Stochastic Sensor Scheduling for Energy Constrained Estimation in Multi-Hop Wireless Sensor Networks

Yilin Mo*, Emanuele Garone†, Alessandro Casavola♯, Bruno Sinopoli*

***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 fusion center. The fusion center computes an estimate of the process state by means of a Kalman filter. In this paper we assume that the WSN admits a tree topology with one fusion center at the root. At each time step only a subset of sensors can be selected to transmit observations to the fusion center due to a limited energy budget. We propose a stochastic sensor selection algorithm that randomly selects a subset of sensors according to a certain probability distribution, which is opportunely designed to minimize the asymptotic expected covariance matrix of the estimation error. We show that the optimal stochastic sensor selection problem can be relaxed into a convex optimization problem and thus efficiently solved. We also provide a possible implementation of our algorithm which does not introduce any communication overhead. The paper ends with some numerical examples that show the effectiveness of the proposed approach.**

***Index Terms*—Wireless Sensor Networks, Optimization, State Estimation.**

I. Introduction

Sensor networks span a wide range of applications, including environmental monitoring and control, health care, home and office automation and traffic control [1]. In these applications, estimation algorithms like Kalman filters can be used to undertake state estimation tasks based on lumped-parameter models of distributed physical phenomena. However, WSN operating constraints, such as power limitations, often make it difficult to collect data from every sensor at the desired sampling rates. These considerations have led to the development of sensor scheduling strategies able to select, at each time step, the subset of reporting sensors that minimizes a certain cost function, usually related to the expected estimation error.

Sensor network energy consumption minimization and, consequently, lifetime maximization problems have been active areas of research over the past few years, as researchers realized that energy limitations constitute one of the major obstacles to the extensive adoption of such a technology. Sensor networks energy minimization is typically accomplished via efficient MAC protocols [2] or via efficient scheduling of sensor states [3], [4]. In [5], Xue and Ganz show that the lifetime of sensor networks is influenced by transmission schemes, network density and transceiver parameters with different constraints on network mobility, position awareness and maximum transmission ranges. Chamam and Pierre [6] propose a sensor scheduling scheme capable of optimally putting sensors in active or inactive modes. Shi et. al [7] propose a sensor tree scheduling algorithm which leads to longer network lifetime.

Conversely, performance optimization of sensor networks under a given energy constraint, which can be seen as the dual problem of network energy minimization, has also been studied by several researchers. Such a constrained optimization problem has been studied for continuous-time linear systems in [8] and [9]. In [10], the author compute the optimal sensor scheduling for the estimation of a Hidden Markov Model. For discrete-time linear systems, methods like dynamic programming [11] or greedy algorithms [12] have been proposed to find the optimal sensor scheduling over long time horizons.

Another important contribution on this topic is the work of Joshi and Boyd [13], where a general single-step sensor selection problem is formulated and solved by means of convex relaxation techniques. Such a paper provides a very general framework that can handle various performance, energy criteria and topology constraints. Following this work, Mo et al. [14], [15] show that multi-step sensor selection problems can also be relaxed into convex optimization problems and thus efficiently solved.

A very different approach with respect to the above deterministic solutions was proposed by Gupta et. al. [16]. There, the authors propose a stochastic sensor selection algorithm in networks endowed with star topology. The algorithm is based on the idea that at each time step the sensors randomly choose whether to send measurements or not according to a certain probability distribution. Therefore, the probability distributions, which are chosen to minimize the expected steady-state error covariance matrix, become the optimization parameters. The authors argue that such a stochastic approach has several advantages over the conventional deterministic approaches. For example, it is easier to take into account random communication channel failures, which is a quite common issue in wireless sensor networks. The most relevant limitation of the results presented in that paper, beside the restriction to star topologies, hinges upon the assumption that only one sensor at a time can transmit its data in each sampling period, which is a strong assumption and requires a precise coordination between sensors.

In the present work, we go further on by proposing a stochastic sensor selection algorithm that not only overcomes the above limitations but also solves the routing problem under the assumption that the wireless sensor network has a tree topology. The proposed approach may be summarized as follows: the sensors are randomly selected according to a certain probability distribution that is designed so as to minimize the expected asymptotic estimation error covariance matrix while maintaining the connectivity of the network. In order to make the determination of the optimal probability distribution tractable, the problem is relaxed into a convex formulation. The advantages of the stochastic schedule over the deterministic one can be summarized as follows:

1) The search space of the stochastic formulation is continuous and convex, while the search space of deterministic formulation is discrete. Hence, the search of the optimal deterministic schedule is formulated as an integer programming problem, which makes the optimization problem potentially harder than its stochastic counterpart.
2) The expected performance of the stochastic formulation can be better than the deterministic one. Moreover, due to the ergodicity of the random Riccati equation, we can prove that under mild assumptions almost every sample path of the stochastic schedule is better than the deterministic one if the system runs long enough.
3) The stochastic schedule can be implemented with the same

*: Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA. Email: ymo@andrew.cmu.edu, brunos@ece.cmu.edu

†: Service d'Automatique et d'Analyse des Systèmes Universitè Libre de Bruxelles, 50 Av. F.D. Roosevelt, B-1050 Brussels, Belgium.egarone@ulb.ac.be

♯: Dipartimento di Elettronica, Informatica e Sistemistica, Università degli Studi della Calabria, Via Pietro Bucci, Cubo 42-c, Rende (CS), 87036, Italy. Email: egarone, casavola@deis.unical.it

This research was 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.

