Quantum Hypothesis Testing with Group Symmetry

Fumio Hiai
Milán Mosonyi
Masahito Hayashi

Vienna, Preprint ESI 2134 (2009) April 16, 2009

Supported by the Austrian Federal Ministry of Education, Science and Culture
Available via http://www.esi.ac.at
Quantum hypothesis testing
with group symmetry

Fumio Hiai, 1 Milán Mosonyi 2 and Masahito Hayashi 3

1, 3 Graduate School of Information Sciences, Tohoku University
Aoba-ku, Sendai 980-8579, Japan
2 Mathematical Institute, Budapest University of Technology and Economics
Egy József u 1., Budapest, 1111 Hungary

Abstract

The asymptotic discrimination problem of two quantum states is studied in the setting
where measurements are required to be invariant under some symmetry group of the
system. We consider various asymptotic error exponents in connection with the problems
of the Chernoff bound, the Hoeffding bound and Stein’s lemma, and derive bounds on
these quantities in terms of their corresponding statistical distance measures. A special
emphasis is put on the comparison of the performances of group-invariant and unrestricted
measurements.

AMS classification: 62F03, 62F05, 94A15, 94A17

Keywords: quantum hypothesis testing, group symmetry, asymptotic error exponent,
Chernoff bound, Hoeffding bound, Stein’s lemma, relative entropy, Rényi relative entropy,
quasi-entropy, fidelity

1 Introduction

In the asymptotic framework of (quantum) state discrimination, one is provided with several
copies of a quantum system and with the knowledge that the state of the system is either
$\rho_0$ (null hypothesis $H_0$) or $\rho_1$ (alternative hypothesis $H_1$). One’s aim is to decide, based on
measurements on the copies, which one the true state is. For simplicity, we will assume here
that the Hilbert space $\mathcal{H}$ of the system is finite dimensional, and hence the states can be
represented by density operators $\hat{\rho}_k$ that satisfy $\rho_k(A) = \text{Tr} \hat{\rho}_k A$ for any observable $A \in \mathcal{B}(\mathcal{H})$
and $k = 1, 2$. A measurement on $n$ copies is given by a binary positive operator valued
measure (POVM) $(T, I_n - T)$ with $T \in \mathcal{B}(\mathcal{H}^\otimes n)$, $0 \leq T \leq I_n$, where $T$ corresponds to
accepting $\rho_0$ and $I_n - T$ to accepting $\rho_1$. An erroneous decision is made if $H_0$ is accepted

---

1 E-mail: hiai@math.is.tohoku.ac.jp
2 E-mail: milan.mosonyi@gmail.com
3 E-mail: hayashi@math.is.tohoku.ac.jp
when it is false (error of the first kind) or the other way around (error of the second kind). The probabilities of these events are given by

\[ \beta_{0,n}(T_n) := \rho_0^{\otimes n}(I_n - T_n), \quad \beta_{1,n}(T_n) := \rho_1^{\otimes n}(T_n). \]

The optimal asymptotic performance in a state discrimination problem can be defined in various ways, depending on whether or not the two hypotheses are treated as of equal importance. Usually, one is interested in the exponential decay rates of the above error probabilities or combinations of them, in the \( n \to \infty \) limit. The most studied quantities are the following:

(i) the optimal exponential decay rate of the sum of the two kinds of error probabilities (Chernoff bound),

(ii) the optimal exponential decay rate of the error probabilities of the second kind under the assumption that the error probabilities of the first kind decay with a given exponential speed (Hoeffding bound),

(iii) the optimal exponential decay rate of the error probabilities of the second kind under the assumption that the error probabilities of the first kind vanish asymptotically (Stein’s lemma).

The quantum problem of Stein’s lemma was solved in [18, 28] (see also [12]), where it was shown that the optimal error bound is equal to the relative entropy of the two states, hence providing an operational interpretation of the relative entropy. Recently, the solution of the quantum problem of the Chernoff bound [3, 26] created a renewed interest in hypothesis testing problems. The techniques developed in [3, 26] were also used in [14, 23] to solve the quantum problem of the Hoeffding bound, improving a weaker bound previously given in [27]. The optimal error bounds in these cases are the Chernoff distance and the Hoeffding distance, respectively. All these results deal with the case where one is allowed to perform any collective measurement to discriminate i.i.d. (independent and identically distributed) extensions of the states \( \rho_0 \) and \( \rho_1 \). Various extensions to non-i.i.d. scenarios were also treated in the works [6, 8, 16, 17, 19, 21, 22]. Note that the present formulation describes only the simple hypothesis testing problem, i.e., when both the null and the alternative hypotheses are a single state of the system. Some results in the case where one of the hypotheses is composite (i.e., a subset of the state space) were obtained e.g., in [7, 8, 9].

The purpose of the present paper is to treat the optimal error exponents (i), (ii), and (iii) in the case where the states to discriminate are still i.i.d. extensions of the two simple hypotheses but measurements are restricted to those invariant under the action of some symmetry group of the system. As symmetries and dynamics are described in the same way in the algebraic formalism, this setting also contains the case where one is only able to measure functions of the energy. Indeed, the group in this case is the dynamical group generated by the Hamiltonian of the system, and invariant measurements are exactly those that commute with the Hamilton operator. Hypothesis testing with group-invariant measurements has applications to the entanglement testing problem, as it was shown in [15].

The structure of the paper is as follows. In Section 2 we give a detailed formulation of the problem. As it was shown in [17], the key to solve the state discrimination problems is to determine the asymptotic Rényi relative entropies. This is carried out for the present scenario in Section 3, and the results are used in Section 4 to give bounds on the various
error exponents. In particular, we provide a complete solution to the problem of Stein’s lemma. In Section 5 we analyze the case where the alternative hypothesis is invariant under the symmetry group and in Section 6 we show some examples to compare the performances of restricted and unrestricted measurements.

2 Formulation of the problem

Let $\mathcal{H}$ be a finite-dimensional Hilbert space with $d := \dim \mathcal{H}$ and let $\text{Tr}$ be the usual trace on $\mathcal{B}(\mathcal{H})$. Let $u : G \to \mathcal{B}(\mathcal{H})$ be a unitary representation $u$ of a group $G$ on $\mathcal{H}$. Since $G$ can be replaced without loss of generality by the closure of $\{u_g : g \in G\}$ in the unitary group of $\mathcal{B}(\mathcal{H})$, we may and do assume that $G$ is a compact group. For each $n \in \mathbb{N}$ consider the $n$-fold tensor product representation $u^{\otimes n} : g \mapsto u_g^{\otimes n} \in \mathcal{B}(\mathcal{H})^{\otimes n}, g \in G$, and define a subalgebra $\mathcal{A}_n$ of $\mathcal{B}(\mathcal{H})^{\otimes n} = \mathcal{B}(\mathcal{H}^{\otimes n})$ as the commutant of $u_g^{\otimes n}, g \in G$, i.e.,

$$\mathcal{A}_n := \{ A \in \mathcal{B}(\mathcal{H})^{\otimes n} : Au_g^{\otimes n} = u_g^{\otimes n} A, g \in G \}.$$ 

That is, $\mathcal{A}_n$ is the fixed point subalgebra $(\mathcal{B}(\mathcal{H})^{\otimes n})^G$ of $\mathcal{B}(\mathcal{H})^{\otimes n}$ under the action $\text{Ad} u_g^{\otimes n} := u_g^{\otimes n} \cdot u_g^{\otimes n}*, g \in G$. Then $\mathcal{A}_1 \subset \mathcal{A}_2 \subset \cdots$ by natural inclusions. Note that $\mathcal{A}_n \otimes \mathcal{A}_m \subset \mathcal{A}_{n+m}$ for any $n, m \in \mathbb{N}$. In particular, $\mathcal{A}_1^{\otimes n} \subset \mathcal{A}_n, n \in \mathbb{N}$.

Let $E_{\mathcal{A}_n}$ be the conditional expectation from $\mathcal{B}(\mathcal{H})^{\otimes n}$ onto $\mathcal{A}_n$ with respect to the trace $\text{Tr}$. Note that $E_{\mathcal{A}_n}$ can be written in the integral form

$$E_{\mathcal{A}_n}(X) = \int_G u_g^{\otimes n} X u_g^{\otimes n} d g, \quad X \in \mathcal{B}(\mathcal{H})^{\otimes n}, \quad (2.1)$$

where $d g$ is the Haar probability measure on $G$. Let $\hat{G}$ denote the representation ring consisting of all unitary equivalence classes of irreducible representations of $G$. For each $n \in \mathbb{N}$ the $n$-fold tensor product representation $u^{\otimes n}$ is decomposed into irreducible components as

$$u^{\otimes n} = m_1^{(n)} u_1^{(n)} \oplus m_2^{(n)} u_2^{(n)} \oplus \cdots \oplus m_k^{(n)} u_k^{(n)},$$

where $u_i^{(n)} \in \hat{G}, 1 \leq i \leq k_n$, are contained in $u^{\otimes n}$ with multiplicities $m_i^{(n)}$. For $1 \leq i \leq k_n$ let $d_i^{(n)}$ be the dimension of $u_i^{(n)}$ so that we have $\sum_{i=1}^{k_n} m_i^{(n)} d_i^{(n)} = d^n$ and we can identify $\mathcal{A}_n$ with

$$\mathcal{A}_n = \bigoplus_{i=1}^{k_n} (M_{m_i^{(n)}} \otimes I_{d_i^{(n)}}), \quad (2.2)$$

where $M_k := \mathcal{B}(\mathbb{C}^k), k \in \mathbb{N}$. Then the conditional expectation $E_{\mathcal{A}_n} : \mathcal{B}(\mathcal{H})^{\otimes n} \to \mathcal{A}_n$ given in (2.1) is rewritten as

$$E_{\mathcal{A}_n}(X) = \sum_{i=1}^{k_n} E_i^{(n)} (P_i^{(n)} X P_i^{(n)}), \quad X \in \mathcal{B}(\mathcal{H})^{\otimes n}, \quad (2.3)$$

where $P_i^{(n)}$ is the orthogonal projection onto the subspace corresponding to $m_i^{(n)} u_i^{(n)}$ in the decomposition (2.2), i.e., $P_i^{(n)}$ is the identity $I_{m_i^{(n)}} \otimes I_{d_i^{(n)}}$ of $M_{m_i^{(n)}} \otimes I_{d_i^{(n)}}$, and $E_i^{(n)}$ is the partial trace or the conditional expectation from $M_{m_i^{(n)}} \otimes M_{d_i^{(n)}}$ onto $M_{m_i^{(n)}} \otimes I_{d_i^{(n)}}$ with
corresponding to the problem of Stein’s lemma with respect to the trace. As is well known (see [29] for a detailed proof), the representation ring of any compact group has polynomial growth so that we have

\[
    \lim_{n \to \infty} \frac{1}{n} \log \left( \sum_{i=1}^{k_n} d_i^{(n)} \right) = 0.
\]  

(2.4)

Consider now the hypothesis testing problem with null hypothesis \( \rho_0 \) and alternative hypothesis \( \rho_1 \), as described in the Introduction. We will be interested in the quantities

\[
    \underline{c}_G(\rho_0, \rho_1) := \inf \left\{ \liminf_{\{T_n\}} \frac{1}{n} \log (\beta_{0,n}(T_n) + \beta_{1,n}(T_n)) \right\},
\]

(2.5)

\[
    \overline{c}_G(\rho_0, \rho_1) := \inf \left\{ \limsup_{\{T_n\}} \frac{1}{n} \log (\beta_{0,n}(T_n) + \beta_{1,n}(T_n)) \right\},
\]

(2.6)

\[
    c_G(\rho_0, \rho_1) := \inf \left\{ \lim_{n \to \infty} \frac{1}{n} \log (\beta_{0,n}(T_n) + \beta_{1,n}(T_n)) \right\},
\]

(2.7)

corresponding to the problem of the Chernoff bound,

\[
    \underline{h}_G(r | \rho_0 \parallel \rho_1) := \inf \left\{ \liminf_{\{T_n\}} \frac{1}{n} \log \beta_{1,n}(T_n) \mid \liminf_{n \to \infty} \frac{1}{n} \log \beta_{0,n}(T_n) < -r \right\}, \ r \geq 0,
\]

(2.8)

\[
    \overline{h}_G(r | \rho_0 \parallel \rho_1) := \inf \left\{ \limsup_{\{T_n\}} \frac{1}{n} \log \beta_{1,n}(T_n) \mid \limsup_{n \to \infty} \frac{1}{n} \log \beta_{0,n}(T_n) < -r \right\}, \ r \geq 0,
\]

(2.9)

\[
    h_G(r | \rho_0 \parallel \rho_1) := \inf \left\{ \lim_{n \to \infty} \frac{1}{n} \log \beta_{1,n}(T_n) \mid \limsup_{n \to \infty} \frac{1}{n} \log \beta_{0,n}(T_n) < -r \right\}, \ r \geq 0,
\]

(2.10)

corresponding to the problem of the Hoeffding bound, and

\[
    \underline{s}_G(\rho_0 \parallel \rho_1) := \inf \left\{ \liminf_{\{T_n\}} \frac{1}{n} \log \beta_{1,n}(T_n) \mid \lim_{n \to \infty} \frac{1}{n} \log \beta_{0,n}(T_n) = 0 \right\},
\]

(2.11)

\[
    \overline{s}_G(\rho_0 \parallel \rho_1) := \inf \left\{ \limsup_{\{T_n\}} \frac{1}{n} \log \beta_{1,n}(T_n) \mid \lim_{n \to \infty} \frac{1}{n} \log \beta_{0,n}(T_n) = 0 \right\},
\]

(2.12)

\[
    s_G(\rho_0 \parallel \rho_1) := \inf \left\{ \lim_{n \to \infty} \frac{1}{n} \log \beta_{1,n}(T_n) \mid \lim_{n \to \infty} \frac{1}{n} \log \beta_{0,n}(T_n) = 0 \right\},
\]

(2.13)

corresponding to the problem of Stein’s lemma. Here, the infima are taken over sequences of \( G \)-invariant measurements \( \{T_n\}_{n \in \mathbb{N}} \) with \( T_n \in \mathcal{A}_n, 0 \leq T_n \leq I_n \). Note that posing \( G \)-invariance on the measurements to distinguish \( \rho_0^{\otimes n} \) from \( \rho_1^{\otimes n} \) is equivalent to considering the discrimination of the \( G \)-invariant states

\[
    \rho_{0,n} := \rho_0^{\otimes n} \circ E_{\mathcal{A}_n}, \quad \rho_{1,n} := \rho_1^{\otimes n} \circ E_{\mathcal{A}_n}
\]

(2.14)

with unrestricted measurements, as we have

\[
    \beta_{0,n}(E_{\mathcal{A}_n}(T_n)) = \rho_0^{\otimes n}(I_n - E_{\mathcal{A}_n}(T_n)) = \rho_0^{\otimes n}(E_{\mathcal{A}_n}(I_n - T_n)) = \rho_{0,n}(I_n - T_n)
\]

and similarly for \( \beta_{1,n} \). Hence, the asymptotic problem with \( G \)-invariant measurements is equivalent to the asymptotic state discrimination problem of the two sequences of \( G \)-invariant states \( \{\rho_{0,n}\}_{n \in \mathbb{N}} \) and \( \{\rho_{1,n}\}_{n \in \mathbb{N}} \). Note also that the families \( \{\rho_{k,n}\}_{n \in \mathbb{N}}, k = 0, 1 \) are compatible
Lemma 3.1. (1) The sequence \( \lim_{n \to \infty} \psi_n(s) \) exists and is orthogonal, then the same holds for any \( \rho \). Let \( G \) be a unitary representation of a compact group \( G \). In this case we will omit the subscript \( G \) from the notations for the error exponents (2.5)–(2.13).

