

Sensor Scheduling for Multiple Parameters Estimation Under Energy Constraint

Yi Wang, Mingyan Liu and Demosthenis Teneketzis
Department of Electrical Engineering and Computer Science
University of Michigan, Ann Arbor, MI
{yiws,mingyan,teneket}@eecs.umich.edu

ABSTRACT

We consider a sensor scheduling problem for estimating Gaussian random variables under an energy constraint. The sensors are described by a linear observation model, and the observation noise is Gaussian. We formulate this problem as a stochastic sequential decision problem. Due to the Gaussian assumption and the linear observation model, the stochastic sequential decision problem is equivalent to a deterministic one. We present a greedy algorithm for this problem, and discover conditions sufficient to guarantee the optimality of the greedy algorithm. Furthermore, we present two special cases of the original scheduling problem where the greedy algorithm is optimal under weaker conditions. We illustrate our result through numerical examples.

I. INTRODUCTION

Advances in integrated sensing and wireless technologies have enabled a wide range of emerging applications, from environmental monitoring to intrusion detection, to robotic exploration, etc. In particular, unattended ground sensors (UGS) have been increasingly used to enhance situational awareness for surveillance and monitoring type of applications.

In this paper we focus on the use of sensors for the purpose of parameter estimation. Specifically, we consider the following scheduling problem. Multiple sensors are sequentially activated by a central controller to take a measurement of one of many parameters, and then transmit the observation data back to the the controller. The latter combines successive measurement data to form an estimate for each parameter. A single parameter may be measured multiple times over time. Each activation incurs a cost (e.g., sensing and communication costs) which may be both sensor and parameter-dependent. This process continues until a certain criterion is satisfied, e.g., when the total estimation error is sufficiently small, when the time period of interest has elapsed, etc. Assuming that sensors may be of different quality (i.e. they may have different signal to noise ratios) and the activation of different sensors may incur different costs, our sensor-scheduling problem is to determine the sequence of

sensors to be activated and the corresponding sequence of parameters to be measured so as to minimize the sum of the total terminal parameter estimation errors and the total sensor activation cost.

In this paper we restrict attention to the case of N stationary scalar parameters, modelled by Gaussian random variables with known mean and variance, measured by M sensors, each described by a linear Gaussian observation model. Without loss of generality, we assume that each sensor can only be used once. This is because multiple uses of the same sensor can be effectively replaced by multiple identical sensors, each with a single use. We formulate the above sensor scheduling problem as a stochastic sequential decision problem. Because of the Gaussian assumption and the linearity of observations, this stochastic sequential decision problem is equivalent to a deterministic one.

Sequential allocation problems have been extensively studied in the literature (see [1]). In general, it is difficult to explicitly determine optimal strategies or even qualitative properties of optimal strategies for sequential allocation problems. The multi-armed bandit problem and its variants (see e.g., [6] and [5]) are one class of sequential allocation problems where the optimal solution has been explicitly determined. In [?](refer technical report) we compare our problem with the multi-armed bandit problem and some of its variants. We show that our problem does not conceptually belong to the class of multi-armed bandits. It appears difficult to determine the nature of an optimal solution for the general problem under consideration. Therefore, to obtain some insight into the nature of the problem, we consider a greedy algorithm, and derive conditions sufficient to guarantee the optimality of this algorithm. We then present two special cases of the general problem under consideration for which the greedy algorithm results in an optimal strategy under conditions weaker than the sufficient conditions mentioned above. Finally we illustrate the nature of our results through a number of numerical examples.

Sensor scheduling problems associated with stationary parameter estimation were also investigated in [3] and

[2]. Our results are different from those of [3] and [2] primarily because the observation model and performance criteria in [3] and [2] are different from ours.

The rest of the paper is organized as follows. In Section II we state the optimization problem and introduce several preliminaries. In Section III we provide the sufficient conditions for the greedy algorithm to be optimal. Special cases are presented in Section IV, and numerical examples are analyzed in Section V. Section VI concludes this paper. Due to the space limitation, all proofs are omitted; they can be founded in [?].

II. PROBLEM FORMULATION

In this section, we formulate the problem of estimating multiple stationary parameters with multiple sensors described in the previous section, and present a number of preliminaries.

