Les Cahiers du GERAD ISSN: 0711–2440

On computing optimal thresholds in decentralized sequential hypothesis testing

C. Cui
A. Mahajan

G–2015–67

July 2015

GERAD HEC Montréal
3000, chemin de la Côte-Sainte-Catherine
Montréal (Québec) Canada H3T 2A7

Tél. : 514 340-6053
Téléc. : 514 340-5665
info@gerad.ca
www.gerad.ca

On computing optimal thresholds in decentralized sequential hypothesis testing

Can Cui
Aditya Mahajan

GERAD & Department of Electrical and Computer Engineering, McGill University, Montréal (Québec) Canada, H3A 0E9

can.cui@mail.mcgill.ca
aditya.mahajan@mcgill.ca

July 2015

Les Cahiers du GERAD
G–2015–67

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.

Key Words: Hypothesis testing, sequential analysis, decentralized systems, dynamic programming, derivative-free optimization.

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

1 Introduction

Sequential hypothesis testing is a classical problem in sequential analysis with applications ranging from clinical trials, quality control, and sensor networks. Broadly speaking, research on sequential hypothesis testing may be classified into three categories: centralized, distributed, and decentralized.

In centralized sequential hypothesis testing, a single decision maker or sensor observes a random process (often assumed to i.i.d.) 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 likelihood ratio test is optimal. See [2–4] for more details.

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 communication decision is made in decentralized manner. Different configurations on *local memory* of distributed sensors and *feedback* from the fusion center are considered. See [5–8] for details and [9, 10] 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 [11–14] for details.

It should be noted that the terms distributed and decentralized are used interchangeably in the literature. We are 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 [11]. 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. It was shown that optimal decision rules are threshold based, where the thresholds of the two decision makers are coupled.

The main contribution of this paper is to present two search procedures that approximately compute the optimal thresholds. In the first search procedure, which we call *orthogonal search* and present in Section 3, we iteratively solve two coupled dynamic programs to identify optimal thresholds. In the second search procedure, which we call *direct search* and present in Section 4, 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.

In Section 5, we present detailed numerical experiments to compare the performance of the two procedures. In most experiments, direct search performs better than orthogonal search. 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.

2 Problem formulation and structure of optimal strategies

2.1 The model

Consider a decentralized sequential hypothesis problem investigated in [11]. 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 $Y_{1:t}^i \coloneqq (Y_1^i \cdots Y_t^i)$.

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.,

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

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:

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

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 he other does not, the loss is 1; and if both users make incorrect stopping decisions, then the loss is L. We will use this loss function in the numerical experiments in Section 5.

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

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

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 [15]. 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.

2.2 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 $y_{1:t}^i$ of $Y_{1:t}^i$, define

$$\pi_t^i := \mathbb{P}(H = h_0 \mid y_{1:t}^i).$$

In addition, define

$$q^i(y_{t+1}^i \mid \pi_t^i) := \pi_t^i f_0^i(y_{t+1}^i) + (1 - \pi_t^i) f_1^i(y_{t+1}^i), \tag{2}$$

$$\phi^i(\pi_t^i, y_{t+1}^i) := \pi_t^i f_0^i(y_{t+1}^i) / q^i(y_{t+1}^i \mid \pi_t^i). \tag{3}$$

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

$$\pi_{t+1}^i = \phi^i(\pi_t^i, y_{t+1}^i). \tag{4}$$

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

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

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 [11] that $\{\pi_t^i\}_{t=1}^{\infty}$ is an information state process for DM^i. In particular:

Lemma 1 ([11]) *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*

$$U_t^i = g_t^i(\pi_t^i). \tag{6}$$

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 $\alpha_t^i, \beta_t^i \in [0, 1]$, $\alpha_t^i \leq \beta_t^i$, such that for any $\pi^i \in [0, 1]$,*

$$g_t^i(\pi^i) = \begin{cases} h_1 & \text{if } \pi^i < \alpha_t^i, \\ \mathsf{C} & \text{if } \alpha_t^i \leq \pi^i \leq \beta_t^i, \\ h_0 & \text{if } \pi^i < \beta_t^i. \end{cases}$$

It was shown in [11] that threshold-based strategies are team optimal. In particular:

Lemma 2 ([11, 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

$$\hat{\ell}(u^i, h_k) := \mathbb{E}[\ell(u^i, U^{-i}, h_k)] = \xi_k^{-i}(h_0, g^{-i}; p)\ell(u^i, h_0, h_k) + \xi_k^{-i}(h_1, g^{-i}; p)\ell(u^i, h_1, h_k).$$

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 $g^i = (g_1^i, g_2^i, \dots)$ is called time invariant if for any $\pi^i \in [0,1]$, $g_t^i(\pi^i)$ does not depend on t.*

For infinite-horizon problems, for a single decision maker, time-invariant strategies are optimal. However, that is not always the case for multiple decision makers. It was shown in [11] that threshold-based time-invariant strategies are PBPO.

Theorem 1 ([11, 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 $\pi^i \in [0,1]$*

$$V^i(\pi^i) = \min\{W_0^i(\pi^i, g^{-i}), W_1^i(\pi^i, g^{-i}), W_{\mathsf{C}}^i(\pi^i, g^{-i})\}, \tag{7}$$

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

$$W_k^1(\pi^1, g^2) = \sum_{u^2 \in \{h_0, h_1\}} \big[\xi_0^2(u^2, g^2; \pi^1) \cdot \pi^1 \cdot \ell(h_k, u^2, h_0) + \xi_1^2(u^2, g^2; \pi^1) \cdot (1-\pi^1) \cdot \ell(h_k, u^2, h_1)\big], \tag{8}$$

W_k^2 is defined similarly, and

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

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 1 gives two coupled dynamic programs, which we write succinctly as

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

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 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 seem 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.

2.3 Computation of PBPO strategies

Assuming that we restrict attention to time-invariant strategies, we propose two methods to compute the optimal thresholds. In the first method, which we call *orthogonal search*, we iteratively solve the coupled dynamic programs (10). A similar approach was used in [11] under the assumption that that f_k^i 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 [1,16]. This method can be extended to the non-Gaussian distributions using the expressions in [17], 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_k^i also needs to be computed to solve these dynamic programs. ξ_k^i 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 [16]); another option is to solve an appropriate Fredholm Integral equation (see [18]). To be consistent with the approximation that we use to solve the dynamic program, we compute ξ_t^i 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 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_t^i, 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 detailed numerical comparison of the two methods in Section 5. In most of our experiments, direct search outperforms orthogonal search; sometimes significantly.

3 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 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.*

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 in [19, 20].

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

3.1 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 [21].

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 $\{\pi^i_t\}_{t=1}^{\infty}$. We will use this to approximately solve the dynamic program of Lemma 1.

3.2 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 $\xi^i_k(h_1, g^i; p)$.

Algorithm 1: Compute transition matrices

input : Discretization size m, DM i
output: P_0^i, P_1^i, 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 $\lambda_q^y, \lambda_{q+1}^y \in [0,1]$ such that
      • $\lambda_q^y + \lambda_{q+1}^y = 1$
      • $s_+ = \lambda_q^y s_q + \lambda_{q+1}^y s_{q+1}$
  forall $q \in \{0, 1, \ldots, m\}$ **do**
    $[P_0^i]_{pq} = \sum_y \lambda_q^y \cdot f_0^i(y) \cdot s_p$
    $[P_1^i]_{pq} = \sum_y \lambda_q^y \cdot f_1^i(y) \cdot (1 - s_p)$
    $[P_*^i]_{pq} = \sum_y \lambda_q^y \cdot q^i(y^i \mid s_p)$

Thus, to approximate ξ_k^i, we can consider the Markov chain with transition probability P_k^i and absorption sets $\mathcal{A}_0^i$ and $\mathcal{A}_1^i$. Let $\hat{P}_k^i$ be the transition matrix of the corresponding absorbing Markov chain. Re-order states so that $\hat{P}_k^i$ may be written in the canonical form

$$\hat{P}_k^i = \begin{pmatrix} Q_k^i & R_k^i \\ 0 & I \end{pmatrix}.$$

Define $B_k^i = (1 - Q_k^i)^{-1} R_k^i$. From standard results in Markov chains, we know that for any $s \in \mathcal{S}_m \setminus (\mathcal{A}_0^i \cup \mathcal{A}_1^i)$ and $b \in \{0,1\}$, $[B_k^i]_{sb}$ is the probability that the Markov chain starting in state s is absorbed in the set $\mathcal{A}_b^i$. Thus,

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

where p^* denotes the index of p in $\mathcal{S}_m \setminus (\mathcal{A}_0^i \cup \mathcal{A}_1^i)$.

