

# Asymptotically Efficient Adaptive Allocation Rules for the Multiarmed Bandit Problem with Switching Cost

RAJEEV AGRAWAL, MANJUNATH V. HEGDE, MEMBER, IEEE, AND DEMOSTHENIS TENEKETZIS

**Abstract**—We consider multiarmed bandit problems with switching cost, define uniformly good allocation rules, and restrict attention to such rules. We present a lower bound on the asymptotic performance of uniformly good allocation rules and construct an allocation scheme that achieves the bound. We discover that despite the inclusion of a switching cost the proposed allocation scheme achieves the same asymptotic performance as the optimal rule for the bandit problem without switching cost. This is made possible by grouping together the samples in a certain fashion. Finally, we illustrate an optimal allocation scheme for a large class of distributions which includes members of the exponential family.

## I. INTRODUCTION

A WELL investigated class of resource allocation problems is the multiarmed bandit problem which in its simplest form can be described as follows. There are  $p \geq 2$  statistical populations whose distributions are parametrized by an unknown  $\theta \in \Theta$ . How should we sample  $x_1, x_2, \dots$  sequentially from the  $p$  populations in order to maximize, in some sense, the expected value of the sum  $J_n = x_1 + x_2 + \dots + x_n$  as  $n \rightarrow \infty$ ?

Various versions of the above problem have been addressed in the recent control and statistics literature. Solutions have been obtained by Lai and Robbins [1], [2], and by Anantharam, Varaiya, and Walrand [5]. To the best of our knowledge none of the formulations of the multiarmed bandit problem presented so far addresses the issue of switching cost.

In this paper we formulate a multiarmed bandit problem with switching cost and present asymptotically optimal allocation schemes. The idea we exploit is the following. Since the introduction of a switching cost obviously discourages frequent switching, we need to sample in blocks. The surprising result we discover is that despite the inclusion of such an additional cost, our allocation schemes achieve the same asymptotic performance as the optimal solutions for the problem with no switching cost.

The paper is organized as follows. In Section II we precisely formulate a multiarmed bandit problem with switching cost and introduce the total regret associated with an allocation rule. In Section III we present a lower bound on the total regret, and in Section IV we construct allocation rules which achieve this lower bound. In Section V we illustrate such an allocation scheme for a large class of distributions which includes some members of the exponential family.

## II. PROBLEM FORMULATION

Let  $\Pi_1, \dots, \Pi_p$  be statistical populations specified, respectively, by univariate density functions  $f(x; \theta_1), \dots, f(x; \theta_p)$  with

Manuscript received May 14, 1987; revised March 23, 1988. Paper recommended by Associate Editor, D. A. Castanon. This work was supported by the National Science Foundation under Grant ECS-8517708.

R. Agrawal and D. Teneketzis are with the Department of Electrical Engineering and Computer Science and the Communications and Signal Processing Laboratory, University of Michigan, Ann Arbor, MI 48109.

M. V. Hegde is with the Department of Electrical and Computer Engineering, Louisiana State University, Baton Rouge, LA 70803.  
IEEE Log Number 8822541.

respect to some measure  $\nu$ , where  $f(\cdot; \cdot)$  is known and the  $\theta_j$ 's are unknown parameters belonging to some set  $\Theta$ . Assume that  $\int_{-\infty}^{\infty} |x|f(x; \theta) d\nu(x) < \infty$  for all  $\theta \in \Theta$ . Let

$$\mu(\theta) := \int_{-\infty}^{\infty} xf(x; \theta) d\nu(x) \tag{2.1}$$

and define

$$\mu^* := \max \{ \mu(\theta_1), \dots, \mu(\theta_p) \} = \mu(\theta^*) = \mu(\theta_{j^*}) \tag{2.2}$$

for some  $j^* \in \{1, \dots, p\}$  where  $\theta^* = \theta_{j^*}$ .

An adaptive allocation rule  $\phi$  consists of a sequence of random variables  $\phi_1, \phi_2, \dots$  taking values in the set  $\{1, \dots, p\}$  such that the event  $\{\phi_n = j\}$  ("sample from  $\Pi_j$  at stage  $n$ ") belongs to the  $\sigma$ -field  $F_{n-1}$  generated by  $\phi_1, x_1, \dots, \phi_{n-1}, x_{n-1}$ . For  $n \leq \infty$  let

$$T_n(j) := \sum_{i=1}^n 1\{\phi_i = j\} \tag{2.3}$$

denote the number of times that the rule  $\phi$  samples from  $\Pi_j$  through stage  $n$ , and let

$$J_n := \sum_{i=1}^n x_i. \tag{2.4}$$

Then by Wald's lemma (cf. [4])

$$EJ_n = \sum_{j=1}^p \mu(\theta_j) ET_n(j). \tag{2.5}$$

Define the sampling regret

$$R'_n(\theta) := n\mu^* - EJ_n = \sum_{j: \mu(\theta_j) < \mu^*} (\mu^* - \mu(\theta_j)) ET_n(j) \tag{2.6}$$

where  $\theta = (\theta_1, \dots, \theta_p)$ .

Also, let

$$S_n(j) := \sum_{i=2}^n 1\{\phi_i = j, \phi_{i-1} \neq j\} \tag{2.7}$$

and define the switching regret

$$SW_n(\theta) := C \sum_{j=1}^p ES_n(j) \tag{2.8}$$

where  $C > 0$  is the fixed switching cost.

Further, define the total regret

$$R_n(\theta) := R'_n(\theta) + SW_n(\theta). \tag{2.9}$$

We want to maximize, in some sense,  $(EJ_n - SW_n)$  which is equivalent to minimizing the total regret  $R_n(\theta)$ . Note that it is impossible to do this uniformly over all parameter configurations

$\theta$ . For example, the rule "always sample from population  $\Pi_1$ " will have zero regret when  $\mu(\theta_j) \leq \mu(\theta_1)$  for all  $j \in \{1, \dots, p\}$ . However, when a parameter configuration has  $\mu(\theta_j) > \mu(\theta_1)$  for some  $j \neq 1$ , this rule will have a regret proportional to  $n$ .

We call a rule uniformly good if for every parameter configuration  $\theta$

$$R_n(\theta) = o(n^a) \quad (2.10)$$

for every  $a > 0$ . Such rules do not allow the total regret to increase very rapidly for any  $\theta$ . We restrict our attention to the class of uniformly good rules, and consider any others as uninteresting. The problem formulation is now complete.

