Chapter 3

Similarity and distance among states and channels

The main focus of this chapter is on quantifiable notions of similarity and distance between quantum states, the task of discrimination among two or more quantum state alternatives, and related notions involving channels.

There are three main sections of the chapter, the first of which discusses the task of *discrimination* between pairs of quantum states, its connection to the trace norm, and generalizations of this task to more than two states. The second section introduces the *fidelity* function and describes some of its basic properties, formulations, and connections to other concepts. The third section discusses the *completely bounded trace norm*, which is a natural analogue of the trace norm for mappings between spaces of operators, and establishes a connection between this norm and the task of discrimination between pairs of quantum channels.

3.1 Quantum state discrimination

It is a natural question to ask how well a given collection of quantum states can be discriminated by means of a measurement. The hypothetical task of *state discrimination* serves as an abstraction through which this question may be considered.

In the simplest formulation of the state discrimination task, one of two known quantum states is selected at random, and a register prepared in that state is made available to a hypothetical individual. This individual's goal

is to determine which of the two states was selected, by means of a measurement performed on the given register. A theorem known as the *Holevo–Helstrom theorem* gives a closed-form expression, based on the trace norm of a weighted difference between the two possible states, for the probability that an optimally chosen measurement correctly identifies the selected state. An explicit description of an optimal measurement may be obtained from the proof of this theorem.

State discrimination may also be considered in the situation where more than two states are to be discriminated. An analysis of this task is more complicated than the two-state case, and simple, closed-form expressions for the optimal success probability are not known in general. It is possible, however, to represent the optimal success probability through the use of semidefinite programming, which provides a valuable analytical tool through which state discrimination may be analyzed. Approximate solutions, together with bounds on their performance, are also considered.

3.1.1 Discriminating between pairs of quantum states

The simplest formulation of the state discrimination task concerns the discrimination between two fixed quantum states $\rho_0, \rho_1 \in \mathrm{D}(\mathcal{X})$ of a given register X. A key aspect of this formulation, together with its analysis, is that it establishes a close connection between the trace norm and the task of state discrimination. Somewhat more generally, one finds that the trace norm provides a natural way of quantifying the "measurable difference" between two quantum states.

Discriminating between pairs of probabilistic states

Before discussing the task of state discrimination between pairs of quantum states, it is appropriate to consider an analogous problem for probabilistic states. To this end, consider the following scenario involving two hypothetical individuals: Alice and Bob.

Scenario 3.1. Let Σ be an alphabet, let X be a classical register with classical state set Σ, and let Y be a classical register with classical state set $\{0,1\}$. Also let $p_0, p_1 \in \mathcal{P}(\Sigma)$ be probability vectors, representing probabilistic states of X, and let $\lambda \in [0,1]$ be a real number. The vectors p_0 and p_1, as well as the number λ, are assumed to be known to both Alice and Bob.

Alice prepares Y in a probabilistic state, so that its value is 0 with probability λ and 1 with probability $1-\lambda$. Conditioned on the value stored in Y, Alice performs one of the following actions:

1. If $\mathsf{Y} = 0$, Alice prepares X in the probabilistic state p_0, and sends X to Bob.
2. If $\mathsf{Y} = 1$, Alice prepares X in the probabilistic state p_1, and sends X to Bob.

Bob's goal is to correctly determine the value of the bit stored in Y, using only the information he can gather from an observation of X.

An optimal strategy in this scenario for Bob, assuming that he wishes to maximize the probability of correctly guessing the value of Y, may be derived from Bayes' theorem, which implies

$$\begin{aligned} \Pr\bigl(\mathsf{Y} = 0 | \mathsf{X} = b\bigr) &= \frac{\lambda p_0(b)}{\lambda p_0(b) + (1-\lambda) p_1(b)} \\ \Pr\bigl(\mathsf{Y} = 1 | \mathsf{X} = b\bigr) &= \frac{(1-\lambda) p_1(b)}{\lambda p_0(b) + (1-\lambda) p_1(b)} \end{aligned} \tag{3.1}$$

for each $b \in \Sigma$. Given the knowledge that $\mathsf{X} = b$, Bob should therefore choose the more likely value for Y: if it holds that $\lambda p_0(b) > (1-\lambda) p_1(b)$, then Bob should guess that $\mathsf{Y} = 0$, while if $\lambda p_0(b) < (1-\lambda) p_1(b)$, then Bob should guess that $\mathsf{Y} = 1$. In the case that $\lambda p_0(b) = (1-\lambda) p_1(b)$, Bob can guess either $\mathsf{Y} = 0$ or $\mathsf{Y} = 1$ arbitrarily without affecting his probability of being correct, as the two values are equally likely in this situation.

The probability that Bob correctly identifies the value of Y using this strategy can be understood by first considering the probability he is correct *minus* the probability he is incorrect. This difference in probabilities is represented by the quantity

$$\sum_{b \in \Sigma} \bigl| \lambda p_0(b) - (1-\lambda) p_1(b) \bigr| = \bigl\| \lambda p_0 - (1-\lambda) p_1 \bigr\|_1. \tag{3.2}$$

It follows that the probability that Bob is correct is given by the quantity

$$\frac{1}{2} + \frac{1}{2} \bigl\| \lambda p_0 - (1-\lambda) p_1 \bigr\|_1. \tag{3.3}$$

This expression makes clear the close connection between probabilistic state discrimination and the vector 1-norm.

Notice that

$$0 \leq \bigl\| \lambda p_0 - (1-\lambda) p_1 \bigr\|_1 \leq 1, \tag{3.4}$$

where the second inequality follows from the triangle inequality. This is consistent with the interpretation of the expression (3.3) as a probability. In the extreme case where

$$\bigl\| \lambda p_0 - (1-\lambda) p_1 \bigr\|_1 = 0, \tag{3.5}$$

which requires $\lambda = 1/2$ and $p_0 = p_1$, Bob is essentially reduced to guessing blindly and will be correct with probability $1/2$. In the other extreme,

$$\bigl\| \lambda p_0 - (1-\lambda) p_1 \bigr\|_1 = 1, \tag{3.6}$$

it must hold that λp_0 and $(1-\lambda) p_1$ have disjoint supports, and thus Bob can determine the value of Y without error. Intermediate values, in which both inequalities in (3.4) hold strictly, correspond to different degrees of certainty in Bob's guess.

Discriminating between pairs of quantum states

The task of discriminating between pairs of quantum states is represented by the following scenario, which is the natural quantum generalization of Scenario 3.1.

Scenario 3.2. Let X be an arbitrary register and let Y be a classical register with classical state set $\{0,1\}$. Also let $\rho_0, \rho_1 \in \mathrm{D}(\mathcal{X})$ be states of X, and let $\lambda \in [0,1]$ be a real number. The states ρ_0 and ρ_1, as well as the number λ, are assumed to be known to both Alice and Bob.

Alice prepares Y in a probabilistic state, so that its value is 0 with probability λ and 1 with probability $1-\lambda$. Conditioned on the value stored in Y, Alice performs one of the following actions:

1. If $\mathsf{Y} = 0$, Alice prepares X in the state ρ_0, and sends X to Bob.
2. If $\mathsf{Y} = 1$, Alice prepares X in the state ρ_1, and sends X to Bob.

Bob's goal is to correctly determine the value of the bit stored in Y, by means of a measurement of X.

The principal goal of the discussion that follows is to establish an analogous connection between this scenario and the trace norm to the one that was shown above to hold between Scenario 3.1 and the vector 1-norm. The following lemma, which happens to concern the spectral norm rather than the trace norm, is useful for establishing this connection. The lemma is stated in greater generality than is required for the purposes of the present section, but the more general form will find uses elsewhere in this book.

Lemma 3.3. *Let $\mathcal{X}$ be a complex Euclidean space, let Σ be an alphabet, let $u \in \mathbb{C}^{\Sigma}$ be a vector, and let $\{P_a : a \in \Sigma\} \subset \mathrm{Pos}(\mathcal{X})$ be a collection of positive semidefinite operators. It holds that*

$$\left\| \sum_{a \in \Sigma} u(a) P_a \right\| \leq \|u\|_\infty \left\| \sum_{a \in \Sigma} P_a \right\|. \tag{3.7}$$

Proof. Define an operator $A \in \mathrm{L}\big(\mathcal{X}, \mathcal{X} \otimes \mathbb{C}^{\Sigma}\big)$ as

$$A = \sum_{a \in \Sigma} \sqrt{P_a} \otimes e_a. \tag{3.8}$$

The spectral norm is submultiplicative with respect to compositions and multiplicative with respect to tensor products, and therefore

$$\begin{aligned} \left\| \sum_{a \in \Sigma} u(a) P_a \right\| &= \left\| \sum_{a \in \Sigma} u(a) A^* (\mathbb{1}_{\mathcal{X}} \otimes E_{a,a}) A \right\| \\ &\leq \|A^*\| \left\| \sum_{a \in \Sigma} u(a) E_{a,a} \right\| \|A\| = \|u\|_\infty \|A\|^2. \end{aligned} \tag{3.9}$$

Finally, by the spectral norm property (1.173), one has

$$\|A\|^2 = \|A^* A\| = \left\| \sum_{a \in \Sigma} P_a \right\|, \tag{3.10}$$

which completes the proof. □

A direct connection between Scenario 3.2 and the trace norm can now be established. The next theorem, known as the Holevo–Helstrom theorem, expresses this connection in mathematical terms.

Theorem 3.4 (Holevo–Helstrom theorem)**.** *Let $\mathcal{X}$ be a complex Euclidean space, let $\rho_0, \rho_1 \in \mathrm{D}(\mathcal{X})$ be density operators, and let $\lambda \in [0,1]$. For every choice of a measurement $\mu : \{0,1\} \to \mathrm{Pos}(\mathcal{X})$, it holds that*

$$\lambda \langle \mu(0), \rho_0 \rangle + (1-\lambda) \langle \mu(1), \rho_1 \rangle \leq \frac{1}{2} + \frac{1}{2} \big\| \lambda \rho_0 - (1-\lambda) \rho_1 \big\|_1. \tag{3.11}$$

Moreover, there exists a projective measurement $\mu : \{0,1\} \to \mathrm{Pos}(\mathcal{X})$ for which equality is achieved in (3.11).

Proof. Define

$$\rho = \lambda \rho_0 + (1-\lambda)\rho_1 \quad \text{and} \quad X = \lambda \rho_0 - (1-\lambda)\rho_1, \tag{3.12}$$

so that

$$\lambda \rho_0 = \frac{\rho + X}{2} \quad \text{and} \quad (1-\lambda)\rho_1 = \frac{\rho - X}{2}, \tag{3.13}$$

and therefore

$$\lambda \langle \mu(0), \rho_0 \rangle + (1-\lambda) \langle \mu(1), \rho_1 \rangle = \frac{1}{2} + \frac{1}{2} \langle \mu(0) - \mu(1), X \rangle. \tag{3.14}$$

By Lemma 3.3, together with the Hölder inequality for Schatten norms, it follows that

$$\begin{aligned} &\frac{1}{2} + \frac{1}{2} \langle \mu(0) - \mu(1), X \rangle \\ &\quad \leq \frac{1}{2} + \frac{1}{2} \big\| \mu(0) - \mu(1) \big\| \, \big\| X \big\|_1 \leq \frac{1}{2} + \frac{1}{2} \big\| X \big\|_1. \end{aligned} \tag{3.15}$$

Combining (3.14) and (3.15) yields (3.11).

To show that equality is achieved in (3.11) for a projective measurement $\mu : \{0,1\} \to \mathrm{Pos}(\mathcal{X})$, one may consider the Jordan-Hahn decomposition

$$X = P - Q, \tag{3.16}$$

for $P, Q \in \mathrm{Pos}(\mathcal{X})$. Define $\mu : \{0,1\} \to \mathrm{Pos}(\mathcal{X})$ as

$$\mu(0) = \Pi_{\mathrm{im}(P)} \quad \text{and} \quad \mu(1) = \mathbb{1} - \Pi_{\mathrm{im}(P)}, \tag{3.17}$$

which is a projective measurement. It holds that

$$\langle \mu(0) - \mu(1), X \rangle = \mathrm{Tr}(P) + \mathrm{Tr}(Q) = \big\| X \big\|_1, \tag{3.18}$$

and therefore

$$\lambda \langle \mu(0), \rho_0 \rangle + (1-\lambda) \langle \mu(1), \rho_1 \rangle = \frac{1}{2} + \frac{1}{2} \big\| X \big\|_1, \tag{3.19}$$

which completes the proof. □

It follows from Theorem 3.4 that an optimal choice of a measurement for Bob in Scenario 3.2 correctly determines the value of Y with probability

$$\frac{1}{2} + \frac{1}{2}\big\| \lambda\rho_0 - (1-\lambda)\rho_1 \big\|_1, \tag{3.20}$$

and, moreover, this optimal probability is achievable with a projective measurement.

One might question the implicit claim that the possible strategies for Bob in Scenario 3.2 are exhausted by the consideration of measurements having 0 and 1 as the only possible outcomes. For instance, Bob could measure X using a measurement with three or more outcomes, and then base his guess for the value of Y on some sort of post-processing of the measurement outcome obtained. However, no generality is introduced by this type of strategy, or any other strategy having access to the register X alone. Any process used by Bob to eventually produce a binary-valued guess for the classical state of Y must define a binary-valued measurement, and Theorem 3.4 may be applied to this measurement.

The following proposition, whose proof has some overlap with the proof of the Theorem 3.4, establishes a useful relationship between the trace norm of an operator and the 1-norm of a vector obtained from that operator's inner products with the measurement operators of any measurement.

Proposition 3.5. *Let $\mathcal{X}$ be a complex Euclidean space, let Σ be an alphabet, let $\mu : \Sigma \to \mathrm{Pos}(\mathcal{X})$ be a measurement, and let $X \in \mathrm{L}(\mathcal{X})$ be an operator. Define a vector $v \in \mathbb{C}^\Sigma$ as*

$$v(a) = \langle \mu(a), X \rangle \tag{3.21}$$

for each $a \in \Sigma$. It holds that $\|v\|_1 \leq \|X\|_1$.

Proof. One has

$$\|v\|_1 = \sum_{a\in\Sigma} \big|\langle \mu(a), X\rangle\big| = \sum_{a\in\Sigma} u(a)\langle \mu(a), X\rangle = \left\langle \sum_{a\in\Sigma} \overline{u(a)}\mu(a), X \right\rangle \tag{3.22}$$

for some choice of a vector $u \in \mathbb{C}^\Sigma$ satisfying $|u(a)| = 1$ for each $a \in \Sigma$. By Lemma 3.3, together with Hölder's inequality for Schatten norms, it follows that

$$\|v\|_1 \leq \left\| \sum_{a\in\Sigma} \overline{u(a)}\mu(a) \right\| \big\|X\big\|_1 \leq \big\|X\big\|_1, \tag{3.23}$$

as required. □

Discriminating between convex sets of quantum states

The task of state discrimination between pairs of quantum states may be generalized to one in which two convex sets of quantum states are to be discriminated. The following scenario describes this task in more precise terms.

Scenario 3.6. Let X be an arbitrary register and let Y be a classical register having classical state set $\{0, 1\}$. Also let $\mathcal{C}_0, \mathcal{C}_1 \subseteq \mathrm{D}(\mathcal{X})$ be nonempty, convex sets of states, and let $\lambda \in [0, 1]$ be a real number. The sets $\mathcal{C}_0$ and $\mathcal{C}_1$, as well as the number λ, are assumed to be known to both Alice and Bob.

Alice prepares Y in a probabilistic state, so that its value is 0 with probability λ and 1 with probability $1 - \lambda$. Conditioned on the value stored in Y, Alice performs one of the following actions:

1. If $\mathsf{Y} = 0$, Alice prepares X in any state $\rho_0 \in \mathcal{C}_0$ of her choice, and sends X to Bob.
2. If $\mathsf{Y} = 1$, Alice prepares X in any state $\rho_1 \in \mathcal{C}_1$ of her choice, and sends X to Bob.

Bob's goal is to predict the value of the bit stored in Y, by means of a measurement of X.

The description of Scenario 3.6 does not specify how Alice is to choose ρ_0 or ρ_1, beyond stating the requirement that $\rho_0 \in \mathcal{C}_0$ and $\rho_1 \in \mathcal{C}_1$. It could be, for instance, that Alice chooses these states randomly according to fixed distributions, or she could choose the states adversarially, even based on a knowledge of the measurement Bob intends to use. What is relevant is that Bob can make no assumptions regarding Alice's choices for ρ_0 and ρ_1, beyond the requirement that she chooses $\rho_0 \in \mathcal{C}_0$ and $\rho_1 \in \mathcal{C}_1$.

One may note that Scenario 3.2 represents a special case of Scenario 3.6 in which $\mathcal{C}_0$ and $\mathcal{C}_1$ are the singleton sets $\{\rho_0\}$ and $\{\rho_1\}$, respectively.

It follows from the Holevo–Helstrom theorem (Theorem 3.4) that Bob cannot hope to succeed in his task in Scenario 3.6 with probability higher than

$$\frac{1}{2} + \frac{1}{2}\big\| \lambda\rho_0 - (1-\lambda)\rho_1 \big\|_1, \tag{3.24}$$

for whichever states $\rho_0 \in \mathcal{C}_0$ and $\rho_1 \in \mathcal{C}_1$ Alice chooses, for this is his optimal success probability when he has the additional knowledge that Alice

chooses either ρ_0 or ρ_1. The following proposition implies that Bob can succeed with probability at least

$$\frac{1}{2} + \frac{1}{2} \inf_{\rho_0,\rho_1} \left\| \lambda\rho_0 - (1-\lambda)\rho_1 \right\|_1, \tag{3.25}$$

where the infimum is taken over all choices of $\rho_0 \in \mathcal{C}_0$ and $\rho_1 \in \mathcal{C}_1$. In light of the limitation imposed by the Holevo–Helstrom theorem, this is necessarily the optimal probability of success in the worst case.

Theorem 3.7. *Let $\mathcal{C}_0, \mathcal{C}_1 \subseteq \mathrm{D}(\mathcal{X})$ be nonempty, convex sets, for $\mathcal{X}$ being a complex Euclidean space, and let $\lambda \in [0,1]$. It holds that*

$$\begin{aligned} &\max_{\mu} \inf_{\rho_0,\rho_1} \Big(\lambda \langle \mu(0), \rho_0 \rangle + (1-\lambda) \langle \mu(1), \rho_1 \rangle \Big) \\ &\qquad = \inf_{\rho_0,\rho_1} \max_{\mu} \Big(\lambda \langle \mu(0), \rho_0 \rangle + (1-\lambda) \langle \mu(1), \rho_1 \rangle \Big) \\ &\qquad = \frac{1}{2} + \frac{1}{2} \inf_{\rho_0,\rho_1} \left\| \lambda\rho_0 - (1-\lambda)\rho_1 \right\|_1, \end{aligned} \tag{3.26}$$

where the infima are over all choices of $\rho_0 \in \mathcal{C}_0$ and $\rho_1 \in \mathcal{C}_1$, and the maxima are over all choices of binary measurements $\mu : \{0,1\} \to \mathrm{Pos}(\mathcal{X})$.

Proof. Define sets $\mathcal{A}, \mathcal{B} \subset \mathrm{Pos}(\mathcal{X} \oplus \mathcal{X})$ as

$$\mathcal{A} = \left\{ \begin{pmatrix} \rho_0 & 0 \\ 0 & \rho_1 \end{pmatrix} : \rho_0 \in \mathcal{C}_0,\, \rho_1 \in \mathcal{C}_1 \right\} \tag{3.27}$$

and

$$\mathcal{B} = \left\{ \begin{pmatrix} \lambda P_0 & 0 \\ 0 & (1-\lambda) P_1 \end{pmatrix} : P_0, P_1 \in \mathrm{Pos}(\mathcal{X}),\, P_0 + P_1 = \mathbb{1}_{\mathcal{X}} \right\}, \tag{3.28}$$

as well as a function $f : \mathcal{A} \times \mathcal{B} \to \mathbb{R}$ as $f(A,B) = \langle A, B \rangle$. It holds that $\mathcal{A}$ and $\mathcal{B}$ are convex, $\mathcal{B}$ is compact, and f is bilinear, so that

$$\inf_{A \in \mathcal{A}} \max_{B \in \mathcal{B}} f(A,B) = \max_{B \in \mathcal{B}} \inf_{A \in \mathcal{A}} f(A,B) \tag{3.29}$$

holds by Sion's min-max theorem (Theorem 1.12). The equation (3.29) is equivalent to the first equality of (3.26), and the second equality in (3.26) follows from Theorem 3.4. □

3.1.2 Discriminating quantum states of an ensemble

The remaining variant of quantum state discrimination to be discussed in this chapter is similar to the one represented by Scenario 3.2, except that more than two possible states, selected from a given ensemble, are to be discriminated. The following scenario describes this task in more precise terms.

Scenario 3.8. Let X be an arbitrary register, let Σ be an alphabet, let Y be a classical register having classical state set Σ, and let $\eta : \Sigma \to \mathrm{Pos}(\mathcal{X})$ be an ensemble of states. Alice prepares the pair (Y, X) in the classical-quantum state

$$\sigma = \sum_{a \in \Sigma} E_{a,a} \otimes \eta(a) \tag{3.30}$$

determined by the ensemble η. Equivalently, the register Y takes each value $a \in \Sigma$ with probability $p(a) = \mathrm{Tr}(\eta(a))$, and conditioned on the event $\mathsf{Y} = a$ the state of X is set to $\eta(a)/\mathrm{Tr}(\eta(a))$, for each $a \in \Sigma$. The register X is sent to Bob, and Bob's goal is to predict the classical state of Y, using only the information he can gather from a measurement of X.

For any measurement $\mu : \Sigma \to \mathrm{Pos}(\mathcal{X})$ chosen by Bob in this scenario, the probability that he correctly predicts the classical state of Y is given by the expression

$$\sum_{a \in \Sigma} \langle \mu(a), \eta(a) \rangle. \tag{3.31}$$

It is therefore natural to consider a maximization of this quantity over all choices of the measurement μ.

More generally, one may substitute an arbitrary function of the form $\phi : \Sigma \to \mathrm{Herm}(\mathcal{X})$ in place of the ensemble $\eta : \Sigma \to \mathrm{Pos}(\mathcal{X})$, and consider a maximization of the quantity

$$\sum_{a \in \Sigma} \langle \mu(a), \phi(a) \rangle \tag{3.32}$$

over all measurements $\mu : \Sigma \to \mathrm{Pos}(\mathcal{X})$. One situation in which this more general optimization problem is meaningful is a variant of Scenario 3.8 in which different payoff values are associated to each pair (a,b), representing the state a of Alice's register Y and Bob's measurement outcome b. If Bob receives a payoff value of $K(a,b)$ for producing the measurement outcome

b when Alice's register Y holds the symbol a, for instance, Bob's expected gain for a given measurement $\mu : \Sigma \to \mathrm{Pos}(\mathcal{X})$ is given by

$$\sum_{a\in\Sigma}\sum_{b\in\Sigma} K(a,b)\langle \mu(b), \eta(a)\rangle = \sum_{b\in\Sigma}\langle \mu(b), \phi(b)\rangle \tag{3.33}$$

for

$$\phi(b) = \sum_{a\in\Sigma} K(a,b)\,\eta(a). \tag{3.34}$$

This sort of hypothetical situation could be further generalized by allowing the classical state set of Alice's register Y and Bob's set of measurement outcomes to disagree.

A semidefinite program for optimal measurements

For any choice of a complex Euclidean space $\mathcal{X}$, an alphabet Σ, and a function $\phi : \Sigma \to \mathrm{Herm}(\mathcal{X})$, define

$$\mathrm{opt}(\phi) = \max_{\mu} \sum_{a\in\Sigma} \langle \mu(a), \phi(a)\rangle, \tag{3.35}$$

where the maximum is over all measurements $\mu : \Sigma \to \mathrm{Pos}(\mathcal{X})$. This optimal value is necessarily achieved for some choice of a measurement, as it is a maximization of a continuous function over the compact set of measurements of the form $\mu : \Sigma \to \mathrm{Pos}(\mathcal{X})$, which justifies the use of the maximum rather than the supremum. It may also be said that a particular choice of a measurement μ is *optimal* for ϕ if the above expression (3.32) coincides with the value $\mathrm{opt}(\phi)$.

There is no closed-form expression that is known to represent the value $\mathrm{opt}(\phi)$ for an arbitrary choice of a function $\phi : \Sigma \to \mathrm{Herm}(\mathcal{X})$. However, it is possible to express the value $\mathrm{opt}(\phi)$ by a semidefinite program, providing a method by which it may be numerically calculated using a computer. A simplified description of the primal and dual problems associated with such a semidefinite program are as follows:

Primal problem (simplified)		Dual problem (simplified)	
maximize:	$\sum_{a\in\Sigma}\langle \mu(a), \phi(a)\rangle$	minimize:	$\mathrm{Tr}(Y)$
subject to:	$\mu : \Sigma \to \mathrm{Pos}(\mathcal{X})$,	subject to:	$Y \geq \phi(a)$ (for all $a \in \Sigma$),
	$\sum_{a\in\Sigma} \mu(a) = \mathbb{1}_{\mathcal{X}}$.		$Y \in \mathrm{Herm}(\mathcal{X})$.

A formal expression of this semidefinite program that conforms to the definition of semidefinite programs presented in Section 1.2.3 is given by the triple $(\Phi, A, \mathbb{1}_{\mathcal{X}})$, where the mapping $\Phi \in \mathrm{T}(\mathcal{Y}\otimes\mathcal{X}, \mathcal{X})$ is defined as the partial trace $\Phi = \mathrm{Tr}_{\mathcal{Y}}$, for $\mathcal{Y} = \mathbb{C}^{\Sigma}$, and the operator A is defined as

$$A = \sum_{a\in\Sigma} E_{a,a} \otimes \phi(a). \tag{3.36}$$

The primal and dual problems associated with the triple $(\Phi, A, \mathbb{1}_{\mathcal{X}})$ are as follows:

Primal problem (formal)		Dual problem (formal)	
maximize:	$\langle A, X\rangle$	minimize:	$\mathrm{Tr}(Y)$
subject to:	$\mathrm{Tr}_{\mathcal{Y}}(X) = \mathbb{1}_{\mathcal{X}}$,	subject to:	$\mathbb{1}_{\mathcal{Y}} \otimes Y \geq A$,
	$X \in \mathrm{Pos}(\mathcal{Y}\otimes\mathcal{X})$.		$Y \in \mathrm{Herm}(\mathcal{X})$.

These problems are equivalent to the simplified primal and dual problems described above. In greater detail, any feasible solution μ to the simplified primal problem described above gives rise to the feasible solution

$$X = \sum_{a\in\Sigma} E_{a,a} \otimes \mu(a) \tag{3.37}$$

to the formal primal problem, in which the same objective value

$$\langle A, X\rangle = \sum_{a\in\Sigma} \langle \mu(a), \phi(a)\rangle \tag{3.38}$$

is achieved. While a feasible solution X to the formal primal problem need not take the form (3.37) in general, one may nevertheless obtain a feasible solution μ to the simplified primal problem from such an operator X by setting

$$\mu(a) = \bigl(e_a^* \otimes \mathbb{1}_{\mathcal{X}}\bigr) X \bigl(e_a \otimes \mathbb{1}_{\mathcal{X}}\bigr) \tag{3.39}$$

for each $a \in \Sigma$. The equality (3.38) again holds, and therefore the two primal problems have the same optimal values. The fact that the two dual problems are equivalent is evident from the observation that the inequality

$$\mathbb{1}_{\mathcal{Y}} \otimes Y \geq \sum_{a\in\Sigma} E_{a,a} \otimes \phi(a) \tag{3.40}$$

is equivalent to the inequality $Y \geq \phi(a)$ holding for every $a \in \Sigma$.

Strong duality holds for this semidefinite program. The operator

$$X = \frac{1}{|\Sigma|} \mathbb{1}_{\mathcal{Y}} \otimes \mathbb{1}_{\mathcal{X}} \tag{3.41}$$

is a strictly feasible primal solution, while $Y = \gamma \mathbb{1}_{\mathcal{X}}$ is a strictly feasible dual solution for any real value $\gamma > \lambda_1(A)$. It follows from Slater's theorem (Theorem 1.18) that the optimal primal and dual values for the semidefinite program are equal, and moreover the optimum value is achieved in both the primal and dual problems.

Criteria for measurement optimality

It may be difficult to obtain an analytic description of a measurement $\mu : \Sigma \to \mathrm{Pos}(\mathcal{X})$ that is optimal for a given function $\phi : \Sigma \to \mathrm{Herm}(\mathcal{X})$, given the lack of a known closed-form expression for such a measurement. In contrast, it is straightforward to verify that an optimal measurement is indeed optimal by means of the following theorem.

