Quantum information theory (MAT4430) **Spring 2021**

Lecture 8: The Fidelity

Lecturer: Alexander Müller-Hermes

The goal of this lecture is to introduce the fidelity, which is one of the most important distance measures in quantum information theory. We will then express this quantity as convex optimization problems in two different ways. By doing so, it will be easy to prove the data-processing inequality of the fidelity, i.e., the inequality

$$F(T(\rho), T(\sigma)) \geqslant F(\rho, \sigma),$$

for all positive and trace-preserving maps $T : B(\mathcal{H}_A) \to B(\mathcal{H}_B)$.

1 Definition and basic properties

Let $\mathcal{H}$ denote a complex Euclidean space. For quantum states $\rho, \sigma \in D(\mathcal{H})$, we define the *fidelity* by

$$F(\rho, \sigma) = \|\sqrt{\rho}\sqrt{\sigma}\|_1,$$

where $\sqrt{\cdot}$ denotes the unique positive square root and $\|\cdot\|_1$ is the trace norm (i.e., the matrix 1-norm). The fidelity measures the distance between the two quantum states, and (using the definition of the trace norm) it can be written as

$$F(\rho, \sigma) = \text{Tr}\left[\left(\rho^{\frac{1}{2}}\sigma\rho^{\frac{1}{2}}\right)^{\frac{1}{2}}\right].$$

Note that it is not necessary to have normalized quantum states in the definition of the fidelity. In general, we define the fidelity

$$F(A, B) = \|\sqrt{A}\sqrt{B}\|_1 = \text{Tr}\left[\left(A^{\frac{1}{2}}BA^{\frac{1}{2}}\right)^{\frac{1}{2}}\right],$$

for any positive operators $A, B \in B(\mathcal{H})^+$. In the exercises we have shown the following elementary properties.

Lemma 1.1 (Elementary properties of the fidelity)**.** *Let $\mathcal{H}$ denote a complex Euclidean space and consider quantum states $\rho, \sigma \in D(\mathcal{H})$. We have the following properties:*

1. *We have $F(\rho, \sigma) = F(\sigma, \rho)$.*
2. *We have $F(\rho, \sigma) \geqslant 0$ with equality if and only if $\rho\sigma = 0$.*
3. *We have $F(\rho, \sigma) \leqslant 1$ with equality if and only if $\rho = \sigma$.*
4. *We have $F(V\rho V^\dagger, V\sigma V^\dagger) = F(\rho, \sigma)$ for any isometry $V : \mathcal{H} \to \mathcal{H}'$ into another complex Euclidean space $\mathcal{H}'$.*
5. *We have $F(|\psi\rangle\langle\psi|, \sigma) = \sqrt{\langle\psi|\sigma|\psi\rangle}$ for any pure quantum state $|\psi\rangle\langle\psi| \in \text{Proj}(\mathcal{H})$.*
6. *We have $F(\rho \otimes \tau, \sigma \otimes \tau) = F(\rho, \sigma)$ for every quantum state $\tau \in D(\mathcal{H}')$.*

Besides these properties, we note that the fidelity is a continuous function in both its inputs, since the operator square root and the 1-norm are continuous.

1.1 Expressing the fidelity as optimization problems

Theorem 1.2. *Let $\mathcal{H}$ denote a complex Euclidean space. For any $A, B \in B(\mathcal{H})^+$ we have*

$$F(A,B) = \max\{|\operatorname{Tr}[X]| \ : \ X \in B(H), \begin{pmatrix} A & X^\dagger \\ X & B \end{pmatrix} \in B(\mathcal{H}\oplus\mathcal{H})^+\}.$$

Proof. By a lemma from the last lecture we have

$$\begin{pmatrix} A & X^\dagger \\ X & B \end{pmatrix} \in B(\mathcal{H}\oplus\mathcal{H})^+,$$

if and only if there exists a $K \in B(H)$ satisfying $\|K\|_\infty \leqslant 1$ such that $X = B^{\frac{1}{2}}KA^{\frac{1}{2}}$. Therefore, we can write

