Probabilistic Performance Bounds for Randomized Sensor Selection in Kalman Filtering

Christopher I. Calle and Shaunak D. Bopardikar

***Abstract*— We consider the problem of randomly choosing the sensors of a linear time-invariant dynamical system subject to process and measurement noise. Each sensor is sampled independently and from the same distribution for the purpose of state estimation by Kalman filtering. Due to our randomized sampling procedure, the estimation error covariance cannot be bounded in a deterministic sense. Using tools from random matrix theory, we derive probabilistic bounds on the steady-state estimation error covariance in the semi-definite sense for an arbitrary sampling distribution. Our bounds are functions of several tunable parameters of interest, such as the number of sampled sensors and the likelihood that our bounds hold. We indirectly improve the performance of our Kalman filter for the maximum eigenvalue metric by finding the optimal sampling distribution. By minimizing the maximum eigenvalue of the upper bound, we are able to minimize the maximum eigenvalue of the steady-state estimation error covariance, the actual metric of interest. We identify the subset of sensors to sample with high frequency through the optimal sampling distribution. We illustrate our findings through several insightful simulations and comparisons with multiple sampling policies.**

***Index Terms*— Sensor selection, Kalman filtering, Random matrix theory**

I. Introduction

In large-scale sensor networks, the problem of allocating the fewest resources possible while simultaneously achieving some minimal performance is computationally expensive and, in some settings, infeasible. Examples of such scenarios include mobile platforms to study the development of severe weather [1], underwater sensing technologies to detect and monitor the dispersal of chemical plumes [2], and smart sensors to monitor traffic [3], to name a few. A common theme is that the quantity of interest evolves in time. In the presence of multiple sensors to choose from, the question naturally arises: *Can one select the sensors in an efficient manner and, simultaneously, provide provable guarantees on the estimation performance?* If the Kalman filter is the estimator for a linear time-invariant (LTI) dynamical system subject to noise, then this paper provides an affirmative answer for randomized sampling *with replacement* policies.

Sensor selection has a rich history in the control literature – refer to [4] or [5] for a survey of early works. Metrics provide a measure on the quality of a sensor selection. Several metrics and efficient algorithms are discussed in [6] and [7]. In the context of large sensor networks, one approach is to sample sensors in a randomized manner. Such a sampling scheme offers computational efficiency at the expense of yielding only probabilistic guarantees on the quality of the sampled sensor selection. Notable early works that studied random sampling of sensors and its effect on the estimation error covariance of the Kalman filter include [8] and [9]. In [10], [11], and [12], the submodularity property of Kalman filter metrics and randomized sampling strategies are addressed.

Recent works employ concentration inequalities to address the sensor selection problem for randomized sampling policies. In [13] and [14], the approximate nature of non-submodular (non-supermodular) metrics in the Kalman filter setting are studied in order to justify the efficacy of a greedy heuristic. The notion of curvature is exploited in [14] in order to quantify how close a metric is to being submodular and bounds on the curvature are established by the Bernstein inequality [15] for a limited setting. In [16] and [17], bounds on observability Gramian metrics are established with high probability using the Ahslwede-Winter inequality [18]. Using the non-trivial result in [19], probabilistic guarantees on nonlinear observability metrics are established in [16] for uniform sampling *without replacement*. In [17], sampling *with replacement* policies are addressed for observability Gramian metrics of LTI dynamical systems. In this work, a Gramian lower bound is derived in terms of the Gramian of the original system for sensor placement at each location. However, such a result comes at a cost because the lower bound is no longer a bound of the original system. Instead, the lower bound holds for a scaled version of the original output system. In other words, guarantees on the observability Gramian are not directly established for the original system and the utility of the Ahlswede-Winter inequality [18] is not fully realized. By focusing on the observability Gramian, [16] and [17] assume that process and measurement noise are either non-existent or negligible. In this paper, we address the more practical estimation problem, where the sensor network and the process it is actively observing are corrupted by Gaussian noise. We extend the application of the Ahlswede-Winter inequality [18] for arbitrary sampling *with replacement* policies. In contrast to [17], we establish estimation performance bounds for the original system via the Kalman filter.

A. Contributions

The contributions of this work are three-fold. First, using the Ahlswede-Winter inequality [18], we derive upper and lower bounds on the estimation error covariance in the

The authors are with the Department of Electrical and Computer Engineering, Michigan State University (MSU), East Lansing, MI 48824, USA. Emails: callechr@msu.edu, shaunak@egr.msu.edu

This work was supported in part by NSF Grant # ECCS-2030556, GAANN Grant # P200A180025, the National GEM Fellowship, and MSU's University Enrichment Fellowship (UEF).

probabilistic sense. To our knowledge, we introduce the first concentration inequality to bound the estimation error covariance in the semi-definite sense for an arbitrary sampling *with replacement* policy. Probabilistic guarantees in the semi-definite sense are appealing since they imply assurances on several metrics of significance in state estimation, such as the maximum eigenvalue, condition number, or trace of the estimation error covariance, to name a few.

