Sensor Selection Strategies for State Estimation in Energy Constrained Wireless Sensor Networks

Yilin Mo [a,1] Roberto Ambrosino [b] Bruno Sinopoli [a]

[a] *Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA, US*

[b] *Dipartimento per le Tecnologie, Università degli Studi di Napoli Parthenope, Centro Direzionale, Isola C4, 80143, Napoli, Italy*

Abstract

Wireless Sensor Networks (WSNs) enable a wealth of new applications where remote estimation is essential. Individual sensors simultaneously sense a dynamic process and transmit measured information over a shared channel to a central base station. The base station computes an estimate of the process state by means of a Kalman filter. In this paper we assume that, at each time step, only a subset of all sensors are selected to send their observations to the fusion center due to channel capacity constraints or limited energy budget. We propose a multi-step sensor selection strategy to schedule sensors to transmit for the next T steps of time with the goal of minimizing an objective function related to the Kalman filter error covariance matrix. This formulation, in a relaxed convex form, defines an unified framework to solve a large class of optimization problems over energy constrained WSNs. We offer some numerical examples to further illustrate the efficiency of the algorithm.

Key words: Estimation problem, convex optimization, wireless sensor networks.

1 Introduction

Design and analysis of systems based on wireless sensor networks (WSNs) involves cross disciplinary research which spans domains within computer science, communications and control theory. A WSN is composed of low power devices that integrate computing with heterogeneous sensing and wireless communication. WSN-based systems are usually embedded in the physical world, with which they interact by collecting, processing and transmitting relevant data.

Sensor networks span a wide range of applications, including environmental monitoring and control, health care, home and office automation and traffic control [6]. In these applications, algorithms like Kalman filters can be used to perform state estimation based on lumped-parameter models of the physical phenomena. However, WSN operating constraints such as power and bandwidth often make it difficult to collect data from every sensor at the sampling rate required for effective monitoring. These considerations have led to the development of sensor scheduling strategies to select, at each time step, the subset of reporting sensors that minimize a certain cost function, usually related to the estimation error. The state estimation problem of a linear time-invariant system over a digital communication channel with a finite bandwidth capacity was introduced by Wong and Brockett [18], [19]. The work by Sinopoli et al. [16] shows the existence of a critical value of the packet loss rate for bounded estimation error covariance. For multi-sensor scenario, Liu and Goldsmith [10] extend the analytical work [16] to a two-sensor case. The trade-off between communication constraints and estimation performance is explored in the works of Ambrosino et al. [1], [2].

Sensor network energy consumption minimization and consequently lifetime maximization problems have been active areas of research over the past few years, as researchers soon realized that energy limitation constitutes one of the major obstacles to the adoption of such technology. Sensor network energy minimization is typically done via efficient MAC protocol design [11], or via efficient scheduling of the sensor states [3, 9, 17]. In [7], Gupta et al. propose a stochastic sensor selection algorithm to minimize the expected steady state estimation

[1] This author is responsible for correspondence. Ph.: +1 412 2689432; fax: +1 412 2683890; e-mail: ymo@andrew.cmu.edu

[2] This research is supported in part by CyLab at Carnegie Mellon under grant DAAD19-02-1-0389 from the Army Research Office. Foundation. The views and conclusions contained here are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either express or implied, of ARO, CMU, or the U.S. Government or any of its agencies.

error covariance. Shi et. al [15] consider sensor energy minimization as a mean to maximize network lifetime while guaranteeing a desired quality of estimation accuracy. They [14] further propose a sensor tree scheduling algorithm which leads to longer network lifetime. In these papers, the authors consider the estimation performance in terms of the steady state value of the error covariance matrix.
Another important contribution comes from the work of Joshi and Boyd [8], where they formulate the general sensor selection problem and solve it by relaxing it to a convex programming problem. By doing that, they provide a very general framework that can handle various performance criteria and energy and topology constraints. However, the main drawback of their approach is that they assume the sensor measurements are uncorrelated. At a first glance, this assumption may seem reasonable for measurements made at the same time. However, when we consider the measurements at different times, they are actually correlated as the states they measure are coupled through the system dynamics. As a result, the independent assumption fails and the approach proposed in [8] is not suited to address sensor selection problems beyond the single step.
In this paper we address the sensor selection problem over a WSN. The proposed approach breaks away from the existing literature in several ways. The novelty of the approach allows 1- formulation of the problem in a way that can address several different performance criteria over a large class of network constraints; 2- solution of the sensor selection problem over arbitrary time steps in an efficient manner; 3- use of correlated measurements. Furthermore, as the complexity increases proportionally with the length of the horizon, we propose a "divide and conquer" approach consisting of dividing a long time horizon into smaller ones and solving the sensor selection problem sequentially over them, making feasible finding solution over otherwise intractable large horizons. While this solution is clearly suboptimal, we are able to bound its performance analytically showing sublinear degradation of performance with respect to the fragmentation of the horizon. This last result is particularly attractive for the designer as it allows to explore the tradeoff between optimality and computation complexity.

We accomplish such results firstly by providing a unified framework capable to address, with the same methodology, a large class of sensor selection problems; secondly by a clever reformulation of the multistep Kalman filter equations that directly relates the estimate at time k to the past measurements, exposing the dependence of the filter on the optimization variables, i.e. the subset of sensors we select at each step. This allows to cast the problem in the standard quadratic optimization framework, after relaxation of the L_0 constraint to its L_1 equivalent.

Simulation results are presented to illustrate the effectiveness of the approach. We provide a real world example of temperature monitoring over a planar field and a comparison with existing results over a single step. Finally we validate the analytical results previously deducted for the proposed "divide and conquer" approach.

The rest of the paper is organized as follows: in Section 2 we introduce a general sensor selection problem (P_0) for the optimization of the state estimation performance involving different energy and topology constraints and performance criteria. In Section 3, we first convert problem (P_0) into an equivalent one (P_0'), which explicitly relates both the objective function and the constraints to the optimization variables. We later relax the formulation (P_0') into a quadratic programming problem (P_1) using the reweighted L_1 approximation [5]. Finally, we provide a bound on the improvement of the estimation performance we can obtain by increasing the window size T in problem (P_0). Section 4 provides several simulation results and illustrates the efficiency of our algorithm comparing it both with the optimal solution when computable, and other sensor selection techniques proposed in the literature. In Section 5, we provide conclusions. Section 6 reports some proofs of the main theorems.

