

Error Exponents of Quantum State Discrimination With Composite Correlated Hypotheses

Kun Fang^{ID}, *Member, IEEE*, and Masahito Hayashi^{ID}, *Fellow, IEEE*

Abstract—We study the error exponents in quantum hypothesis testing between two sets of quantum states, extending the analysis beyond the independent and identically distributed case to encompass composite correlated hypotheses. In particular, we introduce and compare two natural extensions of the quantum Hoeffding divergence and anti-divergence to sets of quantum states, establishing their equivalence or quantitative relations. In the error exponent regime, we generalize the quantum Hoeffding bound to stable sequences of convex, compact sets of quantum states, demonstrating that the optimal Type-I error exponent, under an exponential constraint on the Type-II error, is precisely characterized by the regularized quantum Hoeffding divergence between the sets. In the strong converse exponent regime, we establish a general lower bound on the exponent in terms of the regularized quantum Hoeffding anti-divergence, and we prove a matching upper bound when the null hypothesis is a singleton, under additional assumptions. The generality of these results enables applications in various contexts, including (i) refining the generalized quantum Stein’s lemma by Fang et al. (2024); (ii) exhibiting counterexamples to the continuity of the regularized Petz Rényi divergence and Hoeffding divergence; (iii) obtaining error exponents for adversarial channel discrimination and resource detection problems.

Index Terms—Error exponents, quantum hypothesis testing, composite hypotheses, quantum Hoeffding divergence, quantum Hoeffding anti-divergence.

Received 10 December 2025; revised 10 April 2026; accepted 10 April 2026. Date of publication 16 April 2026; date of current version 21 May 2026. The work of Kun Fang was supported in part by the National Natural Science Foundation of China under Grant 12404569 and Grant 92470113, in part by Shenzhen Science and Technology Program under Grant QNXMB20250701091826036 and Grant JCYJ20240813113519025, in part by Shenzhen International Quantum Academy under Grant SIQA2025KFKT03, in part by Shenzhen Fundamental Research Program under Grant JCYJ20241202124023031, in part by the General Research and Development Projects of 1+1+1 CUHK-CUHK(SZ)-GDST Joint Collaboration Fund under Grant GRDP2025-022, and in part by the University Development Fund under Grant UDF01003565. The work of Masahito Hayashi was supported in part by the General Research and Development Projects of 1+1+1 CUHK-CUHK(SZ)-GDST Joint Collaboration Fund under Grant GRDP2025-022, in part by Guangdong Provincial Quantum Science Strategic Initiative under Grant GDZX2505003, in part by Shenzhen International Quantum Academy under Grant SIQA2025KFKT07, and in part by the National Natural Science Foundation of China under Grant 62171212. (Corresponding author: Kun Fang.)

Kun Fang is with the School of Data Science, The Chinese University of Hong Kong, Shenzhen, Guangdong 518172, China (e-mail: kunfang@cuhk.edu.cn).

Masahito Hayashi is with the School of Data Science, The Chinese University of Hong Kong, Shenzhen, Guangdong 518172, China, also with the International Quantum Academy, Futian, Shenzhen 518048, China, and also with the Graduate School of Mathematics, Nagoya University, Nagoya 464-8602, Japan (e-mail: hmasahito@cuhk.edu.cn).

Communicated by H.-C. Cheng, Associate Editor for Quantum.

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/TIT.2026.3684314>.

Digital Object Identifier 10.1109/TIT.2026.3684314

I. INTRODUCTION

DISTINGUISHABILITY is a central topic in information theory from both theoretical and practical perspectives. A fundamental framework for studying distinguishability is *asymmetric hypothesis testing*. In this setting, a source generates a sample x from one of two probability distributions $p \equiv \{p(x)\}_{x \in \mathcal{X}}$ or $q \equiv \{q(x)\}_{x \in \mathcal{X}}$. The objective of asymmetric hypothesis testing is to minimize the Type-II error (decides p when the fact is q) while keeping the Type-I error (decides q when the fact is p) within a certain threshold. The celebrated Chernoff-Stein’s Lemma [1] states that, for any constant bound on the Type-I error, the optimal Type-II error decays exponentially fast in the number of samples, and the decay rate is exactly the relative entropy,

$$D(p||q) = \sum_{x \in \mathcal{X}} p(x) [\log p(x) - \log q(x)]. \quad (1)$$

In particular, this lemma also states the “strong converse property”, a desirable mathematical property in information theory [2] that delineates a sharp boundary for the tradeoff between the Type-I and Type-II errors in the asymptotic regime: any possible scheme with Type-II error decaying to zero with an exponent larger than the relative entropy will result in the Type-I error converging to one in the asymptotic limit. Therefore, the Chernoff-Stein’s Lemma provides a rigorous operational interpretation of the relative entropy and establishes a crucial connection between hypothesis testing and information theory [3].

A natural question is whether the above result generalizes to the quantum case. Substantial efforts have been made to answer this fundamental question in quantum information community (see, e.g., [4], [5], [6], [7], [8], [9], [10], [11], [12], [13], [14], [15], [16]). Consider the problem of distinguishing between two quantum hypotheses: the system is prepared either in state ρ_n (the null hypothesis) or in state σ_n (the alternative hypothesis). Operationally, the discrimination is carried out using a two-outcome positive operator-valued measure (POVM) $\{M_n, I - M_n\}$, with $0 \leq M_n \leq I$. Then, the Type-I and Type-II errors are, respectively, given by

$$\text{(Type-I)} \quad \alpha(\rho_n, M_n) := \text{Tr}[\rho_n(I - M_n)], \quad (2)$$

$$\text{(Type-II)} \quad \beta(\sigma_n, M_n) := \text{Tr}[\sigma_n M_n]. \quad (3)$$

It is generally impossible to find a quantum measurement that simultaneously makes both errors vanish; thus, one studies the asymptotic behavior of α and β as $n \rightarrow \infty$, expecting a trade-off between minimizing α and minimizing β . The interplay

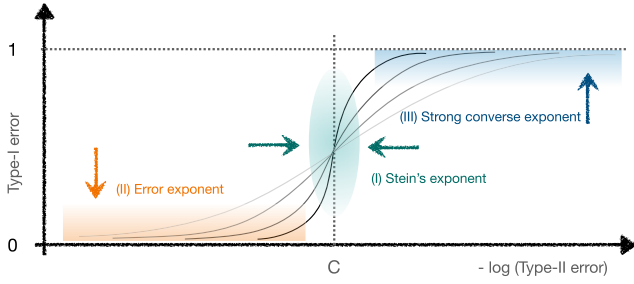


Fig. 1. Illustration of different study regimes of quantum hypothesis testing. Each curve represents the tradeoff between the Type-I and Type-II errors for varying block lengths n , with darker lines corresponding to longer block lengths. (I) represents the Stein exponent regime, (II) represents the error exponent regime, and (III) represents the strong converse exponent regime.

between these errors can be analyzed in various operational regimes (see Figure 1): (I) the **Stein exponent** regime, which focuses on the exponential decay rate of the Type-II error when the Type-I error is below a constant threshold; (II) the **error exponent** regime, which investigates the exponential rate at which the Type-I error vanishes when the Type-II error is required to decay exponentially at a prescribed rate; and (III) the **strong converse exponent** regime, which examines the exponential rate at which the Type-I error converges to one when the Type-II error decays exponentially at a given rate.

A. Stein Exponent

In asymmetric hypothesis testing, one aims to minimize the Type-II error while keeping the Type-I error below a fixed threshold $\varepsilon \in (0, 1)$. The optimal Type-II error is given by

$$\beta_\varepsilon(\rho_n \| \sigma_n) := \min_{0 \leq M_n \leq I} \{\beta(\sigma_n, M_n) : \alpha(\rho_n, M_n) \leq \varepsilon\}. \quad (4)$$

The quantum version of the Chernoff-Stein's Lemma (also known as quantum Stein's lemma) states that the optimal Type-II error decays exponentially with the number of copies n of the states when the Type-I error is restricted below a constant threshold and the optimal exponent is given by the quantum relative entropy $D(\rho \| \sigma) := \text{Tr}[\rho(\log \rho - \log \sigma)]$ [4], [5],

$$\lim_{n \rightarrow \infty} -\frac{1}{n} \log \beta_\varepsilon(\rho^{\otimes n} \| \sigma^{\otimes n}) = D(\rho \| \sigma), \quad \forall \varepsilon \in (0, 1). \quad (5)$$

B. Error Exponent

As a refinement of the quantum Stein's lemma, one can study the optimal Type-I error given that the Type-II error decays with a given exponential speed. One is then interested in the asymptotics of the Type-I error,

$$\alpha_{n,r}(\rho_n \| \sigma_n) := \min_{0 \leq M_n \leq I} \{\alpha(\rho_n, M_n) : \beta(\sigma_n, M_n) \leq 2^{-nr}\}, \quad (6)$$

with a constant $r > 0$. When $r < D(\rho \| \sigma)$, the optimal Type-I error $\alpha_{n,r}(\rho^{\otimes n} \| \sigma^{\otimes n})$, also decays with an exponential speed, as was shown in [17]. The exact decay rate is determined by the quantum Hoeffding bound [7], [8], [18] as

$$\lim_{n \rightarrow \infty} -\frac{1}{n} \log \alpha_{n,r}(\rho^{\otimes n} \| \sigma^{\otimes n}) = H_r(\rho \| \sigma), \quad (7)$$

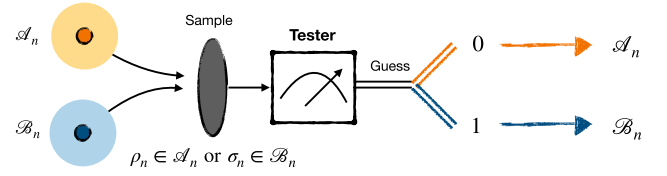


Fig. 2. Quantum hypothesis testing between two sets of quantum states $\mathcal{A}_n$ and $\mathcal{B}_n$.

where the quantum Hoeffding divergence is defined by

$$H_r(\rho \| \sigma) := \sup_{\alpha \in (0,1)} \frac{\alpha - 1}{\alpha} (r - D_{P,\alpha}(\rho \| \sigma)), \quad (8)$$

and $D_{P,\alpha}$ is the Petz Rényi divergence.

C. Strong Converse Exponent

On the other hand, when $r > D(\rho \| \sigma)$, the optimal Type-I error $\alpha_{n,r}(\rho^{\otimes n} \| \sigma^{\otimes n})$ goes to 1 exponentially fast [5]. The rate of this convergence is [11],

$$\lim_{n \rightarrow \infty} -\frac{1}{n} \log(1 - \alpha_{n,r}(\rho^{\otimes n} \| \sigma^{\otimes n})) = H_r^*(\rho \| \sigma), \quad (9)$$

where the Hoeffding anti-divergence is defined by

$$H_r^*(\rho \| \sigma) := \sup_{\alpha > 1} \frac{\alpha - 1}{\alpha} (r - D_{S,\alpha}(\rho \| \sigma)), \quad (10)$$

and $D_{S,\alpha}$ is the sandwiched Rényi divergence.

The results in (7) and (9) provide a comprehensive characterization of the asymptotic trade-off between the Type-I and Type-II error probabilities. In particular, the quantum Stein's lemma emerges as a special case in the limit $r \rightarrow D(\rho \| \sigma)$.

While much of the existing literature has focused on i.i.d. sources, practical scenarios often involve quantum states that are not fully specified (i.e., *composite hypotheses* [9], [16], [19], [20], [21], [22]) such as in adversarial or black-box settings [14], [23] and exhibit correlations that preclude a simple tensor product structure (i.e., *correlated hypotheses* [16], [24], [25], [26]). In this context, the task is to discriminate between two sets of correlated quantum states (see Figure 2). That is, a tester receives samples prepared according to either the set $\mathcal{A}_n$ or $\mathcal{B}_n$, and determines, via a quantum measurement $\{M_n, I - M_n\}$, from which set the samples originate.

As in standard hypothesis testing, two types of errors can occur: a Type-I error, where a sample from $\mathcal{A}_n$ is incorrectly classified as coming from $\mathcal{B}_n$, and a Type-II error, where a sample from $\mathcal{B}_n$ is incorrectly classified as coming from $\mathcal{A}_n$. Since we aim to control the errors for any state within the given sets, regardless of which one is drawn, the (worst-case) Type-I error is defined by

$$\alpha(\mathcal{A}_n, M_n) := \sup_{\rho_n \in \mathcal{A}_n} \text{Tr}[\rho_n(I - M_n)], \quad (11)$$

and the (worst-case) Type-II error is defined by

$$\beta(\mathcal{B}_n, M_n) := \sup_{\sigma_n \in \mathcal{B}_n} \text{Tr}[\sigma_n M_n]. \quad (12)$$

A few variants of quantum Stein's lemma has been shown in this worst-case setting under some structural assumptions

on the sets $\mathcal{A}_n$ and $\mathcal{B}_n$ [15], [27], [28]. Particularly, the Stein's lemma in [15] shows that

$$\lim_{n \rightarrow \infty} -\frac{1}{n} \log \beta_\varepsilon(\mathcal{A}_n \| \mathcal{B}_n) = D^\infty(\mathcal{A} \| \mathcal{B}), \quad \forall \varepsilon \in (0, 1), \quad (13)$$

where

$$\beta_\varepsilon(\mathcal{A}_n \| \mathcal{B}_n) := \min_{0 \leq M_n \leq I} \{\beta(\mathcal{B}_n, M_n) : \alpha(\mathcal{A}_n, M_n) \leq \varepsilon\}, \quad (14)$$

and $D(\mathcal{A}_n \| \mathcal{B}_n) := \inf_{\rho_n \in \mathcal{A}_n, \sigma_n \in \mathcal{B}_n} D(\rho_n \| \sigma_n)$ is the quantum relative entropy between two sets of states and $D^\infty(\mathcal{A} \| \mathcal{B}) := \lim_{n \rightarrow \infty} \frac{1}{n} D(\mathcal{A}_n \| \mathcal{B}_n)$ is the regularization.

In this work, we refine the analysis of asymmetric hypothesis testing for composite correlated hypotheses, as developed in [15], [27], and [28], by extending it to the error exponent and strong converse exponent regimes. Specifically, we consider the optimal Type-I error:¹

$$\begin{aligned} \alpha_{n,r}(\mathcal{A}_n \| \mathcal{B}_n) \\ := \min_{0 \leq M_n \leq I} \{\alpha(\mathcal{A}_n, M_n) : \beta(\mathcal{B}_n, M_n) \leq 2^{-nr}\}, \end{aligned} \quad (15)$$

and seek to determine the error exponent

$$\lim_{n \rightarrow \infty} -\frac{1}{n} \log \alpha_{n,r}(\mathcal{A}_n \| \mathcal{B}_n) = ? \quad (16)$$

and the strong converse exponent

$$\lim_{n \rightarrow \infty} -\frac{1}{n} \log(1 - \alpha_{n,r}(\mathcal{A}_n \| \mathcal{B}_n)) = ? \quad (17)$$

Due to the uncertainty and correlation inherent in composite correlated hypotheses, several significant challenges arise when obtaining precise characterizations. First, discrimination strategies must be state-agnostic, ensuring that error probabilities are universally controlled for every state within the respective sets, regardless of which specific state is encountered. Second, the optimization problem exhibits a minimax structure, forming a competing scenario that requires simultaneous maximization over all possible states in the sets and minimization over all possible quantum measurements. Third, the non-i.i.d. structure significantly complicates the asymptotic analysis, as standard techniques relying on tensor product structures no longer directly apply. Fourth, it becomes necessary to define suitable extensions of the quantum Hoeffding divergence and anti-divergence to sets of quantum states that both recover existing results for i.i.d. sources and precisely capture the essential features of the general composite correlated setting.

For the error exponent regime, we explore two natural approaches to extending the quantum Hoeffding divergence (see Figure 3). The first treats the Hoeffding divergence as a quantum divergence and considers the minimal divergence between the sets, denoted as $H_{n,r}(\mathcal{A}_n \| \mathcal{B}_n)$. The second approach uses the explicit formula for the Hoeffding divergence in terms of the Petz Rényi divergences: we first extend the Rényi divergences to sets of quantum states, $D_{p,\alpha}(\mathcal{A}_n \| \mathcal{B}_n)$, and then define the Hoeffding divergence accordingly, denoted as $\mathfrak{H}_{n,r}(\mathcal{A}_n \| \mathcal{B}_n)$. We show that these two extensions are equivalent for finite n in general, and further show the inequality

¹Such version of definition is similarly used in [10], [21], and [29].

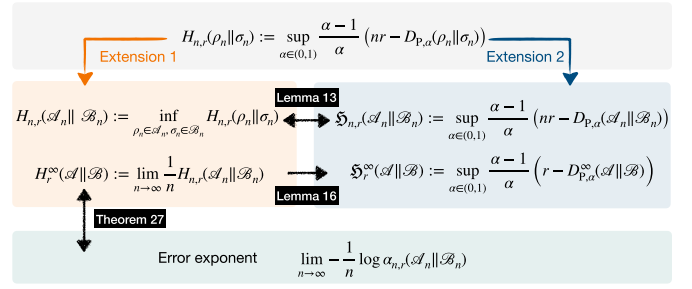


Fig. 3. Summary of results in the error exponent regime. Quantitative relationships between the various quantities are indicated by black arrows: the quantity at the tail of an arrow is always greater than or equal to the one at the head. A double arrow indicates an equality.

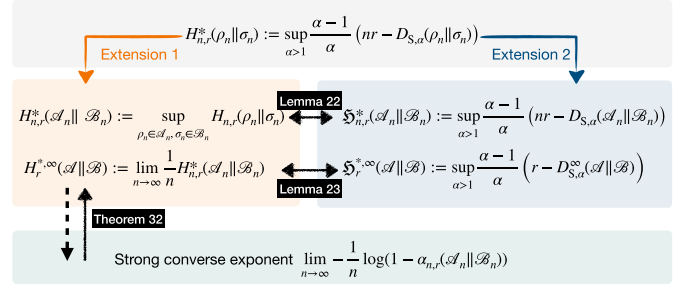


Fig. 4. Summary of results in the strong converse exponent regime. Quantitative relationships between the various quantities are indicated by black arrows: the quantity at the tail of an arrow is always greater than or equal to the one at the head. A double arrow indicates an equality. A dashed arrow indicates partial progress when $\mathcal{A}_n$ is a singleton.

$H_r^\infty(\mathcal{A} \| \mathcal{B}) \geq \mathfrak{H}_r^\infty(\mathcal{A} \| \mathcal{B})$ for their regularizations. As a main result, we prove that the error exponent is completely characterized by the regularized quantum Hoeffding divergence defined via the first approach.

In the strong converse regime, we likewise investigate two natural extensions of the quantum Hoeffding anti-divergence (see Figure 4). The first approach considers the maximal anti-divergence between the sets, denoted $H_{n,r}^*(\mathcal{A}_n \| \mathcal{B}_n)$. The second approach leverages the explicit formula for the anti-divergence in terms of the sandwiched Rényi divergences: we first extend the sandwiched Rényi divergences to sets, $D_{s,\alpha}(\mathcal{A}_n \| \mathcal{B}_n)$, and then define the anti-divergence as $\mathfrak{H}_{n,r}^*(\mathcal{A}_n \| \mathcal{B}_n)$. We establish that these two approaches are equivalent for both finite n and in the asymptotic setting. As a main result, we prove that the strong converse exponent is lower bounded by the regularized quantum Hoeffding anti-divergence in general, and we provide a matching upper bound when the null hypothesis is a singleton, under additional assumptions.

As the error exponent and strong converse exponent regimes provide a finer characterization of the trade-off between Type-I and Type-II errors than the Stein regime, we apply the results above to refine and extend several existing studies. In particular, we recover the quantum Stein's lemma for composite correlated hypotheses as established in [15], and further strengthen it by providing explicit convergence rates for the Type-I error. Note that the Stein exponent only indicates whether the Type-I error decays, but does not quantify how fast it decays. Specifically, under the same assumptions as in

[15], we establish the following refined asymptotic trade-off between Type-I and Type-II errors:

$$\liminf_{n \rightarrow \infty} -\frac{1}{n} \log \alpha_{n,r}(\mathcal{A}_n \| \mathcal{B}_n) \geq \mathfrak{H}_r^\infty(\mathcal{A} \| \mathcal{B}) > 0, \quad (18)$$

for any $0 < r < D^\infty(\mathcal{A} \| \mathcal{B})$, and

$$\liminf_{n \rightarrow \infty} -\frac{1}{n} \log(1 - \alpha_{n,r}(\mathcal{A}_n \| \mathcal{B}_n)) \geq \mathfrak{H}_r^{*,\infty}(\mathcal{A} \| \mathcal{B}) > 0, \quad (19)$$

for any $r > D^\infty(\mathcal{A} \| \mathcal{B})$, where $D^\infty(\mathcal{A} \| \mathcal{B})$ is the regularized relative entropy between the sets. These indicate that any Type-II error exponent r below $D^\infty(\mathcal{A} \| \mathcal{B})$ is achievable, with the corresponding Type-I error decaying exponentially at a rate at least $\mathfrak{H}_r^\infty(\mathcal{A} \| \mathcal{B})$. Conversely, if the Type-II error exponent r exceeds $D^\infty(\mathcal{A} \| \mathcal{B})$, the Type-I error inevitably converges to one exponentially fast, with a rate at least $\mathfrak{H}_r^{*,\infty}(\mathcal{A} \| \mathcal{B})$. Thus, the regularized quantum relative entropy delineates a sharp threshold for the asymptotic trade-off in hypothesis testing between two sets of quantum states.

Besides this, we also provide a few more explicit applications, including: (i) finding counterexamples to the continuity of the regularized Petz Rényi divergence and Hoeffding divergence; (ii) deriving error exponents for adversarial quantum channel discrimination; and (iii) obtaining error exponents for resource detection problems in coherence theory and entanglement theory.

The rest of the paper is organized as follows. In Section II we introduce the notation and review a few quantum divergences, together with their extensions to sets of states. In Section III we introduce two natural extensions of the Hoeffding divergence and anti-divergence to sets and prove their equivalence or quantitative relation. In Section IV we formulate asymmetric hypothesis testing for composite correlated hypotheses and show how the worst-case problem reduces to pairwise optimizations. Section V establishes the quantum Hoeffding bound for composite correlated hypotheses, while Section VI studies the strong converse exponent. Section VII discusses applications and implications of our results. In Section VIII we conclude the paper and discuss a few open questions for future exploration. Some technical lemmas are delegated to the appendices.

II. PRELIMINARIES

A. Notation

Throughout this work, we adopt the following notational conventions. Finite-dimensional Hilbert spaces are denoted by $\mathcal{H}$, with $|\mathcal{H}|$ indicating their dimension. The set of all linear operators on $\mathcal{H}$ is denoted by $\mathcal{L}(\mathcal{H})$, while $\mathcal{H}(\mathcal{H})$ and $\mathcal{H}_+(\mathcal{H})$ denote the sets of Hermitian and positive semidefinite operators on $\mathcal{H}$, respectively. The set of density operators (i.e., positive semidefinite operators with unit trace) on $\mathcal{H}$ is denoted by $\mathcal{D}(\mathcal{H})$. Calligraphic letters such as $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$ are used to represent sets of linear operators. Unless otherwise specified, all logarithms are taken to base two and denoted by $\log(x)$. The positive semidefinite ordering is written as $X \geq Y$ if and only if $X - Y \geq 0$. The absolute value of an operator X is defined by $|X| := (X^\dagger X)^{1/2}$. For a Hermitian operator X with spectral decomposition $X = \sum_i x_i E_i$, the projection onto the

non-negative eigenspaces is denoted by $\{X \geq 0\} := \sum_{x_i \geq 0} E_i$. Similarly, $\{X > 0\} := \sum_{x_i > 0} E_i$.

B. Quantum Divergences

A functional $\mathbb{D} : \mathcal{D} \times \mathcal{H}_+ \rightarrow \mathbb{R}$ is called a *quantum divergence* if it satisfies the data-processing inequality: for any completely positive and trace-preserving (CPTP) map $\mathcal{E}$ and any $(\rho, \sigma) \in \mathcal{D} \times \mathcal{H}_+$, it holds that $\mathbb{D}(\mathcal{E}(\rho) \| \mathcal{E}(\sigma)) \leq \mathbb{D}(\rho \| \sigma)$. In the following, we introduce several quantum divergences that will be used throughout this work. We also define quantum divergences between two sets of quantum states.

Definition 1 (Umegaki relative entropy [30]): For any $\rho \in \mathcal{D}$ and $\sigma \in \mathcal{H}_+$, the Umegaki relative entropy is defined by

$$D(\rho \| \sigma) := \text{Tr}[\rho(\log \rho - \log \sigma)], \quad (20)$$

if $\text{supp}(\rho) \subseteq \text{supp}(\sigma)$ and ∞ otherwise.

Definition 2 (Petz Rényi divergence [31]): Let $\alpha \in (0, 1) \cup (1, \infty)$. For any $\rho \in \mathcal{D}$ and $\sigma \in \mathcal{H}_+$, the Petz Rényi divergence is defined by

$$D_{p,\alpha}(\rho \| \sigma) := \frac{1}{\alpha - 1} \log Q_\alpha(\rho \| \sigma), \quad (21)$$

with $Q_\alpha(\rho \| \sigma) := \text{Tr}[\rho^\alpha \sigma^{1-\alpha}]$, if $\alpha \in (0, 1)$ and $\rho \not\perp \sigma$, or if $\alpha \in (1, \infty)$ and $\text{supp}(\rho) \subseteq \text{supp}(\sigma)$; it is defined by ∞ otherwise.

Definition 3 (Sandwiched Rényi divergence [32], [33]): Let $\alpha \in (0, 1) \cup (1, \infty)$. For any $\rho \in \mathcal{D}$ and $\sigma \in \mathcal{H}_+$, the sandwiched Rényi divergence is defined by

$$D_{s,\alpha}(\rho \| \sigma) := \frac{1}{\alpha - 1} \log \text{Tr} \left[\sigma^{\frac{1-\alpha}{2\alpha}} \rho \sigma^{\frac{1-\alpha}{2\alpha}} \right]^\alpha, \quad (22)$$

if $\alpha \in (0, 1)$ and $\rho \not\perp \sigma$, or if $\alpha \in (1, \infty)$ and $\text{supp}(\rho) \subseteq \text{supp}(\sigma)$; it is defined by ∞ otherwise.

