Universal Scheme for Optimal Search and Stop

Sirin Nitinawarat
Qualcomm Technologies, Inc.
5775 Morehouse Drive
San Diego, CA 92121, USA
Email: sirin.nitinawarat@gmail.com

Venugopal V. Veeravalli
Coordinated Science Laboratory
Department of Electrical and Computer Engineering
University of Illinois at Urbana-Champaign
Urbana, IL, 61801, USA
Email: vvv@illinois.edu

Abstract—The problem of universal search and stop using an adaptive search policy is considered. When the target location is searched, the observation is distributed according to the target distribution, otherwise it is distributed according to the absence distribution. A universal sequential scheme for search and stop is proposed using only the knowledge of the absence distribution, and its asymptotic performance is analyzed. The universal test is shown to yield a vanishing error probability, and to achieve the optimal reliability when the target is present, universally for every target distribution. Consequently, it is established that the knowledge of the target distribution is only useful for improving the reliability for detecting a missing target. It is also shown that a multiplicative gain for the search reliability equal to the number of searched locations is achieved by allowing adaptivity in the search.

I. INTRODUCTION

We study the problem of universal search and stop using an adaptive search policy. When the target location is searched, the observation is assumed to be distributed according to the target distribution, otherwise it is distributed according to the absence distribution. We assume that only the absence distribution is known, and the target distribution can be arbitrarily distinct from the absence distribution. An adaptive search policy specifies the current search location based on the past observations and past search locations. At the stopping time, the target’s location is determined or it is decided that it is missing. The overall goal is to achieve a certain level of accuracy for the final decision using the fewest number of observations. The results in this paper should be regarded as a contribution to the long-studied area of search theory (see, e.g., [1, 2, 3, 8, 10, 18]), in particular, searching for a stationary target in discrete time and space with a discrete search effort (cf. [3][Subsection 4.2]).

Conceptually, a desirable goal of the search at each location should be to determine if the target is there. To this end, a universal sequential test for two hypotheses can be used at each location to collect multiple subsequent observations that will eventually lead to a binary outcome that the target is there or not. To improve reliability for this binary decision at a particular search location, one can use a test that takes more observations at that location. If we insist on using the mentioned sequential binary test at each location as an “inner” test, then it is convenient to select the current search location based on the past binary outcomes of the subsequent binary tests (instead of all the past observational outcomes of all the searches, generally taken multiple times at each of the locations). With this imposition, the search and stop problem can be conceptually reduced to the problem of constructing an “outer” test for the sequential design of such inner experiments. This intuitive decomposition leads to our proposed universal sequential test for search and stop.

Universal sequential testing for two hypotheses was first considered for certain parametric families of distributions for continuous observation spaces in [6, 7, 11, 17], the latest of which employed the concept of time-dependent thresholding. Here in Subsection II-A, we look at a non-parametric family of distributions for a finite observation space, for which we propose a universal test using a suitable time-dependent threshold and analyze its performance.

Sequential design of experiments with a uniform experimental cost was first considered in [5, 4] under a certain positivity assumption for the model, which was successfully dispensed with later in [15, 14]. A generalization of the model with a more complicated memory structure for the experimental outcomes and with non-uniform experimental cost was studied in [16].

We show that when the target is present, the proposed universal test based on the aforementioned decomposition yields a vanishing error probability, and achieves the optimal reliability, in terms of a suitable exponent for the error probability, universally for every target distribution. Consequently, we establish that the knowledge of the target distribution is only useful for improving the reliability for detecting a missing target. We also show that a multiplicative gain for the search reliability equal to the number of searched locations is achieved by allowing adaptivity in the search.

We review the pertinent existing results on universal sequential testing for two hypotheses and sequential design of experiments in Subsections II-A and II-B, respectively. The general model for universal search and stop is set up in Section III. We present the proposed sequential test for search and stop and state the main result pertaining to its performance in Section IV. Due to space limitations, we omit all the proofs.

II. PRELIMINARIES

Throughout the paper, random variables (rvs) are denoted by capital letters, and their realizations are denoted by the corresponding lower-case letters. All rvs are assumed to take values in finite sets, and all logarithms are the natural ones.
For a finite set $\mathcal{X}$, and a probability mass function (pmf) $p$ on $\mathcal{X}$ we write $X \sim p$ to denote that the rv $X$ is distributed according to $p$.

