# Always Choose Second Best: Tracking a Moving Target on a Graph with a Noisy Binary Sensor

Kevin Leahy and Mac Schwager

*Abstract*— In this work, we consider the problem of using a noisy binary sensor to optimally track a target that moves as a Markov Chain in a finite discrete environment. Our approach focuses on one-step optimality because of the apparent infeasibility of computing an optimal policy via dynamic programming. We prove that always searching in the second most likely location minimizes one-step variance while maximizing the belief about the target's location over one step. Simulation results demonstrate the performance of our strategy and suggest the policy performs well over arbitrary horizons.

## I. INTRODUCTION

This work focuses on finding a search strategy that provides the most confident estimate of the location of a target that evolves as a stochastic process in a graph environment. This work is inspired by [1], which proved that the strategy of searching for a stationary target in the first or second most likely locations is optimal in the sense of maximizing the most likely location. Our counterintuitive result shows that under similar assumptions, always searching the second most likely location is the optimal one-step policy in expectation, and that such a policy minimizes the variance of the estimate of the location of a moving target.

As a motivating example, consider the search for an endangered animal that has previously been tagged with a radio tracking device. The animal moves about its environment, and you know the areas it frequents (e.g. shady areas and waterholes), and how often it tends to move from area to area. You pilot a UAV, attempting to locate the animal via its tag, but given the animal's propensity for moving about and radio interference, it is difficult to be certain about where the animal is located. With all of this knowledge, where in the environment should you check to maximize your confidence about the animal's location?

Problems similar to the one we examine arise in a variety of domains, whenever sensing resources are limited and the motion of a phenomenon to be sensed is probabilistic. Such target tracking problems are of considerable interest to the robotics community, and have been well-studied recently, including gathering information about the environment, locating survivors at sea, general pursuit-evasion and search problems [2], and even object recognition and detection [3]. There is also a strong interest for the purpose of target localization with radar [4] or sensor network management [5]. Even anomaly detection in cyberphysical systems has been approached as a problem of this form [6]. For many of these applications, sensing resources are limited, and the computational complexity of optimal policies presents a significant barrier to practical implementation of such solutions. Our method is simple and efficient to implement, and provides a high degree of confidence in such scenarios.

We aim to track a target that is moving among a finite set of states as a Markov Chain. At each time instant, we may search one state for the target. It is known that searching either of the most likely locations for the target is optimal in expectation [1]. However, extending these results to a moving target to find an optimal solution appears infeasible. By examining a one-step optimization, we find a strategy for minimizing the variance of our estimate of the target's location by simply looking in the second most likely location. This policy results in a consistently good estimate of the target location. We demonstrate our strategy via simulation, which suggests it performs well over arbitrary horizons.

This problem falls generally within the class of problems known as $M$-ary Hypothesis Testing, in which a series of observations is used to decide which of $M$ possible statistical situations give rise to those observations. These problems were first studied in the binary ($M = 2$) case in [7]. $M$-ary Hypothesis Testing problems originated in the study of stationary targets, and therefore a decision could be made about the true hypothesis upon receipt of an entire set of observations. On the other hand, the presence of a moving target means our problem is a special case of $M$-ary Hypothesis Testing known as Sequential Hypothesis Testing. In the sequential problem, observations arrive one at a time, with an action performed after each observation to maximize the discriminatory power of the next observation. Early work for detecting moving targets in such sequential problems was performed in the Operations Research community, for both Markovian [8], [9] and non-Markovian targets [10], [11]. In these problem, over a finite horizon, the probability of detecting a target before the time horizon is maximized, using an exponential detection function. Our work focuses instead on infinite horizon search problems, in which we attempt to maximize localization information about a target using binary detection. Many of the results from this area focus on an optimal strategy consisting of a finite sequence of search locations determined *a priori*, based on a prior distribution over possible target locations, such as [12] and [13]. In our work, however, we seek to evaluate the performance of a reactive strategy that can be used efficiently online: always

K. Leahy is with the Department of Mechanical Engineering at Boston University, Boston MA 02215. Email: `kjleahy@bu.edu`

M. Schwager is with the Department of Aeronautics and Astronautics at Stanford University, Stanford CA 94305. Email: `schwager@stanford.edu`

search the second most likely location for the target.

It is the closed-loop structure of the policy we find, as well as the generality of the problem formulation that make it most closely related to [1], although that work considers a stationary target rather than a moving one. Other work has considered index-based policies in this problem domain [14], but not for closed-loop policies or a moving target. Our reactive strategy provides a simple, computationally efficient feedback policy for searching for a target that can be extended to an inifinite horizon, as opposed to the finite horizons considered in those works.

The organization of the rest of this paper is as follows. In Sec. II we introduce the models and assumptions in the problem. In Sec. II-C, we formally state the problem. In Secs. III and IV we present our approach to solving the problem, as well as simulation results. Finally, we discuss our results and conclusion in Sec. V.

## II. PROBLEM FORMULATION

### A. Target Model

For this problem, we consider a discrete environment consisting of a set of states $S$. Moving in this environment is a target whose location at time $k$ is denoted as $s^k$, where $s^k \in S$. The target evolves as a discrete time Markov Chain whose transition probabilities are known a priori. We let $P(s_1, s_2)$ denote the probability of the target moving from a state $s_1$ to a state $s_2$.

### B. Sensing Model

Our goal is to track the target as it moves about the environment. We do so by choosing a location $q^k \in S$ to take an observation $y^k$ at each time step $k$. For this work, we assume no restrictions on the motion of our sensor, although this is an interesting direction for future work. At each time instant $q^k$ may be chosen to be any $s \in S$, regardless of the value of $q^{k-1}$. The measurement model is binary, with measurements $y^k \in Y$, where $Y = \{0, 1\}$. The measurement probabilities are given by:

$$P\left(y^k = 1 \mid q^k, s\right) = \begin{cases} 1 - \beta & \text{if } q^k = s \\ \alpha & \text{otherwise} \end{cases}$$

$$P\left(y^k = 0 \mid q^k, s\right) = \begin{cases} \beta & \text{if } q^k = s \\ 1 - \alpha & \text{otherwise} \end{cases},$$

where $\beta$ represents the probability of missed detection and $\alpha$ represents the probability of false alarm. We assume that $\alpha, \beta < 0.5$. Further, we assume that $\alpha = \beta$. Such an assumption is reasonable in many cases, such as obtaining a binary sensor from continuous measurements with additive Gaussian noise (or other symmetric noise) by using a threshold.

