On computing optimal thresholds in decentralized sequential hypothesis testing

Can Cui and Aditya Mahajan

***Abstract*— There has been a lot of recent progress in understanding the structure of optimal control strategies in decentralized stochastic control, but not much is known about computational methods. In this paper, we propose two finite-state approximation methods for decentralized sequential hypothesis testing. 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 the discretization of the 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.**

I. Introduction

Sequential hypothesis testing was first formulated by Wald [1] for efficient testing of anti-aircraft guns in World War II. Since then, the theory of sequential hypothesis testing has been applied to various applications including sensor networks, intrusion detection in surveillance networks, primary channel detection in cognitive radio, quality control, and clinical trials.

In many of the modern applications, multiple nodes observe noisy information about the system state. Communicating all the observations to a single node is often too expensive and impractical. Thus, decisions need to be made in a decentralized manner by decision makers that share a common objective. Such decentralized problem are investigated using team decision theory.

Decentralized sequential hypothesis testing has received considerable attention in the literature [2]–[6]. The main emphasis is on identifying qualitative properties of optimal decision strategies. In particular, in identifying information state (or sufficient statistics) of the data available at the decision maker and in establishing the structure of optimal decision rules, e.g., showing that threshold-based strategies (similar in spirit to Wald's sequential likelihood ratio test [1]) are optimal. We summarize some of these results below.

A model of decentralized hypothesis testing problems in which multiple decision makers make individual decisions that are coupled through a common loss function was considered in [2], [3]. It was shown that optimal decision rules are characterized by two thresholds. A model in which the decision makers communicate their decision to a centralized fusion center was considered in [4]. The optimal decision rules in this case are also characterized by two thresholds. Models in which the decision makers communicate with one-another have been considered in [5], [6]. It is shown that for some decision makers, the optimal decision rules are characterized by four thresholds.

In spite of the significant advances in understanding the optimality of threshold-based strategies in decentralized hypothesis testing, not much is known about numerical methods to compute these thresholds. This reflects the general state of affairs in decentralized stochastic control. There has been a lot of recent progress of identifying structure of optimal control strategies, but not much is known about numerically identifying optimal strategies that have that structure. In this paper, we revisit the decentralized sequential hypothesis testing model of [2], [3] and develop numerical techniques to compute the optimal thresholds. These numerical techniques might be useful for general decentralized control problems as well.

Can Cui and Aditya Mahajan are with the Department of Electrical and Computer Engineering, McGill University, Montreal, QC, CA. Email: {can.cui@mail.mcgill.ca, aditya.mahajan.mcgill.ca}.

II. Problem Formulation And Structure of Optimal Strategies

A. The Model

Consider a decentralized sequential hypothesis problem investigated in [2]. 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 $H = h_k$, $k \in \{1,2\}$, (i) the observations $\{Y_t^i\}_{t=1}^\infty$ are conditionally i.i.d. with PMF 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 $Y_{1:t}^i := (Y_1^i \cdots Y_t^i)$.

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.,

$$N^i = \min\{t \in \mathbb{Z} > 0 : U^i_t \in \{h_0, h_1\}\}.$$

We denote the terminal decision $U^i_{N^i}$ 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:

$$\ell(u^1, u^2, h) = \begin{cases} 0, & \text{if } u^1 = u^2 = h, \\ L_1, & \text{if } u^1 \neq u^2, \\ L_2, & \text{if } u^1 = u^2 \neq h; \end{cases}$$

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 L_1; and if both DMs make incorrect stopping decisions, then the loss is L_2. We will use this loss function in the numerical experiments in Section V.

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^i_0, f^i_1, 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^i_0, f^i_1, 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 [7]. 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 $i \in \{1, 2\}$, let $-i$ denote the other decision maker. For any realization $y^i_{1:t}$ of $Y^i_{1:t}$, define the information state

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

In addition, define

$$q^i(y^i_{t+1} \mid \pi^i_t) := \pi^i_t f^i_0(y^i_{t+1}) + (1 - \pi^i_t) f^i_1(y^i_{t+1}), \quad (2)$$
$$\phi^i(\pi^i_t, y^i_{t+1}) := \pi^i_t f^i_0(y^i_{t+1}) / q^i(y^i_{t+1} \mid \pi^i_t). \quad (3)$$

Then, by Bayes' rule, the update of the information state is given by

$$\pi^i_{t+1} = \phi^i(\pi^i_t, y^i_{t+1}). \quad (4)$$

For ease of notation, for any $i \in \{1, 2\}$, $k \in \{0, 1\}$, $u^i \in \{h_0, h_1\}$, and $g^i \in \mathcal{G}^i$, define

$$\xi^i_k(u^i, g^i; p) = \mathbb{P}(U^i = u^i \mid H = h_k; g^i, p), \quad (5)$$

which is called 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.

