On Computing Optimal Thresholds in Decentralized Sequential Hypothesis Testing

Can Cui and Aditya Mahajan

***Abstract*—Decentralized sequential hypothesis testing refers to a generalization of Wald's sequential hypothesis testing setup in which multiple decision makers make separate stopping decisions that are coupled through a common loss function. In the simplest such generalization, the stopping decisions are not seen by other decision makers. For this model, it is known that threshold-based stopping strategies are optimal. Two methods are presented for approximately computing the optimal thresholds. The first method, which is called orthogonal search, is an iterative method that approximately solves the coupled dynamic programs proposed in Teneketzis and Ho, Information and Computation, 1987. The second method, which is called direct search, approximates the performance of a threshold-based strategy and then searches over the thresholds using a derivative-free non-convex optimization algorithm. The approximations for both methods are based on discretizing the continuous-state information state process to a finite-state Markov chain and calculating the absorption probabilities and absorption stopping times for appropriately defined absorption sets. The performance of both the methods is compared numerically.**

***Index Terms*—Hypothesis testing, sequential analysis, decentralized systems, dynamic programming, derivative-free optimization.**

I. Introduction

Sequential hypothesis testing is a classical problem in sequential analysis with applications ranging from clinical trials, quality control, and sensor networks. In the basic sequential hypothesis model, a single decision maker or sensor observes a random process (often assumed to i.i.d.) that depends on the underlying binary hypothesis. The decision maker takes observations until a stopping time and then declares a guess for the hypothesis. This problem was posed and solved by Wald [1], who showed that a threshold-based sequential probability ratio test is optimal. See [2]–[4] for more details. Wald's results considered binary hypothesis. Subsequently, the model with multiple hypothesis have been investigated in detail in [5]–[11] where multihypothesis version of sequential probability ratio test are developed. In recent years, various generalizations to controlled sensing (or active sequential hypothesis testing), where the decision maker also has the option to pick the observation channel (out of a finite set of possibilities) have been considered in [12]–[18].

Various approaches have been proposed to compute the optimal thresholds. Wald proposed an approximation to compute the performance of a sequential probability ratio test that provides bounds on the operating characteristics (type I and type II errors) and approximate the expected sampling time [1]. When the cost of observation is significantly smaller than the cost of making a mistake, more acurate "offshoot approximations" based on non-linear renewal theory have been proposed in [2]. For the general case, various approaches have been proposed to obtain the optimal thresholds including approximate dynamic programming [19], [20] and Markov chain approximation [21].

The common feature of the above models is that a single decision maker makes all the observations and decides when to stop. In many applications, there are multiple sensors that observe data. Such models may broadly be classified as distributed sequential hypothesis testing and decentralized sequential hypothesis testing.

In distributed sequential hypothesis testing, multiple sensors observe correlated random processes and send quantized signals to a fusion center that makes the stopping decision. Therefore, the stopping decision is made in centralized manner, while the observations are made in decentralized manner. Different configurations on *local memory* of distributed sensors and *feedback* from the fusion center are considered. See [22]–[25] for details and [26], [27] for asymptotic results.

In decentralized sequential hypothesis testing, multiple decision makers observe correlated random processes and make separate stopping decisions (which may or may not be observed by others). These stopping decisions are coupled through a common loss function. Most of the results in the literature have emphasized on identifying qualitative properties of optimal decision rules i.e., identifying whether threshold-based decision rules are optimal. See [28]–[31] for details.

It should be noted that the terms distributed and decentralized are used interchangeably in the literature. We are using the terminology followed in team-theory and making a distinction based on whether one or multiple stopping decisions are made.

In this paper, we revisit a model of decentralized sequential hypothesis testing of [28]. In this model, two (or multiple) decision makers observe correlated random processes (which, given the hypothesis, are conditionally independent across time and also conditionally independent of each other), and make separate stopping decisions that are coupled through a common loss function. An individual cost is incurred for each observation made by each sensor. It was shown that optimal decision rules are threshold based, where the thresholds of the two decision makers are coupled. The continuous time extension of this model was solved in [29].

A closely related model was investigated in [30]. The key difference in [30] is that instead of individual observation costs, there is a cost associated with the observations made by the sensors as a team. For this model, it was shown that the optimal decision rules are given by time varying convergent thresholds.

Preliminary version of this article is under review at the 2015 IEEE Conference on Decision and Control (CDC).

In [28], Wald approximation was used to compute the operating characteristics of threshold based strategies. Wald approximations are accurate only when the observation cost is much smaller than the loss function. In [30], a Monte Carlo sampling approach was proposed to estimate the operating characteristics and the distribution of the stopping time.

In this paper, we revisit the mode of decentralized sequential hypothesis testing investigated in [28] and present two search procedures that approximately compute the optimal thresholds. In the first search procedure, which we call *orthogonal search*, we iteratively solve two coupled dynamic programs to identify optimal thresholds. In the second search procedure, which we call *direct search*, we propose a method to approximately compute the performance of an arbitrary threshold-based strategy and then optimize over the choice of thresholds using a derivative-free optimization method. Both these approximations are based on approximating a continuous-state Markov process by a finite-state Markov chain. Although the idea of approximately computing the operating characteristics and stopping time by discretizing a continuous state Markov chain was also used in [21], who proposed to quantize the log-likelihood ratio (which takes value in $(-\infty, \infty)$), our approach is different because we quantize the $[0, 1]$-valued belief state.

In Section IV, we present detailed numerical experiments to compare the performance of the two procedures. In most experiments, both methods have similar performance. As far as we are aware, these are the first results on numerical methods for computing optimal thresholds for decentralized sequential hypothesis testing (i.e., when stopping decisions are made by multiple decision makers).

We use the following notation in the paper. Upper case letters (e.g., X, Y, etc.) denote random variables, the corresponding lower case letters (e.g., x, y, etc.) denote their realization and calligraphic letters ($\mathcal{X}$, $\mathcal{Y}$, etc.) denote sets. Superscripts index decision makers and subscripts index time. $Y_{1:t}$ is a short-hand for the vector $(Y_1, \dots, Y_t)$. $\mathbb{P}(\cdot)$ denotes the probability of an event and $\mathbb{E}[\cdot]$ denotes the expectation of a random variable. For a matrix A, $[A]_{nm}$ denotes the (n, m)-th element. Similarly, for a vector A, $[A]_n$ denotes the n-th element.

II. Problem formulation and structure of optimal strategies

A. The Model

Consider a decentralized sequential hypothesis problem investigated in [28]. For ease of exposition, we assume that there are two decision makers that we denote by DM^1 and DM^2; the results generalize to multiple decision makers in a natural manner. The hypothesis H takes two values h_0 and h_1 with *a priori* probability p and $1 - p$.

At time t, the DM^i, $i \in \{1, 2\}$, observes $Y_t^i \in \mathcal{Y}^i$. It is assumed that given the hypothesis $H = h_k$, $k \in \{0, 1\}$, (i) the observations $\{Y_t^i\}_{t=1}^{\infty}$ are conditionally i.i.d. with PMF or PDF f_k^i; and (ii) the observations $\{Y_t^1\}_{t=1}^{\infty}$ and $\{Y_t^2\}_{t=1}^{\infty}$ are conditionally independent.

There is no communication between the decision makers and each decision makers decides which hypothesis is true based on its local observations. In particular, at time t, DM^i, $i \in \{1, 2\}$, takes a decision $U_t^i \in \{h_0, h_1, \mathsf{C}\}$ according to

$$U_t^i = g_t^i(Y_{1:t}^i),$$

where we use the short-hand notation $Y_{1:t}^i \coloneqq (Y_1^i \cdots Y_t^i)$. The collection $g^i = (g_1^i, g_2^i, \dots)$ is called the *decision strategy* of DM^i, $i \in \{1, 2\}$.

