Discrimination of quantum states under locality constraints in the many-copy setting

Hao-Chung Cheng[1], Andreas Winter[2,3], and Nengkun Yu[4]

[1] *Department of Electrical Engineering & Graduate Institute of Communication Engineering, National Taiwan University, Tapei 106, Taiwan (R.O.C.)*
[2] *ICREA—Institució Catalana de Recerca i Estudis Avançats, 08010 Barcelona, Spain*
[3] *Física Teòrica: Informaciò i Fenòmens Quàntics, Departament de Física, Universitat Autònoma de Barcelona, 08193 Barcelona, Spain*
[4] *Centre for Quantum Software and Information & Faculty of Engineering and Information Technology, University of Technology Sydney, Ultimo NSW 2007, Australia*

Abstract. We study the discrimination of a pair of orthogonal quantum states in the many-copy setting. This is not a problem when arbitrary quantum measurements are allowed, as then the states can be distinguished perfectly even with one copy. However, it becomes highly nontrivial when we consider states of a multipartite system and locality constraints are imposed. We hence focus on the restricted families of measurements such as local operation and classical communication (LOCC), separable operations (SEP), and the positive-partial-transpose operations (PPT) in this paper.

We first study asymptotic discrimination of an arbitrary multipartite entangled pure state against its orthogonal complement using LOCC/SEP/PPT measurements. We prove that the incurred optimal average error probability always decays exponentially in the number of copies, by proving upper and lower bounds on the exponent. In the special case of discriminating a maximally entangled state against its orthogonal complement, we determine the explicit expression for the optimal average error probability and the optimal trade-off between the type-I and type-II errors, thus establishing the associated Chernoff, Stein, Hoeffding, and the strong converse exponents. Our technique is based on the idea of using PPT operations to approximate LOCC.

Then, we show an infinite separation between SEP and PPT operations by providing a pair of states constructed from an unextendible product basis (UPB): they can be distinguished perfectly by PPT measurements, while the optimal error probability using SEP measurements admits an exponential lower bound. On the technical side, we prove this result by providing a quantitative version of the well-known statement that the tensor product of UPBs is UPB.

1. Introduction

Testing whether a system has a specified property of interest is a fundamental problem in science. In statistics, this problem is called *hypothesis testing* [1], which has substantial applications in numerous fields, such as information sciences [2, 3, 4, 5, 6, 7, 8], computational learning theory [9, 10, 11, 12], property and distribution testing [13, 14, 15, 16], and differential privacy [17, 18, 19, 20, 21].

The most basic form of hypothesis testing is binary hypothesis testing, which consists of a null hypothesis H_0 and an alternative hypothesis H_1. In quantum computing, the two hypotheses are modeled by quantum states ρ_0 and ρ_1, respectively. To distinguish the two quantum states, one has to perform a test T, or equivalently a two-outcome positive-operator valued measure (POVM) measurement $\{T, \mathbb{1} - T\}$ on the received state, where $0 \leq T \leq \mathbb{1}$ is a quantum observable. Such a test T incurs two types of error probabilities; namely, the type-I error $\alpha(T) := \Pr[\mathsf{H}_1|\mathsf{H}_0] = \operatorname{Tr}[\rho_0(\mathbb{1} - T)]$ is the probability of accepting the H_1 when H_0 is true, and the type-II error $\beta(T) := \Pr[\mathsf{H}_0|\mathsf{H}_1] = \operatorname{Tr}[\rho_1 T]$ is to decide for H_0 when actually H_1 is true. If the prior probabilities of the hypotheses are known, say p and $1-p$, we measure the performance of the decision scheme by calculating the average (Bayes) error probability. We term this the Bayesian approach, and specifically the symmetric setting when the two hypotheses are equally likely,

Date: 20 November 2020.

$p = 1 - p = \frac{1}{2}$. In most practical situations where the prior probabilities are unknown, a Neyman-Pearson approach is to analyze the trade-off between the two types of errors. We call this the asymmetric setting.

In the *Bayesian* setting, the optimal average error probability is then given by

$$P_{\mathrm{e}}(\rho_0, \rho_1; p) := \inf_T \{p\alpha(T) + (1-p)\beta(T)\}$$

for $p := \Pr[\mathsf{H}_0] \in (0;1)$. Helstrom and Hoelvo [22] proved a closed-form expression of P_{e} and showed that the optimal test is achieved by projection onto the positive support of $p\rho_0 - (1-p)\rho_1$. This measurement can be viewed as the quantum generalization of the classical *Neyman–Pearson test* [23, 24, 25, 26].

In the present paper we are concerned with the many-copy and asymptotic behavior of $P_{\mathrm{e}}(\rho_0^{\otimes n}, \rho_1^{\otimes n}; p)$, where n identical copies of states are prepared. The celebrated quantum Chernoff theorem [27, 28, 29] then establishes that

$$\xi_{\mathrm{C}}(\rho_0, \rho_1) := \lim_{n\to\infty} -\frac{1}{n} \log P_{\mathrm{e}}(\rho_0^{\otimes n}, \rho_1^{\otimes n}; p) = -\min_{0\le s\le 1} \log \operatorname{Tr}\left[\rho_0^{1-s}\rho_1^s\right]. \tag{1}$$

That is, the *Chernoff exponent* $\xi_{\mathrm{C}}(\rho_0, \rho_1)$ determines the convergence rate of the error probability. Because of this result, in the asymptotics we consider all the Bayesian error probabilities $P_{\mathrm{e}}(\rho_0^{\otimes n}, \rho_1^{\otimes n}; p)$ as pertaining to the symmetric setting, as long as p does not depend on n.

In the *asymmetric* setting, one aims to study the asymptotic functional dependence between the type-I and type-II errors. In particular, three exponents are the most important (the detailed definitions will be given in Section 2). The *Stein exponent* characterizes the best (i.e. the largest) exponential decaying rate of they type-II error β_n when α_n is bounded by some $\varepsilon \in (0,1)$. The quantum Stein's lemma [30, 31] shows that the exponent is given by the quantum relative entropy [32]. Moreover, it is independent of $\varepsilon \in (0,1)$, which is called the strong converse property. The *Hoeffding exponent* and the *strong converse exponent*, respectively, determine the optimal exponential rate of the type-I error or the type-I success probability when the type-II error exponentially decays at the rate below or above the Stein exponent. They are proved to be given by quantities involving Petz's Rényi divergence [33, 34, 28, 29] and the sandwiched Rényi divergence [35, 36, 37, 38]. Other extensions in the large, moderate, and small deviation regimes have been studied in depth (see e.g. [39, 40, 41, 42, 43, 44, 6, 45, 7, 8]).

Small to immediate scale quantum computers will be available in the near-term future [46, 47]. However, such quantum computers will be built in geographically separated laboratories, which means that only local quantum operations in each lab and mutual classical communications may be available. Those operations constitute a restricted class of measurements and they are termed *local operations and classical communication* (LOCC) [48, 49, 50, 51]. It is thus natural to ask how well LOCC measurements perform in hypothesis testing in comparison to global measurements. For instance, the above-mentioned quantum Neyman–Pearson test is generally not implementable via LOCC due to quantum entanglement and nonlocality [52, 50]. This problem of local discrimination and local hypothesis testing thus gained considerable attention in quantum computation and quantum information recently [53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67]. Unfortunately, only limited results are known due to the complicated mathematical structure of LOCC.

In the present work, our goal is to study the asymptotic behavior of the errors incurred by the restriction to LOCC, and derive the above-mentioned four exponents. We also consider other classes of measurements: the positive-partial-transpose operations (PPT) and separable operations (SEP) [68], mainly out of theoretical interest or as a tool to bound LOCC. It is well-known that strict inclusions hold among them [52], i.e.

$$\text{LOCC} \subset \text{SEP} \subset \text{PPT} \subset \text{ALL}. \tag{2}$$

Although hypothesis testing under LOCC has been studied in manyf papers [53, 54, 55, 56, 57, 58, 59, 60, 62, 63, 65], our results, for the first time, demonstrate an interesting phenomenon in the asymptotic error behavior, which shows that distinguishing a pair of orthogonal states under LOCC indeed exhibits a fundamental difference from the conventional task using global measurements. By definition of the Chernoff exponent given in (1) and the inclusions (2), we have

$$\xi_{\mathrm{C}}^{\mathrm{LOCC}} \le \xi_{\mathrm{C}}^{\mathrm{SEP}} \le \xi_{\mathrm{C}}^{\mathrm{PPT}} \le \xi_{\mathrm{C}}^{\mathrm{ALL}} = -\min_{0\le s\le 1} \log \operatorname{Tr}\left[\rho_0^{1-s}\rho_1^s\right]. \tag{3}$$

Here and subsequently, we put a superscript X on P_{e} and ξ to highlight the class X of measurements allowed. Particularly intriguing examples arising in the context of data hiding [69, 70, 53, 71, 72, 73] are that the underlying states are orthogonal, i.e. $\operatorname{Tr}[\rho_0\rho_1] = 0$. This implies that $\xi_{\mathrm{C}}^{\mathrm{ALL}}(\rho_0, \rho_1) = \infty$. However, whether $\xi_{\mathrm{C}}^{\mathrm{LOCC}}(\rho_0, \rho_1)$ is finite or not for a given pair of orthogonal states has remained open since the early days of quantum information theory. At least, we know it is nonzero, which can be achieved through local tomography and then using the classical Chernoff bound. On the other hand, it is unknown if the definition of the Chernoff exponent with respect to a class $\mathrm{X} \in \{\mathrm{LOCC}, \mathrm{SEP}, \mathrm{PPT}\}$ is faithful in the sense that the sequence $\left(-\frac{1}{n}\log P_{\mathrm{e}}^{\mathrm{X}}(\rho_0^{\otimes n}, \rho_1^{\otimes n}; p)\right)_n$ diverges if and only if $P_{\mathrm{e}}^{\mathrm{X}}(\rho_0^{\otimes n}, \rho_1^{\otimes n}; p) = 0$ for some finite n; cf. [64]. Hence, in this work, we study the case of distinguishing an entangled state (possibly on a multipartite system) and its orthogonal complement using restricted classes of POVMs. Perhaps surprisingly, we show that the Chernoff bound in this case is indeed faithful; we remark that no simple expression for a general multipartite entangled state is known.

Although it has been shown that the inclusion relations (2) are all strict [52], it is not known whether strict inclusion still holds in the many-copy asymptotics. For instance, does any equality hold in (3)? This question arises naturally since SEP and PPT operations are often exploited to approximate the LOCC operations. Hence, one may wonder how differently the restricted classes of measurements perform. Here, we demonstrate that there is a separation between the SEP and PPT operations. Namely, we construct a pair of states such that $\xi_{\mathrm{C}}^{\mathrm{PPT}} = \infty$, while $\xi_{\mathrm{C}}^{\mathrm{SEP}} \leq -\log\mu < \infty$ for some $\mu > 0$.

1.1. **Main Contributions.** In the following, we state our contributions.

(I) We study the optimal error probability of distinguishing an arbitrary multipartite entangled pure state ψ and its orthogonal complement $\psi^\perp := \frac{\mathbb{1}-\psi}{D-1}$ (where D is the dimension of the underlying Hilbert space) in the many-copy scenario.

(i) For ψ being an arbitrary multipartite entangled pure state, we prove that the optimal error probability decays exponentially in the number of copies n (see Theorem 1 of Section 3): there exists constants $0 < a \leq b < \infty$ such that,

$$\mathrm{e}^{-nb} \leq P_{\mathrm{e}}^{\mathrm{LOCC}}\left(\psi^{\otimes n}, (\psi^\perp)^{\otimes n}; p\right) \leq \mathrm{e}^{-na}, \quad \forall n \in \mathbb{N}.$$

Hence, the Chernoff bound is faithful.

Our key technique is to prove this result is to establish an exponential lower bound to the optimal error probability of distinguishing bipartite entangled state ψ on a bipartite system $\mathbb{C}^d \otimes \mathbb{C}^d$ against its orthogonal complement $\psi^\perp$ using PPT POVMs (Proposition 3 of Section 3.1):

$$P_{\mathrm{e}}^{\mathrm{PPT}}\left(\psi^{\otimes n}, (\psi^\perp)^{\otimes n}; p\right) \geq \min\{p, 1-p\} \cdot t^n, \quad \forall n \in \mathbb{N}$$

where $t := \frac{1-\eta}{(d^2-1)\eta} \in (0, \frac{1}{d+1})$ and η denotes the largest squared Schmidt coefficient of ψ.

The above bound also provides a lower bound to the error probability using LOCC. In the case that ψ is a bipartite entangled pure state, we establish a universal upper bound to the optimal error probability using LOCC POVMs (Proposition 2 of Section3):

$$P_{\mathrm{e}}^{\mathrm{LOCC}}\left(\psi^{\otimes n}, (\psi^\perp)^{\otimes n}; p\right) \leq (1-p)\left(\frac{1}{d+1}\right)^n, \quad n \in \mathbb{N}.$$

Moreover, our approach extends to a strong converse bound in the asymmetric setting—the type-I error approaches 1 at the exponential rate $r - \log\frac{1}{t}$ whenever the type-II error decays exponentially at rate r (Proposition 4 of Section 3.2), which also implies that the Stein exponent is upper bounded by the quantity $\log\frac{1}{t}$.

(ii) In the special case where the underlying pair of states are maximally entangled state $\Phi_d := \frac{1}{d}\sum_{i,j=0}^{d-1}|ii\rangle\langle jj|$ on $\mathbb{C}^d \times \mathbb{C}^d$ and its orthogonal complement state $\Phi_d^\perp := (\mathbb{1} - \Phi_d)/(d^2-1)$, we explicitly calculate the optimal average error probability as (Theorem 8 of Section 4.1):

$$P_{\mathrm{e}}^{\mathrm{LOCC}}(\Phi_d^{\otimes n}, (\Phi_d^\perp)^{\otimes n}; p) = \min\left\{(1-p)\left(\frac{1}{d+1}\right)^n, p\right\}.$$

This then immediately gives the Chernoff exponent $\xi_{\mathrm{C}}^{\mathrm{LOCC}}(\Phi_d, \Phi_d^{\perp}) = \log(d+1)$. Further, we establish the optimal trade-off between the type-I and type-II errors (Theorem 10 of Section 4.2), i.e. given any type-I error α_n, the minimum type-II error satisfies

$$\beta_n = \frac{(1-\alpha_n)}{(d+1)^n}. \tag{4}$$

With such functional dependence, we then obtain the associated Stein, Hoeffding, and strong converse exponents, respectively (Corollaries 11, 12, and 13 of Section 4.2).

(iii) Our results apply to case of testing the pure state with uniform non-zero Schmidt coefficients (i.e. $\frac{1}{m}\sum_{i,j=0}^{m-1} |ii\rangle\langle jj|$ for $m \leq d$) and its orthogonal complement (Propositions 15 and 16 of Section 5). Lastly, we extend Matthews and Winter's work [53] of testing completely symmetric state against completely anti-symmetric state [74] to prove the corresponding Stein, Hoeffding, and strong converse exponents (Propositions 17 and 18). In Table 1 below, we summarize the established exponents in various setting of testing.

(II) In Theorem 20 of Section 6, we prove an *infinite separation* between SEP measurements and PPT measurements. Specifically, we consider the null hypothesis to be the uniform mixture of an unextendible product basis (UPB) [76], and the alternative hypothesis to be a state supported on the orthogonal complement of the former. Such a pair of states can be discriminated perfectly by a PPT measurement, while we show that the optimal error probability under SEP measurements possesses an exponential lower bound (Theorem 20 of Section 6). Our key technique to establish this result is the introduction of a novel quantity (Definition 21 of Section 6) that measures how far a UPB is from being an extendible product basis, and its multiplicativity property under the tensor product (Proposition 22). Our result hence gives a quantitative characterization of the property that the tensor product of UPBs is still a UPB [77].

Organization of the paper. Section 2 introduces necessary notation and definitions. In Section 3, we test arbitrary entangled pure state and its orthogonal complement. Section 4 is devoted to a special case of testing a maximally entangled state and its orthogonal complement. In Section 5, we consider the case of testing pure state with equal non-zero Schmidt coefficients. In Section 6, we demonstrate an infinite separation for the SEP and PPT operations. We provide discussions and conclusions of this paper in Section 7.

