Lecture 20: Channel distinguishability and the completely bounded trace norm

This lecture is primarily concerned with the *distinguishability* of quantum channels, along with a norm defined on mappings between operator spaces that is closely related to this problem. In particular, we will define and study a norm called the *completely bounded trace norm* that plays an analogous role for channel distinguishability that the ordinary trace norm plays for density operator distinguishability.

20.1 Distinguishing between quantum channels

Recall from Lecture 3 that the trace norm has a close relationship to the optimal probability of distinguishing two quantum states. In particular, for every choice of density operators $\rho_0, \rho_1 \in \mathrm{D}(\mathcal{X})$ and a scalar $\lambda \in [0,1]$, it holds that

$$\max_{P_0,P_1} [\lambda \langle P_0, \rho_0 \rangle + (1-\lambda) \langle P_1, \rho_1 \rangle] = \frac{1}{2} + \frac{1}{2} \left\| \lambda \rho_0 - (1-\lambda)\rho_1 \right\|_1,$$

where the maximum is over all $P_0, P_1 \in \mathrm{Pos}(\mathcal{X})$ satisfying $P_0 + P_1 = \mathbb{1}_{\mathcal{X}}$, i.e., $\{P_0, P_1\}$ representing a binary-valued measurement. In words, the optimal probability to distinguish (or correctly identify) the states ρ_0 and ρ_1, given with probabilities λ and $1-\lambda$, respectively, by means of a measurement is

$$\frac{1}{2} + \frac{1}{2} \left\| \lambda \rho_0 - (1-\lambda)\rho_1 \right\|_1.$$

One may consider a similar situation involving channels rather than density operators. Specifically, let us suppose that $\Phi_0, \Phi_1 \in \mathrm{C}(\mathcal{X}, \mathcal{Y})$ are channels, and that a bit $a \in \{0,1\}$ is chosen at random, such that

$$\Pr[a = 0] = \lambda \qquad \text{and} \qquad \Pr[a = 1] = 1 - \lambda.$$

A single evaluation of the channel Φ_a is made available, and the goal is to determine the value of a with maximal probability.

One approach to this problem is to choose $\xi \in \mathrm{D}(\mathcal{X})$ in order to maximize the quantity $\|\rho_0 - \rho_1\|_1$ for $\rho_0 = \Phi_0(\xi)$ and $\rho_1 = \Phi_1(\xi)$. If a register X is prepared so that its state is ξ, and X is input to the given channel Φ_a, then the output is a register Y that can be measured using an optimal measurement to distinguish the two possible outputs ρ_0 and ρ_1.

This, however, is not the most general approach. More generally, one may include an *auxiliary register* Z in the process—meaning that a pair of registers (X, Z) is prepared in some state $\xi \in \mathrm{D}(\mathcal{X} \otimes \mathcal{Z})$, and the given channel Φ_a is applied to X. This results in a pair of registers (Y, Z) that will be in either of the states $\rho_0 = (\Phi_0 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})})(\xi)$ or $\rho_1 = (\Phi_1 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})})(\xi)$, which may then be distinguished by a measurement on $\mathcal{Y} \otimes \mathcal{Z}$. Indeed, this more general approach can give a sometimes striking improvement in the probability to distinguish Φ_0 and Φ_1, as the following example illustrates.

Example 20.1. Let $\mathcal{X}$ be a complex Euclidean space and let $n = \dim(\mathcal{X})$. Define channels $\Phi_0, \Phi_1 \in \mathrm{T}(\mathcal{X})$ as follows:

$$\Phi_0(X) = \frac{1}{n+1}\left((\operatorname{Tr} X)\mathbb{1}_{\mathcal{X}} + X^{\mathsf{T}}\right),$$

$$\Phi_1(X) = \frac{1}{n-1}\left((\operatorname{Tr} X)\mathbb{1}_{\mathcal{X}} - X^{\mathsf{T}}\right).$$