The location of the target at time $0$ is unknown, but we assume a prior belief at time $0$, given as $b^0(s)$. This belief evolves over time, and we denote the belief about state $s$ at time $k$ as $b^k(s)$.

We update the belief after taking each measurement according to Bayes rule:

$$b^k(s) = \eta P\left(y^k \mid s, q^k\right) \sum_{s' \in S} P(s', s) b^{k-1}(s') , \quad (1)$$

where $\eta$ is the appropriate normalization factor and $P(s', s)$ is the probability of the target transitioning from state $s'$ to state $s$ as describe in Sec. II-A.

Note that (1) can be written in matrix form as

$$b^k = \eta P\left(Y^k = y | q^k, s\right) \Gamma b^{k-1}, \quad (2)$$

where $\Gamma$ is a column stochastic matrix, whose $i, j^{th}$ entry represents the probability of transitioning to state $i$ from state $j$, $b^k$ is a vector of length $|S|$ representing the belief from states 1 to $|S|$, and $P\left(Y^k = 1 | q^k, s\right)$ and $P\left(Y^k = 0 | q^k, s\right)$ are diagonal matrices of the form

$$P\left(Y^k = 1 | q^k, s\right) = \begin{bmatrix} \alpha & & & & & & \\ & \ddots & & & & \mathbf{0} & \\ & & \alpha & & & & \\ & & & 1-\beta & & & \\ & & & & \alpha & & \\ & \mathbf{0} & & & & \ddots & \\ & & & & & & \alpha \end{bmatrix} \quad (3)$$

$$P\left(Y^k = 0 | q^k, s\right) = \begin{bmatrix} 1-\alpha & & & & & \\ & \ddots & & & \mathbf{0} & \\ & & 1-\alpha & & & \\ & & & \beta & & \\ & & & & 1-\alpha & \\ & \mathbf{0} & & & & \ddots \\ & & & & & 1-\alpha \end{bmatrix}, \quad (4)$$

which we call *observation matrices*. Each observation matrix has a single entry that differs from the others, corresponding to the choice of $q^k$.

It is worth noting that our problem can be seen as choosing the observation matrices at each time step that results in a long product of matrices with the initial belief of the form

$$\eta P\left(Y^k | q^k, s\right) \Gamma P\left(Y^{k-1} | q^{k-1}, s\right) \Gamma \ldots P\left(Y^1 | q^1, s\right) \Gamma b^0 . \quad (5)$$

In general, these matrices do not commute. If $\Gamma$ is the identity matrix, the problem is reduced to the problem of locating a stationary target, which has been heavily studied in the literature. For example, from [1] it is known that when $\alpha = \beta$ (or when an analagous symmetry condition applies in the case of continuous-valued sensor measurements), an optimal policy is to choose either the first or second most likely location (it does not matter which) at each time.

### C. Problem Statement

We consider a multi-objective optimization, in which we want to both maximize the next step probability of locating the object, while simultaneously minimizing the next step variance on the belief that we have located the object. Formally, we have two objectives as follows.

**Objective 1:** First, we wish to maximize the maximum belief of the target's location at each time step. Formally, we seek a strategy to maximize the objective function

$$J_1\left(b^k\right) = \max_{s\in S} b^k\left(s\right) . \tag{6}$$

**Objective 2:** Among all policies that satisfy Objective 1, find that which minimizes the next step expected variance on the maximum belief,

$$J_2\left(b^k\right) = var\left(\max_{s\in S} b^k\left(s\right)\right) . \tag{7}$$

Although we have stated these objectives as finding only the next-step optimal policy, we have observed empirically that this policy performs well over arbitrary horizons, as demonstrated in our simulation results in Sec. IV.

## III. APPROACH

We wish to find the solution for choosing $q^k$ to maximize (6) by taking expectation of $b^k$ with respect to $y^k$, where $b^k$ is the vector of $b^k\left(s\right)$ for $s = \{1, \ldots, |S|\}$. This expectation is written

$$E_{y^k}\left[\max_{s\in S} b^k\left(s\right)\right] , \tag{8}$$

where we place no assumptions on the current belief $b^{k-1}$. We can expand (8) to become

$$\sum_{y\in Y} \max_{s\in S} b^k\left(s\right) P\left(y^k = y \mid y^{1:k-1}\right) . \tag{9}$$

Note that $P\left(y^k = y \mid y^{1:k-1}\right)$ represents the total probability of observing $y^k = y$, which is $1/\eta$. Therefore, we can combine (1) and (9) to get

$$\sum_{y\in Y} \max_{s\in S} P\left(y^k \mid s, q^k\right) \sum_{s'\in S} P\left(s', s\right) b^{k-1}\left(s'\right) . \tag{10}$$

which we would like to maximize over $q^k$.

Thus, our one-step optimization becomes

$$\max_{q^k}\left[\sum_{y\in Y} \max_{s\in S} P\left(y^k \mid s, q^k\right) \sum_{s'\in S} P\left(s', s\right) b^{k-1}\left(s'\right)\right] , \tag{11}$$

which we then expand to

$$\max_{q^k}\left[\left(\max_{s\in S} P\left(y^k = 1 \mid s, q^k\right) \sum_{s'\in S} P\left(s', s\right) b^{k-1}\left(s'\right)\right) + \left(\max_{s\in S} P\left(y^k = 0 \mid s, q^k\right) \sum_{s'\in S} P\left(s', s\right) b^{k-1}\left(s'\right)\right)\right]. \tag{12}$$

Each of the terms in (12) can be rewritten using matrices:

$$\max_{q^k}\left[\left(\max_{s\in S} P\left(y^k = 1\right) \Gamma b^{k-1}\right) + \left(\max_{s\in S} P\left(y^k = 0\right) \Gamma b^{k-1}\right)\right]. \tag{13}$$

We will write the term $\Gamma b^{k-1}$ from (2) as $\hat{b}^k$, representing the estimate at time $k$ before measuring (i.e. accounting for the motion of the target). The maximum entry in $\hat{b}^k$ will be denoted $\pi_1$ and the second highest entry in $\hat{b}^k$ will be denoted $\pi_2$. Without loss of generality, let "1" refer to the location of $\pi_1$ (the location with the highest likelihood of containing the target), and "2" refer to the location of $\pi_2$ (the location with the second highest likelihood of containing the target).