2 Problem Formulation

In this section we derive a general framework that can capture a large number of sensor selection problems. Consider the linear system

$$\begin{aligned} x_{k+1} &= Ax_k + w_k, \\ y_k &= Cx_k + v_k, \end{aligned} \tag{1}$$

where w_k, v_k, x_1 are independent Gaussian random variables, $x_1 \sim \mathcal{N}(\bar{x}_1, \Sigma)$, $w_k \sim \mathcal{N}(0, Q)$ and $v_k \sim \mathcal{N}(0, R)$ [3]. We assume $x_k \in \mathbb{R}^n$ and $y_k = [y_{k,1}, y_{k,2}, \ldots, y_{k,m}]^T \in \mathbb{R}^m$ is the vector of the sensors' measurements where $y_{k,i}$ is the measurement of sensor i at time k. Let us indicate with $\gamma_j, j = 1, \ldots, mT$, the binary variables such that $\gamma_{m(k-1)+i} = 1$ if sensor i transmits at step time k, for $i = 1 \ldots m$ and $k = 1 \ldots T$, and it is 0 otherwise. Thus, at time k, the estimator receives readings $[\gamma_{m(k-1)+1} y_{k,1}, \ldots, \gamma_{m(k-1)+m} y_{k,m}]^T$. The aim of this paper is to compute a sensor selection sequence $\gamma_1, \ldots, \gamma_{mT}$ from time 1 to time T, which provides the optimal estimation performance while satisfying specific energy and topology constraints.
In order to define a multi-step Kalman filter with sensor

[3] Note that we start at time 1 instead of time 0.

selection, let us introduce the following quantities:

$$Y_k \triangleq [y_1^T, y_2^T, \ldots, y_k^T]^T \quad \in\mathbb{R}^{mk}, \tag{2}$$
$$\Gamma_k \triangleq diag(\gamma_1, \ldots, \gamma_{mk}) \quad \in\mathbb{R}^{mk\times mk}, \tag{3}$$
$$\vec{\gamma} \triangleq [\gamma_1, \ldots, \gamma_{mT}]^T \quad \in\mathbb{R}^{mT}, \tag{4}$$
$$\hat{x}^*_{k|k} \triangleq E\left(x_k|Y_k\right) \quad \in\mathbb{R}^{n}, \tag{5}$$
$$P^*_{k|k} \triangleq Cov\left(x_k|Y_k\right) \quad \in\mathbb{R}^{n\times n}, \tag{6}$$
$$\hat{x}_{k|k} \triangleq E\left(x_k|\Gamma_k Y_k\right) \quad \in\mathbb{R}^{n}, \tag{7}$$
$$P_{k|k} \triangleq Cov\left(x_k|\Gamma_k Y_k\right) \quad \in\mathbb{R}^{n\times n}. \tag{8}$$

The multi-step sensor selection problem can be formulated as the following optimization problem

$$\begin{aligned}(P_0): \min_{\vec{\gamma}} \quad & \sum_{k=1}^{T} trace(\mathcal{Q}_k P_{k|k} \mathcal{Q}_k^T), \\ s.t. \quad & H\vec{\gamma} \leq b, \\ & \gamma_i = 0 \text{ or } 1,\ i = 1, \ldots, mT,\end{aligned} \tag{9}$$

where $\mathcal{Q}_k \in \mathbb{R}^{n'\times n}$ is of full row rank, $H \in R^{h\times mT}$ and $b \in R^h$.

Next we will show how this formulation can address several classes of sensor selection problems. In this framework, the matrices $\mathcal{Q}_k$ can be used to define the minimization problem we want to tackle, while the matrix H and the vector b define the constraints we impose on the wireless sensors. In the following paragraphs we outline some of the problems that can be tackled in the proposed framework, by appropriate choice of the matrices $\mathcal{Q}_k s, H$ and the vector b. Let us start with the objective function.

- *Minimization of the final estimation error*
  Assume we want to minimize the estimation error at time T. Thus, we can choose

$$\mathcal{Q}_1, \ldots, \mathcal{Q}_{T-1} = 0,\ \mathcal{Q}_T = I.$$

- *Minimization of the average estimation error*
  Assume we want to minimize the average estimation error from time 1 to time T. Thus, we can choose

$$\mathcal{Q}_k = I, k = 1, \ldots, T.$$

- *Minimization of the estimation error of a single state*
  Assume at time k we are interested in the ith state $x_{k,i}$ of the system. Thus, we can choose

$$\mathcal{Q}_k = [\delta_{1,i}, \ldots, \delta_{n,i}] \in \mathbb{R}^{1\times n},$$

  where $\delta_{i,j} = 1$ if $i = j$, it is 0 otherwise. Notice that $n' = 1 \neq n$ in general, which explains the reason why we do not require $\mathcal{Q}_k$ to be a square matrix.

- *Minimization of the cost function of a finite horizon Linear Quadratic Gaussian (LQG) regulation problem*
  Assume we have a LQG regulation problem where the system dynamic is given by

$$x_{k+1} = Ax_k + Bu_k + w_k,\ y_k = Cx_k + v_k,$$

  and the cost function is

$$J_T = E\left[x_T^T W_T x_T + \sum_{k=1}^{T-1}(x_k^T W_k x_k + u_k^T U_k u_k)\right],$$

  where W_k, U_k are positive semidefinite matrices. Given a sensor selection schedule, it is well known that the optimal controller and estimator, which minimize the LQG cost function, can be designed separately. The optimal control law is

$$u_k = -(B^T S_{k+1} B + U_k)^{-1} B^T S_{k+1} A\hat{x}_{k|k},$$

  where S_k satisfies the backward recursive equations

$$\begin{aligned}S_k = {} & A^T S_{k+1} A + W_k \\ & - A^T S_{k+1} B(B^T S_{k+1} B + U_k)^{-1} B^T S_{k+1} A,\end{aligned}$$

  with $S_T = W_T$. The optimal estimator is still the Kalman filter. The optimal value of the cost function is

$$\begin{aligned}J_T^* = {} & trace(S_1\Sigma) + \sum_{k=1}^{T-1} trace(S_{k+1}Q) \\ & + \sum_{k=1}^{T-1} trace[(A^T S_{k+1} A + W_k - S_k)P_{k|k}].\end{aligned} \tag{10}$$

  Thus, in order to find the best sensor selection schedule which minimizes J_T^*, we can let

$$\mathcal{Q}_k^T \mathcal{Q}_k = A^T S_{k+1} A + W_k - S_k, \mathcal{Q}_T = 0.$$

Next we will show how several network constraints can be formulated within the proposed framework.

- *Fixed number of sensors to be used at each time step*
  Assume at each time step we want to select no more than $p < m$ sensors. Thus, the constraints can be written as

$$\sum_{i=1}^{m} \gamma_{m(k-1)+i} \leq p,\ k = 1, \ldots, T\,. \tag{11}$$

- *Sensor energy constraints*
  Assume each sensor has initial energy $\mathcal{E}_1, \ldots, \mathcal{E}_m$ and each observation consumes $e_1, \ldots, e_m$ energy

units respectively. As a result, assuming for each sensor i the initial energy $\mathcal{E}_i$ is a multiple of e_i, it can do at most $\mathcal{E}_i/e_i$ observations until its battery runs out. Since we cannot use the sensor after the battery dies, the energy constraints can be written as

$$\sum_{k=1}^{T} \gamma_{m(k-1)+i} \leq \mathcal{E}_i/e_i,\ i = 1, \ldots, m\,. \qquad (12)$$

- *Multi-hop sensor networks*
  Consider that the sensor network has a tree structure and the estimator is the root node. The network uses a data aggregation protocol, where each node forwards the observation packets to its parent until the packet reaches the root. Moreover, the nodes on the route of the packet will add their own readings into the packet. As a result, a node needs to be selected to send observations if a child node is selected. Define the set of child nodes of sensor i to be $\mathcal{C}_i$. We can write the topology constraints of the network as

$$\gamma_{m(k-1)+i} \geq \gamma_{m(k-1)+j},\ k = 1, \ldots, T,\ j \in \mathcal{C}_i\,. \quad (13)$$

These examples clearly demonstrate the generality of the formulation and its applicability to a large class of problems.