This paper was supported by the Belgian Network DYSCO (Dynamical Systems, Control and Optimization) funded by the Interuniversity Attraction Poles Program, initiated by the Belgian State, Science Policy Office.

computation and communication cost as the deterministic one.

The rest of the paper is organized as follows. In Section II we define the system and the communication model and introduce both the stochastic and deterministic sensor selection problems. We further present an ergodicity result on the performance of the stochastic sensor scheduling method which shows that the stochastic formulation showcases performance improvements with high probability. In Section III, we relax the stochastic sensor selection algorithm to make the problem solvable and propose an efficient implementation of the algorithm. Numerical examples are provided in Section IV. Section V finally concludes the paper.

II. Stochastic and Deterministic Sensor Selection

A. System Description

Consider the following discrete-time LTI system

$$x_{k+1} = Ax_k + w_k, \tag{1}$$

where $x_k \in \mathbb{R}^n$ represents the state and $w_k \in \mathbb{R}^n$ the disturbance. It is assumed that w_k and x_0 are independent Gaussian random vectors, $x_0 \sim \mathcal{N}(0,\ \Sigma)$ and $w_k \sim \mathcal{N}(0,\ Q)$, where $\Sigma,\ Q > 0$ are positive definite matrices. A wireless sensor network composed of m sensing devices $s_1, \ldots, s_m$ and one fusion center s_0 is used to monitor the state of system (1). The measurement equation is

$$y_k = Cx_k + v_k, \tag{2}$$

where $y_k = [y_{k,1}, y_{k,2}, \ldots, y_{k,m}]' \in \mathbb{R}^m$ is the measurement vector[1]. Each element $y_{k,i}$ represents the measurement of sensor i at time k, $C = [C_1', \ldots, C_m']'$ is the observation matrix and the matrix pair $(C,\ A)$ is assumed observable[2], $v_k \sim \mathcal{N}(0,\ R)$ is the measurement noise, assumed to be independent of x_0 and w_k. We also assume that the covariance matrix $R = diag(r_1, \ldots, r_m)$ is diagonal, which means that the measurement noise at each sensor is independent of all others and nonsingular, that is $r_i > 0, i = 1, ..., m$.

In order to model the communication amongst the nodes we introduce an oriented communication graph $G = \{V, E\}$ where the vertex set $V = \{s_0,\ s_1, \ldots, s_m\}$ contains all sensor nodes, including the fusion center. The set of edges $E \subseteq V \times V$ represents the available connections, i.e. $(i, j) \in E$ implies that the node s_i may send information to the node s_j. Moreover we assume that each node of the sensor network acts as a gateway for a specific number of other nodes, which means that every time it communicates with another node it sends, in a single packet, its own measurements collected together with all data received from the other nodes.

We always assume that, for every sensor in the network, there exists one and only one communication path to the fusion center, i.e. the sensor network has a directed tree topology. Moreover, we assume that each link has an associated weight $c(e_{i,j})$ which indicates the energy consumed when s_i directly transmits a packet to s_j. For the sake of legibility, we sometimes abbreviate $c(e_{i,j})$ as c_i, $i = 1, \ldots, m$ because, in the assumed topology, each sensor node has only one outgoing edge.

Remark 1. *The tree topology assumption may be a restrictive hypothesis for the general cases where one sensor can communicate with several nearby nodes. However, it is worth to remark that typical communication network graphs can be approximated by a collection of "representative" spanning trees (e.g. the first m spanning trees of the spanning tree enumeration [17]). Space constraints are forcing the authors to defer this discussion to future works.*

[1]The $'$ on a matrix always means transpose.

[2]The assumption of observability is without loss of generality since we could perform Kalman decomposition and only consider the observable space even if the system is not observable.

B. Stochastic v.s. Deterministic Sensor Selection

In order to reduce the energy consumption it is desirable to use a subset of sensors at each sampling time because of the redundancy in sensor measurements. However, in a tree topology, we cannot select arbitrary subsets of nodes as we are forced to select nodes to guarantee that there exists a communication path to the fusion center for each selected node. As a result, any possible transmission topology of G is a subtree $T = \{V_T, E_T\}$, with $s_0 \in V_T$, $V_T \subseteq V$ and $E_T \subseteq E$. Hereafter, V_T denotes the selected subset of sensors and E_T the communication links used by the sensors to transmit observations to the fusion center. We also denote by $\mathcal{T}$ the set of all possible transmission topologies T (i.e. the set of all possible subtrees of G containing s_0). It is straightforward to show that, for a transmission tree T, the total transmission energy consumption is given by[3]

$$\mathcal{E}(T) = \sum_{e \in E_T} c(e).$$

Suppose that at each time k we randomly select a tree T from $\mathcal{T}$ and each sensor in T transmits its observation back to the fusion node according to the topology T. Let $\pi_{k,T}$ be the probability that the transmission tree T is selected at time k. Then, we define

