On Sensor Scheduling of Linear Dynamical Systems with Error Bounds

Michael P. Vitus†, Wei Zhang†, Alessandro Abate, Jianghai Hu and Claire J. Tomlin

***Abstract*— Consider a set of sensors estimating the state of a process in which only one of these sensors can operate at each time-step due to constraints on the overall system. The problem addressed here is to choose which sensor should operate at each time-step to minimize a function of the error covariance of the state estimation at each time-step. Previously, the authors developed tractable algorithms to solve for the optimal and suboptimal sensor schedule. The suboptimal algorithm trades off the quality of the solution and the complexity of the problem through a tuning parameter. As the tuning parameter is increased the complexity of the problem significantly decreases but the overall affect on the quality of the solution has not been completely characterized as of yet. This work concentrates on developing an upper bound on the distance from the optimal solution through two different approaches. The first approach exploits the peak estimation error, and the second method decomposes the covariance into two factors to linearly propagate the effects of perturbations. Numerical simulations are also performed to demonstrate the performance of the suboptimal algorithm for various tuning parameters.**

I. Introduction

With the advances of sensor networks and the improvement of unmanned systems for reconnaissance and surveillance missions, the environment is being inundated with sensor networks monitoring external processes. Some examples are presented in [1], [2], [3], [4]. One paradigm for these networks has each node perform some local processing of the data and then transmit it to a central aggregation process. Constraints on the network's communication bandwidth might not allow all of the nodes to communicate at each time-step. Also, each node may only have a limited amount of power and therefore should be turned off to conserve power when its measurement is not required. In addition, sensors may interfere with one another, as with sonar range-finding sensors, and thus cannot operate at the same time. Consequently, the objective is to manage the schedule of nodes' measurements. In this work, the problem of *sensor scheduling* is to select one out of multiple available sensors at each time-step to minimize the sum of all the estimation errors over a certain time-horizon.

In a seminal work, Meier et al. [5] proposed a solution to the discrete time scheduling problem through the use of dynamic programming which enumerates all possible sensor schedules; the combinatorial complexity makes this method intractable for long schedule horizons. A local gradient method was also proposed which is more likely to be computationally feasible, but only provides a suboptimal solution with no performance guarantees. In [6], a relaxed dynamic programming procedure is applied to obtain a suboptimal strategy that is bounded by a prespecified distance to optimality, but is only applicable for an objective function that minimizes the final step estimation error. Another approach [7] switched sensors randomly according to a probability distribution to obtain the best upper bound on the expected steady-state performance but without any performance guarantees. The sensor scheduling problem has also been formulated as a partially observable Markov decision process [8], [9]. In [10], a condition was presented that characterized when the initialization of a sensor schedule is not optimal. Using this condition, three efficient algorithms for finding the optimal and suboptimal solution were developed. Although the algorithms significantly reduced the computational complexity, there was no bound on the distance from the optimal solution.

This work extends our previous contributions of [10] by developing two different upper bounds on the performance of the suboptimal solution. To develop the bounds, both approaches utilize a novel decomposition of the objective function and characterize the effect of small perturbations in the initial covariance matrix of the Riccati mapping. The first method uses the peak estimation error to develop the bound, and although the exact peak estimation error is not readily available, a conservative upper bound is obtained. The second approach uses a factorization of the covariance matrix that enables the combination of the process and measurement updates into a linear transfer function. This decomposition allows the effect of a small perturbation to be exactly determined. To enable the efficient computation of the numerical value of the bound an approximation is proposed; simulations are performed to ensure the consistency of this approximation. Finally, numerical simulations are also conducted to demonstrate the typical performance for various tuning parameters for the suboptimal algorithm.

The paper proceeds as follows. Section II describes the standard sensor scheduling problem formulation. Then, several properties of the objective function are explored and a theorem which is useful for pruning branches in the search tree is presented in Section III. In Section IV, a brief description of two tractable algorithms for determining the optimal and suboptimal solutions is provided, and two bounds for the performance of the suboptimal algorithm are developed in Section V. The paper concludes with directions of future work.

II. Problem Formulation

Consider the following linear stochastic system defined by,

$$x(k+1) = Ax(k) + w(k), \forall k \in T_N, \tag{1}$$

Michael P. Vitus is with the Department of Aeronautics and Astronautics, Stanford University, CA 94305. Email: vitus@stanford.edu

Alessandro Abate is with the Department of Mechanical, Maritime and Materials Engineering, Delft University of Technology, Netherlands. Email: a.abate@tudelft.nl

Jianghai Hu is with the School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47906. Email: jianghai@purdue.edu

Wei Zhang and Claire J. Tomlin are with the Department of Electrical Engineering and Computer Sciences, University of California at Berkeley. Email: {weizhang,tomlin}@eecs.berkeley.edu

†These authors contributed equally to this work.

where $x(k) \in \mathbb{R}^n$ is the state of the system, $w(k) \in \mathbb{R}^n$ is the process noise and $T_N = \{0, \ldots, N-1\}$ is the horizon. The initial state, $x(0)$, is assumed to be a zero mean Gaussian distribution with covariance Σ_0 i.e., $x(0) \sim \mathcal{N}(0, \Sigma_0)$. At each time step, only one sensor is allowed to operate from a set of M sensors. The measurement of the i^{th} sensor is,

$$y_i(k) = C_i x(k) + v_i(k), \forall k \in T_N, \tag{2}$$

where $y_i(k) \in \mathbb{R}^p$ and $v_i(k) \in \mathbb{R}^p$ are the measurement output and noise of the i^{th} sensor at time k, respectively. The process and measurement noise have zero mean Gaussian distributions, $w(k) \sim \mathcal{N}(0, \Sigma_w)$, $v_i(k) \sim \mathcal{N}(0, \Sigma_{v_i})$, $\forall i \in \mathbb{M}$, where $\mathbb{M} \triangleq \{1, \ldots, M\}$ is the set of M sensors. The process noise, measurement noise and initial state are assumed to be mutually independent. Let λ_w^- be the smallest eigenvalue of Σ_w and assume that $\lambda_w^- > 0$. Denote by $\mathbb{M}^t$ the set of all ordered sequences of sensor schedules of length t where $t \leq N$. An element $\sigma = \{\sigma_0, \sigma_1, \ldots, \sigma_{t-1}\} \in \mathbb{M}^t$ is called a (t-horizon) *sensor schedule*. Under a given sensor schedule σ, the measurement sequence is,

$$y(k) = y_{\sigma_k}(k) = C_{\sigma_k} x(k) + v_{\sigma_k}(k), \forall k \in \{0, \ldots, t-1\}.$$