It was shown in [2] that $\{\pi^i_t\}_{t=1}^{\infty}$ is an sufficient statistic for DM^i. In particular:

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

$$U^i_t = g^i_t(\pi^i_t). \quad (6)$$

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

Definition 2 (Threshold based strategy): A strategy of the form (6) is called threshold based if there exists thresholds $\alpha^i_t, \beta^i_t \in [0, 1]$, $\alpha^i_t \leq \beta^i_t$, such that for any $\pi^i \in [0, 1]$,

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

It was shown in [2] that threshold-based strategies are team optimal. In particular:

Proposition 2 ([2, Theorem 3.1]): For any $i \in \{1, 2\}$, and any strategy $g^{-i} \in \mathcal{G}^i$, 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

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

From classical results in sequential hypothesis testing, we know that for any loss function of this form, the optimal strategy is threshold-based.

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 [2] that there are threshold-based time-invariant strategies that are PBPO.

Proposition 3 ([2, 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 $\pi^i \in [0,1]$

$$V^i(\pi^i) = \min\{W^i_0(\pi^i, g^{-i}), W^i_1(\pi^i, g^{-i}), W^i_{\mathsf{C}}(\pi^i, g^{-i})\}, \tag{7}$$

where for $k \in \{0,1\}$,

$$W^1_k(\pi^1, g^2) = \sum_{u^2 \in \{h_0, h_1\}} \big[\xi^2_0(u^2, g^2; \pi^1) \cdot \pi^1 \cdot \ell(h_k, u^2, h_0) + \xi^2_1(u^2, g^2; \pi^1) \cdot (1-\pi^1) \cdot \ell(h_k, u^2, h_1)\big], \tag{8}$$

W^2_k is defined similarly, and

$$W^i_{\mathsf{C}}(\pi^i, g^{-i}) = c^i + [\mathscr{B}^i V^i](\pi^i), \tag{9}$$

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

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

$q(y^i \mid \pi^i)$ and $\phi(\pi^i, y^i)$ are given by (2) and (3).

For ease of notation, denote a threshold-based strategy g^i by the tuple $\langle \alpha^i, \beta^i \rangle$. Theorem 3 gives two coupled dynamic programs, which we write succinctly as

$$\langle \alpha^1, \beta^1 \rangle = \mathcal{D}^1(\langle \alpha^2, \beta^2 \rangle) \quad \text{and} \quad \langle \alpha^2, \beta^2 \rangle = \mathcal{D}^2(\langle \alpha^1, \beta^1 \rangle). \tag{10}$$

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

Note that the above approach cannot be generalized to determine team optimal solutions for the following reason. In Theorem 3, 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 3 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. Computation of PBPO Strategies

In [2], an approximate solution for these coupled dynamic programs was presented under the assumption that f^i_k are Gaussian distributions and

$$c^i \ll \min\{\ell(h_0, h_1, h_0), \ell(h_1, h_0, h_1)\}. \tag{11}$$

Under (11), the stopping time $N^i \gg 1$, and one can use the asymptotic expressions of Type I and Type II errors in [1], [8]. This method can be extended to the non-Gaussian distributions using the expressions in [9], but it is not a good approximation when the stopping cost does not satisfy (11).

We do not assume (11) rather approximately solve the dynamic programs. The information state of these dynamic programs is continuous valued. To solve them numerically, we discretize the state space on a uniform grid. The operating characteristics ξ^i_k also needs to be computed to solve these dynamic programs. ξ^i_k is the probability that a discrete-time Markov process crosses a threshold and there are various methods to compute it approximately. One option is to use the asymptotic expressions of [1] (also see [8]); another option is to solve an appropriate Fredholm Integral equation (see [10]). To be consistent with the approximation that we use to solve the dynamic program, we compute ξ^i_t by approximating the discrete-time continuous-state Markov process by a *finite state* Markov chain.