A. The Measurement Model and Problem Formulation

Consider a set Ω_p of stationary scalar parameters, indexed by $\{1, 2, \dots, N\}$ that have to be estimated. Parameter i is modeled by a Gaussian random variable, denoted by X_i , with mean $\mu_i(0)$ and variance $\sigma_i(0)$. There is a set Ω_s of sensors, indexed by $\{1, 2, \dots, M\}$, which we use to measure the parameters. The measurement of parameter i taken by sensor j is given by

$$Z_{i,j} = H_{i,j}X_i + V_{i,j}, \quad (1)$$

where $Z_{i,j}$ is the observation of parameter i by sensor j , $H_{i,j}$ is a known gain, and $V_{i,j}$ is a Gaussian random variable with $E(V_{i,j}) = 0$, $Var(V_{i,j}) = R_{i,j}$. A non-negative measurement cost $c_{i,j}$ is incurred by activating and using sensor j to measure parameter i .

The available sensors are activated one at a time to take a measurement of a specific parameter, upon request from a control center. The measurement data is then transmitted back to the control center, where the estimate of that parameter and the total accumulated cost are updated. The control center then decides whether to activate another sensor from the set of remaining available sensors, or to terminate the process. This continues until either all M sensors are used, or until the time period of interest T has elapsed, or until the control center decides to terminate the process. For simplicity and without loss of generality, we redefine $T = \min\{M, T\}$, implying that at most T sensors/parameters can be scheduled.

Consequently, the decision/control action at each time instant t is a random vector $U_t := (U_{1,t}, U_{2,t})$, taking values in $\Omega_p \times S^t \cup \{\emptyset, \emptyset\}$, where S^t is the set of sensors available at t , and $U_t = (\emptyset, \emptyset)$ means that no measurement is taken at t . A measurement policy is

defined by

$$\gamma := (\gamma_1, \gamma_2, \dots, \gamma_T),$$

where γ_t is such that

$$\begin{aligned} U_t^\gamma &= (U_{1,t}^\gamma, U_{2,t}^\gamma) \\ &= \gamma_t(\sigma_1(0), \sigma_2(0), \dots, \sigma_N(0), \Omega_p, Z^{\gamma,t-1}, U^{\gamma,t-1}), \end{aligned}$$

where

$$\begin{aligned} Z^{\gamma,t-1} &:= (Z_1^\gamma, Z_2^\gamma, \dots, Z_{t-1}^\gamma), \\ U^{\gamma,t-1} &:= (U_1^\gamma, U_2^\gamma, \dots, U_{t-1}^\gamma), \end{aligned}$$

and the variable Z_t^γ denotes the measurement taken at time t . Since parameters are stationary, not taking a measurement at some time instant will leave the parameters and their estimates unchanged. Thus, without loss of optimality, we can restrict attention to measurement strategies with the following property.

Property 1: For $\forall t, t = 1, \dots, T-1$, if $U_t^\gamma = (\emptyset, \emptyset)$, then for $t' > t$, $U_{t'}^\gamma = (\emptyset, \emptyset)$.

Let Γ be the set of all admissible measurement policies that satisfy this property. The optimization problem is

Problem 1 (P1):

$$\begin{aligned} \min_{\gamma \in \Gamma} J(\gamma) &= \sum_{i=1}^N E \left\{ \left[X_i - \hat{X}_i^\gamma(T) \right]^2 \right\} + E \left\{ \sum_{t=1}^T c_{U_t^\gamma} \right\} \\ s.t. \quad &\begin{cases} \hat{X}_i^\gamma(T) = E[X_i | Z_{U_t^\gamma} \cdot 1(\{U_{1,t}^\gamma = i\}), t = 1, \dots, T] \\ U_{2,t}^\gamma \neq U_{2,t'}^\gamma \text{ if } t \neq t', t, t' = 1, \dots, \tau^\gamma, \end{cases} \end{aligned}$$

where $J(\gamma)$ is the cost of policy $\gamma \in \Gamma$, $\hat{X}_i^\gamma(T)$ is the terminal estimate of parameter i under strategy γ , and $1(A)$ is the indication function such that $1(A) = 1$ if A is true and 0 otherwise.

Denote by $Z_i^{\gamma,t}$ the observation data set for parameter i until time t under strategy γ , $t = 1, 2, \dots, T$. Since X_i is a Gaussian random variable,

$$\begin{aligned} E\{[X_i - \hat{X}_i^\gamma(T)]^2\} &= E \left\{ E \left\{ [X_i - E(X_i | Z_i^{\gamma,T})]^2 | Z_i^{\gamma,T} \right\} \right\} \\ &= E \left\{ [X_i - E(X_i | Z_i^{\gamma,T})]^2 \right\}, \end{aligned}$$

i.e. the error variance is independent of the observation data. Denote the variance of parameter i under strategy γ at time t as

$$\sigma_i^\gamma(t) := E \left\{ \left[X_i - E(X_i | Z_i^{\gamma,t}) \right]^2 \right\}, \quad i = 1, \dots, N.$$

Then at any time instant t , the variance of parameter i evolves as follows.