$$p_{k,i} \triangleq \sum_{T \in \mathcal{T}, s_i \in V_T} \pi_{k,T} \tag{3}$$

as the marginal probability that sensor i is selected at time k. Further, let us define the aggregate vector $\mathbf{p}_k = [p_{k,1}, \ldots, p_{k,m}]'$ and $\pi_k = [\pi_{k,T_1}, \ldots, \pi_{k,T_{|\mathcal{T}|}}]'$ containing all $p_{k,i}$s and $\pi_{k,T}$s respectively. We introduce also a binary random variable $\delta_{k,T}$ such that $\delta_{k,T} = 1$ if the transmission tree T is selected at time k and $\delta_{k,T} = 0$ otherwise. Similarly, the binary random variable $\gamma_{k,i}$ will take value 1 if the i-th sensor is selected at time k and 0 otherwise. It is well known that the Kalman filter is still optimal [16] for the above time-varying system structure. Suppose next that $V_T = \{s_0,\ s_{i_1}, \ldots, s_{i_j}\}$. Then, we can define

$$C_T \triangleq [C_{i_1}',\ C_{i_2}', \ldots, C_{i_j}']', R_T \triangleq diag(r_{i_1}, \ldots, r_{i_j}). \tag{4}$$

It can be shown that the estimation error covariance matrix P_k of the Kalman filter satisfies the following recursive equations:

$$P_k = \left(P_{k|k-1}^{-1} + C_T' R_T^{-1} C_T\right)^{-1}, \tag{5}$$

where $P_{k|k-1} = AP_{k-1}A' + Q$. Let us define $\mathbf{g}_{\pi_k,k}$ as a random operator such that

$$\mathbf{g}_{\pi_k,k}(X) \triangleq \sum_{T \in \mathcal{T}} \delta_{k,T} g_T(X), \tag{6}$$

where $P(\delta_{k,T} = 1) = \pi_{k,T}$, and

$$g_T(X) \triangleq \left[(AXA' + Q)^{-1} + \sum_{s_i \in V_T,\, s_i \neq s_0} \frac{C_i C_i'}{r_i}\right]^{-1}. \tag{7}$$

We have

$$P_k = \mathbf{g}_{\pi_k,k}(P_{k-1}). \tag{8}$$

In this paper we are interested in determining a time-invariant stochastic schedule π, i.e. a schedule where the probability of choosing a tree is constant over time. Hence, let us define

$$\mathbf{g}_\pi^\infty(X) \triangleq \lim_{k \to \infty} \mathbb{E}(g_{\pi,k} \circ g_{\pi,k-1} \circ \cdots \circ g_{\pi,1})(X), \tag{9}$$

[3]Here we assume that $cost(e_{i,j})$ is constant regardless of the number of observations contained in the packet. This is realistic in most of the cases, especially when measurements are of simple type, such as low precision scalar values, and the transmission overhead, e.g. header, handshaking protocol, dominates the payload.

when the limit exists and infinity otherwise. Note that $\mathbf{g}_\pi^\infty$ is a deterministic function which indicates what asymptotic performance of the stochastic sensor selection scheme is to be expected. It is easy to see that

$$\lim_{k\to\infty} \mathbb{E}P_k = \mathbf{g}_\pi^\infty(\Sigma),$$

when π is used and $\mathbf{g}_\pi^\infty(\Sigma) < \infty$.

Because the transmission trees are randomly selected, P_k is a random matrix. Thus, we minimize the asymptotic expected estimation error covariance matrix while requiring that the expected energy consumption does not exceed a designated threshold $\mathcal{E}_d$. Such problem can be formulated as follows:

Problem 1 (Fixed Random Schedule that Optimizes Expected Asymptotic Performance)**.**

$$\begin{aligned} &\underset{\pi}{\textit{minimize}} && trace(\mathbf{g}_\pi^\infty(\Sigma)) \\ &\textit{subject to} && \sum_{T\in\mathcal{T}} \pi_T \mathcal{E}(T) \le \mathcal{E}_d,\ \pi_T \ge 0,\ \sum_{T\in\mathcal{T}} \pi_T = 1. \end{aligned}$$

Please note that if, instead of a stochastic schedule, we want to look for a fixed deterministic one, the problem we need to solve is the same Problem 1 with the further constraints that π_Ts are forced to be either 0 or 1, namely :

Problem 2 (Fixed Deterministic Schedule that Optimizes Asymptotic Performance)**.**

$$\begin{aligned} &\underset{\pi}{\textit{minimize}} && trace(\mathbf{g}_\pi^\infty(\Sigma)) \\ &\textit{subject to} && \sum_{T\in\mathcal{T}} \pi_T \mathcal{E}(T) \le \mathcal{E}_d,\ \pi_T = 0 \textit{ or } 1,\ \sum_{T\in\mathcal{T}} \pi_T = 1. \end{aligned}$$

Remark 2. *In Problem 1 we require that the expected energy consumption does not exceed a certain energy budget. In real applications different constraints may be considered (e.g. requirements on sensor lifetime). However, it can be shown (see e.g. [13]) that many of these constraints can be easily integrated into the present framework.*