The decision $U_t^i = h_0$ (or $U_t^i = h_1$) means that DM^i decides to stop and declare h_0 (or h_1) as the true hypothesis and makes no further observations. The decision $U_t^i = \mathsf{C}$ means that DM^i decides to take an additional observation.

Let N^i denote the stopping time when DM^i decides to stop, i.e.,

$$N^i = \min\{t \in \mathbb{N} : U_t^i \in \{h_0, h_1\}\}.$$

We denote the terminal decision $U_{N^i}^i$ by U^i.

There are two types of cost: (i) cost c^i for each observation at DM^i, and (ii) a stopping cost $\ell(U^1, U^2, H)$, which satisfies the following assumptions:

(A1) $\ell(U^1, U^2, H)$ cannot be decomposed as $\ell(U^1, H) + \ell(U^2, H)$, otherwise, the problem decomposes into two independent sequential hypothesis testing problems with one decision maker.

(A2) For any $m, n \in \{h_0, h_1\}$, $m \neq n$,

$$\ell(m, m, n) \geqslant \ell(n, m, n) \geqslant c^i \geqslant \ell(n, n, n);$$
$$\ell(m, m, n) \geqslant \ell(m, n, n) \geqslant c^i \geqslant \ell(n, n, n).$$

An example of such a loss function is:

$$\ell(u^1, u^2, h) = \begin{cases} 0, & \text{if } u^1 = u^2 = h, \\ 1, & \text{if } u^1 \neq u^2, \\ L, & \text{if } u^1 = u^2 \neq h, \quad 1 < L < \infty \end{cases}$$

This loss function implies that if both DMs make correct stopping decisions, there is no loss; if one DM makes a correct stopping decision but the other does not, the loss is 1; and if both DMs make incorrect stopping decisions, then the loss is L. We will use this loss function in the numerical experiments in Section IV.

Let $\mathcal{G}^i$ denote the set of all strategies for DM^i. Then for any choice $(g^1, g^2) \in \mathcal{G}^1 \times \mathcal{G}^2$, the total cost is

$$J(g^1, g^2; p) = \mathbb{E}[c^1 N^1 + c^2 N^2 + \ell(U^1, U^2, H)]. \quad (1)$$

We are interested in the following optimization problem:

Problem 1: Given the prior probability p, the observation PMFs f_0^i, f_1^i, the observation cost c^i, and the loss function ℓ, find a strategy (g^1, g^2) that minimizes $J(g^1, g^2; p)$ given by (1).

Note that in Problem 1, we are seeking team optimal decision strategies. For team problems, a weaker solution concept is that of person-by-person optimality (PBPO), defined below.

Definition 1 (Person-By-Person Optimality (PBPO)): A strategy (g^1, g^2) is called person-by-person optimal (PBPO) if

$$J(g^1, g^2; p) \leq J(g^1, \tilde{g}^2; p), \quad \forall \tilde{g}^2 \in \mathcal{G}^2,$$

and

$$J(g^1, g^2; p) \leq J(\tilde{g}^1, g^2; p), \quad \forall \tilde{g}^1 \in \mathcal{G}^1.$$

This gives rise to the following relaxation of Problem 1.

Problem 2: Given the prior probability p, the observation PMFs f_0^i, f_1^i, the observation cost c^i, and the loss function ℓ, find a strategy (g^1, g^2) that is person-by-person optimal (PBPO).

In general, a person-by-person optimal strategy need not be team optimal. For an example in the context of hypothesis testing, see [32]. However, very little is known regarding team optimal solutions. For that reason, we concentrate on identifying person-by-person optimal strategies. In the next section, we present qualitative properties of optimal decision rules.

B. Structure of Optimal Decision Rules

For any $i \in \{1, 2\}$, we use the game-theoretic notation and use $-i$ to denote the other decision maker. For any realization $y^i_{1:t}$ of $Y^i_{1:t}$, define

$$\pi^i_t \coloneqq \mathbb{P}(H = h_0 \mid y^i_{1:t}).$$

By Bayes rule, the update of this posterior probability is given by

$$\pi^i_{t+1} = \phi^i(\pi^i_t, y^i_{t+1}). \tag{2}$$

where

$$\phi^i(\pi^i_t, y^i_{t+1}) \coloneqq \pi^i_t f^i_0(y^i_{t+1})/q^i(y^i_{t+1} \mid \pi^i_t), \tag{3}$$

$$q^i(y^i_{t+1} \mid \pi^i_t) \coloneqq \pi^i_t f^i_0(y^i_{t+1}) + (1 - \pi^i_t) f^i_1(y^i_{t+1}). \tag{4}$$

It was shown in [28] that $\{\pi^i_t\}_{t=1}^{\infty}$ is an information state process for DM^i. In particular:

Lemma 1 ([28]): For any $i \in \{1, 2\}$ and any strategy $g^{-i} \in \mathcal{G}^{-i}$ of DM^{-i}, there is no loss of optimality for DM^i to restrict attention to strategies of the form

$$U^i_t = g^i_t(\pi^i_t). \tag{5}$$

To characterize the structure of the optimal strategy, we define the following.

Definition 2 (Threshold based strategy): A strategy of the form (5) is called threshold based if there exists thresholds $\alpha^i_t, \beta^i_t \in [0, 1]$, $\alpha^i_t \le \beta^i_t$, such that for any $\pi^i \in [0, 1]$,

$$g^i_t(\pi^i) = \begin{cases} h_1 & \text{if } \pi^i < \alpha^i_t, \\ \mathsf{C} & \text{if } \alpha^i_t \le \pi^i \le \beta^i_t, \\ h_0 & \text{if } \pi^i > \beta^i_t. \end{cases}$$

In general, the thresholds depend on the *a priori* probability p.

It was shown in [28] that threshold-based strategies are team optimal. In particular:

Lemma 2 ([28, Theorem 3.1]): For any $i \in \{1, 2\}$, and any strategy $g^{-i} \in \mathcal{G}^i$ of DM^{-i}, there is no loss of optimality in restricting attention to threshold-based strategies at DM^i.

The intuition behind the result is as follows: if we arbitrarily fix the strategy g^{-i} of DM^{-i}, then DM^i is solving a centralized sequential hypothesis testing problem with loss function

$$\begin{aligned} \hat{\ell}^i(u^i, h_k; g^{-i}, p) &\coloneqq \mathbb{E}[\ell(u^i, U^{-i}, h_k)] \\ &= \xi^{-i}_k(h_0, g^{-i}; p)\ell(u^i, h_0, h_k) \\ &\quad + \xi^{-i}_k(h_1, g^{-i}; p)\ell(u^i, h_1, h_k), \end{aligned} \tag{6}$$

where

$$\xi^i_k(u^i, g^i; p) \coloneqq \mathbb{P}(U^i = u^i \mid H = h_k; g^i, p) \tag{7}$$

is the *operating characteristic* of the decision strategy and denotes the conditional probability given hypothesis $H = h_k$ and information state p that DM^i using strategy g^i makes a terminal decision u^i. From classical results in sequential hypothesis testing, we know that for any loss function $\hat{\ell}^i(\cdot, \cdot)$, the best-response strategy at DM^i is threshold-based. Since the loss function $\hat{\ell}^i(\cdot, \cdot)$ depends on the *a priori* probability p, so do the thresholds.