3 Main Result

This section is devoted to manipulating and relaxing the optimization problem (P_0) to make it explicit and convex with respect to the optimization variables.

3.1 Reformulation

In this subsection, we will derive the relation between $P_{k|k}$ and $\vec{\gamma}$ explicitly. First let us consider the estimation problem for a linear system without sensor selection, i.e. all the readings from sensors are used. Kalman filter provides the optimal solution which takes the following form:

$$\begin{aligned}
\hat{x}^*_{1|0} &= \bar{x}_1, \quad P^*_{1|0} = \Sigma \qquad (14)\\
\hat{x}^*_{k+1|k} &= A\hat{x}^*_{k|k}, \quad P^*_{k+1|k} = AP^*_{k|k}A^T + Q\\
K^*_k &= P^*_{k|k-1}C^T(CP^*_{k|k-1}C^T + R)^{-1}\\
\hat{x}^*_{k|k} &= \hat{x}^*_{k|k-1} + K^*_k(y_k - C\hat{x}^*_{k|k-1})\\
P^*_{k|k} &= P^*_{k|k-1} - K^*_kCP^*_{k|k-1}.
\end{aligned}$$

This formulation of the Kalman filter explicitly shows the dependence of the estimated state $\hat{x}^*_{k|k}$ from y_k. However the dependence from the "old" measurements is recursive. In the following Lemma we express the estimated state $\hat{x}^*_{k|k}$ given by the Kalman filter in terms of all the observations, represented by Y_k.

Lemma 1 *Consider a linear system as in* (1)*. The state estimation $\hat{x}^*_{k|k}$ given by the Kalman filter is*

$$\hat{x}^*_{k|k} = G^*_k \begin{bmatrix} y_1 \\ \vdots \\ y_{k-1} \\ y_k \end{bmatrix} + H^*_k\bar{x}_1 = G^*_kY_k + H^*_k\bar{x}_1\,, \qquad (15)$$

where $G^*_1 = K^*_1$, $H^*_1 = I - K^*_1C$ and G^*_k, H^*_k can be evaluated recursively as

$$\begin{aligned}
G^*_{k+1} &= [(A - K^*_{k+1}CA)G^*_k,\ K^*_{k+1}]\,,\\
H^*_{k+1} &= (A - K^*_{k+1}CA)H^*_k\,,
\end{aligned} \qquad (16)$$

where K^*_k is the optimal Kalman gain given by (14)*.*

PROOF. We will prove the lemma by induction. When $k = 1$, by (14), we know that

$$\hat{x}_{1|1} = \bar{x}_1 + K^*_1(y_1 - C\bar{x}_1) = K^*_1Y_1 + (I - K^*_1C)\bar{x}_1. \quad (17)$$

Hence $G^*_1 = K^*_1$ and $H^*_1 = I - K^*_1C$. Now suppose that (16) holds for $k = 1, \ldots, N$. When $k = N + 1$, we have

$$\begin{aligned}
\hat{x}^*_{N+1|N} &= A\hat{x}^*_{N|N} = AG^*_NY_N + AH^*_N\bar{x}_1\\
\hat{x}^*_{N+1|N+1} &= \hat{x}^*_{N+1|N} + K^*_{N+1}(y_{N+1} - C\hat{x}^*_{N+1|N})\\
&= K^*_{N+1}y_{N+1} + (A - K^*_{N+1}CA)G^*_NY_k\\
&\quad + (A - K^*_{N+1}CA)H^*_N\bar{x}_1\,.
\end{aligned} \qquad (18)$$

Hence, (16) holds for $k = N + 1$ and, by induction, we conclude the proof.

Now, let us consider the sensor selection problem. In the case of sensor selection, the estimated state $\hat{x}_{k|k} = E(x_k|\Gamma_kY_k)$ can still be written as

$$\begin{aligned}
\hat{x}_{k|k} &= G_k\Gamma_kY_k + H_k\bar{x}_1,\\
P_{k|k} &= Cov(\hat{x}_{k|k} - x_k) = Cov(G_k\Gamma_kY_k + H_k\bar{x}_1 - x_k)\\
&= Cov(G_k\Gamma_kY_k - x_k), \qquad (19)
\end{aligned}$$

where G_k, H_k are the optimal gains under the sensor selection scheme Γ_k. The last equality is true since $H_k\bar{x}_1$ is a deterministic vector. In general, G_k is different from G^*_k and depends on the sensor selection strategy Γ_k we choose. However we will not write G_k explicitly in terms of Γ_k, because that will lead to a Riccati equation, which is hard to analyze in general. In contrast, by exploiting the optimality of Kalman filter, we will optimize over the matrix G_k. In the next Theorem, we propose a new formulation for the minimization problem of $trace(P_{k|k})$ over both G_k and Γ_k.

Theorem 2 *Consider the linear system* (1) *and let* Γ_k *be the sensor selection matrix. The minimization problem over* Γ_k *of the* $trace(P_{k|k})$*, given by the Kalman filter, can be formulated as*

$$\min_{G_k,\Gamma_k} \|G_k\Gamma_k\mathcal{S}_k - G_k^*\mathcal{S}_k\|_F^2\,, \tag{20}$$

where G_k^* *is given in* (16)*,* $\|\cdot\|_F$ *represents the Frobenius norm and*

$$\mathcal{S}_k \in \mathbb{R}^{mk\times mk} \;:\; Cov(Y_k) = \mathcal{S}_k\mathcal{S}_k\,. \tag{21}$$

PROOF. In the case of sensor selection, the covariance matrix can be written as

$$\begin{aligned} P_{k|k} &= Cov(\hat{x}_{k|k} - x_k) = Cov(G_k\Gamma_kY_k - x_k)\\ &= Cov[(G_k\Gamma_k - G_k^*)Y_k + G_k^*Y_k + H_k^*\bar{x}_1 - x_k]\\ &= Cov[(G_k\Gamma_k - G_k^*)Y_k] + Cov(G_k^*Y_k + H_k^*\bar{x}_1 - x_k)\\ &= (G_k\Gamma_k - G_k^*)Cov(Y_k)(G_k\Gamma_k - G_k^*)^T + P_{k|k}^*\,. \end{aligned} \tag{22}$$