2. Notation and Definitions

Let $\mathbb{C}^d$ be a d-dimensional complex Hilbert space. A quantum state (i.e. density operator) on $\mathbb{C}^d$ is a positive semi-definite operator with unit trace. The trace operation is denoted by $\mathrm{Tr}[\,\cdot\,]$. We use $\mathbb{1}_d$ to stand for the identity operator on $\mathbb{C}^d$, and $\mathbb{0}_d$ denotes the zero operation on $\mathbb{C}^d$. If no confusion is possible, we will drop the subscript for simplicity. A positive-operator valued measure (POVM) is a set of positive semi-definite operators that sum to identity. For any density operator ρ on $\mathbb{C}^d$, we denote by $\rho^{\perp}$ the density operator on $\mathbb{C}^d$ (if it exists and is unique) which is orthogonal to ρ. For example, if ρ is pure (i.e. a rank-one projection), then $\rho^{\perp} = (\mathbb{1}_d - \rho)/(d-1)$. We call such a state $\rho^{\perp}$ the *orthogonal compliment* of ρ. The maximally entangled state on $\mathbb{C}^d \otimes \mathbb{C}^d$ is denoted by $\Phi_d := \frac{1}{d}\sum_{i,j=0}^{d-1} |ii\rangle\langle jj|$, and $\Phi_d^{\perp}$ is its orthogonal complement. Let σ_d and $\sigma_d^{\perp}$ denote the completely symmetric and completely anti-symmetric states on $\mathbb{C}^d \otimes \mathbb{C}^d$ [74]. We use $\tau_d := \mathbb{1}_d/d$ to denote the completely mixed state on $\mathbb{C}^d$. The operations $\otimes$ and $\oplus$ mean the tensor product and direct sum, respectively. We use boldface to denote a column vector, e.g. $\mathbf{x} = (x_0 \dots x_n)^T$ with T being the transpose. We use $\mathbb{N}$ to denote natural numbers. For an operator X on a bipartite Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$, we use Γ to denote its partial transpose with respect to $\mathcal{H}_B$, i.e. for some orthonormal basis $\{|i\rangle_A \otimes |j\rangle_B\}_{(i,j)}$, we defined the partial transpose to be

$$\left(|i\rangle_A \otimes |j\rangle_B \langle k|_A \otimes \langle \ell|_B\right)^{\Gamma} := |i\rangle_A \otimes |\ell\rangle_B \langle k|_A \otimes \langle j|_B.$$

Setting \ Exponents	Chernoff	Stein	Hoeffding	Strong converse
$\begin{cases}\mathsf{H}_0 : \Phi_d^{\otimes n} \\ \mathsf{H}_1 : \left(\Phi_d^{\perp}\right)^{\otimes n}\end{cases}$	$\log(d+1)$	$\log(d+1)$	∞	$r - \log(d+1)$
$\begin{cases}\mathsf{H}_0 : \left(\Phi_d^{\perp}\right)^{\otimes n} \\ \mathsf{H}_1 : \Phi_d^{\otimes n}\end{cases}$		∞	$\log(d+1)$	0
$\begin{cases}\mathsf{H}_0 : \sigma_d^{\otimes n} \\ \mathsf{H}_1 : \left(\sigma_d^{\perp}\right)^{\otimes n}\end{cases}$	$\log \dfrac{d+1}{d-1}$ [53]	∞	$\log \dfrac{d+1}{d-1}$	0
$\begin{cases}\mathsf{H}_0 : \left(\sigma_d^{\perp}\right)^{\otimes n} \\ \mathsf{H}_1 : \sigma_d^{\otimes n}\end{cases}$		$\log \dfrac{d+1}{d-1}$	∞	$r - \log \dfrac{d+1}{d-1}$
$\begin{cases}\mathsf{H}_0 : \left(\Phi_m \oplus \mathbb{0}\right)^{\otimes n} \\ \mathsf{H}_1 : \left(\lambda \Phi_m^{\perp} \oplus (1-\lambda)\tau\right)^{\otimes n}\end{cases}$	$\log \dfrac{m+1}{\lambda}$	$\log \dfrac{m+1}{\lambda}$	∞	$r - \log \dfrac{m+1}{\lambda}$
$\begin{cases}\mathsf{H}_0 : \left(\Phi_m^{\perp} \oplus \mathbb{0}\right)^{\otimes n} \\ \mathsf{H}_1 : \left(\lambda \Phi_m \oplus (1-\lambda)\tau\right)^{\otimes n}\end{cases}$	$\max\left\{\log(m+1), \log \dfrac{1}{\lambda}\right\}$	∞	$\begin{cases}\infty & r \leq \log \frac{1}{\lambda} \\ \log(m+1) & r > \log \frac{1}{\lambda}\end{cases}$	0
$\begin{cases}\mathsf{H}_0 : \left(\sigma_m \oplus \mathbb{0}\right)^{\otimes n} \\ \mathsf{H}_1 : \left(\lambda \sigma_m^{\perp} \oplus (1-\lambda)\tau\right)^{\otimes n}\end{cases}$	$\max\left\{\log \dfrac{m+1}{m-1}, \log \dfrac{1}{\lambda}\right\}$	∞	$\begin{cases}\infty & r \leq \log \frac{1}{\lambda} \\ \log \frac{m+1}{m-1} & r > \log \frac{1}{\lambda}\end{cases}$	0
$\begin{cases}\mathsf{H}_0 : \left(\sigma_m^{\perp} \oplus \mathbb{0}\right)^{\otimes n} \\ \mathsf{H}_1 : \left(\lambda \sigma_m \oplus (1-\lambda)\tau\right)^{\otimes n}\end{cases}$	$\log \dfrac{m+1}{\lambda(m-1)}$	$\log \dfrac{m+1}{\lambda(m-1)}$	∞	$r - \log \dfrac{m+1}{\lambda(m-1)}$

TABLE 1. This summarizes our main results of the Chernoff (6), Stein (9), Hoeffding (10), and the strong converse exponents (11) under various settings of binary hypothesis testing via LOCC, SEP, and PPT measurements. The states Φ_d and σ_d are the maximally entangled state and completely symmetric state on $\mathbb{C}^d \otimes \mathbb{C}^d$, and τ is a completely mixed state on another system (whose dimension is irrelevant here).

Given a multipartite system $\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_m$, a (one-way) LOCC POVM [51] is a decision rule based on all the measurement outcomes performed locally on each subsystem. A SEP operation is defined as

$$\mathrm{SEP} := \left\{ \left(E_k^{(1)} \otimes \cdots \otimes E_k^{(m)} \right) : E_k^{(j)} \geq 0, \; \sum_k E_k^{(1)} \otimes \cdots \otimes E_k^{(m)} = \mathbb{1} \right\},$$

and the PPT POVMs are defined as

$$\mathrm{PPT} := \left\{ (E_k)_k \text{ POVM} : \forall 1 \leq j \leq m, \forall k, (E_k)^{\Gamma} \geq 0, \text{ where } \Gamma \text{ is partial transpose on } \mathcal{H}_j. \right\}.$$

Consider a binary hypothesis testing problem as follows:

$$\begin{cases} \mathsf{H}_0 : \rho_0^{\otimes n} \\ \mathsf{H}_1 : \rho_1^{\otimes n} \end{cases}, \forall n \in \mathbb{N}.$$

Given a test T where $(T, \mathbb{1} - T)$ forms a two-outcome POVM, we let $\alpha_n(T) := \operatorname{Tr}\left[\rho^{\otimes n}(\mathbb{1} - T)\right]$ and $\beta_n(T) := \operatorname{Tr}\left[\sigma^{\otimes n} T\right]$. In the symmetric case with prior $0 < p < 1$, we define the optimal error probability of the binary hypothesis testing using the class X of POVMs as:

$$P_{\mathrm{e}}^{\mathrm{X}}(\rho_0^{\otimes n}, \rho_1^{\otimes n}; p) := \inf_{T, (\mathbb{1}-T) \in \mathrm{X}} \left\{ p\alpha(T_n) + (1-p)\beta(T_n) \right\}. \tag{5}$$

The associated Chernoff exponent is

$$\xi_{\mathrm{C}}^{\mathrm{X}}(\rho_0, \rho_1) := \lim_{n \to \infty} -\frac{1}{n} \log P_{\mathrm{e}}^{\mathrm{X}}(\rho_0^{\otimes n}, \rho_1^{\otimes n}; p). \tag{6}$$

We note that the definition given in (5) can be naturally extended to multiple hypothesis testing with priors.

Given density matrices ρ_0 and ρ_1, we define

$$\alpha_\beta^{\mathrm{X}}(\rho_0,\rho_1) := \inf\left\{\alpha : (\alpha,\beta) \in \mathcal{R}^{\mathrm{X}}(\rho_0,\rho_1)\right\}, \tag{7}$$

$$\beta_\alpha^{\mathrm{X}}(\rho_0,\rho_1) := \inf\left\{\beta : (\alpha,\beta) \in \mathcal{R}^{\mathrm{X}}(\rho_0,\rho_1)\right\}, \tag{8}$$

$$\mathcal{R}^{\mathrm{X}}(\rho_0,\rho_1) = \left\{(\alpha,\beta) : \exists T, \mathbb{1}-T \in \mathrm{X} : \alpha = \mathrm{Tr}\left[(\mathbb{1}-T)\rho_0\right],\ \beta = \mathrm{Tr}\left[T\rho_1\right]\right\}.$$

Here, $\mathcal{R}^{\mathrm{X}}(\rho_0,\rho_1)$ is sometimes called the *hypothesis testing region*, and the trade-off between type-I and type-II errors, $\beta_\alpha^{\mathrm{X}}(\rho_0,\rho_1)$, is termed as the *Neyman-Pearson function* or the *trade-off function* (see e.g. [1, Section 3.2], [75], [21, Definition 2.1]).

In the asymmetric case, we define the following exponent functions:

$$\xi_{\mathrm{S}}^{\mathrm{X}}(\rho_0,\rho_1,\varepsilon) := \lim_{n\to\infty} \sup_{T,(\mathbb{1}-T)\in \mathrm{X}} \left\{-\frac{1}{n}\log\beta_n(T) : \alpha_n(T) \le \varepsilon\right\}, \quad \forall \varepsilon \in (0,1); \tag{9}$$

$$\xi_{\mathrm{H}}^{\mathrm{X}}(\rho_0,\rho_1,r) := \lim_{n\to\infty} \sup_{T,(\mathbb{1}-T)\in \mathrm{X}} \left\{-\frac{1}{n}\log\alpha_n(T) : -\frac{1}{n}\log\beta_n(T) \ge r\right\}, \quad \forall r > 0; \tag{10}$$

$$\xi_{\mathrm{SC}}^{\mathrm{X}}(\rho_0,\rho_1,r) := \lim_{n\to\infty} \sup_{T,(\mathbb{1}-T)\in \mathrm{X}} \left\{-\frac{1}{n}\log(1-\alpha_n(T)) : -\frac{1}{n}\log\beta_n(T) \ge r\right\}, \quad \forall r > 0. \tag{11}$$

3. Testing Arbitrary Entangled Pure States

This section is devoted to proving that the Chernoff bound of testing an arbitrary multipartite entangled pure state against its orthogonal complement using LOCC POVMs is faithful. Namely, the associated optimal error probability exponentially decays in the number of copies (see Theorem 1 below).

The main ingredient in establishing this result is to show an exponential lower bound to the optimal error probability using PPT POVMs, which will be delayed to Proposition 3 of Section 3.1. Moreover, our technique also extends to showing a strong converse bound for the Stein exponent and a lower bound to the strong converse exponent in the asymmetric setting (Section 3.2).

Theorem 1 (Faithfulness of the Chernoff bound)**.** *Let ψ be an arbitrary multipartite entangled pure state and $\psi^\perp$ be its orthogonal complement state. The optimal error probability of discriminating them using LOCC POVMs decays exponentially in the number of copies, i.e. for any $0 < p < 1$ there exist $0 < a \le b < \infty$ such that,*

$$\mathrm{e}^{-nb} \le P_{\mathrm{e}}^{\mathrm{LOCC}}\left(\psi^{\otimes n}, (\psi^\perp)^{\otimes n}; p\right) \le \mathrm{e}^{-na}, \quad \forall n \in \mathbb{N}.$$

In other words,

$$0 < a \le \xi_{\mathrm{C}}^{\mathrm{LOCC}}\left(\psi,\psi^\perp\right) \le b < \infty.$$

Proof. The rightmost upper bound to $P_{\mathrm{e}}^{\mathrm{LOCC}}\left(\psi^{\otimes n}, (\psi^\perp)^{\otimes n}; p\right)$ can be proved as follows. Write P and Q as the probability distributions induced by taking an LOCC POVM on ψ and $\psi^\perp$ in each copy, respectively. Then, the classical Chernoff bound gives that $c = \min_{0\le s\le 1}\sum_i P(i)^{1-s}Q(i)^s$. This value c is one if and only if $P = Q$ for all possible LOCC POVMs. However, this is not the case since clearly $\psi \neq \psi^\perp$ and the product basis spans the whole Hilbert space. Hence, there exists an LOCC POVM such that $P \neq Q$.

For the converse, note that any multipartite system can be viewed as a bipartite system with an arbitrary cut through the whole system. Hence, LOCC POVMs on a multipartite system can be achieved by LOCC POVMs on the associated bipartite system, which then incurs smaller errors. Then, without loss of generality, it suffices to apply the exponential lower bound to the error probability using LOCC POVMs on a bipartite system, given in Proposition 3 of Section 3.1 below, to complete the proof. ☐

If ψ and $\psi^\perp$ live in a bipartite Hilbert space, we establish a universal upper bound (which only depends on the dimension of the underlying space) to the optimal average error probability using LOCC POVMs.

Proposition 2 (Achievability in the bipartite case)**.** *For any bipartite pure state ψ on $\mathbb{C}^d \otimes \mathbb{C}^d$, its orthogonal complement $\psi^\perp := \frac{\mathbb{1}-\psi}{d^2-1}$, and any $p \in (0,1)$, the error probability of testing multi-copies of them using LOCC POVMs is upper bounded by*

$$P_{\mathrm{e}}^{\mathrm{LOCC}}\left(\psi^{\otimes n}, (\psi^\perp)^{\otimes n}; p\right) \leq (1-p)\left(\frac{1}{d+1}\right)^n, \quad n \in \mathbb{N}.$$

Proof. Let $\sum_{i=0}^{d-1}\sqrt{\lambda_i}|ii\rangle$ be the Schmidt decomposition of ψ for some Schmidt basis $\{|i\rangle\}_{i=0}^{d-1}$. On each copy we perform the LOCC measurement on the Schmidt basis of ψ to obtain the two probability distributions P and Q:

$$P(x) = \begin{cases} \lambda_i & x = (i,i) \\ 0 & \text{otherwise} \end{cases};$$
$$Q = \frac{d^2}{d^2-1}U - \frac{1}{d^2-1}P,$$

where U is the uniform distribution over $\{0, 1, \ldots, d^2-1\}$.

Then, the classical Chernoff gives an upper bound on the error probability of discriminating P and Q by

$$p^\alpha(1-p)^{1-\alpha}\left(\sum_x P^\alpha(x)Q^{1-\alpha}(x)\right)^n = p^\alpha(1-p)^{1-\alpha}\left(\frac{1}{(d^2-1)^{1-\alpha}}\sum_{i=0}^{d-1}\lambda_i^\alpha(1-\lambda_i)^{1-\alpha}\right)^n, \quad \forall \alpha \in [0,1].$$

We simply choose $\alpha = 0$ to obtain desired upper bound, which completes the proof. □

Remark 3.1. For multipartite state $|\psi\rangle \in \mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2} \otimes \cdots \otimes \mathbb{C}^{d_n}$, we can choose the measurement as follows. We first expand $|\psi\rangle$ in the computational basis of the first $n-2$ system

$$|\psi\rangle = |0\cdots 0\rangle|\phi\rangle_{n-1,n} + \sum_{j \neq 0\cdots 0}|j\rangle|\phi_j\rangle_{n-1,n}.$$

Let $\sum_{i=0}^{d-1}\sqrt{\lambda_i}|ii\rangle$ be the Schmidt decomposition of $\psi_{n-1,n}$ for some Schmidt basis $\{|i\rangle\}_{i=0}^{d-1}$, where $d \leq d_{n-1}, d_n$.

We measure system the first $n-2$ system in computational basis and the last two system in the Schmidt basis corresponding to $\psi_{n-1,n}$. The resulting probability distributions is P and

$$Q = \frac{d_1 d_2 \cdots d_n}{d_1 d_2 \cdots d_n - 1}U - \frac{1}{d_1 d_2 \cdots d_n - 1}P$$

where the support of P is at most $d_1 d_2 \cdots d_n + d - d_{n-1}d_n$. Using the classical Chernoff bound we can obtain that

$$P_{\mathrm{e}}^{\mathrm{LOCC}}\left(\psi^{\otimes n}, (\psi^\perp)^{\otimes n}; p\right) \leq (1-p)\left(1 - \frac{d_{n-1}d_n - \min\{d_{n-1}, d_n\}}{d_1 d_2 \cdots d_n - 1}\right)^n, \quad n \in \mathbb{N}.$$

Remark 3.2. Let us emphasize that if ψ is not entangled, but rather a product state, then the distribution P is a singleton and hence the Chernoff bound becomes ∞, which can be seen by choosing $\alpha = 1/2$.

As we will prove in Theorem 8 of Section 4.1, the upper bound given in Proposition 2 is tight when ψ is maximally entangled.

3.1. **A Converse Bound for Symmetric Hypothesis Testing.**

Proposition 3 (Exponential lower bound for PPT POVMs)**.** *Let ψ be a bipartite entangled pure state on $\mathbb{C}^d \otimes \mathbb{C}^d$ and $\psi^\perp := \frac{\mathbb{1}-\psi}{d^2-1}$ be its orthogonal complement. Then, the error probability of testing multi-copies of ψ and $\psi^\perp$ using LOCC POVMs is lower bounded by*

$$P_{\mathrm{e}}^{\mathrm{LOCC}}\left(\psi^{\otimes n}, (\psi^\perp)^{\otimes n}; p\right) \geq P_{\mathrm{e}}^{\mathrm{PPT}}\left(\psi^{\otimes n}, (\psi^\perp)^{\otimes n}; p\right) \geq \min\{p, 1-p\} \cdot t^n, \quad \forall n \in \mathbb{N} \tag{12}$$

where $t := \frac{1-\eta}{(d^2-1)\eta} \in (0, \frac{1}{d+1})$ and η denotes the largest squared Schmidt coefficient of ψ, i.e. $\eta := \|\operatorname{Tr}_A[\psi]\| \in [\frac{1}{d}, 1)$.