Even within the class of time-invariant strategies, orthogonal search only identifies PBPO strategies. We propose a second method, which we call *direct search*, to approximate team optimal strategies (within the class of time-invariant strategies). To do so, we approximate the expected cost $J(\langle \alpha^1, \beta^1 \rangle, \langle \alpha^2, \beta^2 \rangle; p)$ (which is given by (1)) of an arbitrary threshold-based strategy by approximating the evolution of π^i_t, which is a discrete-time continuous-state Markov process, by a *finite-state* Markov process. In general, this cost is not convex in the threshold. So, we use a derivative-free non-convex optimization method to identify the optimal thresholds. Since the objective function is non-convex, direct search can only identify locally optimal solution (within the class of time-invariant strategies).

Both the proposed methods only guarantee local optimality. So, it is not immediately clear which method is better. We perform a numerical comparison of the two methods in Section V. In most of our experiments, direct search outperforms orthogonal search; sometimes significantly.

III. Method I: Orthogonal search

The coupled dynamic programs of (10) may be solved using *orthogonal search* procedure described below.

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

$$\langle \alpha^2_{(n)}, \beta^2_{(n)} \rangle = \mathcal{D}^2(\langle \alpha^1_{(n-1)}, \beta^1_{(n-1)} \rangle).$$

Note that orthogonal search is conceptually similar to the iterated best response procedure to compute Nash equilibirum and the coordinate descent procedure to compute local mininum of a function.

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

Proof: For any strategy (g^1, g^2), define

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

and

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

Then, by construction, for even n

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

and, for odd n

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

Thus, at every step n,

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

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

There are two difficulties in using orthogonal search. First, at step n, we need to compute ξ^i_k for a threshold-based strategy $\langle \alpha^i_{(n)}, \beta^i_{(n)} \rangle$. Second, the dynamic program at step n is a POMDP. So, we either need to discretize the state-space, or use the point-based methods [11], [12].

We use discretization to approximately compute ξ^i_k and solve the dynamic program as well. In principle, other approaches can be used for both approximations.

A. The discretization procedure

For any $m \in \mathbb{N}$, define $\mathcal{S}_m = \{0, \frac{1}{m}, \frac{2}{m}, \dots, 1\}$. For any $i \in \{0, 1\}$, we approximate the $[0, 1]$-valued Markov process $\{\pi^i_t\}_{t=1}^{\infty}$, which is given by (4), by a $\mathcal{S}_m$-valued Markov chain. We consider three approximations that make different assumptions on probability distribution of Y^i. We denote the corresponding transition probabilities by P^i_0, P^i_1, and P^i_*. For P^i_k, $k \in \{0, 1\}$, we assume that $Y^i \sim f^i_k$; for P^i_* we assume that $Y^i_{t+1} \sim q^i(\cdot \mid \pi^i_t)$, which is given by (2). The discretization procedure is shown in Algorithm 1, which corresponds to the first-order hold method in [13].

Note that the transition probabilities P^i_k, $k \in \{0, 1\}$, approximate the evolution of the $\{\pi^i_t\}_{t=1}^{\infty}$ process when hypothesis $H = h_k$ is true. We will use these to approximate probabilities ξ^i_k. On the other hand, the transition probability P^i_* approximates the uncontrolled evolution of $\{\pi^i_t\}_{t=1}^{\infty}$. This will be used to approximately solve the dynamic program of Proposition 3.

Algorithm 1: Compute transition matrices

input : Discretization size m, DM i
output: P^i_0, P^i_1, P^i_*
forall $s_p \in \mathcal{S}_m$ **do**
  forall $y \in \mathcal{Y}^i$ **do**
    let $s_+ = \phi^i(s, y^i)$
    find $s_q, s_{q+1} \in \mathcal{S}_m$ such that $s_+ \in [s_q, s_{q+1})$
    find $\lambda^y_q, \lambda^y_{q+1} \in [0, 1]$ such that
      • $\lambda^y_q + \lambda^y_{q+1} = 1$
      • $s_+ = \lambda^y_q s_q + \lambda^y_{q+1} s_{q+1}$
  forall $q \in \{0, 1, \dots, m\}$ **do**
    $[P^i_0]_{pq} = \sum_y \lambda^y_q \cdot f^i_0(y) \cdot s_p$
    $[P^i_1]_{pq} = \sum_y \lambda^y_q \cdot f^i_1(y) \cdot (1 - s_p)$
    $[P^i_*]_{pq} = \sum_y \lambda^y_q \cdot q^i(y^i \mid s_p)$