3.3 Approximate solution of the dynamic program

Using the procedure of the previous section, we can approximate $\xi_k^i(\cdot, g^{-i}; \pi^i)$, and therefore approximately compute $W_k^i(\pi^i, g^{-i})$, 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 Lemma 1, 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 [22]). 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 Lemma 1. This, in turn, gives an approximate procedure for finding a PBPO strategy using orthogonal search.

4 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*.

4.1 Performance of an arbitrary strategy

Given an arbitrary strategy (g^1, g^2) and $i \in \{1,2\}$, $k \in \{0,1\}$, define:

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

Note that we index θ_k^i by the *a priori* probability p. Then the total cost (1) is given by

$$\begin{aligned} J(g^1, g^2; p) = p \cdot [c^1 \cdot \theta_0^1(g^1; p) + c^2 \cdot \theta_0^2(g^2; p)] + (1-p) \cdot [c^1 \cdot \theta_1^1(g^1; p) + c^2 \cdot \theta_1^2(g^2; p) \\ + \sum_{u^1, u^2 \in \{h_0, h_1\}^2} \big[p \cdot \xi_0^1(u^1, g^1; p) \cdot \xi_0^2(u^2, g^2; p) \cdot \ell(u^1, u^2, h_0) \\ + (1-p) \cdot \xi_1^1(u^1, g^1; p) \cdot \xi_1^2(u^2, g^2; p) \cdot \ell(u^1, u^2, h_1)\big]. \end{aligned} \tag{13}$$

For an arbitrary strategy, it is difficult to compute θ_k^i and ξ_k^i. However, based on the approximation presented in Section 3.2, when (g^1, g^2) are threshold-based strategies, then ξ_k^i may be approximated by the absorption probabilities of appropriate sets for a Markov chain with transition matrix P_k^i. The same idea can be used to approximate $\theta_k^i(g^i; p)$, as explained below.

Given any threshold based strategy $g^i = \langle \alpha^i, \beta^i \rangle$, define sets $\mathcal{A}_0^i$ and $\mathcal{A}_1^i$, and the matrix $\hat{P}_k^i$ as in Section 3.2. Define $T_k^i = (I - Q_k^i)^{-1}\mathbf{1}$, where $\mathbf{1}$ is a column vector with all entries as 1. From standard results in Markov chain analysis, we know that for any $s \in \mathcal{S}_m \setminus (\mathcal{A}_0^i \cup \mathcal{A}_1^i)$, $[T_k^i]_s$ is the expected stopping time that the Markov chain starting in state s is absorbed in $(\mathcal{A}_b^i \cup \mathcal{A}_1^i)$. Thus,

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

where p^* denotes the index of p in $\mathcal{S}_m \setminus (\mathcal{A}_0^i \cup \mathcal{A}_1^i)$.

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

4.2 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 [15].

In principle, such non-convex optimization problems can be solved using *derivative-free* methods that do no use numerical or analytic gradients (see [23]). In the numerical results we use one of the simplest derivative-free algorithms—Nelder-Mead simplex algorithm (see [24]). 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.

5 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

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

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) and observation distributions $(f_0^1, f_1^1, f_0^2, f_1^2)$ 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.

5.1 Coupled loss case

Let $c^1 = c^2 = 0.05$, $L = 2.5$ and

$$f_0^1 = \begin{bmatrix} 0.25 & 0.75 \end{bmatrix}, \qquad f_0^2 = \begin{bmatrix} 0.80 & 0.20 \end{bmatrix},$$
$$f_1^1 = \begin{bmatrix} 0.60 & 0.40 \end{bmatrix}, \qquad f_1^2 = \begin{bmatrix} 0.30 & 0.70 \end{bmatrix}.$$

The result of orthogonal search and direct search are shown in Table 1. 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 implementations of the algorithms.

Table 1: Comparison of orthogonal search and direct search for the specific parameters presented in Section 5.1.

