

Fixed-Horizon Active Hypothesis Testing

Dhruva Kartik , Ashutosh Nayyar , Senior Member, IEEE, and Urbashi Mitra , Fellow, IEEE

Abstract—Two active hypothesis testing problems are formulated. In these problems, the agent can perform a fixed number of experiments and then decide on one of the hypotheses or declare its experiments inconclusive. The first problem is an asymmetric formulation in which the objective is to minimize the probability of incorrectly declaring a particular hypothesis to be true while ensuring that the probability of correctly declaring that hypothesis is moderately high. This formulation is a generalization of the formulation in the Chernoff–Stein lemma to an active setting. The second problem is a symmetric formulation in which the objective is to minimize the misclassification probability while ensuring that the true hypothesis is declared conclusively with moderately high probability. For these problems, lower and upper bounds on the optimal misclassification probabilities are derived and these bounds are shown to be asymptotically tight. Novel experiment selection strategies are provided. It is shown that these strategies are asymptotically optimal and using numerical experiments, it is demonstrated that these strategies have a significantly better performance in the nonasymptotic regime.

Index Terms—Anomaly detection, Chernoff–Stein lemma, controlled sensing, hypothesis testing.

I. INTRODUCTION

WE FREQUENTLY encounter scenarios wherein we would like to deduce whether one of several hypotheses is true by gathering data or evidence. This problem is referred to as multihypothesis testing. If we have access to multiple candidate experiments or data sources, we can adaptively select more informative experiments to infer the true hypothesis. This leads to a joint control and inference problem commonly referred to as *active* hypothesis testing. There are numerous ways of formulating this problem and the precise mathematical formulation depends on the target application.

Manuscript received November 13, 2019; revised January 27, 2021; accepted May 16, 2021. Date of publication June 21, 2021; date of current version March 29, 2022. This work was supported in part under Grants ONR N00014-15-1-2550, NSF CCF-1817200, ECCS 1750041, NSF CCF-2008927, and ARO W911NF1910269, in part by Cisco Foundation under Grant 1980393, in part under Grant ONR 503400-78050, in part by DOE under Grant DE-SC0021417, in part by Swedish Research Council under Grant 2018-04359, and in part by Okawa Foundation. This paper was presented in part at the 2019 IEEE International Symposium on Information Theory, Maison de la Mutualite, Paris, France, July 2019. Recommended by Associate Editor P. G. Mehta. (Corresponding author: Dhruva Kartik.)

The authors are with the Ming Hsieh Department of Electrical and Computer Engineering, University of Southern California, Los Angeles, CA 90089 USA (e-mail: mokhasun@usc.edu; ashutosn@usc.edu; ubli@usc.edu).

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/TAC.2021.3090742>.

Digital Object Identifier 10.1109/TAC.2021.3090742

In this article, we consider a scenario in which there is an agent that can perform a *fixed* number of experiments. Subsequently, the agent can decide on one of the hypotheses using the collected data. The agent is also allowed to declare the experiments *inconclusive* if needed. In this fixed-horizon setting, we consider two formulations. In the first formulation, we are interested in minimizing the probability of incorrectly declaring a *particular* hypothesis to be true while ensuring that the probability of correctly declaring the same hypothesis is moderately high. Thus, we would like to declare that this hypothesis is true only if we have very strong evidence supporting it. This formulation is intended for applications like anomaly detection wherein incorrectly declaring the system to be safe (i.e., anomaly-free) can be very expensive whereas a moderate number of false alarms can be tolerated. This formulation, and thus our results, can be viewed as a generalization of the classical Chernoff–Stein lemma [1] to an active multihypothesis testing setup.

In the second formulation, we are interested in minimizing the probability of making an *incorrect* inference (misclassification probability) while ensuring that the true hypothesis is declared conclusively with moderately high probability. The key difference between the first and second formulations is that the former is asymmetric, i.e., it focuses on reliably inferring a *particular* hypothesis, whereas the latter formulation is symmetric in the sense that it aims to avoid misclassifying every hypothesis. This symmetric formulation is of particular interest when the penalty for making *any* incorrect inference is significantly higher than the penalty for making no decision. In such cases, it is reasonable for the agent to abstain from drawing conclusions unless there is strong evidence supporting one of the hypotheses.

In both these problems, the agent can select experiments at each time in a data-driven manner. We refer to the strategy used for selecting these experiments as the *experiment selection strategy*. We refer to the strategy used by the agent to make an inference (or to declare its experiments inconclusive) based on all the data collected as the *inference strategy*. Thus, the two problems described above involve optimization over the space of inference and experiment selection strategy pairs.

Our contributions in this article pertaining to these hypothesis testing problems can be summarized as follows.

- 1) We find lower and upper bounds on the optimal misclassification probabilities in our constrained optimization problems. These bounds are asymptotically (w.r.t. the time-horizon) tight under some mild assumptions. Thus, we characterize the optimal misclassification error exponents in each problem.
- 2) We propose a novel approach for designing experiment selection strategies. Unlike the classical approach which

results in randomized and, in some cases, open-loop strategies, this approach allows us to design deterministic and adaptive experiment selection strategies that are asymptotically optimal.

- 3) We demonstrate numerically that the experiment selection strategies designed using our approach, when coupled with appropriate inference strategies, achieve superior nonasymptotic performance in comparison to the classical approaches.

The rest of the article is organized as follows. In Section I-A, we summarize key prior literature on hypothesis testing and discuss how our problem is related to various other formulations. In Section I-B, we describe our notation. We describe our system model in Sections II-A and II-B, we formulate our problems. We state the main results in Section II-C and sketch the proof of our results. We provide a detailed analysis of our hypothesis testing problems in Sections III and IV. In Section V, we discuss an anomaly detection example and provide the results of our numerical experiments. We conclude the article in Section VI.

A. Related Work

Hypothesis testing is a long-standing problem and has been addressed in various settings. Works that are closely related to active hypothesis testing can be broadly classified into the following paradigms.

1) Fixed-Horizon Hypothesis Testing: In the simplest fixed-horizon hypothesis testing setup, we have binary hypotheses and a single experiment. The inference is made based on a fixed number of i.i.d. observations obtained by repeatedly performing this experiment. In this setup, there are two popular formulations: 1) the Neyman–Pearson type asymmetric formulation used in the Chernoff–Stein lemma [1]; 2) the unconstrained symmetric formulation that involves minimizing the Bayesian error probability [1]. While our asymmetric formulation is a generalization of the Neyman–Pearson type formulation, our symmetric formulation is different from Bayesian error probability minimization in [1]. The key difference is that in Bayesian error minimization, the agent is not allowed to declare its experiments inconclusive at the end of the horizon and *must* declare one of the hypotheses to be true. More general works in this paradigm include [2]–[6]. All the aforementioned works are passive in the sense that there is only one experiment and thus, the experiment selection strategy is trivial. Nevertheless, we employ many of the analysis techniques developed in these works all of which are available in the form of lecture notes in [7].

An active fixed-horizon formulation has been considered in [8] in which the objective is to minimize the maximal error probability. This formulation is symmetric and does not allow the inconclusive declaration. Allowing the inconclusive declaration makes the nature of our analysis and results significantly different from the formulation in [8].

2) Sequential Hypothesis Testing: In sequential hypothesis testing, the time horizon is not fixed and the agent can continue to perform experiments until a stopping criterion is met. The objective then is to minimize the *Bayesian risk* which

is a combination of the expected stopping time and the error probability. Inspired by Wald’s sequential probability ratio test (SPRT) [9], Chernoff first addressed the problem of *active* sequential hypothesis testing in [10]. This work was later generalized in [8], [11], and [12]. These works focused on characterizing the asymptotic behavior¹ of the Bayesian risk in the regime where the cost of performing experiments is much lower than the cost of making an incorrect inference.

Our fixed-horizon formulations are most closely related to this sequential hypothesis testing framework. In both these settings, the optimal error rates and the strategies used to achieve them are very similar. Intuitively, this is because in both the sequential setting and our fixed-horizon setting, the agent conclusively declares a hypothesis to be true only if there is strong evidence supporting it. If strong enough evidence is not found, the agent in the sequential setting continues to perform experiments whereas the agent in our setting simply declares the experiments inconclusive. Fixed-horizon formulations are useful in applications with hard time constraints where the agent does not have the luxury to keep performing experiments until strong enough evidence is obtained. A key distinction between sequential and fixed-horizon formulations is that in the sequential setting, much of the analysis is focused on computing the expected stopping time. On the other hand, in our fixed-horizon setting, the stopping time is fixed but there are constraints on the probability of making conclusive correct inference. Proving that the designed strategies satisfy these constraints calls for a different set of mathematical tools.

In all the aforementioned works on the sequential setting, the experiment selection strategy has a randomized component. Although these randomized strategies are asymptotically optimal, their non-asymptotic performance may be poor. Deterministic strategies were proposed in [10] and [13] but in many cases, these strategies are *not* asymptotically optimal. In this article, we develop an approach that helps us in designing deterministic, adaptive and asymptotically optimal experiment selection strategies. Moreover, in some scenarios like anomaly detection, our deterministic strategies have a significantly better non-asymptotic performance. For some specific kinds of sequential hypothesis testing problems, deterministic strategies were designed and analyzed in [14] and [15]. Our strategy design is similar in spirit to these designs. While these works are more general than ours in certain respects such as allowing different experiment costs and infinite observation spaces, they make some other restrictive technical assumptions which are not required in our analysis. These assumptions include the *last* assumption in [14, p. 513] and in [15, Assumption 1.6]. Because of these assumptions, there are settings where our results apply while those of [14], [15] do not. In Section V-A6, we provide one simple example where this is the case. Some of the techniques used to design and analyze our strategy share some similarities with those used in multiarmed bandits [16].

3) Anomaly Detection: Many anomaly detection problems can be viewed as active hypothesis testing problems. In anomaly detection, there are multiple *normal* processes which exhibit a

¹In [12], certain nonasymptotic results are also provided.

certain kind of statistical behavior. Among these processes, there could be an anomaly with statistical characteristics distinct from the normal processes. There are various mechanisms to probe these processes and the objective is to reliably detect the anomaly as quickly as possible. Some recent works in anomaly detection include [17]–[21]. All these works are in the sequential setting. It has been noted in [17]–[19] that deterministic strategies achieve better performance. We believe that these deterministic strategies may be related to ours.

The problem of oddball detection has been considered in [22] and [23]. The approach used in [22] and [23] is similar to Chernoff's approach in [10] but the key innovation is that they do not assume knowledge of the underlying distributions.

4) Variable-Length Coding With Feedback: This problem is concerned with designing variable-length codes for discrete memoryless communication channels with perfect feedback [24], [25]. It can be viewed as a sequential hypothesis testing problem and a detailed discussion on this can be found in [25]. Our framework can be used to formulate a fixed-horizon analogue of the variable-length coding problem wherein the receiver is allowed to declare the transmission inconclusive if needed.

B. Notation

Random variables are denoted by upper case letters (X), their realization by the corresponding lower case letter (x). We use calligraphic fonts to denote sets ($\mathcal{U}$). The probability simplex over a finite set $\mathcal{U}$ is denoted by $\Delta\mathcal{U}$. In general, subscripts denote time indices unless stated otherwise. For time indices $n_1 \leq n_2$, $Y_{n_1:n_2}$ denotes the collection of variables ($Y_{n_1}, Y_{n_1+1}, \dots, Y_{n_2}$). For a strategy g , we use $\mathbb{P}^g[\cdot]$ and $\mathbb{E}^g[\cdot]$ to indicate that the probability and expectation depend on the choice of g . For an hypothesis i , $\mathbb{E}_i^g[\cdot]$ denotes the expectation conditioned on hypothesis i . For a random variable X and an event $\mathcal{E}$, $\mathbb{E}[X; \mathcal{E}]$ denotes $\mathbb{E}[X \mathbb{1}_{\mathcal{E}}]$, where $\mathbb{1}_{\mathcal{E}}$ is the indicator function associated with the event $\mathcal{E}$. The *cross-entropy* and *Kullback-Leibler divergence* between two distributions p and q over a finite space $\mathcal{Y}$ are respectively given by

$$H(p, q) = - \sum_{y \in \mathcal{Y}} p(y) \log q(y). \quad (1)$$

The *Kullback–Leibler divergence* between distributions p and q is given by

$$D(p||q) = \sum_{y \in \mathcal{Y}} p(y) \log \frac{p(y)}{q(y)}. \quad (2)$$

We use the convention that if $x \leq 0$, then $\log x \doteq -\infty$.

II. MINIMUM MISCLASSIFICATION ERROR PROBLEMS

In this section, we will formulate the two active hypothesis testing problems. We will describe our assumptions and state our main results on the asymptotic behavior of optimal misclassification probabilities.

A. System Model

Let $\mathcal{X} = \{0, 1, \dots, M-1\}$ be a finite set of hypotheses and let the random variable X denote the true hypothesis. The

prior probability on X is ρ_1 . Without loss of generality, let us assume that the distribution ρ_1 has full support. At each time $n = 1, 2, \dots$, an agent can perform an experiment $U_n \in \mathcal{U}$ and obtain an observation $Y_n \in \mathcal{Y}$. We assume that the sets $\mathcal{U}$ and $\mathcal{Y}$ are finite. The observation Y_n at time n is given by

$$Y_n = \xi(X, U_n, W_n) \quad (3)$$

where $\{W_n : n = 1, 2, \dots\}$ is a collection of mutually independent and identically distributed *primitive* random variables. The probability of observing y after performing an experiment u under hypothesis i is denoted by $p_i^u(y)$

$$p_i^u(y) \doteq \mathbb{P}(Y_n = y \mid X = i, U_n = u).$$

The time horizon, that is, the total number of experiments performed is fixed *a priori* to $N < \infty$.

