

# On Universal Sequential Classification from Sequentially Observed Empirical Statistics

Chia-Yu Hsu  
Realtek Semiconductor Corp.,  
Hsinchu, Taiwan  
Email: zxc586794a@gmail.com

Ching-Fang Li  
GIEE, National Taiwan University,  
Taipei, Taiwan  
Email: r10921046@ntu.edu.tw

I-Hsiang Wang  
GICE, National Taiwan University,  
Taipei, Taiwan  
Email: ihwang@ntu.edu.tw

**Abstract**—The focus of this paper is on binary classification of a sequentially observed stream of i.i.d. samples, based on sequentially observed empirical statistics. The decision maker (classifier) sequentially observes a sequence of testing data i.i.d. sampled from one of two unknown distributions  $P_0$  and  $P_1$ . In addition, it receives two sequences of training data which are also sequentially sampled from the two unknown distributions in an i.i.d. fashion, respectively. Since the distributions are unknown, it is natural to put a constraint either on the expected stopping times or on the error probabilities that has to be satisfied *universally* over all possible pairs of distributions  $(P_0, P_1)$ . For both settings, we develop tests that are asymptotically optimal within the class of tests that satisfy the respective universality constraints. For expected-stopping-time universality, the optimal error exponents are shown to be the Rényi divergences of order  $\frac{\alpha}{1+\alpha}$ , where  $\alpha$  is the ratio of the length of a training data sequence to that of the testing data sequence. For error-probability universality, the optimal expected stopping times normalized by the logarithm of the error probability are the reciprocal of the  $\alpha$ -weighted generalized Jensen-Shannon (GJS) divergences. The proposed sequential tests are both based on threshold tests of two statistics, each of which is the  $\alpha$ -weighted GJS divergences of the type of the testing sequence from that of a training sequence. Interestingly, to achieve asymptotic optimality, the stopping and decision rules in the two different universality setups take a “matching” and a “discriminating” viewpoint respectively in designing the threshold tests.

## I. INTRODUCTION

For hypothesis testing problems, it is well known that sequentiality in taking samples for making decisions achieves significant enhancement in the reliability. For the binary hypothesis testing problem of an i.i.d. sampled sequence from either  $P_0$  (hypothesis  $\mathcal{H}_0$ ) or  $P_1$  (hypothesis  $\mathcal{H}_1$ ), Wald’s Sequential Probability Ratio Test (SPRT) [1] has been shown to be optimal in the sense that among all sequential tests with the same power, it achieves the smallest expected stopping time [2]. To clearly see the benefit of sequentiality, alternatively one could set a *constraint* on the expected stopping time to be less than equal to a *nominal length*  $n$ , and investigate the error exponents, which are the rate functions of the vanishing-with- $n$  type-I and type-II error probabilities. In sharp contrast to the non-sequential (fixed-length) binary hypothesis testing problem where there is a natural trade-off between the two error exponents [3], in the sequential setting, such a trade-off is completely eradicated, and the achievable error exponents can reach the two extremes, namely, the two KL divergences

$D(P_1||P_0)$  and  $D(P_0||P_1)$ , simultaneously. Such observations have been made in the literature [2], [4], [5].

While  $P_0$  and  $P_1$  are critical for computing the statistics in SPRT, in real-world problems such as machine learning and clinical trials,  $P_0$  and  $P_1$  may no longer be fully known. An interesting question is whether or not the benefit of sequentiality carry to the case when the ground truth distributions ( $P_0$  and  $P_1$ ) are unknown. Since  $P_0$  and  $P_1$  are unknown, the above question should be addressed under a *universal* guarantee on certain performances. The natural choice in this sequential setup is to lay the universality constraint either on the two expected stopping times, or on the two error probabilities.

Towards answering this question, the most closely related work in the literature is [6], in which Haghifam *et al.* consider the classification of a sequentially observed stream of i.i.d. samples (termed the “testing sequence” hereafter) into one of the multiple unknown hypotheses. Each of the hypotheses corresponds to an unknown distribution, based on multiple non-causally observed *fixed-length* sequences of i.i.d. “training” samples (termed the “training sequences” hereafter) respectively drawn from each of these distributions. The problem formulation can be viewed as a *semi-sequential* version of Gutman’s fixed-length setup [7], and the proposed sequential test was shown to improve the Bayesian error exponent over the non-sequential case.

However, Haghifam *et al.* [6] are not concerned about the universal guarantee on the expected stopping time, and it turns out that the expected stopping time of the proposed sequential test in [6] depends implicitly on the unknown pair of distributions  $(P_0, P_1)$ . As a result, sequential tests with a universal guarantee either on the expected stopping times or on the error probabilities over all possible pairs of distributions remain open, let alone the existence of optimal sequential tests subject to respective universality constraints.

In this work, we make progress in a *fully sequential* setup of the binary classification problem with empirically observed statistics [7], where the testing sequence and the two training sequences are all sequentially observed. At each time slot, the number of training samples received so far is assumed to be linearly scaling with that of the testing samples received so far, where the ratio is assumed to be  $\alpha > 0$ . Our contributions are summarized as follows. Firstly, for the setup with expected-stopping-time universality, we propose a

type-based fully sequential test, called the Sequential Type Matching Test (STMT), and show that its expected stopping time is not greater than  $n$  regardless of the ground-truth hypothesis and the pair of ground-truth distributions. The decision rule is inspired by [6], [7], while the main difference is that the decision is declared whenever the  $\alpha$ -weighted GJS divergence of the type of the testing sequence from the type of the training sequence drawn from the *accepted* distribution is *smaller than a time-varying threshold*. Furthermore, it is shown to achieve exponentially vanishing-with- $n$  type-I and type-II error probabilities, with the largest error exponents among all sequential tests satisfying the universal expected stopping time constraint. The optimal error exponents turn out to be Rényi divergences  $D_{\frac{\alpha}{1+\alpha}}(P_1\|P_0)$  and  $D_{\frac{\alpha}{1+\alpha}}(P_0\|P_1)$ .

Secondly, for the setup with error-probability universality, we propose an alternative type-based fully sequential test, called the Sequential Type Discriminating Test (STDT), and show that its error probability is universally upper bounded by a prescribed constant  $\beta$  regardless of the ground-truth hypothesis and the pair of ground-truth distributions. The decision rule is similar to [6]–[8] where the decision is declared whenever the  $\alpha$ -weighted GJS divergence of the type of the testing sequence from the type of the training sequence drawn from the *rejected* distribution is *larger than a time-varying threshold depending on  $\beta$* . Furthermore, it is shown to achieve the minimum asymptotical scaling of the expected stopping times among all sequential tests satisfying the universal error probability constraint. The optimal expected stopping times turn out to be asymptotically  $(\text{GJS}(P_1, P_0, \alpha))^{-1} \log(1/\beta)$  and  $(\text{GJS}(P_0, P_1, \alpha))^{-1} \log(1/\beta)$  as  $\beta \rightarrow 0$  under hypotheses  $\mathcal{H}_0$  and  $\mathcal{H}_1$  respectively.

Finally, to see the benefit of sequentiality, we compare Gutman’s fixed-length test and STMT in terms of their error exponent regions subject to the same nominal  $n \rightarrow \infty$ , similar to what we do in comparing fixed-length hypothesis testing and sequential hypothesis testing. We show that  $D_{\frac{\alpha}{1+\alpha}}(P_1\|P_0) \geq \text{GJS}(P_0, P_1, \alpha)$ , and hence the error exponent region of STMT subsumes that of Gutman’s test.

*Notation:* A finite-length sequence  $(x_1, x_2, \dots, x_n)$  is denoted as  $x^n$ . Logarithms are of base 2 if not specified.  $\mathcal{P}(\mathcal{X})$  is the set of all probability distributions over alphabet  $\mathcal{X}$ . We denote the support of a distribution  $P$  as  $\text{supp}(P)$ .