Remark 3. *It is worth noticing that at each sampling time, the energy cost of the deterministic schedule cannot exceed the designated threshold $\mathcal{E}_d$. Intuitively, this is the reason why stochastic sensor selections, being allowed to use more energy at one single sampling period, can achieve better performance than the above deterministic formulation. It is also worth noticing that this formulation can be used to derive a periodic schedule by enlarging the state space. As a consequence, all the results in this section can be generalized to periodic schedules. However, due to space constraints, Section III focuses exclusively on time-invariant schedules. Such an extension will be discussed in future work.*

Remark 4. *A main difference between Problem 1 and Problem 2 is that, while the search space of deterministic scheduling is discrete, the one for the stochastic problem is continuous and convex. This brings several advantages. First, the deterministic schedule can be seen as a particular kind of random schedule, where π_Ts are binary-valued. As a result, stochastic sensor selection strategies always yield better or equal performance, in the expected sense, when compared to the deterministic one. The second advantage is that the feasible set π_T is convex, which allows us to further manipulate the objective function to obtain a convex optimization problem.*

As mentioned above, the expected performance of the optimal stochastic schedule is equal to or better than its deterministic counterpart. As a consequence, if we denote π^* as the optimal stochastic schedule and π_d^* the optimal deterministic schedule, we can assert that

$$trace(\mathbf{g}_{\pi^*}^\infty(\Sigma)) = \lim_{k\to\infty} \mathbb{E}\, trace(P_k(\pi^*)) \le \lim_{k\to\infty} trace(P_k(\pi_d^*)),$$

which implies that

$$\lim_{N\to\infty} \sum_{k=1}^N \frac{1}{N} \mathbb{E}\, trace(P_k(\pi^*)) \le \lim_{N\to\infty} \sum_{k=1}^N \frac{1}{N} trace(P_k(\pi_d^*)),$$

In other words, the optimal stochastic schedule performs better than the optimal deterministic one in the expected sense. To strengthen this result, the following theorem states that almost surely the stochastic schedule is better than the deterministic schedule almost surely.

Theorem 1. *Suppose that the fixed schedule π^* is the solution of Problem 1. If the linear system and π^* satisfy the following assumptions:*

1) *A is invertible, $(A, Q^{1/2})$ is controllable;*
2) *there exists a transmission topology T with $\pi_T^* > 0$ such that $(C_T,\ A)$ is observable*
3) *the stochastic process $\{P_k\}$ satisfies: $P_k = \mathbf{g}_{\pi^*,k}(P_{k-1})$, $P_0 = \Sigma$,*

then almost surely the following inequality holds

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^N trace(P_k) \le trace(\mathbf{g}_{\pi^*}^\infty(\Sigma)). \tag{10}$$

Proof: It is easy to check that all assumptions in Theorem 3.4 of [18] hold true. As a result, there exists an ergodic stationary process $\{\overline{P}_k\}$ which satisfies $\overline{P}_k = \mathbf{g}_{\pi^*,k}(\overline{P}_{k-1})$. Moreover,

$$\lim_{k\to\infty} \|P_k - \overline{P}_k\| = 0.\ a.s.$$

We want to prove that $\mathbb{E}(trace(\overline{P}_0))$ is lower than or equal to $trace(\mathbf{g}_\pi^\infty(\Sigma))$ and hence bounded. Because $\overline{P}_k$ is ergodic, and P_k converges to $\overline{P}_k$ almost surely, we know that

$$\begin{aligned} &\lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^N \min(trace(P_k), M) = \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^N \min(trace(\overline{P}_k), M) \\ &= \mathbb{E}[\min(trace(\overline{P}_0), M)], \quad a.s. \end{aligned}$$

where $M > 0$ is a constant. By the definition of $\mathbf{g}_\pi^\infty$, we know that

$$\begin{aligned} &trace(\mathbf{g}_\pi^\infty(\Sigma)) \ge \lim_{N\to\infty} \mathbb{E}\left[\frac{1}{N} \sum_{k=1}^N \min(trace(P_k), M)\right] \\ &= \mathbb{E}\left[\lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^N \min(trace(P_k), M)\right] = \mathbb{E}[\min(trace(\overline{P}_0), M)]. \end{aligned}$$

The second equality follows from the Dominated Convergence Theorem. Now, let $M \to \infty$. By the Monotone Convergence Theorem, it results that

$$\mathbb{E}[trace(\overline{P}_0)] = \lim_{M\to\infty} \mathbb{E}[\min(trace(\overline{P}_0), M)] \le trace(\mathbf{g}_\pi^\infty(\Sigma)),$$

which proves that $\mathbb{E}[trace(\overline{P}_0)] \le trace(\mathbf{g}_\pi^\infty(\Sigma))$. Hence, by ergodicity, we obtain

$$\begin{aligned} &\lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^N trace(P_k) = \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^N trace(\overline{P}_k) = \mathbb{E}(trace(\overline{P}_0)) \\ &\le trace(\mathbf{g}_\pi^\infty(\Sigma)), \quad a.s. \end{aligned}$$

■

Remark 5. *Combining Remark 4 with Theorem 1, we conclude that the average performance of almost every sample path of the optimal stochastic schedule has a lower cost than its deterministic counterpart.*

Before moving forward, it is worth noticing that Problem 1 is still numerically intractable for the following reasons:

1) it is usually difficult to express $\mathbb{E}P_\infty$ as an explicit function of $\pi_{1,T}, \ldots, \pi_{k,T}$;[4]
2) since $|\mathcal{T}|$ is large, the number of optimization variables and constraints may be not polynomial with respect to the number of nodes.

