On Sensor Coverage by Mobile Sensors

Vijay Gupta, David E. Jeffcoat and Richard M. Murray

Abstract— **We study the problem of using a small number of mobile sensors to monitor various threats in a geographical area. Using some recent results on stochastic sensor scheduling, we propose a stochastic sensor movement strategy. We present simple conditions under which it is not possible to maintain a bounded estimate error covariance for all the threats. We also study a simple sub-optimal algorithm to generate stochastic trajectories. Simulations are presented to illustrate the results.**

I. INTRODUCTION AND MOTIVATION

In recent years, systems comprised of multiple mobile sensors acting cooperatively have garnered increasing attention. Even though using a network of mobile sensors is more complicated than having one static sensor, the increased accuracy and flexibility often makes it worthwhile. In this work, we look at the problem of monitoring a geographical region using multiple sensors cooperatively. Each sensor can individually sense a limited region, but together the sensors must monitor the entire area. The problem of optimal sensor location in case there are no bounds on the range over which the sensors can sense leads to the problem of Voronoi partitioning of the space and has been solved both in a centralized framework [2] and in a decentralized fashion [3]. The latter reference also discussed imposing sensing range restrictions and proposed altering the cost function in a similar manner to deal with the situation. The problem when there are range (or direction) limitations on the sensors has also been looked at other places in the literature. However most of the approaches that have been proposed are very application specific, see [4] for a typical example. Finally some greedy approaches have also been proposed to determine the optimal sensor trajectory using different cost functions, e.g., in [5], [6].

A different approach was proposed in [1], which showed the formal similarity between the sensor trajectory generation and the sensor scheduling problem. A stochastic algorithm was proposed to solve the problem. The algorithm differs from the other approaches mentioned above in that it is based on the idea of letting the sensors switch randomly according to some optimal probability distribution to obtain the best *expected* steady-state performance. In this note, we explore the algorithm further for the problem of sensor coverage. We consider multiple sensors patrolling a grid of

Division of Engineering and Applied Science, California Institute of Technology, Pasadena, CA 91125, USA vijay@cds.caltech.edu. Research supported in part by the AFOSR grant F49620-01-1-0460 and in part by NSF grant CCR-0326554.
Air Force Research Laboratory, Munitions Directorate, Eglin Air Force Base, FL 32542, USA david.jeffcoat@eglin.af.mil
Division of Engineering and Applied Science, California Institute of Technology, Pasadena, CA 91125, USA murray@caltech.edu.

points and identify conditions when no stochastic strategy would lead to a bounded expected error covariance. We also provide a gradient descent based algorithm to generate the optimal probability distribution.

This paper is organized as follows. We set up the problem in the next section. We then present some results that place a bound on how fast the dynamics of the underlying systems evolve and how many sensors need to be present for the estimation error to be unbounded. Finally we present an algorithm to design an optimal probability distribution.

II. PROBLEM SETUP

Let the geographical region that needs to be monitored be divided into a grid of N points. There are dynamical processes occurring at these points whose state we want to estimate. Denote the state at the i-th point at time k by x_i^k. The process at the i-th point is driven by w_i^k, assumed to be zero mean, white and Gaussian with covariance matrix R_i. We assume that the noises at two distinct points i and j are uncorrelated. We consider two distinct cases:

1) Coupled processes: The processes at points i and j are coupled. Thus the process at a point i evolves as

$$x_i^{k+1} = A_i x_i^k + \sum_{j \neq i} A_{i,j} x_j^k + w_i^k,$$

where all the matrices $A_{i,j}$ are not zero.

2) Uncoupled processes: Processes at distinct points are unaffected by each other. All matrices $A_{i,j}$ are zero.

Denote the state of the entire region obtained by stacking all x_i^k's as x^k. Then x^k evolves according to the equation

$$x^{k+1} = A x^k + w^k,$$

where w^k is the vector formed by stacking w_i^k's. If the processes are uncoupled, A is a (block) diagonal matrix with A_i's along the diagonal.