Definition 4 (Quantum divergence between two sets of states): Let $\mathbb{D}$ be a quantum divergence between two quantum states. Let $\mathcal{H}$ be a finite-dimensional Hilbert space. Then for any sets $\mathcal{A}, \mathcal{B} \subseteq \mathcal{D}(\mathcal{H})$, the quantum divergence between these two sets is defined by

$$\mathbb{D}(\mathcal{A} \| \mathcal{B}) := \inf_{\substack{\rho \in \mathcal{A} \\ \sigma \in \mathcal{B}}} \mathbb{D}(\rho \| \sigma). \quad (23)$$

Let $\mathcal{A} = \{\mathcal{A}_n\}_{n \in \mathbb{N}}$ and $\mathcal{B} = \{\mathcal{B}_n\}_{n \in \mathbb{N}}$ be two sequences of sets of quantum states,² where each $\mathcal{A}_n, \mathcal{B}_n \subseteq \mathcal{D}(\mathcal{H}^{\otimes n})$. the regularized divergence between these sequences is defined by

$$\underline{\mathbb{D}}^\infty(\mathcal{A} \| \mathcal{B}) := \liminf_{n \rightarrow \infty} \frac{1}{n} \mathbb{D}(\mathcal{A}_n \| \mathcal{B}_n), \quad (24)$$

$$\overline{\mathbb{D}}^\infty(\mathcal{A} \| \mathcal{B}) := \limsup_{n \rightarrow \infty} \frac{1}{n} \mathbb{D}(\mathcal{A}_n \| \mathcal{B}_n). \quad (25)$$

If the limit exists, we define the regularized divergence as

$$\mathbb{D}^\infty(\mathcal{A} \| \mathcal{B}) := \lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{D}(\mathcal{A}_n \| \mathcal{B}_n). \quad (26)$$

Remark 1: Note that if $\mathbb{D}$ is lower semicontinuous (which is true for most quantum divergences of interest), and $\mathcal{A}$ and $\mathcal{B}$ are compact sets, the infimum in the above expression is always attained and can thus be replaced by a minimization (see, e.g., [34, Lemma 2.8]).

²We abuse the notation $\mathcal{A}, \mathcal{B}$ to refer both to sets of states and to sequences of such sets, depending on the context.

In many practical scenarios, the sequences of sets of interest are not arbitrary but possess a structure that is compatible with tensor products. This property, known as *stability* (or *closeness*) under tensor product, is formalized as follows.

Definition 5 (Stable sequence): Let $\mathcal{A} \subseteq \mathcal{H}_+(\mathcal{H}_1)$, $\mathcal{B} \subseteq \mathcal{H}_+(\mathcal{H}_2)$, and $\mathcal{C} \subseteq \mathcal{H}_+(\mathcal{H}_1 \otimes \mathcal{H}_2)$. We say that $\{\mathcal{A}, \mathcal{B}, \mathcal{C}\}$ is stable under tensor product if, for any $X_1 \in \mathcal{A}$ and $X_2 \in \mathcal{B}$, it holds that $X_1 \otimes X_2 \in \mathcal{C}$. In short, we write $\mathcal{A} \otimes \mathcal{B} \subseteq \mathcal{C}$. A sequence of sets $\{\mathcal{C}_n\}_{n \in \mathbb{N}}$ with $\mathcal{C}_n \subseteq \mathcal{H}_+(\mathcal{H}^{\otimes n})$ is stable under tensor product if $\mathcal{C}_n \otimes \mathcal{C}_m \subseteq \mathcal{C}_{n+m}$ for all $n, m \in \mathbb{N}$.

Remark 2: If the divergence $\mathbb{D}$ is subadditive under tensor product states, its extension to sets of states is also subadditive for stable sequences of sets [15, Lemma 26]. This implies the existence of the regularized divergence by Fekete's lemma,

$$\mathbb{D}^\infty(\mathcal{A} \parallel \mathcal{B}) = \underline{\mathbb{D}}^\infty(\mathcal{A} \parallel \mathcal{B}) \quad (27)$$

$$= \overline{\mathbb{D}}^\infty(\mathcal{A} \parallel \mathcal{B}) = \inf_{n \in \mathbb{N}} \frac{1}{n} \mathbb{D}(\mathcal{A}_n \parallel \mathcal{B}_n). \quad (28)$$

III. Hoeffding Divergence and Anti-Divergence for Sets of States

In this section, we develop extensions of the quantum Hoeffding divergence and anti-divergence to sets of quantum states. Two natural approaches arise for this purpose. The first approach treats the Hoeffding divergence as a quantum divergence and extends it to sets via Definition 4. The second approach leverages the explicit formula for the Hoeffding divergence in terms of the Petz Rényi divergences. Analogous constructions apply to the Hoeffding anti-divergence.

A. Quantum Hoeffding Divergence

Definition 6: Let $\mathcal{H}$ be a finite-dimensional Hilbert space, $r > 0$ be a real number, and $n \in \mathbb{N}$. Let $\rho_n, \sigma_n \in \mathcal{D}(\mathcal{H}^{\otimes n})$. The quantum Hoeffding divergence is defined by

$$H_{n,r}(\rho_n \parallel \sigma_n) := \sup_{\alpha \in (0,1)} \frac{\alpha - 1}{\alpha} (nr - D_{p,\alpha}(\rho_n \parallel \sigma_n)). \quad (29)$$

Lemma 1 (Subadditivity): For any $\rho_m, \sigma_m \in \mathcal{D}(\mathcal{H}^{\otimes m})$ and $\rho_n, \sigma_n \in \mathcal{D}(\mathcal{H}^{\otimes n})$, it holds that

$$\begin{aligned} H_{(m+n),r}(\rho_m \otimes \rho_n \parallel \sigma_m \otimes \sigma_n) \\ \leq H_{m,r}(\rho_m \parallel \sigma_m) + H_{n,r}(\rho_n \parallel \sigma_n). \end{aligned} \quad (30)$$

Proof: This can be seen as follows:

$$\begin{aligned} H_{(m+n),r}(\rho_m \otimes \rho_n \parallel \sigma_m \otimes \sigma_n) \\ = \sup_{\alpha \in (0,1)} \frac{\alpha - 1}{\alpha} \left\{ (m+n)r \right. \\ \left. - D_{p,\alpha}(\rho_m \otimes \rho_n \parallel \sigma_m \otimes \sigma_n) \right\} \end{aligned} \quad (31)$$

$$\begin{aligned} = \sup_{\alpha \in (0,1)} \frac{\alpha - 1}{\alpha} \left\{ (m+n)r \right. \\ \left. - D_{p,\alpha}(\rho_m \parallel \sigma_m) - D_{p,\alpha}(\rho_n \parallel \sigma_n) \right\} \end{aligned} \quad (32)$$

$$\leq \sup_{\alpha \in (0,1)} \frac{\alpha - 1}{\alpha} \left\{ mr - D_{p,\alpha}(\rho_m \parallel \sigma_m) \right\}$$

$$+ \sup_{\alpha \in (0,1)} \frac{\alpha - 1}{\alpha} \left\{ nr - D_{p,\alpha}(\rho_n \parallel \sigma_n) \right\} \quad (33)$$

$$= H_{m,r}(\rho_m \parallel \sigma_m) + H_{n,r}(\rho_n \parallel \sigma_n), \quad (34)$$

where the second equality uses the additivity of $D_{p,\alpha}$ under tensor product states, and the inequality follows from splitting the supremum over α for each term.

Definition 7 (Quantum Hoeffding divergence between sets of states): Let $\mathcal{H}$ be a finite-dimensional Hilbert space, and $r > 0$ be a real number, $n \in \mathbb{N}$. Let $\mathcal{A}_n, \mathcal{B}_n \subseteq \mathcal{D}(\mathcal{H}^{\otimes n})$ be two sets of quantum states. Two variants of the quantum Hoeffding divergence between these sets are defined by

$$H_{n,r}(\mathcal{A}_n \parallel \mathcal{B}_n) := \inf_{\substack{\rho_n \in \mathcal{A}_n \\ \sigma_n \in \mathcal{B}_n}} H_{n,r}(\rho_n \parallel \sigma_n), \quad (35)$$

$$\mathfrak{H}_{n,r}(\mathcal{A}_n \parallel \mathcal{B}_n) := \sup_{\alpha \in (0,1)} \frac{\alpha - 1}{\alpha} (nr - D_{p,\alpha}(\mathcal{A}_n \parallel \mathcal{B}_n)), \quad (36)$$

where $D_{p,\alpha}(\mathcal{A}_n \parallel \mathcal{B}_n)$ is defined by in Definition 4. Moreover, let $\mathcal{A} = \{\mathcal{A}_n\}_{n \in \mathbb{N}}$ and $\mathcal{B} = \{\mathcal{B}_n\}_{n \in \mathbb{N}}$ be two sequences of sets of quantum states, where each $\mathcal{A}_n, \mathcal{B}_n \subseteq \mathcal{D}(\mathcal{H}^{\otimes n})$. The regularized quantum Hoeffding divergences between these sequences are defined by

$$\underline{H}_r(\mathcal{A} \parallel \mathcal{B}) := \liminf_{n \rightarrow \infty} \frac{1}{n} H_{n,r}(\mathcal{A}_n \parallel \mathcal{B}_n), \quad (37)$$

$$\overline{H}_r(\mathcal{A} \parallel \mathcal{B}) := \limsup_{n \rightarrow \infty} \frac{1}{n} H_{n,r}(\mathcal{A}_n \parallel \mathcal{B}_n), \quad (38)$$

$$\underline{\mathfrak{H}}_r(\mathcal{A} \parallel \mathcal{B}) := \sup_{\alpha \in (0,1)} \frac{\alpha - 1}{\alpha} (r - \underline{D}_{p,\alpha}(\mathcal{A} \parallel \mathcal{B})), \quad (39)$$

$$\overline{\mathfrak{H}}_r(\mathcal{A} \parallel \mathcal{B}) := \sup_{\alpha \in (0,1)} \frac{\alpha - 1}{\alpha} (r - \overline{D}_{p,\alpha}(\mathcal{A} \parallel \mathcal{B})), \quad (40)$$

where $\underline{D}_{p,\alpha}(\mathcal{A} \parallel \mathcal{B})$ and $\overline{D}_{p,\alpha}(\mathcal{A} \parallel \mathcal{B})$ are defined by in Definition 4. If the limits exist, we define the regularized quantum Hoeffding divergence as

$$H_r^\infty(\mathcal{A} \parallel \mathcal{B}) := \lim_{n \rightarrow \infty} \frac{1}{n} H_{n,r}(\mathcal{A}_n \parallel \mathcal{B}_n), \quad (41)$$

$$\mathfrak{H}_r^\infty(\mathcal{A} \parallel \mathcal{B}) := \sup_{\alpha \in (0,1)} \frac{\alpha - 1}{\alpha} (r - D_{p,\alpha}^\infty(\mathcal{A} \parallel \mathcal{B})), \quad (42)$$

where $D_{p,\alpha}^\infty(\mathcal{A} \parallel \mathcal{B})$ is defined by in Definition 4.

Remark 3 (Attainment): Since $D_{p,\alpha}(\rho_n \parallel \sigma_n)$ is lower semi-continuous in (ρ_n, σ_n) for any fixed α [35, Proposition III.11], it follows from Lemma 13 that $H_{n,r}(\rho_n \parallel \sigma_n)$ is also lower semicontinuous in (ρ_n, σ_n) . Consequently, if $\mathcal{A}_n$ and $\mathcal{B}_n$ are compact sets, we know from Lemma 12 that the infimum in the definition of $H_{n,r}(\mathcal{A}_n \parallel \mathcal{B}_n)$ is attained.

Remark 4 (Limit existence): The quantum Hoeffding divergence is subadditive under tensor product by Lemma 1. So this property extends to stable sequences of sets by Remark 2,

$$\begin{aligned} H_{(m+n),r}(\mathcal{A}_{m+n} \parallel \mathcal{B}_{m+n}) \\ \leq H_{m,r}(\mathcal{A}_m \parallel \mathcal{B}_m) + H_{n,r}(\mathcal{A}_n \parallel \mathcal{B}_n). \end{aligned} \quad (43)$$

As a consequence, the regularized quantum Hoeffding divergence exists and satisfies

$$H_r^\infty(\mathcal{A} \parallel \mathcal{B}) = \overline{H}_r^\infty(\mathcal{A} \parallel \mathcal{B}) \quad (44)$$

$$= \underline{H}_r^\infty(\mathcal{A} \parallel \mathcal{B}) = \inf_{n \geq 1} \frac{1}{n} H_{n,r}(\mathcal{A}_n \parallel \mathcal{B}_n). \quad (45)$$

Similarly, due to the additivity of $D_{p,\alpha}$ under tensor product states, the regularized quantum Hoeffding divergence $\mathfrak{H}_r^\infty(\mathcal{A}||\mathcal{B})$ also exists for stable sequences and satisfies

$$\mathfrak{H}_r^\infty(\mathcal{A}||\mathcal{B}) = \overline{\mathfrak{H}}_r^\infty(\mathcal{A}||\mathcal{B}) = \underline{\mathfrak{H}}_r^\infty(\mathcal{A}||\mathcal{B}). \quad (46)$$

The following results establish the relationship between the two variants of the quantum Hoeffding divergence for sets and sequences of sets.

Lemma 2 (Finite equivalence): Let $\mathcal{H}$ be a finite-dimensional Hilbert space, and $r > 0$ be a real number, $n \in \mathbb{N}$. Let $\mathcal{A}_n, \mathcal{B}_n \subseteq \mathcal{D}(\mathcal{H}^{\otimes n})$ be two convex compact sets of quantum states. Then

$$H_{n,r}(\mathcal{A}_n||\mathcal{B}_n) = \mathfrak{H}_{n,r}(\mathcal{A}_n||\mathcal{B}_n). \quad (47)$$

Proof: This was previously established in [20, Lemma II.8].

□

Remark 5 (Computability): For any fixed $\alpha \in (0, 1)$, the function $Q_\alpha(\rho_n||\sigma_n)$ is jointly concave in (ρ_n, σ_n) . Consequently, the quasi-divergence $Q_\alpha(\mathcal{A}_n||\mathcal{B}_n)$ can be efficiently computed using the QICS package [36], provided that $\mathcal{A}_n$ and $\mathcal{B}_n$ admit semidefinite representations. If the sets $\mathcal{A}_n$ and $\mathcal{B}_n$ exhibit additional symmetries, the computational complexity can be further reduced. With this, $\mathfrak{H}_{n,r}(\mathcal{A}_n||\mathcal{B}_n)$ can be efficiently evaluated by scanning over $\alpha \in (0, 1)$.

Remark 6 (Alternative expression): By the duality relation of the Petz Rényi divergences, we have the following alternative expression for the Hoeffding divergence:

$$\begin{aligned} H_{n,r}(\mathcal{A}_n||\mathcal{B}_n) &= \sup_{\alpha \in (0,1)} \frac{\alpha-1}{\alpha} (nr - D_{p,\alpha}(\mathcal{A}_n||\mathcal{B}_n)) \end{aligned} \quad (48)$$

$$= \sup_{\alpha \in (0,1)} \frac{\alpha-1}{\alpha} \left(nr + \frac{\alpha}{\alpha-1} D_{p,1-\alpha}(\mathcal{B}_n||\mathcal{A}_n) \right) \quad (49)$$

$$= \sup_{\alpha \in (0,1)} \left(\frac{\alpha-1}{\alpha} nr + D_{p,1-\alpha}(\mathcal{B}_n||\mathcal{A}_n) \right) \quad (50)$$

$$= \sup_{\alpha \in (0,1)} \left(-\frac{\alpha}{1-\alpha} nr + D_{p,\alpha}(\mathcal{B}_n||\mathcal{A}_n) \right), \quad (51)$$

where we replace α by $1-\alpha$ in the last equality. This alternative formulation is particularly useful when analyzing the Stein's exponent by taking the limit $r \rightarrow 0$.

Lemma 3 (Asymptotic relation): Let $\mathcal{H}$ be a finite-dimensional Hilbert space, and $r > 0$ be a real number. Let $\mathcal{A} = \{\mathcal{A}_n\}_{n \in \mathbb{N}}$ and $\mathcal{B} = \{\mathcal{B}_n\}_{n \in \mathbb{N}}$ be two stable sequences of convex compact sets of quantum states, where each $\mathcal{A}_n, \mathcal{B}_n \subseteq \mathcal{D}(\mathcal{H}^{\otimes n})$. Then for any $n \in \mathbb{N}$,

$$\frac{1}{n} \mathfrak{H}_{n,r}(\mathcal{A}_n||\mathcal{B}_n) \geq H_r^\infty(\mathcal{A}||\mathcal{B}) \geq \mathfrak{H}_r^\infty(\mathcal{A}||\mathcal{B}). \quad (52)$$

Proof: The existence of the regularizations is ensured by Remark 4. Then we have the following chain of inequalities:

$$\begin{aligned} H_r^\infty(\mathcal{A}||\mathcal{B}) &= \inf_{n \geq 1} \frac{1}{n} H_{n,r}(\mathcal{A}_n||\mathcal{B}_n) \end{aligned} \quad (53)$$

$$= \inf_{n \geq 1} \sup_{\alpha \in (0,1)} \frac{\alpha-1}{\alpha} \left(r - \frac{1}{n} D_{p,\alpha}(\mathcal{A}_n||\mathcal{B}_n) \right) \quad (54)$$

$$\geq \sup_{\alpha \in (0,1)} \inf_{n \geq 1} \frac{\alpha-1}{\alpha} \left(r - \frac{1}{n} D_{p,\alpha}(\mathcal{A}_n||\mathcal{B}_n) \right) \quad (55)$$

$$= \sup_{\alpha \in (0,1)} \frac{\alpha-1}{\alpha} \left(r - \inf_{n \geq 1} \frac{1}{n} D_{p,\alpha}(\mathcal{A}_n||\mathcal{B}_n) \right) \quad (56)$$

$$= \sup_{\alpha \in (0,1)} \frac{\alpha-1}{\alpha} (r - D_{p,\alpha}^\infty(\mathcal{A}||\mathcal{B})) \quad (57)$$

$$= \mathfrak{H}_r^\infty(\mathcal{A}||\mathcal{B}), \quad (58)$$

where the first equality follows from Remark 4, the second equality follows from Lemma 2, the inequality follows by minimax inequality, the fourth equality follows from Remark 4, and the last equality follows by definition. Moreover, we have

$$H_r^\infty(\mathcal{A}||\mathcal{B}) = \inf_{n \geq 1} \frac{1}{n} H_{n,r}(\mathcal{A}_n||\mathcal{B}_n) \quad (59)$$

$$\leq \frac{1}{n} H_{n,r}(\mathcal{A}_n||\mathcal{B}_n) = \frac{1}{n} \mathfrak{H}_{n,r}(\mathcal{A}_n||\mathcal{B}_n). \quad (60)$$

This completes the proof.

Remark 7: If the minimax equality in the above proof can be established, then we would have $H_r^\infty(\mathcal{A}||\mathcal{B}) = \mathfrak{H}_r^\infty(\mathcal{A}||\mathcal{B})$. However, this appears to be challenging, as existing minimax theorems typically require at least one of the spaces to be compact—a condition that is not directly satisfied here. Moreover, as we show later that the Petz Rényi divergence and the Hoeffding divergence between two sets can be nonadditive in general (see Section VII-C), which further complicates the situation. Nevertheless, in the special case where the Petz Rényi divergence is additive for all $\alpha \in (0, 1)$, i.e., $\frac{1}{n} D_{p,\alpha}(\mathcal{A}_n||\mathcal{B}_n) = D_{p,\alpha}(\mathcal{A}_1||\mathcal{B}_1)$, we have that

$$H_r^\infty(\mathcal{A}||\mathcal{B}) = \mathfrak{H}_r^\infty(\mathcal{A}||\mathcal{B}) \quad (61)$$

$$= H_{1,r}(\mathcal{A}_1||\mathcal{B}_1) = \mathfrak{H}_{1,r}(\mathcal{A}_1||\mathcal{B}_1). \quad (62)$$

B. Quantum Hoeffding Anti-Divergence

Analogous to the quantum Hoeffding divergence, we can also define the quantum Hoeffding anti-divergence for sets of quantum states.

Definition 8: Let $\mathcal{H}$ be a finite-dimensional Hilbert space, $r > 0$ be a real number, $n \in \mathbb{N}$. Let $\rho_n, \sigma_n \in \mathcal{D}(\mathcal{H}^{\otimes n})$. The quantum Hoeffding anti-divergence is defined by

$$H_{n,r}^*(\rho_n||\sigma_n) := \sup_{\alpha > 1} \frac{\alpha-1}{\alpha} (nr - D_{s,\alpha}(\rho_n||\sigma_n)). \quad (63)$$

Lemma 4 (Subadditivity): For any $\rho_m, \sigma_m \in \mathcal{D}(\mathcal{H}^{\otimes m})$ and $\rho_n, \sigma_n \in \mathcal{D}(\mathcal{H}^{\otimes n})$, it holds that

$$\begin{aligned} H_{(m+n),r}^*(\rho_m \otimes \rho_n||\sigma_m \otimes \sigma_n) &\leq H_{m,r}^*(\rho_m||\sigma_m) + H_{n,r}^*(\rho_n||\sigma_n). \end{aligned} \quad (64)$$

Proof: This can be seen as follows:

$$\begin{aligned} H_{(m+n),r}^*(\rho_m \otimes \rho_n||\sigma_m \otimes \sigma_n) &= \sup_{\alpha > 1} \frac{\alpha-1}{\alpha} \left\{ (m+n)r \right. \\ &\quad \left. - D_{s,\alpha}(\rho_m \otimes \rho_n||\sigma_m \otimes \sigma_n) \right\} \end{aligned} \quad (65)$$

$$= \sup_{\alpha > 1} \frac{\alpha-1}{\alpha} \left\{ (m+n)r - D_{s,\alpha}(\rho_m||\sigma_m) - D_{s,\alpha}(\rho_n||\sigma_n) \right\} \quad (66)$$

$$\leq \sup_{\alpha > 1} \frac{\alpha - 1}{\alpha} \left\{ mr - D_{s,\alpha}(\rho_m \| \sigma_m) \right\} + \sup_{\alpha > 1} \frac{\alpha - 1}{\alpha} \left\{ nr - D_{s,\alpha}(\rho_n \| \sigma_n) \right\} \quad (67)$$

$$= H_{m,r}^*(\rho_m \| \sigma_m) + H_{n,r}^*(\rho_n \| \sigma_n), \quad (68)$$

where the second equality uses the additivity of $D_{s,\alpha}$ under tensor product states, and the inequality follows from splitting the supremum over α for each term. $\square$

Definition 9 (Quantum Hoeffding anti-divergence between sets of states): Let $\mathcal{H}$ be a finite-dimensional Hilbert space, and $r > 0$ be a real number, $n \in \mathbb{N}$. Let $\mathcal{A}_n, \mathcal{B}_n \subseteq \mathcal{D}(\mathcal{H}^{\otimes n})$. Two variants of the quantum Hoeffding anti-divergence between these sets are defined by³

$$H_{n,r}^*(\mathcal{A}_n \| \mathcal{B}_n) := \sup_{\substack{\rho_n \in \mathcal{A}_n \\ \sigma_n \in \mathcal{B}_n}} H_{n,r}^*(\rho_n \| \sigma_n), \quad (69)$$

$$\mathfrak{H}_{n,r}^*(\mathcal{A}_n \| \mathcal{B}_n) := \sup_{\alpha > 1} \frac{\alpha - 1}{\alpha} (nr - D_{s,\alpha}(\mathcal{A}_n \| \mathcal{B}_n)), \quad (70)$$

where $D_{s,\alpha}(\mathcal{A}_n \| \mathcal{B}_n)$ is defined by in Definition 4. Moreover, let $\mathcal{A} = \{\mathcal{A}_n\}_{n \in \mathbb{N}}$ and $\mathcal{B} = \{\mathcal{B}_n\}_{n \in \mathbb{N}}$ be two sequences of sets of quantum states, where each $\mathcal{A}_n, \mathcal{B}_n \subseteq \mathcal{D}(\mathcal{H}^{\otimes n})$. The regularized quantum Hoeffding anti-divergences between these sequences are defined by

$$\underline{H}_r^{*,\infty}(\mathcal{A} \| \mathcal{B}) := \liminf_{n \rightarrow \infty} \frac{1}{n} H_{n,r}^*(\mathcal{A}_n \| \mathcal{B}_n), \quad (71)$$

$$\overline{H}_r^{*,\infty}(\mathcal{A} \| \mathcal{B}) := \limsup_{n \rightarrow \infty} \frac{1}{n} H_{n,r}^*(\mathcal{A}_n \| \mathcal{B}_n), \quad (72)$$

$$\underline{\mathfrak{H}}_r^{*,\infty}(\mathcal{A} \| \mathcal{B}) := \sup_{\alpha > 1} \frac{\alpha - 1}{\alpha} (r - \underline{D}_{s,\alpha}^\infty(\mathcal{A} \| \mathcal{B})), \quad (73)$$

$$\overline{\mathfrak{H}}_r^{*,\infty}(\mathcal{A} \| \mathcal{B}) := \sup_{\alpha > 1} \frac{\alpha - 1}{\alpha} (r - \overline{D}_{s,\alpha}^\infty(\mathcal{A} \| \mathcal{B})). \quad (74)$$

where $\underline{D}_{s,\alpha}^\infty(\mathcal{A} \| \mathcal{B})$ and $\overline{D}_{s,\alpha}^\infty(\mathcal{A} \| \mathcal{B})$ are defined by in Definition 4. If the limits exist, we define the regularized Hoeffding divergence as

$$H_r^{*,\infty}(\mathcal{A} \| \mathcal{B}) := \lim_{n \rightarrow \infty} \frac{1}{n} H_{n,r}^*(\mathcal{A}_n \| \mathcal{B}_n), \quad (75)$$

$$\mathfrak{H}_r^{*,\infty}(\mathcal{A} \| \mathcal{B}) := \sup_{\alpha > 1} \frac{\alpha - 1}{\alpha} (r - D_{s,\alpha}^\infty(\mathcal{A} \| \mathcal{B})), \quad (76)$$

where $D_{s,\alpha}^\infty(\mathcal{A} \| \mathcal{B})$ is defined by in Definition 4.

Remark 8 (Attainment): It is known that $H_{n,r}^*(\rho_n \| \sigma_n)$ is upper semicontinuous in (ρ_n, σ_n) [20, Corollary V.16]. So the supremum in $H_{n,r}^*(\mathcal{A}_n \| \mathcal{B}_n)$ is achieved for any compact sets.

The following results aim to establish the relationship between the two variants of the quantum Hoeffding anti-divergence for sets and sequences of sets.

Lemma 5 (Finite equivalence): Let $\mathcal{H}$ be a finite-dimensional Hilbert space, $r > 0$ be a real number, and $n \in \mathbb{N}$. Let $\mathcal{A}_n, \mathcal{B}_n \subseteq \mathcal{D}(\mathcal{H}^{\otimes n})$. Then it holds that

$$H_{n,r}^*(\mathcal{A}_n \| \mathcal{B}_n) = \mathfrak{H}_{n,r}^*(\mathcal{A}_n \| \mathcal{B}_n). \quad (77)$$

Proof: By definition, we have

$$H_{n,r}^*(\mathcal{A}_n \| \mathcal{B}_n)$$

³The anti-divergence is monotone non-decreasing under CPTP maps, so the extension to sets is based on the supremum rather than the infimum.

$$= \sup_{\substack{\rho_n \in \mathcal{A}_n \\ \sigma_n \in \mathcal{B}_n}} \sup_{\alpha > 1} \frac{\alpha - 1}{\alpha} (nr - D_{s,\alpha}(\rho_n \| \sigma_n)) \quad (78)$$

$$= \sup_{\alpha > 1} \sup_{\substack{\rho_n \in \mathcal{A}_n \\ \sigma_n \in \mathcal{B}_n}} \frac{\alpha - 1}{\alpha} (nr - D_{s,\alpha}(\rho_n \| \sigma_n)) \quad (79)$$

$$= \sup_{\alpha > 1} \frac{\alpha - 1}{\alpha} \left(nr - \inf_{\substack{\rho_n \in \mathcal{A}_n \\ \sigma_n \in \mathcal{B}_n}} D_{s,\alpha}(\rho_n \| \sigma_n) \right) \quad (80)$$

$$= \sup_{\alpha > 1} \frac{\alpha - 1}{\alpha} (nr - D_{s,\alpha}(\mathcal{A}_n \| \mathcal{B}_n)) \quad (81)$$

$$= \mathfrak{H}_{n,r}^*(\mathcal{A}_n \| \mathcal{B}_n), \quad (82)$$

where in the second line we exchange the two suprema. $\square$

It is important to note that $H_{n,r}^*(\mathcal{A}_n \| \mathcal{B}_n)$ is defined by a supremum over the feasible states, which makes its additivity property for stable sequences unclear—even though we know the Hoeffding anti-divergence for states is subadditive. As a result, we cannot directly apply Remark 2 as in previous discussions of Remark 4. Nevertheless, the following result shows that the regularization $H_r^{*,\infty}$ indeed exists for stable sequences and coincides with $\mathfrak{H}_r^{*,\infty}$ in general.

