PHYS 7895: Quantum Information Theory Fall 2015

Lecture 16 — October 19, 2015

Prof. Mark M. Wilde *Scribe: Mark M. Wilde*

1 Overview

In the previous lecture, we continued developing the trace distance in more detail, defined the fidelity, and proved several properties of the fidelity.

In this lecture, we show how a measurement achieves the fidelity, and we show to relate the trace distance to the fidelity. Finally, we discuss the gentle measurement lemma.

2 A Measurement Achieves the Fidelity

There is a classical notion of fidelity for probability distributions, which is sometimes called the classical fidelity or Bhattacharyya overlap. It is defined as follows:

Definition 1 (Classical Fidelity)**.** *Let p and q be probability distributions defined over a finite alphabet $\mathcal{X}$. The classical fidelity $F(p,q)$ is defined as follows:*

$$F(p,q) \equiv \left[\sum_{x\in\mathcal{X}} \sqrt{p(x)q(x)}\right]^2. \tag{1}$$

Exercise 2. *Verify that the classical fidelity is a special case of the quantum fidelity. That is, let p and q be probability distributions defined over a finite alphabet $\mathcal{X}$, and then place the entries of these distributions along the diagonal of commuting matrices ρ and σ, respectively. Show that $F(p,q) = F(\rho,\sigma)$.*

Now suppose that we have two density operators $\rho, \sigma \in \mathcal{D}(\mathcal{H})$, and suppose further that we perform a POVM $\{\Lambda_x\}$ on these states, leading to the following probability distributions:

$$p(x) = \text{Tr}\{\Lambda_x \rho\}, \qquad q(x) = \text{Tr}\{\Lambda_x \sigma\}. \tag{2}$$

We can then compute the classical fidelity of the distributions for the measurement outcomes by using the formula in (1), and this is a measure of distinguishability of the two quantum states, with respect to a particular measurement. From the monotonicity of the quantum fidelity with respect to quantum channels, it follows that the quantum fidelity $F(\rho,\sigma)$ never exceeds this classical fidelity

$$F(\rho,\sigma) \leq \left[\sum_{x\in\mathcal{X}} \sqrt{\text{Tr}\{\Lambda_x \rho\}\,\text{Tr}\{\Lambda_x \sigma\}}\right]^2. \tag{3}$$

In particular, this bound follows from the monotonicity of fidelity, where the channel here is understood to be a measurement channel of the form $\omega \to \sum_x \mathrm{Tr}\{\Lambda_x \omega\} |x\rangle\langle x|$ and then we apply the result of Exercise 2. What is perhaps surprising is that there always exists a measurement that saturates the bound above, leading to the following alternate characterization of fidelity:

Theorem 3 (Measurement Achieves Fidelity)**.** *Let $\rho, \sigma \in \mathcal{D}(\mathcal{H})$. Then*

$$F(\rho,\sigma) = \min_{\{\Lambda_x\}} \left[\sum_{x\in\mathcal{X}} \sqrt{\mathrm{Tr}\{\Lambda_x \rho\} \,\mathrm{Tr}\{\Lambda_x \sigma\}} \right]^2, \tag{4}$$

where the minimization is with respect to all POVMs.

Proof. As justified before the statement of the theorem, the bound in (3) holds for any POVM. So here we construct a specific POVM (known as the Fuchs-Caves measurement) that saturates the bound. First consider the case in which σ is positive definite (and thus invertible). Consider the following operator (known as an operator geometric mean of ρ and σ^{-1}):

$$M = \sigma^{-1/2} \left[\sigma^{1/2} \rho \sigma^{1/2} \right]^{1/2} \sigma^{-1/2}. \tag{5}$$

The operator M is positive semi-definite, and thus has a spectral decomposition:

$$M = \sum_y \lambda_y |y\rangle\langle y|, \tag{6}$$

with $\{\lambda_y\}$ a set of non-negative eigenvalues and $\{|y\rangle\}$ a corresponding set of eigenvectors.

We will prove that the optimal measurement in (4) is $\{|y\rangle\langle y|\}$. We begin by noting that a simple calculation gives

$$M\sigma M = \rho. \tag{7}$$

So now consider the classical fidelity of the measurement $\{|y\rangle\langle y|\}$:

$$\begin{aligned}
\sum_y \sqrt{\mathrm{Tr}\{|y\rangle\langle y|\rho\} \,\mathrm{Tr}\{|y\rangle\langle y|\sigma\}} &= \sum_y \sqrt{\langle y|\rho|y\rangle \ \langle y|\sigma|y\rangle} && (8)\\
&= \sum_y \sqrt{\langle y|M\sigma M|y\rangle \ \langle y|\sigma|y\rangle} && (9)\\
&= \sum_y \sqrt{\langle y|\lambda_y \sigma \lambda_y|y\rangle \ \langle y|\sigma|y\rangle} && (10)\\
&= \sum_y \lambda_y \langle y|\sigma|y\rangle && (11)
\end{aligned}$$

The second equality follows from (7). The third equality follows because $M|y\rangle = \lambda_y|y\rangle$. Continuing, the last line above is equal to

$$\mathrm{Tr}\left\{ \sum_y \lambda_y |y\rangle\langle y|\sigma \right\} = \mathrm{Tr}\left\{ M\sigma \right\} = \mathrm{Tr}\left\{ \left[\sigma^{1/2}\rho\sigma^{1/2} \right]^{1/2} \right\} = \sqrt{F}(\rho,\sigma). \tag{12}$$

The last equality follows from the fact that $\sqrt{F}(\rho,\sigma) = \|\sqrt{\rho}\sqrt{\sigma}\|_1$.