The second equality holds because $H_k^*\bar{x}_1$ is deterministic and furthermore it does not change the covariance. The decomposition of the covariance matrix in the third line of equation (22) is possible since, by the optimality of Kalman filter, $x_k - \hat{x}_{k|k}^* = x_k - G_k^*Y_k - H_k^*\bar{x}_1$ is orthogonal to Y_k. Therefore the cost function can be written as

$$\begin{aligned} &\min_{\Gamma_k} trace(P_{k|k}) = \min_{\Gamma_k} trace(P_{k|k}^*) +\\ &+ \min_{\Gamma_k,G_k} trace\left[(G_k\Gamma_k - G_k^*)Cov(Y_k)(G_k\Gamma_k - G_k^*)^T\right]. \end{aligned} \tag{23}$$

Since the first term does not depend on Γ_k, we will focus only on the second term. Let us decompose $Cov(Y_k)$ as $\mathcal{S}_k\mathcal{S}_k$, where $\mathcal{S}_k$ is symmetric. Since $Cov(Y_k)$ is positive semidefinite, we can always find $\mathcal{S}_k$. Thus

$$\begin{aligned} &(G_k\Gamma_k - G_k^*)Cov(Y_k)(G_k\Gamma_k - G_k^*)^T\\ &= (G_k\Gamma_k\mathcal{S}_k - G_k^*\mathcal{S}_k)(G_k\Gamma_k\mathcal{S}_k - G_k^*\mathcal{S}_k)^T. \end{aligned} \tag{24}$$

As a result, the minimization problem $\min_{\Gamma_k}\ trace(P_{k|k})$ is equivalent to

$$\begin{aligned} &\min_{\Gamma_k,G_k} trace\left[(G_k\Gamma_k\mathcal{S}_k - G_k^*\mathcal{S}_k)(G_k\Gamma_k\mathcal{S}_k - G_k^*\mathcal{S}_k)^T\right] =\\ &\min_{\Gamma_k,G_k} \|G_k\Gamma_k\mathcal{S}_k - G_k^*\mathcal{S}_k\|_F^2\,. \end{aligned}$$

From Theorem 2 we can derive the following corollary.

Corollary 3 *Consider the linear system* (1) *and let* Γ_k *be the sensor selection matrix. The minimization problem over* Γ_k *of the* $trace(\mathcal{Q}_kP_{k|k}\mathcal{Q}_k^T)$ *given by the Kalman filter can be formulated as*

$$\min_{\Gamma_k,G_k} \|\mathcal{Q}_kG_k\Gamma_k\mathcal{S}_k - \mathcal{Q}_kG_k^*\mathcal{S}_k\|_F^2\,, \tag{25}$$

where where G_k^* *is given in* (16) *and*

$$\mathcal{S}_k \in \mathbb{R}^{mk\times mk} \;:\; Cov(Y_k) = \mathcal{S}_k\mathcal{S}_k\,.$$

PROOF. It follows from the proof of Theorem 2.

Now let us define the matrix

$$\mathcal{G}_k \triangleq \mathcal{Q}_kG_k\Gamma_k \in \mathbb{R}^{n'\times mk}\,. \tag{26}$$

The following theorem relates γ_i and $\mathcal{G}_k$.

Theorem 4 *Denote the* i^{th} *column of the matrix* $\mathcal{G}_k$ *as* $\vec{\mathcal{G}}_{k,i}$*. The following inequality holds:*

$$\gamma_i \geq \left\| \sum_{k=k'}^{T} \left\|\vec{\mathcal{G}}_{k,i}\right\|_1 \right\|_0, \tag{27}$$

where $\|\cdot\|_0$ *(called the* L_0 *norm*[4] *) of a scalar is* 0 *if the scalar is* 0 *(it is* 1 *otherwise) and* k' *is the smallest number which satisfies* $mk' > i$*.*

PROOF. It follows directly from the definition of $\mathcal{G}_k$.

Using all the previous arguments, we can conclude that the optimization problem (P_0) is equivalent to

$$(P_0'): \min_{\mathcal{G}_k,\vec{\gamma}} \quad \sum_{k=1}^{T} \|\mathcal{G}_k\mathcal{S}_k - \mathcal{Q}_kG_k^*\mathcal{S}_k\|_F^2\,, \tag{28}$$

$$s.t. \quad H\vec{\gamma} \leq b$$

$$\gamma_i \geq \left\| \sum_{k=k'}^{T} \left\|\vec{\mathcal{G}}_{k,i}\right\|_1 \right\|_0 \quad i = 1\ldots mT. \tag{29}$$

Remark 5 *The main improvement of formulation* (P_0') *over* (P_0) *is that instead of* $trace(\mathcal{Q}_kP_{k|k}\mathcal{Q}_k)$*, we now have an explicit quadratic objective function as well an explicit relation between the said objective function and the* γ_i*s.*

In general, sensor selections are hard combinatorial problems as they involve binary constraints. For large

[4] L_0 is in fact not a norm.

systems the problem becomes computationally infeasible. Solutions in this case are based on heuristic methods. In this paper, the formulation (P_0') is also a combinatorial problem since the last constraint of (P_0') contains an L_0 norm. However, formulation (P_0') allows us to easily relax it to a convex problem, which will be discussed in the next subsection.

3.2 Reweighted L_1 approximation

This subsection describes a reweighted L_1 approximation of problem (P_0').
As discussed in the previous subsection, (P_0') is a combinatorial problem. In principle, to find the optimal solution, we should check all possible sensor selection schedules verifying the constraints. However, if the number of sensors is large, this evaluation becomes infeasible. Hence we have to consider solutions which are suboptimal, but computationally feasible. We will derive suboptimal solution based on a relaxed convex version of the original optimization problem [4].
One standard technique to solve the L_0 optimization problem (P_0') is to replace the L_0 norm with L_1 norm. However, the main drawback of this approach is that in L_1 larger numbers are penalized much more than in L_0, where all the non-zero numbers are treated equally. In [5], the authors provide a reweighted L_1 minimization to address this issue. According to this method, the L_0 norm is substituted with a weighted L_1 norm, where the weights are chosen to avoid the penalization of the bigger coefficients. The authors further propose an iterative algorithm that alternates between a minimization phase and another one involving the redefinition of the weights. The authors find that this reweighted L_1 minimization outperforms traditional L_1 minimization in many situations. In the following we use this algorithm to relax the optimization problem (P_0'). The algorithm is composed of 4 steps:

(1) Set the iteration count l to zero and set the weights vector to $w_i^0 = 1$ for $i = 1, ..., mT$.
(2) Solve the weighted L_1 minimization problem

$$(P_1): \min_{\mathcal{G}_k, \vec{\gamma}} \quad \sum_{k=1}^{T} \|\mathcal{G}_k \mathcal{S}_k - \mathcal{Q}_k G_k^* \mathcal{S}_k\|_F^2 , \qquad (30)$$
$$s.t. \quad H\vec{\gamma} \leq b$$
$$\gamma_i \geq w_i^l \sum_{k=k'}^{T} \left\| \vec{\mathcal{G}}_{k,i} \right\|_1 \quad i = 1 \dots mT.$$

Let the solution be $\gamma_1^l, \dots, \gamma_{mT}^l$.
(3) Update the weights

$$w_i^{l+1} = \frac{1}{\gamma_i^l + \epsilon} \quad i = 1, \dots, mT.$$

(4) Terminate if either l reaches a specified maximum number of iterations l_{max} or the solution has converged. Otherwise, increase l and return to step 2.

