CS 766/QIC 820 Theory of Quantum Information (Fall 2011)

Lecture 4: Purifications and fidelity

Throughout this lecture we will be discussing pairs of registers of the form (X, Y), and the relationships among the states of X, Y, and (X, Y).

The situation generalizes to collections of three or more registers, provided we are interested in bipartitions. For instance, if we have a collection of registers $(\mathsf{X}_1, \ldots, \mathsf{X}_n)$, and we wish to consider the state of a subset of these registers in relation to the state of the whole, we can effectively group the registers into two disjoint collections and relabel them as X and Y to apply the conclusions to be drawn. Other, multipartite relationships can become more complicated, such as relationships between states of $(\mathsf{X}_1, \mathsf{X}_2)$, $(\mathsf{X}_2, \mathsf{X}_3)$, and $(\mathsf{X}_1, \mathsf{X}_2, \mathsf{X}_3)$, but this is not the topic of this lecture.

4.1 Reductions, extensions, and purifications

Suppose that a pair of registers (X, Y) has the state $\rho \in \mathrm{D}(\mathcal{X} \otimes \mathcal{Y})$. The states of X and Y individually are then given by

$$\rho^{\mathsf{X}} = \mathrm{Tr}_{\mathcal{Y}}(\rho) \qquad \text{and} \qquad \rho^{\mathsf{Y}} = \mathrm{Tr}_{\mathcal{X}}(\rho).$$

You could regard this as a definition, but these are the only choices that are consistent with the interpretation that disregarding Y should have no influence on the outcomes of any measurements performed on X alone, and likewise for X and Y reversed. The states ρ^{X} and ρ^{Y} are sometimes called the *reduced states* of X and Y, or the *reductions* of ρ to X and Y.

We may also go in the other direction. If a state $\sigma \in \mathrm{D}(\mathcal{X})$ of X is given, we may consider the possible states $\rho \in \mathrm{D}(\mathcal{X} \otimes \mathcal{Y})$ that are consistent with σ on X, meaning that $\sigma = \mathrm{Tr}_{\mathcal{Y}}(\rho)$. Unless Y is a trivial register with just a single classical state, there are always multiple choices for ρ that are consistent with σ. Any such state ρ is said to be an *extension* of σ. For instance, $\rho = \sigma \otimes \xi$, for any density operator $\xi \in \mathrm{D}(\mathcal{Y})$, is always an extension of σ, because

$$\mathrm{Tr}_{\mathcal{Y}}(\sigma \otimes \xi) = \sigma \otimes \mathrm{Tr}(\xi) = \sigma.$$

If σ is pure, this is the only possible form for an extension. This is a mathematically simple statement, but it is nevertheless important at an intuitive level: it says that a register in a pure state cannot be correlated with any other registers.

A special type of extension is one in which the state of (X, Y) is pure: if $\rho = uu^* \in \mathrm{D}(\mathcal{X} \otimes \mathcal{Y})$ is a pure state for which

$$\mathrm{Tr}_{\mathcal{Y}}(uu^*) = \sigma,$$

it is said that ρ is a *purification* of σ. One also often refers to the vector u, as opposed to the operator uu^*, as being a purification of σ.

The notions of reductions, extensions, and purifications are easily extended to arbitrary positive semidefinite operators, as opposed to just density operators. For instance, if $P \in \mathrm{Pos}(\mathcal{X})$ is

a positive semidefinite operator and $u \in \mathcal{X} \otimes \mathcal{Y}$ is a vector for which

$$P = \mathrm{Tr}_{\mathcal{Y}}(uu^*),$$

it is said that u (or uu^*) is a purification of P.

For example suppose $\mathcal{X} = \mathbb{C}^{\Sigma}$ and $\mathcal{Y} = \mathbb{C}^{\Sigma}$, for some arbitrary (finite and nonempty) set Σ. The vector

$$u = \sum_{a \in \Sigma} e_a \otimes e_a$$

satisfies the equality

$$\mathbb{1}_{\mathcal{X}} = \mathrm{Tr}_{\mathcal{Y}}(uu^*),$$

and so u is a purification of $\mathbb{1}_{\mathcal{X}}$.

