

Decentralized Detection with Signaling

Ashutosh Nayyar and Demosthenis Teneketzis

Abstract— We consider a sequential problem in decentralized detection. Two observers can make repeated noisy observations of a binary hypothesis on the state of the environment. At any time, any of the two observers can stop and send a final message to the other observer or it may continue to take more measurements. After an observer has sent its final message, it stops operating. The other observer is then faced with a different stopping problem. At each time instant, it can decide either to stop and declare a final decision on the hypothesis or take another measurement. At each time, the system incurs an operating cost depending on the number of observers that are active at that time. A terminal cost that measures the accuracy of the final decision is incurred at the end. We show that, unlike in other sequential detection problems, stopping rules characterized by two thresholds on an observer’s posterior belief no longer guarantee optimality in this problem. Thus the potential for signaling among observers alters the nature of optimal policies. We obtain a new parametric characterization of optimal policies for this problem.

I. INTRODUCTION

Decentralized detection problems are motivated by applications in large scale decentralized systems such as sensor networks, power systems and surveillance networks. In such networks, sensors receive different information about the environment but share a common objective, for example to decide if a fault has occurred or not in a power system, or to detect the presence of a target in a surveillance area. Sensors may be allowed to communicate but they are constrained to exchange only a limited amount of information because of energy constraints, data storage and data processing constraints, communication constraints etc.

Decentralized detection problems may be static or sequential. In static problems, sensors make a fixed number of observations about a hypothesis on the state of the environment which is modeled as a random variable H . Sensors may transmit a single message (a quantized version of their observations) to a fusion center which makes a final decision on H . Such problems have been extensively studied since their initial formulation in [1] (See the surveys in [2], [3] and references therein). In most such formulations, it has been shown that person-by-person optimal decision rules (as defined in [4]) for a binary hypothesis detection problem are characterized by thresholds on the likelihood ratio (or equivalently on the posterior belief on the hypothesis).

In sequential problems, the number of observations taken by the sensors is not fixed a priori. In the centralized sequential detection problems, as formulated in [5], a sensor can sequentially make costly observations and, after each

observation, can choose whether to stop and declare its final decision on H or to take more observations. In the decentralized analogue of the sequential detection problem, two (or more) sensors locally decide when to stop taking more measurements and then make a final decision on H . Each sensor pays a penalty for delaying its final decision and a terminal cost that depends on the final decisions of all the sensors and the true value of H is incurred at the end. A version of this problem (called the decentralized Wald problem) was formulated in [6] and it was shown that at each time instant, optimal policies for the sensors are described by two thresholds. The computation of these thresholds requires solution of two coupled sets of dynamic programming equations. Similar results were obtained in a continuous time setting in [7].

A key feature of the decentralized Wald problem is that the individual sensors do not communicate their decisions to each other. That is, the i^{th} sensor is not aware of decisions of other sensors. This implies that if policies of all other sensors are fixed, the i^{th} sensor is faced with a classical sequential detection problem for which two-threshold policies are optimal. In the problem we consider in this paper, each sensor observes the other sensor’s decisions. Hence, in addition to its own measurements of H , the i^{th} sensor can use the decisions made by other sensors (whether they have stopped or not and whether the final decision was 0 or 1) to make its decisions. The final decision of the sensor that stops in the end is taken as the final decision made by the group of sensors. Thus, sensors can convey information to each other through their decisions. The presence of signaling among sensors implies that, even if all other sensors have fixed their strategies, the problem for i^{th} sensor is no longer a classical sequential detection problem. We show that, for this problem, the classical two-threshold policies no longer guarantee optimality. We obtain an alternative parametric characterization of the optimal policies of the sensors. A related sequential detection problem with one-way communication was presented in [8].

The rest of the paper is organized as follows. In Section II, we formulate our problem with two observers. We present the information states for the observers in Section III. A counterexample that shows that classical two-thresholds are not necessarily optimal is presented in Section IV. We derive a parametric characterization of optimal policies in Section V. We conclude in Section VI.

Notation: Throughout this paper, $X_{1:t}$ refers to the sequence X_1, X_2, \dots, X_t . Subscripts are used as time index and the superscripts are used as the index of the sensor. We use capital letters to denote random variable and the

A. Nayyar and D. Teneketzis are with Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor. anayyar@umich.edu, teneket@eecs.umich.edu

corresponding lower case letters for their realizations.

II. PROBLEM FORMULATION

Consider a binary hypothesis problem where the true hypothesis is modeled as a random variable H taking values 0 or 1 with known prior probabilities.

$$P(H = 0) = p_0; \quad P(H = 1) = 1 - p_0$$

Consider two observers: observer 1 (O1) and observer 2 (O2). We assume that each observer can make noisy observations of the true hypothesis. Conditioned on the hypothesis H , the following statements are assumed to be true:

1. The observation of the i^{th} observer at time t , (Y_t^i) (taking values in the set \mathcal{Y}^i), either has a discrete distribution ($P_t^i(\cdot|H)$) or admits a probability density function ($f_t^i(\cdot|H)$).
2. Observations of the i^{th} observer at different time instants are conditionally independent given H .
3. The observation sequences at the two observers are conditionally independent given H .

Observer i ($i = 1, 2$) observes the measurement process

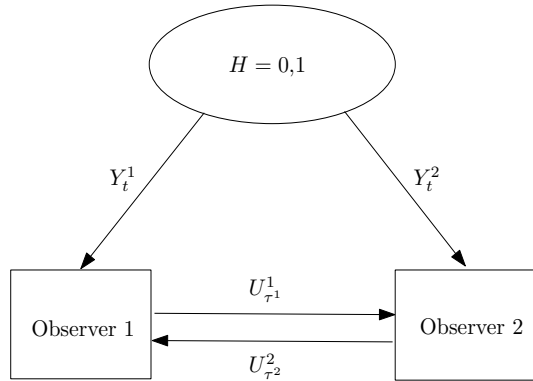


Fig. 1. Decentralized Detection

Y_t^i ($t = 1, 2, \dots$). If no observer has stopped before time t , then at time t , any observer can decide either to stop and send a binary message 0 or 1 to the other observer or to postpone its decision and get another measurement. After an observer has sent its final message, it stops operating. The other observer (the one which has not yet stopped) is then faced with a different stopping problem. At each time instant, it can decide either to stop and declare a final decision on the hypothesis or take another measurement.

