2023 IEEE International Symposium on Information Theory (ISIT) | 978-1-6654-7554-9/23/$31.00 ©2023 IEEE | DOI: 10.1109/ISIT54713.2023.10206736

Hypothesis Testing for Adversarial Channels: Chernoff-Stein Exponents

Eeshan Modak[1], Neha Sangwan[1], Mayank Bakshi[2], Bikash Kumar Dey[3], and Vinod M. Prabhakaran[1]

[1]Tata Institute of Fundamental Research, Mumbai, India
[2]Arizona State University, Tempe, AZ, USA
[3]Indian Institute of Technology Bombay, Mumbai, India

***Abstract*—We study the Chernoff-Stein exponent of the following binary hypothesis testing problem: Associated with each hypothesis is a set of channels. A transmitter, without knowledge of the hypothesis, chooses the vector of inputs to the channel. Given the hypothesis, from the set associated with the hypothesis, an adversary chooses channels, one for each element of the input vector. Based on the channel outputs, a detector attempts to distinguish between the hypotheses. We study the Chernoff-Stein exponent for the cases where the transmitter (i) is deterministic, (ii) may privately randomize, and (iii) shares randomness with the detector that is unavailable to the adversary. It turns out that while a memoryless transmission strategy is optimal under shared randomness, it may be strictly suboptimal when the transmitter only has private randomness.**

1. Introduction

In binary hypothesis testing the goal is to distinguish between two distributions (sources) [1], [2]. When n independent and identically distributed (i.i.d.) observations from the source are available, the Chernoff-Stein lemma [3, Theorem 11.8.3] states that for a fixed false alarm (type-1 error) probability, the optimal missed detection (type-2 error) probability decays exponentially in n with the exponent given by the relative entropy between the distributions.

A variation on this problem is where each observation is from an arbitrarily varying source [4]. There is a set of distributions associated with each hypothesis. Given a hypothesis, the observations are independent, but each observation could be arbitrarily distributed according to any one of the distributions belonging to the set of distributions corresponding to the hypothesis. We may view the choice of distribution as being made by an adversary who is aware of the detection scheme used. Fangwei and Shiyi [5] studied this problem where the adversary's choice may be stochastic. Recently, Brandão, Harrow, Lee, and Peres [6] considered the case with an adaptive adversary who has feedback of the past observations and may use this to choose the distribution of the next observation.

In another variation on the binary hypothesis testing problem, instead of distinguishing between sources, the objective is to distinguish between two channels with the same input and output alphabets [7], [8]. Here, a transmitter, which is unaware of the hypothesis, may choose the inputs to the channels. It can be shown that the optimal Chernoff-Stein error exponent may be attained using a deterministic transmission strategy which sends the input letter for which the relative entropy between the channel output distributions under the two hypotheses is maximized [7]. Hayashi [8] studied the adaptive case where the transmitter has feedback of the channel output when the block length is fixed and showed that feedback does not improve the optimal error exponent. Polyanskiy and Verdú [9] considered the same problem with variable-length transmissions and showed that feedback may improve the error exponent in general.

In this work we study the Chernoff-Stein exponent of the binary hypothesis testing problem for arbitrarily varying channels [10]. Associated with each hypothesis is a set of channels. All channels have the same input and output alphabets. The transmitter, without knowledge of the hypothesis, chooses the vector of inputs to the channel. Given the hypothesis, the adversary chooses a vector of channels where each element belongs to the set of channels associated with the hypothesis. The adversary is aware of the strategy of the transmitter and detector, but not necessarily the choice of channel inputs. The detector observes the outputs resulting from applying the inputs chosen by the transmitter element-wise independently to the channels selected by the adversary. We consider three different settings depending on the nature of randomness unknown to the adversary which is available to the transmitter and detector[1]: (i) deterministic schemes (Section 4), (ii) randomness shared between transmitter and detector (Section 3), and (iii) private randomness at the transmitter (Section 5). We also comment on the role of adaptivity both of the transmitter (under a fixed block length) and of the adversary (Section 6). Wherever omitted, the proofs can be found in the extended version [11].

When the channels are not arbitrarily varying, randomization (and adaptivity in the fixed length case) do not change the optimal Chernoff-Stein exponent which is achieved by the deterministic transmitter strategy of repeating the input symbol for which the channel output distributions under the

E. Modak, N. Sangwan and V. M. Prabhakaran were supported by DAE under project no. RTI4001. N. Sangwan was additionally supported by the TCS Foundation through the TCS Research Scholar Program. The work of M. Bakshi was supported by the National Science Foundation under Grant No. CCF-2107526. The work of B. K. Dey was supported by Bharti Centre for Communication in IIT Bombay. V. M. Prabhakaran was additionally supported by SERB through project MTR/2020/000308.

[1] We allow the adversary to randomize in all cases. The optimal exponent is unaffected by the availability of common randomness known also to the adversary, nor by additional private randomness at the detector.

