INFORMATION AND COMPUTATION 73, 23–44 (1987)

The Decentralized Wald Problem

DEMOSTHENIS TENEKETZIS

Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, Michigan 48109

AND

YU-CHI HO

Division of Applied Sciences, Harvard University, Cambridge, Massachusetts 02138

Two detectors making independent observations must decide which one of two hypotheses is true. The decisions are coupled through a common cost function. It is shown that the detectors' optimal decisions are characterized by thresholds which are coupled and whose computation requires the solution of two coupled sets of dynamic programming equations. An approximate computation of the thresholds is proposed and numerical results are presented.

1. INTRODUCTION

The classical theory of optimal sensor signal processing is based on statistical estimation and hypothesis testing methods (Van Trees, 1969). The salient feature of classical signal processing theory is that all sensor signals are implicitly assumed to be available in one place for processing. In recent years, however, there has been an increasing interest in distributed sensor systems. This interest has been sparked by large-scale systems such as power systems, surveillance systems, etc., where because of considerations such as cost, reliability, survivability, communication bandwidth, compartmentalization, or even problems caused by flooding a central processor with more information than it can process, there is never centralization of information in practice. Thus, extensions are needed to the classical framework of detection theory if it is to be relevant to the design of distributed systems. The purpose of this paper is to attempt a modest step in the direction of a detection theory for distributed sensors.

In this paper we study one of the simplest possible decentralized detection problems. We consider two detectors 1 and 2, and two hypotheses

0890-5401/87 $3.00

$h_0 = 0$ and $h_1 = 1$. The detectors make independent observations and based only on their information they have to decide which hypothesis is true. Each observation is costly. The cost associated with the final decisions u_i ($u_i = 0, 1$, $i = 1, 2$) of the detectors is $J(u_1, u_2, h)$. In general $J(u_1, u_2, h) \neq J_1(u_1, h) + J_2(u_2, h)$ so that *the detectors are coupled through their common cost.* The detectors' objective is to determine the optimal decision rules which minimize the average cost due to their observations and their final decisions.

A similar situation where two or more detectors with different information are coupled through a common cost has been previously considered by Tenney and Sandell (1981) and Lauer and Sandell (1982). However, the problems studied by Tenney and Sandell (1981) and Lauer and Sandell (1982) are considerably simpler than the problem considered here because the detector's final decisions are based on a single observation only. A model of decentralized hypothesis testing and coordination where the detectors are allowed to accumulate more information at some cost has been recently considered by Kushner and Pacut (1982). The presence of the coordinator, as well as the approach taken in Kushner and Pacut (1982) (simulation study), makes that problem essentially different from the problem and the approach presented in this paper. Another model of decentralized detection where the detectors are allowed to accumulate more information at some cost has been considered by Teneketzis and Varaiya (1984) and Teneketzis and Sandell (1985). However, the objective in Teneketzis and Varaiya (1984) and Teneketzis and Sandell (1985) is to detect the time of the jump from one hypothesis to another and not the true hypothesis. The same problem has been considered in Teneketzis (1982). The results presented here are more general as they deal with both the finite and infinite horizon decentralized Wald problem and provide an approximate solution to the problem when the statistics of the observation noise are Gaussian.

The remainder of the paper is organized as follows: The formal model is presented in Section 2. Section 3 is devoted to a proof of the threshold property. An approximate computation of the thresholds is proposed in Section 4, and the numerical results of the proposed computation appear in Section 5.

2. The Model

2.1. *Problem Formulation*

Consider two hypotheses $h_0 = 0$, $h_1 = 1$ and assume that

$$\text{Prob}(h = 0) = p \tag{2.1}$$

Consider two detectors 1 and 2 and make the following assumptions:

(A.1) The ith detector's observation at time t is described by

$$y_i(t) = f_i(h, w_t^i), \qquad i = 1, 2, \tag{2.2}$$

where $\{w_t^i\}$, $i = 1, 2$ are mutually independent i.i.d. sequences which are also independent of the hypothesis h. A typical example is the case of Eq. (4.1), where $y_i(t) = h + w_i(t)$. The probability p, the distributions of w^1, w^2 and the functions f_1, f_2 are known to the designer of the policies.

(A.2) The two detectors do not communicate. Each detector has to decide which hypothesis is true based on its own observations. Thus, if u_i is the decision of detector i, and t is the time this decision is made then

$$u_i(t) = \gamma_i(y_i^t), \tag{2.3}$$

where

$$y_i^t := (y_i(1) \cdots y_i(t)) \tag{2.4}$$

$$u_i = 0, 1, \qquad i = 1, 2. \tag{2.5}$$

(A.3) The cost incurred by the final decisions u_i of the detectors is $J(u_1, u_2, h)$, where h is the true hypothesis. In general, $J(u_1, u_2, h) \neq J_1(u_1, h) + J_2(u_2, h)$. Otherwise the problem decomposes into two standard independent Wald problems (Wald, 1947; Bertsekas, 1976; and Chernoff, 1972). It is the coupling of the detectors through the cost that makes this problem interesting. Furthermore,

$$\begin{aligned} J(0, u_2, h_1) &\geqslant J(1, u_2, h_1) \\ J(1, u_2, h_0) &\geqslant J(1, u_2, h_1) \\ J(1, u_2, h_0) &\geqslant J(0, u_2, h_0) \\ J(0, u_2, h_1) &\geqslant J(0, u_2, h_0). \end{aligned} \tag{2.6}$$

Similar relations hold for u_1. All inequalities in (2.6) imply that at most one mistake is less costly than at least one mistake.

(A.4) Each observation made by each detector costs c.

Let $Y_i^t = \sigma(y_i(s), s \leqslant t)$, let τ_i denote Y_i^t stopping times and let Γ_i $(i = 1, 2)$ denote the set of stopping rules which are measurable functions of the data of detector i. The Decentralized Wald problem is

$$\underset{\{\gamma_i \in \Gamma_i\} i = 1,2}{\text{Minimize}} \; E\{c\tau_1(\gamma_1) + c\tau_2(\gamma_2) + J(\gamma_1(y_1^{\tau_1}), \gamma_2(y_2^{\tau_2}), h)\} \tag{2.7}$$

subject to the assumptions above.

2.2. *Features of the Problem*

The salient features of the problem formulated above are:

1. There are two detectors with *different* information
2. The decisions of the detectors are *coupled* through their common cost.

Since $J(u_1, u_2, h) \neq J_1(u_1, h) + J_2(u_2, h)$, the decentralized Wald problem is a team problem. More specifically, it is a sequential team problem with static information structure. The information structure is static because each detector's information is not affected by the actions of the other detector, (Ho, 1972, and Yoshikawa, 1978). Thus, the decomposition techniques of Yoshikawa (1978) can be used to determine the member by member optimal solutions of the decentralized Wald problem.

3. Analysis

Fix $\gamma_2 \in \Gamma_2$, possibly at the optimum. Then, detector 1's problem is to determine a stopping rule to minimize $E\mathbf{L}(\gamma_1)$, where

$$E\mathbf{L}(\gamma_1) = E\{c\tau_1(\gamma_1) + J(\gamma_1(y_1^{\tau_1}), u_2, h)\}. \tag{3.1}$$

Note that in (3.1) we have used u_2 instead of $\gamma_2(y_2^{\tau_2})$; we will use the same notation as in (3.1) in the sequel, with the understanding that u_2 is a random variable whose statistics depend on the decision rule γ_2.

In extensive form the problem for detector 1 is

$$\underset{u_1 \in \{0,1\}, \tau_1}{\text{Minimize}}\; E\mathbf{L}(u_1, \tau_1) \tag{3.2}$$

where

$$E\mathbf{L}(u_1, \tau_1) = E\{c\tau_1 + J(u_1, u_2, h) \mid Y_1^{\tau_1}\}. \tag{3.3}$$

This problem can be solved by backward induction. We first establish some notation, then we consider a finite horizon T and finally let $T \to \infty$.