4.2 Existence and properties of purifications

A study of the properties of purifications is greatly simplified by the following observation. The vec mapping defined in Lecture 2 is a one-to-one and onto linear correspondence between $\mathcal{X} \otimes \mathcal{Y}$ and $\mathrm{L}(\mathcal{Y}, \mathcal{X})$; and for any choice of $u \in \mathcal{X} \otimes \mathcal{Y}$ and $A \in \mathrm{L}(\mathcal{Y}, \mathcal{X})$ satisfying $u = \mathrm{vec}(A)$ it holds that

$$\mathrm{Tr}_{\mathcal{Y}}(uu^*) = \mathrm{Tr}_{\mathcal{Y}}(\mathrm{vec}(A)\,\mathrm{vec}(A)^*) = AA^*.$$

Therefore, for every choice of complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$, and for any given operator $P \in \mathrm{Pos}(\mathcal{X})$, the following two properties are equivalent:

1. There exists a purification $u \in \mathcal{X} \otimes \mathcal{Y}$ of P.
2. There exists an operator $A \in \mathrm{L}(\mathcal{Y}, \mathcal{X})$ such that $P = AA^*$.

The following theorem, whose proof is based on this observation, establishes necessary and sufficient conditions for the existence of a purification of a given operator.

Theorem 4.1. *Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces, and let $P \in \mathrm{Pos}(\mathcal{X})$ be a positive semidefinite operator. There exists a purification $u \in \mathcal{X} \otimes \mathcal{Y}$ of P if and only if* $\dim(\mathcal{Y}) \geq \mathrm{rank}(P)$.

Proof. As discussed above, the existence of a purification $u \in \mathcal{X} \otimes \mathcal{Y}$ of P is equivalent to the existence of an operator $A \in \mathrm{L}(\mathcal{Y}, \mathcal{X})$ satisfying $P = AA^*$. Under the assumption that such an operator A exists, it is clear that

$$\mathrm{rank}(P) = \mathrm{rank}(AA^*) = \mathrm{rank}(A) \leq \dim(\mathcal{Y})$$

as claimed.

Conversely, under the assumption that $\dim(\mathcal{Y}) \geq \mathrm{rank}(P)$, there must exist operator $B \in \mathrm{L}(\mathcal{Y}, \mathcal{X})$ for which $BB^* = \Pi_{\mathrm{im}(P)}$ (the projection onto the image of P). To obtain such an operator B, let $r = \mathrm{rank}(P)$, use the spectral theorem to write

$$P = \sum_{j=1}^{r} \lambda_j(P) x_j x_j^*,$$

and let

$$B = \sum_{j=1}^{r} x_j y_j^*$$

for any choice of an orthonormal set $\{y_1, \ldots, y_r\} \subset \mathcal{Y}$. Now, for $A = \sqrt{P}B$ it holds that $AA^* = P$ as required. □

Corollary 4.2. *Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces such that* $\dim(\mathcal{Y}) \geq \dim(\mathcal{X})$. *For every choice of $P \in \mathrm{Pos}(\mathcal{X})$, there exists a purification $u \in \mathcal{X} \otimes \mathcal{Y}$ of P.*

Having established a simple condition under which purifications exist, the next step is to prove the following important relationship among all purifications of a given operator within a given space.

Theorem 4.3 (Unitary equivalence of purifications)**.** *Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces, and suppose that vectors $u, v \in \mathcal{X} \otimes \mathcal{Y}$ satisfy*

$$\mathrm{Tr}_{\mathcal{Y}}(uu^*) = \mathrm{Tr}_{\mathcal{Y}}(vv^*).$$

There exists a unitary operator $U \in \mathrm{U}(\mathcal{Y})$ such that $v = (\mathbb{1}_{\mathcal{X}} \otimes U)u$.

Proof. Let $P \in \mathrm{Pos}(\mathcal{X})$ satisfy $\mathrm{Tr}_{\mathcal{Y}}(uu^*) = P = \mathrm{Tr}_{\mathcal{Y}}(vv^*)$, and let $A, B \in \mathrm{L}(\mathcal{Y}, \mathcal{X})$ be the unique operators satisfying $u = \mathrm{vec}(A)$ and $v = \mathrm{vec}(B)$. It therefore holds that $AA^* = P = BB^*$. Letting $r = \mathrm{rank}(P)$, it follows that $\mathrm{rank}(A) = r = \mathrm{rank}(B)$.

