

Asymptotically Optimal Discrimination between Pure Quantum States

Michael Nussbaum^{1,*} and Arleta Szkola²

¹ Department of Mathematics, Cornell University, Ithaca NY, USA

² Max Planck Institute for Mathematics in the Sciences, Leipzig, Germany

Abstract. We consider the decision problem between a finite number of states of a finite quantum system, when an arbitrarily large number of copies of the system is available for measurements. We provide an upper bound on the exponential rate of decay of the averaged probability of rejecting the true state. It represents a generalized quantum Chernoff distance of a finite set of states. As our main result we prove that the bound is sharp in the case of pure states.

Keywords: multiple quantum state discrimination, generalized quantum Chernoff distance, quantum hypothesis testing, error exponents.

1 Introduction

In various branches of quantum theory such as quantum information processing, quantum communication theory or quantum statistics one of the basic problems is to determine the state of a given quantum system. In the simplest case there is a *finite* set of states specifying the possible preparation of the quantum system. In the Bayesian approach of quantum statistics, the likelihood of the different states is determined by an a priori probability distribution. One makes a decision in favor of one of the states following a specified rule based on the outcomes of a generalized measurement -called a quantum test. In the binary case optimal tests, i.e. tests minimizing the averaged probability of rejecting the true state, are known to be given by Holevo-Helstrom projections [5], [4]. These generalize the classical likelihood ratio tests. Here we consider the scenario where there is an arbitrarily large finite number n of copies of the quantum system available for performing a measurement. The corresponding state is then described by an n -fold tensor product of one of the associated density operators. There are two main goals: firstly, to construct a sequence of quantum tests in n which maximize the asymptotic (exponential) rate of decay of the averaged probability of rejecting the true state. The second goal is to determine the corresponding optimal error exponent. It has been shown that in the binary case asymptotically optimal quantum tests, thus in particular the Holevo-Helstrom tests, achieve an exponential rate of decay which is equal to the quantum Chernoff bound, cf. [8], [1] and [2]. Surprisingly, the corresponding questions in the case of $r > 2$ states

* Supported in part by NSF grant DMS-08-05632.

have not yet received a final answer, despite a number of efforts and numerous strong results obtained in relation to multiple quantum state discrimination, see [11], [7], [3], [10] and references therein.

We define a generalized quantum Chernoff distance of a finite set of states as the minimum of the binary quantum Chernoff distances over all possible pairs of different states. The binary quantum Chernoff distance has been introduced in the context of binary quantum hypothesis testing in [8]. Relying on [8] we prove that the generalized quantum Chernoff distance specifies a bound on the achievable asymptotic error exponents in multiple quantum state discrimination. This is in line with results obtained in the context of classical multiple hypothesis testing, cf. [9]. As our main result we prove that in the special case of pure quantum states this bound, indeed, is achievable and hence specifies the optimal asymptotic error exponent. The corresponding asymptotically optimal quantum tests rely on a Gram-Schmidt orthonormalization procedure of the associated state vectors. Similar quantum tests were already considered by Holevo in [6] in the context of quantum minimal error decision problems. However, the question of the corresponding asymptotic error exponent is not addressed in [6].

2 Notations and the Main Results

Let S be a finite quantum system and \mathcal{H} be the associated complex Hilbert space with $\dim \mathcal{H} = d < \infty$. Further denote by \mathcal{A} the algebra of observables of S , i.e. \mathcal{A} is the algebra of linear operators on \mathcal{H} . For each $n \in \mathbb{N}$ denote by $\mathcal{A}^{(n)}$ the algebra of linear operators on the n -fold tensor product Hilbert space $\mathcal{H}^{\otimes n}$. It represents the algebra of observables of a compound quantum system S_n with its n unit systems being of the same type S .

For each $n \in \mathbb{N}$ the set of density operators in $\mathcal{A}^{(n)}$ corresponds one-to-one to the state space $\mathcal{S}(\mathcal{A}^{(n)})$ of $\mathcal{A}^{(n)}$. Recall that a density operator is defined to be a self-adjoint, positive linear operator of trace 1.