We denote by U_t^i the decision of the i^{th} observer at time t . U_t^i belongs to the set $\{0, 1, b\}$, where b denotes a blank, that is, no message or no final decision. At time t , the i^{th} observer makes its decision based on its observations till time t and the messages (blanks or otherwise) exchanged between the two observers till time $t - 1$. We have,

$$U_t^i = \gamma_t^i(Y_{1:t}^i, U_{1:t-1}^1, U_{1:t-1}^2) \quad (1)$$

where the collection of functions $\Gamma^i := (\gamma_t^i, t = 1, 2, \dots)$ constitute the policy of the i^{th} observer. We define the following stopping times:

$$\tau^1 := \min\{t : U_t^1 \neq b\}$$

$$\tau^2 := \min\{t : U_t^2 \neq b\}$$

After the time τ^i , all future observations and decisions of the i^{th} observer are assumed to be in the empty set (\emptyset), that is, the observer stops taking measurements or making decisions. We assume that the final decision on the hypothesis must be made no later than a finite horizon T , hence we have that $\tau^1 \leq T$ and $\tau^2 \leq T$. We define, $\tau^{\min} := \min\{\tau^1, \tau^2\}$, $\tau^{\max} := \max\{\tau^1, \tau^2\}$ and

$$L := \begin{cases} 1 & \text{if } \tau^2 < \tau^1 \\ 2 & \text{if } \tau^2 \geq \tau^1 \end{cases}$$

The system incurs a total cost given by:

$$\begin{aligned} \mathcal{J}(\Gamma^1, \Gamma^2) &:= \mathbb{E}^{\Gamma^1, \Gamma^2} \{K(\tau^{\min} - 1) + k(\tau^{\max} - \tau^{\min}) \\ &\quad + J(U_{\tau^{\max}}^L, H)\} \end{aligned} \quad (2)$$

where $K > k > 0$ are constants and $J(\cdot, \cdot)$ is a non-negative distortion function with $J(0, 0) = J(1, 1) = 0$. (The superscript Γ^1, Γ^2 over the expectation denotes that the expectation is with respect to a measure that depends on the choice of the policies Γ^1, Γ^2 .) The first term in the objective represents the operating costs when both observers are active, the second term represents the operating cost of only one observer. The operating costs incorporate the cost of taking a new measurement, the energy cost of staying on for another time step and/or a penalty for delaying the decision. The last term in the objective represents the accuracy of the final decision. Note that only the decision of the observer that stops later is considered as the final decision. In case of simultaneous decisions, only observer 2's decision is considered as the final decision. We can now formulate the optimization problem as follows:

Problem P: Given the statistics of the binary hypothesis and the observation processes, the cost parameters K, k , the distortion function $J(\cdot, \cdot)$ and a time horizon T , the objective is to select policies Γ^1, Γ^2 that minimize the total expected cost in (2).

III. INFORMATION STATES

In this section, we identify information states for the two observers. We start by fixing the policy of observer 1 to an arbitrary choice and finding a sufficient statistic for observer 2. The nature of this sufficient statistic does not depend on the arbitrary choice of observer 1's policy. This sufficient statistic is the information state for observer 2.

Consider a fixed policy $\Gamma^1 = (\gamma_1^1, \gamma_2^1, \dots, \gamma_T^1)$ for O1. At any time t , we define the following:

Definition 1: Given a fixed policy Γ^1 of observer 1 and functions $\gamma_{1:t-1}^2$, we define observer 2's belief on the hypothesis given all its information at time t .

$$\Pi_t^2 := P(H = 0 | Y_{1:t}^2, U_{1:t-1}^1, U_{1:t-1}^2 = b_{1:t-1}),$$

where $b_{1:t-1}$ denotes a sequence of blank messages from time 1 to $t - 1$. For $t = 0$, we have $\Pi_0^2 = p_0$.

We will show that the pair $(\Pi_t^2, \mathbb{1}_{\{\tau^1 < t\}})$ is the information state for observer 2. We first describe the evolution of $(\Pi_t^2, \mathbb{1}_{\{\tau^1 < t\}})$ in time in the following lemma.

Lemma 1: (i) $\mathbb{1}_{\{\tau^1 < t+1\}} = \mathbb{1}_{\{\tau^1 < t\}} + \mathbb{1}_{\{U_t^1 \neq b\}}$

(ii) With observer 1's policy fixed to Γ^1 , Π_t^2 evolves as follows:

$$\Pi_{t+1}^2 = \begin{cases} f_{t+1}(\Pi_t^2, Y_{t+1}^2) & \text{if } \mathbb{1}_{\{\tau^1 < t\}} = 1 \\ g_{t+1}(\Pi_t^2, Y_{t+1}^2, U_t^1) & \text{if } \mathbb{1}_{\{\tau^1 < t\}} = 0 \end{cases},$$

where f_{t+1} and g_{t+1} are deterministic functions.

Proof: See Appendix I. ■

The optimal policy for observer 2 (for the given choice of Γ^1) can be obtained by means of a dynamic program. We now define the value functions of the dynamic program.

Definition 2: (i) For $\pi \in [0, 1]$ and $a \in \{0, 1\}$, we define

$$V_T(\pi, a) := \min\{\mathbb{E}[J(0, H)|\Pi_T^2 = \pi], \mathbb{E}[J(1, H)|\Pi_T^2 = \pi]\}$$

(ii) For $\pi \in [0, 1]$ and $t = T - 1, T - 2, \dots, 1$, we define

$$V_t(\pi, 1) = \min\{\mathbb{E}[J(0, H)|\Pi_t^2 = \pi], \mathbb{E}[J(1, H)|\Pi_t^2 = \pi], k + \mathbb{E}[V_{t+1}(\Pi_{t+1}^2, 1)|\Pi_t^2 = \pi, \mathbb{1}_{\{\tau^1 < t\}} = 1]\}$$

(iii) For $\pi \in [0, 1]$ and $t = T - 1, T - 2, \dots, 1$, we define $V_t(\pi, 0)$ in equation (3) at the bottom of the page.

Theorem 1: With a fixed policy Γ^1 of observer 1, there is an optimal policy for observer 2 of the form:

$$U_t^2 = \gamma_t^2(\Pi_t^2, \mathbb{1}_{\{\tau^1 < t\}})$$

for $t = 1, 2, \dots, T$. Moreover, this optimal policy can be obtained by the dynamic program described by the value functions in Definition 2. Thus, at time t and for a given π and a , the optimal decision is 0 (or $1/b$) if the first (or second/third) term is the minimum in the definition of $V_t(\pi, a)$.

Proof: See Appendix II. ■

The above arguments can be repeated by interchanging the roles of observer 1 and observer 2 to conclude that for a fixed policy Γ^2 of observer 2, an optimal policy for observer 1 is of the form:

$$U_t^1 = \gamma_t^1(\Pi_t^1, \mathbb{1}_{\{\tau^2 < t\}})$$

where $\Pi_t^1 := P^{1^2}(H = 0|Y_{1:t}^1, U_{1:t-1}^2, U_{1:t-1}^1 = b_{1:t-1})$. Note, however, that the actual dynamic program for observer 1 will differ from that of Theorem 1 because of the

asymmetry in the objective function of equation (2) when both observers make simultaneous decisions to stop. While the value functions V_T and $V_t(\pi, 1)$ in the dynamic program for observer 1 will be the same as in Definition 2, the value function $V_t(\pi, 0)$ is given in equation (4) at the bottom of the next page.