*Theorem 1:* Choosing either $q^k = 1$ or $q^k = 2$ maximizes (8), while any other choice gives a lower value.

*Proof:* This result follows directly from Proposition 8 in [1].

∎

We have shown that the best choice of $q^k$ is $s$ corresponding to the maximum entry in the vector $\Gamma b^{k-1}$ or the next biggest entry in $\Gamma b^{k-1}$. Recalling that we have assumed $\alpha = \beta$, both cases evaluate to

$$\mathbb{E}_{Y^k}\left[\max_{s\in S} b^k\left(s\right)\right] = \left(1 - \beta\right)\pi_1 + \max\left[\beta\pi_1, \left(1 - \beta\right)\pi_2\right] . \tag{14}$$

Further, any other choice results in a lower expected value.

It is tempting to continue our soltion using standard POMDP techniques (as in [15]). Unfortunately, such an approach is not computationally feasible, as the partition of the belief space grows exponentially in time. That is, at each time step, we would need to consider an exponentially branching sequence of search locations and possible observations. Further, the non-linearity of the objective function further complicates the analysis. This suggests that it may not be possible to compute an optimal solution to this problem. Instead, we will continue to examine the results of a one-step optimization for more insight into the infinite horizon problem.

A natural question, given that choosing location 1 or 2 results in the same value of the objective function in expectation, is which choice results in lower variance. In other words, does one choice provide an optimal cost with greater confidence? As usual, we write the variance

$$var\left(\max_{s\in S} b^k\left(s\right)\right) = \mathbb{E}\left[\left(\max_{s\in S} b^k\left(s\right)\right)^2\right] - \left(\mathbb{E}\left[\max_{s\in S} b^k\left(s\right)\right]\right)^2 . \tag{15}$$

We have already computed $\mathbb{E}\left[\max b^k\left(s\right)\right]$, and hence we know the second term in the variance equation, which is the same for both strategies. What remains is to compute the first term,

$$\mathbb{E}\left[\left(\max b^k\left(s\right)\right)^2\right]$$

$$= P\left(Y^k = 1\right) \max_{s\in S}\left(\eta P\left(Y^k = 1 | q^k, s\right)\hat{b}^k\right)^2$$

$$+ P\left(Y^k = 0\right) \max_{s\in S}\left(\eta P\left(Y^k = 0 | q^k, s\right)\hat{b}^k\right)^2 . \tag{16}$$

Again, by noticing that $\frac{1}{\eta}$ is $P\left(Y^k = y\right)$, we can write

$$\mathbb{E}\left[\left(\max b^k\left(s\right)\right)^2\right]$$
$$= \frac{1}{P\left(Y^k = 1\right)} \max_{s \in S} \left(P\left(Y^k = 1 | q^k, s\right)\hat{b}^k\right)^2 \tag{17}$$
$$+ \frac{1}{P\left(Y^k = 0\right)} \max_{s \in S} \left(P\left(Y^k = 0 | q^k, s\right)\hat{b}^k\right)^2 .$$

Then, since $P\left(Y^k = 1\right) = (1-\beta)\hat{b}^k\left(q^k\right) + \beta\left(1 - \hat{b}^k\left(q^k\right)\right)$ and $P\left(Y^k = 0\right) = \beta\hat{b}^k\left(q^k\right) + (1-\beta)\left(1 - \hat{b}^k\left(q^k\right)\right)$ we can evaluate the variance for a given choice of $q^k$.

Since we have already winnowed our strategies for a one-step horizon to looking at the locations of $\pi_1$ and $\pi_2$, it remains to compare the variance for each of these strategies. As noted above, since the strategies are the same in expectation, only $\mathbb{E}\left[\left(\max b^k\left(s\right)\right)^2\right]$ needs to be computed for each strategy.

*Theorem 2:* Choosing $q^k = 2$ minimizes the variance of $\max b^k\left(s\right)$.

*Proof:* There are two cases for which we need to prove the theorem: $(1-\beta)\pi_2 > \beta\pi_1$ and $\beta\pi_1 \geq (1-\beta)\pi_2$.

*Case 1:* $(1-\beta)\pi_2 > \beta\pi_1$

For strategy $q^k = 1$, (17) evaluates to

$$\mathbb{E}\left[\left(\max b^k\left(s\right)\right)^2 | q^k = 1\right] =$$
$$\frac{(1-\beta)^2\pi_1^2}{(1-\beta)\pi_1 + \beta(1-\pi_1)} + \frac{(1-\beta)^2\pi_2^2}{\beta\pi_1 + (1-\beta)(1-\pi_1)} , \tag{18}$$

while for strategy $q^k = 2$, it evaluates to

$$\mathbb{E}\left[\left(\max b^k\left(s\right)\right)^2 | q^k = 2\right] =$$
$$\frac{(1-\beta)^2\pi_2^2}{(1-\beta)\pi_2 + \beta(1-\pi_2)} + \frac{(1-\beta)^2\pi_1^2}{\beta\pi_2 + (1-\beta)(1-\pi_2)} . \tag{19}$$

*Case 2:* $\beta\pi_1 \geq (1-\beta)\pi_2$

For strategy $q^k = 1$, (17) evaluates to

$$\mathbb{E}\left[\left(\max b^k\left(s\right)\right)^2 | q^k = 1\right] =$$
$$\frac{(1-\beta)^2\pi_1^2}{(1-\beta)\pi_1 + \beta(1-\pi_1)} + \frac{\beta^2\pi_1^2}{\beta\pi_1 + (1-\beta)(1-\pi_1)} , \tag{20}$$

while for strategy $q^k = 2$, it evaluates to

$$\mathbb{E}\left[\left(\max b^k\left(s\right)\right)^2 | q^k = 2\right] =$$
$$\frac{\beta^2\pi_1^2}{(1-\beta)\pi_2 + \beta(1-\pi_2)} + \frac{(1-\beta)^2\pi_1^2}{\beta\pi_2 + (1-\beta)(1-\pi_2)} . \tag{21}$$

The strategy that minimizes variance can be found by determining which is smaller (18) or (19) for Case 1, and (20) or (21) for Case 2. For simplicity of notation, we will replace $(1-\beta)\pi_1 + \beta(1-\pi_1)$ with $P_{\pi_1}$ and $(1-\beta)\pi_2 + \beta(1-\pi_2)$ with $P_{\pi_2}$.