3 Asymptotic distance measures

Let \( u \) be a unitary representation of a compact group \( G \) on \( \mathcal{H} \) as given in Section 2, and let \( \rho_0 \) and \( \rho_1 \) be two states on \( \mathcal{B}(\mathcal{H}) \) with the density matrices \( \hat{\rho}_0 \) and \( \hat{\rho}_1 \). We consider the sequences of states \( \{\rho_{0,n}\}_{n \in \mathbb{N}} \) and \( \{\rho_{1,n}\}_{n \in \mathbb{N}} \) as defined in (2.14). Note that the densities \( \hat{\rho}_{k,n} \) of \( \rho_{k,n} \) with respect to \( \text{Tr} \) are given as \( \hat{\rho}_{k,n} = E_{\mathcal{A}_n}(\hat{\rho}_{k,n}^{\otimes n}) \), \( k = 0, 1 \). Define

\[
\psi_n(s) := \log \text{Tr} \hat{\rho}_{0,n}^{s} \hat{\rho}_{1,n}^{1-s}, \quad s \in \mathbb{R}, \ n \in \mathbb{N},
\]

where \( \hat{\rho}_{0,n}^{s} \) and \( \hat{\rho}_{1,n}^{1-s} \) are defined for all \( s \in \mathbb{R} \) with convention \( 0^s = 0 \) for all \( s \in \mathbb{R} \). Also, we write the \( \psi \)-function in the unrestricted setting as

\[
\psi^\circ(s) := \log \text{Tr} \hat{\rho}_{0}^{s} \hat{\rho}_{1}^{1-s}, \quad s \in \mathbb{R}, \ n \in \mathbb{N}.
\]

Furthermore, let

\[
\psi(s) := \lim_{n \to \infty} \frac{1}{n} \psi_n(s)
\]

whenever the limit exists. Note that \( \psi_n \) is finite and convex on \( \mathbb{R} \) as long as \( \text{supp} \rho_{0,n} \) and \( \text{supp} \rho_{1,n} \) are not orthogonal (otherwise, \( \psi_n \) is identically \( -\infty \)). Hence \( \psi \) is convex on any interval where it exists with values in \([-\infty, +\infty)\). Also, note that if \( \rho_{0,n} \) and \( \rho_{1,n} \) are orthogonal, then the same holds for any \( m \geq n \).

Lemma 3.1. (1) The sequence \( \psi_n(s) \), \( n \in \mathbb{N} \), is subadditive for any \( s \in [0,1] \). Hence the limit (3.1) exists and

\[
\psi(s) = \lim_{n \to \infty} \frac{1}{n} \psi_n(s) = \inf_{n \geq 1} \frac{1}{n} \psi_n(s), \quad s \in [0,1].
\]

Furthermore, if \( \rho_0 \) and \( \rho_1 \) are not orthogonal, then \( \psi(s) \) is finite with \( \psi^\circ(s) \leq \psi(s) \leq \psi_1(s) \) for all \( s \in [0,1] \).

(2) Assume that \( \text{supp} \rho_1 \) is \( G \)-invariant, i.e., \( \text{supp} \rho_1 \in \mathcal{A}_1 \). Then the sequence \( \psi_n(s) \), \( n \in \mathbb{N} \), is superadditive for any \( s \in [1,2] \). Hence the limit (3.1) exists and

\[
\psi(s) = \lim_{n \to \infty} \frac{1}{n} \psi_n(s) = \sup_{n \geq 1} \frac{1}{n} \psi_n(s), \quad s \in [1,2].
\]

If, additionally, \( \rho_0 \) and \( \rho_1 \) are not orthogonal, then \( \psi(s) \) is finite with \( \psi^\circ(s) \geq \psi(s) \geq \psi_1(s) \) for all \( s \in [1,2] \).

Proof. Let \( 0 \leq s \leq 1 \). By Lieb’s concavity theorem [20] (see also the Appendix A.1), the function \( (A, B) \mapsto \text{Tr} A^s B^{1-s} \) is jointly concave on the set of \( (A, B) \in \mathcal{B}(\mathcal{H})^{\otimes n+m} \times \cdots \)
Lemma 3.1, so that the continuity at 1 is obvious.

\[ \psi \}

and hence

\[ \psi \]

so that \( \psi \}

for all \( \psi \}

and \( \psi \}

By Lemma 3.1 (1) and the convexity of \( \psi \}

Corollary 3.2. If \( \text{supp} \ \rho_0 \leq \text{supp} \ \rho_1 \) then \( \psi \) is left-continuous at 1 as

\[ \lim_{s \downarrow 1} \psi(s) = \psi(1) = 0. \]

Similarly, \( \text{supp} \ \rho_0 \geq \text{supp} \ \rho_1 \) implies the right continuity of \( \psi \) at 0. If \( \text{supp} \ \rho_1 \) is \( G \)-invariant and \( \text{supp} \ \rho_0 \) and \( \text{supp} \ \rho_1 \) are not orthogonal (in particular, if \( \rho_1 \) is faithful), then \( \psi \) is continuous at 1.

\[ \text{Proof.} \ \text{Assume that } \text{supp} \ \rho_0 \leq \text{supp} \ \rho_1. \ \text{Then } \psi_n(1) = 0 \ \text{for all } n \in \mathbb{N}, \ \text{and hence } \psi(1) = 0. \]

By Lemma 3.1(1) and the convexity of \( \psi \),

\[ 0 = \psi(1) = \lim_{s \downarrow 1} \psi(s) \leq \lim_{s \downarrow 1} \psi(s) \leq \psi(1) = 0, \]

and hence \( \psi \) is left-continuous at 1. The proof of the second assertions is similar. Assume the conditions in the last assertion. Then \( \psi \) is a finite-valued convex function on \([0, 2]\) by Lemma 3.1, so that the continuity at 1 is obvious.

The Rényi relative entropy of order \( \alpha \in \mathbb{R} \setminus \{1\} \) of \( \rho_{0,n} \) with respect to \( \rho_{1,n} \) is defined as

\[ S_\alpha(\rho_{0,n} \parallel \rho_{1,n}) := \frac{1}{\alpha - 1} \log \text{Tr} \hat{\rho}_{0,n}^{\alpha} \hat{\rho}_{1,n}^{1-\alpha} = -\frac{1}{1-\alpha} \psi_n(\alpha). \]

By Lemma 3.1, the Rényi relative entropies with parameter between 0 and 1 are superadditive:

\[ S_\alpha(\rho_{0,n+m} \parallel \rho_{1,n+m}) \geq S_\alpha(\rho_{0,n} \parallel \rho_{1,n}) + S_\alpha(\rho_{0,m} \parallel \rho_{1,m}), \quad n, m \in \mathbb{N}, \ \alpha \in [0, 1), \quad (3.3) \]
and the mean Rényi relative entropy of order $\alpha$ exists for any $\alpha \in [0, 1)$:

$$S_{\alpha,M}(\rho_0 \parallel \rho_1) := \lim_{n \to \infty} \frac{1}{n} S_{\alpha}(\rho_{0,n} \parallel \rho_{1,n}) = \sup_{n \geq 1} \frac{1}{n} S_{\alpha}(\rho_{0,n} \parallel \rho_{1,n}).$$  \hfill (3.4)

Similarly, if $\rho_{1,n}$ is faithful for all $n$, then (3.3) and (3.4) hold for the Rényi relative entropies with parameter $\alpha \in (1, 2]$. One can easily see that

$$\partial^- \psi_n(1) = \lim_{\alpha \to 1} S_{\alpha}(\rho_{0,n} \parallel \rho_{1,n}) = S(\rho_{0,n} \parallel \rho_{1,n}),$$

where $\partial^- \psi_n(1)$ is the left derivative of $\psi$ at 1 and $S(\rho_{0,n} \parallel \rho_{1,n})$ is the relative entropy of $\rho_{0,n}$ with respect to $\rho_{1,n}$, defined as

$$S(\rho_{0,n} \parallel \rho_{1,n}) := \begin{cases} \text{Tr} \hat{\rho}_{0,n}(\log \hat{\rho}_{0,n} - \log \hat{\rho}_{1,n}) & \text{if supp} \rho_{0,n} \subseteq \text{supp} \rho_{1,n}, \\ +\infty & \text{otherwise.} \end{cases}$$

Hence the relative entropy is also superadditive:

$$S(\rho_{0,n+m} \parallel \rho_{1,n+m}) \geq S(\rho_{0,n} \parallel \rho_{1,n}) + S(\rho_{0,m} \parallel \rho_{1,m}),$$  \hfill (3.5)

and the mean relative entropy is given by

$$S_{M}(\rho_0 \parallel \rho_1) := \lim_{n \to \infty} \frac{1}{n} S(\rho_{0,n} \parallel \rho_{1,n}) = \sup_{n \geq 1} \frac{1}{n} S(\rho_{0,n} \parallel \rho_{1,n}).$$

Note that the superadditivity (3.5) can also be shown by the monotonicity of the relative entropy.

**Remark 3.3.** If we choose the maximally mixed state for $\rho_1$ (i.e., $\hat{\rho}_1 = d^{-1}I_d$), then (3.3) gives the subadditivity

$$S_{\alpha}(\rho_{0,n+m}) \leq S_{\alpha}(\rho_{0,n}) + S_{\alpha}(\rho_{0,m}), \quad n, m \in \mathbb{N}, \ \alpha \in [0, 1),$$

where

$$S_{\alpha}(\rho) := \frac{1}{1-\alpha} \log \text{Tr} \hat{\rho}^\alpha$$

is the Rényi entropy of order $\alpha$. This is of some interest since, as is well known [31, Chap. IX, §6], the Rényi entropy of order $\alpha$ is not subadditive in general except for the cases $\alpha = 0$ and $\alpha = 1$ ($S_1$ denotes the von Neumann entropy).

We define the Chernoff distance of $\rho_{0,n}$ and $\rho_{1,n}$ as

$$C(\rho_{0,n}, \rho_{1,n}) := -\min_{0 \leq s \leq 1} \psi_n(s),$$

and their Hoeffding distance with parameter $r \geq 0$ as

$$H(r \mid \rho_{0,n} \parallel \rho_{1,n}) := \sup_{0 \leq t < 1} \frac{-tr - \psi_n(t)}{1-t} = \sup_{0 \leq t < 1} \left\{ S_t(\rho_{0,n} \parallel \rho_{1,n}) - \frac{tr}{1-t} \right\}.$$ 

The mean versions of the above quantities are defined as

$$C_{M}(\rho_0, \rho_1) := \lim_{n \to \infty} \frac{1}{n} C(\rho_{0,n}, \rho_{1,n}),$$

$$H_{M}(r \mid \rho_0, \rho_1) := \lim_{n \to \infty} \frac{1}{n} H(r \mid \rho_{0,n} \parallel \rho_{1,n}).$$
\begin{align}
H_M(r \mid \rho_0 \parallel \rho_1) := \lim_{n \to \infty} \frac{1}{n} H(nr \mid \rho_{0,n} \parallel \rho_{1,n}),
\end{align}

if the limits exist.

We also define the Legendre-Fenchel transforms (or the polar functions)
\begin{align}
\varphi_n(a) := \max_{0 \leq s \leq 1} \{as - \psi_n(s)\} \quad \text{and} \quad \varphi(a) := \sup_{0 \leq s \leq 1} \{as - \psi(s)\}, \quad a \in \mathbb{R}.
\end{align}

Note that
\begin{align}
C(\rho_{0,n}, \rho_{1,n}) = \varphi_n(0).
\end{align}

**Lemma 3.4.** Let \( f_n, n \in \mathbb{N} \), be a superadditive sequence of functions on an interval \( I \). Let \( f(s) := \lim_n \frac{1}{n} f_n(s) = \sup_n \frac{1}{n} f_n(s) \), the existence of which is guaranteed by superadditivity. Then
\begin{align}
\sup_{s \in I} f(s) = \lim_{n \to \infty} \frac{1}{n} \sup_{s \in I} f_n(s) = \sup_n \frac{1}{n} \sup_{s \in I} f_n(s).
\end{align}

**Proof.** Obviously,
\begin{align}
\frac{1}{n} \sup_{s \in I} f_n(s) \geq \frac{1}{n} f_n(t), \quad n \in \mathbb{N}, \ t \in I,
\end{align}
and thus
\begin{align}
\liminf_{n \to \infty} \frac{1}{n} \sup_{s \in I} f_n(s) \geq \liminf_{n \to \infty} \frac{1}{n} f_n(t) = f(t), \quad t \in I,
\end{align}
which yields
\begin{align}
\liminf_{n \to \infty} \frac{1}{n} \sup_{s \in I} f_n(s) \geq \sup_{s \in I} f(s).
\end{align}
On the other hand, we have
\begin{align}
\sup_{s \in I} f(s) \geq f(t) \geq \frac{1}{n} f_n(t), \quad n \in \mathbb{N}, \ t \in I,
\end{align}
and hence
\begin{align}
\sup_{s \in I} f(s) \geq \sup_{n \geq 1} \frac{1}{n} \sup_{s \in I} f_n(s),
\end{align}
implying the assertion. \hfill \Box

**Proposition 3.5.** The sequence \( \frac{1}{n} \varphi_n(na) \) converges for any \( a \in \mathbb{R} \), and
\begin{align}
\varphi(a) = \lim_{n \to \infty} \frac{1}{n} \varphi_n(na) = \sup_n \frac{1}{n} \varphi_n(na), \quad a \in \mathbb{R}.
\end{align}

Moreover, the mean Chernoff bound and the mean Hoeffding bound exist, and
\begin{align}
C_M(\rho_0, \rho_1) = \sup_n \frac{1}{n} C(\rho_{0,n}, \rho_{1,n}) = \varphi(0) = - \inf_{0 \leq s \leq 1} \psi(s),
\end{align}
\begin{align}
H_M(r \mid \rho_0 \parallel \rho_1) = \sup_n \frac{1}{n} H(nr \mid \rho_{0,n} \parallel \rho_{1,n}) = \sup_{0 \leq s < 1} \frac{-sr - \psi(s)}{1 - s}, \quad r \geq 0.
\end{align}
Lemma 3.8. Assume that \( \text{supp} \, \rho_{0,n} \leq \text{supp} \, \rho_{1,n} \) for all \( n \in \mathbb{N} \). Then
\[
H_M(0 \| \rho_0 \| \rho_1) = \partial^- \psi(1) = S_M(\rho_0 \| \rho_1),
\]
where \( \partial^- \psi(1) \) is the left derivative of \( \psi \) at 1.

