Heterogeneous Sequential Hypothesis Testing with Active Source Selection under Budget Constraints

Sung-Wen Lan and I-Hsiang Wang
Graduate Institute of Communication Engineering, National Taiwan University, Taipei, Taiwan
Email: andylan0706@gmail.com,ihwang@ntu.edu.tw

***Abstract*—Sequential binary hypothesis testing from temporally heterogeneously generated random samples with an active decision maker under budget constraints is considered. The problem is motivated from applications in crowdsourced classification and sequential detection from sensory data in IoT networks. In such applications, at each time slot, the source of data may vary from time to time, and the decision on the two possible hypotheses is to be made in a reliable, fast, and cost effective manner. In particular, the active decision maker either takes the current source and collects a sample, or skips the current source and waits for the next time slot. At the end of each time slot, the decision maker either decides to claim the decision on the two hypotheses, or continues to observe the data source for the next time slot. The goal is to design action taking and decision making policies so that the probability of error is minimized under two constraints: one on the total number of samples collected by the decision maker, and the other is on the total number of time slots. In this work, the available data source changes among n possible ones i.i.d. over time, and the two constraints are in expectation. We establish the optimal error exponents of the two types of error probabilities as the two constraints tend to infinity with a fixed proportion. For achievability, a scheme that combines a sequential probability ratio test and an adaptive randomized policy that dynamically switches between two sets of accepting probabilities of the current source, according to the observed samples collected so far, is proposed. Matching upper bounds on the error exponents are developed using data processing inequality and Doob's Optional Stopping Theorem.**

A full version of this paper is accessible at:
http://homepage.ntu.edu.tw/~ihwang/Eprint/isit21sacsc.pdf

I. Introduction

Sequential hypothesis testing has gained increasing attention recent years owing to its usage in a wide range of applications including crowdsourcing, surveillance, spectrum sensing, sensor networks, scientific research, etc.. The problem involves making a decision over a given set of hypotheses based on the samples sequentially collected from the observed data sources. The decision maker can adaptively decide when to stop the observation and make the decision based on the samples collected so far. While its root can be traced back to the seminal works by Wald [1] and Wald and Wolfowitz [2], a variety of extensions and alternatives of the classical sequential hypothesis testing problem have been proposed and explored (to name a few, [3], [4]). See [5] for a glimpse at the vast literature.

Among them, a branch of works initiated by Chernoff [6] (termed *active sequential hypothesis testing* in [7] and *controlled sensing* in [8]) have been focused on taking additional action which affects the quality of the samples. The action typically involves selection from multiple data sources (or statistical experiments in [6]), with applications in sensor selection [9]–[12], medical diagnosis, etc.. The set of available data sources to be selected from, however, is assumed to be fixed throughout the entire process, which is a reasonable assumption in the applications considered in the literature.

Meanwhile, in other applications such as crowdsourced classification and sequential detection from sensory data in IoT networks, the available data sources might not be the same throughout – they may change over time and hence are temporally heterogeneous, rendering the setting considered in Chernoff [6] and follow-up works [7], [9]–[13] not directly applicable. For example, in crowdsourced data annotation, the quality and reliability of available workers may vary over time. Faced with such temporal variability, the action now additionally involves deciding whether to select from the current set of available data sources, or to wait for the next time instance. With *two* constraints – one on the total number of collected samples and the other on the total number of time instances, the goal is to design optimal decision making and action taking policies so that the probability of error is minimized. While there has been a vast amount of existing works that address the policy design and characterization of the optimal performance when the set of available data sources is fixed [7], [9]–[13], for the setting with temporally heterogeneous data sources, such results are missing in the literature.

To address the challenge imposed by such temporal heterogeneity, in this work we consider a canonical setting with a single available data source at each time instance, and the available one changes over time randomly (i.i.d. over time). The active decision maker is faced with two constraints: one on the total number of collected samples, and the other on the total number of time instances. At each time instance, the active decision maker either takes the current source and collects a sample, or skips it and waits for the next time instance. At the end of each time instance, the decision maker either decides to claim the decision on the two hypotheses, or continues to observe the data source for the next time instance. The goal is to design action taking and decision making policies so that the probability of error is minimized under the two aforementioned constraints. These constraints can either be absolute (termed "hard" constraints) or in expectation (termed "soft" constraints).

The focus of this work is on the problem with soft constraints. Under the constraint that the expected number of collected samples is not greater than B and the expected number of time instances is not greater than T, it is natural to employ a randomized selection policy to meet the constraints in expectation. Hence, the key is to determine the probability of accepting each kind of source based on the source distributions and the observations so far. Such a randomized selection policy is specified by a vector of accepting probabilities. We propose an adaptive policy that dynamically switches between two kinds of accepting probabilities, depending on the observed samples collected so far. The proposed policy is shown to be *asymptotically* optimal as T and B both tend to infinity with a fixed proportion, in the sense that the achievable type-I and type-II error probabilities can both vanish exponentially fast with T, and the rate functions (error exponents) are optimal. To characterize the achievable error exponents, proof techniques in [7] based on martingale theory are employed. For the proof of the converse part, data processing inequality along with Doob's Optional Stopping Theorem are leveraged, which is similar to that in [14], [15, Chapter 15.3]. As a result, the optimal type-I and type-II error exponents are characterized by two linear programs that can be solved explicitly, and the resultant error exponents are certain linear combinations of the n KL divergences between the two hypothetical distributions of the n data sources. The coefficients in the linear combinations are proportional to the accepting probabilities of the sources. The fact that both types of error exponents are linear combinations of KL divergences reflects the benefit of adaptivity in the sequential setting since the soft constraints leave some soft margins in time so that the decision maker can wait for more informative sources.

In the literature, the most closely related works lie in the field of sensor selection [10]–[12]. The setting can be viewed as a special case of active sequential binary hypothesis testing [7], [8], where multiple sources are available, and the decision maker has to select a source at each time instance. Bai and Gupta [11] considered heterogeneous sampling cost. They minimized the expected total cost over selection policies under the sequential probability ratio test with fixed stopping boundaries. Bai *et al.* [10] also studied the minimization of the expected total cost when the sampling cost is heterogeneous. The minimization problem is subject to the constraints on the probability of error and the expected usage of each source usage. Policies in their work are prohibited from using the collected samples. Meanwhile, Li *et al.* [12] used the information of collected samples to minimize the expected number of samples, subject to the constraints on probability of error and sensor usage. Notably, none of the above-mentioned existing works consider temporal variability of the sources that are available.