For the case in which σ is not invertible, we repeat the above analysis, replacing ρ with $\Pi_\sigma \rho \Pi_\sigma$, where Π_σ is the projection onto the support of σ. In this case, the geometric mean operator M has its support contained in the support of σ, and one can find a spectral decomposition of M as in (6) so that

$$\sqrt{F}(\Pi_\sigma \rho \Pi_\sigma, \sigma) = \sum_y \sqrt{\operatorname{Tr}\{|y\rangle\langle y|\Pi_\sigma \rho \Pi_\sigma\} \operatorname{Tr}\{|y\rangle\langle y|\sigma\}}. \tag{13}$$

Since the eigenvectors $\{|y\rangle\}$ do not necessarily span the whole space, we can add additional orthonormal vectors all orthogonal to those in $\{|y\rangle\}$, such that all of them taken together form a legitimate measurement. Since both $\Pi_\sigma \rho \Pi_\sigma$ and σ are orthogonal to all of the new vectors, the probabilities for these measurement outcomes are all equal to zero and thus they do not contribute anything to the sum in (13). Finally, we have that

$$F(\Pi_\sigma \rho \Pi_\sigma, \sigma) = F(\rho, \sigma) \tag{14}$$

because $\sigma^{1/2} = \Pi_\sigma \sigma^{1/2} = \sigma^{1/2}\Pi_\sigma$, so that

$$\sqrt{F}(\rho, \sigma) = \operatorname{Tr}\left\{\sqrt{\sigma^{1/2}\rho\sigma^{1/2}}\right\} = \operatorname{Tr}\left\{\sqrt{\sigma^{1/2}\Pi_\sigma \rho \Pi_\sigma \sigma^{1/2}}\right\} = \sqrt{F}(\Pi_\sigma \rho \Pi_\sigma, \sigma). \tag{15}$$

□

3 Relationships between Trace Distance and Fidelity

In quantum Shannon theory, we are interested in showing that a given quantum information-processing protocol approximates an ideal protocol. We might do so by showing that the quantum output of the ideal protocol, say ρ, is close to the quantum output of the actual protocol, say σ. For example, we may be able to show that the fidelity between ρ and σ is high:

$$F(\rho, \sigma) \geq 1 - \varepsilon, \tag{16}$$

where ε is a small, positive real number that determines how well ρ approximates σ according to the above fidelity criterion. Typically, in a quantum Shannon-theoretic argument, we will take a limit to show that it is possible to make ε as small as we would like. As the performance parameter ε becomes vanishingly small, we expect that ρ and σ are becoming approximately equal so that they are identically equal when ε vanishes in some limit.

We would naturally think that the trace distance should be small if the fidelity is high because the trace distance vanishes when the fidelity is one and vice versa. The next theorem makes this intuition precise by establishing several relationships between the trace distance and fidelity.

Theorem 4 (Relations between Fidelity and Trace Distance)**.** *The following bound applies to the trace distance and the fidelity between two quantum states $\rho, \sigma \in \mathcal{D}(\mathcal{H})$:*

$$1 - \sqrt{F(\rho, \sigma)} \leq \frac{1}{2}\left\|\rho - \sigma\right\|_1 \leq \sqrt{1 - F(\rho, \sigma)}. \tag{17}$$

Proof. We first show that there is an exact relationship between fidelity and trace distance for pure states. Let us pick two arbitrary pure states $|\psi\rangle, |\phi\rangle \in \mathcal{H}$. We can write the state $|\phi\rangle$ in terms of the state $|\psi\rangle$ and its orthogonal complement $|\psi^\perp\rangle$ as follows:

$$|\phi\rangle = \cos(\theta)|\psi\rangle + e^{i\varphi}\sin(\theta)|\psi^\perp\rangle. \tag{18}$$

First, the fidelity between these two pure states is

$$F(\psi, \phi) = |\langle\phi|\psi\rangle|^2 = \cos^2(\theta). \tag{19}$$

Now let us determine the trace distance. The density operator $|\phi\rangle\langle\phi|$ is as follows:

$$\begin{aligned}
|\phi\rangle\langle\phi| &= \left(\cos(\theta)|\psi\rangle + e^{i\varphi}\sin(\theta)|\psi^{\perp}\rangle\right)\left(\cos(\theta)\langle\psi| + e^{-i\varphi}\sin(\theta)\langle\psi^{\perp}|\right) && (20)\\
&= \cos^2(\theta)|\psi\rangle\langle\psi| + e^{i\varphi}\sin(\theta)\cos(\theta)\,|\psi^{\perp}\rangle\langle\psi| \\
&\quad + e^{-i\varphi}\cos(\theta)\sin(\theta)\,|\psi\rangle\langle\psi^{\perp}| + \sin^2(\theta)|\psi^{\perp}\rangle\langle\psi^{\perp}|. && (21)
\end{aligned}$$