Now, let $\{x_1, \ldots, x_r\} \subset \mathcal{X}$ be any orthonormal collection of eigenvectors of P with corresponding eigenvalues $\lambda_1(P), \ldots, \lambda_r(P)$. By the singular value theorem, it is possible to write

$$A = \sum_{j=1}^{r} \sqrt{\lambda_j(P)}\, x_j y_j^* \qquad \text{and} \qquad B = \sum_{j=1}^{r} \sqrt{\lambda_j(P)}\, x_j z_j^*$$

for some choice of orthonormal sets $\{y_1, \ldots, y_r\}$ and $\{z_1, \ldots, z_r\}$.

Finally, let $V \in \mathrm{U}(\mathcal{Y})$ be any unitary operator satisfying $Vz_j = y_j$ for every $j = 1, \ldots, r$. It follows that $AV = B$, and by taking $U = V^{\mathsf{T}}$ one has

$$(\mathbb{1}_{\mathcal{X}} \otimes U)u = (\mathbb{1}_{\mathcal{X}} \otimes V^{\mathsf{T}})\,\mathrm{vec}(A) = \mathrm{vec}(AV) = \mathrm{vec}(B) = v$$

as required. □

Theorem 4.3 will have significant value throughout the course, as a tool for proving a variety of results. It is also important at an intuitive level that the following example aims to illustrate.

Example 4.4. Suppose X and Y are distinct registers, and that Alice holds X and Bob holds Y in separate locations. Assume moreover that the pair (X, Y) is in a pure state uu^*.

Now imagine that Bob wishes to transform the state of (X, Y) so that it is in a different pure state vv^*. Assuming that Bob is able to do this without any interaction with Alice, it must hold that

$$\mathrm{Tr}_{\mathcal{Y}}(uu^*) = \mathrm{Tr}_{\mathcal{Y}}(vv^*). \tag{4.1}$$

This equation expresses the assumption that Bob does not touch X.

Theorem 4.3 implies that not only is (4.1) a necessary condition for Bob to transform uu^* into vv^*, but in fact it is sufficient. In particular, there must exist a unitary operator $U \in \mathrm{U}(\mathcal{Y})$ for which $v = (\mathbb{1}_{\mathcal{X}} \otimes U)u$, and Bob can implement the transformation from uu^* into vv^* by applying the unitary operation described by U to his register Y.

4.3 The fidelity function

There are different ways that one may quantify the similarity or difference between density operators. One way that relates closely to the notion of purifications is the *fidelity* between states. It is used extensively in the theory of quantum information.

4.3.1 Definition of the fidelity function

Given positive semidefinite operators $P, Q \in \mathrm{Pos}(\mathcal{X})$, we define the fidelity between P and Q as

$$\mathrm{F}(P, Q) = \left\| \sqrt{P}\sqrt{Q} \right\|_1.$$

Equivalently,

$$\mathrm{F}(P, Q) = \mathrm{Tr}\sqrt{\sqrt{P}Q\sqrt{P}}.$$

Similar to purifications, it is common to see the fidelity defined only for density operators as opposed to arbitrary positive semidefinite operators. It is, however, useful to extend the definition to all positive semidefinite operators as we have done, and it incurs little or no additional effort.

4.3.2 Basic properties of the fidelity

There are many interesting properties of the fidelity function. Let us begin with a few simple ones. First, the fidelity is symmetric: $\mathrm{F}(P, Q) = \mathrm{F}(Q, P)$ for all $P, Q \in \mathrm{Pos}(\mathcal{X})$. This is clear from the definition, given that $\|A\|_1 = \|A^*\|_1$ for all operators A.

Next, suppose that $u \in \mathcal{X}$ is a vector and $Q \in \mathrm{Pos}(\mathcal{X})$ is a positive semidefinite operator. It follows from the observation that $\sqrt{uu^*} = \frac{uu^*}{\|u\|}$ whenever $u \neq 0$ that

$$\mathrm{F}(uu^*, Q) = \sqrt{u^*Qu}.$$

In particular, $\mathrm{F}(uu^*, vv^*) = |\langle u, v\rangle|$ for any choice of vectors $u, v \in \mathcal{X}$.

One nice property of the fidelity that we will utilize several times is that it is multiplicative with respect to tensor products. This fact is stated in the following proposition (which can be easily extended from tensor products of two operators to any finite number of operators by induction).