At time $n = 1, 2, \dots, N$, the information available to the agent, denoted by I_n , is the collection of all experiments performed and the corresponding observations up to time $n-1$

$$I_n \doteq \{U_{1:n-1}, Y_{1:n-1}\}. \quad (4)$$

Let the collection of all possible realizations of information I_n be denoted by $\mathcal{I}_n$. At time n , the agent selects a distribution over the set of actions $\mathcal{U}$ according to an *experiment selection rule* $g_n : \mathcal{I}_n \rightarrow \Delta\mathcal{U}$ and the action U_n is randomly drawn from the distribution $g_n(I_n)$, i.e.,

$$U_n \sim g_n(I_n). \quad (5)$$

For a given experiment $u \in \mathcal{U}$ and information realization $\mathcal{J} \in \mathcal{I}_n$, the probability $\mathbb{P}^g[U_n = u \mid I_n = \mathcal{J}]$ is denoted by $g_n(\mathcal{J} : u)$. The sequence $\{g_n, n = 1, \dots, N\}$ is denoted by g and referred to as the *experiment selection strategy*. Let the collection of all such experiment selection strategies be $\mathcal{G}$.

After performing N experiments, the agent can declare one of the hypotheses to be true or it can declare that its experiments were inconclusive. We refer to this final declaration as the agent's *inference decision* and denote it by $\hat{X}_N$. Thus, the inference decision can take values in $\mathcal{X} \cup \{\mathbb{N}\}$, where $\mathbb{N}$ denotes the inconclusive declaration. Using the information I_{N+1} , the agent chooses a distribution over the set of hypotheses according to an *inference strategy* $f : \mathcal{I}_{N+1} \rightarrow \Delta(\mathcal{X} \cup \{\mathbb{N}\})$ and the inference $\hat{X}_N$ is drawn from the distribution $f(I_{N+1})$, i.e.,

$$\hat{X}_N \sim f(I_{N+1}). \quad (6)$$

For a given inference $\hat{x} \in \mathcal{X} \cup \{\mathbb{N}\}$ and information realization $\mathcal{J} \in \mathcal{I}_{N+1}$, the probability $\mathbb{P}^{f,g}[\hat{X}_N = \hat{x} \mid I_{N+1} = \mathcal{J}]$ is denoted by $f_N(\mathcal{J} : \hat{x})$. Let the set of all inference strategies be $\mathcal{F}$. For an experiment selection strategy g and an inference strategy f , we define the following probabilities.

Definition 1: For $i \in \mathcal{X}$, let $\psi_N(i)$ be the probability that the agent infers hypothesis i given that the true hypothesis is i , i.e.,

$$\psi_N(i) \doteq \mathbb{P}^{f,g}[\hat{X}_N = i \mid X = i]. \quad (7)$$

We refer to $\psi_N(i)$ as the *correct-inference probability* of type- i . Let $\phi_N(i)$ be the probability that the agent infers i given that the true hypothesis is not i , i.e.,

$$\phi_N(i) \doteq \mathbb{P}^{f,g}[\hat{X}_N = i \mid X \neq i]. \quad (8)$$

We refer to $\phi_N(i)$ as the *misclassification probability* of type- i .

We will also be interested in the event that the agent declares an incorrect hypothesis to be true. That is, we will consider

the event $\cup_{i \in \mathcal{X}} \{\hat{X}_N = i, X \neq i\}$. We refer to this event as the *misclassification event*.

Definition 2: Let γ_N be the probability of making an incorrect inference, i.e.,

$$\gamma_N \doteq \mathbb{P}^{f,g}[\cup_{i \in \mathcal{X}} \{\hat{X}_N = i, X \neq i\}]. \quad (9)$$

We will refer to γ_N as the *misclassification probability*.

Remark 1: Note that the misclassification probability γ_N can be expressed in terms of the misclassification probabilities $\phi_N(i)$ of type- i in the following manner:

$$\gamma_N = \sum_{i \in \mathcal{X}} \mathbb{P}^{f,g}[\hat{X}_N = i \mid X \neq i] \mathbb{P}[X \neq i] \quad (10)$$

$$= \sum_{i \in \mathcal{X}} \phi_N(i)(1 - \rho_1(i)). \quad (11)$$

B. Problem Formulation and Preliminaries

We will consider two active hypothesis testing formulations. The first one is an asymmetric formulation in which the focus is on a particular hypothesis i and involves minimizing the misclassification probability $\phi_N(i)$ of type- i . The second formulation is a symmetric formulation that involves minimizing the misclassification probability γ_N .

1) Asymmetric Formulation (P1): In this formulation, we are interested in designing an experiment selection strategy g and an inference strategy f that minimize the misclassification probability $\phi_N(i)$ of type- i subject to the constraint that the correct-inference probability $\psi_N(i)$ of type- i is sufficiently large. In other words, we would like to solve the following optimization problem:

$$\begin{aligned} \inf_{f \in \mathcal{F}, g \in \mathcal{G}} \quad & \phi_N(i) \\ \text{subject to} \quad & \psi_N(i) \geq 1 - \epsilon_N \end{aligned} \quad (P1)$$

where $0 < \epsilon_N < 1$. Let the infimum value of this optimization problem be $\phi_N^*(i)$. Note that this problem is always feasible because the agent can trivially satisfy the correct-inference probability constraint by always declaring hypothesis i . We refer to this problem as the *minimum misclassification error problem of type- i* or simply Problem (P1).

Remark 2: Problem (P1) can be seen as a binary hypothesis testing problem with null hypothesis $\{X = i\}$ and alternate hypothesis $\{X \neq i\}$. We observe that when there is only one experiment and two hypotheses, this formulation is identical to that of the Chernoff–Stein lemma in [1].

This formulation is helpful in modeling scenarios in which the cost of incorrectly declaring a particular hypothesis to be true is very high. For instance, consider a system which can potentially have various types of anomalies. We are interested in testing whether the system has no anomalies (hypothesis $X = i$) or has some anomaly (hypothesis $X \neq i$). In such systems, a few false alarms may be tolerable but declaring that the system is free of anomalies when there is one can be very expensive. Therefore, we would like to minimize the probability of falsely declaring the system to be free of anomalies subject to the constraint that the probability of raising false alarms is sufficiently small. Clearly,

this scenario can be modeled using the asymmetric formulation (P1).

2) Symmetric Formulation (P2): In this formulation, we are interested in designing an experiment selection strategy g and an inference strategy f that minimize the misclassification probability γ_N while satisfying the constraint that the correct-inference probability $\psi_N(i)$ of type- i is sufficiently large for every hypothesis $i \in \mathcal{X}$. In other words, we would like to solve the following optimization problem:

$$\begin{aligned} \min_{f \in \mathcal{F}, g \in \mathcal{G}} \quad & \gamma_N \\ \text{subject to} \quad & \psi_N(i) \geq 1 - \epsilon_N, \forall i \in \mathcal{X} \end{aligned} \quad (P2)$$

where $0 < \epsilon_N < 1$. Let γ_N^* denote the infimum value of this optimization problem. We define $\gamma_N^* \doteq \infty$ if the optimization problem is infeasible. We refer to this problem as the *minimum misclassification error problem* or simply Problem (P2).

The above formulation is intended for scenarios where the penalty for making an incorrect inference is much higher than the penalty for not making any inference. In such cases, it is reasonable for the agent to abstain from drawing conclusions when the evidence is not strong enough. The constraints on type- i correct-inference probabilities $\psi_N(i)$ ensure that the agent does not abstain from drawing conclusions too often. Thus, the optimization problem (P2) aims to find experiment selection and inference strategies that misclassify the least among all those strategies that make the correct inference with high probability.

Definition 3 (Log-likelihood ratio): For an experiment $u \in \mathcal{U}$ and any pair of hypotheses $i, j \in \mathcal{X}$ let

$$\lambda_j^i(u, y) \doteq \log \frac{p_i^u(y)}{p_j^u(y)} \quad (12)$$

be the log-likelihood ratio associated with an observation $y \in \mathcal{Y}$.

We make the following assumptions on our system model.

Assumption 1 (Common support): For any given experiment $u \in \mathcal{U}$, there exists a non-empty set $\mathcal{Y}(u) \subseteq \mathcal{Y}$ such that for every hypothesis $i \in \mathcal{X}$, the support of the distribution p_i^u is $\mathcal{Y}(u)$. In other words, for every hypothesis $i \in \mathcal{X}$, $p_i^u(y) > 0$ if and only if $y \in \mathcal{Y}(u)$.

Let $B > 0$ be a constant such that $|\lambda_j^i(u, y)| < B$ for every experiment $u \in \mathcal{U}$, observation $y \in \mathcal{Y}(u)$ and any pair of hypotheses $i, j \in \mathcal{X}$. Note that the existence of such a constant B is guaranteed because of Assumption 1.

Since our focus is on minimizing the incorrect inference probabilities, we make the following assumption on ϵ_N . This assumption captures the fact that we need the correct-inference probabilities to be large, but not necessarily too large.

Assumption 2: We have that the bound $1 - \epsilon_N$ on the type- i correct-inference probability satisfies $\epsilon_N \rightarrow 0$. Further, there exists a constant $b > 0$ such that $\lim_{N \rightarrow \infty} \epsilon_N N^b = \infty$.

The following assumption is standard in the active hypothesis testing literature and ensures that every pair of hypotheses can be distinguished by performing some experiment.

Assumption 3: For any pair of distinct hypotheses $i, j \in \mathcal{X}$, there exists an experiment $u \in \mathcal{U}$ such that $D(p_i^u || p_j^u) > 0$.

Before stating our main results, we will define some important quantities.

Definition 4 (Max–min Kullback–Leibler divergence): For each hypothesis $i \in \mathcal{X}$, define

$$D^*(i) \doteq \max_{\alpha \in \Delta \mathcal{U}} \min_{j \neq i} \sum_u \alpha(u) D(p_i^u || p_j^u) \quad (13)$$

$$= \min_{\beta \in \Delta \tilde{\mathcal{X}}_i} \max_{u \in \mathcal{U}} \sum_{j \neq i} \beta(j) D(p_i^u || p_j^u) \quad (14)$$

where $\tilde{\mathcal{X}}_i = \mathcal{X} \setminus \{i\}$. Note that α is a distribution over the set of experiments $\mathcal{U}$ and β is a distribution over the set of alternate hypotheses $\tilde{\mathcal{X}}_i$. The max–min Kullback–Leibler divergence $D^*(i)$ can be viewed as the value of a two-player zero-sum game [10], [26]. In this zero-sum game, the maximizing player selects a mixed strategy $\alpha \in \Delta \mathcal{U}$ and the minimizing player selects a mixed strategy $\beta \in \Delta \tilde{\mathcal{X}}_i$. The payoff associated with these strategies is

$$\sum_u \sum_{j \neq i} \alpha(u) D(p_i^u || p_j^u) \beta(j). \quad (15)$$

The equality of the min–max and max–min values follows from the minimax theorem [26] because the sets $\mathcal{U}$ and $\mathcal{X}$ are finite and the Kullback–Leibler divergences are bounded by B due to Assumption 1. Let the max–minimizer in (13) be denoted by α^{i*} and the min–maximizer in (14) be denoted by β^{i*} .

Definition 5 (Posterior belief): The posterior belief ρ_n on the hypothesis X based on information I_n is given by

$$\rho_n(i) = \mathbb{P}[X = i | U_{1:n-1}, Y_{1:n-1}] = \mathbb{P}[X = i | I_n]. \quad (16)$$

Note that given a realization of the experiments and observations until time n , the posterior belief does not depend on the experiment selection strategy g or the inference strategy f .

Definition 6 (Confidence level): For a hypothesis $i \in \mathcal{X}$ and a distribution ρ on $\mathcal{X}$ such that $0 < \rho(i) < 1$, the confidence level $C_i(\rho)$ associated with hypothesis i is defined as

$$C_i(\rho) \doteq \log \frac{\rho(i)}{1 - \rho(i)}. \quad (17)$$

The confidence level is the logarithm of the ratio of the probability (w.r.t. the distribution ρ) that hypothesis i is true versus the probability that hypothesis i is not true.

C. Main Results

We will now state our *three* main results on some asymptotic aspects of Problems (P1) and (P2).

1) Asymptotic Decay-Rate of Optimal Misclassification Probability $\phi_N^*(i)$ in Problem (P1): The following theorem can be viewed as a generalization of the classical Chernoff–Stein lemma [1] to the setting of active hypothesis testing. It states that the optimal value $\phi_N^*(i)$ in Problem (P1) decays exponentially with the horizon N and its asymptotic rate of decay is equal to the max–min Kullback–Leibler divergence $D^*(i)$ defined in Definition 4.

Theorem 1: The asymptotic rate of the optimal misclassification probability in Problem (P1) is given by

$$\lim_{N \rightarrow \infty} \frac{1}{N} \log \phi_N^*(i) = D^*(i). \quad (18)$$

Proof: To prove this result, we use the following approach. We first establish a lower bound on the misclassification probability $\phi_N(i)$ for any pair of experiment selection and inference

strategies (g, f) that satisfy the constraint $\psi_N(i) \geq 1 - \epsilon_N$. The asymptotic decay-rate of this lower bound is equal to the max–min Kullback–Leibler divergence $D^*(i)$. We then construct experiment selection and inference strategies that asymptotically achieve this rate. The details of the proof of this result are included in Section III. ■

The inference strategy constructed in the achievability proof of Theorem 1 has the following threshold structure:

$$f^N(\rho_{N+1} : i) = \begin{cases} 1 & \text{if } C_i(\rho_{N+1}) - C_i(\rho_1) \geq \theta_N \\ 0 & \text{otherwise} \end{cases} \quad (19)$$

where $\theta_N = ND^*(i) - o(N)$. The precise value of the threshold θ_N is provided in Appendix G. The experiment selection strategy used is as follows: at each time n , randomly select an experiment u from the max–min distribution α^{i*} as defined in Definition 4. This strategy design is motivated from the design in [10]. Note that this is a completely open-loop strategy, that is, the experiments are selected without using any of the information acquired in the past.

The open-loop randomized experiment selection strategy described above is asymptotically optimal for Problem (P1). However, in the nonasymptotic regime, there may be other experiment selection strategies that perform significantly better than this open-loop randomized strategy. For some specialized problems such as anomaly detection, it was observed in recent active hypothesis testing literature [17]–[19] that there exist deterministic and adaptive strategies that are asymptotically optimal and also outperform the classical Chernoff-type randomized strategies in the nonasymptotic regime. Even in [10], Chernoff proposed a deterministic and adaptive strategy (see [10, Sec. 7]) but he also presented a counter-example in which the strategy was *not* asymptotically optimal.