Lemma 6 (Asymptotic equivalence): Let $\mathcal{H}$ be a finite-dimensional Hilbert space, and $r > 0$ be a real number. Let $\mathcal{A} = \{\mathcal{A}_n\}_{n \in \mathbb{N}}$ and $\mathcal{B} = \{\mathcal{B}_n\}_{n \in \mathbb{N}}$ be two stable sequences of sets of quantum states, where each $\mathcal{A}_n, \mathcal{B}_n \subseteq \mathcal{D}(\mathcal{H}^{\otimes n})$. Then it holds that

$$H_r^{*,\infty}(\mathcal{A} \| \mathcal{B}) = \mathfrak{H}_r^{*,\infty}(\mathcal{A} \| \mathcal{B}). \quad (83)$$

Proof: We have the following chain of inequalities:

$$\underline{H}_r^{*,\infty}(\mathcal{A} \| \mathcal{B}) = \liminf_{n \rightarrow \infty} \frac{1}{n} H_{n,r}^*(\mathcal{A}_n \| \mathcal{B}_n) \quad (84)$$

$$= \liminf_{n \rightarrow \infty} \frac{1}{n} \mathfrak{H}_{n,r}^*(\mathcal{A}_n \| \mathcal{B}_n) \quad (85)$$

$$= \liminf_{n \rightarrow \infty} \frac{1}{n} \sup_{\alpha > 1} \frac{\alpha - 1}{\alpha} (nr - D_{s,\alpha}(\mathcal{A}_n \| \mathcal{B}_n)) \quad (86)$$

$$\geq \sup_{\alpha > 1} \liminf_{n \rightarrow \infty} \frac{1}{n} \frac{\alpha - 1}{\alpha} (nr - D_{s,\alpha}(\mathcal{A}_n \| \mathcal{B}_n)) \quad (87)$$

$$= \sup_{\alpha > 1} \frac{\alpha - 1}{\alpha} \left(r - \limsup_{n \rightarrow \infty} \frac{1}{n} D_{s,\alpha}(\mathcal{A}_n \| \mathcal{B}_n) \right) \quad (88)$$

$$= \sup_{\alpha > 1} \frac{\alpha - 1}{\alpha} (r - D_{s,\alpha}^\infty(\mathcal{A} \| \mathcal{B})) \quad (89)$$

$$= \mathfrak{H}_r^{*,\infty}(\mathcal{A} \| \mathcal{B}), \quad (90)$$

where the second line follows from Lemma 5, the inequality follows by the fact that for any sequence of numbers $x_{\alpha,n}$, $\liminf_{n \rightarrow \infty} \sup_{\alpha > 1} x_{\alpha,n} \geq \sup_{\alpha > 1} \liminf_{n \rightarrow \infty} x_{\alpha,n}$, the second last line follows from Remark 2 and the stability of the sequences.

In the other direction, we have

$$\overline{H}_r^{*,\infty}(\mathcal{A} \| \mathcal{B}) = \limsup_{n \rightarrow \infty} \frac{1}{n} H_{n,r}^*(\mathcal{A}_n \| \mathcal{B}_n) \quad (91)$$

$$= \limsup_{n \rightarrow \infty} \frac{1}{n} \mathfrak{H}_{n,r}^*(\mathcal{A}_n \| \mathcal{B}_n) \quad (92)$$

$$= \limsup_{n \rightarrow \infty} \frac{1}{n} \sup_{\alpha > 1} \frac{\alpha - 1}{\alpha} (nr - D_{s,\alpha}(\mathcal{A}_n \| \mathcal{B}_n)), \quad (93)$$

where the second line follows from Lemma 5. Note that $D_{s,\alpha}(\mathcal{A}_n\|\mathcal{B}_n)$ is subadditive for stable sequences, so we have $\frac{1}{n}D_{s,\alpha}(\mathcal{A}_n\|\mathcal{B}_n) \geq D_{s,\alpha}^\infty(\mathcal{A}\|\mathcal{B})$ for any $n \in \mathbb{N}$. This gives

$$\begin{aligned} & \frac{1}{n} \sup_{\alpha>1} \frac{\alpha-1}{\alpha} (nr - D_{s,\alpha}(\mathcal{A}_n\|\mathcal{B}_n)) \\ & \leq \sup_{\alpha>1} \frac{\alpha-1}{\alpha} (r - D_{s,\alpha}^\infty(\mathcal{A}\|\mathcal{B})). \end{aligned} \quad (94)$$

Then taking the limit of n , we have

$$\bar{H}_r^{*,\infty}(\mathcal{A}\|\mathcal{B}) \leq \sup_{\alpha>1} \frac{\alpha-1}{\alpha} (r - D_{s,\alpha}^\infty(\mathcal{A}\|\mathcal{B})) \quad (95)$$

$$= \mathfrak{H}_r^{*,\infty}(\mathcal{A}\|\mathcal{B}). \quad (96)$$

Combining (90) and (95), we have the asserted result. $\square$

IV. OPTIMAL TYPE-I ERROR PROBABILITY

The error exponent and strong converse exponent regimes study the optimal behavior of the Type-I error provided that the Type-II error exponentially decays. Recall that the optimal Type-I error for hypothesis testing between two sets of quantum states, $\mathcal{A}_n$ and $\mathcal{B}_n$, is defined by

$$\begin{aligned} & \alpha_{n,r}(\mathcal{A}_n\|\mathcal{B}_n) \\ & := \min_{0 \leq M_n \leq I} \{ \alpha(\mathcal{A}_n, M_n) : \beta(\mathcal{B}_n, M_n) \leq 2^{-nr} \}, \end{aligned} \quad (97)$$

where the measurement M_n is chosen to minimize the worst-case Type-I error $\alpha(\mathcal{A}_n, M_n)$, subject to the constraint that the Type-II error $\beta(\mathcal{B}_n, M_n)$ decays exponentially at a rate r . In other words, the measurement must perform universally well for all states in $\mathcal{A}_n$ and $\mathcal{B}_n$.

The following shows that the optimal Type-I error for hypothesis testing between two sets of quantum states is precisely determined by the most challenging pair of states from these sets. This implies a universal measurement for $\mathcal{A}_n$ and $\mathcal{B}_n$ whose performance matches that of the optimal measurement for the worst-case pair of states, providing a key simplification that enables our subsequent analysis.

Lemma 7: Let $\mathcal{H}$ be a finite-dimensional Hilbert space, $r > 0$ be a real number, and $n \in \mathbb{N}$. Let $\mathcal{A}_n, \mathcal{B}_n \subseteq \mathcal{D}(\mathcal{H}^{\otimes n})$ be two convex sets of quantum states. Then it holds that

$$\alpha_{n,r}(\mathcal{A}_n\|\mathcal{B}_n) = \sup_{\substack{\rho_n \in \mathcal{A}_n \\ \sigma_n \in \mathcal{B}_n}} \alpha_{n,r}(\rho_n\|\sigma_n). \quad (98)$$

Proof: We begin by noting the following symmetry role between Type-I and Type-II errors:

$$\alpha(\mathcal{A}_n, M_n) = \beta(\mathcal{A}_n, I - M_n), \quad (99)$$

$$\beta(\mathcal{B}_n, M_n) = \alpha(\mathcal{B}_n, I - M_n). \quad (100)$$

This allows us to rewrite the optimization in Eq. (97) as

$$\begin{aligned} & \alpha_{n,r}(\mathcal{A}_n\|\mathcal{B}_n) \\ & = \min_{\substack{0 \leq M_n \leq I \\ \alpha(\mathcal{B}_n, I - M_n) \leq 2^{-nr}}} \beta(\mathcal{A}_n, I - M_n) \end{aligned} \quad (101)$$

$$= \min_{\substack{0 \leq M_n \leq I \\ \alpha(\mathcal{B}_n, M_n) \leq 2^{-nr}}} \beta(\mathcal{A}_n, M_n) \quad (102)$$

$$= \sup_{\substack{\rho_n \in \mathcal{A}_n \\ \sigma_n \in \mathcal{B}_n}} \min_{\substack{0 \leq M_n \leq I \\ \alpha(\sigma_n, M_n) \leq 2^{-nr}}} \beta(\rho_n, M_n) \quad (103)$$

$$= \sup_{\substack{\rho_n \in \mathcal{A}_n \\ \sigma_n \in \mathcal{B}_n}} \min_{\substack{0 \leq M_n \leq I \\ \alpha(\sigma_n, I - M_n) \leq 2^{-nr}}} \beta(\rho_n, I - M_n) \quad (104)$$

$$= \sup_{\substack{\rho_n \in \mathcal{A}_n \\ \sigma_n \in \mathcal{B}_n}} \min_{\substack{0 \leq M_n \leq I \\ \beta(\sigma_n, M_n) \leq 2^{-nr}}} \alpha(\rho_n, M_n) \quad (105)$$

$$= \sup_{\substack{\rho_n \in \mathcal{A}_n \\ \sigma_n \in \mathcal{B}_n}} \alpha_{n,r}(\rho_n\|\sigma_n), \quad (106)$$

where the first and fifth equalities use (100), the second and fourth equalities follow by substituting M_n to $I - M_n$ in the optimization, the third equality uses [15, Lemma 31] which allows us to pull out the optimization over $\mathcal{A}_n, \mathcal{B}_n$ for the optimal Type-II error probability when Type-I error is restricted to a constant threshold, and the last equality follows by definition. $\square$

Remark 9: Note that for general sets $\mathcal{A}_n$ and $\mathcal{B}_n$, i.e., not necessarily convex, we have

$$\alpha(\mathcal{A}_n, M_n) = \alpha(\text{conv}(\mathcal{A}_n), M_n), \quad (107)$$

$$\beta(\mathcal{B}_n, M_n) = \beta(\text{conv}(\mathcal{B}_n), M_n), \quad (108)$$

where $\text{conv}(\mathcal{C})$ denotes the convex hull of set $\mathcal{C}$. Therefore, we have the relation that

$$\alpha_{n,r}(\mathcal{A}_n\|\mathcal{B}_n) = \alpha_{n,r}(\text{conv}(\mathcal{A}_n)\|\text{conv}(\mathcal{B}_n)). \quad (109)$$

V. QUANTUM Hoeffding Bound for Composite Correlated Hypotheses

In this section, we establish the quantum Hoeffding bound for composite correlated hypotheses. Notably, the result holds under *minimal and standard* assumptions, which are satisfied by many frameworks of interest, such as those considered in [15], [27], and [28].

Assumption 1 (*Assumptions for sets of quantum states*): We denote the following assumptions for a sequence of sets of quantum states $\mathcal{C} = \{\mathcal{C}_n\}_{n \in \mathbb{N}}$ where each $\mathcal{C}_n \subseteq \mathcal{D}(\mathcal{H}^{\otimes n})$.

- (C1) Convexity: For any $n \in \mathbb{N}$, the set $\mathcal{C}_n$ is convex.
- (C2) Compactness: For any $n \in \mathbb{N}$, the set $\mathcal{C}_n$ is compact.
- (C3) Stability under tensor product: For any $m, n \in \mathbb{N}$, it holds that $\mathcal{C}_m \otimes \mathcal{C}_n \subseteq \mathcal{C}_{m+n}$.

For two sequences of sets of quantum states $\mathcal{A} = \{\mathcal{A}_n\}_{n \in \mathbb{N}}$ and $\mathcal{B} = \{\mathcal{B}_n\}_{n \in \mathbb{N}}$ where each $\mathcal{A}_n, \mathcal{B}_n \subseteq \mathcal{D}(\mathcal{H}^{\otimes n})$, we denote the additional assumption as follows:

- (C4) Finiteness: $D_{p,\alpha}(\mathcal{A}_1\|\mathcal{B}_1) < \infty$ for $\alpha \in (0, 1)$.⁴

Theorem 1 (*Quantum Hoeffding bound for composite correlated hypotheses*): Let $\mathcal{H}$ be a finite-dimensional Hilbert space. Let $\mathcal{A} = \{\mathcal{A}_n\}_{n \in \mathbb{N}}$ and $\mathcal{B} = \{\mathcal{B}_n\}_{n \in \mathbb{N}}$ be two sequences of sets of quantum states, where each $\mathcal{A}_n, \mathcal{B}_n \subseteq \mathcal{D}(\mathcal{H}^{\otimes n})$. Let $0 < r < D^\infty(\mathcal{A}\|\mathcal{B})$ be a real number.

- If $\mathcal{A}$ and $\mathcal{B}$ satisfy assumptions (C1) and (C2):

$$\liminf_{n \rightarrow \infty} -\frac{1}{n} \log \alpha_{n,r}(\mathcal{A}_n\|\mathcal{B}_n) \geq \underline{H}_r^\infty(\mathcal{A}\|\mathcal{B}). \quad (110)$$

- If $\mathcal{A}$ and $\mathcal{B}$ satisfy assumptions (C3) and (C4):

$$\limsup_{n \rightarrow \infty} -\frac{1}{n} \log \alpha_{n,r}(\mathcal{A}_n\|\mathcal{B}_n) \leq H_r^\infty(\mathcal{A}\|\mathcal{B}). \quad (111)$$

⁴This is a mild technical assumption, requiring that there exist $\rho \in \mathcal{A}_1$ and $\sigma \in \mathcal{B}_1$ which are not orthogonal.

Consequently, if $\mathcal{A}$ and $\mathcal{B}$ satisfy (C1), (C2), (C3) and (C4), then the following limit exists and

$$\lim_{n \rightarrow \infty} -\frac{1}{n} \log \alpha_{n,r}(\mathcal{A}_n \| \mathcal{B}_n) = H_r^\infty(\mathcal{A} \| \mathcal{B}). \quad (112)$$

Remark 10: By Remark 9, we can remove the convexity on $\mathcal{A}_n$ and $\mathcal{B}_n$ in Theorem 1 and get

$$\lim_{n \rightarrow \infty} -\frac{1}{n} \log \alpha_{n,r}(\mathcal{A}_n \| \mathcal{B}_n) = H_r^\infty(\text{conv}(\mathcal{A}) \| \text{conv}(\mathcal{B})), \quad (113)$$

where $\text{conv}(\mathcal{C}) := \{\text{conv}(\mathcal{C}_n)\}_{n \in \mathbb{N}}$ represents the sequences of the convex hulls. This result strengthens [20, Eq. (II.65)] by establishing the tightness of the regularized Hoeffding bound. Another related analysis of the Hoeffding bound for composite correlated hypotheses was undertaken in [16, Theorem 2] for the specific case where $\mathcal{A}_n$ is a singleton and $\mathcal{B}_n$ is the set of separable states. However, the lower bound provided therein is not tight in general.

A. Proof of the Lower Bound

Recall that for any $V, W \in \mathcal{H}_+$ and $\alpha \in (0, 1)$, it holds that [37],

$$\text{Tr}[V^\alpha W^{1-\alpha}] \geq \text{Tr}W\{W \leq V\} + \text{Tr}V\{W > V\}. \quad (114)$$

Let $\rho_n \in \mathcal{A}_n$ and $\sigma_n \in \mathcal{B}_n$. Applying the inequality with the choice $V = \rho_n$ and $W = 2^{nR}\sigma_n$ with an arbitrary real number R . Then we have

$$\begin{aligned} \text{Tr}2^{nR}\sigma_n\{2^{nR}\sigma_n \leq \rho_n\} + \text{Tr}\rho_n\{2^{nR}\sigma_n > \rho_n\} \\ \leq 2^{n(1-\alpha)R} Q_\alpha(\rho_n \| \sigma_n). \end{aligned} \quad (115)$$

This implies that

$$\text{Tr}\rho_n\{2^{nR}\sigma_n > \rho_n\} \leq 2^{n(1-\alpha)R} Q_\alpha(\rho_n \| \sigma_n), \quad (116)$$

$$\text{Tr}\sigma_n\{2^{nR}\sigma_n \leq \rho_n\} \leq 2^{-n\alpha R} Q_\alpha(\rho_n \| \sigma_n). \quad (117)$$

Now, letting the constant

$$R = \frac{nr + \log Q_\alpha(\rho_n \| \sigma_n)}{n\alpha}. \quad (118)$$

we get

$$\text{Tr}\rho_n\{2^{nR}\sigma_n > \rho_n\} \leq 2^{\frac{1-\alpha}{\alpha} n(r - \frac{1}{n} D_{p,\alpha}(\rho_n \| \sigma_n))}, \quad (119)$$

$$\text{Tr}\sigma_n\{2^{nR}\sigma_n \leq \rho_n\} \leq 2^{-nr}. \quad (120)$$

Let $M_n = \{2^{nR}\sigma_n \leq \rho_n\}$, which is a valid measurement operator. Then Eq. (120) implies that it is a feasible solution to the optimization problem in Eq. (6). Therefore, we have

$$\alpha_{n,r}(\rho_n \| \sigma_n) \leq 2^{\frac{1-\alpha}{\alpha} n(r - \frac{1}{n} D_{p,\alpha}(\rho_n \| \sigma_n))}. \quad (121)$$

This gives

$$-\frac{1}{n} \log \alpha_{n,r}(\rho_n \| \sigma_n) \geq \frac{\alpha - 1}{\alpha} \left(r - \frac{1}{n} D_{p,\alpha}(\rho_n \| \sigma_n) \right). \quad (122)$$

Taking infimum over $\rho_n \in \mathcal{A}_n$ and $\sigma_n \in \mathcal{B}_n$ on both sides, we have

$$-\frac{1}{n} \log \alpha_{n,r}(\mathcal{A}_n \| \mathcal{B}_n) \geq \frac{\alpha - 1}{\alpha} \left(r - \frac{1}{n} D_{p,\alpha}(\mathcal{A}_n \| \mathcal{B}_n) \right), \quad (123)$$

where we use Lemma 7 and the assumption (C1). Taking supremum over $\alpha \in (0, 1)$, we have

$$-\frac{1}{n} \log \alpha_{n,r}(\mathcal{A}_n \| \mathcal{B}_n) \geq \frac{1}{n} \mathfrak{H}_{n,r}(\mathcal{A}_n \| \mathcal{B}_n) \quad (124)$$

$$= \frac{1}{n} H_{n,r}(\mathcal{A}_n \| \mathcal{B}_n), \quad (125)$$

where the equality follows from Lemma 2 and the assumptions (C1) and (C2).

Taking limit of n , we have

$$\begin{aligned} \liminf_{n \rightarrow \infty} -\frac{1}{n} \log \alpha_{n,r}(\mathcal{A}_n \| \mathcal{B}_n) \\ \geq \liminf_{n \rightarrow \infty} \frac{1}{n} H_{n,r}(\mathcal{A}_n \| \mathcal{B}_n) = \underline{H}_r^\infty(\mathcal{A} \| \mathcal{B}). \end{aligned} \quad (126)$$

B. Proof of the Upper Bound

We can easily prove the upper bound for limit inferior as follows. For any fixed $m \in \mathbb{N}$ any $\rho_m \in \mathcal{A}_m, \sigma_m \in \mathcal{B}_m$, then

$$\begin{aligned} \liminf_{n \rightarrow \infty} -\frac{1}{n} \log \alpha_{n,r}(\mathcal{A}_n \| \mathcal{B}_n) \\ \leq \liminf_{n \rightarrow \infty} -\frac{1}{mn} \log \alpha_{mn,r}(\mathcal{A}_{mn} \| \mathcal{B}_{mn}) \end{aligned} \quad (127)$$

$$\leq \liminf_{n \rightarrow \infty} -\frac{1}{mn} \log \sup_{\substack{\rho_{mn} \in \mathcal{A}_{mn} \\ \sigma_{mn} \in \mathcal{B}_{mn}}} \alpha_{mn,r}(\rho_{mn} \| \sigma_{mn}) \quad (128)$$

$$\leq \liminf_{n \rightarrow \infty} -\frac{1}{mn} \log \alpha_{mn,r}(\rho_m^{\otimes n} \| \sigma_m^{\otimes n}) \quad (129)$$

$$= \frac{1}{m} H_{m,r}(\rho_m \| \sigma_m), \quad (130)$$

where the first inequality follows as the lower limit of a subsequence is no smaller than the lower limit of the sequence, the second inequality holds trivially as $\alpha_{n,r}(\mathcal{A}_n \| \mathcal{B}_n) \geq \alpha_{n,r}(\rho_n \| \sigma_n)$ for any $\rho_n \in \mathcal{A}_n, \sigma_n \in \mathcal{B}_n$ by definition, the third inequality follows by taking a particular feasible solution and the assumption (C3), the equality follows from the quantum Hoeffding bound between two quantum states (see Eq. (7)). As this holds for any $\rho_m \in \mathcal{A}_m, \sigma_m \in \mathcal{B}_m$,

$$\liminf_{n \rightarrow \infty} -\frac{1}{n} \log \alpha_{n,r}(\mathcal{A}_n \| \mathcal{B}_n) \leq \frac{1}{m} H_{m,r}(\mathcal{A}_m \| \mathcal{B}_m). \quad (131)$$

Taking limit of m , we get

$$\begin{aligned} \liminf_{n \rightarrow \infty} -\frac{1}{n} \log \alpha_{n,r}(\mathcal{A}_n \| \mathcal{B}_n) \\ \leq \liminf_{m \rightarrow \infty} \frac{1}{m} H_{m,r}(\mathcal{A}_m \| \mathcal{B}_m) = H_r^\infty(\mathcal{A} \| \mathcal{B}), \end{aligned} \quad (132)$$

where the equality follows from the stability assumption (C3) and Eq. (45).

The above argument only gives the upper bound for limit inferior by choosing a suitable subsequence of i.i.d. states. However, we show that the upper bound can be strengthened further to limit superior by carefully designing a sequence of states whose limit superior is also upper bounded by the regularized Hoeffding divergence. However, as this sequence is not i.i.d. states anymore, its analysis is more challenging and requires the Nussbaum-Szkoła distributions [38] and the Gärtner-Ellis theorem [39].

Let the spectral decompositions of ρ and σ be given by

$$\rho = \sum_{i=1}^d \lambda_i |u_i\rangle\langle u_i| \quad \text{and} \quad \sigma = \sum_{j=1}^d \mu_j |v_j\rangle\langle v_j|, \quad (133)$$

where $|u_i\rangle$ and $|v_j\rangle$ are two orthonormal bases and λ_i and μ_j are the corresponding eigenvalues, respectively. Then the Nussbaum-Szkoła distributions of ρ, σ are defined by

$$(P_{\rho, \sigma})(i, j) = \lambda_i |\langle u_i | v_j \rangle|^2, \quad (134)$$

$$(Q_{\rho, \sigma})(i, j) = \mu_j |\langle u_i | v_j \rangle|^2, \quad (135)$$

where $i, j \in \{1, \dots, d\}$.

In the remaining discussion of this section, let $\log$ be a logarithm with natural base e for simplicity. Given a sequence of random variables $\{X_n\}_{n \in \mathbb{N}}$, the asymptotic cumulant generating function is defined by

$$\Lambda_X(t) := \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E} [\exp(ntX_n)], \quad (136)$$

provided that the limit exists.

We will use the Gärtner–Ellis theorem [39, Theorem 2.3.6] in the following slightly tailored form, which suffices for our later arguments. Compared with the version employed in [21, Proposition 17], we state the bound for the open set (x, ∞) . Since $\Pr\{X_n > x\} \leq \Pr\{X_n \geq x\}$, the same upper bound also applies to the closed set $[x, \infty)$.

Lemma 8: Assume that the asymptotic cumulant generating function $t \mapsto \Lambda_X(t)$ exists for all $t \in \mathbb{R}$, and that Λ_X is strictly convex and continuously differentiable on $\mathbb{R}$ (i.e., $\Lambda_X \in C^1(\mathbb{R})$). Fix an open interval $(a, b) \subseteq \mathbb{R}$. Then, for any $x \in (\Lambda'_X(a), \Lambda'_X(b))$,

$$\limsup_{n \rightarrow \infty} -\frac{1}{n} \log \Pr\{X_n > x\} \leq \sup_{t \in (a, b)} \{tx - \Lambda_X(t)\}. \quad (137)$$

Proof: This is a direct consequence of the Gärtner–Ellis theorem (see, e.g., [39, Theorem 2.3.6]), with minor adaptations. By assumption, Λ_X exists on $\mathbb{R}$ and is $C^1(\mathbb{R})$ and strictly convex; in particular, it is lower semicontinuous and essentially smooth (see, e.g., [40, Page 251]). Hence, by [39, Theorem 2.3.6(c)], the sequence $\{X_n\}$ satisfies the large deviation principle with good rate function

$$\Lambda_X^*(w) := \sup_{t \in \mathbb{R}} \{tw - \Lambda_X(t)\}. \quad (138)$$

Therefore, for the open set (x, ∞) ,

$$\limsup_{n \rightarrow \infty} -\frac{1}{n} \log \Pr\{X_n > x\} \leq \inf_{w > x} \Lambda_X^*(w). \quad (139)$$

Since Λ_X is strictly convex and differentiable, the supremum defining $\Lambda_X^*(w)$ is attained at the unique maximizer t_w satisfying the first-order condition $\Lambda'_X(t_w) = w$, and

$$\Lambda_X^*(w) = t_w w - \Lambda_X(t_w). \quad (140)$$

Moreover, Λ'_X is strictly increasing and continuous, so $w \mapsto t_w = (\Lambda'_X)^{-1}(w)$ is continuous, and consequently Λ_X^* is continuous on the corresponding interval of w 's. Thus,

$$\inf_{w > x} \Lambda_X^*(w) \leq \lim_{k \rightarrow \infty} \Lambda_X^*\left(x + \frac{1}{k}\right) = \Lambda_X^*(x), \quad (141)$$

where the inequality follows as $x + 1/k > x$ and the equality follows from the continuity of Λ_X^* .

