Open access • Journal Article • DOI:10.1103/PHYSREVLETT.98.160501

Discriminating States: the quantum Chernoff bound. — Source link

Kmr Audenaert, John Calsamiglia, Ramon Muñoz-Tapia, Emilio Bagan ...+3 more authors

Institutions: Imperial College London, Autonomous University of Barcelona, University of Cambridge, University of Vienna

Published on: 17 Apr 2007 - Physical Review Letters (American Physical Society)

Topics: Chernoff bound, Binary symmetric channel, Quantum capacity, Trace distance and Quantum algorithm

Discriminating States: The Quantum Chernoff Bound

K. M. R. Audenaert
Institute for Mathematical Sciences, Imperial College London, 53 Prince's Gate, London SW7 2PG, United Kingdom

J. Calsamiglia, R. Muñoz-Tapia, and E. Bagan
Grup de Física Teòrica, Universitat Autònoma de Barcelona, 08193 Bellaterra (Barcelona), Spain

Ll. Masanes
DAMTP, University of Cambridge, Wilberforce Road, Cambridge CB3 0WA, United Kingdom

A. Acin
ICFO-Institut de Ciencies Fotoniques, Mediterranean Technology Park, 08860 Castelldefels (Barcelona), Spain

F. Verstraete
Fakultät für Physik, Universität Wien, Boltzmanngasse 5, 1090 Wien, Austria
(Received 18 October 2006; published 17 April 2007)

We consider the problem of discriminating two different quantum states in the setting of asymptotically many copies, and determine the minimal probability of error. This leads to the identification of the quantum Chernoff bound, thereby solving a long-standing open problem. The bound reduces to the classical Chernoff bound when the quantum states under consideration commute. The quantum Chernoff bound is the natural symmetric distance measure between quantum states because of its clear operational meaning and because it does not seem to share some of the undesirable features of other distance measures.

DOI: 10.1103/PhysRevLett.98.160501 PACS numbers: 03.67.−a

One of the most basic tasks in information theory is the discrimination of two different probability distributions: given a source that outputs variables following one out of two possible probability distributions, determine which one it is with the minimal possible error. In a seminal paper, Chernoff [1] solved this problem in the asymptotic regime and showed that the probability of error P_e in discriminating two probability distributions decreases exponentially in the number of tests n that one can perform: $P_e \sim \exp(-n\xi_{\mathrm{CB}})$. The optimal exponent ξ_{CB} arising in the asymptotic limit is called the Chernoff bound [2]. One of the virtues of the Chernoff bound is that it yields a very natural distance measure between probability distributions; it is essentially the unique distance measure in the ubiquitous situation of independent and identically distributed (i.i.d.) random variables.

A quantum generalization of this result is highly desired. Indeed, the concept of randomness is much more elementary in the field of quantum mechanics than in classical physics. Given the large amount of experimental effort in the context of quantum information processing to prepare and measure quantum states, it is of fundamental importance to have a theory that allows one to discriminate different quantum states. Despite considerable effort, this quantum generalization of the Chernoff bound has until now remained unsolved. The problem is to discriminate two sources that output many identical copies of one out of two different quantum states ρ and σ, and the question is to identify the exponent arising asymptotically when performing the optimal test to discriminate them. This task is so fundamental that it was probably the first problem ever considered in the field of quantum information theory; it was solved in the one-copy case more than 30 years ago [3,4]. In this Letter, we finally identify the asymptotic error exponent when the optimal strategy for discriminating the states is used. A nice feature of such a result is its universality, as it identifies the unique metric quantifying the distance of quantum states in the i.i.d. setting. Note that the related question of the optimal error exponents in asymmetric hypothesis testing in the sense of Stein's lemma was already solved a long time ago [5], leading to the operational meaning of quantum relative entropy. The quantum Chernoff bound can therefore be understood as the symmetric version of the quantum relative entropy.