For Case 1, we will demonstrate that choosing $q^k = 2$ is the optimal choice by showing that

$$\frac{\pi_1^2}{P_{\pi_1}} + \frac{\pi_2^2}{1 - P_{\pi_1}} \geq \frac{\pi_2^2}{P_{\pi_2}} + \frac{\pi_1^2}{1 - P_{\pi_2}} . \tag{22}$$

Denote $\frac{\pi_2}{\pi_1}$ with $\gamma$, so $0 \leq \gamma \leq 1$ covers all valid $\pi_1$, $\pi_2$ pairs. The domain of valid $\pi_1$, $\gamma$ pairs is shown in Fig. 1. We will begin by demonstrating that

$$P_{\pi_2}\left(1 - P_{\pi_1}\right) - \gamma^2 P_{\pi_1}\left(1 - P_{\pi_2}\right) \geq 0 , \tag{23}$$

which is a simple rearrangement of the inequality in (22). Expanding this inequality and replacing $\pi_2$ with $\gamma\pi_1$ yields

$$\pi_1\gamma(1-2\beta) + \beta - (\pi_1(1-2\beta) + \beta)(\pi_1\gamma(1-2\beta) + \beta)$$
$$-\gamma^2\pi_1(1-2\beta) + \beta - (\pi_1(1-2\beta) + \beta)(\pi_1\gamma(1-2\beta) + \beta)$$
$$\geq 0 . \tag{24}$$

We will call the left-hand expression of (24) $f\left(\pi_1, \gamma\right)$. The domain of this function is displayed in Fig. 1. Then, the inequality in (24) will be demonstrated by showing that $f\left(\pi_1, \gamma\right)$ is non-negative along the boundaries of its domain and concave in $\pi_1$ thus proving its overall non-negativity, which proves the inequality in (22).

Starting with the boundary for $\pi_1 = \pi_2$, (i.e. $\gamma = 1$), clearly, $f\left(\pi_1, \gamma\right)$ evaluates to $0$. This further implies that the variance for choosing $\pi_1$ and $\pi_2$ is the same when $\pi_1 = \pi_2$.

We also wish to check the behavior at the other boundaries. Next, we examine the boundary $\pi_1 + \pi_2 = 1$ for $0.5 \leq \pi_1 \leq 1$. Then, $\pi_2 = 1 - \pi_1$, which means $0 \leq \pi_2 \leq 0.5$. We note that for $\pi_1 \geq 0.5$, $\pi_1 \geq P_{\pi_1}$. Thus we may write

$$\pi_1^2\left(1 - 2P_{\pi_1}\right) \geq P_{\pi_1}^2\left(1 - 2\pi_1\right) , \tag{25}$$

which we rearrange as follows:

$$\pi_1^2\left(1 - 2P_{\pi_1}\right) + \pi_1^2 P_{\pi_1}^2 \geq P_{\pi_1}^2\left(1 - 2\pi_1\right) + \pi_1^2 P_{\pi_1}^2 \tag{26}$$
$$\pi_1^2\left(1 - P_{\pi_1}\right)^2 \geq P_{\pi_1}^2\left(1 - \pi_1\right)^2 \tag{27}$$
$$\frac{\left(1 - P_{\pi_1}\right)^2}{P_{\pi_1}^2} \geq \frac{\left(1 - \pi_1\right)^2}{\pi_1^2} . \tag{28}$$

By noting that $1 - \pi_2 = \pi_1$ implies $P_{\pi_2} = 1 - P_{\pi_1}$, we may then further rearrange the expression to

$$\frac{P_{\pi_2}\left(1 - P_{\pi_1}\right)}{P_{\pi_1}\left(1 - P_{\pi_2}\right)} - \frac{\pi_2^2}{\pi_1^2} \geq 0 , \tag{29}$$

and finally

$$P_{\pi_2}\left(1 - P_{\pi_1}\right) - \frac{\pi_2^2}{\pi_1^2} P_{\pi_1}\left(1 - P_{\pi_2}\right) \geq 0 , \tag{30}$$

which is (23). Hence we have demonstrated the non-negativity of $f\left(\pi_1, \gamma\right)$ along the boundary $\pi_1 + \pi_2 = 1$.

Finally, we must check the behavior along the boundary where $\pi_2 = \frac{1 - \pi_1}{|S| - 1}$. We observe the following relationships:

$$\left(\pi_1^2 - \pi_2^2\right) \geq 0 \tag{31}$$
$$P_{\pi_2}\left(1 - P_{\pi_2}\right) \geq 0 , \tag{32}$$

which therefore imply

$$\left(\pi_1^2 - \pi_2^2\right) P_{\pi_2}\left(1 - P_{\pi_2}\right) \geq 0 . \tag{33}$$

We also note that since $\beta \leq 0.5$, then $1 - 2\beta \geq 0$. Then we may also write that

$$\pi_1^2 P_{\pi_2}\left(1 - |S|\pi_2\right)\left(1 - 2\beta\right) \geq 0 \tag{34}$$

and

$$\pi_2^2 \left(1 - P_{\pi_2}\right) \left(1 - |S|\pi_2\right) \left(1 - 2\beta\right) \geq 0 , \qquad (35)$$

since each of the terms in these two expressions is non-negative. The fact that $1 - |S|\pi_2$ is non-negative follows from the fact that this boundary represents a regime in which $\pi_2$ is no bigger than $\frac{1}{|S|}$. The sum is of the previous in equalities is similarly non-negative:

$$\pi_1^2 P_{\pi_2} \left(1 - |S|\pi_2\right) \left(1 - 2\beta\right) + \pi_2^2 \left(1 - P_{\pi_2}\right) \left(1 - |S|\pi_2\right) \left(1 - 2\beta\right) \geq 0 , \qquad (36)$$

and negating that expression flips the inequality to

$$-\pi_1^2 P_{\pi_2} \left(1 - |S|\pi_2\right) \left(1 - 2\beta\right) - \pi_2^2 \left(1 - P_{\pi_2}\right) \left(1 - |S|\pi_2\right) \left(1 - 2\beta\right) \leq 0 . \qquad (37)$$