The region is monitored using n sensors. If the m-th sensor is at point i at time k, it generates the measurement

$$y_m^k = x_i^k + v_m^k, \tag{1}$$

where v_m^k is zero mean white Gaussian noise with covariance R_m, assumed independent of all other noises present. This can be rewritten as an observation of the state x^k as

$$y_m^k = C_i x^k + v_m^k,$$

where C_i is a row vector with zeros everywhere except the i-th element which is replaced by 1.[1] This gives rise to

---

[1]This description of C_i assumes the states x_i^k's to be scalars. The extension to the vector case is obvious. Similarly we can consider a sensing matrix being present in (1).

the concept of a virtual sensor. A *physical* sensor at point i gives rise to a *virtual* sensor being used with sensing matrix C_i. Similarly, if there are physical sensors at points i and j, we will say that a virtual sensor is being used that has a sensing matrix with rows C_i and C_j.

If there is more than one sensor present, we assume that all measurements are exchanged without delay or distortion. Thus based on all the measurements obtained by all the sensors, any sensor can compute an estimate $\hat{x}^k$ of the state x^k.[2] Let P^k denote the covariance of the estimate error,

$$P^k = E\left[(x^k - \hat{x}^k)^T(x^k - \hat{x}^k)\right]. \quad (2)$$

There are two basic problems that arise.

1) Under what conditions does P^k remain bounded?
2) What is the optimal trajectory that minimizes P^k?

The error covariance is a function of the trajectory of the physical sensor, or of the sensor schedule for the virtual sensor. All the possible sensor schedules can be represented by a tree. The depth of any node represents time steps with the root at time zero. The branches correspond to choosing a particular sensor to be active at that time instant. Thus, the path from the root to a node at depth d represents a sensor schedule for time steps 0 to d. We can associate with each node the cost function evaluated using the schedule corresponding to the path from the root to that node. Finding *the* optimal sequence requires traversing all the exponential number of paths from the root to the leaves in the tree. This procedure might place too high a demand on the computational and memory resources of the system.

In this note, we forgo traversing the tree altogether and propose stochastic trajectories, i.e., the sensors choose their positions at any time step at random according to a probability distribution. The probability distribution is chosen so as to minimize the expected steady state error covariance. We cannot calculate the exact value of the error covariance since that will depend on the specific sensor schedule chosen. Hence we optimize the expected value of the error covariance. To characterize the expected error covariance, we use some of the framework developed in [1].

In this work, we assume that the sensor trajectories are designed independent of each other. There are two particular cases of sensor motion that we will study:

1) The choice for the position of the j-th sensor at time step $k+1$ is done in an i.i.d. fashion at each time step with probability q_i of being at the i-th point.
2) The choice is done according to a Markov chain with transition probability matrix Q. This can model physical constraints on the sensor motion, e.g., the probability q_{ij} is 0 if i and j are points that are physically distant from each other.

Note that we have assumed that each sensor chooses its trajectory according to the same parameters (probabilities q_i's or the transition probability matrix Q). We will say

[2]Since every sensor has access to the same information set, they would have identical estimates.

that the problem can be solved if there exists at least one choice of parameters such that beginning from any initial condition, the expected error covariance remains bounded as time progresses; otherwise the problem cannot be solved.

In the above description we have assumed that all the agents are interested in coming up with an estimate for the processes in the area. We can use the same framework if the data is transmitted to a central data processing node. In this case, we can also allow for communication channel imperfections. As an example, if the communication channel loses packets stochastically, we can model the time instants at which data loss occurs as being used up by a *fictitious* (as opposed to physical and virtual) sensor which has sensing matrix 0. The data can be lost in an i.i.d. or a Markovian fashion (e.g., according to the Gilbert-Elliot channel model).

III. MOTION GOVERNED BY I.I.D. CHOICES

In this section we will consider the case when each sensor is choosing the next point to move to in an i.i.d. fashion, with the probability of any sensor being at point i being q_i.

A. Uncoupled processes

From theorem 4 of [1], we can obtain the following result.

