathcal{T}_i = \mathcal{G}_j\cup\{v_1^\star,\dots,v_i^\star\}$, $i=1,\dots,s$, and $v_i^\star$ is the ith element of $\mathcal{X}^\star$. Notice that this holds regardless of the order in which the $v_i^\star$ are taken. Since f is α-supermodular and $\mathcal{G}_j\subseteq\mathcal{T}_i$ for all i, the incremental gains in (52) can be bounded using (9) to get

$$f(\mathcal{X}^\star) \ge f(\mathcal{G}_j) - \alpha^{-1}\sum_{i=1}^{s}\left[f(\mathcal{G}_j) - f(\mathcal{G}_j\cup v_i^\star)\right]. \tag{53}$$

To proceed, we use the fact that $\mathcal{G}_{j+1} = \mathcal{G}_j\cup\{u\}$ is constructed by the greedy procedure in (10) so as to maximize the incremental gains in (53). It therefore holds that

$$f(\mathcal{X}^\star) \ge f(\mathcal{G}_j) - \alpha^{-1}s\left[f(\mathcal{G}_j) - f(\mathcal{G}_{j+1})\right]. \tag{54}$$

By adding and subtracting $f(\mathcal{X}^\star)$ in the brackets of (54), it yields a recursion for the distance to the optimal value $\delta_j = f(\mathcal{X}^\star) - f(\mathcal{G}_j)$ given by

$$\delta_j \ge \alpha^{-1}s\left[\delta_j - \delta_{j+1}\right] \Rightarrow \delta_{j+1} \ge \left(1 - \frac{1}{\alpha^{-1}s}\right)\delta_j.$$

Note that since f is nonpositive, $\delta_j\le 0$ for all j by the optimality of $\mathcal{X}^\star$.

Authorized licensed use limited to: University of Pennsylvania. Downloaded on October 29,2026 at 15:10:48 UTC from IEEE Xplore. Restrictions apply.

The expression in (14) yields directly from this recursion by noticing that since f is normalized, we have that $\delta_0 = f(\mathcal{X}^\star) - f(\emptyset) = f(\mathcal{X}^\star)$. Immediately, since Algorithm 1 is used for r iterations, we obtain

$$f(\mathcal{X}^\star) - f(\mathcal{G}_r) \geq \left(1 - \frac{\alpha}{s}\right)^r f(\mathcal{X}^\star). \tag{55}$$

Using the fact that $1 - x \leq e^{-x}$, the expression in (55) can be rearranged into (14). ■

C. Proof of Theorem 2

Proof: Using the assumption that f is monotone decreasing, we write

$$\begin{aligned} f(\mathcal{X}^\star) &\geq f(\mathcal{X}^\star \cup \mathcal{G}_j) \\ &= f(\mathcal{G}_j) - \sum_{i=1}^{s} \left[f(\mathcal{T}_{i-1}) - f(\mathcal{T}_{i-1} \cup \{v_i^\star\})\right] \end{aligned} \tag{56}$$

where again $\mathcal{T}_0 = \mathcal{G}_j$, $\mathcal{T}_i = \mathcal{G}_j \cup \{v_1^\star, \ldots, v_i^\star\}$, for $i = 1, \ldots, s$, and $v_i^\star$ is the ith element of $\mathcal{X}^\star$. Observe that the order in which the $v_i^\star$ occur in the telescopic sum is irrelevant. Since f is ϵ-supermodular and $\mathcal{G}_j \subseteq \mathcal{T}_i$ for all i, (15) can be used to bound the incremental gains in (56). Explicitly,

$$f(\mathcal{X}^\star) \geq f(\mathcal{G}_j) - \sum_{i=1}^{s} \left[f(\mathcal{G}_j) - f(\mathcal{G}_j \cup \{v_i^\star\}) + \epsilon\right].$$

Since $\mathcal{G}_{j+1} = \mathcal{G}_j \cup \{u\}$ is constructed in (10) so as to minimize $f(\mathcal{G}_{j+1})$, it holds that

$$f(\mathcal{X}^\star) \geq f(\mathcal{G}_j) - s\left[f(\mathcal{G}_j) - f(\mathcal{G}_{j+1}) + \epsilon\right]. \tag{57}$$

Finally, a recursion for the optimality gap $\delta'_j = f(\mathcal{G}_j) - f(\mathcal{X}^\star)$ is obtained from (57) by adding and subtracting $f(\mathcal{X}^\star)$ in the brackets. Explicitly,

$$\delta'_j \leq s\left(\delta'_j - \delta'_{j+1} + \epsilon\right) \Rightarrow \delta'_{j+1} \leq \left(1 - \frac{1}{s}\right)\delta'_j + \epsilon. \tag{58}$$