Proof. The assumption on the supports yields \( \psi_n(1) = 0 \) for all \( n \in \mathbb{N} \) and by the convexity of the \( \psi_n \), the functions \( s \mapsto \psi_n(t)/(t-1) \) are monotonically increasing, and hence,
\[
H(0 \| \rho_{0,n} \| \rho_{1,n}) = \sup_{0 \leq t < 1} \frac{\psi_n(t)}{t-1} = \lim_{t \uparrow 1} \frac{\psi_n(t)}{t-1} = \partial^- \psi_n(1) = S(\rho_{0,n} \| \rho_{1,n}).
\]
Since this holds for all \( n \in \mathbb{N} \), we have
\[
H_M(0 \| \rho_0 \| \rho_1) = S_M(\rho_0 \| \rho_1),
\]
and \( H_M(0 \| \rho_0 \| \rho_1) = \partial^- \psi(1) \) follows again by the convexity of \( \psi \) and that \( \psi(1) = 0 \).
Remark 3.9. Note that \( \rho_{1,n} \) is unchanged if the alternative hypothesis \( \rho_1 \) is replaced by \( \rho_1 \circ \text{Ad} u_g \) for any \( g \in G \). Hence, by (3.2),

\[
\psi^\circ(s \mid \rho_0 \| \rho_1 \circ \text{Ad} u_g) \leq \psi(s), \quad s \in [0, 1], \quad g \in G,
\]

where the above left-hand side denotes \( \psi^\circ(s) \) for the hypotheses \( \rho_0 \) and \( \rho_1 \circ \text{Ad} u_g \). From (3.9) and (3.10) one obtains

\[
C_M(\rho_0, \rho_1) \leq \inf_{g \in G} C(\rho_0, \rho_1 \circ \text{Ad} u_g) \leq C(\rho_0, \rho_1), \quad (3.11)
\]

\[
H_M(r \mid \rho_0 \| \rho_1) \leq \inf_{g \in G} H(r \mid \rho_0 \| \rho_1 \circ \text{Ad} u_g) \leq H(r \mid \rho_0 \| \rho_1), \quad r \geq 0, \quad (3.12)
\]

and also by the monotonicity of the relative entropy,

\[
S_M(\rho_0 \| \rho_1) \leq \inf_{g \in G} S(\rho_0 \| \rho_1 \circ \text{Ad} u_g) \leq S(\rho_0 \| \rho_1). \quad (3.13)
\]

Similar inequalities are valid for the error exponents (2.5)–(2.13) as well. For example,

\[
H(r \mid \rho_0 \| \rho_1) \geq \sup_{g \in G} H(r \mid \rho_0 \| \rho_1 \circ \text{Ad} u_g) \geq \tilde{H}(r \mid \rho_0 \| \rho_1), \quad r \geq 0.
\]

4 Asymptotic error probabilities

For each \( a \in \mathbb{R} \), we define the corresponding minimal asymmetric error probability for the discrimination between \( p_{1,n} \) and \( p_{1,n} \) as

\[
P_{\min}(a \mid \rho_{0,n} : p_{1,n}) := \min_{T_n \in \mathcal{A}_n, 0 \leq T_n \leq I_n} \left\{ e^{-na} \beta_{0,n}(T_n) + \beta_{1,n}(T_n) \right\},
\]

where \( \beta_{0,n}(T_n) := \rho_{0,n}(I_n - T_n) \) and \( \beta_{1,n}(T_n) := \rho_{1,n}(T_n) \) are the error probabilities of the first and the second kinds for a test \( T_n \). One can easily see that

\[
P_{\min}(a \mid \rho_{0,n} : p_{1,n}) = e^{-na} \beta_{0,n}(S_{n,a}) + \beta_{1,n}(S_{n,a})
\]

\[
= \frac{1 + e^{-na}}{2} - \frac{1}{2} \| e^{-na} \rho_{0,n} - \rho_{1,n} \|_1, \quad (4.1)
\]

where \( S_{n,a} := \{ e^{-na} \rho_{0,n} - \rho_{1,n} > 0 \} \) is the spectral projection of the self-adjoint operator \( e^{-na} \rho_{0,n} - \rho_{1,n} \) corresponding to the positive part of its spectrum. \( S_{n,a} \) is called a Neyman-Pearson test or Holevo-Helstrom test. We define the minimal symmetric error probabilities as \( P_{\min}(\rho_{0,n} : p_{1,n}) := P_{\min}(0 \mid \rho_{0,n} : p_{1,n}) \). One can easily see that

\[
\lim_{n \to \infty} \frac{1}{n} \log P_{\min}(\rho_{0,n} : p_{1,n}) = \theta_{\varphi}(\rho_0, \rho_1), \quad (4.2)
\]

\[
\lim_{n \to \infty} \frac{1}{n} \log P_{\min}(\rho_{0,n} : p_{1,n}) = \theta_{\psi}(\rho_0, \rho_1). \quad (4.3)
\]

Lemma 4.1. For any \( a \in \mathbb{R} \),

\[
\lim_{n \to \infty} \frac{1}{n} \log P_{\min}(a \mid \rho_{0,n} : p_{1,n}) \leq -\varphi(a), \quad (4.4)
\]

\[
\lim_{n \to \infty} \frac{1}{n} \log P_{\min}(a \mid \rho_{0,n} : p_{1,n}) \geq \begin{cases} 2\psi(1/2) & \text{if } a \leq 0, \\ 2\psi(1/2) - a & \text{if } a > 0. \end{cases} \quad (4.5)
\]
Proof. Let $A$ and $B$ be positive semidefinite operators $A$ and $B$ on the same Hilbert space. By [3, Theorem 1],
\[
\frac{1}{2} \text{Tr}(A + B) - \frac{1}{2} \|A - B\|_1 \leq \text{Tr} A^t B^{1-t} , \quad t \in [0,1],
\] (4.6)
and by [4, Theorem 7],
\[
(\frac{1}{2} \|A - B\|_1)^2 + (\text{Tr} A^{1/2} B^{1/2})^2 \leq \left( \frac{1}{2} \text{Tr}(A + B) \right)^2 .
\] (4.7)
Let $A := e^{-na} \hat{\rho}_{0,n}$ and $B := \hat{\rho}_{1,n}$. Then, (4.6) together with (4.1) yields (4.4). On the other hand, by (4.7) we have
\[
\frac{1}{2} \|e^{-na} \hat{\rho}_{0,n} - \hat{\rho}_{1,n}\|_1 \leq \sqrt{(\frac{1 + e^{-na}}{2})^2 - e^{-na} (\text{Tr} \hat{\rho}_{0,n}^{1/2} \hat{\rho}_{1,n}^{1/2})^2},
\]
and hence, by (4.1),
\[
P_{\min}(a \mid \rho_{0,n} : \rho_{1,n}) \geq \frac{1 + e^{-na}}{2} - \sqrt{(\frac{1 + e^{-na}}{2})^2 - e^{-na} (\text{Tr} \hat{\rho}_{0,n}^{1/2} \hat{\rho}_{1,n}^{1/2})^2}.
\]
Therefore,
\[
\liminf_{n \to \infty} \frac{1}{n} \log P_{\min}(a \mid \rho_{0,n} : \rho_{1,n}) \geq 2 \lim_{n \to \infty} \frac{1}{n} \log \text{Tr} \hat{\rho}_{0,n}^{1/2} \hat{\rho}_{1,n}^{1/2} + \lim_{n \to \infty} \frac{1}{n} \log \frac{e^{-na}}{1 + e^{-na}},
\]
which yields (4.5). \qed

Note that $-\psi(1/2) \leq \varphi(0) = C_M(\rho_0, \rho_1)$. By taking into account (4.2) and (4.3), Lemma 4.1 with the choice $a = 0$ yields the following:

**Proposition 4.2.**
\[-2C_M(\rho_0, \rho_1) \leq \zeta_2(\rho_0, \rho_1) \leq \tau_c(\rho_0, \rho_1) \leq -C_M(\rho_0, \rho_1).\]

**Proposition 4.3.** For any $r \geq 0$,
\[
\overline{h}_c(r \mid \rho_0 \parallel \rho_1) \leq -H_M(r \mid \rho_0 \parallel \rho_1).
\]

**Proof.** Note that $\max \{e^{-na} \beta_{0,n}(S_{n,a}), \beta_{1,n}(S_{n,a})\} \leq P_{\min}(a \mid \rho_{0,n} : \rho_{1,n})$, and hence (4.4) yields
\[
\limsup_n \frac{1}{n} \log \beta_{0,n}(S_{n,a}) \leq -\varphi(a) + a,
\]
\[
\limsup_n \frac{1}{n} \log \beta_{1,n}(S_{n,a}) \leq -\varphi(a).
\]

11
Therefore,
\[ \overline{h}_G(r \mid \rho_0 \parallel \rho_1) \leq - \sup_{a : \varphi(a) - a > r} \varphi(a) = - \sup_{0 \leq s < 1} \frac{-sr - \psi(s)}{1 - s}, \]
where the last identity was shown, e.g., in the proof of [17, Theorem 4.8]. A detailed proof is given in Lemma A.2 of the Appendix. Finally, the right-hand side of the above inequality is equal to \(-H_M(r \mid \rho_0 \parallel \rho_1)\) by Proposition 3.5. \(\Box\)

The following theorem gives the solution of Stein’s lemma in our setting:

**Theorem 4.4.** Assume that \(\text{supp } \rho_{0,n} \leq \text{supp } \rho_{1,n}\) for all \(n \in \mathbb{N}\). Then
\[ s_G(\rho_0 \parallel \rho_1) = \overline{s}_G(\rho_0 \parallel \rho_1) = s_G(\rho_0 \parallel \rho_1) = -S_M(\rho_0 \parallel \rho_1). \]

**Proof.** We have
\[ -S_M(\rho_0 \parallel \rho_1) \leq s_G(\rho_0 \parallel \rho_1) \leq \overline{s}_G(\rho_0 \parallel \rho_1) \leq s_G(\rho_0 \parallel \rho_1), \]
where the first inequality follows, e.g., from [17, Proposition 5.2], and the rest are obvious by definitions. Note that \(\overline{s}_G(\rho_0 \parallel \rho_1) \leq \overline{h}_G(0 \mid \rho_0 \parallel \rho_1)\) by definition. Taking into account Lemma 3.6, Proposition 4.3 implies that
\[ \overline{s}_G(\rho_0 \parallel \rho_1) \leq \overline{h}_G(0 \mid \rho_0 \parallel \rho_1) \leq -H_M(0 \mid \rho_0 \parallel \rho_1) = -S_M(\rho_0 \parallel \rho_1). \]

By the definition of \(\overline{s}_G(\rho_0 \parallel \rho_1)\), for each \(k \in \mathbb{N}\) there exists a sequence of tests \(T_{n,k}, n \in \mathbb{N}\), such that
\[ \lim_{n \to \infty} \beta_{0,n}(T_{n,k}) = 0, \quad \limsup_{n \to \infty} \frac{1}{n} \log \beta_{1,n}(T_{n,k}) < -S_M(\rho_0 \parallel \rho_1) + \frac{1}{k}. \]

For each \(k\), we can choose an \(n_k \in \mathbb{N}\) such that for every \(n \geq n_k\),
\[ \beta_{0,n}(T_{n,k}) < \frac{1}{k}, \quad \frac{1}{n} \log \beta_{1,n}(T_{n,k}) < -S_M(\rho_0 \parallel \rho_1) + \frac{1}{k}. \]

Here we may assume that \(n_1 < n_2 < \ldots\), and we define \(T^*_n := T_{n,k}\) if \(n_k \leq n < n_{k+1}\), \(k \in \mathbb{N}\). Obviously, for this sequence of tests,
\[ \lim_{n \to \infty} \beta_{0,n}(T^*_n) = 0, \quad \limsup_{n \to \infty} \frac{1}{n} \log \beta_{1,n}(T^*_n) \leq -S_M(\rho_0 \parallel \rho_1). \]

On the other hand, \(-S_M(\rho_0 \parallel \rho_1) \leq \overline{s}_G(\rho_0 \parallel \rho_1)\) yields
\[ -S_M(\rho_0 \parallel \rho_1) \leq \liminf_{n \to \infty} \frac{1}{n} \log \beta_{1,n}(T^*_n), \]
and hence,
\[ \lim_{n \to \infty} \frac{1}{n} \log \beta_{1,n}(T^*_n) = -S_M(\rho_0 \parallel \rho_1). \]

This implies that
\[ s_G(\rho_0 \parallel \rho_1) \leq -S_M(\rho_0 \parallel \rho_1), \]
which completes the proof. \(\Box\)
The problem of Stein’s lemma can also be formulated in a slightly different way that is not completely equivalent to the above formulation. For each \( \varepsilon \in (0, 1) \) and \( n \in \mathbb{N} \), define the quantity

\[
\beta_{\varepsilon}(\rho_{0,n} \parallel \rho_{1,n}) := \min \{ \beta_{1,n}(T_n) : T_n \in \mathcal{A}_n, \ 0 \leq T_n \leq I_n, \ \beta_{0,n}(T_n) \leq \varepsilon \}
\]

and

\[
\underline{s}_{G,\varepsilon}(\rho_0 \parallel \rho_1) := \inf_{\{T_n\}} \left\{ \liminf_{n \to \infty} \frac{1}{n} \log \beta_n(T_n) \Bigg| \alpha_n(T_n) \leq \varepsilon \right\},
\]

\[
\overline{s}_{G,\varepsilon}(\rho_0 \parallel \rho_1) := \inf_{\{T_n\}} \left\{ \limsup_{n \to \infty} \frac{1}{n} \log \beta_n(T_n) \Bigg| \alpha_n(T_n) \leq \varepsilon \right\}.
\]

One can easily see that

\[
\liminf_{n \to \infty} \frac{1}{n} \log \beta_{\varepsilon}(\rho_{0,n} \parallel \rho_{1,n}) = \underline{s}_{G,\varepsilon}(\rho_0 \parallel \rho_1), \quad \limsup_{n \to \infty} \frac{1}{n} \log \beta_{\varepsilon}(\rho_{0,n} \parallel \rho_{1,n}) = \overline{s}_{G,\varepsilon}(\rho_0 \parallel \rho_1),
\]

and

\[
\sup_{\varepsilon} \underline{s}_{G,\varepsilon}(\rho_0 \parallel \rho_1) = \underline{s}_G(\rho_0 \parallel \rho_1), \quad \sup_{\varepsilon} \overline{s}_{G,\varepsilon}(\rho_0 \parallel \rho_1) = \overline{s}_G(\rho_0 \parallel \rho_1).
\]

Hence, Theorem 4.4 implies that if \( \text{supp} \rho_{0,n} \leq \text{supp} \rho_{1,n} \) for all \( n \) then

\[
\underline{s}_{G,\varepsilon}(\rho_0 \parallel \rho_1) \leq \overline{s}_{G,\varepsilon}(\rho_0 \parallel \rho_1) \leq -S_M(\rho_0 \parallel \rho_1) \quad (4.8)
\]

for all \( \varepsilon \in (0, 1) \).