Remark 6 *As in [5], we introduce the parameter $\epsilon > 0$ in step 3 in order to avoid inversion of zero-valued components in $\vec{\gamma}$.*

Remark 7 *We do not make any assumption on the covariance matrices Q, R, Σ during the reformulation and relaxation steps. Hence, the relaxed problem (P_1) can handle correlated measurement noise.*

Remark 8 *Usually the problem takes less than 10 reweightings for the solution to converge*[5] *. For each weighted minimization, it is easy to see that the problem is a quadratic programming(QP) problem. In [13], the authors proposed an interior point algorithm for solving QP with $O(N^{2.5})$ arithmetic operations per iteration, where N is the total number of optimization variables. In our problem, each G_k is a $n \times mk$ matrix and $\vec{\gamma} \in \mathbb{R}^{mT}$. Hence, the total number of optimization variables is*

$$N = nm \sum_{k=1}^{T} k + mT = nmT(T+1)/2 + mT.$$

Since we do not need an accurate solution of QP, we can assume that the number of iterations needed for solving QP is fixed. Hence, the complexity of our algorithm is $O(n^{2.5} m^{2.5} T^5)$.

3.3 Multi-step Analysis

In the previous section we have shown how the multi-step sensor selection problem in case of correlated noise can be relaxed into a convex minimization problem. Although the relaxation technique significantly contributes to reducing the computational complexity, the problem is still intractable for large Ts. In fact, as the window size T increases, the number of optimization variables rapidly grows ($O(T^2)$) as shown in the previous subsection. A natural solution to this problem is to divide the scheduling period T into smaller intervals and solve the optimization problem sequentially for each of these smaller intervals. This is clearly an approximation which will yield suboptimal results. In this section, we will provide a complete analysis of the approximation and we will prove that the gap between the greedy approach and optimal is at most linear with respect to the variation of the window size.
Let us indicate with S_1 the optimal sensor selection strategy given by Problem (P_0) in $[1\,, NT]$ with $\mathcal{Q}_k = I_n$. Let $P_{k|k}(S_1)$ be the estimation error covariance at time k

[5] We do not need to have a very accurate solution because we will threshold $\vec{\gamma}$ and make it binary.

given by the schedule S_1. Consider now the greedy algorithm that first solves problem (P_0) from time 1 to time T and then solves it again for time $T+1$ to $2T$ and so on. Let us indicate with S_2 the sensor selection strategy in $[1\,, NT]$ given by the greedy method and with $P_{k|k}(S_2)$ the corresponding error covariance matrix [6] .

Remark 9 *To compare the optimal sensor selection strategy in* $[1\,, NT]$ *with the solution of the greedy algorithm we assume the constraints of Problem* (P_0) *on* $\gamma_i s$ *in each small window* $[KT+1,\ (K+1)T]$ *are decoupled from each other, which means that the constraints take the following form*

$$\begin{bmatrix} H_1 & & \\ & \ddots & \\ & & H_N \end{bmatrix} \vec{\gamma} \leq \begin{bmatrix} b_1 \\ \vdots \\ b_N \end{bmatrix}, \gamma_i = 0 \text{ or } 1,\ i = 1, \ldots, NmT,$$

where each H_i *has* mT *columns.*

In this following theorem we characterize the loss of performance of the sensor selection strategy S_2 compared to S_1 in terms of objective function of Problem (P_0) in $[1\,, NT]$.

Theorem 10 *The loss of performance of schedule* S_2 *with respect to* S_1 *satisfies the following inequality:*

$$\sum_{k=1}^{NT} trace(P_{k|k}(S_2)) - \sum_{k=1}^{NT} trace(P_{k|k}(S_1)) \leq (N-1)\frac{\rho\beta}{\alpha}$$

where

$$\rho = \inf\{r > 0 | P_{KT|KT}(S_2) \leq (r+1)P_{KT|KT}(S_1), \quad K = 1, \ldots, N-1\} \tag{31}$$

$$\alpha = \sup\{r > 0 | Q \geq r(AP_{k|k}(S_1)A^T) \quad , k = T+1, \ldots, NT\} \tag{32}$$

$$\beta = \sup\{trace(P_{k|k}(S_1)) | k = T+1, \ldots, NT\}. \tag{33}$$

PROOF. See Appendix.

Let S_* be the schedule that selects all the sensors and S_ϕ be the schedule that selects no sensors. For any schedule S, we know that

$$P_{k|k}(S_*) \leq P_{k|k}(S) \leq P_{k|k}(S_\phi)\,. \tag{34}$$

If the system is detectable, then $P_{k|k}(S_*)$ converges. Moreover if we assume that the system is stable also $P_{k|k}(S_\phi)$ converges. It means that $P_{k|k}(S_*)$ and $P_{k|k}(S_\phi)$ are uniformly bounded and hence there exists two positive semidefinite matrices $\overline{P}$ and $\underline{P}$, such that

$$\underline{P} \leq P_{k|k}(S_*) \leq P_{k|k}(S) \leq P_{k|k}(S_\phi) \leq \overline{P}.$$

Since the above inequality is independent of both schedule S and time index k, we can conclude that there exist uniform bounds for α, β, ρ which are independent of the real window size T. Moreover, the loss incurred when dividing the whole window in N parts will always be sublinear with respect to N.
In the meanwhile, solving (P_0) on $[0, NT]$ needs $O(n^{2.5}m^{2.5}(NT)^5)$ arithmetic operations while to compute schedule S_2, we only need $O(n^{2.5}m^{2.5}T^5)$ arithmetic operations for N smaller windows of size T. Hence, reducing the window size by N will reduce the computation complexity to N^{-4} of the original problem. We can conclude that when the system is large, the greedy algorithm is a reasonable approach to reduce computational complexity.

4 Simulation Result

In this section we will illustrate the efficiency of the sensor selection strategy with some numerical examples.
Throughout this section we will consider a system composed of m sensors from which only p of them can be selected at each time. The sensors measurements are used to estimate a dynamical system consisting of n states. We assume to minimize the average estimation error, hence we impose that $\mathcal{Q}_k = I$, $k = 1, \ldots, T$.