Finally, the assumption $x \in (\Lambda'_X(a), \Lambda'_X(b))$ ensures that the unique maximizer t_x satisfying $\Lambda'_X(t_x) = x$ lies in (a, b) . Hence the supremum defining $\Lambda_X^*(x)$ can be restricted to the interval (a, b) :

$$\Lambda_X^*(x) = \sup_{t \in (a, b)} \{tx - \Lambda_X(t)\}. \quad (142)$$

Combining the above relations yields the desired bound. $\square$

Lemma 9: Let $m \in \mathbb{N}$ be any integer. Let $\rho_1 \in \mathcal{A}_1$, $\sigma_1 \in \mathcal{B}_1$ and $\rho_m \in \mathcal{A}_m$, $\sigma_m \in \mathcal{B}_m$ be quantum states such that $D_{P, \alpha}(\rho_1 \| \sigma_1) < \infty$ and $D_{P, \alpha}(\rho_m \| \sigma_m) < \infty$ for any $\alpha \in (0, 1)$. We set $k := \lfloor n/m \rfloor$ and construct quantum states

$$\rho^{(n)} := \rho_1^{\otimes n-km} \otimes \rho_m^{\otimes k}, \quad (143)$$

$$\sigma^{(n)} := \sigma_1^{\otimes n-km} \otimes \sigma_m^{\otimes k}. \quad (144)$$

Let $\phi(s) := \frac{1}{m} \log \text{Tr} \rho_m^{1-s} \sigma_m^s$. If

$$R \in \left(-\frac{1}{m} D(\sigma_m \| \rho_m), \frac{1}{m} D(\rho_m \| \sigma_m) \right), \quad (145)$$

then for any $0 \leq T_n \leq I$,

$$\begin{aligned} \limsup_{n \rightarrow \infty} -\frac{1}{n} \log [\text{Tr} e^{-nR} \rho^{(n)} (I - T_n) + \text{Tr} \sigma^{(n)} T_n] \\ \leq \max_{s \in (0, 1)} (1-s)R - \phi(s). \end{aligned} \quad (146)$$

Proof: Let $P^{(n)}$ and $Q^{(n)}$ be the Nussbaum-Szkoła distributions of $\rho^{(n)}$ and $\sigma^{(n)}$. Let

$$S_n = \{e^{-nR} P^{(n)} > Q^{(n)}\}, \quad (147)$$

be a likelihood ratio test. Consider the random variable

$$X_n(x) := \frac{1}{n} (\log Q^{(n)}(x) - \log P^{(n)}(x)), \quad (148)$$

where x is drawn from the distribution $P^{(n)}$. Then we have the asymptotic cumulant generating function of the random variable X_n as,

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} \log \sum_x P^{(n)}(x) \exp(snX_n(x)) \\ = \lim_{n \rightarrow \infty} \frac{1}{n} \log \text{Tr}(Q^{(n)})^s (P^{(n)})^{1-s} \end{aligned} \quad (149)$$

$$= \lim_{n \rightarrow \infty} \frac{1}{n} \log \text{Tr}(\sigma^{(n)})^s (\rho^{(n)})^{1-s} \quad (150)$$

$$= \lim_{n \rightarrow \infty} \frac{1}{n} ((n-km) \log \text{Tr} \rho_1^{1-s} \sigma_1^s + k \log \text{Tr} \rho_m^{1-s} \sigma_m^s) \quad (151)$$

$$= \phi(s), \quad (152)$$

where the second equality is a simple fact of the Nussbaum-Szkoła distributions (c.f. [7, Proposition 1]). Note that $\phi(s)$ is C^1 continuous and strictly convex, with $\phi'(0) = -\frac{1}{m} D(\rho_m \| \sigma_m)$ and $\phi'(1) = \frac{1}{m} D(\sigma_m \| \rho_m)$ (see Appendix A). Applying the Gärtner–Ellis theorem in Lemma 8 for the random variable X_n , interval $(0, 1)$ and $x = -R$, we have

$$\limsup_{n \rightarrow \infty} -\frac{1}{n} \log \Pr\{X_n \geq -R\} \leq \sup_{s \in (0, 1)} -sR - \phi(s). \quad (153)$$

Similarly, consider the random variable

$$Y_n(x) := \frac{1}{n} (\log P^{(n)}(x) - \log Q^{(n)}(x)), \quad (154)$$

where x is drawn from the distribution $Q^{(n)}$. Then we have the asymptotic cumulant generating function of the random variable Y_n as,

$$\begin{aligned} & \lim_{n \rightarrow \infty} \frac{1}{n} \log \sum_x Q^{(n)}(x) \exp(tnY_n(x)) \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \log \text{Tr}(Q^{(n)})^{1-t} (P^{(n)})^t \end{aligned} \quad (155)$$

$$= \lim_{n \rightarrow \infty} \frac{1}{n} \log \text{Tr}(\sigma^{(n)})^{1-t} (\rho^{(n)})^t \quad (156)$$

$$= \lim_{n \rightarrow \infty} \frac{1}{n} ((n - km) \log \text{Tr} \rho_1^t \sigma_1^{1-t} + k \log \text{Tr} \rho_m^t \sigma_m^{1-t}) \quad (157)$$

$$= \phi(1 - t). \quad (158)$$

Applying again the Gärtner-Ellis theorem in Lemma 8 for the random variable Y_n , interval $(0, 1)$ and $x = R$, we have

$$\limsup_{n \rightarrow \infty} -\frac{1}{n} \log \Pr\{Y_n > R\} \leq \sup_{t \in (0,1)} tR - \phi(1 - t). \quad (159)$$

By direct calculation, we have the relations

$$\begin{aligned} & \limsup_{n \rightarrow \infty} -\frac{1}{n} \log \Pr\{X_n \geq -R\} \\ &= \limsup_{n \rightarrow \infty} -\frac{1}{n} \log \text{Tr} P^{(n)}(I - S_n), \end{aligned} \quad (160)$$

$$\begin{aligned} & \limsup_{n \rightarrow \infty} -\frac{1}{n} \log \Pr\{Y_n > R\} \\ &= \limsup_{n \rightarrow \infty} -\frac{1}{n} \log \text{Tr} Q^{(n)} S_n. \end{aligned} \quad (161)$$

Moreover, the Nussbaum-Szkoła theorem (c.f. [41, Lemma 3.4]) implies that for any test T_n ,

$$\begin{aligned} & \text{Tre}^{-nR} \rho^{(n)}(I - T_n) + \text{Tr} \sigma^{(n)} T_n \\ & \geq \frac{1}{2} (\text{Tre}^{-nR} P^{(n)}(I - S_n) + \text{Tr} Q^{(n)} S_n). \end{aligned} \quad (162)$$

Combining Eqs. (153), (159), (160), (161) and (162), we have

$$\begin{aligned} & \limsup_{n \rightarrow \infty} -\frac{1}{n} \log [\text{Tre}^{-nR} \rho^{(n)}(I - T_n) + \text{Tr} \sigma^{(n)} T_n] \\ & \leq \limsup_{n \rightarrow \infty} -\frac{1}{n} \log \frac{1}{2} [\text{Tre}^{-nR} P^{(n)}(I - S_n) + \text{Tr} Q^{(n)} S_n] \end{aligned} \quad (163)$$

$$\leq \min \left\{ R + \sup_{s \in (0,1)} -sR - \phi(s), \sup_{t \in (0,1)} tR - \phi(1 - t) \right\} \quad (164)$$

$$= \sup_{s \in (0,1)} (1 - s)R - \phi(s), \quad (165)$$

where the last equality follows by replacing t with $1 - s$. Finally, as the objective function in Eq. (165) is concave and its first derivative is given by $-R - \phi'(s)$, there is a unique critical point that achieves the maximum if $R \in (-\phi'(1), -\phi'(0))$. Therefore, the supremum is attained. $\square$

With the above Lemma 9, we are ready to show the upper bound for limit superior. That is, we aim to show that

$$\limsup_{n \rightarrow \infty} -\frac{1}{n} \log \alpha_{n,r}(\mathcal{A}_n \| \mathcal{B}_n) \leq H_r^\infty(\mathcal{A} \| \mathcal{B}). \quad (166)$$

For this, we plan to show for any fixed $m \in \mathbb{N}$,

$$\limsup_{n \rightarrow \infty} -\frac{1}{n} \log \alpha_{n,r}(\mathcal{A}_n \| \mathcal{B}_n) \leq \frac{1}{m} H_{m,r}(\mathcal{A}_m \| \mathcal{B}_m). \quad (167)$$

If $H_{m,r}(\mathcal{A}_m \| \mathcal{B}_m) = \infty$, the upper bound holds trivially. It remains to show Eq. (167) when $H_{m,r}(\mathcal{A}_m \| \mathcal{B}_m) < \infty$. So for any $\delta > 0$, there exist $\rho_m \in \mathcal{A}_m$ and $\sigma_m \in \mathcal{B}_m$ such that

$$H_{m,r}(\rho_m \| \sigma_m) \leq H_{m,r}(\mathcal{A}_m \| \mathcal{B}_m) + \delta < \infty. \quad (168)$$

This implies that $D_{p,\alpha}(\rho_m \| \sigma_m) < \infty$ for any $\alpha \in (0, 1)$. Otherwise, we will have a contradiction to the finiteness of $H_{m,r}(\rho_m \| \sigma_m)$ by definition in Eq. (29). Using this choice of ρ_m and σ_m , we construct the sequence of states $\rho^{(n)}$ and $\sigma^{(n)}$ as in Lemma 9.

Combining Lemma 9 and Lemma 14, we have for any test with $\text{Tr} \sigma^{(n)} T_n \leq e^{-nr}$, that

$$\begin{aligned} & \limsup_{n \rightarrow \infty} -\frac{1}{n} \log [\text{Tre}^{-nR} \rho^{(n)}(I - T_n) + \text{Tr} \sigma^{(n)} T_n] \\ & \geq \limsup_{n \rightarrow \infty} -\frac{1}{n} \log [\text{Tre}^{-nR} \rho^{(n)}(I - T_n) + e^{-nr}] \end{aligned} \quad (169)$$

$$= \min \left\{ r, R + \limsup_{n \rightarrow \infty} -\frac{1}{n} \log \text{Tr} \rho^{(n)}(I - T_n) \right\}, \quad (170)$$

where the equality follows from Lemma 14. Therefore, by Lemma 9, we have

$$\begin{aligned} & \min \left\{ r, R + \limsup_{n \rightarrow \infty} -\frac{1}{n} \log \text{Tr} \rho^{(n)}(I - T_n) \right\} \\ & \leq \max_{s \in (0,1)} (1 - s)R - \phi(s). \end{aligned} \quad (171)$$

Let $r' < r$. From Lemma 16, we know that the optimization

$$\max_{s \in (0,1)} \frac{-\phi(s) - sr'}{1 - s} \quad (172)$$

has a unique maximizer $s_{r'} \in (0, 1)$ such that $r' = (s_{r'} - 1)\phi'(s_{r'}) - \phi(s_{r'})$. We set

$$R_{r'} := \frac{\phi(s_{r'}) + r'}{1 - s_{r'}} = -\phi'(s_{r'}). \quad (173)$$

It is clear that

$$R_{r'} \in (-\phi'(1), -\phi'(0)) \quad (174)$$

$$= \left(-\frac{1}{m} D(\sigma_m \| \rho_m), \frac{1}{m} D(\rho_m \| \sigma_m) \right). \quad (175)$$

By the proof of Lemma 9, $\max_{s \in (0,1)} (1 - s)R - \phi(s)$ is uniquely achieved at point s such that $R = -\phi'(s)$. For $R = R_{r'}$, Eq. (173) implies that the maximum is uniquely achieved at $s = s_{r'}$. That is,

$$\max_{s \in (0,1)} (1 - s)R_{r'} - \phi(s) = (1 - s_{r'})R_{r'} - \phi(s_{r'}) = r'. \quad (176)$$

Then, by Eq. (171) we have

$$\min \left\{ r, R_{r'} + \limsup_{n \rightarrow \infty} -\frac{1}{n} \log \text{Tr} \rho^{(n)}(I - T_n) \right\} \leq r'. \quad (177)$$

Then, we have

$$R_{r'} + \limsup_{n \rightarrow \infty} -\frac{1}{n} \log \text{Tr} \rho^{(n)}(I - T_n) \leq r. \quad (178)$$

Otherwise, it will contradict to the assumption that $r' < r$. Thus, we have

$$\limsup_{n \rightarrow \infty} -\frac{1}{n} \log \text{Tr} \rho^{(n)}(I - T_n) \leq r - R_{r'}. \quad (179)$$

Since s_r is continuous function in r (see Lemma 17) and $\phi(s)$ is a continuous function in s (see Lemma 15), $R_{r'}$ is a continuous function in r' . Therefore,

$$\lim_{r' \rightarrow r^-} R_{r'} = R_r := \frac{\phi(s_r) + r}{1 - s_r}. \quad (180)$$

This gives

$$\limsup_{n \rightarrow \infty} -\frac{1}{n} \log \text{Tr} \rho^{(n)}(I - T_n) \leq r - R_r. \quad (181)$$

By direct calculation, the right-hand side gives

$$r - R_r = \frac{-\phi(s_r) - s_r r}{1 - s_r} \quad (182)$$

$$= \max_{s \in (0,1)} \frac{-\phi(s) - s r}{1 - s} \quad (183)$$

$$= \frac{1}{m} H_{m,r}(\rho_m \| \sigma_m), \quad (184)$$

where the second equality follows from the optimality of s_r . By the choice of ρ_m, σ_m in Eq. (168), we have

$$\begin{aligned} \limsup_{n \rightarrow \infty} -\frac{1}{n} \log \text{Tr} \rho^{(n)}(I - T_n) \\ \leq \frac{1}{m} H_{m,r}(\mathcal{A}_m \| \mathcal{B}_m) + \frac{\delta}{m}. \end{aligned} \quad (185)$$

This implies that

$$\begin{aligned} \limsup_{n \rightarrow \infty} -\frac{1}{n} \log \alpha_{n,r}(\rho^{(n)} \| \sigma^{(n)}) \\ \leq \frac{1}{m} H_{m,r}(\mathcal{A}_m \| \mathcal{B}_m) + \frac{\delta}{m}. \end{aligned} \quad (186)$$

As we have the trivial relation that $\alpha_{n,r}(\mathcal{A}_n \| \mathcal{B}_n) \geq \alpha_{n,r}(\rho_n \| \sigma_n)$ for any $\rho_n \in \mathcal{A}_n, \sigma_n \in \mathcal{B}_n$ by definition, this implies that

$$\begin{aligned} \limsup_{n \rightarrow \infty} -\frac{1}{n} \log \alpha_{n,r}(\mathcal{A}_n \| \mathcal{B}_n) \\ \leq \frac{1}{m} H_{m,r}(\mathcal{A}_m \| \mathcal{B}_m) + \frac{\delta}{m}. \end{aligned} \quad (187)$$

As this holds for any $\delta > 0$, we have

$$\limsup_{n \rightarrow \infty} -\frac{1}{n} \log \alpha_{n,r}(\mathcal{A}_n \| \mathcal{B}_n) \leq \frac{1}{m} H_{m,r}(\mathcal{A}_m \| \mathcal{B}_m). \quad (188)$$

Finally, taking the limit $m \rightarrow \infty$, we have

$$\limsup_{n \rightarrow \infty} -\frac{1}{n} \log \alpha_{n,r}(\mathcal{A}_n \| \mathcal{B}_n) \leq H_r^{\infty}(\mathcal{A} \| \mathcal{B}), \quad (189)$$

where the existence of the limit on the right hand side follows from the stability assumption (C3) and Eq. (45). This finishes the proof of the upper bound.

VI. STRONG CONVERSE EXPONENT FOR COMPOSITE CORRELATED HYPOTHESES

In this section, we extend the strong converse exponent from the i.i.d. setting to the broader context of composite correlated hypotheses and show that the strong converse exponent is lower bounded by a regularized quantum Hoeffding anti-divergence between the sets in general. We also provide a matching upper bound when the null hypothesis is a singleton i.i.d. state under additional assumptions.

For the upper bound, we need to introduce the permutation invariance. Let U_{A^n} denote the natural unitary representation of the permutation group S_n that permutes the subsystems of $A^n := A_1 A_2 \cdots A_n$. An linear operator $X_{A^n} \in \mathcal{L}(A^n)$ is called permutation invariant if it satisfies $X_{A^n} = U_{A^n}(\pi) X_{A^n} U_{A^n}(\pi)^\dagger$ for all $\pi \in S_n$. A set of quantum state $\mathcal{A}_n$ is called permutation invariant if $\rho_n \in \mathcal{A}_n$ implies that $U_{A^n}(\pi) \rho_n U_{A^n}(\pi)^\dagger \in \mathcal{A}_n$ for all $\pi \in S_n$. Let A be a system with $|A| = d$. For all $n \in \mathbb{N}$ there exists a permutation invariant state $\omega_{A^n}^n \in \mathcal{D}(A^n)$, which we call *universal state*, such that for all permutation invariant states $\rho_{A^n} \in \mathcal{D}(A^n)$, we have

$$\rho_{A^n} \leq g_{n,d} \omega_{A^n}^n, \quad \text{with} \quad g_{n,d} \leq (n+1)^{d^2-1}. \quad (190)$$

Note that such a universal state exists and has been explicitly constructed in [21].

Assumption 2 (Assumptions for sets of quantum states): We denote the following assumptions for a sequence of sets of quantum states $\mathcal{C} = \{\mathcal{C}_n\}_{n \in \mathbb{N}}$ where each $\mathcal{C}_n \subseteq \mathcal{D}(\mathcal{H}^{\otimes n})$.

- (C5) Permutation invariance: For any $n \in \mathbb{N}$, the set $\mathcal{C}_n$ is permutation invariant.
- (C6) Universal state containment: For any $n \in \mathbb{N}$, the set $\mathcal{C}_n$ contains a universal state.

Theorem 2 (Strong converse exponent for composite correlated hypotheses): Let $\mathcal{H}$ be a finite-dimensional Hilbert space. Let $\mathcal{A} = \{\mathcal{A}_n\}_{n \in \mathbb{N}}$ and $\mathcal{B} = \{\mathcal{B}_n\}_{n \in \mathbb{N}}$ be two sequences of sets of quantum states, where each $\mathcal{A}_n, \mathcal{B}_n \subseteq \mathcal{D}(\mathcal{H}^{\otimes n})$. Let $r > D^\infty(\mathcal{A} \| \mathcal{B})$ be a real number.

- Without any additional assumptions on $\mathcal{A}$ or $\mathcal{B}$,

$$\liminf_{n \rightarrow \infty} -\frac{1}{n} \log(1 - \alpha_{n,r}(\mathcal{A}_n \| \mathcal{B}_n)) \geq \underline{H}_r^{*,\infty}(\mathcal{A} \| \mathcal{B}). \quad (191)$$

- Suppose that $\mathcal{A}_n = \{\rho^{\otimes n}\}$ is a singleton i.i.d. sequence, and that $\mathcal{B}$ satisfies assumptions (C1), (C5), and (C6). Assume further that $D_{s,\alpha}^{\infty}(\mathcal{A} \| \mathcal{B})$ exists and is differentiable in α for all $\alpha \geq 1$. Then, for $D^\infty(\mathcal{A} \| \mathcal{B}) < r < \sup_{\alpha > 1} D_{s,\alpha}^{\infty}(\mathcal{A} \| \mathcal{B})$,

$$\limsup_{n \rightarrow \infty} -\frac{1}{n} \log(1 - \alpha_{n,r}(\mathcal{A}_n \| \mathcal{B}_n)) \leq \mathfrak{H}_r^{*,\infty}(\mathcal{A} \| \mathcal{B}). \quad (192)$$

Consequently, with the same assumptions for the upper bound and additionally assuming that $\mathcal{B}$ satisfies assumption (C3), then the limit exists and is given by

$$\lim_{n \rightarrow \infty} -\frac{1}{n} \log(1 - \alpha_{n,r}(\mathcal{A}_n \| \mathcal{B}_n)) = H_r^{*,\infty}(\mathcal{A} \| \mathcal{B}). \quad (193)$$

Remark 11: This result applies to a few interesting cases, such as the resource theory of coherence and entanglement, as discuss in Section VII-D. Moreover, by Remark 9, the convexity of $\mathcal{A}_n$ and $\mathcal{B}_n$ can be removed by replacing the right hand side with $H_r^{*,\infty}(\text{conv}(\mathcal{A}) \| \text{conv}(\mathcal{B}))$, where $\text{conv}(\mathcal{C}) := \{\text{conv}(\mathcal{C}_n)\}_{n \in \mathbb{N}}$ denotes the set of convex hulls.

A. Proof of the Lower Bound

For any $\delta > 0$, let $\rho_n \in \mathcal{A}_n, \sigma_n \in \mathcal{B}_n$ such that

$$D_{s,\alpha}(\rho_n \| \sigma_n) \leq D_{s,\alpha}(\mathcal{A}_n \| \mathcal{B}_n) + \delta. \quad (194)$$

By standard arguments, e.g. [10, Lemma 5], we have for any $0 \leq M_n \leq I$, that

$$\begin{aligned} & \frac{1}{n} \log(1 - \text{Tr}[(I - M_n)\rho_n]) \\ & \leq \frac{\alpha - 1}{\alpha} \left(\frac{1}{n} D_{s,\alpha}(\rho_n \| \sigma_n) + \frac{1}{n} \log \text{Tr}[M_n \sigma_n] \right). \end{aligned} \quad (195)$$

Since $\text{Tr}[(I - M_n)\rho_n] \leq \alpha(\mathcal{A}_n, M_n)$ and $\text{Tr}[M_n \sigma_n] \leq \beta(\mathcal{B}_n, M_n)$, it follows that

$$\begin{aligned} & \frac{1}{n} \log(1 - \alpha(\mathcal{A}_n, M_n)) \\ & \leq \frac{\alpha - 1}{\alpha} \left(\frac{1}{n} D_{s,\alpha}(\mathcal{A}_n \| \mathcal{B}_n) + \delta + \frac{1}{n} \log \beta(\mathcal{B}_n, M_n) \right). \end{aligned} \quad (196)$$

As this holds for any $\delta > 0$, we have

$$\begin{aligned} & \frac{1}{n} \log(1 - \alpha(\mathcal{A}_n, M_n)) \\ & \leq \frac{\alpha - 1}{\alpha} \left(\frac{1}{n} D_{s,\alpha}(\mathcal{A}_n \| \mathcal{B}_n) + \frac{1}{n} \log \beta(\mathcal{B}_n, M_n) \right). \end{aligned} \quad (197)$$

For any $0 \leq M_n \leq I$ such that $\beta(\mathcal{B}_n, M_n) \leq 2^{-nr}$, we obtain

$$\begin{aligned} & \frac{1}{n} \log(1 - \alpha(\mathcal{A}_n, M_n)) \\ & \leq \frac{\alpha - 1}{\alpha} \left(\frac{1}{n} D_{s,\alpha}(\mathcal{A}_n \| \mathcal{B}_n) - r \right). \end{aligned} \quad (198)$$

Taking the supremum over all such M_n , we find

$$\begin{aligned} & -\frac{1}{n} \log(1 - \alpha_{n,r}(\mathcal{A}_n \| \mathcal{B}_n)) \\ & \geq \frac{\alpha - 1}{\alpha} \left(r - \frac{1}{n} D_{s,\alpha}(\mathcal{A}_n \| \mathcal{B}_n) \right). \end{aligned} \quad (199)$$

Since this holds for any $\alpha > 1$, we have

$$-\frac{1}{n} \log(1 - \alpha_{n,r}(\mathcal{A}_n \| \mathcal{B}_n)) \geq H_{n,r}^*(\mathcal{A}_n \| \mathcal{B}_n), \quad (200)$$

where we use the finite equivalence in Lemma 5. Taking limit inferior of n on both sides,

$$\liminf_{n \rightarrow \infty} -\frac{1}{n} \log(1 - \alpha_{n,r}(\mathcal{A}_n \| \mathcal{B}_n)) \geq \underline{H}_r^{*,\infty}(\mathcal{A} \| \mathcal{B}). \quad (201)$$

B. Proof of the Upper Bound

Before the proof of the upper bound, we require the following lemma that simplifies the optimization over $\mathcal{B}_n$ via the universal state. Let $X = \sum_{\lambda \in \text{spec}(X)} \lambda E_\lambda$ be the spectral decomposition of the Hermitian operator X , where E_λ are projectors and $\text{spec}(X)$ is its discrete spectrum. Then the pinching map for this spectral decomposition is defined by $\mathcal{P}_X(\cdot) := \sum_i E_i(\cdot)E_i$.

Lemma 10: Let $\mathcal{A}_n = \{\rho_n\}$ be a permutation invariant state and $\mathcal{B}_n$ be a convex and permutation invariant set, which contains a universal state ω_n as defined in Eq. (190). Then for any $\alpha \geq 1/2$,

$$\frac{1}{n} D_{s,\alpha}(\mathcal{P}_{\omega_n}(\rho_n) \| \omega_n) = \frac{1}{n} D_{s,\alpha}(\rho_n \| \mathcal{B}_n) + O\left(\frac{\log n}{n}\right). \quad (202)$$

In particular,

$$\lim_{n \rightarrow \infty} \frac{1}{n} D_{s,\alpha}(\mathcal{P}_{\omega_n}(\rho_n) \| \omega_n) = D_{s,\alpha}^{\infty}(\mathcal{A} \| \mathcal{B}). \quad (203)$$

Proof: By the permutation invariance and the convexity of the sandwiched Rényi divergence, we know that $D_{s,\alpha}(\rho_n \| \mathcal{B}_n)$ can be achieved at some permutation invariant state $\sigma_n \in \mathcal{B}_n$. Note also that $\sigma_n \leq g_{n,d} \omega_n$ for any permutation invariant state σ_n . Therefore, we have

$$D_{s,\alpha}(\mathcal{P}_{\omega_n}(\rho_n) \| \omega_n) \leq D_{s,\alpha}(\rho_n \| \omega_n) \quad (204)$$

$$\leq D_{s,\alpha}(\rho_n \| \sigma_n) + \log g_{n,d}, \quad (205)$$

where the first inequality uses the data processing inequality and the second inequality uses the monotonicity of the sandwiched Rényi divergence under scaling of the second argument. As this holds for any permutation invariant state $\sigma_n \in \mathcal{B}_n$, we have

$$D_{s,\alpha}(\mathcal{P}_{\omega_n}(\rho_n) \| \omega_n) \leq D_{s,\alpha}(\rho_n \| \mathcal{B}_n) + \log g_{n,d}. \quad (206)$$

On the other hand, we have [21, Lemma 3],

$$D_{s,\alpha}(\mathcal{P}_{\omega_n}(\rho_n) \| \omega_n) \geq D_{s,\alpha}(\rho_n \| \omega_n) - 2 \log |\text{spec}(\omega_n)|, \quad (207)$$

where $|\text{spec}(\cdot)|$ denotes the number of distinct eigenvalues. Since $\omega_n \in \mathcal{B}_n$, we have

$$D_{s,\alpha}(\mathcal{P}_{\omega_n}(\rho_n) \| \omega_n) \geq D_{s,\alpha}(\rho_n \| \mathcal{B}_n) - 2 \log |\text{spec}(\omega_n)|. \quad (208)$$

As ω_n is permutation invariant, we have $|\text{spec}(\omega_n)| \leq (n+1)^d$. This completes the proof. $\square$

Remark 12: In the above case, $D_{s,\alpha}^{\infty}(\mathcal{A} \| \mathcal{B})$ is a pointwise limit of $\frac{1}{n} D_{s,\alpha}(\mathcal{P}_{\omega_n}(\rho_n) \| \omega_n)$ and the convergence is uniform in α . Therefore, $D_{s,\alpha}^{\infty}(\mathcal{A} \| \mathcal{B})$ is continuous and monotonically increasing in α . Similarly, the function $t D_{s,1+t}^{\infty}(\mathcal{A} \| \mathcal{B})$ is continuous and convex in t .

