COMPARISONS BETWEEN QUANTUM STATE DISTINGUISHABILITY MEASURES

KOENRAAD M.R. AUDENAERT

Department of Mathematics, Royal Holloway, University of London,
Egham TW20 0EX, United Kingdom

We provide a compendium of inequalities between several quantum state distinguishability measures. For each measure these inequalities consist of the sharpest possible upper and lower bounds in terms of another measure. Some of these inequalities are already known, but new or more general proofs are given, whereas other inequalities are new. We also supply cases of equality to show that all inequalities are indeed the sharpest possible.

Keywords: Trace norm distance, Uhlmann fidelity, relative entropy, Chernoff distance

1 Introduction

Over the years, a large number of quantum state distinguishability measures have been defined and subsequently studied. Prime examples are the trace norm distance, Uhlmann fidelity, relative entropy and Chernoff distance. Each of these measures has different properties, hence in different situations one or the other measure may be most suitable, or practical. It is therefore important to know how the different measures are related to one another. In particular, inequalities providing bounds on a measure in terms of another measure are very useful. One well-known such 'sandwich' bound is the one which says that the Uhlmann fidelity F is bounded above by $\sqrt{1-T^2}$ and bounded below by $1-T$, where T is the trace norm distance (see below for the exact definitions). Is this bound best possible? And what about comparisons of F with other measures?

In this note we present best possible dimension-independent upper and lower bounds (where they exist) between the following well-established distinguishability measures: relative entropy, Chernoff distance, Uhlmann fidelity, trace norm distance and the linear overlap (admittedly, this one is only useful for pure states). In addition, we also include a quantity which is essentially the Renyi relative entropy with parameter $1/2$, henceforth denoted by the letter Q. The reason for including this little-used measure is mathematical convenience, although it may deserve more prominence than it currently enjoys.

With 6 measures, in principle one has to consider 30 pairs (even though not every pair

yields an inequality; for example, there are no upper bounds on the relative entropy in terms of the other measures). Fortunately, most combinations can be derived from a handful of essential inequalities simply by chaining two or more of them together.

After introducing some basic mathematical definitions and tools, we give the definitions of the distinguishability measures that we study in this note in Section 3. The essential inequalities are given in Section 4 together with short proofs. Some of these inequalities have appeared already in the literature; nevertheless, the proofs given here are new unless indicated. Finally, in Section 5 we list all possible lower and upper bounds on each of the distinguishability measures considered, together with a number of cases of equality. All bounds provided are dimension-independent, and they are all best possible, as attested by the cases of equality. We hope that this compendium of relations will be found useful by the quantum information community.

2   Preliminaries

We will denote a diagonal matrix with diagonal elements $x_1, \ldots, x_n$ by $X = \mathrm{Diag}(x_1, \ldots, x_n)$. We will use the convention $0^t := 0$ for all real t and $0 \log 0 = 0$. That is, for positive semidefinite A, A^t and $A \log A$ are defined to be zero outside the support of A. For any matrix X let X^* denote its Hermitian conjugate and $|X|$ its absolute value (modulus); $|X| := \sqrt{X^* X}$. The Jordan decomposition of a Hermitian matrix X into its positive and negative parts is given by $X = X_+ - X_-$, with $X_\pm = (|X| \pm X)/2$, and $|X| = X_+ + X_-$.

The Schatten q-norms are the non-commutative generalisations of the ℓ_q norms. For any matrix A, its q-norm $||A||_q$, for $1 \le q$, is defined as

$$||A||_q = (\mathrm{Tr}\, |A|^q)^{1/q}.$$

One sees that this norm is equal to the ℓ_q norm of the vector of singular values of A. The Schatten 1-norm ($q = 1$) is well-known in quantum information theory as the trace norm: $||A||_1 = \mathrm{Tr}\, |A|$. It is equal to the sum of the singular values of A. In terms of the positive and negative parts of A, $||A||_1 = \mathrm{Tr}\, A_+ + \mathrm{Tr}\, A_- = 2\,\mathrm{Tr}\, A_+ - \mathrm{Tr}\, A$. Basic properties of the q-norm are: if $p < q$ then $||A||_q \le ||A||_p$; for positive semidefinite A and $p > 0$, $||A^p||_q = ||A||_{pq}^p$. We will occasionally need the quantity $||A||_q$ for values of q less than 1; while the definition allows this, it has to be kept in mind that this is no longer a norm but a quasi-norm.