For each $k \leq t$ with $t \leq N$ and each $\sigma \in \mathbb{M}^t$, let $\hat{\Sigma}_k^\sigma$ be the predictor covariance matrix of the optimal estimate of $x(k)$ given the measurements $\{y(0), \ldots, y(k-1)\}$. By a standard result of linear estimation theory, the Kalman filter is the minimum mean square error estimator, and the predictor covariance of the system state estimate evolves according to the Riccati recursion,

$$\begin{aligned}\hat{\Sigma}_{k+1}^\sigma = & A\hat{\Sigma}_k^\sigma A^{\mathrm{T}} + \Sigma_w - \\ & A\hat{\Sigma}_k^\sigma C_{\sigma_k}^{\mathrm{T}} \left(C_{\sigma_k} \hat{\Sigma}_k^\sigma C_{\sigma_k}^{\mathrm{T}} + \Sigma_{v_{\sigma_k}} \right)^{-1} C_{\sigma_k} \hat{\Sigma}_k^\sigma A^{\mathrm{T}}\end{aligned} \tag{3}$$

with initial condition $\hat{\Sigma}_0^\sigma = \Sigma_0$ and $k \leq t$. Let $\mathbb{R}_+$ and $\mathbb{Z}_+$ be the set of nonnegative real numbers and integers, respectively. Define $J(\sigma): \mathbb{M}^t \to \mathbb{R}_+$ as the accrued estimation error under schedule σ, i.e., $J(\sigma) = \sum_{k=1}^t \mathrm{tr}\left(\hat{\Sigma}_k^\sigma\right)$. The sensor scheduling problem is defined as solving the following discrete optimization problem,

$$V = \min_{\sigma \in \mathbb{M}^N} J(\sigma). \tag{4}$$

III. Properties of the Objective Function

Let $\mathcal{A}$ denote the *positive semidefinite cone*, which is the set of all symmetric positive semidefinite matrices. A *Riccati Mapping* $\rho_i : \mathcal{A} \to \mathcal{A}$ can be defined, which maps the current covariance matrix, $\hat{\Sigma}_k$, under a new measurement from sensor $i \in \mathbb{M}$ to the next covariance matrix,

$$\begin{aligned}\rho_i(\hat{\Sigma}_k) = & A\hat{\Sigma}_k A^{\mathrm{T}} - \\ & A\hat{\Sigma}_k C_i^{\mathrm{T}} \left(C_i \hat{\Sigma}_k C_i^{\mathrm{T}} + \Sigma_{v_i} \right)^{-1} C_i \hat{\Sigma}_k A^{\mathrm{T}} + \Sigma_w.\end{aligned} \tag{5}$$

A k-horizon Riccati mapping, $\phi_k^\sigma : \mathcal{A} \to \mathcal{A}$ can similarly be defined, which maps the covariance matrix at time 0, Σ_0, to the covariance matrix at time-step k, using the first k elements of the sensor schedule σ,

$$\phi_k^\sigma(\Sigma_0) = \rho_{\sigma_{k-1}}(\ldots \rho_{\sigma_1}(\rho_{\sigma_0}(\Sigma_0))). \tag{6}$$

Definition 1 (Characteristic Sets): Let $\{\mathcal{H}_k\}_{k=0}^N$ be defined as the characteristic sets as they completely characterize the objective function. Each set in the sequence has elements of the form $(\Sigma, \gamma) \in \mathcal{A} \times \mathbb{R}_+$ and is generated recursively by:

$$\mathcal{H}_{k+1} = h_{\mathbb{M}}(\mathcal{H}_k) \text{ from } \mathcal{H}_0 = \{(\Sigma_0, \mathrm{tr}(\Sigma_0))\} \text{ where}$$

$h_{\mathbb{M}}(\mathcal{H}) = \{(\rho_i(\Sigma), \gamma + \mathrm{tr}(\rho_i(\Sigma))) : \forall i \in \mathbb{M}, \ \forall (\Sigma, \gamma) \in \mathcal{H}\}$.

Let $h_{\mathbb{M}}(\cdot)$ be referred to as the characteristic set mapping. The characteristic sets grow exponentially in size from the singleton set $\{(\Sigma_0, \mathrm{tr}(\Sigma_0))\}$ to the set $\mathcal{H}_N$ consisting of up to M^N pairs each comprising of a positive semidefinite matrix and an accrued cost. These sets characterize the covariance of the estimate and the objective cost at every time-step under every possible sensor schedule.

Let $\mathcal{H}_k(i) = (\Sigma_k(i), \gamma_k(i))$ be the i^{th} element of the set $\mathcal{H}_k$, where $\Sigma_k(i)$ and $\gamma_k(i)$ are the covariance matrix and objective cost, respectively. For any $(\Sigma, \gamma) \in \mathcal{H}_k$, denote by $\sigma(\Sigma, \gamma)$ the corresponding sensor schedule. Also, for any $\sigma \in \mathbb{M}^k$, denote by $(\hat{\Sigma}_k^\sigma, \gamma_k^\sigma)$ the corresponding pair in $\mathcal{H}_k$. If $\mathcal{H} \subset \mathcal{H}_k$, then the set of schedules corresponding to $\mathcal{H}$ is defined by,

$$\mathbb{M}(\mathcal{H}) = \{\sigma \in \mathbb{M}^k : (\hat{\Sigma}_k^\sigma, \gamma_k^\sigma) \in \mathcal{H}\}. \tag{7}$$

The main idea of the proposed solution methods is motivated by the following properties of the Riccati mapping.

Theorem 1 ([10]): For any $i \in \mathbb{M}$ and any $\Sigma_1, \Sigma_2 \in \mathcal{A}$,
(i) **[Monotonicity]** If $\Sigma_1 \preceq \Sigma_2$, then $\rho_i(\Sigma_1) \preceq \rho_i(\Sigma_2)$;
(ii) **[Concavity]** $\rho_i(c\Sigma_1 + (1-c)\Sigma_2) \succeq c\rho_i(\Sigma_1) + (1-c)\rho_i(\Sigma_2)$, $\forall c \in [0, 1]$.

Thus, systems starting with a larger initial covariance, in the positive semidefinite sense, will yield larger covariances at all future time-steps. This result is important because it provides insight on how to reduce the complexity of the scheduling problem. Theorem 1 can be repeatedly applied to result in the following corollary.