In the next section, we devise a way to overcome the above two problems.

III. Relaxation and Implementation

In this section, we first relax Problem 1 to an explicit convex problem. We then propose an implementation of our stochastic schedule without introducing any communication overhead.

A. Relaxation

A lower-bound L_k to $\mathbb{E}P_k$ is first derived by means of the following theorem, whose proof is reported in a technical report [20].

Theorem 2. *Let $L_0 = P_0$ and*

$$L_k = \left(L_{k|k-1}^{-1} + \sum_{i=1}^{m} p_{k,i} \frac{C_i C_i'}{r_i} \right)^{-1}, \tag{11}$$

where $L_{k|k-1} = AL_{k-1}A' + Q$. The following inequalities hold:

$$\mathbb{E}P_k \geq L_k. \tag{12}$$

To further improve the legibility, let us define the function

$$L(X, \mathbf{p}) \triangleq \left[(AXA' + Q)^{-1} + \sum_{i=1}^{m} p_i \frac{C_i C_i'}{r_i} \right]^{-1}, \tag{13}$$

where $X \in \mathbb{R}^{n \times n}$ is positive semidefinite and $\mathbf{p} = [p_1, \ldots, p_m]' \in \mathbb{R}^m$. Moreover, let us define,

$$L^{(1)}(X, \mathbf{p}) = L(X, \mathbf{p}), L^{(k)}(X, \mathbf{p}) = L(L^{(k-1)}(X, \mathbf{p}), \mathbf{p}), \tag{14}$$

with

$$L^\infty(X, \mathbf{p}) = \lim_{k \to \infty} L^{(k)}(X, \mathbf{p}), \tag{15}$$

when the limit exists. Hence (11) can be simplified as

$$L_k = L(L_{k-1}, \mathbf{p}_k). \tag{16}$$

By replacing the objective function in Problem 1 with its lower bound, we obtain

Problem 3 (Asymptotic Lower Bound for Random Transmission Tree Selection)**.**

$$\begin{aligned} \underset{\pi_T, \mathbf{p}}{\textit{minimize}} \quad & trace(L^\infty(\Sigma, \mathbf{p})) \\ \textit{subject to} \quad & \sum_{T \in \mathcal{T}} \pi_T \mathcal{E}(T) \leq \mathcal{E}_d, \\ & \pi_T \geq 0, \sum_{T \in \mathcal{T}} \pi_T = 1, \, p_i = \sum_{s_i \in V_T} \pi_T. \end{aligned}$$

There are two drawbacks in the above formulation: 1) the optimization problem still has the π_T terms as optimization variables, the number of which grows more than polynomially with respect to m in general; 2) L^∞ is still not an explicit function of the optimization variables.

[4]The readers can refer to [19] for more information.

Let us first drop the dependence on π_T. To this end, define the set of feasible $\mathbf{p}$ for Problem 3:

$$\mathcal{P} \triangleq \left\{ \mathbf{p} \in \mathbb{R}^m \,\middle|\, \exists \pi, \sum_{T \in \mathcal{T}} \pi_T \mathcal{E}(T) \leq \mathcal{E}_d, \right. \\ \left. \pi_T \geq 0, \sum_{T \in \mathcal{T}} \pi_T = 1, \, p_i = \sum_{s_i \in V_T} \pi_T \right\}.$$

The following results can be easily proved:

Proposition 1. *The energy cost of a given collection of tree selection probabilities $\pi_{k,T}, \forall T \in \mathcal{T}$ is a linear function of the resulting marginal probability:*

$$\sum_{T \in \mathcal{T}} \pi_T \mathcal{E}(T) = \sum_{i=1}^{m} c_i p_i. \tag{17}$$

Proposition 2. *If $p_i \in [0, 1]$ and if it satisfies*

$$p_i \leq p_j, \qquad \text{if } j \text{ is a parent of } i \tag{18}$$

then there exists at least one collection of tree selection probabilities π, such that

$$\pi_T \geq 0, \; \sum_{T \in \mathcal{T}} \pi_T = 1, \; p_i = \sum_{s_i \in V_T} \pi_T. \tag{19}$$

Conversely, if there exists π_k such that (19) *holds, then $p_{k,i} \in [0, 1]$ and satisfies* (18) .

