Quantum Computation **(CMU 18-859BB, Fall 2015)**

Lecture 17: Discriminating Two Quantum States

November 5, 2015

Lecturer: Ryan O'Donnell *Scribe: Zilin Jiang*

1 Recap and more

1.1 Mixed state

Recall that a *mixed state* $\{p_i, |\psi_i\rangle\}$, where $|\psi_i\rangle \in \mathbb{C}^d$, is represented by *density matrix* $\rho = \sum_i p_i |\psi_i\rangle \langle\psi_i| \in \mathbb{C}^{d\times d}$. Conversely, a density matrix ρ, by definition,

1. is positive semidefinite (PSD), and so it has orthonormal eigenvectors $|v_1\rangle, \ldots, |v_d\rangle$ with nonnegative eigenvalues $\lambda_1, \ldots, \lambda_d$;
2. and it satisfies $\operatorname{tr}(\rho) = 1$, and so $\lambda_1 + \cdots + \lambda_d = 1$.

Therefore ρ represents mixed state $\{\lambda_i, |v_i\rangle\}$.

We shall use repeatedly use the following properties of traces in this lecture.

Fact 1.1. *Here are some basic properties of the trace operator.*

1. $\operatorname{tr}(A) = \sum_i A_{ii} = \sum \lambda_i$, *where* λ_i*'s are the eigenvalues of* A*;*
2. *and* $\operatorname{tr}(AB) = \sum_{i,j} A_{ij}B_{ji} = \operatorname{tr}(BA)$.

Corollary 1.2. *The following properties follow immediately from the second fact above.*

1. $\operatorname{tr}(|\psi\rangle\langle\phi|) = \operatorname{tr}(\langle\phi|\psi\rangle) = \langle\phi|\psi\rangle$;
2. *the trace operator is invariant under orthonormal transformation: given unitary matrix* U*, we have*
$$\operatorname{tr}\left(UMU^\dagger\right) = \operatorname{tr}\left(U^\dagger UM\right) = \operatorname{tr}(M).$$

1.2 General measurement

A general measurement is defined by matrices $M_1, \ldots, M_m$ such that $\sum_i M_i^\dagger M_i = I$. If this general measurement is carried out on a mixed state represented by ρ, in the last lecture we have shown that we will see outcome "j" with probability

$$p_j := \operatorname{tr}\left(M_j^\dagger M_j \rho\right) = \operatorname{tr}\left(M_j \rho M_j^\dagger\right) \tag{1}$$

and the state collapses a mixed stated represented by

$$\frac{M_j \rho M_j^\dagger}{p_j}. \tag{2}$$

If we measure a mixed state represented by ρ in orthonormal basis $|v_1\rangle, \ldots, |v_d\rangle$, or equivalently we carry out a measurement described by $M_i = |v_i\rangle\langle v_i|$, then, from (1) and (2), we see outcome "j" with probability

$$p_j := \langle v_j| \rho |v_j\rangle = \mathrm{tr}\left(|v_j\rangle\langle v_j| \rho\right) \tag{3}$$

and the state collapses to the mixed state represented by

$$\frac{1}{p_j} |v_j\rangle\langle v_j| \rho |v_j\rangle\langle v_j| . \tag{4}$$

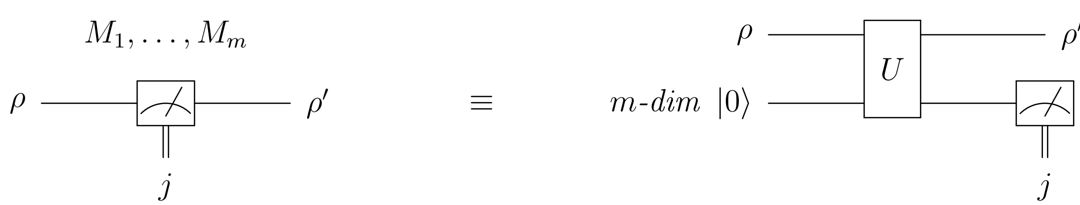
Let $E_i = M_i^\dagger M_i$ in the definition of general measurement. We know that $\{E_i\}$ is a set of PSD matrices such that $\sum_i E_i = I$. Suppose that we are only given E_i. From (1), we will still be able to determine the probability of outcome "j": $p_j = \mathrm{tr}\left(E_i \rho\right)$. This makes this set $\{E_i\}$ intersting.