	Chernoff-Stein exponent	Condition for the exponent to be non-zero
Shared randomness	$\sup_{P_X} \min_{W \in \text{conv}(\mathcal{W}), \overline{W} \in \text{conv}(\overline{\mathcal{W}})} D(W\|\overline{W}\|P_X)$	$\text{conv}(\mathcal{W}) \cap \text{conv}(\overline{\mathcal{W}}) = \emptyset$
Deterministic transmitter	$\max_x \min_{W_x \in \text{conv}(\mathcal{W}_x), \overline{W}_x \in \text{conv}(\overline{\mathcal{W}}_x)} D(W_x\|\overline{W}_x)$	$\text{conv}(\mathcal{W}_x) \cap \text{conv}(\overline{\mathcal{W}}_x) = \emptyset$ for some x
Private randomness	Open (see Theorem 5)	$\text{conv}(\mathcal{W}) \cap \text{conv}(\overline{\mathcal{W}}) = \emptyset$ and $(\mathcal{W}, \overline{\mathcal{W}})$ is not trans-symmetrizable

two hypotheses have the largest relative entropy [7], [8]. With arbitrarily varying channels, we see that randomization improves the exponent in general (Remark 1 and Example 2). This is analogous to the usefulness of randomization in communication over arbitrarily varying channels [12]. We also demonstrate that the optimal exponents under the three different settings are different in general. Our results also show the following interesting phenomenon: When the transmitter has private randomness which is unknown to the adversary, but shares no randomness with the detector, it turns out that a memoryless transmission strategy is strictly sub-optimal in general (Section 5). This is in contrast to the optimality of a memoryless transmission scheme when the transmitter and detector share randomness. Another related work, especially to Section 5 on the private randomness case, is [13] as we discuss there. It considered communication and testing in a similar model though error exponents for testing were not considered there. While the present paper was under review, a work by Bergh, Datta and Salzmann [14] which studies binary composite classical and quantum channel discrimination appeared. The results there on classical composite convex sets are closely related to those of Section 3 on the shared randomness case.

2. Preliminaries

Adversarial Hypothesis Testing. Our achievability proofs use the adversarial Chernoff-Stein lemma from [6] which we briefly describe here. Let $\mathcal{Z}$ be a finite set. Let $\mathcal{P}, \mathcal{Q} \subseteq \mathbb{R}^{\mathcal{Z}}$ be closed, convex sets of probability distributions. The adaptive adversary is specified by $\hat{p}_i : \mathcal{Z}^{i-1} \to \mathcal{P}$ and $\hat{q}_i : \mathcal{Z}^{i-1} \to \mathcal{Q}$ for $i \in [1:n]$. For any $z^n \in \mathcal{Z}^n$, let $\hat{p}(z^n) := \prod_{i=1}^n \hat{p}_i(z^{i-1})(z_i)$ and $\hat{q}(z^n) := \prod_{i=1}^n \hat{q}_i(z^{i-1})(z_i)$. Let $A_n \subseteq \mathcal{Z}^n$ be an acceptance region for $\mathcal{P}$. For $\epsilon > 0$, the type-I and type-II errors are defined to be

$$\alpha_n \overset{\text{def}}{=} \sup_{(\hat{p}_i)_{i=1}^n} \hat{p}(A_n^c), \qquad \beta_n^\epsilon \overset{\text{def}}{=} \min_{A_n : \alpha_n \leq \epsilon} \sup_{(\hat{q}_i)_{i=1}^n} \hat{q}(A_n),$$

and the adversarial Chernoff-Stein exponent is given by

$$\mathcal{E}_{\text{adv}}^\epsilon(\mathcal{P}, \mathcal{Q}) \overset{\text{def}}{=} \lim_{n \to \infty} -\frac{1}{n} \log \beta_n^\epsilon.$$

For any pair $p \in \mathcal{P}, q \in \mathcal{Q}$, since the adversary may (non-adaptively) choose $\hat{p}_i = p$ and $\hat{q}_i = q$ for all $i \in [1:n]$, by the Chernoff-Stein lemma [3, Theorem 11.8.3] it is clear that $\mathcal{E}_{\text{adv}}^\epsilon(\mathcal{P}, \mathcal{Q}) \leq \min_{p \in \mathcal{P}, q \in \mathcal{Q}} D(p\|q)$. In [5] it was shown that this upper bound is achievable if the adversary is non-adaptive. The following theorem states that this remains true even when the adversary is adaptive.

Theorem 1 (Adversarial Chernoff-Stein Lemma [6])**.** Let $\mathcal{Z}$ be a finite domain. For any pair of closed, convex sets of probability distributions $\mathcal{P}, \mathcal{Q} \subseteq \mathbb{R}^{\mathcal{Z}}$,

$$\mathcal{E}_{\text{adv}}^\epsilon(\mathcal{P}, \mathcal{Q}) = \min_{p \in \mathcal{P}, q \in \mathcal{Q}} D(p\|q). \tag{1}$$

Problem Setup. Let $\mathcal{X}$ and $\mathcal{Y}$ be finite sets. A discrete memoryless channel $W(.|.)$ takes an input symbol $x \in \mathcal{X}$ and outputs a symbol $y \in \mathcal{Y}$ with probability $W(y|x)$. Consider two finite sets of channels $\mathcal{W} = \{W(.|.,s) : s \in \mathcal{S}\}$, $\overline{\mathcal{W}} = \{\overline{W}(.|.,\bar{s}) : \bar{s} \in \bar{\mathcal{S}}\}$. The goal is to distinguish between the two sets of channels. In particular, we study the asymmetric hypothesis test between the null hypothesis $H_0 : \mathcal{W}$ and the alternative hypothesis $H_1 : \overline{\mathcal{W}}$. There are three entities involved: (a) the transmitter, (b) the adversary, and (c) the detector. The transmitter is unaware of which hypothesis has been realized and chooses the input symbols. The adversary, depending on which hypothesis is realized, chooses the state symbols (from $\mathcal{S}$ under H_0 and $\bar{\mathcal{S}}$ under H_1). The detector decides between H_0 and H_1 based on everything it knows. We will elaborate this in the coming sections.