In what follows we make extensive use of the Kulback-Leibler number which is defined as

$$I(\theta, \lambda) = \int_{-\infty}^{\infty} [\log(f(x; \theta)/f(x; \lambda))] f(x; \theta) dv(x). \quad (2.11)$$

The Kulback-Leibler number is a well-known distance measure between two distributions.

The main results of the paper, appearing in Sections III and IV are derived under the following technical assumptions.

A.1:  $0 < I(\theta, \lambda) < \infty$  whenever  $\mu(\lambda) > \mu(\theta)$ .

A.2: For every  $\epsilon > 0$  and  $\theta, \lambda \in \Theta$  such that  $\mu(\lambda) \geq \mu(\theta)$ , there exists  $\delta > 0$ , such that

$$|I(\theta, \lambda) - I(\theta, \lambda')| < \epsilon \text{ if } \mu(\lambda) \leq \mu(\lambda') \leq \mu(\lambda) + \delta.$$

A.3:  $\forall \lambda \in \Theta$  and  $\forall \delta > 0$ ,  $\exists \lambda' \in \Theta$  such that  $\mu(\lambda) < \mu(\lambda') < \mu(\lambda) + \delta$ .

A.4: The parameter configuration  $\theta = (\theta_1, \dots, \theta_p)$  is such that  $\mu(\theta_j) < \mu^* = \mu(\theta_j^*)$  for all  $j \neq j^*$ .

The assumption  $I(\theta, \lambda) > 0$  is automatically satisfied whenever  $\mu(\lambda) > \mu(\theta)$ . The condition  $I(\theta, \lambda) < \infty$  implies that the distribution of the samples under the parameter  $\theta$  is absolutely continuous with respect to the distribution of the samples under any parameter  $\lambda$  such that  $\mu(\lambda) > \mu(\theta)$ . Such a condition can be expected to be satisfied for most parametric families of distributions which are mutually absolutely continuous. Assumption A2 is a right continuity condition on  $I(\theta, \lambda)$  for fixed  $\theta$  and  $\mu(\lambda) \geq \mu(\theta)$ . Assumption A3 is a denseness condition on the space  $\Theta$ . Assumptions A2-A3 are needed to obtain the lower bound on the total regret. Assumption A4 implies that there is a unique best population among all of the  $p$  populations. This assumption is essential in obtaining the upper bound on the total regret.

### III. A LOWER BOUND FOR THE TOTAL REGRET

In this section we note the extension of the lower bound obtained by Lai and Robbins [1] to our problem. We state this in the form of Theorem 3.1.

**Theorem 3.1:** Assume that Assumptions A1-A3 hold. Let  $\phi$  be any uniformly good allocation rule, i.e.,  $\phi$  satisfies (2.10). Then

$$i) \liminf_{n \rightarrow \infty} ET_n(\phi) / \log n \geq 1/I(\theta_j, \theta^*) \quad (3.1)$$

for any inferior population  $\Pi_j$ , i.e.,  $\mu(\theta_j) < \mu(\theta^*) = \mu^*$  and consequently

$$ii) \liminf_{n \rightarrow \infty} R_n(\theta) / \log n \geq \sum_{j: \mu(\theta_j) < \mu^*} \frac{(\mu^* - \mu(\theta_j))}{I(\theta_j, \theta^*)}. \quad (3.2)$$

**Proof:** Follows from Theorem 1 of Lai and Robbins [1].  $\square$

We shall call rules that attain the above lower bound asymptotically efficient, i.e.,

$$R_n(\theta) \sim \left( \sum_{j: \mu(\theta_j) < \mu^*} \frac{(\mu^* - \mu(\theta_j))}{I(\theta_j, \theta^*)} \right) \log n. \quad (3.3)$$

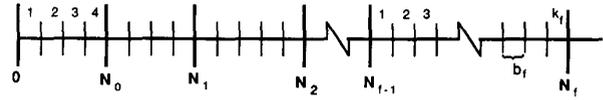


Fig. 1. Block allocation scheme for  $p = 4$ .

### IV. CONSTRUCTION OF ASYMPTOTICALLY EFFICIENT RULES

In the first part of Section IV-A we motivate the idea of block allocation and then introduce a specific block allocation scheme. In this scheme we employ upper confidence bounds that are constructed in Section IV-B. Finally, in Section IV-C we derive an upper bound on the total regret of our allocation scheme. This bound is asymptotically equal to the lower bound of Theorem 3.1. Consequently, the proposed allocation scheme is asymptotically efficient.

#### A. Block Allocation Scheme

In view of Theorem 3.1, if  $\phi$  is an asymptotically efficient rule, then the number of samples that  $\phi$  takes from any inferior population  $\Pi_j$  up to stage  $n$  is about  $(\log n)/I(\theta_j, \theta^*)$ . With no knowledge of the time instants at which the samples are taken from the inferior populations, all we can infer about the contribution from population  $\Pi_j$  to the switching regret up to stage  $n$  is that it is at most about  $(2 \log n)/I(\theta_j, \theta^*)$ . (The largest contribution to the switching cost occurs when every sample from population  $\Pi_j$  involves switching to and from it.) Clearly any asymptotically efficient rule must ensure that the samples chosen from any population are grouped together in blocks in such a fashion that the contribution to the switching cost is much smaller than the above upper bound, in fact  $o(\log n)$ . Furthermore, the block lengths must increase with  $n$ .

With this idea in mind we construct a "block allocation scheme" in two steps. We first determine, *a priori*, intervals of time, and over each interval we sample from the same population. Then, at the beginning of each interval we *adaptively* decide which population to sample from. The intervals are chosen so that if we ensure the expected numbers of samples from each inferior population is  $O(\log n)$ , the expected number of switches is automatically controlled to  $o(\log n)$ .

**Step 1:** To facilitate analysis, time is first divided into "frames" numbered  $0, 1, 2, \dots$ . Each frame  $f$  is further subdivided into "blocks" numbered  $1, 2, 3, \dots$ . All the blocks in a frame are of equal length. Each such block can thus be uniquely identified by  $(f, i)$  where  $f$  is the frame number to which it belongs, and  $i$  is the block number.

Furthermore, let

$N_f$  denote the time instant at the end of frame  $f$ ,

$b_f$  denote the block length of each block in frame  $f$ ,

$k_f$  denote the number of blocks in frame  $f$ .

We choose the block lengths and frame lengths  $(N_f - N_{f-1})$  as follows:

Frame # ( $f$ )	$b_f$	$N_f - N_{f-1}$
0	1	$p$
1	1	$\left[ \frac{2^{1^2} - 2^{0^2}}{1} \right] \cdot p.1 = p$
2	2	$\left[ \frac{2^{2^2} - 2^{1^2}}{2} \right] \cdot p.2 = 14p$
$f$	$f$	$\left[ \frac{2^{f^2} - 2^{(f-1)^2}}{f} \right] \cdot p.f. \quad (4.1)$

Such a block allocation scheme is illustrated in Fig. 1 for  $p = 4$ .