Both Φ_0 and Φ_1 are indeed channels: the fact that they are trace-preserving can be checked directly, while complete positivity follows from a calculation of the Choi-Jamiołkowski representations of these mappings:

$$J(\Phi_0) = \frac{1}{n+1}\left(\mathbb{1}_{\mathcal{X}\otimes\mathcal{X}} + W\right) = \frac{2}{n+1}S,$$

$$J(\Phi_1) = \frac{1}{n-1}\left(\mathbb{1}_{\mathcal{X}\otimes\mathcal{X}} - W\right) = \frac{2}{n-1}R,$$

where $W \in \mathrm{L}(\mathcal{X}\otimes\mathcal{X})$ is the swap operator and $R, S \in \mathrm{L}(\mathcal{X}\otimes\mathcal{X})$ are the projections onto the symmetric and antisymmetric subspaces of $\mathcal{X}\otimes\mathcal{X}$, respectively.

Now, for any choice of a density operator $\xi \in \mathrm{D}(\mathcal{X})$ we have

$$\begin{aligned}\Phi_0(\xi) - \Phi_1(\xi) &= \left(\frac{1}{n+1} - \frac{1}{n-1}\right)\mathbb{1}_{\mathcal{X}} + \left(\frac{1}{n+1} + \frac{1}{n-1}\right)\xi^{\mathsf{T}}\\ &= -\frac{2}{n^2-1}\mathbb{1}_{\mathcal{X}} + \frac{2n}{n^2-1}\xi^{\mathsf{T}}.\end{aligned}$$

The trace norm of such an operator is maximized when ξ has rank 1, in which case the value of the trace norm is

$$\frac{2n-2}{n^2-1} + (n-1)\frac{2}{n^2-1} = \frac{4}{n+1}.$$

Consequently

$$\|\Phi_0(\xi) - \Phi_1(\xi)\|_1 \leq \frac{4}{n+1}$$

for all $\xi \in \mathrm{D}(\mathcal{X})$. For large n, this quantity is small, which is not surprising because both $\Phi_0(\xi)$ and $\Phi_1(\xi)$ are almost completely mixed for any choice of $\xi \in \mathrm{D}(\mathcal{X})$.

However, suppose that we prepare two registers $(\mathsf{X}_1, \mathsf{X}_2)$ in the maximally entangled state

$$\tau = \frac{1}{n}\operatorname{vec}(\mathbb{1}_{\mathcal{X}})\operatorname{vec}(\mathbb{1}_{\mathcal{X}})^* \in \mathrm{D}(\mathcal{X}\otimes\mathcal{X}).$$

We have

$$(\Phi_a \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})})(\tau) = \frac{1}{n}J(\Phi_a)$$

for $a \in \{0, 1\}$, and therefore

$$(\Phi_0 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})})(\tau) = \frac{2}{n(n+1)}S \qquad \text{and} \qquad (\Phi_1 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})})(\tau) = \frac{2}{n(n-1)}R.$$

Because R and S are orthogonal, we have

$$\left\|(\Phi_0 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})})(\tau) - (\Phi_1 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})})(\tau)\right\|_1 = 2,$$

meaning that the states $(\Phi_0 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})})(\tau)$ and $(\Phi_1 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})})(\tau)$, and therefore the channels Φ_0 and Φ_1, can be distinguished *perfectly*.

By applying Φ_0 or Φ_1 to part of a larger system, we have therefore completely eliminated the large error that was present when the limited approach of choosing $\xi \in \mathrm{D}(\mathcal{X})$ as an input to Φ_0 and Φ_1 was considered.

The previous example makes clear that auxiliary systems must be taken into account if we are to understand the optimal probability with which channels can be distinguished.

20.2 Definition and properties of the completely bounded trace norm

With the discussion from the previous section in mind, we will now discuss two norms: the *induced trace norm* and the *completely bounded trace norm*. The precise relationship these norms have to the notion of channel distinguishability will be made clear later in the section.

20.2.1 The induced trace norm

We will begin with the induced trace norm. This norm does not provide a suitable way to measure distances between channels, at least with respect to the issues discussed in the previous section, but it will nevertheless be helpful from a mathematical point of view for us to start with this norm.