The matrix representation of the operator $|\psi\rangle\langle\psi| - |\phi\rangle\langle\phi|$ with respect to the basis $\left\{|\psi\rangle, |\psi^{\perp}\rangle\right\}$ is

$$\begin{bmatrix} 1-\cos^2(\theta) & -e^{-i\varphi}\sin(\theta)\cos(\theta) \\ -e^{i\varphi}\sin(\theta)\cos(\theta) & -\sin^2(\theta) \end{bmatrix}. \tag{22}$$

It is straightforward to show that the eigenvalues of the above matrix are $|\sin(\theta)|$ and $-|\sin(\theta)|$ and it then follows that the trace distance between $|\psi\rangle$ and $|\phi\rangle$ is the absolute sum of the eigenvalues:

$$\left\| |\psi\rangle\langle\psi| - |\phi\rangle\langle\phi| \right\|_1 = 2\,|\sin(\theta)|\,. \tag{23}$$

Consider the following trigonometric relationship:

$$\left(\frac{2\,|\sin(\theta)|}{2}\right)^2 = 1 - \cos^2(\theta). \tag{24}$$

It then holds that the fidelity and trace distance for pure states are related as follows:

$$\left(\frac{1}{2}\left\| |\psi\rangle\langle\psi| - |\phi\rangle\langle\phi| \right\|_1\right)^2 = 1 - F(\psi, \phi), \tag{25}$$

by plugging (19) into the right-hand side of (24) and (23) into the left-hand side of (24). Thus,

$$\frac{1}{2}\left\| |\psi\rangle\langle\psi| - |\phi\rangle\langle\phi| \right\|_1 = \sqrt{1 - F(\psi, \phi)}. \tag{26}$$

To prove the upper bound for mixed states ρ_A and σ_A, choose purifications $|\phi^{\rho}\rangle_{RA}$ and $|\phi^{\sigma}\rangle_{RA}$ of respective states ρ_A and σ_A such that

$$F(\rho_A, \sigma_A) = |\langle\phi^{\sigma}|\phi^{\rho}\rangle|^2 = F(\phi^{\rho}_{RA}, \phi^{\sigma}_{RA}). \tag{27}$$

(Recall that these purifications exist by Uhlmann's theorem.) Then

$$\begin{aligned}
\frac{1}{2}\left\|\rho_A - \sigma_A\right\|_1 &\leq \frac{1}{2}\left\|\phi^{\rho}_{RA} - \phi^{\sigma}_{RA}\right\|_1 && (28)\\
&= \sqrt{1 - F(\phi^{\rho}_{RA}, \phi^{\sigma}_{RA})} && (29)\\
&= \sqrt{1 - F(\rho_A, \sigma_A)}, && (30)
\end{aligned}$$

where the first inequality follows by the monotonicity of the trace distance under the discarding of systems.

To prove the lower bound for mixed states ρ and σ, recall Theorem 3. Also, recall that the trace distance between states ρ and σ is the maximum classical trace distance between two probability distributions resulting from a POVM $\{\Lambda_m\}$ acting on the states ρ and σ:

$$\|\rho-\sigma\|_1 = \max_{\{\Lambda_m\}} \sum_m |p_m - q_m| \,, \tag{31}$$

where

$$p_m \equiv \mathrm{Tr}\{\Lambda_m \rho\}\,, \qquad q_m \equiv \mathrm{Tr}\{\Lambda_m \sigma\}\,. \tag{32}$$

Furthermore, Theorem 3 states that the quantum fidelity is the minimum classical fidelity between two probability distributions p'_m and q'_m resulting from a measurement $\{\Gamma_m\}$ of the states ρ and σ:

$$F(\rho,\sigma) = \min_{\{\Gamma_m\}} \left(\sum_m \sqrt{p'_m q'_m}\right)^2 , \tag{33}$$

where

$$p'_m \equiv \mathrm{Tr}\{\Gamma_m \rho\}\,, \qquad q'_m \equiv \mathrm{Tr}\{\Gamma_m \sigma\}\,. \tag{34}$$

We return to the proof. Suppose that the POVM $\{\Gamma_m\}$ achieves the minimum Bhattacharya distance and results in probability distributions p'_m and q'_m, so that

$$F(\rho,\sigma) = \left(\sum_m \sqrt{p'_m q'_m}\right)^2 . \tag{35}$$

Consider that

$$\begin{aligned}
\sum_m \left(\sqrt{p'_m} - \sqrt{q'_m}\right)^2 &= \sum_m p'_m + q'_m - \sqrt{p'_m q'_m} &(36)\\
&= 2 - 2\sqrt{F(\rho,\sigma)}. &(37)
\end{aligned}$$

It also follows that

$$\begin{aligned}
\sum_m \left(\sqrt{p'_m} - \sqrt{q'_m}\right)^2 &\leq \sum_m \left|\sqrt{p'_m} - \sqrt{q'_m}\right| \left|\sqrt{p'_m} + \sqrt{q'_m}\right| &(38)\\
&= \sum_m |p'_m - q'_m| &(39)\\
&\leq \sum_m |p_m - q_m| &(40)\\
&= \|\rho-\sigma\|_1 \,. &(41)
\end{aligned}$$