II. Problem Formulation

As motivated in the previous section, let us formulate the problem in detail as follows. Consider n heterogeneous data sources, labeled by integers in $[n] \triangleq \{1, ..., n\}$.

A. Statistical model

The statistical model involves two parts: which data source is available at each time instance, and the realization of the collected data samples.

At time instance t, $t \geq 1$, the index $W_t \in [n]$ of the current *available* source follows a distribution P_α, that is,

$$P_\alpha(j) \triangleq \mathbb{P}\{W_t = j\} = \alpha_j, \quad \forall\, j \in [n]\,.$$

In words, the sequence of the indices of available data sources $\{W_t\}_{t\in\mathbb{N}}$ follows P_α in an i.i.d. fashion. For notational convenience, let $\boldsymbol{\alpha} = [\alpha_1, ..., \alpha_n]$ denote this probability vector in the n-dimensional probability simplex.

The goal of the decision maker is to infer the hypothesis $\mathcal{H}_\theta$, $\theta \in \{0, 1\}$ from the collected samples. Under the hypothesis $\mathcal{H}_\theta$, the sample drawn from each source (say source j, $j \in [n]$) follows distribution $P_{\theta,j}$, that is,

$$X_t \sim P_{\theta,j} \quad \text{if } w_t = j.$$

The distributions of all sources under different hypotheses are known to the active decision maker. For notational convenience, let us also denote the alphabet of X_t's as $\mathcal{X}$, and use $\mathbb{P}_\theta\{\cdot\}$ and $\mathbb{E}_\theta[\cdot]$ as the short-hand notations for "the probability" and "the expectation" under hypothesis $\mathcal{H}_\theta$, for $\theta = 0, 1$.

Throughout this manuscript, we make the following technical assumption as in [7]. This technical assumption is used in our achievability and converse proofs as well as to ensure the error probabilities do vanish in the asymptotic regime.

Assumption 1: The one-step log-likelihood ratio (LLR) is bounded, that is, there exists a constant $L > 0$ such that

$$\forall\, j \in [n]\,,\ \sup_{x\in\mathcal{X}} \left| \log \frac{P_{1,j}(x)}{P_{0,j}(x)} \right| \leq L. \tag{1}$$

Moreover, $\exists\, j \in [n]$ such that $D\left(P_{1,j} \,\|P_{0,j}\right) > 0$, and $\exists\, j \in [n]$ such that $D\left(P_{0,j} \,\|P_{1,j}\right) > 0$. Here we denote the KL divergence of distribution P from distribution Q as $D\left(P \,\|Q\right)$.

B. Active source selection and inference

At each time instance, the decision maker first takes the following binary *selection* action: it either takes the current available data source and collects a sample X_t drawn from the source, or skips the current source and waits for the next time instance. The binary action is denoted as $\delta_t \in \{0, 1\}$, where 1 means to take the current source. In our formulation, randomized policy is allowed, that is, δ_t is a random variable, which follows a distribution determined by the decision maker. After the selection action, based on the previous collected samples and their respective data sources, the decision maker decides whether to stop collecting data and infer the underlying hypothesis, or to continue to the next time instance. Hence, the time instance at which the decision maker outputs the inferred result is a *stopping time*, which we denote as τ. The inferred result is denoted as $\phi \in \{0, 1\}$, that is, the inferred hypothesis is $\mathcal{H}_\phi$.

In short, there are three parts of the decision maker's policy: (1) the randomized selection action $\{\delta_t \,|\, t = 1, 2, ...\}$, (2) the stopping rule $\mathbb{1}\{\tau = t\}$, and (3) inference of the hypothesis ϕ. To describe how the three parts depend on the overall collected information by the decision maker, let us denote it (up to time t) as F_t, consisting of the collected samples and their sources so far: $F_t \triangleq \{(X_s, W_s) \,|\, 1 \le s \le t \,:\, \delta_s = 1\}$. Note that there is no need to include the indices of the *skipped* data sources because of the temporal independence. For the selection action, the probability of taking the current source $\mathbb{P}\{\delta_t = 1\}$ is a function of F_{t-1} and the index of the current source W_t. The stopping rule $\mathbb{1}\{\tau = t\}$ and the inference ϕ are both functions of F_t. For notational convenience, let us denote the distribution of F_t under hypothesis $\mathcal{H}_\theta$ as $\mathsf{P}_\theta^{(t)}(\cdot)$, for $\theta = 0, 1$.

As for the performance of the inferred result, let us define the following probability of errors:

$$\text{Type 1: } \pi_{1|0} \triangleq \mathbb{P}_0\{\phi = 1\} \;;\; \text{Type 2: } \pi_{0|1} \triangleq \mathbb{P}_1\{\phi = 0\}.$$

For convenience, let the probabilities of success $\pi_{0|0} \triangleq 1 - \pi_{1|0}$ and $\pi_{1|1} \triangleq 1 - \pi_{0|1}$. For the Bayesian setting, let π_θ denote the prior of hypothesis $\mathcal{H}_\theta$, $\theta \in \{0, 1\}$. The Bayesian error probability is hence $\mathsf{P}_\mathsf{e} \triangleq \pi_0 \pi_{1|0} + \pi_1 \pi_{0|1}$.

C. Constraints and error exponents

The decision maker is faced with two constraints: one on the total number of collected samples (budget constraint), and the other on the total number of time instances (time constraint). In particular, let us denote the total number of collected samples as $\mathcal{B} \triangleq \sum_{t \le \tau} \delta_t$. Meanwhile, the stopping time τ corresponds to the total number of time instances. The two soft constraints are then given as follows:

$$\mathbb{E}_\theta[\tau] \le T \quad \text{and} \quad \mathbb{E}_\theta[\mathcal{B}] \le B, \quad \forall\, \theta \in \{0, 1\}. \tag{2}$$