Definition 3 (Time invariant strategy): A strategy $g^i = (g^i_1, g^i_2, \dots)$ is called time invariant if for any $\pi^i \in [0, 1]$, $g^i_t(\pi^i)$ does not depend on t.

For infinite-horizon problems with a single decision maker, there are time-invariant strategies that are optimal. However, that is not always the case for multiple decision makers. Nonetheless, it was shown in [28] that there are threshold-based time-invariant strategies that are PBPO.

Theorem 1 ([28, Theorem 3.2]): For any $i \in \{1, 2\}$ and any *time-invariant* and *threshold-based* strategy $g^{-i} \in \mathcal{G}^{-i}$, there is no loss of optimality in restricting attention to time-invariant and threshold-based strategies at DM^i. Moreover, the best response strategy at DM^i is given by the solution of the following dynamic program: for any $\pi^i \in [0, 1]$

$$V^i(\pi^i) = \min\{W^i_0(\pi^i, g^{-i}), W^i_1(\pi^i, g^{-i}), W^i_{\mathsf{C}}(\pi^i, g^{-i})\}, \tag{8}$$

where for $i \in \{1, 2\}$ and $k \in \{0, 1\}$,

$$W^i_k(\pi^1, g^2) = \hat{\ell}^i(h_k, h_0; g^{-i}, p)\pi^i + \hat{\ell}^i(h_k, h_1; g^{-i}, p)(1 - \pi^i) \tag{9}$$

and

$$W^i_{\mathsf{C}}(\pi^i, g^{-i}) = c^i + [\mathscr{B}^i V^i](\pi^i), \tag{10}$$

where $\mathscr{B}^i$ is the Bellman operator given by

$$[\mathscr{B}^i V^i](\pi^i) = \sum_{y^i \in \mathcal{Y}^i} q(y^i \mid \pi^i) \cdot V^i(\phi(\pi^i, y^i)), \tag{11}$$

$\phi(\pi^i, y^i)$ and $q(y^i \mid \pi^i)$ are given by (3) and (4).

For ease of notation, denote a threshold-based time-invariant strategy g^i by the tuple $\langle \alpha^i, \beta^i \rangle$. Theorem 1 gives two coupled dynamic programs, which we write succinctly as

$$\langle \alpha^1, \beta^1 \rangle = \mathcal{D}^1(\langle \alpha^2, \beta^2 \rangle) \quad \text{and} \quad \langle \alpha^2, \beta^2 \rangle = \mathcal{D}^2(\langle \alpha^1, \beta^1 \rangle). \tag{12}$$

A solution of these coupled dynamic program determines a PBPO solution for Problem 2.

Note that the terms $W^i_k(\pi^i, g^{-i})$ in the dynamic programs depend on the *a priori* probability p. Therefore, the solution $(\langle \alpha^1, \beta^1 \rangle, \langle \alpha^2, \beta^2 \rangle)$ of (12) also depends on p. This is in contrast to a single sensor setup, where the optimal thresholds do not depend on the *a priori* probability p.

Remark 1: The above approach cannot be generalized to determine team optimal solutions for the following reason. In Theorem 1, it is assumed that DM^{-i} is using a time-invariant strategy but it has not been shown that this assumption is without loss of optimality. If this assumption is removed, then the expected loss function that is seen by DM^i is not time-invariant. Consequently, the optimal strategy of DM^i would not

be time-invariant and we would need to identify optimal time-varying strategies. On the other hand, if we do assume that a time-invariant strategy is being used at DM^{-i}, then Theorem 1 shows that the best response strategy at DM^i is also time-invariant and one can seek to identify the best strategies within the class of time-invariant strategies.

C. Computing optimal threshold strategies

In this paper, we investigate two algorithms to compute person-by-person optimal strategies: orthogonal search and direct search, which are explained below. Orthogonal search is an iterative algorithm that identifies a PBPO strategy while direct search uses a non-convex optimization algorithm that tries to identify team optimal strategy (within the class of time-invariant strategies). Since both the proposed methods do not guarantee global optimality, it is not immediately clear which method is better. We perform a detailed numerical comparison of the two methods in Section IV. In most of our experiments, both algorithms give similar performance.

1) Orthogonal search: Orthogonal search is an iterative algorithm to solve the coupled dynamic programs of (12). It is conceptually similar to the iterated best response procedure to compute Nash equilibrium and the coordinate descent procedure to compute local minimum of a function. In particular, orthogonal search proceeds as follows:

1) Start with an arbitrary threshold-based strategy $(\langle \alpha^1_{(1)}, \beta^1_{(1)} \rangle, \langle \alpha^2_{(1)}, \beta^2_{(1)} \rangle)$.
2) Construct a sequence of strategies as follows:
   a) For even n:
   $$\langle \alpha^1_{(n)}, \beta^1_{(n)} \rangle = \mathcal{D}^1(\langle \alpha^2_{(n-1)}, \beta^2_{(n-1)} \rangle),$$
   and
   $$\langle \alpha^2_{(n)}, \beta^2_{(n)} \rangle = \langle \alpha^2_{(n-1)}, \beta^2_{(n-1)} \rangle.$$
   b) For odd n:
   $$\langle \alpha^1_{(n)}, \beta^1_{(n)} \rangle = \langle \alpha^1_{(n-1)}, \beta^1_{(n-1)} \rangle,$$
   and
   $$\langle \alpha^2_{(n)}, \beta^2_{(n)} \rangle = \mathcal{D}^2(\langle \alpha^1_{(n-1)}, \beta^1_{(n-1)} \rangle).$$

Theorem 2: The orthogonal search procedure described above converges to a time-invariant threshold-based strategy (g^1, g^2) that is person-by-person optimal.

See Appendix for proof.

There are two difficulties in using orthogonal search. First, the state space of the dynamic programs is continuous valued. Second, we need to compute $\xi^{-i}_k(u^{-i}, \langle \alpha^{-i}, \beta^{-i} \rangle; p)$ to compute W^i_k. We develop a procedure to approximately compute ξ^{-i}_k and approximately solve the dynamic program by using discretization.

2) Direct search: The idea behind direct search is to write the total expected cost $J(g^1, g^2; p)$ as a function of the thresholds $(\alpha^1, \beta^1, \alpha^2, \beta^2)$. In particular, for any strategy (g^1, g^2) and $i \in \{1, 2\}$, $k \in \{0, 1\}$, define:

$$\theta^i_k(g^i; p) = \mathbb{E}[N_i \mid H = h_k; g^i, p].$$

Note that θ^i_k depends on the *a priori* probability p. Then the total cost (1) is given by

$$\begin{aligned} J(g^1, g^2; p) &= p \cdot [c^1 \cdot \theta^1_0(g^1; p) + c^2 \cdot \theta^2_0(g^2; p)] \\ &\quad + (1-p) \cdot [c^1 \cdot \theta^1_1(g^1; p) + c^2 \cdot \theta^2_1(g^2; p) \\ &\quad + \sum_{u^1, u^2 \in \{h_0, h_1\}^2} \big[p \cdot \xi^1_0(u^1, g^1; p) \cdot \xi^2_0(u^2, g^2; p) \cdot \ell(u^1, u^2, h_0) \\ &\quad + (1-p) \cdot \xi^1_1(u^1, g^1; p) \cdot \xi^2_1(u^2, g^2; p) \cdot \ell(u^1, u^2, h_1) \big]. \end{aligned} \tag{13}$$

We develop a procedure to approximately compute $\theta^i_k(\langle \alpha^i, \beta^i \rangle; p)$ and $\xi^i_k(u^i, \langle \alpha^i, \beta^i \rangle; p)$. Using this procedure, we can approximately compute $J(\langle \alpha^1, \beta^1 \rangle, \langle \alpha^2, \beta^2 \rangle; p)$ for a given choice of $(\alpha^1, \beta^1, \alpha^2, \beta^2)$. In general, the total cost is not convex in the thresholds. So, we use a derivative-free non-convex optimization method to identify approximately optimal thresholds. Since the objective function is non-convex, direct search may not identify team optimal solution (within the class of time-invariant strategies).