The following technical facts will be useful; their derivations can be found in [9, Chapter 11]. Consider random variables $Y^n = (Y_1, \ldots, Y_n)$ which are independent and identically distributed (i.i.d.) according to a pmf $p$ on $\mathcal{Y}$, i.e., $Y_i \sim p$, $i = 1, \ldots, n$. Let $y^n = (y_1, \ldots, y_n) \in \mathcal{Y}^n$ be a sequence with an empirical distribution $\gamma = \gamma^{(n)}$ on $\mathcal{Y}$. It follows that the probability of such sequence $y^n$, under the i.i.d. assumption according to the pmf $p$, is

$$p(y^n) = e^{-n[D(\gamma||p)+H(\gamma)]},$$

where $D(\gamma||p)$ and $H(\gamma)$ are the relative entropy of $\gamma$ and $p$, and entropy of $\gamma$, defined as

$$D(\gamma||p) \triangleq \sum_{y \in \mathcal{Y}} \gamma(y) \log \frac{\gamma(y)}{p(y)},$$

and

$$H(\gamma) \triangleq -\sum_{y \in \mathcal{Y}} \gamma(y) \log \gamma(y),$$

respectively. Consequently, it holds that for each $y^n$, the pmf $p$ that maximizes $p(y^n)$ is $p = \gamma$, and the associated maximal probability of $y^n$ is

$$\gamma(y^n) = e^{-nH(\gamma)}.$$  

Next, for each $n \geq 1$, the number of all possible empirical distributions from a sequence of length $n$ in $\mathcal{Y}^n$ is upper bounded by $(n+1)^{|\mathcal{Y}|}$. In particular, using this last fact, it can be shown that for any $\epsilon > 0$, it holds that the probability of the i.i.d. sequence $Y^n$ under $p$ satisfies

$$\mathbb{P}[D(\gamma||p) \geq \epsilon] \leq (n+1)^{|\mathcal{Y}|} e^{-n\epsilon}.$$  

We now review the relevant preliminary results on universal sequential testing for two hypotheses, and model-based sequential design of experiments with varying experimental cost in Subsections II-A and II-B, respectively. These results will be key to our proposed universal set for search and stop.

**A. Universal Sequential Testing for Two Hypotheses**

Consider sequential testing between the null hypothesis $H_0$ with i.i.d. observations $Y_k \in \mathcal{Y}$, $k = 1, 2, \ldots$, according to a pmf $\pi$ on $\mathcal{Y}$, and the alternative hypothesis $H_1$ with i.i.d. $Y_k$, $k = 1, 2, \ldots$, according to a pmf $\mu \neq \pi$. We assume that only $\pi$ is known, and nothing is known about $\mu$, i.e., it can be arbitrarily close to $\pi$. We further assume that both $\mu$ and $\pi$ have full support on $\mathcal{Y}$.

For a threshold parameter $a > 1$, we shall employ a sequential test defined in terms of the following (Markov) time:

$$\hat{N}_b \triangleq \arg\min_{n \geq 1} \left[ nD(\gamma||\pi) > \left( \log a + n \frac{b}{\log a} + |\mathcal{Y}| \log (n+1) \right) \right]$$

where $\gamma$ denotes the empirical distribution of the observation sequence $(y_1, \ldots, y_n)$ . The test stops at this time or $|a (\log a)^{\rho_1}|$ for some $\rho_1 > 1$, depending on which one is smaller, i.e., it stops at time $N^b$, where

$$N^b \triangleq \min \left\{ \hat{N}_b, |a (\log a)^{\rho_1}| \right\}.$$  

Correspondingly, the final decision is made according to

$$\delta_b \left( Y^{N^b} \right) = \begin{cases} 1 & \text{if } \hat{N}_b \leq |a (\log a)^{\rho_1}| \\ 0 & \text{if } \hat{N}_b > |a (\log a)^{\rho_1}| \end{cases}.$$  

**Lemma II.1.** With $\mu$ and $\pi$ having full support on $\mathcal{Y}$, for every $a > 1$, the sequential test in (II.4), (II.5), (II.6) yields that

$$\alpha_a \triangleq \mathbb{P}_0 \left[ \delta_b \left( Y^{N^b} \right) = 1 \right] \leq \frac{1}{a}.$$  

