A lower bound on the probability of error in quantum state discrimination

Ashley Montanaro
*Department of Computer Science, University of Bristol, Woodland Road, Bristol, BS8 1UB, U.K.**
(Dated: May 19, 2008)

We give a lower bound on the probability of error in quantum state discrimination. The bound is a weighted sum of the pairwise fidelities of the states to be distinguished.

I. INTRODUCTION

The fact that non-orthogonal states are not perfectly distinguishable is a characteristic feature of quantum mechanics and the basis of the field of quantum cryptography. In this short note, we derive a quantitative lower bound on the indistinguishability of a set of quantum states.

The scenario we consider is that of quantum state discrimination: we are given a quantum system that was previously prepared in one of a known set of states, with known a priori probabilities, and must determine which state we were given with the minimum average probability of error. This fundamental problem was first studied by Helstrom [10] and Holevo [11] in the 1970s, and has since developed a vast literature (see [5] for a survey).

One can use efficient numerical techniques to determine this minimum average probability of error [7], but a general closed-form expression appears elusive. We are therefore led to putting bounds on this probability. Such bounds have been useful in the study of quantum query complexity [6] and in the security evaluation of quantum cryptographic schemes [9]. However, prior to this work no lower bound based on the most natural "local" measure of distinguishability of the quantum states in question – their pairwise fidelities – was known.

The most general strategy for quantum state discrimination is given by a positive operator valued measure (POVM) [13], namely a set of positive operators $M = \{\mu_i\}$ such that $\sum_i \mu_i = I$. The probability of receiving result i from measurement M on input of state ρ is $\mathrm{tr}(\mu_i \rho)$. We define an ensemble $\mathcal{E}$ as a set of quantum states $\{\rho_i\}$, each with a priori probability p_i, and associate measurement outcome i with the inference that we received state ρ_i. The average probability of error is then given by

$$P_E(M, \mathcal{E}) = \sum_{i \neq j} p_j \,\mathrm{tr}(\mu_i \rho_j).$$

We mention some matrix-theoretic notation that we will require; for more details, see [2]. For any matrix M and real $p > 0$, we define $\|M\|_p = (\sum_i \sigma_i(M)^p)^{1/p}$, where $\{\sigma_i(M)\}$ is the set of singular values of M. For $p \geq 1$ this is a matrix norm (known as the Schatten p-norm) and the case $p = 1$ is known as the trace norm. As it only depends on the singular values of M, $\|M\|_p$ is invariant under pre- and post-multiplication by unitaries.

The fidelity (Bures-Uhlmann transition probability) between two mixed quantum states ρ, σ can be defined in terms of the trace norm as $F(\rho, \sigma) = \|\sqrt{\rho}\sqrt{\sigma}\|_1^2$ [12, 15].

We can now state the main result of this paper as the following theorem.

Theorem 1. *Let $\mathcal{E}$ be an ensemble of quantum states $\{\rho_i\}$ with a priori probabilities $\{p_i\}$. Then, for any measurement M,*

$$P_E(M, \mathcal{E}) \geq \sum_{i>j} p_i p_j F(\rho_i, \rho_j).$$

We stress that this bound does not depend on the number of states in $\mathcal{E}$, nor their dimension. Before proving this theorem, we compare the lower bound of this note with some related previous results.

II. PREVIOUS WORK

A classic result of Holevo and Helstrom [10, 11] gives the exact minimum probability of error that can be achieved when discriminating between two states ρ_0 and ρ_1 with a priori probabilities p and $1 - p$:

$$\min_M P_E(M, \mathcal{E}) = \frac{1}{2} - \frac{1}{2}\|p\rho_0 - (1-p)\rho_1\|_1. \qquad (1)$$

However, in the case where we must discriminate between more than two states, no such exact expression for the minimum $P_E(M, \mathcal{E})$ is known. Indeed, it appears that until recently the only known lower bound on $P_E(M, \mathcal{E})$ was a result of Hayashi, Kawachi and Kobayashi that gives a bound in terms of the individual operator norms of the states in $\mathcal{E}$ [9]. A lower bound in terms of pairwise trace distances has very recently been given by Qiu [14].