For any choice of complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$, and for a given mapping $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$, the induced trace norm is defined as

$$\|\Phi\|_1 = \max\{\|\Phi(X)\|_1 : X \in \mathrm{L}(\mathcal{X}),\ \|X\|_1 \leq 1\}. \tag{20.1}$$

This norm is just one of many possible examples of induced norms; in general, one may consider the norm obtained by replacing the two trace norms in this definition with any other choice 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 every norm defined on a complex Euclidean space is continuous and its corresponding unit ball is compact.

Let us note two simple properties of the induced trace norm that will be useful for our purposes. First, for every choice of complex Euclidean spaces $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$, and mappings $\Psi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ and $\Phi \in \mathrm{T}(\mathcal{Y}, \mathcal{Z})$, it holds that

$$\|\Phi\Psi\|_1 \leq \|\Phi\|_1 \|\Psi\|_1. \tag{20.2}$$

This is a general property of every induced norm, provided that the same norm on $\mathrm{L}(\mathcal{Y})$ is taken for both induced norms.

Second, the induced trace norm can be expressed as follows for a given mapping $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$:

$$\|\Phi\|_1 = \max\{\|\Phi(uv^*)\|_1 : u, v \in \mathcal{S}(\mathcal{X})\}, \tag{20.3}$$

where $\mathcal{S}(\mathcal{X}) = \{x \in \mathcal{X} : \|x\| = 1\}$ denotes the unit sphere in $\mathcal{X}$. This fact holds because the trace norm (like every other norm) is a convex function, and the unit ball with respect to the trace norm can be represented as

$$\{X \in \mathrm{L}(\mathcal{X}) : \|X\|_1 \leq 1\} = \operatorname{conv}\{uv^* : u, v \in \mathcal{S}(\mathcal{X})\}.$$

Alternately, one can prove that (20.3) holds by considering a singular value decomposition of any operator $X \in \mathrm{L}(\mathcal{X})$ with $\|X\|_1 \leq 1$ that maximizes (20.1).

One undesirable property of the induced trace norm is that it is not multiplicative with respect to tensor products. For instance, for a complex Euclidean space $\mathcal{X}$, let us consider the transpose mapping $T \in \mathrm{T}(\mathcal{X})$ and the identity mapping $\mathbb{1}_{\mathrm{L}(\mathcal{X})} \in \mathrm{T}(\mathcal{X})$. It holds that

$$\|T\|_1 = 1 = \left\|\mathbb{1}_{\mathrm{L}(\mathcal{X})}\right\|_1$$

but

$$\left\|T \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})}\right\|_1 \geq \left\|(T \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})})(\tau)\right\|_1 = \frac{1}{n}\|W\|_1 = n \tag{20.4}$$

for $n = \dim(\mathcal{X})$, $\tau \in \mathrm{D}(\mathcal{X} \otimes \mathcal{X})$ defined as

$$\tau = \frac{1}{n}\operatorname{vec}(\mathbb{1}_{\mathcal{X}})\operatorname{vec}(\mathbb{1}_{\mathcal{X}})^*,$$

and $W \in \mathrm{L}(\mathcal{X} \otimes \mathcal{X})$ denoting the swap operator, as in Example 20.1. (The inequality in (20.4) is really an equality, but it is not necessary for us to prove this at this moment.) Thus, we have

$$\left\|T \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})}\right\|_1 > \|T\|_1 \left\|\mathbb{1}_{\mathrm{L}(\mathcal{X})}\right\|_1,$$

assuming $n \geq 2$.

20.2.2 The completely bounded trace norm

We will now define the completely bounded trace norm, which may be seen as a modification of the induced trace norm that corrects for that norm's failure to be multiplicative with respect to tensor products.