An important inequality involving the Schatten norms is Hölder's inequality: for all matrices A, B, and indices p, q, r satisfying $p, q, r \ge 1$ and $1/p + 1/q \le 1/r$,

$$||AB||_r \le ||A||_p \ ||B||_q.$$

Extensions to products of more than 2 factors are obvious. More information can be found, e.g. in the standard textbook [1].

3   Distinguishability measures for pairs of states

Let ρ and σ be two density matrices. The following quantities are distinguishability measures between the two states, all of which (except L) have the property that they achieve their extremal (minimal or maximal) values for $\rho = \sigma$ and for $\rho\sigma = 0$, respectively.

$$L(\rho, \sigma) \quad = \quad \mathrm{Tr}\, \rho\sigma$$

$$
\begin{aligned}
T(\rho,\sigma) &= ||\rho-\sigma||_1/2 \\
F(\rho,\sigma) &= ||\rho^{1/2}\sigma^{1/2}||_1 \\
Q_s(\rho,\sigma) &= \mathrm{Tr}(\rho^s\sigma^{1-s}) \\
Q(\rho,\sigma) &= Q_{1/2}(\rho,\sigma) \\
Q_{\min}(\rho,\sigma) &= \min_{0\le s\le 1} Q_s(\rho,\sigma) \\
C(\rho,\sigma) &= -\log Q_{\min}(\rho,\sigma) \\
S(\rho||\sigma) &= \mathrm{Tr}\,\rho(\log\rho-\log\sigma).
\end{aligned}
$$

Here, L is known as the linear fidelity, or overlap, T as the trace norm distance, F as the Uhlmann fidelity, Q_s (essentially) as the Renyi relative entropy of order s, C as the Chernoff distance and S as the relative entropy. It has to be kept in mind that the relative entropy $S(\rho||\sigma)$ becomes infinite whenever the support of σ is not contained in the support of ρ; in that case any upper bound on any of the distinguishability measures in terms of the relative entropy becomes trivial. For commuting ρ and σ, the quantities Q and F coincide; they are two different non-commutative generalisations of a quantity that is alternatively known as the Hellinger affinity or the Bhattacharyya coefficient.

Note that L, T, F, $Q_{\min}$ and C are symmetric in their arguments; that is, the order of ρ and σ does not matter. In the case of F, this is because ρ and σ are Hermitian; for Hermitian A and B and any unitarily invariant norm, $|||AB||| = |||BA|||$.

The definitions of these quantities can be extended to non-normalised states, i.e. positive semidefinite matrices, allowing to absorb prior probabilities into the states. In what follows, let p be the prior probability of ρ and $1-p$ the one of σ, and let $A = p\rho$ and $B = (1-p)\sigma$ so that $A, B \ge 0$ and $\mathrm{Tr}(A+B) = 1$. We redefine the above distinguishability measures in terms of A and B in such a way that setting $p = 1/2$ we recover the original definitions for normalised states:

$$
\begin{aligned}
L(A,B) &= 4\,\mathrm{Tr}\,AB \\
T(A,B) &= ||A-B||_1 \\
F(A,B) &= 2||A^{1/2}B^{1/2}||_1 \\
Q_s(A,B) &= 2\,\mathrm{Tr}(A^sB^{1-s}) \\
Q(A,B) &= 2\,\mathrm{Tr}(A^{1/2}B^{1/2}).
\end{aligned}
$$

For every pair of these quantities one can ask for the sharpest possible upper and lower bounds on one quantity in terms of the other. Of the many possible combinations only a few inequalities turn out to be essential; by chaining them together in various ways all other combinations can easily be obtained.

$$
\begin{aligned}
L(A,B) &\le F^2(A,B) & (1) \\
1 - Q_s(A,B) &\le T(A,B) & (2) \\
T^2(A,B) &\le 1 - F^2(A,B) & (3) \\
F^2(\rho,\sigma) &\le Q_s(\rho,\sigma), \forall 0\le s\le 1 & (4) \\
Q(A,B) &\le F(A,B). & (5)
\end{aligned}
$$

For example, chaining $T^2 \leq 1 - F^2$ with $Q \leq F$ yields $T^2 \leq 1 - Q^2$, and chaining $1 - Q \leq T$ with $Q \leq F$ yields $1 - T \leq F$, neither of which can be improved. On the other hand, chaining $1 - Q \leq T$ with $T^2 \leq 1 - F^2$ yields $1 - \sqrt{1 - F^2} \leq Q$, but this is weaker than $F^2 \leq Q$ (which follows immediately from (4)). We will present all optimal chainings in Section 5, together with cases of equality.