Theorem 3.9 (Holevo–Yuen–Kennedy–Lax)**.** *Let $\phi : \Sigma \to \mathrm{Herm}(\mathcal{X})$ be a function and let $\mu : \Sigma \to \mathrm{Pos}(\mathcal{X})$ be a measurement, for $\mathcal{X}$ being a complex Euclidean space and Σ being an alphabet. It holds that*

$$\sum_{a \in \Sigma} \langle \mu(a), \phi(a) \rangle = \mathrm{opt}(\phi) \tag{3.42}$$

(i.e., the measurement μ is optimal for the function ϕ) if and only if the operator

$$\sum_{a \in \Sigma} \phi(a)\mu(a) \tag{3.43}$$

is Hermitian and satisfies

$$\sum_{a \in \Sigma} \phi(a)\mu(a) \geq \phi(b) \tag{3.44}$$

for every $b \in \Sigma$.

Proof. Define an operator $X \in \mathrm{Herm}(\mathcal{Y} \otimes \mathcal{X})$ as

$$X = \sum_{a \in \Sigma} E_{a,a} \otimes \mu(a). \tag{3.45}$$

Suppose first that μ is an optimal measurement for ϕ, so that X is an optimal primal solution to the semidefinite program $(\Phi, A, \mathbb{1}_{\mathcal{X}})$ representing $\mathrm{opt}(\phi)$, as described previously. As the dual optimum of this semidefinite program is always achieved, one may choose $Y \in \mathrm{Herm}(\mathcal{X})$ to be such a dual-optimal solution. By the property of complementary slackness for semidefinite programs (Proposition 1.19), it necessarily holds that

$$(\mathbb{1}_{\mathcal{Y}} \otimes Y)X = AX. \tag{3.46}$$

Taking the partial trace of both sides of (3.46) over $\mathcal{Y}$, one finds that

$$Y = Y \,\mathrm{Tr}_{\mathcal{Y}}(X) = \mathrm{Tr}_{\mathcal{Y}}(AX) = \sum_{a \in \Sigma} \phi(a)\mu(a). \tag{3.47}$$

The dual-feasibility of Y therefore implies that the operator (3.43) is Hermitian and satisfies (3.44).

To prove the reverse implication, note that if the operator (3.43) is Hermitian and satisfies (3.44) for every $b \in \Sigma$, then

$$Y = \sum_{a \in \Sigma} \phi(a)\mu(a) \tag{3.48}$$

is a dual-feasible solution to the semidefinite program $(\Phi, A, \mathbb{1}_{\mathcal{X}})$ representing $\mathrm{opt}(\phi)$. The operator X defined in (3.45) is a primal-feasible solution to this semidefinite program, simply by virtue of the fact that μ is a measurement. The objective values achieved by X in the primal problem and Y in the dual problem are both equal to

$$\sum_{a \in \Sigma} \langle \mu(a), \phi(a) \rangle. \tag{3.49}$$

The equality between these values implies that both are optimal by the property of weak duality of semidefinite programs. The measurement μ is therefore optimal for ϕ. □

The pretty good measurement

Returning to Bob's task, as described in Scenario 3.8, suppose an ensemble $\eta : \Sigma \to \mathrm{Pos}(\mathcal{X})$ is given, and a measurement $\mu : \Sigma \to \mathrm{Pos}(\mathcal{X})$ maximizing the probability

$$\sum_{a \in \Sigma} \langle \mu(a), \eta(a) \rangle \tag{3.50}$$

of a correct determination of the state of Alice's classical register Y is sought.

In a concrete setting in which an explicit description of η is known, the semidefinite programming formulation of $\mathrm{opt}(\eta)$ allows for an efficient numerical approximation to a measurement μ that is optimal for η. This approach may, however, be unsatisfactory in more abstract settings, such as ones in which it is necessary to view η as being indeterminate. Although Theorem 3.9 allows for a verification that a given optimal measurement is indeed optimal, it does not provide a method to find a measurement that is optimal.

One alternative to searching for an optimal measurement is to consider measurements that are determined from η by closed-form expressions, but that might be sub-optimal. The so-called *pretty good measurement* is an example of such a measurement.

To define the pretty good measurement for a given ensemble η, one first considers the average state

$$\rho = \sum_{a\in\Sigma} \eta(a) \tag{3.51}$$

of η. In the case that ρ is positive definite, the pretty good measurement associated with η is the measurement $\mu : \Sigma \to \mathrm{Pos}(\mathcal{X})$ defined as

$$\mu(a) = \rho^{-\frac{1}{2}} \eta(a) \rho^{-\frac{1}{2}}. \tag{3.52}$$

In general, when ρ is not necessarily invertible, one may use the Moore–Penrose pseudo-inverse of ρ, in place of the inverse of ρ, to define[1] the pretty good measurement associated with η as

$$\mu(a) = \sqrt{\rho^+}\, \eta(a) \sqrt{\rho^+} + \frac{1}{|\Sigma|} \Pi_{\ker(\rho)} \tag{3.54}$$

for every $a \in \Sigma$.

The pretty good measurement will generally not be optimal for a given ensemble. It does, however, achieve a probability of a correct prediction that

[1] It should be noted that, although the equation (3.54) is taken here as the definition of the pretty good measurement, it is somewhat arbitrary in the case that ρ is not invertible. Any measurement $\mu : \Sigma \to \mathrm{Pos}(\mathcal{X})$ satisfying

$$\mu(a) \geq \sqrt{\rho^+}\, \eta(a) \sqrt{\rho^+} \tag{3.53}$$

for all $a \in \Sigma$ would be equivalent with respect to the discussion that follows.

is at least the square of the optimal success probability, as the following theorem states.

Theorem 3.10 (Barnum–Knill). *Let $\mathcal{X}$ be a complex Euclidean space, let Σ be an alphabet, let $\eta : \Sigma \to \mathrm{Pos}(\mathcal{X})$ be an ensemble of states, and let $\mu : \Sigma \to \mathrm{Pos}(\mathcal{X})$ denote the pretty good measurement associated with η. It holds that*

$$\sum_{a\in\Sigma} \langle \mu(a), \eta(a) \rangle \geq \mathrm{opt}(\eta)^2. \tag{3.55}$$

Proof. Let $\rho = \sum_{a\in\Sigma} \eta(a)$ and let $\nu : \Sigma \to \mathrm{Pos}(\mathcal{X})$ be any measurement. For every $a \in \Sigma$ it holds that $\mathrm{im}(\eta(a)) \subseteq \mathrm{im}(\rho)$, and therefore

$$\langle \nu(a), \eta(a) \rangle = \left\langle \rho^{\frac{1}{4}} \nu(a) \rho^{\frac{1}{4}}, \left(\rho^+\right)^{\frac{1}{4}} \eta(a) \left(\rho^+\right)^{\frac{1}{4}} \right\rangle. \tag{3.56}$$

By the Cauchy–Schwarz inequality, it follows that

$$\langle \nu(a), \eta(a) \rangle \leq \left\| \rho^{\frac{1}{4}} \nu(a) \rho^{\frac{1}{4}} \right\|_2 \left\| \left(\rho^+\right)^{\frac{1}{4}} \eta(a) \left(\rho^+\right)^{\frac{1}{4}} \right\|_2 \tag{3.57}$$

for each $a \in \Sigma$. Applying the Cauchy–Schwarz inequality again, this time for vectors of real numbers rather than for operators, one finds that

$$\sum_{a\in\Sigma} \langle \nu(a), \eta(a) \rangle \leq \sqrt{\sum_{a\in\Sigma} \left\| \rho^{\frac{1}{4}} \nu(a) \rho^{\frac{1}{4}} \right\|_2^2} \sqrt{\sum_{a\in\Sigma} \left\| \left(\rho^+\right)^{\frac{1}{4}} \eta(a) \left(\rho^+\right)^{\frac{1}{4}} \right\|_2^2}. \tag{3.58}$$

The first term on the right-hand side of (3.58) is at most 1. To verify that this is so, one may first use the definition of the Frobenius norm to obtain the expression

$$\left\| \rho^{\frac{1}{4}} \nu(a) \rho^{\frac{1}{4}} \right\|_2^2 = \left\langle \rho^{\frac{1}{4}} \nu(a) \rho^{\frac{1}{4}}, \rho^{\frac{1}{4}} \nu(a) \rho^{\frac{1}{4}} \right\rangle = \left\langle \nu(a), \sqrt{\rho} \nu(a) \sqrt{\rho} \right\rangle \tag{3.59}$$

for each $a \in \Sigma$, from which it follows that

$$\left\| \rho^{\frac{1}{4}} \nu(a) \rho^{\frac{1}{4}} \right\|_2^2 \leq \mathrm{Tr}\left(\sqrt{\rho} \nu(a) \sqrt{\rho}\right), \tag{3.60}$$

by virtue of the fact that $\nu(a) \leq \mathbb{1}_{\mathcal{X}}$ and $\sqrt{\rho}\nu(a)\sqrt{\rho} \geq 0$. Summing over all $a \in \Sigma$ yields

$$\sum_{a\in\Sigma} \left\| \rho^{\frac{1}{4}} \nu(a) \rho^{\frac{1}{4}} \right\|_2^2 \leq \sum_{a\in\Sigma} \mathrm{Tr}\left(\sqrt{\rho} \nu(a) \sqrt{\rho}\right) = \mathrm{Tr}(\rho) = 1. \tag{3.61}$$

By the definition of the pretty good measurement, along with a similar computation to the one expressed by (3.59), one has that

$$\left\| \left(\rho^{+}\right)^{\frac{1}{4}} \eta(a) \left(\rho^{+}\right)^{\frac{1}{4}} \right\|_2^2 = \left\langle \sqrt{\rho^{+}}\, \eta(a) \sqrt{\rho^{+}}, \eta(a) \right\rangle \leq \langle \mu(a), \eta(a) \rangle \tag{3.62}$$

for each $a \in \Sigma$, and therefore

$$\sum_{a \in \Sigma} \left\| \left(\rho^{+}\right)^{\frac{1}{4}} \eta(a) \left(\rho^{+}\right)^{\frac{1}{4}} \right\|_2^2 \leq \sum_{a \in \Sigma} \langle \mu(a), \eta(a) \rangle. \tag{3.63}$$

By (3.58), (3.61), and (3.63) it follows that

$$\left(\sum_{a \in \Sigma} \langle \nu(a), \eta(a) \rangle \right)^2 \leq \sum_{a \in \Sigma} \langle \mu(a), \eta(a) \rangle. \tag{3.64}$$

As this inequality holds for all measurements $\nu : \Sigma \to \mathrm{Pos}(\mathcal{X})$, including those measurements that are optimal for η, the proof is complete. □

3.2 The fidelity function

This section introduces the *fidelity function*, which provides a measure of the similarity, or "overlap," between quantum states (and positive semidefinite operators more generally) that will be used extensively throughout this book. It is defined as follows.

Definition 3.11. Let $\mathcal{X}$ be a complex Euclidean space and let $P, Q \in \mathrm{Pos}(\mathcal{X})$ be positive semidefinite operators. The *fidelity* $\mathrm{F}(P, Q)$ between P and Q is defined as

$$\mathrm{F}(P, Q) = \left\| \sqrt{P} \sqrt{Q} \right\|_1. \tag{3.65}$$

The function F is called the *fidelity function*.

The fidelity function is most often considered for density operator inputs, but there is value in defining it more generally, allowing its arguments to range over arbitrary positive semidefinite operators. An alternative expression for the fidelity function is obtained by expanding (3.65) according to the definition of the trace norm:

$$\mathrm{F}(P, Q) = \mathrm{Tr}\left(\sqrt{\sqrt{P} Q \sqrt{P}} \right). \tag{3.66}$$

3.2.1 Elementary properties of the fidelity function

The following proposition establishes several basic properties of the fidelity function.

Proposition 3.12. *Let $\mathcal{X}$ be a complex Euclidean space and let $P, Q \in \mathrm{Pos}(\mathcal{X})$ be positive semidefinite operators. The following facts hold:*

1. *The fidelity function* F *is continuous at* (P, Q).
2. $\mathrm{F}(P, Q) = \mathrm{F}(Q, P)$.
3. $\mathrm{F}(\lambda P, Q) = \sqrt{\lambda}\, \mathrm{F}(P, Q) = \mathrm{F}(P, \lambda Q)$ *for every nonnegative real number* λ.
4. $\mathrm{F}(P, Q) = \mathrm{F}(P, \Pi_{\mathrm{im}(P)} Q \Pi_{\mathrm{im}(P)}) = \mathrm{F}(\Pi_{\mathrm{im}(Q)} P \Pi_{\mathrm{im}(Q)}, Q)$.
5. $\mathrm{F}(P, Q) \geq 0$, *with equality if and only if* $PQ = 0$.
6. $\mathrm{F}(P, Q)^2 \leq \mathrm{Tr}(P)\, \mathrm{Tr}(Q)$, *with equality if and only if P and Q are linearly dependent.*
7. *For every complex Euclidean space $\mathcal{Y}$ with* $\dim(\mathcal{Y}) \geq \dim(\mathcal{X})$ *and every isometry* $V \in \mathrm{U}(\mathcal{X}, \mathcal{Y})$, *it holds that* $\mathrm{F}(P, Q) = \mathrm{F}(VPV^*, VQV^*)$.

Proof. The first three statements follow directly from the definition of the fidelity function: the fidelity function is defined as a composition of continuous functions, and is therefore continuous at every point in its domain; it holds that $\|A\|_1 = \|A^*\|_1$ for any choice of an operator A, and therefore

$$\left\| \sqrt{P} \sqrt{Q} \right\|_1 = \left\| \left(\sqrt{P} \sqrt{Q} \right)^* \right\|_1 = \left\| \sqrt{Q} \sqrt{P} \right\|_1; \tag{3.67}$$

and by the positive scalability of the trace norm, one has

$$\left\| \sqrt{\lambda P} \sqrt{Q} \right\|_1 = \sqrt{\lambda} \left\| \sqrt{P} \sqrt{Q} \right\|_1 = \left\| \sqrt{P} \sqrt{\lambda Q} \right\|_1. \tag{3.68}$$

Moving on to the fourth statement, it follows from the observation

$$\sqrt{P} = \sqrt{P}\, \Pi_{\mathrm{im}(P)} = \Pi_{\mathrm{im}(P)} \sqrt{P} \tag{3.69}$$

that

$$\sqrt{P} Q \sqrt{P} = \sqrt{P} \Pi_{\mathrm{im}(P)} Q \Pi_{\mathrm{im}(P)} \sqrt{P}. \tag{3.70}$$

Through the use of the expression (3.66), it follows that

$$\mathrm{F}(P, Q) = \mathrm{F}\big(P, \Pi_{\mathrm{im}(P)} Q \Pi_{\mathrm{im}(P)}\big). \tag{3.71}$$

This proves the first equality in statement 4, while the second equality follows through a combination of the first equality and statement 2.

Statement 5 follows from the fact that the trace norm is positive definite:

$$\left\| \sqrt{P}\sqrt{Q} \right\|_1 \geq 0, \tag{3.72}$$

with equality if and only if $\sqrt{P}\sqrt{Q} = 0$, which is equivalent to $PQ = 0$.

To prove the sixth statement, observe first that, by (1.177), there must exist a unitary operator $U \in \mathrm{U}(\mathcal{X})$ for which

$$\mathrm{F}(P,Q)^2 = \left\| \sqrt{P}\sqrt{Q} \right\|_1^2 = \left| \left\langle U, \sqrt{P}\sqrt{Q} \right\rangle \right|^2 = \left| \left\langle \sqrt{P}U, \sqrt{Q} \right\rangle \right|^2. \tag{3.73}$$

By the Cauchy–Schwarz inequality, it holds that

$$\left| \left\langle \sqrt{P}U, \sqrt{Q} \right\rangle \right|^2 \leq \left\| \sqrt{P}U \right\|_2^2 \left\| \sqrt{Q} \right\|_2^2 = \mathrm{Tr}(P)\,\mathrm{Tr}(Q), \tag{3.74}$$

which establishes the claimed inequality in statement 6. If it is the case that P and Q are linearly dependent, then it must hold that $P = \lambda Q$ or $Q = \lambda P$ for some choice of a nonnegative real number λ. In either case, it is straightforward to verify that

$$\mathrm{F}(P,Q)^2 = \mathrm{Tr}(P)\,\mathrm{Tr}(Q). \tag{3.75}$$

On the other hand, if P and Q are linearly independent, then so too are $\sqrt{P}U$ and $\sqrt{Q}$ for all unitary operators U; for if it holds that

$$\alpha\sqrt{P}U + \beta\sqrt{Q} = 0 \tag{3.76}$$

for scalars $\alpha, \beta \in \mathbb{C}$, then it follows that $|\alpha|^2 P = |\beta|^2 Q$. The assumption that P and Q are linearly independent therefore implies that a strict inequality occurs in the application of the Cauchy–Schwarz inequality in (3.74), which completes the proof of statement 6.

Finally, to prove statement 7, one may observe first that

$$\sqrt{VPV^*} = V\sqrt{P}V^* \quad \text{and} \quad \sqrt{VQV^*} = V\sqrt{Q}V^* \tag{3.77}$$

for every isometry $V \in \mathrm{U}(\mathcal{X}, \mathcal{Y})$. By the isometric invariance of the trace norm, it follows that

$$\mathrm{F}(VPV^*, VQV^*) = \left\| V\sqrt{P}V^*V\sqrt{Q}V^* \right\|_1 = \left\| \sqrt{P}\sqrt{Q} \right\|_1, \tag{3.78}$$

which proves statement 7. □

Statements 5 and 6 of Proposition 3.12 imply that

$$0 \leq \mathrm{F}(\rho,\sigma) \leq 1 \tag{3.79}$$

for all density operators $\rho, \sigma \in \mathrm{D}(\mathcal{X})$. Moreover, $\mathrm{F}(\rho,\sigma) = 0$ if and only if ρ and σ have orthogonal images, and $\mathrm{F}(\rho,\sigma) = 1$ if and only if $\rho = \sigma$.

The output of the fidelity function is given by a simple formula when one of its input operators has rank equal to 1, as the next proposition states.

Proposition 3.13. *Let $\mathcal{X}$ be a complex Euclidean space, let $v \in \mathcal{X}$ be a vector, and let $P \in \mathrm{Pos}(\mathcal{X})$ be a positive semidefinite operator. It holds that*

$$\mathrm{F}(P, vv^*) = \sqrt{v^*Pv}. \tag{3.80}$$

In particular, for every choice of vectors $u, v \in \mathcal{X}$, it holds that

$$\mathrm{F}(uu^*, vv^*) = |\langle u, v \rangle|. \tag{3.81}$$

Proof. The operator

$$\sqrt{P}\,vv^*\sqrt{P} \tag{3.82}$$

is positive semidefinite and has rank at most 1, which makes its vector of eigenvalues straightforward to calculate:

$$\lambda_1\Big(\sqrt{P}\,vv^*\sqrt{P}\Big) = \mathrm{Tr}\Big(\sqrt{P}\,vv^*\sqrt{P}\Big) = v^*P\,v \tag{3.83}$$

and

$$\lambda_k\Big(\sqrt{P}\,vv^*\sqrt{P}\Big) = 0 \tag{3.84}$$

for $k \geq 2$. It follows that

$$\mathrm{F}(P, vv^*) = \mathrm{Tr}\left(\sqrt{\sqrt{P}\,vv^*\sqrt{P}}\right) = \sqrt{\lambda_1\Big(\sqrt{P}\,vv^*\sqrt{P}\Big)} = \sqrt{v^*P\,v}, \tag{3.85}$$

as claimed. □

The following proposition is representative of another case in which the fidelity function has a simple formula. One corollary of this proposition, known as *Winter's gentle measurement lemma*, is useful in some situations.[2]

[2] The term *gentle measurement* reflects the observation that if a measurement of a particular state yields a particular outcome with very high probability, then a non-destructive analogue of that measurement causes only a small perturbation to the state in the event that the likely outcome is obtained.

Proposition 3.14. *Let $\mathcal{X}$ be a complex Euclidean space and let $P, Q \in \mathrm{Pos}(\mathcal{X})$ be positive semidefinite operators. It holds that*

$$\mathrm{F}(P, QPQ) = \langle P, Q \rangle. \tag{3.86}$$

Proof. It holds that

$$\sqrt{\sqrt{P}QPQ\sqrt{P}} = \sqrt{\left(\sqrt{P}Q\sqrt{P}\right)^2} = \sqrt{P}Q\sqrt{P}, \tag{3.87}$$

and therefore

$$\mathrm{F}(P, QPQ) = \mathrm{Tr}\left(\sqrt{\sqrt{P}QPQ\sqrt{P}}\right) = \mathrm{Tr}\left(\sqrt{P}Q\sqrt{P}\right) = \langle P, Q \rangle, \tag{3.88}$$

as claimed. □

Corollary 3.15 (Winter's gentle measurement lemma)**.** *Let $\mathcal{X}$ be a complex Euclidean space, let $\rho \in \mathrm{D}(\mathcal{X})$ be a density operator, and let $P \in \mathrm{Pos}(\mathcal{X})$ be a positive semidefinite operator satisfying $P \leq \mathbb{1}_{\mathcal{X}}$ and $\langle P, \rho \rangle > 0$. It holds that*

$$\mathrm{F}\left(\rho, \frac{\sqrt{P}\rho\sqrt{P}}{\langle P, \rho \rangle}\right) \geq \sqrt{\langle P, \rho \rangle}. \tag{3.89}$$

Proof. By Proposition 3.14, along with statement 3 of Proposition 3.12, one has

$$\mathrm{F}\left(\rho, \frac{\sqrt{P}\rho\sqrt{P}}{\langle P, \rho \rangle}\right) = \frac{1}{\sqrt{\langle P, \rho \rangle}} \mathrm{F}\left(\rho, \sqrt{P}\rho\sqrt{P}\right) = \frac{\langle \sqrt{P}, \rho \rangle}{\sqrt{\langle P, \rho \rangle}}. \tag{3.90}$$

Under the assumption $0 \leq P \leq \mathbb{1}$, it holds that $\sqrt{P} \geq P$, and therefore $\langle \sqrt{P}, \rho \rangle \geq \langle P, \rho \rangle$, from which the corollary follows. □

Another simple, yet very useful, property of the fidelity function is that it is multiplicative with respect to tensor products.

Proposition 3.16. *Let $P_0, Q_0 \in \mathrm{Pos}(\mathcal{X}_0)$ and $P_1, Q_1 \in \mathrm{Pos}(\mathcal{X}_1)$ be positive semidefinite operators, for complex Euclidean spaces $\mathcal{X}_0$ and $\mathcal{X}_1$. It holds that*

$$\mathrm{F}(P_0 \otimes P_1, Q_0 \otimes Q_1) = \mathrm{F}(P_0, Q_0)\, \mathrm{F}(P_1, Q_1). \tag{3.91}$$

Proof. Operator square roots and compositions respect tensor products, and the trace norm is multiplicative with respect to tensor products, so

$$\begin{aligned} \mathrm{F}(P_0 \otimes P_1, Q_0 \otimes Q_1) &= \left\| \sqrt{P_0 \otimes P_1}\sqrt{Q_0 \otimes Q_1} \right\|_1 \\ &= \left\| \sqrt{P_0}\sqrt{Q_0} \otimes \sqrt{P_1}\sqrt{Q_1} \right\|_1 = \left\| \sqrt{P_0}\sqrt{Q_0} \right\|_1 \left\| \sqrt{P_1}\sqrt{Q_1} \right\|_1 \\ &= \mathrm{F}(P_0, Q_0)\, \mathrm{F}(P_1, Q_1), \end{aligned} \tag{3.92}$$

as claimed. □

3.2.2 Alternative characterizations of the fidelity function

Multiple alternative characterizations of the fidelity function are known; a selection of such alternative characterizations is presented below. Some of these characterizations will allow for further properties of the fidelity function to be established, or will find other uses elsewhere in this book.

Block operator characterization

The first alternate characterization of the fidelity function to be presented is given by the following theorem. This characterization is particularly useful for establishing relevant properties of the fidelity function, including joint concavity in its arguments and monotonicity under the actions of channels, as will be described in the section following this one.

Theorem 3.17. *Let $\mathcal{X}$ be a complex Euclidean space and let $P, Q \in \mathrm{Pos}(\mathcal{X})$ be positive semidefinite operators. It holds that*

$$\mathrm{F}(P, Q) = \max\left\{ \left|\mathrm{Tr}(X)\right| \,:\, X \in \mathrm{L}(\mathcal{X}),\, \begin{pmatrix} P & X \\ X^* & Q \end{pmatrix} \in \mathrm{Pos}(\mathcal{X} \oplus \mathcal{X}) \right\}. \tag{3.93}$$

The following lemma, which will find other uses elsewhere in this book, will be used to prove Theorem 3.17. The lemma is stated in slightly greater generality than is needed in the present context, in that it does not require P and Q to act on the same spaces, but there is no added difficulty in proving it with this greater generality.

Lemma 3.18. *Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces, let $P \in \mathrm{Pos}(\mathcal{X})$ and $Q \in \mathrm{Pos}(\mathcal{Y})$ be positive semidefinite operators, and let $X \in \mathrm{L}(\mathcal{Y}, \mathcal{X})$ be an operator. It holds that*

$$\begin{pmatrix} P & X \\ X^* & Q \end{pmatrix} \in \mathrm{Pos}(\mathcal{X} \oplus \mathcal{Y}) \tag{3.94}$$

if and only if $X = \sqrt{P}K\sqrt{Q}$ for $K \in \mathrm{L}(\mathcal{Y}, \mathcal{X})$ satisfying $\|K\| \leq 1$.