III. Approximately computing optimal strategies

A. Approximately computing operating characteristics and expected stopping times

Each step of orthogonal search involves solving a dynamic program where the cost depends on the operating characteristic of the strategy used by DM. Similarly, irrespective of the choice of the derivative free non-convex optimization method, each step of direct search depends on the operating characteristics and expected stopping times of the strategies used by both users.

In [28], the operating characteristics and expected stopping times were computed using the Wald approximation [33], which is a good approximation only when

$$c^i \ll \min\{\ell(h_0, h_1, h_0), \ell(h_1, h_0, h_1)\}. \tag{14}$$

In this paper, we propose an alternative method to approximate the operating characteristics and expected stopping times that works well for all ranges of c^i.

Given a time-homogeneous strategy $g^i = \langle \alpha^i, \beta^i \rangle$, the information state process $\{\pi^i_t\}_{t \geq 1}$ is a continuous state Markov process. When the hypothesis is h_k, the operating characteristic $\xi^i_k(h_0, g^i; p)$ (respt., $\xi^i_t(h_1, g^i; p)$) is the probability that $\{\pi^i_t\}_{t \geq 1}$ crosses β^i before α^i (respt., crosses α^i before β^i); the expected stopping time θ^i_k is the expected stopping time for crossing the thresholds α^i or β^i. We quantize the continuous state Markov process $\{\pi^i_t\}_{t \geq 1}$ by a discrete state Markov process $\{\hat{\pi}^i_t\}_{t \geq 1}$ defined on a grid $\mathcal{S}_m = \{0, \frac{1}{m}, \frac{2}{m}, \dots, 1\}$, where $m \in \mathbb{N}$. The transition matrix P^i_k of quantized Markov process (when the true hypothesis is h_k) can be computed using a zero-order hold approximation shown in Algorithms 1 and 2. We approximately compute the operating characteristics ξ^i_k and expected stopping times θ^i_k using results from absorbing Markov chains as explained below.

Given any threshold-based strategy $g^i = \langle \alpha^i, \beta^i \rangle$ such that $\alpha^i, \beta^i \in \mathcal{S}_m$, define sets $\mathcal{A}^i_0$, $\mathcal{A}^i_1 \subset \mathcal{S}_m$ as follows:

$$\mathcal{A}^i_0 = \big\{\beta^i, \beta^i + \tfrac{1}{m}, \dots, 1\big\} \quad \text{and} \quad \mathcal{A}^i_1 = \big\{0, \tfrac{1}{m}, \dots, \alpha^i\big\}.$$

Algorithm 1: Compute transition matrices for discrete $\mathcal{Y}$

input : Discretization size m, DM i, Hypothesis h_k
output : P_k^i
Initialize $[P_k^i] = \mathbf{0}_{m\times m}$
forall $s_p \in \mathcal{S}_m$ **do**
  forall $y^i \in \mathcal{Y}^i$ **do**
    let $s_+ = \phi^i(s_p, y^i)$
    find $s_q = \arg\min_{s'\in\mathcal{S}_m} |s' - s_+|$
    let $[P_k^i]_{s_p s_q} = [P_k^i]_{s_p s_q} + f_k^i(y^i)$

Algorithm 2: Compute transition matrices for continuous $\mathcal{Y}$

input : Discretization size m, DM i, Hypothesis h_k, No. of samples J
output : P_k^i
Initialize $[P_k^i] = \mathbf{0}_{m\times m}$
forall $s_p \in \mathcal{S}_m$ **do**
  for J *times* **do**
    sample $y^i \sim f_k^i$ let $s_+ = \phi^i(s_p, y^i)$
    find $s_q = \arg\min_{s'\in\mathcal{S}_m} |s' - s_+|$
    let $[P_k^i]_{s_p s_q} = [P_k^i]_{s_p s_q} + 1$
  forall $s_q \in \mathcal{S}_m$ **do**
    normalize $[P_k^i]_{s_p s_q} = [P_k^i]_{s_p s_q}/J$

$\xi_k^i(h_0, g^i; p)$ is the probability of the event that the Markov process $\{\pi_t^i\}_{t=1}^\infty$ that starts in p, goes above the threshold β^i before it goes below the threshold α^i. This event is approximated by the event that the Markov chain with transition probability P_k^i that starts in p (which is assumed to belong to $\mathcal{S}_m$) gets absorbed in the set $\mathcal{A}_0^i$ before it is absorbed in the set $\mathcal{A}_1^i$. A similar interpretation holds for $\xi_k^i(h_1, g^i; p)$. θ_k^i is the expected stopping time until absorption.

Partition the transition matrix P_k^i as

$$P_k^i = \left[\begin{array}{c|c|c} A_k^i & B_k^i & C_k^i \\ \hline D_k^i & E_k^i & F_k^i \\ \hline G_K^i & H_k^i & K_k^i \end{array}\right]$$

where the dashed lines partition $\mathcal{S}_m$ into $\mathcal{A}_1^i$, $\mathcal{S}_m \setminus (\mathcal{A}_0^i \cup \mathcal{A}_1^i)$, and $\mathcal{A}_0^i$. The transition matrix of the absorbing Markov chain is given by

$$\hat{P}_k^i = \left[\begin{array}{c|c|c} I & 0 & I \\ \hline D_k^i & E_k^i & F_k^i \\ \hline I & 0 & I \end{array}\right].$$

Now, suppose j is the index of p in $\mathcal{S}_m$. Then, by properties of absorbing Markov chains, if $p \in (\alpha^i, \beta^i)$, then

$$\xi_k^i(h_0, \langle\alpha^i, \beta^i\rangle; p) \approx \left[(I - E_k^i)^{-1} F_k^i \mathbf{1}\right]_j \tag{15a}$$
$$\xi_k^i(h_1, \langle\alpha^i, \beta^i\rangle; p) \approx \left[(I - E_k^i)^{-1} D_k^i \mathbf{1}\right]_j \tag{15b}$$
$$\theta_k^i(\langle\alpha^i, \beta^i\rangle; p) \approx \left[(I - E_k^i)^{-1} \mathbf{1}\right]_j \tag{15c}$$

where $\mathbf{1}$ is an all ones column vector of appropriate dimensions. If $p \le \alpha^i$, then

$$\xi_k^i(h_0, \langle\alpha^i, \beta^i\rangle; p) = 0,$$
$$\xi_k^i(h_1, \langle\alpha^i, \beta^i\rangle; p) = 1,$$
$$\theta_k^i(\langle\alpha^i, \beta^i\rangle; p) = 1;$$

and if $p \ge \beta^i$, then

$$\xi_k^i(h_0, \langle\alpha^i, \beta^i\rangle; p) = 1,$$
$$\xi_k^i(h_1, \langle\alpha^i, \beta^i\rangle; p) = 0,$$
$$\theta_k^i(\langle\alpha^i, \beta^i\rangle; p) = 1.$$

B. Orthogonal search

Two approximations are needed to implement orthogonal search. First, for a given *a priori* probability p and time-invariant threshold based strategy g^{-i}, the expected loss function $\hat{\ell}^i$ (given by (6)) can be numerically approximated using (15).