3.1 *Preliminaries*

To write the equations for the backward induction in a more convenient form introduce the statistic

$$\pi_t := P(h = 0 \mid Y_1^t). \tag{3.4}$$

Let $P_i(y_1(t))$ be the probability density of $y_1(t)$ conditioned on $h=i$; define

$$q(y_1(t+1)\mid \pi_t) := \pi_t P_0(y_1(t+1)) + (1-\pi_t)\, P_1(y_1(t+1)) \qquad \forall t. \quad (3.5)$$

$$\phi(\pi_t,\ y_1(t+1)) := \pi_t P_0(y_1(t+1))/q(y_1(t+1)\mid \pi_t) \qquad \forall t. \quad (3.6)$$

A familiar argument using Bayes' rule gives the "updating" formulas

$$P(y_1(t+1)\mid Y_1^t) = q(y_1(t+1)\mid \pi_t) \qquad \forall t. \quad (3.7)$$

$$\pi_{t+1} = \phi(\pi_t,\ y_1(t+1)) \qquad \forall t. \quad (3.8)$$

With this notation we proceed to study a finite horizon problem.

3.2. *Finite Horizon*

Fix $T<\infty$ and consider the problem

$$\underset{\substack{u_1\in\{0,1\}\\ 1\leqslant \tau_1\leqslant T}}{\mathrm{Min}}\ E\mathbf{L}(u_1,\tau_1). \quad (3.9)$$

Define the operator ψ which transforms any function $W_{t+1}(\pi)$, $\pi\in[0,1]$, $t=0,1,2,\ldots$, into

$$[\psi W_{t+1}](\pi) = \int W_{t+1}(\phi(\pi,\ y_1(t+1))\, q(y_1(t+1)\mid \pi)\, dy_1(t+1), \quad (3.10)$$

and define the functions W_t^T by

$$W_T^T(\pi) = \min\{G_0(\gamma_2)\,\pi + K_0(\gamma_2),\ G_1(\gamma_2)\,\pi + K_1(\gamma_2)\} \quad (3.11)$$

$$W_t^T(\pi) = \min\{G_0(\gamma_2)\,\pi + K_0(\gamma_2),\ G_1(\gamma_2)\,\pi + K_1(\gamma_2),\ c + [\psi W_{t+1}^T](\pi)\}, \qquad t=1,2,\ldots,T-1, \quad (3.12)$$

where

$$G_i(\gamma_2) = \sum_{u_2} p(u_2\mid h_0)\, J(i,u_2,h_0) - \sum_{u_2} p(u_2\mid h_1)\, J(i,u_2,h_1) \qquad i=0,1 \quad (3.13)$$

$$K_i(\gamma_2) = \sum_{u_2} p(u_2\mid h_1)\, J(i,u_2,h_1) \qquad i=0,1. \quad (3.14)$$

A dynamic programming argument shows that W_t^T is the value function, that is,

$$W_t^T(\pi) = \min_{\substack{t \leqslant \tau_1 \leqslant T \\ u_1 \in \{0,1\}}} E\{c(\tau_1 - t) + J(u_1, u_2, h) | \pi_t = \pi\}$$
$$= \min_{\substack{t \leqslant \tau_1 \leqslant T \\ u_1 \in \{0,1\}}} E\{c(\tau_1 - t) + J(u_1, u_2, h) | \operatorname{Prob}(h = 0) = \pi\}$$
$$= \min_{\substack{t \leqslant \tau_1 \leqslant T \\ u_1 \in \{0,1\}}} E\{c(\tau_1 - t) + J(u_1, u_2, h) | Y_1^t\}. \tag{3.15}$$

The term $G_0(\gamma_2)\pi + K_0(\gamma_2)$, represents the cost due to stopping at a certain time t and deciding h_0. It is obtained by considering the cost $E\{J(u_1, u_2, h) | Y_1^t\}$ and setting $u_1 = 0$. Then $E\{J(0, u_2, h) | Y_1^t\} = \pi \sum_{u_2} p(u_2 | h_0)\ J(0, u_2, h_0) + (1 - \pi) \sum_{u_2} p(u_2 | h_1)\ J(0, u_2, h_1) = G_0(\gamma_2)\pi + K_0(\gamma_2)$. The term $G_1(\gamma_2)\pi + K_1(\gamma_2)$ represents the cost due to stopping at a certain time t and deciding h_1. It is obtained in exactly the same way as the cost due to stopping and deciding h_0. Finally, the term $c + [\psi W_{t+1}^T](\pi)$ represents the cost due to continuing at time t. It is optimal to stop at time t if and only if the cost due to stopping does not exceed the cost due to continuing, that is if and only if

$$\min_{i \in \{0,1\}} \{G_i(\gamma_2)\pi + K_i(\gamma_2)\} \leqslant c + [\psi W_{t+1}^T](\pi). \tag{3.16}$$

The properties of the optimal stopping rule of detector 1 for a fixed $\gamma_2 \in \Gamma_2$ are based on the following facts:

LEMMA 3.1. *$W_t^T(\pi)$ is a nonnegative concave function of π ($t = 1, 2, \ldots, T$).*

Proof. By (3.11) the assertion is true for $t = T$ as $W_T^T(\pi)$ is the minimum of two affine functions of π. Suppose that $W_{t+1}^T(\pi)$ is concave. Then it can be described as an envelope of a collection I of affine functions $\lambda_i \pi + \mu_i$, $i \in I$, where λ_i, μ_i are constants such that

$$W_{t+1}^T(\pi) = \inf_i \{\lambda_i \pi + \mu_i\}. \tag{3.17}$$

With this representation of $W_{t+1}^T(\pi)$,

$$[\psi W_{t+1}](\pi) = \int \inf_i \{\lambda_i \phi(\pi, y_1(t+1)) + \mu_i\}\ q(y_1(t+1) | \pi)\ dy_1(t+1)$$
$$= \int \inf_i \left\{\lambda_i \frac{\pi P_1(y_1(t+1))}{q(y_1(t+1) | \pi)} + \mu_i\right\} q(y_1(t+1) | \pi)\ dy_1(t+1)$$
$$= \int \inf_i \{\lambda_i P_1(y_1(t+1) | \pi)\pi + \mu_i q(y_1(t+1) | \pi)\}\ dy_1(t+1). \tag{3.18}$$

Consequently $[\psi W_{t+1}](\pi)$ is concave since the term within $\{\ \}$ is affine in π. From (3.14) it follows that $W_t^T(\pi)$ is concave, as it is the minimum of two affine and one concave function of π. ∎

LEMMA 3.2. *At $\pi = 0$ and $\pi = 1$ the following inequalities hold for all t ($t = 1, 2, \ldots, T-1$)*

$$\min_{i \in \{0,1\}} \{G_i(\gamma_2)\,\pi + K_i(\gamma_2)\}|_{\pi=0} < c + [\psi W_{t+1}^T](\pi)|_{\pi=0} \tag{3.19}$$

$$\min_{i \in \{0,1\}} \{G_i(\gamma_2)\,\pi + K_i(\gamma_2)\}|_{\pi=1} < c + [\psi W_{t+1}^T](\pi)|_{\pi=1}. \tag{3.20}$$

Moreover,

$$[\psi W_{t+1}^T](\pi) \geqslant [\psi W_t^T](\pi) \qquad \text{for all} \quad t = 1, 2, \ldots, T. \tag{3.21}$$

Proof. Equations (3.19)–(3.20) follow directly from the definitions of $W_t^T(\pi)$ and $[\psi W_t^T](\pi)$ for $t = 1, 2, \ldots, T$. To prove (3.21) note that for all t

$$W_t^T(\pi) < W_{t+1}^T(\pi) \tag{3.22}$$

(because the set of stopping times increases as the horizon increases). The last inequality and (3.10) prove (3.21). ∎

The threshold property of the optimal stopping rule of detector 1 for fixed $\gamma_2 \in \Gamma_2$ follows from Lemmas 3.1 and 3.2.