Combining (33) and (37), we find

$$\begin{aligned}
\left(\pi_1^2 - \pi_2^2\right) &P_{\pi_2} \left(1 - P_{\pi_2}\right) \\
&\geq -\pi_1^2 P_{\pi_2} \left(1 - |S|\pi_2\right) \left(1 - 2\beta\right) \\
&\quad - \pi_2^2 \left(1 - P_{\pi_2}\right) \left(1 - |S|\pi_2\right) \left(1 - 2\beta\right) . \qquad (38)
\end{aligned}$$

Now we rearrange this expression as follows:

$$\begin{aligned}
\pi_1^2 P_{\pi_2} &\left(1 - P_{\pi_2}\right) + \pi_1^2 P_{\pi_2} \left(1 - |S|\pi_2\right) \left(1 - 2\beta\right) \\
&\geq \pi_2^2 P_{\pi_2} \left(1 - P_{\pi_2}\right) - \pi_2^2 \left(1 - P_{\pi_2}\right) \left(1 - |S|\pi_2\right) \left(1 - 2\beta\right) \qquad (39)
\end{aligned}$$

$$\begin{aligned}
\pi_1^2 P_{\pi_2} &\left(1 - P_{\pi_2} + \left(1 - |S|\pi_2\right) \left(1 - 2\beta\right)\right) \\
&\geq \pi_2^2 \left(P_{\pi_2} - \left(1 - |S|\pi_2\right) \left(1 - 2\beta\right)\right) \left(1 - P_{\pi_2}\right) \qquad (40)
\end{aligned}$$

$$\frac{P_{\pi_2} \left(1 - P_{\pi_2} + \left(1 - |S|\pi_2\right) \left(1 - 2\beta\right)\right)}{\left(P_{\pi_2} - \left(1 - |S|\pi_2\right) \left(1 - 2\beta\right)\right) \left(1 - P_{\pi_2}\right)} - \frac{\pi_2^2}{\pi_1^2} \geq 0 . \qquad (41)$$

By noting that $\pi_2 = \frac{1 - \pi_1}{|S| - 1}$ implies $P_{\pi_1} = P_{\pi_2} - \left(1 - |S|\pi_2\right) \left(1 - 2\beta\right)$ and $\pi_1 = 1 - (|S| - 1)\pi_2$, we finally write (41) as

$$\frac{P_{\pi_2} \left(1 - P_{\pi_1}\right)}{P_{\pi_1} \left(1 - P_{\pi_2}\right)} - \frac{\pi_2^2}{\pi_1^2} \geq 0 \qquad (42)$$

and then

$$P_{\pi_2} \left(1 - P_{\pi_1}\right) - \frac{\pi_2^2}{\pi_1^2} P_{\pi_1} \left(1 - P_{\pi_2}\right) \geq 0 \qquad (43)$$

which is again (23). We note that the term $P_{\pi_1} \left(1 - P_{\pi_2}\right)$ is always positive, hence the direction of the inequality in (41) is always correct. This demonstrates the non-negativity of $f\left(\pi_1, \gamma\right)$ along the boundary $\pi_2 = \frac{1 - \pi_1}{|S| - 1}$.

To complete our proof, we show the concavitiy of $f\left(\pi_1, \gamma\right)$ with respect to $\pi_1$. Taking the second derivative of $f\left(\pi_1, \gamma\right)$ yields

$$\frac{\partial^2 f\left(\pi_1, \gamma\right)}{\partial \pi_1^2} = 2\gamma \left(\gamma^2 - 1\right) \left(1 - 2\beta\right)^2 . \qquad (44)$$

It is clear that the term $\left(1 - 2\beta\right)^2$ and $2\gamma$ are non-negative. Since $0 \leq \gamma \leq 1$, then $\left(\gamma^2 - 1\right) \leq 0$. Hence, the second derivative of our function is non-positive and therefore the function is concave with respect to $\pi_1$.

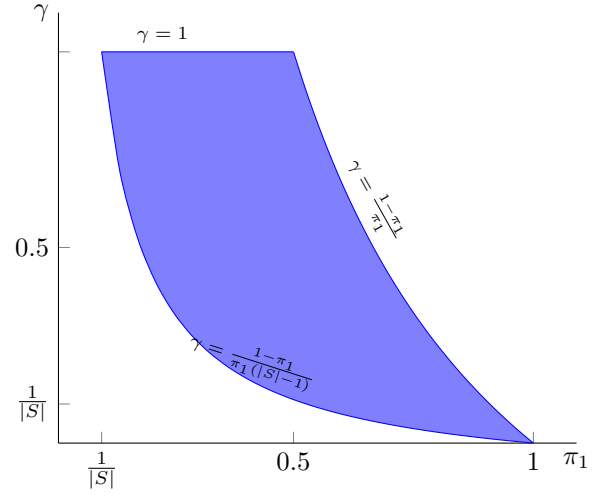Now, we have demonstrated the function is equal to zero for $\pi_1 = \pi_2$, and non-negative along the rest of the



Fig. 1: Colored region marks the $\pi_1, \gamma$ pairs constituting the domain of $f\left(\pi_1, \gamma\right)$. Corresponding constraints on $\pi_1$ and $\gamma$ are labeled for the borders of the region.
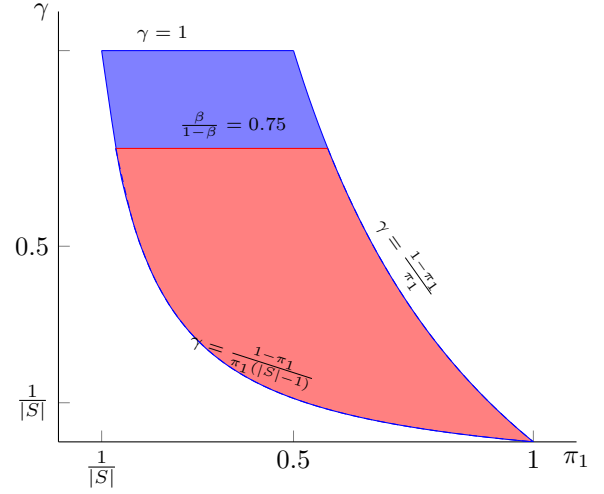


Fig. 2: Red region marks the $\pi_1, \gamma$ pairs constituting the domain of $g\left(\pi_1, \gamma, \beta\right)$ for a given value of $\beta$. We include the boundary for $\beta = 0.75$ as an example. Corresponding constraints on $\pi_1$ and $\gamma$ are labeled for the borders of the region.

boundaries. Further, by demonstrating that the function is concave in $\pi_1$, we know the values of the function between these boundaries must also be non-negative. Hence, the function is greater than or equal to zero for the entire feasible set of values of $\pi_1$, $\pi_2$, and $\beta$.