Notice that since f is nonpositive, $\delta'_j \geq 0$ for all j due to the optimality of $\mathcal{X}^\star$. The solution of (58) after r steps is

$$\delta'_r \leq \left(1 - \frac{1}{s}\right)^r \delta'_0 + \epsilon \sum_{j=0}^{r-1} \left(1 - \frac{1}{s}\right)^j. \tag{59}$$

Evaluating the geometric series in (59) yields

$$\delta'_r \leq \left(1 - \frac{1}{s}\right)^r \delta'_0 + s\left[1 - \left(1 - \frac{1}{s}\right)^r\right]\epsilon$$

and under the assumption that f is normalized, i.e., for $\delta'_0 = f(\emptyset) - f(\mathcal{X}^\star) = -f(\mathcal{X}^\star) \geq 0$, we obtain

$$f(\mathcal{G}_r) \leq \left[1 - \left(1 - \frac{1}{s}\right)^r\right]\left[f(\mathcal{X}^\star) + s \cdot \epsilon\right].$$

Using once again the fact that $1 - x \leq e^{-x}$ yields (17). ■

D. Proof of Lemma 1

Proof: We proceed with the proof case-by-case:

i) From the definition of α-supermodularity in (12), for $\mathcal{A} \subseteq \mathcal{B} \subseteq \mathcal{O}$ and $u \in \mathcal{O} \setminus \mathcal{B}$, it holds that

$$\begin{aligned} g(\mathcal{A}) - g(\mathcal{A} \cup \{u\}) &= \sum_i \theta_i \left[f_i(\mathcal{A}) - f_i(\mathcal{A} \cup \{u\})\right] \\ &\geq \sum_i \alpha_i \theta_i \left[f_i(\mathcal{B}) - f_i(\mathcal{B} \cup \{u\})\right] \end{aligned}$$

where the constant factors of the affine transformations cancel out. Since $\alpha_i \geq \min(\alpha_i)$, we obtain

$$\begin{aligned} &g(\mathcal{A}) - g(\mathcal{A} \cup \{u\}) \\ &\quad \geq \min(\alpha_i) \sum_i \theta_i \left[f_i(\mathcal{B}) - f_i(\mathcal{B} \cup \{u\})\right] \\ &\quad \geq \min(\alpha_i)\left[g(\mathcal{B}) - g(\mathcal{B} \cup \{u\})\right]. \end{aligned}$$

ii) Similarly, using the definition of ϵ-supermodularity in (15), for $\mathcal{A} \subseteq \mathcal{B} \subseteq \mathcal{V}$ and $u \in \mathcal{O} \setminus \mathcal{B}$, we obtain

$$\begin{aligned} &g(\mathcal{A}) - g(\mathcal{A} \cup \{u\}) \\ &\quad \geq \sum_i \theta_i \left[f_i(\mathcal{B}) - f_i(\mathcal{B} \cup \{u\})\right] - \sum_i \theta_i \epsilon_i \\ &\quad = g(\mathcal{B}) - g(\mathcal{B} \cup \{u\}) - \sum_i \theta_i \epsilon_i \end{aligned}$$

where again the constant factors b canceled out.

■

E. Proof of Lemma 3

Proof: Start by simplifying $\Delta_u t(\mathcal{X}) = t(\mathcal{X}) - t(\mathcal{X} \cup \{u\})$ using the matrix inversion lemma. To do so, write

$$\boldsymbol{Y}(\mathcal{X} \cup \{u\}) = \left[\boldsymbol{R}(\mathcal{X}) + \boldsymbol{M}_u\right]^{-1} \tag{60}$$

letting once again $\boldsymbol{R}(\mathcal{X}) = \boldsymbol{M}_\emptyset + \sum_{u \in \mathcal{X}} \boldsymbol{M}_u$. Since $\boldsymbol{R}(\mathcal{X}) \succ 0$, but the matrices $\boldsymbol{M}_u$ need not be invertible, we use an alternative form of the matrix inversion lemma [51] to get

$$\boldsymbol{Y}(\mathcal{X} \cup \{u\}) = \boldsymbol{Y}(\mathcal{X}) - \boldsymbol{Y}(\mathcal{X})\boldsymbol{M}_u\boldsymbol{Y}(\mathcal{X} \cup \{u\}) \tag{61}$$

where we used (60) to write $\boldsymbol{Y}(\mathcal{X}) = \boldsymbol{R}(\mathcal{X})^{-1}$. Using (61) and the linearity of the trace [46], we can write $\Delta_u t(\mathcal{X})$ as

$$\Delta_u t(\mathcal{X}) = \operatorname{Tr}\left[\boldsymbol{Y}(\mathcal{X})\boldsymbol{M}_u\boldsymbol{Y}(\mathcal{X} \cup \{u\})\right]. \tag{62}$$