We now derive lower and upper bounds for  $N_f$ , the time

instants at the end of frame  $f$ . These bounds will be used extensively in subsequent analysis.

$$\begin{aligned} N_f &= \sum_{i=1}^f \left\lceil \frac{2^{i^2} - 2^{(i-1)^2}}{i} \right\rceil pi + p \\ &\geq \sum_{i=1}^f \left( \frac{2^{i^2} - 2^{(i-1)^2}}{i} \right) pi + p \\ &= p2^{f^2} \end{aligned} \quad (4.2)$$

$$\begin{aligned} N_f &\leq \sum_{i=1}^f \left( \frac{2^{i^2} - 2^{(i-1)^2}}{i} + 1 \right) pi + p \\ &= p2^{f^2} + p \sum_{i=1}^f i \\ &\leq p2^{f^2} + pf^2. \end{aligned} \quad (4.3)$$

Therefore,  $N_f \sim p2^{f^2}$ , and  $\log N_f \sim f^2$ . On the other hand, the block lengths  $b_f = f$ .

*Step 2:* To start up our allocation scheme, in frame 0 we sample once from each population. From then on we proceed as follows. The beginning of each block  $(f, i)$  is a comparison instant,  $n(f, i) (= N_{f-1} + (i-1)b_f)$ ; at that time we decide which population to sample from. We sample from the chosen population for the entire block  $(f, i)$ , i.e.,  $b_f$  times. To decide what population to sample from, at each comparison instant  $n$  we employ upper confidence bounds  $U_n(j)$  for  $\mu(\theta_j)$ , the mean of each population  $\Pi_j$ . We define the leader as the population  $\Pi_{j_n}$  for which

$$U_n(j_n) \geq U_n(j) \quad \forall j \in \{1, \dots, p\}.$$

At each comparison instant we choose to sample from the leader. (Ties are resolved by choosing any population among the leaders.)

### B. Construction of Upper Confidence Bounds

To fix ideas, let  $Y_1, Y_2, \dots$ , be i.i.d. random variables with a common density function  $f(y; \theta)$  with respect to some measure  $\nu$ , where  $\theta \in \Theta$  denotes an unknown parameter. Let  $g_{ni}: R^i \rightarrow R$  ( $n = 1, 2, \dots; i = 1, 2, \dots, n$ ) be Borel functions such that for every  $\theta \in \Theta$

$$P_\theta \{g_{ni}(Y_1, \dots, Y_i) \geq \mu(\theta) \mid \text{for all } i \leq n\} = 1 - o(n^{-1}) \quad (4.4)$$

$$\limsup_{n \rightarrow \infty} [E_\theta [\sup \{1 \leq i \leq n \mid g_{ni}(Y_1, \dots, Y_i) \geq \mu(\lambda)\}] / \log n] \leq \frac{1}{I(\theta, \lambda)} \quad (4.5)$$

whenever  $\mu(\lambda) > \mu(\theta)$ , and

$$g_{ni} \text{ is nondecreasing in } n \geq i \text{ for every fixed } i = 1, 2, \dots. \quad (4.6)$$

We now make use of the functions  $g_{ni}$  to define our upper confidence bounds. Let  $Y_{j1}, \dots, Y_{jT_n(j)}$  be the samples drawn from population  $\Pi_j$  up to stage  $n$  (i.e.,  $T_n(j)$  samples out of a total of  $n$  samples). Then at each comparison instant  $n(f, i)$  the upper confidence bound  $U_n(j)$  for  $\mu(\theta_j)$ , the mean of population  $\Pi_j$ , is given by

$$U_n(j) = g_{nT_n(j)}(Y_{j1}, \dots, Y_{jT_n(j)}) \quad (4.7)$$

for each  $j \in \{1, \dots, p\}$ .

Denote by  $\phi^*$  the allocation rule constructed in Sections IV-A and IV-B.

*Heuristics:* Conditions (4.4)–(4.6), which define the upper confidence bounds, play a very crucial role in proving asymptotic efficiency of the proposed block allocation scheme. That is why we discuss their intuitive meaning below.

Conditions (4.4), (4.5) relate the behavior of the sequence of upper confidence bounds  $U_n(j)$  of population  $\Pi_j$  to the actual mean  $\mu(\theta_j)$ . They upper bound the deviation above and below the actual mean.

Condition (4.4) shows that asymptotically the upper confidence bounds  $U_n(j)$  of population  $\Pi_j$  (for any  $j$ ) will very rarely fall below its actual mean.

Conditions (4.5), (4.6) can be interpreted as follows. Consider two populations; a population  $\Pi_j$  whose mean  $\mu(\theta_j)$  is unknown, and another population  $\Pi_k$  whose mean  $\mu(\theta_k) = \mu(\lambda)$  is known. Assume that we employ the following block allocation scheme. At each comparison instant (defined in the same way as before) we compare the upper confidence bound  $U_n(j)$  of  $\Pi_j$  to  $\mu(\lambda)$  and choose to sample from  $\Pi_j$  if  $U_n(j) \geq \mu(\lambda)$ , and from  $\Pi_k$  otherwise. Then, if  $\mu(\lambda) > \mu(\theta_j)$ , the expected number of samples from population  $\Pi_j$  up to stage  $N_f$  is upperbounded by  $\log N_f$  times a factor which measures how statistically distinct the population  $\Pi_j$  and  $\Pi_k$  are. The statistical distance between  $\Pi_j$  and  $\Pi_k$  is measured by the Kulback–Leibler number. This interpretation of (4.5), (4.6) is not only intuitively appealing but is also in agreement with standard results from statistical sequential analysis (Lai [3], Siegmund [6]). Moreover, it motivates the proposed block allocation scheme as follows. In the absence of knowledge of  $\mu(\theta_k)$  it is reasonable to substitute  $\mu(\theta_k)$  by an upper confidence bound that now in addition satisfies (4.4) so that the same results are maintained. These considerations are reflected in the analysis in Section IV-C.