As for the asymptotic performance, our focus is on the case where both B and T tend to infinity with a fixed ratio $\frac{B}{T} = r$, $0 \le r \le 1$. In such an asymptotic regime, we say that the error exponents of the type 1 and the type 2 error probabilities, denoted by E_0 and E_1 respectively, are *achievable* if there exists a sequence of policies under the constraints in (2) such that the two types of error probabilities *both vanish* and satisfy

$$\liminf_{\substack{T\to\infty \\ B=rT}} \left\{\frac{1}{T}\log\frac{1}{\pi_{1|0}}\right\} \ge E_0, \liminf_{\substack{T\to\infty \\ B=rT}} \left\{\frac{1}{T}\log\frac{1}{\pi_{0|1}}\right\} \ge E_1. \tag{3}$$

The collection of all achievable (E_0, E_1) is called the *exponent region* $\mathcal{E}(r)$, where the dependency on the *budget-to-time ratio* r is emphasized. In addition, the optimal Bayesian error exponent is defined as

$$E^*(r) \triangleq \lim_{\substack{T\to\infty \\ B=rT}} \left\{\frac{1}{T}\log\frac{1}{\mathsf{P}_\mathsf{e}^*(T, B)}\right\},$$

where $\mathsf{P}_\mathsf{e}^*(T, B)$ is the minimum Bayesian error probability P_e subject to the constraints in (2).

Remark 1: Our treatment on the asymptotic performance is focused on policies that can guarantee vanishing type-1 and type-2 error probabilities. Note that policies providing a constant guarantee on one type of error probability, namely, those in the "*Neyman-Pearson*" regime, are not considered in this work. Such restriction when studying the asymptotic performance was also taken in previous works [6], [8]. In our main theorem, the characterization of error exponents in this work does not cover those in the Neyman-Pearson regime.

Notations: The symbol $\preceq$ denotes the element-wise less-than-or-equal-to: for $\boldsymbol{\beta}, \boldsymbol{\alpha} \in \mathbb{R}^n$, $\boldsymbol{\beta} \preceq \boldsymbol{\alpha} \iff \beta_j \le \alpha_j,\ \forall\, j \in [n]$. $\|\,.\,\|_1$ denotes the ℓ_1-norm. $\mathbb{R}_+$ denotes the set of non-negative reals.

III. Main Results

Our main result is summarized in the following theorem.

Theorem 1 (Characterization of the Exponent Region): The exponent region subject to the budget-to-time ratio $r \in [0, 1]$, $\mathcal{E}(r)$, as defined in Section II-C, is the set of positive (E_0, E_1) satisfying the following inequalities:

$$E_0 \le \underbrace{\max_{\boldsymbol{\beta}\in\mathbb{R}_+^n} \left\{\sum_{j\in[n]} \beta_j D\left(P_{1,j} \,\|P_{0,j}\right) \,\middle|\, \|\boldsymbol{\beta}\|_1 \le r,\, \boldsymbol{\beta} \preceq \boldsymbol{\alpha}\right\}}_{\mathsf{D}_0^*(r)}$$

$$E_1 \le \underbrace{\max_{\boldsymbol{\beta}\in\mathbb{R}_+^n} \left\{\sum_{j\in[n]} \beta_j D\left(P_{0,j} \,\|P_{1,j}\right) \,\middle|\, \|\boldsymbol{\beta}\|_1 \le r,\, \boldsymbol{\beta} \preceq \boldsymbol{\alpha}\right\}}_{\mathsf{D}_1^*(r)}$$

Sketch of Proof: The converse part is proved in Section V by data processing inequality and Doob's Optional Stopping Theorem. As for the achievability part, note that the maximum of the type 1 and type 2 error exponents, as denoted above by $\mathsf{D}_0^*(r)$ and $\mathsf{D}_1^*(r)$ respectively, are both solutions to a simple linear program that can be solved by a greedy algorithm explicitly. In each of the two cases, the optimal solution of $\boldsymbol{\beta}$ is to greedily raise the "weightings" of taking the most informative source (measured by the respective KL divergence) under the constraints until the sum of weightings reaches the budget-to-time ratio r. Intuitively, the optimal solutions of $\boldsymbol{\beta}$ are proportional to the accepting probabilities of the sources.

Let us denote the optimal solutions to the two linear programs as $\boldsymbol{\beta}_1^*(r)$ and $\boldsymbol{\beta}_0^*(r)$ respectively. Hence, $\boldsymbol{\beta}_1^*(r)$ depends on the ordering of the source indices ranked by KL divergences $\{D(P_{1,j} \,\|P_{0,j}) \mid j \in [n]\}$, while $\boldsymbol{\beta}_0^*(r)$ depends on the ordering of the source indices ranked by $\{D(P_{0,j} \,\|P_{1,j}) \mid j \in [n]\}$. Since KL divergences are not necessarily symmetric, in general $\boldsymbol{\beta}_1^*(r) \ne \boldsymbol{\beta}_0^*(r)$, and a static randomized policy will not work. Instead, inspired by [7], we propose an *adaptive* randomized policy which dynamically switches between two sets of accepting probabilities. As for determining when to stop and make the inference, we employ a sequential probability ratio test (SPRT). The details of the scheme and analysis can be found in Section IV. ∎

As a straightforward corollary of Theorem 1, the optimal Bayesian error exponent is characterized.

Corollary 1: The optimal Bayesian error exponent subject to the budget-to-time ratio $r \in [0, 1]$ is

$$E^*(r) = \min\{\mathsf{D}_0^*(r), \mathsf{D}_1^*(r)\}.$$

Note: For convenience, when the context is clear, in the rest of this manuscript, we may drop the dependency on r in $\mathsf{D}_0^*(r)$, $\mathsf{D}_1^*(r)$, $\boldsymbol{\beta}_0^*(r)$, and $\boldsymbol{\beta}_1^*(r)$.

IV. Proof of Achievability

In this section, we first describe the proposed scheme which comprises an adaptive randomized policy as motivated in the Sketch of Proof of Theorem 1 and a SPRT. The main intuition is that, at a certain stage, if the decision maker is more confident that $\mathcal{H}_\theta$ is correct based on the collected samples so far, it should use the accepting probabilities that *favor* distinguishing from the other hypothesis. We will formalize this intuition later. Then, performance analysis of the proposed policy is carried out to complete the proof of achievability.