Second, we propose a search procedure for finding the optimal sampling distribution that minimizes the maximum eigenvalue of the upper bound for the steady-state estimation error covariance. We confirm that the optimal sampling distribution indirectly minimizes our actual metric of interest, the maximum eigenvalue of the steady-state estimation error covariance.

Third, we investigate the estimation performance and its variability for the optimal sampling distribution. We numerically compare the performance of our optimal sampling distribution against the uniform sampling distribution and a greedy heuristic.

B. Outline of the paper

This paper is organized as follows. Section II outlines the randomized sampling policy and the linear dynamical system under consideration. In Section III, we address the problem of sensor selection in the Kalman filter setting. We derive novel bounds on the steady-state estimation error covariance in the probabilistic sense and propose an optimal sampling distribution that indirectly improves estimation performance for the maximum eigenvalue metric. In Section IV, we present our numerical studies and comparisons with other sampling strategies. Finally, we summarize our findings and identify directions for future research in Section V. Due to constraints, the proofs of all mathematical claims are detailed in the online technical report [20].

II. Problem Formulation

A. Notation

We summarize the notation employed in this paper. Let $\underline{\lambda}(\cdot)$ and $\overline{\lambda}(\cdot)$ denote the minimum and maximum eigenvalue of a Hermitian matrix argument, respectively. We denote $I_n \in \mathbb{R}^{n\times n}$ as the identity matrix and Δ^n as the probability simplex in $\mathbb{R}^n$.

B. Sampling Scheme

We assume each sensor outputs only one measurement $y_{j,t} \in \mathbb{R}$. For clarity, we refer to the individual sensors available for sampling as candidate sensors. If a candidate sensor is modeled by an LTI measurement model corrupted by zero-mean Gaussian noise, i.e.,

$$y_{j,t} = \boldsymbol{c}_j^T x_t + \mu_{j,t},$$

then the pair $(\boldsymbol{c}_j, \boldsymbol{\sigma}_j^2)$ is sufficient in completely describing the measurement properties of the j-th candidate sensor, where $\boldsymbol{c}_j^T$ denotes the sequence that linearly relates the state x_t to the output $y_{j,t}$ and $\boldsymbol{\sigma}_j^2$ is the measurement variance of Gaussian noise $\mu_{j,t}$. Let $\mathcal{X} := \{(\boldsymbol{c}_1, \boldsymbol{\sigma}_1^2), \ldots, (\boldsymbol{c}_{n_c}, \boldsymbol{\sigma}_{n_c}^2)\}$ denote the set of candidate pairs and n_c specify the number of candidate sensors under consideration.

In our sampling scheme, n_s sensors are chosen *with replacement* out of a sampling pool of n_c candidate sensors. In other words, n_s pairs are chosen *with replacement* from distribution $\mathcal{X}$. Each candidate sensor is sampled with some given probability and the list of sampling probabilities is specified by distribution $p \in \Delta^{n_c}$. Let $\mathcal{S} \in \{1, \ldots, n_c\}^{n_s}$ denote the indices of the n_s sampled candidate sensors. Throughout this paper, quantities that are either directly or indirectly dependent on our randomly chosen sensor selection are accompanied by a subscript $\mathcal{S}$ notation. We assume the measurement properties $(\boldsymbol{c}, \boldsymbol{\sigma}^2)$ of each candidate sensor are known prior to sampling.

C. Model

Consider the tuple $(A, C_{\mathcal{S}}, Q, R_{\mathcal{S}}, I_m)$, an LTI state and measurement model subject to Gaussian noise, i.e.,

$$\begin{aligned} x_{t+1} &= A\, x_t + w_t, \\ y_t &= C_{\mathcal{S}}\, x_t + I_m\, v_t, \end{aligned} \tag{1}$$

where $x_t \in \mathbb{R}^m$ is the state vector, $y_t \in \mathbb{R}^{n_s}$ is the output vector, and n_s specifies the total number of measurements at time instant t. Let $A \in \mathbb{R}^{m\times m}$ and $C_{\mathcal{S}} \in \mathbb{R}^{n_s\times m}$ denote the state and output matrix, respectively. Each row of $C_{\mathcal{S}}$ consists of row vector c_j^T, where $c_j \in \mathbb{R}^m$ relates the state x_t to the output $y_{j,t}$ for the j-th sampled sensor. Let $w_t \sim \mathcal{N}(0, Q)$ and $v_t \sim \mathcal{N}(0, R_{\mathcal{S}})$ denote the process and measurement noise, respectively. Assume $\{w_t\}_{t=0}^{\infty}$ and $\{v_t\}_{t=0}^{\infty}$ are uncorrelated, zero-mean, white Gaussian processes. Additional assumptions on the noise properties of w_t and v_t are necessary for subsequent derivations.