3. Shared Randomness

In this setting, the transmitter and detector share randomness which is unknown to the adversary. The input X^n to the channel, which is a function of this randomness, is known to the detector. For a transmitter strategy P_{X^n} and a pair of adversary strategies P_{S^n} and $P_{\bar{S}^n}$, the distribution induced on $\mathcal{X}^n \times \mathcal{Y}^n$ under H_0 is given by

$$Q_{\text{sh}}^n(x^n, y^n) = \sum_{s^n \in \mathcal{S}^n} P_{X^n}(x^n) P_{S^n}(s^n) \prod_{i=1}^n W(y_i|x_i, s_i). \tag{2}$$

A similar expression is obtained for $\bar{Q}_{\text{sh}}^n$ under H_1 where instead of P_{S^n} and W we have $P_{\bar{S}}$ and $\overline{W}$ respectively. The detector uses a (possibly privately randomized) decision rule $f_{\text{sh}} : \mathcal{X}^n \times \mathcal{Y}^n \to \{0,1\}$. Let A_n be the (possibly random) acceptance region for H_0, i.e., $A_n = \{(x^n, y^n) \in \mathcal{X}^n \times \mathcal{Y}^n : f_{\text{sh}}(x^n, y^n) = 0\}$. For a given transmitter and detector strategy, the type-I error is given by

$$\alpha_n^{\text{sh}} = \sup_{P_{S^n}} \mathbb{E}\left[Q_{\text{sh}}^n(A_n^c)\right],$$

where the expectation is over the random choice of A_n. For $\epsilon > 0$, when the type-I error α_n^{sh} is at most ϵ, the optimal type-II error is given by

$$\beta_n^{\epsilon,\text{sh}} \overset{\text{def}}{=} \inf_{P_{X^n}} \inf_{A_n : \alpha_n^{\text{sh}} \leq \epsilon} \sup_{P_{\bar{S}^n}} \mathbb{E}\left[\bar{Q}_{\text{sh}}^n(A_n)\right],$$

where the expectation is over the random A_n set by the inner inf. The Chernoff-Stein exponent is then defined to be

$$\mathcal{E}^{\epsilon}_{\text{sh}}(\mathcal{W},\overline{\mathcal{W}}) \overset{\text{def}}{=} \liminf_{n\to\infty} -\frac{1}{n}\log \beta^{\epsilon,\text{sh}}_{n}, \quad \epsilon > 0.$$

Let $\text{conv}(\mathcal{W})$ and $\text{conv}(\overline{\mathcal{W}})$ be the convex hulls of the channel sets $\mathcal{W}$ and $\overline{\mathcal{W}}$ respectively. i.e.,

$$\text{conv}(\mathcal{W}) \overset{\text{def}}{=} \left\{ \sum_{s\in\mathcal{S}} P_S(s)W(.|.,s) : P_S \in \Delta_{\mathcal{S}} \right\},$$

where $\Delta_{\mathcal{S}}$ is the set of all probability distributions over $\mathcal{S}$. $\text{conv}(\overline{\mathcal{W}})$ is defined similarly with $\bar{S},\overline{W}$ instead of S, W. Let

$$D^*_{\text{sh}} \overset{\text{def}}{=} \sup_{P_X} \min_{\substack{W\in\text{conv}(\mathcal{W})\\ \overline{W}\in\text{conv}(\overline{\mathcal{W}})}} D(W\|\overline{W}|P_X). \tag{3}$$

Since $\text{conv}(\mathcal{W})$, $\text{conv}(\overline{\mathcal{W}})$ are closed, convex sets and $D(.\|.)$ is lower semi-continuous, the minimum exists.

Theorem 2. Let $\mathcal{W}$ and $\overline{\mathcal{W}}$ be two sets of discrete memoryless channels which map $\mathcal{X}$ to $\mathcal{Y}$. For any $\epsilon \in (0,1)$, we have

$$D^*_{\text{sh}} \le \mathcal{E}^{\epsilon}_{\text{sh}}(\mathcal{W},\overline{\mathcal{W}}) \le \frac{D^*_{\text{sh}}}{1-\epsilon}. \tag{4}$$

Proof. Achievability ($\mathcal{E}^{\epsilon}_{\text{sh}}(\mathcal{W},\overline{\mathcal{W}}) \ge D^*_{\text{sh}}$): We argue the achievability for the (stronger) adaptive adversary who has access to previous channel inputs and outputs. The transmitter transmits X^n chosen i.i.d. according to P_X using the shared randomness. This reduces the problem to the adversarial hypothesis testing problem studied in [6]. For any fixed choice of P_X, invoking Theorem 1 with $\mathcal{P} = \{P_X W : W \in \text{conv}(\mathcal{W})\}$ and $\mathcal{Q} = \{P_X\overline{W} : \overline{W} \in \text{conv}(\overline{\mathcal{W}})\}$,

$$\mathcal{E}^{\epsilon}_{\text{sh}}(\mathcal{W},\overline{\mathcal{W}}) \ge \min_{\substack{W\in\text{conv}(\mathcal{W})\\ \overline{W}\in\text{conv}(\overline{\mathcal{W}})}} D(W\|\overline{W}|P_X).$$

Optimizing over P_X completes the proof of achievability.