Let $r \in \mathbb{N}$ and Σ be a set of density operators $\rho_i \in \mathcal{S}(\mathcal{A})$, $i = 1, \dots, r$, representing the possible states of the quantum system S . Assume that for each $n \in \mathbb{N}$ there is a compound quantum system S_n being an n -fold copy of S . This means, in particular, that the corresponding quantum state is in $\Sigma^{\otimes n} := \{\rho_i^{\otimes n}\}_{i=1}^r$, i.e. it is uniquely determined by the index $i \in \{1, \dots, r\}$.

Further, let $E^{(n)} = \{E_i^{(n)}\}_{i=1}^r$ be a positive operator valued measure (POVM) in $\mathcal{A}^{(n)}$, i.e. each $E_i^{(n)}$, $i = 1, \dots, r$, is a self-adjoint element of $\mathcal{A}^{(n)}$ with $E_i^{(n)} \geq 0$ and $\sum_{i=1}^r E_i^{(n)} = \mathbf{1}$. The POVMs $E^{(n)}$ describe quantum tests for discrimination between the r states from $\Sigma^{\otimes n}$, or simply *quantum tests for $\Sigma^{\otimes n}$* , by identifying the measurement outcome corresponding to $E_i^{(n)}$, $i = 1, \dots, r$, with the density operator $\rho_i^{\otimes n}$, respectively. If ρ_i happens to describe the true state of S , and correspondingly $\rho_i^{\otimes n}$ determines the state of S_n , then the associated *individual success probability* is given by

$$\text{Succ}_i(E^{(n)}) := \text{tr} [\rho_i^{\otimes n} E_i^{(n)}].$$

The *individual error probability* refers to the situation when the density operator ρ_i is discarded as possible preparation of S ; it is given by the formula

$$\text{Err}_i(E^{(n)}) := \text{tr} [\rho_i^{\otimes n} (\mathbf{1} - E_i^{(n)})] .$$

Assuming $0 < p_i < 1$, $i = 1, \dots, r$, with $\sum_{i=1}^r p_i = 1$ to be the a priori distribution of the r quantum states from Σ the *averaged error probability* is defined by

$$\text{Err}(E^{(n)}) = \sum_{i=1}^r p_i \text{tr} [\rho_i^{\otimes n} (\mathbf{1} - E_i^{(n)})] .$$

If the limit $\lim_{n \rightarrow \infty} -\frac{1}{n} \log \text{Err}(E^{(n)})$ exists, we refer to it as the *asymptotic error exponent*. Otherwise we have to consider the corresponding lim sup and lim inf expressions.

For two density operators ρ_1 and ρ_2 the *quantum Chernoff distance* is defined by

$$\xi_{QCB}(\rho_1, \rho_2) := -\log \inf_{0 \leq s \leq 1} \text{tr} [\rho_1^{1-s} \rho_2^s] . \quad (1)$$

It specifies the optimal achievable asymptotic error exponent in discriminating between ρ_1 and ρ_2 , compare [8], [1], [2]. Quantum tests with minimal averaged error probability for a pair of different density operators ρ_1 and ρ_2 on the same Hilbert space \mathcal{H} are well-known to be given by the respective *Holevo-Helstrom projectors*

$$\Pi_1 := \text{supp} (\rho_1 - \rho_2)_+, \quad \Pi_2 := \text{supp} (\rho_2 - \rho_1)_+ = \mathbf{1} - \Pi_1 .$$

Here $\text{supp } a$ denotes the support projector of a self-adjoint operator a , while a_+ means its positive part, i.e. $a_+ = (|a| + a)/2$ for $|a| := (a^*a)^{1/2}$, see [5], [4]. As mentioned in the introduction, the Holevo-Helstrom projectors generalize the likelihood ratio tests for two probability distributions. This can be verified by letting ρ_1 and ρ_2 be two commuting density matrices, cf. [8].

For a set $\Sigma = \{\rho_i\}_{i=1}^r$ of density operators on \mathcal{H} , where $r > 2$, we introduce the *generalized quantum Chernoff distance*

$$\xi_{QCB}(\Sigma) := \min\{\xi_{QCB}(\rho_i, \rho_j) : 1 \leq i < j \leq r\} . \quad (2)$$

This is in full analogy to the definition of the generalized Chernoff distance in classical multiple hypothesis testing, where the density operators are replaced by probability distributions on a finite sample space, cf. [9].

Our first theorem is an implication of Theorem 2.2 in [8].