If at $t+1$, parameter i and sensor j are selected by γ , then

$$\sigma_i^\gamma(t+1) = \sigma_i^\gamma(t) - \frac{(\sigma_i^\gamma(t))^2 H_{i,j}^2}{\sigma_i^\gamma(t) H_{i,j}^2 + R_{i,j}}; \quad (2)$$

if at time $t+1$, parameter i is not selected by γ , then

$$\sigma_i^\gamma(t+1) = \sigma_i^\gamma(t). \quad (3)$$

(see [4])

Therefore problem $P1$ can be reformulated as a deterministic problem in the following way.

Denote a scheduling strategy by $\mathbf{g} := (P^{\mathbf{g}}, S^{\mathbf{g}})$, with

$$P^{\mathbf{g}} = \{a_1^{\mathbf{g}}, \dots, a_T^{\mathbf{g}}\}, \text{ and } S^{\mathbf{g}} = \{b_1^{\mathbf{g}}, \dots, b_T^{\mathbf{g}}\},$$

which indicates that under strategy \mathbf{g} , parameter $a_t^{\mathbf{g}}$ is measured by sensor $b_t^{\mathbf{g}}$ at time t , where $a_t^{\mathbf{g}} \in \Omega_p$, $b_t^{\mathbf{g}} \in \Omega_s \cup \{\emptyset\}$. If $b_t^{\mathbf{g}} = \emptyset$, no measurement takes place at time t and $c_{a_t^{\mathbf{g}}, b_t^{\mathbf{g}}} = 0$. Similarly to property 1, we can restrict attention to measurement strategies with the following property.

Property 2: For $\forall t, t = 1, \dots, T-1$, if $b_t^{\mathbf{g}} = \emptyset$, then $b_{t'}^{\mathbf{g}} = \emptyset$, for $\forall t' > t$.

Let \mathbf{G} be the set of all admissible measurement policies with the property 2. The optimization problem is

Problem 2 (P2):

$$\begin{aligned} \min_{\mathbf{g} \in \mathbf{G}} \quad & J(\mathbf{g}) = \sum_{i=1}^N \sigma_i^{\mathbf{g}}(\tau^{\mathbf{g}}) + \sum_{t=1}^{\tau^{\mathbf{g}}} c_{a_t^{\mathbf{g}}, b_t^{\mathbf{g}}} \\ \text{s.t.} \quad & \begin{cases} a_t^{\mathbf{g}} \in \Omega_p, \text{ and } b_t^{\mathbf{g}} \in \Omega_s \cup \{\emptyset\}, \\ b_t^{\mathbf{g}} \neq b_{t'}^{\mathbf{g}} \text{ if } t \neq t', \end{cases} \end{aligned}$$

where $\tau^{\mathbf{g}}$ denotes the number of measurements taken under policy \mathbf{g} .

B. Preliminaries

The following definitions characterize the ‘‘goodness’’ of a sensor in terms of its quality of measurement.

Definition 1: The *index* of sensor j for parameter i is given by $I_{i,j} = \frac{H_{i,j}^2}{R_{i,j}}$.

An index can be viewed as the signal-to-noise ratio (SNR) of sensor j when measuring parameter i . This quantity reflects the accuracy of the measurement; the higher the index/SNR, the more accurate the measurement.

Lemma 1: Assume sensor set A is used to measure parameter i with initial variance $\sigma_i(0)$ and parameter i 's post-measurement variance is $\sigma_i(A)$. Then we have

$$\sigma_i(A) = \frac{\sigma_i(0)}{\sigma_i(0)\hat{I}_{i,A} + 1} \quad (4)$$

where $\hat{I}_{i,A} = \sum_{j \in A} I_{i,j}$. Furthermore, $\sigma_i(A)$ is an increasing function of $\sigma_i(0)$ and a decreasing function of A , i.e. if $A_1 \subset A_2$, then $\sigma_i(A_1) > \sigma_i(A_2)$.