Assumption 1. *Noise covariance matrices Q and $R_{\mathcal{S}}$ are time-invariant and positive definite.*

Assumption 2. *Measurement covariance $R_{\mathcal{S}}$ is diagonal, i.e., v_t consists of n_s uncorrelated random variables, where $v_{j,t}$ and σ_j^2 denote the measurement noise and variance, respectively, corresponding to the j-th sampled sensor.*

Note that the pair (c_j, σ_j^2) is sampled *with replacement* from distribution $\mathcal{X}$ for all $j \in \{1, \ldots, n_s\}$ as outlined in Section II-B.

D. Sensor Selection for Kalman Filtering

Under the assumptions of model linearity and Gaussian noise, the Kalman filter is a minimum mean squared error (MMSE) estimator that computes an optimal estimate of state x_t in the mean-squared sense. If our measurement vector y_t is available at each time instant t for sensor fusion in a centralized manner, then the covariance *information form* of the Kalman filter can be formulated into the following recursive equation,

$$P_{\mathcal{S},t}^{-1} = (AP_{\mathcal{S},t-1}A^T + Q)^{-1} + C_{\mathcal{S}}^T R_{\mathcal{S}}^{-1} C_{\mathcal{S}}, \tag{2}$$

where $P_{\mathcal{S},t}$ denotes the filtered covariance of the state estimation error at time instant t. If $(A, C_{\mathcal{S}})$ and $(A, Q^{1/2})$ are detectable and stabilizable, respectively, then filtered error covariance $P_{\mathcal{S},t}$ converges to a steady-state solution $P_{\mathcal{S}}$ [21]. Note the dependence of $C_{\mathcal{S}}^T R_{\mathcal{S}}^{-1} C_{\mathcal{S}}$ on the row vector c^T and measurement variance σ^2 of each randomly sampled sensor. In order to identify the contribution of each randomly sampled sensor in $C_{\mathcal{S}}^T R_{\mathcal{S}}^{-1} C_{\mathcal{S}}$, Assumption 2 is established. If a symmetric, positive semi-definite, *random* matrix Z_j is generated by the pair (c_j, σ_j^2) of the j-th randomly sampled sensor, i.e.,

$$Z_j = (\sigma_j^{-1} c_j)(\sigma_j^{-1} c_j)^T,$$

then, under Assumption 2, $C_{\mathcal{S}}^T R_{\mathcal{S}}^{-1} C_{\mathcal{S}}$ can be decomposed into a finite sum of independent and identically distributed (i.i.d.) random matrices,

$$C_{\mathcal{S}}^T R_{\mathcal{S}}^{-1} C_{\mathcal{S}} = \sum_{j=1}^{n_s} c_j\, \sigma_j^{-2} c_j^T = \sum_{j=1}^{n_s} Z_j.$$

Let $Z_1, \ldots, Z_{n_s}$ denote independent copies of random variable Z, i.e., independently sampled matrices with the same distribution as Z. The expectation of random matrices Z and $C_{\mathcal{S}}^T R_{\mathcal{S}}^{-1} C_{\mathcal{S}}$ are given by

$$\mathbb{E}[Z] = \sum_{j=1}^{n_c} p_j \mathcal{Z}_j,\ \mathbb{E}[C_{\mathcal{S}}^T R_{\mathcal{S}}^{-1} C_{\mathcal{S}}] = n_s \mathbb{E}[Z].$$

We construct $\mathcal{Z}_j$ by the pair $(\boldsymbol{c}_j, \boldsymbol{\sigma}_j^2)$, i.e.,

$$\mathcal{Z}_j = (\boldsymbol{\sigma}_j^{-1} \boldsymbol{c}_j)(\boldsymbol{\sigma}_j^{-1} \boldsymbol{c}_j)^T$$

for all $j \in \{1, \ldots, n_c\}$. Symmetric, positive semi-definite matrices $\mathcal{Z}_1, \ldots, \mathcal{Z}_{n_c}$ are possible realizations of random variable Z.

E. Problem Statement

Our focus is on the following complementary problems.

Problem 1. *Given an arbitrary sampling distribution p, determine the upper and lower bounds on the steady-state error covariance $P_{\mathcal{S}}$ in the semi-definite sense.*

Problem 2. *Find an optimal sampling distribution p^* that minimizes the maximum eigenvalue of the upper bound P_U on the steady-state error covariance $P_{\mathcal{S}}$.*

Problem 1 asks whether some minimal performance can be guaranteed with high probability, regardless of the sampling distribution under consideration. If such assurances exist, then the next question is whether there exists some ideal sampling scheme that optimizes our state estimation performance. Problem 2 addresses the latter and asks how one can strategically choose a sensor selection to minimize a performance measure, specifically, the maximum eigenvalue of $P_{\mathcal{S}}$. Since $\overline{\lambda}(P_{\mathcal{S}})$ is a random variable, it cannot be directly minimized. Instead, the maximum eigenvalue of upper bound P_U is minimized in order to indirectly influence our actual metric of interest. Solutions to Problem 1 and Problem 2 are found in Section III-A and Section III-B, respectively.