## II. PROBLEM FORMULATION

Consider a finite alphabet  $\mathcal{X}$  with  $|\mathcal{X}| = d \geq 2$  and two distributions  $P_0, P_1 \in \mathcal{P}(\mathcal{X})$ . We assume

$$P_0 \neq P_1, \text{supp}(P_0) = \text{supp}(P_1) = \mathcal{X} \quad (1)$$

to avoid degenerated cases where the divergences may become 0 or  $\infty$ . Furthermore,  $P_0$  and  $P_1$  are unknown to the decision maker. The decision maker observes three mutually independent sequences including a testing sequence and two training sequences. The testing sequence  $\{X_k\}_{k \geq 1}$  consists of i.i.d. samples following  $P_\theta$ , where  $\theta$  equals to either 0 or 1, and  $\mathcal{H}_\theta$  is the ground truth hypothesis. The two training sequences  $\{T_{0,k}\}_{k \geq 1}$  and  $\{T_{1,k}\}_{k \geq 1}$  consist of i.i.d. samples

following  $P_0$  and  $P_1$  respectively. The ratio of the number of training samples to the number of testing samples converges to a positive constant  $\alpha$  as they tend to infinity. More specifically, at time  $k$ , there are  $k$  testing samples and  $N_k = \lceil \alpha k \rceil$  training samples from each distribution, denoted as  $(X^k, T_0^{N_k}, T_1^{N_k})$ . The objective of the decision maker is to decide whether  $\theta$  is 0 or 1, based on the observed samples. Here we first give the formal definition of such tests.

*Definition 1 (Sequential Classification Test):* A sequential classification test is a pair  $\Phi = (\tau, \delta)$  where

- $\tau \in \mathbb{N}$  is a stopping time with respect to the filtration  $\{\mathcal{F}_k\}_{k \geq 1}$ , that is,  $\{\tau \leq k\} \in \mathcal{F}_k = \sigma(X^k, T_0^{N_k}, T_1^{N_k})$ .
- The decision rule  $\delta : (X^\tau, T_0^{N_\tau}, T_1^{N_\tau}) \rightarrow \{0, 1\}$  is a  $\mathcal{F}_\tau$ -measurable function.

To evaluate the performance of a sequential classification test, we consider the error probability and the expected stopping time of a sequential classification test.

- The type-I and type-II error probabilities are

$$\begin{aligned} \pi_{1|0}(\Phi) &= \mathbb{P}_0\{\delta(X^\tau, T_0^{N_\tau}, T_1^{N_\tau}) = 1\} \\ \text{and } \pi_{0|1}(\Phi) &= \mathbb{P}_1\{\delta(X^\tau, T_0^{N_\tau}, T_1^{N_\tau}) = 0\} \end{aligned}$$

respectively.  $\mathbb{P}_\theta$  is the shorthand notation for the joint probability law of the testing samples and the training samples when the ground truth hypothesis is  $\mathcal{H}_\theta$ .

- The expected stopping time is denoted as  $\mathbb{E}_\theta[\tau]$  when the ground truth hypothesis is  $\mathcal{H}_\theta$ .

Since  $P_0$  and  $P_1$  are unknown, it is natural to ask for a universal guarantee on the above performance metrics. We are also interested in the asymptotic performance of a family of tests subject to such a universality constraint. The asymptotes can be compactly described by a pair of error exponents. The two universality constraints and the corresponding error exponents are defined as follows.

1) *Universality Constraint on the Expected Stopping Time:* Given  $n \in \mathbb{N}$  and a sequential classification test  $\Phi$ , we say that  $\Phi$  satisfies the universality constraint on the expected stopping time with  $n$  if  $\forall (P_0, P_1)$  satisfying (1),

$$\mathbb{E}_0[\tau] \leq n \quad \text{and} \quad \mathbb{E}_1[\tau] \leq n. \quad (2)$$

*Definition 2 (Error Exponents subject to Expected-Stopping-Time Universality Constraint):* Consider a family of tests  $\{\Phi_n\}$  with vanishing error probability for all  $(P_0, P_1)$  satisfying (1). If each  $\Phi_n$  satisfies the universality constraint on the expected stopping time with  $n$ , we say that  $\{\Phi_n\}$  satisfy the expected-stopping-time universality constraint. The error exponents of  $\{\Phi_n\}$  are defined as

$$e_0 = \liminf_{n \rightarrow \infty} \frac{-\log \pi_{1|0}(\Phi_n)}{n}, \quad e_1 = \liminf_{n \rightarrow \infty} \frac{-\log \pi_{0|1}(\Phi_n)}{n}.$$

2) *Universality Constraint on the Error Probability:* Given  $\beta \in (0, 1)$  and a sequential classification test  $\Phi$ , we say that  $\Phi$  satisfies the universality constraint on the error probability with  $\beta$  if  $\forall (P_0, P_1)$  satisfying (1),

$$\pi_{1|0}(\Phi) \leq \beta \quad \text{and} \quad \pi_{0|1}(\Phi) \leq \beta. \quad (3)$$

*Definition 3 (Error Exponents subject to Error-Probability Universality Constraint):* For a family of tests  $\{\Phi_\beta\}$  where each  $\Phi_\beta$  satisfies the universality constraint on the error probability with  $\beta$ , we say that  $\{\Phi_\beta\}$  satisfy the error-probability universality constraint. The error exponents of  $\{\Phi_\beta\}$  are defined as

$$\tilde{e}_0 = \liminf_{\beta \rightarrow 0} \frac{-\log \beta}{\mathbb{E}_0[\tau]}, \quad \tilde{e}_1 = \liminf_{\beta \rightarrow 0} \frac{-\log \beta}{\mathbb{E}_1[\tau]}.$$

*Remark 1:* Definition 2 and 3 consider the limits of error probabilities in two different asymptotic regimes. Similar definitions have appeared in the literature of sequential hypothesis testing. For example, [4], [5] consider error exponents similar to Definition 2, while [6], [8]–[10] adopt asymptotic performance metrics similar to Definition 3.

### III. PRELIMINARIES

In this section, let us first briefly revisit the results in sequential hypothesis testing [1], [2] and Gutman’s fixed-length classification problem from empirically observed statistics [7], in terms of the error exponents. They serve as the baselines of our main results. Let us begin with a few necessary definitions for the exposition of the results.

*Definition 4 (Type (Empirical Distribution)):* Consider a sequence  $x^k$ , where each testing sample  $x_i \in \mathcal{X}$ . The type (empirical distribution) of  $x^k$  is denoted as  $\Pi_{x^k}$ , where

$$\Pi_{x^k}(a) = \frac{1}{k} \sum_{i=1}^k \mathbb{1}\{x_i = a\}, \quad a \in \mathcal{X}.$$

*Definition 5 (Generalized Jensen-Shannon Divergence):* Given  $\alpha > 0$  and  $P, Q \in \mathcal{P}(\mathcal{X})$ , the  $\alpha$ -weighted generalized Jensen-Shannon (GJS) divergence of  $Q$  from  $P$  is defined as

$$\text{GJS}(P, Q, \alpha) = \alpha \text{D}\left(P \left\| \frac{\alpha P + Q}{\alpha + 1}\right.\right) + \text{D}\left(Q \left\| \frac{\alpha P + Q}{\alpha + 1}\right.\right).$$

*Definition 6 (Rényi Divergence):* Given  $a \in (0, 1) \cup (1, \infty)$  and  $P, Q \in \mathcal{P}(\mathcal{X})$ , the Rényi Divergence of order  $a$  of  $P$  from  $Q$  is defined as

$$\text{D}_a(P \| Q) = \frac{1}{a-1} \sum_{x \in \mathcal{X}} P(x)^a Q(x)^{1-a}.$$

#### A. Sequential Hypothesis Testing

In the classical sequential hypothesis testing problem where  $P_0$  and  $P_1$  are known, the asymptotic performance of SPRT (the optimal test [2]) could be investigated subject to the expected-stopping-time constraint as in Definition 2 or the error-probability constraint as in Definition 3. It turns out that the corresponding pairs of error exponents are not the same. The following proposition summarizes this observation.

*Proposition 1:* A family of SPRT indexed by  $n \in \mathbb{N}$  with proper selection of the upper and lower threshold so that (2) is satisfied for all  $n \in \mathbb{N}$  achieve

$$e_0 = \text{D}(P_1 \| P_0), \quad e_1 = \text{D}(P_0 \| P_1). \quad (4)$$

A family of SPRT indexed by  $\beta \in (0, 1)$  with proper selection of the upper and lower threshold so that (3) is satisfied for all  $\beta \in (0, 1)$  achieve

$$\tilde{e}_0 = \text{D}(P_0 \| P_1), \quad \tilde{e}_1 = \text{D}(P_1 \| P_0). \quad (5)$$

For the sake of completeness, the proof of this proposition is given in Appendix A of the extended version [11].

Proposition 1 tells that for the sequential hypothesis testing problem, the two pairs of optimal error exponents in the two different regimes are flipped in the sense that  $e_0 = \tilde{e}_1$  and  $e_1 = \tilde{e}_0$ . As shown in our main results, this difference remains in the sequential classification problem. Moreover, the former ( $e_0$  and  $e_1$ ) change from KL divergences to Rényi divergences, while the latter ( $\tilde{e}_0$  and  $\tilde{e}_1$ ) change from KL divergences to GJS divergences.