In this article, we propose a class of deterministic and adaptive strategies that are asymptotically optimal for Problem (P1). To the best of our knowledge, this is the first proof of asymptotic optimality of such strategies in a general active hypothesis testing setting. For a simple anomaly detection example, we also demonstrate numerically in Section V that our deterministic adaptive strategy (DAS) (when paired with an appropriate inference strategy) performs significantly better than the open-loop randomized strategy α^{i*} in the nonasymptotic regime.

2) Asymptotically Optimal Experiment Selection Strategies for Problem (P1): Let the moment generating function of the negative log-likelihood ratios $-\lambda_j^i(u, Y)$ for an experiment u be denoted by $\mu_j^i(u, s)$, i.e.,

$$\mu_j^i(u, s) \doteq \mathbb{E}_i[\exp(-s\lambda_j^i(u, Y))]. \quad (20)$$

Definition 7: For a given experiment $u \in \mathcal{U}$, belief $\rho \in \Delta \mathcal{X}$ and $0 \leq s \leq 1$, define

$$\mathcal{M}_i(u, \rho, s) \doteq \frac{\sum_{j \neq i} (\rho(j))^s \mu_j^i(u, s)}{\sum_{j \neq i} (\rho(j))^s}. \quad (21)$$

Using this definition, we will now define a class of experiment selection strategies.

Criterion 1: For a given time horizon N , consider an experiment selection strategy $g^N \doteq (g_1^N, g_2^N, \dots, g_N^N)$ such that at

each time n , $\alpha_n \doteq g_n^N(I_n) \in \Delta\mathcal{U}$ satisfies

$$\sum_{u \in \mathcal{U}} \alpha_n(u) \mathcal{M}_i(u, \rho_n, s_N) \leq \sum_{u \in \mathcal{U}} \alpha^{i*}(u) \mathcal{M}_i(u, \rho_n, s_N) \quad (22)$$

with probability 1. Here, ρ_n is the posterior belief at time n and

$$s_N \doteq \min \left\{ 1, \sqrt{\frac{2 \log \frac{M}{\epsilon_N}}{NB^2}} \right\}. \quad (23)$$

Observation 1: Criterion 1 captures three experiment selection strategies of interest. These are the following.

- 1) *Open-loop randomized strategy (ORS):* At any time n , $g_n^N(I_n) \doteq \alpha^{i*}$. This is the open-loop randomized strategy discussed earlier in Section II-C1.
- 2) *Deterministic adaptive strategy (DAS):* At each time n , $g_n^N(I_n)$ selects the experiment $u \in \mathcal{U}$ that minimizes $\mathcal{M}_i(u, \rho_n, s_N)$. This is a deterministic and adaptive strategy.
- 3) *DAS with restricted support (DAS-RS):* At each time n , $g_n^N(I_n)$ selects the experiment u from the *support* of α^{i*} that minimizes $\mathcal{M}_i(u, \rho_n, s_N)$. This is also a deterministic and adaptive strategy.

For all these experiment selection strategies, we have the following result.

Theorem 2: Let f^N be as defined in (19) and g^N be an experiment selection strategy that satisfies Criterion 1. Then the class of strategies $\{(f^N, g^N) : N \in \mathbb{N}\}$ is asymptotically optimal. In other words, if $\psi_N(i)$ and $\phi_N(i)$ are the correct-inference and misclassification probabilities associated with the strategy pair (f^N, g^N) , then $\psi_N(i) \geq 1 - \epsilon_N$ for every N and

$$\lim_{N \rightarrow \infty} -\frac{1}{N} \log \phi_N(i) = D^*(i). \quad (24)$$

Proof: See Section III-C.

Remark 3: Theorem 2 holds as long as the value of s_N used in Criterion 1 satisfies $s_N \rightarrow 0$ and $Ns_N \rightarrow \infty$. We chose the value of s_N that maximizes the value of θ_N defined in (126).

Remark 4 (Zero-sum game Interpretation): Note that selecting an experiment that minimizes $\mathcal{M}_i(u, \rho, s)$ over the set $\mathcal{U}$ is equivalent to selecting an experiment that maximizes $(1 - \mathcal{M}_i(u, \rho, s))/s$. When s is small, this function can be approximated as follows:

$$\frac{1 - \mathcal{M}_i(u, \rho, s)}{s} = \frac{\sum_{j \neq i} (\rho(j))^s (1 - \mu_j^i(u, s))}{s \sum_{j \neq i} (\rho(j))^s} \quad (25)$$

$$\approx \frac{\sum_{j \neq i} (\rho(j))^s D(p_i^u \| p_j^u)}{\sum_{j \neq i} (\rho(j))^s} \quad (26)$$

since $(1 - \mu_j^i(u, s))/s \rightarrow D(p_i^u \| p_j^u)$ as $s \rightarrow 0$. Thus, we can interpret the strategy DAS in terms of the zero-sum game discussed earlier in Section II-B after Definition 4. In the zero-sum game, if the minimizing player selects an alternate hypothesis j with probability $\beta(j) = (\rho(j))^s / \sum_{k \neq i} (\rho(k))^s$, then the strategy DAS selects an approximate best-response to the minimizing player's strategy with respect to the payoff function in (15).

Given a horizon N , the strategies DAS and DAS-RS described in this section are time-invariant functions of the posterior belief. However, they depend on the value of s_N and thus, on the horizon N of the problem. In some cases, these strategies turn out to be

independent of the value s_N which results in fully stationary (with respect to the posterior belief ρ_n) strategies. We show that this is the case in the example discussed in Section V. It may be possible to show that for such stationary strategies in [10, Lemma 2] holds and thus, they are asymptotically optimal even in the *sequential* formulation in [10].

3) Asymptotic Decay-Rate of Optimal Misclassification Probability γ_N^* in Problem (P2): Similar to the result in Theorem 1, we can characterize the decay-rate of the optimal misclassification probability γ_N^* as follows.

Theorem 3: The optimal misclassification probability γ_N^* in Problem (P2) decays exponentially with the horizon N and its asymptotic decay-rate is equal to $\min_{i \in \mathcal{X}} D^*(i)$, i.e.,

$$\lim_{N \rightarrow \infty} -\frac{1}{N} \log \gamma_N^* = \min_{i \in \mathcal{X}} D^*(i). \quad (27)$$

Proof: The methodology used for proving this result is very similar to that of Theorem 1. We first obtain a lower bound on the misclassification probability γ_N for any pair of experiment selection and inference strategies (g, f) that satisfy the constraints of Problem (P2). This lower bound is very closely related to the lower bounds established for Problem (P1). Then we construct a class of experiment selection and inference strategies that achieve this lower bound asymptotically. This class of experiment selection strategies includes the randomized strategy proposed in [10] and also, deterministic strategies similar to DAS and DAS-RS introduced in Section II-C2. The derivation of the lower bound and the construction of the experiment selection and inference strategies are discussed in detail in Section IV. ■

III. ANALYSIS OF PROBLEM (P1)

In this section, we analyze the asymmetric formulation (P1). To optimize the misclassification error of type- i in Problem (P1), we need to design both an experiment selection strategy g and an inference strategy f . We will first arbitrarily fix the experiment selection strategy g . For a fixed g , we derive a lower bound on the misclassification probability $\phi_N(i)$ associated with any inference strategy f that satisfies the constraint in Problem (P1). These bounds are obtained using the weak converse approach described in [7]. In these derivations, we will introduce some useful properties of the confidence level defined in Definition 6. Using these properties, we will then weaken the lower bounds to derive a bound that does not depend on the strategy g . Further, we will show that any experiment selection strategy satisfying Criterion 1 defined in Section II-C2, coupled with an appropriate inference strategy, can asymptotically achieve this strategy-independent lower bound. Finally, we will discuss some methods for obtaining better nonasymptotic lower bounds on the misclassification probability $\phi_N(i)$ using the strong converse theorem in [7].

A. Lower Bound for a Fixed Experiment Selection Strategy

In this section, we fix the experiment selection strategy to an arbitrary choice g and analyze the problem of optimizing the inference strategy for this particular experiment selection strategy. This analysis will help us in obtaining a lower bound

on the misclassification probability and in designing inference strategies for Problem (P1). Consider the following optimization problem:

$$\begin{aligned} \min_{f \in \mathcal{F}} \quad & \phi_N(i) \\ \text{subject to} \quad & \psi_N(i) \geq 1 - \epsilon_N. \end{aligned} \quad (\text{P3})$$

To analyze problem (P3), we will first define some useful quantities related to the confidence level in Definition 6.

For the hypothesis i and a strategy $g \in \mathcal{G}$, define the likelihood distributions $P_{N,i}^g$ and $Q_{N,i}^g$ over the set $\mathcal{I}_{N+1}$ as follows:

$$P_{N,i}^g(\mathcal{I}_{N+1}) \doteq \mathbb{P}^g[I_{N+1} = \mathcal{I}_{N+1} \mid X = i] \quad (28)$$

$$Q_{N,i}^g(\mathcal{I}_{N+1}) \doteq \mathbb{P}^g[I_{N+1} = \mathcal{I}_{N+1} \mid X \neq i]. \quad (29)$$

Proposition 1: Under any experiment selection strategy g , with probability 1, we have

$$\log \frac{P_{N,i}^g(I_{N+1})}{Q_{N,i}^g(I_{N+1})} = C_i(\rho_{N+1}) - C_i(\rho_1). \quad (30)$$

Proof: For any instance $\mathcal{I}_{N+1} \in \mathcal{I}_{N+1}$ such that $\mathbb{P}^g[I_{N+1} = \mathcal{I}_{N+1}] > 0$, we have the following using Bayes' rule:

$$\log \frac{P_{N,i}^g(\mathcal{I}_{N+1})}{Q_{N,i}^g(\mathcal{I}_{N+1})} = C_i(\rho_{N+1}) - C_i(\rho_1) \quad (31)$$

where ρ_{N+1} is the posterior belief associated with the instance $\mathcal{I}_{N+1}$. Note that because of Assumption 1, $\mathbb{P}^g[I_{N+1} = \mathcal{I}_{N+1}] > 0$ implies that both the numerator and the denominator in (31) are nonzero and thus, the expression in (31) is well-defined. Since (31) is true for every instance $\mathcal{I}_{N+1}$ with nonzero probability, we have our result. ■

Thus, the increment in confidence level is a log-likelihood ratio.

Definition 8 (Expected confidence rate): We define the expected confidence rate $J_N^g(i)$ as

$$J_N^g(i) \doteq \frac{1}{N} \mathbb{E}_i^g[C_i(\rho_{N+1}) - C_i(\rho_1)]. \quad (32)$$

Remark 5: Due to Proposition 1, the expected confidence rate is the averaged Kullback–Leibler divergence between the distributions $P_{N,i}^g$ and $Q_{N,i}^g$. That is

$$J_N^g(i) = \frac{1}{N} D(P_{N,i}^g \| Q_{N,i}^g).$$

When the experiment selection strategy is fixed, we can view the problem (P1) as a one-shot hypothesis testing problem in which we are trying to infer whether the collection of actions and observations I_{N+1} is drawn from distribution $P_{N,i}^g$ or $Q_{N,i}^g$. This interpretation allows us to use the classical results [1], [7] on one-shot hypothesis testing and derive various properties.

We will first obtain a lower bound on the misclassification probability $\phi_N(i)$ in Problem (P3) using the data-processing inequality of Kullback–Leibler divergences [7]. This is commonly known as the weak converse [5], [7].

Lemma 1 (Weak converse): Let g be any given experiment selection strategy. Then for any inference strategy f such that $\psi_N(i) \geq 1 - \epsilon_N$, we have

$$-\frac{1}{N} \log \phi_N(i) \leq \frac{J_N^g(i)}{1 - \epsilon_N} + \frac{\log 2}{N(1 - \epsilon_N)} \quad (33)$$

where $J_N^g(i)$ is the expected confidence rate. Therefore

$$-\frac{1}{N} \log \phi_N^*(i) \leq \frac{\sup_{g \in \mathcal{G}} J_N^g(i)}{1 - \epsilon_N} + \frac{\log 2}{N(1 - \epsilon_N)} \quad (34)$$

where $\phi_N^*(i)$ is the optimum value in Problem (P1).

Proof: See Appendix B. Note that this lemma is true for every $0 \leq \epsilon_N < 1$.

The bound (34) suggests that we can obtain a strategy-independent lower bound on ϕ_N^* by obtaining upper bounds on the quantity $\sup_{g \in \mathcal{G}} J_N^g(i)$. In the next section, we will focus on obtaining this upper bound.

B. Strategy-Independent Lower Bound

We will first describe some important properties of the confidence level $C_i(\rho)$ which will be used to derive a strategy-independent lower bound on the misclassification probability in problem (P1).

Definition 9: For a given experiment selection strategy g and an alternate hypothesis $j \neq i$, define the total log-likelihood ratio up to time n as

$$\mathbf{z}_n(j) \doteq \sum_{k=1}^n \lambda_j^i(\mathbf{u}_k, \mathbf{y}_k); \quad \bar{Z}_n \doteq \sum_{j \neq i} \beta^{i*}(j) Z_n(j),$$

where the log-likelihood ratio λ_j^i is as defined in equation (ref{llrdef}) and β^{i*} is the min-maximizing distribution in Definition.

Notice that the processes $Z_n(j)$ for each $j \neq i$ and $\bar{Z}_n$ are submartingales with respect to the filtration I_{n+1} when $X = i$. We will now establish the relationship between the total log-likelihood ratios $Z_n(j)$ and the confidence level C_i .

Lemma 2: For any experiment selection strategy g and for each $1 \leq n \leq N$, we have

$$\begin{aligned} C_i(\rho_{n+1}) - C_i(\rho_1) \\ = -\log \left[\sum_{j \neq i} \exp(\log \tilde{\rho}_1(j) - Z_n(j)) \right] \end{aligned} \quad (35)$$

where $\tilde{\rho}_1(j) = \rho_1(j)/(1 - \rho_1(i))$.