By exploiting the above Propositions we can reformulate the feasible set $\mathcal{P}$ as follows

$$\mathcal{P} = \left\{ \mathbf{p} \,\middle|\, p_i \in [0, \, 1], \sum_{i=1}^{m} c_i p_i \leq \mathcal{E}_d, \, p_i \leq p_j, \text{if } j \text{ is parent of } i \right\}, \tag{20}$$

and we can rewrite Problem 3 as

Problem 4 (Asymptotic Lower Bound for Random Transmission Tree Selection)**.**

$$\begin{aligned} \underset{\mathbf{p} \in \mathbb{R}^m}{\textit{mininize}} \quad & trace(L^\infty(\Sigma, \mathbf{p})) \\ \textit{subject to} \quad & \mathbf{p} \in \mathcal{P}. \end{aligned}$$

The main difficulty in solving the above problem is that $L^\infty(X, \mathbf{p})$ is in general not convex in $\mathbf{p}$ and its exact form is unknown. Here we propose the use of the following heuristic:

1) Define $\mathbf{p}_0 = \left(\mathcal{E}_d / \left(\sum_{i=1}^m c_i \right) \right) \mathbf{1}_m$, where $\mathbf{1}_m \in \mathbb{R}^m$ is a vector with '1' in all entries and choose the matrix $L_0 = L^\infty(I_n, \mathbf{p}_0)$.
2) Let L_k and $\mathbf{p}_k$ be the solution of the following optimization problem

   Problem 5 (Random Sensor Selection with Descent Constraint)**.**

$$\begin{aligned} \underset{\mathbf{p}_k \in \mathbb{R}^m}{\textit{minimize}} \quad & trace(L_k) (= trace(L(L_{k-1}, \mathbf{p}_k))) \\ \textit{subject to} \quad & L_k \leq L_{k-1}, \, \mathbf{p}_k \in \mathcal{P}. \end{aligned}$$

3) Choose $\mathbf{p}^*$ as an accumulation point of $\mathbf{p}_k$[5]. Then $L^\infty(X, \mathbf{p}^*) = \lim_{k \to \infty} L_k$ for any $X \geq 0$.

Before proving the feasibility of the above algorithm, we want to point out that our algorithm is greedy. In fact, we try to minimize the lower bound for the next step in the hope of reducing the final asymptotic lower bound. As a result, it is suboptimal by nature. The

[5]An accumulation point of a sequence is the limit of a converging subsequence

following theorem [20] gives a characterization of the main features of the proposed algorithm.

Theorem 3. *The following statements are true for the proposed algorithm:*

1) L_0 *exists.*
2) $L(X, \mathbf{p})$ *is convex with respect to* $\mathbf{p}$ *and it is concave and monotonically increasing with respect to* X.
3) *Problem 5 is always feasible.*
4) $\mathbf{p}^*$ *exists and* $\mathbf{p}^* \in \mathcal{P}$.
5) $L_\infty = \lim_{k\to\infty} L_k$ *exists.*
6) $L_\infty = L^\infty(X, \mathbf{p}^*)$ *for all positive semidefinite* X.

Proof:

1) The proof is reported in [20].
2) The proof is reported in [20].
3) Suppose that the Problem 5 is feasible up to time k. To prove the problem is also feasible at time $k+1$, we only need to find one $\mathbf{p} \in \mathcal{P}$ and $L(L_k, \mathbf{p}) \leq L_k$. If we choose $\mathbf{p} = \mathbf{p}_k$ then, since $\mathbf{p}_k$ is the solution at time k, it follows that $\mathbf{p}_k \in \mathcal{P}$. It remains to show that $L(L_k, \mathbf{p}_k) \leq L_k$, which can be proved by noticing that $L_k = L(L_{k-1}, \mathbf{p}_k) \leq L_{k-1}$ and $L(X, \mathbf{p})$ is monotonically increasing with respect to X. Similarly, Problem 5 is also feasible at time 1 and then, by induction, Problem 5 is always feasible.
4) It is easy to see that $\mathbf{p}_k$ is bounded because $p_{k,i} \in [0, 1]$. By means of the Bolzano-Weierstrass Theorem, this implies that there always exists an accumulation point $\mathbf{p}^*$. Moreover, since $\mathbf{p}_k \in \mathcal{P}$ and $\mathcal{P}$ is closed, $\mathbf{p}^* \in \mathcal{P}$.
5) The limit must exist because $\{L_k\}$ is decreasing and $L_k \geq 0$ for all k,
6) The proof is reported in [20].

■

Remark 6. *Due to the convexity of L and* $\mathcal{P}$*, Problem 5 is a convex optimization problem with* $O(m)$ *optimization variables and* $O(m)$ *constraints. Thus, it can be solved efficiently. For example, if interior-points methods are used, then the complexity is* $O(m^3)$ *[21].*

Remark 7. *It is worth noticing that in general there may exist more than one set of* $\pi_T, \forall T \in \mathcal{T}$ *with the same marginal probabilities. One possible way to determine a feasible* π_T *is as follows:*

1) *Sort the marginal probability* p_i*, suppose that* $p_{i_1} \geq p_{i_2} \geq \ldots \geq p_{i_m}$.
2) *Define* $T_0 = \{s_0\}$, $T_j = T_{j-1} \bigcup \{s_{i_j}\}$.
3) *Choose* $\pi_{T_0} = 1 - p_{i_1}$, $\pi_{T_1} = p_{i_1} - p_{i_2}, \pi_{T_2} = p_{i_2} - p_{i_3}, \ldots, \pi_{T_m} = p_{i_m}$.

One can easily verify that $T_i \in \mathcal{T}$ *and* π_T *are compatible with the marginal probability.*

B. Implementation

In this subsection we discuss a possible implementation of our sensor selection scheme. We assume the optimal marginal probability is $\mathbf{p}$, which is computed off-line in a centralized fashion. Each i-th sensor stores the optimal values p_i and p_j of all its children.