III. Main Results

First, we derive probabilistic bounds on the steady-state error covariance $P_{\mathcal{S}}$ in the semi-definite sense. Next, the expected steady-state solution $\mathbb{E}[P_{\mathcal{S}}]$ and its relation to the bounds P_U and P_L are explored. Lastly, the sampling distribution p^* that optimally minimizes $\overline{\lambda}(P_U)$ is obtained in order to indirectly minimize $\overline{\lambda}(P_{\mathcal{S}})$ and improve our state estimation performance.

A. Steady-State Solution Guarantees

Before establishing bounds on the steady-state solution $P_{\mathcal{S}}$, the filtered error covariance in the deterministic setting is investigated. If we assume n_s sensors are chosen beforehand and not randomly sampled, then the output matrix and measurement covariance of an LTI system are deterministic. Lemma 1 outlines the conditions required to deterministically upper and lower bound the filtered error covariance of an arbitrary LTI system in the semi-definite sense.

Lemma 1. ***(Deterministic Bounds)*** *Consider the following LTI systems, $(A, Y_3^{1/2}, Q, \Pi_3, \Pi_3^{-1/2})$, $(A, \Gamma_2, Q, \Pi_2, I_m)$, and $(A, Y_1^{1/2}, Q, \Pi_1, \Pi_1^{-1/2})$, and define their filtered error covariance matrices, i.e.,*

$$\begin{aligned} P_{1,t}^{-1} &= (AP_{1,t-1}A^T + Q)^{-1} + Y_3, \\ P_{2,t}^{-1} &= (AP_{2,t-1}A^T + Q)^{-1} + Y_2, \\ P_{3,t}^{-1} &= (AP_{3,t-1}A^T + Q)^{-1} + Y_1, \end{aligned}$$

respectively, such that $Y_i = \Gamma_i^T \Pi_i^{-1} \Gamma_i$ for all $i \in \{1, 2, 3\}$. If the following conditions are satisfied,

(C1) $0 \preceq Y_1 \preceq Y_2 \preceq Y_3$, and
(C2) $0 \preceq P_{1,-1} \preceq P_{2,-1} \preceq P_{3,-1}$, and
(C3) $(A, Y_3^{1/2})$, (A, Γ_2), and $(A, Y_1^{1/2})$ are detectable,
then $P_1 \preceq P_2 \preceq P_3$ and

$$P_{1,t-1} \preceq P_{2,t-1} \preceq P_{3,t-1},\ \forall t \geq 0.$$

Lemma 1 in tandem with the Ahlswede-Winter inequality [18], Theorem 3 in the appendix, yields probabilistic bounds in the semi-definite sense on the steady-state solution $P_{\mathcal{S}}$ for an arbitrary sampling distribution p.

Theorem 1. ***(Probabilistic Steady-State Bounds)*** *Suppose $m, n_s \in \mathbb{N}$, $\delta \in (0, 1)$, $\rho \in \left[1, \frac{n_s \epsilon^2}{4 \log(2m/\delta)}\right)$ and*

$$\epsilon = \sqrt{\frac{4\rho}{n_s} \log \frac{2m}{\delta}} \in (0, 1) \tag{3}$$

for specified sampling distribution p, such that $\mathcal{Z}_j \preceq \rho\, \mathbb{E}[Z]$ for all $j \in \{1, \ldots, n_c\}$. Assume $(A, \mathbb{E}[Z]^{1/2})$ and $(A, C_{\mathcal{S}})$ are detectable and $(A, Q^{1/2})$ is stabilizable. If P_U and P_L denote steady-state solutions, i.e.,

$$\begin{aligned} P_U^{-1} &= (AP_U A^T + Q)^{-1} + (1 - \epsilon) n_s \mathbb{E}[Z], \\ P_L^{-1} &= (AP_L A^T + Q)^{-1} + (1 + \epsilon) n_s \mathbb{E}[Z], \end{aligned}$$

and if the steady-state error covariance $P_{\mathcal{S}}$ satisfies

$$P_{\mathcal{S}}^{-1} = (AP_{\mathcal{S}} A^T + Q)^{-1} + C_{\mathcal{S}}^T R_{\mathcal{S}}^{-1} C_{\mathcal{S}},$$

then

$$\mathbb{P}[P_L \preceq P_{\mathcal{S}} \preceq P_U] \geq (1-\delta).$$

A few comments are summarized below.

Remark 1. *If m is fixed, then the only parameters that can be tuned to guarantee $\epsilon \in (0,1)$ are δ, n_s, and ρ. Quantities δ and n_s can be easily tuned since they are user-specified. In contrast, tuning the value of ρ is a non-trivial problem due to its dependence on the sampling distribution p.*

Remark 2. *The detectability assumptions of Theorem 1 can be satisfied by one of the following sufficient conditions:*

(S1) $(A, \mathbf{c}_j)$ is observable for all $j \in \{1, \ldots, n_c\}$, or

(S2) Prior to randomly selecting n_s sensors as outlined in Section II, if n_a additional sensors are first strategically sampled from the sampling pool of candidate sensors and shown to collectively guarantee observability of the system, then the steady-state solutions P_U, P_L, and $P_{\mathcal{S}}$ will always exist, regardless of the sampling distribution p, assuming the stabilizability condition is satisfied.