Now, we are ready to prove the upper bound. The proof follows the similar idea as in [21]. Let $\mathcal{P}_{\omega_n}(\rho_n)$ be the pinching of ρ_n with respect to ω_n . Then $\mathcal{P}_{\omega_n}(\rho_n)$ and ω_n commute. Let $|v_{x_n}\rangle$ be a common orthonormal eigenbasis and we define the classical probability distributions

$$\begin{aligned} P_n(x) &:= \langle v_{x_n} | \mathcal{P}_{\omega_n}(\rho_n) | v_{x_n} \rangle \\ > t, \end{aligned} \quad (209)$$

$$\begin{aligned} Q_n(x) &:= \langle v_{x_n} | \omega_n | v_{x_n} \rangle \\ > t. \end{aligned} \quad (210)$$

Let X_n be a random variable distributed according to P_n and X'_n be a random variable distributed according to Q_n . Define

$$S_n := \{\mathcal{P}_{\omega_n}(\rho_n) \geq \exp(\lambda_n) \omega_n\}. \quad (211)$$

Then we have the relation that

$$\alpha(\rho_n, S_n) = \text{Tr}[\rho_n(I - S_n)] \quad (212)$$

$$= \text{Tr}[\mathcal{P}_{\omega_n}(\rho_n)(I - S_n)] \quad (213)$$

$$= \Pr[P_n(X_n) < \exp(\lambda_n) Q_n(X_n)], \quad (214)$$

where the second equality follows as $\text{Tr} \mathcal{P}_X(Y)X = \text{Tr} YX$ for all Y . Similarly, we have

$$\beta(\mathcal{B}_n, S_n) = \sup_{\sigma_n \in \mathcal{B}_n} \text{Tr}[\sigma_n S_n] \quad (215)$$

$$= \text{Tr}[\sigma'_n S_n] \quad (216)$$

$$\leq \text{Tr}[g_{n,d} \omega_n S_n] \quad (217)$$

$$\leq g_{n,d} \Pr[P_n(X'_n) \geq \exp(\lambda_n) Q_n(X'_n)]. \quad (218)$$

Here we use the fact that S_n is permutation invariant and therefore the optimal solution is achieved at some permutation invariant state σ'_n , which is then upper bounded by $g_{n,d} \omega_n$.

Moreover, by direct calculation, for any $s > 1$, we have

$$\begin{aligned} \Pr[P_n(X'_n) \geq \exp(\lambda_n) Q_n(X'_n)] \\ = \sum_{x: P_n(x) \geq \exp(\lambda_n) Q_n(x)} Q_n(x) \end{aligned} \quad (219)$$

$$\leq \sum_{x: P_n(x) \geq \exp(\lambda_n) Q_n(x)} Q_n(x)^{1-s} [\exp(-\lambda_n) P_n(x)]^s \quad (220)$$

$$\leq \sum_x Q_n(x)^{1-s} [\exp(-\lambda_n) P_n(x)]^s \quad (221)$$

$$= \exp(-s\lambda_n) \text{Tr}[P_n^s Q_n^{1-s}], \quad (222)$$

where the first inequality uses the condition $P_n(x) \geq \exp(\lambda_n) Q_n(x)$, and the second inequality extends the sum to all x . Consequently, we obtain

$$\beta(\mathcal{B}_n, S_n) \leq g_{n,d} \exp(-s\lambda_n) \text{Tr}[P_n^s Q_n^{1-s}]. \quad (223)$$

For arbitrary fixed $s \in (1, \infty)$, we choose λ_n such that

$$\lambda_n := \frac{1}{s} \left[\phi(1 - s|P_n||Q_n) + \log g_{n,d} + nr \right]. \quad (224)$$

Consider a random variable

$$Z_n := \frac{1}{n} (\log P_n(X_n) - \log Q_n(X_n) - \lambda_n). \quad (225)$$

Hence, by the choice of λ_n , we have $\beta(\mathcal{B}_n, S_n) \leq \exp(-nr)$. Moreover,

$$1 - \alpha(\rho_n, S_n) = \Pr[P_n(X_n) \geq \exp(\lambda_n) Q_n(X_n)] \quad (226)$$

$$= \Pr[Z_n \geq 0]. \quad (227)$$

So the optimal Type-I error satisfies

$$1 - \alpha_{n,r}(\rho_n||\mathcal{B}_n) \geq 1 - \alpha(\rho_n, S_n) = \Pr[Z_n \geq 0]. \quad (228)$$

This implies that

$$\begin{aligned} \limsup_{n \rightarrow \infty} -\frac{1}{n} \log(1 - \alpha_{n,r}(\rho_n||\mathcal{B}_n)) \\ \leq \limsup_{n \rightarrow \infty} -\frac{1}{n} \log \Pr[Z_n \geq 0]. \end{aligned} \quad (229)$$

In the following, we will use the Gärtner-Ellis theorem to upper bound the right-hand side. For this, we need to compute the asymptotic cumulant generating function of Z_n .

For $t \geq -\frac{1}{2}$, we have

$$\Lambda_Z(t) := \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E}[e^{ntZ_n}] \quad (230)$$

$$\begin{aligned} = \lim_{n \rightarrow \infty} \frac{1}{n} \log \left\{ \sum_x P_n(x) \cdot \right. \\ \left. \exp(t(\log P_n(x) - \log Q_n(x) - \lambda_n)) \right\} \end{aligned} \quad (231)$$

$$\begin{aligned} = \lim_{n \rightarrow \infty} \frac{1}{n} \left\{ \log \sum_x P_n(x)^{1+t} Q_n(x)^{-t} \right. \\ \left. - \frac{t}{s} [\phi(1 - s|P_n||Q_n) + \log g_{n,d} + nr] \right\} \end{aligned} \quad (232)$$

$$\begin{aligned} = \lim_{n \rightarrow \infty} \frac{1}{n} \left\{ \log \text{Tr} P_n^{1+t} Q_n^{-t} \right. \\ \left. - \frac{t}{s} [\phi(1 - s|P_n||Q_n) + \log g_{n,d} + nr] \right\} \end{aligned} \quad (233)$$

$$\begin{aligned} = \lim_{n \rightarrow \infty} \frac{1}{n} \left\{ \phi(-t|P_n||Q_n) \right. \\ \left. - \frac{t}{s} [\phi(1 - s|P_n||Q_n) + \log g_{n,d} + nr] \right\}. \end{aligned} \quad (234)$$

Denote $g(\alpha) := D_{s,\alpha}^\infty(\mathcal{A}||\mathcal{B})$ for $\alpha > 1$ and $g(1) = D^\infty(\mathcal{A}||\mathcal{B})$. We have the relation that

$$\phi(s|P_n||Q_n) = \log \text{Tr}[P_n^{1-s} Q_n^s] = -sD_{1-s}(P_n||Q_n). \quad (235)$$

This means that

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} \phi(-t|P_n||Q_n) \\ = \lim_{n \rightarrow \infty} \frac{1}{n} tD_{1+t}(P_n||Q_n) \end{aligned} \quad (236)$$

$$= \lim_{n \rightarrow \infty} \frac{1}{n} tD_{s,1+t}(\mathcal{P}_{\omega_n}(\rho_n)||\omega_n) \quad (237)$$

$$= tg(1+t), \quad (238)$$

where the second equality follows as $\mathcal{P}_{\omega_n}(\rho_n)$ and ω_n commute and the last equality follows from Lemma 10. Similarly, we have

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} \phi(1 - s|P_n||Q_n) \\ = \lim_{n \rightarrow \infty} \frac{1}{n} (s-1)D_s(P_n||Q_n) \end{aligned} \quad (239)$$

$$= \lim_{n \rightarrow \infty} \frac{1}{n} (s-1)D_{s,s}(\mathcal{P}_{\omega_n}(\rho_n)||\omega_n) \quad (240)$$

$$= (s-1)g(s). \quad (241)$$

Define the function

$$f(s, t) := t(r - sg(1+t) + (s-1)g(s)). \quad (242)$$

Then we have

$$\Lambda_Z(t) = -\frac{f(s, t)}{s}. \quad (243)$$

As we assume that $g(\alpha)$ is differentiable in α for $\alpha \geq 1$. Then $f(s, t)$ is differentiable in both s and t . Moreover, we have

$$\Lambda'_Z(t) = -\frac{1}{s} (r - sg(1+t) + (s-1)g(s) - tsg'(1+t)). \quad (244)$$

Since $g(\alpha)$ is monotonically increasing and differentiable for $\alpha \geq 1$, we have for any $s > 1$,

$$\lim_{t \rightarrow 0^+} \Lambda'_Z(t) = -\frac{1}{s} (r - sg(1) + (s-1)g(s)) \quad (245)$$

$$= g(1) - \frac{s-1}{s} g(s) - \frac{r}{s} \quad (246)$$

$$\leq g(1) - \frac{s-1}{s} g(1) - \frac{r}{s} \quad (247)$$

$$= \frac{1}{s} (g(1) - r) \quad (248)$$

$$< 0, \quad (249)$$

where we use the assumption that $r > g(1) = D^\infty(\mathcal{A}||\mathcal{B})$ in the last inequality. On the other hand, using the convexity of

$t \mapsto \psi(t) := tg(1+t)$ and $\psi(0) = 0$ (see Remark 12), we have $\psi(\lambda t) \leq \lambda\psi(t)$ for all $\lambda \in (0, 1)$. This implies that

$$\psi'(t) = \lim_{\lambda \rightarrow 1^-} \frac{\psi(t) - \psi(\lambda t)}{t(1-\lambda)} \geq \frac{\psi(t)}{t}, \quad (250)$$

and therefore we have $tg'(1+t) \geq 0$. Let t_0 be such that $r < g(t_0 + 1)$. Then for any $s \leq t_0 + 1$, we find that

$$\lim_{t \rightarrow t_0^-} \Lambda'_Z(t) \geq g(t_0 + 1) - \frac{s-1}{s}g(s) - \frac{r}{s} \quad (251)$$

$$\geq g(t_0 + 1) - \frac{s-1}{s}g(t_0 + 1) - \frac{r}{s} \quad (252)$$

$$= \frac{1}{s}(g(t_0 + 1) - r) > 0, \quad (253)$$

where we have used that $r < g(t_0 + 1)$. Hence, applying the Gärtner-Ellis theorem in [21, Proposition 17] for the random variable Z_n on the interval $(0, t_0)$ and with the threshold value 0, we have

$$\begin{aligned} \limsup_{n \rightarrow \infty} -\frac{1}{n} \log \Pr[Z_n \geq 0] \\ \leq \sup_{t \in (0, t_0)} \{-\Lambda_Z(t)\} \leq \sup_{t \in [0, t_0]} \frac{f(s, t)}{s}. \end{aligned} \quad (254)$$

As this holds for any $s \in (1, t_0 + 1]$, we have from (229) that

$$\begin{aligned} \limsup_{n \rightarrow \infty} -\frac{1}{n} \log(1 - \alpha_{n,r}(\rho_n \| \mathcal{B}_n)) \\ \leq \inf_{s \in (1, t_0 + 1]} \sup_{t \in [0, t_0]} \frac{f(s, t)}{s}. \end{aligned} \quad (255)$$

By the convexity of $tg(1+t)$ (Remark 12), we known that $f(s, t)$ is concave in t and convex in s . We optimize t over a compact convex set and s over a convex set. Therefore, by the minimax theorem in [21, Proposition 21], we can exchange the order of inf and sup. This gives

$$\inf_{s \in (1, t_0 + 1]} \sup_{t \in [0, t_0]} \frac{f(s, t)}{s} = \sup_{t \in [0, t_0]} \inf_{s \in (1, t_0 + 1]} \frac{f(s, t)}{s} \quad (256)$$

$$= \sup_{t \in (0, t_0)} \inf_{s \in (1, t_0 + 1]} \frac{f(s, t)}{s} \quad (257)$$

$$\leq \sup_{t \in (0, t_0]} \frac{f(t+1, t)}{t+1} \quad (258)$$

$$\leq \sup_{t > 0} \frac{f(t+1, t)}{t+1}, \quad (259)$$

where the second line follows as $f(s, 0) = 0$ and the penultimate line follows by choosing $s = 1 + t$. This gives

$$\begin{aligned} \limsup_{n \rightarrow \infty} -\frac{1}{n} \log(1 - \alpha_{n,r}(\rho_n \| \mathcal{B}_n)) \\ \leq \sup_{t > 0} \frac{f(t+1, t)}{t+1} = \mathfrak{H}_r^{*,\infty}(\mathcal{A} \| \mathcal{B}). \end{aligned} \quad (260)$$

VII. APPLICATIONS

Since the error exponent and strong converse exponent regimes provide a finer characterization of the trade-off between Type-I and Type-II errors than the Stein regime, we apply our results to refine and extend several existing studies. Specifically, we: (i) strengthen the generalized quantum Stein's lemma in [15] for hypothesis testing between sets of

states; (ii) provide counterexamples to the continuity of the regularized Petz Rényi divergence and Hoeffding divergence; (iii) derive error exponents for adversarial quantum channel discrimination; and (iv) obtain error exponents for resource detection problems in coherence theory and entanglement theory.

A. Refining Quantum Stein's Lemma Between Two Sets

The following generalized quantum Stein's lemma for hypothesis testing between two sets of quantum states was established in [15, Theorem 32].

Theorem 3 (Generalized quantum Stein's lemma): Let $\mathcal{A} = \{\mathcal{A}_n\}_{n \in \mathbb{N}}$ and $\mathcal{B} = \{\mathcal{B}_n\}_{n \in \mathbb{N}}$ be two sequences of sets of quantum states satisfying [15, Assumption 24] and $\mathcal{A}_n, \mathcal{B}_n \subseteq \mathcal{D}(\mathcal{H}^{\otimes n})$ and $D_{\max}(\mathcal{A}_n \| \mathcal{B}_n) \leq cn$, for all $n \in \mathbb{N}$ and a constant $c \in \mathbb{R}_+$. For any $\varepsilon \in (0, 1)$,

$$\lim_{n \rightarrow \infty} -\frac{1}{n} \log \beta_\varepsilon(\mathcal{A}_n \| \mathcal{B}_n) = D^\infty(\mathcal{A} \| \mathcal{B}). \quad (261)$$

The above result can be both recovered and strengthened as follows.

Theorem 4: Let $\mathcal{A} = \{\mathcal{A}_n\}_{n \in \mathbb{N}}$ and $\mathcal{B} = \{\mathcal{B}_n\}_{n \in \mathbb{N}}$ be two sequences of sets satisfying the same assumptions in Theorem 3. For any $0 < r < D^\infty(\mathcal{A} \| \mathcal{B})$, then

$$\liminf_{n \rightarrow \infty} -\frac{1}{n} \log \alpha_{n,r}(\mathcal{A}_n \| \mathcal{B}_n) \geq \mathfrak{H}_r^\infty(\mathcal{A} \| \mathcal{B}) > 0. \quad (262)$$

For any $r > D^\infty(\mathcal{A} \| \mathcal{B})$, then

$$\liminf_{n \rightarrow \infty} -\frac{1}{n} \log(1 - \alpha_{n,r}(\mathcal{A}_n \| \mathcal{B}_n)) \geq \mathfrak{H}_r^{*,\infty}(\mathcal{A} \| \mathcal{B}) > 0. \quad (263)$$

This result shows that any Type-II error exponent below $D^\infty(\mathcal{A} \| \mathcal{B})$ is achievable, with the corresponding Type-I error decaying exponentially at a rate at least $\mathfrak{H}_r^\infty(\mathcal{A} \| \mathcal{B})$. Conversely, if the Type-II error exponent exceeds $D^\infty(\mathcal{A} \| \mathcal{B})$, the Type-I error inevitably converges to one exponentially, with a rate at least $\mathfrak{H}_r^{*,\infty}(\mathcal{A} \| \mathcal{B})$. Thus, the regularized quantum relative entropy $D^\infty(\mathcal{A} \| \mathcal{B})$ delineates a sharp threshold for the asymptotic trade-off in hypothesis testing between two sets of quantum states. In particular, these results apply to adversarial quantum channel discrimination, which satisfies all the required assumptions [14]. See Section VII-C for more detailed discussion in this setting.

Proof: By the assumptions on the sequences, we have

$$D^\infty(\mathcal{A} \| \mathcal{B}) = \sup_{\alpha \in (0, 1)} D_{M, \alpha}^\infty(\mathcal{A} \| \mathcal{B}) \quad (264)$$

$$\leq \sup_{\alpha \in (0, 1)} D_{P, \alpha}^\infty(\mathcal{A} \| \mathcal{B}) \quad (265)$$

$$\leq D^\infty(\mathcal{A} \| \mathcal{B}), \quad (266)$$

where $D_{M, \alpha}$ refers to the measured Rényi divergence. The first equality follows from [15, Lemmas 27, 28], and the inequalities use that $D_{M, \alpha}(\rho \| \sigma) \leq D_{P, \alpha}(\rho \| \sigma) \leq D(\rho \| \sigma)$ for any $\alpha \in (0, 1)$. This implies

$$\sup_{\alpha \in (0, 1)} D_{P, \alpha}^\infty(\mathcal{A} \| \mathcal{B}) = D^\infty(\mathcal{A} \| \mathcal{B}). \quad (267)$$

Note that $D_{p,\alpha}^\infty(\mathcal{A}||\mathcal{B})$ is monotone increasing in α . Therefore, for any $0 < r < D^\infty(\mathcal{A}||\mathcal{B})$, there exists $\alpha \in (0, 1)$ such that $r < D_{p,\alpha}^\infty(\mathcal{A}||\mathcal{B})$. Then

$$\mathfrak{H}_r^\infty(\mathcal{A}||\mathcal{B}) = \sup_{\alpha \in (0,1)} \frac{\alpha - 1}{\alpha} (r - D_{p,\alpha}^\infty(\mathcal{A}||\mathcal{B})) > 0. \quad (268)$$

By Theorem 1 and Lemma 3, we have

$$\begin{aligned} \liminf_{n \rightarrow \infty} -\frac{1}{n} \log \alpha_{n,r}(\mathcal{A}_n||\mathcal{B}_n) \\ = H_r^\infty(\mathcal{A}||\mathcal{B}) \geq \mathfrak{H}_r^\infty(\mathcal{A}||\mathcal{B}) > 0 \end{aligned} \quad (269)$$

which shows that the Type-I error decays exponentially, and thus r is an achievable rate. This recovers the direct part of the generalized quantum Stein's lemma in Theorem 3.

Since $\inf_{\alpha > 1} D_{s,\alpha}^\infty(\mathcal{A}||\mathcal{B}) = D^\infty(\mathcal{A}||\mathcal{B})$ [15, Lemma 27] and $D_{s,\alpha}^\infty(\mathcal{A}||\mathcal{B})$ is monotone increasing in α , for any $r > D^\infty(\mathcal{A}||\mathcal{B})$, there exists $\alpha > 1$ such that $r > D_{s,\alpha}^\infty(\mathcal{A}||\mathcal{B})$. This implies that

$$\mathfrak{H}_r^{*,\infty}(\mathcal{A}||\mathcal{B}) = \sup_{\alpha > 1} \frac{\alpha - 1}{\alpha} (r - D_{s,\alpha}^\infty(\mathcal{A}||\mathcal{B})) > 0. \quad (270)$$

Applying Theorem 2 and Lemma 6, we obtain

$$\begin{aligned} \liminf_{n \rightarrow \infty} -\frac{1}{n} \log (1 - \alpha_{n,r}(\mathcal{A}_n||\mathcal{B}_n)) \\ \geq H_r^{*,\infty}(\mathcal{A}||\mathcal{B}) = \mathfrak{H}_r^{*,\infty}(\mathcal{A}||\mathcal{B}) > 0. \end{aligned} \quad (271)$$

This shows that the Type-I error converges to one exponentially, and thus r is not an achievable rate, recovering the converse part of the generalized quantum Stein's lemma in Theorem 3. $\square$

It is worth emphasizing that the quantum Hoeffding bound in Theorem 1 and the strong converse exponent in Theorem 2 (Eq. (191)) hold in great generality and do not require, particularly, the polar assumption as used in [15]. However, to recover the Stein's setting from the error exponent regime, one needs the continuity of the regularized Petz Rényi divergences, i.e., $\sup_{\alpha \in (0,1)} D_{p,\alpha}^\infty(\mathcal{A}||\mathcal{B}) = D^\infty(\mathcal{A}||\mathcal{B})$. For this, we require the polar assumption in Eq. (264). Interestingly, a similar situation also arises in the context of best-case channel discrimination, where the continuity of the regularized sandwiched Rényi divergence between channels is sufficient to establish the quantum Stein's lemma for two quantum channels (particularly the strong converse part); see [42, Theorem 21] for further details.

B.: Discontinuity of $D_{p,\alpha}^\infty(\mathcal{A}||\mathcal{B})$ at $\alpha = 1$

As discussed above, our error exponent results can recover the Stein's exponent provided the continuity of the regularized Petz Rényi divergence holds at $\alpha = 1$, i.e.,

$$\sup_{\alpha \in (0,1)} D_{p,\alpha}^\infty(\mathcal{A}||\mathcal{B}) = D^\infty(\mathcal{A}||\mathcal{B}). \quad (272)$$

Therefore, for any sequences of sets $\mathcal{A} = \{\mathcal{A}_n\}_{n \in \mathbb{N}}$ and $\mathcal{B} = \{\mathcal{B}_n\}_{n \in \mathbb{N}}$ that satisfy our error exponent assumptions but violate the Stein's lemma, there must be a discontinuity in $D_{p,\alpha}^\infty(\mathcal{A}||\mathcal{B})$ at $\alpha = 1$, i.e., a violation of Eq. (272). Such examples can be found in [43].

In this section, we provide a more direct analysis of these counterexamples and introduce additional ones. We show that

such discontinuities can occur in both nonconvex and convex settings, as well as in scenarios where a composite hypothesis is tested against a simple hypothesis (see Example 1 and Example 2), and conversely, where a simple hypothesis is tested against a composite hypothesis (see Example 3). Furthermore, the example presented in [44] for the nonadditivity of the Petz Rényi divergence between two sets, makes both sides of Eq. (272) diverge, which is not a valid counterexample for discontinuity.

In the following, we denote $H(p) := -p \log p - (1-p) \log(1-p)$ as the binary entropy function and $D(p||q) := p \log \frac{p}{q} + (1-p) \log \frac{1-p}{1-q}$ as the classical relative entropy between two binary distributions $(p, 1-p)$ and $(q, 1-q)$. Denote the set of types with length n by $\mathcal{P}_n$, which is the set of all empirical distributions. Give $P \in \mathcal{P}_n$, the type class of P , denoted as $\mathcal{T}_n(P)$, is the set of bit strings whose empirical distribution is P . We choose two rational numbers $p_1 = \frac{m_1}{m}$ and $p_2 = \frac{m_2}{m}$ such that $p_1 < p_2 \in (0, p_3)$. Let $P_j = (p_j, 1-p_j)$ be the binary distribution. Consider a two-dimensional Hilbert space $\mathcal{H}$ spanned by $\{|0\rangle, |1\rangle\}$ and define the qubit state

$$\rho(p) := p|0\rangle\langle 0| + (1-p)|1\rangle\langle 1|. \quad (273)$$

We denote the uniform distribution over $\mathcal{T}_{km}(P_j)$ in the space $\mathcal{H}^{\otimes km}$ by

$$\rho_{j,km} := \frac{1}{|\mathcal{T}_{km}(P_j)|} \sum_{x^{km} \in \mathcal{T}_{km}(P_j)} |x^{km}\rangle\langle x^{km}|, \quad (274)$$

for $j = 1, 2$, where $|\mathcal{T}_{km}(P_j)| = \binom{km}{k_{mj}}$ is the size of the type class. Note that $\rho_{1,km}$ and $\rho_{2,km}$ are orthogonal, as their supports are defined by different type classes. We have $\rho(p_j) = \text{Tr}_{km-1} \rho_{j,km}$, where the partial trace can be understood as tracing out any $km-1$ subsystems, as it gives the same marginal state. In the following examples, we regard the Hilbert space $\mathcal{H}^{\otimes m}$ as a one-copy system.

Example 1 (Nonconvex sets, $\mathcal{A}_n$ composite and $\mathcal{B}_n$ simple i.i.d.): Let $\rho(p)$ and $\rho_{j,km}$ defined by Eqs. (273) and (274), respectively. We define the convex mixture,

$$\rho_{km} := \lambda \rho_{1,km} + (1-\lambda) \rho_{2,km}, \quad (275)$$

with the choice of $\lambda \in (0, 1)$ such that

$$\min_{i \in \{1,2\}} \frac{1}{2} \log \frac{2\pi m_i(m-m_i)}{m} \geq H(\lambda). \quad (276)$$

Consider the set of states

$$\begin{aligned} \mathcal{A}_n &:= \left\{ \rho_{k_1 m} \otimes \cdots \otimes \rho_{k_l m} : k_1, \dots, k_l \geq 1, \sum_{i=1}^l k_i = n \right\}, \\ \mathcal{B}_n &:= \left\{ \rho(p_3)^{\otimes mn} \right\}. \end{aligned} \quad (277)$$

Then sequences $\mathcal{A} := \{\mathcal{A}_n\}$ and $\mathcal{B} := \{\mathcal{B}_n\}$ are stable under tensor product. But we have

$$\sup_{\alpha \in (0,1)} D_{p,\alpha}^\infty(\mathcal{A}||\mathcal{B}) < D^\infty(\mathcal{A}||\mathcal{B}). \quad (278)$$

Moreover, we have

$$\lim_{r \rightarrow 0} H_r^\infty(\mathcal{B}||\mathcal{A}) < D^\infty(\mathcal{A}||\mathcal{B}), \quad (279)$$

$$\lim_{r \rightarrow 0} \mathfrak{H}_r^\infty(\mathcal{B} \parallel \mathcal{A}) < D^\infty(\mathcal{A} \parallel \mathcal{B}), \quad (280)$$

indicating that the Hoeffding bound cannot recover the Stein's exponent in this case.