$$\begin{aligned} &\sup\{|\operatorname{Tr}[X]| \ : \ X \in B(H), \begin{pmatrix} A & X^\dagger \\ X & B \end{pmatrix} \in B(\mathcal{H}\oplus\mathcal{H})^+\} \\ &= \sup\{|\operatorname{Tr}\left[B^{\frac{1}{2}}KA^{\frac{1}{2}}\right]| \ : \ K \in B(H), \|K\|_\infty \leqslant 1\} \\ &= \max\{|\operatorname{Tr}\left[UA^{\frac{1}{2}}B^{\frac{1}{2}}\right]| \ : \ U \in \mathcal{U}(H)\}, \end{aligned}$$

where we used, in the last step, that the unitaries are the extreme points of the unit ball of the operator norm $\|\cdot\|_\infty$. Using Hölder's inequality it is easy to see that

$$|\operatorname{Tr}\left[UA^{\frac{1}{2}}B^{\frac{1}{2}}\right]| \leqslant \|A^{\frac{1}{2}}B^{\frac{1}{2}}\|_1 = F(A,B).$$

On the other hand, if $A^{\frac{1}{2}}B^{\frac{1}{2}} = VSW$ is the singular value decomposition with unitaries $V, W \in \mathcal{U}(\mathcal{H})$, then the unitary $U = W^\dagger V^\dagger$ satisfies

$$|\operatorname{Tr}\left[UA^{\frac{1}{2}}B^{\frac{1}{2}}\right]| = \|A^{\frac{1}{2}}B^{\frac{1}{2}}\|_1 = F(A,B).$$

Together, these two statements finish the proof. □

As a simple corollary, we can rewrite this optimization problem slightly:

Corollary 1.3. *Let $\mathcal{H}$ denote a complex Euclidean space. For $A, B \in B(\mathcal{H})^+$ we have*

$$F(A,B) = \max\{Re\left(\operatorname{Tr}[X]\right) \ : \ X \in B(H), \begin{pmatrix} A & X^\dagger \\ X & B \end{pmatrix} \in B(\mathcal{H}\oplus\mathcal{H})^+\}.$$

Proof. Clearly, the right-hand-side is less than the fidelity by Theorem 1.2. To show that they are equal, consider $X \in B(H)$ such that

$$\begin{pmatrix} A & X^\dagger \\ X & B \end{pmatrix} \in B(\mathcal{H}\oplus\mathcal{H})^+,$$

and $|\operatorname{Tr}[X]| = F(A,B)$. Then, for any $\alpha \in \mathbb{R}$ we have that

$$\begin{pmatrix} A & e^{-i\alpha}X^\dagger \\ e^{i\alpha}X & B \end{pmatrix} = \begin{pmatrix} \mathbb{1}_\mathcal{H} & 0 \\ 0 & e^{i\alpha}\mathbb{1}_\mathcal{H} \end{pmatrix} \begin{pmatrix} A & X^\dagger \\ X & B \end{pmatrix} \begin{pmatrix} \mathbb{1}_\mathcal{H} & 0 \\ 0 & e^{-i\alpha}\mathbb{1}_\mathcal{H} \end{pmatrix} \in B(\mathcal{H}\oplus\mathcal{H})^+.$$

Choosing $\alpha \in \mathbb{R}$ such that

$$\operatorname{Re}\left(\operatorname{Tr}\left[e^{i\alpha}X\right]\right) = |\operatorname{Tr}[X]|,$$

finishes the proof. □

Theorem 1.4. *Let $\mathcal{H}$ denote a complex Euclidean space. For $A, B \in B(\mathcal{H})^+$ we have*

$$F(A,B) = \inf\{\frac{1}{2}\langle A, Y\rangle_{HS} + \frac{1}{2}\langle B, Y^{-1}\rangle_{HS} \;:\; Y \in B(\mathcal{H})^{++}\}.$$

Proof. Note that

$$\begin{pmatrix} Y & -\mathbb{1}_{\mathcal{H}} \\ -\mathbb{1}_{\mathcal{H}} & Y^{-1} \end{pmatrix} \in B(\mathcal{H} \oplus \mathcal{H})^+$$

for any $Y \in B(\mathcal{H})^{++}$. If

$$\begin{pmatrix} A & X^\dagger \\ X & B \end{pmatrix} \in B(\mathcal{H} \oplus \mathcal{H})^+,$$