IV. A COUNTEREXAMPLE

In the sequential detection problem with a single observer [5], it is well known that an optimal policy is a function of the observer's posterior belief Π_t and is described by two thresholds at each time t . That is the decision at time t , Z_t , is given as:

$$Z_t = \begin{cases} 1 & \text{if } \Pi_t \leq \alpha_t \\ b & \text{if } \alpha_t < \Pi_t < \beta_t \\ 0 & \text{if } \Pi_t \geq \beta_t \end{cases}$$

where b denotes a decision to continue taking measurement and $\alpha_t \leq \beta_t$ are real numbers in $[0, 1]$. A similar two-threshold structure of optimal policies was also established for the decentralized Wald problem in [6]. We will show by means of a counterexample that such a structure is not necessarily optimal in our problem.

Consider the following instance of Problem P. We have equal prior on H , that is $P(H = 0) = P(H = 1) = 1/2$ and a time horizon of $T = 3$. Assume $k = 1$ and $1 < K < 2$. The observation space of observer 1 is $\mathcal{Y}^1 = \{0, 1\}$ and the observations at time t obey the following conditional probabilities:

<i>Observation</i>	0	1
$P(\cdot H = 0)$	q_t	$(1 - q_t)$
$P(\cdot H = 1)$	$(1 - q_t)$	q_t

where $q_1 = q_2 = 1/2$ and $q_3 = 1$. Thus, the first two observations of observer 1 reveal no information about H while the third observation reveals H noiselessly. The observation space of observer 2 is $\mathcal{Y}^2 = \{0, 1, 2\}$ and the observations at time t obey the following conditional probabilities:

<i>Observation</i>	0	1	2
$P(\cdot H = 0)$	r_t	$(1 - r_t)$	0
$P(\cdot H = 1)$	0	$(1 - r_t)$	r_t

where $r_2 = r_3 = 0$ and $0 < r_1 < 1$. Thus, the second and third observations of observer 2 reveal no information about H . Note that under this statistical model of observations, there exists a choice of policies such that the system makes perfect final decision on the hypothesis and incurs only operational costs (if observer 2 stops at $t = 1$ and observer 1

$$V_t(\pi, 0) = \min\{\mathbb{E}[\mathbb{1}_{\{\tau^1 = t\}}J(0, H) + \mathbb{1}_{\{\tau^1 > t\}}(k(\tau^1 - t) + J(U_{\tau^1}^1, H))|\Pi_t^2 = \pi, \mathbb{1}_{\{\tau^1 < t\}} = 0, U_t^2 = 0], \mathbb{E}[\mathbb{1}_{\{\tau^1 = t\}}J(1, H) + \mathbb{1}_{\{\tau^1 > t\}}(k(\tau^1 - t) + J(U_{\tau^1}^1, H))|\Pi_t^2 = \pi, \mathbb{1}_{\{\tau^1 < t\}} = 0, U_t^2 = 1], \mathbb{E}[\mathbb{1}_{\{\tau^1 = t\}}(k + V_{t+1}(\Pi_{t+1}^2, 1)) + \mathbb{1}_{\{\tau^1 > t\}}(K + V_{t+1}(\Pi_{t+1}^2, 0))|\Pi_t^2 = \pi, \mathbb{1}_{\{\tau^1 < t\}} = 0, U_t^2 = b]\} \quad (3)$$

waits till time $t = 3$, then it can make a perfect decision on H and the system incurs an operational cost of $2k = 2$). We assume that the cost of a mistake in the final decision ($U_{\tau_{max}}^L \neq H$) is sufficiently high so that any choice of policies that makes a mistake in the final decision with non-zero probability will have a performance worse than 2. Thus, any choice of policies that makes a mistake in the final decision with non-zero probability cannot be optimal.

Under the above instance of our problem, if observer 2 is restricted to use a two-threshold rule at time $t = 1$, then the lowest achievable value of the objective is given as:

$$\min\{r_1 + (1 - r_1)(K + 1)\}, \{2 - r_1/2\} \quad (5)$$

where the first term corresponds to the case when γ_1^2 is given as:

$$U_1^2 = \begin{cases} 1 & \text{if } \Pi_1^2 = 0 \\ b & \text{if } 0 < \Pi_1^2 < 1 \\ 0 & \text{if } \Pi_1^2 = 1 \end{cases} \quad (6)$$

and the second term corresponds to γ_1^2 being

$$U_1^2 = \begin{cases} 1 & \text{if } \Pi_1^2 < 1 \\ 0 & \text{if } \Pi_1^2 = 1 \end{cases} \quad (7)$$

Other choices of thresholds for observer 2 at time $t = 1$ do not give a lower value than the expression in (5).

Consider now the following choice of $\gamma_1^{2,*}$:

$$U_1^2 = \begin{cases} 1 & \text{if } \Pi_1^2 = 0 \\ 0 & \text{if } 0 < \Pi_1^2 < 1 \\ b & \text{if } \Pi_1^2 = 1 \end{cases} \quad (8)$$

The lowest achievable expected cost under the above choice of γ_1^2 is $\mathcal{J}^* = 2(1 - r_1) + r_1(K + 1)/2$. It is easy to check that for $1 < K < 2$ and $r_1 < 2/3$,

$$\mathcal{J}^* < \min\{r_1 + (1 - r_1)(K + 1)\}, \{2 - r_1/2\}$$

Thus, $\gamma_1^{2,*}$ outperforms the two-threshold rules.

Discussion: In the above example, observer 1 can always make the correct decision at time $t = 3$. However, this incurs additional operational costs. A good policy should try to enable the observers to make the correct decision before time $t = 3$, whenever possible. If observer 2 gets the observations 0 or 2 at $t = 1$, then it is certain about the true hypothesis. Using the first threshold rule given in (6), observer 2 is able to convey to observer 1 that it is certain about the true hypothesis and what this hypothesis is, thus preventing observer 1 from waiting till time $t = 3$ to make the final decision. However, in the case when observer 2 gets measurement 1, it sends a blank and postpones its decision to stop. This incurs additional operating costs for observer 2

without providing any new information that may prevent observer 1 from delaying its decision to time $t = 3$. By making r_1 small, the contribution of this term in the overall cost is increased. The second threshold (equation (7)) rule attempts to keep the operational costs of observer 2 small but does not always send enough information to observer 1 to enable it to make a decision before $t = 3$ even when observer 1 knows the true value of H . The non-threshold rule (given in (8)), however, keeps the operational cost of observer 2 small (when r_1 is small) while at the same time ensuring that whenever observer 2 is certain about true value of H , the final decision is not postponed to time $t = 3$.