Corollary 1: Let $\sigma \in \mathbb{M}^N$ and $\Sigma_1, \Sigma_2 \in \mathcal{A}$, then $\forall k \in [0, N]$, $\forall c \in [0, 1]$,
(i) If $\Sigma_1 \preceq \Sigma_2$, then $\phi_k^\sigma(\Sigma_1) \preceq \phi_k^\sigma(\Sigma_2)$;
(ii) $\phi_k^\sigma(c\Sigma_1 + (1-c)\Sigma_2) \succeq c\phi_k^\sigma(\Sigma_1) + (1-c)\phi_k^\sigma(\Sigma_2)$.

IV. Solution Methodology

A. Algebraic Redundancy

To enable the solution of larger systems with longer scheduling horizon, it is necessary to prune branches from the search tree that will not lead to the optimal solution.

Definition 2 (Algebraic Redundancy): A pair $(\Sigma, \gamma) \in \mathcal{H}$ is called algebraically redundant with respect to $\mathcal{H} \backslash \{(\Sigma, \gamma)\}$, if there exist nonnegative constants $\{\alpha_i\}_{i=1}^{l-1}$ such that

$$\sum_{i=1}^{l-1} \alpha_i = 1, \quad \text{and} \quad \begin{bmatrix} \Sigma & 0 \\ 0 & \gamma \end{bmatrix} \succeq \sum_{i=1}^{l-1} \alpha_i \begin{bmatrix} \Sigma(i) & 0 \\ 0 & \gamma(i) \end{bmatrix}$$

where $l = |\mathcal{H}|$ and $\{(\Sigma(i), \gamma(i))\}_{i=1}^{l-1}$ is an enumeration of $\mathcal{H} \setminus \{(\Sigma, \gamma)\}$.

Using the results from Corollary 1 and Definition 2, the following theorem provides a condition which characterizes

the branches that can be pruned without eliminating the optimal solution of the sensor scheduling problem.

Theorem 2 ([10]): If the pair $(\Sigma, \gamma) \in \mathcal{H}$ is algebraically redundant, then the branch and all of its descendants can be pruned without eliminating the optimal solution.

In computing the characteristic sets, Definition 2 can be applied to calculate a subset of $\mathcal{H}_k$, $\forall k \in \{1, \ldots, N\}$, which is denoted as $Algo_{AR}$. Theorem 2 guarantees that this subset still contains the optimal solution. This leads to an efficient method for computing the optimal sensor schedule which is outlined in Algorithm 1.

Algorithm 1 Sensor Scheduling for a Finite Horizon

1: $\mathcal{H}_0 = \{(\Sigma_0, \mathrm{tr}\,(\Sigma_0))\}$
2: **for** k = 1,. . .,N **do**
3: $\quad \mathcal{H}_k = Algo_{AR}\left(h_{\mathbb{M}}\left(\mathcal{H}_{k-1}\right)\right)$
4: **end for**
5: $\sigma^* = \underset{j \in \{1,\ldots,|\mathcal{H}_N|\}}{\arg\min} \gamma_N(j)$

Even though this method prunes a large number of branches, the growth of the tree may still become prohibitive for some problems. An approximate solution may be desired.

B. Numerical Redundancy

To further reduce the complexity, the algebraic redundancy concept can be generalized to allow for numerical error. Similar to Definition 2, the following definition provides a condition for testing the ϵ-redundancy of a matrix.

Definition 3 (ϵ-Redundant): A pair $(\Sigma, \gamma) \in \mathcal{H}$ is called ϵ-redundant with respect to $\mathcal{H} \setminus \{(\Sigma, \gamma)\}$, if there exist nonnegative constants $\{\alpha_i\}_{i=1}^{l-1}$ such that

$$\sum_{i=1}^{l-1} \alpha_i = 1, \quad \begin{bmatrix} \Sigma + \epsilon I & 0 \\ 0 & \gamma + \epsilon \end{bmatrix} \succeq \sum_{i=1}^{l-1} \alpha_i \begin{bmatrix} \Sigma(i) & 0 \\ 0 & \gamma(i) \end{bmatrix}$$

where $l = |\mathcal{H}|$ and $\{(\Sigma(i), \gamma(i))\}_{i=1}^{l-1}$ is an enumeration of $\mathcal{H} \setminus \{(\Sigma, \gamma)\}$.

Denote by $Algo_\epsilon\left(\mathcal{H}\right)$ the set of the remaining pairs after removing all the ϵ-redundant pairs in $\mathcal{H}$ that satisfy the conditions given in Definition 3. The following lemma can be proved using Corollary 1 and Definition 3.

Lemma 1 ([10]): For any $\epsilon \in \mathbb{R}_+$, $(\Sigma, \gamma) \in \mathcal{H}_k$ and $\sigma \in \mathbb{M}^{N-k}$, there always exists another pair $\left(\hat{\Sigma}, \hat{\gamma}\right) \in Algo_\epsilon\left(\mathcal{H}_k\right)$ such that $\hat{\gamma} + \sum_{s=1}^{N-k} \mathrm{tr}(\phi_s^\sigma(\hat{\Sigma})) \leq \gamma + \epsilon + \sum_{s=1}^{N-k} \mathrm{tr}(\phi_s^\sigma(\Sigma + \epsilon I))$, for all $k = 1, \ldots, N$.

To determine the ϵ-approximate solution of the sensor scheduling problem, Algorithm 1 can be modified by substituting $Algo_\epsilon$ for $Algo_{AR}$. Define the ϵ-relaxed characteristic sets $\{\mathcal{H}_k^\epsilon\}_{k=0}^N$ by

$$\mathcal{H}_k^\epsilon = Algo_\epsilon(h_{\mathbb{M}}\left(\mathcal{H}_{k-1}^\epsilon\right)), \text{with } \mathcal{H}_0^\epsilon = \{(\Sigma_0, \mathrm{tr}\,(\Sigma_0))\}. \quad (8)$$

The set $\mathcal{H}_N^\epsilon$ typically contains many fewer pairs than $\mathcal{H}_N$ and is much easier to compute. To simplify the computation, the schedule that minimizes $V_N(\sigma)$ among all the schedules in $\mathbb{M}\left(\mathcal{H}_N^\epsilon\right)$ can be used as an alternative to the optimal schedule. While the suboptimal algorithm drastically reduces the computational complexity, it might sacrifice the quality of the solution too much. Consequently, an upper bound on the distance from the optimal solution is needed.