Proof: Using the standard result in type method [45, Theorem 11.1.2], we have

$$\begin{aligned} D(\rho_{i,km} \parallel \rho(p_3)^{\otimes km}) \\ = -\log \binom{km}{km_i} + km(D(p_i \parallel p_3) + H(p_i)), \end{aligned} \quad (281)$$

for $i = 1, 2$. As $\rho_{1,km}$ and $\rho_{2,km}$ are orthogonal, we have

$$\begin{aligned} D(\rho_{km} \parallel \rho(p_3)^{\otimes km}) &= \lambda D(\rho_{1,km} \parallel \rho(p_3)^{\otimes km}) \\ &+ (1 - \lambda) D(\rho_{2,km} \parallel \rho(p_3)^{\otimes km}) - H(\lambda). \end{aligned} \quad (282)$$

Note that the binomial coefficient $\binom{n}{l}$ is evaluated as [46]

$$\binom{n}{l} \leq \sqrt{\frac{n}{2\pi l(n-l)}} 2^{nH(l/n)}, \quad (283)$$

for all $1 \leq l \leq n-1$. Thus, we have

$$\begin{aligned} -\log \binom{km}{km_i} + kmH(p_i) - H(\lambda) \\ \geq \frac{1}{2} \log \frac{2\pi km_i(m - m_i)}{m} - H(\lambda) \geq 0, \end{aligned} \quad (284)$$

for $i \in \{1, 2\}$, where the last inequality follows from the choice of λ in Eq. (276). Discarding these two positive terms, we have

$$D(\rho_{km} \parallel \rho(p_3)^{\otimes km}) \geq km((1 - \lambda)D(p_2 \parallel p_3) + \lambda D(p_1 \parallel p_3)). \quad (285)$$

By the additivity of quantum relative entropy, we have

$$\begin{aligned} D(\rho_{k_1m} \otimes \cdots \otimes \rho_{k_l m} \parallel \rho(p_3)^{\otimes nm}) \\ \geq nm((1 - \lambda)D(p_2 \parallel p_3) + \lambda D(p_1 \parallel p_3)). \end{aligned} \quad (286)$$

This implies that

$$D(\mathcal{A}_n \parallel \mathcal{B}_n) \geq nm((1 - \lambda)D(p_2 \parallel p_3) + \lambda D(p_1 \parallel p_3)), \quad (287)$$

$$D^\infty(\mathcal{A} \parallel \mathcal{B}) \geq m((1 - \lambda)D(p_2 \parallel p_3) + \lambda D(p_1 \parallel p_3)). \quad (288)$$

Using $k = n$ in Eq. (282), we obtain

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} D(\rho_{nm} \parallel \rho(p_3)^{\otimes nm}) \\ = m((1 - \lambda)D(p_2 \parallel p_3) + \lambda D(p_1 \parallel p_3)), \end{aligned} \quad (289)$$

where some terms vanish because of the Stirling's approximation [45, Example 11.1.3]

$$\frac{1}{n+1} 2^{nH(k/n)} \leq \binom{n}{k} \leq 2^{nH(k/n)}. \quad (290)$$

Note that $\rho_{nm} \in \mathcal{A}_n$ by choosing $l = 1$, so we have $D(\mathcal{A}_n \parallel \mathcal{B}_n) \leq D(\rho_{nm} \parallel \rho(p_3)^{\otimes nm})$. Combined with Eq. (289), this implies that

$$D^\infty(\mathcal{A} \parallel \mathcal{B}) \leq m((1 - \lambda)D(p_2 \parallel p_3) + \lambda D(p_1 \parallel p_3)). \quad (291)$$

Therefore, we have from Eqs. (288) and (291) that

$$D^\infty(\mathcal{A} \parallel \mathcal{B}) = m((1 - \lambda)D(p_2 \parallel p_3) + \lambda D(p_1 \parallel p_3)). \quad (292)$$

We now estimate the Petz Rényi divergence. For $\alpha \in (0, 1)$, we have

$$\begin{aligned} \text{Tr}(\rho_{i,km}^\alpha (\rho(p_3)^{\otimes km})^{1-\alpha}) \\ = \left[\binom{km}{km_i} 2^{-km(D(p_i \parallel p_3) + H(p_i))} \right]^{1-\alpha}, \end{aligned} \quad (293)$$

for $i = 1, 2$, by the type method [45, Theorem 11.1.2]. Since $\rho_{1,km}$ and $\rho_{2,km}$ are orthogonal, we have

$$\begin{aligned} \text{Tr}(\rho_{km}^\alpha (\rho(p_3)^{\otimes km})^{1-\alpha}) \\ = \lambda^\alpha \text{Tr}(\rho_{1,km}^\alpha (\rho(p_3)^{\otimes km})^{1-\alpha}) \\ + (1 - \lambda)^\alpha \text{Tr}(\rho_{2,km}^\alpha (\rho(p_3)^{\otimes km})^{1-\alpha}). \end{aligned} \quad (294)$$

Using Lemma 14 and the fact that $D(p_2 \parallel p_3) < D(p_1 \parallel p_3)$, we have for any $\alpha \in (0, 1)$,

$$D_{p,\alpha}^\infty(\mathcal{A} \parallel \mathcal{B}) \leq \lim_{n \rightarrow \infty} \frac{1}{n} D_{p,\alpha}(\rho_{nm} \parallel \rho(p_3)^{\otimes nm}) \quad (295)$$

$$= mD(p_2 \parallel p_3), \quad (296)$$

where some terms vanish because of the Stirling's approximation in Eq. (290).

Combining (292) and (296), we have the strict inequality,

$$\sup_{\alpha \in (0,1)} D_{p,\alpha}^\infty(\mathcal{A} \parallel \mathcal{B}) < D^\infty(\mathcal{A} \parallel \mathcal{B}). \quad (297)$$

Now, we further analyze the Hoeffding bound. For $r > 0$, we have from the alternative expression of the Hoeffding bound (similar to Remark 6) that

$$\begin{aligned} H_{n,r}(\rho(p_3)^{\otimes nm} \parallel \rho_{nm}) \\ = \sup_{\alpha \in (0,1)} \left(-\frac{\alpha}{1-\alpha} nr + D_{p,\alpha}(\rho_{nm} \parallel \rho(p_3)^{\otimes nm}) \right). \end{aligned} \quad (298)$$

By Eq. (289), we have for any $\delta > 0$,

$$\begin{aligned} \frac{1}{n} D_{p,\alpha}(\rho_{nm} \parallel \rho(p_3)^{\otimes nm}) \\ \leq \frac{1}{n} D(\rho_{nm} \parallel \rho(p_3)^{\otimes nm}) \\ \leq m((1 - \lambda)D(p_2 \parallel p_3) + \lambda D(p_1 \parallel p_3)) + \delta =: z, \end{aligned} \quad (299)$$

$$\leq m((1 - \lambda)D(p_2 \parallel p_3) + \lambda D(p_1 \parallel p_3)) + \delta =: z, \quad (300)$$

for sufficiently large n . Then for any $\alpha > t_r := \frac{z}{r+z} \in (0, 1)$, we can check that

$$-\frac{\alpha}{1-\alpha} nr + D_{p,\alpha}(\rho_{nm} \parallel \rho(p_3)^{\otimes nm}) < 0. \quad (301)$$

So for $r > 0$, we have

$$\begin{aligned} \sup_{\alpha \in (0,1)} \left(-\frac{\alpha}{1-\alpha} nr + D_{p,\alpha}(\rho_{nm} \parallel \rho(p_3)^{\otimes nm}) \right) \\ = \sup_{\alpha \in (0,t_r]} \left(-\frac{\alpha}{1-\alpha} nr + D_{p,\alpha}(\rho_{nm} \parallel \rho(p_3)^{\otimes nm}) \right). \end{aligned} \quad (302)$$

Therefore, we have

$$\begin{aligned} H_{n,r}(\rho(p_3)^{\otimes nm} \parallel \rho_{nm}) \\ = \sup_{\alpha \in (0,t_r]} \left(-\frac{\alpha}{1-\alpha} nr + D_{p,\alpha}(\rho_{nm} \parallel \rho(p_3)^{\otimes nm}) \right). \end{aligned} \quad (303)$$

This implies that

$$\lim_{n \rightarrow \infty} \frac{1}{n} H_{n,r}(\rho(p_3)^{\otimes nm} \parallel \rho_{nm})$$

$$= \lim_{n \rightarrow \infty} \frac{1}{n} \sup_{\alpha \in (0,1)} \left(-\frac{\alpha}{1-\alpha} nr + D_{p,\alpha}(\rho_{nm} \|\rho(p_3)^{\otimes nm}) \right) \quad (304)$$

$$\leq \lim_{n \rightarrow \infty} \frac{1}{n} \sup_{\alpha \in (0,1)} D_{p,\alpha}(\rho_{nm} \|\rho(p_3)^{\otimes nm}) \quad (305)$$

$$= \lim_{n \rightarrow \infty} \frac{1}{n} D_{p,\alpha}(\rho_{nm} \|\rho(p_3)^{\otimes nm}) \quad (306)$$

$$= mD(p_2 \| p_3), \quad (307)$$

where the inequality follows by throwing away a negative term, the second equality follows by the monotonicity of Petz Rényi divergence in α and the last equality follows from Eq. (296).

This implies that

$$H_r^\infty(\mathcal{B} \|\mathcal{A}) \leq \lim_{n \rightarrow \infty} \frac{1}{n} H_{n,r}(\rho(p_3)^{\otimes nm} \|\rho_{nm}) \quad (308)$$

$$\leq mD(p_2 \| p_3). \quad (309)$$

Similarly, we have

$$\mathfrak{H}_r^\infty(\mathcal{B} \|\mathcal{A}) = \sup_{\alpha \in (0,1)} \left(-\frac{\alpha}{1-\alpha} r + D_{p,\alpha}(\mathcal{A} \|\mathcal{B}) \right) \quad (310)$$

$$\leq \sup_{\alpha \in (0,1)} \left(-\frac{\alpha}{1-\alpha} nr + mD(p_2 \| p_3) \right) \quad (311)$$

$$= mD(p_2 \| p_3), \quad (312)$$

where the inequality follows from Eq. (296).

Thus, combining Eqs. (292), (309), and (312), we have

$$\lim_{r \rightarrow 0} H_r^\infty(\mathcal{B} \|\mathcal{A}) < D^\infty(\mathcal{A} \|\mathcal{B}), \quad (313)$$

$$\lim_{r \rightarrow 0} \mathfrak{H}_r^\infty(\mathcal{B} \|\mathcal{A}) < D^\infty(\mathcal{A} \|\mathcal{B}). \quad (314)$$

This completes the proof. $\square$

Here we note that, through additional analysis, we can also establish that $D_{p,\alpha}(\mathcal{A} \|\mathcal{B}) = mD(p_2 \| p_3)$ for $\alpha \in (0,1)$ satisfying

$$\lambda^\alpha 2^{-(1-\alpha)mD(p_1 \| p_3)} \leq (1 - (1-\lambda)^\alpha) 2^{-(1-\alpha)mD(p_2 \| p_3)}. \quad (315)$$

At least, when α is sufficiently close to 1, the above condition holds. Since

$$\binom{l_1 + l_2}{l} = \sum_{l'=0}^l \binom{l_1}{l-l'} \binom{l_2}{l'} \geq \binom{l_1}{l-l''} \binom{l_2}{l''} \quad (316)$$

for $l'' \in [0, l]$, we have

$$\binom{km}{km_j}^l \leq \binom{lk m}{lk m_j} \leq 2^{klmH(p_j)}, \quad (317)$$

for $j = 1, 2$, where the second inequality follows from Eq. (283). Then, we have

$$\lambda^\alpha \left[\binom{km}{km_1} 2^{-km(D(p_1 \| p_3) + H(p_1))} \right]^{1-\alpha} \leq \lambda^\alpha 2^{-(1-\alpha)kmD(p_1 \| p_3)} \quad (318)$$

$$\leq (1 - (1-\lambda)^\alpha) 2^{-(1-\alpha)mD(p_2 \| p_3)} 2^{-(1-\alpha)(k-1)mD(p_1 \| p_3)} \quad (319)$$

$$\leq (1 - (1-\lambda)^\alpha) 2^{-(1-\alpha)kmD(p_2 \| p_3)}, \quad (320)$$

where the first inequality follows from Eq. (317) and the second inequality follows from Eq. (315) and the fact that $D(p_2 \| p_3) < D(p_1 \| p_3)$. Similarly, we have

$$\left[\binom{km}{km_2} 2^{-km(D(p_2 \| p_3) + H(p_2))} \right]^{1-\alpha} \leq 2^{-(1-\alpha)kmD(p_2 \| p_3)}. \quad (321)$$

Thus, we have

$$\begin{aligned} & \text{Tr}(\rho_{km})^\alpha (\rho(p_3)^{\otimes km})^{1-\alpha} \\ & \leq (1 - (1-\lambda)^\alpha) 2^{-(1-\alpha)kmD(p_2 \| p_3)} \\ & \quad + (1-\lambda)^\alpha 2^{-(1-\alpha)kmD(p_1 \| p_3)} \end{aligned} \quad (322)$$

$$= 2^{-(1-\alpha)kmD(p_2 \| p_3)}, \quad (323)$$

where the inequality follows from Eqs. (294), (320) and (321). By the additivity of Petz Rényi divergence, we have

$$D_{p,\alpha}(\rho_{k_1 m} \otimes \cdots \otimes \rho_{k_l m} \|\rho(p_3)^{\otimes nm}) \geq nmD(p_2 \| p_3). \quad (324)$$

This implies that

$$D_{p,\alpha}^\infty(\mathcal{A} \|\mathcal{B}) \geq mD(p_2 \| p_3). \quad (325)$$

Combining Eqs. (296) and (325), we have

$$D_{p,\alpha}^\infty(\mathcal{A} \|\mathcal{B}) = mD(p_2 \| p_3), \quad (326)$$

for Example 1 when α satisfies Eq. (315).

An illustration for the discontinuities of the regularized Petz Rényi divergence and the Hoeffding divergence is shown in Figure. 5, where we can clearly see the discontinuities of both regularized quantities at $\alpha = 1$ and $r = 0$, respectively.

The second example follows the construction in [43, Appendix C], though that reference omits the technical details which we provide here.

Example 2 (Convex sets, $\mathcal{A}_n$ composite and $\mathcal{B}_n$ simple i.i.d): Let $\rho(p_3)$ and ρ_{km} defined in Example 1. Consider the sets $\mathcal{S}_n := \{\text{Tr}_{km-n} \rho_{km} : k \in \mathbb{N}\}$, where Tr_{km-n} is the partial trace over the initial $km - n$ subsystems. Define

$$\begin{aligned} \mathcal{A}_n &:= \text{conv} \left\{ \rho_{n_1} \otimes \cdots \otimes \rho_{n_l} : n_1, \dots, n_l \geq 1, \right. \\ & \quad \left. \rho_{n_j} \in \mathcal{S}_{n_j}, \sum_{j=1}^l n_j = n \right\}, \end{aligned} \quad (327)$$

and $\mathcal{B}_n = \{\rho(p_3)^{\otimes n}\}$. Then the sequences $\mathcal{A} := \{\mathcal{A}_n\}$ and $\mathcal{B} := \{\mathcal{B}_n\}$ are convex compact and stable under tensor product. But we have

$$\sup_{\alpha \in (0,1)} D_{p,\alpha}^\infty(\mathcal{A} \|\mathcal{B}) < D^\infty(\mathcal{A} \|\mathcal{B}). \quad (328)$$

Moreover, we have

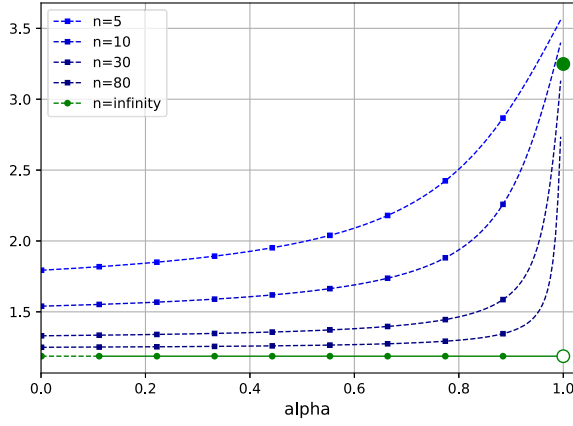
$$\lim_{r \rightarrow 0} H_r^\infty(\mathcal{B} \|\mathcal{A}) < D^\infty(\mathcal{A} \|\mathcal{B}), \quad (329)$$

$$\lim_{r \rightarrow 0} \mathfrak{H}_r^\infty(\mathcal{B} \|\mathcal{A}) < D^\infty(\mathcal{A} \|\mathcal{B}), \quad (330)$$

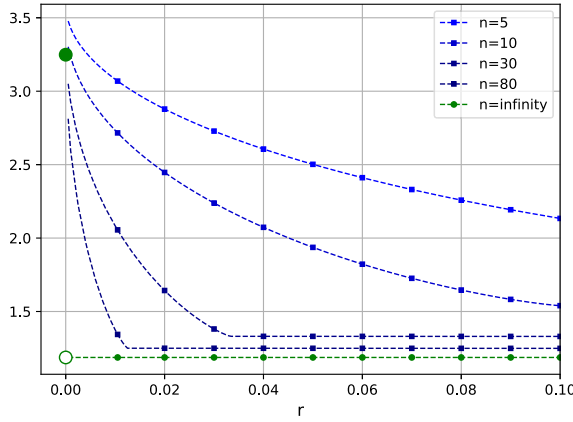
indicating that the Hoeffding bound cannot recover the Stein's exponent in this case.

Proof: Let ρ be any quantum state on n subsystems. We can easily check the identity

$$D\left(\rho \left\| \bigotimes_{i=1}^n \sigma_i\right.\right) = D\left(\rho \left\| \bigotimes_{i=1}^n \rho_i\right.\right) + \sum_{i=1}^n D(\rho_i \|\sigma_i), \quad (331)$$



(a) Petz Rényi divergence



(b) Hoeffding divergence

Fig. 5. Discontinuities in the regularized Petz Rényi divergence and Hoeffding divergence given in Example 1. We choose $m_1 = 1$, $m_2 = 3$, $m = 10$, $p_3 = \lambda = 0.5$. Panel (a) shows the upper bound for Petz Rényi divergence $\frac{1}{n} D_{P,\alpha}(\rho_{nm} \| \rho(p_3)^{\otimes nm})$ for different values of n as a function of α , which can be computed from Eq. (294). The solid line indicates tightness of the upper bound as analyzed in Eq. (326), and the marked dot at $\alpha = 1$ corresponds to the regularized relative entropy in Eq. (292). Panel (b) displays the upper bound for Hoeffding divergence $\frac{1}{n} H_{n,r}(\rho(p_3)^{\otimes nm} \| \rho_{nm})$ from Eq. (298) for different n as a function of r ; for $n = \infty$, the upper bound from Eq. (309) is used. The marked dot at $r = 0$ again indicates the regularized relative entropy from Eq. (292).

where ρ_i is the reduced state of ρ on the i -th subsystem. Then we have for any $\rho \in \mathcal{A}_n$,

$$D(\rho \| \rho(p_3)^{\otimes n}) = D(\rho \| (\text{Tr}_{n-1} \rho)^{\otimes n}) + D((\text{Tr}_{n-1} \rho)^{\otimes n} \| \rho(p_3)^{\otimes n}). \quad (332)$$

Since $\text{Tr}_{n-1} \rho = \rho(\lambda p_1 + (1 - \lambda) p_2)$, we have

$$D(\rho \| \rho(p_3)^{\otimes n}) \geq n D(\rho(\lambda p_1 + (1 - \lambda) p_2) \| \rho(p_3)) \quad (333)$$

$$= n D(\lambda p_1 + (1 - \lambda) p_2 \| p_3). \quad (334)$$

This implies that

$$D(\mathcal{A}_n \| \mathcal{B}_n) \geq n D(\lambda p_1 + (1 - \lambda) p_2 \| p_3). \quad (335)$$

On the other hand, since $\rho(\lambda p_1 + (1 - \lambda) p_2)^{\otimes n} \in \mathcal{A}_n$ and $\rho(p_3)^{\otimes n} \in \mathcal{B}_n$, the above bound is achievable at these states and therefore, we have

$$D(\mathcal{A}_n \| \mathcal{B}_n) = n D(\lambda p_1 + (1 - \lambda) p_2 \| p_3). \quad (336)$$

Thus, we have

$$D^\infty(\mathcal{A} \| \mathcal{B}) = D(\lambda p_1 + (1 - \lambda) p_2 \| p_3). \quad (337)$$

Since $\rho_{km} \otimes \rho_{n-km} \in \mathcal{A}_n$ with $0 \leq n - km < m$, we have

$$D_{P,\alpha}(\mathcal{A}_n \| \mathcal{B}_n) \leq D_{P,\alpha}(\rho_{km} \| \rho(p_3)^{\otimes km}) + D_{P,\alpha}(\rho_{n-km} \| \rho(p_3)^{\otimes n-km}). \quad (338)$$

Using the same calculation as in Eq. (296), for $\alpha > 0$,

$$D_{P,\alpha}^\infty(\mathcal{A} \| \mathcal{B}) \leq D(p_2 \| p_3). \quad (339)$$

Therefore, we have

$$\sup_{\alpha \in (0,1)} D_{P,\alpha}^\infty(\mathcal{A} \| \mathcal{B}) \leq D(p_2 \| p_3) \quad (340)$$

$$< D(\lambda p_1 + (1 - \lambda) p_2 \| p_3) \quad (341)$$

$$= D^\infty(\mathcal{A} \| \mathcal{B}). \quad (342)$$

Finally, the statement for the Hoeffding bounds can be shown similarly as in Example 1. This completes the proof. $\square$

The next example demonstrates an extreme case where the regularized Petz Rényi divergence vanishes while the regularized relative entropy diverges.

Example 3 (Convex sets, $\mathcal{A}_n$ simple i.i.d. and $\mathcal{B}_n$ composite): Let $\rho(p_2)$ and $\rho_{2,km}$ defined by Eqs. (273) and (274), respectively. Consider the sets $\mathcal{A}_n := \{\rho(p_2)^{\otimes nm}\}$, and

$$\mathcal{B}_n := \text{conv} \left\{ \rho_{2,k_1 m} \otimes \cdots \otimes \rho_{2,k_l m} : k_1, \dots, k_l \geq 1, \sum_{i=1}^l k_i = n \right\}. \quad (343)$$

Then the sequences $\mathcal{A} := \{\mathcal{A}_n\}$ and $\mathcal{B} := \{\mathcal{B}_n\}$ are convex compact and stable under tensor product. But we have

$$\sup_{\alpha \in (0,1)} D_{P,\alpha}^\infty(\mathcal{A} \| \mathcal{B}) = 0 < \infty = D^\infty(\mathcal{A} \| \mathcal{B}). \quad (344)$$

Proof: Note that the support of the state $\rho_{2,k_1 m} \otimes \cdots \otimes \rho_{2,k_l m}$ with $\sum_{i=1}^l k_i = n$ is included in the typical subspace $\mathcal{T}_{nm}(P_2)$. The support $\rho(p_2)^{\otimes nm}$ is full space. Thus, we have

$$D(\mathcal{A}_n \| \mathcal{B}_n) = \min_{\sigma \in \mathcal{B}_n} D(\rho(p_2)^{\otimes nm} \| \sigma) = \infty, \quad (345)$$

and therefore,

$$D^\infty(\mathcal{A} \| \mathcal{B}) = \infty. \quad (346)$$

For the Petz Rényi divergence, we have

$$D_{P,\alpha}(\rho(p_2)^{\otimes km} \| \rho_{2,km}) = \frac{1}{\alpha - 1} \log \left[2^{-km H(p_2)} \binom{km}{km_2} \right]^\alpha, \quad (347)$$

by the standard result in type method as in [45, Eq. (11.15)]. By the Stirling's approximation in Eq. (290), we have

$$\lim_{n \rightarrow \infty} \frac{1}{n} D_{P,\alpha}(\rho(p_2)^{\otimes nm} \| \rho_{2,nm}) = 0. \quad (348)$$

Thus, we have for any $\alpha \in (0, 1)$,

$$D_{P,\alpha}^\infty(\mathcal{A} \| \mathcal{B}) \leq \lim_{n \rightarrow \infty} \frac{1}{n} D_{P,\alpha}(\rho(p_2)^{\otimes nm} \| \rho_{2,nm}) = 0. \quad (349)$$

As the divergence is always nonnegative, the above equality is attained. Finally, the combination of Eqs. (346) and (349) implies that

$$\sup_{\alpha \in (0,1)} D_{P,\alpha}^\infty(\mathcal{A} \parallel \mathcal{B}) = 0 < \infty = D^\infty(\mathcal{A} \parallel \mathcal{B}). \quad (350)$$

This completes the proof. $\square$

The last example is presented in [44] to demonstrate the nonadditivity of the Petz Rényi divergence between two sets of quantum states. However, we show that this example does not give a violation of Eq. (272) as both sides diverge.

Example 4 (Convex sets): Let $\mathcal{A}_n := \text{SEP}_n$ be the set of bipartite separable states on n subsystems. Let $\mathcal{B}_n := \{\rho^{\otimes n}\}$ with $\rho := \frac{1}{d(d-1)} \left(I - \sum_{i,j=1}^d |ij\rangle\langle ij| \right)$ being the antisymmetric Werner state on $\mathbb{C}^d \otimes \mathbb{C}^d$. Then the sequences $\mathcal{A} := \{\mathcal{A}_n\}$ and $\mathcal{B} := \{\mathcal{B}_n\}$ are convex compact and stable under tensor product. But we have

$$\sup_{\alpha \in (0,1)} D_{P,\alpha}^\infty(\mathcal{A} \parallel \mathcal{B}) = D^\infty(\mathcal{A} \parallel \mathcal{B}) = \infty. \quad (351)$$

Proof: It has been shown in [44, Eq. (E5)] that for any $n \in \mathbb{N}$ and $\alpha \in (0, 1)$,

$$D_{P,\alpha}(\mathcal{A}_n \parallel \mathcal{B}_n) = \frac{\alpha}{1-\alpha} D(\mathcal{B}_n \parallel \mathcal{A}_n). \quad (352)$$

This implies

$$D_{P,\alpha}^\infty(\mathcal{A} \parallel \mathcal{B}) = \frac{\alpha}{1-\alpha} D^\infty(\mathcal{B} \parallel \mathcal{A}), \quad (353)$$

by taking regularization on both sides. Note that we have an explicit lower bound $D^\infty(\mathcal{B} \parallel \mathcal{A}) \geq \log_2 \sqrt{4/3} > 0$ [47, Corollary 3]. Taking supremum over $\alpha \in (0, 1)$, we have

$$\sup_{\alpha \in (0,1)} D_{P,\alpha}^\infty(\mathcal{A} \parallel \mathcal{B}) = \sup_{\alpha \in (0,1)} \frac{\alpha}{1-\alpha} D^\infty(\mathcal{B} \parallel \mathcal{A}) = \infty. \quad (354)$$

However, as we always have the direction,

$$\sup_{\alpha \in (0,1)} D_{P,\alpha}^\infty(\mathcal{A} \parallel \mathcal{B}) \leq D^\infty(\mathcal{A} \parallel \mathcal{B}), \quad (355)$$

this implies that

$$\sup_{\alpha \in (0,1)} D_{P,\alpha}^\infty(\mathcal{A} \parallel \mathcal{B}) = D^\infty(\mathcal{A} \parallel \mathcal{B}) = \infty. \quad (356)$$

$\square$

C. Error Exponents in Adversarial Channel Discrimination

The adversarial quantum channel discrimination was recently proposed and analyzed in [14], where a tester interacts with an untrusted quantum device that generates quantum states upon request. The device guarantees that the states are produced by either a quantum channel $\mathcal{N}$ or a quantum channel $\mathcal{M}$. More formally, let $\mathcal{N}_{A \rightarrow B}$ and $\mathcal{M}_{A \rightarrow B}$ be the two quantum channels to be distinguished, and let $\mathcal{U}_{A \rightarrow BE}$ and $\mathcal{V}_{A \rightarrow BE}$ denote their respective Stinespring dilations, with E representing the environmental system. Let $\text{CPTP}(X: Y)$ denote the set of completely positive and trace-preserving maps from input system X to output system Y .