Theorem 1. *Let $r \in \mathbb{N}$ and $\Sigma = \{\rho_i\}_{i=1}^r$ be a set of pairwise different density operators on \mathcal{H} with corresponding a priori probability distribution $\{p_i\}_{i=1}^r$. For any sequence $E^{(n)}$, $n \in \mathbb{N}$, of quantum tests for $\Sigma^{\otimes n}$, respectively, it holds*

$$\limsup_{n \rightarrow \infty} -\frac{1}{n} \log \text{Err}(E^{(n)}) \leq \xi_{QCB}(\Sigma) ,$$

where $\xi_{QCB}(\Sigma)$ is the generalized quantum Chernoff distance defined by (2).

It turns out that the generalized quantum Chernoff distance is achievable as an asymptotic error exponent in the case of pure states. This is the statement of our main theorem below.

Theorem 2. *Let $r \in \mathbb{N}$ and $\Sigma = \{\rho_i\}_{i=1}^r$ be a set of pairwise different pure states of a quantum system S . Then there exists a sequence $\{E^{(n)}\}_{n \in \mathbb{N}}$ of quantum tests for $\Sigma^{\otimes n}$, respectively, with*

$$\lim_{n \rightarrow \infty} -\frac{1}{n} \log \text{Err}(E^{(n)}) = \xi_{QCB}(\Sigma) ,$$

i.e. the generalized quantum Chernoff distance is an achievable asymptotic error exponent in multiple pure state discrimination.

3 Generalized Quantum Chernoff Bound in Multiple Quantum State Discrimination

In this section we give a proof of Theorem 1 stating that the generalized quantum Chernoff distance specifies a bound on the asymptotically achievable error exponents in multiple quantum state discrimination. It relies on its binary version presented in Theorem 2.2 in [8].

Proof (Theorem 1). Fix any two indices $1 \leq i < j \leq r$. For $n \in \mathbb{N}$ let $A^{(n)}, B^{(n)} \in \mathcal{A}^{(n)}$ be two positive operators such that $A^{(n)} + B^{(n)} = \mathbf{1} - E_i^{(n)} - E_j^{(n)}$. Then the positive operators $\tilde{E}_i^{(n)} := E_i^{(n)} + A^{(n)}$ and $\tilde{E}_j^{(n)} := E_j^{(n)} + B^{(n)}$ represent a POVM $\tilde{E}^{(n)}$ in $\mathcal{A}^{(n)}$, which we consider a quantum test for the pair $\{\rho_i^{\otimes n}, \rho_j^{\otimes n}\}$. For the individual error probabilities of the modified quantum test $\tilde{E}^{(n)}$ we obtain the upper bounds

$$\text{Err}_i(\tilde{E}^{(n)}) = \text{tr} [\rho_i^{\otimes n}(\mathbf{1} - \tilde{E}_i^{(n)})] \leq \text{tr} [\rho_i^{\otimes n}(\mathbf{1} - E_i^{(n)})] = \text{Err}_i(E^{(n)}) ,$$

and similarly $\text{Err}_j(\tilde{E}^{(n)}) \leq \text{Err}_j(E^{(n)})$. It follows a lower bound on the average error probability with respect to the original tests $\{E_i^{(n)}\}_{i=1}^r$:

$$\begin{aligned} \text{Err}(E^{(n)}) &= \sum_{k=1}^r p_k \text{Err}_k(E^{(n)}) \geq \left(p_i \text{Err}_i(E^{(n)}) + p_j \text{Err}_j(E^{(n)}) \right) \\ &\geq \left(p_i \text{Err}_i(\tilde{E}^{(n)}) + p_j \text{Err}_j(\tilde{E}^{(n)}) \right) \\ &\geq p_{\min} \left(\text{Err}_i(\tilde{E}^{(n)}) + \text{Err}_j(\tilde{E}^{(n)}) \right) , \end{aligned}$$

where $p_{\min} := \min\{p_i : 1 \leq i \leq r\}$. The above bound implies

$$\begin{aligned} \limsup_{n \rightarrow \infty} -\frac{1}{n} \log \text{Err}(E^{(n)}) &\leq \limsup_{n \rightarrow \infty} -\frac{1}{n} \log p_{\min} \\ &\quad + \limsup_{n \rightarrow \infty} -\frac{1}{n} \log \left(\text{Err}_i(\tilde{E}^{(n)}) + \text{Err}_j(\tilde{E}^{(n)}) \right) \\ &= \limsup_{n \rightarrow \infty} -\frac{1}{n} \log \frac{1}{2} \left(\text{Err}_i(\tilde{E}^{(n)}) + \text{Err}_j(\tilde{E}^{(n)}) \right) \\ &\leq \xi_{QCB}(\rho_i, \rho_j) . \end{aligned}$$