Second, we need to numerically solve a dynamic program for a POMDP at each step of orthogonal search. We solve these by discretizing the continuous valued state space π^i. In particular, we approximate the Bellman operator $\mathscr{B}^i$ in (11) by using a zero-order hold as follows:

$$[\hat{\mathscr{B}}^i V^i](s) = c^i + \sum_{s_+\in\mathcal{S}_m} [P_*^i]_{s s_+} V(s_+).$$

where P_*^i is the transition matrix obtained by quantizing the continuous state Markov kernel $q(y^i \mid \pi^i)$ (which may be done by replacing f_k^i with $q^i(y^i|s_p)$ in Algorithms 1 and 2). The approximation $\hat{\mathscr{B}}^i$ is similar to the discretization described in [34], where the corresponding error bounds corresponding error bounds were also derived.

C. Direct search

For a given *a priori* probability p, the optimal time-invariant threshold based strategy is given by

$$\arg\min_{(\alpha^1,\beta^1,\alpha^2,\beta^2)} J(\langle\alpha^1, \beta^1\rangle, \langle\alpha^2, \beta^2\rangle; p)$$

where $J(\langle\alpha^1, \beta^1\rangle, \langle\alpha^2, \beta^2\rangle; p)$ is given by (13) and can be numerically approximated using (15). In general, team problems are non-convex in strategy space; so we expect $J(\langle\alpha^1, \beta^1\rangle, \langle\alpha^2, \beta^2\rangle; p)$ to be non-convex in the parameters $(\alpha^1, \beta^1, \alpha^2, \beta^2)$. For an example, see [32].

In principle, such non-convex optimization problems can be solved using *derivative-free* methods that do no use numerical or analytic gradients (see [35]). Since the optimization problem is non-convex, such algorithms cannot guarantee convergence to a globally optimal solution. In the numerical results we use Nelder-Mead simplex algorithm [36], which is implemented as `fminsearch` function in MATLAB.

Note that the transition probabilities P_k^i, $k \in \{0, 1\}$, approximate the evolution of the $\{\pi_t^i\}_{t=1}^\infty$ process when hypothesis $H = h_k$ is true. We will use these to approximate probabilities ξ_k^i. On the other hand, the transition probability P_*^i approximates the uncontrolled evolution of $\{\pi_t^i\}_{t=1}^\infty$. We will use this to approximately solve the dynamic program of Lemma 1.

D. Computational complexity

Although an exact analysis of the complexity of the proposed algorithms is difficult, we explain the complexity of some of the steps.

Given a grid $\mathcal{S}_m$ of size m, computing the operating characteristics and stopping time involves matrix inversion, or more precisely, solving a linear systems of equations. The complexity of this operating is $O(m^3)$.

For a fixed p, the initialization step (computing $\hat{\ell}$) of orthogonal search requires computing the operating characteristics. So, it has $O(m^3)$ complexity. In each step of the orthogonal search, a dynamic program with state space m needs to be solved, which involves $O(nm^2)$ operations, where n is an estimate of number of steps needed for value iteration to converge. The exact bound on n depends on the structure of the transition matrix P^i_* (see [37]). We are not aware of any bounds on the number of steps needed for orthogonal search to converge.

For a fixed p, each step of direct search requires computing the operating characteristics and stopping times. So, it has $O(m^3)$ complexity. There are bounds on the number of steps needed for direct search to converge [38], but in practice it depends on the shape of the function being minimized.

IV. Numerical comparison of the two methods

Both the approaches presented in this paper do not guarantee global optimality. In this section, we compare their performance on a benchmark system in which $\mathcal{Y}^1 = \mathcal{Y}^2 = \{0, 1\}$ and the loss function is of the form

$$\ell(u^1, u^2, h) = \begin{cases} 0, & \text{if } u^1 = u^2 = h, \\ 1, & \text{if } u^1 \neq u^2, \\ L, & \text{if } u^1 = u^2 \neq h \end{cases} \tag{16}$$

where $L \geq 1$. We select several prior probability p and compare the performance of direct search and orthogonal search. For both methods, we use $m = 1000$.

Note that the choice of parameters (c^1, c^2, L) and observation distributions $(f^1_0, f^1_1, f^2_0, f^2_1)$ completely specifies the model. We compare the performance $J_{OS}(p)$ of orthogonal search with the performance $J_{DS}(p)$ of direct search: first for a specific values of parameters and distributions and then for randomly chosen values. For each randomly chosen instance, we identify a threshold-based strategy using orthogonal and direct search for each $p \in [0 : 0.05 : 1]$. We compare $J_{OS}(p)$ and $J_{DS}(p)$ in two different ways: L^2 norm and L^∞ norm. In particular, we evaluate

$$\Delta J^2_{OS} = \frac{\|J_{OS} - J_{DS}\|_2}{\|J_{OS}\|_2}, \quad \Delta J^2_{DS} = \frac{\|J_{DS} - J_{OS}\|_2}{\|J_{DS}\|_2},$$
$$\Delta J^\infty_{OS} = \frac{\|J_{OS} - J_{DS}\|_\infty}{\|J_{OS}\|_\infty}, \quad \Delta J^\infty_{DS} = \frac{\|J_{DS} - J_{OS}\|_\infty}{\|J_{DS}\|_\infty}.$$

and plot their histogram for comparison.

A. General Loss Case

We first consider specific values of all the parameters: $c^1 = c^2 = 0.05$, $L = 4$ and

$$f^1_0 = \begin{bmatrix} 0.25 & 0.75 \end{bmatrix}, \qquad f^2_0 = \begin{bmatrix} 0.80 & 0.20 \end{bmatrix},$$
$$f^1_1 = \begin{bmatrix} 0.60 & 0.40 \end{bmatrix}, \qquad f^2_1 = \begin{bmatrix} 0.30 & 0.70 \end{bmatrix}.$$

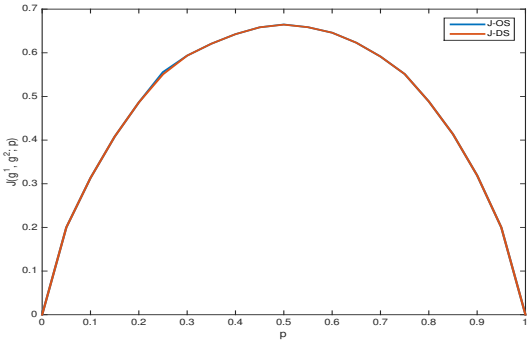


Fig. 1. The plots of $J_{OS}(p)$ and $J_{DS}(p)$ for a specific choice of parameters

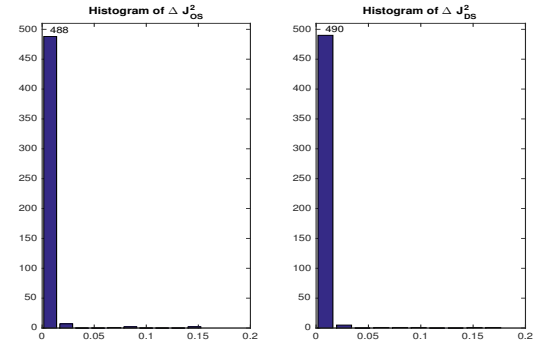


Fig. 2. Plot of ΔJ^2_{OS} and ΔJ^2_{DS} for 500 random values of parameters

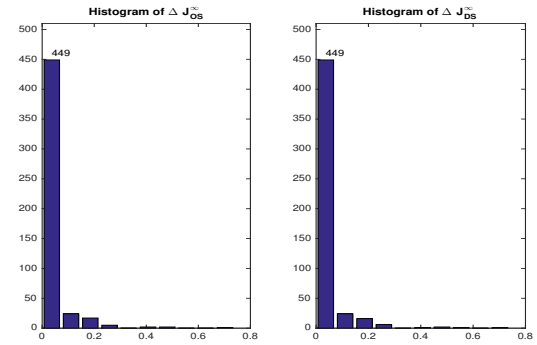


Fig. 3. Plot of ΔJ^∞_{OS} and ΔJ^∞_{DS} for 500 random values of parameters

For each $p \in [0 : 0.05 : 1]$, we identify a threshold-based strategy using orthogonal search and direct search. The plots of the performance $J_{OS}(p)$ and $J_{DS}(p)$, given in Fig. 1, show that the two methods yield same performance for all p.