At each time k, we have to select sensors according to the marginal probabilities $\mathbf{p}$. However, we don't want the fusion center to query the nodes because this would increase the communication overhead, defying the purpose of sensors selection. To overcome this problem, we propose the following algorithm:

1) Every sensor is equipped with the same random number generator and the same seed.
2) At time k, each sensor draws a random number α_k from the random number generator.
3) If sensor i has no children, then it compares α_k with p_i. If $\alpha_k \leq p_i$, then it transmits the measurement to its parent. Otherwise, it does not transmit anything.
4) If sensor i has children, then it compares α_k with p_j, where j is the index of its child node. If $\alpha_k \leq p_j$, then sensor i knows that child j will forward an observation packet to him. After the node i receives all the observation packets from its children, it merges all packets and its own observations into a single packet and forwards it to its parent. If $\alpha_k > p_j$ for all j-th child of i, then the node i compares α_k with p_i. If $\alpha_k \leq p_i$, then sensor i transmits its measurements to its parent. Otherwise, it does not transmit anything.

Since all sensors are equipped with the same random number generator and the same seed, every sensor gets the same α_k at time k. Hence, the above algorithm guarantees that all sensors agree on the same transmission topology T which satisfies the marginal distribution $\mathbf{p}$. It is worth remarking that in such a scheme the only required data exchange amongst nodes is the transmission of the observation packets and no communication overhead for coordination purposes is needed.

Remark 8. *It is worth mentioning that, because all sensors agree on the same* α_k*, it is very easy to implement a Time Division Multiple Access (TDMA) protocol to avoid wireless interference.*

IV. Simulation Result

In order to show the effectiveness of the proposed method we apply our stochastic sensor selection algorithm to a numerical example in which a sensor network is deployed to monitor a diffusion process in a $l \times l$ planar closed region, whose model is given by

$$u_t = \alpha \nabla^2 u. \tag{21}$$

where ∇^2 is the Laplace operator. The term $u(t, x_1, x_2)$ denotes the temperature at location (x_1, x_2) at time t and α indicates the speed of the diffusion process.

We use the finite difference method to discretize this model by gridding the region into $1\ meter \times 1\ meter$ tiles and time steps of $1s$. If we group all temperature values at time k in the vector $U_k = [u(k,0,0), \ldots, u(k,0,N-1), u(k,1,0), \ldots, u(k,N-1,N-1)]^T$, we can write the evolution of the discretized system as $U_{k+1} = AU_k$, where the A matrix can be computed from discretization. If we introduce process noise, U_k will evolve according to $U_{k+1} = AU_k + w_k$, where $w_k \in \mathcal{N}(0,\ Q)$ is the process noise.

We suppose that the fusion center is located in the bottom left corner at position $(0, 0)$. We assume that m sensors are randomly distributed in the region and that each sensor measures a linear combination of temperature on the grid vertices around it[6]. In particular, if we suppose the location of sensor l of coordinates (a_1, a_2) is in the cell $[i, j]$, i.e. $a_1 \in [i, i+1)$ and $a_2 \in [j, j+1)$, the measurement of this sensor is given by

$$y_{k,l} = [\ (1-\Delta a_1)(1-\Delta a_2)u(k,i,j) + \Delta a_1(1-\Delta a_2)u(k,i+1,j) + (1-\Delta a_1)\Delta a_2 u(k,i,j+1) + \Delta a_1 \Delta a_2 u(k,i+1,j+1)\ ]/h^2 + v_{k,l}.$$

where $\Delta a_1 = a_1 - i$, $\Delta a_2 = a_2 - j$ and $v_{k,l}$ is the measurement noise of sensor l at time k. Indicating with Y_k the vector of all measurements at time k, it follows that: $Y_k = CU_k + v_k$, where v_k denotes the measurement noise at time k assumed to have normal distribution $\mathcal{N}(0, R)$ and C is the observation matrix. Finally, we assume that the sensor network admits a minimum spanning tree topology with communication cost from sensor i to j given by

$$cost(e_{i,j}) = c + d_{i,j}^2,$$

[6]We do not require the sensors to be placed at grid points

where d_{ij} is the Euclidean distance from sensor i to sensor j and c is a constant related to the energy consumption for sensing[7]. For the simulations, we impose the following parameters: $l = 3\ meters$, $m = 16$, $\alpha = 0.1$, $Q = I = R = I \in \mathbb{R}^{16\times 16}$, $\Sigma = 4I \in \mathbb{R}^{16\times 16}$, $\mathcal{E}_d = 6$,$c = 1$.

We compare the performance of the proposed fixed stochastic schedule with the optimal fixed deterministic one found by exhaustive search. Figure 1 shows the histogram of the ratio between $trace(P_\infty)$ of the deterministic schedule and $trace(EP_\infty)$ of the stochastic one computed used the proposed algorithm, computed over 100 random sensor placements. For each placement we computed the empirical expectation. The blue dashed line is the average ratio. It can be seen that the deterministic schedule is always worse than the stochastic one, which yields over 35% improvement over the deterministic optimum in the average. Figure 2 shows the trace of P_k for the optimal deterministic fixed schedule, together with the trace of P_k from a sample path and EP_k of the stochastic fixed schedule for one placement. The figure clearly shows how the average savings are due to the fact that the stochastic schedule is not constrained to stay within the energy budget at each step.