The first inequality holds because $\left|\sqrt{p'_m} - \sqrt{q'_m}\right| \leq \left|\sqrt{p'_m} + \sqrt{q'_m}\right|$. The second inequality holds because the distributions p'_m and q'_m minimizing the Bhattacharya distance in general have Kolmogorov distance less than the distributions p_m and q_m that maximize the Kolmogorov distance. Thus, the following inequality results

$$2 - 2\sqrt{F(\rho,\sigma)} \leq \|\rho-\sigma\|_1 \,, \tag{42}$$

and the lower bound in the statement of the theorem follows. □

Theorem 4 allows us to complete our understanding of the extreme values of trace distance and fidelity. We have already argued that two states $\rho, \sigma \in \mathcal{D}(\mathcal{H})$ have trace distance equal to zero if and only if $\rho = \sigma$. Theorem 4 allows us to conclude that $F(\rho, \sigma) = 1$ if and only if $\rho = \sigma$. Similarly, we have argued already that $F(\rho, \sigma) = 0$ if and only if the support of ρ is orthogonal to that of σ. Theorem 4 allows us to conclude that $\|\rho - \sigma\|_1 = 2$ if and only if the support of ρ is orthogonal to that of σ.

The following two corollaries are simple consequences of Theorem 4.

Corollary 5. *Let $\rho, \sigma \in \mathcal{D}(\mathcal{H})$ and fix $\varepsilon \in [0, 1]$. Suppose that ρ is ε-close to σ in trace distance:*

$$\|\rho - \sigma\|_1 \leq \varepsilon. \tag{43}$$

Then the fidelity between ρ and σ is greater than $1 - \varepsilon$:

$$F(\rho, \sigma) \geq 1 - \varepsilon. \tag{44}$$

Corollary 6. *Let $\rho, \sigma \in \mathcal{D}(\mathcal{H})$ and fix $\varepsilon \in [0, 1]$. Suppose the fidelity between ρ and σ is greater than $1 - \varepsilon$:*

$$F(\rho, \sigma) \geq 1 - \varepsilon. \tag{45}$$

Then ρ is $2\sqrt{\varepsilon}$-close to σ in trace distance:

$$\|\rho - \sigma\|_1 \leq 2\sqrt{\varepsilon}. \tag{46}$$

Exercise 7. *Let $\rho, \sigma \in \mathcal{D}(\mathcal{H})$. Prove the following lower bound on the probability of error p_e in a quantum hypothesis test to distinguish ρ from σ:*

$$p_e \geq \frac{1}{2}\left(1 - \sqrt{1 - F(\rho, \sigma)}\right). \tag{47}$$

4 Gentle Measurement

The gentle measurement and gentle operator lemmas are particular applications of Theorem 4, and they concern the disturbance of quantum states. We generally expect in quantum theory that certain measurements might disturb the state which we are measuring. For example, suppose a qubit is in the state $|0\rangle$. A measurement along the X direction gives $+1$ and -1 with equal probability while drastically disturbing the state to become either $|+\rangle$ or $|-\rangle$, respectively. On the other hand, we might expect that the measurement does not disturb the state by very much if one outcome is highly likely. For example, suppose that we instead measure the qubit along the Z direction. The measurement returns $+1$ with unit probability while causing no disturbance to the qubit. The "gentle measurement lemma" below quantitatively addresses the disturbance of quantum states by demonstrating that a measurement with one outcome that is highly likely causes only a little disturbance to the quantum state that we measure (hence, the measurement is "gentle" or "tender").

Lemma 8 (Gentle Measurement)**.** *Consider a density operator ρ and a measurement operator Λ where $0 \leq \Lambda \leq I$. The measurement operator could be an element of a POVM. Suppose that the measurement operator Λ has a high probability of detecting state ρ:*

$$\operatorname{Tr}\{\Lambda\rho\} \geq 1 - \varepsilon, \tag{48}$$

where $1 \geq \varepsilon > 0$ (the probability of detection is high if and only if ε is close to zero). Then the post-measurement state

$$\rho' \equiv \frac{\sqrt{\Lambda}\rho\sqrt{\Lambda}}{\text{Tr}\left\{\Lambda\rho\right\}} \tag{49}$$

is $2\sqrt{\varepsilon}$-close to the original state ρ in trace distance:

$$\left\|\rho - \rho'\right\|_1 \leq 2\sqrt{\varepsilon}. \tag{50}$$

Thus, the measurement does not disturb the state ρ by much if ε is small.

Proof. Suppose first that ρ is a pure state $|\psi\rangle\langle\psi|$. The post-measurement state is then

$$\frac{\sqrt{\Lambda}|\psi\rangle\langle\psi|\sqrt{\Lambda}}{\langle\psi|\Lambda|\psi\rangle}. \tag{51}$$

The fidelity between the original state $|\psi\rangle$ and the post-measurement state above is as follows:

$$\begin{aligned}
\langle\psi|\left(\frac{\sqrt{\Lambda}|\psi\rangle\langle\psi|\sqrt{\Lambda}}{\langle\psi|\Lambda|\psi\rangle}\right)|\psi\rangle &= \frac{\left|\langle\psi|\sqrt{\Lambda}|\psi\rangle\right|^2}{\langle\psi|\Lambda|\psi\rangle} &(52)\\
&\geq \frac{\left|\langle\psi|\Lambda|\psi\rangle\right|^2}{\langle\psi|\Lambda|\psi\rangle} &(53)\\
&= \langle\psi|\Lambda|\psi\rangle &(54)\\
&\geq 1 - \varepsilon. &(55)
\end{aligned}$$