In the other direction, Barnum and Knill [1] developed a useful upper bound on the error probability, which is given by

$$\min_M P_E(M, \mathcal{E}) \leq 2 \sum_{i>j} \sqrt{p_i p_j} \sqrt{F(\rho_i, \rho_j)}.$$

It was pointed out by Harrow and Winter [8] that this leads to a worst-case upper bound on the number of copies required to achieve a specified probability of suc-

*Electronic address: montanar@cs.bris.ac.uk

cess of discriminating between a set of states whose pairwise fidelities are known to be bounded above by some constant. Similarly, Theorem 1 can be used to lower bound the number of copies required in an average-case setting. For example, assume that each pair of states (ρ_i, ρ_j) has $F(\rho_i, \rho_j) \geq F$ for some F, that there are $n \geq 2$ equiprobable states to discriminate, and that we have m copies of the state to test. Then

$$P_E(M, \mathcal{E}) \geq \frac{1}{n^2} \sum_{i>j} F(\rho_i, \rho_j)^m \geq \frac{(n-1)F^m}{2n},$$

so in order to achieve a error probability of at most ϵ, we need to have access to at least

$$m \geq \frac{\log_2(1/\epsilon) - 2}{\log_2(1/F)}$$

copies of the test state.

Finally, we mention a related quantum state discrimination scenario that has been considered in the literature: *unambiguous* state discrimination [5]. In this scenario, our measurement process is not allowed to make a mistake. That is, it is required that the measurement result is i only if the input was state i. This can be achieved by allowing the possibility of failure, i.e. of outputting "don't know". Define $P_E^u(M, \mathcal{E})$ as the failure probability of an unambiguous measurement M on ensemble $\mathcal{E}$. Zhang et al. have given a lower bound on this probability of failure in terms of the pairwise fidelity and n, the number of states to be discriminated [16].

$$P_E^u(M, \mathcal{E}) \geq \frac{2}{n-1} \sum_{i>j} \sqrt{p_i p_j} |\langle \psi_i | \psi_j \rangle|.$$

Now let us turn to the proof of our main result.

III. PROOF OF THEOREM 1

We start by noting the following characterisation of a measurement based on that of Barnum and Knill [1]. Decompose each state (weighted by its a priori probability) in terms of its eigenvectors as $p_i \rho_i = \sum_j |e_{ij}\rangle\langle e_{ij}|$, where we fix the norm of each eigenvector $|e_{ij}\rangle$ as the square root of its corresponding eigenvalue λ_{ij}. Then define the matrix $S_i = \sum_j |e_{ij}\rangle\langle j|$, and form the overall block matrix S by writing the S_i matrices in a row. That is, $S = \sum_{i,j} |e_{ij}\rangle\langle i|\langle j|$. If the states are not of equal rank, pad each matrix S_i with zero columns so all the blocks are the same size.

Now perform the same task on an arbitrary measurement M. Perform the eigendecomposition of each measurement operator $\mu_i = \sum_j |f_{ij}\rangle\langle f_{ij}|$ (again, the norm of each eigenvector is given by the square root of its corresponding eigenvalue), and form the matrix N_i whose j'th column is $|e_{ij}\rangle$ (again, padding with zero columns if necessary). Write these matrices in a row to give $N = \sum_{i,j} |f_{ij}\rangle\langle i|\langle j|$. As $\sum_i \mu_i = I$, it is immediate that $NN^\dagger = I$.