THEOREM 3.1. *For fixed $\gamma_2 \in \Gamma_2$ the optimal stopping rule of detector 1 is described by thresholds $m_T, \alpha_1, \beta_1, \alpha_2, \beta_2, \ldots, \alpha_{T-1}, \beta_{T-1}$. The optimal stopping time for detector 1 is*

$$\tau_1 = \min\{t : \alpha_t \geqslant \pi_t \text{ or } \beta_t \leqslant \pi_t\}. \tag{3.23}$$

Proof. For $t = T$ the threshold property of the optimal stopping rule follows from (3.11). The threshold m_T is determined by the solution of the equation

$$G_0(\gamma_2)\,m_T + K_0(\gamma_2) = G_1(\gamma_2)\,m_T + K_1(\gamma_2). \tag{3.24}$$

For $t = 1, 2, \ldots, T-1$ the threshold property follows from (3.12) the concavity of $W_t^T(\pi)$, $[\psi W_{t+1}^T](\pi)$, and (3.19)–(3.21). The thresholds $\alpha_1, \beta_1, \alpha_2, \beta_2, \ldots, \alpha_{T-1}, \beta_{T-1}$ are defined by (3.16). More specifically $\alpha_1, \alpha_2, \ldots, \alpha_{T-1}$ are determined by the solution of

$$G_1(\gamma_2)\,\alpha_t + K_1(\gamma_2) = c + [\psi W_{t+1}^T](\alpha_t), \qquad t = 1, 2, \ldots, T-1, \tag{3.25}$$

and $\beta_1, \beta_2, ..., \beta_{T-1}$ are determined by the solution of

$$G_0(\gamma_2)\,\beta_t + K_0(\gamma_2) = c + [\psi W^T_{t+1}](\beta_t), \qquad t = 1, 2, ..., T-1 \quad (3.26)$$

(see Fig. 1). Detector 1 stops as soon as the cost due to stopping does not exceed the cost due to continuing; the cost due to stopping does not exceed the cost due to continuing if and only if $\pi_t \leqslant \alpha$ or $\pi_t \geqslant \beta$ (see Fig. 1). Hence τ_1 satisfies (3.23). ∎

3.3. *Infinite Horizon*

To minimize (3.2) take $T \to \infty$ in (3.9). So let W^T_t denote the value functions defined by (3.11), (3.12). Since the set of stopping times τ_1, $\{\tau_1 \leqslant T\}$, increases with T it follows that $W^{T+1}_t(\pi) \leqslant W^T_t(\pi)$, therefore the following limit is defined:

$$W_t(\pi) = \lim_{T \to \infty} W^T_t(\pi) = \inf_T W^T_t(\pi) = W(\pi). \quad (3.27)$$

The last equality in (3.27) follows from (3.15); for all t, by a time-shift, we can obtain W_t by

$$\min_{\tau_1 \geqslant 0} E\{c\tau_1 + J(u_1, u_2, h) \mid \text{Prob}(h=0) = \pi\}.$$

It is possible to extend the results of Section 3.2 to obtain the following properties of $W(\pi)$ and the optimal stopping rule of detector 1 for the infinite horizon problem:

THEOREM 3.2. *The value function $W(\pi)$ is a nonnegative concave function of π which satisfies the equation*

$$W(\pi) = \min\{G_0(\gamma_2)\,\pi + K_0(\gamma_2),\ G_1(\gamma_2)\,\pi + K_1(\gamma_2),\ c + [\psi W](\pi)\}. \quad (3.28)$$

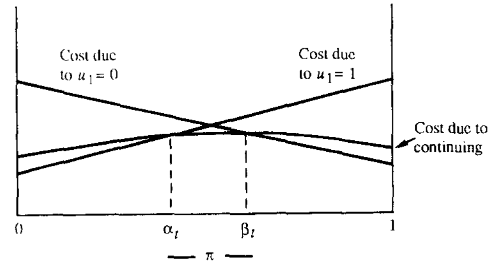

FIGURE 1

The optimal stopping rule of detector 1 *is characterized by thresholds* α, β *which are determined by*

$$c + [\psi W](\pi) = \min_{i \in \{0,1\}} \{G_i(\gamma_2)\, \pi + K_i(\gamma_2)\}. \tag{3.29}$$

The optimal stopping time for detector 1 *is*

$$\tau_1 = \min\{t : \pi \leqslant a \text{ or } \pi \geqslant \beta\}. \tag{3.30}$$

Proof. The nonnegativity and concavity of $W(\pi)$ follow from Lemma 3.1. Equation (3.28) follows from (3.12). Inequalities similar to (3.19)–(3.20) also hold because of Lemma 3.2. The threshold property of the optimal decision rule of detector 1 follows from (3.28), the concavity of $[\psi W](\pi)$, and (3.19)–(3.20) (see Fig. 1). The thresholds α, β are determined by (3.29). The threshold α is determined by the solution of

$$c + [\psi W](\alpha) = G_1(\gamma_2)\, \alpha + K_1(\gamma_2) \tag{3.31}$$

and the threshold β is determined by the solution of

$$c + [\psi W](\beta) = G_0(\gamma_2)\, \beta + K_0(\gamma_2). \tag{3.32}$$

Finally (3.30) can be obtained in exactly the same way as (3.23). ∎

It is interesting to note the uniqueness of the solution to (3.23).

LEMMA 3.3. *The value function* $W(\pi)$ *gives the unique solution to*

$$V(\pi) = \min\{G_0(\gamma_2)\, \pi + K_0(\gamma_2), G_1(\gamma_2)\, \pi + K_1(\gamma_2), c + [\psi V](\pi)\}. \tag{3.33}$$

Proof. To show that $V(\pi) \leqslant W(\pi)$ consider $W_t^T(\pi)$, $t \leqslant T$. Then, from (3.11) we get

$$W_T^T(\pi) \geqslant V(\pi).$$

Suppose

$$W_{t+1}^T(\pi) \geqslant V(\pi).$$

Then

$$\begin{aligned} W_t^T(\pi) = \min\{&G_0(\gamma_2)\, \pi + K_0(\gamma_2), G_1(\gamma_2)\, \pi \\ &+ K_1(\gamma_2), c + [\psi W_{t+1}^T](\pi)\} \\ \geqslant \min\{&G_0(\gamma_2)\, \pi + K_0(\gamma_2), G_1(\gamma_2)\, \pi + K_1(\gamma_2), \\ &c + [\psi V](\pi)\} = V(\pi). \end{aligned} \tag{3.34}$$

The inequality in (3.34) follows from the fact that $f_1 \geqslant f_2$ implies $\psi f_1 \geqslant \psi f_2$. Letting $T \to \infty$ proves $W(\pi) \geqslant V(\pi)$. To show $V(\pi) \geqslant W(\pi)$ fix π_t and define the stopping time $\tau \geqslant t$ by

$$\tau = \min\{s \geqslant t \mid \min_i \{G_i(\gamma_2)\,\pi_s + K_i(\gamma_2)\} \leqslant c + [\psi V](\pi_s)\}. \tag{3.35}$$