Definition 1.3. A *positive-operator valued measure (POVM)* is a set of PSD matrices $E_1, \ldots, E_m$ such that $E_1 + E_2 + \cdots + E_m = I$.

Specifying E_i's gives probability distribution of the outcome, but implementation of such a measurement still needs M_i. Moreover, from (2), we can see that POVM does not tell what the state collapses to.

Though we have generalized the definitions of quantum state and measurement, we haven't actually added anything new at all. On one hand, one can still reason about mixed states through conditional probability. On the other hand, density matrices are not new due to Naimark's (Neumark's) theorem. The theorem will appear as a homework problem.

Theorem 1.4. *A general measurement described by $M_1, \ldots, M_m$ is equivalent to a quantum circuit with m-dimensional ancilla and a unitary gate U followed by a partial (simple) measurement on the ancilla bits.*



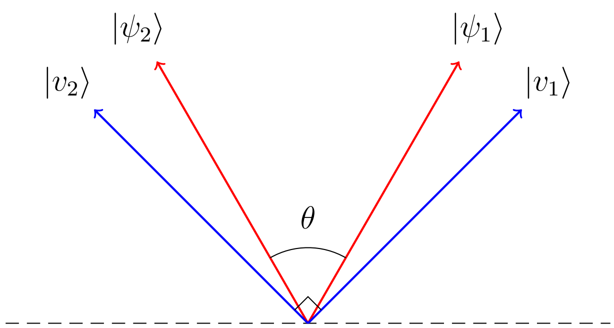
So the point of the generalization should really be thought as a new mathematical way to describe quantum states and measurements nicely.

2 Discriminating two pure states

Given a pure d-dimensional state $|\psi\rangle$ known to be either $|\psi_1\rangle$ or $|\psi_2\rangle$. You must guess which state $|\psi\rangle$ really is. There are two kinds of errors: the probability p_1 we guess wrong when it's $|\psi_1\rangle$ and the probability p_2 we guess wrong when it's $|\psi_2\rangle$. The goal is to find a strategy to minimize $\max(p_1, p_2)$.

Assume with out loss of generality the angle between $|\psi_1\rangle$ and $|\psi_2\rangle$, say θ, is between 0 and $\pi/2$. Otherwise, we replace $|\psi_1\rangle$ by $-|\psi_1\rangle$.

Theorem 2.1. *The best strategy is to do the projective measurement with $\{|v_1\rangle, |v_2\rangle\}$, where $|v_1\rangle, |v_2\rangle$ are in the span of $|\psi_1\rangle$ and $|\psi_2\rangle$ such that $\langle v_1|v_2\rangle = 0$, they are symmetric with respect to the angle bisector of $|\psi_1\rangle$ and $|\psi_2\rangle$, and $|v_i\rangle$ is closer to $|\psi_i\rangle$ for $i = 1, 2$. On outcome "$|v_i\rangle$", we guess $|\psi_i\rangle$.*

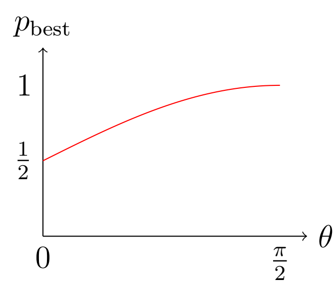


Remark 2.2. The strategy makes sense as measurements in the space perpendicular to $|\psi_1\rangle$ and $|\psi_2\rangle$ does not reveal information. Moreover, the best strategy should be symmetric with respect to $|\psi_1\rangle$ and $|\psi_2\rangle$.

The probability of success using the best strategy is

$$|\langle\psi_1|v_1\rangle|^2 = \cos^2(\angle(\psi_1, v_1)) = \cos^2\left(\frac{\pi/2 - \theta}{2}\right) = \frac{1}{2} + \frac{1}{2}\cos\left(\frac{\pi}{2} - \theta\right) = \frac{1}{2} + \frac{1}{2}\sin\theta.$$

The best probability distribution of success is plotted below.