	$g^1 = \langle \alpha^1, \beta^1 \rangle$	$g^2 = \langle \alpha^2, \beta^2 \rangle$	$\hat{J}(g^1, g^2)$	iters.	runtime
OS[1]	$\langle 0.326, 0.73 \rangle$	$\langle 0.07, 0.931 \rangle$	0.455	5	1.45s
DS[1]	$\langle 0.287, 0.726 \rangle$	$\langle 0.14, 0.863 \rangle$	0.436	45	6.05s

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

5.2 Decomposable case

In general, the team optimal solution is not known and both orthogonal search and direct search converge to a local optimal solution. In the special case when $L = 2$, the total stopping cost $\ell(U^1, U^2, H)$ is not coupled because it can 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. In this case, the solution $g^i = \langle \alpha^i, \beta^i \rangle$ can be obtained by separately solving the two centralized sequential hypothesis testing problems. We can use value iteration to find the optimal threshold-based policy for two centralized sequential hypothesis testing problems and compare the result with orthogonal search and direct search. We refer to this solution as *centralized solution*.

Let $L = 2$ and set $c^1, c^2, f_0^1, f_1^1, f_0^2, f_1^2$ to be the same as the values in Section 5.1. The results of orthogonal search, direct search, and the centralized solution is shown in Table 2. For this particular example, direct search gives the same solution as the centralized solution while orthogonal search converges to a solution with poorer performance.

Table 2: Comparison of decomposable orthogonal search and global search with centralized solution.

	$g^1 = \langle \alpha^1, \beta^1 \rangle$	$g^2 = \langle \alpha^2, \beta^2 \rangle$	$\hat{J}(g^1, g^2)$
OS	$\langle 0.318, 0.686 \rangle$	$\langle 0.089, 0.913 \rangle$	0.428
DS	$\langle 0.3053, 0.7055 \rangle$	$\langle 0.1845, 0.8218 \rangle$	0.406
CS[1]	$\langle 0.305, 0.705 \rangle$	$\langle 0.184, 0.822 \rangle$	0.406

[1] CS stands for centralized solution.

5.3 General performance for coupled loss cases

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) and $(f_0^1, f_1^1, f_0^2, f_1^2)$. Specifically, we use $c^1, c^2 \sim \text{unif}[0, 0.05]$, $L \sim \text{unif}[1, 4]$. We pick f_k^i by picking a random number $\delta_k^i \sim \text{unif}[0, 1]$ and setting $f_k^i = [\delta_k^i, 1 - \delta_k^i]$.

Let J_{OS} and J_{DS} denote the performance of the solution obtained by orthogonal search and direct search. Define $\Delta J_{OS} = (J_{OS} - J_{DS})/J_{OS}$ and $\Delta J_{DS} = (J_{DS} - J_{OS})/J_{DS}$ as the relative difference between the performance of orthogonal search and direct search. The histograms of ΔJ_{OS} and ΔJ_{DS} are shown in Figure 1. Note that for 488 of the 500 cases, $J_{OS} > J_{DS} + 10^{-4}$; therefore, for most scenarios, direct search performs better than orthogonal search.

5.4 General performance for decomposable cases

It is also interesting to look at the general performance of two approaches when the team optimal solution is known. In this experiment, we set $L = 2, c^1 = c^2 = 0.05$ and pick $f_0^1, f_1^1, f_0^2, f_1^2$ randomly as specified in Section 5.3. Let J_{OS}, J_{DS} denote the performance of the solution obtained by orthogonal search and direct search. Let J^* denote the centralized solution. Define the relative errors $E_{OS} = (J_{OS} - J^*)/J^*$ and $E_{DS} = (J_{DS} - J^*)/J^*$. The histograms of E_{OS} and E_{DS} are shown in Figure 2. From the plot, we can see that the errors of both approaches are within tolerable range and that the performance of direct search is better than the performance of orthogonal search in most cases.

The nature of the solution remains the same when $c^i \ll L$. For example, for $L = 2, c^1 = c^2 = 5 \times 10^{-4}$, and f_h^i chosen randomly as in Section 5.3, the histograms of E_{OS} and E_{DS} are shown in Figure 3.