Here the last inequality is by Theorem 2.2 in [8], which represents the statement of our Theorem 1 in its binary version corresponding to the special case $r = 2$. Since the pair of indices (i, j) was chosen arbitrary, the statement of the theorem follows. \square

4 Asymptotically Optimal Pure State Discrimination

In this section we provide a constructive proof for Theorem 2. Roughly speaking, our quantum tests, which can be shown to achieve an asymptotic error exponent equal to the generalized quantum Chernoff distance of Σ , are obtained from a Gram-Schmidt orthonormalization procedure of the unit vectors associated to the pure states in Σ .

Proof (Theorem 2). Observe that in view of Theorem 1 it is sufficient to construct quantum tests for which we can verify

$$\liminf_{n \rightarrow \infty} -\frac{1}{n} \log \text{Err}(E^{(n)}) \geq \xi_{QCB}(\Sigma) .$$

For each $1 \leq i \leq r$ let v_i be a unit vector in \mathcal{H} such that $|v_i\rangle\langle v_i| = \rho_i$.

1. We assume that the set $V(\Sigma) := \{v_i\}_{i=1}^r$ is linearly independent and start with the case $n = 1$, where no tensor products are included. We define for each $k = 1, \dots, r$, a $(d \times k)$ matrix Ψ_k

$$\Psi_k := (v_1, \dots, v_k) , \tag{3}$$

i.e. the columns of Ψ_k are equal to the state vectors v_i , $1 \leq i \leq k$. We refer to the $(k \times k)$ -matrix

$$\Psi_k^* \Psi_k =: \Gamma_k$$

as a Gram matrix of $\{v_1, \dots, v_k\}$. By the assumption of linear independence of the set $V(\Sigma)$ for each $k \in \{1, \dots, r\}$ the operator

$$P_k := \Psi_k (\Psi_k^* \Psi_k)^{-1} \Psi_k^* = \Psi_k \Gamma_k^{-1} \Psi_k^* ,$$

represents an orthogonal projector onto a k -dimensional subspace of \mathcal{H} , which is spanned by the k state vectors v_1, \dots, v_k . Further, we set $P_0 = 0$ and define for $1 \leq k \leq r$

$$E_k := P_k - P_{k-1} .$$

The E_k represent one-dimensional orthogonal projectors, which are mutually orthogonal. With $e_k := \frac{1}{\|E_k v_k\|} E_k v_k$ we can write $E_k = |e_k\rangle\langle e_k|$, and the set $\{e_k\}_{k=1}^r$ represents a Gram-Schmidt orthonormalization of the linearly independent set $V(\Sigma)$ of unit vectors v_k , $k = 1, \dots, r$.

Observe that by construction $\sum_{i=1}^r E_i \leq \mathbf{1}$. If $E_0 := \mathbf{1} - \sum_{i=1}^r E_i \neq 0$, we redefine E_1 to be $E_1 + E_0$, such that $\sum_{i=1}^r E_i = \mathbf{1}$ is satisfied. By identifying E_i , $i = 1, \dots, r$, with ρ_i , respectively, we obtain a quantum test $E^{(1)} = \{E_i\}_{i=1}^r$ for Σ .

For $1 \leq i \leq r$ the corresponding individual success probability reads

$$\text{Succ}_i(E^{(1)}) = \text{tr} [\rho_i E_i] = \text{tr} [|v_i\rangle\langle v_i| E_i] = \langle v_i | P_i - P_{i-1} | v_i \rangle . \quad (4)$$