Weak Converse ($\mathcal{E}^{\epsilon}_{\text{sh}}(\mathcal{W},\overline{\mathcal{W}}) \le \frac{D^*_{\text{sh}}}{1-\epsilon}$): We show this converse result for an adaptive transmitter who has feedback of the outputs. Given an adaptive transmitter, we construct an adversarial strategy to show the upper bound on the exponent. Specifically, we consider a memoryless strategy (not necessarily i.i.d.) for the adversary, i.e. $P_{S^n} = \prod_{i=1}^{n} P_{S_i}$ and $P_{\bar{S}^n} = \prod_{i=1}^{n} P_{\bar{S}_i}$ where P_{S_i} and $P_{\bar{S}_i}$ will be specified in course of the proof. Let Q^n and $\bar{Q}^n$ denote the joint distributions on $\mathcal{X}^n \times \mathcal{Y}^n$ under H_0 and H_1 respectively. They are given by

$$Q^n(x^n,y^n) = \prod_{i=1}^{n} \vec{Q}_i(x_i|x^{i-1},y^{i-1}) \left(\sum_{s_i\in\mathcal{S}} P_{S_i}(s_i) W(y_i|x_i,s_i) \right) \tag{5}$$

and

$$\bar{Q}^n(x^n,y^n) = \prod_{i=1}^{n} \vec{Q}_i(x_i|x^{i-1},y^{i-1}) \left(\sum_{\bar{s}_i\in\bar{\mathcal{S}}} P_{\bar{S}_i}(\bar{s}_i) \overline{W}(y_i|x_i,\bar{s}_i) \right). \tag{6}$$

Here, $\vec{Q}_i(x_i|x^{i-1},y^{i-1})$ denotes the transmitter strategy at the i^{th} timestep. We now try to get an upper bound on $D(Q^n\|\bar{Q}^n)$.

$$\begin{aligned} D(Q^n\|\bar{Q}^n) &= \sum_{i=1}^{n} D(Q_{X_i,Y_i|(X,Y)^{i-1}}\|\bar{Q}_{X_i,Y_i|(X,Y)^{i-1}}|Q^{i-1}) \\ &= \sum_{i=1}^{n} \big(D(\vec{Q}_i\|\vec{Q}_i|Q^{i-1}) \\ &\quad + D(Q_{Y_i|X^i,Y^{i-1}}\|\bar{Q}_{Y_i|X^i,Y^{i-1}}|Q^{i-1}\vec{Q}_i) \big) \end{aligned}$$

Observe that all the $D(\vec{Q}_i\|\vec{Q}_i|Q^{i-1})$ terms are zero. Furthermore, from (5), (6), we can see that $Q_{Y_i|X^i,Y^{i-1}} = Q_{Y_i|X_i}$, $\bar{Q}_{Y_i|X^i,Y^{i-1}} = \bar{Q}_{Y_i|X_i}$. Thus,

$$\begin{aligned} D(Q^n\|\bar{Q}^n) &= \sum_{i=1}^{n} D(Q_{Y_i|X_i}\|\bar{Q}_{Y_i|X_i}|Q^{i-1}\vec{Q}_i) \\ &= \sum_{i=1}^{n} D(Q_{Y_i|X_i}\|\bar{Q}_{Y_i|X_i}|Q_{X_i}) \end{aligned} \tag{7}$$

It is easy to see that $(P_{S_1}, P_{\bar{S}_1})$ can be chosen such that $D(Q_{Y_1|X_1}\|\bar{Q}_{Y_1|X_1}|\vec{Q}_1) \le D^*_{\text{sh}}$. We then recursively specify $(P_{S_i}, P_{\bar{S}_i})$ such that each term in (7) is upper bounded by D^*_{sh}. Thus,

$$D(Q^n\|\bar{Q}^n) \le nD^*_{\text{sh}}. \tag{8}$$

With this upper bound in place, we may follow a standard approach via the data processing inequality to complete the proof (e.g., see [8, Section VI]). See Appendix A, [11] where we complete these steps. □

The following theorem characterizes the pairs of $(\mathcal{W},\overline{\mathcal{W}})$ for which $\mathcal{E}^{\epsilon}_{\text{sh}} > 0$.

Theorem 3. $\mathcal{E}^{\epsilon}_{\text{sh}}(\mathcal{W},\overline{\mathcal{W}}) > 0 \iff \text{conv}(\mathcal{W}) \cap \text{conv}(\overline{\mathcal{W}}) = \emptyset$.

Proof. The if ($\Leftarrow$) part follows from Theorem 2. To see the (contrapositive of the) only if ($\Rightarrow$) direction, notice that under hypothesis H_0 (resp., H_1), the adversary may induce any channel from $\text{conv}(\mathcal{W})$ (resp., $\text{conv}(\overline{\mathcal{W}})$) from the transmitter to the detector. Hence, when the intersection is non-empty, the adversary may induce the same channel under both hypotheses so that no transmission strategy (including an adaptive one) can distinguish between the hypotheses. □

4. Deterministic Transmitter

In this setting, the transmitter strategy is completely deterministic and is defined by a fixed tuple $(x_1, x_2, \ldots, x_n)$. The distribution on $\mathcal{Y}^n$ under H_0 and H_1 are similar to (2) with P_{X^n} as a point mass on $(x_1, x_2, \ldots, x_n)$. The definitions of decision rule f_{det} and acceptance region A_n are similar to those in Section 3 except that the observation space is $\mathcal{Y}^n$ instead of $\mathcal{X}^n \times \mathcal{Y}^n$. The definitions of $\alpha^{\text{det}}_{n}, \beta^{\epsilon,\text{det}}_{n}$ and $\mathcal{E}^{\epsilon}_{\text{det}}$ are also similar except that the inf is over the input symbols in the expression for $\beta^{\epsilon,\text{det}}_{n}$.