Distance measures between quantum states have been used in a wide variety of applications in quantum information theory. The most popular such measure seems to be Uhlmann's fidelity [6], which happens to coincide with the quantum Chernoff bound when one of the states is pure. The trace distance has a more natural operational meaning, but lacks monotonicity under taking tensor powers of its arguments. The problem is that one can easily find states ρ, σ, ρ', σ' such that $\mathrm{Tr}|\rho - \sigma| < \mathrm{Tr}|\rho' - \sigma'|$ but $\mathrm{Tr}|\rho^{\otimes 2} - \sigma^{\otimes 2}| > \mathrm{Tr}|\rho'^{\otimes 2} - \sigma'^{\otimes 2}|$. The quantum Chernoff bound exactly characterizes the exponent arising in the asymptotic behavior of the trace distance in the case of many identical copies, and therefore does not suffer from this problem. Note that a similar situation happens

in the case of one-copy entanglement versus the asymptotic entanglement entropy.

In this Letter, we give an upper bound for the probability of error for discriminating two arbitrary states. In the particular case of a large number of identical copies, this result nicely complements the recent work of Nussbaum and Szkoła [7], where a lower bound for the asymptotic error exponent was found and hence a lower bound for the probability of error. These respective upper and lower bounds coincide in the asymptotic limit and hence give the exact expression for the error exponent. The conjecture of Ogawa and Hayashi concerning the quantum Chernoff bound raised in [8] is thus solved.

Our Letter is organized as follows. After the mathematical formulation of the problem, we prove a nontrivial and fundamental inequality relating the trace distance to the quantum Chernoff bound. Finally, we discuss some interesting properties of the quantum Chernoff bound.

The optimal error probability of discriminating two quantum states ρ_0 and ρ_1 has been identified a long time ago by Helstrom [3]. We consider the two hypotheses H_0 and H_1 that a given quantum system is prepared either in the state ρ_0 or in the state ρ_1, respectively. Since the (quantum) Chernoff bound arises in a Bayesian setting, we supply the prior probabilities π_0 and π_1, which are positive quantities summing up to 1 (the degenerate cases $\pi_0 = 0$ or $\pi_1 = 0$ are excluded).

Physically discriminating between these hypotheses corresponds to performing a generalized (POVM) measurement on the quantum system with two outcomes, 0 and 1. This POVM consists of the two elements $\{E_0, E_1\}$, where $E_0 + E_1 = \mathbb{1}$, $E_i \geq 0$. The symmetric distinguishability problem consists in finding those E_0 and E_1 that minimize the total error probability P_e, which is given by $P_e = \pi_0 \mathrm{Tr}[E_1\rho_0] + \pi_1 \mathrm{Tr}[E_0\rho_1] = \pi_1 - \mathrm{Tr}[E_1(\pi_1\rho_1 - \pi_0\rho_0)]$. This problem can be solved using some basic linear algebra. Let us first introduce some basic notations. Abusing terminology, we will use the term "positive" for "positive semidefinite" (denoted $A \geq 0$). We employ the positive semidefinite ordering throughout, $A \geq B$ iff $A - B \geq 0$. The absolute value $|A|$ is defined as $|A| := (A^*A)^{1/2}$. The Jordan decomposition of a self-adjoint operator A is given by $A = A_+ - A_-$, where A_+ and A_- are the positive and negative part of A, respectively, and are defined by $A_+ := (|A| + A)/2$ and $A_- := (|A| - A)/2$. Both parts are positive by definition, and $A_+A_- = 0$. The error probability P_e has to be minimized over all operators E_1 that satisfy $0 \leq E_1 \leq \mathbb{1}$. The result is that E_1 has to be the projector on the range of the positive part of $(\pi_1\rho_1 - \pi_0\rho_0)$, leading to

$$P_{e,\min} = \frac{1}{2}(1 - \|\pi_1\rho_1 - \pi_0\rho_0\|_1),$$

where $\|A\|_1 = \mathrm{Tr}|A|$ is the trace norm.