Proposition 4.5. *Let $P_1, Q_1 \in \mathrm{Pos}(\mathcal{X}_1)$ and $P_2, Q_2 \in \mathrm{Pos}(\mathcal{X}_2)$ be positive semidefinite operators. It holds that*

$$\mathrm{F}(P_1 \otimes P_2, Q_1 \otimes Q_2) = \mathrm{F}(P_1, Q_1)\,\mathrm{F}(P_2, Q_2).$$

Proof. We have

$$\begin{aligned}
\mathrm{F}(P_1 \otimes P_2, Q_1 \otimes Q_2) &= \left\| \sqrt{P_1 \otimes P_2}\sqrt{Q_1 \otimes Q_2} \right\|_1 = \left\| \left(\sqrt{P_1} \otimes \sqrt{P_2}\right)\left(\sqrt{Q_1} \otimes \sqrt{Q_2}\right) \right\|_1 \\
&= \left\| \sqrt{P_1}\sqrt{Q_1} \otimes \sqrt{P_2}\sqrt{Q_2} \right\|_1 = \left\| \sqrt{P_1}\sqrt{Q_1} \right\|_1 \left\| \sqrt{P_2}\sqrt{Q_2} \right\|_1 = \mathrm{F}(P_1, Q_1)\,\mathrm{F}(P_2, Q_2)
\end{aligned}$$

as claimed. □

4.3.3 Uhlmann's theorem

Next we will prove a fundamentally important theorem about the fidelity, known as Uhlmann's theorem, which relates the fidelity to the notion of purifications.

Theorem 4.6 (Uhlmann's theorem)**.** *Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces, let $P, Q \in \mathrm{Pos}(\mathcal{X})$ be positive semidefinite operators, both having rank at most $\dim(\mathcal{Y})$, and let $u \in \mathcal{X} \otimes \mathcal{Y}$ be any purification of P. It holds that*

$$\mathrm{F}(P, Q) = \max\left\{ |\langle u, v\rangle| \,:\, v \in \mathcal{X} \otimes \mathcal{Y},\ \mathrm{Tr}_{\mathcal{Y}}(vv^*) = Q \right\}.$$

Proof. Given that the rank of both P and Q is at most $\dim(\mathcal{Y})$, there must exist operators $A, B \in \mathrm{L}(\mathcal{X}, \mathcal{Y})$ for which $A^*A = \Pi_{\mathrm{im}(P)}$ and $B^*B = \Pi_{\mathrm{im}(Q)}$. The equations

$$\mathrm{Tr}_{\mathcal{Y}}\left(\mathrm{vec}\left(\sqrt{P}A^*\right)\mathrm{vec}\left(\sqrt{P}A^*\right)^*\right) = \sqrt{P}A^*A\sqrt{P} = P$$
$$\mathrm{Tr}_{\mathcal{Y}}\left(\mathrm{vec}\left(\sqrt{Q}B^*\right)\mathrm{vec}\left(\sqrt{Q}B^*\right)^*\right) = \sqrt{Q}B^*B\sqrt{Q} = Q$$

follow, demonstrating that

$$\mathrm{vec}\left(\sqrt{P}A^*\right) \quad \text{and} \quad \mathrm{vec}\left(\sqrt{Q}B^*\right)$$

are purifications of P and Q, respectively. By Theorem 4.3 it follows that every choice of a purification $u \in \mathcal{X} \otimes \mathcal{Y}$ of P must take the form

$$u = (\mathbb{1}_{\mathcal{X}} \otimes U)\,\mathrm{vec}\left(\sqrt{P}A^*\right) = \mathrm{vec}\left(\sqrt{P}A^*U^{\mathsf{T}}\right),$$

for some unitary operator $U \in \mathrm{U}(\mathcal{Y})$, and likewise every purification $v \in \mathcal{X} \otimes \mathcal{Y}$ of Q must take the form

$$v = (\mathbb{1}_{\mathcal{X}} \otimes V)\,\mathrm{vec}\left(\sqrt{Q}B^*\right) = \mathrm{vec}\left(\sqrt{Q}B^*V^{\mathsf{T}}\right)$$

for some unitary operator $V \in \mathrm{U}(\mathcal{Y})$.