B. Approximately computing ξ^i_k

Fix a decision maker i, $i \in \{1, 2\}$, and the threshold $H = h_k$, $k \in \{0, 1\}$. 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:

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

Note that $\xi^i_k(h_0, g^i; p)$ corresponds to the event that the Markov process $\{\pi^i_t\}_{t=1}^{\infty}$ 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^i_k that starts in p (which is assumed to belong to $\mathcal{S}_m$) gets absorbed in the set $\mathcal{A}^i_0$ before it is absorbed in the set $\mathcal{A}^i_1$. A similar interpretation holds for $\xi^i_k(h_1, g^i; p)$.

Thus, to approximate ξ^i_k, we can consider the Markov chain with transition probability P^i_k and two absorption sets: $\mathcal{A}^i_0$ and $\mathcal{A}^i_1$. Let $\hat{P}^i_k$ be the transition matrix of the corresponding absorbing Markov chain. Re-order states so that $\hat{P}^i_k$ may be written in the canonical form

$$\hat{P}^i_k = \begin{pmatrix} Q^i_k & R^i_k \\ 0 & I \end{pmatrix}.$$

Define $B^i_k = (1 - Q^i_k)^{-1} R^i_k$. From standard results in Markov chain analysis, we know that for any $s \in \mathcal{S}_m \setminus (\mathcal{A}^i_0 \cup \mathcal{A}^i_1)$ and $b \in \{0, 1\}$, $[B^i_k]_{sb}$ is the probability that the Markov chain starting in state s is absorbed in the set $\mathcal{A}^i_b$. Thus,

$$\xi^i_k(h_b, \langle \alpha^i, \beta^i \rangle; p) \approx [B^i_k]_{p^* b}, \quad b \in \{0, 1\}, \tag{12}$$

where p^* denotes the index of p in $\mathcal{S}_m \setminus (\mathcal{A}^i_0 \cup \mathcal{A}^i_1)$.

C. Approximate solution of the dynamic program

Using the procedure of the previous section, we can approximate $\xi^i_k(\cdot, g^{-i}; \pi^i)$, and therefore approximately compute $W^i_k(\pi^i, g^{-i})$, for any $\pi^i \in \mathcal{S}_m$ and any threshold-based strategy $g^{-i} = \langle \alpha^{-i}, \beta^{-i} \rangle$. To approximately solve the dynamic program of Proposition 3, we also need to approximate the Bellman operator $\mathscr{B}^i$. Define an approximate Bellman

operator using the first-order hold transition matrix P^i_* as follows:

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

Then $\hat{\mathscr{B}}^i$ corresponds to the discretization of $\mathscr{B}^i$ on $\mathcal{S}_m$ and performing linear interpolation on points outside $\mathcal{S}_m$ (see [14]). Hence, it may be used to approximately compute $W_{\mathsf{C}}(\pi^i, g^{-i})$.

Combing all these, we get an approximate procedure to solve the dynamic program of Proposition 3. This, in turn, gives an approximate procedure for finding a PBPO strategy using orthogonal search.

IV. Method II: Direct search

In this section, we describe a second approach to compute team optimal threshold-based strategies. The main idea of the approach is to approximately compute the performance of a generic threshold-based strategy $(\langle\alpha^1, \beta^1\rangle, \langle\alpha^2, \beta^2\rangle)$, and then optimize over the thresholds. For this reason, we call this approach *direct search*.

A. Performance of an arbitrary strategy

Given an arbitrary 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 we index θ^i_k by the *a priori* probability p. Then the total cost (1) is given by

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

For an arbitrary strategy, it is difficult to compute θ^i_k and ξ^i_k. However, based on the approximation presented in Section III-B, when (g^1, g^2) are threshold-based strategies, then ξ^i_k may be approximated by the absorption probabilities of appropriate sets for a Markov chain with transition matrix P^i_k. The same idea can be used to approximate $\theta^i_k(g^i; p)$, as explained below.