To proceed, let $\tilde{\boldsymbol{M}}_u = \boldsymbol{M}_u + \epsilon\boldsymbol{I} \succ 0$ for $\epsilon > 0$ and define the perturbed version of (62) as

$$\widetilde{\Delta}_u t(\mathcal{X}) = \operatorname{Tr}\left[\boldsymbol{Y}(\mathcal{X})\tilde{\boldsymbol{M}}_u\left(\boldsymbol{R}(\mathcal{X}) + \tilde{\boldsymbol{M}}_u\right)^{-1}\right]. \tag{63}$$

Notice that $\widetilde{\Delta}_u t \to \Delta_u t$ as $\epsilon \to 0$. Using the invertibility of $\tilde{\boldsymbol{M}}_u$ and $\boldsymbol{R}(\mathcal{X})$, we obtain

$$\widetilde{\Delta}_u t(\mathcal{X}) = \operatorname{Tr}\left[\boldsymbol{Y}(\mathcal{X})\left(\boldsymbol{Y}(\mathcal{X}) + \tilde{\boldsymbol{M}}_u^{-1}\right)^{-1}\boldsymbol{Y}(\mathcal{X})\right].$$

Since $\boldsymbol{Y}(\mathcal{X}) \succ 0$, its square-root $\boldsymbol{Y}(\mathcal{X})^{1/2}$ is well-defined and unique [46]. We can therefore use the circular commutation property of the trace to get

$$\widetilde{\Delta}_u t(\mathcal{X}) = \operatorname{Tr}\left[\boldsymbol{Y}(\mathcal{X})\boldsymbol{Z}(\mathcal{X}, u)\right] \tag{64}$$

Authorized licensed use limited to: University of Pennsylvania. Downloaded on October 29,2026 at 15:10:48 UTC from IEEE Xplore. Restrictions apply.

with $\boldsymbol{Z}(\mathcal{X},u)=\boldsymbol{Y}(\mathcal{X})^{1/2}[\boldsymbol{Y}(\mathcal{X})+\tilde{\boldsymbol{M}}_u^{-1}]^{-1}\boldsymbol{Y}(\mathcal{X})^{1/2}$. Since both matrices in (64) are positive definite, we can use the bound from [52] to get

$$\lambda_{\min}\left[\boldsymbol{Y}(\mathcal{X})\right]\operatorname{Tr}\left(\boldsymbol{Z}\right)\leq\widetilde{\Delta}_u t(\mathcal{X})\leq\lambda_{\max}\left[\boldsymbol{Y}(\mathcal{X})\right]\operatorname{Tr}\left(\boldsymbol{Z}\right).$$

Reversing the manipulations used to obtain (64) yields

$$\lambda_{\min}\left[\boldsymbol{Y}(\mathcal{X})\right]\operatorname{Tr}\left[\tilde{\boldsymbol{M}}_u\left(\boldsymbol{Y}(\mathcal{X})+\tilde{\boldsymbol{M}}_u\right)^{-1}\right]\leq\widetilde{\Delta}_u t(\mathcal{X})$$

$$\leq\lambda_{\max}\left[\boldsymbol{Y}(\mathcal{X})^{-1}\right]\operatorname{Tr}\left[\tilde{\boldsymbol{M}}_u\left(\boldsymbol{Y}(\mathcal{X})+\tilde{\boldsymbol{M}}_u\right)^{-1}\right].$$

The result in (22) is obtained by continuity as $\epsilon\to 0$. ■

REFERENCES

[1] L. F .O. Chamon, G. J. Pappas, and A. Ribeiro, "The mean square error in Kalman filtering sensor selection is approximately supermodular," in *Proc. Conf. Decis. Control*, 2017, pp. 343–350.

[2] Y. Mo, R. Ambrosino, and B. Sinopoli, "Sensor selection strategies for state estimation in energy constrained wireless sensor networks," *Automatica*, vol. 47, no. 7, pp. 1330–1338, 2011.