Then

$$\begin{aligned} V(\pi_t) &= c + E\{V(\pi_{t+1}) \mid Y_1^t\} \\ V(\pi_{t+1}) &= c + E\{V(\pi_{t+2}) \mid Y_1^{t+1}\} \\ &\vdots \\ V(\pi_{\tau-1}) &= c + E\{V(\pi_\tau) \mid Y_1^{\tau-1}\} \\ V(\pi_\tau) &= \min_{i \in \{0,1\}} \{G_i(\gamma_2)\,\pi_\tau + K_i(\gamma_2)\}. \end{aligned}$$

Adding and taking expectations conditioned on Y_1^t gives

$$\begin{aligned} V(\pi_t) &= \min_{i \in \{0,1\}} E\{c(\tau - t) + G_i(\gamma_2)\,\pi_\tau + K_i(\gamma_2) \mid Y_1^t\} \\ &= \min_{u_1 \in \{0,1\}} E\{c(\tau - t) + J(u_1, u_2, h) \mid Y_1^t\} \\ &\geqslant W(\pi_t), \end{aligned} \tag{3.36}$$

since

$$W(\pi_t) = \min_{\substack{u_1 \in \{0,1\} \\ \tau_1}} E\{c(\tau_1 - t) + J(u_1, u_2, h) \mid Y_1^t\}. \quad \blacksquare \tag{3.37}$$

For fixed $\gamma_2 \in \Gamma_2$ the analysis of detector 1's problem is now complete. Based on the analysis above we can conclude the following about the member-by-member optimal (mbmo) solutions of the decentralized Wald problem.

3.4. *Qualitative Properties of the mbmo Solutions of the Decentralized Wald Problem*

THEOREM 3.3. *The mbmo stopping rules of the detectors are characterized by time-invariant thresholds α^{1*}, β^{1*}, α^{2*}, β^{2*}. These thresholds are coupled and their computation requires the solution of the following coupled sets of dynamic programming equations:*

$$\alpha^{i*} F_1(\alpha^{j*}, \beta^{j*}) + Q_1(\alpha^{j*}, \beta^{j*}) = c + [\psi W^i](\alpha^{i*}), \tag{3.38}$$

$$\beta^{i*} F_0(\alpha^{j*}, \beta^{j*}) + Q_0(\alpha^{j*}, \beta^{j*}) = c + [\psi W^i](\beta^{i*}), \tag{3.39}$$

where $i \neq j$, $i, j = 1, 2$;

$$F_l(\alpha^{j*}, \beta^{j*}) = G_l(\gamma_j^*), \qquad j = 1, 2, \ l = 0, 1, \tag{3.40}$$

$$Q_l(\alpha^{j*}, \beta^{j*}) = K_l(\gamma_j^*), \qquad j = 1, 2, \ l = 0, 1. \tag{3.41}$$

$G_l(\gamma_j^*)$, $K_l(\gamma_j^*)$ are given by (3.13) and (3.14), respectively, and W^i refers to the value function of detector i. The optimal stopping times of the detectors have the property

$$\tau_i^* = \min\{t : \pi_t^i \leqslant \alpha^{i*} \text{ or } \pi_\tau^i \geqslant \beta^{i*}\}, \qquad i = 1, 2. \tag{3.42}$$

Proof. Since Theorem 3.2 holds for any stopping rule $\gamma_2 \in \Gamma_2$ of the second detector, it also holds for a mbmo γ_2^* (the existence of such γ_2^*'s will be discussed below). Thus, the mbmo stopping rules of the first detector are characterized by thresholds α^{1*}, β^{1*}. By symmetry, the mbmo stopping rules of the second detector are also characterized by thresholds α^{2*}, β^{2*}. These thresholds are coupled because the terms $G_i(\gamma_j^*)$ and $K_i(\gamma_j^*)$ that appear in the dynamic program of detector i ($i, j = 1, 2$, $i \neq j$) depend on the decision u_j^* of detector j which in turn depend on the thresholds α^{j*}, β^{j*}. Hence (3.38)–(3.41) result from the argument above, (3.29), and the properties of the value function $W(\pi)$ described by Lemmas 3.1 and 3.2. The property of the optimal stopping times τ_1^*, τ_2^* can be obtained by arguments similar to those of Theorems 3.1 and 3.2. ∎

Remarks. 1. It should be clear that the thresholds that satisfy (3.38) and (3.39) guarantee only member-by-member optimality. To prove that member-by-member optimal solutions exist one may argue as follows: Define stopping times $\{\tau_1(n) : n \geqslant 1\}$, $\{\tau_2(n) : n \geqslant 1\}$ and sequences $\{G_0^1(n), G_1^1(n), K_0^1(n), K_1^1(n), W^1(n)\}$, $\{G_0^2(n), G_1^2(n), K_0^2(n), K_1^2(n), W^2(n)\}$, for $n \geqslant 1$ recursively as follows: Define $G_i^1(n)$, $K_i^1(n)$, $i = 0, 1$, as the functions $t \to G_{it}^1(n) = G_i^1(n)$ and $t \to K_{it}^1(n) = K_i^1(n)$ in (3.13) and (3.14), respectively, using the rule γ_2 defined by $\tau_2(n)$. Let $W^1(n)$ be the value function $(\pi, t) \to W_t^1(\pi) = W^1(\pi)$ defined from the functions $G_i^1(n)$, $K_i^1(n)$, $i = 0, 1$, above and (3.28). Using (3.28), let $\tau_1(n)$ be the stopping time defined by $W^1(n)$ and $G_i^1(n)$, $K_i^1(n)$, $i = 0, 1$. For detector 2 define $G_i^2(n+1)$, $K_i^2(n+1)$, $i = 0, 1$, as the functions $t \to G_{it}^2(n+1) = G_i^2(n+1)$ and $t \to K_{it}^2(n+1) = K_i^2(n+1)$, in the same way as in (3.13) and (3.14), respectively, using the rule γ_1 defined by $\tau_1(n)$ Similarly, for detector 2 let $W^2(n+1)$ be the value function $(\pi, t) \to W_t^2(\pi) = W^2(\pi)$ defined from $G_i^2(n+1)$, $K_i^2(n+1)$, $i = 0, 1$, and an equation similar to (3.28). Define $\tau_2(n+1)$ as the stopping time resulting from $W^2(n+1)$, $G_i^2(n+1)$, $K_i^2(n+1)$, $i = 0, 1$. Note that $t \to G_{it}^j(n)$, $t \to K_{it}^j(n)$, and $t \to W_t^j(n)$, $i = 0, 1$, $j = 1, 2$, have a compact range. Sequential compactness of $G_i^1(n)$, $G_i^2(n)$, $K_i^1(n)$, $K_i^2(n)$, $i = 0, 1$, $W^1(n)$ and $W^2(n)$ then follow from Tychonoff's theorem (Kelly, 1975). Consequently, there will be a subsequence along which $G_i^1(n)$, $G_i^2(n)$, $K_i^1(n)$, $K_i^2(n)$, $i=0, 1$, $W^1(n)$, $W^2(n)$ converge. These limit functions define a person-by-person optimal pair.

2. When the finite horizon decentralized Wald problem is considered the mbmo thresholds of the detectors are time varying. In this case, if T is the horizon, one has to solve $4T-2$ nonlinear algebraic equations of the form (3.24) (for time T) and (3.38)–(3.39) (for time $t=1, 2,..., T-1$) in $4T-2$ unknowns (the thresholds) to determine the mbmo stopping rules of the detectors.

3. The results presented in this section hold for the case where there are two hypotheses h_0, h_1 and M detectors ($M>2$) coupled through their common cost.