**Theorem 4.5.** If \( \text{supp} \rho_1 \) is \( G \)-invariant and \( \text{supp} \rho_0 \) and \( \text{supp} \rho_1 \) are not orthogonal (in particular, if \( \rho_1 \) is faithful), then

\[
-\partial^+ \psi(1) \leq \underline{s}_{G,\varepsilon}(\rho_0 \parallel \rho_1) \leq \overline{s}_{G,\varepsilon}(\rho_0 \parallel \rho_1) \leq -\partial^- \psi(1).
\]

**Proof.** In exactly the same way as in [28], one can show that

\[
\beta_{1,n}(T_n) \geq e^{-na} \left( 1 - \varepsilon - e^{-\tilde{\varphi}(na)} \right)
\]

for any test \( T_n \) that satisfies \( \beta_{0,n}(T_n) \leq \varepsilon \). By Lemma 3.8,

\[
\lim_{n \to \infty} \frac{1}{n} \tilde{\varphi}(na) = \tilde{\varphi}(a) = \max_{1 \leq s \leq 3/2} \{ a(s - 1) - \psi(s) \}.
\]

The latter is strictly positive if and only if \( a > \partial^+ \psi(1) \), and in this case \( \tilde{\varphi}(na) > (n/2)\tilde{\varphi}(a) \) for every large enough \( n \), and hence \( \lim_n \tilde{\varphi}(na) = +\infty \). Hence,

\[
\liminf_{n \to \infty} \frac{1}{n} \log \beta_{1,n}(T_n) \geq -a + \lim_{n \to \infty} \frac{1}{n} \log \left( 1 - \varepsilon - e^{-\tilde{\varphi}(na)} \right) = -a.
\]

Since this is true for all \( a > \partial^+ \psi(1) \), we get \( \underline{s}_{G,\varepsilon}(\rho_0 \parallel \rho_1) \geq -\partial^+ \psi(1) \). The rest of the inequalities are just a restatement of (4.8) thanks to Lemma 3.6. \qed

Theorem 4.5 together with Lemma 3.6 yields immediately the following:
Corollary 4.6. If $\rho_1$ is faithful and $\psi$ is differentiable at 1, then

$$\lim_{n \to \infty} \frac{1}{n} \log \beta_\epsilon(\rho_{0,n} \parallel \rho_{1,n}) = -S_M(\rho_0 \parallel \rho_1)$$

for all $\epsilon \in (0, 1)$.

We close this section with the following:

Remark 4.7. The analysis in [17] shows that if $\psi$ exists and is differentiable on the whole real line and $\partial^\psi(1) = S_M(\rho_0 \parallel \rho_1)$ (see Lemma 3.6) then

$$c_G(\rho_0, \rho_1) = c_G(\rho_0, \rho_1) = c_G(\rho_0, \rho_1) = -C_M(\rho_0, \rho_1),$$
$$h_G(r | \rho_0 \parallel \rho_1) = h_G(r | \rho_0 \parallel \rho_1) = h_G(r | \rho_0 \parallel \rho_1) = -H_M(r | \rho_0 \parallel \rho_1), \quad r \geq 0,$$
$$s_G(\rho_0 \parallel \rho_1) = s_G(\rho_0 \parallel \rho_1) = s_G(\rho_0 \parallel \rho_1) = -S_M(\rho_0 \parallel \rho_1).$$

(Actually, it is enough to require the existence and differentiability of $\psi$ on the open interval $(0, 1)$ to show the above identities based on the Gärtner-Ellis theorem.) Note that (4.2) and (4.3) imply in this case that

$$\lim_{n \to \infty} \frac{1}{n} \log P_{\min}(\rho_{0,n} : \rho_{1,n}) = -C_M(\rho_0, \rho_1). \quad (4.9)$$

In Section 6 we will show some examples where $\psi$ can be explicitly computed and shown to be differentiable on $\mathbb{R}$, and hence the above identities hold.

5 Asymptotic distance measures for an invariant alternative hypothesis

As the examples of Section 6 will show, the performance of the $G$-invariant and the unrestricted measurements can be very different in general. In particular, the states might be perfectly distinguishable by unrestricted measurements, while completely indistinguishable by $G$-invariant ones. As our following discussion shows, this cannot happen if the alternative hypothesis is invariant under the symmetry group. In the first part, we show that in this case $G$-invariant measurements perform just as well as unrestricted ones in the setting of Stein's lemma. This follows immediately from Theorem 5.1, thanks to Theorem 4.4. Although the same is not true for the settings of the Chernoff and the Hoeffding bounds (see Example 6.2), it is still possible to establish a strong relation between the different performances in the setting of the Chernoff bound as is shown the second part of this section.

5.1 Mean relative entropy

We prove the following partial extension of [18, Theorem 2.1], improving the arguments in [18] based on (2.4). In the proof, we use the same notations as defined in Section 2 for the irreducible decompositions of the tensor powers of the representation $u$.

Theorem 5.1. If $\rho_1$ is $G$-invariant (i.e., $\hat{\rho}_1 \in \mathcal{A}_1$), then

$$S(\rho_0 \parallel \rho_1) = S_M(\rho_0 \parallel \rho_1)$$

$$= \lim_{n \to \infty} \frac{1}{n} \sup \left\{ S(\rho_{0,n} | B \parallel \rho_{1,n} | B) : B \text{ is an abelian subalgebra of } \mathcal{A}_n \right\}.$$
Proof. The monotonicity of the relative entropy implies that
\[
S(\rho_0^{\otimes n} \parallel \rho_1^{\otimes n}) \geq S(\rho_{n,0} \parallel \rho_{n,1}) \geq S(\rho_{0,n}|_B \parallel \rho_{1,n}|_B)
\]
for any subalgebra \(B \subset A_n\) and hence,
\[
S(\rho_0 \parallel \rho_1) \geq S_M(\rho_0 \parallel \rho_1) \geq \lim_{n \to \infty} \frac{1}{n} \sup \{S(\rho_{0,n}|_B \parallel \rho_{1,n}|_B) : B \text{ is an abelian subalgebra of } A_n\}.
\]
The assumption \(\hat{\rho}_1 \in \mathcal{A}_1\) implies that \(\hat{\rho}_1^{\otimes n} \in A_n\) for all \(n \in \mathbb{N}\). By (2.2),
\[
\hat{\rho}_{1,n} = \hat{\rho}_1^{\otimes n} = \sum_{i=1}^{k_n} \left( D_i^{(n)} \otimes I_{d_i^{(n)}} \right) \text{ with } D_i^{(n)} \in M_{m_i^{(n)}}^{(n)}.
\]
With the spectral decomposition \(D_i^{(n)} = \sum_{j=1}^{\lambda_{ij}^{(n)}} \lambda_{ij}^{(n)} P_{ij}^{(n)}\) for \(1 \leq i \leq k_n\), we define
\[
\mathcal{E}_n(A) := \sum_{i=1}^{k_n} \sum_{j=1}^{\lambda_{ij}^{(n)}} (P_{ij}^{(n)} \otimes I_{d_i^{(n)}}) \mathcal{E}_{A_n}(A)(P_{ij}^{(n)} \otimes I_{d_i^{(n)}}), \quad A \in \mathcal{B}(\mathcal{H})^{\otimes n}.
\]
Let \(\mathcal{B}_n\) denote the abelian subalgebra of \(A_n\) generated by
\[
(P_{ij}^{(n)} \otimes I_{d_i^{(n)}}) \hat{\rho}_0^{\otimes n} (P_{ij}^{(n)} \otimes I_{d_i^{(n)}}), \quad 1 \leq i \leq k_n, \quad 1 \leq j \leq \lambda_{ij}^{(n)}.
\]
Then, as in the proof of [18, Lemma 3.1], we have
\[
nS(\rho_0 \parallel \rho_1) = S(\rho_0^{\otimes n} \parallel \rho_1^{\otimes n}) = S(\rho_0^{\otimes n}|_{\mathcal{B}_n} \parallel \rho_1^{\otimes n}|_{\mathcal{B}_n}) + S(\rho_0^{\otimes n} \circ \mathcal{E}_n) - S(\rho_0^{\otimes n}). \quad (5.2)
\]
Similarly to [18, Lemma 3.2] we next prove that
\[
S(\omega \circ \mathcal{E}_n) - S(\omega) \leq d \log(n+1) + 2 \log \left( \sum_{i=1}^{k_n} d_i^{(n)} \right) \quad (5.3)
\]
for any state \(\omega\) on \(\mathcal{B}(\mathcal{H})^{\otimes n}\). Note that \(S(\omega \circ \mathcal{E}_n) - S(\omega) = S(\omega \parallel \omega \circ \mathcal{E}_n)\) and hence, by the joint convexity of the relative entropy, it is enough to show (5.3) for pure states. Assume thus that \(\omega = |\psi\rangle \langle \psi|\) with some unit vector \(\psi \in \mathcal{H}^{\otimes n}\). By (2.3) we write
\[
\mathcal{E}_{A_n}(|\omega\rangle) = \sum_{i=1}^{k_n} E_i^{(n)}(A_i^{(n)}) \text{ with } A_i^{(n)} := |P_i^{(n)} \psi\rangle \langle P_i^{(n)} \psi| \in M_{m_i^{(n)}}^{(n)} \otimes M_{d_i^{(n)}}^{(n)}
\]
and \(E_i^{(n)}(A_i^{(n)}) = B_i^{(n)} \otimes I_{d_i^{(n)}}\) with \(B_i^{(n)} \in M_{m_i^{(n)}}^{(n)}\). The rank of \(B_i^{(n)}\) is equal to the Schmidt rank of the vector \(P_i^{(n)} \psi\), which is upper bounded by \(d_i^{(n)}\). Hence the rank of \(E_i^{(n)}(A_i^{(n)})\) is at most \((d_i^{(n)})^2\). One can also see that the number of different eigenvalues of \(\hat{\rho}_1^{\otimes n}\) is upper bounded by \((n+1)^d\) and hence \(d_i^{(n)} \leq (n+1)^d\). Hence, by (5.1), the rank of
\( \mathcal{E}_n(\psi\langle\psi|) \) is upper bounded by \((n+1)^d \sum_{i=1}^{k_n} (d_i^{(n)})^2 \) and therefore, \( S(\omega\mathcal{E}_n) - S(\omega) = S(\omega\mathcal{E}_n) \) is dominated by

\[
\log \left( (n+1)^d \sum_{i=1}^{k_n} (d_i^{(n)})^2 \right) \leq d \log(n+1) + 2 \log \left( \sum_{i=1}^{k_n} d_i^{(n)} \right). 
\]

Now we apply (5.3) to (5.2) with the choice \( \omega := \rho_0 \otimes \rho \) and use (2.4) to obtain

\[
S(\rho_0 \parallel \rho_1) \leq \liminf_{n \to \infty} \frac{1}{n} S(\rho_0,n \parallel \rho_1,n),
\]

which completes the proof. \( \square \)

### 5.2 Fidelity and Chernoff bound

In this section we will discuss the minimal symmetric error probability based on the relation between the fidelity and the Chernoff bound in the setting with group symmetry. The **fidelity** of two states \( \rho \) and \( \sigma \) on a matrix algebra is given by

\[
F(\rho, \sigma) := \text{Tr} \left| \hat{\rho}^{1/2} \hat{\sigma}^{1/2} \right| = \text{Tr}(\hat{\sigma}^{1/2} \hat{\rho}^{1/2})^{1/2} = \text{Tr}(\hat{\rho}^{1/2} \hat{\rho}^{1/2})^{1/2},
\]

which is used as a distinguishability measure in quantum hypothesis testing. It is known (see [10, Theorem 1], [25, (9.110)]) that

\[
1 - \sqrt{1 - F(\rho, \sigma)}^2 \leq P_{\min}(\rho : \sigma) \leq \frac{F(\rho, \sigma)}{2}.
\]

Furthermore, it is also well known (see [25, Theorem 9.6], [30, Theorem 6.2]) that the fidelity \( F(\rho, \sigma) \) is monotone increasing under trace-preserving completely positive maps. Hence we have

\[
F(\rho_0,n, \rho_1,n) \geq F(\rho_0 \otimes n, \rho_1 \otimes n) = F(\rho_0, \rho_1)^n. \quad (5.5)
\]

The following inequality was proved in [4, Theorem 6]. Here, we provide an alternative proof.

**Lemma 5.2.** For every states \( \rho \) and \( \sigma \) on \( B(\mathcal{H}) \) and for every \( 0 \leq s \leq 1 \),

\[
\text{Tr} \hat{\rho}^s \hat{\sigma}^{1-s} \geq F(\rho, \sigma)^2.
\]

**Proof.** By the fidelity formula with purifications due to Uhlmann (see [13, Lemma 8.2], [25, Theorem 9.4]), there are purifications \( |\varphi\rangle\langle\varphi| \) and \( |\psi\rangle\langle\psi| \) of \( \rho \) and \( \sigma \), respectively, such that

\[
F(\rho, \sigma) = |\langle\varphi, \psi|\rangle|.
\]

By Lieb’s concavity theorem (see the proof of Lemma 3.1 for details), we then have

\[
\text{Tr} \hat{\rho}^s \hat{\sigma}^{1-s} \geq \text{Tr}(|\varphi\rangle\langle\varphi|^s(|\varphi\rangle\langle\varphi|)^{1-s} = |\langle\varphi, \psi|\rangle|^2 = F(\rho, \sigma)^2.
\]

We need only the following (5.7) with \( s = 1/2 \) for later use, but the extended inequalities are of some interest in themselves.
Theorem 5.3. If $\rho_1$ is $G$-invariant (i.e., $\hat{\rho}_1 \in \mathcal{A}_1$), then

$$\lim inf \frac{1}{n} \log \text{Tr} |\hat{\rho}_{0,n}^{s} \hat{\rho}_{1,n}^{1-s}| \geq \log \text{Tr} |\hat{\rho}_{0}^{s} \hat{\rho}_{1}^{1-s}| \quad \text{if } 0 \leq s \leq 1/2, \quad (5.6)$$

$$\lim sup \frac{1}{n} \log \text{Tr} |\hat{\rho}_{0,n}^{s} \hat{\rho}_{1,n}^{1-s}| \leq \log \text{Tr} |\hat{\rho}_{0}^{s} \hat{\rho}_{1}^{1-s}| \quad \text{if } 1/2 \leq s \leq 1. \quad (5.7)$$

Proof. With the same notations as in Section 2, $\hat{\rho}_{0,n}$ is written as

$$\hat{\rho}_{0,n} = E_{\mathcal{A}_n}(\hat{\rho}_0^{\otimes n}) = \sum_{i=1}^{k_n} E_i^{(n)}(P_i^{(n)} \hat{\rho}_0^{\otimes n} P_i^{(n)}),$$

(see (2.3)) while $\hat{\rho}_{1,n} = \hat{\rho}_1^{\otimes n} \in \mathcal{A}_n$ by the assumption $\hat{\rho}_1 \in \mathcal{A}_1$.