Proposition 1: (Theorem 4 from [1]) Let a process with evolution matrix A be observed at each time step by one among n sensors with sensing matrices $C_1, C_2, \cdots, C_n$ such that the i-th sensor is chosen at any time step with probability q_i. Denote $\lambda_{max}(A_i)$ as the eigenvalue with the maximum magnitude of the unobservable part of A when the pair (A, C_i) is put in the observable canonical form. Then a sufficient condition for the expected estimate error covariance to diverge from at least one initial value is

$$q_i|\lambda_{max}(A_i)|^2 > 1,$$

for any $i = 1, 2, \cdots, n$.
Let λ_i be the eigenvalue with the maximum magnitude of matrix A_i. Without loss of generality, we can assume that

$$|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_N|. \quad (3)$$

We can prove the following result.

Proposition 2: Consider the sensor coverage problem when N points are to be patrolled by one sensor. If the processes are uncoupled and (3) holds, then the problem cannot be solved if for any $1 \leq m \leq N - 1$,

$$\frac{\binom{N-1}{m}}{|\lambda_1|^2} + \frac{\binom{N-2}{m-1}}{|\lambda_2|^2} + \cdots + \frac{\binom{N-1-m}{0}}{|\lambda_{m+1}|^2} < \binom{N-1}{m-1}. \quad (4)$$

Proof: The m-th sufficient condition is obtained by considering all virtual sensors formed by considering sets of m points. We say that a virtual sensor is used if the physical sensor is present at any point in the set of m points that the virtual sensor represents.

1) There are $\binom{N}{m}$ such virtual sensors.
2) Denote the probability of choosing the j-th virtual sensor by π_j. For a virtual sensor with the set of m points denoted by $\mathcal{M}$, $\pi_j = \sum_{t \in \mathcal{M}} q_t$.

For each virtual sensor, denote the lowest i which is not included in its set of m points by i_{min}. Then the condition for stability when that virtual sensor is used is $\pi_j|\lambda_{i_{min}}|^2 < 1$. Simple algebra yields that λ_t occurs in $\binom{N-t}{m-t+1}$ such inequalities. Adding all the inequalities together, we obtain that at least one inequality will be violated if

$$\frac{\binom{N-1}{m}}{|\lambda_1|^2} + \frac{\binom{N-2}{m-1}}{|\lambda_2|^2} + \cdots + \frac{\binom{N-1-m}{0}}{|\lambda_{m+1}|^2} < \binom{N-1}{m-1}.$$

Considering different values of m, we obtain the result. $\blacksquare$
For specific values of m, the condition in equation (4) looks as follows. For $m = 1$, the condition is

$$\frac{N-1}{|\lambda_1|^2} + \frac{1}{|\lambda_2|^2} < 1. \qquad (5)$$

For $m = N - 1$, the condition is

$$\sum_{i=1}^{N} \frac{1}{|\lambda_i|^2} < N - 1. \qquad (6)$$

Neither of the conditions is more general. As an example, for a system with $\lambda_1 = 2$, $\lambda_2 = \sqrt{3}$, $\lambda_3 = 1/\sqrt{2}$ the problem is predicted to be unsolvable by (5) but not by (6). The opposite is true for a system with $\lambda_1 = \lambda_2 = \lambda_3 = \sqrt{2}$.

We now move on to the case when there is more than one physical sensor, i.e., $n > 1$. To begin with, consider the case of only two points to be patrolled, i.e., $N = 2$.

Proposition 3: If $N = 2$ points have to be patrolled by n sensors with the assumptions stated above, the sensor coverage problem cannot be satisfied if

$$\frac{1}{|\lambda_1|^{\frac{2}{n}}} + \frac{1}{|\lambda_2|^{\frac{2}{n}}} < 1.$$

Proof: There are 2^n virtual sensors in this case, corresponding to the n physical sensors being present at either of the two points. When both the points are covered by at least one physical sensor, the entire system matrix A is observed. There are two cases when A is not observed

1) all the physical sensors are located at the first point. This event occurs with a probability $(q_1)^n$; or
2) all the physical sensors are located at the second point. This occurs with a probability $(q_2)^n$ or $(1 - q_1)^n$.