Hence, it follows that $\xi_{\mathrm{C}}^{\mathrm{LOCC}}\left(\psi,\psi^{\perp}\right)<\infty$*.*

Proof. For the leftmost lower bound of (12), our key technique is to succinctly formulate the PPT-distinguishability norm [71] into its dual representation in terms of Schatten 1-norm. Then, we find a feasible solution to the dual problem to prove our claim.

Since

$$P_{\mathrm{e}}^{\mathrm{PPT}}\left(\psi^{\otimes n},(\psi^{\perp})^{\otimes n};p\right)\geq 2\min\{p,1-p\}\cdot P_{\mathrm{e}}^{\mathrm{PPT}}\left(\psi^{\otimes n},(\psi^{\perp})^{\otimes n};\tfrac{1}{2}\right),$$

it suffices to consider exponential lower bound for the scenario with equiprobable prior distribution. Let $\rho=\psi^{\otimes n}$ and $\sigma=(\psi^{\perp})^{\otimes n}$. The error probability using PPT POVM is

$$\begin{aligned}P_{\mathrm{e}}^{\mathrm{PPT}}\left(\psi^{\otimes n},(\psi^{\perp})^{\otimes n};\tfrac{1}{2}\right)&:=\inf_{M,(\mathbb{1}-M)\in\mathrm{PPT}}\frac{1}{2}\operatorname{Tr}\left[\rho(\mathbb{1}-M)\right]+\frac{1}{2}\operatorname{Tr}\left[\sigma M\right] &(13)\\&=\inf_{M,(\mathbb{1}-M)\in\mathrm{PPT}}\frac{1}{4}+\frac{1}{2}\operatorname{Tr}\left[(\rho-\sigma)\left(2M-\mathbb{1}\right)\right]\\&=\frac{1}{2}\left(1-\frac{1}{2}\left\|\rho-\sigma\right\|_{\mathrm{PPT}}\right), &(14)\end{aligned}$$

where the PPT-distinguishability norm for Hermitian matrices [71] is defined as

$$\|H\|_{\mathrm{PPT}}:=\sup_{-\mathbb{1}\leq M,M^{\Gamma}\leq\mathbb{1}}\operatorname{Tr}\left[HM\right]. \tag{15}$$

Since there are four linear inequality constraints in the above optimization, we write the Lagrangian L by introducing the corresponding dual variables $A,B,C,D\geq 0$ as

$$\begin{aligned}L&=\operatorname{Tr}\left[HM\right]+\operatorname{Tr}\left[A(\mathbb{1}-M)\right]+\operatorname{Tr}\left[B(\mathbb{1}+M)\right]+\operatorname{Tr}\left[C(\mathbb{1}-M^{\Gamma})\right]+\operatorname{Tr}\left[D(\mathbb{1}+M^{\Gamma})\right]\\&=\operatorname{Tr}\left[A+B+C+D\right]+\operatorname{Tr}\left[M\left(H-A+B-C^{\Gamma}+D^{\Gamma}\right)\right].\end{aligned}$$

To derive the dual program of (15), we maximize L over all Hermitian matrices M without constraints. Clearly, this is infinite unless $H=A-B+(C-D)^{\Gamma}$. For such a case, the maximum is simply $\operatorname{Tr}\left[A+B+C+D\right]$. Hence, the dual formulation for the PPT-distinguishability norm is

$$\|H\|_{\mathrm{PPT}}=\inf_{A,B,C,D\geq 0}\left\{\operatorname{Tr}\left[A+B+C+D\right]:H=A-B+(C-D)^{\Gamma}\right\}.$$

Moreover, we can rewrite the above in terms of 1-norm as follows by denoting Hermitian matrices $X=A-B$ and $Y=(C-D)^{\Gamma}$:

$$\|H\|_{\mathrm{PPT}}=\inf_{X,Y}\left\{\|X\|_1+\left\|Y^{\Gamma}\right\|_1:H=X+Y\right\}. \tag{16}$$

This is because the optimal decomposition of the absolute value of a matrix is into its positive and negative parts. That is,

$$\begin{aligned}\|X\|_1&=\sup_{-\mathbb{1}\leq M\leq\mathbb{1}}\operatorname{Tr}\left[MX\right]\\&=\inf_{A,B\geq 0}\sup_{M}\left\{\operatorname{Tr}[MX]+\operatorname{Tr}\left[A\left(\mathbb{1}-M\right)\right]+\operatorname{Tr}\left[B\left(\mathbb{1}+M\right)\right]\right\}\\&=\inf_{A,B\geq 0}\left\{\operatorname{Tr}\left[A+B\right]:X=A-B\right\}.\end{aligned}$$

Next, we find a feasible solution to the dual program (16) to obtain a lower bound to the error probability. Let $\eta:=\|\operatorname{Tr}_A[\psi]\|\in[\frac{1}{d},1)$. We choose

$$\begin{aligned}X&=(1-t^n)\rho;\\Y&=t^n\rho-\sigma;\\t&=\frac{1-\eta}{(d^2-1)\eta}\in\left(0,\frac{1}{d+1}\right].\end{aligned} \tag{17}$$

Since $t^n < 1$ and $\frac{\psi^\Gamma}{\eta} \le \mathbb{1}$, we have

$$\begin{aligned} X &= (1-t^n)\psi^{\otimes n} \ge 0; \\ Y^\Gamma &= \frac{1}{(d^2-1)^n} \cdot \left[\left(\frac{\psi^\Gamma}{\eta} - \psi^\Gamma \right)^{\otimes n} - \left(\mathbb{1} - \psi^\Gamma \right)^{\otimes n} \right] \le 0, \end{aligned} \tag{18}$$

which, in turn, implies that

$$\begin{aligned} \|X\|_1 &= \operatorname{Tr}[X] = (1-t^n); \\ \left\|Y^\Gamma\right\|_1 &= -\operatorname{Tr}\left[Y^\Gamma\right] = -\operatorname{Tr}[Y] = (1-t^n). \end{aligned} \tag{19}$$

Combining (13), (14), (16), (17), and (19) then gives

$$P_{\mathrm{e}}^{\mathrm{PPT}} \ge \frac{1}{2}\left(1 - (1-t^n)\right) = \frac{1}{2} t^n$$

as desired. □

3.2. A Strong Converse Bound for Asymmetric Hypothesis Testing.

Proposition 4 (Lower bound to the strong converse exponent). *Let ψ be a pure bipartite entangled state on $\mathbb{C}^d \otimes \mathbb{C}^d$ and $\psi^\perp := \frac{\mathbb{1}-\psi}{d^2-1}$ be its orthogonal complement. Let $t := \frac{1-\eta}{(d^2-1)\eta} \in (0, \frac{1}{d+1})$ and $\eta := \|\operatorname{Tr}_A[\psi]\| \in [\frac{1}{d}, 1)$. Then, for any $r > \log\frac{1}{t}$, we have*

$$\alpha_{\mathrm{e}^{-nr}}^{\mathrm{PPT}}\left(\psi^{\otimes n}, (\psi^\perp)^{\otimes n}\right) \ge 1 - \mathrm{e}^{-n\left[r - \log\frac{1}{t}\right]}, \quad \forall n \in \mathbb{N}.$$

In other words,

$$\xi_{\mathrm{SC}}^{\mathrm{PPT}}\left(\psi, \psi^\perp, r\right) \ge r - \log\tfrac{1}{t}, \quad \forall r > \log\tfrac{1}{t}.$$

Proof. Firstly, we claim that

$$\alpha_\mu^{\mathrm{X}}(\rho,\sigma) = \max_{\lambda \ge 0} \left\{ \frac{1}{2}\left(1 + \lambda - \|\rho - \lambda\sigma\|_{\mathrm{X}}\right) - \lambda\mu \right\}, \tag{20}$$

where α_μ^{X} is defined in (7) of Section 2. This can be proved by the following:

$$\begin{aligned} \alpha_\mu^{\mathrm{X}}(\rho,\sigma) &:= \inf_{T\in\mathrm{X}} \left\{ \operatorname{Tr}\left[\rho(\mathbb{1}-T)\right] : \operatorname{Tr}\left[\sigma T\right] \le \mu \right\} \\ &= \max_{\lambda\ge 0} \inf_{T\in\mathrm{X}} \left\{ \operatorname{Tr}\left[\rho(\mathbb{1}-T)\right] + \lambda \operatorname{Tr}\left[\sigma T\right] - \lambda\mu \right\} \\ &= \max_{\lambda\ge 0} \inf_{T\in\mathrm{X}} \left\{ \frac{1}{2}\left(1 + \lambda - \operatorname{Tr}\left[(\rho-\lambda\sigma)(2T-\mathbb{1})\right]\right) - \lambda\mu \right\} \\ &= \max_{\lambda\ge 0} \left\{ \frac{1}{2}\left(1 + \lambda - \|\rho-\lambda\sigma\|_{\mathrm{X}}\right) - \lambda\mu \right\}, \end{aligned}$$

where we have used the identity $\rho(\mathbb{1}-T) + \lambda\sigma T = \frac{1}{2}\left(\rho + \lambda\sigma - (\rho-\lambda\sigma)(2T-\mathbb{1})\right)$, and the PPT-distinguishability norm given in (15).

Now let $\rho = \psi^{\otimes n}$ and $\sigma = (\psi^\perp)^{\otimes n}$. We invoke the dual formulation for the PPT-distinguishability norm given in (16):

$$\|\rho - \lambda\sigma\|_{\mathrm{PPT}} = \inf_{X,Y} \left\{ \|X\|_1 + \left\|Y^\Gamma\right\|_1 : \rho - \lambda\sigma = X + Y \right\}. \tag{21}$$

To obtain a lower bound to $\alpha_\mu^{\mathrm{PPT}}(\rho,\sigma)$, we choose

$$\begin{aligned} X &= (1-\lambda t^n)\rho; \\ Y &= \lambda t^n \rho - \lambda\sigma; \\ t &= \frac{1-\eta}{(d^2-1)\eta} \in \left(0, \frac{1}{d+1}\right]; \\ \lambda &= t^{-n}. \end{aligned}$$

As the derivation given in (18), we have

$$X = (1-\lambda t^n)\psi^{\otimes n} = 0$$
$$Y^\Gamma = \frac{\lambda}{(d^2-1)^n}\cdot\left[\left(\frac{\psi^\Gamma}{\eta}-\psi^\Gamma\right)^{\otimes n} - \left(\mathbb{1}-\psi^\Gamma\right)^{\otimes n}\right] \leq 0,$$

and therefore,

$$\begin{aligned} &\|X\|_1 = 0; \\ &\left\|Y^\Gamma\right\|_1 = -\operatorname{Tr}\left[Y^\Gamma\right] = -\operatorname{Tr}[Y] = \lambda(1-t^n). \end{aligned} \tag{22}$$

Combining (20), (21), and (22), we have

$$\begin{aligned} \alpha^{\mathrm{PPT}}_{\exp\{-nr\}}\left(\psi^{\otimes n},(\psi^\perp)^{\otimes n}\right) &\geq \frac{1}{2}\left[1+\lambda-\lambda(1-t^n)\right] - \lambda\,\mathrm{e}^{-nr} \\ &= 1 - t^{-n}\,\mathrm{e}^{-nr} \end{aligned}$$

as claimed. □

From Proposition 4, the type-I error converges to 1 exponentially fast as long as the exponential decay rate of the type-II error exceeds $\log\frac{1}{t}$. This directly gives a strong converse bound for the Stein exponent for distinguishing $\psi^{\otimes n}$ against $(\psi^\perp)^{\otimes n}$.

Corollary 5 (Strong converse bound for the Stein exponent)**.** *Let ψ be a pure bipartite entangled state on $\mathbb{C}^d\otimes\mathbb{C}^d$ and $\psi^\perp := \frac{\mathbb{1}-\psi}{d^2-1}$ be its orthogonal complement. Let $\eta := \|\operatorname{Tr}_A[\psi]\| \in [\frac{1}{d},1)$. We have*

$$\xi^{\mathrm{PPT}}_{\mathrm{S}}\left(\psi^{\otimes n},(\psi^\perp)^{\otimes n},\varepsilon\right) \leq \log\frac{(d^2-1)\eta}{1-\eta} \in [\log(d+1),\infty)\,, \quad \forall\varepsilon\in(0,1).$$

4. Optimal LOCC Protocol for Testing Maximally Entangled States

This section aims to test the following hypotheses of maximally entangled states against its orthogonal complement:

$$\begin{cases} \mathsf{H}_0 : \rho_0^{\otimes n} = \Phi_d^{\otimes n} \\ \mathsf{H}_1 : \rho_1^{\otimes n} = \left(\Phi_d^\perp\right)^{\otimes n} := \left(\dfrac{\mathbb{1}_{d^2}-\Phi_d}{d^2-1}\right)^{\otimes n} \end{cases}, \quad \forall n\in\mathbb{N}. \tag{23}$$

In Section 4.1, we establish the optimal average error probability (Theorem 8) and the Chernoff exponent (Corollary 9) in the symmetric setting. In Section 4.2, we obtain the optimal trade-off between the two types of errors (Theorem 10) and the Stein, Hoeffding, and strong converse exponents (Corollaries 11, 12, and 13).

4.1. Symmetric Hypothesis Testing.

In this section, we consider the binary hypotheses given in (23) with the prior probability $\Pr[\mathsf{H}_0] = p \in (0,1)$ and $\Pr[\mathsf{H}_1] = 1-p$. In Theorem 8 below, we derive the expression for the optimal average error probabilities under the optimal PPT, SEP, and LOCC POVMs, i.e.

$$P_{\mathrm{e}}^{\mathrm{PPT}}(\rho_0^{\otimes n},\rho_1^{\otimes n};p) = P_{\mathrm{e}}^{\mathrm{SEP}}(\rho_0^{\otimes n},\rho_1^{\otimes n};p) = P_{\mathrm{e}}^{\mathrm{LOCC}}(\rho_0^{\otimes n},\rho_1^{\otimes n};p) = \min\left\{(1-p)\left(\frac{1}{d+1}\right)^n, p\right\}.$$

To obtain this result, we first provide an upper bound using LOCC (i.e. achievability) in Proposition 6, and then prove a lower bound using PPT (i.e. converse) in Proposition 7.

Proposition 6. *Consider the binary hypothesis given in (23). There is an LOCC protocol such that*

$$P_{\mathrm{e}}^{\mathrm{LOCC}}(\rho_0^{\otimes n},\rho_1^{\otimes n};p) \leq \min\left\{(1-p)\left(\frac{1}{d+1}\right)^n, p\right\}.$$

Proof. For each copy, we perform the two-outcome measurement by

$$\left\{ G_d = \sum_{i\neq j}^{d-1} |ij\rangle\langle ij|,\ \mathbb{1} - G_d = \sum_{i=0}^{d-1} |ii\rangle\langle ii| \right\}.$$

Since the states in the hypotheses are both $U \otimes U^*$-invariant (where '$*$' means the complex conjugate), after the twirling operation, we have the measurement

$$\left\{ M_d = \Phi_d + \frac{1}{d+1}\left(\mathbb{1} - \Phi_d\right),\ \mathbb{1} - M_d = \frac{d}{d+1}\left(\mathbb{1} - \Phi_d\right) \right\}. \tag{24}$$

It can be verified that such choice of M_d is implementable by a (one-way) LOCC protocol. Moreover,

$$M_d^{\otimes n} = \sum_{k=0}^{n} \left(\frac{1}{d+1}\right)^{n-k} T_k,$$

where T_k denotes the sum of all elements of $\{\Phi_d, \mathbb{1} - \Phi_d\}^{\otimes n}$ which have k copies of Φ_d. Hence,

$$\begin{aligned} P_{\mathrm{e}}^{\mathrm{LOCC}} &= p \operatorname{Tr}\left[\left(\mathbb{1} - M_d^{\otimes n}\right) \Phi_d^{\otimes n}\right] + (1-p) \operatorname{Tr}\left[M_d^{\otimes n} \left(\frac{\mathbb{1} - \Phi_d}{d^2-1}\right)^{\otimes n}\right] \\ &= (1-p) \operatorname{Tr}\left[\left(\frac{1}{d+1}\right)^{n} T_0 \left(\frac{\mathbb{1} - \Phi_d}{d^2-1}\right)^{\otimes n}\right] \\ &= (1-p) \left(\frac{1}{d+1}\right)^{n}. \end{aligned} \tag{25}$$

Lastly, if $P_{\mathrm{e}}^{\mathrm{LOCC}} > p$, we just choose $\rho_1^{\otimes n} = \left(\frac{\mathbb{1}-\Phi_d}{d^2-1}\right)^{\otimes n}$. □

Next, we move on to the lower bound via PPT POVMs. Our approach is inspired by [53] to formulate the average error probability as a linear programming, and find a feasible solution to its dual problem.