Since the P_i 's are constructed as orthogonal projectors onto $\text{span}\{v_1, \dots, v_i\}$ it holds $|v_i\rangle\langle v_i| \leq P_i$ and as a consequence $\langle v_i | P_i | v_i \rangle = 1$. Then from the relation $\text{Err}_i(E^{(1)}) = 1 - \text{Succ}_i(E^{(1)})$ we obtain

$$\begin{aligned} \text{Err}_i(E^{(1)}) &= \langle v_i | P_{i-1} | v_i \rangle \\ &= \langle v_i | \Psi_i(\Gamma_{i-1})^{-1} \Psi_{i-1}^* | v_i \rangle \\ &\leq \frac{1}{\lambda_{\min}(\Gamma_{i-1})} \langle v_i | \Psi_{i-1} \Psi_{i-1}^* | v_i \rangle \\ &= \frac{1}{\lambda_{\min}(\Gamma_{i-1})} \|\Psi_{i-1}^* v_i\|^2 , \end{aligned} \quad (5)$$

where $\lambda_{\min}(\cdot)$ denotes the minimal eigenvalue of a self-adjoint matrix. By definition (3) of Ψ_i we have

$$\|\Psi_{i-1}^* v_i\|^2 = \sum_{j=1}^{i-1} |\langle v_j | v_i \rangle|^2, \quad i = 2, \dots, r . \quad (6)$$

Inserting expression (6) into (5) we obtain the upper bound

$$\text{Err}_i(E^{(1)}) \leq \sum_{j=1}^{i-1} \frac{|\langle v_j | v_i \rangle|^2}{\lambda_{\min}(\Gamma_{i-1})} . \quad (7)$$

Recall that the density operators ρ_i , $i = 1, \dots, r$, are expected to appear with probability p_i , respectively. Then the averaged error probability can be estimated from above as follows

$$\begin{aligned} \text{Err}(E^{(1)}) &= \sum_{i=1}^r p_i \text{Err}_i(E^{(1)}) \leq \sum_{i=1}^r \text{Err}_i(E^{(1)}) \\ &\leq \sum_{i=2}^r \sum_{j=1}^{i-1} \frac{|\langle v_j | v_i \rangle|^2}{\lambda_{\min}(\Gamma_{i-1})} , \end{aligned} \quad (8)$$

where in the second line we have applied (7).

2. Let $n > 1$. Notice that still assuming that $V(\Sigma)$ is a set of r linearly independent unit vectors, the same remains true for $V(\Sigma^{\otimes n})$ consisting of the n -fold tensor product state vectors $v_i^{\otimes n}$, $i = 1, \dots, r$. Hence we can adopt the construction of the quantum test $E^{(1)}$ for Σ as it stands for the tensor product

case. In particular, we define $\Psi_{j,n}$, $1 \leq j \leq r$, analogously to (3) as the $(d^n \times j)$ -matrix

$$\Psi_{j,n} := (v_1^{\otimes n}, \dots, v_j^{\otimes n}) ,$$

respectively. Then the corresponding averaged error probability $\text{Err}(E^{(n)})$ can be upper bounded similarly to (8):

$$\text{Err}(E^{(n)}) \leq \sum_{i=2}^r \sum_{j=1}^{i-1} \frac{|\langle v_j^{\otimes n} | v_i^{\otimes n} \rangle|^2}{\lambda_{\min}(\Gamma_{i-1,n})} = \sum_{i=2}^r \sum_{j=1}^{i-1} \frac{(|\langle v_j | v_i \rangle|^2)^n}{\lambda_{\min}(\Gamma_{i-1,n})} ,$$

where $\Gamma_{i-1,n} := \Psi_{i-1,n}^* \Psi_{i-1,n}$.

Observe that each Gram matrix $\Gamma_{j,n} = \Psi_{j,n}^* \Psi_{j,n}$, $j = 1, \dots, r$, is a square matrix of fixed dimension j , respectively. Further, note that the diagonal entries $\gamma_{kk}^{(j,n)}$, $k = 1, \dots, j$ of $\Gamma_{j,n}$ are given by $\langle v_k^{\otimes n} | v_k^{\otimes n} \rangle$, respectively, and hence are all equal to 1. Since for $k \neq l$ it holds $|\langle v_k | v_l \rangle| < 1$, the off-diagonal entries $\gamma_{k,l}^{(j,n)} = \langle v_k^{\otimes n} | v_l^{\otimes n} \rangle = \langle v_k | v_l \rangle^n$ tend to 0 as n goes to infinity. It follows for every $1 \leq j \leq r$

$$\Gamma_{j,n} \rightarrow I_j \quad \text{as } n \rightarrow \infty ,$$

where I_j denotes the identity matrix of dimension j . By continuity of the minimal eigenvalue this implies

$$\lambda_{\min}(\Gamma_{j,n}) \rightarrow 1 \quad \text{as } n \rightarrow \infty .$$