Next, we compare the performance of the two methods for 500 randomly chosen values of the parameters (c^1, c^2, L) and $(f^1_0, f^1_1, f^2_0, f^2_1)$. Specifically, we use $c^1, c^2 \sim \text{unif}[0, 0.05]$, $L \sim \text{unif}[1, 4]$. We pick f^i_k by picking a random number $\delta^i_k \sim \text{unif}[0, 1]$ and setting $f^i_k = [\delta^i_k, 1 - \delta^i_k]$.

The histograms of (J^2_{OS}, J^2_{DS}) and $(J^\infty_{OS}, J^\infty_{DS})$ are plotted in Figs. 2 and 3. For more than 97% of the cases, $J_{OS}(\cdot)$ and $J_{DS}(\cdot)$ are within 0.01% of each other in L^2 norm. For almost 90% of the cases, they are within 0.05% of each other in L^∞ norm.

The fact that these two methods converge to similar solutions

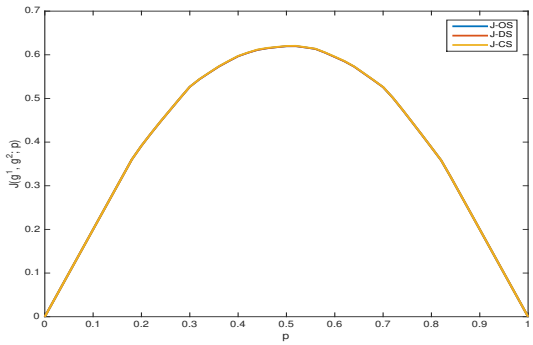


Fig. 4. The plots of $J_{OS}(p)$, $J_{DS}(p)$, and $J_{CS}(p)$ for a specific choice of parameters

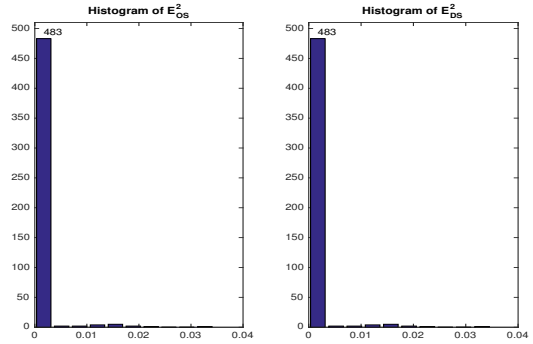


Fig. 5. Plot of E^2_{OS} and E^2_{DS} for 500 random values of parameters

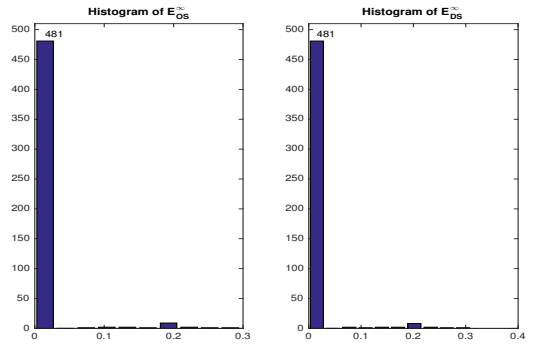


Fig. 6. Plot of E^∞_{OS} and E^∞_{DS} for 500 random values of parameters

strongly suggests that the resulting solution is also team optimal (within the class of time-invariant threshold-based strategies). However, such a claim cannot be checked because there is no general method to compute team optimal solution. In the following, we compare our results with two other methods: the centralized solution with $L = 2$ and the Wald approximation.

B. Decomposable Case

We consider the special case when $L = 2$ and, therefore, the total cost $\ell(U^1, U^2, H)$ may be written as $\ell(U^1, H)+\ell(U^2, H)$. Hence the decentralized sequential hypothesis testing problem decomposes into two independent centralized sequential hypothesis testing problem. We call this *the decomposable case*. In this case, the team optimal solution can be obtained by separately solving the two separate centralized sequential hypothesis testing problems, which can be solved using value iteration. We refer to this solution as *centralized solution* and denote it by $J_{CS}(p)$.

We first consider specific values of all parameters: $L = 2$ and $c^1, c^2, f^1_0, f^1_1, f^2_0, f^2_1$ are the same as the values in Section IV-A. For each $p \in [0 : 0.05 : 1]$, we identify a threshold-based strategy using orthogonal search, direct search, and centralized solution. The plots of the performance $J_{OS}(p)$, $J_{DS}(p)$ and $J_{CS}(p)$, given in Fig. 4, show that the three methods yield same performance for all p.

Next, we compare the two methods with the centralized solution for 500 randomly chosen values of the parameters. We set $L = 2$ and choose the other parameters as in Sec. IV-A. To compare the solution, define relative errors

$$E^2_{OS} = \frac{\|J_{CS} - J_{OS}\|_2}{\|J_{CS}\|_2}, \qquad E^2_{DS} = \frac{\|J_{CS} - J_{DS}\|_2}{\|J_{CS}\|_2},$$
$$E^\infty_{OS} = \frac{\|J_{CS} - J_{OS}\|_\infty}{\|J_{CS}\|_\infty}, \qquad E^\infty_{DS} = \frac{\|J_{CS} - J_{DS}\|_\infty}{\|J_{CS}\|_\infty}.$$

The histograms of (E^2_{OS}, E^2_{DS}) and $(E^\infty_{OS}, E^\infty_{DS})$ are plotted in Figs. 5 and 6. For more than 96% of the cases, $J_{OS}(\cdot)$ and $J_{DS}(\cdot)$ are within 0.005% of $J_{CS}(\cdot)$ in L^2 norm and are within 0.05% of $J_{CS}(\cdot)$ in L^∞ norm. Thus, both orthogonal search and direct search have the same performance as the centralized (optimal) solution.