The *variance reduction* of parameter i with initial variance $\sigma_i(0)$ by using sensor set A , denoted by $R_i(\sigma_i(0), A)$, and given by

$$R_i(\sigma_i(0), A) := \sigma_i(0) - \sigma_i(A) = \frac{\sigma_i^2(0)\hat{I}_{i,A}}{\sigma_i(0)\hat{I}_{i,A} + 1}. \quad (5)$$

Under any policy \mathbf{g} , we can re-write the objective function $J(\mathbf{g})$ as follows.

$$\begin{aligned} J(\mathbf{g}) &= \sum_{t=1}^{\tau^{\mathbf{g}}} \left\{ c_{a_t^{\mathbf{g}}, b_t^{\mathbf{g}}} - \left[\sigma_{a_t^{\mathbf{g}}}^{\mathbf{g}}(t-1) - \sigma_{a_t^{\mathbf{g}}}^{\mathbf{g}}(t) \right] \right\} + \sum_{i=1}^N \sigma_i(0) \\ &= \sum_{t=1}^{\tau^{\mathbf{g}}} P_{a_t^{\mathbf{g}}}(\sigma_{a_t^{\mathbf{g}}}^{\mathbf{g}}(t-1), b_t^{\mathbf{g}}) + \sum_{i=1}^N \sigma_i(0), \end{aligned} \quad (6)$$

where $P_i(\sigma, j)$ is given by:

$$P_i(\sigma, j) = c_{i,j} - R_i(\sigma, j) = c_{i,j} - \frac{\sigma^2 I_{i,j}}{\sigma I_{i,j} + 1}. \quad (7)$$

The quantity $P_i(\sigma, j)$ is referred to as the *step cost* of using sensor j to measure parameter i , when the variance of parameter i before the measurement is σ . Thus the total cost to be minimized is the sum of initial variance of each parameter and the step costs.

Definition 2: The *threshold* of a sensor j for parameter i is given by $TH_{i,j} = \frac{1}{2} \cdot (c_{i,j} + \sqrt{c_{i,j}^2 + 4 \cdot c_{i,j}/I_{i,j}})$.

With this definition, we have that

$$\text{when } \sigma = TH_{i,j}, \quad P_i(\sigma, j) = c_{i,j} - \frac{\sigma^2 I_{i,j}}{\sigma I_{i,j} + 1} = 0; \quad (8)$$

$$\text{when } \sigma > TH_{i,j}, \quad P_i(\sigma, j) = c_{i,j} - \frac{\sigma^2 I_{i,j}}{\sigma I_{i,j} + 1} < 0. \quad (9)$$

Therefore sensor j 's threshold for parameter i can be viewed as the break-even point in parameter i 's variance. That is, using sensor j to take a measurement of parameter i , whose variance σ is equal to $TH_{i,j}$, results in zero step cost. If the current variance of parameter i exceeds the threshold $TH_{i,j}$, then using sensor j will result in a negative step cost (i.e. benefit), and vice versa.

For sensors with the same index, lower threshold is equivalent to smaller measurement cost; for sensors with the same measurement cost, lower threshold is equivalent to higher index. Therefore, the threshold combines measurement quality and measurement cost and reflects the overall ‘‘goodness’’ of a sensor: the lower the threshold of a sensor, the better the sensor's quality.

III. SUFFICIENT CONDITIONS FOR THE OPTIMALITY OF A GREEDY POLICY

We decompose the sensor-selection parameter-estimation sequential decision problem into two subproblems. The first one is to determine the order in which sensors should be used regardless of which parameter is measured. The second problem is to determine which parameter should be measured at each time instant given the order in which sensors are used. Such a decomposition is not always optimal. We present conditions that guarantee the optimality of the aforementioned decomposition. Specifically, we determine two conditions under which it is optimal

to use the sensors in non-increasing order of their indices (regardless of which parameter is measured). Having guaranteed the optimality of the proposed decomposition, we propose a greedy algorithm for the selection of parameters. We determine a condition sufficient to guarantee the optimality of the greedy algorithm.

A. The Optimal Sensor Sequence

Condition 1: The sensors can be ordered into a sequence s_1, s_2, \dots, s_M such that

$$I_{j,s_1} \geq I_{j,s_2} \geq \dots \geq I_{j,s_M}, \quad \forall j = 1, 2, \dots, N. \quad (10)$$

This condition says that if we order the sensors in non-increasing order of their quality for a particular parameter, then this order remains the same for all other parameters. For the rest of our discussion we will denote s_j as the j -th sensor in this ordered set.