#### B. Fixed-Length Classification

In the fixed-length setting, the length of the testing sequence is  $n$ , and the length of each training sequence is  $N = \lceil \alpha n \rceil$ . Let  $\lambda \in \mathbb{R}_+$ , Gutman’s decision rule [7]

$$\Phi_\lambda^{(\text{Gut})}(x^n, t_0^N, t_1^N) = \begin{cases} 0 & \text{if } \text{GJS}\left(\Pi_{t_0^N}, \Pi_{x^n}, \alpha\right) \leq \lambda, \\ 1 & \text{if } \text{GJS}\left(\Pi_{t_0^N}, \Pi_{x^n}, \alpha\right) > \lambda, \end{cases}$$

achieves error exponents  $e_0(\Phi_\lambda^{(\text{Gut})}) = \lambda$ ,  $e_1(\Phi_\lambda^{(\text{Gut})}) = F(\alpha, \lambda)$  where

$$F(\alpha, \lambda) = \min_{\substack{(Q_0, Q_1) \in \mathcal{P}(\mathcal{X})^2 \\ \text{GJS}(Q_0, Q_1, \alpha) \leq \lambda}} \alpha \text{D}(Q_0 \| P_0) + \text{D}(Q_1 \| P_1).$$

Furthermore, Gutman’s test is asymptotically optimal in the sense that among all  $(e_0, e_1)$ -achievable tests with  $e_0 \geq \lambda$  for all pairs of distinct distributions  $(P_0, P_1)$ ,  $e_1 \leq e_1(\Phi_\lambda^{(\text{Gut})})$ . In other words, the universality constraint is set on the type-I error exponent, and Gutman’s test attains *instance-optimality* for the type-II error exponent, that is, different pairs of  $(P_0, P_1)$  give different optimal type-II error exponents.

One may view  $\{(\lambda, F(\alpha, \lambda))\}$  as the optimal trade-off between the two error exponents, reminiscent of that in binary hypothesis testing [3]. See Figure 1 for an illustration. Note that as observed in [7], [12], the supremum of  $\lambda$  that allows a constant type-II error probability less than 1 is  $\text{GJS}(P_0, P_1, \alpha)$  (the square in Figure 1), while the supremum of  $F(\alpha, \lambda)$  over  $\lambda > 0$  is  $\text{D}_{\frac{\alpha}{1+\alpha}}(P_0 \| P_1)$  (the diamond in Figure 1).

As shown in our main results, Rényi divergences appear in the instance-optimal error exponents when the universality constraint is set on the expected stopping times, while GJS divergences show up when the universality constraint is set on the error probabilities. It is interesting to note that the  $\alpha$ -weighted GJS divergence has the following variational form:

$$\text{GJS}(P, Q, \alpha) = \min_{V \in \mathcal{P}(\mathcal{X})} \{\alpha \text{D}(P \| V) + \text{D}(Q \| V)\}. \quad (6)$$

Meanwhile, in the objective function above, if  $P, V$  are swapped in the first KL divergence and  $Q, V$  are swapped in the second, the Rényi Divergence of order  $\frac{\alpha}{1+\alpha}$  emerges:

$$\text{D}_{\frac{\alpha}{1+\alpha}}(P \| Q) = \min_{V \in \mathcal{P}(\mathcal{X})} \{\alpha \text{D}(V \| P) + \text{D}(V \| Q)\}. \quad (7)$$

Such duality has been observed in the literature of fixed-length classification problems [12], [13]. It turns out that  $\text{GJS}(\cdot, \cdot, \alpha)$  and  $\text{D}_{\frac{\alpha}{1+\alpha}}(\cdot \| \cdot)$  play a dual role as well in the sequential setup.

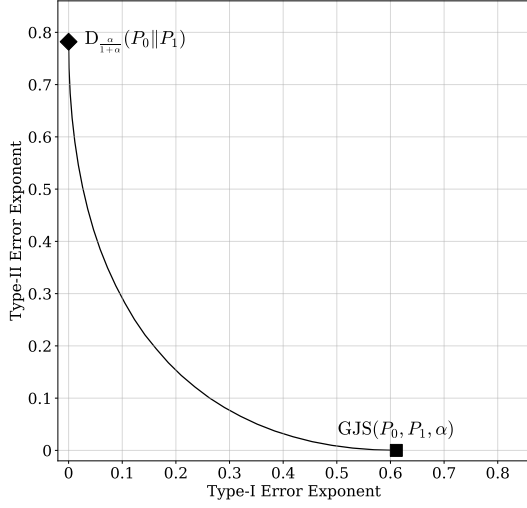


Fig. 1: The error exponents of Gutman's test under ground truth distributions  $P_0 = [0.4, 0.6]$ ,  $P_1 = [0.9, 0.1]$ , and  $\alpha = 3$ .

#### IV. MAIN RESULTS

Our main result is the characterization of the optimal error exponents of the sequential classification problem under either the expected-stopping-time universality as in Definition 2, or the error-probability universality as in Definition 3. In particular, two type-based tests (tests that declare stopping and make decision only based on the testing type and the training types) are proposed. The two proposed tests satisfy (2) and (3) respectively, no matter what the ground truth distributions are. Moreover, they achieve the optimal error exponents. The converse part (asymptotic optimality) of the two main theorems are proved in Section V. The proofs of the achievability parts are based on the method of types [14] and left in Appendix B and C of the extended version [11].

##### A. Expected-Stopping-Time Universality

*Definition 7 (Sequential Type Matching Test):* Given integer  $n \geq 2$ , define STMT  $\Phi_n^{(M)} = (\tau, \delta)$  as

$$\tau = \inf\{k \geq n - 1 : \text{GJS}(\Pi_{t_0^k}, \Pi_{x^k}, \alpha) \leq f(k) \\ \text{or } \text{GJS}(\Pi_{t_1^k}, \Pi_{x^k}, \alpha) \leq f(k)\},$$

$$\delta(x^\tau, t_0^{N_\tau}, t_1^{N_\tau}) = \begin{cases} 0 & \text{if } \text{GJS}(\Pi_{t_0^{N_\tau}}, \Pi_{x^\tau}, \alpha) \leq f(\tau), \\ 1 & \text{if } \text{GJS}(\Pi_{t_1^{N_\tau}}, \Pi_{x^\tau}, \alpha) \leq f(\tau), \end{cases}$$

where  $f(k) = \frac{2d \log(k+1)}{k} + \frac{d \log(N_k+1)}{k}$ .

Intuitively, STMT declares the decision if the type of the testing sequence is “close enough” to the type of one of the two training sequences, where GJS divergence measures “how far” two distributions are. Note that when the number of samples is still small, the resolution of a type is too coarse and the threshold is loose so that the GJS divergences have higher probability to get below the threshold at the wrong side. To avoid stopping too early and making wrong decisions, STMT always first observes a fixed number of samples.

The asymptotic optimality of STMT subject to the expected-stopping-time universality constraint is summarized in the following theorem.

*Theorem 1:* Let  $(P_0, P_1) \in \mathcal{P}(\mathcal{X})^2$  be the ground truth distributions satisfying (1). For any family of tests  $\{\Phi_n\}$  satisfying the expected-stopping-time universality constraint,

$$e_0 \leq D_{\frac{\alpha}{1+\alpha}}(P_1||P_0) \text{ and } e_1 \leq D_{\frac{\alpha}{1+\alpha}}(P_0||P_1).$$

Moreover, STMT  $\Phi_n^{(M)}$  satisfies the universality constraint on the expected stopping time with  $n$ , and  $\{\Phi_n^{(M)}\}_{n \geq 2}$  achieves these upper bounds.

*Remark 2:* As  $\alpha \rightarrow \infty$ , the optimal point converges to  $(D(P_1||P_0), D(P_0||P_1))$ , the same as (4).

*Remark 3:* It is worth noting that STMT does not stop earlier than  $n - 1$ . Meanwhile, it still satisfies the expected-stopping-time universality constraint with  $n$  and improves the error exponents comparing to the fixed-length tests. This is because in the fixed length setting, the instances that result in error are rare. Spending longer time on these instances would not add much to the expected stopping time. The details can be found in Appendix B of the extended version [11].

*Remark 4 (Benefit of sequentiality):* STMT achieves  $e_0 = D_{\frac{\alpha}{1+\alpha}}(P_1||P_0)$ ,  $e_1 = D_{\frac{\alpha}{1+\alpha}}(P_0||P_1)$ . Compared with Gutman's result in Section III-B, one could assert the gain in error exponents by sequentiality if  $D_{\frac{\alpha}{1+\alpha}}(P_1||P_0) \geq \text{GJS}(P_0, P_1, \alpha)$ . Indeed this is true and proved in Appendix D of the extended version [11].