An *adaptive strategy* for adversarial discrimination proceeds as follows. Suppose the device operates as channel $\mathcal{N}$. Initially, the adversary prepares a quantum state via an operation

$\mathcal{P}^1 \in \text{CPTP}(R_0 E_0 : A_1 R_1)$, where R_0 and E_0 are trivial ($|R_0| = |E_0| = 1$), and sends system A_1 through the channel $\mathcal{U}$, generating the output state $\mathcal{U} \circ \mathcal{P}^1$ and returning system B_1 to the tester. In the next round, the adversary performs an internal update $\mathcal{P}^2 \in \text{CPTP}(E_1 R_1 : A_2 R_2)$, utilizing information stored in the quantum memory R_1 and the environmental system E_1 from the previous round. The adversary then sends system A_2 through the channel $\mathcal{U}$ again, producing the output state $\mathcal{U} \circ \mathcal{P}^2 \circ \mathcal{U} \circ \mathcal{P}^1$ and returning system B_2 to the tester. This process can be repeated for n rounds. A *non-adaptive strategies* for adversarial discrimination is a subclass of adaptive strategies that disregards the environmental systems E_i and performs no updates between rounds, that is, taking the operations $\mathcal{P}_i, \mathcal{Q}_i$ ($i \geq 2$) simply as identity maps, with the choice $R_i = A_{i+1} \cdots A_n$.

After n rounds of state generation, the tester obtains an overall state on $B_1 \cdots B_n$ in their possession as:

$$\rho[\{\mathcal{P}^i\}_{i=1}^n] := \text{Tr}_{R_n E_n} \prod_{i=1}^n [\mathcal{U}_{A_i \rightarrow B_i E_i} \circ \mathcal{P}_{R_{i-1} E_{i-1} \rightarrow A_i R_i}^i].$$

Similarly, if the device is governed by $\mathcal{M}$ and the internal operations by the adversary are given by $\mathcal{Q}^i$, then the overall state is given by

$$\sigma[\{\mathcal{Q}^i\}_{i=1}^n] := \text{Tr}_{R_n E_n} \prod_{i=1}^n [\mathcal{V}_{A_i \rightarrow B_i E_i} \circ \mathcal{Q}_{R_{i-1} E_{i-1} \rightarrow A_i R_i}^i].$$

The tester needs to perform a binary quantum measurement $\{M_n, I - M_n\}$ on systems $B_1 \cdots B_n$ to determine which channel was used inside the black box.

Due to limited knowledge of the device's internal workings and the adversary's strategies, the tester only gets partial information, knowing that the state belongs to one of two sets:

$$\begin{aligned} \mathcal{A}_n &:= \{\rho[\{\mathcal{P}^i\}_{i=1}^n] : \mathcal{P}^i \in \text{CPTP}(R_{i-1} E_{i-1} : A_i R_i), \forall R_i, \forall i\}, \\ \mathcal{B}_n &:= \{\sigma[\{\mathcal{Q}^i\}_{i=1}^n] : \mathcal{Q}^i \in \text{CPTP}(R_{i-1} E_{i-1} : A_i R_i), \forall R_i, \forall i\}, \end{aligned}$$

where the adversary's internal memory R_i may have arbitrarily large dimension. In particular, for non-adaptive strategies, where the adversary ignores the environmental systems E_i and performs no updates between rounds, the sets $\mathcal{A}_n$ and $\mathcal{B}_n$ reduce to

$$\mathcal{A}'_n := \{\mathcal{N}^{\otimes n}(\rho_n) : \rho_n \in \mathcal{D}(A^{\otimes n})\}, \quad (357)$$

$$\mathcal{B}'_n := \{\mathcal{M}^{\otimes n}(\sigma_n) : \sigma_n \in \mathcal{D}(A^{\otimes n})\}. \quad (358)$$

It has been shown in [14] that the optimal Stein exponent for both adaptive and non-adaptive strategies is given by the regularized quantum relative entropy between the two channels:

$$D^{\text{inf},\infty}(\mathcal{N} \parallel \mathcal{M}) := \lim_{n \rightarrow \infty} \frac{1}{n} D^{\text{inf}}(\mathcal{N}^{\otimes n} \parallel \mathcal{M}^{\otimes n}), \quad (359)$$

where the quantum relative entropy between two channels is defined by

$$D^{\text{inf}}(\mathcal{N} \parallel \mathcal{M}) := \inf \{D(\mathcal{N}(\rho) \parallel \mathcal{M}(\sigma)) : \rho, \sigma \in \mathcal{D}(A)\}. \quad (360)$$

1) *Error Exponents:* We can check that $\{\mathcal{A}_n\}, \{\mathcal{A}'_n\}, \{\mathcal{B}_n\}, \{\mathcal{B}'_n\}$ are all stable sequences of convex and compact sets of quantum states (the convexity of $\{\mathcal{A}_n\}, \{\mathcal{B}_n\}$ is proved in [14]). Therefore, our result directly applies to adversarial quantum channel discrimination. In particular, for any $0 < r < D^{\text{inf},\infty}(\mathcal{N}||\mathcal{M})$, the optimal Type-I error exponent under an exponential constraint on the Type-II error is given by

$$\lim_{n \rightarrow \infty} -\frac{1}{n} \log \alpha_{n,r}(\mathcal{A}_n||\mathcal{B}_n) = H_r^\infty(\mathcal{A}||\mathcal{B}), \quad (361)$$

$$\lim_{n \rightarrow \infty} -\frac{1}{n} \log \alpha_{n,r}(\mathcal{A}'_n||\mathcal{B}'_n) = H_r^\infty(\mathcal{A}'||\mathcal{B}') > 0, \quad (362)$$

where we can ensure the strict positivity as the continuity of the regularized Petz Rényi divergence holds in this case [14]. It would be interesting to further investigate whether adaptive strategies can outperform non-adaptive strategies in terms of the error exponent, i.e., whether $H_r^\infty(\mathcal{A}||\mathcal{B}) < H_r^\infty(\mathcal{A}'||\mathcal{B}')$ for some channels $\mathcal{N}, \mathcal{M}$ and rate r .

The authors in [21] consider the error exponent and strong converse exponents for distinguishing the null hypothesis $\mathcal{A}''_n = \{\rho_{AB}^{\otimes n}\}$ from the composite alternative $\mathcal{B}''_n = \{\rho_A^{\otimes n} \otimes \sigma_{B^n} : \sigma \in \mathcal{D}(B^n)\}$, which reflects the scenario of correlation detection. This can be modelled as an adversarial channel discrimination problem as well. In particular, let $\mathcal{R}_\rho$ be a replacer channel that always outputs ρ regardless of the input. Then the above $\mathcal{A}''_n$ can be regarded as the image set of the replacer channel $(\mathcal{R}_{\rho_{AB}})^{\otimes n}$ and $\mathcal{B}''_n$ is the image set of the channel $(\mathcal{R}_{\rho_A} \otimes I_B)^{\otimes n}$. All assumptions in Theorem 1 and Theorem 2 are satisfied in this case. Therefore, we can recover the results in [21] as:

$$\begin{aligned} \lim_{n \rightarrow \infty} -\frac{1}{n} \log \alpha_{n,r}(\mathcal{A}''_n||\mathcal{B}''_n) \\ = \sup_{\alpha \in (0,1)} \frac{\alpha - 1}{\alpha} (r - I_{p,\alpha}(A:B)_\rho), \end{aligned} \quad (363)$$

for $0 < r < I(A:B)_\rho$, where $I(A:B)_\rho := D(\mathcal{A}''_1||\mathcal{B}''_1)$ and $I_{p,\alpha}(A:B)_\rho := D_{p,\alpha}(\mathcal{A}''_1||\mathcal{B}''_1)$ denote, respectively, the (Umegaki) mutual information and the Petz Rényi mutual information (which is additive), and

$$\begin{aligned} \lim_{n \rightarrow \infty} -\frac{1}{n} \log(1 - \alpha_{n,r}(\mathcal{A}''_n||\mathcal{B}''_n)) \\ = \sup_{\alpha > 1} \frac{\alpha - 1}{\alpha} (r - I_{s,\alpha}(A:B)_\rho), \end{aligned} \quad (364)$$

for $I(A:B)_\rho < r < \sup_{\alpha > 1} I_{s,\alpha}(A:B)_\rho$, where $I_{s,\alpha}(A:B)_\rho := D_{s,\alpha}(\mathcal{A}''_1||\mathcal{B}''_1)$ denotes the sandwiched Rényi mutual information (which is also additive).

2) *Examples for Nonadditivity:* The adversarial channel discrimination also provides a scenario where we can find nonadditivity examples for the Petz Rényi divergence and Hoeffding divergence between two sets of quantum state. For this, we consider two qutrit quantum channels. Let $\mathcal{N}(\cdot) = \text{Tr}[\cdot]\rho$ to be the replacer channel with

$$\rho = 0.9 \cdot |\psi\rangle\langle\psi| + 0.1 \cdot \frac{I}{3}, \quad (365)$$

where $|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |2\rangle)$. Let $\mathcal{M}$ be the platypus channel [48], $\mathcal{M}(X) = M_0 X M_0^\dagger + M_1 X M_1^\dagger$ with Kraus operators

$$M_0 = \begin{bmatrix} \sqrt{p} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \quad M_1 = \begin{bmatrix} 0 & 0 & 0 \\ \sqrt{1-p} & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (366)$$

In this case, we have the image sets of these channels as

$$\mathcal{A}'_n = \{\rho^{\otimes n}\}, \quad (367)$$

$$\mathcal{B}'_n = \{\mathcal{M}^{\otimes n}(\sigma_n) : \sigma_n \in \mathcal{D}((\mathbb{C}^3)^{\otimes n})\}. \quad (368)$$

As these sets are given by semidefinite constraints, we can efficiently evaluate the Petz Rényi divergence and Hoeffding divergence between them via semidefinite programming (see Remark 5). More explicitly, for fixed $\alpha \in (0, 1)$, we can efficiently evaluate

$$D_{p,\alpha}(\mathcal{A}'_n||\mathcal{B}'_n) = \inf_{\sigma_n \in \mathcal{D}} D_{p,\alpha}(\rho^{\otimes n}||\mathcal{M}^{\otimes n}(\sigma_n)), \quad (369)$$

by the QICS package [36]. Moreover, we can also evaluate the Hoeffding divergence

$$\begin{aligned} H_{n,r}(\mathcal{A}'_n||\mathcal{B}'_n) \\ = \mathfrak{H}_{n,r}(\mathcal{A}'_n||\mathcal{B}'_n) \end{aligned} \quad (370)$$

$$= \sup_{\alpha \in (0,1)} \frac{\alpha - 1}{\alpha} (nr - D_{p,\alpha}(\mathcal{A}'_n||\mathcal{B}'_n)), \quad (371)$$

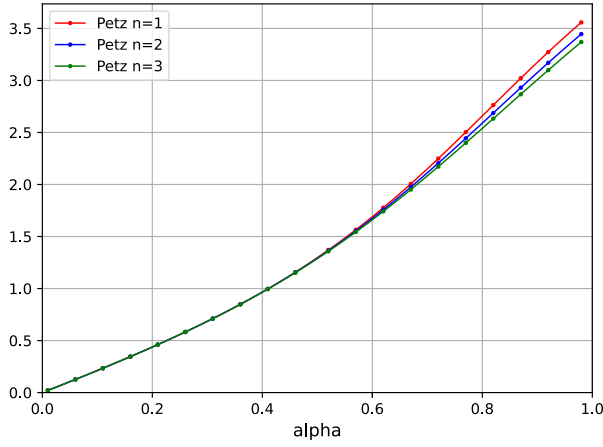
by scanning the parameter $\alpha \in (0, 1)$ with fine grid.

The numerical results are shown in Figure 6 where we use channel parameter $p = 0.01$ and scan $\alpha \in (0, 1)$ with step size 0.01. Panel (a) displays the Petz Rényi divergence $D_{p,\alpha}(\mathcal{A}'_n||\mathcal{B}'_n)/n$ for $n = 1, 2, 3$ as a function of α , while panel (b) shows the objective function of the Hoeffding divergence versus α , together with its maximum (i.e., the Hoeffding divergence $H_{n,r}(\mathcal{A}'_n||\mathcal{B}'_n)/n$) for $n = 1, 2, 3$. Here we choose the rate $r = 1$. The plots exhibit a clear separation between different numbers of copies for both the Petz Rényi and Hoeffding divergences, illustrating non-additivity in this example and justifying the necessity of regularization in our main results.

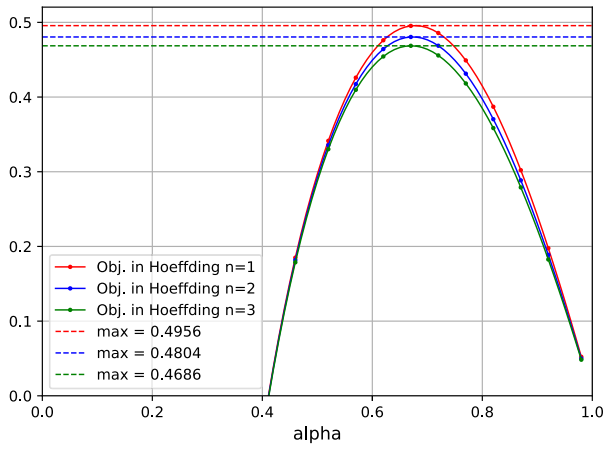
D. Error Exponents in Quantum Resource Detection

As our results only require minimal assumptions, they can be readily applied to the detection problem in various quantum resource theories (see e.g. [16]), where the stability of the set of free states is naturally satisfied [49], [50], including entanglement [9], [51], [52], [53], coherence [54], [55], [56], [57], [58], asymmetry [59], thermodynamics [60] and purity [61] and magic [62], [63], [64], [65]. Our refined analysis of error exponents can thus provide a better understanding of the trade-offs between Type-I and Type-II errors in hypothesis testing within these resource theories.

We showcase coherence theory and entanglement theory as two illustrative examples with explicit expressions. These examples have been previously studied in [16], but the error exponents obtained therein are not tight. Here we provide tight characterizations using our results.



(a) Petz Rényi divergence



(b) Hoeffding divergence

Fig. 6. Non-additivity for the Petz Rényi divergence and the Hoeffding divergence between the image set of two channels. Here, we consider the replacer channel $\mathcal{N}$ with output state $\rho = 0.9 \cdot |\psi\rangle\langle\psi| + 0.1 \cdot I/3$, where $|\psi\rangle = (|0\rangle + |2\rangle)/\sqrt{2}$, and the platypus channel $\mathcal{M}$ with channel parameter $p = 0.01$. We plot (a) the Petz Rényi divergence $D_{p,\alpha}(\mathcal{A}'_n\|\mathcal{B}'_n)/n$ and (b) the objective function in Hoeffding divergence for $n = 1, 2, 3$ as functions of $\alpha \in (0, 1)$, respectively. The Hoeffding divergence $H_{n,1}(\mathcal{A}'_n\|\mathcal{B}'_n)/n$ is indicated by the maximum value in dashed lines.

1) *Quantum Coherence Detection*: In the resource theory of coherence, the set of free states is given by the set of incoherent states $\mathcal{I}_n$, which is convex, compact, and stable under tensor products. Moreover, the Petz Rényi divergence between a quantum state and the set of incoherent states is known to be additive [66, Theorem 3]. Therefore, our results yield explicit expressions for the error exponent in hypothesis testing between a quantum state and the set of incoherent states

$$\lim_{n \rightarrow \infty} -\frac{1}{n} \log \alpha_{n,r}(\rho^{\otimes n} \|\mathcal{I}_n) = \sup_{\alpha \in (0,1)} \frac{\alpha - 1}{\alpha} (r - C_\alpha(\rho)), \quad (372)$$

for $0 < r < C(\rho) := D(\rho \|\mathcal{I}_1)$ where $C_\alpha(\rho) := D_{p,\alpha}(\rho \|\mathcal{I}_1)$ has closed form expression [67] as

$$C_\alpha(\rho) = \frac{\alpha}{\alpha - 1} \log \text{Tr} [(\Delta(\rho^\alpha))^{1/\alpha}]. \quad (373)$$

where $\Delta(\cdot)$ is the completely dephasing operation in the reference basis.

As for the strong converse exponent, the sandwiched Rényi divergence between these sets is also additive [66, Theorem 3] and is given by

$$C_\alpha^*(\rho) := D_{s,\alpha}(\rho \|\mathcal{I}_1) = \inf_{\sigma_B \in \mathcal{D}(B)} D_{s,\alpha}(\rho_{\text{mc}} \| I_A \otimes \sigma_B), \quad (374)$$

where $\rho_{\text{mc}} := \sum_{i,j} \rho_{ij} |ii\rangle\langle jj|$ is the associated maximally correlated state if $\rho = \sum_{i,j} \rho_{ij} |i\rangle\langle j|$. The differentiability of $C_\alpha^*(\rho)$ with respect to α has been established in [21, Proposition 11]. Therefore, our results yield the strong converse exponent in this case as well:

$$\begin{aligned} \lim_{n \rightarrow \infty} -\frac{1}{n} \log(1 - \alpha_{n,r}(\rho^{\otimes n} \|\mathcal{I}_n)) \\ = \sup_{\alpha > 1} \frac{\alpha - 1}{\alpha} (r - C_\alpha^*(\rho)), \end{aligned} \quad (375)$$

for $C(\rho) < r < \sup_{\alpha > 1} C_\alpha^*(\rho)$.

2) *Quantum Entanglement Detection*: In the resource theory of entanglement, the set of free states is the set of separable states SEP_n , which is convex, compact, and stable under tensor products. For a maximally correlated state $\rho_{\text{mc}} := \sum_{i,j} \rho_{ij} |ii\rangle\langle jj|$ —which includes many important cases of interest, e.g., pure states and mixtures of Bell states—the Petz and sandwiched Rényi divergences relative to the set of separable states are known to be additive. The error exponent and strong converse exponent are then analogous to those in the resource theory of coherence. In particular, the error exponent is given by

$$\begin{aligned} \lim_{n \rightarrow \infty} -\frac{1}{n} \log \alpha_{n,r}(\rho_{\text{mc}}^{\otimes n} \|\text{SEP}_n) \\ = \sup_{\alpha \in (0,1)} \frac{\alpha - 1}{\alpha} (r - E_\alpha(\rho_{\text{mc}})), \end{aligned} \quad (376)$$

for $0 < r < E(\rho_{\text{mc}}) := D(\rho_{\text{mc}} \|\text{SEP}_1)$, where $E_\alpha(\rho_{\text{mc}}) := D_{p,\alpha}(\rho_{\text{mc}} \|\text{SEP}_1) = D_{p,\alpha}(\rho \|\mathcal{I}_1)$ is the Rényi relative entropy of entanglement, which has the closed-form expression [66, Corollary 1]:

$$E_\alpha(\rho_{\text{mc}}) = \frac{\alpha}{\alpha - 1} \log \text{Tr} [(\Delta(\rho_{\text{mc}}^\alpha))^{1/\alpha}]. \quad (377)$$

Similarly, the strong-converse exponent is given by

$$\begin{aligned} \lim_{n \rightarrow \infty} -\frac{1}{n} \log(1 - \alpha_{n,r}(\rho_{\text{mc}}^{\otimes n} \|\text{SEP}_n)) \\ = \sup_{\alpha > 1} \frac{\alpha - 1}{\alpha} (r - E_\alpha^*(\rho_{\text{mc}})), \end{aligned} \quad (378)$$

for $E(\rho_{\text{mc}}) < r < \sup_{\alpha > 1} E_\alpha^*(\rho_{\text{mc}})$, where $E_\alpha^*(\rho_{\text{mc}}) := D_{s,\alpha}(\rho_{\text{mc}} \|\text{SEP}_1) = D_{s,\alpha}(\rho \|\mathcal{I}_1)$ is the sandwiched Rényi relative entropy of entanglement and we can choose a universal state as the tensor product of the universal state on A^n and B^n , respectively. Finally, the same result works if we consider the set of PPT states as the free states, as the Petz and sandwiched Rényi divergences are the same as those relative to the set of separable states for maximally correlated states [66].

VIII. DISCUSSION

We have established a framework for analyzing error exponents in quantum hypothesis testing between two sets of quantum states, extending beyond the conventional i.i.d. setting to encompass composite correlated hypotheses. Our main

contributions include a generalization of the quantum Hoeffding bound and strong converse exponent to stable sequences of convex, compact sets of quantum states. These results provide a finer characterization than the Stein regime, yielding deeper insights into the fundamental trade-offs between Type-I and Type-II errors in quantum state discrimination with composite correlated hypotheses. The broad applicability of our framework is demonstrated through concrete applications to adversarial quantum channel discrimination and various quantum resource theories, highlighting its potential for further advances across quantum information theory.

Several open questions remain. While we showed $H_r^\infty(\mathcal{A}||\mathcal{B}) \geq \mathfrak{H}_r^\infty(\mathcal{A}||\mathcal{B})$, proving equality in full generality remains open and would likely require a minimax theorem suited to the regularized setting. It would also be interesting, as a first step, to establish equality under some physically natural assumptions. Extending the strong converse upper bound beyond the current technical restrictions is another important challenge. Resolving these issues would complete the asymptotic characterization of quantum state discrimination for composite, correlated hypotheses. Finally, since the generalized quantum Stein's lemma implies asymptotic reversibility in the corresponding resource theory [15], [27], [28], our refined error exponent analysis may shed new light on conversion rates and convergence properties of resource inter-conversion, enabling finer control of asymptotic reversibility and efficiency. We leave these questions for future work.

APPENDIX

The following lemma is a minimax theorem that accounts for the infinity values of the function. Let X be a convex set in a linear space. A function $f : X \rightarrow (-\infty, \infty]$ said to be convex, if $f(px + (1-p)y) \leq pf(x) + (1-p)f(y)$, the multiplication $0 \cdot f(x)$ is interpreted as 0 and $p \cdot \infty = \infty$ for $p \neq 0$. Similar definition holds for concave functions.

Lemma 11 [68, Theorem 5.2]: Let X be a compact, convex subset of a Hausdorff topological vector space and Y be a convex subset of the linear space. Let $f : X \times Y \rightarrow (-\infty, \infty]$ be lower semicontinuous on X for fixed $y \in Y$, and assume that f is convex in the first and concave in the second variable. Then we have

$$\sup_{y \in Y} \inf_{x \in X} f(x, y) = \inf_{x \in X} \sup_{y \in Y} f(x, y). \quad (379)$$

The following lemmas are standard results in mathematical analysis and will be used frequently in our proofs. For detailed proofs, see, e.g., [34, Lemma 2.8, 2.9].

Lemma 12: Let X be a nonempty compact topological space, and let $f : X \rightarrow \mathbb{R}$ be a function. Then if f is upper semicontinuous, it attains its maximum, meaning there is some $x \in X$ such that for all $x' \in X$, $f(x') \leq f(x)$. Similarly, if f is lower semicontinuous, it attains its minimum.

Lemma 13: Let X be a topological space, let I be a set, and let $\{f_i\}_{i \in I}$ be a collection of functions $f_i : X \rightarrow \mathbb{R}$. Then if each f_i is upper semicontinuous, the function $f(x) = \inf_{i \in I} f_i(x)$ is also upper semicontinuous. Similarly, if each f_i is lower semicontinuous, the pointwise supremum is lower semicontinuous.

Lemma 14: For two scalar sequences $x_n, y_n > 0$ with $\limsup_{n \rightarrow \infty} -\frac{\log x_n}{n} = x_R$ and $\lim_{n \rightarrow \infty} -\frac{\log y_n}{n} = y_R$, then $\limsup_{n \rightarrow \infty} -\frac{1}{n} \log(x_n + y_n) = \min\{x_R, y_R\}$.

Proof: For any scalars $x, y > 0$, we can verify that [7],

$$\max\{\log x, \log y\} \leq \log(x + y) \leq \max\{\log x, \log y\} + 1. \quad (380)$$

Therefore, we have

$$\limsup_{n \rightarrow \infty} -\frac{1}{n} \log(x_n + y_n) \leq \limsup_{n \rightarrow \infty} -\frac{1}{n} \log x_n = x_R, \quad (381)$$

$$\limsup_{n \rightarrow \infty} -\frac{1}{n} \log(x_n + y_n) \leq \limsup_{n \rightarrow \infty} -\frac{1}{n} \log y_n = y_R. \quad (382)$$

This gives $\limsup_{n \rightarrow \infty} -\frac{1}{n} \log(x_n + y_n) \leq \min\{x_R, y_R\}$. For the other direction, by the definition of x_R , for any $\varepsilon > 0$, there exists an infinite subsequence n_m such that $-\frac{1}{n_m} \log x_{n_m} \geq x_R - \varepsilon$. As the limit of y_n exists, we know that for this subsequence $-\frac{1}{n_m} \log y_{n_m} \geq y_R - \varepsilon$. This implies that

$$\min\left\{-\frac{1}{n_m} \log x_{n_m}, -\frac{1}{n_m} \log y_{n_m}\right\} \geq \min\{x_R, y_R\} - \varepsilon \quad (383)$$

Therefore, we have

$$\begin{aligned} \limsup_{n \rightarrow \infty} -\frac{1}{n} \log(x_n + y_n) \\ \geq \limsup_{n \rightarrow \infty} -\frac{1}{n} \max\{\log x_n, \log y_n\} \end{aligned} \quad (384)$$

$$= \limsup_{n \rightarrow \infty} \min\left\{-\frac{1}{n} \log x_n, -\frac{1}{n} \log y_n\right\} \quad (385)$$

$$\geq \limsup_{m \rightarrow \infty} \min\left\{-\frac{1}{n_m} \log x_{n_m}, -\frac{1}{n_m} \log y_{n_m}\right\} \quad (386)$$

$$\geq \min\{x_R, y_R\} - \varepsilon. \quad (387)$$

As this holds for any $\varepsilon > 0$, we have

$$\limsup_{n \rightarrow \infty} -\frac{1}{n} \log(x_n + y_n) \geq \min\{x_R, y_R\}. \quad (388)$$

This completes the proof. $\square$

Note that the existence of the limit for one sequence is necessary in Lemma 14. Consider the case where $x_n = 1$ if n is odd and $x_n = 2^n$ if n is even; and $y_n = 2^n$ if n is odd and $y_n = 1$ if n is even. Then we have $\limsup_{n \rightarrow \infty} -\frac{1}{n} \log x_n = \limsup_{n \rightarrow \infty} -\frac{1}{n} \log y_n = 0$, but $\limsup_{n \rightarrow \infty} -\frac{1}{n} \log(x_n + y_n) = -1$. This gives a counterexample to Lemma 14 if the limit does not exist for either sequence.

In this section, we discuss some useful properties of the function⁵ $\phi(s|\rho||\sigma) := \log \text{Tr} \rho^{1-s} \sigma^s$. These results can be found in the book [41, Chapter 3] and we list them here for completeness. Let $P_{\rho, \sigma}$ and $Q_{\rho, \sigma}$ denote the Nussbaum-Szkoła distributions for ρ and σ . These classical distributions preserve many important information of the original states, including

$$D(\rho||\sigma) = D(P_{\rho, \sigma}||Q_{\rho, \sigma}), \quad (389)$$

$$\phi(s|\rho||\sigma) = \phi(s|P_{\rho, \sigma}||Q_{\rho, \sigma}). \quad (390)$$

⁵Note that this definition differs slightly from the one in the main text Lemma 9, but they are equivalent up to a constant factor. Thus, all subsequent results work as the same.