The maximization in the statement of the theorem is therefore equivalent to

$$\max_{V \in \mathrm{U}(\mathcal{Y})} \left|\left\langle \mathrm{vec}\left(\sqrt{P}A^*U^{\mathsf{T}}\right), \mathrm{vec}\left(\sqrt{Q}B^*V^{\mathsf{T}}\right)\right\rangle\right|,$$

which may alternately be written as

$$\max_{V \in \mathrm{U}(\mathcal{Y})} \left|\left\langle U^{\mathsf{T}}\overline{V}, A\sqrt{P}\sqrt{Q}B^*\right\rangle\right| \tag{4.2}$$

for some choice of $U \in \mathrm{U}(\mathcal{Y})$. As $V \in \mathrm{U}(\mathcal{Y})$ ranges over all unitary operators, so too does $U^{\mathsf{T}}\overline{V}$, and therefore the quantity represented by equation (4.2) is given by

$$\left\| A\sqrt{P}\sqrt{Q}B^* \right\|_1.$$

Finally, given that A^*A and B^*B are projection operators, A and B must both have spectral norm at most 1. It therefore holds that

$$\left\| \sqrt{P}\sqrt{Q} \right\|_1 = \left\| A^*A\sqrt{P}\sqrt{Q}B^*B \right\|_1 \leq \left\| A\sqrt{P}\sqrt{Q}B^* \right\|_1 \leq \left\| \sqrt{P}\sqrt{Q} \right\|_1$$

so that

$$\left\| A\sqrt{P}\sqrt{Q}B^* \right\|_1 = \left\| \sqrt{P}\sqrt{Q} \right\|_1 = \mathrm{F}(P, Q).$$

The equality in the statement of the theorem therefore holds. □

Various properties of the fidelity follow from Uhlmann's theorem. For example, it is clear from the theorem that $0 \leq \mathrm{F}(\rho, \xi) \leq 1$ for density operators ρ and ξ. Moreover $\mathrm{F}(\rho, \xi) = 1$ if and only if $\rho = \xi$. It is also evident (from the definition) that $\mathrm{F}(\rho, \xi) = 0$ if and only if $\sqrt{\rho}\sqrt{\xi} = 0$, which is equivalent to $\rho\xi = 0$ (i.e., to ρ and ξ having orthogonal images).

Another property of the fidelity that follows from Uhlmann's theorem is as follows.

Proposition 4.7. *Let $P_1, \ldots, P_k, Q_1, \ldots, Q_k \in \mathrm{Pos}(\mathcal{X})$ be positive semidefinite operators. It holds that*

$$\mathrm{F}\left(\sum_{i=1}^{k} P_i, \sum_{i=1}^{k} Q_i\right) \geq \sum_{i=1}^{k} \mathrm{F}\left(P_i, Q_i\right).$$

Proof. Let $\mathcal{Y}$ be a complex Euclidean space having dimension at least that of $\mathcal{X}$, and choose vectors $u_1, \ldots, u_k, v_1, \ldots, v_k \in \mathcal{X} \otimes \mathcal{Y}$ satisfying $\mathrm{Tr}_{\mathcal{Y}}\left(u_i u_i^*\right) = P_i$, $\mathrm{Tr}_{\mathcal{Y}}\left(v_i v_i^*\right) = Q_i$, and $\langle u_i, v_i \rangle = \mathrm{F}(P_i, Q_i)$ for each $i = 1, \ldots, k$. Such vectors exist by Uhlmann's theorem. Let $\mathcal{Z} = \mathbb{C}^k$ and define $u, v \in \mathcal{X} \otimes \mathcal{Y} \otimes \mathcal{Z}$ as

$$u = \sum_{i=1}^{k} u_i \otimes e_i \qquad \text{and} \qquad v = \sum_{i=1}^{k} v_i \otimes e_i.$$

We have

$$\mathrm{Tr}_{\mathcal{Y}\otimes\mathcal{Z}}\left(uu^*\right) = \sum_{i=1}^{k} P_i \qquad \text{and} \qquad \mathrm{Tr}_{\mathcal{Y}\otimes\mathcal{Z}}\left(vv^*\right) = \sum_{i=1}^{k} Q_i.$$