V. PARAMETRIC CHARACTERIZATION OF OPTIMAL POLICIES

An important advantage of the threshold rules in the case of the centralized or the decentralized Wald problem is that it modifies the problem of finding the globally optimal policies from a sequential functional optimization problem to a sequential parametric optimization problem. Even though we have established that a classical threshold rule does not hold for our problem, it is still possible to get a finite parametric characterization of optimal policies. Such a parametric characterization provides significant computational advantage in finding optimal policies, for example by reducing the search space for an optimal policy.

In Theorem 1, we have established that for an arbitrarily fixed choice of observer 1's policy, the optimal policy for observer 2 can be determined by a dynamic program using the value functions $V_t(\pi, a)$, $t = T, \dots, 2, 1$. We have the following lemma.

Lemma 2: With a fixed (but arbitrary) choice of Γ^1 , the value function at T can be expressed as:

$$V_T(\pi, a) := \min\{l^0(\pi), l^1(\pi)\} \quad (9)$$

where l^0 and l^1 are affine functions of π . Also, the value functions at time t can be expressed as:

$$V_t(\pi, 1) := \min\{l^0(\pi), l^1(\pi), G_t(\pi)\} \quad (10)$$

where G_t is a concave function of π , and

$$V_t(\pi, 0) := \min\{L_t^0(\pi), L_t^1(\pi), H_t(\pi)\} \quad (11)$$

where L_t^0 and L_t^1 are affine functions of π and H_t is a concave function of π (the actual form of these functions depends on the choice of Γ^1).

Proof: See Appendix III. ■

$$\begin{aligned} V_t(\pi, 0) = & \min\{\mathbb{E}[k(\tau^2 - t) + J(U_{\tau^2}^2, H)] | \Pi_t^1 = \pi, \mathbb{1}_{\{\tau^2 < t\}} = 0, U_t^1 = 0\}, \\ & \mathbb{E}[k(\tau^2 - t) + J(U_{\tau^2}^2, H)] | \Pi_t^1 = \pi, \mathbb{1}_{\{\tau^2 < t\}} = 0, U_t^1 = 1\}, \\ & \mathbb{E}[\mathbb{1}_{\{\tau^2 = t\}}(k + V_{t+1}(\Pi_{t+1}^1, 1)) \\ & + \mathbb{1}_{\{\tau^2 > t\}}(K + V_{t+1}(\Pi_{t+1}^1, 0)) | \Pi_t^1 = \pi, \mathbb{1}_{\{\tau^2 < t\}} = 0, U_t^1 = b\}] \end{aligned} \quad (4)$$

Theorem 2: For any fixed policy Γ^1 of observer 1, an optimal policy for observer 2 can be characterized as follows:

$$U_T^2 = \begin{cases} 1 & \text{if } \Pi_T^2 \leq \alpha_T \\ 0 & \text{if } \Pi_T^2 > \alpha_T \end{cases}$$

where $0 \leq \alpha_T \leq 1$. For $t = 1, 2, \dots, T-1$, if $\mathbb{1}_{\{\tau^1 < t\}} = 1$,

$$U_t^2 = \begin{cases} 1 & \text{if } \Pi_t^2 \leq \alpha_t(1) \\ b & \text{if } \alpha_t(1) < \Pi_t^2 \leq \beta_t(1) \\ 0 & \text{if } \Pi_t^2 > \beta_t(1) \end{cases}$$

where $0 \leq \alpha_t(1) \leq \beta_t(1) \leq 1$, and if $\mathbb{1}_{\{\tau^1 < t\}} = 0$,

$$U_t^2 = \begin{cases} b & \text{if } \Pi_t^2 < \alpha_t(0) \\ 1 & \text{if } \alpha_t(0) \leq \Pi_t^2 \leq \beta_t(0) \\ b & \text{if } \beta_t(0) < \Pi_t^2 < \delta_t(0) \\ 0 & \text{if } \delta_t(0) \leq \Pi_t^2 \leq \theta_t(0) \\ b & \text{if } \Pi_t^2 > \theta_t(0) \end{cases}$$

where $0 \leq \alpha_t(0) \leq \beta_t(0) \leq \delta_t(0) \leq \theta_t(0) \leq 1$.

Proof: From lemma 2, we know that the value functions can be written as minimum of affine and concave functions. Since taking minimum of two straight lines and a concave function can partition the interval $[0, 1]$ into at most five intervals, this gives a four threshold characterization of optimal policy where the thresholds signify the boundaries of these intervals. At time T or when $\mathbb{1}_{\{\tau^1 < t\}} = 1$, observer 2's decision of 0 or 1 is the final decision ($U_{\tau_{max}}^L$) on H . In these cases, if observer 2 is certain about H (that is its belief is 0 or 1), then it should clearly choose the correct value of H . This fact reduces the number of thresholds for time T and when $\mathbb{1}_{\{\tau^1 < t\}} = 1$. ■

Discussion: It is instructive to compare our problem with the decentralized Wald problem studied in [6]. Both problems involve two observers that make repeated observations of H and decide when to stop. Unlike the problem in this paper, in the decentralized Wald problem the sensors do not have access to each other's decisions, that is, an observer's policy is restricted to be of the form:

$$U_t^i = \gamma_t^i(Y_{1:t}^i, U_{1:t-1}^i)$$

The optimality of classical two-threshold rules for the decentralized Wald problem was established in [6]. In this paper, we allowed each observer to observe other's decisions and showed that the two threshold rules are no longer optimal.

Both the decentralized Wald problem and the problem formulated in this paper are *team problems*. That is, they involve more than one decision maker with a common objective. However, in the decentralized Wald problem, a decision-maker's decisions do not influence the information available to other decision-makers. This is the essential criterion for *static* team problems [9]. The problem formulated in this paper is a *dynamic* team problem since a decision-maker's past decisions are a part of the information available to other decision-makers. It is the dynamic aspect of this team problem that allows for signaling between decision-makers and changes the nature of optimal policies.

VI. CONCLUSIONS

We considered a sequential problem in decentralized detection problem with signaling. Two observers make separate costly measurements of a binary hypothesis and decide when to stop. The observers can observe each other's decisions (whether the other observe has stopped or not and whether the final decision was 0 or 1). The final decision of the observer that stops in the end is taken as the final decision made by the group. Thus, observers can convey information to each other through their decisions. We identified information states for the two observers and showed that classical two threshold rules no longer guarantee optimality. However, a finite parametric characterization of optimal policies is still possible.

APPENDIX I PROOF OF LEMMA 1

Proof: Part (i) follows from definition of τ^1 .