Proof: This is a consequence of simple algebraic manipulations. See Appendix C for details. ■

Note that for any vector z , we have $-\log \sum_j \exp(-z(j)) \approx \min_j z(j)$. Thus, the interpretation of this lemma is that the increment in confidence level $C_i(\rho_{n+1}) - C_i(\rho_1)$ approximately represents the smallest total log-likelihood ratio $\min_{j \neq i} Z_n(j)$. Therefore, this lemma can be seen as the first step towards establishing the relationship between the average expected increment in confidence J_N^g and the max–min Kullback–Leibler divergence $D^*(i)$. To formally establish this relationship, we use Lemma 2 to decompose the increment in confidence into a nonpositive cross-entropy term and a sub-martingale. This decomposition will be used in deriving strategy-independent lower bounds, both weak and strong, on the misclassification probability in Problem (P1).

Lemma 3 (Decomposition): For any experiment selection strategy g , we have

$$C_i(\rho_{n+1}) - C_i(\rho_1) = -H(\beta^{i*}, \tilde{\rho}_{n+1}) + \bar{Z}_n + H(\beta^{i*}, \tilde{\rho}_1).$$

Here, $\tilde{\rho}_n(j) = \rho_n(j)/(1 - \rho_n(i))$. As a result of the nonnegativity of cross entropy, we have

$$\mathcal{C}_i(\rho_{n+1}) - \mathcal{C}_i(\rho_1) \leq \bar{Z}_n + H(\beta^{i*}, \tilde{\rho}_1). \quad (36)$$

Proof: This is an algebraic consequence of Lemma 2. See Appendix D for details.

Using Lemma 3, we will now establish the relationship between the confidence rate $J_N^g(i)$ and the max–min Kullback–Leibler divergence $D^*(i)$ defined in (13). This, in conjunction with Lemma 1, will give us a strategy-independent lower bound on $\phi_N^*(i)$.

Lemma 4: For any experiment selection strategy g , we have

$$J_N^g(i) \leq D^*(i) + \frac{H(\beta^{i*}, \tilde{\rho}_1)}{N} \quad (37)$$

where $\tilde{\rho}_1(j) = \rho_1(j)/(1 - \rho_1(i))$. Further, using Lemma 1 and Assumption 2, we can conclude that

$$\limsup_{N \rightarrow \infty} -\frac{1}{N} \log \phi_N^*(i) \leq D^*(i). \quad (38)$$

Proof: See Appendix E.

C. Achievability of Decay-Rate $D^*(i)$ in Problem (P1)

We will now construct inference and experiment selection strategies that satisfy the constraint on hit probability $\psi_N(i)$ and asymptotically achieve misclassification decay-rate of $D^*(i)$. We will begin with the construction and analysis of the inference strategy.

The following is an upper bound on the misclassification probability associated with a deterministic confidence-threshold based inference strategy of the form discussed in Section II-C. Incidentally, this bound does not depend on the experiment selection strategy g . Similar upper bounds on error probabilities are commonly used in hypothesis testing [1], [10].

Lemma 5: Let f be a deterministic inference strategy in which hypothesis i is decided only if $\mathcal{C}_i(\rho_{N+1}) - \mathcal{C}_i(\rho_1) \geq \theta$. Then $\phi_N(i) \leq e^{-\theta}$.

Proof: See Appendix F.

The inference strategy f^N is constructed as follows:

$$f^N(\rho_{N+1} : i) = \begin{cases} 1 & \text{if } \mathcal{C}_i(\rho_{N+1}) - \mathcal{C}_i(\rho_1) \geq \theta_N \\ 0 & \text{otherwise} \end{cases} \quad (39)$$

where $\theta_N = ND^*(i) - o(N)$ and its precise value is provided in (142) in Appendix G. Due to Lemma 5, we can say that under the inference strategy f^N , any experiment selection strategy g achieves $\phi_N(i) \leq e^{-\theta_N}$. However, the inference strategy and the experiment selection strategy must also satisfy the constraint $\psi_N(i) \geq 1 - \epsilon_N$. In Section II-C2, we discussed experiment selection strategies that satisfy Criterion 1. Let g^N be any such experiment selection strategy that satisfies Criterion 1. Let the type- i correct-inference and misclassification probabilities associated with the strategy pair (g^N, f^N) be $\psi_N(i)$ and $\phi_N(i)$, respectively. We can show that this strategy pair satisfies the constraint $\psi_N(i) \geq 1 - \epsilon_N$. The proof is in Appendix G. Using the result (38), Lemma 5 and the fact that $\theta_N/N \rightarrow D^*(i)$ as $N \rightarrow \infty$, we can say that

$$\lim_{N \rightarrow \infty} \frac{1}{N} \log \frac{1}{\phi_N(i)} = D^*(i). \quad (40)$$

Thus, the experiment selection strategies that satisfy Criterion 1 (including ORS, DAS, and DAS-RS discussed in Section II-C2), when used in conjunction with the inference strategy f^N described above, are feasible solutions for the optimization problem (P1) and asymptotically achieve a type- i misclassification probability decay rate of $D^*(i)$. This concludes the proof of Theorem 2.

Since the inference strategy f^N and the experiment selection strategy g^N described above are feasible strategies with respect to the optimization problem in (P1), we can say that $\phi_N^*(i) \leq \phi_N(i) \leq e^{-\theta_N}$. And since under Assumption 2, $\theta_N/N \rightarrow D^*(i)$ as $N \rightarrow \infty$, we have

$$\liminf_{N \rightarrow \infty} -\frac{1}{N} \log \phi_N^*(i) \geq D^*(i). \quad (41)$$

Combining this result with the upper bound on the asymptotic decay rate of $\phi_N^*(i)$ in Lemma 4, we have Theorem 1.

Remark 6: We observe that our proof methodology shares some similarities with that of prior works [8], [10], [12] on active hypothesis testing. Nonetheless, many of our techniques are novel, especially those used to prove the asymptotic optimality of strategies that satisfy Criterion 1.

D. Tighter Nonasymptotic Lower Bounds

In this section, we will provide an alternate approach to finding lower bounds on the misclassification probability $\phi_N(i)$. This approach can be used to obtain tight lower bounds in some special cases. We will later illustrate this procedure with the help of an example.

For any given pair of inference and experiment selection strategies f, g that are feasible in Problem (P1), recall that the increment in confidence can be viewed as a log-likelihood ratio (30). Therefore for this strategy pair f, g , we have the following for every $\chi \in \mathbb{R}$:

$$-\log \phi_N(i) \quad (42)$$

$$\stackrel{a}{\geq} \chi - \log(\psi_N(i) - \mathbb{P}_i^g[\mathcal{C}_i(\rho_{N+1}) - \mathcal{C}_i(\rho_1) > \chi]) \quad (43)$$

$$\stackrel{b}{\geq} \chi - \log(1 - \epsilon_N - \mathbb{P}_i^g[\mathcal{C}_i(\rho_{N+1}) - \mathcal{C}_i(\rho_1) > \chi]) \quad (44)$$

$$= \chi - \log(\mathbb{P}_i^g[\mathcal{C}_i(\rho_{N+1}) - \mathcal{C}_i(\rho_1) \leq \chi] - \epsilon_N). \quad (45)$$

Inequality (a) is a consequence of the strong converse theorem in [7]. Inequality (b) holds because $\psi_N(i) \geq 1 - \epsilon_N$. However, much like the weak converse in Lemma 1, this lower bound on $\phi_N(i)$ depends on the experiment selection strategy g . We made use of the decomposition in Lemma 3 to obtain a strategy-independent lower bound in Lemma 4. We will follow a similar approach here. We have

$$\mathbb{P}_i^g[\mathcal{C}_i(\rho_{N+1}) - \mathcal{C}_i(\rho_1) \leq \chi] \quad (46)$$

$$\stackrel{a}{=} \mathbb{P}_i^g[-H(\beta^{i*}, \tilde{\rho}_{N+1}) + \bar{Z}_N + H(\beta^{i*}, \tilde{\rho}_1) \leq \chi] \quad (47)$$

$$\stackrel{b}{\geq} \mathbb{P}_i^g[\bar{Z}_N + H(\beta^{i*}, \tilde{\rho}_1) \leq \chi]. \quad (48)$$

Equality (a) is a consequence of Lemma 3, and since $H(\beta^{i*}, \tilde{\rho}_{N+1}) \geq 0$, we have that the event

$$\{-H(\beta^{i*}, \tilde{\rho}_{N+1}) + \bar{Z}_N + H(\beta^{i*}, \tilde{\rho}_1) \leq \chi\} \quad (49)$$

$$\supseteq \{\bar{Z}_N + H(\beta^{i*}, \tilde{\rho}_1) \leq \chi\} \quad (50)$$

which results in the inequality (b). Combining (45) and (48) leads us to the following lemma.

Lemma 6 (Stong converse): For any given pair of inference and experiment selection strategies f, g that are feasible in Problem (P1), we have for every $\chi \in \mathbb{R}$

$$-\log \phi_N(i) \leq \chi - \log(\mathbb{P}_i^g[\bar{Z}_N + H(\beta^{i*}, \tilde{\rho}_1) \leq \chi]) - \epsilon_N$$

with the convention that $\log x \doteq -\infty$ if $x \leq 0$.

Note that this lower bound is also dependent on the strategy g . However, it may be easier to derive strategy-independent lower bounds using the bound in Lemma 6. This is because the process $\bar{Z}_n - nD^*(i)$ is a super-martingale given $X = i$. In fact, if every experiment in $\mathcal{U}$ is in the support of α^{i*} (which is the case in many problems), then the process $\bar{Z}_n - nD^*(i)$ is a martingale given $X = i$. Thus, a lower bound on $\phi_N(i)$ may be obtained using a strategy-independent lower bound on the tail probability $\mathbb{P}_i^g[\bar{Z}_N + H(\beta^{i*}, \tilde{\rho}_1) \leq \chi]$ [28]. In some special cases, it may even occur that the evolution of the process $\bar{Z}_n$ is completely independent of the strategy g . We will discuss an example that satisfies this condition in Section V.

IV. ANALYSIS OF PROBLEM (P2)

In this section, we will analyze Problem (P2) and prove Theorem 3. Let f, g be inference and experiment selection strategies that satisfy the constraints in Problem (P2), i.e. $\psi_N(i) \geq 1 - \epsilon_N$ for every $i \in \mathcal{X}$, where $\psi_N(i)$ is the type- i correct-inference probability associated with strategies f, g . Let $\phi_N(i)$ be the type- i misclassification probability associated with strategies f, g . Since the strategy pair f, g satisfies $\psi_N(i) \geq 1 - \epsilon_N$, we can use Lemma 1 to obtain the following inequality:

$$-\frac{1}{N} \log \phi_N(i) \leq \frac{J_N^g(i) + \frac{\log 2}{N}}{1 - \epsilon_N} \quad (51)$$

for every $i \in \mathcal{X}$. Let γ_N be the misclassification probability associated with strategy pair f, g . Then, we have

$$\gamma_N = \sum_{i \in \mathcal{X}} (1 - \rho_1(i)) \phi_N(i) \quad (52)$$

$$\geq \sum_{i \in \mathcal{X}} (1 - \rho_1(i)) \exp\left(-\frac{(N J_N^g(i) + \log 2)}{1 - \epsilon_N}\right). \quad (53)$$

Rearranging the terms above, we have

$$-\frac{1}{N} \log \gamma_N \quad (54)$$

$$\leq -\frac{1}{N} \log \sum_{i \in \mathcal{X}} (1 - \rho_1(i)) e^{-\frac{(N J_N^g(i) + \log 2)}{1 - \epsilon_N}} \quad (55)$$

$$-\frac{1}{N} \log \sum_{i \in \mathcal{X}} e^{-\left(\frac{(N J_N^g(i) + \log 2)}{1 - \epsilon_N} + \log(1 - \rho_1(i))\right)} \quad (56)$$

$$\stackrel{a}{\leq} -\frac{1}{N} \max_{i \in \mathcal{X}} \left(-\frac{(N J_N^g(i) + \log 2)}{1 - \epsilon_N} + \log(1 - \rho_1(i)) \right) \quad (57)$$

$$= \min_{i \in \mathcal{X}} \left(\frac{J_N^g(i) + \frac{\log 2}{N}}{1 - \epsilon_N} - \frac{\log(1 - \rho_1(i))}{N} \right) \quad (58)$$

where inequality (a) is because $\log \sum_i \exp(x_i) \geq \max_i(x_i)$. Using the definition of γ_N^* , we have

$$-\frac{1}{N} \log \gamma_N^* \quad (59)$$

$$= \sup_{f, g: \psi_N(i) \geq 1 - \epsilon_N, i \in \mathcal{X}} \left[-\frac{1}{N} \log \gamma_N \right] \quad (60)$$

$$\stackrel{b}{\leq} \min_{i \in \mathcal{X}} \left(\frac{\sup_g J_N^g(i) + \frac{\log 2}{N}}{1 - \epsilon_N} - \frac{\log(1 - \rho_1(i))}{N} \right) \quad (61)$$

$$\stackrel{c}{\leq} \min_{i \in \mathcal{X}} \left(\frac{D^*(i) + \frac{H(\beta^{i*}, \tilde{\rho}_1) + \log 2}{N}}{1 - \epsilon_N} - \frac{\log(1 - \rho_1(i))}{N} \right). \quad (62)$$

Inequality (b) is due to the result in (58) and inequality (c) follows from Lemma 4. Thus, we can conclude that

$$\limsup_{N \rightarrow \infty} -\frac{1}{N} \log \gamma_N^* \leq \min_{i \in \mathcal{X}} D^*(i). \quad (63)$$

This establishes an upper bound on the asymptotic decay rate of the optimal misclassification probability γ_N^* in Problem (P2). We will now show that this rate is asymptotically achievable by constructing appropriate experiment selection and inference strategies.

For a given horizon N and for each $i \in \mathcal{X}$, let $g^{N,i}$ be an experiment selection strategy that satisfies Criterion 1 with respect to hypothesis i . Let the *maximum likelihood* (ML) estimate at time n be

$$\bar{i}_n \doteq \arg \max_{i \in \mathcal{X}} \bar{\rho}_n(i) \quad (64)$$

where ties are broken arbitrarily in the $\arg \max$ operator and $\bar{\rho}_n$ is the posterior belief at time n formed using uniform prior at time 1 instead of ρ_1 . Let $\mathcal{N} \doteq \{\lceil a^l \rceil : l \in \mathbb{N}, \lceil a^l \rceil \leq N\}$, where $a > 1$ is a constant and is defined in Appendix H.