For $x \in \mathcal{X}$, let $\text{conv}(\mathcal{W}_x)$ and $\text{conv}(\overline{\mathcal{W}}_x)$ be the convex hulls of the conditional distributions $W(.|x,s)$ and $\overline{W}(.|x,\bar{s})$.

$$\text{conv}(\mathcal{W}_x) \overset{\text{def}}{=} \left\{ \sum_{s\in\mathcal{S}} P_S(s)W(.|x,s) : P_S \in \Delta_{\mathcal{S}} \right\},$$

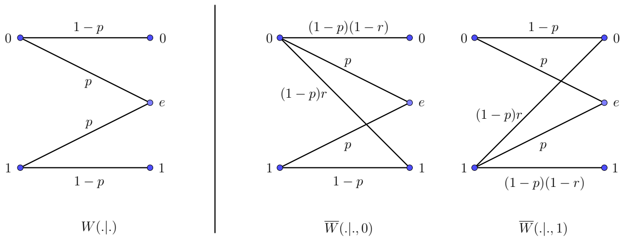

Fig. 1. An example in which, for a privately randomized transmitter, the hypotheses cannot be distribguished using memoryless transmission schemes, but a scheme with 2-step memory yields a positive Chernoff-Stein exponent.

$\text{conv}(\overline{\mathcal{W}}_x)$ is defined similarly with $\bar{S}, \overline{W}$ instead of S, W. Define D^*_{det} to be

$$D^*_{\text{det}} := \max_x \min_{\substack{W_x \in \text{conv}(\mathcal{W}_x) \\ \overline{W}_x \in \text{conv}(\overline{\mathcal{W}}_x)}} D(W_x \| \overline{W}_x) \tag{9}$$

Theorem 4. Let $\mathcal{W}$ and $\overline{\mathcal{W}}$ be two sets of discrete memoryless channels which map $\mathcal{X}$ to $\mathcal{Y}$. For any $\epsilon \in (0,1)$, we have

$$D^*_{\text{det}} \le \mathcal{E}^{\epsilon}_{\text{det}}(\mathcal{W}, \overline{\mathcal{W}}) \le \frac{D^*_{\text{det}}}{1-\epsilon}. \tag{10}$$

The proof is on similar lines as Theorem 2. We also show that (10) holds when both the transmitter and the adversary are adaptive. A characterization theorem analogous to Theorem 3 can also be shown. We omit these in the interest of space.

5. Private Randomness

We now consider the case where the transmitter may choose the channel input X^n randomly, but the realization of X^n is unavailable to the detector and the adversary. We may define the optimal Chernoff-Stein exponent $\mathcal{E}^{\epsilon}_{\text{pvt}}(\mathcal{W}, \overline{\mathcal{W}})$ along the same lines as earlier sections. Note that the decision function is now a (possibly random) partition of $\mathcal{Y}^n$. By the discussion leading up to Theorem 1, if the transmitter adopts an i.i.d. P_X strategy, the best possible exponent (irrespective of whether the adversary is adaptive or not) is

$$D_{\text{pvt,iid}} = \sup_{P_X} \min_{\substack{Q_Y \in \mathcal{Q}(P_X) \\ \bar{Q}_Y \in \bar{\mathcal{Q}}(P_X)}} D(Q_Y \| \bar{Q}_Y),$$

where $\mathcal{Q}(P_X)$ (resp. $\bar{\mathcal{Q}}(P_X)$) is the set of (single-letter) channel output distributions that can be induced by the adversary when the input is distributed as P_X under hypothesis H_0 (resp. H_1), i.e., $\mathcal{Q} \overset{\text{def}}{=} \left\{ \sum_{x,s} P_S(s) P_X(x) W(.|x,s) : P_S \in \Delta_{\mathcal{S}} \right\}$. It turns out that in general the optimal exponent $\mathcal{E}^{\epsilon}_{\text{pvt}}(\mathcal{W}, \overline{\mathcal{W}})$ could be strictly larger that $D_{\text{pvt,iid}}$. In the following example, $\mathcal{E}^{\epsilon}_{\text{pvt}}(\mathcal{W}, \overline{\mathcal{W}}) > 0$ for all $\epsilon > 0$ even though $D_{\text{pvt,iid}} = 0$.

Example 1. $H_0 : \mathcal{W} = \{W(.|.)\}$ consists of a binary erasure channel (BEC) with parameter $p < 1$ and $H_1 : \overline{\mathcal{W}} = \{\overline{W}(.|.,0), \overline{W}(.|.,1)\}$ consists of two modified BEC(p) channels where one of the symbols flips with probability $(1-p)r$, $r > 0$ as shown in Figure 1. Here, $\mathcal{X} = \{0,1\}, \mathcal{Y} = \{0,1,e\}, \mathcal{S} = \{0\}, \bar{\mathcal{S}} = \{0,1\}$. Note that $\mathcal{Q}$ is a singleton. It is easy to verify that, under H_1, if the adversary sets $P_{\bar{S}}(0) = 1 - P_X(0)$, the induced channel output distribution will be the same as the one under H_0. Hence, $\mathcal{Q} \subset \bar{\mathcal{Q}}$ and therefore $D_{\text{pvt,iid}} = 0$.