Thus the conditions for covariance of the error to diverge are for any one of the following inequalities to be true,

$$\begin{aligned} (q_1)^n|\lambda_2|^2 &> 1 \\ (1-q_1)^n|\lambda_1|^2 &> 1. \end{aligned}$$

Adding the inequalities completes the proof. $\blacksquare$
Combining the proof technique of Propositions 2 and 3 immediately leads to the generalization stated below.

Proposition 4: The sensor coverage problem when N physical points are to be patrolled using n sensors, but otherwise the same assumptions hold as above, cannot be solved if for any $1 \leq m \leq N - 1$,

$$\frac{\binom{N-1}{m}}{|\lambda_1|^{\frac{2}{n}}} + \frac{\binom{N-2}{m-1}}{|\lambda_2|^{\frac{2}{n}}} + \cdots + \frac{\binom{N-1-m}{0}}{|\lambda_{m+1}|^{\frac{2}{n}}} < \binom{N-1}{m-1}. \qquad (7)$$

B. Coupled processes

Let the process at the i-th point evolve as

$$x_i^{k+1} = A_i x_i^k + \sum_{j \neq i} A_{i,j} x_j^k + w_i^k.$$

As long as all $A_{i,j}$'s are not zero, it is possible to obtain information about x_j^k even though the sensor is at point i. Moreover, the eigenvalues of the unobservable modes when considering two physical sensors located at points i and j may have no relation to the eigenvalues when the sensors are at points i and k. Thus the analysis is more involved in this case. We will now consider virtual sensors formed by sets of m physical points and say that a virtual sensor is used if none of the physical sensors are located outside the specified m points. Denote the set of all such virtual sensors by $\mathcal{S}_m$. For any member M of this set, consider the sensing matrix C_M formed by stacking all the C_i's such that i belongs to the set of m points corresponding to M. Denote by α_M the eigenvalue with the maximum magnitude of the unobservable part of A when the pair (A, C_M) is put in the observer canonical form.

Proposition 5: The sensor coverage problem for N physical points and n sensors with the above assumptions cannot be satisfied if for any m such that $1 \leq m \leq N - 1$,

$$\sum_{M \in \mathcal{S}_m} \frac{1}{|\alpha_M|^{\frac{2}{n}}} < \binom{N-1}{m-1}.$$

Proof: Proof follows exactly along the lines of that of proposition 4 and is omitted. $\blacksquare$

IV. MOTION GOVERNED BY A MARKOV CHAIN

We now consider the more general case where each sensor decides its position at time step $k + 1$ according to its position at time step k by using a transition probability matrix Q. This can also model the case when data loss is occurring according to a Markov chain. We still assume that the various sensors act independently. We will use the following result from [1].

Proposition 6: (Theorem 5 from [1]) Let a process with evolution matrix A be observed at each time step by one among n sensors with sensing matrices C_1, C_2, $\cdots$, C_n such that the i-th sensor is chosen at any time step according to a Markov chain with transition probability matrix $Q = [q_{ij}]$. Denote $\lambda_{max}(A_i)$ as the eigenvalue with the maximum magnitude of the unobservable part of A when the pair (A, C_i) is put in the observable canonical form. Then a sufficient condition for the expected estimate error covariance to diverge from at least one initial value is

$$q_{ii}|\lambda_{max}(A_i)|^2 > 1,$$

for any $i = 1, 2, \cdots, n$.

We assume that the Markov chain is positive recurrent and irreducible, thus there exists a unique stationary distribution. Let π_i denote the stationary probability of being in the i-th state. The general result statement along the lines of proposition 5 is now presented. Define α_M as before.

Proposition 7: The sensor coverage problem for N physical points and n sensors with the above assumptions cannot be satisfied if for any m such that $1 \leq m \leq N-1$, any of the following $\binom{N}{m}$ conditions are satisfied

$$\frac{1}{\sum_{j \in M} \pi_j} \left[\sum_{k \in M} \sum_{j \in M} q_{kj} \pi_k \right] > \frac{1}{|\alpha_M|^{\frac{2}{n}}}.$$