1) Experiment Selection Strategy: If $n \in \mathcal{N}$, then the experiment U_n is selected randomly according to the uniform distribution over $\mathcal{U}$. If $n \notin \mathcal{N}$ and $\bar{i}_n = i$, then U_n is selected randomly with distribution $g^{N,i}(I_n)$. We denote this experiment selection strategy by $\bar{g}^N$ (or $\bar{g}$). Note that if $g^{N,i}$ is ORS, then the strategy $\bar{g}^N$ is identical to the one in [8].

2) Inference Strategy: Consider the following deterministic inference strategy $\bar{f}^N$ where for each $i \in \mathcal{X}$:

$$\bar{f}^N(\rho_{N+1} : i) = \begin{cases} 1 & \text{if } \mathcal{C}_i(\rho_{N+1}) - \mathcal{C}_i(\rho_1) \geq \theta_N(i) \\ 0 & \text{otherwise} \end{cases}$$

where $-\mathcal{C}_i(\rho_1) < \theta_N(i) = ND^*(i) - o(N)$ and is precisely defined in Appendix H. Notice that if $\bar{f}^N(\rho_{N+1} : i) = 0$ for every $i \in \mathcal{X}$, then $\bar{f}^N(\rho_{N+1} : \mathbb{N}) = 1$.

Since $\theta_N(i) > 0$ for every $i \in \mathcal{X}$, the threshold condition in $\bar{f}^N$ can be satisfied by at most one hypothesis. Thus, $\bar{f}^N$ declares hypothesis i if and only if the confidence increment associated with i exceeds the threshold $\theta_N(i)$. Hence, for each hypothesis $i \in \mathcal{X}$, the inference strategy $\bar{f}^N$ admits the structure required for Lemma 5 and thus, using Lemma 5, we can conclude that for each hypothesis $i \in \mathcal{X}$, $\phi_N(i) \leq e^{-\theta_N(i)}$. Therefore, under the strategies $(\bar{f}^N, \bar{g}^N)$, we have

$$\gamma_N \leq \sum_{i \in \mathcal{X}} (1 - \rho_1(i)) \exp(-\theta_N(i)). \quad (65)$$

TABLE I
CONDITIONAL PROBABILITIES $\mathbb{P}[Y = 1 | X, U]$: IN OUR NUMERICAL
EXPERIMENTS $\nu = 0.6$ WHICH INDICATES THAT THE OBSERVATIONS FROM
THESE SENSORS ARE VERY NOISY

	$X = 0$	$X = 1$	$X = 2$
$U = \mathcal{A}$	$1 - \nu$	ν	$1 - \nu$
$U = \mathcal{B}$	$1 - \nu$	$1 - \nu$	ν

In Appendix H, we show that there exists an integer $\bar{N}$ such that for every $N \geq \bar{N}$, the strategy pair $(\bar{f}^N, \bar{g}^N)$ also satisfies all the type- i correct-inference probability constraints in problem (P2). Thus, for every $N \geq \bar{N}$, we have

$$\gamma_N^* \leq \gamma_N \leq \sum_{i \in \mathcal{X}} (1 - \rho_1(i)) \exp(-\theta_N(i))$$

and hence

$$\liminf_{N \rightarrow \infty} -\frac{1}{N} \log \gamma_N^* \quad (66)$$

$$\geq \lim_{N \rightarrow \infty} -\frac{1}{N} \log \sum_{i \in \mathcal{X}} (1 - \rho_1(i)) \exp(-\theta_N(i)) \quad (67)$$

$$\geq \lim_{N \rightarrow \infty} -\frac{1}{N} \log \left(M \max_{i \in \mathcal{X}} \{ (1 - \rho_1(i)) \exp(-\theta_N(i)) \} \right)$$

$$= \lim_{N \rightarrow \infty} -\frac{1}{N} \left(\max_{i \in \mathcal{X}} \{ \log((1 - \rho_1(i)) \exp(-\theta_N(i))) \} \right) \quad (68)$$

$$= \min_{i \in \mathcal{X}} \left\{ \lim_{N \rightarrow \infty} \frac{\theta_N(i) - \log(1 - \rho_1(i))}{N} \right\} \quad (69)$$

$$= \min_{i \in \mathcal{X}} D^*(i). \quad (70)$$

Using the results (63) and (66), we can conclude that

$$\lim_{N \rightarrow \infty} -\frac{1}{N} \log \gamma_N^* = \min_{i \in \mathcal{X}} D^*(i). \quad (71)$$

This concludes the proof of Theorem 3.

V. EXAMPLE: ANOMALY DETECTION

Consider a system with two sensors $\mathcal{A}$ and $\mathcal{B}$. These sensors can detect an anomaly in the system in their proximity. The system state X can take three values $\{0, 1, 2\}$ where $X = 0$ indicates that the system is safe, i.e., there is no anomaly in the system. On the other hand, $X = 1$ indicates that there is an anomaly near sensor $\mathcal{A}$ and $X = 2$ indicates that there is an anomaly near sensor $\mathcal{B}$. The prior belief ρ_1 is uniform over the set $\{0, 1, 2\}$. There is a controller that can activate one of these sensors at each time to obtain an observation. Thus, the collection of actions that the controller can select from is $\mathcal{U} = \{\mathcal{A}, \mathcal{B}\}$. The observations are binary, i.e., $\mathcal{Y} = \{0, 1\}$. The conditional probabilities $\mathbb{P}[Y = 1 | X, U]$ associated with the observations given various states and actions are given in Table I.

1) Formulation: After collecting N observations from these sensors, we are interested in determining whether the system is safe or unsafe. We consider a setting where incorrectly declaring the system to be safe can be very expensive while a few false alarms can be tolerated. In this setting, the inconclusive decision $\aleph$ is treated as an alarm. Therefore, we would like to design an experiment (sensor) selection strategy g and an inference

strategy f that minimize the probability $\phi_N(0)$ of incorrectly declaring the system to be safe subject to the condition that the probability $\psi_N(0)$ of correctly declaring the system to be safe is sufficiently high. This can be formulated as

$$\begin{aligned} & \min_{f \in \mathcal{F}, g \in \mathcal{G}} \phi_N(0) \\ & \text{subject to} \quad \psi_N(0) \geq 1 - \epsilon_N \end{aligned} \quad (P4)$$

where $\epsilon_N = \min\{0.05, 10/N\}$ in our numerical experiments. Notice that this fits the formulation of Problem (P1).

2) Asymptotically Optimal Rate and Weak Converse:

Using Theorem 1, we can conclude that the asymptotically optimal misclassification rate is

$$D^*(0) = \max_{\alpha \in \Delta \mathcal{U}} \min_{j \neq 0} \sum_u \alpha(u) D(p_0^u || p_j^u) = (\nu - 1/2) \log \frac{\nu}{1 - \nu}.$$

The max-minimizer $\alpha^{0*}(\mathcal{A}) = \alpha^{0*}(\mathcal{B}) = 0.5$ and the min-maximizer $\beta^{0*}(1) = \beta^{0*}(2) = 0.5$. For convenience, we will refer to α^{0*} as α^* and β^{0*} as β^* . Also, notice that the bound on the log-likelihood ratios $B = \log \frac{\nu}{1 - \nu}$, $\nu > 0.5$. Therefore using the weak converse in Lemma 1 and Lemma 4, we have

$$\frac{1}{N} \log \frac{1}{\phi_N^*(0)} \leq \left(\frac{\nu - 1/2}{1 - \epsilon_N} \right) \log \frac{\nu}{1 - \nu} + \frac{2 \log 2}{N(1 - \epsilon_N)}.$$

3) Strong Converse: We will now use the lower bound in Lemma 6 to derive a strategy-independent lower bound on $\phi_N(0)$. Define

$$L_n \doteq \beta^*(1) \lambda_1^0(U_n, Y_n) + \beta^*(2) \lambda_2^0(U_n, Y_n). \quad (72)$$

Notice that given $X = 0$, for *either* experiment $u \in \mathcal{U}$, we have

$$\beta^*(1) \lambda_1^0(u, Y) + \beta^*(2) \lambda_2^0(u, Y) = \begin{cases} \frac{1}{2} \log \frac{\nu}{1 - \nu} & \text{w.p. } \nu \\ \frac{1}{2} \log \frac{1 - \nu}{\nu} & \text{w.p. } 1 - \nu. \end{cases}$$

Let the moment generating function of the variable above be $\bar{\mu}(s)$. Therefore, for any strategy g , we have

$$\mathbb{E}_0^g \left[\exp \left(\sum_{k=1}^n s_k L_k \right) \right]$$

$$= \mathbb{E}_0^g \left[\mathbb{E}_0 \exp \left(\sum_{k=1}^n s_k L_k \right) \mid I_n \right] \quad (73)$$

$$= \mathbb{E}_0^g \left[\exp \left(\sum_{k=1}^{n-1} s_k L_k \right) \mathbb{E}_0 [\exp(s_n L_n) \mid I_n] \right]$$

$$= \mathbb{E}_0^g \left[\exp \left(\sum_{k=1}^{n-1} s_k L_k \right) \right] \bar{\mu}(s_n) = \Pi_{k=1}^n \bar{\mu}(s_k). \quad (74)$$

Thus, L_n is an i.i.d. sequence and the process $\bar{Z}_n$ is the sum of these i.i.d. random variables. We can also exploit the fact that the observations are binary valued. Let the number of 0's in N observations be K . Then K has binomial distribution with parameters N, ν . Further

$$\bar{Z}_N \leq \chi - \log 2 \iff \left(K - \frac{N}{2} \right) \log \frac{\nu}{1 - \nu} \leq \chi - \log 2. \quad (75)$$

Define $\chi^* = (\mathcal{Q}(2\epsilon_N) - N/2) \log \frac{\nu}{1 - \nu} + \log 2$, where $\mathcal{Q}$ is the quantile function of the binomial distribution with parameters

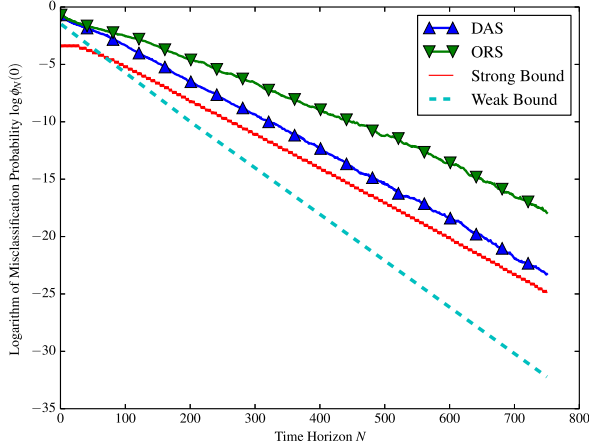


Fig. 1. Plot depicts the performance of strategies ORS and DAS. Both are asymptotically optimal but DAS is better in the nonasymptotic regime. When the horizon $N = 500$, we see a 13 dB improvement in the misclassification probability with DAS. Also notice that the strong bound on misclassification probability is very close to the performance of DAS.

N, ν . Using the relation (75), we can conclude that

$$\mathbb{P}_0^g[\bar{Z}_N + H(\beta^{i*}, \bar{\rho}_1) \leq \chi^*] \geq 2\epsilon_N \quad (76)$$

under any experiment selection strategy g . Finally, using Lemma 6, we have for any pair of inference and experiment selections strategies that satisfy the constraints in Problem (P4)

$$\log \frac{1}{\phi_N(0)} \leq \chi^* - \log(\epsilon_N). \quad (77)$$

4) Deterministic Adaptive Strategy: If the sensor $\mathcal{A}$ is selected, then the log-likelihood ratio $\lambda_2^0(\mathcal{A}, Y)$ is identically 0. Therefore, $\mu_2^0(\mathcal{A}, s) = 1$ for every s . Further, it can easily be verified that $\mu_1^0(\mathcal{A}, s) = 1$ at $s = 0$ and $s = 1$. This fact combined with the convexity of $\mu_1^0(\mathcal{A}, s)$ implies that $\mu_1^0(\mathcal{A}, s) \leq 1$ for every $0 \leq s \leq 1$. Similarly, $\mu_1^0(\mathcal{B}, s) = 1$ for every s and $\mu_2^0(\mathcal{B}, s) \leq 1$ for every $0 \leq s \leq 1$. Because of this, the deterministic adaptive experiment selection strategy (DAS) in Section II-C2 reduces to the following:

$$g_n(\rho_n : \mathcal{A}) = \begin{cases} 1 & \text{if } \rho_n(1) \geq \rho_n(2) \\ 0 & \text{otherwise.} \end{cases} \quad (78)$$

From Theorem 2, we know this strategy is asymptotically optimal. Note that the strategy described above is independent of the time-horizon N . This strategy happens to coincide with the one designed in [14], [15].

5) Numerical Results: We compare the performance of the open-loop randomized strategy α^{0*} and the DAS described above in Fig. 1. We observe that the performance of DAS is significantly better in the nonasymptotic regime. We also plot the weak and strong bounds established earlier. We observe that the strong bound is very close to the performance of DAS. In our numerical experiments, the inference strategy is a confidence threshold based strategy. Instead of computing the threshold, we empirically find the best threshold using binary search.

6) Chernoff's Deterministic Strategy: In [10], Chernoff described a fully deterministic strategy for *sequential* hypothesis testing but gave an example scenario for which his strategy was

TABLE II
CONDITIONAL PROBABILITIES $\mathbb{P}[Y = 1 | X, U]$ FOR THE PROBLEM SETUP IN SECTION V-8.

	$X = 0$	$X = 1$	$X = 2$
$U = \mathcal{A}$	0.400	0.600	0.400
$U = \mathcal{B}$	0.400	0.400	0.600
$U = \mathcal{C}$	0.402	0.598	0.280
$U = \mathcal{D}$	0.402	0.280	0.598

not asymptotically optimal [10, Sec. 7]. We will now demonstrate that even in such pathological scenarios, our strategies designed based on Criterion 1 are asymptotically optimal and also, have a better performance in the nonasymptotic regime.