Condition 2: For each parameter i , $TH_{i,s_1} \leq TH_{i,s_2} \leq \dots \leq TH_{i,s_M}$, where $s_i, i = 1, \dots, N$ is defined in Condition 1.

If Conditions 1 and 2 both hold, then they imply that the ordering of sensors with respect to their sensing quality is the same as their ordering when cost is also taken into account. Furthermore, both orderings are parameter invariant.

The next theorem shows that the optimal sequencing of sensors is according to non-increasing order of their indices.

Theorem 1: Under Conditions 1 and 2, assume that an optimal selection strategy is $\mathbf{g} = (P^{\mathbf{g}}, S^{\mathbf{g}})$, where $P^{\mathbf{g}} = \{p_1, p_2, \dots, p_{\tau^{\mathbf{g}}}\}, S^{\mathbf{g}} = \{b_1, b_2, \dots, b_{\tau^{\mathbf{g}}}\}$. Then for each parameter $i, \forall b_k \in S^{\mathbf{g}}, \forall a \in \Omega_s - S^{\mathbf{g}}$, we have $I_{i,b_k} \geq I_{i,a}$.

The intuition behind this theorem is that although using different sensors may incur different costs, so long as the costs are such that they do not change the relative quality of the sensors (represented by their indices), the best way to use the sensors is by non-increasing order of their indices.

The performance of an allocation strategy is completely determined by the set of sensors allocated to each parameter; the order in which the sensors are used for a parameter is irrelevant. Thus, strategies that result in the same association between sensors and parameters may be viewed as *equivalent strategies*. From Theorem 1, we conclude that for any optimal strategy, there exists one equivalent strategy, under which sensors are used in non-increasing order of their indices. Therefore, without loss of optimality, we will only consider strategies that use sensors in non-increasing order of their indices.

Parameter Selection Algorithm L:

```

1:  $t := 0$ 
2: while  $t < T$  do
3:    $k := \arg \min_{i=1, \dots, N} P_i(\sigma_i(t), s_{t+1})$ 
4:   if  $P_k(\sigma_k(t), s_{t+1}) < 0$  then
5:      $p_{t+1} := k$ 
6:      $\sigma_k(t+1) := \frac{\sigma_k(t)}{\sigma_k(t)I_{k,s_{t+1}} + 1}$ 
7:     for  $i := 1$  to  $M$  do
8:       if  $i \neq k$  then
9:          $\sigma_i(t+1) := \sigma_i(t)$ 
10:      end if
11:    end for
12:  else
13:    BREAK
14:  end if
15:   $t := t + 1$ 
16: end while
17: return  $\tau := t$  and  $P := \{p_1, \dots, p_{\tau}\}$ 

```

Fig. 1. A greedy algorithm to determine the parameter sequence.

Consequently, problem $P2$ is reduced to determining the stopping time τ and the parameter sequence corresponding to the sensor sequence $S = \{s_1, s_2, \dots, s_{\tau}\}$.

B. A Greedy Algorithm

We consider the parameter selection algorithm, given in Figure 1.

Under Conditions 1 and 2, this algorithm computes a sequence of parameters, P , by sequentially selecting a parameter that provides the minimum step cost (i.e. the maximum benefit) among all parameters. The algorithm terminates when the minimum step cost becomes non-negative, or the time horizon T is reached. The termination time is the stopping time τ . The parameter selection strategy resulting from this algorithm, combined with the given sensor sequence, is denoted by $\mathbf{g} := (P, S)$, where

$$P = \{p_1, \dots, p_{\tau^{\mathbf{g}}}\}, \text{ and } S = \{s_1, \dots, s_{\tau^{\mathbf{g}}}\}.$$

This algorithm is greedy in nature in that it always uses the best available sensor (in terms of its index), and for that sensor it always selects the parameter whose measurement provides the maximum gain (minimum step cost). In the next subsection, we investigate conditions under which this greedy scheduling algorithm is optimal for problem $P2$.

C. Optimality of Algorithm L

In this section, our objective is to determine conditions sufficient to guarantee the optimality of the greedy algorithm given in Figure 1.

Below is a list of notations used.

$\sigma_i(t)$: The variance of parameter i at time t . This variance depends on the initial variance $\sigma_i(0)$ and the set of sensors used for parameter i up until time t .

$\sigma_i(t, A)$: The variance of parameter i after sensor set A is used to measure parameter i after time t .