Thus, again using Uhlmann's theorem, we have

$$\mathrm{F}\left(\sum_{i=1}^{k} P_i, \sum_{i=1}^{k} Q_i\right) \geq \left|\langle u, v \rangle\right| = \sum_{i=1}^{k} \mathrm{F}\left(P_i, Q_i\right)$$

as required. □

It follows from this proposition is that the fidelity function is *concave* in the first argument:

$$\mathrm{F}(\lambda\rho_1 + (1-\lambda)\rho_2, \xi) \geq \lambda\, \mathrm{F}(\rho_1, \xi) + (1-\lambda)\, \mathrm{F}(\rho_2, \xi)$$

for all $\rho_1, \rho_2, \xi \in \mathrm{D}\left(\mathcal{X}\right)$ and $\lambda \in [0,1]$, and by symmetry it is concave in the second argument as well. In fact, the fidelity is *jointly concave*:

$$\mathrm{F}(\lambda\rho_1 + (1-\lambda)\rho_2, \lambda\xi_1 + (1-\lambda)\xi_2) \geq \lambda\, \mathrm{F}(\rho_1, \xi_1) + (1-\lambda)\, \mathrm{F}(\rho_2, \xi_2).$$

for all $\rho_1, \rho_2, \xi_1, \xi_2 \in \mathrm{D}\left(\mathcal{X}\right)$ and $\lambda \in [0,1]$.

4.3.4 Alberti's theorem

A different characterization of the fidelity function is given by Alberti's theorem, which is as follows.

Theorem 4.8 (Alberti)**.** *Let $\mathcal{X}$ be a complex Euclidean space and let $P, Q \in \mathrm{Pos}\left(\mathcal{X}\right)$ be positive semidefinite operators. It holds that*

$$\left(\mathrm{F}(P, Q)\right)^2 = \inf_{R \in \mathrm{Pd}(\mathcal{X})} \langle R, P \rangle \, \langle R^{-1}, Q \rangle.$$

When we study semidefinite programming later in the course, we will see that this theorem is in fact closely related to Uhlmann's theorem through semidefinite programming duality. For now we will make due with a different proof. It is more complicated, but it has the value that it illustrates some useful tricks from matrix analysis. To prove the theorem, it is helpful to start first with the special case that $P = Q$, which is represented by the following lemma.

Lemma 4.9. *Let $P \in \mathrm{Pos}(\mathcal{X})$. It holds that*

$$\inf_{R \in \mathrm{Pd}(\mathcal{X})} \langle R, P \rangle \langle R^{-1}, P \rangle = (\mathrm{Tr}(P))^2.$$

Proof. It is clear that

$$\inf_{R \in \mathrm{Pd}(\mathcal{X})} \langle R, P \rangle \langle R^{-1}, P \rangle \leq (\mathrm{Tr}(P))^2,$$

given that $R = \mathbb{1}$ is positive definite. To establish the reverse inequality, it suffices to prove that

$$\langle R, P \rangle \langle R^{-1}, P \rangle \geq (\mathrm{Tr}(P))^2$$

for any choice of $R \in \mathrm{Pd}(\mathcal{X})$. This will follow from the simple observation that, for any choice of positive real numbers α and β, we have $\alpha^2 + \beta^2 \geq 2\alpha\beta$ and therefore $\alpha\beta^{-1} + \beta\alpha^{-1} \geq 2$. With this fact in mind, consider a spectral decomposition

$$R = \sum_{i=1}^{n} \lambda_i u_i u_i^*.$$

We have

$$\begin{aligned}
\langle R, P \rangle \langle R^{-1}, P \rangle &= \sum_{1 \leq i,j \leq n} \lambda_i \lambda_j^{-1} (u_i^* P u_i)(u_j^* P u_j) \\
&= \sum_{1 \leq i \leq n} (u_i^* P u_i)^2 + \sum_{1 \leq i < j \leq n} (\lambda_i \lambda_j^{-1} + \lambda_j \lambda_i^{-1})(u_i^* P u_i)(u_j^* P u_j) \\
&\geq \sum_{1 \leq i \leq n} (u_i^* P u_i)^2 + 2 \sum_{1 \leq i < j \leq n} (u_i^* P u_i)(u_j^* P u_j) \\
&= (\mathrm{Tr}(P))^2
\end{aligned}$$

as required. □

Proof of Theorem 4.8. We will first prove the theorem for P and Q positive definite. Let us define $S \in \mathrm{Pd}(\mathcal{X})$ to be

$$S = \left(\sqrt{P} Q \sqrt{P}\right)^{-1/4} \sqrt{P} R \sqrt{P} \left(\sqrt{P} Q \sqrt{P}\right)^{-1/4}.$$