for some $X \in B(\mathcal{H})$, then we have

$$0 \leqslant \mathrm{Tr}\left[\begin{pmatrix} A & X^\dagger \\ X & B \end{pmatrix}\begin{pmatrix} Y & -\mathbb{1}_{\mathcal{H}} \\ -\mathbb{1}_{\mathcal{H}} & Y^{-1} \end{pmatrix}\right] = \langle A, Y\rangle_{HS} + \langle B, Y^{-1}\rangle_{HS} - \mathrm{Tr}\left[X\right] - \mathrm{Tr}\left[X^\dagger\right],$$

and, by Corollary 1.3, we conclude that

$$\inf\{\frac{1}{2}\langle A, Y\rangle_{HS} + \frac{1}{2}\langle B, Y^{-1}\rangle_{HS} \;:\; Y \in B(\mathcal{H})^{++}\} \geqslant F(A,B),$$

for all $A, B \in B(\mathcal{H})^+$. To show that the infimum coincides with the fidelity, we start with the case where $A, B \in B(\mathcal{H})^{++}$ are invertible. Then, we can define

$$Y = A^{-\frac{1}{2}}\left(A^{\frac{1}{2}}BA^{\frac{1}{2}}\right)^{\frac{1}{2}}A^{-\frac{1}{2}} \in B(\mathcal{H})^{++}$$

and check that

$$\langle A, Y\rangle_{HS} = \mathrm{Tr}\left[AY\right] = \mathrm{Tr}\left[\left(A^{\frac{1}{2}}BA^{\frac{1}{2}}\right)^{\frac{1}{2}}\right] = F(A,B),$$

and

$$\langle B, Y^{-1}\rangle_{HS} = \mathrm{Tr}\left[BY^{-1}\right] = \mathrm{Tr}\left[BA^{\frac{1}{2}}\left(A^{\frac{1}{2}}BA^{\frac{1}{2}}\right)^{-\frac{1}{2}}A^{\frac{1}{2}}\right] = \mathrm{Tr}\left[\left(A^{\frac{1}{2}}BA^{\frac{1}{2}}\right)^{\frac{1}{2}}\right] = F(A,B).$$

We conclude that the infimum is attained and coincides with the fidelity in this case.

For general $A, B \in B(\mathcal{H})^+$ we will use a continuity argument. For every $\epsilon > 0$ and any $Y \in B(\mathcal{H})^{++}$ we have

$$\frac{1}{2}\langle A, Y\rangle_{HS} + \frac{1}{2}\langle B, Y^{-1}\rangle_{HS} \leqslant \frac{1}{2}\langle A + \epsilon\mathbb{1}_{\mathcal{H}}, Y\rangle_{HS} + \frac{1}{2}\langle B + \epsilon\mathbb{1}_{\mathcal{H}}, Y^{-1}\rangle_{HS}.$$

Taking the infimum over $Y \in B(\mathcal{H})^{++}$ on both sides and using the previous argument, we find that

$$\inf\{\frac{1}{2}\langle A, Y\rangle_{HS} + \frac{1}{2}\langle B, Y^{-1}\rangle_{HS} \;:\; Y \in B(\mathcal{H})^{++}\} \leqslant F(A + \epsilon\mathbb{1}_{\mathcal{H}}, B + \epsilon\mathbb{1}_{\mathcal{H}}).$$

Taking the limit $\epsilon \to 0$ finishes the proof. □

In the exercises you will be asked to prove the following corollary:

Corollary 1.5 (Alberti's theorem)**.** *Let $\mathcal{H}$ denote a complex Euclidean spaces and $\rho, \sigma \in D(\mathcal{H})$ two quantum states. We have*

$$F(\rho,\sigma)^2 = \inf\{\langle \rho, Y\rangle_{HS}\langle \sigma, Y^{-1}\rangle_{HS} \;:\; Y \in B(\mathcal{H})^{++}\}.$$

1.2 Data-processing inequality and joined concavity of the fidelity

Theorem 1.6 (Data-processing inequality for the fidelity)**.** *For any positive and trace-preserving map $P : B(\mathcal{H}_A) \to B(\mathcal{H}_B)$ we have*

$$F(P(\rho), P(\sigma)) \geqslant F(\rho, \sigma),$$

for all quantum states $\rho, \sigma \in D(\mathcal{H}_A)$.