##### B. Error-Probability Universality

*Definition 8 (Sequential Type Discriminating Test):* Given  $\beta \in (0, 1)$ , define STDT  $\Phi_\beta^{(D)} = (\tau, \delta)$  as

$$\tau = \inf\{k \in \mathbb{N} : \text{GJS}(\Pi_{t_0^k}, \Pi_{x^k}, \alpha) > g(\beta, k) \\ \text{or } \text{GJS}(\Pi_{t_1^k}, \Pi_{x^k}, \alpha) > g(\beta, k)\},$$

$$\delta(x^\tau, t_0^{N_\tau}, t_1^{N_\tau}) = \begin{cases} 0 & \text{if } \text{GJS}(\Pi_{t_1^{N_\tau}}, \Pi_{x^\tau}, \alpha) > g(\beta, \tau), \\ 1 & \text{if } \text{GJS}(\Pi_{t_0^{N_\tau}}, \Pi_{x^\tau}, \alpha) > g(\beta, \tau), \end{cases}$$

where  $g(\beta, k) = \frac{-\log(\beta(d-1))}{k} + f(k)$ .

In contrast to STMT, STDT declares the decision if the type of the testing sequence is “far enough” from the type of the other training sequence. To decide how far is enough, we use a threshold  $g(\beta, k)$ , which is time-dependent and grows as  $\beta$  decreases. The threshold is chosen judiciously to achieve the target error probabilities.

The asymptotic optimality of STDT subject to the error-probability universality constraint is summarized in the following theorem.

*Theorem 2:* Let  $(P_0, P_1) \in \mathcal{P}(\mathcal{X})^2$  be the ground truth distributions satisfying (1). For any family of tests  $\{\Phi_\beta\}$  satisfying the error-probability universality constraint,

$$\tilde{e}_0 \leq \text{GJS}(P_1, P_0, \alpha) \text{ and } \tilde{e}_1 \leq \text{GJS}(P_0, P_1, \alpha).$$

Moreover, STDT  $\Phi_\beta^{(D)}$  satisfies the universality constraint on error probability with  $\beta$ , and  $\{\Phi_\beta^{(D)}\}_{\beta \in (0,1)}$  achieves these upper bounds.

*Remark 5:* As  $\alpha \rightarrow \infty$ , the optimal point converges to  $(D(P_0\|P_1), D(P_1\|P_0))$ , the same as (5).

*Remark 6:* STDT is modified from the test proposed in [8], and the achievability proof is essentially the same. In [8], they consider an outlier detection problem where the distributions are unknown. The test plugs in the empirical distribution to the likelihood ratio test. Through some calculations, the likelihood ratio in [8] can be written as  $k$  times the subtraction of two GJS divergences between the empirical distributions. Moreover, it turns out that we can omit the second GJS divergence. The details are given in Appendix C of the extended version [11].

Figure 2 gives an illustration about how the optimal exponents in the two settings grow as  $\alpha$  increases.

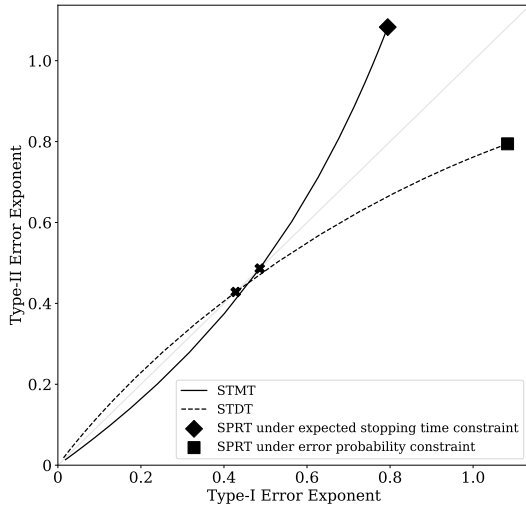


Fig. 2: The error exponents of STMT and STDT with different  $\alpha$  under ground truth distributions  $P_0 = [0.4, 0.6]$ ,  $P_1 = [0.9, 0.1]$ . The error exponents of STMT and STDT approach the error exponents of SPRT under two constraints respectively, as  $\alpha$  increases from  $1.5^{-10}$  to  $1.5^{19}$ . The two marked crosses in the middle correspond to  $\alpha = 1$ .

## V. PROOF OF CONVERSE

We first present two lemmas that will be used in proving the converse of Theorem 1 and Theorem 2. Proofs of the two lemmas are given in Appendix E of the extended version [11].

*Lemma 1:* Let  $\theta \in \{0, 1\}$  and  $(P'_0, P'_1), (P''_0, P''_1) \in \mathcal{P}(\mathcal{X})^2$  be two pairs of distributions satisfying (1). Consider a sequential classification test  $\Phi = (\tau, \delta)$ . For any  $\mathcal{E} \in \mathcal{F}_\tau$ ,

$$\begin{aligned} & d(\mathbb{P}'_\theta(\mathcal{E}), \mathbb{P}''_{1-\theta}(\mathcal{E})) \\ & \leq \mathbb{E}'_\theta[\tau] D(P'_\theta\|P''_{1-\theta}) + \mathbb{E}'_\theta[N_\tau] (D(P'_0\|P''_0) + D(P'_1\|P''_1)), \end{aligned}$$

where  $d(p, q) = p \log \frac{p}{q} + (1-p) \log \frac{1-p}{1-q}$  is the binary KL-divergence and  $\mathbb{E}'_\theta$  is the expectation taken under ground truth distributions  $(P'_0, P'_1)$  and hypothesis  $\mathcal{H}_\theta$ .

Lemma 1 is proved by the data processing inequality of KL divergences and the optional stopping theorem, similar to the converse part in [15, Theorem 15.3]

*Lemma 2:* Define the function  $h : (0, 1) \times (0, 1) \rightarrow \mathbb{R}$  as

$$h(p, q) = \frac{d(1-p, q)}{-\log q}.$$

Then,  $\lim_{(p,q) \rightarrow (0,0)} h(p, q) = 1$ .

Due to the space constraint, here we only sketch the proof of the upper bound for one error exponent. The other one can be proved by a similar argument. Details can be found in Appendix E of the extended version [11].

### A. Under Expected-Stopping-Time Universality Constraint

Consider any family of tests  $\{\Phi_n\}$  satisfying the expected-stopping-time universality constraint. For each  $\Phi_n = (\delta_n, \tau_n)$ , let us employ Lemma 1 with  $\theta = 1, P''_0 = P_0, P'_1 = P_1$ , and  $\mathcal{E} = \{\delta_n(X^{\tau_n}, T_0^{N_{\tau_n}}, T_1^{N_{\tau_n}}) = 1\}$ . By the universality constraint that  $\mathbb{E}'_1[\tau] \leq n$ , we have

$$\begin{aligned} & d(1 - \pi'_{0|1}(\Phi_n), \pi_{1|0}(\Phi_n)) \\ & \leq \mathbb{E}'_1[\tau_n] D(P'_1\|P_0) + \mathbb{E}'_1[N_{\tau_n}] (D(P'_0\|P_0) + D(P'_1\|P_1)) \\ & \leq n [D(P'_1\|P_0) + \alpha D(P'_0\|P_0) + \alpha D(P'_1\|P_1)] + o(1). \end{aligned}$$

Since  $\pi'_{0|1}(\Phi_n) \rightarrow 0, \pi_{1|0}(\Phi_n) \rightarrow 0$  as  $n \rightarrow \infty$ , we can apply Lemma 2 and get

$$e_0 \leq D(P'_1\|P_0) + \alpha D(P'_0\|P_0) + \alpha D(P'_1\|P_1).$$

Since this is true for all  $(P'_0, P'_1)$  satisfying (1), we can minimize the RHS with respect to  $(P'_0, P'_1)$  and leverage (7) to get  $e_0 \leq D_{\frac{\alpha}{1+\alpha}}(P_1\|P_0)$ .