In Part (ii), if $\tau^1 < t$, then by definition, we have

$$\begin{aligned} \Pi_t^2 &:= P(H = 0 | Y_{1:t}^2, U_{1:t-1}^1, U_{1:t-1}^2 = b_{1:t-1}) \\ &= P(H = 0 | Y_{1:t}^2, U_{1:\tau^1}^1) \end{aligned} \quad (12)$$

where we removed redundant terms from the conditioning (terms which are constants or functions of other terms). Similarly,

$$\Pi_{t+1}^2 = P(H = 0 | Y_{1:t+1}^2, U_{1:\tau^1}^1),$$

which, on using Bayes' rule gives,

$$\begin{aligned} \Pi_{t+1}^2 &= \frac{P(Y_{t+1}^2 | H = 0) \Pi_t^2}{P(Y_{t+1}^2 | H = 0) \Pi_t^2 + P(Y_{t+1}^2 | H = 1)(1 - \Pi_t^2)} \\ &=: f_{t+1}(\Pi_{t+1}^2, Y_{t+1}^2) \end{aligned} \quad (13)$$

If $\tau^1 \geq t$, then $U_{1:t-1}^1 = b_{1:t-1}$ (that is all decisions of observer 1 are blanks till time $t-1$) and

$$\Pi_t^2 := P(H = 0 | Y_{1:t}^2, U_{1:t-1}^1 = b_{1:t-1}, U_{1:t-1}^2 = b_{1:t-1})$$

Also,

$$\begin{aligned} \Pi_{t+1}^2 &:= P(H = 0 | Y_{1:t+1}^2, U_{1:t-1}^1 = b_{1:t-1}, U_t^1, U_{1:t}^2 = b_{1:t}) \\ &= \frac{P(Y_{t+1}^2, U_t^1, H = 0 | Y_{1:t}^2, U_{1:t-1}^1 = b_{1:t-1}, U_{1:t}^2 = b_{1:t})}{\sum_{h \in \{0,1\}} P(Y_{t+1}^2, U_t^1, h | Y_{1:t}^2, U_{1:t-1}^1 = b_{1:t-1}, U_{1:t}^2 = b_{1:t})} \end{aligned} \quad (14)$$

The numerator in (14) can be written as:

$$\begin{aligned} &P(Y_{t+1}^2 | H = 0) \cdot \\ &\{P(U_t^1 | H = 0, Y_{1:t}^2, U_{1:t-1}^1 = b_{1:t-1}, U_{1:t}^2 = b_{1:t})\} \Pi_t^2 \end{aligned} \quad (15)$$

We now focus on the second term in (15).

Claim: Consider a realization $u_t^1, y_{1:t}^2$. Then,

$$\begin{aligned} &P(u_t^1 | H = 0, y_{1:t}^2, U_{1:t-1}^1 = b_{1:t-1}, U_{1:t}^2 = b_{1:t}) \\ &= P(u_t^1 | H = 0, U_{1:t-1}^1 = b_{1:t-1}, U_{1:t}^2 = b_{1:t}) \end{aligned} \quad (16)$$

Moreover, under the given choice of Γ^1 , the probability on the right hand side of (16) is a function only of u_t^1 .

Proof of claim: Using Bayes' rule,

$$\begin{aligned} & P(u_t^1 | H = 0, y_{1:t}^2, U_{1:t-1}^1 = b_{1:t-1}, U_{1:t}^2 = b_{1:t}) \\ &= \frac{P(u_t^1, H = 0, y_{1:t}^2, U_{1:t-1}^1 = b_{1:t-1}, U_{1:t}^2 = b_{1:t})}{\sum_{u'} P(U_t^1 = u', H = 0, y_{1:t}^2, U_{1:t-1}^1 = b_{1:t-1}, U_{1:t}^2 = b_{1:t})} \end{aligned} \quad (17)$$

Consider the joint probability in the numerator in (17)

$$P(u_t^1, H = 0, y_{1:t}^2, b_{1:t-1}^1, b_{1:t}^2)$$

where we use $b_{1:t-1}^1, b_{1:t}^2$ as shorthand notations for $U_{1:t-1}^1 = b_{1:t-1}$ and $U_{1:t}^2 = b_{1:t}$ respectively. This probability can be further written as:

$$\begin{aligned} &= \sum_{y_{1:t}^1} P(u_t^1, H = 0, y_{1:t}^2, b_{1:t-1}^1, b_{1:t}^2, y_{1:t}^1) \\ &= \sum_{y_{1:t}^1} [P(u_t^1 | y_{1:t}^1, b_{1:t-1}^1, b_{1:t}^2) \\ &\quad \cdot P(U_t^2 = b | y_{1:t}^2, b_{1:t-1}^1, b_{1:t}^2) \\ &\quad \cdot P(y_t^1 | H = 0) P(y_t^2 | H = 0) \\ &\quad \cdot \prod_{k=1}^{t-1} \{P(U_k^1 = b | y_{1:k}^1, b_{1:k-1}^1, b_{1:k-1}^2) \\ &\quad \cdot P(U_k^2 = b | y_{1:k}^2, b_{1:k-1}^1, b_{1:k-1}^2) \\ &\quad \cdot P(y_k^1 | H = 0) P(y_k^2 | H = 0)\} \cdot p_0 \end{aligned} \quad (18)$$

Rearranging the summation in (18), we get

$$\begin{aligned} & P(U_t^2 = b | y_{1:t}^2, b_{1:t-1}^1, b_{1:t}^2) P(y_t^2 | H = 0) \cdot p_0 \\ &\quad \cdot \prod_{k=1}^{t-1} \{P(U_k^2 = b | y_{1:k}^2, b_{1:k-1}^1, b_{1:k-1}^2) P(y_k^2 | H = 0)\} \\ &\quad \cdot \sum_{y_{1:t}^1} [P(u_t^1 | y_{1:t}^1, b_{1:t-1}^1, b_{1:t}^2) P(y_t^1 | H = 0) \\ &\quad \cdot \prod_{k=1}^{t-1} \{P(U_k^1 = b | y_{1:k}^1, b_{1:k-1}^1, b_{1:k-1}^2) P(y_k^1 | H = 0)\}] \end{aligned} \quad (19)$$

Expressions similar to (19) hold for each term in the denominator of (17) and the terms outside the summation over $y_{1:t}^1$ cancel out in the numerator and the denominator. We note that the summation over $y_{1:t}^1$ in (19) does not depend on $y_{1:t}^2$. Hence, the conditional probability in the left hand side of (17) does not depend on $y_{1:t}^2$. This establishes equation (16). We also note that under the fixed policy Γ^1 of observer 1, the summation over $y_{1:t}^1$ in (19) is a function only of u_t^1 . Thus, the probability on the right hand side of (16) is a function only of u_t^1 . This concludes the proof of the claim.