In addition, for any $\nu < 1$, $\mu \neq \pi$ and every $a > a^*$ ($\nu, \mu, \pi$), the test also yields that

$$c_a \triangleq \mathbb{E}_1 \left[ N^b \right] \leq \mathbb{E}_1 \left[ \hat{N}_b \right] \leq \frac{\log a}{\nu D(\mu||\pi)},$$

$$\kappa_a \triangleq \mathbb{E}_0 \left[ N^b \right] \leq |a (\log a)^{\rho_1}|,$$

$$\beta_a \triangleq \mathbb{P}_1 \left[ \delta_b \left( Y^{N^b} \right) = 0 \right] = \mathbb{P}_1 \left[ \hat{N}_b > |a (\log a)^{\rho_1}| \right]$$

$$\leq \frac{\mathbb{E}_1 \left[ \hat{N}_b \right]}{a (\log a)^{\rho_1}} \leq \frac{1}{\nu D(\mu||\pi) a (\log a)^{(\rho_1-1)}}.$$  

**B. Sequential Design of Experiments with Varying Experimental Cost**

Now we turn our attention to another sequential decision-making problem (a model-based one this time). Consider the problem of sequential design of experiments to facilitate the eventual testing for $H$ hypotheses. We assume a (conditionally) memoryless model for the outcome conditioned on the currently chosen experiment. In particular, under the $i$-th hypothesis, $i \in \{1, \ldots, H\} = [H]$, and conditioned on the current experiment $u_t = u \in \mathcal{U}$, at time $t = 1, 2, \ldots$, the current outcome of the experiment, denoted by $Z_t$, is assumed to be conditionally independent of all past outcomes and past experiments $U^{t-1}, Z^{t-1}$, and to be conditionally distributed according to a pmf $p^u_{i^*}$ on $\mathcal{Z}$. There is a cost function $c : [H] \times \mathcal{U} \to \mathbb{R}^+$, and the current experiment $u_t$ is assumed to incur a cost of $c(i, u_t)$ under the $i$-th hypothesis. We assume that for every $i = 1, \ldots, H$, $u \in \mathcal{U}$, $z \in \mathcal{Z}$, $p^u_{i^*}(z) > 0$, $c(i, u) > 0$. A test consists of a adaptive policy $\phi$ that chooses each experiment as a suitable (possibly randomized) function of past experiments and their outcomes, a stopping time $\tau$, and a final decision rule $\delta$ that outputs a guess of a hypothesis in $[H]$. The goal is to design a test to optimize the tradeoff between the cost accumulated up to
the final decision, as measured by \( \sum_{t=1}^{\tau} c(i, U_t), i \in [H] \), and the accuracy of the final decision, as measured by \( P_{\text{max}} \equiv \max_{i=1,\ldots,H} \mathbb{P}_i [\delta (Z^*) \neq i] \). The problem is model-based: all the (conditional) distributions \( p_h, i \in [H], u \in \mathcal{U} \), and the cost function \( c \) are assumed to be known.

For each hypothesis \( i \in [H] \), let

\[
q^*_i(u) = \arg\max_{q} \frac{\min_{j \neq i} \sum_u q(u) D(p^n_j \| p^n_i)}{\sum_u q(u) c(i, u)}. \tag{II.12}
\]

Then an asymptotically optimal test can be specified based on these distributions as follows. At each time \( t \geq 1 \), the ML estimate of the true hypothesis \( \hat{i} \) can be computed based on past experiments and their outcomes \( u^{t-1}, z^{t-1} \) using the model \( p^n_{\hat{i}}, i \in [H], u \in \mathcal{U} \) (ties are broken arbitrarily). For \( b > 0 \), during the sparse occasions \( t = \lfloor e^{bt} \rfloor \), \( k = 0, 1, \ldots \), the experiment is selected to explore all possible options in \( \mathcal{U} \) in a round-robin manner independently of \( \hat{i} \): for \( \mathcal{U} = \{u_1, \ldots, u_{|\mathcal{U}|}\} \),

\[
u_t = u(k \mod |\mathcal{U}|) + 1. \tag{II.13}
\]

At all other times, the current (random) experiment is selected as \( U_t \sim q^*_i \). Denote the joint distribution under the \( i \)-th hypothesis of all experiments and their outcomes up to time \( t \) (induced by the control policy) by \( p_i(z^t, u^t) \). For a threshold \( a' > 1 \), the test stops at time \( \tau^* \) and decides in favor of the ML hypothesis according to the rule \( \delta^* \), where

\[
\tau^* \triangleq \arg\min_{t} \min_{j \neq i} \frac{p_i(z^t, u^t)}{p_j(z^t, u^t)} > a', \quad \delta^*(z^\tau, u^{\tau^*}) = \hat{i}. \tag{II.14}
\]

Note that as the \( q^*_i, i \in [H] \), are, in general, not point-mass distributions, in addition to the realization of all experimental outcomes \( z^t \), we also need to account for the realization of the experiments \( u^t \) as well in the instantaneous computation of the ML hypothesis and the stopping criterion (II.14). If the experiments have been chosen deterministically at all times, we can just use the joint distributions of all experimental outcomes \( p_i(z^t), i \in [H], t = 1, 2, \ldots \) in these computations. The resulting test is asymptotically optimal and its performance is characterized in Proposition II.1 as follows.