4 Essential inequalities, with proofs

We now present the essential inequalities as theorems, and provide short proofs.

Theorem 1 *For $0 \leq s \leq 1$, $1 - T(A, B) \leq Q_s(A, B)$.*

This inequality was first proven for all $s \in [0, 1]$ in [2], although the special case for $s = 1/2$ had already been proven much earlier by Powers and Størmer [3], with a generalisation to von Neumann algebras by Araki [4]. In these works, the inequality appears in the form $||A^{\frac{1}{2}} - B^{\frac{1}{2}}||_2^2 \leq ||A - B||_1$. The book [5] by Hayashi also contains a proof of the $s = 1/2$ case. A much shorter proof of (2), due to N. Ozawa, appeared recently in [6], and allowed generalisation of this inequality to von Neumann algebras [7].

Proof: (Ozawa). Let $\Delta = A - B$ and define $C = B + \Delta_+ = A + \Delta_-$. Clearly, $0 \leq A, B \leq C$. Hence, by operator monotonicity of the fractional power x^s (for $0 \leq s \leq 1$), $0 \leq A^s, B^s \leq C^s$ and $0 \leq A^{1-s}, B^{1-s} \leq C^{1-s}$. Therefore, we have the chain of inequalities

$$
\begin{aligned}
\mathrm{Tr}(A - B^{1-s}A^s) &= \mathrm{Tr}(A^{1-s} - B^{1-s})A^s \\
&\leq \mathrm{Tr}(C^{1-s} - B^{1-s})A^s \\
&\leq \mathrm{Tr}(C^{1-s} - B^{1-s})C^s \\
&= \mathrm{Tr}(C - B^{1-s}C^s) \\
&= \mathrm{Tr}\,\Delta_+ + \mathrm{Tr}(B - B^{1-s}C^s) \\
&= \mathrm{Tr}\,\Delta_+ + \mathrm{Tr}\,B^{1-s}(B^s - C^s) \\
&\leq \mathrm{Tr}\,\Delta_+.
\end{aligned}
$$

Thus, $1 - T(A, B) = 2\,\mathrm{Tr}(A - \Delta_+) \leq 2\,\mathrm{Tr}\,B^{1-s}A^s = Q_s(A, B)$. □

Theorem 2 *$L(A, B) \leq F^2(A, B)$.*

Proof: This follows from the inequality $||X||_1 \leq ||X||_{1/2}$ applied to $X = B^{1/2}AB^{1/2}$ and noting that $\mathrm{Tr}(B^{1/2}AB^{1/2}) = \mathrm{Tr}(AB) = L(A, B)/4$ and $||B^{1/2}AB^{1/2}||_{1/2} = ||A^{1/2}B^{1/2}||_1^2 = F^2(A, B)/4$. □

Inequalities (4) and (5) also appeared in [2] (its Theorem 6 and equation (28), respectively).

Theorem 3 *For $0 \leq s \leq 1$, $F(\rho, \sigma) \leq (\mathrm{Tr}\,\rho^s \sigma^{1-s})^{1/2}$.*

Proof: We rewrite $\rho^{1/2}\sigma^{1/2}$ as a product of three factors

$$\rho^{1/2}\sigma^{1/2} = \rho^{(1-s)/2}(\rho^{s/2}\sigma^{(1-s)/2})\sigma^{s/2},$$

apply Hölder's inequality on the trace norm of this product, and exploit the relation $||X^p||_q = ||X||_{pq}^p$ (for $X \geq 0$):

$$
\begin{aligned}
||\rho^{1/2}\sigma^{1/2}||_1 &= ||\rho^{(1-s)/2}(\rho^{s/2}\sigma^{(1-s)/2})\sigma^{s/2}||_1 \\
&\leq ||\rho^{(1-s)/2}||_{2/(1-s)} \, ||\rho^{s/2}\sigma^{(1-s)/2}||_2 \, ||\sigma^{s/2}||_{2/s} \\
&= (\mathrm{Tr}\,\rho)^{(1-s)/2} \, ||\rho^{s/2}\sigma^{(1-s)/2}||_2 \, (\mathrm{Tr}\,\sigma)^{s/2} \\
&= (\mathrm{Tr}\,\rho^s \sigma^{1-s})^{1/2}.
\end{aligned}
$$

□

Theorem 4 $Q(A,B) \leq F(A,B)$.