So far we have determined the *qualitative properties* of the mbmo stopping rules for the decentralized Wald problem. To compute the mbmo thresholds for the infinite horizon problem we have to solve a coupled set of equations like (3.38)–(3.39) to determine α^{1*}, β^{1*}, α^{2*}, β^{2*}. In the next section we present an approximate computation of the thresholds when the observation noise for both detectors has Gaussian statistics and discuss the features of the solution.

4. An Approximate Computation of the Thresholds

Consider the decentralized Wald problem formulated in Section 2 and assume that:

1. The detectors' observations are described by

$$y_i(t)=h+w_i(t), \tag{4.1}$$

where $\{w_i(t)\}$ ($i=1, 2$) are zero mean white Gaussian noise sequences with variance σ; $\{w_1(t)\}$ and $\{w_2(t)\}$ are independent of each other and independent of the hypothesis h.

2. The cost $J(u_1, u_2, h)$ incurred by the decisions u_1, u_2 of the detectors is

$$J(u_1, u_2, h)=\begin{cases} 0 & \text{if } u_1=u_2=h \\ 1 & \text{if } u_1\neq u_2 \\ k & \text{if } u_1=u_2\neq h, k>1, k<\infty. \end{cases} \tag{4.2}$$

In this section we propose an approximate solution of the decentralized Wald problem with observations and terminal cost given by (4.1)–(4.2). To achieve this solution we combine the main results of Section 3 with results from standard sequential analysis.

The idea of the solution is the following: Let δ_1 (resp. ε_1) be the probability of error type 1 for detector 1 (resp. detector 2) (that is, the probability that if $h=0$ detector 1(2) will declare $h=1$); similarly let δ_2 (resp. ε_2) be the probability of error of type 2 for detector 1 (resp. detector 2) (that is, the probability that if $h=1$ is true detector 1 (2) will declare $h=0$). We shall write the cost (2.7) as a function of these four quantities and then we shall minimize the cost jointly over $\delta_1, \delta_2, \varepsilon_1, \varepsilon_2$. After $\delta_1, \delta_2, \varepsilon_1, \varepsilon_2$ are determined, standard results from statistical sequential analysis will be used to determine the thresholds for the two detectors, and the final decisions u_1, u_2 of the detectors will be determined graphically.

From statistical sequential analysis (Chernoff, 1972; Wald, 1947) it is known that the average number of observations required to reach a decision with errors δ_1 and δ_2 is approximately

$$\bar{\eta}^1(0) = -2\sigma\left[\delta_1 \log\frac{1-\delta_2}{\delta_1} + (1-\delta_1)\log\frac{\delta_2}{1-\delta_1}\right] \tag{4.3}$$

when the hypothesis h_0 is true, and

$$\bar{\eta}^1(1) = 2\sigma\left[(1-\delta_2)\log\frac{1-\delta_2}{\delta_1} + \delta_2\log\frac{\delta_2}{1-\delta_1}\right] \tag{4.4}$$

when the hypothesis h_1 is true. Relations similar to (4.3) and (4.4) hold for detector 2 with ε_1 and ε_2 in place of δ_1 and δ_2, respectively. Using (4.3), (4.4), and (2.1) we can approximately write the cost to be minimized as

$$\begin{aligned}
&E\{c\tau_1 + c\tau_2 + J(u_1, u_2, h)\}\\
&\quad= 2\sigma c(1-p)\left[(1-\delta_2)\log\frac{1-\delta_2}{\delta_1} + \delta_2\log\frac{\delta_2}{1-\delta_1} + \varepsilon_2\log\frac{\varepsilon_2}{1-\varepsilon_1}\right.\\
&\quad\quad \left. + (1-\varepsilon_2)\log\frac{1-\varepsilon_2}{\varepsilon_1}\right] - 2\sigma cp\left[\delta_1\log\frac{1-\delta_2}{\delta_1} + (1-\delta_1)\log\frac{\delta_2}{1-\delta_1}\right.\\
&\quad\quad \left. + \varepsilon_1\log\frac{1-\varepsilon_2}{\varepsilon_1} + (1-\varepsilon_1)\log\frac{\varepsilon_2}{1-\varepsilon_1}\right] + (1-\delta_2)\,\varepsilon_2(1-p)\\
&\quad\quad + \delta_2(1-\varepsilon_2)(1-p) + \delta_1(1-\varepsilon_1)\,p + (1-\delta_1)\,\varepsilon_1\,p + k\delta_1\varepsilon_1\,p\\
&\quad\quad + k\delta_2\varepsilon_2(1-p) := L(\delta_1, \delta_2, \varepsilon_1, \varepsilon_2).
\end{aligned} \tag{4.5}$$

Note that $L(\delta_1, \delta_2, \varepsilon_1, \varepsilon_2)$ is a nonconvex function of $\delta_1, \delta_2, \varepsilon_1, \varepsilon_2$ so that the minimization of $L(\cdot)$ with respect to $\delta_1, \delta_2, \varepsilon_1, \varepsilon_2$ can only guarantee a local minimum. Let δ_1^*, δ_2^*, ε_1^*, ε_2^* correspond to a local minimum. Then the definitions

$$A_1 = \log \frac{1-\delta_2^*}{\delta_1^*} \tag{4.6}$$

$$A_2 = \log \frac{\delta_2^*}{1-\delta_1^*} \tag{4.7}$$

$$B_1 = \log \frac{1-\varepsilon_2^*}{\varepsilon_1^*} \tag{4.8}$$

$$B_2 = \log \frac{\varepsilon_2^*}{1-\varepsilon_1^*} \tag{4.9}$$

from standard sequential analysis (Chernoff, 1972; Wald, 1947) can be used to compute the mbmo thresholds of the detectors. Afterwards, the decisions u_1, u_2 of the dectectors can be determined graphically as follows (Wald, 1947): At any time t the sum $S = \sum_{s=1}^{t} y_1(s)$ of the observations up to that time is a sufficient statistic for detector 1. As long as this sum remains between the two parallel lines l_2, l_2 (Fig. 2) detector 1 continues to take measurements. The first instant of time the sum S is above l_1 or below l_2 detector 1 stops and accepts h_1 if S is above l_1 and $h=0$ if S lies below l_2. Similar results hold for detector 2.

The thresholds A_1, A_2, B_1, B_2 determined by (4.6)–(4.9), result when the log likelihood ratio $\log(p(h_1 \mid y_i^t)/p(h_0 \mid y_i^t)) = \log(1-\pi/\pi)$ is used as a sufficient statistic for decision making instead of π. Thus, the thresholds

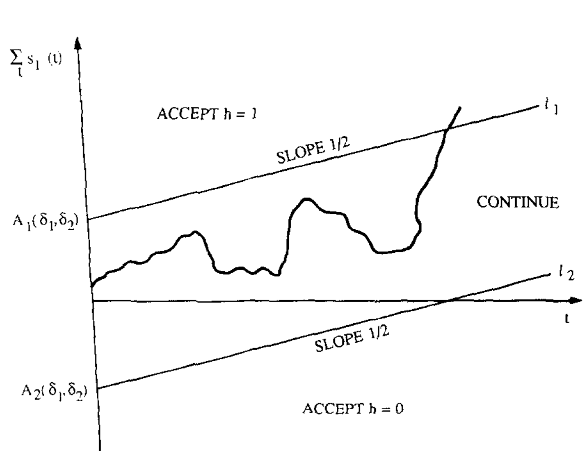

FIG. 2. Detector 1's decision rule.