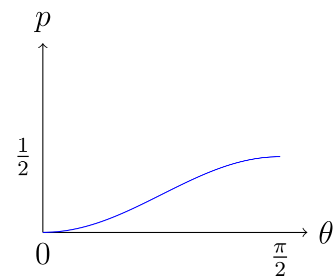
The optimality of the strategy follows from Theorem 3.4 in the next subsection. Over there, we actually prove the average error probability $(p_1 + p_2)/2 \geq \frac{1}{2} - \frac{1}{2}\sin\theta$ and so $\max(p_1, p_2) \geq \frac{1}{2} - \frac{1}{2}\sin\theta$.

2.1 Unambiguous discrimination

Now you can guess which state $|\psi\rangle$ or say "don't know", but you may never be wrong. One simple strategy is to do the projective measurement with $\left\{|\psi_1\rangle, |\psi_1^\perp\rangle\right\}$. If the outcome is $|\psi_1\rangle$, we say "don't know", otherwise we are sure that it's $|\psi_2\rangle$. This strategy gives

$$\mathbf{Pr}\left[\text{“don't known”}\right] = \begin{cases} 1 & \text{if it is } |\psi_1\rangle \\ \cos^2\theta & \text{if it is } |\psi_2\rangle \,. \end{cases}$$

A slightly cleverer strategy is to do, with probability 1/2, the projective measurement with $\left\{|\psi_1\rangle, |\psi_1^\perp\rangle\right\}$ and do, with probability 1/2, the projective measurement with $\left\{|\psi_2\rangle, |\psi_2^\perp\rangle\right\}$. Overall, the probability of "don't know" is $\frac{1}{2} + \frac{1}{2}\cos^2\theta$ and the probability of success is $\frac{1}{2} - \frac{1}{2}\cos^2\theta = \frac{1}{2}\sin^2\theta$. Its probability distribution is plotted below.



Apparently, we can do better in the case when $\theta = \pi/2$. In fact, we are able to decide deterministically $|\psi\rangle$ in this case. This suggests that the slightly cleverer strategy might not be the best strategy.

The strategy above can be thought as a mixture of a projection onto $|\psi_i^\perp\rangle$ for $i = 1, 2$. Set $\Pi_i = I - |\psi_i\rangle\langle\psi_i|$ for $i = 1, 2$. You might want to try the general measurement described by $M_i = \Pi_i$ or POVM with $E_i = M_i^\dagger M_i = \Pi_i$. This is not OK since $E_1 + E_2$ may not equal to I. If $E_1 + E_2 \preceq I$[1], then $E_0 = I - E_1 - E_2$ is PSD and $\{E_0, E_1, E_2\}$ is a POVM. However, in general $E_1 + E_2 \preceq I$ might not be the case.

Here is the way to fix this idea. Say maximum eigenvalue of $\Pi_1 + \Pi_2$ is c. So $\Pi_1 + \Pi_2 \preceq cI$. We will let $E_1 = \Pi_1/c, E_2 = \Pi_2/c, E_0 = I - E_1 - E_2 \succeq 0$.

Claim 2.3. *The maximum eigenvalue, c, of $\Pi_1 + \Pi_2$, is $1 + \cos\theta$.*

Proof. Here's the proof[2]. We have that c is the maximum, over all unit $|v\rangle$, of

$$\langle v|(\Pi_1 + \Pi_2)|v\rangle = \langle v|\Pi_1|v\rangle + \langle v|\Pi_2|v\rangle = \sin^2\theta_1 + \sin^2\theta_2,$$

where θ_i is the angle $|v\rangle$ makes with $|\psi_i\rangle$, and we used the most basic geometry. It's clear that this is maximized when $|v\rangle$ is in the same 2-dimensional plane as $|\psi_1\rangle, |\psi_2\rangle$. So we're now maximizing

$$\sin^2\theta_1 + \sin^2\theta_2 = 1 - \frac{1}{2}(\cos(2\theta_1) + \cos(2\theta_2))$$

[1] If $X - Y$ is PSD, we write $X \succeq Y$.