For any choice of complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$, and a mapping $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$, we define the *completely bounded trace norm* of Φ to be

$$|||\Phi|||_1 = \left\|\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})}\right\|_1.$$

(This norm is also commonly called the *diamond norm*, and denoted $\|\Phi\|_\diamond$.) The principle behind its definition is that tensoring Φ with the identity mapping has the effect of *stabilizing* its induced trace norm. This sort of stabilization endows the completely bounded trace norm with many nice properties, and allows it to be used in contexts where the induced trace norm is not sufficient. This sort of stabilization is also related to the phenomenon illustrated in the previous section, where tensoring with the identity channel had the effect of amplifying the difference between channels.

The first thing we must do is to explain why the definition of the completely bounded trace norm tensors Φ with the identity mapping on $\mathrm{L}(\mathcal{X})$, rather than some other space. The answer is that $\mathcal{X}$ has sufficiently large dimension—and replacing $\mathcal{X}$ with any complex Euclidean space with larger dimension would not change anything. The following lemma allows us to prove this fact, and includes a special case that will be useful later in the lecture.

Lemma 20.2. *Let $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$, and let $\mathcal{Z}$ be a complex Euclidean space. For every choice of unit vectors $u, v \in \mathcal{X} \otimes \mathcal{Z}$ there exist unit vectors $x, y \in \mathcal{X} \otimes \mathcal{X}$ such that*

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

In case $u = v$, we may in addition take $x = y$.

Proof. The lemma is straightforward when $\dim(\mathcal{Z}) \leq \dim(\mathcal{X})$; for any choice of a linear isometry $U \in \mathrm{U}(\mathcal{Z}, \mathcal{X})$ the vectors $x = (\mathbb{1}_{\mathcal{X}} \otimes U)u$ and $y = (\mathbb{1}_{\mathcal{X}} \otimes U)v$ satisfy the required conditions. Let us therefore consider the case where $\dim(\mathcal{Z}) > \dim(\mathcal{X}) = n$.

Consider the vector $u \in \mathcal{X} \otimes \mathcal{Z}$. Given that $\dim(\mathcal{X}) \leq \dim(\mathcal{Z})$, there must exist an orthogonal projection $\Pi \in \mathrm{L}(\mathcal{Z})$ having rank at most $n = \dim(\mathcal{X})$ that satisfies $u = (\mathbb{1}_{\mathcal{X}} \otimes \Pi)u$ and therefore there must exist a linear isometry $U \in \mathrm{U}(\mathcal{X}, \mathcal{Z})$ such that $u = (\mathbb{1}_{\mathcal{X}} \otimes UU^*)u$. Likewise there must exist a linear isometry $V \in \mathrm{U}(\mathcal{X}, \mathcal{Z})$ such that $v = (\mathbb{1}_{\mathcal{X}} \otimes VV^*)v$. Such linear isometries U and V can be obtained from Schmidt decompositions of u and v.

Now let $x = (\mathbb{1} \otimes U^*)u$ and $y = (\mathbb{1} \otimes V^*)v$. Notice that we therefore have $u = (\mathbb{1} \otimes U)x$ and $v = (\mathbb{1} \otimes V)y$, which shows that x and y are unit vectors. Moreover, given that $\|U\| = \|V\| = 1$, we have

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

as required. In case $u = v$, we may take $U = V$, implying that $x = y$. □

The following theorem, which explains the choice of taking the identity mapping on $\mathrm{L}(\mathcal{X})$ in the definition of the completely bounded trace norm, is immediate from Lemma 20.2 together with (20.3).

Theorem 20.3. *Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces, let $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ be any mapping, and let $\mathcal{Z}$ be any complex Euclidean space for which* $\dim(\mathcal{Z}) \geq \dim(\mathcal{X})$*. It holds that*

$$\left\| \Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})} \right\|_1 = |||\Phi|||_1 .$$

One simple but important consequence of this theorem is that the completely bounded trace norm is multiplicative with respect to tensor products.