Remark 3. *From the dependence in* (3)*, we conclude that for the analysis to be applicable, we require*

$$n_s \geq \frac{4\rho}{\epsilon^2}\log\frac{2m}{\delta}.$$

Thus, the number of samples show a logarithmic dependence on m and $1/\delta$, which is reasonable. However, the $1/\epsilon^2$ dependence is a consequence of sampling with replacement and the central limit theorem which is the key result used in the proof of Theorem 3, the Ahlswede-Winter inequality [18], that is employed to establish Theorem 1.

In order to measure the average state estimation performance of our sampling scheme in Section II, we introduce an analytical lower bound on the expectation of the steady-state solution $P_{\mathcal{S}}$ that closely approximates it.

Lemma 2. ***(Analytical Lower Bound)*** *If $(A, \mathbb{E}[Z]^{1/2})$ and $(A, Q^{1/2})$ are detectable and stabilizable, respectively, and if L denotes the solution to the following,*

$$L^{-1} = (ALA^T + Q)^{-1} + n_s\mathbb{E}[Z],$$

then $L \preceq \mathbb{E}[P_{\mathcal{S}}]$.

Lemma 2 is used to explain how the bounds P_U and P_L of Theorem 1 are related to the expected steady-state error covariance $\mathbb{E}[P_{\mathcal{S}}]$.

Remark 4. *In the limit as ϵ tends to* 0*, P_U and P_L tend to identical solutions, i.e.,*

$$\lim_{\epsilon\to 0} P_U^{-1} = (AP_UA^T + Q)^{-1} + n_s\mathbb{E}[Z],$$
$$\lim_{\epsilon\to 0} P_L^{-1} = (AP_LA^T + Q)^{-1} + n_s\mathbb{E}[Z].$$

For this special case, P_U and P_L are denoted by P_I, and

$$P_I^{-1} = (AP_IA^T + Q)^{-1} + n_s\mathbb{E}[Z]. \tag{4}$$

A simple comparison of Lemma 2 and (4) *shows that L and P_I are identical. This implies that P_U and P_L bound a lower bound of $\mathbb{E}[P_{\mathcal{S}}]$, denoted as L, for all $\epsilon \in (0,1)$. As ϵ decreases (increases), P_U and P_L converge towards (diverge from) L in the semi-definite sense.*

B. Optimal Sampling Scheme

In Section III-A, the steady-state error covariance is bounded in the probabilistic sense for an arbitrary sampling distribution p. Due to its dependence on a randomly chosen sensor selection $\mathcal{S}$, the steady-state solution $P_{\mathcal{S}}$ cannot directly be influenced. Instead, the bounds P_U and P_L of Theorem 1 must be used to indirectly affect state estimation performance. In this section, the maximum eigenvalue of the steady-state solution $P_{\mathcal{S}}$ is the performance metric of interest. By minimizing the maximum eigenvalue of the upper bound P_U, $\overline{\lambda}(P_{\mathcal{S}})$ is similarly minimized with high probability.

Theorem 2. ***(Optimal Sampling Distribution)*** *A sampling distribution $p_\rho^* = \{p_i^*\}_{i=1}^{n_c}$ that optimally minimizes $\overline{\lambda}(P_U)$ with respect to a selected $\rho \in \left[1, \frac{n_s\epsilon^2}{4\log(2m/\delta)}\right)$ and an arbitrarily small $\eta > 0$ is computed by solving the following semi-definite program (SDP).*

$$\begin{aligned}
\max_{\lambda, X, \{p_i\}_{i=1}^{n_c}} \quad & \lambda \\
\text{s.t.} \quad & \lambda > 0,\ X \succeq \eta I_m,\ \{p_i\}_{i=1}^{n_c} \in \Delta^{n_c} \\
& \epsilon = \sqrt{\frac{4\rho}{n_s}\log\frac{2m}{\delta}} \in (0,1) \\
& \begin{bmatrix} X + A^TQ^{-1}A & (A^TQ^{-1}) \\ (A^TQ^{-1})^T & \Gamma_3 \end{bmatrix} \succeq 0 \\
& \Gamma_3 = Q^{-1} + (1-\epsilon)n_s\sum_{j=1}^{n_c} p_j\mathcal{Z}_j - \lambda I_m \\
& \mathcal{Z}_i \preceq \rho\sum_{j=1}^{n_c} p_j\mathcal{Z}_j,\ \forall i \in \{1, \ldots, n_c\}
\end{aligned}$$

Theorem 2 computes the sampling distribution p_ρ^* for a selected ρ and η. In order to find the sampling distribution p^* that optimally minimizes $\overline{\lambda}(P_U)$, irrespective of ρ, a search algorithm is necessary.