The first inequality follows because $\sqrt{\Lambda} \geq \Lambda$ when $\Lambda \leq I$. The second inequality follows from the hypothesis of the lemma. Now let us consider when we have mixed states ρ_A and ρ'_A. Suppose $|\psi\rangle_{RA}$ and $|\psi'\rangle_{RA}$ are respective purifications of ρ_A and ρ'_A, where

$$\left|\psi'\right\rangle_{RA} \equiv \frac{I_R \otimes \sqrt{\Lambda_A}|\psi\rangle_{RA}}{\sqrt{\langle\psi|I_R \otimes \Lambda_A|\psi\rangle_{RA}}}. \tag{56}$$

Then we can apply monotonicity of fidelity and the above result for pure states to show that

$$F(\rho_A, \rho'_A) \geq F(\psi_{RA}, \psi'_{RA}) \geq 1 - \varepsilon. \tag{57}$$

We finally obtain the bound on the trace distance $\|\rho_A - \rho'_A\|_1$ by exploiting Corollary 6. $\square$

The following lemma is a variation on the gentle measurement lemma that we sometimes exploit.

Lemma 9 (Gentle Operator)**.** *Consider a density operator ρ and a measurement operator Λ where $0 \leq \Lambda \leq I$. The measurement operator could be an element of a POVM. Suppose that the measurement operator Λ has a high probability of detecting state ρ:*

$$\text{Tr}\left\{\Lambda\rho\right\} \geq 1 - \varepsilon, \tag{58}$$

where $1 \geq \varepsilon > 0$ (the probability is high only if ε is close to zero). Then $\sqrt{\Lambda}\rho\sqrt{\Lambda}$ is $2\sqrt{\varepsilon}$-close to the original state ρ in trace distance:

$$\left\|\rho - \sqrt{\Lambda}\rho\sqrt{\Lambda}\right\|_1 \leq 2\sqrt{\varepsilon}. \tag{59}$$

Proof. Consider the following chain of inequalities:

$$\begin{aligned}
&\left\|\rho - \sqrt{\Lambda}\rho\sqrt{\Lambda}\right\|_1 \\
&= \left\|\left(I - \sqrt{\Lambda} + \sqrt{\Lambda}\right)\rho - \sqrt{\Lambda}\rho\sqrt{\Lambda}\right\|_1 &(60)\\
&\leq \left\|\left(I - \sqrt{\Lambda}\right)\rho\right\|_1 + \left\|\sqrt{\Lambda}\rho\left(I - \sqrt{\Lambda}\right)\right\|_1 &(61)\\
&= \mathrm{Tr}\left|\left(I - \sqrt{\Lambda}\right)\sqrt{\rho}\cdot\sqrt{\rho}\right| + \mathrm{Tr}\left|\sqrt{\Lambda}\sqrt{\rho}\cdot\sqrt{\rho}\left(I - \sqrt{\Lambda}\right)\right| &(62)\\
&\leq \sqrt{\mathrm{Tr}\left\{\left(I - \sqrt{\Lambda}\right)^2\rho\right\}\mathrm{Tr}\left\{\rho\right\}} + \sqrt{\mathrm{Tr}\left\{\Lambda\rho\right\}\mathrm{Tr}\left\{\rho\left(I - \sqrt{\Lambda}\right)^2\right\}} &(63)\\
&\leq \sqrt{\mathrm{Tr}\left\{\left(I - \Lambda\right)\rho\right\}} + \sqrt{\mathrm{Tr}\left\{\rho\left(I - \Lambda\right)\right\}} &(64)\\
&= 2\sqrt{\mathrm{Tr}\left\{\left(I - \Lambda\right)\rho\right\}} &(65)\\
&\leq 2\sqrt{\varepsilon}. &(66)
\end{aligned}$$

The first inequality is the triangle inequality. The second equality follows from the definition of the trace norm and the fact that ρ is a positive semi-definite operator. The second inequality is essentially the Cauchy–Schwarz inequality for the Hilbert–Schmidt inner product. Let U be a unitary such that $\left|A^\dagger B\right| = UA^\dagger B$ (recall the polar decomposition), where A and B are square operators acting on the same Hilbert space. Applying Cauchy–Schwarz gives

$$\begin{aligned}
\mathrm{Tr}\left\{\left|A^\dagger B\right|\right\} &= \mathrm{Tr}\left\{UA^\dagger B\right\} &(67)\\
&\leq \sqrt{\mathrm{Tr}\left\{UA^\dagger AU^\dagger\right\}\mathrm{Tr}\left\{B^\dagger B\right\}} &(68)\\
&= \sqrt{\mathrm{Tr}\left\{A^\dagger A\right\}\mathrm{Tr}\left\{B^\dagger B\right\}}. &(69)
\end{aligned}$$

The third inequality follows because $(1-\sqrt{x})^2 \leq 1-x$ for $0 \leq x \leq 1$, $\mathrm{Tr}\{\rho\} = 1$, and $\mathrm{Tr}\{\Lambda\rho\} \leq 1$. The final inequality follows from applying (58) and because the square root function is monotone increasing. □

Exercise 10. *Show that the gentle operator lemma holds for subnormalized positive semi-definite operators ρ (operators ρ such that $\mathrm{Tr}\{\rho\} \leq 1$).*

Below is another variation on the gentle measurement lemma that applies to ensembles of quantum states.