V. Performance Analysis

To aid the discussion, the relaxed value function V_k^ϵ is introduced and defined as,

$$V_k^\epsilon = \min_{\sigma \in \mathbb{M}(\mathcal{H}_k^\epsilon)} J_k\left(\sigma\right) = \min_{(\Sigma,\gamma) \in \mathcal{H}_k^\epsilon} \gamma. \quad (9)$$

The goal of this section is to derive an upper bound for the error, $V_N^\epsilon - V_N$, incurred by the relaxation in Eqn. (8). For each $j = 1, \ldots, N$, let $\{\mathcal{H}_k^\epsilon\}_{k=0}^N$ be generated by

$$\mathcal{H}_k^{\epsilon,j} = \begin{cases} Algo_\epsilon(h_{\mathbb{M}}\left(\mathcal{H}_{k-1}^{\epsilon,j}\right)) & \forall k < j \\ h_{\mathbb{M}}\left(\mathcal{H}_{k-1}^{\epsilon,j}\right) & \forall k \geq j \end{cases} \quad (10)$$

with $\mathcal{H}_0^{\epsilon,j} = \{(\Sigma_0, \mathrm{tr}\,(\Sigma_0))\}$. The set $\mathcal{H}_k^{\epsilon,j}$ evolves (as a function of k) according to the ϵ-relaxed iteration for the first j steps and then evolves according to the non-relaxed iteration for the remaining steps. When $j = 0$, define $\mathcal{H}_k^{\epsilon,0} = \mathcal{H}_k^\epsilon$ for $k = 0, \ldots, N$. For any $j = 1, \ldots, N$, $\mathcal{H}_k^{\epsilon,j}$ coincides with $\mathcal{H}_k^\epsilon$ for all $k < j$, and the set $\mathcal{H}_N^{\epsilon,j}$ decreases monotonically to $\mathcal{H}_N^\epsilon$ as $j \uparrow N$. Let $\sigma^{\epsilon,j}$ be the "optimal" schedule within $\mathbb{M}(\mathcal{H}_N^{\epsilon,j})$ defined by

$$\sigma^{\epsilon,j} = \underset{\sigma \in \mathbb{M}(\mathcal{H}_N^{\epsilon,j})}{\arg\min} J_N(\sigma), \text{ for } j = 0, \ldots, N. \quad (11)$$

It should be noted that $\sigma^{\epsilon,0}$ coincides with the optimal sensor schedule within $\mathbb{M}^N$ and $\sigma^{\epsilon,N}$ is the best sensor schedule among the set $\mathbb{M}(\mathcal{H}_N^\epsilon)$. For each $j = 1, \ldots, N-1$, the performance of the schedule $\sigma^{\epsilon,j}$ is sandwiched between the solutions corresponding to $\sigma^{\epsilon,0}$ and $\sigma^{\epsilon,N}$. For each j, $k = 0, \ldots, N$, define

$$V_k^{\epsilon,j} = \min_{\sigma \in \mathbb{M}(\mathcal{H}_k^{\epsilon,j})} J_k\left(\sigma\right) = \min_{(\Sigma,\gamma) \in \mathcal{H}_k^{\epsilon,j}} \gamma. \quad (12)$$

Thus the total error can be decomposed by,

$$V_N^\epsilon - V_N = \sum_{j=1}^{N} \left(V_N^{\epsilon,j} - V_N^{\epsilon,j-1}\right). \quad (13)$$

Note that due to the monotonicity of $V_N^{\epsilon,j}$, each term in the summation is nonnegative.

To develop an analytical expression for the bound of the error $V_N^\epsilon - V_N$, the effect of a perturbation of the initial covariance on all future covariances must be determined. Suppose this effect of the perturbation can be quantified through the matrix-valued error function $\Theta_k^\sigma\left(\epsilon, \phi_k^\sigma\left(\Sigma\right)\right)$ such that $\phi_k^\sigma\left(\Sigma + \epsilon I\right) \preceq \phi_k^\sigma\left(\Sigma\right) + \Theta_k^\sigma\left(\epsilon, \phi_k^\sigma\left(\Sigma\right)\right)$ then the total error incurred can be obtained in the following theorem.

Theorem 3: For any $N \in \mathbb{Z}_+$, the error is bounded by

$$V_N^\epsilon - V_N \leq \sum_{j=1}^{N} \left(\epsilon + \sum_{k=1}^{N-j} \mathrm{tr}\left(\Theta_{k+j-1}^{\sigma^{\epsilon,j}}\left(\epsilon, \phi_{k+j-1}^{\sigma^{\epsilon,j}}\left(\Sigma_0\right)\right)\right)\right).$$

Proof: From Eqn. (13), it suffices to bound $V_N^{\epsilon,j} - V_N^{\epsilon,j-1}$ for all $j = 1, \ldots, N$. Recall that $\mathcal{H}_k^{\epsilon,j} = \mathcal{H}_k^{\epsilon,j-1}$ for all $k \leq j-1$ and that $\mathcal{H}_k^{\epsilon,j} = Algo_\epsilon\left(\mathcal{H}_k^{\epsilon,j-1}\right)$. Using the definition of $\phi_k^\sigma\left(\Sigma\right)$ in Eqn. (6),

$$V_N^{\epsilon,j} = \min_{\substack{(\Sigma,\gamma)\in Algo_\epsilon\left(\mathcal{H}_k^{\epsilon,j-1}\right)\\ \sigma\in\mathbb{M}^{N-j}}} \gamma + \sum_{k=1}^{N-j} \mathrm{tr}\left(\phi_k^\sigma\left(\Sigma\right)\right).$$

From Lemma 1,

$$V_N^{\epsilon,j} \leq \min_{\substack{(\Sigma,\gamma)\in\mathcal{H}_j^{\epsilon,j-1}\\ \sigma\in\mathbb{M}^{N-j}}} \gamma + \epsilon + \sum_{k=1}^{N-j} \mathrm{tr}\left(\phi_k^\sigma\left(\Sigma+\epsilon I\right)\right).$$

If the term, $\phi_k^\sigma\left(\Sigma+\epsilon I\right)$, can be upper bounded such that $\phi_k^\sigma\left(\Sigma+\epsilon I\right) \preceq \phi_k^\sigma\left(\Sigma\right) + \Theta_k^\sigma\left(\epsilon, \phi_k^\sigma\left(\Sigma\right)\right)$ then,

$$V_N^{\epsilon,j} \leq \min_{\substack{(\Sigma,\gamma)\in\mathcal{H}_j^{\epsilon,j-1}\\ \sigma\in\mathbb{M}^{N-j}}} \begin{array}{c}\gamma + \epsilon + \sum_{k=1}^{N-j}\left[\mathrm{tr}\left(\phi_k^\sigma\left(\Sigma\right)\right) + \right.\\ \left.\mathrm{tr}\left(\Theta_k^\sigma\left(\epsilon, \phi_k^\sigma\left(\Sigma\right)\right)\right)\right]\end{array}$$

$$\leq V_N^{\epsilon,j-1} + \epsilon + \sum_{k=j+1}^{N} \mathrm{tr}\left(\Theta_k^{\sigma^{\epsilon,j}}\left(\epsilon, \phi_k^{\sigma^{\epsilon,j}}\left(\Sigma_0\right)\right)\right).$$

The desired result then follows from Eqn. (13) by summing over all $j = 1, \ldots, N$. ∎

According to the previous theorem, once the error function Θ_k^σ is obtained, the total error can easily be determined. In the following sections, two different approaches are explored to bound Θ_k^σ which leads to two different total error bounds.