Proof. Suppose first that $X = \sqrt{P}K\sqrt{Q}$ for $K \in \mathrm{L}(\mathcal{Y}, \mathcal{X})$ being an operator for which $\|K\| \leq 1$. It follows that $KK^* \leq \mathbb{1}_{\mathcal{X}}$, and therefore

$$0 \leq \begin{pmatrix} \sqrt{P}K \\ \sqrt{Q} \end{pmatrix} \begin{pmatrix} K^*\sqrt{P} & \sqrt{Q} \end{pmatrix} = \begin{pmatrix} \sqrt{P}KK^*\sqrt{P} & X \\ X^* & Q \end{pmatrix} \leq \begin{pmatrix} P & X \\ X^* & Q \end{pmatrix}. \tag{3.95}$$

For the reverse implication, assume

$$\begin{pmatrix} P & X \\ X^* & Q \end{pmatrix} \in \mathrm{Pos}(\mathcal{X} \oplus \mathcal{Y}), \tag{3.96}$$

and define

$$K = \sqrt{P^+}X\sqrt{Q^+}. \tag{3.97}$$

It will be proved that $X = \sqrt{P}K\sqrt{Q}$ and $\|K\| \leq 1$. Observe first that, for every Hermitian operator $H \in \mathrm{Herm}(\mathcal{X})$, the block operator

$$\begin{pmatrix} H & 0 \\ 0 & \mathbb{1} \end{pmatrix} \begin{pmatrix} P & X \\ X^* & Q \end{pmatrix} \begin{pmatrix} H & 0 \\ 0 & \mathbb{1} \end{pmatrix} = \begin{pmatrix} HPH & HX \\ X^*H & Q \end{pmatrix} \tag{3.98}$$

is positive semidefinite. In particular, for $H = \Pi_{\ker(P)}$ being the projection onto the kernel of P, one has that the operator

$$\begin{pmatrix} 0 & \Pi_{\ker(P)}X \\ X^*\Pi_{\ker(P)} & Q \end{pmatrix} \tag{3.99}$$

is positive semidefinite, which implies that $\Pi_{\ker(P)}X = 0$, and therefore $\Pi_{\mathrm{im}(P)}X = X$. Through a similar argument, one finds that $X\Pi_{\mathrm{im}(Q)} = X$. It therefore follows that

$$\sqrt{P}K\sqrt{Q} = \Pi_{\mathrm{im}(P)}X\Pi_{\mathrm{im}(Q)} = X. \tag{3.100}$$

Next, note that

$$\begin{pmatrix} x^*Px & x^*Xy \\ y^*X^*x & y^*Qy \end{pmatrix} = \begin{pmatrix} x^* & 0 \\ 0 & y^* \end{pmatrix} \begin{pmatrix} P & X \\ X^* & Q \end{pmatrix} \begin{pmatrix} x & 0 \\ 0 & y \end{pmatrix} \geq 0 \tag{3.101}$$

for every choice of vectors $x \in \mathcal{X}$ and $y \in \mathcal{Y}$. Setting

$$x = \sqrt{P^+}u \quad \text{and} \quad y = \sqrt{Q^+}v \tag{3.102}$$

for arbitrarily chosen unit vectors $u \in \mathcal{X}$ and $v \in \mathcal{Y}$, one finds that

$$\begin{pmatrix} 1 & u^*Kv \\ v^*K^*u & 1 \end{pmatrix} \geq \begin{pmatrix} u^*\Pi_{\mathrm{im}(P)}u & u^*Kv \\ v^*K^*u & v^*\Pi_{\mathrm{im}(Q)}v \end{pmatrix} \geq 0 \tag{3.103}$$

and therefore $|u^*Kv| \leq 1$. As this inequality holds for all unit vectors u and v, it follows that $\|K\| \leq 1$, as required. □

Proof of Theorem 3.17. By Lemma 3.18, the expression on the right-hand side of the equation (3.93) may be written as

$$\max\left\{ \left| \mathrm{Tr}\left(\sqrt{P}K\sqrt{Q}\right) \right| \,:\, K \in \mathrm{L}(\mathcal{X}),\, \|K\| \leq 1 \right\}, \tag{3.104}$$

which is equivalent to

$$\max\left\{ \left| \left\langle K, \sqrt{P}\sqrt{Q} \right\rangle \right| \,:\, K \in \mathrm{L}(\mathcal{X}),\, \|K\| \leq 1 \right\}. \tag{3.105}$$

By the duality of the trace and spectral norms, as expressed by (1.168), one has

$$\left\| \sqrt{P}\sqrt{Q} \right\|_1 = \mathrm{F}(P, Q), \tag{3.106}$$

which completes the proof. □

Remark 3.19. For any choice of operators $P, Q \in \mathrm{Pos}(\mathcal{X})$ and $X \in \mathrm{L}(\mathcal{X})$, and a scalar $\alpha \in \mathbb{C}$ satisfying $|\alpha| = 1$, it holds that

$$\begin{pmatrix} P & X \\ X^* & Q \end{pmatrix} \in \mathrm{Pos}(\mathcal{X} \oplus \mathcal{X}) \tag{3.107}$$

if and only if

$$\begin{pmatrix} P & \alpha X \\ \overline{\alpha}X^* & Q \end{pmatrix} \in \mathrm{Pos}(\mathcal{X} \oplus \mathcal{X}). \tag{3.108}$$

This fact follows from Lemma 3.18. Alternatively, one may conclude that (3.107) implies (3.108) through the equation

$$\begin{pmatrix} \mathbb{1} & 0 \\ 0 & \alpha\mathbb{1} \end{pmatrix}^* \begin{pmatrix} P & X \\ X^* & Q \end{pmatrix} \begin{pmatrix} \mathbb{1} & 0 \\ 0 & \alpha\mathbb{1} \end{pmatrix} = \begin{pmatrix} P & \alpha X \\ \overline{\alpha}X^* & Q \end{pmatrix}, \tag{3.109}$$

while the reverse implication is obtained similarly, through the equation

$$\begin{pmatrix} \mathbb{1} & 0 \\ 0 & \alpha\mathbb{1} \end{pmatrix} \begin{pmatrix} P & \alpha X \\ \overline{\alpha} X^* & Q \end{pmatrix} \begin{pmatrix} \mathbb{1} & 0 \\ 0 & \alpha\mathbb{1} \end{pmatrix}^* = \begin{pmatrix} P & X \\ X^* & Q \end{pmatrix}. \tag{3.110}$$

For any two positive semidefinite operators $P, Q \in \mathrm{Pos}(\mathcal{X})$, it therefore holds that the fidelity $\mathrm{F}(P, Q)$ is given by the expression

$$\max\left\{ \Re(\mathrm{Tr}(X)) \,:\, X \in \mathrm{L}(\mathcal{X}),\, \begin{pmatrix} P & X \\ X^* & Q \end{pmatrix} \in \mathrm{Pos}(\mathcal{X} \oplus \mathcal{X}) \right\}. \tag{3.111}$$

Moreover, there must exist an operator $X \in \mathrm{L}(\mathcal{X})$ such that

$$\begin{pmatrix} P & X \\ X^* & Q \end{pmatrix} \in \mathrm{Pos}(\mathcal{X} \oplus \mathcal{X}) \tag{3.112}$$

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

The characterization of the fidelity function established by Theorem 3.17 provides an expression of the fidelity $\mathrm{F}(P, Q)$ corresponding to the optimal value of a semidefinite program, as will now be explained. First, define a map $\Phi \in \mathrm{T}(\mathcal{X} \oplus \mathcal{X})$ as

$$\Phi \begin{pmatrix} X_0 & \cdot \\ \cdot & X_1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} X_0 & 0 \\ 0 & X_1 \end{pmatrix} \tag{3.113}$$

for every $X_0, X_1 \in \mathrm{L}(\mathcal{X})$, where the dots represent elements of $\mathrm{L}(\mathcal{X})$ that have no influence on the output of this map. One may verify that the map Φ is self-adjoint: $\Phi = \Phi^*$. Then, for a given choice of $P, Q \in \mathrm{Pos}(\mathcal{X})$, define Hermitian operators $A, B \in \mathrm{Herm}(\mathcal{X} \oplus \mathcal{X})$ as

$$A = \frac{1}{2} \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} \quad \text{and} \quad B = \frac{1}{2} \begin{pmatrix} P & 0 \\ 0 & Q \end{pmatrix}. \tag{3.114}$$

The primal and dual optimization problems associated with the semidefinite program (Φ, A, B), after minor simplifications, are as follows:

Primal problem		Dual problem	
maximize:	$\frac{1}{2}\mathrm{Tr}(X) + \frac{1}{2}\mathrm{Tr}(X^*)$	minimize:	$\frac{1}{2}\langle P, Y_0 \rangle + \frac{1}{2}\langle Q, Y_1 \rangle$
subject to:	$\begin{pmatrix} P & X \\ X^* & Q \end{pmatrix} \geq 0,$	subject to:	$\begin{pmatrix} Y_0 & -\mathbb{1} \\ -\mathbb{1} & Y_1 \end{pmatrix} \geq 0,$
	$X \in \mathrm{L}(\mathcal{X}).$		$Y_0, Y_1 \in \mathrm{Herm}(\mathcal{X}).$

The optimal primal value of this semidefinite program is equal to $\mathrm{F}(P, Q)$, as it is in agreement with the expression (3.111).

The primal problem is evidently feasible, as one may simply take $X = 0$ to obtain a primal feasible solution. The dual problem is strictly feasible: for any choice of $Y_0 > \mathbb{1}$ and $Y_1 > \mathbb{1}$, one has that the operator

$$\begin{pmatrix} Y_0 & -\mathbb{1} \\ -\mathbb{1} & Y_1 \end{pmatrix} \tag{3.115}$$

is positive definite. Strong duality therefore follows by Slater's theorem (Theorem 1.18).

Alberti's theorem

Given that the semidefinite program for the fidelity described above possesses the property of strong duality, its dual optimum must be equal to the primal optimum $\mathrm{F}(P, Q)$. The following theorem is a consequence of this observation.

Theorem 3.20. *Let $\mathcal{X}$ be a complex Euclidean space and let $P, Q \in \mathrm{Pos}(\mathcal{X})$ be positive semidefinite operators. It holds that*

$$\mathrm{F}(P, Q) = \inf\left\{ \frac{1}{2}\langle P, Y \rangle + \frac{1}{2}\langle Q, Y^{-1} \rangle \,:\, Y \in \mathrm{Pd}(\mathcal{X}) \right\}. \tag{3.116}$$

Proof. Through the use of Lemma 3.18, it is routine to verify that the block operator

$$\begin{pmatrix} Y_0 & -\mathbb{1} \\ -\mathbb{1} & Y_1 \end{pmatrix} \tag{3.117}$$

is positive semidefinite, for a given choice of $Y_0, Y_1 \in \mathrm{Herm}(\mathcal{X})$, if and only if both Y_0 and Y_1 are positive definite and satisfy $Y_1 \geq Y_0^{-1}$. Because Q is positive semidefinite, it holds that $\langle Q, Y_1 \rangle \geq \langle Q, Y_0^{-1} \rangle$ provided $Y_0 > 0$ and $Y_1 \geq Y_0^{-1}$, so the dual problem associated to the semidefinite program (Φ, A, B) defined from P and Q as above is equivalent to a minimization of

$$\frac{1}{2}\langle P, Y \rangle + \frac{1}{2}\langle Q, Y^{-1} \rangle \tag{3.118}$$

over all positive definite operators $Y \in \mathrm{Pd}(\mathcal{X})$. As the optimal solution to this problem is equal to $\mathrm{F}(P, Q)$, the theorem follows. □

Theorem 3.20 implies the following corollary, which states a fact that is known as Alberti's theorem.[3]

Corollary 3.21 (Alberti's theorem)**.** *Let $\mathcal{X}$ be a complex Euclidean space and let $P, Q \in \mathrm{Pos}(\mathcal{X})$ be positive semidefinite operators. It holds that*

$$\mathrm{F}(P,Q)^2 = \inf\Big\{ \big\langle P, Y \big\rangle \big\langle Q, Y^{-1} \big\rangle \;:\; Y \in \mathrm{Pd}(\mathcal{X}) \Big\}. \tag{3.119}$$

Proof. If either of P or Q is zero, the corollary is trivial, so it may be taken as an assumption that neither P nor Q is zero for the remainder of the proof.

The arithmetic-geometric mean inequality implies that

$$\frac{1}{2}\langle P, Y \rangle + \frac{1}{2}\langle Q, Y^{-1} \rangle \geq \sqrt{\langle P, Y \rangle \langle Q, Y^{-1} \rangle} \tag{3.120}$$

for every operator $Y \in \mathrm{Pd}(\mathcal{X})$. By Theorem 3.20, one concludes that

$$\inf\Big\{ \big\langle P, Y \big\rangle \big\langle Q, Y^{-1} \big\rangle \;:\; Y \in \mathrm{Pd}(\mathcal{X}) \Big\} \leq \mathrm{F}(P,Q)^2. \tag{3.121}$$

On the other hand, for any choice of $Y \in \mathrm{Pd}(\mathcal{X})$, it holds that

$$\sqrt{\langle P, Y \rangle \langle Q, Y^{-1} \rangle} = \sqrt{\langle P, \alpha Y \rangle \langle Q, (\alpha Y)^{-1} \rangle} \tag{3.122}$$

for every nonzero real number $\alpha \in \mathbb{R}$. In particular, for

$$\alpha = \sqrt{\frac{\langle Q, Y^{-1} \rangle}{\langle P, Y \rangle}}, \tag{3.123}$$

which has been selected so that $\langle P, \alpha Y \rangle = \langle Q, (\alpha Y)^{-1} \rangle$, one has

$$\begin{aligned} \sqrt{\langle P, Y \rangle \langle Q, Y^{-1} \rangle} &= \sqrt{\langle P, \alpha Y \rangle \langle Q, (\alpha Y)^{-1} \rangle} \\ &= \frac{1}{2}\langle P, \alpha Y \rangle + \frac{1}{2}\langle Q, (\alpha Y)^{-1} \rangle \geq \mathrm{F}(P,Q), \end{aligned} \tag{3.124}$$

and therefore

$$\inf\Big\{ \big\langle P, Y \big\rangle \big\langle Q, Y^{-1} \big\rangle \;:\; Y \in \mathrm{Pd}(\mathcal{X}) \Big\} \geq \mathrm{F}(P,Q)^2, \tag{3.125}$$

which completes the proof. □

[3] One may also prove that Corollary 3.21 implies Theorem 3.20, so the two facts are in fact equivalent.

It is possible to prove Theorem 3.21 directly, without making use of semidefinite programming duality, as the following proof demonstrates.

Alternative proof of Theorem 3.21. The special case in which $P = Q$ will be considered first. It this case, one aims to prove

$$\inf\left\{ \frac{1}{2}\langle Y, P \rangle + \frac{1}{2}\langle Y^{-1}, P \rangle \;:\; Y \in \mathrm{Pd}(\mathcal{X}) \right\} = \mathrm{Tr}(P). \tag{3.126}$$

As $Y = \mathbb{1}$ is positive definite, it is evident that the infimum in (3.126) is at most $\mathrm{Tr}(P)$, so it suffices to prove

$$\frac{1}{2}\langle Y, P \rangle + \frac{1}{2}\langle Y^{-1}, P \rangle \geq \mathrm{Tr}(P) \tag{3.127}$$

for every choice of $Y \in \mathrm{Pd}(\mathcal{X})$. As the operator

$$\frac{Y + Y^{-1}}{2} - \mathbb{1} = \frac{1}{2}\Big(Y^{\frac{1}{2}} - Y^{-\frac{1}{2}}\Big)^2 \tag{3.128}$$

is the square of a Hermitian operator, it must be positive semidefinite, and therefore

$$\frac{1}{2}\big\langle Y + Y^{-1}, P \big\rangle \geq \langle \mathbb{1}, P \rangle = \mathrm{Tr}(P) \tag{3.129}$$

for every positive semidefinite operator P. This proves that equation (3.126) holds, and therefore proves the theorem in the special case $P = Q$.

Toward the proof of the general case, suppose that P and Q are positive definite operators. Let

$$R = \sqrt{\sqrt{P} Q \sqrt{P}}, \tag{3.130}$$

and define a mapping $\Phi \in \mathrm{CP}(\mathcal{X})$ as

$$\Phi(X) = R^{-\frac{1}{2}} \sqrt{P} X \sqrt{P} R^{-\frac{1}{2}} \tag{3.131}$$

for every $X \in \mathrm{L}(\mathcal{X})$. For $Y \in \mathrm{Pd}(\mathcal{X})$ being any positive definite operator, it holds that

$$\big\langle \Phi(Y), R \big\rangle = \langle Y, P \rangle \quad \text{and} \quad \big\langle \Phi(Y)^{-1}, R \big\rangle = \big\langle Y^{-1}, Q \big\rangle, \tag{3.132}$$

and therefore

$$\inf_{Y \in \mathrm{Pd}(\mathcal{X})} \frac{\langle Y, P \rangle + \langle Y^{-1}, Q \rangle}{2} = \inf_{Y \in \mathrm{Pd}(\mathcal{X})} \frac{\langle \Phi(Y), R \rangle + \langle \Phi(Y)^{-1}, R \rangle}{2}. \tag{3.133}$$

Observing that, as Y ranges over all positive definite operators, so too does $\Phi(Y)$, one has that

$$\inf_{Y \in \mathrm{Pd}(\mathcal{X})} \frac{\langle Y, P\rangle + \langle Y^{-1}, Q\rangle}{2} = \mathrm{Tr}(R) = \mathrm{F}(P, Q) \tag{3.134}$$

by the special case considered in the initial part of the proof.

Finally, in the most general case in which $P, Q \in \mathrm{Pos}(\mathcal{X})$ may not be invertible, the theorem follows from a continuity argument. In greater detail, for every positive real number $\varepsilon > 0$, one has

$$\frac{1}{2}\langle Y, P\rangle + \frac{1}{2}\langle Y^{-1}, Q\rangle \leq \frac{1}{2}\langle Y, P + \varepsilon\mathbb{1}\rangle + \frac{1}{2}\langle Y^{-1}, Q + \varepsilon\mathbb{1}\rangle \tag{3.135}$$

for every choice of $Y \in \mathrm{Pd}(\mathcal{X})$. Taking the infimum over all positive definite operators $Y \in \mathrm{Pd}(\mathcal{X})$ yields the inequality

$$\inf_{Y \in \mathrm{Pd}(\mathcal{X})} \frac{\langle Y, P\rangle + \langle Y^{-1}, Q\rangle}{2} \leq \mathrm{F}(P + \varepsilon\mathbb{1}, Q + \varepsilon\mathbb{1}), \tag{3.136}$$

which holds by virtue of the fact that $P + \varepsilon\mathbb{1}$ and $Q + \varepsilon\mathbb{1}$ are necessarily positive definite. As this inequality holds for all $\varepsilon > 0$, it follows from the continuity of the fidelity function that

$$\inf_{Y \in \mathrm{Pd}(\mathcal{X})} \frac{\langle Y, P\rangle + \langle Y^{-1}, Q\rangle}{2} \leq \mathrm{F}(P, Q). \tag{3.137}$$

On the other hand, for each choice of $Y \in \mathrm{Pd}(\mathcal{X})$, one has

$$\frac{1}{2}\langle Y, P + \varepsilon\mathbb{1}\rangle + \frac{1}{2}\langle Y^{-1}, Q + \varepsilon\mathbb{1}\rangle \geq \mathrm{F}(P + \varepsilon\mathbb{1}, Q + \varepsilon\mathbb{1}) \tag{3.138}$$

for all $\varepsilon > 0$, and therefore the inequality

$$\frac{1}{2}\langle Y, P\rangle + \frac{1}{2}\langle Y^{-1}, Q\rangle \geq \mathrm{F}(P, Q) \tag{3.139}$$

follows from the continuity of the expressions on the two sides of this inequality. This is so for all $Y \in \mathrm{Pd}(\mathcal{X})$, and therefore

$$\inf_{Y \in \mathrm{Pd}(\mathcal{X})} \frac{\langle Y, P\rangle + \langle Y^{-1}, Q\rangle}{2} \geq \mathrm{F}(P, Q), \tag{3.140}$$

which completes the proof. □

Uhlmann's theorem

Uhlmann's theorem establishes a link between the fidelity function and the notion of a purification of a state (or of a positive semidefinite operator more generally), providing a characterization of the fidelity function that finds many uses in the theory of quantum information. The elementary lemma that follows will be used to prove this theorem.

Lemma 3.22. *Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces and let $A, B \in \mathrm{L}(\mathcal{Y}, \mathcal{X})$ be operators. It holds that*

$$\mathrm{F}\big(AA^*, BB^*\big) = \big\|A^*B\big\|_1. \tag{3.141}$$

Proof. Using the polar decomposition, one may write

$$A = PU \quad \text{and} \quad B = QV, \tag{3.142}$$

for $P, Q \in \mathrm{Pos}(\mathcal{X})$ being positive semidefinite operators and $U, V \in \mathrm{U}(\mathcal{X})$ being unitary operators. Applying the unitary invariance of the trace norm to the definition of the fidelity function, one finds that

$$\mathrm{F}\big(AA^*, BB^*\big) = \mathrm{F}\big(P^2, Q^2\big) = \big\|PQ\big\|_1 = \big\|U^*PQV\big\|_1 = \big\|A^*B\big\|_1, \tag{3.143}$$

as required. □

Theorem 3.23 (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 having rank at most $\dim(\mathcal{Y})$, and let $u \in \mathcal{X} \otimes \mathcal{Y}$ satisfy $\mathrm{Tr}_{\mathcal{Y}}(uu^*) = P$. It holds that*

$$\mathrm{F}(P, Q) = \max\big\{|\langle u, v\rangle| \,:\, v \in \mathcal{X} \otimes \mathcal{Y},\ \mathrm{Tr}_{\mathcal{Y}}(vv^*) = Q\big\}. \tag{3.144}$$

Proof. Let $A \in \mathrm{L}(\mathcal{Y}, \mathcal{X})$ be the operator for which $u = \mathrm{vec}(A)$, let $w \in \mathcal{X} \otimes \mathcal{Y}$ be a vector satisfying $Q = \mathrm{Tr}_{\mathcal{Y}}(ww^*)$, and let $B \in \mathrm{L}(\mathcal{Y}, \mathcal{X})$ be the operator for which $w = \mathrm{vec}(B)$. It follows by the unitary equivalence of purifications (Theorem 2.11) that

$$\begin{aligned}
&\max\big\{|\langle u, v\rangle| \,:\, v \in \mathcal{X} \otimes \mathcal{Y},\ \mathrm{Tr}_{\mathcal{Y}}(vv^*) = Q\big\} \\
&\qquad = \max\big\{|\langle u, (\mathbb{1}_{\mathcal{X}} \otimes U)w\rangle| \,:\, U \in \mathrm{U}(\mathcal{Y})\big\} \\
&\qquad = \max\big\{\big|\langle A, BU^{\mathsf{T}}\rangle\big| \,:\, U \in \mathrm{U}(\mathcal{Y})\big\} \\
&\qquad = \max\big\{\big|\langle \overline{U}, A^*B\rangle\big| \,:\, U \in \mathrm{U}(\mathcal{Y})\big\} \\
&\qquad = \big\|A^*B\big\|_1.
\end{aligned} \tag{3.145}$$

By Lemma 3.22, it holds that

$$\|A^*B\|_1 = \mathrm{F}(AA^*, BB^*) = \mathrm{F}(P, Q), \tag{3.146}$$

which completes the proof. □

It will be convenient later in the chapter to make use of the following corollary, which is essentially a rephrasing of Lemma 3.22.

Corollary 3.24. *Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces and let $u, v \in \mathcal{X} \otimes \mathcal{Y}$ be vectors. It holds that*

$$\mathrm{F}\big(\mathrm{Tr}_{\mathcal{Y}}(uu^*), \mathrm{Tr}_{\mathcal{Y}}(vv^*)\big) = \big\|\mathrm{Tr}_{\mathcal{X}}(vu^*)\big\|_1. \tag{3.147}$$

Remark 3.25. Note that the partial traces on the left-hand side of (3.147) are taken over the space $\mathcal{Y}$, while the partial trace on the right-hand side is taken over $\mathcal{X}$.

Proof of Corollary 3.24. Let $A, B \in \mathrm{L}(\mathcal{Y}, \mathcal{X})$ be the operators for which it holds that $u = \mathrm{vec}(A)$ and $v = \mathrm{vec}(B)$. By Lemma 3.22, one has

$$\begin{aligned} \mathrm{F}\big(\mathrm{Tr}_{\mathcal{Y}}(uu^*), \mathrm{Tr}_{\mathcal{Y}}(vv^*)\big) &= \mathrm{F}(AA^*, BB^*) \\ &= \|A^*B\|_1 = \|(A^*B)^{\mathsf{T}}\|_1 = \big\|\mathrm{Tr}_{\mathcal{X}}(vu^*)\big\|_1 \end{aligned} \tag{3.148}$$

as required. □

Bhattacharyya coefficient characterization

The last characterization of the fidelity function to be described in this section is based on a quantity known as the *Bhattacharyya coefficient*. For any alphabet Σ, and for vectors $u, v \in [0, \infty)^\Sigma$ having nonnegative real number entries, the Bhattacharyya coefficient $\mathrm{B}(u, v)$ of u and v is defined as

$$\mathrm{B}(u, v) = \sum_{a \in \Sigma} \sqrt{u(a)}\sqrt{v(a)}. \tag{3.149}$$

The Bhattacharyya coefficient relates to the fidelity between commuting operators in a straightforward way, as the following proposition describes.

Proposition 3.26. *Let $\mathcal{X}$ be a complex Euclidean space and let $P, Q \in \mathrm{Pos}(\mathcal{X})$ be positive semdefinite operators satisfying $[P, Q] = 0$. It holds that*

$$\mathrm{F}(P, Q) = \mathrm{B}(\lambda(P), \lambda(Q)). \tag{3.150}$$

Proof. Let $n = \dim(\mathcal{X})$. Given that P and Q commute, and are positive semidefinite (and therefore normal) operators, Theorem 1.5 implies that there must exist an orthonormal basis $\{x_1, \ldots, x_n\}$ of $\mathcal{X}$ for which

$$P = \sum_{k=1}^{n} \lambda_k(P)\, x_k x_k^* \quad \text{and} \quad Q = \sum_{k=1}^{n} \lambda_k(Q)\, x_k x_k^*. \tag{3.151}$$

It therefore holds that

$$\sqrt{P}\sqrt{Q} = \sum_{k=1}^{n} \sqrt{\lambda_k(P)}\sqrt{\lambda_k(Q)}\, x_k x_k^*, \tag{3.152}$$

so that

$$\mathrm{F}(P, Q) = \left\|\sqrt{P}\sqrt{Q}\right\|_1 = \sum_{k=1}^{n} \sqrt{\lambda_k(P)}\sqrt{\lambda_k(Q)} = \mathrm{B}(\lambda(P), \lambda(Q)), \tag{3.153}$$

as required. □

There exists a more interesting connection between the Bhattacharyya coefficient and the fidelity function, concerning the measurement statistics generated from arbitrary pairs of states. The following notation is helpful when explaining this connection: for a measurement $\mu : \Sigma \to \mathrm{Pos}(\mathcal{X})$ and positive semidefinite operators $P, Q \in \mathrm{Pos}(\mathcal{X})$, one defines

$$\mathrm{B}(P, Q \,|\, \mu) = \sum_{a \in \Sigma} \sqrt{\langle \mu(a), P\rangle}\sqrt{\langle \mu(a), Q\rangle}. \tag{3.154}$$

Equivalently,

$$\mathrm{B}(P, Q \,|\, \mu) = \mathrm{B}(u, v) \tag{3.155}$$

for $u, v \in [0, \infty)^\Sigma$ being the vectors defined as

$$u(a) = \langle \mu(a), P\rangle \quad \text{and} \quad v(a) = \langle \mu(a), Q\rangle \tag{3.156}$$

for each $a \in \Sigma$.

Theorem 3.27. *Let $\mathcal{X}$ be a complex Euclidean space, let Σ be an alphabet, and let $P, Q \in \mathrm{Pos}(\mathcal{X})$ be positive semidefinite operators. For every choice of a measurement $\mu : \Sigma \to \mathrm{Pos}(\mathcal{X})$, it holds that*

$$\mathrm{F}(P, Q) \leq \mathrm{B}(P, Q \,|\, \mu). \tag{3.157}$$

Moreover, if it is the case that $|\Sigma| \geq \dim(\mathcal{X})$, then there exists a measurement $\mu : \Sigma \to \mathrm{Pos}(\mathcal{X})$ for which equality holds in (3.157).

Proof. Assume first that $\mu : \Sigma \to \mathrm{Pos}(\mathcal{X})$ is an arbitrary measurement, and let $U \in \mathrm{U}(\mathcal{X})$ be a unitary operator satisfying

$$\mathrm{F}(P,Q) = \left\| \sqrt{P}\sqrt{Q} \right\|_1 = \left\langle U, \sqrt{P}\sqrt{Q} \right\rangle. \tag{3.158}$$

By the triangle inequality followed by the Cauchy–Schwarz inequality, one finds that

$$\begin{aligned} \mathrm{F}(P,Q) &= \left\langle U, \sqrt{P}\sqrt{Q} \right\rangle = \sum_{a\in\Sigma} \left\langle U, \sqrt{P}\mu(a)\sqrt{Q} \right\rangle \\ &\leq \sum_{a\in\Sigma} \left| \left\langle \sqrt{\mu(a)}\sqrt{P}U, \sqrt{\mu(a)}\sqrt{Q} \right\rangle \right| \\ &\leq \sum_{a\in\Sigma} \sqrt{\langle \mu(a), P\rangle}\sqrt{\langle \mu(a), Q\rangle} = \mathrm{B}(P,Q\,|\,\mu). \end{aligned} \tag{3.159}$$

Next, it will be proved, under the assumption $|\Sigma| \geq \dim(\mathcal{X})$, that there exists a measurement $\mu : \Sigma \to \mathrm{Pos}(\mathcal{X})$ for which $\mathrm{F}(P,Q) = \mathrm{B}(P,Q\,|\,\mu)$. It suffices to prove that there is a measurement $\mu : \{1,\ldots,n\} \to \mathrm{Pos}(\mathcal{X})$ for which $\mathrm{F}(P,Q) = \mathrm{B}(P,Q\,|\,\mu)$, for $n = \dim(\mathcal{X})$.

Consider first the case in which P is invertible. Define

$$R = P^{-\frac{1}{2}} \left(\sqrt{P}Q\sqrt{P} \right)^{\frac{1}{2}} P^{-\frac{1}{2}}, \tag{3.160}$$

and let

$$R = \sum_{k=1}^{n} \lambda_k(R) u_k u_k^* \tag{3.161}$$

be a spectral decomposition of R. One may verify that $Q = RPR$, from which it follows that

$$\begin{aligned} &\sum_{k=1}^{n} \sqrt{\langle u_k u_k^*, P\rangle}\sqrt{\langle u_k u_k^*, Q\rangle} = \sum_{k=1}^{n} \sqrt{\langle u_k u_k^*, P\rangle}\sqrt{\langle u_k u_k^*, RPR\rangle} \\ &\quad = \sum_{k=1}^{n} \lambda_k(R)\langle u_k u_k^*, P\rangle = \langle R, P\rangle = \mathrm{Tr}\left(\sqrt{\sqrt{P}Q\sqrt{P}} \right) = \mathrm{F}(P,Q). \end{aligned} \tag{3.162}$$

The measurement $\mu : \{1,\ldots,n\} \to \mathrm{Pos}(\mathcal{X})$ defined by

$$\mu(k) = u_k u_k^* \tag{3.163}$$

for each $k \in \{1,\ldots,n\}$ therefore satisfies $\mathrm{F}(P,Q) = \mathrm{B}(P,Q\,|\,\mu)$.