Now to see that $\mathcal{E}^{\epsilon}_{\text{pvt}}(\mathcal{W}, \overline{\mathcal{W}}) > 0$, consider a transmission scheme with 2-step memory: $n/2$ i.i.d. pairs are sent where each pair is distributed as $P_{X_1,X_2}(0,0) = P_{X_1,X_2}(1,1) = 0.5$. The effective channel is now a random map from $\mathcal{X}^2$ to $\mathcal{Y}^2$. The new state space for the (non-adaptive) adversary under H_0 is $\mathcal{S}^2$ (which is still a singleton), and $\bar{\mathcal{S}}^2$ under H_1. Let $\mathcal{Q}_2$ (resp. $\bar{\mathcal{Q}}_2$) be the set of (two-letter) channel output distributions that can be induced by the adversary when the input is distributed according to P_{X_1,X_2} under H_0 (resp. H_1). Since $\mathcal{Q}_2$ is a singleton, let the member be denoted by Q_{Y_1,Y_2}. If we show that $Q_{Y_1,Y_2} \notin \bar{\mathcal{Q}}_2$, we may conclude that $\mathcal{E}^{\epsilon}_{\text{pvt}}(\mathcal{W}, \overline{\mathcal{W}}) > 0$. Assume for contradiction that this is not the case, i.e., suppose there exists $P_{\bar{S}_1,\bar{S}_2}$ such that the resulting $\bar{Q}_{Y_1,Y_2}$ is the same as Q_{Y_1,Y_2}. Since the marginals also have to be equal, we have $Q_{Y_1} = \bar{Q}_{Y_1}$. This forces P_{S_1} to be uniform. Now, observe that $Q_{Y_1,Y_2}(0,1) = 0$ while, irrespective of $P_{S_2|S_1}$, we have $\bar{Q}_{Y_1,Y_2}(0,1) > 0$ since $r > 0$. This is a contradiction and hence $Q_{Y_1,Y_2} \notin \bar{\mathcal{Q}}_2$. Therefore, $\mathcal{E}^{\epsilon}_{\text{pvt}}(\mathcal{W}, \overline{\mathcal{W}}) > 0$ by Theorem 1.

The above argument does not account for an adaptive adversary. In Appendix E, [11] we show that even with an adaptive adversary the above transmission scheme leads to a positive exponent.

Remark 1. For the above example, $\mathcal{E}^{\epsilon}_{\text{det}}\left(\mathcal{W}, \overline{\mathcal{W}}\right) < \mathcal{E}^{\epsilon}_{\text{pvt}}(\mathcal{W}, \overline{\mathcal{W}})$. This follows from $D^*_{\text{det}} \le D_{\text{pvt,iid}}$ which is a consequence of the fact that for P_X such that $P_X(x) = 1$ for some $x \in \mathcal{X}$, the corresponding $\mathcal{Q}(P_X)$ and $\bar{\mathcal{Q}}(P_X)$ are $\text{conv}(\mathcal{W}_x)$ and $\text{conv}(\overline{\mathcal{W}}_x)$ respectively.

In the rest of this section, we give an achievable lower bound on the error exponent $\mathcal{E}^{\epsilon}_{\text{pvt}}(\mathcal{W}, \overline{\mathcal{W}})$ and characterize the pairs $\left(\mathcal{W}, \overline{\mathcal{W}}\right)$ for which it is positive[2]. If $\text{conv}(\mathcal{W}) \cap \text{conv}(\overline{\mathcal{W}}) \ne \emptyset$, then $\mathcal{E}^{\epsilon}_{\text{pvt}}(\mathcal{W}, \overline{\mathcal{W}}) = 0$ (by Theorem 3). This follows from the fact that the adversary can choose S^n and $\bar{S}^n$ i.i.d. so that a channel in the intersection may be induced which renders the hypotheses indistinguishable irrespective of the transmission scheme. It turns out that when the transmitter only has private randomness, a more carefully chosen adversary strategy which now depends on the transmission scheme may render $\mathcal{E}^{\epsilon}_{\text{pvt}}(\mathcal{W}, \overline{\mathcal{W}}) = 0$ for a larger class of $\left(\mathcal{W}, \overline{\mathcal{W}}\right)$ pairs.

Definition 1 ([13, eq. (2)])**.** The pair $\left(\mathcal{W}, \overline{\mathcal{W}}\right)$ is *trans-symmetrizable* if there exist conditional distributions $P_{S|X}, P_{\bar{S}|X}$ such that, for every $x, \tilde{x} \in \mathcal{X}$ and $y \in \mathcal{Y}$,

$$\sum_{s \in \mathcal{S}} P_{S|X}(s|x) W(y|\tilde{x}, s) = \sum_{\bar{s} \in \bar{\mathcal{S}}} P_{\bar{S}|X}(\bar{s}|\tilde{x}) W(y|x, \bar{s}). \tag{11}$$

[2] This characterization is implicit in [15, Corollary 1]. Note that the "deterministic coding" transmitter there has access to the message which serves as a source of private randomness for the testing problem.