A. Error Bound Through the Peak Estimation Error

1) Perturbation: An upper bound for the perturbed Riccati mapping is first derived. Let for each $i \in \mathbb{M}$ and $\Sigma \in \mathcal{A}$,

$$\bar{A}_i(\Sigma) \triangleq A - AK_i(\Sigma)C_i, \tag{14}$$

where $K_i(\Sigma)$ is the Kalman gain defined as,

$$K_i(\Sigma) = \Sigma C_i^{\mathrm{T}}(C_i\Sigma C_i^{\mathrm{T}} + \Sigma_{v_i})^{-1}.$$

Lemma 2: For each $i \in \mathbb{M}$ and any $\Sigma, Q \in \mathcal{A}$,

$$\left.\frac{d\rho_i(\Sigma+\epsilon Q)}{d\epsilon}\right|_{\epsilon=0} = \bar{A}_i(\Sigma)Q\bar{A}_i(\Sigma)^{\mathrm{T}},$$

where $\bar{A}_i(\Sigma)$ is defined in (14).

Proof: Let $i \in \mathbb{M}$, $\Sigma \in \mathcal{A}$ and $Q \in \mathcal{A}$ be arbitrary but fixed. Define $f(\epsilon) = C_i(\Sigma+\epsilon Q)C_i^{\mathrm{T}} + \Sigma_{v_i}$. It can be easily shown that,

$$\frac{df^{-1}(\epsilon)}{d\epsilon} = -f^{-1}(\epsilon)C_iQC_i^{\mathrm{T}}f^{-1}(\epsilon).$$

Taking the derivative of $\rho_i(\Sigma+\epsilon Q)$ with respect to ϵ and letting $\epsilon = 0$ yields

$$\left.\frac{d\rho_i(\Sigma+\epsilon Q)}{d\epsilon}\right|_{\epsilon=0} = A\left[(I - \Sigma C_i^{\mathrm{T}}f^{-1}(0)C_i)Q(I - C_i^{\mathrm{T}}f^{-1}(0)C_i\Sigma)\right]A^{\mathrm{T}}.$$

Noting that $f^{-1}(0) = \left(C_i\Sigma C_i^{\mathrm{T}} + \Sigma_{v_i}\right)^{-1}$ and by the definition of $\bar{A}_i(\Sigma)$, the desired result is obtained. ∎

By the concavity of the Riccati mapping (Theorem 1), it can be easily verified that the mapping $\mu_{i,\Sigma,Q} : \mathbb{R}_+ \to \mathcal{A}$ defined by $\mu_{i,\Sigma,Q}(\epsilon) = \rho_i(\Sigma+\epsilon Q)$, $\forall \epsilon \in \mathbb{R}_+$ is also concave for any $i \in \mathbb{M}$, $\Sigma \in \mathcal{A}$ and $Q \in \mathcal{A}$. Thus $\mu_{i,\Sigma,Q}(\epsilon)$ can be upper bounded by an affine function of ϵ, namely, $\mu_{i,\Sigma,Q}(0) + \mu'_{i,\Sigma,Q}(0)\epsilon$, which implies,

$$\begin{array}{c}\rho_i(\Sigma+\epsilon Q) \preceq \rho_i(\Sigma) + \left(\bar{A}_i(\Sigma)Q\bar{A}_i(\Sigma)^{\mathrm{T}}\right)\epsilon,\\ \forall \epsilon \in \mathbb{R}_+, i \in \mathbb{M} \text{ and } \Sigma, Q \in \mathcal{A}.\end{array} \tag{15}$$

Suppose that at some generic time k, the covariance matrix is perturbed from Σ to $\Sigma + \epsilon I$. An upper bound for the k-step effect of this perturbation, $\phi_k^\sigma(\Sigma + \epsilon I) - \phi_k^\sigma(\Sigma)$, is desired.

Lemma 3: For each $k = 1, \ldots, N$ and any $\Sigma \in \mathcal{A}$,

$$g_k^\sigma(\Sigma) \triangleq \left.\frac{d\phi_k^\sigma(\Sigma+\epsilon I)}{d\epsilon}\right|_{\epsilon=0} = \prod_{t=k-1}^{0}\left(\bar{A}_{\sigma(t)}(\phi_t^\sigma(\Sigma))\right)\prod_{t=0}^{k-1}\left(\bar{A}_{\sigma(t)}(\phi_t^\sigma(\Sigma))\right)^{\mathrm{T}}.$$

Proof: For simplicity, let $\hat{A}_t = \bar{A}_{\sigma(t)}(\phi_t^\sigma(\Sigma))$. The case $k = 1$ follows from Lemma 2. Suppose that the result holds for a general $k \leq N-1$,

$$\phi_k^\sigma(\Sigma+\epsilon I) = \phi_k^\sigma(\Sigma) + \left[\prod_{t=k-1}^{0}\hat{A}_t\prod_{t=0}^{k-1}\hat{A}_t^{\mathrm{T}}\right]\epsilon + o(\epsilon),$$

where $o(\epsilon)$ satisfies $o(\epsilon)/\epsilon \to 0$ as $\epsilon \to 0$. Now, it has to be shown that it is also true for $k+1$. Notice that,

$$\begin{aligned}\phi_{k+1}^\sigma(\Sigma+\epsilon I) &= \rho_{\sigma(k)}(\phi_k^\sigma(\Sigma+\epsilon I))\\ &= \rho_{\sigma(k)}\left(\phi_k^\sigma(\Sigma) + \left[\prod_{t=k-1}^{0}\hat{A}_t\prod_{t=0}^{k-1}\hat{A}_t^{\mathrm{T}}\right]\epsilon + o(\epsilon)\right).\end{aligned}$$

Applying Lemma 2 to the right-hand side will yield the desired result. ∎

Similar to Eqn. (15), an affine upper bound for $\phi_k^\sigma(\Sigma+\epsilon I)$ can be obtained using Lemma 3.