Theorem 20.4. *For every choice of mappings $\Phi_1 \in \mathrm{T}(\mathcal{X}_1, \mathcal{Y}_1)$ and $\Phi_2 \in \mathrm{T}(\mathcal{X}_2, \mathcal{Y}_2)$, it holds that*

$$|||\Phi_1 \otimes \Phi_2|||_1 = |||\Phi_1|||_1 \, |||\Phi_2|||_1 .$$

Proof. Let $\mathcal{W}_1$ and $\mathcal{W}_2$ be complex Euclidean spaces with $\dim(\mathcal{W}_1) = \dim(\mathcal{X}_1)$ and $\dim(\mathcal{W}_2) = \dim(\mathcal{X}_2)$, so that

$$\begin{aligned}
|||\Phi_1|||_1 &= \left\| \Phi_1 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{W}_1)} \right\|_1 , \\
|||\Phi_2|||_1 &= \left\| \Phi_2 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{W}_2)} \right\|_1 , \\
|||\Phi_1 \otimes \Phi_2|||_1 &= \left\| \Phi_1 \otimes \Phi_2 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{W}_1 \otimes \mathcal{W}_2)} \right\|_1 .
\end{aligned}$$

We have

$$\begin{aligned}
|||\Phi_1 \otimes \Phi_2|||_1 &= \left\| \Phi_1 \otimes \Phi_2 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{W}_1 \otimes \mathcal{W}_2)} \right\|_1 \\
&= \left\| \left(\Phi_1 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y}_2)} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{W}_1 \otimes \mathcal{W}_2)} \right) \left(\mathbb{1}_{\mathrm{L}(\mathcal{X}_1)} \otimes \Phi_2 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{W}_1 \otimes \mathcal{W}_2)} \right) \right\|_1 \\
&\leq \left\| \left(\Phi_1 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y}_2)} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{W}_1 \otimes \mathcal{W}_2)} \right) \right\|_1 \left\| \left(\mathbb{1}_{\mathrm{L}(\mathcal{X}_1)} \otimes \Phi_2 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{W}_1 \otimes \mathcal{W}_2)} \right) \right\|_1 \\
&= |||\Phi_1|||_1 \, |||\Phi_2|||_1 ,
\end{aligned}$$

where the last inequality follows by Theorem 20.3.

For the reverse inequality, choose operators $X_1 \in \mathrm{L}(\mathcal{X}_1 \otimes \mathcal{W}_1)$ and $X_2 \in \mathrm{L}(\mathcal{X}_2 \otimes \mathcal{W}_2)$ such that $\|X_1\|_1 = \|X_2\|_1 = 1$, $|||\Phi_1|||_1 = \|(\Phi_1 \otimes \mathbb{1}_{\mathcal{W}_1})(X_1)\|_1$, and $|||\Phi_2|||_1 = \|(\Phi_2 \otimes \mathbb{1}_{\mathcal{W}_2})(X_2)\|_1$. IT holds that $\|X_1 \otimes X_2\|_1 = 1$, and so

$$\begin{aligned}
|||\Phi_1 \otimes \Phi_2|||_1 &= \left\| \Phi_1 \otimes \Phi_2 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{W}_1 \otimes \mathcal{W}_2)} \right\|_1 \\
&\geq \left\| \left(\Phi_1 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{W}_1)} \otimes \Phi_2 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{W}_2)} \right) (X_1 \otimes X_2) \right\|_1 \\
&= \left\| \left(\Phi_1 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{W}_1)} \right) (X_1) \otimes \left(\Phi_2 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{W}_2)} \right) (X_2) \right\|_1 \\
&= \left\| \left(\Phi_1 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{W}_1)} \right) (X_1) \right\|_1 \left\| \left(\Phi_2 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{W}_2)} \right) (X_2) \right\|_1 \\
&= |||\Phi_1|||_1 \, |||\Phi_2|||_1
\end{aligned}$$

as required. □

The final fact that we will establish about the completely bounded trace norm in this lecture concerns input operators for which the value of the completely bounded trace norm is achieved. In particular, we will establish that for Hermiticity-preserving mappings $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$, there must exist a unit vector $u \in \mathcal{X} \otimes \mathcal{X}$ for which

$$|||\Phi|||_1 = \left\| (\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})})(uu^*) \right\|_1 .$$

This fact will give us the final piece we need to connect the completely bounded trace norm to the distinguishability problem discussed in the beginning of the lecture.