### B. Under Error-Probability Universality Constraint

Consider any family of tests  $\{\Phi_\beta\}$  satisfying the error-probability universality constraint. For each  $\Phi_\beta = (\delta_\beta, \tau_\beta)$ , let us employ Lemma 1 with  $\theta = 0, P'_0 = P_0, P'_1 = P_1$ , and  $\mathcal{E} = \{\delta_\beta(X^{\tau_\beta}, T_0^{N_{\tau_\beta}}, T_1^{N_{\tau_\beta}}) = 0\}$ . Then we have

$$\begin{aligned} & d(1 - \pi_{1|0}(\Phi_\beta), \pi''_{0|1}(\Phi_\beta)) \\ & \leq \mathbb{E}_0[\tau_\beta] D(P_0\|P''_1) + \mathbb{E}_0[N_{\tau_\beta}] (D(P_0\|P''_0) + D(P_1\|P''_1)). \end{aligned}$$

It is clear that  $\mathbb{E}_0[\tau_\beta] \rightarrow \infty$  as  $\beta \rightarrow 0$ . Moreover,  $\pi''_{0|1}(\Phi_\beta) \leq \beta$  by the universality constraint. By Lemma 2, we have

$$\tilde{e}_0 \leq D(P_0\|P''_1) + \alpha D(P_0\|P''_0) + \alpha D(P_1\|P''_1).$$

Since this is true for all  $(P''_0, P''_1)$  satisfying (1), we can minimize the RHS with respect to  $(P''_0, P''_1)$  and leverage (6) to get  $\tilde{e}_0 \leq \text{GJS}(P_1, P_0, \alpha)$ .

## ACKNOWLEDGMENT

This work was supported by National Science and Technology Council (formerly Ministry of Science and Technology) of Taiwan under Grant MOST 110-2628-E-002-006.

## 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] R. E. Blahut, "Hypothesis testing and information theory," *IEEE Transactions on Information Theory*, vol. 20, no. 4, pp. 405–417, July 1974.
- [4] Y. Li and V. Y. F. Tan, "Second-order asymptotics of sequential hypothesis testing," *IEEE Transactions on Information Theory*, vol. 66, no. 11, pp. 7222–7230, November 2020.
- [5] S.-W. Lan and I.-H. Wang, "Heterogeneous sequential hypothesis testing with active source selection under budget constraints," *Proceedings of the 2021 IEEE International Symposium on Information Theory*, pp. 178–183, Melbourne, July 2021.
- [6] M. Haghifam, V. Y. F. Tan, and A. Khisti, "Sequential classification with empirically observed statistics," *IEEE Transactions on Information Theory*, vol. 67, no. 5, pp. 3095–3113, May 2021.
- [7] M. Gutman, "Asymptotically optimal classification for multiple tests with empirically observed statistics," *IEEE Transactions on Information Theory*, vol. 35, no. 2, pp. 401–408, March 1989.
- [8] Y. Li, S. Nitinawarat, and V. V. Veeravalli, "Universal sequential outlier hypothesis testing," *Sequential Analysis*, vol. 36, no. 3, pp. 309–344, 2017.
- [9] M. Naghshvar and T. Javidi, "Active sequential hypothesis testing," *The Annals of Statistics*, vol. 41, no. 6, pp. 2703–2738, 2013.
- [10] 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.
- [11] C.-Y. Hsu, C.-F. Li, and I.-H. Wang, "On universal sequential classification from sequentially observed empirical statistics," *Preprint*, 2022, <http://homepage.ntu.edu.tw/~ihwang/Eprint/itw22usc.pdf>.
- [12] L. Zhou, V. Y. F. Tan, and M. Motani, "Second-order asymptotically optimal statistical classification," *Information and Inference: A Journal of the IMA*, vol. 9, no. 1, pp. 81–111, March 2020.
- [13] H.-W. Hsu and I.-H. Wang, "On binary statistical classification from mismatched empirically observed statistics," in *2020 IEEE International Symposium on Information Theory (ISIT)*, 2020, pp. 2533–2538.
- [14] I. Csiszar, "The method of types," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2505–2523, 1998.
- [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](http://people.lids.mit.edu/yp/homepage/data/itlectures_v5.pdf)

APPENDIX

A. Proof of Proposition 1

Recall that a sequential probability ratio test  $\Phi = (\tau, \delta)$  is defined as

$$\tau = \inf\{k \in \mathbb{N} : S_k \geq B \text{ or } S_k \leq -A\}$$

$$\delta(x^\tau) = \begin{cases} 0 & \text{if } S_\tau \geq B, \\ 1 & \text{if } S_\tau \leq -A, \end{cases}$$

where  $S_k = \sum_{i=1}^k \log \frac{P_0(x_i)}{P_1(x_i)}$  is the logarithm of the probability ratio, and  $A, B > 0$  are two thresholds. Following the proof of [15, Theorem 15.3], it can be shown that if  $\mathbb{E}_0[\tau] < \infty$ , then

$$\mathbb{E}_0[S_\tau] = \mathbb{E}_0[\tau]D(P_0\|P_1), \quad (8)$$

and for every event  $\mathcal{E} \in \mathcal{F}_\tau$ , we have

$$\mathbb{E}_0[\mathbb{1}\{\mathcal{E}\}] = \mathbb{E}_1[2^{S_\tau} \mathbb{1}\{\mathcal{E}\}].$$

Hence

$$\pi_{1|0}(\Phi) = \mathbb{P}_0\{S_\tau \leq -A\} = \mathbb{E}_1[2^{S_\tau} \mathbb{1}\{S_\tau \leq -A\}] \leq 2^{-A}. \quad (9)$$

Let  $\tau_0 = \inf\{k \in \mathbb{N} : S_k \geq B\}$  and observe that  $\tau \leq \tau_0$ . By (8),

$$\mathbb{E}_0[\tau] \leq \mathbb{E}_0[\tau_0] = \frac{\mathbb{E}_0[S_{\tau_0}]}{D(P_0\|P_1)} \leq \frac{B + c_0}{D(P_0\|P_1)}, \quad (10)$$

where  $c_0 = \max_{x \in \mathcal{X}} \left| \log \frac{P_0(x)}{P_1(x)} \right|$ . Similarly we can show  $\pi_{0|1}(\Phi) \leq 2^{-B}$  and  $\mathbb{E}_1[\tau] \leq \frac{A + c_0}{D(P_1\|P_0)}$ .

For (4), consider a family of SPRT indexed by  $n \in \mathbb{N}$  with  $A_n = nD(P_1\|P_0) - c_0$  and  $B_n = nD(P_0\|P_1) - c_0$ . By (10), (2) is satisfied. Using (9), we have

$$e_0 \geq \liminf_{n \rightarrow \infty} \frac{nD(P_1\|P_0) - c_0}{n} = D(P_1\|P_0),$$

$$e_1 \geq \liminf_{n \rightarrow \infty} \frac{nD(P_0\|P_1) - c_0}{n} = D(P_0\|P_1).$$

For (5), consider a family of SPRT indexed by  $\beta \in (0, 1)$  with  $A_\beta = B_\beta = -\log \beta$ . By (9), (3) is satisfied. Using (10), we have

$$\tilde{e}_0 \geq \liminf_{\beta \rightarrow 0} \frac{-\log \beta}{-\log \beta + c_0} D(P_0\|P_1) = D(P_0\|P_1),$$

$$\tilde{e}_1 \geq \liminf_{\beta \rightarrow 0} \frac{-\log \beta}{-\log \beta + c_0} D(P_1\|P_0) = D(P_1\|P_0).$$

B. Achievability with Expected-Stopping-Time Universality

1) *Method of Types*: First we state some notations and properties about the method of types [14]. The collection of all possible realization sequences  $x^k$  is denoted as  $\mathcal{P}_d^k$ , and

$$|\mathcal{P}_d^k| \leq (k+1)^d. \quad (11)$$

The collection of all length- $k$  sequences that have the same type as  $\Pi_{x^k}$  is denoted as  $T(\Pi_{x^k})$ , so-called type class of  $\Pi_{x^k}$ . If  $x^k$  consists of i.i.d. samples following  $P$ ,

$$P^{\otimes k} \{X^k \in T(\Pi_{x^k})\} \leq 2^{-kD(\Pi_{x^k}\|P)}. \quad (12)$$

2) *Expected-Stopping-Time Universality*: Using these properties, we show that  $\Phi_n^{(M)}$  satisfies the universality constraint on expected stopping time with  $n$ . Express the expected stopping time as

$$\mathbb{E}_0[\tau] = \sum_{k=1}^{\infty} \mathbb{P}_0\{\tau \geq k\} \leq n - 1 + \sum_{k=n-1}^{\infty} \mathbb{P}_0\{\tau > k\}. \quad (13)$$

Next we aim to upper bound  $\mathbb{P}_0\{\tau > k\}$  such that the sum is less than 1.