A. The proposed scheme

The proposed scheme comprises two parts: (1) an adaptive randomized policy for determining the selection action, and (2) a SPRT for determining when to stop and make the inference. For both parts, the *log-likelihood ratio* (LLR) at each time instance t, defined as follows, plays a crucial role:

$$S_t \triangleq \log \frac{\mathsf{P}_1(F_t)}{\mathsf{P}_0(F_t)} \tag{4}$$

First, let us describe the adaptive randomized selection policy. It depends on the solutions to the linear programs charactering the error exponents. Denote the re-ordering of the indices $[n]$ ranked by $D\left(P_{1,j} \,\|P_{0,j}\right)$ and $D\left(P_{0,j} \,\|P_{1,j}\right)$ as σ_1, σ_0 respectively:

$$\begin{aligned} &D\left(P_{1,\sigma_1(1)} \left\|P_{0,\sigma_1(1)}\right.\right) \geq ... \geq D\left(P_{1,\sigma_1(n)} \left\|P_{0,\sigma_1(n)}\right.\right) \\ &D\left(P_{0,\sigma_0(1)} \left\|P_{1,\sigma_0(1)}\right.\right) \geq ... \geq D\left(P_{0,\sigma_0(n)} \left\|P_{1,\sigma_0(n)}\right.\right). \end{aligned}$$

Recall that $\boldsymbol{\beta}_1^*$ denote the optimal solution of achieving D_0^* and $\boldsymbol{\beta}_0^*$ denote that of D_1^*. They can be explicitly solved as follows: for $\theta = 0, 1$, $j \in [n]$, the $\sigma_\theta(j)$-th entry of $\boldsymbol{\beta}_\theta$ is

$$\beta^*_{\theta,\sigma_\theta(j)} = \begin{cases} \alpha_{\sigma_\theta(j)} & \text{if } j < \tilde{w}(\sigma_\theta, r) \\ r - \sum_{j=1}^{\tilde{w}(\sigma_\theta,r)-1} \alpha_{\sigma_\theta(j)} & \text{if } j = \tilde{w}(\sigma_\theta, r) \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

where $\tilde{w}(\sigma_\theta, r)$ is the threshold such that

$$\textstyle\sum_{j=1}^{\tilde{w}(\sigma_\theta,r)-1} \alpha_{\sigma_\theta(j)} < \ r \ \leq \sum_{j=1}^{\tilde{w}(\sigma_\theta,r)} \alpha_{\sigma_\theta(j)}. \tag{6}$$

One can recognize the greedy algorithm for solving the linear programs from the solutions above.

We are now ready to describe the adaptive randomized policy for the take-it-or-not action:

$$\mathbb{P}\{\delta_t^* = 1 | w_t = j\} = \begin{cases} \beta^*_{1,j}/\alpha_j & \text{if } S_t \geq 0 \\ \beta^*_{0,j}/\alpha_j & \text{otherwise.} \end{cases} \tag{7}$$

The intuition of the policy is that, if one has more confidence that one of the hypotheses (for example, $\mathcal{H}_1$) is true, one should take the action that maximizes the expected absolute value of the LLR (which is a linear combination of KL divergences $\{D\left(P_{1,j} \,\|P_{0,j}\right) \mid j \ \in \ [n]\}$) to distinguish the hypothesis faster from the other hypothesis.

As for the SPRT that determines when to stop collecting samples and make the inference, the stopping time

$$\tau = \min\nolimits_{t \in \mathbb{N}} \{t \,|\, S_t \geq A_1 \text{ or } S_t \leq -A_0\}, \tag{8}$$

where $A_1 \triangleq T\mathsf{D}_0^* - C_1$, $A_0 \triangleq T\mathsf{D}_1^* - C_0$, and C_1, C_0 are constants not depending on T. The decision rule that gives the inferred result is

$$\phi = \begin{cases} 1 & \text{if } S_\tau \geq A_1 \\ 0 & \text{if } S_\tau \leq -A_0 \end{cases}. \tag{9}$$

B. Proof of the achievability part of Theorem 1

Let us now analyze the performance of the proposed scheme and complete the proof of the achievability part. The key is how to bound the expected stopping time by analyzing the behavior of the LLR's $\{S_t\}$ under the proposed adaptive policy. Once the expected stopping time is successfully bounded, the expected budget can be bounded via martingale method and Doob's Optional Stopping Theorem, and the rest of the proof just follows the same procedure of bounding the error probabilities as in SPRT.

To get a handle on the behavior of the LLR under the proposed policy, we leverage the following lemma which is a slight extension of [7, Lemma 4] to our case. Its proof is provided in Appendix B of the full version.

Lemma 1: Suppose the following three conditions are satisfied for constants K_1, K_2, K_3:

$$\begin{aligned} &\mathbb{E}_1\left[S_{t+1} \,|S_t\right] \geq S_t + K_1 \quad \text{if } S_t < 0 \\ &\mathbb{E}_1\left[S_{t+1} \,|S_t\right] \geq S_t + K_2 \quad \text{if } S_t \geq 0 \\ &|S_{t+1} - S_t| \leq K_3 \end{aligned} \tag{10}$$

Then, for the stopping time $\tau_{\text{upper}} = \min\{t : S_t \geq A\}$, its expectation is upper bounded as follows:

$$\mathbb{E}_1\left[\tau_{\text{upper}}\right] \leq \frac{A + K_3}{K_2} + \frac{K_3}{1 - e^{-K_3}}\left(\frac{1}{K_1} - \frac{1}{K_2}\right). \tag{11}$$

Similarly, suppose the following three conditions are satisfied for constants K_4, K_5, K_6:

$$\begin{aligned} &\mathbb{E}_0\left[-S_{t+1} \,|-S_t\right] \geq -S_t + K_4 \quad \text{if } -S_t \leq 0 \\ &\mathbb{E}_0\left[-S_{t+1} \,|-S_t\right] \geq -S_t + K_5 \quad \text{if } -S_t > 0 \\ &|S_{t+1} - S_t| \leq K_6 \end{aligned} \tag{12}$$

Then, for the stopping time $\tau_{\text{lower}} = \min\{t : -S_t \geq A\}$, its expectation is upper bounded as follows:

$$\mathbb{E}_0\left[\tau_{\text{lower}}\right] \leq \frac{A + K_6}{K_5} + \frac{K_6}{1 - e^{-K_6}}\left(\frac{1}{K_4} - \frac{1}{K_5}\right) \tag{13}$$

We are ready to analyze the behavior of LLR's. Note that the stopping time τ in (8) can be written as $\tau = \min\{\tau_1, \tau_0\}$, where

$$\tau_1 = \min_{t \in \mathbb{N}} \{t \,|\, S_t \geq A_1\} \text{ and } \tau_0 = \min_{t \in \mathbb{N}} \{t \,|\, S_t \leq -A_0\}.$$

In the following, we are going to invoke Lemma 1 to find appropriate C_1 and C_0 so that $\mathbb{E}_1\left[\tau_1\right] \leq T$ and $\mathbb{E}_0\left[\tau_0\right] \leq T$. Since $\mathbb{E}_1\left[\tau_1\right] \geq \mathbb{E}_1\left[\tau\right]$ and $\mathbb{E}_0\left[\tau_0\right] \geq \mathbb{E}_0\left[\tau\right]$, the temporal constraints are satisfied.