Proof: Proof is along the lines of proposition 5. The only trick is in the calculation of the probability q_{ii} for the i-th virtual sensor. For a Markov chain with transition probability matrix P and a set of states S, the probability that the state at time $k+1$ belongs to S given that the state at time k belonged to S is given by the expression

$$\frac{1}{\sum_{j \in S} \pi_j^k} \left[\sum_{t \in S} \sum_{j \in S} p_{tj} \pi_t^k \right],$$

where π_t^k is the probability of being in state t at time k. Since the Markov chain reaches a stationary distribution, $\pi_t^k \to \pi_t$ as k increases. Thus q_{ii}'s can be evaluated. ∎

Remarks:

1) One special case is when the sensors are chosen in an i.i.d. fashion. Thus $q_{kj} = q_j$ for all pairs (k, j) and the conditions in proposition 7 reduce to

$$\sum_{j \in M} q_j > \frac{1}{|\alpha_M|^{\frac{2}{n}}}.$$

Summing all $\binom{N}{m}$ inequalities obtained by the various choices of M yields the condition in proposition 5.

2) When the processes at various points are uncoupled, the terms α_M are expressible in terms of the eigenvalues with the maximum magnitude λ_i of the processes at the various points. Let the points be numbered such that $|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_N|$. Let $\bar{M}$ be the set of points $i \in \{1, 2, \cdots, N\}$ that are not contained in M. Then $\alpha_M = \lambda_i$, where $i = \min_{j \in \bar{M}}(j)$.

V. SENSOR SCHEDULE DESIGN

In this section, we look at designing the probabilities such that the sensor coverage is done in some optimal fashion. As a metric, we would consider the trace of the steady-state value of the estimate error covariance P^k as defined in (2). Since the trajectories we consider are stochastic, the error covariance is a random quantity. We will adopt the trace of the steady state expected error covariance as the metric to be minimized. We begin by characterizing this quantity.

Consider a process with evolution matrix A being driven by white noise with mean 0 and covariance matrix Q. Let the process be observed at each time step by one among n sensors, with the i-th sensor characterized by the sensing matrix C_i and measurement noise with mean 0 and covariance R_i. Further let all the noises be independent of each other. Let the choice of sensor be done randomly at each time step. Then the minimum mean square estimator is a Kalman filter with a time-varying sensor. If the i-th sensor

was chosen at time step k, the error covariance at time step $k + 1$, denoted by P^{k+1}, is given by $P^{k+1} = f_i \left(P^k \right)$, where $f_i \left(. \right)$ is the Riccati operator defined as

$$f_i \left(X \right) = AXA^T + Q - AXC_i^T \left(C_i X C_i^T + R_i \right)^{-1} C_i X A^T.$$

Because of the random schedule of the sensors, the error covariance is a stochastic quantity. We consider its mean value which evolves as

$$E \left[P^{k+1} \right] = E \left[f_i \left(P^k \right) \right]. \tag{8}$$

However, exact evaluation of the above quantity seems intractable. Instead, we consider an upper bound on the expected error covariance, using the results from [1].

A. *Sensor motion being governed by i.i.d. choices*

Denote the probability of sensor m being at point i at time step k by q_i. We start by obtaining an upper bound for the mean error covariance in (8) using Theorem 3 of [1].

Proposition 8: (Theorem 3 of [1]) The mean error covariance $E \left[P^k \right]$ evolving as in (8) is upper bounded by Δ^k, where Δ^k evolves according to the recursion

$$\Delta^{k+1} = A\Delta^k A^T + Q$$
$$- \sum_{i=1}^{n} q_i A\Delta^k C_i^T \left(C_i \Delta^k C_i^T + R_i \right)^{-1} C_i \Delta^k A^T, \tag{9}$$

with $\Delta^0 = P^0$. Further if there exist matrices $K_1, K_2, \cdots, K_n$ and a positive definite matrix P such that

$$P > Q + \sum_{i=1}^{n} q_i \left((A + K_i C_i) P (A + K_i C_i)^T + K_i R_i K_i^T \right),$$

then the above recursion converges for all initial conditions $P^0 > 0$ and the limit X is the unique positive semi-definite solution of the equation

$$X = AXA^T + Q$$
$$- \sum_{i=1}^{n} q_i AXC_i^T \left(C_i X C_i^T + R_i \right)^{-1} C_i X A^T. \tag{10}$$