Remark 5. *One should expect that minimizing ρ will minimize $\overline{\lambda}(P_U)$ upon inspection of constraint* (3)*, since minimizing ρ minimizes ϵ and, subsequently, tightens the bounds outlined in Theorem 1. Though this heuristic can be used to identify a relatively minimal $\overline{\lambda}(P_U)$, it cannot be guaranteed to find the global minimum. Instead, the optimal ρ^* that globally minimizes $\overline{\lambda}(P_U)$ can be found incrementally. By employing a binary or bisection search procedure throughout the feasible regime of ρ, Theorem 2 can be consecutively applied to find the ρ that globally minimizes $\overline{\lambda}(P_U)$ within a predefined constant γ of the optimal ρ^*.*

IV. Simulation Results

In this section, the optimal sampling distribution found in Section III-B is demonstrated to substantially minimize the maximum eigenvalue of upper bound P_U relative to a trivial uniform sampling distribution. We also show that relative to a greedy heuristic the estimation performance of our optimal sampling distribution is consistently better on-average for the maximum eigenvalue metric.

In our numerical analysis, the state dimension $m = 3$, the number of candidate sensors $n_c = 200$, and $\delta = 0.10$. We assume the process covariance matrix $Q = 0.5\, I_m$ and the measurement noise variance of each candidate sensor is identical, such that $\boldsymbol{\sigma}_j^2 = 0.5$ for all $j \in \{1, \dots, n_c\}$. The entries of state matrix A and output vector $\boldsymbol{c}$ for each candidate sensor are chosen independently and uniformly at random out of the interval $[0, 1]$. Detectability conditions of Theorem 1 are satisfied by verifying that the synthetically generated pair $(A, \boldsymbol{c}_j)$ is observable for all $j \in \{1, \dots, n_c\}$.

In Figure 1, the optimal sampling distribution p_ρ^* and its corresponding $\overline{\lambda}(P_U)$ are computed using Theorem 2 for varying ρ values and a fixed number of sampled sensors $n_s = 100$. Figure 1 confirms our discussion in Remark 5, such that minimizing ρ tends to minimize $\overline{\lambda}(P_U)$ in general. Furthermore, the nature of the maximum eigenvalue curve over the regime of feasible ρ values leads us to conjecture that $\overline{\lambda}(P_U)$ is a convex function of ρ. If proven true, the search procedure outlined in Section III-B would be obsolete and the globally minimum $\overline{\lambda}(P_U)$ could be solved directly with minor alterations to Theorem 2.

In Figure 2, the optimal sampling distribution p^* is plotted. Note that sampling distribution p^* is sparse and Figure 2 identifies the small subset of candidate sensors collectively responsible for minimizing $\overline{\lambda}(P_U)$ by the greatest margin.

In Figure 3, the sampling distribution p^* that globally minimizes $\overline{\lambda}(P_U)$ is computed for varying number of sampled sensors and compared against the $\overline{\lambda}(P_U)$ curve for a trivial uniform sampling distribution as a benchmark. Figure 3 shows that the uniform sampling distribution is only applicable for a limited regime of sampled sensors. In fact, if too few sensors are sampled, then the probabilistic guarantees of Theorem 1 no longer hold. In contrast, the $\overline{\lambda}(P_U)$ curve for the optimal sampling distribution p^* requires significantly fewer sampled sensors to substantially minimize the maximum eigenvalue of upper bound P_U.

In Figure 4, quantities $\overline{\lambda}(P_U)$, $\overline{\lambda}(P_L)$, and $\overline{\lambda}(P_{\mathcal{S}})$ are plotted for varying number of sampled sensors and compared against the $\overline{\lambda}(P)$ obtained via a greedy heuristic. For each n_s the maximum eigenvalue of bounds P_U and P_L are computed using their corresponding optimal sampling distribution p^*. Similarly, for each n_s and corresponding optimal distribution p^*, the average maximum eigenvalue of steady-state solution $P_{\mathcal{S}}$ is estimated by 100 Monte Carlo trials. In the greedy heuristic, n_s sensors are sampled *with replacement* out of the sampling pool of candidate sensors. At each sampling instant, the candidate sensor that minimizes $\overline{\lambda}(P)$ is greedily chosen. Figure 4 shows that the average $\overline{\lambda}(P_{\mathcal{S}})$ generated by sampling distribution p^* is consistently smaller than the $\overline{\lambda}(P)$ of the greedy heuristic. In other words, the average $\overline{\lambda}(P_{\mathcal{S}})$ consistently outperforms the greedily chosen $\overline{\lambda}(P)$.