Finally, the case in which $r = \mathrm{rank}(P) < n$ will be considered. Let $\Pi = \Pi_{\mathrm{im}(P)}$ denote the projection onto the image of P. By restricting one's attention to this subspace, the argument above may be seen to imply the existence of an orthonormal basis $\{u_1,\ldots,u_r\}$ for $\mathrm{im}(P)$ that satisfies

$$\mathrm{F}(P,\Pi Q\Pi) = \sum_{k=1}^{r} \sqrt{\langle u_k u_k^*, P\rangle}\sqrt{\langle u_k u_k^*, \Pi Q\Pi\rangle}. \tag{3.164}$$

Let $\{u_1,\ldots,u_n\}$ be any orthonormal basis of $\mathcal{X}$ that is obtained by completing the orthonormal set $\{u_1,\ldots,u_r\}$. As $\langle u_k u_k^*, P\rangle = 0$ for $k > r$ and $\langle u_k u_k^*, \Pi Q\Pi\rangle = \langle u_k u_k^*, Q\rangle$ for $k \leq r$, it follows that

$$\begin{aligned} &\sum_{k=1}^{n} \sqrt{\langle u_k u_k^*, P\rangle}\sqrt{\langle u_k u_k^*, Q\rangle} \\ &\quad = \sum_{k=1}^{r} \sqrt{\langle u_k u_k^*, P\rangle}\sqrt{\langle u_k u_k^*, \Pi Q\Pi\rangle} = \mathrm{F}(P,\Pi Q\Pi) = \mathrm{F}(P,Q), \end{aligned} \tag{3.165}$$

where the final equality holds by statement 4 of Proposition 3.12. Once again, the measurement $\mu : \{1,\ldots,n\} \to \mathrm{Pos}(\mathcal{X})$ defined by (3.163) for each $k \in \{1,\ldots,n\}$ satisfies $\mathrm{F}(P,Q) = \mathrm{B}(P,Q\,|\,\mu)$, which completes the proof. □

3.2.3 Further properties of the fidelity function

Various properties of the fidelity function can be established by means of the alternative characterizations presented in Section 3.2.2.

Joint concavity and monotonicity under the action of channels

The next theorem may be proved using the block operator characterization of the fidelity function (Theorem 3.17). As a corollary of this theorem, one finds that the fidelity function is *jointly concave* in its arguments.

Theorem 3.28. *Let $P_0, P_1, Q_0, Q_1 \in \mathrm{Pos}(\mathcal{X})$ be positive semidefinite operators, for $\mathcal{X}$ being a complex Euclidean space. It holds that*

$$\mathrm{F}(P_0 + P_1, Q_0 + Q_1) \geq \mathrm{F}(P_0, Q_0) + \mathrm{F}(P_1, Q_1). \tag{3.166}$$

Proof. By Theorem 3.17 (together with Remark 3.19), one may choose operators $X_0, X_1 \in \mathrm{L}(\mathcal{X})$ such that the block operators

$$\begin{pmatrix} P_0 & X_0 \\ X_0^* & Q_0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} P_1 & X_1 \\ X_1^* & Q_1 \end{pmatrix} \tag{3.167}$$

are both positive semidefinite, and such that

$$\mathrm{Tr}(X_0) = \mathrm{F}(P_0, Q_0) \quad \text{and} \quad \mathrm{Tr}(X_1) = \mathrm{F}(P_1, Q_1). \tag{3.168}$$

The sum of two positive semidefinite operators is positive semidefinite, and therefore

$$\begin{pmatrix} P_0 + P_1 & X_0 + X_1 \\ (X_0 + X_1)^* & Q_0 + Q_1 \end{pmatrix} = \begin{pmatrix} P_0 & X_0 \\ X_0^* & Q_0 \end{pmatrix} + \begin{pmatrix} P_1 & X_1 \\ X_1^* & Q_1 \end{pmatrix} \tag{3.169}$$

is positive semidefinite. Applying Theorem 3.17 again, one finds that

$$\mathrm{F}(P_0 + P_1, Q_0 + Q_1) \geq |\mathrm{Tr}(X_0 + X_1)| = \mathrm{F}(P_0, Q_0) + \mathrm{F}(P_1, Q_1), \tag{3.170}$$

as required. □

Corollary 3.29 (Joint concavity of the fidelity function)**.** *Let $\mathcal{X}$ be a complex Euclidean space, let $\rho_0, \rho_1, \sigma_0, \sigma_1 \in \mathrm{D}(\mathcal{X})$ be density operators, and let $\lambda \in [0, 1]$. It holds that*

$$\begin{aligned} &\mathrm{F}(\lambda\rho_0 + (1-\lambda)\rho_1, \lambda\sigma_0 + (1-\lambda)\sigma_1) \\ &\qquad \geq \lambda\,\mathrm{F}(\rho_0, \sigma_0) + (1-\lambda)\,\mathrm{F}(\rho_1, \sigma_1). \end{aligned} \tag{3.171}$$

Proof. By Theorem 3.28, together with statement 3 of Proposition 3.12, it holds that

$$\begin{aligned} &\mathrm{F}(\lambda\rho_0 + (1-\lambda)\rho_1, \lambda\sigma_0 + (1-\lambda)\sigma_1) \\ &\qquad \geq \mathrm{F}(\lambda\rho_0, \lambda\sigma_0) + \mathrm{F}((1-\lambda)\rho_1, (1-\lambda)\sigma_1) \\ &\qquad = \lambda\,\mathrm{F}(\rho_0, \sigma_0) + (1-\lambda)\,\mathrm{F}(\rho_1, \sigma_1), \end{aligned} \tag{3.172}$$

as claimed. □

The joint concavity of the fidelity function implies that the fidelity function is concave in each of its arguments individually:

$$\mathrm{F}(\lambda\rho_0 + (1-\lambda)\rho_1, \sigma) \geq \lambda\,\mathrm{F}(\rho_0, \sigma) + (1-\lambda)\,\mathrm{F}(\rho_1, \sigma) \tag{3.173}$$

for all $\rho_0, \rho_1, \sigma \in \mathrm{D}(\mathcal{X})$ and $\lambda \in [0, 1]$, and similar for concavity in the second argument rather than the first.

The *monotonicity* of the fidelity function under the action of channels is another fundamental property that may be established using the block operator characterization.

Theorem 3.30. *Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces, let $P, Q \in \mathrm{Pos}(\mathcal{X})$, and let $\Phi \in \mathrm{C}(\mathcal{X}, \mathcal{Y})$ be a channel. It holds that*

$$\mathrm{F}(P, Q) \leq \mathrm{F}(\Phi(P), \Phi(Q)). \tag{3.174}$$

Proof. By Theorem 3.17, one may choose $X \in \mathrm{L}(\mathcal{X})$ so that

$$\begin{pmatrix} P & X \\ X^* & Q \end{pmatrix} \tag{3.175}$$

is positive semidefinite and satisfies $|\mathrm{Tr}(X)| = \mathrm{F}(P, Q)$. By the complete positivity of Φ, the block operator

$$\begin{pmatrix} \Phi(P) & \Phi(X) \\ \Phi(X^*) & \Phi(Q) \end{pmatrix} = \begin{pmatrix} \Phi(P) & \Phi(X) \\ \Phi(X)^* & \Phi(Q) \end{pmatrix} \tag{3.176}$$

is positive semidefinite as well. Invoking Theorem 3.17 again, and using the fact that Φ is trace-preserving, it follows that

$$\mathrm{F}(\Phi(P), \Phi(Q)) \geq |\mathrm{Tr}(\Phi(X))| = |\mathrm{Tr}(X)| = \mathrm{F}(P, Q), \tag{3.177}$$

as required. □

Fidelity between extensions of operators

Suppose, for a given choice of complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$, that $P_0, P_1 \in \mathrm{Pos}(\mathcal{X})$ are positive semidefinite operators and $Q_0 \in \mathrm{Pos}(\mathcal{X} \otimes \mathcal{Y})$ is a positive semidefinite operator that extends P_0, meaning that $\mathrm{Tr}_{\mathcal{Y}}(Q_0) = P_0$. For every operator $Q_1 \in \mathrm{Pos}(\mathcal{X} \otimes \mathcal{Y})$ satisfying $\mathrm{Tr}_{\mathcal{Y}}(Q_1) = P_1$, it follows from Theorem 3.30 that

$$\mathrm{F}(Q_0, Q_1) \leq \mathrm{F}(\mathrm{Tr}_{\mathcal{Y}}(Q_0), \mathrm{Tr}_{\mathcal{Y}}(Q_1)) = \mathrm{F}(P_0, P_1). \tag{3.178}$$

It is natural, in some situations, to consider the maximum value that the fidelity $\mathrm{F}(Q_0, Q_1)$ may take, over all choices of an operator $Q_1 \in \mathrm{Pos}(\mathcal{X} \otimes \mathcal{Y})$ extending P_1. As the following theorem establishes, this maximum value is necessarily equal to $\mathrm{F}(P_0, P_1)$, irrespective of the choice of Q_0.

Theorem 3.31. *Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces, let $P_0, P_1 \in \mathrm{Pos}(\mathcal{X})$ be positive semidefinite operators, and let $Q_0 \in \mathrm{Pos}(\mathcal{X} \otimes \mathcal{Y})$ satisfy $\mathrm{Tr}_{\mathcal{Y}}(Q_0) = P_0$. It holds that*

$$\max\bigl\{\mathrm{F}(Q_0, Q_1) : Q_1 \in \mathrm{Pos}(\mathcal{X} \otimes \mathcal{Y}),\ \mathrm{Tr}_{\mathcal{Y}}(Q_1) = P_1\bigr\} = \mathrm{F}(P_0, P_1). \tag{3.179}$$

Proof. Let $\mathcal{Z}$ be a complex Euclidean space with $\dim(\mathcal{Z}) = \dim(\mathcal{X} \otimes \mathcal{Y})$, and choose any vector $u_0 \in \mathcal{X} \otimes \mathcal{Y} \otimes \mathcal{Z}$ satisfying

$$\operatorname{Tr}_{\mathcal{Z}}(u_0 u_0^*) = Q_0. \tag{3.180}$$

As Q_0 is an extension of P_0, it follows that

$$\operatorname{Tr}_{\mathcal{Y}\otimes\mathcal{Z}}(u_0 u_0^*) = P_0. \tag{3.181}$$

By Uhlmann's theorem (Theorem 3.23), there exists a vector $u_1 \in \mathcal{X} \otimes \mathcal{Y} \otimes \mathcal{Z}$ so that

$$\operatorname{Tr}_{\mathcal{Y}\otimes\mathcal{Z}}(u_1 u_1^*) = P_1 \quad \text{and} \quad |\langle u_0, u_1 \rangle| = \mathrm{F}(P_0, P_1). \tag{3.182}$$

By setting

$$Q_1 = \operatorname{Tr}_{\mathcal{Z}}(u_1 u_1^*) \tag{3.183}$$

and applying Theorem 3.30 (for the channel being the partial trace over $\mathcal{Z}$), one has

$$\begin{aligned} \mathrm{F}(Q_0, Q_1) &= \mathrm{F}(\operatorname{Tr}_{\mathcal{Z}}(u_0 u_0^*), \operatorname{Tr}_{\mathcal{Z}}(u_1 u_1^*)) \\ &\geq \mathrm{F}(u_0 u_0^*, u_1 u_1^*) = |\langle u_0, u_1 \rangle| = \mathrm{F}(P_0, P_1). \end{aligned} \tag{3.184}$$

This demonstrates that the maximum in (3.179) is at least $\mathrm{F}(P_0, P_1)$. The maximum is at most $\mathrm{F}(P_0, P_1)$ by (3.178), and so the proof is complete. □

A sum-of-squares relationship for fidelity

The next theorem states a useful fact relating the fidelity between two fixed states and the sum of the squared-fidelities between these two states and a third.

Theorem 3.32. *Let $\mathcal{X}$ be a complex Euclidean space and let $\rho_0, \rho_1 \in \mathrm{D}(\mathcal{X})$ be density operators. It holds that*

$$\max_{\sigma \in \mathrm{D}(\mathcal{X})} \Big(\mathrm{F}(\rho_0, \sigma)^2 + \mathrm{F}(\rho_1, \sigma)^2 \Big) = 1 + \mathrm{F}(\rho_0, \rho_1). \tag{3.185}$$

Proof. The proof will make use of the fact that, for any two unit vectors u_0 and u_1, chosen from an arbitrary complex Euclidean space, there is a simple closed-form expression for the largest eigenvalue of the sum of the rank-one projections corresponding to these vectors:

$$\lambda_1\big(u_0 u_0^* + u_1 u_1^*\big) = 1 + \big|\langle u_0, u_1 \rangle\big|. \tag{3.186}$$

There are two steps of the proof, both of which combine the expression (3.186) with Uhlmann's theorem (Theorem 3.23). The first step proves the existence of a density operator $\sigma \in \mathrm{D}(\mathcal{X})$ such that

$$\mathrm{F}(\rho_0, \sigma)^2 + \mathrm{F}(\rho_1, \sigma)^2 \geq 1 + \mathrm{F}(\rho_0, \rho_1). \tag{3.187}$$

Let $\mathcal{Y}$ be a space with $\dim(\mathcal{Y}) = \dim(\mathcal{X})$, and let $u_0, u_1 \in \mathcal{X} \otimes \mathcal{Y}$ be vectors satisfying the following equations:

$$\begin{aligned} \operatorname{Tr}_{\mathcal{Y}}(u_0 u_0^*) &= \rho_0, \\ \operatorname{Tr}_{\mathcal{Y}}(u_1 u_1^*) &= \rho_1, \\ |\langle u_0, u_1 \rangle| &= \mathrm{F}(\rho_0, \rho_1). \end{aligned} \tag{3.188}$$

The fact that there exists such a choice of vectors follows from Uhlmann's theorem.

Let $v \in \mathcal{X} \otimes \mathcal{Y}$ be a unit eigenvector of the operator $u_0 u_0^* + u_1 u_1^*$ that corresponds to its largest eigenvalue, so that

$$v^*\big(u_0 u_0^* + u_1 u_1^*\big) v = 1 + |\langle u_0, u_1 \rangle|, \tag{3.189}$$

and let

$$\sigma = \operatorname{Tr}_{\mathcal{Y}}(v v^*). \tag{3.190}$$

Using Uhlmann's theorem again, one has

$$\mathrm{F}(\rho_0, \sigma) \geq |\langle u_0, v \rangle| \quad \text{and} \quad \mathrm{F}(\rho_1, \sigma) \geq |\langle u_1, v \rangle|, \tag{3.191}$$

so that

$$\begin{aligned} \mathrm{F}(\rho_0, \sigma)^2 + \mathrm{F}(\rho_1, \sigma)^2 &\geq v^*\big(u_0 u_0^* + u_1 u_1^*\big) v \\ &= 1 + \big|\langle u_0, u_1 \rangle\big| = 1 + \mathrm{F}(\rho_0, \rho_1), \end{aligned} \tag{3.192}$$

which proves the required inequality.

The second step of the proof is to establish that the inequality

$$\mathrm{F}(\rho_0, \sigma)^2 + \mathrm{F}(\rho_1, \sigma)^2 \leq 1 + \mathrm{F}(\rho_0, \rho_1) \tag{3.193}$$

holds for every $\sigma \in \mathrm{D}(\mathcal{X})$. Again, let $\mathcal{Y}$ be a complex Euclidean space with $\dim(\mathcal{Y}) = \dim(\mathcal{X})$, let $\sigma \in \mathrm{D}(\mathcal{X})$ be chosen arbitrarily, and choose $v \in \mathcal{X} \otimes \mathcal{Y}$ to be any unit vector satisfying

$$\operatorname{Tr}_{\mathcal{Y}}(v v^*) = \sigma. \tag{3.194}$$

Also let $u_0, u_1 \in \mathcal{X} \otimes \mathcal{Y}$ be unit vectors satisfying the following equations:

$$\begin{aligned} \operatorname{Tr}_{\mathcal{Y}}(u_0 u_0^*) &= \rho_0, \\ \operatorname{Tr}_{\mathcal{Y}}(u_1 u_1^*) &= \rho_1, \\ |\langle u_0, v\rangle| &= \mathrm{F}(\rho_0, \sigma), \\ |\langle u_1, v\rangle| &= \mathrm{F}(\rho_1, \sigma). \end{aligned} \tag{3.195}$$

As in the first step of the proof, the existence of such vectors is implied by Uhlmann's theorem. As v is a unit vector, it holds that

$$\begin{aligned} v^*(u_0 u_0^* + u_1 u_1^*)v &\leq \lambda_1(u_0 u_0^* + u_1 u_1^*) \\ &= 1 + |\langle u_0, u_1\rangle| \leq 1 + \mathrm{F}(\rho_0, \rho_1), \end{aligned} \tag{3.196}$$

where the last inequality is, once again, implied by Uhlmann's theorem. Therefore, one has

$$\mathrm{F}(\rho_0, \sigma)^2 + \mathrm{F}(\rho_1, \sigma)^2 = v^*(u_0 u_0^* + u_1 u_1^*)v \leq 1 + \mathrm{F}(\rho_0, \rho_1), \tag{3.197}$$

as required. □

Fidelity between inputs and outputs of completely positive maps

With respect to the storage and transmission of quantum information, the identity map represents an ideal quantum channel, as this channel causes no disturbance to the quantum states it acts upon. For this reason, it may be desirable to measure the similarity between a given channel of the form $\Phi \in \mathrm{C}(\mathcal{X})$ and the identity channel $\mathbb{1}_{\mathrm{L}(\mathcal{X})}$ in some settings.

One setting in which such a comparison is made arises in connection with quantum source coding (to be discussed in Section 5.3.2). Here, one is interested in the fidelity between the input and output states of a given channel $\Phi \in \mathrm{C}(\mathcal{X})$, under the assumption that the channel acts on a state $\sigma \in \mathrm{D}(\mathcal{X} \otimes \mathcal{Y})$ that extends a known fixed state $\rho \in \mathrm{D}(\mathcal{X})$. The *mapping fidelity*, which is specified by the following definition, is representative of this situation when σ is taken as a purification of the state ρ.

Definition 3.33. Let $\mathcal{X}$ be a complex Euclidean space, let $\Phi \in \mathrm{CP}(\mathcal{X})$ be a completely positive map, and let $P \in \mathrm{Pos}(\mathcal{X})$ be a positive semidefinite operator. The *mapping fidelity* of Φ with respect to P is defined as

$$\mathrm{F}(\Phi, P) = \mathrm{F}\big(uu^*, \big(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})}\big)(uu^*)\big) \tag{3.198}$$

for $u = \operatorname{vec}\big(\sqrt{P}\big)$.

The mapping fidelity is also called the *channel fidelity* when Φ is a channel and $P = \rho$ is a density operator. (It is also commonly called the *entanglement fidelity* in this case, although that terminology will not be used in this book.)

An explicit formula for the mapping fidelity $\mathrm{F}(\Phi, P)$, from any Kraus representation of the mapping Φ, is given by the following proposition.

Proposition 3.34. *Let $\mathcal{X}$ be a complex Euclidean space, let $\Phi \in \mathrm{CP}(\mathcal{X})$ be a completely positive map, and assume that a Kraus representation of Φ is given:*

$$\Phi(X) = \sum_{a \in \Sigma} A_a X A_a^* \tag{3.199}$$

for all $X \in \mathrm{L}(\mathcal{X})$, for $\{A_a : a \in \Sigma\} \subset \mathrm{L}(\mathcal{X})$ being a collection of operators indexed by some alphabet Σ. For every operator $P \in \mathrm{Pos}(\mathcal{X})$, it holds that

$$\mathrm{F}(\Phi, P) = \sqrt{\sum_{a \in \Sigma} \big|\langle P, A_a\rangle\big|^2}. \tag{3.200}$$

Proof. Using Proposition 3.13, one may evaluate the expression (3.198) to obtain

$$\begin{aligned} \mathrm{F}(\Phi, P) &= \sqrt{\sum_{a \in \Sigma} \big|\operatorname{vec}\big(\sqrt{P}\big)^* (A_a \otimes \mathbb{1}_{\mathcal{X}}) \operatorname{vec}\big(\sqrt{P}\big)\big|^2} \\ &= \sqrt{\sum_{a \in \Sigma} \big|\big\langle \sqrt{P}, A_a \sqrt{P}\big\rangle\big|^2} = \sqrt{\sum_{a \in \Sigma} \big|\langle P, A_a\rangle\big|^2}, \end{aligned} \tag{3.201}$$

as required. □

As the next proposition implies, the purification $u = \operatorname{vec}\big(\sqrt{P}\big)$ taken in the definition of the mapping fidelity is representative of a worst case scenario. That is, for an arbitrary state $\sigma \in \mathrm{D}(\mathcal{X} \otimes \mathcal{Y})$ that extends a known fixed state $\rho \in \mathrm{D}(\mathcal{X})$, the fidelity

$$\mathrm{F}\big(\sigma, (\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})})(\sigma)\big) \tag{3.202}$$

can be no smaller than the mapping fidelity $\mathrm{F}(\Phi, \rho)$.

Proposition 3.35. *Let $\mathcal{X}$ be a complex Euclidean space, let $\Phi \in \mathrm{CP}(\mathcal{X})$ be a completely positive map, and let $P \in \mathrm{Pos}(\mathcal{X})$ be a positive semidefinite operator. Suppose further that $\mathcal{Y}$ and $\mathcal{Z}$ are complex Euclidean spaces, $u \in \mathcal{X} \otimes \mathcal{Y}$ is a vector satisfying $\operatorname{Tr}_{\mathcal{Y}}(uu^*) = P$, and $Q \in \mathrm{Pos}(\mathcal{X} \otimes \mathcal{Z})$ is an operator satisfying $\operatorname{Tr}_{\mathcal{Z}}(Q) = P$. It holds that*

$$\mathrm{F}(Q, (\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})})(Q)) \geq \mathrm{F}(uu^*, (\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})})(uu^*)). \tag{3.203}$$

Proof. By Proposition 2.29, there must exist a channel $\Psi \in \mathrm{C}(\mathcal{Y}, \mathcal{Z})$ such that

$$\big(\mathbb{1}_{\mathrm{L}(\mathcal{X})} \otimes \Psi\big)(uu^*) = Q. \tag{3.204}$$

By Theorem 3.30, one has

$$\begin{aligned} &\mathrm{F}\big(uu^*, (\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})})(uu^*)\big) \\ &\quad \leq \mathrm{F}\big(\big(\mathbb{1}_{\mathrm{L}(\mathcal{X})} \otimes \Psi\big)(uu^*), (\Phi \otimes \Psi)(uu^*)\big) \\ &\quad = \mathrm{F}(Q, (\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})})(Q)), \end{aligned} \tag{3.205}$$

which completes the proof. □

It is also evident from this proposition that taking any other purification of P in place of $u = \mathrm{vec}\big(\sqrt{P}\big)$ in Definition 3.33 would yield precisely the same value.

Fuchs–van de Graaf inequalities

The final property of the fidelity function to be established in this section concerns its connection to the trace distance between quantum states. This is an important relationship, as it allows for an approximate conversion between the more operationally motivated trace distance and the often more analytically robust fidelity function evaluated on a given pair of states.

Theorem 3.36 (Fuchs–van de Graaf inequalities)**.** *Let $\mathcal{X}$ be a complex Euclidean space and let $\rho, \sigma \in \mathrm{D}(\mathcal{X})$ be density operators. It holds that*

$$1 - \frac{1}{2}\big\|\rho - \sigma\big\|_1 \leq \mathrm{F}(\rho, \sigma) \leq \sqrt{1 - \frac{1}{4}\big\|\rho - \sigma\big\|_1^2}. \tag{3.206}$$

Equivalently,

$$2 - 2\,\mathrm{F}(\rho, \sigma) \leq \big\|\rho - \sigma\big\|_1 \leq 2\sqrt{1 - \mathrm{F}(\rho, \sigma)^2}. \tag{3.207}$$

Proof. The two inequalities in (3.207) will be established separately, beginning with the first. By Theorem 3.27, there exists an alphabet Σ and a measurement $\mu : \Sigma \to \mathrm{Pos}(\mathcal{X})$ such that

$$\mathrm{F}(\rho, \sigma) = \mathrm{B}(\rho, \sigma \,|\, \mu). \tag{3.208}$$

Fix such a measurement, and define probability vectors $p, q \in \mathcal{P}(\Sigma)$ as

$$p(a) = \langle \mu(a), \rho \rangle \quad \text{and} \quad q(a) = \langle \mu(a), \sigma \rangle \tag{3.209}$$

for each $a \in \Sigma$, so that $\mathrm{B}(p, q) = \mathrm{F}(\rho, \sigma)$. By Proposition 3.5, together with the observation that

$$\big(\sqrt{\alpha} - \sqrt{\beta}\big)^2 \leq |\alpha - \beta| \tag{3.210}$$

for every choice of nonnegative real numbers $\alpha, \beta \geq 0$, it follows that

$$\begin{aligned} \big\|\rho - \sigma\big\|_1 &\geq \big\|p - q\big\|_1 = \sum_{a \in \Sigma} |p(a) - q(a)| \\ &\geq \sum_{a \in \Sigma} \Big(\sqrt{p(a)} - \sqrt{q(a)}\Big)^2 = 2 - 2\,\mathrm{B}(p, q) = 2 - 2\,\mathrm{F}(\rho, \sigma). \end{aligned} \tag{3.211}$$

The first inequality in (3.207) is therefore proved.

Next, the second inequality in (3.207) will be proved. Let $\mathcal{Y}$ be a complex Euclidean space with $\dim(\mathcal{Y}) = \dim(\mathcal{X})$. It follows by Uhlmann's theorem (Theorem 3.23) that there exists a choice of unit vectors $u, v \in \mathcal{X} \otimes \mathcal{Y}$ satisfying the equations

$$\begin{aligned} \mathrm{Tr}_{\mathcal{Y}}(uu^*) &= \rho, \\ \mathrm{Tr}_{\mathcal{Y}}(vv^*) &= \sigma, \end{aligned} \tag{3.212}$$

and

$$|\langle u, v \rangle| = \mathrm{F}(\rho, \sigma). \tag{3.213}$$

By the identity (1.181), it holds that

$$\big\|uu^* - vv^*\big\|_1 = 2\sqrt{1 - |\langle u, v \rangle|^2} = 2\sqrt{1 - \mathrm{F}(\rho, \sigma)^2}. \tag{3.214}$$

Consequently, by the monotonicity of the trace norm under partial tracing (1.178), one has

$$\big\|\rho - \sigma\big\|_1 \leq \big\|uu^* - vv^*\big\|_1 = 2\sqrt{1 - \mathrm{F}(\rho, \sigma)^2}. \tag{3.215}$$

The second inequality in (3.207) has been established, which completes the proof. □

The use of the Bhattacharyya coefficient characterization of the fidelity (Theorem 3.27) in the above proof may be substituted by the following operator norm inequality, which is a useful inequality in its own right.

Lemma 3.37. *Let $\mathcal{X}$ be a complex Euclidean space and let $P_0, P_1 \in \mathrm{Pos}(\mathcal{X})$ be positive semidefinite operators. It holds that*

$$\big\|P_0 - P_1\big\|_1 \geq \Big\|\sqrt{P_0} - \sqrt{P_1}\Big\|_2^2. \tag{3.216}$$

Proof. Let

$$\sqrt{P_0} - \sqrt{P_1} = Q_0 - Q_1, \tag{3.217}$$

for $Q_0, Q_1 \in \mathrm{Pos}(\mathcal{X})$, be the Jordan-Hahn decomposition of $\sqrt{P_0} - \sqrt{P_1}$, and let $\Pi_0 = \mathrm{im}(Q_0)$ and $\Pi_1 = \mathrm{im}(Q_1)$. The operator $\Pi_0 - \Pi_1$ has spectral norm at most 1, and therefore

$$\left\| P_0 - P_1 \right\|_1 \geq \left\langle \Pi_0 - \Pi_1, P_0 - P_1 \right\rangle. \tag{3.218}$$

Through the use of the operator identity

$$A^2 - B^2 = \frac{1}{2}(A - B)(A + B) + \frac{1}{2}(A + B)(A - B), \tag{3.219}$$

one finds that

$$\begin{aligned}
&\langle \Pi_0 - \Pi_1, P_0 - P_1 \rangle \\
&\quad = \frac{1}{2}\Big\langle \Pi_0 - \Pi_1, \Big(\sqrt{P_0} - \sqrt{P_1}\Big)\Big(\sqrt{P_0} + \sqrt{P_1}\Big)\Big\rangle \\
&\qquad\quad + \frac{1}{2}\Big\langle \Pi_0 - \Pi_1, \Big(\sqrt{P_0} + \sqrt{P_1}\Big)\Big(\sqrt{P_0} - \sqrt{P_1}\Big)\Big\rangle \\
&\quad = \frac{1}{2}\,\mathrm{Tr}\Big((Q_0 + Q_1)\Big(\sqrt{P_0} + \sqrt{P_1}\Big)\Big) \\
&\qquad\quad + \frac{1}{2}\,\mathrm{Tr}\Big(\Big(\sqrt{P_0} + \sqrt{P_1}\Big)(Q_0 + Q_1)\Big) \\
&\quad = \Big\langle Q_0 + Q_1, \sqrt{P_0} + \sqrt{P_1}\Big\rangle.
\end{aligned} \tag{3.220}$$

Finally, as Q_0, Q_1, $\sqrt{P_0}$, and $\sqrt{P_1}$ are positive semidefinite, one has

$$\begin{aligned}
&\Big\langle Q_0 + Q_1, \sqrt{P_0} + \sqrt{P_1}\Big\rangle \\
&\qquad \geq \Big\langle Q_0 - Q_1, \sqrt{P_0} - \sqrt{P_1}\Big\rangle = \Big\| \sqrt{P_0} - \sqrt{P_1} \Big\|_2^2,
\end{aligned} \tag{3.221}$$

which completes the proof. □

Alternative proof of Theorem 3.36. For the first inequality in (3.207), one has

$$\begin{aligned}
\left\| \rho - \sigma \right\|_1 &\geq \left\| \sqrt{\rho} - \sqrt{\sigma} \right\|_2^2 = \mathrm{Tr}\left(\sqrt{\rho} - \sqrt{\sigma}\right)^2 \\
&= 2 - 2\,\mathrm{Tr}\left(\sqrt{\rho}\sqrt{\sigma}\right) \geq 2 - 2\,\mathrm{F}(\rho, \sigma)
\end{aligned} \tag{3.222}$$

by Lemma 3.37. The second inequality in (3.207) is proved as before. □

3.3 Channel distances and discrimination

The trace norm induces a notion of distance between quantum states that is closely related to the task of state discrimination, as established by the Holevo–Helstrom theorem (Theorem 3.4). The present section discusses an analogous notion of distance for channels, induced by a norm known as the *completely bounded trace norm*, along with a similar connection to the task of *channel discrimination*.