Using the result of the claim in (15) and using similar arguments for the denominator in (14), we get

$$\begin{aligned} & \Pi_{t+1}^2 \\ &= \frac{P(Y_{t+1}^2 | H = 0) P(U_t^1 | H = 0, b_{1:t-1}^1, b_{1:t}^2) \Pi_t^2}{P(Y_{t+1}^2 | H = 0) P(U_t^1 | H = 0, b_{1:t-1}^1, b_{1:t}^2) \Pi_t^2 + P(Y_{t+1}^2 | H = 1) P(U_t^1 | H = 1, b_{1:t-1}^1, b_{1:t}^2) (1 - \Pi_t^2)} \\ &=: g_{t+1}(\Pi_{t+1}^2, Y_{t+1}^2, U_t^1) \end{aligned} \quad (20)$$

■

APPENDIX II

PROOF OUTLINE OF THEOREM 1

We provide an outline of the proof of Theorem 1. The general idea is to show that at each time t , the value functions of Definition 2 represent the optimal future costs. Therefore, a policy that for each realization of $\Pi_t^2, \mathbb{1}_{\{\tau^1 < t\}}$ selects the minimizing term in the corresponding value function achieves the optimal cost. Thus, an optimal policy can be found that depends only on $\Pi_t^2, \mathbb{1}_{\{\tau^1 < t\}}$. We start from time T .

If the observer 2 is active at the terminal time T , it can only make one of two decisions: 0 or 1. The expected future cost of choosing $u \in \{0, 1\}$ for observer 2 is

$$\begin{aligned} & E[J(u, H) | Y_{1:T}^2, U_{1:T-1}^1, U_{1:T-1}^2 = b_{1:T-1}] \\ &= J(u, 0) \Pi_T^2 + J(u, 1) (1 - \Pi_T^2) \\ &= E[J(u, H) | \Pi_T^2] \end{aligned} \quad (21)$$

Thus, the value function at time T is the minimum of the expected future costs incurred by choosing 0 or 1. Hence, it represents the optimal expected future cost for observer 2 at time T . Proceeding backwards, we assume that the value functions at time $t+1, t+2, \dots, T$ represent optimal future costs at the respective times and consider two cases at each time $t < T$.

Case A: $\tau^1 < t$ If observer 1 has already stopped before t , then observer 2's stopping problem is the same as the centralized Wald problem and the value function $V_t(\pi, 1)$ is same as the value function in the dynamic program for the Wald problem.

Case B: $\tau^1 \geq t$ We now consider the case when observer 1 has not stopped before time t . If observer 2 decides to stop and chooses $u \in \{0, 1\}$ at time t , then the expected future cost will be

$$\begin{aligned} & E[\mathbb{1}_{\{\tau^1 = t\}} J(u, H) \\ &\quad + \mathbb{1}_{\{\tau^1 > t\}} (k(\tau^1 - t) + J(U_{\tau^1}^1, H)) \Big| \left. \begin{array}{l} Y_{1:t}^2, b_{1:t-1}^1, \\ b_{1:t-1}^2, U_t^2 = u \end{array} \right] \end{aligned} \quad (22)$$

Claim: The expectation in (22) is same as:

$$\begin{aligned} & E[\mathbb{1}_{\{\tau^1 = t\}} J(u, H) + \\ &\quad \mathbb{1}_{\{\tau^1 > t\}} (k(\tau^1 - t) + J(U_{\tau^1}^1, H)) | \Pi_t^2, \mathbb{1}_{\{\tau^1 < t\}} = 0, U_t^2 = u] \end{aligned} \quad (23)$$

Proof of claim: For each realization $y_{1:t}^2$ of observer 2's observations, the expectation in (22) depends on the conditional distribution of the following random variables: $H, \tau^1, U_{\tau^1}^1$ given the realization of the random variables $y_{1:t}^2, b_{1:t-1}^1, b_{1:t-1}^2, U_t^2 = u$. Note that under the fixed policy Γ^1 of observer 1, $\tau^1, U_{\tau^1}^1$ are functions of observer 1's observation sequence $Y_{1:T}^1$ and the terms $b_{1:t-1}^1, b_{1:t-1}^2, U_t^2 = u$ fixed in the conditioning. Hence the conditional belief

$$P(H, \tau^1, U_{\tau^1}^1 | y_{1:t}^2, b_{1:t-1}^1, b_{1:t-1}^2, U_t^2 = u)$$

is a deterministic transformation of the belief

$$P(H, Y_{1:T}^1 | y_{1:t}^2, b_{1:t-1}^1, b_{1:t-1}^2, U_t^2 = u).$$

We will show that the above probability is same as

$$P(H, Y_{1:T}^1 | \pi_t^2, \mathbb{1}_{\{\tau^1 < t\}} = 0, U_t^2 = u)$$

and hence the conditional expectation in (22) is same as

$$\mathbb{E}[\mathbb{1}_{\{\tau^1 = t\}} J(u, H) + \mathbb{1}_{\{\tau^1 > t\}} (k(\tau^1 - t) + J(U_{\tau^1}^1, H)) | \pi_t^2, \mathbb{1}_{\{\tau^1 < t\}} = 0, U_t^2 = u]$$

which corresponds to the first two terms in the minimization in $V_t(\pi_t^2, 0)$ in equation (3).

$$\begin{aligned} & \text{Consider } P(H = 0, y_{1:T}^1 | y_{1:t}^2, b_{1:t-1}^1, b_{1:t-1}^2, U_t^2 = u) \\ &= P(y_{1:T}^1 | H = 0, y_{1:t}^2, b_{1:t-1}^1, b_{1:t-1}^2, U_t^2 = u) \pi_t^2 \end{aligned} \quad (24)$$

Similarly,

$$\begin{aligned} & P(H = 0, y_{1:T}^1 | \pi_t^2, \mathbb{1}_{\{\tau^1 < t\}} = 0, U_t^2 = u) \\ &= P(y_{1:T}^1 | H = 0, \pi_t^2, \mathbb{1}_{\{\tau^1 < t\}} = 0, U_t^2 = u) \pi_t^2 \end{aligned} \quad (25)$$

We now compare the first terms in (24) and (25). Consider the first term in (24), which can be written as

$$= \frac{P(H = 0, y_{1:T}^1, y_{1:t}^2, b_{1:t-1}^1, b_{1:t-1}^2, U_t^2 = u)}{\sum_{\tilde{y}_{1:T}^1} P(H = 0, \tilde{y}_{1:T}^1, y_{1:t}^2, b_{1:t-1}^1, b_{1:t-1}^2, U_t^2 = u)} \quad (26)$$

The numerator can be written as:

$$\begin{aligned} & P(y_{t+1:T}^1 | H = 0) P(U_t^2 = u | y_{1:t}^2, b_{1:t-1}^1, b_{1:t-1}^2) \\ & \cdot P(y_t^1 | H = 0) P(y_t^2 | H = 0) \\ & \cdot \prod_{k=1}^{t-1} \{P(U_k^1 = b | y_{1:k}^1, b_{1:k-1}^1, b_{1:k-1}^2) \\ & \cdot P(U_k^2 = b | y_{1:k}^2, b_{1:k-1}^1, b_{1:k-1}^2) \\ & \cdot P(y_k^1 | H = h) P(y_k^2 | H = h)\} \cdot p_0 \end{aligned} \quad (27)$$