4.1 Examples: Temperature Monitoring

Let us consider a numerical example in which a sensor network is deployed to monitor the temperature in a planar closed region. We model the heat process in the region of side l as

$$\frac{\partial u}{\partial t} = \alpha \left(\frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} \right), \tag{35}$$

with boundary conditions

$$\left.\frac{\partial u}{\partial x_1}\right|_{t,0,x_2} = \left.\frac{\partial u}{\partial x_1}\right|_{t,l,x_2} = \left.\frac{\partial u}{\partial x_2}\right|_{t,x_1,0} = \left.\frac{\partial u}{\partial x_2}\right|_{t,x_1,l} = 0, \tag{36}$$

where x_1, x_2 indicate the coordinates of the region; $u(t, x_1, x_2)$ denotes the temperature at time t at location (x_1, x_2) and α indicates the speed of the diffusion process. In [12] the model has been discretized in space with a $N \times N$ grid and in time with frequency of 1 Hz yielding the following linear discrete time model

$$U_{k+1} = AU_k + w_k\,,$$

[6] For simplicity we just choose $\mathcal{Q}_k = I$

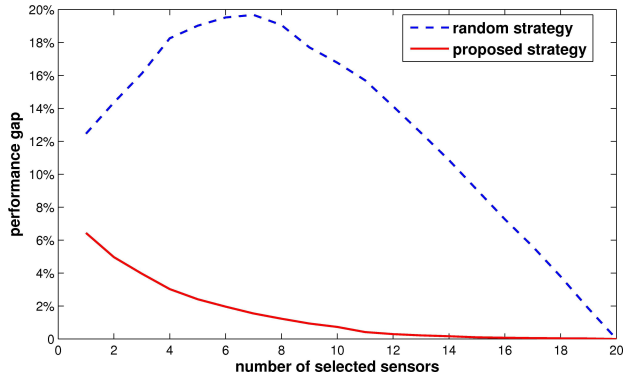

Fig. 1. Average relative estimation performance gap of the proposed sensor selection strategy (red solid line) and the random strategy (blue dashed line) with respect to the optimal strategy.

where the vector $U_k \in \mathbb{R}^{N^2}$ indicates the temperature at each point of the grid, $A \in \mathbb{R}^{N^2 \times N^2}$ takes into account the diffusion process and w_k is the process noise.

Let us assume that m sensors are randomly distributed in the region and each sensor measures a linear combination of the temperature of the grid around it (see [12]) so that the measurement equation becomes

$$Y_k = CU_k + v_k \,,$$

where $Y_k \in \mathbb{R}^m$ is the measurement vector, v_k is the measurement noise and $C \in \mathbb{R}^{m \times N^2}$ is defined in [12].

For the simulations, we impose $\alpha = 0.1\ m^2/s$, $l = 4\ m$ and $N = 3$ (hence the grid size is $h = 2\ m$). Moreover, we assume the WSN is composed by $m = 20$ nodes. In this section we consider the problem of minimizing the next step estimation error, i.e. $T = 1$, and we also assume uncorrelated noise with identity covariance matrix both for the process and the measurement noise. To show the quality of our sensor selection strategy, we compare it in terms of $trace(P_{k|k})$ with the real optimal sensor selection scheme. The optimal sensor selection strategy is evaluated with an exhaustive search of the solution set, which is composed by $\binom{m}{p}$ elements. It is straightforward to recognize that it can be computed just for small networks. We assume that $1 \leq p \leq 20$ and, for each value of p, we consider the estimation performance both of our strategy and a random strategy and we normalize them with respect to the optimal one. We average the simulation results for each p over 50 random placements of the sensors in the planar region. In Figure 1, we show the performance gap for $1 \leq p \leq 20$. The figure clearly illustrates that the proposed sensor selection strategy assures near optimal performance for all values of p with a gap that rapidly decreases to zero as p increases. On the other hand, the performance of the random method reaches a maximum gap of 20% compared to the optimal strategy and it is always sensibly lower than the performance of proposed strategy.

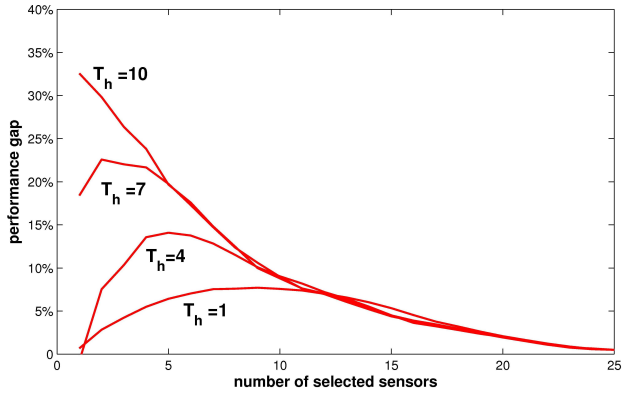

Fig. 2. Relative estimation performance gap of the sensor selection strategy in [8] with respect to the proposed strategy.

4.2 Comparison to state of the art

In this section we consider again the minimization problem of the next step estimation error with uncorrelated noise and we compare our algorithm to the one proposed in [8] for a WSN of $m = 50$ nodes[7]. To perform the comparison we consider random generated systems with $n = 20$ states where A, C are chosen randomly with each entry independently and identically distributed (iid) on $[0, 1]$. The matrix Q is computed by multiplying a randomly generated matrix with its transpose (to make them positive definite), where each entry is also iid on $[0, 1]$. We choose $\Sigma = I_n$. In Figure 2 we plot the performance gap between the proposed algorithm and the one in [8]. In particular, for each value of p, we evaluate the estimation performance of both approaches over a time horizon of T_h steps by applying the one step sensor selection method step by step. To achieve a statistic performance analysis, we average the simulation results over 50 random generated systems for each p. In Figure 2 we show the performance gap of the strategy in [8] with respect to our sensor selection strategy for $T_h = 1, 4, 7, 10$.
Figure 2 shows that our sensor selection strategy performs better than the one in [8] for all the values of p. Moreover the gap increases with the time horizon T_h and it is approximately 33% in case of $T_h = 10$ and $p = 1$. However, it is important to recognize that our sensor selection method is computationally more intensive than the one in [8]. In particular, the number of arithmetic operations for the algorithm in [8] is $O(m^3)$ while for our one step sensor selection, is $O(n^{2.5}m^{2.5})$.
Finally, we want to remark that the one step sensor selection strategy with uncorrelated measurements represents just a possible application of the proposed approach. We do not perform any comparison for either the multi-step case or the one with correlated noises as the optimal one is practically computationally infeasible

[7] In this comparison we do not implement the local optimization algorithm proposed in [8] to improve the sensor selection performance, since it is a general technique that can be applied to any sensor selection schemes, including ours.