Proof. Without loss of generality we may assume that $P^*(X)$ is invertible for any $X \in B(\mathcal{H}_A)^+$. If this is not the case, we can consider the positive maps $P_\epsilon : B(\mathcal{H}_A) \to B(\mathcal{H}_B)$ given by $P_\epsilon(X) = (1-\epsilon)P(X) + \epsilon \operatorname{Tr}[X] \mathbb{1}_{\mathcal{H}_B}$ for $0 < \epsilon \leqslant 1$ instead, and in the end of the proof take the limit $\epsilon \to 0$ using continuity of the fidelity.

By Theorem 1.4 we have

$$\begin{aligned} F(P(\rho), P(\sigma)) &= \inf\{\frac{1}{2}\langle P(\rho), Y\rangle_{HS} + \frac{1}{2}\langle P(\sigma), Y^{-1}\rangle_{HS} \ : \ Y \in B(\mathcal{H})^{++}\} \\ &= \inf\{\frac{1}{2}\langle \rho, P^*(Y)\rangle_{HS} + \frac{1}{2}\langle \sigma, P^*(Y^{-1})\rangle_{HS} \ : \ Y \in B(\mathcal{H})^{++}\} \\ &\geqslant \inf\{\frac{1}{2}\langle \rho, P^*(Y)\rangle_{HS} + \frac{1}{2}\langle \sigma, P^*(Y)^{-1}\rangle_{HS} \ : \ Y \in B(\mathcal{H})^{++}\} \\ &\geqslant F(\rho, \sigma). \end{aligned}$$

Here, we have used Choi's inequality from the exercises in the second-to-last step (note that P^* is unital whenever P is trace-preserving) and Theorem 1.4 for the final inequality. □

Next, we can prove the joined concavity of the fidelity:

Theorem 1.7 (Joined concavity)**.** *For quantum states $\rho_1, \rho_2, \sigma_1, \sigma_2 \in D(\mathcal{H})$ and $\lambda \in [0, 1]$ we have*

$$F((1-\lambda)\rho_1 + \lambda\rho_2, (1-\lambda)\sigma_1 + \lambda\sigma_2) \geqslant (1-\lambda)F(\rho_1, \sigma_1) + \lambda F(\rho_2, \sigma_2).$$

Proof. Consider the quantum states

$$\rho = (1-\lambda)\rho_1 \otimes |1\rangle\langle 1| + \lambda\rho_2 \otimes |2\rangle\langle 2| \in D(\mathcal{H} \otimes \mathbb{C}^2),$$

and

$$\sigma = (1-\lambda)\sigma_1 \otimes |1\rangle\langle 1| + \lambda\sigma_2 \otimes |2\rangle\langle 2| \in D(\mathcal{H} \otimes \mathbb{C}^2),$$

and the completely positive and trace-preserving partial trace map $\operatorname{Tr}_B = \operatorname{id}_{\mathcal{H}} \otimes \operatorname{Tr}$. By Theorem 1.6 we have

$$F(\operatorname{Tr}_B(\rho), \operatorname{Tr}_B(\sigma)) \geqslant F(\rho, \sigma).$$

Now, note that

$$F(\operatorname{Tr}_B(\rho), \operatorname{Tr}_B(\sigma)) = F((1-\lambda)\rho_1 + \lambda\rho_2, (1-\lambda)\sigma_1 + \lambda\sigma_2),$$

and

$$\begin{aligned} F(\rho, \sigma) = \|\sqrt{\rho}\sqrt{\sigma}\|_1 &= \|(1-\lambda)\sqrt{\rho_1}\sqrt{\sigma_1} \otimes |1\rangle\langle 1| + \lambda\sqrt{\rho_2}\sqrt{\sigma_2} \otimes |2\rangle\langle 2|\|_1 \\ &= (1-\lambda)\|\sqrt{\rho_1}\sqrt{\sigma_1}\|_1 + \lambda\|\sqrt{\rho_2}\sqrt{\sigma_2}\|_1 \\ &= (1-\lambda)F(\rho_1, \sigma_1) + \lambda F(\rho_2, \sigma_2). \end{aligned}$$

□

1.3 Uhlmann's theorem

The following lemma characterizes the fidelity between two quantum states by using their purifications:

Theorem 1.8 (Uhlmann's theorem). *Let $\mathcal{H}_A, \mathcal{H}_E$ denote complex Euclidean spaces and $\rho, \sigma \in D(\mathcal{H}_A)$ two quantum states. For any normalized vector $|\psi_{EA}\rangle \in \mathcal{H}_E \otimes \mathcal{H}_A$ satisfying $\mathrm{Tr}_E[|\psi_{EA}\rangle\langle\psi_{EA}|] = \rho$, we have*

$$F(\rho, \sigma) = \max\{|\langle\psi_{EA}|\phi_{EA}\rangle| \; : \; \mathrm{Tr}_E[|\phi_{EA}\rangle\langle\phi_{EA}|] = \sigma\}.$$

Proof. Given normalized vectors $|\psi_{EA}\rangle, |\phi_{EA}\rangle \in \mathcal{H}_E \otimes \mathcal{H}_A$ such that

$$\mathrm{Tr}_E[|\psi_{EA}\rangle\langle\psi_{EA}|] = \rho, \quad \text{and} \quad \mathrm{Tr}_E[|\phi_{EA}\rangle\langle\phi_{EA}|] = \sigma,$$

we can use the data-processing inequality from Theorem 1.6 (for $P = \mathrm{Tr}_E$) to conclude

$$|\langle\psi_{EA}|\phi_{EA}\rangle| = F(|\psi_{EA}\rangle\langle\psi_{EA}|, |\phi_{EA}\rangle\langle\phi_{EA}|) \leqslant F(\rho, \sigma).$$

To show that it is attained, consider a unitary $U \in \mathcal{U}(\mathcal{H}_A)$ such that

$$F(\rho, \sigma) = \mathrm{Tr}\left[U\sqrt{\rho}\sqrt{\sigma}\right].$$

Writing $|\psi_{EA}\rangle = \mathrm{vec}(X)$ for some $X \in B(\mathcal{H}_E, \mathcal{H}_A)$ we find that $\rho = XX^\dagger$. Using the singular value decomposition $X = VSW$ we find that

$$\begin{aligned} F(\rho, \sigma) &= \mathrm{Tr}\left[U\sqrt{\rho}\sqrt{\sigma}\right] \\ &= \mathrm{Tr}\left[U\sqrt{VS^2V^\dagger}\sqrt{\sigma}\right] \\ &= \mathrm{Tr}\left[UVSV^\dagger\sqrt{\sigma}\right] \\ &= \mathrm{Tr}\left[UVW(W^\dagger SV^\dagger)\sqrt{\sigma}\right] \\ &= \mathrm{Tr}\left[UVWX^\dagger\sqrt{\sigma}\right] \\ &= \langle X, \sqrt{\sigma}UVW\rangle_{HS} = \langle\psi_{EA}|\phi_{EA}\rangle, \end{aligned}$$

where we introduced

$$|\phi_{EA}\rangle := \mathrm{vec}\left(\sqrt{\sigma}UVW\right).$$

We can check that

$$\mathrm{Tr}_E[|\phi_{EA}\rangle\langle\phi_{EA}|] = (\sqrt{\sigma}UVW)(\sqrt{\sigma}UVW)^\dagger = \sigma,$$

and the proof is finished. □

1.4 Relationship between fidelity and trace distance

The following theorem is quite useful:

Theorem 1.9 (Fuchs-van de Graaf inequalities). *Let $\mathcal{H}$ denote a complex Euclidean space and $\rho, \sigma \in D(\mathcal{H})$ two quantum states. Then, we have*

$$1 - \frac{1}{2}\|\rho - \sigma\|_1 \leqslant F(\rho, \sigma) \leqslant \sqrt{1 - \frac{1}{4}\|\rho - \sigma\|_1^2}.$$

Proof. Clearly, the statement from the theorem is equivalent to

$$2 - 2F(\rho,\sigma) \leqslant \|\rho - \sigma\|_1 \leqslant 2\sqrt{1 - F(\rho,\sigma)^2}. \tag{1}$$