Consider the same anomaly detection setup and formulation as in Section V-A1 with two additional experiments $\mathcal{C}$ and $\mathcal{D}$. The conditional distributions of the observations associated with these experiments are provided in Table II. Chernoff's approach to deterministic strategy design is as follows. Let $\bar{j}_n$ be the most-likely *alternate* hypothesis at time n , i.e.,

$$\bar{j}_n \doteq \arg \max_{j \neq 0} \bar{\rho}_n(j) \quad (79)$$

where $\bar{\rho}_n$ is the posterior belief at time n formed using a uniform prior. Then at time n , perform the experiment that maximizes $D(p_0^u || p_{\bar{j}_n}^u)$. In other words, select the experiment that can best distinguish the most-likely alternate hypothesis from the hypothesis of interest (in this case, it is $X = 0$). In our setup, Chernoff's deterministic strategy reduces to the following:

$$U_n = \begin{cases} \mathcal{A} & \text{if } \rho_n(1) \geq \rho_n(2) \\ \mathcal{B} & \text{otherwise.} \end{cases} \quad (80)$$

It can be shown that this strategy is *not* asymptotically optimal. Chernoff's randomized strategy (ORS) in this case reduces to selecting either experiment $\mathcal{C}$ or $\mathcal{D}$ with probability 0.5.

On the other hand, with some simple calculations, we can show that our strategy DAS-RS described in Section II-C2 reduces to

$$U_n = \begin{cases} \mathcal{C} & \text{if } \rho_n(1) \geq \rho_n(2) \\ \mathcal{D} & \text{otherwise.} \end{cases} \quad (81)$$

Because of Theorem 2, we know that both ORS and DAS-RS are asymptotically optimal. The performances of ORS, DAS-RS and Chernoff's deterministic strategy are shown in Fig. 2. In this case, the strategy DAS-RS does not depend on the time-horizon N whereas the strategy DAS depends on N . We note that the *last* assumption in [14, p. 513] and in [15, Assumption 1.6] do not apply for the above example.

VI. CONCLUSION

We formulated two fixed horizon active hypothesis testing problems (asymmetric and symmetric) in which the agent can decide on one of the hypotheses or declare its experiments inconclusive. Using information-theoretic techniques, we obtained lower bounds on optimal misclassification probability in these problems. We also derived upper bounds by constructing appropriate strategies and analyzing their performance. We proposed

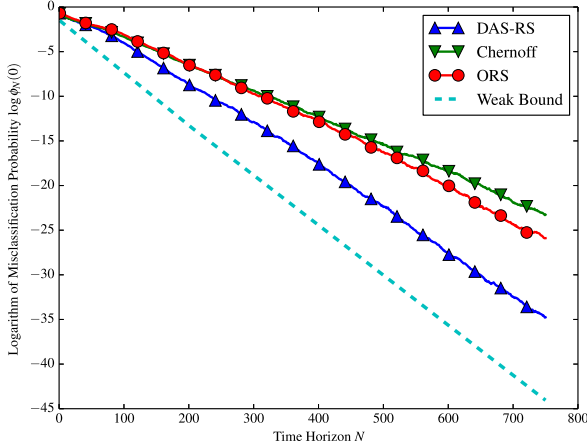


Fig. 2. Type-0 misclassification probability associated with strategies DAS-RS, ORS, and Chernoff's deterministic strategy. Both ORS and DAS-RS are asymptotically optimal and Chernoff's deterministic strategy is not. Notice that DAS-RS outperforms all the other strategies in the nonasymptotic regime.

a novel approach to designing deterministic and adaptive strategies for these active hypothesis testing problems. We proved that these deterministic strategies are asymptotically optimal, and through numerical experiments, demonstrated that they have significantly better performance in the nonasymptotic regime in some problems of interest.

In this article, our analysis was restricted to the setting where the observation space is finite. We believe that our analysis can be extended to cases where the observation space is infinite and the log-likelihood ratios are sub-Gaussian. We hope to address this extension in future work. Another direction for future work is to consider potentially infinite composite hypothesis spaces which arise in a variety of problems like oddball detection [22] and drug efficacy testing [10].

APPENDIX A AUXILIARY RESULTS

The proposition and its corollary stated below are consequences of simple algebraic manipulations.

Proposition 2: Let $i \in \mathcal{X}$ be a hypothesis. For any $j \neq i$ and at each time n , let $\tilde{\rho}_n(j) = \rho_n(j)/(1 - \rho_n(i))$. Then for any experiment selection strategy g , we have

$$\tilde{\rho}_n(j) = \frac{e^{(\log \tilde{\rho}_1(j) - Z_{n-1}(j))}}{\sum_{k \neq i} e^{(\log \tilde{\rho}_1(k) - Z_{n-1}(k))}} \quad (82)$$

with probability 1. Here, $Z_n(j)$ is the total log-likelihood ratio defined in Definition 9.

Corollary 1: Under the setting of Proposition 2, we have for each $0 \leq s \leq 1$ and $j \neq i$

$$\frac{(\rho_n(j))^s}{\sum_{k \neq i} (\rho_n(k))^s} = \frac{e^{(s \log \tilde{\rho}_1(j) - s Z_{n-1}(j))}}{\sum_{k \neq i} e^{(s \log \tilde{\rho}_1(k) - s Z_{n-1}(k))}}. \quad (83)$$

Lemma 7: Let $i, j \in \mathcal{X}$. If $0 \leq s \leq 1$, then under any strategy g and for every n , we have

$$\mathbb{E}_i^g[\exp(s \log \tilde{\rho}_1(j) - s Z_n(j))] \leq 1. \quad (84)$$

Proof: We have

$$\begin{aligned} & \mathbb{E}_i^g[\exp(s \log \tilde{\rho}_1(j) - s Z_n(j))] \\ &= \mathbb{E}_i^g[\exp(s \log \tilde{\rho}_1(j) - s Z_{n-1}(j))] \mathbb{E}_i^g[\mu_j^i(U_n, s) | I_n] \\ &\stackrel{a}{\leq} \mathbb{E}_i^g[\exp(s \log \tilde{\rho}_1(j) - s Z_{n-1}(j))] \\ &\stackrel{b}{\leq} \mathbb{E}_i^g[\exp(s \log \tilde{\rho}_1(j))] \leq 1. \end{aligned}$$

Inequality (a) is because for any experiment u , $\mu_j^i(u, s)$ is convex and $\mu_j^i(u, 0) = \mu_j^i(u, 1) = 1$. Inequality (b) is obtained by inductively applying the same arguments.

APPENDIX B PROOF OF LEMMA 1

Define a random variable $X^\dagger$ as follows: $X^\dagger \doteq 1$ if $\hat{X}_N = i$ and $X^\dagger \doteq 0$ otherwise. Thus, we have

$$\psi_N(i) = \mathbb{P}^{f,g}[X^\dagger = 1 | X = i] \quad (85)$$

$$\phi_N(i) = \mathbb{P}^{f,g}[X^\dagger = 1 | X \neq i]. \quad (86)$$

In this proof, let us denote $\psi_N(i)$ with ψ and $\phi_N(i)$ with ϕ for convenience. Notice that under strategies g and f , the variables X, I_{N+1} and $X^\dagger$ form a Markov chain. That is

$$\mathbb{P}^{f,g}[X^\dagger = 1 | X, I_{N+1}] = \mathbb{P}^{f,g}[X^\dagger = 1 | I_{N+1}]. \quad (87)$$

Therefore, using the data-processing property of relative entropy [5], [7], we can conclude that

$$\begin{aligned} D(P_{N,i}^g || Q_{N,i}^g) &\geq \psi \log \frac{\psi}{\phi} + (1 - \psi) \log \frac{1 - \psi}{1 - \phi} \\ &\stackrel{a}{\geq} -\psi \log \phi + \psi \log \psi + (1 - \psi) \log (1 - \psi) \\ &\stackrel{b}{\geq} -\psi \log \phi - \log 2 \\ &\stackrel{c}{\geq} -(1 - \epsilon_N) \log \phi - \log 2. \end{aligned}$$

Inequality (a) follows from the fact that $1 - \phi \leq 1$. Inequality (b) holds because $-\psi \log \psi - (1 - \psi) \log (1 - \psi)$ is a binary entropy and can at most be $\log 2$. Inequality (c) follows from our assertion that $\psi_N(i) \geq 1 - \epsilon_N$. Therefore, we have

$$-\frac{1}{N} \log \phi_N(i) \leq \frac{J_N^g(i)}{1 - \epsilon_N} + \frac{\log 2}{N(1 - \epsilon_N)}. \quad (88)$$

APPENDIX C PROOF OF LEMMA 2

We have

$$\log \frac{\rho_{n+1}(i)}{1 - \rho_{n+1}(i)} - \log \frac{\rho_1(i)}{1 - \rho_1(i)} \quad (89)$$

$$\stackrel{a}{=} \log \frac{\rho_1(i) \prod_{m=1}^n p_i^{U_m}(Y_m)}{\sum_{j \neq i} \rho_1(j) \prod_{m=1}^n p_j^{U_m}(Y_m)} - \log \frac{\rho_1(i)}{1 - \rho_1(i)} \quad (90)$$

$$= \log \frac{\prod_{m=1}^n p_i^{U_m}(Y_m)}{\sum_{j \neq i} \tilde{\rho}_1(j) \prod_{m=1}^n p_j^{U_m}(Y_m)} \quad (91)$$

$$= -\log \sum_{j \neq i} \exp \left(\log \tilde{\rho}_1(j) + \sum_{m=1}^n \lambda_i^j(U_m, Y_m) \right)$$

$$= -\log \left[\sum_{j \neq i} \exp(\log \tilde{\rho}_1(j) - Z_n(j)) \right]. \quad (92)$$

Equality (a) follows from the fact that the observation Y_m is independent of the past $I_m = \{U_{1:m-1}, Y_{1:m-1}\}$ conditioned on the hypothesis X and the current experiment U_m .

APPENDIX D PROOF OF LEMMA 3

Using the definition of cross-entropy, we have

$$\begin{aligned} -H(\beta^{i*}, \tilde{\rho}_{n+1}) &= \sum_{j \neq i} \beta^{i*}(j) \log \tilde{\rho}_{n+1}(j) \\ &\stackrel{a}{=} \sum_{j \neq i} \beta^{i*}(j) \log \left(\frac{\tilde{\rho}_1(j) e^{-Z_n(j)}}{\sum_{k \neq i} \tilde{\rho}_1(k) e^{-Z_n(k)}} \right) \\ &= -H(\beta^{i*}, \tilde{\rho}_1) - \bar{Z}_n - \log \left[\sum_{k \neq i} \exp(\log \tilde{\rho}_1(k) - Z_n(k)) \right] \\ &\stackrel{b}{=} -H(\beta^{i*}, \tilde{\rho}_1) - \bar{Z}_n + C_i(\rho_{n+1}) - C_i(\rho_1). \end{aligned}$$

Equality (a) follows from Proposition 2 in Appendix A and equality (b) is a consequence of Lemma 2.

APPENDIX E PROOF OF LEMMA 4

Using the definition of expected confidence rate, we have

$$\begin{aligned} J_N^g(i) &= \frac{1}{N} \mathbb{E}_i^g [C_i(\rho_{N+1}) - C_i(\rho_1)] \\ &\stackrel{a}{\leq} \frac{H(\beta^{i*}, \tilde{\rho}_1)}{N} + \frac{1}{N} \mathbb{E}_i^g \left[\sum_{j \neq i} \beta^{i*}(j) \sum_{n=1}^N \lambda_j^i(U_n, Y_n) \right] \\ &\stackrel{b}{=} \frac{H(\beta^{i*}, \tilde{\rho}_1)}{N} + \frac{1}{N} \mathbb{E}_i^g \left[\sum_{j \neq i} \beta^{i*}(j) \sum_{n=1}^N D(p_i^{U_n} \| p_j^{U_n}) \right] \\ &= \frac{H(\beta^{i*}, \tilde{\rho}_1)}{N} + \frac{1}{N} \mathbb{E}^g \left[\sum_{n=1}^N \sum_{j \neq i} \beta^{i*}(j) D(p_i^{U_n} \| p_j^{U_n}) \right] \\ &\stackrel{c}{\leq} \frac{H(\beta^{i*}, \tilde{\rho}_1)}{N} + \frac{1}{N} \mathbb{E}_i^g \left[\sum_{n=1}^N D^*(i) \right] \\ &= \frac{H(\beta^{i*}, \tilde{\rho}_1)}{N} + D^*(i). \end{aligned} \quad (93)$$

Equality (a) is a consequence of Lemma 3. Equality (b) follows from the fact that

$$\mathbb{E}_i^g \sum_{n=1}^N \lambda_j^i(U_n, Y_n) = \mathbb{E}_i^g \sum_{n=1}^N \mathbb{E}_i[\lambda_j^i(U_n, Y_n) | U_n] \quad (94)$$

$$= \mathbb{E}_i^g \sum_{n=1}^N D(p_i^{U_n} \| p_j^{U_n}). \quad (95)$$

Inequality (c) follows from the definition of the min-max distribution β^{i*} . Combining inequalities (93) and (34) from Lemma 1 gives us (38).