Consider a trans-symmetrizable pair $\left(\mathcal{W},\overline{\mathcal{W}}\right)$ and a (non-adaptive[3]) transmission scheme $\hat{P}$. We will demonstrate (non-adaptive) adversary strategies under which the detector is unable to distinguish between the hypotheses. Under hypothesis H_1, the adversary, independent of the transmitter, samples a $\tilde{X}^n$ according to $\hat{P}$ and passes it through the (memoryless) channel $P_{\bar{S}|X}$ of Definition 1 to produce its $\bar{S}^n$. This induces the following distribution on the channel output vector:

$$\sum_{x^n,\bar{s}^n} \hat{P}(x^n)\left[\sum_{\tilde{x}^n}\hat{P}(\tilde{x}^n)\prod_{i=1}^{n}\left(P_{\bar{S}|X}(\bar{s}_i|\tilde{x}_i)\right)\right] W^n(y^n|x^n,\bar{s}^n)$$
$$= \sum_{x^n,\tilde{x}^n}\hat{P}(x^n)\hat{P}(\tilde{x}^n)\prod_{i=1}^{n}\left[\sum_{\bar{s}_i\in\bar{\mathcal{S}}} P_{\bar{S}|X}(\bar{s}_i|\tilde{x}_i)W(y_i|x_i,\bar{s}_i)\right]$$
$$\overset{(a)}{=} \sum_{\tilde{x}^n,x^n}\hat{P}(\tilde{x}^n)\hat{P}(x^n)\prod_{i=1}^{n}\left[\sum_{s_i\in\mathcal{S}}P_{S|X}(s_i|x_i)W(y_i|\tilde{x}_i,s_i)\right]$$
$$= \sum_{\tilde{x}^n,s^n}\hat{P}(\tilde{x}^n)\left[\sum_{x^n}\hat{P}(x^n)\prod_{i=1}^{n}\left(P_{S|X}(s_i|x_i)\right)\right]W^n(y^n|\tilde{x}^n,s^n)$$

where (a) follows from (11). This is identical to the channel output distribution under hypothesis H_0 if the adversary samples from $\hat{P}$ (independent of the transmitter) and passes through the channel $P_{S|X}$ of Definition 1 to produce its S^n. Thus, $\mathcal{E}^{\epsilon}_{\text{pvt}}(\mathcal{W},\overline{\mathcal{W}})=0$ if $\left(\mathcal{W},\overline{\mathcal{W}}\right)$ is trans-symmetrizable. The example below establishes a separation between shared and private randomness.

Example 2 ([13, Example 1])**.** Let $\mathcal{X}=\mathcal{S}=\bar{\mathcal{S}}=\{0,1\}$ and $\mathcal{Y}=\{0,1\}^2$. Suppose W deterministically outputs $Y=(X,S)$ while $\bar{W}$ outputs $Y=(\bar{S},X)$. Clearly, $\text{conv}(\mathcal{W})\cap\text{conv}(\overline{\mathcal{W}})=\emptyset$. Hence, by Theorem 3, $\mathcal{E}^{\epsilon}_{\text{sh}}>0$. However, $\left(\mathcal{W},\overline{\mathcal{W}}\right)$ is trans-symmetrizable since $P_{S|X}(x|x)=P_{\bar{S}|X}(x|x)=1$ for all $x\in\mathcal{X}$ satisfies (11). Hence $\mathcal{E}^{\epsilon}_{\text{pvt}}(\mathcal{W},\overline{\mathcal{W}})=0$.

Our lower bound on $\mathcal{E}^{\epsilon}_{\text{pvt}}(\mathcal{W},\overline{\mathcal{W}})$ is in terms of the following quantitative measure of how far the pair $\left(\mathcal{W},\overline{\mathcal{W}}\right)$ is from being trans-symmetrizable and/or having a non-empty intersection of their convex hulls; Lemma 1 and its proof in Appendix F, [11] makes this connection concrete.

Definition 2. For a distribution P over $\mathcal{X}$, we define $\eta(P)$ as the set of triples (η_1,η_2,η_3) for which there exists $\delta>0$ such that there is no joint distribution $P_{XX'\bar{S}SY}$ with $P_X=P_{X'}=P$ satisfying

1) $I(X;\bar{S})<\eta_1$, $I(X';S)<\delta$,
2) $D(P_{X\bar{S}Y}||P_{X\bar{S}}W)<\eta_2$,
3) $D(P_{X'SY}||P_{X'S}W)<\delta$, and
4) if $P_{XX'}(X'\neq X)>0$,
(i) $I(X';XY|\bar{S})<\eta_3$, and (ii) $I(X;X'Y|S)<\delta$.

Our main theorem for this section is the following:

Theorem 5. Let $\epsilon>0$.

$\mathcal{E}^{\epsilon}_{\text{pvt}}(\mathcal{W},\overline{\mathcal{W}})=0$ if $\left(\mathcal{W},\overline{\mathcal{W}}\right)$ is trans-symmetrizable or

$$\text{conv}(\mathcal{W})\cap\text{conv}(\overline{\mathcal{W}})\neq\emptyset$$

$$\mathcal{E}^{\epsilon}_{\text{pvt}}(\mathcal{W},\overline{\mathcal{W}})\geq \max\left\{\max_{P,(\eta_1,\eta_2,\eta_3)\in\eta(P)}\min\left\{\eta_1,\eta_2,\frac{\eta_3}{3}\right\},D^*_{\text{det}}\right\}$$

Lemma 1. If $\left(\mathcal{W},\overline{\mathcal{W}}\right)$ is not trans-symmetrizable and $\text{conv}(\mathcal{W})\cap\text{conv}(\overline{\mathcal{W}})=\emptyset$, there exists an input distribution P with $(\eta_1,\eta_2,\eta_3)\in\eta(P)$ such that $\eta_1,\eta_2,\eta_3>0$.