A_1, A_2, B_1, B_2 are related to the thresholds a^{1*}, a^{2*}, β^{1*}, β^{2*} defined in Section 3 by

$$A_1 = \log \frac{1-\beta^{1*}}{\beta^{1*}}, \quad A_2 = \log \frac{1-\alpha^{1*}}{\alpha^{1*}}, \qquad B_1 = \log \frac{1-\beta^{2*}}{\beta^{2*}}, \quad B_2 = \log \frac{1-\alpha^{2*}}{\alpha^{2*}}.$$

Note that the mbmo thresholds are coupled because they depend on δ_1^*, δ_2^*, ε_1^*, ε_2^*, which are determined by joint optimization for the two detectors. The optimization problem whose solution determines δ_1^*, δ_2^*, ε_1^*, ε_2^* is simple as it only requires the minimization of (4.5) with respect to $\delta_1, \delta_2, \varepsilon_1, \varepsilon_2$; furthermore the numerical results of the next section obtained by the approach proposed here are intuitively appealing. The only approximation in the proposed solution appears in Eqs. (4.3) and (4.4). These equations are derived by considering the log-likelihood ratio $\log(p(h_1 \mid y_i^t)/p(h_0 \mid y_i^t))$ $(i = 1, 2)$ as a sufficient statistic for decision making for each detector. When the sequential process is terminated and a decision is reached by detector 1 it is assumed that if $u_1 = 1$ then the value of $\log(p(h_1 \mid y_1^t)/p(h_0 \mid y_1^t)) = A_1$; if $u_1 = 0$ then it is assumed that the value of $\log(p(h_1 \mid y_1^t)/p(h_0 \mid y_1^t)) = A_2$. Similar assumptions hold for detector 2; that is, if $u_2 = 1$ then $\log(p(h_1 \mid y_2^t)/p(h_0 \mid y_2^t)) = B_1$ and if $u_2 = 0$ then $\log(p(h_1 \mid y_2^t)/p(h_0 \mid y_2^t)) = B_2$. Since the excess of $\log(p(h_1 \mid y_i^t)/p(h_0 \mid y_i^t))$ over the thresholds A_1, A_2, B_1, B_2, is neglected when the sequential process is terminated, (4.3) and (4.4) are only approximate expressions for the average number of observations. A detailed derivation of (4.3) and (4.4) as well as a more complete discussion about the computation of $\bar{\eta}^1(0)$ and $\bar{\eta}^1(1)$ is given in (Wald, 1947, Chap. 3.5 and Appendix A.3).

5. Numerical Results

In this section we present the numerical results obtained by the implementation of the solution approach proposed in Section 4. The probabilities of error $\delta_1, \delta_2, \varepsilon_1, \varepsilon_2$ as well as the thresholds $A_1(\delta_1, \delta_2)$, $A_2(\delta_1, \delta_2)$, $B_1(\varepsilon_1, \varepsilon_2)$, $B_2(\varepsilon_1, \varepsilon_2)$ are computed for various values of the following parameters:

1. The prior probability $p = \text{Prob}(h = 0)$.
2. The variance σ of the observation noise.
3. The cost c of the observations.
4. The penalty k arising when both detectors' decisions are wrong.

We present each one of our parametric studies separately and interpret the results obtained by these studies. As pointed out in Section 4 the cost

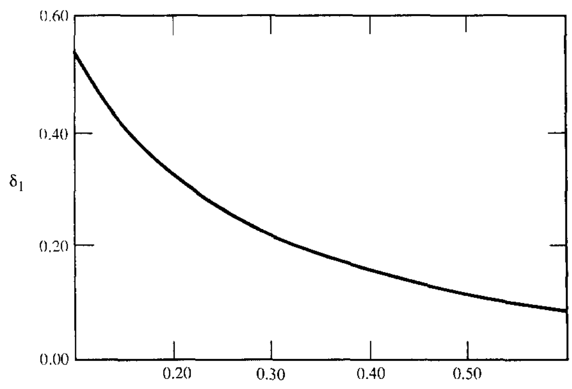

FIG. 3. Type 1 error versus p ($k = 4$, $c = 0.1$, $\sigma = 0.5$).

$L(\delta_1, \delta_2, \varepsilon_1, \varepsilon_2)$ is a nonconvex function of $(\delta_1, \delta_2, \varepsilon_1, \varepsilon_2)$, consequently the values of $(\delta_1, \delta_2, \varepsilon_1, \varepsilon_2)$ determined by the minimization of (4.5) correspond to local minima. The result of the minimization of (4.5) depends on the initial guess of $(\delta_1, \delta_2, \varepsilon_1, \varepsilon_2)$. Some of the local minima of (4.5) result in $\delta_1 = \varepsilon_1$ and $\delta_2 = \varepsilon_2$. Such local minima are obtained when the minimization of (4.5) is initiated with $\delta_1^{in} = \delta_2^{in} = \varepsilon_1^{in} = \varepsilon_2^{in}$. The numerical results we present below correspond to minima for which $\delta_1 = \varepsilon_1$, $\delta_2 = \varepsilon_2$.

5.1. *The Variation of* $\delta_1, \delta_2, \varepsilon_1, \varepsilon_2$, $A_1(\delta_1, \delta_2)$, $A_2(\delta_1, \delta_2)$, $B_1(\varepsilon_1, \varepsilon_2)$, $B_2(\varepsilon_1, \varepsilon_2)$ *as a Function of* p

Figures 3 and 4 present the variation of the proabilities of error of types 1 and 2 as a function of p for fixed c, k, σ. These figures show that as

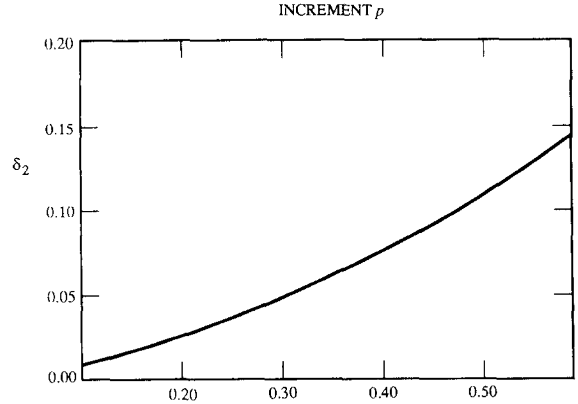

FIG. 4. Type 2 error versus p ($k = 4$, $c = 0.1$, $\sigma = 0.5$).

p increases, δ_1 and ε_1 decrease whereas δ_2 and ε_2 increase. Such a variation of $\delta_1, \delta_2, \varepsilon_1, \varepsilon_2$ (as a function of p) is also predicted by the qualitative properties of the mbmo stopping rules. As p increases the probability of the set of measurements y_1^t (y_2^t) that would cause p to drop below β_1^* (β_2^*) decreases, thus decreasing the probability of error of type 1. On the contrary, as p increases the probability of the set of measurements that would result in $\pi > \alpha_1^*$ (α_2^*) increases, thus increasing the probability of error of type 2.

Figure 5 presents the variation of the thresholds $A_1(\delta_1, \delta_2)$, $A_2(\delta_1, \delta_2)$, $B_1(\varepsilon_1, \varepsilon_2)$, $B_2(\varepsilon_1, \varepsilon_2)$ as a function of p for fixed c, k, σ. The figure shows that as p increases the thresholds $A_1(\delta_1, \delta_2)$ ($B_1(\varepsilon_1, \varepsilon_2)$) and $A_2(\delta_1, \delta_2)$ ($B_2(\varepsilon_1, \varepsilon_2)$) increase. Such a behavior of the thresholds is also intuitively expected because as p increases each detector would be biased more and more towards declaring $h = 0$. Therefore, the area where $h = 0$ is accepted in Fig. 2 would get larger, and the area where $h = 1$ is accepted in Fig. 2 would get smaller. Consequently all thresholds (A_1, A_2, B_1, B_2) should increase.