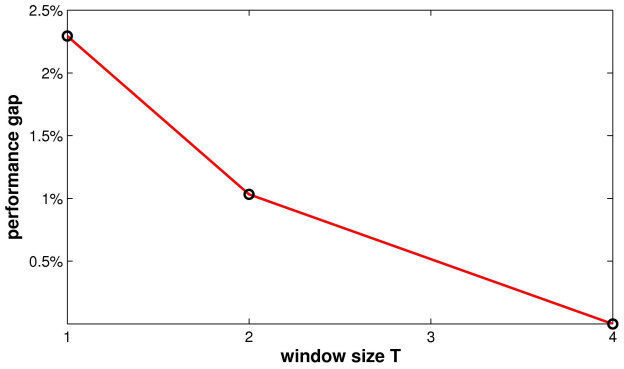

Fig. 3. Relative gap of $trace(P_{1|1} + P_{2|2} + P_{3|3} + P_{4|4})$ of solving the optimization problem in subintervals of $T = 1$, $T = 2$ or $T = 4$, normalized to the case of $T = 4$.

and we are unaware of any other method in the literature.

4.3 Multi-step sensor selection: a "divide and conquer" approach

In theorem 10 we bounded the performance loss due to dividing the horizon T into smaller intervals and performing sensor scheduling over each of them. We consider the problem of minimizing the estimation error over a window size of 4 steps, i.e. we assume to minimize $trace(P_{1|1} + P_{2|2} + P_{3|3} + P_{4|4})$. We wish to compare the solution of problem (P_0) for $T = 4$ with a greedy algorithm that solves (P_0) over subintervals $T' = 1$ and 2 until the time window $T = 4$ is reached. We assume a WSN of $m = 10$ nodes is used to monitor a random generated system of dimension $n = 10$ and we select $p = 2$ sensor each step. In figure 3 we indicate the performance gap, in terms of $trace(P_{1|1} + P_{2|2} + P_{3|3} + P_{4|4})$, solving the optimization problem in subintervals of $T = 1, 2, 4$, normalized to the case of $T = 4$. The simulation results show that the performance loss is quasi linear and not significant, confirming that the "divide and conquer" approach can be a suitable approximation to the multi-step approach for large time horizons.

5 Conclusions

In this paper we considered the problem of sensor scheduling for state estimation of a dynamical system over a sensor network. Scheduling sensor's readings is crucial to save valuable power and bandwidth in energy-constrained sensor networks. We provide a general optimization framework to address a large class of sensor scheduling problems, with several different cost functions and network constraints over a predetermined time horizon. We significantly advance the state of the art as existing approaches are limited to scheduling at most one step ahead and can not handle either correlated measurement or general network constraints.
We provide a solution to the optimization problem by reformulating it using a convex relaxation based on a reweighted L_1 approximation. Finally we analyze the loss of performance related to the choice of the horizon and we show that loss of performance is sublinear with the ratio between different horizons. Numerical examples are provided to illustrate the effectiveness of the proposed algorithm in providing quasi-optimal performance while reducing the high computational complexity associated with the sensor selection problem.

6 Appendix

In order to prove Theorem 10, we need some preliminary lemmas.
Let us define the following functions:

$$g(X, S) = X + S \quad , \quad h(X, S) = (X^{-1} + S)^{-1}$$

where $X, S \in \mathbb{R}^{n\times n}$.

Lemma 11 *Suppose that $X, Y \in \mathbb{R}^{n\times n}$ are positive semidefinite and $X \geq Y$, then* [8]

$$g(X, S) \geq g(Y, S) \quad , \quad h(X, S) \geq h(Y, S).$$

PROOF. The proof easily follows from the definition of positive semidefinite matrix.

Lemma 12 *Assume $X, S \in \mathbb{R}^{n\times n}$ are positive semidefinite and let $\alpha > 0$ be a scalar such that $S \geq \alpha X$. Then*

$$g((1 + \rho)X, S) \leq (1 + \frac{\rho}{1 + \alpha})g(X, S) \quad \forall \rho > 0\,. \quad (37)$$

PROOF.

$$g((1 + \rho)X, S) - (1 + \frac{\rho}{1 + \alpha})g(X, S) =$$
$$(1 + \rho)X + S - (1 + \frac{\rho}{1 + \alpha})(X + S) = \frac{\rho(\alpha X - S)}{1 + \alpha} \leq 0\,.$$

Lemma 13 *Assume $X, S \in \mathbb{R}^{n\times n}$ are positive semidefinite. Then*

$$h((1 + \rho)X, S) \leq (1 + \rho)h(X, S), \forall \rho > 0\,. \quad (38)$$

PROOF. Since X, S are both positive semidefinite, we know that there exists a matrix $U \in \mathbb{R}^{n\times n}$ that can diagonalize X^{-1} and S simultaneously, i.e.

$$X^{-1} = U^{-1}(U^{-1})^T,\ S = U^{-1}\Lambda(U^{-1})^T, \quad (39)$$

[8] All the comparisons between matrices in this section are in the positive semidefinite sense

where $\Lambda = diag(\lambda_1, \ldots, \lambda_n)$ and $\lambda_i \geq 0$. Hence

$$\begin{aligned} &h((1+\rho)X, S) - (1+\rho)h(X, S) \\ &= U^T \left\{ \left[(1+\rho)^{-1} I + \Lambda\right]^{-1} - (1+\rho)(I+\Lambda)^{-1} \right\} U \\ &= U^T diag(\ldots, \frac{1+\rho}{1+(1+\rho)\lambda_i} - \frac{1+\rho}{1+\lambda_i}, \ldots) U \,. \end{aligned}$$

It is easy to see that the diagonal matrix in the equation above is negative semidefinite. Thus,

$$h((1+\rho)X, S) \leq (1+\rho)h(X, S), \forall \rho > 0 \,.$$

We are now ready to prove Theorem 10.

PROOF. Due to space limits, we will only prove the theorem for $N = 2$ as the proof for arbitrary N follows the same reasoning. Let us define a third sensor scheduling S_3 as:

$$\begin{aligned} S_3[1\ ,\ T] &= S_2[1\ ,\ T] \\ S_3[T+1\ ,\ 2T] &= S_1[T+1\ ,\ 2T] \end{aligned}$$

We denote the estimation error covariance of schedule S_3 at time k to be $P_{k|k}(S_3)$.
By definition of the sensor scheduling S_2, we know that from time 1 to T it represents the optimal scheduling, which means that

$$\begin{aligned} \sum_{k=1}^{T} trace(P_{k|k}(S_3)) &= \sum_{k=1}^{T} trace(P_{k|k}(S_2)) \\ &\leq \sum_{k=1}^{T} trace(P_{k|k}(S_1)). \end{aligned} \quad (40)$$

Moreover, from time $T+1$ to time $2T$, the schedule S_2 is the optimal for initial condition $P_{T|T}(S_2)$. Since the schedule S_3 has the same initial condition $P_{T|T}(S_3) = P_{T|T}(S_2)$, we have

$$\sum_{k=T+1}^{2T} trace(P_{k|k}(S_3)) \geq \sum_{k=T+1}^{2T} trace(P_{k|k}(S_2)). \quad (41)$$

Combining (40) and (41), we get

$$\begin{aligned} &\sum_{k=1}^{2T} trace(P_{k|k}(S_2)) - \sum_{k=1}^{2T} trace(P_{k|k}(S_1)) \\ &\leq \sum_{k=T+1}^{2T} trace(P_{k|k}(S_3)) - \sum_{k=T+1}^{2T} trace(P_{k|k}(S_1)). \end{aligned}$$