For Case 2, to show that the optimal choice of $q^k$ is 2, we must show that

$$\frac{\left(1 - \beta\right)^2 \pi_1^2}{P_{\pi_1}} + \frac{\beta^2 \pi_1^2}{1 - P_{\pi_1}} \geq \frac{\beta^2 \pi_1^2}{P_{\pi_2}} + \frac{\left(1 - \beta\right)^2 \pi_1^2}{1 - P_{\pi_2}} . \qquad (45)$$

We will follow the same process that was followed in Case 1 to prove this inequality. We begin by rearranging the

inequality to

$$\frac{(1-\beta)^2}{P_{\pi_1}} + \frac{\beta^2}{1-P_{\pi_1}} \geq \frac{\beta^2}{P_{\pi_2}} + \frac{(1-\beta)^2}{1-P_{\pi_2}} . \qquad (46)$$

Now we make common denominators:

$$\frac{(1-P_{\pi_2})(1-\beta)^2}{P_{\pi_1}(1-P_{\pi_2})} + \frac{P_{\pi_2}\beta^2}{P_{\pi_2}(1-P_{\pi_1})}$$
$$\geq \frac{(1-P_{\pi_1})\beta^2}{P_{\pi_2}(1-P_{\pi_1})} + \frac{P_{\pi_1}(1-\beta)^2}{P_{\pi_1}(1-P_{\pi_2})} . \qquad (47)$$

We can group terms to get

$$\frac{(1-P_{\pi_1}-P_{\pi_2})(1-\beta)^2}{P_{\pi_1}(1-P_{\pi_2})} \geq \frac{(1-P_{\pi_1}-P_{\pi_2})\beta^2}{P_{\pi_2}(1-P_{\pi_1})} . \qquad (48)$$

For the boundary where $1 = \pi_1 + \pi_2$, we note that

$$1 = P_{\pi_1} + P_{\pi_2} , \qquad (49)$$

which implies that both sides of (48) evaluate to 0. To evaluate along the other two boundaries, we note that elsewhere, $1 - P_{\pi_1} - P_{\pi_2} > 0$ and we continue simplifying to get

$$\frac{P_{\pi_2}(1-P_{\pi_1})}{P_{\pi_1}(1-P_{\pi_2})} \geq \frac{\beta^2}{(1-\beta)^2} , \qquad (50)$$

and finally rearranging to get

$$P_{\pi_2}(1-P_{\pi_1}) - \frac{\beta^2}{(1-\beta)^2}P_{\pi_1}(1-P_{\pi_2}) \geq 0 . \qquad (51)$$

We will call the left-hand side of (51) $g(\pi_1, \gamma, \beta)$. We will demonstrate that $g(\pi_1, \gamma, \beta)$ is non-negative for the remaining two boundaries, thus proving the inequality in (45).

We note that we have shown the non-negativity of $f(\pi_1, \gamma)$ over its entire domain in Case 1. Then the boundary of $g(\pi_1, \gamma, \beta)$ where $\frac{\beta}{1-\beta} = \gamma$ is $f(\pi_1, \gamma)$ with $\beta = \frac{\gamma}{1+\gamma}$. This boundary has therefore already been shown to be non-negative, since it is contained in Case 1.

Therefore, we must only check the behavior along the boundary where $\pi_2 = \frac{1-\pi_1}{|S|-1}$. As in Case 1, we observe:

$$\left((1-\beta)^2 - \beta^2\right) \geq 0 \qquad (52)$$
$$P_{\pi_2}(1-P_{\pi_2}) \geq 0 , \qquad (53)$$

which therefore imply

$$\left((1-\beta)^2 - \beta^2\right) P_{\pi_2}(1-P_{\pi_2}) \geq 0 . \qquad (54)$$

We also note that since $\beta \leq 0.5$, then $1 - 2\beta \geq 0$. Then we may also write that

$$(1-\beta)^2 P_{\pi_2}(1-|S|\pi_2)(1-2\beta) \geq 0 \qquad (55)$$

and

$$\beta^2(1-P_{\pi_2})(1-|S|\pi_2)(1-2\beta) \geq 0 , \qquad (56)$$

since each of the terms in these two expressions is non-negative. As before, the fact that $1 - |S|\pi_2$ is non-negative follows from the fact that this boundary represents a regime

in which $\pi_2$ is no bigger than $\frac{1}{|S|}$. The sum is of the previous in equalities is similarly non-negative:

$$(1-\beta)^2 P_{\pi_2}(1-|S|\pi_2)(1-2\beta) + \beta^2(1-P_{\pi_2})(1-|S|\pi_2)(1-2\beta) \geq 0 , \qquad (57)$$

and negating that expression flips the inequality to

$$-(1-\beta)^2 P_{\pi_2}(1-|S|\pi_2)(1-2\beta) - \beta^2(1-P_{\pi_2})(1-|S|\pi_2)(1-2\beta) \leq 0 . \qquad (58)$$

Combining (54) and (58), we find

$$\left((1-\beta)^2 - \beta^2\right) P_{\pi_2}(1-P_{\pi_2})$$
$$\geq -(1-\beta)^2 P_{\pi_2}(1-|S|\pi_2)(1-2\beta)$$
$$- \beta^2(1-P_{\pi_2})(1-|S|\pi_2)(1-2\beta) . \qquad (59)$$