The basic problem to be solved now is to identify how the error probability P_e behaves in the asymptotic limit, i.e., when one has to discriminate between the hypotheses H_0 and H_1 corresponding to either n copies of ρ_0 having been produced or n copies of ρ_1. To do so, we need to study the quantity $P_{e,\min,n} := (1 - \|\pi_1\rho_1^{\otimes n} - \pi_0\rho_0^{\otimes n}\|_1)/2$, and it turns out that it exponentially decreases in n:

$$P_{e,\min,n} \sim \exp(-n\xi_{\mathrm{QCB}}).$$

We will prove that the exponent ξ_{QCB} is given by the following quantity, which can therefore be called the *quantum Chernoff bound*:

$$\xi_{\mathrm{QCB}} = \lim_{n\to\infty} -\frac{\log P_{e,\min,n}}{n} \tag{1}$$

$$= -\log \min_{0\leq s\leq 1} \mathrm{Tr}(\rho^s\sigma^{1-s}). \tag{2}$$

Note that the quantity $\mathrm{Tr}(\rho^s\sigma^{1-s})$ is well-defined and guaranteed to be positive. As should be, this expression for the quantum Chernoff bound reduces to the usual definition of the classical Chernoff bound ξ_{CB} when ρ and σ commute: for classical distributions p_0 and p_1,

$$\xi_{\mathrm{CB}} = -\log \min_{0\leq s\leq 1} \sum_i p_0(i)^s p_1(i)^{1-s}. \tag{3}$$

The fact that ξ_{QCB} is lower bounded by the expression on the right hand side of (2) was proven very recently in [7]. The fact that this is also an upper bound can be inferred from the following theorem:

Theorem 1.—Let A and B be positive operators, then for all $0 \leq s \leq 1$,

$$\mathrm{Tr}\,[A^sB^{1-s}] \geq \mathrm{Tr}[A + B - |A - B|]/2. \tag{4}$$

Indeed, let $A = \pi_1\rho_1^{\otimes n}$ and $B = \pi_0\rho_0^{\otimes n}$, then the upper bound trivially follows from the fact that the logarithm of the left hand side of the inequality (4) becomes $\log(\pi_0^s\pi_1^{1-s}) + n\log(\mathrm{Tr}[\rho_0^s\rho_1^{1-s}])$. Upon dividing by n and taking the limit $n \to \infty$, we obtain the quantum Chernoff bound ξ_{QCB}, independently of the priors π_0, π_1 (as long as the priors are not degenerate).

Note that it was already known that $P_{e,\min,n}$ is upper bounded by $\exp(-n\log\mathrm{Tr}(\rho^{1/2}\sigma^{1/2}))$ ([13], Lemma 3.2).

Inequality (4) is also very interesting from a purely matrix analytic point of view, as it relates the trace norm to a multiplicative quantity that is highly nontrivial and very useful. Note that the optimal measurement to discriminate the two sources enforces the use of joint measurements. As pointed out by A. Harrow, the particular permutational symmetry of N-copy states guarantees that the optimal collective measurement can be implemented efficiently (with a polynomial-size circuit) [10], and hence that the minimum probability of error is achievable with reasonable resources.

Let us now move on to prove Theorem 1. The proof that we present here goes through in infinite dimensions.

The proof relies on the following Lemma:

Lemma 1.—Let $A, B \geq 0$. Let $0 \leq t \leq 1$, and let P be the projector on the range of $(A - B)_+$. Then

$$\mathrm{Tr}\,[PB(A^t - B^t)] \geq 0. \tag{5}$$

Proof.—We exploit the integral representation [12]

$$a^t = \frac{\sin(t\pi)}{\pi} \int_0^{+\infty} dx \frac{ax^{t-1}}{a+x}, \qquad a \geq 0, \qquad 0 \leq t \leq 1, \tag{6}$$

which can be extended to positive operators in the usual way. Under the integral sign, $A^t - B^t$ contributes a factor $A(A+x)^{-1} - B(B+x)^{-1} = x\{(B+x)^{-1} - (A+x)^{-1}\}$. We next use the obvious relations

$$\frac{1}{b} - \frac{1}{a} = \int_0^1 dy \frac{d}{dy} \frac{-1}{b + (a-b)y}, \qquad \frac{d}{dy}\frac{-1}{c} = \frac{1}{c}\frac{dc}{dy}\frac{1}{c},$$

which hold for arbitrary invertible operators a, b, and c. By introducing the notation $\Delta = y(A-B)$ and $V = (B + \Delta + x)^{-1}$, we can write $A^t - B^t = \pi^{-1} \sin t\pi \int_0^\infty dx x^t \times \int_0^1 dy y^{-1} V \Delta V$. Hence, to prove the lemma, we just need to show that $\mathrm{Tr} PBV\Delta V = \mathrm{Tr} P(V^{-1} - \Delta - x)V\Delta V \geq 0$ for $0 \leq x$ and $0 \leq y \leq 1$.