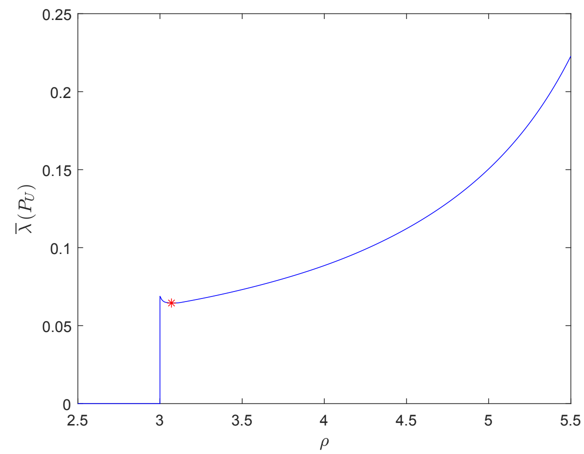

Fig. 1. Maximum eigenvalue of the upper bound P_U for a limited regime of ρ values. Non-zero values of $\overline{\lambda}(P_U)$ indicate feasible ρ values. The red asterisk locates the optimal ρ^* value that globally minimizes $\overline{\lambda}(P_U)$.

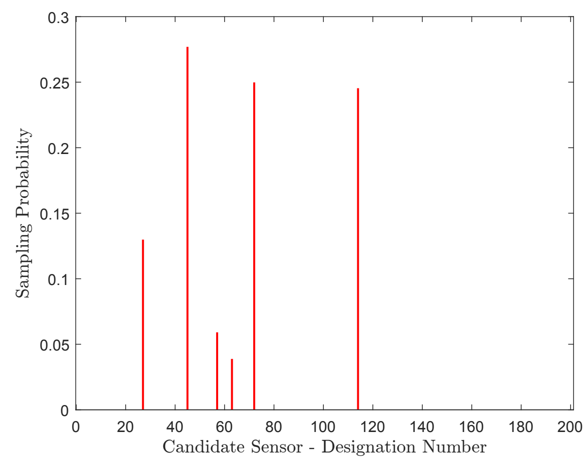

Fig. 2. Optimal sampling distribution p^* that globally minimizes $\overline{\lambda}(P_U)$.

V. Conclusion

In this paper, we consider the sensor selection problem in the context of state estimation for the Kalman filter. Novel bounds on the steady-state error covariance of a randomly sampled sensor selection were derived in the probabilistic sense. We confirmed that the sampling distribution that minimizes the maximum eigenvalue of the upper bound indirectly minimizes the maximum eigenvalue of the steady-state error covariance. Our simulations demonstrated that the optimal sampling distribution significantly outperforms the trivial uniform sampling distribution in terms of the maximum eigenvalue of the upper bound. A numerical analysis showed that the maximum eigenvalue of the steady-state error covariance generated by the optimal sampling distribution consistently outperforms on-average a greedy heuristic. Our results are expected to be significant in the analysis of large sensor networks, since manually choosing the sampling distribution that minimizes a non-trivial objective function is infeasible.

Future directions include extending our analytical guarantees on estimation performance to the constrained setting, where each candidate sensor is limited in availability and cannot be sampled *with replacement* indefinitely.

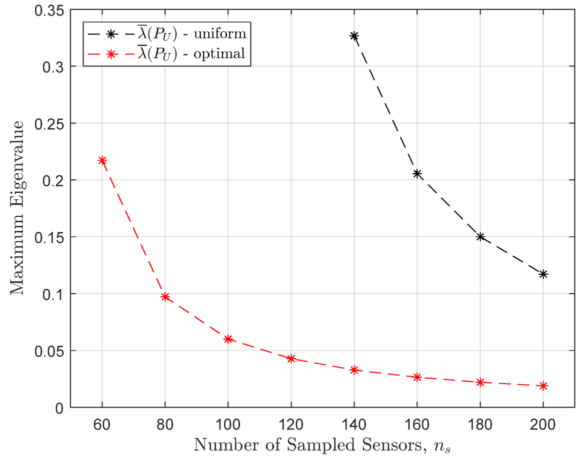

Fig. 3. A numerical comparison between the uniform (black-star) and optimal (red-star) sampling distribution for the maximum eigenvalue of the upper bound P_U for varying number of sampled sensors n_s.

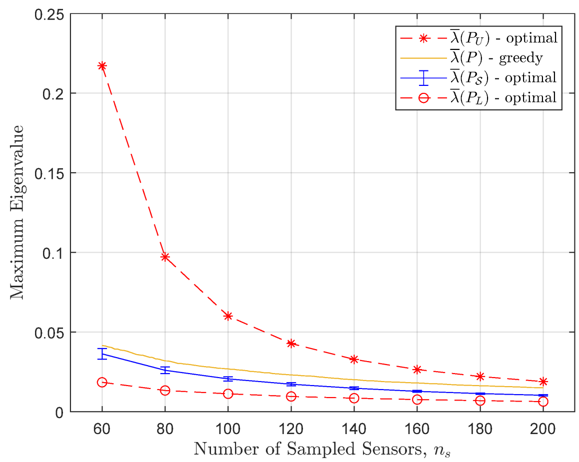

Fig. 4. Maximum eigenvalue for varying number of sampled sensors n_s, such that the maximum eigenvalue of P_U and P_L are indicated by the (red-star) and (red-circle) curve, respectively. The maximum eigenvalue of the steady-state error covariance obtained per a greedy heuristic is indicated by the (yellow-line) curve. The average maximum eigenvalue of P_S is indicated by the (blue-line) curve and variability of $\overline{\lambda}(P_S)$ is captured by the standard deviation. The error bars indicate $\pm$ one standard deviation.