Proposition 1: For any $\Sigma \in \mathcal{A}$, $\epsilon \in \mathbb{R}_+$ and $k = 0, \ldots, N$, the k-step effect of a perturbation, $\Sigma + \epsilon I$, can be upper bounded by $\phi_k^\sigma(\Sigma+\epsilon I) \preceq \phi_k^\sigma(\Sigma) + g_k^\sigma(\Sigma)\epsilon$.

The function $g_k^\sigma(\Sigma)$ quantifies how a perturbation error incurred at some generic time t will affect the error covariance matrix at k iterations later provided that no further perturbation is applied after step t. By utilizing this function, an upper bound on the overall error incurred through the numerical redundancy pruning algorithm can now be determined.

2) Error Bound: For each $j = 1, \ldots, N$, to quantify the error $V_N^{\epsilon,j} - V_N^{\epsilon,j-1}$, an arbitrary N-horizon schedule $\sigma \in \mathbb{M}^N$ is decomposed into a j-horizon schedule σ' and an $(N-j)$-horizon schedule σ'', i.e., $\sigma = \{\sigma', \sigma''\}$. Also, let G_j^σ be defined as,

$$G_j^\sigma \triangleq \sum_{k=1}^{N-j} g_k^{\sigma''}(\phi_j^{\sigma'}(\Sigma_0)). \tag{16}$$

It will become clear later that to obtain the desired bound, it suffices to bound G_j^σ along the "optimal" schedule within $\mathbb{M}(\mathcal{H}_N^{\epsilon,j})$. For each $\sigma \in \mathbb{M}^N$, define the peak estimation error by $\mathcal{E}^\sigma \triangleq \max_{k=1,\ldots,N} \mathrm{tr}(\phi_k^\sigma(\Sigma_0))$. Define the peak estimation error over all schedules $\{\sigma^{\epsilon,j}\}_{j=0}^N$ as,

$$\beta = \max_{j=0,\ldots,N} \mathcal{E}^{\sigma^{\epsilon,j}}. \tag{17}$$

Using the peak estimation error β an upper bound for $G_j^{\sigma^{\epsilon,j}}$ can be derived.

Lemma 4: Let β be the constant given in Eqn. (17). Then for each $j = 0, \ldots, N$, the term $G_j^{\sigma^{\epsilon,j}}$ can be upper bounded by,

$$G_j^{\sigma^{\epsilon,j}} \leq \frac{n\beta\eta}{\lambda_w^-(1-\eta)},$$

where $\sigma^{\epsilon,j}$ is defined in (11) and

$$\alpha = \frac{\lambda_w^-}{\|A\|^2\beta^2 + \lambda_w^-\beta}, \; \eta = \frac{1}{1+\alpha\lambda_w^-}.$$

Proof: See [11]. ■

Theorem 4: Let β, α and η be given as in Lemma 4. For any $N \in \mathbb{Z}_+$,

$$\frac{1}{N}\left(V_N^\epsilon - V_N\right) \leq \left(\frac{n\beta\eta}{\lambda_w^-(1-\eta)} + 1\right)\epsilon.$$

Proof: From Theorem 3,

$$\begin{aligned} V_N^{\epsilon,j} \leq & \min_{\substack{(\Sigma,\gamma)\in\mathcal{H}_j^{\epsilon,j-1} \\ \sigma\in\mathbb{M}^{N-j}}} \gamma + \epsilon + \textstyle\sum_{k=1}^{N-j} [\mathrm{tr}(\phi_k^\sigma(\Sigma)) \\ & \qquad + \mathrm{tr}(g_k^\sigma(\Sigma))\epsilon] \\ \leq & V_N^{\epsilon,j-1} + \epsilon + G_j^{\sigma^{\epsilon,j}}\epsilon, \end{aligned}$$

where $G_j^{\sigma^{\epsilon,j}}$ and $\sigma^{\epsilon,j}$ are defined in Eqn. (16) and Eqn. (11), respectively. The desired result follows from Lemma 4. ■

Remark 1: The error bound derived above depends on the quantity β. If (A, C_i) is detectable for some $i \in \mathbb{M}$, then it can be easily shown that β is bounded from above by $\beta_i N$, where β_i denotes the peak estimation error corresponding to the schedule $\sigma = \{i, \ldots, i\}$, i.e., $\beta_i = \mathcal{E}^\sigma$.

Let C be defined as $C = [C_1^{\mathrm{T}}, \ldots, C_M^{\mathrm{T}}]^{\mathrm{T}}$ and let Σ_v be defined as $\Sigma_v = \mathrm{diag}\left(\Sigma_{v_1}, \ldots, \Sigma_{v_M}\right)$ where diag places the elements along the diagonal. The covariance matrix $\hat{\Sigma}_k^{\mathrm{all}}$ is obtained by using all the sensors at each time-step and evolves according to,

$$\hat{\Sigma}_{k+1}^{\mathrm{all}} = A\hat{\Sigma}_k^{\mathrm{all}}A^{\mathrm{T}} - A\hat{\Sigma}_k^{\mathrm{all}}C^{\mathrm{T}}\left(C\hat{\Sigma}_k^{\mathrm{all}}C^{\mathrm{T}} + \Sigma_v\right)^{-1}C\hat{\Sigma}_k^{\mathrm{all}}A^{\mathrm{T}} + \Sigma_w$$

with $\hat{\Sigma}_0^{\mathrm{all}} = \Sigma_0$. It can be easily verified that $\mathrm{tr}(\hat{\Sigma}_k^{\mathrm{all}}) \leq \mathrm{tr}(\hat{\Sigma}_k^\sigma)\ \forall\sigma \in \mathbb{M}^N$ and $\forall k \in [0, \ldots, N]$.

Proposition 2: Using the previous insights, the quantity β can be upper bounded by, $\beta \leq J(\sigma^{\epsilon,N}) - \sum_{k=0}^{N-1}\mathrm{tr}\left(\hat{\Sigma}_k^{\mathrm{all}}\right)$.

Figure 1 shows the bound for $\log\left(G_j^\sigma\right)$ for randomly generated systems with $n = 4$ states, $M = 3$ sensors and a time-horizon of $N = 14$. In generating the random systems, each pair (A, C_i), $\forall i \in \mathbb{M}$, was restricted to be unobservable to coerce the optimal solution to switch between sensors, while if all the sensors are used at once then the system is fully observable. The function $\epsilon\left(\exp(G) + 1\right)N$ transforms any value G, from the plot, to the upper bound for the performance of the suboptimal algorithm. The actual maximum realized error percentage was 0.5% and consequently the bound is very conservative in predicting the performance of the suboptimal algorithm.