C. Comparison with Wald approximation

In [28], the Wald approximation was used to compute person-by-person optimal strategies. Given the *a priori* probability p and a threshold based strategy $g^i = \langle \alpha^i, \beta^i \rangle$, define

$$A^i = \frac{p}{1-p}\frac{1-\beta^i}{\beta^i} \quad \text{and} \quad B^i = \frac{p}{1-p}\frac{1-\alpha^i}{\alpha^i}.$$

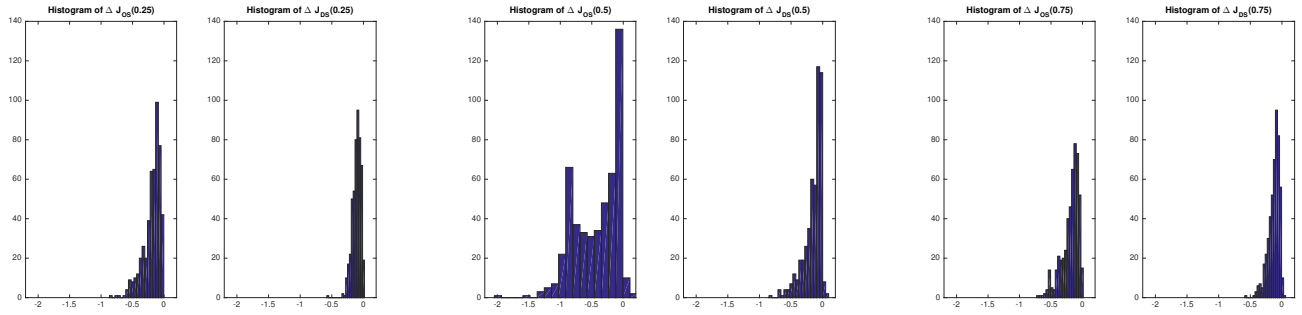
Then, according to the Wald approximation [1], [33],

$$\xi^i_0(h_1, g^i; p) \approx \frac{1 - A^i}{B^i - A^i} \tag{17a}$$
$$\xi^i_1(h_0, g^i; p) \approx \frac{A^i(B^i - 1)}{B^i - A^i} \tag{17b}$$
$$\theta^i_0(g^i; p) \approx -\frac{(\log A^i)(B^i - 1) + (\log B^i)(1 - A^i)}{(B^i - A^i)D(f^i_0 \| f^i_1)} \tag{17c}$$
$$\theta^i_1(g^i; p) \approx \frac{A^i(\log A^i)(B^i - 1) + B^i(\log B^i)(1 - A^i)}{(B^i - A^i)D(f^i_1 \| f^i_0)} \tag{17d}$$

where $D(P\|Q)$ is the Kullback-Leibler divergence between distributions P and Q.

Instead of the proposed Markov chain approximation (15), one can use the Wald approximations (17) in both orthogonal and direct search. Let $J^W_{OS}(p)$ and $J^W_{DS}(p)$ denote the corresponding performance. In this section, we compare the proposed Markov chain approximations with the Wald approximation for both methods.

We choose 500 random values for (c^1, c^2, L) and $(f^1_0, f^1_1, f^2_0, f^2_1)$ as described in Sec. IV-A. To compare the performance of the Wald approximation with the Markov chain



(a) $p = 0.25$ (b) $p = 0.5$ (c) $p = 0.75$

Fig. 7. Comparison of the performance of orthogonal and direct search using Markov chain approximation and using Wald approximation for different values of p

approximation, we use the following metrics: for an arbitrary $p \in [0, 1]$,

$$\Delta J_{OS}(p) = \frac{J_{OS}(p) - J^W_{OS}(p)}{J_{OS}(p)}$$
$$\Delta J_{DS}(p) = \frac{J_{DS}(p) - J^W_{DS}(p)}{J_{DS}(p)}.$$

The histograms of $\Delta J_{OS}(p)$ and $\Delta J_{DS}(p)$ for different values of p are shown in Fig. 7. For almost all cases, the Markov chain approximation performs better than the Wald approximation. Note that the random values are chosen in a range where (c^1, c^2) are about a factor of 100 smaller than L. So, one expects that the Markov chain approximation will perform significantly better when (c^1, c^2) are of the same order as L.

D. Computational Complexity

The proposed search algorithms consist of two parts: an initial step of computing the transition matrix (Algorithm 1) and an iterative step of using the transition matrix to compute the thresholds. The complexity of the first step is linear in $|\mathcal{Y}|$, but the complexity of the second iterative step does not depend on $|\mathcal{Y}|$. Therefore, one would not expect a significant increase in the run-time with an increase in the size of the observations $|\mathcal{Y}|$. This is confirmed numerically as well.

We consider four cases: $|\mathcal{Y}| = 2$, $|\mathcal{Y}| = 4$, $|\mathcal{Y}| = 8$, $|\mathcal{Y}| = 16$. For each case, we run 100 simulations where c^1, c^2, L are chosen randomly as described in Sec. IV-A and $p = [0 : 0.05 : 1]$. To choose f^i_h, for $|\mathcal{Y}| = m$, we pick m random numbers $(\delta^i_{k0}, \dots, \delta^i_{km}) \sim \text{unif}[0, 1]$ and set $f^i_h = [\delta^i_{k0}, \dots, \delta^i_{km}]/S^i_k$, where $S^i_k = \sum_{j=1}^m \delta^i_{kj}$. The median run-time for orthogonal search and direct search is shown in Table I. As expected, increasing $|\mathcal{Y}|$ has a small effect on runtime.

TABLE I
RUNTIME WITH RESPECT TO NUMBER OF OBSERVATIONS.

	$\lvert\mathcal{Y}\rvert = 2$	$\lvert\mathcal{Y}\rvert = 4$	$\lvert\mathcal{Y}\rvert = 8$	$\lvert\mathcal{Y}\rvert = 16$
OS median	24.9412s	29.8904s	34.9856s	39.7705s
DS median	25.0533s	36.1206s	51.1771s	62.7652s

V. DISCUSSION

In this paper, we proposed two methods to approximately compute the optimal threshold-based strategies in decentralized sequential hypothesis testing. Both these methods are based on discretization of the continuous-valued information state process by a finite-valued Markov chain. The orthogonal search method computes PBPO strategies while the direct search methods attempts to compute team optimal strategies. Direct search involves solving a non-convex optimization problem, so in practice, it will also not converge to a global optimum. In our numerical study, both algorithms identify strategies with similar performance.

These results generalize naturally to multiple hypothesis and multiple decision makers, but, as expected, accompanied by an increase in computational complexity. An interesting future direction is to develop procedures to compute optimal thresholds for more general models of sequential hypothesis testing such as those considered in [31], [39].

ACKNOWLEDGMENTS

This work was partly supported by Natural Science and Engineering Research Council of Canada (NSERC) Discovery Grant 402753-11.

APPENDIX A
PROOF OF THEOREM 2

The dynamic program of Theorem 1 is from the point of view of a single DM; it does not minimize $J(g^1, g^2; p)$ rather for a given strategy g^{-i} of DM^{-i}, it minimizes

$$J^i(g^1, g^2; p) \coloneqq \mathbb{E}[c^i N^i + \ell(U^1, U^2, H)].$$

Note that

$$J(g^1, g^2; p) = J^i(g^1, g^2; p) + \mathbb{E}[c^{-i} N^{-i}]$$

where the second term only depends on g^{-i} (which is fixed).

For even n, $g^1_{(n)} = \mathcal{D}^1(g^2_{(n)})$ and $g^2_{(n)} = g^2_{(n-1)}$. Therefore,

$$J^1(g^1_{(n)}, g^2_{(n)}; p) = J^1(g^1_{(n)}, g^2_{(n-1)}; p) \le J^1(g^1_{(n-1)}, g^2_{(n-1)}; p),$$

or, equivalently,

$$J(g^1_{(n)}, g^2_{(n)}; p) \le J(g^1_{(n-1)}, g^2_{(n-1)}; p).$$

Similarly, for odd n

$$J(g^1_{(n)}, g^2_{(n)}; p) \leq J(g^1_{(n-1)}, g^2_{(n-1)}; p).$$

Thus, at every step n,

$$J(g^1_{(n)}, g^2_{(n)}; p) \leq J(g^1_{(n-1)}, g^2_{(n-1)}; p),$$

Therefore, the sequence $\{J(g^1_{(n)}, g^2_{(n)}; p)\}$ is a decreasing sequence lower bounded by 0. Hence, a limit exists. By definition, the limiting strategy is PBPO.