Notice that as R ranges over all positive definite operators, so too does S. We have

$$\begin{aligned}
\left\langle S, \left(\sqrt{P} Q \sqrt{P}\right)^{1/2} \right\rangle &= \langle R, P \rangle, \\
\left\langle S^{-1}, \left(\sqrt{P} Q \sqrt{P}\right)^{1/2} \right\rangle &= \langle R^{-1}, Q \rangle.
\end{aligned}$$

Therefore, by Lemma 4.9, we have

$$\begin{aligned}
\inf_{R \in \mathrm{Pd}(\mathcal{X})} \langle R, P \rangle \langle R^{-1}, Q \rangle &= \inf_{S \in \mathrm{Pd}(\mathcal{X})} \left\langle S, \left(\sqrt{P} Q \sqrt{P}\right)^{1/2} \right\rangle \left\langle S^{-1}, \left(\sqrt{P} Q \sqrt{P}\right)^{1/2} \right\rangle \\
&= \left(\mathrm{Tr} \sqrt{\sqrt{P} Q \sqrt{P}}\right)^2 \\
&= (\mathrm{F}(P, Q))^2.
\end{aligned}$$

To prove the general case, let us first note that, for any choice of $R \in \mathrm{Pd}(\mathcal{X})$ and $\varepsilon > 0$, we have

$$\langle R, P \rangle \langle R^{-1}, Q \rangle \leq \langle R, P + \varepsilon \mathbb{1} \rangle \langle R^{-1}, Q + \varepsilon \mathbb{1} \rangle.$$

Thus,

$$\inf_{R \in \mathrm{Pd}(\mathcal{X})} \langle R, P \rangle \langle R^{-1}, Q \rangle \leq (\mathrm{F}(P + \varepsilon \mathbb{1}, Q + \varepsilon \mathbb{1}))^2$$

for all $\varepsilon > 0$. As

$$\lim_{\varepsilon \to 0^+} \mathrm{F}(P + \varepsilon \mathbb{1}, Q + \varepsilon \mathbb{1}) = \mathrm{F}(P, Q)$$

we have

$$\inf_{R \in \mathrm{Pd}(\mathcal{X})} \langle R, P \rangle \langle R^{-1}, Q \rangle \leq (\mathrm{F}(P, Q))^2.$$

On the other hand, for any choice of $R \in \mathrm{Pd}(\mathcal{X})$ we have

$$\langle R, P + \varepsilon \mathbb{1} \rangle \langle R^{-1}, Q + \varepsilon \mathbb{1} \rangle \geq (\mathrm{F}(P + \varepsilon \mathbb{1}, Q + \varepsilon \mathbb{1}))^2 \geq (\mathrm{F}(P, Q))^2$$

for all $\varepsilon > 0$, and therefore

$$\langle R, P \rangle \langle R^{-1}, Q \rangle \geq (\mathrm{F}(P, Q))^2.$$

As this holds for all $R \in \mathrm{Pd}(\mathcal{X})$ we have

$$\inf_{R \in \mathrm{Pd}(\mathcal{X})} \langle R, P \rangle \langle R^{-1}, Q \rangle \geq (\mathrm{F}(P, Q))^2,$$

which completes the proof. □

4.4 The Fuchs–van de Graaf inequalities

We will now state and prove the Fuchs–van de Graaf inequalities, which establish a close relationship between the trace norm of the difference between two density operators and their fidelity. The inequalities are as stated in the following theorem.

Theorem 4.10 (Fuchs–van de Graaf)**.** *Let $\mathcal{X}$ be a complex Euclidean space and assume that $\rho, \xi \in \mathrm{D}(\mathcal{X})$ are density operators on $\mathcal{X}$. It holds that*

$$1 - \frac{1}{2} \|\rho - \xi\|_1 \leq \mathrm{F}(\rho, \xi) \leq \sqrt{1 - \frac{1}{4} \|\rho - \xi\|_1^2}.$$

To prove this theorem we first need the following lemma relating the trace norm and Frobenius norm. Once we have it in hand, the theorem will be easy to prove.