In addition, conditions (4.4)–(4.6) allow us to substantiate the notion of an asymptotically efficient rule, which extends the notion of a consistent rule, proposed by Robbins in [7]. Briefly, consistent rules are rules that achieve optimality in an average cost sense, i.e.,  $\lim_{n \rightarrow \infty} (EJ_n - SW_n)/n = \mu^*$ . Consistent rules have been constructed in [7]. Consider only two populations  $a, b$  and let  $\{a_i\}_{i=1}^\infty$  and  $\{b_i\}_{i=1}^\infty$  be two fixed disjoint increasing sequences of positive integers such that  $a_n/n \rightarrow \infty$  and  $b_n/n \rightarrow \infty$ . At the time instants corresponding to either sequence the allocation rule forces the selection of the respective population  $a$  or  $b$ . At other time instants the rule selects the population with the largest sample mean. Consistency of the above rule is easily established using the strong law of large numbers and the sparseness of the sequences  $\{a_i\}$  and  $\{b_i\}$ . If asymptotic efficiency of the allocation rule is desired we should ensure that the normalized total regret  $n^{-1}R_n(\theta)$  tends to zero at the fastest possible rate as  $n \rightarrow \infty$ . This can only be achieved if the selection of the time instants  $\{a_i\}, \{b_i\}$  is not done *a priori* but *adaptively* from the data. Conditions (4.4)–(4.6) ensure that the sequences  $\{a_i\}$  and  $\{b_i\}$  are indeed generated adaptively at the best possible rate.

### C. Upper Bound on the Total Regret

*Theorem 4.1:* Under the block allocation rule  $\phi^*$ , for all  $\theta$  and every  $j$  such that  $\mu(\theta_j) < \mu^*$

$$i) E_\theta T_n(j) \leq \left( \frac{1}{I(\theta_j, \theta^*)} + o(1) \right) \log n, \quad (4.8)$$

$$ii) E_\theta S_n(j) \leq o(\log n), \quad (4.9)$$

and consequently, for all  $\theta$  satisfying Assumption A4

$$iii) \limsup_{n \rightarrow \infty} R_n(\theta) / \log n \leq \sum_{j: \mu(\theta_j) < \mu^*} (\mu^* - \mu(\theta_j)) / I(\theta_j, \theta^*). \quad (4.10)$$

*Proof:*

*Proof of 4.1 i):* We shall first prove i) for  $n = N_f$ , the end of

frame  $l$ , i.e., we shall show that

$$E_{\theta} T_{N_l}(j) \leq \left( \frac{1}{I(\theta_j, \theta^*)} + o(1) \right) \log N_l.$$

Let  $j^*$  be such that  $\mu(\theta_{j^*}) = \mu(\theta^*) = \mu^*$ . (In case there are more than one  $j^*$ , choose any one.) Let  $\#A$  denote the number of elements of a set  $A$ . Then for any fixed  $j$  such that  $\mu(\theta_j) < \mu^*$

$$\begin{aligned} T_{N_l(j)} &= \sum_{f=0}^l b_f [\# \{1 \leq i \leq k_f : \phi^* \text{ samples from } \Pi_j \text{ at} \\ &\quad \text{comparison instant } n(f, i)\}] \\ &= 1 + \sum_{f=1}^l b_f [\# \{1 \leq i \leq k_f : \phi^* \text{ samples from } \Pi_j \text{ at} \\ &\quad \text{comparison instant } n(f, i), U_n(j^*) \geq \mu(\theta_{j^*})\} \\ &\quad + \# \{1 \leq i \leq k_f : \phi^* \text{ samples from } \Pi_j \text{ at} \\ &\quad \text{comparison instant } n(f, i), U_n(j^*) < \mu(\theta_{j^*})\}] \\ &\leq 1 + \sum_{f=1}^l b_f [\# \{1 \leq i \leq k_f : U_n(j) \geq \mu(\theta^*), \text{ and } \phi^* \\ &\quad \text{samples from } \Pi_j \text{ at comparison instant } n(f, i)\}] \\ &\quad + \sum_{f=1}^l b_f [\# \{1 \leq i \leq k_f : U_n(j^*) < \mu(\theta_{j^*})\}] \\ &= 1 + \text{Term 1} + \text{Term 2 (say)}. \end{aligned}$$

*Claim 1:*

$$\limsup_{l \rightarrow \infty} \left( \frac{E_{\theta} [\text{Term 1}]}{\log N_l} \right) \leq \frac{1}{I(\theta_j, \theta^*)}.$$

*Claim 2:*  $E_{\theta} [\text{Term 2}] \leq o(1) \log N_l$ .

*Proof of Claim 1:*

$$\begin{aligned} \text{Term 1} &= \sum_{f=1}^l b_f [\# \{1 \leq i \leq k_f : U_n(j) \geq \mu(\theta^*), \text{ and } \phi^* \\ &\quad \text{samples from } \Pi_j \text{ at comparison instant } n(f, i)\}] \\ &= \sum_{f=1}^l b_f [\# \{1 \leq i \leq k_f : g_{nT_n(j)}(Y_{j1} \cdots Y_{jT_n(j)}) \geq \mu(\theta^*), \\ &\quad \text{and } \phi^* \text{ samples from } \Pi_j \text{ at comparison instant} \\ &\quad n(f, i)\}] \quad \text{by (4.7)} \\ &= \sum_{f=1}^l b_f [\# \{1 \leq i \leq k_f : g_{N_l T_n(j)}(Y_{j1}, \cdots, Y_{jT_n(j)}) \geq \mu(\theta^*), \\ &\quad \text{and } \phi^* \text{ samples from } \Pi_j \text{ at comparison instant} \\ &\quad n(f, i)\}] \quad \text{by (4.6)} \\ &\leq \sup \{1 \leq i \leq N_l | g_{N_l i}(Y_1, \cdots, Y_i) \geq \mu(\theta^*)\} + b_l. \end{aligned}$$

Then, by (4.1) and (4.2)

$$\frac{E_{\theta} [\text{Term 1}]}{\log N_l} \leq E_{\theta_j} [\sup \{1 \leq i \leq N_l | g_{N_l i}(Y_1, \cdots, Y_i) \geq \mu(\theta^*)\}] / \log N_l + l / \log(p2^{l^2}).$$

Hence, by (4.5)

$$\limsup_{l \rightarrow \infty} \left( \frac{E_{\theta} [\text{Term 1}]}{\log N_l} \right) \leq \frac{1}{I(\theta_j, \theta^*)}.$$

*Proof of Claim 2:*

$$\begin{aligned} \text{Term 2} &= \sum_{f=1}^l b_f [\# \{1 \leq i \leq k_f : U_n(j^*) < \mu(\theta_{j^*})\}] \\ &= \sum_{f=1}^l b_f \sum_{i=1}^{k_f} 1 \{U_n(j^*) < \mu(\theta_{j^*})\}. \end{aligned}$$