References

[1] "The National Severe Storms Laboratory." https://www.nssl.noaa.gov/tools/observation/.

[2] J.-C. Chiny, I.-H. Houz, J. C. Houz, C. May, N. S. Rao, M. Saxenay, M. Shankar, Y. Yangz, and D. K. Yau, "A sensor-cyber network testbed for plume detection, identification, and tracking," in *Proceedings of the 6th international conference on Information processing in sensor networks*, pp. 541–542, 2007.

[3] J. Zhang, Y. Lu, Z. Lu, C. Liu, G. Sun, and Z. Li, "A new smart traffic monitoring method using embedded cement-based piezoelectric sensors," *Smart Materials and Structures*, vol. 24, no. 2, p. 025023, 2015.

[4] P. Müller and H. Weber, "Analysis and optimization of certain qualities of controllability and observability for linear dynamical systems," *Automatica*, vol. 8, no. 3, pp. 237–246, 1972.

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

[6] T. H. Summers, F. L. Cortesi, and J. Lygeros, "On submodularity and controllability in complex dynamical networks," *IEEE Transactions on Control of Network Systems*, vol. 3, no. 1, pp. 91–101, 2015.

[7] S. Pequito, S. Kar, and A. P. Aguiar, "Minimum cost input/output design for large-scale linear structural systems," *Automatica*, vol. 68, pp. 384–391, 2016.

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

[9] 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.

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

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

[12] 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.

[13] L. F. Chamon, G. J. Pappas, and A. Ribeiro, "Approximate supermodularity of kalman filter sensor selection," *IEEE Transactions on Automatic Control*, vol. 66, no. 1, pp. 49–63, 2020.

[14] A. Hashemi, M. Ghasemi, H. Vikalo, and U. Topcu, "Randomized greedy sensor selection: Leveraging weak submodularity," *IEEE Transactions on Automatic Control*, 2020.

[15] J. A. Tropp, "An introduction to matrix concentration inequalities," *Foundations and Trends® in Machine Learning*, vol. 8, no. 1-2, pp. 1–230, 2015.

[16] S. D. Bopardikar, O. Ennasr, and X. Tan, "Randomized sensor selection for nonlinear systems with application to target localization," *IEEE Robotics and Automation Letters*, vol. 4, no. 4, pp. 3553–3560, 2019.

[17] S. D. Bopardikar, "A randomized approach to sensor placement with observability assurance," *Automatica*, vol. 123, p. 109340, 2021.

[18] R. Ahlswede and A. Winter, "Strong converse for identification via quantum channels," *IEEE Transactions on Information Theory*, vol. 48, no. 3, pp. 569–579, 2002.

[19] D. Gross and V. Nesme, "Note on sampling without replacing from a finite collection of matrices," 2010. [Online]. Available: https://arxiv.org/abs/1001.2738.

[20] C. I. Calle and S. D. Bopardikar, "Probabilistic performance bounds for randomized sensor selection in Kalman filtering," 2021. [Online]. Available: https://arxiv.org/abs/2103.11182.

[21] B. D. Anderson and J. B. Moore, *Optimal filtering*. Courier Corporation, 2012.

[22] R. Qiu and M. Wicks, "Sums of matrix-valued random variables," in *Cognitive Networked Sensing and Big Data*, pp. 85–144, Springer, 2014.

Appendix

A non-trivial inequality from random matrix theory, known as the Ahlswede-Winter inequality [18], allows us to bound sums of independent positive semi-definite matrices.

Theorem 3. ***(Ahlswede-Winter Inequality)*** *Let Z be a random, symmetric, positive semi-definite $m \times m$ matrix. Define $U = \mathbb{E}[Z]$ and suppose that $Z \preceq \rho U$ almost surely, for some scalar $\rho \geq 1$. Let $Z_1, \dots, Z_{n_s}$ denote independent copies of Z, i.e., independently sampled matrices with the same distribution as Z. For any $\epsilon \in (0,1)$, we have*

$$\mathbb{P}\left[(1-\epsilon)\,U \preceq \frac{1}{n_s}\sum_{j=1}^{n_s} Z_j \preceq (1+\epsilon)\,U\right] \geq 1 - 2m\,\mathrm{e}^{-\frac{\epsilon^2 n_s}{4\rho}}.$$

Theorem 3 is Corollary 2.2.2 in [22] with minor alterations to the notation.