$$\mathbb{P}_0\{\tau > k\} \leq \mathbb{P}_0 \left\{ \bigcap_{i=1}^k \left( \text{GJS} \left( \Pi_{T_0^{N_i}}, \Pi_{X^i}, \alpha \right) \geq f(i) \right) \right\}$$

$$\begin{aligned}
&\leq \mathbb{P}_0 \left\{ \text{GJS} \left( \Pi_{T_0^{N_k}}, \Pi_{X^k}, \alpha \right) \geq f(k) \right\} \\
&= \sum_{\substack{\left( \Pi_{x^k}, \Pi_{t_0^{N_k}} \right) \in \mathcal{P}_d^k \times \mathcal{P}_d^{N_k} \\ \text{GJS} \left( \Pi_{t_0^{N_k}}, \Pi_{x^k}, \alpha \right) \geq f(k)}} \mathbb{P}_0 \left\{ X^k \in T \left( \Pi_{x^k} \right) \right\} \mathbb{P}_0 \left\{ T_0^{N_k} \in T \left( \Pi_{t_0^{N_k}} \right) \right\} \\
&\leq \sum_{\substack{\left( \Pi_{x^k}, \Pi_{t_0^{N_k}} \right) \in \mathcal{P}_d^k \times \mathcal{P}_d^{N_k} \\ \text{GJS} \left( \Pi_{t_0^{N_k}}, \Pi_{x^k}, \alpha \right) \geq f(k)}} 2^{-kD(\Pi_{x^k} \| P_0)} 2^{-N_k D \left( \Pi_{t_0^{N_k}} \| P_0 \right)} \quad (\text{by (12)}) \\
&\leq \sum_{\substack{\left( \Pi_{x^k}, \Pi_{t_0^{N_k}} \right) \in \mathcal{P}_d^k \times \mathcal{P}_d^{N_k} \\ \text{GJS} \left( \Pi_{t_0^{N_k}}, \Pi_{x^k}, \alpha \right) \geq f(k)}} 2^{-kf(k)} \quad (\text{by the definition of the GJS divergence}) \\
&\leq (k+1)^d (N_k+1)^d (k+1)^{-2d} (N_k+1)^{-d} \quad (\text{by (11)}) \\
&= (k+1)^{-d}
\end{aligned}$$

Since  $(k+1)^{-d}$  is decreasing, we have for all  $n \geq 2$ ,

$$\sum_{k=n-1}^{\infty} \mathbb{P}_0 \{ \tau > k \} \leq \sum_{k=n-1}^{\infty} (k+1)^{-d} \leq \int_{n-2}^{\infty} (u+1)^{-d} du = \frac{(n-1)^{-(d-1)}}{d-1} \leq 1. \quad (14)$$

Combining (13) and (14) gives us  $\mathbb{E}_0[\tau] \leq n$ . Using the same argument, we can show  $\mathbb{E}_1[\tau] \leq n$ .

3) *Error Exponents:* First we upper bound the error probability.

$$\begin{aligned}
\pi_{1|0}(\Phi_n^{(M)}) &= \mathbb{P}_0 \left\{ \delta(X^\tau, T_0^{N_\tau}, T_1^{N_\tau}) = 1 \right\} \\
&= \mathbb{P}_0 \left\{ \bigcup_{k=n-1}^{\infty} \left( \text{GJS} \left( \Pi_{T_1^{N_k}}, \Pi_{X^k}, \alpha \right) \leq f(k) \right) \right\} \\
&\leq \sum_{k=n-1}^{\infty} \mathbb{P}_0 \left\{ \text{GJS} \left( \Pi_{T_1^{N_k}}, \Pi_{X^k}, \alpha \right) \leq f(k) \right\} \\
&\leq \sum_{k=n-1}^{\infty} \sum_{\substack{\left( \Pi_{x^k}, \Pi_{t_1^{N_k}} \right) \in \mathcal{P}_d^k \times \mathcal{P}_d^{N_k} \\ \text{GJS} \left( \Pi_{t_1^{N_k}}, \Pi_{x^k}, \alpha \right) \leq f(k)}} 2^{-kD(\Pi_{x^k} \| P_0)} 2^{-N_k D \left( \Pi_{t_1^{N_k}} \| P_1 \right)} \quad (\text{by (12)}) \quad (15)
\end{aligned}$$

Define

$$\Delta(n, P_0, P_1) = \min_{(V, V_1) \in \mathcal{P}(\mathcal{X})^2: \text{GJS}(V, V_1, \alpha) \leq f(n-1)} \alpha D(V \| P_1) + D(V_1 \| P_0).$$

Using (11) and the fact that  $f(k)$  is decreasing, we can upper bound (15) by

$$\sum_{k=n-1}^{\infty} 2^{-k} \left[ \Delta(n, P_0, P_1) - \frac{d \log n}{n-1} - \frac{d \log(N_{n-1} + 1)}{n-1} \right]. \quad (16)$$

Note that  $\text{GJS}(V, V_1, \alpha)$  is jointly convex in  $(V, V_1)$  [6, Lemma 4], and the objective function is strictly jointly convex in  $(V, V_1)$ . By the continuity of optimization problem [13, Corollary D.1],

$$\Delta(n, P_0, P_1) \rightarrow D_{\frac{\alpha}{1+\alpha}}(P_1 \| P_0) \text{ as } f(n-1) \rightarrow 0.$$

Moreover,

$$\frac{d \log n}{n-1} + \frac{d \log(N_{n-1} + 1)}{n-1} \rightarrow 0 \text{ and } f(n-1) \rightarrow 0 \text{ as } n \rightarrow \infty.$$

Hence, for any  $\epsilon > 0$ , we can find  $N > 0$  such that

$$\forall n \geq N, \left| \Delta(n, P_0, P_1) - \frac{d \log n}{n-1} - \frac{d \log(N_{n-1} + 1)}{n-1} - D_{\frac{\alpha}{1+\alpha}}(P_1 \| P_0) \right| \leq \epsilon.$$



Now pick some  $\epsilon < D_{\frac{\alpha}{1+\alpha}}(P_1\|P_0)$ , then for all  $n \geq N$ , (16) can be further upper bounded by

$$\sum_{k=n-1}^{\infty} 2^{-k(D_{\frac{\alpha}{1+\alpha}}(P_1\|P_0)-\epsilon)} = \frac{2^{-(n-1)(D_{\frac{\alpha}{1+\alpha}}(P_1\|P_0)-\epsilon)}}{1 - 2^{-(D_{\frac{\alpha}{1+\alpha}}(P_1\|P_0)-\epsilon)}}.$$

Take  $n$  to infinity and get  $e_0 \geq D_{\frac{\alpha}{1+\alpha}}(P_1\|P_0) - \epsilon$ . Since  $\epsilon$  can be made arbitrarily close to zero, we have  $e_0 \geq D_{\frac{\alpha}{1+\alpha}}(P_1\|P_0)$ . Similarly,  $e_1 \geq D_{\frac{\alpha}{1+\alpha}}(P_0\|P_1)$ .

### C. Achievability with Error-Probability Universality

1) *Error-Probability Universality*: First we show that  $\Phi_{\beta}^{(D)}$  satisfies the universality constraint on error probability with  $\beta$ .

$$\begin{aligned} \pi_{1|0}(\Phi_{\beta}^{(D)}) &= \mathbb{P}_0 \left\{ \delta(X^{\tau}, T_0^{N_{\tau}}, T_1^{N_{\tau}}) = 1 \right\} \\ &\leq \sum_{k=1}^{\infty} \mathbb{P}_0 \left\{ \tau = k, \delta(X^k, T_0^{N_k}, T_1^{N_k}) = 1 \right\} \\ &\leq \sum_{k=1}^{\infty} \mathbb{P}_0 \left\{ \text{GJS} \left( \Pi_{T_0^{N_k}}, \Pi_{X^k}, \alpha \right) \geq g(\beta, k) \right\} \\ &\leq \sum_{k=1}^{\infty} (k+1)^d (N_k+1)^d 2^{-kg(\beta, k)} \quad (\text{by the same method as in Appendix B2}) \\ &\leq \beta(d-1) \sum_{k=1}^{\infty} (k+1)^{-d} \\ &\leq \beta(d-1) \int_0^{\infty} (u+1)^{-d} du = \beta \end{aligned}$$

2) *Error Exponents*: For  $\Phi_{\beta}^{(D)}$ , define

$$\tau_0 = \inf \left\{ k \in \mathbb{N} : \text{GJS} \left( \Pi_{T_1^{N_k}}, \Pi_{X^k}, \alpha \right) > g(\beta, k) \right\}.$$

*Lemma 3*:  $\exists N > 0, c > 0$  such that for  $\Phi_{\beta}^{(D)}$ ,  $\mathbb{P}_0 \{ \tau_0 \geq k \} \leq \frac{1}{\beta} 2^{-ck}$  for any  $k \geq N$ .