[2] The proof below is based on a Piazza post by Ryan O'Donnell.

subject to $\theta_1 + \theta_2 = \theta$, where I used the double-angle formula $\cos(2\alpha) = 1 - 2\sin^2\alpha$. Using $\cos(\alpha) + \cos(\beta) = 2\cos\left(\frac{\alpha+\beta}{2}\right)\cos\left(\frac{\alpha-\beta}{2}\right)$, we have that

$$\frac{1}{2}(\cos(2\theta_1) + \cos(2\theta_2)) = \cos(\theta_1 + \theta_2)\cos(\theta_1 - \theta_2) = \cos\theta\cos(\theta_1 - \theta_2) \geq -\cos\theta.$$

Hence $c \leq 1 + \cos\theta$ and equality occurs when $\theta_1 = \theta/2 - \pi/2$ and $\theta_2 = \theta/2 + \pi/2$. □

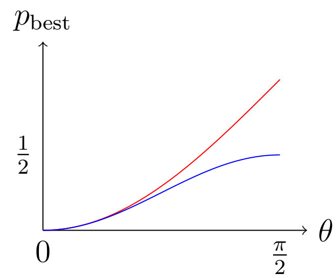
Fact 2.4. *The best strategy to unambiguously discriminate $|\psi_1\rangle$ and $|\psi_2\rangle$ is to carry out the POVM $\{E_0, E_1, E_2\}$. We guess $|\psi_2\rangle$ if the outcome is 1, we guess $|\psi_1\rangle$ if the outcome is 2, and we say "don't know" if the outcome is 0.*

For a proof of the optimality of the strategy, see [Per88, Iva87, Die88]. Here we only compute the best success probability. Say the state is $|\psi_2\rangle$. Then density matrix is $\rho = |\psi_2\rangle\langle\psi_2|$. We compute the probability of each outcome:

$$\begin{aligned}
\mathbf{Pr}[\text{outcome } 1] &= \operatorname{tr}(E_1\rho) \\
&= \operatorname{tr}\left(\frac{\Pi_1}{c}|\psi_2\rangle\langle\psi_2|\right) \\
&= \frac{1}{c}\operatorname{tr}\left((I - |\psi_1\rangle\langle\psi_1|)|\psi_2\rangle\langle\psi_2|\right) \\
&= \frac{1}{c}\operatorname{tr}\left(|\psi_2\rangle\langle\psi_2| - |\psi_1\rangle\langle\psi_1||\psi_2\rangle\langle\psi_2|\right) \\
&= \frac{1}{c}\left(\operatorname{tr}(|\psi_2\rangle\langle\psi_2|) - \operatorname{tr}(|\psi_1\rangle\langle\psi_1||\psi_2\rangle\langle\psi_2|)\right) \\
&= \frac{1}{c}\left(\langle\psi_2|\psi_2\rangle - |\langle\psi_1|\psi_2\rangle|^2\right) \\
&= \frac{1}{c}(1 - \cos^2\theta) \\
&= 1 - \cos\theta; \\
\mathbf{Pr}[\text{outcome } 2] &= \operatorname{tr}(E_2\rho) \\
&= \operatorname{tr}\left(\frac{\Pi_2}{c}|\psi_2\rangle\langle\psi_2|\right) \\
&= 0; \\
\mathbf{Pr}[\text{outcome } 0] &= \cos\theta.
\end{aligned}$$

The probability of success is plotted below (in red) in comparison with the previous strategy (in blue).

In general, if we want to unambiguously discriminate > 2 pure states, it is not clear how to find expression of the best probability is closed form. Using semidefinite programming in general gives good bounds on the best probability.



3 Discriminating two mixed states

Given a mixed state ρ known to be either ρ_1 or ρ_2. You must guess which state ρ really is. Assume it's ρ_1 with probability 1/2 and ρ_2 with probability 1/2. Our goal is to come up with a strategy to minimize the error probability.

Suppose ρ_1 represents ensemble $\{p_i, |\psi_i\rangle\}_{i=1}^d$, where p_i's are the eigenvalues of ρ_1 and $|\psi_i\rangle$'s are the eigenvectors. Similarly, ρ_2 represents ensemble $\{q_i, |\phi_i\rangle\}_{i=1}^d$.