Then,

$$\begin{aligned} E_{\theta} [\text{Term 2}] &= \sum_{f=1}^l b_f \sum_{i=1}^{k_f} P_{\theta} \{U_n(j^*) < \mu(\theta_{j^*})\} \\ &\leq \sum_{f=1}^l b_f \sum_{i=1}^{k_f} P_{\theta_{j^*}} \{g_{nt}(Y_{j^*1}, \cdots, Y_{j^*t}) < \mu(\theta_{j^*}) \\ &\quad \text{for some } t \leq n\} \\ &= \sum_{f=1}^l b_f \sum_{i=1}^{k_f} o(n^{-1}) \quad \text{by (4.4)} \\ &= o \left( \sum_{f=1}^l b_f \sum_{i=1}^{k_f} \frac{1}{n(f, i)} \right) \\ &= o(\log N_l) \quad \text{(by Appendix } R_1). \end{aligned}$$

Thus, by Claims 1 and 2

$$E_{\theta} T_{N_l}(j) \leq \left( \frac{1}{I(\theta_j, \theta^*)} + o(1) \right) \log N_l.$$

Now we extend this result for any arbitrary  $n$ . Let  $l$  be such that  $N_{l-1} < n \leq N_l$ .

Clearly

$$\begin{aligned} \frac{E_{\theta} T_n(j)}{\log n} &\leq \frac{E_{\theta} [T_{N_l}(j)]}{\log N_{l-1}} \\ &\leq \left( \frac{1}{I(\theta_j, \theta^*)} + o(1) \right) \frac{\log(p2^{l^2} + p^{l^2})}{\log(p2^{(l-1)^2})} \\ &= \left( \frac{1}{I(\theta_j, \theta^*)} + o(1) \right). \end{aligned}$$

This completes the proof of i).

*Proof of 4.1 ii):* Let  $l$  be such that  $N_{l-1} < n \leq N_l$ . Then,

$$\begin{aligned} S_n(j) &\leq 1 + \sum_{f=1}^l \frac{T_{N_l}(j) - T_{N_{f-1}}(j)}{f} \\ &= 1 + \sum_{f=1}^l \frac{T_{N_l}(j)}{f} - \sum_{f=0}^{l-1} \frac{T_{N_f}(j)}{f+1} \\ &= \frac{T_{N_l}(j)}{l} + \sum_{f=1}^{l-1} T_{N_f}(j) \left[ \frac{1}{f} - \frac{1}{f+1} \right] \\ &\leq \frac{T_{N_l}(j)}{l} + \sum_{f=1}^{l-1} T_{N_f}(j) \left[ \frac{1}{f^2} \right]. \end{aligned}$$

By (4.8)

$$\limsup_{l \rightarrow \infty} \frac{E_{\theta} T_{N_l}(j)}{\log N_l} \leq 1/I(\theta_j, \theta^*).$$

Therefore, for any given  $\epsilon > 0 \exists f_0$  such that  $\forall f \geq f_0$

$$\frac{E_{\theta} T_{N_f}(j)}{\log N_f} \leq 1/I(\theta_j, \theta^*) + \epsilon.$$

Hence, by (4.3)

$$\begin{aligned} E_{\theta} T_{N_f}(j) &\leq (1/I(\theta_j, \theta^*) + \epsilon) \log(p2^{f^2} + pf^2) \\ &\leq K(\epsilon)f^2 \quad \text{for some } K(\epsilon). \end{aligned}$$

Thus, for  $n > N_{f_0}$

$$\begin{aligned} E_{\theta}(S_N(j)) &\leq \frac{E_{\theta} T_{N_l}(j)}{l} + \sum_{f=1}^{l-1} \frac{E_{\theta} T_{N_f}(j)}{f^2} \\ &\leq \frac{K(\epsilon)l^2}{l} + \sum_{f=f_0}^{l-1} K(\epsilon) + \sum_{f=1}^{f_0-1} \frac{E_{\theta} T_{N_f}(j)}{f^2} \\ &\leq K(\epsilon)2l + M(\epsilon) \end{aligned}$$

where

$$M(\epsilon) = \sum_{f=1}^{f_0-1} \frac{E_{\theta} T_{N_f}(j)}{f^2}.$$

Consequently,

$$\frac{E_{\theta} S_n(j)}{\log n} \leq \frac{E_{\theta} S_n(j)}{\log N_{l-1}} \leq \frac{K(\epsilon)2l + M(\epsilon)}{(l-1)^2} = o(1) \quad \text{proving ii).}$$

*Proof of 4.1 iii):* By (4.8) and (2.6) it follows that

$$R_n^*(\theta) = \left[ \sum_{j: \mu(\theta_j) < \mu^*} (\mu^* - \mu(\theta_j))/I(\theta_j, \theta^*) + o(1) \right] \log n. \quad (4.11)$$

Under the additional Assumption A4 we have from (2.8)

$$\begin{aligned} SW_n(\theta) &= C \left[ \sum_{j \neq j^*} ES_n(j) + ES_n(j^*) \right] \\ &\leq C \left[ 2 \sum_{j \neq j^*} ES_n(j) + 1 \right] \\ &= o(\log n). \end{aligned} \quad (4.12)$$

Hence, by (2.9), (4.11), and (4.12), iii) follows.  $\square$

In view of Theorems 3.1 and 4.1 the block allocation scheme  $\phi^*$  that we propose in this section is asymptotically efficient, i.e.,

$$R_n(\theta) \sim \left[ \sum_{j: \mu(\theta_j) < \mu^*} (\mu^* - \mu(\theta_j))/I(\theta_j, \theta^*) \right] \log n.$$

Thus, despite the imposition of a switching cost we are able to recapture the same asymptotically optimal performance as Lai and Robbins [1] achieve in the nonswitching cost case. The block allocation scheme proposed in this section is crucial in achieving this performance. By grouping together samples from each inferior population in blocks, we manage to maintain the number of samples from each inferior population at about  $\log n / I(\theta_j, \theta^*)$  and to limit the number of switches to  $o(\log n)$ .

## V. UPPER CONFIDENCE BOUNDS FOR SPECIAL DISTRIBUTIONS

In this section we construct upper confidence bounds for a family of distributions under the following assumptions: A.1–A.3

and

$$\Theta = R, \quad (5.1)$$

$$\mu(\theta) = \int_{-\infty}^{\infty} xf(x; \theta) d\nu(x) \quad (5.2)$$

is a strict monotone increasing function of the parameter  $\theta$

$$\log f(x; \theta) \text{ is concave in } \theta \text{ for each fixed } x, \quad (5.3)$$

$$\int_{-\infty}^{\infty} x^2 f(x; \theta) d\nu(x) < \infty \quad \text{for each } \theta \in \Theta. \quad (5.4)$$

Conditions (5.1)–(5.4) are satisfied by some members of the exponential family of distributions, for instance, the normal distribution. We prove that the upper confidence bounds we construct satisfy conditions (4.4)–(4.6). Lemmas 5.1 and 5.2 are needed for the proof.