5.5 Computational complexity

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

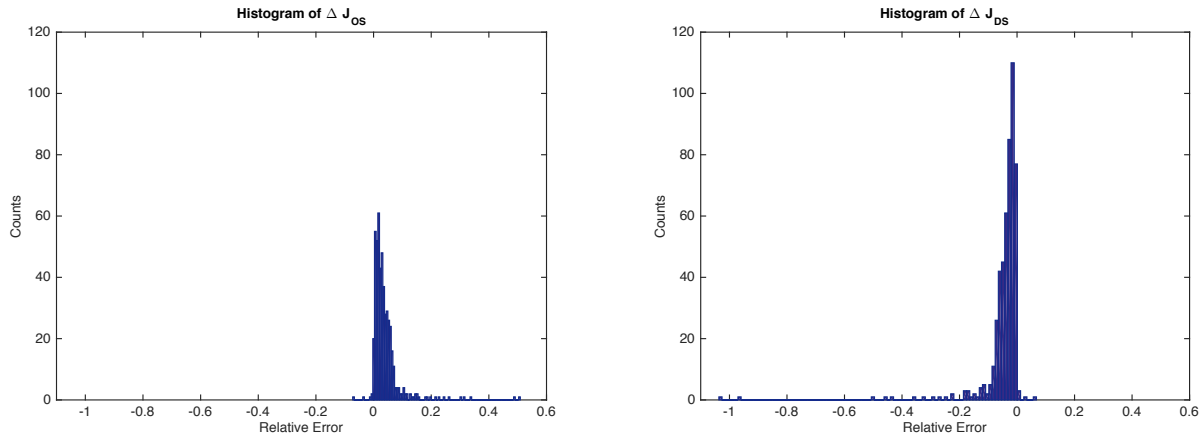

Figure 1: Histograms of ΔJ_{OS} and ΔJ_{DS}.

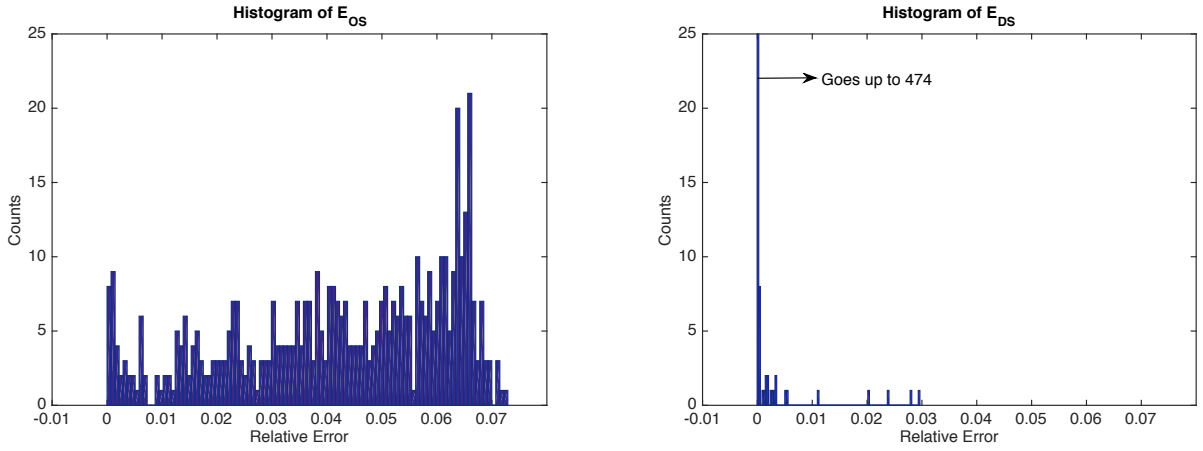

Figure 2: Histograms of E_{OS} and E_{DS}.