Proof: This follows from the inequality $|\operatorname{Tr} X| \leq ||X||_1$ applied to $X = A^{1/2}B^{1/2}$, which has positive trace, hence $|\operatorname{Tr} X| = \operatorname{Tr} X$. □

Theorem 5 $F^2(A,B) + T^2(A,B) \leq 1$.

This inequality was first proven by Fuchs and van de Graaf [8], albeit only for a uniform prior. We provide a general and shorter proof; again this proof can be generalised to the von Neumann algebra setting [4].

Proof: First note that, for any unitary matrix U, we have

$$2(A - B) = (A^{\frac{1}{2}} + UB^{\frac{1}{2}})^*(A^{\frac{1}{2}} - UB^{\frac{1}{2}}) + (A^{\frac{1}{2}} - UB^{\frac{1}{2}})^*(A^{\frac{1}{2}} + UB^{\frac{1}{2}}).$$

Exploiting first the triangle inequality for the trace norm, and then the Cauchy-Schwarz inequality yields

$$||A - B||_1 \leq ||A^{\frac{1}{2}} + UB^{\frac{1}{2}}||_2 \, ||A^{\frac{1}{2}} - UB^{\frac{1}{2}}||_2.$$

Taking the square and simplifying the right hand side then gives

$$T(A,B)^2 = ||A - B||_1^2 \leq (\operatorname{Tr} A + \operatorname{Tr} B)^2 - (2\operatorname{Re}\operatorname{Tr}(A^{\frac{1}{2}}UB^{\frac{1}{2}}))^2 = 1 - (2\operatorname{Re}\operatorname{Tr}(A^{\frac{1}{2}}UB^{\frac{1}{2}}))^2.$$

As this holds for any unitary matrix U, the sharpest bound is obtained by minimising over all U. Noting that $\max_U \operatorname{Re}\operatorname{Tr}(A^{\frac{1}{2}}UB^{\frac{1}{2}}) = ||A^{\frac{1}{2}}B^{\frac{1}{2}}||_1 = F(A,B)$ yields the inequality of the theorem. □

Theorem 6 $-\log Q(\rho,\sigma) \leq \frac{1}{2}S(\rho||\sigma)$.

Proof: By Lemma 2.1 in [9], the function $s \mapsto \psi(s) := \log \operatorname{Tr} \rho^s \sigma^{1-s}$ is convex over $[0,1]$, for any choice of states ρ and σ. This implies the following inequality:

$$\frac{\psi(1) - \psi(1/2)}{1 - 1/2} \leq \psi'(1).$$

Since $\psi(1) = 0$, $\psi(1/2) = \log Q(\rho,\sigma)$ and $\psi'(1) = S(\rho||\sigma)$, we immediately get the inequality of the Theorem. □

The sharpest lower bound on the relative entropy in terms of the trace norm distance was found by Hiai, Ohya and Tsukuda [10]:

Theorem 7 (H-O-T) $S(\rho||\sigma) \geq s(||\rho - \sigma||_1/2)$.

Here, $s(x)$ is a special function defined for $0 \leq x < 1$ as

$$\begin{aligned}
s(x) &:= \min_{0 < r < 1-x} S(\operatorname{Diag}(r + x, 1 - r - x)||\operatorname{Diag}(r, 1 - r)) \\
&= \min_{x < r < 1} S(\operatorname{Diag}(r - x, 1 - r + x)||\operatorname{Diag}(r, 1 - r)).
\end{aligned}$$

In [11], the first three non-zero terms in its series expansion around $x = 0$ were determined:

$$s(x) = 2x^2 + \frac{4}{9}x^4 + \frac{32}{135}x^6 + O(x^8). \tag{6}$$

One can easily prove [12] that the lowest order expansion $2x^2$ is actually a lower bound, known as Pinsker's bound. For values of $x > 4/5$, $s(x)$ is well approximated by its upper bound

$$\begin{aligned}
s(x) &\leq \lim_{r \to 1-x} S(\operatorname{Diag}(r + x, 1 - r - x)||\operatorname{Diag}(r, 1 - r)) \\
&= -\log(1 - x).
\end{aligned}$$

5   All inequalities, with cases of equality

In the following, we consider four special families of pairs of states.