The proof will show this statement. For the second inequality of (1), we use Uhlmann's theorem to find normalized vectors $|\psi_{EA}\rangle, |\phi_{EA}\rangle \in \mathcal{H}_E \otimes \mathcal{H}_A$ satisfying $F(\rho,\sigma) = |\langle\psi_{EA}|\phi_{EA}\rangle|$. Then, we have

$$\|\rho - \sigma\|_1 = \| \operatorname{Tr}_E [|\psi_{EA}\rangle\langle\psi_{EA}|] - \operatorname{Tr}_E [|\phi_{EA}\rangle\langle\phi_{EA}|] \|_1 \leqslant \||\psi_{EA}\rangle\langle\psi_{EA}| - |\phi_{EA}\rangle\langle\phi_{EA}|\|_1,$$

by the Russo-Dye theorem. Using a result from the exercises, we conclude that the right-hand-side equals

$$\||\psi_{EA}\rangle\langle\psi_{EA}| - |\phi_{EA}\rangle\langle\phi_{EA}|\|_1 = 2\sqrt{1 - |\langle\psi_{EA}|\phi_{EA}\rangle|^2} = 2\sqrt{1 - F(\rho,\sigma)^2}.$$

For the other inequality in (1) let us restrict first to the case where ρ and σ are invertible. Recall, the operator

$$Y = \rho^{-\frac{1}{2}} \left(\rho^{\frac{1}{2}}\sigma\rho^{\frac{1}{2}}\right)^{\frac{1}{2}} \rho^{-\frac{1}{2}} \in B(\mathcal{H})^{++}$$

from the proof of Theorem 1.4 satisfying $\sigma = Y\rho Y$, and

$$\langle Y, \rho\rangle_{HS} = \langle Y^{-1}, \sigma\rangle_{HS} = F(\rho,\sigma).$$

By the spectral decomposition there are normalized vectors $|v_i\rangle \in \mathcal{H}$ and numbers $\lambda_i > 0$ for any $i \in \{1, \ldots \dim(\mathcal{H})\}$ such that

$$Y = \sum_{i=1}^{\dim(\mathcal{H})} \lambda_i |v_i\rangle\langle v_i|.$$

After defining probability distributions $p, q \in \mathcal{P}(\{1, \ldots, \dim(\mathcal{H})\})$ by

$$p_i = \langle v_i|\rho|v_i\rangle \quad \text{and} \quad q_i = \langle v_i|\sigma|v_i\rangle,$$

we compute

$$\begin{aligned}\sum_{i=1}^{\dim(\mathcal{H})} \sqrt{p_i}\sqrt{q_i} &= \sum_{i=1}^{\dim(\mathcal{H})} \sqrt{\langle v_i|\rho|v_i\rangle}\sqrt{\langle v_i|Y\rho Y|v_i\rangle} \\ &= \sum_{i=1}^{\dim(\mathcal{H})} \lambda_i \langle v_i|\rho|v_i\rangle = \langle Y, \rho\rangle_{HS} = F(\rho,\sigma).\end{aligned}$$

Finally, note that the linear map $T : B(\mathcal{H}) \to B(\mathcal{H})$ given by

$$T(X) = \sum_{i=1}^{\dim(\mathcal{H})} \langle v_i|X|v_i\rangle |v_i\rangle\langle v_i|,$$

is a quantum channel, and we have

$$\|\rho - \sigma\|_1 \geqslant \|T(\rho) - T(\sigma)\|_1 = \|p - q\|_1 \geqslant 2 - 2\sum_{i=1}^{\dim(\mathcal{H})} \sqrt{p_i}\sqrt{q_i} = 2 - 2F(\rho,\sigma),$$

where we used first the Russo-Dye theorem and in the last step an exercise from sheet 1. Finally, for general quantum states $\rho, \sigma \in D(\mathcal{H})$, we may consider the invertible quantum states $\rho_\epsilon, \sigma_\epsilon \in D(\mathcal{H})$ given by

$$\rho_\epsilon = (1 - \epsilon)\rho + \epsilon\frac{\mathbb{1}_{\mathcal{H}}}{\dim(\mathcal{H})},$$

and

$$\sigma_\epsilon = (1-\epsilon)\sigma + \epsilon \frac{\mathbb{1}_{\mathcal{H}}}{\dim(\mathcal{H})},$$

for $\epsilon \in (0,1)$. From the previous argument, we have

$$2 - 2F(\rho_\epsilon, \sigma_\epsilon) \leqslant \|\rho_\epsilon - \sigma_\epsilon\|_1$$

for any $\epsilon \in (0,1)$ and taking the limit $\epsilon \to 0$ finishes the proof.

□