[3] H. Rowaihy *et al.*, "A survey of sensor selection schemes in wireless sensor networks," in *Proc. Int. Soc. Opt. Photon.*, vol. 6562, 2007, pp. 1–13.

[4] F. Zhao, J. Shin, and J. Reich, "Information-driven dynamic sensor collaboration," *IEEE Signal Process. Mag.*, vol. 19, no. 2, pp. 61–72, Mar. 2002.

[5] H. Zhang, J. M. F. Moura, and B. Krogh, "Dynamic field estimation using wireless sensor networks: Tradeoffs between estimation error and communication cost," *IEEE Trans. Signal Process.*, vol. 57, no. 6, pp. 2383–2395, 2009.

[6] A. O. Hero and D. Cochran, "Sensor management: Past, present, and future," *IEEE Sensors J.*, vol. 11, no. 12, pp. 3064–3075, Dec. 2011.

[7] L. Chen, N. Li, L. Jiang, and S. H. Low, *Optimal Demand Response: Problem Formulation and Deterministic Case*. New York, NY, USA: Springer, 2012, pp. 63–85.

[8] Y.-Y. Liu, J.-J. Slotine, and A.-L. Barabási, "Observability of complex systems," *Proc. Nat. Acad. Sci.*, vol. 110, no. 7, pp. 2460–2465, 2012.

[9] G. Thatte and U. Mitra, "Spatio-temporal sensor management for environmental field estimation," *Signal Process.*, vol. 128, pp. 369–381, 2016.

[10] A. Krause, A. Singh, and C. Guestrin, "Near-optimal sensor placements in Gaussian processes: Theory, efficient algorithms and empirical studies," *J. Mach. Learn. Res.*, vol. 9, pp. 235–284, 2008.

[11] A. Das and D. Kempe, "Submodular meets spectral: Greedy algorithms for subset selection, sparse approximation and dictionary selection," in *Proc. Int. Conf. Mach. Learn.*, 2011, pp. 1057–1064.

[12] J. Ranieri, A. Chebira, and M. Vetterli, "Near-optimal sensor placement for linear inverse problems," *IEEE Trans. Signal Process.*, vol. 62, no. 5, pp. 1135–1146, Mar. 2014.

[13] V. Tzoumas, M. A. Rahimian, G. J. Pappas, and A. Jadbabaie, "Minimal actuator placement with bounds on control effort," *IEEE Trans. Control Netw. Syst.*, vol. 3, no. 1, pp. 67–78, Mar. 2016.

[14] A. Olshevsky, "On (non)supermodularity of average control energy," *IEEE Trans. Control Netw. Syst.*, vol. 5, no. 3, pp. 1177–1181, Sep. 2018.

[15] H. Zhang, R. Ayoub, and S. Sundaram, "Sensor selection for Kalman filtering of linear dynamical systems: Complexity, limitations and greedy algorithms," *Automatica*, vol. 78, pp. 202–210, 2017.

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

[17] T. H. Summers, F. L. Cortesi, and J. Lygeros, "On submodularity and controllability in complex dynamical networks," *IEEE Trans. Control Netw. Syst.*, vol. 3, no. 1, pp. 91–101, Mar. 2016.

[18] C. Commault and J.-M. Dion, "Input addition and leader selection for the controllability of graph-based systems," *Automatica*, vol. 49, no. 11, pp. 3322–3328, 2013.

[19] S. Pequito, S. Kar, and A. P. Aguiar, "On the complexity of the constrained input selection problem for structural linear systems," *Automatica*, vol. 62, pp. 193–199, 2015.

[20] S. Pequito, S. Kar, and A. P. Aguiar, "A framework for structural input/output and control configuration selection in large-scale systems," *IEEE Trans. Autom. Control*, vol. 61, no. 2, pp. 303–318, Feb. 2016.

[21] J. E. Weimer, B. Sinopoli, and B. H. Krogh, "A relaxation approach to dynamic sensor selection in large-scale wireless networks," in *Proc. Int. Conf. Distrib. Comput. Syst.*, 2008, pp. 501–506.

[22] S. Joshi and S. Boyd, "Sensor selection via convex optimization," *IEEE Trans. Signal Process.*, vol. 57, no. 2, pp. 451–462, Feb. 2009.

[23] N. Moshtagh, L. Chen, and R. Mehra, "Optimal measurement selection for any-time Kalman filtering with processing constraints," in *Proc. Conf. Decis. Control*, 2009, pp. 5074–5079.