Similar expressions hold for the denominator in (26) and the terms that depend on $y_{1:t}^2$ will cancel in the numerator and the denominator. Therefore,

$$\begin{aligned} & P(y_{1:T}^1 | H = 0, y_{1:t}^2, b_{1:t-1}^1, b_{1:t-1}^2, U_t^2 = u) \\ &= P(y_{1:T}^1 | H = 0, b_{1:t-1}^1, b_{1:t-1}^2, U_t^2 = u) \end{aligned} \quad (28)$$

Now consider the first term in (25) which can be written as:

$$\begin{aligned} & P(y_{1:T}^1 | H = 0, \pi_t^2, \mathbb{1}_{\{\tau^1 < t\}} = 0, U_t^2 = u) \\ &= \sum_{y_{1:t}^2} [P(y_{1:T}^1 | y_{1:t}^2, H = 0, \pi_t^2, \mathbb{1}_{\{\tau^1 < t\}} = 0, U_t^2 = u) \\ & \cdot P(y_{1:t}^2 | H = 0, \pi_t^2, \mathbb{1}_{\{\tau^1 < t\}} = 0, U_t^2 = u)] \\ &= \sum_{y_{1:t}^2} [P(y_{1:T}^1 | y_{1:t}^2, H = 0, b_{1:t-1}^1, b_{1:t-1}^2, U_t^2 = u) \\ & \cdot P(y_{1:t}^2 | H = 0, \pi_t^2, \mathbb{1}_{\{\tau^1 < t\}} = 0, U_t^2 = u)] \end{aligned} \quad (29)$$

The first term inside the summation in (29) is same as LHS of (28). Using (28) in (29) gives

$$\begin{aligned} & \sum_{y_{1:t}^2} [P(y_{1:T}^1 | H = 0, b_{1:t-1}^1, b_{1:t-1}^2, U_t^2 = u) \\ & \cdot P(y_{1:t}^2 | H = 0, \pi_t^2, \mathbb{1}_{\{\tau^1 < t\}} = 0, U_t^2 = u)] \\ &= P(y_{1:T}^1 | H = 0, b_{1:t-1}^1, b_{1:t-1}^2, U_t^2 = u) \end{aligned} \quad (30)$$

which is same as RHS of (28). Thus the probabilities in RHS of (24) and (25) are equal. Similar conclusions hold for $H = 1$ in (24) and (25). This implies the equality of expectations and completes the proof of the claim.

As a consequence of the claim, the first two terms in the minimization in the definition of $V_t(\pi, 0)$ (equation 3) correspond to the expected future cost of choosing 0 or 1 at time t . On the other hand, if observer 2 decides to continue at time t , then by the fact that value functions at $t+1$ represent the expected future costs at $t+1$, we can write the expected future cost as:

$$\begin{aligned} & \mathbb{E} \left[\mathbb{1}_{\{\tau^1 = t\}} (k + V_{t+1}(\Pi_{t+1}, 1)) \right. \\ & \left. + \mathbb{1}_{\{\tau^1 > t\}} (K + V_{t+1}(\Pi_{t+1}, 0)) \right] \Bigg|_{\substack{Y_{1:t}^2, b_{1:t-1}^1 \\ b_{1:t-1}^2, U_t^2 = b}} \end{aligned} \quad (31)$$

Using lemma 1 and the fact that under fixed policy Γ^1 of observer 1, U_t^1 is a function of observer 1's observation sequence $Y_{1:t}^1$ and the terms $b_{1:t-1}^1, b_{1:t-1}^2$ fixed in the conditioning, one can conclude that for each realization of $y_{1:t}^2$ this expectation in (31) is a function of the following conditional probability:

$$P(Y_{1:t}^1, Y_{t+1}^2 | y_{1:t}^2, b_{1:t-1}^1, b_{1:t-1}^2, U_t^2 = b)$$

Using arguments similar to those in the claim above, it can be shown that the above conditional probability is same as:

$$P(Y_{1:t}^1, Y_{t+1}^2 | \pi_t^2, \mathbb{1}_{\{\tau^1 < t\}} = 0, U_t^2 = b)$$

This shows that the third term in the minimization in the definition of $V_t(\pi, 0)$ (equation 3) is the expected future cost of making a decision to continue at time t . Thus, $V_t(\Pi_t^2, 0)$ is the minimum of the future costs incurred by choosing 0, 1 or b . Hence, it represents the optimal future cost at time t , if observer 1 has not already stopped before time t .

APPENDIX III

PROOF OUTLINE OF LEMMA 2

We define the following functions

$$\begin{aligned} l^0(\pi) &:= J(0, 0)\pi + J(0, 1)(1 - \pi) \\ &= \mathbb{E}[J(0, H) | \Pi_t^2 = \pi] \\ l^1(\pi) &:= J(1, 0)\pi + J(1, 1)(1 - \pi) \\ &= \mathbb{E}[J(1, H) | \Pi_t^2 = \pi] \end{aligned}$$

For the value function at time T , the result of the lemma follows from the definitions of $l^0(\pi), l^1(\pi)$ and $V_T(\pi, a)$. Since, for each $a \in \{0, 1\}$, $V_T(\pi, a)$ is the minimum of two affine functions of π , it implies that, for each $a \in \{0, 1\}$, $V_T(\pi, a)$ is a concave function of π . Now, assume that $V_{t+1}(\pi, a)$ is concave in π for each $a \in \{0, 1\}$. The concavity of the value functions at time $t+1$ implies that they can be written as infimum of affine functions of π . In particular, we have

$$V_{t+1}(\pi, 1) = \inf_i \{a_i \pi + b_i\} \quad (32)$$

and

$$V_{t+1}(\pi, 0) = \inf_i \{c_i \pi + d_i\} \quad (33)$$

Now consider $V_t(\pi, 1)$. The first two terms in the definition of $V_t(\pi, 1)$ are affine in π (see Definition 2). We need to show that the third term -

$$k + \mathbb{E}[V_{t+1}(\Pi_{t+1}^2, 1) | \Pi_t^2 = \pi, \mathbb{1}_{\{\tau^1 < t\}} = 1] \quad (34)$$

is a concave function of π . From equation (13) in the proof of lemma 1, we know that Π_{t+1} can be written as:

$$\begin{aligned} \Pi_{t+1}^2 &= \frac{P(Y_{t+1}^2 | H = 0) \Pi_t^2}{P(Y_{t+1}^2 | H = 0) \Pi_t^2 + P(Y_{t+1}^2 | H = 1)(1 - \Pi_t^2)} \\ &= \frac{P(Y_{t+1}^2 | H = 0) \Pi_t^2}{P(Y_{t+1}^2 | \Pi_t^2)} \end{aligned} \quad (35)$$