Now we rearrange this expression as follows:

$$(1-\beta)^2 P_{\pi_2}(1-P_{\pi_2}) + (1-\beta)^2 P_{\pi_2}(1-|S|\pi_2)(1-2\beta)$$
$$\geq \beta^2 P_{\pi_2}(1-P_{\pi_2}) -$$
$$\beta^2(1-P_{\pi_2})(1-|S|\pi_2)(1-2\beta) \qquad (60)$$

$$(1-\beta)^2 P_{\pi_2}(1-P_{\pi_2} + (1-|S|\pi_2)(1-2\beta))$$
$$\geq \beta^2(P_{\pi_2} - (1-|S|\pi_2)(1-2\beta))(1-P_{\pi_2}) \qquad (61)$$

$$\frac{P_{\pi_2}(1-P_{\pi_2}+(1-|S|\pi_2)(1-2\beta))}{(P_{\pi_2}-(1-|S|\pi_2)(1-2\beta))(1-P_{\pi_2})} - \frac{\beta^2}{(1-\beta)^2} \geq 0 . \qquad (62)$$

By noting that $\pi_2 = \frac{1-\pi_1}{|S|-1}$ implies $P_{\pi_1} = P_{\pi_2} - (1-|S|\pi_2)(1-2\beta)$ and $\pi_1 = 1 - (|S|-1)\pi_2$, we finally write (62) as

$$\frac{P_{\pi_2}(1-P_{\pi_1})}{P_{\pi_1}(1-P_{\pi_2})} - \frac{\beta^2}{(1-\beta)^2} \geq 0 \qquad (63)$$

and then

$$P_{\pi_2}(1-P_{\pi_1}) - \frac{\beta^2}{(1-\beta)^2}P_{\pi_1}(1-P_{\pi_2}) \geq 0 \qquad (64)$$

which is again (51). We note that the term $P_{\pi_1}(1-P_{\pi_2})$ is always positive, hence the direction of the inequality in (62) is always correct. This demonstrates the non-negativity of $g(\pi_1, \gamma, \beta)$ along the boundary $\pi_2 = \frac{1-\pi_1}{|S|-1}$.

Finally, we compute the second derivative of $g(\pi_1, \gamma, \beta)$

$$\frac{\partial^2 g(\pi_1, \gamma, \beta)}{\partial \pi_1^2} = \frac{2\gamma(2\beta-1)^3}{(1-\beta)^2} . \qquad (65)$$

We know that $\gamma \geq 0$ and $(1-\beta)^2 > 0$. Further, $(2\beta - 1) < 0$, since $\beta < 0.5$, hence (65) is negative, and $g(\pi_1, \gamma, \beta)$ is concave in $\pi_1$. We have then demonstrated that $g(\pi_1, \gamma, \beta)$ is non-negative along its boundaries and concave in $\pi_1$, and is therefore non-negative over its entire domain.

The proofs for Case 1 and Case 2 together show that choosing $q^k = 2$ is always optimal for a one-step optimization.
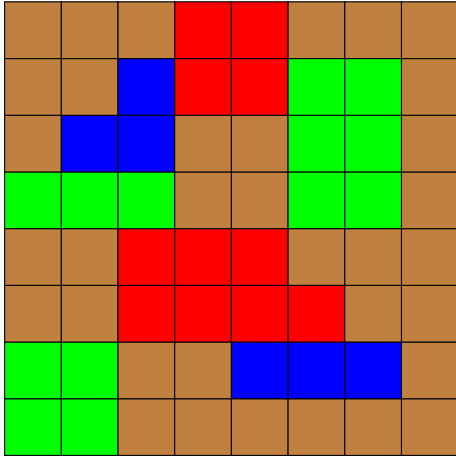
∎

Fig. 3: Environment used in the case study. Green indicates grassland, blue indicates water sources, red indicates predator habitat, and brown indicates plains.

### A. Comparison with POMDPs

The problem considered in this paper is similar to many Partially Observable Markov Decision Process (POMDP) models. In contrast to solving Markov Decision Processes (MDP), which involves optimizing with respect to a cost function $J(x)$ over a set of states $x \in \{1, \ldots, N\}$, solving POMDPs involves optimizing with respect to the expectation of a cost function over a belief. That is, for a belief $b(s)$ over a set of states, $S$, optimize $\mathbb{E}_b[J(s)] = \sum_s J(s)b(s)$. Note that this cost function is linear in the belief, $b$. Smallwood and Sondik [15] showed that optimal solutions can be computed for such problems by repeatedly partitioning the belief space with hyperplanes, and performing a recursive dynamic program. This recursion leverages the fact that the cost is linear in $b$. Even despite this linearity, in many problems, computational complexity renders such an approach infeasible. In contrast, in our problem both of the objective functions $J_1(b)$ from (6) and $J_2(b)$ from (7) are non-linear in $b$, and hence they do not fit in to the typical POMDP problem setup. Therefore known solution techniques for POMDPs involving recursive linear partitioning of the belief space do not apply in our case.

Instead, this work focuses on finding policies that optimize a non-linear cost function with respect to *belief*, rather than a linear cost function over states. While finding an optimal policy in such a problem appears to be impossible in our case, we present results showing that a simple index based approach has predictable and desirable results, even over an infinite horizon. Moreover, the nature of an index based policy allows for rapid implementation, due to the simple nature of the decision rule.

### IV. Simulations and Results

To demonstrate our results from Sec. III, we simulate the wildlife tracking problem introduced in Sec. I. The environment we consider is a discrete $8 \times 8$ grid, consisting of four types of cells: grassland, water sources, predator habitat, and plains. The probability of the animal staying in each of these types of cells is $0.9$, $0.75$, $0.1$ and $0.4$, respectively. That is, the animal prefers to hide in grassland and spend time near water, while spending less time in open areas and predator territory. If the animal does not stay in its current cell, it is equally likely to move to any of its neighboring cells.

Simulations were performed to compare three strategies: always choose $q^k = 1$, and always choose $q^k = 2$, and minimize the expected one-step entropy. We include the expected one-step entropy minimization to demonstrate the efficacy of both strategies with respect to a different choice of simple one-step strategy. For each strategy, 1000 simulations were performed, with a horizon of 100 time steps. The value of $\alpha$ and $\beta$ was set to $0.05$ for each simulation. The target location was initialized randomly based on the steady state distribution for the target Markov chain for each simulation, and the initial belief was set to the steady state distribution.

Figure 4 shows the mean and standard error for $max_{s \in S} b^k(s)$ for each of the three strategies. In the mean, the performance of all three scenarios is comparable. Any small differences in the mean performance of the $q^k = 1$ and $q^k = 2$ strategies are due to random variations in the Monte Carlo simulations. However $q^k = 2$ performs much better than $q^k = 1$ in terms of variance, as is expected. These results suggest a much more confident estimate of target location for strategies that result in minimum variance. Figure 4 suggests that although our policy is based on a one-step optimization, it performs well over arbitrary horizons.