*Lemma 5.1:* Let  $S_n = X_1 + \dots + X_n$  where  $X_1, X_2, \dots$  are i.i.d.,  $EX_1 > 0$  and let  $N = \sum_{n=1}^{\infty} 1(S_n \leq 0)$ ,  $L = \sum_{n=1}^{\infty} 1(\inf_{i \geq n} S_i \leq 0)$ .

Then the following are equivalent.

- $E(|X_1|^2 1(X_1 \leq 0)) < \infty$
- $EN < \infty$
- $EL < \infty$ .

*Proof:* See Hogan [7].  $\square$

*Lemma 5.2:* Let  $S_n = X_1 + X_2 + \dots + X_n$  where  $X_1, X_2, \dots$  are i.i.d.,  $0 < EX_1 < \infty$ . Given  $A > 0$ , let  $L_A = \sum_{n=1}^{\infty} 1(\inf_{i \geq n} S_i \leq A)$ . If  $E(|X_1|^2 1(X_1 \leq 0)) < \infty$ , then

$$\limsup_{A \rightarrow \infty} \frac{EL_A}{A} \leq \frac{1}{EX_1}.$$

*Proof:* For  $\epsilon > 0$

$$L_A - \sum_{n=1}^{\infty} 1 \left( \inf_{i \geq n} \left( S_i - \frac{iEX_1}{1+\epsilon} \right) \leq 0 \right) \leq \frac{A(1+\epsilon)}{EX_1}.$$

Now we need to consider the i.i.d. r.v.'s  $Z_i = X_i - EX_1/(1+\epsilon)$ .

$$\begin{aligned} &E\{|Z_1|^2 1(Z_1 \leq 0)\} \\ &\leq 2E \left\{ \left[ |X_1|^2 + \left( \frac{EX_1}{1+\epsilon} \right)^2 \right] 1 \left( X_1 \leq \frac{EX_1}{1+\epsilon} \right) \right\} \\ &\leq 2E\{|X_1|^2 1(X_1 \leq 0)\} \\ &\quad + 2E \left\{ |X_1|^2 1 \left( 0 < X_1 \leq \frac{EX_1}{1+\epsilon} \right) \right\} \\ &\quad + 2 \left( \frac{EX_1}{1+\epsilon} \right)^2 \\ &< \infty. \end{aligned}$$

Thus, by Lemma 5.1 it follows that for some constant  $K(\epsilon)$

$$EL_A \leq \frac{A(1+\epsilon)}{EX_1} + K(\epsilon)$$

$$\limsup_{A \rightarrow \infty} \frac{EL_A}{A} \leq \frac{1+\epsilon}{EX_1}.$$

By letting  $\epsilon \rightarrow 0$  we get the desired result.

Using Lemmas 5.1 and 5.2 we construct upper confidence bounds in two steps described by Theorems 5.1 and 5.2.

**Theorem 5.1:** Let  $Y_1, Y_2, \dots$  be a sequence of i.i.d. samples from a population. Let

$$W_i(\theta) = \int_{-\infty}^0 \prod_{t=1}^i \frac{f(Y_t; \theta + w)}{f(Y_t; \theta)} l(w) dw$$

where  $l: (-\infty, 0) \rightarrow R_+$  is a positive, continuous function with  $\int_{-\infty}^0 l(w) dw = 1$ . For any  $K > 0$  let

$$V(i, Y_1, \dots, Y_i, K) = \inf \{ \theta \mid W_i(\theta) \geq K \}. \quad (5.5)$$

Then, for all  $\lambda > \theta$

$$i) P_\theta \{ \theta \leq V(i, Y_1, \dots, Y_i, K) \text{ for all } i \geq 1 \} \geq 1 - \frac{1}{K},$$

$$ii) \lim_{K \rightarrow \infty} \frac{1}{\log K} E_\theta \{ \sup \{ i \geq 1 \mid V(i, Y_1, \dots, Y_i, K) \geq \lambda \} \} = 1/I(\theta, \lambda).$$

*Proof:* By (5.3) it follows that  $W_i(\theta)$  is nondecreasing in  $\theta$ , so that

$$V(i, Y_1, \dots, Y_i, K) < \theta \Rightarrow W_i(\theta) \geq K.$$

Now

$$\{ V(i, Y_1, \dots, Y_i, K) < \theta \text{ for some } i \geq 1 \} \subseteq \{ W_i(\theta) \geq K \text{ for some } i \geq 1 \}.$$

$W_i(\theta)$  is a nonnegative Martingale under  $\theta$  with mean 1. Thus, by the sub-Martingale inequality (cf. [4])

$$P_\theta \{ W_i(\theta) \geq K \text{ for some } i \geq 1 \} \leq 1/K$$

establishing i).

Let  $L_K = \sum_{i=1}^{\infty} 1(\inf_{t \geq i} W_t(\lambda) \leq K)$ .

Given  $\epsilon > 0$ , choose  $0 < \delta < \lambda - \theta$  so that

$$|I(\theta, w)| < \epsilon \quad \text{if } \theta < w < \theta + \delta.$$

[This is possible due to Assumptions A2, A3, and (5.2).]

We have

$$\begin{aligned} \{ W_i(\lambda) \leq K \} &= \left\{ \log \int_{-\infty}^0 \prod_{t=1}^i \frac{f(Y_t; \lambda + w)}{f(Y_t; \lambda)} l(w) dw \leq \log K \right\} \\ &\subseteq \left\{ \log \int_{\theta < w < \theta + \delta} \prod_{t=1}^i \frac{f(Y_t; w)}{f(Y_t; \lambda)} l(w - \lambda) dw \leq \log K \right\} \\ &= \left\{ \log \int_{\theta < w < \theta + \delta} \prod_{t=1}^i \frac{f(Y_t; w)}{f(Y_t; \lambda)} l^\circ(w) dw \leq \log K - \log A \right\} \end{aligned}$$

where

$$A = \int_{\theta < w < \theta + \delta} l(w - \lambda) dw, \quad l^\circ(w) = \frac{l(w - \lambda)}{A}.$$