REFERENCES

[1] A. Wald, *Sequential hypothesis testing*. Wiley, 1947.

[2] D. Siegmund, *Sequential analysis: tests and confidence intervals*. Springer-Verlag, 1985.

[3] R. S. Lipster and A. N. Shiryayev, *Statistics of Random Processes, Vol. II:Applications*. Springer-Verlag, 1977.

[4] A. N. Shiryaev and A. Aries, *Optimal stopping rules*. Springer-Verlag, 2007, vol. 8.

[5] H. Chernoff, "Sequential design of experiments," *The annals of mathematical statistics*, pp. 755–770, Sep. 1959.

[6] ——, *Sequential Analysis and Optimal Design*. SIAM, 1972.

[7] J. Kiefer and J. Sacks, "Asymptotically optimum sequential inference and design," *The Annals of Mathematical Statistics*, pp. 705–750, 1963.

[8] C. Baum and V. Veeravalli, "A sequential procedure for multihypothesis testing," *IEEE Transactions on Information Theory*, vol. 40, no. 6, pp. 1994–2007, 1994. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=340472

[9] V. Veeravalli and C. Baum, "Asymptotic efficiency of a sequential multihypothesis test," *IEEE Transactions on Information Theory*, vol. 41, no. 6, pp. 1994–1997, 1995. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=476323

[10] V. Draglia, A. G. Tartakovsky, and V. V. Veeravalli, "Multihypothesis sequential probability ratio tests. i. asymptotic optimality," *IEEE Transactions on Information Theory*, vol. 45, no. 7, pp. 2448–2461, 1999.

[11] V. Dragalin, A. Tartakovsky, and V. Veeravalli, "Multihypothesis sequential probability ratio tests. ii. accurate asymptotic expansions for the expected sample size," *IEEE Transactions on Information Theory*, vol. 46, no. 4, pp. 1366–1383, 2000. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=850677

[12] S. Nitinawarat, G. Atia, and V. Veeravalli, "Controlled sensing for multihypothesis testing," *IEEE Transactions on Automatic Control*, vol. 58, no. 10, pp. 2451–2464, 2013. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6513292

[13] M. Naghshvar, T. Javidi *et al.*, "Active sequential hypothesis testing," *The Annals of Statistics*, vol. 41, no. 6, pp. 2703–2738, 2013.

[14] M. Naghshvar and T. Javidi, "Sequentiality and adaptivity gains in active hypothesis testing," *IEEE Journal of Selected Topics in Signal Processing*, vol. 7, no. 5, pp. 768–782, 2013. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6512566

[15] S. Nitinawarat and V. V. Veeravalli, "Controlled sensing for sequential multihypothesis testing with controlled markovian observations and non-uniform control cost," *Sequential Analysis*, vol. 34, no. 1, pp. 1–24, 2015.

[16] K. Cohen and Q. Zhao, "Asymptotically optimal anomaly detection via sequential testing," *IEEE Transactions on Signal Processing*, vol. 63, no. 11, pp. 2929–2941, 2015. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=7067439

[17] ——, "Active hypothesis testing for anomaly detection," *IEEE Transactions on Information Theory*, vol. 61, no. 3, pp. 1432–1450, 2015. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=7001595

[18] N. K. Vaidhiyan and R. Sundaresan, "Active search with a cost for switching actions," *arXiv preprint arXiv:1505.02358*, 2015.

[19] D. P. Bertsekas, *Dynamic Programming and Optimal Control*. Belmont, MA: Athena Scientific, 1995, vol. 1.

[20] K. Cohen and A. Leshem, "Enery-efficient detection in wireless sensor networks using likelihood ratio and channel state information," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 5, pp. 1671–1683, 2011.

[21] W. H. Woodall and M. R. Reynolds Jr, "A discrete markov chain representation of the sequential probability ratio test," *Sequential Analysis*, vol. 2, no. 1, pp. 27–44, 1983.

[22] H. Hashemi and I. B. Rhodes, "Decentralized sequential detection," *Information Theory, IEEE Transactions on*, vol. 35, no. 3, pp. 509–520, 1989.

[23] V. V. Veeravalli, "Comments on "decentralized sequential detection"," *IEEE Transactions on Signal Processing*, vol. 38, no. 4, pp. 1428–1429, 1992.

[24] V. V. Veeravalli, T. Başar, and H. V. Poor, "Decentralized sequential detection with a fusion center performing the sequential test," *IEEE Transactions on Information Theory*, vol. 39, no. 2, pp. 433–442, 1993.

[25] J. N. Tsitsiklis, "On threshold rules in decentralized detection," in *IEEE Conference on Decision and Control*. IEEE, 1986, pp. 232–236.

[26] Y. Mei, "Asymptotic optimality theory for decentralized sequential hypothesis testing in sensor networks," *IEEE Transactions on Information Theory*, vol. 54, no. 5, pp. 2072–2089, May 2008.

[27] G. Fellouris and G. V. Moustakides, "Decentralized sequential hypothesis testing using asynchronous communication," *IEEE Transactions on Information Theory*, vol. 57, no. 1, pp. 534–548, Jan. 2011.

[28] D. Teneketzis and Y. Ho, "The decentralized Wald problem," *Information and Computation (formerly Information and Control)*, vol. 73, no. 1, pp. 23–44, Apr. 1987.

[29] A. LaVigna, A. M. Makowski, and J. S. Baras, "A continuous-time distributed version of wald's sequential hypothesis testing problem," in *Analysis and Optimization of Systems*, ser. Lecture Notes in Control and Information Sciences, A. Bensoussan and J. Lions, Eds. Springer Berlin Heidelberg, 1986, vol. 83, pp. 533–543.

[30] V. V. Veeravalli, T. Başar, and H. V. Poor, "Decentralized sequential detection with sensors performing the sequential test," *Mathematics of Control, Signals and Systems*, vol. 7, no. 4, pp. 292–305, 1994.

[31] A. Nayyar and D. Teneketzis, "Sequential problems in decentralized detection with communication," *IEEE Transactions on Information Theory*, vol. 57, no. 8, pp. 5410–5435, Aug. 2011.

[32] R. R. Tenney and N. R. Sandell Jr, "Detection with distributed sensors," *IEEE Transactions on Aerospace Electronic Systems*, vol. 17, pp. 501–510, 1981.

[33] M. DeGroot, *Optimal Statistical Decisions*. McGraw Hills, 1970.

[34] D. P. Bertsekas, "Convergence of discretization procedures in dynamic programming," *IEEE Transactions on Automatic Control*, vol. 20, pp. 415–419, Jun. 1975.

[35] L. M. Rios and N. V. Sahinidis, "Derivative-free optimization: a review of algorithms and comparison of software implementations," *Journal of Global Optimization*, vol. 56, no. 3, pp. 1247–1293, 2013.

[36] J. A. Nelder and R. Mead, "A simplex method for function minimization," *The Computer Journal*, vol. 7, no. 4, pp. 308–313, 1965.

[37] L. K. Michael Littman, Thomas Dean, "On the complexity of solving markov decision problems," in *Proc. of Uncertainity in Artificial Intelligence (UAI)*, 1995.

[38] S. Singer and S. Singer, "Complexity analysis of nelder-mead search interations," in *Proc. Conf. App. Math. and Comp.*, 1999.

[39] A. Nayyar and D. Teneketzis, "Signaling in sensor networks for sequential detection," *IEEE Trans. Control of Network Systems*, vol. 2, no. 1, pp. 36–46, Mar. 2015.