- **Family (a):**

$$\rho = \mathrm{Diag}(1,0), \quad \sigma = \mathrm{Diag}(t, 1 - t), 0 \le t \le 1.$$

For these states we have

$$L = F^2 = Q_{\min} = Q^2 = 1 - T = \exp(-S) = t.$$

- **Family (b):** Both ρ and σ are pure. For these states we have

$$L = F^2 = Q_{\min} = Q = 1 - T^2.$$

The relative entropy is infinite except for $\rho = \sigma$, in which case it is 0.
- **Family (c):**

$$\rho = \mathrm{Diag}(1 - t, t, 0), \quad \sigma = \mathrm{Diag}(1 - t, 0, t), 0 \le t \le 1.$$

For these states we have

$$\sqrt{L} = F = Q_{\min} = Q = 1 - T = 1 - t.$$

The relative entropy is infinite except for $t = 0$, in which case it is 0.
- **Family (d):**

$$\rho = \mathrm{Diag}(1 - t, t), \quad \sigma = \mathrm{Diag}(t, 1 - t), 0 \le t \le 1.$$

For these states we have

$$F = Q_{\min} = Q = 2\sqrt{t(1 - t)} = \sqrt{2L}; \quad T = |1 - 2t| = \sqrt{1 - F^2}.$$

The relative entropy is given by $S = (2t - 1)\log(t/(1 - t))$.

The basic inequalities that were proven in the above can be chained together in various ways. Below we present, for each of the considered distinguishability measures, the optimal lower and upper bounds (if there are any) in terms of the other measures. The letters above the inequality signs indicate for which of the four families of states considered above equality is obtained.

We do not attempt to provide all cases of equality; the goal here is merely to show that the presented inequalities cannot be improved. For example, the fact that all inequalities in the chain $L \le F^2 \le Q_{\min} \le Q$ become equalities for the same family of states (in this case family (b)) shows that $L \le Q$ is the best possible upper bound on L in terms of Q.

To shorten the formulas, we define the function

$$v(x) := \sqrt{1 - x^2}.$$

- L:

$$L \stackrel{a,b,c}{\leq} F^2 \quad \begin{cases} \stackrel{a,b}{\leq} Q_{\min} \stackrel{b,c,d}{\leq} Q \\[2mm] \stackrel{b,d}{\leq} 1 - T^2 \end{cases}$$

There are no lower bounds on L in terms of the other measures because L is unable to indicate equality of the states when not both of them are pure; indeed, by taking $\rho = \sigma$ every other measure would give 0 (or 1), but L could assume any value between 0 and 1 depending on the purity of ρ.

There is no upper bound on L in terms of S. For the choice $\rho = \mathrm{Diag}(1 - \epsilon, \epsilon)$ and $\sigma = \mathrm{Diag}(1, 0)$, $L = 1 - \epsilon$ and $S = \infty$, whereas for the choice $\rho = \sigma = \mathrm{Diag}(1 - x, x)$ with $x = \epsilon/2$, $L \approx 1 - \epsilon$ too but $S = 0$.

- F:

$$\left. \begin{array}{c} 1 - T \stackrel{c}{\leq} Q_{\min} \stackrel{b,c,d}{\leq} \\[2mm] \exp(-S/2) \stackrel{a}{\leq} \end{array} \right\} \stackrel{a,c,d}{Q \leq} \left. \begin{array}{c} \\ \sqrt{L} \stackrel{a,b,c}{\leq} \end{array} \right\} F \begin{cases} \stackrel{a,b}{\leq} \sqrt{Q_{\min}} \stackrel{b,c,d}{\leq} \sqrt{Q} \\[2mm] \stackrel{b,d}{\leq} v(T) \end{cases}$$

From this diagram one obtains, for example, the known sandwich bound $1 - T \leq F \leq \sqrt{1 - T^2}$. There is no upper bound on F in terms of S, as it would imply an upper bound on S in terms of F; however, S can be infinite for any allowed value of F.