First, we prove (5.7). Since $1 \leq 2s \leq 2$, note that $x^{2s}$ is an operator convex function on $[0, +\infty)$. Hence we have $\hat{\rho}_{0,n}^{2s} \leq E_{\mathcal{A}_n}((\hat{\rho}_0^{\otimes n})^{2s})$ so that

$$\text{Tr} |\hat{\rho}_{0,n}^{s} \hat{\rho}_{1,n}^{1-s}| = \text{Tr}(\hat{\rho}_{1,n}^{1-s} \hat{\rho}_{0,n}^{s} \rho_{1,n}^{1-s})^{1/2} \leq \text{Tr}(E_{\mathcal{A}_n}(\hat{\rho}_1^{\otimes n})^{2s} \rho_{1,n}^{1-s})^{1/2} = \text{Tr} \left( \sum_{i=1}^{k_n} E_i^{(n)}(P_i^{(n)} \hat{\rho}_{1,n}^{s} \rho_{1,n}^{2s} \rho_{1,n}^{1-s} P_i^{(n)}) \right)^{1/2}. \quad (5.8)$$

Set $A_i^{(n)} := P_i^{(n)} \hat{\rho}_{1,n}^{s} \rho_{1,n}^{2s} \rho_{1,n}^{1-s} P_i^{(n)}$ for $1 \leq i \leq k_n$. By Lemma A.3 of the Appendix, for $1 \leq i \leq k_n$ there are unitaries $U_{i,j}^{(n)} \in I_{m_i^{(n)}} \otimes M_{d_i^{(n)}}$, $1 \leq j \leq (d_i^{(n)})^2$, such that

$$E_i^{(n)}(A_i^{(n)}) = \frac{1}{(d_i^{(n)})^2} \sum_{j=1}^{(d_i^{(n)})^2} U_{i,j}^{(n)} A_i^{(n)} U_{i,j}^{(n)*}.$$ Hence we have

$$\text{Tr} \left( \sum_{i=1}^{k_n} E_i^{(n)}(A_i^{(n)}) \right)^{1/2} \leq \text{Tr} \left( \sum_{i=1}^{k_n} \sum_{j=1}^{(d_i^{(n)})^2} U_{i,j}^{(n)} A_i^{(n)} U_{i,j}^{(n)*} \right)^{1/2} \leq \sum_{i=1}^{k_n} \sum_{j=1}^{(d_i^{(n)})^2} \text{Tr}(U_{i,j}^{(n)} A_i^{(n)} U_{i,j}^{(n)*})^{1/2} = \sum_{i=1}^{k_n} (d_i^{(n)})^2 \text{Tr}(A_i^{(n)})^{1/2}. \quad (5.9)$$

In the above, the first inequality is just removing $1/(d_i^{(n)})^2$, and for the second inequality, see [5, Eq. (12)] (or an extended result in [2]). Combining (5.8) and (5.9) yields

$$\text{Tr} |\hat{\rho}_{0,n}^{s} \hat{\rho}_{1,n}^{1-s}| \leq \sum_{i=1}^{k_n} (d_i^{(n)})^2 \text{Tr}(P_i^{(n)} \hat{\rho}_{1,n}^{s} \rho_{1,n}^{2s} \rho_{1,n}^{1-s} P_i^{(n)})^{1/2} = \sum_{i=1}^{k_n} (d_i^{(n)})^2 \text{Tr}(\hat{\rho}_{1,n}^{s} \rho_{1,n}^{2s} \rho_{1,n}^{1-s} P_i^{(n)} (\hat{\rho}_{1,n}^{s} \rho_{1,n}^{2s} \rho_{1,n}^{1-s})^{1/2})^{1/2}.$$
\[
\leq \left( \sum_{i=1}^{k_n} d_i^{(n)} \right)^2 \text{Tr}(\rho_{1,n}^{1-s}(\rho_0^{\otimes n})^{2s}\rho_{1,n}^{1-s})^{1/2} \\
= \left( \sum_{i=1}^{k_n} d_i^{(n)} \right)^2 \text{Tr}|(\hat{\rho}_0^{\otimes n})^s(\hat{\rho}_1^{\otimes n})^{1-s}| \\
= \left( \sum_{i=1}^{k_n} d_i^{(n)} \right)^2 (\text{Tr} |\hat{\rho}_0^{\otimes n}\hat{\rho}_1^{\otimes n}|)^{1/2} 
\]

Thanks to (2.4) we obtain inequality (5.7).

Next, we prove (5.6). Since \(x^{2s}\) is operator concave on \([0, +\infty)\) thanks to \(0 \leq 2s \leq 1\), we have \(\rho_{0,n}^{2s} \geq E_{A_n}(\rho_0^{\otimes n})^{2s}\). Hence inequality (5.8) is reversed. Inequality (5.9) is also reversed as follows:

\[
\text{Tr}\left( \sum_{i=1}^{k_n} E_i^{(n)}(A_i^{(n)}) \right)^{1/2} = \text{Tr} \sum_{i=1}^{k_n} (E_i^{(n)}(A_i^{(n)}))^{1/2} \\
\geq \text{Tr} \sum_{i=1}^{k_n} E_i^{(n)}((A_i^{(n)})^{1/2}) \\
\geq \text{Tr} \sum_{i=1}^{k_n} P_i^{(n)}(\hat{\rho}_{1,n}^{1-s}(\rho_0^{\otimes n})^{2s}\rho_{1,n}^{1-s})^{1/2} P_i^{(n)} \\
= (\text{Tr} |\hat{\rho}_0^{\otimes n}\hat{\rho}_1^{\otimes n}|)^{1/2} 
\]

In the above, the first inequality follows by the operator concavity of the square root function, and the second one follows from [11]. Hence we have inequality (5.6).

Theorem 5.3 for \(s = 1/2\) together with (5.5) yields

**Corollary 5.4.** If \(\hat{\rho}_1 \in A_1\) then

\[
\lim_{n \to \infty} \frac{1}{n} \log F(\rho_{0,n}, \rho_{1,n}) = \inf_{n \geq 1} \frac{1}{n} \log F(\rho_{0,n}, \rho_{1,n}) = \log F(\rho_0, \rho_1). 
\]

Note that the logarithmic fidelity \(\log F(\cdot, \cdot)\) is a generalized relative entropy in the sense that (i) it takes strictly positive values on unequal states and zero if its arguments are equal, (ii) it is monotonically decreasing under trace-preserving completely positive maps, and (iii) it is jointly convex in its arguments. In view of this, Corollary 5.4 is a direct analogue of Theorem 5.1. Example 6.1 of the next section shows again that assuming the \(G\)-invariance of \(\rho_1\) is essential for Theorem 5.3 and Corollary 5.4.

As Example 6.2 shows, the \(G\)-invariance of \(\rho_1\) does not imply the same asymptotics for the restricted and the unrestricted minimal error probabilities. However, one can still obtain the following non-trivial bound:

**Theorem 5.5.** If \(\hat{\rho}_1 \in A_1\) then

\[
\lim_{n \to \infty} \frac{1}{n} \log P_{\min}(\rho_0^{\otimes n} : \rho_1^{\otimes n}) \leq \lim \inf_{n \to \infty} \frac{1}{n} \log P_{\min}(\rho_{0,n} : \rho_{1,n}) \\
\leq \lim \sup_{n \to \infty} \frac{1}{n} \log P_{\min}(\rho_{0,n} : \rho_{1,n}) 
\]
\[ \frac{1}{2} \lim_{n \to \infty} \frac{1}{n} \log P_{\min}(\rho_0^n : \rho_1^n) \leq \frac{1}{2} \lim_{n \to \infty} \frac{1}{n} \log P_{\min}(\rho_0 : \rho_1). \]

**Proof.** By Lemma 5.2 and Corollary 5.4 we have

\[ \frac{1}{2} \lim_{n \to \infty} \frac{1}{n} \log P_{\min}(\rho_0^n : \rho_1^n) = \frac{1}{2} \min_{0 \leq s \leq 1} \log \text{Tr} \rho_0^s \rho_1^{1-s} \]
\[ \geq \log F(\rho_0, \rho_1) \]
\[ = \lim_{n \to \infty} \frac{1}{n} \log F(\rho_{0,n}, \rho_{1,n}) \]
\[ \geq \limsup_{n \to \infty} \frac{1}{n} \log P_{\min}(\rho_{0,n} : \rho_{1,n}) \]

thanks to (5.4). Hence the last inequality follows, and the others are obvious. \qed

By Remark 4.7, the inequalities of the above theorem can be rewritten as

\[ -C(\rho_0, \rho_1) \leq C_G(\rho_0, \rho_1) \leq C_G(\rho_0, \rho_1) \leq -\frac{1}{2} C(\rho_0, \rho_1). \]

Comparing these with the inequalities of Proposition 4.2 and also taking into account (4.9), we have the following:

**Corollary 5.6.** Assume that \( \rho_1 \) is \( G \)-invariant. Then

\[ \frac{1}{4} C(\rho_0, \rho_1) \leq C_M(\rho_0, \rho_1) \leq C(\rho_0, \rho_1). \]

Moreover, \( \frac{1}{2} C(\rho_0, \rho_1) \leq C_M(\rho_0, \rho_1) \) holds whenever \( \psi \) is differentiable on \((0, 1)\).

**Remark 5.7.** The constant \( 1/2 \) in Theorem 5.5 is actually the best possible, as will be seen in Remark 6.4 of the next section. This also shows that in the case of a differentiable \( \psi \), the bound \( \frac{1}{2} C(\rho_0, \rho_1) \leq C_M(\rho_0, \rho_1) \leq C(\rho_0, \rho_1) \) is the best possible.

### 6 Restricted vs. unrestricted measurements: examples

In this section, we will illustrate, through some examples, the difference between the performance of \( G \)-invariant measurements and that of unrestricted ones. As the following example shows, the difference can be as extreme as possible even in the classical situation where the densities corresponding to the null and the alternative hypotheses are commuting. This also shows that the assumption that \( \rho_1 \) is \( G \)-invariant cannot be removed in Theorem 5.1.

**Example 6.1.** (Two commuting states with \( \mathbb{Z}_2 \)-symmetry) Let \( G := \mathbb{Z}_2 = \{ \pm 1 \} \) and \( u \) be the representation of \( G \) on \( \mathcal{H} := \mathbb{C}^2 \) with \( u_{-1} := \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \). Then the \( \mathbb{Z}_2 \)-fixed point subalgebra of \( B(\mathcal{H})^\otimes n \) is \( \mathcal{A}_n = M_{2^n-1} \oplus M_{2^n-1} \). Consider a commuting set of states \( \sigma_\lambda \), \( 0 \leq \lambda \leq 1 \), given by

\[ \hat{\sigma}_\lambda := \lambda |+\rangle\langle +| + (1 - \lambda) |-\rangle\langle -| = \lambda \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix} + (1 - \lambda) \begin{bmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \end{bmatrix}, \]
where \(|+\rangle := (1/\sqrt{2}, 1/\sqrt{2})\) and \(|-\rangle := (1/\sqrt{2}, -1/\sqrt{2})\). We write

\[
\tilde{\sigma}_\lambda^{\otimes n} = \sum_{i=0}^{n} \lambda^i (1 - \lambda)^{n-i} E_{n,i} \quad \text{with} \quad E_{n,i} := \sum_{e_j = \pm, \#\{e_j = +\} = i} \otimes_{j=1}^{n} |e_j\rangle\langle e_j|.
\]

Since \(u_1^{\otimes n} u_1^{\otimes n} = E_{n,n-i}\), we have

\[
E_{A,n}(\tilde{\sigma}_\lambda^{\otimes n}) = \frac{1}{2}(\tilde{\sigma}_\lambda^{\otimes n} + u_1^{\otimes n} \tilde{\sigma}_\lambda^{\otimes n} u_1^{\otimes n}) = \sum_{i=0}^{n} \lambda^i (1 - \lambda)^{n-i} + \lambda^{n-i} (1 - \lambda)^{i} E_{n,i}. \tag{6.1}
\]

Now let \(\rho_0 := \sigma_\lambda\) and \(\rho_1 := \sigma_\mu\) with any \(\lambda, \mu \in [0, 1]\). When \(0 \leq s \leq 1\), noting that \(2^{s-1}(a^s + b^s) \leq (a + b)^s \leq a^s + b^s\) for all \(a, b \geq 0\) and that \(\text{Tr} E_{n,i} = \binom{n}{i}\), we see from (6.1) that \(\text{Tr} \rho_0^{\otimes n} \rho_1^{1-s} n\) is upper bounded by

\[
\frac{1}{2} \sum_{i=0}^{n} \frac{1}{n} \left\{ (\lambda^s(1 - \lambda)^{1-s})^i + (\lambda^{1-s}(1 - \lambda)^s)^i \right\} \binom{n}{i} = \frac{1}{2} \sum_{i=0}^{n} \frac{1}{n} \left\{ (\lambda^s(1 - \lambda)^{1-s})^i + (\lambda^{1-s}(1 - \lambda)^s)^i \right\} \binom{n}{i} = (\lambda^s(1 - \lambda)^{1-s} + (\lambda^{1-s}(1 - \lambda)^s)^n) \tag{6.2}
\]

and also lower bounded by \(1/2\) times (6.2). Therefore,

\[
\psi(s) = \max \left\{ \log(\lambda^s(1 - \lambda)^{1-s}) + (\lambda^s(1 - \lambda)^{1-s} + (\lambda^{1-s})^{s s^{1-s}}) \right\}
\]

where \(\psi(s \mid \sigma_\lambda \| \sigma_\mu)\) denotes \(\psi(s)\) for \(\rho_0 = \sigma_\lambda\) and \(\rho_1 = \sigma_\mu\). Note that

\[
\psi(s) = \begin{cases} 
\psi(0 \mid \sigma_\lambda \| \sigma_\mu) & \text{if } (1/2 - \lambda)(1/2 - \mu) \geq 0, \\
\psi(0 \mid \sigma_\lambda \| \sigma_1 - \mu) & \text{if } (1/2 - \lambda)(1/2 - \mu) < 0.
\end{cases} \tag{6.3}
\]

In particular, \(\psi\) is differentiable on \((0, 1)\). Note that \(\sigma_1 - \mu = \sigma_\mu \circ \text{Ad} u_1\), and (6.3) yields

\[
C_M(\rho_0, \rho_1) = \min \{ C(\sigma_\lambda, \sigma_\mu), C(\sigma_\lambda, \sigma_1 - \mu) \}, \quad H_M(r \| \rho_0 \| \rho_1) = \min \{ H(r \mid \sigma_\lambda \| \sigma_\mu), H(r \mid \sigma_\lambda \| \sigma_1) \}, \quad S_M(\rho_0 \| \rho_1) = \min \{ S(\sigma_\lambda \| \sigma_\mu), S(\sigma_\lambda \| \sigma_1 - \mu) \},
\]

and hence the first inequality in each of (3.11)–(3.13) hold with equality. On the other hand, (6.3) shows that if \((1/2 - \lambda)(1/2 - \mu) < 0\) then

\[
\psi(s) = \psi^0(s \mid \sigma_\lambda \| \sigma_1 - \mu) > \psi^0(s \mid \sigma_\lambda \| \sigma_\mu) = \psi(s)
\]

so that for any \(r \geq 0\),

\[
C_M(\rho_0, \rho_1) < C(\rho_0, \rho_1), \quad H_M(r \| \rho_0 \| \rho_1) < H(r \| \rho_0 \| \rho_1), \quad S_M(\rho_0 \| \rho_1) < S(\rho_0 \| \rho_1). \tag{6.4}
\]

The differentiability of \(\psi\) on \((0, 1)\) implies that the identities of Remark 4.7 hold (as long as \(0 < \mu < 1\)), and hence (6.4) shows that \(G\)-invariant measurements perform strictly worse than
unrestricted ones in all of the settings of the Chernoff and Hoeffding bounds and of Stein’s lemma. In particular, in the extremal case where \( \rho_0 = \sigma_0 \) and \( \rho_1 = \sigma_1 \), the two states have orthogonal supports and hence unrestricted measurements yield a perfect distinguishability, while one can easily see that \( \rho_{0,n} = \rho_{1,n} \) for all \( n \) so that the states are completely indistinguishable with \( G \)-invariant measurements. Note that in this case \( \psi(s) = 0 \) while \( \psi^\circ(s) = -\infty \) for all \( s \in \mathbb{R} \).