Proposition 7. *Consider the binary hypothesis given in (23). Then, for all PPT POVMs, it satisfied that*

$$P_{\mathrm{e}}^{\mathrm{PPT}}(\rho_0^{\otimes n}, \rho_1^{\otimes n}; p) \geq \min\left\{ (1-p)\left(\frac{1}{d+1}\right)^{n}, p \right\}.$$

Proof. Since the states $\rho_0^{\otimes n}$ and $\rho_1^{\otimes n}$ are invariant under permutations of each copy and under $U \otimes U^*$ transformations of the individual copies, we can impose such symmetry on the two-outcome POVM as the following without loss of generality:

$$\left\{ \sum_{k=0}^{n} x_k B_k, \sum_{k=0}^{n} (1 - x_k) B_k \right\}, \tag{26}$$

where B_k denotes the sum of all elements of $\{\Phi_d, \mathbb{1} - \Phi_d\}^{\otimes n}$ which have k copies of Φ_d. Note that the constraints $0 \leq x_k \leq 1$ for $k = 0, \ldots, n$ are necessary and sufficient for the test to be valid. The minimum average error probability is then

$$P_{\mathrm{e}}^{\mathrm{PPT}}(\rho_0^{\otimes n}, \rho_1^{\otimes n}; p) = p \operatorname{Tr}\left[\sum_{k=0}^{n} (1-x_k) B_k \rho_0^{\otimes n}\right] + (1-p) \operatorname{Tr}\left[\sum_{k=0}^{n} x_k B_k \rho_1^{\otimes n}\right] = p + (1-p)\left(x_0 - \frac{p}{1-p} x_n\right). \tag{27}$$

Since the partial transpose of the flip operator F is $d\Phi_d$, we have

$$\Phi_d^\Gamma = \frac{F}{d} = \frac{1}{d}\Pi_s - \frac{1}{d}\Pi_a,$$
$$(\mathbb{1} - \Phi_d)^\Gamma = \mathbb{1} - \Phi_d^\Gamma = \left(1 - \frac{1}{d}\right)\Pi_s + \left(1 + \frac{1}{d}\right)\Pi_a.$$

Here, we denote Π_s and Π_a the projections onto the symmetric and anti-symmetric subspaces. Hence, we can write the operators B_k^Γ as linear combinations of operators from the set of 2^n orthogonal operators $\{\Pi_s, \Pi_a\}^{\otimes n}$.

Let S_k^n denote the subset of strings in $\{0,1\}^n$ which have k zeros. Then,

$$\begin{aligned} B_k^\Gamma &= \sum_{v \in S_k^n} \bigotimes_{i=1}^{n} \left(\left(v_i + (-1)^{v_i}\frac{1}{d}\right)\Pi_s + \left(v_i - (-1)^{v_i}\frac{1}{d}\right)\Pi_a \right) \\ &= \sum_{l=0}^{n} \sum_{0 \le j \le l,k} \binom{l}{j}\binom{n-l}{k-j}\left(\frac{1}{d}\right)^j \left(1-\frac{1}{d}\right)^{l-j}\left(-\frac{1}{d}\right)^{k-j}\left(1+\frac{1}{d}\right)^{n-l-k+j} A_l, \end{aligned}$$

where A_l is the sum over all elements of $\{\Pi_s, \Pi_a\}^{\otimes n}$ which have l copies of Π_s. Then, a necessary and sufficient condition for the two-outcome POVM to be PPT is given by the following inequalities:

$$\begin{aligned} \sum_{k=0}^{n} x_k \sum_{0 \le j \le l,k} \binom{l}{j}\binom{n-l}{k-j}\left(\frac{1}{d}\right)^j \left(1-\frac{1}{d}\right)^{l-j}\left(-\frac{1}{d}\right)^{k-j}\left(1+\frac{1}{d}\right)^{n-l-k+j} &\ge 0 \quad \text{for } l = 0, \ldots, n, \\ \sum_{k=0}^{n} (1-x_k) \sum_{0 \le j \le l,k} \binom{l}{j}\binom{n-l}{k-j}\left(\frac{1}{d}\right)^j \left(1-\frac{1}{d}\right)^{l-j}\left(-\frac{1}{d}\right)^{k-j}\left(1+\frac{1}{d}\right)^{n-l-k+j} &\ge 0 \quad \text{for } l = 0, \ldots, n. \end{aligned} \tag{28}$$

In the following, we rewrite the constraints in (28) for the convenience of the linear program. Denote an $(n+1) \times (n+1)$ matrix Q with its elements

$$Q_{lk} = \sum_{0 \le j \le l,k} \binom{l}{j}\binom{n-l}{k-j}\left(\frac{1}{d}\right)^j \left(1-\frac{1}{d}\right)^{l-j}\left(-\frac{1}{d}\right)^{k-j}\left(1+\frac{1}{d}\right)^{n-l-k+j}. \tag{29}$$

It can be verified that

$$\begin{aligned} \sum_{k=0}^{n} Q_{lk} &= \left(\frac{d+1}{d}\right)^n \left(\frac{d-1}{d+1}\right)^l \sum_{m=0}^{n-l}\binom{n-l}{m}\left(\frac{-1}{d+1}\right)^m \sum_{j=0}^{l}\binom{l}{j}\left(\frac{1}{d-1}\right)^j \\ &= \left(\frac{d+1}{d}\right)^n \left(\frac{d-1}{d+1}\right)^l \sum_{m=0}^{n-l}\binom{n-l}{m}\left(\frac{-1}{d+1}\right)^m \left(1+\frac{1}{d-1}\right)^l \\ &= \left(\frac{d+1}{d}\right)^n \left(\frac{d-1}{d+1}\right)^l \left(1-\frac{1}{d+1}\right)^{n-l}\left(\frac{d}{d-1}\right)^l \\ &= 1. \end{aligned}$$

Then, letting

$$c_i = \delta_{0i} - \frac{p}{1-p}\delta_{ni}$$
$$b_i = \begin{cases} 0 & i = 0, \ldots, n, \\ -1 & i = n+1, \ldots, 2n+1, \\ -1 & i = 2n+2, \ldots, 3n+2, \end{cases}$$
$$P = \begin{pmatrix} Q \\ -Q \\ -\mathbb{1} \end{pmatrix},$$

we can write (27) as a linear program as follows:

$$\min \left\{ \mathbf{c}^T \cdot \mathbf{x} | P \cdot \mathbf{x} \geq \mathbf{b}, \mathbf{x} \geq \mathbf{0} \right\}.$$

The dual problem is

$$P_{\mathrm{e}}^{\mathrm{PPT}}(\rho_0^{\otimes n}, \rho_1^{\otimes n}; p) = p + (1-p)\mathbf{c}^T \cdot \mathbf{x} \geq p + (1-p)\mathbf{b}^T \cdot \mathbf{y}, \tag{30}$$

subject to $\mathbf{P}^T \cdot \mathbf{y} \leq \mathbf{c}, \mathbf{y} \geq \mathbf{0}$.

Letting $\mathbf{y} = \mathbf{u} \oplus \mathbf{v} \oplus \mathbf{w}$, the dual problem can be rewritten as

$$\max_{\mathbf{y}} \left\{ -\sum_{i=1}^{n} v_i - \sum_{i=1}^{n} w_i \middle| \mathbf{u}, \mathbf{v}, \mathbf{w} \geq \mathbf{0}, Q^T \cdot \mathbf{u} - Q^T \cdot \mathbf{v} - \mathbf{w} \leq \mathbf{c} \right\}. \tag{31}$$

Now, we consider a point $\mathbf{y}^\star = \mathbf{u}^\star \oplus \mathbf{v}^\star \oplus \mathbf{w}^\star$, where

$$\begin{aligned} u_i^\star &= \frac{1}{2^n}\left[1 - \left(-\frac{d-1}{d+1}\right)^{n-i}\right], \\ v_i^\star &= 0, \\ w_i^\star &= \begin{cases} 0 & i = 0, \ldots, n-1 \\ \max\left\{\frac{p}{1-p} - \left(\frac{1}{d+1}\right)^n, 0\right\} & i = n \end{cases}. \end{aligned} \tag{32}$$

In Appendix A, we show that $\mathbf{y}^\star$ is a feasible solution to (31). Then, the dual objective function at $\mathbf{y}^\star = \mathbf{u}^\star \oplus \mathbf{v}^\star \oplus \mathbf{w}^\star$ is

$$-\sum_{i=0}^{n} v_i^\star - \sum_{i=0}^{n} w_i^\star = \min\left\{-\frac{p}{1-p} + \left(\frac{1}{d+1}\right)^n, 0\right\},$$

which by combining with (30) gives the desired lower bound to the minimum average error probability under PPT POVMs:

$$P_{\mathrm{e}}^{\mathrm{PPT}}(\rho_0^{\otimes n}, \rho_1^{\otimes n}; p) \geq \min\left\{(1-p)\left(\frac{1}{d+1}\right)^n, p\right\}.$$

□

Combining (3), Propositions 6 and 7, we explicitly calculate the minimum average error probability for discriminating the maximally entangled state and its orthogonal complement. (We do not prove because they follow from the preceding discussion).

Theorem 8 (Optimal average error probability)**.** *Consider the binary hypotheses given in* (23)*. It holds that*

$$P_{\mathrm{e}}^{\mathrm{PPT}}(\rho_0^{\otimes n}, \rho_1^{\otimes n}; p) = P_{\mathrm{e}}^{\mathrm{SEP}}(\rho_0^{\otimes n}, \rho_1^{\otimes n}; p) = P_{\mathrm{e}}^{\mathrm{LOCC}}(\rho_0^{\otimes n}, \rho_1^{\otimes n}; p) = \min\left\{(1-p)\left(\frac{1}{d+1}\right)^n, p\right\}.$$

By the definition given in (6), we obtain of our main result of the Chernoff exponents:

Corollary 9 (Chernoff exponent)**.** *Consider the binary hypothesis given in* (23)*. For every* $0 < p < 1$*, we have*

$$\xi_{\mathrm{C}}^{\mathrm{PTT}}(\rho_0, \rho_1) = \xi_{\mathrm{C}}^{\mathrm{SEP}}(\rho_0, \rho_1) = \xi_{\mathrm{C}}^{\mathrm{LOCC}}(\rho_0, \rho_1) = \log(d+1).$$

Remark 4.1. Theorem 8 shows that our result of exponential lower bound for testing arbitrary multipartite entangled pure state against its orthogonal complement given in Proposition 3 of Section 3.1 is tight for maximally entangled state Φ_d.

4.2. **Asymmetric Hypothesis Testing.** In this section, we consider the asymmetric setting of the binary hypotheses (23). Our main result is first to obtain the optimal trade-ff between the type-I and type-II errors (Theorem 10), and then apply it to obtain the corresponding Stein, Hoeffding, and strong converse exponents (Corollaries 11, 12, and 13).

In Theorem 10 below, we calculate the β_α for the three classes of POVMs.

Theorem 10 (Optimal trade-off)**.** *Consider the binary hypotheses given in* (23)*. It holds that*

$$\beta_\alpha^{\mathrm{PPT}}(\rho_0^{\otimes n}, \rho_1^{\otimes n}) = \beta_\alpha^{\mathrm{SEP}}(\rho_0^{\otimes n}, \rho_1^{\otimes n}) = \beta_\alpha^{\mathrm{LOCC}}(\rho_0^{\otimes n}, \rho_1^{\otimes n}) = \frac{1-\alpha}{(d+1)^n}.$$

Proof. We will first prove an upper bound:

$$\beta_\alpha^{\mathrm{LOCC}}(\rho_0^{\otimes n}, \rho_1^{\otimes n}) \leq \frac{1-\alpha}{(d+1)^n}, \tag{33}$$

and next an lower bound:

$$\beta_\alpha^{\mathrm{PPT}}(\rho_0^{\otimes n}, \rho_1^{\otimes n}) \geq \frac{1-\alpha}{(d+1)^n}. \tag{34}$$

On the other hand, from the definition given in (8), it follows that

$$\beta_\alpha^{\mathrm{LOCC}} \geq \beta_\alpha^{\mathrm{SEP}} \geq \beta_\alpha^{\mathrm{PPT}}.$$

Combining the above three inequalities thus completes the proof.

We begin with the upper bound (33), i.e. the achievability. For every $\alpha \in [0,1]$, we choose the test

$$T_n := (1-\alpha)\cdot M_d^{\otimes n} = (1-\alpha)\left[\Phi_d + \frac{1}{d+1}\left(\mathbb{1} - \Phi_d\right)\right]^{\otimes n},$$

where M_d is defined in (24) of Section 4.1. As in the proof of Proposition 7, the chosen test and its complement are implementable by an (one-way) LOCC protocol. Then, we calculate the following

$$\begin{aligned}
\alpha_n(T_n) &= \mathrm{Tr}\left[\left(\mathbb{1} - T_n\right)\Phi_d^{\otimes n}\right] \\
&= (1-\alpha)\cdot \mathrm{Tr}\left[\left(\mathbb{1} - M_d^{\otimes n}\right)\Phi_d^{\otimes n}\right] + \alpha\cdot\mathrm{Tr}\left[\mathbb{1}\cdot\Phi_d^{\otimes n}\right] \\
&= (1-\alpha)\cdot \mathrm{Tr}\left[\left(\mathbb{1} - \Phi_d^{\otimes n}\right)\Phi_d^{\otimes n}\right] + \alpha \\
&= \alpha.
\end{aligned} \tag{35}$$

On the other hand,

$$\begin{aligned}
\beta_n(T_n) &= \mathrm{Tr}\left[T_n(\Phi_d^{\perp})^{\otimes n}\right] \\
&= (1-\alpha)\cdot\mathrm{Tr}\left[M_d^{\otimes n}(\Phi_d^{\perp})^{\otimes n}\right] \\
&= \frac{1-\alpha}{(d+1)^n},
\end{aligned} \tag{36}$$

where we have used (25) in the last line.

Equations (35) and (36) imply that $(\alpha, \beta_n(T_n)) \in \mathcal{R}(\rho_0^{\otimes n}, \rho_1^{\otimes n})$. Hence, we have $\beta_\alpha^{\mathrm{LOCC}}(\rho_0^{\otimes n}, \rho_1^{\otimes n}) \leq \beta_n(T_n) = \frac{1-\alpha}{(d+1)^n}$ as desired.

Next, we move on to prove the lower bound (34), i.e. the converse. Following the proof of Proposition 7, we let the test to of the form in (26),

$$T_n := \sum_{k=0}^{n} x_k B_k,$$

where $0 \le x_k \le 1$, and B_k means the sum of all elements of $\{\Phi_d, \mathbb{1} - \Phi_d\}^{\otimes n}$ which have k copies of Φ_d. From (27), we have

$$\alpha_n(T_n) = (1 - x_n); \quad \beta_n(T_n) = x_0.$$

Since the test (and its complement) must satisfy the PPT condition, the constraint in (28) still hold. Hence, we formulate our problem as a linear program:

$$\beta_\alpha^{\mathrm{PTT}}(\rho_0^{\otimes n}, \rho_1^{\otimes n}) = \min_{\mathbf{x}} \left\{ \mathbf{c}^T \cdot \mathbf{x} | P \cdot \mathbf{x} \ge \mathbf{b}, \mathbf{x} \ge \mathbf{0} \right\},$$

where

$$\begin{aligned} c_i &= \delta_{0i}, \\ b_i &= \begin{cases} 0 & i = 0, \dots, n, \\ -1 & i = n+1, \dots, 2n+1, \\ -1 & i = 2n+2, \dots, 3n+2, \\ 1-\alpha & i = 3n+3, \end{cases} \\ P &= \begin{pmatrix} Q \\ -Q \\ -\mathbb{1} \\ \mathbf{h}^T \end{pmatrix}, \\ h_i &= \delta_{ni}, \end{aligned}$$

and the matrix Q is introduced in (29).

The dual problem is then

$$\max_{\mathbf{x}} \left\{ \mathbf{b}^T \cdot \mathbf{y} | P^T \cdot \mathbf{y} \le \mathbf{c}, \mathbf{y} \ge \mathbf{0} \right\}.$$

Letting $\mathbf{y} = \mathbf{u} \oplus \mathbf{v} \oplus \mathbf{w} \oplus z$ for convenience, we have

$$\begin{aligned} &\min_{\mathbf{x}} \left\{ \mathbf{c}^T \cdot \mathbf{x} | P \cdot \mathbf{x} \ge \mathbf{b}, \mathbf{x} \ge \mathbf{0} \right\} \\ &\ge \max_{\mathbf{x}} \left\{ -\mathbf{1}^T \cdot \mathbf{v} - \mathbf{1}^T \cdot \mathbf{w} + (1-\varepsilon) z | \, Q^T \cdot \mathbf{u} - Q^T \cdot \mathbf{v} - \mathbf{w} + z\mathbf{h} \le \mathbf{c}, \mathbf{u}, \mathbf{v}, \mathbf{w} \ge \mathbf{0}, z \ge 0 \right\}. \end{aligned} \tag{37}$$

Now, we choose a point $\mathbf{y}^\star = \mathbf{u}^\star \oplus \mathbf{v}^\star \oplus \mathbf{w}^\star \oplus z^\star$, with $\mathbf{u}^\star$ be given in (32), $\mathbf{v}^\star = \mathbf{w}^\star = \mathbf{0}$, and $z^\star = \frac{1}{(d+1)^n}$. Then, by (53) in Appendix A, we have

$$\left(Q^T \cdot \mathbf{u}^\star - Q^T \cdot \mathbf{v}^\star - \mathbf{w}^\star + z^\star \mathbf{h} \right)_k = \delta_{0k} - \delta_{nk} \frac{1}{(d+1)^n} + \delta_{nk} \frac{1}{(d+1)^n} = \delta_{0k} = c_k.$$

This along with (52) implies that the chosen $\mathbf{y}^\star$ is a feasible solution to the dual problem, (37). Hence,

$$\beta_\varepsilon^{\mathrm{PTT}}(\rho_0^{\otimes n}, \rho_1^{\otimes n}) \ge -\mathbf{1}^T \cdot \mathbf{v}^\star - \mathbf{1}^T \cdot \mathbf{w}^\star + (1-\alpha) z^\star = (1-\varepsilon) z^\star = \frac{1-\alpha}{(d+1)^n},$$

which proves our claim. □

The optimal trade-off between type-I and type-II errors (Theorem 10) directly gives us the results of the Stein exponent (Corollary 11), the Hoeffding exponent (Corollary 12), and strong converse exponent (Corollary 13) as below.