3.1 Easy case

An easy case is when $|\psi_i\rangle = |\phi_i\rangle$ for all $i \in [d]$. Now ρ_1 and ρ_2 are simultaneously diagonalizable. Without loss of generality, we can assume that $|\psi_i\rangle = |\phi_i\rangle = |i\rangle$ for all $i \in [d]$ and

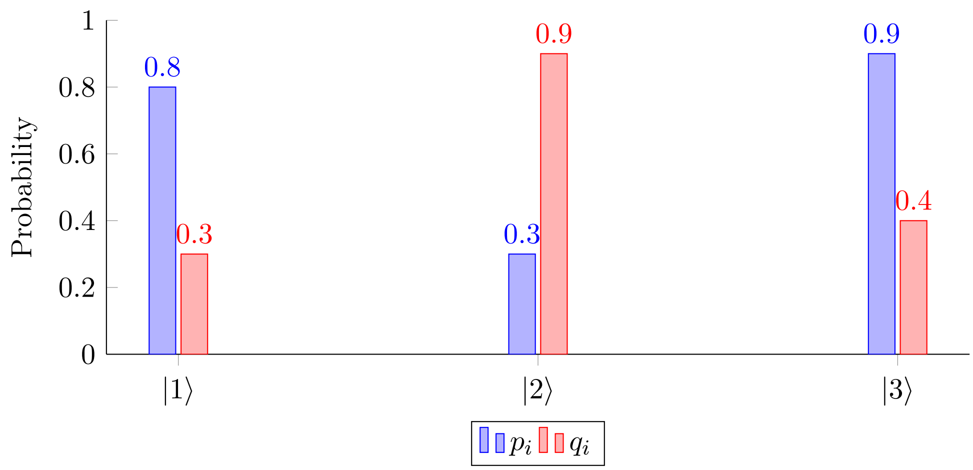
$$\rho_1 = \begin{pmatrix} p_1 & & \\ & \ddots & \\ & & p_d \end{pmatrix} \quad \text{and} \quad \rho_2 = \begin{pmatrix} q_1 & & \\ & \ddots & \\ & & q_d \end{pmatrix}.$$

The optimal strategy is to measure in the standard basis and the problem of discriminating ρ_1 and ρ_2 reduces to a purely classical problem. Now we know that if the state is ρ_1, then with probability p_i you see $|i\rangle$, and if the state is ρ_2, then with probability q_i you see $|i\rangle$.

For example, suppose $d = 3$ and the p_i's and q_i's are displayed in the histogram below. If you see "$|1\rangle$", you should guess ρ_1; if you see "$|2\rangle$", you should guess ρ_2; and if you see "$|3\rangle$", you should guess ρ_1.

Let $A = \{i : p_i \geq q_i\}$. It is optimal to guess ρ_1 if and only if the outcome is in A.

Definition 3.1. The statistical distance / total variation of probability distributions $(p_1, \ldots, p_d)$ and $(q_1, \ldots, q_d)$ is defined by $d_{\text{TV}}(\{p_i\}, \{q_i\}) := \frac{1}{2} \sum_i |p_i - q_i|$



Now we compute the success probability:

$$\begin{aligned}\mathbf{Pr}\left[\text{success}\right] &= \frac{1}{2}\mathbf{Pr}\left[i \in A \mid \rho = \rho_1\right] + \frac{1}{2}\mathbf{Pr}\left[i \notin A \mid \rho = \rho_2\right] \\ &= \frac{1}{2}\sum_{i\in A} p_i + \frac{1}{2}\sum_{i \notin A} q_i \\ &= \frac{1}{2}\sum_i \max\left(p_i, q_i\right) \\ &= \frac{1}{2}\sum_i \left(\frac{p_i + q_i}{2} + \frac{|p_i - q_i|}{2}\right) \\ &= \frac{1}{2} + \frac{1}{2} d_{\mathrm{TV}}(\{p_i\}, \{q_i\}).\end{aligned}$$

3.2 General case

In order to state the result in the general case, we first define the p-norm of a matrix:

Definition 3.2. Suppose matrix A is Hermitian. The p-norm of A is $\|A\|_p := (\sum_i |\lambda_i|^p)^{1/p}$, where λ_i's are the eigenvalues of A. The ∞-norm of A is $\|A\|_\infty := \max_i(\lambda_i)$

In the easy case, the success probability can be rewritten as $\frac{1}{2} + \frac{1}{2}(\frac{1}{2}\|\rho_1 - \rho_2\|_1)$.

Remark 3.3. $\frac{1}{2}\|\rho_1 - \rho_2\|_1$ is called the trace distance between ρ_1 and ρ_2. There are different notations for $\frac{1}{2}\|\rho_1 - \rho_2\|_1$, such as $D, \delta, T, d_{\mathrm{tv}}, d_{\mathrm{TV}}$.