Finally, we show for completeness that \( \psi(s) \) exists and is differentiable on the whole real line. Replacing \( \lambda \) by \( 1 - \lambda \), \( \mu \) by \( 1 - \mu \) and \( s \) by \( 1 - s \) if necessary, we may assume that \( 0 < \lambda < \mu \leq 1/2 \), since the cases \( \lambda = 0 \) and \( \lambda = \mu \) are easy to verify. The \( \psi(s) \) for \( s \in [0, 1] \) has been computed above. When \( s \leq 0 \), since

\[
\frac{\lambda^i(1 - \lambda)^{n-i} + \lambda^{n-i}(1 - \lambda)^i}{2} \leq \lambda^i(1 - \lambda)^{n-i}, \quad 1 \leq i \leq \lfloor n/2 \rfloor,
\]

we see from (6.1) that \( \text{Tr} \rho_{0,n}^s \rho_{1,n}^{1-s} \) is lower bounded by

\[
\sum_{i=0}^{\lfloor n/2 \rfloor} \left( \frac{\lambda^i(1 - \lambda)^{n-i}}{2} \right)^s (\mu^i(1 - \mu)^{n-i})^{1-s} \binom{n}{i},
\]

It is also upper bounded by

\[
\sum_{i=0}^{\lfloor n/2 \rfloor} \left( \frac{\lambda^i(1 - \lambda)^{n-i}}{2} \right)^s (\mu^i(1 - \mu)^{n-i})^{1-s} \binom{n}{i} + \sum_{i=\lfloor n/2 \rfloor + 1}^{n} \left( \frac{\lambda^{n-i}(1 - \lambda)}{2} \right)^s (\mu^{n-i}(1 - \mu))^i^{1-s} \binom{n}{i}
\leq 2 \sum_{i=0}^{\lfloor n/2 \rfloor} \left( \frac{\lambda^i(1 - \lambda)^{n-i}}{2} \right)^s (\mu^i(1 - \mu)^{n-i})^{1-s} \binom{n}{i}.
\]

Hence, by Lemma A.5 of the Appendix we have

\[
\psi(s) = \lim_{n \to \infty} \frac{1}{n} \log \sum_{i=0}^{\lfloor n/2 \rfloor} \binom{n}{i} (\lambda^s \mu^{1-s})^i (1 - \lambda)^s (1 - \mu)^{1-s} n^{-i}
= \begin{cases} \log(\lambda^s \mu^{1-s}) + (1 - \lambda)^s (1 - \mu)^{1-s} & \text{if } s^* \leq s \leq 0, \\ \frac{s}{2} \log \lambda(1 - \lambda) + \frac{1-s}{2} \log \mu(1 - \mu) + \log 2 & \text{if } s \leq s^*, \end{cases}
\]

where \( s^* \in (-\infty, 0] \) satisfies

\[
\lambda^{s^*} \mu^{1-s^*} = (1 - \lambda)^{s^*} (1 - \mu)^{1-s^*} \quad \text{or} \quad \left( \frac{(1 - \lambda) \mu}{\lambda(1 - \mu)} \right)^{s^*} = \frac{\mu}{1 - \mu}.
\]

When \( s \geq 1 \), the computation using Lemma A.5 is similar. Summing up all, we write

\[
\psi(s) = \begin{cases} \log(\lambda^s \mu^{1-s}) + (1 - \lambda)^s (1 - \mu)^{1-s} & \text{if } s \geq s^*, \\ \frac{s}{2} \log \lambda(1 - \lambda) + \frac{1-s}{2} \log \mu(1 - \mu) + \log 2 & \text{if } s \leq s^*, \end{cases}
\]

which shows the differentiability of \( \psi(s) \) at any \( s \in \mathbb{R} \) including \( s = s^* \).
The next example shows that in the settings of the Chernoff and the Hoeffding bounds the restricted measurements may yield a strictly worse performance even if \( \rho_1 \) is \( G \)-invariant.

**Example 6.2. (A pure state vs. an invariant mixed state with \( \mathbb{T} \)-symmetry)** Let \( \mathcal{H} := \mathbb{C}^2 \) and the states to discriminate be \( \hat{\rho}_0 := \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix} \) and \( \hat{\rho}_1 := \begin{bmatrix} \alpha & 0 \\ 0 & 1 - \alpha \end{bmatrix} \) with \( 0 < \alpha < 1 \). In the unrestricted scenario, we have

\[
\psi^\circ(s) := \log \text{Tr} \hat{\rho}_0^s \hat{\rho}_1^{1-s} = \log \frac{\alpha^{1-s} + (1 - \alpha)^{1-s}}{2}.
\]

Since
\[
d\frac{d}{ds}(\alpha^{1-s} + (1 - \alpha)^{1-s}) = -\alpha^{1-s} \log \alpha - (1 - \alpha)^{1-s} \log(1 - \alpha) > 0,
\]
we get
\[
C(\rho_0, \rho_1) = - \min_{0 \leq s \leq 1} \psi^\circ(s) = -\psi^\circ(0) = \log 2.
\]

Hence, by (4.9) (in the unrestricted setting [3, 26]),
\[
\lim_{n \to \infty} \frac{1}{n} \log P_{\min}(\rho_0^{\otimes n} : \rho_1^{\otimes n}) = -C(\rho_0, \rho_1) = -\log 2. \tag{6.5}
\]

Now let \( G := \mathbb{T} = \{ \zeta \in \mathbb{C} : |\zeta| = 1 \} \) and define the unitary representation \( u_\zeta := \begin{bmatrix} 1 & 0 \\ 0 & \zeta \end{bmatrix}, \zeta \in \mathbb{T}, \) on \( \mathcal{H} \). It is easy to see that \( u_\zeta^{\otimes n} \) is diagonal with \( 1, \zeta, \ldots, \zeta^n \) standing in the diagonal entries, and \( \zeta^i \) appears exactly \( \binom{n}{i} \) times. Hence,
\[
\mathcal{A}_n = \bigoplus_{i=0}^{n} M_{\binom{n}{i}},
\]
and one can also see that the Bratteli diagrams of the inclusions \( \mathcal{C} \mathcal{I} \subset \mathcal{A}_1 \subset \mathcal{A}_2 \subset \cdots \) form the Pascal triangle. Note that \( \rho_1 \) is \( G \)-invariant. The \( G \)-invariant reductions of \( \rho_0^{\otimes n} \) and \( \rho_1^{\otimes n} \) are given by

\[
\hat{\rho}_{0,n} = \mathcal{E}_{\mathcal{A}_n}(\hat{\rho}_0^{\otimes n}) = \sum_{i=0}^{n} \binom{n}{i} \frac{1}{2^n} P_{n,i}, \quad \hat{\rho}_{1,n} = \hat{\rho}_1^{\otimes n} = \sum_{i=0}^{n} \alpha^i (1 - \alpha)^{n-i} E_{n,i}, \tag{6.6}
\]

where \( E_{n,i} \) is the identity of \( M_{\binom{n}{i}} \) with \( \sum_{i=0}^{n} E_{n,i} = I_n \), and \( P_{n,i} \) is a rank one projection with \( P_{n,i} \leq E_{n,i} \). Therefore,

\[
\text{Tr} \hat{\rho}_{0,n}^{s} \hat{\rho}_{1,n}^{1-s} = \sum_{i=0}^{n} \binom{n}{i} \frac{1}{2^n} (\alpha^i (1 - \alpha)^{n-i})^{1-s}, \quad s \in \mathbb{R}.
\]

Take \( a = \alpha^{1-s} \) and \( b = (1 - \alpha)^{1-s} \) in Lemma A.4 of the Appendix to obtain

\[
\psi(s) = \begin{cases} 
  s \log(\alpha^{1-s} + (1 - \alpha)^{1-s}) - s \log 2 & \text{if } s > 0, \\
  (1 - s) \log \max\{\alpha, 1 - \alpha\} - s \log 2 & \text{if } s \leq 0.
\end{cases}
\]

It is obvious that \( \psi \) is differentiable at any \( s \neq 0 \). To check the differentiability at \( s = 0 \), assume \( \alpha > 1 - \alpha \) and set \( \beta := (1 - \alpha)/\alpha \in (0, 1) \). Then, for \( s > 0 \) we have

\[
\psi(s) = (1 - s) \log \alpha + s \log(1 + \beta^{1-s}) - s \log 2,
\]

22
and the differentiability at \( s = 0 \) follows from

\[
\lim_{s \searrow 0} \frac{s \log(1 + \beta^{\frac{1-s}{x}})}{s} = \lim_{s \searrow 0} \log(1 + \beta^{\frac{1-s}{x}}) = 0.
\]

The case \( \alpha < 1 - \alpha \) goes in the same way, and the case \( \alpha = 1/2 \) is easy to verify. Consequently, \( \psi \) exists and is differentiable on the whole real line, and hence the identities of Remark 4.7 hold. Moreover, by Lemma 3.6 and Theorem 5.1,

\[
S_M(\rho_0 \parallel \rho_1) = S(\rho_0 \parallel \rho_1) = \psi'(1) = -\frac{\log \alpha + \log(1 - \alpha)}{2}
\]

so that the error exponents for Stein’s lemma are the same in the unrestricted and the \( G \)-invariant cases.

If \( \alpha = 1/2 \) (i.e., \( \hat{\rho}_1 = 2^{-1}I_2 \)), then \( \psi(s) = -(1-s) \log 2 = \psi^\circ(s) \) for all \( s \geq 0 \) so that \( C_M(\rho_0, \rho_1) = C(\rho_0, \rho_1) \) and \( H_M(r \mid \rho_0 \parallel \rho_1) = H(r \mid \rho_0 \parallel \rho_1) \) for all \( r \geq 0 \). Therefore, by (4.9) and (6.5),

\[
\lim_{n \to \infty} \frac{1}{n} \log P_{\min}(\rho_{0,n} : \rho_{1,n}) = -\log 2 = \lim_{n \to \infty} \frac{1}{n} \log P_{\min}(\rho_0^\otimes n : \rho_1^\otimes n).
\]

Assume for the rest that \( \alpha \neq 1/2 \). Then

\[
\psi(0) = \log \max\{\alpha, 1 - \alpha\} > -\log 2 = \psi^\circ(0), \quad \psi(1) = 0 = \psi^\circ(1),
\]

and for any \( s \in (0, 1) \),

\[
\psi(s) = \log \left( \frac{\alpha^{\frac{1-s}{x}} + (1-\alpha)^{\frac{1-s}{x}}}{2} \right)^{s} > \log \frac{\alpha^{1-s} + (1-\alpha)^{1-s}}{2} = \psi^\circ(s)
\]

thanks to the strict concavity of \( x \mapsto x^s, x \geq 0 \). Therefore, again by (4.9) and (6.5),

\[
\lim_{n \to \infty} \frac{1}{n} \log P_{\min}(\rho_{0,n} : \rho_{1,n}) = -C_M(\rho_0, \rho_1) > -C(\rho_0, \rho_1) = \lim_{n \to \infty} \frac{1}{n} \log P_{\min}(\rho_0^\otimes n : \rho_1^\otimes n).
\]

Furthermore, for any \( r > -\psi(1) = 0 \) we have \( H_M(r \mid \rho_0 \parallel \rho_1) = (-s_0r - \psi(s_0))/(1 - s_0) \) for some \( s_0 \in [0, 1) \) (see the proof of Lemma A.2), and hence

\[
H_M(r \mid \rho_0 \parallel \rho_1) < \frac{-s_0r - \psi^\circ(s_0)}{1 - s_0} \leq H(r \mid \rho_0 \parallel \rho_1)
\]

by (3.10). Remark 4.7 then implies that

\[
h_c^G(r \mid \rho_0 \parallel \rho_1) = \overline{h}_c^G(r \mid \rho_0 \parallel \rho_1) = h_c^G(r \mid \rho_0 \parallel \rho_1) = -H_M(r \mid \rho_0 \parallel \rho_1)
\]

\[
> h(r \mid \rho_0 \parallel \rho_1) = \overline{h}(r \mid \rho_0 \parallel \rho_1) = h(r \mid \rho_0 \parallel \rho_1) = -H(r \mid \rho_0 \parallel \rho_1), \quad r > 0.
\]

That is, in both settings of the Chernoff and the Hoeffding bounds, the optimal performance of the \( G \)-invariant measurements is strictly worse than that of the unrestricted ones. Moreover, since the alternative hypothesis in this example is \( G \)-invariant, one sees immediately that the inequalities

\[
C_M(\rho_0, \rho_1) \leq \inf_{g \in G} \mathcal{C}(\rho_0, \rho_1 \circ \text{Ad } u_g), \quad H_M(r \mid \rho_0 \parallel \rho_1) \leq \inf_{g \in G} H(r \mid \rho_0 \parallel \rho_1 \circ \text{Ad } u_g)
\]

of (3.11) and (3.12) cannot hold as an equality in general.
**Remark 6.3.** Note that if we reduce the group $G$ in the above example, then the difference between the optimal performances of the restricted and the unrestricted measurements may disappear. Indeed, consider the subgroup $G := \mathbb{Z}_2 = \{\pm 1\}$ of $\mathbb{T}$ with the same representation as in Example 6.1. Let $\rho_0$ and $\rho_1$ be the same as in Example 6.2. Then $\mathcal{A}_n = M_{2^n-1} \oplus M_{2^n-1}$ as in Example 6.1 (but with a different arrangement of basis) and $\rho_{1,n}$ is the same as in (6.6) while $\rho_{0,n}$ is given by

$$\hat{\rho}_{0,n} = \left( \frac{1}{2n} J_{2^n-1} \right) \oplus \left( \frac{1}{2n} J_{2^n-1} \right),$$

where $J_{2^n-1}$ is the $2^{n-1} \times 2^{n-1}$ matrix of all entries equal to one. Since

$$\hat{\rho}_{0,n}^s = \left( \frac{1}{2^{s+n-1}} J_{2^n-1} \right) \oplus \left( \frac{1}{2^{s+n-1}} J_{2^n-1} \right),$$

we have

$$\operatorname{Tr} \hat{\rho}_{0,n}^s \rho_{1,n} = \frac{1}{2^{s+n-1}} \sum_{i=0}^{n} \left( \frac{n}{i} \right) (\alpha^{n-i}(1-\alpha)^i)^{1-s} = \frac{(\alpha^{1-s} + (1-\alpha)^{1-s})^n}{2^{s+n-1}}.$$

Hence the $\psi$-function in this case is equal to $\psi^\circ$:

$$\psi(s) = \lim_{n \to \infty} \frac{1}{n} \log \operatorname{Tr} \hat{\rho}_{0,n}^s \rho_{1,n} = \log \frac{\alpha^{1-s} + (1-\alpha)^{1-s}}{2} = \psi^\circ(s),$$

and therefore, by Remark 4.7,

$$\lim_{n \to \infty} \frac{1}{n} \log P_{\min}(\rho_{0,n} : \rho_{1,n}) = \lim_{n \to \infty} \frac{1}{n} \log P_{\min}(\rho_{0,n}^\circ : \sigma_0^\circ).$$

**Remark 6.4.** In the setting of Example 6.2 we have seen that

$$\min_{0 \leq s \leq 1} \psi(s) \leq \psi(1/2) = \frac{1}{2} \log 2 = \frac{1}{2} \min_{0 \leq s \leq 1} \psi^\circ(s).$$

On the other hand,

$$\psi'(s) = \log(\alpha^{\frac{1-s}{s}} + (1-\alpha)^{\frac{1-s}{s}}) - \frac{1}{s} \frac{\alpha^{\frac{1-s}{s}} \log \alpha + (1-\alpha)^{\frac{1-s}{s}} \log (1-\alpha)}{\alpha^{\frac{1-s}{s}} + (1-\alpha)^{\frac{1-s}{s}}} - \log 2$$

and hence

$$\psi'(1/2) = -2(\alpha \log \alpha + (1-\alpha) \log (1-\alpha)) - \log 2.$$

This shows that there exists an $\alpha^* \in (0, 1)$ such that $\psi'(1/2) = 0$ for $\alpha = \alpha^*$ (a numerical computation shows $\alpha^* \approx 0.11$). Hence, if $\hat{\rho}_1 = \begin{bmatrix} \alpha^* & 0 \\ 0 & 1-\alpha^* \end{bmatrix}$, then $\psi'(1/2) = 0$, and $\psi$ takes the minimum at $s = 1/2$ so that

$$C_M(\rho_0, \rho_1) = \min_{0 \leq s \leq 1} \psi(s) = \frac{1}{2} \min_{0 \leq s \leq 1} \psi^\circ(s) = \frac{1}{2} C(\rho_0, \rho_1).$$

Comparing this with (4.9), we see that the constant $1/2$ in Theorem 5.5 is the best possible.