Theorem 20.5. *Suppose that $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ is Hermiticity-preserving. It holds that*

$$|||\Phi|||_1 = \max \left\{ \left\| (\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})})(xx^*) \right\|_1 \,:\, x \in \mathcal{S}(\mathcal{X} \otimes \mathcal{X}) \right\}.$$

Proof. Let $X \in \mathrm{L}(\mathcal{X} \otimes \mathcal{X})$ be an operator with $\|X\|_1 = 1$ that satisfies

$$|||\Phi|||_1 = \left\| (\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})})(X) \right\|_1 .$$

Let $\mathcal{Z} = \mathbb{C}^{\{0,1\}}$ and let

$$Y = \frac{1}{2} X \otimes E_{0,1} + \frac{1}{2} X^* \otimes E_{1,0} \in \mathrm{Herm}(\mathcal{X} \otimes \mathcal{X} \otimes \mathcal{Z}).$$

We have $\|Y\| = \|X\| = 1$ and

$$\begin{aligned}
\left\| (\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X} \otimes \mathcal{Z})})(Y) \right\|_1 &= \frac{1}{2} \left\| (\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})})(X) \otimes E_{0,1} + (\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})})(X^*) \otimes E_{1,0} \right\|_1 \\
&= \frac{1}{2} \left\| (\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})})(X) \otimes E_{0,1} + ((\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})})(X))^* \otimes E_{1,0} \right\|_1 \\
&= \left\| (\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})})(X) \right\|_1 \\
&= |||\Phi|||_1 ,
\end{aligned}$$

where the second equality follows from the fact that Φ is Hermiticity-preserving. Now, because Y is Hermitian, we may consider a spectral decomposition

$$Y = \sum_j \lambda_j u_j u_j^*.$$

By the triangle inequality we have

$$|||\Phi|||_1 = \left\| (\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X}\otimes\mathcal{Z})})(Y) \right\| \leq \sum_j |\lambda_j| \left\| (\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X}\otimes\mathcal{Z})})(u_j u_j^*) \right\|_1 .$$

As $\|Y\| = 1$, we have $\sum_j |\lambda_j| = 1$, and thus

$$\left\| (\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X}\otimes\mathcal{Z})})(u_j u_j^*) \right\|_1 \geq |||\Phi|||_1$$

for some index j. We may therefore apply Lemma 20.2 to obtain a unit vector $x \in \mathcal{X} \otimes \mathcal{X}$ such that

$$\left\| (\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})})(xx^*) \right\|_1 = \left\| (\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X}\otimes\mathcal{Z})})(u_j u_j^*) \right\|_1 \geq |||\Phi|||_1 .$$

Thus

$$|||\Phi|||_1 \leq \max \left\{ \left\| (\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})})(xx^*) \right\|_1 \,:\, x \in \mathcal{S}(\mathcal{X} \otimes \mathcal{X}) \right\}.$$

It is clear that the maximum cannot exceed $|||\Phi|||_1$, and so the proof is complete. □

Let us note that at this point we have established the relationship between the completely bounded trace norm and the problem of channel distinguishability that we discussed at the beginning of the lecture; in essence, the relationship is an analogue of Helstrom's theorem for channels.

Theorem 20.6. *Let $\Phi_0, \Phi_1 \in \mathrm{C}(\mathcal{X}, \mathcal{Y})$ be channels and let $\lambda \in [0,1]$. It holds that*

$$\sup_{\xi, P_0, P_1} \left[\lambda \left\langle P_0, (\Phi_0 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})})(\xi) \right\rangle + (1-\lambda) \left\langle P_1, (\Phi_1 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})})(\xi) \right\rangle \right] = \frac{1}{2} + \frac{1}{2} |||\lambda\Phi_0 - (1-\lambda)\Phi_1|||_1 ,$$

where the supremum is over all complex Euclidean spaces $\mathcal{Z}$, density operators $\xi \in \mathrm{D}(\mathcal{X} \otimes \mathcal{Z})$, and binary valued measurements $\{P_0, P_1\} \subset \mathrm{Pos}(\mathcal{Y} \otimes \mathcal{Z})$. Moreover, the supremum is achieved for $\mathcal{Z} = \mathcal{X}$ and $\xi = uu^*$ being a pure state.