*Proof*:

$$\begin{aligned} \mathbb{P}_0 \{ \tau_0 \geq k \} &\leq \mathbb{P}_0 \left\{ \text{GJS} \left( \Pi_{T_1^{N_{k-1}}}, \Pi_{X^{k-1}}, \alpha \right) \leq g(\beta, k-1) \right\} \\ &\leq \mathbb{P}_0 \left\{ \text{GJS} \left( \Pi_{T_1^{N_{k-1}}}, \Pi_{X^{k-1}}, \alpha \right) \leq g(\beta, k-1) \text{ and } D \left( \Pi_{T_1^{N_{k-1}}} \| P_1 \right) \leq \epsilon \text{ and } D \left( \Pi_{X^{k-1}} \| P_0 \right) \leq \epsilon \right\} \\ &\quad + \mathbb{P}_0 \left\{ D \left( \Pi_{T_1^{N_{k-1}}} \| P_1 \right) > \epsilon \right\} + \mathbb{P}_0 \left\{ D \left( \Pi_{X^{k-1}} \| P_0 \right) > \epsilon \right\} \\ &\stackrel{(i)}{\leq} \mathbb{P}_0 \left\{ \begin{array}{l} \text{GJS} \left( \Pi_{T_1^{N_{k-1}}}, \Pi_{X^{k-1}}, \alpha \right) \leq g(\beta, k-1) + \text{GJS} \left( \Pi_{T_0^{N_{k-1}}}, \Pi_{X^{k-1}}, \alpha \right) \\ \text{and } D \left( \Pi_{T_1^{N_{k-1}}} \| P_1 \right) \leq \epsilon \text{ and } D \left( \Pi_{X^{k-1}} \| P_0 \right) \leq \epsilon \end{array} \right\} \\ &\quad + \mathbb{P}_0 \left\{ D \left( \Pi_{T_1^{N_{k-1}}} \| P_1 \right) > \epsilon \right\} + \mathbb{P}_0 \left\{ D \left( \Pi_{X^{k-1}} \| P_0 \right) > \epsilon \right\} \\ &\stackrel{(ii)}{\leq} \mathbb{P}_0 \left\{ \text{GJS} \left( \Pi_{T_0^{N_{k-1}}}, \Pi_{X^{k-1}}, \alpha \right) \geq a - g(\beta, k-1) \right\} + k^d 2^{-(k-1)\epsilon} + (N_{k-1}+1)^d 2^{-N_{k-1}\epsilon} \\ &\leq \frac{1}{\beta(d-1)} k^{3d} (N_{k-1}+1)^{2d} 2^{-(k-1)a} + k^d 2^{-(k-1)\epsilon} + (N_{k-1}+1)^d 2^{-N_{k-1}\epsilon} \\ &\leq \frac{1}{\beta} 2^{-ck} \quad \forall k \geq N \text{ for some } c, N > 0 \text{ independent of } n \end{aligned}$$

where (i) is because  $\text{GJS} \left( \Pi_{T_0^{N_{k-1}}}, \Pi_{X^{k-1}}, \alpha \right)$  is nonnegative. For (ii), since  $P_0 \neq P_1$ , when  $\epsilon$  is chosen to be sufficiently small, it follows that  $\forall Q_0, Q_1 \in \mathcal{P}(\mathcal{X})$ , if  $D(Q_0\|P_0) \leq \epsilon$  and  $D(Q_1\|P_1) \leq \epsilon$ , then  $\text{GJS}(Q_1, Q_0, \alpha) \geq a$ . ■

*Lemma 4*: As  $\beta \rightarrow 0$ ,

$$-\frac{\tau_0}{\log \beta} \xrightarrow{\text{a.s.}} \frac{1}{\text{GJS}(P_1, P_0, \alpha)}.$$

*Proof:* By the strong law of large numbers, as  $k \rightarrow \infty$ ,

$$\Pi_{X^k} \xrightarrow{\text{a.s.}} P_0, \Pi_{T_0^{N_k}} \xrightarrow{\text{a.s.}} P_0, \Pi_{T_1^{N_k}} \xrightarrow{\text{a.s.}} P_1.$$

With the continuity of KL-divergence, we have

$$\text{GJS}\left(\Pi_{T_1^{N_k}}, \Pi_{X^k}, \alpha\right) \xrightarrow{\text{a.s.}} \text{GJS}(P_1, P_0, \alpha) \text{ as } k \rightarrow \infty.$$

By Lemma 3,  $\tau_0$  is finite a.s. Hence with probability 1 under  $\mathbb{P}_0$ ,

$$\text{GJS}\left(\Pi_{T_1^{N_{\tau_0}}}, \Pi_{X^{\tau_0}}, \alpha\right) > g(\beta, \tau_0) = -\frac{\log \beta + \log(d-1)}{\tau_0} + f(\tau_0), \quad (17)$$

$$\text{GJS}\left(\Pi_{T_1^{N_{\tau_0-1}}}, \Pi_{X^{\tau_0-1}}, \alpha\right) \leq g(\beta, \tau_0 - 1) = -\frac{\log \beta + \log(d-1)}{\tau_0 - 1} + f(\tau_0 - 1). \quad (18)$$

Next we show that  $\tau_0 \xrightarrow{\text{a.s.}} \infty$  as  $\beta \rightarrow 0$ :

$$\begin{aligned} \mathbb{P}_0\{\tau_0 \leq k\} &= \mathbb{P}_0\left\{\text{GJS}\left(\Pi_{T_1^{N_{\tau_0}}}, \Pi_{X^{\tau_0}}, \alpha\right) > g(\beta, \tau_0); \tau_0 \leq k\right\} \\ &\leq \mathbb{P}_0\left\{k \left(\min_V \left[\alpha D\left(\Pi_{T_1^{N_{\tau_0}}}\|V\right) + D\left(\Pi_{X^{\tau_0}}\|V\right)\right]\right) > \log \frac{1}{\beta(d-1)}\right\} \\ &\leq \mathbb{P}_0\left\{k\tilde{\alpha} \left(\min_V \left[D\left(\Pi_{T_1^{N_{\tau_0}}}\|V\right) + D\left(\Pi_{X^{\tau_0}}\|V\right)\right]\right) > \log \frac{1}{\beta(d-1)}\right\} \quad \text{where } \tilde{\alpha} = \max\{\alpha, 1\} \\ &\leq \mathbb{P}_0\left\{2k\tilde{\alpha} > \log \frac{1}{\beta(d-1)}\right\} \quad (\text{Since Jensen-Shannon divergence} \leq 1) \\ &= 0 \quad \forall k < -\frac{\log \beta + \log(d-1)}{2\tilde{\alpha}} \end{aligned}$$

Taking  $\beta \rightarrow 0$ , we have  $\tau_0 \xrightarrow{\text{a.s.}} \infty$ . The result follows from (17) and (18). ■

*Lemma 5:*  $\left\{-\frac{\tau_0}{\log \beta}\right\}_{\beta \in (0, 0.9]}$  is uniformly integrable.

*Proof:* Given any  $\epsilon > 0$ ,

$$\begin{aligned} &\mathbb{E}_0\left[-\frac{\tau_0}{\log \beta} \mathbb{1}\left\{-\frac{\tau_0}{\log \beta} \geq v\right\}\right] \\ &= -\frac{1}{\log \beta} \mathbb{E}_0[\tau_0 \mathbb{1}\{\tau_0 \geq -v \log \beta\}] \\ &= -\frac{1}{\log \beta} \sum_{k=1}^{\infty} \mathbb{P}_0\{\tau_0 \mathbb{1}\{\tau_0 \geq -v \log \beta\} \geq k\} \\ &\leq -\frac{1}{\log \beta} (-v \log \beta) \mathbb{P}_0\{\tau_0 \geq \lceil -v \log \beta \rceil\} - \frac{1}{\log \beta} \sum_{k=\lceil -v \log \beta \rceil}^{\infty} \mathbb{P}_0\{\tau_0 \geq k\} \\ &\leq \frac{v}{\beta} 2^{-c\lceil -v \log \beta \rceil} - \frac{1}{\beta \log \beta} \sum_{k=\lceil -v \log \beta \rceil}^{\infty} 2^{-ck} \quad (\text{by Lemma 3 if we choose } v \text{ such that } \lceil -v \log 0.9 \rceil \geq N) \\ &\leq v\beta^{cv-1} - \frac{1}{\beta \log \beta} \cdot \frac{\beta^{cv}}{1-2^{-c}} \\ &\leq 0.9^{cv-1} \left(v - \frac{1}{\log 0.9(1-2^{-c})}\right) \quad \text{if } v \text{ is large enough such that } cv-1 > 0 \\ &\leq \epsilon \quad \text{if } v \text{ is large enough} \end{aligned}$$

By Lemma 4 and Lemma 5, we have ■

$$\lim_{\beta \rightarrow 0} \mathbb{E}_0 \left[ \left| \frac{\tau_0}{-\log \beta} - \frac{1}{\text{GJS}(P_1, P_0, \alpha)} \right| \right] = 0.$$

By definition,  $\tau \leq \tau_0$ , hence  $e_0 \geq \text{GJS}(P_1, P_0, \alpha)$ .