$R_i(\sigma_i(t), A)$: The variance reduction of parameter i resulting from the use of sensor set A with initial variance $\sigma_i(t)$, s.t. $R_i(\sigma_i(t), A) = \sigma_i(t) - \sigma_i(t, A)$.

$R_i(\sigma_i(t, A), B)$: The variance reduction of parameter i resulting from the use of sensor set B with initial variance $\sigma_i(t, A)$, s.t.

$$R_i(\sigma_i(t, A), B) = R_i(\sigma_i(t), A \cup B) - R_i(\sigma_i(t), A) \\ = \sigma_i(t, A) - \sigma_i(t, A \cup B). \quad (11)$$

Assume at some time instant i , the available sensor set is $\{s_{i+1}, \dots, s_M\}$. Then for any sensor subset $E \subseteq \{s_{i+1}, \dots, s_M\}$, we define the *advantage* of using sensor s_i to measure p_i at time i followed by E , denoted by $B_i(p_i, E)$, as follows:

$$B_i(p_i, E) := R_{p_i}(\sigma_{p_i}(i-1), \{s_i\} \cup E) \\ - R_{p_i}(\sigma_{p_i}(i-1), E) - c_{p_i, s_i}. \quad (12)$$

The advantage is essentially the additional variance reduction resulting from sensor s_i measuring parameter p_i , after it has been measured by sensor set E , minus the observation cost.

Because of (11), $B_i(p_i, E)$ can be rewritten as $B_i(p_i, E) = R_{p_i}(\sigma_{p_i}(i-1), \{s_i\}) - c_{p_i, s_i} + \Delta_{p_i}(E)$, (13) where

$$\Delta_{p_i}(E) := R_{p_i}(\sigma_{p_i}(i-1), \{s_i\}, E) - R_i(\sigma_{p_i}(i-1), E).$$

We have the following property for $\Delta_{p_i}(E)$.

Lemma 2: Consider the available sensor set $A = \{s_{i+1}, \dots, s_M\}$ after stage i of the sequential allocation process. Let $E_1 = \{s_{i+1}, s_{i+2}, \dots, s_k\}$, $E_2 = \{s_{i+1}, s_{i+2}, \dots, s_j\}$, where $j < k \leq M$. Then

$$\Delta_{p_i}(A) \leq \Delta_{p_i}(E_1) < \Delta_{p_i}(E_2) \leq 0. \quad (14)$$

(?)Based on Lemma 2 and (13), we can define upper bound $B_{u,i}(p_i)$ and lower bound $B_{l,i}(p_i)$ on the aforementioned advantage as follows:

$$B_{u,i}(p_i) := R_{p_i}(\sigma_{p_i}(i-1), \{s_i\}) - c_{p_i, s_i} \quad (15) \\ = R_{p_i}(\sigma_{p_i}(i-1), \{s_i\}) - c_{p_i, s_i} + \max_{E \subseteq A} \Delta_{p_i}(E) \\ \geq R_{p_i}(\sigma_{p_i}(i-1), \{s_i\}) - c_{p_i, s_i} + \Delta_{p_i}(E) \\ = B_i(p_i, E),$$

where the equality holds when $E = \emptyset$, $B_{u,i}(p_i)$ is the upper bound of $B_i(p_i, E)$, and

$$B_{l,i}(p_i) := R_{p_i}(\sigma_{p_i}(i-1), \{s_i\}) - c_{p_i, s_i} + \Delta_{p_i}(A) \quad (16) \\ = R_{p_i}(\sigma_{p_i}(i-1), \{s_i\}) - c_{p_i, s_i} + \min_{E \subseteq A} \Delta_{p_i}(E) \\ \leq R_{p_i}(\sigma_{p_i}(i-1), \{s_i\}) - c_{p_i, s_i} + \Delta_{p_i}(E) \\ = B_i(p_i, E)$$

where the equality holds when $E = A$ and $B_{u,i}(p_i)$ is the lower bound of $B_i(p_i, E)$.

Therefore, $B_{l,i}(p_i) \leq B_i(p_i, E) \leq B_{u,i}(p_i)$.

Note that $-B_{u,i}(p_i)$ is the same as the step cost $P_{p_i}(\sigma_{p_i}(i-1), s_i)$.

The use of the above upper and lower bounds allows us to obtain the following result.