Let Δ have the Jordan decomposition $\Delta = \Delta_+ - \Delta_-$. Thus P is the projector on the range of Δ_+. We choose the basis in which Δ is diagonal, and hence Δ, P, and V can be partitioned as

$$\Delta = \begin{pmatrix} \Delta_+ & 0 \\ 0 & -\Delta_- \end{pmatrix}, \quad P = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad V = \begin{pmatrix} V_{++} & V_{+-} \\ V_{-+} & V_{--} \end{pmatrix}.$$

For ease of notation, the subscript ij will henceforth refer to the (ij)-th block of an operator valued expression. We can rewrite $\mathrm{Tr} PBV\Delta V$ as

$$\mathrm{Tr}\, PBV\Delta V = \mathrm{Tr}[\Delta_+ V_{++} - (\Delta_+ + x)(V\Delta V)_{++}].$$

By the fact that V_{+-} and V_{-+} are each other's adjoint, the latter expression is positive, which finally proves the statement of the Lemma. □

Proof of theorem.—We apply Lemma 1 to the case $t = s/(1-s)$, $A = a^{1-s}$, and $B = b^{1-s}$, where a, b are positive operators and $0 \leq s \leq 1/2$. With P the projector on the range of $(a^{1-s} - b^{1-s})_+$, this yields

$$\mathrm{Tr}\,[Pb^{1-s}(a^s - b^s)] \geq 0.$$

Subtracting both sides from $\mathrm{Tr}[P(a-b)]$ then yields

$$\mathrm{Tr}\,[a^s P(a^{1-s} - b^{1-s})] \leq \mathrm{Tr}[P(a-b)].$$

Since P is the projector on the range of the positive part of $(a^{1-s} - b^{1-s})$, the LHS can be rewritten as $\mathrm{Tr}[a^s(a^{1-s} - b^{1-s})_+]$. Because $a^s \geq 0$, this is lower bounded by $\mathrm{Tr}[a^s(a^{1-s} - b^{1-s})] = \mathrm{Tr}[a - a^s b^{1-s}]$.

On the other hand, the RHS is upper bounded by $\mathrm{Tr}[(a-b)_+]$; this is because for any self-adjoint H, $\mathrm{Tr}[H_+]$ is the maximum of $\mathrm{Tr}[QH]$ over all self-adjoint projectors Q. We thus have

$$\mathrm{Tr}\,[a - a^s b^{1-s}] \leq \mathrm{Tr}[(a-b)_+] = \mathrm{Tr}[(a-b) + |a-b|]/2.$$

Subtracting both sides from $\mathrm{Tr}[a]$ finally yields (4) for $0 \leq s \leq 1/2$. The remaining case $1/2 \leq s \leq 1$ obviously follows by interchanging the roles of a and b. □

In the remainder of this Letter, we discuss the main properties of the nonlogarithmic variety of the quantum Chernoff bound, which we denote here by $Q(\rho, \sigma) := \min_{0 \leq s \leq 1} \mathrm{Tr}[\rho^s \sigma^{1-s}]$.

The following upper and lower bounds on Q in terms of the trace norm distance $T(\rho, \sigma) := ||\rho - \sigma||_1/2$ exist [9]:

$$1 - Q \leq T \leq \sqrt{1 - Q^2}. \tag{9}$$

Based on these bounds, the following properties of the Q-quantity and the Chernoff bound can be derived:

Inverted measure.—The maximum value Q can attain is 1, and this is reached when $\rho = \sigma$. This follows, for example, from the upper bound $Q^2 + T^2 \leq 1$. The minimal value is 0, and this is only attained for pairs of orthogonal states, i.e., states such that $\rho\sigma = 0$. This implies that the Chernoff bound is infinite if the states are orthogonal; this has to be contrasted with the asymmetric error exponents occurring in the context of relative entropy, where infinite values are obtained whenever the states have a different support.