Lemma 11 (Gentle Measurement for Ensembles). *Let $\{p_X(x), \rho_x\}$ be an ensemble with average density operator $\overline{\rho} \equiv \sum_x p_X(x)\rho_x$. Given a positive semi-definite operator Λ with $\Lambda \leq I$ and $\mathrm{Tr}\{\overline{\rho}\Lambda\} \geq 1-\varepsilon$ where $\varepsilon \leq 1$, then*

$$\sum_x p_X(x)\left\|\rho_x - \sqrt{\Lambda}\rho_x\sqrt{\Lambda}\right\|_1 \leq 2\sqrt{\varepsilon}. \tag{70}$$

Proof. We can apply the same steps in the proof of the gentle operator lemma to get the following inequality, holding for all x:

$$\left\|\rho_x - \sqrt{\Lambda}\rho_x\sqrt{\Lambda}\right\|_1^2 \leq 4\left(1 - \mathrm{Tr}\left\{\Lambda\rho_x\right\}\right). \tag{71}$$

Taking the expectation over both sides produces the following inequality:

$$\sum_x p_X(x) \left\| \rho_x - \sqrt{\Lambda}\rho_x\sqrt{\Lambda} \right\|_1^2 \leq 4\left(1 - \text{Tr}\left\{\Lambda\rho\right\}\right) \tag{72}$$

$$\leq 4\varepsilon. \tag{73}$$

Taking the square root of the above inequality gives the following one:

$$\sqrt{\sum_x p_X(x) \left\| \rho_x - \sqrt{\Lambda}\rho_x\sqrt{\Lambda} \right\|_1^2} \leq 2\sqrt{\varepsilon}. \tag{74}$$

Concavity of the square root then implies the result:

$$\sum_x p_X(x) \sqrt{\left\| \rho_x - \sqrt{\Lambda}\rho_x\sqrt{\Lambda} \right\|_1^2} \leq 2\sqrt{\varepsilon}. \tag{75}$$

□

5 Fidelity of a Quantum Channel

It is useful to have measures that determine how well a quantum channel $\mathcal{N}$ preserves quantum information. We developed static distance measures, such as the trace distance and the fidelity, in previous sections. We would now like to exploit those measures in order to define dynamic measures.

A "first guess" measure of this sort is the minimum fidelity $F_{\min}(\mathcal{N})$, where

$$F_{\min}(\mathcal{N}) \equiv \min_{|\psi\rangle} F(\psi, \mathcal{N}(\psi)). \tag{76}$$

This measure seems like it may be a good one because we generally do not know the state that Alice inputs to a noisy channel before transmitting to Bob.

It may seem somewhat strange that we chose to minimize over pure states in the definition of the minimum fidelity. Are not mixed states the most general states that occur in the quantum theory? It turns out that joint concavity of the root fidelity and monotonicity of the square function implies that we do not have to consider mixed states for the minimum fidelity. Consider the following sequence of inequalities:

$$\sqrt{F}\left(\rho, \mathcal{N}(\rho)\right) = \sqrt{F}\left(\sum_x p_X(x)|x\rangle\langle x|, \mathcal{N}\left(\sum_x p_X(x)|x\rangle\langle x|\right)\right) \tag{77}$$

$$= \sqrt{F}\left(\sum_x p_X(x)|x\rangle\langle x|, \sum_x p_X(x)\mathcal{N}(|x\rangle\langle x|\right)) \tag{78}$$

$$\geq \sum_x p_X(x)\sqrt{F}(|x\rangle\langle x|, \mathcal{N}(|x\rangle\langle x|)) \tag{79}$$

$$\geq \sqrt{F}(|x_{\min}\rangle\langle x_{\min}|, \mathcal{N}(|x_{\min}\rangle\langle x_{\min}|)). \tag{80}$$

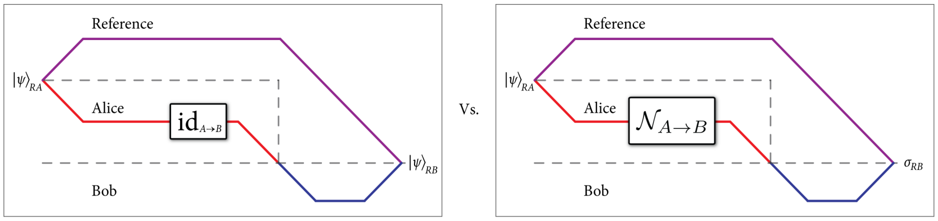

Figure 1: The entanglement fidelity compares the output of the ideal scenario (depicted on the left) and the output of the noisy scenario (depicted on the right).

The first equality follows by expanding the density operator ρ with a spectral decomposition. The second equality follows from linearity of the quantum operation $\mathcal{N}$. The first inequality follows from joint concavity of the root fidelity, and the last inequality follows because there exists some pure state $|x_{\min}\rangle$ (one of the eigenstates of ρ) with fidelity never larger than the expected fidelity in the previous line.