5.2. *The Variation of $\delta_1, \delta_2, \varepsilon_1, \varepsilon_2$, $A_1(\delta_1, \delta_2)$, $A_2(\delta_1, \delta_2)$, $B_1(\varepsilon_1, \varepsilon_2)$, $B_2(\varepsilon_1, \varepsilon_2)$ as a Function of k*

Figure 6 shows the variation of the probabilities of error of type 1 and type 2 as a function of the terminal cost k, incurred by two errors, for fixed c, σ and $p = 0.5$ (when $p = 0.5$ some of the local minima result in $\delta_1 = \delta_2 = \varepsilon_1 = \varepsilon_2$ when the minimization of (4.5) is initiated with $\delta_1^{in} = \delta_2^{in} = \varepsilon_1^{in} = \varepsilon_2^{in}$; Fig. 6 presents such a local minimum). It is seen that as k increases the error probabilities $\delta_1, \delta_2, \varepsilon_1, \varepsilon_2$ decrease. Such a variation is also intuitively

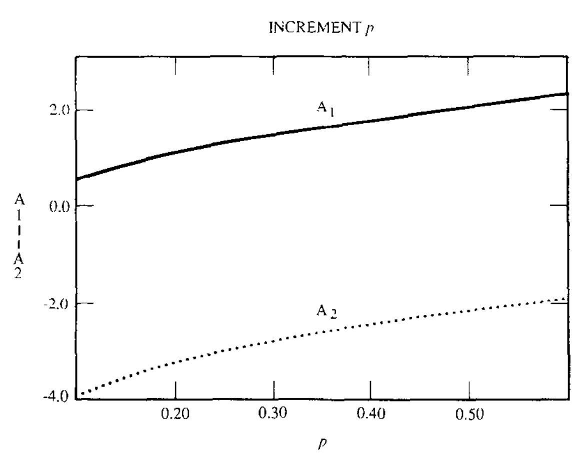

FIG. 5. Thresholds versus p ($k = 4$, $c = 0.1$, $\sigma = 0.5$).

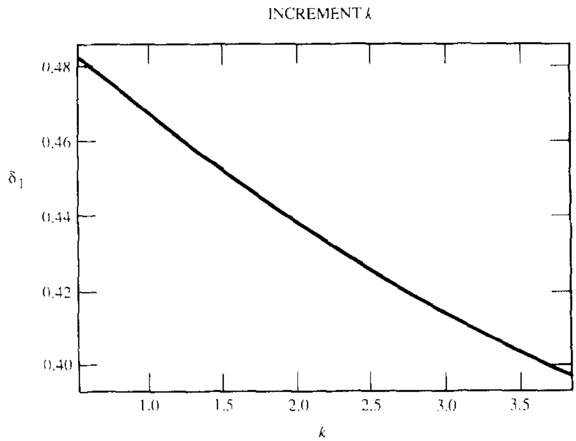

FIG. 6. Type 1 error versus k ($p = 0.5$, $c = 0.99$, $\sigma = 0.5$).

expected, because as k increases the detectors tend to become more conservative and more cautious, hence they tend to base their decisions on more reliable information. Thus, the probability of error decreases.

The variation of the thresholds A_1, A_2, B_1, B_2 as a function of k is shown in Fig. 7. Since the detectors become more conservative as k increases, the areas where $h = 0$ and $h = 1$ are accepted in Fig. 2 should get smaller. Consequently A_1 (B_1) should increase and A_2 (B_2) should decrease. This behavior is indeed shown by Fig. 7.

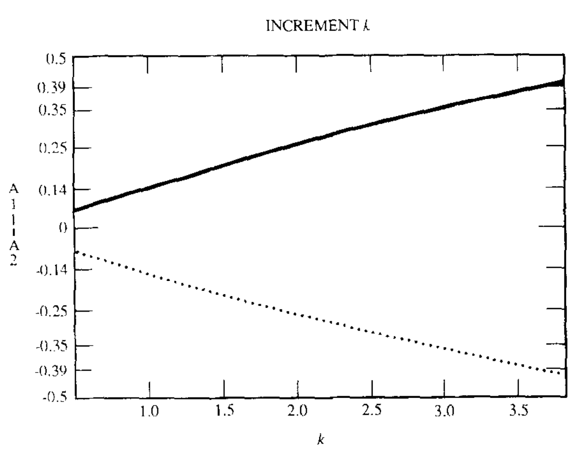

FIG. 7. Threshold versus k ($p = 0.5$, $c = 0.99$, $\sigma = 0.5$).

5.3. *The Variation of* $\delta_1, \delta_2, \varepsilon_1, \varepsilon_2$, $A_1(\delta_1, \delta_2)$, $A_2(\delta_1, \delta_2)$, $B_1(\varepsilon_1, \varepsilon_2)$, $B_2(\varepsilon_1, \varepsilon_2)$ *as a Function of* c

Figure 8 shows the variation of the probabilities of error of type 1 and 2 as a function of the cost c of observations for fixed σ, k and $p = 0.5$ (as pointed out before, when $p = 0.5$ some of the local minima result in $\delta_1 = \delta_2 = \varepsilon_1 = \varepsilon_2$ when the minimization of (4.5) is initiated with $\delta_1^{in} = \delta_2^{in} = \varepsilon_1^{in} = \varepsilon_2^{in}$, Fig. 8 presents such a local minimum). As the cost of observation increases, the detectors tend to take less observations before making a final decision, hence the quality of information, upon which the final decision is made gets worse with increasing c, and one would expect $\delta_1, \delta_2, \varepsilon_1, \varepsilon_2$ to increase with increasing c. This behavior is shown by Fig. 8.

The variation of the thresholds A_1, A_2, B_1, B_2 as a function of c is shown in Fig. 9. Since the detectors would tend to make a final decision more quickly as c increases, we would expect the areas of Fig. 2 where $h = 0$ and $h = 1$, are accepted to get larger with increasing c. Hence, we would expect the lower thresholds A_2, B_2 to increase and the upper thresholds A_1, B_1 to decrease. The variation of the thresholds shown by Fig. 9 confirms this intuition.

5.4. *The Variation of* $\delta_1, \delta_2, \varepsilon_1, \varepsilon_2$, $A_1(\delta_1, \delta_2)$, $A_2(\delta_1, \delta_2)$, $B_1(\varepsilon_1, \varepsilon_2)$ $B_2(\varepsilon_1, \varepsilon_2)$ *as a Function of* σ

Figure 10 shows the variation of the probabilities of error of types 1 and 2 as a function of the noise variance σ for fixed k, c, p. We set $p = 0{,}5$; then some of the local minima result in $\delta_1 = \delta_2 = \varepsilon_1 = \varepsilon_2$ when the minimization of (4.5) is initiated with $\delta_1^{in} = \delta_2^{in} = \varepsilon_1^{in} = \varepsilon_2^{in}$. Such minima are shown in Figs. 10 and 11. It is intuitively expected that as the noise variance σ

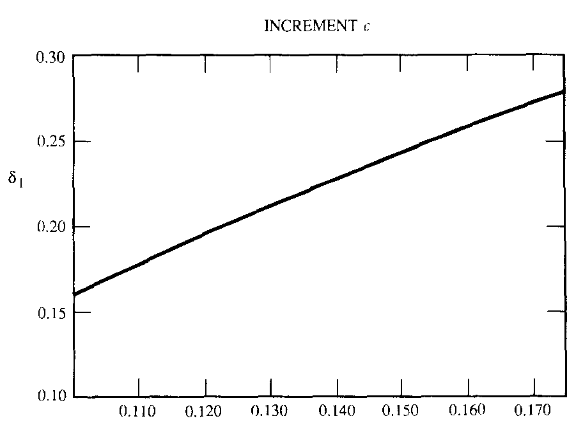

FIG. 8. Type 1 error versus c ($p = 0.5$, $k = 1$, $\sigma = 0.5$).