APPENDIX F PROOF OF LEMMA 5

Let $\mathcal{A}_{N+1}$ be the region in which the inference policy f described in Lemma 5 selects hypothesis i , that is

$$\mathcal{A}_{N+1} := \{\mathcal{J} : f(\mathcal{J} : i) = 1, \mathbb{P}^{f,g}[I_{N+1} = \mathcal{J}] \neq 0\}.$$

We have

$$\begin{aligned} \mathbb{P}^{f,g}[\hat{X}_{N+1} = i, X \neq i] &= \mathbb{P}^g[I_{N+1} \in \mathcal{A}_{N+1}, X \neq i] \\ &= \sum_{\mathcal{J} \in \mathcal{A}_{N+1}} \mathbb{P}^g[I_{N+1} = \mathcal{J}, X = i] e^{\left[-\log \frac{\mathbb{P}^g[I_{N+1} = \mathcal{J}, X = i]}{\mathbb{P}^g[I_{N+1} = \mathcal{J}, X \neq i]} \right]} \\ &\stackrel{a}{=} \sum_{\mathcal{J} \in \mathcal{A}_{N+1}} \mathbb{P}^g[I_{N+1} = \mathcal{J}, X = i] \exp[-C_i(\rho)] \end{aligned} \quad (96)$$

$$\stackrel{b}{\leq} \sum_{\mathcal{J} \in \mathcal{A}_{N+1}} \mathbb{P}^g[I_{N+1} = \mathcal{J}, X = i] \exp[-(\theta + C_i(\rho_1))] \quad (97)$$

$$\stackrel{c}{\leq} \rho_1(i) e^{-(\theta + C_i(\rho_1))} = (1 - \rho_1(i)) e^{-\theta}. \quad (98)$$

In equality (a), ρ is the posterior belief on X given information $\mathcal{J}$. Equality (a) follows from the definition of confidence level. Inequality (b) follows from the fact that $C_i(\rho) \geq \theta + C_i(\rho_1)$ for every $\mathcal{J} \in \mathcal{A}_{N+1}$. And inequality (c) is simply because $\mathbb{P}^g[I_{N+1} = \mathcal{J}, X = i] \leq \mathbb{P}[X = i]$. Therefore

$$\phi_N(i) = \mathbb{P}^{f,g}[\hat{X}_{N+1} = i | X \neq i] \leq e^{-\theta}. \quad (99)$$

APPENDIX G PROOF OF THEOREM 2

Let us fix the horizon N . In this proof, we will drop the superscript from g^N and simply refer to it as g for convenience. Since the inference strategy has a threshold structure [see (19)], proving that $\psi_N(i) \geq 1 - \epsilon_N$ is equivalent to proving that

$$\mathbb{P}_i^g[C_i(\rho_{N+1}) - C_i(\rho_1) < \theta_N] \leq \epsilon_N.$$

To this end, we will begin with obtaining upper bounds on the moment-generating function (MGF) of the confidence increment and then obtain a Chernoff bound based on these upper bounds.

1) Confidence Level and Log-Likelihood Ratios: Let $0 < s \leq 1$. Using Lemma 2, we have

$$\begin{aligned} \mathbb{E}_i^g \exp[-s(C_i(\rho_{N+1}) - C_i(\rho_1))] \\ = \mathbb{E}_i^g \left(\sum_{j \neq i} \exp(\log \tilde{\rho}_1(j) - Z_N(j)) \right)^s \end{aligned} \quad (100)$$

$$= \mathbb{E}_i^g \left(\sum_{j \neq i} (\exp(s \log \tilde{\rho}_1(j) - s Z_N(j)))^{1/s} \right)^s$$

$$\stackrel{a}{\leq} \mathbb{E}_i^g \left[\sum_{j \neq i} \exp(s \log \tilde{\rho}_1(j) - s Z_N(j)) \right] \quad (101)$$

$$= \sum_{j \neq i} \mathbb{E}_i^g \exp(s \log \tilde{\rho}_1(j) - sZ_N(j)). \quad (102)$$

Inequality (a) holds because $\|\cdot\|_{1/s} \leq \|\cdot\|_1$. Inequality (102) provides an upper bound on the MGF of the confidence increment in terms of the MGFs of the log-likelihood ratios.

2) Open-Loop Randomized Experiment Selection At Time $n+1$: Let $n < N$. Consider a scenario in which all the experiments upto time n are selected using strategy g but the experiment U_{n+1} is randomly selected using the distribution α^{i*} instead of using the strategy g . Under this modified strategy (say $\tilde{g}$), for some alternate hypothesis $j \neq i$, we have

$$\mathbb{E}_i^{\tilde{g}} \exp(s \log \tilde{\rho}_1(j) - sZ_{n+1}(j)) \quad (103)$$

$$\begin{aligned} &= \mathbb{E}_i^g[\mathbb{E}_i^{\tilde{g}}[\exp(s \log \tilde{\rho}_1(j) - sZ_{n+1}(j)) \mid I_{n+1}]] \\ &= \mathbb{E}_i^g[\exp(s \log \tilde{\rho}_1(j) - sZ_n(j)) \\ &\quad \times \mathbb{E}_i^{\tilde{g}}[\exp(-s\lambda_j^i(U_{n+1}, Y_{n+1})) \mid I_{n+1}]]. \end{aligned} \quad (104)$$

Let us analyze the term $\mathbb{E}_i^{\tilde{g}}[\exp(-s\lambda_j^i(U_{n+1}, Y_{n+1})) \mid I_{n+1}]$ in (110) separately. Since the observation Y_{n+1} is conditionally independent of I_{n+1} given the experiment U_{n+1} and the true hypothesis X [see (3)], we have

$$\begin{aligned} &\mathbb{E}_i^{\tilde{g}}[\exp(-s\lambda_j^i(U_{n+1}, Y_{n+1})) \mid I_{n+1}] \\ &= \sum_u \alpha^{i*}(u) \mathbb{E}_i[\exp(-s\lambda_j^i(u, Y))] = \sum_u \alpha^{i*}(u) \mu_j^i(u, s). \end{aligned} \quad (105)$$

Also, under the same strategy α^{i*} , the random variable $\lambda_j^i(U_{n+1}, Y_{n+1})$ has mean $D_j^i \doteq \sum_u \alpha^{i*}(u) D(p_j^u \parallel p_j^i)$ and is bounded by B with probability 1 because of Assumption 1. Bounded variables are sub-Gaussian (see Hoeffding's lemma in [29]), and thus, we have

$$\begin{aligned} &\mathbb{E}_i^\alpha[\exp(-s\lambda_j^i(U_{n+1}, Y_{n+1})) \mid I_{n+1}] \\ &\leq \exp(-sD_j^i + s^2B^2/2) \end{aligned} \quad (106)$$

$$\leq \exp(-sD^*(i) + s^2B^2/2) \quad (107)$$

where the last inequality follows from the fact that $D_j^i \geq D^*(i)$ for all $j \neq i$. Combining the results (105) and (107), we can say that when U_{n+1} is selected using α^{i*}

$$\begin{aligned} &\sum_{j \neq i} \exp(s \log \tilde{\rho}_1(j) - sZ_n(j)) \\ &\quad \times \mathbb{E}_i^{\tilde{g}}[\exp(-s\lambda_j^i(U_{n+1}, Y_{n+1})) \mid I_{n+1}] \end{aligned} \quad (108)$$

$$= \sum_{j \neq i} \exp(s \log \tilde{\rho}_1(j) - sZ_n(j)) \sum_u \alpha^{i*}(u) \mu_j^i(u, s) \quad (109)$$

$$\leq \sum_{j \neq i} \exp(s \log \tilde{\rho}_1(j) - sZ_n(j)) \exp(-sD^*(i) + s^2B^2/2). \quad (110)$$

We will use the result in (110) to inductively obtain a bound on the MGF of the confidence increment under the strategy g .

3) Inductive Step: Having established the results (104) and (110), we will now prove the following key Lemma.

Lemma 8: For any experiment selection strategy g that satisfies Criterion 1, we have

$$\sum_{j \neq i} \mathbb{E}_i^g \exp(s \log \tilde{\rho}_1(j) - sZ_{n+1}(j)) \quad (111)$$

$$\leq \left(\sum_{j \neq i} \mathbb{E}_i^g \exp(s \log \tilde{\rho}_1(j) - sZ_n(j)) \right) \quad (112)$$

$$\times \exp(-sD^*(i) + s^2B^2/2). \quad (113)$$

Proof: Since g satisfies Criterion 1, using Corollary 1 in Appendix A, we have

$$\begin{aligned} &\frac{\sum_{j \neq i} \sum_u (\rho_{n+1}(j))^s \alpha_{n+1}(u) \mu_j^i(u, s)}{\sum_{j \neq i} (\rho_{n+1}(j))^s} \\ &\leq \frac{\sum_{j \neq i} \sum_u (\rho_{n+1}(j))^s \alpha^{i*}(u) \mu_j^i(u, s)}{\sum_{j \neq i} (\rho_{n+1}(j))^s} \\ &\iff \sum_{j \neq i} \exp(s \log \tilde{\rho}_1(j) - sZ_n(j)) \sum_u \alpha_{n+1}(u) \mu_j^i(u, s) \\ &\leq \sum_{j \neq i} \exp(s \log \tilde{\rho}_1(j) - sZ_n(j)) \sum_u \alpha^{i*}(u) \mu_j^i(u, s). \end{aligned} \quad (114)$$

Recall that $\alpha_{n+1} \in \Delta\mathcal{U}$ is the distribution selected by the strategy g at time $n+1$. Thus, we have

$$\sum_{j \neq i} \mathbb{E}_i^g \exp(s \log \tilde{\rho}_1(j) - sZ_{n+1}(j)) \quad (115)$$

$$= \sum_{j \neq i} \mathbb{E}_i^g[\mathbb{E}_i^g[\exp(s \log \tilde{\rho}_1(j) - sZ_{n+1}(j)) \mid I_{n+1}]] \quad (116)$$

$$= \sum_{j \neq i} \mathbb{E}_i^g[\exp(s \log \tilde{\rho}_1(j) - sZ_n(j)) \quad (117)$$

$$\times \mathbb{E}_i^g[\exp(-s\lambda_j^i(U_{n+1}, Y_{n+1})) \mid I_{n+1}]]$$

$$\stackrel{a}{=} \mathbb{E}_i^g \sum_{j \neq i} \exp(s \log \tilde{\rho}_1(j) - sZ_n(j)) \sum_u \alpha_{n+1}(u) \mu_j^i(u, s)$$

$$\stackrel{b}{\leq} \mathbb{E}_i^g \sum_{j \neq i} \exp(s \log \tilde{\rho}_1(j) - sZ_n(j)) \sum_u \alpha^{i*}(u) \mu_j^i(u, s)$$

$$\stackrel{c}{\leq} \mathbb{E}_i^g \sum_{j \neq i} \exp(s \log \tilde{\rho}_1(j) - sZ_n(j)) \times \exp(-sD^*(i) + s^2B^2/2) \quad (118)$$

where equality (a) follows from the fact that the observation Y_{n+1} is conditionally independent of I_{n+1} given α_{n+1} . Inequality (b) follows from the result in (114) and inequality (c) is a consequence of the result (110).

Using the result in (102) and applying Lemma 8 inductively, we have

$$\mathbb{E}_i^g \exp[-s(\mathcal{C}_i(\rho_{N+1}) - \mathcal{C}_i(\rho_1))] \quad (119)$$

$$\leq \sum_{j \neq i} (\tilde{\rho}_1(j))^s \exp(-sND^*(i) + s^2NB^2/2) \quad (120)$$

$$\leq M \exp(-sND^*(i) + s^2NB^2/2). \quad (121)$$

4) Chernoff Bound: We can use the Chernoff bound [27] to conclude that

$$\mathbb{P}_i^g[\mathcal{C}_i(\rho_{N+1}) - \mathcal{C}_i(\rho_1) < \theta] \quad (122)$$

$$\leq \mathbb{E}_i^g \exp[-s(\mathcal{C}_i(\rho_{N+1}) - \mathcal{C}_i(\rho_1) - \theta)] \quad (123)$$

$$\stackrel{a}{\leq} M \exp(s\theta - sND^*(i) + s^2NB^2/2) \quad (124)$$

$$\stackrel{b}{=} \epsilon_N. \quad (125)$$

Inequality (a) follows from the result in (121). Equality (b) is obtained by substituting $\theta = \theta_N$ and $s = s_N$ where

$$\theta_N \doteq ND^*(i) - \frac{s_N NB^2}{2} - \frac{1}{s_N} \log \frac{M}{\epsilon_N} \quad (126)$$

and s_N is as defined in (23). Under Assumptions 2, one can easily verify that $\theta_N/N \rightarrow D^*(i)$ as $N \rightarrow \infty$. Thus, we have shown that for the strategy pair (f^N, g^N) , $\psi_N(i) \geq 1 - \epsilon_N$.

APPENDIX H

FEASIBILITY OF STRATEGY IN SECTION IV-A1 FOR PROBLEM (P2)

Let T be the smallest time index such that ML hypothesis $\bar{i}_n = X$ for every $n \geq T$. That is

$$T = \min\{n' : \bar{i}_n = X \quad \forall n \geq n'\}. \quad (127)$$

Under Assumption 3, it was shown in [8, Appendix B-2] that for a constant a sufficiently close to 1, there exists a constant $K > 0$ such that for every $i \in \mathcal{X}$, we have $\mathbb{P}_i^g[T > n] \leq Kn^{-b}$. Let

$$N' \doteq \left\lceil \left(\frac{2K}{\epsilon_N} \right)^{1/b} \right\rceil. \quad (128)$$

This ensures that $\mathbb{P}_i^g[T > N'] \leq \epsilon_N/2$. Fix a hypothesis i . Define the following event for each $n \geq N'$

$$\mathcal{Z}_n = \{\bar{i}_k = X, N' \leq k \leq n\}. \quad (129)$$

Clearly, the events $\mathcal{Z}_n$ are decreasing with n . Also, we have $\{T \leq N'\} \subseteq \mathcal{Z}_n$, for every $n \geq N'$.

Due to the threshold structure of the inference strategy $\bar{f}^N$, proving that $\psi_N(i) \geq 1 - \epsilon_N$ is equivalent to showing that

$$\mathbb{P}_i^g[\mathcal{C}_i(\rho_{N+1}) - \mathcal{C}_i(\rho_1) < \theta_N(i)] \leq \epsilon_N.$$

To do so, we will use a Chernoff-bound based approach similar to the approach in Appendix G. We have

$$\mathbb{P}_i^g[\mathcal{C}_i(\rho_{N+1}) - \mathcal{C}_i(\rho_1) < \theta_N(i)] \quad (130)$$

$$= \mathbb{P}_i^g[\mathcal{C}_i(\rho_{N+1}) - \mathcal{C}_i(\rho_1) < \theta_N(i), T > N'] \quad (131)$$

$$+ \mathbb{P}_i^g[\mathcal{C}_i(\rho_{N+1}) - \mathcal{C}_i(\rho_1) < \theta_N(i), T \leq N'] \quad (132)$$

$$\leq \epsilon_N/2 + \mathbb{P}_i^g[\mathcal{C}_i(\rho_{N+1}) - \mathcal{C}_i(\rho_1) < \theta_N(i), T \leq N'] \quad (133)$$

where the last inequality follows from the definition of N' .