We conclude

$$\text{Err}(E^{(n)}) \leq \sum_{i=2}^r \sum_{j=1}^{i-1} (|\langle v_j | v_i \rangle|^2)^n (1 + o(1)) .$$

As n tends to infinity the largest term dominates. As a consequence we have

$$\begin{aligned} \frac{1}{n} \log \text{Err}(E^{(n)}) &\leq \max\{\log |\langle v_j | v_i \rangle|^2 : 1 \leq j < i \leq r\} + o(1) \\ &= -\min\{\xi_{QCB}(\rho_i, \rho_j), 1 \leq j < i \leq r\} + o(1) \\ &= -\xi_{QCB}(\Sigma) + o(1) , \end{aligned} \tag{9}$$

where in the second line we have used the fact that in the case of two different pure states on \mathcal{H} , say $\rho = |v\rangle\langle v|$ and $\sigma = |w\rangle\langle w|$, the corresponding (binary) quantum Chernoff distance $\xi_{QCB}(\rho, \sigma)$ takes the simple form $-\log |\langle v | w \rangle|^2$, cf. [8]. The last identity is by definition (2) of the generalized quantum Chernoff distance. The proof is complete under the assumption of linear independence of the set of eigenvectors of Σ .

3. Finally, notice that even if $V(\Sigma)$ is not linearly independent, the set $V(\Sigma^{\otimes N})$ consisting of N -fold tensor product vectors becomes linearly independent for N large enough. Then, for every $n \geq N$ we can adopt the construction of quantum tests $E^{(n)}$ for $\Sigma^{\otimes n}$ as presented in parts 1 and 2 of the proof, and the asymptotic relation (9) remains valid. \square

Acknowledgments. The work of M. N. has been supported in part by NSF Grant DMS-08-05632. A. S. wishes to thank the research groups of Prof. Jost and Nihat Ay at the MPI MiS for their interest and helpful discussions.

References

1. Audenaert, K.M.R., Casamiglia, J., Muñoz-Tapia, R., Bagan, E., Masanes, L.I., Acín, A., Verstraete, F.: Discriminating States: The Quantum Chernoff Bound. *Phys. Rev. Lett.* 98, 160501 (2007)
2. Audenaert, K.M.R., Nussbaum, M., Szkola, A., Verstraete, F.: Asymptotic Error Rates in Quantum Hypothesis Testing. *Commun. Math. Phys.* 279, 251–283 (2008)
3. Barrett, S., Croke, S.: On the conditions for discrimination between quantum states with minimum error. *J. Phys. A: Math. Theor.* 42 (2009)
4. Helstrom, C.W.: *Quantum Detection and Estimation Theory*. Academic Press, New York (1976)
5. Holevo, A.: Investigations in the general theory of statistical decisions. *Trudy Mat. Inst. Steklov* 124 (in Russian) (English translation in *Proc. Steklov Inst. of Math.* 3. Amer. Math. Soc., Providence) (1978)
6. Kholevo, A.: On asymptotically optimal hypothesis testing in quantum statistics. *Theor. Probab. Appl.* 23, 411–415 (1978)
7. König, R., Renner, R., Schaffner, C.: The operational meaning of min- and max-entropy. *IEEE Trans. Inf. Th.* 55(9) (2009)
8. Nussbaum, M., Szkola, A.: The Chernoff lower bound for symmetric quantum hypothesis testing. *Ann. Stat.* 37(2), 1040–1057 (2009)
9. Salikhov, N.P.: On one generalisation of Chernov’s distance. *Theory Probab. Appl.* 43(2), 239–255 (1997)
10. Tyson, J.: Two-sided estimates of minimum-error distinguishability of mixed quantum states via generalized Holevo-Curlander bounds. *J. Math. Phys.* 50, 32106 (2009)
11. Yuen, H.P., Kennedy, R.S., Lax, M.: Optimum testing of Multiple Hypotheses in Quantum Detection Theory. *IEEE Trans. Inform. Theory* IT-21(2), 125–134 (1975)