- T:

$$\left. \begin{array}{c} 1 - F \stackrel{a,c,d}{\leq} 1 - Q \stackrel{b,c,d}{\leq} 1 - Q_{\min} \stackrel{a,c}{\leq} \\[2mm] s^{-1}(S) \leq \end{array} \right\} T \begin{cases} \stackrel{b,d}{\leq} v(F) \stackrel{a,c,d}{\leq} v(Q) \stackrel{b,c,d}{\leq} v(Q_{\min}) \\[2mm] \stackrel{a}{\leq} \sqrt{1 - L} \end{cases}$$

Equality in the H-O-T bound is achieved for certain 2×2 diagonal states.

- Q:

$$\left. \begin{array}{c} L \stackrel{a,b,c}{\leq} F^2 \stackrel{a,b}{\leq} \\[2mm] 1 - T \stackrel{c}{\leq} \end{array} \right\} \stackrel{b,c,d}{Q_{\min} \leq} \left. \begin{array}{c} \\ \exp(-S/2) \stackrel{a}{\leq} \end{array} \right\} Q \stackrel{a,c,d}{\leq} F \begin{cases} \stackrel{a,b}{\leq} \sqrt{Q_{\min}} \\[2mm] \stackrel{b,d}{\leq} v(T) \end{cases}$$

- $Q_{\min}$:

$$\left. \begin{array}{c} L \stackrel{a,b,c}{\leq} \\[2mm] \exp(-S) \stackrel{a}{\leq} Q^2 \stackrel{a,c,d}{\leq} \end{array} \right\} \left. \begin{array}{c} F^2 \stackrel{a,b}{\leq} \\[2mm] 1 - T \stackrel{a,c}{\leq} \end{array} \right\} Q_{\min} \stackrel{b,c,d}{\leq} Q \stackrel{a,c,d}{\leq} F \stackrel{b,d}{\leq} v(T)$$

Noting that the Chernoff divergence C is just minus the logarithm of $Q_{\min}$ we immediately obtain bounds on C. For example: $C \leq S$ and $C \leq -2\log F \leq -2\log Q \leq 2C$.

Acknowledgements

References

1. R. Bhatia, *Matrix Analysis*, Springer, (1997).
2. K.M.R. Audenaert, M. Nussbaum, A. Szkoła, F. Verstraete, *Asymptotic error rates in quantum hypothesis testing*, Commun. Math. Phys. **279**, 251–283, (2008).
3. R.T. Powers and E. Størmer, *Free states of the canonical anticommutation relations*, Commun. Math. Phys. **16**, 1–33 (1970).
4. H. Araki, *Some Properties of Modular Conjugation Operator of von Neumann Algebras and a Non-Commutative Radon-Nikodym Theorem with a Chain Rule*, Pacific J. Math. **50**, 309–354 (1974).
5. M. Hayashi, *Quantum Information: an Introduction*, Springer (2006).
6. V. Jaksic, Y. Ogata, C-A. Pillet and R. Seiringer, *Quantum hypothesis testing and non-equilibrium statistical mechanics*, Rev. Math. Phys. **24**, 1–67 (2012).
7. Y. Ogata, *A generalization of Powers-Størmer inequality*, Lett. Math. Phys. **97**, 339–346 (2011).
8. C.A. Fuchs and J. van de Graaf, *Cryptographic distinguishability measures for quantum-mechanical states*, IEEE Trans. Inf. Theory **45**, 1216 (1999).
9. M. Mosonyi, K.M.R. Audenaert and F. Verstraete, *State discrimination bounds for finite sample size*, J. Math. Phys. **53**, 122205 (2012).
10. F. Hiai, M. Ohya, and M. Tsukada, *Sufficiency, KMS Condition and Relative Entropy in von Neumann Algebras*, Pacific J. Math. **96**, 99–109 (1981).
11. K.M.R. Audenaert and J. Eisert, *Continuity bounds on the quantum relative entropy*, J. Math. Phys. **46**, 102104 (2005).
12. I. Csiszar, *Information-type measures of difference of probability distributions and indirect observations*, Studia Sci. Math. Hungar. **2**, 299–318 (1967).