1) Bounds on the MGF of Confidence Increment: For some $0 \leq s \leq 1$, define

$$\tau_n(s) \doteq \begin{cases} 1 & \text{if } n \in \mathcal{N} \\ \exp(-sD^*(i) + s^2B^2/2) & \text{otherwise.} \end{cases} \quad (134)$$

Consider the following:

$$\mathbb{E}_i^g \exp[-s(\mathcal{C}_i(\rho_{N+1}) - \mathcal{C}_i(\rho_1)); T \leq N'] \quad (135)$$

$$\stackrel{a}{\leq} \sum_{j \neq i} \mathbb{E}_i^g[\exp(s \log \tilde{\rho}_1(j) - sZ_N(j)); T \leq N'] \quad (136)$$

$$\stackrel{b}{\leq} \sum_{j \neq i} \mathbb{E}_i^g[\exp(s \log \tilde{\rho}_1(j) - sZ_N(j)); \mathcal{Z}_N] \quad (137)$$

$$\stackrel{c}{\leq} \sum_{j \neq i} \mathbb{E}_i^g[\exp(s \log \tilde{\rho}_1(j) - sZ_{N-1}(j)); \mathcal{Z}_N] \tau_N(s) \quad (138)$$

$$\stackrel{d}{\leq} \sum_{j \neq i} \mathbb{E}_i^g[\exp(s \log \tilde{\rho}_1(j) - sZ_{N-1}(j)); \mathcal{Z}_{N-1}] \tau_N(s) \quad (139)$$

$$\times (\exp(-sD^*(i) + s^2B^2/2))^{N-N'-N^\dagger+1} \quad (140)$$

$$\stackrel{f}{\leq} M(\exp(-sD^*(i) + s^2B^2/2))^{N-N'-N^\dagger+1}. \quad (141)$$

Here, $N^\dagger$ is the number of indices in $\mathcal{N}$ that are at least N' . Note that $N^\dagger < \lceil \log_a N \rceil$. Inequality (a) is a consequence of the result in (102). Inequality (b) holds because the event $\{T \leq N'\} \subseteq \mathcal{Z}_N$. We consider two cases for obtaining inequality (c). 1) If $N \in \mathcal{N}$, then we select U_N uniformly and inequality (c) follows from the same arguments used to prove Lemma 7 in Appendix A. 2) Notice that under the events $\mathcal{Z}_N$ and $\{X = i\}$, we have $\bar{i}_N = i$, and by the construction of strategy $\bar{g}^N$, the experiment U_N is selected using the control law $g_N^{N,i}$ if $N \notin \mathcal{N}$. This control law at time N satisfies Criterion 1 which is the condition required for using Lemma 8. Thus, inequality (c) can be obtained from the same arguments used to prove Lemma 8. Inequality (d) holds because $\mathcal{Z}_N \subseteq \mathcal{Z}_{N-1}$. Inequality (e) is obtained by inductively applying the arguments (b) – (d). Inequality (f) is due to Lemma 7 in Appendix A.

2) Chernoff Bound: Let $N'' \doteq N - N' - N^\dagger + 1$ and let $\zeta > 0$ be any small constant. Define

$$\begin{aligned} &\theta_N(i) \\ &\doteq \max \left\{ \zeta - \mathcal{C}_i(\rho_1), N''D^*(i) - \frac{s_N N'' B^2}{2} - \frac{1}{s_N} \log \frac{2M}{\epsilon_N} \right\} \end{aligned} \quad (142)$$

where s_N is as defined in (23). Under Assumption 2, one can verify that $\theta_N(i)/N \rightarrow D^*(i)$. Further, we can say that there exists an integer $\bar{N}$ such that $N'' > 0$ and for every $i \in \mathcal{X}$ and $N \geq \bar{N}$, we have $\theta_N(i) > \zeta - \mathcal{C}_i(\rho_1)$. Thus, for every $N \geq \bar{N}$, using the Chernoff bound [27], we have

$$\begin{aligned} &\mathbb{P}_i^g[\mathcal{C}_i(\rho_{N+1}) - \mathcal{C}_i(\rho_1) < \theta_N(i), T \leq N'] \\ &\leq \mathbb{E}_i^g \exp[-s(\mathcal{C}_i(\rho_{N+1}) - \mathcal{C}_i(\rho_1) - \theta_N(i)); T \leq N'] \\ &\stackrel{a}{\leq} M(\exp(s\theta_N(i) - sN''D^*(i) + s^2N''B^2/2)) \stackrel{b}{\leq} \epsilon_N/2 \end{aligned}$$

where inequality (a) follows from (141) and inequality (b) is obtained by substituting the values of $\theta_N(i)$ and s_N . Combining this result with (133), we have

$$\mathbb{P}_i^g[\mathcal{C}_i(\rho_{N+1}) - \mathcal{C}_i(\rho_1) \leq \theta_N(i)] \leq \epsilon_N.$$

Therefore, the strategy pair $(\bar{f}^N, \bar{g}^N)$ defined in Section IV-A1 satisfies the constraints in Problem (P2).

REFERENCES

- [1] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Hoboken, NJ, USA: Wiley, 2012.
- [2] R. Blahut, "Hypothesis testing and information theory," *IEEE Trans. Inf. Theory*, vol. 20, no. 4, pp. 405–417, Jul. 1974.
- [3] R. Blahut, "Information bounds of the Fano-Kullback type," *IEEE Trans. Inf. Theory*, vol. 22, no. 4, pp. 410–421, Jul. 1976.
- [4] E. Tuncel, "On error exponents in hypothesis testing," *IEEE Trans. Inf. Theory*, vol. 51, no. 8, pp. 2945–2950, Aug. 2005.
- [5] J. Liu, R. van Handel, and S. Verdú, "Second-order converses via reverse hypercontractivity," *Math. Statist. Learn.*, vol. 2, no. 2, pp. 103–163, 2020.
- [6] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Channel coding rate in the finite blocklength regime," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2307–2359, May 2010.
- [7] Y. Polyanskiy and Y. Wu, "Lecture notes on information theory," *MIT (6.441), UIUC (ECE 563)*, Yale (STAT 664), 2012–2017, p. 7, 2014. [Online]. Available: http://people.lids.mit.edu/yp/homepage/data/itlectures_v5.pdf
- [8] S. Nitinawarat, G. K. Atia, and V. V. Veeravalli, "Controlled sensing for multihypothesis testing," *IEEE Trans. Autom. Control*, vol. 58, no. 10, pp. 2451–2464, Oct. 2013.
- [9] A. Wald, *Sequential Analysis*. Chelmsford, MA, USA: Courier Corporation, 1973.
- [10] H. Chernoff, "Sequential design of experiments," *Ann. Math. Statist.*, vol. 30, no. 3, pp. 755–770, 1959.
- [11] S. A. Bessler, "Theory and applications of the sequential design of experiments, k-actions and infinitely many experiments," Department of Statistics, Stanford University, Stanford, CA, USA, Rep. SOLONR55, 1960.
- [12] M. Naghshvar *et al.*, "Active sequential hypothesis testing," *Ann. Statist.*, vol. 41, no. 6, pp. 2703–2738, 2013.
- [13] M. Naghshvar and T. Javidi, "Extrinsic Jensen-Shannon divergence with application in active hypothesis testing," in *Proc. IEEE Int. Symp. Inf. Theory*, 2012, pp. 2191–2195.
- [14] R. Keener *et al.*, "Second order efficiency in the sequential design of experiments," *Ann. Statist.*, vol. 12, no. 2, pp. 510–532, 1984.
- [15] S. Lalley and G. Lorden, "A control problem arising in the sequential design of experiments," *Ann. Probability*, vol. 14, no. 1, pp. 136–172, 1986.
- [16] P. Auer, N. Cesa-Bianchi, Y. Freund, and R. E. Schapire, "The non-stochastic multiarmed bandit problem," *SIAM J. Comput.*, vol. 32, no. 1, pp. 48–77, 2002.
- [17] B. Huang, K. Cohen, and Q. Zhao, "Active anomaly detection in heterogeneous processes," *IEEE Trans. Inf. Theory*, vol. 65, no. 4, pp. 2284–2301, Apr. 2019.
- [18] D. Chen, Q. Huang, H. Feng, Q. Zhao, and B. Hu, "Active anomaly detection with switching cost," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, 2019, pp. 5346–5350.
- [19] A. Tsopelakos, G. Fellouris, and V. V. Veeravalli, "Sequential anomaly detection with observation control," in *Proc. IEEE Int. Symp. Inf. Theory*, 2019, pp. 2389–2393.
- [20] A. Gurevich, K. Cohen, and Q. Zhao, "Sequential anomaly detection under a nonlinear system cost," *IEEE Trans. Signal Process.*, vol. 67, no. 14, pp. 3689–3703, Jul. 2019.
- [21] C. Wang, K. Cohen, and Q. Zhao, "Active hypothesis testing on a tree: Anomaly detection under hierarchical observations," in *Proc. IEEE Int. Symp. Inf. Theory*, 2017, pp. 993–997.
- [22] N. K. Vaidhiyan and R. Sundaresan, "Learning to detect an oddball target," *IEEE Trans. Inf. Theory*, vol. 64, no. 2, pp. 831–852, Feb. 2018.
- [23] N. K. Vaidhiyan, S. P. Arun, and R. Sundaresan, "Neural dissimilarity indices that predict oddball detection in behaviour," *IEEE Trans. Inf. Theory*, vol. 63, no. 8, pp. 4778–4796, Aug. 2017.
- [24] M. V. Burnashev, "Data transmission over a discrete channel with feedback. Random transmission time," *Problemy Peredachi Informatsii*, vol. 12, no. 4, pp. 10–30, 1976.
- [25] M. Naghshvar, T. Javidi, and M. Wigger, "Extrinsic Jensen-Shannon divergence: Applications to variable-length coding," *IEEE Trans. Inf. Theory*, vol. 61, no. 4, pp. 2148–2164, Apr. 2015.
- [26] M. J. Osborne and A. Rubinstein, *A Course in Game Theory*. Cambridge, MA, USA: MIT Press, 1994.
- [27] J. A. Gubner, *Probability and Random Processes for Electrical and Computer Engineers*. Cambridge, U.K.: Cambridge Univ. Press, 2006.
- [28] D. A. Freedman *et al.*, "On tail probabilities for martingales," *Ann. Probability*, vol. 3, no. 1, pp. 100–118, 1975.
- [29] P. Massart, "Concentration inequalities and model selection," *Ecole d'Été de Probabilités de Saint-Flour XXXIII – 2003*, J. Picard, Ed., Berlin, Heidelberg: Springer, Eds. 1, 2007, pp. XIV, 343, doi: [10.1007/978-3-540-48503-2](https://doi.org/10.1007/978-3-540-48503-2).

Dhruva Kartik received the B.Tech. degree in electronics and communication engineering from the Indian Institute of Technology, Guwahati, India, in 2015.

He joined the Ming Hsieh Department of Electrical and Computer Engineering, University of Southern California, Los Angeles, CA, USA, in 2015 where he is currently a doctoral candidate. His research interests include decentralized stochastic control, decision-making in sensing and communication systems, game theory, and reinforcement learning.

Mr. Kartik was an Annenberg Fellow from 2015 to 2019.

Ashtosh Nayyar (Senior Member, IEEE) received the B.Tech. degree in electrical engineering from the Indian Institute of Technology, Delhi, India, the M.S. degree in electrical engineering and computer science, the M.S. degree in applied mathematics, and the Ph.D. degree in electrical engineering and computer science, all from the University of Michigan, Ann Arbor, MI, USA, in 2006, 2008, 2011, and 2011, respectively.

He was a Postdoctoral Researcher with the University of Illinois at Urbana-Champaign, Champaign, IL, USA, and at the University of California, Berkeley, Berkeley, CA, USA, before joining the University of Southern California, Los Angeles, CA, USA, in 2014. His research interests include decentralized stochastic control, decentralized decision-making in sensing and communication systems, reinforcement learning, game theory, mechanism design and electric energy systems.

Urbashi Mitra (Fellow, IEEE) received the B.S. and the M.S. degrees from the University of California at Berkeley, Berkeley, CA, USA, and the Ph.D. degree from Princeton University, Princeton, NJ, USA, in 1987, 1989, and 1994.

She is currently the Gordon S. Marshall Professor in Engineering with the University of Southern California. She has held visiting appointments with King's College, London, U.K., Imperial College, London, U.K., the Delft University of Technology, Delft, Netherlands, Stanford University, Stanford, CA, USA, Rice University, Houston, TX, USA, and the Eurecom Institute, Biot, France. Her research interests include wireless communications, communication and sensor networks, biological communication systems, detection and estimation and the interface of communication, sensing and control.

Dr. Mitra was the inaugural Editor-in-Chief for the IEEE TRANSACTIONS ON MOLECULAR, BIOLOGICAL AND MULTI-SCALE COMMUNICATIONS. She has been a member of the IEEE Communication Society's Board of Governors from 2018–2020, the IEEE Information Theory Society's Board of Governors (2002–2007, 2012–2017), the IEEE Signal Processing Society's Technical Committee on Signal Processing for Communications and Networks (2012–2016), the IEEE Signal Processing Society's Awards Board (2017–2018), and the Chair/Vice Chair of the IEEE Communications Society, Communication Theory Technical Committee (2019–2020, 2017–2018). She was the recipient of 2021 Viterbi School of Engineering Senior Research Award, the 2017 IEEE Women in Communications Engineering Technical Achievement Award, a 2015 U.K. Royal Academy of Engineering Distinguished Visiting Professorship, a 2015 US Fulbright Scholar Award, a 2015–2016 U.K. Leverhulme Trust Visiting Professorship, IEEE Communications Society Distinguished Lecturer, 2012 Globecom Signal Processing for Communications Symposium Best Paper Award, 2012 US National Academy of Engineering Lillian Gilbreth Lectureship, the 2009 DCSS Applications & Systems Best Paper Award, 2001 Okawa Foundation Award, 2000 Ohio State University's College of Engineering Lumley Award for Research, 1997 Ohio State University's College of Engineering MacQuigg Award for Teaching, and a 1996 National Science Foundation CAREER Award. She has been an Associate Editor for multiple IEEE publications.