Lemma 4.11. *Let $\mathcal{X}$ be a complex Euclidean space and let $P, Q \in \mathrm{Pos}(\mathcal{X})$ be positive semidefinite operators on $\mathcal{X}$. It holds that*

$$\|P - Q\|_1 \geq \left\| \sqrt{P} - \sqrt{Q} \right\|_2^2.$$

Proof. Let

$$\sqrt{P} - \sqrt{Q} = \sum_{i=1}^{n} \lambda_i u_i u_i^*$$

be a spectral decomposition of $\sqrt{P} - \sqrt{Q}$. Given that $\sqrt{P} - \sqrt{Q}$ is Hermitian, it follows that

$$\sum_{i=1}^{n} |\lambda_i|^2 = \left\| \sqrt{P} - \sqrt{Q} \right\|_2^2.$$

Now, define

$$U = \sum_{i=1}^{n} \operatorname{sign}(\lambda_i) u_i u_i^*$$

where

$$\operatorname{sign}(\lambda) = \begin{cases} 1 & \text{if } \lambda \geq 0 \\ -1 & \text{if } \lambda < 0 \end{cases}$$

for every real number λ. It follows that

$$U\left(\sqrt{P} - \sqrt{Q}\right) = \left(\sqrt{P} - \sqrt{Q}\right) U = \sum_{i=1}^{n} |\lambda_i| \, u_i u_i^* = \left| \sqrt{P} - \sqrt{Q} \right|.$$

Using the operator identity

$$A^2 - B^2 = \frac{1}{2}((A - B)(A + B) + (A + B)(A - B)),$$

along with the fact that U is unitary, we have

$$\begin{aligned} \|P - Q\|_1 &\geq |\operatorname{Tr}((P - Q)U)| \\ &= \left| \frac{1}{2} \operatorname{Tr}\left((\sqrt{P} - \sqrt{Q})(\sqrt{P} + \sqrt{Q})U\right) + \frac{1}{2} \operatorname{Tr}\left((\sqrt{P} + \sqrt{Q})(\sqrt{P} - \sqrt{Q})U\right) \right| \\ &= \operatorname{Tr}\left(\left| \sqrt{P} - \sqrt{Q} \right| \left(\sqrt{P} + \sqrt{Q} \right) \right). \end{aligned}$$

Now, by the triangle inequality (for real numbers), we have that

$$u_i^* \left(\sqrt{P} + \sqrt{Q} \right) u_i \geq \left| u_i^* \sqrt{P} u_i - u_i^* \sqrt{Q} u_i \right| = |\lambda_i|$$

for every $i = 1, \ldots, n$. Thus

$$\operatorname{Tr}\left(\left| \sqrt{P} - \sqrt{Q} \right| \left(\sqrt{P} + \sqrt{Q} \right) \right) = \sum_{i=1}^{n} |\lambda_i| \, u_i^* \left(\sqrt{P} + \sqrt{Q} \right) u_i \geq \sum_{i=1}^{n} |\lambda_i|^2 = \left\| \sqrt{P} - \sqrt{Q} \right\|_2^2$$

as required. □

Proof of Theorem 4.10. The operators ρ and ξ have unit trace, and therefore

$$\left\| \sqrt{\rho} - \sqrt{\xi} \right\|_2^2 = \operatorname{Tr}\left(\sqrt{\rho} - \sqrt{\xi} \right)^2 = 2 - 2 \operatorname{Tr}\left(\sqrt{\rho} \sqrt{\xi} \right) \geq 2 - 2 \operatorname{F}(\rho, \xi).$$

The first inequality therefore follows from Lemma 4.11.

To prove the second inequality, let $\mathcal{Y}$ be a complex Euclidean space with $\dim(\mathcal{Y}) = \dim(\mathcal{X})$, and let $u, v \in \mathcal{X} \otimes \mathcal{Y}$ satisfy $\operatorname{Tr}_{\mathcal{Y}}(uu^*) = \rho$, $\operatorname{Tr}_{\mathcal{Y}}(vv^*) = \xi$, and $\operatorname{F}(\rho, \xi) = |\langle u, v \rangle|$. Such vectors exist as a consequence of Uhlmann's theorem. By the monotonicity of the trace norm we have

$$\|\rho - \xi\|_1 \leq \|uu^* - vv^*\|_1 = 2\sqrt{1 - |\langle u, v \rangle|^2} = 2\sqrt{1 - \operatorname{F}(\rho, \xi)^2},$$

and therefore

$$\mathrm{F}(\rho, \xi) \leq \sqrt{1 - \frac{1}{4} \left\| \rho - \xi \right\|_1^2}$$

as required. □