X is an upper bound on the steady state expected error covariance. In the sequel, we will adopt trace(X) as the metric to be minimized as an approximation to minimizing the expected error covariance itself. Divergence of the upper bound is a necessary condition for the divergence of the expected error covariance; hence the design can be expected to be conservative in this sense. Let there be p virtual sensors present. The design problem is

$$\min_{q_i} \ \text{trace}(X) \tag{11}$$

$$\text{s.t. } X = \sum_{i=1}^{p} q_i f_i \left(X \right)$$

$$\sum q_i = 1 \qquad 0 \leq q_i \leq 1 \qquad X \geq 0.$$

For a problem of small size, a brute force search suffices to find the optimal probabilities. However, we can also use

a gradient descent algorithm to solve the problem. For ease of notation, we adopt the following notation. Define

$$g_q(X) = \sum_{i=1}^{p} q_i f_i(X),$$

where q is the vector of q_i's. The cost function of our problem is $\mathrm{trace}(X)$ or $\mathrm{trace}(g_q(X))$. From now on, we will refer to any vector γ whose components γ_i's are non-negative and sum to 1 as a valid probability vector. The algorithm proceeds as follows:

1) Initialize at step $k = 1$, with an arbitrary valid probability vector γ^1 and calculate the positive semi-definite matrix X^1 that satisfies $X^1 = g_{\gamma^1}(X^1)$.
2) At every step k, do the following :
   - Calculate $\gamma_{\min}$ as a valid probability vector that minimizes $\mathrm{trace}\left(g_\gamma(X^k)\right)$.
   - Calculate $\bar{\gamma}^k = \gamma^k + \delta\left(\gamma_{\min} - \gamma^k\right)$, where δ is the step size parameter between 0 and 1.
   - Obtain γ^{k+1} by projecting $\bar{\gamma}^k$ on the set of valid probability vectors.
   - Calculate $X^{k+1} = g_{\gamma^{k+1}}(X^k)$.
   - If $\gamma^k = \gamma^{k+1}$ (within a prescribed tolerance) then break else repeat the loop.
3) Output γ^{k+1} as the minimizing vector and $\mathrm{trace}\left(X^{k+1}\right)$ as the minimum cost function.

$\gamma_{\min}$ is obtained through an optimization of the form

$$\arg\min_{\gamma}\ \mathrm{trace}\left(\sum_{i=1}^{p} \gamma_i f_i(X)\right)$$
$$\sum \gamma_i = 1 \qquad 0 \le \gamma_i \le 1,$$

where X is a given positive semi-definite matrix. This is a linear program and can be solved efficiently. The projection step in the algorithm is required since $\bar{\gamma}^k$ may have individual components that are negative or greater than 1, even though they sum up to 1. Even though the optimal projection would be to find out the vector that is a valid probability vector and minimizes the Euclidean distance from the original vector, in practice, heuristics such as setting the negative components to 0 and redistributing their weight to all the other components seem to work well. Also since we have not proven anything about the convexity of the problem, the minima may not be global. Finally note that we can consider additional constraints placed on the probability vector. As an example if packets are being dropped with a probability λ, then the other $p - 1$ sensors are being used with a probability $q_1(1 - \lambda)$, $q_2(1 - \lambda)$ and so on, where the q_i's still sum to 1.

B. Sensor motion being governed by Markovian choices

Denote the probability of sensor m being at point i at time step k by π_i^k and the probability of the sensor moving from point i to point j as q_{ij}. We again obtain an upper bound for the mean error covariance using a result from [1].