Convexity in s.—The function to be minimized in Q is $s \mapsto \mathrm{Tr}[\rho^s \sigma^{1-s}]$. It is important to realize that this function is convex in $s \in [0, 1]$ because that means that the minimization has only one local minimum, and therefore this local minimum is automatically the global minimum. This is an important benefit in actual calculations.

Indeed, the function $s \mapsto x^s y^{1-s}$ is convex for positive scalars x and y, as one easily confirms by calculating the second derivative $x^s y^{1-s}(\log x - \log y)^2$, which is nonnegative. Consider then a basis in which ρ is diagonal and given by $\rho = \mathrm{Diag}(\lambda_1, \lambda_2, \ldots)$. Let the eigenvalue decomposition of σ (in that basis) be given by $\sigma = U\mathrm{Diag}(\mu_1, \mu_2, \ldots)U^*$, where U is a unitary. Then $\mathrm{Tr}[\rho^s \sigma^{1-s}] = \sum_{i,j} \lambda_i^s \mu_j^{1-s} |U_{ij}|^2$. As this is a sum with positive weights of convex terms $\lambda_i^s \mu_j^{1-s}$, the sum itself is also convex.

Joint concavity in (ρ, σ).—By Lieb's theorem [11], $\mathrm{Tr}[\rho^s \sigma^{1-s}]$ is jointly concave in (ρ, σ). Since the quantum Chernoff bound is the point-wise minimum of $\mathrm{Tr}[\rho^s \sigma^{1-s}]$ (over a fixed set, namely, over $s \in [0, 1]$), it is itself jointly concave as well. The Chernoff bound is therefore jointly convex, just like the relative entropy.

Monotonicity under CPT maps.—From the joint concavity, one easily derives the following monotonicity property: for any completely positive trace preserving (CPT) map Φ,

$$Q(\Phi(\rho), \Phi(\sigma)) \geq Q(\rho, \sigma). \tag{10}$$

For a proof, see [13].

Continuity.—By the lower bound $Q + T \geq 1$, $1 - Q$ is continuous in the sense that states that are close in trace norm distance are also close in $1 - Q$ distance: $0 \leq 1 - Q \leq T$.

Relation to fidelity.—If one of the states is pure, then Q equals the Uhlmann fidelity. Indeed, assume that $\rho_1 = |\psi\rangle\langle\psi|$ is pure, then the minimum of the expression $\mathrm{Tr}(\rho_1^s \rho_2^{1-s})$ is obtained for $s = 0$ and reduces to $\langle\psi|\rho_2|\psi\rangle$. From inequality (9), the fidelity is always an upper bound to Q.

Relation to the relative entropy.—Just as in the classical case, there is a nice connection between the quantum relative entropy and the Chernoff bound. By differentiating the expression $\mathrm{Tr}(\rho^s \sigma^{1-s})$ with relation to s, one observes that the minimum (which is unique due to convexity) is obtained when

$$\mathrm{Tr}\,(\rho^s \sigma^{1-s} \log\rho) = \mathrm{Tr}(\rho^s \sigma^{1-s} \log\sigma).$$

One easily verifies that this is equivalent to the condition that

$$S(\tau_s \| \rho) = S(\tau_s \| \sigma)$$

with $S(A\|B)$ the quantum relative entropy $\mathrm{Tr}(A \log A - A \log B)$ and τ_s defined as

$$\tau_s = \frac{\rho^s \sigma^{1-s}}{\mathrm{Tr}\rho^s \sigma^{1-s}}. \tag{11}$$

Note that τ_s is not a state because it is not even self-adjoint (except in the commuting case). Nevertheless, as it is basically the product of two positive operators, it has positive spectrum, and its entropy and the relative entropies used in (11) are well-defined. The value of s for which both relative entropies coincide is the optimal value s^*. This τ_{s^*} can be considered the quantum generalization of the Hellinger arc and interpolates between two different quantum states, albeit in a rather special (unphysical) way.