5.1 Expected Fidelity of a Quantum Channel

In general, the minimum fidelity is less useful than other measures of quantum information preservation over a quantum channel. The difficulty with the minimum fidelity is that it requires an optimization over the potentially large space of input states. Since it could be somewhat difficult to manipulate and compute in general, we introduce other ways to determine the performance of a quantum channel.

We can simplify our notion of fidelity by instead restricting the states that Alice sends and averaging the fidelity over this set of states. That is, suppose that Alice is transmitting states from an ensemble $\{p_X(x), \rho_x\}$ and we would like to determine how well a quantum channel $\mathcal{N}$ is preserving this source of quantum information. Sending a particular state ρ_x through a quantum channel $\mathcal{N}$ produces the state $\mathcal{N}(\rho_x)$. The fidelity between the transmitted state ρ_x and the received state $\mathcal{N}(\rho_x)$ is $F(\rho_x, \mathcal{N}(\rho_x))$ as defined before. We define the *expected fidelity* of the ensemble as follows:

$$\overline{F}(\mathcal{N}) \equiv \mathbb{E}_X \left[F(\rho_X, \mathcal{N}(\rho_X))\right] \tag{81}$$

$$= \sum_x p_X(x) F(\rho_x, \mathcal{N}(\rho_x)). \tag{82}$$

The expected fidelity indicates how well Alice is able to transmit the ensemble on average to Bob. It again lies between zero and one, just as the usual fidelity does.

5.2 Entanglement Fidelity

We now consider a different measure of the ability of a quantum channel to preserve quantum information. Suppose that Alice would like to transmit a quantum state with density operator ρ_A. It admits a purification $|\psi\rangle_{RA}$ to a reference system R. Sending the A system of $|\psi\rangle_{RA}$ through the identity channel id_A gives back $|\psi\rangle_{RA}$. Sending the A system of $|\psi\rangle_{RA}$ through a quantum channel $\mathcal{N} : \mathcal{L}(\mathcal{H}_A) \to \mathcal{L}(\mathcal{H}_A)$ gives the state $\sigma_{RA} \equiv (\mathrm{id}_R \otimes \mathcal{N}_A)(\psi_{RA})$. The entanglement fidelity is defined as follows:

Definition 12 (Entanglement Fidelity)**.** *For ρ, $\mathcal{N}$, σ, and $|\psi\rangle$ as defined above, the entanglement fidelity is given by $F_e(\rho,\mathcal{N}) \equiv \langle\psi|\sigma|\psi\rangle$.*

It is a measure of how well the quantum channel $\mathcal{N}$ preserves the entanglement with another system. Figure 1 visually depicts the two states that the entanglement fidelity compares.

The entanglement fidelity is invariant with respect to which purification of the input that we pick. This follows simply because all purifications are related by an isometry acting on the purifying system. That is, let $|\psi\rangle_{R_1A}$ be one purification of ρ_A, let $|\varphi\rangle_{R_2A}$ be a different one, and let $U_{R_1\to R_2}$ be an isometry that relates them via $|\varphi\rangle_{R_2A} = U_{R_1\to R_2}\,|\psi\rangle_{R_1A}$. Then

$$\begin{aligned}
&\langle\varphi|_{R_2A}\left(\mathrm{id}_{R_2}\otimes\mathcal{N}_A\right)\left(\varphi_{R_2A}\right)|\varphi\rangle_{R_2A} \\
&= \langle\psi|_{R_1A}U^\dagger_{R_1\to R_2}\left[\left(\mathrm{id}_{R_2}\otimes\mathcal{N}_A\right)\left(U_{R_1\to R_2}\psi_{R_1A}U^\dagger_{R_1\to R_2}\right)\right]U_{R_1\to R_2}|\psi\rangle_{R_1A} && (83)\\
&= \langle\psi|_{R_1A}U^\dagger_{R_1\to R_2}U_{R_1\to R_2}\left[\left(\mathrm{id}_{R_2}\otimes\mathcal{N}_A\right)\left(\psi_{R_1A}\right)\right]U^\dagger_{R_1\to R_2}U_{R_1\to R_2}\,|\psi\rangle_{R_1A} && (84)\\
&= \langle\psi|_{R_1A}\left[\left(\mathrm{id}_{R_1}\otimes\mathcal{N}_A\right)\left(\psi_{R_1A}\right)\right]|\psi\rangle_{R_1A}, && (85)
\end{aligned}$$

where the second equality follows because the isometry commutes with the identity map id_{R_2} and the last follows because $U_{R_1\to R_2}$ is an isometry so that $U^\dagger_{R_1\to R_2}U_{R_1\to R_2} = I_{R_1}$.

One of the benefits of considering the task of entanglement preservation is that it implies the task of quantum communication. That is, if Alice can devise a protocol that preserves the entanglement with another system, then this same protocol will also be able to preserve quantum information that she transmits.

The following theorem gives a simple way to represent the entanglement fidelity in terms of the Kraus operators of a given noisy quantum channel.