20.3 Distinguishing unitary and isometric channels

We started the lecture with an example of two channels $\Phi_0, \Phi_1 \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ for which an auxiliary system was necessary to optimally distinguish Φ_0 and Φ_1. Let us conclude the lecture by observing that this phenomenon does not arise for mappings induced by linear isometries.

Theorem 20.7. *Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces, let $U, V \in \mathrm{U}(\mathcal{X}, \mathcal{Y})$ be linear isometries, and suppose*

$$\Phi_0(X) = UXU^* \quad \textit{and} \quad \Phi_1(X) = VXV^*$$

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

$$\| \Phi_0(uu^*) - \Phi_1(uu^*) \|_1 = |||\Phi_0 - \Phi_1|||_1 .$$

Proof. Recall that the numerical range of an operator $A \in \mathrm{L}(\mathcal{X})$ is defined as

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

Let us define $\nu(A)$ to be the *smallest* absolute value of any element of $\mathcal{N}(A)$:

$$\nu(A) = \min\{|\alpha| \,:\, \alpha \in \mathcal{N}(A)\}.$$

Now, for any choice of a unit vector $u \in \mathcal{X}$ we have

$$\|\Phi_0(uu^*) - \Phi_1(uu^*)\|_1 = \|Uuu^*U^* - Vuu^*V^*\|_1 = 2\sqrt{1 - |u^*U^*Vu|^2}.$$

Maximizing this quantity over $u \in \mathcal{S}(\mathcal{X})$ gives

$$\max\{\|\Phi_0(uu^*) - \Phi_1(uu^*)\|_1 \,:\, u \in \mathcal{S}(\mathcal{X})\} = 2\sqrt{1 - \nu(U^*V)^2}.$$

Along similar lines, we have

$$|\!|\!|\Phi_0 - \Phi_1|\!|\!|_1 = \max_{u \in \mathcal{S}(\mathcal{X}\otimes\mathcal{X})} \left\|(\Phi_0 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})})(uu^*) - (\Phi_1 \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})})(uu^*)\right\|_1 = 2\sqrt{1 - \nu(U^*V \otimes \mathbb{1}_{\mathcal{X}})^2},$$

where here we have also made use of Theorem 20.5.

To complete the proof it therefore suffices to prove that

$$\nu(A \otimes \mathbb{1}_{\mathcal{X}}) = \nu(A)$$

for every operator $A \in \mathrm{L}(\mathcal{X})$. It is clear that $\nu(A \otimes \mathbb{1}_{\mathcal{X}}) \leq \nu(A)$, so we just need to prove $\nu(A \otimes \mathbb{1}_{\mathcal{X}}) \geq \nu(A)$. Let $u \in \mathcal{X} \otimes \mathcal{X}$ be any unit vector, and let

$$u = \sum_{j=1}^{r} \sqrt{p_j}\, x_j \otimes y_j$$

be a Schmidt decomposition of u. It holds that

$$u^*(A \otimes \mathbb{1}_{\mathcal{X}})u = \sum_{j=1}^{r} p_j x_j^* A x_j.$$

For each j we have $u_j^* A u_j \in \mathcal{N}(A)$, and therefore

$$u^*(A \otimes \mathbb{1}_{\mathcal{X}})u = \sum_{j=1}^{r} p_j x_j^* A x_j \in \mathcal{N}(A)$$

by the Toeplitz–Hausdorff theorem. Consequently

$$|u^*(A \otimes \mathbb{1}_{\mathcal{X}})u| \geq \nu(A)$$

for every $u \in \mathcal{S}(\mathcal{X} \otimes \mathcal{X})$, which completes the proof. □