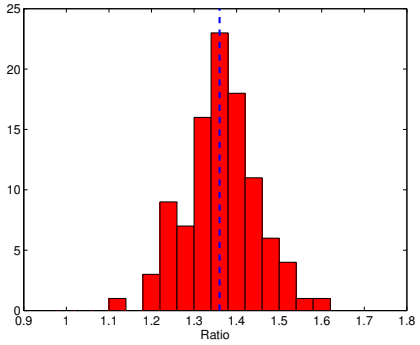

Fig. 1. Histogram of the ratio between $trace(P_\infty)$ of deterministic schedule and $trace(EP_\infty)$ of stochastic schedule

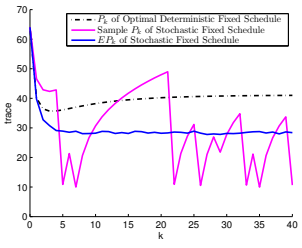

Fig. 2. Evolution of $trace(P_k)$

[7] c models the fact that as the distance goes to zero the communication cost does not

V. Conclusions

In this paper, we propose a stochastic sensor selection algorithm for a wireless sensor network with a tree topology. We solve the optimal stochastic sensor selection problem after relaxation by means of convex optimization. We show that a stochastic formulation is preferable over a deterministic one in terms of both performance and computational complexity. We also provide an implementation of our random sensor selection algorithm that doesn't introduce any communication overhead, making the results proposed in this paper appealing for practical implementations. Simulations results validate the effectiveness of the proposed approach. Future work will include extensions to arbitrary graphs.

References

[1] N. P. Mahalik, *Sensor Networks and Configuration*. Springer, 2007.

[2] 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*, vol. 4, no. 3, pp. 228 – 242, May-June 2005.

[3] S. Liang, Y. Tang, and Q. Zhu, "Passive wake-up scheme for wireless sensor networks," in *Second International Conference on Innovative Computing, Information and Control*, Sept 2007.

[4] 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*, Jan 2006.

[5] Q. Xue and A. Ganz, "Maximizing sensor network lifetime: analysis and design guidelines," in *IEEE Conference on Military Communications*, Nov 2004.

[6] A. Chamam and S. Pierre, "Optimal scheduling of sensors' states to maximize network lifetime in wireless sensor networks," in *IEEE Internatonal Conference on Mobile Adhoc and Sensor Systems*, Oct 2007.

[7] 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*, June 2008.

[8] B. M. Miller and W. J. Runggaldier, "Optimization of observations: A stochastic control approach," *SIAM J. Control Optim.*, vol. 35, no. 3, pp. 1030–1052, May 1997.

[9] R. K. Mehra, "Optimiztion of measurement schedules and sensor designs for linear dynamic systems," *IEEE Trans. Automatic Control*, vol. 21, no. 1, pp. 55–64, Feburary 1976.

[10] V. Krishnamurthy, "Algorithms for optimal scheduling and management of hidden markov model sensors'," *IEEE Trans. Signal Processing*, vol. 50, no. 6, pp. 1382–1397, June 2002.

[11] W. We and A. Arapostatish, "Optimal sensor querying: General markovian and lqg models with controlled observations," *IEEE Trans. Automatic Control*, vol. 53, no. 6, pp. 1392–1405, July 2008.

[12] Y. Oshman, "Optimal sensor selection strategy for discrete-time state estimators," *IEEE Trans. Aerospace and Electronic Systems*, vol. 30, no. 2, pp. 307–314, May 1994.

[13] S. Joshi and S. Boyd, "Sensor Selection via Convex Optimization," *IEEE Transactions on Signal Processing*, vol. 57, no. 2, pp. 451–462, February 2009.

[14] Y. Mo, R. Ambrosino, and B. Sinopoli, "Network energy minimization via sensor selection and topology control," in *First IFAC Workshop on Estimation and Control of Networked Systems*, Venice, Italy, September 2009.

[15] ——, "A convex optimization approach of multi-step sensor selection under correlated noise," in *Proceedings of the Forty-Seventh Annual Allerton Conference on Communication, Control, and Computing*, Monticello, IL, September 2009.

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

[17] K. Sorensen and G. K. Janssens, "An algorithm to generate all spanning trees of a graph in order of increasing cost," *Pesqui. Oper. [online]*, vol. 25, pp. 219–229, 2005.

[18] P. Bougerol, "Some results on the filtering riccati equation with random parameters," in *Applied Stochastic Analysis*, ser. Lecture Notes in Control and Information Sciences. Springer Berlin, 1992, vol. 177, pp. 30–37.

[19] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. Jordan, and S. Sastry, "Kalman Filtering with Intermittent Observations," *IEEE Trans. on Automatic Control*, vol. 49, no. 9, pp. 1453–1464, September 2004.

[20] Y. Mo, E. Garone, A. Casavola, and B. Sinopoli, "Stochastic sensor scheduling for energy constrained estimation in multi-hop wireless sensor networks," 2011, http://arxiv.org/abs/1107.1752.

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