Consider the case where the ground truth is $\mathcal{H}_1$. Based on the proposed policy and equation (1) in Assumption 1, we have

$$\mathbb{E}_1\left[S_{t+1}\,|S_t\right] = S_t + \textstyle\sum_{j\in[n]} \beta^*_{0,j} D\left(P_{1,j}\,\|P_{0,j}\right) \quad \text{for } S_t < 0$$
$$\mathbb{E}_1\left[S_{t+1}\,|S_t\right] = S_t + \mathsf{D}^*_0 \quad \text{for } S_t \geq 0$$
$$|S_{t+1} - S_t| \leq L$$

Let us identify the constants in (10) as follows:

$$K_1 \leftarrow \sum_{j\in[n]} \beta^*_{0,j} D\left(P_{1,j}\,\|P_{0,j}\right);\; K_2 \leftarrow \mathsf{D}^*_0;\; K_3 \leftarrow L. \quad (14)$$

By equation (11) in Lemma 1, the constraint $\mathbb{E}_1\left[\tau_1\right] \leq T$ can be satisfied by solving C_1 from the following equation:

$$T = \frac{A + K_3}{K_2} + \frac{K_3}{1 - e^{-K_3}}\left(\frac{1}{K_1} - \frac{1}{K_2}\right)$$

with the identification in equation (14) and

$$A \leftarrow A_1 = T\mathsf{D}^*_0 - C_1.$$

The resulting C_1 is indeed a constant independent of T:

$$C_1 = L + \tfrac{L\mathsf{D}^*_0}{1-e^{-L}}\left(\tfrac{1}{\sum_{j\in[n]}\beta^*_{0,j} D(P_{1,j}\,\|P_{0,j})} - \tfrac{1}{\mathsf{D}^*_0}\right).$$

As for the case where the ground truth is $\mathcal{H}_0$, a similar approach can be applied to find a constant C_0 that is independent of T such that $\mathbb{E}_0\left[\tau_0\right] \leq T$. The resulting constant

$$C_0 = L + \tfrac{L\mathsf{D}^*_1}{1-e^{-L}}\left(\tfrac{1}{\sum_{j\in[n]}\beta^*_{1,j} D(P_{0,j}\,\|P_{1,j})} - \tfrac{1}{\mathsf{D}^*_1}\right).$$

Hence, with the above choices of C_1 and C_0, we have shown that the proposed policy satisfies the temporal constraint $\mathbb{E}_\theta\left[\tau\right] \leq T$ for $\theta = 0, 1$.

Now, let us show that the proposed policy satisfies the budget constraint $\mathbb{E}_\theta\left[\mathcal{B}\right] \leq B$ for $\theta = 0, 1$. Note that under the proposed policy described in (5) – (7), the probability of taking the source is r:

$$\mathbb{P}\{\delta^*_t = 1\} = \textstyle\sum_{j\in[n]} \alpha_j \mathbb{P}\{\delta^*_t = 1 | w_t = j\} = r.$$

A martingale is then constructed to bound the expected budget using Doob's Optional Stopping Theorem. Let $\mathcal{B}_t \triangleq \sum_{k\leq t}\delta_k$ denote the total spent budget at time t. Under the proposed policy, the random process $\{M_t = \mathcal{B}_t - rt\,|\,t \geq 0\}$ is a martingale. By Doob's Optional Stopping Theorem,

$$0 = M_0 = \mathbb{E}_\theta\left[M_\tau\right] = \mathbb{E}_\theta\left[\mathcal{B}_\tau\right] - r\mathbb{E}_\theta\left[\tau\right] = \mathbb{E}_\theta\left[\mathcal{B}\right] - r\mathbb{E}_\theta\left[\tau\right].$$

Since the temporal constraints are satisfied, $\mathbb{E}_\theta\left[\tau\right] \leq T$, and hence $\mathbb{E}_\theta\left[\mathcal{B}\right] = r\mathbb{E}_\theta\left[\tau\right] \leq rT = B$.

Finally, let us upper bound the probability of errors by the martingale method. The proof is similar to that of SPRT. Due to space constraint, we leave the details in Appendix A of the full version. Eventually, we are able to show that $\pi_{0|1} \leq e^{-A_0}$ and $\pi_{1|0} \leq e^{-A_1}$. Hence,

$$\liminf_{\substack{T\to\infty\\B=rT}}\left\{\frac{1}{T}\log\frac{1}{\pi_{0|1}}\right\} \geq \lim_{T\to\infty}\frac{A_0}{T} = \mathsf{D}^*_1,$$
$$\liminf_{\substack{T\to\infty\\B=rT}}\left\{\frac{1}{T}\log\frac{1}{\pi_{1|0}}\right\} \geq \lim_{T\to\infty}\frac{A_1}{T} = \mathsf{D}^*_0,$$

and the achievability proof is now complete. ■

V. Proof of Converse

Let us first introduce the following notation: for $p, q \in [0, 1]$,

$$d\left(p\,\|q\right) \triangleq p\log\tfrac{p}{q} + (1-p)\log\tfrac{1-p}{1-q} = D\left(\mathrm{Ber}(p)\,\|\mathrm{Ber}(q)\right),$$
$$h(p) \triangleq p\log\tfrac{1}{p} + (1-p)\log\tfrac{1}{1-p} = H(\mathrm{Ber}(p)).$$

Our goal is to show that, for any sequence of schemes with vanishing $\pi_{0|1}$ and $\pi_{1|0}$ and satisfying (3) as $T \to \infty$ with $B = rT$, $E_0 \leq \mathsf{D}^*_0(r)$ and $E_0 \leq \mathsf{D}^*_1(r)$.