Lemma 3: Consider any two strategies $\mathbf{g}_1 = (S_1, P_1)$ and $\mathbf{g}_2 = (S_2, P_2)$, with

$$S_1 = S_2 = \{s_1, s_2, \dots, s_t\}, \\ P_1 = \{p_1, \dots, p_{i-1}, p_i, p_{i+1}, \dots, p_t\}, \\ P_2 = \{p_1, \dots, p_{i-1}, p'_i, p_{i+1}, \dots, p_t\}.$$

If $B_{l,i}(p_i) > B_{u,i}(p'_i)$, then $J(\mathbf{g}_1) < J(\mathbf{g}_2)$.

The intuition behind this lemma is that regardless of which allocation strategy will follow after time i , using sensor s_i to measure parameter p_i at time i will result in better performance than using sensor s_i to measure parameter p'_i .

The result of Lemma 3 allows us to obtain the following condition (?) which, together with Conditions 1 and 2, are sufficient to guarantee the optimality of the greedy algorithm described in Figure 1.

Condition 3: At some time instant t , there exists some parameter p_t^* , such that for any parameter p'_t not equal to p_t^* , $B_{l,t}(p_t^*) \geq B_{u,t}(p'_t)$, where $B_{l,t}(p_t^*)$ and $B_{u,t}(p'_t)$ are defined in a manner similar to (16) and (15) respectively.

Note that if Condition 3 holds at time instant t , p_t^* is unique. Furthermore, since

$$B_{u,t}(p_t^*) \geq B_{l,t}(p_t^*) \geq B_{u,t}(p'_t),$$

and $-B_{u,t}(p_t^*)$ is equal to the step cost, p_t^* is the parameter, which will result in the smallest step cost, measured by sensor s_t .

Theorem 2: If Conditions 1 and 2 hold and Condition 3 is satisfied at each time instant $1 \leq t \leq \tau$, then Algorithm L results in an optimal strategy for Problem (P2).

IV. SPECIAL CASES

We present two special cases of the general formulation given in Section II. In the first case, there is only one parameter to be estimated, which means the second subproblem in the decomposition of $P2$ does not exist. In the second case, M sensors are identical, which means the first subproblem in the decomposition of $P2$ does not exist. For both cases, we show that the greedy strategy is optimal under conditions weaker than those in Section III.

A. 1 Parameter and M Different Sensors

Consider problem $P2$, when only one static parameter has to be estimated. Then the observation model for sensor j is

$$Z_j = H_j X + V_j \quad (17)$$

The cost of using sensor j is c_j .

In this case we only need to determine which sensors should be used to measure the parameter. Thus, the second subproblem of the decomposition in Section III does not exist. Furthermore, Condition 1 is satisfied automatically. Then under Conditions 1 and 2, it is optimal to use the sensors according to non-increasing order of their indices by Theorem 1. Note that if the observation cost for every sensor is equal, i.e. $c_j = c$, $\forall j = 1, \dots, M$, Condition 2 is equivalent to Condition 1. In this situation, it is optimal to use the sensors according to non-increasing order of their indices without any constraint.

B. N Parameters and M Identical Sensors with Sensor-Independent Observation Model

Consider problem $P2$ in the case where the M sensors are identical. Then the observation model for parameter i is sensor-independent, that is,

$$Z_i = H X_i + V, \text{ for all sensors.} \quad (18)$$

The cost of measuring parameter i by any sensor j is c_i .

Since the sensors are identical, Conditions 1 and 2 are satisfied automatically. Therefore, in this case we only concern the second subproblem of the decomposition in Section III. Thus, we can view the M identical sensors as one processor which can be used at most M times, and the N different parameters as N independent machines. The state of every machine/parameter is its variance. At every time instant t , we must select one machine/parameter a_t to process/ estimate. The variance of machine/parameter a_t is updated and all the other machines'/parameters' state/variance is frozen. The reward at each time instant t is the variance reduction of parameter a_t minus the observation cost c_{a_t} . Viewed this way, problem $P2$ is a finite horizon multi-armed bandit problem with discount factor equal to one.