Set $A = N^\dagger S$. A is made up of blocks $A_{ij} = N_i^\dagger S_j$. It is easy to verify that the probability of error of the measurement is completely determined by A:

$$\|A_{ij}\|_2^2 = \mathrm{tr}((N_i N_i^\dagger)(S_j S_j^\dagger)) = p_j \, \mathrm{tr}(\mu_i \rho_j),$$

so the squared 2-norm $\|A_{ij}\|_2^2$ gives the probability of receiving state j and identifying it as state i, and we have $P_E(M, \mathcal{E}) = \sum_{i \neq j} \|A_{ij}\|_2^2$.

Our proof rests on the fact that on the one hand $A^\dagger A = S^\dagger N N^\dagger S = S^\dagger S$, and on the other the pairwise fidelities of the states in $\mathcal{E}$ can also be obtained from $S^\dagger S$. Indeed, consider the (i,j)'th block of this matrix, $(S^\dagger S)_{ij} = S_i^\dagger S_j$. If the states in $\mathcal{E}$ are all pure (say $\rho_i = |\psi_i\rangle\langle\psi_i|$), then each block is a 1×1 matrix $(S^\dagger S)_{ij} = \sqrt{p_i}\sqrt{p_j}\langle\psi_i|\psi_j\rangle$. That is, $S^\dagger S$ is the Gram matrix of the states in $\mathcal{E}$ [2], scaled by their a priori probabilities.

More generally, we have $S_i S_i^\dagger = p_i \rho_i$. This implies that, by the polar decomposition of S_i, $S_i = \sqrt{p_i \rho_i}\, U$ for some unitary U. Thus, for some unitary U and V,

$$\begin{aligned}\|S_i^\dagger S_j\|_1^2 &= \|U^\dagger \sqrt{p_i \rho_i}\sqrt{p_j \rho_j}\, V\|_1^2 = p_i p_j \|\sqrt{\rho_i}\sqrt{\rho_j}\|_1^2 \\ &= p_i p_j F(\rho_i, \rho_j),\end{aligned}$$

where the second equality follows from the unitary invariance of the trace norm.

Our approach, following [1], will be to use these facts to lower bound the sum $\sum_{j \neq i} \|A_{ij}\|_2^2$ for a fixed i in terms of the entries of $A^\dagger A$, and then to sum over i. We will require two matrix norm inequalities. The first appears to be new, and the second was proven by Bhatia and Kittaneh using a duality argument [3]; we give a simple direct proof for completeness.

Lemma 2. *Let A, B, C, D be square matrices of the same dimension. Then*

$$\|AB + CD\|_1^2 \leq (\|A\|_2^2 + \|D\|_2^2)(\|B\|_2^2 + \|C\|_2^2).$$

Proof. Perform the polar decomposition $CD = PU$ for some positive semidefinite P and unitary U. Then

$$\begin{aligned}\|AB + CD\|_1 &= \|AB + PU\|_1 = \|AB + P^\dagger U\|_1 \\ &= \|ABU^\dagger + P^\dagger\|_1 = \|ABU^\dagger + UD^\dagger C^\dagger\|_1,\end{aligned}$$

where the third equality follows from the unitary invariance of the trace norm. Writing this as the product of two block matrices,

$$\begin{aligned}
&\|AB + CD\|_1^2 \\
&= \| \begin{pmatrix} A & UD^\dagger \end{pmatrix} \begin{pmatrix} BU^\dagger & C^\dagger \end{pmatrix}^{\mathrm{T}} \|_1^2 \\
&\le \|AA^\dagger + UD^\dagger DU^\dagger\|_1 \|UB^\dagger BU^\dagger + CC^\dagger\|_1 \\
&\le (\|AA^\dagger\|_1 + \|UD^\dagger DU^\dagger\|_1)(\|UB^\dagger BU^\dagger\|_1 + \|CC^\dagger\|_1) \\
&= (\|A\|_2^2 + \|D\|_2^2)(\|B\|_2^2 + \|C\|_2^2),
\end{aligned}$$

where the first inequality is the Cauchy-Schwarz inequality for unitarily invariant norms [2] and the second is the triangle inequality. □