Let us deal with E_0 first. From (3), the key is to upper bound $\log\frac{1}{\pi_{1|0}}$. From the definition of the above notations,

$$\log\tfrac{1}{\pi_{1|0}} = \tfrac{1}{\pi_{1|1}}(d\left(\pi_{1|1}\,\|\pi_{1|0}\right) + h(\pi_{1|1}) + \pi_{0|1}\log\pi_{0|0}). \quad (15)$$

By the data processing inequality,

$$d\left(\pi_{1|1}\,\|\pi_{1|0}\right) \leq D\left(\mathsf{P}^{(\tau)}_1\,\middle\|\,\mathsf{P}^{(\tau)}_0\right) \quad (16)$$
$$= \mathbb{E}_1\left[S_\tau\right] = \textstyle\sum_{j\in[n]}\mathbb{E}\left[S_{\tau,j}\right], \quad (17)$$

where

$$S_{t,j} \triangleq \sum_{s=1}^{t}\mathbb{1}\{W_s = j\}\delta_s\log\frac{P_{1,j}(X_s)}{P_{0,j}(X_s)}$$

denotes the LLR with only samples collected from source j.

Now, by the Optional Stopping Theorem, we have

$$\mathbb{E}_1\left[S_{\tau,j}\right] = \mathbb{E}_1\left[\textstyle\sum_{s\in\{1,\dots,\tau\}}\mathbb{1}\{W_s = j\}\delta_s\right]\mathbb{E}_1\left[\log\tfrac{P_{1,j}(X)}{P_{0,j}(X)}\right]$$
$$= \mathbb{E}_1\left[\mathcal{B}_j\right]\,D\left(P_{1,j}\,\|P_{0,j}\right),$$

where we denote the number of collected samples from source $j \in [n]$ as $\mathcal{B}_j \triangleq \sum_{s\leq\tau}\mathbb{1}\{W_s = j\}\delta_s$. Note that $\mathbb{E}_1\left[\mathcal{B}_j\right] \leq \mathbb{E}_1\left[\tau\right]\alpha_j \leq T\alpha_j$, and $\sum_{j\in[n]}\mathbb{E}_1\left[\mathcal{B}_j\right] = \mathbb{E}_1\left[\mathcal{B}\right] \leq B$. As a result, with the above notations, we can find the following inequalities which are necessary conditions for the original budget and temporal constraints:

$$\mathbb{E}_1\left[\mathcal{B}_j\right] \leq T\alpha_j, \quad \forall\, j \in [n] \quad (18)$$
$$\textstyle\sum_{j\in[n]}\mathbb{E}_1\left[\mathcal{B}_j\right] \leq B. \quad (19)$$

Hence, we can derive an upper bound on $\mathbb{E}_1\left[S_\tau\right]$ as follows:

$$\begin{aligned}&\mathbb{E}_1\left[S_\tau\right]\\ &= \textstyle\sum_{j\in[n]}\mathbb{E}\left[S_{\tau,j}\right]\\ &\leq \max_{\substack{\mathbb{E}_1[\mathcal{B}_1],\dots,\mathbb{E}_1[\mathcal{B}_n]\\ \text{satisfying (18) and (19)}}}\left\{\sum_{j\in[n]}\mathbb{E}_1\left[\mathcal{B}_j\right]D\left(P_{1,j}\,\|P_{0,j}\right)\right\}. \quad (20)\end{aligned}$$

Combining (15) – (17) and (20), an upper bound on $\frac{1}{T}\log\frac{1}{\pi_{1|0}}$ is reached:

$$\tfrac{1}{T}\log\tfrac{1}{\pi_{1|0}} \leq \tfrac{1}{T\pi_{1|1}}(h(\pi_{1|1}) + \pi_{0|1}\log\pi_{0|0} + (20)). \quad (21)$$

Taking $T \to \infty$ with $B = rT$, since $\pi_{0|1}$ and $\pi_{1|0}$ both vanish, the upper bound (21) asymptotically becomes

$$\max_{\boldsymbol{\beta}\,:\,\|\boldsymbol{\beta}\|\leq r,\,0\preceq\boldsymbol{\beta}\preceq\boldsymbol{\alpha}}\;\textstyle\sum_{j\in[n]}\beta_j D\left(P_{1,j}\,\|P_{0,j}\right)$$

which is exactly $\mathsf{D}^*_0(r)$. This shows that $E_0 \leq \mathsf{D}^*_0(r)$.

As for E_1, a similar argument shows that $E_1 \leq \mathsf{D}^*_1(r)$, and the converse proof is now complete. ■

References

[1] A. Wald, "Sequential tests of statistical hypotheses," *The Annals of Mathematical Statistics*, vol. 16, no. 2, pp. 117–186, June 1945.

[2] A. Wald and J. Wolfowitz, "Optimum character of the sequential probability ratio test," *The Annals of Mathematical Statistics*, vol. 19, no. 3, pp. 326–339, September 1948.

[3] P. Armitage, "Sequential analysis with more than two alternative hypotheses, and its relation to discriminant function analysis." *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 12, no. 1, pp. 137–144, 1950.

[4] V. P. Dragalin, A. G. Tartakovsky, and V. V. Veeravalli, "Multihypothesis sequential probability ratio tests—part i: Asymptotic optimality," *IEEE Transactions on Information Theory*, vol. 45, no. 7, pp. 2448–2461, November 1999.

[5] A. Tartakovsky, I. Nikiforov, and M. Basseville, *Sequential Analysis: Hypothesis Testing and Changepoint Detection*. CRC Press, 2014.

[6] H. Chernoff, "Sequential design of experiments," *The Annals of Mathematical Statistics*, vol. 30, no. 3, pp. 755–770, 1959.

[7] M. Naghshvar and T. Javidi, "Active sequential hypothesis testing," *The Annals of Statistics*, vol. 41, no. 6, pp. 2703–2738, 2013.

[8] S. Nitinawarat, K. Atia, George, and V. Veeravalli, Venugopal, "Controlled sensing for multihypothesis testing," *IEEE Transactions on automatic control*, vol. 58, no. 10, pp. 2451–2464, October 2013.

[9] D. Bajovic, B. Sinopoli, and J. Xavier, "Sensor selection for event detection in wireless sensor networks," *IEEE Transactions on signal processing*, vol. 59, no. 10, pp. 4938–4953, 2011.