[24] Y. Wang, M. Sznaier, and F. Dabbene, "A convex optimization approach to worst-case optimal sensor selection," in *Proc. Conf. Decis. Control*, 2013, pp. 6353–6358.

[25] N. K. Dhingra, M. R. Jovanović, and Z. Q. Luo, "An ADMM algorithm for optimal sensor and actuator selection," in *Proc. Conf. Decis. Control*, 2014, pp. 4039–4044.

[26] S. Liu, S. P. Chepuri, M. Fardad, E. Maşazade, G. Leus, and P. K. Varshney, "Sensor selection for estimation with correlated measurement noise," *IEEE Trans. Signal Process.*, vol. 64, no. 13, pp. 3509–3522, Jul. 2016.

[27] M. Shamaiah, S. Banerjee, and H. Vikalo, "Greedy sensor selection: Leveraging submodularity," in *Proc. Conf. Decis. Control*, 2010, pp. 2572–2577.

[28] V. Tzoumas, A. Jadbabaie, and G. J. Pappas, "Sensor placement for optimal Kalman filtering: Fundamental limits, submodularity, and algorithms," in *Proc. Amer. Control Conf.*, 2016, pp. 191–196.

[29] G. E. P. Box, J. S. Hunter, and W. G. Hunter, *Statistics for Experimenters: Design, Innovation, and Discovery*. New York, NY, USA: Springer, 2005.

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

[31] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*. Englewood Cliffs, NJ, USA: Prentice-Hall, 2000.

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

[33] A. Krause and D. Golovin, "Submodular function maximization, " in, *Tractability: Practical Approaches to Hard Problems*. Cambridge: Cambridge University Press, 2014.

[34] F. Bach, "Learning with submodular functions: A convex optimization perspective," *Found. Trends Mach. Learn.*, vol. 6, no. 2–3, pp. 145–373, 2013.

[35] A. Das and D. Kempe, "Algorithms for subset selection in linear regression," in *Proc. ACM Symp. Theory Comput.*, 2008, pp. 45–54.

[36] P. Singh, M. Chen, L. Carlone, S. Karaman, E. Frazzoli, and D. Hsu, "Supermodular mean squared error minimization for sensor scheduling in optimal Kalman filtering," in *Proc. Amer. Control Conf.*, 2017, pp. 5787–5794.

[37] H. Avron and C. Boutsidis, "Faster subset selection for matrices and applications," *SIAM J. Matrix Anal. Appl.*, vol. 34, no. 4, pp. 1464–1499, 2013.

[38] L. F. O. Chamon and A. Ribeiro, "Approximate supermodularity bounds for experimental design," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, pp. 5403–5412.

[39] T. Horel and Y. Singer, "Maximization of approximately submodular functions," in *Proc. Adv. Neural Inf. Process. Syst.*, 2016, pp. 3045–3053.

[40] B. Lehmann, D. Lehmann, and N. Nisan, "Combinatorial auctions with decreasing marginal utilities," *Games Econ. Behav.*, vol. 55, no. 2, pp. 270–296, 2006.

[41] M. Sviridenko, J. Vondrák, and J. Ward, "Optimal approximation for submodular and supermodular optimization with bounded curvature," in *Proc. SIAM Symp. Discrete Algorithms*, 2014, pp. 1134–1148.

[42] L. F. O. Chamon and A. Ribeiro, "Near-optimality of greedy set selection in the sampling of graph signals," in *Proc. Global Conf. Signal Inf. Process.*, 2016, pp. 1265–1269.

[43] A. Bian, J. M. Buhmann, A. Krause, and S. Tschiatschek, "Guarantees for greedy maximization of non-submodular functions with applications," in *Proc. Int. Conf. Mach. Learn.*, 2017, pp. 498–507.

[44] T. Summers and M. Kamgarpour, "Performance guarantees for greedy maximization of non-submodular set functions in systems and control," in *Proc. Europ. Control Conf.*, 2019, pp. 2796–2801.

[45] A. Krause and V. Cevher, "Submodular dictionary selection for sparse representation," in *Proc. Int. Conf. Mach. Learn.*, 2010, pp. 567–574.

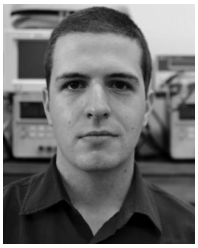
[46] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge: Cambridge University Press, 2013.

[47] R. Bhatia, *Matrix Analysis*. New York, NY, USA: Springer, 1997.