Due to the second equality, we may simplify the notation as $\phi(s)$ when there is no confusion. Moreover, the Hoeffding divergence can also be expressed as

$$H_{1,r}(\rho||\sigma) = \max_{s \in (0,1)} \frac{-\phi(s) - sr}{1-s}. \quad (391)$$

Lemma 15: Let $\rho, \sigma \in \mathcal{H}_+$. Then the first and second derivatives of $\phi(s|\rho||\sigma)$ with respect to s are given by [41, Exercise 3.5],

$$\phi'(s|\rho||\sigma) = \frac{\text{Tr} \rho^{1-s} \sigma^s (\log \sigma - \log \rho)}{\text{Tr} \rho^{1-s} \sigma^s}, \quad (392)$$

$$\begin{aligned} \phi''(s|\rho||\sigma) &= \frac{\text{Tr} \rho^{1-s} (\log \sigma - \log \rho) \sigma^s (\log \sigma - \log \rho)}{\text{Tr} \rho^{1-s} \sigma^s} \\ &\quad - \frac{(\text{Tr} \rho^{1-s} \sigma^s (\log \rho - \log \sigma))^2}{(\text{Tr} \rho^{1-s} \sigma^s)^2}. \end{aligned} \quad (393)$$

In particular, $\phi'(0||\rho||\sigma) = -D(\rho||\sigma)$ and $\phi'(1||\rho||\sigma) = D(\sigma||\rho)$. Moreover, using Schwarz inequality, we know that $\phi''(s|\rho||\sigma) \geq 0$ and therefore $\phi(s|\rho||\sigma)$ is convex in $s \in \mathbb{R}$. Moreover, $\phi(s|\rho||\sigma)$ is strictly convex in $s \in \mathbb{R}$ if ρ and σ are not proportional.

Proof: For the first derivative, we have

$$\phi'(s|\rho||\sigma) = \frac{(\text{Tr} \rho^{1-s} \sigma^s)'}{\text{Tr} \rho^{1-s} \sigma^s} \quad (394)$$

$$= \frac{\text{Tr}[(\rho^{1-s} \sigma^s)']}{\text{Tr} \rho^{1-s} \sigma^s} \quad (395)$$

$$= \frac{\text{Tr}[(\rho^{1-s})' \sigma^s + \rho^{1-s} (\sigma^s)']}{\text{Tr} \rho^{1-s} \sigma^s} \quad (396)$$

$$= \frac{\text{Tr}[-\rho^{1-s} (\log \rho) \sigma^s + \rho^{1-s} \sigma^s (\log \sigma)]}{\text{Tr} \rho^{1-s} \sigma^s} \quad (397)$$

$$= \frac{\text{Tr}[-(\log \rho) \rho^{1-s} \sigma^s + \rho^{1-s} \sigma^s (\log \sigma)]}{\text{Tr} \rho^{1-s} \sigma^s} \quad (398)$$

$$= \frac{\text{Tr} \rho^{1-s} \sigma^s (\log \sigma - \log \rho)}{\text{Tr} \rho^{1-s} \sigma^s} \quad (399)$$

For the second derivative, we have first that

$$\begin{aligned} &(\text{Tr} \rho^{1-s} \sigma^s (\log \sigma - \log \rho))' \\ &= \text{Tr}[(\rho^{1-s} \sigma^s (\log \sigma - \log \rho))'] \end{aligned} \quad (400)$$

$$= \text{Tr}[(\rho^{1-s} \sigma^s)' (\log \sigma - \log \rho)] \quad (401)$$

$$= \text{Tr}[-\rho^{1-s} (\log \rho) \sigma^s + \rho^{1-s} \sigma^s (\log \sigma)] (\log \sigma - \log \rho) \quad (402)$$

$$= \text{Tr}[-\rho^{1-s} (\log \rho) \sigma^s + \rho^{1-s} (\log \sigma) \sigma^s] (\log \sigma - \log \rho) \quad (403)$$

$$= \text{Tr}[\rho^{1-s} (\log \sigma - \log \rho) \sigma^s (\log \sigma - \log \rho)]. \quad (404)$$

Then we have the second derivative as

$$\begin{aligned} &\phi''(s|\rho||\sigma) \\ &= \frac{X}{(\text{Tr} \rho^{1-s} \sigma^s)^2} \end{aligned} \quad (405)$$

$$\begin{aligned} &= \frac{\text{Tr}[\rho^{1-s} (\log \sigma - \log \rho) \sigma^s (\log \sigma - \log \rho)]}{\text{Tr} \rho^{1-s} \sigma^s} \\ &\quad - \frac{(\text{Tr} \rho^{1-s} \sigma^s (\log \rho - \log \sigma))^2}{(\text{Tr} \rho^{1-s} \sigma^s)^2}, \end{aligned} \quad (406)$$

where $X = (\text{Tr} \rho^{1-s} \sigma^s (\log \sigma - \log \rho))' (\text{Tr} \rho^{1-s} \sigma^s) - (\text{Tr} \rho^{1-s} \sigma^s (\log \sigma - \log \rho)) (\text{Tr} \rho^{1-s} \sigma^s)'$. By the Schwarz inequality $|\text{Tr}(XY)|^2 \leq |\text{Tr}(XX^*)| |\text{Tr}(YY^*)|$, we have

$$\begin{aligned} &(\text{Tr} \rho^{1-s} \sigma^s (\log \rho - \log \sigma))^2 \\ &\leq \text{Tr}[\rho^{1-s} (\log \sigma - \log \rho) \sigma^s (\log \sigma - \log \rho)] \text{Tr}[\rho^{1-s} \sigma^s], \end{aligned} \quad (407)$$

by considering $X = \rho^{(1-s)/2} (\log \sigma - \log \rho) \sigma^{s/2}$ and $Y = \rho^{(1-s)/2} \sigma^{s/2}$. This implies that $\phi''(s|\rho||\sigma) \geq 0$ and therefore $\phi(s|\rho||\sigma)$ is convex in $s \in \mathbb{R}$. Note that the Schwarz inequality is saturated if and only if X and Y are linearly dependent, which is equivalent to the condition that $\rho^{(1-s)/2} (\log \sigma - \log \rho) \sigma^{s/2}$ and $\rho^{(1-s)/2} \sigma^{s/2}$ are linearly dependent. This is then equivalent to $\rho = \lambda \sigma$ for some $\lambda \in \mathbb{R}$. So in case ρ and σ are not proportional, we have $\phi''(s|\rho||\sigma) > 0$ for all $s \in \mathbb{R}$, which means that $\phi(s|\rho||\sigma)$ is strictly convex in $s \in \mathbb{R}$. $\square$

The following result can be found in [41, Exercise 3.45].

Lemma 16: Let $r < D(\rho||\sigma)$. There exists a unique maximum achieved at $s_r \in (0, 1)$ for

$$\sup_{0 \leq s \leq 1} f(s), \quad \text{with } f(s) := \frac{-sr - \phi(s|\rho||\sigma)}{1-s}. \quad (408)$$

Moreover, the optimal solution s_r satisfies

$$r = (s_r - 1)\phi'(s_r) - \phi(s_r). \quad (409)$$

Proof: Let P, Q be the Nussbaum-Szkoła distributions and let $P_s(x) := P(x)^{1-s} Q(x)^s e^{-\phi(s)}$. Then the derivative of $f(s)$ is given by

$$f'(s) = \frac{-r + (s-1)\phi'(s) - \phi(s)}{(1-s)^2}. \quad (410)$$

The sign of $f'(s)$ depends on the sign of its nominator $g(s) := -r + (s-1)\phi'(s) - \phi(s)$. By taking the derivative of $g(s)$, we have

$$g'(s) = (s-1)\phi''(s) < 0, \quad \text{for } s \in (0, 1). \quad (411)$$

This means that $g(s)$ is strictly decreasing in $s \in (0, 1)$. We also have [41, Exercise 3.45] that,

- $D(P_s||P_1) = (s-1)\phi'(s) - \phi(s)$ and $D(P_s||P_0) = s\phi'(s) - \phi(s)$.
- $\frac{d}{ds} D(P_s||P_1) = (s-1)\phi''(s) < 0$ and $\frac{d}{ds} D(P_s||P_0) = s\phi''(s) > 0$ for $s \in (0, 1)$.
- Let $r < D(P||Q)$. Then there uniquely exists $s_r \in (0, 1)$ such that $D(P_{s_r}||P_1) = r$.

The last item implies that there exists a unique critical point $s_r \in (0, 1)$ such that $g(s_r) = -r + D(P_{s_r}||P_1) = 0$. Together with the monotonicity of $g(s)$, we know that $g(s) > 0$ for $s < s_r$ and $g(s) < 0$ for $s > s_r$. This implies that $f'(s) > 0$ for $s < s_r$ and $f'(s) < 0$ for $s > s_r$. So $f(s)$ is increasing for $s < s_r$ and decreasing for $s > s_r$. This shows that $f(s)$ attains its maximum at s_r . Moreover, we can show that $f(s_r) = D(P_{s_r}||P)$. $\square$

The following shows that the optimal solution s_r is differentiable in r .

Lemma 17: The optimal solution s_r is differentiable in r and [41, Exercise 3.51],

$$\frac{ds_r}{dr} = \frac{1}{(s_r - 1)\phi''(s_r)}. \quad (412)$$

Proof: We have the following

$$1 = \frac{dr}{dr} \quad (413)$$

$$= \frac{d}{dr}((s_r - 1)\phi'(s_r) - \phi(s_r)) \quad (414)$$

$$= \frac{ds_r}{dr} \cdot \frac{d}{ds}((s - 1)\phi'(s) - \phi(s))|_{s=s_r} \quad (415)$$

$$= \frac{ds_r}{dr} \cdot ((s_r - 1)\phi''(s_r)), \quad (416)$$

where the third line uses the chain rule of differentiation. This implies the desired result. $\square$

REFERENCES

- [1] H. Chernoff, "A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations," *Ann. Math. Statist.*, vol. 23, no. 4, pp. 493–507, Dec. 1952.
- [2] J. Wolfowitz, "Coding theorems of information theory," *Math. Comput.*, vol. 274, no. 1, p. 61, 1978.
- [3] R. Blahut, "Hypothesis testing and information theory," *IEEE Trans. Inf. Theory*, vol. IT-20, no. 4, pp. 405–417, Jul. 1974.
- [4] F. Hiai and D. Petz, "The proper formula for relative entropy and its asymptotics in quantum probability," *Commun. Math. Phys.*, vol. 143, no. 1, pp. 99–114, Dec. 1991.
- [5] H. Nagaoka and T. Ogawa, "Strong converse and Stein's lemma in quantum hypothesis testing," *IEEE Trans. Inf. Theory*, vol. 46, no. 7, pp. 2428–2433, Feb. 2000.
- [6] M. Hayashi, "Optimal sequence of quantum measurements in the sense of Stein's lemma in quantum hypothesis testing," *J. Phys. A: Math. Gen.*, vol. 35, no. 50, pp. 10759–10773, Dec. 2002.
- [7] K. M. R. Audenaert, M. Nussbaum, A. Szkola, and F. Verstraete, "Asymptotic error rates in quantum hypothesis testing," *Commun. Math. Phys.*, vol. 279, no. 1, pp. 251–283, Apr. 2008.
- [8] M. Hayashi, "Error exponent in asymmetric quantum hypothesis testing and its application to classical-quantum channel coding," *Phys. Rev. A, Gen. Phys.*, vol. 76, no. 6, Dec. 2007, Art. no. 062301.
- [9] F. G. S. L. Brandão and M. B. Plenio, "A generalization of quantum Stein's lemma," *Commun. Math. Phys.*, vol. 295, no. 3, pp. 791–828, May 2010.
- [10] T. Cooney, M. Mosonyi, and M. M. Wilde, "Strong converse exponents for a quantum channel discrimination problem and quantum-feedback-assisted communication," *Commun. Math. Phys.*, vol. 344, no. 3, pp. 797–829, Jun. 2016.
- [11] M. Mosonyi and T. Ogawa, "Quantum hypothesis testing and the operational interpretation of the quantum Rényi relative entropies," *Commun. Math. Phys.*, vol. 334, no. 3, pp. 1617–1648, Mar. 2015.
- [12] X. Wang and M. M. Wilde, "Resource theory of asymmetric distinguishability," *Phys. Rev. Res.*, vol. 1, no. 3, Dec. 2019, Art. no. 033170.
- [13] X. Wang and M. M. Wilde, "Resource theory of asymmetric distinguishability for quantum channels," *Phys. Rev. Res.*, vol. 1, no. 3, Dec. 2019, Art. no. 033169.
- [14] K. Fang, H. Fawzi, and O. Fawzi, "Adversarial quantum channel discrimination," *Phys. Rev. Lett.*, vol. 135, no. 26, Dec. 2025, Art. no. 260201.
- [15] K. Fang, H. Fawzi, and O. Fawzi, "Generalized quantum asymptotic equipartition," 2024, *arXiv:2411.04035*.
- [16] M. Hayashi and Y. Ito, "Entanglement measures for detectability," *IEEE Trans. Inf. Theory*, vol. 71, no. 6, pp. 4385–4405, Jun. 2025.
- [17] T. Ogawa and M. Hayashi, "On error exponents in quantum hypothesis testing," *IEEE Trans. Inf. Theory*, vol. 50, no. 6, pp. 1368–1372, Jun. 2004.
- [18] H. Nagaoka, "The converse part of the theorem for quantum Hoeffding bound," 2006, *arXiv: quant-ph/0611289*.
- [19] M. Berta, F. G. S. L. Brandão, and C. Hirche, "On composite quantum hypothesis testing," *Commun. Math. Phys.*, vol. 385, no. 1, pp. 55–77, Jul. 2021.
- [20] M. Mosonyi, Z. Szilagy, and M. Weiner, "On the error exponents of binary state discrimination with composite hypotheses," *IEEE Trans. Inf. Theory*, vol. 68, no. 2, pp. 1032–1067, Feb. 2022.
- [21] M. Hayashi and M. Tomamichel, "Correlation detection and an operational interpretation of the Rényi mutual information," *J. Math. Phys.*, vol. 57, no. 10, 2016, Art. no. 102201.
- [22] A. Dasgupta, N. A. Warsi, and M. Hayashi, "Universal tester for multiple independence testing and classical-quantum arbitrarily varying multiple access channel," *IEEE Trans. Inf. Theory*, vol. 71, no. 5, pp. 3719–3765, May 2025.
- [23] K. Watanabe and R. Takagi, "Black box work extraction and composite hypothesis testing," *Phys. Rev. Lett.*, vol. 133, no. 25, Dec. 2024, Art. no. 250401.
- [24] F. Hiai, M. Mosonyi, and T. Ogawa, "Large deviations and Chernoff bound for certain correlated states on a spin chain," *J. Math. Phys.*, vol. 48, no. 12, Dec. 2007, Art. no. 123301.
- [25] F. Hiai, M. Mosonyi, and T. Ogawa, "Error exponents in hypothesis testing for correlated states on a spin chain," *J. Math. Phys.*, vol. 49, no. 3, Mar. 2008, Art. no. 032112.
- [26] M. Mosonyi and T. Ogawa, "Two approaches to obtain the strong converse exponent of quantum hypothesis testing for general sequences of quantum states," *IEEE Trans. Inf. Theory*, vol. 61, no. 12, pp. 6975–6994, Dec. 2015.
- [27] M. Hayashi and H. Yamasaki, "The generalized quantum Stein's lemma and the second law of quantum resource theories," *Nature Phys.*, vol. 21, no. 12, pp. 1988–1993, Dec. 2025.
- [28] L. Lami, "A solution of the generalized quantum Stein's lemma," *IEEE Trans. Inf. Theory*, vol. 71, no. 6, pp. 4454–4484, Jun. 2025.
- [29] M. M. Wilde, M. Berta, C. Hirche, and E. Kaur, "Amortized channel divergence for asymptotic quantum channel discrimination," *Lett. Math. Phys.*, vol. 110, no. 8, pp. 2277–2336, Aug. 2020.
- [30] H. Umegaki, "Conditional expectation in an operator algebra," *Tohoku Math. J., 2nd Ser.*, vol. 6, nos. 2–3, pp. 177–181, 1954.
- [31] D. Petz, "Quasi-entropies for finite quantum systems," *Rep. Math. Phys.*, vol. 23, no. 1, pp. 57–65, Feb. 1986.
- [32] M. Müller-Lennert, F. Dupuis, O. Szechr, S. Fehr, and M. Tomamichel, "On quantum Rényi entropies: A new generalization and some properties," *J. Math. Phys.*, vol. 54, no. 12, Dec. 2013, Art. no. 122203.
- [33] M. M. Wilde, A. Winter, and D. Yang, "Strong converse for the classical capacity of entanglement-breaking and Hadamard channels via a sandwiched Rényi relative entropy," *Commun. Math. Phys.*, vol. 331, no. 2, pp. 593–622, Oct. 2014.
- [34] S. Ben-David and E. Blais, "A new minimax theorem for randomized algorithms," *J. ACM*, vol. 70, no. 6, pp. 1–58, Dec. 2023.
- [35] M. Mosonyi and F. Hiai, "Some continuity properties of quantum Rényi divergences," *IEEE Trans. Inf. Theory*, vol. 70, no. 4, pp. 2674–2700, Apr. 2024.
- [36] K. He, J. Saunderson, and H. Fawzi, "QICS: Quantum information conic solver," 2024, *arXiv:2410.17803*.
- [37] K. M. R. Audenaert et al., "Discriminating states: The quantum Chernoff bound," *Phys. Rev. Lett.*, vol. 98, no. 16, Apr. 2007, Art. no. 160501.
- [38] M. Nussbaum and A. Szkola, "The Chernoff lower bound for symmetric quantum hypothesis testing," *Ann. Statist.*, vol. 37, no. 2, pp. 1040–1057, Apr. 2009.
- [39] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*. Berlin, Germany: Springer, 2010.
- [40] R. T. Rockafellar, *Convex Analysis*. Princeton, NJ, USA: Princeton Univ. Press, 1970.
- [41] M. Hayashi, "Quantum information theory," in *Graduate Texts in Physics*. Berlin, Germany: Springer, 2017.
- [42] K. Fang, G. Gour, and X. Wang, "Towards the ultimate limits of quantum channel discrimination and quantum communication," *Sci. China Inf. Sci.*, vol. 68, no. 8, Aug. 2025, Art. no. 180509.
- [43] M. Hayashi, "General detectability measure," *Commun. Math. Phys.*, vol. 406, no. 12, p. 289, Dec. 2025.
- [44] L. Lami, M. Berta, and B. Regula, "Asymptotic quantification of entanglement with a single copy," *Nature Phys.*, vol. 22, no. 3, pp. 439–445, Mar. 2026.
- [45] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, vol. 1. Hoboken, NJ, USA: Wiley, 2006.
- [46] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*, vol. 16, 3rd ed., Amsterdam, The Netherlands: North Holland, 1981.
- [47] M. Christandl, N. Schuch, and A. Winter, "Entanglement of the anti-symmetric state," *Commun. Math. Phys.*, vol. 311, no. 2, pp. 397–422, Apr. 2012.

- [48] F. Leditzky, D. Leung, V. Siddhu, G. Smith, and J. A. Smolin, "The platypus of the quantum channel zoo," *IEEE Trans. Inf. Theory*, vol. 69, no. 6, pp. 3825–3849, Jun. 2023.
- [49] F. G. S. L. Brandão and G. Gour, "Reversible framework for quantum resource theories," *Phys. Rev. Lett.*, vol. 115, no. 7, Aug. 2015, Art. no. 070503.
- [50] E. Chitambar and G. Gour, "Quantum resource theories," *Rev. Modern Phys.*, vol. 91, no. 2, 2019, Art. no. 025001.
- [51] K. Fang, X. Wang, M. Tomamichel, and R. Duan, "Non-asymptotic entanglement distillation," *IEEE Trans. Inf. Theory*, vol. 65, no. 10, pp. 6454–6465, Oct. 2019.
- [52] B. Regula, K. Fang, X. Wang, and M. Gu, "One-shot entanglement distillation beyond local operations and classical communication," *New J. Phys.*, vol. 21, no. 10, Oct. 2019, Art. no. 103017.
- [53] J. Theurer, K. Fang, and G. Gour, "Single-shot entanglement manipulation of states and channels revisited," *IEEE Trans. Inf. Theory*, vol. 71, no. 7, pp. 5330–5360, May 2023, doi: [10.1109/TIT.2023.3570770](https://doi.org/10.1109/TIT.2023.3570770).
- [54] A. Winter and D. Yang, "Operational resource theory of coherence," *Phys. Rev. Lett.*, vol. 116, no. 12, Mar. 2016, Art. no. 120404.
- [55] M. G. Díaz et al., "Using and reusing coherence to realize quantum processes," *Quantum*, vol. 2, p. 100, Oct. 2018.
- [56] K. Fang, X. Wang, L. Lami, B. Regula, and G. Adesso, "Probabilistic distillation of quantum coherence," *Phys. Rev. Lett.*, vol. 121, no. 7, Aug. 2018, Art. no. 070404.
- [57] B. Regula, K. Fang, X. Wang, and G. Adesso, "One-shot coherence distillation," *Phys. Rev. Lett.*, vol. 121, no. 1, Jul. 2018, Art. no. 010401.
- [58] M. Hayashi, K. Fang, and K. Wang, "Finite block length analysis on quantum coherence distillation and incoherent randomness extraction," *IEEE Trans. Inf. Theory*, vol. 67, no. 6, pp. 3926–3944, Jun. 2021.
- [59] I. Marvian and R. W. Spekkens, "The theory of manipulations of pure state asymmetry: I. basic tools, equivalence classes and single copy transformations," *New J. Phys.*, vol. 15, no. 3, Mar. 2013, Art. no. 033001.
- [60] F. Brandão, M. Horodecki, N. Ng, J. Oppenheim, and S. Wehner, "The second laws of quantum thermodynamics," *Proc. Nat. Acad. Sci. USA*, vol. 112, no. 11, pp. 3275–3279, 2015.
- [61] M. Horodecki, P. Horodecki, and J. Oppenheim, "Reversible transformations from pure to mixed states and the unique measure of information," *Phys. Rev. A, Gen. Phys.*, vol. 67, no. 6, Jun. 2003, Art. no. 062104.
- [62] S. Bravyi and A. Kitaev, "Universal quantum computation with ideal Clifford gates and noisy ancillas," *Phys. Rev. A, Gen. Phys.*, vol. 71, no. 2, Feb. 2005, Art. no. 022316.
- [63] K. Fang and Z.-W. Liu, "No-go theorems for quantum resource purification," *Phys. Rev. Lett.*, vol. 125, no. 6, Aug. 2020, Art. no. 060405.
- [64] K. Fang and Z.-W. Liu, "No-go theorems for quantum resource purification: New approach and channel theory," *PRX Quantum*, vol. 3, no. 1, Mar. 2022, Art. no. 010337.
- [65] K. Fang and Z.-W. Liu, "One-shot distillation with constant overhead using catalysts," 2024, *arXiv:2410.14547*.
- [66] H. Zhu, M. Hayashi, and L. Chen, "Coherence and entanglement measures based on Rényi relative entropies," *J. Phys. A, Math. Theor.*, vol. 50, no. 47, Nov. 2017, Art. no. 475303.
- [67] E. Chitambar and G. Gour, "Comparison of incoherent operations and measures of coherence," *Phys. Rev. A, Gen. Phys.*, vol. 94, no. 5, Nov. 2016, Art. no. 052336.
- [68] B. Farkas and S. G. Révész, "Potential theoretic approach to rendezvous numbers," *Monatshefte für Math.*, vol. 148, no. 4, pp. 309–331, Aug. 2006.

Kun Fang (Member, IEEE) received the B.S. degree in mathematics from Wuhan University in 2015 and the Ph.D. degree in quantum information from the University of Technology Sydney in 2018. After that, he was a Post-Doctoral Researcher at the University of Cambridge and the University of Waterloo from 2018 to 2020. He served as a Senior Researcher and a Tech Lead at the Institute for Quantum Computing, Baidu, from 2020 to 2023. He is currently an Assistant Professor with the School of Data Science, The Chinese University of Hong Kong, Shenzhen. His research interests lie in understanding the capabilities and limitations of quantum resources, and their usage in quantum computation and quantum communications.

Masahito Hayashi (Fellow, IEEE) was born in Japan, in 1971. He received the B.S. degree from the Faculty of Sciences, Kyoto University, Japan, in 1994, and the M.S. and Ph.D. degrees in mathematics from Kyoto University in 1996 and 1999, respectively. He was with Kyoto University as a Research Fellow of Japan Society of the Promotion of Science (JSPS) from 1998 to 2000 and the Laboratory for Mathematical Neuroscience, Brain Science Institute, RIKEN, as a Researcher, from 2000 to 2003. He was with the ERATO Quantum Computation and Information Project, Japan Science and Technology Agency (JST), as the Research Head, from 2003 to 2006, and the ERATO-SORST Quantum Computation and Information Project, JST, as the Group Leader, from 2006 to 2007. He was also with the Superrobust Computation Project Information Science and Technology Strategic Core (21st Century COE by MEXT) Graduate School of Information Science and Technology, The University of Tokyo, as an Adjunct Associate Professor, from 2004 to 2007. He was with the Graduate School of Information Sciences, Tohoku University, as an Associate Professor, from 2007 to 2012. In 2012, he joined the Graduate School of Mathematics, Nagoya University, as a Full Professor. He was with Shenzhen Institute for Quantum Science and Engineering, Southern University of Science and Technology, Shenzhen, China, as the Chief Research Scientist, from 2020 to 2023. In 2023, he joined the School of Data Science, The Chinese University of Hong Kong, Shenzhen (CUHK-Shenzhen), as a Full Professor, and joined the International Quantum Academy (SIQA) as the Chief Research Scientist. He also has the title of Presidential Chair Professor with CUHK-Shenzhen in 2023. Also, he was with the Centre for Quantum Technologies, National University of Singapore, as a Visiting Research Associate Professor from 2009 to 2012, and a Visiting Research Professor from 2012 to 2024. He was with the Center for Advanced Intelligence Project, RIKEN, as a Visiting Scientist, from 2017 to 2020. He was with Shenzhen Institute for Quantum Science and Engineering, Southern University of Science and Technology, as a Visiting Professor, from 2018 to 2020, and the Center for Quantum Computing, Peng Cheng Laboratory, Shenzhen, as a Visiting Professor, from 2019 to 2020. He has been with the Department of Information Engineering, CUHK-Shenzhen, as an Adjunct Professor, since 2024. In 2006, he published the book *Quantum Information: An Introduction* (Springer), whose revised version was published as *Quantum Information Theory: Mathematical Foundation* from Graduate Texts in Physics (Springer) in 2016. In 2016, he published other two books *Group Representation for Quantum Theory* and *A Group Theoretic Approach to Quantum Information* (Springer). His research interests include classical and quantum information theory and classical and quantum statistical inference. In 2022, he was elected as an IMS Fellow and an Asia-Pacific Artificial Intelligence Association (AAPAI) Fellow. He received the Information Theory Society Paper Award for Information-Spectrum Approach to Second Order Coding Rate in Channel Coding in 2011. In 2016, he received Japan Academy Medal from Japan Academy and JSPS Prize from Japan Society for the Promotion of Science. He is on the Editorial Board of *New Journal of Physics* and *International Journal of Quantum Information*.