**Proposition II.1** ([16]). \(^1\) For \( b > 0 \), in (II.13) chosen to be sufficiently small, and as \( a' \to \infty \), the test in (II.12), (II.13), (II.14) yields a vanishing error probability \( P_{\text{max}} \to 0 \), and satisfies for each \( i = 1, \ldots, H \), that

\[
E_i \left[ \sum_{t=1}^{\tau^*} c(i, U_t) \right] \leq \frac{-\log P_{\text{max}}}{\min_q \sum_u q(u) D(p^n_i \| p^n_j)} \left(1 + o(1)\right), \tag{III.1}
\]

1The result in [16] was proven for the model in which the cost function depends only on the experiment; however, the proof generalizes to the current setting when the cost function also depends on the hypothesis.

In addition, the proposed test is asymptotically optimal in the sense that any sequence of tests \( (\phi, \tau, \delta) \) that achieve \( P_{\text{max}} \to 0 \) must satisfy

\[
E_i \left[ \sum_{t=1}^{\tau} c(i, U_t) \right] \geq \frac{-\log P_{\text{max}}}{\max_{q, i} \sum_u q(u) D(p^n_i \| p^n_j)} \left(1 + o(1)\right), \tag{III.2}
\]

for every \( i = 1, \ldots, H \).

### III. Model for Search and Stop

Consider searching for a single target located in one of the \( M \) locations. At each time \( k \geq 1 \), if a location without the target is searched, then the observation \( Y_k \in \mathcal{Y} \) is assumed to be conditionally independent of all past observations and past search locations, and to be conditionally distributed according to the absence distribution \( \pi \). The distribution \( \pi \) represents pure noise, and we shall assume that this distribution is known to the searcher. On the other hand, if the target location is searched, then the observation would be conditionally distributed according to the “target” distribution \( \mu \) on \( \mathcal{Y} \) (and would be conditionally independent of past observations and search locations). We assume that both \( \mu \) and \( \pi \) have full support on \( \mathcal{Y} \).

We also allow for the possibility of an absent target. In this latter case, the observations at all locations are distributed according to \( \pi \). Denote the search location at time \( k \geq 1 \) by \( U_k \in [M] \), which is allowed to be any function of all past observations \( Y_{k-1} = (Y_1, \ldots, Y_{k-1}) \) and past search locations \( U_{k-1} = (U_1, \ldots, U_{k-1}) \).

It is interesting to note that the most basic search problem with an overlook probability \( \alpha > 0 \) (see, e.g., Chapters 4, 5 of [18] and Section 4.2 of [3]) that is uniform over all locations, corresponds to a special case of our general model wherein \( \mathcal{Y} = \{0,1\} \), \( \mu(0) = \alpha, \pi(0) = 1 \). In contrast, our model allows for any general (finite) observation space \( \mathcal{Y} \), but assumes that both \( \mu \) and \( \pi \) have full supports. The degeneracy in the model for the classic search problem as mentioned affords the construction of a search plan that is more efficient than that for our model (with the assumption of full support). The main concern for the classic search problem has been to come up with the search plan that is absolutely optimal (non-asymptotically), whereas our main concern is to construct a universal test that is asymptotically efficient in the regime of vanishing error probability.