### V. Discussion

It might seem counter-intuitive that it is best to look in the second most likely place, but this is indeed the best strategy, as is clearly demonstrated in our simulation results. This result stems from the fact that a negative measurement for the target in the second most likely location makes the most likely location more likely, while a positive measurement in that location results in most of the mass of the belief pmf being moved to that location. In contrast, looking in the most likely location and getting a negative measurement results in a more uniform belief pmf, while a positive measurement produces largely the same result as it would when looking in the second most likely location.

The results of this work demonstrate that we have found an effective, computationally efficient method for tracking a moving target with high confidence. There is great opportunity for future work in this area. For example, considering scenarios with multiple searchers or targets as well as motion constraints on the searcher are natural extensions of this problem. Further, looking at variations in false alarm and missed detection rates may yield different policies. Likewise, restrictions on the types of target motion and environment topology may allow even simpler policies to be found, as in the special case for searching on a grid. The great variety of
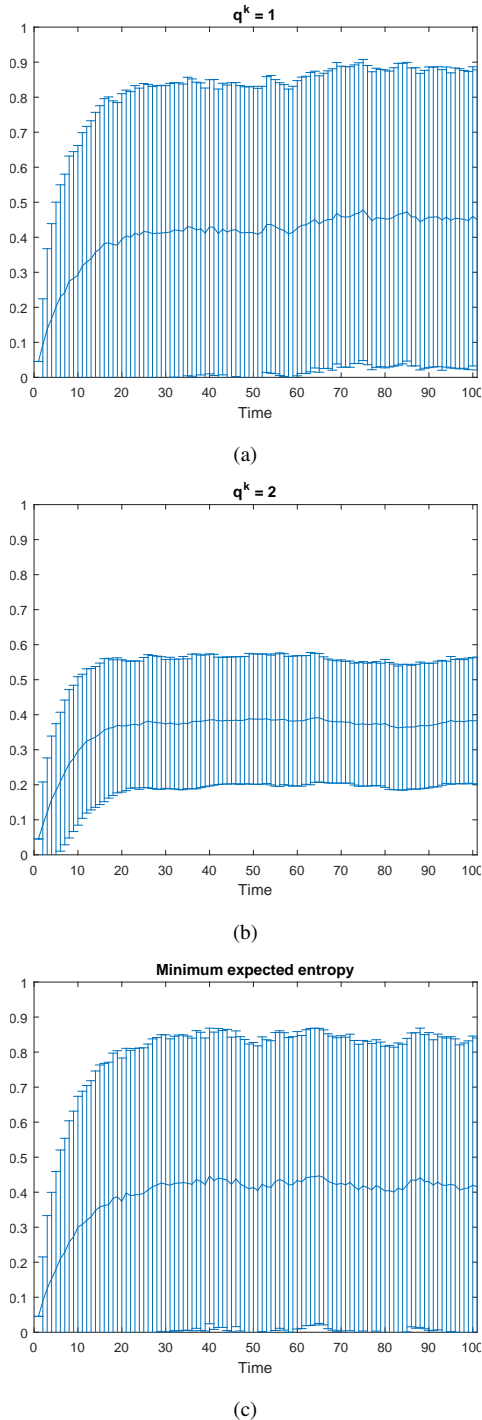
(a)



(b)



(c)

Fig. 4: Mean and standard error results for maximum belief from 1000 simulations with time horizon 100.

potential applications for such work suggests that pursuing these research avenues will prove worthwhile.

REFERENCES

[1] D. Castanon. Optimal search strategies in dynamic hypothesis testing. *Systems, Man and Cybernetics, IEEE Transactions on*, 25(7):1130–1138, 1995.
[2] T. H. Chung, G. A. Hollinger, and V. Isler. Search and pursuit-evasion in mobile robotics. *Autonomous robots*, 31(4):299–316, 2011.
[3] N. Atanasov, B. Sankaran, J. Le Ny, G. J. Pappas, and K. Daniilidis. Nonmyopic view planning for active object classification and pose estimation. *Robotics, IEEE Transactions on*, 30(5):1078–1090, 2014.
[4] A. Abdel-Samad, A. H. Tewfik, et al. Search strategies for radar target localization. In *Image Processing, 1999. ICIP 99. Proceedings. 1999 International Conference on*, volume 3, pages 862–866. IEEE, 1999.
[5] Shuchin Aeron, Venkatesh Saligrama, David Castañón, et al. Energy efficient policies for distributed target tracking in multihop sensor networks. In *Decision and Control, 2006 45th IEEE Conference on*, pages 380–385. IEEE, 2006.
[6] K. Cohen and Q. Zhao. Quickest anomaly detection: a case of active hypothesis testing. In *Information Theory and Applications Workshop (ITA), 2014*, pages 1–5. IEEE, 2014.
[7] A. Wald and J. Wolfowitz. Optimum character of the sequential probability ratio test. *The Annals of Mathematical Statistics*, pages 326–339, 1948.
[8] S. S. Brown. Optimal search for a moving target in discrete time and space. *Operations Research*, 28(6):1275–1289, 1980.
[9] A. R. Washburn. Search for a moving target: The fab algorithm. *Operations research*, 31(4):739–751, 1983.
[10] L. D. Stone. Necessary and sufficient conditions for optimal search plans for moving targets. *Mathematics of Operations Research*, 4(4):431–440, 1979.
[11] W. R. Stromquist and L. D. Stone. Constrained optimization of functionals with search theory applications. *Mathematics of Operations Research*, 6(4):518–529, 1981.
[12] J. B. Kadane. Optimal whereabouts search. *Operations Research*, 19(4):894–904, 1971.
[13] D. Assaf and S. Zamir. Optimal sequential search: a bayesian approach. *The Annals of Statistics*, pages 1213–1221, 1985.
[14] K. Cohen, Q. Zhao, and A. Swami. Optimal index policies for anomaly localization in resource-constrained cyber systems. *Signal Processing, IEEE Transactions on*, 62(16):4224–4236, 2014.
[15] R. D. Smallwood and E. J. Sondik. The optimal control of partially observable markov processes over a finite horizon. *Operations Research*, 21(5):1071–1088, 1973.