3.3.1 Channel discrimination

The task of discriminating between pairs of channels is represented by the scenario that follows.

Scenario 3.38. Let X and Y be registers, let Z be a classical register having classical state set $\{0, 1\}$, let $\Phi_0, \Phi_1 \in \mathrm{C}(\mathcal{X}, \mathcal{Y})$ be channels, and let $\lambda \in [0, 1]$ be a real number. The channels Φ_0 and Φ_1, as well as the number λ, are assumed to be known to both Alice and Bob.

Alice prepares Z in a probabilistic state, so that its state is 0 with probability λ and 1 with probability $1 - \lambda$. Conditioned on the state of Z, Alice interacts with Bob in one of the following two ways:

1. If $\mathsf{Z} = 0$, Alice receives X from Bob, transforms X into Y according to the action of Φ_0, and sends Y to Bob.
2. If $\mathsf{Z} = 1$, Alice receives X from Bob, transforms X into Y according to the action of Φ_1, and sends Y to Bob.

Bob's goal is to determine the classical state of Z, by means of an interaction with Alice.

One approach Bob may choose to take in this scenario is to select a state $\sigma \in \mathrm{D}(\mathcal{X})$ that maximizes the quantity

$$\left\| \lambda\rho_0 - (1 - \lambda)\rho_1 \right\|_1, \tag{3.223}$$

for $\rho_0 = \Phi_0(\sigma)$ and $\rho_1 = \Phi_1(\sigma)$. If he prepares the register X in the state σ and sends it to Alice, he will receive Y in either of the states ρ_0 or ρ_1, and can then measure Y using an optimal measurement for discriminating ρ_0 and ρ_1 given with probabilities λ and $1 - \lambda$, respectively.

This, however, is not the most general approach. More generally, Bob may make use of an *auxiliary* register W in the following way. First, he prepares the pair of registers (X,W) in some chosen state $\sigma \in \mathrm{D}(\mathcal{X}\otimes\mathcal{W})$, and then he allows Alice to transform X into Y according to Φ_0 or Φ_1. This results in the pair (Y,W) being in one of the two states

$$\rho_0 = \big(\Phi_0 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{W})}\big)(\sigma) \quad \text{and} \quad \rho_1 = \big(\Phi_1 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{W})}\big)(\sigma), \tag{3.224}$$

with probabilities λ and $1-\lambda$, respectively. Finally, he measures the pair (Y,W) in order to discriminate these two states. This more general approach can, in some cases, result in a striking improvement in the probability to discriminate Φ_0 and Φ_1, as the following example illustrates.

Example 3.39. Let $n \geq 2$, let Σ be an alphabet with $|\Sigma| = n$, and let X be a register having classical state set Σ. Define two channels $\Phi_0, \Phi_1 \in \mathrm{C}(\mathcal{X})$ as follows:

$$\begin{aligned} \Phi_0(X) &= \frac{1}{n+1}\big((\operatorname{Tr} X)\mathbb{1} + X^{\mathsf{T}}\big), \\ \Phi_1(X) &= \frac{1}{n-1}\big((\operatorname{Tr} X)\mathbb{1} - X^{\mathsf{T}}\big), \end{aligned} \tag{3.225}$$

for all $X \in \mathrm{L}(\mathcal{X})$.

The maps Φ_0 and Φ_1, which are sometimes called the *Werner–Holevo channels*, are indeed channels. These maps are evidently trace preserving, and the fact that they are completely positive follows from a calculation of their Choi representations:

$$J(\Phi_0) = \frac{\mathbb{1}\otimes\mathbb{1} + W}{n+1} \quad \text{and} \quad J(\Phi_1) = \frac{\mathbb{1}\otimes\mathbb{1} - W}{n-1}. \tag{3.226}$$

Here, $W \in \mathrm{L}(\mathcal{X}\otimes\mathcal{X})$ is the swap operator, which satisfies $W(u\otimes v) = v\otimes u$ for every $u, v \in \mathcal{X}$. As W is unitary and Hermitian, the operators $J(\Phi_0)$ and $J(\Phi_1)$ are both positive semidefinite.

Now, consider the channels Φ_0 and Φ_1, along with the scalar value

$$\lambda = \frac{n+1}{2n}, \tag{3.227}$$

in Scenario 3.38. It holds that

$$\lambda\Phi_0(X) - (1-\lambda)\Phi_1(X) = \frac{1}{n}X^{\mathsf{T}} \tag{3.228}$$

for every $X \in \mathrm{L}(\mathcal{X})$, and therefore

$$\big\|\lambda\Phi_0(\sigma) - (1-\lambda)\Phi_1(\sigma)\big\|_1 = \frac{1}{n} \tag{3.229}$$

for every choice of a density operator $\sigma \in \mathrm{D}(\mathcal{X})$. This quantity is relatively small when n is large, which is consistent with the observation that $\Phi_0(\sigma)$ and $\Phi_1(\sigma)$ are both close to the completely mixed state for any choice of an input $\sigma \in \mathrm{D}(\mathcal{X})$. If Bob prepares X in some state σ, and elects not to use an auxiliary register W, his probability to correctly identify the classical state of Z is therefore at most

$$\frac{1}{2} + \frac{1}{2n}. \tag{3.230}$$

On the other hand, if Bob makes use of an auxiliary register, the situation is quite different. In particular, suppose that W is a register sharing the same classical state set Σ as X, and suppose that Bob prepares the pair (X,W) in the state $\tau \in \mathrm{D}(\mathcal{X}\otimes\mathcal{W})$ defined as

$$\tau = \frac{1}{n}\sum_{a,b\in\Sigma} E_{a,b}\otimes E_{a,b}. \tag{3.231}$$

The actions of the channels Φ_0 and Φ_1 on this state are as follows:

$$\begin{aligned} \big(\Phi_0 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{W})}\big)(\tau) &= \frac{\mathbb{1}\otimes\mathbb{1} + W}{n^2+n}, \\ \big(\Phi_1 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{W})}\big)(\tau) &= \frac{\mathbb{1}\otimes\mathbb{1} - W}{n^2-n}. \end{aligned} \tag{3.232}$$

These are orthogonal density operators, following from the calculation

$$\langle \mathbb{1}\otimes\mathbb{1} + W, \mathbb{1}\otimes\mathbb{1} - W\rangle = \operatorname{Tr}\big(\mathbb{1}\otimes\mathbb{1} + W - W - W^2\big) = 0. \tag{3.233}$$

It is therefore the case that the states $\big(\Phi_0 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})}\big)(\tau)$ and $\big(\Phi_1 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})}\big)(\tau)$ can be discriminated without error: for every $\lambda \in [0,1]$, one has

$$\big\|\lambda\big(\Phi_0 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{W})}\big)(\tau) - (1-\lambda)\big(\Phi_1 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{W})}\big)(\tau)\big\|_1 = 1. \tag{3.234}$$

By making use of an auxiliary register W in this way, Bob can therefore correctly discriminate the channels Φ_0 and Φ_1 without error.

This example makes clear that auxiliary registers must be taken into account when considering the optimal probability with which channels can be discriminated.

3.3.2 The completely bounded trace norm

This section defines a norm on the space $\mathrm{T}(\mathcal{X},\mathcal{Y})$, for $\mathcal{X}$ and $\mathcal{Y}$ being any two complex Euclidean spaces, known as the *completely bounded trace norm*, and establishes some of its properties. The precise connection between this norm and the task of channel discrimination will be explained in the section following this one, but it will be evident from its definition that this norm is motivated in part by the discussion from the previous section stressing the importance of auxiliary systems in the task of channel discrimination.

The induced trace norm

When introducing the completely bounded trace norm, it is appropriate to begin with the definition of a related norm known as the *induced trace norm*.

Definition 3.40. Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces. For each map $\Phi \in \mathrm{T}(\mathcal{X},\mathcal{Y})$, the *induced trace norm* of Φ is defined as

$$\big\|\Phi\big\|_1 = \max\Big\{\big\|\Phi(X)\big\|_1 \,:\, X \in \mathrm{L}(\mathcal{X}),\ \ \big\|X\big\|_1 \leq 1\Big\}. \tag{3.235}$$

True to its name, this norm is an example of an *induced norm*; in general, one may consider the norm obtained by replacing the two trace norms in this definition with any other choices of norms that are defined on $\mathrm{L}(\mathcal{X})$ and $\mathrm{L}(\mathcal{Y})$. The use of the maximum, rather than the supremum, is justified in this context by the observation that the norm defined on $\mathrm{L}(\mathcal{Y})$ is continuous and the unit ball with respect to the norm defined on $\mathrm{L}(\mathcal{X})$ is compact.

Generally speaking, the induced trace norm fails to provide a physically well-motivated measure of distance between channels. It will, nevertheless, be useful to consider some basic properties of this norm, for many of these properties will be inherited by the completely bounded trace norm, to be defined shortly.

The first property of the induced trace norm to be observed is that the maximum in Definition 3.40 is always achieved by a rank-one operator X.

Proposition 3.41. *Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces and let $\Phi \in \mathrm{T}(\mathcal{X},\mathcal{Y})$ be a map. It holds that*

$$\big\|\Phi\big\|_1 = \max_{u,v\in\mathcal{S}(\mathcal{X})} \big\|\Phi(uv^*)\big\|_1. \tag{3.236}$$

Proof. Every operator in $X \in \mathrm{L}(\mathcal{X})$ satisfying $\|X\|_1 \leq 1$ can be written as a convex combination of operators of the form uv^*, for $u, v \in \mathcal{S}(\mathcal{X})$ being unit vectors. The equation (3.236) follows from the fact that the trace norm is a convex function. □

Under the additional assumption that the mapping under consideration is positive, one has that the maximum in Definition 3.40 is achieved by a rank-one projection, as the following theorem states.

Theorem 3.42 (Russo–Dye)**.** *Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces and let $\Phi \in \mathrm{T}(\mathcal{X},\mathcal{Y})$ be a positive map. It holds that*

$$\big\|\Phi\big\|_1 = \max_{u\in\mathcal{S}(\mathcal{X})} \mathrm{Tr}\big(\Phi(uu^*)\big). \tag{3.237}$$

Proof. Using the duality of the trace and spectral norms, along with the identity (1.177), one finds that

$$\big\|\Phi\big\|_1 = \max_{U\in\mathrm{U}(\mathcal{Y})} \big\|\Phi^*(U)\big\|. \tag{3.238}$$

Consider an arbitrary unitary operator $U \in \mathrm{U}(\mathcal{Y})$, and let

$$U = \sum_{k=1}^{m} \lambda_k \Pi_k \tag{3.239}$$

be the spectral decomposition of U. Define an operator $P_k = \Phi^*(\Pi_k)$ for each index $k \in \{1,\ldots,m\}$. As Φ is positive, it holds that Φ^* is also positive (by Proposition 2.17), and therefore $P_1,\ldots,P_m \in \mathrm{Pos}(\mathcal{X})$. By Lemma 3.3, along with the obervation that the eigenvalues $\lambda_1,\ldots,\lambda_m$ all lie on the unit circle, it follows that

$$\big\|\Phi^*(U)\big\| = \left\|\sum_{k=1}^{m} \lambda_k P_k\right\| \leq \left\|\sum_{k=1}^{m} P_k\right\| = \big\|\Phi^*(\mathbb{1}_{\mathcal{Y}})\big\|. \tag{3.240}$$

Consequently, as $\mathbb{1}_{\mathcal{Y}}$ is itself a unitary operator, one has

$$\|\Phi\|_1 = \big\|\Phi^*(\mathbb{1}_{\mathcal{Y}})\big\|. \tag{3.241}$$

Finally, as $\Phi^*(\mathbb{1}_{\mathcal{Y}})$ is necessarily positive semidefinite, it follows that

$$\big\|\Phi^*(\mathbb{1}_{\mathcal{Y}})\big\| = \max_{u\in\mathcal{S}(\mathcal{X})} \big\langle uu^*, \Phi^*(\mathbb{1}_{\mathcal{Y}})\big\rangle = \max_{u\in\mathcal{S}(\mathcal{X})} \mathrm{Tr}\big(\Phi(uu^*)\big), \tag{3.242}$$

which completes the proof. □

Corollary 3.43. *Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces and let $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ be a positive and trace-preserving map. It holds that $\|\Phi\|_1 = 1$.*

The next proposition establishes three basic properties of the induced trace norm: submultiplicativity under compositions, additivity of channel differences under compositions, and unitary invariance.

Proposition 3.44. *For every choice of complex Euclidean spaces $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$, the following facts regarding the induced trace norm hold:*

1. *For all maps $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ and $\Psi \in \mathrm{T}(\mathcal{Y}, \mathcal{Z})$, it holds that*

$$\|\Psi\Phi\|_1 \leq \|\Psi\|_1 \|\Phi\|_1. \tag{3.243}$$

2. *For all channels $\Phi_0, \Psi_0 \in \mathrm{C}(\mathcal{X}, \mathcal{Y})$ and $\Phi_1, \Psi_1 \in \mathrm{C}(\mathcal{Y}, \mathcal{Z})$, it holds that*

$$\|\Psi_1\Psi_0 - \Phi_1\Phi_0\|_1 \leq \|\Psi_0 - \Phi_0\|_1 + \|\Psi_1 - \Phi_1\|_1. \tag{3.244}$$

3. *Let $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ be a map, let $U_0, V_0 \in \mathrm{U}(\mathcal{X})$ and $U_1, V_1 \in \mathrm{U}(\mathcal{Y})$ be unitary operators, and let $\Psi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ be defined as*

$$\Psi(X) = U_1\Phi(U_0XV_0)V_1 \tag{3.245}$$

for all $X \in \mathrm{L}(\mathcal{X})$. It holds that $\|\Psi\|_1 = \|\Phi\|_1$.

Proof. To prove the first fact, one may observe that $\|\Psi(Y)\|_1 \leq \|\Psi\|_1\|Y\|_1$ for every $Y \in \mathrm{L}(\mathcal{Y})$, and therefore

$$\|\Psi(\Phi(X))\|_1 \leq \|\Psi\|_1\|\Phi(X)\|_1 \tag{3.246}$$

for every $X \in \mathrm{L}(\mathcal{X})$. Taking the maximum over all $X \in \mathrm{L}(\mathcal{X})$ with $\|X\|_1 \leq 1$ yields the inequality (3.243).

To prove the second fact, one may apply the triangle inequality, the inequality (3.243), and Corollary 3.43, to obtain

$$\begin{aligned}
\|\Psi_1\Psi_0 - \Phi_1\Phi_0\|_1 &\leq \|\Psi_1\Psi_0 - \Psi_1\Phi_0\|_1 + \|\Psi_1\Phi_0 - \Phi_1\Phi_0\|_1 \\
&= \|\Psi_1(\Psi_0 - \Phi_0)\|_1 + \|(\Psi_1 - \Phi_1)\Phi_0\|_1 \\
&\leq \|\Psi_1\|_1\|\Psi_0 - \Phi_0\|_1 + \|\Psi_1 - \Phi_1\|_1\|\Phi_0\|_1 \\
&= \|\Psi_0 - \Phi_0\|_1 + \|\Psi_1 - \Phi_1\|_1.
\end{aligned} \tag{3.247}$$

Finally, by the unitary invariance of the trace norm, it follows that

$$\begin{aligned}
\|\Psi(X)\|_1 &= \|U_1\Phi(U_0XV_0)V_1\|_1 = \|\Phi(U_0XV_0)\|_1 \\
&\leq \|\Phi\|_1\|U_0XV_0\|_1 = \|\Phi\|_1\|X\|_1
\end{aligned} \tag{3.248}$$

for all $X \in \mathrm{L}(\mathcal{X})$, and therefore $\|\Psi\|_1 \leq \|\Phi\|_1$. By observing that

$$\Phi(X) = U_1^*\Psi(U_0^*XV_0^*)V_1^* \tag{3.249}$$

for all $X \in \mathrm{L}(\mathcal{X})$, one finds that $\|\Phi\|_1 \leq \|\Psi\|_1$ through a similar argument, which proves the third fact. □

One undesirable property of the induced trace norm is that it fails to be multiplicative with respect to tensor products, as the following example (which is closely related to Example 3.39) illustrates.

Example 3.45. Let $n \geq 2$, let Σ be an alphabet with $|\Sigma| = n$, let $\mathcal{X} = \mathbb{C}^\Sigma$, and consider the transpose map $\mathrm{T} \in \mathrm{T}(\mathcal{X})$, defined as $\mathrm{T}(X) = X^\mathsf{T}$ for all $X \in \mathrm{L}(\mathcal{X})$. It is evident that $\|\mathrm{T}\|_1 = 1$, as $\|X\|_1 = \|X^\mathsf{T}\|_1$ for every operator $X \in \mathrm{L}(\mathcal{X})$, and it holds that $\|\mathbb{1}_{\mathrm{L}(\mathcal{X})}\|_1 = 1$. On the other hand, one has

$$\|\mathrm{T} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})}\|_1 = n. \tag{3.250}$$

To verify this claim, one may first consider the density operator

$$\tau = \frac{1}{n}\sum_{a,b\in\Sigma} E_{a,b} \otimes E_{a,b} \in \mathrm{D}(\mathcal{X} \otimes \mathcal{X}), \tag{3.251}$$

which has trace norm equal to 1. It holds that

$$\|(\mathrm{T} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})})(\tau)\|_1 = \frac{1}{n}\|W\|_1 = n \tag{3.252}$$

for $W \in \mathrm{U}(\mathcal{X} \otimes \mathcal{X})$ denoting the swap operator, and therefore

$$\|\mathrm{T} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})}\|_1 \geq n. \tag{3.253}$$

To prove that $\|\mathrm{T} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})}\|_1$ is no larger than n, one may first observe that the relationship (1.163) between the trace and Frobenius norms implies

$$\|(\mathrm{T} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})})(X)\|_1 \leq n\|(\mathrm{T} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})})(X)\|_2 \tag{3.254}$$

for every operator $X \in \mathrm{L}(\mathcal{X} \otimes \mathcal{X})$. As the entries of the operators X and $(\mathrm{T} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})})(X)$ are equal, up to being shuffled by the transposition mapping, one has that

$$\big\| (\mathrm{T} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})})(X) \big\|_2 = \big\| X \big\|_2. \tag{3.255}$$

Finally, by (1.164) it holds that $\| X \|_2 \leq \| X \|_1$, from which it follows that

$$\big\| \mathrm{T} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})} \big\|_1 \leq n. \tag{3.256}$$

Definition of the completely bounded trace norm

The *completely bounded trace norm* is defined below. In words, its value for a given map is simply the induced trace norm of that map tensored with the identity map on the same input space as the mapping itself.

Definition 3.46. For any choice of complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$, the *completely bounded trace norm* of a mapping $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ is defined as

$$\big|\!\big|\!\big| \Phi \big|\!\big|\!\big|_1 = \big\| \Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})} \big\|_1. \tag{3.257}$$

As the discussion in Section 3.3.1 has suggested, this is the more relevant norm, when compared with the induced trace norm, within the context of the channel discrimination task. In essence, the completely bounded trace norm quantifies the effect that a map may have when it acts on just one tensor factor of a tensor product space (or, in more physical terms, just one part of a compound system), as opposed to the action of that map on its input space alone. As it turns out, this definition not only yields a norm that is more relevant to the channel discrimination task, but also one possessing many interesting and desirable properties (including multiplicativity with respect to tensor products).

The specific choice to take the identity mapping on $\mathrm{L}(\mathcal{X})$, as opposed to $\mathrm{L}(\mathcal{Y})$, or $\mathrm{L}(\mathcal{Z})$ for some other complex Euclidean space $\mathcal{Z}$, is explained in greater detail below. In simple terms, the space $\mathcal{X}$ is sufficiently large, and just large enough in the worst case, that the value (3.257) does not change if the identity mapping on $\mathrm{L}(\mathcal{X})$ is replaced by the identity mapping on $\mathrm{L}(\mathcal{Z})$, for any complex Euclidean space $\mathcal{Z}$ having dimension at least as large as the dimension of $\mathcal{X}$.

Basic properties of the completely bounded trace norm

The proposition that follows, which is immediate from Propositions 3.41 and 3.44 and Corollary 3.43, summarizes some of the basic properties that the completely bounded trace norm inherits from the induced trace norm.

Proposition 3.47. *For every choice of complex Euclidean spaces $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$, the following facts regarding the completely bounded trace norm hold:*

1. *For all maps $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$, it holds that*

$$\big|\!\big|\!\big| \Phi \big|\!\big|\!\big|_1 = \max\Big\{ \big\| (\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})})(uv^*) \big\|_1 \,:\, u, v \in \mathcal{S}(\mathcal{X} \otimes \mathcal{X}) \Big\}. \tag{3.258}$$

2. *For all channels $\Phi \in \mathrm{C}(\mathcal{X}, \mathcal{Y})$, it holds that $\big|\!\big|\!\big| \Phi \big|\!\big|\!\big|_1 = 1$.*
3. *For all maps $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ and $\Psi \in \mathrm{T}(\mathcal{Y}, \mathcal{Z})$, it holds that*

$$\big|\!\big|\!\big| \Psi\Phi \big|\!\big|\!\big|_1 \leq \big|\!\big|\!\big| \Psi \big|\!\big|\!\big|_1 \big|\!\big|\!\big| \Phi \big|\!\big|\!\big|_1. \tag{3.259}$$

4. *For all channels $\Phi_0, \Psi_0 \in \mathrm{C}(\mathcal{X}, \mathcal{Y})$ and $\Phi_1, \Psi_1 \in \mathrm{C}(\mathcal{Y}, \mathcal{Z})$, it holds that*

$$\big|\!\big|\!\big| \Psi_1\Psi_0 - \Phi_1\Phi_0 \big|\!\big|\!\big|_1 \leq \big|\!\big|\!\big| \Psi_0 - \Phi_0 \big|\!\big|\!\big|_1 + \big|\!\big|\!\big| \Psi_1 - \Phi_1 \big|\!\big|\!\big|_1. \tag{3.260}$$

5. *Let $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ be a map, let $U_0, V_0 \in \mathrm{U}(\mathcal{X})$ and $U_1, V_1 \in \mathrm{U}(\mathcal{Y})$ be unitary operators, and let $\Psi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ be defined as*

$$\Psi(X) = U_1 \Phi(U_0 X V_0) V_1 \tag{3.261}$$

for all $X \in \mathrm{L}(\mathcal{X})$. It holds that $\big|\!\big|\!\big| \Psi \big|\!\big|\!\big|_1 = \big|\!\big|\!\big| \Phi \big|\!\big|\!\big|_1$.

The next lemma will be used multiple times to establish further properties of the completely bounded trace norm.

Lemma 3.48. *Let $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ be complex Euclidean spaces and let $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ be a map. For every choice of unit vectors $x, y \in \mathcal{X} \otimes \mathcal{Z}$ there exist unit vectors $u, v \in \mathcal{X} \otimes \mathcal{X}$ such that*

$$\big\| (\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})})(xy^*) \big\|_1 = \big\| (\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})})(uv^*) \big\|_1. \tag{3.262}$$

If it is the case that $x = y$, then the equality (3.262) *holds under the additional requirement that $u = v$.*

Proof. In the case that $\dim(\mathcal{Z}) \leq \dim(\mathcal{X})$, the lemma is straightforward: for any choice of a linear isometry $U \in \mathrm{U}(\mathcal{Z}, \mathcal{X})$, the vectors $u = (\mathbb{1}_{\mathcal{X}} \otimes U)x$ and $v = (\mathbb{1}_{\mathcal{X}} \otimes U)y$ satisfy the required conditions.

In the case that $\dim(\mathcal{Z}) > \dim(\mathcal{X})$, one may consider Schmidt decompositions

$$x = \sum_{k=1}^{n} \sqrt{p_k}\, x_k \otimes z_k \quad \text{and} \quad y = \sum_{k=1}^{n} \sqrt{q_k}\, y_k \otimes w_k \tag{3.263}$$

of x and y, for $n = \dim(\mathcal{X})$, from which a suitable choice for the vectors u and v is given by

$$u = \sum_{k=1}^{n} \sqrt{p_k}\, x_k \otimes x_k \quad \text{and} \quad v = \sum_{k=1}^{n} \sqrt{q_k}\, y_k \otimes y_k. \tag{3.264}$$

For linear isometries $U, V \in \mathrm{U}(\mathcal{X}, \mathcal{Z})$ defined as

$$U = \sum_{k=1}^{n} z_k x_k^* \quad \text{and} \quad V = \sum_{k=1}^{n} w_k y_k^*, \tag{3.265}$$

it holds that $x = (\mathbb{1}_{\mathcal{X}} \otimes U)u$ and $y = (\mathbb{1}_{\mathcal{X}} \otimes V)v$, and therefore

$$\begin{aligned} \big\| (\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})})(xy^*) \big\|_1 &= \big\| (\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})})((\mathbb{1} \otimes U)uv^*(\mathbb{1} \otimes V^*)) \big\|_1 \\ &= \big\| (\mathbb{1} \otimes U)(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})})(uv^*)(\mathbb{1} \otimes V^*) \big\|_1 \\ &= \big\| (\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})})(uv^*) \big\|_1, \end{aligned} \tag{3.266}$$

as required. In case $x = y$, one may take the same Schmidt decomposition for x and y in (3.263), implying that $u = v$. □

With Lemma 3.48 in hand, the following theorem may be proved. The theorem implies a claim that was made earlier: the identity map on $\mathrm{L}(\mathcal{X})$ in Definition 3.46 could be replaced by the identity map on $\mathrm{L}(\mathcal{Z})$, for any space $\mathcal{Z}$ having dimension at least that of $\mathcal{X}$, without changing the value of the norm.

Theorem 3.49. *Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces, let $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ be a map, and let $\mathcal{Z}$ be a complex Euclidean space. It holds that*

$$\big\| \Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})} \big\|_1 \leq \big\| \Phi \big\|_1, \tag{3.267}$$

with equality holding under the assumption that $\dim(\mathcal{Z}) \geq \dim(\mathcal{X})$.