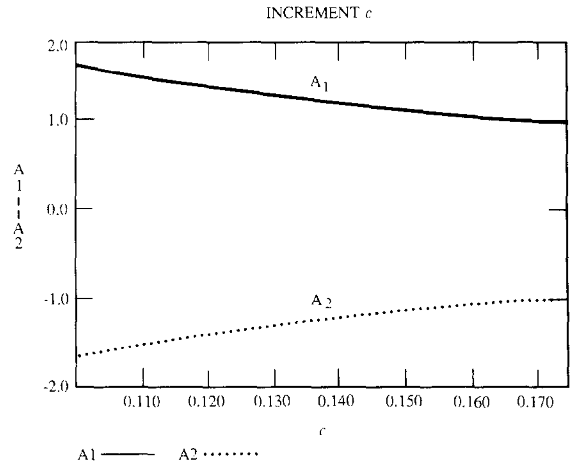

FIG. 9. Thresholds versus c ($p=0.5$, $k=1$, $\sigma=0.5$).

increases the quality of information of the detectors gets worse, thus $\delta_1, \delta_2, \varepsilon_1, \varepsilon_2$ increase. This behavior is actually shown in Fig. 10. Note that for $\sigma \geqslant 20$ the information from the observations is practically useless for the detectors.

The variation of the thresholds A_1, A_2, B_1, B_2 as a function of σ is shown in Fig. 11. As the quality of information received by the observations gets worse the detectors tend to rely more on their prior information, thus they tend to make decisions more quickly. Consequently, as σ increases the areas of Fig. 2 where $h=0$ and $h=1$ are accepted will get larger; hence the upper thresholds A_1 and B_1 will decrease and the lower

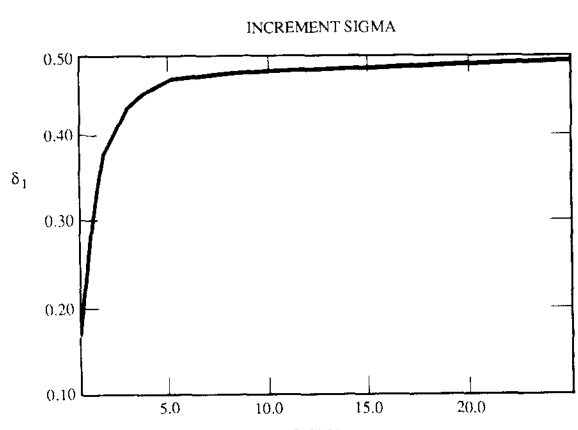

FIG. 10. Type 1 error versus σ ($p=0.5$, $k=1$, $c=0.1$).

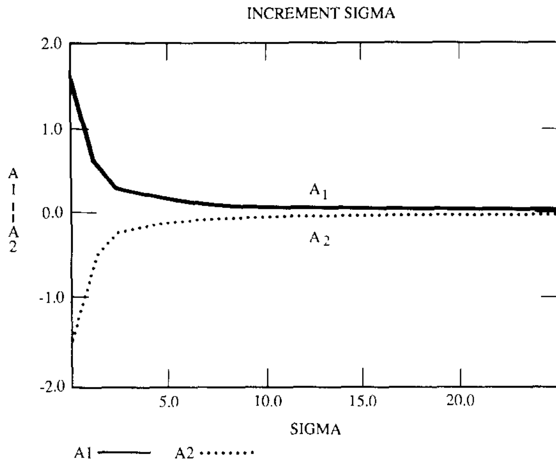

FIG. 11. Thresholds versus σ ($p=0.5$, $k=1$, $c=0.1$).

thresholds A_2 and B_2 will increase. This is seen in Fig. 11. Note, as before, that for $\sigma \geqslant 20$ the information from the observations is practically useless, therefore the thresholds $A_1(B_1)$ and $A_2(B_2)$ approach very close to each other because the detectors make decisions based practically on their prior information.

So far the numerical results presented in this section correspond to local minima for which $\delta_1=\varepsilon_1$, $\delta_2=\varepsilon_2$. There are local minima of (4.5) other than the symmetric ones. We present below such a local minimum.

5.5. *Nonsymmetric Member-by-Member Optimal Thresholds*

When the initial values of δ_1, δ_2, ε_1, ε_2 used in the minimization of (4.5) are $\delta_1^{in} \neq \delta_2^{in} \neq \varepsilon_1^{in} \neq \varepsilon_2^{in}$ then the resulting local minima of (4.5) and the corresponding mbmo thresholds are not symmetric. For example, for $p=0.9$, $k=4$, $c=0.05$, and initial guess,

$$\delta_1^{in}=0.2, \qquad \delta_2^{in}=0.5, \qquad \varepsilon_1^{in}=0.74, \qquad \varepsilon_2^{in}=0.3,$$

the resulting local minimum of (4.5) is

$$\delta_1=0.000164, \qquad \varepsilon_1=0.1150,$$
$$\delta_2=0.999457, \qquad \varepsilon_2=0.5694,$$

and the corresponding mbmo thresholds are

$$A_1=1.1972429, \qquad A_2=-0.0003791,$$
$$B_1=2.791737, \qquad B_2=-0.4410045.$$

6. Conclusions

In this paper we formulated a simple decentralized detection problem which is the decentralized version of Wald's problem. Even in this simple case the coupling induced by the cost structure causes considerable complexity in the computation of the optimal stopping rules. However, the qualitative properties of the mbmo stopping rules obtained in this paper suggest some simple approximate rules for the decentralized Wald problem. Such a simple approximate rule has been proposed in this paper: it was shown that results obtained by that rule are intuitively appealing.

Acknowledgments

This research was supported in part by DOE Contract DEAC01-80-RA50418, AFOSR Contract F49620-81-C-0015, and by a Rackham Research Grant from the Graduate School of the University of Michigan. The programming effort of Mr. Steven M. Basile is gratefully acknowledged.

Received December 1985; accepted July 15, 1986

References

Bertsekas, D. P. (1976), "Dynamic Programming and Stochastic Control," Academic Press, New York.

Chernoff, H. (1972), "Sequential Analysis and Optimal Design," Soc. Indus. Appl. Math., Philadelphia.

Ho, Y. C. (1980), Team decision theory and information structures, *Proc. IEEE* **68**, No. 6.

Kelly, J. L. (1975), "General Topology," Springer-Verlag, Berlin/New York.

Kushner, H. J. and Pacut A. (1982), A simulation study of a decentralized detection problem, *IEEE Trans. Automat. Control* **AC-27**, No. 5.

Lauer, G. S. and Sandell, N. R., Jr. (1982), "Decentralized Detection Given Waveform Observations,"Technical Paper TP-131, ALPHATECH, Inc., Burlington, MA.

Teneketzis, D. (1982), The decentralized Wald problem, *in* "Proceedings, Large-Scale Systems Symposium, Virginia Beach, VA.

Teneketzis, D., and Sandell N. R., Jr. (1985), The decentralized disruption problem with linear penalties on false alarms, *Large Scale Systems*, in press.

Teneketzis, D. and Varaiya, P. (1984), The decentralized quickest detection problem, *IEEE Trans. Automat. Control* **AC-29**, No. 7.

Tenney, R. R., and Sandell, N. R., Jr. (1981), Detection with distributed sensors, *IEEE Trans. Aerospace Electron. Systems* **AES-17**, No 4.

Van Trees, H. (1969), "Detection Estimation and Modulation Theory, Vols. I–III," Wiley, New York.

Wald, A. (1947), "Sequential Analysis," Wiley, New York.

Yoshikawa, T. (1978), Decomposition of dynamic team decision problem, *IEEE Trans. Automat. Control* **AC-23**, No. 4.