Corollary 11 (Stein exponent)**.** *The following Stein exponents hold.*

(a) Consider the binary hypothesis testing: $\rho_0 = \Phi_d$ and $\rho_1 = \Phi_d^\perp$. Then,

$$\xi_{\rm S}^{\rm PTT}(\rho_0,\rho_1,\varepsilon) = \xi_{\rm S}^{\rm SEP}(\rho_0,\rho_1,\varepsilon) = \xi_{\rm S}^{\rm LOCC}(\rho_0,\rho_1,\varepsilon) = \log(d+1), \quad \forall \varepsilon \in [0,1).$$

(b) Consider the binary hypothesis testing: $\rho_0 = \Phi_d^\perp$ and $\rho_1 = \Phi_d$. Then,

$$\xi_{\rm S}^{\rm PTT}(\rho_0,\rho_1,\varepsilon) = \xi_{\rm S}^{\rm SEP}(\rho_0,\rho_1,\varepsilon) = \xi_{\rm S}^{\rm LOCC}(\rho_0,\rho_1,\varepsilon) = \infty, \quad \forall \varepsilon \in [0,1).$$

Proof. (a) By Theorem 10, we have, for every $\varepsilon \in [0,1)$,

$$\xi_{\rm S}^{\rm LOCC}(\Phi_d^{\otimes n}, (\Phi_d^\perp)^{\otimes n}, \varepsilon) = \lim_{n\to\infty} \beta_\varepsilon^{\rm X}(n) = \lim_{n\to\infty} \log(d+1) - \frac{1}{n}\log(1-\varepsilon) = \log(d+1)$$

as desired.

(b) For every $\varepsilon \in (0,1)$, we choose the test $T_n := \mathbb{1} - M_d^{\otimes n}$. From (25), (35) and (36), we have

$$\alpha_n(T_n) = \frac{1}{(d+1)^n}; \quad \beta_n(T_n) = 0.$$

Hence, it follows that $\lim_{n\to 0}\alpha_n(T_n) \le \varepsilon$, and

$$\xi_{\rm S}^{\rm X}(\rho_0,\rho_1,\varepsilon) \ge \lim_{n\to\infty} -\frac{1}{n}\log\beta_n(T_n) = \infty.$$

This completes the proof. ☐

Remark 4.2. In [54], Brandão *et al.* proved that the Stein exponent is given by the regularized relative entropy between the measurement outcome when $\varepsilon = 0$, i.e. for every class of measurements 'X',

$$\lim_{\varepsilon\to 0}\xi_{\rm S}^{\rm X}(\rho_0,\rho_1,\varepsilon) = \lim_{n\to\infty}\sup_{\mathcal{M}\in {\rm X}} \frac{D(\mathcal{M}(\rho_0^{\otimes n})\|\mathcal{M}(\rho_1^{\otimes n}))}{n},$$

where D is the quantum relative entropy [32], and $\mathcal{M}$ is any POVM in the class 'X'. Firstly, calculating such a regularized quantity is computationally intractable. Secondly, the author did not prove whether the strong converse property holds. Namely, if $\xi_{\rm S}^{\rm X}(\rho_0,\rho_1,\varepsilon)$ is dependent on $\varepsilon \in (0,1)$. Hence, in Corollary 11, we establish a single-letter formula for testing Φ_d against $\Phi_d^\perp$, and further prove the strong converse property.

Corollary 12 (Hoeffding exponent)**.** *The following Hoeffding exponents hold.*

(a) Consider the binary hypothesis testing: $\rho_0 = \Phi_d$ and $\rho_1 = \Phi_d^\perp$. Then,

$$\xi_{\rm H}^{\rm PTT}(\rho_0,\rho_1,r) = \xi_{\rm H}^{\rm SEP}(\rho_0,\rho_1,r) = \xi_{\rm H}^{\rm LOCC}(\rho_0,\rho_1,r) = \infty, \quad \forall r \le \log(d+1).$$

(b) Consider the binary hypothesis testing: $\rho_0 = \Phi_d^\perp$ and $\rho_1 = \Phi_d$. Then,

$$\xi_{\rm H}^{\rm PTT}(\rho_0,\rho_1,r) = \xi_{\rm H}^{\rm SEP}(\rho_0,\rho_1,r) = \xi_{\rm H}^{\rm LOCC}(\rho_0,\rho_1,r) = \log(d+1), \quad \forall r > 0.$$

Proof. (a) From Theorem 10, we know that if the type-II error is allowed to decay exponentially at the speed of $\log(d+1)$, then the type-I error is zero for all n. From the definition of the Hoeffding exponent given in Eq. (10), the type-I is always zero if the type-II error decays slower, which shows that the Hoeffding exponent is infinite.

(b) Fix an arbitrary n. From Theorem 10, (35), (36), choosing the test as $\mathbb{1} - (1-\varepsilon)M_d^{\otimes n}$, it follows that

$$\alpha_n(T_n) = \frac{1-\varepsilon}{(d+1)^n}, \quad \beta_n(T_n) = \varepsilon.$$

Note that the test is optimal for every $\varepsilon \in [0,1)$. Now, letting $\varepsilon = \exp\{-nr\}$ for any $r > 0$, we have $\alpha_n(T_n) = (1-\exp\{-nr\})$. Since this holds for every n, then

$$\xi_{\rm H}^{\rm LOCC}(\rho_0,\rho_1,r) = \lim_{n\to\infty} -\frac{1}{n}\log\alpha_n(T_n) = \log(d+1) - \lim_{n\to\infty}\frac{1}{n}\log(1-\exp\{-nr\}) = \log(d+1),$$

and we are done. ☐

Corollary 13 (Strong converse exponent)**.** *The following strong converse exponents hold.*

(a) *Consider the binary hypothesis testing:* $\rho_0 = \Phi_d$ *and* $\rho_1 = \Phi_d^{\perp}$*. Then,*

$$\xi_{\mathrm{SC}}^{\mathrm{PTT}}(\rho_0, \rho_1, r) = \xi_{\mathrm{SC}}^{\mathrm{SEP}}(\rho_0, \rho_1, r) = \xi_{\mathrm{SC}}^{\mathrm{LOCC}}(\rho_0, \rho_1, r) = r - \log(d+1), \quad \forall r > \log(d+1).$$

(b) *Consider the binary hypothesis testing:* $\rho_0 = \Phi_d^{\perp}$ *and* $\rho_1 = \Phi_d$*. Then,*

$$\xi_{\mathrm{SC}}^{\mathrm{PTT}}(\rho_0, \rho_1, r) = \xi_{\mathrm{SC}}^{\mathrm{SEP}}(\rho_0, \rho_1, r) = \xi_{\mathrm{SC}}^{\mathrm{LOCC}}(\rho_0, \rho_1, r) = 0, \quad \forall r \geq 0.$$

Proof. (a) By denoting $\alpha_\beta^{\mathrm{X}} := \inf\left\{\alpha : (\alpha, \beta) \in \mathcal{R}^{\mathrm{X}}(\rho_0^{\otimes n}, \rho_1^{\otimes n})\right\}$, Theorem 10 implies that

$$\beta = \frac{1 - \alpha_\beta^{\mathrm{X}}}{(d+1)^n}.$$

Letting $\beta = \exp\{-nr\}$, we have the desired strong converse exponent.

(b) By Corollary 12-(b), the type-I error will exponentially decay when the type-II error exponentially decay at any finite rate r. This thus implies that the strong converse exponent is zero for all finite rate r since $\lim_{n\to\infty} \frac{1}{n} \log[1 - \frac{1}{(d+1)^n}] = 0$. □

We remark that by following our argument in Theorem 10, one can also obtain the optimal trade-off:

$$\beta_\alpha^{\mathrm{X}}\left(\sigma_d^{\otimes n}, (\sigma_d^{\perp})^{\otimes n}\right) = (1-\alpha)\left(\frac{d-1}{d+1}\right)^n,$$

for testing completely symmetric state σ_d and completely anti-symmetric state $\sigma_d^{\perp}$ [74]. Here, the optimal LOCC measurement is achieved by $(1-\alpha) \cdot \bar{M}_m^{\otimes n}$, where $\bar{M}_m := \frac{m-1}{m+1}\Pi_s + \Pi_a$ as in equation (13) of [53]; i.e. Π_s and Π_a are the projections onto the support of the symmetric and anti-symmetric subspaces, respectively.

Accordingly, the associated Stein, Hoeffding, and the strong converse exponents can be established similarly as in Corollaries 11, 12, and 13. We summaries those results in Table 1.

Remark 4.3. From Corollaries 11 and 13, we know that shows that our converse bounds for testing arbitrary multipartite entangled pure state against its orthogonal complement given in Proposition 4 and Corollary 5 of Section 3.2 are both tight for maximally entangled state Φ_d.

5. Testing Pure States with Uniform Non-Zero Schmidt Coefficients

In this section, we aim to test pure entangled states with equal positive Schmidt coefficients, i.e.

$$\rho_0 := \frac{1}{m}\sum_{i,j=0}^{m-1} |ii\rangle\langle jj| \quad \text{for } m \leq d$$

against its orthogonal complement $(\mathbb{1}_{d^2} - \rho_0)/(d^2-1)$. Equivalently, such state can be written as a maximally entangled state Φ_m on $\mathbb{C}^m \otimes \mathbb{C}^m$ that is embedded in a higher dimension Hilbert space $\mathbb{C}^d \otimes \mathbb{C}^d$. In other words, we consider the binary hypotheses

$$\begin{cases} \mathsf{H}_0 : \rho_0^{\otimes n} = (\Phi_m \oplus \mathbb{O}_{d^2-m^2})^{\otimes n} \\ \mathsf{H}_1 : \rho_1^{\otimes n} = \left(\dfrac{\mathbb{1}_{d^2} - \Phi_m \oplus \mathbb{O}_{d^2-m^2}}{d^2-1}\right)^{\otimes n} \end{cases}, \quad \forall n. \tag{38}$$

In this case, the analysis of Proposition 7 in the previous section does not directly apply. To circumvent this problem, we prove a useful converse bound below, which will be used for our analysis later.

Lemma 14. *Fix dimensions d and d'. Let ρ and σ be density operators on $\mathbb{C}^d \otimes \mathbb{C}^d$. Then, for any class of two-outcome POVM,* $\mathrm{X} \in \{\mathrm{LOCC}, \mathrm{SEP}, \mathrm{PPT}, \mathrm{ALL}\}$*, any $p, \lambda \in [0,1]$ and any natural number n, it follows that*

$$P_{\mathrm{e}}^{\mathrm{X}}\left((\rho \oplus \mathbb{O}_{d'^2})^{\otimes n}, (\lambda\sigma \oplus (1-\lambda)\tau_{d'^2})^{\otimes n}; p\right) \geq P_{\mathrm{e}}^{\mathrm{X}}\left(\rho^{\otimes n}, (\lambda\sigma)^{\otimes n}; p\right).$$

Proof. Recalling the definition of optimal average error probability for any (probably sub-normalized) operators, i.e.

$$P_{\mathrm{e}}(\rho_0,\rho_1;p) := \inf_T \left\{p\alpha(T) + (1-p)\beta(T)\right\},$$

it follows that

$$\begin{aligned}
&P_{\mathrm{e}}^{\mathrm{X}}\left((\rho\oplus \mathbb{0}_{d'^2})^{\otimes n}, \left(\lambda\rho^{\perp}\oplus(1-\lambda)\tau_{d'^2}\right)^{\otimes n}; p\right)\\
&= \inf_{T_n\in\mathrm{X}} \left\{p\operatorname{Tr}\left[(\mathbb{1}-T_n)(\rho\oplus\mathbb{0}_{d'^2})^{\otimes n}\right] + (1-p)\operatorname{Tr}\left[T_n\left(\lambda\sigma\oplus(1-\lambda)\tau_{d'^2}\right)^{\otimes n}\right]\right\}\\
&\geq \inf_{T_n\in\mathrm{X}} \left\{p\operatorname{Tr}\left[(\mathbb{1}-T_n)\rho^{\otimes n}\right] + (1-p)\operatorname{Tr}\left[T_n\left(\lambda\sigma\right)^{\otimes n}\right]\right\}\\
&\quad+ \inf_{G_n\in\mathrm{X}} \left\{p\operatorname{Tr}\left[(\mathbb{1}-G_n)\mathbb{0}_{d'^2}\right] + (1-p)\operatorname{Tr}\left[G_n\sum_{k=0}^{n-1}\lambda^k(1-\lambda)^{1-k}B_k\right]\right\},
\end{aligned} \tag{39}$$

where we have used super-additivity of infimum, and B_k denotes the sum of all elements of $\{\sigma,\tau_{d'^2}\}^{\otimes n}$ which have k copies of σ. Note that the second term in (39) is zero since one allows to choose the projection $\mathbb{1}_{d'^2}$ onto the copy of $\tau_{d'^2}$ as the POVM and, further, is implementable by all the four classes of POVMS. This then completes the proof. □

In the following proposition, we consider a more general case compared to Eq. (38), since there are non-unique orthogonal complements in high-dimensional systems.

Proposition 15. *Consider the hypotheses*

$$\begin{cases}\mathsf{H}_0:\rho_0^{\otimes n} = (\Phi_m\oplus\mathbb{0}_{d^2-m^2})^{\otimes n}\\ \mathsf{H}_1:\rho_1^{\otimes n} = \left(\lambda\Phi_m^{\perp}\oplus(1-\lambda)\tau_{d^2-m^2}\right)^{\otimes n}\end{cases},\quad \forall n. \tag{40}$$

Then, it holds that

$$P_{\mathrm{e}}^{\mathrm{PPT}}(\rho_0^{\otimes n},\rho_1^{\otimes n};p) = P_{\mathrm{e}}^{\mathrm{SEP}}(\rho_0^{\otimes n},\rho_1^{\otimes n};p) = P_{\mathrm{e}}^{\mathrm{LOCC}}(\rho_0^{\otimes n},\rho_1^{\otimes n};p) = \min\left\{(1-p)\left(\frac{\lambda}{m+1}\right)^n, p\right\}.$$

Proof. We first prove the achievability, i.e. the "≥" direction. We choose the POVM as

$$\begin{aligned}
&\left\{(M_m\oplus\mathbb{0}_{d^2-m^2})^{\otimes n}, \mathbb{1}_{d^{2n}} - (M_m\oplus\mathbb{0}_{d^2-m^2})^{\otimes n}\right\}\\
&= \left\{M_m^{\otimes n}\oplus\mathbb{0}_{d^{2n}-m^{2n}}, \left(\mathbb{1}_{m^{2n}} - M_m^{\otimes n}\right)\oplus\mathbb{1}_{d^{2n}-m^{2n}}\right\},
\end{aligned}$$

where $M_m = \Phi_m + \frac{1}{m+1}(\mathbb{1}_{m^2}-\Phi_m)$ as in (24).

Then, the average error probability for the chosen LOCC protocol is

$$\begin{aligned}
P_{\mathrm{e}}^{\mathrm{LOCC}} &= p\operatorname{Tr}\left[\left(\mathbb{1}_{m^{2n}} - M_m^{\otimes n}\right)\oplus\mathbb{1}_{d^{2n}-m^{2n}}\cdot\Phi_m^{\otimes n}\oplus\mathbb{0}_{d^{2n}-m^{2n}}\right]\\
&\quad+(1-p)\operatorname{Tr}\left[M_m^{\otimes n}\oplus\mathbb{0}_{d^{2n}-m^{2n}}\cdot\left(\lambda\Phi_m^{\perp}\oplus(1-\lambda)\tau_{d^2-m^2}\right)^{\otimes n}\right]\\
&= (1-p)\operatorname{Tr}\left[M_m^{\otimes n}\left(\lambda\Phi_m^{\perp}\right)^n\right]\\
&= (1-p)\left(\frac{m^2-1}{d^2-1}\right)^n\left(\frac{\lambda}{m+1}\right)^n.
\end{aligned}$$

Similar to the proof of Proposition 6, we choose $\rho_1^{\otimes n} = \left(\frac{\mathbb{1}_{d^2}-\Phi_m\oplus\mathbb{0}_{d^2-m^2}}{d^2-1}\right)^{\otimes n}$ whenever $P_{\mathrm{e}}^{\mathrm{LOCC}} > p$. Hence, we complete the proof of achievability.

Next, we move on to prove the converse, i.e. the "≥" direction. Invoking Lemma 14, we have

$$P_{\mathrm{e}}^{\mathrm{PPT}}(\rho_0^{\otimes n},\rho_1^{\otimes n};p) \geq P_{\mathrm{e}}^{\mathrm{PPT}}(\Phi^{\otimes n},(\lambda\Phi_m^{\perp})^{\otimes n};p).$$

Following a similar argument as in Proposition 7 obtains the desired result. □

Remark 5.1. Choosing $\lambda = \frac{m^2-1}{d^2-1}$ in (40), it can be verified that the single copy in the alternative hypothesis H_1 coincides the canonical orthogonal complement in the alternative hypothesis of (38), namely $\frac{\mathbb{1}_{d^2}-\Phi_m \oplus \mathbb{0}_{d^2-m^2}}{d^2-1}$. This answers the optimal average error probability of binary hypotheses (38) considered at the beginning of this section, i.e. $\min\left\{(1-p)\left(\frac{m-1}{d^2-1}\right)^n, p\right\}$.