[48] L. Lováz, "Submodular functions and convexity," in, *Mathematical Programming: The State of the Art*. Berlin, Germany: Springer-Verlag, 1982.

Authorized licensed use limited to: University of Pennsylvania. Downloaded on October 29,2026 at 15:10:48 UTC from IEEE Xplore. Restrictions apply.

[49] M. S. Grewal and A. P. Andrews, *Kalman Filtering: Theory and Practice Using MATLAB*. Hoboken, NJ, USA: Wiley, 2014.
[50] F. Muller, F. Seyler, and J.-L. Guyot, "Utilisation d'imagerie radar (ROS) JERS-1 pour l'obtention de réseaux de drainage. Exemple du Rio Negro (Amazonie)," in *Hydrological and Geochemical Processes in Large Scale River Basins*, 1999. Available: http://www.ore-hybam.org/index.php/eng/Data/Cartography/Amazon-basin-hy drography.
[51] H. V. Henderson and S. R. Searle, "On deriving the inverse of a sum of matrices," *SIAM Rev.*, vol. 23, no. 1, pp. 53–60, 1981.
[52] B. Wang and F. Zhang, "Some inequalities for the eigenvalues of the product of positive semidefinite Hermitian matrices," *Linear Algebra Appl.*, vol. 160, pp. 113–118, 1992.



Luiz F. O. Chamon (Student Member, IEEE) received the B.Sc. and M.Sc. degrees in electrical engineering from the University of São Paulo, São Paulo, Brazil, in 2011 and 2015, respectively. In 2009, he was an Undergraduate Exchange Student at the Masters in Acoustics of the École Centrale de Lyon, Lyon, France. He is currently working toward the Ph.D. degree in electrical and systems engineering at the University of Pennsylvania, Philadelphia, USA.

In 2009, he was an Assistant Instructor and Consultant on nondestructive testing with the INSACAST Formation Continue. From 2010 to 2014, he worked as a Signal Processing and Statistical Consultant on a project with the EMBRAER. His research interests include signal processing, optimization, statistics, and control.



George J. Pappas (Fellow, IEEE) received the Ph.D. degree in electrical engineering and computer sciences from the University of California, Berkeley, CA, USA, in 1998.

He is currently the Joseph Moore Professor and Chair of the Department of Electrical and Systems Engineering, University of Pennsylvania, Philadelphia, PA, USA. He also holds a secondary appointment with the Department of Computer and Information Sciences and the Department of Mechanical Engineering and Applied Mechanics. He is a member of the GRASP Lab and the PRECISE Center. He had previously served as the Deputy Dean for Research with the School of Engineering and Applied Science. His research interests include control theory and, in particular, hybrid systems, embedded systems, cyberphysical systems, and hierarchical and distributed control systems, with applications to unmanned aerial vehicles, distributed robotics, green buildings, and biomolecular networks.

Dr Pappas has received various awards, such as the Antonio Ruberti Young Researcher Prize, the George S. Axelby Award, the Hugo Schuck Best Paper Award, the George H. Heilmeier Award, the National Science Foundation PECASE Award, and numerous best student papers awards at ACC, CDC, and ICCPS.



Alejandro Ribeiro received the B.Sc. degree in electrical engineering from the Universidad de la Republica Oriental del Uruguay, Montevideo, in 1998, and the M.Sc. and Ph.D. degrees in electrical engineering from the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, in 2005 and 2007.

From 1998 to 2003, he was a member of the technical staff with the Bellsouth Montevideo. He joined the University of Pennsylvania, Philadelphia, where he is currently the Rosenbluth Associate Professor with the Department of Electrical and Systems Engineering. His research interests include the applications of statistical signal processing to the study of networks and networked phenomena, and on structured representations of networked data structures, graph signal processing, network optimization, robot teams, and networked control.

Dr Ribeiro received the 2014 O. Hugo Schuck Best Paper Award, and paper awards at CDC 2017, 2016 SSP Workshop, 2016 SAM Workshop, 2015 Asilomar SSC Conference, ACC 2013, ICASSP 2006, and ICASSP 2005. His teaching has been recognized with the 2017 Lindback Award for Distinguished Teaching and the 2012 S. Reid Warren, Jr. Award presented by Penn's undergraduate student body for outstanding teaching. He is a Fulbright scholar class of 2003 and a Penn Fellow class of 2015.

Authorized licensed use limited to: University of Pennsylvania. Downloaded on October 29,2026 at 15:10:48 UTC from IEEE Xplore. Restrictions apply.