Finally, we consider the discrimination problem of two pure states of a spin-$\frac{1}{2}$ system with $G$-invariant measurements, with $G$ given as in Example 6.2.
Example 6.5. (Two pure states with \(\mathbb{T}\)-symmetry)  In the same setting as in Example 6.2, let \(\rho_0\) and \(\rho_1\) be pure states with densities

\[
\hat{\rho}_0 := \begin{bmatrix} \frac{\lambda}{\sqrt{\lambda(1-\lambda)}} & \frac{\sqrt{\lambda(1-\lambda)}}{1-\lambda} \\ \frac{\sqrt{\lambda(1-\lambda)}}{1-\lambda} & \frac{\mu}{\sqrt{\mu(1-\mu)}} & \frac{\sqrt{\mu(1-\mu)}}{1-\mu} \end{bmatrix}, \quad \hat{\rho}_1 := \begin{bmatrix} \frac{\mu}{\sqrt{\mu(1-\mu)}} & \frac{\sqrt{\mu(1-\mu)}}{1-\mu} \\ \frac{\sqrt{\mu(1-\mu)}}{1-\mu} & \frac{\lambda}{\sqrt{\lambda(1-\lambda)}} & \frac{\sqrt{\lambda(1-\lambda)}}{1-\lambda} \end{bmatrix},
\]

where \(\lambda, \mu \in (0,1), \lambda \neq \mu\). By looking at how \(A_n\) in Example 6.2 is decomposed into the direct summands \(M(n)_i\) for \(0 \leq i \leq n\) and also by looking at the entries of \(\hat{\rho}_0\otimes^n\) and \(\hat{\rho}_1\otimes^n\), it is easy to see that

\[
\hat{\rho}_{0,n} = \bigoplus_{i=0}^{n} \lambda^{n-i}(1-\lambda)^i J(n)_i, \quad \hat{\rho}_{1,n} = \bigoplus_{i=0}^{n} \mu^{n-i}(1-\mu)^i J(n)_i,
\]

where \(J(n)_i\) is the \((n)\times(n)_i\) matrix of all entries equal to one. Therefore, for any \(s \in \mathbb{R}\),

\[
\text{Tr} \hat{\rho}_{0,n} \hat{\rho}_{1,n}^{1-s} = \sum_{i=0}^{n} (\lambda^{n-i}(1-\lambda)^i)(\mu^{n-i}(1-\mu)^i)^{1-s} \text{ Tr} J(n)_i
\]

\[
= \sum_{i=0}^{n} \binom{n}{i} (\lambda^s \mu^{1-s})^{n-i}(1-\lambda)^s(1-\mu)^{1-s})^i
\]

\[
= (\lambda^s \mu^{1-s} + (1-\lambda)^s(1-\mu)^{1-s})^n.
\]

Hence we have

\[
\psi(s) = \frac{1}{n} \log \text{Tr} \hat{\rho}_{0,n} \hat{\rho}_{1,n}^{1-s} = \log(\lambda^s \mu^{1-s} + (1-\lambda)^s(1-\mu)^{1-s})
\]

for all \(s \in \mathbb{R}\) and all \(n \in \mathbb{N}\). Consequently, \(\psi\) is differentiable on \(\mathbb{R}\) and by (4.9),

\[
\lim_{n \to \infty} \frac{1}{n} \log P_{\min}(\rho_{0,n} : \rho_{1,n}) = \log \min_{0 \leq s \leq 1} (\lambda^s \mu^{1-s} + (1-\lambda)^s(1-\mu)^{1-s}).
\]

On the other hand, since

\[
\psi^\circ(s) := \log \text{Tr} \hat{\rho}_{0}^{\circ n} \hat{\rho}_{1}^{\circ n} = \log \text{Tr} \hat{\rho}_0 \hat{\rho}_1 = \log(\sqrt{\lambda\mu} + \sqrt{(1-\lambda)(1-\mu)} )^2,
\]

we have

\[
\lim_{n \to \infty} \frac{1}{n} \log P_{\min}(\rho_{0}^{\circ n} : \rho_{1}^{\circ n}) = 2 \log(\sqrt{\lambda\mu} + \sqrt{(1-\lambda)(1-\mu)} )^2.
\]

We notice that

\[
\min_{0 \leq s \leq 1} (\lambda^s \mu^{1-s} + (1-\lambda)^s(1-\mu)^{1-s}) > (\sqrt{\lambda\mu} + \sqrt{(1-\lambda)(1-\mu)} )^2 \quad (6.7)
\]

so that

\[
\lim_{n \to \infty} \frac{1}{n} \log P_{\min}(\rho_{0,n} : \rho_{1,n}) > \lim_{n \to \infty} \frac{1}{n} \log P_{\min}(\rho_{0}^{\circ n} : \rho_{1}^{\circ n})
\]

similarly to Example 6.2. An elementary proof of (6.7) is as follows: Since

\[
(\lambda^s \mu^{1-s} + (1-\lambda)^s(1-\mu)^{1-s})(\lambda^{1-s} \mu^s + (1-\lambda)^{1-s}(1-\mu)^s)
\]

\[
- (\sqrt{\lambda\mu} + \sqrt{(1-\lambda)(1-\mu)} )^2
\]
Note also that we have $\psi(0) = \psi(1) = 0$ and $\psi'(0) = -S((\mu, 1 - \mu) \parallel (\lambda, 1 - \lambda))$, $\psi'(1) = S((\lambda, 1 - \lambda) \parallel (\mu, 1 - \mu))$, where $S((\lambda, 1 - \lambda) \parallel (\mu, 1 - \mu))$ is the relative entropy of $(\lambda, 1 - \lambda)$ and $(\mu, 1 - \mu)$. We have

$$S(\rho_{0,n} \parallel \rho_{1,n}) = \sum_{k=0}^{n} \binom{n}{k} \lambda^{n-k}(1 - \lambda)^{k}\log \frac{(n)}{(k)} \lambda^{n-k}(1 - \lambda)^{k}$$

$$= \sum_{k=0}^{n} \binom{n}{k} \lambda^{n-k}(1 - \lambda)^{k}\left( (n-k) \log \frac{\lambda}{\mu} + k \log \frac{1 - \lambda}{1 - \mu}\right)$$

$$= n\lambda \log \frac{\lambda}{\mu} + n(1 - \lambda) \log \frac{1 - \lambda}{1 - \mu}$$

for all $n \in \mathbb{N}$ and hence,

$$S_M(\rho_0 \parallel \rho_1) = S((\lambda, 1 - \lambda) \parallel (\mu, 1 - \mu)) = \psi'(1).$$

On the other hand, $S(\rho_0 \parallel \rho_1) = +\infty > S_M(\rho_0 \parallel \rho_1)$. This shows again that Theorem 5.1 cannot generally hold, that is, the assumption $\rho_1 \in A_1$ cannot be removed.

7 Concluding remarks and problems

The asymptotic binary state discrimination problem can be formulated in a very general way, with two sequences $\{\rho_{0,n}\}_{n \in \mathbb{N}}$ and $\{\rho_{1,n}\}_{n \in \mathbb{N}}$ to be discriminated. Here, $\rho_{0,n}$ and $\rho_{1,n}$ are states on $B(\mathcal{H}_n)$ of some Hilbert space $\mathcal{H}_n$, where $\mathcal{H}_{n+m} = \mathcal{H}_n \otimes \mathcal{H}_m$ need not be assumed; see, e.g., [24, 17]. To get a complete solution of the problems of the Chernoff and the Hoeffding bounds and of Stein’s lemma, i.e., to show the identities of Remark 4.7, the key point is to show that (4.4) holds with equality for a suitable range of parameters $\alpha$. This can be done, for instance, by showing that the states satisfy a certain factorization property [16, 17] or that the function $\psi$ exists and is differentiable on the interval $(0, 1)$ [17, 21, 22]. As the examples of Section 6 suggest, one can expect the differentiability of $\psi$ to hold in our setting of group-invariant state discrimination. The main open question of the present work is to show that this is indeed the case.

Once the $\psi$-function (and, in an optimal situation, its differentiability) is obtained, the identities of the error exponents and the corresponding asymptotic statistical distances follow quite automatically. In this sense, the results of Section 4 do not depend much on our present setting of group-symmetric measurements. On the other hand, the determination of the $\psi$-function requires significantly different techniques in the different scenarios; see, e.g., [17, 21, 22] for example. The key technical tool that we used in the present setting is the monotonicity of quasi-entropies under stochastic maps. It is worthwhile to note that our
analysis works whenever the states to discriminate are defined on a sequence of algebras \( \{A_n\}_{n \in \mathbb{N}} \) that satisfy \( A_n \otimes A_m \subset A_{n+m} \).

Stein’s lemma can be considered as an extremal point of the family of the Hoeffding bounds. In this sense, Stein’s lemma is the most asymmetric scenario, which gives a heuristic support for the validity of Theorem 5.1 which says that \( G \)-invariance of the measurements does not mean a real restriction as long as the alternative hypothesis is also \( G \)-invariant. It is, however, somewhat surprising that such an asymmetric condition can yield the bound of Theorem 5.5 in the totally symmetric discrimination problem of the Chernoff bound. On the other hand, we conjecture that if the symmetry group \( G \) is finite and the alternative hypothesis is \( G \)-invariant, then the function \( \psi \) is actually equal to \( \psi^0 \) of the unrestricted setting, and hence \( G \)-invariant measurements yield the same optimal error exponents as the unrestricted ones. This would of course also implies that all the inequalities of (3.11)–(3.13) hold with equality.

**Acknowledgments**

Partial funding was provided by the Grant-in-Aid for Scientific Research (B) 17340043 (F.H.); the Grant-in-Aid for JSPS Fellows 18-06916, the Hungarian Research Grant OTKA T068258 (M.M.); the Grant-in-Aid for Scientific Research on Priority Area “Deepening and Expansion of Statistical Mechanical Informatics (DEX-SMI)” 18079014, the MEXT Grant-in-Aid for Young Scientists (A) 20686026 (M.H.). This work was also partially supported by the JSPS Japan-Hungary Joint Project (F.H. & M.H.). Part of this work was done while M.M. was a Junior Research Fellow at the Erwin Schrödinger Institute for Mathematical Physics in Vienna.

**Appendix A**

The appendix supplies, for the reader’s convenience, the details on some technical lemmas used in the main body of the paper.

**A.1 Lieb’s concavity and Ando’s convexity**

Let \( B(\mathcal{H})_+ \) be the set of \( A \geq 0 \) in \( B(\mathcal{H}) \) and \( B(\mathcal{H})_{++} \) the set of strictly positive \( A \in B(\mathcal{H}) \). Lieb’s concavity theorem says that the function

\[
(A, B) \mapsto \text{Tr} \ X^* A^s X B^{1-s}
\]

(A.1)

is jointly concave on \( B(\mathcal{H})_+ \times B(\mathcal{H})_+ \) for any \( X \in B(\mathcal{H}) \) when \( 0 \leq s \leq 1 \). On the other hand, Ando proved in [1, §4] that the map \( (A, B) \mapsto A^s \otimes B^{1-s} \) is jointly convex on \( B(\mathcal{H})_{++} \times B(\mathcal{H})_{++} \) when \( 1 \leq s \leq 2 \). This can equivalently be stated that the function (A.1) is jointly convex on \( B(\mathcal{H})_{++} \times B(\mathcal{H})_{++} \) for any \( X \in B(\mathcal{H}) \) when \( 1 \leq s \leq 2 \), as easily verified from the identification of \( B(\mathcal{H}) \otimes B(\mathcal{H}) \) with \( B(\mathcal{H}) \), where \( \mathcal{H} \) is the Hilbert space \( B(\mathcal{H}), (\cdot, \cdot)_{\text{HS}} \) with the Hilbert-Schmidt inner product. We slightly extend this to the following:

**Lemma A.1.** Let \( X \in B(\mathcal{H}) \) be arbitrary and \( 1 \leq s \leq 2 \). Let \( Q \in B(\mathcal{H}) \) be an orthogonal projection. Then the function (A.1) is jointly convex on \( B(\mathcal{H})_+ \times \{B \in B(\mathcal{H})_+: \text{supp} \ B = Q\} \).
Proof. Since the case $s = 1$ is trivial, assume that $1 < s \leq 2$. For $A_1, B_1, A_2, B_2 \in \mathcal{B}(\mathcal{H})_+$ with $\text{supp } B_1 = \text{supp } B_2 = Q$, apply the above joint convexity to $(A_k + \varepsilon I, B_k + \varepsilon^{-1}(I - Q))$, $k = 1, 2$, and take the limit as $\varepsilon \searrow 0$. \hfill \square

### A.2 Mean Hoeffding distance

**Lemma A.2.** Let $\psi$ be a convex function on the interval $[0, 1]$ and $\varphi$ be its Legendre-Fenchel transform, i.e., $\varphi(a) := \sup_{0 \leq s \leq 1} \{as - \psi(s)\}$, $a \in \mathbb{R}$. Then for every $r \geq 0$,

$$\sup_{a : \varphi(a) - a > r} \varphi(a) = \sup_{0 \leq s \leq 1} \frac{-sr - \psi(s)}{1 - s}. \quad (A.2)$$

**Proof.** Define

$$\hat{\varphi}(a) := \varphi(a) - a = \sup_{0 \leq s \leq 1} \{a(s - 1) - \psi(s)\}, \quad a \in \mathbb{R}.$$ 

The following properties can easily be seen by definitions:

(i) $\varphi$ and $\hat{\varphi}$ are convex and continuous on $\mathbb{R}$,

(ii) $\varphi$ is increasing and $\hat{\varphi}$ is decreasing on $\mathbb{R}$ with $\varphi(a) \to +\infty$, $\hat{\varphi}(a) \to -\psi(0)$ as $a \to +\infty$ and $\varphi(a) \to -\psi(1)$, $\hat{\varphi}(a) \to +\infty$ as $a \to -\infty$. Moreover, $\hat{\varphi}$ is strictly decreasing on $(-\infty, \partial^{-}\psi(1))$. (See [17, Lemma 4.1].)