By Jensen's inequality

$$\{ W_i(\lambda) \leq K \} \subseteq \left\{ \sum_{t=1}^i \int_{\theta < w < \theta + \delta} \log \left( \frac{f(Y_t; w)}{f(Y_t; \lambda)} \right) l^\circ(w) dw \leq \log K - \log A \right\}.$$

Thus, we must examine the sum of i.i.d. variables

$$X_t = \int_{\theta < w < \theta + \delta} \log \frac{f(Y_t; w)}{f(Y_t; \lambda)} l^\circ(w) dw$$

where  $Y_t$  has the distribution  $f(x; \theta)$ . These random variables have mean

$$\begin{aligned} EX_1 &= E_\theta \left[ \log \frac{f(Y_1; \theta)}{f(Y_1; \lambda)} + \int_{\theta < w < \theta + \delta} \log \frac{f(Y_1; w)}{f(Y_1; \theta)} l^\circ(w) dw \right] \\ &\geq I(\theta, \lambda) - \epsilon > 0 \quad \text{for } \epsilon \text{ sufficiently small.} \end{aligned}$$

Also  $EX_1 < \infty$ . Next we prove that  $E[|X_1|^2 1(X_1 \leq 0)] < \infty$  so that we can use Lemma 5.2 to upperbound  $E_\theta L_K$ .

Note that

$$\begin{aligned} 0 \geq X_1 1(X_1 \leq 0) &\geq \int_{\theta < w < \theta + \delta} \left( \log \frac{f(Y_1; w)}{f(Y_1; \lambda)} \right) \\ &\quad \cdot 1 \left( \frac{f(Y_1; w)}{f(Y_1; \lambda)} \leq 1 \right) l^\circ(w) dw \end{aligned}$$

$$E_\theta [ (X_1 1(X_1 \leq 0))^2 ]$$

$$\leq \int_{\theta < w < \theta + \delta} E_\theta \left[ \left( \log \frac{f(Y_1; w)}{f(Y_1; \lambda)} \right) 1 \left( \frac{f(Y_1; w)}{f(Y_1; \lambda)} \leq 1 \right) \right]^2 l^\circ(w) dw.$$

Now

$$\begin{aligned} &\int f(x; \theta) \left[ \log \frac{f(x; w)}{f(x; \lambda)} \right]^2 1 \left( \frac{f(x; w)}{f(x; \lambda)} \leq 1 \right) dv \\ &= \int \frac{f(x; \theta) f(x; \lambda) f(x; w)}{f(x; w) f(x; \lambda)} \left[ \log \frac{f(x; w)}{f(x; \lambda)} \right]^2 \\ &\quad \cdot 1 \left( \frac{f(x; w)}{f(x; \lambda)} \leq 1 \right) dv. \end{aligned} \quad (5.6)$$

Observe that

$$a) x [\log x]^2 \leq \frac{4}{e^2} \text{ on } \{ x \leq 1 \},$$

$$b) \text{ since } \lambda > w > \theta, \text{ there is } 0 < \alpha < 1 \text{ such that } w = \alpha\theta + (1 - \alpha)\lambda.$$

$$(5.7)$$

By (5.3) for each  $x$ ,  $f(x; w) \geq f(x; \theta)^\alpha f(x; \lambda)^{(1-\alpha)}$ . Hence,

$$\frac{f(x; \theta) f(x; \lambda)}{f(x; w)} \leq f(x; \theta)^{(1-\alpha)} f(x; \lambda)^\alpha. \quad (5.8)$$

Let  $w^\circ = \alpha\lambda + (1 - \alpha)\theta$ . By (5.3) again

$$f(x; \theta)^{(1-\alpha)} f(x; \lambda)^\alpha \leq f(x; w^\circ). \quad (5.9)$$

Combining (5.6)–(5.9), we obtain

$$E_\theta [X_1 1(X_1 \leq 0)]^2 \leq \frac{4}{e^2}.$$

We may now use Lemma 5.2 to conclude that  $E_\theta L_K < \infty$  and

$$\limsup_{K \rightarrow \infty} \frac{E_\theta L_K}{\log K} \leq \frac{1}{I(\theta, \lambda) - \epsilon}.$$

Letting  $\epsilon \rightarrow 0$  gives

$$\limsup_{K \rightarrow \infty} \frac{E_\theta L_K}{\log K} \leq \frac{1}{I(\theta, \lambda)}. \quad (5.10)$$

We now bound  $E_\theta L_K$  from below. Define the stopping time

$$T_K = \inf \{i \geq 1 \mid W_i(\lambda) \geq K\}.$$

Notice that  $L_K \geq T_K - 1$ . Thus,  $E_\theta T_K < \infty$ .

Write

$$W_i(\lambda) = \prod_{t=1}^i \frac{f(Y_t, \theta)}{f(Y_t, \lambda)} \int_{-\infty}^0 \prod_{t=1}^i \frac{f(Y_t, \lambda + w)}{f(Y_t, \theta)} l(w) dw$$

$$= \Lambda_i M_i$$

where

$$\Lambda_i = \prod_{t=1}^i \frac{f(Y_t, \theta)}{f(Y_t, \lambda)}, \quad M_i = \int_{-\infty}^0 \prod_{t=1}^i \frac{f(Y_t, \lambda + w)}{f(Y_t, \theta)} l(w) dw$$

and  $M_i$  is a Martingale under  $\theta$  of mean 1. We observe that

$$\log K \leq E_\theta \log W_{T_K}(\lambda) = E_\theta \log \Lambda_{T_K} + E_\theta \log M_{T_K}$$

$$\leq E_\theta \log \Lambda_{T_K} + \log E_\theta M_{T_K}$$

$$= I(\theta, \lambda) E_\theta T_K \leq I(\theta, \lambda) E_\theta L_K$$

which along with (5.10) establishes ii).  $\square$

**Theorem 5.2:** Let  $g_{ni}(Y_1, \dots, Y_i) = \mu[V(i, Y_1, \dots, Y_i, n(\log n)^p)]$  for some  $p > 1$ . Then for any  $\lambda > \theta$

i)  $P_\theta \{g_{ni}(Y_1, \dots, Y_i) > \mu(\theta) \text{ for all } i \leq n\} = 1 - O(n^{-1}(\log n)^{-p})$

ii)  $\limsup_{n \rightarrow \infty} [E_\theta \{\sup \{1 \leq i \leq n: g_{ni}(Y_1, \dots, Y_i) \geq \mu(\lambda)\} / \log n\}] \leq 1/I(\theta, \lambda)$

iii)  $g_{ni}$  is nondecreasing in  $n \geq i$  for fixed  $i$ .