Proof. By Proposition 3.41, there exist unit vectors $x, y \in \mathcal{X} \otimes \mathcal{Z}$ such that

$$\big\| \Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})} \big\|_1 = \big\| \big(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})}\big)(xy^*) \big\|_1. \tag{3.268}$$

Therefore, by Lemma 3.48, there exist unit vectors $u, v \in \mathcal{X} \otimes \mathcal{X}$ such that

$$\big\| \Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})} \big\|_1 = \big\| \big(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})}\big)(uv^*) \big\|_1, \tag{3.269}$$

which implies

$$\big\| \Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})} \big\|_1 \leq \big\| \Phi \big\|_1. \tag{3.270}$$

To prove that equality holds under the assumption $\dim(\mathcal{Z}) \geq \dim(\mathcal{X})$, one may take $V \in \mathrm{U}(\mathcal{X}, \mathcal{Z})$ to be any isometry, and observe that

$$\begin{aligned} \big\| \big(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})}\big)(X) \big\|_1 &= \big\| (\mathbb{1}_{\mathcal{Y}} \otimes V)\big(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})}\big)(X)(\mathbb{1}_{\mathcal{Y}} \otimes V)^* \big\|_1 \\ &= \big\| \big(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})}\big)\big((\mathbb{1}_{\mathcal{X}} \otimes V)X(\mathbb{1}_{\mathcal{X}} \otimes V)^*\big) \big\|_1 \\ &\leq \big\| \Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})} \big\|_1 \big\| (\mathbb{1}_{\mathcal{X}} \otimes V)X(\mathbb{1}_{\mathcal{X}} \otimes V)^* \big\|_1 \\ &= \big\| \Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})} \big\|_1 \big\| X \big\|_1 \\ &\leq \big\| \Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})} \big\|_1, \end{aligned} \tag{3.271}$$

for every operator $X \in \mathrm{L}(\mathcal{X} \otimes \mathcal{X})$ with $\|X\|_1 \leq 1$, by the isometric invariance of the trace norm. It therefore holds that

$$\big\| \Phi \big\|_1 \leq \big\| \Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})} \big\|_1 \tag{3.272}$$

as required. □

Corollary 3.50. *Let $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ be complex Euclidean spaces and let $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ be a map. It holds that*

$$\big\| \Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})} \big\|_1 = \big\| \Phi \big\|_1. \tag{3.273}$$

The fact that the completely bounded trace norm is multiplicative with respect to tensor products may now be proved.

Theorem 3.51. *Let $\Phi_0 \in \mathrm{T}(\mathcal{X}_0, \mathcal{Y}_0)$ and $\Phi_1 \in \mathrm{T}(\mathcal{X}_1, \mathcal{Y}_1)$ be maps, for $\mathcal{X}_0$, $\mathcal{X}_1$, $\mathcal{Y}_0$, and $\mathcal{Y}_1$ being complex Euclidean spaces. It holds that*

$$\big\| \Phi_0 \otimes \Phi_1 \big\|_1 = \big\| \Phi_0 \big\|_1 \big\| \Phi_1 \big\|_1. \tag{3.274}$$

Proof. By Proposition 3.47 and Corollary 3.50, it follows that

$$\begin{aligned}\bigl|\!\bigl|\!\bigl|\Phi_0 \otimes \Phi_1\bigr|\!\bigr|\!\bigr|_1 &= \bigl|\!\bigl|\!\bigl|(\Phi_0 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y}_1)})(\mathbb{1}_{\mathrm{L}(\mathcal{X}_0)} \otimes \Phi_1)\bigr|\!\bigr|\!\bigr|_1 \\ &\leq \bigl|\!\bigl|\!\bigl|\Phi_0 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y}_1)}\bigr|\!\bigr|\!\bigr|_1 \bigl|\!\bigl|\!\bigl|\mathbb{1}_{\mathrm{L}(\mathcal{X}_0)} \otimes \Phi_1\bigr|\!\bigr|\!\bigr|_1 = \bigl|\!\bigl|\!\bigl|\Phi_0\bigr|\!\bigr|\!\bigr|_1 \bigl|\!\bigl|\!\bigl|\Phi_1\bigr|\!\bigr|\!\bigr|_1.\end{aligned} \tag{3.275}$$

It remains to prove the reverse inequality.

First, choose an operator $X_0 \in \mathrm{L}(\mathcal{X}_0 \otimes \mathcal{X}_0)$ such that $\|X_0\|_1 = 1$ and

$$\bigl|\!\bigl|\!\bigl|\Phi_0\bigr|\!\bigr|\!\bigr|_1 = \bigl\|\bigl(\Phi_0 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X}_0)}\bigr)(X_0)\bigr\|_1, \tag{3.276}$$

as well as an operator $X_1 \in \mathrm{L}(\mathcal{X}_1 \otimes \mathcal{X}_1)$ such that $\|X_1\|_1 = 1$ and

$$\bigl|\!\bigl|\!\bigl|\Phi_1\bigr|\!\bigr|\!\bigr|_1 = \bigl\|\bigl(\Phi_1 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X}_1)}\bigr)(X_1)\bigr\|_1. \tag{3.277}$$

The trace norm is multiplicative with respect to tensor products, and therefore $\|X_0 \otimes X_1\|_1 = 1$.

Next, observe that

$$\begin{aligned}\bigl|\!\bigl|\!\bigl|\Phi_0 \otimes \Phi_1\bigr|\!\bigr|\!\bigr|_1 &= \bigl\|\Phi_0 \otimes \Phi_1 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X}_0 \otimes \mathcal{X}_1)}\bigr\|_1 \\ &= \bigl\|\Phi_0 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X}_0)} \otimes \Phi_1 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X}_1)}\bigr\|_1.\end{aligned} \tag{3.278}$$

The second equality follows from the unitary invariance of the induced trace norm (the third statement of Proposition 3.44), which implies that this norm is invariant under permuting the ordering of tensor factors of maps. Again using the multiplicativity of the trace norm with respect to tensor products, it follows that

$$\begin{aligned}\bigl|\!\bigl|\!\bigl|\Phi_0 \otimes \Phi_1\bigr|\!\bigr|\!\bigr|_1 &\geq \bigl\|\bigl(\Phi_0 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X}_0)} \otimes \Phi_1 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X}_1)}\bigr)(X_0 \otimes X_1)\bigr\|_1 \\ &= \bigl\|\bigl(\Phi_0 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X}_0)}\bigr)(X_0)\bigr\|_1 \bigl\|\bigl(\Phi_1 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X}_1)}\bigr)(X_1)\bigr\|_1 \\ &= \bigl|\!\bigl|\!\bigl|\Phi_0\bigr|\!\bigr|\!\bigr|_1 \bigl|\!\bigl|\!\bigl|\Phi_1\bigr|\!\bigr|\!\bigr|_1,\end{aligned} \tag{3.279}$$

which completes the proof. □

3.3.3 Distances between channels

This section explains the connection between the completely bounded trace norm and the task of channel discrimination that was alluded to above, and discusses other aspects of the notion of distance between channels induced by the completely bounded trace norm.

The completely bounded trace norm of Hermiticity-preserving maps

For a given map $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$, one has that

$$\bigl|\!\bigl|\!\bigl|\Phi\bigr|\!\bigr|\!\bigr|_1 = \bigl\|(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})})(uv^*)\bigr\|_1 \tag{3.280}$$

for some choice of unit vectors $u, v \in \mathcal{X} \otimes \mathcal{X}$. The stronger condition that

$$\bigl|\!\bigl|\!\bigl|\Phi\bigr|\!\bigr|\!\bigr|_1 = \bigl\|(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})})(uu^*)\bigr\|_1 \tag{3.281}$$

for a single unit vector $u \in \mathcal{X} \otimes \mathcal{X}$ does not generally hold; without any restrictions on Φ, this could not reasonably be expected.

When the map Φ is Hermiticity-preserving, however, there will always exist a unit vector $u \in \mathcal{X} \otimes \mathcal{X}$ for which (3.281) holds. This fact is stated as Theorem 3.53 below, whose proof makes use of the following lemma.

Lemma 3.52. *Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces, let $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ be a Hermiticity-preserving map, and let $\mathcal{Z}$ be any complex Euclidean space with $\dim(\mathcal{Z}) \geq 2$. There exists a unit vector $u \in \mathcal{X} \otimes \mathcal{Z}$ such that*

$$\bigl\|\bigl(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})}\bigr)(uu^*)\bigr\|_1 \geq \bigl\|\Phi\bigr\|_1. \tag{3.282}$$

Proof. Let $X \in \mathrm{L}(\mathcal{X})$ be an operator for which it holds that $\|X\|_1 = 1$ and $\|\Phi(X)\|_1 = \|\Phi\|_1$. Let $z_0, z_1 \in \mathcal{Z}$ be any to orthogonal unit vectors, define a Hermitian operator $H \in \mathrm{Herm}(\mathcal{X} \otimes \mathcal{Z})$ as

$$H = \frac{1}{2} X \otimes z_0 z_1^* + \frac{1}{2} X^* \otimes z_1 z_0^*, \tag{3.283}$$

and observe that $\|H\|_1 = \|X\|_1 = 1$. Moreover, one has

$$\begin{aligned}\bigl(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})}\bigr)(H) &= \frac{1}{2}\Phi(X) \otimes z_0 z_1^* + \frac{1}{2}\Phi(X^*) \otimes z_1 z_0^* \\ &= \frac{1}{2}\Phi(X) \otimes z_0 z_1^* + \frac{1}{2}\Phi(X)^* \otimes z_1 z_0^*,\end{aligned} \tag{3.284}$$

where the second equality follows from Theorem 2.25, together with the assumption that Φ is a Hermiticity-preserving map. It is therefore the case that

$$\bigl\|\bigl(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})}\bigr)(H)\bigr\|_1 = \bigl\|\Phi(X)\bigr\|_1 = \bigl\|\Phi\bigr\|_1. \tag{3.285}$$

Now consider a spectral decomposition

$$H = \sum_{k=1}^{n} \lambda_k u_k u_k^* \tag{3.286}$$

for $n = \dim(\mathcal{X} \otimes \mathcal{Z})$. By the triangle inequality, one has

$$\big\| \big(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})}\big)(H) \big\|_1 \leq \sum_{k=1}^{n} |\lambda_k| \, \big\| \big(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})}\big)(u_k u_k^*) \big\|_1. \tag{3.287}$$

As $\|H\|_1 = 1$, the expression on the right-hand side of the inequality (3.287) is a convex combination of the values

$$\big\| \big(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})}\big)(u_k u_k^*) \big\|_1, \tag{3.288}$$

ranging over $k \in \{1, \ldots, n\}$. There must therefore exist $k \in \{1, \ldots, n\}$ for which the inequality

$$\big\| \big(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})}\big)(u_k u_k^*) \big\|_1 \geq \big\| \big(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})}\big)(H) \big\|_1 = \big\| \Phi \big\|_1 \tag{3.289}$$

is satisfied. Setting $u = u_k$ completes the proof. □

Theorem 3.53. *Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces and let $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ be a Hermiticity-preserving map. It holds that*

$$\big\| \Phi \big\|_1 = \max_{u \in \mathcal{S}(\mathcal{X} \otimes \mathcal{X})} \big\| \big(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})}\big)(uu^*) \big\|_1. \tag{3.290}$$

Proof. For every unit vector $u \in \mathcal{X} \otimes \mathcal{X}$, it holds that

$$\big\| \big(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})}\big)(uu^*) \big\|_1 \leq \big\| \Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})} \big\|_1 = \big\| \Phi \big\|_1, \tag{3.291}$$

so it suffices to prove that there exists a unit vector $u \in \mathcal{X} \otimes \mathcal{X}$ for which

$$\big\| \big(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})}\big)(uu^*) \big\|_1 \geq \big\| \Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})} \big\|_1 = \big\| \Phi \big\|_1. \tag{3.292}$$

Let $\mathcal{Z} = \mathbb{C}^2$. By Lemma 3.52 there exists a unit vector $x \in \mathcal{X} \otimes \mathcal{X} \otimes \mathcal{Z}$ such that

$$\big\| \big(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})}\big)(xx^*) \big\|_1 \geq \big\| \Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})} \big\|_1, \tag{3.293}$$

and by Lemma 3.48 there must exists a unit vector $u \in \mathcal{X} \otimes \mathcal{X}$ such that

$$\big\| \big(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})}\big)(uu^*) \big\|_1 = \big\| \big(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})}\big)(xx^*) \big\|_1. \tag{3.294}$$

For such a choice of u, one has (3.292), which completes the proof. □

A channel analogue of the Holevo–Helstrom theorem

The next theorem represents an analogue of the Holevo–Helstrom theorem (Theorem 3.4) for channels rather than states, with the completely bounded trace norm replacing the trace norm accordingly.

Theorem 3.54. *Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces, let $\Phi_0, \Phi_1 \in \mathrm{C}(\mathcal{X}, \mathcal{Y})$ be channels, and let $\lambda \in [0, 1]$. For any choice of a complex Euclidean space $\mathcal{Z}$, a density operator $\sigma \in \mathrm{D}(\mathcal{X} \otimes \mathcal{Z})$, and a measurement $\mu : \{0, 1\} \to \mathrm{Pos}(\mathcal{Y} \otimes \mathcal{Z})$, it holds that*

$$\begin{aligned} \lambda \langle \mu(0), (\Phi_0 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})})(\sigma) \rangle + (1 - \lambda) \langle \mu(1), (\Phi_1 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})})(\sigma) \rangle \\ \leq \frac{1}{2} + \frac{1}{2} \big\| \big\| \lambda \Phi_0 - (1 - \lambda) \Phi_1 \big\| \big\|_1. \end{aligned} \tag{3.295}$$

Moreover, for any choice of $\mathcal{Z}$ satisfying $\dim(\mathcal{Z}) \geq \dim(\mathcal{X})$, equality is achieved in (3.295) *for some choice of a pure state σ and a projective measurement μ.*

Proof. By the Holevo–Helstrom theorem (Theorem 3.4), the quantity on the left-hand side of (3.295) is at most

$$\frac{1}{2} + \frac{1}{2} \big\| \lambda (\Phi_0 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})})(\sigma) - (1 - \lambda)(\Phi_1 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})})(\sigma) \big\|_1. \tag{3.296}$$

This value is upper-bounded by

$$\frac{1}{2} + \frac{1}{2} \big\| (\lambda \Phi_0 - (1 - \lambda) \Phi_1) \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})} \big\|_1, \tag{3.297}$$

which is at most

$$\frac{1}{2} + \frac{1}{2} \big\| \big\| \lambda \Phi_0 - (1 - \lambda) \Phi_1 \big\| \big\|_1 \tag{3.298}$$

by Theorem 3.49.

The mapping $\lambda \Phi_0 - (1 - \lambda) \Phi_1$ is Hermiticity-preserving, by virtue of the fact that Φ_0 and Φ_1 are completely positive and λ is a real number. By Theorem 3.53, there must therefore exist a unit vector $u \in \mathcal{X} \otimes \mathcal{X}$ for which

$$\begin{aligned} \big\| \lambda (\Phi_0 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})})(uu^*) - (1 - \lambda)(\Phi_1 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})})(uu^*) \big\|_1 \\ = \big\| \big\| \lambda \Phi_0 - (1 - \lambda) \Phi_1 \big\| \big\|_1. \end{aligned} \tag{3.299}$$

Under the assumption that $\dim(\mathcal{Z}) \geq \dim(\mathcal{X})$, one therefore has

$$\begin{aligned} \big\| \lambda (\Phi_0 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})})(\sigma) - (1 - \lambda)(\Phi_1 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})})(\sigma) \big\|_1 \\ = \big\| \big\| \lambda \Phi_0 - (1 - \lambda) \Phi_1 \big\| \big\|_1 \end{aligned} \tag{3.300}$$

for

$$\sigma = (\mathbb{1}_{\mathcal{X}} \otimes V)uu^*(\mathbb{1}_{\mathcal{X}} \otimes V^*), \tag{3.301}$$

for an arbitrary choice of an isometry $V \in \mathrm{U}(\mathcal{X}, \mathcal{Z})$.

Finally, by the Holevo–Helstrom theorem, there must exist a projective measurement $\mu : \{0,1\} \to \mathrm{Pos}(\mathcal{Y} \otimes \mathcal{Z})$ such that

$$\begin{aligned} &\lambda \langle \mu(0), (\Phi_0 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})})(\sigma) \rangle + (1-\lambda) \langle \mu(1), (\Phi_1 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})})(\sigma) \rangle \\ &\quad = \frac{1}{2} + \frac{1}{2} \big\| \lambda (\Phi_0 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})})(\sigma) - (1-\lambda)(\Phi_1 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})})(\sigma) \big\|_1 \\ &\quad = \frac{1}{2} + \frac{1}{2} \big\| \lambda \Phi_0 - (1-\lambda)\Phi_1 \big\|_1, \end{aligned} \tag{3.302}$$

which completes the proof. □

Distances between networks of channels

Many computations and interactions that arise in the study of quantum information and computation can be represented as *networks* of channels. Here, one supposes that a collection of channels $\Phi_1, \ldots, \Phi_N$ having varying input and output spaces are arranged in an acyclic network, as suggested by the example depicted in Figure 3.1. The completely bounded trace norm is well-suited to analyses concerning errors, inaccuracies, and noise that may occur in such networks.

By composing the channels $\Phi_1, \ldots, \Phi_N$ in a manner consistent with the network, a single channel Φ is obtained. Under the assumption that registers $\mathsf{X}_1, \ldots, \mathsf{X}_n$ are treated as inputs to the network and registers $\mathsf{Y}_1, \ldots, \mathsf{Y}_m$ are output, the channel Φ representing the composition of the channels $\Phi_1, \ldots, \Phi_N$ takes the form

$$\Phi \in \mathrm{C}(\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_n, \mathcal{Y}_1 \otimes \cdots \otimes \mathcal{Y}_m). \tag{3.303}$$

Now suppose that $\Psi_1, \ldots, \Psi_N$ are channels whose input spaces and output spaces agree with $\Phi_1, \ldots, \Phi_N$, respectively, and that Ψ_k is substituted for Φ_k for all $k = 1, \ldots, N$. Equivalently, the channels $\Psi_1, \ldots, \Psi_N$ are composed in a manner that is consistent with the description of the network, yielding a single channel

$$\Psi \in \mathrm{C}(\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_n, \mathcal{Y}_1 \otimes \cdots \otimes \mathcal{Y}_m) \tag{3.304}$$

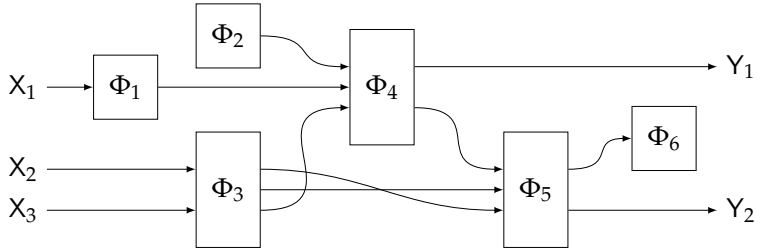

Figure 3.1: A hypothetical example of an acyclic network of channels. The arrows represent registers, and one assumes that the input and output spaces of the channels (represented by rectangles in the figure) are compatible with the registers represented by the arrows. For instance, the channel Φ_1 transforms the register X_1 into some other register (not explicitly named in the figure), which is the second of three inputs to the channel Φ_4. By composing the channels $\Phi_1, \ldots, \Phi_6$, one obtains a single channel $\Phi \in \mathrm{C}(\mathcal{X}_1 \otimes \mathcal{X}_2 \otimes \mathcal{X}_3, \mathcal{Y}_1 \otimes \mathcal{Y}_2)$.

in place of Φ. It is natural to ask how much Φ and Ψ may differ, as a function of the differences between Φ_k and Ψ_k for $k = 1, \ldots, N$. It could be, for instance, that $\Phi_1, \ldots, \Phi_N$ represent ideal channels that are specified by a protocol or algorithm while $\Psi_1, \ldots, \Psi_N$ represent slightly noisy or corrupted variants of $\Phi_1, \ldots, \Phi_N$.

The following upper bound on the difference between Φ and Ψ, as a function of the differences between Φ_k and Ψ_k (for $k = 1, \ldots, N$), is obtained by induction from statement 4 of Proposition 3.47 along with Corollary 3.50:

$$\big\| \Phi - \Psi \big\|_1 \leq \sum_{k=1}^{N} \big\| \Phi_k - \Psi_k \big\|_1. \tag{3.305}$$

Thus, irrespective of the specific properties of the network under consideration, the differences between the channels Φ_k and Ψ_k for $k = 1, \ldots, N$ can only accumulate additively when they are composed into a network.

Discrimination between pairs of isometric channels

As Example 3.39 illustrates, it is necessary in some instances of Scenario 3.38 for Bob to use an auxiliary register W in order to optimally discriminate a given pair of channels. One interesting case in which it is *not* necessary for

Bob to make use of an auxiliary register in this scenario is when the two channels are isometric channels, defined as

$$\Phi_0(X) = V_0 X V_0^* \quad \text{and} \quad \Phi_1(X) = V_1 X V_1^* \tag{3.306}$$

for all $X \in \mathrm{L}(\mathcal{X})$, for some choice of isometries $V_0, V_1 \in \mathrm{U}(\mathcal{X}, \mathcal{Y})$. The fact that an auxiliary register is not needed for an optimal discrimination in this case is proved below. The proof makes use of the notion of the *numerical range* of an operator.

Definition 3.55. Let $\mathcal{X}$ be a complex Euclidean space and let $A \in \mathrm{L}(\mathcal{X})$ be an operator. The *numerical range* of A is the set $\mathcal{N}(A) \subset \mathbb{C}$ defined as follows:

$$\mathcal{N}(A) = \bigl\{ u^* A u \,:\, u \in \mathcal{S}(\mathcal{X}) \bigr\}. \tag{3.307}$$

In general, every eigenvalue of a given operator A is contained in $\mathcal{N}(A)$, and one may prove that $\mathcal{N}(A)$ is equal to the convex hull of the eigenvalues of A in the case that A is normal. For non-normal operators, however, this will not generally be the case. It is, however, always the case that $\mathcal{N}(A)$ is compact and convex, which is the content of the following theorem.

Theorem 3.56 (Toeplitz–Hausdorff theorem)**.** *For any complex Euclidean space $\mathcal{X}$ and any operator $A \in \mathrm{L}(\mathcal{X})$, the set $\mathcal{N}(A)$ is compact and convex.*

Proof. The function $f : \mathcal{S}(\mathcal{X}) \to \mathbb{C}$ defined by $f(u) = u^* A u$ is continuous, and the unit sphere $\mathcal{S}(\mathcal{X})$ is compact. Continuous functions map compact sets to compact sets, implying that $\mathcal{N}(A) = f(\mathcal{S}(\mathcal{X}))$ is compact.

It remains to prove that $\mathcal{N}(A)$ is convex. Fix any choice of $\alpha, \beta \in \mathcal{N}(A)$ and a real number $\lambda \in [0,1]$. It will be proved that

$$\lambda\alpha + (1-\lambda)\beta \in \mathcal{N}(A), \tag{3.308}$$

which suffices to prove the theorem. It will be assumed hereafter that $\alpha \neq \beta$, as the assertion is trivial in the case that $\alpha = \beta$.

By the definition of the numerical range, one may choose unit vectors $u, v \in \mathcal{S}(\mathcal{X})$ such that $u^* A u = \alpha$ and $v^* A v = \beta$. It follows from the fact that $\alpha \neq \beta$ that the vectors u and v are linearly independent.

Next, define

$$B = \frac{-\beta}{\alpha - \beta} \mathbb{1}_{\mathcal{X}} + \frac{1}{\alpha - \beta} A \tag{3.309}$$

so that $u^* B u = 1$ and $v^* B v = 0$. Let

$$X = \frac{B + B^*}{2} \quad \text{and} \quad Y = \frac{B - B^*}{2i} \tag{3.310}$$

represent the Hermitian and anti-Hermitian parts of B. It follows that

$$\begin{aligned} u^* X u &= 1, \qquad & v^* X v &= 0, \\ u^* Y u &= 0, \qquad & v^* Y v &= 0. \end{aligned} \tag{3.311}$$

Without loss of generality, it may be assumed that $u^* Y v$ is purely imaginary (i.e., has real part equal to 0), for otherwise v may be replaced by $e^{i\theta} v$ for an appropriate choice of θ without changing any of the previously observed properties.

As u and v are linearly independent, the vector $tu + (1-t)v$ is nonzero for every choice of $t \in \mathbb{R}$. Thus, for each $t \in [0,1]$, one may define a unit vector

$$z(t) = \frac{tu + (1-t)v}{\| tu + (1-t)v \|}. \tag{3.312}$$

Because $u^* Y u = v^* Y v = 0$ and $u^* Y v$ is purely imaginary, it follows that $z(t)^* Y z(t) = 0$ for every $t \in [0,1]$, and therefore

$$z(t)^* B z(t) = z(t)^* X z(t) = \frac{t^2 + 2t(1-t)\Re(v^* X u)}{\| tu + (1-t)v \|}. \tag{3.313}$$

The expression on the right-hand side of (3.313) is a continuous real-valued function mapping 0 to 0 and 1 to 1. Consequently, there must exist at least one choice of $t \in [0,1]$ such that $z(t)^* B z(t) = \lambda$. Let $w = z(t)$ for such choice of t, so that $w^* B w = \lambda$. It holds that w is a unit vector, and

$$\begin{aligned} w^* A w &= (\alpha - \beta)\left(\frac{\beta}{\alpha - \beta} + w^* B w \right) \\ &= \beta + \lambda(\alpha - \beta) = \lambda\alpha + (1-\lambda)\beta. \end{aligned} \tag{3.314}$$

It has therefore been shown that $\lambda\alpha + (1-\lambda)\beta \in \mathcal{N}(A)$ as required. □

Theorem 3.57. *Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces, let $V_0, V_1 \in \mathrm{U}(\mathcal{X}, \mathcal{Y})$ be isometries, and let $\Phi_0, \Phi_1 \in \mathrm{C}(\mathcal{X}, \mathcal{Y})$ be channels defined as*

$$\Phi_0(X) = V_0 X V_0^* \quad \textit{and} \quad \Phi_1(X) = V_1 X V_1^* \tag{3.315}$$

for all $X \in \mathrm{L}(\mathcal{X})$. *There exists a unit vector* $u \in \mathcal{X}$ *such that*

$$\big\| \lambda \Phi_0(uu^*) - (1-\lambda)\Phi_1(uu^*) \big\|_1 = \big\| \lambda \Phi_0 - (1-\lambda)\Phi_1 \big\|_1 \tag{3.316}$$

for every $\lambda \in [0,1]$.

Proof. Using the identity (1.179), one finds that

$$\begin{aligned} &\big\| \lambda \Phi_0(uu^*) - (1-\lambda)\Phi_1(uu^*) \big\|_1 \\ &\qquad = \sqrt{1 - 4\lambda(1-\lambda)\left| u^* V_0^* V_1 u \right|^2}, \end{aligned} \tag{3.317}$$

for every unit vector $u \in \mathcal{X}$, and similarly

$$\begin{aligned} &\big\| \lambda\big(\Phi_0 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})}\big)(vv^*) - (1-\lambda)\big(\Phi_1 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})}\big)(vv^*) \big\|_1 \\ &\qquad = \sqrt{1 - 4\lambda(1-\lambda)\left| v^*\big(V_0^* V_1 \otimes \mathbb{1}_{\mathcal{Z}}\big) v \right|^2} \end{aligned} \tag{3.318}$$

for every complex Euclidean space $\mathcal{Z}$ and unit vector $v \in \mathcal{X} \otimes \mathcal{Z}$. Taking $\mathcal{Z}$ be a complex Euclidean space with $\dim(\mathcal{Z}) = \dim(\mathcal{X})$, it follows from (3.318) (along with Theorem 3.53) that there exists a unit vector $v \in \mathcal{X} \otimes \mathcal{Z}$ such that

$$\big\| \lambda \Phi_0 - (1-\lambda)\Phi_1 \big\|_1 = \sqrt{1 - 4\lambda(1-\lambda)\left| v^*\big(V_0^* V_1 \otimes \mathbb{1}_{\mathcal{Z}}\big) v \right|^2}. \tag{3.319}$$

Now, one may observe that

$$v^*\big(V_0^* V_1 \otimes \mathbb{1}_{\mathcal{Z}}\big) v = \big\langle \rho, V_0^* V_1 \big\rangle \tag{3.320}$$

for $\rho = \mathrm{Tr}_{\mathcal{Z}}(vv^*)$. By considering a spectral decomposition of the density operator ρ, one finds that the value represented by (3.320) is a convex combination of values of the form

$$w^* V_0^* V_1 w, \tag{3.321}$$

in which $w \in \mathcal{X}$ ranges over a set of unit eigenvectors of ρ. Each of these values is contained in the numerical range of $V_0^* V_1$, and therefore, by the Toeplitz–Hausdorff theorem (Theorem 3.56), there must exist a unit vector $u \in \mathcal{X}$ such that

$$u^* V_0^* V_1 u = \big\langle \rho, V_0^* V_1 \big\rangle. \tag{3.322}$$

By (3.317), it follows that

$$\big\| \lambda \Phi_0(uu^*) - (1-\lambda)\Phi_1(uu^*) \big\|_1 = \big\| \lambda \Phi_0 - (1-\lambda)\Phi_1 \big\|_1. \tag{3.323}$$

Observing that the vector u does not depend on λ, the proof is complete. □

The completely bounded trace distance from a channel to the identity

Returning once again to Example 3.39, one has that the Werner–Holevo channels can be perfectly discriminated through the use of a sufficiently large auiliary register, but are nearly indistinguishable (when the channels themselves are defined with respect to sufficiently large registers) without the use of an auxiliary register. The Werner–Holevo channels happen to have another feature, which is that they are highly noisy channels; their outputs are close to the completely mixed state for every possible input state.