Starting at time T, the prediction at time $T+1$ is

$$\begin{aligned} P_{T+1|T}(S_3) &= AP_{T|T}(S_3)A^T + Q = g(AP_{T|T}(S_3)A^T, Q) \\ P_{T+1|T}(S_1) &= AP_{T|T}(S_1)A^T + Q = g(AP_{T|T}(S_1)A^T, Q). \end{aligned}$$

Hence, by Lemmas 11 and 12,

$$\begin{aligned} P_{T+1|T}(S_3) &= g(AP_{T|T}(S_3)A^T, Q) \\ &\leq g((1+\rho)AP_{T|T}(S_1)A^T, Q) \\ &\leq (1 + \frac{\rho}{1+\alpha}) g(AP_{T|T}(S_1)A^T, Q) \\ &= (1 + \frac{\rho}{1+\alpha}) P_{T+1|T}(S_1), \end{aligned}$$

where ρ and α are defined in (31) and (32). At time $T+1$, S_1 and S_3 choose the same sensors. Suppose the corresponding measurement is

$$y_{T+1} = C_{T+1} x_{T+1} + v_{T+1}, \quad (42)$$

where $v_{T+1} \sim \mathcal{N}(0, R_{T+1})$. Using information filter, we can write the measurement update at time $T+1$ as

$$\begin{aligned} P_{T+1|T+1}(S_3) &= \left[(P_{T|T}(S_3))^{-1} + C_{T+1}^T R_{T+1}^{-1} C_{T+1}\right]^{-1} \\ &= h(P_{T+1|T}(S_3), C_{T+1}^T R_{T+1}^{-1} C_{T+1}) \\ P_{T+1|T+1}(S_1) &= h(P_{T+1|T}(S_1), C_{T+1}^T R_{T+1}^{-1} C_{T+1}) \end{aligned}$$

Hence, from Lemmas 11 and 13,

$$\begin{aligned} &P_{T+1|T+1}(S_3) = h(P_{T+1|T}(S_3), C_{T+1}^T R_{T+1}^{-1} C_{T+1}) \\ &\leq (1 + \frac{\rho}{1+\alpha}) h(P_{T+1|T}(S_1), C_{T+1}^T R_{T+1}^{-1} C_{T+1}) \\ &= (1 + \frac{\rho}{1+\alpha}) P_{T+1|T+1}(S_1) \,. \end{aligned}$$

By induction, we know that

$$P_{T+k|T+k}(S_3) \leq \left[1 + \frac{\rho}{(1+\alpha)^k}\right] P_{T+k|T+k}(S_1) \,, \quad k = 1, \ldots, T. \quad (43)$$

Hence,

$$\begin{aligned} &\sum_{k=T+1}^{2T} trace(P_{k|k}(S_3)) - \sum_{k=T+1}^{2T} trace(P_{k|k}(S_1)) \leq \\ &\sum_{k=1}^{T} \frac{\rho}{(1+\alpha)^k} trace(P_{T+k|T+k}(S_1)) \leq \sum_{k=1}^{T} \frac{\rho\beta}{(1+\alpha)^k} \leq \frac{\rho\beta}{\alpha}, \end{aligned}$$

which concludes the proof.

References

[1] R. Ambrosino, B. Sinopoli, and K. Poolla. Optimal sensor scheduling for remote estimation over wireless sensor

networks. In *Modelling, Estimation and Control of Networked Complex Systems*. Springer, 2009.

[2] R. Ambrosino, B. Sinopoli, K. Poolla, and S. S. Sastry. Optimal sensor density for remote estimation overwireless sensor networks. In *Forty-Sixth Annual Allerton Conference*, pages 599 – 606, Sept 2008.

[3] E. Bitar, E. Baeyens, and K. Poolla. Energy-aware sensor selection and scheduling in wireless sensor networks. In *1st IFAC Workshop on Estimation and Control of Networked Systems*, Sept 2009.

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

[5] E. J. Candes, M. B. Wakin, and S. Boyd. Enhancing sparsity by reweighted l1 minimization. *Journal of Fourier Analysis and Applications*, 14(5):877–905, December 2008.

[6] Nitaigour Premchand Mahalik (Ed.). *Sensor Networks and Configuration*. Springer, 2007.

[7] V. Gupta, T. Chung, B. Hassibi, and R. M. Murray. On a stochastic sensor selection algorithm with applications in sensor scheduling and dynamic sensor coverage. *Automatica*, 42(2):251–260, February 2006.

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

[9] Shan Liang, Yunjian Tang, and Qin Zhu. Passive wake-up scheme for wireless sensor networks. In *Second International Conference on Innovative Computing, Information and Control*, page 507, Sept 2007.

[10] X. Liu and A. Goldsmith. Kalman filtering with partial observation losses. In *IEEE Conference on Decision and Control*, pages 4180–4186, 2004.

[11] M. J. Miller and N. H. Vaidya. A mac protocol to reduce sensor network energy consumption using a wakeup radio. *IEEE Transactions on Mobile Computing*, 4(3):228 – 242, May-June 2005.

[12] Y. Mo, R. Ambrosino, and B. Sinopoli. Network energy minimization via sensor selection and topology control. In *1st IFAC Workshop on Estimation and Control of Networked Systems*, Sept 2009.

[13] Renato D. C. Monteiro and Ilan Adler. Interior path following primal-dual algorithms. part ii: Convex quadratic programming. *Mathematical Programming*, 44(1-3):43–66, May 1989.

[14] L. Shi, A. Capponi, K.H. Johansson, and R.M. Murray. Network lifetime maximization via sensor trees construction and scheduling. In *Third International Workshop on Feedback Control Implementation and Design in Computing Systems and Networks*, 2008.

[15] L. Shi, K.H. Johansson, and R. M. Murray. Change sensor topology when needed: How to efficiently use system resources in control and estimation over wireless networks. In *Proceedings of the 46th IEEE Conference on Decision and Control*, pages 5478 – 5485, 2007.

[16] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. Jordan, and S. Sastry. Kalman filtering with intermittent observations. *IEEE Trans. on Automatic Control*, 49(9):1453–1464, Sept 2004.

[17] N. A. Vasanthi and S. Annadurai. Sleep schedule for fast and efficient control of parameters in wireless sensor-actor networks. In *First International Conference on Communication System Software and Middleware*, pages 1–6, Jan 2006.

[18] W. S. Wong and R. W. Brockett. Systems with Finite Communication Bandwidth-Part I: State Estimation Problems. *IEEE Trans. Automatic Control*, 42(9):1294–1299, Sept 1997.

[19] W. S. Wong and R. W. Brockett. Systems with Finite Communication Bandwidth-Part II: Stabilization with Limited Information Feedback. *IEEE Trans. Automatic Control*, 44(5):1049–1053, May 1999.