#### D. Benefit of Sequentiality

Our goal is to show the following inequality for all  $P, Q$  and  $\alpha > 0$ :

$$D_{\frac{\alpha}{1+\alpha}}(P\|Q) \geq \text{GJS}(Q, P, \alpha).$$

First we introduce Hellinger distance of order  $a$

$$\text{Hel}_a(P\|Q) = \mathbb{E}_{X \sim Q} \left[ \frac{\left( \frac{P(X)}{Q(X)} \right)^a - 1}{a - 1} \right].$$

By definition, Rényi divergence of order  $a$  is related to Hellinger distance as

$$D_a(P\|Q) = \frac{1}{a-1} \log(1 + (a-1)\text{Hel}_a(P\|Q)).$$

Hence

$$D_{\frac{\alpha}{1+\alpha}}(P\|Q) = -(\alpha+1) \log \left( 1 - \frac{1}{1+\alpha} \text{Hel}_{\frac{\alpha}{1+\alpha}}(P\|Q) \right).$$

Since  $-\log(1-x) \geq (\log e)x$  for  $x > 0$ ,

$$D_{\frac{\alpha}{1+\alpha}}(P\|Q) \geq (\log e) \text{Hel}_{\frac{\alpha}{1+\alpha}}(P\|Q).$$

It remains to show that

$$(\log e) \text{Hel}_{\frac{\alpha}{1+\alpha}}(P\|Q) \geq \text{GJS}(Q, P, \alpha).$$

Note that  $\text{Hel}_{\frac{\alpha}{1+\alpha}}(P\|Q)$  is a  $f$ -divergence with

$$f_{\text{H}}(t) = -(\alpha+1) \left( t^{\frac{\alpha}{\alpha+1}} - 1 \right) + \alpha(t-1),$$

and  $\text{GJS}(Q, P, \alpha)$  is a  $f$ -divergence with

$$f_{\text{G}}(t) = \alpha \log \frac{\alpha+1}{\alpha+t} + t \log \frac{(\alpha+1)t}{\alpha+t}.$$

The proof is complete if we show  $(\log e)f_{\text{H}}(t) \geq f_{\text{G}}(t)$  for all  $t > 0$ . To do this, we calculate the derivatives of  $f_{\text{H}}(t)$  and  $f_{\text{G}}(t)$ :

$$f'_{\text{H}}(t) = \alpha \left( 1 - t^{-\frac{1}{\alpha+1}} \right) \tag{19}$$

$$f''_{\text{H}}(t) = \frac{\alpha}{\alpha+1} t^{-\frac{\alpha+2}{\alpha+1}} \tag{20}$$

$$f'_{\text{G}}(t) = \log \frac{(\alpha+1)t}{\alpha+t} \tag{21}$$

$$f''_{\text{G}}(t) = (\log e) \left( \frac{1}{t} - \frac{1}{\alpha+t} \right) \tag{22}$$

Consider  $\rho(t) = (\log e)f_{\text{H}}(t) - f_{\text{G}}(t)$ . It's clear that  $\rho(1) = 0$ . So it suffices to show  $\rho(1) = 0$  is a global minimum of  $\rho(t)$ . Combining (19) and (21), we have  $\rho'(1) = 0$ . Next we show that  $\rho(t)$  is convex by examining its second derivative. By (20) and (22),

$$\begin{aligned} \rho''(t) &= (\log e) \left( \frac{\alpha t^{-\frac{\alpha+2}{\alpha+1}}}{\alpha+1} - \frac{1}{t} + \frac{1}{\alpha+t} \right) \\ &= \frac{\alpha(\log e) t^{-\frac{\alpha+2}{\alpha+1}}}{(\alpha+1)(\alpha+t)} \left( \alpha+t - (\alpha+1)t^{\frac{1}{\alpha+1}} \right). \end{aligned}$$

It is obvious that  $\frac{\alpha(\log e) t^{-\frac{\alpha+2}{\alpha+1}}}{(\alpha+1)(\alpha+t)}$  is positive for  $t \in (0, \infty)$ . Let  $\eta(t) = \alpha+t - (\alpha+1)t^{\frac{1}{\alpha+1}}$ . We have  $\eta(1) = 0$ ,  $\eta'(1) = 0$  and  $\eta''(t) = \frac{\alpha}{\alpha+1} t^{-\frac{2\alpha+1}{\alpha+1}} > 0$  for  $t \in (0, \infty)$ . Hence  $\eta(t)$  is convex and  $\eta(1) = 0$  is a global minimum. Using the fact that  $\eta(t)$  is non-negative, we obtain  $\rho''(t) \geq 0$ . Therefore  $\rho(t)$  is convex and  $\rho(1) = 0$  is a global minimum.

E. Converse

1) *Proof of Lemma 1:* The proof is similar to the converse part in [15, Theorem 15.3]. Using data processing inequality of divergence:

$$\begin{aligned} d(\mathbb{P}'_\theta(\mathcal{E}), \mathbb{P}''_{1-\theta}(\mathcal{E})) &\leq D(\mathbb{P}'_\theta \| \mathbb{P}''_{1-\theta}) |_{\mathcal{F}_\tau} \\ &= \mathbb{E}'_\theta \left[ \sum_{k=1}^{\tau} \log \frac{P'_\theta(X_k)}{P''_{1-\theta}(X_k)} + \sum_{k=1}^{N_\tau} \log \frac{P'_0(T_{0,k})}{P''_0(T_{0,k})} + \sum_{k=1}^{N_\tau} \log \frac{P'_1(T_{1,k})}{P''_1(T_{1,k})} \right] \\ &= \mathbb{E}'_\theta [\tau] D(P'_\theta \| P''_{1-\theta}) + \mathbb{E}'_\theta [N_\tau] (D(P'_0 \| P''_0) + D(P'_1 \| P''_1)), \end{aligned}$$

where the last equality is by Doob's Optional Stopping Theorem.

2) *Proof of Lemma 2:* Given any  $\epsilon > 0$ , we aim to find  $\delta > 0$  such that for all  $p, q \in (0, \delta)$ ,  $|h(p, q) - 1| \leq \epsilon$ . For  $p, q \in (0, \delta)$ , if  $\delta < 0.5$ ,

$$\begin{aligned} d(1-p, q) &\geq d(1-\delta, q) \\ &= (1-\delta) \log \frac{1-\delta}{q} + \delta \log \frac{\delta}{1-q} \\ &= (1-\delta) \log \frac{1}{q} + (1-\delta) \log(1-\delta) + \delta \log \delta + \delta \log \frac{1}{1-q} \\ &\geq (1-\delta) \log \frac{1}{q} - \frac{2}{e \ln 2} \quad (\text{since } \forall x \in (0, 1), x \log x \geq -\frac{2}{e \ln 2}) \end{aligned}$$

For the upper bound,

$$d(1-p, q) = (1-p) \log \frac{1-p}{q} + p \log \frac{p}{1-q} \leq \log \frac{1}{q} + \delta \log \frac{1}{1-\delta}.$$

Hence

$$\frac{d(1-p, q)}{\log \frac{1}{q}} \geq 1 - \delta - \frac{2}{e \ln 2} \frac{1}{\log \frac{1}{q}} \geq 1 - \delta - \frac{2}{e \ln 2} \frac{1}{\log \frac{1}{\delta}},$$

and

$$\frac{d(1-p, q)}{\log \frac{1}{q}} \leq 1 + \frac{\delta \log \frac{1}{1-\delta}}{\log \frac{1}{q}} \leq 1 + \frac{\delta \log \frac{1}{1-\delta}}{\log \frac{1}{\delta}}.$$

It suffices to choose  $\delta > 0$  small enough such that

$$\delta - \frac{2}{e \ln 2} \frac{1}{\log \delta} < \epsilon \text{ and } \frac{\delta \log(1-\delta)}{\log \delta} < \epsilon.$$