One may ask if a similar phenomenon, in which an auxiliary register has a dramatic effect on the optimal probability of successfully discriminating channels, occurs when one of the channels is the identity channel. This is a natural question, as the closeness of a given channel to the identity channel may be a highly relevant figure of merit of that channel in some situations. The following theorem demonstrates that the phenomenon suggested above is limited in this setting. In particular, the theorem demonstrates that the potential advantage of using an auxiliary register in discriminating a given channel from the identity channel is dimension-independent.

Theorem 3.58. *Let* $\mathcal{X}$ *be a complex Euclidean space, let* $\Phi \in \mathrm{C}(\mathcal{X})$ *be a channel, let* $\varepsilon \in [0,1]$, *and suppose that*

$$\big\| \Phi(\rho) - \rho \big\|_1 \leq \varepsilon \tag{3.324}$$

for every density operator $\rho \in \mathrm{D}(\mathcal{X})$. *It holds that*

$$\big\| \Phi - \mathbb{1}_{\mathrm{L}(\mathcal{X})} \big\|_1 \leq 2\sqrt{\varepsilon}. \tag{3.325}$$

Proof. It is evident from the assumptions of the theorem that, for every unit vector $u \in \mathcal{X}$, one has

$$\big\| \Phi(uu^*) - uu^* \big\|_1 \leq \varepsilon, \tag{3.326}$$

and therefore

$$\big| \big\langle uu^*, \Phi(uu^*) - uu^* \big\rangle \big| \leq \varepsilon. \tag{3.327}$$

The first main step of the proof will be to establish a bound of a similar nature:

$$\big| \big\langle uv^*, \Phi(uv^*) - uv^* \big\rangle \big| \leq \varepsilon, \tag{3.328}$$

for every pair of orthogonal unit vectors $u, v \in \mathcal{X}$. Toward this goal, assume that $u, v \in \mathcal{X}$ are orthogonal unit vectors, and define a unit vector

$$w_k = \frac{u + i^k v}{\sqrt{2}} \tag{3.329}$$

for each $k \in \{0,1,2,3\}$. From the observation that

$$uv^* = \frac{1}{2}\sum_{k=0}^{3} i^k w_k w_k^*, \tag{3.330}$$

it follows that

$$\Phi(uv^*) - uv^* = \frac{1}{2}\sum_{k=0}^{3} i^k \bigl(\Phi(w_k w_k^*) - w_k w_k^*\bigr). \tag{3.331}$$

Because the spectral norm of a traceless Hermitian operator is at most one-half of its trace norm, it follows that

$$\begin{aligned} \bigl\|\Phi(uv^*) - uv^*\bigr\| &\leq \frac{1}{2}\sum_{k=0}^{3}\bigl\|\Phi(w_k w_k^*) - w_k w_k^*\bigr\| \\ &\leq \frac{1}{4}\sum_{k=0}^{3}\bigl\|\Phi(w_k w_k^*) - w_k w_k^*\bigr\|_1 \leq \varepsilon. \end{aligned} \tag{3.332}$$

This implies the desired bound (3.328).

Now, let $z \in \mathcal{X}\otimes\mathcal{X}$ be a unit vector, expressed in the form of a Schmidt decomposition

$$z = \sum_{a\in\Sigma} \sqrt{p(a)}\, x_a \otimes y_a, \tag{3.333}$$

for Σ being an alphabet, $\{x_a : a \in \Sigma\}$ and $\{y_a : a \in \Sigma\}$ being orthonormal subsets of $\mathcal{X}$, and $p \in \mathcal{P}(\Sigma)$ being a probability vector. It holds that

$$\bigl\langle zz^*, \bigl(\Phi\otimes\mathbb{1}_{\mathrm{L}(\mathcal{X})}\bigr)(zz^*)\bigr\rangle = \sum_{a,b\in\Sigma} p(a)p(b)\bigl\langle x_a x_b^*, \Phi(x_a x_b^*)\bigr\rangle, \tag{3.334}$$

and therefore, by the triangle inequality and the bounds (3.327) and (3.328) from above,

$$\begin{aligned} 1 - \bigl\langle zz^*, \bigl(\Phi\otimes\mathbb{1}_{\mathrm{L}(\mathcal{X})}\bigr)(zz^*)\bigr\rangle &= \bigl|\bigl\langle zz^*, \bigl(\Phi\otimes\mathbb{1}_{\mathrm{L}(\mathcal{X})}\bigr)(zz^*) - zz^*\bigr\rangle\bigr| \\ &\leq \sum_{a,b\in\Sigma} p(a)p(b)\bigl|\bigl\langle x_a x_b^*, \Phi(x_a x_b^*) - x_a x_b^*\bigr\rangle\bigr| \leq \varepsilon. \end{aligned} \tag{3.335}$$

Using the expression of the fidelity function when one of its arguments has rank equal to one, as given by Proposition 3.13, it follows that

$$\mathrm{F}\bigl(\bigl(\Phi\otimes\mathbb{1}_{\mathrm{L}(\mathcal{X})}\bigr)(zz^*), zz^*\bigr) \geq \sqrt{1-\varepsilon}. \tag{3.336}$$

Therefore, by one of the Fuchs–van de Graaf inequalities (Theorem 3.36), it follows that

$$\begin{aligned} &\bigl\|\bigl(\Phi\otimes\mathbb{1}_{\mathrm{L}(\mathcal{X})}\bigr)(zz^*) - zz^*\bigr\|_1 \\ &\quad\leq 2\sqrt{1 - \mathrm{F}\bigl(\bigl(\Phi\otimes\mathbb{1}_{\mathrm{L}(\mathcal{X})}\bigr)(zz^*), zz^*\bigr)^2} \leq 2\sqrt{\varepsilon}. \end{aligned} \tag{3.337}$$

Because $\Phi - \mathbb{1}_{\mathrm{L}(\mathcal{X})}$ is a Hermiticity preserving map, the theorem follows by Theorem 3.53. □

3.3.4 Properties of the completely bounded trace norm

This section discusses additional facts concerning the completely bounded trace norm. A few alternative characterizations of the completely bounded trace norm are presented, along with a theorem concerning the completely bounded trace norm of maps having bounded Choi rank.

The maximum output fidelity between completely positive maps

It is possible to characterize the completely bounded trace norm of a map in terms of the *maximum output fidelity* between two completely positive maps derived from the given map. The maximum output fidelity is defined as follows.

Definition 3.59. Let $\Psi_0, \Psi_1 \in \mathrm{CP}(\mathcal{X},\mathcal{Y})$ be completely positive maps, for $\mathcal{X}$ and $\mathcal{Y}$ being complex Euclidean spaces. The *maximum output fidelity* between Ψ_0 and Ψ_1 is defined as

$$\mathrm{F}_{\max}(\Psi_0,\Psi_1) = \max_{\rho_0,\rho_1\in\mathrm{D}(\mathcal{X})} \mathrm{F}\bigl(\Psi_0(\rho_0), \Psi_1(\rho_1)\bigr). \tag{3.338}$$

For any choice of vectors of the form $u, v \in \mathcal{X}\otimes\mathcal{Y}$, for $\mathcal{X}$ and $\mathcal{Y}$ being arbitrary complex Euclidean spaces, Corollary 3.24 states that

$$\bigl\|\mathrm{Tr}_{\mathcal{Y}}\bigl(vu^*\bigr)\bigr\|_1 = \mathrm{F}\bigl(\mathrm{Tr}_{\mathcal{X}}\bigl(uu^*\bigr), \mathrm{Tr}_{\mathcal{X}}\bigl(vv^*\bigr)\bigr). \tag{3.339}$$

It is an extension of this fact that provides the link between the maximum output fidelity and the completely bounded trace norm. In considering this extension, it is convenient to isolate the fact represented by the lemma that follows.

Lemma 3.60. *Let $A_0, A_1 \in \mathrm{L}(\mathcal{X}, \mathcal{Y} \otimes \mathcal{Z})$ be operators, for $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ being complex Euclidean spaces, and let $\Psi_0, \Psi_1 \in \mathrm{CP}(\mathcal{X}, \mathcal{Z})$ and $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ be the maps defined as*

$$\begin{aligned} \Psi_0(X) &= \mathrm{Tr}_{\mathcal{Y}}\big(A_0 X A_0^*\big), \\ \Psi_1(X) &= \mathrm{Tr}_{\mathcal{Y}}\big(A_1 X A_1^*\big), \end{aligned} \tag{3.340}$$

and

$$\Phi(X) = \mathrm{Tr}_{\mathcal{Z}}\big(A_0 X A_1^*\big), \tag{3.341}$$

for every operator $X \in \mathrm{L}(\mathcal{X})$. Also let $\mathcal{W}$ be a complex Euclidean space and let $u_0, u_1 \in \mathcal{X} \otimes \mathcal{W}$ be vectors. It holds that

$$\big\| \big(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{W})}\big)\big(u_0 u_1^*\big) \big\|_1 = \mathrm{F}\big(\Psi_0\big(\mathrm{Tr}_{\mathcal{W}}\big(u_0 u_0^*\big)\big), \Psi_1\big(\mathrm{Tr}_{\mathcal{W}}\big(u_1 u_1^*\big)\big)\big). \tag{3.342}$$

Proof. Let $W \in \mathrm{U}(\mathcal{Y} \otimes \mathcal{Z} \otimes \mathcal{W}, \mathcal{Z} \otimes \mathcal{Y} \otimes \mathcal{W})$ be the operator defined by the equation

$$W(y \otimes z \otimes w) = z \otimes y \otimes w, \tag{3.343}$$

holding for all $y \in \mathcal{Y}$, $z \in \mathcal{Z}$, and $w \in \mathcal{W}$. In other words, W represents a reordering of tensor factors, from $\mathcal{Y} \otimes \mathcal{Z} \otimes \mathcal{W}$ to $\mathcal{Z} \otimes \mathcal{Y} \otimes \mathcal{W}$. It is evident that one has

$$\begin{aligned} \big(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{W})}\big)\big(u_0 u_1^*\big) &= \mathrm{Tr}_{\mathcal{Z}}\Big(\big(A_0 \otimes \mathbb{1}_{\mathcal{W}}\big) u_0 u_1^* \big(A_1^* \otimes \mathbb{1}_{\mathcal{W}}\big)\Big) \\ &= \mathrm{Tr}_{\mathcal{Z}}\Big(W\big(A_0 \otimes \mathbb{1}_{\mathcal{W}}\big) u_0 u_1^* \big(A_1^* \otimes \mathbb{1}_{\mathcal{W}}\big) W^*\Big). \end{aligned} \tag{3.344}$$

Applying Corollary 3.24, one has

$$\begin{aligned} &\big\| \big(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{W})}\big)\big(u_0 u_1^*\big) \big\|_1 \\ &\quad = \mathrm{F}\Big(\mathrm{Tr}_{\mathcal{Y}\otimes\mathcal{W}}\big(W\big(A_0 \otimes \mathbb{1}_{\mathcal{W}}\big) u_0 u_0^* \big(A_0^* \otimes \mathbb{1}_{\mathcal{W}}\big) W^*\big), \\ &\qquad\qquad \mathrm{Tr}_{\mathcal{Y}\otimes\mathcal{W}}\big(W\big(A_1 \otimes \mathbb{1}_{\mathcal{W}}\big) u_1 u_1^* \big(A_1^* \otimes \mathbb{1}_{\mathcal{W}}\big) W^*\big)\Big) \\ &\quad = \mathrm{F}\Big(\mathrm{Tr}_{\mathcal{Y}}\big(A_0\, \mathrm{Tr}_{\mathcal{W}}\big(u_0 u_0^*\big) A_0^*\big), \mathrm{Tr}_{\mathcal{Y}}\big(A_1\, \mathrm{Tr}_{\mathcal{W}}\big(u_1 u_1^*\big) A_1^*\big)\Big) \\ &\quad = \mathrm{F}\big(\Psi_0\big(\mathrm{Tr}_{\mathcal{W}}\big(u_0 u_0^*\big)\big), \Psi_1\big(\mathrm{Tr}_{\mathcal{W}}\big(u_1 u_1^*\big)\big)\big), \end{aligned} \tag{3.345}$$

as required. □

Theorem 3.61. *Let $A_0, A_1 \in \mathrm{L}(\mathcal{X}, \mathcal{Y} \otimes \mathcal{Z})$ be operators, for $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ being complex Euclidean spaces, and let $\Psi_0, \Psi_1 \in \mathrm{CP}(\mathcal{X}, \mathcal{Z})$ and $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ be the maps defined as*

$$\begin{aligned} \Psi_0(X) &= \mathrm{Tr}_{\mathcal{Y}}\big(A_0 X A_0^*\big), \\ \Psi_1(X) &= \mathrm{Tr}_{\mathcal{Y}}\big(A_1 X A_1^*\big), \end{aligned} \tag{3.346}$$

and

$$\Phi(X) = \mathrm{Tr}_{\mathcal{Z}}\big(A_0 X A_1^*\big), \tag{3.347}$$

for every operator $X \in \mathrm{L}(\mathcal{X})$. It holds that

$$\big|\!\big|\!\big|\Phi\big|\!\big|\!\big|_1 = \mathrm{F}_{\max}(\Psi_0, \Psi_1). \tag{3.348}$$

Proof. Let $\mathcal{W}$ be a complex Euclidean space with $\dim(\mathcal{W}) = \dim(\mathcal{X})$. By Proposition 3.47 and Lemma 3.60, one has

$$\begin{aligned} \big|\!\big|\!\big|\Phi\big|\!\big|\!\big|_1 &= \max_{u_0, u_1 \in \mathcal{S}(\mathcal{X}\otimes\mathcal{W})} \big\| \big(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{W})}\big)\big(u_0 u_1^*\big) \big\|_1 \\ &= \max_{u_0, u_1 \in \mathcal{S}(\mathcal{X}\otimes\mathcal{W})} \mathrm{F}\big(\Psi_0\big(\mathrm{Tr}_{\mathcal{W}}\big(u_0 u_0^*\big)\big), \Psi_1\big(\mathrm{Tr}_{\mathcal{W}}\big(u_1 u_1^*\big)\big)\big) \\ &= \max_{\rho_0, \rho_1 \in \mathrm{D}(\mathcal{X})} \mathrm{F}\big(\Psi_0(\rho_0), \Psi_1(\rho_1)\big) \\ &= \mathrm{F}_{\max}(\Psi_0, \Psi_1), \end{aligned} \tag{3.349}$$

as required. □

Remark 3.62. The proof of Theorem 3.61 establishes a connection between those choices of density operators $\rho_0, \rho_1 \in \mathrm{D}(\mathcal{X})$ achieving the maximal value in the expression

$$\mathrm{F}_{\max}(\Psi_0, \Psi_1) = \max_{\rho_0, \rho_1 \in \mathrm{D}(\mathcal{X})} \mathrm{F}\big(\Psi_0(\rho_0), \Psi_1(\rho_1)\big) \tag{3.350}$$

and the choices of vectors $u_0, u_1 \in \mathcal{S}(\mathcal{X} \otimes \mathcal{W})$ achieving the maximal value in the expression

$$\big|\!\big|\!\big|\Phi\big|\!\big|\!\big|_1 = \max_{u_0, u_1 \in \mathcal{S}(\mathcal{X}\otimes\mathcal{W})} \big\| \big(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{W})}\big)\big(u_0 u_1^*\big) \big\|_1. \tag{3.351}$$

Specifically, for any choice of unit vectors $u_0, u_1 \in \mathcal{S}(\mathcal{X} \otimes \mathcal{W})$, one may take

$$\rho_0 = \mathrm{Tr}_{\mathcal{W}}\big(u_0 u_0^*\big) \quad \text{and} \quad \rho_1 = \mathrm{Tr}_{\mathcal{W}}\big(u_1 u_1^*\big), \tag{3.352}$$

and conversely, for any choice of density operators $\rho_0, \rho_1 \in \mathrm{D}(\mathcal{X})$, one may take $u_0, u_1 \in \mathcal{S}(\mathcal{X} \otimes \mathcal{W})$ to be arbitrary purifications of ρ_0, ρ_1, respectively, with equal values being obtained in the above expressions in both cases.

By combining Theorem 3.61 with the multiplicativity of the completely bounded trace norm with respect to tensor products (Theorem 3.51), one finds that the maximum output fidelity is also multiplicative with respect to tensor products.

Corollary 3.63. *Let $\mathcal{X}_0$, $\mathcal{X}_1$, $\mathcal{Y}_0$, and $\mathcal{Y}_1$ be complex Euclidean spaces and let $\Phi_0, \Psi_0 \in \mathrm{CP}(\mathcal{X}_0, \mathcal{Y}_0)$ and $\Phi_1, \Psi_1 \in \mathrm{CP}(\mathcal{X}_1, \mathcal{Y}_1)$ be completely positive maps. It holds that*

$$\mathrm{F}_{\max}(\Phi_0 \otimes \Phi_1, \Psi_0 \otimes \Psi_1) = \mathrm{F}_{\max}(\Phi_0, \Psi_0)\, \mathrm{F}_{\max}(\Phi_1, \Psi_1). \tag{3.353}$$

This corollary states a fact that is simple but not necessarily obvious: the maximum output fidelity between two completely positive product maps is achieved for product state inputs. It may be contrasted with some other quantities of interest (such as the minimum output entropy of a quantum channel, to be discussed in Chapter 7) that fail to respect tensor products in this way.

A semidefinite program for maximum output fidelity

It is natural to ask if the value $|||\Phi|||_1$ of the completely bounded trace norm of a given map $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ can be efficiently calculated. While there is no closed-form expression that is known to represent this value, it is equal to the optimal value of a semidefinite program that has a simple description in terms of the mapping Φ. In particular, when Theorem 3.61 is combined with the semidefinite program for the fidelity function discussed in Section 3.1.2, a semidefinite program for the completely bounded trace norm is obtained. This allows for an efficient calculation of the value $|||\Phi|||_1$ using a computer, as well as an efficient method of verification through the use of semidefinite programming duality.

In greater detail, let $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ be a map, for complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$, and assume that a Stinespring representation of Φ is known:

$$\Phi(X) = \mathrm{Tr}_{\mathcal{Z}}\bigl(A_0 X A_1^*\bigr) \tag{3.354}$$

for all operators $X \in \mathrm{L}(\mathcal{X})$, for $\mathcal{Z}$ being a complex Euclidean space and $A_0, A_1 \in \mathrm{L}(\mathcal{X}, \mathcal{Y} \otimes \mathcal{Z})$ being operators. Define completely positive maps $\Psi_0, \Psi_1 \in \mathrm{CP}(\mathcal{X}, \mathcal{Z})$ as

$$\Psi_0(X) = \mathrm{Tr}_{\mathcal{Y}}\bigl(A_0 X A_0^*\bigr) \quad \text{and} \quad \Psi_1(X) = \mathrm{Tr}_{\mathcal{Y}}\bigl(A_1 X A_1^*\bigr) \tag{3.355}$$

for all $X \in \mathrm{L}(\mathcal{X})$, and consider the semidefinite program whose primal problem is as follows:

$$\begin{array}{rl} & \underline{\text{Primal problem}} \\ \text{maximize:} & \dfrac{1}{2}\mathrm{Tr}(Y) + \dfrac{1}{2}\mathrm{Tr}(Y^*) \\ \text{subject to:} & \begin{pmatrix} \Psi_0(\rho_0) & Y \\ Y^* & \Psi_1(\rho_1) \end{pmatrix} \geq 0 \\ & \rho_0, \rho_1 \in \mathrm{D}(\mathcal{X}),\ Y \in \mathrm{L}(\mathcal{Z}). \end{array}$$

Such a semidefinite program may be expressed with greater formality, with respect to the definition of semidefinite programs presented in Section 1.2.3, in the following way. First, one defines a Hermiticity-preserving map

$$\Xi : \mathrm{L}(\mathcal{X} \oplus \mathcal{X} \oplus \mathcal{Z} \oplus \mathcal{Z}) \to \mathrm{L}(\mathbb{C} \oplus \mathbb{C} \oplus \mathcal{Z} \oplus \mathcal{Z}) \tag{3.356}$$

as

$$\begin{aligned} &\Xi \begin{pmatrix} X_0 & \cdot & \cdot & \cdot \\ \cdot & X_1 & \cdot & \cdot \\ \cdot & \cdot & Z_0 & \cdot \\ \cdot & \cdot & \cdot & Z_1 \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} \mathrm{Tr}(X_0) & 0 & 0 & 0 \\ 0 & \mathrm{Tr}(X_1) & 0 & 0 \\ 0 & 0 & Z_0 - \Psi_0(X_0) & 0 \\ 0 & 0 & 0 & Z_1 - \Psi_1(X_1) \end{pmatrix} \end{aligned} \tag{3.357}$$

for all $X_0, X_1 \in \mathrm{L}(\mathcal{X})$ and $Z_0, Z_1 \in \mathrm{L}(\mathcal{Z})$, and where the dots represent operators on appropriately chosen spaces upon which Ξ does not depend. Then, one may define Hermitian operators $A \in \mathrm{Herm}(\mathcal{X} \oplus \mathcal{X} \oplus \mathcal{Z} \oplus \mathcal{Z})$ and $B \in \mathrm{Herm}(\mathbb{C} \oplus \mathbb{C} \oplus \mathcal{Z} \oplus \mathcal{Z})$ as

$$A = \frac{1}{2} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathbb{1} \\ 0 & 0 & \mathbb{1} & 0 \end{pmatrix} \quad \text{and} \quad B = \frac{1}{2} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{3.358}$$

It is evident that the primal problem specified above is equivalent to the maximization of the quantity $\langle A, X \rangle$ over all choices of

$$X = \begin{pmatrix} X_0 & \cdot & \cdot & \cdot \\ \cdot & X_1 & \cdot & \cdot \\ \cdot & \cdot & Z_0 & Y \\ \cdot & \cdot & Y^* & Z_1 \end{pmatrix} \in \mathrm{Pos}(\mathcal{X} \oplus \mathcal{X} \oplus \mathcal{Z} \oplus \mathcal{Z}) \tag{3.359}$$

obeying the constraint $\Xi(X) = B$.

The adjoint mapping to Ξ is given by

$$\begin{aligned} &\Xi^* \begin{pmatrix} \lambda_0 & \cdot & \cdot & \cdot \\ \cdot & \lambda_1 & \cdot & \cdot \\ \cdot & \cdot & Z_0 & \cdot \\ \cdot & \cdot & \cdot & Z_1 \end{pmatrix} \\ &\qquad = \frac{1}{2} \begin{pmatrix} \lambda_0 \mathbb{1}_{\mathcal{X}} - \Psi_0^*(Z_0) & 0 & 0 & 0 \\ 0 & \lambda_1 \mathbb{1}_{\mathcal{X}} - \Psi_1^*(Z_1) & 0 & 0 \\ 0 & 0 & Z_0 & 0 \\ 0 & 0 & 0 & Z_1 \end{pmatrix}, \end{aligned} \tag{3.360}$$

so the dual problem corresponding to the semidefinite program (Ξ, A, B) is to minimize the quantity $(\lambda_0 + \lambda_1)/2$ subject to the conditions

$$\lambda_0 \mathbb{1}_{\mathcal{X}} \geq \Psi_0^*(Z_0) \quad \text{and} \quad \lambda_1 \mathbb{1}_{\mathcal{X}} \geq \Psi_1^*(Z_1), \tag{3.361}$$

for $Z_0, Z_1 \in \mathrm{Herm}(\mathcal{Z})$ being Hermitian operators satisfying

$$\begin{pmatrix} Z_0 & 0 \\ 0 & Z_1 \end{pmatrix} \geq \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}. \tag{3.362}$$

Observing that Z_0 and Z_1 must be positive definite in order for (3.362) to be satisfied, along with the fact that Ψ_0^* and Ψ_1^* are positive, one obtains the following statement of the dual problem:

Dual problem

$$\begin{aligned} \text{minimize:} \quad & \frac{1}{2}\|\Psi_0^*(Z_0)\| + \frac{1}{2}\|\Psi_1^*(Z_1)\| \\ \text{subject to:} \quad & \begin{pmatrix} Z_0 & -\mathbb{1}_{\mathcal{Z}} \\ -\mathbb{1}_{\mathcal{Z}} & Z_1 \end{pmatrix} \geq 0 \\ & Z_0, Z_1 \in \mathrm{Pd}(\mathcal{Z}). \end{aligned}$$

To prove that strong duality holds for this semidefinite program, one may observe that the primal problem is feasible and the dual problem is strictly feasible. With respect to the formal specification of the semidefinite program just described, one has that the operator

$$\begin{pmatrix} \rho_0 & 0 & 0 & 0 \\ 0 & \rho_1 & 0 & 0 \\ 0 & 0 & \Psi_0(\rho_0) & 0 \\ 0 & 0 & 0 & \Psi_1(\rho_1) \end{pmatrix} \tag{3.363}$$

is primal feasible, for an arbitrary choice of density operators $\rho_0, \rho_1 \in \mathrm{D}(\mathcal{X})$. One also has that the dual problem is strictly feasible: the operator

$$\begin{pmatrix} \lambda_0 & 0 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & \mathbb{1}_{\mathcal{Z}} & 0 \\ 0 & 0 & 0 & \mathbb{1}_{\mathcal{Z}} \end{pmatrix} \tag{3.364}$$

is strictly dual feasible provided that $\lambda_0 > \|\Psi_0^*(\mathbb{1}_{\mathcal{Z}})\|$ and $\lambda_1 > \|\Psi_1^*(\mathbb{1}_{\mathcal{Z}})\|$. It follows by Slater's theorem (Theorem 1.18) that the primal and dual optimal values are equal, and moreover the primal optimal value is achieved for some choice of a primal feasible operator.

The fact that the optimal value of the semidefinite program is in agreement with the completely bounded norm $\|\!|\!|\Phi|\!|\!|_1$ follows from Theorem 3.61 together with Theorem 3.17.

It may be noted that the dual problem stated above may be further simplified as follows:

Dual problem (simplified)

$$\begin{aligned} \text{minimize:} \quad & \frac{1}{2}\|\Psi_0^*(Z)\| + \frac{1}{2}\|\Psi_1^*(Z^{-1})\| \\ \text{subject to:} \quad & Z \in \mathrm{Pd}(\mathcal{Y}). \end{aligned}$$

To verify that this problem has the same optimal value as the dual problem stated above, one may first observe that

$$\begin{pmatrix} Z_0 & -\mathbb{1} \\ -\mathbb{1} & Z_1 \end{pmatrix} \tag{3.365}$$

is positive semidefinite if and only if Z_0 and Z_1 are both positive definite and satisfy $Z_1 \geq Z_0^{-1}$. For any such choice of Z_0 and Z_1, the inequality

$$\big\|\Psi_1^*(Z_1)\big\| \geq \big\|\Psi_1^*(Z_0^{-1})\big\| \tag{3.366}$$

holds by the positivity of Ψ_1^*, implying that there is no loss of generality in restricting one's attention to operators $Z_0 = Z$ and $Z_1 = Z^{-1}$ for $Z \in \mathrm{Pd}(\mathcal{Z})$.

The observation that the simplified dual problem above has optimal value $\|\!|\!|\Phi|\!|\!|_1$ may be stated in the form of a theorem as follows.

Theorem 3.64. *Let $A_0, A_1 \in \mathrm{L}(\mathcal{X}, \mathcal{Y} \otimes \mathcal{Z})$ be operators, for $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ being complex Euclidean spaces, and let $\Psi_0, \Psi_1 \in \mathrm{CP}(\mathcal{X}, \mathcal{Z})$ and $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ be the maps defined as*

$$\begin{aligned} \Psi_0(X) &= \mathrm{Tr}_{\mathcal{Y}}\big(A_0 X A_0^*\big), \\ \Psi_1(X) &= \mathrm{Tr}_{\mathcal{Y}}\big(A_1 X A_1^*\big), \end{aligned} \tag{3.367}$$

and

$$\Phi(X) = \mathrm{Tr}_{\mathcal{Z}}\big(A_0 X A_1^*\big), \tag{3.368}$$

for every operator $X \in \mathrm{L}(\mathcal{X})$. It holds that

$$\|\!|\!|\Phi|\!|\!|_1 = \inf_{Z \in \mathrm{Pd}(\mathcal{Z})} \left(\frac{1}{2}\big\|\Psi_0^*(Z)\big\| + \frac{1}{2}\big\|\Psi_1^*(Z^{-1})\big\| \right). \tag{3.369}$$

Spectral norm characterization of the completely bounded trace norm

Consider a map $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$, for complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$. One has, by Theorem 2.22, that a given complex Euclidean space $\mathcal{Z}$ admits a Stinespring representation

$$\Phi(X) = \mathrm{Tr}_{\mathcal{Z}}\big(A_0 X A_1^*\big) \tag{3.370}$$

of Φ, for some choice of operators $A_0, A_1 \in \mathrm{L}(\mathcal{X}, \mathcal{Y} \otimes \mathcal{Z})$, if and only if the dimension of $\mathcal{Z}$ is at least as large as the Choi rank of Φ. An equivalent condition to (3.370) holding for all operators $X \in \mathrm{L}(\mathcal{X})$ is that

$$J(\Phi) = \mathrm{Tr}_{\mathcal{Z}}\big(\mathrm{vec}(A_0)\,\mathrm{vec}(A_1)^*\big). \tag{3.371}$$

As the next theorem states, the completely bounded trace norm of Φ is equal to the infimum value of the product $\|A_0\|\|A_1\|$, ranging over all such choices of A_0 and A_1.