Proposition 9: (Theorem 5 of [1]) The mean error co-variance P^k evolving as in (8) when sensors are chosen in a Markovian fashion is upper bounded by Δ^k, where Δ^k evolves according to the recursion

$$\Delta^{k+1} = \sum_{j=1}^{n} \pi_j^k \Delta_j^{k+1} \tag{12}$$
$$\pi_j^k \Delta_j^{k+1} = \sum_{i=1}^{n} f_j\left(\Delta_i^k\right) q_{ij} \pi_i^k,$$

with the initial conditions $\Delta_i^0 = P^0$. Further, assume that the Markov chain transition probability matrix Q is such that the states reach a stationary probability distribution with the probability of the j-th sensor being used as $\pi_j > 0$. If there exist n positive definite matrices $X_1, X_2, \cdots, X_n$ and n^2 matrices $K_{11}, K_{12}, \cdots, K_{1n}, K_{2,1}, \cdots, K_{n,n}$ such that

$$\pi_j X_j > \sum_{i=1}^{n}\Big((A + K_{ij}C_j)X_i(A + K_{ij}C_j)^* \\ + Q + K_{ij}R_j K_{ij}^*\Big)q_{ij}\pi_i,$$

then (12) converges for all initial conditions $X_0 > 0$ and the limit $\bar{X}_j$ is the unique positive semi-definite solution of

$$\pi_j X_j = \sum_{i=1}^{n} f_j(X_i) q_{ij} \pi_i. \tag{13}$$

The upper bound for the error covariance is given by

$$\bar{X} = \sum_{j=1}^{n} \pi_j \bar{X}_j.$$

Similar to the i.i.d. case then, if we assume p virtual sensors to be present, we can pose the following optimization problem to solve for the transition probability matrix.

$$\min_{q_{ij}}\ \mathrm{trace}(X) \tag{14}$$

s.t. $\quad X = \sum_{j=1}^{p} \pi_j X_j \qquad \pi_j X_j = \sum_{i=1}^{p} f_j(X_i) q_{ij} \pi_i$

$$\sum_j q_{ij} = 1 \qquad 0 \le q_{ij} \le 1$$

$$X_j \ge 0 \qquad \pi_i = \sum_{j=1}^{p} q_{ji}\pi_j.$$

This can again be solved by an algorithm similar to the one proposed above for the i.i.d. case. The step of finding the minimizing q_{ij}'s remains a linear program.

VI. EXAMPLE

We now illustrate our results with the help of some simple examples. As the first example, consider a grid of 6 points such that the value at each point represents a flow traveling from the first node towards the sixth node. Thus the dynamic equation at points 2 through 6 is given by

$$x_i^{k+1} = x_{i-1}^k + w_i^k,$$

while for point 1 it is given by $x_1^{k+1} = w_1^k$. We assume that the covariance matrix of the noise w_i^k is $R_i = 0.5$. Consider only one sensor of the form (1) with the matrix $R_m = 0.1$ that chooses its position independently from one time step to the next. There are 6 virtual sensors with the sensing matrix of the i-th sensor, C_i, being a row vector with all zeros except a 1 at the i-th place. The noise covariance matrix for all the sensors is R_m. The process matrix A is a 6×6 identity matrix. Let q_i denote the probability of its being at the i-th point. Naively, we may assume that the optimal probability distribution would be either to spend a lot of time at the source, i.e., the first node or equally among all the nodes. However if we optimize the probability distribution (using a brute-force search), it turns out that for the optimal distribution, $q_3 = 1$. The optimal cost is 6.28. If the sensor spends all its time at the source node, the cost is 8.42 while for a strategy of spending time with the same probability at all the points it is 8.23. If we use a greedy strategy in which the sensor moves to minimize the cost at every time-step, it leads to the sensor spending all its time at the fourth point, leading to a cost of 6.69. Thus our strategy performs better than heuristic or greedy algorithms.

As our second example we choose a ring network of 4 agents in which each agent is trying to calculate the average of the values of all the agents. Thus for the i-th agent

$$x_i^{k+1} = x_i^k - h(2x_i^k - x_{i+1}^k - x_{i-1}^k) + w_i^k,$$

where the addition in the agent number i is done modulo 4 and h is a positive constant. For a small enough value of h, the agents will calculate the average if no noise were present. We assume the noises w_i^k to be independent of each other and with variance $R_1 = R_4 = 1$ and $R_2 = R_3 = 0.8$. We again assume there is only one sensor that is choosing its position in an i.i.d. fashion. We will consider $h = 0.2$. We use the gradient descent algorithm with initial probability distribution $q_1 = q_2 = 0.5$ and a step size of 0.01. On optimizing the distribution, the values turn out to $q_1 = q_4 = 0.3$ and $q_2 = q_3 = 0.2$ with an optimal cost (upper bound) of 5.81. Indeed if we run 10000 random runs of the system generating sensor switching with this probability, we obtain a mean steady state error covariance trace of 5.8. Hence the upper bound is pretty tight at least in this example.