We seek to design a universal sequential test to search the target (or to decide that it is missing). Precisely speaking, a test consists of a sequential search policy, a stopping rule \( \tau \) and a decision rule \( \delta \). The stopping rule defines a random stopping time, denoted by \( N \), which is the number of searches taken until the final decision is made. At the stopping time, the final decision for the target location is made based on the decision rule \( \delta : \mathcal{Y}^N \times [M]^N \to \{0,1,\ldots,M\} \), where the 0 output corresponds to the final decision for a missing target. The overall goal is to achieve a certain level of accuracy for the final decision using the fewest number of observations, universally for all \( \mu \neq \pi \).
A. Fundamental Performance Limit

When both $\mu$ and $\pi$ are known, the search and stop problem falls under the umbrella of sequential design of experiments with a uniform experimental cost [5]. In particular, there are $M + 1$ hypotheses: $0, 1, \ldots, M$, where the null (0-th) hypothesis corresponds to the possibility of a missing target. Each $i$-th hypothesis, $i = 1, \ldots, M$, corresponds to a possible location of the present target. The experiment set corresponds to $\mathcal{U} = [M]$; and the model $p_i(y), i = 0, \ldots, M$, for sequential design of experiments can be identified as

$$
p_i^u = \mu, \quad u = i, \quad p_i^u = \pi, \quad u \neq i, \quad i = 1, \ldots, M,
\quad p_0^u = \pi, \quad u = 1, \ldots, M. \quad \text{(III.1)}$$

Then in this idealistic situation when the probabilistic model (both $\mu$ and $\pi$) is known, by particularizing the characterization of the asymptotically optimal performance in Proposition II.1 to our search and stop problem using (III.1) and $c(i, u) = 1$, for each $i = 0, \ldots, M$, for each $u \in \mathcal{U}$, we get that as the error probability

$$P_{\text{max}} = \max_{i=0,\ldots,M} \mathbb{P}_i[\delta(Y^N, U^N) \neq i]$$

is driven to zero, the optimal asymptotes of $\mathbb{E}_i[N], \quad i = 0, \ldots, M$, can be characterized as follows.

**Proposition III.1.** There exists a sequence of tests to search the target that satisfy $P_{\text{max}} \to 0$ and yield

$$\mathbb{E}_i[N] = \begin{cases} 
\frac{-\log P_{\text{max}}}{2\log(M)} (1 + o(1)), & i = 0, \\
\frac{-\log P_{\text{max}}}{D(\mu||\pi)} (1 + o(1)), & i = 1, \ldots, M.
\end{cases} \quad \text{(III.2)}$$

Furthermore, the asymptotic performance in (III.2) (each term in the denominators) is optimal for every $i = 0, \ldots, M$ simultaneously.

Of course, the asymptotic performance in Proposition III.1 is idealistic, as it requires the knowledge of $\mu$ (with $\pi$ being already known). When $\mu$ is not known, since $\mu$ can be arbitrarily close to $\pi$, this asymptotic performance cannot be achieved universally. Nevertheless, our main contribution (Theorem IV.1) described below shows that one can design a universal test (without the knowledge of $\mu$) that drives the error probability to zero and achieves the optimal exponent of $D(\mu||\pi)$ under all the non-null hypotheses universally for any $\mu \neq \pi$.

IV. Proposed Universal Scheme for Search and Stop Performance

A. Motivation

Intuitively speaking, a desirable goal of the search at each location should be to determine if the target is there. To this end, the universal sequential test for two hypotheses in Subsection II-A can be used at each location to collect multiple subsequent observations that will eventually lead to a binary outcome (say 1 if it is guessed that the target is there, and 0 otherwise). To improve reliability for this binary decision at a particular search location, one can increase the threshold $\alpha$ in (II.4), (II.5), (II.6) with the cost of taking more observations at that location.

If we use the mentioned sequential binary test at each location as the “inner” test, then it is convenient to select the current search location based on the past binary outcomes of the subsequent binary tests (instead of the past $\mathcal{Y}$-ary outcomes of every search, generally taken multiple times at each of the locations). With this imposition, the search and stop problem can be reduced to a problem of constructing an “outer” test for the sequential design of such inner experiments, each of which has a binary outcome.

Mathematically speaking, we have reduced the original problem of sequential design of $\mathcal{Y}$-ary-output experiments specified by the (conditional) distributions as in (III.1) to one of sequential design of binary-output experiments specified as

$$\mu_b(0) = 1 - \mu_b(1) = \beta_a, \quad \pi_b(1) = 1 - \pi_b(0) = \alpha_a, \quad \text{(IV.1)}$$

and

$$p_i^u = \mu_b, \quad u = i, \quad p_i^u = \pi_b, \quad u \neq i, \quad i = 1, \ldots, M, \quad p_0^u = \pi_b, \quad u = 1, \ldots, M \quad \text{(IV.2)}$$

where $\alpha_a, \beta_a$ are as defined in (II.7), (II.10). On the other hand, each binary-output experiment will not have the same cost as for the original $\mathcal{Y}$-ary-output experiment. In particular, the cost of each binary-output experiment can be specified as

$$c(i, u) = c_a, \quad u = i, \quad c(i, u) = \kappa_a, \quad u \neq i, \quad i = 1, \ldots, M, \quad c(0, u) = \kappa_a, \quad u = 1, \ldots, M, \quad \text{(IV.3)}$$

where $c_a, \kappa_a$ are as defined in (II.8) and (II.9), respectively.