Given any threshold-based strategy $g^i = \langle\alpha^i, \beta^i\rangle$, define sets $\mathcal{A}^i_0$ and $\mathcal{A}^i_1$, and the matrix $\hat{P}^i_k$ as in Section III-B. Define $T^i_k = (I - Q^i_k)^{-1}\mathbf{1}$, where $\mathbf{1}$ is a column vector with all entries as 1. From standard results in Markov chain analysis, we know that for any $s \in \mathcal{S}_m \setminus (\mathcal{A}^i_0 \cup \mathcal{A}^i_1)$, $[T^i_k]_s$ is the expected stopping time that the Markov chain starting in state s is absorbed in $(\mathcal{A}^i_b \cup \mathcal{A}^i_1)$. Thus,

$$\theta^i_k(\langle\alpha^i, \beta^i\rangle; p) \approx [T^i_k]_{p^*}, \tag{14}$$

where p^* denotes the index of p in $\mathcal{S}_m \setminus (\mathcal{A}^i_0 \cup \mathcal{A}^i_1)$.

By substituting the approximate values of θ^i_k from (14) and ξ^i_k from (12) in (13), we can approximately compute $J(\langle\alpha^1, \beta^1\rangle, \langle\alpha^2, \beta^2\rangle; p)$ when $p, \alpha^1, \beta^1, \alpha^2, \beta^2 \in \mathcal{S}_m$.

B. Approximate search over all threshold-based strategies

Although there is no analytic expression for $J(\langle\alpha^1, \beta^1\rangle, \langle\alpha^2, \beta^2\rangle; p)$, it can be numerically approximated using the method proposed above. 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 [7].

In principle, such non-convex optimization problems can be solved using *derivative-free* methods that do no use numerical or analytic gradients (see [15]). In the numerical results we use one of the simplest derivative-free algorithms—Nelder-Mead simplex algorithm (see [16]). This step can be replaced by more sophisticated algorithms to obtain better results. However, it is not possible to guarantee that such algorithms will converge to team optimal solution. Thus, in practice, the direct search algorithm converges to a locally optimal solution.

To reduce the dependence of the numerical results on the choice of the *a priori* probability p, we pick multiple values of p in a finite set $\mathcal{P} \subset [0, 1]$ and use

$$\hat{J}(\langle\alpha^1, \beta^1\rangle, \langle\alpha^2, \beta^2\rangle) = \frac{1}{|\mathcal{P}|} \sum_{p \in \mathcal{P}} J(\langle\alpha^1, \beta^1\rangle, \langle\alpha^2, \beta^2\rangle; p)$$

as the objective function for the non-convex optimization algorithm. If $J(\langle\alpha^1, \beta^1\rangle, \langle\alpha^2, \beta^2\rangle, p)$ was computed exactly, then such an averaging will not affect the result of the optimization algorithm because the optimal strategy (g^1, g^2) does not depend on the choice of p.

V. Numerical comparison of the two methods

Both the approaches presented in this paper only guarantee local 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

$$\ell(u^1, u^2, h) = \begin{cases} 0, & \text{if } u^1 = u^2 = h, \\ L_1, & \text{if } u^1 \neq u^2, \\ L_2, & \text{if } u^1 = u^2 \neq h. \end{cases} \tag{15}$$

For both methods, we use $m = 1000$ and in direct search, we use $\mathcal{P} = \mathcal{S}_m$.

Note that the choice of parameters (c^1, c^2, L_1, L_2) and observation distributions $(f^1_0, f^1_1, f^2_0, f^2_1)$ completely specifies the model. We first consider an example where we pick specific values for the parameters and the distributions. Then we compare the performance of the two methods when all the parameters are chosen at random.

A. A specific example

Let $c^1 = c^2 = 1$, $L_1 = 20$, $L_2 = 50$ and

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

The result of orthogonal search and direct search are shown in Table I. Both approaches converge to a locally optimal solution. For this particular example, direct search converges to a slightly better solution that orthogonal search.

Although orthogonal search converges in significantly fewer number of iterations, each iteration of orthogonal search involves solving an infinite horizon dynamic program (which was solved using value iteration with convergence threshold 10^{-3}). The running time of both the algorithms is reported, but it should be noted that we did not attempt to optimize the Matlab implementation of the algorithms.

TABLE I

COMPARISON OF ORTHOGONAL SEARCH AND DIRECT SEARCH FOR THE SPECIFIC PARAMETERS PRESENTED IN SECTION V-A.