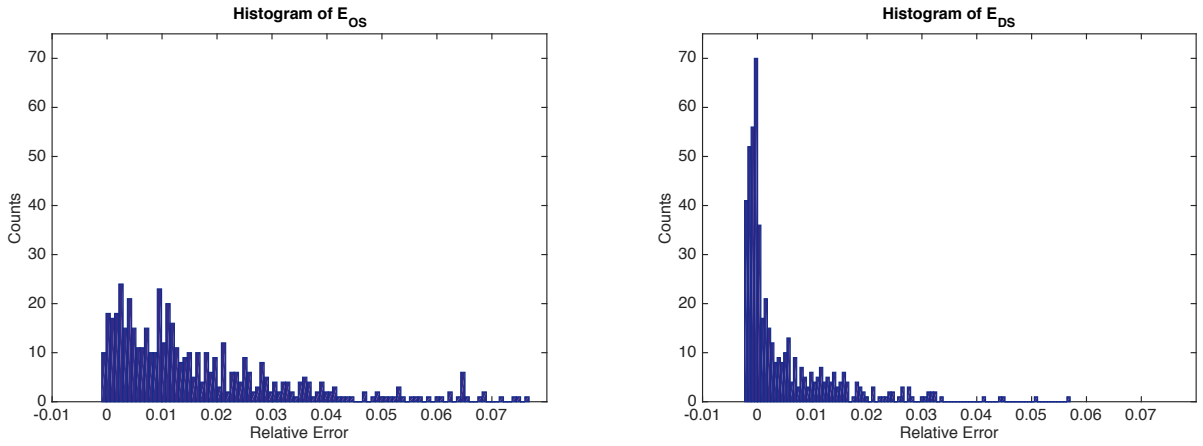

Figure 3: Histograms of E_{OS} and E_{DS} when $c^i \ll L$.

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 $|\mathcal{Y}|$. 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 Section 5.3. To choose f_h^i, for $|\mathcal{Y}| = m$, we pick m random numbers $(\delta_{k0}^i, \ldots, \delta_{km}^i) \sim \text{unif}[0,1]$ and set $f_h^i = [\delta_{k0}^i, \ldots, \delta_{km}^i]/S_k^i$, where $S_k^i = \sum_{j=1}^m \delta_{kj}^i$. The average run-time for orthogonal search and direct search is shown in Table 3. As expected, the run-time does not increase accordingly with increasing $|\mathcal{Y}|$, it remains at a steady state. This suggests that the proposed approaches should work well even when $\mathcal{Y}$ is continuous valued.

Table 3: Running time 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 average	5.497s	3.952s	6.692s	5.266s
DS average	3.560s	2.558s	2.380s	2.574s

6 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 converge to a local optimal. In our numerical investigation of the two algorithms, direct search performs better than orthogonal search; sometimes, significantly better.

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 [14, 25].

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. Hashemi and I.B. Rhodes, Decentralized sequential detection, Information Theory, IEEE Transactions on, 35(3), 509–520, 1989.

[6] V.V. Veeravalli, Comments on "decentralized sequential detection", IEEE Transactions on Signal Processing, 38(4), 1428–1429, 1992.

[7] 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, 39(2), 433–442, 1993.

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

[9] Y. Mei, Asymptotic optimality theory for decentralized sequential hypothesis testing in sensor networks, IEEE Transactions on Information Theory, 54(5), 2072–2089, May 2008.

[10] G. Fellouris and G.V. Moustakides, Decentralized sequential hypothesis testing using asynchronous communication, IEEE Transactions on Information Theory, 57(1), 534–548, Jan. 2011.

[11] D. Teneketzis and Y. Ho, The decentralized Wald problem, Information and Computation (formerly Information and Control), 73(1), 23–44, Apr. 1987.

[12] 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, 83, 533–543.

[13] V.V. Veeravalli, T. Başar, and H.V. Poor, Decentralized sequential detection with sensors performing the sequential test, Mathematics of Control, Signals and Systems, 7(4), 292–305, 1994.

[14] A. Nayyar and D. Teneketzis, Sequential problems in decentralized detection with communication, IEEE Transactions on Information Theory, 57(8), 5410–5435, Aug. 2011.

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

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

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

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

[19] 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, 406–415.

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

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

[22] W.S. Lovejoy, Computationally feasible bounds for partially observed Markov decision processes, Operations Research, 39(1), 162–175, 1991.

[23] L.M. Rios and N.V. Sahinidis, Derivative-free optimization: a review of algorithms and comparison of software implementations, Journal of Global Optimization, 56(3), 1247–1293, 2013.

[24] J.A. Nelder and R. Mead, A simplex method for function minimization, The Computer Journal, 7(4), 308–313, 1965.

[25] A. Nayyar and D. Teneketzis, Signaling in sensor networks for sequential detection, IEEE Trans. Control of Network Systems, 2(1), 36–46, Mar. 2015.