There is still a large gap in turning the motivation described above into a “working” test for search and stop. To this end, there are two major challenges. First, the optimal test for sequential design of experiments in (II.12), (II.13), (II.14), achieving the performance stated Proposition III.1, requires precise knowledge of the model. In contrast, the induced model for sequential design of binary-output experiments in (IV.1), (IV.2) is a complicated function of the inner threshold $a$ for the sequential binary test in (II.4), (II.5), (II.6). Only an estimate of this “true” induced model is available through the bounds for $\alpha_a, \beta_a$ stated in (II.7), (II.11) of Lemma II.1. Second, as the threshold $a'$ for the optimal test in (II.12), (II.13), (II.14) increases, the model for sequential design of experiments in Proposition III.1 remains fixed. In contrast, in our proposed test, the “outer” threshold for the test for sequential design of binary-output experiments increases together with the inner threshold $a$, the latter of which determines the induced model in (IV.1), (IV.2). Consequently, the analysis leading to Proposition III.1 does not apply to our proposed test. Our main technical contribution are precisely, first, to overcome these challenges through the proposed test described below in Subsection IV-B, employing an outer threshold, which is an appropriate function of the inter threshold, and, second, to provide the analysis for its performance, stated in Theorem IV.1 below.
B. Proposed Universal Test

As mentioned in the previous subsection, the “true” induced model for sequential design of the binary-output experiments in (IV.1), (IV.2) is a complicated function of the inner threshold \( \alpha \) and is not available to us. Nevertheless, Lemma II.1 yields that for \( \alpha \) sufficiently large (as a function of \( \nu \), in (II.8) and \( \mu, \pi \))

\[
\alpha_\alpha, \beta_\alpha \leq \frac{1}{\alpha}.
\]

Our idea would be to use a mismatched model defined in terms of \( \bar{\pi}_b, \bar{\pi}_b \), where

\[
\bar{\pi}_b(u) = 1 - \bar{\pi}_b(1) = \bar{\pi}_b(1) = 1 - \bar{\pi}_b(0) = \frac{1}{\alpha}
\]

to perform the sequential design of the binary-output experiments. Specifically, instead of (IV.1), (IV.2), consider the following mismatched model for sequential design of binary-output experiments

\[
\pi^i_{u} = \pi^i, \ u = i, \quad \pi^i_{u} = \pi^i, \ u \neq i, \ i = 1, \ldots, M,
\]

Heuristically speaking, by (IV.4), this mismatched model is “more noisy” than the true model (for large \( \alpha \)); hence, the test designed based on this mismatched model should be conservative enough to work well for the true model as well. This intuition will be proven to be correct.

With the mismatched model specified in (IV.6), we can now describe our universal test as follows. At each time \( t \geq 1 \), we compute the estimate of the true hypothesis \( \hat{i} \) based on past searched locations and their binary outcomes \( u^{t-1}, z^{t-1} \) using the (mismatched) model \( \pi^i_{u} \), \( i = 0, \ldots, M, u \in [M] \) in (IV.6). Denote \( N(i, 1), N(i, 0), \ i \in [M] \), as the number of times the \( i \)-th location was searched and the sequential binary test in (II.4), (II.5), (II.6) decides that the target is there, and that the target is not there, respectively. By the reciprocity of \( \pi^0 \) and \( \pi^b \), the computation of this estimate can be simplified as

\[
\hat{i} = \begin{cases} 
\arg\max_{i \in [M]} N(i, 1) - N(i, 0) & \text{if } \max_{i \in [M]} N(i, 1) - N(i, 0) > 0; \\
0 & \text{if } \max_{i \in [M]} N(i, 1) - N(i, 0) \leq 0.
\end{cases}
\]

The estimation in (IV.7) is quite intuitive, as the difference between the numbers of “searched-and-found” and “searched-and-not-found” at the \( i \)-th location: \( N(i, 1) - N(i, 0) \), \( i \in [M] \), should approximate the likelihood that the target is there. When all these numbers are negative, it is most likely that the target is missing.