The binary hypotheses considered in Eq. (40) have a simple variant (by interchanging Φ_m and $\Phi_m^\perp$), for which we can immediately calculate its optimal average error.

Proposition 16. *Consider the binary hypotheses*

$$\begin{cases} \mathsf{H}_0 : \rho_0^{\otimes n} = \left(\Phi_m^\perp \oplus \mathbb{0}_{d^2-m^2}\right)^{\otimes n}, \\ \mathsf{H}_1 : \rho_1^{\otimes n} = \left(\lambda \Phi_m \oplus (1-\lambda)\tau_{d^2-m^2}\right)^{\otimes n} \end{cases}, \quad \forall n,$$

where $\lambda \in [0,1]$. *Then, it holds that*

$$P_{\mathrm{e}}^{\mathrm{PPT}}(\rho_0^{\otimes n}, \rho_1^{\otimes n}; p) = P_{\mathrm{e}}^{\mathrm{SEP}}(\rho_0^{\otimes n}, \rho_1^{\otimes n}; p) = P_{\mathrm{e}}^{\mathrm{LOCC}}(\rho_0^{\otimes n}, \rho_1^{\otimes n}; p) = \min\left\{p\left(\frac{1}{m+1}\right)^n, (1-p)\lambda^n\right\}.$$

Proof. We choose the POVM as

$$\left\{\left(\mathbb{1}_{m^{2n}} - M_m^{\otimes n}\right) \oplus \mathbb{0}_{d^{2n}-m^{2n}}, M_m^{\otimes n} \oplus \mathbb{1}_{d^{2n}-m^{2n}}\right\}, \tag{41}$$

where $M_m = \Phi_m + \frac{1}{m+1}(\mathbb{1}_{m^2} - \Phi_m)$ as defined in (24).

The average error probability for the chosen LOCC protocol is

$$\begin{aligned} P_{\mathrm{e}}^{\mathrm{LOCC}} &= p \operatorname{Tr}\left[M_m^{\otimes n} \oplus \mathbb{1}_{d^{2n}-m^{2n}} \cdot (\Phi_m^\perp)^{\otimes n} \oplus \mathbb{0}_{d^{2n}-m^{2n}}\right] \\ &\quad + (1-p)\operatorname{Tr}\left[\left(\mathbb{1}_{m^{2n}} - M_m^{\otimes n}\right) \oplus \mathbb{0}_{d^{2n}-m^{2n}} \cdot \left(\lambda\Phi_m \oplus (1-\lambda)\tau_{d^2-m^2}\right)^{\otimes n}\right] \\ &= p \operatorname{Tr}\left[M_m^{\otimes n}(\Phi_m^\perp)^{\otimes n}\right] + (1-p)\operatorname{Tr}\left[\left(\mathbb{1}_{m^{2n}} - \Phi_m^{\otimes n}\right)\cdot(\lambda\Phi_m)^{\otimes n}\right] \\ &= p \operatorname{Tr}\left[M_m^{\otimes n}(\Phi_m^\perp)^{\otimes n}\right] + (1-p)\lambda^n \\ &= p\left(\frac{1}{m+1}\right)^n. \end{aligned} \tag{42}$$

On the other hand, we can also choose the POVM as

$$\left\{\mathbb{1}_{m^{2n}} \oplus \mathbb{0}_{d^{2n}-m^{2n}}, \mathbb{0}_{m^{2n}} \oplus \mathbb{1}_{d^{2n}-m^{2n}}\right\}. \tag{43}$$

Then, the corresponding average error probability is

$$\begin{aligned} P_{\mathrm{e}}^{\mathrm{LOCC}} &= p \operatorname{Tr}\left[\mathbb{0}_{m^{2n}} \oplus \mathbb{1}_{d^{2n}-m^{2n}} \cdot (\Phi_m^\perp)^{\otimes n} \oplus \mathbb{0}_{d^{2n}-m^{2n}}\right] \\ &\quad + (1-p)\operatorname{Tr}\left[\mathbb{1}_{m^{2n}} \oplus \mathbb{0}_{d^{2n}-m^{2n}} \cdot \left(\lambda\Phi_m \oplus (1-\lambda)\tau_{d^2-m^2}\right)^{\otimes n}\right] \\ &= (1-p)\operatorname{Tr}\left[\mathbb{1}_{m^{2n}} \cdot (\lambda\Phi_m)^{\otimes n}\right] \\ &= (1-p)\lambda^n. \end{aligned} \tag{44}$$

Since both the chosen POVMs in (41) and (43) are implementable by LOCC protocols, we minimize the average error probabilities given in (42) and (44) to arrive at the "$\leq$" direction of our result.

For the other direction, i.e. "$\geq$", we use Lemma 14, to obtain

$$P_{\mathrm{e}}^{\mathrm{PPT}}(\rho_0^{\otimes n}, \rho_1^{\otimes n}; p) \geq P_{\mathrm{e}}^{\mathrm{PPT}}((\Phi_m^\perp)^{\otimes n}, (\lambda\Phi_m)^{\otimes n}; p).$$

Following a similar argument as in Proposition 7 obtains the desired result. □

In the following Propositions 17 and 18, we apply similar techniques as before to calculate the binary hypothesis of a symmetric state embedded in a high-dimensional system and its orthogonal complement.

Proposition 17. *Consider the binary hypothesis*

$$\begin{cases} \mathsf{H}_0 : \rho_0^{\otimes n} = (\sigma_m \oplus \mathbb{0}_{d^2-m^2})^{\otimes n}, \\ \mathsf{H}_1 : \rho_1^{\otimes n} = \left(\lambda\sigma_m^{\perp} \oplus (1-\lambda)\tau_{d^2-m^2}\right)^{\otimes n} \end{cases}, \quad \forall n,$$

where $\lambda \in [0,1]$ is arbitrary. Then, it holds that

$$P_{\mathrm{e}}^{\mathrm{LOCC}}(\rho_0^{\otimes n}, \rho_1^{\otimes n}; p) = \min\left\{ p\left(\frac{m-1}{m+1}\right)^n, (1-p)\lambda^n \right\}.$$

Proof. The proof follows similar reasoning as in Theorem 15.

We choose the POVM as

$$\left\{ \left(\mathbb{1}_{m^{2n}} - \bar{M}_m^{\otimes n}\right) \oplus \mathbb{0}_{d^{2n}-m^{2n}}, \bar{M}_m^{\otimes n} \oplus \mathbb{1}_{d^{2n}-m^{2n}} \right\},$$

where $\bar{M}_m := \frac{m-1}{m+1}\Pi_s + \Pi_a$ as in equation (13) of [53]; Π_s and Π_a are the projections onto the support of the symmetric and anti-symmetric subspaces, respectively.

Then, the average error probability for the chosen LOCC protocol is

$$\begin{aligned} P_{\mathrm{e}}^{\mathrm{LOCC}} &= p \operatorname{Tr}\left[\bar{M}_m^{\otimes n} \oplus \mathbb{0}_{d^{2n}-m^{2n}} \cdot \sigma_m^{\otimes n} \oplus \mathbb{1}_{d^{2n}-m^{2n}}\right] \\ &\quad + (1-p)\operatorname{Tr}\left[\left(\mathbb{1}_{m^{2n}} - M_m^{\otimes n}\right) \oplus \mathbb{0}_{d^{2n}-m^{2n}} \cdot \left(\lambda\sigma_m^{\perp} \oplus (1-\lambda)\tau_{d^2-m^2}\right)^{\otimes n}\right] \\ &= p \operatorname{Tr}\left[\bar{M}_m^{\otimes n}\sigma_m^{\otimes n}\right] \\ &= p\left(\frac{m-1}{m+1}\right)^n \end{aligned}$$

where we invoke equation (17) of [53] in the last line. On the other hand, we can also choose the POVM as in (43) to obtain $P_{\mathrm{e}}^{\mathrm{LOCC}} = (1-p)\lambda^n$. We then choose the minimum of the two to complete the achievability.

The converse follows from Lemma 14 and [53, Proposition 3]. □

Proposition 18. *Consider the binary hypothesis*

$$\begin{cases} \mathsf{H}_0 : \rho_0^{\otimes n} = \left(\sigma_m^{\perp} \oplus \mathbb{0}_{d^2-m^2}\right)^{\otimes n}, \\ \mathsf{H}_1 : \rho_1^{\otimes n} = (\lambda\sigma_m \oplus (1-\lambda)\tau_{d^2-m^2})^{\otimes n} \end{cases}, \quad \forall n,$$

for some $\lambda \in [0,1]$. Then, it holds that

$$P_{\mathrm{e}}^{\mathrm{LOCC}}(\rho_0^{\otimes n}, \rho_1^{\otimes n}; p) = \min\left\{ (1-p)\left(\lambda\frac{m-1}{m+1}\right)^n, p \right\}.$$

Proof. We choose the POVM as

$$\left\{ \bar{M}_m^{\otimes n} \oplus \mathbb{0}_{d^{2n}-m^{2n}}, \left(\mathbb{1}_{m^{2n}} - \bar{M}_m^{\otimes n}\right) \oplus \mathbb{1}_{d^{2n}-m^{2n}} \right\},$$

where $\bar{M}_m := \frac{m-1}{m+1}\Pi_s + \Pi_a$ as in equation (13) of [53] and the proof of Theorem 17.

Then, the average error probability for the chosen LOCC protocol is

$$\begin{aligned} P_{\mathrm{e}}^{\mathrm{LOCC}} &= p \operatorname{Tr}\left[\left(\mathbb{1}_{m^{2n}} - \bar{M}_m^{\otimes n}\right) \oplus \mathbb{1}_{d^{2n}-m^{2n}} \cdot \sigma_m^{\otimes n} \oplus \mathbb{0}_{d^{2n}-m^{2n}}\right] \\ &\quad + (1-p)\operatorname{Tr}\left[\bar{M}_m^{\otimes n} \oplus \mathbb{0}_{d^{2n}-m^{2n}} \cdot \left(\lambda\sigma_m \oplus (1-\lambda)\tau d^2 - m^2\right)^{\otimes n}\right] \\ &= (1-p)\operatorname{Tr}\left[\bar{M}_m^{\otimes n}\lambda^n\sigma_m^{\otimes n}\right] \\ &= (1-p)\left(\lambda\frac{m-1}{m+1}\right)^n. \end{aligned}$$

Lastly, we choose $\rho_1^{\otimes n}$ whenever $P_{\mathrm{e}}^{\mathrm{LOCC}} > p$.

The converse follows Lemma 14 and [53, Proposition 3]. □

Remark 5.2. The Stein, Hoeffding, and the strong converse exponents can be obtained by following the same arguments in Section 5. We refer the readers to Table 1 for the summary of those results.

6. Infinite separation between SEP POVMs and PPT POVMs

In this section, we prove that there is an infinite separation between the optimal average error probabilities using SEP POVMs and PPT POVMs.

Definition 19 (Unextendible product basis [76, 77])**.** Consider two Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$. A set $S = \{|\alpha_i\rangle \otimes |\beta_i\rangle : 1 \le i \le N\} \subseteq \mathcal{H}_A \otimes \mathcal{H}_B$ is called an *unextendible product basis*[1] (UPB) if it satisfies the following:

$$\begin{cases} |\alpha_i\rangle \otimes |\beta_i\rangle \perp |\alpha_j\rangle \otimes |\beta_j\rangle, \quad \forall i \neq j \\ \{|\phi\rangle \otimes |\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B : \forall i \in \{1, \dots, N\}, |\phi\rangle \otimes |\psi\rangle \perp |\alpha_i\rangle \otimes |\beta_i\rangle\} = 0 \end{cases}.$$

For a UPB $S = \{|\alpha_i\rangle \otimes |\beta_i\rangle : 1 \le i \le N\} \subseteq \mathcal{H}_A \otimes \mathcal{H}_B$, we consider the following binary hypotheses:

$$\begin{cases} \mathsf{H}_0 : \rho^{\otimes n} := \left[\frac{1}{N}\left(\sum_{i=1}^{N} |\alpha_i\beta_i\rangle\langle\alpha_i\beta_i|\right)\right]^{\otimes n} \\ \mathsf{H}_1 : \sigma^{\otimes n} \end{cases}, n \in \mathbb{N}. \tag{45}$$

Here, σ is a state in $\mathcal{H}_A \otimes \mathcal{H}_B$ that is orthogonal to ρ. We show below that the optimal exponential decaying rate of using SEP POVMs is strictly worse than that of using PPT POVMs.

Theorem 20 (Separation)**.** *There is an infinite separation between SEP POVMs and PPT POVMs. That is, for testing the states given in* (45)*, there exists a* $\mu > 0$ *such that*

$$\begin{cases} P_{\mathrm{e}}^{\mathrm{PPT}}(n) = 0 \\ P_{\mathrm{e}}^{\mathrm{SEP}}(n) \ge \frac{\mu^n}{2} \end{cases}, \quad \forall n \in \mathbb{N}.$$

The key ingredient to establish Theorem 20 is to introduce a novel quantity that characterizes the "richness" of a product basis:

Definition 21. Given a product basis $S = \{|\alpha_i\rangle \otimes |\beta_i\rangle : 1 \le i \le N\} \subseteq \mathcal{H}_A \otimes \mathcal{H}_B$, we define

$$\delta_S := \min_{\rho_A, \rho_B} \max_{1 \le i \le N} \langle\alpha_i|\rho_A|\alpha_i\rangle\langle\beta_i|\rho_B|\beta_i\rangle, \tag{46}$$

where ρ_A (resp. ρ_B) ranges over all density matrices on $\mathcal{H}_A$ (resp. $\mathcal{H}_B$).

Note that the minimum in (46) is always attained for some ρ_A and ρ_B, since the objective function on the righthand side is continuous in $\rho_A \otimes \rho_B$ and the set of product states $\rho_A \otimes \rho_B$ is compact.

For an unextendible product basis S [77], one immediately has $\delta_S > 0$.

Proposition 22. *The quantity* δ_S *is multiplicative. That is,*

$$\delta_{S_1 \otimes S_2} = \delta_{S_1}\delta_{S_2}$$

for any two product bases S_1 *and* S_2*.*

Remark 6.1. Note that a product basis S is unextendible if and only if $\delta_S > 0$. Hence, the quantity δ_S indicates how far a UPB is from an extendible product basis. Then Proposition 22 implies that the tensor product of any two UPBs S_1 and S_2 enjoys the property that $\delta_{S_1 \otimes S_2} = \delta_{S1}\delta_{S_2} > 0$. This thus gives a quantitative characterization of the well-known fact that the tensor product of UPBs is also UPB [77].

Proof of Proposition 22. Let two product bases be given as

$$S_1 = \{|\alpha_i\rangle \otimes |\beta_i\rangle : 1 \le i \le N_1\} \subseteq \mathcal{H}_{A_1} \otimes \mathcal{H}_{B_1},$$
$$S_2 = \{|\psi_j\rangle \otimes |\phi_j\rangle : 1 \le j \le N_2\} \subseteq \mathcal{H}_{A_2} \otimes \mathcal{H}_{B_2}.$$

From the definition given in Eq. (46), one can directly verify that

$$\delta_{S_1 \otimes S_2} \le \delta_{S_1}\delta_{S_2}.$$

[1]In this paper we only considered unextendible product bases on a bipartite Hilbert spaces. However, the Definition 19 naturally extends to the multipartite scenario.

It remains to prove the other direction.

Let $\rho_{A_1A_2}$ on $\mathcal{H}_{A_1} \otimes \mathcal{H}_{A_2}$ and $\rho_{B_1B_2}$ on $\mathcal{H}_{B_1} \otimes \mathcal{H}_{B_2}$ attain the minimization in $\delta_{S_1\otimes S_2}$, i.e.

$$\delta_{S_1\otimes S_2} = \max_{\substack{1\leq i\leq N_1\\ 1\leq j\leq N_2}} \langle \alpha_i\psi_j|\rho_{A_1A_2}|\alpha_i\psi_j\rangle\langle \beta_i\phi_j|\rho_{B_1B_2}|\beta_i\phi_j\rangle. \tag{47}$$

Further, for any $1 \leq i \leq N_1$, we let

$$\langle\alpha_i|\rho_{A_1A_2}|\alpha_i\rangle = \lambda_i\sigma_{i,A_2},$$
$$\langle\beta_i|\rho_{B_1B_2}|\beta_i\rangle = \mu_i\sigma_{i,B_2},$$

where

$$\lambda_i := \mathrm{Tr}\left[\langle\alpha_i|\rho_{A_1A_2}|\alpha_i\rangle\right] = \langle\alpha_i|\rho_{A_1}|\alpha_i\rangle \geq 0,$$
$$\mu_i := \mathrm{Tr}\left[\langle\beta_i|\rho_{B_1B_2}|\beta_i\rangle\right] = \langle\beta_i|\rho_{B_1}|\beta_i\rangle \geq 0.$$

By the definition of δ_1, there always exists some i_0 such that

$$\lambda_{i_0}\mu_{i_0} \geq \delta_{S_1}. \tag{48}$$

For such choice of i_0, for any $1 \leq j \leq N_2$, we have

$$\langle\alpha_{i_0}\psi_j|\rho_{A_1A_2}|\alpha_{i_0}\psi_j\rangle\langle\beta_{i_0}\phi_j|\rho_{B_1B_2}|\beta_{i_0}\phi_j\rangle = \lambda_{i_0}\mu_{i_0}\langle\psi_j|\sigma_{i_0,A_2}\psi_j\rangle\langle\phi_j|\sigma_{i_0,B_2}\phi_j\rangle.$$

By the definition of δ_{S_2}, there exists some j_0 such that

$$\langle\psi_{j_0}|\sigma_{i_0,A_2}|\psi_{j_0}\rangle\langle\phi_{j_0}|\sigma_{i_0,B_2}|\phi_{j_0}\rangle \geq \delta_{S_2}. \tag{49}$$

Combining (47), (48), and (49) gives

$$\delta_{S_1\otimes S_2} \geq \delta_{S_1}\delta_{S_2},$$

which completes the proof. □

Now, we are ready to prove the main result, Theorem 20, of this section.