Metric.—The quantum Chernoff bound (or its nonlogarithmic variety) between two infinitesimally close states ρ and $\rho - d\rho$ induces a monotone metric that gives a geometrical structure to the state space [14]. The metric is given by [9]

$$ds^2 = 1 - \min_{0\le s\le 1} \mathrm{Tr}[\rho^s (\rho - d\rho)^{1-s}] = \frac{1}{2} \sum_{ij} \frac{|\langle i|d\rho|j\rangle|^2}{(\sqrt{\lambda_i} + \sqrt{\lambda_j})^2}$$

where $\rho = \sum_i \lambda_i |i\rangle\langle i|$ is the eigenvalue decomposition of ρ.

In conclusion, we have identified the exact expression of the quantum generalization of the Chernoff bound, which allows us to quantify the asymptotic behavior of the error in the context of Bayesian discrimination of different sources of quantum states. This resolves a long-standing open question. Our main theorem (Theorem 1), which gives a computable lower bound to the trace norm difference of two states in the many-copy regime, may also find other relevant applications in and outside the field of state discrimination [15].

F. V. and K. A. thank the hospitality of the Max Planck Institute for Quantum Optics where part of this work was done. K. A. was supported by The Leverhulme Trust (Grant No. F/07 058/U), by the QIP-IRC (www.qipirc.org) supported by EPSRC (No. GR/S82176/0), by EU Integrated Project QAP, and by the Institute of Mathematical Sciences, Imperial College London. We acknowledge financial support from CIRIT, Project No. SGR-00185, and from the Spanish MEC under Ramón y Cajal program (A. A. and J. C.); Projects Nos. FIS2004-05639-C02-02 and FIS2005-01369; and Consolider-Ingenio 2010, Project "QOIT." We are grateful to Montserrat Casas, Juli Céspedes, Alex Monràs, Sandu Popescu, and Andreas Winter for discussions, and to A. Harrow for pointing out that the optimal measurement can be implemented efficiently.

[1] H. Chernoff, Ann. Math. Stat. **23**, 493 (1952).
[2] The quantity which in this paper is referred to as the "Chernoff bound" also goes under the alternative names of Chernoff distance, Chernoff divergence, and Chernoff information. While the term "Chernoff bound" is also used for a bound on tail probabilities for statistical distributions (of which, actually, the Chernoff information is a derived quantity), we decided, against common sense, to follow common terminology.
[3] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press, New York, 1976).
[4] A. S. Holevo, Theory Probab. Appl. **23**, 411 (1979).
[5] F. Hiai and D. Petz, Commun. Math. Phys. **143**, 99 (1991); T. Ogawa and H. Nagaoka, IEEE Trans. Inf. Theory **46**, 2428 (2000); T. Ogawa and H. Nagaoka, IEEE Trans. Inf. Theory **46**, 2428 (2000).
[6] A. Uhlmann, Rep. Math. Phys. **9**, 273 (1976).
[7] M. Nussbaum and A. Szkoła, quant-ph/0607216.
[8] T. Ogawa and M. Hayashi, IEEE Trans. Inf. Theory **50**, 1368 (2004).
[9] K. Audenaert *et al.*, quant-ph/0610027.
[10] D. Bacon, I. Chuang, and A. Harrow, Phys. Rev. Lett. **97**, 170502 (2006).
[11] E. H. Lieb, Advances in Mathematics **11**, 267 (1973).
[12] R. Bhatia, *Matrix Analysis* (Springer, Berlin, 1997).
[13] M. Hayashi, *Quantum Information: An Introduction* (Springer Verlag, Berlin, 2006).
[14] D. Petz, Linear Algebra Appl. **244**, 81 (1996).
[15] And, indeed, after the appearance of the first draft of the manuscript, our main inequality has been used by Hayashi to prove achievability of the quantum Hoeffding bound; see M. Hayashi, quant-ph/0611013.