Lemma 3 (Bhatia and Kittaneh [3]). *Let M be a block matrix $M = (M_1 \dots M_n)$. Then $\|M\|_1^2 \ge \sum_i \|M_i\|_1^2$.*

Proof. Let N_i be the matrix given by replacing all blocks in M other than block i with zeroes. Then it is easy to see that

$$M^\dagger M = \sum_i N_i^\dagger N_i$$

and also that $\|M\|_1 = \|\sqrt{M^\dagger M}\|_1$, $\|M_i\|_1 = \|\sqrt{N_i^\dagger N_i}\|_1$. Thus

$$\begin{aligned}
\|M\|_1^2 &= \|\sqrt{\sum_i N_i^\dagger N_i}\|_1^2 = \|\sum_i N_i^\dagger N_i\|_{1/2} \\
&\ge \sum_i \|N_i^\dagger N_i\|_{1/2} = \sum_i \|\sqrt{N_i^\dagger N_i}\|_1^2 = \sum_i \|M_i\|_1^2,
\end{aligned}$$

where the inequality in the second line can be proven easily by a majorisation argument [2], and is given explicitly as Lemma 1 of [4]. □

We now return to the proof of Theorem 1. Group the blocks of A into four "super-blocks" as follows:

$$A = \begin{pmatrix} \begin{pmatrix} A_{11} \\ A_{21} \\ \vdots \\ A_{n2} \end{pmatrix} & \begin{pmatrix} A_{12} & \dots & A_{1n} \\ A_{22} & \dots & A_{2n} \\ \vdots & \ddots & \vdots \\ A_{n2} & \dots & A_{nn} \end{pmatrix} \end{pmatrix}.$$

Now define a new 2×2 block matrix B by setting block B_{ij} to be the corresponding super-block in the above decomposition of A, appending rows and/or columns of zeroes to each of these blocks such that each block in B is square. Super-block A_{12} is thus the first row of block B_{12}. Consider the product $B^\dagger B$ with the same block structure. One can verify that the first row of the block $(B^\dagger B)_{12}$ is equal to the submatrix of $A^\dagger A$ defined as $T = ((A^\dagger A)_{12} \dots (A^\dagger A)_{1n})$, and the remaining rows in this block are zero. We therefore have $\|(B^\dagger B)_{12}\|_1 = \|T\|_1$. Using Lemma 3 followed by Lemma 2 gives

$$\begin{aligned}
\sum_{i>1} \|(A^\dagger A)_{1i}\|_1^2 \le \|T\|_1^2 &= \|B_{11}^\dagger B_{12} + B_{21}^\dagger B_{22}\|_1^2 \\
&\le (\|B_{11}\|_2^2 + \|B_{22}\|_2^2)(\|B_{12}\|_2^2 + \|B_{21}\|_2^2) \\
&\le \|B_{12}\|_2^2 + \|B_{21}\|_2^2 \\
&= \sum_{i>1} \|A_{1i}\|_2^2 + \|A_{i1}\|_2^2,
\end{aligned}$$

where we use the fact that $\sum_{i,j} \|B_{ij}\|_2^2 = \sum_{i,j} \|A_{ij}\|_2^2 = 1$ in the final inequality. We may now proceed to obtain corresponding inequalities for the other rows of A by permuting its rows and columns. Summing these inequalities, and noting that each off-diagonal element of A appears twice in total, gives

$$\begin{aligned}
P_E(M, \mathcal{E}) &= \sum_{i \ne j} \|A_{ij}\|_2^2 \ge \sum_{i>j} \|(A^\dagger A)_{ij}\|_1^2 \\
&= \sum_{i>j} \|(S^\dagger S)_{ij}\|_1^2 = \sum_{i>j} p_i p_j F(\rho_i, \rho_j)
\end{aligned}$$

and the proof is complete.