Proof of Theorem 20. Let us consider the binary hypotheses given in (45) and let

$$P = \sum_{i=1}^{N} |\alpha_i\beta_i\rangle\langle\alpha_i\beta_i|, \quad \rho = \frac{1}{N}P.$$

By construction, both P and $\mathbb{1} - P$ are PPT operators [77]. Hence, the pair of states (ρ, σ) can be perfectly distinguished by PPT POVMs. Hence, it remains to show an exponential lower bound to the error probability for distinguishing $(\rho^{\otimes n}, \sigma^{\otimes n})_{n\in\mathbb{N}}$ using SEP POVMs.

Let us consider uniform prior. Then,

$$P_{\mathrm{e}}^{\mathrm{SEP}}\left(\rho^{\otimes n}, \sigma^{\otimes n}; \tfrac{1}{2}\right) = 1 - \sup_{M,(\mathbb{1}-M)\in\mathrm{SEP}} \left\{\frac{1}{2}\mathrm{Tr}\left[\rho^{\otimes n}M\right] + \frac{1}{2}\mathrm{Tr}\left[\sigma^{\otimes n}(\mathbb{1}-M)\right]\right\}.$$

Note that above can be written as a primal problem of a *semi-definite program*:

$$\begin{aligned} \text{maximize}: \quad & \frac{1}{2}\mathrm{Tr}\left[\rho^{\otimes n}\Pi_0\right] + \frac{1}{2}\mathrm{Tr}\left[\sigma^{\otimes n}\Pi_1\right]; \\ \text{subject to}: \quad & \Pi_0 + \Pi_1 = \mathbb{1}, \\ & \Pi_i \in \mathrm{SEP} \quad \forall i \in \{0,1\}. \end{aligned}$$

To derive its dual problem, we introduce the *dual cone* to SEP as follows:

$$\mathrm{SEP}^*(A^n : B^n) := \left\{H : H^\dagger = H,\ \mathrm{Tr}\left[\Pi H\right] \geq 0, \quad \forall \Pi \in \mathrm{SEP}\right\}.$$

Such dual cone is also known as the set of *block-positive operators* [78, Section 2], i.e.

$$\mathrm{SEP}^*(A^n : B^n) = \left\{H : H^\dagger = H,\ (\mathbb{1}_{X^n} \otimes \langle y|)H(\mathbb{1}_{X^n} \otimes |y\rangle) \geq 0, \quad \forall |y\rangle \in \mathcal{H}_B^{\otimes n}\right\}.$$

Then the associated dual problem is:

$$\begin{aligned} \text{minimize :} \quad & \mathrm{Tr}[H]; \\ \text{subject to :} \quad & H - \frac{1}{2}\rho^{\otimes n} \in \mathrm{SEP}^*(A^n : B^n), \\ & H - \frac{1}{2}\sigma^{\otimes n} \in \mathrm{SEP}^*(A^n : B^n), \\ & H = H^\dagger. \end{aligned}$$

By the weak duality, we have

$$P_{\mathrm{e}}^{\mathrm{SEP}}\left(\rho^{\otimes n}, \sigma^{\otimes n}; \tfrac{1}{2}\right) \geq 1 - \inf_{H^\dagger = H} \left\{ \mathrm{Tr}[H] : H - \frac{1}{2}\rho^{\otimes n}, H - \frac{1}{2}\sigma^{\otimes n} \in \mathrm{SEP}^*(A^n : B^n) \right\}. \tag{50}$$

Now, we choose a Hermitian operator H as

$$\begin{aligned} H &= \frac{1}{2}\rho^{\otimes n} + \left(\frac{1}{2} - \frac{\mu^n}{2}\right)\sigma^{\otimes n}, \\ \mu &:= \frac{\delta_S}{N} \in (0,1), \end{aligned}$$

where δ_S was introduced in Definition 21, and it is clear that $\delta_S \leq 1$. We aim to show that the operator H is a feasible solution to (50) to complete the proof. Firstly, it is not hard to see that

$$H - \frac{1}{2}\rho^{\otimes n} = \left(\frac{1}{2} - \frac{\mu^n}{2}\right)\sigma^{\otimes n} \in \mathrm{SEP}^*(A^n : B^n),$$

since positivity implies block-positivity. Secondly, we have

$$\begin{aligned} H - \frac{1}{2}\sigma^{\otimes n} &= \frac{1}{2}\left(\rho^{\otimes n} - \mu^n \sigma^{\otimes n}\right) \\ &\geq \frac{1}{2}\left(\rho^{\otimes n} - \mu^n \mathbb{1}^{\otimes n}\right) \\ &= \frac{1}{2N^n}\left(P^{\otimes n} - \delta_S^n \mathbb{1}^{\otimes n}\right). \end{aligned}$$

To show that the quantity $P^{\otimes n} - \delta_S^n \mathbb{1}^{\otimes n}$ is block-positive, we will invoke the definition of δ_S given in Definition 21 and the multiplicativity of δ_S established in Proposition 22. Note that for all τ_A on $\mathcal{H}_A$ and τ_B on $\mathcal{H}_B$, we have

$$\mathrm{Tr}\left[P(\tau_A \otimes \tau_B)\right] \geq \delta_S > 0,$$

Hence, for any $\tau_{A_1 A_2 \ldots A_n} \otimes \tau_{B_1 B_2 \ldots B_n}$,

$$\mathrm{Tr}\left[P^{\otimes n} \tau_{A_1 A_2 \ldots A_n} \otimes \tau_{B_1 B_2 \ldots B_n}\right] \geq \delta_{S^{\otimes n}} = \delta_S^n > 0$$

In other words, $P^{\otimes n} - \delta_S^n \mathbb{1}^{\otimes n}$ is block-positive, which proves our claim. □

7. Discussions and Conclusions

We studied the hypothesis testing between entangled pure state against its orthogonal complement in the many-copy scenario. Our motivation is, firstly, because quantum entanglement is a valuable resource in quantum computation and other important quantum information-theoretic protocols. How to distinguish a state which possesses entangled bits is of fundamental significance. Secondly, whether the Chernoff exponent for pairs of orthogonal states under LOCC is faithful (i.e. finite) is a long-term open problem. This problem is challenging because there is no known simple mathematical structure of LOCC. Moreover, a general state could be entangled in a multipartite quantum system. Without any simple mathematical expression for it makes such hypothesis testing using LOCC more notorious.

In this paper, we show that the optimal average error probability for testing an arbitrary multipartite entangled pure state against its orthogonal complement decays exponentially in the number of copies, which then implies that the associated Chernoff exponent is faithful. In the special case of a maximally entangled state, we explicitly derive its optimal average error probability in the symmetric setting and show that the best LOCC protocol is achieved by a single-copy measurement. This then directly leads to an explicit expression of the Chernoff exponent.

In the asymmetric setting, we obtain the optimal trade-off between the type-I error α_n and the type-II error β_n. This allows us to fully characterize their asymptotic properties. The associated Stein's exponent, namely the optimal exponential rate of β_n when α_n is at most a constant, is proved. When β_n decays at a rate below or above the Stein exponent, we also establish the optimal exponential rate of α_n and $1-\alpha_n$, respectively. Our results show that the negative logarithm of the considered error does not diverge, which then guarantees that all the definitions of the four exponents are faithful. In other words, the asymptotic behavior of the errors satisfies the so-called *large deviation principle* [79]. It is worth mentioning that the second-order asymptotics under LOCC have a distinctive difference from that under the global measurements. This signifies that LOCC indeed has an exceptional structure. To be more specific, we recall the second-order expansion of the optimal exponential rate of type-II error with type-I error no larger than an constant ε under global measurements [43, 44], i.e.

$$D(\rho_0\|\rho_1) + \sqrt{\frac{V(\rho_0\|\rho_1)}{n}}Q^{-1}(\varepsilon) + O\left(\frac{\log n}{n}\right), \tag{51}$$

where D and V are the quantum relative entropy and the relative entropy variance, and Q is the cumulative distribution function of standard normal distribution. As pointed out in [44, Section 5], $V(\rho_0\|\rho_1) = 0$ implies $D(\rho_0\|\rho_1) = 0$ for any pair of quantum states ρ_0, ρ_1. In other words, the first-order term in (51) disappears whenever the second-order term vanish. On the other hand, the obtained result for LOCC, (4), shows that the optimal exponential rate of type-II error is given by

$$-\frac{1}{n}\log\beta_n = \log(d+1) - \frac{\log(1-\varepsilon)}{n}.$$

Here, the second-order term is missing; however, the first-order term $\log(d+1)$ is strictly positive. This implies that such asymptotic expansion indeed has a different second-order term from that in (51). We note that in a previous work by Hayashi and Owari [55, 56, 57], the asymptotic bounds for the case of distinguishing a bipartite pure state and a completely mixed state under LOCC admit a similar second-order expansion as in (51).

Finally, we establish an infinite separation between the SEP and PPT operations in the many-copy scenario. Our result shows that indeed there is a gap between the SEP and PPT operations no matter how many copies of states are provided. Our technique is a multiplicativity property—a quantitative characterization of the tensor product of unextendible product bases, which might be of independent interest. On the other hand, whether there is an infinite separation between the SEP and LOCC operations is an interesting open problem and is left for future work. We believe that our analysis and results might have applications in data hiding or the studies of other important sets of orthogonal states.

Acknowledgements

H.-C. Cheng is supported by the Young Scholar Fellowship (Einstein Program) of the Ministry of Science and Technology in Taiwan (R.O.C.) under grant number MOST 109-2636-E-002-001, and is supported by the Yushan Young Scholar Program of the Ministry of Education in Taiwan (R.O.C.) under grant number NTU-109V0904. A. Winter acknowledges financial support by the Spanish MINECO (projects FIS2016-86681-P and PID2019-107609GB-I00) with the support of FEDER funds, and the Generalitat de Catalunya (project CIRIT 2017-SGR-1127). N. Yu is supported by ARC Discovery Early Career Researcher Award DE180100156 and ARC Discovery Program DP210102449.

Appendix A. Proof that (32) in the proof of Proposition 7 is a feasible solution

In this section, we prove that the $\mathbf{y}^\star = \mathbf{u}^\star \oplus \mathbf{v}^\star \oplus \mathbf{w}^\star$ (32) is a feasible solution to the linear program (31).

Since

$$u_i^\star = \frac{1}{2^n}\left[1-\left(-\frac{d-1}{d+1}\right)^{n-i}\right]$$
$$\geq \frac{1}{2^n}\left[1-\left(\frac{d-1}{d+1}\right)^{n-i}\right]$$
$$\geq 0 \tag{52}$$

for all $i = 0, \ldots, n$, we have $\mathbf{u}^\star \geq 0$. Further, $\mathbf{v}^\star \geq 0$ and $\mathbf{w}^\star \geq 0$ by the choice in (32). Hence, the feasible solution $\mathbf{y}^\star$ satisfies the first three inequalities in (31).

It remains to verify that

$$Q^T \cdot \mathbf{u}^\star - Q^T \cdot \mathbf{v}^\star - \mathbf{w}^\star \leq \mathbf{c}.$$

To that end, we will show

$$(Q^T \cdot \mathbf{u}^\star)_k = \delta_{0k} - \delta_{nk}\left(\frac{1}{d+1}\right)^n, \tag{53}$$

which in turn implies that the constraint in (31) is satisfied:

$$\left(Q^T \cdot \mathbf{u}^\star - Q^T \cdot \mathbf{v}^\star - \mathbf{w}^\star\right)_k = \delta_{0k} - \delta_{nk}\left(\frac{1}{d+1}\right)^n - \delta_{nk}\max\left\{\frac{p}{1-p} - \left(\frac{1}{d+1}\right)^n, 0\right\} \leq c_k.$$

Now, we write

$$(Q^T \cdot \mathbf{u}^\star)_k$$
$$= \frac{1}{2^n}\sum_{l=0}^{n}\sum_{0\leq j\leq l,k}\binom{n-l}{k-j}\binom{l}{j}\binom{n}{l}\left(\frac{1}{d}\right)^j\left(1-\frac{1}{d}\right)^{l-j}\left(-\frac{1}{d}\right)^{k-j}\left(1+\frac{1}{d}\right)^{n-l-k+j}\left[1-\left(-\frac{d-1}{d+1}\right)^{n-l}\right]$$
$$= \left(\frac{d+1}{2d}\right)^n\left(-\frac{1}{d+1}\right)^k\sum_{l=0}^{n}\sum_{0\leq j\leq l,k}\binom{n-l}{k-j}\binom{l}{j}\binom{n}{l}\left(\frac{d-1}{d+1}\right)^{l-j}(-1)^{-j}\left[1-\left(-\frac{d-1}{d+1}\right)^{n-l}\right]$$
$$= s_1(d,n;k) - s_2(d,n;k), \tag{54}$$

where

$$s_1(d,n;k) = \left(\frac{d+1}{2d}\right)^n\left(-\frac{1}{d+1}\right)^k\sum_{l=0}^{n}\sum_{0\leq j\leq l,k}\binom{n-l}{k-j}\binom{l}{j}\binom{n}{l}\left(\frac{d-1}{d+1}\right)^{l-j}(-1)^{-j}$$
$$= \left(\frac{d+1}{2d}\right)^n\left(-\frac{1}{d+1}\right)^k\sum_{l=0}^{n}\sum_{0\leq j\leq l,k}\binom{n-l}{k-j}\binom{l}{j}\binom{n}{l}\left(\frac{d-1}{d+1}\right)^{l-j}(-1)^{j}, \tag{55}$$

$$s_2(d,n;k) = \left(\frac{d+1}{2d}\right)^n\left(-\frac{1}{d+1}\right)^k\sum_{l=0}^{n}\sum_{0\leq j\leq l,k}\binom{n-l}{k-j}\binom{l}{j}\binom{n}{l}\left(\frac{d-1}{d+1}\right)^{l-j}(-1)^{-j}\left(-\frac{d-1}{d+1}\right)^{n-l}$$
$$= \left(-\frac{d-1}{2d}\right)^n\left(-\frac{1}{d+1}\right)^k\sum_{l=0}^{n}\sum_{0\leq j\leq l,k}\binom{n-l}{k-j}\binom{l}{j}\binom{n}{l}(-1)^{l-j}\left(\frac{d+1}{d-1}\right)^j. \tag{56}$$

Letting $m = l - j$, we rewrite (55) as

$$\begin{aligned}
s_1(d,n;k) &= \left(\frac{d+1}{2d}\right)^n \left(-\frac{1}{d+1}\right)^k \sum_{m=0}^{n-k}\sum_{j=0}^{k} \binom{n-m-j}{k-j}\binom{m+j}{j}\binom{n}{m+j}\left(\frac{d-1}{d+1}\right)^m (-1)^j \\
&= \left(\frac{d+1}{2d}\right)^n \left(-\frac{1}{d+1}\right)^k \sum_{m=0}^{n-k}\sum_{j=0}^{k} \frac{n!}{(k-j)!(n-m-k)!m!j!}\left(\frac{d-1}{d+1}\right)^m (-1)^j \\
&= \left(\frac{d+1}{2d}\right)^n \left(-\frac{1}{d+1}\right)^k \sum_{m=0}^{n-k}\sum_{j=0}^{k} \frac{n!}{(n-m)!m!}\frac{(n-m)!}{(n-m-k)!k!}\frac{k!}{(k-j)!j!}\left(\frac{d-1}{d+1}\right)^m (-1)^j \\
&= \left(\frac{d+1}{2d}\right)^n \left(-\frac{1}{d+1}\right)^k \sum_{m=0}^{n-k}\binom{n}{m}\binom{n-m}{k}\left(\frac{d-1}{d+1}\right)^m \sum_{j=0}^{k}\binom{k}{j}(-1)^j.
\end{aligned}$$

Note the summation over j equals 0 unless $k = 0$. Hence,

$$s_1(d,n;k) = \delta_{0k}\left(\frac{d+1}{2d}\right)^n \sum_{m=0}^{n-k}\binom{n}{m}\left(\frac{d-1}{d+1}\right)^m = \delta_{0k}\left(\frac{d+1}{2d}\right)^n\left(1+\frac{d-1}{d+1}\right)^n = \delta_{0k}. \tag{57}$$

Similarly, (56) can be rewritten as

$$\begin{aligned}
s_2(d,n;k) &= \left(-\frac{d-1}{2d}\right)^n \left(-\frac{1}{d+1}\right)^k \sum_{m=0}^{n-k}\sum_{j=0}^{k} \binom{n}{m+j}\binom{n-m-j}{k-j}\binom{m+j}{j}(-1)^m\left(\frac{d+1}{d-1}\right)^j \\
&= \left(-\frac{d-1}{2d}\right)^n \left(-\frac{1}{d+1}\right)^k \sum_{m=0}^{n-k}\sum_{j=0}^{k} \frac{n!}{(k-j)!(n-m-k)!m!j!}(-1)^m\left(\frac{d+1}{d-1}\right)^j \\
&= \left(-\frac{d-1}{2d}\right)^n \left(-\frac{1}{d+1}\right)^k \sum_{m=0}^{n-k}\sum_{j=0}^{k} \frac{n!}{(n-k)!k!}\frac{(n-k)!}{(n-k-m)!m!}\frac{k!}{(k-j)!j!}(-1)^m\left(\frac{d+1}{d-1}\right)^j \\
&= \left(-\frac{d-1}{2d}\right)^n \left(-\frac{1}{d+1}\right)^k \binom{n}{k}\sum_{m=0}^{n-k}\binom{n-k}{m}(-1)^m\sum_{j=0}^{k}\binom{k}{j}\left(\frac{d+1}{d-1}\right)^j \\
&= \left(-\frac{d-1}{2d}\right)^n \left(-\frac{1}{d+1}\right)^k \binom{n}{k}\sum_{m=0}^{n-k}\binom{n-k}{m}(-1)^m\left(1+\frac{d+1}{d-1}\right)^k.
\end{aligned}$$

Again, the summation over m equals 0 except $n - k = 0$. As a result,

$$s_2(d,n;k) = \delta_{nk}\left(-\frac{d-1}{2d}\right)^n\left(-\frac{1}{d+1}\right)^n\left(1+\frac{d+1}{d-1}\right)^n = \delta_{nk}\left(\frac{1}{d+1}\right)^n. \tag{58}$$

Combining (54), (57) and (58) gives our claim in (53), and thus the proof is completed. □

References

[1] E. L. Lehmann, *Testing Statistical Hypotheses.* Springer New York, 1986.