	$g^1 = \langle \alpha^1, \beta^1 \rangle$	$g^2 = \langle \alpha^2, \beta^2 \rangle$	$\hat{J}(g^1, g^2)$	iters.	runtime
OS[1]	$\langle 0.326, 0.73 \rangle$	$\langle 0.07, 0.931 \rangle$	9.109	5	1.45s
DS[1]	$\langle 0.287, 0.726 \rangle$	$\langle 0.14, 0.863 \rangle$	8.719	45	6.05s

[1] OS stands for orthogonal search and DS stands for direct search.

B. General Performance for random instance

To compare the performance of the two methods, we test both of them over 500 randomly generated instances of the parameters (c^1, c^2, L_1, L_2) and $(f_0^1, f_1^1, f_0^2, f_1^2)$. Specifically, we use $c^1, c^2 \sim$ unif$[0, 1]$, $L_1 \sim$ unif$[2, 40]$, and $L_2 \sim$ unif$[40, 80]$. We pick f_k^i by picking a random number $\delta_k^i \sim$ unif$[0, 1]$ and setting $f_k^i = [\delta_k^i, 1 - \delta_k^i]$.

Let J_1 denote the performance of the strategy obtained using orthogonal search and J_2 denote the performance of the strategy obtained using direct search. To compare the performance of the two methods, we pick $\epsilon = 10^{-2}$ and check the number of instances for which $J_1 > J_2 + \epsilon$ or $J_2 > J_1 + \epsilon$. When $J_a > J_b + \epsilon$, we define the maximum percentage improvement as the maximum of $(J_a - J_b)/J_a$ over all random instances and the average percentage improvement as the average of $(J_a - J_b)/J_a$ over all random instances. These results are presented in Table II.

TABLE II

COMPARISON OF ORTHOGONAL AND DIRECT SEARCH OVER RANDOMLY GENERATED INSTANCES

	$J_1 > J_2 + \epsilon$	$J_2 > J_1 + \epsilon$
Number of cases	475	5
Max improvement	50.9%	6.75%
Avg improvement	4.16%	2.03%

From the result, we see that in over 90% of the 500 cases, direct search does better than orthogonal search. There are instances where direct search does significantly better.

VI. DISCUSSION

The proposed search algorithms consist of two parts: computing the transition matrix (Algorithm 1) and using the transition matrix to compute the thresholds. The complexity of the first part is linear in $|\mathcal{Y}|$, but the complexity of the second part 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 $|\mathcal{Y}|$.

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 globally optimal strategies. Direct search involves solving a non-convex optimization problem, so in practice, it will also converge to a local optimal. In our numerical investigation of the two algorithms, direct search performs better than orthogonal search; sometimes, significantly better.

A future direction is to generalize the approximation methods developed in this paper to more general decentralized sequential hypothesis models of [4]–[6].

REFERENCES

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

[2] 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.

[3] 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.

[4] V. V. Veeravalli, T. Başar, and H. V. Poor, "Decentralized sequential detection with a fusion center performing the sequential test," vol. 39, no. 2, pp. 433–442, 1993.

[5] A. Nayyar and D. Teneketzis, "Sequential problems in decentralized detection with communication," vol. 57, no. 8, pp. 5410–5435, Aug. 2011.

[6] ——, "Signaling in sensor networks for sequential detection," *IEEE Trans. Control of Network Systems*, vol. 2, no. 1, pp. 36–46, Mar. 2015.

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

[8] H. Chernoff, *Sequential Analysis and Optimal Design*. SIAM, 1972.

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

[10] A. Tartakovsky, I. Nikiforov, and M. Basseville, *Sequential analysis: Hypothesis testing and changepoint detection*. CRC Press, 2014.

[11] A. Y. Ng and M. I. Jordan, "PEGASUS: A policy search method for large MDPs and POMDPs," in *Proceedings of the 16th Conference on Uncertainty in Artificial Intelligence*, ser. UAI '00, 2000, pp. 406–415.

[12] M. T. J. Spaan and N. Vlassis, "Perseus: Randomized point-based value iteration for POMDPs," *Journal of Artificial Intelligence Research*, vol. 24, pp. 195–220, 2005.

[13] A. Pieter, "Discretization of continuous state space mdps," http://www.cs.berkeley.edu/ pabbeel/cs287-fa12/slides/discretization.pdf.

[14] W. S. Lovejoy, "Computationally feasible bounds for partially observed Markov decision processes," *Operations Research*, vol. 39, no. 1, pp. 162–175, 1991.

[15] 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.

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