For finite horizon multi-armed bandit problems, the Gittins Index rule(see [1]) is not generally optimal. However, in the problem under consideration, the reward sequence for each machine/parameter is deterministic and non-decreasing with time. Thus, for each machine/parameter, the Gittins Index is always achieved at $\tau = 1$, which coincides with the one-step look-ahead policy resulting from Algorithm L described in Section

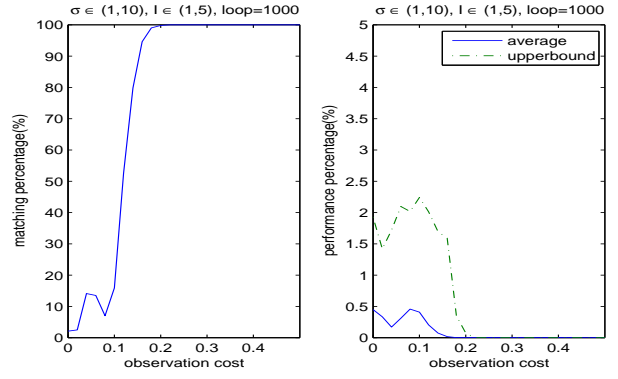


Fig. 2. Performance of Greedy Algorithm.

III. Consequently, the Gittins Index rule is optimal for this special case.

V. NUMERICAL EXAMPLES

We illustrate the performance of Algorithm L with numerical experiments(examples).

We denote by \mathbf{g}_L the “greedy strategy” corresponding to Algorithm L , and by \mathbf{g}_o the optimal strategy. We define the **performance deviation (PD)** of strategy \mathbf{g} as

$$PD(\mathbf{g}) := \frac{J(\mathbf{g}) - J(\mathbf{g}_o)}{J(\mathbf{g}_o)}. \quad (19)$$

The setup of the numerical experiment is as follows. There are 7 sensors and 3 parameters. The observation cost is a constant for all the sensors and parameters; 51 observation costs are incremented from 0 to 0.5 with incremental size 0.01. For each cost selection, we run the experiment 1000 times with 7 indices, each chosen according to a uniform distribution on (1,5) and 3 initial variances, each chosen according to a uniform distribution on (1,10). We adopt the following three performance criteria.

- 1) Matching Rate: $\frac{\text{the number that } \mathbf{g}_L = \mathbf{g}_o}{1000}$;
- 2) Average Deviation: $\frac{\sum PD(\mathbf{g}_L)}{1000}$;
- 3) Maximum Deviation: $\max PD(\mathbf{g}_L)$.

The results are shown in Figure 2. When the observation cost is sufficient large, strategy \mathbf{g}_L is always optimal. This is consistent with the intuition. When the observation cost is large, each parameter can be measured at most once. In this case one can show that using sensor with the largest index to measure the parameter with the largest variance at present is an optimal strategy. When strategy \mathbf{g}_L is not optimal, the average deviation and the maximum deviation are always below 5%. We observe very similar results when the experiment is repeated with different values of sensors’ indices and parameters’ initial variance.

From the numerical experiments, we can conclude that Algorithm L is a very good approximating algorithm,

which can be optimal when the observation cost is sufficient large.

VI. CONCLUSION

We considered a sensor scheduling problem for multiple parameter estimation under an energy constraint. We decompose the sequential decision problem into two subproblems. The first one is to determine the sequence of the sensors to be used, which is independent of the parameter selection, and the second one is to determine the sequence of parameters to be measured for a given sensor sequence. We identified conditions sufficient to guarantee that a greedy policy, defined by Algorithm *L*, is optimal for the problem under consideration. The numerical examples we considered indicate that for large values of the measurement cost, the greedy policy performs well. We presented an explanation as to why such behavior of the greedy policy should be expected for large measurement cost.

REFERENCES

- [1] J. C. Gittins. Bandit process and dynamic allocation indices. *Journal of the Royal Statistical Society. Series B (Methodological)*, 41:pp. 148–177, 1979.
- [2] G. P. H. Wang, K. Yao and D. Estrin. Entropy-based sensor selection heuristic for target localization. In *Proceedings of The Third International Symposium on IPSN*, pages pp. 36–45.
- [3] V. Isler and R. Bajcsy. The sensor selection problem for bounded uncertainty sensing models. In *Proceedings of The Fourth International Symposium on IPSN*, pages pp. 151–158.
- [4] P. R. Kumar and P. Varaiya. *Stochastic Systems: Estimation, Identification and Adaptive Control*. Prentice Hall, 1986.
- [5] J. C. W. Pravin P. Varaiya and C. Buyukkoc. Extensions of the multiarmed bandit problem: The discounted case. *IEEE Transactions on Automatic Control*, 30:pp. 426–439, 1985.
- [6] P. Whittle. Multi-armed bandits and the gittins index. *Journal of the Royal Statistical Society. Series B (Methodological)*, 42:pp. 143–149, 1980.