Substituting (35) in (34) and evaluating the expectation gives:

$$\begin{aligned} k + \sum_{y \in \mathcal{Y}} P(Y_{t+1}^2 = y | \Pi_t^2 = \pi, \mathbb{1}_{\{\tau^1 < t\}} = 1) \\ V_{t+1} \left(\frac{P(Y_{t+1}^2 = y | H = 0) \pi}{P(Y_{t+1}^2 = y | \pi)}, 1 \right) \end{aligned} \quad (36)$$

Now using the characterization of $V_{t+1}(\pi, 1)$ from (32), we get

$$\begin{aligned} k + \sum_{y \in \mathcal{Y}} P(Y_{t+1}^2 = y | \Pi_t^2 = \pi) \\ \left[\inf_i \{a_i \left(\frac{P(Y_{t+1}^2 = y | H = 0) \pi}{P(Y_{t+1}^2 = y | \pi)} \right) + b_i\} \right] \\ = k + \sum_{y \in \mathcal{Y}} \inf_i \{a_i (P(Y_{t+1}^2 = y | H = 0) \pi) \\ + b_i P(Y_{t+1}^2 = y | H = 0) \pi \\ + b_i P(Y_{t+1}^2 = y | H = 1)(1 - \pi)\} \end{aligned} \quad (37)$$

Each term in the summation over $y \in \mathcal{Y}$ is infimum of affine functions of π , hence each term in the summation is a concave function of π . Thus, the third term of $V_t(\pi, 1)$ is a concave function of π .

Next consider $V_t(\pi, 0)$ defined in (3). The conditional expectation for the first (or second) term in the minimization in RHS of (3) is an affine function of the conditional probability of the random variables $H, Y_{1:T}^1$. Using arguments from Appendix II, this conditional probability can be written as:

$$\begin{aligned} P(y_{1:T}^1, H = 0 | \pi_t^2, b_{1:t-1}^1, b_{1:t-1}^2, U_t^2 = 0) \\ = P(y_{1:T}^1 | H = 0, b_{1:t-1}^1, b_{1:t-1}^2, U_t^2 = 0) \pi_t^2 \end{aligned} \quad (38)$$

and

$$\begin{aligned} P(y_{1:T}^1, H = 1 | \pi_t^2, b_{1:t-1}^1, b_{1:t-1}^2, U_t^2 = 0) \\ = P(y_{1:T}^1 | H = 1, b_{1:t-1}^1, b_{1:t-1}^2, U_t^2 = 0)(1 - \pi_t^2) \end{aligned} \quad (39)$$

Thus, the conditional probability $P(Y_{1:T}^1, H | \pi_t^2, b_{1:t-1}^1, b_{1:t-1}^2, U_t^2 = 0)$ is an affine function of π_t^2 . This establishes the affine nature of the first two

terms of RHS of (3). The third term in $V_t(\pi, 0)$ can be written as:

$$\begin{aligned} \mathbb{E} \left[\mathbb{1}_{\{U_t^1 \neq b\}} (k + V_{t+1}(g_{t+1}(\pi, Y_{t+1}^2, U_t^1), 1)) \right. \\ \left. + \mathbb{1}_{\{U_t^1 = b\}} (K + V_{t+1}(g_{t+1}(\pi, Y_{t+1}^2, b), 0)) \right] \begin{matrix} \Pi_t^2 = \pi, \\ \mathbb{1}_{\{\tau^1 < t\}} = 0, \\ U_t^2 = b \end{matrix} \quad (40) \end{aligned}$$

Consider the first term in the summation in (40). Evaluating the expectation, we get

$$\begin{aligned} \sum_{u_t^1, y_{t+1}^2} [\mathbb{1}_{\{u_t^1 \neq b\}} (k + V_{t+1}(g_{t+1}(\pi, y_{t+1}^2, u_t^1), 1)) \\ \cdot P(u_t^1, y_{t+1}^2 | \Pi_t^2 = \pi, \mathbb{1}_{\{\tau^1 < t\}} = 0, U_t^2 = b)] \end{aligned} \quad (41)$$

Using the characterization of g_{t+1} from (14) and (20), the characterization of $V_{t+1}(\pi, 1)$ from (33) and arguments from Appendix II, (41) can be shown to be equal to

$$\begin{aligned} \sum_{u_t^1 = 0, 1} \sum_{y_{t+1}^2} [k \\ + \inf_i \{a_i P(y_{t+1}^2 | H = 0) P(u_t^1 | H = 0, b_{1:t-1}^1, b_{1:t}^2) \pi \\ + b_i P(y_{t+1}^2 | H = 0) P(u_t^1 | H = 0, b_{1:t-1}^1, b_{1:t}^2) \pi \\ + b_i P(y_{t+1}^2 | H = 1) P(u_t^1 | H = 1, b_{1:t-1}^1, b_{1:t}^2) (1 - \pi)\}] \end{aligned} \quad (42)$$

which is concave in π (since it is infimum of affine functions in π). Similar arguments can be made for the second term in (40) to conclude the concavity of third term in $V_t(\pi, 0)$.

ACKNOWLEDGMENTS

This research was supported in part by NSF Grant CCR-0325571 and NASA Grant NNX09AE91G.

REFERENCES

- [1] R. R. Tenney and N. R. Sandell Jr., "Detection with distributed sensors," *IEEE Trans. Aerospace Electron. Systems*, vol. AES-17, no. 4, pp. 501–510, July 1981.
- [2] J. N. Tsitsiklis, "Decentralized detection," in *Advances in Statistical Signal Processing*. JAI Press, 1993, pp. 297–344.
- [3] P. K. Varshney, *Distributed Detection and Data Fusion*. Springer, 1997.
- [4] R. Radner, "Team decision problems," *The Annals of Math. Statistics*, vol. 33, no. 3, pp. 857–881, Sept. 1962.
- [5] A. Wald, *Sequential Analysis*. Wiley, New York, 1947.
- [6] D. Teneketzis and Y. C. Ho, "The decentralized wald problem," *Information and Computation*, 73, pp. 23–44, 1987.
- [7] A. LaVigna, A.M. Makowski, and J.S. Baras, "A continuous-time distributed version of the wald's sequential hypothesis testing problem," *Lecture Notes in Control and Information Sciences*, vol. 83, pp. 533–543, 1986.
- [8] A. Nayyar and D. Teneketzis, "A sequential problem in decentralized detection with communication," *IEEE Trans. Info. Theory*, submitted.
- [9] Y. C. Ho, "Team decision theory and information structures," in *Proceedings of the IEEE*, vol. 68, no. 6, 1980, pp. 644–654.