Corollary 1. $\mathcal{E}^{\epsilon}_{\text{pvt}}(\mathcal{W},\overline{\mathcal{W}})>0$ if and only if $\left(\mathcal{W},\overline{\mathcal{W}}\right)$ is not trans-symmetrizable and $\text{conv}(\mathcal{W})\cap\text{conv}(\overline{\mathcal{W}})=\emptyset$.

This recovers [15, Corollary 1] which gave the same characterization for $\left(\mathcal{W},\overline{\mathcal{W}}\right)$ which allow hypothesis testing with vanishing probability of error when the transmitter has private randomness (in the form a random message). Our proof (in Appendix F, [11]) of the lower bound to $\mathcal{E}^{\epsilon}_{\text{pvt}}(\mathcal{W},\overline{\mathcal{W}})$ in Theorem 5, which is inspired by [15], entails significant careful modifications to the detector and the error analysis there.

6. On the Role of Adaptivity

1) With shared randomness: It turns out that our results hold even if the transmitter and/or adversary is adaptive. We proved the achievability part of Theorem 2 assuming that the adversary is adaptive and the converse assuming the transmitter is adaptive.

2) Deterministic schemes: Here the optimal exponent remains unchanged even if the adversary is adaptive (irrespective of whether the transmitter is adaptive or not). This is also the case if both the adversary and the transmitter are adaptive. These follow from our achievability proof which is shown assuming an adaptive adversary and the converse which is shown when (a) both the transmitter and adversary are non-adaptive and (b) when both are adaptive (see Appendix D, [11]). It is also easy to see that, in general, if the transmitter is adaptive and the adversary is not, the exponent could be improved. The transmitter and detector may extract some randomness unknown to the adversary from the channel output feedback of, say, the first half of the block, and use this to implement a scheme with shared randomness during the second half. Since there are channels for which deterministic exponent is zero while the exponent under shared randomness is positive (for instance, see Example 2), these (possibly augmented by an independent random channel output component which provide additional shared randomness) serve as examples where such an improvement is feasible.

3) With private randomness: If the adversary is non-adaptive and the transmitter is adaptive, improved exponents are possible along the lines of the above discussion. This follows from the fact that there are channels where the exponent with shared randomness is positive, while that with private randomness is zero (specifically, trans-symmetrizable but with $\text{conv}(\mathcal{W})\cap\text{conv}(\overline{\mathcal{W}})=\emptyset$; see Example 2). We also showed that memoryless schemes may be strictly sub-optimal even if the adversary is adaptive (Appendix E, [11]). Also, the impossibility result in Theorem 5 can be shown when both the transmitter and adversary are adaptive.

[3] This discussion can be modified to handle an adaptive transmission scheme if the adversary is also adaptive. This is omitted in the interest of space.

References

[1] H. Chernoff, "A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations," *The Annals of Mathematical Statistics*, pp. 493–507, 1952.

[2] W. Hoeffding, "Asymptotically optimal tests for multinomial distributions," *The Annals of Mathematical Statistics*, pp. 369–401, 1965.

[3] T. M. Cover, *Elements of information theory*. John Wiley & Sons, 1999.

[4] V. Strassen, "Meßfehler und information," *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, vol. 2, pp. 273–305, 1964.

[5] F. Fangwei and S. Shiyi, "Hypothesis testing for arbitrarily varying source," *Acta Mathematica Sinica*, vol. 12, no. 1, pp. 33–39, 1996.

[6] F. G. Brandão, A. W. Harrow, J. R. Lee, and Y. Peres, "Adversarial hypothesis testing and a quantum Stein's lemma for restricted measurements," *IEEE Transactions on Information Theory*, vol. 66, no. 8, pp. 5037–5054, 2020.

[7] R. Blahut, "Hypothesis testing and information theory," *IEEE Transactions on Information Theory*, vol. 20, no. 4, pp. 405–417, 1974.

[8] M. Hayashi, "Discrimination of two channels by adaptive methods and its application to quantum system," *IEEE Transactions on Information Theory*, vol. 55, no. 8, pp. 3807–3820, 2009.

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

[10] D. Blackwell, L. Breiman, and A. Thomasian, "The capacities of certain channel classes under random coding," *The Annals of Mathematical Statistics*, vol. 31, no. 3, pp. 558–567, 1960.

[11] E. Modak, N. Sangwan, M. Bakshi, B. K. Dey, and V. M. Prabhakaran, "Hypothesis testing for adversarial channels: Chernoff-Stein exponents," *arXiv preprint arXiv:2304.14166*, 2023.

[12] A. Lapidoth and P. Narayan, "Reliable communication under channel uncertainty," *IEEE transactions on Information Theory*, vol. 44, no. 6, pp. 2148–2177, 1998.

[13] S. Chaudhuri, N. Sangwan, M. Bakshi, B. K. Dey, and V. M. Prabhakaran, "Compound arbitrarily varying channels," in *2021 IEEE International Symposium on Information Theory (ISIT)*, pp. 503–508, IEEE, 2021.

[14] B. Bergh, N. Datta, and R. Salzmann, "Composite classical and quantum channel discrimination," *arXiv preprint arXiv:2303.02016*, 2023.

[15] S. Chaudhuri, N. Sangwan, M. Bakshi, B. K. Dey, and V. M. Prabhakaran, "Compound arbitrarily varying channels," *arXiv preprint arXiv:2105.03420*, 2021.