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

[3] E. A. Haroutunian, M. E. Haroutunian, and A. N. Harutyunyan, "Reliability criteria in information theory and in statistical hypothesis testing," *Foundations and Trends® in Communications and Information Theory*, vol. 4, no. 2–3, pp. 97–263, 2007.

[4] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems (2nd ed.).* Cambridge University Press (CUP), 2011.

[5] Y. Polyanskiy, H. V. Poor, and S. Verdu, "Channel coding rate in the finite blocklength regime," *IEEE Trans. Inform. Theory*, vol. 56, no. 5, pp. 2307–2359, May 2010.
[6] H.-C. Cheng and M.-H. Hsieh, "Moderate deviation analysis for classical-quantum channels and quantum hypothesis testing," *IEEE Transactions on Information Theory*, vol. 64, no. 2, pp. 1385–1403, Feb 2018.
[7] H.-C. Cheng, M.-H. Hsieh, and M. Tomamichel, "Quantum sphere-packing bounds with polynomial prefactors," *IEEE Transactions on Information Theory*, vol. 65, no. 5, pp. 2872–2898, May 2019.
[8] H.-C. Cheng, "Error exponent analysis in quantum information theory," *PhD Thesis (University of Technology Sydney)*, 2018.
[9] O. Goldreich, S. Goldwasser, and D. Ron, "Property testing and its connection to learning and approximation," *Journal of the ACM*, vol. 45, no. 4, pp. 653–750, Jul 1998.
[10] M. Kearns and D. Ron, "Testing problems with sub-learning sample complexity," in *Proceedings of the eleventh annual conference on Computational learning theory - COLT' 98*. ACM Press, 1998.
[11] D. Ron, "Property testing: A learning theory perspective," *Foundations and Trends® in Machine Learning*, vol. 1, no. 3, pp. 307–402, 2007.
[12] ——, "Algorithmic and analysis techniques in property testing," *Foundations and Trends® in Theoretical Computer Science*, vol. 5, no. 2, pp. 73–205, 2009.
[13] E. Fischer, "The art of uninformed decisions: A primer to property testing," in *Current Trends in Theoretical Computer Science*. Wolrd Scientific, Apr 2004, pp. 229–263.
[14] J. Acharya, C. Daskalakis, and G. Kamath, "Optimal testing for properties of distributions," in *Advances in Neural Information Processing Systems*, 2015, pp. 3591–3599.
[15] C. L. Canonne, "A survey on distribution testing: Your data is big. but is it blue?" in *Electronic Colloquium on Computational Complexity (ECCC)*, vol. 22, no. 63, 2015, pp. 1–1.
[16] N. Yu, "Efficient independence testing for quantum states." arXiv:1904.03218 [quant-ph].
[17] L. Wasserman and S. Zhou, "A statistical framework for differential privacy," *Journal of the American Statistical Association*, vol. 105, no. 489, pp. 375–389, 2010.
[18] J. C. Duchi, M. I. Jordan, and M. J. Wainwright, "Local privacy and statistical minimax rates," in *2013 IEEE 54th Annual Symposium on Foundations of Computer Science*. IEEE, Oct 2013.
[19] P. Kairouz, S. Oh, and P. Viswanath, "The composition theorem for differential privacy," *IEEE Transactions on Information Theory*, vol. 63, no. 6, pp. 4037–4049, Jun 2017.
[20] B. Balle, G. Barthe, M. Gaboardi, J. Hsu, and T. Sato, "Hypothesis testing interpretations and Rényi differential privacy." arXiv:1905.09982 [cs.LG].
[21] J. Dong, A. Roth, and W. J. Su, "Gaussian differential privacy." arXiv:1905.02383 [cs.LG].
[22] C. W. Helstrom, *Quantum Detection and Estimation Theory*. Elsevier, 1976.
[23] J. Neyman and E. S. Pearson, "On the problem of the most efficient tests of statistical hypotheses," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 231, no. 694-706, pp. 289–337, Jan 1933.
[24] H. Chernoff, "A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations," *The Annals of Mathematical Statistics*, vol. 23, no. 4, pp. 493–507, Dec 1952.
[25] ——, "Large-sample theory: Parametric case," *The Annals of Mathematical Statistics*, vol. 27, no. 1, pp. 1–22, Mar 1956.
[26] W. Hoeffding, "Asymptotically optimal tests for multinomial distributions," *The Annals of Mathematical Statistics*, vol. 36, no. 2, pp. 369–401, Apr 1965.
[27] K. M. R. Audenaert, J. Calsamiglia, R. Muñoz-Tapia, E. Bagan, L. Masanes, A. Acin, and F. Verstraete, "Discriminating states: The quantum Chernoff bound," *Physical Review Letters*, vol. 98, p. 160501, Apr 2007.
[28] K. M. R. Audenaert, M. Nussbaum, A. Szkoła, and F. Verstraete, "Asymptotic error rates in quantum hypothesis testing," *Communications in Mathematical Physics*, vol. 279, no. 1, pp. 251–283, Feb 2008.
[29] M. Nussbaum and A. Szkoła, "The Chernoff lower bound for symmetric quantum hypothesis testing," *Annals of Statistics*, vol. 37, no. 2, pp. 1040–1057, Apr 2009.

[30] F. Hiai and D. Petz, "The proper formula for relative entropy and its asymptotics in quantum probability," *Communications in Mathematical Physics*, vol. 143, no. 1, pp. 99–114, Dec 1991.

[31] T. Ogawa and H. Nagaoka, "Strong converse and Stein's lemma in quantum hypothesis testing," *IEEE Transaction on Information Theory*, vol. 46, no. 7, pp. 2428–2433, 2000.

[32] H. Umegaki, "Conditional expectation in an operator algebra. IV. entropy and information," *Kodai Mathematical Seminar Reports*, vol. 14, no. 2, pp. 59–85, 1962.

[33] D. Petz, "Quasi-entropies for finite quantum systems," *Reports on Mathematical Physics*, vol. 23, no. 1, pp. 57–65, Feb 1986.

[34] H. Nagaoka, "The converse part of the theorem for quantum Hoeffding bound," 2006. arXiv:quant-ph/0611289 [quant-ph].

[35] ——, "Strong converse theorems in quantum information theory," in *Asymptotic Theory of Quantum Statistical Inference*. World Scientific, Feb 2005, pp. 64–65.

[36] M. Mosonyi and T. Ogawa, "Quantum hypothesis testing and the operational interpretation of the quantum Rényi relative entropies," *Communications in Mathematical Physics*, vol. 334, no. 3, pp. 1617–1648, Dec 2014.

[37] M. Müller-Lennert, F. Dupuis, O. Szehr, S. Fehr, and M. Tomamichel, "On quantum Rényi entropies: A new generalization and some properties," *Journal of Mathematical Physics*, vol. 54, no. 12, p. 122203, 2013.

[38] 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," *Communications in Mathematical Physics*, vol. 331, no. 2, pp. 593–622, Jul 2014.

[39] M. Hayashi, "Optimal sequence of quantum measurements in the sense of Stein's lemma in quantum hypothesis testing," *Journal of Physics A: Mathematical and General*, vol. 35, no. 50, pp. 10 759–10 773, Dec 2002.

[40] I. Bjelakovic and R. Siegmund-Schultze, "Quantum Stein's lemma revisited, inequalities for quantum entropies, and a concavity theorem of Lieb," 2003. arXiv:quant-ph/0307170.

[41] I. Bjelaković, J.-D. Deuschel, T. Krüger, R. Seiler, R. Siegmund-Schultze, and A. Szkoła, "A quantum version of Sanov's theorem," *Communications in Mathematical Physics*, vol. 260, no. 3, pp. 659–671, 2005.

[42] H. Nagaoka and M. Hayashi, "An information-spectrum approach to classical and quantum hypothesis testing for simple hypotheses," *IEEE Transactions on Information Theory*, vol. 53, no. 2, pp. 534–549, Feb 2007.

[43] M. Tomamichel and M. Hayashi, "A Hierarchy of Information Quantities for Finite Block Length Analysis of Quantum Tasks," *IEEE Transactions on Information Theory*, vol. 59, no. 11, pp. 7693–7710, Nov. 2013.

[44] K. Li, "Second-order asymptotics for quantum hypothesis testing," *The Annals of Statistics*, vol. 42, no. 1, pp. 171–189, Feb 2014.

[45] C. T. Chubb, V. Y. F. Tan, and M. Tomamichel, "Moderate deviation analysis for classical communication over quantum channels," *Communications in Mathematical Physics*, vol. 355, no. 3, pp. 1283–1315, Nov 2017.

[46] S. Boixo, S. V. Isakov, V. N. Smelyanskiy, R. Babbush, N. Ding, Z. Jiang, M. J. Bremner, J. M. Martinis, and H. Neven, "Characterizing quantum supremacy in near-term devices," *Nature Physics*, vol. 14, no. 6, pp. 595–600, Apr 2018.

[47] J. Preskill, "Quantum computing in the NISQ era and beyond," *Quantum*, vol. 2, p. 79, Aug 2018.

[48] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, "Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels," *Physical Review Letters*, vol. 70, no. 13, pp. 1895–1899, Mar 1993.

[49] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*. Cambridge University Press, 2009.

[50] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, "Quantum entanglement," *Reviews of Modern Physics*, vol. 81, no. 2, pp. 865–942, Jun 2009.

[51] E. Chitambar, D. Leung, L. Mančinska, M. Ozols, and A. Winter, "Everything you always wanted to know about LOCC (but were afraid to ask)," *Communications in Mathematical Physics*, vol. 328, no. 1, pp. 303–326, Mar 2014.

[52] C. H. Bennett, D. P. DiVincenzo, C. A. Fuchs, T. Mor, E. Rains, P. W. Shor, J. A. Smolin, and W. K. Wootters, "Quantum nonlocality without entanglement," *Physical Review A*, vol. 59, no. 2, p. 1070, 1999.

[53] W. Matthews and A. Winter, "On the Chernoff distance for asymptotic LOCC discrimination of bipartite quantum states," *Communications in Mathematical Physics*, vol. 285, no. 1, pp. 161–174, Jul 2008.

[54] F. G. Brandão, A. W. Harrow, J. R. Lee, and Y. Peres, "Adversarial hypothesis testing and a quantum stein's lemma for restricted measurements," in *Proceedings of the 5th conference on Innovations in theoretical computer science.* ACM, 2014, pp. 183–194.

[55] M. Owari and M. Hayashi, "Asymptotic local hypothesis testing between a pure bipartite state and the completely mixed state," *Physical Review A*, vol. 90, no. 3, Sep 2014.

[56] ——, "Local hypothesis testing between a pure bipartite state and the white noise state," *IEEE Transactions on Information Theory*, vol. 61, no. 12, pp. 6995–7011, Dec 2015.

[57] M. Hayashi and M. Owari, "Tight asymptotic bounds on local hypothesis testing between a pure bipartite state and the white noise state," *IEEE Transactions on Information Theory*, vol. 63, no. 6, pp. 4008–4036, Jun 2017.

[58] J. Calsamiglia, R. Muñoz-Tapia, L. Masanes, A. Acín, and E. Bagan, "Quantum Chernoff bound as a measure of distinguishability between density matrices: Application to qubit and gaussian states," *Physical Review A*, vol. 77, no. 3, Mar 2008.

[59] J. Calsamiglia, J. I. de Vicente, R. Muñoz-Tapia, and E. Bagan, "Local discrimination of mixed states," *Physical Review Letters*, vol. 105, no. 8, Aug 2010.

[60] M. Nathanson, "Testing for a pure state with local operations and classical communication," *Journal of Mathematical Physics*, vol. 51, no. 4, p. 042102, Apr 2010.

[61] N. Yu, R. Duan, and M. Ying, "Four locally indistinguishable ququad-ququad orthogonal maximally entangled states," *Physical Review Letters*, vol. 109, no. 2, Jul 2012.

[62] ——, "Distinguishability of quantum states by positive operator-valued measures with positive partial transpose," *IEEE Transactions on Information Theory*, vol. 60, no. 4, pp. 2069–2079, Apr 2014.

[63] Y. Li, X. Wang, and R. Duan, "Indistinguishability of bipartite states by positive-partial-transpose operations in the many-copy scenario," *Physical Review A*, vol. 95, no. 5, May 2017.

[64] N. Yu and L. Zhou, "Chernoff bound for quantum operations is faithful." arXiv:1705.01642 [quant-ph].

[65] S. Akibue and G. Kato, "Bipartite discrimination of independently prepared quantum states as a counterexample to a parallel repetition conjecture," *Physical Review A*, vol. 97, no. 4, Apr 2018.

[66] H.-C. Cheng, N. Datta, and C. Rouzé, "Strong converse bounds in quantum network information theory: distributed hypothesis testing and source coding," 2019. arXiv:1905.00873 [quant-ph].

[67] H.-C. Cheng, N. Datta, and C. Rouzé, "Strong Converse for Classical-Quantum Degraded Broadcast Channels," 2019. arXiv:1905.00874 [quant-ph].

[68] E. Rains, "A semidefinite program for distillable entanglement," *IEEE Transactions on Information Theory*, vol. 47, no. 7, pp. 2921–2933, 2001.

[69] B. M. Terhal, D. P. DiVincenzo, and D. W. Leung, "Hiding bits in Bell states," *Physical Review Letters*, vol. 86, no. 25, pp. 5807–5810, Jun 2001.

[70] D. DiVincenzo, D. Leung, and B. Terhal, "Quantum data hiding," *IEEE Transactions on Information Theory*, vol. 48, no. 3, pp. 580–598, Mar 2002.

[71] W. Matthews, S. Wehner, and A. Winter, "Distinguishability of quantum states under restricted families of measurements with an application to quantum data hiding," *Communications in Mathematical Physics*, vol. 291, no. 3, pp. 813–843, Aug 2009.

[72] C. Lancien and A. Winter, "Distinguishing multi-partite states by local measurements," *Communications in Mathematical Physics*, vol. 323, no. 2, pp. 555–573, Aug 2013.

[73] L. Lami, C. Palazuelos, and A. Winter, "Ultimate data hiding in quantum mechanics and beyond," *Communications in Mathematical Physics*, vol. 361, no. 2, pp. 661–708, Jun 2018.

[74] R. F. Werner, "Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model," *Physical Review A*, vol. 40, no. 8, pp. 4277–4281, Oct 1989.

[75] Y. Polyanskiy, "Saddle point in the minimax converse for channel coding," *IEEE Transactions on Information Theory*, vol. 59, no. 5, pp. 2576–2595, May 2013.

[76] C. H. Bennett, D. P. DiVincenzo, T. Mor, P. W. Shor, J. A. Smolin, and B. M. Terhal, "Unextendible product bases and bound entanglement," *Physical Review Letters*, vol. 82, no. 26, pp. 5385–5388, Jun 1999.

[77] D. P. DiVincenzo, T. Mor, P. W. Shor, J. A. Smolin, and B. M. Terhal, "Unextendible product bases, uncompletable product bases and bound entanglement," *Communications in Mathematical Physics*, vol. 238, no. 3, pp. 379–410, Jul 2003.

[78] S. Bandyopadhyay, A. Cosentino, N. Johnston, V. Russo, J. Watrous, and N. Yu, "Limitations on separable measurements by convex optimization," *IEEE Transactions on Information Theory*, vol. 61, no. 6, pp. 3593–3604, Jun 2015.

[79] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*. Springer, 1998.