First, assume that $r < -\psi(1)$ or that $r = -\psi(1)$ and $\partial^{-}\psi(1) = +\infty$. Then $\hat{\varphi}(a) > r$ for every $a \in \mathbb{R}$, and the left-hand side of (A.2) is $+\infty$ since $\lim_{a \to +\infty} \varphi(a) = +\infty$. On the other hand, since

$$\frac{-sr - \psi(s)}{1 - s} = \frac{\psi(1) - \psi(s)}{1 - s} - \frac{s}{1 - s} (r + \psi(1)) - \psi(1),$$

we have

$$\lim_{s \searrow 1} \frac{-sr - \psi(s)}{1 - s} = +\infty$$

and hence the right-hand side of (A.2) is also equal to $+\infty$.

If $r = -\psi(1)$ and $\partial^{-}\psi(1) < +\infty$ or if $r > -\psi(1)$, then there exists an $a_r$ such that $\hat{\varphi}(a_r) = r$ and $\hat{\varphi}$ is strictly decreasing on $(-\infty, a_r]$. Thus, the left-hand side of (A.2) is $\varphi(a_r)$. Assume that $r = -\psi(1)$ and $\partial^{-}\psi(1) < +\infty$. Then $a_r = \partial^{-}\psi(1)$ and $\varphi(a_r) = \partial^{-}\psi(1) - \psi(1)$. On the other hand, by (A.3) and the convexity of $\psi$, we have

$$\sup_{0 \leq s \leq 1} \frac{-sr - \psi(s)}{1 - s} = \lim_{s \searrow 1} \frac{-sr - \psi(s)}{1 - s} = \partial^{-}\psi(1) - \psi(1).$$

Next, assume that $r > -\psi(1)$, and define $s_r := \arg\max_{0 \leq s \leq 1} \{a_r s - \psi(s)\}$. Then $s_r < 1$ and

$$r = a_r (s_r - 1) - \psi(s_r) \geq a_r (s - 1) - \psi(s), \quad 0 \leq s \leq 1,$$

so that $a_r \geq (-r - \psi(s))/(1 - s)$ for any $0 \leq s < 1$ with equality for $s = s_r$. Therefore,

$$\varphi(a_r) = a_r s_r - \psi(s_r) \geq a_r s - \psi(s) \geq \frac{-sr - \psi(s)}{1 - s}, \quad 0 \leq s < 1,$$

and equality holds for $s = s_r$. Hence we see that the right-hand side of (A.2) is also equal to $\varphi(a_r)$. \hfill \square
A.3 Conditional expectation with discrete Weyl operators

**Lemma A.3.** Let $m, d \in \mathbb{N}$ and $E$ be the conditional expectation (i.e., the partial trace) from $M_m \otimes M_d$ onto $M_m \otimes I_d$ with respect to the trace. Then there are unitaries $U_1, \ldots, U_{d^2}$ in $I_m \otimes M_d$ such that

$$E(A) = \frac{1}{d^2} \sum_{j=1}^{d^2} U_j AU_j^*$$

for all $A \in M_m \otimes M_d$.

**Proof.** Let $\mathbb{Z}_d$ denote the additive group of $\{0, \ldots, d-1\}$ with the modulo $d$ addition. Let $\{e_j, j \in \mathbb{Z}_d\}$ be a basis in $\mathbb{C}^d$ and define $Ue_j := e_{j+1}$ and $Ve_j := w^j e_j$ for $j \in \mathbb{Z}_d$, where $w := e^{i 2 \pi / d}$. The so defined operators satisfy the commutation relation $VU = wUV$. Let

$$W_k := \bar{w}^{k_1 k_2 / 2} V^{k_1} U^{k_2}, \quad k = (k_1, k_2) \in (\mathbb{Z}_d)^2,$$

which are the so-called discrete Weyl operators. It can easily be seen that

$$W_0 = I, \quad W_k^* = W_{-k}, \quad W_k W_l = w^{(k_1 l_2 - k_2 l_1) / 2} W_{k+l},$$

and $\text{Tr} W_k = \delta_{k,0}$ for all $k, l \in \mathbb{Z}_d^2$. Hence the Weyl operators are unitaries, and moreover $W := \{d^{-1/2} W_k : k \in (\mathbb{Z}_d)^2\}$ is an orthonormal base for $M_d$ with respect to the Hilbert-Schmidt inner product. As a consequence, the commutant of $W$ is $\mathbb{C} I$. One can see by a straightforward computation that for any $A \in M_d$,

$$\tilde{E}(A) := \frac{1}{d^2} \sum_{k \in \mathbb{Z}_d^2} W_k AW_k^*$$

is in the commutant of $W$ and hence it is a constant multiple of the identity. Since $\text{Tr} \tilde{E}(A) = \text{Tr} A$, we have $\tilde{E}(A) = d^{-1}(\text{Tr} A) I$. Setting $U_k := I_m \otimes W_k$ yields the assertion. \(\square\)

A.4 Limiting formulas

**Lemma A.4.** For any $s \in \mathbb{R}$ and any $a, b \geq 0$,

$$\lim_{n \to \infty} \left( \sum_{i=0}^{n} \binom{n}{i} s^i a^i b^{n-i} \right)^{1/n} = \begin{cases} (a^{1/s} + b^{1/s})^s & \text{if } s > 0, \\ \max\{a, b\} & \text{if } s \leq 0. \end{cases}$$

**Proof.** If $a = 0$ or $b = 0$, then the equality is trivial. If $a, b > 0$ then we may assume $b = 1$ by homogeneity. Since

$$\left( \max_{0 \leq i \leq n} \binom{n}{i} a^i \right)^{1/n} \leq \left( \sum_{i=0}^{n} \binom{n}{i} s^i \right)^{1/n} \leq (n + 1)^{1/n} \left( \max_{0 \leq i \leq n} \binom{n}{i} s^i \right)^{1/n},$$

what we have to prove is that

$$\lim_{n \to \infty} \left( \max_{0 \leq i \leq n} \binom{n}{i} a^i \right)^{1/n} = \begin{cases} (a^{1/s} + 1)^s & \text{if } s > 0, \\ \max\{a, 1\} & \text{if } s \leq 0. \end{cases}$$
that is,
\[
\lim_{n \to \infty} \max_{0 \leq i \leq n} \left( \frac{s}{n} \log \binom{n}{i} + \frac{i}{n} \log a \right) = \begin{cases} 
\log(a^{1/s} + 1) & \text{if } s > 0, \\
\max\{\log a, 0\} & \text{if } s \leq 0.
\end{cases}
\]

Since the Stirling formula gives
\[
\frac{s}{n} \log \binom{n}{i} + \frac{i}{n} \log a = s \left( \frac{i}{n} \log \frac{i}{n} - (1 - \frac{i}{n}) \log (1 - \frac{i}{n}) + o(1) \right) + \frac{i}{n} \log a,
\]
we may show that
\[
\max_{0 \leq x \leq 1} h_s(x) = \begin{cases} 
s \log(a^{1/s} + 1) & \text{if } s > 0, \\
\max\{\log a, 0\} & \text{if } s \leq 0,
\end{cases}
\]
for
\[
h_s(x) := s(-x \log x - (1-x) \log(1-x)) + x \log a, \quad 0 \leq x \leq 1.
\]

Notice that
\[
h'_s(x) = s \log \frac{1-x}{x} + \log a, \quad 0 < x < 1.
\]
When \(s > 0\), the maximizer \(x_0\) of \(h_s(x)\) satisfies
\[
s \log \frac{1-x_0}{x_0} = -\log a \quad \text{or} \quad \frac{1}{x_0} = 1 + \frac{1}{a^{1/s}},
\]
and the maximum is
\[
h_s(x_0) = sx_0 \log \frac{1-x_0}{x_0} - s \log(1-x_0) + x_0 \log a
\]
\[
= -s \log(1-x_0) = s \log \frac{1}{x_0} + \log a = s \log(a^{1/s} + 1).
\]
When \(s \leq 0\), \(h_s(x)\) takes the maximum at either \(x = 0\) or \(x = 1\), so that the maximum is \(\max\{\log a, 0\}\). \(\Box\)

**Lemma A.5.** For any \(a, b \geq 0\),
\[
\lim_{n \to \infty} \left( \sum_{i=0}^{[n/2]} \binom{n}{i} a^i b^{n-i} \right)^{1/n} = \begin{cases} 
a + b & \text{if } a \leq b, \\
2\sqrt{ab} & \text{if } a \geq b.
\end{cases}
\]

*Proof.* Since the case \(a = 0\) or \(b = 0\) is trivial, we assume that \(a, b > 0\). As in the proof of Lemma A.4, letting \(b = 1\) we may prove that
\[
\lim_{n \to \infty} \left( \max_{0 \leq i \leq [n/2]} \binom{n}{i} a^i \right)^{1/n} = \begin{cases} 
a + 1 & \text{if } a \leq 1, \\
2\sqrt{a} & \text{if } a \geq 1.
\end{cases}
\]

Hence it suffices to show that
\[
\max_{0 \leq x \leq 1/2} h(x) = \begin{cases} 
\log(a + 1) & \text{if } a \leq 1, \\
\log 2\sqrt{a} & \text{if } a \geq 1,
\end{cases}
\]
where \(h(x)\) denotes \(h_s(x)\) with \(s = 1\) in the proof of Lemma A.4. It is indeed seen since the maximum of \(h(x)\) on \([0, 1/2]\) is taken at \(x = a/(a+1)\) if \(a \leq 1\) and at \(x = 1/2\) if \(a \geq 1\). \(\Box\)
References

[1] T. Ando  
Concavity of certain maps and positive definite matrices and applications to Hadamard products; Linear Algebra Appl. 26, 203–241 (1979).

[2] T. Ando, X. Zhan:  
Norm inequalities related to operator monotone functions; Math. Ann. 315, 771–780 (1999).

[3] K.M.R. Audenaert, J. Calsamiglia, Ll. Masanes, R. Munoz-Tapia, A. Acin, E. Bagan, F. Verstraete:  
Discriminating states: the quantum Chernoff bound; Phys. Rev. Lett. 98, 160501 (2007).

[4] K.M.R. Audenaert, M. Nussbaum, A. Szkola, F. Verstraete:  
Asymptotic error rates in quantum hypothesis testing; Comm. Math. Phys. 279, 251–283 (2008).

[5] R. Bhatia, F. Kittaneh:  
Norm inequalities for positive operators; Lett. Math. Phys. 43, 225–231 (1998).

[6] I. Bjelaković, R. Siegmund-Schultze:  
An ergodic theorem for the quantum relative entropy; Comm. Math. Phys. 247, 697–712 (2004).

[7] I. Bjelakovic, J.-D. Deuschel, T. Krüger, R. Seiler, Ra. Siegmund-Schultze, A. Szkola:  
a quantum version of Sanov’s theorem; Comm. Math. Phys. 260, 659–571 (2005).

[8] I. Bjelakovic, J.-D. Deuschel, T. Krüger, R. Seiler, Ra. Siegmund-Schultze, A. Szkola:  
Typical support and Sanov large deviations of correlated states; Comm. Math. Phys. 279, 559–584 (2008).

[9] F.G.S.L. Brandao, M.B. Plenio:  
A Generalization of Quantum Stein’s lemma; arxiv:0904.0281.

[10] C.A. Fuchs, J. van de Graaf:  
Cryptographic distinguishability measures for quantum mechanical states; IEEE Trans. Inform. Theory 45, 1216–1227 (1999).

[11] F. Hansen:  
An operator inequality; Math. Ann. 246, 249–250 (1980).

[12] M. Hayashi:  
Optimal sequence of quantum measurements in the sense of Stein’s lemma in quantum hypothesis testing; J. Phys. A: Math. Gen. 35, 10759–10773 (2002).

[13] M. Hayashi:  
Quantum Information: An Introduction; Springer, Berlin-Heidelberg-New York, 2006.

[14] M. Hayashi:  
Error exponent in asymmetric quantum hypothesis testing and its application to classical-quantum channel coding; Phys. Rev. A 76, 062301 (2007).

[15] M. Hayashi:  
Group theoretical study of LOCC-detection of maximally entangled state using hypothesis testing; arXiv:0810.3380; to appear in New Journal of Physics.

[16] F. Hiai, M. Mosonyi, T. Ogawa:  
Large deviations and Chernoff bound for certain correlated states on the spin chain; J. Math. Phys. 48, 123301 (2007).

[17] F. Hiai, M. Mosonyi, T. Ogawa:  
Error exponents in hypothesis testing for correlated states on a spin chain; J. Math. Phys. 49, 032112 (2008).
[18] F. Hiai, D. Petz: The proper formula for relative entropy and its asymptotics in quantum probability; Comm. Math. Phys. 143, 99–114 (1991).

[19] F. Hiai, D. Petz: Entropy densities for algebraic states; J. Funct. Anal. 125, 287–308 (1994).

[20] E.H. Lieb: Convex trace functions and the Wigner-Yanase-Dyson conjecture; Adv. Math. 11, 267–288 (1973).

[21] M. Mosonyi, F. Hiai, T. Ogawa, M. Fannes: Asymptotic distinguishability measures for shift-invariant quasi-free states of fermionic lattice systems; J. Math. Phys. 49, 072104 (2008).

[22] M. Mosonyi: Hypothesis testing for Gaussian states on bosonic lattices; J. Math. Phys. 50, 032104, (2009).

[23] H. Nagaoka: The converse part of the theorem for quantum Hoeffding bound; preprint; quant-ph/0611289.

[24] H. Nagaoka, M. Hayashi: An information-spectrum approach to classical and quantum hypothesis testing for simple hypotheses; IEEE Trans. Inform. Theory 53, 534–549 (2007).

[25] M. A. Nielsen, I. L. Chuang: Quantum Computation and Quantum Information; Cambridge University Press, Cambridge, 2000.

[26] M. Nussbaum, A. Szkola: A lower bound of Chernoff type for symmetric quantum hypothesis testing; Ann. Statist. 37, 1040–1057, (2009).

[27] T. Ogawa, M. Hayashi: On error exponents in quantum hypothesis testing; IEEE Trans. Inform. Theory 50, 1368–1372 (2004).

[28] T. Ogawa, H. Nagaoka: Strong converse and Stein’s lemma in quantum hypothesis testing; IEEE Trans. Inform. Theory 47, 2428–2433 (2000).

[29] H. Ohno: Dynamical entropy of generalized quantum Markov chains on gauge-invariant C*-algebras; Lett. Math. Phys. 78, 111–124 (2006).

[30] D. Petz: Quantum Information Theory and Quantum Statistics, Springer, 2008.

[31] A. Rényi: Probability Theory, North-Holland, Amsterdam-London, 1970.