Theorem 13. *Let $\rho_A \in \mathcal{D}(\mathcal{H}_A)$ and let $\mathcal{N} : \mathcal{L}(\mathcal{H}_A) \to \mathcal{L}(\mathcal{H}_A)$ be a quantum channel. Suppose that $\{K^m\}$ is a set of Kraus operators for $\mathcal{N}$. Then the entanglement fidelity $F_e(\rho,\mathcal{N})$ is equal to the following expression:*

$$F_e(\rho,\mathcal{N}) = \sum_m \left|\mathrm{Tr}\left\{\rho_A K^m\right\}\right|^2. \tag{86}$$

Proof. Given that the entanglement fidelity is invariant with respect to the choice of purification, we can simply use the canonical purification $|\psi\rangle_{RA}$ of ρ_A, i.e.,

$$|\psi\rangle_{RA} = \left(I_R \otimes \sqrt{\rho_A}\right)|\Gamma\rangle_{RA}, \tag{87}$$

where $|\Gamma\rangle_{RA}$ is the unnormalized maximally entangled vector. We then find that

$$\begin{aligned}
&\langle\psi|_{RA}\left(\mathrm{id}_R\otimes\mathcal{N}_A\right)\left(\psi_{RA}\right)|\psi\rangle_{RA} \\
&= \langle\psi|_{RA}\sum_m \left(I_R\otimes K^m_A\right)|\psi\rangle\langle\psi|_{RA}\left(I_R\otimes (K^m_A)^\dagger\right)|\psi\rangle_{RA} && (88)\\
&= \sum_m \langle\psi|_{RA}\left(I_R\otimes K^m_A\right)|\psi\rangle\langle\psi|_{RA}\left(I_R\otimes (K^m_A)^\dagger\right)|\psi\rangle_{RA} && (89)\\
&= \sum_m \left|\langle\psi|_{RA}\left(I_R\otimes K^m_A\right)|\psi\rangle\right|^2. && (90)
\end{aligned}$$

Then consider that

$$
\begin{aligned}
&\langle\psi|_{RA} (I_R \otimes K_A^m) |\psi\rangle_{RA} \\
&= \langle\Gamma|_{RA} (I_R \otimes \sqrt{\rho_A}) (I_R \otimes K_A^m) (I_R \otimes \sqrt{\rho_A}) |\Gamma\rangle_{RA} &(91)\\
&= \langle\Gamma|_{RA} (I_R \otimes \sqrt{\rho_A} K_A^m \sqrt{\rho_A}) |\Gamma\rangle_{RA} &(92)\\
&= \mathrm{Tr}\left\{\sqrt{\rho_A} K_A^m \sqrt{\rho_A}\right\} &(93)\\
&= \mathrm{Tr}\left\{\rho_A K_A^m\right\}. &(94)
\end{aligned}
$$

So we conclude that

$$\langle\psi|_{RA} (\mathrm{id}_R \otimes \mathcal{N}_A) (\psi_{RA}) |\psi\rangle = \sum_m |\mathrm{Tr}\{\rho_A K^m\}|^2. \tag{95}$$

□

Exercise 14. *Let $\rho_1, \rho_2 \in \mathcal{D}(\mathcal{H})$ and let $\mathcal{N} : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H})$ be a quantum channel. Fix $\lambda \in [0,1]$. Show that the entanglement fidelity is convex in the input state:*

$$F_e(\lambda\rho_1 + (1-\lambda)\rho_2, \mathcal{N}) \leq \lambda F_e(\rho_1, \mathcal{N}) + (1-\lambda) F_e(\rho_2, \mathcal{N}). \tag{96}$$

(Hint: The result of Theorem 13 is useful here.)

Exercise 15. *Prove that the entanglement fidelity does not depend upon the particular choice of Kraus operators for a given channel. (Hint: Recall that there always exists an isometry that relates two different Kraus representations of a quantum channel, i.e., for a set $\{K^m\}$ of Kraus operators and another set $\{L^n\}$, we have that*

$$K^m = \sum_n u_{mn} L^n, \tag{97}$$

where u_{mn} are the entries of a unitary matrix.)

5.3 Relationship between Expected Fidelity and Entanglement Fidelity

The entanglement fidelity and the expected fidelity provide seemingly different methods for quantifying the ability of a noisy quantum channel to preserve quantum information. Is there any way that we can show how they are related?

It turns out that they are indeed related. First, consider that the entanglement fidelity is a lower bound on the channel's fidelity for preserving the state ρ:

$$F_e(\rho, \mathcal{N}) \leq F(\rho, \mathcal{N}(\rho)). \tag{98}$$

The above result follows simply from the monotonicity of fidelity under partial trace. We can show that the entanglement fidelity is always less than the expected fidelity in (81) by combining convexity of entanglement fidelity (Exercise 14) and the bound in (98):

$$
\begin{aligned}
F_e\left(\sum_x p_X(x)\rho_x, \mathcal{N}\right) &\leq \sum_x p_X(x) F_e(\rho_x, \mathcal{N}) &(99)\\
&\leq \sum_x p_X(x) F(\rho_x, \mathcal{N}(\rho_x)) &(100)\\
&= \overline{F}(\mathcal{N}). &(101)
\end{aligned}
$$

Thus, any channel that preserves entanglement with some reference system preserves the expected fidelity of an ensemble. In most cases, we only consider the entanglement fidelity as the defining measure of performance of a noisy quantum channel.