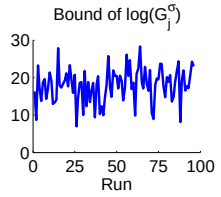


Fig. 1. A plot of the bound of the logarithm of G_j^σ.

B. Error Bound Through Linear Covariance Factors

It has been shown that the estimator error covariance can be factored such that it can be propagated linearly [12]. Similarly, if A is invertible, then the predictor covariance can also be linearly factored via, $\hat{\Sigma}_k = E_k F_k^{-1}$ where $E_0 = \Sigma_0$ and $F_0 = I$. This factorization allows the predictor covariance factors to evolve via the following linear system,

$$\Psi_k = \begin{bmatrix} E_k \\ F_k \end{bmatrix} = \xi(i)\begin{bmatrix} E_{k-1} \\ F_{k-1} \end{bmatrix}$$

for each $i \in \mathbb{M}$ and where $\xi(i)$ is defined as,

$$\xi(i) = \begin{bmatrix} \Sigma_w A^{\mathrm{T}} C_i^{\mathrm{T}}\Sigma_{v_i}^{-1}C_i + A & \Sigma_w A^{-\mathrm{T}} \\ A^{-\mathrm{T}}C_i^{\mathrm{T}}\Sigma_{v_i}^{-1}C_i & A^{-\mathrm{T}} \end{bmatrix}.$$

1) Perturbations: Let Ψ_k^p be the perturbed covariance factors at time-step k, and let $\bar{\xi}_k$ be the transfer function for the covariance factors at time-step k starting from the initial time-step. The transfer function $\bar{\xi}_k^\sigma$ is defined as, $\bar{\xi}_k^\sigma = \xi(\sigma_{k-1})\ldots\xi(\sigma_1)\xi(\sigma_0)$. Assume that the covariance matrix is perturbed at a single time-step, i.e. $\Sigma + \epsilon I$, then the perturbed linear covariance factors can be calculated by

$$\Psi_k^p = \bar{\xi}_k^\sigma\begin{bmatrix} \Sigma + \epsilon I \\ I \end{bmatrix} = \bar{\xi}_k^\sigma\left(\begin{bmatrix} \Sigma \\ I \end{bmatrix} + \begin{bmatrix} \epsilon I \\ 0 \end{bmatrix}\right) \tag{18}$$

which can be decomposed into, $\Psi_k^p = \Psi_k^{np} + \Psi_k^\epsilon$ where

$$\Psi_k^{np} = \bar{\xi}_k^\sigma\begin{bmatrix} \Sigma \\ I \end{bmatrix} = \bar{\xi}_k^\sigma\Psi_0^{np} \quad \text{and} \quad \Psi_k^\epsilon = \bar{\xi}_k^\sigma\begin{bmatrix} \epsilon I \\ 0 \end{bmatrix} = \bar{\xi}_k^\sigma\Psi_0^\epsilon.$$

The terms Ψ_k^{np} and Ψ_k^ϵ are the unperturbed and perturbation covariance factors at time-step k, respectively. Also, the terms $\Sigma_k^{np} = E_k^{np}F_k^{np-1}$ and $\Sigma_k^\epsilon = E_k^\epsilon F_k^{\epsilon-1}$ are the unperturbed and perturbation covariance matrices, respectively, if the inverses exist. Consequently, the original and perturbation factors can be propagated independently to determine the effect. The perturbed covariance matrix at time-step k is calculated through,

$$\hat{\Sigma}_k^p = E_k^{np}\left(F_k^{np} + F_k^\epsilon\right)^{-1} + E_k^\epsilon\left(F_k^{np} + F_k^\epsilon\right)^{-1}.$$

Using the matrix inversion lemma for the term $\left(F_k^{np} + F_k^\epsilon\right)^{-1}$ yields,

$$\hat{\Sigma}_k^p = \hat{\Sigma}_k^{np} + \Delta_k^{\epsilon,\sigma}, \tag{19}$$

where

$$\begin{aligned} \Delta_k^{\epsilon,\sigma} = & \hat{\Sigma}_k^\epsilon - E_k^\epsilon F_k^{\epsilon-1}\left(F_k^{np-1} + F_k^{\epsilon-1}\right)^{-1}F_k^{\epsilon-1} \\ & - E_k^{np}F_k^{np-1}\left(F_k^{np-1} + F_k^{\epsilon-1}\right)^{-1}F_k^{np-1}. \end{aligned} \tag{20}$$

The only constraint on the linear covariance factors is that $E_0^{np}F_0^{np-1} = \Sigma_0$. Consequently, there is some freedom in choosing the initial covariance factors, E_0^{np} and F_0^{np}, which can be used to our advantage to simplify the $\Delta_k^{\epsilon,\sigma}$ term. If the initial unperturbed covariance factors can be chosen such that $E_k^{np} = \hat{\Sigma}_k^{np}$ and $F_k^{np} = I$ then it will simplify the expression for $\Delta_k^{\epsilon,\sigma}$. The initial covariance factors can be calculated

by solving the following linear system, $\bar{\xi}_k^{\sigma} \begin{bmatrix} E_0^{np} \\ F_0^{np} \end{bmatrix} = \begin{bmatrix} \hat{\Sigma}_k^{np} \\ I \end{bmatrix}$ with the constraint that $E_0^{np} F_0^{np-1} = \Sigma_0$. There is also a dependence on the perturbed/perturbation covariance factors and the unperturbed covariance factors. Therefore the initial perturbation covariance factors must be determined from the unperturbed initial covariance factors but again there is some freedom in choosing them as well. One such possibility is: $E_0^{\epsilon} = \epsilon F_0^{np}$, $F_0^{\epsilon} = 0$ and another is: $E_0^{\epsilon} = \left(\Sigma + \epsilon I - E_0^{np}\left(F_0^{np} + I\right)^{-1}\right)\left(F_0^{np} + I\right)$, $F_0^{\epsilon} = I$. Substituting the previous simplifications into Eqn. (20) yields,

$$\Delta_k^{\epsilon,\sigma} = \left(\hat{\Sigma}_k^{\epsilon} - \hat{\Sigma}_k^{np}\right)\left(I + F_k^{\epsilon\,-1}\right)^{-1}. \tag{21}$$

Even though the perturbation term cannot be readily computed because the term $\hat{\Sigma}_k^{np}$ is not known, trends can still be inferred about the overall effect of the suboptimal algorithm.

2) Error Bound: Using the results from the previous section, an upper bound on the performance of the numerical redundancy algorithm can be established.