For \( b > 0 \), during the sparse occasions \( t = [e^{k}b], \ k = 0, 1, \ldots, \) the experiment is selected to explore all locations in a round-robin manner as

\[
u_t = (k \mod M) + 1
\]

independently of \( \hat{i} \). At all the other times, if \( \hat{i} \neq 0 \), we shall search at the \( i \)-th location, i.e.,

\[
u_t = \hat{i}, \ \text{if } \hat{i} \neq 0.
\]

If \( \hat{i} = 0 \), we search among all locations with equal frequency, namely,

\[
u_t = (i_{t'} \mod M) + 1,
\]

where \( i_{t'} \) was the search location at the last time \( t' < t \) such that \( \hat{i} = 0 \). Denote the joint (mismatched) distribution under the \( i \)-th hypothesis of all binary searched outcomes up to time \( t \) (induced by the above control policy) by \( \bar{\pi}_i(z^t) \). The test stops at time \( \tau \) and decides in favor of the current estimate of the hypothesis as:

\[
\tau = \arg\min_{t \neq 0} \left[ \min_{j \neq \hat{i}} \left( \frac{\bar{\pi}_j(z^t)}{\bar{\pi}_i(z^t)} \right) > e^{a^2(\log a)^{\rho_1}} \right],
\]

where \( a \) is the inner threshold for the binary test in (II.4), (II.5), (II.6) and some \( \rho_2 > 1 \). As clarified at the end of the paragraph preceding Proposition II.1, since the search policy in (IV.8), (IV.9), (IV.10) specifies the search location as a deterministic function of the current estimate of the hypothesis at all times, it suffices to work with the joint distribution \( \bar{\pi}_i(z^t) \) instead of \( \bar{\pi}_i(z^t, u^t) \), i.e., \( u^t \) can be written as a deterministic function of \( z^t \).

Using (IV.5), (IV.6), we can simplify (IV.11) as

\[
\tau = \min (\tau', \tau_0),
\]

where

\[
\tau' = \arg\min_{t \neq 0} \left[ \min_{j \neq \hat{i}} \left( \frac{N(\hat{i}, 1) - N(\hat{i}, 0)}{N(j, 1) - N(j, 0)} \right) > \frac{a^{\rho_2} (\log a)^{\rho_1}}{\log (\alpha - 1)} \right],
\]

\[
\tau_0 = \arg\min_{t \neq 0} \left[ \min_{i \in [M]} (N(i, 0) - N(i, 1)) > \frac{a^{\rho_2} (\log a)^{\rho_1}}{\log (\alpha - 1)} \right].
\]

Of course, the Markov time in (IV.13) is reached first when \( \delta(z^\tau) = \hat{i} \neq 0 \), whereas the other Markov time in (IV.14) is reached first when \( \delta(z^\tau) = \hat{i} = 0 \).

Note that the total number of \( \mathcal{Y} \)-ary-output observations \( N \) used to produce the search result is related to the stopping time \( \tau \) above as

\[
N = \sum_{t=1}^{\tau} N_t^b,
\]
where each \( N^b_i, \ t = 1, \ldots, \tau, \) is the number of observations taken at each location until the sequential test in (II.4), (II.5), (II.6) produces a binary result \( Z_i. \) Consequently, we get from successive uses of the property of conditional expectation and (IV.3) that under the true hypothesis \( i = 0, \ldots, M, \) it holds that

\[
\mathbb{E}_i [N] = \mathbb{E}_i \left[ \sum_{t=1}^{\tau-1} N^b_t + \mathbb{E}_i \left[ N^b_\tau \mid Y \left( \sum_{t=1}^{\tau-1} N^b_t \right) \right] \right] = \mathbb{E}_i \left[ \sum_{t=1}^{\tau-1} N^b_t \right] + \mathbb{E}_i \left[ N^b_\tau \mid U_\tau \right] = \mathbb{E}_i \left[ \sum_{t=1}^{\tau-1} N^b_t \right] + c(i, U_\tau) = \mathbb{E}_i \left[ \sum_{t=1}^{\tau} c(i, U_t) \right].
\]

**C. Performance of Proposed Test**

**Theorem IV.1.** For any \( \nu < 1 \) in (II.8) and for \( b > 0 \) used in (IV.8) chosen to be sufficiently small, as \( a \to \infty, \) the test in (IV.7), (IV.8), (IV.9), (IV.10), (IV.11) yields a vanishing error probability \( P_{\text{max}} \to 0 \) and also satisfies

\[
\mathbb{E}_i [N] = \mathbb{E}_i \left[ \sum_{t=1}^{\tau-1} c(i, U_t) \right] \leq -\log P_{\text{max}} \nu D (\mu \parallel \pi) (1 + o(1)), \quad (IV.15)
\]