We can also impose the restriction that the sensor can only move from one physical point to its neighbors. Thus the sensor positions are chosen according to a Markov chain. Let us assume no packet losses for simplicity. Because of the symmetry of the system, we look for transition probability matrices of the form

$$\begin{bmatrix} 1 - 2\lambda_1 & \lambda_1 & 0\lambda_1 & \\ \lambda_2 & 1 - 2\lambda_2 & \lambda_2 & 0 \\ 0 & \lambda_2 & 1 - 2\lambda_2 & \lambda_2 \\ \lambda_1 & 0 & \lambda_1 & 1 - 2\lambda_1 \end{bmatrix}.$$

Then the optimal parameters turn out to be $\lambda_1 = \lambda_2 = 0.5$. As we vary the value of h, the system becomes more or less

stable. Thus to keep the error covariance bounded, we need different number of sensors. Figure 1 shows a bound on the number of sensors required, as predicted by Proposition 5.
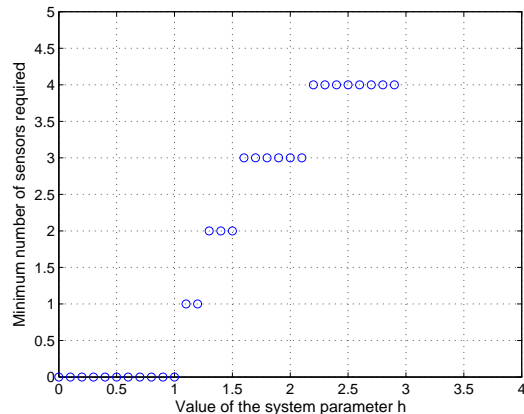

Fig. 1.   Lower bound on the number of sensors required.

VII. CONCLUSIONS AND FUTURE WORK

We considered the problem of monitoring a geographical region using a small number of mobile sensors. We proposed a stochastic strategy to determine the trajectories of the sensors. We identified conditions under which no probability distribution can lead to a bounded expected error covariance and presented an algorithm to generate optimal probability distributions for choosing the sensor trajectories.

The work can be extended in multiple ways. In this work, we assumed that the motion for the n sensors was independent of each other. If the sensors are able to plan their motion together, they probably can be represented as one sensor with a bigger transition probability matrix. We are currently working on the details of this case. Another possible avenue for future work is to determine a relaxation such that the optimization problem stated in the paper can be solved efficiently (possibly as a convex program).

REFERENCES

[1] 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, February 2006, Pages: 251-260.
[2] Q. Du, V. Faber and M. Gunzburger, "Centroidal Voronoi Tesselations: Applications and Algorithms", *SIAM Review*, vol. 41, No. 4, 1999, page(s):637-676.
[3] J. Cortes, S. Martinez, T. Karatas and F. Bullo, "Coverage Control for Mobile Sensing Networks", *IEEE Transactions on Robotics and Automation*, vol. 20, No. 2, 2004, page(s): 243-255.
[4] H. Mori, "Active Sensing in Vision-based Stereotyped Motion", *Proc. IEEE Int. Workshop on Intelligent Robots and Systems*, Vol. 1, 1990, page(s): 167-174.
[5] T. Mukai and I. Ishikawa, "An Active Sensing Method using Estimated Errors for Multisensor Fusion Systems", *IEEE Transactions on Industrial Electronics*, Vol. 43, 1996, page(s): 380-386.
[6] T. H. Chung, V. Gupta, J. W. Burdick and R. M. Murray, "On a Decentralized Active Sensing Strategy using Mobile Sensor Platforms in a Network", *Proc. IEEE Conf. on Decision and Control*, Vol. 2, Page(s): 1914-1919, 2004.