Theorem 3.65 (Smith). *Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces, let $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ be a map, let $\mathcal{Z}$ be a complex Euclidean space satisfying $\dim(\mathcal{Z}) \geq \mathrm{rank}(J(\Phi))$, and let*

$$\begin{aligned} \mathcal{K}_\Phi = \big\{ (A_0, A_1) \in \mathrm{L}(\mathcal{X}, \mathcal{Y} \otimes \mathcal{Z}) \times \mathrm{L}(\mathcal{X}, \mathcal{Y} \otimes \mathcal{Z}) \,:\, \\ J(\Phi) = \mathrm{Tr}_{\mathcal{Z}}\big(\mathrm{vec}(A_0)\,\mathrm{vec}(A_1)^*\big) \big\}. \end{aligned} \tag{3.372}$$

It holds that

$$\|\!|\!|\Phi|\!|\!|_1 = \inf_{(A_0, A_1) \in \mathcal{K}_\Phi} \|A_0\|\|A_1\|. \tag{3.373}$$

Proof. There exists a pair of unit vectors $u, v \in \mathcal{X} \otimes \mathcal{X}$ such that, for any pair of operators $(A_0, A_1) \in \mathcal{K}_\Phi$, one has

$$\|\!|\!|\Phi|\!|\!|_1 = \big\|\mathrm{Tr}_{\mathcal{Z}}\big((A_0 \otimes \mathbb{1}_{\mathcal{X}}) u v^* (A_1 \otimes \mathbb{1}_{\mathcal{X}})^*\big)\big\|_1. \tag{3.374}$$

By the monotonicity of the trace norm under partial tracing (1.178) and the multiplicativity of the spectral norm with respect to tensor products, it follows that

$$\begin{aligned} \|\!|\!|\Phi|\!|\!|_1 &\leq \big\|(A_0 \otimes \mathbb{1}_{\mathcal{X}}) u v^* (A_1 \otimes \mathbb{1}_{\mathcal{X}})^*\big\|_1 \\ &= \big\|(A_0 \otimes \mathbb{1}_{\mathcal{X}}) u\big\|\big\|(A_1 \otimes \mathbb{1}_{\mathcal{X}}) v\big\| \\ &\leq \|A_0 \otimes \mathbb{1}_{\mathcal{X}}\|\|A_1 \otimes \mathbb{1}_{\mathcal{X}}\| \\ &= \|A_0\|\|A_1\|. \end{aligned} \tag{3.375}$$

As this inequality holds for every pair $(A_0, A_1) \in \mathcal{K}_\Phi$, it follows that

$$\|\!|\!|\Phi|\!|\!|_1 \leq \inf_{(A_0, A_1) \in \mathcal{K}_\Phi} \|A_0\|\|A_1\|. \tag{3.376}$$

It remains to prove the reverse inequality. To this end, fix any pair of operators $(B_0, B_1) \in \mathcal{K}_\Phi$, and define $\Psi_0, \Psi_1 \in \mathrm{CP}(\mathcal{X}, \mathcal{Z})$ as

$$\Psi_0(X) = \mathrm{Tr}_{\mathcal{Y}}\big(B_0 X B_0^*\big) \quad \text{and} \quad \Psi_1(X) = \mathrm{Tr}_{\mathcal{Y}}\big(B_1 X B_1^*\big) \tag{3.377}$$

for all $X \in \mathrm{L}(\mathcal{X})$, so that

$$\Psi_0^*(Z) = B_0^*(\mathbb{1}_{\mathcal{Y}} \otimes Z) B_0 \quad \text{and} \quad \Psi_1^*(Z) = B_1^*(\mathbb{1}_{\mathcal{Y}} \otimes Z) B_1 \tag{3.378}$$

for every $Z \in \mathrm{L}(\mathcal{Z})$. By Theorem 3.64, the expression (3.369) holds. For any choice of a positive real number $\varepsilon > 0$, there must therefore exist a positive definite operator $Z \in \mathrm{Pd}(\mathcal{Z})$ so that

$$\frac{1}{2}\big\|\Psi_0^*(Z)\big\| + \frac{1}{2}\big\|\Psi_1^*(Z^{-1})\big\| < \|\!|\!|\Phi|\!|\!|_1 + \varepsilon. \tag{3.379}$$

By the arithmetic-geometric mean inequality, it follows that

$$\sqrt{\|\Psi_0^*(Z)\|}\sqrt{\|\Psi_1^*(Z^{-1})\|} < \left|\!\left|\!\left|\Phi\right|\!\right|\!\right|_1 + \varepsilon. \tag{3.380}$$

Setting

$$A_0 = \left(\mathbb{1}_{\mathcal{Y}} \otimes Z^{\frac{1}{2}}\right) B_0 \quad \text{and} \quad A_1 = \left(\mathbb{1}_{\mathcal{Y}} \otimes Z^{-\frac{1}{2}}\right) B_1, \tag{3.381}$$

one has that $(A_0, A_1) \in \mathcal{K}_\Phi$ by the cyclic property of the trace. Moreover, it holds that

$$\begin{aligned} \|A_0\|\|A_1\| &= \sqrt{\|A_0^* A_0\|}\sqrt{\|A_1^* A_1\|} \\ &= \sqrt{\|\Psi_0^*(Z)\|}\sqrt{\|\Psi_1^*(Z^{-1})\|} < \left|\!\left|\!\left|\Phi\right|\!\right|\!\right|_1 + \varepsilon. \end{aligned} \tag{3.382}$$

As it has been established that, for any choice of $\varepsilon > 0$, there exists a pair of operators $(A_0, A_1) \in \mathcal{K}_\Phi$ satisfying the inequality (3.382), it follows that

$$\inf_{(A_0,A_1)\in\mathcal{K}_\Phi} \|A_0\|\|A_1\| \leq \left|\!\left|\!\left|\Phi\right|\!\right|\!\right|_1, \tag{3.383}$$

which completes the proof. □

The completely bounded trace norm of maps with bounded Choi rank

For a given map $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ and a complex Euclidean space $\mathcal{Z}$, it holds (by Theorem 3.49) that

$$\left\|\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})}\right\|_1 \leq \left|\!\left|\!\left|\Phi\right|\!\right|\!\right|_1, \tag{3.384}$$

with equality under the condition that $\dim(\mathcal{Z}) \geq \dim(\mathcal{X})$. If it is the case that $\dim(\mathcal{Z}) < \dim(\mathcal{X})$, then equality may fail to hold. For instance, the transpose map $\mathrm{T}(X) = X^{\mathsf{T}}$ on an arbitrary complex Euclidean space $\mathcal{X}$ is such that

$$\left\|\mathrm{T} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})}\right\|_1 = \min\bigl\{\dim(\mathcal{X}), \dim(\mathcal{Z})\bigr\} \tag{3.385}$$

for every complex Euclidean space $\mathcal{Z}$.

It is the case, however, that equality holds in (3.384) under a different and generally incomparable assumption, which is that the dimension of $\mathcal{Z}$ is at least as large as the Choi rank of Φ, as the following theorem states.

Theorem 3.66 (Timoney). *Let $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ be complex Euclidean spaces, let $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ be a map, and assume $\dim(\mathcal{Z}) \geq \operatorname{rank}(J(\Phi))$. It holds that*

$$\left|\!\left|\!\left|\Phi\right|\!\right|\!\right|_1 = \left\|\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})}\right\|_1. \tag{3.386}$$

The proof of Theorem 3.66 to be presented below makes use of the following lemma.

Lemma 3.67. *Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces, let $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ be a positive map, and let $P \in \mathrm{Pos}(\mathcal{Y})$ be a nonzero positive semidefinite operator satisfying $P = \Phi(\rho)$ for some choice of a density operator $\rho \in \mathrm{D}(\mathcal{X})$. There exists a density operator $\sigma \in \mathrm{D}(\mathcal{X})$ with $\operatorname{rank}(\sigma) \leq \operatorname{rank}(P)$ that satisfies $P = \Phi(\sigma)$.*

Proof. Define a set

$$\mathcal{C} = \bigl\{\xi \in \mathrm{D}(\mathcal{X}) \,:\, \Phi(\xi) = P\bigr\}. \tag{3.387}$$

The set $\mathcal{C}$ is nonempty by the assumptions of the lemma, and it is evidently both compact and convex. There must therefore exist an extreme point of $\mathcal{C}$. Let σ be such an extreme point and let $r = \operatorname{rank}(\sigma)$. It will be proved that $r \leq \operatorname{rank}(P)$, which suffices to prove the lemma.

Let $n = \dim(\mathcal{X})$ and $m = \operatorname{rank}(P)$, and let $\Pi = \Pi_{\operatorname{im}(P)}$. Define a linear mapping $\Psi : \mathrm{Herm}(\mathcal{X}) \to \mathrm{Herm}(\mathcal{Y} \oplus \mathbb{C})$ as

$$\Psi(H) = \begin{pmatrix} \Pi\Phi(H)\Pi & 0 \\ 0 & \langle \mathbb{1}_{\mathcal{Y}} - \Pi, \Phi(H)\rangle \end{pmatrix} \tag{3.388}$$

for all $H \in \mathrm{Herm}(\mathcal{X})$. The image of Ψ has dimension at most $m^2 + 1$, and therefore the kernel of Ψ is a subspace of $\mathrm{Herm}(\mathcal{X})$ having dimension at least $n^2 - m^2 - 1$. Also define a subspace $\mathcal{W} \subseteq \mathrm{Herm}(\mathcal{X})$ as

$$\mathcal{W} = \bigl\{H \in \mathrm{Herm}(\mathcal{X}) \,:\, \operatorname{im}(H) \subseteq \operatorname{im}(\sigma) \text{ and } \operatorname{Tr}(H) = 0\bigr\}. \tag{3.389}$$

The dimension of $\mathcal{W}$ is equal to $r^2 - 1$.

Now consider any operator $H \in \ker(\Psi) \cap \mathcal{W}$. As $\operatorname{im}(H) \subseteq \operatorname{im}(\sigma)$ and σ is positive semidefinite, there must exist a positive real number $\varepsilon > 0$ for which $\sigma + \varepsilon H$ and $\sigma - \varepsilon H$ are both positive semidefinite. As H is traceless, it follow that $\sigma + \varepsilon H$ and $\sigma - \varepsilon H$ are density operators. By the assumption that $H \in \ker(\Psi)$, one has $\langle \mathbb{1}_{\mathcal{Y}} - \Pi, \Phi(H)\rangle = 0$, and therefore

$$\langle \mathbb{1}_{\mathcal{Y}} - \Pi, \Phi(\sigma + \varepsilon H)\rangle = \langle \mathbb{1}_{\mathcal{Y}} - \Pi, P + \varepsilon\Phi(H)\rangle = 0. \tag{3.390}$$

By the positivity of Φ, it follows that

$$\Phi(\sigma + \varepsilon H) = \Pi\Phi(\sigma + \varepsilon H)\Pi = P + \varepsilon\,\Pi\Phi(H)\Pi = P. \tag{3.391}$$

By similar reasoning, $\Phi(\sigma - \varepsilon H) = P$. It has therefore been proved that $\sigma + \varepsilon H$ and $\sigma - \varepsilon H$ are both elements of $\mathcal{C}$; but given that σ was chosen to be an extreme point of $\mathcal{C}$ and

$$\frac{1}{2}\big(\sigma + \varepsilon H\big) + \frac{1}{2}\big(\sigma - \varepsilon H\big) = \sigma, \tag{3.392}$$

it follows that $H = 0$. Consequently, the subspace $\ker(\Psi) \cap \mathcal{W}$ must have dimension 0.

Finally, given that $\mathrm{Herm}(\mathcal{X})$ has dimension n^2, $\ker(\Psi) \subseteq \mathrm{Herm}(\mathcal{X})$ has dimension at least $n^2 - m^2 - 1$, $\mathcal{W} \subseteq \mathrm{Herm}(\mathcal{X})$ has dimension $r^2 - 1$, and $\ker(\Psi) \cap \mathcal{W}$ has dimension 0, it follows that

$$\big(n^2 - m^2 - 1\big) + \big(r^2 - 1\big) \leq n^2, \tag{3.393}$$

and therefore

$$r^2 \leq m^2 + 2. \tag{3.394}$$

As r and m are positive integers, it follows that $r \leq m$, which completes the proof. □

Proof of Theorem 3.66. One may choose operators $A_0, A_1 \in \mathrm{L}(\mathcal{X}, \mathcal{Y} \otimes \mathcal{Z})$ such that

$$\Phi(X) = \mathrm{Tr}_{\mathcal{Z}}\big(A_0 X A_1^*\big) \tag{3.395}$$

for all $X \in \mathrm{L}(\mathcal{X})$, by Theorem 2.22. By Theorem 3.61, it follows that

$$\big|\!\big|\!\big|\Phi\big|\!\big|\!\big|_1 = \mathrm{F}_{\max}(\Psi_0, \Psi_1) \tag{3.396}$$

for $\Psi_0, \Psi_1 \in \mathrm{CP}(\mathcal{X}, \mathcal{Z})$ being the completely positive maps defined by

$$\Psi_0(X) = \mathrm{Tr}_{\mathcal{Y}}\big(A_0 X A_0^*\big) \quad \text{and} \quad \Psi_1(X) = \mathrm{Tr}_{\mathcal{Y}}\big(A_1 X A_1^*\big) \tag{3.397}$$

for all $X \in \mathrm{L}(\mathcal{X})$. Let $\rho_0, \rho_1 \in \mathrm{D}(\mathcal{X})$ be density operators that satisfy

$$\mathrm{F}\big(\Psi_0(\rho_0), \Psi_1(\rho_1)\big) = \mathrm{F}_{\max}(\Psi_0, \Psi_1) = \big|\!\big|\!\big|\Phi\big|\!\big|\!\big|_1. \tag{3.398}$$

The operators $P_0 = \Psi_0(\rho_0)$ and $P_1 = \Psi_1(\rho_1)$ are elements of $\mathrm{Pos}(\mathcal{Z})$, and therefore their rank cannot exceed the dimension of $\mathcal{Z}$. It follows from

Lemma 3.67 that there exist density operators $\sigma_0, \sigma_1 \in \mathrm{D}(\mathcal{X})$, whose rank also does not exceed the dimension of $\mathcal{Z}$, for which it holds that

$$\Psi_0(\sigma_0) = P_0 \quad \text{and} \quad \Psi_1(\sigma_1) = P_1. \tag{3.399}$$

Therefore, one has that

$$\mathrm{F}\big(\Psi_0(\sigma_0), \Psi_1(\sigma_1)\big) = \big|\!\big|\!\big|\Phi\big|\!\big|\!\big|_1. \tag{3.400}$$

Because σ_0 and σ_1 have rank at most the dimension of $\mathcal{Z}$, there must exist unit vectors $u_0, u_1 \in \mathcal{X} \otimes \mathcal{Z}$ satisfying

$$\sigma_0 = \mathrm{Tr}_{\mathcal{Z}}\big(u_0 u_0^*\big) \quad \text{and} \quad \sigma_1 = \mathrm{Tr}_{\mathcal{Z}}\big(u_1 u_1^*\big). \tag{3.401}$$

By Lemma 3.60, one has that

$$\big\|\big(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})}\big)(u_0 u_1^*)\big\|_1 = \mathrm{F}\big(\Psi_0(\sigma_0), \Psi_1(\sigma_1)\big) = \big|\!\big|\!\big|\Phi\big|\!\big|\!\big|_1, \tag{3.402}$$

which establishes that

$$\big\|\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})}\big\|_1 \geq \big|\!\big|\!\big|\Phi\big|\!\big|\!\big|_1. \tag{3.403}$$

As the reverse inequality holds by Theorem 3.49, the proof is complete. □

Corollary 3.68. *Let $\Phi_0, \Phi_1 \in \mathrm{C}(\mathcal{X}, \mathcal{Y})$ be channels, for complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$, and let $\mathcal{Z}$ be any complex Euclidean space with*

$$\dim(\mathcal{Z}) \geq 2\,\mathrm{rank}(J(\Phi_0 - \Phi_1)), \tag{3.404}$$

There exists a unit vector $u \in \mathcal{X} \otimes \mathcal{Z}$ such that

$$\big\|\big(\Phi_0 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})}\big)(uu^*) - \big(\Phi_1 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})}\big)(uu^*)\big\|_1 = \big|\!\big|\!\big|\Phi_0 - \Phi_1\big|\!\big|\!\big|_1. \tag{3.405}$$

Proof. The theorem is vacuous when $\Phi_0 = \Phi_1$, so it will be assumed that this is not the case. Let $\mathcal{W}$ be a complex Euclidean space having dimension equal to $\mathrm{rank}(J(\Phi_0 - \Phi_1))$. By Theorem 3.66, it holds that

$$\big|\!\big|\!\big|\Phi_0 - \Phi_1\big|\!\big|\!\big|_1 = \big\|\Phi_0 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{W})} - \Phi_1 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{W})}\big\|_1 \tag{3.406}$$

By Lemma 3.52, it follows that there exists a unit vector $v \in \mathcal{X} \otimes \mathcal{W} \otimes \mathcal{V}$, for $\mathcal{V}$ being any complex Euclidean space with dimension equal to 2, such that

$$\big\|\big(\Phi_0 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{W}\otimes\mathcal{V})}\big)(vv^*) - \big(\Phi_1 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{W}\otimes\mathcal{V})}\big)(vv^*)\big\|_1 \geq \big|\!\big|\!\big|\Phi_0 - \Phi_1\big|\!\big|\!\big|_1. \tag{3.407}$$

Now, under the assumption that $\dim(\mathcal{Z}) \geq 2\,\mathrm{rank}(J(\Phi_0 - \Phi_1))$, there must exist a linear isometry of the form $V \in \mathrm{U}(\mathcal{W} \otimes \mathcal{V}, \mathcal{Z})$. One may set

$$u = (\mathbb{1}_{\mathcal{X}} \otimes V)v \tag{3.408}$$

to obtain

$$\begin{aligned}
&\big\| \big(\Phi_0 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})}\big)(uu^*) - \big(\Phi_1 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})}\big)(uu^*) \big\|_1 \\
&\quad = \big\| (\mathbb{1}_{\mathcal{Y}} \otimes V)\big(\big(\Phi_0 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{W}\otimes\mathcal{V})}\big)(vv^*) \\
&\qquad\qquad - \big(\Phi_1 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{W}\otimes\mathcal{V})}\big)(vv^*)\big)(\mathbb{1}_{\mathcal{Y}} \otimes V^*)\big\|_1 \\
&\quad = \big\| \big(\Phi_0 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{W}\otimes\mathcal{V})}\big)(vv^*) - \big(\Phi_1 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{W}\otimes\mathcal{V})}\big)(vv^*) \big\|_1 \\
&\quad \geq \big\| \Phi_0 - \Phi_1 \big\|_1
\end{aligned} \tag{3.409}$$

by the isometric invariance of the trace norm together with (3.407). As the reverse inequality holds for all unit vectors $u \in \mathcal{X} \otimes \mathcal{Z}$ by Theorem 3.49, the proof is complete. □

3.4 Exercises

3.1. Let $\mathcal{X}$ be a complex Euclidean space, let $\rho_0, \rho_1 \in \mathrm{D}(\mathcal{X})$ be states, and let $\delta = \mathrm{F}(\rho_0, \rho_1)$. Also let n be a positive integer and define two new density operators as follows:

$$\begin{aligned}
\sigma_0 &= \frac{1}{2^{n-1}} \sum_{\substack{a_1,\ldots,a_n \in \{0,1\} \\ a_1 + \cdots + a_n \text{ even}}} \rho_{a_1} \otimes \cdots \otimes \rho_{a_n}, \\
\sigma_1 &= \frac{1}{2^{n-1}} \sum_{\substack{a_1,\ldots,a_n \in \{0,1\} \\ a_1 + \cdots + a_n \text{ odd}}} \rho_{a_1} \otimes \cdots \otimes \rho_{a_n}.
\end{aligned} \tag{3.410}$$

Prove that

$$\mathrm{F}(\sigma_0, \sigma_1) \geq 1 - \exp\left(-\frac{n\delta^2}{2}\right). \tag{3.411}$$

3.2. Suppose $\mathcal{X}$ and $\mathcal{Y}$ are complex Euclidean spaces, $P, Q \in \mathrm{Pos}(\mathcal{X})$ are positive semidefinite operators, and $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ is a trace-preserving and positive (but not necessarily completely positive) map. Prove that

$$\mathrm{F}(P, Q) \leq \mathrm{F}(\Phi(P), \Phi(Q)). \tag{3.412}$$

3.3. Find an example of two channels $\Phi_0, \Phi_1 \in \mathrm{C}(\mathcal{X}, \mathcal{Y})$, for some choice of complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$, such that

$$\big\| \Phi_0(\rho) - \Phi_1(\rho) \big\|_1 < \big\| \Phi_0 - \Phi_1 \big\|_1 \tag{3.413}$$

for every density operator $\rho \in \mathrm{D}(\mathcal{X})$.

3.4. Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces and let $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ be a map. Prove that

$$\big\|\!\big| \Phi \big|\!\big\|_1 = \max_{\rho_0, \rho_1 \in \mathrm{D}(\mathcal{X})} \big\| \big(\mathbb{1}_{\mathcal{Y}} \otimes \sqrt{\rho_0}\big) J(\Phi) \big(\mathbb{1}_{\mathcal{Y}} \otimes \sqrt{\rho_1}\big) \big\|_1. \tag{3.414}$$

3.5. Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces with $\dim(\mathcal{X}) = n$ and let $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$. Prove that

$$\big\|\!\big| \Phi \big|\!\big\|_1 \leq \| J(\Phi) \|_1 \leq n \big\|\!\big| \Phi \big|\!\big\|_1. \tag{3.415}$$

3.6. Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces and let $H \in \mathrm{Herm}(\mathcal{Y} \otimes \mathcal{X})$ be a Hermitian operator. Consider the problem of maximizing the value

$$\langle H, J(\Phi) \rangle \tag{3.416}$$

over all choices of a channel $\Phi \in \mathrm{C}(\mathcal{X}, \mathcal{Y})$. Prove that a channel $\Phi \in \mathrm{C}(\mathcal{X}, \mathcal{Y})$ satisfies

$$\langle H, J(\Phi) \rangle = \sup\{ \langle H, J(\Psi) \rangle \,:\, \Psi \in \mathrm{C}(\mathcal{X}, \mathcal{Y}) \} \tag{3.417}$$

if and only if the operator $\mathrm{Tr}_{\mathcal{Y}}(HJ(\Phi))$ is Hermitian and satisfies

$$\mathbb{1}_{\mathcal{Y}} \otimes \mathrm{Tr}_{\mathcal{Y}}(HJ(\Phi)) \geq H. \tag{3.418}$$

3.5 Bibliographic remarks

The task of quantum state discrimination was evidently first formulated (in abstract terms) by Helstrom [100], although instances of the task with respect to specific quantum physical systems had certainly been considered earlier. Theorem 3.4 was proved by Helstrom [100] for the restricted case of projective measurements, and by Holevo [104] for general measurements.

Theorem 3.7 was proved by Gutoski and Watrous [84], and a slightly weaker form of the theorem (for finite sets of states) was proved by Jain [124] around the same time, using Sion's minmax theorem. Jain's proof

extends easily to the more general case, and this is the proof that has been presented in this chapter.

Theorem 3.9 is attributed to Holevo [104, 105] and Yuen, Kennedy, and Lax [234, 235]. The semidefinite programming formulation that has been used in the proof of this theorem is due to Yuen, Kennedy, and Lax, although it was not recognized as a semidefinite program in their work (as their work predates much of the development of semidefinite programming). Eldar, Megretski, and Verghese [70] recognized this optimization problem as a semidefinite program. The pretty good measurement was so-named and popularized by Hausladen and Wootters [93]—it is one among a family of measurements introduced earler by Belavkin [26] and considered in other works (such as Eldar and Forney [69]). Theorem 3.10 is due to Barnum and Knill [20].

The fidelity function was introduced by Uhlmann [209], who referred to it as the *transition probability*. (Uhlmann defined the transition probability as the square of the fidelity function, as it has been defined in this book. Many authors follow the convention of referring to the square of the fidelity function as the fidelity function.) Uhlmann also proved Theorem 3.23 and observed several elementary properties of the fidelity function in the same paper. Corollary 3.21 is due to Alberti [5], and Theorem 3.30 is due to Alberti and Uhlmann [7]. The term *fidelity* was first introduced by Jozsa [127], who presented a simplified proof of Uhlmann's theorem.

A variant of Corollary 3.15 was proved by Winter in [230], stated in terms of the trace distance rather than the fidelity. Theorem 3.27 is due to Fuchs and Caves [74], Theorem 3.32 is due to Spekkens and Rudolph [196], and Theorem 3.36 is due to Fuchs and van de Graaf [75]. Theorem 3.17 and the semidefinite program associated with that theorem was independently found by Killoran [128] and Watrous [223].

The channel fidelity was introduced by Schumacher [186], who named it the entanglement fidelity and established some basic results about it, including a derivation of its expression as represented by Proposition 3.34 and the fact that it is invariant under the choice of the purification used to define it. A proof of Proposition 3.35 appears in Nielsen [166].

The relevance of the completely bounded trace norm to the theory of quantum information and computation appears to have first been realized by Kitaev [129], who took the spectral norm characterization (Theorem 3.65) as the definition and proved its equivalence to Definition 3.46. Several

basic properties of the completely bounded trace norm, including those summarized in Proposition 3.47, appear in the same paper, as well as in the work of Aharonov, Kitaev, and Nisan [3]. Kitaev used the notation $\|\cdot\|_\diamond$ rather than $|||\cdot|||_1$ when referring to the completely bounded trace norm, which led to its being referred to as the "diamond norm." This norm's close relationship to the completely bounded norm used in the study of operator algebras later came to be realized; in finite dimensions, the norm known as the completely bounded norm is essentially equivalent to the completely bounded trace norm, with the definition being based on the spectral norm rather than the trace norm. The book of Paulsen [168] provides an overview of the properties, uses, and history of this norm in the subject of operator algebras.

Example 3.39 appears to be somewhat of a folklore result. A variation of this example appears in Kretschmann, Schlingemann, and Werner [136], and had been recognized by others (including this author) a couple of years earlier. A different example having a similar character, but not giving as sharp a separation, appears in Kitaev, Shen, and Vyalyi [130]. (See Example 11.1 and the text immediately following in that book.) Underlying all of these examples is the observation that the transpose mapping provides a separation between the induced trace norm and the completely bounded trace norm; an equivalent example goes back (at least) to Arveson [14].

Theorem 3.42 is equivalent to a theorem of Russo and Dye [179]. Results similar to Lemma 3.48 appear in Gilchrist, Langford, and Nielsen [78] and Watrous [219], and a fact equivalent to Lemma 3.52 appears in Rosgen and Watrous [177]. Theorem 3.57 is stated in Aharonov, Kitaev, and Nisan [3] for the case of unitary channels, and an equivalent statement was proved by Childs, Preskill, and Renes [46]. The extension of this statement to isometric channels through the use of the Toeplitz–Hausdorff theorem can reasonably be described as being routine. A similar bound to the one in Theorem 3.58 appears in Kretschmann and Werner [137].

Theorem 3.61 appears (as an exercise, together with a solution) in Kitaev, Shen, and Vyalyi [130]. Theorem 3.65 is due to Smith [195]. Theorem 3.66 is due to Timoney [201], and the proof of this theorem given in this chapter is from Watrous [220].

The completely bounded trace norm can be expressed as a semidefinite program in a few different ways, as was proved by Watrous [222, 223]. Ben-Aroya and Ta-Shma [28] independently proved that the completely

bounded trace norm can be efficiently computed through the use of convex programming techniques, and a similar statement is made for the somewhat simpler task of computing the completely bounded trace norm of the difference between two channels by Gilchrist, Langford, and Nielsen [78]. Other computational methods for evaluating the completely bounded trace norm, but not accompanied by proofs of their computationally efficiency, were devised by Zarikian [236] and Johnston, Kribs, and Paulsen [126].