\( i = 1, \ldots, M, \) universally for every \( \mu \neq \pi. \)

**Remark IV.1.** Compared to the ideally optimal performance (when \( \mu \) is known) in Proposition III.1, it is interesting to note that our universal test is universally asymptotically optimal, except only when the target is missing. In other words, the knowledge of \( \pi \) is not useful in improving reliability for detecting the missing target. This consequence of our result is directly relevant in practical settings, wherein the knowledge of the target distribution \( \pi \) would be lacking before the target is found.

**D. Comparison with Universal Non-Adaptive Scheme for Search and Stop**

Our main result in Theorem IV.1 illustrates that one can construct a test with adaptive search policy, using only the knowledge of \( \pi, \) that yields a vanishing error probability and achieves the exponent of \( D (\mu \parallel \pi) \) universally for every \( \mu \neq \pi \) when the target is present. A natural question that arises is how much can be gained by employing such an adaptive search policy beyond a non-adaptive one. A non-adaptive search policy \( \overline{\phi} \) has to specify the sequence of search locations at the outset and cannot adapt to the outcomes of the instantaneous searches. By the symmetry of the problem, there is no reason for a non-adaptive search policy to favor any location. Consequently, the only non-adaptive search policy that should be considered in the universal setting is the one that search all locations with equal frequency:

\[
u_k = (k \mod M) + 1, \ k \geq 0. \quad (IV.16)
\]

We denote this non-adaptive search policy by \( \overline{\phi}^* \). With this search policy, an efficient universal test has been constructed in [13], which we now describe.

For each time \( k = \ell M, \ \ell = 1, 2, \ldots, \) let \( \gamma_i, \ i = 1, \ldots, M \) denote the empirical distribution of the observations when the \( i \)-th location is searched, namely, \( \gamma_i = \left( y_i, y_{M+i}, \ldots, y_{(\ell-1)M+i} \right), \ i = 1, \ldots, M. \) Next, denote the estimate of the target location \( \hat{i} \) as

\[
\hat{i} = \arg\max_{i \in [M]} D (\gamma_i \parallel \pi). \quad (IV.17)
\]

With the non-adaptive search policy \( \overline{\phi}^* \), consider the stopping rule defined in terms of the following Markov time:

\[
\mathbb{N}' \equiv M \times \arg\min_{\ell \geq 1} \left[ \left( D (\gamma_i \parallel \pi) - \max_{j \neq i} D (\gamma_j \parallel \pi) \right) \right] > \log \pi + M |Y| \log (\ell + 1). \quad (IV.18)
\]

The test stops at time \( \mathbb{N}', \) where

\[
\mathbb{N}' \equiv \min \left( \mathbb{N}', |\bar{\pi} \log \bar{\pi}| \right). \quad (IV.19)
\]

Correspondingly, the final decision is made according to

\[
\delta (y \mathbb{N}') = \left\{ \begin{array}{ll}
\hat{i} & \text{if } \mathbb{N}' \leq \pi \log \pi \\
0 & \text{if } \mathbb{N}' > \pi \log \pi.
\end{array} \right. \quad (IV.20)
\]

The performance of this test with the non-adaptive search scheme follows from the result in [13].

**Proposition IV.1 ([13]).** With the non-adaptive search policy \( \overline{\phi}^* \) in (IV.16), the test in (IV.17), (IV.18), (IV.19), (IV.20) yields a vanishing error probability \( P_{\text{max}} \to 0 \) and also satisfies

\[
\mathbb{E}_i [\mathbb{N}] \leq -\log P_{\text{max}} \frac{D (\mu \parallel \pi)}{M} (1 + o(1)), \ i = 1, \ldots, M,
\]

universally for every \( \mu \neq \pi. \)

In summary, adaptivity offers a multiplicative gain of \( M \) for search reliability beyond non-adaptive searching. This gain increases with the size of the area to be searched.

**Acknowledgements**

This work was supported by the Air Force Office of Scientific Research (AFOSR) under the Grant FA9550-10-1-0458 through the University of Illinois at Urbana-Champaign and by the U.S. Defense Threat Reduction Agency through subcontract 147755 at the University of Illinois from prime award HDTRA1-10-1-0086.

**References**