[10] C.-Z. Bai, V. Katewa, V. Gupta, and Y. Huang, "A stochastic sensor selection scheme for sequential hypothesis testing with multiple sensors," *IEEE Transactions on signal processing*, vol. 63, no. 14, pp. 3687–3699, July 2015.

[11] C.-Z. Bai and V. Gupta, "An on-line sensor selection algorithm for SPRT with multiple sensors," *IEEE Transactions on automatic control*, vol. 62, no. 7, pp. 3532–3539, July 2017.

[12] S. Li, X. Li, X. Wang, and J. Liu, "Sequential hypothesis test with online usage-constrained sensor selection," *IEEE Transactions on information theory*, vol. 65, no. 7, pp. 4392–4410, July 2019.

[13] B. Huang, K. Cohen, and Q. Zhao, "Active anomaly detection in heterogeneous processes," *IEEE Transactions on information theory*, vol. 65, no. 4, pp. 2284–2301, April 2019.

[14] Y. Polyanskiy and S. Verdu, "Binary hypothesis testing with feedback," in *Information Theory and Applications Workshop (ITA)*, 2011.

[15] Y. Polyanskiy and Y. Wu. (2017) Lecture notes on information theory. [Online]. Available: http://people.lids.mit.edu/yp/homepage/data/itlectures_v5.pdf

APPENDIX

A. Error probability upper bounds for the achievability proof

Note that

$$\begin{aligned}
&\mathbb{E}_1\left[e^{-S_{t+1}}\middle|e^{-S_t}\right]\\
&=\mathbb{E}_1\left[e^{-S_{t+1}}\middle|S_t\right]\\
&=e^{-S_t}\,\mathbb{E}_1\left[e^{-(S_{t+1}-S_t)}\middle|S_t\right]\\
&=e^{-S_t}\,\mathbb{E}_1\left[\exp\left(-\delta_{t+1}\log\frac{P_{1,W_{t+1}}(X_{t+1})}{P_{0,W_{t+1}}(X_{t+1})}\right)\middle|S_t\right].
\end{aligned}$$

Since $\forall\, j\in[n]$,

$$\begin{aligned}
&\mathbb{E}_1\left[e^{-\delta_{t+1}\log\frac{P_{1,W_{t+1}}(X_{t+1})}{P_{0,W_{t+1}}(X_{t+1})}}\middle|S_t,W_{t+1}=j,\delta_{t+1}=0\right]=1,\\
&\mathbb{E}_1\left[e^{-\delta_{t+1}\log\frac{P_{1,W_{t+1}}(X_{t+1})}{P_{0,W_{t+1}}(X_{t+1})}}\middle|S_t,W_{t+1}=j,\delta_{t+1}=1\right]\\
&=\mathbb{E}_1\left[\frac{P_{0,j}(X_{t+1})}{P_{1,j}(X_{t+1})}\middle|S_t,W_{t+1}=j,\delta_{t+1}=1\right]=1,
\end{aligned}$$

we have $\mathbb{E}_1\left[\exp\left(-\delta_{t+1}\log\frac{P_{1,W_{t+1}}(X_{t+1})}{P_{0,W_{t+1}}(X_{t+1})}\right)\middle|S_t\right]=1$. Hence,

$$\mathbb{E}_1\left[e^{-S_{t+1}}\middle|e^{-S_t}\right]=e^{-S_t}.$$

Following a similar derivation, it can also be shown that

$$\mathbb{E}_0\left[e^{S_{t+1}}\middle|e^{S_t}\right]=e^{S_t}.$$

Therefore, $\{e^{-S_t}\,|\,t\geq 1\}$ is a martingale under hypothesis $\mathcal{H}_1$ and $\{e^{S_t}\,|\,t\geq 1\}$ is a martingale under hypothesis $\mathcal{H}_0$. By the Optional Stopping Theorem, we have

$$1=e^{-S_0}=\mathbb{E}_1\left[e^{-S_\tau}\right],\ 1=e^{S_0}=\mathbb{E}_0\left[e^{S_\tau}\right].$$

Recall the inference rule in equation (9), we have

$$\begin{aligned}
1&=\mathbb{E}_1\left[e^{-S_\tau}\right]\\
&=\pi_{0|1}\mathbb{E}_1\left[e^{-S_\tau}\middle|\phi=0\right]+\pi_{1|1}\mathbb{E}_1\left[e^{-S_\tau}\middle|\phi=1\right]\\
&\geq\pi_{0|1}e^{A_0}+\pi_{1|1}e^{-A_1-L}\\
&\geq\pi_{0|1}e^{A_0},
\end{aligned}$$

implying that $\pi_{0|1}\leq e^{-A_0}$. Similarly, we have

$$1=\mathbb{E}_0\left[e^{S_\tau}\right]\geq\pi_{1|0}e^{A_1}\implies\pi_{1|0}\leq e^{-A_1}.$$

B. Proof of Lemma 1

The most of the proof here follows the proof in [7] (See "Proof of Lemma 4" in supplement to [7]). But we slightly adjust the lemma to directly fit our setting by consider the cases of $\{S_t\leq 0\}$ and $\{S_t>0\}$.

Proof: First note that due to the time-varying sources and adaptive policy, $\frac{S_t}{K_2}-t$ is not martingale. But we still have:

$$\begin{cases}\mathbb{E}_1\left[\frac{S_{t+1}}{K_1}-(t+1)\middle|S_t\right]\geq\frac{S_t}{K_1}-t & \text{if } S_t\leq 0\\ \mathbb{E}_1\left[\frac{S_{t+1}}{K_2}-(t+1)\middle|S_t\right]\geq\frac{S_t}{K_2}-t & \text{if } S_t>0\end{cases}\tag{22}$$

and e^{-S_t} is still martingale under hypothesis $\mathcal{H}_1$.