IV. CONCLUDING REMARKS

We have given a lower bound on the probability of error in quantum state discrimination that depends only on the pairwise fidelities of the states in question and is appealingly similar to a known upper bound of Barnum and Knill [1]. We close by commenting on the tightness of this bound.

It can be seen by comparing Theorem 1 with the Helstrom bound (1) that the lower bound of this paper is not always tight, even for two states, but is nevertheless close to optimal (in some sense). Consider a pair of identical states $\rho_0 = \rho_1 = \rho$ for some arbitrary ρ. Then, by (1),

$$\min_M P_E(M, \mathcal{E}) = \frac{1}{2} - \frac{1}{2}\|(p - (1-p))\rho\|_1 = \frac{1}{2} - |p - \frac{1}{2}|,$$

whereas Theorem 1 guarantees only a weaker lower bound of

$$\min_M P_E(M, \mathcal{E}) \ge p(1-p).$$

On the other hand, this lower bound cannot be improved by any constant factor $\alpha > 1$ without violating (1).

Note added. Following the completion of this work, I became aware of recent work by Qiu [14], which obtains a lower bound on $P_E(M, \mathcal{E})$ in terms of pairwise trace distances. For an ensemble of 2 states, this bound reduces to the Holevo-Helstrom quantity (1).

Acknowledgements

This work was supported by the EC-FP6-STREP network QICS. I would like to thank Richard Jozsa for helpful discussions on the subject of this paper, and Daowen Qiu for bringing reference [14] to my attention.

---

[1] H. Barnum and E. Knill. Reversing quantum dynamics with near-optimal quantum and classical fidelity. *J. Math. Phys.*, 43(5):2097–2106, 2002. `quant-ph/0004088`.

[2] R. Bhatia. *Matrix Analysis*. Springer-Verlag, 1997.

[3] R. Bhatia and F. Kittaneh. Norm inequalities for partitioned operators and an application. *Mathematische Annalen*, 287(1):719–726, 1990.

[4] R. Bhatia and F. Kittaneh. Cartesian decompositions and Schatten norms. *Linear Algebra and its Applications*, 318:109–116, 2000.

[5] A. Chefles. Quantum state discrimination. *Contemporary Physics*, 41(6):401–424, 2001. `quant-ph/0010114`.

[6] A. Childs, L. Schulman, and U. Vazirani. Quantum algorithms for hidden nonlinear structures. In *Proc. 47^{th} Annual Symp. Foundations of Computer Science*, 2007. `arXiv:0705.2784`.

[7] Y. C. Eldar, A. Megretski, and G. Verghese. Designing optimal quantum detectors via semidefinite programming. *IEEE Trans. Inform. Theory*, 49(4):1007–1012, 2003. `quant-ph/0205178`.

[8] A. Harrow and A. Winter. How many copies are needed for state discrimination?, 2006. `quant-ph/0606131`.

[9] M. Hayashi, A. Kawachi, and H. Kobayashi. Quantum measurements for hidden subgroup problems with optimal sample complexity, 2006. `quant-ph/0604174`.

[10] C. W. Helstrom. *Quantum detection and estimation theory*. Academic Press, New York, 1976.

[11] A. S. Holevo. Statistical decision theory for quantum systems. *Journal of Multivariate Analysis*, 3:337–394, 1973.

[12] R. Jozsa. Fidelity for mixed quantum states. *Journal of Modern Optics*, 41(12):2315–2323, 1994.

[13] A. Peres. *Quantum theory: concepts and methods*. Kluwer, 1995.

[14] D. Qiu. Minimum-error discrimination between mixed quantum states. *Phys. Rev. A.*, 77(1):012328, 2008. `arXiv:0707.3970`.

[15] A. Uhlmann. The “transition probability” in the state space of a *-algebra. *Rep. Math. Phys.*, 9(2):273–279, 1976.

[16] S. Zhang, Y. Feng, X. Sun, and M. Ying. Upper bound for the success probability of unambiguous discrimination among quantum states. *Phys. Rev. A.*, 64:062103, 2001.