Lemma 5: Let $\sigma^{\epsilon,j}$ be decomposed into a j-horizon schedule $\sigma^{j'}$ and an $(N-j)$-horizon schedule $\sigma^{j''}$, i.e. $\sigma^{\epsilon,j} = \{\sigma^{j'}, \sigma^{j''}\}$ For any $N \in \mathbb{Z}_+$, the error is bounded by,

$$V_N^{\epsilon} - V_N \leq \sum_{j=1}^{N}\left(\epsilon + \sum_{k=1}^{N-j} \mathrm{tr}\left(\Delta_k^{\epsilon,\sigma^{j''}}\right)\right).$$

Proof: From Theorem 3,

$$V_N^{\epsilon,j} \leq V_N^{\epsilon,j-1} + \epsilon + \sum_{k=1}^{N-j} \mathrm{tr}\left(\Theta_k^{\sigma^{j''}}\left(\epsilon, \phi_{k+j}^{\sigma^{\epsilon,j}}\left(\Sigma_0\right)\right)\right).$$

From Eqn. (19), the error function $\Theta_k^{\sigma^{j''}}\left(\epsilon, \phi_{k+j}^{\sigma^{\epsilon,j}}\left(\Sigma_0\right)\right)$ is equal to $\Delta_k^{\epsilon,\sigma^{j''}}$. Therefore,

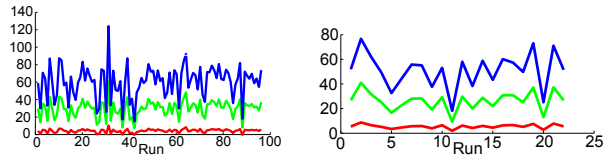
$$V_N^{\epsilon,j} \leq V_N^{\epsilon,j-1} + \epsilon + \sum_{k=1}^{N-j} \mathrm{tr}\left(\Delta_k^{\epsilon,\sigma^{j''}}\right).$$

■

One thing to note about the bound is that $\Delta_k^{\epsilon,\sigma}$ is proportional to $\Sigma_k^{\epsilon} - \Sigma_k^{np}$ which means that under most circumstances a small perturbation only has an affect during the first few time-steps. This is due to the fact that $\Sigma_k^{\epsilon} \to \Sigma_k^{np}$ because Σ_k^{np} should converge to a (cyclic) steady-state covariance if the system can be fully observed through switching between sensors.

The bound cannot be directly computed using the results from the algorithm since it requires knowing $\sigma^{\epsilon,j}$ for all j; thus, the optimal sensor schedule needs to be computed since it is by definition $\sigma^{\epsilon,0}$. Fortunately, a good approximation can be obtained by using $\sigma^{\epsilon,j} = \sigma^{\epsilon,N}$ for all j which is the sensor schedule returned from the numerical redundancy pruning algorithm. Under this approximation the exact representation of $\Delta_k^{\epsilon,\sigma}$ in Eqn. (19) can be used to calculate the bound. Through this approximation, it is no longer guaranteed to be a valid upper bound but simulations were performed in which the optimal sensor schedule could be computed and the approximate bound was consistent in all cases.

Figure 2(a) shows a comparison of the approximate bound for the same set of randomly generated systems as shown before in Figure 1 for the peak estimation error bound except that it is plotted as the percentage of the total cost. The approximate bound is much less conservative in predicting the performance of the suboptimal solution. Figure 2(b) shows a comparison of the approximate bound for randomly generated systems with a time-horizon of $N = 50$. The approximate bound for the longer time-horizon exhibits the same trends as with the shorter time-horizon.



(a) (b)

Fig. 2. Two comparisons of the estimated bound for $\epsilon = [0.1, 0.5, 1.0]$ which are colored red, green and blue respectively. The estimated bound is plotted as the percentage of the total cost.

VI. Conclusions

One area of future work that the authors wish to explore is in analyzing periodic schedules. It has been previously noticed that the sensor schedules tend to be periodic for the non-transient portion of the schedule. The authors would like to analyze this behavior to determine conditions for the periodicity and a bound for the objective function if the periodic schedule were used.

References

[1] S. Arai, Y. Iwatani, and K. Hashimoto, "Optimal sensor scheduling of sensors in a sensor network for mobile robot navigation," in *2007 American Control Conference*, (New York City, NY), July 2007.

[2] S. Martinez and F. Bullo, "Optimal sensor placement and motion coordination for target tracking," vol. 42, no. 4, pp. 661–668, 2006.

[3] S. Gao, C. T. Vu, and Y. Li, "Sensor scheduling for k-coverage in wireless sensor networks," in *MSN*, vol. 4325 of *Lecture Notes in Computer Science*, pp. 268–280, Springer, 2006.

[4] Y. Li, L. W. Krakow, E. K. P. Chong, and K. N. Groom, "Approximate stochastic dynamic programming for sensor scheduling to track multiple targets," *Digital Signal Processing*, vol. 19, no. 6, pp. 978–989, 2009.

[5] L. Meier III, J. Peschon, and R. M. Dressler, "Optimal control of measurement subsystems," *IEEE Transactions on Automatic Control*, vol. 12, pp. 528–536, October 1967.

[6] P. Alriksson and A. Rantzer, "Sub-optimal sensor scheduling with error bounds," in *Proceedings of the 16th IFAC World Congress*, (Prague, Czech Republic), July 2005.

[7] V. Gupta, T. Chung, B. Hassibi, and R. M. Murray, "On a stochastic sensor selection algorithm with applications in sensor scheduling and sensor coverage," *Automatica*, 2004.

[8] M. Rezaeian, "Sensor scheduling for optimal observability using estimation entropy," in *Fifth Annual IEEE International Conference on Pervasive Computing and Communications Workshops*, 2007.

[9] Y. He and K. Chung, "Sensor Scheduling for Target Tracking in Sensor Networks," in *Proceedings of the 43rd IEEE Conference on Decision and Control*, (Atlantis, Paradise Island, Bahamas), December 2004.

[10] M. P. Vitus, W. Zhang, A. Abate, J. Hu, and C. J. Tomlin, "On efficient sensor scheduling for linear dynamical systems," in *Proceedings of the 2010 American Control Conference*, (Baltimore, Maryland), June 2010.

[11] M. P. Vitus, W. Zhang, A. Abate, J. Hu, and C. J. Tomlin, "Efficient sensor scheduling for linear dynamical systems," *to be submitted for journal publication*, 2010.

[12] S. Prentice and N. Roy, "The belief roadmap: Efficient planning in linear POMDPs by factoring the covariance," in *Proceedings of the 13th International Symposium of Robotics Research (ISRR)*, (Hiroshima, Japan), November 2007.