$$
\begin{aligned}
&\mathbb{E}_1\left[e^{-S_{t+1}}\middle|S_t\right]\\
=&e^{-S_t}\ \mathbb{E}_1\left[e^{-(S_{t+1}-S_t)}\middle|S_t\right]\\
=&e^{-S_t}\ \mathbb{E}_1\left[e^{-\delta_{t+1}\log\frac{\mathbb{P}_{1,W_{t+1}}}{\mathbb{P}_{0,W_{t+1}}}}\middle|S_t\right]\\
=&e^{-S_t}\ \sum_{j\in[n]}\sum_{x\in\mathcal{X}}P_\alpha(j)\mathbb{P}_{1,j}(x)\\
&\qquad\Bigg(\mathbb{P}\left\{\delta_{t+1}=0\,|F_t,w_{t+1}=j\right\}\\
&+\mathbb{P}\left\{\delta_{t+1}=1\,|F_t,w_{t+1}=j\right\}\ \left(\frac{\mathbb{P}_{0,j}(x)}{\mathbb{P}_{1,j}(x)}\right)\Bigg)\\
=&e^{-S_t}\ \sum_{j\in[n]}P_\alpha(j)\Bigg(\mathbb{P}\left\{\delta_{t+1}=0\,|F_t,w_{t+1}=j\right\}\\
&\quad+\ \mathbb{P}\left\{\delta_{t+1}=1\,|F_t,w_{t+1}=j\right\}\ \Bigg)\\
=&e^{-S_t}
\end{aligned}
\tag{23}
$$

A submartingale can be constructed by linear combination of these term. Let

$$
\xi_t=\begin{cases}-C+\frac{S_t}{K_1}-t & \text{if } S_t\le 0\\ -Ce^{-S_t}+\frac{S_t}{K_2}-t & \text{if } S_t>0\end{cases}
\tag{24}
$$

where $C=(\frac{K_3}{1-e^{-K_3}})(\frac{1}{K_1}-\frac{1}{K_2})>0$. We shall prove that the sequence $\{\xi_t\}$ forms a submartingale, i.e. $\mathbb{E}_1\left[\xi_{t+1}\,|\xi_t\right]\ge\xi$.

For case 1 $S_t\le 0$: If $S_{t+1}\le 0$, then

$$
\xi_{t+1}=-C+\frac{S_{t+1}}{K_1}-(t+1)
\tag{25}
$$

If $S_{t+1}>0$, then

$$
\begin{aligned}
\xi_{t+1}&=-Ce^{-S_{t+1}}+\frac{S_{t+1}}{K_2}-(t+1)\\
&\ge -C+\frac{S_{t+1}}{K_1}-(t+1)
\end{aligned}
\tag{26}
$$

The last inequality is due to the concavity of $-Ce^{-S_{t+1}}+\frac{S_{t+1}}{K_2}$ over $S_{t+1}\in(0,K_3]$. Note that C is carefully picked to let $-Ce^{-S_{t+1}}+\frac{S_{t+1}}{K_2}\big|_{S_{t+1}=K_3}=-C+\frac{S_{t+1}}{K_1}\big|_{S_{t+1}=K_3}$ and $-Ce^{-S_{t+1}}+\frac{S_{t+1}}{K_2}\big|_{S_{t+1}=0}=-C+\frac{S_{t+1}}{K_1}\big|_{S_{t+1}=0}$. Hence we have for $S_t\le 0$

$$
\begin{aligned}
&\mathbb{E}_1\left[\xi_{t+1}\,|S_t\right]\\
\ge&\mathbb{E}_1\left[-C+\frac{S_{t+1}}{K_1}-(t+1)\middle|S_t\right]\\
\ge&-C+\frac{S_t}{K_1}-t=\xi_t
\end{aligned}
\tag{27}
$$

For case 2 $S_t>0$: If $S_{t+1}>0$, then

$$
\xi_{t+1}=-Ce^{-S_{t+1}}\frac{S_{t+1}}{K_2}-(t+1)
\tag{28}
$$

If $S_{t+1}\le 0$, then

$$
\begin{aligned}
\xi_{t+1}&=-C+\frac{S_{t+1}}{K_1}-(t+1)\\
&\ge -Ce^{-S_{t+1}}+\frac{S_{t+1}}{K_2}-(t+1)
\end{aligned}
\tag{29}
$$

The last inequality is due to the concavity of $-Ce^{-S_{t+1}} + \frac{S_{t+1}}{K_2}$ over $S_{t+1} \in (-\infty, K_3]$. Hence we have for $S_t > 0$

$$\begin{aligned}
&\mathbb{E}_1 \left[\xi_{t+1} \left| S_t \right.\right] \\
\geq &\mathbb{E}_1 \left[-Ce^{-S_{t+1}} + \frac{S_{t+1}}{K_2} - (t+1) \middle| S_t \right] \\
\geq &- C\mathbb{E}_1 \left[e^{-S_{t+1}} \left| S_t \right.\right] + \mathbb{E}_1 \left[\frac{S_{t+1}}{K_2} - (t+1) \middle| S_t \right] \\
\geq &- Ce^{-S_t} + \frac{S_t}{K_2} - t
\end{aligned} \tag{30}$$

The last inequality is by martingale e^{-S_t} under hypothesis $\mathcal{H}_1$ (See Equation 24) and by $\mathbb{E}_1 \left[S_{t+1} - S_t \left| S_t > 0 \right.\right] \geq K_2$.

We conclude that $\{\xi_t\}$ is submartingale. Hence by optional stopping theorem, we have

$$-C = \xi_0 \leq \mathbb{E}_1 \left[\xi_\tau\right] = \mathbb{E}_1 \left[-Ce^{-S_\tau} + \frac{S_\tau}{K_2} - \tau \right] \tag{31}$$

Recall the stopping rule is $\tau : \min\{S_t \geq A\}$. The LLR at stopping time is upper bounded. $S_\tau = S_{\tau-1} + (S_\tau - S_{\tau-1}) \leq A + K_3$. Therefore,

$$\mathbb{E}_1 \left[-Ce^{-S_\tau} + \frac{S_\tau}{K_2} - \tau \right] \leq -Ce^{-A-K_3} + \frac{A+K_3}{K_2} - \mathbb{E}_1 \left[\tau\right] \tag{32}$$

Plug-in the value of $C = \frac{K_3}{1-e^{-K_3}} \left(\frac{1}{K_1} - \frac{1}{K_2} \right)$, the expected stopping time is upper bounded.

$$\begin{aligned}
\mathbb{E}_1 \left[\tau\right] &\leq C \left(1 - e^{-A-K_3}\right) + \frac{A+K_3}{K_2} \\
&< \frac{K_3}{1-e^{-K_3}} \left(\frac{1}{K_1} - \frac{1}{K_2} \right) + \frac{A+K_3}{K_2}
\end{aligned} \tag{33}$$

The proof of the lemma is complete. ■