Now we drop the assumption that ρ_1 and ρ_2 are simultaneously diagonalizable, and consider the general case. We have the following result (see [Hol73, Hel69]).

Theorem 3.4 (Holevo–Helstrom)**.** *In general, the best success probability to discriminate two mixed states represented by ρ_1 and ρ_2 is given by $\frac{1}{2} + \frac{1}{2}(\frac{1}{2}\|\rho_1 - \rho_2\|_1)$.*

To prove the theorem, we need the Hölder's inequality for matrices.

Lemma 3.5. *Given two Hermitian matrices* A, B*. We have that*

$$\operatorname{tr}(AB) \leq \|A\|_p \|B\|_q, \forall p, q \in [1, \infty] \ s.t. \ \frac{1}{p} + \frac{1}{q} = 1.$$

Proof of Theorem 3.4 assuming Lemma 3.5. Say we have any POVM $\{E_1, E_2\}$ such that E_1, E_2 are PSD and $E_1 + E_2 = I$. Our strategy is to guess ρ_i when the outcome is "i". The success probability can be expressed as

$$\frac{1}{2}\operatorname{tr}(E_1\rho_1)+\frac{1}{2}\operatorname{tr}(E_2\rho_2) = \frac{1}{4}\operatorname{tr}((E_1+E_2)(\rho_1+\rho_2))+\frac{1}{4}\operatorname{tr}((E_1+E_2)(\rho_1+\rho_2)) =: \frac{1}{2}T_1+\frac{1}{2}T_2,$$

the first equality of which uses the linearity of the trace operator. On one hand, because $E_1 + E_2 = I$ and $\operatorname{tr}(\rho_i) = 1$, $T_1 = \frac{1}{2}\operatorname{tr}(\rho_1 + \rho_2) = \frac{1}{2}\operatorname{tr}(\rho_1) + \frac{1}{2}\operatorname{tr}(\rho_2) = 1$. On the other hand, by Lemma 3.5, we have

$$T_2 = \frac{1}{2}\operatorname{tr}((E_1 - E_2)(\rho_1 - \rho_2)) \leq \frac{1}{2}\|E_1 - E_2\|_\infty \|\rho_1 - \rho_2\|_1.$$

Since $0 \preceq E_1, E_2 \preceq I$, $\|E_1 - E_2\|_\infty \leq 1$ and so $T_1 + T_2 \leq \frac{1}{2} + \frac{1}{2}\|\rho_1 - \rho_2\|_1$ gives an upper bound of the best success probability. This bound is achievable. Suppose $\Delta = \rho_1 - \rho_2$ has some eigenvalues $\lambda_1, \ldots, \lambda_m \geq 0$ and $\mu_1, \ldots, \mu_{m'} < 0$. Let P be the eigenspace associated to $\lambda_1, \ldots, \lambda_m$ and Q be the eigenspace associated to $\mu_1, \ldots, \mu_{m'}$. Let E_1 be the projection into P and let E_2 be the projection into Q. Note that $E_1\Delta$ has eigenvalues $\lambda_1, \ldots, \lambda_m$ and $E_2\Delta$ has eigenvalues $\mu_1, \ldots, \mu_{m'}$. We check that $T_2 \leq \frac{1}{2}\|\Delta\|_1$ can achieve equality.

$$T_2 = \frac{1}{2}\operatorname{tr}((E_1 - E_2)\Delta) = \frac{1}{2}\operatorname{tr}(E_1\Delta) - \frac{1}{2}\operatorname{tr}(E_2\Delta) = \frac{1}{2}\sum \lambda_i - \frac{1}{2}\sum \mu_i = \frac{1}{2}\|\Delta\|_1.$$

□

Proof of Lemma 3.5. Since A, B are Hermitian, we can write

$$A = \sum_{i=1}^{d} p_i |u_i\rangle \langle u_i|, B = \sum_{j=1}^{d} q_j |v_j\rangle \langle v_j|,$$

where both $\{|u_i\rangle\}$ and $\{|v_i\rangle\}$ are orthonormal bases. Now we rewrite the left hand side of the inequality in terms of p_i, q_j, u_i, v_j and then apply the (usual) Hölder's inequality:

$$\begin{aligned}
\operatorname{tr}(AB) &= \operatorname{tr}\left(\sum_i p_i |u_i\rangle\langle u_i| \sum_j q_j |v_j\rangle\langle v_j|\right) \\
&= \sum_{i,j} p_i q_j \operatorname{tr}(|u_i\rangle\langle u_i|v_j\rangle\langle v_j|) \\
&= \sum_{i,j} p_i q_j |\langle u_i|v_j\rangle|^2 \\
&= \sum_{i,j} \left(p_i |\langle u_i|v_j\rangle|^{2/p}\right)\left(q_j |\langle u_i|v_j\rangle|^{2/p}\right) \\
&\le \left(\sum_{i,j} p_i^p |\langle u_i|v_j\rangle|^2\right)^{1/p}\left(\sum_{i,j} q_i^q |\langle u_i|v_j\rangle|^2\right) && \text{(the Hölder's inequality)} \\
&= \left(\sum_i p_i^p \sum_j |\langle u_i|v_j\rangle|^2\right)^{1/p}\left(\sum_j q_j^q \sum_i |\langle u_i|v_j\rangle|^2\right) \\
&= \left(\sum_i p_i^p\right)^{1/p}\left(\sum_j q_j^q\right)^{1/q} && \text{(the Pythagorean Theorem)} \\
&= \|A\|_p \|B\|_q .
\end{aligned}$$

□

Finally, Theorem 2.1 is a corollary to Theorem 3.4. Consider the case when $\rho_i = |\psi_i\rangle\langle\psi_i|$ for $i = 1, 2$ and $\langle\psi_1|\psi_2\rangle = \langle\psi_2|\psi_1\rangle = \cos\theta$. The best (average) success probability is $\frac{1}{2} + \frac{1}{2}\|\rho_1 - \rho_2\|_1$. Observe that $\rho_1 - \rho_2 = |\psi_1\rangle\langle\psi_1| - |\psi_2\rangle\langle\psi_2|$ has $(d-2)$ zero eigenvalues in the subspace perpendicular to both $|\psi_1\rangle$ and $|\psi_2\rangle$. The other two eigenvalues are associated to eigenvectors of the form $|\psi\rangle = c_1|\psi_1\rangle + c_2|\psi_2\rangle$. If $|\psi\rangle = c_1|\psi_1\rangle + c_2|\psi_2\rangle$ is an eigenvector, then

$$(\rho_1 - \rho_2)|\psi\rangle = (|\psi_1\rangle\langle\psi_1| - |\psi_2\rangle\langle\psi_2|)(c_1|\psi_1\rangle + c_2|\psi_2\rangle) = (c_1 - c_2\cos\theta)|\psi_1\rangle + (c_1\cos\theta - c_2)|\psi_2\rangle .$$

So the associated eigenvalue λ satisfies $\lambda c_1 = c_1 - c_2\cos\theta$ and $\lambda c_2 = c_1\cos\theta - c_2$. In other words, the homogeneous system of linear equations

$$(\lambda - 1)c_1 + (\cos\theta)c_2 = 0, \quad (-\cos\theta)c_1 + (\lambda + 1)c_2 = 0.$$

has a non-trivial solution, and so

$$\det\begin{pmatrix} \lambda - 1 & \cos\theta \\ -\cos\theta & \lambda + 1 \end{pmatrix} = \lambda^2 - 1 + \cos^2\theta = \lambda^2 - \sin^2\theta = 0.$$

Therefore, the only two nonzero eigenvalues are $\pm\sin\theta$, hence the best success probability is $\frac{1}{2} + \frac{1}{4}\|\rho_1 - \rho_2\|_1 = \frac{1}{2} + \frac{1}{4}(\sin\theta + \sin\theta) = \frac{1}{2} + \frac{1}{2}\sin\theta$.

References

[Die88] Dennis Dieks. Overlap and distinguishability of quantum states. *Physics Letters A*, 126(5):303–306, 1988.

[Hel69] Carl W Helstrom. Quantum detection and estimation theory. *Journal of Statistical Physics*, 1(2):231–252, 1969.

[Hol73] Alexander S Holevo. Statistical decision theory for quantum systems. *Journal of Multivariate Analysis*, 3(4):337–394, 1973.

[Iva87] Igor D Ivanovic. How to differentiate between non-orthogonal states. *Physics Letters A*, 123(6):257–259, 1987.

[Per88] Asher Peres. How to differentiate between non-orthogonal states. *Physics Letters A*, 128(1):19, 1988.