*Proof:* i) follows from i) and ii) from ii) of Theorem 5.1, iii) follows from the form of  $V(i, Y_1, \dots, Y_i, K)$  and the assumption that  $\mu(\theta)$  is monotonically increasing in  $\theta$ .  $\square$

Theorem 5.2 shows that upper confidence bounds  $g_{ni}$  constructed above satisfy the required conditions (4.4)–(4.6). This concludes our construction.

### VI. CONCLUSIONS

Despite the inclusion of a switching cost, our allocation scheme achieves the same asymptotic performance as the optimal solutions for the case without switching cost. This is made possible by grouping together samples into blocks of increasing sizes, thereby reducing the number of switches to  $o(\log n)$ .

Notice that the block length and frame lengths are prescribed in advance and not generated adaptively from the data. With our block scheme if we can ensure that the number of samples from an inferior population is  $O(\log n)$ , then we automatically control the number of switches to  $o(\log n)$ .

Although in our problem formulation we consider a fixed switching cost, we can equally well handle switching costs which vary with time and with the pair of populations between which switching occurs, provided the switching cost is bounded.

Assumption A4 is essential to obtain asymptotic efficiency. If we do not have a unique best population, then the number of switches among superior populations can be arbitrarily large.

### APPENDIX

$$R1: \sum_{f=1}^l b_f \sum_{i=1}^{k_f} \frac{1}{n(f, i)} \quad (b_f = f)$$

$$\leq \sum_{t=1}^{N_l} \frac{1}{t} + \sum_{f=1}^l \frac{1}{N_f - b_f} \quad (\text{see explanation below})$$

$$\leq \sum_{t=1}^{N_l} \frac{1}{t} + \sum_{f=2}^l \frac{1}{p2^{f^2-f}} + \frac{1}{N_l - 1}$$

$$\leq \sum_{t=1}^{N_l} \frac{1}{t} + \frac{1}{p} \sum_{f=2}^l \left(\frac{1}{2}\right)^f + \frac{1}{N_l - 1}$$

$$\leq \log N_l + 1 + \frac{1}{p} + \frac{1}{N_l - 1}$$

$$\sum_{f=1}^l b_f \sum_{i=1}^{k_f} \frac{1}{n(f, i)}$$

$$= \frac{1}{N_0} + \frac{1}{N_0 + 1} + \dots + \frac{1}{N_l - 1} + \frac{1}{N_l} + \frac{1}{N_l}$$

$$+ \frac{1}{N_l + 2} + \frac{1}{N_l + 2} + \dots + \frac{1}{N_2 - 2} + \frac{1}{N_2 - 2}$$

$$+ \frac{1}{N_2} + \frac{1}{N_2} + \frac{1}{N_2} + \dots$$

$$\leq \frac{1}{N_0} + \frac{1}{N_0 + 1} + \dots + \frac{1}{N_l - 1} + \frac{1}{N_l - 1}$$

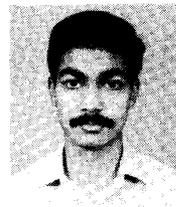
$$+ \frac{1}{N_l} + \frac{1}{N_l + 1} + \frac{1}{N_l + 2} + \dots + \frac{1}{N_2 - 3} + \frac{1}{N_2 - 2}$$

$$+ \frac{1}{N_2 - 2} + \frac{1}{N_2 - 1} + \frac{1}{N_2} + \dots$$

$$= \sum_{t=N_0}^{N_l - 1} \frac{1}{t} + \sum_{f=1}^l \frac{1}{N_f - b_f}.$$

### REFERENCES

- [1] T. L. Lai and H. Robbins, "Asymptotically efficient adaptive allocation rules," *Adv. Appl. Math.*, vol. 6, pp. 4–22, 1985.
- [2] T. L. Lai and H. Robbins, "Asymptotically efficient allocation of treatments in sequential experiments," in *Design of Experiments*, T. J. Santner and A. C. Tamhane, Eds. New York: Marcel Dekker, pp. 127–142.
- [3] T. L. Lai, "Some thoughts on stochastic adaptive control," in *Proc. 23rd IEEE Conf. Decision Contr.*, Las Vegas, NV, Dec. 1984, pp. 51–56.
- [4] S. M. Ross, *Stochastic Processes*. New York: Wiley, 1983.
- [5] V. Ananthraman, P. Varaiya, and J. Walrand, "Asymptotically efficient allocation rules for multi-armed bandit problem with multiple plays—Part I: I.I.D. Rewards; Part II: Markovian rewards," *IEEE Trans. Automat. Contr.*, vol. AC-32, pp. 968–982, Nov. 1987.
- [6] D. Siegmund, *Sequential Analysis*. New York: Springer-Verlag, 1985.
- [7] M. Hogan, "Moments of the minimum of a random walk and complete convergence," Tech. Rep. 21, Dep. Statistics, Stanford Univ., Jan. 1983.
- [8] H. Robbins, "Some aspects of the sequential design of experiments," *Bull. Amer. Math. Soc.*, vol. 55, pp. 527–535, 1952.



**Rajeev Agrawal** was born on December 1, 1963 in Lucknow, India. He received the B.Tech degree in electrical engineering from the Indian Institute of Technology, Kanpur, in May 1985. In 1987 he received the M.S. degree in electrical engineering-systems from the University of Michigan, Ann Arbor, where he is currently completing the requirements for the Ph.D. degree.

At the University of Michigan he has been a Research/Teaching Assistant since 1985. His current research interests are in stochastic systems, stochastic adaptive control, resource allocation problems, stochastic scheduling, communication networks, and communication systems.

Mr. Agrawal received the Outstanding Graduate Student Award for electrical engineering-systems in March 1988 from the College of Engineering at the University of Michigan.



**Manjunath V. Hegde** (S'84-M'87) received the B.Tech. degree in electrical engineering from the Indian Institute of Technology, Bombay, India, in 1979, the M.M.S. degree in finance from the University of Bombay, Bombay, India, in 1981, the M.S. degree in statistics from the University of Toledo, Toledo, OH, in 1983, and the Ph.D. degree in computer, information, and control engineering from the University of Michigan, Ann Arbor, in 1987.

From May 1981 to August 1982, he worked as a Financial Consultant for 20th Century Financial Services P. Ltd., Bombay, where he was involved with Eurodollar syndications. Since August 1987 he

has been an Assistant Professor of Electrical and Computer Engineering at Louisiana State University, Baton Rouge. His research interests include communication and information theory, coding theory, game theory, and stochastic optimization.

**Demosthenis Teneketzis**, for a photograph and biography, see p. 523 of the June 1988 issue of this TRANSACTIONS.

---